Order Assignment Strategy on a Luxury Digital Marketplace

Maria do Carmo Guedes Oliveira Ferreira da Silva

Master Thesis

Supervisor at FEUP: Prof. Bernardo Almada Lobo

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Master in Engineering and Industrial Management

2019-07-01

Abstract

As online marketplaces popularity continues to grow, finding ways to improve the matching algorithm between buyers and sellers in this sector is becoming increasingly important. Relative to brick-and-mortar retailers, online retailers have the potential to offer more options to their customers, with respect to both inventory and delivery times. In this thesis, the impact of smarter, forward-looking fulfillment decisions in an online marketplace environment is studied. This project aims at developing a decision-making model that will improve the current order assignment process at an optimal cost to deliver on customers' increasing expectations. The focus is on the fulfillment problem, namely, deciding the seller and shipping type to fulfill a customer's order when several options exist. This project provides a new approach to develop a successful business case using business analytic and programming techniques, which helps decision-makers to take more informed fact based decisions. The results of this project were obtained through discrete simulation of order assignments, which helped the company identifying potential revenue and efficiency gains otherwise unobtainable. First, the current order assignment decision process is assessed. Second, ranking and selection methods are explored, as well as key variables that can be included in the decision making of the novel algorithm. To design a decision-model that satisfies all business requirements, research and interviews were conducted in order to include all necessary variables and restrictions. Third, different algorithm compositions are tested for the order assignment problem. Fourth, estimate the impact of the new design solution against current process. At the end, results are discussed and a next steps plan is proposed.

During execution three main challenges were identified: 1) historical records of past decisions are not available and need to be reconstructed, in order to make a proper diagnosis of the AS-IS situation; 2) this is a classic multi-criteria problem, in which decision-makers are confronted with conflicting business objectives. The ambition to balance profitability and customer experience requires/presupposes a clear and measurable definition of service quality and how it is valued among customers; 3) there is resistance to change, as it is unknown how sellers will be affected. Clearly, the new criteria for order assignment identified through literature reviews and benchmark in this thesis are very important in a marketplace context. The proposed method classifies the criteria into three groups (profitability, customer experience, seller success). Some perspectives, such as profitability and delivery speed were not contemplated prior to this project in the seller selection process. The analysis conducted during this project reveal that using a different selection strategy can lead to dynamics that result in revenue increase and service quality improvement. A simulation program, that replicates past decisions, developed in the scope of this dissertation shows that approximately 50% of the products ordered are located in more than one single supplier, and on average each order that can be fulfilled by more than one alternative is disputed by 5.3 competitors. After showing that the order assignment tool

is valid by replicating past decisions different strategies were tested on company historical data, showing a potential of 1.69% increment in net results as compared to the AS-IS process. Other proposed strategies show small improvements in both profitability and service metrics.
With regards to practical implications, first, the order assignment tool provides solid understanding on how order allocation process currently works. With this tool it is possible to evaluate the impact of this project and to predict the effect of changing the business rules of the decision process on the sellers' performance. This tool helps managers to make more informed strategic decisions through what-if analysis. The level of sophistication and detail of the tool allows to estimate the impact of different scenarios with high accuracy and reliability. The proposed strategies have shown significant improvements in terms of profitability and delivery time (service related metric), which led to budget approval from the company for project roll out next year. This tool also provides important business insights to be used by different business departments. The exhaustive study of the order allocation process results in a compilation of next steps for the implementation of the recommended improvements and future works initiatives that can further sophisticate the decision.

Estudo da estratégia de alocação de encomendas num marketplace de luxo

Com o aumento da popularidade dos marketplaces nos últimos anos, arranjar formas de melhorar o algoritmo de associação do comprador ao vendedor mais adequado torna-se cada vez mais importante. Relativamente às lojas físicas, o retalho online tem o potencial de oferecer mais opções aos seus clientes, quer relativamente à oferta quer relativamente a prazos de entrega. O objetivo desta dissertação é desenvolver um modelo de tomada de decisão que melhore o atual processo de alocação de encomendas a um custo ótimo, sem comprometer as expetativas cada vez mais exigentes dos clientes. O foco do projecto é a alocação de encomendas, decidir qual deverá ser o vendedor e o tipo de shipping quando existem várias opções possíveis. Este projecto oferece uma nova abordagem de desenvolvimento de um business case utilizando técnicas de análise e programação que permitiram aos decisores tomar decisões mais informadas e fundamentadas em factos. Os resultados deste projecto foram obtidos através da simulação discreta de alocações de encomendas, o permitiu à empresa identificar ganhos potenciais de receita e de eficiência que de outra forma seriam dificilmente estimáveis. Em primeiro lugar, analisou-se o estado do algoritmo atual. Em segundo lugar, foram explorados métodos de ordenação e selecção e foi feito um levantamento das variáveis chave para avaliar vendedores. Em terceiro lugar, foram testadas diferentes composições do algoritmo de alocação de encomendas. Em quarto lugar, foi estimado o impacto da nova solução relativamente ao modelo atual. Finalmente, faz-se uma discussão dos resultados obtidos e propõe-se um plano de próximos passos.

Durante a execução foram identificados três desafios. 1) Não existiam registos históricos das tomadas de decisão as quais tiveram de ser reconstruídas por forma a fazer um diagnostico correto da situação atual (AS-IS); 2) Este é um problema clássico de tomada de decisão com múltiplos critérios em que o decisor se confronta com objetivos de negócio conflituantes. 3) Existe resistência à mudança, dada à incerteza relativamente ao impacto nos vendedores.

Os novos critérios identificados nesta dissertação através da bibliografia revista e de benchmark efetuado são muito importantes no contexto do marketplace. O método proposto classifica os critérios em três grupos: rentabilidade, experiência do cliente e sucesso do vendedor. A ótica de rentabilidade e rapidez de entrega não eram consideradas antes deste projecto no processo de escolha dos vendedores. A análise efetuada durante o projecto evidência que uma estratégia de seleção diferente pode gerar uma dinâmica que se traduzira em aumento da receita e a melhoria da qualidade do serviço. Um programa, que replica decisões passadas desenvolvido no âmbito desta dissertação mostra que aproximadamente 50% dos artigos encomendados se encontram em mais do que um vendedor e que em média cada encomenda é disputada por 5.3 vendedores. Depois de demonstrar que o simulador é válido - através da replicação de decisões passadas, testaram-se, com base nos novos critérios identificados, diferentes estratégias de alocação nos dados históricos da empresa, que evidenciaram um aumento potencial até 1.69% dos resultados liquidos da empresa relativamente à política atual de alocação. Outras estratégias sugeridas permitem aumentar lucros, bem como outras métricas de qualidade de serviço e tempo de entrega.

No que diz respeito a implicações práticas o programa oferece uma sólida compreensão de como funciona atualmente o processo de alocação de encomendas. Com esta ferramenta é possível avaliar o efeito das alterações na evolução do negócio dos participantes do marketplace e e prever o impacto deste projecto. Esta ferramenta permite aos gestores tomarem decisões estratégicas mais fundamentas através de análises what-if. A sofisticação e o detalhe da ferramenta permitem estimar o impacto de diferentes cenários na rentabilidade e em métricas de serviço, com elevada precisão e fiabilidade. As estratégias propostas evidenciaram ganhos significativos em termos de rentabilidade e prazos de entrega o que levou à aprovação orçamental por parte da empresa para implementação do projeto no próximo ano. O simulador também oferece informação de gestão relevante que pode ser utilizada por várias áreas da empresa. O estudo exaustivo do processo de alocação de encomendas levou à identificação dos próximos passos para a implementação das melhorias recomendadas e de iniciativas de trabalho futuras que podem sofisticar ainda mais o processo de decisão.

Acknowledgements

First and foremost, I would like to thank my advisor Professor Bernardo Almada Lobo for all his support and insights throughout the project. I truly admire his willingness and ability to discuss all kinds of problems, always striving for the best. To my supervisor at Farfetch, Filipe Aguilar, I am grateful for the opportunity to work on this exciting challenge and for his inspiring and valuable feedbacks. I would like to thank Martinho Brito for his great input. I truly appreciate his wisdom, enthusiasm and friendship at key points of the project. A special thanks to Fábio Moreira, who was leading the project at the company, for all the knowledge shared, his ideas and patience to teach. The three were indispensable and great mentors. The many problem-solving meetings with them allowed this project to be a great challenge and remarkable learning opportunity for me. Thanks to all my teachers at Universidade do Porto and colleagues at Farfetch for providing me the academic and business knowledge as well as helping me develop technical and soft skills. A special thank to Operations Strategy team for all their friendship, support and good humor. Farfetch is a great place to work, filled with amazing people. At last, I thank my parents and family for instilling in me by example the ambition to be better and for all their extraordinary support throughout my academic years. To all my good friends thank you for your interest and always being there for me.

Contents

Acronyms and Symbols

AHP	Analytic Hierarchy Process
CAC	Customer Acquisition Cost
CLV	Customer Lifetime Value
DDP	Deliver Duty Paid
DDU	Delivery Duty Unpaid
DES	Discrete Event Simulation
EDD	Estimated-Delivery-Date
FF	Farfetch
KPI	Key Performance Indicator
MCDM	Multicriteria Decision Making
NPS	Net Promoter Score
NS	No Stock
PDP	Product Display Page
PLP	Product List Page
QFD	Quality Function Deployment
SKU	Stock Keeping Unit
SOS	Speed of Sending
TIT	Time in Transit
VAT	Value Added Tax
WI	Wrong Item

List of Figures

List of Tables

Chapter 1

Introduction

An online marketplace is a website or app that facilitates shopping from many different suppliers (Hagiu, 2009). As the popularity of this business model continues to grow, finding ways to improve the matching algorithm between buyers and sellers is becoming increasingly important. In this thesis, the impact of smarter, forward-looking fulfillment decisions in an online marketplace environment is studied. This project aims at developing a decision-making model that will improve the current order assignment process at an optimal cost to deliver on customers' increasing expectations. The focus is on the fulfillment problem, namely, deciding from which seller(s) to fulfill a customer's order when several options exist.

One of the biggest advantages of a marketplace is that the company manages to offer a much broader assortment than any other traditional retailer because there is no need to invest for stocking the products that might eventually sell. This risk of purchasing and storing products is distributed over small boutiques that integrate the platform (Hänninen et al., 2017). With this great competitive advantage big challenges come along, that need to be addressed with cutting-edge technology and disruptive thinking. On the other side, the marketplace operator has, in theory, limited control of supply volume and distribution which increases the need for operational excellence. Another common downside of marketplaces is that products are being offered by many different boutiques and the delivery quality and speed is not uniform.

This work was motivated by the company Farfetch, hereafter referred as FF. The company has more than 3000 employees with most of its operations established in Portugal. One of the business units of Farfetch is the online marketplace for fashion. This marketplace is transactional, meaning that revenues only come once an exchange is made. Farfetch, as operator of the marketplace, does not own any inventory and is creating value by presenting the inventory from several fashion boutiques, brands and department stores located across the world to online customers and facilitating the transaction (Hänninen et al., 2017). The delivery options, including same-day delivery in 19 major global cities and store-to-door in 90 minutes (F90), are outsourced to carriers that are integrated in the platform of the company. Apart from order cycle management, which takes care of information flow and coordination between sellers, carriers, payment and fraud check services, this business unit is responsible for customer/seller support and content creation (imagery and product descriptions) for Farfetch website. Farfetch Marketplace platform accounts for over 90% of company revenue today and has created a community of 2.3 million customers in 190 countries and more than 1,000 luxury goods sellers, which represent around 1,600 stock points spread in 52 countries (Danziger, 2018). Farfetch platform offers the opportunity for boutiques, that used to be limited to their geographic market, to scale and absorb demand across the world. For that exposure boutiques renounce to part of their profit margin and pay a fixed commission to FF. Farfetch has a diverse set of competitors in the online luxury landscape.

1.1 Problem Description

Farfetch is considered a homogeneous marketplace as opposed to search marketplaces, as customers do not decide from which boutique they receive the orders (Golden, 2017). When a client opens on the website the product list page or product display page to select a size, a single product with a unique price is offered. However, the exact same product online can be available in multiple stock points across the world, which might be privately-owned boutiques or brands' warehouses that price items differently. In this case, an algorithm has to run in real time in order to search for the available fulfillment options and select from which seller the item will be shipped.

The challenge for the company is to decide for every order, when the desired item exists at more than one location, from which vendor the item should be shipped and guarantee a standardized quality. Moazed and Johnson (2016) define matchmaking as one of the core functions of a platform business, where the right consumer has to be connected with the right seller. The process of deciding which seller should be favored in case there is more than one seller offering the SKU is the DNA of a marketplace business. The process of selecting a seller to fulfill an order will be referred as order assignment process.

Currently, the business is selecting the seller filtering sequentially the possible boutiques based on five criteria: price, geographic location, a calculated score that considers fulfillment time, probability of order rejection and customer feedback rate and finally the product creation date. This logic runs for every item of the product list page so that a final price can be displayed when the customer is navigating and searching for products. The company's strategy is designed around customer centricity being one major area of focus the improvement of the delivery experience. However, specially for a listed company, where investors expect positive results, it is important to create conditions for achieving sustained growth by maximizing profitability. The choice of the seller can influence both factors: the seller location and ability to prepare orders faster impacts delivery speed. At the same time, costs are different depending on shipping and duties costs.There are inefficiencies and key concerns raised around the current algorithm that need to be explored. In the presence of multiple solutions, it is worth evaluating how different decision-models can impact business results and stakeholders' satisfaction. For a small business the performance of this algorithm was not the main concern, as there was little supply and a simple and intuitive logic was capable of making obvious decisions. Currently, the marketplace has a network of over one thousand suppliers across fifty two countries, consequently, the decision process must be improved. The ambition is to have a smart selection process that evaluates the alternatives and returns the best possible option for a specific order.

There are several hurdles of the current decision process. First, there is no record of past decisions to evaluate the algorithm performance. Second, by changing the currently used dispatching rules for profitability maximization strategies the results of the company can be improved, but negative or eventual positive impact on customers are not easily quantifiable. Third, testing and changing the algorithm's criteria may negatively impact some sellers' growth and until this project the company did not have simulation capabilities. The fear of drastic change on boutiques sales and unbalanced workload has blocked optimization initiatives. Fourth, at this point, nobody at the company can quantify accurately how many of the items ordered were located in a single stock point for example and it is unclear which criteria are decisive for a seller to capture a certain sale. Lastly, current technological infrastructure is not ready to change and integrate the new criteria. Since there are many initiatives on the road-map and the impact of this initiative could not be measured until this project, changes to the order assignment algorithm were not planned for the next years.

The main task is to assess the current algorithm performance given the current supply and demand and deliver an analysis on the amount of orders that follow this decision path and quantify the size of the opportunity. The ambition is to give visibility over possible outcomes based on certain scenarios of demand and supply. The final goal of this project is to raise problems and propose a method for the design of the order allocation.

1.2 Project Goal and its Relevance for the Company

The challenge of Farfetch's Supply Chain Team is to guarantee that delivery operations are run efficiently in order to achieve simultaneously top service quality, high operating profits and better management of the distribution network and supply. Therefore, orders' fulfillment management and optimized network design is crucial for success. Most initiatives are associated with shipping and return costs reductions. Other teams are responsible for projects that boost demand generation or guaranty the quality in the end-to-end shopping experience, which maximizes client satisfaction. The order assignment process affects all areas at Farfetch – from Customer Service, Commercial, Sales, Finance and Marketing to Fulfillment Teams. The expected outcome of this project is the implementation of an efficient decision-process that maximizes Farfetch profit, customer retention and supplier satisfaction. Naturally the sellers have to be engaged and committed to the business in order to take the risk of purchasing items in advance to merchandise on the online platform. Changes to the current algorithm logic may impact the distribution of sales volume among sellers, and ultimately, their profitability. Given the above mentioned roll-out barriers a simulation tool to predict the impact of the proposed changes in the business is of great importance in order to convince all stakeholders of the benefit, obtain budget and capacity for the project and foster a successful implementation. This project proposes a data-driven Discrete-Event Simulation (DES) approach to provide decision support to C-level-management concerning the strategy deployment of the routing algorithm. The developed tool will allow managers to compare different strategies and make more informed decisions.

1.3 Methodology

The common steps of a problem solving process are depicted in Figure 1.1. First, understand and define the problem in the business context and think about the impact. Second, structure some hypotheses and raise the key elements of the problem. Third, prioritize the initiatives. Fourth, plan the project, collect data and prepare the analysis of the AS-IS situation (Diagnosis-phase). Based on previous gathered information design a solution proposal and predict impact and implications. Finally, develop a recommendation listing the requirements for implementation.

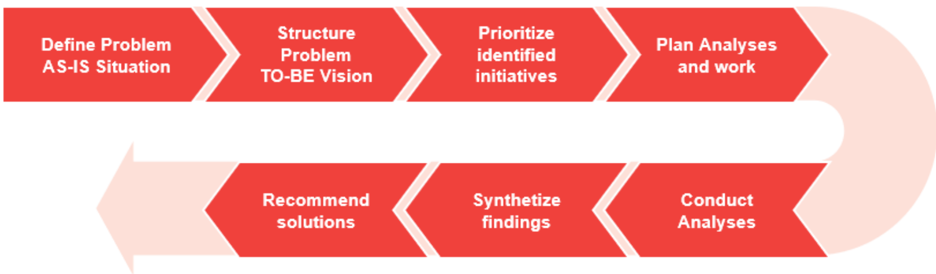


Figure 1.1: Based on McKinsey seven step problem-solving process

Following this approach, efforts were made to design a solution for the order allocation problem aligned with the current business needs. To deliver an adequate solution, research and meetings with the involved teams were conducted to find out which stakeholders are affected, and the underlying reasons. For example the algorithm response time has to be high as it affects the online website performance. Another important aspect is that forecasts that are sent to the boutiques using historical data will no longer be valid if the logic changes. Another important step of the project was to become acquainted with the data warehouses and business metrics in other to perform the required analysis that lead to a solution proposal. To guarantee the successful implementation of the proposed strategy and convince all the stakeholders of its merits, part of the project was to create a process and the appropriate metrics to measure and communicate the efficiency of the solution and respective KPIs.

