



**HAL**  
open science

# Basket data-driven forecasting and inventory management for omnichannel supply chain

Haytham Omar Abdelakher Mohamed

► **To cite this version:**

Haytham Omar Abdelakher Mohamed. Basket data-driven forecasting and inventory management for omnichannel supply chain. Automatic Control Engineering. Université de Bordeaux, 2023. English. NNT : 2023BORD0352 . tel-04429514

**HAL Id: tel-04429514**

**<https://theses.hal.science/tel-04429514>**

Submitted on 31 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE  
POUR OBTENIR LE GRADE DE  
**DOCTEUR**  
**DE L'UNIVERSITÉ DE BORDEAUX**

ECOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGÉNIEUR  
SPÉCIALITÉ : AUTOMATIQUE, PRODUCTIQUE, SIGNAL ET IMAGE,  
INGENIERIE COGNITIVE

Par **Haytham OMAR**

**PREVISION ET GESTION DE STOCKS BASES  
SUR LES DONNEES DE PANIER DANS LA  
SUPPLY CHAIN OMNICANAL**

Sous la direction de : **Yves Ducq**  
Co-directeur : **Walid Klibi**

Soutenue le 1 décembre 2023

Membres du jury :

M. Yacine Rekik	Professeur	ESCP Business School	Rapporteur et présidente du jury
M. Konstantinos Nikolopoulos	Professeur	Durham University	Rapporteur
M. Bahman Rostami-Tabar	Professeur Associé	Cardiff University	Examinateur
M. Yves Ducq	Professeur	Université de Bordeaux	Directeur
M. Walid Klibi	Professeur	Kedge Business School	Co-directeur
M. Zied Babai	Professeur	Kedge Business School	Invité



# Acknowledgment

I would like to express my gratitude to my father, my mother, my sisters, and, my grandmother.

Also, my professors and supervisors, Prof. Yves Ducq, Prof. Walid Klibi, and, Prof. Zied Babai who, although not formally enlisted as co-director of the thesis, equally participated in the supervision of the thesis and its contributions. I am grateful to have them as directors/co-directors of this thesis.

# Contents

<b>List of Figures</b>	<b>7</b>
<b>List of Tables</b>	<b>8</b>
<b>Publications</b>	<b>9</b>
<b>Introduction</b>	<b>10</b>
<b>1 Problem statement and literature review</b>	<b>17</b>
1.1 Introduction and Problem Statement . . . . .	17
1.1.1 Definitions . . . . .	17
1.2 Research background . . . . .	21
1.2.1 Introduction . . . . .	21
1.2.2 Omni-channel retailing . . . . .	22
1.2.3 Anticipatory shipping . . . . .	28
1.2.4 Demand Forecasting in Retail . . . . .	30
1.2.5 Forecasting aggregation in supply chain . . . . .	36
1.3 Conclusion . . . . .	37
<b>2 Basket Data driven approach for omnichannel demand forecasting</b>	<b>38</b>
2.1 Introduction . . . . .	38
2.1.1 General introduction and motivation . . . . .	38
2.1.2 Business context . . . . .	39
2.1.3 Contributions and organization of the chapter . . . . .	40
2.2 Market Baskets . . . . .	40
2.3 Basket data-driven forecasting approach . . . . .	41
2.3.1 Network characterization and attributes extraction . . . . .	42
2.3.2 Forecasting methods . . . . .	45
2.3.3 Inventory and fulfillment performance . . . . .	46
2.4 Empirical investigation . . . . .	47
2.4.1 Supply chain and data description . . . . .	47
2.4.2 Network and basket analysis . . . . .	47
2.4.3 Substitutability set tests . . . . .	48
2.4.4 Attributes testing and validation results . . . . .	54
2.4.5 Forecast accuracy based on different network horizons . . . . .	56
2.4.6 Forecasting accuracy analysis . . . . .	58
2.4.7 Statistical significance test for forecasting methods . . . . .	61
2.4.8 Omnichannel inventory performance . . . . .	62
2.5 Conclusion . . . . .	65

<b>3</b>	<b>A new basket data-driven approach for forecast-based anticipatory shipping</b>	<b>67</b>
3.1	Introduction . . . . .	67
3.1.1	General introduction and motivation . . . . .	67
3.1.2	Business context . . . . .	68
3.2	Two-step advancement shipping approach . . . . .	71
3.2.1	Basket Prediction . . . . .	71
3.2.2	Prediction-based basket retrieval algorithm . . . . .	78
3.2.3	Retrieval algorithm variations . . . . .	79
3.3	Performance Evaluation- Fulfillment and Costing . . . . .	79
3.4	Empirical investigation . . . . .	84
3.4.1	Supply chain and data description . . . . .	84
3.4.2	Arc modeling and feature analysis . . . . .	84
3.5	Conclusion . . . . .	90
<b>4</b>	<b>Basket data impact on forecasting aggregation and inventory fulfillment</b>	<b>92</b>
4.1	Introduction . . . . .	92
4.1.1	General introduction and motivation . . . . .	92
4.1.2	Business context . . . . .	93
4.1.3	Contributions and organization of the chapter . . . . .	93
4.2	Hierarchical forecasting approaches . . . . .	94
4.2.1	Approaches' definitions . . . . .	94
4.2.2	Findings on the approaches' performance . . . . .	95
4.3	Considered basket data attributes and methods . . . . .	96
4.3.1	Basket data attributes . . . . .	96
4.3.2	Forecasting models . . . . .	97
4.3.3	Forecast accuracy measures . . . . .	98
4.3.4	Inventory performance . . . . .	98
4.3.5	Empirical investigation . . . . .	99
4.4	Forecast accuracy results . . . . .	100
4.5	Inventory performance results . . . . .	107
4.5.1	Outliers . . . . .	112
4.6	Conclusion . . . . .	113
	<b>Conclusion and Perspectives</b>	<b>116</b>

# List of Figures

- 1.1 A graph of products . . . . . 19
- 1.2 different retail formats . . . . . 22
- 1.3 Customer journey . . . . . 25
- 1.4 Different distribution design - [Chopra and Meindl \(2016\)](#) . . . . . 27
- 1.5 Layout of forecast aggregation approaches from [Babai et al. \(2022\)](#) . . . . . 37
  
- 2.1 Omnichannel fulfillment . . . . . 39
- 2.2 Illustration of the forecasting approach . . . . . 42
- 2.3 Basket data representation in a multi-category graph . . . . . 43
- 2.4 Illustrative example of a multi-category graph and the associated calculation of the attributes . . . . . 44
- 2.5 Illustration of the demand classification proposed by Syntetos et al. (2005) 50
- 2.6 Percentage of products in each class in the online channel . . . . . 50
- 2.7 Percentage of single-product baskets and multiple-product baskets . . . . . 51
- 2.8 The number of times (percentage wise) each product was sold with other products . . . . . 51
- 2.9 Degree distribution of products per class of products . . . . . 53
- 2.10 Relation between number of substitutes and average sales (left handside), distribution of complementarity (right handside) . . . . . 53
- 2.11 Strength distribution of products per class of products . . . . . 54
- 2.12 Baskets characteristics for the omnichannel case . . . . . 61
- 2.13 Efficiency curves of the different forecasting methods . . . . . 64
- 2.14 Efficiency curves: holding Versus backordering volumes of the fulfillment scenarios . . . . . 65
  
- 3.1 Omni-channel fulfillment- AS-IS VS TO-BE . . . . . 69
- 3.2 Illustration of the prediction approach . . . . . 72
- 3.3 Basket data representation in a weighted graph . . . . . 73
- 3.4 Nodes to the arc attributes . . . . . 73
- 3.5 Arc attributes . . . . . 74
- 3.6 Probability/Bayes attributes . . . . . 75
- 3.7 Basket data representation in a weighted graph/Arc attributes . . . . . 78
- 3.8 Retrieval algorithm variations . . . . . 80
- 3.9 Multi-period 8-week simulation . . . . . 80
- 3.10 Basket data size frequencies and histogram . . . . . 85
- 3.11 Future occurrences of historical exact baskets . . . . . 85
- 3.12 Correlation between attributes,BS2 and BS>2 . . . . . 86
- 3.13 Conservation threshold effect on prediction . . . . . 88
  
- 4.1 Hierarchical grouped structure . . . . . 97
- 4.2 Hierarchical grouped structure . . . . . 100
- 4.5 Store level error quartiles . . . . . 105
- 4.3 Regional level error quarterlies . . . . . 106

4.6	Store level error quartiles by class . . . . .	106
4.4	Regional level error quartiles by class . . . . .	107
4.7	Average lost sales vs average safety stock- Distribution Centre . . . . .	111
4.8	Average lost sales vs average safety stock-Store . . . . .	112
4.10	Inventory performance of outliers . . . . .	113
4.9	Outliers in lumpy Vs standard lumpy class . . . . .	114

# List of Tables

- 2.1 Results of the substitutability set tests . . . . . 49
- 2.2 Attributes significance . . . . . 55
- 2.3 Attributes significance per class under the polynomial regression model . . 56
- 2.4 Forecasting accuracy results for different horizon lengths-ARIMAX . . . . 57
- 2.5 Forecasting accuracy results for different horizon lengths-LinReg-LS . . . . 57
- 2.6 Forecasting accuracy results . . . . . 59
- 2.7 Forecasting accuracy results per category of products . . . . . 60
- 2.8 Omnichannel forecasting results . . . . . 61
- 2.9 Z-test scores - absolute, squared and symmetric absolute percentage errors 62
- 2.10 Empirical goodness-of-fit results . . . . . 63
- 2.11 Inventory performance results . . . . . 63
  
- 3.1 Comparison between No advancement(baseline) and baskets advancement . 82
- 3.2 Notation . . . . . 83
- 3.3 ML Models performance on Panel B . . . . . 87
- 3.4 Forecasting samples from test set . . . . . 88
- 3.5 Fulfillment/Costing -BS2 . . . . . 89
- 3.6 Fulfillment/Costing -BS2 . . . . . 90
  
- 4.1 Percentage of products obtained after aggregation . . . . . 100
- 4.2 Aggregated top-level accuracy . . . . . 101
- 4.3 Regional level accuracy . . . . . 102
- 4.4 Regional level accuracy percentiles-RMSSE . . . . . 103
- 4.5 Regional level accuracy percentile-Lumpy-RMSSE . . . . . 104
- 4.7 Store level accuracy . . . . . 104
- 4.6 Regional level accuracy percentile-Erratic-RMSSE . . . . . 105
- 4.8 Store level accuracy percentile-RMSSE . . . . . 108
- 4.11 Inventory performance - regional level . . . . . 108
- 4.9 Store level accuracy percentile-Lumpy-RMSSE . . . . . 109
- 4.12 Inventory Performance - store level . . . . . 109
- 4.10 Store level accuracy percentile-Erratic-RMSSE . . . . . 110

# Publications

## Journals

1. Omar, H., Klibi, W., Babai, M. Z., and Ducq, Y. Basket data-driven approach for omnichannel demand forecasting. *International Journal of Production and Economics* – Elsevier – Vol. 257 – Mars 2023 -DOI: 10.1016/j.ijpe.2022.108748

## International Conferences

1. Omar, H., Klibi, W., Babai, M. Z., and Ducq, Y. (2022), "On the inventory performance of omnichannel fulfillment strategies" *Proceedings of the Production and Operations Management Society Conference, College of Business Administration, Florida International University- April 21 2022*
2. Omar, H., Klibi, W., Babai, M. Z., and Ducq, Y. (2021), "A Two-step Demand Forecasting Method for Omnichannel Retailing" , *Proceedings of the Production and Operations Management Society Conference, College of Business Administration, Florida International University-April 30 2021*

# Introduction

## Business context and motivation

Omnichannel retailing is of growing importance. The omnichannel business seeks a seamless shopping experience, allowing customers to order anytime from anywhere, in person or through digital devices, and to have their purchase delivered at their preferred time and location (Bell et al. 2014). It offers flexibility to consumers in shopping through stores or online and the convenience of picking and delivery. Before the internet boom in the 2000s, customers' accessibility to information was limited and required physical effort to compare between products. This was not the case after the 2000s as customers now expect more service, convenience, and less tolerance to lack of virtual accessibility. As Chopra (2018) pointed out, retailers who failed to adapt to online presence slowly went out of business or lost a big portion of their market share. For example, Barnes and Noble against AMAZON's early online bookstore business, Barnes and Noble had a physically intensive infrastructure across the United States. At the same time, AMAZON only relied on small storage hubs at the time to sell books, which led to significant price differences in the price of books and more convenient delivery options of books in favor of AMAZON. Customers now expect the option to buy from physical stores their basic needs and shop through online platforms when buying more sophisticated items. In addition, customers tend to shop physically for new items to get familiarized with their features, afterwards, if they like the items, they can buy them again more conveniently through online platforms. In terms of delivery, customers can shop online and ask for products to be delivered door to door or go to a nearby store for pick up.

This change in customer behavior and retailer dynamics has led to the existence of three types of retail channels: i) a traditional retail channel that carries a high level of inventory and is capital-intensive. Such examples can be department stores, jewelry stores, and supermarkets as well as small convenience stores and small specialty stores. This type of service offers faster fulfillment and better customer experience, yet except for hyper stores, small to medium traditional retail stores have less assortment and offerings than online stores due to the restriction of physical space, ii) the second category is pure online stores which only offer through online platforms such as Blue Nile a specialty jewelry online store that offers more assortment than the traditional jewelry store and is also less expensive due to the economics of scale and inventory aggregation. However, transportation cost is higher as the main hub is distant from different destinations where customer orders, iii) finally, an omnichannel retailer by default invests in both capital-intensive physical stores and online fulfillment capabilities. Not only this, the supply chain operations must be adaptable to the cross-channel flows when it comes to information and physical flow of products, for example, buy online pick up from store (BOPS) where a customer order online and pick from a store, this implies more sophisticated technological resource planning systems and more collaboration between channels.

Another important customer behavior encountered nowadays in the retail context is that customers tend to buy various items in one basket for several reasons. For example,

the trip-to-store cost and time, lead to buying more items to save on cost. Another incentive is the promotions offered by retailers such as buy one get one or minimum order value for a discount or when buying online to be eligible for free delivery. Note that the baskets are often multiple-item baskets. Moreover, the items within the baskets usually show some form of association whether positive in the case of items complementing each other, neutral where items are not related to each other, or negative in the case of the existence of competitive characteristics between these items. Interestingly, other than marketing and product recommendations, such abundant and available characteristics are often not considered in demand forecasting or omnichannel fulfillment.

Furthermore, the availability of information online and standardization of products has led to increased competition over price and service level to customers. As a result, retailers differentiate themselves through bundling and loyalty programs as well as campaigns and promotions in terms of pricing. As for service level, the margin of differentiation includes after-sales support, returns program, ease of platform use in terms of omnichannel retailers and pure online retailers, improved delivery convenience, and availability of products consistently with less exposure to stock-outs. In addition, logistics providers have boomed with the evolution of omnichannel offering logistical support to retailers about packaging, storage, consolidated transportation, and last-mile delivery. This has also led to extra complexity in terms of information flow between the customer, the retailer, and the logistics service provider.

One of the most important aspects of service level is delivery time to customers, due to competition, customers expect their orders faster in the standard delivery option. For example, AMAZON, Walmart, and Instacart are now delivering within a two-hour window. This entails that orders with a quick delivery option should be available near the customer. In 2011, AMAZON issued a patent named speculative shipping [Bensinger \(2014\)](#) that uses real-time information from customers, along with collaboration with third-party suppliers, they can package the orders for unnamed customers before they make an order as an anticipation of their order beforehand. If the customer does not finally buy the order, they can circulate the packaged order to other customers at a discounted rate. Application of such a strategy would significantly reduce delivery time for customers without the need for more fast and expensive options of delivery.

It should also be stressed that even though omnichannel retailers enjoy higher market exposure to customers, there is a need for collaborative strategic planning, buying, transportation, assortment, and allocation. This is what has led to retailers adopting collaborative space sharing of inventory between traditional retail and other channels in some stores and regional hubs. As for buying, omni-channel benefits from the drop shipping concept where products remain at the supplier and are shipped when a customer makes an order reducing inventory space and the need to pre-purchase products from the supplier in some cases. Retailers also can benefit from the reduction of purchasing costs because of the increased quantity of buying for multiple channels. Omnichannel retailers can use channel integration by fulfilling online orders from the store's backroom. The order's item can be advanced from the online fulfillment center or served from store inventories. Retailers consider the implementation of "buy online pick-up at store" (BOPS) and "ship-from-store" (SFS) policies through the online channel to ensure a competitive response time that is also supported via the fulfillment center. SFS takes advantage of the existing physical network by turning certain store locations to ship-from points for online sales. The inventory for SFS can come from a BM store replenishment stock that happens to be available online as well. However, SFS requires online fulfillment operations in-store.

In addition, we attempt to answer a question about where the inventory of each chan-

nel should be allocated, and if they should be merged. The same question is valid for forecasting: should each channel be forecast separately or jointly? These questions face omnichannel retailers constantly as customers expect more than ever products delivered or available to them in a short lead time given the nature of the multi-channel environment they experience. Omnichannel retailers have to choose which products to offer on online platforms, which to offer in BM stores, and which to offer in all channels. [Chopra and Meindl \(2016\)](#) stated that retailers should keep slow-moving high-value items in a centralized location for online offerings to reduce inventory cost and risk of obsolescence and advance fast-moving items to stores. This is more significant when centralized fulfillment locations are far from BM stores. Furthermore, demand planning becomes more complex as the demand is generated from different streams, this poses the challenge of aggregating demand across channels and doing segregation afterward or having a separate division for demand planning for each channel. Demand planning has a direct impact on inventory volume, aggregating demand and inventory of both channels is a question yet to be investigated.

Omnichannel retailers are dealing with these challenges by advancing internet-based sales like home delivery or buying online and pick up from the store (BOPS) to physical store locations. This leads to a level of consolidation that would reduce transportation costs and at the same time reduce the delivery time to the customers, providing them with instant gratification and pick-up of the product at the time of their choice [Rooderkerk et al. \(2023\)](#). However, the preparation and the packaging of the order can significantly increase complexity if it is handled in a physical store instead of a fulfillment center with extra operations and backroom store space required. This operation is commonly known as a ship-from-store (SFS). Another variant is BOPS where the customer picks the basket from the store instead of home delivery. All of this has led retailers to ask the question of how to plan inventory for these integrated channels noting that research has been silent on forecasting in an omnichannel context [Rooderkerk and K ok \(2019\)](#), [Rooderkerk et al. \(2023\)](#).

With this challenging strategic planning, there is an additional dimension specific to the retail industry, customers tend to buy several items in a basket. Often times products are associated with each other, or they are competing with each other, and in some cases, they have no association or a distinctive pattern. This complicates the pre-packaging of online baskets for anticipatory shipping. There is also another possibility that part of the basket is available in a nearby store and the rest of the basket is in a centralized location, which leads to online orders suffering idle time and being delayed to customers if online fulfillment was made from the store. In addition, online orders are usually sent from a centralized hub where packaging capabilities are economically viable. However, this is done from a distance not close to every demand zone leading to high single shipments transportation costs due to irregular times of shipment and not benefiting from consolidation or economics of scale like the physical channel replenishment. Furthermore, although product associations and competitiveness were researched from a marketing perspective ([Berry and Linoff 2004](#), [Ghoshal and Sarkar 2014](#), [Kim et al. 2012](#), [Manchanda et al. 1999](#)) the information related to product association is lacking from the forecasting research stream.

With this in mind, the motivation behind this work is to fill the gap in research concerning basket data in the omnichannel supply chain. Basket data has always been a tool for marketing and sales operations to recommend products to customers. AMAZON and similar online retailers use basket association rules to recommend complementary products to customers and Netflix uses the same methodology to recommend movies to subscribers. In Both cases, basket data and product associations are used to derive more sales, and more convenience for users and customers. We aim to use product associations

from basket data to motivate more accurate forecasting in an omnichannel environment. The shopping basket notion has gained increasing importance in omnichannel retailing. It has become a crucial vehicle to capture sales through continuous promotional periods and various bundling offers. For instance, the data of a major cosmetics retailer used in this work reveals that, out of 2.2 million online orders over the year, 50% of them occurred in baskets with two distinct products or more, and there were 2.4 ordered products on average per basket. This raises the question of how data from multiple product baskets could be considered by operations managers to improve the accuracy of demand forecasting in omnichannel retailing. In an omnichannel environment, customers expect to order products either by visiting stores or online interfaces such as websites or third-party applications, and more often than not, orders made by customers are in bundles, and more than one item in an order, this motivates using basket data for forecasting.

## Objectives

This thesis aims to address these questions through the following objectives:

- Omnichannel forecasting and inventory performance using basket data. We intend to explore how basket data can be used to gain information that helps more accurate forecasting. Basket data can inform about associations between products and the competition between products among each other. It is more likely when a product has many substitutes that it will have fewer sales than a product with fewer substitutes. In addition, a product that usually comes in orders of size more than one is likely to have more sales than a product that usually comes in a single-item basket. We also aim to test forecasting the omnichannel by focusing on the BM sales and online sales of one region and see if there is an improvement in forecasting accuracy and inventory performance through channel aggregation.
- We aim to use basket data for anticipatory shipping and investigate if basket data can be used to predict basket content. Retailers face increased pressure to reduce lead time to deliver orders to customers. Utilizing store space for online orders for either home delivery or BOPS is a leverage that needs more specialized and dedicated advancement approaches that take into account order heterogeneity. We attempt to provide an approach to tackle this problem by looking at product associations.
- Basket data impact on forecasting aggregation and inventory performance. aggregation of products on a natural level or temporal level attempts to capture information hierarchies from all the levels of the series. However, information is lost during aggregation to upper levels while basket data is information that is compatible with many levels of the hierarchy for example; at the store level, there is basket data as well at the regional level. We aim to utilize basket data through a select of models to forecast store sales, regional sales, and top-level sales.

## Thesis content and outline

This work attempts to tackle these challenges in three different folds. First, for product association integration and omnichannel forecasting, we investigate developing a forecasting approach using basket data and graph theory that extracts attributes to predict product relations and interdependence. Graph theory is a modeling approach that is used to formalize the connections between items within a network and analyze the magnitude of their connectivity. Two variables are often used in the literature to characterize these

connections, namely: the degree of connectivity and the strength. These two variables can also be used to reflect the complementarity level of products since they measure the magnitude of connections between products that are often sold together within the same category or across categories. We also consider complementarity and substitutability for within-category associations. Following the recommendation of the M5 competition on the benefit of considering exogenous/explanatory variables when forecasting the demand in the retail context [Makridakis et al. \(2020\)](#), we consider the four cited above variables for the sales forecasting using baskets data in the omnichannel context.

In the second fold, for anticipatory shipping, we develop a two-step advanced inventory approach that develops a strategy for basket advancement near customers before an order is made, we also develop a profit model that captures all the related costs when applying such strategies, costs such as packaging cost, inventory cost, transportation cost, shortage cost, unpacking cost, and salvage cost. Unlike the first fold, here the approach tackles predicting the arc, not the nodes and thus graph theory is used to extract attributes related to the arcs (bi-product connection). Mainly three sets of attributes; *nodes to the arc* attributes, *arc related* attributes, and *probability/Bayes* related attributes. It is worth noting that to the best of our knowledge, no research explicitly considered basket data to forecast the content of a basket in advance. We endeavor to fill this gap in this work.

In the third fold, we investigate the impact of basket data attributes on forecasting aggregation. This has also not been investigated before in the literature and this work attempts to fill the gap. The work focuses on one city’s transnational data from our industry partner. Basket data are extracted and used as regressors with a statistical model and a machine learning model across different grouped-natural aggregation levels such as product/store level, product/region level, and aggregated sales level. The resulting forecasts are then simulated with inventory policy to measure inventory performance noting that we are among the first that measure the performance of hierarchical forecasting approaches on inventory performance in a retail context.

Finally, we give the conclusion and perspectives of the work and future avenues that could be explored to extend the work carried out in this thesis.

## Main contributions

The main contributions of this dissertation are listed as follows:

- **Contribution 1**
  - Building on findings from marketing and graph theory, we propose a novel forecasting approach for online and store sales that is driven by data on customers’ shopping baskets. It builds on a multi-category graph-learning approach from the basket data using an original set of cross- and within-categories regressors. Using this approach, four forecasting methods are proposed based on linear and polynomial regressions and machine learning. To the best of our knowledge, we are among the first contributors to the literature that considers the shopping basket in demand forecasting.
  - Through the analysis of more than 2 million orders in the online channel of a major cosmetics retailer, we characterize the demand patterns, and we provide empirical evidence of the dominance of cross-category complementarity and the high level of within-category substitutability of products in the market baskets.

- Based on a dataset of an assortment of 24,029 products in both the online and store channels, we empirically show the outperformance of four proposed forecasting methods compared to two benchmark forecasting methods, namely: Croston’s method and the autoregressive integrated moving average (ARIMA) method, which are often used in the retail context as shown in (Syntetos et al. 2010, Babai et al. 2013, Fildes et al. 2019). The outperformance is shown to be substantial for the class of lumpy demands due to the high connectivity of products. We also empirically show the benefit of joint forecasting of online and store sales. This holds for all the proposed forecasting methods.
- Through an investigation based on the order-up-to-level inventory control policy, we provide empirical evidence that using joint forecasting and shared inventory in an omnichannel context leads to a reduction of inventory shortages and an increase in the achieved service levels. This is shown by comparing the inventory performance of three fulfillment scenarios where the sales of both channels are forecasted separately or jointly and the inventory is either dedicated to each channel or shared by both channels. We are also among the first that consider channel hierarchical aggregation

- **Contribution 2**

- Building on findings from graph theory and Bayes theory, we propose a two-step approach to advancing baskets near customers before an order is made. In the first step, we suggest an arc prediction method instead of the traditional individual product forecasting method using three sets of attributes; arc attributes, nodes to the arc attributes, and probability/Bayes attributes. In the second step, we propose a policy and an algorithm based on the size of the basket, Baskets that contain two items or baskets of size 2 (BS2) are advanced consolidated and packaged while baskets of size more than 2 (BS>2) are sent not packaged. BS>2 consist of more than one arc which requires the algorithm to retrieve baskets from arcs and reshape arcs into baskets of different sizes.
- A multi-period profit model is proposed to evaluate the performance of the considered approach. The model is used to compare the performance of our proposed approach (TO-BE) and the baseline (AS-IS) with various costing components including packaging, inventory, handling, and transportation costs. Key performance indicators that evaluate the approach from a service-level perspective are also considered.
- Based on the dataset of baskets from one city from a major cosmetics retailer based in France. We formulate and run a multi-period simulation on basket advancements that is suitable for the proposed two-step approach as it uniquely tackles bi-products advancement which is not compatible with the traditional inventory policies such as (R,S) policies. The resulting outcome of the simulations is assessed through proposed fulfillment key performance indicators and a multi-period costing model. We also provide empirical evidence that arc-proposed attributes utilized by machine learning methods successfully predict arcs that will be ordered in the future and eliminate arcs that will not be ordered.

- **Contribution 3**

- Building on findings from basket data, cross-category attributes; degree and strength, and within-category attributes; substitutability, and complementarity, we assess the impact of the aforementioned basket attributes on hierarchical aggregation approaches.
- The forecasts generated are used to simulate inventory planning in stores and the warehouse of the region and then are compared for inventory shortage and inventory volume which to the best of our knowledge, we are among the first to to investigate inventory performance with hierarchical forecasting approaches.

# Chapter 1

## Problem statement and literature review

### 1.1 Introduction and Problem Statement

This chapter provides an overall academic perspective, the objectives of this work, and the steps required to meet them. First, we define some key terms that are used in the work in relation to graph theory, omnichannel, forecasting, inventory, and demand aggregation. Following that, we explain the business context, the research background, and we provide an overview of the research objectives. We elaborate on our contribution to the state of the art in chapters 3 to 5. The following section includes the definitions of the key terms that are used in this work.

#### 1.1.1 Definitions

Omnichannel

[Chopra \(2018\)](#) defined omnichannel retailing as using various channels to interact with customers and fulfill their orders. The interaction between a customer and a retail channel is primarily in terms of three flows – information, product, and funds. First, the retailer provides product and pricing information to the customer who then places an order. Following that, the order information is used by the retailer to move the product to the customer. Finally, payment is transferred from the customer to the retailer. From an operations point of view, the various channels offered to customers require integration of the supply chain in terms of transportation, inventory, packaging, and fulfillment.

Time series

[Durlauf and Blume \(2016\)](#) defined time series as any series of observations ordered along a single dimension, such as time, may be thought of as a time series. The emphasis in time series analysis is on studying the dependence among observations at different points in time. What distinguishes time series analysis from general multivariate analysis is precisely the temporal order imposed on the observations. In this work context, a time series refers to the sales or the demand for a product over time. These products can be categorized by store, by the fulfillment center, by channel, or by multiple channels aggregated.

A forecasting model

Hyndman and Athanasopoulos (2018) defined forecasting as predicting the future as accurately as possible, using all of the information available, including historical data and knowledge of any future events that may impact the forecasts. The models can be statistical or machine learning in nature, extrapolated, or causal Petropoulos et al. (2022). In this work, we rely on two statistical methods: ARIMA and Croston as a benchmark of widely used models in retail supply chains. As for machine learning and causal methods, in chapters 3 and 5, we rely on LightGBM as a top-performing model in machine learning methods as well as linear regression as a standard used model to capture relations between dependant and independent variables. In Chapter 3 where we predict basket content, multiple machine-learning models are tested and compared.

### Forecasting aggregation

Forecasting aggregation relates to the aggregation of several data time points. This can be a temporal aggregation such as aggregating daily demand data to weekly for a particular product. Moreover, the aggregation can be cross-sectional, for example in a retail context, it can be an aggregation of multiple product revenue under a specific category. Besides, it can be a natural aggregation by location for example (store/region/country/continent). In this work, we implement a grouped natural structure aggregation where each product is not aggregated at the store level, sales of each product are aggregated at a regional level Babai et al. (2022).

### Accuracy measures

Accuracy measures the errors of the forecasts where the latter is the difference between observed and predicted value. It is important to note that an error is not a mistake but an unpredictable part of the forecast Hyndman and Athanasopoulos (2018). Accuracy measures are used to explain how well a model predicts new time data points not fitted by the model. They are considered performance measurements that measure and compare characteristics in the forecast in relation to the resulting error. For example, mean error (ME) is used to calculate the bias of forecast if it is usually overshooting and optimistic in its prediction or undershooting and pessimistic in its prediction while mean squared error (MSE) is used to magnify the error.

Accuracy measures can be divided into three categories; the first category is scale-dependant errors which measure directly the error and this works well when all the time series have the same unit scale, volume, and frequency or when the comparison is between several models on the same time-series. From scale-dependant errors are the mean absolute error (MAE) which is equal to  $mean(|e_t|)$  and the root mean squared error (RMSE) which is equal to  $\sqrt{mean(e_t^2)}$  where  $e_t$  is the error at period  $t$ .

The second category is scale in-dependant errors which does not depend on the unit scale of the time series. as it compares a forecast prediction and a simple forecasting method from the training sample Hyndman and Koehler (2006), Hyndman and Athanasopoulos (2018). The widely used metrics in this category are the root mean squared scaled error (RMSSE) and mean absolute scaled error (MASE). The error for MASE is calculated as follows.

$$q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-1}|} \quad (1.1)$$

and accordingly the MASE is calculated as  $mean(|q_i|)$ . While the error for RMSSE is calculated as follows.

$$q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T (y_t - y_{t-1})^2} \quad (1.2)$$

and accordingly the RMSSE is calculated as  $\sqrt{mean(q_i^2)}$ .

The last category is percentage errors, where the error is calculated as a percentage of the observation. It is useful when forecasting performance is compared against different time series. The most common metric is the mean absolute percentage error which is equal to  $mean(|100e_t/y_t|)$ . However, this metric is prone to infinite results and is not interpreted when  $y_t$  at any period is equal to zero, and the Scaled mean absolute percentage error which is calculated as  $mean(\frac{200|y_t-f_t|}{y_t+f_t})$ , where  $f_t$  is the forecast at time  $t$ .

## Graph theory

Graph theory is a modeling approach that is used to formalize the connections between items within a network and analyze the magnitude of their connectivity [Gross and Yellen \(2005\)](#), [Dooley et al. \(2019\)](#). A graph is made of nodes and arcs. A node is an object that is considered in this work's context as a product, while an arc is a connection between two products. In this work's context it is two products that were bought together at least once in a basket. Furthermore, two variables are often used in the literature to characterize connections between two entities in the graph, namely: the degree of connectivity and the strength. In Figure 1.1, a sample of products is presented where nodes model products (A,B,C,D) and they are connected through arcs (A-B, A-C, A-D, B-D, D-C) as the graph is un-directed: arc A-C is equal to arc C-A, and arc A-C has a frequency of 2 which means that this connection appeared twice in two baskets.

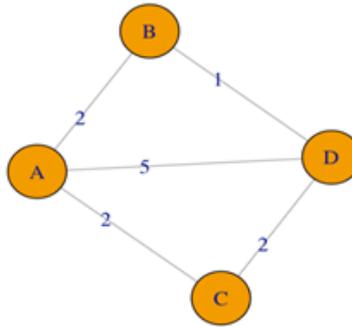


Figure 1.1: A graph of products

## Network horizon

In our context, "Network horizon" is the number of historical collective periods considered in a graph of products. The product or arc attributes are collected based on the time length (duration) of the graph. The attributes extracted from the graph are utilized as regressors with a forecasting model to predict products or to predict baskets.

## Inventory policy

An inventory policy is a policy that contains a set of thresholds and parameters that guide the quantity and the time of replenishment. There is a plethora of inventory policies used in industries and applied in the literature [Babai and Dallery \(2009\)](#). Some of these policies are periodic such as the periodic (R,S) policy which defines a review period at which a product is ordered to a defined level (max), and (R,s,S) policy at which a product is ordered every review period to a defined level (max) but only when the inventory position is below min at the time of the review period. Besides, a base stock policy is a periodic review period where a base is set and the quantity sold is ordered every time period. From the continuous policies where the stock level is always monitored, there is the (s,Q) policy, where a fixed quantity is ordered when the inventory position reaches s (min), also known as re-order point. Moreover, there is the (s,S) policy, where an offset between inventory position and max is ordered when the inventory position reaches min. We use the periodic review policy (R,S), in this work, as it matches the setup employed by the industry partner.

### Industry partner

The industry partner -hereinafter referred to as the retailer- of this work is a multi-national cosmetics retailer that has stores globally. It is based in France. The retailer sells a broad range of assortment that includes perfumes, skincare, lotions, and beauty products. The retailer not only sells its brand's products, but also brands from other third-party partners. They own and operate a dynamic and integrated omnichannel where products sold through their website can be stored in Brick and Mortar stores and where a portion of their offering is sold through third-party sellers. They also have exclusive products that are only sold in their physical stores and products that are sold online as well as products that are sold in both channels. This work uses retail transaction data of a particular city in France that is segregated by channel and investigates the contribution of this work to the provided data through empirical investigation. The retailer in France has a fulfillment where most online orders are sent to all cities in France and neighboring countries. Physical store replenishment is made through scheduled transportation deliveries from regional warehouse hubs.

### Transaction data

In this work's context, transaction data refers to the daily purchase data of online and physical stores. The data includes the date of the order, the products purchased within the order, and their quantities, the order number and description of the products. Each store has its transaction data, the daily transactions data. Besides, the daily transactions can be aggregated into daily, weekly, and monthly as a form of temporal aggregation, and also, can be aggregated by city, cluster, and country as a form of natural aggregation. Finally, it can be aggregated by channel; BM and online channel.

### Cycle service level (CSL)

Cycle service level is a performance metric that evaluates the performance of the inventory policy. The CSL relates to the number of periods where demand was satisfied in full from the total periods of the simulation [Babai and Dallery \(2009\)](#).

### Profit model

A profit model is a model that includes all the related costs and revenue set up for a specific order fulfillment setup [Meixell and Gargeya \(2005\)](#). Costs include for example the cost of packaging a product either in a fulfillment center or a store, the cost of handling, the cost of transportation of an order either consolidated or sent by freight forwarding, the penalty cost of not having the order on time, and inventory cost for keeping a product either in fulfillment center or store. The model serves as a comparison approach for different fulfillment scenarios as explained in detail in Chapter 4. As for the revenue in the profit model, it relates to the selling price of a basket as well as the bonus resulting from having the order available on time for a customer.

## 1.2 Research background

In this chapter, an overview of the literature is provided that draws on research in the areas of omnichannel retailing, demand forecasting for retail, forecasting aggregation, inventory performance in retail, and anticipatory shipping.

### 1.2.1 Introduction

The retail industry is a broad industry that includes many different functions and utilities. It ranges from electronics retail, consumer goods retail, and specialty retailers like Ikea, Home Dept, and Zara to small convenience stores and kiosks that sell snacks and desserts. Retailers provide a wide range of services: they can enable access to information, gives products' accessibility to customers, promote products of suppliers, organize the assortment for customers, allow for product timely introduction to the market, and create a presentable and acceptable user experience for customers [Gauri et al. \(2021\)](#). Retailers that provide an assortment that serves the same utility to customers differ in their service offering which justifies a customer's willingness to pay an extra margin against the extra service offering. For instance, Tesco provides a nice ambiance for customers, a tidy assortment organization, and accordingly a little more extra margin than COSTCO which depends more on the wholesale price offering and bulk buying. Retailers have come a long way since the middle of the nineteenth century. In fact, retailers used to be small stores that were owned as family businesses where crops and general merchandise were sold. Then, department stores were introduced in the nineteenth century followed by malls and super stores, limited assortment stores and convenience stores. The biggest revolution in retail in the last thirty years is the commencement of online retailing using the World Wide Web. Retailers, before the internet, used to reach the customer through a multi-channel approach. They often extend their reach to customers either through catalogs that are delivered by mail or through delivery by phone. On the other hand, show-rooming stores also gained popularity, customers visit the show-rooming store and check the products which can be delivered to them by home delivery.

As for the online channel evolution, it already evolved significantly in the last thirty years, with the flourishing of AMAZON in the online retail space. Other retailers, also, started to focus exclusively on selling online through websites such as (E-bay). where they sell products not necessarily new, it can be used or antiques through the general public and the notion of marketplace started to evolve. After that, pure online stores were introduced such as Blue Nile, which sells jewelry at a considerably reduced price, and Indochino which sells customized suites ([Chopra and Meindl \(2016\)](#)). The online space, also, availed non-retailers to craft products and merchandise through specialized crafting marketplaces like ETSY ([Chandna and Salimath \(2018\)](#)). In addition, online sales

integrated solution provider, like Shopify, allows entrepreneurs to sell merchandise they make or consolidate through their own customized website. In the food retail business, the emergence of food aggregation applications significantly impacted the way restaurants and grocery stores deliver products to customers and served as a complimentary marketplace to restaurants like Delivery Hero, Deliveroo, and Uber Eats [Farah et al. \(2022\)](#). Food aggregators became much more popular such that shadow kitchens (multiple brands in one kitchen) that have multiple brand labels serve customers only through food aggregators. Specialty retailers and original equipment manufacturing (OEM) started to take notice early on in the online revolution and online stores such as DELL, FORD, and APPLE started to establish their online presence. DELL even shifted its strategy drastically from 2005 to 2012 to sell online their laptop assortment online exclusively through its online platform in the United States following a make-to-order strategy where a customer can customize their laptop and receive it via home delivery within 3 to 7 days [Chopra and Meindl \(2016\)](#). [Gauri et al. \(2021\)](#) presented a summary of the evolution of retail and the integration between physical stores and online stores in Figure 1.2. This has subsequently become omni-channel retailing.

While there is a constant argument that the enormous boom of online retail is here to stay and that there is an impressive growth of online sales year by year, BM is still here to say contradicting some consensus that it might disappear soon [Gao and Su \(2019\)](#). Besides, physical retail stores offer a tangible ambiance that online stores do not offer, for instance, big stores like FNAC in France, IKEA offer food court and coffee shops inside their stores which offer a sort of social shopping to customers.

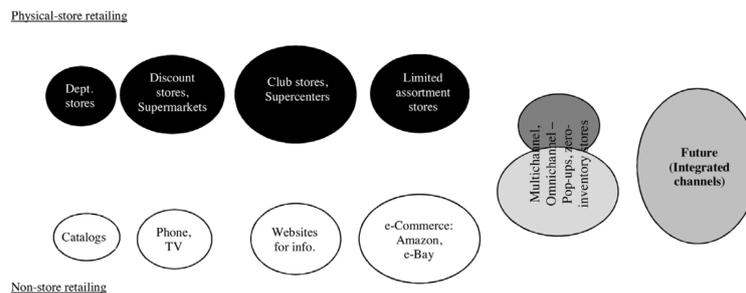


Figure 1.2: different retail formats

## 1.2.2 Omni-channel retailing

The term omnichannel was introduced by [Strang \(2013\)](#) as a boundary-less retail experience to customers where a customer can research a product online or through a catalog, mobile applications, showrooms, or physical stores and then decide to buy through one of these channels. This is congruent with most of the definitions referring to omnichannel retailing as the use of a variety of channels to interact with customers and fulfill their orders ([Bell et al. 2014](#), [Chopra 2018](#), [Marty 2022](#), [Jasin et al. 2019](#), [Melacini et al. 2018](#), [Cai and Lo 2020](#)). [Gallino and Moreno \(2014\)](#) are among the first to cover the integration of online and offline channels in retail. They discussed the important effects of the implementation of a buy online pick up from store (BOPS) channel, especially the cross-selling effect and the channel-shift effect. Gradually, Brick & Mortar BM started to gain market entry and have an online presence, which led to gradually blurring the line between BM and web shops [Brynjolfsson et al. \(2013\)](#), [Hübner et al. \(2016\)](#). It is evident that the emergence of online channels led to direct competition with physical stores. However, it, also, presents an opportunity for expansion of market share [Goedhart et al. \(2023\)](#). Recent contributions on network design ([Arslan et al. 2021](#), [Guerrero-Lorente et al. 2020](#)), inven-

tory management (Saha and Bhattacharya 2021, Hu et al. 2021) and fulfillment (Gao and Su 2016, Bayram and Cesaret 2021), showed the benefit of using ship-from-store (SFS) channel and/or of using BOPS channel. Such benefit was also discussed in Gallino et al. (2019) and Jasin et al. (2019), where the authors argued that adapting an SFS model can increase the service level perceived by customers.

Goic and Olivares (2019) reported that in a data-driven omnichannel framework, the boundaries between online and offline data are disappearing and the methodologies to analyze these data are converging. From a consumer point of view, Ailawadi and Farris (2017) underlined the heterogeneity in consumer reasons for buying online and the unknown path to purchasing across channels, which is challenging from a forecasting perspective.

In the last years, several studies in the literature rely on empirical studies to explore this novel topic.

Hübner et al. (2016) provided an exploratory study with 33 retailers and reported that inventory integration and allocation, are among the most important areas for fulfilling distribution requirements. Ishfaq et al. (2016) underlined the role of store-based retailing on the omnichannel fulfillment strategy, which would be improved by forward placement of inventory. Gallino et al. (2016) empirically studied how the deployment of ship-to-store policy at a major US retailer generated sales dispersion and affected inventory-ordering models with an increase of cycle and safety inventories. Similarly, Gallino and Moreno (2014) studied the importance of sharing reliable available inventory information with customers when channels are integrated. All these empirical investigations highlight the need to integrate store and online channel operations and to investigate cost-service trade-offs when channels are planned and operated jointly.

Furthermore, Rooderkerk and Kök (2019) appealed to revoking the classical predictive and prescriptive modeling approaches in an omnichannel context that all work by a separate demand stream and by location (stores or fulfillment centers). The authors reported that academic research has been largely silent on the challenges of forecasting the demand streams of the various types of omnichannel flows. To the best of our knowledge, there is no research work in the literature that considers forecasting and inventory planning to operate an omnichannel distribution network under online and store demand.

Asmare and Zewdie (2022) argued that omnichannel retailing has made the customer touch points more holistic. Jasin et al. (2019) segregated the types of retailers into five categories; a single channel retailer, for example, Shopify sellers, the retailers of this type have a big opportunity to optimize their supply channel. The multi-channel retailer runs a separate channel for each business (almost). For example, Walmart started an online business where they had a dedicated warehouse for online baskets. They argued that complexity is low as they are separate channels but they can achieve economies of scale with aggregate buying. Cross-channel where the supply chain is integrated between multiple channels this is abundantly evident when a customer orders a basket online and picks it up from the store or the retailer chooses to send an item to a customer directly from the store and finally an omnichannel retailer which is a completely seamless experience where the customer can surf online, check the product in a showroom and have it delivered it to home. As most retailers now are moving towards an omnichannel environment, delivery time has become more of a determining factor in retailer selection. AMAZON was one of the first movers through AMAZON Prime to offer 2-day delivery. Other retailers started to compete within the same arena. For example, Target started to offer 2-hour delivery (Jasin et al. (2019)). they attempt to achieve reduced lead-time time slots by utilizing store space or establishing mini hubs near expected demand zones.

When a customer purchases an item online, the item's processing takes a different route than the Brick and Mortar fulfillment. More Generally, when an omnichannel retailer has

a strong online retail business, and a dedicated fulfillment warehouse is established and designed to be as efficient as possible. In such cases, online orders would be shipped from online fulfillment centres to customers. Activities of online fulfillment include processing order requests, picking, sorting, and packing. When these activities are made inside a dedicated online fulfillment warehouse, the opportunity of the economics of scale is achieved. However, there is a high risk of delaying customer orders could be distant from potential demand zones. Having online orders stored in a store near a customer location in large quantities risks interrupting the shopping experience of walk-in customers. In addition, shipping from the store requires manual picking operations that divide the tasks of store staff and take extra backroom space that not all retail stores might have. There is also the constraint of multiple items in a basket and not all of the requested items being available near the demand zone which would potentially lead to a delay in order to the customer. However, generally, shipping from the store lead to reduced shipping cost due to demand pooling. A plethora of models were developed to deal with this dynamic allocation problem. For instance, [Mahar et al. \(2009\)](#) developed a dynamic model using real-time data. They argued that having a dynamic assignment may reduce inventory cost, back-ordering cost, and transportation by 8.2 percent. [Jasin et al. \(2019\)](#) developed a split assignment linear deterministic model where orders allow for multiple items in a basket. [Gallino and Moreno \(2014\)](#) are among the first to cover the integration of online and offline channels in retail. Recent contributions on network design ([Arslan et al. 2021](#), [Guerrero-Lorente et al. 2020](#)), inventory management ([Saha and Bhattacharya 2021](#), [Hu et al. 2021](#)) and fulfillment ([Gao and Su 2016](#), [Bayram and Cesaret 2021](#)) showed the benefit of using SFS channel and/or of using BOPS channel. In terms of assortment planning, customer behavior has shifted dramatically with the emergence of digital and social media which consequently affects the approach retailers offer and profile their assortment in each channel, the models of [Rooderkerk and K ok \(2019\)](#), [Bijmolt et al. \(2019\)](#) for the customer journey is adapted and it depicts a traditional omnichannel customer journey from recognition to evaluation or review of order.

