The visualization of efficiency and performance with the use of dashboards

Pol, S.J.B. van der (Stijn, Student B-IEM)

Supervisor 1: Derya Demirtas (University of Twente)
Supervisor 2: Patricia Rogetzer (University of Twente)
Supervisor 3: Max Fortuin (Peddler)

PREFACE

Dear reader,

This bachelor thesis is the final part of the completion of my bachelor of Industrial Engineering and Management at the University of Twente. The company where this research was done is a startup in the south of Amsterdam called Peddler. This research named '' The visualization of efficiency and performance with the use of dashboards'', is thriving to help Peddler gain more insight into their logistic operations.

I want to take this opportunity to thank the People at the university and at Peddler who have helped me write this thesis and do this research. A special thanks to Derya Demirtas for being my first supervisor and for guiding me through this process. It was a nice experience to work with you, and I hope this will not be the last time. Also, I would like to thank Patricia Rogetzer for being my second supervisor. I really appreciate all the tips and criticism that you have given me. You were always very accurate and precise when providing me with feedback. Last I would like to thank Max Fortuin from Peddler. You have helped me get to know Peddler, and you were always ready to help when I had a question. It was a great experience.

Enjoy reading,

Kindest regards,

Stijn van der Pol
Enschede, September 2022

MANAGEMENT SUMMARY

Introduction

Peddler is a start-up which helps local shops to catch up with the large innovative multinationals by creating a platform to sell their products online. The implementation of an online platform and the collaboration between Peddler and these local stores help these local stores to reach a target group that would otherwise be missed. Peddler also provides a delivery service by picking up online orders and delivering them to customers with the use of an eco-friendly cargo bike. It is also possible to only make use of the delivery service. In this case, stores replace their normal delivery company with Peddler to deliver online orders.

Currently, Peddler is not as aware as they like to be about their performance in the logistics department. To summarize, they want more insights into delivery efficiency and performance of their operating regions, stores and riders. This research puts the focus on the visualization of Peddler's, request which is done with the use of multiple dashboards.

Approach

In order to sufficiently develop the dashboards, a systematic literature review was carried out. Here multiple ways of analyzing data were discovered, which helped to correctly prepare the data before it was imported into the tool. In addition, multiple interviews were conducted to obtain the right knowledge about the company, which helped to determine suitable and necessary KPIs. After this was done, five different dashboards which, each of their own focus and factors, were created. These dashboards were then analyzed, resulting in helpful advice based on the discovered insights.

The dashboards

As mentioned above, five dashboards were created in total. The selected tool to work with is Qlik, with the main reason being that employees are already working with it, so it is easier to work with for them than another tool. The following dashboards were created:

- Location performance: The first created dashboard is focused, as the name implies, on geographical performance. The dashboards contain two maps which help to visualize the performance of regions, cities and postal codes. The measures are the total distribution of orders and the percentage of orders that have been delivered on time.
- Performance over time (including timeslot performance): The second dashboard focuses on the performance of Peddler over time (cohort analysis). The available data covers an entire year, so this dashboard helps to provide Peddler with knowledge of their performance during this year. It also includes knowledge about the performance of each possible timeslot.
- Timeslot performance per postal code: The reasoning behind the existence of the third dashboard is to see how each timeslot performs per postal code. Based on this dashboard, a preferable timeslot can be appointed to each postal code.
- Store performance: with the use of multiple pivot tables, a line chart and a scatter plot, Peddler obtains knowledge about the performance of each collaborating store with the use of the fourth dashboard.
- Rider performance: the performance of each employed rider has been visualized in dashboard five. Within this dashboard, it is possible to compare riders with each other and see how the performance has been over time.

Analysis and results

With the use of the five dashboards, a list of points of interest can be formulated. To summarize, the most important findings are listed:

- An analysis based on dashboard three has been carried out, which resulted in a list of all operating postal codes, and their new preferred timeslot has been made. This list was focused on the most on-time deliveries of each timeslot per postal code. For all 330 postal codes, an optimal timeslot has been allocated, including an insight into the increase in on-time orders compared to the second most optimal timeslot.
- The best performing city with more than 100 orders in total is Rotterdam. It has been able to deliver 82,5% of orders. City The Hague is performing poorly, with an average of 47,1% of orders being delivered on time. Where the hub of The Hague is located quite poorly, the hub of Rotterdam is located quite nice around the center of the city. Even more impressive considering that it has a 6,79% share of the total orders and a 9,16% share of orders that have been delivered on time. The worst performing city is Delft, having only 22,6% of their orders delivered on time.
- Looking at regions, provides more insight in the contribution of each postal code. In Amsterdam, for example, the ten lowest buying postal codes combined have bought 0.225% of the ten highest buying postal codes. The total number of orders in the region of Amsterdam is 43,172. The top ten postal codes are responsible for 35.05% of this number, while the bottom ten are only 0.079% responsible for this.

Recommendations

A list of active postal codes and their new preferred timeslot has been sent to Peddler. This helps them to make it easier for them to deliver packages on time. They can implement this in their customer checkout phase to let the customer only select the preferred timeslot instead of all possibilities. It lets the inclusion of timeslots be more of a tool to benefit from instead of having it work against you. This increases the number of on-time delivery orders.

The five dashboards that have been created can be of great use for Peddler. It gives them insight into their performance and lets them see the direct consequences of certain strategic decisions. Differences can be measured over time, regions can be compared, and timeslots can be optimized. All dashboards can function as a strategic tool that could benefit Peddler in the future.

TABLE OF CONTENTS

1. INTRODUCTION

The chapter functions as an overview of what is to be expected in this research. It provides a small description of the company and the core problem they are facing. It also describes how this problem is analyzed. This analysis consists of a research design, determined research questions and the scope of the research.

1.1 BACKGROUND AND THE IDENTIFICATION OF PROBLEMS

1.1.1 BACKGROUND OF THE COMPANY

A problem many local stores have is that they have to compete against big supermarket or huge brands that offer a lot more services than them. Most of these stores get their customers from people who walk by the store or already know its existence. Many of these stores are too small to create a website or think it is too much trouble. This is where Peddler.com comes in. Peddler is a company that provides two services for local stores: an online platform for stores to sell their products and a delivery service to make sure that their sold products actually arrive at the customer. This creates a much larger target audience and helps these local stores compete against these big brands. The local stores that can enter into this cooperation range from a local butcher to a local toy store. When a customer buys a product from a store on Peddlers platform, the order will be added to the route of a Peddler delivery rider which will deliver it to their doorstep. Peddler riders are delivering orders on electric cargo bikes which means that the transport of the product will be eco-friendly as well.

Peddler is currently active in five cities in the Netherlands: Rotterdam, Amsterdam, The Hague, Utrecht and Groningen (Peddler, 2022). Since their start in Amsterdam a bit more than a year ago, they have expanded with the intention to create more brand awareness.

1.1.2 MANAGEMENT PROBLEM

Currently, Peddler offers two different types of deliveries (Peddler, 2022): Same day delivery without a time indication (or a different day if it is not possible to deliver it on the same day) and same day delivery in a specified time window. Same day delivery without a time indication means that the delivery can be delivered to the customer at any time of the day if it is delivered at the mentioned day. This service accounts for about 80% of the total orders. However, there are also orders that are being booked with a certain time of arrival. The way this order works is that customers can select a time slot in which they can expect the delivery. These orders have the highest priority, so an efficient delivery route is necessary.

Right now, Peddler is struggling to deliver the orders within the agreed timeslots. They want to see how these agreed time indications can be met with their current delivery planning system. Given that Peddler is still in the early phase of existence, the current brand awareness is smaller than it will be over five years. They are growing rapidly. Currently, there are new openings in many Dutch cities with the potential to expand to Germany in the future. For this reason, a potential increase in demand is expected in the near future. To prevent the system from being overloaded when this increase arrives, preventive measures need to be taken to improve the route planning.

1.1.3 PROBLEM IDENTIFICATION

As mentioned above, there are two different types of orders with different types of priority for Peddler to work with. Peddler has stated that scheduled orders which are bounded by a certain time slot, have priority over orders that can be delivered at any time of the day. The quantity of strict time bounded orders is a lot less than that of the day delivery orders. This, however, may change over time since the position of the company in the market has not been fully established yet. This, combined with the possible increase in demand overall, means

that Peddler needs to be prepared for these changes without the loss of efficiency. Loss of efficiency in this case means that other orders arrive too late, too many or too few employees are used, unnecessary number of extra miles are travelled or the pickup at the store takes too long because an order has not been prepared properly yet. These are all potential causes for an order arriving too late.

The routes consist of connected pickup and delivery points that are suggested automatically by an online platform. This platform creates the route by calculating the shortest distance. However, these routes do not have any constraints such as times slots. The adjusting of this route to fit these criteria is done manually. Route planners of Peddler react to new orders or changes in the current route. An example of a problem that could be present now where Peddler has no inside in, is that a store could structurally fail to have their orders ready for delivery on the agreed time. Currently, there is a lot of historical data available, but Peddler is not using its full potential. There is a possibility that current decisions harm their efficiency and this needs to be looked into.

Core problem: *Peddler currently has no clear view of how their decisions in route planning impact their efficiency and performance.*

1.1.4 GAP BETWEEN NORM AND REALITY

As defined in the core problem, reality is that Peddler has no view how their choices influence the efficiency and performance of their company. A dashboard brings more insight into the results of these choices. Based on key performance indicators (KPIs), efficiency and performance can be visualized to see where changes should be made. The norm for this research and for Peddler is a tool which helps to substantiate strategic decision-making. The reality is the absence of such a tool. The creation and use of a substantiated dashboard which helps to measure performance and shows points of attention, fills this gap.

1.2 RESEARCH QUESTION AND PROBLEM-SOLVING APPROACH

In this section of the report, the main research question will be discussed. In addition to this, the chosen problem-solving approach to solve the core problem of Peddler will be explained as well. From this problem-solving approach, certain sub-questions can be derived.

1.2.1 RESEARCH QUESTION

The main research question is based on the core problem Peddler currently has. As discussed above, Peddler currently has no clear view what the impact of their decisions is. There is a lot of data that is unused which could provide this knowledge. The goal of this research is to give Peddler more insight how they are currently performing, visualizing their efficiency and to form a basis to improve future decision-making substantiated by decisive indicators. To clarify this, the performance of Peddler is how effective they are in achieving their goals. Efficiency is much of their resources is used in order to achieve this goal.

Research question: *How can Peddler gain more insight and have a better view of the impact on efficiency and performance as a result of the choices they have made?*

1.2.2 PROBLEM SOLVING APPROACH

When looking at the main research question, crucial parts of the problem solving approach are identified. First of all, there needs to be an analysis of the current transportation system to get a better understanding how Peddler operates and how their routes are being calculated. In order to do this, a data set has been provided by the company containing all sorts of historical orders. This data set is analyzed which shows where certain improvements can be made. The analysis is the foundation for a dashboard that has the purpose to identify weak points in performance and efficiency. In order to define this efficiency and make sure it is aligned with the

vision, norms and values of Peddler, certain KPIs need to be determined. These KPIs can be calculated from the data. When finished, this dashboard is evaluated to make sure expectations are met.

The Managerial Problem-Solving Method (MPSM) of Heerkens & Van Winden (2021) could give a guideline for this research since it guides the researchers through a number of steps that need to be taken in order to deal with an identified problem. However, like mentioned before, the research consists of a lot of data mining and processing it. A research method which is more aligned with this problem approach is the Cross Industry Standard Process for Data Mining better known as CRISP-DM (Wirth & Hipp, 2000). It is a process model that can be seen as an in-depth version of the MPSM model which is more focused on research that is related to data mining. It consists of six phases: Business understanding, data understanding, data preparation, modelling, evaluation and deployment. To make this process clearer, the sequence of these phases can be seen in Appendix A1 and are described below in more detail.

PHASE 1: BUSINESS UNDERSTANDING

Peddler is a start-up and still has to develop and adjust a lot of their strategy and management. It is necessary that the research is aligned with the project objectives and requirements from Peddler. In this case, there are multiple stakeholders involved and a part of them have a saying in decisions that need to be made. They also put pressure on certain deadlines and milestones which can influence approaches to problems. It is important to know the vision and approach from everyone involved. These stakeholders need to be identified.

Opinions and visions from multiple stakeholders can be seen as qualitative data and are of high importance to make sure that the research fits with the organization of the company.

PHASE 2 & 3: DATA UNDERSTANDING AND DATA PREPARATION

These two phases are quite closely related in the case of this research. Data understanding means that the researcher obtains knowledge about the available data. This means that during this phase, the researcher gets familiar with the data, can discover some insights into the data, can already make segmentations and cleanse certain parts of the data.

In the case of Peddler, there is a big data set containing all historical orders. This quantitative data is analyzed and processed. In order to get a better understanding about the data, it needs to be segmented. This is done by translating the objectives and requirements of Peddler into KPIs and variables. Interesting findings are documented, and important factors are taken into account. Apart from this data set, there is also qualitative data that has been acquired from the company. These are used to see how the current transportation network performs which can help to already spot potential improvements. It makes it easier to segment the data set when it is known where points of attention are.

When all of this is done, the data preparation phase can begin. This consists out of the construction of a final segmented dataset based on the needed KPIs and variables which form the foundation for the dashboard.

PHASE 4: MODELING

In this phase, the dashboard is created. This is accompanied by the right constraints, input and output. During the construction of the dashboard, multiple perspectives are looked at. In other words, multiple scenarios need to be taken into account. In this case, a dashboard needs to be created which creates a visualization of the findings based on the segmented data set. The base for this visualization consists out of the previously determined KPIs.

PHASE 5 & 6: EVALUATION AND DEPLOYMENT

When the dashboard is constructed and finished, it needs to be evaluated to see if expectations are met. The quality of the dashboard is tested, and feedback is gained. During this phase, it is also discussed if some issues

within the company are not addressed properly or missed completely in the final design. After everything is checked and approved, the dashboard is implemented in the organization of the company. This is done in the deployment phase. The created dashboard during this bachelor thesis is evaluated to see if the quality is met but a proper implementation is outside of the scope of this research.

1.2.3 MANAGEMENT LEVEL

The created dashboard functions as a tool which can be used for the identification of knowledge. It mainly plays a role for the organization on a tactical level, but it can also serve as a substantiation on a strategic level. It serves as a way to allocate knowledge within the firm in an effective and efficient way. It is not something that is used every day so it does not function at an operational level. Interesting findings based on the created dashboards, could influence the strategy of Peddler. These findings could be a solid argument for a change in their business approach which could mean that the dashboard also functions on a strategic level. At the moment, which is the completion of this research, Peddler has not indicated to do so.

1.2.4 SUB-QUESTIONS

In addition to the main research question, certain sub-questions are created as well:

Sub-question 1: How is the current route planning system of Peddler organized?

- Who are the stakeholders in Peddler?
- What methods are helpful to analyze the current system of Peddler?
- How can useful data be gathered?
- What are important KPIs that are already measured by Peddler?
- Which new KPIs can be added?
- What are the opinions of employees about the current route scheduling?
- What are factors of importance for feasible route planning according to employees?
- How can the determined factors help to segment data?

Sub-question 2: What changes can be made to make the system adaptable to an increase in demand?

- What shortcomings can be identified during the making of the route?
- What shortcomings of the monitored operations can be identified?
- Do employees have recommendations for improvements?
- How can the determined KPIs be implemented in the route planning policy?
- What opportunities would occur if demand was increased?

Sub-question 3: Based on the acquired and segmented data, how can a dashboard help to visualize findings?

- What acquired data can serve as input?
- What type of dashboards are suited for the acquired data?
- What KPIs are important to be visualized?

Sub-questions 4: Based on the created dashboard, what recommendations can be made?

- What limitations does this dashboard have?
- What interesting findings can be derived from this dashboard?
- How can Peddler improve discovered shortcomings?

1.2.5 SCOPE

The dashboard is evaluated and a short-term implementation is made to see if it fits with the management of the company. However, in order to implement it successfully, an evaluation cycle needs to be used. Due to the limited time, resources and expertise that is available, this is beyond the scope of this bachelor thesis.

2. LITERATURE REVIEW

In this research, there is a lot of quantitative and qualitative data that needs to be gathered, processed and analyzed. As described in the research design, both types of data will have to be gathered in different way. After this gathering is done, there is a lot of data that needs to be analyzed. Before this analysis starts, certain methods are needed to make the size of this data manageable. This chapter provides an explanation of the gathering of the different types of data and the methods that are used to analyze it.

2.1 QUALITATIVE DATA

An important deliverable of this research is the dashboard which closes the gap for Peddler and provides them a better view in performance and efficiency. The foundation for this dashboard consist a lot out of quantitative data. However, without the knowledge how this quantitative data exists and what it means, it is of no use. Therefore, qualitative data is needed to gain information about the management and goals of Peddler. The qualitative data is gathered through an observational study and by having multiple interviews.

2.1.1 OBSERVATIONAL STUDY

The observational study has gathered information about the management and structure of the company. For this research, the goal is to give Peddler a better understanding what the consequences of their decisions are. The main focus of these decisions lies in their performance and efficiency. An important start for this research is to gain knowledge about all the relevant choices that are being made by different departments. The opportunity has been provided to experience the gathering of this data by working alongside with different departments.

The relevant departments for this research are the route planners and logistics. The route planning and logistic department make sure that promises are met which means that orders are on time and that everything runs smooth behind the scenes. In addition, they need to inform customers if this is not the case. A helpful method to analyze this gathered data is by expressing findings and choices in variables.

The first type of variable that make the data easier to work with, is the explanatory variables. These variables can be seen as some type of factor that has been changed by the researcher during the study (Cox & Snell, 2016). In the case of Peddler, these are the decisions that have been made within the relevant departments that have possible impact on the outcome of performance. Examples of these changed factors can be the choice of a different timeslot or the allocation of an order to a new rider.

The obtained data also contains control variables, variables that have influence on the explanatory variables (Rohrer, 2014). Apart from the researcher, there are also other entities that can influence certain factors during the observational study. This results in different outcomes. Within Peddler, the logistics and route planning department constantly make decisions based on the scenarios they are facing. The outcome from the change in some explanatory variables can have a big impact on performance.

