

Analysis of Fashion E-commerce Data to Promote Process Strategical Improvements

Marta Melo Coelho Ferreira

Master Dissertation

FEUP Supervisor: Prof. Mário Amorim Lopes



Integrated Master in Mechanical Engineering

2018-01-11

“Be Human, Be Brilliant, Be Revolutionary, Think Global, Todos Juntos.” Farfetch’s value

Abstract

Digital commerce is one of the most addressed topics in luxury businesses. The innovation of this sector can and must be a path to obtain high profits. Farfetch, which is currently a worldwide success, was created upon the realization of the rising interest and fascination regarding electronic commerce.

Nowadays, with an increasing competitive market in this sector, it becomes essential to adopt strategies to value companies in relation to their competitors. In the case of large companies, which are responsible for a huge amount of data, it is interesting to convert these in useful knowledge to the business and adopt strategies to improve processes.

The present dissertation was developed in the centre of digital production of an online luxury sale company. After data analysis referring to previous years, it was observed that stock times between the different phases of the process were too expensive, that the computer systems had defects with impact on cycle times and that there was a high variability regarding the productivity of collaborators.

The focus of this dissertation was thus to perform a more detailed analysis of all the processes and to suggest strategical improvements of the latter. Two different studies were performed to achieve such. One focused on understanding how the cycle time of the station with higher impact on the process fluctuated, and how it could be reduced. The second aimed to reuse the high stock time to perform activities in parallel, that although obligatory, can be executed at any stage of the process.

Using a multivariate linear regression in the first study, it was understood that adding a new item with a different category to a same slot increases the cycle time of the station in 2.960 minutes, while adding an extra item increases it in 2.899 minutes. It was suggested that the items started to be organized by categories and not by slot, so that each work post would produce two different categories at most. This would necessarily create the need for a triage station to reorganize the items by slot again, and allow them to proceed to the boutiques.

The second study allowed to realise that the stations' stock times can be used to enter the information of the items at the company's software, therefore reducing the process time of the items up until 13%. One of the already existing stations would then have to be divided in two new ones: one that could be executed at any time of the process and another that would necessarily have to be performed after the items' shooting.

Análise de dados de uma empresa de comércio eletrônico de moda para promover melhorias estratégicas nos processos

Resumo

O comércio digital é um dos tópicos mais abordados nos negócios de luxo. A inovação neste setor pode e deve ser um caminho para a obtenção de ganhos elevados. A empresa, que é, até ao momento, um sucesso a nível mundial, foi criada quando se percebeu o crescente interesse e fascínio pelo comércio eletrônico.

Hoje em dia, com o mercado cada vez mais competitivo neste sector, torna-se essencial adotar estratégias que valorizem as empresas em relação à concorrência. No caso de grandes empresas, responsáveis por uma grande quantidade de dados, é interessante converter os mesmo em conhecimentos úteis para o negócio e adotar estratégias para melhorar os processos.

A presente dissertação foi realizada no centro de produção digital de uma empresa de venda de artigos de luxo online. Após análises dos dados referentes a anos anteriores, observou-se que os tempos de stock entre as diferentes fases do processo eram demasiado elevados, os sistemas informáticos possuíam falhas que impactavam os tempos de ciclo e havia uma elevada variabilidade na produtividade dos colaboradores.

Neste sentido, o principal foco da dissertação foi a análise mais detalhada de todos os processos e a sugestão de melhorias estratégicas nos mesmos. Para isso realizaram-se dois estudos diferentes. Um em que o principal foco era perceber como variava o tempo de ciclo da estação com maior impacto no processo e como este poderia ser reduzido. O último estudo passa pelo reaproveitamento dos elevados tempos de stock para realizar, em paralelo, atividades que apesar de terem obrigatoriamente de ser realizadas, poderem ser realizadas em qualquer altura do processo.

Com o primeiro estudo, através de uma regressão linear multivariada, percebeu-se que o aumento de uma categoria contida no mesmo *slot*, aumenta 2.960 minutos o tempo de ciclo da estação, já o aumento de um item aumenta 2.899 minutos. Sugeriu-se que os artigos comesçassem a ser organizados por categorias em vez de por *slot*, de forma a que em cada posto de trabalho sejam produzidas no máximo 2 categorias diferentes. Para tal, seria necessário criar uma estação de triagem para reorganizar os artigos novamente por *slot* para seguirem para as boutiques.

No segundo estudo foi possível perceber que efetivamente o tempo de espera nas estações pode ser aproveitado para realizar a inserção dos dados dos artigos no software da empresa, reduzindo desta forma o tempo de processamento dos artigos em até 13%. Para tal seria necessária a divisão de uma das estações já existentes em duas novas, uma que podia ser realizada em qualquer altura e outra que teria obrigatoriamente de ser realizada no final após os artigos serem fotografados. Criaram-se 3 cenários diferentes em que a principal diferença entre eles eram a percentagem de artigos que eram abrangidos por esta divisão de atividades.

Acknowledgements

"Alone we can do so little; together we can do so much." Helen Keller

First and foremost, I would like to thank Farfetch for the opportunity provided for the execution of this project, for the excellent working conditions that were offered and for the comprehension and flexibility during this period. A sincere thank you to Pedro Rocha, for the patience he had and for being always available when needed, always helping me to discuss my ideas. I would also like to thank all other people that closely accompanied my every day and were always available to help: Marco Teixeira, Renata Miranda, Silvana Mendes, Miguel Gama, Maria Morais and Maria Silva.

I would like to leave an appreciation word to my dissertation colleagues, Ana Rita Ferreira, Rita Fernandes and Joana Sousa. They made everything easier.

To my faculty supervisor, Professor Mário Amorim Lopes, thank you for the attention, guidance, availability and readiness in answering my questions.

To all the professors that contributed to my academic formation.

To my childhood friends, especially to Verónica, Ana, Carolina and Joana, a profound thank you for the support they always demonstrated on this step of my life to be carried out with success.

To Ana Sá, for being a tireless friend with a talent beyond normal and for helping me to write this dissertation. We have been walking together for 18 years, we have been through a lot but always alongside each other.

To my friends, for accompanying me throughout this academic path and rendering it an unforgettable period. A big thank you to Catarina Teles for being an exceptional friend.

To my family, mother, father, brother and sister, for always believing that I was capable and for giving me all the tools and strength to be who I am today. Thank you for everything!

Ultimately, I would like to thank my grandmother Fernanda, who unfortunately departed during this phase of my life, leaving too much longing. She was my role model and I know how much she would be happy to accompany my success.

Contents

1	Introduction	1
1.1	Farfetch Overview	1
1.2	Project Motivation and Goals	4
1.3	Methodology	4
1.4	Dissertation Structure	5
2	Literature review	6
2.1	Process Mapping.....	6
2.1.1	BPMN: Business Process Management Notation	7
2.2	Existing methods for cycle time estimation.....	8
2.2.1	Simulation Methods	8
2.2.2	Statistical Analysis methods.....	8
2.2.3	Analytical Methods	8
2.3	Multivariate Analysis.....	9
2.4	Lean Methodology	11
2.4.1	Push System	12
2.4.2	Pull System.....	12
3	Overview of Current Production Processes	13
3.1	Production Process.....	13
3.1.1	Duplicated items.....	17
3.1.2	Repetitions.....	18
3.2	Data collection	18
3.3	Push System.....	23
3.4	Problems Observed in the Current Process.....	24
4	Case Study.....	25
4.1	Decrease of Cycle Time in Live Model Women Station.....	25
4.1.1	Multivariate Regression with 2 Independent Variables	26
4.1.2	Data Preparation.....	27
4.1.3	Verification of the Assumptions and Realization of Regression	28
4.1.4	Analysis of Results.....	38
4.2	Stock Time Improvements.....	40
4.2.1	“Split Scan Out” Test.....	41
4.2.2	Data-Driven Insights and Opportunities to Improve.....	44
4.2.3	Comparison between the Different Scenarios Studied.....	49
5	Conclusions and Future Work.....	50
5.1	Conclusions.....	50
5.2	Future Work.....	51
6	References	53
	APPENDIX A: Number of items produced per hour in workstations.....	55

APPENDIX B: Process Maps of different workstations	56
APPENDIX C: “Split Scan Out” Test Excel Sheet for each study day	63
APPENDIX D: Tasks that add value or not to Data Entry and Packing	64
APPENDIX E: Spreadsheet with the values obtained in the test performed.....	65
APPENDIX F: Excel Sheet with analysis of stock time in Stills and Data Entry times	66
APPENDIX G: Excel Sheet with analysis of stock time in Live Model Women and Data Entry times	67
APPENDIX H: Excel Sheet with analysis of stock time in Live Model Men and Data Entry times.....	68
APPENDIX I: Excel Sheet with analysis of stock time in Flat Kids and Data Entry times.....	69
APPENDIX J: Spreadsheet with the lead times and different scenarios	70

Nomenclature

BN: Billion

BPMN: *Business Process Model and Notation*

COO: Cook's Distance

FL: Flat

IN: Scan In

JIT: Just-In-Time

KPI: Key Performance Indicators

LEV: Leverage Value

LM: Live Model

OUT: Scan Out

PK: Packing

PRE: Unstandardized Predicted Value

SDR: Studentized Deleted Residuals

SQL: Structured Query Language

SRE: Studentized Residuals

ST: Stills

UML: *Unified Modelling Language*

WIP: Work in Process

List of Figures

Figure 1 - Overview of the distribution of Farfetch photography studios	2
Figure 2 - General process of the company	2
Figure 3 - Scheme with the different stages of production of Farfetch	3
Figure 4 - Standardized notation for process diagrams	7
Figure 5 - Linear relations examples (LaerdStatistics 2015)	10
Figure 6 - Different types of homoscedasticity (LaerdStatistics 2015)	10
Figure 7 - Example of outlier points (LaerdStatistics 2015).....	11
Figure 8 - Production process of the company according to the category of the items	13
Figure 9 - Example of a photo produced in Live Model studios	15
Figure 10 - Example of a photo produced in Flat studios.....	15
Figure 11 - Example of a photo produced in Stills studios.....	16
Figure 12 - Percentage of produced items and duplicates in 2017	17
Figure 13 - Percentage of repetitions according to the respective motif in 2017	18
Figure 14 - Company databases	19
Figure 15 - Percentage of number of items produced by category	19
Figure 16 - Average period in days that an item of clothing takes to return to the boutique ...	20
Figure 17 - Cycle time of the different workstations in minutes	21
Figure 18 - Number of items produced per hour in IN on a given day.....	21
Figure 19 - Number of items produced per hour in LMW on a given day	22
Figure 20 - Number of Items produced per hour in ST on a given day	22
Figure 21 - Stock time at the main workstations in hours	22
Figure 22 - Classification of the variables in SPSS software.	27
Figure 23 - Data view in SPSS	28
Figure 24 - Auxiliary variables created in the multivariate regression.....	29
Figure 25 - Partial regression between independent variables and dependent variable	30
Figure 26 - ScatterPlot of the standardized residues and the standardized predicted values....	31
Figure 27 - Data sorted in descending order of SDR_1	33
Figure 28 - Data sorted in descending order of LEV_1	34
Figure 29 - Data sorted in descending order of COO_1	34
Figure 30 - Histogram of standardized residues	35
Figure 31 - Normal P-P plot of regression standardized residual.....	35
Figure 32 - ScatterPlot of SRE and PRE of the new regression	36
Figure 33 - New partial regression between independent variables and dependent variable ...	37
Figure 34 - Histogram and P-P Plot of the new regression.....	37

Figure 35 - Scan Out process mapping.....	41
Figure 36 - Layouts created for the test “Split Scan Out” [m].....	42
Figure 37 - Scan Out layout.....	45

List of Tables

Table 1 - General data about slots produced in 2017.....	23
Table 2 - Model summary: Durbin-Watson test	30
Table 3 - Test of homogeneity of variance (Levene test)	32
Table 4 - Pearson correlation	32
Table 5 - Collinearity statistics	32
Table 6 - Model summary: Durbin-Watson test of the new regression.....	37
Table 7 - Collinearity statistics of the new regression.....	37
Table 8 - Test of homogeneity of variance of the new regression.....	38
Table 9 - Regression summary table.....	38
Table 10 - ANOVA table.....	39
Table 11 - Cycle time of the 3 slots with different characteristics	40
Table 12 - Verification of the assumptions for the “Split Scan Out” test.....	42
Table 13 - Summary table with the values reached by season and day	43
Table 14 - Calculated cycle times in the test by gender and category of items	44
Table 15 - Calculated data for the current scenario	46
Table 16 - Data Entry table of items that go directly to ST	46
Table 17 - Table of Data Entry and Packing data in the first scenario	47
Table 18 - Calculated data for Data Entry of items that go past LM station	47
Table 19 - Table of Data Entry and Packing data in the second scenario	48
Table 20 - Table of Data Entry and Packing data in the third scenario	49
Table 21 - Comparison between scenarios	49

1 Introduction

The recent technology developments have been so significant that the online market is one of the most leveraged by luxury brands (Remy, *et al.* 2015). The accurate exploitation of the luxury online market will bring high profits to the sector, since the competition is restricted and the growth perspectives are encouraging. This growth has been noticeable in the year of 2015, when online sales of luxury items reached an amount of 23 Bn euros. However, the trend remains, and the forecasts point to an increase in sales up to 70 Bn of euros in 2025 (Remy, *et al.* 2015). A recent study elaborated by McKinsey explains that about 40% of luxury sales are somehow influenced by the clients' digital experience and that is why all online luxury trade companies have as their main objective to surprise customers (Dauriz, *et al.* 2014).

Nowadays, the increasingly competitive market makes it important to adopt management strategies that allow companies to differentiate themselves above the competition. In the case of large companies, responsible for a great amount of data that grows exponentially, it is necessary to often critically considerate all this information, converting it to useful knowledge for the business. This will consequently lead to the possibility of adopting strategies that allow not only to attract new clients, but also to achieve customer retention, while also improving the companies' processes.

This dissertation, developed in the department of process engineering of the Digital Production Centre of Farfetch in Guimarães, focuses on an analysis and critical consideration of data regarding the production processes of the company, while aiming to optimize these processes with the final goal to provide improvements which are able to reduce lead times.

1.1 Farfetch Overview

Farfetch is a global platform of fashion electronic commerce that sells luxury products of boutiques and brands from all over the world. Farfetch founded in London in 2008, since then the company has experienced an exponential grows, with offices in 11 cities worldwide. In the beginning of 2015, Farfetch was valued in 1 billion dollars.

The CEO José Neves started by creating Farfetch.com but the company now comprises more diversified projects that add value to this start up (e.g. Browns Boutique, Black&White, the Mobile Creative Operations Project and the Project of the Future). Nowadays, the company serves more than 510 boutiques with 340 000 different items distributed over 190 countries. It gathers 58 studios spread worldwide to carry out the digital production of its items, as it can be seen in Figure 1.



Figure 1 - Overview of the distribution of Farfetch photography studios

The company is a B2C online platform that bases its business on the online sale of items from different boutiques, gaining a commission for each transaction. The most distinctive feature of Farfetch is that it does not hold stocks i.e.: the company serves as only an intermediary between the boutiques or brands and the final consumer, since the distribution is performed directly from the boutique. One of the advantages of this business model is that it benefits from economies of scale and does not carry the direct logistical burden of holding stock. However, the supply through multiple boutiques can be complex and it requires a flexible and efficient organization. It also depends on the level of satisfaction of its clients, thus one of the focus of the company is precisely the customer service.

At the heart of the company's vision is what retail strategists call 'omnichannel,' the integration of online and offline commerce. In Figure 2, we can see the normal trajectory of an item since it leaves the boutique until it is commercialized.

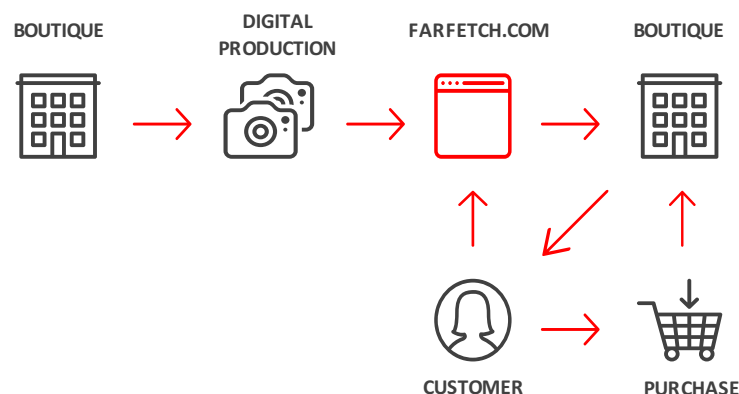


Figure 2 - General process of the company

Farfetch is divided in several departments, among which Account Management, Business Development, Customer Service, Finance, Human Resources, Marketing, Merchandising, Operations, Office Operations, Partner Services, Production and Technology.

This dissertation project was carried out in the Faculty of Engineering of University of Porto and in Avepark – Gandra industrial park. This office represents the main centre of Farfetch's digital production, through which 90% of the items commercialized on the website pass through. The digital production is divided in sub departments such as Logistics, Production, Post Production, Planning, Project Management and Process Engineering.

Each item must be digitally produced and go through different stages according to its category before being commercialized on the website. Different pictures of the product are taken according to its type, and the details regarding the item are displayed so that the client knows exactly what he is purchasing; these include the price, designer ID, size, washing instructions, among others. All photos go through a post-production department, where they are edited, and their quality is verified so that they can be displayed online with the adequate description. The process regarding the different stations by which an item goes through is depicted in Figure 3.

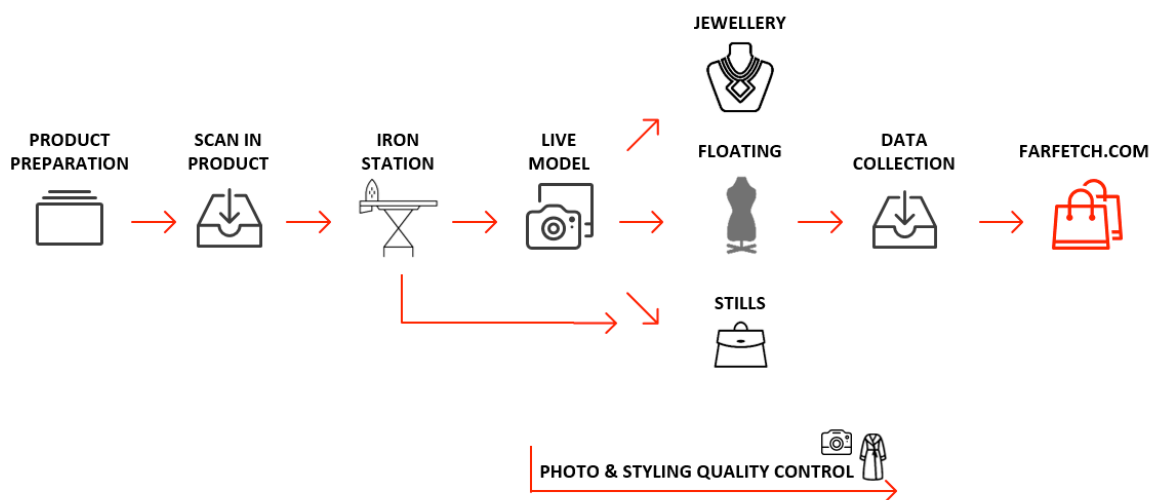


Figure 3 - Scheme with the different stages of production of Farfetch

The Production Centre in Guimarães receives a daily average request of 3000 items daily, distributed in different slots. The strong growth observed in the last years prevented the capability to provide a quick and efficient response to every boutique's request to commercialize products on the website. In order to keep up with this growth and due to the difficulty of receiving such a high number of items every day, the company initiated the construction's project of a new production building.

After a more detailed analysis of all the production process and of the data retrieved from the production of thousands of items in the previous months, the company realized that the time each item remains in the production centre could have been reduced, with the goal of rendering all processes more efficient and improve customer satisfaction. Each product stayed too long in stock areas. The reduction of this lead time will allow an earlier item to be released to sale on Farfetch's website and it will also allow the products to return to the partner boutique to be exposed and sold quicker in the store.

The study of return slot¹ and go online² time reduction is therefore of extreme importance for Farfetch and its partners, since it can lead to significant profit increases for both entities.

¹ "Return Slot" Time – Time it takes to a slot return to the boutique that sent it. Does not include post-production time.

1.2 Project Motivation and Goals

As a company focused on the continuous improvement of its processes and on allowing its clients to experience the best online luxury sale experience, there is a need for constant change and adaptation to the current market needs. It was then understood that the earlier online placement of products, relatively to what takes place currently, can bring significant business benefits to Farfetch.

The main goal of this project is to analyse statistically the data regarding the current production, to ascertain the time that each slot spends in the production centre, how much is spent in each phase of the process and how much time the slot remains in stock. It is important to realize, according to the item's category, where is the bottleneck of the process. It is also vital to understand in detail each phase of the production process and what may lead to an increase of the time cycle, as it is the case of duplicate products, items to reshoot or even items that are defective. For this purpose, a mapping of the company's processes was carried out. After this assessment, it is necessary to develop improvements that can be easily implemented, allowing the fast and efficient placements of products in the online market.

1.3 Methodology

Considering the goals presented in the previous subchapter, a methodical approach was implemented. The objective of this methodology is to create the opportunity to identify the process problematic areas and the chance to improve.

The first step of any project is to clearly understand its context. Accordingly, the first phase involves the interpretation of each area addressed in this dissertation. Defining the processes, using process mapping techniques, is a good way to get a clearer view of the crucial activities. However, it is important to emphasize that the objective is to get the big picture of the workflow, not redesigning the processes. Also, having a representation of the main activities eases communication with specialists and managers of a certain area and enables the identification of problematic nodes. Assessing and analysing current Farfetch KPIs (cycle times, percentage of items produced in different workstations, variability of items produced per working hours, stock times, etc.), targets and other points of reference is also a task to be addressed in this first phase.

The analysis of the obtained data led to the necessity of promoting strategical improvements regarding the production processes of the company. In order to do so, a statistical analysis method was applied – multivariate linear regression – to estimate the cycle time of the stations with higher impact on the process, and comprehend how it varies taking different variables into account. After validating all the assumptions, the obtained results were analysed with the intention of improving this station and increasing its efficiency.

At last, to reuse the stock times of the different stations, the number of resource and workstations were calculated, regarding the necessity of executing a specific activity that could be performed at any stage of the process and which was chosen to be performed in unexploited periods between stations.

² “Go Online” Time – Time an item takes to be placed online on the online platform to be sold. Includes editing time and descriptions.

1.4 Dissertation Structure

The structure of this dissertation is divided in 6 chapters, developed in the document that will be explained briefly:

Chapter 1	Introductory chapter that contextualizes the project and the defined objectives, as well as the methodology and structure used in the Master Dissertation in question. It briefly describes Farfetch.
Chapter 2	Chapter destined to the revision of theoretical concepts related to the methodologies and tools to be implemented.
Chapter 3	Chapter designated to the characterization of the initial situation regarding data collection and their respective analysis, in order to find possible improvements that can be implemented.
Chapter 4	Chapter that presents two different solutions and their respective analyses, performed with the objective of optimizing processes and resolve possible existing issues.
Chapter 5	Conclusions of the Master Dissertation are presented. New projects and future works are also suggested in this chapter aiming for the continuity of the work.

2 Literature review

A brief literature review of the concepts and methodologies used for the development of this dissertation is presented on the current chapter. The theme of process mapping and an notation type used to elaborate process maps (BPMN) are presented. Methods to estimate cycle time are also addressed, as well as the statistical method used in the dissertation – multivariate regression. Lean methodology is briefly reported at last, used for the analysis of the current situation of the company.

2.1 Process Mapping

Traditionally, a process is defined as a series of actions, alterations or functions that culminated in an end or a specific result (Mifflin, 1975). This definition has gaps when the processes present as a goal the improvement of the performed tasks. Due to that, Wesner, Hiatt e Trimble (1994) defined a process as a group of one or more tasks that transform a series of inputs to outputs specific to other person or process, through the combination of people, procedures and tools. Both definitions allow a better understanding of what processes are.