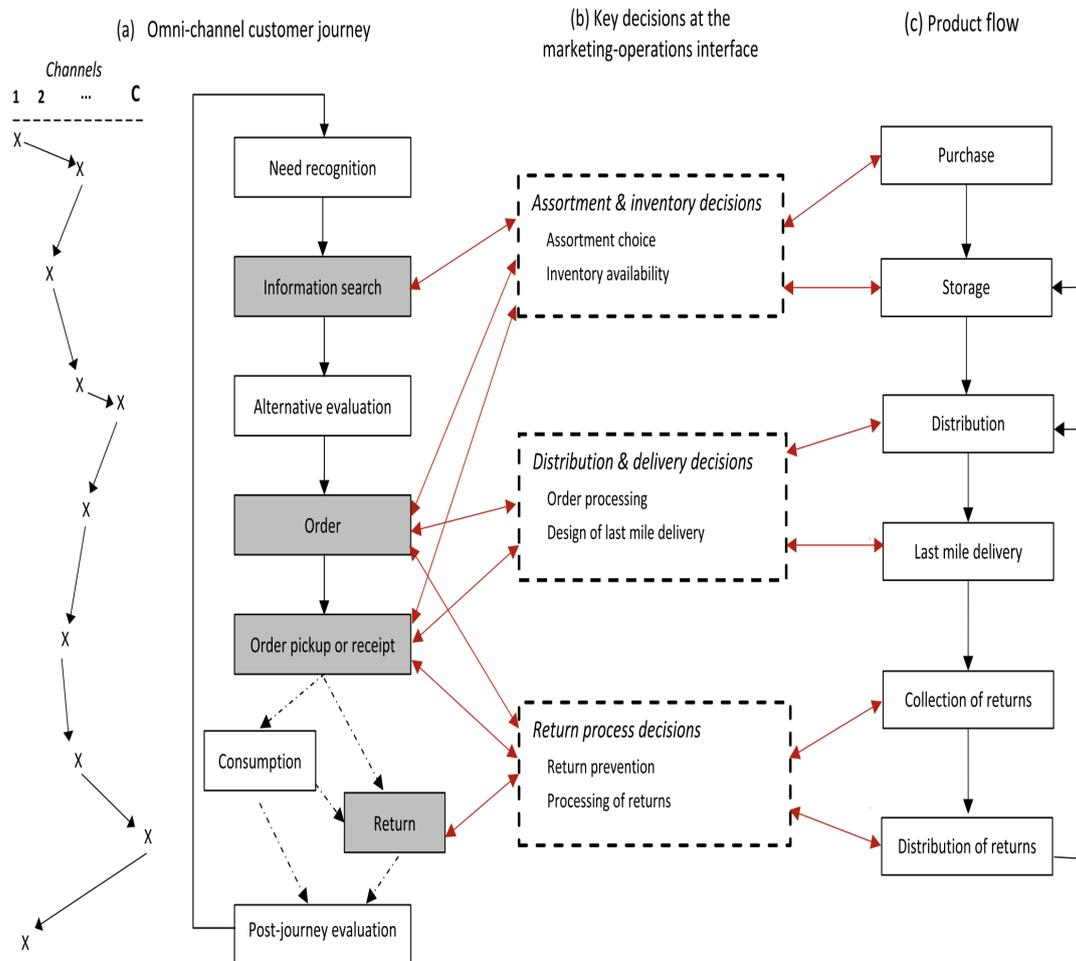


Fig. 1. Connecting the omni-channel customer journey to the product flow: Key decision areas at the marketing-operations interface. Panel (a) depicts the omni-channel customer journey: consumers migrate between channels throughout their journey, between and within stages. Panel (c) depicts the firm's product flow. The two panels are interconnected through firm decisions that affect both journey (panel (a)) and product flow (panel (c)). These decisions, indicated in panel (b), are located at the marketing-operations interface. The red bidirectional arrows indicate how the decisions interact with the customer journey on the one hand (between panel (a) and (b)) and with the product flow on the other (between panel (b) and (c)). The dotted-dash lines in panel (a) indicate that multiple paths are possible. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 1.3: Customer journey

In their flow model, an illustration is made between the interaction between the customer journey, the decisions that the retailers make in reaction to the customer, and the respective product flow. Central panel b demonstrates the decision areas that relate to the customer journey in panel A and the product flow in panel C.

In the customer journey (panel C), the retailer has to ensure a seamless experience for a customer from his/her need of recognition, either by targeted advertising or other media of advertisement. Followed by the previous options, a customer could consider gathering information about the product either by web search or physical showrooming or tutorials ...etc. Followed by price comparisons either by similar products from the same retailer physically or through their online website or from price comparison websites. Food aggregators are well-established examples of price comparisons. The ordering phase can be through the BM store, the retailer's website, the retailer's application, and third-party sellers among other options. Picking up can be at a designated point, at a store, or third-party logistics provider's office among others. Finally the return of the product is either at a specific designation, near the store, or via third-party partners. As shown, each decision in the customer journey can have multiple channels to choose from.

In the case of product flow, which is facilitated by the operations and design team, a

customer has the ability to buy an item through digital and social media, or by visiting the store. This step starts with an order step by the customer, the retailer previously-stored the items of the order in some location and the item is distributed near a demand zone of the customer if not already available in the demand zone via a store or if the retailer does not support ship from store capabilities in this location. Then, the last mile delivery is usually made by a third-party logistics provider in case the customer does not pick it up followed by the return process and distribution of return in case of defect or other policy allowances.

Finally, panel B which is the retailer marketing and operations decisions, note that these decisions affect both the customer journey and product flow, which summarizes into three key decisions: (1) inventory and assortment decisions, (2) distribution and delivery decisions which include order processing, and design of last-mile delivery, and (3) return process decisions. The red arrows illustrate at which stage the marketing/operations coincide with the customer journey and product flow. Interestingly, the retailer choice with regard to assortment design would dictate where the product is shipped from. If the products inside the basket are not included in the BM assortment, it is shipped from the online fulfillment center.

In the assortment choice, the breadth and the depth of the assortment have a direct impact on retail operations and the depth has significantly increased with the advent of online retail operations [Rooderkerk and K ok \(2019\)](#), [Bijmolt et al. \(2019\)](#). As the depth of assortment increases, the inventory volume naturally becomes more complex and potentially increases. In addition, the chance of a particular product not being available in multiple locations increases, and reduces the flexibility of the retailer to offer a complete basket near the customer. While for the customer it can cause overload and stress. On the other side, marketing and sales derive higher depth and breadth of assortment for the sake of an increase in customer retention and sales. Furthermore, to provide exclusivity for each channel and eagerness by the customer to check each channel. Retailers might overcome the burden of keeping extensive depth by having consignment options by which the retailer keeps the inventory but only pays for it when it is sold. Hence, operations and marketing decisions are jointly established for inventory and assortment decisions.

Consequently, retailers reduce inventory overload especially if they accept products from multiple suppliers to offer a drop shipping strategy for online retail fulfillment [Cheng et al. \(2016\)](#) or by consignment options which free cash flow for retailers, although it increases space utilization. However, depending entirely on the supplier's assortment might cause retailers to lose control over the supply chain. This is why some retailers offer their own assortment side by side with the supplier's offering. A lot of examples are in this space such as AMAZON and CARREFOUR. In general, more depth provides an opportunity for up-selling while more breadth provides an opportunity for cross-selling [Bijmolt et al. \(2019\)](#). In terms of inventory, the choice of which strategy to adopt with suppliers implies the inventory and operational costs that can be incurred. The retailer owning the inventory would increase the inventory and operational cost but would provide more control over where to allocate the inventory. Selling on a consignment basis provides flexibility to the retailer on where to allocate products as well as dynamically assigning products to home delivery, delivery from the fulfillment center, and BOPS. In terms of locating inventory positions for each channel, there was a lot of research on the subject [Ishfaq et al. \(2016\)](#), [Goedhart et al. \(2023\)](#). [Govindarajan et al. \(2018\)](#) considered inventory and fulfillment optimization on a retailer network facing two demands—online and in-store for a single product—but demand by channel was kept separate. [Gao and Su \(2019\)](#) illustrated that the recent trend is to move more inventory to the store for seamless omnichannel integration and to offer BOPS. In addition, offering BOPS increases the sales area of the retailer and offers the customer more opportunity to check other

products in the store [Gallino et al. \(2017\)](#). [Ishfaq and Raja \(2018\)](#) demonstrated that it is more efficient to pool demand in the fulfillment center than to send from stores due to extra time suffered by staff costs for picking and packing.

As for distribution decisions, traditionally the BM replenishment is sent with consolidated shipments with a full truckload and less truck load and this depends on the size of the store and the quantity replenished by the supplier if it is a direct delivery from the supplier. The strategy of the retailer defines greatly how the design is implemented. [Chopra and Meindl \(2016\)](#) argued that specific customer attributes are defined by the distribution strategy; 1) response time, 2) product variety, 3) product availability, 4) customer experience, 5) time to market, 6) order visibility, and 7) returnability. On the other hand, operational attributes are 1) inventory 2) transportation, 3) facility and handling, and 4) information. Figure 1.4 shows the performance of each design with the attributes that impact the customer journey 1 being the strongest performance.

	<b>Retail Storage with Customer Pickup</b>	<b>Manufacturer Storage with Direct Shipping</b>	<b>Manufacturer Storage with In-Transit Merge</b>	<b>Distributor Storage with Package Carrier Delivery</b>	<b>Distributor Storage with Last-Mile Delivery</b>	<b>Manufacturer Storage with Pickup</b>
Response time	1	4	4	3	2	4
Product variety	4	1	1	2	3	1
Product availability	4	1	1	2	3	1
Customer experience	Varies from 1 to 5	4	3	2	1	5
Time to market	4	1	1	2	3	1
Order visibility	1	5	4	3	2	6
Returnability	1	5	5	4	3	2
Inventory	4	1	1	2	3	1
Transportation	1	4	3	2	5	1
Facility and handling	6	1	2	3	4	5
Information	1	4	4	3	2	5

Key: 1 corresponds to the strongest performance and 6 the weakest performance.

Figure 1.4: Different distribution design - [Chopra and Meindl \(2016\)](#)

The BM Retailers fall in the first category "Retail Storage with Customer Pickup" where a customer goes to the store and buys products such as Walmart and Target. Response time to the customer is highest but with extra hassle cost on the customer, product availability and variety is low as it is constrained by the shelf space, time to market is long as it is far from the source, order visibility is high as the customer can instantly see the product in stock or out of stock and Returnability is considerable easy for the customer. Interestingly when Walmart decided to enter the online space, established a "Distributor strategy with last-mile delivery" along with the already existing BM supply chain. AMAZON established a supply network that includes three different distribution networks; "Manufacturer Storage with Direct shipping" for their drop shipping business where a customer orders an item that is stored by agreement at the supplier premise, AMAZON

processes the order, and then the supplier ships it to the customer. If a customer orders an item and there is no local fulfillment center, last-mile delivery is not an option and the item is sent by carrier delivery to the customer. Finally, AMAZON keeps the usual in-demand assortment in the local fulfillment center following a "Distributor Storage with Last-Mile Delivery". DELL, when they were selling desktop computers, they were following "manufacturer storage with in-transit merge" when they were compiling the CPU manufactured by DELL and the monitor by SONY [Chopra and Meindl \(2016\)](#). Notably, with the current dominance of omnichannel retail, it can be observed that a single-channel retailer applies one distribution network while a multi-channel or omnichannel retailer applies more than one distribution network. [Lim et al. \(2018\)](#) illustrated that there are three categories of last-mile delivery for omnichannel retailers: a push last mile delivery system, where the customer plays no role in the product flow of the delivery process and delivery is handled either by the omnichannel retailer or a third-party logistics provider. The second category is a pull last mile delivery system where a customer makes the effort of going to the store or directly to pick-up point. Usually, in this scenario, transportation cost is saved due to customer contribution. Finally, a hybrid system where the omnichannel retailer or the logistics provider ships the products to a hub point near customer's location and the customer picks the item from the location. This strategy is a balance between the push system and the pull system. However, the contribution of the customer does not result always in savings of cost for the total supply chain cost due to schedule conflicts that might hinder the customer from receiving the product at home or picking it up from a named location which results in returns or shipping the products again.

For returns decisions, which are the last marketing-operations in panel b ([Bijmolt et al. \(2021\)](#)), it is a major factor in supply chain cost, especially in the online channel from the omnichannel as customers do not have the chance to try the products if it was apparel or fashion. In addition, the online sales channel depends heavily on third-party suppliers the product is not entirely controlled by the retailer. Any retailer, if they have a return policy, determines when the customer could return a product, for what reasons, and the method of this return. In general, a customer plays a part when the customer returns a product, the customer might go to a nearby store, or package the item and leave it at a destination point, or pay for the whole delivery to the main hub of the retailer. In a study made by [de Leeuw et al. \(2016\)](#), they showed that consumer returns could reach 50 % of the fashion market. The location of where the retailer collects products can vary supply chain costs significantly but it can also from a marketing perspective be a determinant if a customer would purchase from the same retailer again. For instance, a customer would prefer a returns location that is more accessible and has fewer returning procedures. [de Leeuw et al. \(2016\)](#) argued that returning a product to a store from an online purchase, particularly from omnichannel retailers is appreciated by customers and might facilitate the quick return of an item to be in stock and assortment again. Items returned can either be recycled and enter into the making of other products, refurbished and sold in the secondary market which is a big market considering the amount of returns and finally it can be re-sold if it was returned for reasons other than malfunction. In general, it is particularly effective to return the product quickly back to the supply chain if the product is intended to be resold as products lose value over time especially if they are resold.

### 1.2.3 Anticipatory shipping

Anticipatory shipping is a recent discipline of research that closely impacts omnichannel retail, it can be described as the planning and partial shipping of baskets close to the customer in anticipation of a customer order. The first notable literature in this area, to

the best of our knowledge, is the research and development made by AMAZON. [Spiegel et al. \(2014\)](#), [Pöchhacker and Nyckel \(2020\)](#) introduced a concept that suggests real-time tracking of customer preferences on their website, collaboration with suppliers, and use of quantitative models to package in advance a basket to the customer demand zone without a name tag or a specified address. So, as when the customer actually orders the basket, it is delivered within hours to his/her home, in case the basket is not eventually bought by the him/her, it is promoted at a discounted price to other customers. This strategy would allow for same-day delivery and next-day delivery for a lot of baskets that had products in a distant fulfillment center or at an OEM supplier. Such a strategy when implemented would increase the retailer's competitiveness. For example, consumer goods that are similar in utility and approximately similar in price, such similarity would give room for service level, bundling, and pricing to differentiate a retailer.

The work on anticipatory shipping is rather limited and still in its early stages. [Lee \(2017\)](#) proposed a genetic algorithm optimization model complemented by baskets data association rules to anticipate the basket in advance on clustered historical data. A close approach was proposed by [Viet et al. \(2020\)](#), where they used basket data association rules with the addition of a time constraint for perishable food items for the agri-food industry. Baskets data are heavily used by retailers, especially in marketing. It is a vital input to increase basket value and size through investigations of associations between different products of the retail assortment [Manchanda et al. \(1999\)](#), [Cachon and Kök \(2007\)](#). Walmart entered a strategic partnership with JingDong to provide customers with a one-hour delivery service. The customer can order online through the Walmart website, and then JingDong goes to the nearby physical store, picks up the items, and delivers them to customers [Ren et al. \(2023\)](#). For a purely online retailer, anticipatory shipping can be achieved by advancing baskets either packaged or not packaged to a close point near the demand zone. While omnichannel retailers can advance online baskets to nearby stores in the back room to avail one-day delivery and next-day delivery. From the recent research on the omnichannel, there is the ship from store concept (SFS), [Goedhart et al. \(2023\)](#) focused on the use of physical stores for online fulfillment as it can significantly reduce lead time to customers. This strategy would allow BM retailers to compete against large online retailers. They developed an exact solution approach to solve the online allocation of baskets to physical stores as well as formulating a model that incorporates revenue and costs realized from such operations.. However, in [Goedhart et al. \(2023\)](#) proposed to satisfy the demand for online baskets in advance, not to anticipate customer baskets in advance and ship them to physical stores. [Xiao et al. \(2009\)](#) were among the first to propose fulfillment of online orders through stores, they used dynamic programming to assign inventory to stores compared to a static policy where inventory is automatically allocated to the closest store. They found that the static policy leads to an imbalance of inventory and would cause revenue loss while the dynamic programming policy balances inventory. The dynamic problem becomes more complicated as the number of stores increases and it becomes more computationally expensive, this suggests that the use of heuristics is recommended for online order allocation to store problems [Bretthauer et al. \(2010\)](#)

Furthermore, [Ren et al. \(2022\)](#) introduced a forecasting algorithm that recommends items to be advanced based on product demand data with individual product-by-product demand prediction. This strategy works well in availing baskets if they are single basket items. They proposed an integrated approach to advance products to nearby customer hubs. In addition, [Lee \(2017\)](#) proposed a Genetic algorithm optimization model complemented by baskets data association rules to anticipate the basket in advance on clustered historical data. Baskets' data are heavily used by retailers, especially in marketing, associations between products provide a variety of insights for marketing and operations in

online and BM retail.

Mainly, through the use of association rules and recommendation algorithms, retailers like AMAZON and ALIBABA propose to customers products that complement the basket of a customer through similar historical purchases. Weingarten and Spinler (2021) proposed a customer-product anticipatory shipping approach where they were provided with customer clicks data. A classification model is developed to predict if a customer is likely to buy a product. This effectively focuses on customer real-time click stream data and the prediction model is customer-product based which is different from the general allocation problems where product demand and inventory are concerned.

In conclusion, anticipatory shipping is a growing field of research still in its early stages, it can prove vital to the success of omnichannel retailers to compete with large online retailers who have large fulfillment centers across the globe, particularly since online retail is growing rapidly, especially after the COVID-19 pandemic, where the growth of online sales grew at a much higher rate than physical retail. We can divide the existing literature on anticipatory shipping into two folds, a fold that focuses on the dynamic allocation of inventory that is targeted to be consumed by traditional walk-in customers and customers who order online and a fold that focuses and advancing a particular set of products exclusively intended for online sales and anticipation of customer order. It is worth mentioning that such strategies would lead to overage cost if the advanced baskets are not required by customers leading to discounted sales or obsolescence, and underage costs if the advanced quantities are less than what is requested by online customers or if the advanced basket is incomplete.

In Omar et al. (2022), the dimension of quantity (frequency) and the time frame were addressed through modeling historical basket sales as a graph where products are nodes and association frequency are arcs. The features extracted from the graph are then used as relational features to the products that improve the forecast of individual products in an omnichannel context. Such methodology will, also, fail to predict basket content as it focuses on the nodes of the graph and improves the forecasts of individual products. Accordingly, in Chapter 3, we divert the focus on the arcs of the graph and the nodes related to these arcs. This is because every arc represents a bi-product combination that has a cretin frequency from previous purchases which can be forecast and each graph represents a time frame of historical bi-product combinations. In addition, modeling the arcs of the graph presents an opportunity to extract the historical basket size of bi-product combinations. For example, one arc can likely be purchased in a basket of size 2, while another arc is usually purchased in a basket of size 6. As a result, the basket size for each arc historically can be added as lags to the arc. Furthermore, addressing basket forecasting as an arc modeling approach (bi-product combinations) instead of node (product) modeling will allow for probability inference measures related to the arc with relation to the entire graph structure which proves to be highly significant to the predictive power of machine learning models as explained later on in empirical analysis.

#### 1.2.4 Demand Forecasting in Retail

A plethora of forecasting methods have been studied since the 1980s (Syntetos et al. 2016, Petropoulos et al. 2020). Such methods include simple extrapolative methods, such as exponential smoothing and moving averages (Gardner Jr 1985, Svetunkov and Petropoulos 2018), ARIMA-type models (Gilbert 2005, Babai et al. 2013), machine learning methods (Zhang and Qi 2005, Punia et al. 2020), and judgmental methods (Petropoulos et al. 2016, 2018). A considerable amount of research work has been dedicated to analyzing and comparing the performance of such methods through empirical investigations and international forecasting competitions using supply chain data (Petropoulos and Makridakis

2020). The M5 competition was built on the case of a Walmart retail supply chain with more than 30,000 products (Makridakis et al. 2020). It showed the outperformance of machine learning methods compared to standard statistical forecasting methods, especially the LightGBM-based method. The M5 competition has also shown the ability to improve the accuracy of demand forecasting by considering exogenous/explanatory variables. The inclusion of such variables in the forecasting approach as well as the machine learning methods will be part of this research.

Demand forecasting is considered the starting point in many retail and non-retail, strategic, tactical, and operational decisions. Retail companies rely on forecasts across different dimensional spaces, and the position in the supply chain is in the store, in the distribution center, or at the OEM. The product hierarchy level such as product /brand/ section /subfamily/family, such level is important for communicating with suppliers as the granularity level and time aggregation. For instance, a forecast of stores is usually provided in days and in distribution centers at the weekly level Syntetos et al. (2016). Inaccurate forecasts cause a lot of problems for retailers, it may lead to out-of-stock periods and accordingly, loss of sales, loss of customers for future purchases, excessive inventory costs, and Bullwhip effect Lee et al. (1997).

In addition, retail sales are hardly stable especially when products are seasonal or if the retailer applies different kinds of promotions, if there are unseen events and pandemics. For example, during the COVID-19 pandemic, BM sales suffered considerably especially for luxury items while everyday food and non-food products suffered from inventory shortages due to high spikes in demand. Besides, online sales boomed, which led pure BM retailers to seek online channels to compensate for the loss of sales in BM retail. Furthermore, Fildes et al. (2019) argued that retail sales at the product and daily levels are characterized by a high degree of intermittence, that is, frequent zero sales. This demand characteristic attracted a considerable amount of research, and several forecasting methods were proposed to deal with demand intermittence. Syntetos and Boylan (2005) developed a commonly used framework to categorize demand into four categories; smooth, lumpy, erratic, and intermittent. For a recent overview of the research dealing with intermittent demand, readers are referred to Nikolopoulos (2020).

Another problem that exists in retail sales, is the bias forecasting in out-of-stock periods. In this case, the demand series do not show if the item has no sales due to out-of-stock instances or customers simply not wanting to buy the products. One advantage of causal methods is that they allow taking into consideration factors such as promotions, out-of-stock instances, and holiday sales.

Although the sales of a particular product in the same area via online and BM should exhibit the same pattern, there are extra features that can be extracted from the online channel that are not available in the retail channel such as click stream data and baskets previously purchased and association between baskets. With this dynamic interaction between online, BM and the BOPS, a lot of questions arise about aggregate forecasting of channels also known as omnichannel sales or forecasting each channel separately Omar et al. (2022). Promotions by itself play an important role in deriving sales and footfall for BM or click volume in terms of online. Retailers usually have a promotional calendar where they plan for long to medium-term different kinds of promotions. This can be done along with pricing and product associations while doing forecasting. For example, there are products that complement each other like a tortilla and tortilla sauce, and products that are competing with each other like two different brands of the same product like chocolate spread. In fact, a chocolate spread and a peanut butter spread can be complementary products and competing products based on customer's behavior and preference Manchanda et al. (1999), Cachon and Kök (2007). Accordingly, promotional models are used to forecast product data Ma et al. (2016), Trapero et al. (2015).

In the omnichannel retail context, [Armstrong \(2016\)](#) argued that forecasting strongly depends on the retailer’s omnichannel strategy, noting that both forecasts of the online orders and sales at stores are important for fulfillment decisions. This issue can be addressed through hierarchical forecasting. [Fildes et al. \(2019\)](#) presented an overview of the practice and research of retail forecasting. They reported that, at the product level, the time horizon of forecasting changes across the supply chain from quarterly to monthly to daily depending on the location (e.g., monthly forecasting at the distribution center level and quarterly forecasting at the factory level). They highlighted the complexity of product forecasting due to different physical attributes such as color, size, and packaging. They also highlighted the importance of product mix across product categories and they recommended that category management starts with forecasting the category level with a judgmental approach, taking into account inter-category purchasing behavior or the product mix. [Ma and Fildes \(2021\)](#) presented a meta-learning framework based on deep convolution neural networks to forecast retail sales. Based on weekly data of a grocery and drug chain related to a sample of 50 stores and 30 product categories, they showed the superior forecasting performance of the proposed meta-learner. They recommended, for forecasting retail sales, building a pool of base forecasters using both individual and pooled forecasting methods, to target finding the best combination of forecasts instead of the best individual method. [Shang et al. \(2020\)](#) showed that online retailers, especially in fashion, face a much bigger forecasting issue compared to physical retailers, which relates to product returns.

There has been extensive review and research on the field of forecasting [Petropoulos et al. \(2022\)](#), it has seen significant growth in theory and practice due to the increase in business and academic operations over time. Thus, the need for a forecast for planning, the technological advances in the field of big data, computing, and cloud storage which also led to exponential growth and innovation in the field of forecasting and machine learning. Computing particularly has enabled the analysis of larger and more complex data sets which was not possible before. Computer science and computing power and speed increase have led to applications of complex models on big data sets such as neural networks random forests, XGBOOST, LightGBM, and many others, these kinds of machine learning models have gained big attention in both theory and practice [Petropoulos and Makridakis \(2020\)](#) while computing as well enabled the processing and application of complex statistical models such as Bayesian regression, Linear regression and Lasso Regression and ARIMAX models. There is also growing research on the branch of judgmental forecasting which relies on insights proposed by subject matter experts that either complement quantitative methods or become stand-alone models used in practice, judgmental methods are particularly useful when there is not enough historical data or there are unrecognized patterns of future events expected that is not recognized by historical data among other reasons [Petropoulos et al. \(2018\)](#).

Moreover, there is a mixture of models that can be used also, known as a combination of models or ensemble models. These approaches have been known to provide more accurate results in forecasting especially since models tend to perform well on some aspects, so combining multiple models could naturally lead to better results. Besides, growing literature and research do not show for certain if statistical models are better than machine learning models or vice versa. Prior to implementing quantitative forecasting, pre-processing techniques could be applied to segregate patterns in historical data and provide a clear focus on the pattern of the time series. For example, to reduce outliers in data, practitioners apply BOX-COX transformation or variations of it ([Box et al. \(1994\)](#), [Atkinson et al. \(2021\)](#)). After modeling, these forecasts are back-transformed to the actual sales of observed data. In addition, there is the decomposition of time series data [Macauley \(1931\)](#)

to the trend component, seasonality component, and remainder component. In the field of forecasting, anomaly detection is also researched as anomalies can occur due to many reasons; as an example an unexpected event in historical data such as distastes, wars, and pandemics and those anomalies should be handled when modeling for forecasting, another type of anomalies is anomalies that appear due to missing data, wrong entries, extremes, and duplicates [Talagala et al. \(2020\)](#). In addition to anomalies, there could be outliers in data such as extreme observations that can appear in weather conditions or sales. Treating outliers is also recommended when forecasting as it biases the coefficients of the models and decreases the efficiency of predictions [Bianco et al. \(2001\)](#).

Exogenous variables are also variables that are included in forecasting. Naturally, if they impact the independent variable or are co-related to it, it leads to improvement in prediction. For instance, when we predict sales in a particular store, the weather conditions might be a factor in the increase or decrease of sales of this product if it is impacted by weather, for example, umbrellas, this is considered a continuous variable or a promotion that will be initiated by the store for umbrellas which in this case is considered a categorical variable, there are also time series components that can be added to the independent variables such as lags, rolling averages, rolling demand interval and sales of other associated products. Not only so, but the interaction of different variables and polynomials is also known to capture non-linear effects and is useful for linear models. However, preparing for independent variables can be extensive and requires extra effort. In addition, independent variables can lead to biased predictions if regressors are co-related to each other and have an inverse effect of what it is intended for. Such events can be handled by approaches like principle component analysis (PCA), which is used as a method for linear dimension reduction. Note that this correlation problem does not impact non-linear models that depend on ranking such as tree-based models [Petropoulos et al. \(2022\)](#), [Bishop and Nasrabadi \(2006\)](#)

As a result, when we develop forecasting models based on exogenous variables, practitioners tend to do variable selection methods that rank the independent variables based on their effect on the target forecast. Reducing the features, based on a ranking, would reduce the dimension of data in the model. Note that regression models tend to perform poorly when there is a lot of explanatory variables ([Bishop and Nasrabadi \(2006\)](#), [Petropoulos et al. \(2022\)](#)). Effective and popular models such as Least Absolute and Selection Operator (LASSO) are known to penalize having too many coefficients in a model, thus reducing the number of coefficients. This reduces the error resulting from the model ([Tibshirani \(1996\)](#)). There was another update to the LASSO procedure, particularly by [Rapach et al. \(2019\)](#). In addition, a notable extension for forecasting is the complete subset regression (CSR) model where all the possible combinations of features are tested and the best possible features are selected [Elliott et al. \(2015\)](#).

From the known statistical models used in retail and used in other fields as well as exponential smoothing, which is one of the most flexible models and always ranks high in benchmarks ([Petropoulos et al. \(2022\)](#), [Makridakis et al. \(2020\)](#)). The method's idea is simple yet effective. It relies on the weighted average of past observations, these weights decrease exponentially for older observations as we move back in time, based on the pattern of data, there could be variations of exponentially smoothing that includes trends such as Holt model that incorporates exponential smoothing with the trend and Holt-Winters models which incorporates exponential smoothing with trend and seasonality [Gardner Jr and McKenzie \(1985\)](#). Exponential smoothing can be denoted as :

$$f_{t+1} = \alpha y_t + (1 - \alpha) f_t \quad (1.3)$$

Simple exponential smoothing forecasting (SES) must use first an initial forecast and a parameter  $\alpha \in [0, 1]$ , deduced by minimizing the sum of squares of one step ahead

forecast errors. [Hyndman et al. \(2002\)](#) introduced a variation of traditional exponential smoothing that is widely used in practice as it allowed the point forecast of exponential smoothing to have a predication interval which allows for future probabilistic forecasting.

Another model that is widely used in retail is Auto-regressive integrated moving average (ARIMA) models [Box et al. \(1994\)](#) where it gains information from the previous lags of the time series based on the orders of auto-regression  $p$ , integration or differencing  $d$  and moving average error  $q$ . For example an ARIMA (1,0,1) model would be denoted as [Petropoulos et al. \(2022\)](#) :

$$y_t = c + \phi y_{t-1} + \epsilon_t + \theta \epsilon_{t-1} \quad (1.4)$$

where  $\epsilon_t$  is the zero mean uncorrelated process,  $\phi$  is the coefficient of the autoregression for previous observation and  $\theta$  is the coefficient of the moving average error, there are various ways to identify the  $(p, d, q)$  either by observing the auto-correlation lags or the partial auto-correlation lags and observe the significant autocorrelation lags, apply a grid of combination of orders and identify the one that provides the lowest Akaike information criterion (AIC). There is also the `auto.arima` function in Python and R statistical software (programming language), that applies an algorithm to quickly identify, not only the ARIMA's orders  $(p, d, q)$  but also the seasonal orders if any  $(P, D, Q)$  which are the seasonal lags. In fact, It is known to be quick and very effective ([Hyndman et al. \(2002\)](#), [Petropoulos et al. \(2022\)](#)). For retail data, where demand sporadically appears, with no demand and all in some periods [Croston \(1972\)](#), forecasting is very challenging. The demand characteristics for such timer series can be classified into two categories; parametric and non-parametric. The parametric supposes that the distribution of demand can be presented in the time series such as Poisson, and negative binomial while for non-parametric, demand distribution is unrecognized.

Simple exponential smoothing is known to forecast intermittent demand data but as Syntetos mentioned in [Petropoulos et al. \(2022\)](#), SES has a positive bias issue and no tolerance for sporadic data, it cannot capture the inter-demand intervals and the size of the demand. [Croston \(1972\)](#) developed the Croston method where the demand size and demand interval are forecasted separately using simple exponential smoothing and the ratio of demand size divided by the demand interval forms the mean demand. It is the only method that is available in major ERP-type solutions such as SAP and specialized forecasting software packages. The M5 competition is the first forecasting competition that has included the Croston method to deal with intermittent demand. However, it is worth noting that for a high degree of intermittence and lumpiness, the Croston method often leads to low forecasting accuracy ([Syntetos et al. 2015](#), [Makridakis et al. 2020](#)). Croston's method is included in Chapter 2 as a benchmark for the performance analysis of the proposed approach.

A modification of the Croston method was proposed by [Syntetos and Boylan \(2001\)](#) that deals with the inversion bias that the Croston method does not overcome. The forecast method is known as The Syntetos and Bolyan Approximation (SBA). For non-parametric forecasting approaches for intermittent demand, there are two common methods; the bootstrapping approach where a re-sampling of the historical demand is done to deduce empirically the distribution of the demand [Rubin \(1981\)](#), [Hasni et al. \(2019\)](#). Development of the method was proposed by [Hasni et al. \(2019\)](#), where they assumed that the first day of demand lead time is a non-zero demand. As for the method of aggregation with overlapping and non-overlapping blocks, the overlapping blocks divide the time series into time buckets where the oldest observation is dropped and a new one is added creating an overlap of buckets while for the non-overlapping block, the demand series is divided into non-overlapping blocks at a time [Porrás and Dekker \(2008\)](#).

Forecasting methods fall broadly under two categories; statistical forecasting and machine learning models (ML). While there is no clear distinction between them, it can be assumed that ML is hard to explain and a black box compared to statistical forecasting models. For example in machine learning models, generally, we cannot understand the coefficients resulting from the fitted model, while in statistical models, the coefficients are explainable (Makridakis et al. (2018), Barker (2020)). ML models are easier to adapt to different time series and they can deal well with noisy data, for example, linear regression models are prone to over-fitting especially if there a lot of independent variables included. In addition, machine learning models capture complex relationships that linear models fail to capture, one popular method to capture non-linear relationships is neural networks Barron (1994). However, they require sufficient data to be fitted opposing to exponential smoothing for example. ML models also are usually more computationally expensive compared to statistical models Makridakis et al. (2020) which might be challenging in a retail environment where there are a lot of products to be foretasted. This can give preference to aggregation approaches as top-down or middle-out. In addition, ML models require more intensive training than statistical models as their parameters need to be tuned using a grid search.

As forecasting in retail is directly linked with inventory, the areas of performance management and bullwhip effect have been studied in line with different forecasting approaches and methods Petropoulos et al. (2022). The bullwhip effect occurs due to the variability of demand in different stock points across the supply chain, for example, a retailer might have an overstock of products in the supply facilities and fulfillment centers based on the previous sudden spikes in demands at the retail stores causing accumulation of non utilized stocks across the supply chain as well as incorrect or slow processing of information across various echelons of the retailer Lee et al. (1997). There was further research on the subject indicating that information sharing is important to mitigate the Bullwhip effect unless the wholesaler or the upstream side directly forecasts the retailer orders or the demand of the retailer is available on a real-time basis for the wholesaler or the manufacturing facility Ali et al. (2017), Graves (1999). When the lead time is more than one period, a forecast of multiple periods ahead is required and the variance of forecast errors is needed to determine the level of safety stock for this particular product. Traditionally, the forecast error is multiplied by the lead time period with a desired service level cover. However, recent approaches recommend that the forecast error be smoothed over multiple periods instead of directly multiplying by lead time period Boylan et al. (2006) to measure inventory performance with regard to multiple forecasting comparisons to conclude which method is better, an exchange curve is recommended where it can be noted the effect of sequentially increasing service levels on the inventory holding volume and inventory shortage. The forecast with a lower curve on both axes is a more suitable one.

Another important body of literature in retail forecasting deals with demand aggregation. In fact, demand in a retail supply chain can be aggregated at several levels, such as the product level (e.g., stock-keeping unit, family, etc.), geographic location (e.g., store, region, etc.), or time (e.g., day, week, month, quarter, etc.). These different aggregation strategies can be either hierarchical and/or temporal and necessitate different forecasting methods (Rostami-Tabar et al. 2013). However, it should be noted that demand aggregation in an omnichannel context has never been analyzed in the literature. In Chapter 2, we bridge this gap in the literature by analyzing the forecasting and inventory management performance of demand aggregation at the product level across channels.

### 1.2.5 Forecasting aggregation in supply chain

Retail data is known to be numerous and diverse, depending on the type of retailer, it can generate up to millions of product's sales records per year. Researchers have long tackled the forecasting of products in the retail environment through aggregation. In fact, time series can be aggregated to many levels. First of all, geographical aggregation is one common aggregation scheme in retail especially with the abundance of many stores in different places across the globe, as each region can have its own taste and preference. Secondly, another kind of aggregation is cross-sectional aggregation. For example, the product level can be aggregated to the sub-family level and the latter level can be aggregated to the family level, all the way to the department/segment level. Finally, Another approach is temporal aggregation where we derive low-frequency series from high frequency ones (daily/weekly/monthly) [Babai et al. \(2022\)](#) . This aggregation is also applicable in retail as retailers need long-term planning for purchasing products, planning for manpower, negotiating contracts, and hedging product purchases if needed. Temporal aggregation is also vital for the retailer to model business growth and budgeting. In terms of tactical and operational planning, a retailer uses daily forecasting for replenishment while distribution centers may use weekly and monthly time series for their replenishment ([Oliveira and Ramos \(2019\)](#), [Mircetic et al. \(2022\)](#), [Babai et al. \(2022\)](#)). Hierarchical forecasts are said to be coherent when the bottom-level series sum up to the forecast of the higher levels. If every level is forecasted independently and directly, this means that the hierarchy is not coherent ([Hyndman and Athanasopoulos \(2018\)](#), [Babai et al. \(2022\)](#)). The most common approaches in hierarchical forecasting are bottom-up, middle-out, and top-down. In the bottom-up approach, the bottom series are summed up to the higher level. This procedure is iterated until the highest level is reached. The bottom level is recognized as the base forecast. Since the process starts from the bottom level, no information is lost during the aggregation. However, if the base forecast error is significant, the error is amplified at the upper levels. In the the top-down approach, the base forecast is in the upper level, this level is, then, disaggregated to lower levels using approaches such as average historical proportions or forecasting proportions.

Finally, the middle-out approach starts in the middle level of the hierarchy and is forecasted directly which is then recognized as the base forecast ([Hyndman and Athanasopoulos \(2018\)](#), [Gross and Sohl \(1990\)](#)). It is then summed up to higher levels and dis-aggregated to lower levels based on a procedure such as average historical proportions of the lowest level compared to the mid-level or forecast proportions which is the proportion of the h-step ahead lowest level forecast to the h-step ahead mid-level forecast. [Hyndman et al. \(2011\)](#) stated that top-down forecasting using average proportions may lead to biased forecasts at the bottom level. While bottom-up forecasting could be biased to the high variability of the low-level data. Accordingly, popular methods were introduced to capture variability from all the hierarchical levels. Reconciled forecasts are a common method to obtain relations from the entire hierarchy, where each level forecast is obtained separately (bottom/middle/top) and they are adjusted to obtain coherent forecasts [Hyndman and Koehler \(2006\)](#), [Wickramasuriya et al. \(2019\)](#). A layout of the aggregation procedures as provided by [Babai et al. \(2022\)](#) is provided in Figure 1.5

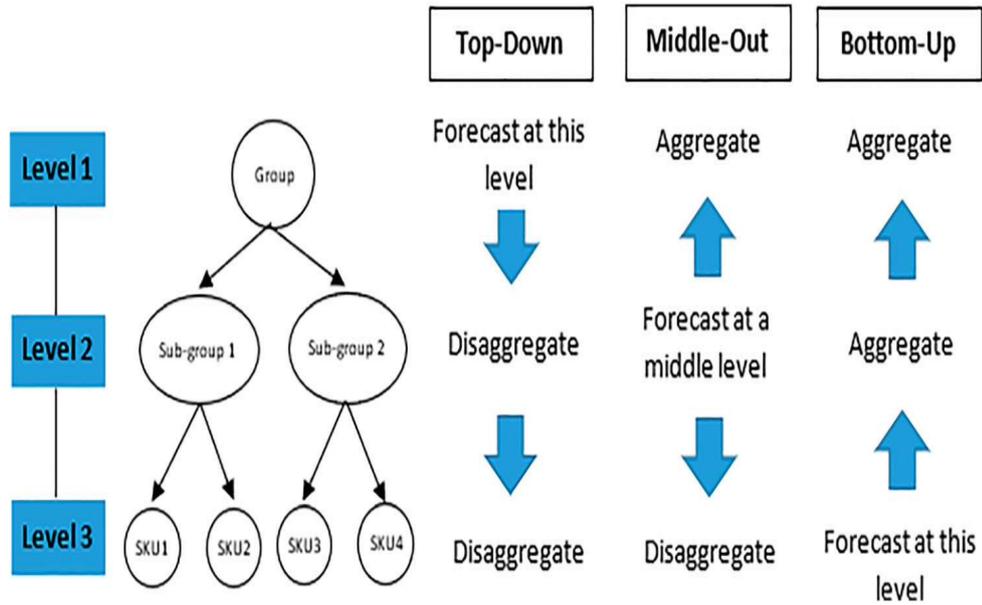


Figure 1.5: Layout of forecast aggregation approaches from Babai et al. (2022)

Another simpler approach is the combination (COMB) approach where several aggregation approaches are averaged at a level of choice in the hierarchy and then summed up or disaggregated accordingly Mircetic et al. (2022). For example, if the COMB approach starts at the bottom level, the forecast of TD and BU at the low level can be averaged and then summed up through the entire hierarchy. More details are provided in Chapter 4.

In General, hierarchical approaches can be efficient for retailers if they provide good accuracy in forecasting and less computing time. Note that in the M5 competition, a lot of ranking competitors used hierarchical approaches as the data itself was a geographical hierarchy.

### 1.3 Conclusion

We have outlined the research work on omnichannel fulfillment, forecasting in retail, anticipatory shipping, and hierarchical aggregation. There is a clear gap in the use of basket data or lack thereof in omnichannel fulfillment even-though a basket ordered by a customer is often, more than not, a multi-item or omni-item basket. We will explore the use of this information in all of the related disciplines presented in this chapter. Starting with omnichannel forecasting and inventory fulfillment, followed by omnichannel fulfillment, cost minimization, and anticipatory shipping. Finally, its use of basket data in forecasting aggregation.

# Chapter 2

## Basket Data driven approach for omnichannel demand forecasting

### 2.1 Introduction

#### 2.1.1 General introduction and motivation

Omnichannel retailing uses a variety of channels to interact with customers and fulfill their orders in a seamless shopping experience, enabling customers to order anytime from anywhere, in person or through digital devices, and to have their purchase be delivered at their preferred time and location (Strang 2013, Bell et al. 2014, Chopra 2018). From an operations management perspective, omnichannel retailing implies the integration of different demand streams, which raises several key issues, related to pricing optimization, assortment planning, demand forecasting, inventory management, and network design. These issues are far from being resolved today.

Byrne (2016) highlighted that omnichannel retailing adds complexity and that part of the change retailers would need to make involves their forecasting and planning processes, which work well for shelf sales but not necessarily for online order fulfillment. Rooderkerk and Kök (2019) argued that inventory replenishment traditionally relies on a separate forecasting demand of channels, whereas the demand cannot be truly captured because of the omnichannel customer's journey between showrooming and webrooming (discussed in Bell et al. (2018)). Moreover, Byrne (2016) argued that among the most common trends in omnichannel distribution is the recognition of the need for a shared view of inventory across all channels.

Furthermore, the shopping basket notion has gained an increasing importance in omnichannel retailing. It has become a crucial vehicle to capture sales through continuous promotional periods and various bundling offers. For instance, the data panel of a major cosmetics retailer used in this work reveals that, out of 2.2 million online orders over the year, 50% of them occurred in baskets with two distinct products or more, and there were 2.4 ordered products on average per basket. This raises the question of how data from multiple product baskets could be considered by operations managers to improve the accuracy of demand forecasting in omnichannel retailing. Boone et al. (2019) highlighted the existence of influences between products purchased within customers' baskets and they stated the use of basket data in forecasting as one of the research avenues that are yet to be explored. It is worth noticing that little research has been devoted to demand forecasting within an omnichannel context, and to the best of our knowledge, no research explicitly considering basket data in a forecasting method has been previously proposed in the literature. We endeavor to fill this gap in this chapter.

This research work builds on some findings and evidence from both marketing and operations. In fact, the marketing literature highlights the existence of an interdependence of products’ sales within single and multiple category purchases in market baskets (Russell et al. 1999, Manchanda et al. 1999, Russell and Petersen 2000). Products’ complementarity and substitution have emerged in this literature as important aspects of this products’ interdependence. In fact, complementarity could be for example the result of a marketing activity (price or promotion) of a product that influences the consumers’ purchase of another product. The substitution can be viewed as a negative complementarity, i.e. purchases between a pair of products are negatively related. Furthermore, graph theory is another modeling approach that is used to formalize the connections between items within a network and analyze the magnitude of their connectivity. Two variables are often used in the literature to characterize these connections, namely: the degree of connectivity and the strength. These two variables can also be used to reflect the complementarity level of products since they measure the magnitude of connections between products that are often sold together within the same category or across categories. Following the recommendation of the M5 competition on the benefit of considering exogenous/explanatory variables when forecasting the demand in the retail context (Makridakis et al. 2020), we consider the four cited above variables for the sales forecasting using baskets data in the omnichannel context.

### 2.1.2 Business context

Our research work is part of a project in collaboration with a large retailer in the cosmetics industry that sells a broad catalog of products through stores and online channels. The project aims at the integration of in-store and online channels from demand forecasting and fulfillment perspectives. Figure 2.1 illustrates the retailer’s downstream supply chain schema composed of a set of customers and their locations, a selected store, a retail warehouse, and a fulfillment center.