2.1.2 INTERVIEWS

Apart from the choices and experience, input and advice for the plan of approach for the quantitative data is also needed. Interviews with employees of Peddler help to set guidelines for this plan of approach. The questions that were asked, are showcased in Appendix A1. The obtained qualitative data could be analyzed with the use of the Non-numerical Unstructured Data–Indexing, Searching and Theorizing (NUDIST) method (Lillis, 1999). To get a clear overview of the correlating answers, themes are added to them to classify them. The Miles and Huberman method implies that data first needs to be reduced. It stimulates to create sufficient bodies of data from unedited amount of text (Miles & Huberman, 1994). The interviews gave a better insight in the priority of certain global KPIs. The request to let the interviewee translate the norms and values of Peddler into a KPI, made it clearer what the most important KPI were to them. The combination of both methods made for a clear overview of concrete information. Information such as shortcomings that can be further analyzed with the use of new KPIs.

2.2 QUANTITATIVE DATA

Apart from the analysis of the interviews and the observational study, there is also a lot of quantitative data that needs to be analyzed to get a better understanding of the performance of Peddler. This quantitative data is shaped in an Excel file containing tens of thousands of rows. These cells contain information about historical orders which should enlighten the current statistical performance of Peddler. Because of the size of this dataset, multiple methods are used to make it manageable to calculate the needed KPIs.

2.2.1 CLASSIFICATION ANALYSIS

Within the given dataset, the size of the set makes it harder to manage and give it a structure. Because there are multiple stakeholders for each order, the number of variables for a single order increases with a lot. Classification focuses on the segmentation of data based on certain characteristics that belong to a classified object (Tufféry, 2011). These characteristics can be seen as explanatory variables. There have been instances where objects are divided into smaller objects with adjusted explanatory variables (De'ath & Fabricius, 2000).

With a file of this size, the presence of overlap was almost inevitable. The overlap is used to establish relations between objects which improves the structure of the model. An example of such a relation is an order which connects the objects customer and store.

2.2.2 CLUSTER ANALYSIS

Secondly, a cluster analysis is used. A clustering analysis can be seen as a family of methods that want to identify groups or clusters with distinctive characteristics from a set of heterogeneous data (Rapkin & Luke, 1993). Homogeneous groups are created out of heterogeneous data (Tufféry, 2011). These needed properties were determined before the clustering started. The clustering is mostly based on two variables. An example of this based on the data of Peddler, is be the postal code of a group of customers and the number of orders they have ordered. This created cluster is used as the base for a related KPI.

A common form of clustering is hierarchical clustering (Wedel & Shi, 2010). Clusters are formed and their hierarchy is showcased in a dendrogram. For the dendrogram, shown in Figure 1, agglomerative clustering (Hexmoor, 2015) is used. This bottom-up hierarchical cluster approach connects nodes to create clusters and increases these clusters over time. This method is very helpful when working with a dashboard.

A cluster for book delivery orders in Amsterdam that were delivered on time was created. This cluster on its own did not contain enough specific information so different postal codes were added. The merging of the postal code delivery cluster in Amsterdam with the on-time book delivery cluster, created a base cluster that was linked to a KPI.

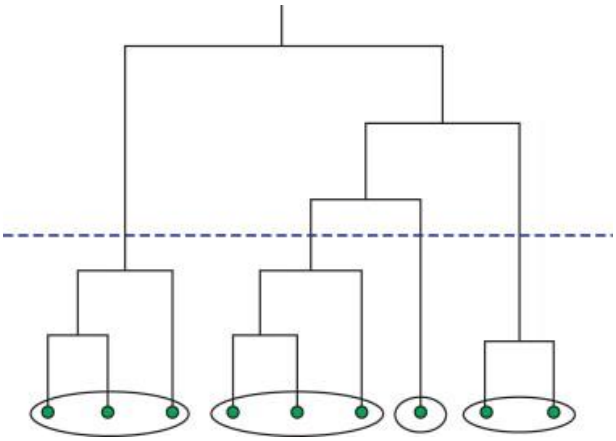


Figure 1 Dendogram as a result of hierarchical clustering (Hexmoor, 2015)

2.2.3 DIFFERENCE CLASSIFICATION AND CLUSTERING

Both methods to analyze quantitative data sound very much the same when looking at the definition of the method. However, there are some important differences in characteristics between these two methods. With classification, the intent is to label data while clustering is used to cluster groups with similar data (Lin et al., 2014). The biggest difference with this is that, for clustering, the number of clusters is not known yet. This is the case for classification.

To make things clearer, an example where both methods can help to identify a specific KPI, is made. . With classification, a customer with all relevant characteristics are determined. When clustering is applied, the number of customers whose order arrived too late and live in postal code area 1010 is calculated.

2.2.4 COHORT ANALYSIS

The last type of analysis that will be used is the cohort analysis. A cohort analysis is *''quantitative research using a measure of the concept of cohort and studying its effect on some outcome variable''* (Glenn, 2013). A cohort analysis is used for this research when looking at a KPI over time. This was used during this research as an indication how many orders were delivered on time over a year period. This can be split up into months or even weeks. This has provided a clear view of increase or decrease in performance over this given time period.

2.3 CHAPTER 2 SUMMARY

There are multiple analysis used in this research in order to find interesting KPIs for the company. Multiple methods are used to analyze quantitative and qualitative data.

- **Qualitative data:** gives insight in the management of Peddler and how the quantitative data came to be.
    - **Observational study:** explanatory and control variables are used to help create a better overview of different scenarios that influence the performance and efficiency.
    - **Interviews:** interviews with different departments identified multiple global KPIs.
- **Quantitative data:** A large Excel file containing information of historical orders that Peddler has received last year.
    - **Classification:** a method used for the segmentation of the data based on the explanatory variables of a classified object.
    - **Clustering:** a method used for the segmentation of data based on clusters with similar data.
    - **Cohorting:** a method which analyzes the segmented data over time.

3. THE ROUTE PLANNING SYSTEM OF PEDDLER

This chapter provides an insight in the organization of Peddler. Within this chapter, the platforms, tools, variables, stakeholders are brought to light to get a better understanding how Peddler manages their logistics. This information about Peddler is the result of qualitative data gathering. As mentioned in the research design above, the gathering of this data has been done through interviews and observation.

3.1 CURRENT SITUATION

First, it is important to see how the route planning of Peddler is done in practice on a daily basis. According to the CRISP-DM method, this is the first phase of this research. It all starts with the business understanding. This brings more inside how variables are determined. One of these variables are the different stakeholders of Peddler.

3.1.1 STAKEHOLDERS

Peddler is currently active in five cities in which they have multiple stakeholders. The stakeholders that are of importance for the logistics department, can be classified as the route planners, the stores that are working with them and the riders.

Riders of Peddler are the employees on electric cargo bikes who deliver the orders. The number of riders needed depends on the city. Each city has a different demand so the number of needed riders can range between five and fifty. Route planners make sure that every order is handled in a professional matter. The number of route planners that are needed are based on scheduled orders. Currently, Peddler is working together with a lot of different stores. Most stores use Peddler on a daily basis, but this is not the case for all of them. Because of this, the number of orders can differ a bit. The number of needed route planners is based on the following number of predicted orders:

- 250 orders: 1 route planner
- 400 orders: 2 route planners
- 500 orders or more: 3 route planners

3.1.2 DIFFERENCE IN ORDERS

It is also of importance to separate two types of orders. Apart from the time slot determination, there is also a difference in the type of service when the order is created. Peddler provides customers with the option to order from their platform and have it delivered at their doorstep. These orders that are ordered from the website arrive at in the admin of the website as a Product Order. However, Peddler also provides a delivery service for stores. Stores have the option to have their product range available on the website of Peddler, but it is not an obligation for a business collaboration with Peddler. It is also possible to only make use of the delivery service they provide.

Normally, in a world without Peddler, a company makes use of PostNL or another delivery service to have their products delivered to the customer. In this world however, packages arrive to the customer within two to three business days. It has to be delivered to a pickup point which is chosen by the delivery company, has to be processed to see to which area it needs to be delivered and has to be scheduled to the right deliverer. Peddler provides the option to companies to have the products be picked up at an agreed timeslot and have them delivered on the same day if it is requested. These orders arrive in the admin of the website as Delivery Orders. In other words, these are requests by companies to have their products delivered to their customers. For these orders, there is also the option to have it delivered in a selected time slot or on the same day. The system will automatically provide a time slot. However, based on agreements with the store, the route planner can decide

whether this can be transformed into a same day delivery in order to optimize their route planning. For example, deliveries from bookstores will be processed as same day deliveries because books don't have a high priority to be delivered within a certain timeslot.

3.1.3 TIME SLOTS

Before the rider gets to see the route that needs to be followed, several steps preceded it. The planning and creation of the route is done semi-automatically and consists out of multiple actions.

The creation of the route all starts with the customer. When a customer is selecting products and adding them to their shopping cart, it is recognized and stored by the admin of the website. As soon as the customer has completed their payment, they get the opportunity to select a time slot in which they prefer the order to be delivered. There are four possible choices for time slots that can be selected:

- 12:00 – 14:00
- 14:00 – 16:00
- 16:00 – 18:00
- 18:00 – 20:00

There is also a delivery timeslot from 08:00 to 12:00. However, this is not visible for the customer because it is used to deliver backorders of the previous days. It is also for products from stores that have subscription-based business model. Subscription-based orders are fixed orders, so they are already known before the day starts. To avoid an overloaded route planning system in the afternoon, these orders can be delivered in the morning.

The selection of a time slot depends on the time the payment of the order was completed. Each order has a time span of around four hours. The store where the order has been placed gets a notification when the order has been paid. From that moment, they have two hours to prepare the order to make it ready for pickup. After that, as mentioned above, the delivery time slots are two hours as well.

3.1.4 THE ADMINISTRATION OF THE WEBSITE OF PEDDLER

Created product orders arrive in the administration part of the website of Peddler. In there, the status of all these orders is visible. The status of these orders is crucial for route planners, because most created orders are irrelevant. As soon as a customer adds something to their online shopping bag, it is seen by the system as a created order with the status ''in process''. Orders with this status require no immediate actions due to multiple reasons. Most customers delete their shopping bag because they change their mind, or they wait a long time with the payment because they have not decided yet. A payment needs to be executed for an order to be processed. This order then arrives at the store where it can be prepared for the drivers. The route planners receive a notification from the store when the order is completed so it can be included in a route. In other words, route planners will only start to take action when the pickup is ready. The admin of the site is helpful for them to see the number of orders that have to be included in their route later on. It helps them to be pro-active and to act preventive to possible threats. It is also a tool to make sure that stores have their pickups ready within the appointed time slot. If there is a chance that this feasibility is threatened, route planners can call the stores to check the reason for the delay.

3.1.5 DISTRICTS

Even though Peddler is active in five large cities, it is also operates and delivers orders to their smaller villages that just outside the city. In other words, Peddler does not only operate in for example the city of Amsterdam but in the region of Amsterdam. In practice this means that riders in Amsterdam also deliver orders to customers living in Diemen. To get a better understanding in the model that is made, the following active regions with their correlated cities are cited below:

- Amsterdam
    - Amsterdam
    - Diemen
    - Zaanstad
    - Zwaanshoek
    - Amstelveen
    - Haarlemmermeer
- The Hague
    - The Hague
    - Wassenaar
    - Scheveningen
    - Delft
    - Leidschendam
    - Voorburg
    - Wateringen
    - Leiden
    - Westland
    - Badhoevedorp
    - Stompwijk
- Groningen
    - Groningen
    - Haren
- Rotterdam
    - Rotterdam
- Utrecht
    - Utrecht
    - Vleuten
    - Hof van Twente
    - De Meern

3.1.6 THE ROUTE PLANNING

When a store has prepared an order to be picked up, the route planners get a notification and can start including it in a route. Route planners of Peddler create and operate their routes in a program called X. Here, they have an overview of all orders shown on a map of the city. Orders are divided into three different categories: unassigned, assigned and completed. The name speaks for itself, but unassigned orders are the orders that still need to be included in a route. Within X, route planners are constantly updated about the location of the riders. It helps them to efficiently add orders to an already created route. This helps Peddler to prevent a huge increase in extra kilometers by having to create an entire new route. In other words, existing routes are adaptive which allows for efficient route planning.

The route planning itself is semi-automatic. The calculation of the shortest distance when adding a new location is done automatically. However, the adding of the order is done manually. The reason for this is also that orders with a higher priority can be added first. X is not capable of ranking orders based on criteria so for now this is done by the route planners. An example of orders with a lower priority are book orders. Books can be delivered any time of the day so the urgency to add the pickup in the route is lower than other type of deliveries. As with most manual things, mistakes are more likely to occur during this phase of the process than during the automatic phases.

After the route is created or modified, the rider needs to be notified about these changes. The route planners are constantly in contact with the riders through WhatsApp. Here, the delivers notify them about potential changes, broken attributes or other potential reasons why the feasibility of the current route could be threatened. Cargo bikes are attached to the riders of which the location is shown on a map. This allows for route planners to stay aware of riders and provides the opportunity to let them know if something is not according to plan.

3.2 CHAPTER 3 SUMMARY

For this research, relevant stakeholders for Peddler are their route planners, the stores that they are working with and their riders. These stakeholders are all important factors for the processing of orders. Peddler works with two types of orders.

- **Delivery orders:** these are orders that have been booked in by stores. These stores use Peddler as a delivery service.
- **Platform orders:** these are orders of customers that are ordered from the platform of Peddler.

Peddler works with timeslots in which these orders are delivered. An example of such a timeslot is 12:00-14:00. Customers have the option to select a timeslot that fits their schedule. Peddler is operative in five regions which each contain multiple cities.

4. DATA ANALYSIS

In chapter four, the selected methods of chapter two are put into practice to determine the needed KPIs. In other words, the chapter contains an analysis of both the qualitative and quantitative data. This results in a list of multiple helpful KPIs that are used as input for the dashboards.

4.1 QUALITATIVE DATA ANALYSIS

During phase two of the CRISP-DM research method, the data understanding phase, the gathered data needs to be analyzed. The transformation of qualitative data to quantitative already gives some good clarification for the input and purpose of the dashboard. As with the business understanding phase, this data has been gathered through observation and interviews. With the use of the NUDIST method, the following points of interest were derived from the answers of the interviewees.

4.1.1 NORMS AND VALUES

One of the norms and values of Peddler is the quality of their delivery service. “Peddler will always go one step further when comparing it to other delivery instances’’ (Logistics, 2022). These norms and values can partly be translated to KPIs which enable the measurement of quality. Interviewing riders to get their opinion how they are treated is not something that is relevant for this research. However, the quality of the delivery is. The previously mentioned importance of quality mostly derives from the time the order was delivered. In order words, if the delivery was within the agreed timeslot. The feasibility of the delivery time indication is determined from the quantity and the distance. These two can be seen as important variables for the following determined KPIs:

Overall:
Overall Percentage of orders delivered on time
Percentage of orders delivered on time per city
Percentage of orders delivered on time per postal code
Percentage of orders delivered on time per area
Overall Percentage of orders delivered on time in proportion with the total orders
Percentage of orders delivered on time per timeslot

Platform orders:
Percentage of orders delivered on time per store

Delivery orders:
Percentage of orders delivered on time per delivery package (Alcohol, goods, etc.)

4.1.2 SHORTCOMINGS

It is impossible for a manual procedure to be perfect so, during the creation and support of a route, certain shortcomings are present in the current route planning process of Peddler. These shortcomings can be the result of manual mistakes or flaws in the tool peddler is currently using. The flaws in the tool are not relevant for this research because the advice that is formulated based on the created dashboard will not solve these shortcomings.

Other identified shortcomings are the optimization of the space of the container, unfortunate evens during the route and the skills and communication of riders. The space of the container must be optimized for deliveries.

Each delivery has a maximum in terms of dimensions, but it is unclear how big the package will actually be. This can be solved by adding multiple scales of measurements in which the package could fit.

During the delivery, multiple shortcomings can occur which have an impact on the efficiency. New instructions from the route planner to the rider can be incorrect, something could happen with the cargo bike during the delivery, or a store could be too late with their pickup. The first two, again, are not directly solvable with the advice from the dashboard. The last one however, can be tackled with the use of the dashboard. Stores which consistently fail to have the pickup ready in the agreed timeslot can be held accountable for their actions.

Riders also have a big influence on the quality of the delivery. Their delivery speed, knowledge of the city, flexibility, communication skills and efficiency are all important factors that contribute to Peddler meeting their promised delivery timeslot. As with the same as other businesses, some employees perform better than others. Riders are currently scored by the number of stops they had per hour. Even though Peddler should not start changing their norms and values, this method of reviewing riders is still too broad. Riders can be pointed out on their actions by looking at their arrival at the pickup and the final delivery address. It could be that a rider achieves their requested amount of stops per hour, but their orders are consistently delivered too late.

Overall:
Number of failed pickups
Percentage of orders delivered on time per rider
Number of on time pickups per rider

4.1.3 INPUT FOR THE COHORT ANALYSIS

Measurement of performance is only useful for a company if the measuring is done over a time period. If the outcomes of certain KPIs are not consistently positively over time, they cannot be seen as selling points. A peak in on time delivered orders by one rider could have other reasons than their own capabilities. These capabilities of this rider can be proven if this peak is consistently achieved over time. Looking at the performance and efficiency over a time period serves as input for the cohort analysis.

During this analysis, a starting point can be selected and various KPIs can be tracked over a determined time period. For example, an analysis about the number of orders that were delivered on time starting from January. This KPI can later be used to see the performance over time for one store or one ride for example. Useful KPIs that will be helpful for the cohort analysis are the following:

Overall:
Percentage of orders delivered on time per week
Percentage of orders delivered on time per month

4.1.4 DERIVING NEW KPIS

KPIs derived from the qualitative data provided by interviews and observational study serve as starting points for the processing of the quantitative data. Variables such as store, rider, postal code, city and timeslot help to make clear what data is useful what data is not. These determined KPIs are not fixed and instead can be combined with other KPIs to form new KPIs. The goal is that, within the dashboard, the user is allowed to narrow down their search. This can result the following KPI:

Overall:
Percentage of orders delivered on time per city per store per rider per month

4.2 QUANTITATIVE DATA ANALYSIS

The transformation from qualitative data to KPIs functions as helpful starting points for the quantitative data. Like with the qualitative data, within the data understanding phase, is suited to get more acquainted with the data the research is working with. Within this phase, multiple analysis are executed to get a better understanding. A classification and cluster analysis is used to give more structure to the data. Analyzing and processing the data using these methods, help to correctly prepare the data to serve as input for the model. This is all part of the third phase named data preparation.