To support the diagnosis phase and help designing a solution, a tool capable of recreating order assignments was developed. This allowed to experiment different strategies and to decide on the best approach to be implemented based on the scenarios outcome. A significant share of this project was spent in developing this tool and test the proposed alternative solutions for the order assignment. This simulator requires high level of sophistication as it is only useful in case the decision scenarios are realistically generated. This means that for each simulated order all the sellers, that in reality had available stock on that moment, have to be considered in the assignment decision. The waterfall methodology was used to develop the order assignment tool (see Figure 1.2). In an early stage of the project the requirements had to be determined and the program infrastructure designed. For implementation data collection and treatment plus variables modeling was necessary. Several tests were made to assure the accuracy of the data and prioritization rules were included in the program. Additionally, performance tests to increase productivity were conducted.

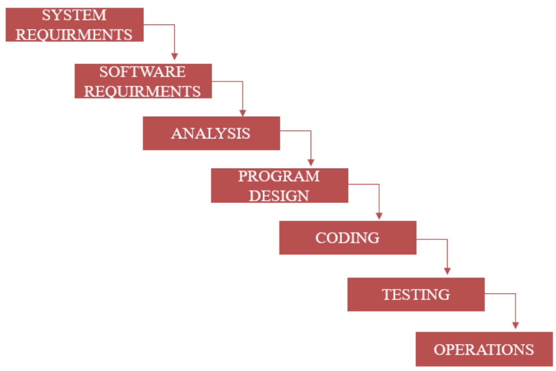


Figure 1.2: Waterfall Methodology. Source: Royce (1970)

1.4 Thesis outline

This thesis is organized into six chapters as follows. Chapter one introduces the project, the company that motivated it, and the context. Chapter two presents the results of theoretical and practical research considered relevant to deliver a solution to the challenge: The first three sections aim to understand the consumer and the challenges that arise in the luxury industry and e-commerce context. The last section of the literature review looks at the previous work developed for adopting order allocation strategies in similar business contexts. Concluding this Chapter, the opportunity to use simulation to address similar problems is explored. Moreover, Chapter three makes a diagnosis of the as-is situation to identify improvement opportunities and the requirements to meet the business vision for this problem are exhaustively presented. In Chapter four a solution approach for the addressed problem is proposed and a detailed overview of the processes and modelling issues involved in the simulator construction is offered. Chapter five presents the obtained results with the simulation tool and proposed solution approach. Finally, Chapter six is a summary and a discussion of the findings of this thesis and identifies future works.

Chapter 2

Literature Review

The following chapter is the result of the research conducted with the ambition of better understanding the clients needs and requirements of e-commerce order fulfillment in the luxury industry.

2.1 Business Overview

2.1.1 Luxury customers and their expectations

For premium markets, where value is associated with experience, the customer needs become more relevant. Luxury retail has typically been associated with physical stores that offer an exquisite shopping experience. Nevertheless, the market landscape for luxury goods is changing and online sales are expected to grow from 8% in 2016 to 19% in 2025 (in Linda Dauriz, 2014). Driving this change is the shift in the consumers demography: "Collectively Millennials and Generation Z will represent 40 per cent of the overall luxury goods market by 2025" (Deloitte, 2018), making them the most important audience to attract. There is now an expectation that consumers can access luxury goods in a similar way to how they access other goods (Arnett, 2019). However, there is a major challenge to replicate the luxury shopping experience online and convince customers to switch: as opposed to mainstream goods consumers studies, Kiang (2012) findings reveal that design and product information have a significant effect on service quality perception for luxury consumers, because they value contact with sales associates, see/touch the product and are afraid of counterfeit items (Google, 2013). Additionally, one of the main reasons for choosing traditional over online stores is the possibility to leave the store with the desired product. Despite many reasons to buy offline, one of the main motivations underlying customer inclinations to adopt online shopping is convenience. There are many types of shopping convenience, such as access, search, evaluation, transaction, post-purchase and possession(Jiang et al., 2011). The improvement of order assignment algorithm in this project can influence the perception of possession convenience. According to Jiang et al.

(2011) this kind of convenience is perceived by delivery offered, on-time delivery, delivery time uncertainty, delivery change notification, damages in the product, attitude and performance of deliverymen. Price inconsistency is also mentioned as hurdle of the transaction convenience. All this aspects should be considered in the design phase of the new assignment algorithm. Marketplaces, due to their business model characteristics, have an opportunity to offer substantially speedier deliveries globally (due to fragmented supplier base) and thus have a major advantage against monobrand's or even department stores' online business (Arnett, 2019). On the other hand, there is less control over product conditions, order accuracy and delivery speed consistency. For this reason, order routing and seller selection process plays an important role on achieving convenient e-service delivery.

In addition, luxury customer's tend to be more sensitive to the unboxing experience and thus the ability of a seller to prepare orders in a pleasant way is crucial for retention. Packaging has been used as a powerful e-commerce marketing tool. Specially for luxury consumers, that usually have great attention to detail, receiving the desired product in an attractive packaging improves shopping experience. It has been proven that people are naturally drawn to products and packaging that looks good. A study from Hubert et al. (2013) found that attractive packaging stimulates the reward-seeking areas of the brain that are associated with impulse purchasing.

According to an Exploratory study conducted by Jiang et al. (2011) managers put a lot of emphasis on the high cost of providing convenient online shopping. Efforts to meet or exceed expectations, creating unique value propositions and investing in marketing activities to retain and attract new customers are heavy costs for luxury e-commerce players. The challenge is to find the balance between convenience and operational cost, which must be included when designing the selection-model.

2.1.2 Customer satisfaction value – building up trust and loyalty

One of the strategies followed by leaders in e-commerce business is the customer-centric strategy. Customer centricity is defined "as the ecosystem and operating model that enables an organization to design a unique and distinctive customer experience. This architecture enables the business to acquire, retain and develop targeted customers efficiently for the benefit of employees, customers and stakeholders". (Leather, 2013) This companies derive revenue from the creation and sustenance of long-term relationship with their customers. The ambition is to maximize the customer life time value (denoted as CLV) that is generally defined as expressed in equation 2.1 (Gupta et al., 2006). CLV is positively influenced by the price paid by a consumer (p) and by the probability of repeat purchase (r). On the other hand the direct costs of servicing that customer (c) and acquisition costs reduce the lifetime value.

$$\mathrm{CLV} = \sum_{t=1}^{T} \frac{(p_t - c_t) \times r_t}{(1+i)^t} - \mathrm{AC} \tag{2.1}$$

There are many studies that have looked at the impact of customer satisfaction on repeat purchase, retention, brand loyalty and positive worth of mouth. Customer satisfaction influences repurchase intentions while dissatisfaction is the leading reason for customer churn. Table 2.1 reveals some statistics that translate the value of customer satisfaction. A positive experience can enhance the potential of future transactions and higher order frequency. On the other hand, a bad experience can result in a lost sale, in customer churn and in negative word-of-mouth.

Table 2.1: Value of Customer Satisfaction. Source: Leather (2013)

86% of customers would pay more for a better experience
73% of the customers will reward a superior experience by increasing purchases by 10% or more in such a case
A negative customer experience will make a customer leave in 89% of the times. (2011)
26% of customers would post a negative social media comment after a negative experience
58% of customers who have superior experience would refer the company to their friends and connections.

Reichheld (2006) explains how customer satisfaction (promoters) and dissatisfaction (detractors) bring value to a company. According to the author, there are five factors that have to be measured to relate satisfaction and value.

First, the retention rate, as "detractors generally defect at higher rates than promoters, which means that they have shorter and less profitable relationships with a company".

Secondly, the author points out the fact that satisfied customers lead to higher margins. Promoters are usually less price-sensitive than other customers, while detractors are more price-sensitive". The author recommends that a company examines the market basket of goods or services purchased by promoters and detractors over a six-to twelve-month period.

Thirdly, the annual spend is influenced by customer satisfaction as promoters "tend to consolidate more of their category purchases with their favorite supplier. A company's share of wallet increases as promoters upgrade to higher-priced products and respond to cross-selling efforts. Promoters' interest in new product offerings and brand extensions far exceeds that of detractors or passives".

The last two factors referred by the researcher are lower customer-service resources consumption as detractors tend to complain less and lower marketing costs as satisfied customer have longer relationships duration, have a tendency to engage in positive word-of-mouth and generate referrals. However, this factors are more difficult to quantify.

Taking on Reichheld (2006) perspective, in order to decide how much the company should be willing to spend for delivering a better service, it is important to know how much repeat business results from those actions. As described in the previous section order fulfillment impacts directly customer satisfaction. Rao et al. (2011) investigates the relationship of performance along the dimensions of order fulfillment satisfaction and its cost to purchase satisfaction and customer retention in online retailing. The results indicate “that satisfaction with the physical distribution quality and cost are positively related with customer’s purchase satisfaction and customer retention”. This project can therefore contribute to the maximization of the CLV.

Yet it is difficult to quantify the expected future profits generated by the customer due to a certain service improvement like delivery speed.

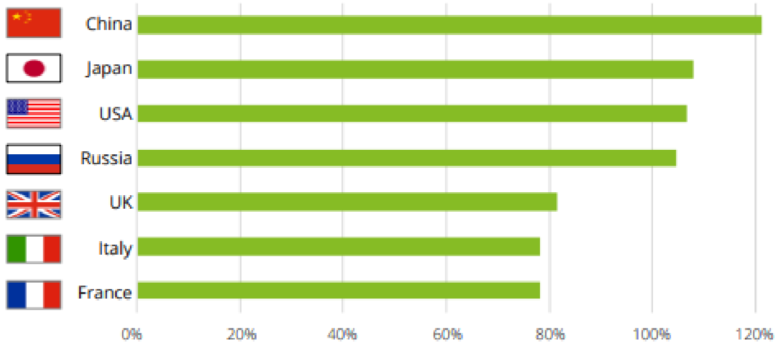
Concluding, benefits can be obtained from offering e-shopping service, such as increasing sales revenue and strengthening competitive advantage. By enhancing convenience or usability it is possible to expand the existing customer base. Gurley (2012) findings show that there is a relationship between convenience and repurchase intention. For this reason, the impact on convenience will be regarded in the design of the final solution for the presented problem.

2.1.3 Pricing and conversion in the online luxury industry

Pricing research was included here because, as opposed to conventional offline firms (business-to-consumer) the marketplace business model has a more complex and unstable pricing model. It is important to understand how customers react to price and how prices vary across globally spread vendors.

Price-sensitivity differs across industries, channels and type of customer. Luxury consumers are commonly associated with less responsiveness to price fluctuations. On the other hand, millennials "have grown up with technology and have learned to leverage it in more sophisticated ways" to truly understand the quality and value they receive for the money they spend. "As a greater portion of luxury consumers shopping online are young people and seeking value", it is important to convince them they had managed to get a good deal (Jacobson et al., 2011). Online shoppers can compare prices across competitors more easily than those shopping in a brick and-mortar location because the Web allows the customer just to move to a different Web site with the click of a mouse" (Razi et al., 2003). The ability to compare prices before buying is one of the top reason U.S. luxury consumers give for purchasing products online (Schmidt J., 2015).

The study from Deloitte’s luxury pricing analytic department summarized in Figure 2.1, shows that "despite increasing internationalization" prices are on average 50 per cent higher in China than in Italy or France. "A premium is charged in Asian markets for all brands" (Deloitte, 2017). This suggests that it is worth exploring the opportunity of differentiating prices by market in the luxury industry, which is specially interesting for marketplaces that have buyers and sellers globally present.



Index values represent the median US$-equivalent price relative to the global average. Created by BenchMarque

Figure 2.1: Price Index for full price like-for-like products in key markets. Source: Global Powers of Luxury Goods Report Deloitte (2017)

2.1.4 E-commerce order fulfillment operational challenges

Razi et al. (2003) defines e-commerce order fulfillment as "a series of coordinated steps [...] required to bridge the gap between a customer browsing an e-commerce site and actually purchasing and receiving an ordered product". The author points out five challenges of online retailing: demand nature, prompt delivery, reverse logistics, accuracy and supply chain.

First and foremost, demand in e-commerce is very unpredictable "since anyone sitting with a computer connected to the Internet is a potential customer, the customer base is huge". As marketplaces do not purchase inventory this does not represent a major challenge.

Following, modern customers expect rapid delivery, free shipping and free returns.This phenomenon, driven mainly by Amazon and its huge pushes to deliver products cheaper and faster, is placing increased pressure on managing order's fulfillment. For traditional retailers that integrate online channels, rapid delivery can be achieved via expedited freight or by placing inventory in the local markets. Since the additional high cost of expedited freight is not a viable long-term solution, the companies' challenge is to analyze which inventory should be placed tactically in each warehouse or establish business in the local market.

Specially for the apparel online industry managing returns efficiently and having a strong reverse logistic is crucial for success. According to Boyajian (2017) thirty percent of all products ordered online are returned, compared to 8.9 percent in brick-and-mortar stores. Customers need to touch and try on the products in order to decide if they really want to keep them. A study conducted by Nav reveals that specially luxury consumers bracket [1]. On top of that, it is proven that 78% of shoppers will buy more in the long-run if a retailer has free returns (Klarna, 2017). Although efforts have been made to reduce the

[1] Bracketing: buying multiple versions of an item, trying on at home, and returning those that do not fit.

bracketing rate (such as better fit and color information during the consideration phase), return costs will always be a challenge for e-commerce operators. To mitigate the costs of returns it is important to understand that there is a controllable and uncontrollable rate of returns. It is possible to reduce the probability of return by acting on the controllable reasons. Still other strategies could be deployed to reduce the costs that result from natural returns. Many traditional e-retailers offer in-store returns, which reduces shipping and handling return costs but also drive customers to visit the stores. Yet for a marketplace there are other opportunities to be explored. For instance, if the return reason is attributable to the seller, then sellers with better metrics should be favored, as it is detailed in the Chapters three and four. Furthermore, shipping costs are impacted directly by the route selection as further explained below. It is possible to verify that shipping costs and duties vary a lot.

Another important aspect that complicates the execution of online order fulfillment is guaranteeing that the right product is shipped. For retailers that need to process many orders this can represent a big effort. Shipping wrong items not only negatively impacts customer retention but also implicates extra operating costs.

Moreover, managing correctly order fulfillment is very demanding because all parts of the supply chain must be fully integrated and coordinated in order to deliver the items on time. Every order should be monitored closely to guarantee that the customer receives the desired product according to his expectations. For marketplaces this includes having an integrated order creation system, strong tracking systems and overseeing merchant selection, product offering and customer support (Nautiyal, A. et al., 2019). In order to support the above-mentioned high customer acquisition costs of the luxury e-tailing, improving order's fulfillment management and maximizing profit per order without jeopardizing customer experience is crucial. For monobrand businesses the complexity and expense of shipping logistics is considerable, due to several reasons, such as accurate inventory holding and warehouses management. Marketplaces have an opportunity to take the lead but need to manage operating costs as efficiently as possible.

Finally, in conventional e-commerce business if some of the items requested by a customer are not available, normally a backorder [2] is created, until the fulfiller's inventory is replenished. For marketplaces this situation is very common as the items are in the shelves of the physical store and therefore can be sold "offline". As marketplaces do not own the inventory and can not control whether the item will be replenished they try to assign the order to other possible seller or must cancel the order.

2.1.5 Order assignment strategies in different business models

Most academic research conducted on order allocation strategies in online retail deals with allocating inventory in warehouses or, in case of a make-to-order business, the choice

[2] Backorder: is an order for a good or service that cannot be filled at the current time due to a lack of available supply (Kenton, 2019).

of the production center to fulfill an order. Some articles around dynamic pricing models for time slot management exist in literature commonly applied to the grocery market, where delivery routes can be coordinated (Chen and Chen, 2014). Since all orders are processed individually and shipped from different vendors with third party logistic providers (e.g DHL or UPS 3-PL), the vehicle routing problem does not apply to the present problem of this thesis. As opposed to traditional retailers, marketplaces do not own inventory: it is no longer a classic straightforward exercise of optimizing the stock allocation in distribution centers and transportation routes to minimize total costs. Instead, it is about selecting the seller that should get the sale. As a matching marketplace, the challenge for the company is to guarantee a standardized quality service independently from the seller that is chosen to fulfill the order (Golden, 2017). The online luxury marketplace industry has different requirements. First, shipping costs are not the only driver, as in the case of low-value items, such as groceries, sold by large conventional retailers where profit margins are low, and thus transportation costs are a key determinant for business profitability. In the luxury sector, where retail margins are considerably higher, other factors have to be considered. Second, a pure marketplace without warehouses can not move supply from one place to another according to forecasted demand. It is possible to make purchase recommendations but supply distribution is totally dependent on the choice of the products that the seller decides to upload on the platform. However, in traditional business a similar challenge arises when a supplier has to be selected. The main difference is that this decision process has most of the times long-term implications and involves establishing a contractual relationship. For procurement purposes decision-makers are faced with many criteria in supplier selection. The selection of the best suppliers can greatly improve the supply chain performance. The criteria identified in literature to evaluate suppliers, namely price, quality, service, supplier's profile and global risk, are good inputs to evaluate sellers in a marketplace but miss to cover the marketplace necessities. For instance, a literature review from Weber et al. (1991) lists net price, delivery and quality as the most referenced criteria for selecting sellers. Interestingly, the packaging ability of the suppliers is considered a criterion with average importance. It is worth noting that geographic location is number twenty on the criteria ranking, which is obviously not transferable to the e-commerce order fulfillment problem. Although some of the criteria to asses vendors are identical, the importance given to each criterion is business specific and can therefore not be applied to solve the problem described in Chapter 1.

