Demand Forecast for Short Life Cycle Products: Zara Case Study

by
Tatiana Bonnefoi
B.S. Mechanical Engineering, ITESM, Mexico City, 2005

Submitted to the MIT Sloan School of Management and the MIT School of Engineering in Partial Fulfillment of the Requirements for the Degrees of
Master of Business Administration
and
Master of Science in Mechanical Engineering

In conjunction with the Leaders for Global Operations Program at the Massachusetts Institute of Technology

June 2012

Signature of Author

May 11, 2012
MIT School of Engineering, MIT Sloan School of Management

Certified by

David Simchi-Levi, Thesis Supervisor
Professor of Engineering Systems Division and Department of Civil & Environmental

Certified by

Stephen C. Graves Thesis Supervisor
Professor of Management Science, Mechanical Engineering & Engineering Systems

Certified by

Warren Seering, Thesis Reader
Professor of Mechanical Engineering and Engineering Systems

Accepted by

David E. Hardt, Chair of Department Committee on Graduate Students
Professor of Mechanical Engineering

Accepted by

Maura M. Herson, Director of MBA Program
MIT Sloan School of Management

This page has been intentionally left blank.

Demand Forecast for Short Life Cycle Products: Zara Case Study
by
Tatiana Bonnefoi

Submitted to the MIT Sloan School of Management and the MIT School of Engineering on May 11, 2012
Fulfillment of the Requirements for the Degrees of
Master of Business Administration and
Master of Science in Mechanical Engineering

Abstract

The problem of optimally purchasing new products is common to many companies and industries. This thesis describes how this challenge was addressed at Zara, a leading retailer in the "fast fashion" industry. This thesis discusses the development of a methodology to optimize the purchasing process for seasonal, short life-cycle articles. The methodology includes a process to develop a point forecast of demand of new articles, the top-down forecast at the color and size level and an optimization module to produce recommendations to define the optimal quantity to purchase and the optimal origin to source from.

This thesis is the first phase of a two phases purchasing optimization process. The focus of this thesis is: a) the outline of an enhanced purchasing methodology b) the development of the most important input in the system: a point forecast of demand at the article, color, and size level, and c) the development of an IT prototype to automatically manage the purchasing methodology. The second phase of the purchasing optimization process focuses on the optimization module. The optimization module is beyond the reach of this thesis.

Thesis Supervisor: Jérémie Gallien
Associate Professor of Management Science and Operations
London Business School

Thesis Supervisor: Stephen C. Graves
Professor of Management Science, Mechanical Engineering & Engineering Systems, MIT

Thesis Supervisor: David Simchi-Levi
Professor of Engineering Systems Division and Department of Civil & Environmental, MIT

This page has been intentionally left blank.

Table of Contents

List of Equations

List of Figures

List of Tables

This page has been intentionally left blank.

Acknowledgements

I would like to thank Zara (Inditex SA Corporation) and the Leaders for Global Operations program at MIT for the opportunity to spend six months working in the fashion industry. I would also like to thank Zara for allowing me to work on a process so critical to their organization and for giving me all the support I needed to be successful.

The work presented in this thesis was made possible by the tireless efforts of the Zara Purchasing and Information Technology staff and the support of my advisors, Jeremie Gallien, Adam Mersereau, Stephen C. Graves and David Simchi-Levi.

Within Zara I would like particularly like to thank Miguel Diaz and Francisco Babio, the buyers, country managers, controllers, and IT staff without their help understanding the underlying business, the project would not have occurred.

Finally, I would like to thank my family Marco, Lourdes, and Catalina for their endless support and encouragement in following my dreams.

This page has been intentionally left blank.

1. Overview

1.1 Project overview

The operation strategy of the companies in the retail industry is delimited by their availability in maintaining the flow of raw materials and finished goods in their supply chains. The continuous operation in the retail sector also depends on the companies' ability to deliver the correct set of products, in the correct place, at the correct time to improve the sales performance in the stores. Therefore, developing and improving the companies' sales forecasting techniques is important in the retail sector. To this end, researchers have developed increasingly sophisticated forecasting techniques, believed to more accurately model the complexities of marketplace conditions[1].

This thesis is based on an internship conducted at Zara, the flagship brand of Inditex SA Corporation (2010). The main purpose of this thesis is to propose a standardized improved purchasing methodology that combines the art of buying with the science of sales forecasting. This enhanced methodology is accompanied by a set of tools that facilitates the decision making process in the purchasing departments. Previous research have consistently show that qualitative forecasting practice (art) are more widely used than quantitative forecasting techniques (science), even though there is an extensive research supporting the superiority of the quantitative forecasting methods in most situations (Dalrymple, 1987; McCarthy et al., 2006; Mentzer & Cox, 1984;Mentzer & Kahn, 1994; Sparkes & McHugh, 1984; Davis & Mentzer, 2007) they proposed a hybrid purchasing methodology that combines the best of the qualitative and quantitative forecast.

1.2 Company Overview

Inditex SA Corporation, a Spanish company based near the port of La Coruña, Spain, is the world's largest fashion retailer. It has eight store brand names: Zara, Pull & Bear, Massimo Dutti, Bershka, Stradivarius, Oysho, Zara Home, and Uterque. With operations in 82 countries, it has 5,527 stores and more than 92,000 employees. Inditex SA's unique management model, based on innovation and

---

[1] (Fildes & Hastings, 1994)

flexibility, and its vision of fashion, based on creativity and quality designs, together with the capacity to react quickly to market demands, have enabled it to enjoy rapid international expansion and an outstanding reaction to its various commercial concepts.

In May 2001, Inditex turned into a publicly traded company, valued at $8 billion (€9 billion at the time)[2]. When the public offer occurred, the company made a major effort to make itself known to investors, explaining its structure, its business model, and its organization. The investors showed their confidence by participating in a public share offering that prompted very high levels of oversubscription.

The fashion retail industry is a volatile environment in which globalization, customization, and speed are the key drivers for success. Within today's financial environment, traditional companies' valuation practices have evolved to more sophisticated and data-driven conventions. In the past, including the operational data, such as inventory data, to valuate a company was not a common practice[3]. However, investors today consider that the earnings per share movements in retail companies are strongly correlated with their operations and supply chain performance.

A main goal of this thesis is to standardize the purchasing process in all the sourcing departments of the company, with the future intention of optimizing the overall stocking level of the company. The current situation of the inventory levels at Inditex SA can be summarized as follows: the inventory turnover of the company in 2010 was 4.2 times a year, a 14% decrease compared to 2009; also in 2010 the obtained markups were 145%, an increase of 8% compared to 2009. The inventory turnover and markups combined reflect that the company made more profits per sale on average. The inventory turnover metric can be used as a performance reference of a retailer, as shown in "An Econometric

---

[2] ("Inside Zara," *Forbes Global*)
[3] Fishers and Raman. The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance. (2010)

Table 1: 2009-2010 Inditex Group Income Statement Analysis[4]

Inditex SA Consolidated Income Statement Analysis (In thousands of euros)			
	2010	**2009**	**Comparison 2010-2009**
Net Sales	€ 12,526,595	€ 11,083,514	12%
Cost of Merchandise (COGS)	€ 5,104,573	€ 4,755,505	7%
Gross Margin	€ 7,422,022	€ 6,328,009	15%
Gross Margin/ Sales	59%	57%	4%
Markup (Gross Margin/COGS)	145%	133%	8%
Net Income	€ 1,741,280	€ 1,322,137	24%
Earnings per share Euro Cents	€ 277.90	€ 211.40	24%
End Year Inventory	1,214,623	992,570	18%
Inventory Turnover (COGS/Inventory)	4.20	4.80	-14%
Total Stores	€ 5,044	€ 4,607	9%
Average Sales per store	€ 2,483	€ 2,406	3%
Property, Plant & Equipment (PPE)	€ 3,397,083	€ 3,293,535	3%
Sales/PPE	3.687	3.365	9%

Table 1 summarizes the most important financial figures from the Inditex SA income statement. From a financial perspective", Inditex SA is clearly a growing and healthy company. A 12% increase in operating revenue has combined with an extraordinary net income performance increase of 24% compared to 2009. All of these growth ratios show the financial stability of the company. One ratio that measures the operative performance of the purchasing group is the profit margin, in 2009 it was required 42% of the profits to cover the COGS, in 2010 the requirement was reduced to 40%. However, a company growing at an accelerated pace faces several operative challenges. The main operative challenges that Inditex SA is experiencing are the following: maintaining a good relationship with suppliers; creating supply chain processes with adequate inventory controls; staying current with

[4] (Inditex SA Consolidated Income Statement 2009-2010, www.inditex.com)

IT technology and document policies; sharing best practices and procedures; and, finally, sustaining the same customer focus while growing the staff inorganically[5].

1.2 Zara Overview

The Inditex SA Group is made up of more than 100 companies operating in textile design, manufacturing and distribution. Zara is the best-known brand of the group and the flagship of the company; it represents 70% of the total sales of the company and it has operations in 82 countries, with a network of more than 1,600 stores, mostly owned and operated by the group.

Their latest addition, the new store in downtown Sydney, opened in 2011, breaking the €40M record on the opening night and giving Zara exposure for the first time in Oceania. Until 2010 the only selling channel for Zara was at the retail outlets located on five continents: America, Europe, Asia, Africa and now Oceania. In September 2010, it opened up to online retail; this format is only offered in selected countries, based on the Internet penetration and distribution capabilities, and e-commerce operations in China started in September 2011. In the near future, the online business will become one of the main retailing channels for the company.[6]

The core business strategy of Zara is: "Cutting-edge Fashion at Affordable Prices,"[7] and their main operation strategy is a combination of Just-in-time manufacturing and a highly vertically integrated structure. The planning cycle for Zara, as for many other fashion retailers, starts a year in advance of the corresponding season. Zara's designing team is formed by 250 designers[8], each one of them will create over a 100 designs a year, just half of them will make it to Zara's windows, and just a few designs will become top sellers. Designing appealing garments is just one side of the successful

---

[5] (Fast Growing Companies – Challenges and Solutions, Steve Y. Lehrer, 2011)
[6] (http://www.inditex.com/en/who_we_are/timeline)
[7] (The Fast-Fashion Business Model: An Overview Based on the Zara Case, F. Caro, 2008)
[8] (Zara, Case Study. K. Ferdows, M. Lewis, and J.A.D. Machuca, 2003)

business model of Zara; after preparation of the final assortment that will fill the shelves of the store in the next season, the tangible operation of creating more than 14,000 articles a year will take place.

To offer cutting-edge products at affordable prices, Zara follows a "fast fashion" philosophy. That entails, a combination of two factors: first, a short production and distribution lead-times and second, a highly fashionable product design[9]. In addition, the firm exerts a strong control over almost all the supply chain: design, purchasing, production, distribution, and retailing.

As companies struggle to increase customer value by improving performance, many companies are turning attention to purchasing and supply management. Getting the materials and inputs services from suppliers has a major impact on companies' ability to met customers needs[10]. Purchasing is one of the main cost-savings activities in the entire supply management. The ratio COGS (5.105B €) to Net Sales (12.527B €) for Zara was 40% in 2010[11]. That is, for every Euro collected from the sales operations, forty cents go back to the suppliers. Thus, the amount of time and effort that the buyer (person in charge of the purchasing, procuring, and sourcing activities in Zara) spends bargaining over a better deal or building a stronger relation with the suppliers, impacts directly the profitability of the company.

The sourcing strategy at Zara can be subdivided into four different types. The first type is the *procurement of fabrics*. This type of procurement refers to the purchase of large volume of fabrics that are constantly used in the company. These fabrics, are commodities and their market prices commove with the oil price i.e. cotton, cold wool. The commitments for these textiles are made roughly 6-8 months prior to the delivery in stores, and they can be carried as inventory if a strong movement in price is expected. This type of sourcing strategy gives Zara a relevant economy of scales edge. In addition, by having available raw material in their inventories their production time is shorter.

---

[9] (The Value of Fast Fashion: Quick Response, Enhanced Design and Strategic Consumer Behavior, G.P. Cachon and R. Swinney, 2010)
[10] (Purchasing and Supply Chain Management, R. M. Monczka, R. B. Hanfield, and L. Giunipero, 2009)
[11] (Inditex SA Consolidated Income Statement 2009-2010, www.inditex.com)

The second sourcing strategy is *in house manufacturing.* The 35% of Zara's total production is done in the Inditex owned factories, located next to the company's headquarters. Once a design is completed, the garment's patterns cut on site. Then, all the sewing is subcontracted to a network of 400 smaller firms in the surrounding areas of Galicia. Finally, the pieces of cloth are returned to the Inditex factories to be quality-inspected and prepared for shipment. This is the most efficient sourcing method of all, from start to finish the process can take one or two weeks, assuming there is enough fabric in stock, and there are available resources at the sewing contractors. The final quality of the products is very high; the most complicated designs are done in house. Thus, this sourcing method is the most expensive of all. [12]

The third sourcing strategy is the *proximity sourcing.* Between 40-55% of the total production is sourced using the proximity strategy. Most commonly Zara procures the fabric and gives the design specifications and the sewing patterns. A closely monitored third party subcontractor cuts, sews finalizes the assembly, and the inbound transportation of the clothes. The proximity subcontractors are primarily located in Portugal, Turkey and Spain. This type of sourcing strategy will take from start to delivery four to six weeks[13].

The fourth sourcing strategy is called *long circuit.* It is a fully subcontracted operation, Zara provides the design and quality inspection, but the third party provides the fabric, patterning, sewing, and inbound transportation of the final assembly of the product. This strategy is used for high volume low fashion design articles i.e. knitwear cardigans. The subcontractors are located in Asia (China, India and Vietnam). The delivery of the final product can take up to 12 weeks and the cost per unit that can be negotiated is the lowest of all sourcing types.

The four sourcing strategies are owned and highly monitored by the different purchasing groups in the company. The first sourcing strategy looks for the most important raw material in the system, the

---

[12] (N. Fraiman, M.Singh, and C. Paris. Zara Case. Columbia Business School, 2002)
[13] (N. Fraiman, M.Singh, and C. Paris. Zara Case. Columbia Business School, 2002)

textile. This sourcing type can be combined with the other three strategies. The responsible for the fabrics sourcing are specialized textile buyers that are constantly looking for good quality/price deals. The remaining three sourcing strategies are methods to acquire final goods. The buyers make the decision of making in house (second sourcing strategy) or outsource from suppliers (third and fourth sourcing strategies). The decision is based on four criteria, ranked in order of importance: production speed, supplier experience (quality of the final product), cost-efficiency and available capacity. Depending on the product mix in the different groups, and the final products available in the stores the buyers will strategically place the final goods orders. That is, buyers will decide to make in house those designs that requires special care or that they need to be deliver into the stores very fast. But they will select to outsource the easy design and high volume products that can bring cost reductions through economies of scales. Figure 1 shows the main activities develop in the buyer's role. The figure also illustrates the trade offs between the different final goods' sourcing origins.

1. Negotiate the costs of articles
2. Visit suppliers
3. Show the possible mix of articles in the stores
4. Decide how much to buy from where

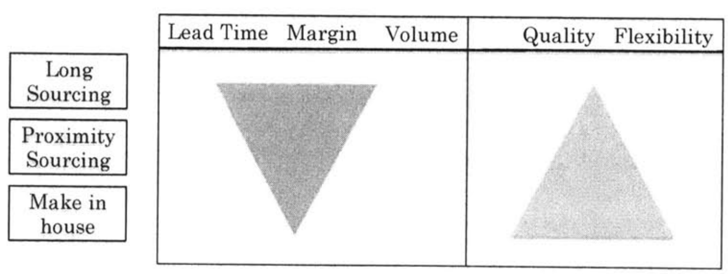


Figure 1: The role of a buyer in the organization

Zara is subdivided into three operative departments: "Children" (12% of the total sales), "Men" (15% of the total sales) and "Women" (73% of the total sales). After the products are manufactured, all the items are inbound transported and stored in one of the two Zara's Distribution Centers (DCs) (La

Coruna Spain and Zaragoza Spain). The inbound transportation methods are three: truck, air or ship. If the sourcing strategy selected by the buyer is making in house, the required inbound transportation method is truck, and the products are same-day-deliver from the local sewing subcontractors to Inditex own factories. If the buyer selected proximity sourcing method, there are two options truck or airfreight. The recommended transportation method for this sourcing origin is truck. The products that are delivered from Portugal take approximately one week to be delivered. If the product origin is Turkey or Morocco, the journey can take between two to three weeks. Finally, if the sourcing type is long circuit the two options are ship freight and airfreight. The suggested transportation method is ship and it will take between six and eight weeks. These differences in time are key in understanding the differences in Zara's product mix; high fashion items are mostly sourced from proximate suppliers or they will be produced in-house, whereas basic items (with a simple design, easy to manufacture patter and low fashion trendiness), such as a white T-shirt, will most likely come from Asia. Once the items are at the DCs, the manner through which the references get to the stores is the same regardless of where they are sourced from; references are sorted by store at the DCs and trucked to European stores shipped or air freighted to the rest of the world.[14]

Zara commits only a portion of its season inventory months in advance. [15]. Once the selling season starts, the total budget is split between different sourcing origins. The rest is in season production. [16] Most of this reminder will be produced in house, just after the ramping up collection in the stores. Figure 2 schematizes the ongoing business cycle that Zara faces after observing the customers' reactions to the first products of the collection. Figure 3 shows how the two most important business cycles in the company, buying and selling are related. Zara is an organization highly driven by sales performance. The in-season production is the most flexible asset of the company, and it allows the

[14] (Myraida Vega, 2005)
[15] (N. Fraiman, M.Singh, and C. Paris. Zara Case. Columbia Business School, 2002)
[16] (N. Fraiman, M.Singh, and C. Paris. Zara Case. Columbia Business School, 2002)

modify the assortment according to the market demand, and they will contain the profit loss of articles that might be not as successful as others; i.e., if a long sleeve T-shirt is not selling well during the summer, they might not distribute to the rest of the world until the next season. On the other hand, if a product shows a strong demand, the sourcing departments will be triggered to look for more volume of that reference, adjusting ad-hoc to stores' necessity.

