Master's thesis

Warehouse, operatives and driver optimization process in a parcel delivery company

Autor: Antonio Castaño Blanco
Director: Ginés Alarcón

Submit date: September 2020

Escola Tècnica Superior
d'Enginyeria Industrial de Barcelona

Index

ETSEIB

Index of figures

ETSEIB

1. Introduction

Ecommerce is a growing industry. Since Amazon delivered its first book in 1995, to encompass the 95% of all purchases by 2040 according to the marketing agency 99 firms. But that is still far from the current situation. With a growth of 19% over 2019, there is a lot still to be discovered, which leaves a door open for entrepreneurs to find a place in the market. Such growth is leading to a turn on events that will change (and it is currently doing so) the global economic markets.

Distribution centres are a fundamental part for this market's growth, being (at least nowadays) a necessary piece on the whole process. From the moment the client orders something until it gets delivered there are two main processes regarding logistics:

- The retailer's facilities: Where the product must be located and loaded into a truck to be sent to the distribution centre.
- Distribution centre: Most commonly known as warehouse. Its purpose is to receive and process all the freight received by many retailers, and prepare it for the last mile delivery, where vans follow a route delivering all the parcels received from the retailers.

A distribution centre must be able to face the yearly increase of demand by also increasing its capacity and efficiency, maintaining or decreasing its costs.

The competitiveness of a sector such as distribution centres, where its yearly revenues have increased a 47.6% in the last ten years (figure 1), comes along with a profit subjected to both the final client's high demands regarding fast and personalized delivery and the retailer's volume of parcels fluctuations. This narrows the profit margins left for the warehouse company. How to enhance it while fulfilling the mentioned requirements is key to evolve and adapt within the market's growth. The growth factor might also cause problems if some of the main variables of parcel delivery such as parcel control, warehouse capacity and operatives' availability in a short period of time are not faced in time.

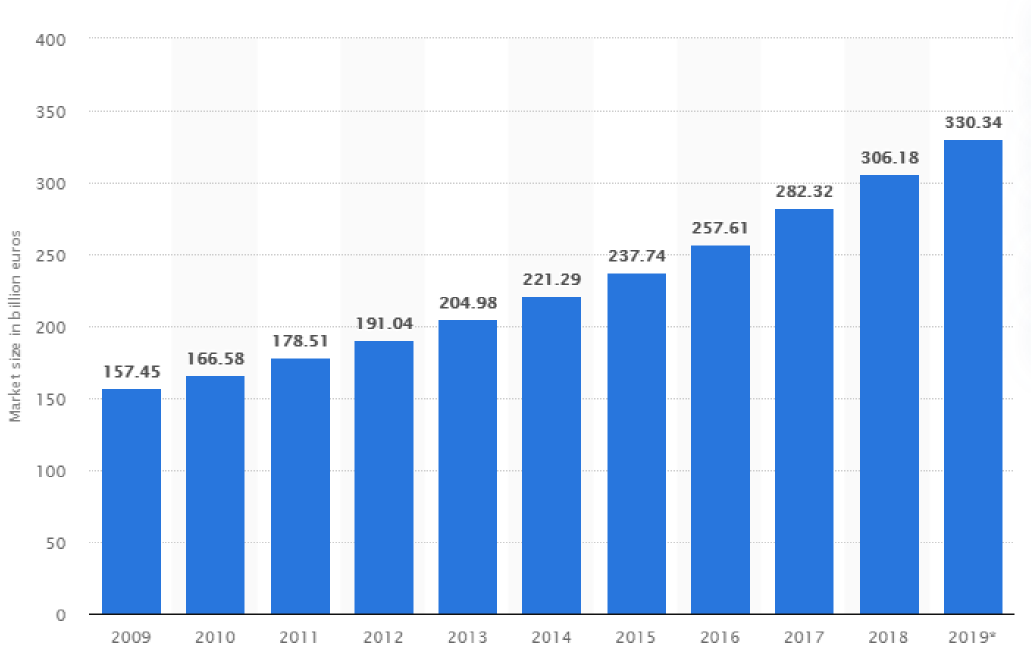

Figure 1: Parcel delivery market size worldwide (Source: statista)

In addition, European Union green industry policies such as industrial renewal, competitiveness, sustainability and eco-efficiency must be considered as milestones for the future of delivery industry.

1.1. Objectives

The main objective of this project is to analyse the delivery company performance regarding the dispatchment process so to study how the current processes and profit margins interrelate, to finally improve those margins to a minimum of the average of the market, by optimizing the processes considered as productivity/cost bottleneck.

This analysis will be carried out from the moment the warehouse receives a parcel to the instant the parcel is delivered.

Once the analysis is performed and conclusions are taken from it, a series of proposals will be analysed so to be able to enhance the warehouse productivity ratios at a feasible cost.

To do so, two different approaches will be taken:

- Provide a warehouse automation and control system.
- Improving the already existing processes.

But first, in order to be able to reach those approaches, we must perform a complete study of the company, its warehouse layout and all its processes.

The aim of this project is to find a new warehouse model that optimizes the parcel processing, reducing the costs and improving the service offered to the final client.

After the project is implemented, the warehouse will be an adaptive environment that can successfully re define its layout within short notice. Through developing a new model, we will be able to keep up with the company's growth rate, and to change variable costs to fixed costs independently to the volume of parcels.

1.2. Scope of the project

It is crucial to correctly establish the scope so its boundaries are set and we can focus on what is within the scope.

In Scope: The parameters considered as inside the scope are mainly everything that takes place or operates in the warehouse:

- Labour related: From the cost of the operators to setting the number of workers.
- Processes: They will be analysed and changed in case the overall productivity can be enhanced.
- Layout: The way the warehouse is distributed can be essential to be able to process and control as much volume as possible within the lowest cost.

Out of scope: These parameters are considered as fixed, and its change won't even be considered:

- Out of warehouse: not a single process nor parameter taking place out of the warehouse will be considered in this project.
- Warehouse facilities: changing the warehouse to a new one that fits better with the processes is out of the scope, since due to a contract doing so is not possible for several years.
- Retailer: Anything related to changing the conditions with the retailers such finding an agreement to improve the conditions or getting new clients.
- Labour cost: The costs associated to labour work are fixed.

2. Warehouse design

Before any further analysis on the current warehouse and its processes, it is essential to understand what defines the design of a warehouse, so to be able to detect later on its weak spots.

In order to set the layout of a warehouse properly, we must define the processes the parcels go through. As mentioned previously, this thesis only studies the process from the trucks departing from the retailer's dependencies to their arrival to the client.

2.1. Picker

Before going into further detail about the states, it is important to also introduce the picker. A picker is a feature that is assigned to an order the moment it gets firstly scanned at the warehouse. Thanks to this feature, we can move a vast quantity of items through different states, making sure they arrived to the warehouse without the need of having to manually check it. After a parcel is scanned, it gets assigned the picker "Picker BCN Warehouse". A parcel may get a picker assigned before or after it is routed since routing and picking are processed that are carried out in parallel. This way it is easier to keep control of the parcels' location. Currently, a scanned parcel gets the same picker regardless the state it is in, this will be a subject of study and improvement further on.

2.2. States of a parcel

There are eight states a parcel goes through since it appears in the system until it is delivered (or sent back to its retailer). These states are key to establish a proper communication between the parcel, the server, the workers and the consumers. This constitutes a basic IoT relationship, needed to track all parcels that are delivered every day, and especially for those that get lost through the process. The current states are the following:

- Draft
- Pending
- On course
- Picked-up
- Delivered
- Return in progress
- Undelivered

2.2.1. Draft

The order has entered to the company's server. This occurs before the parcel even leaves the retailer's facilities due to an input received from the retailer containing the parcel's id and delivery date. The retailer must fill the company in with the quantity of orders that are to be delivered for the next day so the company can ask the agencies for the appropriate number of drivers and operatives. In case not all the retailers can provide that information before the company has to ask the agency, the priority is set on the retailer whose number of orders are higher, so the rest can be forecasted by using previous data.

2.2.2. Pending

The order is in the warehouse. A parcel may be stated as pending when it firstly arrives to the WH or when it couldn't be delivered and was returned from the driver. Moving the orders to pending take place after they have been scanned. Note that only the orders that have a picker assigned and are in draft must be moved to pending. Otherwise, orders currently routed and in on course would be moved to pending.

2.2.3. On course

The order has been routed, and is to be delivered that same day. Routes are not linked to a specific driver, but to a pre driver that groups al the parcel with the same routes together.

2.2.4. Picked-up

When dispatching is taking place, and once the drivers arrive to the WH, routes are linked to them. Once the driver has received his/her route, its status is changed by the lead manager to picked-up. Furthermore, a notification is sent to the final client indicating that the driver has the parcel.

2.2.5. Delivered

This status is set by the driver via app the moment the client received the parcel. The driver must also attach a picture with the building's number in case there is an incidence reported.

2.2.6. Return in progress

When an order cannot be delivered after an established number of attempts, its status is set to return in progress. Once a week a truck goes to the retailer's facilities to return all the orders that couldn't be delivered.

2.2.7. Undelivered

An order is set to undelivered the moment a return in progress parcel is sent back to the retailer.

All these states reflect the parcel situation in real time, and though switching from one state to another isn't done automatically, it remains a key solution for managing thousands of parcels on a daily basis.

Along with these states stand the orders' process flow, which starts with the order number being received by the company's servers and ends when the final client gets it (or when the parcel is returned to its retailer due to an incidence).

2.3. Routing

Consists on assigning the following day parcels to a specific delivery route. This is carried out by a lead operative during the night shift, the moment the order is added to the system along with a delivery date set for the following date. Routing takes place independently of the parcel having arrived or not at the warehouse, since routing all the orders can take some time and there would be no margin of operation in case the operative would have to wait for all the freight to be picked.

2.4. Sorting

The sorting process starts right after all the orders that are to be delivered that day are routed. Sorting is carried out by scanning every parcel's barcode with a scanner paired with a phone app. Every time a parcel is scanned, the app indicates to which route does it belong to.

2.5. Roundcheck

Once all the parcels are sorted into different boxes according to their route, roundcheck takes place. It consists on scanning the parcels in each route to make sure all parcels that were put into the same box belong to the same route. This process is carried out due to the big volume the warehouse daily handles, resulting in some misplaced parcels due to a wrong sorting. Roundcheck is performed using a specific app mode that identifies the route of the first scanned parcel, generating a list with all the parcels attached to that route (except for the already scanned one) so they get removed whenever they are scanned. In case a wrongly sorted parcel is scanned, it is indicated by sound and visual signals in the app. This also allows to find if a parcel is not in the route box, since once the roundcheck is done in the whole box, the route list should be empty, meaning that all parcels where round checked. In case there is still parcels in the list, they were misplaced or not received by the retailer.

2.6. Dispatch

It starts when the first drivers arrive to the parking area. As they arrive, they get a route assigned. Operatives move the routes to the parking area. Each driver must count the number of parcels she/he has, and compare it to the quantity that appears in their app. This is done so from that moment they are responsible for the parcel they have attached.

2.7. Order's process

Figure 2 shows the states the parcel goes through since the order is received to the system until it leaves the warehouse for delivery.

The colour green indicates the parcel is in the retailer's facilities whereas the brown ones indicate the order being in the warehouse. In addition, the brackets coloured in a darker shade mean that the state is carried out automatically while the lighters are manually taken care of. However, it doesn't show the steps that don't involve a change of state in the parcel, such as sorting and roundcheck.

The problem lays at the bifurcation, when routing takes place there are no guarantees that every routed parcel will have already arrived to the warehouse, nor that the order will even arrive (the parcel may have been sent to a different location by mistake), meaning that in order not to lose the parcels' track, the process has to remain halted until all the orders are scanned and routed.

Once the parcels have been routed, they are manually moved to On Course.

An order is proven to be at the warehouse when it gets a picker assigned, which happens automatically the moment it is firstly scanned at the warehouse. When a parcel is scanned, a new tag is printed with the company's parcel id. In addition to it, the orders are sorted into their routes. The second phase of the sorting process is the roundcheck, which purpose is to make sure all the parcels where put into their right route and that the routes have all the orders.

The need of waiting for all the parcels to be routed to start the scanning and to check the routes have all their parcels suppose a delay in the process and more operatives.