According to Pojasek (2014), many companies do not fully assess their processes and how they interact. These often are overlapping, present inefficiencies which can be easily corrected when well documented, giving rise to more efficient, optimized and structured processes. All companies must have as a fundamental objective the improvement of their processes, therefore benefiting the consumers and all the company's hierarchies.

The process mapping, defined by the understanding and representation of processes in flowcharts, presents as the most practical and simple tool to identify inefficiencies and promote improvements (White and Miers, 2008). Depending on the interest, a map of processes can focus on a whole organization, a business unit, a specific function or even on an employee. Normally, the map of processes shows when, where and what will be done, and who will do it (Marrelli, 2005).

The distinction between different types of process maps relates essentially to the level of detail the diagrams need to present. The most informal ones allow simply the managers to have a better perception of processes, with a limited level of detail. On the contrary, the most formal present a high complexity and they aim to give an overview of all activities in a thorough manner (Harmon 2010).

The process mapping consists of the identification of all the different process steps, either they add value or not (Muda), eventual operational losses, and input/output parameters. Therefore, after the characterization of these parameters it is possible to gather data and identify the sources of variation in the process (Feld, 2000).

The *Business Process Model and Notation*, BPMN, and the *Unified Modelling Language*, UML, are graphic annotations normalized to the moulding of processes. The objective of these solutions resides essentially in the facilitation of understanding descriptive diagrams.

2.1.1 BPMN: Business Process Management Notation

The BPMN model provides companies with a greater capacity to improve the assessment of their internal processes' through a graphical mapping. This facilitates the agreement between the several departments of the company, allowing a quicker adaption to the new internal circumstances of the business (Gagné and Ringuette, 2008).

The set of moulding resources of this type of graphic solution allows the simple development through Flowchart, through various graphic elements illustrated in they are used to provide additional information on the process; there are three types of artifacts: Data Object, Group and Annotation.

- **Flow Objects:** the main graphic elements to define the behaviour of certain processes; these can be Events, Activities and Gateways.
- **Connecting Objects:** connect the different flow objects, to support the skeleton of the project. There are three connecting objects: (1) Sequence Flow, (2) Message Flow and the (3) Association.
- **Swimlanes:** provide the capacity of organizing activities in separate categories if necessary, to illustrate the different responsibilities; these can be Pools and Lanes.
- **Artifacts:** used to provide additional information on the process; there are three types of artifacts: (1) Data Object, (2) Group and (3) Annotation.

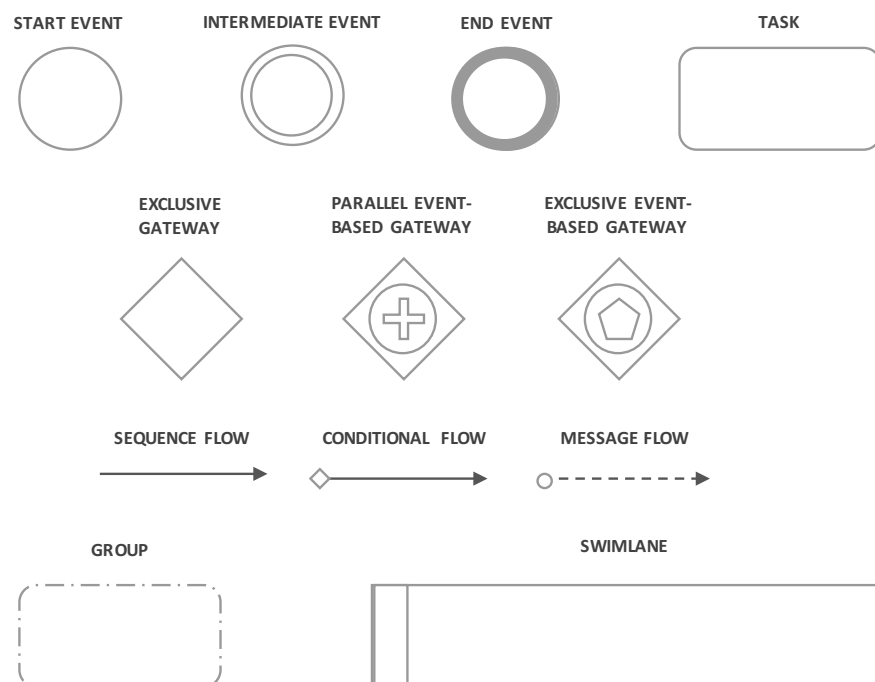


Figure 4 - Standardized notation for process diagrams

2.2 Existing methods for cycle time estimation

Accurate cycle time is an essential planning basis for many diverse applications in production. It has a prominent importance in applications such as due date commitments, analyses of performance metrics, planning and scheduling (Pearn, *et al.* 2009).

During the last few years, there has been an increase in the attention paid to different methods of cycle time estimation, since this is an interesting topic for industrial professionals and researchers (Tai, *et al.* 2012).

Chung and Huang (2002) classified the existing methods to estimate cycle time in four categories: simulation, analytical methods, statistical analysis methods and hybrid methods. The soft computing methods were more recently applied in the estimation of these times, being able to be considered the fifth category (Tai, *et al.* 2012).

The hybrid methods combine different methods to produce a cycle time estimation. (Chung and Huang, 2002)

2.2.1 Simulation Methods

In the present days, simulation is a highly used technique in the exploration, design and optimization of complex systems. It consists on the creation of a model that depicts a dynamic process, with the main goal of withdrawing conclusions and allowing an analysis that may be transported to reality. With the simulation method, it becomes possible to modulate real systems in computer programs, enabling thus the analysis and description of their behaviour when the simulation parameters are altered (Nyhuis, *et al.* 2005).

Simulation has as main advantages the possibility of adapting models, allow the analysis of discrete systems and get a large level of acceptance. It implies, however, considerable effort during the stage of implementation, and it does not enable to obtain general conclusions and validations (Nyhuis, *et al.* 2005).

2.2.2 Statistical Analysis methods

Statistics is applied in many diverse areas of knowledge. The use of different statistical methods allows professionals to be better prepared to organize and analyse the needed data.

A regression analysis or other methods of statistical analysis are applied to determine the relation between cycle time and different parameters related to it. Since the historical conditions are used to make forecasts, the bigger the changes in the systems, the less needed these statistical analysis methods will be (Chung and Huang, 2002).

Backus *et al.* (2006) applied a statistical method, the data-mining approach, and provided nonlinear predictor variables to estimate factory cycle time.

2.2.3 Analytical Methods

Analytical methods are based mainly on the Queuing Theory or on other similar mathematical models to estimate cycle time and its deviation (Chung and Huang, 2002).

The models that use Queuing Theory consist of analytical systems to describe interdependency between indicators of logistic performance. These models consider the

stochastic characteristics of processes that arrive or that are active in a server processing objects in a queue (Nyhuis, *et al.* 2005).

These models have as an advantage the simple application and the fact that its parameters are inferred from fundamental principles. They imply, however, a large effort when defining the system: they are only valid to stable operational stages, its adaptability is reduced and the parameters may not match to the practical reality (Nyhuis, *et al.* 2005).

2.3 Multivariate Analysis

To specify the probability of occurrence of many events of practical interest, it is usually necessary to make probabilistic assertions about several correlated random variables simultaneously.

Multivariate analysis is a statistics branch that dedicates to the study of variables correlated with the variable in question. If two variables are correlated, knowing the behaviour of one provides information on the behaviour of the other. In practice, to describe the incidence probability of certain events it is frequent to consider several random and correlated variables simultaneously.

This analysis consists in performing a simple linear regression and it is used to predict the value of a variable (dependent variable) depending on the value of two or more variables (independent or explicative variables). It is possible to know the adjustment degree of the model to the supplied data, as well as the relative contribution of each variable. This allows to conclude which independent variables explain better the dependent variable's variation to be studied. *SPSS Statistics* allows this analysis to be performed in a simple manner, while supplying a report more or less detailed regarding the performed test (LaerdStatistics, 2015).

A statistical model tends to become more complex as the number of independent variables increases. According to Baur and Lamnek (2007), there are several stages to be followed after data retrieval. These are:

1. **Framing the research question** so that it can be mathematically modulated.
2. **Selection of the most adequate statistical model.** There are several methods for these type of analyses, and new approaches commonly arise. Different models can give rise to different results, which makes the selection of the best model to serve the objectives an essential step of the statistical analysis.
3. **Verify that the assumptions of the model are fulfilled.** The data set needs to follow some assumptions so that the regression can be performed and originate an accurate result.
4. **Preparation of the data for analysis.** A special case of data preparation is data mining, which specializes in extracting variables from complex data banks for statistical analysis.
5. **Model computation** using a specific software of statistical analysis as it is the case of *SPSS*.
6. **Interpretation of the analysis results.** The model may help to interpret data, but it does not prove theories.

When using a multiple regression to analyse data, it is known that a part of the process involves verifying some premises which guarantee that these data can be accurately analysed.

There are 8 premises which must be validated for the regression to provide an accurate result (LaerdStatistics, 2015).

1. **Dependent variables are measured in a continuous scale.** If such variables were to be measured in an ordinal scale, an ordinal regression would have to be performed, instead of a multiple regression.
2. **There are two or more independent variables that can be continuous or categorical.**
3. **There is independency of the observations.** This premise can be verified by performing a Durbin-Watson Statistic test.
4. **There is a linear relation between (a) the dependent variable and each one of the independent variables and between (b) the dependent variable and the independent variables collectively.** Several options exist to ascertain these linear relations; one of them implies visual inspection of the scatterplots and partial regression plots to verify linearity. In the next figure, there are examples of linear and non-linear relations.

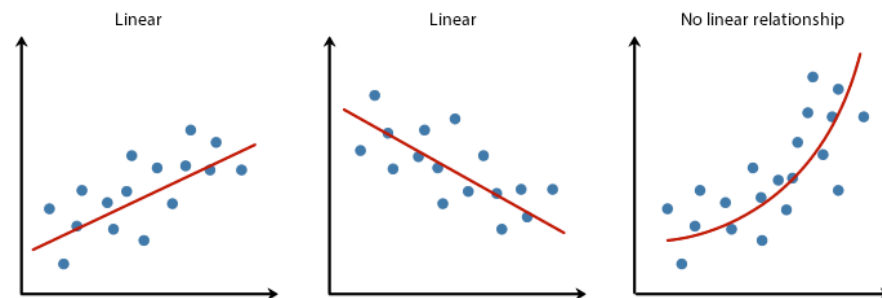


Figure 5 - Linear relations examples (LaerdStatistics 2015)

5. **The data must have homoscedasticity of residues.** Homoscedasticity is a term used to illustrate the constant experimental error variance for distinct observations of the dependent variable. In other words, the error variances maintain a similar pattern along the regression line. In the next figure, it is possible to observe the different types of homoscedasticity.

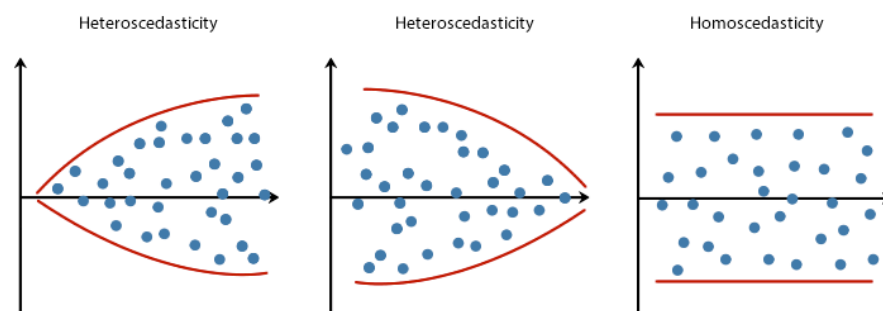


Figure 6 - Different types of homoscedasticity (LaerdStatistics 2015)

6. **The data are not multicollinear.** When two or more variables are highly correlated, problems can arise in understanding which independent variables contribute to the

variation of the dependent variable; technical problems can also arise in the calculation of the multiple regression model.

7. **There should be no significant outliers, high leverage points or highly influential points.** An outlier consists in an observation that is somehow unusual and that can distort the results when obtaining the multiple regression. Its existence may reduce the precision in result prediction, as well as the statistical significance. In Figure 7, two outliers are shown as an example.

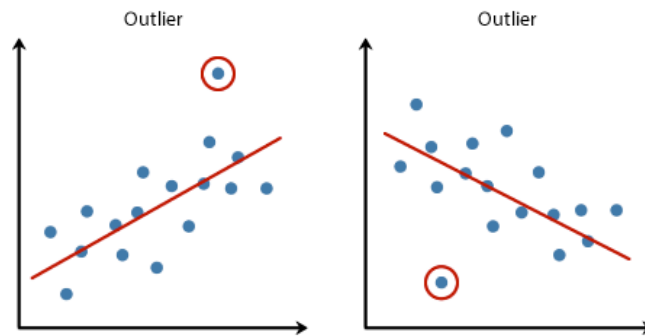


Figure 7 - Example of outlier points (LaerdStatistics 2015)

8. **The residues must be normally distributed.** This premise may be validated through the observation and analysis of a histogram and a normal PP-Plot or normal Q-Q plot of the studentized residuals.

2.4 Lean Methodology

The world is experiencing an industrial revolution that presents as an opportunity for many companies to rethink their strategy and redefine their productive processes. The market pressure to produce increasingly more quantity and more personalized items has led many companies to implement Lean methodologies to solve their problems.

According to Hicks (2007), the Lean concept comprises five fundamental principles:

1. **Identification of the value streams/flows** for each product or product family;
2. **Specific value** through activity specification that generates value in the customer's perspective;
3. **Value generation** by taking actions that contribute to this objective;
4. **Pull Strategy** that occurs when the production is triggered by the request, making use of Just-in-Time³;
5. **Chasing perfection** by continuously removing wastes at every level of the value chain.

One of these methodologies is the implementation of the Pull system, based on the market inputs, to the detriment of the traditional Push system.

³ Just-In-Time: A system of production that makes and delivers just what is needed, just when it is needed, and just in the amount needed.

2.4.1 Push System

The traditional management system Push is based on the release of products as quickly as possible to the next station, to avoid starvation in the upcoming stations (Puchkova, *et al.* 2016).

This is particularly interesting to the suppliers, since it allows them to define the pace and control the production stream, as well as to exert pressure on the market. This stream uses market forecasts and a Work in Process⁴ (WIP) storage.

A Push system may deal with several problems, among which the high inventory costs, few inventory turns, high risk of stock obsolescence and reduced service levels (Puchkova, *et al.* 2016).

2.4.2 Pull System

The Pull principle states that manufacturers should only produce what clients want and when they want it, thus minimizing the intermediary stock. The products arrive to the next station following a JIT methodology and complete the process rapidly (Hicks 2007).

According to Hopp and Spearman (2004), although the improvements at the environmental level are an influence, there are three main reasons to suggest a performance improvement through Pull systems:

1. Less congestion;
2. Easy control of WIP material;
3. Prevention of a WIP explosion.

The organization in Pull Flow promotes an optimization of material and information streams, as well as a decrease in the inventory, since only the necessary amount required by the customer are produced: there is no production for stock.

⁴ Work in Process: the goods that are on the factory floor. The manufacturing of these goods has begun but has not yet been completed.

3 Overview of Current Production Processes

Farfetch had an actual necessity of process optimization and time waste reduction, as previously mentioned. It was necessary to reduce the time that each slot⁵ spends in the production centre and consequently the time necessary for the items to be placed for sale on the website. It was understood that the reduction of these time periods would bring advantages not only to the company but also to the business partners. Coupled to this necessity, a collection of data was gathered regarding all processes and the company data until the beginning of the project.

3.1 Production Process

Each product arrives to Farfetch sent by the boutiques so that the photos are taken, and the required details are retrieved to place it to sell on the company's website. In Figure 8, the productive process is described in more detail.

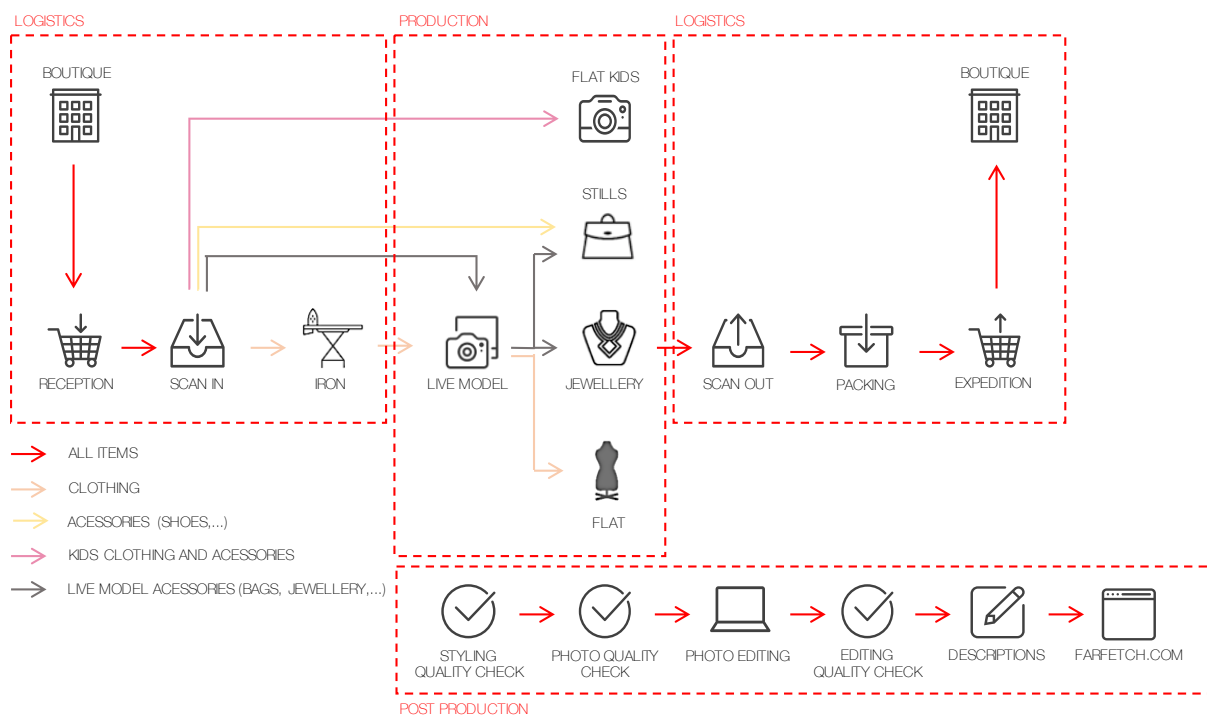


Figure 8 - Production process of the company according to the category of the items

⁵ Slot – box with the items sent by boutiques. It may or may not match only one volume/box. Each slot contains a maximum of 50 items

As depicted in Figure 8, the process of digital production of each item is divided in three sub departments with different functions and teams: Logistics, Production and Post-Production.

The process begins when the boutiques send their products to Farfetch using a transport company. Each boutique can send several slots every day as long as there has been prior planning; however, each slot has a maximum capacity of 50 products. The reception station receives the slots sent by the boutique and sets the respective priorities of production, previously defined by the planning team. After the reception employee sets the priorities in the slots, he opens each one of them to understand if there is any problem; if everything is in order, he places them in the “done zone” organized by priority. Thereafter, the Mizu⁶ takes these slots and distributes them by the different stations in Scan In.

In Scan In, the products and their sheets are withdrawn from the transportation boxes and they are introduced in the system, to ascertain the product type, if it is a duplicated item or an item to reshoot. It is also verified if the product has any defects and if so, the defects are photographed together with the product sheet to prove that these did not occur within the facility of the production centre. After this, the products that are properly packaged, are placed in boxes or rails according to the category and type. These boxes/rails then proceed to a specific stock area to be taken by the Mizu's to the next phase. These transportation systems are adequately identified with a labelling that identifies priority, the boutique that owns the item, the day it arrived and its category and type. If the rail is composed by 100% of duplicates it will not be sent to production; instead, it enters a specific area for duplicates. The accessories that do not go through Live Model or Flat are taken directly to the station of Stills and do not enter the industrial lift with the remaining products.

In the Iron station, the items that need (in the absence of information to the contrary) ironing with professional equipment before entering the station of Live Model and start being shot.

Hereupon, in the Live Model Preparation, the stylist's study the products contained in the rails and create outfits ready to be photographed, either composed of other articles from the same rail or of items from the ToolKit⁷. Still regarding the design of the looks, there are some requirements imposed by the brands or boutiques that must be taken into consideration. After the first outfit preparation, the rails proceed to the studios. In the studios, there are always a photographer, a stylist assistant, the stylist that prepared the rail and one or two models according to the necessary pace. For each product, 3 views are photographed (front, $\frac{3}{4}$ and back) and a more detailed photo is also taken, as illustrated in Figure 10.

All the photos are sent to the Quality Department, so the quality is checked and assured, since sometimes there are photos that do not present the minimum requisites imposed by this department. There are two types of Live Model Studios, which differ in the type of photographed products, model genre and period of cycle stations: Live Model Men studios generally present shorted cycle periods.

⁶ Mizu – Normally defined as Mizusumashi, logistic train, in the company is defined by a collaborator who makes the transport of the items between different workstations.

⁷ ToolKit – Set of basic clothes and other items purchased by the company for the production of outfits.



Figure 9 - Example of a photo produced in Live Model studios

Regarding the category of the item that needs to be produced, some proceed directly to the Flat station, as it happens with clothing items, with the aim of creating a 3-dimensional photo that allows the visualization of the product exactly as it is, including the tag, as shown in Figure 10. For the elaboration of this photo, several materials are used, such as transparent and plastic manikins and sleeves to provide shape to the item.



Figure 10 - Example of a photo produced in Flat studios

There are some accessories (e.g. shoes, glasses, jewellery, amongst others) that do not carry out the Flat photo but that go through the Still station. In this station, 4 different views of an isolated item are photographed (Figure 11). To photograph jewellery, there are two specific Still stations due to the greater details of the product, which leads to a greater difficulty in photo production. The number of registered views in this station depends on the item in question. Before the rail/boxes proceed to the Scan Out stations, it is necessary to validate all the photos taken in the previous stages. This validation is given by the Check-in team of Quality Control. The team needs to be efficient and quick to allow the Scan Out team not to stop while waiting for the quality check. This department rigorously checks the quality of the photos, and when it verifies that one does not fulfil the minimum required requisites, an order is sent to repeat them.



Figure 11 - Example of a photo produced in Stills studios

In Flat Kids station, the products follow the same procedure of the Flat Stations; however, the photos are two dimensional and the photographers use magnets to allow the item to remain completely straightened and symmetrical.

Before the rail/boxes proceed to the Scan Out station it is necessary to validate all the photos taken on the previous steps; meanwhile, they await in an intermediary stock station until all products of the same slot have a positive photographic feedback. This validation is given by the Check In team from Quality Control. The team needs to be efficient so that the Scan Out teams does not stop while it awaits the quality control check. When the box that contains the items already processed is full, it proceeds in the modular structure of rollers (Scan Out line) to go to Packing.

Finally, in the stations of Packing and Expedition, the products are packaged by slot in the boxes from the boutiques, which were stored after the Scan In process, if these are in suitable condition. The reuse of the transportation carton boxes allows the company to save money and time. In the Expedition process, these same boxes are sealed and weighed, receiving also a freight bill. They are then forwarded to a stock area while awaiting the arrival of the transportation company.

After this, the photos are edited by the editing team, where the background and the shadows are corrected, where the Flat photo is assembled and where the faces of the models are removed concerning the Live Model photos. There is a Quality Control team destined to verify the quality of this editing. The product does not have to be still in the facility if a negative feedback is given: it is only necessary to correct the photo editing. The Descriptions station, performed simultaneously with image editing, is responsible for the product description with all the information derived from Scan Out.

3.1.1 Duplicated items

Each partner boutique can send items to be produced, however, it often happens that many of those have already been digitally produced for other associate. Although a previous control to avoid these duplicates exists, they remain a challenge to Farfetch. It is understood by the observation of Figure 12 that in 2017 the percentage of duplicated items was very high, which results in high costs both for the boutiques that pay the transportation fees without any need and for the company.