Since 1966 vendor selection criteria and methods have been discussed in literature, supporting the fact that decisions should be systematized and based on the analysis of facts rather than purely on intuition. In the online industry automated algorithms can support and enhance data-driven decision making by bringing in more data points from many sources, breaking down information asymmetries and making the process instantaneous. "As the sources of data grow richer and more diverse, there are many ways to use the resulting insights to make decisions faster, more accurate, more consistent, and

Table 2.2: Vendor selection criteria discussed in literature for procurement purposes. Source: Weber et al. (1991)

Criteria	Nr of articles	Dickson's Evaluation of Importance
Net price	61	Considerable
Delivery	44	Extreme
Quality	40	Extreme
Production facilities and capacity	23	Considerable
Geographic location	16	Average
Technical capability	15	Considerable
Management and organization	10	Considerable
Reputation and position in industry	8	Considerable
Financial position	7	Considerable
Performance history	7	Extreme
Repair service	7	Average
Attitude	6	Average
Packaging ability	3	Average
Operational controls	3	Considerable
Tracking aids	2	Average
Bidding procedural compliance	2	
Labor relations record	2	Average
Communication system	2	Considerable
Reciprocal arrangements	2	Slight
Impression	2	Average
Desire for business	1	Considerable
Amount of past business	1	Average
Warranties and claims	0	Extreme

From European Journal of Operational Research Volume 50, Issue 1, 7 January 1991, Page 12

more transparent" (Henke,N.,Bughin, J., Chui,M., Manyika, J., Saleh,T. , Wiseman, B. , Sethupathy, G., 2016). Many platform algorithms highlight products with a better value or experience, which results in better placement and increased competitiveness. Little scientific research is found on sellers evaluation in a marketplace, but a significant number of practical reports on that topic exist from companies that try to boost sales on Amazon and other marketplaces. According to Amazon's documentation, as well as speculation from 3P sellers Grasso (2018), Amazon weights multiple customer-centric criteria to define the "Buy Box" winner when there are multiple sellers offering the same product (see Table 2.3). According to Grasso (2018) "as a customer-centric company, Amazon's goal has always been to offer the best possible experience to its customers. The Buy Box was created with the objective of comparing multiple offerings of the same product to determine which will give the customer the highest level of satisfaction". This includes sales volume, response time to customer inquiries, rate of returns and refunds, and shipping times. By means of an empirical analysis of algorithmic pricing, Chen et al. (2016) were able to achieve high prediction accuracy on the buy box solution, suggesting the following seller features to be considered: landed price, customer rating, positive feedback, feedback count, fulfillment by amazon (FBA) or seller is Amazon.

Table 2.3: Vendor selection criteria in Amazon marketplace. Source: Grasso (2018)

Importance	Criteria
Very High	Fulfillment Method Landed Price
High	Availability Order Defect Rate Valid Tracking Rate
Medium	Late Shipment Rate On time delivery Feedback score Customer Response Time Feedback Count
Low	Inventory Depth and Sales Volume Cancellation and Refund Rate

Amazon considers its own fulfillment service to have perfect scores for multiple variables, including shipping-time, on-time delivery rate, and inventory depth and thus gives prioritization to sellers that use FBA. In this case this is also a business strategic measure because Amazon charges for that service. The simplest metric looked at by the Buy Box is the time in which the seller promises to ship the item to the customer. This mechanism should work as an incentive for sellers to cope with the business and deliver a standard quality service. It is in the marketplace best interest to have fragmented supply, making sellers' growth evenly distributed. Amazon example indicates many practical criteria that can be adopted in the seller selection process discussed in this thesis. Similarly to Amazon, Farfetch has Fulfillment by Farfetch, a logistics service for the marketplace sellers supported by three warehouses globally present. The design of the assignment algorithm has to be comprehensive to be able to capture this kind of effects and prioritize the fulfillment from one of this warehouses in case it is more profitable for FF and/or it guarantees higher customer satisfaction.

2.2 Methodologies

2.2.1 Multicriteria Decision Making

Multicriteria Decision Making, hereafter referred as MCDM is a term used to describe a subfield in Operations Research and Management Science. MCDM is generally defined as a tool to solve decision problems that involve multiple (sometimes conflicting) objectives. Marler and Arora (2004) compile several examples of methods for multi-objective problem solving. The theory behind the existing sequential employed logic at the company is the lexicographic method, where the objective functions are arranged in order of importance. What distinguishes the theoretical to the practical approach is that some of the filters are qualitative and therefore can not be defined as classic multi-objective optimization process.

To address the problem of combining different objective functions in the supplier selection context different approaches were identified in the literature review. Hu et al. (2017) uses an integration of mixed integer linear programming with weighted sum method, which requires defining weights and normalize values. This method is associated with the planning phase. However, marketplaces decisions are made on the fly (discrete alternatives), where every order is an isolated event and thus there is a chance to evaluate the alternatives and select the optimal without considering global constraints. As it is not possible to anticipate subsequent orders, it is difficult to impose hard constraints like a minimum order quantity for vendors at the time of the order. Furthermore, currently the commission that FF charges is calculated individually independently from the total quantity sold which gives no immediate economic incentive to balance the order allocation. Therefore the usage of for instance lexicographic goal programming (Çebi and Bayraktar, 2013) is not obvious under this circumstances.

A lot of research exists around ranking methodologies considering multiple criteria for the supplier selection problem. Bhattacharya et al. (2010) propose an "integrating analytic hierarchy process (AHP) with quality function deployment (QFD) in combination with cost factor measure [...] to rank and subsequently select candidate-suppliers under multiple, conflicting-in-nature criteria environment". The output of AHP is a prioritized ranking of the decision alternatives based on the overall preferences expressed by the decision maker. "QFD is a well-structured, cross-functional planning technique that is used to hear the customers' voice throughout the product planning, development, engineering and manufacturing stages of any product" (in Bhattacharya et al., 2010). Vahdani et al. (2008) use a multi-criteria-decision model approach considering vendor attributes such as profitability, relationship closeness, technological capability, conformance quality and conflict resolution. A ranking matrix is built, indicating the frequency of the relative superiority of suppliers by making pair wise comparison for each defined criteria.

AHP could only be used to determine the weights in case of a single objective formulation, but not to reach the preferred solution for the order as this approach requires too much computational capacity to do a pairwise comparison between all suppliers for every product. Moreover, it uses verbal judgment which rules would have to be programmed in order to be scaled to the website. Another downside of these methods is that they are not easy to communicate and require the decision maker to provide judgments about the relative importance of each criterion and to specify a preference for each decision alternative using each criterion, which is not the aim of the project.

2.2.2 Simulation for decision-making problems

A commonly used alternative to solve decision-making problems is simulation. Simulation tools are being used to facilitate and minimize the risk of making wrong decisions in the various business process phases. Using simulation tools for business challenges enables organizations to save time and money. Simulation is seen as a powerful tool for testing operations in different industries and have been widely used in supply chain network design to examine and verify scenarios as for example design the best layout for logistics networks (Anylogic, 2019).

According to Ramadan et al. (2015) there are several advantages in using simulation. Built models can be used repeatedly for different analysis. Computer simulation can enhance the qualitative analysis by stress-testing the business model under many component variations and scenarios of external developments. Simulation can estimate performance measures and evaluate the effects of any changes that may occur to the system operational parameters. Simulation can just describe different scenarios and “what-if” question results. Another positive aspect of simulation is the fact that sophisticated models can be easily integrated to introduce, for example, variability of processes and reproduce system dynamic interactions.

There are also many challenges that can be faced when using simulation. “Building simulation models is costly and requires lots of runs for one model to achieve reliable results. Accurate results can not be given if the inputs are inaccurate“ (Ramadan et al., 2015). Problems can not be solved by simulation: it can only provide information that helps getting or deriving solutions.

2.2.3 Real-time assignments vs re-evaluation

As mentioned before, most research focuses on conventional e-commerce. Considering online versus offline order assignment, most examine the benefits from periodically re-evaluating real-time assignments in order to achieve an overall optimized solution. "When a customer order occurs, the retailer assigns immediately the order to the warehouse that minimizes procurement and transportation costs. However, this assignment is necessarily myopic as it does not account for any subsequent customer orders or future inventory replenishment"(Xu, 2005). The authors developed "a near-optimal heuristic for the re-assignment of a large set of customer orders, which minimizes the total number of shipments without compromising customer experience by keeping the promise-to-sip date commitment for any customer order". Another great example that illustrates the complexity in the online retail environment is the paper “Better Fulfillment and Replenishment Decisions” from Acimovic (2013). Korolko et al. (2018) for example discuss this topic for Ride-Hailing Platforms, where re-evaluation occurs by changing the assigned driver in case another customer starts a ride request and it is globally better to swap drivers. The re-evaluation of the decisions in time buckets is a possible alternative for the company to explore. After the system assigns the order to a seller, there is a step in the order cycle to check stock availability. The order might wait several hours before the stock is confirmed by the seller and info given to a third-party carrier to deliver the package, which gives margin to improve. The real-time decision is myopic because future customer orders or inventory replenishment are not being considered. By re-evaluating the real-time assignment decisions, without compromising estimated-delivery-date and price for any customer order it would be possible to make better decisions globally. However, this will not be further explored as the focus of the project is the order assignment decision process.

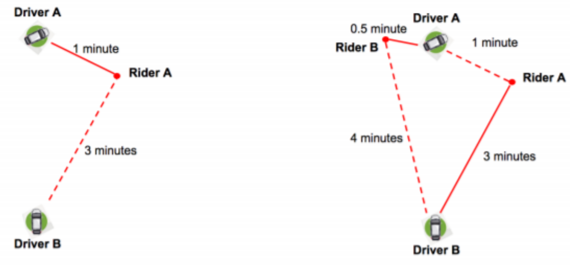


Figure 2.2: An illustration of efficiency loss in myopic matching algorithms. Source: Korolko et al. (2018)page 8

2.3 Final considerations

Having explored the business context and alternative methods to solve similar problems we are now in position to transfer some knowledge and adapt the studied techniques to the presented problem. Several business concepts and identified trends will be taken into consideration. As for the present problem, theoretical approaches of MCDM do not present solutions, what-if analysis will be conducted to evaluate which changes to the current process can deliver more value. In the next chapters the focus is to analyze the strategy for order allocation and develop the capabilities to estimate the impact of deploying that strategy. Re-evaluation can only occur after the "myopic" decision process is optimized and therefore is a next step of the project.

Chapter 3

The Order Assignment Problem

To improve the order assignment process, first the AS-IS situation and how the seller is being currently selected must be analyzed. Secondly, a TO-BE vision of the order assignment decision making process and all the criteria to consider in the selection of the seller have to be developed. The third step consists in developing a tool that mimics the exact decision process that occurs currently at the company for past orders. This is extremely important to assure that the simulation tool is valid, and that all fulfillment alternatives for a certain order are being considered when recreating a parallel reality. With that knowledge it becomes possible to analyze the size of the opportunity of improvement and evaluate with precision the inefficiencies of the decision-process. Since no record of the decisions made exists, the first effort is to create a database to feed a simulation tool able to reproduce past order allocations with real stock levels, prices and other relevant data. Having that, the simulation tool must be programmed considering all business rules. Simulation is the only way to mirror the behavior of the real system as it runs forward in real time because former decisions impact future ones. For example, if a store has only one available SKU, if it is sold to a first customer, naturally a second customer could not receive the item from that same store. It is a system where past decisions affect future ones. The output of this step is a list of possible solutions and the decision for every order at the time it was placed. Finally, changes can be made to the program using the same structure and impacts of alternative processes can be measured.

3.1 AS-IS Diagnosis

The foundation of the order experience is the order cycle illustrated in Figure 3.1, which determines the workflow of every order execution. The order cycle is divided in six major steps. After an order is placed, the selected seller has to confirm the stock availability while systems verify fraudulent buyers. Step 2 and 3 describe the order fulfillment tasks of the sellers to prepare products to be shipped. Finally, package is ready to be collected by the carrier at the store. The final step (6) is the transportation of the package to the customer location. This steps are used to determine the performance of sellers and carriers and will be mentioned henceforward. The order assignment is an automated procedure that precedes this cycle. Assigning the order to a specific vendor not only determines the price displayed to the customer visiting the portal, but also impacts the efficiency of the following steps, which are controlled by the sellers or executed by third-party logistic providers.

Figure 3.1: Order Work-flow Overview

Every time a customer visits the website, there is an algorithm that selects for every product (and size) that is online the best boutique to fulfill the order. This is done based on the delivery country selected by the customer on the web-page and product availability. Depending on the selected boutique, which does not necessarily have all the sizes, the algorithm returns a price in the customer currency to charge and display on the product listing page (PLP). To provide a price consistent shopping experience after this moment the price will not change, and the item will be shipped from the selected seller. Only in case the selected boutique does not have the size desired by the customer, by changing his preference on the product display page (PDP) the algorithm will return a new solution that might be of a higher price. In its core, the algorithm is designed around customer experience. All the eligible merchants are listed and submitted to a successive filtering technique. Eligible merchants are stores that have the item available for online sale and contain distribution to the customer country. Figure 3.2 shows the select criteria in order of importance. Each filter reduces the number of possible solutions until there is just one seller left. Filter price selects only sellers with a landed price within a certain interval, then boutiques are filtered based on geography, then score and finally if there are still ties the seller who uploaded the item first wins the sale.

Since the seller is selected upfront, the origin of the items in the customer shopping bag remains unchanged not considering any subsequent products that the customer might order. A re-evaluation of the selected sellers does not take place in the checkout step to evaluate consolidation potential.

The first step to redesign this selection procedure was to look at the current logic and understand the criteria and the order of priority defined by the decision maker to evaluate the suppliers. Figure 3.2 reveals the ambition to favor suppliers capable of delivering

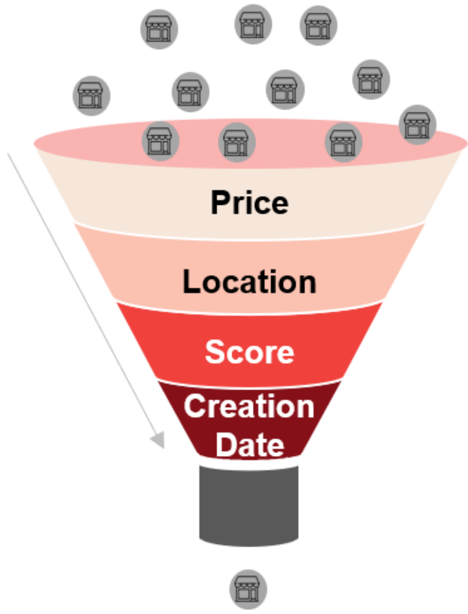


Figure 3.2: Successive criteria filters used to select stock point

more value to customers. The idea is to provide the best available price option that is closer to the customer location offering the best service. However, this algorithm can be further improved as there are some business particularities that are not being considered in the decision-making process. In this chapter the above-mentioned sequential filtering technique and used criteria will be discussed.

3.1.1 Price Criterion

From all listed alternatives that have stock available to fulfill an order to the country of the customers, only the boutiques selling the SKU with a price between the minimum price among the sellers and an upper bound of 5% can be chosen. Sellers selling above this upper limit are automatically excluded from being selected, even if they are the closest to the customer.

The current business model deals with two different pricing schemes, the standard pricing and the geographical pricing, introduced afterwards.

Standard Pricing is when the price is defined as the producer price plus seller's margin as illustrated in Figure 3.3. In return for managing the transaction Farfetch captures part of the seller's margin. For this reason, in case of standard pricing, the price is variable depending on seller's margin rate in the physical channel and VAT of the store country.

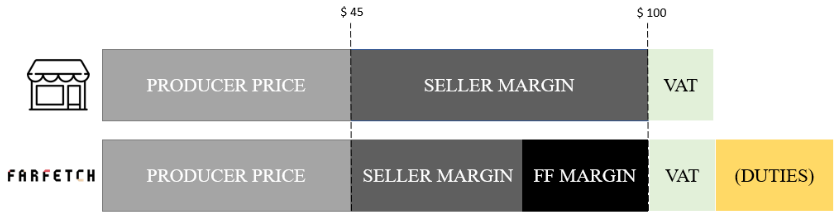


Figure 3.3: Illustrative Industry Economics

The fact that boutiques are being filtered first based on price raises the need to make a sensitivity analysis of the price threshold to understand if the value can be captured just by enlarging the filter. This aspect leads to the discussion of price sensitivity and how price impacts conversion rate. Price shown to the customer will not be explored in this initial phase of the project as the key goal is to improve order fulfillment customer experience and business profitability without influencing the conversion rate. For this reason, the results in Chapter 5 assume that the price charged to the customer will always be the minimum price possible. Looking at the example in Table 3.1 even if the price charged is the lowest one, it might be preferable to order from a more expensive boutique as shipping or other relevant costs might counterweight the price difference. With the current logic there is a risk of shipping from another country just due to small price discrepancies, which is environmentally unfriendly. On top of the price without VAT a fee can be applied to non-domestic routes (shipping subsidy), which mitigates the risk of shipping from another country just due to small price discrepancies. Additionally an exchange rate fee can be charged to accommodate currency fluctuations.

Table 3.1: Example of seller selection problem for an Italian customer

Store	Store Country	Customer country	vendor online price	Shipping Cost
A	Italy	Italy	106 EUR	8.58 EUR
B	UK	Italy	100 EUR	13.28 EUR

Moreover, Figure 2.1 in the Literature Review Chapter shows that prices are very different depending on the market. This makes sense as the standard cost of living varies and boutiques have therefore different retail margins. For products following the standard pricing scheme, FF as a marketplace can profit from shipping from countries that have a lower price level to markets that are used to higher prices offline. Currently the option is to offer the lowest price to every customer, as the company is investing in gaining market share. However, in the future there is an opportunity to capture extra margins by meeting the market price and shipping from less expensive countries. As an example, a german customer can afford to pay more for a product than a portuguese does. Therefore, the portuguese boutique charges naturally a lower price than the german boutique. For this reason it might compensate to ship the product from Portugal and charge the german price. Clearly, from an environmental point of view this should be penalized.

The other pricing scheme occurs if some brands establish a fixed value required to sell in a specific market. This concept is called within the industry **geographical pricing**. Farfetch respects that strategy followed by some brands. In those cases, a positive or negative margin, named geomargin, is captured by Farfetch partially or totally depending on the terms agreed in the contract.

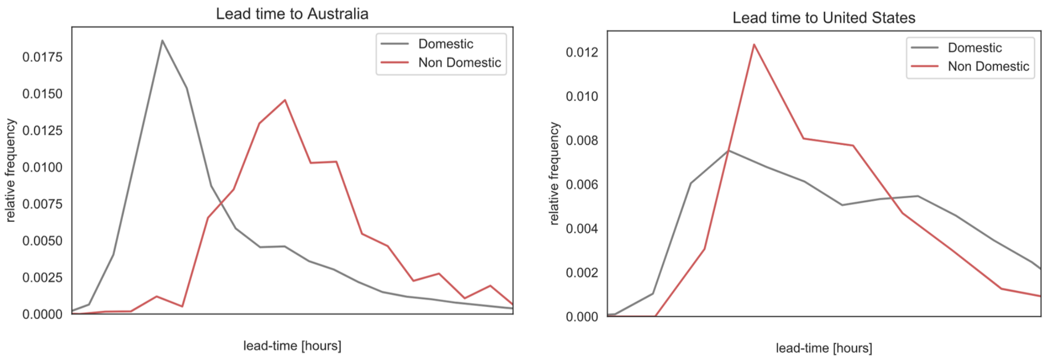
Interesting here is the fact that there is a great opportunity to drive profitability by maximizing the margin obtained by the company. The base value for calculation of the effective commission captured in these cases is the net retail price. To estimate the size of the opportunity of improvement a comparative study on the net retail prices of items ordered, where Geo-price policy is applied, was made.