4.2.1 HISTORICAL ORDERS

The provided quantitative data consists out of historical orders extracted from the administration of their database. Provided datasets are equal to the orders within a month. To clarify, the foundation of the dashboard consists in this case out of multiple Excel files. Each of these files contains the data of all historical order for a given month. From the creation of the shopping card until the delivery of the order, data is created and stored. During this process, new variables are created depending on the position of the order within the process. For instance, when a customer is filling in their name and postal code, variables with the equivalent names are created and stored. This procedure occurs every time a new action is taken. This results in a total of 161 different variables for just the month May that are created for just one order. This number of different variables makes the cleansing of the dataset absolutely necessary.

The data is structured with the use of rows and columns. Each row in the Excel file contains information about an order. All the information of this order is spread out over 161 columns. Each of these columns contain a variable which is either filled in manually or automatically.

A lot of data within the Excel files is information about the order that needed to be filled in by the customer or the rider. This is all done manually which means it vulnerable for mistakes and inconsistency. Information about the delivery address and contact details of the customer is filled in by the customer. Every new customer needs to do this separately. A result of this manual work from a lot of different people, is that the data is not consistent. The best example of this is the variables house number, apartment number and address. The variables speak for themselves, but a lot of customers mix them or combine them. The same goes for the riders. A rider needs to select when an order is picked up, when it is back on route after the pickup and when the order is delivered. Some riders tend to forget this which leads to impossible values between the pickup and delivery time. This inconvenience will be taken into account in the discussion.

Looking at the determined KPIs, important variables are identified which serve as a guideline for the cleansing of the provided dataset. These variables, in combination with the classification and cluster analysis, help to give the data more structure. These steps, combined with the cleansing of the data, take place in the data preparation phase and form the foundation for the dashboard.

4.2.2 CLASSIFICATION ANALYSIS

In order to get a better understanding of the data this research is dealing with; a classification analysis is used. Classification helps to segment the data in smaller groups. This analysis focusses on classified objects with certain characteristics that support it and determine how it is classified (Tufféry, 2011). These characteristics can be seen as explanatory variables (De'ath & Fabricius, 2000) which help to explain the foundation of an object. Relations between these objects make for a clear and cohesive model which will form the base for the dashboard.

Looking at the KPIs, it becomes clear which objects and variables will be necessary. Some of these variables can be seen as explanatory variables of determined objects. An example of an object is a store because it can be divided into various explanatory variables, or characteristics, which are unique for each store. Suited

explanatory variables are the products they sell, their address and postal code and name. Within the determined KPIs, relevant variables and object are underlined.

Overall:
Percentage of orders delivered on time per city
Percentage of orders delivered on time per postal code (per district)
Percentage of orders delivered on time per timeslot
Percentage of orders delivered on time per rider
Percentage of orders delivered on time per week
Percentage of orders delivered on time per month
Percentage of orders delivered on time per city per store per rider per month

Platform orders:
Percentage of orders delivered on time per store

Delivery orders:
Percentage of orders delivered on time per delivery package (Alcohol, goods, etc.)

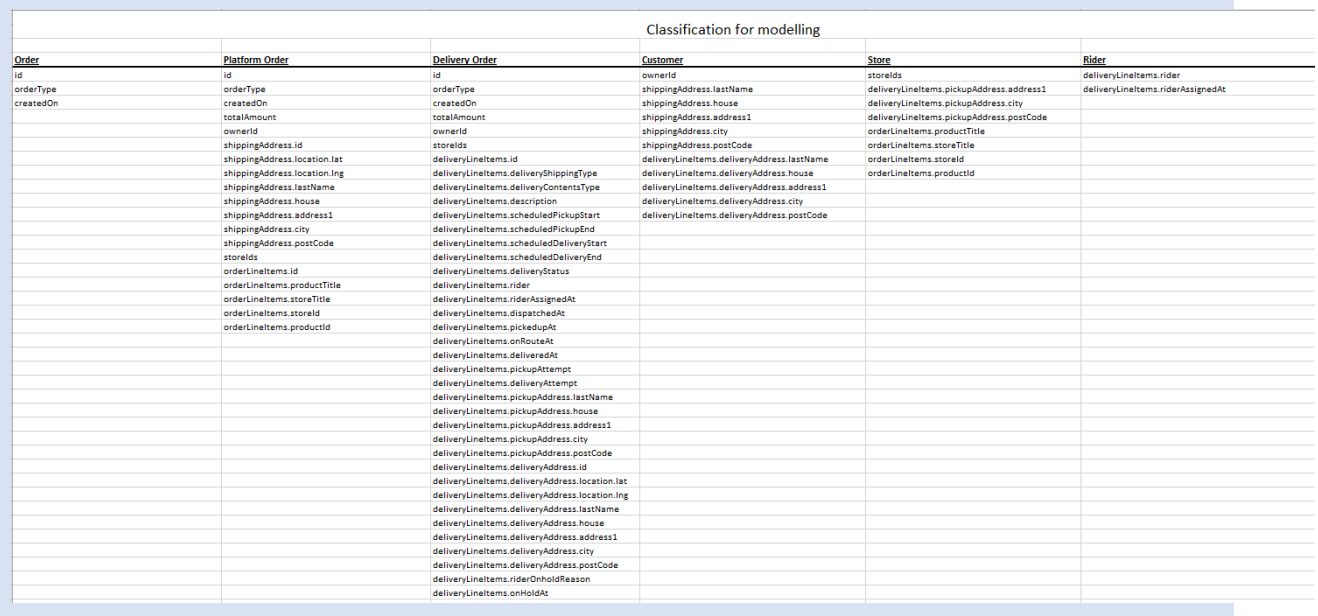
CLASSES

With the use of the classification analysis, the data file containing 161 (for only May) different variables could be cleansed. Different objects and sub objects combined with their explanatory variables could be determined. These objects are shown in Figure 2. The objects are shown above the line with their suited characteristics below them. Objects and sub objects derived from the dataset are:

- Order
    - Platform Order
    - Delivery Order
- Customer
- Store
- Rider

In this case, order is the main object consisting of just three different variables. Due to the difference in characteristics, sub objects platform orders and delivery orders have been labeled as separate sub objects.

The total number of explanatory variables is 81. The reason for this not being 161 is that a lot of variables could be removed. These variables don't contribute to the calculation of the KPIs which makes them irrelevant to add as input for the model.



Classification for modelling

Order	Platform Order	Delivery Order	Customer	Store	Rider
id	id	id	ownerId	storeIds	deliveryLineItems.rider
orderType	orderType	orderType	shippingAddress.lastName	deliveryLineItems.pickupAddress.address1	deliveryLineItems.riderAssignedAt
createdOn	createdOn	createdOn	shippingAddress.house	deliveryLineItems.pickupAddress.city	
	totalAmount	totalAmount	shippingAddress.address1	deliveryLineItems.pickupAddress.postCode	
	ownerId	ownerId	shippingAddress.city	orderLineItems.productTitle	
	shippingAddress.id	storeIds	shippingAddress.postCode	orderLineItems.storeTitle	
	shippingAddress.location.lat	deliveryLineItems.id	deliveryLineItems.deliveryAddress.lastName	orderLineItems.storeId	
	shippingAddress.location.lng	deliveryLineItems.deliveryShippingType	deliveryLineItems.deliveryAddress.house	orderLineItems.productId	
	shippingAddress.lastName	deliveryLineItems.deliveryContentsType	deliveryLineItems.deliveryAddress.address1		
	shippingAddress.house	deliveryLineItems.description	deliveryLineItems.deliveryAddress.city		
	shippingAddress.address1	deliveryLineItems.scheduledPickupStart	deliveryLineItems.deliveryAddress.postCode		
	shippingAddress.city	deliveryLineItems.scheduledPickupEnd			
	shippingAddress.postCode	deliveryLineItems.scheduledDeliveryStart			
	storeIds	deliveryLineItems.scheduledDeliveryEnd			
	orderLineItems.id	deliveryLineItems.deliveryStatus			
	orderLineItems.productTitle	deliveryLineItems.rider			
	orderLineItems.storeTitle	deliveryLineItems.riderAssignedAt			
	orderLineItems.storeId	deliveryLineItems.dispatchedAt			
	orderLineItems.productId	deliveryLineItems.pickedupAt			
		deliveryLineItems.onRouteAt			
		deliveryLineItems.deliveredAt			
		deliveryLineItems.pickupAttempt			
		deliveryLineItems.deliveryAttempt			
		deliveryLineItems.pickupAddress.lastName			
		deliveryLineItems.pickupAddress.house			
		deliveryLineItems.pickupAddress.address1			
		deliveryLineItems.pickupAddress.city			
		deliveryLineItems.pickupAddress.postCode			
		deliveryLineItems.deliveryAddress.id			
		deliveryLineItems.deliveryAddress.location.lat			
		deliveryLineItems.deliveryAddress.location.lng			
		deliveryLineItems.deliveryAddress.lastName			
		deliveryLineItems.deliveryAddress.house			
		deliveryLineItems.deliveryAddress.address1			
		deliveryLineItems.deliveryAddress.city			
		deliveryLineItems.deliveryAddress.postCode			
		deliveryLineItems.riderOnholdReason			
		deliveryLineItems.onHoldAt			

Figure 2 Classification

4.2.3 CLUSTER ANALYSIS

In order to be able to give advice based on the findings in the dashboard, a cluster analysis is needed. The goal of this analysis is to identify groups with distinctive characteristics (Rapkin & Luke, 1993) which have certain matching properties. With the created objects and their corresponding variables, clusters can be made to start analyzing and processing the formulated KPIs. Adding criteria to these clusters can help to narrow down the search to which help to obtain the requested results. Based on these results, advice can be formulated.

Looking at the KPIs which were derived from the qualitative data, clusters can already be identified. The first KPI shown below is still very broad but consists out of a cluster: orders that were delivered on time. The cluster orders with pickups that failed are the foundation for the second KPI. A KPI can also give constraints that help to create a cluster. This is the case for the third KPI. Here the KPI serves as a guideline for the dashboard to gather results. An example of a cluster for this KPI could be orders from store A in city B that were delivered by rider C in month D.

Overall:
Overall Percentage of orders delivered on time
Number of failed pickups
Percentage of orders delivered on time per city per store per rider per month

4.2.4 COHORT ANALYSIS

Clusters created during the during the cluster analysis need to be monitored over a certain period of time in order to carry out a cohort analysis. The cohort analysis serves to see the change in results of a desired KPI within such a time period. Such a cohort can also be seen as a cluster. An example of a cohort which is also a cluster is: all orders that were delivered on time by rider A during month B but shown per day. Within this cohort, the performance based on orders delivered on time of rider A is shown during month B. This could be plotted in a line graph to see if rider A has improved during month B.

4.3 CHAPER 4 SUMMARY

With the use of interview and observational study, the norms and values and the shortcomings in the logistics department of Peddlers can be transformed into a list of KPIs. This list of KPIs was used as a guideline to analyze the quantitative data. Classification, a cohort and cluster analysis has helped to prepare the data. This meant cleansing the Excel file and segmenting the data. The following KPIs are the foundation for the created dashboards.

Overall:
Overall Percentage of orders delivered on time
Percentage of orders delivered on time per city
Percentage of orders delivered on time per postal code
Percentage of orders delivered on time per area
Overall Percentage of orders delivered on time in proportion with the total orders
Percentage of orders delivered on time per timeslot
Number of failed pickups
Percentage of orders delivered on time per rider
Number of on time pickups per rider
Percentage of orders delivered on time per week
Percentage of orders delivered on time per month

Platform orders:
Percentage of orders delivered on time per store

Delivery orders:
Percentage of orders delivered on time per delivery package (Alcohol, goods, etc.)

5. MODELING

Chapter 5 is about phase 3 and 4 of CRISP-DM: the data preparation and the modelling. In practice, this means that the Excel file has to be prepared to serve as input for the dashboards. In addition, the tool in which these dashboards are build is constructively chosen. The modelling phase has been executed and the created dashboards are shown and explained.

5.1 DATA PREPARATION

With the help of the quantitative data analysis, the data has been rightfully prepared to be imported in the data visualizing tool used for this research. During the data preparation, it became clear that Peddler is still in its start-up phase. In practice, this meant that the company does not have a fixed amount of variables that they want to measure. In other words, every month has a new set of variables that measure something new.

The preparation of all the Excel files was done semi-automatically. Due to the different number of variables and the changing location of these variables, the files needed to be cleansed manually. After manually moving columns to their destined place in the script, several created macros were triggered resulting in automatic calculations and formulation. In other words, only the cleansing was done manually. Inserting the script in a VBA macro workbook makes sure Excel automatically cleanses, improves and merges the files in order to form a solid foundation for the tool to work with. One of the most important additions for the data is the introduction of a new variable named “Appointed timeslot”. This variable enables the calculation of the KPI *Percentage of orders delivered on time per timeslot.* Based on the already existing variable “Scheduled End”, a timeslot is assigned to an order due to the code in Figure 3.

```
Range("H2").Select

Range("H2").Value = "08:00:00 - 12:00:00"
Range("H3").Value = "12:00:01 - 14:00:00"
Range("H4").Value = "14:00:01 - 16:00:00"
Range("H5").Value = "16:00:01 - 18:00:00"
Range("H6").Value = "18:00:01 - 20:00:00"


Range("I2").Select
ActiveCell.Formula2R1C1 = _
    "=INDEX(R2C8:R6C8,MATCH(MOD(RC[-2],1),LEFT(R2C8:R6C8,FIND(""-"",R2C8:R6C8)-1)+0))"
Range("I2").Select
Selection.AutoFill Destination:=Range("I2:I30000")
```

Figure 3 Code for the timeslot variable

After all these steps were completed, one final step in the data preparation phase needed to be carried out before the files were ready to be visualized. Due to manually filled in data, certain cities were written differently but needed to be seen as the same. Qlik Sense allows the user to group values and give them an umbrella name.

5.2 TOOL SELECTION

The selected tool for this research in which the data is visualized, is Qlik Sense. The biggest motivation behind this choice is that Peddler is already using this program is their main tool to visualize performance. The company uses it to get a better understanding of their performance in other departments than Logistics. In other words, the biggest reason for the selection of this tool is that employees are already familiar with it. In addition, for those who are not familiar with Qlik Sense or Business Intelligence dashboards in general, it is very accessible user-friendly for new users.

Therefore, Qlik sense serves as the ideal tool to fill the gap between norm and reality for Peddler. It will help to solve the core problem which enables Peddler to get a better understanding of their performance and efficiency.

5.3 DATA VISUALIZATION

In order to sufficiently solve the core problem and make sure the dashboards are able to visualize and answer the needed KPIs, five different dashboards were created. All these dashboards are connected with each other and have common dimension filters. The addition of these filters allow the user to specify the data base by narrowing their search field. The filters that allow the user to make their search more specific are the following:

- Order type: allows to only show data for delivery or platform orders;
- Month: the user is able to select a month from June 2021 to June 2022;
- Week number: allows to specify the search to a certain week ranging between one and fifty-two;
- Appointed timeslots: allows to only select date for one of the four timeslots, or orders that were done in the morning from 08:00 to 12:00;
- Rider: the name of anyone who has delivered an order for Peddler can be selected;
- Delivery city: a desired city in which Peddler operates can be selected;
- Pickup attempts: allows the user to see what a failed pickup does with the data;
- Delivery content: lets the user specify the content of the order (alcohol, food etc).

The visualized performance is made structurally by dividing and implementing the desired KPIs over five separate dashboards. Again, these dashboards are all connected which each other. In practice, this means that the filtered dimensions are automatically applied to all dashboards when chosen. It allows the user to be provided with a clear overview of all KPIs connected by the selected filters.

5.3.1 DASHBOARD 1: LOCATION PERFORMANCE

The purpose of the first dashboard, location performance (Figure 4), is to visualize the KPIs which, as the name implies, focus on geographical performance. The geographical performance of each of the different locations in which Peddler operates in shown. This dashboard focuses on four measures: the percentage of orders that has been delivered on time, the number of orders that was delivered on time, the number of orders that was delivered too late and total number of orders. This has been done with the use of two maps, multiple bar charts, a pivot table and a KPI shown two measures. The maps help to visualize the scope of the selected area such as shown in Appendix A3 where Amsterdam has been selected as city. To let the user get a better and faster understanding of the performance of each city and postal codes, colors are used. Each color indicates a range bounded by limits.

The four bar graphs to the right allow the user to compare different cities and postal codes based on the four measures. The pivot table provides the option to have a total overview of the performance of each city. Last, the KPI to the gives the user the total value from the selected dimensions of both measures.