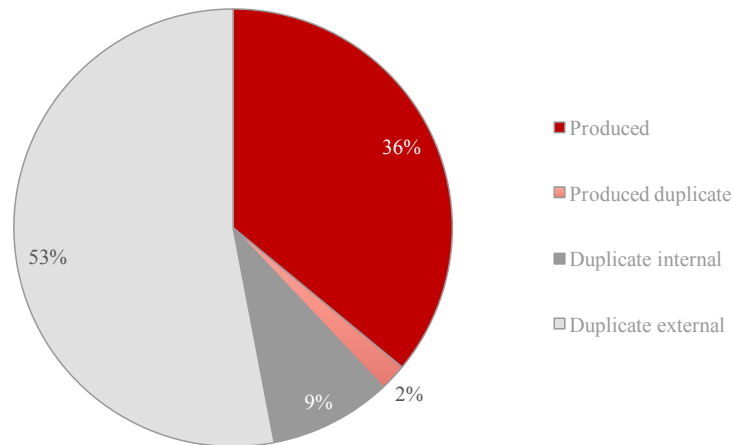


Figure 12 - Percentage of produced items and duplicates in 2017

With the goal of reducing the number of duplicates sent by the boutiques, a team dedicated only to product identification was built. The team is split regarding the stage of the process in which the items were identified.

The main stage consists in the identification of already produced items while they are still in the boutique. The so-called “Pre-Association” is performed by the duplicates team, and these duplicated items are detected even before they reach the production centre, using tools created for matching product’s ID. The Farfetch association is an identification stage of duplicates when these are in the Scan In station; they are not produced but they are already in the physical space of the production centre.

When these items enter digital production they only represent an extra cost that could be avoided, since they have already been produced and they are only occupying stock area. Some duplicates can, however, be used to create outfits if they belong to the same rail of the article to be produced, increasing its usefulness.

There is also a small percentage of products that have already been produced but in which this was not detected; these items are processed again and, in some cases, released once again on the website, something that brings even higher costs to the company.

3.1.2 Repetitions

At Farfetch, there is a station responsible for verifying the quality of the photos taken in Live Model, Stills and Flat. The Quality Control station analyses rigorously all of the photos taken of an item and verifies if they fulfil the minimum requisites to proceed to the editing team for posterior editing and placement on the online platform. Although the Live Model, Stills and Flat teams perform quick verifications of the photos' quality before submitting, there are flaws which are not detected, and which compromise the quality of these. When a defect is detected, the teams order the photos to be repeated.

This team only performs the quality check when at least 94% of the items composing the slot have already been photographed and are already in stock, waiting to proceed to the Scan Out station. Repetitions have therefore a negative impact on the processing time of products.

As an example, when the Quality Control team detects a flaw in a Live Model photo, the quality assistant introduces in the system a new request for a photo and the Mizu must understand where the product is at the moment and transport it to the Live Model station to allow the repetition. All this process can take a significant period of time, even because the Live Model station is located on a different floor from the remaining departments.

During the year of 2017 (until September), it was necessary to repeat one of the photos for about 12% of items due to defects. There are several reasons that lead to a repetition request, either because of a flaw in the photos taken in each station, or because the duplicates were not photographed, or even because the chosen look is not the most adequate for the item in question. Figure 13 displays the percentage of repetitions according to the respective reason.

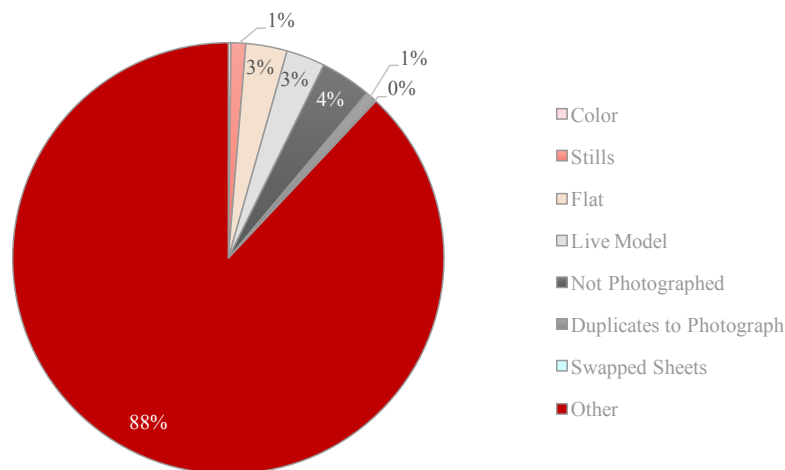


Figure 13 - Percentage of repetitions according to the respective motif in 2017

3.2 Data collection

Aiming to understand the current situation of Farfetch, the production data was collected since January of 2017 until September of the same year, at the beginning of this dissertation project. This diagnosis allowed to identify problems and to understand where to possibly act in the future. Data were retrieved from different databases for this analysis (Figure 14) and field measurements were also necessary in some situations. All the relevant information was then compiled in Excel sheets.

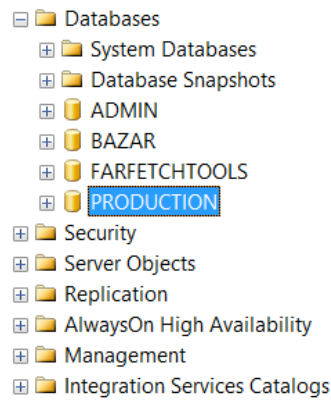


Figure 14 - Company databases

In order to filter all the retrieved information from the Excel databases, lines of code were added to exclude all products produced outside of Europe, all duplicates and every photo repetition due to lack of product. Only the items that followed a production process considered normal were used for the statistical studies.

As the data withdrawn from the databases often presented inconsistencies with uneven values and very high standard deviations, these data were treated to eliminate waste and possible outliers, having as an objective to render them as reliable as possible.

One of the variables with most impact on this project is the relevance of each item category, as it can be observed in Figure 15. Each one of the main categories follows a path along the productive process.

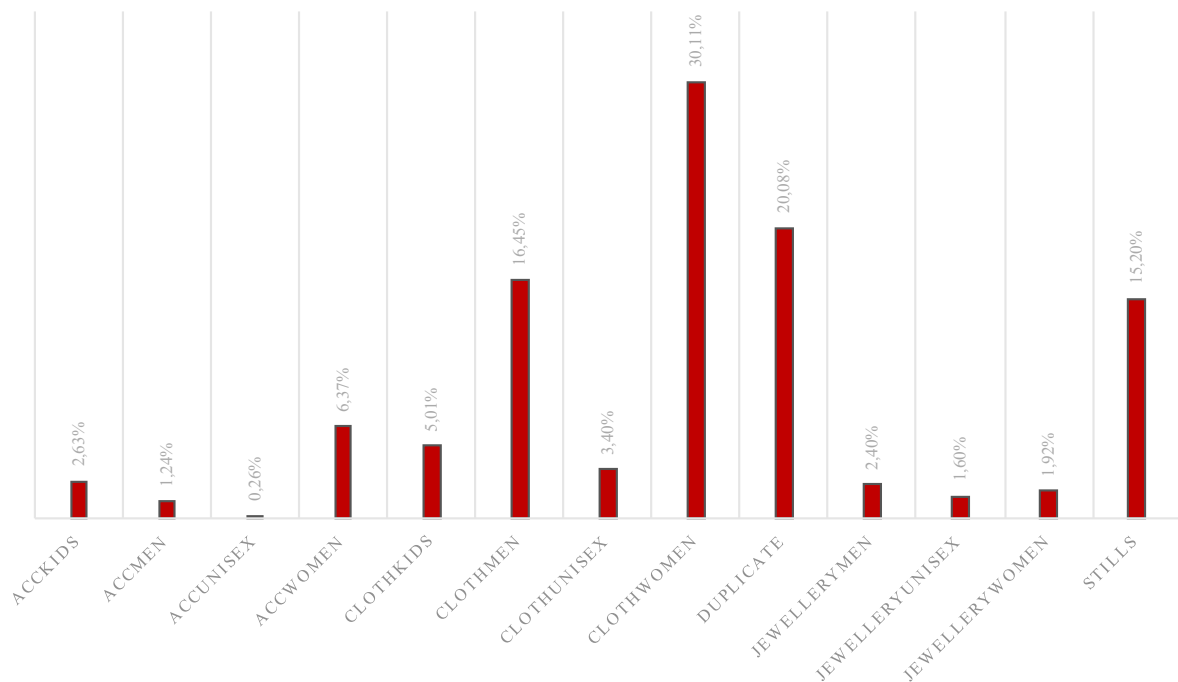


Figure 15 - Percentage of number of items produced by category

When analysing the previous table, it is understood that stations like Live Model Women, Live Model Man, Flat and Stills present highly increased arrival rates of items to be produced when compared to stations such as Flat Kids and Stills Jewellery.

According to their categories, the products proceed to different workstations, so it was then possible to identify four main production lines. It was also understood that despite the categories, there are common stages in the process of digital production, such as Reception, Scan In, Scan Out, Quality Check, Packing and Expedition. It was also seen that all men's and women's clothing go through the Live Model and Flat stations. Regarding accessories, it was realized that some were photographed in Live Model and Stills (Stills or Stills Jewellery), as it is the case of handbags, scarfs and Jewellery. The rest were only photographed in Stills (shoes, key-rings, phone cases and other types of accessories). As it is not feasible to have children as models due to the implied paperwork, the junior products only go through the Flat Kids stations. In line with these observations, four lines of production were highlighted: Clothing, Accessories directly to Stills, Accessories with Live Model and Children products.

After data retrieval and outlier exclusion, it was then possible to calculate the average time a slot took to return to the boutique in the different months of 2017, as well as the time it took for a product to be placed in the online platform, on average and according to the respective line of production. Figure 16 represents the average period that an item of clothing (independently of being masculine or feminine clothing) takes to return to the respective boutique, for each month studied.

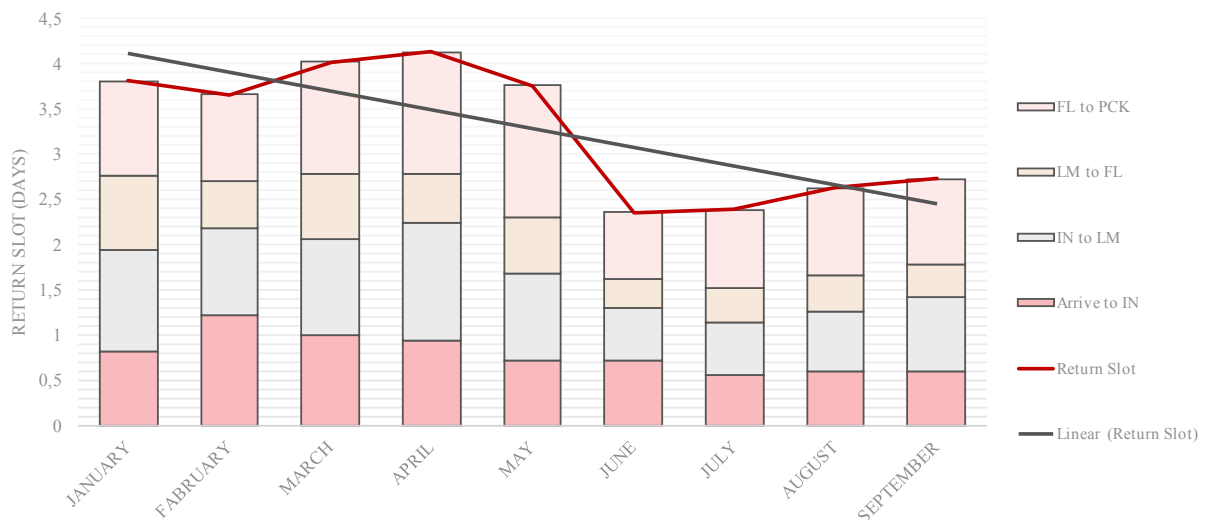


Figure 16 - Average period in days that an item of clothing takes to return to the boutique

The observation of the previous graph allows to perceive a decrease in the period of slot return in the month of June. This decrease may be explained by a test performed between these dates, in which the companies' resources were increased for a higher amount relatively to what is necessary, aiming to understand if this increase would influence the processing times. However, a decrease in the periods of slot return in the year of 2017 is visible, due to improvements implemented in the processes and also to a higher efficacy in production planning.

Another analysis was performed using the daily productivity and efficiency hours of the staff, which were inserted in a programme specific for team, in order to realize the cycle periods for each station. Figure 17 summarizes these periods, in minutes, in all the phases of item production.

The observation of the following table enables to notice that the Jewellery station presents a cycle time considerably superior to the remaining stations. This can be explained by the complexity of the pieces that are photographed here.

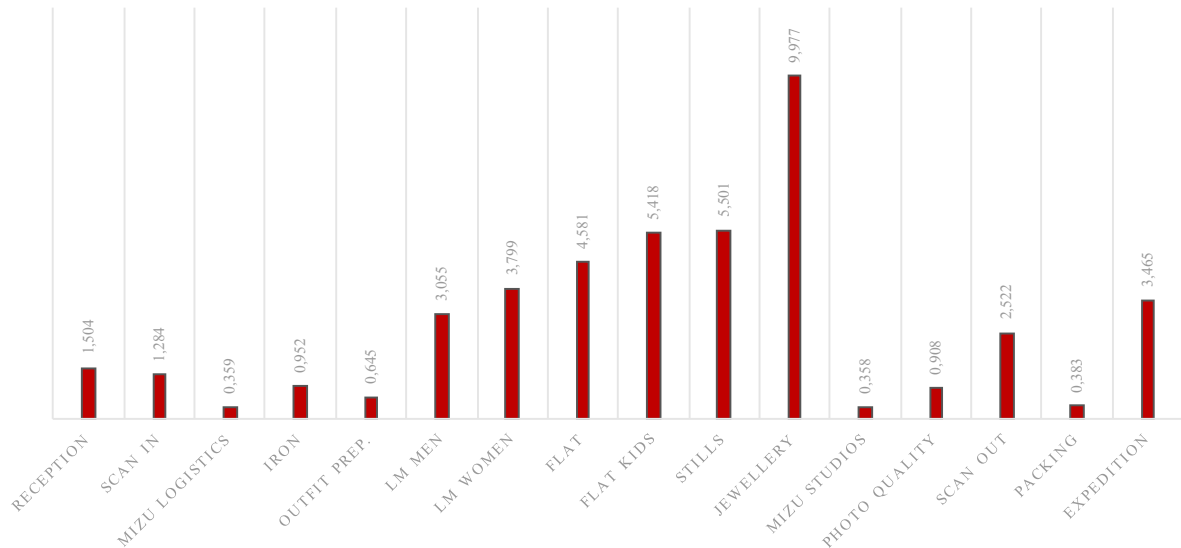


Figure 17 - Cycle time of the different workstations in minutes

It is noteworthy that these time periods present high variability, since there are no machines and the cycle times depend exclusively on the category and detail of the item and on the employee's productivity. People with more training contribute positively to the cycle times in the better part of cases, while younger employees or with less experience can take more time to produce the items, having a negative influence on these periods of each station.

The following figures illustrate the high variability of the items produced during an hour of work in a specific day of production in the stations of Scan In, Live Model Women and Stills. The graphs concerning the variability of the remaining stations can be observed in Appendix A.

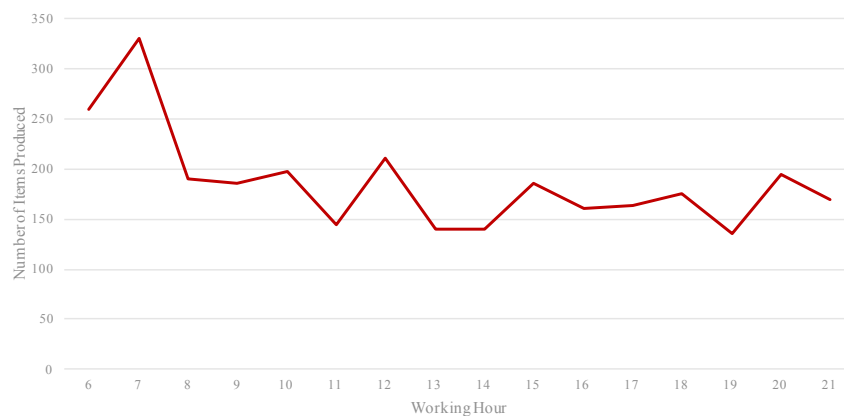


Figure 18 - Number of items produced per hour in IN on a given day



Figure 19 - Number of items produced per hour in LMW on a given day



Figure 20 - Number of Items produced per hour in ST on a given day

The stock times (in hours) were retrieved from the available database, regarding each one of the stations with impact on the process, and these can be seen in Figure 21. The hours in which the employees did not work (night and weekends) were excluded from the stock times. After the calculations, it was observed that there were some lead time values influencing negatively the average. High stock times may be explained by errors in the scanning of the code bar, products that were not processed due to defects, among others.

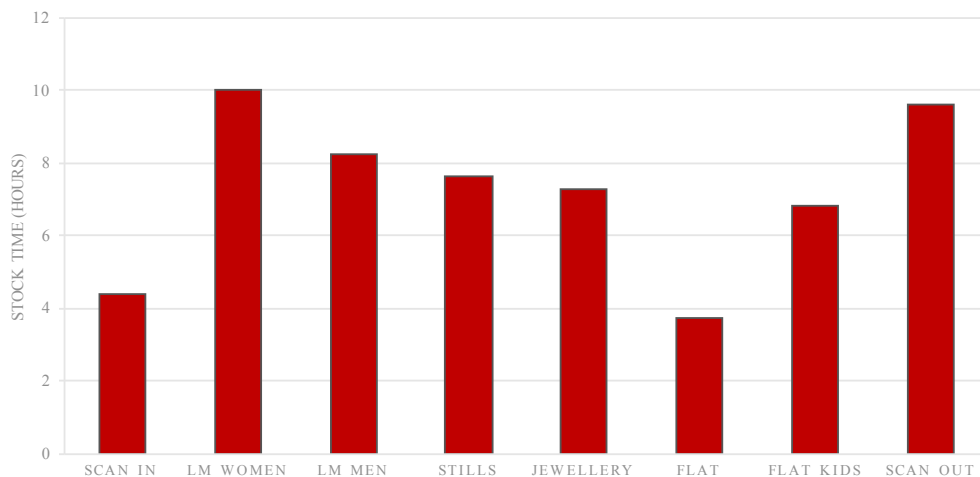


Figure 21 - Stock time at the main workstations in hours

Stations like Live Model can be considered the bottleneck of the process due to the great stock time and the number of items awaiting to be processed. The implementation of improvements that have in sight decreasing or taking more advantage of these periods may have a positive impact on the process. Scan Out has to wait for all rail items to get a quality check so they can proceed, and this is why the waiting time is so high, since it encompasses the phase of quality inspection of the previously taken photos.

To comprehend some of the important data regarding the slots that arrived to production centre during 2017, some parameters (presented in Table 1) were studied. Aiming to eliminate all the waste present in these data, some premises were assumed:

1. Only items produced in Europe were considered;
2. Repetitions were not considered;
3. Process periods superior to 6 days were not considered;
4. Weekends were not considered;
5. It was considered that all the items of slots followed standard processes.

Table 1 - General data about slots produced in 2017

	Average	Standard Deviation	Minimum	Maximum
Number of duplicates	8.4	7.0	0	41
Number of items (No duplicates)	25.8	12.1	0	50
Lead Time (hours)	61.7	21.9	0.8	144.0
Process Time (hours)	9.3	3.9	0.2	20.0

At last, after all the analysis previously performed and gathering all the retrieved information, it was understood that a slot spends on average 16.8% of the time being processed, and the remaining time waiting or in preparation.

3.3 Push System

The current digital production of the company follows a Push system throughout the whole process. This is not the ideal system for the productive process. However, since the goal is to avoid stock the product does not belong to the company, thus the more time it remains in stock, the higher the associated costs are.

The company's process is very dependent on its collaborators; as it is not a highly mechanized process, this complicate the forecast regarding absence due to station failure. By being a process with many production peaks, an absence in a peak period renders the process chaotic, causing workstations to stop. These stops can be due to blockage or lack of internal supply, that is to say, the stations could not produce because the downstream process had products in excess, therefore lacking stock area, or due to the contrary situation, in which the stations were stopped because the upstream process lacked capability to produce enough.

The stop of a station brings elevated costs to the production process, causing frequently the arrest of 5 or more workstations. In a future situation of possible growth, the types of problems would lead to the arrest of 10 or more workstations, which would mean duplicating the costs by station.

A very common example is when the stations of Scan Out and Live Model stop due to lack of product or lack of space in stock area in the Flat station. In Scan Out, the 6-minute daily arrest in 8 stations has an average cost of 5€. The problem can get bigger, however, in the case of a 6-minute daily arrest in the Live Model station, in which this delay per 6 stations has a cost of 32€. If these costs are converted to a monthly amount, and considering a month with 22 workdays, a cost of about 814€ is verified only in the Scan Out and Live Model Stations, due to excess of items in the Flat station. These costs do not correspond to reality, but are an approximation of actual costs.

3.4 Problems Observed in the Current Process

There are improvements that can be implemented having in sight the earlier placement of the products online, thus decreasing the process time and rendering the production more efficient.

After an analysis of the practical functioning of the processes, it is understood that in each station, the workers need to use diverse programs to register information, send photos, and verify duplicates, among other activities. These programs are not integrated, something that often potentiates the considerable increase of cycle times.

It was also noteworthy that the toolkit used in the Live Model station is manifestly insufficient when all the studios are functioning, since there is only a piece of each item. The stylist assistants have to wait until the photos are taken in a different studio to get the toolkit item to conjugate with the piece to be photographer. All of this has a negative impact on the production process.

There is a too high rate of duplicates. Since each partner boutique sends its items to be produced and as the same item can be sold by different entities, it is not uncommon that several pieces of clothing received have already been previously produced. This entails unnecessary costs of both transportation and labour for the boutiques and the company. These products occupy stock areas, and there is also a waste of time in the process of confirming whether they are duplicates or not. It was observed, however, that there has been an attempt in reducing this rate through the creation of a team prepared exclusively to the previous detection of these items.

There is also a problem in the current building. The production is divided in two different floors due to lack of space. This brings implications not only in terms of lead times, but also regarding repetitions requests. The process period extends because there is a waste of time in the ascending and descending of items through the elevators; this renders repetition requests as complications due to the shifts of the process. In the new building the main aim will be to render the process more efficient, in terms of occupied space and time. For this, the optimization of the space will be a vital step by avoiding the passage of the productive process through 2 floors; instead, it will be gathered in a single one.

The planning is made in a single Push system as previously mentioned. In other words, the orders are produced in each section without taking into consideration the necessities of internal customers and leading to an increased WIP, flow entropy and internal logistics. This system is not ideal when the main objective is to avoid stock. The well-performed implementation of a Pull system could allow the earlier placement of items to sell on the website and their return to the boutique sooner than expected.

4 Case Study

The analysis of the results from the completed studies regarding the current situation of the company allow to perceive a necessity for a change in the processes. There are slight changes that are easily executable, resulting in gains not only in terms of processing times but also of productivity and collaborators' efficiency. These minor modifications are, for example, the creation of integrated and more dynamic software's. It was also observable that cycle times can be reduced, but for this, there is a necessity to understand what really influences these periods. This study will be detailed in subchapter 4.1, through a statistical analysis method – multivariate regression. After the calculation of stock times between the different stations that compose the process, it was finally possible to verify that these are too high, representing about 83% of the total period a slot spends in the production centre. The study performed in subchapter 4.2 enables a strategical overview on how these periods can be used to execute activities in parallel, which are obligatory but can be performed at any stage of the process.

4.1 Decrease of Cycle Time in Live Model Women Station

As observed in the previous chapter, Live Model Women stations receives clothes, jewellery and some accessories (e.g., bags). Women clothing, jewellery and these women accessories represent 30.11%, 1.92% and 6.37% of the total, respectively, which means that 38.4% of the items that arrive to the centre of digital production of the company go through this station. As this rate is high when compared to the remaining stations, and considering that this is the stations (together with Live Model Men) that brings higher costs, there is a necessity to understand if these cycle times can be reduced, with the goal of increasing the efficiency of this station.

