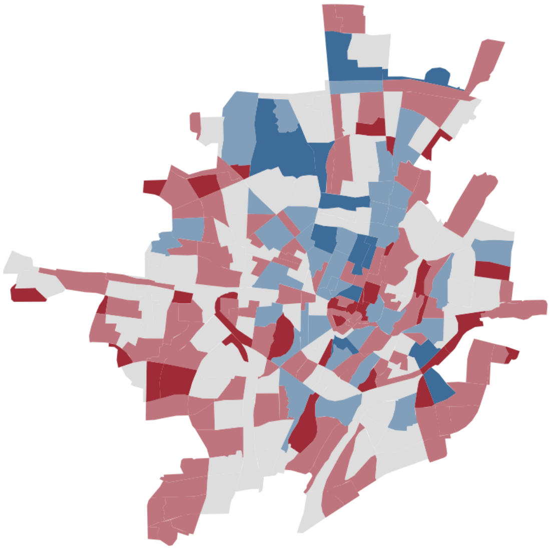
M.Sc. Thesis
Master of Science in Engineering



Transport DTU
Centre for Transport Research

Chair of Transportation Systems Engineering
TUM Department of Civil, Geo and Environmental Engineering
Technical University of Munich

Data-driven time series demand forecasting of Car-Sharing services

The case study of *DriveNow* in Munich, Germany

Shahrom Hosseini Sohi (s161159)

Munich - Kongens Lyngby 2019

DTU Transport
Department of Management Engineering
Technical University of Denmark
In collaboration with
Chair of Transportation Systems Engineering
Ingenieurfakultät Bau Geo Umwelt
Technische Universität München

Denmark,

DTU Management Engineering
Bygningstorvet 116B
2800 Kgs. Lyngby, Denmark
Phone.: +45 4525 1516
E-mail: transport@transport.dtu.dk
http://www.transport.dtu.dk

Germany,

Transportation Systems Engineering
Arcisstraße 21
80333 Munich
Phone.: +49 89 289 22443
Fax: +49 89 289 22333
E-mail: ts-info@vt.bgu.tum.de
https://www.bgu.tum.de/transportation/

Executive Summary

Sharing-vehicle systems are one of the tools that municipalities are trying to apply in order to resolve environmental and traffic issues. The trend of the usage of car-sharing is increasing with high rate of new subscribers each year. Therefore, better performed operations in car-sharing systems lead to a higher acceptance as mode of transport from inhabitants of urban areas. This research proposes the methodology to apply Machine Learning technique of Neural Networks to forecast the Demand of car-sharing vehicles in the case study of Munich. The study analyses the figures and trends of the share-mobility in Munich and investigates on the Business of the DriveNow in Munich. The thesis reveals that major influenced areas within the city (e.g. the Airport) and applies a reclassification of the areas based on district level in order to ease the computation of the model. The results of the Neural Network are compared with traditional ARIMA time-series forecasting describing pros and cons of both techniques.

Preface

This Master thesis was prepared at both the Chair of Transportation Systems Engineering at Technical University of Munich and at the Department of Transport at the Technical University of Denmark in fulfilment of the requirements for acquiring a Master degree in Transport and Logistic.

Munich - Kongens Lyngby, January 17, 2019

Shahrom Hosseini Sohi (s161159)

Acknowledgements

I would first like to thank my thesis supervisors Prof. Dr. Constantinos Antoniou and Dr. Tao Ma from Techincal University of Munich and Prof. Dr. Jeppe Rich from Technical University of Denmark who instructed me wisely throughout this path.

I would also like to thank David Durán Rodas for his wonderful collaboration.The conceptualisation of the thesis work and several interesting talks added precious value to this work thank to his help.

I am also very grateful to: Emmanouil Chaniotakis and Mohamed Abouelela Besheer for their advises. Moreover, I would like to thank my ”greater family” of the people at the Chair of Transportation Systems Engineering: Margit Stötzel, Dr. Christos Katrakazas, Dr. Roja Ezzati Amini, Ralph Harfouche and Moeid Qurashi. I would also like to thank Cat Silva for her precious help in the very last days.

I would thank Tapio Schmidt-Achert for his endless support and being a great friend.

Finally I would like to thank my family, my dad Mahmoud, my mum Clara and my brothers Daniele and Manuel that always encouraged me and sustained since I started this journey.