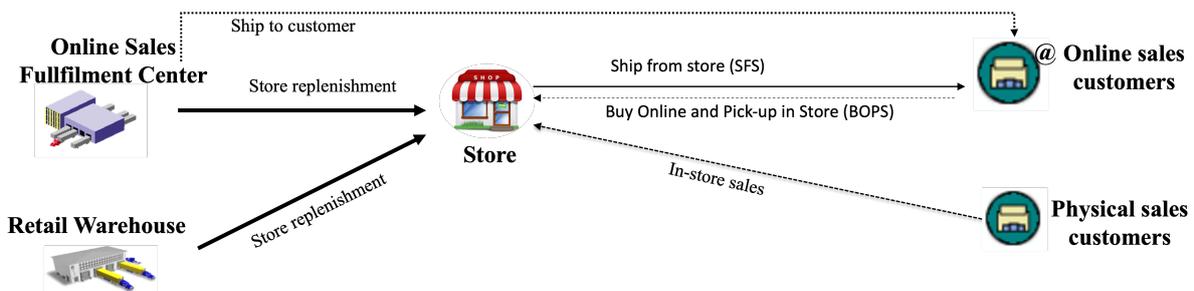


Figure 2.1: Omnichannel fulfillment

The retailer has been operating its in-store sales for a long time, relying on the deployment of a brick-and-mortar store network, replenished periodically from a central warehouse (denoted BMS). With the advent of e-commerce, the retailer started offering online sales based on a ship-to-customer delivery process from a dedicated fulfillment center. Recently, the retailer considered the implementation of “buy online pick-up at store” (BOPS) and “ship-from-store” (SFS) policies through the online channel to ensure a competitive response time that is also supported via the fulfillment center. SFS takes advantage of the existing physical network by turning certain store locations to ship-from points for online sales. Figure 1 depicts the replenishment-storage-shipment flows at the store under the omnichannel context, with dual product flow replenishment to the store (bold arrows), ship-to customer flows for SFS (a regular arrow), and customer moves to

the store for BMS and for BOPS options (dotted arrows). Hence, a key question tackled in this work is, how can omnichannel demand forecasting be used to improve inventory management and fulfillment performance in retailing?

### 2.1.3 Contributions and organization of the chapter

Our chapter contributes to the literature in four ways:

1. Building on findings from marketing and graph theory, we propose a novel forecasting approach for online and store sales that is driven by data on customers' shopping baskets. It builds on a multi-category graph-learning approach from the basket data using an original set of cross- and within-categories regressors. Using this approach, four forecasting methods are proposed based on linear and polynomial regressions and machine learning. To the best of our knowledge, our chapter is among the first contributions to the literature that consider the shopping basket in demand forecasting.
2. Through the analysis of more than 2 million orders in the online channel of a major cosmetics retailer, we characterize the demand patterns, and we provide empirical evidence of the dominance of cross-category complementarity and the high level of within-category substitutability of products in market baskets.
3. Based on a dataset of an assortment of 24,029 products in both the online and store channels, we empirically show the outperformance of four proposed forecasting methods compared to two benchmark forecasting methods, namely: Croston's method and the autoregressive integrated moving average (ARIMA) method, which are often used in the retail context as shown in (Syntetos et al. 2010, Babai et al. 2013, Fildes et al. 2019). The outperformance is shown to be substantial for the class of lumpy demands due to the high connectivity of products. We also empirically show the benefit of joint forecasting of online and store sales. This holds for all the considered forecasting methods.
4. Through an investigation based on the order-up-to-level inventory control policy, we provide empirical evidence that using joint forecasting and shared inventory in an omnichannel context leads to a reduction of inventory shortages and an increase in the achieved service levels. This is shown by comparing the inventory performance of three fulfillment scenarios where the sales of both channels are forecasted separately or jointly and the inventory is either dedicated to each channel or shared by both channels.

The remainder of the chapter is organized as follows. Section 2 is dedicated to a review of market baskets. In Section 3, we present our basket data-driven forecasting approach and the considered alternative forecasting methods. We empirically assess the performance of the alternative forecasting methods in Section 4, along with an analysis of the impact of omnichannel forecasting on inventory and fulfillment performance. Finally, in Section 5, we present the conclusions of the chapter in addition to suggesting some avenues for further research.

## 2.2 Market Baskets

Market baskets have been largely considered in the marketing and operations literature to analyze consumer purchase behavior in the retail context and to make marketing and planning decisions (Manchanda et al. 1999, Cachon and Kök 2007). One of the main observations that arises from this literature when analyzing basket data is the existence of single-category and multiple-category purchases in market baskets (Russell et al. 1999). Moreover, this literature points out the existence of interdependence between products. Such interdependence occurs within a single category and across categories. This is obvious since consumer preference for a product in one category is contingent on the sales,

price or promotions of products in the same and in other categories (Russell and Petersen 2000).

Among the important features of the interdependence of products within market baskets, complementarity and substitutability (also called negative complementarity) have emerged in the literature. Manchanda et al. (1999) empirically showed, when analyzing the case of more than 17000 sales for households in the grocery retail context in the US, that for example the detergent and softener are complementary as well as cake mix and frosting. Harlam and Lodish (1995) developed a consumer choice model in the retail grocery context where product complementarity in the same category is considered to link choices of purchased products. Manchanda et al. (1999) and Ma et al. (2012) among others incorporated aspects of complementarity and substitution in modeling multicategory consumer purchase decisions in market baskets. Hence, in this chapter, we build on this literature to consider both complementarity and substitution as explanatory variables to forecast the sales of a product within a particular category of products.

Furthermore, there are other measures that emerge from graph theory to characterize the complementarity of products sold together within a shopping basket across categories. Two of these measures that are often used in the literature are: Strength and Degree (Berry and Linoff 2004). These two attributes are also included in our work as explanatory variables to forecast the sales of a product across different product categories. Note that multivariate forecasting methods could have been used to take into account the interdependencies between products when forecasting their sales. However, as discussed in (Koop 2013, Wang et al. 2019), such methods have dimensionality issues when dealing with a large number of SKUs, which is the case in the retailing context.

To conclude this section, it is worth pointing out that despite the richness of the forecasting literature in the retail context and the integration of different exogenous variables, attributes of shopping baskets, such as degree, strength, complementarity, and substitutability, have never been considered in terms of demand forecasting. This gap in the literature motivates our research to consider these variables when forecasting sales in the omnichannel retailing context.

## 2.3 Basket data-driven forecasting approach

Starting from historical sales, we rely explicitly on graph theory in order to model a graph-like structure that captures cross-categories and within-categories characteristics. Such a multi-category graph modeling approach captures the relationship between sold products on various dimensions such as the baskets, the days, and/or the categories. Next, based on the findings of the literature described above, the four identified attributes (i.e. degree, strength, substitutability, complementarity) will be used as exogenous variables in a forecasting method. The proposed forecasting approach can be summarized into two steps and is illustrated in Figure 2.2. The two steps of the proposed approach are detailed in the following subsections.

- **Step 1.** a) Shopping baskets data (from the sales history) are used to build the network, b) the network is then considered to calculate the four attributes: *Degree*, *Strength*, *Substitutability*, and *Complementarity*.
- **Step 2.** The four attributes are used as regressors along with historical sales data within a forecasting method. Such a method should be selected from the set of forecasting methodologies that allow for the incorporation of exogenous variables, which are mainly regression-type models, ARIMAX models, and machine learning models (Poyil 2019, Sheppard 2019). Hence, four forecasting methods are proposed

in this chapter integrating the basket-data-based regressors, namely: (1) an ARIMAX method, (2) a linear regression method with one lag of sales, hereafter denoted by LinReg-LS, (3) a polynomial regression method with one lag of sales, hereafter denoted by PolyReg-LS, and (4) a machine learning method based on lightGBM.

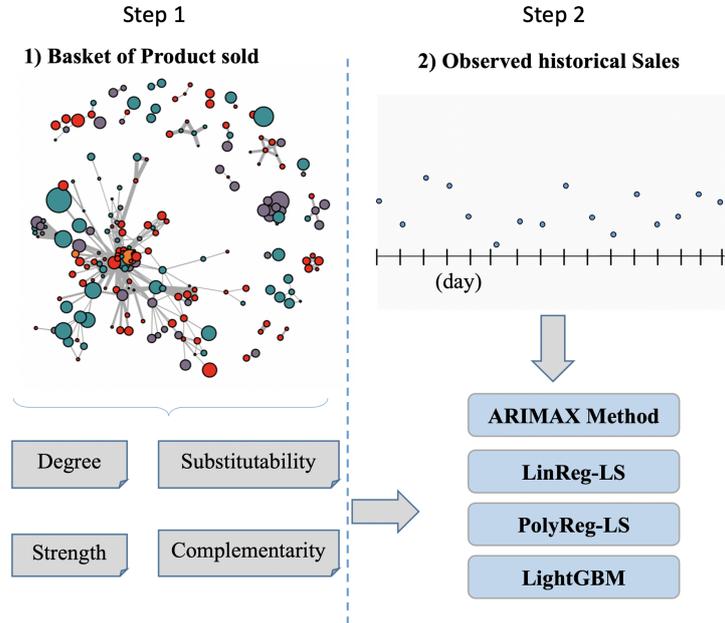


Figure 2.2: Illustration of the forecasting approach

### 2.3.1 Network characterization and attributes extraction

Graph theory is widely applied to conceptualize and analyze complex networks in supply chains (Gross and Yellen 2005, Kim et al. 2015, Dooley et al. 2019). In our context, we rely explicitly on graph theory in order to: 1) model a graph-like structure from historical sales that captures cross-categories and within-categories characteristics, 2) explore a graph-learning from the basket data using an original set of regressors. More specifically, we consider a weighted graph to model the relationship between sold products pertaining to various categories (i.e., a multi-category graph). The nodes correspond to the products in the considered assortment, and when there is a pair of products sold together within at least one basket (we refer to this as a connection between the two products), the edge linking the two products is associated with the number of baskets in which the two products are sold together. This number is referred to as the frequency of this connection. We note that in our graph representation, we assume that the relational attributes of products are the only associations and there is no causal relationship; thus, the network is undirected. Figure 2.3 illustrates a typical basket data representation in a multi-category graph. The data is extracted as follows, each basket’s data is transformed into every two possible pairs, this process is done on all baskets for a period of seven consecutive days (the graph horizon), attributes are then calculated on those pairs and they would be effectively the attributes values for the next day of forecast. This process was implemented on two panels of data: the first panel is one year of online sales data of many cities and the second panel is one year of physical retail data of one city. Figure 2.3 represents a graph subset of seven days period graph from panel A. One can note how pairs of products are from different substitutability sets/categories.

With this in mind, we identify the four attributes of the product considered in the proposed forecasting approach: (1) *Degree* of the product, the number of connections (or

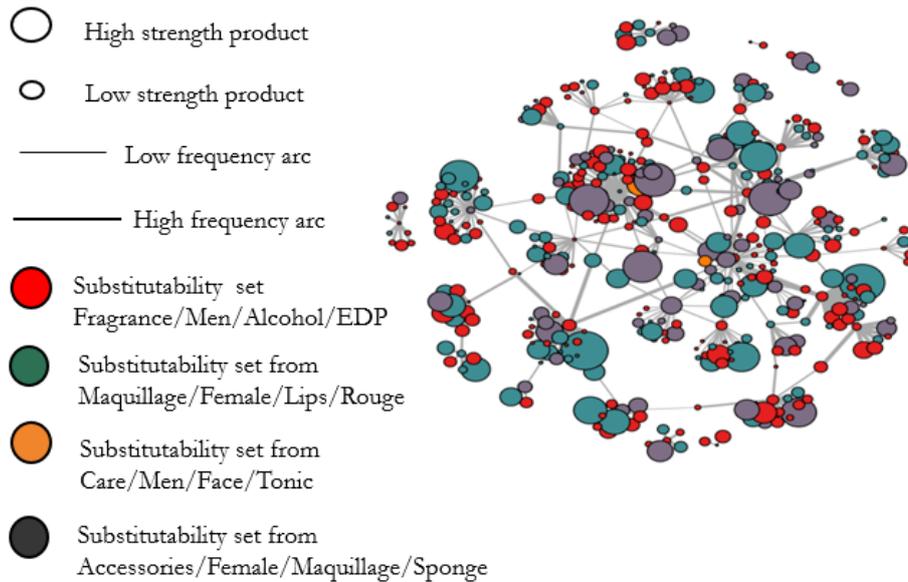


Figure 2.3: Basket data representation in a multi-category graph

arrows) with the other products of the multi-category graph; (2) *Strength* of the product, the total frequency with all connected products; (3) *Substitutability* of the product, the chance to be sold as a substitute among all the products within its substitutability set; and (4) *Complementarity* of the product, the proportion between the number of connections within the substitutability set and the number of connections with the other products of the assortment (i.e., *Degree*). We notice that the first and second attributes capture cross-category features, whereas the third and fourth attributes are focused on within-category characteristics.

Note that the substitutability set is a subjective managerial decision based on assortment breadth and depth, and also on the sector in which the omnichannel retailer operates. In this research, the substitutability set consists of the lowest level category, which is the level above the product level (also defined as category breadth for some retailers). An example from our empirical study is the products from the fragrance department, which may be categorized based on gender (i.e., Men, Women), the concentration of alcohol (i.e. "Eau de toilette" (EDT), "Eau de parfum" (EDP) which is known to have much higher concentration than EDT, and "Eau de cologne" (EDC) which has less concentration than "EDT"). Accordingly, the lowest level category that could constitute a substitutability set may be (Fragrance/Men/Alcohol/EDT) or (Fragrance/Men/Alcohol/EDP), or (Men/Fragrance/Alcohol/EDC). To select the lowest level category in this chapter, we worked closely with supply chain leaders from the company and leveraged company data to make an empirical test in the empirical investigation section. Other retailers may follow the same empirical approach to determine the appropriate category level to define their substitutability sets. They could also end up with a different categorization (a higher level category) and substitutability set such as (Fragrance/Men/Alcohol), which fits, straightforwardly, in our forecasting approach.

In order to illustrate the graph and the characterization of the attributes, we consider an example of six products (A, B, C, D, E, F). In this example, two substitutability sets are distinguished: (A, C, E, F) in grey and (B, D) in blue. Here, for instance, product A is sold nine times. It is sold in two baskets with product F, in two baskets with product C, and in five baskets with product D. The graph in Figure 2.4 presents the network

structure of this example, and the table in Figure 2.4 shows the calculation of the four attributes for the six products. For example, the degree of product A is 3 because this product is sold with three other products, whereas the strength is 9, which is the sum of the frequency of the arrows. The substitutability of product A is  $1/4$  because the substitutability set of this product is composed of four products (A, C, E, F). Finally, the within-category complementarity is  $2/3$  as it is connected to two other products (C, F) in its substitutability set among all the connected products (C, D, F).

Product	Degree	Strength	Substitutability	Complementarity
A	3	9	$1/4$	$2/3$
B	1	4	$1/2$	$0/1$
C	2	4	$1/4$	$1/2$
D	3	8	$1/2$	$0/3$
E	2	5	$1/4$	$0/2$
F	1	2	$1/4$	$1/1$

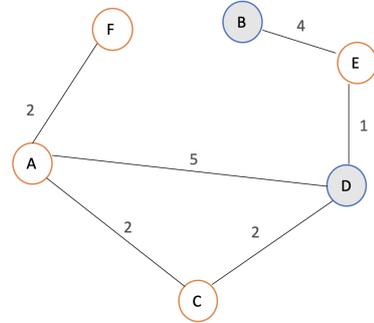


Figure 2.4: Illustrative example of a multi-category graph and the associated calculation of the attributes

From a graph-theoretic perspective, we can conceptualize the network of connections between products in baskets as follows. Let  $t \in T$  be the set of historical sales periods (days) and  $b \in B$  be the set of sold baskets, where  $B^t$  is the subset of baskets sold in period  $t \in T$ . Let  $p \in P$  be the set of products, where  $P_b$  is the subset of products included in sold basket  $b$ . Let  $S_p \subset P$  be the subset of substitutable products for a given product  $p \in P$ , that could be built based on the product categories. We define  $G$  as an undirected graph, denoted by  $G = (N, E)$ , composed of  $N$  nodes, representing the number of products ( $N = P$ ) and  $E$  edges representing pairs of products, where  $E : (p, p'), \forall p, p' \in P$ . Accordingly, the graph  $G^t$  is defined for a given period  $t$ , and is composed by a set of subgraphs  $G_b^t$  per basket  $b \in B^t$ .

The network is produced on a daily basis with a rolling time, and the attributes are determined for each product and for each forecasting period  $t$  (using basket data on periods  $t-1, t-2, t-3, \dots$ ). Let  $h(t)$  be the historical horizon used to build the network (i.e. 7, 14, 21, or 28 days prior to period  $t$ ) related to the estimation at period  $t$  (i.e. period where the forecast is made), composed by  $[t-1, \dots, t-h]$ , for the collection of baskets  $B^{h(t)}$ . Consequently,  $G^{h(t)}$  defines the network graph associated with the set of periods in  $h(t)$  such that  $G^{h(t)} = \cup_{[t-1, \dots, t-h]} G^t$ . For a given period  $t$ , based on the historical horizon  $h(t)$  data, the four attributes are computed as follows.

We introduce  $\lambda_{b(p,p')}^t$  as a counter that takes 1 if edge  $(p, p') \in G_b^t$  (i.e., when both products are in the same basket), and 0 otherwise. It is then used to compute  $\lambda_{(p,p')}^t = \sum_{t \in h(t)} \sum_{b \in B^{h(t)}} \lambda_{b(p,p')}^t, \forall p, p' \in P$ , which assesses the number of occurrences of a given edge  $(p, p')$  (i.e. a pair of products), in the set of baskets, over the historical horizon  $h(t)$ . The *Strength* attribute in period  $t$  is based on the number of incident edges for product  $p$  in basket  $b$ , which is computed as  $Strength_{pt} = \sum_{p' \in P} \lambda_{(p,p')}^t$ . The *Degree* attribute in period  $t$  of product  $p$ , is expressed as  $Degree_{pt} = \sum_{p' \in P} \min(1, \lambda_{(p,p')}^t)$ . Let  $S_p^{h(t)}$  be the subset of substitutable products for a given product  $p \in P$  that appears in the set of baskets, over the historic horizon  $h(t)$ . The *Substitutability* attribute in period  $t$  of product  $p$ , is expressed by  $Substitutability_{pt} = 1/|S_p^{h(t)}|$ , where  $|S_p^{h(t)}|$  is the cardinality of  $S_p^{h(t)}$ . The *Complementarity* attribute in period  $t$  of product  $p$  is based on the number

of incident arcs within the substitutability subset  $S_p$  and the  $Degree_{pt}$ , that is expressed by  $Complementarity_{pt} = \sum_{p' \in S_p} \min(1, \lambda_{(p,p')}^t) / Degree_{pt}$ . Note that the product quantity information is captured by the historical sales of the product and it is not represented in the graph.

### 2.3.2 Forecasting methods

We present in this subsection the four forecasting methods considered in this chapter: the ARIMAX method, the linear regression method with one lag of sales, the polynomial regression method with one lag of sales, and the lightGBM machine learning method. The reason we choose these methods is that first; they allow for using regressors; second; compare linear and non-linear methods and third; evaluate the impact of adding regressors on ARIMA methods and finally; lightGBM is known to be a top-performing machine learning model in forecasting competitions [Makridakis et al. \(2020\)](#).

At any time period  $t$  and for each product  $p$ , the forecast, denoted by  $y_{pt}$ , is calculated for a network historic horizon  $h(t)$  using the *degree*, *strength*, *substitutability*, and *complementarity* regressors, denoted by  $Degree_{pt}$ ,  $Strength_{pt}$ ,  $Substitutability_{pt}$ , and  $Complementarity_{pt}$  respectively (regressors calculated using basket data on periods  $t - 1, t - 2, \dots, t - h$ ), in each proposed forecasting method as follows:

- In the ARIMAX method:

$$y_{pt} = aDegree_{pt} + bStrength_{pt} + cSubstitutability_{pt} + dComplementarity_{pt} + n_{pt} \quad (2.1)$$

where  $a$ ,  $b$ ,  $c$ , and  $d$  are the regressors respective coefficients. We assume that  $n_{pt}$  is given by an ARIMA( $r, k, q$ ) model, where  $r$  is the auto-regressive order,  $k$  is the differencing order, and  $q$  is the moving average order.

- In the linear regression method, LinReg-LS:

$$y_{pt} = aDegree_{pt} + bStrength_{pt} + cSubstitutability_{pt} + dComplementarity_{pt} + eD_{pt-1} \quad (2.2)$$

where  $D_{pt-1}$  is the lag of sales of product  $p$  at time period  $t - 1$  and  $a$ ,  $b$ ,  $c$ ,  $d$  and  $e$  are the regressors respective coefficients

- In the polynomial regression method, PolyReg-LS:

$$y_{pt} = \sum_{n=1}^{n_1} a_n Degree_{pt}^n + \sum_{n=1}^{n_2} b_n Strength_{pt}^n + \sum_{n=1}^{n_3} c_n Substitutability_{pt}^n + \sum_{n=1}^{n_4} d_n Complementarity_{pt}^n + eD_{pt-1} \quad (2.3)$$

where  $D_{pt-1}$  is the lag of sales of product  $p$  at time period  $t - 1$ ;  $a_n$ ,  $b_n$ ,  $c_n$  and  $d_n$  are the polynomial regressors' coefficients; and  $n_1$ ,  $n_2$ ,  $n_3$  and  $n_4$  are the polynomial regressors' powers.

- The LightGBM machine learning method:

We recall that LightGBM is a gradient-boosting model based on the decision tree algorithm. We use degree, strength, substitutability, complementarity, and lag sales of product  $p$  at time period  $t - 1$  as inputs of this method to forecast the sales at time period  $t$ . For the model selection and tuning, we make a grid search that consists of a number of leaves between 1 and 100, maximum depth between 1 and 30, and minimum data points in leaf between 20 and 200. We split the training set into training and validation sets.

The model with the lowest absolute error against the grid search is then selected to forecast the test set. More details on the implementation of the LightGBM method using exogenous variables can be found in [Abolghasemi and Esmaeilbeigi \(2021\)](#).

The forecast accuracy of the four proposed forecasting methods is empirically analyzed further in the chapter using the sales and basket shopping data of the considered cosmetics retailer. The performance of the proposed methods is compared to that of two benchmark methods: Croston’s method and the ARIMA method, which are commonly used in the intermittent demand forecasting literature ([Gamberini et al. 2010](#), [Shenstone and Hyndman 2005](#), [Murray et al. 2018](#)).

### 2.3.3 Inventory and fulfillment performance

An important question is whether the proposed forecasting approach can be used to improve the inventory and fulfillment performance in the omnichannel network, in addition to the forecast accuracy. To answer this question, we first analyze the inventory and fulfillment performance of the different forecasting methods. Then, under a proposed forecasting method, we analyze the fulfillment performance at the store using a separate versus a joint forecasting method and a dedicated versus a shared inventory. Recall that, as illustrated in Figure 1, the business context considers that the store has a dual source replenishment (a central warehouse and a fulfillment center) and that inventories kept at the store are for online and store sales. Hence, the three scenarios are as follows:

- **Scenario 1:** Dedicated inventory and separate forecasting scenario (referred to as the DISF scenario). In this baseline scenario, the store inventory is split by sales channel and the forecast is separated by channel.
- **Scenario 2:** Shared inventory and separate forecast scenario (referred to as the SISF scenario). In this scenario, the two sales channels share the same inventory at the store. However, the replenishment system depends separately on the forecast of demand of the store channel and the online channel.
- **Scenario 3:** Shared inventory and joint forecasting scenario (referred to as the SIJF scenario). In this scenario, the store and online channels share the same inventory at the store. However, the forecast is made jointly using the aggregated data from both channels.

The inventory performance of a forecasting method or a fulfillment scenario is assessed by measuring the resulting stock on hand, the inventory backorder and the achieved service level. To do so, we consider a periodic order-up-to-level inventory control policy, where the order-up-to level is calculated to satisfy a target cycle service level (CSL, the fraction of replenishment periods in which all of the demand can be met from stock). This policy is among the state-of-the-art inventory policies that are considered in the omnichannel context ([Hu et al. 2021](#), [Saha and Bhattacharya 2021](#)). Recall that under this policy, each day, the inventory position is reviewed and an order is triggered if it is found to be below the order-up-to level to raise it up to the order-up-to level. The order arrives after a lead time, and any demand that is not satisfied from stock on hand is backordered.

To calculate the order-up-to-level at any period  $t$ , as commonly recommended in the literature ([Syntetos et al. 2015](#), [Van der Auweraer and Boute 2019](#)), the mean and variance of the lead-time demand distribution, denoted by  $\mu_{LT}$  and  $Var_{LT}$  respectively, are estimated as:  $\mu_{LT} = (L + 1) * y_{pt}$  and  $Var_{LT} = (L + 1) * MSE_{pt}$  where  $L$  is the lead-time,  $y_{pt}$  is the forecast of sales at period  $t$  and  $MSE_{pt}$  is the mean squared forecast error at period  $t$  ( $L+1$  periods are considered instead of  $L$  to account for the protection interval under a periodic review). Note also that  $MSE_{pt}$  is given by:  $MSE_{pt} = \alpha[y_{pt} - D_{pt}]^2 + (1 - \alpha)MSE_{p(t-1)}$ ,

where  $\alpha$  is a smoothing constant between 0 and 1. The order-up-to-level at any period  $t$ , denoted by  $S_{pt}$ , is then calculated as:  $S_{pt} = F_{LTD}^{-1}(CSL)$  where  $F_{LTD}(\cdot)$  is the cumulative distribution function of the lead-time demand, having the mean and variance equal to  $\mu_{LT}$  and  $Var_{LT}$ . More details on the calculation of the order-up-to-level and the implementation of the simulation model of the inventory policy are given in (Hasni et al. 2019, Babai et al. 2019).

## 2.4 Empirical investigation

### 2.4.1 Supply chain and data description

We empirically analyze the performance of the proposed forecasting approach by considering real data of a major French cosmetics retailer. We use an assortment of 24,029 products that are divided into six product families: perfume, care, makeup, bath, hair, and accessories. The retailer is globally deployed in more than 25 countries and is specifically well established in France with hundreds of stores in many cities and an online sales platform. As illustrated in Figure 3.1, the retailer operated a retail warehouse and an online fulfillment platform, both located in France. The retailer data used for the purpose of this research relates to sales in 2018. The data contain the product reference, the description of the product, the sold quantity, the selling date, the delivery date, and the invoice (order) number. Based on this information, two datasets are built. The first dataset, referred here to as Panel A, contains the online sales in all of France for the year 2018 with about 2.2 million orders (single- and multiple-item baskets). Panel A is used to make the network analysis and to show the accuracy of the proposed forecasting approach. The second dataset, referred to here as Panel B, represents the orders of the online and store channels in one of the largest region in France in 2018. Panel B is used to assess the omnichannel forecasting accuracy and inventory fulfillment performance.

### 2.4.2 Network and basket analysis

We start by presenting the empirical results of the network analysis and the exploratory analysis of the sold products and baskets.

#### 2.4.2.1 Characteristics of the network

We empirically analyze the obtained networks through some characteristics, namely, the density, the assortativity, and the average path length. The density of the network is the ratio between the number of connections in the network and all the possible connections. This characteristic enables us to show if the products in the network have a high connectivity (i.e., degree). The empirical results show that the density of the network is equal to 0.3%, which proves that the network is characterized by a low density. However, from our analysis, we find that a few products have a very high degree of connectivity, whereas most of the products have a low degree. This shows that the network is heavily skewed in terms of degree distribution to some products with high connectivity. The assortativity of the network enables testing whether the baskets are composed of products from the same product category. A network is highly assortative if its correlation coefficient is close to 1, non-assortative if it is 0, and disassortative if it has a coefficient less than 0 (Newman 2002, Noldus and Van Mieghem 2015). The empirical results show that the overall network assortativity is 0.23, which means that 77 % of the retail baskets are cross-category orders. Finally, we investigate if the omnichannel retail network follows a small-world network (Newman 2003). To do so, we calculate the average path length of

the network. Note that small-world networks are often characterized by a short average path length. The empirical results show a high average path length of 2.85 compared with the case of randomized networks in which the average path length is 2.8 (for the latter, we used 50 randomly generated networks with the same number of nodes and density). This empirical result shows that this retail network is not considered a small-world network. Furthermore, in the following section, we present the investigation of network density and assortativity. Note that density and network assortativity are not attributes included in the methodology but are important characteristics to investigate on the retail network,

### 2.4.3 Substitutability set tests

As the substitutability set definition may be context-dependent, we provide a simple approach to determine the category level that can be used to define the substitutability set using the basket data. To illustrate the approach, we consider categories of products within the case study organization. In fact, we focus on products that belong to the fragrance category for men sold in one big store. We test different categories at different levels to identify the lowest-level category. Let  $S_1$  be a higher level category (Fragrance/Men/Alcohol) where the alcohol concentration can be "Eau de toilette" (EDT), "Eau de parfum" (EDP) or "Eau de cologne" (EDC). Let  $S_2, S_3, S_4$  be lower level categories which are (Fragrance/Men/Alcohol/EDT), (Fragrance/Men/Alcohol/EDC), and (Fragrance/Men/Alcohol/EDP), respectively. Our approach assumes that a customer is less likely to buy two products in a basket from the same substitutability set as they are competing with each other. Accordingly, the category with the lowest percentage of orders where two products within the same basket are bought would be more appropriate to be the substitutability set.

In Table 2.1, we report for the four substitutability sets (SS),  $S_1, S_2, S_3, S_4$ , the number of times two items from the same SS appeared in a basket, the number of times only one item from the same SS appeared in a basket, the percentage of times two items from the same SS appeared in a basket and the number of substitutes in the SS. Note that the results relate to four randomly sampled days from the historical sales.

Substitutability set (SS)	Sampled day from historical sales	Number of times two items from the same SS appeared in a basket	Number of times only one item from the same SS appeared in a basket	Percentage to which two items from the same SS appeared in a basket	Number of substitutes
$S_1$	Sample day -Jan	28	201	12.23%	134
$S_2$	Sample day -Jan	20	169	10.58%	104
$S_3$	Sample day -Jan	00	06	0%	5
$S_4$	Sample day -Jan	00	42	0%	025
$S_1$	Sample day -Apr	26	222	10.48%	136
$S_2$	Sample day -Apr	18	167	9.73%	101
$S_3$	Sample day -Apr	00	11	0%	6
$S_4$	Sample day -Apr	02	55	3.51%	28
$S_1$	Sample day -Sep	36	238	13.14%	151
$S_2$	Sample day -Sep	22	185	10.63%	112
$S_3$	Sample day -Sep	00	16	0%	9
$S_4$	Sample day -Sep	04	55	6.78%	28
$S_1$	Sample day -Nov	139	727	16.05%	247
$S_2$	Sample day -Nov	74	551	11.84%	174
$S_3$	Sample day -Nov	02	18	10.00%	9
$S_4$	Sample day -Nov	31	241	11.40%	63

Table 2.1: Results of the substitutability set tests

The results show that the lowest percentages where two products are ordered together in a basket correspond to the lower level categories  $S_2, S_3, S_4$ , whereas the highest percentages and number of times two items from the same SS appeared in a basket correspond to the substitutability set  $S_1$ . The results also show that the number of substitutes can also be high in the substitutability sets. Hence, it is clear that the lowest level categories can be defined as the substitutability sets.

#### 2.4.3.1 Exploratory analysis of products and sold baskets

We start by analyzing the sales patterns and the degree of intermittence. To do so, we analyze the data by using the demand classification proposed by Syntetos and Boylan (2005), referred to hereafter as SBC. Recall that in the SBC scheme, the demand is classified based on the average demand interval (ADI) and squared coefficient of variation of demand sizes ( $CV^2$ ) with the cut-off values of  $ADI = 1.32$  and  $CV^2 = 0.49$  (as illustrated in Figure 2.5 in Figure 2.5). Four classes are identified: smooth ( $ADI < 1.32$  and  $CV^2 < 0.49$ ), lumpy ( $ADI > 1.32$  and  $CV^2 > 0.49$ ), erratic ( $ADI < 1.32$  and  $CV^2 > 0.49$ ), and intermittent ( $ADI > 1.32$  and  $CV^2 < 0.49$ ). The class of smooth demand is the class where the demand has a high number of demand occurrences (i.e. low demand intervals) and a low variation of demand sizes. The erratic demand is also characterized by a high number of demand occurrences but it has high demand sizes' variability. The intermittent demand has few demand occurrences and a low variation of demand sizes. Finally, the class of lumpy demands is characterized by few demand occurrences but a high variation of demand sizes. This classification is used further in the chapter to identify the appropriate forecasting method in each class and the related inventory performance. We report in Figure 2.6 the percentage of products within each class.

Figure 2.5 illustrates the classification proposed by Syntetos et al. (2005) for a demand

over 12 months (from January to December). It shows the four classes: smooth, erratic, intermittent, and lumpy, with the two cut-off values between the classes,  $ADI = 1.32$  and  $CV^2 = 0.49$ .

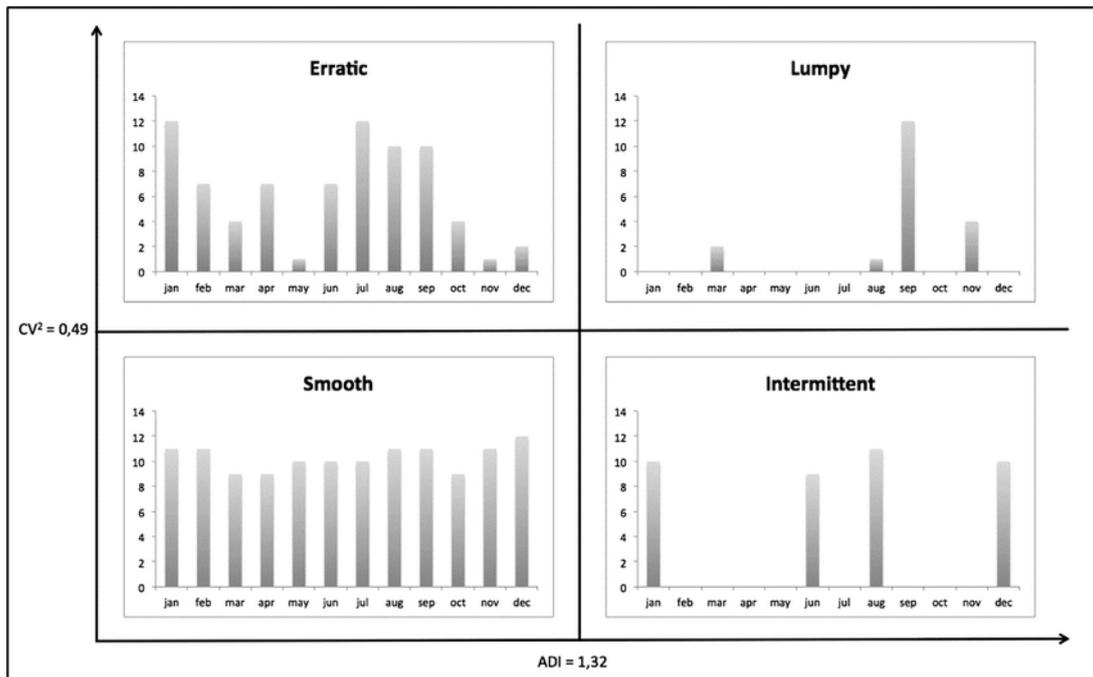


Figure 2.5: Illustration of the demand classification proposed by Syntetos et al. (2005)

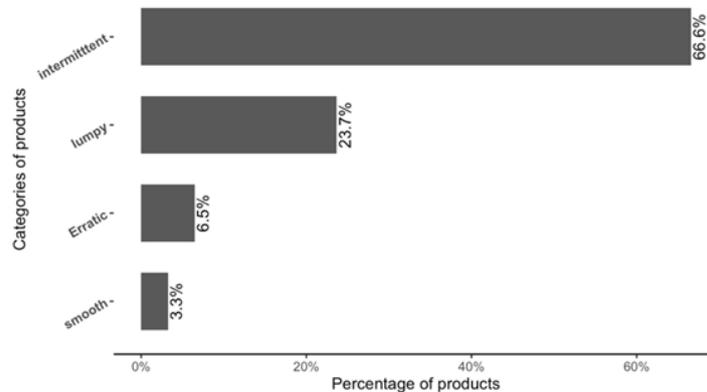


Figure 2.6: Percentage of products in each class in the online channel

The results in Figure 2.6 show that the majority of products (66.6%) are characterized by an intermittent demand pattern. This confirms that the online sales considered in our case are in line with the case presented by Fildes et al. (2019) in the physical retail space, where 52.2% of the sold products had an intermittent demand pattern. Figure 2.6 shows as well that products with lumpy and erratic demand patterns represent more than 30%, whereas those with a smooth demand constitute only 3.3% of the assortment. A focus on the average demand intervals of the products in each class shows that a product with an intermittent demand is sold on average once every 14 days, whereas a product with a smooth demand is sold on average every 1.1 days.

We now analyze the composition of orders made for the considered assortment of products. We present in Figure 2.7 the percentage of orders with the number of products in these orders. The results in Figure 2.7 show that 48.5% of the total baskets have only one

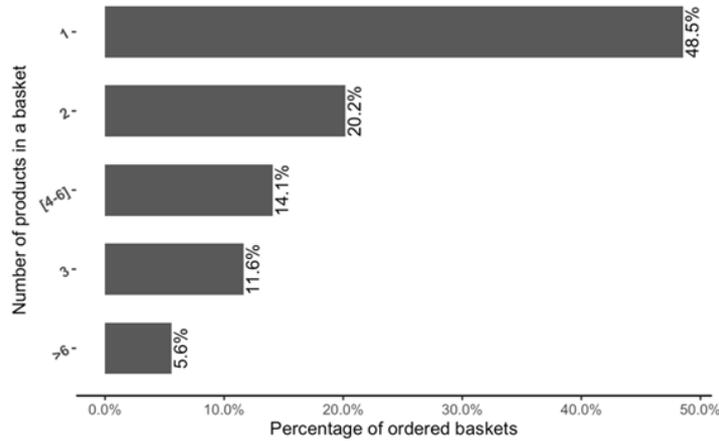


Figure 2.7: Percentage of single-product baskets and multiple-product baskets

product, whereas 51.5% of the baskets have two products or more. More precisely, the average number of products in a basket is 2.4; a third quartile equals 3, and a maximum equals 90. This concurs with our initial inference that there could be a singularity of buying for certain products, and effectively these products will have low degree and strength, whereas many products will have higher degrees, which means higher connectivity to other products as well as a higher strength.

We now check whether a product is usually sold in a single-product basket or in a multiple-product basket. Our initial intuition is that the new forecasting approach would work well for a product that is sold at least once with another product in the training horizon. Obviously, a product that is sold all the time as a single product in a basket would be equally well forecasted using the proposed ARIMAX or a basic ARIMA approach. To conduct this analysis, we report in Figure 2.8 the percentage of products and the percentage of times (across all orders) each product is sold in multiple-product baskets.

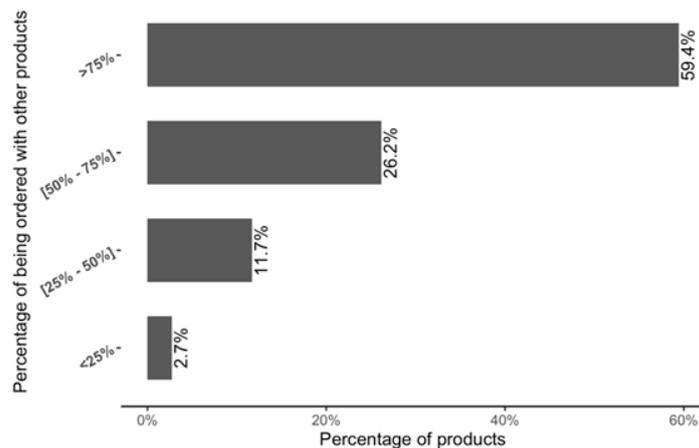


Figure 2.8: The number of times (percentage wise) each product was sold with other products

The results in Figure 2.8 show that only 2.7% of the products are sold individually, whereas 59.4% of the products are sold more than 75% of the time with other products. This confirms that the majority of products are sold with other products in a basket more than once, which also shows the importance of degree and strength attributes when forecasting demand for these products. Finally, we report in Figure 2.9 the degree distribution

of the four classes. The results show that the majority of products in the intermittent class have a weekly average degree of 4.08 and a weekly average strength of 4.9 (see Figure 2.11, which implies that there is a low frequency of products being ordered together across this class and that this behavior is translated as coincidental buying, while in the lumpy class, the ratio of strength to degree is higher, almost 2 to 1, which implies that there is a regularity of certain products being ordered together.

As for the erratic and smooth classes, they exhibit a much higher degree of connectivity with an average weekly of 71 connections per product in the erratic class and 105 connections per product in the smooth class. It is also noticed that they have an average weekly strength of 96.6 for the erratic class and 136 for the smooth one, which implies that within those classes, there is a habit of buying certain products together with high frequency among many customers. Note that the degree and strength attributes are indicators of the cross-category connectivity and complementarity of products as evidenced by the assortativity indicator presented earlier.

As for the within-category complementarity represented in Figure 2.10, we can notice that the average within category complementarity 0.0675. Recall that the within-category complementarity is the proportion of connections of a product to products within the same category. This average low measure implies that most of the products when they are bought in multiple product baskets, they are bought with different products from other categories. Only a few products are bought with other products from the same category. For example, based on our empirical findings, it was found that perfumes are usually bought with products from other categories while lipstick is bought many times with other lipsticks in the same category. Products like lipstick exhibit a high within-category complementarity. This concurs with the findings of the marketing literature regarding cross-category purchasing behavior [Cachon and Kök \(2007\)](#).

As for substitutability, it is the probability of a customer buying a product from a set of product substitutes given the number of product substitutes available. For example in our empirical study of panel A, it was found that the average substitutes for a given product is 89, with a min of 1, a median of 13, and a maximum of 2500 substitutes within a category. It is expected that as the number of substitutes in a category increases, the average sales for a product in this category will decrease as the customers will have more options to choose from, and subsequently the substitutability attribute captures this information. The relation between the number of substitutes in a category and the average sales for a product in this category is shown in Figure 2.10.

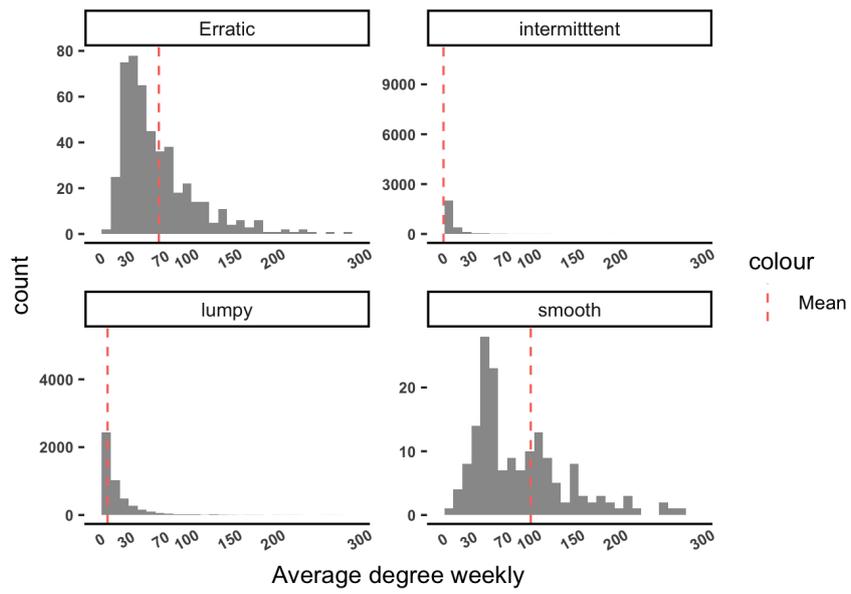


Figure 2.9: Degree distribution of products per class of products

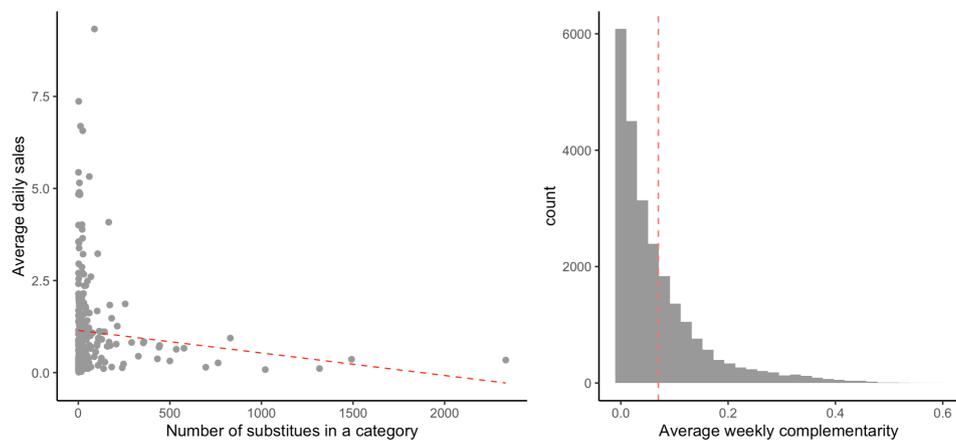


Figure 2.10: Relation between number of substitutes and average sales (left handside), distribution of complementarity (right handside)

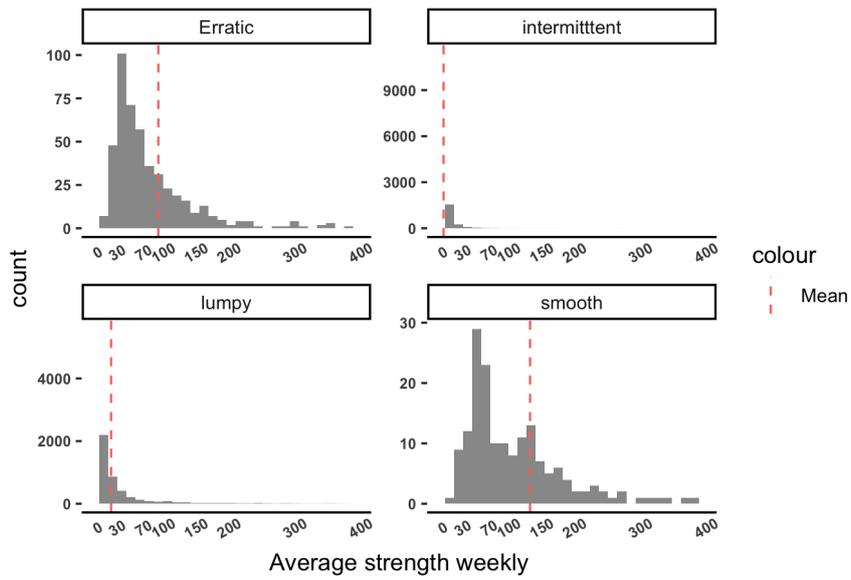


Figure 2.11: Strength distribution of products per class of products

#### 2.4.4 Attributes testing and validation results

Table 2.2 reports the  $R^2$  of the regression models, the root mean squared error (RMSE), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). It is shown that  $R^2$  improves as attributes are added as well as a decrease in the residual error.