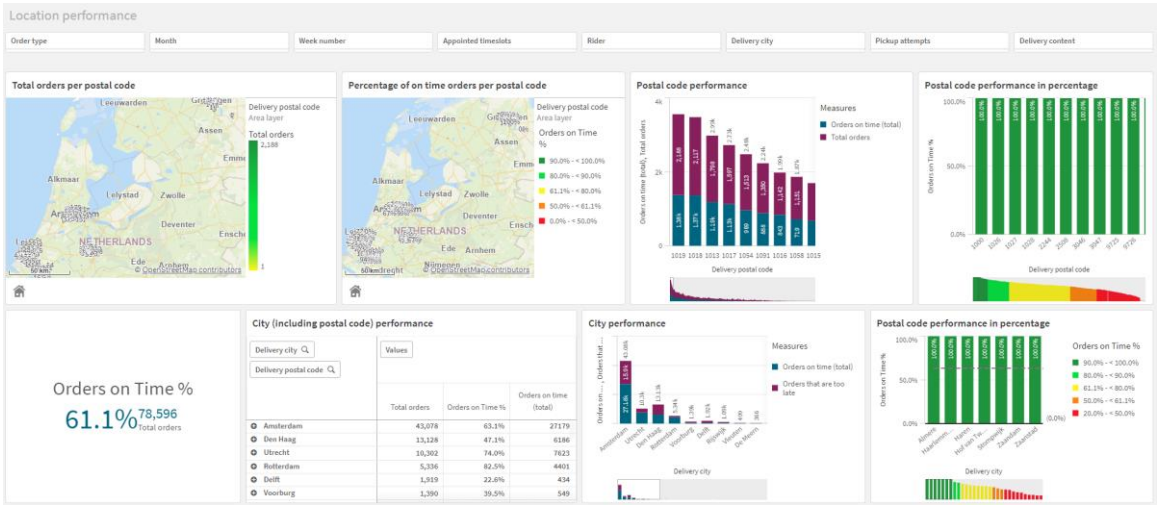


Figure 4 Dashboard 1: Location performance

All graphs shown in this dashboard are interactive. This means that zooming into a process can be carried out from each of these graphs. For example with the map, when a desired area in the map is selected, the entire dashboard adjusts its data in order to have it aligned with the required search field.

5.3.2 DASHBOARD 2: PERFORMANCE OVER TIME (INCLUDING TIMESLOT PERFORMANCE)

In order to be able to successfully carry out a cohort analysis, the option to measure dimensions over time needs to be present in the visualization (Figure 5). Line charts and bar charts fulfil these requirements sufficiently and help to visualize the performance of timeslots and orders over time. This overview takes place on the left side of the dashboard. The right is focused on the performance of the different timeslots. A pie chart (in the middle) shows the distribution of on time orders per timeslot. This allows the user to see how many orders per time slot have been delivered on time and what percentage of this is of the total on time orders. The bar charts make it possible to see the performance per month, week, and timeslot. The map makes it easier for the user to only select certain postal codes and see their performance over time. It also provides the opportunity to see how each timeslot performs per postal code. This will go more in depth in the third dashboard.

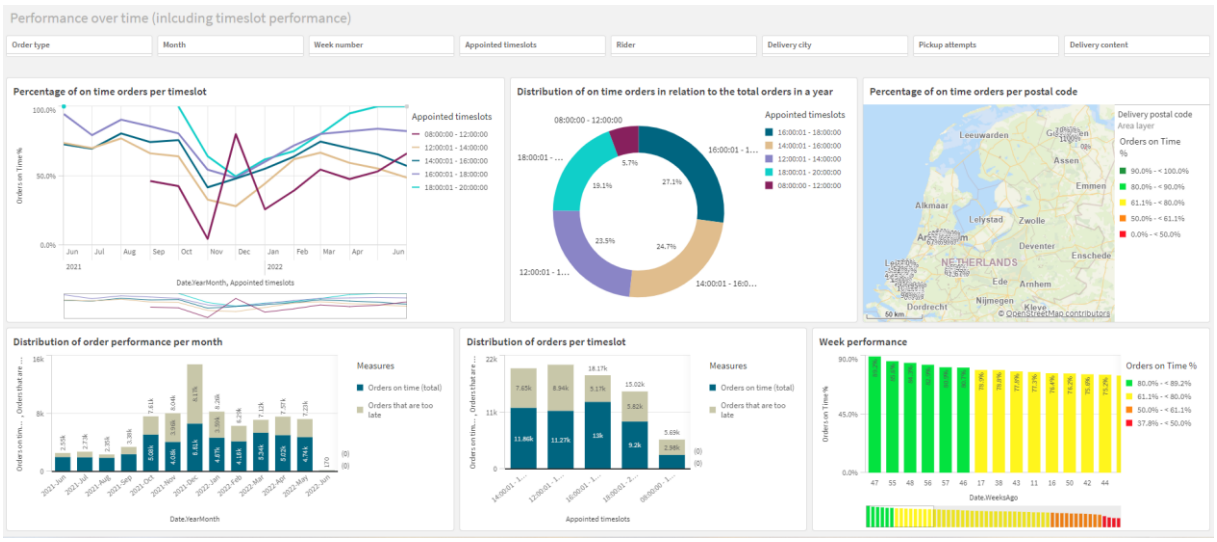


Figure 5 Dashboard 2: Performance over time

5.3.3 DASHBOARD 3: TIMESLOT PERFORMANCE PER POSTAL CODE

The third dashboard consists, apart from the common filters, only of one pivot table (Figure 6). The reason for the existence of this dashboard is that it functions as the foundation for the cluster KPI ''*Percentage of orders delivered on time per postal code per timeslot''*. This KPI allows Peddler to get a better understanding of the overall performance of each timeslot in a postal code. Due to the geographical location of certain postal codes, it is possible that on time deliveries within certain timeslots aren't feasible. This dashboard, based on the historical orders of an entire year, shows the overall performance of each timeslot within these postal codes. Due to the small number of orders in some postal codes, the total orders within the postal codes for the given timeslot is added to get a better indication if the performance is sufficiently substantiated.

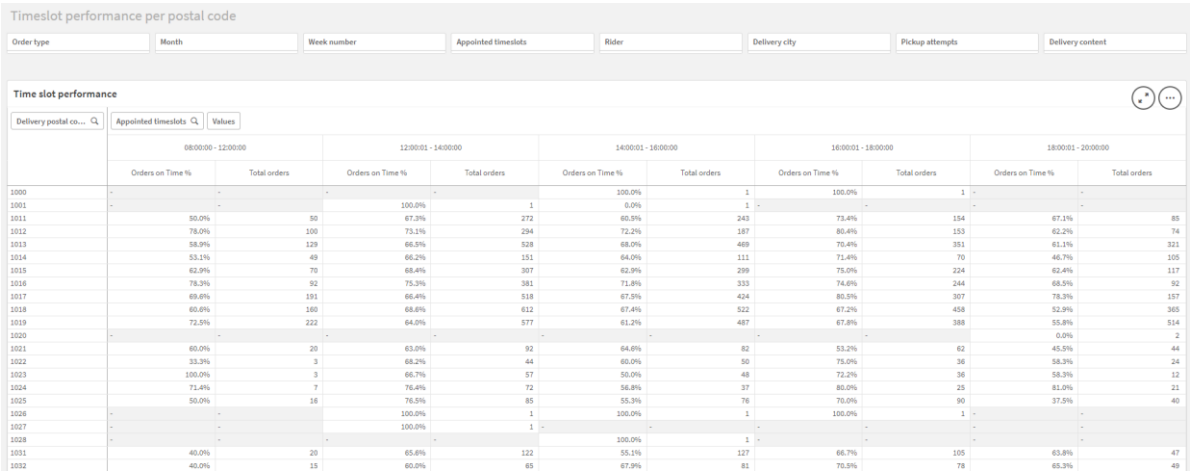


Figure 6 Dashboard 3: Timeslot performance per postal code

5.3.4 DASHBOARD 4: STORE PERFORMANCE

The fourth dashboard is focused on KPIs which have to do with the performance of the stores Peddler is collaborating with (Figure 7). This has based on the KPI: *Percentage of orders delivered on time per store*. It provides an overview of all the stores and their overall performance in the pivot table on the right. This performance of each store is plotted over time in the line chart in the middle. The large pivot table to left and the scatter plot in the middle function as a small market research for Peddler. Here, it shows in which postal codes have the largest share in the total orders for each store. This share is further showcased in the scatter plot. This indication of the share in orders each postal code has, allows Peddler to have more customer knowledge. Due to privacy reasons, the names of the stores have been covered (Figure 7).

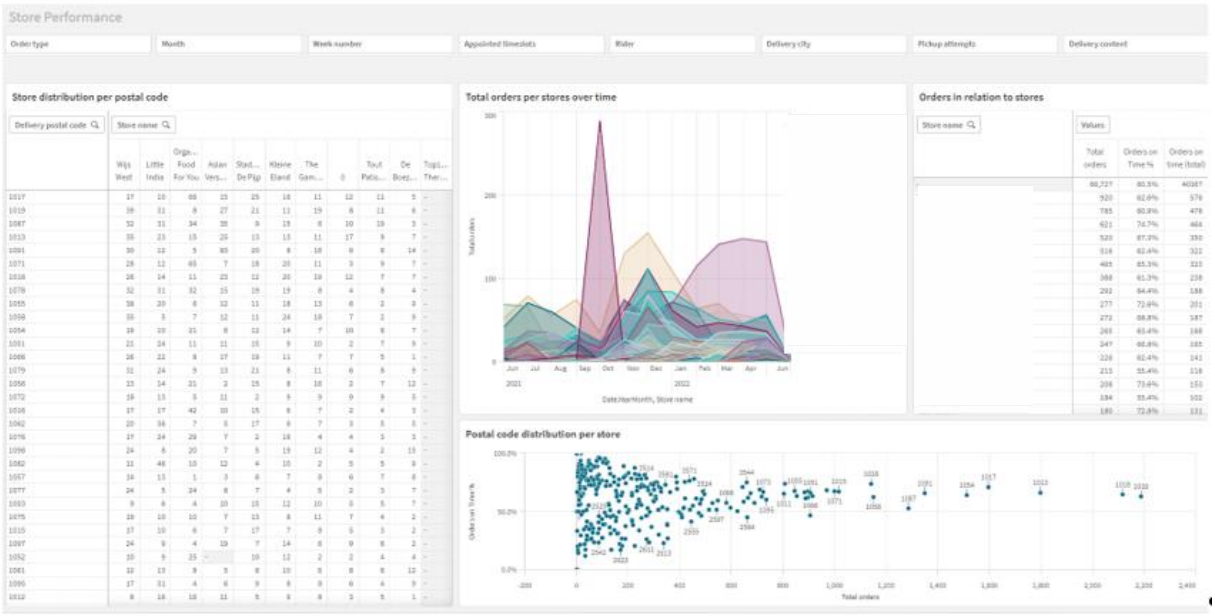


Figure 7 Dashboard 4: Store performance

5.3.5 DASHBOARD 5: RIDER PERFORMANCE

The fifth and final dashboard visualizes the performance of all the riders that have been employed by Peddler (Figure 8). Riders are an important factor that play a role in the number of orders that are delivered on time. The fifth dashboard provides more in depth and objective insights in the performance of each rider. It consists of a pivot table that shows every riders that has delivered at least one order in the last year. This table shows a ranking based on the total number of orders delivered. The scatterplot below shows the performance of riders based on the total number of delivered orders and the percentage of which orders were delivered on time. Here, it becomes clear which riders perform better than others. To make it fair, the map allows the user to see the performance of each rider in the same location. A pie chart compares the selected riders and shows each rider's share in the total on time delivered orders. Their performance over time is plotted with the use of a line graph. Due to privacy reasons, the names of the riders have been covered (Figure 8).

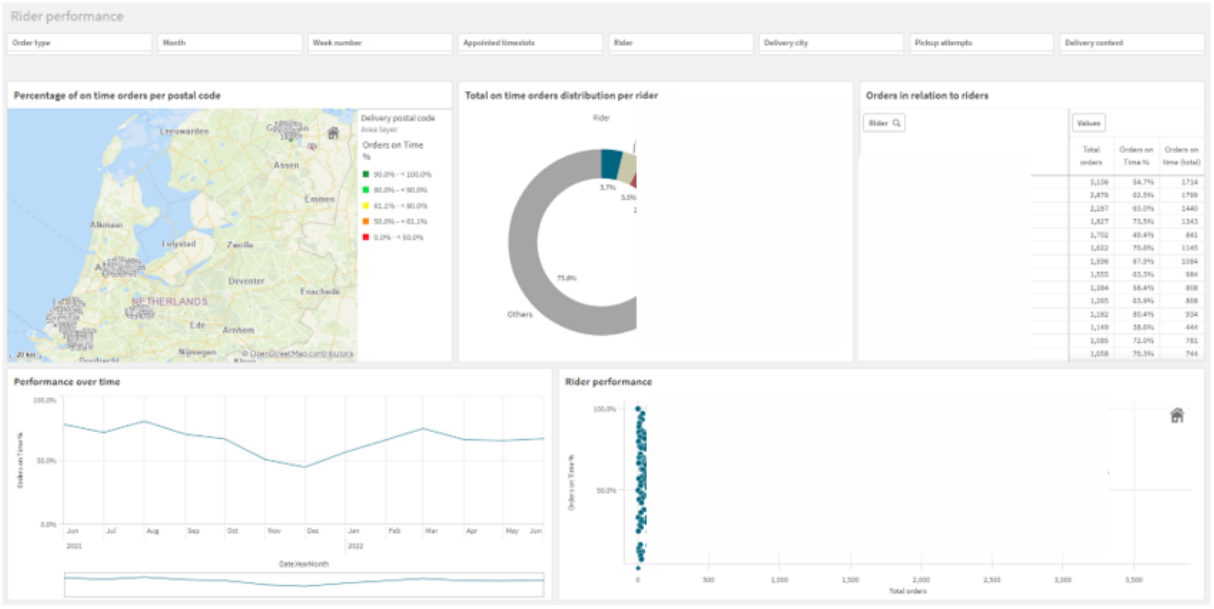


Figure 8 Dashboard 5: Rider performance

5.4 DATA VALIDATION

Because the created dashboard serve as a foundation for possible adjustments in strategy or create an explanation for some consequences, it is of high importance that the data is validated. In order to do this properly, white-box and black-box validations are used (Robinson, 2014). This way of validating looks at the visualization in two different ways, the small parts individually and the final visualization that comes from these smaller parts. White-box validation looks at the calculations of the smaller data that serve as input for the dashboard and Black-box looks at the final products and determines if the results are sufficient enough.

Because almost all calculations of the dashboard were done in the Excel files themselves, the White-box validation becomes more manageable. In addition, the formulas used to calculate the needed data are not very complex. This Excel file has been shown to the employees of Peddler to see if the results are realistic. After the dashboards were created, they were shown to Peddler for the Black-box validation. Representatives of almost all departments were present to see if the created dashboards were realistic and accurate according to their knowledge and data.

5.5 CHAPTER 5 SUMMARY

During the data preparation and modeling phase, the Excel file was cleansed, the right calculations were made with the use of macros, the tool was selected and the dashboards were created. The tool for the visualization of the data is Qlik Sense. The reason for this is that Peddler is already using the tool for other purposes so employees are already familiar with it.

In total, there are five dashboards that have been created with each their own purpose. All five dashboards are interactive and connected with each other. They have multiple filter options with variables such as the month and the delivery city. This helps the user to narrow down their search field. The five created dashboards are:

- **Location performance:** focuses on geographical performance.
- **Performance over time (including timeslot performance):** focuses on the cohort analysis.
- **Timeslot performance per postal code:** provides an overview of how each timeslots performs per postal code in terms of on-time deliveries.
- **Store performance:** focuses on the performance of stores peddler is cooperating with.
- **Rider performance:** focuses on the performance of the riders Peddler has or had employed.

In addition, the dashboard has been validated using the **White-box** and **Black-box** validation.

6. RESEARCH RESULTS

When no filters are selected, all dashboards show the performance of all regions, riders, orders and stores for an entire year. This year takes place from June 2021 to the beginning of June 2022. This results in a total of 78,000 orders with an average of 61.1% orders that have been delivered on time. To determine how these two numbers came to be and what factors were most important, different analyses have been carried out from which the following results were concluded. This chapter focuses on the different results that were found after the analysis were processed.

6.1 LOCATION BASED RESULTS

As described in Section 5.2, the first created dashboard named location performance focuses on the KPIs that use location driven dimensions. This dashboard functions as the substantiation for the results for the following KPIs:
Percentage of orders delivered on time per city
Percentage of orders delivered on time per postal code
Percentage of orders delivered on time per area

6.1.1 CITIES

Starting with the dimension city, the created pivot table can be exported to Excel resulting in the table shown in (Figure 9) . Here, each city is ranked based on the total orders that have been delivered. Additional relevant columns containing measures help to give an idea how the city has been performing compared with other operating cities. The values shown in the table are based on all available data which means all orders from last year. Colors help to indicate great and weak performing cities when it comes to orders that have been delivered on time. Apart from the five big cities in which Peddler operates, their smaller surrounding suburbs are present as well. Due to the small number of orders in these areas, the ranking has been based on the total number of orders and not the percentage of on time orders. The following points of interest were found:

- Amsterdam's share in total orders is more than 50%. It has been Peddlers operating city since they have started the operation but, as shown in Figure 9, it still is the most important and bestselling place they are operating in.
- City The Hague is performing poorly with an average of 47.1% of order being delivered on time. The reason for this is that the hub of The Hague is located quite poorly leading to a lot of insufficient starting points for the created routes. The Hague is currently Peddler's second biggest operating city so this should be a point of interest.
- Comparing The Hague with the third biggest city, Utrecht, it should even more be a point of interest. Utrecht has an average of 74% of orders being delivered on time. Utrecht has, with 3000 orders less in total, around more 600 orders that have been delivered on time than The Hague. This results in Utrecht's share in total on time delivered orders is 3% more than that of The Hague.
- The best performing city with more than 100 orders in total is Rotterdam. It has been able to deliver 82.5% of orders. Where the hub of The Hague is located quite poorly, the hub of Rotterdam is located quite nice around the center of the city. Even more impressive considering that it has 6.79% share of the total orders and a 9.16% share of orders that have been delivered on time (Appendix A5).
- The worst performing city is Delft, having only 22.6% of their orders delivered on time. This mostly has to do with geographical reasons. Delft's location is just between The Hague and Rotterdam making at a distant travel location for both. It is in favor for Rotterdam that this has appointed to region The Hague (Appendix A5).