Contents

Executive Summary **i**

Preface **iii**

Acknowledgements **v**

Contents **vii**

List of Figures **ix**

List of Tables **xi**

1 Introduction **1**
1.1 Problem Statement 1
1.2 Need 2
1.3 Objectives 3
1.4 Report Structure 4

2 Literature review **7**
2.1 Share Mobility and the Car's Triple Revolution 7
2.1.1 Car's Triple Revolution 7
2.1.2 Share Mobility 9
2.2 Types of Sharing Vehicles 10
2.3 Shared Mobility in Munich 11
2.3.1 The overview 11
2.3.2 The Multimodal Mobility Station 12
2.4 Car-Sharing 13
2.4.1 Technology and Operational setting 15
2.4.2 History of Car-Sharing and current figures in Germany 15
2.5 Studies and Research on Car-Sharing 17
2.5.1 Impacts of Car-sharing 17
2.5.2 Modelling Car-Sharing 18
2.5.2.1 Optimisation 18
2.5.2.2 Factors 19
2.5.3 Demand Forecasting 19

3 Data analysis and visualisation **21**
3.1 Structure of the data 21
3.1.1 Data handling 22
3.1.2 Definition of zones 25
3.1.3 DriveNow in Munich 25
3.2 Statistical analysis 27
3.2.1 Temporal trend 27
3.2.2 Spatial trend 30
3.2.3 Strategic Zones 31
3.3 Visualisation 33
3.3.1 Map of Departures and Arrivals 33
3.3.2 ”Specific circumstances” visualisation 34

4 Demand forecast, methodology and model specifications **37**
4.1 Introduction 37
4.2 Aim of prediction 37
4.3 Configuration of Models 38
4.3.1 Neural Network 38
4.3.2 Parameters 39
4.4 Implementation of Neural Network model for DriveNow 41
4.5 Model Specifications 43
4.5.1 Space and time aggregation 47
4.5.2 Variable selection 48
4.5.3 Neural Network settings for the test 50

5 Results Summary **55**

6 Conclusion **59**
6.1 Conclusions 59
6.2 Recommendations and future works 60

Bibliography **61**

A Appendix **66**

List of Figures

1.1 Research framework 6

2.1 Level of automation for AV (credits to Yole.fr) 8
2.2 From Michelberger (2018), ”Future States of mobility” 9
2.3 Multimodal Mobility Station map from Montserrat et al. (2017) 13
2.4 Car-sharing vehicles per 1.000 inhabitants according to system (yellow SB, blue FF) 17

3.1 Time series of departures of DriveNow 23
3.2 Process of data-tidying 24
3.3 Sum of Departures of 2016; Top Left: Top 4 areas, Bottom Left: sum departures based on hour, Right sum of departures based on hour. 24
3.4 Definition of the zones for the study, red colour areas where missing geographical correlation between shape file and data-set 26
3.5 Example of the matrix of efficiency of the fleet 27
3.6 Departures by day-hour of the week 29
3.7 Trend daily in 2016 - particular on the Oktoberfest 30
3.8 Departures on Munich based on time-space 34
3.9 Arrivals on Munich based on time-space 35
3.10 Oktoberfest bookings 36
3.11 Average Departures 18:00-23:00 36
3.12 Departures on 22/07/16 between 18:00-23:00 36
3.13 Particular in the Olympiaeinkaufzentrum area 36

4.1 The perceptrons elemental structure. 39
4.2 Multi-layers Neural Networks a) without bias b) with bias 41
4.3 Example of spatio-temporal Neural Network 42
4.4 Flow chart in model specification 44
4.5 Departures according to the time of the day - DriveNow, 2016 45
4.6 Box-plot, Average hourly-daily departures of DriveNow in 2016 46
4.7 Zone Configuration based on parking regulation 47
4.8 New Zone Aggregation based on districts (in German Stadtbezirke) . . . 48
4.9 Process of space and time aggregation 49
4.10 Configuration of the Input/Output of the Neural Network 50

4.11 Sensitivity Analysis Neurons 51
4.12 Sensitivity Analysis Time 52
4.13 Proportion data-set between the training, testing and forecasting 54

5.1 Representation of the Model created, larger resolution in Figure A.4 . . . 56

A.3 Tails of the chart for the Balance of the areas 67
A.1 Box-plot departures-hourly-monthly-weekly 68
A.2 Leer values outliers 69
A.4 Representation of the Model created 82

List of Tables

2.1 List of share mobility services in the Bavarian Capital 12
2.2 Car-sharing Vehicles per 1.000 inhabitants 16

3.1 Sample of first configuration of the data-set 22
3.2 Figures of Number of departures and arrivals in for every month of 2016 . 28
3.3 Departures on day of the week 28
3.4 Linear trend model 30
3.5 First 13 Areas with higher concentration of Departure and Arrivals 32

4.1 Normalisation effects on computational time 49
4.2 Multi-layers test 52
4.3 K-fold cross validation with configuration of 56 neurons 53