As observed in the field, a slot can contain a maximum of 50 items. However, there is no asserted limit of categories. It may contain women, men, children or unisex clothes. When in Scan In, the slot is divided in several rails or boxes according to the categories of its items and of the stations it will go through.

A previous preparation of the outfits is made when a rail arrives to the Live Model stations; thereafter, these proceed to a studio to be photographed. The model assigned to the studio dresses the items one at a time, as well as the chosen sets to combine with these, and he or she is photographed. Each time the model finishes photographing one item, he or she muss undress and put on the next set of clothes.

It was verified that whenever a slot is composed by a single category (trousers, for example), the cycle time of the Live Model station decreases, since the models often maintain the same top part and only change the piece to be photographed.

A study was performed to understand how the cycle time of the Live Model Women, the station with more impact, is influenced by different variables, applying a multivariate regression. A multivariate linear regression was chosen since it enables to predict the value of

a dependent variable according to more than one independent variable, while being a relatively simple statistical method.

In order to retrieve from the SQL the data needed for the regression, a series of constraints were imposed when creating the query used. The Live Model date could not be null, the category of the articles could not be null, the production of the articles was performed in Europe and the date of passing through the Live Model station had to be higher than 01/01/2016.

4.1.1 Multivariate Regression with 2 Independent Variables

From the various possible methods to use and predict the cycle time, a statistical method was used – multivariate regression. This regression aimed to predict the cycle times for the Live Model station.

It is necessary to define firstly what are the variables to consider in order to perform the multivariate regression. The dependent variable is the cycle time (min), since it is the one whose behaviour we aim to predict, according to the independent variables. A set of variables that were predicted to influence the dependent one was defined:

- Number of items: number of items contained in each slot that go through the Live Model Women station;
- Number of categories: number of different categories that the items belong to;
- Season: season (low or high) in which the items were produced.

In multiple regression, the model takes the form of Eq.1 and in that equation, there are several unknown parameters (the b-values). The b-values tell about the relationship between Live Model Women Cycle Time and each predictor.

The equation generated by the regression presents the following structure:

$$T_{ciclo} = b_0 + (b_1 \times NItems) + (b_2 \times NCategories) + (b_3 \times Season) \quad (1)$$

where:

b_0 is a constant

b_1 is the slope coefficient for the variable “Number of Items”

b_2 is the slope coefficient for the variable “Number of categories”

b_3 is the slope coefficient for the variable “Season”

All the values of the independent and dependent variables mentioned above were retrieved from the databases. Beyond these variables, it is expected that the cycle time also depends on the experience of the team composed by the photographer, stylist, stylist assistant and model. There are certain rules to be followed regarding the position of the model and the clothes. Teams with less experience probably take more time managing the piece to be photographed perfectly. It was not possible, however, to distinguish the teams and understand how this variable influences the cycle time.

After the retrieval of these data, the sample obtained was composed of 14 684 observations which were registered since the beginning of the year 2016. The multivariate analysis was performed after using the statistical software SPSS.

4.1.2 Data Preparation

With the purpose of performing the multivariate regression using SPSS, the first step of the process consisted in inserting the data in the program. To do so, all the previously described variables were inserted with their respective values, with the exception of the variable ‘Team experience’, since it was not possible to be quantified. In addition, a variable called ID was created to help organize the data.

The type of variable was decided for each one of them; all were classified as scalar or nominal. Scalar variables correspond to variables of measurable characteristics that can be measured in a continuous scale, as it is the case of cycle time (min). Nominal variables have qualitative instead of quantitative characteristic, as it is the case of categorical variables that represent a classification. These can be exemplified by the variable Season, which can be low or high. In Figure 22 it is possible to observe the classification of the different variables studied in the software.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	0		None	None	8	Right	Nominal	None
2	CycleTime	Numeric	3	0	Cycle Time	None	None	12	Right	Scale	None
3	NºItems	Numeric	2	0	Nº Items	None	None	12	Right	Scale	None
4	NºCategory	Numeric	1	0	Nº Category	None	None	12	Right	Scale	None
5	Season	String	5	0	Season	{1, Low}...	None	11	Left	Nominal	None

Figure 22 - Classification of the variables in SPSS software.

It was necessary to encode the values that the nominal variable can assume. As in the Excel file, in the season column had the information regarding if the photos were taken in low or high season, corresponding this to each line of cycle time. A dummy variable was matched to each season, and these are indicated as “_d”. It was assumed that the photos produced in the months of January, April, May, June, October, November and December corresponded to the low season, in other words, months with a reduced number of produced items when comparing to the others. The remaining months corresponded to the high season, since they presented higher levels of production.

A dummy variable is an explicative variable that can assume only two values: 0 (absence of an attribute) and 1 (presence of an attribute). This allows the inclusion of several qualitative variables in a regression model. It is important to outline that if a variable presents K categories, the number of dummy variables necessary to encode it is K-1.

After all the variables were correctly introduced and defined in the programme, the data table looked like the one in Figure 23.

	ID	CycleTime	N°Items	N°Category	High_d	Low_d
1	1	9	4	1	1	0
2	2	11	3	1	1	0
3	3	84	22	6	1	0
4	4	60	12	4	1	0
5	5	95	25	2	0	1
6	6	43	12	3	0	1
7	7	52	14	2	0	1
8	8	29	10	4	1	0
9	9	26	10	1	1	0
10	10	31	12	6	0	1
11	11	51	11	4	1	0
12	12	87	18	4	0	1
13	13	50	18	4	1	0
14	14	31	13	1	0	1
15	15	98	32	5	1	0

Figure 23 - Data view in SPSS

4.1.3 Verification of the Assumptions and Realization of Regression

The assumptions that need to be verified and fulfilled when performing a multivariate analysis are mentioned in subchapter 2.3. The process performed for this verification will be detailed in the present subchapter.

The first two assumptions that state that **the dependent variable must be measured in a continuous scale** and that **there must be two or more independent variables which may be continuous or categorical** can be verified even before performing the regression. The remaining assumptions need to be verified posteriorly. A reverse logic was adopted, in other words, although there is a need for validating the assumptions before performing the regression, only by carrying it out and generating some outputs could the majority of these prerequisites be verified. To begin with, the fulfilment of the first two assumptions was established; the multivariate analysis was then performed and, at last, the remaining premises were validated, always keeping in mind that the non-verification of any one of them would invalidate the regression.

The used method of variable selection when performing the regression was “Enter”, which is the predefined method in SPSS. The standard method of entry is simultaneous and all independent variables are entered into the equation at the same time. This is an appropriate analysis when the researcher does not know which independent variables will create the best prediction equation. In addition to the existing variables, 5 auxiliary ones were created to use for the verification of the assumptions. These variables that are presented on Figure 24, contain the predicted non-standardized value (PRE_1), the studentized residues (SRE_1), the deleted studentized residues (SDR_1), the values of Cook’s Distance (COO_1) and the leverage values (LEV_1).

PRE_1	SRE_1	SDR_1	COO_1	LEV_1
26.11402	-.65630	-.65628	.00003	.00017
23.12945	-.46515	-.46514	.00001	.00017
92.81024	-.33787	-.33786	.00001	.00023
57.77496	.08532	.08532	.00000	.00010
89.31808	.21792	.21791	.00001	.00039
53.11352	-.38783	-.38782	.00001	.00013
56.48785	-.17210	-.17210	.00000	.00016
51.80583	-.87456	-.87455	.00004	.00014
44.02142	-.69110	-.69109	.00003	.00018
60.89794	-1.14678	-1.14679	.00021	.00058
54.79040	-.14535	-.14535	.00000	.00012
73.61573	.51326	.51325	.00001	.00013
75.68236	-.98482	-.98482	.00003	.00005
50.90848	-.76350	-.76349	.00005	.00027
120.06110	-.84603	-.84602	.00005	.00020
125.64047	-.63816	-.63815	.00003	.00025
146.82210	-.02282	-.02281	.00000	.00000

Figure 24 - Auxiliary variables created in the multivariate regression

Assumptions Verification

- 1. The dependent variables are measured in a continuous scale.**
- 2. There are two or more independent variables that can be continuous or categorical.**

As explained in the present chapter, these two assumptions are the first to be verified and validated; in case they are not fulfilled, the multivariate analysis would not be an adequate statistical test to analyse the data.

Cycle time is a measurable variable in this case, assuming value in a continuous scale. As previously mentioned in subchapter 4.1.1, the independent variables make up a total of 3 and all are continuous (defined in the programme as scalar variables) or categorical (defined as nominal). This first two assumptions were then validated.

- 3. There is independency of the observations.**

The assumption of observation independence in a multivariate regression implies that there is no first order correlation between the observations: adjacent observations should not be correlated. The equation used to estimate the standard error is valid only if observations are independent. The standard error is used to compute confidence intervals and significance tests, so if the assumption of independence is violated, then confidence intervals and significance tests will be invalid.

The data were retrieved from the databases and the slots were produced in different times of the day and year; some were produced in the first shift, others in the next shift, in low and high seasons. The collected sample was therefore as random as possible. The only imposed restriction was that the values of cycle time were centred between 0 and 240 minutes, since it was considered that from that point on, the high cycle time was due to external problems, as it is the case of an item that was stopped due to a default and that was photographed later than the rest of the slot's items.

To confirm if the assumption was indeed fulfilled, a Durbin-Watson test was performed. This test is performed to detect a possible correlation in the retrieval of data. The value of the test can range from 0 to 4; for the non-existence of a correlation, the test must assume a value of approximately 2. As it is observed in Table 2, the obtained value was 1.723 and this specific assumption was verified.

Table 2 - Model summary: Durbin-Watson test

Model	R	R ²	Adjusted R ²	Std. Error of the Estimate	Durbin-Watson
1	0.797	0.755	0.706	26.080	1.723

4. There is a linear relation between (a) the dependent variable and each of the independent variables and between (b) the dependent variable and the independent variables collectively.

According to Andy Field (2017), this assumption is the most important because if it is not true then even if all other assumptions are met, the model is invalid because it was incorrectly described. This assumption of linearity means that the outcome variable is, in fact, linearly related to any predictors and that if you have several predictors, then their combined effect is the best described by adding their effects together.

The validation of the linearity assumption in the multivariate regression was divided in two parts, one dedicated to the verification of (a) and another to the verification of (b).

For the validation of the condition (a), partial regressions were used between each one of the independent and dependent variables; the categorical variable may be ignored. The observation of Figure 25 leads to the conclusion that both presented regressions follow an approximately equal distribution to the ones shown in subchapter 2.3. Condition (a) is therefore fulfilled.

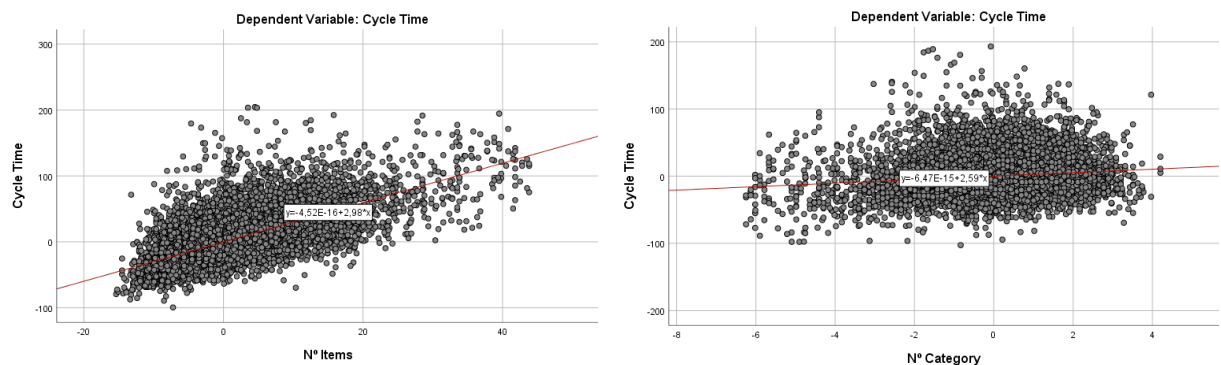


Figure 25 - Partial regression between independent variables and dependent variable

To validate condition (b), the scatterplot observed in Figure 26 was performed, using the studentized residues (SRE_1) and the predicted non-standardized values (PRE_1). If the residues have an approximately horizontal distribution, it can be concluded that there is strong evidence pointing towards a linear and collective relation between the dependent and independent variables.

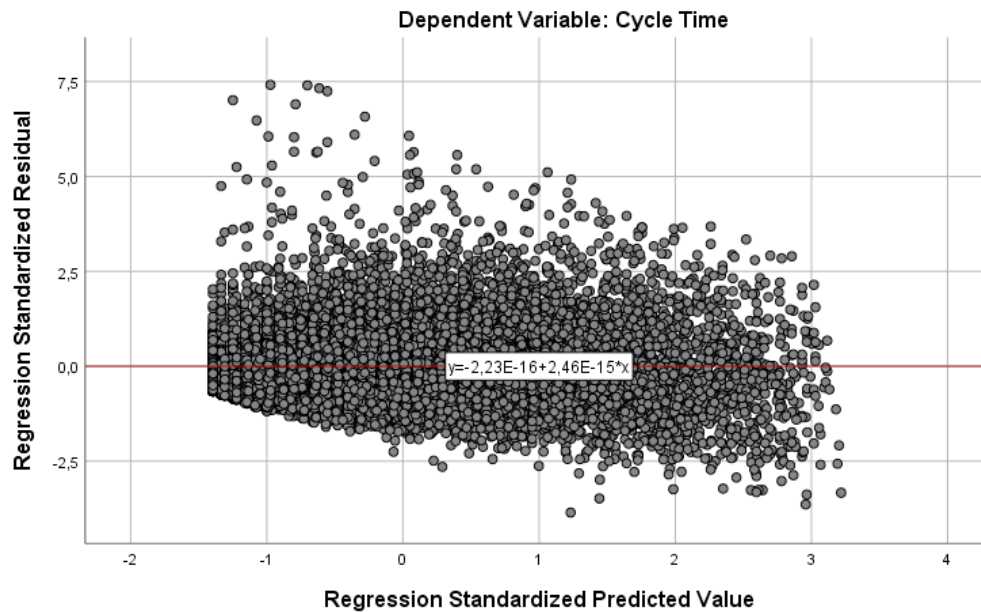


Figure 26 - ScatterPlot of the standardized residues and the standardized predicted values

The analysis of the previous ScatterPlot can lead to the uncertain conclusion that the residues' distribution is approximately horizontal. Due to this reason, condition (b) and consequently this premise were both validated.

5. The data must have homoscedasticity of residues.

Homoscedasticity is the term used to designate a constant variance of experimental errors for distinct observations.

The scatterplot of Figure 26 was used to verify the assumption of residues' homoscedasticity. For the validation, the dispersion of residues must not increase nor decrease as one moves along the axis of the predicted non-standardized values. By observing the graph, it is not clear that the dispersion remains constant along the x-axis. Levene's Test was then performed to test the null hypothesis that states the variance in different groups to be equal.

- If the p-value of Levene's Test is equal or superior to 0.05, it is concluded that the variances are significantly different. The assumption of variance homogeneity is violated.
- If Levene's Test is not significant ($p > 0.05$), the variances are approximately equal and the assumption is sustainable.

The observation of Table 3 generated by SPSS allows to ascertain that the p-value (highlighted in red) has a value highly superior to the level of significance. Because of that, the assumption is then verified.

Table 3 - Test of homogeneity of variance (Levene test)

		Levene Statistic	df1	df2	Sig.
Cycle Time	Based on Mean	0.200	1	14682	0.655
	Based on Median	0.001	1	14682	0.977
	Based on Median and with adjusted df	0.001	1	14677.93	0.977
	Based on trimmed mean	0.069	1	14682	0.793

6. The data do not reveal multicollinearity.

The correlation table generated when executing the regression (presented in Table 4) was used to verify this assumption. The correlation values should be inferior to 0.7 for it to be validated.

Table 4 - Pearson correlation

	Cycle Time	Number of Items	Number of Categories	Dummy for High Season	Dummy for Low Season
Cycle Time	1.000	0.793	0.603	0.040	-0.040
Number of Items	0.793	1.000	0.690	0.017	-0.017
Number of Categories	0.603	0.690	1.000	0.45	-0.045
Dummy for High Season	0.040	0.017	0.45	1.000	-1.000
Dummy for low Season	-0.040	-0.017	-0.45	-1.000	1.000

The values highlighted in red in the table above were analysed, and it was after verified that all independent variables fulfilled this assumption.

The validation is also dependent on the verification that the tolerance values, present in the statistics of collinearity, are superior to 0.1. As observed in Table 5, that contains data retrieved from the tables generated by multivariate analysis, the assumption is verified.

Table 5 - Collinearity statistics

Model	Unstandardized Coef.		t	Sigma	Tolerance	VIF
	B	Std. Error				
(Constant)	11.581	0.497	23.318	0.000		
Nitems	2.985	0.029	104.649	0.000	0.524	1.910
Ncategories	2.595	0.171	15.140	0.000	0.523	1.913
Dummy for Low Season	-2.067	0.447	-4.626	0.000	0.998	1.002

7. There are no outliers.

An outlier is a point that presents a different behaviour from the remainders. If it is influent, it interferes with the adjusted function of regression. An observation considered an outlier is not always, however, looked at as an influent point.

To verify the existence of outlier points, the auxiliary variables SDR_1, COO_1 and LEV_1 were used.

Initially, it was verified that some values of the SDR_1 variable were superior to ± 3 standard deviations; should this happen, it would mean that the corresponding observation was a potential outlier. For the sake of simplicity, the data was exported to an Excel file and ordered in decreasing order, as illustrated in Figure 27. The standard deviation of SDR_1 was then calculated, as well as the value range in which the observation was not considered a potential outlier.

	ID	CycleTime	N°Items	N°Category	High_d	Low_d	PRE_1	ZRE_1	SDR_1
1	10972	226	6	2	0	1	32,61132	7,41530	7,42987
2	7751	235	10	1	0	1	41,95478	7,40213	7,41686
3	2983	236	11	1	0	1	44,93935	7,32603	7,34037
4	11527	236	11	1	1	0	47,00598	7,24679	7,26048
5	14161	206	3	1	1	0	23,12945	7,01199	7,02438
6	6696	219	9	1	0	1	38,97021	6,90306	6,91509
7	10975	228	14	2	0	1	56,48785	6,57646	6,58670
8	2266	198	5	1	1	0	29,09859	6,47636	6,48616
9	11112	213	14	1	0	1	53,89305	6,10080	6,10944
10	2607	226	16	4	0	1	67,64659	6,07190	6,07998
11	13513	190	6	1	1	0	32,08315	6,05516	6,06323
12	10973	196	8	2	0	1	38,58045	6,03610	6,04407
13	11528	201	11	1	1	0	47,00598	5,90475	5,91233
14	3290	192	10	2	0	1	44,54959	5,65384	5,66043
15	7071	186	8	2	0	1	38,58045	5,65266	5,65926

Figure 27 - Data sorted in descending order of SDR_1

Knowing that the calculated standard deviation was 1.00023, the value range of non-potential outliers was:

$$[-3.007 ; 3.007]$$

It was concluded that there were 195 observations with SDR_1 values not contained within the calculated range that constitute approximately 1.35% of total of registered observations. The observations could constitute potential outliers, therefore, they should be paid special attention in the subsequent verifications.

The verification regarding observations with values of leverage was then performed. The leverage points present an unusual value of the explicate variable. The leverage values can be considered:

- Safe: inferior to 0,2;
- Risky: between 0,2 and 0,5;
- Dangerous: superior to 0,5.

To perform this evaluation, the data were organized in decreasing order regarding the value of the variable LEV_1. No observation had a leverage value considered dangerous or even risky, as depicted in Figure 28.

	N°Category	High_d	Low_d	PRE_1	ZRE_1	SDR_1	SRE_1	COO_1	LEV_1
1	1	0	1	161,33744	-1,85345	-1,85598	-1,85582	,00221	,00249
2	1	0	1	161,33744	-,16631	-,16652	-,16653	,00002	,00249
3	1	1	0	163,40407	-2,12441	-2,12736	-2,12710	,00286	,00246
4	1	1	0	163,40407	-,59065	-,59139	-,59140	,00022	,00246
5	1	1	0	163,40407	-,47562	-,47621	-,47622	,00014	,00246
6	1	0	1	158,35287	-1,35557	-1,35727	-1,35724	,00113	,00238
7	1	1	0	160,41951	-,47621	-,47678	-,47679	,00014	,00235
8	1	1	0	160,41951	-,28449	-,28483	-,28484	,00005	,00235
9	1	0	1	155,36830	-2,81323	-2,81721	-2,81655	,00467	,00228
10	1	0	1	155,36830	,06257	,06264	,06264	,00000	,00228
11	1	0	1	155,36830	,36932	,36974	,36975	,00008	,00228
12	1	1	0	157,43494	-2,35566	-2,35877	-2,35840	,00323	,00225
13	1	1	0	157,43494	-1,51209	-1,51392	-1,51385	,00133	,00225
14	1	1	0	157,43494	-1,01362	-1,01480	-1,01480	,00060	,00225
15	1	1	0	157,43494	-,51515	-,51574	-,51575	,00015	,00225

Figure 28 - Data sorted in descending order of LEV_1

The next step consisted in verifying the level of influence of the several points. The variable COO_1 was used for such, a variable that stores the values of Cook's Distance for each observation. A point can be considered influent if its exclusion from the regression adjustment induces a substantial change in the regression analysis. This measures the influence that a given observation has on all the adjusted values. Similarly to what was previously done, the data were organized in decreasing order to facilitate the observation concerning the value assumed a higher value. It is important to take into account that points with a value superior to 1 may be considered problematic. As it is possible to determine in Figure 29, no observation had an influence level considered dangerous.

	N°Category	High_d	Low_d	PRE_1	ZRE_1	SDR_1	SRE_1	COO_1	LEV_1
1	1	0	1	161,33744	-1,85345	-1,85598	-1,85582	,00221	,00249
2	1	0	1	161,33744	-,16631	-,16652	-,16653	,00002	,00249
3	1	1	0	163,40407	-2,12441	-2,12736	-2,12710	,00286	,00246
4	1	1	0	163,40407	-,59065	-,59139	-,59140	,00022	,00246
5	1	1	0	163,40407	-,47562	-,47621	-,47622	,00014	,00246
6	1	0	1	158,35287	-1,35557	-1,35727	-1,35724	,00113	,00238
7	1	1	0	160,41951	-,47621	-,47678	-,47679	,00014	,00235
8	1	1	0	160,41951	-,28449	-,28483	-,28484	,00005	,00235
9	1	0	1	155,36830	-2,81323	-2,81721	-2,81655	,00467	,00228
10	1	0	1	155,36830	,06257	,06264	,06264	,00000	,00228
11	1	0	1	155,36830	,36932	,36974	,36975	,00008	,00228
12	1	1	0	157,43494	-2,35566	-2,35877	-2,35840	,00323	,00225
13	1	1	0	157,43494	-1,51209	-1,51392	-1,51385	,00133	,00225
14	1	1	0	157,43494	-1,01362	-1,01480	-1,01480	,00060	,00225
15	1	1	0	157,43494	-,51515	-,51574	-,51575	,00015	,00225

Figure 29 - Data sorted in descending order of COO_1

After all the verifications, it was statistically concluded that 195 observations, constituting 1.35% of the total, were considered outliers. They don't present, however, a significant influence, and therefore they do not cause a drastic variation in the multivariate regression model. In spite of this, as a security measure, the assumption was not validated.

8. The residues (errors) are normally distributed.

The standardized residue(s) histogram presented in Figure 30 was used to verify this assumption. As observed, the distribution of residues is approximately normal.