If the GeoPrice policy is forced, the price filter of the existing algorithm is not active. In the last months, on average 35% of the marketplace ordered items follow the Geo-Price policy.

3.1.2 Location Criterion

In case there is more than one boutique that charges a price within the above-mentioned interval (from minimum price to minimum price + 5%), the algorithm looks for boutiques in the country of the customer. If there is no boutique in the delivery country, boutiques in the continent of the customer are filtered. In case there is no boutique matching this requirement, all boutiques that have prices within this range are considered. This criterion was selected to minimize lead time, possible duties and shipping costs. This technique is not precise enough to make optimal decisions in terms of lead-time as speed is highly affected by the response time of the boutique, the so-called speed of sending that is explained in the next section. Furthermore, in some countries covering big areas like the United States, sometimes domestic routes take more time than shipping from another country as ground transport is slower than air. Concluding, the location criterion is not favoring necessarily sellers that offer the best lead-time to the customer.

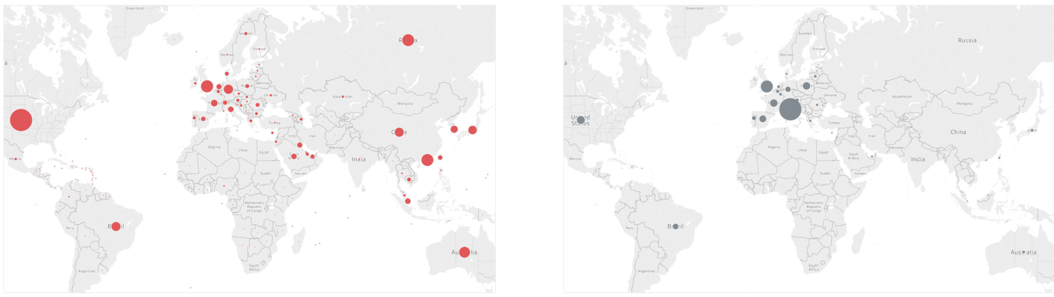
If we analyze the number of boutiques per customer country, we can conclude that 60% of the stores are located in Italy, United States, Brazil and UK and the supply is highly concentrated in European countries (IT,UK). Italy leads the list, being responsible for fulfilling around 45% of orders globally (2018-2019) . This leads to the conclusion that the country filter is not being effective as we have clients in more than 190 countries. Having that, in most of the cases, distance criteria is not being considered. Summing up, the geography filter is not always benefiting the customer and favoring the optimal solution. The variability of lead-time in domestic shipments and high density of European supply emphasizes the importance and interest of using a better strategy to select the fastest delivery option. However, the possibility of increased duties and shipping costs that can



(a) (b)

Figure 3.4: Impact of domestic routes in lead-time

result from not leveraging domestic shipments - as the current process does - should be controlled.



(a) Demand (b) Supply

Figure 3.5: Supply and Demand Distribution

3.1.3 Score Criterion

A score is calculated for each seller every two weeks based on the historical orders (last 15 days). Seller score is composed of multiple performance metrics: the average days of speed of sending excluding weekends (SoS), No Stock Rate (NS) and Net Promoter Score (NPS). NPS stands for a customer order satisfaction rating in a scale of 1 to 10. Speed of sending is the time between order placement and moment where the item is ready for the carrier to pick it up (end of step 4 in Figure 3.1). No stock is given by the following formula:

$$\text{No Stock Rate} = \frac{\sum(\text{items canceled due to stock out})}{\sum \text{items sent}} \tag{3.1}$$

The weights of each criteria are worth challenging. Currently, points to a maximum of 1 in case of NPS and 13 in case of SoS and NS are attributed based on previously defined intervals.

Table 3.2: Boutique Score Points Scheme

	SoS	NSR	NPS
Points Interval	[0 - 13]	[0 - 13]	[0 - 1]

The introduction of these variables in the algorithm makes sense because the business wants to reward sellers that contribute for a high-standard service level. In the year of 2014 around 3.3% of the orders were unfulfilled as a result of no stock. Later, a scheme of incentives and penalties was instituted and the percentage dropped to approximately 1.4% (U.S. Securities and Exchange Commission 2018). Yet there are other variables that are also considered relevant to measure the quality of the service as there is a scheme of incentives that also rewards sellers for not sending wrong items (measured as wrong item rate) and sellers that follow the packaging recommendation. In the future, if all these variables are integrated in the order assignment algorithm and if it is properly communicated it could be possible to reduce the incentive plan costs, as a natural incentive from being prioritized in the order allocation algorithm exists. Sellers should compete to have a better ranking position by improving those metrics as it is seen in other marketplaces (e.g. Amazon mentioned in Chapter 2, page 15) Furthermore, the way these variables are calculated have some shortcomings that are worth mentioning, as explained later in Chapter 4.

The NPS is a score, where clients evaluate multiple parts of the experience, some of them not influenced by order fulfillment and delivery (e.g. customer support and website navigation). Furthermore, the rating response rate is very low. Moreover, the NPS is attributed to a portal order that can be associated to more than one boutique. An analysis of the NPS data reveals that it is highly impacted by wrong-item. The speed of sending is sometimes longer, due to fraud checks that need to be internally processed by FF teams. This should be reviewed as it is unfair for sellers to be harmed for that reason. Also, if the boutique rejects the order, the time until that answer is obtained is not included in the calculation of the SOS. There are more sophisticated ways to predict no-stock and lead-time instead of using an average from the past two weeks, such as time series forecasting, including seasonality. The question that arises from the current process is if whether the current weight given to each variable makes sense and what other variables should be included in the model. Naturally, the impact of these variables in both customer retention and operational efficiency would have to be predicted and assessed.

3.1.4 Creation Date Criterion

The last criteria to decide which seller gets the order is the date in which the product was created, that is the date when the seller decides to upload the item online and send it to the photography studios of Farfetch. This filter works like the First In First Out

(FIFO) dispatching rule. The idea here is to reward sellers that provide "inventory" to Farfetch earlier, which increases the opportunity for Farfetch to capture more sales and gain competitive advantage, and also introduces a sense of fairness.

3.2 TO-BE Business Vision

As remarked in the previous chapter, the sequentially used filter criteria can lead to sub-optimal solutions. To develop a vision for the order allocation it is important to find out which business areas and stakeholders are impacted by the algorithm and the respective reasons. Order assignment impacts profitability, customer relationship and sellers' sales. Furthermore, the order assignment process translates and puts company strategy into action. Today, the way the customer journey and order cycle are designed, impose some limitations to what the ideal scenario of the order allocation process could be.

In case it would be possible to redesign the order cycle work-flow or customer interactions with the platform two alternatives appear to fit the business requirements. As time is one of the most important aspects to consider for improving e-commerce experience and gaining competitive advantage, an auction model - first boutique to confirm availability - could be suited for a marketplace business. To keep control of the process the marketplace company would still have to determine the terms of the offer based on the inventory availability. This has the potential to accelerate many orders that are delayed due to time-zone differences or inefficient response system of some sellers. Despite all, in the short time this is not possible to adopt as it would require extra work streams to develop the technology, which are not planed in this year road-map. Additionally, this strategy would catalyze disruptive changes in the sellers order management process, which would require extra workforce to minimize the introduced instability and support a successful implementation.

On the other hand, a Time-Slot Pricing scheme could be considered. Ideally when a customer places an order on the website, an estimated delivery date in real-time is assigned to the order depending on the boutique and customer country. This service could be used to calculate expected time of all eligible sellers and serve as an output to present delivery speed options to the customer. In this scenario, instead of minimizing lead time, the choice for faster or slower delivery is shifted to the customer that can chose between more than one time-window option. Depending on the customer choice a different price is charged accordingly. With this service a customer expectation is created, and the business is committed to deliver in that period. Deviations on that expectation will deeply harm the business image and thus estimated delivery date variability must be penalized in the future solution for the order assignment process. This would define clear targets for vendors and carriers to meet and delays could be monitored and accordingly penalized. This solution would lead to increased customer satisfaction as it pleases both customers that value speed and are willing to pay for it as well as customers that are more price-sensitive.

Given the current infrastructure, the goal is to move from a sequential pipeline with prioritization to a one ranking formula that maximizes value. Every order allocation should take into account customer satisfaction, profitability, and seller performance.

This thesis presents an exhaustive study of the variables worth including in a ranking formula and the impact of each one in the business performance. The first step was to understand the business requirements and identify possible initiatives. On an order by order basis, for a certain item given a certain price, the new algorithm will optimize the order allocation by finding the right balance between three key areas, namely, the customer experience, the seller success and marketplace profitability (see Figure 3.6).

Figure 3.6: Perspectives to consider in the order assignment process

3.2.1 Customer Experience

The literature review performed in Chapter 2 on this topic suggests that offering convenient e-service can impact business profitability in the future, as customer satisfaction is related to higher purchase intentions and brand loyalty. Transferring the factors raised in the Literature Review to the order fulfillment of the studied company, customer satisfaction can be mainly impacted by: price, lead time, wrong item/damaged, cancellation Farfetch fault, packaging and time to refund. Each of these factors will be discussed afterwards.

Having the luxury customer in mind, the ambition is to deliver the best possible alternative to the customer, as the company is convicted that satisfied customers are reliable future profit sources. The business is looking to balance between offering the best customer service and minimizing operating costs. However, there is no common understanding among decision-makers in order to give weights to the different variables. Ideally, all the variables would be transformed in USD dollars in order to have a single objective problem.

3.2.1.1 Price

According to FF commitment to its customer the prices displayed on the product list page should be the minimum price possible. Currently this is guaranteed by the price criterion. This restriction has to be considered in the design of the new solution. Another aspect to consider is the price consistency, as the literature review points out that price inconsistency throughout the site navigation is negatively perceived.

3.2.1.2 Delivery-time

In order to include the delivery speed in the seller selection process, this variable has to be modeled. The lead time is the time since the customer placed the order online until it gets to the customer door. In order to properly compare sellers lead-time in the order assignment process it is important to use the date of the carriers first attempt to deliver instead of received date.

$$\text{Lead time} = \text{Speed of Sending} + \text{Time in Transit} \tag{3.2}$$

Speed of sending (SoS) is the time elapsed since the order is placed and the item is shipped. Time in transit (TiT) is the time spent in step 6 of the order cycle (see Figure 3.1).

Aligning with the company understanding about the importance of the delivery speed to stay competitive, the goal is to test how much profitability will be lost if lead-time is prioritized over other criteria. Figure 3.7 shows the results of the analysis performed to understand how retention rate (3 months) is impacted by delivery speed. From the analysis results there is evidence that a faster order increases retention, which is vital for marketplace survival.

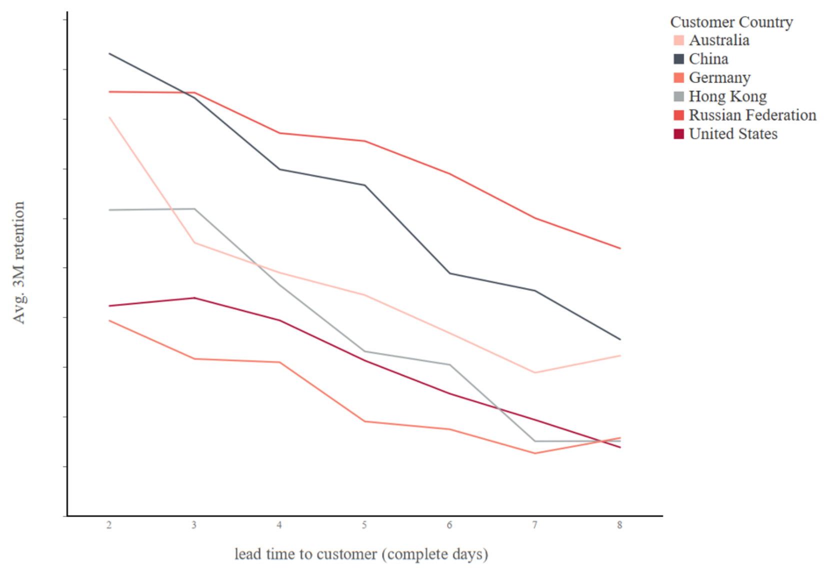


Figure 3.7: Retention 3 Months in Key Markets

To value the delivery speed in dollars three possible alternatives should be considered. First, define a maximum delivery speed to be experienced by a specific customer, which is aligned with the company's strategy to deliver real-time personalized customer management. For this approach the usage of the CLV model presented in Chapter 2 is appropriate. Second, there is an option to create "business-modes", where lead time is used as a second criteria before or after profitability (lexicographic optimization). Depending on market circumstances the business can activate a "profitability mood" or a "delivery speed mood" for acquiring new customers or delighting existing ones. Third, model the order frequency and next orders value to get an expected long-term profit for a certain delivery speed experience by the customer.

3.2.1.3 Other factors of customer experience directly impacted by sellers

Apart from fulfillment errors such as No-stock and sending wrong products (explored in the next section) for which customers are compensated with vouchers, there are two other factors related to sellers performance that affect the customer experience. First is the packaging quality. The literature review, revealed that the online consumer values the unboxing experience, specially in the luxury industry. Farfetch is focused in providing a better unboxing experience to its customer although the packaging (step 3 in 3.1) is controlled by the sellers. Therefore, a variable that measures the customer satisfaction towards packaging quality should be included in the order assignment process as it highly impacts the brand image. By communicating that sellers with higher rates related to the unboxing experience are more prone to sell on FF portal it is expected that sellers get motivated to improve their service.

Another important factor that is directly related to sellers performance is the time to refund. When a customer initiates a refund, the boutiques has to accept the product in order for the process to proceed. The waiting-time is negatively percevied by the customers and therefore should be penalized.

3.2.2 Seller Success

Order assignment algorithm impacts each of the sellers' sales and probably their future growth. One of the main obstacle of this project is the uncertainty about the impact of changes to this algorithm in the sellers performance. In order to monitor sellers satisfaction it is important to evaluate how much of the invested inventory is sold during the season. Expression 3.3 presents the KPI used to monitor vendors satisfaction, which is related to their inventory turnover. This formula is more complex than the traditional formula, as sellers have the opportunity to sell the products through the physical channel.

$$\text{Sell-Through} = \frac{\text{FF Portal Sales Units} + \text{Online Units Sold at Physical Stores}}{\text{Uploaded Inventory since begin of the Season}} \tag{3.3}$$

This metric is specially important at the end of the season. However, sellers are sensitive to it during the season as it is better to sell before discounts period. There is no contractual commitment with sellers, where Farfetch determines a minimum sell-through. Nevertheless, on the commercial side for seller relationship management and marketplace sustained growth it is crucial that sellers are engaged and motivated. The marketplace lives from the sellers inputs, since it does not own any inventory. This relationship has to be managed, in order to achieve better results for all the stakeholders of the business. Some exploratory work was developed to translate the value from sellers engagement. Two options to include this perspective in the selection process were discussed. The first is to categorize sellers by growth potential and design specific healthy sell-through curves over the course of the season (weekly or monthly points) and minimize their distance from optimal curve. With this strategy, one may control sellers growth. Segmentation is commonly used to manage relationships with clients. Differentiating seller levels involves creating levels of value for more engaged sellers that show commitment. Some important aspects to assess sellers' commitment to the business are size, stock quality, risk aversion, speed to market and offline-sales rate. This option limits the solution space as restrictions of minimum sell-through are imposed. Another downside of this solution is that many vendors joined the platform recently and there is no sufficient data to automatically assign a curve to each partner without human intervention. Second option, and more disruptive, is to monitor with simulation capabilities the effect on sellers and use the information obtained to help and advise sellers on possibilities to improve their positioning. This way we instigate competition and sellers become therefore more efficient and bring more value to the marketplace. This model based on fairness and meritocratic guidelines needs to be communicated correctly. When confronted with lower unexpected levels of sales, sellers will complain and demand solutions. The order assignment tool has great impact, because testing the strategy previously allows to anticipate the consequences and plan a proper reaction.

Since the second option has higher improvement potential in this first iteration, sellers' perspective will not be integrated in the model. However, a sellers' impact dashboard was developed in the scope of this project to mitigate the risk of surprising outcomes and to give insights for growth slowdown prevention.

Ideally sales should be balanced across partners, as there is a risk that some suppliers can be overloaded with orders meaning that their capacity will be exceeded. Although there is no obvious economic incentive to balance demand across partners, as the commission that FF charges is calculated individually independently from the total quantity sold (apart from rare exceptions), balancing demand across suppliers to minimize the risk of exceeding sellers' capacity and to lose some of their engagement is important. If there are ties after considering profitability and customer experience it makes sense to assign the order to the seller that sold the least during the season.

3.2.3 Marketplace Profitability

To address the profitability perspective, it is important to contemplate revenue and cost drivers caused by operational activities, which are explicit in Table 3.3. Revenue comes from the seller side and from customer side. On the seller side profit comes in form of a blended commission on the net sales value. Besides charging a commission on each portal transaction, Farfetch has other revenue incomes to cover operating costs. Offering free Returns, for instance, is supposed to boost sales. Since sellers are profiting from it, for each transaction a free return contribution fee is charged to the seller to support free returns service. Although the final customer pays shipping fees, it is not enough to cover shipping expenses.

Table 3.3: Revenue and Cost Drivers controlled by Ops in 2018

Profit and Loss Line	Impact	Driver
Commission	1	Revenue
Shipping	2	Cost
Duties Refund	3	Cost
Shipping Free Returns	4	Cost
Item Swap	5	Cost

The drivers impacting business revenue are proportional to total merchandised value. In order to increase profitability it is important to maximize conversion. Moreover, it is important to highlight that there is a different kind of commission structure for different kind of vendors. It is important to acknowledge that, as commission is the biggest profitability driver.

Looking at the Profit and Loss Statement it is possible to identify improvement opportunities related to the order assignment policy. The choice of the route influences directly shipping costs and potential duties payments in case of returns. In addition, the performance of the chosen seller can influence returns, due to wrong or faulty items. The price choice has direct impact on conversion rate. Moreover sending a wrong item or cancelling the order for no stock reasons can hurt conversion and therefore collected fees and profitability. Finally, as Farfetch follows a sellers commission scheme, assigning sales to vendors whose commission rate is higher, can improve financial performance.

3.2.3.1 Minimizing shipping costs

Some e-commerce business like Amazon can leverage economies of scale in shipping and profit from distributing their inventory optimally. Unfortunately, this is not the case for Farfetch marketplace, as shipment is charged individually. What influences shipping cost value is mainly route and box weight. Therefore order assignment plays an important role in reducing shipping costs.