Delivery city	Total orders	Orders on Time %	Orders on time (total)	Share of total orders	Share of on time orders
Total	78.587	61,1%	48041	100%	100%
Amsterdam	43.078	63,1%	27179	54,82%	56,57%
Den Haag	13.128	47,1%	6186	16,71%	12,88%
Utrecht	10.302	74,0%	7623	13,11%	15,87%
Rotterdam	5.336	82,5%	4401	6,79%	9,16%
Delft	1.919	22,6%	434	2,44%	0,90%
Voorburg	1.390	39,5%	549	1,77%	1,14%
Rijswijk	1.093	33,9%	371	1,39%	0,77%
Leidschendam	598	35,8%	214	0,76%	0,45%
Vleuten	499	65,7%	328	0,63%	0,68%
De Meern	366	66,4%	243	0,47%	0,51%
Groningen	309	77,0%	238	0,39%	0,50%
Wateringen	219	23,3%	51	0,28%	0,11%
Wassenaar	106	68,9%	73	0,13%	0,15%
Diemen	82	48,8%	40	0,10%	0,08%
Filled in incorrectly	55	65,5%	36	0,07%	0,07%
Haren	40	100,0%	40	0,05%	0,08%
Voorschoten	15	26,7%	4	0,02%	0,01%
Scheveningen	11	54,5%	6	0,01%	0,01%
Amstelveen	10	70,0%	7	0,01%	0,01%
Duivendrecht	7	85,7%	6	0,01%	0,01%
Leiden	5	20,0%	1	0,01%	0,00%
Valkenburg Zh	5	20,0%	1	0,01%	0,00%
Westland	5	60,0%	3	0,01%	0,01%
Badhoevedorp	2	50,0%	1	0,00%	0,00%
Almere	1	100,0%	1	0,00%	0,00%
Haarlemmermeer	1	100,0%	1	0,00%	0,00%
Hof van Twente	1	100,0%	1	0,00%	0,00%
Stompwijk	1	100,0%	1	0,00%	0,00%
Zaandam	1	100,0%	1	0,00%	0,00%
Zaanstad	1	100,0%	1	0,00%	0,00%
Zwaanshoek	1	0,0%		0,00%	0,00%

Figure 9 City performance

The extraction of this distribution has been placed in Appendix A4 to get an even better understanding of the distribution between the different cities.

6.1.2 REGIONS AND POSTAL CODES

The distribution of postal codes can be helpful for three departments within Peddler. Their use depends on the measures that are used in the ranking. Ranking postal codes based on the total orders help the Sales department to identify which postal codes have a high Peddler brand awareness due to the popularity of Peddler in those areas. When ranking them on the percentage of on time delivered order, the Logistics department can identify weak point in efficiency but also postal codes that can serve as good examples.

RANKING BASED ON TOTAL NUMBER OF ORDERS

As mentioned in chapter 3, there is a difference between the cities and the regions Peddler operates in. For the total performance of the different locations, the regions have been analyzed as well. In Figure 10 the top ten best and worst ordering postal codes are analyzed and compared with each other. This is purely focused on the total orders each of these postal codes has had the last year (Appendix A6). The goal of this analysis is to give Peddler an insight if their total orders are evenly spread out over each postal codes within the region or not. An example of how this looks on the map for Groningen is shown in Appendix A8. The following findings and points of interest are derived:

- The ten lowest buying postal codes combined have bought 0.225% of the ten highest buying postal codes. The total number of orders in the region of Amsterdam is 43,172. This means that the top ten postal codes are responsible for 35.05% of this number while the bottom ten are only 0.079% responsible for this. An explanation for this are the geographical locations of some of these postal codes. 1026, 1027 and 1028 are all above Amsterdam North which is an area with a small density. 1027 and 1028 both have had only order in an entire year. Most of the top ten postal codes are more towards the center.
- Regions like Utrecht and Rotterdam are more evenly spread out with the bottom and top ten of both regions with Utrecht having more than six times as many orders in the bottom ten than the region of Amsterdam.

Top 10 best postal codes in each region	Total orders	On Time %	Total on time orders
Amsterdam	15.132	65,0%	9838
Den Haag	5.164	49,8%	2570
Utrecht	4.240	73,0%	3097
Rotterdam	2.146	80,1%	1718
Groningen	190	81,0%	128
Top 10 worst postal codes in each region			
Amsterdam	34	58,8%	16
Den Haag	29	42%	13
Utrecht	209	70,1%	151
Rotterdam	98	79,6%	78
Groningen	23	48,0%	13

Figure 10 City ranking based on orders

RANKING BASED ON PERCENTAGE OF ORDERS DELIVERED ON ON-TIME DELIVERIES

Ranking the top and bottom postal codes of a region based on their average percentage of on time delivered orders help to identify strong and weak points in regions. Route planners can have a list of postal codes that are known to have bad efficiency percentages and let their routes be influenced by this knowledge. In order to find a reasoning behind the ranking, two additional area analyses have been carried out which will be explained below. The ranking of the postal codes of Amsterdam has been placed in Appendix 10 as an example.

The following points of interest based on the percentage of on-time orders were found:

- Within the bottom eleven scoring postal codes in the region of Amsterdam, two belong to Diemen. This remarkable because Diemen only has three postal codes of which 1112 is the worst performing of region Amsterdam with an on-time percentage of 40%. Due to these reasons, is not active in Diemen anymore.
- The average of the top performing postal codes with more than ten orders combined, is still only 59.6%. The national average is 61.1%.
- The bottom postal codes of region The Hague that have more than 10 orders are all outside the city of The Hague. This list consists of Wateringen, Leidschendam, Rijkwijk and, most frequently present, Delft. Of the 12 postal codes of Delft in total, 5 are in the bottom 15. Even worse, if the bottom 23 postal codes with more than ten orders are listed, all Delfts postal codes are in there. Of the 1,919 orders, only 22.6% was on time. The reason for this position in the ranking, is that Peddler only sends a rider to Delft when there are enough orders to make it worth the costs of the rider. It's a choice between costs and efficiency and with Peddler still being in the startup phase, they choose cost above efficiency. This results in many orders being delivered too late in Delft.
- The top 15 postal codes in Rotterdam all have an average percentage above the 90% which is better than any other region.
- The worst performing postal code of Rotterdam with more than ten orders has an average of 63.2%. This postal code would have ranked fourth in the top five postal codes of The Hague with more than ten orders.
- The top three of Groningen postal codes with more than 5 orders (adjusted because the total number of orders in Groningen is much lower) consists completely out of orders from Haren. This is not a big surprise considering that 39 out of the 40 in total were books.

The comparison between the top and bottom fifteen most active postal codes in Groningen already showcased the influence of the geographical location of certain postal codes. First, an analysis based on the size of the area of the postal have been carried out to see if more area is in expense of the efficiency of that postal code (Appendix A8). Second, an analysis that focusses on the geographical location of the postal code has been carried out (Appendix A9). Both analyses have been laid next to the efficiency performance of the top and bottom ranked postal codes to see if a trend occurs. An example of this very small research has been placed in (Appendix A10), this is an extension of the ranking mentioned above.

The following points of interest were found based on these two analyses:

- All of Amsterdam's and Utrecht's top ranked postal codes are located in the center of the region.
- All of Amsterdam's, The Hague's, and Utrecht's (except one postal code) bottom ranked postal codes are located in the suburbs of the region. This is especially clear for Utrecht (Appendix A11)
- None of Amsterdam's, The Hague's, and Utrecht's bottom ranked postal codes are considered to cover a small area. A few of these postal codes do cover a large area.
- None of Amsterdam's, The Hague's (except one postal code), and Utrecht's top ranked postal codes are considered to cover a large area.
- Most of Groningen's top ranked postal codes are located in the suburbs. Again, most orders in these postal codes were books.
- None of Rotterdam's top ranked postal codes cover a small area and none of the bottom postal codes cover a large area.

6.2 TIME BASED RESULTS

Apart from the geographical factor for performance efficiency, the factor of time or time period is also of high importance for the performance and efficiency of an order. These analyses focus primarily on the second dashboard which is focused on the dimensions date and timeslot. A cohort analysis has been carried out to measure the performance of peddler during the last year. It focuses on the following KPIs:

Percentage of orders delivered on time per week
Percentage of orders delivered on time per month
Percentage of orders delivered on time per timeslot

6.2.1 PERFORMANCE OVER TIME

As mentioned before, the bar charts in the second dashboard provide an insight in the performance of Peddler over time. Looking at performance of the last year of Peddle, it is clearly visible when Peddler started to gain more brand awareness. As shown in Figure 11, the total number of orders increases rapidly as the holiday season starts to come closer. It all starts in October when the number over orders doubles compared to August. This number increases again with 33% in November and arrives at its peak in December. A total of 14,760 orders have been delivered in the month of December. This is an increase of 434.91% in orders when comparing it to September. In other words, the holidays have been extremely beneficial for Peddler in terms of sales. After December it slowly drops down back to 6,280 orders in February which it keeps. December managed to double the amount of orders going from 3,380 in September to 6,280 in February.

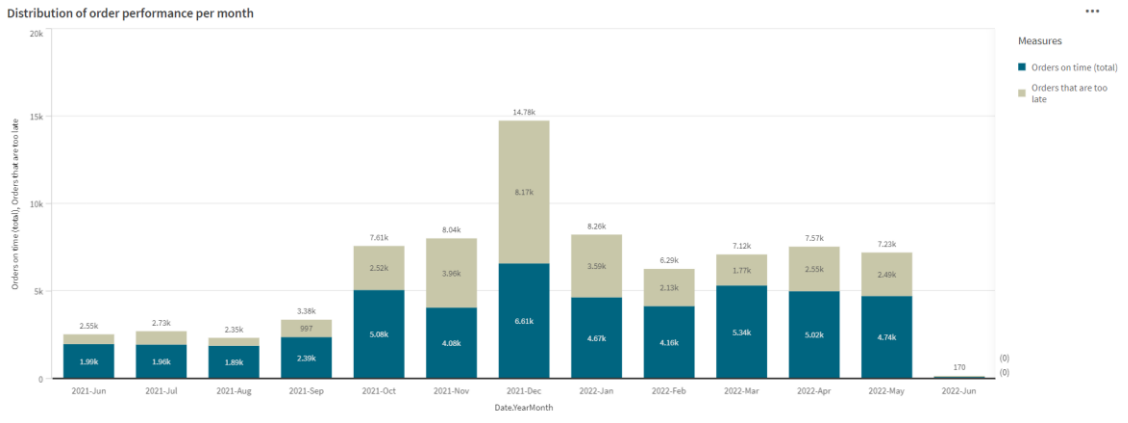


Figure 11 Orders over time

In terms of efficiency, December was not a great month for Peddler. Figure 12 is the opposite of Figure 11, showing a decrease in performance as Christmas comes closer and closer. December has been the worst performing month of Pedder of the last year with only 44.7% of orders arriving on time. The explanation for this is that the number of orders that needed to be delivered was larger than the capacity of Peddlers riders resulting in many delays.

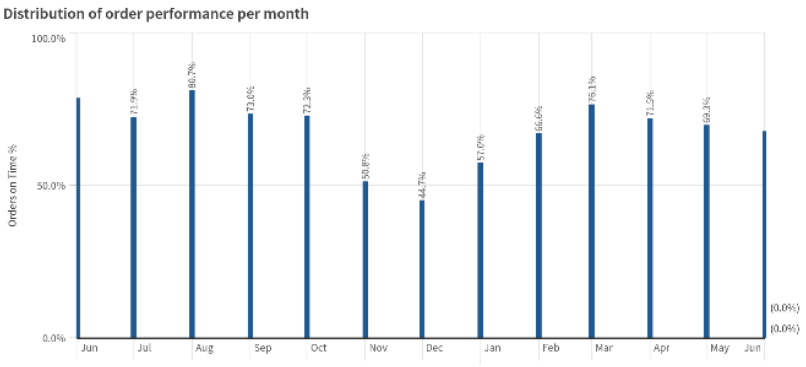


Figure 12 Distribution of order performance per month

Looking at the performance of last year in terms of weeks, it is aligned with the graphs above ranking week number 33, 32 and 34 as one of the best performing weeks with an average of 89,2%, 84,3% and 80,7%. All these weeks are in the month of August resulting in August being the best performing month until now (Figure 12). Looking at Figure 13, an explanation can be found. The number of orders from June to August did not fluctuate that much. In August, Peddler had the correct structure to deal with this number of orders.

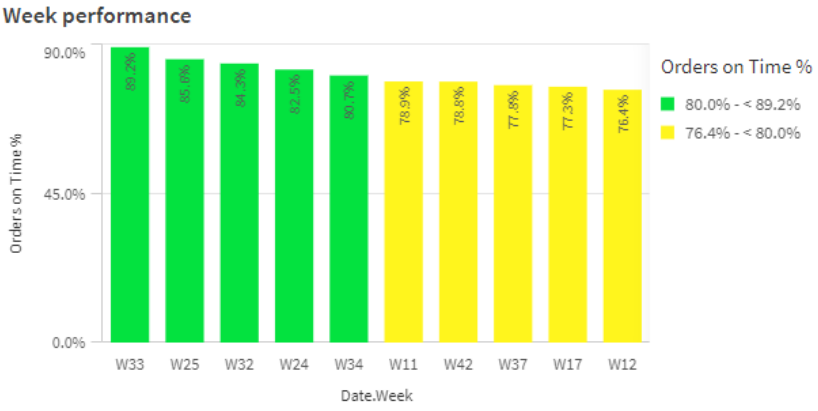


Figure 13 Week performance

6.2.2 TIMESLOT PERFORMANCE

As mentioned before, one of the biggest factors for the feasibility of a route is in which timeslot the order has been placed. The results of section 6.1 were purely focused on the geographical location but this is not the complete picture to fully determine the performance. Figure 14 shows the performance of each timeslot based on the percentage of on time orders. Aligning this with Figure 12 and it becomes clear how the curve of these lines came to be. The timeslot for 12:00-14:00 is performing the weakest with a peak of 77% in August and 28.1% as its lowest point in December. A preventive measure for the month of December saw the introduction of the 18:00-20:00 timeslot in October, giving Peddler more possibilities to spread the large number of orders over the day. Looking at the performance over time for these timeslot, it shows the following fluctuations in performance:

- 12:00-14:00: going from 64% in October to 28.1% (-35.9%) in December back to 55.3% (+27.2%) in May.
- 14:00-16:00: going from 75.7% in October to 48% (-27.7%) in December back to 65.5% (+17.5%) in May.
- 16:00-18:00: going from 80.5% in October to 48.5% (-32%) in December back to 83.9% (+35.4%) in May.
- 18:00-20:00: going from 100% in October to 49.6% (-50.4%) in December back to 100% (+50.4%) in May .

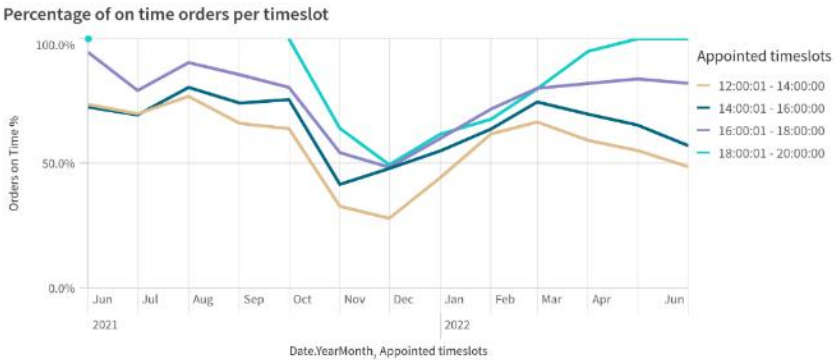


Figure 14 Percentage of on time order per timeslot

Only the timeslot 16:00-18:00 has fully recovered and even increased after December by growing an additional 3.4%. The timeslot from 12:00-14:00 is proven to be the weakest link again by only increasing back to 55,3%. When adjusting the time frame to only the months October 2021 to May 2022, it becomes clear how weak this timeslot is performing on average. Looking at Figure 15, it shows that, even though it has the smallest number of orders of all timeslots, it has the highest number of orders that were delivered too late. This results in them only having a small share in the distribution of on time orders per timeslot (Figure 16). The 16:00 – 18:00 timeslot is proven to be more efficient by having a 29.3% share in the total number of on time order while not having the most orders in total.

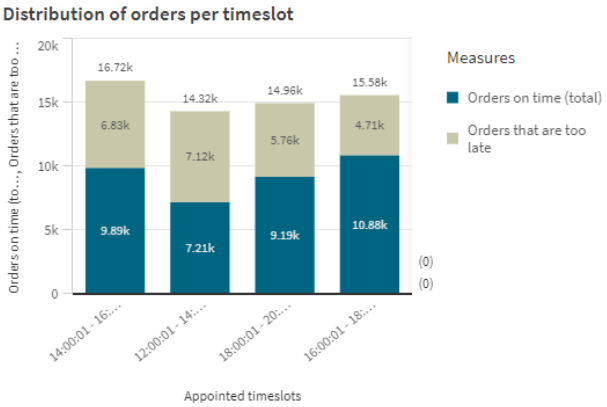


Figure 15 Distribution of orders per timeslot

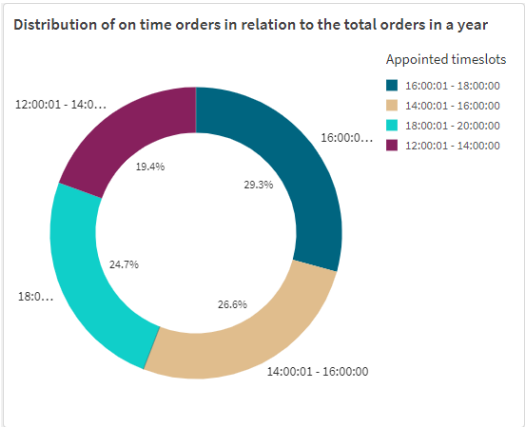


Figure 16 On time order percentage distribution

This proves that timeslots are not yet used efficiently and can use some improvements. The following section provides a solution for this problem.

6.2.3 TIMESLOT PREFERENCE FOR POSTAL CODES

Combining these results, it becomes clear that the weight, in this case the amount of orders, is not balanced between the different timeslots. Because of this, a small research has been carried out to see how each timeslot performs per postal code. Currently, there is not a consistent data substantiated formula for the division of timeslots per postal code. In other words, when customers order something from the website they still have the possibility to select from every timeslot as long as the order has been placed four hours before it. This small research functions as timeslot preference advice for each postal code to improve the performance and efficiency of each postal code and therefor the performance of Peddler itself.