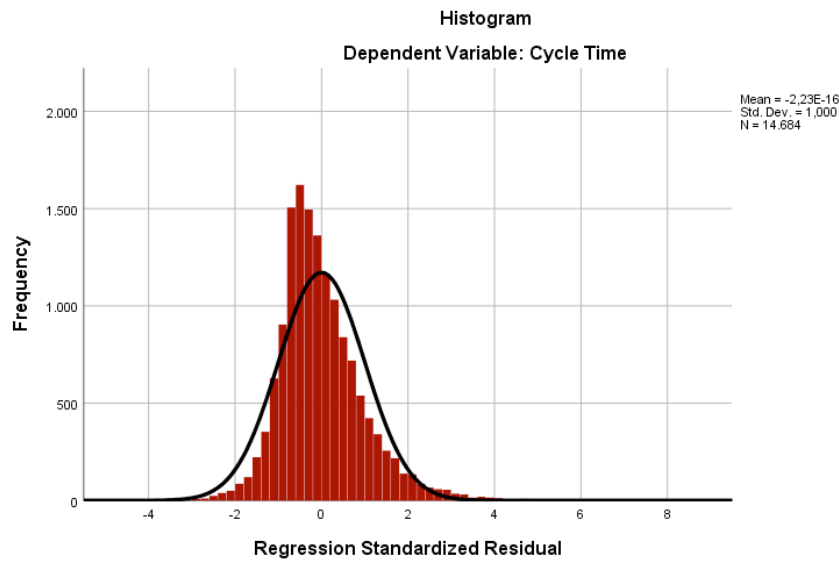


Figure 30 - Histogram of standardized residues

The normal PP-Plot of standardized residues regression was used to confirm this conclusion. If the residues are approximately normalized, the points will be aligned with the diagonal straight line present in the graph. By observing Figure 31, it is perceptible that although the points are not perfectly aligned with the line, their alignment is sufficiently close to determine the residues to be approximately normal. The multivariate regression is robust enough to “allow” some deviations from the normal line, and that is why this result is perfectly acceptable and does not violate the assumption of residues’ normality.

The central limit theorem means that as sample sizes get larger, the less the assumption of normality matters because the sampling distribution will be normal regardless of what our population data look like. So, the problem is that in large samples, where we don’t need to worry about normality, a test of normality is more likely to be significant, and therefore likely to make us worry about and correct for something that doesn’t need to be corrected or worried about.

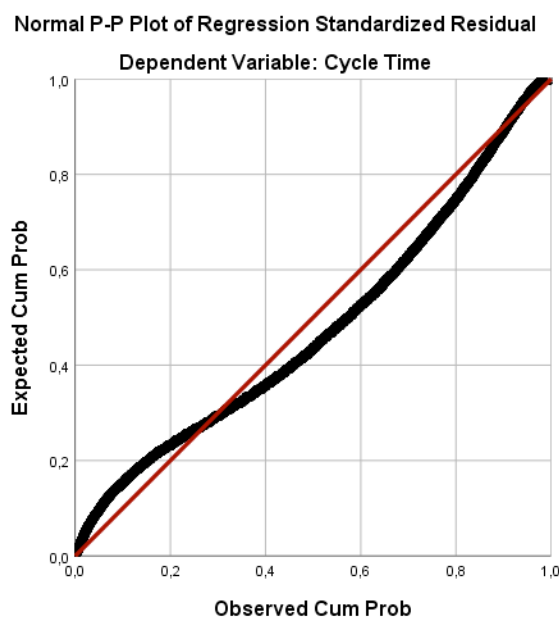


Figure 31 - Normal P-P plot of regression standardized residual

Summary of Assumptions' Analysis

After executing all the procedure of assumptions' verification previously described, it can be concluded that:

- The verification of assumption 5 left some doubts, however, after performing Levene's test, the assumption was verified;
- Assumption 8 was validated through as using the Central Limit Theorem, since the graphs were not perfectly clear;
- Assumption 7 was not validated, as the existence of 195 outliers was observed.

Although these outliers do not influence greatly the regression model, it was chosen to execute a new regression without their presence. A new verification of all the assumptions was then performed. To remove the outliers, the data were organized once more in decreasing order of SDR_1, excluding all the observations that presented a value of $SDR_1 \pm 3$. After this, the data were ordered in increasing order of ID to guarantee the independence of observations.

Verification of assumptions after outlier removal

Following an identical procedure to the one previously described, the verification of assumptions was performed once more.

The outliers were removed because, after an observation of the Live Model Women process, it was realized that there are sometimes flaws on the process due to several motives: unavailability of the chosen clothes, failure of the used software during certain periods, failure of clothing barcode scanning, among others. These influence negatively the cycle time of the workstation and the outliers correspond mainly to these type of situations.

The observation of Figure 32, Figure 33 and Figure 34 and following tables allows the confirmation that all the assumptions were validated; the analysis of obtained results followed.

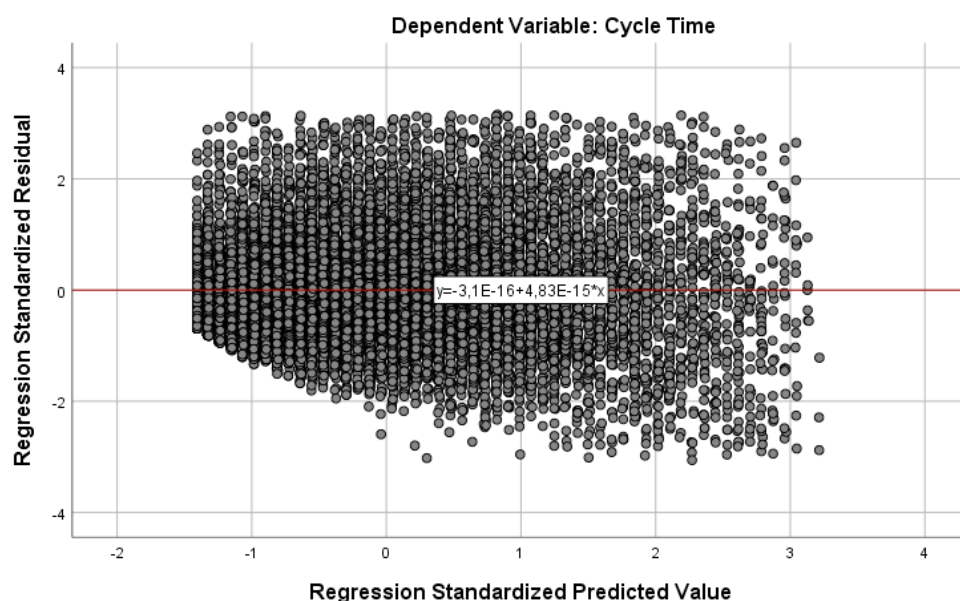


Figure 32 - ScatterPlot of SRE and PRE of the new regression

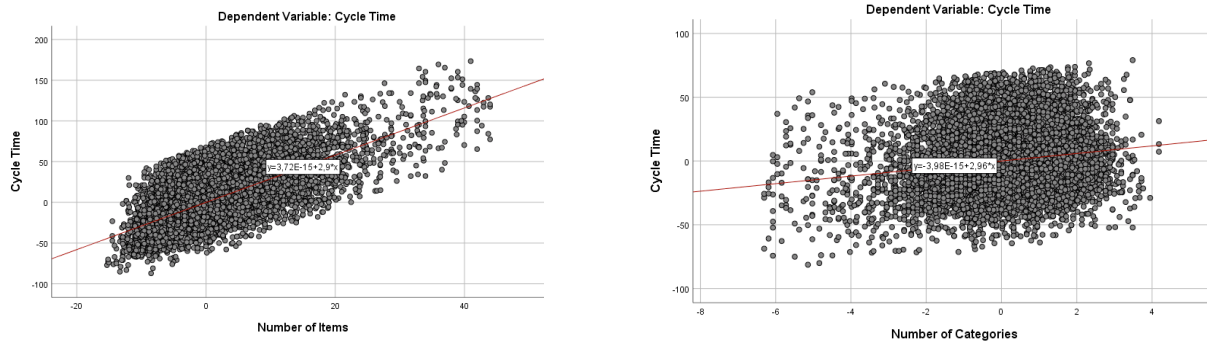


Figure 33 - New partial regression between independent variables and dependent variable

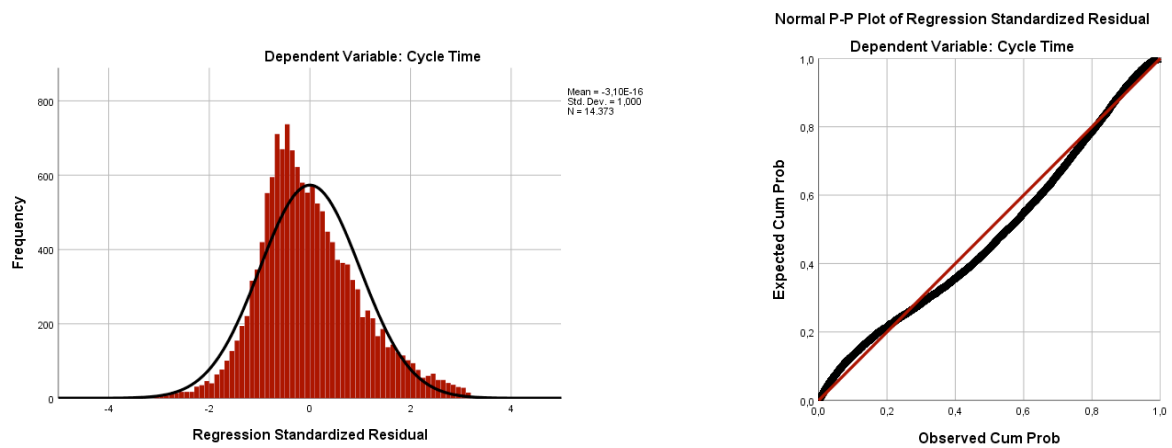


Figure 34 - Histogram and P-P Plot of the new regression

Table 6 - Model summary: Durbin-Watson test of the new regression

Model	R	R ²	Adjusted R ²	Std. Error of the Estimate	Durbin-Watson
1	0.836	0.700	0.700	22.215	1.862

Table 7 - Collinearity statistics of the new regression

Model	Unstandardized Coeff.		t	Sig.	Tolerance	VIF
	B	Std. Error				
(Constant)	10.165	0.427	23.792	0.000		
Nitems	2.899	0.025	116.514	0.000	0.515	1.941
Ncategories	2.960	0.149	19.863	0.000	0.514	1.944
Dummy for Low Season	-2.623	0.385	-6.813	0.000	0.998	1.002

Table 8 - Test of homogeneity of variance of the new regression

		Levene Statistic	df1	df2	Sig.
Cycle Time	Based on Mean	1.031	1	14371	0.310
	Based on Median	1.789	1	14371	0.181
	Based on Median and with adjusted df	1.789	1	14368.00	0.181
	Based on trimmed mean	1.252	1	14371	0.263

4.1.4 Analysis of Results

After executing the regression, there are a series of outputs that contain all the generated information by the program, which must be interpreted to ascertain the validity of the regression and understand its statistical significance. This analysis of results may be divided in two different stages:

1. Verification of the regression's adequacy to the data
2. Interpretation of regression coefficients

1. Verification of the regression's adequacy to the data

Some coefficients observed in the summary table of the model were interpreted to understand if the regression is adequate to the existing data.

The first analysed element was the **coefficient of multiple correlation (R)**. This coefficient measures the strength of linear association among variables. Its value can range between 0 and 1; the higher the value, the stronger the linear association. As observed in Table 9, the value of this coefficient for the performed regression was $R = 0.836$, indicating that there is a strong level of linear association among the variables.

Table 9 - Regression summary table

Model	R	R ²	Adjusted R ²	Std. Error of the Estimate
1	0.836	0.700	0.700	22.215

The element analysed after was the proportion of variation explained (R^2) that measures from 0 to 1 the proportion of dependent variance explained by the independent variables. $R^2 = 0.700$, as observed in Table 9. The percentage that lacks explaining, 30%, may be due to the fact that it was not possible to create a variable to quantify the experience of the production team. The scarce automation of the process (performed integrally by people, with the photo camera being the only machine) also contributed to cycle time variability.

The adjusted R^2 value corrects positive bias to provide a value that would be expected in the population. The fact that the adjusted R^2 has a value equal to R^2 can be explained by the large sample size.

The statistical significance of the regression was verified using the ANOVA table presented in Table 10.

Table 10 - ANOVA table

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16522944.86	3	5507648.29	11160.610	0.000
	Residual	7090956.50	14369	493.490		
	Total	23613901.37	14372			

For a statistical significance of $p < 0$ (or 5%), the p-value indicated in the column “Sig.” must be inferior to 0.05. Since the value of “Sig.” is 0.000, it can be affirmed that we are before a statistically significant result.

2. Interpretation of Regression Coefficients

The statistical significance of each one of the variables was evaluated making use of the coefficient table (Table 7), and the coefficients of the column “Sig.” were analysed. Once again, for the variables to present statistical significance at a $p < 0.05$ level and be considered in the regression, their “Sig.” value must be inferior to 0.05. It was observed that this condition is satisfied by the three variables, therefore, all variables present statistical significance and explain the variation of the dependent variable.

When replacing the values of the constants b_0 , b_1 , b_2 e b_3 of equation (1), an equation that allows to estimate the cycle time of Live Model Women workstations according to the other variables is obtained. The values of these constants are retrieved from the column denominated B present in the coefficient table (Table 7). The final equation for cycle is obtained by replacing in equation (1):

$$T_{ciclo} = 10.165 + (2.899 \times NItems) + (2.960 \times NCategories) - (2.623 \times Season) \quad (2)$$

It is important to recollect that if the season is low, the variable assumes the value 1; otherwise, in case of high season, the variable must assume the value 0.

There are two predictors that have positive b-values indicating positive relationships. So, as the number of items increases, cycle time increases; as the number of categories increases, so do cycle time. Finally, if the season is low, cycle time decreases, but if the season is high, cycle time remains unchanged. The b-values tell what degree each predictor affects the outcome if the effects of all other predictors are held constant:

- Number of items ($b = 2.899$): as number of items increases by one unit, cycle time increases by 2.899 units.
- Number of categories ($b = 2.960$): as the number of categories increases by one, cycle time increases by 2.960 units.
- Season ($b = -2.623$): if the season is Low, cycle time decreases by 2.623 units, but if season is high, cycle time doesn't change.

To be able to quantify the percentage of increase in cycle time on the workstation, three different slots were created, all photographed in low season:

- Slot 1: 50 items in which all are women's trousers;
- Slot 2: 50 items that encompass women's trousers and skirts;
- Slot 3: 50 items of 9 different categories.

The cycle times of these three slots can be consulted in Table 11. Introducing a new category on the slot increases cycle time in 1.9%. The introduction of new categories aside from the one already existent implies an increase of 13.2% of cycle time, a value that starts to be significant.

Table 11 - Cycle time of the 3 slots with different characteristics

	Slot 1	Slot 2	Slot 3
Number of Items	50	50	50
Number of Categories	1	2	9
Cycle Time [min]	155	158	179
Cycle Time [hours]	2.60	2.64	3.00

After developing this study, the necessity of diminishing the number of categories in a slot is evident. A single slot does not bring a very significant profit to the company, but in the long-term, highly positive profits can be obtained regarding the time of item processing.

A scenario was considered in which all the work posts would photograph two different categories at most, and where the items would not be organized by slot but accordingly to their category. In this case, a new work post after Live Model would have to be available to perform the items' triage with the objective to reorganize them by slot, proceeding after to the boutique without flaws in the process. Although the periods resulting from the introduction of this new triage station at the end of the process were ignored, it was understood that it would be possible to spare, at the very least, 3 minutes by item if the number of categories photographed by post was reduced to 2. Considering that a month has 22 work days on average, and performing the calculations, it is perceptible that about 73 hours are spared monthly with this change, again without including the required times to execute triage. 837 hours are estimated to be saved in a year.

4.2 Stock Time Improvements

An improvement proposal will be presented in the present subchapter, consisting in the reuse of stock times between stations to do, in parallel, activities that can be performed at any time of the process. As ascertained in the previous chapter, stock times between stations present an impact excessively significant on the times of products' processing. On average, about 83.2% of the time a slot remains in the production centre is spent in preparations or in stock areas. An attempt regarding the reduction or reutilization of this time can bring savings, even if small. The execution and analysis of the process' general mapping and the detailed mapping of each main workstation present in Appendix B and in Figure 35 allowed to understand the normal flow each station follows, and to realize the existence of two distinct activities in the Scan Out stations. The latter are termed: **Data Entry** and **Packing**.

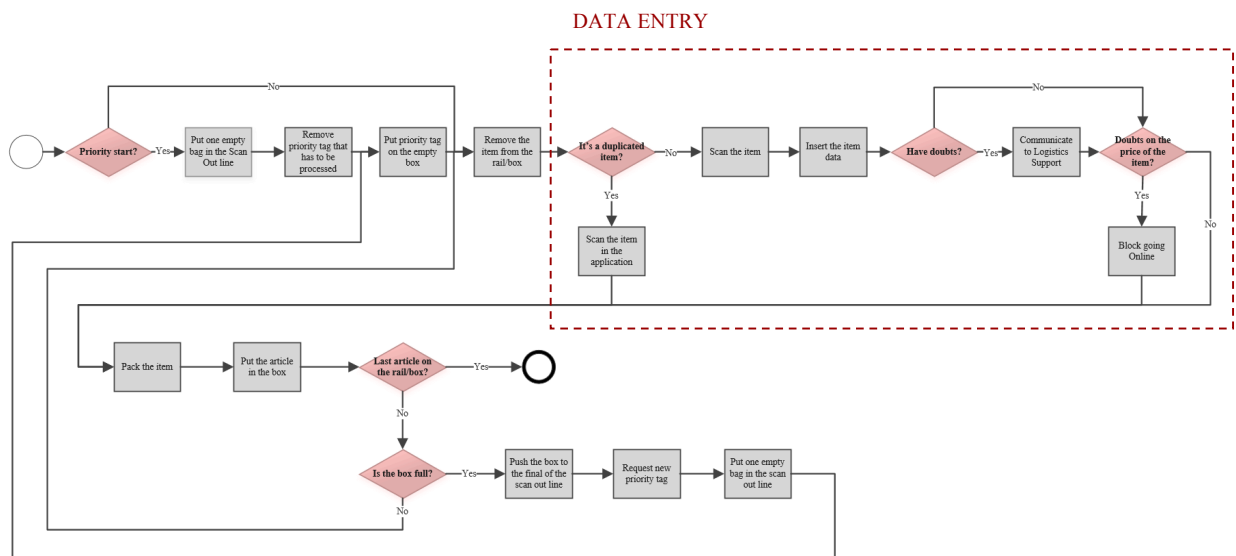


Figure 35 - Scan Out process mapping

At the stage denominated Data Entry, the collaborator takes the product, inserts and confirms all the information relative to the same. These data are, for example, product ID, size, brand, price, washing instructions, among others. At present, the product measurements are still not performed, so that the customer actually understands which the measures are. There is, however, an objective to perform a study of the impact of such and possible inclusion of this new activity at the stage of Data Entry.

Packing, at the stage, the collaborator removes the products from the rail, inspects them to determine the existence or non-existence of defects, places the tags and packs again the items to be place in the box together with the remaining products and proceed to the respective boutique.

After a detailed analysis, it became clear that the process of Data Entry can be performed at any time of the process, since it does not require any other stage to be concluded for data insertion into the system. It was then noticeable the need to study the impact of separating the Scan Out workstation in two distinct workstations, Data Entry and Packing, as well as to study the impact of this separation on lead times, taking advantage of the high stock times of other workstations for executing Data Entry.

4.2.1 “Split Scan Out” Test

A “Split Scan Out” test was performed in the production to realize which would be the cycle times after separating Scan Out in two new workstations. From subchapter 3.2, we have seen that an item spends 2 minutes and 30 seconds on average in this workstation. However, it was unknown how much of this time was spent on Data Entry, Packing and how splitting the workstations would influence the cycle time in question.

The test consisted on the retrieval of data relative to the necessary periods to perform Data Entry and Packing of all the items that went through the workstations. It lasted about 8 work days and occurred on the morning shift, from 6h to 14h. It was intended to get a representative sample of the periods, so that a sustained and reasoned analysis could be performed.

The layouts were created for the workstations of Data Entry and Packing. The layouts can be observed in Figure 36. Packing would be expected to have a reduced cycle time when compared with Data Entry, so it was opted to create, for the study, two workstations for Data Entry to feed and avoid the lack of product of the only station in Packing. It was therefore guaranteed that the three workstations did not stop due to lack of product.

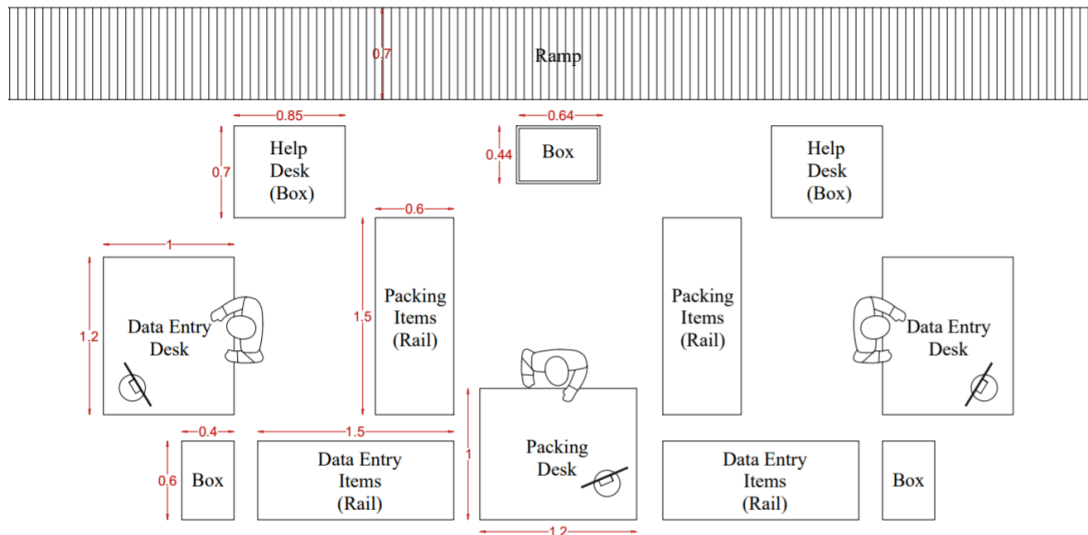


Figure 36 - Layouts created for the test “Split Scan Out” [m]

To ensure this test had no major impact on production, the planning team established that, if certain assumptions were not fulfilled every day, the test would have to be suspended. These assumptions were:

1. The WIP of Scan Out stations should not be superior to 900 items;
2. The number of items in process in which the Return Slot time was equal or superior to 3 should not exceed 800.

The analysis of Table 12 allows the conclusion that the assumptions were verified on each of the study days; therefore, there was no need to suspend it. There was not a day in which the work in process was superior to 900 items nor any day in which the number of items in process for more than 3 days in the production centre was superior to 800.

Table 12 - Verification of the assumptions for the “Split Scan Out” test

Day	1	2	3	4	5	6	7	8
Average WIP	541	408	425	360	365	647	553	440
Maximum WIP	842	504	685	515	762	857	859	595
Average Return Slot	0	3	1	10	1	193	180	65
Maximum Return Slot	0	4	2	20	1	219	320	119

After the test's execution, the times were organized and the average cycle times per day at Data Entry and Packing were calculated, as well as the targets achieved by collaborator. For this purpose, all the periods superior to 12 minutes were removed from the calculations, since it was observed that superior values corresponded to isolated errors that occurred during the process. Table 13 summarizes all the calculated values per test day. The spreadsheet used on each study day to record the test values can be seen in the Appendix C.

Table 13 - Summary table with the values reached by season and day

Day	1	2	3	4	5	6	7	8
Data Entry								
Minute per Article	2.44	2.28	1.96	1.93	1.95	2.43	2.26	2.13
Articles per Person	177	190	220	224	221	178	191	203
Packing								
Minute per Article	1.47	1.68	1.57	1.35	1.35	1.36	1.44	1.38
Articles per Person	293	257	275	320	320	318	300	313

Also through the test, it was possible to understand which manual tasks added value or not to the process, their respective periods and the movement periods of each collaborator. The study of tasks that added value to the process and their times can be consulted in Appendix D.

Finally, the current cycle times and the times after separation in Data Entry and Packing were organized according to genre, category of items; these can be observed in Table 14.

