

Identification and Optimisation of Key Logistic Cost Drivers in a Luxury Fashion Marketplace

André Fernandes de Carvalho

Dissertação de Mestrado

Orientador na FEUP: Prof. Bernardo Almada Lobo



Mestrado em Engenharia e Gestão Industrial

2022-07-11

All you need in this life is ignorance and confidence

Then success is sure.

-Mark Twain

Abstract

In a world that increasingly demands having the right product at the right time and place, logistics cost management is one of the key parts of a company's success. Given this importance, there is a continuous search for optimization opportunities in this area that can contribute to greater efficiency of the supply chain as a whole. Farfetch is no exception. The company has a business model extremely dependent on its supply chain, and it is of great importance for the company to control and optimize logistics costs.

The main goal of this dissertation is to create a dashboard to monitor the drivers of logistics costs, as well as to build opportunities to optimize them.

The project is divided mainly into three distinct phases. The first consisted of the identification phase of the main drivers and their combinations that affect Farfetch's logistics costs. This phase of the project was characterized by communication between the various teams of the Logistics department within Farfetch in order to outline which costs and drivers should be optimized. Furthermore, it is the most analytical phase of the project, in which two models developed in the R language were used in order to evaluate the influence of the drivers and combinations of them identified.

The next phase of the project consisted in the idealization and creation of a dashboard with the purpose of following the evolution of the identified drivers. The dashboard was developed in the Looker tool and is one of the main outputs of the project.

The final phase of the project involved the idealization and planning of two initiatives to optimize logistics costs previously defined. The first initiative, which took less project time, consists in the optimization of shipping costs in the United States for a certain combination of drivers. The second initiative, the main one of the project, aimed at optimizing the return costs associated with returned orders in which the customer considered the item to be too big or too small.

Resumo

Num mundo que exige cada vez mais ter o produto certo, no momento e local certo, a gestão de custos logísticos é uma das partes fundamentais para o sucesso de uma empresa. Dada esta importância, há uma procura contínua de oportunidades de otimização nesta área que podem contribuir para uma maior eficiência da cadeia de abastecimento como um todo. A Farfetch não é exceção. A empresa tem um modelo de negócio extremamente dependente da sua cadeia de abastecimento, sendo de grande importância para empresa o controlo e otimização dos custos logísticos.

Esta dissertação tem como principal objetivo a criação de uma *dashboard* de acompanhamento dos *drivers* dos custos de logística, bem como construção de oportunidades de otimização dos mesmos.

O projeto divide-se principalmente em três fases distintas. A primeira consistiu na fase de identificação dos principais *drivers* e suas combinações que afetam os custos logísticos da Farfetch. Esta fase do projeto foi caracterizada por uma comunicação entre as várias equipas do departamento de Logística dentro da Farfetch de modo a delinear quais os custos e *drivers* a ser otimizados. Para além disso, trata-se da fase mais analítica do projeto, na qual foram utilizados dois modelos desenvolvidos na linguagem R de modo a avaliar a influência dos *drivers* e combinações dos mesmos identificados.

A próxima fase do projeto consistiu na idealização e criação de uma *dashboard* com o intuito de acompanhar a evolução dos *drivers* identificados. A dashboard foi desenvolvida na ferramenta Looker e trata-se de um dos principais *outputs* do projeto.

A fase final do projeto prendeu-se com a idealização e planeamento de duas iniciativas de otimização dos custos logísticos previamente definidos. A primeira iniciativa, que ocupou um menor tempo de projeto, consiste na otimização dos custos de *shipping* nos Estados Unidos para uma certa combinação de *drivers*. A segunda iniciativa, a principal do projeto, teve como objetivo a otimização dos custos de *return* associados a encomendas devolvidas nas quais o cliente considerou que o item era demasiado grande ou demasiado pequeno.

Acknowledgments

After 5 academic years, a journey full of mixed emotions is coming to an end. The present dissertation is dedicated to all the ones that cross my path, and somehow supported me and helped me to achieve my goals.

To my parents, grandmothers and brother for all the love, support, and care.

To my faculty friends for sharing the best time of my life with me, since the best to the worst moments.

To my hometown friends, for all the experiences I have lived until now, and for the greatest teaching they have given me: Bee Kind.

To all the professors that, in a certain way, shaped my path at FEUP.

To my supervisor at FEUP, Prof. Bernardo Almada Lobo, for all the shared suggestions during this dissertation.

To Gonalo Moreira, for being my mentor in my first work experience and for the fundamental role he played in the development of this dissertation.

To Francisca Marinho, Joana Guerreiro and Pedro Hatazima for all advice during the developed project.

To Farfetch for giving me the opportunity to develop my dissertation in such an incredible company.

Contents

1	Introduction.....	1
1.1	Project Background and Motivation.....	1
1.2	Company Presentation.....	2
1.3	Project Objectives	3
1.4	Method Followed in The Project.....	4
1.5	Dissertation Structure.....	5
2	Theoretical Background	6
2.1	The Concept of Logistics.....	6
2.2	Logistics Strategy and Globalization	7
2.3	Porter's Value Chain	8
2.3.1	Advantage By Cost.....	8
2.3.2	Advantage By Differentiation	9
2.4	Cost and Management Accounting	9
2.4.1	Cost Concepts	10
2.4.2	Methods For Selecting Cost Drivers	11
2.5	Data Visualization	11
2.6	Final Considerations	13
3	Problem Description	15
3.1	Farfetch Business Model.....	15
3.1.1	Order Processing.....	15
3.1.2	Returns Overview	20
3.2	Financial Structure of The Company.....	21
3.3	Data Visualization Tool Used at The Company.....	23
4	Driver Identification, Analysis and Dashboard Creation.....	25
4.1	Logistic Costs Identification.....	25
4.2	Main Drivers Identification	25
4.2.1	Time Frame Identification	25
4.2.2	Service Type.....	26
4.2.3	Courier	26
4.2.4	Route	27
4.2.5	Box Size.....	30
4.3	Driver Evaluation	31
4.3.1	Linear Regression Model	32
4.3.2	Decision Tree Model.....	34
4.4	Dashboard Creation	36
5	Initiatives Analysis	38
5.1	Stadium Goods Shipping Cost Optimization	38
5.2	Items To Big & To Small Optimization.....	38
5.2.1	Initiative Planning.....	39
5.2.2	Definition of Thresholds	39
5.2.3	Data Collection and Treatment	41
5.2.4	Monetary Value of The Initiative	43
5.2.5	Scenario Analysis	44
5.2.6	Implementation	46
6	Discusssion and Future Work	47
APPENDIX A:	Gantt Chart.....	52
APPENDIX B:	Route Driver Analysis Country Aggregation Map	53
APPENDIX C:	Looker Dashboard	55

Acronyms and Symbols

PO	Portal Order
BO	Boutique Order
NS	No Stock
AWB	Air Waybill
RTO	Return to Origin
DR	Direct Return
IR	Indirect Return
GTV	Gross Transaction Value
ATV	Actual Merchandising Value
GMV	Gross Merchandise Value
OC	Order Contribution
COS	Cost of Sale
COGS	Cost of Goods Sold
SC	Shipping Cost
RC	Return Cost
CP	Complexity Parameter
RMSE	Root Mean Square Error
LM	Last Month
LY	Last Year
CY	Current Year
MoM	Month on Month
YoY	Year on Year
YTD	Year to Date
TB	Too Big
TS	Too Small

List of Figures

Figure 1 - Porter's value chain (Porter, 1985).....	8
Figure 2 - Farfetch Order Processing Sequence	15
Figure 3 - Diagram with the different types of Farfetch Orders.....	16
Figure 4 - Diagram showing the stages of Step 1 of the Order Processing	17
Figure 5 - Diagram showing the stages of Step 2 of the Order Processing	17
Figure 6 - Diagram showing the stages of Step 3 of the Order Processing	18
Figure 7 - Diagram showing the stages of Step 4 of the Order Processing	18
Figure 8 - Diagram showing the stages of Step 5 of the Order Processing	19
Figure 9 - Diagram showing the stages of Step 6 of the Order Processing	19
Figure 10 - Pie Chart demonstrating the distribution of the Return Reasons for 2021	20
Figure 11 - Simplified diagram of Farfetch's Profits and Losses Report.....	21
Figure 12 - Example of a Decision Tree End Node	34
Figure A.1 - Project Gantt Chart.....	49
Figure B.1 – Store Country Aggregation Map.....	50
Figure B.2 – Shipping Country Aggregation Map.....	51
Figure C.1 - Complete Looker Dashboard.....	52
Figure C.2 – Section One of Looker Dashboard.....	53
Figure C.3 – Section Two of Looker Dashboard.....	54
Figure C.4 – Section Three of Looker Dashboard.....	55

List of Tables

Table 1 - Task Diagram	4
Table 2 - Shipping Cost Distribution by Service Type	26
Table 3 - Return Cost Distribution by Service Type	27
Table 4 - Shipping Cost Distribution by Store Country	27
Table 5 - Return Cost Distribution by Shipping Country.....	28
Table 6 - Shipping Cost Distribution by Shipping Country	29
Table 7 - Return Cost Distribution by Shipping Country.....	29
Table 8 - Variable Route Creation when analysing the Shipping Cost.....	30
Table 9 - Variable Route Creation when analysing the Return Cost.....	30
Table 10 - Shipping Cost Distribution by Product Category	30
Table 11 - Return Cost Distribution by Product Category	31
Table 12 - Simplified data collection Table	31
Table 13 - Most important Shipping Cost Drivers based on the VarImp function	33
Table 14 - Most important Return Cost Drivers based on the VarImp function	33
Table 15 - Most important combination of drivers in the Shipping Cost.....	35
Table 16 - Most important combination of drivers in the Return Cost	35
Table 17 - Distribution of Items TS & TS by Category 1st	42
Table 18 - Scenario analysis for the TB & TS Initiative for the year 2022.....	45
Table 19 - Scenario analysis for the TB & TS Initiative for the year 2023.....	45

1 Introduction

1.1 Project Background and Motivation

In today's competitive market, the products are to be looked for and managed in the most effective manner to minimise costs while satisfying service level requirements (Choudhary & Shankar, 2013). This way, Logistics is a relevant and essential business activity used for enterprises throughout the world; it aids in determining if products arrive at the customer on time and containing the required quality requirements and at the lowest possible cost (Campos-Garcia, Garcia-Vidales, & Gonzalez-Gomez, 2011).

The logistics conception as a strategic factor to reach competitive advantage is a relatively new phenomenon, in other words, logistics related issues are now incorporated to the company strategies, regarding viable means to overcome the competitive market and to exceed the expectations of all interested parties, even clients (Qureshi, Kumar, & Kumar, 2008; Havenga, 2010).

Therefore, logistics consists of planning, performance and control of an efficient flow of raw materials for production, finished products and relevant information from the point of origin to the point of consumption for the purpose of satisfying the customer's requirements (Kubon & Krasnodębski, 2010).

In addition, the logistics activities are inter-dependent, requiring accurate allocation of resources to achieve service goals and to reduce loss in the supply chain such as idle time, and efforts duplication. Logistics deal with the processes of planning, implementing, and controlling the efficient flow and raw materials storage, in-process inventory, final products, services, and related information from origin points to consumption points while conforming to customer requirements (Lai, Wong, & Cheng, 2008).

Within this context, are included logistics cost, which cover significant proportion of management costs. Depending on the company's sector, such costs can reach proportion at least 10% of total cost. Besides, it is very important to the strengthening of a supply chain, because it represents a factor to be considered within the performance (Song & Wang, 2009; Engblom *et al*, 2012).

The knowledge about logistic costs is used to display how much they represent to the company total cost, and the measures to be taken so this cost does not become too high. It is also important to emphasise that many factors can increase logistics costs, so it is necessary to identify which are the most sensitive elements on this impact, to use control measures, which can be provided by efficient logistics cost management (Botes, Jacobs, & Pienaar, 2006; Fang & Ng, 2011).

Logistics cost management is a complex and extended concept, which covers many logistics process activities such as transportation, stocking, distribution, handling, reprocessing and many others. It acts identifying operations' inefficiency, which result in a huge information flow, causing companies some difficulties in measuring logistics costs in a precisely and efficient way (Zhao & Huang, 2009; Zou, 2011).

Logistics costs and its Management have attracted companies' attention, namely Farfetch, and appeared as a matter of interest in academic environment (Song & Wang, 2009). In this sense, it is justified the accomplishment of theoretical works, that also result of the need to form a reference frame which can base in this important topic. Therefore, since Farfetch is increasingly more concerned with controlling its costs, this dissertation has as its main objective the creation of value for the company through the optimization of its logistics costs.

1.2 Company Presentation

Farfetch is an e-commerce platform for luxury fashion that emerged from the visionary spirit of a Portuguese entrepreneur, José Neves, who had two companies in different sectors: one in programming and the other in footwear. He has been involved in the fashion world from an early age, since the 1990s, when he launched the SWEAR shoe brand. Later, he created B Store, a fashion company with a physical store selling young brands and emerging designers. It was during B Store's promotion at the Paris Fashion Week in October 2007 that Farfetch started (Maisey, 2019).

With several start-ups already on his resume, José Neves realized the booming of the online commerce. It was then that the opportunity arose to create a marketplace, a junction of online stores, which would allow him to have a business on a global scale while helping the stores he worked with to sell more. The idea was to allow small independent stores to compete in the market while keeping their physical stores but creating their own identity (Farfetch, 2021).

One year later, in 2008, Farfetch was born. The concept was apparently simple: it was about creating an online bridge between consumers and stores regardless of their location. For example, a customer in Lisbon could buy a luxury item in Hong Kong and receive his order at home. Farfetch would handle the whole process (Farfetch, 2021).

The concept has remained the same. The company creates value through commissions it charges its partners and through logistical revenue paid directly by the customer. A more detailed overview of the Farfetch's business model, as well as the order processing and the company's financial structure will be presented in chapter 3, Problem Description, of this dissertation.

In the period from 2010 to 2017 there were a total of eight investment rounds in Farfetch. In July 2010, a \$4.5M investment was made by Advent Venture Partners. About two years later, in January 2012, it was the turn of three other investors led by Index Ventures to invest about \$18M. Less than 1 year later, in March 2013, the number of investors was growing, as was the amount of investment made. That year, there were four investors, including Condé Nast, for a total amount of \$20M. The following year, in May, Vitruvian Partners along with 7 other investors also decided to bet on Farfetch, investing \$66M (Farfetch, 2021).

The company continued to grow, the investments kept coming and 2015 was no exception, with a total investment of \$86M made by 8 investors led by DST Global. 2016 was also a year of much investment in this platform, with Eurazeo and IDG Capital Partner together with 6 more investors to apply a funding of \$110M. The year of 2017 would break all records by recording a \$397M investment coming from a single investor, JD.com (Farfetch, 2021).

On the 21st of September 2018, the company went public on the New York Stock Exchange with over 44 million shares at \$20 each (€23). Less than ten minutes after trading began, the value of each share was already hovering around \$28.5. The first Portuguese tech company to join the New York Stock Exchange (NYSE) is now recognized on the exchange under the abbreviation "FTCH" (Farfetch, 2021).

In January 2019, Farfetch completed the acquisition of 100% of the outstanding equity interests of Stadium Enterprises LLC, which does business as "Stadium Goods," a premium sneaker and streetwear marketplace, for the total consideration of \$230.9 million (Farfetch, 2021)

Continuing its growth, in March 2021, Farfetch launched its luxury shopping channel on Alibaba's platform, Tmall Luxury Pavilion. This new channel has the potential to expand the reach of Farfetch global luxury platform to approximately 779 million active consumers on Alibaba's China retail marketplaces, offering luxury brands a multi-brand solution in China through a single integration with Farfetch (Farfetch, 2022).

The online retail platform has offices in London, Porto, Braga, Guimarães, Lisbon, New York, Los Angeles, São Paulo, Tokyo, Hong Kong, Shanghai and Moscow. All pieces available on the platform are photographed and catalogued in another company building located either in Guimarães, Hong Honk or Porto (Farfetch, 2022).

The company follows values that it considers essential to its innovative spirit: Be Brilliant, Be Human, All Together, Think Global, Be Revolutionary and Amaze Customers. (Farfetch, 2022).

1.3 Project Objectives

Characterised by its business model and philosophy, Farfetch is constantly looking for ways to invest and optimise its factors. Inherent in the business model itself, a large portion of the company's costs are associated with logistics costs.

Therefore, the company is very interested not only in understanding the main factors that influence each cost associated with logistics, but also in ordering them according to their importance. Furthermore, it is also of great interest to Farfetch to be able to understand which combinations of these drivers influence costs the most, in order to be able to analyse and optimise them.

Looking ahead, it is in Farfetch's best interest to be able to follow the variation of the main drivers of each cost line, both on a short term and long-term level. Observing these variations in an easy and intuitive way will help Farfetch to detect problems at the logistics level, as well as let it know which drivers it should focus on in order to reduce costs.

In this sense, the main objectives of the project are the following:

- Literature review for understanding and evaluation of relevant issues surrounding the project: analysis of studies conducted on logistics, cost and management accounting and data visualization.
- Identification, analysis and understanding the main drivers affecting each logistics cost, as well as identifying the most impacting combinations of drivers;
- Development of a tool to monitor the evolution of the drivers identified in the form of a dashboard;
- Identification and analysis of potential optimisation opportunity and creation of a business cases;

Therefore, this is a project with several stages and its main objective is to reduce the logistics costs of the company.

1.4 Method Followed in The Project

Considering the objectives exposed and outlined in the chapter 1.3, the methodology implemented in the project was rigorously planned through a set of tasks and respective forecast for their duration, considering the total time available, as shown in Table 1. Whenever a task was started and finished, the respective dates were recorded, as well as whether or not the deadline for its execution was met. The evolution of the tasks necessary for the realization of the project can also be visualized graphically through a Gantt chart, shown in Figure A.1 of Appendix A.