Figure 2: In-season selling loop

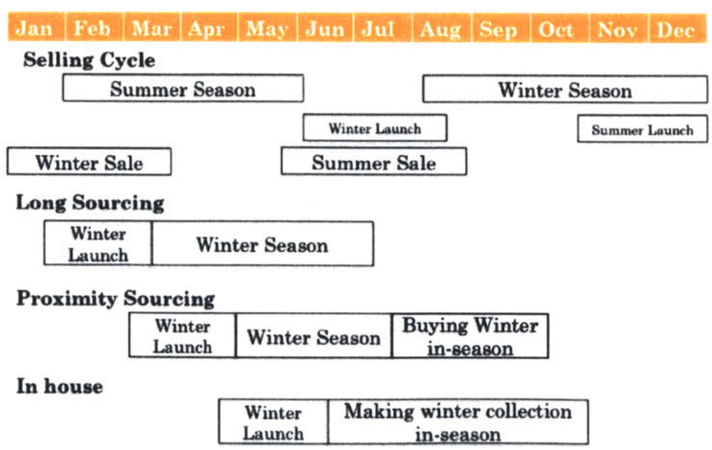


Figure 3: Selling and buying cycles

The "Projects Team" led the project on which this thesis is based; this group serves as the systems integrator in the organization. The team is closely related to the three most important functional areas in the company: operations top management (logistics and budgeting decision planning), product management (buying, assortment, and overall store performance planning), and IT systems. This project is part of a continuous improvement group of projects; the end result will be incorporated into a platform that will record the historical performance of all transactions of the purchasing department.

1.3 Chapter Summary

Inditex SA is one of the lead fashion retailers in the world. Zara, the flagship of the company, accounts for the majority of the success of the group. In the past 10 years the successful implementation of their policy of "Cutting-edge Fashion at Affordable Prices" has penetrated into many geographical regions of the world. The company is growing at an accelerated pace, and one of the biggest challenges is to keep filling the stores with the right products at the right time. The buying teams in the company are responsible for deciding where to buy, how much to buy and when to buy. The main objective of the buying groups is to maximize the limited resources: time, production capacity and budget to satisfying their customers' demands to maximize profits.

2. Project Overview

The focus of this thesis is to define a standardized purchasing methodology for retail short life-cycle articles i.e., fashion cloths. In addition, this thesis defines a process to obtain a point forecast of demand; this forecast is the main input to generate the optimal decision on how much to buy. The point forecast of demand uses historical sales data as baseline information as baseline information. The historical sales are converted into demand data by extracting out of stocks and other supply chain inefficiencies. In addition, the demand data is modified to better represent the market conditions at the moment of the purchase. Finally, this thesis proposes the system architecture to automate the purchasing methodology.

2.1 Project Background

The cumulative volume sourced and purchased of a clothing item is the bloodstream of Zara, and when Zara started operations the owner itself made the three fundamental decisions: how much to buy, where to source from and when to buy. However, as Zara matured it was impossible for one person to make all the decisions by himself, pushing the organization to create a central purchasing/sourcing department. At the time when Zara served only stores in Spain, the Inditex owed factories produced all the references, and the purchasing group was devoted to sourcing only fabrics and raw material. As Zara gradually evolved into the global retailer that is today, the use of different sourcing origins arose as a cost-reduction strategy, and now an organization of twenty buyers and eight product managers is responsible for all the sourcing activities and suppliers' evaluation to ensure the smooth flow of items to the final customer.[19] The product managers serve as the systems integrators in the company, they are responsible to work with the designers and country managers in selecting the best assortment of products to be launched in the stores, but they are also responsible for trigger the purchasing activities, right after completing the products selections the product manager will coordinate with the buyers to start the sourcing activities. In addition to the increasing headcount in the buying departments, the number of suppliers and sourcing origins has increased rapidly, and the international agreements and the specific normative specifics between suppliers and the company have evolved into very complex contracts that require legal assistance and careful thought to ensure the continuous flow of materials around the world.[20]

The global reach of Zara's has increased the company's complexity; however, in essence the main job of a buyer is the same as 25 years ago, when it was one person's responsibility. They analyze the features and specifications of the product to be buy, based on the attributes of the article the buyer select the appropriate supplier, to establish the quantity to purchase, and--more important--to bargain the cost and the terms of the delivery of the final product.

---

[19] (Inditex Timeline, 2010)

The data used in this project come from the Zara "Woman" Section; this section is subdivided into 5 different sourcing groups, and each group faces different challenges.

The different groups develop and sell different products. The most representative classification of articles can be bracketed into four main dimensions: fashion trendiness, sourcing origin, launching date, and customer target. This categorization allows the buyers to estimate the purchase quantity required to fulfill the requirements of the different purchasing departments. The different categorization of articles combined with the intrinsic characteristics of each purchasing department, creates divers purchasing options for the buyers. The buyers analyze the available sourcing options and, the intrinsic characteristics of the products to be acquired, to make rough estimates of the volumes required in each category. Associating a product with a combination of categories help the buyer estimates the potential acceptance and life of the product in the market, this two characteristics are important when defining the quantity to buy.

The estimation of the acceptance and expected life of the product is done by associating the new article with a historical article that belonged to the same set of categories. The number of combinations that can be extracted from the different classifications is vast, and each one results in a different bargaining strategy. However, some of the combinations are impossible to execute. That is, to produce in-season an article that targets fashionistas, cannot be produced at China, the article would be delivered to the stores with a delay that the fashionistas is not willing to wait and this can lead to customer dissatisfactions and in the long term to a bad brand perception. The operation of buying at Zara directly affects the bottom line of the organization, and the top management has observed that is impossible for a buyer to remember all the articles offered at Zara. In addition, the buyers are most likely to remember only successful past items. The observations reflect the opportunity to improve the process by creating a corporate tool that will help buyers decide how much to buy and will keep track of all the historical transactions of the different groups. Now is the perfect opportunity to define a standard purchasing methodology. A well-understood standard procedure can be captured in an automated tool that will relieve the buyers of the no-value-adding activities such as: the use of multiple

The estimation of the acceptance and expected life of the product is done by associating the new article with a historical article that belonged to the same set of categories. The number of combinations that can be extracted from the different classifications is vast, and each one results in a different bargaining strategy. However, some of the combinations are impossible to execute. That is, to produce in-season an article that targets fashionistas, cannot be produced at China, the article would be delivered to the stores with a delay that the fashionistas is not willing to wait and this can lead to customer dissatisfactions and in the long term to a bad brand perception. The operation of buying at Zara directly affects the bottom line of the organization, and the top management has observed that is impossible for a buyer to remember all the articles offered at Zara. In addition, the buyers are most likely to remember only successful past items. The observations reflect the opportunity to improve the process by creating a corporate tool that will help buyers decide how much to buy and will keep track of all the historical transactions of the different groups. Now is the perfect opportunity to define a standard purchasing methodology. A well-understood standard procedure can be captured in an automated tool that will relieve the buyers of the no-value-adding activities such as: the use of multiple sources of information i.e., the buyers manage several spreadsheets that contains information of articles from different categories; the use of large databases to extract aggregated numbers i.e. obtaining the aggregated demand for each store in a country requires the retrieval of several files from the database, this activity need to be done overnight to avoid overloading the database; Manual adjustments to capture changes in stores i.e., if the buyer knows that a store with large selling potential will be open soon, he or she will manually include that calculation. By eliminating these non-value-added activities the buyers will have more time focus on the negotiation strategy that directly affects the bottom-line of the company.

2.2 Problem Statement and Project Objectives

Each year Zara purchases more than 14,000 articles. For each one of them, the buyers' decision of how much to purchase has a substantial impact on sales, store assortment, and costs.

Although this task is arguably one of the company's most important operational decisions, at present buyers tend to use relatively informal guidelines for making decisions. This project looks for a more consistent, explicit, and optimum manner in which to decide on purchase volumes.

Given the rapid and inorganic growth of Zara's purchasing departments, it is necessary to capture the essential steps and sequence of the purchasing process in a standardized process. Identify a standard buying process across the organization will facilitate the communication about how the buying process operates. In addition, this standardization will help reduce the variations in the process. And once all the stakeholders in the process understand the steps and the sequence of the process, they can identify the non-value-adding activities and inefficiencies in the process, and contribute to continuously improve the buying process.[20]

This thesis discusses the results of a qualitative analysis of the different purchasing groups dynamics and proposes a standardized purchasing methodology. The decision-making process at the different purchasing groups at Zara is heterogeneous. That is, the identification, development and selection of alternatives made by the buyer are different across the purchasing departments.[21] The main reasons for these differences are: the seniority of the group of buyer, the complexity and variability inherent to each purchasing group, the set of tools available for each group, the buyers background diversity and the different accessed permissions to top-level information

Having a standardized methodology and a good understanding of the purchasing process can positively impact the following areas:

- Identify, document, and spread best purchasing decision-making practices
- Use IT to relieve buyers of time-consuming and variable manual or spreadsheet-based analysis
- Leverage IT and computing power to perform new analyses, leading to better purchasing decisions
- Help buyers identify and focus their time on the most time-sensitive and difficult decisions

[20] (T. H. Davenport. The Benefits of Business Process Standards, 2005)

[21] (S. Abbas, M. Iqbal, and A. Zaheer. The effect of workgroup heterogeneity on decision making, 2010)

work. Chapter 4 opens with the potential areas of improvement required to standardize and automate the process, followed by a description of the proposed new system architecture for the purchasing process. Chapter 5 explains the prototype architecture of a corporate tool that will automate the purchasing process. Chapter 6. Wrap up the thesis with a summary of the purchasing methodology and the conclusion.

2.4 Literature review

Numerous studies analyze and describe the optimization of different process in the retail industry. In the specific space of demand forecast in the fast fashion retail industry some authors have analyze and document their results. Namely Caro and Gallien (2007) studied the problem of optimizing the products assortment of a store as the season progress and more information is available, in this paper they propose a quantitative model to forecast demand reflecting the real situation at the market place. Garro (2011) studied how to forecast demand of new products to optimize the distribution of the assortment in the stores; he proposes a point forecast currently use at Zara's distribution and logistic department. In this document Garro also defines a clustering methodology to group stores, relevant for forecasting demand with incomplete information. Vega Gonzalez (2009) proposes an inventory management optimization model; she is proposing a prototype tool to be used at the distribution centers in Zara after the purchasing quantity has been defined by the purchasing teams. Fishers and Raman (2010) in the book The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance document and exemplify how simple analytic methods can improve the retail operations and have a major impact in the profitability of the companies. Caro and Gallien (2010) quantitatively analyze the inventory behavior of apparel short-lifecycle products including the broken assortment inventory policies at Zara. The focus of this thesis is to propose a standardized purchasing methodology for short-lifecycle and seasonal products. The main input for this methodology is a point forecast of demand, that will be used to define the optimal quantity to purchase to improve the profitability of a fast fashion company (the optimal purchasing quantity and the profitability assessment are beyond the scope of this study).

3. The current buying process

The first step of this project is to conduct a series of interviews with the buyers from the different purchasing departments. The main objective is to understand the steps, sequence, inputs and outputs of the purchasing process utilized at the different departments. At first glance, each of the departments' processes seemed to have an independent strategy. However, the desired outcome of the process was the same for the groups. The interviews were designed to highlight the similarities and differences among the groups to extract the underling methodology and the particularities for each group.

3.1 Current State

The first question asked to the interviewees was: what is the objective of your work? This question aimed to interpret the desired output of the purchasing process. What are the main deliverables, to whom do they go, and how is the process to be evaluated as successful?[24]

All the buying departments agreed that there are two main output of the purchasing process. The first and most important output of the system is the order quantity **Q**. This output answers the basic question of *how much to buy* of a given article; that is, the estimated quantity that they will need to source to cover the existing demand for the article in the stores. The estimated quantity to order of a given article is the number that the buyer will bring to the negotiation table with the supplier; it is the lever of their bargaining position. In addition, knowing the quantity to order will help the buyers decide the required capacity, the amount of fabric and raw materials that they will need to source, and, most important, the percentage of their budget that will be invested in article. They understand the cost

[24] (The high-velocity edge, Steven J. Spear)

of sourcing as an opportunity cost; once they have used a portion of the budget and the capacity, those resources cannot be use for producing other articles.

The total quantity to order will be passed to suppliers using disaggregated specifications. Therefore, the two secondary outputs of the system are the following:

- *Estimate the disaggregated quantity per color.* Once the capacity requirements are met, the buyers will need to give the supplier detailed specifications of the batch to be produced. Often, the lot to be produced requires dying the same pattern in different colors or tones. For some fabrics this operation can be performed at the end of the manufacturing process, but for many others the dying process needs to be done before cutting the fabric. The buyers are responsible to define the amount of units to order in each color.
- *Estimate the disaggregated quantity per size.* Just as with the color disaggregation, specifying a production batch requires breaking down the quantities required in each size. This disaggregation is a two-part output. First it is necessary to decide how many sizes will be offered. Depending on the cutting pattern, expected fit and targeted market, the number of offered sizes can vary from only one up to seven different sizes. The second part is to decide the exact amount required for each size.

The second output of the purchasing process is to define the *optimal sourcing strategy*; the buyers need to decide *where to buy*. The available capacity, the scalability, and the sewing skills of the different suppliers are not constant in time. As a company resource says, "The production capacity of good suppliers is one of the most scarce resources in the world,"[25] so it is necessary to book suppliers' capacity at least 6 months in advance. Therefore, having a good estimate of the required capacity will provide the buyer with more information to make the optimal selection of suppliers. A buyer is always hunting for new factories and suppliers that can accommodate their increasing requirements at affordable prices.

---

[25] (Inditex Product Management, 2010)

Generating the two principal outputs discussed above requires integrating different pieces of information. The inputs required in the process can be classified into three major groups: historical information, store conditions, and intrinsic characteristics of the new article.

Critical Inputs:

Historical information refers to information from previous seasons that can indicate the potential future performance of a new article in the store.

- *A comparable reference*. This is an item that was sold previously at the Zara stores, for which the historical performance is known and the past sales data are available in Zara's database. The comparable item is similar to the new article, and each group has a different range of attributes that make an item similar to another; however, in essence all the groups will use this information as the baseline sales prediction.

Store Conditions: relevant information about the store that can modify the required quantity.

- *The Network*. When deciding how much to buy, the buyers will consider the stores in the network that will receive the new reference. Given the characteristics of the articles (retail price, fabric, style), part of the assortment is not suitable for Zara's entire network; this input drastically changes the limits of the purchasing ranges.
- *Display in the stores*. The merchandizing specifications of a reference are important when deciding how much to buy. There are two main options: first, a reference located on the walls of the store (usually these articles will be hung and the space is limited) and second, a reference located on tables, (tables can fit a large amount of stock). To provide an attractive display in the store, the stores need a minimum *exposition stock* (inventory that will allow them to initially display the article, a minimum of one per size per color). This information will give the buyer a volumetric idea of how much space in the store a given reference needs to cover.

Some characteristics of the article can influence the order quantity. This list can be extensive but the most influential characteristics for the quantity to be order are three:

- *Expected life of the item.* The estimated life of a reference can vary depending on the introduction point in the season, i.e., a reference designed to be featured only at the ramp-up of the season will require less quantity than one intended for the entire length of the season. Some references are considered essentials, and buyers will want them to be in stores for more than one period; therefore the initial quantity will increase.
- *Reorder point option.* Along with the introduction point, some of the references will have the option of being reordered. If this is the case, buyers will tend to buy the minimum quantity to see the performance in the store and then decide if a second ordering point is necessary.
- *Trendiness.* If an article contains an attribute that follows the new trend, this can increase its demand, so this input needs to be taken into account when deciding the optimal purchase quantity. This is a subjective multiplier that the buyers will apply if they consider the product to be more desired than the rest.

Figure 4 describes the steps follow by buyer from the buying department to decide how much to buy of an article that will be part of new collection. After deciding that an article is going to be bought, the buyer will identify the specific characteristics of the design and amount of time needed for production.

Continuing with the purchasing process, the next step is to select a comparable: a reference with historical sales information and with similar attributes. At this stage it is difficult to find an article with the specific characteristics of the new model. There are two alternatives. The first is using the average total sales information of a group of articles. The second alternative is to use historical data from a single comparable reference. If the buyer selects a comparable, the total quantity sold total quantity purchase for the new article, ignoring the exposition stock required to fill the stores (potential new openings) and also ignoring that sales of the comparable reference represent only the portion of demand that Zara was able to capture; the expected quantity to be purchased for the new reference would be equal to the average of the total sales of the group of articles.

The buyers understand that the comparable reference is not always perfect; typically, the buyers will tweak the baseline quantity to account for the variability associated with the network of stores receiving the new article, but they do it without applying a standard quantitative methodology. The first source of variability is the number of stores that will receive the new article. The number of stores receiving the new article compared to the baseline reference can be different. The buyers understand that to fulfill the inventory requirements of the additional stores they need to "bump-up" to the baseline quantity. They also know that the need to order in batches rounded to the thousands. Contrary, if the network of stores receiving the new reference is smaller than the comparable, buyers would not apply any factor.