The algorithm that is currently being used favors domestic and inter region shipments by selecting sellers in the same country or same region as delivery country. However, it is important to quantify the costs for each order, so that better decisions can be taken considering other profitability factors. According to FF internal operational reports international shipments are 2.7 times higher than domestic shipments. Compared to intra European shipments these are 1.5 times higher than domestic transports. Changing the order assignment criteria might increase international orders and therefore it is important to control shipping costs.

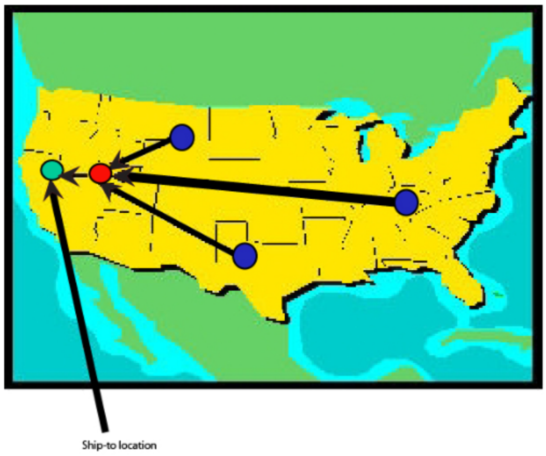
Besides selecting cheaper routes, consolidating orders, favouring partners that use smaller boxes are two other alternatives to reduce shipping costs.

Orders Consolidation

In 2018 only 13.9% of the multi item orders from Farfetch were consolidated. According to Aaron Cheris and Davis-Peccou (2017) doubling the average number of items purchased per e-commerce transaction and eliminating split shipments would reduce average per-item emissions by 30% and cut shipping costs more than 50%. Besides logistic cost reduction, environmental and management costs are to be considered. If the boutique that has only one of the items of the order delivers a higher commission, is it worth to consolidate the items in another lower margin boutique, and give up the higher margin? If both items are available in the best boutique, then the order is consolidated. If on the other hand, both items exist in a boutique that is less cost efficient than the boutique that has just one item then the order should be split. The new order assignment algorithm should analyze if it is worth consolidating and send from the less profitable boutique in terms of commission and geomargin. This means that the spared shipping costs from consolidating have to outweigh the profitability loss.

Consolidating has another advantage that is delivering to the customer all the articles at the same time. One could argue that consolidation provides a better customer experience as the client receives only one package at once. However, this is difficult to quantify as some customer might value receiving at least one product fast. Nevertheless, this approach is definitely more sustainable which is starting to strongly influence consumers' purchasing decisions. Specially the new generations that are taking over the market with their new values, ethics and distancing themselves from concepts like excessive consumerism, extremely high costs and guilty pleasures are becoming very sensitive to this aspects. "Sustainability is key to win new customer generations' hearts" (Deloitte, 2019). It is therefore advised to reconsider in the check-out step if shipping from a less profitable or less efficient boutique is not better than shipping twice. Ideally, this initiative should be implemented as a ranking algorithm of the product list page: after a customer has a product in the shopping cart then suggest at the top of the listing page products that are available in that boutique.

Another alternative to implement consolidation, even if products are not available in the same merchant, is the "merge in transit" concept (see Figure 3.8). It is worth calculating if it is advantageous to ship and consolidate all the products in one single location (seller), and only then ship to the final customer.



From IBM Sterling Selling and Fulfillment Suite V9.2.0 documentation

Figure 3.8: Merge in transit concept. Source: IBM (2013)

Packaging accuracy

Sellers are responsible for the packaging (step 3 in Figure 3.1). However, they do not pay shipping and thus have no incentive to use box space optimally. Favouring partners that use boxes according to FF recommendation, i.e. smaller boxes, is one of the alternatives to reduce shipping costs.

3.2.3.2 Minimizing return costs

In e-commerce returns are always a big share of costs. There is a percentage of return rate that is considered natural and very difficult to influence. Nevertheless, quality issues and wrong item are controllable by selecting a seller that has better metrics on these two aspects. As there is a natural return rate that is not possible to influence by the seller selection (e.g. customer changed his/her mind), the route that has on average lower return costs should be favoured. Returns are processed differently depending if they are intra EU or cross-border. In the cross-border returns Farfetch does not pay return clearance duties as the company does it via UK. It is foreseen by law that there is a duty relief from returned customs (UK Government, 2018). Return shipping costs are therefore in some cases not the same as sending shipping costs and have to be calculated differently. The cost that has more impact though is the duties refund. There are two common Incoterms applicable in e-commerce transports. The Deliver Duty Paid (DDP) where the seller is responsible for handling the risks and costs of the shipment, including import duty and any other charges related to delivery and the Deliver Duty Unpaid (DDU), where the customer is responsible for settling all charges in order for customs to release the shipment (Trade Finance Global, 2019). At Farfetch more than 80% are DDP, which means that the price charged to the customer includes import duties. Because FF refunds the total price to the customer and sending customs duties are not refundable, FF will incur this expense if the customer returns the product. For countries in the EU VAT is charged only if the purchase occurs, meaning that VAT will not be a cost if the customer returns the good. Opposed to that are the duties included in the customer final price that will be refunded in full if the customer returns the item. The problem is that once the customs clearance has been paid, that value is not refundable even if the purchase does not occur. Specially for luxury items this cost can be extremely heavy since Farfetch transacts high-priced items and risk of incurring this cost should be minimized. Moreover, some countries tax luxury products with higher taxes than others. Considering a product with a sales price of 1000 Euros and 40% duty rate, if a return occurs Farfetch loses 400Euros. To weight this cost in the formula a probability of return must be considered. The aim is to include the expected probability of return cost in the formula.

It must be noted that changing the order assignment criteria might select routes with higher duties. Considering the duties refund described in the paragraph before this can represent high costs and therefore it is important to control that effect.

3.2.3.3 Maximizing conversion: Wrong-item lost sale

The probability of a client giving up on the order in case a different item or size than ordered is received was analyzed. This means that besides all logistic costs of sending and returning the item in vain, the transaction will not proceed, which means that a lost sale has just been created. The average wrong item rate is at 3.3% of the returns for H1 2019. Looking at sales from 2016 to 2019 every 10 customers that receive a wrong item, only 1.5 will order the same item again. There are multiple possible factors. The customer can turn to another FF competitor or the item was for a special occasion and will not arrive on time. Moreover, it is possible that by the time of the complaint the stock was no longer available. To include this in the profitability formula, probability of wrong item of the store times estimated order profit value needs to be discounted from the order revenue. Since the order profit is highly dependent on the future allocation, for this purpose we consider an average profit per order to weight (10%) in the wrong-item probability in the formula.

Currently, there is no sophisticated prediction model to calculate the probability of a store sending a wrong product. It is known that after initiating an incentive plan that penalized vendors for giving no-stocks, some of them tried to ship wrong items to escape the penalty. Although some measures were taken to mitigate this effect, this indicates that the wrong-item probability is highly dependent on the store. On top of that, the likelihood to send a wrong item depends on the operational efficiency and quality control of the vendors. As mentioned above the average wrong-item probability is around 3% of 25% of returns, which means, 0,75%. In comparison with the other metrics this will not have much weight in the formula. Forcing this variable to be heavily penalized is not the best strategy, as profitability opportunity can be lost. Instead, the company should cooperate with the vendors and help them organize their operations in order to reduce this percentage.

3.2.3.4 Maximizing conversion and minimizing item swap costs: No stock cancellations

A item swap event is not mentioned in the reviewed literature as it is uncommon in a normal e-commerce business. As a marketplace does not own the inventory, it is necessary to check the stock availability every time an order is placed. The sellers can just decide to keep the stock or can eventually have sold it in the meanwhile at the physical store. Every time there is a no-stock Farfetch tries to find another seller available to replace the order, named internally as item swap. For this reason, in the algorithm the probability of lost sale due to no stock event is naturally lower than the average no stock probability of the business. For this not to happen, every other competing seller that had stock at the time of the order had to sell the items until the process finalizes the step of stock-check.

If a no-stock occurs, and a replacement store exists, Farfetch will "cover" any price difference to a maximum value. This way the customer will not perceive the change and experience is not affected. To achieve that, Farfetch incurs on costs. Although the customer is not notified, lead-time will increase. The average No stock rate is 1,3% (2019). From all stock outs, 23% can be fulfilled by other stores, which means that there is a risk of having to cover extra expenses in 23% of the cases and 77% of the cases originate a lost sale.

Chapter 4

Solution Approach Proposal

The following Chapter details the used techniques to deliver a solution to the presented problem. In the first Section the programming requirements and the design of the order assignment tool are described. The second chapter explains the new defined variables to be used in the tool to evaluate sellers. The focus will be on the order revenue margin, shipping costs and delivery time as it is expected that this initiatives will deliver most value to the company.

4.1 Building a Parallel Reality to Test Order Assignment Strategies

Following the waterfall methodology the development of the simulation program required a preparation phase in which all the requirements were determined and detailed. Afterwards all the features were coded and data warehouses integrated. The whole program is properly documented and automated to be used for other business needs and stakeholders. In 4.2.1 we introduce the data gathering and preparation phase. Then in 4.1.2 every step of the program developed in the scope of this thesis is presented in detail. Appendix A documents the order assignment tool.

4.1.1 Re-Creating Real-World Instances

Today, the candidates list and the used criterion for the order assignment is not being stored. Moreover there is no record of inventory levels at the time of the order. The aim of this project is to develop a simulation tool that returns for every order the selected boutique given a set of previously defined criteria. To make a strong business case it is relevant to present realistic and comparable results. To achieve accurate results, the inputs must be consistent with reality and have to be permanently updated, because the output of this simulation depends purely on external data that varies with time. Moreover, to validate and gain confidence in the performance of the tool, simulated decision results must be comparable with the real ones if the same business rules are applied.

The information required to create the test-instances was not readily available for analysts to use. It is spread across multiple data warehouses and tables. Data Wrangling is the process of data transformation so that information is ready for consumption by the data analysis tools and techniques. This process comprises activities of converting, modifying or merging the data in a certain format to generate well-defined streams of data. According to Rattenburg et al. (2017) "50 to 80 percent of an analyst's time is spent wrangling data to get it to the point at which this kind of analysis is possible. [...] Data wrangling comprises activities like understanding what data is available, choosing what data to use and at what level of detail, understanding how to meaningfully combine multiple sources of data and deciding how to distill the results to a size and shape that can drive downstream analysis." Different queries in Microsoft SQL Server and Google BigQuery were written during this project for extracting and cleaning the necessary data. The work developed provides the structure for any user to chose the time period of the required simulation and the test instances are automatically extracted to load into the simulation tool. Phyton Pandas was used in this project, as it can present data in a way that is suitable for data analysis via its Series and DataFrame data structures. This package contains multiple methods for convenient data filtering and has a variety of utilities to perform Input/Output and search operations with high-performance. It can read data from a variety of formats such as CSV, Excel, and SQL Bigquery.

The first step to make the decision is to know the available stock at the time one order is placed. Available stock means, number of SKUs (combination of product and size) in every single stock point that can fulfill the order in the country of the customer. There are many possible reasons for a certain store to have stock inflows and outflows. Sellers are allowed to sell in the physical store (offline) the items they have uploaded in the marketplace or simply decide to remove stock. On the other hand, boutiques' stock can be uploaded at any time of the day. Adding to this, when a sale on the online marketplace takes place the stock level has to be updated accordingly. To save the level of stock available in each stock point for every SKU a snapshot is made at the end of every day. Depending on the time of the season in one day there can be over 3 million records. To have the stock level at the time of the order we would have to have a snapshot every second, which is impossible.

The approach to overcome the information unavailability was to use a data table that keeps track of all stock movements that occur during the day. Whenever there is an upload or a sale, this information is registered on a table with logs, identifying the product, the size and the stock point ID. For simulation purposes, a day zero has to be defined and a snapshot of the system state at midnight is needed. To know the level of stock in a specific store at the time t of the order we need to add all uploads (including new inventory or returns), eventual offline sales and online sales that happened since stock units were counted at t0 as described in Equation 4.1.

$$\text{Available Stock}_{ij_t} = Stock_{ij_{t0}} + \sum_{t0}^{t} \text{uploads}_{ij} - \sum_{t0}^{t} \text{offline sales}_{ij} - \sum_{t0}^{t} \text{online sales}_{ij} \quad (4.1)$$

$$i = 1, ..., \text{number of stores and } j = 1, ..., \text{number of products}$$

To rebuild a real scenario of the past, the stock movements associated to online sales occurred during the analyzed period must be deleted. Unfortunately, the origin of movements (sale online/offline) are not properly identified. Since the stock is only removed after the check-stock step (see figure 3.1), there is a time gap between the date when the customer places the order and the date of the boutique confirmation. A logic was implemented to find for each order the corresponding record on the table (see Appendix A.5) .

The presumption that boutiques are responsive to sales, meaning that if they sell an item they will upload more stock afterwards is valid. For test purposes we assume that the supply is the same during the evaluated period, independent from order allocation. Real historical data is considered an input and it is assumed that the supply will be the same without accounting for changes in the boutiques behaviour caused by different assignment criteria.

Besides stock level, the price that each seller charges online, for all the products sold in the currency of the client is needed to make decisions. Since this information requires too much storage and processing memory, Google Big Query had to be used. Additionally, the store characteristics (country, region) as well as performance metrics (SOS, no stock, wrong item, packaging accuracy, lead time) on a daily basis are required.

Anticipating future features and requirements of the simulation tool, other variables besides price should be included. For example, to calculate geomargin this data set must include the retail price, taxes/duties, discounts and fees. While preparing the data sets, we had to include not only the variables that are used to make the decision but also new designed variables and variables to measure the impact of the different solutions (e.g. to evaluate the impact it is important to have always the daily exchange rate).

After creating the test instances, different algorithm variants can be applied, and the outcome can be predicted. This work is not only valuable for the project purpose but in general stakeholders would benefit from having that information on demand. For example, to be able to answer why some sellers are not selling although they have inventory.

4.1.2 Programming of dynamic order assignment process

To mimic the AS-IS system behaviour, a discrete event simulation was used, which steps are detailed in the Flowchart in Appendix A). An order triggers an event where a seller is selected based on previously defined criteria and stock level is accordingly updated.

For that purpose, a file listing the customer orders from a whole month with all characteristics has to be uploaded. Before applying the logic it is necessary to list all possible stores and fulfillment characteristics, which means that in every loop it is required to do search operations. The best suited method to implement that with Phyton programming language is Pandas' apply, which is a method takes functions and applies them along an axis of a Data Frame (all rows, in this case). For this purpose, the list of orders must be sorted chronologically.

Cancelled orders, due to fraud check, and itemswap orders do not trigger an order in the simulation. In this first phase of the project the following assumption is made: the simulator assumes that no store rejects the orders and therefore orders originated due to itemswaps are not included in the data set. It would be interesting to model in the future wrong items and no-stocks probability and given a certain distribution trigger itemswap-orders.

To validate the tool an accuracy measure was created, as expressed in Equation 4.2.

$$\text{Simulation Accuracy} = \frac{\text{Number of successful assignments}}{\text{Number of simulated orders}} \tag{4.2}$$

As the current algorithm has many exceptions to what was presented before, it was very complex to capture all the business rules and deliver a good accuracy. To fully capture the running algorithm, we had to iteratively add more rules until we had an accurate solution and could use it as a baseline for comparison. This was a good approach to deeply understand how the current decision process works.

Besides business rules, pricing rules had to be built-in, because depending on the product brand, customer tier and country different prices and discounts are offered. Moreover, having stock does not necessarily mean that the item is available for every customer. There are blocking rules, such the case of some stock points that can only fulfill orders in specific markets or certain products that are not available in certain markets due to offensive content. Every time the algorithm returns the winner boutique the level of stock has to be updated based on the decision. If a boutique in simulation sold an item online but, in reality had sold it offline moments later we assume that Farfetch would have captured that sale and that the offline sale won't take place.

Having built these foundations, removing the current sequential filters and apply a formula to calculate the most profitable allocation can be easily coded. As the simulation goes forward in time the online sales on the marketplace will be registered according to selection model. To measure the "entropy" introduced by changing the assignment rules the number of different assignments in comparison to the baseline was measured (see Equation 4.3).

$$\text{Sellers' Impact} = \frac{\text{Number of different assignments}}{\text{Number of simulated orders}} \tag{4.3}$$

The order assignment tool allocates the order to the selected seller and assumes that it will accept and fulfill the order.

4.2 Design of the Variables to Rank the Sellers

After describing in Chapter 3 the business characteristics and requirements, potential variables to include in the order assignment process were collected. In order to make decisions based on new variables, these will have to be either calculated, in case the required data records exist, or estimated for each eligible merchant. As the scope of this project is wide, for sake of prioritization, the activities were categorized in four categories: quick wins, major projects, fill ins and time-wasters, as illustrated in Figure 4.1.

Figure 4.1: Impact-Effort Matrix of initiatives that impact bottom-line profitability

The impact of each initiative is a ballpark estimation, because the outcome is very susceptible to demand dynamics, sellers assortment and performance. To prepare all these variables to go live would take a long time and require extra capacity. Additionally, order assignment is a core activity that needs to be stable and predictable. To experiment and explore different strategies live would generate to much instability. Hence, to support this project a robust simulator capable of testing on historical data different allocation rules was developed.

The focus will be on the revenue margin, shipping costs and delivery time as it is expected that this initiatives will deliver most value to the company. Packaging quality and consolidation would require extra effort to implement and the expected impact is lower than the other variables, as explained in Chapter 3. Profitability is at the top of the priorities list, as it one major challenge faced by the company today. The delivery speed is considered a quick win, because it is in line with other business initiatives to improve delivery speed. In the future, after the design of the order assignment strategy is optimized, it is worth analyzing the impact of re-evaluation of the assignments in bulks, as explained in Chapter 2.

4.2.1 Order Revenue Margin

A good shopping experience offers a consistent price throughout the customer journey. The business decided to offer the customer the best possible price, and therefore price is an input variable in this project. However, this does not necessarily mean that boutiques with higher prices can not be better alternatives as there are other costs to consider. For this reason, it is important to calculate the margin for each alternative. Moreover, if the product is geopriced, meaning that the price for a customer in a specific market is fix independently of the vendor, and vendors charge different prices, it is relevant to evaluate the margin obtained.