Table 1 - Task Diagram

Task	Expected duration	Start date	Termination date	Accomplished in the expected duration?
Introduction to Farfetch	7 days	28 of February	6 of March	Yes
Introduction to Analytics	12 days	7 of March	18 of March	Yes
Driver Identification	21 days	19 of March	8 of April	Yes
Statistical Model	21 days	28 of March	18 of April	Yes
Looker Dashboard Creation	19 days	18 of April	6 of May	Yes
Exploring Main Drivers	19 days	2 of May	20 of May	Yes
Simulation Phase	14 days	21 of May	3 of June	Yes
Business Case Creation	21 days	4 of June	24 of June	Yes
Dissertation Adjustments	12 days	20 of June	1 of July	Yes

The initial task, the realization of an induction program (Farfetch Plug-In Program), occurred in the first week and consisted in obtaining knowledge about the company's business model. The week was characterized by a "Meet the Business" session which served to increase knowledge about the company both at the operational level and the company's philosophy. In addition, there were also daily "Digital Learning" sessions in which technical and company knowledge was deepened on a stand-alone basis. Finally, there were also "Build the Business" sessions in which the student interns were separated into groups and worked on a real case study about the company's operational model.

The second task was characterised by a deepening of the technical component in both Looker and SQL using Big Query. This stage was also distinguished by a deeper knowledge of the various departments of the company by talking and meeting with several people from different departments, understanding how they interact and what their roles are, therefore facilitating the integration in the company.

Once the first two tasks were accomplished, the focus shifted completely to the project, starting with the identification of the main logistics costs and then its main drivers. After the identification of the drivers, the project goes through the creation of a statistical model of their classification and the definition of combinations of high importance in R language.

The next phase is the creation of a dashboard in Looker, which relies on their own programming language, LookML, to enable the creation of SQL queries against a specific database, with the intent of visualizing variations in the drivers identified and assist in finding opportunities for improvement over time.

The next step is to explore the drivers and identify opportunities for improvement with the help of the dashboard created and the combinations identified. After the identification of improvement opportunities, the next step is the analytical analysis of those opportunities and, finally, the creation of a Business Case, that can be put into practice by the company and will result in the reduction of Farfetch's logistics costs.

The last task, the writing of the dissertation, occurred simultaneously with the above-mentioned tasks. Nevertheless, in this last phase, adjustments and revisions have been put in place.

1.5 Dissertation Structure

In addition to the introduction, this dissertation is organized into six distinct chapters.

In chapter 2 the state of the art is described, starting with the presentation of some useful and important concepts in the scope of Logistics, Cost and Management Accounting and Data Visualization with emphasis on the development of dashboards.

An overview of Farfetch's business model is presented in chapter 3, focusing on the crucial entities involved in the logistics process. The cost structure is also analysed, focusing on the costs associated with logistics. In addition, the data-visualization tool used at Farfetch, Looker, is presented.

In chapter 4, the solution found is described in terms of the importance of the identified drivers and the high value combinations for the company. This chapter will also present the dashboard that accompanies the drivers in Looker.

Chapter 5 presents two optimisation opportunities found and worked on during the project.

Chapter 6 makes some considerations and highlights about the project, discusses the results obtained and suggests some advances and future work according to what emerged throughout the project.

2 Theoretical Background

This chapter presents a literature review of the topics covered in the project. First, it will discuss what logistics is and what it is used for as a solution to support business initiatives and understand the potential benefits of its applicability that a company can develop by implementing it in its strategy. The next section provides an overview of cost accounting, including what it is, how it is used and how to identify costing drivers. Finally, in the final section of the state of the art, the topic of dashboards will be addressed, presenting the various types of dashboards, as well as their utilities.

For Baptista and Sousa (2011) a literature search involves the collection of information and is influenced by information written by others about the topic thus substantiating the state of the art.

2.1 The Concept of Logistics

The Council of Logistics Management (2016), defines logistics "as the part of the supply chain responsible for planning, implementing and efficiently controlling the flow of stock of goods and/or services and information", from the origin to the point of consumption.

For Carvalho (2012) the "logistics activities as including inbound and outbound management in terms of transportation (inbound transport and outbound transport), fleet management, warehousing management, materials management and their handling, order response management, logistics network design, inventory management, supply and demand planning, and management of logistics service providers.". He concludes by stating that "physical and information flows are managed to provide the products, services, or solutions offered to customers with logistical attributes that can be considered as value drivers."

For Mangan *et al* (2008) "logistics involves obtaining the right product, in the right way, in the right quantity and quality, at the right place and time to the right customer at the right. According to Langley *et al* (2008) there are some notions of logistics:

- In a **logic of inventory and inventory management**, logistics deals with the issues of materials management, whether final goods, semi-finished products or raw materials from entering the factory to its output whether they are in motion or stopped (stock). In other words, it is responsible for managing the physical flow of products (Langley *et al*, 2008);
- In a **customer logic**, logistics aims to get the right product, to the right customer, in the right quantity and condition, at the right time and place and especially at the right cost. This logic leads us to try to combine the best service with the lowest cost. Since the best service (finding the right customer, the right product, in the right quantity and condition, and at the right place and time) tends to carry a higher cost. To manage this issue efficiently requires management based on trade-offs and trade-offs to achieve the best fit between service level and cost (Langley *et al*, 2008);
- In a **logic of utility and value**, the logistics of products and services aims to make logistical characteristics available in the market in such a way as to convince customers to be willing to pay for them (value) in exchange for products or services that are also valued logistically (Langley *et al*, 2008);

- In a **functional logic or functional management**, logistics is presented as a set of activities ranging from determining the requirements of the materials (components) that the company will need, products or raw materials being manufactured or finished products, storage of these products, handling, packaging, locations of facilities or points of distribution and sale, reverse logistics, information management of the order cycle and customer support (Langley *et al* , 2008);.

For Beth *et al* (2006), a very important concept linked to logistics is Supply Chain Management, which involves the planning and management of sourcing and procurement activities and all logistics activities. Correct Supply Chain Management ensures that all partners involved in the channel, both upstream and downstream, such as intermediaries, service providers, and customers, have a productive relationship. According to the same author, "The Supply Chain has become a strategic and competitive variable in almost every business sector. It has an impact on all aspects of shareholder value, including costs, customer service, return on assets, and income ".

2.2 Logistics Strategy and Globalization

According to Calmanovici (2011) in an increasingly globalized world, the ability to innovate is crucial for the competitiveness of companies and countries and its development the "only sustainable way forward".

As mentioned by Cosimato and Troisi (2015), "globalization has allowed adjustments in strategy for world companies to balance their performance and achieve sustainable development. Under these conditions, environmental issues, have been brought into question, so that logistics has to adopt greener programs using green technologies in order to gain or maintain a concrete competitive advantage."

For Jingping, (2017) "economic globalization has contributed significantly to world growth. In fact, it is an irreversible historical trend". Although he reminded that trade needs to be "more open, more inclusive, more balanced, more equitable and with more benefits for all". According to Lacombe (2009) "globalization implies an increasing integration of all markets (financial, products, services, labour, etc.), as well as the means of communication and transportation of all countries. Globalization tends to require greater cultural and professional preparation from all players."

Globalization is one of the great benefits for the markets, through the opening and interaction of local economies in large international blocks. This was only possible with the fall of barriers to international trade, namely with the creation of the European Union, which promoted competition and a higher quality of the products produced. For this to happen, it is necessary that the strategy be local, but thought out and defined on a global scale, positively affecting logistics, because logistical issues such as transport, storage, and provisions reach a transactional scale (Lacombe, 2009). As mentioned by Carvalho, (2010) "...globalization brings numerous advantages (stronger growth, more favourable prices for consumers, increase in the number and size of international transactions concerning goods and services, technologies and financial flows) and benefits (increased challenges and opportunities, favouring productivity and efficiency, causing a better allocation of resources, inducing progress and technological innovation...").

2.3 Porter's Value Chain

A company's primary goal is to create value for its stakeholders. "In competitive terms, value is the amount that customers are willing to pay for what a company has to offer them," writes Porter (1989). Representing total revenue in this way, that is, the result of product sales at the price that customers were willing to pay. The term "profitability" is used to describe the difference between the product's sales value and its total cost.

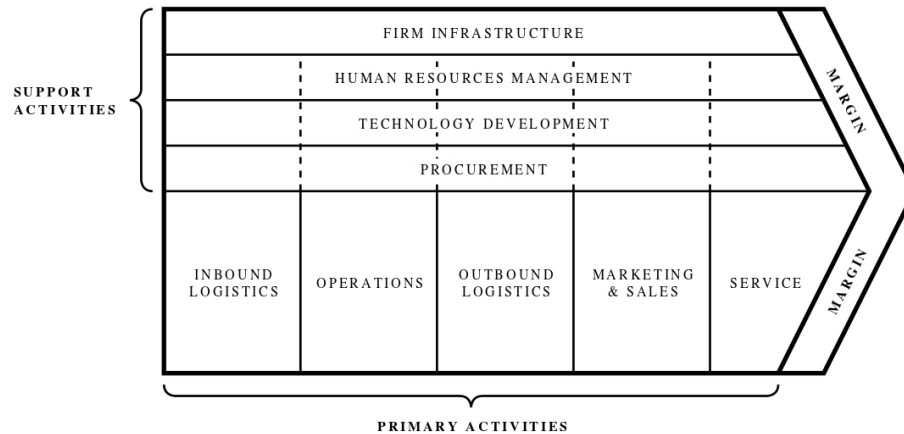


Figure 1 - Porter's value chain in Porter (1985)

According to Porter (1985), the concept of a firm's value chain can be extended both upstream and downstream. Customers, distributors, and suppliers all have their own value chains. It presupposes the creation and delivery of products to the manufacturing company, which can influence the latter's performance; it presupposes the use of the distributor's value chain from these channels until the products are forwarded to customers in the case of suppliers.

As a result, Porter's value chain emerges as a tool for disaggregating a company's most strategically important activities in order to understand the costs and potential sources of differentiation. Porter (1985) claims that "A competitive advantage is gained by a company that performs these activities efficiently and/or better than its competitors. This is critical because, contrary to what some companies believe, the competitive position is gained through the value that customers place on the product, not by the product's higher production costs".

In Bertaglia's (2009) perspective "for a company to gain a competitive advantage in a volatile market, companies need to be able to maintain a balance between demand and supply. In this context, good planning of sales and production operations allows companies to manage their supply chain more efficiently, bringing overall benefits to the company".

2.3.1 Advantage By Cost

The business reality that began to be experienced as of the 1980s, where competition intensified, undoubtedly forced changes in the production process, management and throughout the organisation, so that companies would be more competitive in relation to market competitors. A company can have a cost advantage when the total cost of all activities is lower than the total cost of all the activities of its competitors. Thus, the cost advantage tends to last over time when it presents barriers to entry, namely in terms of the difficulty of imitating or copying the company's cost sources, taking into account the maintenance of superior performance. Thus, it is particularly important to bear in mind that this cost advantage is not eliminated by the need to sell at a lower price than the competition (Porter, 1985).

According to Porter (1985) there are two important ways to obtain a cost advantage:

- The first one is by **controlling the cost drivers**, that is, achieving an advantage with respect to the efficient control of the main costs of value activities of the firm. Knowing which money is affected and for which (main) activities and making a significant proportion of their total costs;
- The second one is by **reorganise the value chain**. The whole process of production, distribution and marketing can be carried out differently and more efficiently by reducing waste and saving costs.

These two ways of reaching a cost advantage can be combined with each other and in a way that further enhances the possibility of gaining a cost advantage by multiple combination, in the longest period of time possible. In other words, obtaining a cost advantage requires a rigorous analysis of each of the company's activities in the quest for maximum efficiency in terms of the cost associated with the activity, without quality or differentiation capacity being negatively affected (Porter, 1985).

According to Carvalho (2011) "a company gains a competitive advantage when it performs activities of strategic relevance in such a way that costs are lower, or even, when the quality of services provided is superior to those offered by the competition".

2.3.2 Advantage By Differentiation

The technological advance and the emergence of new markets producing products that are increasingly sophisticated and balanced in terms of quality, innovation and design require companies to define increasingly well-defined strategies. Namely differentiation strategies.

According to Porter (1985) "A firm differentiates itself from its competitors if it can be or offer something unique that is valuable to its customers".

For Porter (1985) the main drivers of differentiation are:

- policy choices (choice of technology used or qualifications of people included in the manufacturing process);
- linkages (relationships within the value chain, for example between suppliers and distributors);
- location (location of support infrastructure);
- timeliness (in terms of time to start/perform an activity);
- inter-relationships (an activity may be shared with other business units) ;
- scale (an activity carried out on a large scale may allow it to be carried out in a singular and isolated way).

2.4 Cost and Management Accounting

This section provides an overview of cost accounting, what it is and what it is used for. The most relevant concepts required to understand the proposal will be emphasized throughout this section.

The process of collecting, classifying, assigning, and analysing the costs associated with an organization's activity is known as cost accounting (or costing) (Blocher, 2005).

Cost accounting, according to Vanderbeck (2012), provides management with the detailed cost information needed to control current operations and plan for the future. The goal of cost accounting is to collect all available data and organize it so that management can make decisions and assess the organization's performance.

Management Accounting, on the other hand, is an internal business development function led by accounting and finance professionals. Management accounting's main goals include assisting in the design and evaluation of business processes, budgeting and forecasting, implementing and monitoring internal controls, and aggregating data to aid in the organization's economic value creation (Vanderbeck, 2012).

Cost Accounting Information Systems are information systems designed to assist with cost accounting. Recording transaction data and calculating the cost of an organization's outputs are two of the main goals of these systems (Vanderbeck, 2012).

2.4.1 Cost Concepts

Cost insight is required for financial, operational, and strategic management, which have vastly different demands. The information also needs to be viewed by various dimensions relating to the resources acquired, activities completed, and associated cost objects (Blocher, 2005). The following concepts are essential in order to understand the cost accounting process:

- **Cost** – When an organization uses a resource for a specific purpose, it incurs a cost. For example, every time a raw material is used in production or manufacturing labour time is consumed, an organization that produces manufactured products incurs a cost. Several factors can be used to categorize costs: Traceability (direct if economically traceable or indirect if not); Nature (labour, material, or expenses); Behaviour (fixed if total cost does not vary with volume (e.g., rental) or variable if volume affects total cost (e.g. material) (Blocher, 2005, Vanderbeck, 2012).
- **Cost Pools** – Meaningful groups in which costs are often assigned (Blocher, 2005).
- **Cost Driver** – Any factor that has the effect of altering the total cost amount. A cost driver, for example, is any factor that influences the cost of an activity. The management of key cost drivers is critical for an organization that competes on the basis of cost leadership. Allocation bases are another name for cost drivers (Blocher, 2005).
- **Cost Object** – Costs are assigned to any product, service, customer, activity, or organizational unit for management purposes. Products, services, and customers are all cost objects, whereas manufacturing departments are either cost pools or cost objects, depending on whether management is more concerned with product costs or manufacturing costs. Decision-making, performance measurement, and strategy implementation all require cost objects (Blocher, 2005).
- **Cost Assignment** – The process of allocating resource costs to cost pools and/or cost objects. Direct tracing and allocation are the two types of assignment. Direct costs are assigned to the first, while indirect costs are assigned to the second. In terms of the cost pool or cost object under consideration, costs are either direct or indirect (Blocher, 2005).
- **Direct Cost** – A cost that can be traced directly to a cost pool or a cost object in a convenient and cost-effective manner. For example, the cost of raw materials used to make a specific product is a direct cost because it can be traced back to the product (Blocher, 2005).
- **Indirect Cost** – A cost that cannot be traced to a cost pool or from the cost pool to the cost object in a convenient or cost-effective manner. Costs associated with management personnel or material handling are examples of indirect costs that cannot be traced back to specific products. Because these costs cannot be traced back to a cost pool or cost object, cost drivers are used to assign these costs. Cost allocation is the term for this procedure (Blocher, 2005).

- **Product Costs** – Only the costs required to finish the product are included: direct materials, direct labour, and product overheads (Blocher, 2005).
- **Period Costs** – General costs, selling costs, and administrative costs are all non-product expenses for managing the firm and selling the products that are not involved in the manufacturing process (Blocher, 2005).

2.4.2 Methods For Selecting Cost Drivers

According to Lee (2002), aiming to compare costing practices at international level have brought the results that there are no uniform methods for selecting cost drivers. The methodology varies mainly with the industries and their branches, but some differences between implementations in different countries can also be observed (Lee, 2002).

The applicable methods can be grouped into two main clusters:

- methods based on statistics theory - This cluster has as its main method of analysis the use of statistical tools such as the correlation and regression analysis;
- methods based on operations research theory - This cluster is more subjective, it relies on professional experience and has as its main method of analysis the use of the analytic hierarchy process – AHP.

In the first case sound data bases of detailed cost and performance time series shall be made available. The longer time series can be used – in general – the more accurate estimations of cost-performance relations can be derived. Here all activity costs in the reference time intervals are examined by a variety of regression functions: time series of activity costs and performance indicators are coupled for the examined time period. The calculations are carried out for each performance indicator separately. In case of each activity the best correlation (R^2) is to be found. The relationships between costs and performances can then be described through the proper regression functions (Lee, 2002).

When too many activities shall be examined it is worth grouping them in order to simplify the selection process by finding general performance indicators (as joint cost drivers) for the groups. Naturally grouping will probably result in worse correlation values. According to empirical calculations less than 10% of the reliability of final evaluation statements is lost because of such kind of simplifications (Hungarian Association of Logistics, 2009).