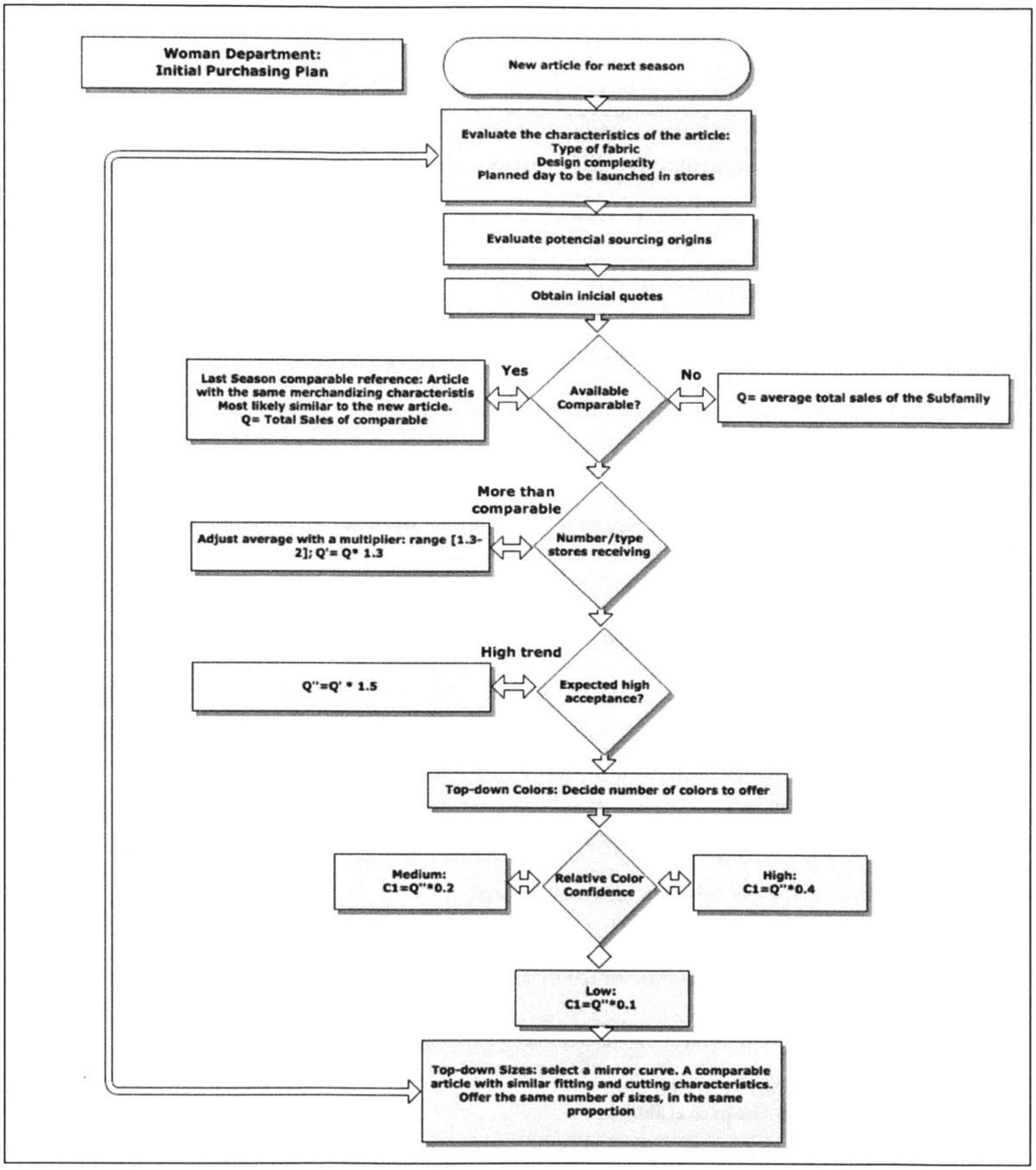
The next step is to introduce an acceptance factor; the commercial intuition and the consumer understanding of the buyers allows them to sense the expected level of acceptance of a garment, and they bet high or low for the level of acceptance in the article.

To place an order to the suppliers, a buyer must break down the total quantity Q'' into the color specifics. To make this decision, the buyer will assign a relative weight to each color. Buyers have developed a color categorization trough empirical knowledge. The buyer breaks-down the total quantity to be purchased the first break down is basic higher than non-basic. Then within the group of basic colors the buyer usually assigns equal weights to all the colors.

Finally, to select the number of sizes to offer, a "Mirror Curve" will be used. This is a table that indicates the number of sizes offered for a given cutting/volume style and the relative weight for each size. The data comes from a reference with similar fitting characteristics (it is not necessary to have the same expected selling behavior) but it was offered in previous seasons. The "Mirror Curve" is outdated. In the last year Zara's has grown mainly in the emerging markets[26]. This growth influences the sizing pattern changes by using an outdated reference these changes are not reflected. Similarly to the POS data used to estimate the purchasing quantity at the article level. The size level information by size only reflects the portion of demand that Zara was able to

[26] (http://www.inditex.com/en/shareholders_and_investors/investor_relations/annual_reports)

capture.



Woman Department:
Initial Purchasing Plan
New article for next season
Evaluate the characteristics of the article:
Type of fabric
Design complexity
Planned day to be launched in stores
Evaluate potencial sourcing origins
Obtain inicial quotes
Yes
No
Last Season comparable reference: Article with the same merchandizing characteristis Most likely similar to the new article.
Q= Total Sales of comparable
Available Comparable?
Q= average total sales of the Subfamily
More than comparable
Adjust average with a multiplier: range [1.3-2]; Q'= Q* 1.3
Number/type stores receiving
High trend
Q''=Q' * 1.5
Expected high acceptance?
Top-down Colors: Decide number of colors to offer
Medium:
C1=Q''*0.2
Relative Color Confidence
High:
C1=Q''*0.4
Low:
C1=Q''*0.1
Top-down Sizes: select a mirror curve. A comparable article with similar fitting and cutting characteristics. Offer the same number of sizes, in the same proportion

Figure 4: Purchasing Department Initial Purchasing Process Map

Figure 4 shows the general process follow by the different purchasing departments. The different purchasing departments have different necessities that can drive to different specifications in the purchasing process.

3.2 Chapter summary

In the interviews with the purchasing departments, people stated that the two desired outputs of the buying process are the initial quantity to order and, based on this quantity, creating the most appropriate cost and time-efficient sourcing strategy. The inputs of the three mapped processes are similar in essence, but the use of the appropriate historical information combined with the commercial understanding of the customer makes each purchasing process successful and different. The individual purchasing departments at Zara have specific characteristics that make them source final goods and raw material using different methodologies. The different groups have heterogeneous metrics to measure success but all of the departments are focus on final goods availability using the least possible amount of budget.

4. Improving the purchasing methodology

Analyzing the current state of the different purchasing methodologies leaves room for improvement in the process. This chapter proposes an improved methodology.

4.1 Room for improvement

Section 3.1 explains in detail the different purchasing processes at Zara. Each purchasing department follows its own process and has its own priorities. Besides the heterogeneity in the methodologies, the fundamentals of the purchasing processes can be improved to better reflect the realities of the marketplace and the supply chain. The observed opportunities for improvements are eight:

1. *Forecast demand.* None of the three aforementioned purchasing processes involve any explicit forecasting calculation. The processes systematically use the historical POS data as proxy for demand. That is, the aggregated sales of the comparable reference are used as the point forecast for the new article.
2. *Optimal purchasing quantity.* There is no evidence in the processes described above that the buyers recognize that the optimal purchasing quantity might be different than the point forecast. The purchasing process occurs in an uncertain environment, the acceptance level of the article in the market, the macroeconomic conditions in the countries, the fashion trend, among others factors can modify the demand level for an article. These uncertain conditions are nowhere recognized in the purchasing process. Likewise, the profitability of an article depends on two factors: the expected revenues and the potential costs. In the process there is no evidence of balance between the costs of underage and overage costs to select the optimal purchasing quantity.[27]
3. *Convert point of sales (POS) data into demand data.* Zara's POS data is not demand data. The POS data represents the portion of the demand that Zara was able to capture in the selling season. This captured demand differs from the total demand that Zara could have captured. This difference in the aggregate significantly changes the sales potential of an article. Sixty articles sold in the stores were used as data sample set. Table 2 shows the results of this comparison. The total actual amount sold was compared to an estimated sales potential. The estimated sales potential is the amount that an article could have sold if the inventory would have been infinite (Section 4.3 explains the proposed methodology to estimate sales potential). The current methodology estimates the future sales expectations based on the historical POS data, not the demand data, leaving aside a good part of the sales potential. It is necessary to correct the observed sales for the out-of-stock situations and other inefficiencies in the supply chain
4. Table 2: Estimated sales potential with infinite inventory (60 articles sample)

[27] (The new science of retailing. M. Fisher, and A. Raman, 2010. p.p. 67-76)

Summary: Relative Error Results 100* [(estimated-actual)/actual]	
Mean	41%
Standard Deviation	35%
Maximum	138%
Minimum	0%
Sample size	60
Summary: Absolute Error Results	
Total Sales [individual pieces of garment]	2,045,339
Total Estimated Sales [individual pieces of garment]	2,678,870
Lost selling potential [individual units of garment]	**633,532**
Lost selling potential as percentage of sales	**31%**
Average price (Euro)	€ 29.95
Potential Loss in Sales (Euro)	**€ 18,974,275**

5. *Account for growth.* In the past 7 years, Zara alone has grown from 1,000 stores to 1,600 stores around the world. The stores can be classified based on their selling potential into five categories. In descending order they are Top Seller, A, B, C and D. The Top Seller stores built in prime locations around the world. They have a very high rate of sales and they need large amounts of stock to be able to start selling; they close their daily performance summary with weekly sales balances in the tens thousands of Euros and they carry the latest trends of the season, 95 out of the 1,600 stores are considered TOP and they account for the 10% of the total sales. In contrast, D stores are small stores, most of them located in small towns in different countries. They enjoy market power in the shopping malls because they are the only selling option in terms of fashionable and affordable clothing. The D stores serve a population with a low average income and most likely with a more utilitarian taste. The weekly sales balance in the thousands of Euros; Having these significant differences in stores performance, inventory requirements, and consumers' tastes makes it necessary to account appropriately for growth when deciding the initial purchasing quantity. In the current use process the buyers account for growth by using the most up-to-date available sales comparable. They claim that using the latest sales data will account for growth in the stores. If they consider the historical sales-comparable is not reflecting the real growth of Zara, they will boost the expected sales by a fix percentage. This method accounts for growth in a suboptimal way; using a fix growth quantity (i.e. 10%) is not reflecting the real selling

potential of the new stores. That is, the selling potential of three D stores is not the same as the selling potential of three Top stores. To account for growth in the company correctly, it is necessary to estimate the selling potential of the sales comparable in the new stores.

6. *Selecting the specific network that will receive the article.* For many reasons the network of stores receiving the new article can be different that the comparable reference. Each set of stores has different selling potentials. Using an incorrect selling potential as baseline might result in over or under stock for the new reference. The Woman Collection at the pilot stores has the primarily objective of scouting the performance of the article to identify the trends of the season. Using historical sales data that erroneously indicates sales in all the stores of Zara's network can introduce error into testing the trend in the piloted stores. After observing the operation, I can state that the "controlled pilot group of stores concept" has evolved into an "all the stores that we can cover" concept. In 2010, 1,428 stores out of 1,600 received at least one article from the Woman Department. Mainly because buyers cannot focus only on the targeted stores, they end-up having excess initial inventory that they send to a larger group of stores. If this new reference is used as baseline for a future reference, the number of stores will be automatically excessive. An automatic tool that selection of the specific network of stores that will receive the new article, associated with the selling potential of each store can be implemented to address this issue.
7. *Extrapolate initial sales observations taking into account seasonal effects.* The retail industry has experiences seasonal variations that can result from natural forces, i.e., seasons in a year and human decisions or customs. Many types of economic activities depend on the fluctuations of this seasonality.[28] As described above, in the current methodology buyers will extrapolate daily sales to weekly sales by using simplified factors. Similarly they will extrapolate the seasonal factors from the weekly sales without taking into account the importance of the peaks and valleys of the different weeks. In reality, selling 100 units of an article on the first day of sales if that day happens to be a Saturday is not the same as selling them on a Tuesday. In addition, the current

[28] (Inventory Management and Production Planning and Scheduling. Silver, Pyke and Peterson 1998)

methodology does not take into account the time of the year when the comparable article was launched. That is, the buyer will use the aggregated sales of the historical comparable reference as the baseline for the new item regardless of the weeks when the comparable item was sold. Most likely, the new article might be launch in a different week of the year. Therefore, when using the sales of the comparable item to forecast selling potential for a new article, the comparable sales need to be deseasonalized and reseasonalized for the new article launching date. The seasonal effects have a significant impact on Zara's performance, and they need to be taken into account when deciding how much to buy.

8. *Calculate the error terms.* A main aim of this case study is to identify those areas that the purchasing departments can constantly improve. Management at Zara is constantly experimenting to learn more about the work they do. One measurement of improvement in the purchasing process is how successful the buyers are at forecasting demand; this metric has not been implemented but is approximated by using the number of units that need to be marked down.
9. *Automate the process.* Zara is updating and robustly enhancing all their IT technologies. Most of the areas of improvement previously proposed require enhancing the database access and the user interface design that the buying teams use. These system improvements will be incorporated into a corporate tool that is being developed in-house. Chapter 5 describes the preliminary format planned for the purchasing tool.

4.2 Proposed standard methodology: Point forecast to estimate demand

Figure 9 shows the purchase optimization top-level process flowchart. The purchasing optimization process contains 3 modules, inputs, outputs, and the update module. The inputs module collects the required information for the optimization model. It contains four sub-modules, historical information, demand forecast, exposition stock, and suppliers and transportation costs. The most important of this input is the demand forecast for the article. This thesis focuses specifically on defining a process for the forecast sub-module. The forecast process addresses six out of the eight improvement areas described in Section 4.1. The output module answers the three most important questions for a buyer, obtains the intra-week seasonality per country. The intra-week seasonality represents the observe peaks and valleys in the weekly selling pattern. That is, the days of the week when a given country has more selling potential. Similarly, the second input is the inter-week seasonality per country. The inter-week seasonality input evaluates the peaks and valleys observed in the yearly selling pattern. That is, the weeks of the year when a country has the highest selling potential. The third input module is the clustering stores module. The main function of this input is to group together the stores with the most similar selling behavior. Section 4.3 describes in detail the calculation of the three inputs modules.

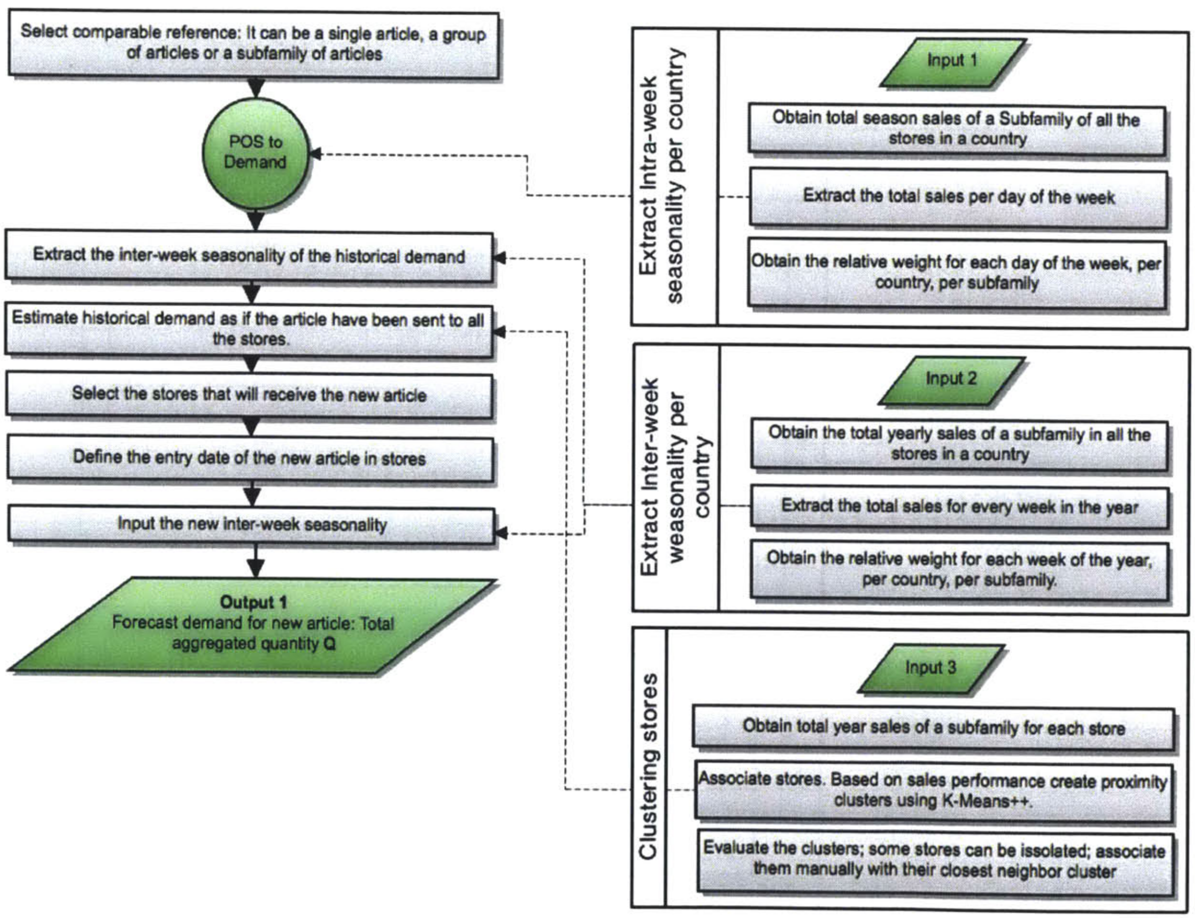


Figure 6: Proposed standard methodology: Initial purchasing quantity process map

The left portion of Figure 10 shows the point forecast process. After selecting the historical information use as baseline data, the process in Figure 10 highlights the sub-module POS to Demand. Figure 11 explains in detail the sub-module that converts the POS data into demand data. The output of this sub-module is the demand dataset for the new article. Once obtained the demand dataset, it is modify to include the stores' situation at the launching date of the article. The number and type of stores receiving the new article, the intra-and inter-week seasonality affecting during the launching period of the new item characterize the stores' conditions. These conditions give each article a different selling potential their calculations is explain in detail bellow in this section.