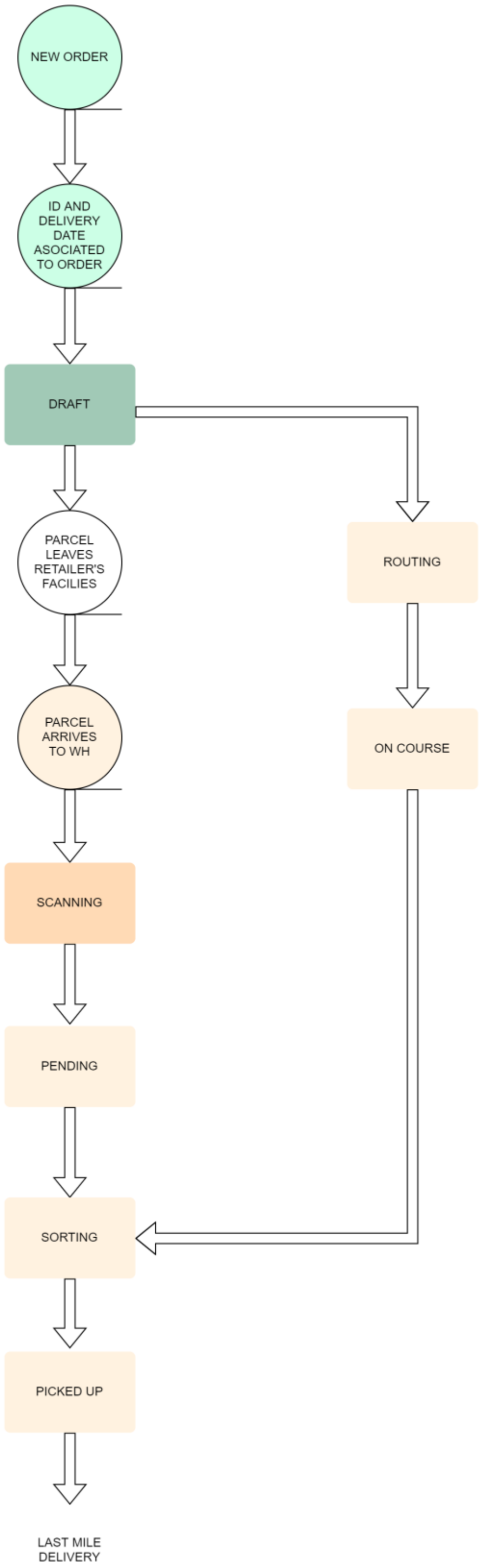

Figure 2: Order's process diagram

3. Operations

The main operation of a distribution centre is to act as an intermediary between retailer and client, delivering the ordered parcels. But there are other operations that a warehouse must carry out on a daily basis.

Right now, the WH’s layout is established according to the dispatching process, since that is the main operation.

3.1. Dispatching

It consists on receiving all the orders that are to be delivered for the day and processing them, labelling and sorting them into pallets assigned to a particular route. Once all the freight has been sorted, the routes are given to the drivers, who will proceed to the last mile delivery process.

3.2. Inventory

This process is complementary to dispatch. In order to keep achieving a successful dispatching process in the long run, the necessity of keeping an inventory up to date seems crucial so there is no lack of consumable items such as labels or film nor gadgets like phones and portable printers.

4. Current warehouse layout

Once the processes have been laid out, the current layout can be explained and discussed.

The second most relevant process regarding costs is the one involving parcel sorting and dispatching, both activities taking place in the warehouse. To set the proper KPIs and to measure variables such as operative's productivity, average and peak volume and retailer truck's arrival time is essential so the warehouse performance can be studied and improved. A proper performance comes along with the correct number of workers and timetable so to enhance the margins obtained from delivering a parcel.

But before getting a close look at the warehouse's processes, it is important to study the current warehouse size, distribution and location.

The current Barcelona warehouse is located in *Parc Logistic*, a logistic area in *Zona Franca* with a total of 120,000 m^2 of facilities aimed for the logistic and e-commerce sector.

It is a one floor building of 1920 m^2, distributed in 6 different areas:

- Load/Onload dock: It is a six-dock area, where the trucks load and unload their freight. Pre-sorting process also takes place in this area.
- Sorting area: Where the parcels are stored before dispatching. They are sorted by the route the driver will follow to deliver them.
- Parking area: Area where drivers park their vehicles to intake the parcels.
- Office: Mainly for routing and solving issues that take place during dispatching.
- Parcels to return area.
- Restroom.

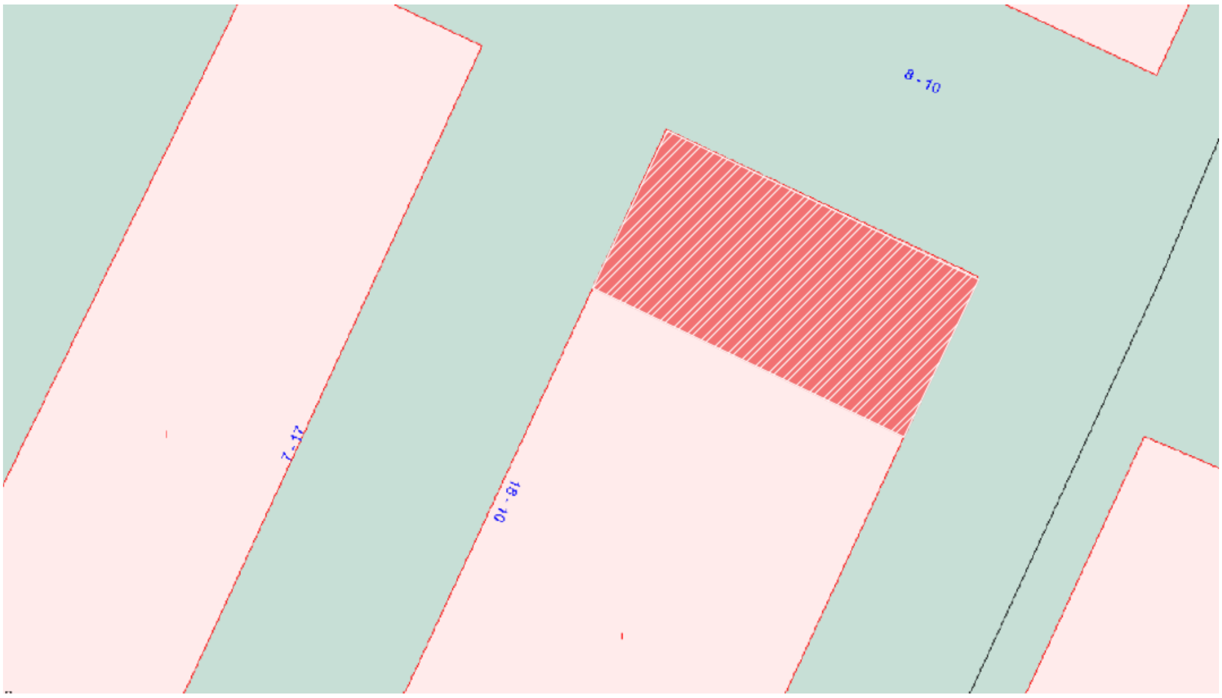

Figure 3: WH Area cadastre

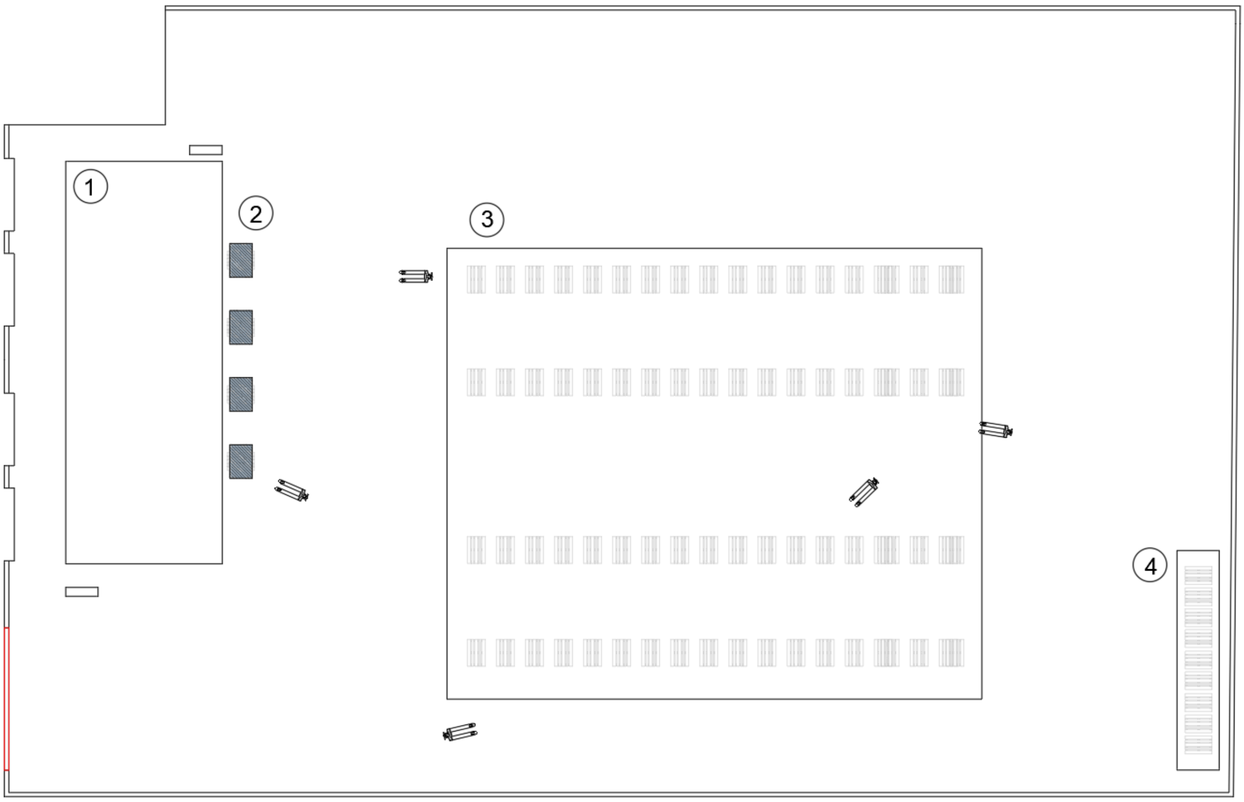

Figure 4: WH current layout

WH Operation areas:

1. Load/Unload and dispatchment area.

2. Pre-sorting boxes.

3. Sorting pallets.

4. To be returned orders.

4.1. Organizational policies

WH organizational policies' aim are to define the warehouse layout according to the following standards:

- Organizational policies: Formed by the assignment policies establishing the allocation of trucks to docks and other vehicles such as vans.

- Storage policies: Irrelevant for our case of study since the freight manage in a parcel delivery warehouse isn't generally stored for more than one day.

- Order picking policies: They refer to the picking and storing hierarchy.

- Operators assignment policies: Related to the number of operatives working in each shift.

The facility has the workers, delivery drivers and truck drivers entrance located in the same side. This affects the WH layout process flow, forcing it to be bidirectional and limiting the organizational policies regarding the vehicle assignment policies: The freight arrives at the same place where the drivers come and pick up the routes. This forces operation related to trucks and related to vans to be coordinated so no in load/outload freight and dispatchment take place at the same time.

The storage area is at the bottom of the warehouse, and the parking area is just between the storage area and the in load/unload area.

Once the freight is out loaded from the truck, picking and sorting take place sequentially, first picking the parcels and then sorting them into their routes.

Because the big volume fluctuations the warehouse process on a daily basis, the operatives, which come from an external agency, are called upon the forecasted weekly volume.

4.2. WH Design problem

There are three main approaches regarding the warehouse design problem. These approaches depend on each other according to how it affects the WH performance in terms of time, being the longer-term decision a constrain for the lower ones. These approaches are classified into strategic, tactical and operational.

4.2.1. Strategic level

Where long term impact decisions are considered. Topics such as process flow design and selecting the machinery WH systems are being considered taking into account WH technical capabilities and economic considerations.

4.2.1.1. Process flow design

Regarding process flow design, both the input and the output material and products must be described. Figure 5 shows and order's flow, since it leaves the retailer's facilities until it is delivered. The input are parcels coming from different retailers, whose delivery time interval is set by the costumer. In the other hand, the output are routes of parcels that are designed basing on the delivery postal code and the delivery time interval so when a driver is dispatching the route, the dispatching order and delivery location make sense regarding their time table and location.