Selection of appropriate cost drivers from the larger set of available candidate drivers is based mainly on human judgment supported by analyses using simple accounting techniques or more sophisticated correlation techniques from statistics. These methods are limited as they are not able to evaluate multiple decision-making criteria. AHP may be a solution to these problems. It is typically used in decision situations which involve selecting one or more decision alternatives from several candidate decision alternatives based on multiple decision criteria. Research have shown that AHP helps achieve consistency in selection problems where the decision criteria are subjective measures based on professional experience (Schniederjans, 1997).

2.5 Data Visualization

Business and financial organizations have relied on information technology to analyse data for operational and strategic insights that can provide a competitive advantage for the past thirty years. Performance management has become commonplace in business and financial organizations as these technologies become more capable of processing large amounts of more complex data and producing more sophisticated results (Pauwels *et al*, 2009).

Performance management is a general term used to describe the integration of strategic plans, goals and objectives, financial reporting, budgeting processes, and business forecasting into a centralized reporting system that produces metrics of operational status and health of the organization. Dashboards are only the most recent approach in the brief history of performance management (Eckerson, 2006).

Dashboard can be defined as an information delivery system that distributes information, insights and alerts to users on demand. Few combined both form and function, defining dashboards as a visual display of the most important information needed to achieve one or more objectives; consolidated and organized on a single screen so that information can be monitored at once. It was also noted that dashboards should be intuitive to end users and should emphasize visual simplicity rather than excessive colours and graphics (Few, 2006).

According to Eckerson (2006), companies and financial organizations have relied on information technology to analyse data in search of operational and strategic insights that can lead to competitive advantages in recent decades. Performance management has become more common in corporations as these technologies have evolved and become capable of processing large amounts of more complex data and producing more sophisticated outputs.

Dashboards have emerged as a functional and visual way to extend performance management. Whereas the performance indicator allows business leaders to view summarized status and progress for specific purposes on a regular basis, the introduction of dashboards allowed the entire enterprise to view this data in real or near real time and draw the necessary conclusions for their activities (Niven, 2006)

Few (2006) defined a dashboard as a visual presentation of the most relevant information that is needed to obtain one or more objectives in a consolidated and ordered manner on a single screen so that the information can be monitored at a glance. Meanwhile, Yigitbasioglu and Velcu-Laitinen (2012) argued that dashboards are tools developed to provide a quick version of the organization's performance to busy managers.

Dashboards have recently evolved from the primary purpose of monitoring corporate performance to more advanced analytical purposes, according to Yigitbasioglu and Velcu-Laitinen (2012), incorporating new features such as scenario analysis, drill-down capabilities, and presentation format flexibility. Pauwels et al. (2009) propose the following list of dashboard uses:

- Assisting executives to access information at any time;
- Helping management to define tasks and important information;
- Setting goals and expectations for specific individuals or groups;
- Assisting employees in comprehending the most critical issues;
- Encouraging specific actions at the right time;
- Highlight exceptions and provide alerts in case of problems;
- Communicate progress and success;
- Provide a common interface for interaction and analysis of important data.

To build a dashboard that can serve the uses presented in the list above, Kerzner (2017) demonstrate that a dashboard can have four possible purposes:

- **Monitoring:** A dashboard aids in performance monitoring and refers to the daily evaluation of metrics that should lead to corrective action. Monitoring may be the most fundamental function of the dashboard.

- **Consistency:** A dashboard ensures that measures and measurement procedures are consistent across departments and business units.
- **Planning:** A dashboard can be used for planning if its functionality includes scenario analysis.
- **Communication:** A dashboard communicates both the performance and the values of an organization to its stakeholders.

There are three types of dashboards, according to different authors: operational, tactical, and strategic dashboards. The definitions presented below are followed by a variety of researchers, despite some nomenclature differences (Eckerson, 2010; Few, 2013; Kerzner, 2017; Yigitbasioglu e Velcu-Laitinen, 2012).

Operational Dashboards

Monitor the main operational processes, which are used primarily by employees and their supervisors who deal directly with clients or oversee the creation and delivery of organizational products and services. This type of dashboard provides detailed information that is only summarized. Furthermore, the majority of metrics in an operational dashboard are updated on an ongoing basis, ranging from minutes to hours, depending on the application. As a result, operational dashboards emphasize monitoring over analysis and management.

Tactical dashboards

Departmental processes and projects that are of interest to a specific segment of the organization or a small group of people are implemented. Dashboards are used by business executives and analysts to compare the performance of their departments or projects to budget plans, forecasts, or previous period results.

Strategic dashboards

Monitor the execution of strategic objectives, which are frequently implemented using a Balanced Scorecard approach, though other methods may also be used. The goal of a strategic dashboard is to align the organization around strategic goals and ensure that all groups are moving in the same direction. To accomplish this, organizations distribute personalized performance indicators to all organizational groups and, on occasion, to each individual. These performance indicators, which are typically used on a weekly or monthly basis, provide powerful tools to executives for communicating strategy, increasing visibility of operations, and identifying key drivers of performance and commercial value. Strategic dashboards emphasize management over monitoring and analysis. This will be the type of dashboard created in the developed project.

2.6 Final Considerations

The aim of any literature review is to summarize and synthesize the arguments and ideas of existing knowledge in a particular field without adding any new contributions. Being built on existing knowledge they help the researcher to even turn the wheels of the topic of research. It is possible only with profound knowledge of what is wrong in the existing findings in detail to overpower them. (Pubrica, 2019)

The literature review conducted proved to be fundamental to the realization of the project. With regard to the first three sub-chapters (“The Concept of Logistics”, “Logistics Strategy and Globalization” and “Porter’s Value Chain”), the literature review allowed for the understanding of fundamental concepts in the logistics field. In addition, it proved to be fundamental in shaping the design thinking and approach of the project.

The literature review of “Cost and Management Accounting” was also important since it provided cost definitions that proved to be fundamental, such as the definition of cost driver and cost allocation.

Finally, the literature review of the last sub-chapter, “Data Visualization”, proved to be fundamental in creating the dashboard to monitor the drivers identified. The literature review of this subject allowed not only to define the purpose of the dashboard, but also the best way to structure it.

3 Problem Description

3.1 Farfetch Business Model

There are two main aspects to consider in the setting of the company where this project is taking place. First, for the majority of the items sold by boutiques, Farfetch does not hold stock of the products, this being the responsibility of the partner. However, on some occasions, Farfetch is responsible for holding the stock in the company's own warehouses. Additionally, it caters to extremely particular clients with refined habits in the luxury market it operates in.

In its warehouses, the company keeps the stock not only of its stores, such as Browns and Stadiums Goods, but also of some partners that Farfetch is responsible for managing its logistics processes. Farfetch currently has warehouses in China, Italy, United States, United Kingdom and Netherlands.

A product only ever goes through Farfetch Operations before a season starts. A sample of the item is sent to one of the company's offices called the production centre so that it can be photographed and have its details added to the database and, subsequently, to the website. Farfetch acts as the link between the sellers, either boutiques or brands, and the customers, handling all the logistics processes that result in the transportation of the product from where the item is in stock to the final destination.

3.1.1 Order Processing

There are a few concepts about how partners handle orders and the steps they go through before their box is dispatched that are essential to comprehend in order to understand how every order is placed and where problems may arise.

This is especially important to the company because the entire operation revolves around the six steps, showed in Figure 2, devised in the early stages of the company to ensure that orders were fulfilled efficiently. This breakdown is important not only for organizational reasons, but also to improve efficiency by measuring and controlling each step and serving as cornerstones for the operational departments' responsibilities.



Figure 2 - Farfetch Order Processing Sequence

When a customer adds more than one product to its basket on the website, the products can come from different boutiques. This means that both products are being stored in different points and will most probably reach client's house at different times. Therefore, two different types of orders need to be defined: Portal order and Boutique order. In Figure 3, an example that illustrates the two concepts is given in which an order is made by a single customer of three different products. The portal order presented is composed of two different boutique orders, the first one with one product, the purse, and the second with two, the dress and the earrings.

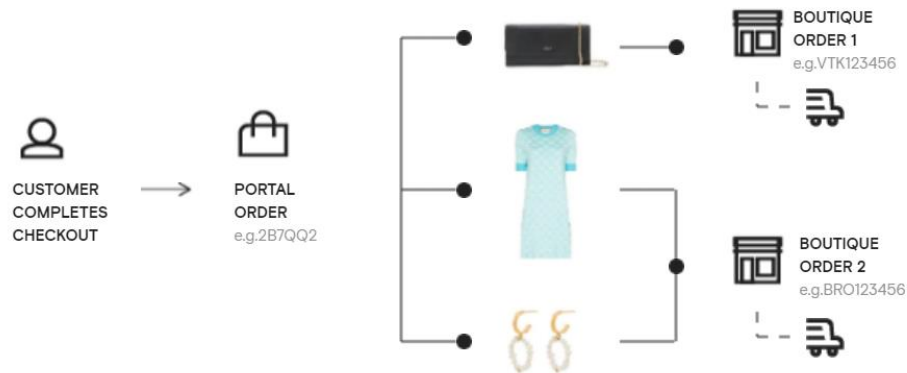


Figure 3 - Diagram with the different types of Farfetch Orders

Step 1 – Order Validation

Firstly, the allocated partner for a specific order confirms their stock availability. The boutiques check if they physically have the product requested online by the customer in stock. This step can have two outcomes, either the boutique confirms that they own stock and the order progresses to the following steps of the workflow, or there were some stocks sync issues and the partner does not have the order item in stock. An example of the origin of these problems is the fact that partners also sell in-store, so the stock update may not have been performed yet when a customer tries to buy a product on the platform. This is an outcome that will penalize the partners performance and, in this situation, Farfetch checks if another partner can fulfil the customer's request. When successful, the boutique order is allocated to another partner and the order restarts in step one. In cases where it is not possible to find an alternative, the order is cancelled, and the customer will be refunded. For ease of understanding, Figure 4 shows a diagram representing the stage in Step 1 of the Order Processing.

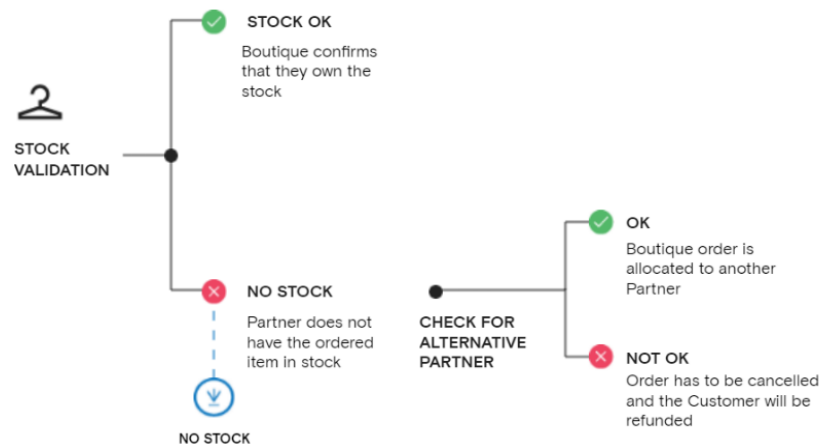


Figure 4 - Diagram showing the stages of Step 1 of the Order Processing

Step 2 – Fraud Check

The next step on Farfetch's order processing sequence is the fraud check. The representation of this step in the form of a diagram can be seen in Figure 5. In this step, Farfetch evaluates the fraud risk associated with the order. To do this, the company relies on external service providers that support internal fraud detection teams with a fraud analysis. The first analysis is done automatically and can have three possible outcomes: the order can be accepted if it is considered non fraudulent, it can be cancelled if it is considered very suspicious, or it can be sent for manual revision in cases where there is some doubt. In the manual revision process, an agent from the fraud team can contact the customer to verify their personal details such as confirming billing or address details to make sure everything matches, and it is a legitimate customer.

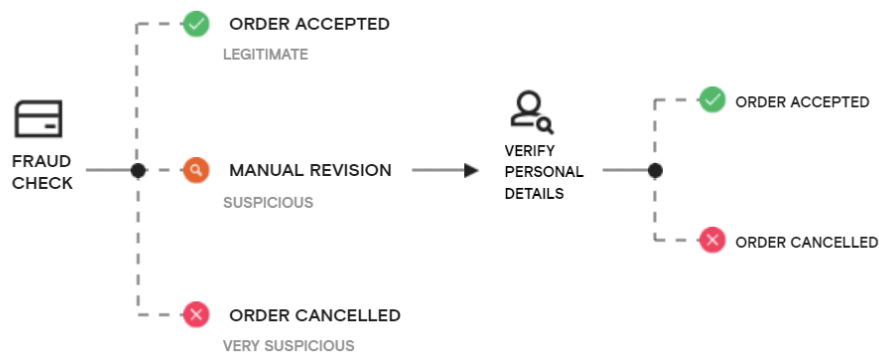


Figure 5 - Diagram showing the stages of Step 2 of the Order Processing

Step 3 – Product Packaging

After the two previous stages being completed successfully, the item's packaging must be decided. In this step, the partner begins packaging the product by selecting the box in which the items will be shipped to the customer. Farfetch's current packaging supply chain is composed of several suppliers distributed worldwide.

When an order is placed, Farfetch recommends a specific box for each product based on its dimensions and considering the best customer experience and delivery cost. As of 25th June 2022, Farfetch has 12 sizes of boxes available for partners to ship their products in. Partners can either accept the recommendation, or they can decide to reject them and change the box size. To do this, the partners create an exception and justify the reason for that decision. This step is also monitored by a key performance indicator called “Packaging Accuracy”. In order to facilitate understanding, Figure 6 shows a diagram representing the stages in Step 3 of the Order Processing.

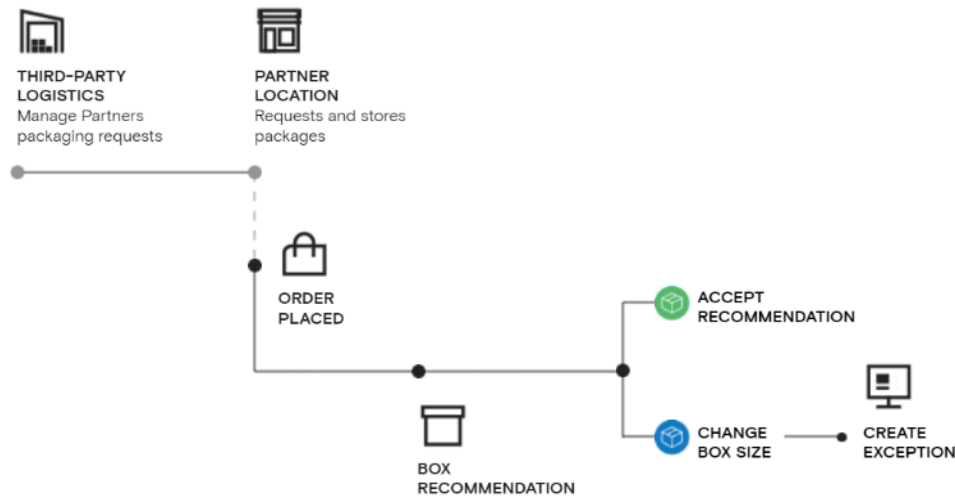


Figure 6 - Diagram showing the stages of Step 3 of the Order Processing

Step 4 – Shipping Preparation

This is the printing step, when the boutique issues the documents necessary for the couriers and the client. Partners order the creation of the shipping label in the platform which complies with the specifications demanded by the courier chosen for that order. Farfetch generates the AWB (air waybill) with a detail of the courier and the address to where the parcel should be delivered. During this step, if there is any problem like an error in the shipping details or a legal restriction, the Delivery support team is automatically notified to try to solve the problem and allow the order to be shipped. The diagram representation of this step can be found in Figure 7.

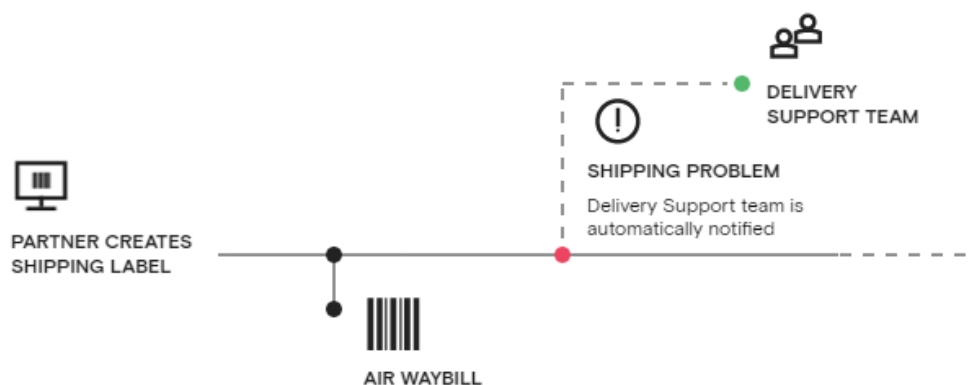


Figure 7 - Diagram showing the stages of Step 4 of the Order Processing

Step 5 – Parcel Pickup

As the shipping label is ready, the order is pushed forward to Step 5 to be collected by the courier. In this step, the partner closes the box, which is ready to be picked up by the courier. Pickups are performed once a day for partners with a low volume of orders and twice a day for the most important partners. The shipment is considered in-transit from the time the parcel is scanned by the courier, and the courier is responsible for delivering it within a specified time window in accordance with service level agreements. During this phase, if the partner experiences any problems with a courier or the shipping, they can create an exception to avoid being penalized in the performance metric that evaluates this process which is called the “Speed of Sending”. To facilitate the understanding of this step, Figure 8 shows a diagram representing the details of the steps in Step 5 of the Order Processing.