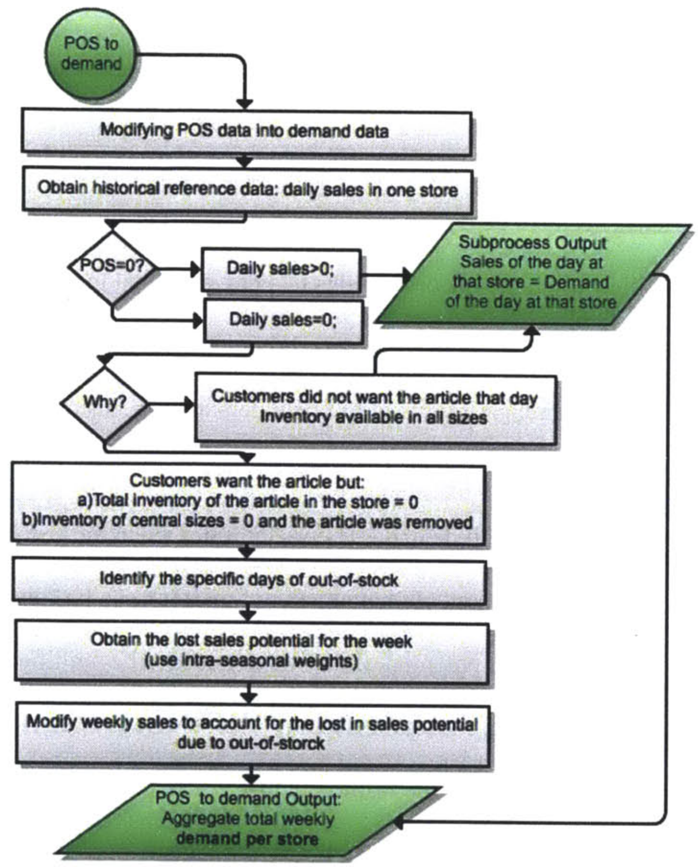


Figure 7: Proposed standard methodology: Converting POS data into demand data process map

The first step in the improved methodology is to obtain the baseline information. All the purchasing departments at Zara use the historical sales information of a comparable reference as the baseline demand forecast for a new article. Following the same logic, a comparable historical reference will be selected as a baseline, but the sales figures need to be adjusted for the occasions when out-of-stock was this sub-module is the demand dataset for the new article. Once obtained the demand dataset, it is modify to include the stores' situation at the launching date of the article. The number and type of stores receiving the new article, the intra-and inter-week seasonality affecting during the launching period of the new item characterize the stores' conditions. These conditions give each article a different selling potential their calculations is explain in detail bellow in this section.

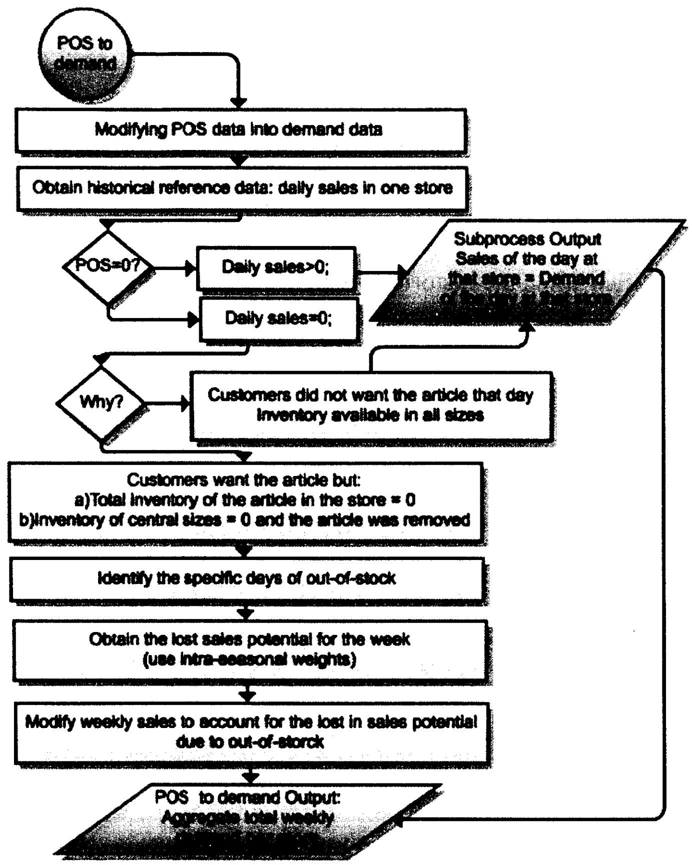


Figure 7: Proposed standard methodology: Converting POS data into demand data process map

The first step in the improved methodology is to obtain the baseline information. All the purchasing departments at Zara use the historical sales information of a comparable reference as the baseline demand forecast for a new article. Following the same logic, a comparable historical reference will be selected as a baseline, but the sales figures need to be adjusted for the occasions when out-of-stock was registered. This approach is a better predictor for the new item's demand because it does not include the supply chain inefficiencies from the past.

Figure 11 shows the changes required to modify the POS data from a historical reference into demand data. The disaggregated (store level) daily sales dataset will be analyzed to observe the inter actions between the daily sales and the on-hand-inventory. This analysis will identify and correct the sales pattern for those days when out-of-stock occurred. The corrected pattern will simulate the potential sales performance of a historical reference with infinite stock availability. The pattern will be corrected by inflating the weekly sales when observing out-of-stock to correct for the potential lost sales.

Inflating the sales needs to be done only when the demand is not satisfied. A stock-out position in the dataset is identified when the historical sales in a day are equal to zero. Having no sales in a given store in one day can mean that the customers did not want to buy the article that day. That is, even when the shelves had enough inventory to fulfill demand, the customer decided not to buy the article. Alternatively, the customers might have wanted to buy the item, but the inventory in the store might not have been enough to fulfill demand. Having no sales when the article is available requires no modification in the sales pattern, but having an out-of-stock can be considered as a potential loss in sales, and therefore the sales potential of that week needs to be modified.

An out-of-stock occurs when the total inventory of an article is zero. However, the inventory display policy at Zara is more severe. In the fashion industry a SKU will identify a specific cutting pattern, the color, and the size of article. For merchandizing and logistic purposes, Zara has subdivided the offered sizes into two groups, central and minor sizes. The central sizes are the ones that continuously have a larger demand, i.e., the Medium size (M) in T-shirts accounts for 40% of the total sales in the world, while the remaining 60% can be divided in fours sizes: Large, Small, Extra Small, and Extra Large. To maximize sales-per-square-foot, Zara inventory display policy requires the store managers to remove an article from the display area if at least one of its central sizes is out-of-stock. This requirement directly affects the potential demand of the article. If an article was removed from the shelves, automatically the recorded sales are zero and it will be necessary to boost the sales pattern for this condition. Equation 1 shows how to count the number of days (per day type) not displaying an article due to inventory out-of-stock in a store: DND_{rj}^{d}. In Equation 1 the set of sizes for reference r are subdivided in two subsets: the central sizes S_r^+ and the minor sizes S_r^-. That is $S_r = S_r^+ \cup S_r^-$ and S_r^+. The first part of Equation 1 captures the days where the total inventory for article r in store j was equal to zero. The next part sums the days when the inventory of at least one of the central sizes S_r^+ was zero and the observed sales for the article were zero. Equation 1 follows the principles developed by Caro and Gallien (2010), they were optimizing the distribution of fashion items once the purchased quantity is defined; with a similar mathematical model they count the number of days in a week when a given size of a given article was not displayed in the store, this is only paper available in literature that consider the broken assortment policies at Zara.

$$DND_{rj}^{d} = \left\{I_{rj}^{d} = 0\right\} or \left\{min_{\tilde{s} \in S_r^+}\ I_{r\tilde{s}j}^{d} = 0\ and\ max_{s \in S(r)}\ Sales_{rj}^{d} = 0\right\}$$

Equation 1: Days not displaying a reference due to out-of-stock in a store[29]

A store j in a week w with infinite inventory has a selling potential of 100%. In reality it is impossible to carry infinite inventory in the stores; some weeks have less than 100% selling potential. Given the intra-week seasonality in the Zara stores, it is necessary to deduct from the selling potential the specific weight of the days with out-of-stock. Section 4.3 describes the calculation of the intra-weekly seasonal factors for a subfamily R and a country P ($\delta_d^{P\,R}$) for each day of the week. If the historical article r belongs to the subfamily R and the store j belong to country P, the intra-week seasonal factors for the subfamily-country are the same for the article-store: $\delta_d^{P\,R} = \delta_d^{j\,r}, \forall\, j \in P\ and\ r \in R$. Equation 2 shows that the captured selling potential will be obtained by combining the days not displaying an item and their corresponding weight. Equation 2 follow the same principles developed by Caro and Gallien (2010) where they construct a data series that provides an estimate of the uncensored customer demand, that is, the sales that would have been observed if the stores have all the merchandize

[29] F. Caro and J. Gallien. Inventory Management of a Fast-Fashion Retail Network (2009). p.p. 265-266

displayed and without any stock-out. That application was used in the context of merchandize distribution, a similar approach is proposed in this document.

$$SP_{rj}^{w} = \sum_{d \forall D} \delta_{d}^{j\,r} * (1 - 1_{DND_{rj}^{d}})$$

Equation 2: Weekly selling potential for a given article in a given store [30]

Equation 3 estimates the weekly demand for the reference by taking the ratio of the sales in week w over the captured selling potential in the same week.

$$Demand_{rj}^{w} = \frac{Sales_{rj}^{w}}{SP_{rj}^{w}}$$

Equation 3: Weekly demand for a given article in a given store

The diagram in Figure 11 explains the steps to generate the output: $Demand_{rj}^{w}$. This pattern has modified the observed sales to capture the potential loss in demand coming from insufficient inventory in the system. The next step is to adapt the historical reference environment to simulate an environment similar to the environment expected for the new item. The first main difference between the comparable reference and a new article is that the launching dates for the articles differ. The strong seasonal influences of the different introduction dates need to be adjust by removing seasonality from the historical demand dataset. Section 4.3 explains how to estimate the inter- week seasonal factors for a subfamily using information from earlier seasons. The resulting input arrays will be used at this point of the methodology. Once the POS data of a historical reference is converted into demand, the corresponding inter-week seasonality is extracted by dividing each component of the demand dataset by the corresponding inter-week seasonal factor of that week. Similar to the intra-week seasonal factors, the inter-week factors are obtained per subfamily and per country. All the stores that belong to the country will share the same inter-week seasonal factors: $\tau_{R\,P}^{w} = \tau_{r\,j}^{w}, \forall\, j \in P \; and \; r \in R$; the logic behind this thought is that usually a given geographical region shares the same socio-political

---

[30] F. Caro and J. Gallien. Inventory Management of a Fast-Fashion Retail Network (2009). p.p. 264-265

behavior, and therefore consumers' shopping habits are similar. Equation 4 shows the required calculations to deseasonalize the demand pattern, obtaining the array $DesDemand_{rj}^{wl}$ that has two columns. The first one counts the number of weeks from the launching date wl. The second column has the corresponding demand in an average week.

$$DesDemand_{rj}^{wl} = \frac{Demand_{rj}^{w\,wl}}{\tau_{r\,j}^{w}}$$

Equation 4: Deseasonalizing the demand dataset

The second main difference between the historical reference and the new item is the network of stores that retail the articles. The historical reference might not have been sent to 100% of the current stores. Two main reasons exist for a store not receiving the historical reference r. First, a portion of the stores did not exist or were under restoration at the time of distributing article r. The second reason is insufficient inventory: only stores with high selling potential for that type of article received article r. Since the article might reach a different subset of stores than the network of stores planned for the new article, it is necessary to give the buyers the flexibility of choosing from the whole network of stores. Therefore, the demand pattern for the historical article needs to be extended to 100% of the stores available today.

To extrapolate the demand dataset to the stores that did not receive the historical reference r, the output of the clustering analysis described in Section 4.3 is necessary. That output is the segmentation of all the existing stores j belonging to Zara's network J into k clusters. In a cluster there are two types of stores: the subset $\breve{k}$ that receive article r sub-grouped and the rest that did not receive article r. The total volume sold by the stores that form a cluster represents the total selling potential of the cluster. Similarly to the seasonality factors calculation, the selling potential differs between stores; within a cluster, the volume sold of the stores varies around 30% from the centroid, as explained in Section 4.3. Therefore, it is reasonable to say that each store has a different relative weight in the cluster. Equation 5 shows that the weight of a store in a cluster W_k^j is the ratio of total sales of store j divided by the total sales registered in the cluster. Each store will contribute a different percentage to the clusters' performance.

$$W_k^j = \frac{\sum_{r\in R, w\in Y} Sales_{rj}^{w}}{\sum_{r\in R, j\in k, w\in Y} Sales_{rj}^{w}}$$

Equation 5: Store's contribution to the cluster

The store's contribution to the cluster is the main source of information when estimating the selling potential of a store that did not receive a historical reference *r*. To do this estimation, it will be assumed that the selling potential of an article *r* that belongs to subfamily *R* is the same as the store's average selling contribution to the cluster. Equation 6 shows how to use a store's selling contribution to a cluster to estimate the historical sales of the stores that did not receive an article *r*.

$$DesDemand_{r,j\notin \bar{k}}^{wl} = W_{j\notin \bar{k}} \frac{\sum_{j\in \bar{k}} DesDemand_{r,j}^{wl}}{\sum_{j\in \bar{k}} W_j}$$

Equation 6: Estimated deseasonalized demand of a store that did not receive a historical reference

After the estimated demand for the stores that did not receive the historical article is obtained, the total demand of an article for the whole network of stores can be aggregated. Equation 7, the total deseasonalized demand for article *r* in a cluster, is the sum of the observed demand for the stores that receive the article and the estimated demand of those stores that did not receive it. The total deseasonalized demand for the article for the whole network of stores is the aggregation of the demand for all the clusters.

$$DesDemand_{r,k(j)}^{wl} = \sum_{j\in k} 1_{\{j\notin \bar{k}\}} * DesDemand_{r,j}^{wl} + \sum_{j\in k} 1_{\{j\in \bar{k}\}} * DesDemand_{r,j}^{wl}$$

Equation 7: Estimated demand for a cluster

The next steps in the standardized methodology incorporate the known information about the new article. To avoid confusion with the historical reference, the new article is referenced as the article *i*. It represents a piece of cloth that will be introduced to the stores and uses the sales information of article *r* as a baseline demand forecast. The most important pieces of information that determine the stores' situation for the new article are the planned network of stores that will display the new article and the launching date of the article. The planned network of stores receiving the new article i might differ from the network that displayed the historical reference. The set of stores receiving the new article might be a subset of the whole network, but this subset might include the new openings (excluding the closures) that the company faces. The expected launching date of article i in the stores is a necessary piece of information that determines the inter-week seasonality factors for the new article in the stores. The buyers propose the launching date based on the suppliers' production schedules and the estimated inbound transportation lead-times. They also select the network of stores that will receive article i based on the characteristics of the article (ramp-up, winter, trendy, etc.), taking into account the characteristics of each purchasing department (Woman Department will pilot in just few stores). The deseasonalized demand for each store $DesDemand_{rj}^{wl}$ is seasonalized with the inter-week seasonal factors corresponding to the launch date of the new article. Equation 8 shows the calculation for the pattern of the new article i: $Demand_{i,j}^{w,wl}$. Similarly to the inter-week seasonality factors for the historical reference r, the factors for the new article are extracted from the output generated in Section 4.3. The inter-week factors for the new item might differ from the calculated factors of the historical reference; if new information is available (a full season is completed), the average factors per country and per family will be updated. Including the newest information in the demand dataset of the new item i, will reflect the most updated information of the stores' situation. That is, if the shopping habits of people have changed, the new item will be able to capture this new trend. The stores j that belongs to a country P has the same inter-week factors. These factors are obtained per subfamily, this point applies for all the articles that belong to subfamily R: $\tau *_{R\,P}^{w} = \tau_{i\,j}^{w}, \forall\, j \in P \; and \; i \in R$ where $\tau *_{R\,P}^{w}$ are the most up-to-date seasonal factors. These seasonal factors are integrated into the deseasonalized demand pattern of the historical reference to obtain the re-seasonalized demand pattern for the new article i. Equation 8 represents the relation between the re-seasonalized demand and the deseasonalized demand.

$$Demand_{i,j}^{w,wl} = \tau_{i\,j}^{w} * DesDemand_{r,j}^{wl}$$

Equation 8: Seasonalized demand for a new article

The total seasonalized demand pattern for the new article *i* is aggregated over the selected network of stores. This pattern contains the required quantity per week; the initial estimated life of the new article is the lesser of the number of observed selling weeks (*wl*) for the historical reference and the remaining number of weeks in the season. In other words, if the historical reference was sold during 20 weeks but the season is halfway started and only 16 weeks remain to the beginning of the markdowns, it is necessary to purchase enough inventory of article *i* to satisfy demand for only16 weeks. The life expectation of an article is a new concept for buyers.

4.3 Proposed standard methodology: Inputs

The similarities and differences between the purchasing groups in the company contributed to the outline of a standardized purchasing methodology. Section 4.1 highlights areas for improvement. Section 4.2 explained the point forecast process to estimate demand. In addition, Section 4.2 explains how to modify POS data into demand data. Section 4.3 continues incorporating improvements to the methodology. Section 4.3 describes the input section used in the process described in Figure 10. These inputs are related to stores' situation at the moment of launching the new article: seasonality, growth, and number of stores selling the new article. This methodology incorporates the operations cross-shared by all the different purchasing groups and adds some of the best practices that each group uses individually to enhance the process. The inputs can be calculated in a different moment, on a different server, and the resulting output document can be recall from the database, the main purpose is to reduce the waiting time when calculating the inputs; three different input modules need to be estimated. The calculation details are described below:

a) *Intra-week seasonality per country per subfamily.* The first main difference between the comparable reference and a new article is that the launching dates for the articles differ. The strong seasonal influences of the different introduction dates need to be adjust by removing seasonality from the historical demand dataset. The main focus of this sub-module is to extract the relative weight of each day of the week (Monday, Tuesday, etc.) per geographical region (France, Mexico, etc.). Different cultures around the world exhibit different economic behaviors and consumer practices: in some countries the stores remain open seven days a week and consumers tend to be more active during the weekends. Other cultures have the stores open only for six days a week, and the consumers are more active towards the middle of the week. These differences in shopping habits are extremely important when using historical information to predict the future performance of an article, and failing to incorporate them can result in an over- or under-estimation of the total required quantity of an article. The proposed methodology allows the system to calculate the seasonal factors in a separate process, that is, the calculation can be performed overnight and use the resulting factors without waiting for the output document. The main idea is to use the total aggregated data from the closing season to extract the steady performance of each day in each country. This calculation will require the database to work overnight collecting the information, aggregating the data, and evaluating the selling potential of each day compared to the rest of the days of the week. The IT systems at Zara have the required capabilities and structure to perform this task. This seasonal calculation will be performed only when the selling season ends. Equation 9 shows the calculations required to obtain the intra-week seasonal factors $\delta_d^{P\,R}$. The intra-week seasonal factor is how much the demand for a particular day compares to the rest of the days. The idea of a different selling potential in different days of the week, is widely accepted not only in Zara, but also in the retail industry in general. The two traditional approaches to generate these intra-weekly seasonal factors are: linear regression analysis and averaging method[31]. The linear regression approach requires a linear regression of the data. Then, to obtain the seasonal factors divides each point of the actual demand by the linear regression for the period. Finally, the factors for each period are averaged to obtain the seasonal factor per type of day. The problem with this method is that Zara's data hardly fits a straight line. Therefore, it was selected the modified averages method described in Equation 9.