<b>Linear Model</b>					
Attributes	$R^2$	<i>Adjusted R<sup>2</sup></i>	AIC	BIC	RMSE
Degree and strength	0.099	0.095	583	597	1.63
Degree , strength and substitutability	0.118	0.105	575	593	1.61
Degree , strength, substitutability and complementarity	0.119	0.120	570	592	1.60
<b>Linear Model with one lag of sales</b>					
Attributes	$R^2$	<i>Adjusted R<sup>2</sup></i>	AIC	BIC	RMSE
Degree and strength	0.136	0.130	563	582	1.52
Degree , strength and substitutability	0.145	0.137	560	581	1.51
Degree , strength, substitutability and complementarity	0.159	0.148	554	580	1.50
<b>Polynomial Model with one lag of sales</b>					
Attributes	$R^2$	<i>Adjusted R<sup>2</sup></i>	AIC	BIC	RMSE
Degree and strength	0.258	0.200	-9.64	76.6	1.134
Degree , strength and substitutability	0.297	0.234	-237	-122.3	1.036
Degree , strength, substitutability and complementarity	0.367	0.254	-1286.541	-1116.663	0.942

Table 2.2: Attributes significance

Table 2.3 reports the results of the tests of the regression models under the four demand classes, which shows the higher  $R^2$  values for the lumpy and erratic classes.

<b>Erratic Class</b>					
Attributes	$R^2$	$Adjusted R^2$	AIC	BIC	RMSE
Degree and strength	0.358	0.303	2068	2217	4.908
Degree , strength and substitutability	0.395	0.312	2077.47	2288.06	4.74
Degree , strength, substitutability and complementarity	0.431	0.314	2074	2325	. 4.59
<b>Lumpy Class</b>					
Attributes	$R^2$	$Adjusted R^2$	AIC	BIC	RMSE
Degree and strength	0.380	0.319	-413.99	-321.89	1.86
Degree , strength and substitutability	0.419	0.365	-882	-759.322	1.662
Degree , strength, substitutability and complementarity	0.4932	0.384	-3021.290	-2849.333	1.477
<b>Intermittent Class</b>					
Attributes	$R^2$	$Adjusted R^2$	AIC	BIC	RMSE
Degree and strength	0.166	0.109	187.59	266.499	0.433
Degree , strength and substitutability	0.204	0.136	123.185	228.65	0.412
Degree , strength, substitutability and complementarity	0.274	0.156	-185.826	-21.48	0.388
<b>Smooth Class</b>					
Attributes	$R^2$	$Adjusted R^2$	AIC	BIC	RMSE
Degree and strength	0.229	0.152	1906.253	2060.123	3.558
Degree , strength and substitutability	0.273	0.161	1916.30	2132.62	3.443
Degree , strength, substitutability and complementarity	0.308	0.163	-1916.854	2171.73	3.341

Table 2.3: Attributes significance per class under the polynomial regression model

## 2.4.5 Forecast accuracy based on different network horizons

We have made some preliminary tests using store retail data from panel B (related to the biggest store of the case study organization) to make a comparison between variations of the proposed methods where the length of the network horizon is equal to 7, 14, 21, and 28 days. Tables D.1 - D.2 show the average forecast accuracy (across all products) as well as the percentage of products where the method variation is better. The results relate to two of our proposed methods: ARIMAX( $h$ ) and LinReg-LS( $h$ ) where  $h$  is the network horizon length. The results show clearly that the network horizon of 7 days leads to the best forecast accuracy as compared to the other horizons.

		ARIMAX(7)	ARIMAX(14)	ARIMAX(21)	ARIMAX(28)
sMAPE	Average error	196	195	195	195
	% of products	54.4%	14.13%	12.36%	19.07%
MASE	Average error	1.46	1.46	2.02	2.51
	% of products	40.41%	19.87%	18.02%	21.69%
RMSE	Average error	0.287	0.313	0.372	0.378
	% of products	39.93%	21.08%	17.84%	8.56%
RMSSE	Average error	0.741	0.742	0.877	0.969
	% of products	39.93%	21.08%	17.84%	8.56%

Table 2.4: Forecasting accuracy results for different horizon lengths-ARIMAX

		LinReg-LS(7)	LinReg-LS(14)	LinReg-LS(21)	LinReg-LS(28)
sMAPE	Average error	196	196	196	196
	% of products	56.13%	11.15%	10.81%	21.91%
MASE	Average error	1.24	1.44	1.75	2.27
	% of products	43.03%	17.23%	15.22%	14.52%
RMSE	Average error	0.313	0.336	0.379	0.412
	% of products	40.26%	11.1%	10.81%	21.91%
RMSSE	Average error	0.683	0.741	0.807	0.934
	% of products	40.26%	11.1%	10.81%	21.91%

Table 2.5: Forecasting accuracy results for different horizon lengths-LinReg-LS

### 2.4.5.1 Forecasting methods' settings

We test the significance of the four attributes to predict sales using regression models. We start with degree and strength as independent variables, then degree, strength, and substitutability, and finally the four attributes as independent variables. Different regression models are considered, namely: (1) a linear regression model including only the attributes, (2) a linear regression model including one lag of sales and the attributes, and (3) a polynomial regression model including one lag of sales and the attributes. The tests are performed on every product from panel A. The average results (across all SKUs) are reported in Table 2.2, where the  $R^2$ , the adjusted  $R^2$ , the root mean squared error (RMSE), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC) of the different regression models are reported. The results show that, regardless of the regression model, the  $R^2$  increases (and the residual errors decrease) as the attributes are added in the regression models, which shows that using the four attributes of degree, strength, substitutability and complementarity provides the best fit. The results also show that when a non-linear regression model is used, the four attributes are more significant compared to the linear one. Finally, we show that higher  $R^2$  values are obtained when the lumpy and Erratic product classes are considered as reported in Table 2.3 under the polynomial model. Note that showing the value of considering the four attributes as well as the non-linear model, is behind the motivation to propose the LinReg-LS, PolyReg-LS, and lightGBM forecasting methods. Note also that the correlation between all the regressors was tested within the models and across all SKUs to ensure that the correlation is not significant.

Furthermore, for the implementation of Croston, LinReg-LS, PolyReg-LS, and lightGBM we coded the fitting and forecasting functions in R language while for the ARIMA and ARIMAX models, we used the R package Fable developed by Rob J.Hyndman. In order to have a fair comparison to the other forecasting methods, we parametrized the R package to use stationary ARIMA models. With regard to the network regressors in ARIMAX,

we developed a function that extracts the attributes from basket data. Note that the computational complexity is relatively low as, for our graph consisting of 24000 products, the algorithm takes less than two minutes to extract the attributes for all products each day. Finally, since the proposed forecasting approach requires a network horizon to determine the exogenous variables, we tested a network horizon with 7 days, 14 days, 21 days and 28 days. We recall that for every day of the forecast, a new network is built using the predetermined historical horizon since every horizon has a different set of products, orders, and accordingly attributes. As shown in section 2.5.6, the preliminary results reveal that the 7-day horizon overall provides the best performance, and will be further used in the chapter.

## 2.4.6 Forecasting accuracy analysis

The forecasting accuracy of the proposed approach is first evaluated using the Panel A data. We use two forecasting methods as a benchmark to show the performance of the proposed forecasting methods, namely, the ARIMA model and the Croston method. As argued in the literature review, these methods are commonly used in the retail context when dealing with products with intermittent demand patterns. The performance of the three methods is evaluated using a training/testing split of the demand history. We consider a within-sample (i.e., the training set of periods of the history) from January 1 to October 19, 2018, and we use the out-of-sample (i.e., the testing set of periods from October 20 to December 31, 2018, to evaluate the performance of the forecasting methods. The smoothing coefficients of Croston are selected such that they optimize the mean squared error (MSE) over the within sample. We evaluate the accuracy of the forecasting methods by means of four measures that are commonly considered in the intermittent demand context: the symmetric mean absolute percentage error (sMAPE), the mean absolute scaled error (MASE), the root mean square error (RMSE) and the root mean squared scaled error (RMSSE). MASE, sMAPE, and RMSSE are scale-free measures of forecasting accuracy. In both MASE and RMSSE, the scaling of the errors is based on the in-sample mean absolute error of the naïve forecasting method (Hyndman and Koehler 2006, Steinker et al. 2017, Babai et al. 2020, Siddiqui et al. 2022). In sMAPE, the scaling of the error in each period in the out-of-sample is based on an average of the actual demand and its forecast. Note that the RMSSE measure has been recommended in the M5 forecasting competition as it is known to be robust with series that have intermittent sales or series with higher demand values and the fact that it does not rely on values close to zero as the other relative measures do (Makridakis et al. 2020). RMSE is an indication of the standard deviation of forecast errors. This measure does suffer from a scale dependence issue but it is used in this chapter because it is the criterion used in optimizing the parameters of some forecasting methods.

### 2.4.6.1 Forecasting accuracy results of the online channel

Table 2.6 shows the empirical results using the data of Panel A. For each forecasting method and accuracy measure, we report two figures: the upper one represents the average error (across all products) and the one below represents the percentage of products where the forecasting method is the best for the accuracy measure. The lower forecasting errors are highlighted in bold font.

		Croston	ARIMA	LinReg-LS	PolyReg-LS	ARIMAX	LightGBM
sMAPE	Average error	185.909	186.773	<b>183.716</b>	184.182	184.161	188.317
	% of products	27.1%	31.4%	10.4%	15.8%	13.1%	2.2%
MASE	Average error	2.613	1.267	1.283	1.243	1.150	<b>0.723</b>
	% of products	1.34%	8%	2.3%	2.1%	5.9%	80%
RMSE	Average error	1.229	0.603	<b>0.491</b>	0.510	0.517	<b>0.491</b>
	% of products	21%	15%	11.6%	10.9%	11.4%	30.1%
RMSSE	Average error	0.776	<b>0.587</b>	0.621	0.676	0.610	0.611
	% of products	21%	15%	11.6%	10.9%	11.4%	30.1%

Table 2.6: Forecasting accuracy results

The results in Table 2.6 show that the proposed forecasting methods lead to the best performance when compared to the two benchmarks in terms of both the forecast accuracy and the highest number of products within the assortment. The method that is associated with the lowest forecasting accuracy is the Croston method. This is not surprising since the relatively low empirical performance of Croston when dealing with highly intermittent and lumpy demand data is shown in the literature (Babai et al. 2014, Syntetos et al. 2015). Note that LightGBM provides the best accuracy results in MASE and RMSE and in RMSSE (for the highest number of products within the assortment). However, for sMAPE it is the least performing method, which is due to the fact that for most of the SKUs with intermittent demand, LightGBM produces forecasts that are equal or close to zero, which leads to high values of sMAPE (since both demands and forecasts are close to zero in the denominator of the sMAPE measure). Such an issue with the sMAPE measure when forecasting intermittent demands using methods producing many zero forecasts is well known in the literature (Teunter and Duncan 2009, Makridakis et al. 2020).

In order to gain additional insights on the comparative performance of the forecasting methods without having the scaling and metric-dependence issue, the performance of the methods was statistically measured for significant differences through pair-wise comparisons of the absolute error, the squared error, and the symmetric absolute percentage error between sales and forecast per product and per day. The results of the tests are reported in Section 2.5.8. The results show that all our proposed methods perform significantly (at the 1% and 5% levels) better than Croston’s method. It is also shown that both ARIMAX and LightGBM perform significantly better than ARIMA.

Furthermore, to better understand the performance of the proposed approach (and therefore the four forecasting methods) with respect to the connectivity of products, we analyze the forecasting accuracy results per demand class according to the SBC classification scheme. The results for the four classes are reported in Table 2.7 and the lower forecast errors are highlighted by a bold font.

Lumpy					Smooth				
	sMAPE	MASE	RMSE	RMSSE		sMAPE	MASE	RMSE	RMSSE
Croston	180.67	4.468	2.438	0.798	Croston	88.299	0.912	2.386	2.612
ARIMA	181.627	1.473	0.956	<b>0.554</b>	ARIMA	<b>86.958</b>	0.867	2.332	0.407
LinReg-LS	<b>171.392</b>	1.026	0.799	0.557	LinReg-LS	130.053	0.442	1.153	<b>0.381</b>
PolyReg-LS	173.051	1.083	0.820	0.589	PolyReg-LS	129.669	0.437	<b>1.151</b>	0.383
ARIMAX	175.975	1.097	<b>0.713</b>	0.560	ARIMAX	88.488	0.823	2.116	0.400
LightGBM	173.762	<b>0.828</b>	0.793	0.573	LightGBM	95.462	<b>0.406</b>	1.202	0.393
Intermittent					Erratic				
	sMAPE	MASE	RMSE	RMSSE		sMAPE	MASE	RMSE	RMSSE
Croston	189.943	1.596	0.539	0.724	Croston	92.672	0.812	2.361	1.423
ARIMA	190.582	1.156	0.384	<b>0.607</b>	ARIMA	93.089	0.815	2.377	0.337
LinReg-LS	189.139	1.189	<b>0.370</b>	0.652	LinReg-LS	136.979	0.388	1.141	0.307
PolyReg-LS	189.334	1.372	0.390	0.717	PolyReg-LS	137.170	0.375	<b>1.101</b>	<b>0.304</b>
ARIMAX	<b>188.987</b>	1.183	0.394	0.637	ARIMAX	<b>91.334</b>	0.737	2.187	0.329
LightGBM	194.935	<b>0.707</b>	0.372	0.634	LightGBM	114.568	<b>0.335</b>	1.134	0.307

Table 2.7: Forecasting accuracy results per category of products

The results in Table 2.7 show that in the lumpy and erratic demand classes, the comparative performance significantly increases in favor of our proposed methods compared to the benchmark methods. The results per demand class show that for example the relative error reduction by using ARIMAX for the lumpy class is much higher compared to the results for all classes, going up to 3.1% when compared to ARIMA (under sMAPE) and 2.6% when compared to Croston. Under MASE, the error reduction is equal to 26% when compared to ARIMA and goes up to 75% when compared to Croston. The results also show that for the erratic demand class, the relative error reduction by using PolyReg-LS reaches 9.8% when compared to ARIMA and 78.6% when compared to Croston (under RMSSE). Under RMSE and MASE, the relative reduction is even higher than in RMSSE. This high relative reduction also holds for LightGBM and LinReg-LS. The outperformance of the proposed approach under these classes can be explained by the fact that under lumpy and erratic demands, demand sizes are higher and the attributes-based forecasting approach better captures the connection between the sold products, which improves its forecasting accuracy. This is in line with the findings of Section 4.2, where we have shown that the connectivity of products is high under the lumpy class. Moreover, as expected, the results reveal that for the intermittent demand class, the relative performance of Croston improves. The Croston method has been developed for the specific case of intermittent demand with stationary demand intervals and demand sizes, which means that when the degree of lumpiness or erraticness increases, the method loses its performance.

#### 2.4.6.2 Omnichannel forecasting results

For the purpose of omnichannel forecasting, we consider the data of Panel B, which relates to the online and store channels. We first analyze the characteristics of the baskets in the omnichannel case. The results are shown in Figure 2.12. On the left, we report the percentage of single-product baskets and multiple-product Baskets and on the right, we report the number of times (percentage-wise) each product was sold with other products. The results in Figure 2.12 show that 62% of the total baskets have two products or more. This exceeds the percentage in the online case (reported in Figure 2.7), which further endorses the importance of the shopping basket behavior in the omnichannel case. Moreover, the results in Figure 2.12 show that 86.6% of the products are sold more than

75% of the time with other products. This percentage is higher than the online channel case (reported in Figure 2.8). This further confirms that the majority of products are sold with other products in a basket more than once, which is more pronounced in the omnichannel case.

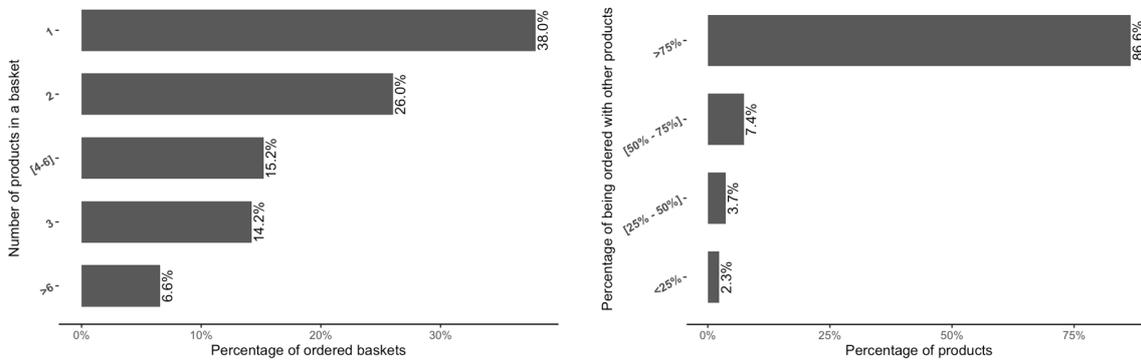


Figure 2.12: Baskets characteristics for the omnichannel case

Next, a joint sales forecast is made after aggregating the daily online and store sales (i.e. at any day  $t$ , the aggregate sale used to forecast is the sum of the sales from both channels on the same day). We analyze the forecasting accuracy of all the forecasting methods using the online sales channel, the sales of the store channel, and the omnichannel aggregated sales. We report in Table 2.8 the average sMAPE results for the three cases. Results are shown using only sMAPE because this is a relative error measure that enables a fair comparison among the three cases.

	Channel	Croston	ARIMA	LinReg-LS	PolyReg-LS	ARIMAX	LightGBM
sMAPE	Online sales	199.778	199.807	198.772	198.659	199.788	198.824
	Store sales	197.413	197.525	188.756	188.898	197.301	194.240
	Omnichannel sales	<b>197.329</b>	<b>197.525</b>	<b>186.339</b>	<b>186.583</b>	<b>197.245</b>	<b>193.260</b>

Table 2.8: Omnichannel forecasting results

The results in Table 2.8 show that by using the omnichannel case, the sMAPE error decreases, which means that the forecasting accuracy improves by forecasting the sales based on an aggregation of the two channels. It is important to note that the improvement of the performance with the joint forecasting of both sales channels is realized under all the forecasting methods. Table 2.8 also shows that although the ARIMAX method is not the most accurate for the online sales channel, it is better than the two benchmarks in the case of the omnichannel sales. Furthermore, it should be noted that under this panel B data and omnichannel case, the accuracy of the LightGBM method improves and it becomes higher than that of the two benchmarks, even under the sMAPE measure. This shows that in the case of the omnichannel sale, the connectivity between products is better captured, which improves the performance of LightGBM and also of our other proposed forecasting methods.

## 2.4.7 Statistical significance test for forecasting methods

The Z-test significance for proportions is used to measure the significant differences between methods through pair-wise comparisons on the absolute error (AE), the squared

error (SE), and the symmetric absolute percentage error (sAPE) between sales and forecast per product and per day. Note that the comparison using these errors per period and per product enables the comparative performance of the forecasting methods without having the scaling issue resulting from the average across SKUs. It should be noted that for each product MAD, MSE, and sMAPE are the mean values of AE, SE, and sAPE respectively. Moreover, comparing SE per period of two methods is sufficient to have the comparative performance of RMSSE since the Naïve mean error in the denominator of this measure is the same between the two methods. Let  $p_1$  denote the percentage of days on which method 1 has a higher error than method 2 (i.e., the latter performs better than the former) and  $n$  be the size of the considered sample (total number of days and products). The Z-test statistic is then given by:

$$Z = \frac{p_1 - 0.5}{\sqrt{\frac{25}{n}}} \quad (2.4)$$

The critical values at 5% and at 1% level are equal to 1.6449 and 2.3263, respectively. Note that if the error (i.e AE, SE, and sAPE) for two given methods are equal, they are eliminated from proportion calculation. The Z-test statistic values are indicated, for all pair-wise comparisons of the considered methods for the AE, SE, and sAPE errors, in Table 2.9. Positive values indicate that method 1 is more accurate than method 2.

Method 1	Method 2	AE		SE		sAPE	
		$p_1$	Z	$p_1$	Z	$p_1$	Z
ARIMA	Croston	71.0%	495.784	71.0%	497.936	63.7%	161.239
ARIMA	ARIMAX	47.0%	-68.560	47.1%	-67.726	35.4%	-196.197
ARIMA	LightGBM	14.7%	-799.205	14.9%	-793.723	13.9%	-791.125
ARIMA	LinReg-LS	54.9%	115.948	55.4%	127.711	59.1%	110.559
ARIMA	PolyReg-LS	54.5%	105.853	55.0%	117.370	58.4%	103.567
Croston	ARIMAX	31.5%	-436.403	31.5%	-438.893	24.7%	-345.757
Croston	LightGBM	13.1%	-874.175	13.2%	-871.468	12.8%	-854.194
Croston	LinReg-LS	27.9%	-523.274	28.2%	-517.514	42.6%	-74.586
Croston	PolyReg-LS	27.4%	-534.978	27.2%	-529.320	42.5%	-78.096
ARIMAX	LightGBM	15.8%	-734.378	16%	-729.812	15.5%	-716.095
ARIMAX	LinReg-LS	62.0%	281.211	62.6%	294.698	71.2%	284.49
ARIMAX	PolyReg-LS	58.3%	194.559	58.8%	207.192	70.1%	278.420
LightGBM	LinReg-LS	86.3%	846.676	86.3%	846.676	86.7%	830.980
LightGBM	PolyReg-LS	86.4%	845.529	86.4%	845.528	86.9%	831.519
LinReg-LS	PolyReg-LS	46.2%	-90.352	46.2%	-90.352	50.2%	-2.465

Table 2.9: Z-test scores - absolute, squared and symmetric absolute percentage errors

## 2.4.8 Omnichannel inventory performance

We start the inventory performance investigation by conducting a goodness-of-fit analysis to test the fit of different distributions with the demand in the different channels. This is needed in order to select the best demand distribution for the lead-time demand. We consider five distributions: normal, Poisson, gamma, negative binomial distribution (NBD), and stuttering Poisson (StPoisson). These distributions are commonly considered in modeling intermittent and non-intermittent demand data (Syntetos et al. 2013, Turrini and Meissner 2019). We show in Table 2.10 the results obtained with a Kolmogorov-Smirnov test at a 5% significance level. For each distribution, the "Fit" percentage shown in Table

2.10 corresponds to the percentage of products for which the hypothesis that the demand data fits the distribution (i.e., the p-value is less than the critical value calculated using the K-S statistical table for 5%), is not rejected.

		Poisson	NBD	Normal	Gamma	StPoisson
Online sales	Fit	99.23%	100.00%	94.04%	99.65%	99.98%
	No fit	0.77%	0.00%	5.96%	0.35%	0.02%
Store sales	Fit	98.54%	99.80%	93.49%	97.77%	99.78%
	No fit	1.30%	0.04%	6.35%	2.07%	0.07%
Omnichannel sales	Fit	97.83%	99.87%	90.97%	96.84%	99.81%
	No fit	2.17%	0.13%	9.03%	3.16%	0.19%

Table 2.10: Empirical goodness-of-fit results

The results show the strong empirical fit of the NBD followed by the StPoisson distribution, with normal being the distribution associated with the lowest fit. These results are expected knowing that the demand data are characterized by a high degree of intermittence.

Therefore, for the purpose of the inventory performance investigation, NBD is selected to model the lead-time demand. The lead time is fixed to 2 days. For the purpose of the analysis, we fix three target CSLs—90%, 95%, and 99%—and for target CSL and each scenario, we measure the average inventory holding volumes, the average back ordering volumes, and the achieved CSLs (averages calculated over the evaluation period and across all products). Note that the last 73 days (from October 20, 2018, to December 31, 2018) are used to evaluate the performance. The inventory performance results of the different forecasting methods are reported in Table 2.11.

		Holding Volumes	Backordering Volumes	Achieved CSL (%)
Croston	CSL=90%	1.159	0.045	91.42
	CSL=95%	1.691	0.027	94.87
	CSL=99%	3.721	0.005	99.19
ARIMA	CSL=90%	0.671	0.030	93.25
	CSL=95%	1.101	0.017	96.07
	CSL=99%	2.762	0.007	98.22
ARIMAX	CSL=90%	0.723	0.033	92.51
	CSL=95%	1.150	0.019	95.42
	CSL=99%	2.754	0.008	98.05
LinReg-LS	CSL=90%	0.872	0.020	95.24
	CSL=95%	1.637	0.011	97.58
	CSL=99%	3.493	0.002	99.66
PolyReg-LS	CSL=90%	1.329	0.024	94.96
	CSL=95%	1.918	0.012	97.,50
	CSL=99%	3.773	0.002	99.58
LightGBM	CSL=90%	0.051	0.101	79.06
	CSL=95%	0.066	0.101	79.18
	CSL=99%	0.120	0.099	79.45

Table 2.11: Inventory performance results

The results in Table 2.11 show that except lightGBM, all the forecasting methods are able to achieve the 90% target CSL whereas, only Croston’s method, LinReg-LS and PolyReg-LS over-achieves the 99% target CSL. Note that LightGBM is associated with a very poor inventory performance. By looking at this method’s performance closely,

we have noticed that it results in many zero forecasts for several SKUs which leads to low holding volumes, low achieved CSLs and high backordering volumes. This inventory performance behavior is similar to that shown in the literature when for example the zero forecast method is used (Teunter and Duncan 2009). In fact, a high forecast accuracy for the zero forecast method is associated with a very low inventory performance. The results also show that our proposed LingReg-LS method provides the highest achieved CSL for all target CSLs. However, it should be noted that these results cannot lead to a conclusive comparative performance between the methods since in many cases, higher holding volumes lead to higher achieved CSLs and lower backordering volumes. In order to have a fair comparison of the forecasting methods. We plot their efficiency curves showing: (1) inventory holding versus backordering volumes. and (2) inventory holding versus achieved CSLs. In the former curves, the method that has its curve closer to the x-axis for a certain inventory holding volume implies a lower backordering volume and thus higher efficiency. In the latter, the method that has its efficiency curve further to the x-axis (i.e. on the top) for a certain inventory holding volume implies a higher achieved CSL. The efficiency curves of the different forecast methods are reported in Figure 2.13. Note that the efficiency curve of LightGBM does not appear in the figure due to the huge performance gap with the other efficiency curves, knowing that we used this scale for readability purposes. Note also that all these performance results are based on Panel B data.

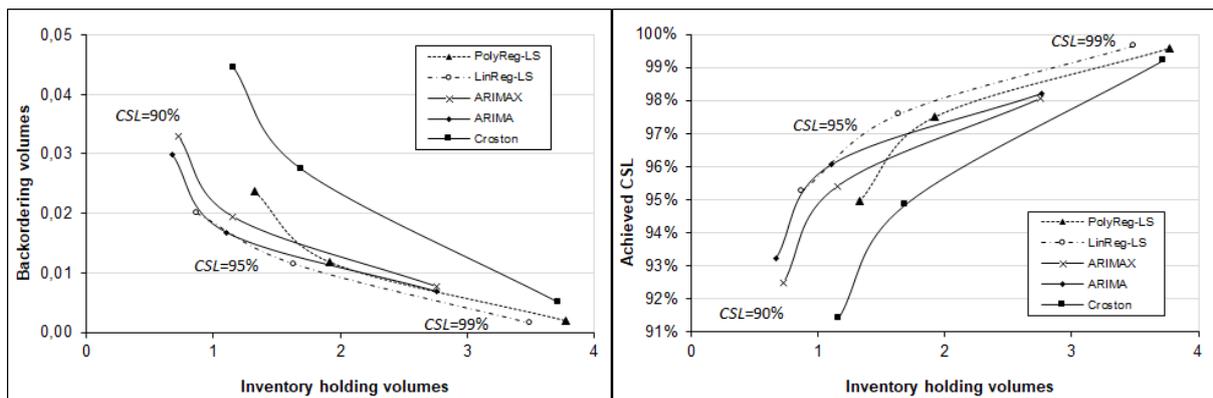


Figure 2.13: Efficiency curves of the different forecasting methods

The results in Figure 2.13 show that the LinReg-LS forecasting method leads to the highest inventory efficiency both when holding volumes-Backordering volumes and holding volumes-achieved CSL trade-offs are considered, whereas Croston’s method leads to the lowest efficiency. The over-achievement of the CSL at 90% when Croston is used despite its low forecast accuracy and inventory efficiency may be attributed to its positive bias as often reported in the literature (Babai et al. 2020). The efficiency curves also show that ARIMA leads to a higher inventory efficiency than ARIMAX. This is not in line with the comparative forecast accuracy of the two methods but this is not surprising since it has been shown in the literature that a higher forecast accuracy does not necessarily lead to a higher inventory performance (Syntetos et al. 2016). This discrepancy between forecast accuracy and inventory performance may be partially due to the difference between the forecast horizon considered to show the accuracy and the replenishment lead time used in the inventory control policy. Tackling this issue constitutes an interesting avenue for further research.

We now report in Figure 2.14 the efficiency curves contrasting the three fulfillment scenarios. We do so for both the ARIMAX and LinReg-LS methods (since those are the best-proposed methods in terms of both forecast accuracy and inventory performance).

We do not report the achieved CSLs in these figures because for the DISF scenario, the CSL cannot be reported and compared to that of the two other scenarios. The results in Figure 2.14 show that the shared inventory and joint forecast scenario leads to the lowest inventory backordering for fixed inventory holding volumes, which indicates that this scenario is associated with the highest inventory efficiency.

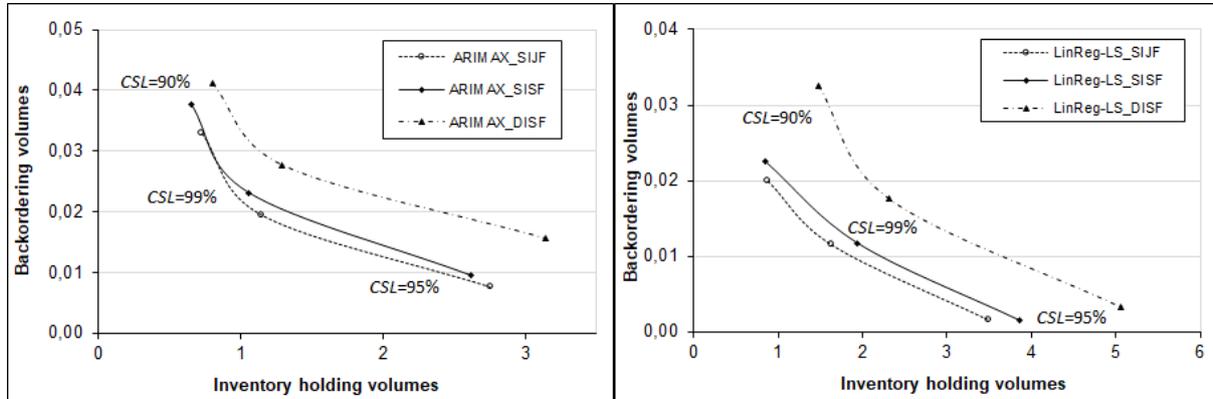


Figure 2.14: Efficiency curves: holding Versus backordering volumes of the fulfillment scenarios

## 2.5 Conclusion

There is an agreement in the omnichannel retail literature that forecasting omnichannel sales is a challenging task, and research devoted to dealing with this issue is lacking. This study has proposed a new approach to forecasting demand in an omnichannel retail context using data on customers' shopping baskets. The forecasting approach builds on findings from the marketing literature and graph theory to identify attributes of the products, which can be used with a time-series forecasting model. We have conducted an empirical analysis of the proposed forecasting approach and other benchmark forecasting methods commonly used in the retail context. A dataset of an assortment of more than 24,000 products in the online and store channels from a global leading cosmetics retailer was used for this purpose. Our study is the first to make an empirical analysis of both online and store sales and the first to use information on shopping baskets to forecast demand.

We have characterized the empirical behavior of more than 2 million online orders. Our investigation has shown that more than 95% of the sold products are characterized by an intermittent demand pattern, with 30% of them having a high lumpiness. By analyzing the composition of orders, we find that more than 50% of products are sold in baskets with more than two products per basket. Our findings reveal as well that the network of omnichannel retail sales is characterized by low density and a high average path length, and thus cannot be considered to be a small-world network. However, the connectivity (i.e., degree) distribution is heavily skewed to a few products in the network. Through a regression analysis, our study also reveals the importance of considering the degree, strength, substitutability, and complementarity of the network as good regressors for forecasting purposes.

The empirical assessment of the performance of the forecasting methods shows that one of the most popular methods for intermittent demand forecasting (the Croston method) leads to poor forecasting accuracy in the omnichannel context. This underperformance is accentuated in the product category with lumpy demand patterns. Our proposed forecasting approach, which uses the degree, strength, substitutability and complementarity

attributes within four forecasting methods, namely: ARIMAX, LinReg-LS, PolyReg-LS and LightGBM, tackles this issue by improving the forecasting accuracy, especially for lumpy demand patterns due to the high connectivity of products in this demand class. These empirical findings highlight the value of considering attributes of the linkages between sold products in baskets for omnichannel forecasting purposes. The results also show that the machine learning method, LightGBM, is associated with a high forecast accuracy but leads though to poor inventory performance.

Furthermore, the empirical study also shows that joint forecasting based on the sales of both online and store sales leads to higher forecasting accuracy and inventory performance. Such an omnichannel forecasting approach is recommended to consolidate inventories at stores, likely leading to a considerable reduction of inventory shortages. These findings enable an omnichannel network designer to gain valuable insights on how to deploy inventories in a set of fulfillment centers and on how stores could play a major role in efficient urban fulfillment. The findings also provide interesting insights on omnichannel assortment planning based on the data-driven network analysis of the baskets.

In summary, omnichannel retailing has changed the purchasing behavior of customers in recent years, especially in online shopping, which has led to higher complexity in supply chain demand forecasting. Nowadays customers buy a variety of products in baskets that do not share similar characteristics and across various channels. In this chapter, we propose a new approach to forecasting demand, driven by data on customers' shopping baskets. Drawing on network graph theory and findings from the marketing literature, we identify for a given product four attributes to promote the connectivity with other products sold together in a basket: *degree* and *strength* for cross-categories connection, *substitutability* and *complementarity* for within-categories connection. These attributes are used as predictor variables within four proposed methods: an autoregressive integrated moving average model with exogenous variables (ARIMAX), a linear and a polynomial regression with one lag of sales and a machine learning method. We conduct an empirical investigation using online and physical sales related to an assortment of 24,000 products from a major cosmetics retailer in France. We provide empirical evidence that using the shopping basket data with the proposed forecasting methods improves the forecasting accuracy and the stock control performance in omnichannel retailing. We also show that there is a benefit from joint forecasting of the online and store channels and a benefit of shared inventory between both channels in terms of shortage reduction. In the next chapter, we reduce the focus on the product and increase the focus on the basket and how it can be used to minimize the cost of omnichannel fulfillment and reduce the waiting time for order delivery to customers.

# Chapter 3

## A new basket data-driven approach for forecast-based anticipatory shipping

### 3.1 Introduction

#### 3.1.1 General introduction and motivation

In the previous chapter, omnichannel fulfillment was under the assumption that inventory of online and physical retail can be held in one place mainly the store (shared inventory strategy). While this is applicable and implemented in many omnichannel contexts, the space inside the retail store could be limited to a select number of online products and not for all the online products in a demand zone of this store. In case store or mini-hub space is limited, such strategy of shared inventory as stated in the previous chapter will not work especially for high-demand zones. With this in mind, customers now expect products to be delivered to them faster than before and they have less tolerance for stock-outs [Gallino et al. \(2019\)](#). For such cases, this creates the need for a more efficient advancement of baskets to stores/mini-hubs of online demand zones.

[Fisher et al. \(2019\)](#) deduced by an empirical investigation of an omnichannel retailer that there is an expected 1.45% increase in sales for every business day delivery reduction. AMAZON is one of the pioneers to offer 2-day delivery to customers through their prime service for example and BM retailers are catching up [Jasin et al. \(2019\)](#). Not only AMAZON, Google, and Instacart are also developing strategies to deliver physical products to customers within 2 hours from online orders [Chopra \(2018\)](#).

For omnichannel retailers, utilizing store space to offer SFS (ship from store) and BOPS (Buy online pick from store) are means to achieve such optimistic deadlines [Rooderkerk et al. \(2023\)](#). Such initiatives save on capital investments in fulfillment centers, reduce transport cost because of consolidation, and increase customer satisfaction because of the reduction of delivery times. Yet packaging of baskets and finalizing order shipping details is still more economical in fulfillment centers due to pooling and economic of scales [Alptekinoglu and Tang \(2005\)](#), [Bretthauer et al. \(2010\)](#).

The data panel of a cosmetics retailer used in this work reveals that 65 % of ordered baskets by customers have two distinct products or more, and there were 2.5 ordered products on average per basket. This increases the complexity of advancing online baskets near to customer's location before the customer makes the order as the content of the basket is challenging to forecast in advance due to all the bi-product combinations that could constitute a basket.

This raises the question of how data from multiple product baskets could be considered

by operations managers to improve the fulfillment of basket orders made by customers by predicting the basket content in advance thus reducing the delivery time to customers and benefiting from economies of scale in transportation, picking and packaging. [Boone et al. \(2019\)](#) highlighted the existence of influences between products purchased within customers' baskets and they stated the use of basket data in forecasting as one of the research avenues that are yet to be explored. Unlike [Omar et al. \(2022\)](#) where they aimed at utilizing the basket data to improve forecasting of individual products and consequently inventory performance on omnichannel demand streams. In this research, we explore predicting the whole basket before the customer makes the order to optimize many aspects of order fulfillment and increase customer satisfaction.

There has been limited and recent literature on anticipatory shipping where retailers attempt to predict customers' orders, however, to the best of our knowledge, there is no previous research that used basket data to forecast and advance the basket itself with its content -not only the product level- in supply chains. While there is recent literature on predicting customer demand and advancing products near to the customers. For example; [Ren et al. \(2022\)](#) introduced a forecasting algorithm that recommends items to be advanced based on product demand data with the individual product-by-product demand prediction. In addition, [Lee \(2017\)](#) proposed a Genetic algorithm optimization model complemented by baskets data association rules to anticipate the basket in advance on clustered historical data. A close approach was proposed by [Viet et al. \(2020\)](#) where they used basket data association rules with the addition of a time constraint for perishable food items for the agri-food industry. Our work adds to the existing literature on a number of levels. First, we take advantage of the rich quantitative attributes of graph theory compared to association rules where association rules are limited to confidence and support attributes and do not have an explicit quantity measure. Second, we model the association between products as an arc in a graph which gives flexibility in adding time constraints to the graph by defining the duration of the graph from historical baskets data as well as adjusting the future forecast horizon, for example, one day, one week, one month which suits perishable and non-perishable items. Third, our approach sources only from product data which are abundantly available for most retailers compared to customer information and click stream data, also, the prediction outcome in our proposed approach is a regression variable that allows for high quantities prediction compared to classification problems. Finally, we predict the content of the baskets which allows for pre-processing activities of baskets such as allocation, price tagging, and packaging before a customer makes an order.

Baskets being correctly predicted will enable retailers to maximize utilization in both fulfillment centers and stores near the customer and at the same time reduce delivery time and overstocking or obsolescence of baskets. Furthermore, It will reduce the load on store staff with regard to packaging if shipped from the point of sale. This motivates our research work to propose a new approach that predicts the demand using basket data for the purpose of anticipatory shipping.

### 3.1.2 Business context

While this project focuses on omnichannel retailers that have store space that can be used for online fulfillment, the work also applies to using warehouse or mini-hub space to satisfy demand zones near the customer. This project is in collaboration with a cosmetics retailer that owns its Brick and Mortar stores, has a prominent website for online shopping, and entered into partnerships with other BM/online retailers where the retailer is not operating or does not have a distribution infrastructure. The company has a broad assortment of cosmetics and perfumes. The project focuses on availing baskets to online customers

in 2 hours to 48 hours through anticipatory shipping. The retailer has been operating traditionally through retail stores with a periodic replenishment to stores via a centralized retail warehouse. The frequency of replenishment depends on the size and popularity of stores. With the introduction of online sales at the beginning of the 2000s, online orders were fulfilled through a Central online fulfillment center. Recently, the retailer has been considering 3 approaches for fulfilling online orders besides sending to customers directly from the centralized fulfillment center. One of them we are proposing in this work is ship-to-store (STS), and the other two are previously discussed in Chapter 2 which are Ship-from-store (SFS) and Buy-online pick-up from store (BOPS). In the new proposed STS concept, the baskets are sent consolidated and packaged to the store of the demand zone before the customers' orders and no further operations are required from store staff. We anticipate that this will work well for certain cases illustrated in the proposed fulfillment strategies of basket advancements as opposed to sending the baskets not packaged. Note that for SFS, the demand is served from store stock or sent from the fulfillment center then packaged and processed in-store. Figure 3.1 illustrates the retailer's downstream AS-IS versus the TO-BE supply chain schema composed of a set of customers and their locations, a selected store, a retail warehouse, and a fulfillment center. The dashed arrows represent decisions while the complete arrows represent flows. Note that the upper decision schema in Figure 3.1 relates to the fulfillment decision-making without applying anticipatory shipping (AS-IS) While the lower-level decision schema represents the anticipatory (TO-BE) strategy.

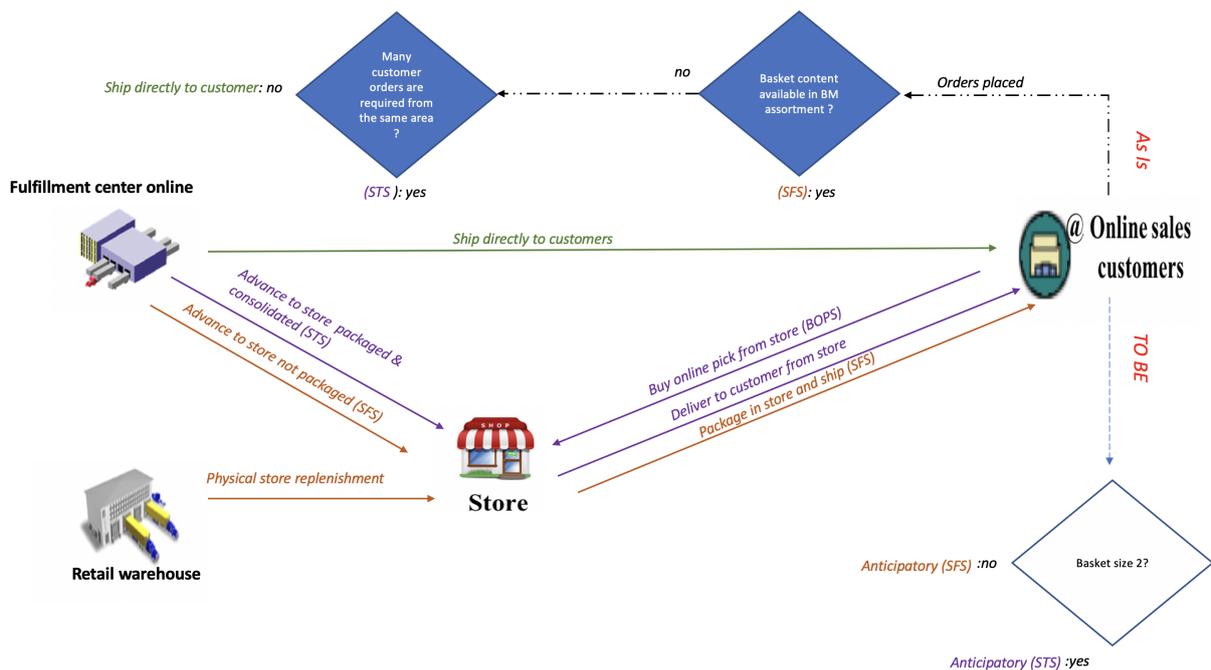


Figure 3.1: Omni-channel fulfillment- AS-IS VS TO-BE

SFS takes advantage of the existing physical network by turning certain store locations to ship-from points for online sales. The stock for SFS can come from a BM store replenishment stock that happens to be available online as well. However, SFS requires online fulfillment operations in-store. Note that STS/SFS are not distinguished by customer but it's a fulfillment strategy on where the basket fulfillment is finalized. However, BOPS requires extra effort from customers for fulfillment. The average delivery lead time for the ship to a customer from a central fulfillment center is 4 days while if the online order items are available in a nearby store with packaging capabilities is 2 hours-2 days,

noting that most of the shipments are shipped with the former option.

In the AS-IS strategy, if a customer places an order online and the content of the basket is not available in the nearby store, it is shipped directly from the fulfillment center (green arrow), while if there is a high volume of customers ordering online from the same region. The baskets could be packaged, consolidated, and shipped to a store near the customer demand zone following an STS approach (purple arrow). This reduces shipping costs compared to shipping directly to the customer but increases the time to delivery for the customer. The third flow for AS-IS fulfillment is SFS where the content of the basket is already available in-store through regular store replenishment and the store has packaging, picking, and shipping capabilities and performs last-mile delivery to the customer (orange arrow) or the customer may pick the basket up from the store following BOPS. Note that in the AS-IS traditional flow, no decision is made before the order of a customer. Hence, unless the online basket content is already available in a store near the customer from BM replenishment, online orders suffer long lead times customers.

With this AS-IS strategy, it is difficult to package baskets in a fulfillment center before customers order baskets as they usually buy online baskets that contain several products [Omar et al. \(2022\)](#). There is also another possibility that part of the basket is available in a nearby store and the rest of the basket is in a centralized location, which leads to online orders suffering idle time and being delayed to customers. In addition, online orders are usually sent from a centralized hub where packaging capabilities are economical but from a distance not close to every demand zone leading to high single shipments transportation costs due to irregular times of shipment and not benefiting from consolidation or economics of scale like the physical channel replenishment. Only in the case where the total basket is already available in a nearby store then it be delivered within 2 to 48 hours provided that the store has packaging capabilities which are a few orders from the total online baskets.

In the TO-BE strategy that this work is proposing in the lower decision schema in Figure 3.1, baskets are advanced to store before an order is placed. Baskets are classified into baskets of size 2 (BS2) which are 2 items ordered in a basket and baskets of size more than 2 (BS>2). This classification is important due to the natural structure of baskets as a basket of size 2 consists of one arc and can be packaged without high uncertainty while a basket of size more than 2 consists of 2 arcs or more increasing uncertainty with regards to which arc combinations make a basket of size more than 2. Thus, further processing is needed for baskets of size >2. We propose a two-step basket prediction approach that uses bi-product (arc) attributes to predict the quantity of the bi-product links for BS2 and BS>2 separately. Hence, every bi-product link (arc) has two continuous prediction targets (BS2 and BS>2). For BS>2, an algorithm is triggered to retrieve BS>2 baskets from arcs. Note that this is not required for BS2 as BS2 has only one arc.