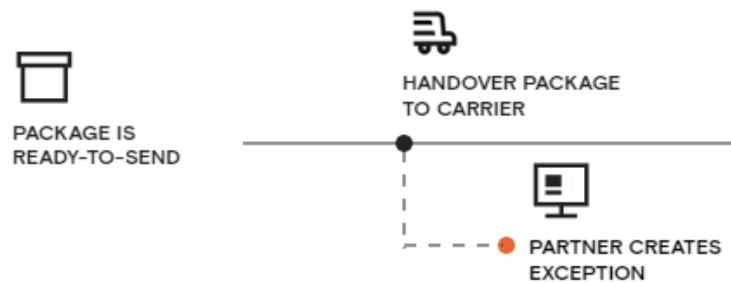


Figure 8 - Diagram showing the stages of Step 5 of the Order Processing

Step 6 – Order Delivery

The final step of the order processing sequence starts when the package is picked up and scanned by the courier. This process is managed by Farfetch, partnering with more than 15 couriers globally like DHL, UPS and FedEx which ensures all orders arrive at their destination in the estimated delivery window.

Upon parcel scanning by the courier, package is considered “in transit” and the customer is given the tracking number in order to monitor the parcel’s journey. The delivery teams also track the parcel in order to tackle issues that might occur during this stage. The diagram representation of the step 6 can be found in Figure 9.



Figure 9 - Diagram showing the stages of Step 6 of the Order Processing

3.1.2 Returns Overview

Returns are a typical component of eCommerce platforms. It is necessary to make a distinction between RTOs ("Return to Origin") and Returns by Users in order to fully understand the scope of Farfetch returns.

RTOs

The term "Return to Origin" (RTO) refers to a situation in which the shipment is cancelled and, as a result, a return request is made before the customer receives the product. A post-order reason needs to be triggered for an RTO to be started. The package is identified and tagged to prevent the delivery from being made and to allow it to be sent back to its origin location after the appropriate delivery provider has been notified. However, there are times when the RTO happens too close to delivery and there is not enough time between the trigger and the parcel identification to stop the delivery process, resulting in the incorrect delivery of the item.

In any case, a delivery performed after an RTO is issued is a situation that should be avoided. In case of customer request, it might raise doubts about the efficiency of delivery processes; if the cancellation was ordered by a boutique or brand, it might leave customers with a flawed image of both Farfetch and the partner; if the reason for the RTO was fraud, then this implies the total loss of the item value and the refund of the customer that was victim of fraud. Despite the existence of improvement opportunities in the RTO management process, this return type is of a much smaller number when compared to returns created by users.

returns created by users.

Returned by User

As opposed to returns by RTO, these returns are characterised by being created by the customer - meaning that the parcel did, in fact, reach its final destination in terms of the outbound order perspective. These types of returns account for the majority of returns when compared with RTOs. Orders returned by users represent around 16% of total orders for 2021. The return rate is a challenge in itself and is being tackled by multiple teams and departments inside Farfetch. Customers may return products without necessarily having a reason to do so. Nonetheless, users are prompted to leave a reason upon return creation. The distribution of these reasons for 2021 is detailed in the Figure 10.

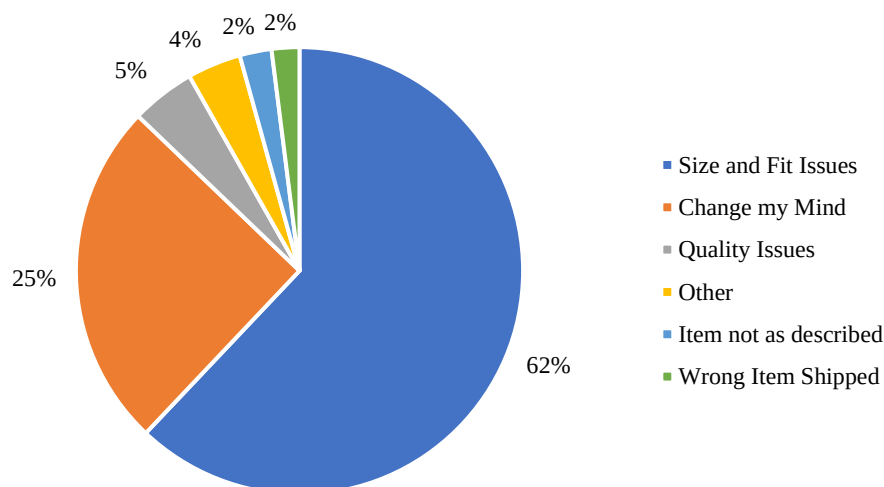


Figure 10 - Pie Chart demonstrating the distribution of the Return Reasons for 2021

As it can be observed, the main reason for item return is due to sizing issues. This is commonly the issue generating most returned items (Reid *et al*, 2016). With the majority of products being wearables, it is often hard to achieve the perfect fit with an online purchase. Elaborate sizing guides help the choice of the most indicated size, however this issue it is still prevalent when selling clothes online compared to brick-and-mortar retail.

In terms of the type of routes that Returns take, they can be Single Leg (Direct Returns) or Double Leg (Indirect Returns). The most common Returns are Single Leg Returns, in which the product goes from the customer's country to the boutique directly. However, the European countries of Belgium, Bulgaria, Finland, France, Greece, Italy, Latvia and Monaco have the type of Indirect Returns for products coming from outside the European Union. For these countries, the products pass through an intermediary EU country before reaching the destination country. This happens because these countries charge extra taxes, namely duties, on products coming from outside the European Union since they don't consider them as returns. So Farfetch uses an intermediary country, like the Netherlands, so the return can be made intra-EU and the final country does not charge the extra taxes.

Due to the fact that RTOs will not be a matter of study in this dissertation, and their introduction was merely to expose the different types of returns, "Returns by User" will be referred to solely as "returns" in order to simplify future analysis.

3.2 Financial Structure of The Company

It is crucial, for the project, to comprehend Farfetch's financial structure after learning about its business model. The financial report Profits and Loss Statement, or P&L, will be presented in order to comprehend the company's financial structure and which costs the project will incur. The schematic representation of the main components of Farfetch's P&L is shown in Figure 11.

The Profit and Loss Statement (P&L), or Income Statement, reflects the effect of management's operating decisions on business performance and the resulting accounting profit or loss over a specified period of time. It is a high-level approach to the financial issues of a company, as it gives a clear idea, without going into details, whether the Organization makes or loses money. It is a fundamental difference to the economical elements because most of them represent a single moment. The fact that the P&L is a high-level tool, when used correctly "can stimulate questions about the effectiveness of management strategies as well as the quality of operational decisions" (Helfert, 2001).

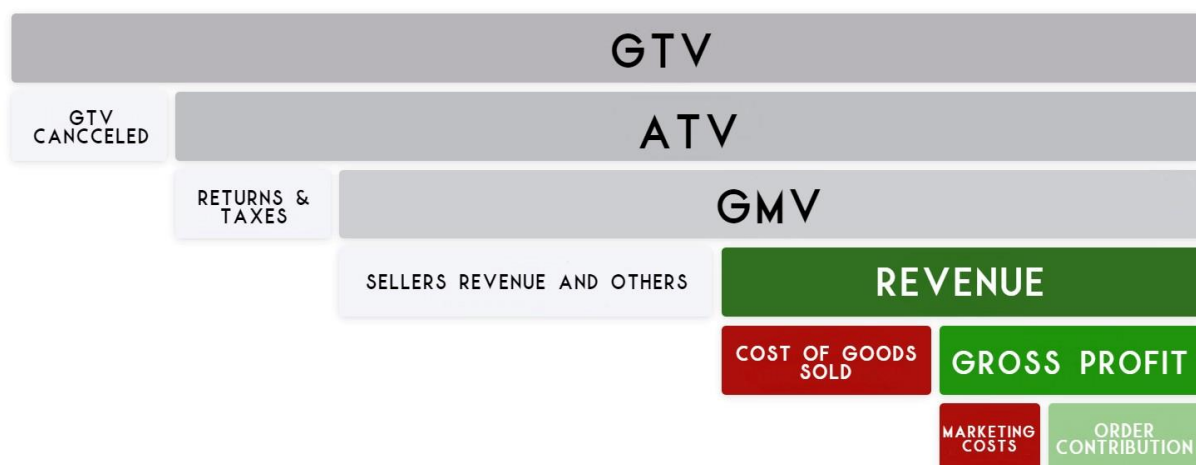


Figure 11 - Simplified diagram of Farfetch's Profits and Losses Report

The GTV, Gross Transaction Value, corresponds to the sum value of all the orders of the items for which the purchase process was triggered by the customer.

Concerning ATV, Actual Merchandising Value, also known as GTV Delivered, it corresponds to the sum of the value of all items that were actually delivered. As mentioned before, the Order Processing is complex and, as such, events may occur at any step of the process and led to the product not being delivered to the consumer. The reasons identified by the company that may lead to a non-delivery of the product are:

- No Stock. This type of cancellation corresponds to an error in step 1, the Stock Validation step, in which a boutique, when performing the Stock Check, verifies that it does not have the requested product in stock.
- Payment Refused. This type of cancellation occurs in step 2, the Fraud Check step of the Order Processing sequence. This type of cancellation occurs when any form of payment fraud is detected on the customer end.
- Other Cancellations. This type of cancellation corresponds to all the other cancellations that are not due to No Stock or Payment Refused. This type of occasional cancellations can happen either by Farfetch fault or by the customer.

In addition to the items not delivered, the value of the Yet to be Shipped products is also not considered in the ATV. These are the products that are still in the order processing stage, that is, they correspond to products associated with orders that were not cancelled but not yet delivered. Therefore, as these items have not yet been delivered to the customer, their value is not taken into account when calculating the ATV. In total, for the year 2021, the ATV corresponded to 89.4% of the GTV.

The GMV, Gross Merchandise Value, corresponds to the value of the ATV from which is subtracted the value of the ATV Returns and the Sales Tax. The ATV Return is the value of the items returned. The value of the Sales Tax is that of the sales taxes on the items that were not returned. For the year 2021, the value of the GMV corresponded to 67% of the GTV.

Revenue corresponds to the value of the GMV from which is subtracted the amount paid to the Boutiques. Another way to obtain the Revenue consists of adding the Fulfilment Revenue and the Platform Revenue and subtracting the cost of the site's promotions. Since the project focuses on costs, it will not go into much detail in describing the types of revenue. However, it should be noted that Platform Revenue is a revenue that comes from the client and is already included in the amount the client pays for the product. On the other hand, Fulfilment Revenue is associated with the partners. The largest proportion of this revenue is linked with the commission required by Farfetch based on the net sales of the products sold by the partners on the website.

The Gross Profit corresponds to the value of the Revenue minus the value of the Costs of Goods Sold (CGS). For the year 2021, the Gross Profit equalled 13.1% of the GTV. The Costs of Goods Sold are divided into 6 types of costs:

- Shipping Cost. This includes what Farfetch pays to couriers, such as DHL, FedEx and UPS, to deliver the products. This cost does not include the costs that Farfetch pays to couriers for returns processing.
- Duties Cost. This cost corresponds to the amount Farfetch pays in Duties to the competent authorities due to the sale of the products. This cost does not include the Duties paid by Farfetch for Returns.

- **Return Shipping Cost.** This cost is similar to the Shipping Cost. However, it refers to the Returns process. Furthermore, it corresponds to the amount that Farfetch pays to the couriers for them to process returns.
- **Return Duties Cost.** These include the duties paid on the return process. It is a value that Farfetch pays to the competent entities.
- **Credit Card Costs.** This cost corresponds to what Farfetch pays to the credit card providers for them to process the sales. It is divided into a cost per transaction plus a fee, according to the payment amount.
- **Packaging Costs.** This is the total value paid by Farfetch to the box suppliers. The boxes are later sold to the partners.

This cost structure can be further broken down into 19 different, more specific costs. Some of these costs are region-specific, such as **Other COS Brazil's Tax**, which is a tax applied exclusively to items shipped to Brazil. In the analysis performed throughout the dissertation, this is the level of detail to which the cost structure will be analysed.

Lastly, the Order Contribution can be obtained by subtracting the Marketing Costs from the Gross Profit. The Marketing Costs are related with the costs the company incurs in relation to demand generation. For the year 2021, Order Contribution corresponded to 7.8% of the GTV.

Within Farfetch, the development and maintenance of the P&L report is the responsibility of the Operations Strategy team, which was consulted during the development of the dissertation.

3.3 Data Visualization Tool Used at The Company

Data visualization allows the company to extract business insights from their sets of data. Farfetch benefits from identifying new patterns and errors in the data because it helps them paying attention to areas that indicate red flags or progress.

The company is currently making an effort to make their data more visible and easily accessible. To do this, they use Looker to help creating dashboards and other reports. Looker is a powerful business intelligence (BI) tool that can help a company develop insightful visualizations. It offers an easy-to-use workflow, is completely browser-based (eliminating the need for desktop software) and facilitates collaboration on dashboards. Among other benefits, users can create interactive, dynamic dashboards, schedule and automate report distribution, set custom parameters for receiving alerts, and use built-in analytics. Looker relies on their own programming language, LookML, to enable the creation of SQL queries against a specific database. To better understand how it works, it is necessary to clarify some of the common file types in a LookML project (Looker, s.d.):

- A **model** is a customized portal into the database, designed to provide intuitive data exploration for specific business users. Multiple models can exist for the same database connection in a single LookML project. In another words, a model is the semantic layer in Looker that controls logic and gates data access for users. This is created by developers as a model file within the LookML project and contains information about which tables to use and how they should be joined together. Multiple models can exist for the same database connection in a single LookML project, and each model can expose different data to different users.
- A **view** declaration defines a list of fields (dimensions or measures) and their linkage to an underlying table or derived table. In LookML, a view typically references an underlying database table, but it can also represent a derived table. ~

- An **Explore** is a view that users can query. It can be considered as the starting point for a query. In Farfetch, there are several Explores that are maintained by different teams. In each team's Explore, the list of fields are related to their daily tasks. For example, in the case of the Business Intelligence team, the Explore "Sales Price Components" is their responsibility, which has the fields linked to the team's operations such as Shipping Cost, Return Cost, Shipping Country, among others.

In addition to the file types in a LookML project, when developing a dashboard in Looker it is also necessary to clarify some terms (Looker, s.d.):

- A **Look** is a single table or visualization saved as its own individual report. Looks can be added to dashboards, scheduled, shared, and made public. Any changes made to a Look will be reflected in any dashboards that contain it.
- A **dashboard** is essentially a collection of one or more saved Looks, displayed as visualization or text tiles together on one page.

4 Driver Identification, Analysis and Dashboard Creation

4.1 Logistic Costs Identification

After planning the project, the first step was to define which lines of Farfetch's P&L the project would focus on. The main objective of the project was to find the main drivers of logistics costs, and to monitor them using Looker. Consequently, the definition of which cost lines to work on followed two criteria:

- The costs had to be the responsibility of Farfetch's logistics department;
- The costs would need to be encapsulated in the Looker Explore in which the subsequent dashboard will be created.

Therefore, based on the first criteria, the Operations Strategy team was consulted, and it was pointed out all the P&L lines impacted by the logistics department. As a result, it was verified that the cost lines for which the logistics department is responsible are: **Shipping Final Customer COS, Packaging COGS, Shipping COGS Free Returns (Retailers) and Others Item Swap**.

However, not all logistics lines are present in the Looker Explore in which the project was developed. As a consequence, of the four cost lines delimited previously the focus was on the following: **Shipping Final Customer COS** and **Shipping COGS Free Returns (Retailers)**. These costs correspond respectively to the Shipping Cost (SC), and the Return Cost (RC) which corresponds to the nomenclature that will be adopted throughout the dissertation.

4.2 Main Drivers Identification

After identifying which cost lines the project will focus on, the next step was to determine the main drivers that affected each identified cost line. With regard to both Shipping Cost and Return Cost, these are under the scope of the Delivery team. So, after meeting with the team and discussing with them how the ship of Farfetch products works, it was concluded that the main drivers of the selected costs are:

- Service Type, which corresponds to the type of service performed. Currently at Farfetch it may be "Standard" or "Express";
- Courier, which corresponds to the company used to transport the order;
- Route, which corresponds to the route taken between the location of the stock and the final customer;
- Box size, which corresponds to the volumetric size of the box to be delivered;

4.2.1 Time Frame Identification

Before carrying out a study of the drivers identified, it was necessary to establish the timeframe in which the study would take place. In recent years, Farfetch has been constantly changing in order to position itself advantageously and keep up with the changes taking place in the global economic environment. During the final months of 2021 Farfetch gradually moved the location of its main warehouse from the United Kingdom to the Netherlands.

In addition, at the beginning of 2022 Russia was Farfetch's third largest market in terms of GMV. However, due to the political situation in the country, the company suspended its operations with Russia in February 2022.

Hence, the time interval of the drivers' analysis followed the following considerations:

- Since an analysis of the current panorama of the identified drivers is to be carried out, the time interval selected should be such that the weight of the location of the warehouse in the United Kingdom is minimized. A too extended time interval would result in routes originating in the United Kingdom having a bigger influence which does not correspond to what Farfetch's current situation looks like;
- As with the previous consideration, the time interval selected should be such that the weight of orders bound for Russia is minimized. Therefore, due to the uncertain situation in the country, and in order to represent the most current situation of the company, the time interval considered should try to minimise the impact of Russia on the Routes driver;
- The timeframe should be wide enough to ensure that the results obtained are reliable. An analysis with a short timeframe is more prone to anomalies and exceptions that do not correctly represent the company's current environment.

As a result, based on the considerations listed above, the time interval of the analysis was the first 4 months of 2022. That is, January, February, March and April 2022.

4.2.2 Service Type

The Service Type corresponds to the type of service ordered to the courier for the transportation the product. For the period under consideration, around 71% of the Shipping Costs were associated with orders with Service Type Express and the remaining 29% were associated to orders with Standard Service Type.

As far as Return Costs are concerned, around 62% of all Return Costs were associated with Service Type Express and the remaining 38% were associated with Shipping Type Standard.