The net retail price is the amount charged by a seller in the physical without local VAT or taxes. The difference between price and margin can be fully or partially captured by Farfetch, depending on the margin capture rate (MCR), determined in the contract. The take value is a blended commission rate (CR) charged on each transaction. Integrating this variable while considering the different alternatives to fulfill the order is extremely relevant to drive, profitability, as this is the most important revenue driver for the company. The formula 4.4 calculates the total revenue obtained by FF from selling a certain product of a seller for a given price. The dissociation of price and seller selection in the order assignment process expands the solution space of the problem. In the AS-IS situation, the price charged online is inevitably associated to the seller selection. Furthermore, this formula is applicable to both standard and geographical pricing schemes, maximizing geomargin and eventual existing fees such as shipping subsidy and exchange-rate fee (see Price Criterion Section in Chapter 3).

$$\begin{aligned}\text{Order Revenue}_i = {} & [\text{Price Charged} - \text{Net Retail Price}_i - \text{Taxes}_i] \times \text{MCR}_i \\ & + \text{Net Retail Price}_i \times \text{CR}_i \\ & + \text{Price Charged} \times \text{Credit Card Charges}_i \\ & + \text{Price Charged} \times \text{Free Returns Fee}_i\end{aligned} \tag{4.4}$$

4.2.2 Shipping Costs

Farfetch offers flat shipping fee for orders above specific thresholds and makes many free shipping campaigns. Sometimes the company makes profit with shipping charges, other times loses money. All this creates increasing pressure to optimize shipping costs.

In order to include the transport costs in the order assignment decision it is necessary to estimate the shipping cost for each of the alternative sellers. Currently, the customer can choose from different delivery type options. Depending on the selected boutique, only the available services for that route are displayed to the customer. Farfetch works with more than one carrier to ship globally. As every order is unique, shipping costs will vary depending on the type of service, carrier, size, weight and shipping route. As premium services still represent a small percentage of the shipments, only Express and Standard shipping options, which make up to 99.5% of the orders, will be modelled. From Jan 2018 to Mar 2019 5160 possible routes were registered.

Instead of using a multiple regression, which would require to compute 242 dummy variables (52 and 190 shipping and delivery countries respectively), a simple linear regression is modelled for every possible combination of origin country, destiny country and shipping type service. The main advantage of this approach is that only existing routes can be considered. The risk of the multiple regression formula is that impossible combinations, as is the case of standard service Italy-US could be computed. Some of the results of this analysis are summarized in Figure 4.2 and Table 4.1. A deeper analysis on the shipping costs model can be found in Appendix B. Since prices of carriers are determined depending on route, billable weight and type of service, by looking at historical data little variability of shipping cost was identified as expected for the same carrier, same weight and same route. To arrive to this results one has to consider that, depending on the carrier, the weight can be in kilograms (kg) or in pounds (lbs) and billed value in different currencies. From all combinations only 3467 linear regressions could be plotted, as 1583 routes only occurred once and the others do not have weight information (routes from brazilian and asian carriers). Only 200 of the models have a coefficient of determination (R^2) lower than 90%.

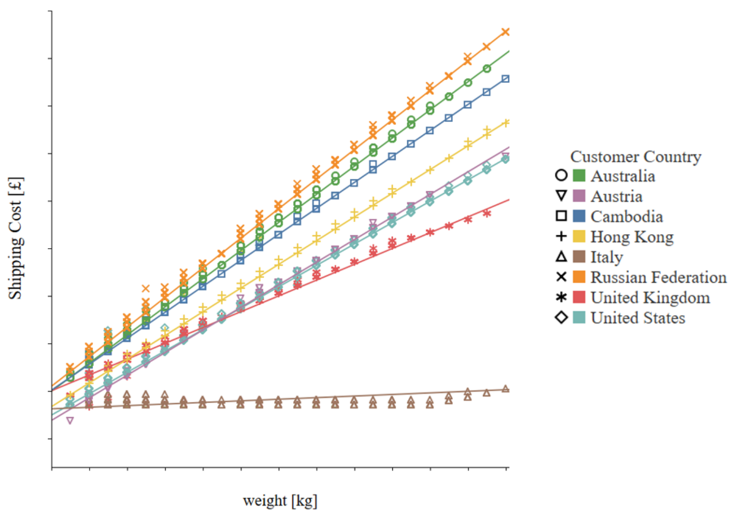


Figure 4.2: Example of shipping cost analysis from Italy by delivery country

Table 4.1: Linear regression models to predict Italian Express Shipments

Route	Slope	R^2	n	MAPE
Italy-United States	2.25	99.3%	260942	-0.0024
Italy-Russia	3.11	99.8%	99647	-0.007
Italy-Hong Kong	2.49	99.8%	80522	-0.091
Italy-Italy	0.17	99.8%	6177	0.2927

During the data cleaning and preparation phases, some clear patterns in the shipping costs were recognized. First, a recurring deviation of 6.5 USD associated with surcharges for remote area was noted. Deliveries from certain stores are charged an extra remote area pick fee whenever they ship an order to the client. It can be argued that this cost should be weighted in the algorithm ranking formula, which will consequently penalize sellers located in remote areas. Considering the business model of Farfetch, if the seller has potential to grow this fee should be renegotiated with the carriers instead of penalizing the sellers. To properly model the shipping costs the remote area fee was removed from the data, which increased the coefficient of determination (R^2). The stores located in remote areas were registered in a file, in order to add this fee to the predicted shipping cost in the cases, where these sellers are possible candidates to win an order. Another important conclusion of the analysis performed on the shipping data the variability of US domestic shipments as illustrated in Figure B.3. Naturally, it is different to have an order from New York to Alaska (117 USD) or New York to New York (6.5 USD). To boost profitability and make better decision, it is advisable to include, at least, the information of the state of the customer in the decision.

Moreover, the weight of the packaged used to ship the order needs to be known. It is hard to predict the weight when an order is placed because it greatly depends on the item and box choice of the seller. Complicating this is the fact that the billable weight is the larger between the package's actual weight and dimensional weight.

$$\text{Dimensional weight} = \frac{(length \times width \times height)}{5000} \tag{4.5}$$

For every product the production department, where online products' photo shoot takes place, makes a recommendation for the best suited packaging box and measures have been introduced to make sellers comply. For the comparison of the shipping costs the most frequent weight for the combination item category and brand is used. In cases where the brand product type does not exist, the most frequent weight for the category is used. Using this model the mean average percentual error of a sample of 268 orders is 3.75%. On average the prediction of the weight with this method fails by -0.5 kg. As this will be used to compare sellers for the same product shipment this will have a very low impact and therefore this approach is considered valid.

4.2.3 Lead-time

As identified in Chapter 3 the fulfillment time of an order depends on both the assigned seller and its characteristics, and type of shipping service. Enough observations between most of the vendors and customer country exist to predict the delivery time using historical information. Analysis show that SoS is mainly influenced by the day of the week, store city, store and carrier. As for TiT, analysis show that variations are mainly explained by carrier (DHL mostly Express and UPS standard services), customer country, customer state store and store location. Interesting is that the store city has more impact than the store country, as certain locations have more sophisticated delivery networks.

Figure 4.3 shows that stores in the same country have different lead-times. Therefore, it is important to consider the pairs (customer country;store country).

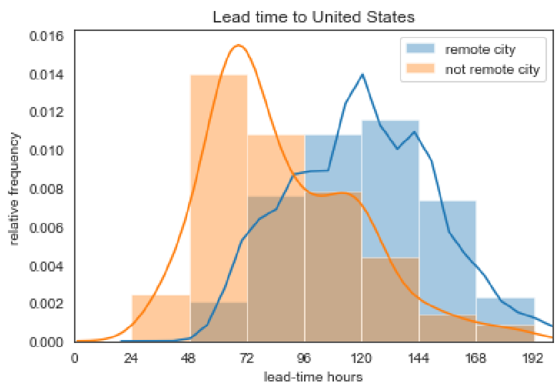


Figure 4.3: Impact of Store in Delivery Time

After identifying the variables that can impact the lead-time, we assumed that using the average lead-time between store and customer-country, excluding outliers (i.e., considering only values within the interval ([mean - 3 $\times\sigma$;mean + 3 $\times\sigma$]) , is acceptable for estimating the project impact. To model properly the lead-time it is highly important to know the customer state and city. As this data do not have sufficient quality to be used, the average seems to be a good proxy to evaluate the impact of the seller selection in the order fulfillment time. The estimation of the delivery date is obviously associated with a degree of uncertainty that has to be taken into account. In a more advanced state of the project, the results from a more robust model will be integrated in the simulation tool to guarantee a successful implementation.

Reichheld (2006) recommends to measure the annual spend of a customer to understand if promoters increase their order quantity and frequency. To transform the variable in a comparable unit to include in a comprehensive formula for the order assignment process, the total amount spent in the following 12 months was measured to try to calculate the impact of a faster or slower deliver experience. The analysis of historical data of the past three years suggest that for a sufficient number of observations customers tend to purchase more in the following twelve months if they had a better delivery time experience.

One has to take into account that different markets react differently to service improvements and therefore a curve has to be designed for each customer country. This analysis does not consider customer city due to low quality of the data, nor VIP customers as they are considered already "loyal" to the company and have a different behaviour. However, this could be misleading, as the hypothesis that customers closer to luxury boutiques have necessarily more purchasing power (and therefore spend more money) is not explored. To remove that effect the relative amount spent was analyzed in Figure 4.4.

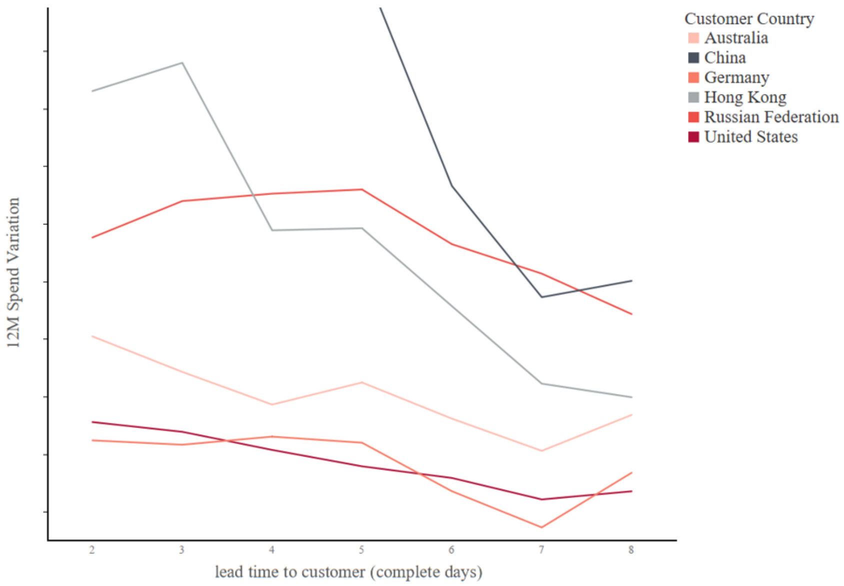


Figure 4.4: Relative amount spend in the next 12 Months after a specific lead-time experience

As the results reveal that for a significant number of observations, there is a clear reduction in the amount and relative amount spent in the next 12 months after a delivery experience. Consequently, a linear regression over these series is estimated (4.4). The expression 4.6 will assign better lead-time service to customers that purchase higher-priced items.

$$\begin{aligned}\text{12M Lead-time Value} = (a + b \times \text{Lead-Time}) \times \text{Order Take-Value},\\ \text{where Order Take-Value= Price of the Order} \times \text{Average Take-Rate per Order}\\ \text{a = intercept of estimated regression from series in Fig.4.4}\\ \text{b= slope of estimated regression from series in Fig.4.4}\end{aligned} \tag{4.6}$$

4.2.4 Performance Metrics - No stock and Wrong-item

Currently the business is using the average no stock (see 3.1 in Chapter 3). We have enough historical data to estimate in a more sophisticated manner how stores behave in terms of no-stock and wrong-item. Seasonality, stock depth as well as number of items sold are possible factors impacting vendors performance. It is very likely to assume that sales that have just one SKU available originate more frequently a no stock. Furthermore, it is worth studying for how long the item has been online. Probably if the product has been uploaded one year ago it might give a no stock. In case the item has been uploaded very recently then there can be some synchronization problems that might cause a no-stock. As estimating the the no stock probability is not the final purpose of this project the average no-stock of the past two weeks will be used as a ranking factor and sellers with higher stock depth will be favoured. If communicated properly to sellers, the use of the average historical performance can motivate sellers to have better performances. Instead of predicting the best solution it rewards vendors for past quality service.

As mentioned in Chapter 3, the from all stock outs, 23% can be fulfilled by other stores, which means that there is a risk of having to cover extra expenses in 23% of the cases and 77% of the cases originate a lost sale. Although, this probabilities might be slightly different considering only products ordered in more than one location, it is considered a good proxy to include the performance perspective in the decision process, as expressed in expression 4.8. The same logic applies to the wrong item probabilities of repurchase event or lost sale (see Chapter 3).

$$\begin{aligned}\text{Expected No Stock Loss} = {} & \text{Order Revenue} \times \text{P(NoStock)} \times 77\% \text{ (lost sale)} \\ & + \text{Average Item-Swap Cost} \times \text{P(NoStock)} \times 23\% \text{ (item swap)}\end{aligned} \tag{4.7}$$

$$\begin{aligned}\text{Expected Wrong Item Loss} = {} & \text{Order Revenue} \times \text{P(wrong item)} \times 85\% \text{ (lost sale)} \\ & + 2 \times \text{Shipping Costs} \times \text{P(Wrong Item)} \times 15\% \text{ (repurchase)}\end{aligned} \tag{4.8}$$

Chapter 5

Results

In this chapter the results of the different identified initiatives supported by the order assignment tool described in Chapter 4 are presented and analyzed. All scenarios are comparable as they apply different algorithm variants to the exact same data from the same month. The first section of this Chapter describes the results from the AS-IS scenario, which mimics the system behavior in the past, giving visibility over the current decision process and quantifying the improvement potential of this project. Moreover, the results validate the order assignment tool. In the second section the results from different order assignment strategies are presented. As mentioned before, in order to have comparable scenarios it is important to assume that the price charged to the customer will always be the minimum price possible, regardless of the criteria used to select the seller. As the company assumes a commitment of offering the best price, the scenarios below calculate the margin assuming that the price displayed to the customer is the minimum plus 5%. This can be further adjusted depending on the pricing strategy that the company wants to follow. At the end, a final recommendation is proposed.

5.1 Current Decision Process (AS-IS) and Tool Validation

In order validate the order assignment tool 15000 real orders of April where used. The ambition at first was to run the simulation for a whole season in order to capture all effects of full-price and discount phase. Unfortunately, due to data quality in the year of 2018 and performance reasons this was not possible. Despite efforts to optimize the code, 15000 orders take 4 hours to run due to limited CPU performance.

The results show an accuracy over 90% calculated according to the expression in 4.2. There are some reasons that explain the 10% error. First the simulation does not include all the orders. For example, if in reality a boutique gives a no-stock an itemswap is possible. However, the developed tool assumes that every order is fulfilled and therefore itemswap orders are not included, which means that stock that was in reality assigned to fulfill itemswaps is available in the tool to fulfill other orders. Additionally, after an order is finalized by the customer the boutique has to confirm stock. If a stock hold occurs, the product is temporarily unavailable. The simulator does not capture that effect and therefore, it is impossible to have 100% accuracy. Also some temporary item specific blocking rules are not included in the program, as effort to include them are higher than the gains obtained from that precision. In case stock was recently uploaded, it might be the case that operational/live systems do not receive information immediately and therefore discrepancies between reality and simulation is possible. It is worth mentioning that in cases were an asynchronism of information was detected at the check-out step it automatically switches to another seller, which is also impossible to identify. At check-out other information such as seller is on holiday is confirmed. All this factors negatively influence the accuracy rate. Despite of all these simplifications, the real winner is most of the times on the list to be considered to fulfill the order. It only does not win, because a better one still has stock, which means that the built scenario is realistic enough to deliver high-quality insights about the AS-IS business performance and evaluate the impact of changing the order assignment process. The evaluation of the impact is made by always comparing the results with the AS-IS obtained with the order assignment tool instead of real ones to mitigate the risk of introducing other variation effects. Since the comparison is made between simulated results, the calculated variations are authentic and reliable.

By examining the results, one can draw the conclusion that this algorithm only impacts 47% of the orders (see Table 5.1), as more than half of the products orders are located in one single stock point. Considering a Gross Merchandise Value of $1.4 Billion in 2018 means that $0.7 Billion USD actual sales go through this decision algorithm. The simulation tool shows that on average each order is disputed by around 5.2 competitors. Table 5.1 shows that price filter reduces the solution space to one option in 19.6% of the orders, without considering any performance aspect or costs. Nevertheless, most of the decisions are made at the score-level, which shows that there is room for improvement and that even if the price logic is kept, changes can have great impact on the company order fulfillment performance. As expected, due to supply concentration (see Figure 3.5 in Chapter 3) the criteria location has low impact and most of the decisions do not consider lead-time. The last filter, selecting the oldest product has little relevance. During execution, some exceptional rules of the algorithm were identified and are therefore grouped as "other" in Table 5.1.

It is worth mentioning that simulation was tested in different periods of the season and the decision breakdown was consistent. Looking at the results this project has the potential to change 47% of the order assignments as 53% are unique products. It is important to refer that the exceptional rules can not be removed for strategical reasons and therefore will not be impacted.

Table 5.1: Pipeline decision steps of the AS-IS algorithm and % of orders decided at each step

Criteria	Percentage of decisions taken by the criterion
Unique seller	**53.0%**
Price	**19.6%**
Location	**2.4%**
Country	1.7%
Region	0.7%
Score	**18.0%**
Country	0.8%
Region	6.0%
World	11.2%
Product creation date	**2.6%**
Country	0.05%
Region	.85%
World	1.70%
Other	**4.4%**

5.2 Different Order Assignment Strategies

5.2.1 Scenarios Overview

In order to design the best solution for the order assignment process **six strategies** were tested on the same orders of April 2019.