After selecting all regions and comparing the performance of each timeslot per region (Appendix A12), an overview can be extracted from the third dashboard. As mentioned in Section 5.2, this dashboard consists out of an interactive pivot table that creates the possibility to get insight in this needed performance. After extracting this data from the dashboard, an overview as shown in Appendix A13 is created. With the use of some conditional formatting rules in which the highest percentage of on time orders is highlighted with green, each postal code is given a timeslot preference.

Extracting this to a new table, a form is created containing some new important variables substantiating the timeslot preference (Appendix A2). Apart from the postal code, time slot preference and correlated performance measured by on time orders, the increase in percentage when comparing this preference to the second best timeslot preference is also added. This gives Peddler an idea how much they are improving by choosing this timeslot over the second best choice. Last, the total number of orders that have been ordered in this timeslot has been added as well. This gives a better understanding of the credibility of the substantiation.

The postal codes shown in Appendix A2 are the outliers of this new recommendation of timeslots per postal codes. They are selected based on their increase in performance percentage compared with the second timeslot preference and their number of orders which needs to be more than ten. The following points of interest are derived from this table:

- The biggest increase in the percentage of orders being delivered on time is postal code 2512 with the preference of the 18:00-20:00 timeslot. The difference between the first and second preference is 31.2%.
- These are the 53 postal codes with the highest increases in performance. The total number of operating postal codes is 373. The total increase in performance when comparing the preference with the second preference from these 53 postal codes alone, would increase the performance with 872%. This is purely based on the percentage of orders delivered on time. The total number of orders has not been implemented as a weight for this calculation.
- Some closely located postal codes are within this list of Appendix A2. For instance:
    - 1095 & 1096: increasing 9.6% and 11.5% having a total number of 112 and 173 orders.
    - 2241, 2242 & 2243: increasing 20%, 25%, and 10.6% having a total number of 10, 27 and 17 orders.
    - 2282 & 2283: increasing 14.9% and 31.1% having a total number of 45 and 21 orders.
    - 2581, 2582, 2584 & 2586: increasing 19%, 14.1%, 13.6% and 20.1% having a total number of 18, 56, 43 and 29 orders.
    - 3071, 3072 & 3073: increasing 16.6%, 28.2% and 11.1% having a total number of 107, 66 and 27 orders.
- Looking at Appendix A2, the preferences are as follows:
    - 12:00-14:00: 3 postal codes
    - 14:00 – 16:00: 7 postal codes
    - 16:00 – 18:00: 36 postal codes
    - 18:00 – 20:00: 7 postal codes
- Looking at the total list of 330, the preferences are as follows. These numbers do not add up to 330 because some postal codes did not have enough orders to suffices the calculation:
    - 12:00-14:00: 47 postal codes
    - 14:00 – 16:00: 63 postal codes
    - 16:00 – 18:00: 160 postal codes
    - 18:00 – 20:00: 58 postal codes

6.3 STORE AND RIDER BASED RESULTS

Lastly, the two important factors that also have a large influence on the efficiency and performance are the stores and riders Peddler is currently collaborating with. Dashboards four and five focus on these two factors and help to give Peddler a more elaborate insight in the influence some of their collaborations have on the region they are operating in. As has been done in chapter 5, the names of the stores and riders have been removed due to privacy reasons. These results will have an impact on the following KPIs:

Number of failed pickups
Percentage of orders delivered on time per rider
Number of on time pickups per rider
Percentage of orders delivered on time per store

6.3.1 STORES

Looking at dashboard four, the main focus lies on the performance of the stores Peddler is collaborating with. It helps them to see what share these stores have in the total market of Peddler. For instance, selecting the top ten biggest stores when it comes to total orders and seeing how these orders are distributed over the postal codes gives an insight what postal codes have the biggest share in the success of some of these stores. For instance the postal code of 1091 is responsible for 83 of the 520 total orders of one store. This is due to

geographical reasons, the store itself also has 1091 as its postal code. It comes to no surprise that 88% of orders from this store to this postal code are delivered on time.

Plotting these stores with total orders as measure over time and it becomes more clear how these stores have been performing last year .The following points of interest have been identified:

- The top nine stores of top ten best stores are all from Amsterdam indicating once again how big the role of this region is for Peddler.
- A lot of postal codes that have a big share in the total number of orders, are the same postal code of the store itself or are located very close to postal code of the store.

Other findings are kept confidential due to the privacy rights of the stores that are ranked and compared.

6.3.2 RIDERS

The fifth and final dashboard that has been created functions as a visualization of the performance of riders that are employed by Peddler. Comparing the riders with each other without filtering per city would be unfair due to the different circumstances each city faces. Riders operating in Delft will be scored much lower compared to Rotterdam but that does not immediately depend on the capability of the rider. Due to privacy reasons, this research won't go too much into depth with the performance of each individual rider. The insights that are to gain or are for Peddler themselves.

A point of interest is the influence a city has on a rider. Looking at the scatter plot of the dashboard, there are hardly to no riders from the regions of Rotterdam, Utrecht and Groningen that have delivered over a 100 orders and have an average on time percentage below the 50%. Looking at Groningen in general, all riders are above the 50%. However, Rotterdam takes the lead here again. All riders with over the 100 orders are above 74%. Taking a look Delft again, only one rider with more than twenty orders has an average above 50%. All the others, which is in this case 30 riders, are below this percentage.

6.4 CHAPTER 6 SUMMARY

This chapter provided all the results that were obtained from the five created dashboards. It has been split up into three different types of result.

- **Location based results:** results of the different regions, cities and postal codes.
    - Cities are compared and scored based on the number of orders and what's the percentage of these orders has been delivered on time.
    - The mentioned cities are linked to the regions they are located in. This provides a better overview how regions perform.
    - Postal codes in the regions are compared. In addition, the postal codes have been ranked with regards to their location and surface.
- **Time based results:** results of the location based results measured over time and timeslot results.
    - Performance of months and weeks are compared with each other to see if there are any outliers last year.
    - Results of the performance of different timeslots over the year. With the change of demand over time, these results showed how each timeslot was effected.
    - Results show new timeslot preferences for each postal code.
- **Store and rider based results:** results of the different stores and riders.

7. CONCLUSION

The main motivation for this research is to help Peddler with their core problem: get a clear view of how their decisions in route planning impact their efficiency and performance. This has been solved by closing the gap between norm and reality, which resulted in the creation of multiple dashboards. The visualization tool that has been selected for this research is Qlik to make it easier for employees because Peddler was already using this tool for other performance measurements. The creation of five dashboards, with each providing Peddler more insights into their logistics performance. During this research and the creation of these dashboards, the following findings, recommendations and limitations were present.

7.1 MAIN FINDINGS

As described in chapter 6, all five dashboards have been a solid foundation for multiple findings. Regions, timeslots, postal codes and months have been analyzed and multiple points of interest have been found. Below are the main findings of this research.

Looking at the geographical results, a few outliers have been identified. One of the most positive findings was the performance of the region of Rotterdam. The top 15 postal codes in Rotterdam all have an average percentage above the 90%, which is better than any other region. The city of Rotterdam also performed amazingly. It has been able to deliver 82.5% of orders. This is all very impressive. Unfortunately, some negative outliers have been identified as well. The worst performing city is Delft, having only 22.6% of their orders delivered on time. This mostly has to do with geographical reasons. Delft's location is just between The Hague and Rotterdam, making it a distant travel location for both. It is in favor of Rotterdam that this has appointed to region The Hague.

It was also discovered that in some regions, the contribution of each postal code differed a lot. In Amsterdam for example, the ten lowest buying postal codes combined have bought 0.225% of the ten highest buying postal codes. The total number of orders in the region of Amsterdam is 43,172. The top ten postal codes are responsible for 35.05% of this number, while the bottom ten are only 0.079% responsible for this. This identified difference in contribution helps Peddler to see in which parts of the region they are known and in which parts they can improve their brand awareness.

The main finding as well was the performance of the month December. December has been great in terms of order quantities. However, as shown in this research, it came at the cost of orders that were delivered on time. Peddler could not handle this great increase in demand. This started to show in October already. This year they have a better expectation of the possible increase in demand, and they can anticipate it.

One of the biggest findings of this research is the description of timeslot performance. All four optional timeslots have been compared with each other to see if there are outliers. It has been proven that the 12:00 – 14:00 timeslot is the most inefficient. It shows that, even though it has the smallest number of orders of all timeslots, it has the highest number of orders that were delivered too late. To expand on this, a small investigation was done to see how these timeslots could be used more efficiently. A rapport which functions as advice for Peddler has been made. Here, each postal code has been given a preferable timeslot based on the number of on-time orders this timeslot has had.

7.2 LIMITATIONS

Given that Peddler is still an evolving company, some limitations occurred during the process of this research. These shortcomings were present due to choices by Peddler that have been made later in time. Some of these choices or shortcomings were minor inconveniences but some had a big influence on the report. The following limitations were present in this research:

- Because Peddler is still in its startup phase, they are experimenting with the type and the number of variables they want to track in their administration. In practice, this means that the extraction of both order types differed for each month resulting in variables being added, relocated in the file or removed. For example, the extraction of the delivery orders for March, April and May had 161, 168 and 163 respectively, different variables. The location of all these variables also differed per each file. This prevented the complete automation of macros because a fixed column would be needed for this. Due to these different columns, the Excel files needed to be prepared first by placing the right variable in the right column. After that was done, the macros could run and further prepare the data.
- An example of these experimentations with variables that had a big influence on the research, is the tracking of pickup and delivery attempts. Peddler introduced these two variables in February 2022 so there is not much data for these variables. Unfortunately for Peddler and the created dashboards, this means that orders before February could have had attempts that were bigger than one that are now considered as too late.
- Another example is the introduction of variable *store name* in September 2021. Before this month, only the *storeid* was stored in the admin. In order to still get the correlating store names for the months of July 2021 and August 2021, a VLOOKUP function in Excel was used with a file containing all *storeids* and *store names*.
- Another big limitation is the ability for customers and riders to fill in variables manually. This creates the possibility of errors that could have been avoided with the implementation of a selection list with options to pick from. Below are some examples how these limitations were an inconvenience for this research:
    - The variable “delivery city” has been one the most influential variables for this research and unfortunately it needed some adjustments before it could be used in Qlik. The reason for this was the number of variations each city has. The ability to let the customer fill in their city manually allows for spelling mistakes resulting in variables named “Amstredam”, “AMsterdam”, “Amstream” and ”amsterdam”. In total, there were 51 formulation versions for the city of Amsterdam alone. Luckily, Qlik has the option to group these versions together but it could have been prevented in the first place.
    - In addition, sometimes the variable delivery city was filled in completely wrong making the data unusable. For example, customers wrote their cell phone in this field and it somehow got excepted. All of these unusable rows have been grouped as the delivery city “filled in incorrectly”.
    - There were so many blanks for the variables “Delivery state” and “Storeid” (only for delivery orders) that these variables were left out during the preparation of the data.
    - Some postal codes had different cities attached to them even though they belong to only one them. For these postal codes, the number of orders was leading for the decision to which city it belongs. Most of the other cities only had less than five orders within this postal code. For example:
        - 2264: Leidschendam and The Hague;
        - 2584: The Hague and Scheveningen;
        - 2597: Voorburg and The Hague;
        - 3451: Utrecht, Vleuten and Hof van Twente.

7.3 DISCUSSION

The weakest point of this research is the inclusion of manually adjusting the Excel files in order to prepare them for Qlik. It will be discussed in Section 7.6, but the addition of real live data synchronization would increase the credibility of this research. The exclusion of manual adjustments makes the chance of mistakes a lot smaller. It would prevent Peddler from reacting to incorrect performance data. It also prevents mistakes during a handover. Right now, the researcher is only capable of adding data to Qlik. This means that after July 2022, no data has been added to the dashboards. A guideline on how to do it has been made, but this does not fully exclude the possibility of errors. This can be prevented with the addition of live synchronization.

In addition, the importance of the rider and store performance is less groundbreaking than the data in the other three dashboards. If the region of The Hague is performing poorly because of a geographical inefficiently located hub, the performance of stores and riders will not be great as well. These dashboards only serve as an identification tool for unusual outliers instead of overall performance. A rider in Rotterdam performs better than a rider in The Hague, but that does not immediately mean that the rider from Rotterdam is better. They do become interesting when there are outliers under the same conditions.

7.4 RECOMMENDATIONS

The gap between norm and reality has been filled by providing Peddler insight in their performance with the use of five different dashboards. Analyzing these created dashboards led to the results mentioned in section 6. Based on these results, the following recommendations can be given to Peddler.

One of the biggest deliverables of this research apart from the dashboards are the results that were derived from the second and third dashboard. Timeslots were not performing up to their full potential which led to a small investigation to see how Peddler can improve the use of these timeslots. After the analysis of the performance of each timeslot per postal code, a list with the preferable timeslot for each of the 330 different active postal codes has been made. This list has the potential to help Peddler increase their efficient use of these timeslots for each of their operating postal codes. Currently, the customer has the option to select any timeslot they want as long as the order has been created four hours in advance of the delivery. If this option were to be replaced with a preferable timeslot from Peddler to the customer, the performance in efficiency would increase.

Next, a recommendation for Peddler is to take a critical look at the structure of their operation regions based on dashboard one. This dashboard provides an overview for the performance of each city and their correlated postal codes. Postal codes that are located in the suburbs which have a low total number of orders and a low percentage of on time delivered orders, can be removed from the operating region. There needs to be a balance between revenues and costs for these postal codes. In this case, cost being the loss in efficiency and the salary of riders. If the revenue does not add up to these costs due to the lack of orders, suitable measures can be taken. The most important example of this is the way the region The Hague is structured. Currently, it has a lot of smaller cities attached to it making it hard to create efficient routs. This results in an extremely low performance of some of these out layers, Delft being the biggest example here. The opening of a new hub somewhat closer to these areas could help to increase the performance for these postal codes. If this is too expensive, the decision to drop some of these poorly performing postal codes would increase the performance as well.

Peddler could use dashboard four to increase their knowledge in store performance under the same conditions. For example , this means comparing and scoring stores in the same region. Updating the dashboard the upcoming months helps them to gain more insight in reoccurring pickup failures for example. If a store is hardly sells any products and constantly fails to have the pickup ready at the agreed time, Peddler can choose to come with sanctions or other solutions to fix this. In addition, they can adjust their marketing for each postal

code or time period. It has been shown that stores that sell products that can serve as gifts increase in popularity in the months of October, November and December. Peddler has also gotten an insight in the share of each postal code for each store. This way, they can see postal codes are more interested in certain types of stores.

Another recommendation is the introduction of a list with options to select from for certain variables instead of letting the customer filling it in manually. As shown in Section 7.2, it could save a lot of trouble. It is also less effort for the customer which is almost always an improvement.

7.5 SCIENTIFIC CONTRIBUTION

Given the practical nature of the problem this thesis is based around, the scientific contribution is limited. This is not the first time a dashboard has been made, and it will not be the last time. This paper does, however, provide a step-by-step data-focused research method and how you can apply this method to a logistical problem using a combination of quantitative and qualitative data analysis. It also shows how to determine a solid foundation for the visualization of data using multiple KPIs.

7.6 FUTURE RESEARCH

Since this was the first time Peddler experimented with the visualization of data in the Logistics department, there is quite some for this research to expand on. Due to the scope of this research, a complete implementation was left out. Long term research about the effects of this implementation would be important for Peddler as well. The dashboards themselves can be improved as well. An example of this is the addition of real live synchronization. This means that as soon as an order is delivered, the dashboard is refreshed, and the order is added. In other words, the dashboard needs to be connected to the admin of Peddler, which allows the company to add newly stored orders automatically. This would remove the need for Excel files completely, which is a big improvement. During this research, macros were used, and Excel files needed to be adjusted. This part would be skipped entirely, which means there is less room for mistakes.

More variables can be added, resulting in more KPIs and possible more dashboards. This would expand the amount of knowledge Peddler has in their performance. Another part that can be researched in the future is the result in performance after the timeslot preferences have been implemented. It would be interesting to see if the more efficient allocation of timeslots has a large impact on the overall performance of Peddler.

REFERENCES

Cox, D. R., & Snell, E. J. (2016). The Choice of Variables in Observational Studies on JSTOR. Jstor.org. ttps://www.jstor.org/stable/2347053?seq=5

De'ath, G., & Fabricius, K. E. (2000). CLASSIFICATION AND REGRESSION TREES: A POWERFUL YET SIMPLE TECHNIQUE FOR ECOLOGICAL DATA ANALYSIS. Ecology, 81(11), 3178–3192. https://doi.org/10.1890/0012-9658(2000)081[3178:cartap]2.0.co;2

Fernando, J. (2022). Learn What Stakeholders Are and the Roles That They Play. Investopedia. https://www.investopedia.com/terms/s/stakeholder.asp

Glenn, N. D. (2013). Cohort Analysis. Google Books. https://books.google.nl/books?hl=nl&lr=&id=KZrEevjbE80C&oi=fnd&pg=PP7&dq=cohort+analysis&ots=LZ2uqzF-3b&sig=RS3o9Uy6NLwMioWEQGd38r1ML1U#v=onepage&q=cohort%20analysis&f=false

Hexmoor, H. (2015). Computational Network Science. Elsevier. https://doi.org/10.1016/c2013-0-19272-0

Lillis, A. M. (1999). A framework for the analysis of interview data from multiple field research sites. Accounting and Finance, 39(1), 79–105. https://doi.org/10.1111/1467-629x.00018

Lin, Y.-S., Jiang, J.-Y., & Lee, S.-J. (2014). A Similarity Measure for Text Classification and Clustering. IEEE Transactions on Knowledge and Data Engineering, 26(7), 1575–1590. https://doi.org/10.1109/tkde.2013.19

Miles, M. B., & Huberman, A. M. (1994). APA PsycNet. Apa.org. https://psycnet.apa.org/record/1995-97407-000

Rapkin, B. D., & Luke, D. A. (1993). Cluster analysis in community research: Epistemology and practice. American Journal of Community Psychology, 21(2), 247–277. https://doi.org/10.1007/bf00941623

Robinson, S. (2014). Verification, validation and confidence. In S. Robinson, Simulation | The practice of model development and use (p. 358). New York: Palgrave Macmillan.