[31] (R. Tibben-Lembke. Forecasting with Seasonality (2003))

$$\delta_d^{P\,R} = \frac{M_{m(d)}^{P\,R}}{\sum_{m(d)\in M} M_m^{P\,R}}; \forall\, d \in D.$$

Equation 9: Intra-week seasonal factors[32]

Where $M_{m\,(d)}^{P\,R} = \frac{\sum_{d\in D\; r\in R\; j\in P} 1_{\{m(d)=m\}}\, Sales_{rj}^{d}}{\sum_{d\in D} 1_{\{m(d)=m\}}}$; where m is the type of day resulting from the calendar date.

The indicator function $1_{\{m(d)=m\}} = \begin{cases} 1\; if\; m(d) = m \\ 0\; if\; m(d) \neq m \end{cases}$. The indicator function compares the calendar type of day *m* to the indexed type of day *d* that numbers the days in a week from 1 to 7. When the function in the bracket become true, that is when the calendar day matches the indexed day, then the function equals one; it will be zero when the function in the brackets is false.

Each factor is the average selling potential of a given type of day *d* ($d \in D$ is the set of days in a week) in country *P* (where j ∈ P is the set of stores located in country P). The intermediate factor $M_{m\,(d)}^{P\,R}$ obtains the selling potential of each type of day by adding the daily sales ($Sales_{rj}^{d}$ where r ∈ R are the set of articles that belong to the subfamily *R* and d ∈ D are the specific days being analyzed) observed on a given type of day and dividing them by the number of a given type of days available in the season (open store days), i.e., 10,000 units sold in 10 Mondays. Each factor $M_{m\,(d)}^{P\,R}$is normalized by dividing it by total sum of $M_{m\,(d)}^{P\,R}$, the total selling potential of all the days in a week Figure 12 shows the resulting intra-week factors for Spain and the Netherlands. The most relevant differences occur on Thursday and Saturday. The shopping hours on Thursdays in the Netherlands are extended from 11:30 am-18:00 pm to 11:30am -21:00 pm, this increase the selling potential in the stores. The factors for Spain are similar to most of the European countries.

---

[32] (S. Makridakis, S. C. Wheelwright & R. J. Hyndman. Forecasting Methods and applications. John Wiley & Sons (1998) p.p. 88-93.)

	Intra-week seasonal factors		
	Day type	Spain	Netherlands
1	Monday	0.12	0.09
2	Tuesday	0.09	0.12
3	Wednesday	0.13	0.15
4	Thursday	0.15	0.21
5	Friday	0.19	0.19
6	Saturday	0.23	0.13
7	Sunday	0.09	0.11

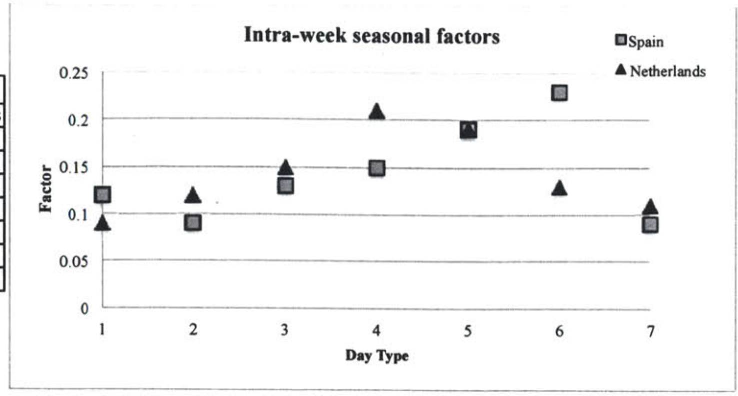


Figure 8: Example of two countries' intra-week seasonality factors using 2010 data

The final array stacks the results of the estimated factors from all the stores in the network; having these outputs recorded allows the buyers to recall and use the weight information in fractions of seconds. To incorporate the relative weight for each type of day in each country of the world without taking time from the buyers is a great benefit for the purchasing methodology.

b) *Inter-week seasonality, per country, per subfamily.* The fashion apparel industry is characterized by a well-defined seasonal cycle by weather conditions in the world. The two main selling seasons are the winter season, from August to December (in the northern portion of the world), and spring season, from February until May. Each season is followed by two months of markdowns. In addition to weather-related cycles, the apparel industry cycles are strongly correlated with the social and cultural holidays around the world. Zara has operations in more than 80 countries. The most relevant holidays or promotions that affect the overall consumer behavior are: Christmas, the Chinese New Year, Mothers Day, Ramadan, Easter, and Thanksgiving. Other socio-political events that might have a negative effect on consumer behavior are strikes, economic crises, war, earthquakes, etc. These events are difficult to predict, but their effects can be translated into drastic actions such as store closures, having an important impact on the demand for apparel articles. In conclusion, positive and negative cyclical events are directly correlated with consumer behavior and can vary drastically from week to week. Therefore, extracting the pattern that corresponds to the specific dates of the selling period of the historical reference is important; the main purpose is to isolate the selling behavior of the historical reference as if all the sales have happened on an average week. This average pattern will be updated with the expected inter-seasonal pattern derived from the introduction date of the new reference. Similar to the intra-week seasonal input module, this estimation is separated from the main process and will be computed yearly. The output will be a series of arrays that will pair each week in a year with its corresponding weight (relative to an average week of sales); these results will be broken down per subfamily and per country.

The inter-week seasonal factor is how much of the demand for a particular week tends to be above (or bellow) the average demand of the year. The required inputs to compute the inter-seasonal factors are the total weekly sales (all the articles sold per subfamily) observed in a particular country in a window of one year. This dataset is an array of two columns; the first one indexes the week of the year ($w \in Y$ refer each weeks in the fiscal year, and Y is either 52 or 53 depending on the fiscal year). The second part of the array is the total sales of the subfamily in the country during that specific week of the year ($\sum_{r \in R, j \in P} Sales_{rj}^{w}$; where $r \in R$ and r is an article that belong to a R and $j \in P$ with j being all the stores that belong to a country P).

Equation 10 shows the calculation required to obtain the inter-seasonal factor $\tau_{R\,P}^{W}$. These factors are obtained as the ratio of the total sales on a given week divided by the average total sales for that year; the denominator computes the total sales over the year divided by the number of weeks in the year. The final output will provide the relative weight of each week compared to an average performance week.

$$\tau^{w}_{R\,P} = \frac{\left(\sum_{r\in R, j\in P} Sales^{w}_{rj}\right)}{\left(\frac{\sum_{r\in R, j\in P, w\in Y} Sales^{w}_{rj}}{Y}\right)}$$

Equation 10: Inter-week seasonal factors per year, per subfamily and per country[33]

Figure 13 shows an example of the inter-week seasonal factors for France in 2010. Near week 18 the stores start having an enhanced selling potential compared to the average; this is an expected peak due to the spring break/Easter period. The valley observed in the last week of February comes from the number of days the month (only 4 Saturdays); in February the stores are showing the first advance of the spring collection, the advance tend to be expensive. Thus, consumers that are sensitive to price will wait to get the item.

Year Week	Week	Inter-week factor
201101	1	1.16
201102	2	1.13
201103	3	1.11
201104	4	1.03
201105	5	1.13
201106	6	1.01
201107	7	0.46
201108	8	0.34
201109	9	0.44
201110	10	0.53
201111	11	0.74
201112	12	0.87
201113	13	0.96
201114	14	1.04
201115	15	1.18
201116	16	1.24
201117	17	1.27
201118	18	1.27
201119	19	1.55
201120	20	1.78
201121	21	1.68
201122	22	1.46
201123	23	1.49
201124	24	1.57
201125	25	1.44
201126	26	1.35
201127	27	1.27
201128	28	1.15
201129	29	0.98
201130	30	0.83
201131	31	0.64
201132	32	0.58
201133	33	0.18

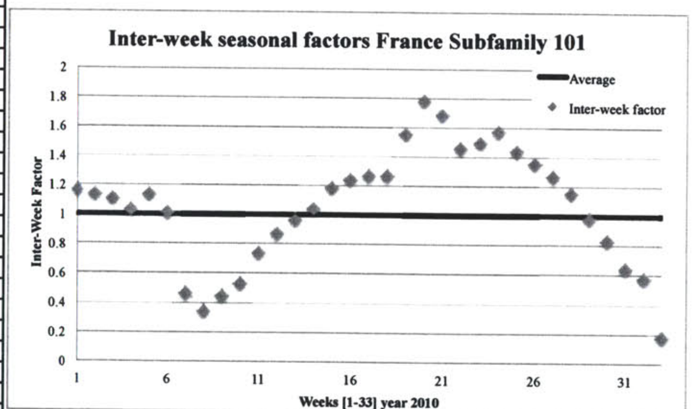


Figure 9: Example of inter- seasonal factors for France using 2010 data

c) *Clusters of stores.* Clustering analysis is the process of grouping similar data together. Currently, Zara has subdivided the 1,600 stores in the network into five different types of stores (Top Seller, A, B,

---

33 (S. Makridakis, S. C. Wheelwright & R. J. Hyndman. Forecasting Methods and applications. John Wiley & Sons (1998) p.p. 88-93.)

C, and D) based on the total amount sold and the inventory turnover observed in the previous season. This simple grouping method helps the logistics group to prioritize their shipment. However, in terms of purchasing quantity, the grouping method is static and not precise enough to determine sales patterns, the groups are too large.

The number of operating stores at Zara is constantly changing (opening and closing stores). For buyers it is common to use historical references that were sold in a different subgroup of stores instead of the subgroup of stores planned for the new article. This difference in the store network can be translated into different initial purchasing quantity requirements. To solve for this inconsistent network size and to give flexibility to the buyers when they decide what subgroup of stores will receive the new article, it is necessary to estimate a hypothetical scenario for the historical reference. In this scenario all the existing stores in the network would receive the historical reference and would have historical sales data. If a store did not receive the historical reference, cluster analysis estimates its performance based on the performance of similar stores.

The clustering method selected to perform the segmentation is called k-means++. This method is a variation of the classical k-means method. The standard k-means method is a hard clustering algorithm that assigns each element to only one of the clusters. The process starts by segmenting the input points, in this case stores' end of the season sales dataset per subfamily, into *k* number of initial groups, also called seeds. The initial groups are selected uniformly at random. Then, the algorithm calculates the mean point (centroid) of each group. The next step is to obtain the Euclidean distance from every point to the neighboring centroids. If is necessary, the points are reassigned to a group with a closer centroid. A new iteration will start using the new set of points to obtain the new centroids. From this point the algorithm will iterate until the points do not switch clusters. The variation between k-means and k-means++ lies in the prescreening of the initial seeds. Using initial seeds selected uniformly at random can lead to a suboptimal grouping. To overcome this suboptimal calculation, the algorithm k-means++

assigns a weighted probability to each point, increasing the chances of selecting the optimal segmentation.[34] The implemented algorithm appears in Appendix 1.

The input required to run the k-mean++ algorithm is the total sales (units) dataset per store and per subfamily of the previous selling season. The total volume sold in a given store is the x element of cluster k. The matrix of distances between the optimal seeds is calculated using the Euclidean distance between point x and the nearest centroid. Depending on where the comparison begins, the measured distances and center points of the clusters will be different. The algorithm uses an optimization method to assure the least squared distance. The distance between the points and the centroids will be minimized, while the inter-cluster distance is maximized. In other words, the algorithm tries to group the stores with the most similar selling behavior.

The resulting output of the *clusters of stores* inputs module is a list of arrays that partition all the stores into k clusters. The array contains the cluster number (ID) and the stores that integrate that cluster; the array also records the average performance of the cluster by extracting the characteristics of the cluster's center points. The extracted information for each of the center points is used as the estimated performance of a store that did not receive an article. That is, if a member (store) of cluster number 1 did not receive a historical reference, it is inferred that its behavior will be similar to the average cluster behavior. Using the k-means algorithm and the closing aggregated sales of 2009, the existing 1,300 stores in 2010 were distributed among 134 clusters for the different subfamilies. The clusters are updated when estimating the total demand for an historical article and new information is available. The algorithm uses POS data that has been archived in one of Zara's databases. The algorithm takes 56 seconds to run, using historical datasets from one closed season. [35]

In the baseline methodology, the inputs described above were fixed into the sales data, and extracting them requires simple mathematical algorithms. However, the datasets must be retrieved from the database archives and formatted in a very specific way. With this clustering procedure, the *clusters of*

---

[34] (K-means++: The Advantages of Careful Seeding, David Arthur and Sergei Vassilvitskii)
[35] (Zara, 2009 total sales information)

stores input can be complete overnight by the IT department, and the purchasing teams can easily recall the results. The modular design of the three input modules save significant amounts of time and nearly eliminate the calculation errors.

4.4 Top-Down forecast

Section 4.2 describes a methodology that generates a point forecast of demand at the article level. However, suppliers need the specifics of the purchased lot. These are the number of garments that they need to dye in a color and the number of units to cut in a given size.

4.4.1 Color Top-Down forecast

Section 3.1 explains how buyers decide how much to buy of each color. Most of their decisions come from empirical knowledge and trendiness expectations. When deciding the final breakdown of the total expected demand of an article into the demand per color, buyers have no other option but to use simple multipliers to get to the color level. After interviewing the different buying groups, they show interest in including the specific information of the network of stores that are planned to receive a particular color.

The color's performance in a store is highly dependent on the selling behavior of the store itself and the specific color preferences in each country. In other words, a Top-selling stores in Shanghai will sell more pink skirts than a B store in Mexico City. First, the Top store is more likely to have a larger volume of sales than the B store, and second, currently in China the color pink is in high demand. Buyers would like to have the option to break down the quantity in a more accurate way.

The standard process recommended doing the Top-Down color forecast uses the demand pattern for the new article *i* as the initial quantity to break down. The buyers input the number of colors that article *i* will offer based on the trend expectations for colors in the upcoming season.

An analysis to obtain the average number of colors offered at Zara and the strength of each color was conducted using two years of historical demand information (2009-2010). The analysis includes information from all the purchasing departments., It showed that the number of different colors offered, index $c\ (i)$, ranges between one and eleven. Appendix 2 shows the results of this analysis and the relative strength of each color when offered with a number of other colors. The first step in the Top-Down forecast of color is to estimate the strength factors per color per number of colors $W_{c\ (i)}$, the first breakdown of the demand pattern for the new article. These factors are obtained considering that 100% of the stores received the article.

The second step is to capture the percentage of each type of stores that are planned to receive article *i* in a given color. The type of store q is the stores' segmentation used by the Logistics Department to decide how to distribute the purchased quantity of an article among the different stores *j*. The buyers will input the specific subset of stores that will receive article *i* and the percentage of the stores. The percentage of stores that will not receive the article will be deducted from the 100% of stores of that type. The stores *j* that belongs to store type q and receive article *i* in color c are the subset of stores $\tilde{q}$.

The third step is to estimate the portion of demand for which each of the store types q is responsible. The stores' partition is mutually exclusive and collectively exhaustive; each type of store accounts for a different percentage of the total demand of a subfamily W_q. Similarly to the intra- and inter- week seasonal factors, the relative strength per store type will be calculated using historical data. The factor for the store type q, $W_q = \frac{\sum_{r \in R\ j \in q\ w \in Y} Demand_{r\ j}^{w}}{\sum_{r \in R\ j \in Q\ w \in Y} Demand_{r\ j}^{w}}$, where each of the five factors $W_q$represents the relative selling strength per store type. It will be obtained by dividing the aggregated yearly demand for a store type q per subfamily over the total yearly demand of the subfamily in all the stores.

The fourth step is to combine the strength per color with the percentage of the store type network receiving the article and the strength per store type to calculate the final confidence in the colors offered, $CC_{c\ (i)\ q}$. Equation 11 shows the calculation to combine the factors into the confidence in each color offered for article *i*.

$$CC_{c(i)\,q} = W_{c(i)} * W_q * \frac{\sum_{j\in q} 1_{j\in \tilde{q}}}{\sum_{j\in q} 1_{j\in q}}$$

Equation 11: Color Confidence for an article per store type

The final step is to break down the purchased quantity per article into the different colors, multiplying every confidence factor $CC_{c\,(i)\,q}$ times the sum of the obtained demand pattern for article *i* per store type q $\sum_{j\in q} Demand_{i,j}^{w,wl}$.

4.4.2 Size Top-Down forecast

Similarly to the breakdown required at the color level, garment suppliers require the specific number of units to cut for each size. As was explained in Section 4.1, buyers use as a guideline a so-called "Mirror Curve" to decide the portion of the article that will be produced in each size. The Mirror Curve is a historical reference for the fitting pattern of the garment, and it contains the historical sales information of an article that has similar proportions to the new article for which buyers plan. Similar to the demand estimation for the article level, the historical sales information for each size reflects only the portion of demand that the stores captured with their available inventory.