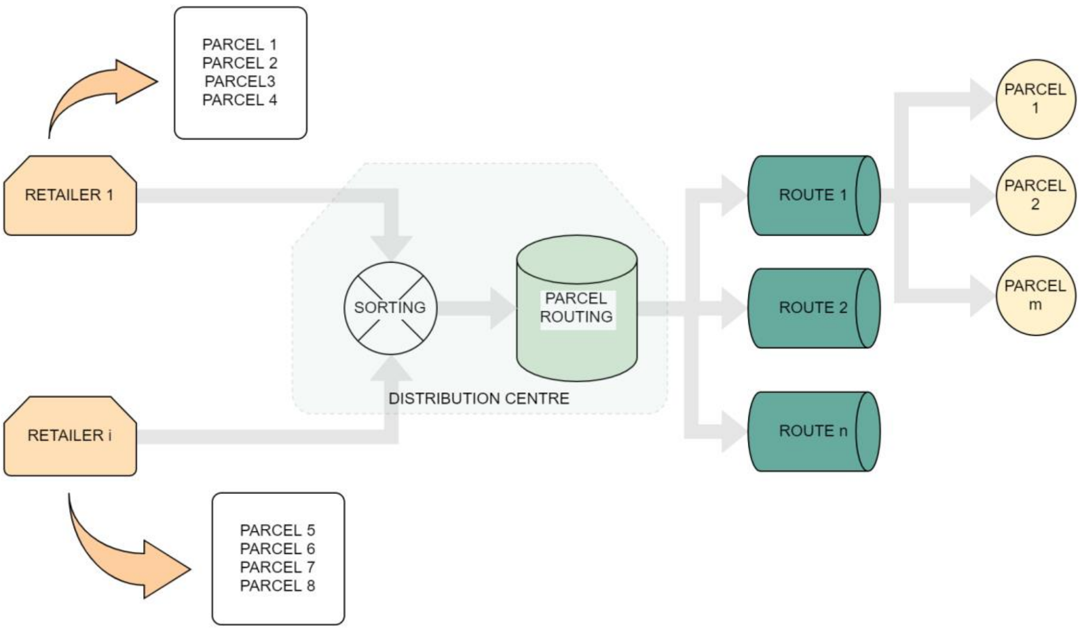

Figure 5: Orders flow chart

So, the main design problems here these three:

- Are the truck arrival times standardized?
- How are the parcels sorted after the trucks unload?
- How are the routes distributed? This question is out of the scope of the thesis.

In terms of choosing the WH systems, a business case regarding the possibility of automating the warehouse entirely or just some selected process must be carried out. Currently, there are not many automated processes regarding parcel processing and delivery, apart from the order getting a picker assigned. This will be further discussed in the following sections.

The required equipment of a distribution centre doesn't go beyond the standard non-automated warehouse tools.

4.2.2. Tactical level

After setting the process flow and the warehousing systems, we have to dimension the WH different zones (sorting, storing, dispatching, etc) and its resources such as the work items, the operatives and the processing systems (in case there are any).

In order to set this up correctly, the average and peak volume processed on a daily basis must be established, and size parameters such as the width of the alleys. This is crucial to minimize the over costs that might be applied due to an over dimension of the WH size or operatives' quantity.

A cost analysis will be carried out later in order to see which are the costs and if they fit the gross margins set by the company.

4.2.3. Operational level

Carried out within the constraints set by strategic and tactical levels. It studies the assignment and control problems of operatives and equipment, in addition to the routing system.

5. Operations' parameters

Before starting the analysis that will lead to future conclusions and decisions, we have to define and quantify the variables that are relevant during the operations. These variables may constitute a wide variety and quantity. It has major importance which variables are selected to measure the warehouse activity throughput.

Most of the parameters change within time, so we have set the month of march to set our variables, although in a real approach they should be analysed after the last updated data inputs.

5.1. Volume related

- Volume: 3500 parcels/day. Although volume is a fluctuating variable that depends on the time of the year, the average volume avoiding high volume periods (for this case, a different study must be carried out) has been calculated taking into account the 2020 orders received by retailers except Black Friday and Christmas days. 90% of these parcels are to be delivered during the morning shift, whereas the other 10% belongs to the afternoon shift.
- Parcels per route: This is a constant value set at 40, constraint by the van capacity and the uncertainty of the parcels' volume, which forces the van to leave an extra space in case a route's orders volume exceed the average.
- Number of drivers: 88 This parameter comes from the two mentioned above, being the number of drivers required for delivery, the volume divided by the parcels per route (since the route refers to the number of parcels a driver delivers in a shift sequentially). 79 of those drivers work in in the morning shift (as mentioned, 90% of the parcels are delivered in the morning) and 9 in the afternoon.

Volume related	Quantity	Units
Volume	3500	parcels
Driver Capacity	40	parcels/route
Morning volume	90%	
Afternoon volume	10%	

Figure 5: WH parameters. Volume related

5.2. Time related

- Sorting productivity: 120 parcels/hour/operative. This parameter is out of the thesis' scope. It measures the number of parcels per hour that an operative should process under a correct performance during sorting.
- Roundcheck productivity: 540 parcels/hour/operative. It is a much faster process compared to sorting productivity, since this phase of the process consists only in double checking that all parcels are put in their route. Although it is way quicker than sorting, it is not as relevant, it is performed due to the high volume of wrongly sorted parcels. Later on, alternatives to this process will be further discussed.

To both productivities, there is an extra margin that should be applied when calculating the timings, due to operatives' extra time that might be spent moving around the WH or solving daily incidences.

TIME RELATED	Quantity	Unit
Sorting productivity	120	parcels/h/op
Sorting capacity	30	s/parcel
Roundcheck productivity	540	parcels/h/op
Margin for incidences	20%	

Figure 6: WH Parameters. Time related

- Already set timings
  - Truck arrivals: 4:30 am. This is the starting point for the operatives. Once the trucks arrive, the operative onload the freight and start sorting. The time this process takes depends on the number of operatives working at the same time.
  - Dispatching start time: 8:50 am. At this time, drivers arrive at the WH to load their routes and start delivering.

5.3. Labour related

To be able to fit the timings, which are already restricted by the truck arrivals and the earliest time range delivery, at 9 am, we have to fix the proper number of operatives for each operation before dispatching: Sorting and roundcheck.

The current number of operatives per operation depends only on the volume of parcels, regardless of the difference of productivity that both processes (sorting and roundcheck present)

The truck arrivals and dispatching start time sets a restriction for sorting and roundcheck timings, hence, also for the number of operatives. The margin between both processes is of

4 hours and 20 minutes. So, the total hours along with the number of operators must fit that space.

	Sorting	Roundcheck		
n Op	Time [h]		Time total [h]	Cost [€]
5	5,3	1,4	6,5	400
6	4,4	1,2	5,5	406
7	3,8	1,0	5	431
8	3,3	0,9	4	394
9	2,9	0,8	3,5	387
10	2,6	0,7	3,5	431
11	2,4	0,6	3	406
12	2,2	0,6	3	443
13	2,0	0,5	2,5	400
14	1,9	0,5	2,5	431
15	1,8	0,5	2	369

Figure 7: Labour costs

Figure 7 shows the total hours of labor required depending on the number of operators working at the operative, along with the total time and the daily cost attached. The time total column is calculated by rounding up the sum of both the sorting and roundcheck time since the shifts' sensitivity can't be lower than half an hour. Another constraint set by the temporary agencies is that the operators working hours cannot be under three hours and a half.

Currently, the number of operators used for an average morning shift is 10, which is not the optimal quantity since the number of operators can be lowered to 9 while meeting the requirement mentioned above.

In case the volume forecast indicates the contrary, this quantity must change accordingly (by just changing the volume field in the excel file). However, taking into account the productivity of sorting and roundcheck, it might be worth taking into consideration adjusting the number of operators for each operation.

The costs attached to operators may vary depending on the shift, being the normal shift salary around 12€/h to 15€/h during night shift and extra hours.

Parameters related to the drivers are described in Figure 8.

Driver related		
Number of drivers	88	drivers
Morning	79	drivers
Afternoon	9	drivers
Cycle time	15	min/van
Capacity	4,0	van/h
Max vans	20	vans
Total Capacity	80	vans/h
Load	59	min

Figure 8: WH Parameters. Driver related

The number of drivers is directly related to the volume and the driver capacity (Figure 5)

$$Number\ of\ drivers = \frac{Volume}{Driver\ Capacity}$$

Out of the total volume, the 10% is usually delivered in the afternoon whereas the other 90% is delivered in the morning, the number of drivers operating in both shifts are similarly distributed.

The total capacity indicates the number of vans that can be dispatched per hour taking into account the maximum number of vans that can fit into the parking area.

$$Total\ Capacity = Max.Vans\ \times Capacity$$

The load time has been calculated taking into account the morning shift numbers of drivers for it is the bottle neck of the process.

$$Load\ time = \frac{Number\ of\ drivers(morning)}{Total\ Capacity}$$

6. Cost analysis

To be able to improve the shipping process, it is necessary to have a clear view of the costs, so to focus on the ones that seem higher than expected. Costs are divided into fixed and variable. Fixed costs are those that remain constant independently of the volume, drivers, etc. On the other hand, variable costs change as one of the previously described parameters change.

6.1. Fixed costs

These types of costs are divided by the days the WH is operating, so the margins can be studied on a daily basis. This study was carried out during the month of May, with a total of 24 operating days.

- Warehouse rent: 1480 €/month, a daily cost of 61.66 €/day.

6.2. Variable costs

- Drivers: The highest of the variable costs. It depends directly upon the volume of orders to be delivered that day. As stated in *figure 5,* every driver may take up to 40-45 parcels per route, so the drivers needed for a day are exactly the volume of orders divided by the driver's parcel capacity. Drivers currently come from an external agency, and the quantity required is forecasted on a weekly basis.
  Because they have attached the highest cost, they have the highest responsibility into delivering the parcels in the time range they have assigned.
  Every fleet of drivers has its own conditions attached, being the most relevant the price paid to the fleet per parcel delivered within its time slot.

- Operators: Also, from external agencies. Similar to drivers, the cost is directly related to the volume and to the operative's productivity, which has been set to 100 parcels/hour/operative for sorting plus presorting and of 540 parcels/hour/operative. It is also essential to know the trucks' arrival time so to establish the available time for the operation to be done.

- Parcel pick-ups: All the costs related to picking up the parcels, whether from the retailer's facilities or directly from the store, are charged in here. This cost varies within the number of trailers required, which change depending the volume of

7. WH Current constraints

A WH process does not only depend upon the internal structure and way of working but also on external factors. These factors usually cannot be approached directly in order to change whatever that might be distancing the process from its ideal, but have to be addressed in an indirect way, like adapting the internal process in order to improve the overall throughput of the operation.

7.1. External factors

The main external factors that nowadays affect the warehouse are not unknown to distribution centres. Factors associated to the retailer such as the truck's arrival time (or delay) to WH or the volume fluctuations that each retailer may present might affect deeply to the gross margins at the end of the day.

Another factor linked to the previous ones mentioned are the temporary agency workers availability, which may be limited during a peak volume period leaving the process without operatives and/or drivers enough to provide a proper service.

External factor that may cause such a deep effect to the service must be taken into account when developing the WH improvement plan. But before, an investigation to find all the casuistry that may affect the operative in a direct or indirect way has to be carried out carefully.

8. WH Current margins

Now that the costs attached to a parcel have been addressed, we can go further on and provide an analysis of the margins obtained per parcel. To do that, all the parameters related to the margins must be studied and scrutinized.

8.1. Fleet parameters

As mentioned in the previous section depending on the driver's performance, the payment may oscillate:

- Price per parcel up to the minimum load: The price the driver gets paid when he/she delivers up to the quantity of parcels considered as the minimum load.
- Price per parcel once the minimum load has been surpassed: Once the driver has exceeded the minimum load, the amount of money per parcel decreases.
- Bonuses for delivering most of the parcels, while delivering them within its time slot.

There are also penalties associated to a faulted delivery service:

- The driver didn't try to deliver the parcel
- The parcel was delivered out of its time slot
- The total amount of delivered parcels within their time slot are below 90% of the total amount of parcels.
- The driver didn't show up at the warehouse for work.

These penalties may not be applied in case there is a justified explanation for such results, such as a traffic retention due to a demonstration or a street that has been closed for any reason.