Data Entry of any item takes around 130 seconds to be processed, without distinguishing between different categories and genres. Each collaborator manages to process 28 products per hour, on average, according to the test.

The packing of different items takes 73 seconds on average to be performed. This value was calculated without distinguishing the periods between different categories and genres. Each collaborator of the Packing workstation can produce about 50 items per hour.

Table 14 - Calculated cycle times in the test by gender and category of items

		Actual	“Split Scan Out” Test		
		Scan Out	Data Entry	Packing	Total
Kids	Accessories	100	136	52	188
	Cloth	103	103	54	157
Men	Accessories	213	158	56	214
	Cloth	119	123	73	196
	Jewellery	176	79	40	119
	Stills	162	132	56	188
Women	Accessories	195	183	94	277
	Cloth	127	127	87	214
	Jewellery	148	158	52	210
	Stills	165	138	53	191
Unisex	Accessories	317	153	61	214
	Cloth	153	142	73	215
	Jewellery	270	110	54	164
	Stills	185	95	52	147

4.2.2 Data-Driven Insights and Opportunities to Improve

Separating the Scan Out in two different workstations presents several advantages, among which:

1. Specialized people enhancing their productivity;
2. Reduce the human error;
3. Simplify the process of formation;
4. Simplify and improve the recruitment;
5. Leveraging high stock times;
6. Possibility of creation, in the future, of a Size & Fit Specialist Team.

To realize the number of necessary resources, as well as the number of work posts and also if the existing stock times so far are enough to perform the activity, an analysis of the current scenario was performed. Three different and fictitious scenarios were created, being the difference between them the types of products that executed Data Entry before the normal time. The scenarios in study are:

Current Scenario

All the items execute Scan Out at the end of the process, and the latter does not present distinction between Data Entry and Packing.

First Scenario	Only the items that go directly from Scan In to Stills execute Data Entry after Scan In. All the remainder will get their data inserted in the software only after quality check.
Second Scenario	All the items that go through Live Model station during their course execute Data Entry before this station. The remaining items follow the regular process and execute Data Entry together with Packing.
Third Scenario	All processed items execute Data Entry after Scan In, reusing stock times for the next station. Packing is performed after quality check.

It was considered that only 78% of the items are effectively processed for all the scenarios in study. The remaining 22% correspond to duplicated items that do not execute Data Entry but need to be packed so they can return in good conditions to the respective boutique. When high peak is referenced in all the scenarios, it means that 2500 items arrived per day; low peak corresponds to 1500 items arrived per day.

Current Scenario

In the current scenario, as previously mentioned, all the items wait after being photographed for the photos' quality check in a specific stock area. After this, they proceed to Scan Out workstation to exit from the production centre and posteriorly be expedited to the respective boutique. At Scan Out, data insertion and product packing are executed at the same station by the same collaborator. The station occupies 105.12 m² and the layout may be observed in Figure 37.

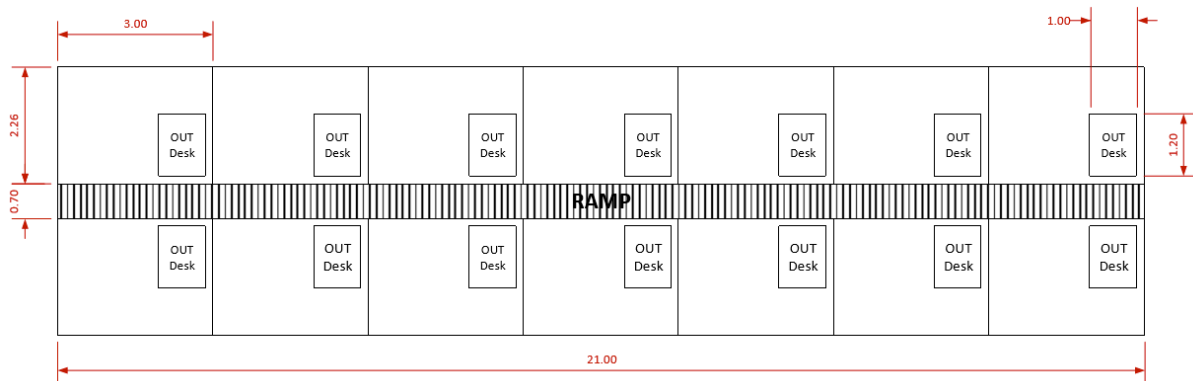


Figure 37 - Scan Out layout

Some variables of the current scenario were calculated, which are necessary for comparing between different scenarios. These variables are: the necessary time to perform this stage of the process, the number of necessary resources to perform it during the available period (2 shifts of 7.2 hours) and the number of necessary work posts. Scan Out's average cycle time in seconds was retrieved, knowing that the verified target per day was of 153 items per collaborator. Table 15 presents these data for the two peaks in question.

Table 15 - Calculated data for the current scenario

	Items Produced/Day	Average Cycle Time (Sec)	Time expended (h)	Resources/day	WorkStations
High Peak	2500	170	118.1	17	9
Low Peak	1500	170	70.8	10	5

First Scenario

A scenario with less impact but with some profits for the company was thought. In order to do so, a Data Entry workstation between Scan In and Stills was created, and its impact on production was studied. All accessory items that go directly to Stills went through this workstation, without passing through Live Model studios. The observation of **Erro! A origem da referência não foi encontrada.**, determines that the stock time of Stills workstation is 7.63 hours on average. The scenario takes advantage of this average period to execute Data Entry of the items before these are photographed in Stills.

The items that pass directly from Scan In to Stills were considered to represent 18.8% of the total. When performing the “Split Scan Out” test, it was verified that Data Entry of these types of items takes 133 seconds on average, and Packing takes 69 seconds.

A study to understand how many stations and resources would be necessary to perform Data Entry of these items was performed, considering the necessary period described in the previous chapter. The data were then organized on Table 16. In this case, only 18.8% of the items that were effectively produced (they are not duplicates) went through this station.

Table 16 - Data Entry table of items that go directly to ST

	Items Produced/Day	Average Cycle Time (Sec)	Time expended (h)	Resources/day	WorkStations
High Peak	368	133	13.5	2	1
Low Peak	219	133	8.1	2	1

The Scan Out post was necessarily reorganized, since the items that had already inserted information in the system only needed to be packed, while all the remainder had to perform all the process of data insertion in the system, analysis of possible defects and packing. The possibility of creating two lines at the space occupied currently by the Scan Out workstation was studied: one dedicated to the packing of Stills items and another destined to executing data entry and packing of the remaining items. The necessary number of resources and stations is presents on Table 17. By observing the two tables it is realized that, in total, 18 resources and 9 work posts are needed in high peak. In low peak, 11 resources and 6 work posts are necessary in total.

Table 17 - Table of Data Entry and Packing data in the first scenario

	Items Produced/Day	Average Cycle Time (Sec)	Time expended (h)	Resources/day	WorkStations
Packing of items directly to Stills					
High Peak	469	69	9.0	2	1
Low Peak	281	69	5.4	1	1
Data Entry and Packing of the remaining items					
High Peak	2031	170	95.9	14	7
Low Peak	1219	170	57.5	8	4

An analysis was performed to verify if the stock time of each slot in Stills stations is sufficient to execute Data Entry, considering that each item of the slot would take 133 seconds on average. This period was sufficient in 97% of the cases. The spreadsheet used to complete the analysis can be consulted in Appendix F.

Second Scenario

A new study was performed, involving all items that go through the station of Live Model. It is known that this station presents high stock times: the items wait 10.0 hours in stock areas of Live Model Women and 8.27 hours in Live Model Men, on average. The objective of the scenario is, once again, to take advantage of this periods to submit item information on the platform, therefore reducing the Scan Out cycle time, and consequently the processing time of products. The items that go through this station are all the clothes, jewellery and some accessories, representing 70.61% of the total produced items.

The study performed on the previous scenarios was repeated to understand how many work posts and resources would be necessary to execute Data Entry, considering that the items take around 134 seconds to execute this step (Table 18). This value was calculated using the following equation:

$$Cycle\ Time_{Data\ Entry}^{LM\ items} = \frac{\sum Total\ Sec/Day_{Data\ Entry}^{Clothing} + \sum Total\ Sec/Day_{Data\ Entry}^{Jewellery} + \sum Total\ Sec/Day_{Data\ Entry}^{Accessories}}{Items\ Produced * Volume\ (\%)_{LM\ items}} \quad (3)$$

where,

$$Total\ Sec/Day_{Data\ Entry} = Time_{Insertion} \times Items\ Produced \times Volume(\%) \quad (4)$$

It is noteworthy that the sum is composed by all the total times of Data Entry per day of the different genders: female, male and unisex. The values used in the respective calculation may be consulted in Appendix E.

Table 18 - Calculated data for Data Entry of items that go past LM station

	Items Produced/Day	Average Cycle Time (Sec)	Time expended (h)	Resources/day	WorkStations
High Peak	1384	134	51.5	8	4
Low Peak	830	134	30.9	5	3

Although all items need to be packed, the items that went through Live Model during the process did not need to execute Data Entry at the ending, thus the Scan Out post was necessarily reorganized once again. As in the previous scenario, two lines were thought, one dedicated to packing of the items that executed LM and another where the remaining items executed Data Entry and Packing combined before being expedited to the respective boutique (Table 19). After observing the results of the performed test and executing the calculations using the equations (5) and (6), it was considered that the packing of items that go through Live Model takes 82 seconds on average.

$$Cycle\ Time_{Packing}^{LM\ Items} = \frac{\sum Total\ Sec/Day_{Packing}^{Clothing} + \sum Total\ Sec/Day_{Packing}^{Jewellery} + \sum Total\ Sec/Day_{Packing}^{Accessories}}{Items\ Arrived * Volume\ (\%)_{LM\ Items}} \quad (5)$$

where,

$$Total\ Sec/Day_{Packing} = Time_{Packing} \times Items\ Arrived \times Volume(\%) \quad (6)$$

Again, the total sum encompasses all the total times of Packing per day of the different genres: female, male and unisex.

Table 19 - Table of Data Entry and Packing data in the second scenario

	Items Produced/Day	Average Cycle Time (Sec)	Time expended (h)	Resources/day	WorkStations
Packing of all Live Model items					
High Peak	1765	82	40.0	6	3
Low Peak	1059	82	24.0	4	2
Data Entry and Packing of the remaining items					
High Peak	735	170	34.7	5	3
Low Peak	441	170	20.8	3	2

Lastly, an analysis was performed to verify if the stock time of each slot in Live Model would be sufficient to introduce the item information on the software, considering that each item of the slot would take 134 seconds on average. The observation of the spreadsheets shown in Appendixes G and H led to determine that, in 99% of the observed cases, the stock time in LM men would be sufficient, an identical percentage to the stock time in LM women.

Third Scenario

The complete separation of the Scan Out station in two new stations – Data Entry and Packing – was studied in the last scenario. Data Entry of all processed and non-duplicated items would be executed after Scan In, making use of the long stock times that exist between this station and the remainder. As packing needs a quality check and the full production process to be complete, it would remain at the end of the process, immediately before expedition. This scenario allows to specialize collaborators in specific functions and it simplifies the process of recruiting and training, since the collaborator is responsible only for Data Entry or Packing.

Data Entry and Packing workstations receive per day, in high peak, 1960 and 2500 items on average, respectively. Duplicated items also need to be verified for the existence of defects and to be packed.

The study performed in the previous scenarios was repeated to understand how many work posts and resources would be necessary to execute Data Entry and Packing of the items, considering that the items take 130 seconds to complete the first step and 73 seconds to complete the second, on average. These periods were estimated using the performed test, without distinguishing different types of categories and genres (Table 20).

Table 20 - Table of Data Entry and Packing data in the third scenario

	Items Produced/Day	Average Cycle Time (Sec)	Time expended (h)	Resources/day	WorkStations
Data Entry					
High Peak	1960	130	70.6	10	5
Low Peak	1176	130	42.4	6	3
Packing					
High Peak	2500	73	50.3	7	4
Low Peak	1500	73	30.2	5	3

Finally, an analysis was performed to verify if the stock period of each slot in the workstations of Stills, Live Model and Flat Kids would be sufficient to introduce the items' information on the software, considering that each product takes 130 seconds on average. The observation of Appendixes F, G, H and I enable the determination that in 98.7% of the observed cases during the year of 2017, the stock time after Scan In would be sufficient to execute the activity of item Data Entry in parallel.

4.2.3 Comparison between the Different Scenarios Studied

In the subchapter 4.2.2, a comparison was made between the different scenarios. The processing time of all the slots produced in 2017 was studied, by comparing the original time with the time in each of the scenarios, and the percentage of time saved was calculated. These performance indicators can be seen in Table 21. The Excel sheet used to make this comparison can be seen in Appendix J.

Table 21 - Comparison between scenarios

	Current Scenario	First Scenario	Second Scenario	Third Scenario
Process Time [h]	8.88	8.44	8.20	7.74
Saving [%]	-	5.3 %	7.8 %	13.0 %

5 Conclusions and Future Work

5.1 Conclusions

The work developed in the present Master's Dissertation had as a primary objective the statistical analysis of data concerning the digital production of Farfetch in previous months, studying strategical improvements of processes.

A process mapping of general production was initially performed, and for each workstation a more detailed mapping, to identify improvement opportunities. In addition, analyses of stock and processing periods were performed, as well as an analysis of variability. After the retrieval of data regarding the initial situation, the existence of some issues was proved.

A study in the Live Model workstation was after performed to assess how the increase in the number of items, categories and season influenced the cycle time. There were other variables that influenced these periods, however, they were not quantifiable. Among the diverse methods used to estimate cycle time, a statistical method was chosen – multivariate linear regression. It was possible to determine that the number of items and categories contained in a slot increases the cycle time of Live Model Women, as already expected. It is possible to get more reduced periods in low season comparing to high season, with the same conditions. Using this model of lead time forecast, it was found that a reduction of category types in the slot can lead to a decrease of 1.9% in the cycle time within the workstation. If the slot goes from one category to nine, the maximum number of categories that go through this station, the cycle time can increase about 13.2%.

A solution was thought in which each work post would be responsible for the photo production of two different types of categories at most. In this case, instead of being organized by slot, the items would be organized by categories. To achieve the goal, a new triage station at the end of the process would have to be created to separate once more the clothes by slots and guarantee that there were no flaws in the delivery of the items to the respective boutique.

After an attentive observation of the processing maps previously elaborated, it was shown that there were activities, at least in the Scan Out workstation, that could be performed at any stage of the process. Scan Out was then divided in two new stations: Data Entry, where the items' information would be inserted in the system; and Packing station, where a possible existence of pieces' defects would be verified, and also where the items would be packed to proceed to the boutique.

A "Split Scan Out" test was performed during 8 work days, in which the periods of these two new stations were studied. A layout was created for the study in question, being the two stations of Data Entry responsible for supplying the Packing station. It was decided that when the Work in Process in this station was superior to 900 or the item number with a return slot superior to 3 days surpassed 800, the tests would be suspended for a determined period, so as

not to impact production. Cycle times were calculated for the two stations with this test, by category and genre of the pieces.

Three different scenarios were studied and compared to the current scenario in terms of stations number and necessary resources. The objective of this study was to execute an activity in parallel while the items waited to be photographed, thus optimizing the stock times and reducing the processing time.

The first scenario proposed that only the items that went directly from Scan In to Stills would execute Data Entry in advance. These items represented 18.8% of the total stock. The second scenario proposed quite the same, but only with the items that went through Live Model, i.e., 70.6% of the total. The third and last scenario, that supposedly represented a greater saving of processing times, proposed the complete separation of Data Entry and Packing. Data Entry would be always executed after Scan In, while Packing would be executed after quality check of the previously produced photos.

5.2 Future Work

Following the work performed during the current dissertation, it is considered that, in a near future, the first step shall include executing tests to understand the true impact of these two possible improvements, and studies of their viability.

After developing the project, the performed studies demonstrated that there are some issues that can and must be explored in the future. Some studies and improvements that can be explored are explained below:

1. Initially, it was of extreme importance to execute a study on how all the existing programs could be integrated by creating one single software, simple to understand by all collaborators and that could be used by them through all stages of the process.
2. There are stations that require an additional caution, since they have been demonstrated as vital within the processing times. These are Flat and Live Model Women/Men stations. In terms of cycle time, they are considered the bottlenecks of the process, due to the amount of time spent and the number of items being produced in each one of them. Live Model Women and Men stations are the ones bringing more costs and more collaborators, therefore it would be of interest to realize if their performance allows to add functions to the ones currently executed.
3. The study of ways to diminish the variability of the process may be of great value. Although there are greater standard deviations in terms of variability in Scan Out station, this is not the only one suffering from this, regarding both cycle times and number of produced items.
4. As the company presents production peaks too pronounced, there is a need for some extra collaborators only in certain times of the year. It is understandable that the collaborators' training has a great impact in terms of their productivity. After observing this, it is necessary to elaborate a study in which, during a certain period of time, only experienced collaborators would be producing, and quantify the respective earnings. From this point, a cost-benefit analysis of having a full-time collaborator could be performed, even if he presents significant periods of work stopped during certain months.
5. When studying the current situation of the company, it was possible to realize that there are approximately 20 brands representing around 25% of the production volume. These brands represent a sales volume highly significant when compared with the

others. Following this, the idea of creating two lines of production arose. One considered normal that would maintain the already existing terms (3 work days of return slot on average) and a new quicker line to these 20 brands that allowed these products to return at the earliest to the boutiques and the items to be placed for sale sooner on the online platform. It would thus be necessary to study the amount of resources required to decrease production times, the costs of such resources and which changes would be necessary in term of logistics.

6 References

- Backus, Phillip, Mani Janakiram, Shahin Mowzoon, George C. Runger, e Amit Bhargava. 2006. "Factory cycle-time prediction with a data-mining approach." *IEEE Transactions on Semiconductor Manufacturing*.
- Baur, Nina, e Siegfried Lamnek. 2007. "Multivariate analysis." Em *The Blackwell Encyclopedia of Sociology*, de Nina Baur e Siegfried Lamnek. Blackwell Publishing.
- Chung, Shu Hsing, e Hung Wen Huang. 2002. "Cycle time estimation for wafer fab with engineering lots." *IIE Transactions (Institute of Industrial Engineers)*.
- Dauriz, Linda, Nathalie Remy, e Nicola Sandri. 2014. "Luxury shopping in the digital age." 6.
- Fayyad, Usama, Gregory Piatetsky-Shapiro, e Padhraic Smyth. 1996. "From Data Mining to Knowledge Discovery in Databases."
- Feld, William M. 2000. *Lean Manufacturing: Tools, Techniques and How to Use Them*. St. Lucie Press.
- Field, Andy. 2017. *Discovering Statistics Using SPSS Statistics*. SAGE Publications.
- Gagné, Denis, e Simon Ringuette. 2008. *Object Management Group: Business Process Model and Notation*.
- Harmon, Paul. 2010. *Business Process Change: A Guide for Business Managers and BPM and Six Sigma Professionals*. 2nd. Morgan Kaufmann.
- Hicks, B. J. 2007. "Lean information management: Understanding and eliminating waste." *International Journal of Information Management*.
- Hopp, Wallace J., e Mark L. Spearman. 2004. "To Pull or Not to Pull: What Is the Question?" *Manufacturing & Service Operations Management*.
- LaerdStatistics. 2015. *Multiple regression using SPSS Statistics*. <https://statistics.laerd.com/>.
- Marrelli, Anne F. 2005. "The Performance Technologist's Toolbox: Process Mapping." *Performance Improvement* 44 (5): 5.
- Mifflin, Houghton. 1975. *The American Heritage Dictionary of the English Language*. Editado por William Morris.
- Nyhuis, Peter, Gregor Von Cieminski, Andreas Fischer, e Klaus Feldmann. 2005. "Applying simulation and analytical models for logistic performance prediction." *CIRP Annals - Manufacturing Technology*.
- Pearn, W. L., Yu Ting Tai, e J. H. Lee. 2009. "Statistical approach for cycle time estimation in semiconductor packaging factories." *IEEE Transactions on Electronics Packaging Manufacturing*.
- Pojasek, Robert B. 2014. "Understanding processes with hierarchical process mapping." 46.

- Puchkova, Alena, Julien Le Romancer, e Duncan McFarlane. 2016. "Balancing Push and Pull Strategies within the Production System." *IFAC-PapersOnLine*.
- Remy, Nathalie Remy, Marco Catena, e Benjamin Durand-Servoingt. 2015. "Digital Inside: Get wired for the ultimate luxury experience." *McKinsey & Company* 17.
- Tai, Y. T., W. L. Pearn, e J. H. Lee. 2012. "Cycle time estimation for semiconductor final testing processes with Weibull-distributed waiting time." *International Journal of Production Research*.
- Wesner, John, Jeffrey Hiatt, e David Trimble. 1994. *Winning with Quality*. MA: Addison Wesley Longman.
- White, Stephen A., e Derek Miers. 2008. *BPMN: Modeling and Reference Guide*. Future Strategies Inc.