4.2.3 Courier

The driver Courier corresponds to the company hired to transport the product from where it is stored to the final customer. In general, Farfetch works with 3 Couriers: DHL, FedEx and UPS. With regards to the Shipping Cost, about 77% of the costs were associated with orders shipped by DHL, 12% are associated with orders shipped by FedEx and 9% are associated with orders shipped by UPS. Table 2 presents a summary of the shipping cost split between the different Couriers.

Table 2 - Shipping Cost Distribution by Service Type

Courier	Shipping Cost
DHL	77%
Fedex	12%
UPS	9%

With regards to the Shipping Cost, about 70% of all Return Costs were associated with parcels transported by DHL, 20% were associated with parcels transported by FedEx and 8% were associated with parcels transported by UPS. Table 3 presents a summary of the Return Costs split between the different Couriers.

Table 3 - Return Cost Distribution by Service Type

Courier	Return Cost
DHL	70%
Fedex	20%
UPS	8%

The remaining Couriers used by Farfetch are not part of the analysis due to their low contribution to the company's Return Cost. This driver is important as each Courier has a different way of calculating their costs.

4.2.4 Route

The Route driver corresponds to the combination of the product and customer's location. So, firstly, the analysis carried out separated the Store Country, which corresponds to the country where the product is stored, and the Shipping Country, which corresponds to the customer's country.

Store Country

For the period considered, Farfetch worked with partners from 53 different countries. However, for the analysis, the 4 main countries were considered individually. The remaining were grouped according to the geographic regions considered by the operational teams of the company.

As far as the Store Country of the Shipping Cost analysis is concerned orders from Italy corresponded to 37% of the total Shipping Cost, followed by the Netherlands with 12% and the United States and United Kingdom both with 11%. Therefore, these 4 countries were considered individually for the analysis, while the remaining countries were grouped in the different geographic regions. The aggregation of the countries in the different regions is detailed in the Figure A.1 of Appendix A. The values that a Store Country can take and its percentage impact on the Shipping Cost are detailed on table 4:

Table 4 - Shipping Cost Distribution by Store Country

Store Country	Shipping Cost
Italy	37%
Europe	19%
Netherlands	12%
United Kingdom	11%
United States	11%
Eastern Asia	4%
Western Asia	2%

The analysis of the Store Country for the Return Cost followed the same reasoning as for the Shipping Cost. However, regarding the return process, there is the possibility of Indirect Returns, that is, Double Leg Returns. Referring to the aggregation of the countries, there was a separation of the countries that have the Direct Returns system and the Indirect Returns system, namely for the European countries mentioned in the Problem Description.

As for the Shipping Cost, the 4 main Store Countries are Italy, Netherlands United States and United Kingdom, which represent 42%, 13%, 9% and 8% of the Return Cost, respectively. The aggregation of the countries in the different regions followed the same reasoning as for the Shipping Cost. The only difference was the separation of the European countries in EuropeDR for countries with Direct Returns and EuropeIR for countries with Indirect Returns. The values that a Store Country can take and the percentage impact on the Return Cost are shown in Table 5.

Table 5 - Return Cost Distribution by Shipping Country

Store Country	Shipping Cost
Italy	42%
EuropeIR	13%
Netherlands	13%
United States	9%
United Kingdom	8%
EuropeDR	7%
Eastern Asia	3%
Western Asia	2%

Shipping Country

Focusing on the Shipping Cost, during the period under analysis, Farfetch sold products to 198 different countries. As in the previous analysis, the 4 main countries were considered individually, while the rest were grouped in the different geographic regions considered by Farfetch.

Regarding the Shipping Costs analysis, the country with the biggest Shipping Costs is the United States which accounts for 22% of the total Shipping Costs, followed by China, Russia and the United Kingdom with 9%, 6% and 5%, respectively. The aggregation of the rest of the countries in the different regions is detailed in the Figure A.2 of the Appendix A. The values that a Shipping Country can take and its percentage impact on the Shipping Cost are detailed in table 6.

Table 6 - Shipping Cost Distribution by Shipping Country

Shipping Country	Shipping Cost
United States	22%
Eastern Asia	20%
Europe	17%
China	9%
Russia	6%
Central and South America	6%
United Kingdom	5%
Australia and New Zealand	4%
Western Asia	4%
Northern America	1%

Following the same method used up to this stage, the four destination countries that were not aggregated into regions and that contribute the most to the Return Costs are the United States, China, the United Kingdom and Russia which represent 26%, 11%, 7% and 6% of the Return Costs, respectively. The aggregation of the countries in the different regions followed the same criteria as for the Shipping Costs. The table 7 shows the values that a Shipping Country can take and its percentage impact on the Return Cost:

Table 7 - Return Cost Distribution by Shipping Country

Shipping Country	Return Cost
United States	26%
Europe	19%
China	11%
Eastern Asia	11%
Western Asia	8%
Russia	7%
United Kingdom	6%
Central and South America	4%
Australia and New Zealand	4%
Northern America	2%

After determining the values that the Store Country and the Shipping Country can take for the Shipping Cost and the Return Cost analysis, they were combined in order to create the Route driver. In the case of the Shipping Cost analysis the Route corresponds to the Store Country followed by the Shipping country. In the case of the Return Cost analysis, the Route consists of the Shipping Country followed by the Store Country, as this is the route when a Return happens. The example of a case for both is described in Table 8 for the Shipping Cost and in Table 9 for the Return Cost.

Table 8 - Variable Route Creation when analysing the Shipping Cost

Store Country	Shipping Country	Route
Italy	United States	Italy-United States

Table 9 - Variable Route Creation when analysing the Return Cost

Store Country	Shipping Country	Route
Italy	United States	United States-Italy

4.2.5 Box Size

The final driver identified is the size of the box to be shipped. The size of the box is a decisive factor in the cost of the order since it is one of the criteria by which couriers calculate the price of the delivery. However, in the Explore in which the dashboard was developed, it is not possible to filter through the size of the box used. Therefore, the best approximation to the size of the box is the type of product to sold and, for this, the variable "Product Category" was used. This variable divides products into 6 categories: Accessories, Bags, Clothing, Kids, Shoes and Other. Besides, optimisations in what concerns the product, like a small change on the website product description, could be more accessible and instantaneous than the optimisation of the boxes. The values which the Product Family Group can take and their impact on the Shipping Cost and Return Cost are shown in the Table 10 and Table 11, respectively.

Table 10 - Shipping Cost Distribution by Product Category

Product Catagory	Shipping Cost
Clothing	39%
Shoes	27%
Kids	9%
Bags	8%
Acessories	6%
Other	11%

Table 11 - Return Cost Distribution by Product Category

Product Category	Return Cost
Clothing	50%
Shoes	27%
Kids	7%
Bags	5%
Accessories	5%
Other	6%

With regard to the products classified as "Other", these are products from a wide variety of categories, the monitoring of which will not result in initiatives and opportunities for improvement. For this reason, products classified as "Other" were excluded from the analysis.

After identifying the drivers of the main cost lines, the next step was to obtain the data to be analysed using the Big Query coding tool which uses the SQL programming language. Therefore, a query was developed in order to obtain the Shipping Cost and the number of Boutique Orders for each different combination of the drivers that were previously analysed. Additionally, in another query, for each different combination of drivers, the number of returned Boutique Orders and the Return Cost were collected. With these values, it was possible to calculate the Shipping Cost per BO, which is the Shipping Cost divided by the number of BO. In the same manner, it is possible to calculate the Return Cost per BO, which is the Return Cost divided by the number of BO. At the end of the data collection, a table with the format of Table 12 was produced.

Table 12 - Simplified data collection Table

Service Type	Courier	Product Family	Route	Shipping Cost (USD)	Boutique Orders	Shipping Cost per BO (USD)
Express	DHL	Clothing	Italy-United States	110,410	3,544	31.15

The subsequent analysis focused on the Shipping Cost per BO and Return Cost per BO values. Consequently, in order to have the representativeness of the number of BO in the model, a new table was created in which each row corresponded to one BO. For the Shipping Cost analysis, 2,753,927 Boutique Orders were analysed and for the Return Cost analysis 425,424 Boutique Orders were analysed. Each Boutique Order corresponds to one row in the table.

4.3 Driver Evaluation

Subsequently, statistical models were applied in order to assess the impact of each driver on the Shipping Cost and Return Cost. The statistical models had the goal of describing the driver's current situation and not to predict future values. In order to evaluate and rank the impact that each driver has on both the Shipping Cost and the Return Cost, linear regression models were used. Furthermore, in order to find critical driver combinations with a high impact on costs, the Decision Tree method was used.

4.3.1 Linear Regression Model

The first model developed was the linear regression and its main goal was to rank the drivers so that Farfetch could perceive which drivers had the greatest impact on the logistics cost structure. Therefore, using the R programming language, a linear regression model was developed. Then the impact of each driver on the objective variable was evaluated.

The dataset was divided into training set and testing set. Due to the size of the dataset and the limitations of computational power, it was decided that 70% would represent the training set while the remaining 30% would represent the testing set. The mentioned partition was conducted randomly, resulting in 2,028,813 data instances for training and 869,513 for testing for the Shipping Cost and 297,797 data instances for training and 127,627 for testing for the Return Cost.

A data reduction process had to be performed due to the computational demanding nature of one the model used combine with a large dataset. As described, the training set is exploited by the model to capture the inner relationships between variables and constantly validate it in an unbiased manner. To accomplish that, the K-fold cross-validation technique was applied avoiding overfitting and, hence, creating the foundations for reliable results. A ten-fold cross-validation was selected. Finally, the testing set was used later to assess the model performance.

Consequently, two linear regressions were performed that had as the objective variable the Shipping Cost per BO and Return Cost per BO. In both analyses, the independent variables were all categorical, so it was necessary to create a set of dummy variables that would represent the categorical attributes. Therefore, one-hot encoding process was employed to perform the mentioned task.

As for the Shipping Cost, the linear regression revealed a Multiple R-Squared of 0.784 and an adjusted R-Squared of 0.784, meaning that the model explains approximately 78.4% of the variability in the response data. In order to evaluate the model performance, the Residual Mean Squared Error (RMSE) was used, which is the standard deviation of the residuals (prediction errors). Regarding the model developed, this revealed an RMSE of 2.67.

However, the most important part of this analysis is to assess the impact of each variable on the Shipping Cost per BO. Therefore, in order to evaluate the influence of each variable in the model, the **VarImp** function from the **caret** package was used. This function uses the absolute value of the t-statistic for each parameter in the model. The t statistic is the coefficient divided by its standard error. The standard error is an estimate of the standard deviation of the coefficient, the amount it varies across cases. It can be thought of as a measure of the precision with which the regression coefficient is measured, that is, the larger the t-statistic, the more certainty there is that the coefficient is not zero (Princeton University, 2007).

It was decided to perform a t-statistic analysis rather than a beta weight analysis. The beta weight analysis revealed which drivers had a higher marginal contribution to costs. However, the output revealed mostly high cost routes that had no significant impact on the company's cost structure. In other words, these are routes that are commonly not used and, as such, their optimization would not reveal a large reduction in logistics costs. Therefore, after conferring with the Logistics Optimization team, it was concluded that the method to be followed would be the t-statistic analysis.

Therefore, applying the **VarImp** function to the developed model the ten drivers that have the most importance in defining the Shipping Cost per BO are presented in the Table 13.

Table 13 - Most important Shipping Cost Drivers based on the VarImp function

Driver	VarImp
Courier DHL	910
Route Italy-United States	894
Route United Kingdom-United States	838
Product Category Shoes	785
Route Western Asia-United States	774
Route Netherlands-United States	773
Route Western Europe-United States	746
Route Italy-Western Asia	740
Route Italy-Russian Federation	725
Route Eastern Asia-United States	696

As for the Return Cost, the linear regression revealed a Multiple R-Squared of 0.891 and an adjusted R-Squared of 0.891, meaning that the model explains approximately 89.1% of the variability in the response data. Regarding the model developed, this revealed an RMSE of 2.54.

As for the importance of the identified drivers, the **VarImp** function of the package **caret** was applied again. The 10 most important variables in defining the Return Cost per BO, according to the function, are shown in the Table 14:

Table 14 - Most important Return Cost Drivers based on the VarImp function

Driver	VarImp
Route United States-Italy IR	543
Route China-Italy	450
Route United States-Europe DR	437
Route Eastern Asia-Italy	431
Route United States-Europe IR	426
Route Europe-United States	422
Route Western Asia-United States	361
Route Central and South America-Italy	361
Route Eastern Asia-United States	352
Route Eastern Asia-Europe IR	344

Based on the results obtained, it can be concluded that for the Shipping Cost analysis of the ten main drivers influencing its value, one is a Courier, one is a Product Category and eight are Routes. The main driver influencing the Shipping Cost is the DHL Courier, which is to be expected since it represented 77% of all volume shipped by Farfetch in the first 4 months of 2022. In addition, it is worth noting the impact not only of the Italy-United States route, but also the United States as a Shipping Country. Of the eight routes in the top ten, six of them have the United States as the destination Country, which proves the impact that the most important country has on the company's cost structure.

Regarding the Return Cost, contrary to what happened for the Shipping Cost, all the drivers that have the biggest impact in the cost are Routes. As it happened with the Shipping Cost, the route that impacts the most is the one where Italy is the Store Country and the United States is the Shipping Country. Opposed to the Shipping Cost, there are more routes that impact the Return Cost. However, the presence of the United States both as a Store Country and as a Shipping Country is noteworthy.

4.3.2 Decision Tree Model

Once the drivers with the greatest impact on both Shipping Cost and Return Cost are known, it is very useful to find combinations of drivers that show a great impact on Shipping Cost and Return Cost. Therefore, a Decision Tree model, the Regression Tree, was applied, since the outcome variable is numeric (Shipping Cost per BO or Return Cost per BO). This machine learning model aimed to find combinations of drivers that had a strong influence on the costs.

Contrary to the usual use of the Decision Tree machine learning model, the objective of using this application is descriptive, not predictive. In other words, the goal of the algorithm is not to create a rule and make predictions, but to identify the main contributors to the logistics cost structure.

Again the data was split into two subsets. A training subset with 70% of the data and a test subset with 30% of the data. In what concerns the Decision Tree constructed, it had as split criteria the ANOVA method, used in Regression Trees, which has as objective the minimization of the variance in each Node. Regarding the remaining parameters of the Decision Tree, the only one that suffered changes was the complexity parameter (CP). If the cost of adding another variable to the decision tree from the current node is above the value of CP, then tree building does not continue. As already mentioned, the goal of this Decision Tree is not to have the optimal decision tree, but to create one with enough depth that the final nodes consist of driver combinations with high detail. Therefore, a CP of 0.0005 was set for both analyses, which for the Shipping Cost resulted in a Decision Tree with 1256 Nodes and an RMSE of 1.57 and for Return Cost resulted in a Decision Tree with 946 Nodes and an RMSE of 1.65.

By analysing the final nodes of the Decision Trees, it was possible to find combinations with a high impact on costs. A combination with a high impact is a combination with a high cost per BO and a high frequency. Let us take for example figure 12, which corresponds to one of the most critical driver combinations in the Shipping Cost analysis.

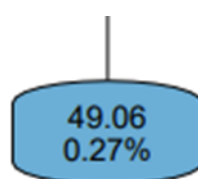


Figure 12 - Example of a Decision Tree End Node

This final Node corresponds to the Products of Category Shoes, shipped with the Service Type Express by the Courier DHL on the Route Italy-United States. In summary, this combination of drivers, which corresponds to 0.27% of all orders analysed for the period in question, has an average shipping cost per BO of \$49.06.

Following this reasoning, it was possible to build the Table 15 and Table 16 which indicate the driver combinations with a higher impact on both Shipping Cost and Return Cost respectively.

Table 15 - Most important combination of drivers in the Shipping Cost

Store Country	Shipping Country	Courier	Product Category	Service Type	Shipping Cost per BO	Percentage total BO's
Italy	United States	DHL	Shoes	Express	\$49.06	0.27%
UK	United States	DHL	Shoes	Express	\$41.17	0.15%
UK	United States	DHL	Kids	Express	\$74.10	0.02%
Netherlands	Eastern Asia	DHL	Shoes	Express	\$29.39	0.35%
Italy + UK	United States	DHL	Clothing	Express	\$31.22	0.25%

Table 16 - Most important combination of drivers in the Return Cost

Shipping Country	Store Country	Courier	Product Category	Shipping Type	Return Cost per BO	Percentage total BO's
Central and South America + China + US	Italy	DHL+ Fedex	Shoes	Express	\$34.61	1.08%
Northern America	EuropeIR + Italy	DHL+ Fedex	Shoes	Express	\$39.25	0.18%
United States	Italy	DHL+ Fedex	Accessories + Clothing	Express	\$39.86	0.16%
United States	EuropeDR	Fedex	Shoes + Bags	Express	\$31.76	0.59%
Northern America + Eastern Asia	Netherlands	DHL	Shoes	Express + Standard	\$30.05	0.33%

The results obtained allowed Farfetch to assess which driver combinations impact the Shipping and Return costs the most. By analysing the results, we observe that they are in accordance with the results of the drivers' ranking obtained in the linear regression. In terms of Shipping Cost, there is a great dependence on the United States as Shipping Country, as well as on the DHL Courier, which is present in all combinations found. Furthermore, it is also worth mentioning Italy and the United Kingdom as important Store Countries and the fact that, in all combinations obtained, the service type is always Express.

When it comes to the analysis of Return Cost, as in the results obtained through Linear Regression, these are more scattered. It is still possible to verify the importance of the United States as Shipping Country. However, it does not have the same importance that it has for the Shipping Cost. Regarding the Courier, it is also not that much impacted by DHL, with FedEx having a more critical role. In the Product Category, it can be seen that Shoes are more dominant and, in the Service type, there is one combination of Express and Standard. Overall, the drivers are more diverse in the combinations that influence the Return Cost.