Recall that in the AS-IS strategy, the baskets are shipped from the store only if the basket content is available in-store through BM replenishment which is minimal compared to the volume of online baskets shipped directly from the fulfillment center. In the TO-BE strategy, the predicted baskets take a separate fulfillment path based on size. If the size of the predicted basket is 2, then the basket is advanced packaged from the fulfillment center one week before the actual order is to be made. If the basket is of size more than 2, then the basket content is shipped not packaged one week in advance from the fulfillment center to the store near the customer, then packaged and fulfilled from the store, as a form of basket assemble-to-order. Finally, if a basket is not picked by the two-step approach, i.e. the prediction is zero, then traditional direct delivery of the basket from the fulfillment center is triggered. The benefit of this strategy is that first, only the most probable baskets are shipped in advance to the store which reduces overstock and space utilization, second, it achieves the economics of scale by packaging in the fulfillment center and reducing the workload on store staff. Third it reduces the delivery time from

the traditional 4 days to possibly 2-24 hours, especially for BS2. Finally, it reduces the cost of transportation by consolidating several anticipated baskets rather than express or traditional delivery from the fulfillment center to the customer.

## 3.2 Two-step advancement shipping approach

### 3.2.1 Basket Prediction

Starting from historical sales, we rely explicitly on graph theory in order to model a graph-like structure that captures the attributes of bi-product links. These attributes will be used as features in state-of-the-art machine learning models to predict the future sales of the arcs. These arcs will, later on, formulate baskets. The response targets (BS2 and BS>2) are continuous variable in this research. However, this approach can also be implemented with ordered classification targets.

The proposed arc prediction step can be summarized in Figure 3.2. The steps of the proposed approach are detailed in the following subsections.

- a) Shopping baskets data (from the sales history) are used to build the network, b) the network is then considered to calculate attributes that are arc related namely three sets of attributes; arc attributes, nodes to the arc attributes, and arc probability/Bayes attributes, where every observation is an arc that appeared in historical sales of baskets. Each observation has two targets; how many times an arc will appear within a given time period in a basket of size 2 and how many times an arc will appear in an order of size more than 2.
- b) Each target is binned to ordered classes for under-sampling purposes only namely; no event which is an arc that has a response variable of zero, one event which is an arc that has a response variable of one, two events, and more than two events. The resulting new classes are likely to be imbalanced as shown from our empirical investigation with more frequencies on the "no event" and decrease as the number of events increases. Thus, a supplementary procedure of under-sampling followed by duplication of low frequencies is recommended for the training and tuning of machine learning models only. No such operations are necessary for testing or prediction.
- c) After under-sampling and removal of the order binning classes, The attributes are used as regressors along with historical basket sales data within a machine learning method. Note that the two targets to predict; basket of size 2 (BS2) and basket of size more than 2 (BS>2) are numerical. The ordered class binning is only for under-sampling and duplication purposes but is not included in model training.

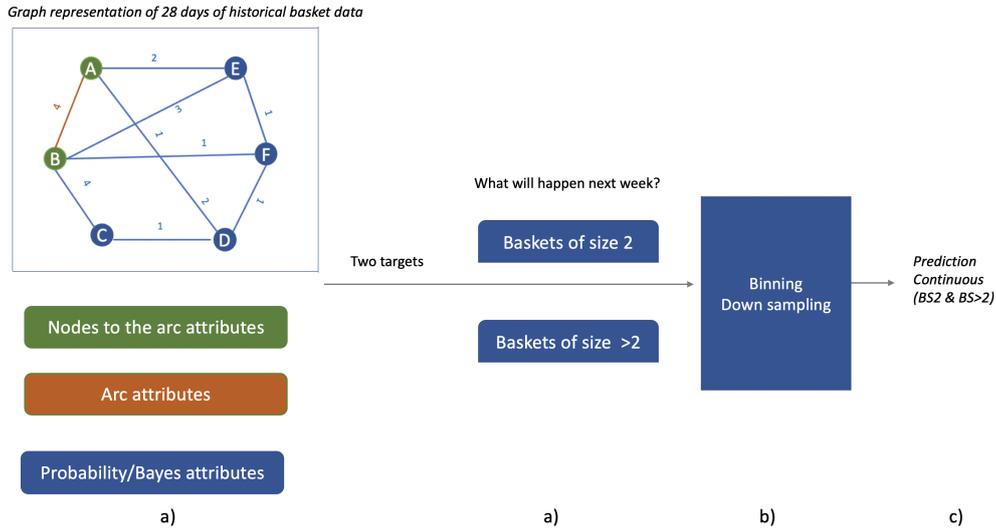


Figure 3.2: Illustration of the prediction approach

### 3.2.1.1 Network characterization and attributes extraction

Using what we have concluded in Chapter 2, we further utilize the graph attributes by focusing on the arcs and the nodes of the network to shape the baskets ordered by customers and not only the forecasting of individual products (Gross and Yellen 2005, Kim et al. 2015, Dooley et al. 2019). In this context, we rely explicitly on graph theory in order to: 1) model a graph-like structure from historical sales that captures the characteristics of the products (node) and the bi-product purchases within baskets (arc) knowing that every arc can appear in a basket where only two items are bought (BS2) or a basket where more than two products are bought. If a customer for example buys a perfume product with two units of identical mascara, it is considered as a basket of more than size 2 (BS>2) 2) Explore a graph-learning from the basket data using an original set of attributes. The arc linking the two products is associated with the number of baskets in which the two products are sold together. This number is referred to as the frequency of this connection. We note that in our graph representation, we assume that the relational attributes of products are the only associations and there is no causal relationship; thus, the network is undirected. Figure 3.3 illustrates a typical basket data representation in a graph. The data is extracted as follows, each basket data is transformed into every two possible pairs, this process is done on all baskets for a period of twenty-eight consecutive days (the graph horizon), and attributes are then calculated on those pairs Figure 3.3 represents a subset of a graph of twenty-eight days period graph. With this in mind, we identify three sets of attributes that use the graph and historical basket data in the proposed forecasting approach; nodes to the arc attributes, arc attributes, and probability/Bayes attributes. Noting that we will focus on arc (A-B) to illustrate these attributes

### 3.2.1.2 Nodes to the arc

For the nodes related to every arc, we identify six attributes of the product considered in the proposed forecasting approach: (1) *Degree* of the product, (2) *Strength* of the product, refer to their definitions in Chapter 2 (3) *Weighted transitivity* of the node is the probability that the adjacent nodes of a node are connected. This is sometimes also called the clustering coefficient Barrat et al. (2004), (4) *closeness centrality* which measures how many steps are required to access every other node from a given node Freeman (1978), (5) *triangles* which counts how many triangles the node is a part of and

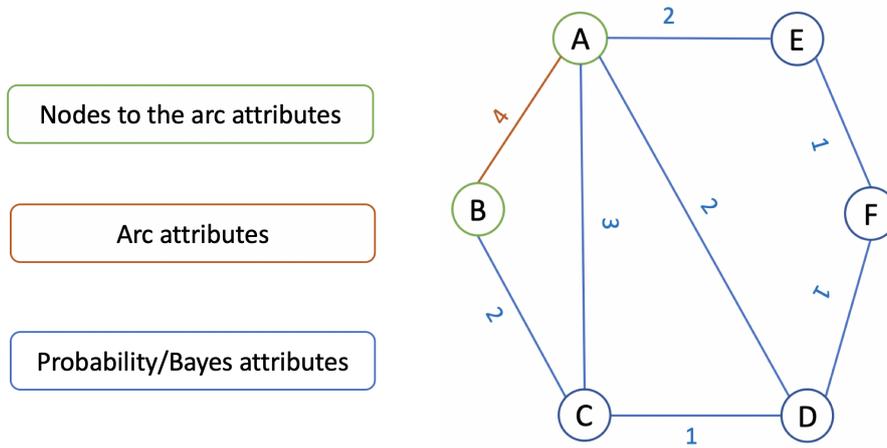


Figure 3.3: Basket data representation in a weighted graph

(5) *weighted betweenness* which is how many weighted shortest path passed through the node, the weights being the frequency of the arc Brandes (2001). So in total, every arc will have twelve node-related attributes. In Figure 3.4, arc(A-B) has two related products,

Product	Degree	Strength	Weighted transitivity	Closeness	Count triangles	Betweenness
A (node 1)	4	11	0.36	0.83	2	1
B (node 2)	2	6	1.00	0.56	1	0
E	2	3	0.00	0.63	0	0.5
F	2	2	0.00	0.56	0	2
D	3	4	0.38	0.71	1	4
C	3	8	0.69	0.71	2	2

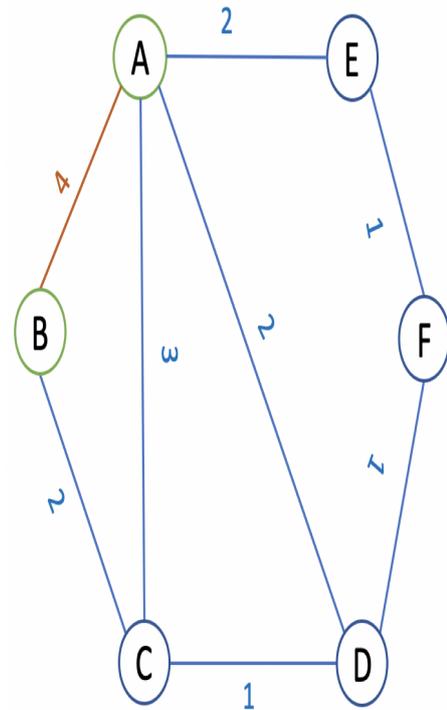


Figure 3.4: Nodes to the arc attributes

product A and product B. Product A has connections with four connections (B, C, D, E) and the sum of frequencies of those connections is eleven. Thus, the degree and strength of product A and B are four and eleven respectively. While product A has 4 adjacent products, it only has two closed triangles; (A, B, C) and (A, C, D). Applying weighted transitivity on product A will result in 0.36 which is the average weight of every connected arc where product A is centered, conditioned that it's part of a closed triangle divided by all the triplets where product A is centered. Product A also exhibits the highest closeness centrality as it has the highest degree of connection with other products. For product A, the attribute count triangles count how many closed triads of products the product

is a part of. As noted before, product A is a part of two triangles of products. Finally, betweenness is the sum of the weighted distances the shortest paths pass through product A. The six product attributes of Product A (node 1) and Product B (node 2) together are the nodes to the arc attributes set to arc (A-B)

### 3.2.1.3 Arc attributes

In the arc attributes set, attributes explicitly related to the arc are proposed to be extracted; 1) *frequency* of the arc, which is the number of times an arc appeared in all baskets, 2) *arc betweenness* which is how many weighted shortest path are passing through the arc, 3) *frequency count for basket of size n* where n is a subset of all basket sizes (number of products inside basket) that the arc appeared in from historical basket data. Note that in Figure 3.5, as we only took a subset of a graph from our empirical analysis *frequency for basket size n* have only three cases; basket size 2, basket size 3, and basket size 4, while in our empirical analysis, panel data have up to 11 cases from basket of size 2 to basket of size 12. And finally, 4) *total baskets size >2* which is the sum of  $n_3$  to  $n_x$  where x is the highest basket size for this arc. This measure could be considered as a lag for the target prediction  $BS > 2$ . And basket size 2 attribute could be considered as the lag for the prediction target of BS2 In Figure 3.5, it is noted that some arcs appear in more baskets of size 2 such as Arc (A-D) while other arcs appear more in high-size baskets as arc (A-E).

Arc	Frequency	Edge betweenness	basket size 2	Basket size 3	Basket size 4	Total baskets >2
A-B	4	2	2	0	2	2
B-C	2	3	0	0	2	2
C-D	1	5.5	0	1	0	1
D-F	1	5.5	1	0	0	0
F-E	1	1	1	0	0	0
A-E	2	2.5	0	0	2	2
A-D	2	2	2	0	0	0
A-C	3	0.5	1	1	1	2

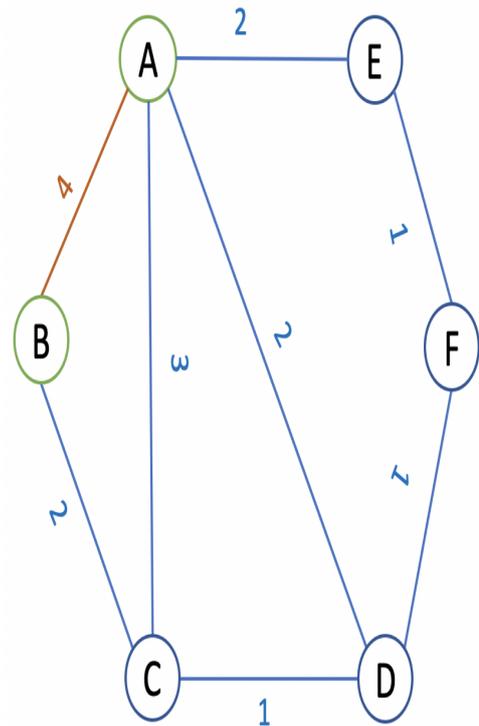


Figure 3.5: Arc attributes

### 3.2.1.4 Probability/Bayes attributes

The last set of attributes extracts information from the nodes to the arc and the arc itself with relation to graph(t), recall that graph (t) is a graph that represents historical basket data in the last 28 days. We rely on conditional probability and Bayes rule to infer the chance that a specific arc from the graph will be likely purchased again in a future basket

given prior knowledge (Prior) about specific characteristics of the graph and basket data and then update the probability with new information (posterior) Reich and Ghosh (2019). Continuing on the previous sub-graph, an extraction of probability/Bayes attributes of arc (A-B) is demonstrated in Figure 3.6. First, *node 1 proportion* is the proportion of arc (A-B) frequency to the total frequency that node 1 (A) participates in, *total basket size 2 node 1* is the total orders which node 1 participates in all of the basket size 2 orders in graph(t), *total basket size >2 node 1* is the total orders which node 1 participates in all of basket size more than 2 orders in graph(t), *Bayes node 1* is calculated as the probability of basket size 2 of arc (A-B) to the total basket size 2 that node 1 participates in multiplied by the likelihood of basket size 2 for node 1 divided by node 1 probability. And *probability basket size 2 given node 1 total baskets* which is the conditional probability that a new basket size 2 from the arc (A-B) that node 1 participates in given that the total baskets that node 1 participates in are eleven baskets. *Node 2 proportion, total basket size 2 node 2, total basket size >2 node 2, Bayes node 2* And *probability basket size 2 given node 2 total baskets* are the same but for the baskets that node 2 participates in for graph(t). Finally, the arc attributes which are related to the graph, *arc proportion* which is the proportion of arc (A-B) with relation to all the arcs in graph(t), *total baskets size 2* are the total baskets of size 2 in graph(t), *total baskets size >2* are the total baskets of size more than two in graph(t) and *prop arc size 2 given all baskets* which is the probability that arc (A-B) is of size 2 given the total baskets of graph(t).

Arcs	node 1 proportion	total basket size 2 node 1	total basket size >2 node 1	Bayes node 1	prop basket size 2 given node 1 total baskets	node 2 proportion	total basket size 2 node 2	total basket size >2 node 2	Bayes node 2	prop basket size 2 given node 2 total baskets	arc proportion	total basket size 2 arcs	total basket size >2 arcs	prop basket size 2 given total baskets
A-B	0.36	5	6	0.5	0.18	0.5	3	5	0.5	0.25	0.2	8	10	0.11
A-E	0.18	5	6	0	0	1	0	2	0	0	0.1	8	10	0
E-F	1	1	0	1	1	1	1	0	1	1	0.1	8	10	0.06
F-D	1	1	0	1	1	0.33	3	0	1	0.33	0.1	8	10	0.06
D-C	1	0	1	0	0	0.25	1	3	0	0	0.1	8	10	0
C-B	1	1	3	0.25	0.25	0.5	3	5	0.25	0.13	0.2	8	10	0.06
A-D	0.18	5	6	1	0.18	0.67	3	0	1	0.67	0.1	8	10	0.11
A-C	0.27	5	6	0.33	0.09	0.75	1	3	0.33	0.25	0.2	8	10	0.06

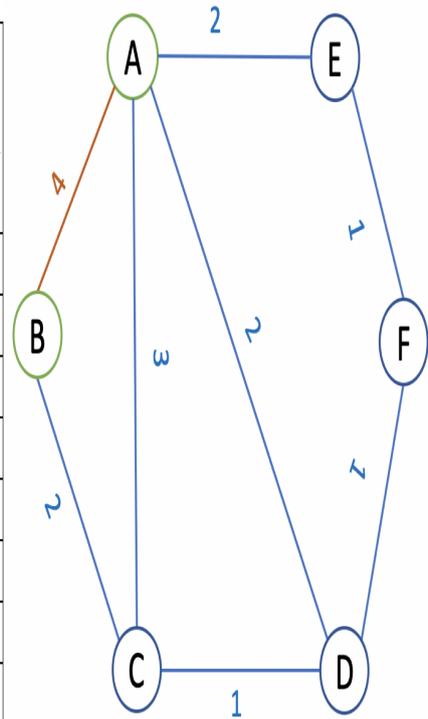


Figure 3.6: Probability/Bayes attributes

From a graph-theoretic perspective, we can conceptualize the network of connections between products in baskets as follows. Let  $t \in T$  be the set of historical sales periods (days) and  $b \in B$  be the set of sold baskets, where  $B^t$  is the subset of baskets sold in period  $t \in T$ . Let  $p \in P$  be the set of products, where  $P_b$  is the subset of products included in sold basket  $b$ . Accordingly, the graph  $G^t$  is defined for a given period  $t$ , and is composed of a set of subgraphs  $G_b^t$  per basket  $b \in B^t$ .

The network is produced on a daily basis with a rolling time, and the attributes are determined for each arc and for each forecasting period  $t$  (using basket data on periods  $t - 1, t - 2, t - 3, \dots$ ). Let  $h(t)$  be the historical horizon used to build the network (i.e. 7, 14, 21, or 28 days prior to period  $t$ ) related to the estimation at period  $t$  (i.e. period where the forecast is made), composed by  $[t - 1, \dots, t - h]$ , for the collection of baskets  $B^{h(t)}$ . Consequently,  $G^{h(t)}$  defines the network graph associated with the set of periods in  $h(t)$  such that  $G^{h(t)} = \cup_{[t-1, \dots, t-h]} G^t$ . For a given period  $t$ , based on the historical horizon  $h(t)$  data.

### 3.2.1.5 Machine learning methods for basket prediction

Machine learning methods have become mainstream in forecasting operations and have been shown to outperform statistical models. This has been demonstrated in recent forecasting competitions. In addition, they provide the advantage of adding exogenous variables [Petropoulos et al. \(2022\)](#). We apply several machine learning models (AutoML) to the training data. The models that provide the best metrics accuracy after 10-fold cross-validation are then selected for further simulations and costing analysis. The graph and arc modeling are made using IGRAPH package in R and the search space for state-of-the-art machine learning models as well as model accuracy results is made using PYCARET library in Python which is a low code library that provides state-of-the-art machine learning and feature engineering capabilities.

The nature of this new approach of arc modeling to forecast produces a vast amount of data as input observations/records to the model, as an example from our empirical investigation. An online or a physical store that processes around 600 orders per day, for a graph of 28 consecutive days, the arcs resulting observations from this graph may reach 60000 different arcs. As a result, to properly train the machine learning models, an intermediate process of binning the two targets and then under-sampling the zero sales instances is recommended to present the models with sufficient samples of different target frequencies and also reduce the training load/time when doing cross-validation. However, this intermediate process of binning and downsampling does not take part in arc modeling or prediction. It is only for providing a moderate informative sample to the prediction model

While the graph is produced on a sequential daily basis to forecast one week ahead of baskets and there is a time dependency on the sales volume/frequency of every arc, this approach is not considered a traditional time series where there is an index time stamp and a key arc that appears over time, this is due to the fact that most of the arcs do not exhibit a long term repetitive nature. However, the time stamp is sequential in nature and every day has the attributes for the sets of arcs that appeared in the preceding 28 days. This characteristic does not allow for relative error measures traditionally used for forecasting such as MASE, RMSSE and SMAPE. Those metrics need the availability of the arc key on every week of prediction. In addition, we also investigate adding the predictions of several models as a form of ensemble along with the arc features to reduce prediction errors as investigated by [Tyrallis et al. \(2021\)](#). While there are many forms of model ensemble, we only use the weighted average of best-performing models in some of the simulation runs as later illustrated in empirical investigation.

We present below the significant machine learning models tested via AutoML on the proposed arc prediction approach, the models that showed low forecasting error and a good inventory performance while the other AutoML results are shown in Appendix 2:

Random forest is a tree-based ensemble method, it consists of a large number of weak decision tree learners which are grown in parallel. Each tree takes a subset of features

and training samples randomly. As a final step, the weak decision learners are aggregated together to reduce the variance and the bias of the model [Rodriguez-Galiano et al. \(2015\)](#).

Extra trees, extremely randomized trees are an extension of the random forest. It employs the same logic of random forest and chooses a random set of features to train every tree, however, it chooses the best available feature and threshold for every split in the tree [Ahmad et al. \(2018\)](#).

Catboost is one of the flexible tree models that often produce high accuracy results without the need for tuning, Like Random forest, Catboost combines several weak predictors but unlike random forest, they are not random and are fitted sequentially using gradient boosting [Lainder and Wolfinger \(2022\)](#)

Linear regression is considered the standard machine learning model where the independent variables are fitted linearly against the target variable, while Linear Regression is not generally the best-performing machine learning model, it is one of the fast and intuitive models to use that works well with highly correlated data [Ciulla and D'Amico \(2019\)](#).

XGBoost or extreme gradient boosting system is a widely used machine learning algorithm and is known to outperform in both forecasting and machine learning competitions on Kaggle and also the M4 forecasting competition. Although, it is similar to the bagged tree and random forest in the sense that a group of weak classifiers builds a stronger classifier. It is a sequential iterative approach where each predictor reduces the prior predictor's error. The wrong predictors have higher weights while the lighter predictors have lower weights and they are the ones used after training. The trees are built in parallel using modern computation systems [Chen and Guestrin \(2016\)](#).

LightGBM is another popular gradient boost decision tree implementation. Unlike XGBoost, LightGBM is a leaf-based tree-splitting mechanism. LightGBM focuses on reducing the training sample size using gradient-based one-side sampling (GOSS) which puts more importance on samples with the highest residual errors and Exclusive feature bundling (EFB) which reduces the number of effective features that are mostly mutually exclusive. Together, they provide a more efficient gradient boosting algorithm on big-sized data with high feature dimensional [Ke et al. \(2017\)](#).

Finally, as BS2 and BS>2 are continuous target variables and because in this modeling approach, the majority of arcs do not appear in the forecast week ahead as later presented in the empirical investigation, we propose a conservation threshold that excludes any target observation that is below it. And rounds any target threshold that is above it. This increases the certainty of the prediction and reduces the wasteful advancement of baskets. As the conservation threshold increases, the number of baskets advanced decreases, and the certainty increases. And evidently, some models will be better than others in selecting which baskets to advance.

After discussions with supply chain leaders from the retail organization, we suggest a complimentary judgmental approach to calibrate the results of the ML models. As an example, if the model predicts that for a specific arc, 4 sales for a BS2 will occur. A decision maker can be pessimistic and advance 2-3 packaged baskets, or optimistic and advance four baskets or more. Such a decision could be based on basket value and the cost of advancing. We acknowledge that such judgmental interpretations are useful to calibrate the basket forecasting model over time [Alroomi et al. \(2022\)](#). This is also why a conservation threshold is proposed to calibrate the outcomes of the models when needed. Another dimension that is subjective and based on discussion with supply chain leaders is the forecasting and graph horizon. As previously stated, the forecasting horizon for every graph of 28 days is one week. We chose the 28-day graph/7-day forecast for two reasons; first, a longer forecasting horizon gives decision-makers more time to advance baskets to stores in anticipation of demand which will be difficult for a one-day duration. Second,

both our preliminary findings and previous research suggest that temporal aggregation provides more accuracy [Kourentzes et al. \(2017\)](#), [Babai et al. \(2012\)](#), our investigation shows that this case holds true with the proposed two-step approach.

### 3.2.2 Prediction-based basket retrieval algorithm

Step 2 of our approach is deciding which arcs to advance for BS2 and BS>2 from step 1 and applying the policy of the T0-BE strategy. An important question is how the predictions for BS>2 could be advanced to a near store using the arcs. We do not have this problem for a basket of size 2 as a basket of size 2 has only one arc. While a basket of size 5 has 9 different arcs and cannot be sent to stores directly in its arc form as every node in a basket of size 5 is duplicated 4 times in its arc form, hence we propose the following basket retrieval algorithm that has the two following advantages :

- Remove duplication caused by arc modeling for BS>2. For example, if a basket has three items A, B, and C, this will consist of arc (A-B),(A-C), (B-C) which means that every node is repeated twice in all the arcs.
- Retrieve the products from arcs to baskets of different sizes instead of arcs.

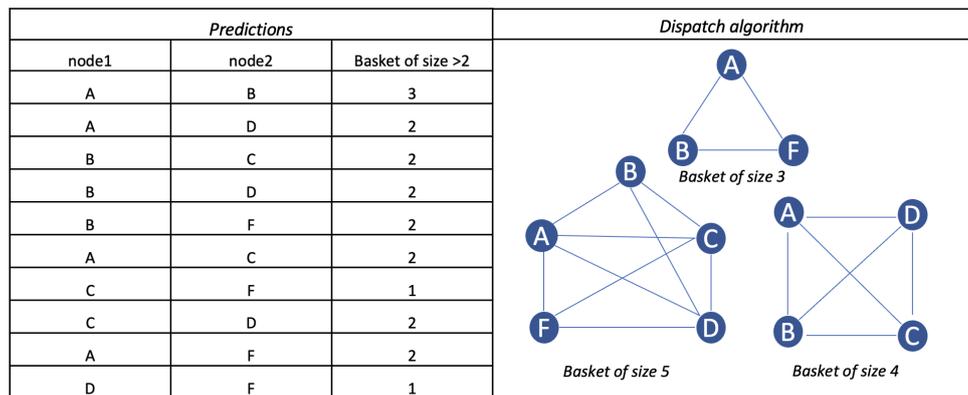


Figure 3.7: Basket data representation in a weighted graph/Arc attributes

In Figure 3.7, an illustration of the proposed basket retrieval algorithm will translate the arcs back into baskets, as noted, every arc in every basket has a frequency of one. As a result, with every arc that is dispatched to a basket, one of the arc predictions is reduced by one. This operation is iterated until all of the basket predictions of size >2 become zero. The algorithm logic is that in order to make a basket, all the products of the basket must be connected to each other. In other words, any basket should be a clique in its graph form [Eppstein et al. \(2010\)](#). Our assumption is that the algorithm always chooses a basket with the maximal clique first. It is also noted that the algorithm is not exact. As there can be different basket shapes from the same arc set. We provide an example of two different baskets in section 3.2.3. This is why we recommend advancing baskets of size >2 not packaged for potential minimal change at the store afterward. The algorithm starts with initiating an arc counter a basket counter and an empty basket list where the products in the cliques are stored. Effectively, we are reducing the arc predictions by distributing them into baskets until the remaining arc predictions become zero. The starting point would be products with the highest connectivity. i.e highest strength, the maximal clique this product is a part of will constitute the first basket. The algorithm is run week by week where the output is the BS>2 forecasts of next week. The algorithm can be presented as follows :

**Algorithm 1:** Prediction-based Basket Retrieval Algorithm

**Input:**  $F_{ij}^{arcs} = [f_{11}^{arc}, f_{12}^{arc}, f_{13}^{arc}, \dots, f_{nj}^{arc}]$ ;  
 $P_i = [p_1, p_2, p_3, \dots, n]$ ;  
 $SP_i = [sp_1, sp_2, sp_3, sp_4, \dots, n]$  at time  $t$   
**Output:**  $List_{Basket_i} = [f_1^{basket}, f_2^{basket}, f_3^{basket}, \dots, f_n^{basket}]$  at time  $t$   
 $C_{f_{ij}^{arc}} \leftarrow 0$ ;  $C_{B_i} \leftarrow 0$   
 $List_{Basket_i} \leftarrow []$  Arcs  $\leftarrow \sum \sum F_{ij}^{arcs}$   
 Arrange  $P_i$  by decreasing  $SP_i$   
**While** Arcs  $> C_{f_{ij}^{arc}}$ :  
   **for**  $p_i$  in  $P_i$ :  
      $F_{ij}^{arcs} \leftarrow \text{Remove } f_{ij}^{arc} = 0$   
      $Q \leftarrow \text{Filter } f_{ij}^{arc}$  **if**  $ij$  **in**  $p_i$   
      $M \leftarrow \text{Get Maximal clique from } Q$  where  $Q$  has  $p_i$   
      $C_{B_i} \leftarrow C_{B_i} + 1$   
     Add distinct  $p_i$  in  $M$  to  $List_{Basket_i}[C_{B_i}]$   
      $C_{f_{ij}^{arc}} \leftarrow C_{f_{ij}^{arc}} + \sum f_{ij}^{arc}$  in  $M$   
      $F_{ij}^{arcs} \leftarrow F_{ij}^{arcs} - f_{ij}^{arc}$  in  $M$  s.t.  $f_{ij}^{arc} \geq 0$

### 3.2.3 Retrieval algorithm variations

We provide an example where the dispatch algorithm does not provide the exact predicted baskets but is close to the optimal basket shapes. Recall that every order is a graph clique and the algorithm rotates over every node and gathers the cliques which every node is a part of forming the baskets and eliminating used arcs in the process. However, the solution might change based on the starting node, this is why there might be changes in some of the output baskets. In Figure 3.8, we give an example of this variation. The table represents the predictions coming from the baskets of more than size 2 model. Per the algorithm, there are 4 nodes with the same highest strength that are eligible to be a starting point, nodes A, B, C, and D. Accordingly the algorithm starts with node A forming the basket cliques around it. Based on that node A has four adjacent nodes, nodes B, D, C, and F. And since B, D, C, and F are all connected to each other and to A, then together they are the maximal clique available for node A and the first basket is formed, basket A, B, C, D, F. Note that the first basket from the algorithm is also available in the actual baskets. Afterwards, the consumed arcs are eliminated and the next maximal cliques for node A is also A, B, C, D, and F as each of the corresponding arcs to these nodes has a prediction of more than one so they are formed again. Note that the second basket formed by the algorithm is not in the actual baskets. The reason is that in the actual baskets, Basket A, B, C, D are forming a clique and F is connected to A and B in basket A, B, F and C, D in basket E, C, D, F which essentially from the algorithm logic of maximal cliques, the second basket that should be formed is A, B, C, D, F. However, the algorithm removes accurately all the duplicates of nodes caused by the arcs even if the shape of the baskets is different.

## 3.3 Performance Evaluation- Fulfillment and Costing

The proposed approach is evaluated for eight consecutive weeks which is also the out-of-sample panel that is indicated later in the empirical investigation. As a result, inventory cost is calculated week by week while salvage, penalty costs, and unpackaging costs are applied at the end of the eighth week. Knowing that a basket that is advanced and not sold as expected in the following week has the potential to be sold in the remaining weeks

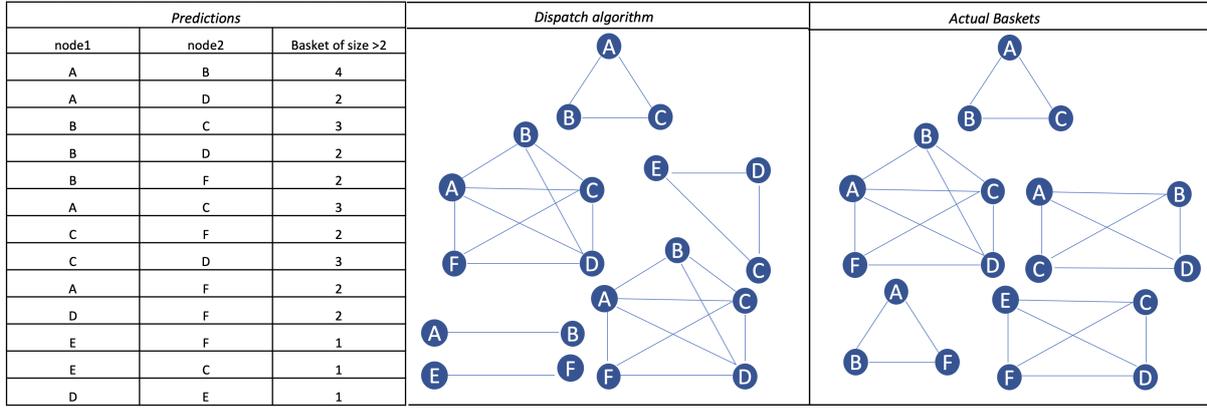


Figure 3.8: Retrieval algorithm variations

of the simulation.

The strategy works as follows; as the forecast is a continuous variable it is converted to a discrete variable based on a conservation threshold, for example, if the threshold is 3.5 and the prediction is 3.51, then the prediction is converted to 4 else if the prediction is 3.49 it is zero. The conservation threshold value is sequentially tested against inventory and costing simulation metrics using a grid of 500 values from 0 to 4. As the threshold increases, a model becomes more conservative and selects fewer arcs to be advanced.

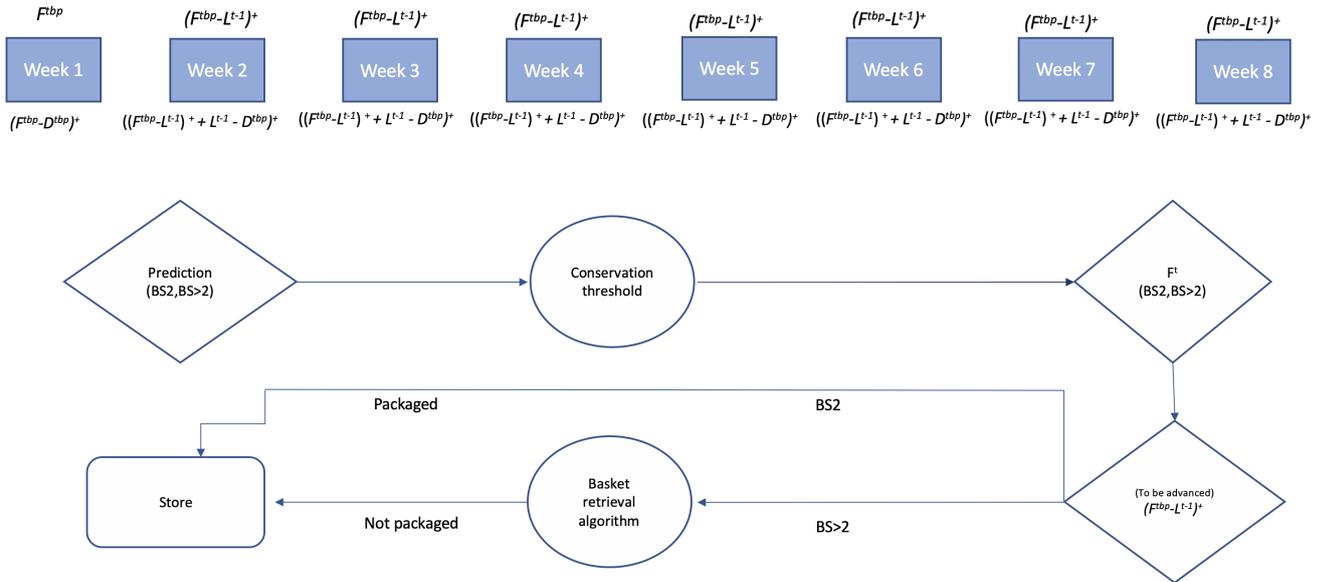


Figure 3.9: Multi-period 8-week simulation

Figure 3.9 illustrates how the inventory simulation procedure works. The inventory simulation of BS2 is carried out separately from BS>2. In the first week, threshold calibrated forecasts are advanced to store, and then if there are basket leftovers from a particular basket that was not sold during the first week, the forecast of the following week will be offset with leftovers from the previous week. Consequently, some of the baskets of the following weeks are served from leftovers. This procedure is carried out for the entire eight-week duration of the simulation. For every week a unit from BS2 or BS>2 is leftover or unsold at the end of the week, an inventory holding cost is incurred and it is more expensive than the main fulfillment center.

For BS2, if a basket is advanced and not sold- while it is not always the case, we assume

that the basket will be sold in later weeks. If it is not sold in a later week, the packaging is removed and the product is salvaged at the end of the eight weeks. This is because the basket is already packaged and its options are constrained. While for BS>2, a particular basket that is not sold is not lost sales due to the flexibility of sending the basket in a tote not packaged, this improves both the fulfillment rate and customer satisfaction with BS>2 assemble-to-order strategy at the store. However, even though packaging is more expensive in-store than in the fulfillment center, having the right basket in a tote would reduce the picking time by store staff. Accordingly, the fulfillment rate from BS>2 is the result of the items advanced being sold regardless of the shape and size of baskets as long as the basket content is within advanced items. Every week of BS>2 forecast is calibrated with conservation threshold and then baskets are retrieved from arcs forecast through basket retrieval algorithm. Note that inventory remaining at the end of every week  $L^t$  is  $((F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp} - D^{tbp})^+$  where  $(F^{tbp} - L^{(t-1)bp})^+$  is the advanced quantity of baskets after deducting leftovers and  $L^{(t-1)bp}$  is the leftovers of the previous week.

To be able to compare the viability of our approach, there are two main criteria to consider, cost and fulfillment KPIs. For fulfillment, the below KPIs are measured against each model and also compared to determine which model is better for BS2 and BS>2, noting that the KPIs are measured against the grid of conservation thresholds :

- *fulfillment rate* the percentage of baskets fulfilled from baskets advanced. Consequently, the leftovers rate is 1- (quantity fulfilled/ quantity advanced)
- *elimination rate* the percentage of baskets that appeared in historical basket data and correctly predicted that it will not appear in the future week ahead.
- *percentage not advanced* the percentage of baskets that are ordered online but not picked by the model.

For Costing, the models are compared against a baseline which is fulfilling online orders from the main fulfillment center. We do not assume there is an order lost if the content of the basket is not advanced. It is assumed that it will arrive within 3-4 days to the customer as a freight forwarded package through a third-party logistics provider and is not a lost order. For example, in [Ren et al. \(2023\)](#), they assumed that there would be emergency shipping in case a basket is not available at a pickup point near the customer.

One important question is how the inventory performance assessment can be measured for the proposed omnichannel strategy especially since traditional inventory policies are not compatible with the basket advancements approach. In addition, to compare between the decision to advance and not to advance is subject to multiple factors with regard to demand satisfaction, customer satisfaction, and operational cost. In Table 3.1 we present the implications of the baseline strategy which is shipping exclusively from the fulfillment center directly to the customer against advancements of BS2 and BS>2 [Chopra \(2018\)](#). Lead time is within desired targets in the proposed advancement strategies as the baskets are near the customer from where the basket is ordered. For the baseline, the shipping cost is high as the shipments are sent in the form of single shipments while in the proposed strategies shipments are consolidated as traditional physical shipments. As for packaging cost, both BS2 and baseline are packaged in the fulfillment center reducing packaging cost and workload on store staff, while BS>2 packaging cost is high as it is made in-store as a form of assemble-to-order for basket content. However, BS>2 enjoys the highest level of flexibility and customer satisfaction as the advanced products can be assembled in many basket sizes in case of leftovers. Note that in BS>2, the bi-product rule is not strict or constrained by packaging which gives the opportunity to sell different other

combinations as long as items are available from baskets advanced or leftovers. The selling price and cost per unit are available and extracted from the data set used in this work. All logistics cost such as packaging cost in the fulfillment center, packaging cost in the store, shipping single baskets, shipping consolidated baskets, salvage, and package removal cost are calculated on average per unit basis based on costs/flows of this region at the year of the dataset. Upon discussion with supply chain leaders in the company, what is the most likely scenario for a basket that is advanced before the customer makes an order and not sold at the week of the forecast? They advised that a basket that is not sold would be un-packaged, displayed in the store, sold as BM, and go through the normal markdown cycle if necessary. In most cases, the content of the basket would be on 20 % promotional discount on the selling price during promotional days and events. We also conclude that proceeding with the basket advancement options would lead to a 4 % increase in sales on average due to the lead time reduction of 1 to 3 days on average [Fisher et al. \(2019\)](#). As a result, to be able to capture this revenue increase in the model, we assume there is a 4 % increase in revenue as a premium paid by the customer for any basket fulfilled through anticipatory advancement.

	<b>Strategies</b>		
	Not Advanced (Baseline)	Advanced BS2	Advanced BS>2
Lead time	2-4 days	2 hours to 24 hours	2 hours to 24 hours
Shipping Cost/unit	High (1.63 Euro )	low (0.14 Euro )	low (0.14 Euro )
Packaging location	Fulfillment centre	Fulfillment centre	In store
Packaging cost/unit	Low (0.3 Euro )	Low (0.3 Euro )	High(1.3 Euro )
Unpackaging cost/unit	Not applicable	0.7 Euro	Not applicable
Inventory holding cost per week/unit	1.2 Euro	2 Euro	2 Euro
Salvage	Not applicable	80 % of selling price	80 % of selling price
Customer satisfaction	Low	High	High
Percentage of Increase in sales	0	(4 % extra Revenue )	(4% extra Revenue )

Table 3.1: Comparison between No advancement(baseline) and baskets advancement

Table 3.2 below shows the notation used for the cost model for the baseline scenario, BS2 and BS>2.

Table 3.2: Notation

Sets	
$\mathcal{P}$	Set of products indexed by $p$
$\mathcal{T}$	Set of weeks periods indexed by $t$ with $T^{t+1}$ being the week after simulation weeks
$\mathcal{B}$	Set of baskets indexed by $b$
Parameters	
$C^{pb}$	Unit cost of product $p$ in basket $b$
$C_{ps}^{pb}/C_{pf}^{pb}$	Unit packing cost of product $p$ in-store/ fulfillment centre
$C_{hs}^{tpb}/C_{hf}^{tpb}$	Unit holding cost of product $p$ in-store/ fulfillment centre in advanced basket $b$ at week $t$
$C_{ts}^{tpb}$	Unit transportation cost of product $p$ in basket $b$ at time $t$ directly to customer
$C_{tc}^{tpb}$	Unit transportation cost of product $p$ in basket $b$ at time $t$ from fulfillment centre consolidated to store
$C_{us}^{tpb}$	unpacking cost of product $p$ in basket $b$ at $T^{t+1}$ at store
$R_c^{tpb}$	Retail price of product $p$ in basket $b$
$S^{pb}$	Salvage value from product $p$ in basket $b$ at $T^{t+1}$
$D^{tb}$	Online demand at time $t$ for basket $b$
$L^{tbp}$	Leftovers of basket $b$ at end of week $t$
$L^{(T+1)bp}$	Leftovers of basket $b$ at end of week $t$
$B^{bp}$	Unit penalty cost of product $p$ in advanced basket $b$
$F_{z'}^{tb}$	forecast of basket $b$ at time $t$
$\pi$	profit function for BS2,BS>2 and baseline.

For BS2, the equation can be presented in equation 3.1:

$$\pi(BS2) = \sum_{bp} \sum_t [1.04R^{pb} \min(D_{z'}^{tb}, (F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp}) - \quad (3.1)$$

$$(F^{tbp} - L^{(t-1)bp})^+ C^{pb} - \quad (3.2)$$

$$(F^{tbp} - L^{(t-1)bp})^+ (C_{tc}^{tpb} + C_{pf}^{pb}) - \quad (3.3)$$

$$((F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp} - D^{tbp})^+ C_{hs}^{tpb} - \quad (3.4)$$

$$L^{(T+1)bp} (B^{bp} + C_{us}^{tpb}) + L^{(T+1)bp} S^{pb} \quad (3.5)$$

while for BS>2, there are consolidated shipments, packaging in store after the customer makes the order, and no unpacking cost, it can be presented in equation 3.6 as follows:

$$\pi(BS > 2) = \sum_{bp} \sum_t [1.04R^{pb} \min(D_{z'}^{tb}, (F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp}) - \quad (3.6)$$

$$(F^{tbp} - L^{(t-1)bp})^+ C^{pb} - \quad (3.7)$$

$$(F^{tbp} - L^{(t-1)bp})^+ (C_{tc}^{tpb} + C_{ps}^{pb}) - \quad (3.8)$$

$$((F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp} - D^{tbp})^+ C_{hs}^{tpb} - \quad (3.9)$$

$$L^{(T+1)bp} B^{bp} + L^{(T+1)bp} S^{pb} \quad (3.10)$$

Finally, in the baseline scenario we assume the profit of baskets advanced for BS2 and BS2 had they remained in the fulfillment center and not advanced, they would incur the holding cost of the fulfillment center, transport cost for shipping directly to the customer, and packaging cost of the fulfillment center. Note that there is no penalty, unpacking cost

or salvage or remaining baskets as we assume they can be sold in other regions and have a higher opportunity to be sold in other regions at full price :

$$\pi(Baseline) = \sum_{bp} \sum_t [R^{pb} \min(D_{z'}^{tb}, (F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp}) - \quad (3.11)$$

$$(F^{tbp} - L^{(t-1)bp})^+ C^{pb} - \quad (3.12)$$

$$(F^{tbp} - L^{(t-1)bp})^+ (C_{ts}^{tpb} + C_{pf}^{pb}) - \quad (3.13)$$

$$((F^{tbp} - L^{(t-1)bp})^+ + L^{(t-1)bp} - D^{tbp})^+ C_{hf}^{tpb} ] + \quad (3.14)$$

$$L^{(T+1)bp} R^{pb} \quad (3.15)$$

## 3.4 Empirical investigation

### 3.4.1 Supply chain and data description

The basket data in focus is the online orders of one city in France for a major cosmetics leader. The data contains the transaction data of baskets including the date of purchase, order number, and list of items inside the order (Basket). On average, 600 baskets are purchased per day with different size variations. The data length is one year from January 1 to December 31, 2018. To understand the nature of basket sizes purchased, we analyzed the different basket sizes as shown in Figure 3.10. The results motivate our approach as baskets of size 2 are 29 % of the total baskets ordered in the city and baskets of more than size 2 are 37 % of the baskets. Furthermore, this result confirms the tendency of customers to buy more than one item in a basket as baskets of size 1 are only 35 % of the baskets. Also, we can note from the distribution of different size lengths that the average basket size is 2.55 items in a basket. This further strengthens the need to study relational features through graphical representations. For the one-year duration data set, from the first of January to 30 of October is used for initial model training (Panel A), from 1 November 31 of December ( Panel B) is used for two purposes, it is used to test the different models that are using the three sets of attributes. In addition, as every forecasting day represents a week ahead forecast, eight days are extracted from Panel B, specifically the days at the beginning of each week to represent the eight weeks of simulation from 1 November to 31 December. Panel A is under-sampled while panel B is unbalanced, not under-sampled, and not seen by any of the models before forecast tasting or inventory/costing simulation.