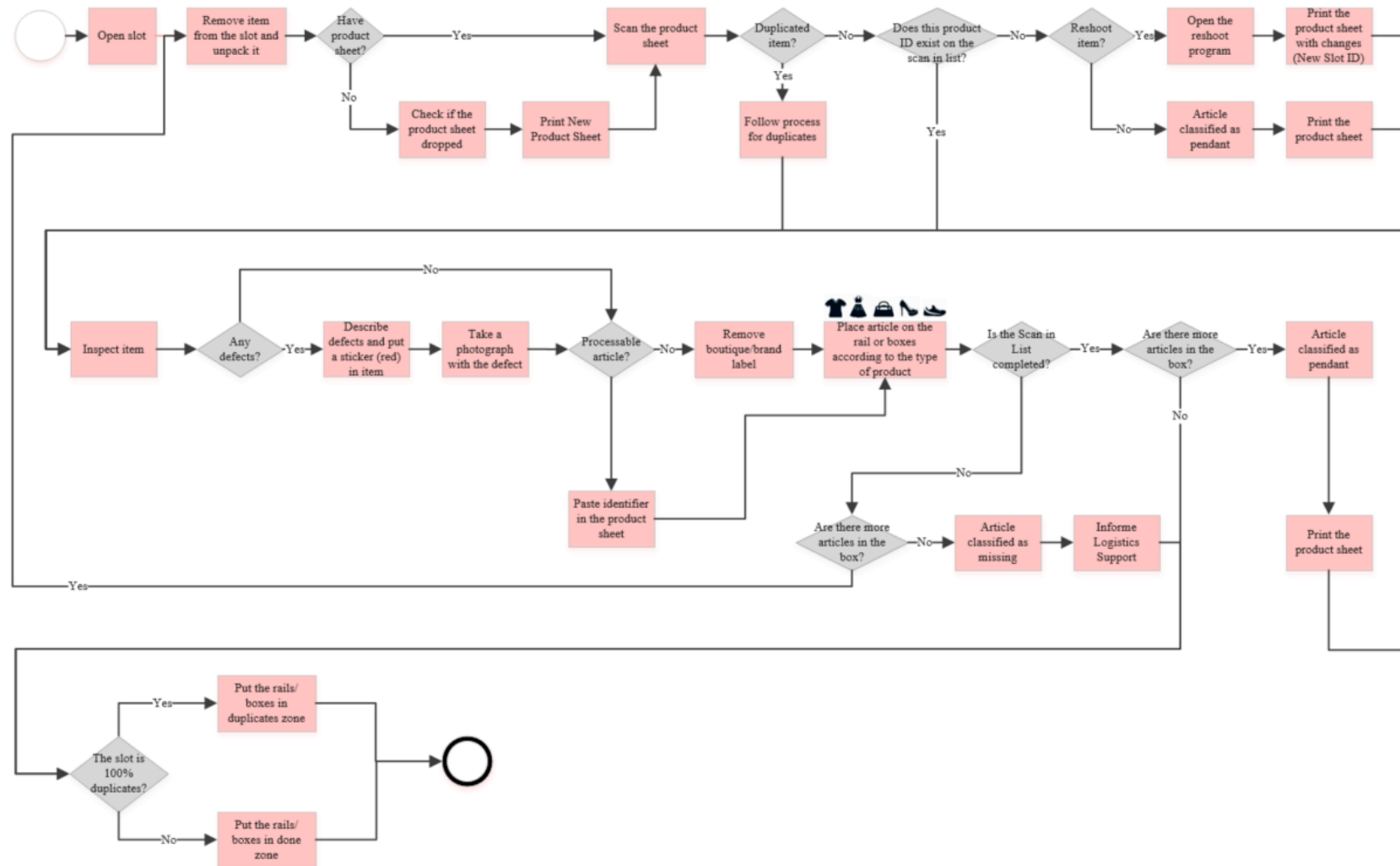
APPENDIX A: Number of items produced per hour in workstations

Variability in different workstations of the production centre on a given day

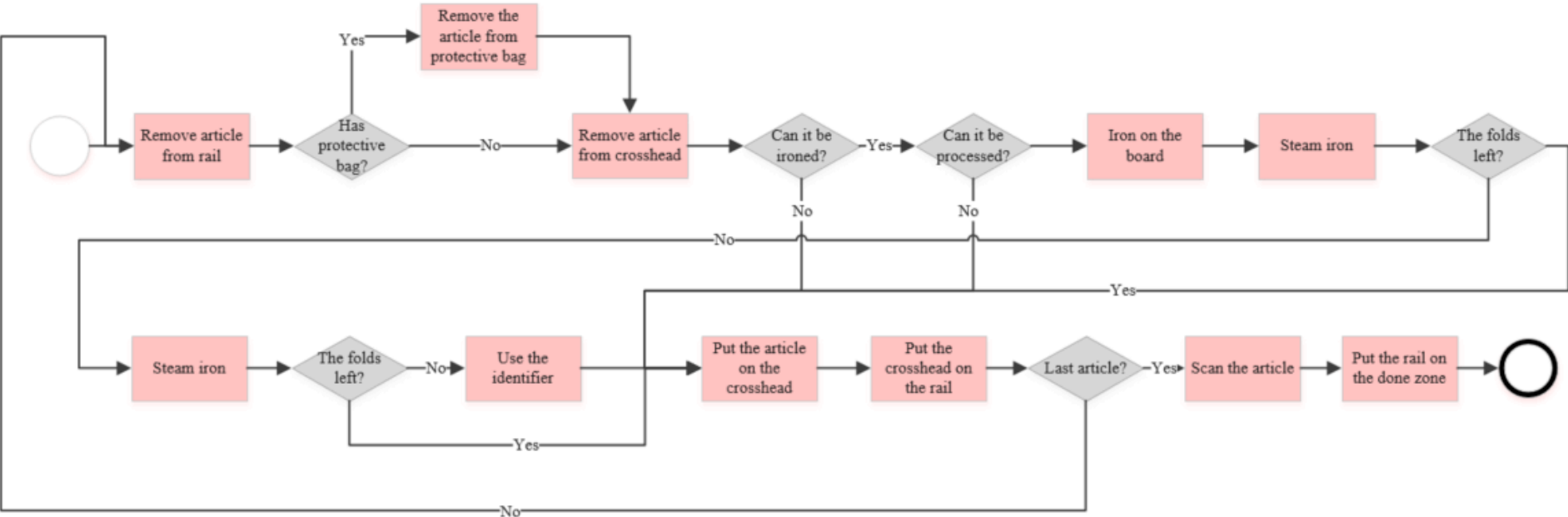


APPENDIX B: Process Maps of different workstations

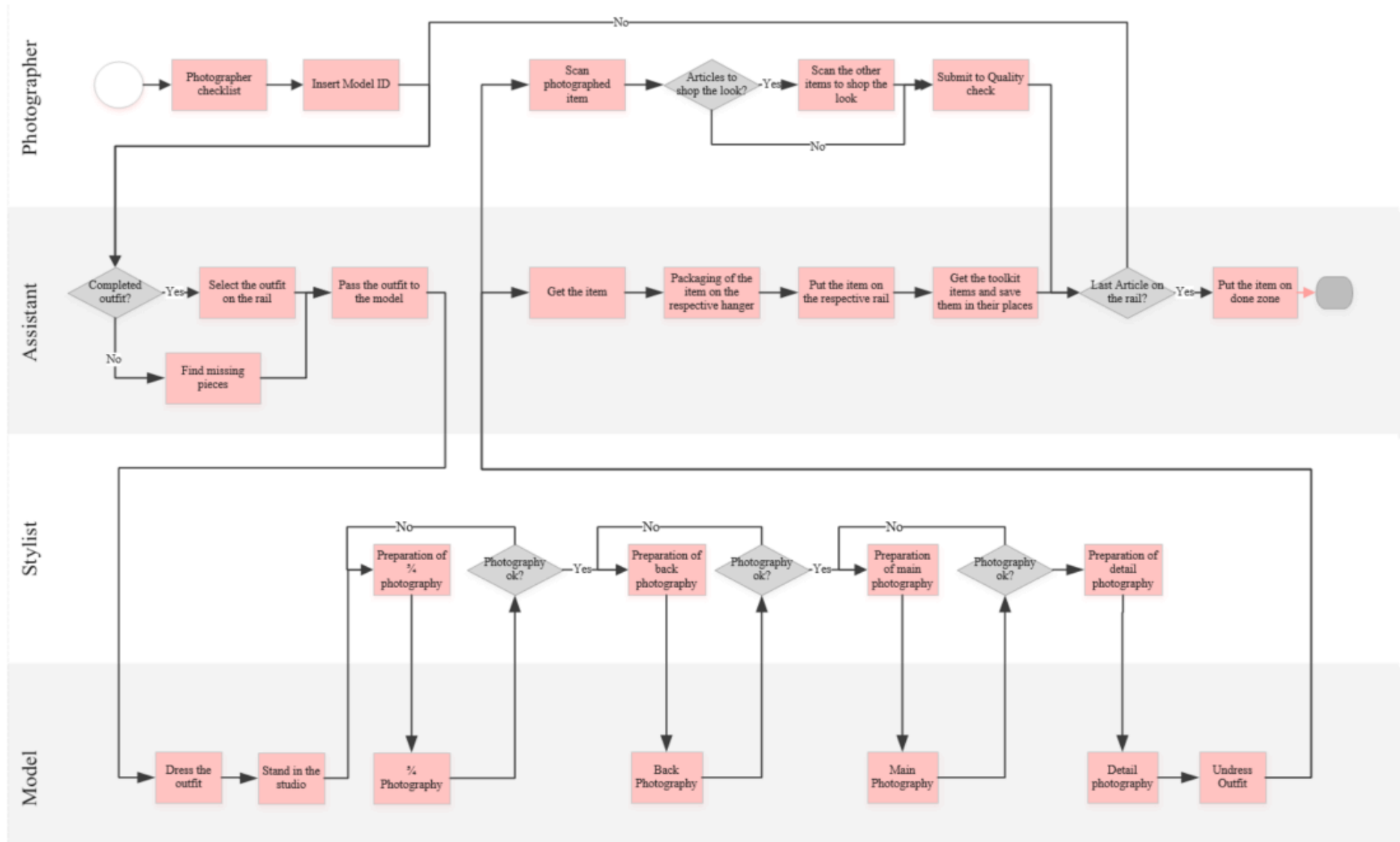
Scan In Process Map



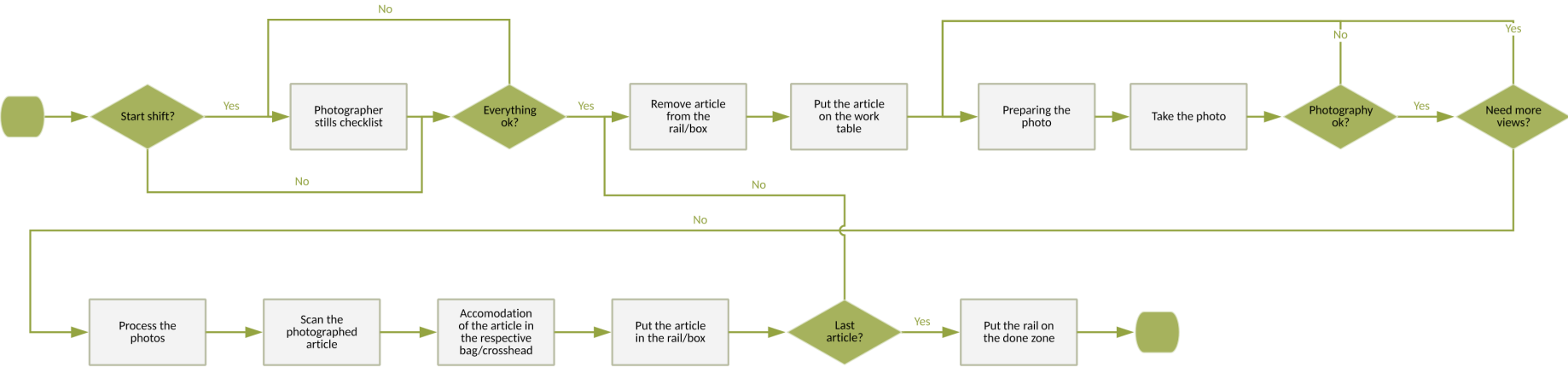
Iron Process Map



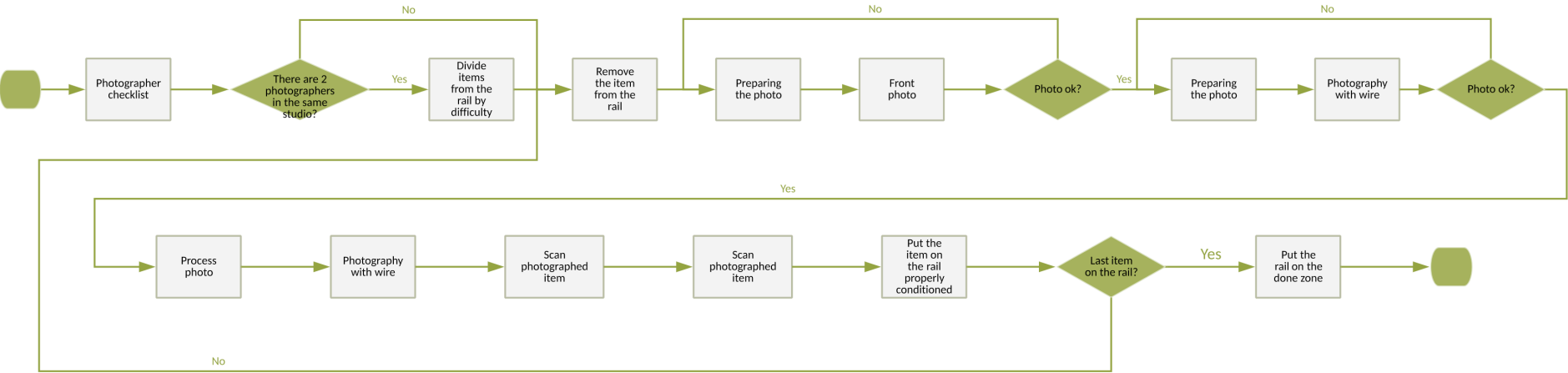
Live Model Process Map



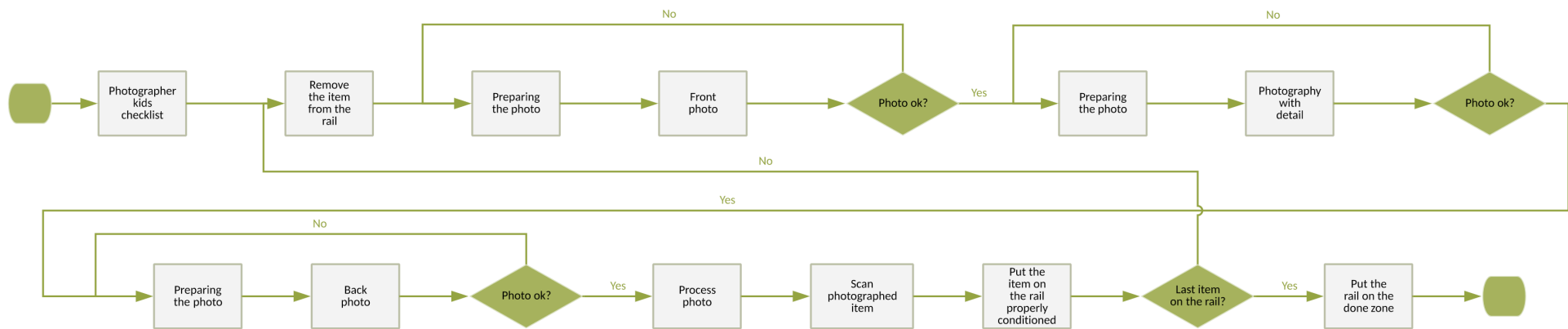
Stills Process Map



Flat Process Map



Kids Process Map



APPENDIX C: “Split Scan Out” Test Excel Sheet for each study day

(Day 4 Example)

Ready for process (MÁX=900)		Return Slot 3ºdia (MÁX=800)		Nº de rails em espera			Tempo de inserção		
Row Labels	Average of WIP	Average	10,125	Average	17	2	Considerado	Sim	
		Máx	20	Máx	21	7	Hour	(All)	
2017-10-12	360								
00:00:00.0000000	410								
00:10:00.0000000	410								
00:20:00.0000000	410								
00:30:00.0000000	410								
00:40:00.0000000	410								
00:50:00.0000000	410								
01:00:00.0000000	410								
01:10:00.0000000	410								
01:20:00.0000000	410								
01:30:00.0000000	410								
01:40:00.0000000	410								
01:50:00.0000000	410								
02:00:00.0000000	410								
02:10:00.0000000	410								
02:20:00.0000000	410								
02:30:00.0000000	410								
02:40:00.0000000	410								
02:50:00.0000000	410								
03:00:00.0000000	410								
03:10:00.0000000	410								
03:20:00.0000000	410								
03:30:00.0000000	410								
03:40:00.0000000	410								
03:50:00.0000000	410								
04:00:00.0000000	410								
04:10:00.0000000	410								
04:20:00.0000000	410								
04:30:00.0000000	410								
04:40:00.0000000	410								
04:50:00.0000000	410								
05:00:00.0000000	410								
05:10:00.0000000	410								
05:20:00.0000000	410								
05:30:00.0000000	410								
05:40:00.0000000	410								
05:50:00.0000000	410								
06:00:00.0000000	410								
06:10:00.0000000	410								
06:20:00.0000000	404								
06:30:00.0000000	372								
06:40:00.0000000	350								
06:50:00.0000000	325								
07:00:00.0000000	303								
07:10:00.0000000	263								
07:20:00.0000000	253								
07:30:00.0000000	238								
07:40:00.0000000	206								

Artigos produzidos por pessoa		
Hour	(All)	
Row Labels	Count of ArtigoID	Nº bloqueados
Pessoa 1	214	6
Pessoa 2	222	1
Grand Total	436	7

Hour	(blank)	
PACKING		Embalados Registrados
Row Labels	Count of ArtigoID	Pessoa 3
Pessoa 3	9	Mafalda (89)+ Maria (13)
Grand Total	9	2º Turno
		429

Produtividade							
Colaborador	Nº artigos Real	Horas produtivas		Artigos/Hora	Eficiencia/ Tableau	Utilização/ Tableau	Produtividade/ Target/ Tableau
Pessoa 1	214	7:06:00	7.10	30	114%	99	214/190
Pessoa 2	221	6:52:00	6.87	32	121%	96	221/95
Pessoa 3	309	6:57:00	6.95	44	168%	97	309/190

Inserção de dados	435	13:58:00	13.97	31
Embalagem	309	6:57:00	6.95	44

	Hora de Inicio	Hora de Fim (Pausa)	Tempo Trabalhado 1	Hora de Inicio	Hora de Fim	Tempo Trabalhado 2	Hora de Inicio	Hora de Fim	Tempo Trabalhado 3
Pessoa 1	06:11	08:15	2:04:00	08:28	11:15	2:47:00	11:35	11:50	0:15:00
Pessoa 2	06:19	08:15	1:56:00	08:34	11:15	2:41:00	11:35	11:50	0:15:00
Pessoa 3	06:18	08:15	1:57:00	08:30	11:15	2:45:00	11:35	11:50	0:15:00

	Hora de Inicio	Hora de Fim	Tempo Trabalhado 4	Hora de Inicio	Hora de Fim	Tempo Trabalhado 5	Hora de Inicio	Hora de Fim	Tempo Trabalhado 6
Pessoa 1	11:50	13:50	2:00:00			0:00:00			0:00:00
Pessoa 2	11:50	13:50	2:00:00			0:00:00			0:00:00
Pessoa 3	11:50	13:50	2:00:00			0:00:00			0:00:00

	1ª Hora	2ª Hora	3ª Hora	4ª Hora	5ª Hora (45+15)	6ª Hora	7ª Hora
Pessoa 1	24	30	32	32	28	28	40
Pessoa 2	34	27	26	35	36	26	37
Pessoa 3	39	38	39	45	43	51	54

Para produtividade no programa somar 48 minutos (20 de almoço 10 de pausa, 18 minutos para reuniões)

APPENDIX D: Tasks that add value or not to Data Entry and Packing

DATA ENTRY														
Workstation	Data Entry	Article Type	All	Cycle Time (Sec)	129	Quantity (Shift)	2500							
		Working Hours	7,2	Target/person	201	Resources Number	13							
								Manual Tasks Times (Seconds)						
								Value Add	Non Value Add	Walk	Number of Articles	Value Add	Non Value Add	Walk
Description														
1	Reach the rail/box of the stock area (10 items per rail/box)									3	10	0	0,0	0,3
2	Move the rail to the processed item area									3	10	0	0,0	0,3
3	Pick up items that are not processed								5		1	0	5,0	0,0
4	Put the item on the work desk								3		1	0	3,0	0,0
5	Check the label and register the information on Retail							113		1	113	0,0	0,0	
6	Pick up the item of the work table								3		1	0	3,0	0,0
7	Put the item in the processed articles rail/box								3		1	0	3,0	0,0
8	Identify the processed articles rail/box								5		10	0	0,5	0,0
9	Pick up the processed articles rail/box								3		10	0	0,3	0,0
10	Place the rail/box in the complete work area									3	10	0	0,0	0,3
11												0	0,0	0,0
12												0	0,0	0,0
13												0	0,0	0,0
14												0	0,0	0,0
15												0	0,0	0,0
16												0	0,0	0,0
												113	15	1

FARFETCH

Standard Work

PACKING														
Workstation	Packing	Article Type	All	Cycle Time (Sec)	86	Quantity (Shift)	2500							
		Working Hours	7,2	Target/Person	301	Resources Number	9							
								Manual Tasks Time (Seconds)						
Description								Value Add	Non Value Add	Walk	Number of Articles	Value Add	Non Value Add	Walk
1	Reach rail/box with items that need to be packed from the waiting area (10 items per rail/box)													
2	Move rail to the processing area													
3	Reach item to pack of the rail / box													
4	Put the item on the work desk													
5	Involve paper in order to protect the item													
6	Place the item in its plastic bag													
7	Close the plastic bag with tape													
8	Pick up the item that is on the work desk													
9	Place the packed item in the final box													
10	Identify and place the box on the ramp													
11	Reach the empty rail/box of the processing area													
12	Move the rail/box to the empty rails/box area													
13														
14														
15														
16														
												70	15	2

FARFETCH

Standard Work

APPENDIX E: Spreadsheet with the values obtained in the test performed

Phase 1 - Split Scan - Out (Data Entry/Packing)

		Actual	"Split Scan Out" Teste				Items Arrived		2500			
		Month 7 and 8					Items Produced		1960	78%		
		ScanOut (sec)	Insertion (sec)	Packing (sec)	Total (sec)	Difference (sec)	% Volume 9 month 2017	Items Number Arrive / Day	Total Sec / Day Today	Total Sec / Day DATA Proposal	Total Sec / Day PACKING Proposal	Target DATA Proposal
Kids	Accessories	100	136	52	188	-88	2,92%	73	7298	7780	3795	190
	Clothing	103	103	54	157	-54	5,98%	149	15393	12066	8070	251
Men	AccMen	213	158	56	214	-1	1,6%	39	8337	4848	2192	164
	Cloth	119	123	73	196	-77	20,6%	516	61371	49726	37648	210
	Jewellery	176	79	40	119	57	0,3%	7	1182	416	267	328
	Stills	162	132	56	188	-26	6,5%	162	26323	16814	9099	196
Women	Accessories	195	183	94	277	-82	8,51%	213	41502	30531	20006	141
	ClothWomen	127	127	87	214	-87	36,74%	919	116659	91449	79916	204
	Jewellery	148	158	52	210	-62	1,97%	49	7272	6086	2555	164
	Stills	165	138	53	191	-26	10,81%	270	44604	29243	14327	187
Unisex	AccUnisex	317	153	61	214	103	0,32%	8	2529	957	487	169
	ClothUnisex	153	142	73	215	-62	0,42%	11	1624	1182	775	182
	JewelleryUnisex	270	110	54	164	106	0,18%	4	1192	381	238	235
	Stills	185	95	52	147	38	1,45%	36	6693	2694	1881	272
		Sec/Article	% Rentabilidad	Target	Target T.	Article/Hour						
DATA Insertion		130	0,92	200	180	28						
Packing		73	0,79	358	280	50						

TARGET BY TYPE OF ITEM

	% Volume 9 month 2017	DATA ENTRY Sec/Article	DATA ENTRY Target	DATA ENTRY Target T.	DATA ENTRY Article/Hour
DATA Acc	13,32%	169	153	140	21
DATA Cloth	63,77%	124	209	190	29
DATA Jewellery	2,41%	146	177	160	25
DATA Stills	18,76%	133	195	175	27

	PACKING Sec/Article	PACKING Target	PACKING Target T.	PACKING Article/Hour
101	255	200	35	
101	256	235	36	
65	400	365	56	
69	376	295	52	

TARGET BY GENDER

	% Volume 9 month 2017	DATA ENTRY Sec/Article	DATA ENTRY Target	DATA ENTRY Target T.	DATA ENTRY Article/Hour
DATA Kids	8,90%	114	227	205	32
DATA Men	28,96%	127	204	185	28
DATA Women	58,03%	138	187	170	26
DATA Unisex	2,37%	112	230	210	32

	PACKING Sec/Article	PACKING Target	PACKING Target T.	PACKING Article/Hour
	53	485	380	67
	68	381	350	53
	81	321	295	45
	57	453	355	63

APPENDIX F: Excel Sheet with analysis of stock time in Stills and Data Entry times