8.2. Gross margins

It is key to analyze only the margin that are within the scope of our case of study in order to be able to detect which parts of the process are the ones diminishing the margins the most. The process has been divided into two, according to the nature of the operation:

- GM1: Includes the part of the dispatching where the drivers are involved, which is the last mile delivery.

- GM2: Covers the tasks related with the receiving and processing of the freight until is it set for dispatching.

These margins have been studied using April's parcels dataset, they have been detailed in *Figure 9*. The table has the following given parameters:

- Sales: They account for the daily gross profit received by the retailer for delivering their parcels. There is a deal for every retailer, which is out of the scope of this project, so the table only shows the overall gain, without differencing between the retailers.

- Drivers: This parameter stands for the costs charged directly to the drivers due to the last mile delivery. The cost has been calculated taking into account the parameters previously detailed at section 6. Again, there are many fleets operating for the warehouse, each one of them with a different economic deal with the company, so the cost shown in the table reflects the overall.

- Operators: Cost charged to the operatives working in all the processes carried out at the warehouse.

- Freight pick-up: The freight might come from another warehouse or directly picked-up at the store. Freight pick-up stands for the costs related to bringing the orders to the warehouse.

	1-abr.	2-abr.	3-abr.	4-abr.	6-abr.	7-abr.	8-abr.	9-abr.	13-abr.	14-abr.	15-abr.	16-abr.	17-abr.	18-abr.	20-abr.	21-abr.	22-abr.	23-abr.	24-abr.	25-abr.	27-abr.	28-abr.	29-abr.	30-abr.	Total
Sales	11.080	8.061	13.770	9.895	11.233	13.901	14.676	16.399	13.693	13.690	16.641	14.262	15.746	13.762	12.471	17.787	20.693	23.115	20.033	14.928	21.216	23.283	18.696	17.456	376.484
Drivers	8.561	5.567	9.414	7.370	7.842	9.447	7.758	11.173	8.943	9.933	10.728	9.567	10.900	9.618	8.615	11.803	13.541	14.217	13.430	10.797	15.023	16.229	11.982	11.318	253.776
GM1	2.519	2.494	4.355	2.525	3.392	4.454	6.917	5.226	4.750	3.757	5.913	4.694	4.846	4.144	3.856	5.984	7.152	8.897	6.603	4.130	6.193	7.054	6.715	6.138	122.708
GM1%	23%	31%	32%	26%	30%	32%	47%	32%	35%	27%	36%	33%	31%	30%	31%	34%	35%	38%	33%	28%	29%	30%	36%	35%	32,59%
Operatives	928	668	1.022	1.187	1.086	1.381	1.036	1.390	844	1.131	1.488	1.215	1.082	1.160	824	1.452	1.683	1.863	1.750	1.530	1.740	1.734	1.650	1.545	31.389
Freight pick-up	899	833	1.141	673	1.292	1.259	989	1.037	1.014	1.712	921	1.365	1.067	974	944	1.357	989	2.061	1.694	1.411	1.020	1.524	1.081	1.245	28.501
WH Rent	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	642	15.418
GM2	8.611	5.917	10.964	7.393	8.212	10.619	12.008	13.330	11.192	10.205	13.589	11.040	12.954	10.985	10.061	14.335	17.379	18.549	15.947	11.344	17.812	19.383	15.323	14.024	301.176
GM2%	78%	73%	80%	75%	73%	76%	82%	81%	82%	75%	82%	77%	82%	80%	81%	81%	84%	80%	80%	76%	84%	83%	82%	80%	80%
GM	50	349	1550	22	371	1172	4250	2157	2250	273	2861	1473	2054	1368	1446	2532	3838	4331	2518	546	2790	3153	3341	2706	47.400
GM%	0%	4%	11%	0%	3%	8%	29%	13%	16%	2%	17%	10%	13%	10%	12%	14%	19%	19%	13%	4%	13%	14%	18%	16%	13%

Figure 9: Monthly margins

8.2.1. GM1

Currently, fleet cost is the factor that narrows the margins the most. Gross Margin 1 (GM1) reflects the margins after subtracting the costs related to the fleet.

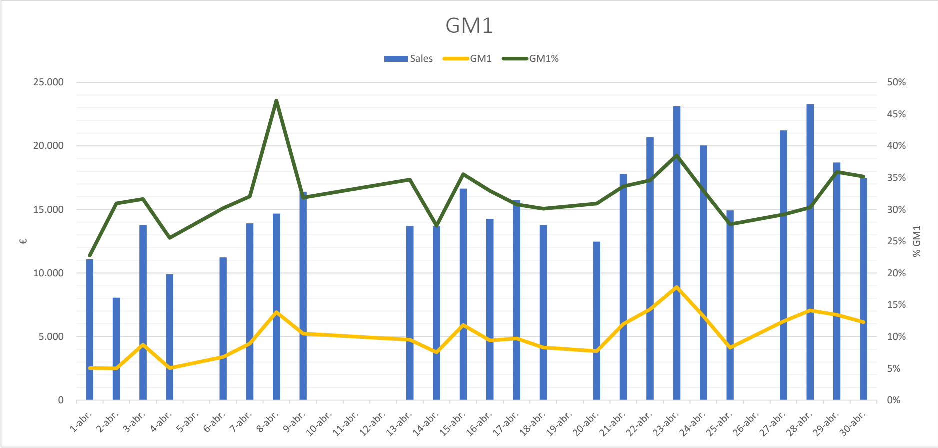

Figure 10: GM1 vs Sales

As it can be observed in figure 10, which compares the gross profit made by the delivery orders with the GM1 both in units and as a percentage. The gross margin left after fleet costs are charged have a peak of a 47% and an average of a 33%, meaning that two thirds of the profit made per parcel delivered goes to the fleets.

8.2.2. GM2

In addition to GM1, we have to also discount the costs related to the warehouse operation, which constitutes the sum of the following parameters:

- Operatives
- Freight pick-up
- WH rent: Divided between the days it is operating so the it can be charged on the daily margin.

Once all the costs are charged, GM2 is obtained.

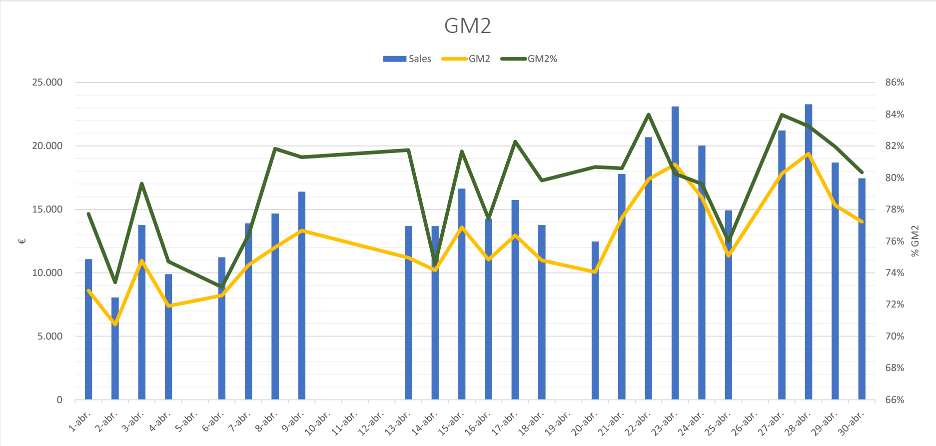

Figure 11: GM2 vs Sales

GM2 presents a higher margin's ratio than GM1, almost doubling it with an 84.5% at its peak and an average of 80%. This shows that the greatest losses come from the last mile delivery phase. However, although GM1 is way lower, GM2 present more factors and a higher ratio of improvement in terms of productivity and process design.

8.2.3. GM

Figure 12 shows the total gross margin, subtracting GM2 to GM1, and comparing it to the daily sales. The overall gross margin is of 47400 €, which compared to the total sales makes for a profit of a 13%.

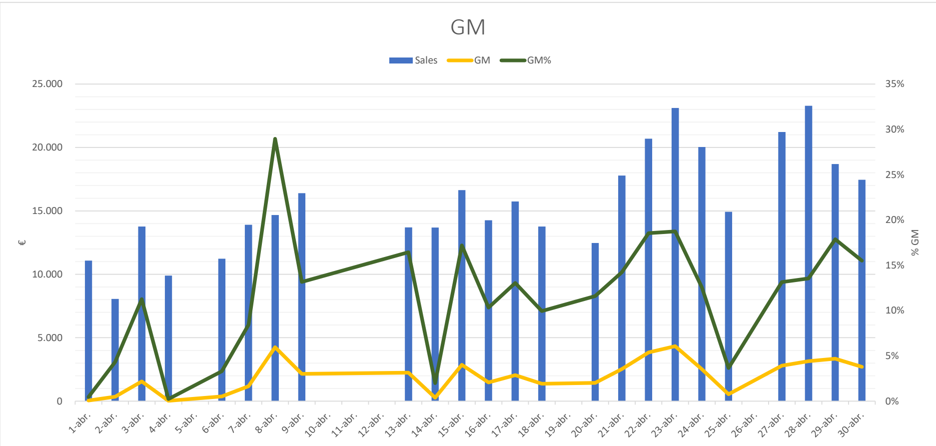

Figure 12: Total GM vs Sales

So, how are the numbers compared other models in the market? According to Capgemini institute study, carried out in 2018, the overall average cost, which accounts for warehouse processing and last-mile delivery accounts for a 28.7%, two times the company's overall gross margin.

It is fundamental to re define the shipping model and design a strategy in order to gradually cut distance between both numbers and, once achieve the 28.7% of profit find new ways to increase it.

In the following sections, different new projects are approached to reach the established margins, starting for the processes related to GM2 and then following to the ones involving the Last-Mile delivery, always having in mind the main goal, which is to get to the average gross margin per parcel within a time period of two years.

9. Warehouse volume

The warehouse labor system is not fixed, it varies directly within the daily demand. Basing upon the expected volume, the warehouse asks the work agencies in case of the operatives and to the fleet companies for the drivers for a specific quantity of workers. A shift on volume has to rapidly translate into a shift on labor demand. An accurate volume prediction is crucial to identify and apply those shifts fast enough without over nor under scaling, which means an increase on the daily gross margin per parcel.

To have a clear view on demand, historical volume data will come in handy. Unfortunately, for companies with a short run other options should be considered and studied. For this case of study, the volume dataset contains 52 days, as indicated in *Figure 13.*

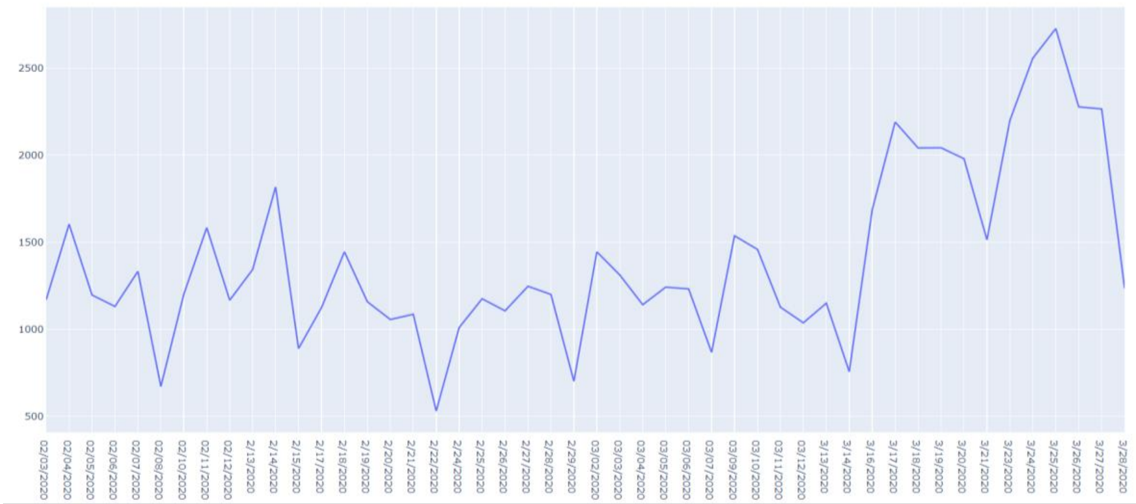

Figure 13: Barcelona daily volume March – April

To get a proper forecast it is crucial that it follows a trend or a pattern, this way the forecasting process can identify and predict basing upon them.