### 3.4.2 Arc modeling and feature analysis

Modeling baskets data as arcs results in bigger data size than the original data. As an example, the original online data of the retailer for this city is 540000 rows of tabular transactions data resulting in 180000 million rows of arc-modeled data. Recall that every day of forecast represents the arc that occurred in the previous 28 days to forecast what will happen in the next 7 days. And most of the arcs that appeared in history will not appear in the forecast horizon. From the reasons that lead to this observation possible that many of the baskets bought by customers are uncorrelated and triggered by impulsive complimentary buying, the availability of many substitutes for a customer to choose from [Omar et al. \(2022\)](#). Also, in retail, the turnaround of merchandise is higher than in other industries and shows a high rate of seasonality especially for cosmetics retailers. Finally, the items previously bought can be out of stock and cannot be bought in the forecast horizon. In Figure 3.11 we show that on an average day of forecast, 95 % of historical arcs are likely not to be bought again in the following seven days while only 5 % of the arcs will

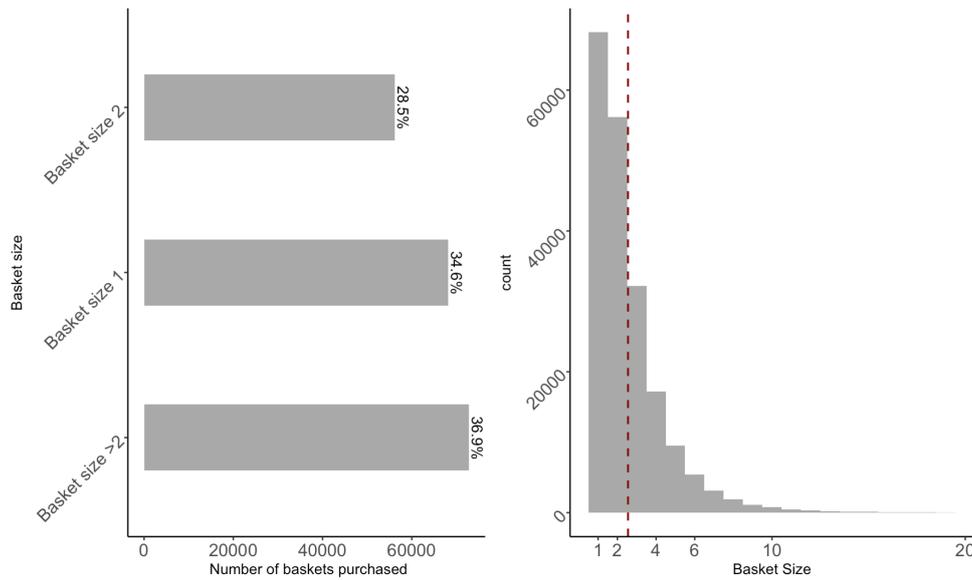


Figure 3.10: Basket data size frequencies and histogram

be bought one time or more in a basket of size 2 or a basket of size more than 2. This is why the down-sampling preprocessing step is important before applying cross-validation to the training data. In addition, the nominated forecasting model should be good at eliminating the no sales basket future occurrences as well as detecting the fewer basket sales occurrences.

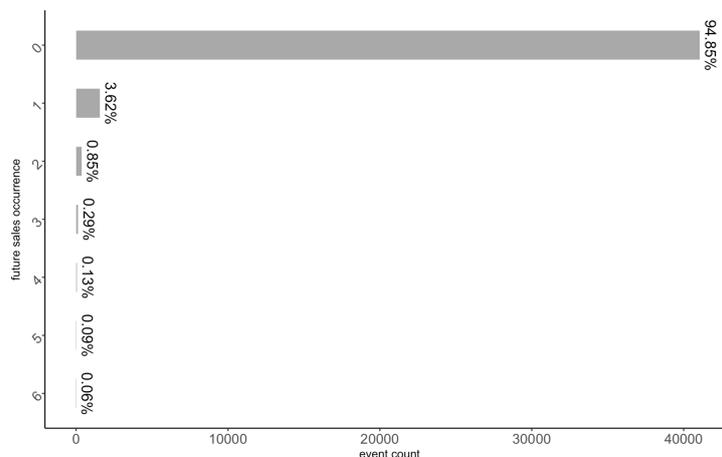


Figure 3.11: Future occurrences of historical exact baskets

Another important analysis is the correlation analysis between the proposed attributes and the two targets baskets of size 2 and baskets of size more than 2. One important preprocessing step is improving the correlation between the features and the target variables either by doing dimension reduction or adding synthetic features such as polynomials for independent variables and/or feature interactions. These steps improve the information gain and accuracy of the models. In Figure 3.12, it is noted that the target basket of size 2 has significant positive correlations with arc frequency, historical baskets of size 2, historical baskets of size 3, arc probability, and probability basket size 2 given total baskets. It has mild correlations with the other features and a notable negative correlation with node1 transitivity and node2 transitivity. While the target basket of size more than 2 has stronger correlations with historical baskets of more than size 2 compared to historical baskets of size 2, the negative probability with the transitivity of the nodes and strong

positive probability with frequency, arc probability, and probability basket size 2 given total baskets. Also, it can be noted there are implicit correlations that indirectly correlate with the targets. For instance, Bayes node 1 and Bayes node 2 have strong positive correlations with historical baskets of size 2 and strong negative correlations with other historical basket sizes even though the Bayesian features have weak correlations with the targets directly. The same applies to the negative correlations between transitivity from one side and degree, strength, and closeness, even though that degree, strength, and closeness have weak direct correlations with the targets directly. This analysis clearly shows the impact of proposed features on future basket predictions. While a number of features more than others, this will be further strengthened by applying polynomial transformation of degree 2 on the features and feature interaction as shown in the section of model analysis.

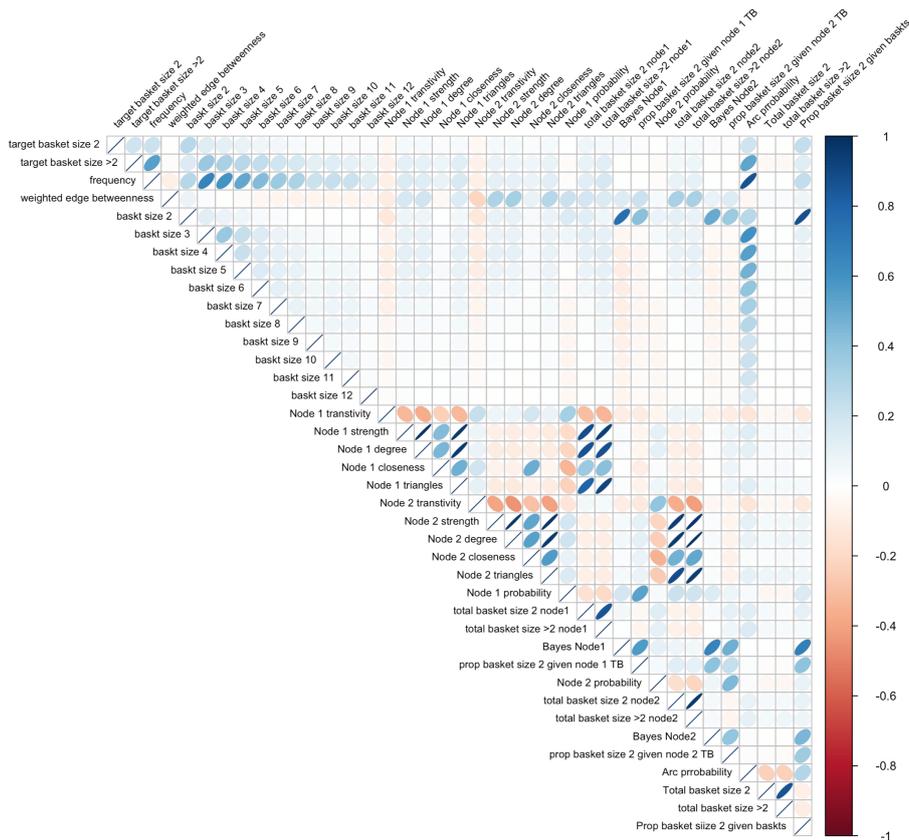


Figure 3.12: Correlation between attributes,BS2 and BS>2

### 3.4.2.1 Models predictive performance against conservation thresholds

Panel A is used for training all the models in PYCARET, other than model tuning, we have found no significant difference in doing preprocessing steps such as normalization or extraction of polynomial/ interactive features. we present all test results on Panel B from the AUTOML procedure in PYCARET which depends on the sci-kit-learn library for machine learning in Python. The results show that LightGBM achieves the lowest MAE, MSE and RMSE, and RMSLE together with XGBoost, Linear Regression, and Catboost Regressor they are vastly better than the other models. The same process is applied to baskets of size >2 where the data for the training horizon is under-sampled for the data of the training horizon. After the models are fitted on the training data, the results on unseen unbalanced future data are presented in Table 3.3. The results show that XGBoost and LightGBM are providing the best results. LightGBM is the lowest

error in terms of MAE and XGBoost achieves the lowest error in terms of MSE, RMSE, and RMSLE.

Model	BS2				BS>2			
	MAE	MSE	RMSE	RMSLE	MAE	MSE	RMSE	RMSLE
Extra Trees Regressor	0.0472	0.0166	0.129	0.096	0.3268	0.3731	0.6108	0.327
Random Forest Regressor	0.0545	0.0196	0.1399	0.1056	0.3546	0.4287	0.6547	0.3473
CatBoost Regressor	0.0162	0.00444	0.0666	0.1135	0.0882	0.0559	0.2248	0.1375
Light Gradient Boosting Machine	<b>0.0115</b>	<b>0.0040</b>	<b>0.0630</b>	<b>0.0404</b>	<b>0.0750</b>	0.0518	0.2776	0.1281
Extreme Gradient Boosting	0.0158	0.0051	0.07171	0.047	0.0876	<b>0.0398</b>	<b>0.1195</b>	<b>0.1262</b>
Linear Regression	0.0434	0.0191	0.1046	0.0760	0.2653	0.24479	0.4979	0.2604

Table 3.3: ML Models performance on Panel B

In Table 4.9, We present 10 samples for observations from Panel B, 5 samples where there are sales on arcs of both targets and 5 samples where there were no sales on arcs of both targets for the sake of comparison between target and predictions. The first 5 arcs have sales on both targets, the prediction is approaching 1 when the sales are 1. However, if this rule is applied, Then sample 5 for the target basket of size 2 does not follow the criteria as it is reaching 0. Also, it is noted when the arc sales are more than 1 as in sample 3 for example, both predictions capture this up trend in sales. The last 5 samples are the samples where there is no sales on both targets, it can be noted that all the predictions are approaching 0. This concurs with our findings that the proposed attributes are important for accurate predictions. However, as noted in Figure 3.11, there will be a lot of bi-product combinations that will have no sales occurrences in the forecasting week So elimination of non-reoccurring baskets is as equally important as detecting reoccurring baskets. With this observation, we apply the conservation threshold on each model to reduce the number of arcs that are predicted as having sales. in Figure 3.13, the effect of the conservation threshold is illustrated against arc fulfillment rate, the number of resulting arcs from the approach that is requested in the week of the forecast, the number of arcs that were requested but not picked by the models and arc leftovers when advanced and not requested by customers. For BS2, every arc is effectively a basket while for BS>2, the conservation threshold is applied before the implementation of the basket retrieval algorithm. It is shown that as the conservation threshold increases, the number of arcs recommended decreases leading to a reduction of inventory leftovers but also a reduction of arcs advanced and subsequently increase in fulfillment rate. As the conservation threshold decreases, the number of arcs advanced increases. With that note, LightGBM is performing better than other models for BS2 across the four measures while XGBOOST is significantly worse even though XGBOOST is the second best model in terms of accuracy measures for BS2. For BS>2, XGBOOST is performing much better than LightGBM even though LightGBM is the second-best model in accuracy measures for BS>2. It is also noted that simple linear regression is performing consistently well on BS2 and BS>2. Accordingly, there is no evidence that a low-error ML model based on traditional accuracy metrics would result in an effective advancement strategy. Hence, in the following section, we rely on fulfillment KPIs and profit models to determine model efficacy.

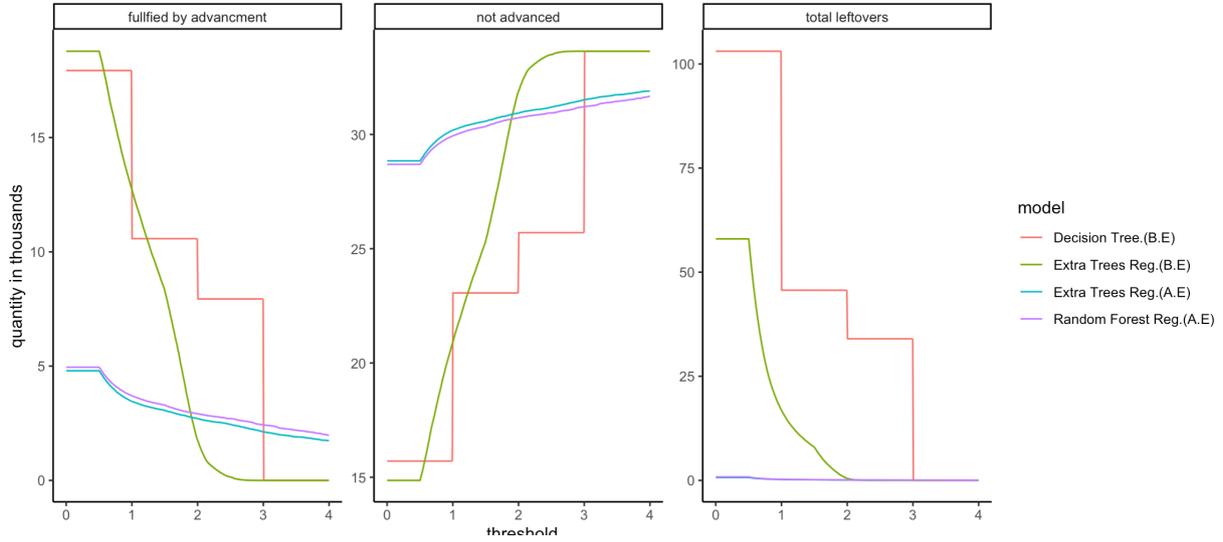


Figure 3.13: Conservation threshold effect on prediction

arc	product 1	product2	target basket size 2	prediction basket size 2	target basket size >2	prediction basket size >2
1	54995	354410	1	0.93	1	1.46
2	67221	285138	1	0.62	2	1.04
3	125586	263476	4	<b>3.27</b>	4	5.21
4	125586	420731	1	1.15	3	2.36
5	203564	321594	1	0.14	1	<b>0.62</b>
6	1	262101	0	0.02	0	0.3
7	1	292884	0	0.02	0	0.19
8	1	397873	0	0.02	0	0.12
9	1	435785	0	0.01	0	0.19
10	1	442550	0	0.01	0	0.13

Table 3.4: Forecasting samples from test set

### 3.4.2.2 Fulfillment KPIs and Costing results

As the ML models are trained and tested, the results of the simulation of 8 consecutive weeks are presented below, we only outline some of the significant models and discuss the output of these models. The simulation was made over the grid of different thresholds from 1 to 4 with 5000 runs. Note that the forecasts from the second to the eighth week are adjusted with remaining basket inventories from the previous weeks. The results in Table 3.5 show the basket strategy results for BS2. The conclusion from the table shows there are trade-offs that should be addressed by decision-makers for example the Ensemble model that uses a weighted average of three models at a conservation threshold of 2.79 achieves a 100 % fulfillment rate which means that all advanced baskets are sold through the eight weeks, yet only 17 baskets were advanced from 2997 baskets requested during the eight weeks. This model is suitable for a conservative high-basket value retail supply chain. While the LightGBM of threshold 0.1 achieves a much more fulfillment quantity of up to 29 % of baskets ordered. However with much more leftovers at the end of the eight-week period. A retailer can afford to keep this much of leftovers if the basket value is low, it is a fast-moving item or competitive advantage and customer satisfaction is the deciding criterion. A balanced model between the high fulfillment rate/low fulfilled

quantity and the low fulfillment rate/high fulfilled quantity is for example the Linear Regression model of threshold 1. Where the fulfillment rate is up to 43 % and the fulfilled quantity is 310 baskets or the Ensemble model of conservation threshold of 2.51. We can note that there is a trade-off between fulfilled quantity and the fulfillment rate. As for costing results based on the proposed model, the first six models which exhibit high advanced quantities are less in terms of profit than the baseline scenario, however, there is still realized profit despite high advanced quantities as shown in Random Forest, Catboost, and Linear Regression. This concurs with our conclusion that these models are suitable when customer satisfaction is the deciding criterion. As The models become more conservative, the BS2 strategy becomes more profitable than the baseline scenario as there are fewer leftovers, inventory holding costs, and a reduction in selling price than the first six models.

Model	threshold	Leftovers	Advanced	Fulfilled	F.R	P.NA	Baseline	BS2
Light GBM	0.1	3133	4001	868	0.22	0.71	23010.92	-17048.32
Random Forest	0.9	1101	1447	346	0.24	0.88	15742.39	707.85
Extra Trees	0.7	1161	1606	445	0.28	0.85	14610.32	-700.24
Catboost	1.1	311	434	123	0.28	0.96	5500.02	1355.8
Linear Regression	1.16	162	283	121	0.43	0.96	3548	2136.89
Linear Regression	1	181	672	310	0.46	0.9	8114.75	5171.77
Ensemble(LGBM-RF-XGB)	2.51	12	47	35	0.74	0.99	989.53	999.69
Light GBM	2.38	5	20	15	0.75	0.99	518.87	526.78
Ensemble(LGBM-RF-XGB)	2.55	7	29	22	0.76	0.99	765.4	778.8
Light GBM	2.42	5	24	19	0.79	0.99	628.53	656.15
Ensemble(LGBM-RF-XGB)	2.79	0	17	17	1	0.99	235.27	303.99

Table 3.5: Fulfillment/Costing -BS2

As for BS>2, after the application of the basket retrieval algorithm, and as the baskets are not packaged and a basket can contribute to the sales of another basket, the results are product-based, not basket-based. In general, the fulfilled baskets are much higher in BS>2 than BS2 for two main reasons; first, they have higher demand as shown in Figure 9, and second, the assemble to order in store technique has led to much more flexibility when customizing the basket for a customer's order. In Table 3.6, there are strict models that achieve a 100 % fulfillment rate, for example, CatBoost with a threshold of 2.6 results in having all of the baskets advanced effectively to be fulfilled in purchased baskets. However, the percentage of baskets advanced is 99 %. This model would work well for a decision-maker who does not want any leftovers from the basket advancement operations. Note that with all the models there is no loss incurred in costing even with the model of the highest leftovers such as XGboost of threshold 1.21, almost half of the advanced quantities are sold during the eight-week period, and the other half are sold after the season. Starting from a fulfillment rate of 91 %, the BS>2 models start to become more profitable than the baseline scenario. One consideration here for the decision maker is the inventory capacity of the store that might decide on which model to choose. We can deduce as well that simple linear regression is providing a balanced advancement approach results for BS2 and BS>2 which is in line with our conclusion in the accuracy measure section.

model	threshold	advanced	fulfilled	leftovers	F.R	P.NA	baseline	BS>2
XGBoost	1.21	164113	85108	79005	0.52	0.13	773885.81	374465.3
CatBoost	0.88	69502	52578	16924	0.76	0.46	322531.12	294753.87
Linear Regression	0.63	22306	16887	5419	0.76	0.83	72420.85	69494.36
Linear Regression	1.11	10626	9696	930	0.91	0.9	19783.23	25227.44
Linear Regression	1.68	8087	7537	550	0.93	0.92	9790.94	14033.24
Extra Tree Regrsson	2	3339	3151	188	0.94	0.97	7209.55	9066.13
Linear Regression	2.93	5874	5586	288	0.95	0.94	3172.69	6207.7
LightGBM	1.91	4150	3987	163	0.96	0.96	6022.73	8707.89
LightGBM	2.04	3117	3018	99	0.97	0.97	4183.24	6337.35
CatBoost	2.56	774	774	0	1	0.99	828.29	1416.56

Table 3.6: Fulfillment/Costing -BS2

### 3.5 Conclusion

Anticipatory shipping can result in significant cost savings, more satisfied customers, an increase in sales, and a competitive edge for retailers. Yet, it is challenging to have the perfect order available to a customer beforehand as the baskets ordered by customers enjoy high variety in terms of quantity and diversity. In addition, assortment in certain retail industries enjoys high turnovers with some of the product offerings lasting for a short period of time which presents a further challenge in anticipating the content of future baskets. In this research, we attempted to tackle this challenge by studying the associations between products separately and collectively by means of arc analysis and attribute extraction using the rich literature on graph theory. Through the nodes to the arc attributes and the arc attributes as these features concern, every two specific products and their sales in a specific historical time period features such as degree, strength, edge betweenness, and frequency of the arc. And collectively through arc probability and Bayes attributes which look at every two products in relation to the whole graph or the whole baskets ordered during a specific time horizon. This approach is flexible to adjust the graph span based on industry and what works for each assortment seasonal life span. We have shown through correlation analysis the impact of these proposed attributes on the prediction of an arc that will come in a basket of size 2 and a basket of more than size 2. However, due to the different nature of arcs in baskets of size more than 2, where a product contributes as a node for more than one arc, the integer prediction forecast has to go to an additional basket retrieval algorithm to reshape the arcs into baskets and remove redundancy because of product repetition. While a basket of size 2 has only one arc and does not need redundancy removal or basket reshaping. Also, forecasting the arcs of baskets of size 2 separately provides more certainty on the content of the basket, and further action of packaging and fulfillment could be done at the fulfillment center where economies of scale is achieved. As for baskets of size more than 2, advancing the baskets not packaged and as predicted modular bundles mitigates the uncertainty, reduces the transportation cost, and offers more flexibility in satisfying a customer demand from advanced modular baskets. The shortcomings of the approach are the amount of leftovers that could result from this strategy, however, we have shown that specific models can have up to 100 % fulfillment of baskets advanced which is a good strategy if the basket is of high value. Or retailers can choose more forthcoming models that advance More baskets with higher left-overs quantity. This is why we recommend judgmental complimentary decision-making post-model selection. The forecasting accuracy results signified the accuracy of LightGBM models as the best model for Baskets of size 2 while XGBoost provides better accuracy for baskets of size more than 2. Upon applying simulation of baskets

across the forecasting models, however, more accurate models do not correspond to better advancement strategies. For example, in baskets of size 2 tuned LightGBM leads to the highest fulfillment rate but the least number of baskets advancement compared to a less accurate model like linear regression where more advanced quantities are fulfilled with higher leftovers. Also, for baskets of size more than 2, XGBoost is considered the best-performing model, yet it has the highest leftovers during the eight-week duration amongst all the simulated models and as well the highest advanced arc quantities.

# Chapter 4

## Basket data impact on forecasting aggregation and inventory fulfillment

### 4.1 Introduction

#### 4.1.1 General introduction and motivation

Supply chain operations in omnichannel retail are characterized by product hierarchies. The natural hierarchies can be structural, for example, when an omnichannel retailer operates in different segments such as appliances, decorations, and furniture. It can be geographically based on regions of sales or it can be cross-sectional such as section, product family, subfamily, and brand. As for the temporal hierarchy, naturally, a daily time series can be summed up into weekly, monthly, bi-monthly, quarterly, or annually or divided from aggregated temporal series to disaggregated temporal series. The upper levels tend to have lower variability than the lower dis-aggregated levels but lack details. A hierarchical forecast is considered coherent if the sum lower levels forecast is equal to the forecast of the upper levels [Babai et al. \(2022\)](#).

Hierarchical forecasting in supply chain management serves a variety of purposes in planning inventory, capacity planning, buying products/ raw materials, and budgeting. It complements and enforces decisions on multiple organizational levels. In the case of an omnichannel retailer, the lower levels of the forecasting hierarchy enforce the inventory planning inside the stores while a level up to the bottom level derives regional planning of inventory and distribution of products to these regions passing through country planning and reaching top-level organizational forecasting for budgeting and buying purposes. As a consequence, time-series hierarchies have natural hierarchical structures and temporal hierarchical structures [Syntetos et al. \(2016\)](#).

Furthermore, retail sales show the complementarity of specific products being bought together, in addition, customers tend to buy multiple-item baskets rather than one item per basket [Omar et al. \(2022\)](#). Our collaboration with an international cosmetics multi-channel retailer showed that a customer on average buys 2.5 items per order and that 65 % of baskets are baskets of items more than size 1. Moreover, When a customer chooses to buy a product, most likely the product has alternatives. The more alternatives during the recent historical period, the less the chance that this product will be bought. The attribute to measure this behavior is substitutability which is derived from basket data, as for complementarity, some products are usually bought with other products from the same category for example socks of different brands. We consider those attributes within category attributes. As for cross-category attributes, we use degree which is the number of distinct connections with other products within orders, and strength which is the sum of these connections per product [Omar et al. \(2022\)](#).

These findings raise questions on how can attributes from basket data improves the performance of forecasting methods on different structural hierarchical levels, especially since the basket attributes differ from a store level to a region level to an aggregate level.

### 4.1.2 Business context

Our research is with the same global omnichannel cosmetics retailer. They have large warehouses for brick-and-mortar (BM) products and international fulfillment centers for online channels. The retailer requires inventory and transportation management for every warehouse/fulfillment center/ store they operate. This mandates forecasting and planning for products from a short operational period for inventory, transportation, and promotion and a longer strategic period for strategic buying, resource availing, and allocation. For that purpose, the retailer applies a temporal forecasting hierarchy that is generated quarterly and divided into monthly, weekly, and daily. Furthermore, the forecast can be grouped by geographical dimension by store, warehouse, and central. Such flexibility is required by multinational retailers almost on a daily basis or planning. As such, an important question for such temporal/hierarchical planning is at which geographical level the main forecast should be applied with which forecasting model, and which aggregation/dis-aggregation approach should be followed. Note that retailers have an abundance of information on the association between products and that products influence each other positively by complimentary buying or negatively by substituting each other. As a result, in this project, we aim to test this hypothesis by incorporating basket data attributes to various hierarchical approaches and report their influence on both forecasting accuracy and inventory performance. In our research, we focus only on one city from this global supply chain where we test the impact of forecasting on the local warehouse (regional level) of this city and three stores (store level). Four different models are tested. AutoARIMA model where the orders and seasonal orders are automatically optimized (ARIMA), ARIMAX model where the basket data is used as regressors (ARIMA BASKET). A LightGBM model (LGBM) where only the lag of sales is used as regressor (LGBM-Lag), and finally, a LightGBM model where basket data attributes and the lag of sales are used (LGBM-BASKET).

### 4.1.3 Contributions and organization of the chapter

In this research work, we build on our previous work on modeling historical sales as a graph where products are the nodes and links between products are the baskets, we use the graph to extract degree, strength, complementarity, and substitutability from each product and use them as exogenous variables for forecasting each product at the store level and regional level. At the regional level, the sales of every product at the store level are aggregated as well as the graph for baskets. In other words, the entire basket sales of the region for a given time period is considered as one graph. Hence, this chapter contributes to the literature in two ways:

1. We assess the impact of the aforementioned basket sales attributes on the forecast accuracy of hierarchical aggregation approaches. To the best of our knowledge, none of the earlier research works has considered basket data in hierarchical forecasting
2. We are among the first to assess the inventory performance of hierarchical forecasting methods. We analyze the impact of hierarchical forecasting methods on the service level, safety stock, and lost sales using real data from the cosmetics retailer considered in the previous chapters.

The remainder of the chapter is organized as follows. Section 2 is dedicated to a review of the hierarchical forecasting approaches. In Section 3, we present our empirical analysis for forecast accuracy and inventory performance of the considered hierarchical forecasting approaches.

## 4.2 Hierarchical forecasting approaches

Demand forecasts are essential for planning in supply chain management. A considerable part of forecasting literature has been focused on natural level forecasts or single level forecasts [Rostami-Tabar et al. \(2015\)](#). Natural level forecasts do not always provide the best results as they do not take all information from lower levels or higher levels. In the framework proposed by [Babai et al. \(2022\)](#), they suggested that the level of granularity is determined by the decision impact and level of supply chain organization, the higher the level of the decision, the lower the granularity on both the natural and the temporal structure. In general, there has been an extensive body of literature on aggregation of forecasting levels since the 1950s mainly on the econometrics literature. There are several hierarchical forecasting approaches that have been proposed within the supply chain forecasting literature. Among these approaches, are the top-down (TD), the bottom-up (BU), the middle-out (MO), and the combination (COMB) approach.

### 4.2.1 Approaches' definitions

- Bottom-Up approach

In the BU forecasting approach, the lower levels are summed to make the aggregated levels thus constructing coherent forecasts. It can be illustrated as an S matrix that can show how the higher level of the hierarchical structure is calculated as:

$$y_{h,t} = S \cdot y_{b,t} \quad (4.1)$$

Where  $y_{b,t}$  is the bottom-level forecasts and  $y_{h,t}$  is the vector of hierarchical forecasts in the hierarchical structure,  $S$  is the matrix that includes the ownership of the lower level to the higher level, and in this approach, the base forecast is the lower level forecast which amongst the BU, TD, and MO is considered the approach with the highest cost, yet it is the approach that captures the most variability and characteristics From the lowest level.

- Top-Down approach

In the top-down approach, the upper-level forecast is disaggregated into the lower levels using calculated proportions. This will generate coherent forecasts in the hierarchy. There are many methods discussed in the literature for dis-aggregation, they are methods that use historical periods of the series and there are methods that use the forecasting proportions, some of the disaggregation methods use averages, lagged averages, and lagged proportions and combined proportions [Mircetic et al. \(2022\)](#). The most two common top-down approaches are average historical proportions and proportions of the historical averages [Gross and Sohl \(1990\)](#). In average historical proportions, the proportions can be denoted as :

$$p_j = \frac{1}{T} \sum_{t=1}^T \frac{y_{j,t}}{y_t} \quad (4.2)$$

where  $p_j$  is the proportion of the product and  $y_{j,t}$  is the forecast of product  $j$  at time  $t$  and  $y_{j,t}$  is the base forecast of the upper level. While the proportions of the historical averages are calculated by :

$$p_j = \frac{\sum_{i=1}^T \frac{y_{i,t}}{T}}{\sum_{i=1}^T \frac{y_t}{T}} \quad (4.3)$$

The forecasting proportion methods have the advantage of using forward-looking proportions as they change over time. In addition, top-down approaches based on historical averages tend to produce less accurate forecasts than bottom-up approaches [Athanasopoulos et al. \(2009\)](#). An h-step forecast is calculated at the bottom level for each series. These h-step forecasts are aggregated and the proportion for each series is calculated as the ratio of the h-step forecast of the series to the aggregated forecast. It can be summarized as :

$$p_j = \prod_{l=0}^{K-1} \frac{y_{j,h}^l}{S_{j,h}^l} \quad (4.4)$$

where  $y_{j,h}^l$  is the h-step initial forecast of the series that corresponds to the node which is  $l$  levels above node  $j$  and  $S_{j,h}^l$  is the sum of h-step initial forecasts.

- Middle-Out approach

The middle-out approach is a mix of the BU approach and the TD approach. The approach is only applicable when the structure has more than two levels as is the case of this work (product-store/product-region-total) or more levels are also possible. In this case, a middle level is chosen and the base forecast would be the level chosen, BU is implemented at the upper levels while TD is implemented at the lower levels [Hyndman and Athanasopoulos \(2018\)](#).

- Combination approach

The combination approach is the average combination of several approaches of the same model as an approach of ensemble. There is also the weighted combination approach where models are combined by a specified weight. The level of choice of the combination can be the middle level, bottom level, and top level. After the combination of levels of choice is selected, the forecast is either aggregated to upper levels or disaggregated to lower levels to generate coherent forecasts. Another weighted average approach is that the weight of each variation will be determined by the inverse of the error of each forecasting variation [Mircetic et al. \(2022\)](#), [Ballou \(2004\)](#). The weight of each product can be calculated as

$$y_{COMBw,h} = \sum_{m=1}^M W^m \cdot y_{B,h_m} \quad (4.5)$$

## 4.2.2 Findings on the approaches' performance

Having accurate forecasts is particularly important in supply chain operational and strategic planning. In strategic planning, forecasts generate budgeting and growth projections which in turn translate to material requirements, new contracts with suppliers and manpower, etc. While in operational planning, forecasts of demand lead to replenishment management, transport planning, and production planning. The decision on which forecast to select is based on the accuracy of the forecast model. In fact, having high forecast

errors can lead to higher /lower inventory, high transportation costs, higher wastes, and unnecessary promotions [Chopra and Meindl \(2016\)](#). As a result, hierarchical forecasting needs to be accurate on all levels not only the lower level or the higher level [Fliedner \(1999\)](#).

Researchers have a majority that BU forecasts lead to better overall performance compared to TD and MO. However, there is merit in using different approaches as evidenced by [Hyndman et al. \(2011\)](#). For instance, [Vogel \(2013\)](#) argued that lower-level data is more subject to errors and would lead to higher errors on higher levels. In addition, the cost of implementing a TD is significantly lower than BU or COMB approach and works well when the demand at the lower levels is stable [Weatherford et al. \(2001\)](#). As for BU proponents, they suggest that when there is a distinction in lower-level demand variability, the BU approach is more accurate. [Rostami-Tabar et al. \(2015\)](#) made an analytical investigation to identify when BU is superior to TD and vice versa. They used real demand data from a European flagship store. They concluded that when the moving average of the lower levels is identical, there is no significant difference between TD and BU at the upper level. While the TD performs better when the moving average of lower level products is of opposing sign amongst each other and they exhibit a high correlation among each other. While if the correlation is negative or low positive, BU performs better. In general, the BU performance becomes superior when the cross-correlation on the lower level decreases.

When forecasting demand for hierarchical structures, practitioners should determine first the forecasting model to use, second the structure of the hierarchy how it should be if grouped, hierarchical or temporal structure, this is based on the needs of the decision-makers and the level of the decision. Finally, the approach to be used should provide the most accurate forecasting [Mircetic et al. \(2022\)](#). It is considered that hierarchical forecasting is more challenging than direct traditional forecasting in general because they must be coherent. For example, in temporal forecasting, the higher monthly level must correspond to the combined values of the lower weekly levels and consequently, the weekly must correspond to the daily levels. Same for hierarchical cross-sectional forecasting, the forecast of one product in a region must correspond to the forecast of the same product in every store. In this work, we investigate the impact of utilizing basket attributes on a real data set on accuracy and inventory performance. We note that the granularity of the data is daily. The structure of the hierarchy is considered to be a grouped hierarchical structure (Store-Product),(Region-Product) and the top level. The bottom level consists of the product sales in each store and this level will be used later on for inventory replenishment inside the stores. there are three stores in the structure (level 0), the upper level is the total sales of a product in the region, this level is used to plan for stocks of the chapter in the regional warehouse and the frequent regional transportation from the main fulfillment center. While the top-level forecast is the aggregate daily sales in the region. The comparison of performance between the different forecasting approaches in this research will be between BU, MO, TD (proportions of the historical averages), and COMB. The models used in the comparison are ARIMA, ARIMAX with basket attributes (ARIMA BASKET), Light gradient boosting model with a lag of sales (LGBM), and with a lag of sales and basket attributes (LGBM BASKET).

## 4.3 Considered basket data attributes and methods

### 4.3.1 Basket data attributes

To illustrate basket data attributes, we refer to Figure 3.10 from chapter 2 which illustrates a sample network of products where basket data is calculated. In this work, four basket

attributes are identified, based on the shopping basket data from sales history at the store level: (1) *Degree* of the product, the number of connections (or arrows) with the other products of the multi-category graph; (2) *Strength* of the product, the total frequency with all connected products; (3) *Substitutability* of the product, the chance to be sold as a substitute among all the products within its substitutability set; (4) *Complementarity* of the product, the proportion between the number of connections within the substitutability set and the number of connections with the other products of the assortment (i.e., *Degree*). The four attributes are used as regressors with historical sales within a forecasting model. This forecasting model can be any model that accepts regressors.

As the basket data is directly related to basket information, it can also be extracted from higher levels, not only the store level. Practitioners have two choices when calculating higher-level basket data regressors. Either by aggregating the regressor values of the bottom level to form the higher level or by constructing graphs directly on higher levels to extract basket attributes. For example, at the regional level, the whole baskets in the region can be used to construct the graph, where basket attributes are extracted and the same for higher levels as the country. In Figure 4.1, we present the hierarchy of the work. Note that this is a grouped hierarchical structure as level 1 (Region/Product). It is still segregated by the product contrary to traditional hierarchy where the aggregate sales of the region per time period is constructed. The reason is that operationally the forecast of every product by region is required for inventory and transportation planning to this region. As shown, the structure is only a part of the hierarchy, for instance, the total sales of the region by product are further processed at higher levels from regional fulfillment centers to manufacturing facilities and/or suppliers.

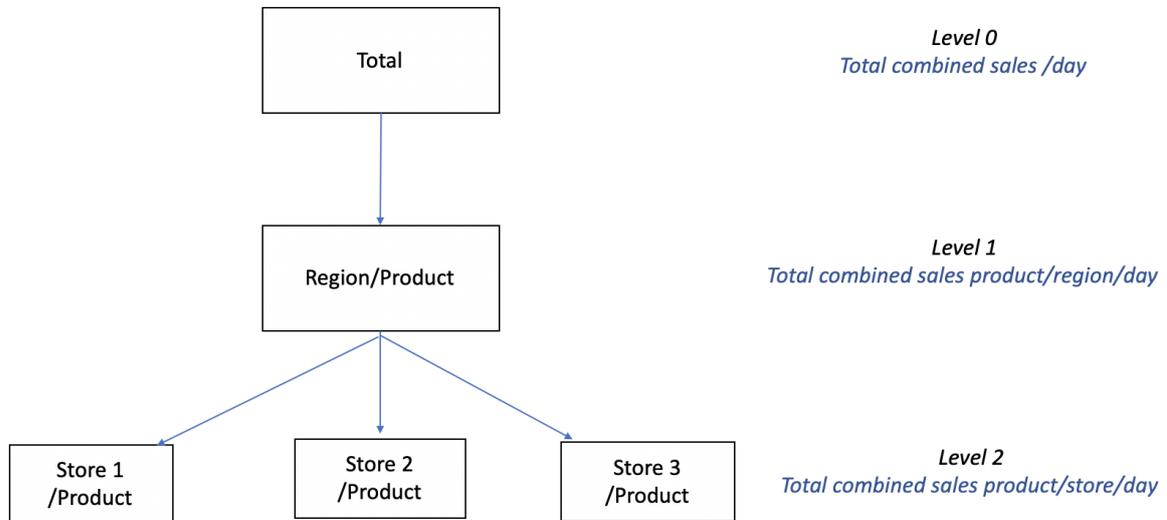


Figure 4.1: Hierarchical grouped structure

### 4.3.2 Forecasting models

To be able to assess the impact of basket data on hierarchical forecasting, four different models are tested. AutoARIMA model, where the orders and seasonal orders are automatically optimized (ARIMA), ARIMAX model, where the basket data is used as regressors (ARIMA BASKET). A LightGBM model, where only lag of sales is used as regressor (LGBM), and finally, a LightGBM model, where basket data attributes and the lag of sales are used (LGBM-BASKET).

At any time period  $t$  and for each product  $p$ , the forecast, denoted by  $y_{pt}$ , is calculated for a network historic horizon  $h(t)$  using the *degree*, *strength*, *substitutability*,

and *complementarity* regressors, denoted by  $Degree_{pt}$ ,  $Strength_{pt}$ ,  $Substitutability_{pt}$ , and  $Complementarity_{pt}$  respectively (regressors calculated using basket data on periods  $t - 1$ ,  $t - 2, \dots, t - h$ ), in each proposed forecasting model as follows:

- The ARIMAX model:

$$y_{pt} = aDegree_{pt} + bStrength_{pt} + cSubstitutability_{pt} + dComplementarity_{pt} + n_{pt} \quad (4.6)$$

where  $a$ ,  $b$ ,  $c$ , and  $d$  are the regressors respective coefficients. We assume that  $n_{pt}$  is given by an ARIMA( $r, k, q$ ) model, where  $r$  is the auto-regressive order,  $k$  is the differencing order, and  $q$  is the moving average order.

- The LightGBM machine learning model:

We recall that LightGBM is a gradient-boosting model based on the decision tree algorithm. We use degree, strength, substitutability, complementarity, transitivity, and lag sales of product  $p$  at time period  $t - 1$  as inputs of this model to forecast the sales at time period  $t$  (LGBM BASKET). While LGBM is using only the lag of sales as a regressor (LGBM). The forecasts are all performed using the SKTIME library in Python.

### 4.3.3 Forecast accuracy measures

In order to evaluate the forecasting performance of each model/approach, each time series is divided into training and out-of-sample. Training data starts from the beginning of the year to 18 October and testing from 19 October to 31 December 2018. The root mean squared scaled error (RMSSE), mean error (ME) and mean absolute scaled error (MASE) will be used to report the accuracy of each mode/approach. The RMSSE and MASE are used to calculate the scaled errors dependent on the scale of the time series as recommended by [Hyndman and Koehler \(2006\)](#). The accuracy will be calculated at each time series and then averaged per model/approach. there will also be a separate accuracy reporting on the bottom level segregated by the classification of [Syntetos and Boylan \(2005\)](#) as will be later presented in the following section. The errors considered will be the MASE and the RMSE.

### 4.3.4 Inventory performance

The resulting forecasting results from all of the hierarchical structures are tested on inventory performance using a periodic order up-to-level policy (R,S), where an order is made for replenishment frequently every given time period noting that this policy adheres to the policy implemented by the retailer in stores and local fulfillment centers. The time from order to delivery is one day in both the stores and local fulfillment centers. To calculate the order-up-to-level at any period  $t$ , as commonly recommended in the literature ([Syntetos et al. 2015](#), [Van der Auweraer and Boute 2019](#)), the mean and variance of the lead-time demand distribution, denoted by  $\mu_{LT}$  and  $Var_{LT}$  respectively, is estimated as  $\mu_{LT} = (L + 1) * y_{pt}$  and  $Var_{LT} = (L + 1) * MSE_{pt}$  where  $L$  is the lead-time,  $y_{pt}$  is the forecast of sales at period  $t$  and  $MSE_{pt}$  is the mean squared forecast error at period  $t$  ( $L + 1$  periods are considered instead of  $L$  to account for the protection interval under a periodic review). Note also that  $MSE_{pt}$  is given by:  $MSE_{pt} = \alpha[y_{pt} - D_{pt}]^2 + (1 - \alpha)MSE_{p(t-1)}$ , where  $\alpha$  is a smoothing constant between 0 and 1. The order-up-to-level at any period  $t$ , denoted by  $S_{pt}$ , is then calculated as  $S_{pt} = F_{LTD}^{-1}(CSL)$  where  $F_{LTD}(\cdot)$  is the cumulative distribution function of the lead-time demand, having the mean and variance equal to

$\mu_{LT}$  and  $Var_{LT}$ . More details on the calculation of the order-up-to-level and the implementation of the simulation model of the inventory policy are given in (Hasni et al. 2019, Babai et al. 2019).

The resulting outcome will be compared on two levels, level 2 where the demand is disaggregated by store, and level 1 where demand is aggregated at local fulfillment centers for each model; ARIMA, ARIMA BASKET, LGBM, and LGBM BASKET, for each hierarchical forecasting approach TD, BU, MO and COMB. The comparison shows the performance of each model/approach on average inventory holding quantity and backlog quantity across 5 points of CSL from 90% to 99%.

### 4.3.5 Empirical investigation

This work is part of a collaboration project with a multinational cosmetics retailer based in France. The retailer operates in multiple channels that started with brick-and-mortar stores and then online channels at the beginning of the 2000s and integration between both channels such as buy online pick from store (BOPS) where a customer buys online and picks the product from the store. The data of focus in this project is three stores in one major city in France. The data contains the transaction data of baskets including the date of purchase, order number, and list of items inside the order (Basket). On average, 600 baskets are purchased per day with different size variations. The data length is one year from January 1 to December 31, 2018. To be able to formulate the hierarchies, a group of products was selected from the total offerings. In our attempt to filter the products in the study we adhered to the following conditions : 1- It must be sold in the three stores. 2- A sufficient sample and representation of each classification of product demand based on Syntetos and Boylan (2005) referred to hereafter as SBC.

While adhering to the criteria, not all classes have the same number in each store, there is a domination of the Intermittent class, followed by Lumpy, then erratic, and finally smooth. Recall that in the SBC scheme, the demand is classified based on the average demand interval (ADI) and squared coefficient of variation of demand sizes ( $CV^2$ ) with the cut-off values of  $ADI = 1.32$  and  $CV^2 = 0.49$ . Four classes are identified: smooth ( $ADI < 1.32$  and  $CV^2 < 0.49$ ), lumpy ( $ADI > 1.32$  and  $CV^2 > 0.49$ ), erratic ( $ADI < 1.32$  and  $CV^2 > 0.49$ ), and intermittent ( $ADI > 1.32$  and  $CV^2 < 0.49$ ). The class of smooth demand is the class where the demand has a high number of demand occurrences (i.e. low demand intervals) and a low variation of demand sizes. The erratic demand is also characterized by a high number of demand occurrences but it has high demand sizes' variability. The intermittent demand has few demand occurrences and a low variation of demand sizes. Finally, the class of lumpy demands is characterized by few demand occurrences but a high variation of demand sizes. This classification is used further in the chapter to identify the appropriate forecasting method in each class and the related inventory performance.

In Figure 4.2, the ADI is plotted against  $CV^2$  for the three stores and the aggregated city, we notice there are slight differences in pattern between the aggregated city and the stores. For example, the aggregation of sales of a particular product across the three stores leads to a reduction of the average demand interval in the aggregated/city. We also note that the majority of products are either lumpy or intermittent with few products that are smooth or erratic. Also, it is interesting to see if there is a rate of change between classification from one store to another store and from stores to aggregated product sales of the city. It is noted in Table 4.1 that aggregation leads to a reduction of intermittent classification and an increase of smooth, erratic, and lumpy by a small percentage.