	H	I	J	K	L	M	N	Q	R	S	T
1		Slot ID	Number of Items	Consider?	Waiting Time (h)	Data Entry (s)	Data Entry (h)	Enough?			
2		35705	20	S	1,50	2652	0,74	YES		Nº of slots that the stock time is enough	2648
3		41277	10	S	31,69	1326	0,37	YES		Nº of slots that the stock time is not enough	70
4		43034	9	S	31,83	1193	0,33	YES		Number of slots in study	2718
5		44112	13	S	18,54	1724	0,48	YES			
6		44147	16	S	7,15	2122	0,59	YES			
7		44567	30	S	33,60	3978	1,11	YES			
8		44568	16	S	5,28	2122	0,59	YES			
9		44569	16	S	39,71	2122	0,59	YES			
10		44726	10	S	-	1326	0,37	-			
11		44762	16	S	33,39	2122	0,59	YES			
12		44804	22	S	9,79	2917	0,81	YES			
13		44896	26	S	-	3448	0,96	-			
14		44897	43	S	-	5702	1,58	-			
15		44898	50	S	-	6630	1,84	-			
16		45167	47	S	8,29	6232	1,73	YES			
17		45168	31	S	1,95	4111	1,14	YES			
18		45169	36	S	7,53	4774	1,33	YES			
19		45314	13	S	9,51	1724	0,48	YES			
20		45321	29	S	15,84	3846	1,07	YES			
21		45337	30	S	12,30	3978	1,11	YES			
22		45338	49	S	3,08	6498	1,80	YES			
23		45339	30	S	24,48	3978	1,11	YES			
24		45340	24	S	7,97	3183	0,88	YES			
25		45412	12	S	1,69	1591	0,44	YES			
26		45537	13	S	36,67	1724	0,48	YES			
27		45693	9	S	0,85	1193	0,33	YES			
28		45706	11	S	-	1459	0,41	-			
29		46246	10	S	16,19	1326	0,37	YES			
30		46544	9	S	2,51	1193	0,33	YES			
31		46553	21	S	4,82	2785	0,77	YES			
32		46634	13	S	31,04	1724	0,48	YES			
33		46709	11	S	9,43	1459	0,41	YES			
34		46790	10	S	4,79	1326	0,37	YES			
35		46792	10	S	9,22	1326	0,37	YES			
36		46797	14	S	2,54	1856	0,52	YES			
37		46799	11	S	0,84	1459	0,41	YES			
38		46882	10	S	40,07	1326	0,37	YES			
39		46883	10	S	40,07	1326	0,37	YES			
40		46884	10	S	40,07	1326	0,37	YES			
41		46885	10	S	40,07	1326	0,37	YES			
42		46886	10	S	40,07	1326	0,37	YES			
43		46887	10	S	40,07	1326	0,37	YES			
44		46888	10	S	40,07	1326	0,37	YES			
45		46889	10	S	40,07	1326	0,37	YES			
46		46890	10	S	40,07	1326	0,37	YES			
47		46891	10	S	40,07	1326	0,37	YES			
48		46892	10	S	40,07	1326	0,37	YES			
49		46893	10	S	40,07	1326	0,37	YES			
50		46894	10	S	40,07	1326	0,37	YES			
51		46895	10	S	40,07	1326	0,37	YES			
52		46896	10	S	40,07	1326	0,37	YES			
53		46897	10	S	40,07	1326	0,37	YES			
54		46898	10	S	40,07	1326	0,37	YES			
55		46899	10	S	40,07	1326	0,37	YES			
56		46900	10	S	40,07	1326	0,37	YES			
57		46901	10	S	40,07	1326	0,37	YES			
58		46902	10	S	40,07	1326	0,37	YES			
59		46903	10	S	40,07	1326	0,37	YES			
60		46904	10	S	40,07	1326	0,37	YES			
61		46905	10	S	40,07	1326	0,37	YES			
62		46906	10	S	40,07	1326	0,37	YES			
63		46907	10	S	40,07	1326	0,37	YES			
64		46908	10	S	40,07	1326	0,37	YES			
65		46909	10	S	40,07	1326	0,37	YES			
66		46910	10	S	40,07	1326	0,37	YES			
67		46911	10	S	40,07	1326	0,37	YES			
68		46912	10	S	40,07	1326	0,37	YES			
69		46913	10	S	40,07	1326	0,37	YES			
70		46914	10	S	40,07	1326	0,37	YES			
71		46915	10	S	40,07	1326	0,37	YES			
72		46916	10	S	40,07	1326	0,37	YES			
73		46917	10	S	40,07	1326	0,37	YES			
74		46918	10	S	40,07	1326	0,37	YES			
75		46919	10	S	40,07	1326	0,37	YES			
76		46920	10	S	40,07	1326	0,37	YES			
77		46921	10	S	40,07	1326	0,37	YES			
78		46922	10	S	40,07	1326	0,37	YES			
79		46923	10	S	40,07	1326	0,37	YES			
80		46924	10	S	40,07	1326	0,37	YES			
81		46925	10	S	40,07	1326	0,37	YES			
82		46926	10	S	40,07	1326	0,37	YES			
83		46927	10	S	40,07	1326	0,37	YES			
84		46928	10	S	40,07	1326	0,37	YES			
85		46929	10	S	40,07	1326	0,37	YES			
86		46930	10	S	40,07	1326	0,37	YES			
87		46931	10	S	40,07	1326	0,37	YES			
88		46932	10	S	40,07	1326	0,37	YES			
89		46933	10	S	40,07	1326	0,37	YES			
90		46934	10	S	40,07	1326	0,37	YES			
91		46935	10	S	40,07	1326	0,37	YES			
92		46936	10	S	40,07	1326	0,37	YES			
93		46937	10	S	40,07	1326	0,37	YES			
94		46938	10	S	40,07	1326	0,37	YES			
95		46939	10	S	40,07	1326	0,37	YES			
96		46940	10	S	40,07	1326	0,37	YES			
97		46941	10	S	40,07	1326	0,37	YES			
98		46942	10	S	40,07	1326	0,37	YES			
99		46943	10	S	40,07	1326	0,37	YES			
100		46944	10	S	40,07	1326	0,37	YES			
101		46945	10	S	40,07	1326	0,37	YES			
102		46946	10	S	40,07	1326	0,37	YES			
103		46947	10	S	40,07	1326	0,37	YES			
104		46948	10	S	40,07	1326	0,37	YES			
105		46949	10	S	40,07	1326	0,37	YES			
106		46950	10	S	40,07	1326	0,37	YES			
107		46951	10	S	40,07	1326	0,37	YES			
108		46952	10	S	40,07	1326	0,37	YES			
109		46953	10	S	40,07	1326	0,37	YES			
110		46954	10	S	40,07	1326	0,37	YES			
111		46955	10	S	40,07	1326	0,37	YES			
112		46956	10	S	40,07	1326	0,37	YES			
113		46957	10	S	40,07	1326	0,37	YES			
114		46958	10	S	40,07	1326	0,37	YES			
115		46959	10	S	40,07	1326	0,37	YES			
116		46960	10	S	40,07	1326	0,37	YES			
117		46961	10	S	40,07	1326	0,37	YES			
118		46962	10	S	40,07	1326	0,37	YES			
119		46963	10	S	40,07	1326	0,37	YES			
120		46964	10	S	40,07	1326	0,37	YES			
121		46965	10	S	40,07	1326	0,37	YES			
122		46966	10	S	40,07	1326	0,37	YES			
123		46967	10	S	40,07	1326	0,37	YES			
124		46968	10	S	40,07	1326	0,37	YES			
125		46969	10	S	40,07	1326	0,37	YES			
126		46970	10	S	40,07	1326	0,37	YES			
127		46971	10	S	40,07	1326	0,37	YES			
128		46972	10	S	40,07	1326	0,37	YES			
129		46973	10	S	40,07	1326	0,37	YES			
130		46974	10	S	40,07	1326	0,37	YES			
131		46975	10	S	40,07	1326	0,37	YES			
132		46976	10	S	40,07	1326	0,37	YES			
133		46977	10	S	40,07	1326	0,37	YES			
134		46978	10	S	40,07	1326	0,37	YES			
135		46979	10	S	40,07	1326	0,37	YES			
136		46980	10	S	40,07	1326	0,37	YES			
137		46981	10	S	40,07	1326	0,37	YES			
138		46982	10	S	40,07	1326	0,37	YES			
139		46983	10	S	40,07	1326	0,37	YES			
140		46984	10	S	40,07	1326	0,37	YES			
141		46985	10	S	40,07	1326	0,37	YES			
142		46986	10	S	40,07	1326	0,37	YES			
143		46987	10	S	40,07	1326	0,37	YES			
144		4698									

APPENDIX G: Excel Sheet with analysis of stock time in Live Model Women and Data Entry times

	A	B	C	D	E	F	H	I	J	K
1	Slot ID	Number of Items	Consider?	Stock Time [h]	Data Entry (s)	Data Entry (h)	Enough?			
2	35823	32	S	9,93	4426,05	1,23	YES		Data Entry Time [s]	138
3	37999	17	S	16,37	2351,34	0,65	YES		N° of slots that the stock time is enough	3391
4	38050	20	S	5,04	2766,28	0,77	YES		N° of slots that the stock time is not enough	12
5	38260	31	S	14,34	4287,74	1,19	YES		Number of slots in study	3403
6	42214	19	S	21,91	2627,97	0,73	YES			
7	42619	24	S	22,19	3319,54	0,92	YES			99,6%
8	42936	21	S	10,14	2904,60	0,81	YES			
9	42960	17	S	3,80	2351,34	0,65	YES			
10	42966	36	S	9,58	4979,31	1,38	YES			
11	43035	24	S	8,35	3319,54	0,92	YES			
12	43528	44	S	26,22	6085,82	1,69	YES			
13	43674	21	S	21,33	2904,60	0,81	YES			
14	44610	46	S	15,78	6362,45	1,77	YES			
15	44762	20	S	21,38	2766,28	0,77	YES			
16	44789	19	S	13,36	2627,97	0,73	YES			
17	45245	22	S	10,92	3042,91	0,85	YES			
18	45478	17	S	15,83	2351,34	0,65	YES			
19	45479	27	S	25,27	3734,48	1,04	YES			
20	45525	34	S	17,46	4702,68	1,31	YES			
21	45526	20	S	17,39	2766,28	0,77	YES			
22	45529	27	S	13,22	3734,48	1,04	YES			
23	45542	33	S	7,88	4564,37	1,27	YES			
24	45611	26	S	18,15	3596,17	1,00	YES			
25	45612	44	S	5,59	6085,82	1,69	YES			
26	45707	29	S	11,00	4011,11	1,11	YES			
27	45708	38	S	15,27	5255,94	1,46	YES			
28	45709	37	S	11,88	5117,62	1,42	YES			
29	45907	20	S	24,95	2766,28	0,77	YES			
30	46543	21	S	27,18	2904,60	0,81	YES			
32	46546	37	S	20,08	5117,62	1,42	YES			
33	46548	22	S	NORMAL	3042,91	0,85	YES			
34	46549	20	S	16,56	2766,28	0,77	YES			
35	46552	26	S	24,60	3596,17	1,00	YES			
36	46634	23	S	21,08	3181,23	0,88	YES			
37	46640	31	S	15,52	4287,74	1,19	YES			
38	46642	38	S	10,29	5255,94	1,46	YES			
39	46709	35	S	1,50	4841,00	1,34	YES			

Note: In 99,6% of the cases, the stock time for Live Model Women is sufficient to do DATA ENTRY of the items that go to Live Model

APPENDIX H: Excel Sheet with analysis of stock time in Live Model Men and Data Entry times

	A	B	C	D	E	F	H	I	J	K
	Slot ID	Number of Item	Consider	Stock Time (h)	Data Entry (s)	Data Entry (s)	Enough?			
1										
2	34255	39	S	10,79	4933,65	1,37	YES		Data Entry Time [s]	127
3	38661	18	S	21,57	2277,07	0,63	YES		Nº of slots that the stock time is enough	3155
4	38663	19	S	7,36	2403,57	0,67	YES		Nº of slots that the stock time is not enough	22
5	41679	19	S	17,35	2403,57	0,67	YES		Number of slots in study	3177
6	42207	27	S	23,23	3415,60	0,95	YES			
7	42214	19	S	4,30	2403,57	0,67	YES			
8	42960	13	S	3,81	1644,55	0,46	YES			99,3%
9	43034	20	S	4,95	2530,08	0,70	YES			
10	43378	14	S	8,24	1771,05	0,49	YES			
11	43675	28	S	3,08	3542,11	0,98	YES			
12	44098	35	S	-	4427,63	1,23	YES			
13	44111	19	S	22,99	2403,57	0,67	YES			
14	44114	13	S	14,96	1644,55	0,46	YES			
15	44146	27	S	12,54	3415,60	0,95	YES			
16	44148	43	S	18,60	5439,66	1,51	YES			
17	44149	33	S	4,01	4174,62	1,16	YES			
18	44726	12	S	-	1518,05	0,42	YES			
19	45081	30	S	14,43	3795,11	1,05	YES			
20	45123	15	S	-	1897,56	0,53	YES			
21	45125	24	S	15,46	3036,09	0,84	YES			
22	45244	39	S	14,53	4933,65	1,37	YES			
23	45245	21	S	13,43	2656,58	0,74	YES			
24	45409	33	S	22,29	4174,62	1,16	YES			
25	45410	32	S	4,92	4048,12	1,12	YES			
26	45412	15	S	14,16	1897,56	0,53	YES			
27	45526	20	S	9,25	2530,08	0,70	YES			
28	45527	32	S	5,26	4048,12	1,12	YES			
29	45687	26	S	9,57	3289,10	0,91	YES			
30	45688	37	S	0,67	4680,64	1,30	NO			
31	45689	22	S	18,77	2783,08	0,77	YES			
32	45690	31	S	7,73	3921,62	1,09	YES			
33	45691	27	S	0,38	3415,60	0,95	NO			
34	45692	35	S	17,97	4427,63	1,23	YES			
35	45693	24	S	2,20	3036,09	0,84	YES			
36	45694	16	S	17,59	2024,06	0,56	YES			
37	46028	15	S	-	1897,56	0,53	YES			
38	46102	29	S	-	3668,61	1,02	YES			

Note: In 99,3% of the cases, the stock time for Live Model Men is sufficient to do DATA ENTRY of the items that go to Live Model

APPENDIX I: Excel Sheet with analysis of stock time in Flat Kids and Data Entry times

	A	B	C	D	E	F	G	H	I
1	SlotId	Number of items	Stock Time (h)	Data Entry (s)	Data Entry (h)	Enough?			
2	73051	3	9,51	390	0,11	SIM		Data Entry Time [s]	130
3	73050	5	5,51	650	0,18	SIM		Nº of slots that the stock time is enough	842
4	72903	2	7,21	260	0,07	SIM		Nº of slots that the stock time is not enough	103
5	72882	46	1,81	5980	1,66	SIM		Number of slots in study	945
6	72881	49	1,63	6370	1,77	NAO			
7	72880	47	1,30	6110	1,70	NAO			89,1%
8	72879	46	3,01	5980	1,66	SIM			
9	72878	43	1,16	5590	1,55	NAO			
10	72815	48	1,90	6240	1,73	SIM			
11	72758	10	3,64	1300	0,36	SIM			
12	72751	17	2,04	2210	0,61	SIM			
13	72562	15	6,87	1950	0,54	SIM			
14	72541	37	7,41	4810	1,34	SIM			
15	72431	9	9,61	1170	0,33	SIM			
16	72209	8	5,29	1040	0,29	SIM			
17	72145	23	12,95	2990	0,83	SIM			
18	72144	46	14,22	5980	1,66	SIM			
19	72081	24	3,52	3120	0,87	SIM			
20	72080	40	4,56	5200	1,44	SIM			
21	72079	32	6,13	4160	1,16	SIM			
22	72078	27	1,71	3510	0,98	SIM			
23	72077	35	9,18	4550	1,26	SIM			
24	72076	21	12,85	2730	0,76	SIM			
25	72075	29	9,14	3770	1,05	SIM			
26	72074	25	7,10	3250	0,90	SIM			
27	72073	19	10,62	2470	0,69	SIM			
28	72032	44	1,32	5720	1,59	NAO			
29	71989	29	11,15	3770	1,05	SIM			
30	71936	11	1,84	1430	0,40	SIM			
31	71927	3	2,79	390	0,11	SIM			
32	71778	47	1,95	6110	1,70	SIM			
33	71777	43	3,57	5590	1,55	SIM			
34	71776	46	5,10	5980	1,66	SIM			
35	71775	46	6,98	5980	1,66	SIM			
36	71774	47	9,70	6110	1,70	SIM			
37	71773	47	4,71	6110	1,70	SIM			
38	71772	41	0,50	5330	1,48	NAO			
	◀ ...	First Cenario	Second Cenario	Third Cenario	AnalysisStillsNLM	AnalysisLMM	AnalysisLMW	Analysis FlatKids	⊕

APPENDIX J: Spreadsheet with the lead times and different scenarios

	A	B	C	D	E	F	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AG	A
	slot	number of items	MinScanInDate	MaxDepartureDate	Work Days	Consider?	Normal?	Nº of Non-duplicates	Nº Duplicates	CKIDS	CMEN	CUNISEX	CWOMEN	OTHKIDS	OTHMEN	OTHUNISEX	OTHWOMEN	WELLERYMEN	WELLERYUNISEX	WELWOMEN	WILLS	Sum of Cycle Times (h)	First Scenario	1st Scenario (Savings)	Second Scenario	2nd Scenario (Savings)	Third Scenario	3rd Scenario (Savings)	Lead Time (h)	Useful Time (%)	Duplicates	
1	51177	2	2017-12-29 21:38:07	2017-12-11 16:41:00	0	S	ANORMAL	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0,53	0,49	8%	0,53	0%	0,46	13%			-	
2	78406	31	2017-12-29 20:45:51	2018-01-03 17:16:00	2	S	NORMAL	31	0	0	0	0	1	0	10	0	5	0	0	0	3	12	9,76	9,50	3%	9,34	4%	8,65	11%	68,50	14%	-
3	72899	44	2017-12-29 20:33:20	2018-01-03 17:27:00	2	S	NORMAL	34	10	0	0	0	1	0	21	0	9	0	0	0	0	3	12,60	12,12	4%	11,51	9%	11,03	12%	68,89	18%	-
4	75969	36	2017-12-29 19:58:23	2018-01-03 17:17:00	2	S	NORMAL	22	14	13	0	0	0	9	0	0	0	0	0	0	0	0	8,09	7,50	7%	7,50	7%	6,81	16%	69,31	12%	-
5	72917	29	2017-12-29 19:52:53	2018-01-03 17:18:00	2	S	NORMAL	21	8	0	0	0	2	0	0	0	18	0	0	0	1	8,29	7,93	4%	7,51	9%	7,25	12%	69,42	12%	-	
6	79133	6	2017-12-29 19:52:35	2018-01-03 17:22:00	2	S	NORMAL	6	0	0	0	0	0	0	6	0	0	0	0	0	0	1,97	1,97	0%	1,84	7%	1,75	11%	69,49	3%	-	
7	78550	32	2017-12-29 19:14:40	2018-01-03 17:30:00	2	S	NORMAL	32	0	0	0	0	0	0	0	32	0	0	0	0	0	10,88	10,88	0%	10,18	6%	9,74	10%	70,26	15%	-	
8	71664	43	2017-12-29 19:08:56	2018-01-03 17:06:00	2	S	NORMAL	43	0	2	0	0	0	41	0	0	0	0	0	0	0	12,00	12,00	0%	12,00	0%	10,47	13%	69,95	17%	-	
9	79028	47	2017-12-29 18:42:31	2018-01-03 17:22:00	2	S	NORMAL	47	0	0	0	0	0	0	47	0	0	0	0	0	0	15,40	15,40	0%	14,38	7%	13,72	11%	70,66	22%	-	
10	72900	10	2017-12-29 18:41:17	2018-01-03 17:27:00	2	S	NORMAL	10	0	0	0	0	3	0	0	0	0	0	0	0	7	2,88	2,72	5%	2,81	2%	2,52	12%	70,76	4%	-	
11	70486	47	2017-12-29 18:25:02	2018-01-03 17:22:00	2	S	NORMAL	36	11	0	0	0	0	0	22	0	0	0	0	0	14	12,60	11,84	6%	11,66	7%	10,93	13%	70,95	18%	-	
12	77774	49	2017-12-29 18:15:38	2018-01-03 17:20:00	2	S	NORMAL	24	25	0	0	0	12	0	0	0	8	0	0	0	4	11,67	10,53	10%	10,19	13%	9,92	15%	71,07	16%	-	
13	63479	30	2017-12-29 17:54:58	2018-01-03 17:29:00	2	S	NORMAL	28	2	0	0	0	1	0	0	0	27	0	0	0	0	9,83	9,74	1%	9,13	7%	8,76	11%	71,57	14%	-	
14	71859	46	2017-12-29 17:26:07	2018-01-03 17:17:00	2	S	NORMAL	28	18	0	0	0	8	0	0	0	20	0	0	0	0	12,26	11,51	6%	10,90	11%	10,62	13%	71,85	17%	-	
15	78722	45	2017-12-29 15:57:07	2018-01-03 17:32:00	2	S	NORMAL	39	6	0	1	0	5	0	1	0	16	0	0	4	12	13,61	13,09	4%	12,77	6%	12,00	12%	73,58	18%	-	
16	74369	50	2017-12-29 15:33:06	2018-01-03 17:20:00	2	S	NORMAL	29	21	12	0	0	0	17	0	0	0	0	0	0	0	11,13	10,25	8%	10,25	8%	9,35	16%	73,78	15%	-	
17	78830	12	2017-12-29 15:21:05	2018-01-03 17:24:00	2	S	NORMAL	12	0	0	0	0	0	0	0	0	0	0	0	0	5	2,40	1,99	17%	2,10	12%	1,97	18%	74,05	3%	-	
18	78778	20	2017-12-29 14:59:54	2018-01-03 17:12:00	2	S	NORMAL	14	6	0	0	0	0	0	5	0	7	0	0	0	2	5,46	5,17	5%	4,95	9%	4,75	13%	74,20	7%	-	
19	76277	42	2017-12-29 14:54:38	2018-01-03 17:28:00	2	S	NORMAL	36	6	0	0	0	0	0	8	0	20	0	0	3	5	12,91	12,55	3%	11,98	7%	11,41	12%	74,56	17%	-	
20	78707	44	2017-12-29 14:13:46	2018-01-03 17:27:00	2	S	NORMAL	36	8	0	0	0	0	0	5	0	31	0	0	0	0	13,40	13,06	3%	12,28	8%	11,83	12%	75,22	18%	-	
21	78706	31	2017-12-29 14:11:21	2018-01-03 17:27:00	2	S	NORMAL	25	6	0	7	0	3	0	4	0	7	0	0	0	4	8,98	8,64	4%	8,27	8%	7,87	12%	75,26	12%	-	
22	78723	47	2017-12-29 13:01:48	2018-01-03 17:32:00	2	S	NORMAL	37	10	0	1	0	5	0	19	0	4	0	0	0	8	13,26	12,66	4%	12,21	8%	11,58	13%	76,50	17%	-	
23	78721	48	2017-12-29 12:40:06	2018-01-03 17:32:00	2	S	NORMAL	43	5	0	0	0	3	0	13	0	19	0	0	1	7	14,85	14,49	2%	13,86	7%	13,14	12%	76,87	19%	-	
24	78715	48	2017-12-29 12:17:14	2018-01-03 17:31:00	2	S	NORMAL	37	11	0	0	0	7	0	16	0	1	0	0	0	13	13,09	12,34	6%	12,10	8%	11,38	13%	77,23	17%	-	
25	78717	42	2017-12-29 12:13:21	2018-01-03 17:31:00	2	S	NORMAL	34	8	0	1	0	10	0	5	0	3	0	0	4	11	12,80	12,22	4%	11,96	7%	11,30	12%	77,29	17%	-	
26	78123	40	2017-12-29 12:06:59	2018-01-03 17:31:00	2	S	NORMAL	40	0	0	5	0	0	0	22	0	1	0	0	0	12	12,37	12,11	2%	11,76	5%	10,94	12%	77,40	16%	-	
27	78713	43	2017-12-29 11:13:46	2018-01-03 17:31:00	2	S	NORMAL	41	2	0	0	0	4	0	25	0	7	0	0	0	5	13,56	13,37	1%	12,69	6%	12,03	11%	78,29	17%	-	
28	78719	47	2017-12-29 10:56:55	2018-01-03 17:32:00	2	S	NORMAL	35	12	0	1	0	4	0	1	0	17	0	0	5	7	13,60	12,94	5%	12,49	8%	11,92	12%	78,58	17%	-	
29	73826	45	2017-12-29 10:53:40	2018-01-03 17:31:00	2	S	NORMAL	39	7	0	0	0	0	0	28	0	10	0	0	0	0	13,64	13,35	2%	12,52	9%	12,04	12%	78,46	17%	-	

	Average	Stand. Deviation	Minimum	Maximum
Nº of duplicates	7,78	7,07	0	43
Nº of non-duplicates	24,66	12,66	0	50
Lead Times [h]	62,54	22,16	0,82	143,96
Σ Cycle Times [h]	8,88	4,13	0,15	20,03
Useful Time [%]	15,8%	9,5%	0,1%	86,4%

Assumptions	
We only consider items produced in Europe	
We do not consider repetitions	
We assume that all items in the slot follow standard processes	

	Average	Stand. Deviation	Minimum	Maximum
Tempo útil LM (h)	1,12	0,64	0,05	3,17
Tempo útil LM (min)	67,31	38,42	3,06	189,95

	Current Scenario	First Scenario	Second Scenario	Third Scenario
Σ Cycle Times [h]	8,88	8,43	8,20	7,73
Saving [%]	0%	5,3%	7,8%	13,0%