Similar to the article level demand estimations, it is necessary to convert the historical sales dataset into demand information. The first step is to count the out-of-stock days of the historical reference, in a given store, in a particular size. Using the same principle used in Equation 1, it is necessary to know if the article was displayed in the store or not. However, this condition is not enough to understand if a particular size was displayed or not. The resulting output from Equation 1 identifies those days when the article was not displayed in the store. For those days it is necessary to evaluate the on-hand-inventory for each particular size. Equation 12 explains the two required conditions to consider a particular size out-of-stock, and it refers to the days not displaying a particular size s: $DND_{r\,j\,s}^{d}$. The first condition looks for those days whereas customer wanted the article but the article was not display; Equation 3 satisfies that condition. However, the article might have been removed because some of the central sizes were out of stock. If that is the case, some of the sizes had enough stock to be displayed but the display inventory policy prevented them from being on the shelves. It is considered an out-of-stock at the size level only if the inventory-on-hand for the specific size being analyzed equals zero $I^{d}_{r\,j\,s} = 0$. The second part of Equation 12 looks for those days when the article was displayed but the on-hand-inventory for the size being evaluated was zero. Even when part of the demand was captured some customers for a particular size might be unsatisfied; this is an over-simplified assumption, where no size substitution is allowed.

$$DND^{d}_{r\,j\,s} = \{DND^{d}_{r\,j} = 0\}\ and\ \{I^{d}_{r\,j\,s} = 0\}\ or\ \{DND^{d}_{r\,j} = 1\}\ and\ \{I^{d}_{r\,j\,s} = 0\}$$

Equation 12: Out-of-stock days for a particular size

The second step of the size Top-Down forecast is to estimate the selling potential of the days when an out-of-stock event for a particular size was registered. The output from Equation 9, the intra-week seasonality factors, will be integrated into those days with out-of stock to obtain the forgone selling potential $SP^{w}_{rj\,s}$ for each size s. Equation 13 shows the calculation required to obtain the selling potential of each size.

$$SP^{w}_{rj\,s} = \sum_{d \forall D} \delta^{j\,r}_{d} * (1 - 1_{DND^{d}_{rj\,s}})$$

Equation 13: Selling potential of a particular size

Equation 14 shows the final step to convert historical reference sales dataset per size into demand per size. The weekly demand dataset per size is obtained by dividing the historical weekly sales pattern by the calculated selling potential for each week for each size per store.

$$Demand^{w}_{rj\,s} = \frac{Sales^{w}_{rj\,s}}{SP^{w}_{rj\,s}}$$

Equation 14: Weekly demand for a particular size in a given store

The total demand for the article will be aggregated over all the stores and over all the sizes. To define the portion to cut in each size, Equation 15 shows the Top-Down partition of the total demand of article r into the portion required for each size $W_{r\,s}$.

$$W_{r\,s} = \frac{\sum_{j \in J\; w \in Y} Demand_{rj\,s}^{w}}{\sum_{j \in J\; w \in Y\; s \in S} Demand_{rj\,s}^{w}}$$

Equation 15: Top-Down forecast at size level

4.5 Chapter Summary

Section 4.1 identifies the potential areas for improvement in Zara's currently used purchasing methodology. Sections 4.2 and 4.3 propose a standardized purchasing methodology, focusing on the development point forecast to estimate demand. The architecture of the proposed methodology obtains the purchasing quantity at the article level; it uses a combination of historical information and the buyer's expertise as inputs into the system. The pathways followed by this new methodology allow the buyers to enhance their decision-making process in three different ways. The first improvement is to use demand information in lieu of POS data. By incorporating into the sales data the portion of demand that Zara was not able to capture, buyers will have more accurate baseline information to predict future demand. The second improvement automates the introduction of inter- and intra- week seasonality into demand patterns based on the stores' situation at the moment of launching the new article. The third improvement refines the selection of the network of stores that will receive the new article, with the purpose of avoiding over-production if the article is not meant for all the stores. Finally, Section 4.4 shows the required calculations to Top-Down forecast the demand of the article into the color and size level. Together these four improve the current purchasing methodology, with a positive impact in the accuracy and time reduction on the purchasing process. Throughout this document there is no numerical evidence supporting the actual improvement. It is necessary to incorporate the changes to an automatic purchasing tool to see the benefits of this assessment that is beyond the limits of this thesis.

5. Architectural design of a tool to automate the proposed purchasing methodology

Every year Zara's purchasing departments source and buy the raw material and the finished goods for more than 14,000 SKUs. Each transaction requires not only the skills of the buyers and the product managers but also the systems preparation of the IT department to include the new references into

Zara's database. As was described in previous chapters, one key necessity at the purchasing departments is to automate the purchasing decision-making process. This chapter proposes the planned architecture for a web-base GUI for the purchasing departments to automate their decision making process.

The prototype interface is divided in three deductive menus; the first menu refers to the broad information of the season "Campaña". This part of the prototype deals with all the articles plans to be displayed in a season. The second menu is at the article level "Artículo". This part of the interface works with single articles. The article menu has the historical databases of historical references, the predictions for future sales, and a summary of the required quantity to purchase of a new reference. The third menu is the category "Categoría". This part of the prototype is devoted to the interaction between multiple items displayed at the same time in the stores.

5.1 Purchasing tool prototype: The selling season

One of the biggest challenges for the end users (buyers) of the purchasing tool is to manage large amounts of data from different selling seasons. The accessibility and the organization of the information are crucial to select the correct historical information and to keep track of the articles to be source.

The menu "Campaña" of the proposed purchasing prototype tool is a receptacle for articles that a purchasing department managed in past campaign or that is planning to manage in the remaining time of the current season. The main objective of this part of the purchasing tool is to give the buyers a top-level view of the season. The functionality of this menu is divided into two sub-menus. Figure 15 shows the first sub-menu "Status del Artículo" this is a to-do list for the buyers. In this part of the interface the buyer interact with the information of the planned articles for the season and track the progress of purchasing process for the new articles. The list shows how far the buyer is in planning and executing the purchasing process for a new article. The buyer needs to complete two processes to have a complete purchase plan. First, the buyer needs to forecast the new article's demand; by completing this step the interface will put a checkmark under the "Prediccion" tab. That means: the most important input in the purchasing optimization tool is ready to be use. Second, the buyer executes the actual purchasing plan. The purchasing-plan uses three inputs: The demand forecast for the new article ("Prediccion"), the estimation of the required exposition stock in stores, and the suppliers and transportation costs information. Then, using these three inputs, the optimization tool will be executed to generate a purchasing-plan proposal. Once the purchasing-plan proposal is completed a checkmark will be reflected under the tab "Plan de Compra" purchasing-plan.

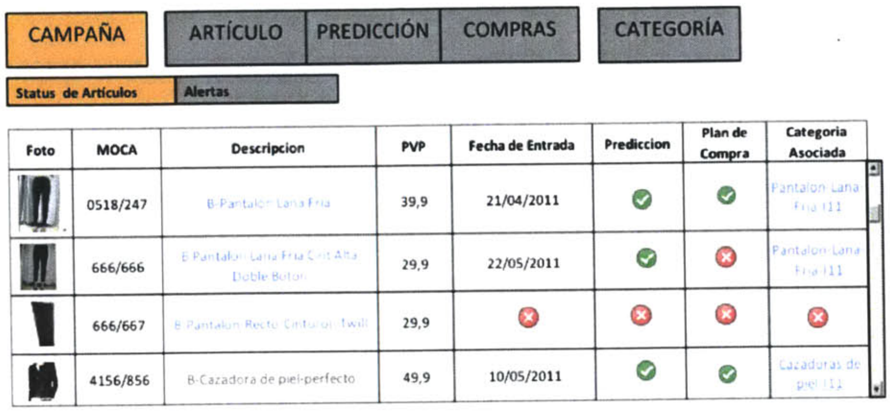


Figure 10: Purchasing tool prototype. Status of the articles offered in the season

Figure 15 shows the second sub-menu "Alertas", that is an alerts list. This list of alerts is linked to the suppliers' information: lead-times and capacity constraints. The list will rank the articles in urgency order. The buyers can review the list of alerts and anticipate the required work for the upcoming due-dates. The alerts sub-menu prevents the buyers from missing sourcing options. That is, the buyer is aware that the last day to order from the optimal sourcing origin is approaching; therefore, the purchasing-plan needs to be completed to lock-in the best sourcing option.

CAMPAÑA | ARTÍCULO | PREDICCIÓN | COMPRAS | CATEGORÍA

Status de Artículos | Alertas

MOCA	Descripcion	Proveedor	Circuito	Cantidad	Ultimo Dia Para Ordenar	Urgencia: dias restantes
666/666	B-Falda LANA FRIA CINT ALTA DOBLE BOT	5033	ASIA	60000	15-abr-11	11
999/999	B-CHAQUETA CORTA RAYON	6500	TURQUIA	35000	02-may-11	28
555/555	B-PANT RECTO CINTURON TWILL	7520	TURQUIA	35000	05-jul-11	92
111/111	B-PANT CHINO RAYAS	5556	FABRICAS	15000	01-ago-11	119
111/112	B-CAZADORA DE PIEL	5551	TURQUIA	15000	02-ago-11	120

Figure 11: Purchasing tool prototype: Season alerts

5.2 Purchasing tool prototype: The article

The article section of the prototype contains all the information related to a new article. This part of the prototype is use as a planning tool when buying a new article. The article section is subdivided into three menus. The first one defines a new article, the second has all the relevant information to forecast demand of the new article, and the third has the relevant information to develop a purchasing plan for the new article. This part of the prototype interface follows the purchase optimization process described in Figure 9.

Figure 16 shows the prototype screen to define a new article in the database. In this part of the tool, the buyer will upload all the available information for the new article. The first step is to assign a number to identify the new article in the database, the convention at Zara is to assign a number to the fitting pattern (Modelo) and a backslash (/) before a second number identifying the origin (Calidad). The first block information is describes the article, the purchasing department, the specific fabric characteristics of the article and the selling season. The main objective of this screen is to collect as much information as possible about the new article. The second section will link the article to a category. A category is a group of articles that share the same attributes: it can be a fabric, or a cutting style, or a feature that provides the article with a specific selling behavior. The third section selects the Mirror Curve for the new article and automatically displays the new article's offered sizes. The fourth section lists the planned colors for the new article. Finally, the sixth section selects the historical reference or references *r* for the new article. These historical references are also called comparable references; they refer to similar articles that can serve as a baseline in the demand forecast for the article. Once a historical reference is selected it the archived sales pattern will be pulled from the database.

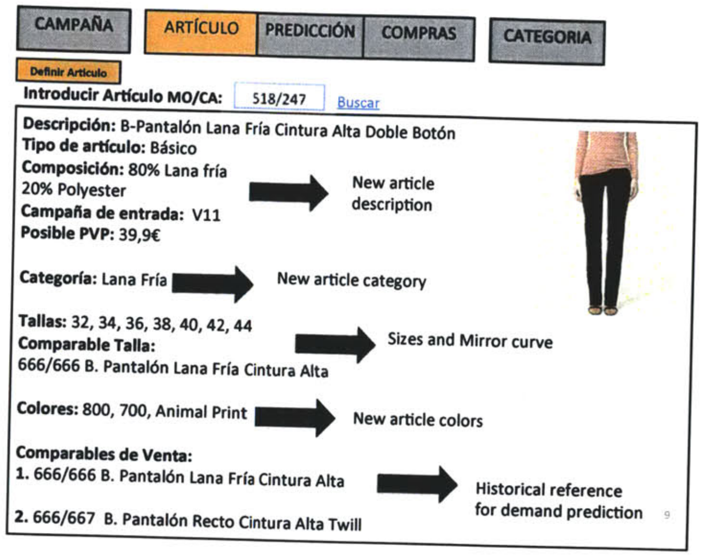


Figure 12: Purchasing tool prototype: Defining a new article

The prediction section has all the information required to estimate the demand for the new article. Section 4.2 explains the process for generating the point forecast of demand of new article. Figure 17 shows an example of the stages to convert the comparable historical reference POS data into demand data; once the demand data set is obtain the demand will be extrapolated to the 100% of the stores. The first stage shows only the POS pattern, which summarizes the observed sales aggregated by week. The first stage is the aggregated POS is the baseline information for the demand forecast of the new article. The second stage corrects the sales pattern to add the selling potential in the out-of-stock observations. Equation 1 and Equation 2 explain how to construct the demand dataset: $Demand_{rj}^{w}$; the demand dataset contains the POS sales dataset corrected for out-of-stock. The third stage, using Equations 5, 6, and 7 extrapolate the corrected demand pattern as if 100% of the existing stores at Zara would have received the historical reference: $\sum_{\forall\, k(j)} DesDemand_{r,k(j)}^{wl}$. On this screen, the buyer will observe the sales evolution and the inventory behavior of the historical reference. Using clustering analysis the demand information for all the stores will be extracted. Once the demand for all the stores is estimated buyers have the option to see the demand patterns for an average week. That is the demand patters without the inter-week seasonality from the historical reference Figure 17 shows the three stages that the buyer will see in the screen. The output of this screen is the demand dataset used as a baseline for the forecast for the new article.

Figure 13: Purchasing tool prototype: Historical reference

Equation 10 in Section 4.2 defines the methods to extract the inter-week seasonal factors $\tau_{R\,P}^{W}$. On the prediction screen the buyer can see the selling potential per week as if they all the weeks where average. The deseasonalized demand dataset of an article $DesDemand_{r,j\notin k}^{wl}$ is generated using Equation 6. Figure 18 shows an example of the interface that the buyer will when deseasonalizing the demand pattern.

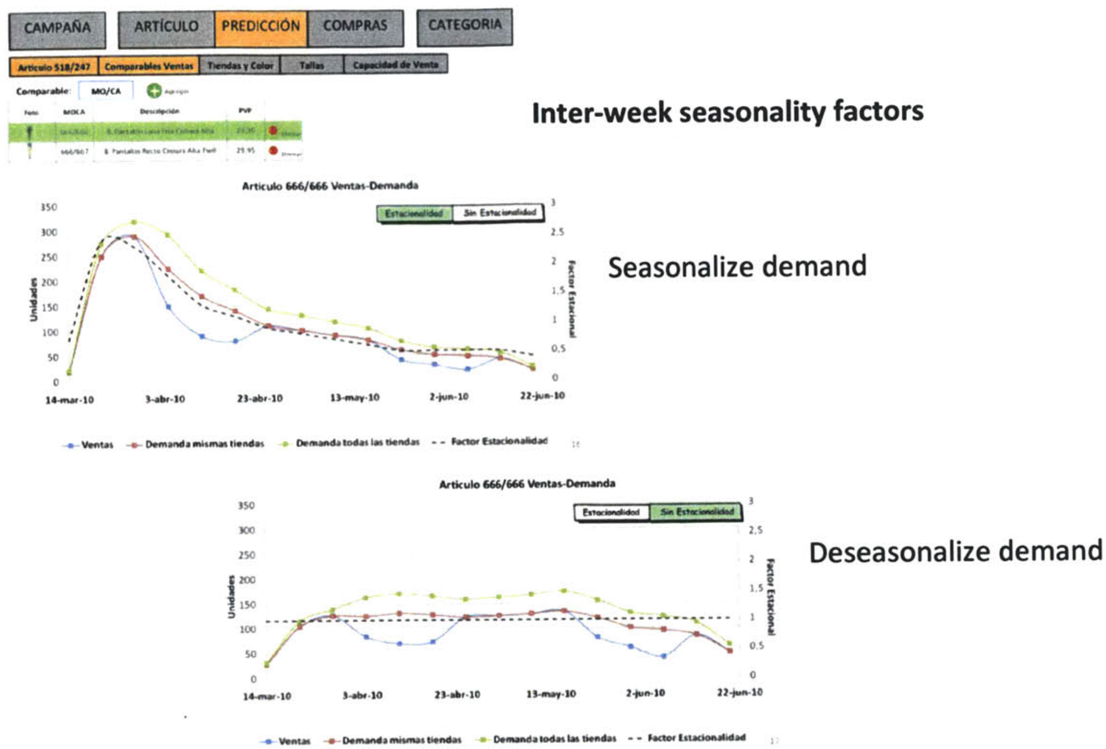
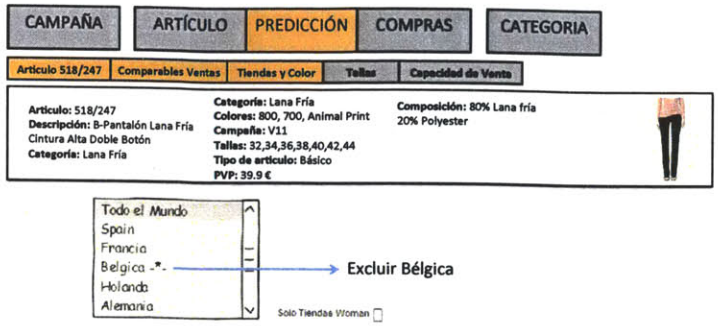


Figure 14: Purchasing tool prototype: Deseasonalized demand of historical reference

Section 4.4.1 explains the top-down forecast for color, the main is to partition the article level forecast into a color level forecast. The next sub-menu in the prediction section is the store and color sub-menu. This part of the tool automates the top-down forecast at the color level. On this screen the buyer will start updating the historical demand pattern with the specific information for the new article. This screen is an interactive screen, and the buyer will input the confidence level in each color and will select the network of stores that will receive the article in a given color. Figure 19 shows a graphical representation of the screen. The buyer can exclude the stores of a country or can predetermine a network of stores used frequently. The tool automatically partitions the article forecast using Equation11 and the information of the network of stores selected by the buyer.



	Alta	Media	Baja	Alta	Media	Baja	Alta	Media	Baja
	50%			30%			20%		
	800-Black			700-Pink			Animal Print		
TOP	☑	50%		☑	30%		☑	20%	
A	☑	50%		☑	30%		☑	20%	
B	☑	71%		☐	0%		☑	29%	
C	☑	62%		☑	38%		☐	0%	
D	☑	100%		☐	0%		☐	0%	

Figure 15: Purchasing tool prototype: Stores and colors

Similarly, Section 4.4.2 explains the top-down forecast for sizes. Figure 20 shows the interface to top-down forecast the sizes decision. The tool gives the buyer two options. First, the tool uses the historical sizes and cutting percentage of the subfamily. Some new articles bring new fittings to the stores; this new trends might not have a comparable historical fitting pattern (Mirror Curve). In this case, the buyer needs to use a "generic curve". The generic curve is the average sales for each size, using the information of all the articles in a subfamily and combined with the selected sizes for the new article, the percentage required to satisfy the expected demand for each size will be obtained.