In this case, there is a pattern that repeats no matter the higher or lower volume trend, that is, two peaks and one valley every week. The peaks belong to the beginning and the end of the week, being one fixed on Friday and another distributed between Monday and Tuesday depending on the week. The valley is always on Saturday, where volume week after week reaches the lowest in the week. This may change within time and then so will do the algorithm estimating the volume to process.

Figure 14 shows the distribution of daily volume depending on the day of the week, being 1 Monday and 6 Saturday (the warehouse remains close on Sundays). The graphic shows that, in spite of the dispersion that specific days show, around three per day, there is also a pattern trend following that dispersion. Wednesday is the most irregular day of the week, which can be observed both through figure 5 and 6.

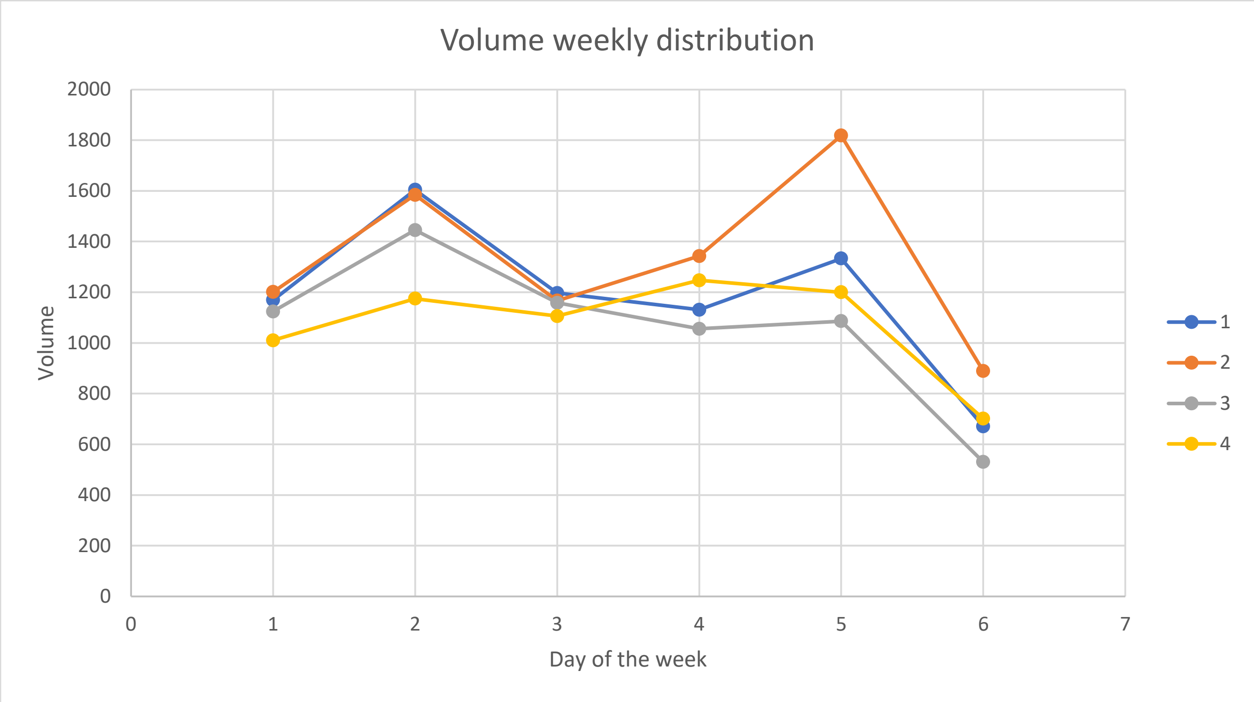

Figure 14: Volume distribution

There are many ways to approach a solution, from solutions that only take into account the weekly volume and it trend towards past weeks to considering the holidays and other special days such as sales in order to apply a scale factor.

It remains clear the need of a forecast whose results vary upon time, and that takes into account factors such as the week day and the trend of the curve. Taking into account these requirements, the solution we've come up with is building a forecast model using neuronal networks that are time related.

Artificial neuronal networks are mathematical models that emulate how the brain behaves and process data. They are especially useful for non-linear dataset interpretation; however, they may also be used for linear data sets [1].

A neuronal network is composed different types of neurons:

- Predictors: They are the input data that is to be processed by the network. They form the bottom layer of the network. Every predictor has a weight associated to them, depending on how their presence reflects on the output. The weights are selected in the neural network framework using a “learning algorithm” that minimises a “cost function” such as the MSE [1].

- Hidden neurons: Forming the intermediate layer. They are used for nonlinear data and the number of hidden neurons may vary upon the nonlinearity degree. Time related inputs don’t require hidden neurons since they are (in most of the cases) linear.

- Output layer: It processes all the predictors and outputs the forecast value.

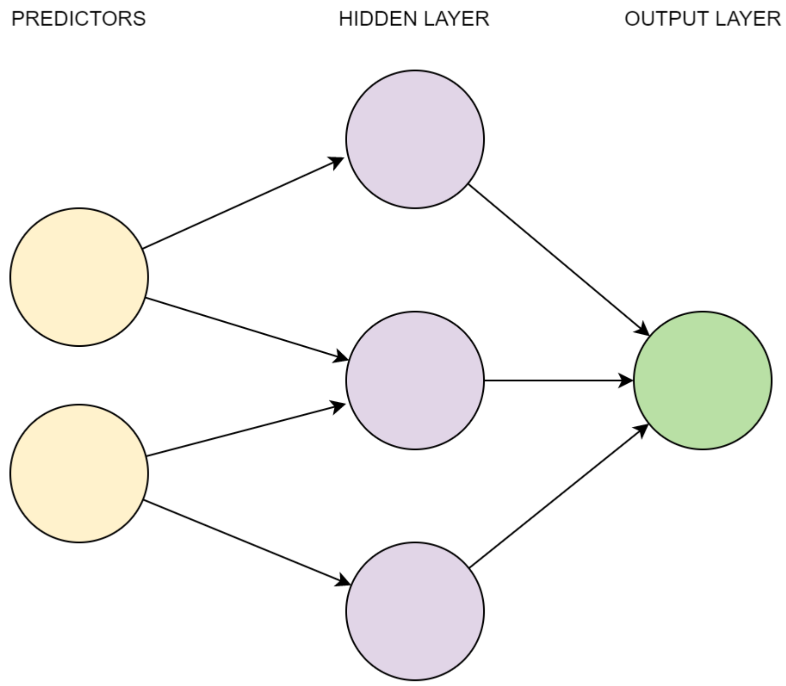

Figure 15: Neuronal network architecture

For time series, the previous dates may be used as input data in form of lagged vales in an autoregressive model just to get the following day as an output of all the lagged inputs taking into account the season ability of the data along with the trend line.

There are some factors that affect how the forecast’s accuracy [2]:

- Trend: Defined by the long-term gradual increase of volume.

- Cycle: Defined by how the volume varies following a pattern.

- Seasonality: It controls the short-term fluctuations of volume.
- Error: Similar to noise on a signal, a time series may have fluctuations that cannot be explained by the model.

9.1.1. LSTM Network

To be able to control these factors and build a model that outputs a proper volume estimation, we have decided to use Long Short-Term Memory networks (LSTM). LSTM process the input data as sequential, so all the inputs are treated as dependent one from each other, another feature of these network is the capacity of remembering long-term sequences being able to associate and process them in order to output a trustworthy output.

The neuronal network has been built up using Python programming environment, since it can provide many useful tools for designing and implementing the LSTM model.

In order to prove that the model estimations are valid a train and test method has been developed. Data will be split into two groups, so the neuronal network is trained using the first group and then it can be tested by forecasting the values from the second group. This way, we can compare the forecast to the real values registered and decide whether the network has been properly trained or not.

9.1.2. Architecture

Every neuronal network must have set its proper architecture, so the forecast is as accurate as possible. These processes can follow some parameters to get the optimal one, although many times the best option is acquired by trying different parameters. The following parameters have been set by a combination of both.

- Inputs: 6. According to the number of days the warehouse is operating.
- Output: 1, for the number of days we want to forecast.
- Hidden neurons: 10. To estimate the number of hidden neurons the following formula has been applied.

$$hidden\ neurons = m + n + \frac{m+n}{2} = 1 + 6 + \frac{1+6}{2} = 10 \sim 11$$

- Activation function: tanh. Used to introduce non linearities to the output network [3].
- Epochs: 1000. It defines the number times that the learning algorithm will work

through the entire training dataset.

For the program to run we have made used of the libraries Keras and Pandas, which are especially designed for neuronal network processing.

Figure 16 shows the results from the training process. The red points are the current values while the green ones are the trained ones, which are trying to approach the real values. The closer they get, the better the train. This is not a closed training, but a continuous one. Within time, the dataset will increase and the neuronal network's train will improve.

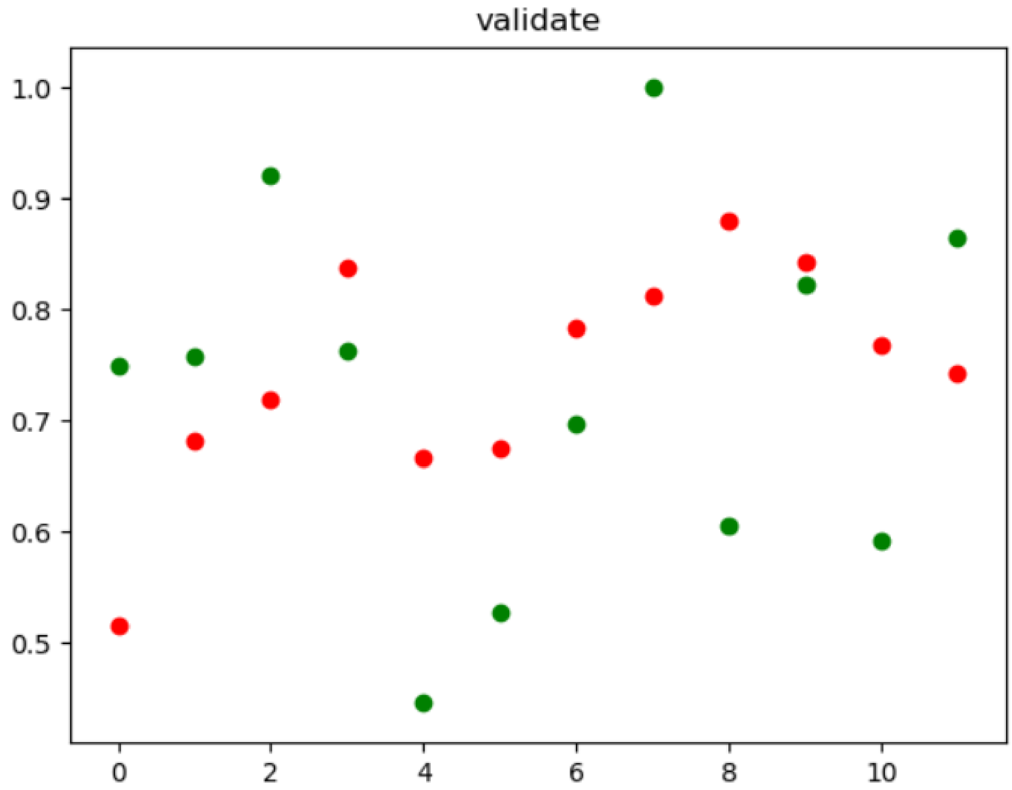

Figure 16: Training process validation

From Figures 17 and 18 the difference between the beginning and the end of the training process is noticeable, from the first epoch error and mean square values to the last one. The last epoch presents pretty low loss and mse values.

```
Epoch 1/1000
2/2 [==============================] - 1s 352ms/step - loss: 1.2154 - mse: 1.7437
```

Figure 17: Training function at epoch 1

```
Epoch 1000/1000
2/2 [==============================] - 0s 19ms/step - loss: 0.0969 - mse: 0.0172
```

Figure 18: Training function at epoch 1000

Now, results from figures 17 and 18 will be compared:

- Loss: 92.03% reduction.
- MSE: 99.02% reduction.

The epochs effect is clear, both the loss and the minimum square error are hugely reduced.