After identifying which combinations have the greatest impact on the logistics costs, it is important to understand that their identification is only the first step in the optimization process. Despite knowing which combinations affect costs the most, it does not mean that they are not optimized properly. Therefore, it is important to have a deep knowledge of the business model and communication between the various teams within Farfetch's logistics department in order to understand how to optimize the identified combinations. The study revealed which combinations could be improved in order to have a greater impact on the company's costs, however, finding optimization opportunities is another process in which two examples will be demonstrated further in the dissertation.

4.4 Dashboard Creation

After the identification of the main drivers of Shipping Cost and Return Cost, the project included the creation of a strategic dashboard so that it could be possible to monitor them over time and find opportunities for optimisation. Therefore, the tool Looker was used, which operates as a part of Google Cloud Platform and offers a wide variety of tools for relational database work, business intelligence, and other related services. This tool uses the LookML language which is a language for describing dimensions, aggregations, calculations, and data relationships in a SQL database. Looker relies on their own programming language, LookML, to enable the creation of SQL queries against a specific database.

The dashboard created uses two-time intervals in order to evaluate the temporal evolution of the drivers. The first-time interval considered was Year on Year (YoY), which consists of comparing the costs in the year the dashboard is used with the same period last year. Therefore, for the first 4 months of 2022, the YoY cost evolution allows to see how the sum of costs for the first 4 months of 2022 compares with the sum of costs for the first 4 months of 2021. This comparison is highly useful as it allows the company to compare performance variations on a larger time interval.

The next time interval considered was Month on Month (MoM), which consists in comparing the costs of the last complete month when using the dashboard with the costs of the previous month. Therefore, if the dashboard was used in May 2022, it would allow the costs for April 2022 to be compared with those for March 2022. This comparison is also of great importance as it allows the identification of any ongoing problems that may be occurring in a driver from one month to the next.

The dashboard created is divided into 3 distinct sections in order to allow a detailed follow-up of the drivers' evolution. In order to better present it, the Shipping Cost dashboard will be used as an example, but all the information presented subsequently is the same for the Return Cost dashboard. The full Shipping Cost analysis dashboard and its division in different sections can be found in the Figure C.1 of the Appendix C.

In the first section, the costs are evaluated without being filtered for any driver. That is, it displays the total evolution of the Shipping Costs. To do this, the Shipping Cost and Shipping Cost per BO information are presented at the Year to Date (YTD) level. This means from the beginning of the year to the last completed month. The YoY evolution of the Shipping Cost and Shipping Cost per BO is also presented. In addition, the Shipping Cost and Shipping Cost per BO for the last complete month (LM) is presented and, finally, the MoM evolution of the same costs is displayed. The Figure C.2 of the Appendix C represents the section one of the dashboard created.

The next section corresponds to a Look in the form of a chart that allows the user to compare the Shipping Cost evolution of the current year with last year at the monthly level. Using the Dashboard filters corresponding to each driver identified, this section allows the user to track the Shipping Cost and the Shipping Cost per BO throughout the year. Therefore, it is possible to know the monthly evolution of costs for a driver, or set of drivers, compared with last year's values. This chart complements the other two sections of the dashboard since it allows the user to follow the monthly evolution of costs, which is not possible to obtain in the other sections due to the information aggregation limitation. The Figure C.3 of the Appendix C shows the section two of the dashboard created filtered for the Italy-United States Route.

The final section corresponds to the monitoring of each identified driver in a Look table format. In order to better explain it, the Routes will be used as an example, it is important to note, however, that the dashboard contains one table as such for each driver identified. Therefore, for each Route previously identified, the user can see the value of Shipping Cost YTD and Shipping Cost LM, and the percentual value that these account for. In addition, the YoY and MoM evolution of both the Shipping Cost and the Shipping Cost per BO is displayed. This section of the dashboard is especially useful since it allows the user to sort which Routes have suffered a greater cost increase both in terms of YoY and MoM, allowing them to create possible actions in order to understand the increase of these costs. The Figure C.4 of the Appendix C shows the section three of the dashboard created for the Routes driver.

5 Initiatives Analysis

After developing the dashboard that allows for the evolution of the main drivers of shipping cost and return cost to be monitored, the next step was to find opportunities for optimisation in these areas that could result in a reduction of the company's logistics costs.

5.1 Stadium Goods Shipping Cost Optimization

The first possible optimisation opportunity was identified based on the use of the tools developed, namely the driver's monitoring dashboard. Using the dashboard, a significant increase was noticed in the shipping cost per BO of orders within the United States, that is, with both the Store Country and Shipping Country as United States. Compared to the year 2021, there was an increase in the Shipping Cost at the YoY level of about \$1.75M which corresponds to a 23% increase. The Shipping Cost per BO at the YoY level increased by \$4.5 per BO, which is equivalent to a 7% increase. Therefore, from a first analysis, it was concluded that the Shipping Cost increase on this route was due to two factors:

- An increase in the total volume sold on this route, meaning the number of orders increased;
- An increase in the price per Boutique Order.

The next step was to understand which drivers were responsible for the increase in the Shipping Cost per BO. Consequently, the dashboard was filtered so that all drivers were representing the United States-United States Route. It was verified that, for this Route, about 61% of all volume sold is associated with the store Stadium Goods. It was found that, for this Route and this Store, the Shipping Cost YoY increased by about \$1.1M.

Following the same method, the next step was to filter all the other drivers for the Route United States-United States and Store Stadium Goods. Therefore, there was an increase in the Shipping Cost associated with orders with UPS Courier and a decrease in the Shipping Cost associated with orders with FedEx Courier. This is due to an increase in the number of orders transported by UPS Courier and this Courier has higher shipping costs per BO than FedEx. It can be concluded that the main driver responsible for the increase in Shipping Cost per BO on the Route United States-United States with Store Stadium Goods is the Courier.

As next steps, the Delivery team, which is responsible for monitoring the processes associated with Delivery at Farfetch, was notified. After a meeting with the team, it was decided that they would be responsible to address the problem encountered.

5.2 Items Too Big & Too Small Optimization

The next optimisation opportunity was not based on the dashboard created, but rather on the business knowledge of the Logistic Optimisation Team. So, having as a final objective the reduction of Return Costs, a brainstorming session with the Logistic Optimisation team was carried out in order to come up with ideas for possible initiatives supported by an analytical component. The first step was to look at the main reasons why customers make returns. Using Looker, it was discovered that, for 2021, Farfetch's Return Rate was around 16.2%. Furthermore, around 62% of returns were due to size and fit issues.

After that, the returns associated with Size and Fit problems were analysed in order to understand their root causes. It was found that, out of the 62% of returns due to size and fit, about 39.1% are associated with customer reviews that considered the item "Too Big" (TB), and about 37.7% are associated with customer reviews that considered the item "Too Small" (TS). Therefore, around 24.3% of all returns are associated with reviews that consider the Item TB and 23.4% of all returns are associated with reviews that consider the Item IS.

5.2.1 Initiative Planning

After understanding that there was an opportunity for improvement, we worked on defining exactly what makes an item too big or too small. An Item TB is an item that has a high number of returns with customer reviews mentioning the size being too big. These returns may happen due to errors in the size guide of the product or due to looser fit of the item. When it comes to an Item TS, it is an item that generates a high number of returns associated with customer reviews mentioning its small size. As with the Items TB, these returns may occur due to errors in the size guide of the product, the item not having a human model as a reference on the product page or due to a tight fit of the item.

Therefore, with the objective of reducing the Return Rate of these products and, consequently the Return Cost, and based on the information from Returns TB & TS, an initiative was defined consisting of the following steps:

- Define Thresholds in order to correctly identify Items TB & TS;
- Group the items and classify them according to their characteristics;
- Suggest possible actions depending on the characteristics of the Item;
- Evaluate the monetary impact of the initiative for 2022 and 2023.

The first step to consider in the data collection stage was to decide the time interval. After some discussion, it was decided that the analysis would cover January to April 2022, making it 4 complete months in the year. The selection of this time interval was made due to the following reasons:

- The fact that the Product ID, also considered to be the primary key that classifies the product displayed on the FF website, is updated annually; and using Looker, it is not possible to track the evolution of the number of sales and returns from one year to the next for a certain item, so the analysis had to be carried out in a single year.
- The aim of the analysis is to be as current as possible in order to filter products that are effectively TB and TS. An analysis with a wider time interval might include items no longer considered TB or TS due to possible changes already made.
- The analysis must have a sufficient time interval to allow the customer to purchase a product and make a return.

5.2.2 Definition of Thresholds

After defining the time interval of the analysis, the next step was to analyse the returns in this time interval and define the possible thresholds that allow filtering the items that are effectively Items TB or Items TS.

Minimum Number of Returns TB & TS

The first threshold defined was the Total Number of Returns TB & TS. For the first 4 months, the number of Returns due to the Item being TB was 139,424. In addition, about 87,518 different Items were returned due to the Item being considered TB (number of different Product ID). Therefore, in the first 4 months of 2022, the items returned for being TB had an average of 1.59 Returns per item.

Following the same thought for the Items TS, in the first 4 months of 2022, the number of Returns due to the Item being TS was 136,570. When it comes to the number of different Items, about 83,533 different items were returned due to the items being TS. In the first 4 months of 2022, the items returned for being TS had an average of 1.63 returns per item.

Note that the average number of returns per item is quite similar for both Items TB and TS, with a difference of about 2%. As for the definition of the threshold, it was considered that it would be set at a minimum number of TB or TS returns of 5. This means that, for an item to be considered an Item TB, it must have at least 5 Returns associated to TB reviews. Likewise, for an Item to be considered an Item Too Small, it must have a minimum of 5 Returns associated with TS reviews.

Minimum Return Rate TB & TS

The next threshold to be set was the minimum return rate for TB or TS items. To calculate the Return Rate TB for an item, we divide the number of Returns associated with TB reviews by the total number of sales of the item for the period under consideration. The Equation (1) explains how to calculate an item's Return Rate TB.

$$\text{Return Rate TB} = \frac{\text{Number of Returns TB}}{\text{Total Quantity Sold}} \quad (1)$$

Following the same thought, to calculate the Return Rate TS of an item, divide the number of returns associated with TS reviews by the total number of sales of the item for the period considered, as explained in the Equation (1.1).

$$\text{Return Rate TS} = \frac{\text{Number of Returns TS}}{\text{Total Quantity Sold}} \quad (1.1)$$

For the first 4 months of 2022, the average Return Rate TB of items with at least 1 return per TB was 8,9%. For each 100 units sold, about 9 were returned due to the Item being TB. Regarding the items considered Too Small, their average Return Rate TS with at least 1 return per TS was 8,5%. For every 100 units sold around 9 were returned due to the Item being TS.

For the definition of threshold, it was decided to consider a minimum return rate TB & TS of about 20%. This means that, for an item to be considered too big, it must have a Return Rate TB of at least 20% and, to be considered too small, it must have a Return Rate TS of at least 20%. As an example, for a random item with 100 sales, the item must have at least 20 Returns TB to be considered an Item TB or at least 20 Returns TS to be considered an Item TS.

Maximum Opposite Return Rate TB & TS

Finally, the last threshold considered was the definition of a maximum value of the Return Rate TS for an item to be considered TB and a maximum value of the Return Rate TB for an item to be considered TS.

Following the definition of the first two thresholds, the data was collected and, after a preliminary analysis, it was found that there were items that met the necessary requirements to be considered both an Item TB and an Item TS. However, such items are not relevant to the initiative since they are not encompassed in any specific category and, therefore, the possible actions suggested do not address these type of products.

The first solution proposed was to consider that, for an item to be considered TB, it could not present any return associated to the Item being TS, and an Item TS could not present any returns associated with the Item being TB. However, this threshold was ineffective because significantly reduced the number of items considered. This happens because items with a high number of returns tend to have both TB and TS returns, even if one of the types is considerably superior to the other.

Therefore, it was defined that, for an item to be considered TB it could not have a Return Rate TS higher than 3%, and the opposite to be considered TS. This applied threshold was more balanced and allowed for a more correct identification of TB and TS items.

During this phase of the project all thresholds were defined following advice from the Logistic Optimization team, which confirmed their usefulness in detecting Items TB & TS.

5.2.3 Data Collection and Treatment

After the definition of thresholds, the data was collected using Looker. For this, two distinct Looks were created, one to obtain the data relative to Items TB and the other to obtain the data relative to Items TS.

For each product that was in accordance with the defined thresholds, 5 fields were extracted with it:

- Product ID, which is a numeric variable and is the primary key of the product within the Farfetch system;
- Category, which is a string variable that characterises the product at a general level
- Return Units, which is a numeric variable that indicates the number of units returned TB, when the Item considered TB ; and the number of units returned TS, when the Item is considered TS;
- Return Rate, which is a numeric variable, is the Return Rate TS when the item is TS and is the Return Rate TB when the item is TB
- ATV Return, which is a numeric variable that indicates the amount of ATV Returned resulting from product returns.

Taking into consideration the defined thresholds, a total of 1319 different items were found to meet the defined requirements. Of these 1319 items, about 768 (58%) were classified as Item TB and about 551 (42%) were classified as Item TS. The division of items according to their classification between Item TB and Item TS and their classification according to the Category are presented in Table 17.

Table 17 - Distribution of Items TS & TS by Category 1st

Category 1st	TB	TS
Accessories	7	15
Activewear	9	9
Baby Boy Clothing	0	3
Baby Boy Shoes	0	7
Baby Girl Clothing	1	3
Baby Girl Shoes	0	1
Bags	1	10
Boys Accessories	0	3
Boys Clothing	3	3
Boys Shoes	1	2
Clothing	568	343
Girls Accessories	0	1
Girls Clothing	10	12
Girls Shoes	1	4
Homeware	1	2
Jewellery	2	4
Lingerie & Nightwear	0	3
Shoes	155	122
Teen Boy Clothing	4	0
Teen Girl Accessories	0	1
Teen Girl Clothing	5	1
Underwear & Socks	0	2

After knowing which items were being addressed by the initiative, a session of product analysis and brainstorming with the Logistic Optimization team followed, in order to come up with possible actions that would result in a decrease in the return rate. Therefore, after analysing the products, it was noted that some of them have a tool called "Find Your Size" on the item description page at Farfetch. This tool suggests a product size to the customer based on information provided by the customer. This information can be the size that the customer wears of a particular brand or even physiological data such as height or weight. Furthermore, all the products have an assistant called "Size guide", which corresponds to a table that converts the size of the piece from the country of the boutique to other distinct sizes worldwide. Therefore, the items were classified according to 3 factors:

- Unique Size Item are pieces that are not graduated in terms of size, meaning that there is only one possible model. An example of this type of item is Bags.

- Item With Find your Size are pieces of various sizes that have the tool "Find Your Size" on the Item page.
- Items Without Find Your Size are pieces of various sizes that feature the "Find Your Size" tool on the Item page.

In the case of "Unique Size Items", the suggested actions are to first check if the dimensions presented by Farfetch in the "Size and Fit" section of the product are correct. If Farfetch provides incorrect information regarding the size of the product, it is likely that there will be a large number of Returns, since the dimensions of the item do not correspond to the customer's expectations. The next suggestion is to create a warning on the item page such as "1 in 5 customers find this item oversized" in order to warn the customer of the product's dimensions and decrease the number of returns. The final suggestion regarding Unique Size Items is to ensure that the critical items have a human model as a reference on the product page. Some identified items do not have an image to be used by a model and so the customer has low awareness of their size.

Regarding the items classified as "Item With Find your Size", the suggested actions are the following: first, to check if the "Size guide" is in accordance with the one provided by the brand; the next suggestion is to advise the customer, through the "Find Your Size" tool to buy a size bigger than the one previously advised if the item is as Item TS, or to buy a size below the one previously advised if it is an Item TB.

In the case of items classified as "Item Without Find Your Size" the suggested actions are, as described in the previous paragraph, to check if the size guide is in accordance with the one provided by the brand. The next suggestion is to advise the customer, directly in the "Size & Fit" section, to buy a size above the one they usually buy if the item is classified as Item TS, or to buy a size below if the item is classified as Item TB.

The final suggestion is to change the form the customer fills out when making a return. It is important for Farfetch to understand if the item returned was due to the unwanted fit or if the customer ordered the wrong size. In other words, for an item classified as Item TB, it is important for Farfetch to understand if the item was returned because the fit was loose or if the customer ordered a size bigger than they should have ordered. This information can be of great use to Farfetch in the future as the company is able to identify the products returned due to the customer ordering a size above or below the ideal and create actions on them. So, the general approach is to put the following options on the returns form: "The item was too loose fitting", "The item was too tight fitting", "I ordered a size up", "I ordered a size down".

5.2.4 Monetary Value of The Initiative

As for the monetary value of this initiative, it corresponds to the sum of the ATV Return of the selected items, which Farfetch retains by decreasing the return rate. For the company, when presenting the initiatives, they consider that the value of the initiative should be the highest possible hypothetical value. Therefore, the calculation of the value of the initiative considers that the return rate of these products has gone to 0%. Consequently, to calculate the ATV Return of the Items for the first 4 months of 2022, we proceeded to sum the individual ATV Returns of each item in the first 4 months. For the first 4 months of 2022, the ATV Return of the initiative encompassing the 1319 Items TB & TS was \$4,5 million. For the same period, the total ATV Return for the company was \$290 million.

So, the next step was to forecast the value of the initiative for 2022 and 2023. It was assumed that the proportion of the ATV Return due to Items TB & TS to the ATV Return of all items would remain constant over the year. Therefore, in order to know the value of the initiative for the full year 2022, the Equation (2) was applied.

$$\frac{ATV \text{ Return (4 months 2022)}}{ATV \text{ Return TB \& TS (4 months 2022)}} = \frac{ATV \text{ Return (2022)}}{ATV \text{ Return TB \& TS (2022)}} \quad (2)$$

Knowing that the value of ATV Return (4 months 2022) was \$250 million and the value of ATV Return TB & TS (4 months 2022) was \$5 million, it is still necessary to know the value of ATV Return (2022) in order to calculate the ATV Return TB & TS (2022). For that, we consulted the ATV Return forecast for 2022 made by the Operation Strategy team. According to their forecasts, the value of the ATV Return for 2022 will be close to \$1 billion. After having these three values, it is possible to estimate the value of the initiative for 2022, which would result in \$18,6 million in savings.