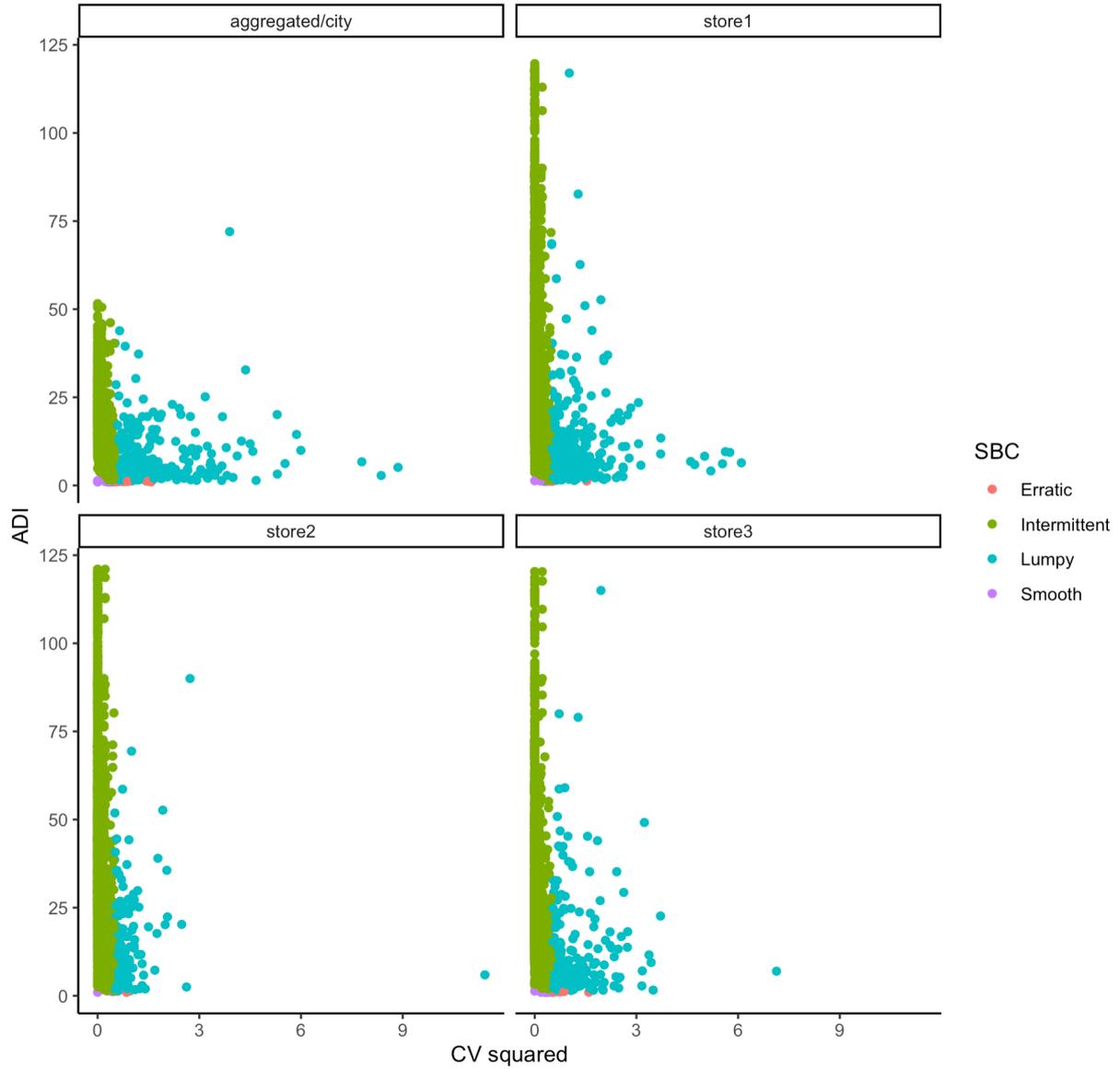


Figure 4.2: Hierarchical grouped structure

	aggregated/region	store1	store2	store3
Erratic	1.5%	0.1%	0.1%	1.5%
Intermittent	88.9%	90.3%	95.0%	90%
Lumpy	9.2%	8.9%	4.8%	8.2%
Smooth	1.5%	0.7%	0.1%	0.13%

Table 4.1: Percentage of products obtained after aggregation

## 4.4 Forecast accuracy results

In this section, the results of the different accuracy measures across all levels are presented. There are four forecasting models with four aggregation forecasting approaches: the ARIMA model with its aggregation variations; MO, BU, TD, COMB which is a simple average of the former 3, it is the same with ARIMA BASKET, the LGBM model with a lag of sales only included, and finally, LGBM model with a lag of sales degree, strength, substitutability, and complementarity. In Table 4.2, the results of the top aggregated

level are presented. We notice the out-performance of the ARIMA BASKET aggregations over the other model/ aggregation approaches particularly the ARIMA BASKET COMB and the ARIMA BASKET TD, if RMSSE is to be the determinant. ARIMA BASKET COMB is the better-performing model across all variations. Note that in the machine learning comparison, LGBM COMB is the better-performing model across all LGBM variations. In addition, we notice that basket attributes with LGBM are not better in performance compared to all the variations of aggregation approaches. On the contrary, ARIMA basket (ARIMAX) variations are more accurate than their ARIMA counterparts. For example, there is a 12% reduction in error between ARIMA BASKET BU and ARIMA BU for MASE, an 18% reduction in error between ARIMA BASKET COMB and ARIMA COMB and a much more significant reduction of 40% between the ARIMA BASKET TD and ARIMA TD. Similar performance is achieved by the RMSSE metric error with 21%, 16%, and 34 % respectively.

	model	MASE	RMSSE
1	ARIMA BASKET BU	1.124	1.106
2	ARIMA BASKET COMB	<b>1.081</b>	1.070
3	ARIMA BASKET MO	1.175	1.112
4	ARIMA BASKET TD	1.101	1.073
5	ARIMA BU	1.186	1.162
6	ARIMA COMB	1.332	1.285
7	ARIMA MO	1.167	1.153
8	ARIMA TD	1.864	1.636
9	LGBM BASKET BU	1.288	1.243
10	LGBM BASKET COMB	1.187	1.181
11	LGBM BASKET MO	1.262	1.217
12	LGBM BASKET TD	1.213	1.157
13	LGBM BU	1.202	1.180
14	LGBM COMB	1.129	1.119
15	LGBM MO	1.211	1.177
16	LGBM TD	1.094	<b>1.060</b>

Table 4.2: Aggregated top-level accuracy

For the Region/product level, we present the accuracy measures over all models in Table 4.3 and the quartile distribution starting from the 10<sup>th</sup> percentile to the 90<sup>th</sup> percentile in Figure 4.3. The results show, while calculating the aggregate average in Table 4.3, ARIMA models are generally better than ARIMA basket models. For machine learning models, LGBM basket variations remain the better-performing model among all variations. Consequently, the best performing models are ARIMA TD for MASE and ARIMA COMB for RMSSE. However, based on the the box plot in Figure 4.3, the ARIMA BASKET TD is the best-performing model across the distribution of RMSSE the difference is not significant across all models. When a deeper analysis considering the SBC classification in Figure 4.4, we can note that there is a significant difference in error distribution particularly in the lumpy, smooth, and erratic classes. For example in the smooth category, the ARIMA BASKET variations are consistently outperforming the ARIMA variations across TD, BU, and MO with a significant difference in the range of the distribution. This concurs with [Omar et al. \(2022\)](#) that ARIMA basket attributes perform well with the lumpy category. For the erratic category, the ARIMA MO is a distinct winner for having a lower median and a lower third quartile. For the intermittent category, there is no clear winner as all models are in line with each other. However, the

model with the lowest median error is ARIMA BASKET TD. For the smooth category, the model with the lowest median is ARIMA BASKET COMB. While the average accuracy measures in Table 4.3 shows that ARIMA basket is not a top performer model, the average measure tends to be skewed by high error values. Accordingly, we adhere to [Mircetic et al. \(2022\)](#) when comparing distribution across all errors via Figure 4.3 and Figure 4.4. Upon investigation, we notice that the basket methods result in high errors when there are no sales observations during the initial training window, accordingly, there are no attributes sufficient enough for the model to be well trained on the pattern on demand. However, these products are outliers to the data such as limited promotion products that are only introduced during the testing period. A sample of the demand pattern of these products is presented in Figure 4.9

	model	MASE	RMSSE
1	ARIMA BASKET BU	2.408	1.025
2	ARIMA BASKET COMB	3.523	1.167
3	ARIMA BASKET MO	4.590	1.368
4	ARIMA BASKET TD	1.745	0.864
5	ARIMA BU	1.789	0.837
6	ARIMA COMB	1.718	<b>0.835</b>
7	ARIMA MO	1.911	0.853
8	ARIMA TD	<b>1.595</b>	0.867
9	LGBM BASKET BU	1.788	0.873
10	LGBM BASKET COMB	1.794	0.870
11	LGBM BASKET MO	1.787	0.880
12	LGBM BASKET TD	1.814	0.875
13	LGBM BU	1.810	0.885
14	LGBM COMB	1.809	0.873
15	LGBM MO	1.826	0.905
16	LGBM TD	1.802	0.870

Table 4.3: Regional level accuracy

To further illustrate the performance of the forecasting methods, we present the percentiles overall and for the erratic and lumpy categories at the regional level. Overall in Table 4.4, the only consistent performance across the percentiles between a basket method and a method that is not using basket attributes on the aggregation approach is LGBM BASKET MO against LGBM MO for the RMSSE, while for all other comparisons, there is no significant difference. For the lumpy category, note the out-performance of the ARIMA BASKET BU compared to the ARIMA BU and the out-performance of this particular model compared to all the other models in this category, same for the erratic category the ARIMA BASKET BU is better than the ARIMA BU up to the 70<sup>th</sup> percentile. But overall, the best model in this category is the ARIMA MO which is consistently lower than all the other models.

For the bottom-level accuracy, on the average measure the ARIMA COMB, ARIMA TD, and ARIMA BU are performing better than their counterparts. This is due to the existence of outliers that derive the average higher for other models as explained in the regional forecasting performance, we follow suit of [Mircetic et al. \(2022\)](#) and present the error distribution using a box plot to understand more the performance of the models. In Figure 4.5, there is no significant accuracy top performer across all models overall. However, the model with the lowest median is ARIMA BASKET TD with a median of 0.615, if it is compared with ARIMA TD, the median is 0.618. This is also shown in detail

model	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
1 ARIMA BASKET BU	0.04	0.20	0.47	0.59	0.69	0.78	0.89	1.04	1.35	8.03
2 ARIMA BASKET COMB	0.09	0.22	0.46	0.58	0.67	0.77	0.87	1.02	1.31	9.06
3 ARIMA BASKET MO	0.00	0.20	0.47	0.59	0.69	0.78	0.89	1.04	1.34	10.50
4 ARIMA BASKET TD	0.15	0.32	0.47	0.57	0.67	0.76	0.86	1.01	1.29	6.52
5 ARIMA BU	0.05	0.21	0.46	0.58	0.68	0.77	0.88	1.02	1.29	5.86
6 ARIMA COMB	0.07	0.19	0.45	0.57	0.67	0.77	0.88	1.02	1.29	6.09
7 ARIMA MO	0.00	0.21	0.46	0.58	0.68	0.78	0.88	1.02	1.31	7.45
8 ARIMA TD	0.10	0.24	0.45	0.57	0.69	0.79	0.89	1.05	1.34	6.54
9 LGBM BASKET BU	0.14	0.33	0.49	0.59	0.68	0.78	0.88	1.02	1.31	6.55
10 LGBM BASKET COMB	0.16	0.33	0.48	0.58	0.68	0.77	0.87	1.02	1.30	6.54
11 LGBM BASKET MO	0.14	0.34	0.50	0.61	0.69	0.79	0.89	1.04	1.31	6.55
12 LGBM BASKET TD	0.18	0.35	0.49	0.59	0.68	0.77	0.87	1.02	1.30	6.51
13 LGBM BU	0.14	0.34	0.50	0.62	0.70	0.80	0.89	1.04	1.31	6.55
14 LGBM COMB	0.16	0.34	0.49	0.60	0.68	0.77	0.88	1.02	1.29	6.54
15 LGBM MO	0.15	0.36	0.54	0.64	0.73	0.82	0.92	1.05	1.32	6.55
16 LGBM TD	0.17	0.35	0.49	0.58	0.67	0.76	0.86	1.01	1.28	6.51

Table 4.4: Regional level accuracy percentiles-RMSSE

in Table 4.8, it also has the lowest error up to the 99<sup>th</sup> percentile. Note that the median is calculated with outliers included. When a deeper analysis of the SBC is made in Figure 4.6, we note the big difference in the lumpy category distribution for the ARIMA BASKET BU compared to the other models, it is a clear top-performing model in this category. In the erratic category, the ARIMA BU with a wider distribution has the lowest median. As for intermittent and smooth, there is no clear difference in which model is a better performing model. This is also shown in Table 4.9 where on the 70<sup>th</sup> percentile, where the performance improvement between the ARIMA BASKET BU and the ARIMA BU is 47%. As for the erratic category shown in Table 4.10, there is also an improvement across percentiles between the forecasting methods using basket attributes and the forecasting methods without basket attributes. For example, the performance improvement at the median level between ARIMA BASKET BU and ARIMA BU is 46 %.

model	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
1 ARIMA BASKET BU	0.00	0.00	0.00	0.01	0.03	0.08	0.19	1.06	3.05	25.21
2 ARIMA BASKET COMB	0.00	0.02	0.04	0.06	0.10	0.14	0.28	0.87	3.33	27.83
3 ARIMA BASKET MO	0.00	0.00	0.00	0.00	0.00	0.10	0.27	0.94	2.96	32.25
4 ARIMA BASKET TD	0.00	0.06	0.10	0.16	0.23	0.29	0.41	0.81	2.95	25.15
5 ARIMA BU	0.00	0.00	0.00	0.02	0.07	0.14	0.25	0.83	3.23	23.37
6 ARIMA COMB	0.00	0.01	0.03	0.06	0.09	0.15	0.23	0.83	2.93	23.91
7 ARIMA MO	0.00	0.00	0.00	0.00	0.03	0.15	0.29	0.83	3.23	25.14
8 ARIMA TD	0.00	0.04	0.07	0.11	0.15	0.20	0.29	0.77	3.16	25.23
9 LGBM BASKET BU	0.00	0.01	0.07	0.11	0.20	0.31	0.46	0.86	2.90	25.25
10 LGBM BASKET COMB	0.00	0.03	0.08	0.14	0.22	0.31	0.45	0.87	2.86	25.22
11 LGBM BASKET MO	0.00	0.01	0.06	0.11	0.20	0.33	0.49	1.13	2.81	25.25
12 LGBM BASKET TD	0.00	0.06	0.12	0.18	0.25	0.33	0.46	0.82	2.93	25.16
13 LGBM BU	0.00	0.01	0.06	0.12	0.21	0.31	0.51	1.12	2.81	25.25
14 LGBM COMB	0.00	0.03	0.08	0.15	0.22	0.32	0.47	1.08	2.83	25.22
15 LGBM MO	0.00	0.01	0.05	0.12	0.22	0.35	0.58	1.22	2.82	25.25
16 LGBM TD	0.00	0.06	0.11	0.18	0.25	0.32	0.45	0.81	2.93	25.15

Table 4.5: Regional level accuracy percentile-Lumpy-RMSSE

	model	MASE	RMSSE
1	ARIMA BASKET BU	1.855	0.822
2	ARIMA BASKET COMB	2.261	0.851
3	ARIMA BASKET MO	2.709	0.929
4	ARIMA BASKET TD	1.474	0.733
5	ARIMA BU	1.553	0.724
6	ARIMA COMB	1.468	<b>0.716</b>
7	ARIMA MO	1.611	0.725
8	ARIMA TD	<b>1.300</b>	0.727
9	LGBM BASKET BU	1.509	0.739
10	LGBM BASKET COMB	1.519	0.736
11	LGBM BASKET MO	1.507	0.739
12	LGBM BASKET TD	1.543	0.741
13	LGBM BU	1.530	0.751
14	LGBM COMB	1.536	0.739
15	LGBM MO	1.548	0.753
16	LGBM TD	1.533	0.738

Table 4.7: Store level accuracy

model	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
1 ARIMA BASKET BU	0.00	0.00	0.01	0.07	0.18	0.52	0.59	0.81	1.34	1.66
2 ARIMA BASKET COMB	0.06	0.20	0.23	0.25	0.44	0.61	0.68	0.74	1.13	1.72
3 ARIMA BASKET MO	0.00	0.00	0.01	0.02	0.36	0.67	0.81	0.93	1.12	1.69
4 ARIMA BASKET TD	0.17	0.51	0.53	0.59	0.62	0.65	0.69	0.83	1.20	1.80
5 ARIMA BU	0.00	0.00	0.01	0.02	0.30	0.58	0.63	0.64	1.45	1.91
6 ARIMA COMB	0.04	0.14	0.15	0.16	0.34	0.58	0.63	0.69	1.43	1.98
7 ARIMA MO	0.00	0.00	0.00	0.00	0.28	0.59	0.65	0.84	1.41	2.01
8 ARIMA TD	0.11	0.37	0.41	0.43	0.51	0.64	0.65	0.69	1.45	2.05
9 LGBM BASKET BU	0.02	0.05	0.08	0.16	0.56	0.60	0.71	0.99	1.43	1.97
10 LGBM BASKET COMB	0.07	0.26	0.28	0.31	0.55	0.61	0.71	0.91	1.35	1.93
11 LGBM BASKET MO	0.01	0.03	0.05	0.08	0.56	0.64	0.71	1.04	1.44	1.96
12 LGBM BASKET TD	0.19	0.61	0.64	0.69	0.71	0.76	0.82	0.95	1.26	1.90
13 LGBM BU	0.01	0.03	0.07	0.17	0.56	0.58	0.67	1.07	1.29	1.77
14 LGBM COMB	0.06	0.25	0.27	0.30	0.55	0.59	0.63	0.99	1.26	1.82
15 LGBM MO	0.00	0.01	0.02	0.03	0.59	0.68	0.70	1.07	1.36	1.94
16 LGBM TD	0.18	0.57	0.59	0.64	0.70	0.74	0.77	0.93	1.20	1.83

Table 4.6: Regional level accuracy percentile-Erratic-RMSSE

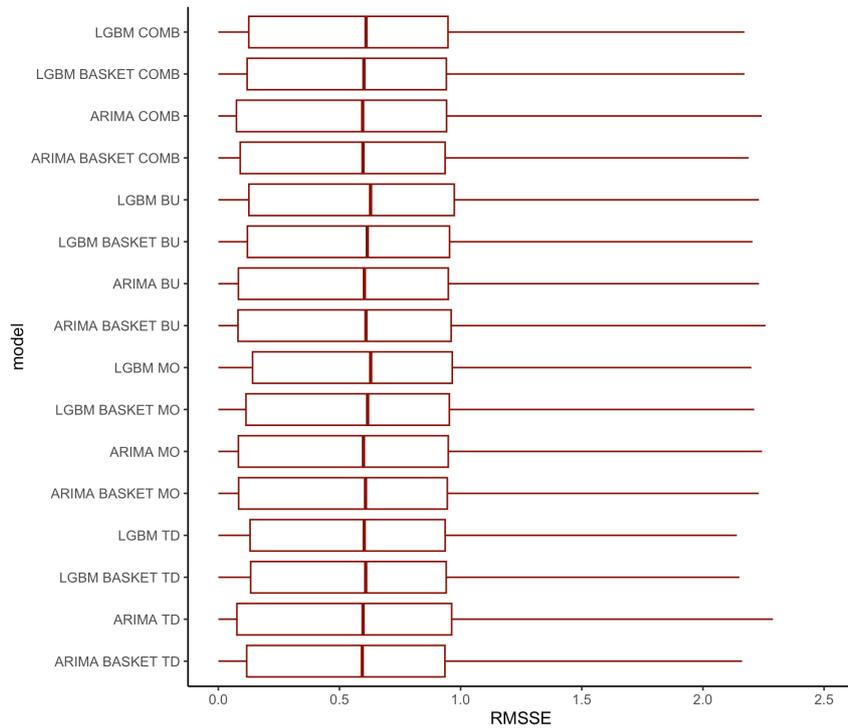


Figure 4.5: Store level error quartiles

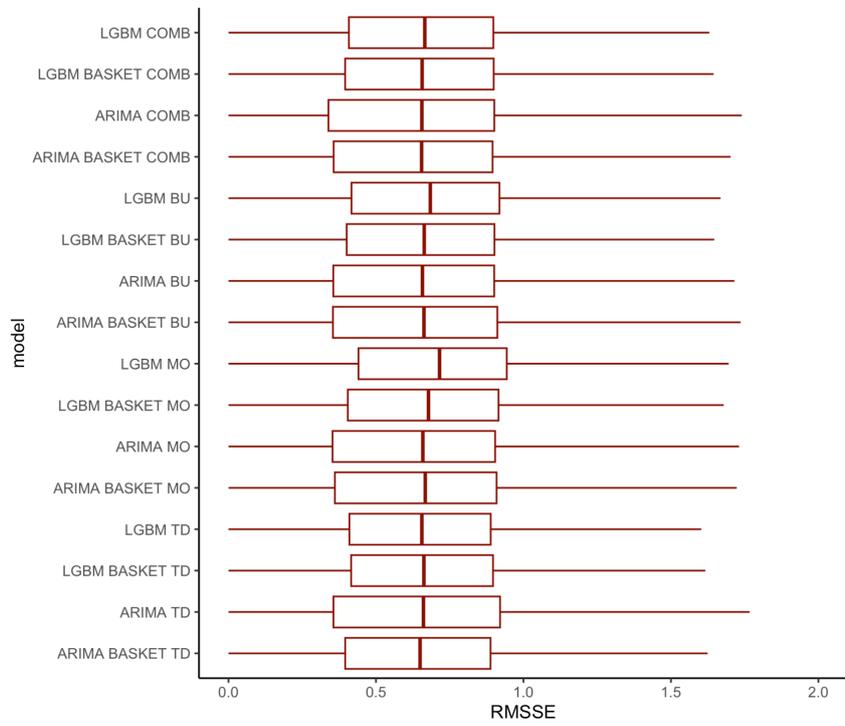


Figure 4.3: Regional level error quarterlies

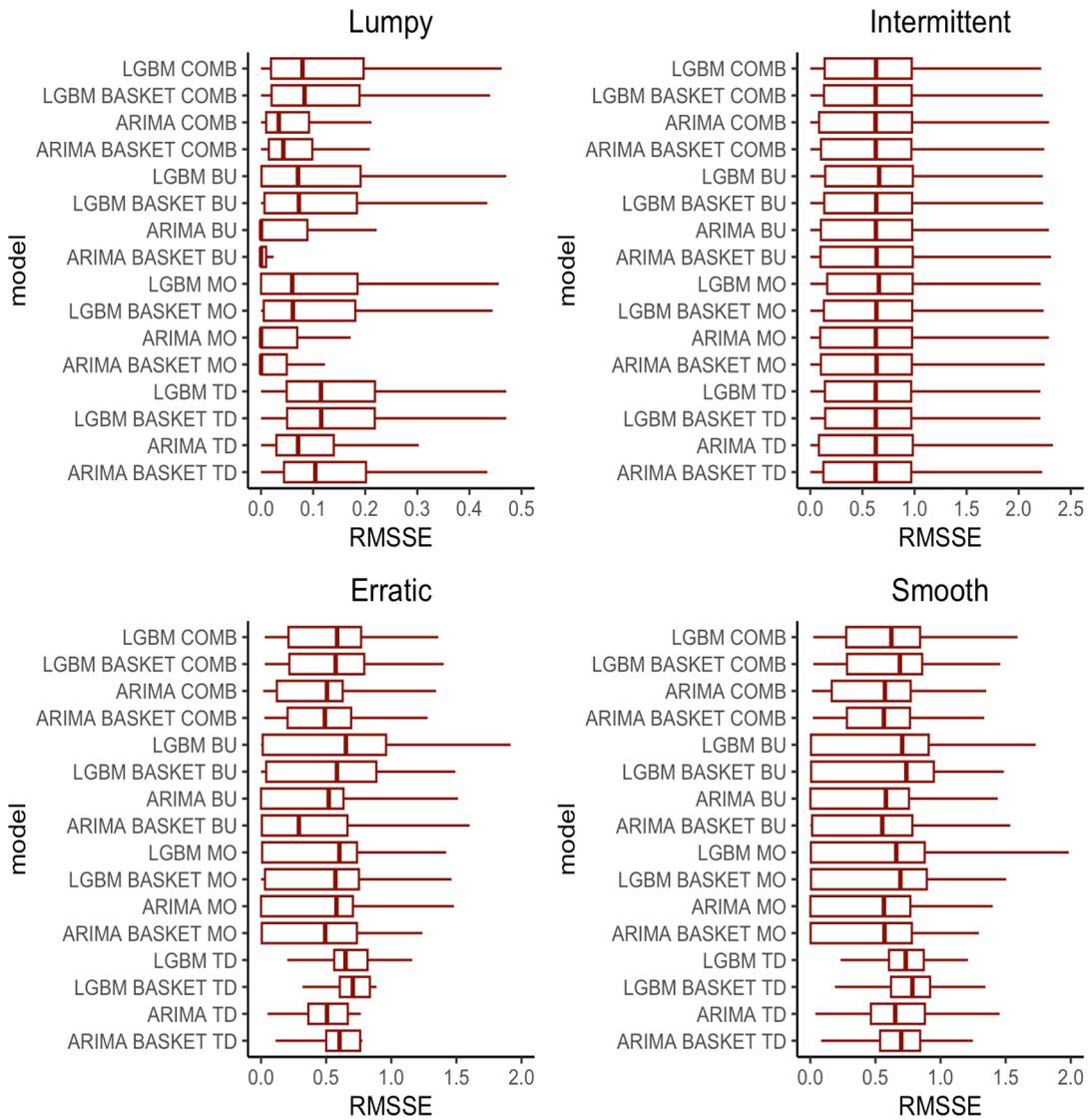


Figure 4.6: Store level error quartiles by class

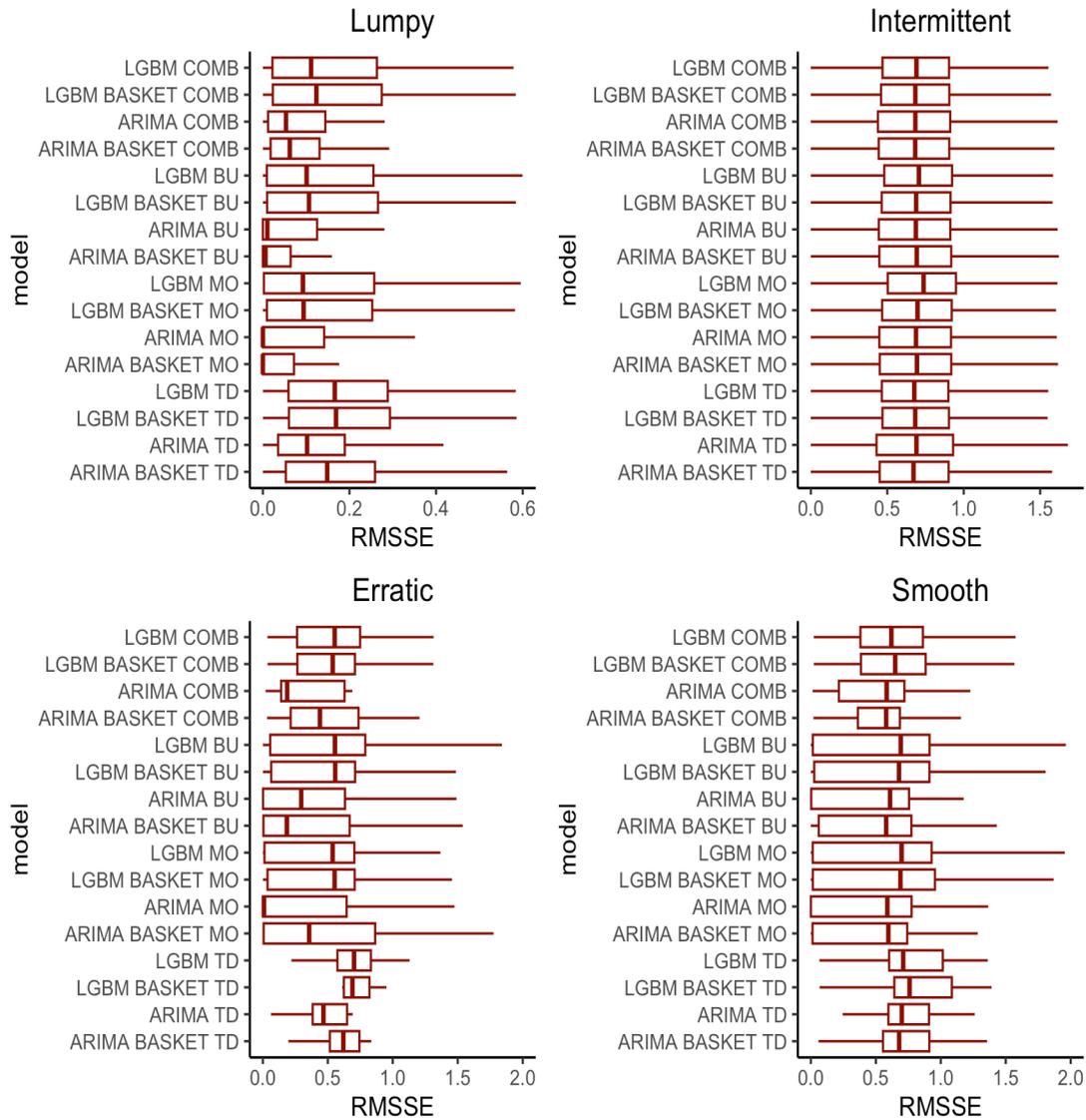


Figure 4.4: Regional level error quartiles by class

## 4.5 Inventory performance results

For the purpose of the analysis, we fix three target CSLs—90%, to 99%—and for target CSL and each scenario, we measure the average inventory holding volumes, the average back-ordering volumes, and the achieved CSLs (averages calculated over the evaluation period and across all products). Note that the last 73 days (from October 20, 2018, to December 31, 2018) are used to evaluate the performance. The lead time  $L$  is 2 days, and the smoothing error is 0.7. we first present the aggregate level results on the mid-level distribution center and the bottom-level stores. In Table 4.11, we give the average inventory performance per model at three service level points 0.9, 0.95, and 0.99. The average is calculated up to the 95 percentile to remove outliers which means that average calculations represent 95 % of the product simulations. Inventory volume ( I.V) represents the average inventory holding per period per product under each model implementation. Lost sales are the average of the total lost sales per product when no there is no inventory at the time of demand. Csl is the average realized cycle service level and safety stock (S.S) is the average safety stock per period per product under each model. It can be noted at the distribution center overall, the lower the average safety stock, the more accurate the forecast as the error is less and thus the safety stock quantity. Yet, it does not necessarily result in better inventory performance. On the other hand, combination models result in

model	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
1 ARIMA BASKET BU	0.00	0.06	0.11	0.47	0.62	0.75	0.89	1.07	1.41	4.67
2 ARIMA BASKET COMB	0.04	0.07	0.12	0.46	0.62	0.74	0.88	1.05	1.40	4.69
3 ARIMA BASKET MO	0.00	0.07	0.12	0.47	0.62	0.74	0.88	1.06	1.40	5.23
4 ARIMA BASKET TD	0.06	0.10	0.17	0.47	0.61	0.73	0.87	1.04	1.40	3.87
5 ARIMA BU	0.00	0.07	0.12	0.47	0.62	0.74	0.88	1.05	1.41	4.22
6 ARIMA COMB	0.04	0.06	0.10	0.45	0.62	0.74	0.88	1.05	1.40	3.92
7 ARIMA MO	0.00	0.06	0.11	0.46	0.62	0.74	0.88	1.05	1.40	3.98
8 ARIMA TD	0.04	0.06	0.11	0.45	0.62	0.74	0.88	1.07	1.40	3.94
9 LGBM BASKET BU	0.06	0.10	0.17	0.47	0.62	0.74	0.88	1.06	1.40	3.95
10 LGBM BASKET COMB	0.07	0.10	0.17	0.47	0.62	0.74	0.88	1.05	1.40	3.94
11 LGBM BASKET MO	0.06	0.09	0.17	0.47	0.62	0.74	0.88	1.06	1.40	3.95
12 LGBM BASKET TD	0.07	0.11	0.19	0.47	0.62	0.74	0.87	1.04	1.40	3.86
13 LGBM BU	0.06	0.10	0.18	0.50	0.64	0.76	0.90	1.07	1.40	3.95
14 LGBM COMB	0.07	0.10	0.18	0.48	0.62	0.74	0.88	1.05	1.40	3.93
15 LGBM MO	0.06	0.11	0.21	0.50	0.64	0.76	0.89	1.07	1.40	3.95
16 LGBM TD	0.07	0.11	0.19	0.47	0.62	0.73	0.87	1.04	1.40	3.85

Table 4.8: Store level accuracy percentile-RMSSE

better fulfillment, specifically the ARIMA BASKET COMB which has the best fulfillment at the cycle service level points of 0.9 and 0.99. In addition, if we compare the ARIMA basket BU and the ARIMA BU, the ARIMA BASKET BU is consistently less in safety stock and less in lost sales at the cycle service level point of 0.9 and 0.95. The same consistently prevails on store level where ARIMA BASKET COMB is better in fulfillment at the cycle service level 0.95 and 0.99.

model	csl= 90%			csl= 95%			csl= 99%		
	lost sales	csl	S.S	lost sales	csl	S.S	lost sales	csl	S.S
ARIMA BASKET TD	2.72	95	1.53	1.86	96	2.30	0.71	98	4.32
ARIMA TD	<b>1.51</b>	93	<b>1.05</b>	<b>2.18</b>	95	<b>1.64</b>	1.23	97	3.45
LGBM BASKET TD	<b>2.18</b>	96	1.63	1.47	97	6.50	0.56	98	4.51
LGBM TD	2.39	96	1.65	1.60	97	6.24	0.63	98	4.52
ARIMA BASKET MO	2.39	96	1.35	1.74	97	1.95	0.99	98	3.35
ARIMA MO	2.38	96	1.44	1.58	97	2.15	0.63	98	3.96
LGBM BASKET MO	3.00	95	1.50	2.16	96	2.24	1.04	97	4.08
LGBM MO	2.87	95	1.37	2.22	96	1.92	1.42	97	<b>3.18</b>
ARIMA BASKET BU	2.21	96	1.49	1.46	97	2.21	0.62	98	4.01
ARIMA BU	2.32	96	1.52	1.50	97	2.26	0.57	98	4.16
LGBM BASKET BU	2.87	95	1.58	2.01	96	2.36	0.99	97	4.27
LGBM BU	2.69	95	1.56	1.94	96	2.33	0.96	97	4.25
ARIMA BASKET COMB	2.25	96	1.58	<b>1.44</b>	97	2.34	<b>0.55</b>	98	4.27
ARIMA COMB	2.81	96	1.46	1.83	97	2.22	0.64	98	4.25
LGBM BASKET COMB	2.61	96	1.62	1.79	96	6.20	0.75	98	4.45
LGBM COMB	2.54	96	1.64	1.75	96	2.43	0.74	98	4.45

Table 4.11: Inventory performance - regional level

model	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
1 ARIMA BASKET BU	0.00	0.00	0.00	0.00	0.00	0.02	0.11	0.72	2.58	18.51
2 ARIMA BASKET COMB	0.00	0.01	0.03	0.04	0.07	0.10	0.19	0.67	2.51	19.52
3 ARIMA BASKET MO	0.00	0.00	0.00	0.00	0.00	0.06	0.18	0.76	2.49	23.13
4 ARIMA BASKET TD	0.00	0.05	0.07	0.11	0.15	0.23	0.32	0.65	2.55	18.13
5 ARIMA BU	0.00	0.00	0.00	0.00	0.00	0.10	0.21	0.72	2.65	17.66
6 ARIMA COMB	0.00	0.01	0.02	0.04	0.06	0.10	0.17	0.68	2.52	16.54
7 ARIMA MO	0.00	0.00	0.00	0.00	0.00	0.08	0.19	0.74	2.63	17.96
8 ARIMA TD	0.00	0.03	0.05	0.07	0.10	0.15	0.23	0.67	2.64	18.18
9 LGBM BASKET BU	0.00	0.01	0.03	0.08	0.13	0.21	0.36	0.81	2.54	18.24
10 LGBM BASKET COMB	0.00	0.02	0.06	0.09	0.14	0.21	0.34	0.72	2.51	18.20
11 LGBM BASKET MO	0.00	0.01	0.04	0.07	0.12	0.21	0.36	0.87	2.50	18.23
12 LGBM BASKET TD	0.00	0.05	0.08	0.13	0.17	0.25	0.36	0.70	2.55	18.12
13 LGBM BU	0.00	0.01	0.03	0.08	0.14	0.23	0.40	0.99	2.54	18.24
14 LGBM COMB	0.00	0.02	0.06	0.09	0.15	0.23	0.37	0.77	2.52	18.20
15 LGBM MO	0.00	0.00	0.03	0.07	0.14	0.25	0.44	0.97	2.56	18.23
16 LGBM TD	0.00	0.05	0.08	0.12	0.17	0.25	0.35	0.69	2.53	18.12

Table 4.9: Store level accuracy percentile-Lumpy-RMSSE

model	csl= 90%			csl= 95%			csl= 99%		
	lost sales	csl	S.S	lost sales	csl	S.S	lost sales	csl	S.S
ARIMA BASKET TD	1.12	97	0.74	0.69	98	1.20	0.21	99	2.34
ARIMA TD	1.83	96	<b>0.47</b>	1.25	97	<b>0.81</b>	0.36	98	1.76
LGBM BASKET TD	<b>0.89</b>	98	0.82	0.54	98	1.28	0.16	99	2.47
LGBM TD	0.98	97	0.82	0.61	98	1.29	0.19	99	2.48
ARIMA BASKET MO	1.06	98	0.59	0.76	98	0.93	0.38	99	1.80
ARIMA MO	1.01	98	0.68	0.64	98	1.10	0.20	99	2.14
LGBM BASKET MO	1.17	97	0.71	0.81	98	1.13	3.04	98	2.19
LGBM MO	1.27	97	0.54	0.99	97	0.84	0.54	0.98	<b>1.61</b>
ARIMA BASKET BU	1.16	98	0.57	0.83	98	0.92	0.52	98	1.71
ARIMA BU	1.04	98	0.69	0.67	98	1.11	0.28	99	2.12
LGBM BASKET BU	1.16	97	0.74	0.80	98	1.15	0.35	98	2.26
LGBM BU	1.25	97	0.60	0.98	97	0.94	0.51	98	1.83
ARIMA BASKET COMB	0.92	98	0.76	<b>0.53</b>	98	1.22	<b>0.14</b>	99	2.37
ARIMA COMB	1.15	98	0.69	0.69	98	1.13	0.19	99	2.27
LGBM BASKET COMB	1.02	97	0.81	0.63	98	1.28	0.19	99	2.46
LGBM COMB	0.98	97	0.81	0.62	98	1.28	0.20	99	2.45

Table 4.12: Inventory Performance - store level

We further investigate the inventory performance of the SBC in both the store and the distribution center. The results in Figure 4.7 show the ARIMA model performance on average safety stock per product on the x-axis, average lost orders per product on the y-axis, and the same for stores in 4.8. The results show the out-performance of ARIMA BASKET BU compared to other models in the lumpy and erratic categories as on the three points of cycle service level it consistently provides lower safety stock quantity and lower number of average lost sales. This finding coincides with the findings in forecasting shown in Figures 4.4 and 4.6 respectively. Lower error leads to lower safety stock and better performance on fulfillment. For the smooth and intermittent category, there is no clear winner. For example, for the smooth category in the distribution center and in the

model	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
1 ARIMA BASKET BU	0.00	0.00	0.02	0.11	0.29	0.51	0.60	0.76	1.22	1.83
2 ARIMA BASKET COMB	0.07	0.18	0.22	0.26	0.49	0.59	0.64	0.71	1.19	1.78
3 ARIMA BASKET MO	0.00	0.00	0.01	0.03	0.49	0.66	0.72	0.84	1.18	1.75
4 ARIMA BASKET TD	0.20	0.50	0.52	0.56	0.60	0.63	0.71	0.78	1.26	1.86
5 ARIMA BU	0.00	0.00	0.00	0.02	0.54	0.59	0.63	0.69	1.46	2.01
6 ARIMA COMB	0.04	0.11	0.14	0.17	0.52	0.59	0.62	0.68	1.32	2.04
7 ARIMA MO	0.00	0.00	0.00	0.01	0.59	0.63	0.69	0.80	1.32	2.04
8 ARIMA TD	0.13	0.34	0.41	0.48	0.51	0.60	0.66	0.74	1.45	2.08
9 LGBM BASKET BU	0.01	0.03	0.05	0.51	0.63	0.66	0.81	1.06	1.36	2.02
10 LGBM BASKET COMB	0.09	0.20	0.26	0.52	0.59	0.66	0.75	0.90	1.30	1.97
11 LGBM BASKET MO	0.01	0.01	0.07	0.54	0.58	0.68	0.74	0.92	1.37	2.00
12 LGBM BASKET TD	0.23	0.58	0.61	0.66	0.70	0.75	0.81	0.88	1.28	1.93
13 LGBM BU	0.00	0.01	0.02	0.56	0.65	0.70	0.81	1.18	1.35	1.91
14 LGBM COMB	0.07	0.19	0.25	0.51	0.59	0.62	0.74	1.09	1.32	1.86
15 LGBM MO	0.00	0.00	0.02	0.55	0.60	0.68	0.72	1.18	1.37	1.99
16 LGBM TD	0.22	0.56	0.58	0.64	0.65	0.69	0.79	0.87	1.26	1.86

Table 4.10: Store level accuracy percentile-Erratic-RMSSE

store, the best model by a very small margin is ARIMA BASKET MO at one unit less in shortage on average.

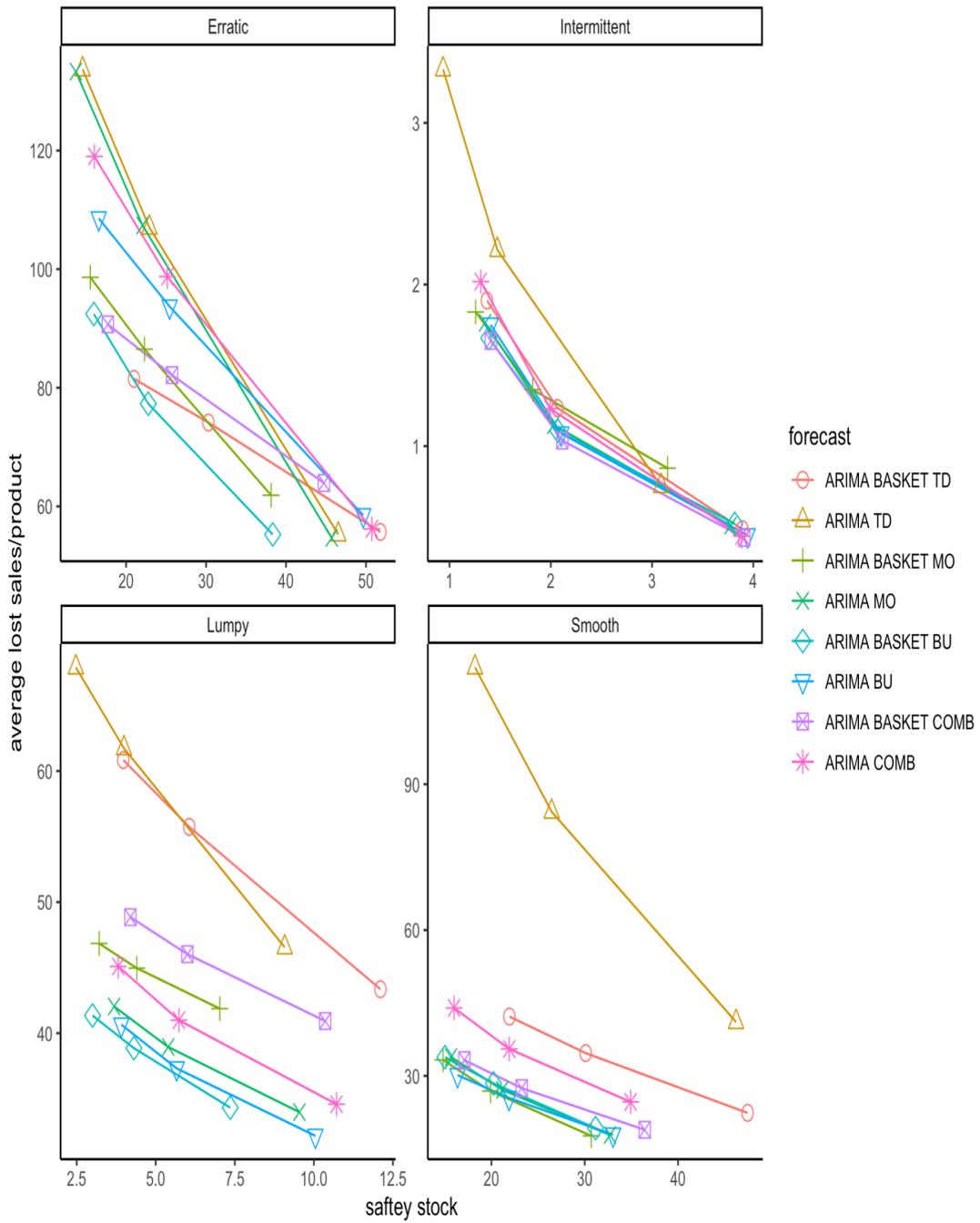


Figure 4.7: Average lost sales vs average safety stock- Distribution Centre

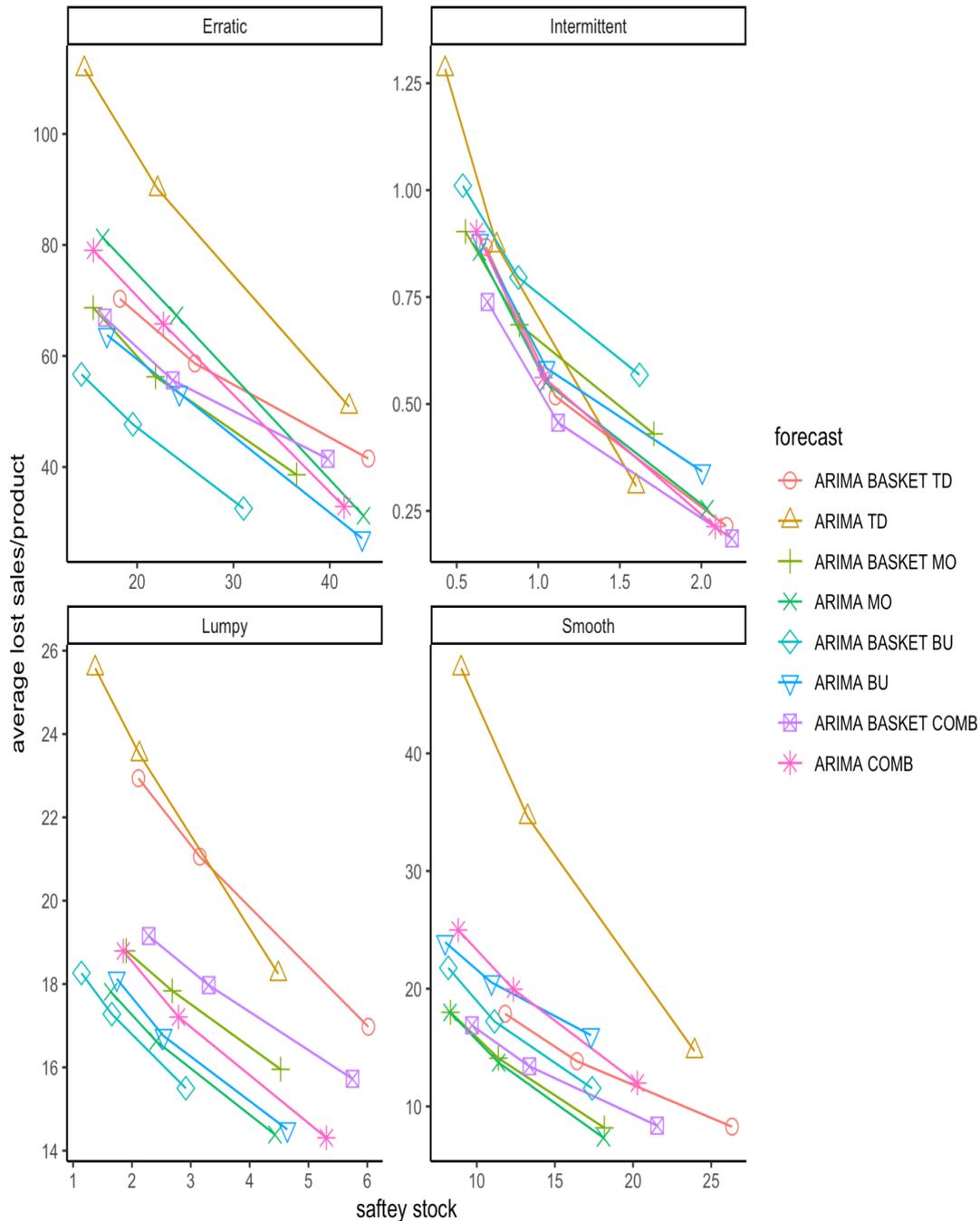


Figure 4.8: Average lost sales vs average safety stock-Store

### 4.5.1 Outliers

As noted in the Empirical investigation, outliers are removed from the inventory performance. Upon investigation, it was concluded that high loss of sales is directly related to the forecast error, especially in the lumpy category as high errors affect its performance. Upon investigation, the outliers show time series with high interrupted demand that occurs in the test period or near the test period which leads to models not efficiently trained on data, This leads to high errors in the basket forecasts. In Figure 4.9, the time series on the left are samples with high errors in the lumpy category, it can be seen there is not enough data to be trained while the samples on the right are samples within the 90 percentile of errors which leads to improved forecasting errors and effectively better inventory performance. Finally, we present the outlier's inventory performance in 4.10. Note the high level of lost orders and safety stock, for those products, we propose ARIMA

COMB.