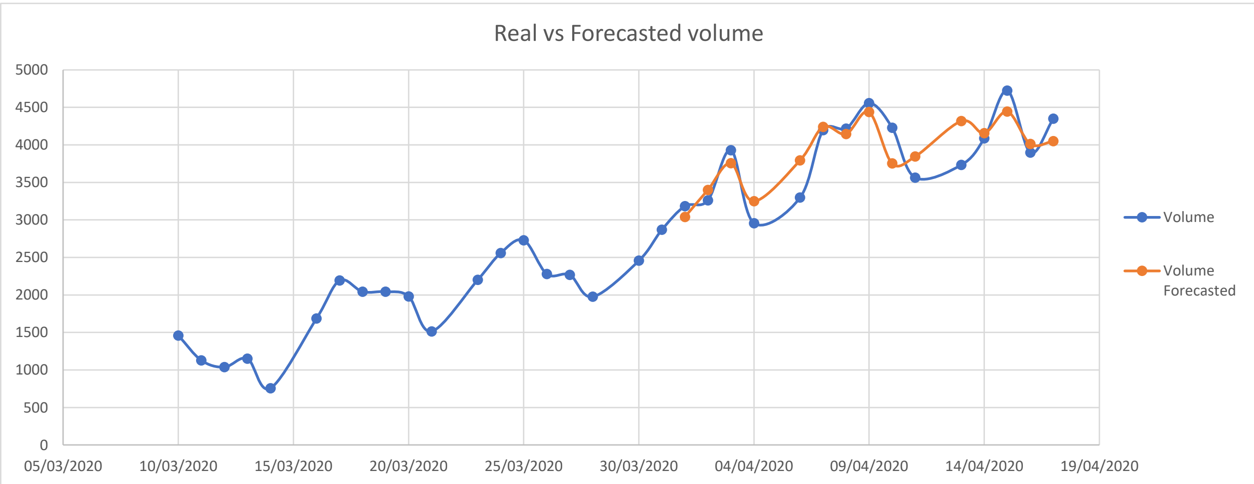

Figure 19: Forecast values compared to the real ones

Real and trained trend lines are compared in Figure 19, where the orange line approach the blue one, although the difference between values could be stretched (the bigger the dataset the better) the trend is correct in every value. This proves the accuracy of the forecasting.

Once the network is trained, now the volume is to be forecasted. For the network to work with the tanh activator, the dataset is changed to oscillate between -1 and 1, and the dataset transformed into this structure:

	var1(t-6)	var1(t-5)	var1(t-4)	var1(t-3)	var1(t-2)	var1(t-1)	var1(t)
6	-0.695206	-0.487718	-0.681851	-0.713332	-0.616981	-0.932745	-0.679943
7	-0.487718	-0.681851	-0.713332	-0.616981	-0.932745	-0.679943	-0.497257
8	-0.681851	-0.713332	-0.616981	-0.932745	-0.679943	-0.497257	-0.696160
9	-0.713332	-0.616981	-0.932745	-0.679943	-0.497257	-0.696160	-0.612211
10	-0.616981	-0.932745	-0.679943	-0.497257	-0.696160	-0.612211	-0.385643
..	...	...	...	...	...	...	...
62	0.758168	0.919866	0.763415	0.446697	0.527307	0.696637	1.000000
63	0.919866	0.763415	0.446697	0.527307	0.696637	1.000000	0.605533
64	0.763415	0.446697	0.527307	0.696637	1.000000	0.605533	0.821607
65	0.446697	0.527307	0.696637	1.000000	0.605533	0.821607	0.591700
66	0.527307	0.696637	1.000000	0.605533	0.821607	0.591700	0.864059

Figure 20: Table with transformed values

This way, the forecasted value will be the outcome of var1(t+1) at the last row.

The forecasted value for the current dataset is 4123, corresponding to the 20th of April 2020.

10. Late truck arrivals

A truck getting later than expected, and therefore unloading outside the allocated time slot affects directly to the routing and sorting process, leading to a late dispatchment, and consequently to delayed or failed delivery.

So, this raises the question of how to proceed when a truck arrives later than expected, which outcome is better in terms of profit and in terms of client satisfaction. There are three possible approaches to this question:

- To wait for the delayed truck: This solution enhances the probability of delivering more parcels for that day. The whole process would be delayed, waiting for the delayed trucks to be sorted and moved to their route areas. Every route containing a parcel from the delayed truck would end up delaying the delivery of all the parcels, which would reduce the margins per parcel, for a parcel delivered out of its time range has a penalization from the retailer.
- To deliver without the delayed parcels. In this case, the non-delivered parcels due to the truck's late arrival are for the retailer to blame, so no penalization is imputed to the company. However, this approach harms the client directly, and the company's reputation.
- To set a quantity of parcels that may be sorted and dispatched from the late trucks, depending the total volume and the arrival time. This proposal is halfway between the other two, delivering as many parcels as possible without compromising the margins, the time slots nor the company's reputation.

10.1. Number of parcels processed depending on the late truck arrival time

To set the number of parcels that may be processed we have to set a number of operators, since both data are correlated. The values this study has been carried out upon are the ones depicted previously in the parameters' section.

The number of parcels that can be processed without interfering with the last mile delivery are directly related to the timings of the whole process. The first time slots attempts are set at 10:20, so this is the first constraint regarding the delivery. It has also been stated the average time taken at arriving to the first destination, 20 minutes and an extra time for any kind of incidence (Figure 20).

Timings		
Start time	5:32	am
Fin Sorting	7:38:10	am
RC Start time	8:07	am
Dispatching Start time	8:50	am
Driver leaving WH time	10:00	am
1st att (drivers)	10:20	am
Roundcheck available time	72	min
Buffer time(drivers)	10	min
First delivery travel time	20	min

Figure 21: Dispatching timings

RC ST (Roundcheck start time): The latest time the roundcheck may take place to be able to process the whole freight in time for the first driver attempt to be achieved.

This restriction set the timings for the rest of the processes, as indicated in figure 20, being the driver leaving warehouse time set at 10 am and so the rest of the process according the time needed per task. The cells coloured in light yellow are those that have been calculated from the other parameters. Down below, these parameters are calculated:

$$RC\ ST = Dispatching\ ST - \frac{Volume}{RC\ productivity * number\ operatives * 24h} = \frac{3500}{540 * 9 * 24}$$
$$= 8{:}07\ am$$

OP ST (Operative starting point)

$$OP\ ST = RC\ ST - \frac{Volume}{Sorting\ productivity * number\ operatives * 24h} = \frac{3500}{150 * 9 * 24} = 5{:}32\ am$$

$$Operative\ ST = RC\ ST - \frac{Volume}{Sorting\ productivity * number\ operatives * 24h} = \frac{3500}{150 * 9 * 24}$$
$$= 5{:}32\ am$$

$$Sorting\ ET = \frac{Sorting\ total\ hours}{24\ hours} + Operative\ ST = \frac{2.33}{24} + 5{:}32\ am = 7{:}52\ am$$

$$RC\ AT = \frac{Dispatching\ ST}{Storting\ ET} = 8{:}50\ am - 7{:}52\ am = 58\ minutes$$

ETSEIB

$$Dispatching\ ST = Driver\ 1st\ attempt - \frac{Driver\ Load\ Time + Buffer\ time + 1st\ delivery\ attempt\ time}{24 * 60} = 10{:}20 - \frac{59 + 10 + 20}{24 * 60} = 8{:}50$$

So, assuming the volume remains stable, figure 21 presents the parcels from the late trucks that can be processed and dispatched along with the already sorted ones depending on the late trucks' arrival time. This data will change along with the daily volume.

By using this method, the losses are minimized and most of the clients get the proper service.

LATE TRUCK ARRIVALS		
Truck arrival	Roundcheck time *	Max parcels **
8:00	9:00	1352
7:45	8:56	1616
7:30	8:53	1880
7:15	8:50	2145
7:00	8:47	2409

Figure 22: Quantity of parcels that can be processed depending on the truck arrival

* already sorted parcels + delayed

**so the driver can deliver the first attempt in time

The late truck arrivals table is achieved after an iterating process through Roundcheck time and Max parcels parameters that iterates until both equations converge (separately).

RC time (Roundcheck time) indicates the time margin for the roundcheck depending on the driver departing time and the max parcels.

$$RC\ time = driver\ departing\ time - \frac{Volume + Max\ parcels}{\frac{RC\ Productivity}{60}} * n\ OP * \frac{1}{\frac{24}{60}}$$

Max parcels from a truck that arrives late (time arrival indicated in figure 7) a warehouse can process.

$$Max\ parcels = (RC\ time - Late\ Truck\ AT) * 24 * 60 * n\ OP * Sorting\ Productivity * \frac{1}{60}$$

By using this method, and daily updating the table field Volume, operatives at the warehouse can make decisions of how many parcels can they processed from a truck that arrives late so the losses are minimized along with the final client's delivery satisfaction enhanced.

11. Warehouse automation

Warehouse automation is a broad term. It may go from automating the tiniest part of a process carried out at the warehouse, such as operatives clocking in and out at their arrival and leaving so to have a control of their work and productivity, but it could also refer to automating the whole parcel processing.

To lean towards one extreme or the other, we have to study the automation level that better fits our process needs and the fixed objectives in addition to study the viability in economic, time and space terms. Automation levels can be classified as [Warehouse Automation Debunked]:

- Level 1: Systems automation. The warehouse main activity is still carried out by human labour. Tools like Warehouse management system (WMS), thought to optimize the warehouse dispatching process in terms of supply chain control and management. They go from stats reporting to inventory tracking and labor management, to IOT implementations by using sensors and scanners. However, the sole implementation of level 1 automation wouldn't necessarily translate into a decrease of labour, at least not a significative one.

  Nowadays, this is the most common level reached by distribution centres due to the low cost required for setting this model and the benefits it can procure.

  This is the model the warehouse from this case of study has reached, although not at its finest. The current parcel sorting process includes the scanning of the parcels so to sort them into different boxes depending on the route they have been assigned.

- Level 2: Mechanized automation. It constitutes the automation of horizontal movement tools, being the most common one the conveyor, and other mechanical tools such as automated film wrapping machine. This kind of solutions, although significatively more expensive than the ones proposed in level 1, mean a relevant reduce in labour costs.

- Level 3: Semi automation. Based on Automated storage/Retrieval systems. The labour gets significantly reduced, but it does it at the expense of the elevated installation costs.

- Level 4: Fully automation. When the level is reached, no worker is needed to process and dispatch an order. They feature high processing speed, high density AS/RS, extensive conveyor and sortation.

Level 1 is compatible and complementary to the other three levels, so prior to decide the grade of automation systems (from level 2 to level 4), the current level 1 system should be improved.

The following steps will study implementing systems automation along with different automation tools based on the depicted automation levels, so to decide which level is the most efficient for the project.

11.1. IOT Systems automation

One of the biggest constraints currently worsening the whole process is the fact that the state of the parcel has to be set manually, as mentioned in section 3.2.

Being in control of the real time state of every order can lead to chaos taking into account the vast volume that is daily processed coming from different retailers at different times, in addition to the parcels that weren't delivered the previous day due to an incidence and are set to be delivered along with the coming orders.

Getting a real (or at least close to real) time track of an order state in addition to its location is necessary to avoid incidences which may cost money to the company and would in the end worsen the clients purchasing experience. The solution to such a problem may be found through the Internet ff Things (IoT) technology.

IoT endows objects with a unique representation in the digital environment, with a series of features attached to this ID and connecting them through a network. This allows the objects to interact with the network and with each other (machine to machine interaction) in an environment where human labor is not required, allowing moving a vast quantity of data within milliseconds.

This technology adapts perfectly to our requirements, and implementing it could result into orders changing their state and location automatically. However, for an object to have an ID attached and be able to communicate through a network, it needs a microprocessor attached to it. But parcels are disposable envelopes, and adding a microcomputer to each one of them would resolve into a huge unaffordable cost increase in addition to an inefficient solution since the volume of parcels daily delivered would require a huge processor and servers in order to achieve a proper tracking. So, if tracking them by units is out of the table, how can the orders be real time tracked within a reliable method? Batch tracking and scanners.

Parcels follow a one-way flux line; from the moment they arrive to the warehouse until they are dispatched to be delivered. Sorting is the stage where the orders are prone to get lost, due to the vast number of parcels that are to be distributed through the route packs, most of the cases due to the parcel to be placed into a wrong route (due to this, round check must be carried out to make sure no parcels where misplaced, which constitutes one of the main bottle necks in terms of productivity and cost related to labor).

The following proposals will be analyzed and studied based on the WH design problem, introduced at the beginning of the project.