For 2023, it was again considered that the proportion of the ATV Return due to Items TB and TS as a function of the ATV Return of all items would remain the same for the full year 2023. Therefore, in order to know the value of the initiative for the full year 2024 the Equation (2.1) was applied.

$$\frac{ATV \text{ Return (4 months 2022)}}{ATV \text{ Return TB \& TS (4 months 2022)}} = \frac{ATV \text{ Return (2023)}}{ATV \text{ Return TB \& TS (2023)}} \quad (2.1)$$

As was done for 2022, the ATV Return forecast figure for 2023 has again been provided by the Operations Strategy team. According to their forecast, the value of the ATV Return for 2023 will be \$1,4 billion. By replacing the values in the expression, the initiative value for 2023 is \$23,7 million, an increase of 27% compared to 2022.

5.2.5 Scenario Analysis

After defining the initiative expected value for 2022 and 2023, the next step was to develop scenarios with different thresholds. Due to the short duration of the project, it was not possible to evaluate the results obtained, so the creation of several scenarios with the variation of the defined thresholds is important as it allows to evaluate how the items categorisation changes at stricter thresholds. Therefore, the stricter the thresholds, the more certain it is that the selected items are effectively Items TB & TS.

Therefore, it was decided to change the threshold of the total number of Returns TB & TS and the threshold of the Return Rate TB & TS. The total number of returns was changed between the values of 5 (original), 10, 15 and 20. As for the value of the return rate TB & TS, it varied between the values of 20% (original), 25%, 30%, 35% and 40%.

For these different threshold values, the number of items covered by the initiative and its monetary expected value were calculated for both 2022 and 2023. The values obtained are described in Table 18 for the year 2022 and Table 19 for the year 2023. For each combination of thresholds, the ATV Return and the number of items covered are presented.

Table 18 - Scenario analysis for the TB & TS Initiative for the year 2022

		Min Return Rate of TB & TS				
		20%	25%	30%	35%	40%
Min TB & TS Returns	5	\$ 18.7 M	\$ 12.3 M	\$ 7.9 M	\$ 5.6 M	\$ 3.6 M
		1319	893	596	425	294
	10	\$ 6.7 M	\$ 3.9 M	\$ 1.8 M	\$ 1.0 M	\$ 0.4 M
		267	169	90	47	26
	15	\$ 3.7 M	\$ 1.9 M	\$ 0.7 M	\$ 0.4 M	\$ 0.1 M
		92	55	27	16	3
	20	\$ 2.5 M	\$ 1.3 M	\$ 0.2 M	\$ 0.1 M	\$ -
		46	27	10	6	0

Table 19 - Scenario analysis for the TB & TS Initiative for the year 2023

		Min Return Rate of TB & TS				
		20%	25%	30%	35%	40%
Min TB & TS Returns	5	\$ 23.7 M	\$ 18.4 M	\$ 4.7 M	\$7.2 M	\$ 4.6 M
		1319	893	596	425	294
	10	\$ 8.6 M	\$ 5.5 M	\$ 2.3 M	\$1.3 M	\$ 0.6 M
		267	169	90	47	26
	15	\$ 4.6 M	\$ 2.5 M	\$ 0.8 M	\$ 0.5 M	\$ 0.1 M
		92	55	27	16	3
	20	\$ 3.1 M	\$ 1.5 M	\$ 0.2 M	\$ 0.2 M	\$ -
		46	27	10	6	0

As expected, as the thresholds increase, the ATV Return and the number of items covered by the initiative decreases considerably. As for the Min Return Rate, an increase in the threshold from 20% to 25% results in a 33% decrease in the ATV Return and a 32% decrease in the number of items considered to be Items TB & TS. With respect to the Min Return Rate, an increase from 5 to 10 results in a 63% decrease in the ATV Return and a 79% decrease in the number of items covered by the initiative.

5.2.6 Implementation

Finally, the next step was to present the initiative to the Catalogue Team, which is responsible for managing the products on the Farfetch online website. The initiative was later presented to the Farfetch Catalogue team which, at the same time, was developing a similar project. The team presented the initiative, in which they filtered items by the Item Return Rate and % of Reorders in different sizes, that is, they evaluated which items are more often repurchased in different sizes.

After the presentation of the TB & TS initiative, the project was discussed with the Catalogue team and a possible combination of the two initiatives was evaluated since they have similar approaches. Although the Catalogue team's initiative is more effective at detecting items with sizing problems, it does not include one-size-fits-all items as the customer cannot repurchase in a different size. A meeting was held a few weeks later in which it was concluded that it would be the scope of the Catalogue Team to understand how the two Initiatives could be combined.

While waiting for the meeting with the Catalogue Team, the Product Team responsible for the Size and Fit Product descriptions was contacted to see if it was possible to run some tests in order to understand the impact the initiative could have. A list of the 10 most critical TB and TS items was shared with them in order to assess whether they could have their information updated. At the time of the submission of this dissertation, the Product team was still analysing the items provided.

6 Discussion and Future Work

The company's operations must be supported and strengthened by structured and consistent processes in order to maintain a profitable and sustainable growth. As a result, Farfetch must constantly look for ways to optimise the current procedures by redesigning them to be even more automated and cost efficient.

Furthermore, taking into account Farfetch's business model, it is extremely important for the company to control its logistics costs. For this, during the course of this dissertation, the main logistics cost drivers of Shipping Costs and Return Costs were presented. In total, 4 drivers of these costs were identified. The first one is the Service Type, which corresponds to the type of transportation practiced by the Couriers, which, in the case of Farfetch, can be Standard or Express. The next driver identified was the Route taken which corresponds to the route between the Store Country and the Shipping Country. The next driver identified is the Courier used in the following order: DHL, FedEx and UPS. The last driver identified was the Size of the Box which, in the analysis performed, was approximated to the Category of the product sold.

Nevertheless, it is not enough to know which are the cost drivers. It is also important to understand which of these drivers and driver combinations have the greatest impact on the company's logistics costs. Therefore, two different models were applied: one of them was a linear regression to which the VarImp function was applied in order to understand which drivers have a greater influence in the costs; In addition, a Decision Tree model was developed in order to understand which are the main combinations of drivers that affect the logistic cost structure of the company. Regarding the main drivers, obtained through linear regression, these are mostly Routes, that is, it is the Routes that have the greatest impact in defining the Shipping Cost (80% of all drivers) and the Return Cost (100% of all drivers) of an order in Farfetch's business model. A deep dive analysis of the routes highlighted the importance of the United States as both a store and shipping country. When it comes to the combinations of drivers identified, these are more diversified. However, there is again a large impact of the United States country and the DHL courier on the cost structure.

However, something useful for the company is the monitoring of the drivers over time, which means being able to verify if there are any variations on the identified drivers and understand why these occurred. To this end, a dashboard was developed in Looker in order to understand the evolution of the selected drivers. The dashboard developed can compare the values of the year under analysis with those of the previous year in order to understand how they are evolving over a longer period of time. In addition, in the developed dashboard, it is also possible to make a comparison at the last month level. This comparison allows the user to identify cost increases in a shorter time interval and thus, understand in a timely manner the origin of such cost increments.

When it comes to defining optimisation opportunities, despite knowing the main drivers of the logistics costs, it is important to know how the company operates. In other words, it is not by knowing that the United States has a big accountability in the company's logistic cost structure that one can find optimisation opportunities. It is important to ask, how to optimize the drivers identified. Therefore, the analysis carried out aims to help Farfetch in its creation of initiatives but, just by analysing it, it is not possible to identify optimisation opportunities. For this, a deep knowledge of the business and inter-team communication within the company is required in order to find opportunities for improvement. Brainstorming of possible ideas, accompanied by an analytical component is fundamental in identifying chances for development.

In the project, two optimisation opportunities affecting logistics costs were assessed. The first one related to the use of the dashboard and was characterized by the increase in shipping costs verified on the United States-United States route. Therefore, through the use of the dashboard it was possible to identify the core of the problem, which corresponds to the increase in the volume of UPS Courier orders from the partner Stadium Goods. This information was passed to the team responsible for controlling the delivery part of the order processing.

The second opportunity to optimize logistics costs was identified through brainstorm sessions supported by an analytical component. When analysing the reason for the returns, a great part of it was caused by Size and Fitting issues. Subsequently, the "Items TB & TS" initiative was developed with the aim of tackling the products return for being either too big or too small. Its objective was to decrease the return rate value of these items and, consequently, decrease the company's return costs. After defining thresholds, it was possible to identify the most critical items and estimate the monetary impact of the initiative. The initiative was later presented to the Farfetch Catalogue team, which, in parallel, was working on a similar project. The analysis and outputs were provided to the Catalogue Team, who were responsible for understanding how the two initiatives could be combined. In addition to the Catalogue team, the initiative was also discussed with the team responsible for Size and Fit product information. A list of the 10 most critical items in both the TB and TS categories was shared with this team. They had the responsibility of understanding whether or not the identified items represented opportunities in case their descriptions were changed.

As far as future actions of the analysis are concerned, especially for the TB & TS initiative, these are related to making the changes on the website and verifying how the number of returns evolves. Therefore, after having the data on how the number of returns changes, it is advisable to do hypothesis tests in order to verify if they suffered changes in their number of Returns in a significative way and, consequently, were correctly identified as TB or TS.

Furthermore, it would also be interesting to conduct a study of Items TB and TS at the level of different routes. After discussing with the Product team responsible for the Size and Fit information of the product, it was verified that the number of returns of a product varies according to the Store Country and Shipping Country combination. Therefore, it would be extremely important to analyse how certain products behave in certain routes in order to find sizing problems related to the direct conversion of a size from one country to another.

As far as monitoring the drivers, a monthly dashboard monitoring is advised. In the monthly analysis, an observation should be made of the drivers at the YoY level in order to understand how they are evolving over a longer analysis time interval. In addition, an analysis should be also done on their MoM evolution in order to verify variations in a shorter time interval.

To conclude, in addition to the value created through the two optimisation opportunities identified, the Logistics Optimisation team can now benefit from an analytical analysis of the main drivers and combinations of drivers of logistics costs, as well as a dashboard to monitor their evolution over time.

Bibliography

- Baptista, C., & Sousa, M. (2011). Como fazer Investigação, Dissertações, Teses e Relatórios, segundo Bolonha. Em C. Baptista, & M. Sousa, *Como fazer Investigação, Dissertações, Teses e Relatórios, segundo Bolonha* (p. 24). Lisboa: Pactor.
- Bertaglia, P. R. (2009). Logística e Gerenciamento da Cadeia de Abastecimento. Em P. R. Bertaglia, *Logística e Gerenciamento da Cadeia de Abastecimento* (p. 375). São Paulo: Saraiva.
- Beth, S., Burt, D., Copacino, W., Gopal, C., Lee, H., Lynch, R., & Morris, S. (2006). *Supply chain challenges: Building relationships*. Boston: Harvard: Harvard Business Review on Supply Chain Management.
- Blocher, E. (2005). *Cost Management: A Strategic Emphasis*. McGraw-Hill College.
- Botes, F., Jacobs, C., & Pienaar, W. (2006). A model to calculate the cost of logistics at a macro level: a case study of South Africa: research article. *Southern African Business Review*, 1-18.
- Calmanovici, C. E. (2011). A inovação, a competitividade e a projeção mundial das empresas brasileiras. Em C. E. Calmanovici, *A inovação, a competitividade e a projeção mundial das empresas brasileiras* (pp. 190-203). São Paulo: Revista USP.
- Campos-Garcia, R., Garcia-Vidales, M., & Gonzalez-Gomez, O. (2011). Relationship between cost and logistic practices in small and medium enterprises: a case study of Queretaro, Mexico. *African Journal of Business Management*, 1245-1252.
- Carvalho, J. C. (2010). Logística e Gestão. Em J. C. Carvalho, *Logística e Gestão* (p. 23). Lisboa: Edições Sílabo, 1ª edição,.
- Carvalho, J. C. (2012). Logística e Gestão. Em J. C. Carvalho, *Logística e Gestão* (p. 24). Lisboa: Edições Sílabo.
- Carvalho, R. M. (2011). Parcerias - Como criar valor com a internacionalização. Em R. M. Carvalho, *Parcerias - Como criar valor com a internacionalização* (pp. 124-127). Lisboa: Deplano.
- Choudhary, D., & Shankar, R. (2013). Joint decision of procurement lot-size, supplier selection, and carrier selection. *Journal of Purchasing & Supply Management*, 16-26.
- Cosimato, & Troisi. (2015). Green supply chain management: Practices and tools for logistics competitiveness. *The TQM Journal*, 256-276.
- Council of Logistics Management. (1 de May de 2016). *Definition of Logistics*. Obtido de <http://www.clm1.org>.
- Eckerson, W. (2006). *Performance dashboards Measuring, monitoring and managing your business*. John Wiley & Sons. Inc.: New Jersey.
- Engblom, J., Solakivi, T., Toyli, J., & Ojala, L. (2012). Multiple-method analysis of logistics costs. *International Journal Production Economics*, 29-35.
- Fang, Y., & Ng, S. (2011). Applying activity-based costing approach for construction logistics cost analysis. *Construction Innovation: Information, Process, Management*, 259-281.
- Farfetch. (2021). *Annual Report*.
- Farfetch. (2022). *Annual Report*.
- Few, S. (2006). *Information dashboard design: The effective visual communication of data*. Italy: O'Reilly Media. Inc.

- Havenga, J. (2010). Logistics costs in South Africa – the case for macroeconomic measurement. *South African Journals of Economics*, 460-478.
- Helfert, E. (2001). *Financial analysis tools and techniques: a guide for managers*. McGraw-Hill.
- Hungarian Association of Logistics. (2009). Activity Based Infrastructure Cost Calculation in Rail Transport. *Yearbook of Logistics 2009*, 63-67.
- Jingping, X. (2017). *Globalization - Trump's protectionism and Chinese globalization clash in speeches at Apec*. RFI, Portuguese ed. Paris.
- Kerzner, J. (2017). *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*. John Wiley & Sons.
- Kubon, M., & Krasnodębski, A. (2010). Logistic costs in competitive strategies of enterprises. *Agricultural Economics*, 397-402.
- Lacombe, F. (2009). Teoria geral da administração. Em F. Lacombe, *Teoria geral da administração* (p. 291). São Paulo: Saraiva.
- Lai, K.-H., Wong, C., & Cheng, T. (2008). A coordination-theoretic investigation of the impact of electronic integration on logistics performance. *Information & Management*, 10-20.
- Langley, C. J., Coyle, J. J., Gibson, B. J., Novack, R. A., & Bardi, E. J. (2008). *Managing supply chain: A*. South-Western Cengage learning.
- Lee, J. Y. (2002). An examination of international differences in adoption and theory development of activity-based costing. *Advances in International Accounting*, 65-77.
- Looker. (s.d.). *Building LookML dashboards*. Obtido de <https://docs.looker.com/dashboards/creating-lookml-dashboards>
- Looker. (s.d.). *Understanding model and view files*. Obtido de <https://docs.looker.com/data-modeling/getting-started/model-development>
- Maisey, S. (23 de May de 2019). *Meet Jose Neves, the man behind one of fashion's quietest revolutions*. Obtido de TheNationalNews: <https://www.thenationalnews.com/lifestyle/meet-jose-neves-the-man-behind-one-of-fashion-s-quietest-revolutions-1.865577>
- Mangan, J., Lalwani, C., & Butcher, C. T. (2008). *Global logistics and supply chain management*. Wiley Higher Education.
- Moura, B. (2006). Logística: conceitos e tendências. Em B. Moura, *Logística: conceitos e tendências* (p. 15). Rio de Janeiro: Centro Atlântico.
- Niven, P. R. (2006). *Balanced Scorecard Step-by-Step: Maximizing Performance and Maintaining Results* by Paul R. Niven. John Wiley and Sons.
- Pauwels, K., Ambler, T., Bruce, H., P., L., D., R., Skiera, B., . . . Wiesel, T. (2009). Dashboards as a service: why, what, how, and what research is needed? *J. Serv. Res.* 12, 89-175.
- Porter, M. (1985). Competitive Advantage. Em M. Porter, *Competitive Advantage* (pp. 33-45).
- Porter, M. (1985). Technology and Competitive Advantage. Em M. Porter, *Technology and Competitive Advantage* (pp. 60-78). Journal of Business Strategy, Vol. 5 Issue: 3.
- Porter, M. (1985). The value chain and competitive advantage. Em M. Porter, *The value chain and competitive advantage* (p. 45).
- Porter, M. (1989). The value chain and competitive advantage. Em M. Porter, *The value chain and competitive advantage* (p. 34).

- Princeton University. (2007). *Interpreting Regression Output*. Obtido de Princeton University Library: https://dss.princeton.edu/online_help/analysis/interpreting_regression.htm
- Pubrica. (8 de August de 2019). *Why is it important to do a literature review in research?* Obtido de Pubrica: <https://pubrica.com/academy/research/why-is-it-important-to-do-a-literature-review-in-research/>
- Qureshi, M., Kumar, D., & Kumar, P. (2008). An integrated model to identify and classify the key criteria and their role in the assessment of 3PL services providers. *Asia Pacific Journal of Marketing and Logistics*, 227-249.
- Reid, Stuart, Miller, G., Prouse, E., & Hofmann, W. (2016). Omnichannel Retail Survey 2016. *KPMG*.
- Schniederjans, M. J. (1997). Using the Analytic Hierarchy Process and multiobjective programming for the selection of cost drivers in activity-based costing. *European Journal of Operational Research*, 72-80.
- Song, H., & Wang, L. (2009). The status and development of logistics cost management: evidence from Mainland China. *Benchmarking: An International Journal*, 657-670.
- Vanderbeck, E. J. (2012). *Principles of Cost Accounting*. Cengage Learning.
- Yigitbasioglu, O., & Velcu, O. (2012). The use of dashboards in performance Management: Evidence from Sales Managers. *International Journal of Digital Accounting Research*, 39-58.
- Zhao, L., & Huang, X. (2009). Research on the application of business intelligence in logistics management. *International Conference on Management and Service Science*, 1-4.
- Zou, Y. (2011). Study on logistics operation cost control Based on the DEA model. *International Conference on Management Science and Industrial Engineering*, 1025-1028.