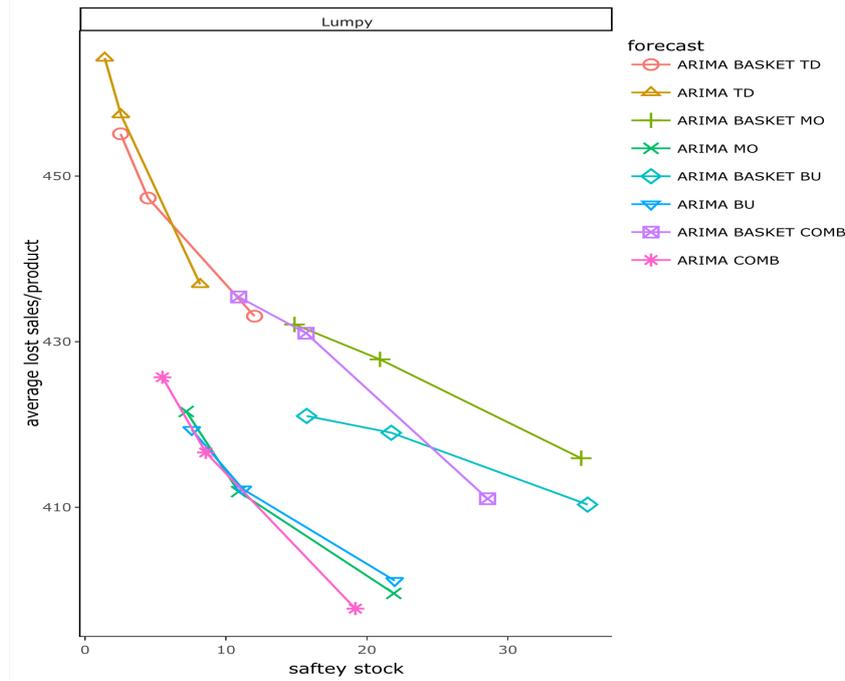


Figure 4.10: Inventory performance of outliers

## 4.6 Conclusion

In this work, we have investigated the impact of forecasting aggregation on supply chain inventory. The case was focused on a retail supply chain in which the sales data of 2018 in one city was used to simulate the physical flow of products from OEM to DC in the city and from DC to three stores based on an order up-to-level policy. In addition, Basket data attributes proposed in [Omar et al. \(2022\)](#) are added to both the ARIMA model and LGBM model with four forecasting aggregation approaches; mainly middle-out, bottom-up, top-down, and COMB. The data included 4000 products that are mostly sold in all three stores at least once and are delivered from the DC to the stores. In the middle level at the DC (Region/Product), the basket data are calculated based on the entire basket orders of the region while in the top non-grouped level which is a single time series, the basket data attributes are the result of the summation of the regional basket attributes. The level forecasting result shows a better performance for the basket attributes for the ARIMA models but not for the LGBM models. The best model in terms of MASE is the ARIMA BASKET COMB, and the best model in terms of RMSE is the LGBM TD. For Mid-level (Region/Product) accuracy, ARIMA COMB and ARIMA MO are best in terms of accuracy. However, the average measure tends to be skewed by high error values. If we look at the error quartiles, the results show better performance for basket models with the model having the lowest median in terms of RMSSE being the ARIMA BASKET TD. As for the SBC classification, the ARIMA BASKET variations are outperforming the ARIMA variations across TD, BU, and MO with a significant difference in the range of the distribution. Especially ARIMA BASKET MO and ARIMA BASKET BU for the lumpy category and a slightly better margin of improvement with ARIMA BASKET TD in the smooth category. For the Bottom store level, the best-performing model is ARIMA BASKET TD, based on error distribution. For SBC, ARIMA BASKET BU is much better performing in the lumpy category and also the model with the lowest median in the erratic category, As for intermittent and smooth, there is no clear difference in

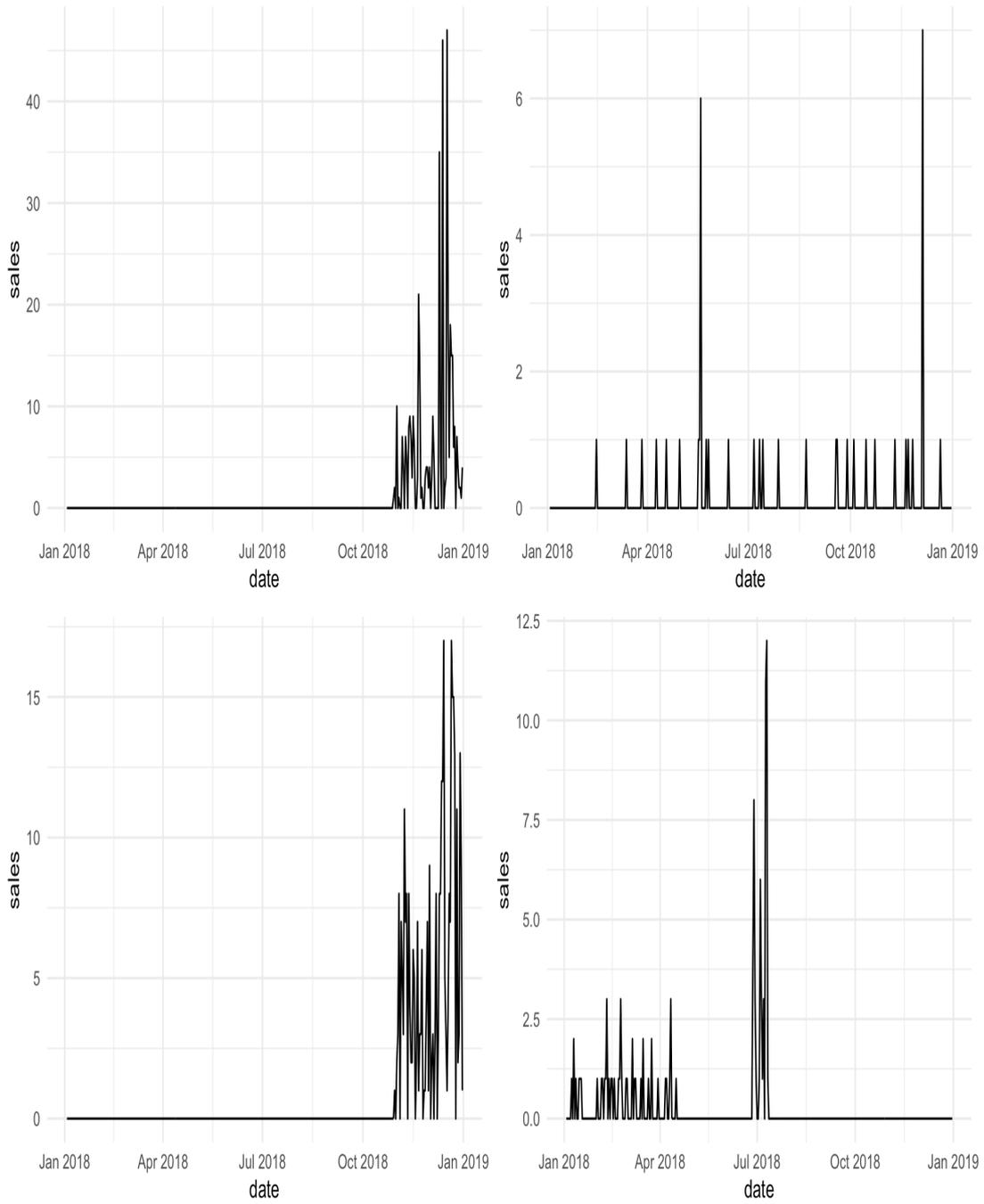


Figure 4.9: Outliers in lumpy Vs standard lumpy class

which model is a better performing model. The inventory results coincide well with the forecasting results, especially for safety stock volume against the lost sales, especially for the Lumpy and erratic categories where ARIMA BASKET BU is the model holding the lowest safety stock quantity and the lowest lost sales on average. In addition, ARIMA BASKET COMB performs well in the intermittent category. We, thus, recommend using basket attributes for low-level hierarchy and mid-level hierarchy.

# Conclusion and Perspectives

In this dissertation, we have focused on the utilization of basket data for the omnichannel supply chain. We have found that basket data leads to significant improvement in operational performance and customer satisfaction which are utilized beyond the current application of basket data for marketing and product recommendations. Furthermore, modeling basket data in a multi-period graph enables quantifying significant attributes of connectivity and relations among products more comprehensively than approaches such as market basket analysis and apriori algorithm. Basket data can have significant improvements on the allocation of products to stores which we did not discuss in detail in this dissertation and could be an avenue that could be explored in future research. Improving the forecasting can be a result also of basket data associations. In this dissertation, we have used basket data to improve the forecasting of individual products either through direct forecasting or by form of aggregation. And in both cases, it proved to be significantly better in the lumpy and erratic category. We have also considered using basket data and graph theory in omnichannel anticipatory shipping with promising results on both service and cost. Our contribution from Chapter 2 through Chapter 4 can be concluded as follows:

- For chapter 2, our investigation in this chapter has shown that 95 % of products show an intermittent or lumpy demand pattern and that 50 % of products are sold in a basket of two items or more, we have also demonstrated that a retail network is characterized by being a low-density network which proves that select products have high connectivity and are usually bought with other products many times while the remaining products have very low connectivity. We have also empirically through regression modeling proved the significance of the basket data attributes to demand prediction. Accordingly, We have developed a basket data-driven forecasting approach to improve the accuracy of forecasting and inventory performance based on the analysis of the significance of these attributes. Upon evaluating the performance of forecasting models, the Croston method which is known to work well for intermittent demand exhibited poor performance due to the lumpiness of demand in the retail context, contrary to our proposed forecasting approach which improves the forecasting accuracy as it takes into account the lumpiness of demand in the retail context. Finally, our investigation also showed that our proposed forecasting approach improves the accuracy of forecasts in lumpy and erratic demand and that forecasting both the online and physical retail channels together leads to better forecasting accuracy and inventory performance.
- In chapter 3, we have developed a two-step approach for anticipatory shipping, It is complicated to have the order available to the customer beforehand as a customer's basket may include different quantities of different products, one practical approach for such complexity is modeling basket data as a graph and deriving attributes and predictions from them. The basket with the lowest uncertainty is a basket of one item, then a basket of two items, and then more than two items will have much higher complexity. We tackled this challenge by studying the associations between

products separately and collectively using arc analysis and attribute extraction using the rich literature on graph theory. Through the nodes to the arc attributes and the arc attributes as these features concern, every two specific products and their sales in a specific historical time period features such as degree, strength, edge betweenness, and frequency of the arc. And collectively through arc probability and Bayes attributes which look at every two products in relation to the whole graph or the whole baskets ordered during a specific time horizon. This approach is flexible to adjust the graph span based on industry and what works for each assortment seasonal life span. It also shows quantity prediction as a continuous variable which is more flexible and realistic than classification prediction in this context. Furthermore, the same attributes are used to predict two targets; baskets of size 2 and baskets of size more than 2 reducing redundancy and duplication. While a basket of size 2 has only one arc and does not need redundancy removal or basket reshaping. Baskets of size more than 2 are passed through a basket retrieval algorithm to retrieve the baskets from arcs. The input of this 2-step approach is used to develop a strategy for efficient moving of online sales through the stores ahead of customer orders to reduce lead time, save on transportation costs, and packaging, and increase revenue and competitiveness utilizing basket data. Forecasting the arcs of baskets of size 2 separately provides more certainty on the content of the basket, and further action of packaging and fulfillment could be done at the fulfillment center where economies of scale are achieved (STS). As for baskets of size more than 2, advancing the baskets not packaged and as predicted modular bundles mitigates the uncertainty, reduces the transportation cost, and offers more flexibility in satisfying a customer demand from advanced modular baskets (SFS). We have shown that specific models can have up to 100 % fulfillment of baskets advanced which is a good strategy if the basket is of high value. Or retailers can choose less conservative models that advance More baskets with higher left-overs quantity if the chance of selling this product as a leftover is high (fast-moving items).

- As for Chapter 4, we have worked on incorporating basket data into hierarchical forecasting approaches. Basket data is an important dimension that if added to hierarchical forecasting approaches will result in improvement in the forecasting accuracy across the different hierarchy levels and also for inventory performance. Our investigation has shown that ARIMA BASKET BU is superior in the Lumpy and erratic category both in forecasting and inventory performance. However, Basket forecasting approaches do not result in good performance for outliers. Outliers are products that are only sold during the test period and with no sales during the training period. for these products, a traditional ARIMA model will work better. Overall, in terms of inventory performance, ARIMA BASKET COMB works best for intermittent demand, ARIMA MO works best for smooth demand, and by a higher margin, ARIMA BASKET BU works best for lumpy and erratic demand.

To extend this dissertation, some interesting avenues for further research could be considered:

- We have used basket data to improve forecasting in an omnichannel context and advancement of products close to customers mainly through developing graphs of products based on historical sales. Another aspect could be retail assortment allocation inside stores based on graph attributes. In addition, traditionally product recommendations use market basket analysis while graph theory provides richer attributes and more clearer association of bi-product links and product communities. This could be interesting to investigate and compare with market basket analysis.

- Another interesting avenue for further research would be to deepen the analysis using findings from marketing and graph theory to identify other attributes (ex: brand loyalty, promotions, etc.) than those that can be considered for forecasting purposes in the omnichannel context. Such attributes can also be considered within machine learning approaches, which have been recommended in the literature for their superior performance when using data from exogenous variables. More specifically, machine learning methods that have a graph-like structure, such as random forest or neural networks could be worthy of exploring for basket data as well as gradient boosting methods such as xgboost and light GBM which have been proven to provide good forecasting results in forecasting competitions (Makridakis et al. 2020). It would also be interesting to broaden the empirical investigation of forecast accuracy by testing other forecast horizons, including a horizon equal to the replenishment lead time. This can reduce the discrepancy between the forecast accuracy and the inventory performance of the analyzed methods. Another extension of this work would be to investigate alternative inventory rationing and fulfillment scenarios that integrate joint forecasting. It is worth pointing out that the data set considered in the empirical investigation in this chapter is rich enough to be representative of general contexts; however, for a generalization of the results and to draw more conclusive findings, it would be interesting to broaden the empirical investigation using other demand data sets and theoretically generated data. Finally, another avenue for future research could be to study the impact of product assortments at stores on customer channel choice using the linkages between sold product baskets.
- In chapter 3, we have explored basket data to advance baskets either packaged when they are baskets of size 2 and not packaged when they are baskets of size more than 2. We relied on machine learning methods and a conservation threshold that converts the continuous prediction of the arc to an integer if above or equal to the conservation threshold and zero if less, it could an area to explore is how to optimize the value of the conservation threshold through an algorithm to minimize the arcs being advanced and not sold at the week of advancement. Another aspect we explored is advancing online orders from fulfillment center to store in anticipation of customer orders, the same methodology could be explored with moving baskets between stores in anticipation of customer orders as it can further reduce transportation costs and obsolescence of products not sold. This would be based on basket attributes in every demand zone.
- Applying forecast reconciliation on the hierarchical aggregation of basket data, forecasting reconciliation has been shown to improve the hierarchy accuracy performance as it calculates the error of forecasting on each level and adjusts the forecast based on the error of forecasting at each level. Such an approach could potentially further improve the accuracy of the forecast using basket data attributes.
- Evaluating other graph attributes that can potentially improve the accuracy further for the product prediction approach, for example, we explored the notion of substitutability which measures how likely a customer would buy a product based on the range of products in this category, a complement to it could be the contribution of product in sales to all products in the range of its category.



# Bibliography

- Abolghasemi, M. and Esmailbeigi, R. (2021). State-of-the-art predictive and prescriptive analytics for ieeecis 3rd technical challenge. *arXiv preprint arXiv:2112.03595*.
- Ahmad, M. W., Reynolds, J., and Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal of cleaner production*, 203:810–821.
- Ailawadi, K. L. and Farris, P. W. (2017). Managing multi-and omni-channel distribution: metrics and research directions. *Journal of retailing*, 93(1):120–135.
- Ali, M. M., Babai, M. Z., Boylan, J. E., and Syntetos, A. A. (2017). Supply chain forecasting when information is not shared. *European Journal of Operational Research*, 260(3):984–994.
- Alptekinoglu, A. and Tang, C. S. (2005). A model for analyzing multi-channel distribution systems. *European Journal of Operational Research*, 163(3):802–824.
- Alroomi, A., Karamatzanis, G., Nikolopoulos, K., Tilba, A., and Xiao, S. (2022). Fathoming empirical forecasting competitions’ winners. *International Journal of Forecasting*, 38(4):1519–1525.
- Armstrong, C. (2016). Omnichannel retailing and demand planning. *The Journal of Business Forecasting*, 35(4):10.
- Arslan, A. N., Klibi, W., and Montreuil, B. (2021). Distribution network deployment for omnichannel retailing. *European Journal of Operational Research*, 294(3):1042–1058.
- Asmare, A. and Zewdie, S. (2022). Omnichannel retailing strategy: a systematic review. *The International Review of Retail, Distribution and Consumer Research*, 32(1):59–79.
- Athanasopoulos, G., Ahmed, R. A., and Hyndman, R. J. (2009). Hierarchical forecasts for australian domestic tourism. *International Journal of Forecasting*, 25(1):146–166.
- Atkinson, A. C., Riani, M., and Corbellini, A. (2021). The box–cox transformation: Review and extensions.
- Babai, M., Tsadiras, A., and Papadopoulos, C. (2020). On the empirical performance of some new neural network methods for forecasting intermittent demand. *IMA Journal of Management Mathematics*, 31(3):281–305.
- Babai, M. Z., Ali, M. M., Boylan, J. E., and Syntetos, A. A. (2013). Forecasting and inventory performance in a two-stage supply chain with arima (0, 1, 1) demand: Theory and empirical analysis. *International Journal of Production Economics*, 143(2):463–471.
- Babai, M. Z., Ali, M. M., and Nikolopoulos, K. (2012). Impact of temporal aggregation on stock control performance of intermittent demand estimators: Empirical analysis. *Omega*, 40(6):713–721.
- Babai, M. Z., Boylan, J. E., and Rostami-Tabar, B. (2022). Demand forecasting in supply chains: a review of aggregation and hierarchical approaches. *International Journal of Production Research*, 60(1):324–348.
- Babai, M. Z. and Dallery, Y. (2009). Dynamic versus static control policies in single stage production-inventory systems. *International Journal of Production Research*, 47(2):415–433.
- Babai, M. Z., Dallery, Y., Boubaker, S., and Kalai, R. (2019). A new method to forecast intermittent demand in the presence of inventory obsolescence. *International Journal of Production Economics*, 209:30–41.

- Babai, M. Z., Syntetos, A., and Teunter, R. (2014). Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. *International Journal of Production Economics*, 157:212–219.
- Ballou, R. (2004). Business logistics-supply chain management: planning, organizing and controlling.
- Barker, J. (2020). Machine learning in m4: What makes a good unstructured model? *International Journal of Forecasting*, 36(1):150–155.
- Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the national academy of sciences*, 101(11):3747–3752.
- Barron, A. R. (1994). Approximation and estimation bounds for artificial neural networks. *Machine learning*, 14:115–133.
- Bayram, A. and Cesaret, B. (2021). Order fulfillment policies for ship-from-store implementation in omni-channel retailing. *European Journal of Operational Research*, 294(3):987–1002.
- Bell, D. R., Gallino, S., and Moreno, A. (2014). How to win in an omnichannel world. *MIT Sloan Management Review*, 56(1):45.
- Bell, D. R., Gallino, S., and Moreno, A. (2018). Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*, 64(4):1629–1651.
- Bensinger, G. (2014). Amazon wants to ship your package before you buy it. *The Wall Street Journal*, 17.
- Berry, M. J. and Linoff, G. S. (2004). *Data mining techniques: for marketing, sales, and customer relationship management*. John Wiley & Sons.
- Bianco, A. M., Garcia Ben, M., Martinez, E., and Yohai, V. J. (2001). Outlier detection in regression models with arima errors using robust estimates. *Journal of Forecasting*, 20(8):565–579.
- Bijmolt, T. H., Broekhuis, M., De Leeuw, S., Hirche, C., Rooderkerk, R. P., Sousa, R., and Zhu, S. X. (2019). Challenges at the marketing–operations interface in omni-channel retail environments. *Journal of Business Research*.
- Bijmolt, T. H., Broekhuis, M., De Leeuw, S., Hirche, C., Rooderkerk, R. P., Sousa, R., and Zhu, S. X. (2021). Challenges at the marketing–operations interface in omni-channel retail environments. *Journal of Business Research*, 122:864–874.
- Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.
- Boone, T., Ganeshan, R., Jain, A., and Sanders, N. R. (2019). Forecasting sales in the supply chain: Consumer analytics in the big data era. *International Journal of Forecasting*, 35(1):170–180.
- Box, G. E., Jenkins, G. M., and Reinsel, G. C. (1994). Time series analysis, forecasting and control. englewood cliffs.
- Boylan, J. E., Syntetos, A. A., et al. (2006). Accuracy and accuracy-implication metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting*, 4:39–42.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177.
- Bretthauer, K. M., Mahar, S., and Venakataramanan, M. (2010). Inventory and distribution strategies for retail/e-tail organizations. *Computers & Industrial Engineering*, 58(1):119–132.
- Brynjolfsson, E., Hu, Y. J., and Rahman, M. S. (2013). Competing in the age of omnichannel retailing. *MIT sloan management Review*.
- Byrne, T. M. M. (2016). Omnichannel: How will it impact retail forecasting and planning processes? *The Journal of Business Forecasting*, 35(4):4.

- Cachon, G. P. and Kök, A. G. (2007). Category management and coordination in retail assortment planning in the presence of basket shopping consumers. *Management Science*, 53(6):934–951.
- Cai, Y.-J. and Lo, C. K. (2020). Omni-channel management in the new retailing era: A systematic review and future research agenda. *International Journal of Production Economics*, 229:107729.
- Chandna, V. and Salimath, M. S. (2018). Peer-to-peer selling in online platforms: A salient business model for virtual entrepreneurship. *Journal of Business Research*, 84:162–174.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.
- Cheng, Y., Li, B., and Li, Z. (2016). Analysis of the optimal threshold policy of the e-tailer with mixture strategy in e-fulfillment. *International Journal of Information Systems and Supply Chain Management (IJISSCM)*, 9(2):21–34.
- Chopra, S. (2018). The evolution of omni-channel retailing and its impact on supply chains. *Transportation Research Procedia*, 30:4–13.
- Chopra, S. and Meindl, P. (2016). Global edition.
- Ciulla, G. and D’Amico, A. (2019). Building energy performance forecasting: A multiple linear regression approach. *Applied Energy*, 253:113500.
- Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Journal of the Operational Research Society*, 23(3):289–303.
- de Leeuw, S., Minguela-Rata, B., Sabet, E., Boter, J., and Sigurðardóttir, R. (2016). Trade-offs in managing commercial consumer returns for online apparel retail. *International Journal of Operations & Production Management*, 36(6):710–731.
- Dooley, K. J., Pathak, S. D., Kull, T. J., Wu, Z., Johnson, J., and Rabinovich, E. (2019). Process network modularity, commonality, and greenhouse gas emissions. *Journal of Operations Management*, 65(2):93–113.
- Durlauf, S. and Blume, L. (2016). *Macroeconometrics and time series analysis*. Springer.
- Elliott, G., Gargano, A., and Timmermann, A. (2015). Complete subset regressions with large-dimensional sets of predictors. *Journal of Economic Dynamics and Control*, 54:86–110.
- Eppstein, D., Löffler, M., and Strash, D. (2010). Listing all maximal cliques in sparse graphs in near-optimal time. *arXiv preprint arXiv:1006.5440*.
- Farah, M. F., Ramadan, Z., and Kanso, J. (2022). Satisfying the online food crave: The case of online food aggregators. *International Journal of Consumer Studies*, 46(4):1413–1427.
- Fildes, R., Ma, S., and Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*.
- Fisher, M. L., Gallino, S., and Xu, J. J. (2019). The value of rapid delivery in omnichannel retailing. *Journal of Marketing Research*, 56(5):732–748.
- Fliedner, G. (1999). An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation. *Computers & operations research*, 26(10-11):1133–1149.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Gallino, S. and Moreno, A. (2014). Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information. *Management Science*, 60(6):1434–1451.
- Gallino, S., Moreno, A., and Rooderkerk, R. P. (2019). Omnichannel fulfillment dilemmas: Customer preferences and manager perceptions. *Available at SSRN 3399664*.
- Gallino, S., Moreno, A., and Stamatopoulos, I. (2016). Channel integration, sales dispersion, and inventory management. *Management Science*, 63(9):2813–2831.

- Gallino, S., Moreno, A., and Stamatopoulos, I. (2017). Channel integration, sales dispersion, and inventory management. *Management Science*, 63(9):2813–2831.
- Gamberini, R., Lolli, F., Rimini, B., and Sgarbossa, F. (2010). Forecasting of sporadic demand patterns with seasonality and trend components: an empirical comparison between holt-winters and (s) arima methods. *Mathematical Problems in Engineering*, 2010.
- Gao, F. and Su, X. (2016). Omnichannel retail operations with buy-online-and-pick-up-in-store. *Management Science*, 63(8):2478–2492.
- Gao, F. and Su, X. (2019). New functions of physical stores in the age of omnichannel retailing. In *Operations in an Omnichannel World*, pages 35–50. Springer.
- Gardner Jr, E. S. (1985). Exponential smoothing: The state of the art. *Journal of forecasting*, 4(1):1–28.
- Gardner Jr, E. S. and McKenzie, E. (1985). Forecasting trends in time series. *Management science*, 31(10):1237–1246.
- Gauri, D. K., Jindal, R. P., Ratchford, B., Fox, E., Bhatnagar, A., Pandey, A., Navallo, J. R., Fogarty, J., Carr, S., and Howerton, E. (2021). Evolution of retail formats: Past, present, and future. *Journal of Retailing*, 97(1):42–61.
- Ghoshal, A. and Sarkar, S. (2014). Association rules for recommendations with multiple items. *INFORMS Journal on Computing*, 26(3):433–448.
- Gilbert, K. (2005). An arima supply chain model. *Management Science*, 51(2):305–310.
- Goedhart, J., Haijema, R., Akkerman, R., and de Leeuw, S. (2023). Replenishment and fulfillment decisions for stores in an omni-channel retail network. *European Journal of Operational Research*.
- Goic, M. and Olivares, M. (2019). Omnichannel analytics. In *Operations in an Omnichannel World*, pages 115–150. Springer.
- Govindarajan, A., Sinha, A., and Uichanco, J. (2018). Joint inventory and fulfillment decisions for omnichannel retail networks. *Working Paper Ross School of Business*, (1341).
- Graves, S. C. (1999). A single-item inventory model for a nonstationary demand process. *Manufacturing & Service Operations Management*, 1(1):50–61.
- Gross, C. W. and Sohl, J. E. (1990). Disaggregation methods to expedite product line forecasting. *Journal of forecasting*, 9(3):233–254.
- Gross, J. L. and Yellen, J. (2005). *Graph theory and its applications*. CRC press.
- Guerrero-Lorente, J., Gabor, A. F., and Ponce-Cueto, E. (2020). Omnichannel logistics network design with integrated customer preference for deliveries and returns. *Computers & Industrial Engineering*, 144:106433.
- Harlam, B. A. and Lodish, L. M. (1995). Modeling consumers’ choices of multiple items. *Journal of Marketing Research*, 32(4):404–418.
- Hasni, M., Aguir, M., Babai, M. Z., and Jemai, Z. (2019). On the performance of adjusted bootstrapping methods for intermittent demand forecasting. *International Journal of Production Economics*, 216:145–153.
- Hu, M., Xu, X., Xue, W., and Yang, Y. (2021). Demand pooling in omnichannel operations. *Management Science*.
- Hübner, A., Holzapfel, A., and Kuhn, H. (2016). Distribution systems in omni-channel retailing. *Business Research*, 9(2):255–296.
- Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., and Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational statistics & data analysis*, 55(9):2579–2589.
- Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688.

- Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of forecasting*, 18(3):439–454.
- Ishfaq, R., Defee, C. C., Gibson, B. J., and Raja, U. (2016). Realignment of the physical distribution process in omni-channel fulfillment. *International Journal of Physical Distribution & Logistics Management*.
- Ishfaq, R. and Raja, U. (2018). Evaluation of order fulfillment options in retail supply chains. *Decision Sciences*, 49(3):487–521.
- Jasin, S., Sinha, A., and Uichanco, J. (2019). Omnichannel operations: Challenges, opportunities, and models. In *Operations in an Omnichannel World*, pages 15–34. Springer.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- Kim, H. K., Kim, J. K., and Chen, Q. Y. (2012). A product network analysis for extending the market basket analysis. *Expert Systems with Applications*, 39(8):7403–7410.
- Kim, Y., Chen, Y.-S., and Linderman, K. (2015). Supply network disruption and resilience: A network structural perspective. *Journal of operations Management*, 33:43–59.
- Koop, G. M. (2013). Forecasting with medium and large bayesian vars. *Journal of Applied Econometrics*, 28(2):177–203.
- Kourentzes, N., Rostami-Tabar, B., and Barrow, D. K. (2017). Demand forecasting by temporal aggregation: Using optimal or multiple aggregation levels? *Journal of Business Research*, 78:1–9.
- Lainder, A. D. and Wolfinger, R. D. (2022). Forecasting with gradient boosted trees: augmentation, tuning, and cross-validation strategies: Winning solution to the m5 uncertainty competition. *International Journal of Forecasting*, 38(4):1426–1433.
- Lee, C. (2017). A ga-based optimisation model for big data analytics supporting anticipatory shipping in retail 4.0. *International Journal of Production Research*, 55(2):593–605.
- Lee, H. L., Padmanabhan, V., and Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management science*, 43(4):546–558.
- Lim, S. F. W., Jin, X., and Srari, J. S. (2018). Consumer-driven e-commerce: A literature review, design framework, and research agenda on last-mile logistics models. *International Journal of Physical Distribution & Logistics Management*, 48(3):308–332.
- Ma, S. and Fildes, R. (2021). Retail sales forecasting with meta-learning. *European Journal of Operational Research*, 288(1):111–128.
- Ma, S., Fildes, R., and Huang, T. (2016). Demand forecasting with high dimensional data: The case of sku retail sales forecasting with intra-and inter-category promotional information. *European Journal of Operational Research*, 249(1):245–257.
- Ma, Y., Seetharaman, P., and Narasimhan, C. (2012). Modeling dependencies in brand choice outcomes across complementary categories. *Journal of Retailing*, 88(1):47–62.
- Macaulay, F. R. (1931). Introduction to "the smoothing of time series". In *The Smoothing of Time Series*, pages 17–30. NBER.
- Mahar, S., Bretthauer, K. M., and Venkataramanan, M. (2009). The value of virtual pooling in dual sales channel supply chains. *European Journal of Operational Research*, 192(2):561–575.
- Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PloS one*, 13(3):e0194889.
- Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2020). The m5 accuracy competition: Results, findings and conclusions. *Int J Forecast*.
- Manchanda, P., Ansari, A., and Gupta, S. (1999). The “shopping basket”: A model for multi-category purchase incidence decisions. *Marketing science*, 18(2):95–114.

- Marty, J. (2022). Consumer/user/customer integration in supply chain management: a review and bibliometric analysis. *Supply Chain Forum: An International Journal*, 23:181–196.
- Meixell, M. J. and Gargeya, V. B. (2005). Global supply chain design: A literature review and critique. *Transportation Research Part E: Logistics and Transportation Review*, 41(6):531–550.
- Melacini, M., Perotti, S., Rasini, M., and Tappia, E. (2018). E-fulfilment and distribution in omni-channel retailing: a systematic literature review. *International Journal of Physical Distribution & Logistics Management*, 48(4):391–414.
- Mircetic, D., Rostami-Tabar, B., Nikolicic, S., and Maslaric, M. (2022). Forecasting hierarchical time series in supply chains: an empirical investigation. *International Journal of Production Research*, 60(8):2514–2533.
- Murray, P. W., Agard, B., and Barajas, M. A. (2018). Forecast of individual customer’s demand from a large and noisy dataset. *Computers & industrial engineering*, 118:33–43.
- Newman, M. E. (2002). Assortative mixing in networks. *Physical review letters*, 89(20):208701.
- Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.
- Nikolopoulos, K. (2020). We need to talk about intermittent demand forecasting. *European Journal of Operational Research*.
- Noldus, R. and Van Mieghem, P. (2015). Assortativity in complex networks. *Journal of Complex Networks*, 3(4):507–542.
- Oliveira, J. M. and Ramos, P. (2019). Assessing the performance of hierarchical forecasting methods on the retail sector. *Entropy*, 21(4):436.
- Omar, H., Klibi, W., Babai, M. Z., and Ducq, Y. (2022). Basket data-driven approach for omnichannel demand forecasting. *International Journal of Production Economics*, page 108748.
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Taieb, S. B., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., et al. (2022). Forecasting: theory and practice. *International Journal of Forecasting*.
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., et al. (2020). Forecasting: theory and practice. *arXiv preprint arXiv:2012.03854*.
- Petropoulos, F., Fildes, R., and Goodwin, P. (2016). Do ‘big losses’ in judgmental adjustments to statistical forecasts affect experts’ behaviour? *European Journal of Operational Research*, 249(3):842–852.
- Petropoulos, F., Kourentzes, N., Nikolopoulos, K., and Siemsen, E. (2018). Judgmental selection of forecasting models. *Journal of Operations Management*, 60:34–46.
- Petropoulos, F. and Makridakis, S. (2020). The m4 competition: Bigger. stronger. better. *International Journal of Forecasting*, 36(1):3–6.
- Pöschhacker, N. and Nyckel, E.-M. (2020). Logistics of probability: Anticipatory shipping and the production of markets.
- Porras, E. and Dekker, R. (2008). An inventory control system for spare parts at a refinery: An empirical comparison of different re-order point methods. *European Journal of Operational Research*, 184(1):101–132.
- Poyil, H. O. (2019). Time series forecasting incorporating exogenous information. In *Thesis, KTH Royal Institute of Technology*. Sweden.
- Punia, S., Nikolopoulos, K., Singh, S. P., Madaan, J. K., and Litsiou, K. (2020). Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail. *International journal of production research*, 58(16):4964–4979.
- Rapach, D. E., Strauss, J. K., Tu, J., and Zhou, G. (2019). Industry return predictability: A machine learning approach. *The Journal of Financial Data Science*, 1(3):9–28.
- Reich, B. J. and Ghosh, S. K. (2019). *Bayesian statistical methods*. Chapman and Hall/CRC.

- Ren, X., Gong, Y., REKIK, Y., and Xu, X. (2022). Data-driven analysis on anticipatory shipping for pickup point inventory. *IFAC-PapersOnLine*, 55(10):714–718.
- Ren, X., Gong, Y., Rekik, Y., and Xu, X. (2023). Anticipatory shipping versus emergency shipment: data-driven optimal inventory models for online retailers. *International Journal of Production Research*, pages 1–18.
- Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., and Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, 71:804–818.
- Rooderkerk, R., de Leeuw, S., and Hübner, A. (2023). Advancing the marketing-operations interface in omnichannel retail. *Journal of Operations Management*.
- Rooderkerk, R. P. and Kök, A. G. (2019). Omnichannel assortment planning. In *Operations in an Omnichannel World*, pages 51–86. Springer.
- Rostami-Tabar, B., Babai, M. Z., Ducq, Y., and Syntetos, A. (2015). Non-stationary demand forecasting by cross-sectional aggregation. *International Journal of Production Economics*, 170:297–309.
- Rostami-Tabar, B., Babai, M. Z., Syntetos, A., and Ducq, Y. (2013). Demand forecasting by temporal aggregation. *Naval Research Logistics (NRL)*, 60(6):479–498.
- Rubin, D. B. (1981). The bayesian bootstrap. *The annals of statistics*, pages 130–134.
- Russell, G. J. and Petersen, A. (2000). Analysis of cross category dependence in market basket selection. *Journal of Retailing*, 76(3):367–392.
- Russell, G. J., Ratneshwar, S., Shocker, A. D., Bell, D., Bodapati, A., Degeratu, A., Hildebrandt, L., Kim, N., Ramaswami, S., and Shankar, V. H. (1999). Multiple-category decision-making: Review and synthesis. *Marketing Letters*, 10(3):319–332.
- Saha, K. and Bhattacharya, S. (2021). ‘buy online and pick up in-store’: Implications for the store inventory. *European Journal of Operational Research*, 294(3):906–921.
- Shang, G., McKie, E. C., Ferguson, M. E., and Galbreth, M. R. (2020). Using transactions data to improve consumer returns forecasting. *Journal of Operations Management*, 66(3):326–348.
- Shenstone, L. and Hyndman, R. J. (2005). Stochastic models underlying croston’s method for intermittent demand forecasting. *Journal of Forecasting*, 24(6):389–402.
- Sheppard, K. (2019). Forecasting with exogenous regressors. In *Arch documentation*. Sweden.
- Siddiqui, R., Azmat, M., Ahmed, S., and Kummer, S. (2022). A hybrid demand forecasting model for greater forecasting accuracy: the case of the pharmaceutical industry. *Supply Chain Forum: An International Journal*, 23:124–134.
- Spiegel, J., McKenna, M., Lakshman, G., and Nordstrom, P. (2014). Amazon us patent anticipatory shipping. *Amazon Technologies Inc*, 12.
- Steinker, S., Hoberg, K., and Thonemann, U. W. (2017). The value of weather information for e-commerce operations. *Production and Operations Management*, 26(10):1854–1874.
- Strang, R. (2013). Retail without boundaries. *Supply Chain Management Review*, 17(6).
- Svetunkov, I. and Petropoulos, F. (2018). Old dog, new tricks: a modelling view of simple moving averages. *International Journal of Production Research*, 56(18):6034–6047.
- Syntetos, A., Babai, M., Davies, J., and Stephenson, D. (2010). Forecasting and stock control: A study in a wholesaling context. *International Journal of Production Economics*, 127(1):103–111.
- Syntetos, A., Lengu, D., and Babai, M. Z. (2013). A note on the demand distributions of spare parts. *International Journal of Production Research*, 51(21):6356–6358.
- Syntetos, A. A., Babai, M. Z., and Gardner Jr, E. S. (2015). Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping. *Journal of Business Research*, 68(8):1746–1752.

- Syntetos, A. A., Babai, Z., Boylan, J. E., Kolassa, S., and Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap and the future. *European Journal of Operational Research*, 252(1):1–26.
- Syntetos, A. A. and Boylan, J. E. (2001). On the bias of intermittent demand estimates. *International journal of production economics*, 71(1-3):457–466.
- Syntetos, A. A. and Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of forecasting*, 21(2):303–314.
- Talagala, P. D., Hyndman, R. J., Smith-Miles, K., Kandanaarachchi, S., and Munoz, M. A. (2020). Anomaly detection in streaming nonstationary temporal data. *Journal of Computational and Graphical Statistics*, 29(1):13–27.
- Teunter, R. H. and Duncan, L. (2009). Forecasting intermittent demand: a comparative study. *Journal of the Operational Research Society*, 60(3):321–329.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.
- Trapero, J. R., Kourentzes, N., and Fildes, R. (2015). On the identification of sales forecasting models in the presence of promotions. *Journal of the operational Research Society*, 66(2):299–307.
- Turrini, L. and Meissner, J. (2019). Spare parts inventory management: New evidence from distribution fitting. *European Journal of Operational Research*, 273(1):118–130.
- Tyralis, H., Papacharalampous, G., and Langousis, A. (2021). Super ensemble learning for daily streamflow forecasting: Large-scale demonstration and comparison with multiple machine learning algorithms. *Neural Computing and Applications*, 33(8):3053–3068.
- Van der Auweraer, S. and Boute, R. (2019). Forecasting spare part demand using service maintenance information. *International journal of production economics*, 213:138–149.
- Viet, N. Q., Behdani, B., and Bloemhof, J. (2020). Data-driven process redesign: anticipatory shipping in agro-food supply chains. *International Journal of Production Research*, 58(5):1302–1318.
- Vogel, S. (2013). *Demand fulfillment in multi-stage customer hierarchies*. Springer Science & Business Media.
- Wang, D., Lian, H., Zheng, Y., and Li, G. (2019). High-dimensional vector autoregressive time series modeling via tensor decomposition. Available at SSRN: <https://ssrn.com/abstract=3453802> or <http://dx.doi.org/10.2139/ssrn.3453802>.
- Weatherford, L. R., Kimes, S. E., and Scott, D. A. (2001). Forecasting for hotel revenue management: Testing aggregation against disaggregation. *Cornell hotel and restaurant administration quarterly*, 42(4):53–64.
- Weingarten, J. and Spinler, S. (2021). Shortening delivery times by predicting customers’ online purchases: A case study in the fashion industry. *Information Systems Management*, 38(4):287–308.
- Wickramasuriya, S. L., Athanasopoulos, G., and Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526):804–819.
- Xiao, Y., Chen, F. Y., and Chen, J. (2009). Optimal inventory and dynamic admission policies for a retailer of seasonal products with affiliate programs and drop-shipping. *Naval Research Logistics (NRL)*, 56(4):300–317.
- Zhang, G. P. and Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European journal of operational research*, 160(2):501–514.



# TITRE : PREVISION ET GESTION DE STOCKS BASES SUR LES DONNEES DE PANIER DANS LA SUPPLY CHAIN OMNICANAL

## Résumé

La vente au détail omnicanal est devenue la stratégie permettant d'obtenir un avantage concurrentiel pour la plupart des détaillants qui s'appuient sur un réseau de magasins physiques. La pandémie de COVID-19 a accéléré la tendance avec de nouvelles préférences des consommateurs et une refonte du secteur de la vente au détail axée sur les données et les nouvelles technologies. Dans un environnement omnicanal, le client dispose de plusieurs options pour inspecter, commander et recevoir la commande via une livraison ou un retrait des produits. Avec ce parcours client ultime, tout détaillant omnicanal est particulièrement confronté à des défis en termes de décisions d'inventaire et de réapprovisionnement, de planification des transports et de prévisions, qui sont décuplés lorsque les ventes se font principalement dans des paniers et pas seulement dans des articles individuels. La thèse s'appuie sur une approche basée sur les données de panier pour répondre à certains de ces défis en matière de prévision et de planification des stocks. Premièrement, nous proposons une nouvelle approche de prévision omnicanal utilisant les données de panier qui améliore la précision des prévisions et la performance des stocks. Deuxièmement, une stratégie d'expédition anticipée est conçue et testée, qui s'appuie sur la prédiction des données du panier, ce qui permet d'améliorer les délais de livraison aux clients et une réduction des coûts. Troisièmement, nous développons et comparons la performance de plusieurs approches de prévision hiérarchique prenant en compte les données de panier. Toutes ces contributions s'appuient méthodologiquement sur des techniques avancées de prévision et d'apprentissage automatique, la théorie des graphes et la simulation des stocks. De plus, ils impliquent des données provenant d'un grand détaillant européen de l'industrie cosmétique et fournissent ainsi des recommandations managériales pour la pratique.

**Mots clés:** données de panier, prévision, expédition anticipée, modèle de profit, agrégation, théorie des graphes, omnicanal.

**TITLE: Basket Data-Driven Forecasting and Inventory Management for Omnichannel Supply chain.**

## Abstract

Omni-channel retailing has become the strategy for a competitive edge for most of the retailers relying on a network of brick-and-mortar stores. The COVID-19 pandemic accelerated the trend with new consumer preferences along with a technology-driven and data-driven redesign of the retail industry. In an omnichannel environment, the customer has several options to inspect, to order, and to receive an ordered product via a delivery or a pick up of the product. With this ultimate customer journey, any omnichannel retailer is particularly challenged in terms of inventory and replenishment decisions, transportation planning, and forecasting, which increase tenfold when sales are mainly in baskets and not only individual items. Accordingly, the thesis builds on a basket data-driven approach to answer some of these challenges when it comes to forecasting and inventory planning. First, we propose a novel omnichannel forecasting approach using basket data that provides an improvement in forecasting accuracy and inventory performance. Second, an anticipatory shipping strategy is designed and tested that builds on basket data prediction, which provides an improvement in the delivery lead time to customers and reduces costs. Third, we develop and compare the performance of several hierarchical forecasting approaches considering basket data. All these contributions, build methodologically on advanced forecasting and machine learning techniques, graph theory, and inventory simulation. In addition, they involve data from a large European retailer in the cosmetics industry and thus provide valuable insights for practice.

**Keywords:** basket data, forecasting, anticipatory shipping, profit model, aggregation, graph theory, omnichannel.