The strategies are summarized in Figure 5.1. **Profitability 1.0** (S1) is the maximization of the Order Revenue (see expression 4.4). In case more than one seller exists with the maximum revenue then lead-time and stock depth are used as tie breakers. **Profitability 2.0** (S2) and **Profitability 3.0** (S3) follow the same structure replacing the revenue maximization by the expressions in 5.1 and 5.2 respectively.

$$\text{Order Profit} = \text{Order Revenue} - \text{Shipping Costs} \tag{5.1}$$

$$\text{Expected Profit} = \text{Order Profit} - \text{Loss from WI and NS} \tag{5.2}$$

Balanced 1.0 (S4) strategy selects sellers that offer values of Expected Profit (see 5.2) included in the fourth quartile of the alternatives and afterwards selects the seller with the lowest lead-time. This is a alternative to find a balance between lead-time and profitability without having to express the value of lead-time in dollars. This approach makes unnecessary to establish a relationship of value and delivery time. The alternative **Balanced 2.0 (S5)** uses the analysis to transform lead-time in USD$ performed in Chapter 4 (see expression 4.6).

Finally, **Delivery Speed (S6)** serves to test the full potential of the business in terms of lead-time, assigning the order to the fastest alternative. This can be seen as a "mood" to gain competitive advantage, that would be activated for a specific period during the season to react to competitors, like Amazon launching 1-day shipping.

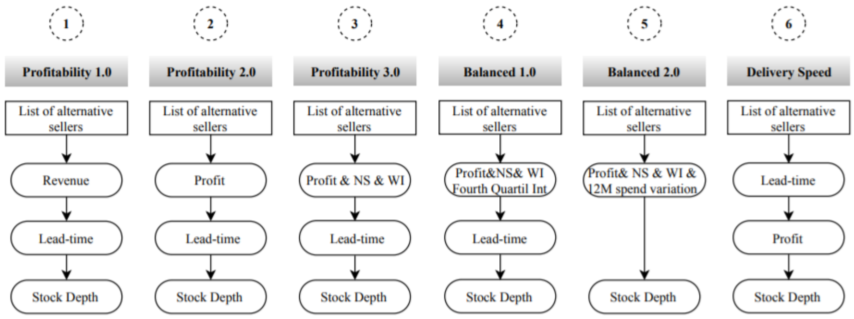


Figure 5.1: Tie breakers of the six strategies tested for the order assignment

The impact on the most important KPI by applying this dispatching rules is detailed in Table 5.2. Strategy Profitability 1.0 (S1) shows that the order revenue can be maximized up to 2.5% relative to the AS-IS. In order to be comparable the margin of the AS-IS considers that price charged to the customer is the minimum price, although the assignment is based on the sequential filters of the AS-IS process. In this scenario, as in all other explored strategies, the duty rate increases. This is the effect of not leveraging domestic shipments as it is done in the location criterion of the current algorithm. As the price charged to the customer is the same, this will not affect conversion. Since order revenue deducts duties this does not impact the net profit of the solution. However, as mentioned in Chapter 3, duties refunds have to be considered. In case all the orders assigned to routes that have heavier duties are returned, more costs are incurred as in the AS-IS scenario. Comparing duties increase cost considering the average return rate, the net profitability is still positive. As expected around 17% of the assignments are changed.

As the previous scenario shows a significant increase of the shipping costs it is important to design a second scenario (S2) to include them in the decision process. As this is the only difference to the former scenario, this allows to examine the "isolated" effect of the inclusion of this variable. By adding the shipping costs the revenue is not affected but the shipping costs are reduced comparing to S1, as more decisions are made at this level and less in the lead-time criteria resulting in a slower solution (see Table 5.3). This shows that the shipping costs variations are small in comparison to variations of the revenue. Although shipping costs are important to consider, specially for small boxes (1 to 2 kg which represent most of the orders from FF) the cost difference between routes is not as significant as profitability discrepancies. As the order of magnitude of revenue is higher than the shipping costs the order profit is positive, showing a net result of 2% increment in net order profit as compared to the AS-IS process. Considering duties refund this means 1.69%. Note that the no-stock and wrong-item rate are also negatively impacted. Specially the NS is affected, as 20% of the decisions in the AS-IS scenario are made at the score criterion level (which considers NS Rate).

In the third scenario (S3), the NS and WI terms added to the formula lead to an increased number of decision made at the margin level, which leads to lower lead-time reduction. Using the average rate of the past two weeks and weighting both factors as a probability of lost sale or extra costs, improvements in these metrics can be achieved. Here, the impact on sellers is lower as it accounts for the NS metric that is present in the AS-IS criteria.

The fourth scenario (S4) shows satisfactory results improving net profit as well as lead-time, NS and WI rate. Curiously, the strategy Balanced 2.0 (S5) shows slightly the better results in terms of profitability and lead-time than the previous strategy. However, the performance metrics of NS and WI have stronger reductions, which reveals that the calculated value of lead-time has higher impact than the weight of these metrics. Strategy S6 shows the best results in terms of lead-time as expected. However, the loss in net profitability is significantly higher than the other scenarios to be even considered.

The last strategy (S6) shows a potential reduction of 4.5% in lead-time. However, this comes at cost of sacrificing 5.5% revenue and increasing shipping costs resulting in a profitability decrease of over 7%. Moreover, duties increase by 12.2% which means that the potential duties refund cost is higher than in any other scenario.

Table 5.2: Impact of replacing the assignment method for order different strategies

KPI	S1	S2	S3	S4	S5	S6
Net Result[a]	+1.60%	+1.69%	+1.59%	+0.66%	+0.67%	-8.00%
Order Profit	+1.96%	**+2.06%**	+1.98%	+1.10%	1.07%	-7.16%
Order Revenue	**+2.54%**	+2.50%	+2.45%	+1.7%	1.63%	-5.5%
Shipping Costs	+4.32%	+3.28%	+3.39%	+3.52%	**+3.26%**	+4.2%
Duties	+10.8%	+11.0%	+11.4%	+11.7%	**+10.53%**	+12.2%
Efficiency						
Lead-time	-0.4%	-0.1%	+0.2%	-1.0%	-1.14%	**-4.5%**
No-Stock	+8.1%	+9.5%	**-1.8%**	-1.4%	+0.4%	+6.0%
Wrong-item	+1.8%	+2.1%	**-3.6%**	-2.6%	-1.85%	+2.2%
Sellers Impact						
% of Changed Assignments	17.5%	17%	**16.8%**	17.7%	18.1%	24%

[a]The calculation of the net result is conservative as it is considering a probability of duties refund

$$\text{Net Result} = \text{Order Revenue} - \text{Shipping Costs} - \text{P(Return)} \times \text{Duty Rate} \times \frac{\text{Price}}{(1 + \text{Duty Rate})} \tag{5.3}$$

As expected the percentage of unique sellers impacts 53% of the orders (see Table 5.3). However, the "other" filter has increased. This is because those criteria where applied after price and in the scenarios simulation it was decided to prioritize this rules above all criteria. However, profitability loss is not significant to negatively influence the results.

Table 5.3: Pipeline decision steps of order assignment strategies and % of orders decided at each step

Criteria	S1	S2	S3	S4	S5	S6
Unique location	53%	53%	53%	53%	53%	53%
Others	7%	7%	7%	7%	7%	7%
Revenue/(Expected) Profit	31%	33%	39%	27%	38.7%	4%
Lead-time	7%	5%	0.2%	12%	-	34%
Stock-depth	1%	1%	0.4%	0.3%	0.3%	1%
Random	1%	1%	0.4%	0.7%	1%	1%

As expected the number of domestic shipments was reduced in comparison to the AS-IS.

Table 5.4: Routes Analysis

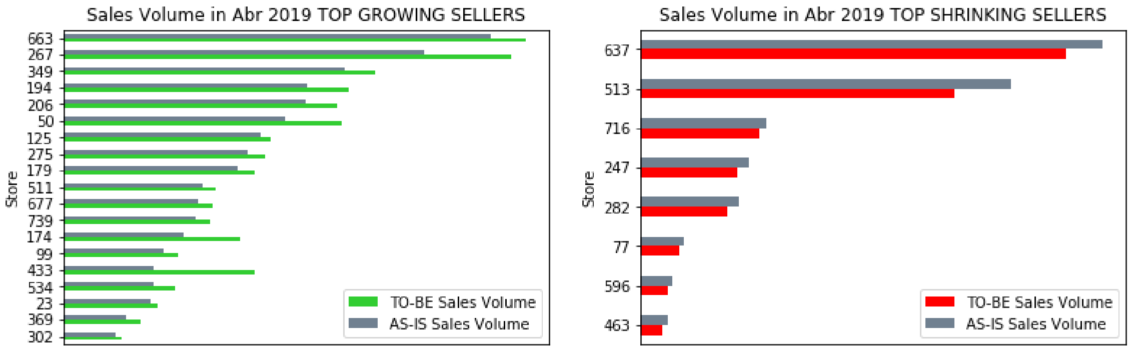
	AS-IS	S1	S2	S3	S4	S5	S6
Standard	15%	1%	1.8%	0.9%	0.9%	0.8%	0.2%
Express	85%	99%	98.2%	99.1%	99.9%	99.2%	99.8%
International	60%	62%	62%	61%	62%	62%	60%
Intra Region	31%	31%	31%	31%	31%	31%	32%
Domestic	9%	7%	7%	7%	7%	7%	8%

5.3 Proposed Strategy Analysis

As identified in the previous scenarios, delivering faster comes at cost of less profitability. Nevertheless, as analysis in this thesis providing a faster delivery option can encourage customers to buy more and become more loyal. Therefore, both balanced scenarios seem to be a prudent compromise that balances profitability and customer experience. As the net result and lead-time do not differ much, as Balanced 1.0 improves NS and WI performance, this approach is preferred. If the company wishes to drive profitability the proposed interval (4th quartile) can be re-adjusted to get closer to Profitability 2.0 (S3). The proposed solution improves profitability (revenue and shipping costs) as well as lead-time, no-stock and wrong-item rate.

Although this strategy does not compromise customer service, the seller's sell-through is affected, which is represented in Figure 5.2.

This shift in sales is particularly worrying if the partners that are positively impacted have a low sell-through history (see Expression 3.3), which means that there is less demand than supply and partners that sold less in the simulated scenario will have a lower sell-through than expected at the end of the season in case the algorithm is implemented. Otherwise, the only change will be the timing when the sellers sell the items. Due to promotions, this can still represent less profit for the partners that sell will sell later according to the new criteria.



(a) Positive impact (b) Negative impact

Figure 5.2: Impact on the sales distribution of sellers of type T0 - differences above 10%

This scenario returns for every order the most profitable allocation. The advantage of this approach is that we can assign to each seller a value for each lost sale. The decision maker can therefore evaluate if it's worth protecting some sellers at the cost of sacrificing profitability. For example store 637 in Figure 5.2 b) lost 8% of the sales. Comparing the metrics of the sales that used to be performed by this seller, it is possible to conclude that the past assignments to that store costs the company on average 3$ USD and 0.1 day per order. On the other hand, seller 267 in (a) increase its sales by 24% because it generates on average more 3$ USD than the assignments according to the AS-IS process. The positive impacts should be carefully analyzed, in order to guarantee that this new solution does not exceed some sellers capacity, which can negatively impact overall performance. As each order is processed individually, smaller partners might have to add extra personnel to fulfill all the orders. This tool is of practical value to anticipate such cases.

Chapter 6

Conclusion and Future Works

In this chapter the main achievements of this project are presented.

6.1 Conclusion

This project aimed at identifying effective strategies for order assignment decision process, that requires managing a complex distribution network involving more decisions than a traditional retailer. The goal was to explore key variables that can be included in the future order assignment decision process and evaluate the impact of a new proposed solution.

The current criteria of price, location, score and creation date applied sequentially are not sufficient to find the best match between customer and seller. Through literature reviews, benchmark analysis and acquired business knowledge new variables to select sellers in a marketplace context were designed. Based on a quantitative and qualitative analysis of the variables impacting the order fulfillment, it can be concluded that profitability, customer experience and seller success are important perspectives to consider when designing the order assignment process of an e-marketplace player.

The current order assignment decision process was assessed and different algorithm compositions for the order assignment problem using a virtual assignment program were tested, which allowed to estimate the impact of the new design solution against the current process.

Before this project three main limitations to improving the order assignment process existed. First, historical records of past decisions were not available and needed to be reconstructed, in order to make a proper diagnosis of the AS-IS situation. Second, as the process contains a classic multi-criteria problem, in which decision-makers are confronted with conflicting business objectives, there were difficulties assessing the weights of the criteria. Third, there was resistance to change the current algorithm, as the impact on sellers sales was unknown. To address these hurdles an order assignment tool was developed to give visibility of the alternatives and of the effect of changing the business rules of the algorithm on the sellers' performance, and predict potential improvement.

This project raised awareness of the inefficiencies of the AS-IS process. The developed tool enables a detailed assessment of the order assignment process. This project suggests the adoption of new evaluation criteria and the use of a comprehensive formula, that assesses a collection of relevant variables that impact order fulfillment profitability and business performance. The impact in the key business metrics was measured for six different strategies, showing a potential of 1.69% increment in net results (see S2) as compared to the AS-IS process or a potential reduction of 4.5% in lead-time, which is unacceptable in the present context as profitability would decrease by more than 7%. An intermediary solution is proposed balancing profitability, efficiency and sellers impact. The results indicate that there is potential to improve profitability and efficiency metrics simultaneously (see S4 and S5). Very satisfactory insights were obtained on the full potential of profitability increase, assuming that the price charged to the customer will always be the minimum price possible.

The level of sophistication and detail of the tool allows to estimate the impact of different what-if scenarios on profitability and service metrics with high accuracy and reliability. The proposed strategy has shown significant improvement in terms of profitability and delivery time, which led to budget approval for project roll out next year, creating a sense of urgency in improving the algorithm performance. This tool helps managers to make more informed strategic decisions. Any algorithm changes considering new business requirements that may arise can now be predicted just by changing the criteria. The exhaustive study of the order assignment process results in a compilation of possible alternatives for the future order assignment process.

All in all, it is possible to know the full potential of the project and anticipated results, which makes all stakeholders comfortable with change. This solution includes an automated process to monitor seller's sales evolution in the different scenarios to overcome any limitations imposed by the commercial team. This project can deliver a list of the most penalized sellers and the reason for that loss, which can now be analyzed by commercial teams in order to take actions. Either motivate those sellers to improve their operational performance (Lead-Time, No Stock, Wrong Item) or renegotiate commissions. If shipping costs are the cause of some strategies can be designed around that problematic.

This tool also provides important business insights to be used by different business departments. The visibility over the decisions that did not exist prior to this project, is very valuable to communicate with sellers and explain the reasons for them not selling as expected.

Concluding, this tool helped the company identifying the potential revenue and efficiency gains without having to take extraordinary measures and therefore demonstrating that Discrete Event Simulation can be an effective tool for making more informed decisions.

6.2 Future Works

There are three work streams that were raised by this project. First, the **order assignment decision process** can be further sophisticated by investigating the probability of no-stock (or wrong-item) and by integrating the lead-time prediction model. In order to include the return costs of a given order a prediction model to estimate the likelihood of a given customer to return an order is required. This is extremely important as all the proposed strategies increase duties and therefore duties refunds. Moreover, the recommendation to include the shipping state from the US. Finally, other methodologies to determine the weights of the criteria can be used and other variables referred in this thesis, such as packaging accuracy or sellers' success metrics, can be included in the formula according to business requirements.

The second area of improvement is the **order assignment tool**. Besides optimizing the tool run time in order to be able to predict the impacts in a whole season, it would be interesting to simulate different sellers' performance concerning no stocks and wrong items and trigger item swaps in the simulation. Moreover, as the benefits of consolidating one order in the same store have been studied in this project this feature should be included in the order assignment tool to make a stronger business case. Finally, changing the algorithm will have major implications that need to be communicated in order to guarantee a smooth and successful implementation. There are many forecasts in the Supply Chain team that rely on historical data (impacted by the current assignment policy) to forecast inventory replenishment and distribution. Also, sellers purchase based on the sales of the past year. Since the structure exists, more than what if analysis can be conducted. If demand and supply are properly modeled this can be easily used to forecast sales distribution or routes density for example. This is possible as FF has valuable historical data.

Third, in what concerns **customer experience improvements** in order to maximize overall satisfaction among customers and optimize order assignments, several delivery time options could be offered and the choice for faster or slower delivery shifted to the client, who could choose between more than one time-window option. Depending on the customer choice different prices are charged according to the calculated order profit. This solution would lead to increased customer satisfaction as it pleases both customers that value speed and are willing to pay for it as well as customers that are more price-sensitive. With this service a customer expectation is created, and the business is committed to deliver in that period. This would define clear targets for vendors and carriers to meet and delays could be monitored and accordingly penalized. Estimated delivery date deviations could be used as a variable in the order assignment process. The operations can be further improved by integrating an auction model, where the first to answer the call will receive the sale, after identifying a group of sellers, which profitability and estimated time in transit are identical. This can enhance the speed of sending and customer will consequently receive the products as soon as possible.

Bibliography

Aaron Cheris, Casey Taylor, J. H. and Davis-Peccou, J. (2017). How to cut carbon emissions as e-commerce soars. Retrieved from Bain Company Website. https://www.bain.com/insights/how-to-cut-carbon-emissions-as-ecommerce-soars/ (accessed May, 2019).

Acimovic, J. (2013). Lowering outbound shipping costs in an online retail environment by making better fulfillment and replenishment decisions. *Thesis (Ph. D.)–Massachusetts Institute of Technology, Sloan School of Management, Operations Research Center.*

Anylogic (2019). in url:https://www.anylogic.com/use-of-simulation/. (accessed March 2019).

Arnett, G. (2019). Luxury brands lose out to online marketplaces with super-fast delivery times. *Vogue Business*.

Bhattacharya, A., Geraghty, J., and Young, P. (2010). Supplier selection paradigm: An integrated hierarchical QFD methodology under multiple-criteria environment. *Applied Soft Computing*, 10(4):1013–1027.

Boyajian, L. (2017). The Real Cost of Returns for Retailers. url:https://retailnext.net/en/blog/the-real-cost-of-returns-for-retailers/. (accessed May, 2019).

Chen, C.-C. and Chen, Y.-J. (2014). A Dynamic Pricing Model for Time Slot Management in Attended Home Delivery Logistics. *Transport Research Arena 2014, Paris.*

Chen, L., Mislove, A., and Wilson, C. (2016). An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace. *International World Wide Web Conference Committee*, pages 1339–1349. doi:10.1145/2872427.2883089.

Danziger, P. N. (2018). Why Farfetch Could Command A $6 Billion IPO. url = https://www.forbes.com/sites/pamdanziger/2018/08/21/why-farfetch-is-worth-6b-in-its-pending-ipo/ (accessed May, 2019).

Deloitte (2017). The new luxury consumer. Retrieved from Global Powers of Luxury Goods Annual Report.