Rohrer, J. E. (2014, June 12). Ten Important Elements for Observational Studies in Primary Care and Community Health. ResearchGate; SAGE Publications. https://www.researchgate.net/publication/263099711_Ten_Important_Elements_for_Observational_Studies_in_Primary_Care_and_Community_Health

Tufféry, S. (2011). Cluster Analysis. Data Mining and Statistics for Decision Making, 235–286. https://doi.org/10.1002/9780470979174.ch9

Wedel, M., & Shi (Savannah), W. (2010). Cluster Analysis. Wiley International Encyclopedia of Marketing. https://doi.org/10.1002/9781444316568.wiem02018

APPENDICES

APPENDIX A1: INTERVIEW

Multiple departments were interviewed to see how the vision of everybody within the company aligns:

- Route planning
- Logistics

These were the questions for the Logistics department for example:

- What exactly is done at Logistics? How does your day/week look like?
- What are your ideas for the future? In other words, where do you like to see this company go to?
- How is the schedule for number of needed riders and route planners determined?
- What do think about the current way of planning routes?
- When creating a route, what factors influence the feasibility of the route?
- What are the norms and values (core values) of Peddler?
- How do you see these norms and values translated into Key Performance Indicators?
- If there is a difference in priority; what Key Performance Indicators are more important to the company, your department or your function?
- Do the previously mentioned factors influence this priority?
- Are there new Key Performance Indicators that you wish to see added?
- Which data should be collected to make these indicators calculable?
- What actions do you take if the monitored results are too low?
- In what intervals are the indicators reviewed?
- Which elements can't be monitored yet, but do you wish to control in the future?
- In addition to the last question, do you have some improvements for the system?

APPENDIX A2: TIMESLOT PREFFERENCE PER POSTAL CODE

Postal code	Best scoring time slot	Percentage of on-time orders	Improvement of on-time orders compared with second best scoring timeslot	Number of orders
1034	14:00:01 - 16:00:00	75,9%	14,1%	58
1060	14:00:01 - 16:00:00	75,6%	12,7%	45
1069	14:00:01 - 16:00:00	71,4%	13,5%	77
1074	12:00:01 - 14:00:00	76,8%	12,3%	112
1095	16:00:01 - 18:00:00	72,8%	9,6%	173
1095	16:00:01 - 18:00:00	72,8%	9,6%	173
1096	16:00:01 - 18:00:00	69,7%	11,5%	122
2241	12:00:01 - 14:00:00	70,0%	20,0%	10
2242	14:00:01 - 16:00:00	88,9%	25,3%	27
2243	12:00:01 - 14:00:00	70,6%	10,6%	17
2282	16:00:01 - 18:00:00	57,8%	14,9%	45
2283	16:00:01 - 18:00:00	61,9%	31,1%	21
2497	16:00:01 - 18:00:00	54,5%	21,2%	11

2512	18:00:01 - 20:00:00	82,9%	31,2%	35
2522	18:00:01 - 20:00:00	85,7%	35,7%	14
2526	14:00:01 - 16:00:00	88,9%	13,9%	18
2541	16:00:01 - 18:00:00	54,5%	10,1%	11
2546	16:00:01 - 18:00:00	75,0%	17,1%	28
2554	16:00:01 - 18:00:00	60,0%	11,7%	25
2561	16:00:01 - 18:00:00	73,5%	11,6%	34
2574	18:00:01 - 20:00:00	67,5%	13,0%	40
2581	16:00:01 - 18:00:00	83,3%	19,0%	18
2582	18:00:01 - 20:00:00	78,6%	14,1%	56
2584	16:00:01 - 18:00:00	65,1%	13,6%	43
2586	16:00:01 - 18:00:00	75,9%	20,1%	29
2591	14:00:01 - 16:00:00	61,5%	21,1%	26
2594	18:00:01 - 20:00:00	69,6%	11,2%	23
2596	16:00:01 - 18:00:00	67,2%	14,2%	67
2611	18:00:01 - 20:00:00	41,0%	12,1%	39
2628	18:00:01 - 20:00:00	43,3%	23,0%	30
3011	16:00:01 - 18:00:00	97,7%	14,5%	133
3012	16:00:01 - 18:00:00	87,8%	12,8%	41
3032	16:00:01 - 18:00:00	100,0%	17,1%	21
3034	16:00:01 - 18:00:00	96,7%	27,4%	60
3035	16:00:01 - 18:00:00	100,0%	15,8%	33
3038	16:00:01 - 18:00:00	100,0%	20,0%	44
3039	16:00:01 - 18:00:00	96,1%	12,7%	127
3054	16:00:01 - 18:00:00	94,4%	27,8%	18
3055	16:00:01 - 18:00:00	96,8%	16,8%	31
3062	16:00:01 - 18:00:00	96,8%	11,1%	63
3063	16:00:01 - 18:00:00	98,6%	14,0%	71
3071	16:00:01 - 18:00:00	94,4%	16,6%	107
3072	16:00:01 - 18:00:00	97,0%	28,2%	66
3073	16:00:01 - 18:00:00	88,9%	11,1%	27
3081	16:00:01 - 18:00:00	94,1%	10,8%	17
3083	16:00:01 - 18:00:00	96,7%	13,3%	30
3521	16:00:01 - 18:00:00	91,9%	20,2%	74
3524	16:00:01 - 18:00:00	79,1%	15,5%	153
3527	16:00:01 - 18:00:00	87,0%	11,4%	100
3533	16:00:01 - 18:00:00	87,6%	11,5%	145
3555	16:00:01 - 18:00:00	85,2%	10,4%	81
3582	16:00:01 - 18:00:00	81,3%	10,3%	64
9737	14:00:01 - 16:00:00	63,6%	23,6%	11

APPENDIX A3: AMSTERDAM PERFORMANCE

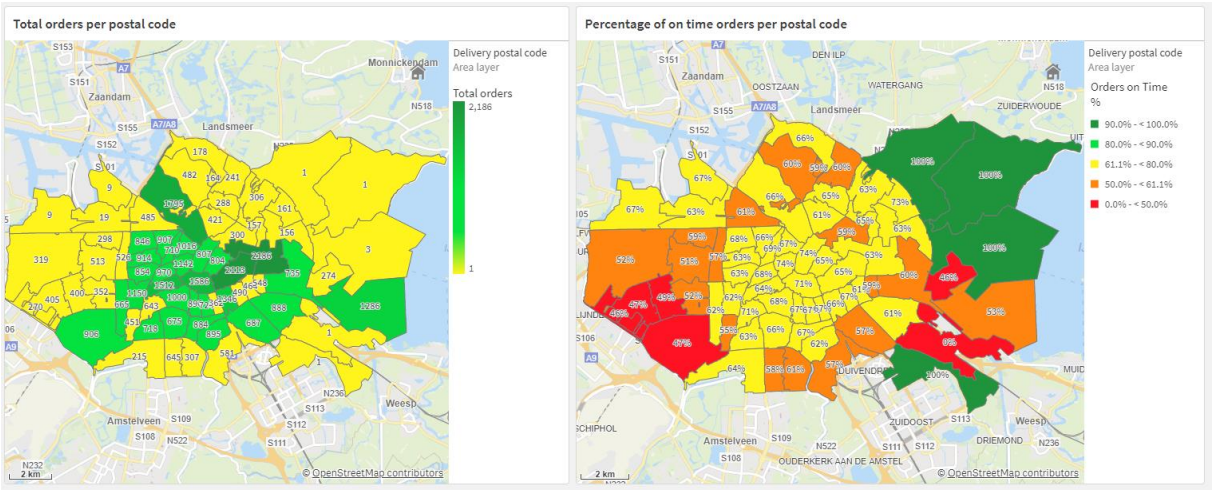


Figure 17 The selection of city Amsterdam

APPENDIX A4: TOTAL CITY PERFORMANCE

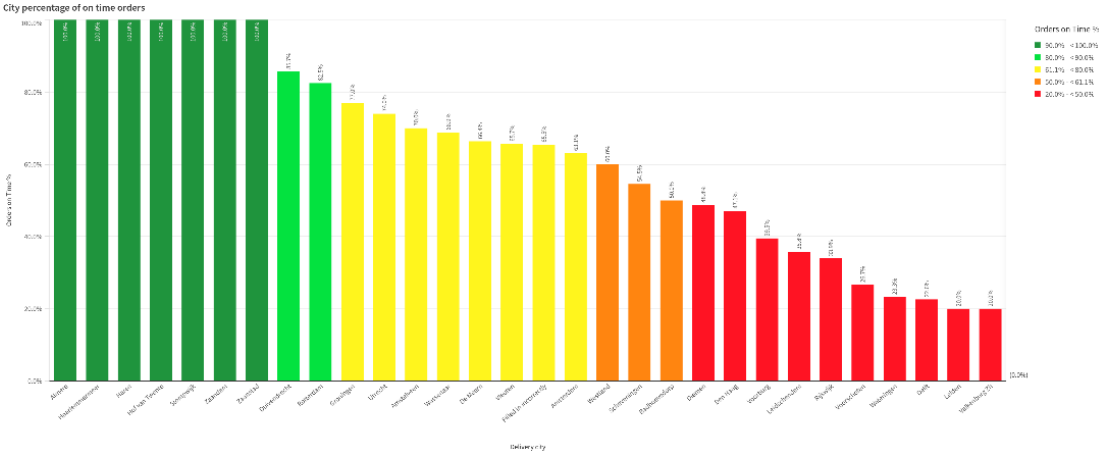


Figure 18 City comparison based on percentage of on-time orders

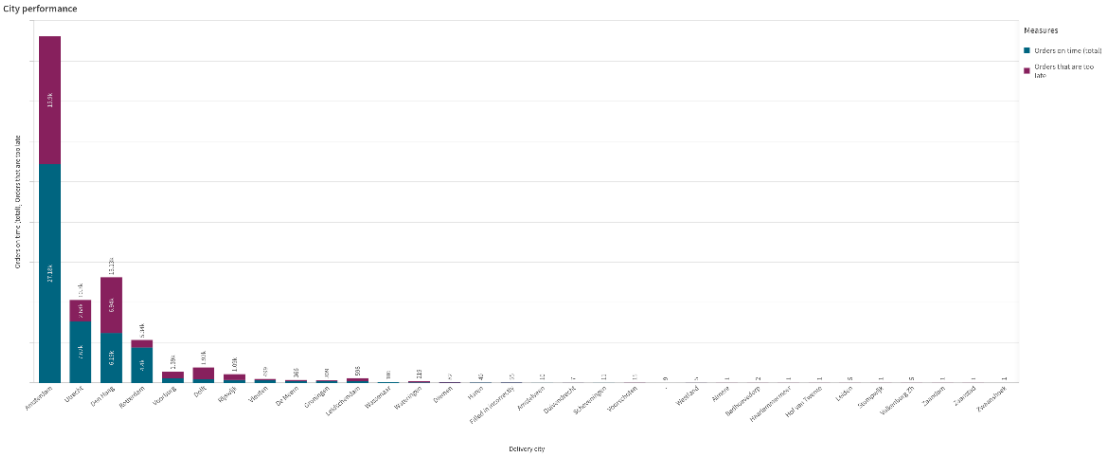


Figure 19 City comparison based on order quantities

APPENDIX A5: EXAMPLES OF ROTTERDAM AND DELFT

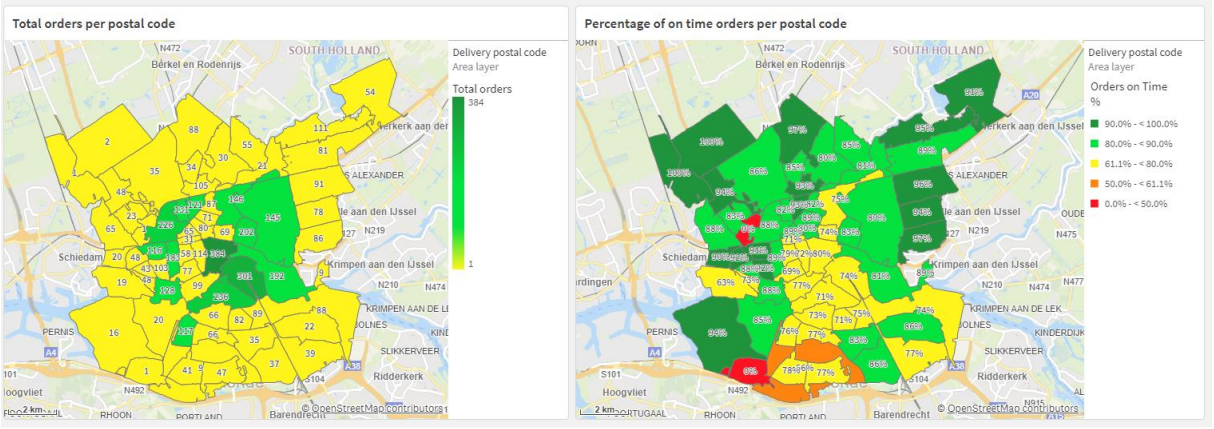


Figure 20 Rotterdam order quantities and percentage of on-time orders

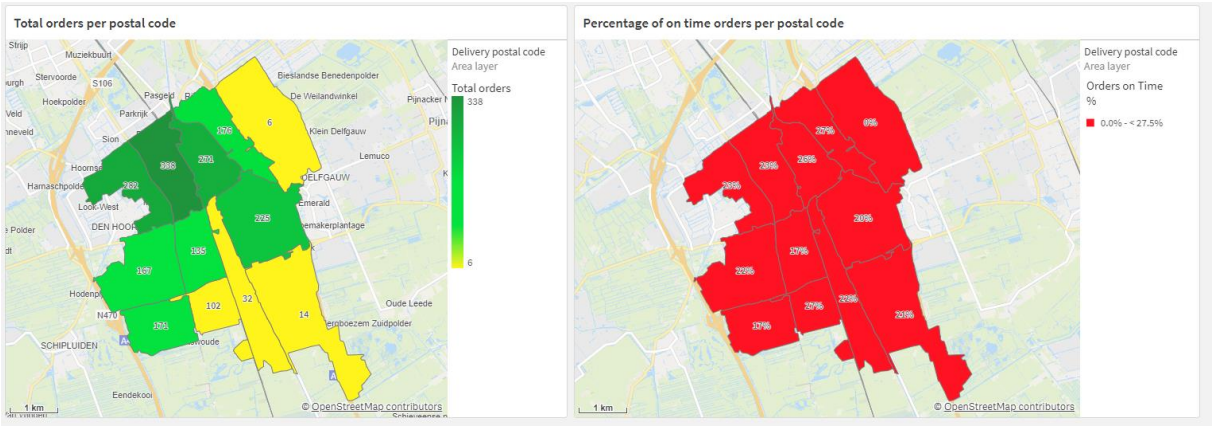


Figure 21 Delft order quantities and percentage of on-time orders

APPENDIX A6: RANKING OF POSTAL CODES BASED ON NUMBER OF ORDERS

Postal codes with the most orders				
Delivery city	Delivery postal code	Total orders	Orders on Time %	Orders on time (total)
Amsterdam	Totals	15.132	65,0%	9838
	1019	2.186	63,0%	1377
	1018	2.113	64,7%	1367
	1013	1.795	66,1%	1186
	1017	1.586	71,3%	1131
	1054	1.512	64,0%	968
	1091	1.346	65,9%	887
	1087	1.286	52,8%	679
	1058	1.150	62,4%	718
	1016	1.142	73,8%	843
	1015	1.016	67,1%	682

Postal codes with the least orders				
Delivery city	Delivery postal code	Total orders	Orders on Time %	Orders on time (total)
Amsterdam	Totals	34	58,8%	16
	1027	1	100,0%	1
	1028	1	100,0%	1
	1000	2	100,0%	
	1001	2	50,0%	
	1020	2	0,0%	2
	1070	2	0,0%	1
	1099	3	66,7%	2
	1026	3	100,0%	
	1042	9	66,7%	6
	1046	9	66,7%	3

Figure 22 Ranking of Amsterdam postal codes based on number of orders

APPENDIX A7: GRONINGEN EXAMPLE WITH HIGHEST AND LOWEST NUMBER OF ORDERS

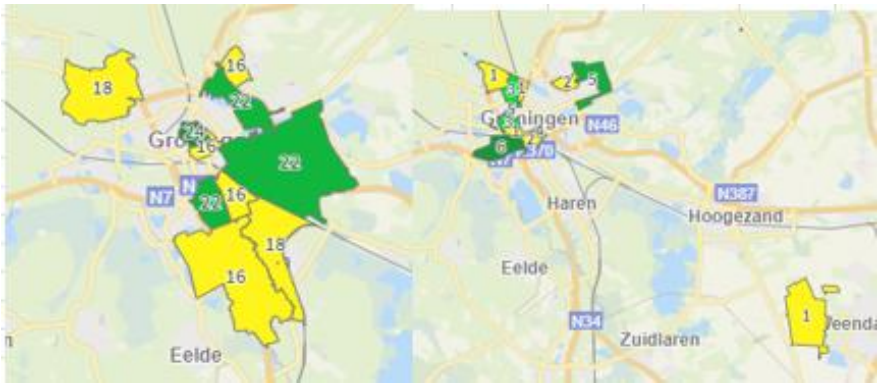


Figure 23 Groningen postal codes with the highest (left) and lowest (right) number of orders