11.1.1. Check point

So, it only makes sense to set a check point at the sorting process the moment a parcel gets placed into the route box. This way, the operatives will be able to check if the parcel were placed into the right route and if any route has missing parcels. It only takes a QR reader per route box to improve the process to the point on of its most time-consuming phases can be removed from it.

This changes the way the sorting process was initially thought, adding an extra step before putting a parcel in its box, which consists on scanning the parcel's QR twice, first with the mobile phone app as usual, and then at the route box attached scanner. This way, scanning it twice pairs the parcel to that route.

The new sorting is carried out sequentially, as shown in *Figure 22*:

- Unload truck: The freight is unloaded from the track and placed at the unload area.
- Parcels are initially scanned, get a picker assigned and tagged with a QR containing information regarding the route it has been assigned to.
- Parcel moved to pre-route box: Due to the route boxes being away from the unload area, they are pre-sorted into the pre-route boxes, which divide the parcels into groups whose routes are adjacent to each other so the sorting gets done faster.
- Once the unload area is empty and all the parcels are pre routed, each pre-sorting box is moved to each of their sorting area. A sorting area constitutes four routes box from one alley plus the four in front of them.
- Double scan: Scan the parcel with the phone and at its assigned route box station and place it at the box. In case the parcel doesn't belong to the route it was meant to be put, a sound alarm will activate.

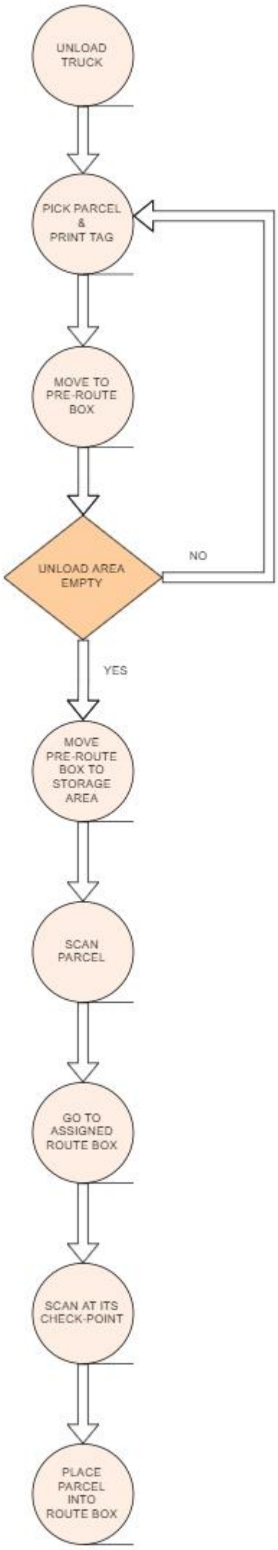

Figure 23: New sorting process

When the double scan is carried out, the parcel is added to the route's check point. Once all the parcels are sorted, the operatives can check out at the check point route list that all its parcels were sorted into the box.

Figure 23 shows how the scanner would be placed relative to the pallet.

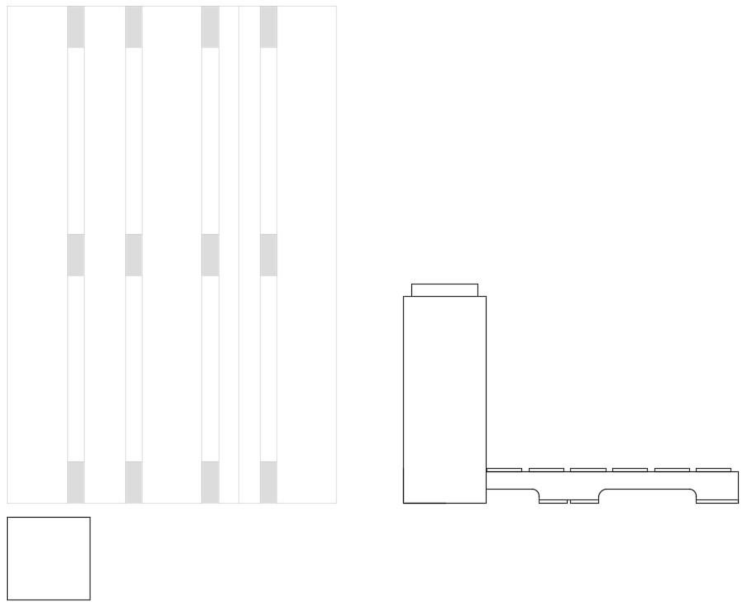

Figure 24: Scanner layout

11.1.2. Equipment

To implement this service, different possibilities have been studied. The main points taken into account when looking for solutions were the following:

- Robustness: the solution implemented must be robust, so after installing the whole layout there are no problems regarding the service.

- Affordable: automating a process is usually y high cost investment. In order to get profit within a low ROI, the solution must come with an affordable price or else it won't be implemented.

The chosen scanner model is YK-EP3000 2d barcode scanner module Wiegand, y fixed horizontal scanner in the shape of a box so it can be placed by the route box. It is a simple device that apart from the barcode reader has a USB port so it can be connected to another device to exchange data

Although the average number of routes is 88, the purchased quantity of scanners will be oversized by a 15% in case the volume increases or a scanner gets broken, meaning that the number of scanners to buy is 101. The cost for each scanner has been set to 33.15€.

Figure 25: Scanner model

A device able to collect the information from the scanner to process it and communicate with the collector devices is required. A good solution for this, that also meets the requirements set for the scanner is the ESP32 controller. The ESP32 is currently one of the most complete low-cost micro controller solutions in the market. Here are some of its features:

- Small size (25.5 x 18.0 x 2.8mm).
- Low cost: each microcontroller has a cost of 3.06 €.
- WIFI and Bluetooth low energy integration: Allowing the device to communicate with the server and directly with the operator phone.
- Low energy consuming: Due to the Ultra-Low Processor, the device can enter the Sleep Mode, which allows the device to be off (light sleep mode), waiting for the scanner to detect a barcode.

Figure 26: ESP32

Integrating both devices so they are connected to each other is not enough. To be able to share and store the information obtained by the barcode scanner it is necessary set the machine to machine (M2M) connection so to establish the IoT environment.

To connect an object to the IoT, it is necessary to have internet connection. For this project, WIFI has been chosen as the network connection method due to the routes changing nature (the routes layout may vary depending upon the volume to process). The object must connect to a server in order to set the communication between devices. All the changes affecting the objects connected to the IoT must be updated to the server. This allows real time communication between all devices.

Data transmission between server and client is carried out through HTTP Requests (Hypertext Transfer Protocol Request), which follows the request-answer protocol between devices.

These communications must be encrypted by an SSL protocol (Super Socket Layer), so no external device can have access to the data.

11.1.3. Cost analysis

Although the double scan decreases the sorting productivity per parcel, it erases the round check operative. The next steps focus on studying the throughput of this new process compared to the original one.

Because the original sorting process already included time for moving the scanned order to its route box, the time the has to be added to the sorting productivity is only the one involving the second scan. Scanning at the fixed route-box scanner takes an extra of three seconds per parcel, which are accountable for scanning it and waiting for the sound alarm approving or denying the check point pairing.

The new sorting numbers would be the following:

TIME RELATED	Quantity	Unit
Sorting capacity	27	s/parcel
Sorting productivity	133	parcels/h/op

Figure 27: Sorting productivity

This changes the timings related to the total hours required depending on the operators along with the processes timing restrictions set to be able to deliver the parcels on time. Figure _ shows the comparison between the new and the old process in terms of number of hours required depending on the operators working.

To fill the requirements set from the temporary agencies establishing that the workers shift must be above 3.5 hours. Old sorting required 3.5 hours and 9 workers to fulfil the hours needed for processing the orders, whereas the new sorting needs just 8 workers and 3.5 hours to accomplish the job at a much lower cost.

Overall, the savings related to removing the roundcheck process are of 44 € per day, which represents a 11% less than the old sorting.

	New Sorting		Old Sorting	
n Op	Time [h]	Cost [€]	Time [h]	Cost [€]
5	6,00	369	400	6,50
6	5,00	369	406	5,50
7	4,00	344	431	5,00
8	3,50	344	394	4,00
9	3,00	332	387	3,50
10	3,00	369	431	3,50
11	2,50	338	406	3,00
12	2,50	369	443	3,00
13	2,00	320	400	2,50
14	2,00	344	431	2,50
15	2,00	369	369	2,00

Figure 28: New sorting vs Old sorting

11.2. Conveyor

The next logical point automation wise after enabling the IoT communication in the warehouse is to study the possibility of installing a conveyor capable to do the sorting into route automatically. By installing a conveyor most of the dispatching process would be automated, with the exception of the trucks' unload, printing the company's barcode and moving the parcels to the sorter.

The conveyor would carry out the following tasks:

- Scan the parcel.
- Sort the parcel into the assigned route box.

The parcels' flow through the conveyor can be observed in figure 29. The orders are scanned at the beginning, so they are output when they pass by the outlet leading to the route-box the parcel has been attached to.

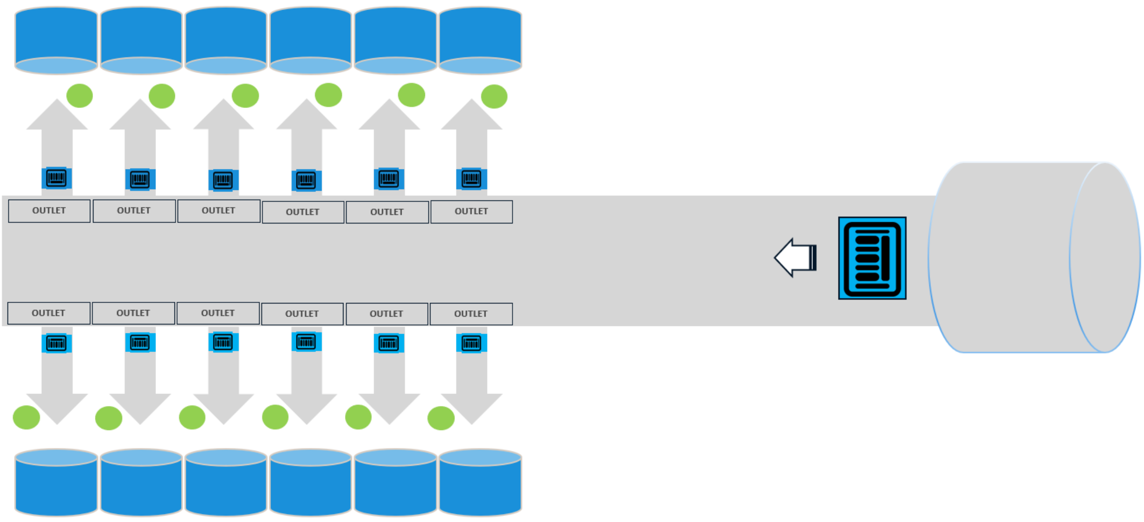

Figure 29: Sorter flow chart

There are currently many conveyor systems in the market, so we must choose the one that fits better our requirements without getting an over-featured one that may increase the price and reduce our margins.

The solution we have come up with uses the Cross-belt Sortation Conveyor for automating the warehouse's layout.

Cross Belt conveyors used to sort parcels, apparel, and small items, at high speed [4].

The parcels are carried out in individual carts, which are small bi-directional conveyor belts. Each spot is a belt that gets activated when it lines up with the parcel's route-box destination. This method main advantage is the individual place for each parcel, so the orders will be separated and can't get mixed nor missorted. When the tray reaches the destination position, the mechanical belt is activated and the parcel ejected to its box. Thanks to the control system, the sorter is able to output the parcel in the assigned route-box.

Figure 30: Cross belt sorter

The new sortation process is described in Figure 31, the tasks colored in yellow are the one manually performed whereas the green ones are processed by the sorter. Parcels are sorted automatically.

This new layout reduces the tasks carried out by labor, so a new cost analysis must be performed to establish the viability of the automation project. In addition, it increases the total productivity, due to the sorters' throughput, which is of 15 pieces/s.

This conveyor may be used to process small-medium parcels and bags, the 80% of the daily freight received by the retailers (this is assured by a deal with the retailers), the other 20% must be handed by the operatives.