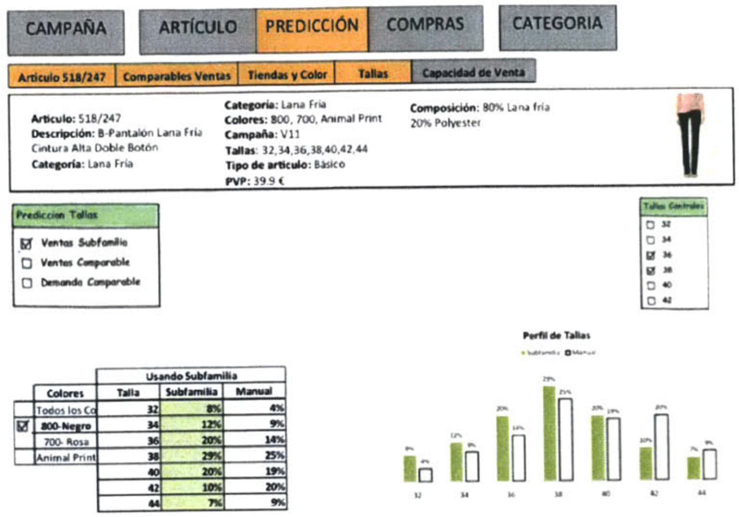


Figure 16: Purchasing tool prototype: Sizes partition Generic curve

The second option is to use a Mirror Curve from a specific historical reference. This option will convert the sales dataset for each size into demand for each size. Equations 11 and 12 converts= the observed sales of the historical reference per size into demand data $Demand^{w}_{rj\,s}$. Then, Equation 13 extracts from the sales pattern of that reference the specific weight for each size $W_{r\,s}$. These factors will top-down the forecast demand at the size level. Figure 21 shows the prototype screen for this part of the tool.

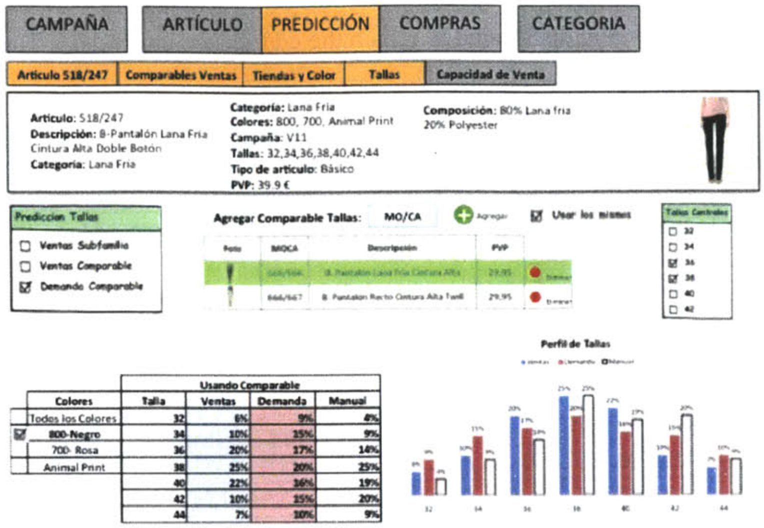


Figure 17: Purchasing tool prototype: Sizes partition Mirror Curve

Equation 8 from Section 4.2 creates the demand data set of the point forecast for a new article, including the seasonality of the launching date. Figure 22 shows the last screen of the prediction section. It summarizes the collected historical information into a demand prediction for the new article. The demand pattern combines expected sales with the most recent information about stores' situation. This interface helps the buyers simulate the best launching date for the new article. Based on the inter-week seasonality factors affecting the demand pattern of the new article and considering the number of articles of the same category, a potential cannibalization factor for the new reference can be obtained. This part of the prototype allows the buyers to rate the level of confidence in the calculation, this level is a qualitative rate based on the available information and how accurate the buyer thinks the prediction is. In addition to just providing a point forecast of demand for the new article, it is good to provide a statistical estimate of how much the actual demand is likely to vary around this single number. The most traditional approach in literature uses the MSE (Mean Square Error). MSE provides an estimate of the variance of the one-step forecast error. That is, one period ahead the last point. The error terms are considered normally distributed to obtain the confidence intervals. [36] The main purpose of this rating is to give the buyer some boundaries in the purchasing decision. This confidence intervals are important because they represent the upper and lower limit that demand can take; if the buyer finds a limit closer to zero, is important to question the results of the demand prediction, changing the launching date might provide different confidence intervals, this part of the tool allow the buyer to plan automatically for different launching dates.

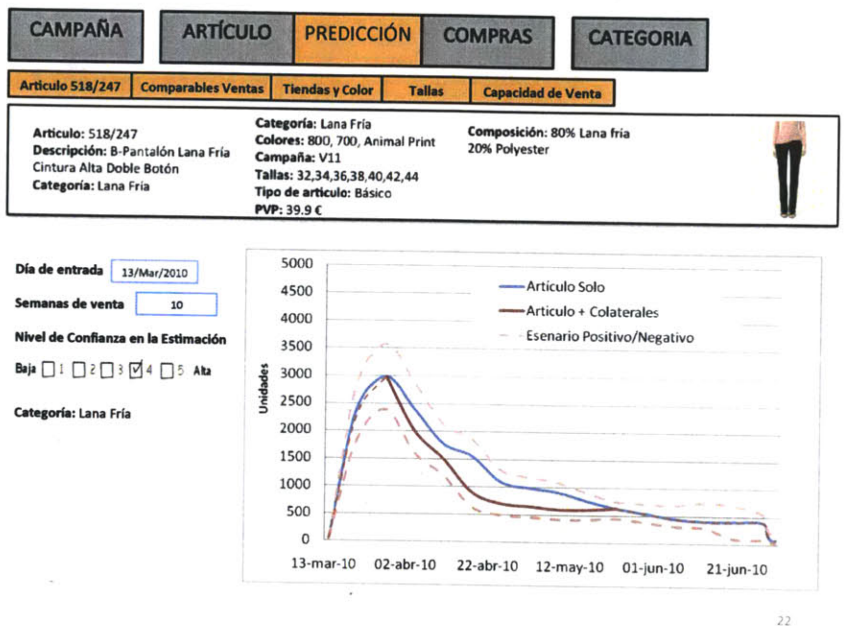
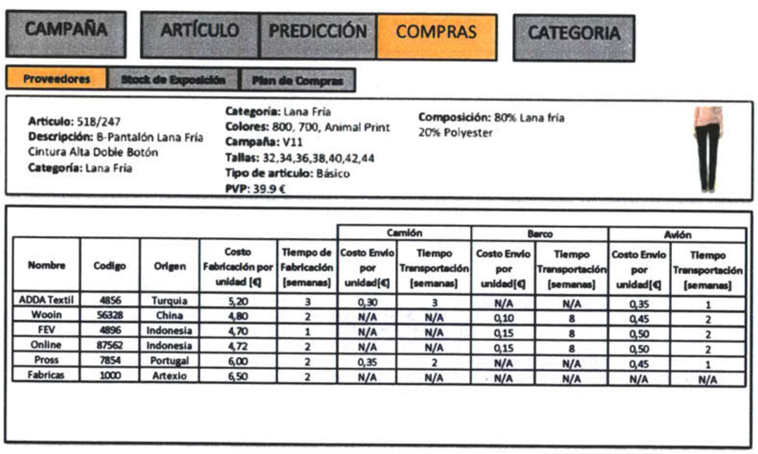


Figure 18: Purchasing tool prototype: Selling capacity

The next section is the purchasing plan section, the purchase-plan recommendation are to be determine using an optimization module that is beyond the scope if this thesis. In this part of the tool the buyer can record the different transactions held with suppliers, the relevant costs information and the output of the optimized purchases plan, that contains information of how much to buy from where. The purchase plan is subdivided into three sub-menus: the first sub-menu is the suppliers' information. The second sub-menu is the exposition stock. The third sub-menu is the purchases plan.

[36] (S. Makridakis, S. C. Wheelwright & R. J. Hyndman. Forecasting Methods and applications. John Wiley & Sons (1998) p.p. 88-93.)

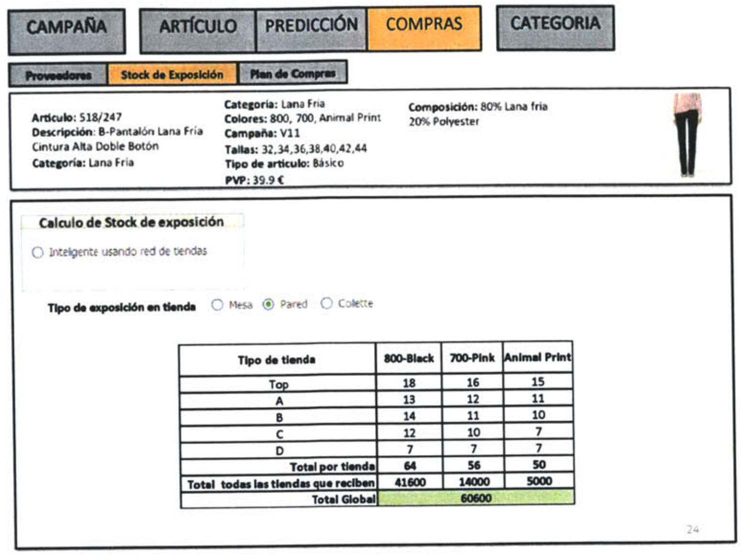
The first sub-menu: suppliers, on this screen the buyer can record all the transactions held with the suppliers. The basic information to subscribe a new supplier is the identification code of the supplier, and the country of origin of the article. Once the supplier is uploaded in the system, the buyer can fill the information about the interaction with supplier. The tradeoff with suppliers is between the total costs per unit and the delivering time. On this screen the buyer will record the potential transportation cost for each of the transportation methods. Using this system the buyers can visually understand the costs differences coming from the different transportation methods. Figure 23 shows the prototype for this screen.



					Camión		Barco		Avión	
Nombre	Codigo	Origen	Costo Fabricación por unidad [€]	Tiempo de Fabricación [semanas]	Costo Envio por unidad[€]	Tiempo Transportación [semanas]	Costo Envio por unidad[€]	Tiempo Transportación [semanas]	Costo Envio por unidad[€]	Tiempo Transportación [semanas]
ADDA Textil	4856	Turquia	5,20	3	0,30	3	N/A	N/A	0,35	1
Wooin	56328	China	4,80	2	N/A	N/A	0,10	8	0,45	2
FEV	4896	Indonesia	4,70	1	N/A	N/A	0,15	8	0,50	2
Online	87562	Indonesia	4,72	2	N/A	N/A	0,15	8	0,50	2
Pross	7854	Portugal	6,00	2	0,35	2	N/A	N/A	0,45	1
Fabricas	1000	Artexio	6,50	2	N/A	N/A	N/A	N/A	N/A	N/A

Figure 19: Purchasing tool prototype: Suppliers

The next sub-menu is exposition stock. This section interacts with the recorded information about the subset of stores receiving the article, and obtains the amount of initial stock required to fill the tables or walls that will display the new article. Figure 24 shows the prototype tool for this part, as it was explained in previous chapters, there is a minimum stock required to launch the selling season in the stores, however, buyers do not compute the required quantity, they only use a predefine number. This tool will compute the specific exposition stock requirement using the predefine specifications of each type of store: number of table or walls to cover, square-feet to cover with the new article, etc. This automatic computation will give the buyers the exact requirements per store type and per color to fill the stores.



Tipo de tienda	800-Black	700-Pink	Animal Print
Top	18	16	15
A	13	12	11
B	14	11	10
C	12	10	7
D	7	7	7
Total por tienda	64	56	50
Total todas las tiendas que reciben	41600	14000	5000
Total Global		60600	

Figure 20: Purchasing tool prototype: Exposition stock

The last sub-menu in the purchase plan is the purchase plan screen. This part of tool has two main objectives. The first objective is to summarize the profits obtained with the merchandizing of this article. The second is to summarize and record the purchasing decisions: suppliers, origin, total cost, transportation methods, and delivering dates. This part of the tool can be access at any time in the future and will serve as a guideline for future purchases. Figure 25 shows the prototype of this screen.

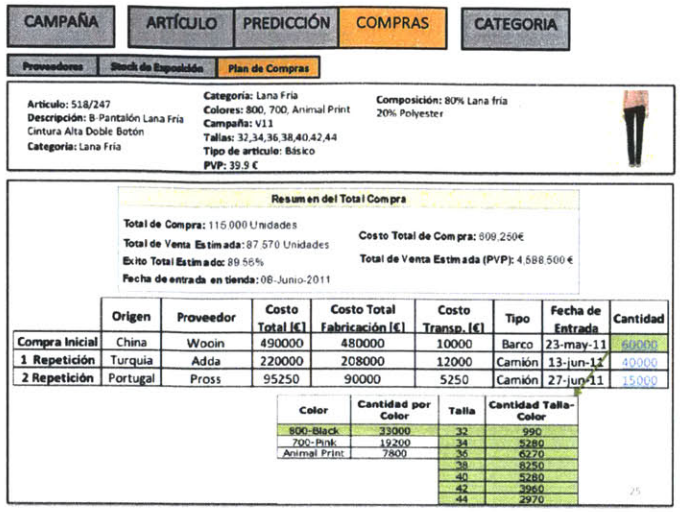


Figure 21: Purchasing tool prototype: Purchasing plan

5.3 Purchasing tool prototype: The category

The last section is the Category, "Categoria". A category is defined as a group of articles for which a buyer needs to do a top-down forecast. In Zara, it is common to define the forecast of a single item based on the total requirements of a category. A group of articles that share a characteristic in common (wool, golden, flower patterns, etc.) is most likely to have a similar group of articles sold in the past. The group of similar articles sold in the past is called a comparable historical category. This comparable historical category is the baseline to forecast the sales potential of the new season's category. The baseline demand of the historical category is modified to include the expected growth for that category. That is, if the fashion trend for the upcoming season emphasizes the use of cotton articles. The next season buyers expect an increase in the cotton demand. Therefore, the demand for the individual items that belong to the cotton category will increase. This part of the buyers' tool kit automates the estimation of the future demand for a category. The "Categoria" section is subdivided into three sub-menus. The first one is defining a Category. The second one is predicting the future performance of the category. The third one is the planning for the future of the category.

The first sub-menu is defining a category. This screen allows the buyer to link a group of articles together. That is, if two or more articles share a given selling attribute and they can be Top-Down forecast they need to be linked to the same category. Figure 26 shows the prototype for the category definition sub-menu. The buyers have the option of including or removing articles in the category and they can save their changes with a different category name if required.

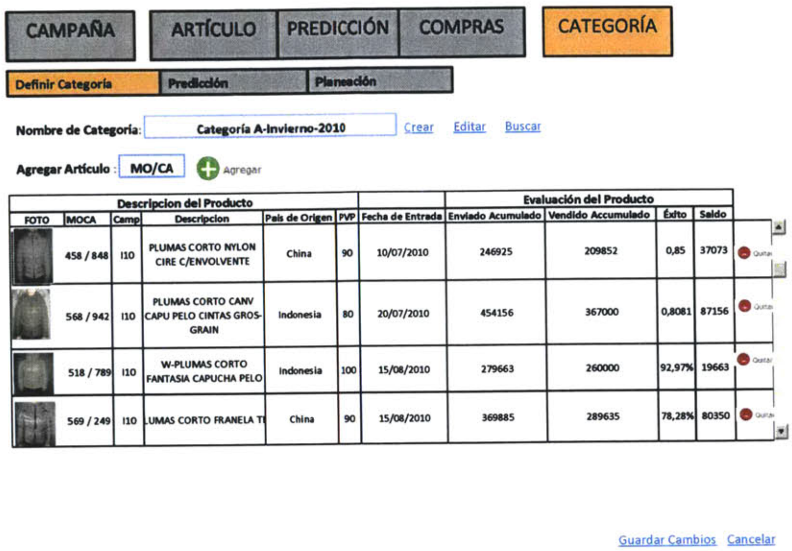


Figure 22: Purchasing tool prototype: Defining a category

The second sub-menu is planning for the future season of the category. Top management will decide the percentage increase aimed for each category. The tool gives the buyers two options: first use historical sales of all the articles and project the extra percentage on the sales figures. The second option use the corrected demand figures and project the expected growth percentage. Having obtained the individual demand figures for each article, this step is only an aggregation step. Figure 27 shows the expected screen to perform this estimation.