APPENDIX A8: POSTAL CODE AREA SIZE ANALYSIS

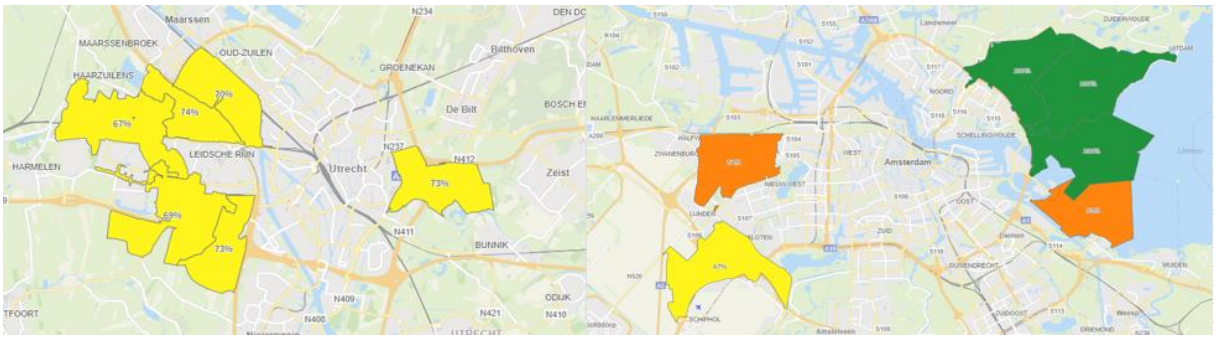


Figure 24 Utrecht (left) and Amsterdam (right) postal codes with a large surface

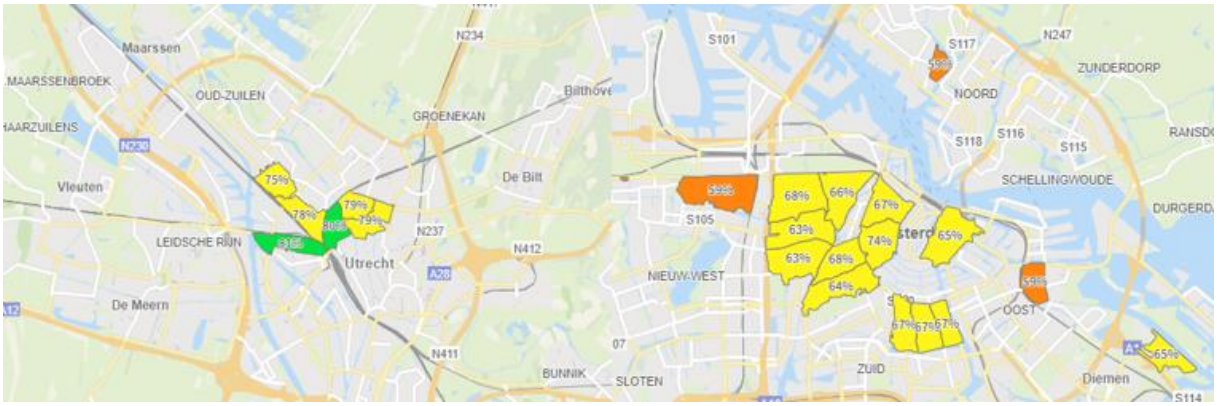


Figure 25 Utrecht (left) and Amsterdam (right) postal codes with a small surface

APPENDIX A9: POSTAL CODE CENTER OR SUBURB ANALYSIS

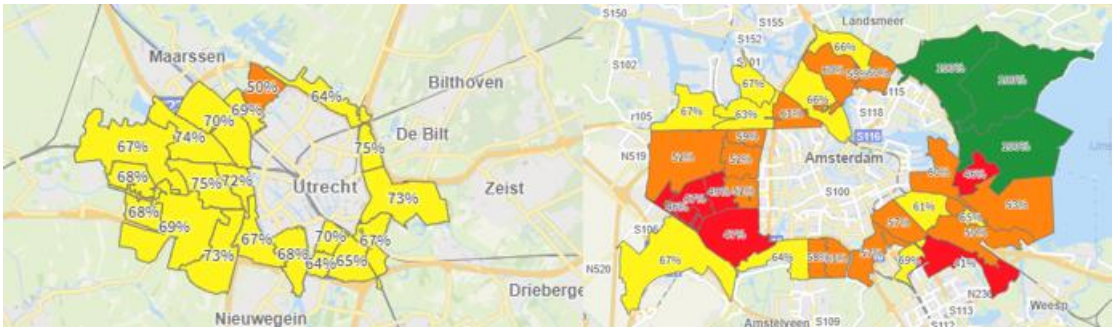


Figure 26 Utrecht (left) and Amsterdam (right) postal codes located in the suburbs

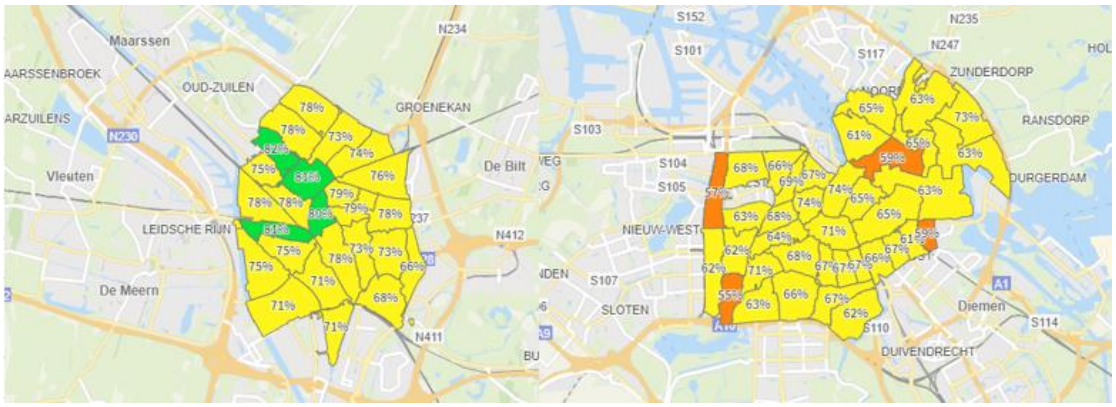


Figure 27 Utrecht (left) and Amsterdam (right) postal codes located in the center

APPENDIX A10:RANKING OF POSTAL CODES BASED ON PERCENTAGE OF ON TIME ORDERS

Rows marked with red are postal codes that don't have more than ten orders in total making them not a very good substantiation for a possible reoccurring trend. Within this research, clusters are made to see if a reoccurrence can be identified.

Top 15

Delivery city	Delivery postal code	Total orders	Orders on Time %	Orders on time (total)	More than 10 total orders?	Postal code covers a large area	Postal code covers a small area	Suburb?	Center?	Suburb and Larger area	Suburb and small area	Center and Larger area	Center and small area
Amsterdam	1026	3	100,0%	3	0	1	0	1	0	1	0	0	0
Amsterdam	1000	2	100,0%	2	0	0	0	0	0	0	0	0	0
Amsterdam	3011	2	100,0%	2	0	0	0	0	1	0	0	0	0
Amsterdam	1027	1	100,0%	1	0	1	0	1	0	1	0	0	0
Amsterdam	1028	1	100,0%	1	0	1	0	1	0	1	0	0	0
Amsterdam	1012	807	73,9%	596	1	0	0	0	1	0	0	0	0
Amsterdam	1016	1.142	73,8%	843	1	0	1	0	1	0	0	0	1
Amsterdam	1024	161	72,7%	117	1	0	0	0	1	0	0	0	0
Amsterdam	1017	1.586	71,3%	1131	1	0	0	0	1	0	0	0	0
Amsterdam	1075	643	71,1%	457	1	0	0	0	1	0	0	0	0
Amsterdam	1052	710	68,9%	489	1	0	0	0	1	0	0	0	0
Amsterdam	1053	970	68,4%	663	1	0	1	0	1	0	0	0	1
Amsterdam	1055	846	67,7%	573	1	0	1	0	1	0	0	0	1
Amsterdam	1071	1.000	67,5%	675	1	0	0	0	1	0	0	0	0
Amsterdam	1072	896	67,4%	604	1	0	1	0	1	0	0	0	1

Bottom 15

Delivery city	Delivery postal code	Total orders	Orders on Time %	Orders on time (total)	More than 10 total orders?	Postal code covers a large area	Postal code covers a small area	Suburb?	Center?	Suburb and Larger area	Suburb and small area	Center and Larger area	Center and small area
Amsterdam	1020	2	0,0%	0	0	0	0	1	0	0	0	0	0
Amsterdam	1070	2	0,0%	0	0	0	0	0	1	0	0	0	0
Diemen	1112	40	40,0%	16	1	0	0	1	0	0	0	0	0
Amsterdam	1060	270	46,3%	125	1	0	0	1	0	0	0	0	0
Amsterdam	1086	274	46,4%	127	1	0	0	1	0	0	0	0	0
Amsterdam	1066	906	46,6%	422	1	0	0	1	0	0	0	0	0
Amsterdam	1069	405	47,2%	191	1	0	0	1	0	0	0	0	0
Amsterdam	1068	400	48,8%	195	1	0	0	1	0	0	0	0	0
Amsterdam	1001	2	50,0%	1	0	0	0	0	1	0	0	0	0
Amstelveen	1115	6	50,0%	3	0	0	0	1	0	0	0	0	0
Amsterdam	1064	513	51,5%	264	1	0	0	1	0	0	0	0	0
Amsterdam	1067	319	51,7%	165	1	1	0	1	0	1	0	0	0
Amsterdam	1065	352	52,0%	183	1	0	0	1	0	0	0	0	0
Diemen	1111	25	52,0%	13	1	0	0	1	0	0	0	0	0
Amsterdam	1087	1.286	52,8%	679	1	1	0	1	0	1	0	0	0

Figure 28 Amsterdam region postal codes ranked based on the percentage of on-time orders

APPENDIX A11: EXAMPLE OF UTRECHT BOTTOM RANKED POSTAL CODES

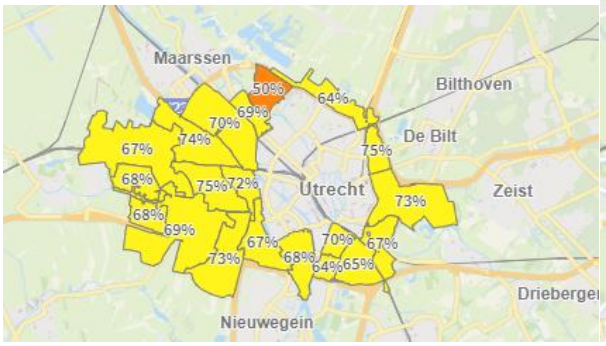


Figure 29 Utrechts postal codes located in the suburbs

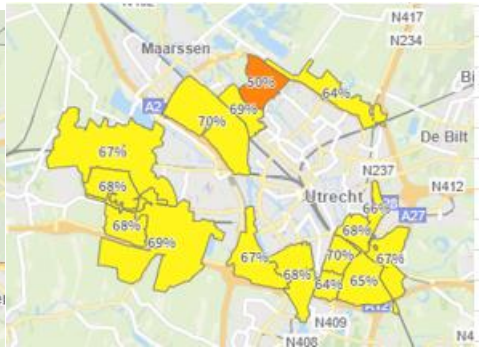


Figure 30 Utrechts bottom ranked postal codes

APPENDIX A12: TIMESLOT PERFORMANCE AMSTERDAM EXAMPLE

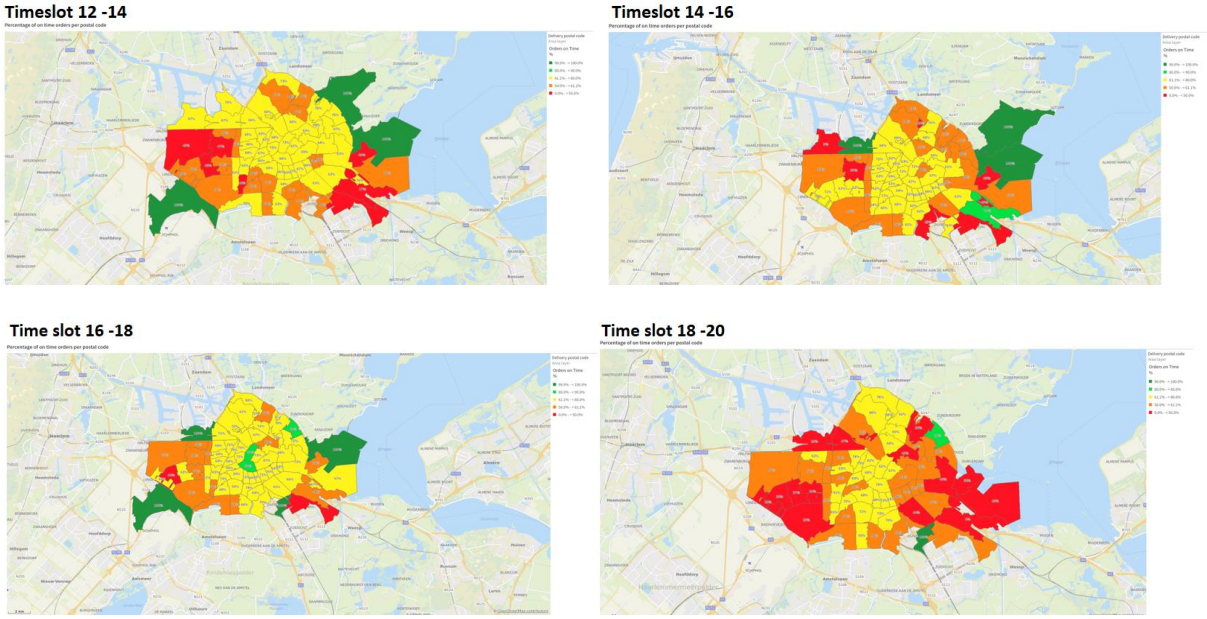


Figure 31 Timeslot performance in Amsterdam

APPENDIX A13: TIME SLOT RANKING ANALYSIS

	12:00:01 - 14:00:00		14:00:01 - 16:00:00		16:00:01 - 18:00:00		18:00:01 - 20:00:00	
Delivery postal code	Orders on Time %	Total orders	Orders on Time %	Total orders	Orders on Time %	Total orders	Orders on Time %	Total orders
1000	-	-	100,0%	1	100,0%	1	-	-
1001	100,0%	1	0,0%	1	-	-	-	-
1011	67,3%	272	60,5%	243	73,4%	154	67,1%	85
1012	73,1%	294	72,2%	187	80,4%	153	62,2%	74
1013	66,5%	528	68,0%	469	70,4%	351	61,1%	321
1014	66,2%	151	64,0%	111	71,4%	70	46,7%	105
1015	68,4%	307	62,9%	299	75,0%	224	62,4%	117
1016	75,3%	381	71,8%	333	74,6%	244	68,5%	92
1017	66,4%	518	67,5%	424	80,5%	307	78,3%	157
1018	68,6%	612	67,4%	522	67,2%	458	52,9%	365
1019	64,0%	577	61,2%	487	67,8%	388	55,8%	514
1020	-	-	-	-	-	-	0,0%	2
1021	63,0%	92	64,6%	82	53,2%	62	45,5%	44
1022	68,2%	44	60,0%	50	75,0%	36	58,3%	24
1023	66,7%	57	50,0%	48	72,2%	36	58,3%	12
1024	76,4%	72	56,8%	37	80,0%	25	81,0%	21
1025	76,5%	85	55,3%	76	70,0%	90	37,5%	40
1026	100,0%	1	100,0%	1	100,0%	1	-	-
1027	100,0%	1	-	-	-	-	-	-
1028	-	-	100,0%	1	-	-	-	-
1031	65,6%	122	55,1%	127	66,7%	105	63,8%	47
1032	60,0%	65	67,9%	81	70,5%	78	65,3%	49
1033	60,8%	166	53,8%	119	62,0%	108	66,1%	62
1034	52,4%	82	75,9%	58	60,7%	56	61,8%	34
1035	72,6%	73	51,2%	41	65,8%	38	76,2%	21
1036	54,3%	46	45,2%	42	70,6%	34	67,6%	34
1042	75,0%	4	-	-	-	-	-	-
1043	66,7%	3	100,0%	2	100,0%	5	28,6%	7
1046	66,7%	6	0,0%	1	-	-	-	-
1051	62,8%	269	62,2%	254	71,4%	217	73,5%	117
1052	61,9%	218	65,4%	208	79,4%	180	75,4%	65
1053	66,8%	286	69,0%	252	68,0%	203	69,3%	75
1054	65,3%	460	63,2%	421	68,8%	314	57,6%	170
1055	65,4%	231	71,6%	285	69,3%	205	61,1%	95
1056	65,6%	247	62,2%	254	63,6%	242	62,7%	134
1057	64,3%	263	63,4%	224	63,7%	204	58,8%	97
1058	63,0%	332	62,3%	318	62,9%	294	61,4%	153
1059	46,0%	161	64,2%	109	57,7%	104	65,2%	46
1060	60,6%	33	75,6%	45	62,9%	62	24,6%	122
1061	59,7%	139	56,0%	125	60,9%	133	54,9%	102
1062	61,9%	160	60,2%	211	66,7%	153	58,7%	109
1063	55,6%	63	58,6%	58	57,6%	59	63,4%	101
1064	47,2%	108	49,6%	129	53,8%	117	55,6%	151
1065	54,2%	72	62,2%	82	57,1%	84	40,2%	102
1066	50,8%	179	55,3%	220	59,8%	194	27,9%	294

Figure 32 Timeslot ranking analysis

APPENDIX A14: STORE DISTRIBUTION PER POSTAL CODE

Store distribution per postal code

Delivery postal code	Store name									
1091	30	12	5	83	20	8	16	6	14	-
1087	32	31	34	35	9	15	6	19	3	-
1071	28	12	65	7	18	20	11	9	7	-
1017	17	10	66	15	25	16	11	11	5	-
1019	39	31	8	27	21	11	19	11	6	-
1078	32	31	32	15	19	19	8	8	4	-
1013	35	23	15	25	13	13	11	9	7	-
1018	26	14	11	23	12	20	19	7	7	-
1079	31	24	9	13	21	8	11	8	9	-
1055	38	20	6	12	11	18	13	2	9	-
1058	35	5	7	12	11	24	18	2	9	-
1016	17	17	42	10	15	6	7	4	3	-
1051	21	24	11	11	15	9	10	7	9	-
1066	26	22	8	17	19	11	7	5	1	-
1098	24	8	20	7	5	19	12	2	15	-
1082	11	46	10	12	4	10	2	5	9	-
1056	13	14	21	2	15	8	16	7	12	-
1062	20	36	7	5	17	6	7	5	5	-
1076	17	24	29	7	2	16	4	3	3	-
1054	19	10	21	6	12	14	7	8	7	-
1095	17	31	4	6	9	8	8	4	9	-

Figure 33 Store distribution per postal code