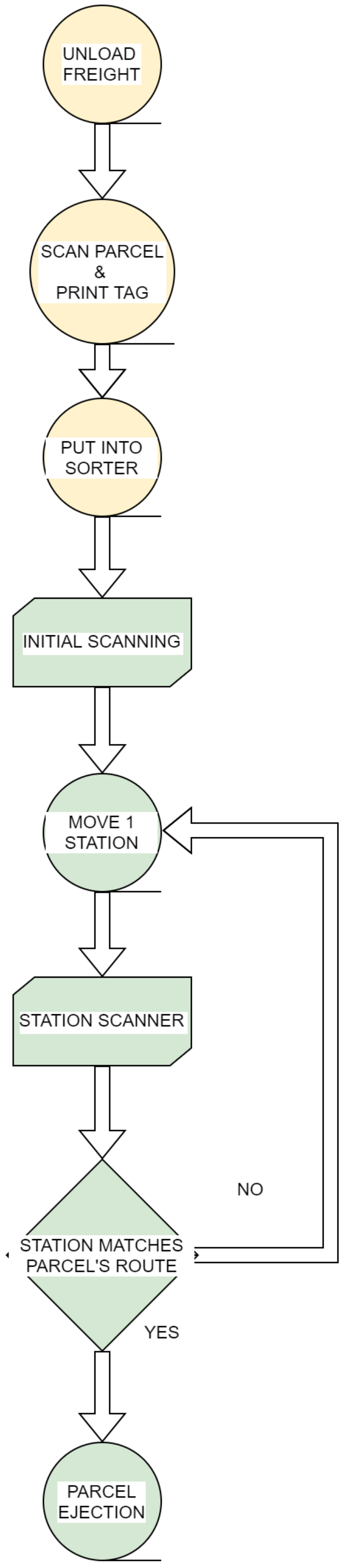

Figure 31: Sorting process

11.2.1. Cost analysis

To make a decision on whether to automate the warehouse by setting a conveyor we must compare the both the current layout and the check-point based model to the conveyor to decide which of the three is the best in terms of fixed and variable costs, labor costs and productivity.

The cross-belt cost is 9500 €/set, which consists on the whole layout formed by: 2 belts with 45 routes each, which makes 90 routes in total.

		Parcel size	
	Total	Small/medium	Large
Volume	3500	2520	630

Figure 32: Volume division by the size of the parcels

Two operatives will carry out the sorting process for the large parcels and another two will be in charge of watching out for the cross-belt sorter not to fail and placing the parcels into the sorter.

Large volume	Sorting productivity	n OP
630 parcels	150 parcels/op/h	2

Figure 33: Large parcels operators assigned

The gross yearly savings exceed the machine acquisition, so the sorter conveyor is worth a purchase for the warehouse automation, which will lead to a higher productivity and lower costs.

Checkpoint	8	
Conveyor	2	
Diff	6	
x n hours	3.5	
x op cost/hour	12.3	
Daily savings	258.3	€/day
x operating days	24	
x 12 months	12	
Yearly savings	74390	€/year

11.3. Next steps

The studied solutions only cover for the viability of launching the project, to implement it furtherly data that is out of the scope should be taken into account, such as how long does it take for installing everything, for instructing the operatives and what kind of agreement does the company have with the external agencies.

The semi automation solution chosen is apt for the current volumes. This solution stands between levels two and three, the storage has been automated but not the retrieval. This may be extended in case the volumes keep increasing and a whole AS/RS is needed.

In addition, the IoT model has also been implemented, and the orders changing of state is now carried out automatically. This could also be further extended into a more complex IoT system, due to the constant improvements this technology is developing.

12. Conclusions

Ecommerce is growing fast and steady, and so is the technology surrounding it. To be able to adapt to such change, a warehouse must be prepared to take vast amounts of parcels and process them in a short period.

The solutions that have been proposed in this project constitute a first approach that solve primarily problems such as real time automated parcel controlling and improving the dispatching process at a reasonable and low cost, which will translate in a better service and a higher adaptive capacity. These are low steps for achieving the automation that big companies are using nowadays, but it is the starting point for reaching it.

Both the Late truck arrival file and the artificial neuronal network are very easy tools to implement and, if so, may reduce over costs in terms of bad decision makings and having more operators than needed, which would impact importantly the total gross margins of the dispatching process.

The control IoT system and the semi automating solution by installing a conveyor have a bigger risk attached due to budget, and would require the warehouse layout to adapt to the new constraints, but the throughput of the process would increase potentially and the costs recovered within the first year.

It is worth noting that every proposal has the capacity of being extended and modified easily, so to adapt to the new requirements the market may establish in the future.

13. Costs

In this section the total costs attached to the project are detailed.

Firstly, the labour costs have been calculated. These tasks were carried out by two industrial engineers, so the hours reflect the work of both engineers. Every project has been divided in the main tasks that needed to be performed in order to implement it. The salary has been set to 25 €/h.

Project	Task	Time [h]
Neuronal network	Conceptual development	45
	Script development	150
	Release	30
Late truck arrival	Draft	28
	Measuring the processes	70
	Developing the late truck file	14
IoT automation	Study current process	7
	Conceptual development	30
	Programming the microcontroller	100
	Setting up the scanners in the WH	14
	Release	7
Warehouse automation	Layout requirements study	14
	Choosing a conveyor	4
	Designing new WH layout	14
	Installing conveyor and changing layout	28
	Total	555

Engineer cost €/h	25
Subtotal	13875
IVA (21%)	2913,75
TOTAL COST	16788,75

Figure 34: Labour costs

Only the automation project has material costs attached, since the neuronal network and the late truck arrivals were implemented upon already existing tools and materials. Both the ESP32 and the Scanner cost reflect the overall purchase, which is 101 per material. The crossbelt conveyor budget is fixed, and has been negotiated directly with the supplier.

Material	Cost [€]
ESP32	309,1
Scanner	331,5
Crossbelt conveyor	9500
TOTAL COST	10140,56

Figure 35: Material costs

The overall cost for implementing these solutions would be the sum of the labour related cost plus the materials.

TOTAL COST	
Type	Cost
Labour	16788,75
Material	10140,56
TOTAL	26929,31

Figure 36:Total cost

14. Annex

14.1. Neuronal network code

In this section, the code implemented to achieve the forecast of the daily volume of parcels to process is explained.

14.1.1. Libraries

To be able to implement the code, external libraries have been used. They are extremely helpful for specific tasks.

- Pandas library is used to process csv files, be able to read them and format them as a dataset.
- Matplotlib is a common library in python, used for potting data.
- Keras is the library used to create the neuronal network given the parameters and the dataset.
- MinMaxScaler is used to scale the given dataset between -1 and 1.

```
import pandas as pd
import matplotlib.pylab as plt
from keras.models import Sequential
from keras.layers import Dense,Activation,Flatten
from sklearn.preprocessing import MinMaxScaler
```

14.1.2. Creating methods

Here, the methods that are to be used during the program execution are defined.

14.1.2.1. make_into_t_series

The make_into_t_series method transforms the dataset into a table with n_in columns with a t-1 | t-2 | … | t-n architecture. This is used as an input for training the neuronal network.

```
# convert series to supervised learning
def make_into_t_series(data, n_in=1, n_out=1, dropnan=True):
    n_vars = 1 if type(data) is list else data.shape[1]
```

```
    df = pd.DataFrame(data)
    cols, names = list(), list()
    # input sequence (t-n, ... t-1)
    for i in range(n_in, 0, -1):
        cols.append(df.shift(i))
        names += [('var%d(t-%d)' % (j + 1, i)) for j in range(n_vars)]
    # forecast sequence (t, t+1, ... t+n)
    for i in range(0, n_out):
        cols.append(df.shift(-i))
        if i == 0:
            names += [('var%d(t)' % (j + 1)) for j in range(n_vars)]
        else:
            names += [('var%d(t+%d)' % (j + 1, i)) for j in
range(n_vars)]
    # put it all together
    agg = pd.concat(cols, axis=1)
    agg.columns = names
    # drop rows with NaN values
    if dropnan:
        agg.dropna(inplace=True)
    return agg
```

14.1.2.2. create_neuronal_model

This method builds up the neuronal network according to the given architecture (explained in the LSTM section).

```
#creates the LSTM model
def create_neuronal_model():
    model = Sequential()
    model.add(Dense(10, input_shape=(1, INPUTS), activation= 'tanh'))
    model.add(Flatten())
    model.add(Dense(1, activation='tanh'))
    model.compile(loss='mean_absolute_error', optimizer='Adam',
metrics=["mse"])
    model.summary()
    return model
```

14.1.3. Input data processing

In this part of the code the input dataset is modified so it fits the input requirements of the neuronal network. The Pandas library reads the csv file containing the dataset, separating it in columns by the ‘;’ delimiters of the file.

```
#data file with the current data
df = pd.read_csv('volbcn.csv',delimiter=';',  parse_dates=[0],
header=0,index_col=0, squeeze=True,names=['fecha','unidades'])
'df.head()'

INPUTS = 6

# load dataset
values = df.values
```

```
# ensure all data is float
values = values.astype('float32')
# normalize features
scaler = MinMaxScaler(feature_range=(-1, 1))
values = values.reshape(-1, 1)
scaled = scaler.fit_transform(values)
# frame as supervised learning
reframed = make_into_t_series(scaled, INPUTS, 1)
```

14.1.4. Neuronal network implementation

In this part of the code the training and testing is defined and implemented. The train days are designed as the 80% of the total of the days in the dataset.

```
#Neuronal network definition
values = reframed.values
n_train_days = int(len(values)*0.8
train = values[:n_train_days, :]
test = values[n_train_days:, :]
# split into input and outputs
x_train, y_train = train[:, :-1], train[:, -1]
x_val, y_val = test[:, :-1], test[:, -1]
# reshape input to be 3D [samples, timesteps, features]
x_train = x_train.reshape((x_train.shape[0], 1, x_train.shape[1]))
x_val = x_val.reshape((x_val.shape[0], 1, x_val.shape[1]))
print(x_train.shape, y_train.shape, x_val.shape, y_val.shape)
```

The number of epochs is set to 1000, and the model is created and the training started.

```
#set number of epochs
EPOCHS=1000
model = create_neuronal_model()
history=model.fit(x_train,y_train,epochs=EPOCHS,validation_data=(x_val,y_
val),batch_size=None)
```

Here, the results from the training are tested and compare with the real ones to check if the training was successful.

```
#Test the training
results=model.predict(x_val)
plt.scatter(range(len(y_val)),y_val,c='g')
plt.scatter(range(len(results)),results,c='r')
plt.title('validate')
plt.show()

values = df[-18:].values
values = values.astype('float32')
```

```
# normalize features
values=values.reshape(-1, 1)
scaled = scaler.fit_transform(values)
reframed = make_into_t_series(scaled, INPUTS, 1)
reframed.drop(reframed.columns[[0]], axis=1, inplace=True)


values = reframed.values
x_test = values[len(values)-1:, :]
x_test = x_test.reshape((x_test.shape[0], 1, x_test.shape[1]))
y = model.predict(x_test)

inverted = scaler.inverse_transform(y)


reframed = make_into_t_series(scaled, INPUTS, 1)
reframed.drop(reframed.columns[[5]], axis=1, inplace=True)


values = reframed.values
x_test = values[len(values)-1:, :]
x_test = x_test.reshape((x_test.shape[0], 1, x_test.shape[1]))
forecast_value = model.predict(x_test)


forecasted_val = scaler.inverse_transform(forecast_value)

#PRINTS THE RESULT
print(forecasted_val)
```

15. Bibliography

[1] Hyndman, R.J & Athanasopoulos, G. (2020, 08 16). *Forecasting: principles and practice, 2nd edition*. Retrieved from https://otexts.com/fpp2/nnetar.html

[2] Lazzeri, F. (2020, 07 03). *Medium*. Retrieved from https://medium.com/

[3] Jain, S. (2020, 07 22). *Activation functions in Neural Networks*. Retrieved from https://www.geeksforgeeks.org/activation-functions-neural-networks

[4] Cross Belt Sorters. (2020, 09 01). Retrieved from https://www.bastiansolutions.com/

[5] Brownlee, J. (2020, July 24). *Machine Learning Maestry*. Retrieved from https://machinelearningmastery.com/

15.1. Complementary bibliography

Brownlee, J. (2020, July 30). *Multi-Step LSTM Time Series Forecasting Models for Power Usage*. Retrieved from https://machinelearningmastery.com/how-to-develop-lstm-models-for-multi-step-time-series-forecasting-of-household-power-consumption/

Keras. (2020, July 19). Retrieved from https://keras.io/api/layers/recurrent_layers/lstm/

N. Kolban, Kolban’s Book on ESP32, USA: Leanpub, 2017.