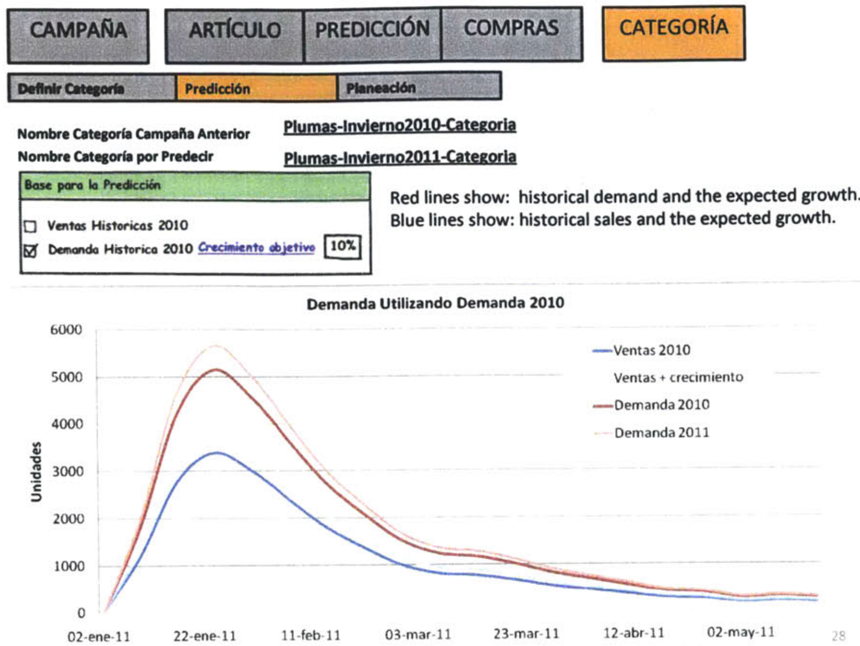


Figure 23: Purchasing tool prototype: Expected growth prediction for a category

The last sub-menu in the category section is the planning sub-menu. This is an interactive screen that allows buyers to graphically see the when is the best point in the season to introduce a new article for that category. Additionally it allows the buyers to see which categories are falling behind the expected growth, if this is the case they can introduce new articles or they can re-think the expected growth. This screen is also designed for allowing the articles to interact in the category. That is, if the articles have saturated the category expectation for a given week all the articles in that category in that week will be penalized for cannibalization between each other. In the screen each of the bar graphs represent an article. Using the selected launching date in the Article section, each is located in its corresponding week launching week in the graph. Depending on the forecasted demand of each article the size of the bar will be represented. The line graph represents the expected demand for the category. Figure 28 shows this part of the tool.

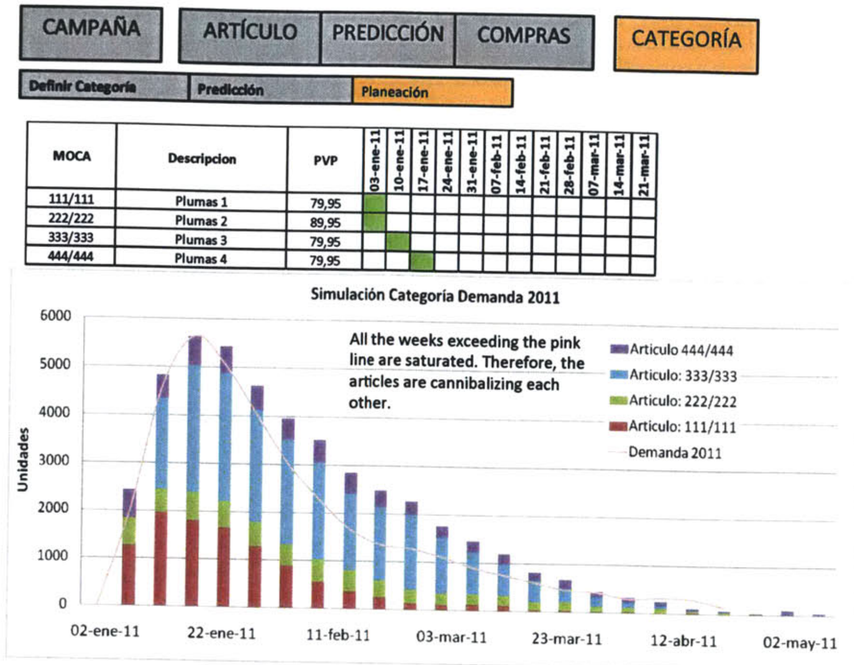


Figure 24: Purchasing tool prototype: Articles' interaction in a category

5.4 Chapter Summary

Chapter 5 describes the prototype purchase tool. The tool is a friendly-user interface that allows the buyers to control and interact with the database information related to the articles in the stores. The information is divided in a logical way: First it is evaluated the information at the season level. That is, the buyer will capture the overall perspective of the entire season. In this section the buyer manages all the articles and the purchasing progress for each article. The system will alert of the most urgent purchases to fulfill the requirements in the stores. The second section is the article. It captures the information related to the article. In this section the buyer defines the new articles and using historical information as it was described in Sections 4.2 and 4.3 of this document, the buyers will forecast the expected demand for the article. This prediction will be Top-Down forecasted at the color-store and the size level, using the process defined in Section 4.4 of this document. Finally the tool gives the buyers the expected purchasing limits based on the confidence in the prediction of the article. The last part of the article menu is the definition the purchasing plan. This purchasing plan utilizes an optimization model to generate recommendation to answer the questions: how much to buy, form where and when to buy; the optimization module is beyond the reach of this document. The third section is the category level. It is address the interaction between articles and allows the buyers to plan the launching dates of individual articles minimizing the expected cannibalization between articles in the same category.

Currently at Zara is building the infrastructure required to support this new implementation. As part of this case of study it was delivered a fully working prototype for a small sample of articles.

6. Conclusions

The purchasing and sourcing activities are two of the main competencies in the fashion industry. A fashion buyer is responsible for choosing the products that the company sells. The buyers oversee the development of clothes, which are targeted towards a particular market and price range. They need to communicate effectively with clothing suppliers and manufacturers. They need to visualize the budget and the profits that the company will make as well as how much they need to spend and to gain. All of these activities are characterized by a nice blend between art and science. The main objective of this case study for Zara is to provide the buyer's team with a set of tools to facilitate the scientific part of their job.

The purchasing process was reviewed, analyzed, and broken down into the smallest activity units. With the valuable input of buyers, managers, IT developers, and country managers, an enhanced and standard purchasing methodology was found. The proposed new methodology is a modularized process that, based on past performance of articles, predicts the future performance of a new article in the stores.

The first enhancement to the purchasing methodology is the conversion of POS datasets into demand information. By automatically extracting the out-of-stock information from the recorded sales data, the methodology enables buyers to base their estimations on demand information. This enhancement will reflect the real selling potential of an article, as if the stores displaying the article would have received an infinite amount of inventory. This calculation will correct the sales pattern for the supply chain errors observed in the past. The predicting performance of a new article will take into account those days for which the inventory was not enough. This insufficiency will be translated into purchase of a more accurate quantity, reducing the overall possibility of losing a sale for under-stocking and enhancing the brand recognition among the final consumers.

The second enhancement in the purchasing methodology is to update the obtained demand pattern with the specific expected stores' conditions at the purchasing time for the new article. That update involves first extracting the inter-week seasonality from the historical reference. Second, using cluster analysis, the deseasonalized demand pattern will be extrapolated as if all the stores had received the article. The third step is selecting the network of stores that will receive the new article. The fourth step is to re-seasonalize the demand pattern using the planned launching date of the new article. This enhancement allows the buyer to obtain the demand prediction corresponding to the exact subset of stores receiving the new article, incorporating the store openings and closures. In general, the aim is to capture the most up-to-date situation of the stores. This level of accuracy in the demand estimation will reduce the initial purchasing costs. In addition, having a good estimate of the required initial quantity will enhance the interaction between the logistics/distribution department and the purchasing/sourcing departments.

The third enhancement in the purchasing methodology is the Top-down forecast for the article's demand pattern. That is, the enhancement includes estimating the demand for all the colors and sizes of the offered article. Using the historical behavior per country per Sub-family and the expected demand pattern of the new article, the selling potential is broken down from the article level to the color and size levels. This breakdown will help the buyers give the supplier an immediate and accurate quantity to cut for each color and size. The Top-down forecast includes the most up-to-date information and captures the evolution of sizes and color preferences of the final consumers around the world.

The fourth enhancement in the purchasing process integrates the capabilities of the IT systems with the previously described standard methodology. The designed prototype of the purchasing GUI is described in Chapter 5. This web-based system gives the buyers the technological platform required to optimize their decision-making process. This enhancement represents an important technological advance in the decision-making process at the purchasing departments at Zara. The accuracy, efficiency, and general ability to use the database will significantly improve the performance of the purchasing group.

All the enhancements in the purchasing process are intended to relieve the buyers of all the burdensome activities that can be performed by a computer. Such enhancements leave buyers with time to focus on the strategic art of purchasing.

7. Bibliography

- F. Caro and J. Gallien, M. Díaz, J. García , J. M. Corredoira, M. Montes, J. A. Ramos and J. Correa. Zara Uses Operations Research to Reengineer Its Global Distribution Process. *Interfaces* Vol. 40, No.1, (2010), pp.77-84

- M. Fisher and A. Raman (2010). The New Science of Retailing: a how analytics are transforming the supply chain and improving performance. Harvard Business Press, (2010), pp. 5-23.

- M. Fisher and A. Raman. Reducing the cost of demand uncertainty through accurate response toearly sales. Operations Research, 44(1): 87-99, 1996

- S. Makridakis, S. C. Wheelwright and R. J. Hyndman (1998). *Forecasting Methods and Applications*. John Wiley & Sons. pp. 1-14, 186-194

- F. Caro and J. Gallien. Inventory management of a fast-fashion retail network. Operations Research, 58(2):257-273, 2010.

- J. A. Hartigan and M. A. Wong. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, Vol. 28, No. 1 (1979), pp. 100-108

- G. R. Murray, Jr. and E. A. Silver. A Bayesian Analysis of the Style Goods Inventory Problem *Management Science*, Vol. 12, No. 11, Series A, Sciences (Jul., 1966), pp. 785-797

- A. D. Cox and J. O. Summers. Heuristics and Biases in the Intuitive Projection of Retail Sales. *Journal of Marketing Research*, Vol. 24, No. 3 (Aug., 1987), pp. 290-297

- J. Song and P. Zipkin. Inventory Control in a Fluctuating Demand Environment. Operations Research, Vol. 41, No. 2 (Mar. - Apr., 1993), pp. 351-370

- J. E. Boylan and F. R. Johnston. Relationships between Service Level Measures for Inventory Systems. *The Journal of the Operational Research Society*, Vol. 45, No. 7 pp. 838-844, 1994.

- S.C. Graves and S.P. Willems. Strategic Inventory Placement in Supply Chains: Non-stationary Demand. Manufacturing & Sevice Operations Management, 10(2), 278-287, 2008

- T. Idia and P. H. Zipkin Approximate Solutions of a Dynamic Forecast-Inventory Model. Manufacturing and Service Operations Management, 8(4),407-425, 2006
- S.J Spear. (2009). *The High-Velocity Edge.* United States: Mc. Graw Hill
- E. A. Silver, D. F. Pike, & R. Peterson, (1998). *Inventory Management and Production Planning and Scheduling.* John Wiley & Son.
- D.B. Rosenfield. Demand Forecasting. The Logistics Handbook, pages 327-351, 1994.
- N. Tokatli. Global sourcing: insights from the global clothing industry the case of Zara, a fast fashion retailer. Journal of Economic Geography, 8(1): 21-38, 2008.
- D. F. Davis & J. T. Mentzer. Organizational factors in sales forecasting management. *International Journal of Forecasting.* Volume 23 Pages 475–495, 2007.
- A. F. Garro. New product demand forecasting and distribution optimization: a case study at Zara. MIT Thesis, 2010.
- M. A.Vega. Stock level optimization at the distribution center through improved supply management practices. MIT Thesis, 2009.

8. Appendix

8.1 Appendix 1: Stores clustering algorithm

The java routine show bellow describes the stores clustering algorithm implemented at Zara.

```
package com.inditex.logistica.zara1.hadoop.reducer;
import java.io.IOException;
import java.util.ArrayList;
import java.util.HashSet;
import java.util.List;
import java.util.Random;
import java.util.Set;
import org.apache.commons.math.stat.clustering.Cluster;
import org.apache.commons.math.stat.clustering.KMeansPlusPlusClusterer;
import org.apache.commons.math.stat.descriptive.moment.Mean;
import org.apache.commons.math.stat.descriptive.rank.Percentile;
import org.apache.hadoop.io.NullWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Reducer;
import com.inditex.logistica.zara1.hadoop.model.cluster.TiendaSubfamilia;
import com.inditex.logistica.zara1.hadoop.writables.VentaTiendaWritable;
public class ClusterizacionSubfamiliaReducer
extends
Reducer<Text, VentaTiendaWritable, NullWritable, Text> {
private static final int NUM_CLUSTERS = 133;
private enum KEY_FIELDS {
ID_CAMPANA, ID_TIPO_SECCION, ID_SUBFAMILIA;}
@Override
public void setup(Context context) {}
```

```
@Override
public void finalize() {}
@Override
public void reduce(Text key,
Iterable<VentaTiendaWritable> values,
Context context) throws IOException, InterruptedException {
Cluster<TiendaSubfamilia> clusterPercentil05 = null;
Cluster<TiendaSubfamilia> clusterPercentil95 = null;
//Lista de las tiendas que se van a clusterizar. Se han quitado los percentiles 5 y 95
ArrayList<TiendaSubfamilia> arrayTiendasClusterizar = new ArrayList<TiendaSubfamilia>();
//Lista de todas las tiendas
ArrayList<TiendaSubfamilia> arrayTiendas = new ArrayList<TiendaSubfamilia>();
List<Cluster<TiendaSubfamilia>> listaClusters = null;
ArrayList<Double> ventasTiendas = new ArrayList<Double>()
clusterPercentil05 = null;
clusterPercentil95 = null;
for (VentaTiendaWritable vTienda : values) {
arrayTiendas.add(new TiendaSubfamilia(vTienda.getIdTienda(), vTienda.getVenta()));
ventasTiendas.add(vTienda.getVenta());}
Percentile percentil = new Percentile();
Set<Double> numPosibleClusters = new HashSet<Double>();
double[] arrayVentas = new double[ventasTiendas.size()];
for (int i = 0; i < ventasTiendas.size(); i++) {
arrayVentas[i] = ventasTiendas.get(i);}
int percentil05 = (int) percentil.evaluate(arrayVentas, 5d);
int percentil95 = (int) percentil.evaluate(arrayVentas, 95d);
for (int i = 0; i < arrayTiendas.size(); i++) {
if (arrayTiendas.get(i).getVentas() < percentil95
&& arrayTiendas.get(i).getVentas() > percentil05) {
arrayTiendasClusterizar.add(arrayTiendas.get(i));
} else {if (arrayTiendas.get(i).getVentas() <= percentil05) {
if (clusterPercentil05 == null) {
clusterPercentil05 = new Cluster<TiendaSubfamilia>(
arrayTiendas.get(i));
clusterPercentil05.addPoint(arrayTiendas.get(i));
} else {
clusterPercentil05
.addPoint(arrayTiendas.get(i)); }}
if (arrayTiendas.get(i).getVentas() >= percentil95) {
if (clusterPercentil95 == null) {
clusterPercentil95 = new Cluster<TiendaSubfamilia>(
arrayTiendas.get(i));
clusterPercentil95.addPoint(arrayTiendas.get(i));
} else {
clusterPercentil95
.addPoint(arrayTiendas.get(i));}}}}
for (int i = 0; i < arrayTiendasClusterizar.size(); i++) {
numPosibleClusters.add(arrayTiendasClusterizar.get(i)
.getVentas());}
if (numPosibleClusters.size()==0) {
return;}
if (numPosibleClusters.size() <= NUM_CLUSTERS) {
if (numPosibleClusters.size() > 10) {
KMeansPlusPlusClusterer<TiendaSubfamilia> clusterer = new KMeansPlusPlusClusterer<TiendaSubfamilia>(
new Random());
listaClusters = clusterer.cluster(
arrayTiendasClusterizar,
numPosibleClusters.size() / 10, 1000000000);
} else {
```

```
if (numPosibleClusters.size() >= 1) {
KMeansPlusPlusClusterer<TiendaSubfamilia> clusterer = new KMeansPlusPlusClusterer<TiendaSubfamilia>(
new Random());
listaClusters = clusterer.cluster(
arrayTiendasClusterizar,
1, 1000000000);
}}} else {
KMeansPlusPlusClusterer<TiendaSubfamilia> clusterer = new KMeansPlusPlusClusterer<TiendaSubfamilia>(
new Random());
listaClusters = clusterer.cluster(arrayTiendasClusterizar,
NUM_CLUSTERS, 1000000000);}
listaClusters.add(clusterPercentil95);
listaClusters.add(clusterPercentil05);
int contadorCluster = 1;
for (Cluster<TiendaSubfamilia> cluster : listaClusters) {
if (cluster==null) {
continue;}
List<TiendaSubfamilia> listaTiendas = cluster.getPoints();
double[] ventasCluster =new double[listaTiendas.size()];
int j = 0;
for (TiendaSubfamilia tiendaSubfamilia : listaTiendas) {
ventasCluster[j] = tiendaSubfamilia.getVentas();
        j++;}
Mean m = new Mean();
double media = m.evaluate(ventasCluster);
for (TiendaSubfamilia tiendaSubfamilia : listaTiendas) {
context.write(NullWritable.get(), new Text(key.toString() +
","+contadorCluster+","+tiendaSubfamilia.getIdTienda()+","+
tiendaSubfamilia.getVentas()/media))}
contadorCluster++;}}}
```

8.2 Appendix 2: Weight per number of colors.

The table bellow shows the summary of number of colors and color confidence per number of colors conducted using the 2010 demand for all the articles sold during the spring-summer collection.

No. De Colores	Nivel de confianza historico	
1	1	1,00
2	1	0,42
	2	0,58
3	1	0,25
	2	0,33
	3	0,42
4	1	0,17
	2	0,22
	3	0,27
	4	0,34
5	1	0,13
	2	0,16
	3	0,20
	4	0,23
	5	0,28
6	1	0,09
	2	0,11
	3	0,14
	4	0,18
	5	0,21
	6	0,26
7	1	0,07
	2	0,09
	3	0,10
	4	0,14
	5	0,17
	6	0,19
	7	0,24
8	1	0,05
	2	0,07
	3	0,10
	4	0,11
	5	0,12
	6	0,13
	7	0,17
	8	0,24
9	1	0,02
	2	0,05
	3	0,08
	4	0,10
	5	0,11
	6	0,13
	7	0,14
	8	0,16
	9	0,21
10	1	0,04
	2	0,05
	3	0,06
	4	0,07
	5	0,09
	6	0,10
	7	0,11
	8	0,12
	9	0,13
	10	0,23
11	1	0,04
	2	0,05
	3	0,06
	3	0,06
	4	0,07
	5	0,09
	6	0,10
	7	0,11
	7	0,11
	8	0,14
	9	0,17