APPENDIX A: Gantt Chart

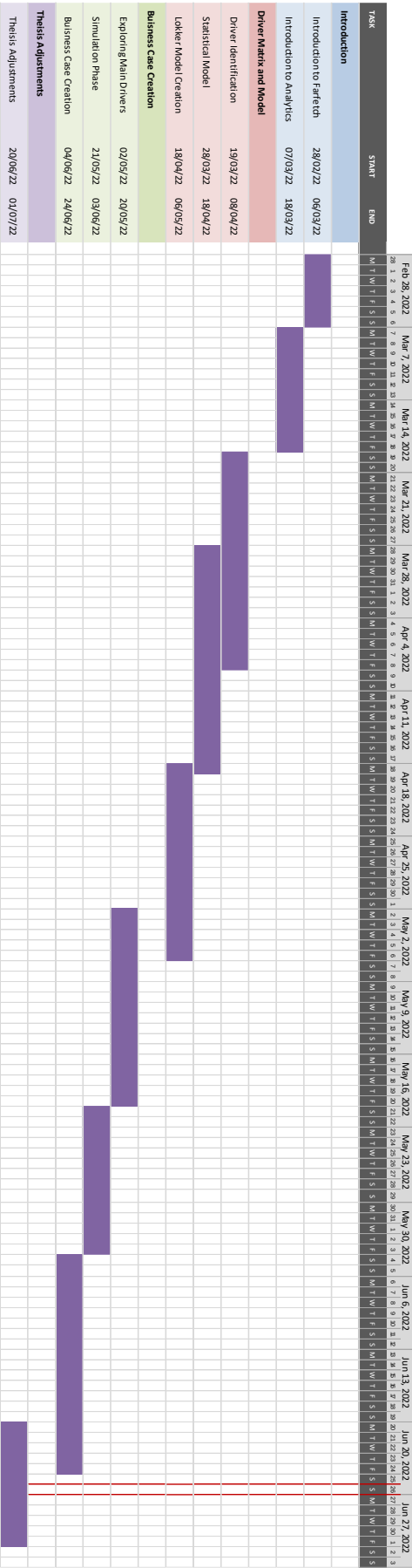


Figure A.1 - Project Gantt Chart

APPENDIX B: Route Driver Analysis Country Aggregation Map

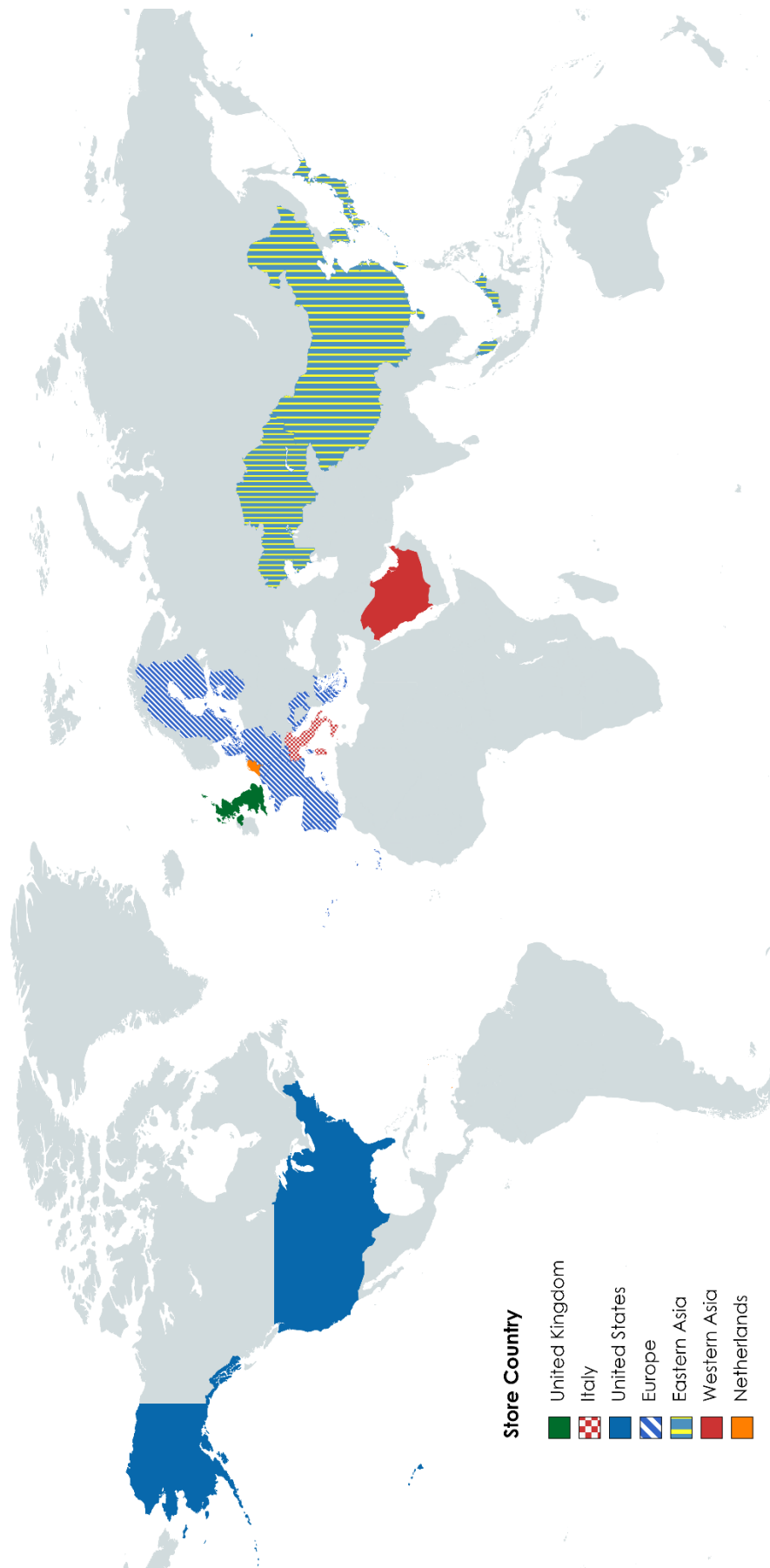


Figure B.1 – Store Country Aggregation Map

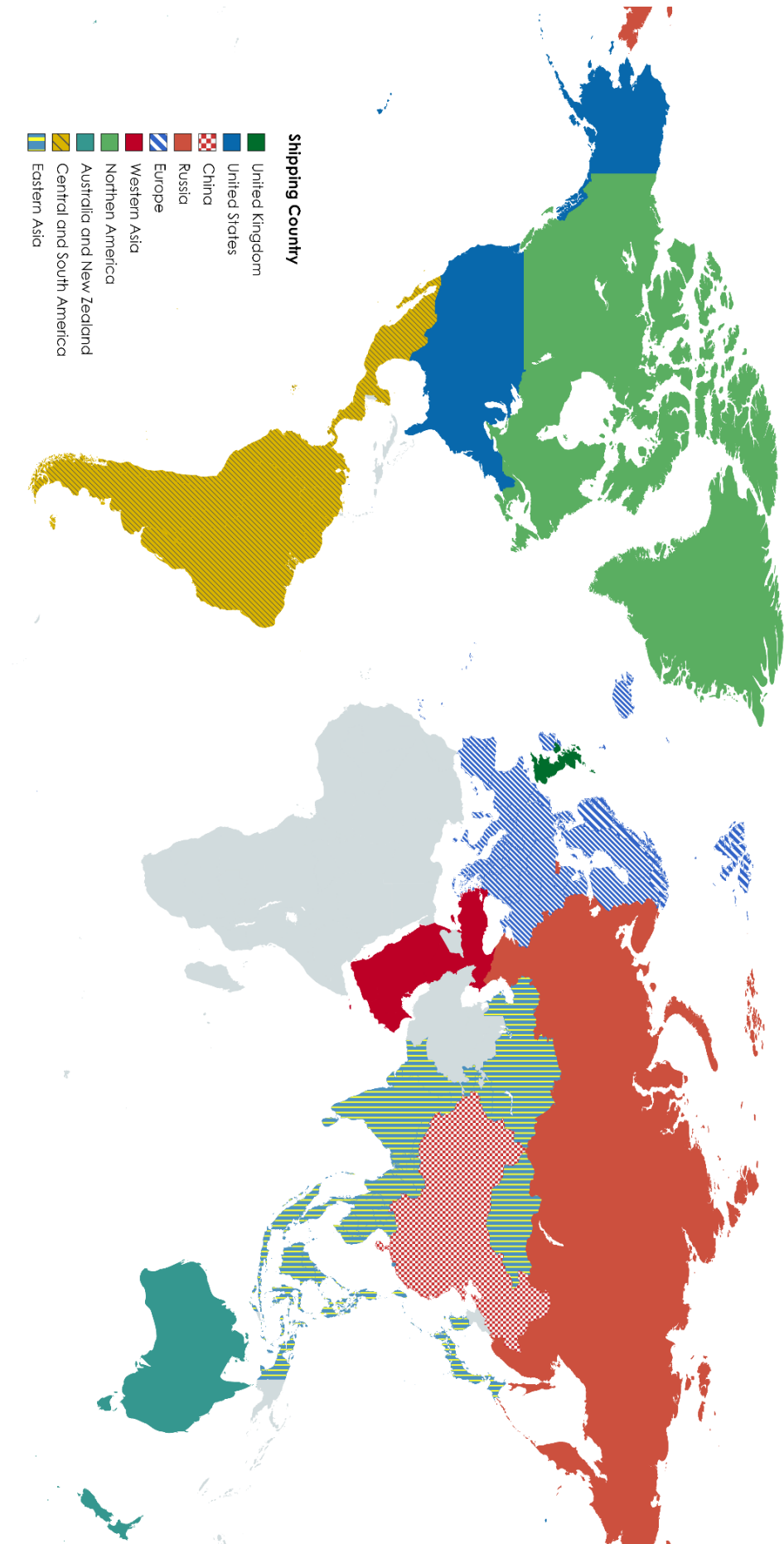


Figure B.2 – Shipping Country Aggregation Map

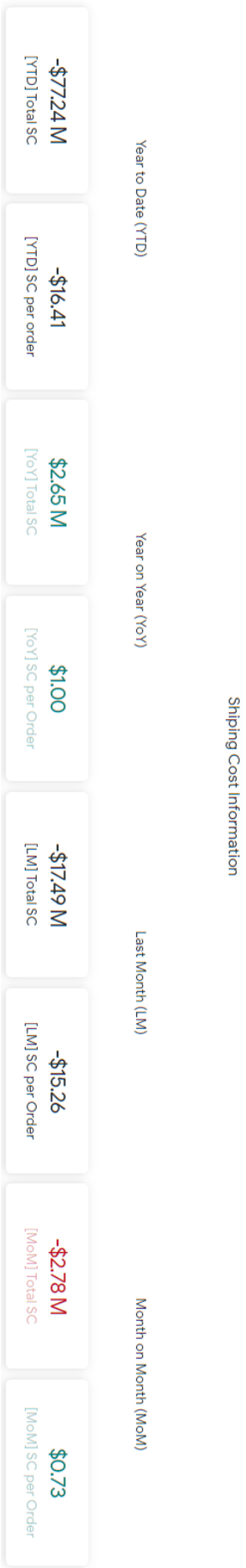


Figure C.2 – Section One of Looker Dashboard

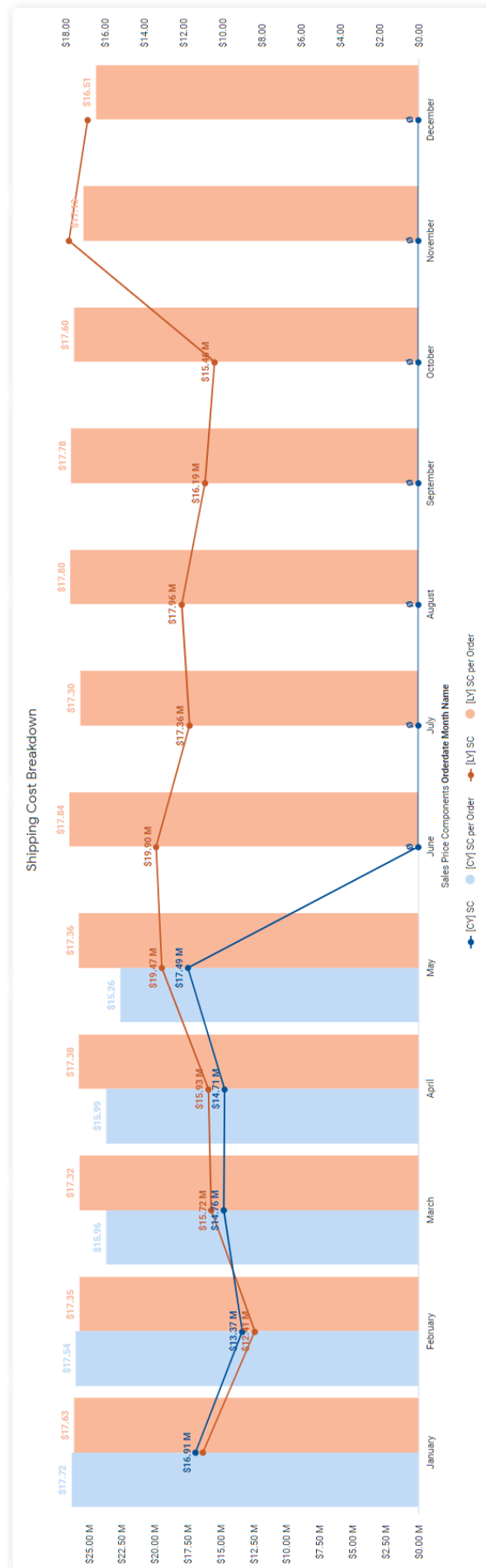


Figure C.3 – Section Two of Looker Dashboard

Driver: Route

It should be noted that Italy, the United Kingdom, the Netherlands are not aggregated in Europe when referring to the Store Country, being individually broken down. In addition, the United States is also not aggregated in Northern America, China is also not aggregated in East Asia when it comes to Shipping Country. Finally, Russian and United Kingdom are not aggregated in Europe when it comes to Shipping Country

Italy - Europe				United Kingdom - Western Asia			
Main Route				Main Route SC			
Route				Route			
Route	[YTD] SC	[YTD] % of Total SC	[YTD] SC Yoy	[YTD] SC per Order	[YTD] SC per Order Yoy	[LM] SC	[LM] % of Total SC
Italy - Europe	-\$5.88 M	7.62%	-\$120.3 k	-\$12.02	-\$0.52	-\$151 M	8.64%
Italy - United States	-\$4.87 M	6.31%	\$620.5 k	-\$14.88	\$1.70	-\$102 M	5.86%
Italy - Eastern Asia	-\$4.52 M	5.85%	\$49.9 k	-\$18.03	\$0.83	-\$115 M	6.60%
United States - United States	-\$4.40 M	5.70%	-\$1754.0 k	-\$12.85	-\$0.11	-\$0.89 M	5.07%
Italy - Western Asia	-\$4.10 M	5.30%	-\$347.7 k	-\$21.51	\$1.05	-\$0.90 M	5.17%
Other	-\$3.51 M	4.54%	\$1983.8 k	-\$17.92	\$2.26	-\$0.60 M	3.42%
Europe - Europe	-\$2.99 M	3.87%	\$468.1 k	-\$12.25	\$0.87	-\$0.66 M	3.79%
Italy - China	-\$2.63 M	3.41%	\$628.3 k	-\$19.78	\$1.87	-\$0.60 M	3.41%
Europe - Eastern Asia	-\$2.45 M	3.18%	-\$340.9 k	-\$18.46	\$0.07	-\$0.53 M	3.02%
Europe - United States	-\$2.14 M	2.77%	\$272.8 k	-\$14.84	\$3.27	-\$0.48 M	2.73%
United Kingdom - United States	-\$1.95 M	2.52%	\$1754.9 k	-\$21.10	-\$0.24	-\$0.50 M	2.84%
Europe - Western Asia	-\$1.94 M	2.51%	-\$201.6 k	-\$21.53	\$0.63	-\$0.42 M	2.38%
Italy - Central and South America	-\$1.79 M	2.32%	-\$474.5 k	-\$20.43	\$0.38	-\$0.46 M	2.64%
Netherlands - Eastern Asia	-\$1.70 M	2.20%	-\$1541.1 k	-\$20.41	-\$0.77	-\$0.53 M	3.06%
Netherlands - China	-\$1.62 M	2.09%	-\$1533.8 k	-\$19.56	\$1.43	-\$0.55 M	3.13%
Netherlands - Europe	-\$1.56 M	2.03%	-\$1366.1 k	-\$12.26	\$3.35	-\$0.41 M	2.37%
Italy - Russia	-\$1.42 M	1.84%	\$321.0 k	-\$19.22	\$1.18	-\$0.25 M	1.45%
Italy - Russia	-\$1.40 M	1.82%	\$205.4 k	-\$23.24	-\$0.50	0	0

Figure C.4 – Section Three of Looker Dashboard