Deloitte (2018). Shaping the future of the luxury industry. Retrieved from Global Powers of Luxury Goods Annual Report.

Deloitte (2019). Bridging the gap between the old and the new. Retrieved from Global Powers of Luxury Goods Annual Report.

Golden, J. (2017). Four Questions Every Marketplace Startup Should Be Able to Answer.

Google (2013). How Affluent Shoppers Buy Luxury Goods. https://www.thinkwithgoogle.com/consumer-insights/fashion-online-affluent-shoppers-luxury/ (accessed May 2019).

Grasso, C. (2018). The 2018 Buy Box Bible. https://feedvisor.com/resources/industry-news/2018-buy-box-bible/(accessed March 2019).

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., and Sriram, S. (2006). Modeling Customer Lifetime Value. *Journal of Service Research - J SERV RES*, 9:139–155.

Gurley, B. (2012). All Markets Are Not Created Equal: 10 Factors To Consider When Evaluating Digital Marketplaces. http://abovethecrowd.com/2012/11/13/all-markets-are-not-created-equal-10-factors-to-consider-when-evaluating-digital-marketplaces/ (accessed May 2019).

Hagiu, A. (2009). Multi-Sided Platforms: From Microfoundations to Design and Expansion Strategies. SSRN Scholarly Paper ID 955584, Social Science Research Network, Rochester, NY.

Henke,N.,Bughin, J., Chui,M., Manyika, J., Saleh,T. , Wiseman, B. , Sethupathy, G. (2016). The age of Analytics competing in a data-driven world. Retrieved from McKinsey Global Insitute 25th Anniversary.

Hu, Y., Gu, Q., Wen, J., and Tang, Y. (2017). A supplier selection and order allocation method for online to offline (O2o)e-commerce markets. In *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 1359–1364, Siem Reap. IEEE.

Hubert, M., Hubert, M., Florack, A., Linzmajer, M., and Kenning, P. (2013). Neural Correlates of Impulsive Buying Tendencies during Perception of Product Packaging. *Psychology & Marketing*, 30(10):861–873.

Hänninen, M., Smedlund, A., and Mitronen, L. (2017). Digitalization in retailing: multi-sided platforms as drivers of industry transformation. *Baltic Journal of Management*, 13(2):152–168.

IBM (2013). *IBM Sterling Sellingand Fulfillment Suite V9.2.0 documentation.*

Jacobson, T., Florio, R., and Salvador, T. (2011). The Premium of Value: Pricing and Promoting Luxury to the New Consumer. Retieved from Accenture Website. https://www.accenture.com/us-en (accessed May 2019).

Jiang, L., Jiang, N., and Liu, S. (2011). Consumer Perceptions of E-Service Convenience: An Exploratory Study. *Procedia Environmental Sciences*, 11:406–410.

Kenton, W. (2019). Investopedia. Retrieved from url:https://www.investopedia.com/terms/b/backorder.asp.

Kiang, P. M. (2012). Measuring Service Quality in Online Luxury Goods Retailing. *Journal of Electronic Commerce Research*, 13(1):16.

Klarna (2017). Re-thinking returns: A strategic advantage. Retrieved from https://www.klarna.com/uk/wp-content/uploads/sites/11/2017/10/Klarna-Confidential-Returns.pdf (accessed May 2019).

Korolko, N., Woodard, D., Yan, C., and Zhu, H. (2018). Dynamic Pricing and Matching in Ride-Hailing Platforms. DOI 10.2139/ssrn.3258234.

Leather, D. (2013). *Customer-centric blueprint: building and leading the 21st century organisation*. ISBN 978-0620558341.

Linda Dauriz, Nathalie Remy, N. S. (2014). Luxury shopping in the digital age. https://www.mckinsey.com/industries/retail/our-insights/luxury-shopping-in-the-digital-age (accessed May 2019).

Marler, R. and Arora, J. (2004). Survey of Multi-Objective Optimization Methods for Engineering. *Structural and Multidisciplinary Optimization*, 26:369–395. DOI 10.1007/s00158-003-0368-6.

Moazed, A. and Johnson, N. L. (2016). *Modern Monopolies: What It Takes to Dominate the 21st Century Economy*. St. Martin's Press, New York, N.Y.

Nautiyal, A. et al. (2019). Online marketplace and merchant-centric services as a driver of the platform model. Retrieved from Roland Berger and GSMA Association in https://www.gsma.com/mobilefordevelopment/resources/online-marketplace-and-and-merchant-centric-services-as-a-driver-of-the-platform-model.

Ramadan, N., Ajami, R., Mohamed, N., and Lazarova-Molnar, S. (2015). Modeling and simulation for enterprise decision-making: Successful projects and approaches. In *2015 International Conference on Industrial Engineering and Operations Management (IEOM)*, page 3, Dubai. IEEE.

Rao, S., Goldsby, T. J., Griffis, S. E., and Iyengar, D. (2011). Electronic Logistics Service Quality (e-LSQ): Its Impact on the Customer's Purchase Satisfaction and Retention: *Impact of e-LSQ on Customer's Satisfaction. Journal of Business Logistics*, 32(2):167–179.

Rattenburg, T., Hellerstein, J., Heer, J., Kandel, S., and Carreras, C. (2017). Principles of Data Wrangling. ISBN 978-1491938928.

Razi, M. A., Wen, H. J., Perez, A. A., and Tarn, J. M. (2003). E-fulfillment: the strategy and operational requirements. *Logistics Information Management*, 16(5):350–362.

Reichheld, F. (2006). The microeconomics of customer relationships. *MITSloan Management Review*.

Royce, W. (1970). Managing the development of large software systems. page 329.

Schmidt J., Dörner K., B. A. S. T. B. K. (2015). The opportunity in online luxury fashion. Retrieved from McKinsey Website.

Trade Finance Global (2019). Incoterms Rules.

UK Government (2018). Notice 236:Returned Goods Relief, (accessed march 2019).

Vahdani, B., Zandieh, M., and Alem-Tabriz, A. (2008). Supplier Selection by Balancing and Ranking Method. *Journal of Applied Sciences*, 8(19):3467–3472.

Weber, C. A., Current, J. R., and Benton, W. (1991). Vendor selection criteria and methods. *European Journal of Operational Research*, 50(1):2–18.

Xu, P. J. (2005). *Order fulfillment in online retailing : what goes where.* Thesis, Massachusetts Institute of Technology.

Çebi, F. and Bayraktar, D. (2013). An integrated approach for supplier selection. *Logistics Information Management.*

Appendix A

Order Assignment Program

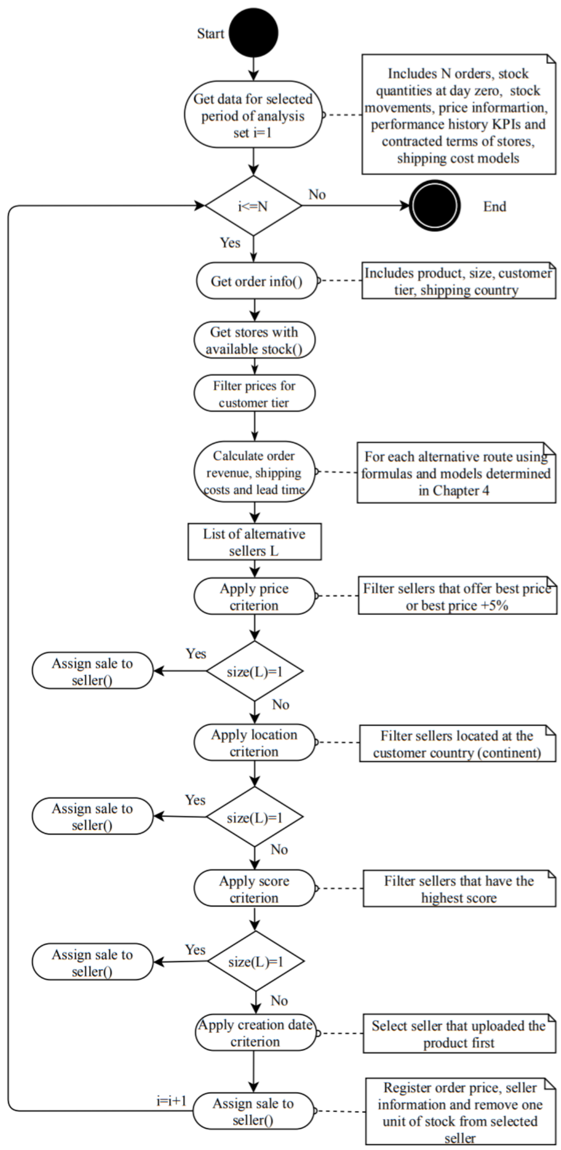


Figure A.1: Order Assignment Recreation Flowchart

The structure of the Order Assignment Program is as following. Order by order (row by row of the dataframe Orders sorted sequentially by date) an algorithm logic is applied.

```
In [ ]: %%time

        restart()
        # sort orders by order date to run the algorithm cronologically
        orders=orders.sort_values(by='orderdate_gmt')

        # for each order, apply the function algorithm
        output=orders[:].apply(algorithm, axis=1)

        1
        2
        1 day after...
        3
        4
        5
        6
        7
        8
        9
        10
        11
        12
        13
        14
        15
        16
        17
        18
```

Figure A.2: Event triggered by new order

The algorithm is a sequence of functions: get stock() at the time of the order (see Figure A.5) and apply order assignment policy, for example price criterion as illustrated in A.4. All the variables required to control the impact of the solution are registered. Note that the case of negative stock level, means that FF got a sale that in reality took place later at the physical store. This is important when testing other algorithm strategies.

```
In [229]: def update_stock(row):
              # This function returns the stock + pricing info each store has of a specific product at the moment of the order
              # Input: row - an orderline

              global stocklog_updated_aux2
              global stocklog_updated_aux

              # stock logs of the product ordered (productid + size)
              stocklog_updated_aux2 = stocklog_updated_aux.loc[(stocklog_updated_aux['productid']==row['productid']) &
                                              (stocklog_updated_aux['position']==row['position'])]

              # stock logs before the order is made
              stocklog_updated_aux2= stocklog_updated_aux2.loc[ (stocklog_updated_aux2['stock_date']<=row['orderdate_gmt'])]

              # sum the number of stock movements for each store to know how many products do we have at the moment of the order
              stocklog_updated_aux2 =stocklog_updated_aux2.groupby(['storeid','stockpointid','productid', 'position'], as_index=False)['st

              # get the price information for that product in that store
              stocklog_updated_aux2 = stocklog_updated_aux2.merge(pricing_aux2, on= ['storeid', 'stockpointid', 'productid'])

              # store in an auxiliar variable if the stock goes negative for a specific store (because of a different order allocation)
              stocklog_nega= stocklog_updated_aux2.loc[stocklog_updated_aux2['stockmovement']<0]

              # in case there is a store which has negative stock, add a stock log with the date of the order and with the number of
              # movements equal to the absolute of number of "stocks negativity" to that store
              # e.g. if the sum of stock -1 for Browns, it adds a stock log of +1 to this store
              if stocklog_nega['stockmovement'].shape[0]>0:
                  stocklog_nega['stockmovement'].abs()
                  stocklog_nega['stock_date']= row['orderdate_gmt']
                  stocklog_updated_aux=stocklog_updated_aux.append(stocklog_nega)


              return stocklog_updated_aux2 if stocklog_updated_aux2.shape[0]==0 else
                     stocklog_updated_aux2.loc[(stocklog_updated_aux2['stockmovement']>0)]
```

Figure A.3: Logic to get stock at the time of the order

```python
# If there is only one store available, it wins. This is the case when the item is not a duplicate.
# Information about the store is stored in the orderline and returned
if competitors_df.shape[0] == 1:
    row['rule_winner'] = 'Not Duplicate'
    row['rule_winner_storeid'] = competitors_df['storeid'].values[0]
    row['rule_winner_stockpoint'] = competitors_df['stockpointid'].values[0]
    row['rule_winner_name'] = competitors_df['storename'].values[0]
    row['price_winner'] = competitors_df['priceonline_scurrency'].values[0]
    row['edd']= competitors_df['avg_leadtime'].values[0]
    row['oc']=competitors_df['oc'].values[0]
    row['ffcake']= competitors_df['avg_ffcake_ltv'].values[0]
    row['avg_shippingcost']= competitors_df['shippingcost'].values[0]
    row['margin']= competitors_df['margin'].values[0]
    return row

# filters the available stores which prices higher than the minimim price + 5%
competitors_df = competitors_df.loc[(competitors_df['priceonline_scurrency'] <=
                                     competitors_df['priceonline_scurrency'].min()*1.05)]

# stores the number of available stores after price filter in the order line
row['nr of competitors_price']=competitors_df.shape[0]

# If there is only one store available, it wins. This is the case when the store wins because of items' price.
# Information about the store is stored in the orderline and returned
if competitors_df.shape[0] == 1:
    row['rule_winner'] = 'Price'
    row['rule_winner_storeid'] = competitors_df['storeid'].values[0]
    row['rule_winner_stockpoint'] = competitors_df['stockpointid'].values[0]
    row['rule_winner_name'] = competitors_df['storename'].values[0]
    row['price_winner'] = competitors_df['priceonline_scurrency'].values[0]
    row['edd']= competitors_df['avg_leadtime'].values[0]
    row['oc']=competitors_df['oc'].values[0]
    row['ffcake']= competitors_df['avg_ffcake_ltv'].values[0]
    row['avg_shippingcost']= competitors_df['shippingcost'].values[0]
    row['margin']= competitors_df['margin'].values[0]
    return row
```

Figure A.4: Unique items identification and Price Criterion Application

```python
In [19]: i=0
orders_aux=orders.copy()

# check condition in the function below and at the end filter only offline and stock uploads
stocklog_updated=stocklog.apply(update_stock_log, axis=1)
stocklog_updated=stocklog_updated[stocklog_updated['aux']==1]

#FUNCTION
def update_stock_log(row):
    global i
    #aim is to remove real occured orders from the stock movements list, as in the simulation this stock has to be available
    i += 1
    print(i)
    #if is possible online sale or offline sale and online sale exists in the orders list then remove order used to update table
    if  ((row['stockmovement']==-1)
        &  (orders_aux.loc[(row['stockpointid']==orders_aux['winner_stockpointid'])
                          & (row['productid']==orders_aux['productid']) &
                    (row['position']==orders_aux['position']) &
                    (row['stock_date']>=orders_aux['orderdate_gmt'])].shape[0]>0)):
        orders_aux.drop(orders_aux.loc[(row['stockpointid']==orders_aux['winner_stockpointid']) &
                    (row['productid']==orders_aux['productid']) &
                    (row['position']==orders_aux['position']) &
                    (row['stock_date']>=orders_aux['orderdate_gmt'])]
                           .sort_values(by=['orderdate_gmt'])
                           .index[0], inplace=True)
        row['aux']= 0
        return row
    row['aux']=1
    return row

print("{} orders successsfully updated. \n{} orders didn't have an associated log".format(stocklog.shape[0]-stocklog_updated.sha

1
2
3
```

Figure A.5: Logic to remove online sales from the table with stock movements

Appendix B

Shipping Model Details

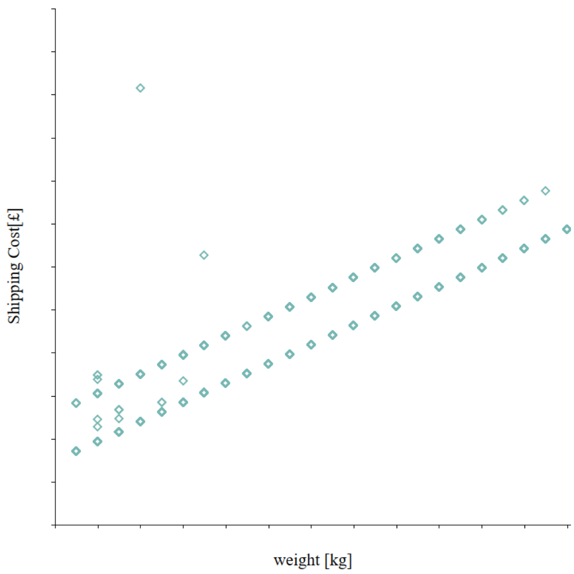


Figure B.1: Visualization of remote area impact in shipping costs

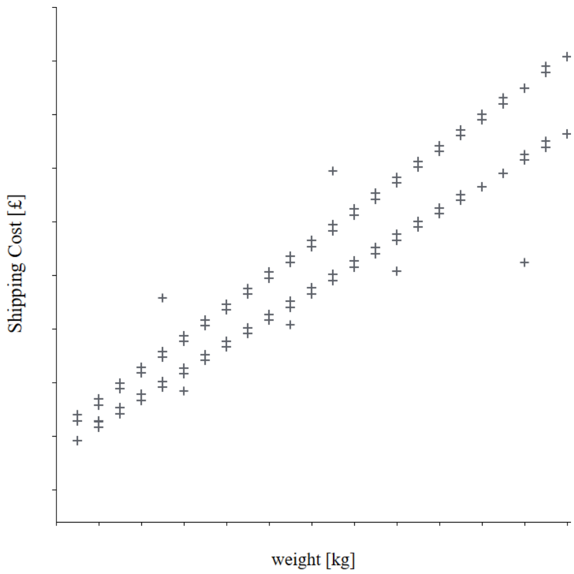


Figure B.2: Visualization of Italy - Hong Kong Renegotiated Costs in Nov 2018

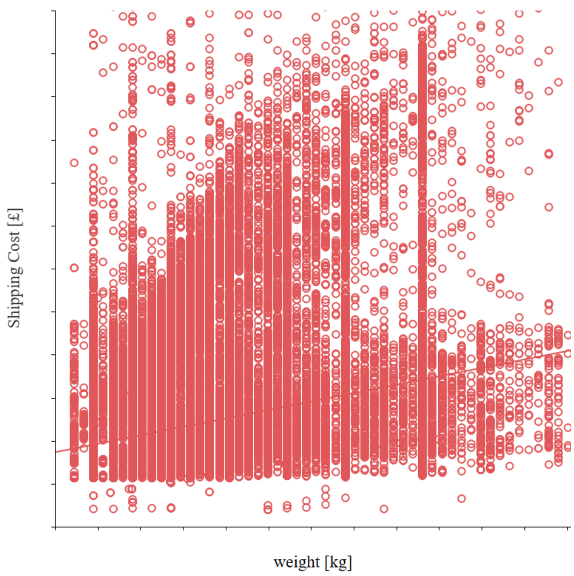


Figure B.3: Visualization of US-US standard shipping variability