5.1 Summary Results Working Days Neural Network model 55
5.2 Summary Results Saturday Neural Network model 55

A.1 Figures of departures and arrivals, classification of the areas 70
A.1 Figures of departures and arrivals, classification of the areas 71
A.1 Figures of departures and arrivals, classification of the areas 72
A.1 Figures of departures and arrivals, classification of the areas 73
A.1 Figures of departures and arrivals, classification of the areas 74
A.1 Figures of departures and arrivals, classification of the areas 75
A.1 Figures of departures and arrivals, classification of the areas 76
A.1 Figures of departures and arrivals, classification of the areas 77
A.1 Figures of departures and arrivals, classification of the areas 78
A.1 Figures of departures and arrivals, classification of the areas 79
A.1 Figures of departures and arrivals, classification of the areas 80
A.2 Sensitivity Analysis number of neurons Neural Network with single layer . 81

CHAPTER 1

Introduction

1.1 Problem Statement

The United Nation Department of Economic and Social Affairs has announced that the growth of the world's population has reached 7.6 billions in the mid-2017. The expansion of the population is expected at least until 2050. The total amount of people will grow from 8.4 to 8.7 billion in 2030, between 9.4 and 10.2 billion in 2050 and reaching the record between 9.6 and 13.2 billion in 2100 with a major concentration into urban areas (United Nations Department of Economic and Social Affairs, 2018). With this increment, also the distribution of the population changes. While in the 1950, there was 30 % of people living in urban areas, nowadays it is 55 %. By 2050, it is expected 68 % of the world's population in urban settlements (United Nations Department of Economic and Social Affairs, 2018). Furthermore, the concentration of the people living in urban areas has reached the 55% of the world's population in 2018. The figure In 1950, was only the 30 % and by 2050, 68% will be concentrated in the urban settlements (United Nations Department of Economic and Social Affairs, 2018). As consequence to this trend, an important global effect due to this enlargement is the major increase of demand of mobility and transport of people and goods.

The overall transportation sector share of global emission amounts to 17.5% (International Energy Agency, 2017) and among 2010 and 2015 transport emissions grew by 2.5% per year. Recent international agreements such as the "Paris Agreement", are trying to delimit warming to an average of two degrees Celsius (2°C) and targeting the 1.5 °C. The main objective is to strongly introduce low-carbon policies in the transport field by 2050 and communities are called to take part to this global commitment. Along with the environmental challenges transportation in the last decades is facing the rise of new issues: an increased level of congestion that leads to a greater idle-time for travellers, lack of parking space, air and noise pollution and these effect will be directly increase with the population growth. The public sector in particular, is taking action with different solutions such as congestion charge, enhancing cycling and use of public transport.

On the other hand, the car business industry and private sector are taking into account the incoming transformations and how to redesign the concept of mobility itself. Economically speaking car industry has been - and still - is the driving power of many national economies and the new demand of products (and mobility) must align the global necessities and discover the new trends in the industry. Moreover the industry is facing another challenge: Driver-licences issued are decreasing, suggesting that the value of the ownership of a car is decreasing for the new young-adults. The causes are both social and economical, the status symbol of the car does not affect people like in the past and the assumption of owning a private car is gradually losing importance. (Delbosc et al., 2013). Furthermore, the car represents one of the most underused object that people use. According to Cohen et al. (2008) cars are parked between the 93.5% and 95%, per day, drivers spend on average 73 minutes per day. These major economic, environmental, and social aspects pave the way to the new sustainable "sharing economies" and "sharing mobility", where the ends of sharing resource, saving money is allowed (or justified) by the mean of the new technologies (Cohen and Shaheen, 2016). Therefore, mobility-sharing providers (where sometimes are the same vehicle-producer themselves under a different name) are arising in all major cities worldwide intercepting the trend of people that are willing to renounce to a self-owned vehicle, or a single mode of transport in order to share the usage and the burden of the ownership.

Last but not least, the advance level of mobile technology has innovated the way that people get access to transport information. People can get instantly check the time-table of the public transport, the traffic situation on the way, the availability of sharing mobility service or call a taxi with the same tool comparing the cost, travel time and comfort at the same exact time. Furthermore this technological progress has brought about an immense and precise user information as well. Therefore, the increasing calculation capacity of modern computers with the very vast and detailed amount of data generated (in a very short time) by people opened to new different analysis techniques. Based on these three characteristics of the data generated velocity, veracity and value (Khan and Machemehl, 2017) more and more data-driven tools are developed and more precised and deeper understanding about transportation the research has reached. The new "Data Era modelling" contributes to better understanding of mobility patterns and people transport modal choices. Indeed, it contributes to create a detailed transport offer tailored to most precise need of the user even before they are actively looking for it.

1.2 Need

The need for institutions and policy makers is to develop better transportation systems in order to tackle the well-known problems such as overpopulation, pollution, congestion and limited land space use. On the other hand, the need for the industry is to best integrate new data-driven strategies to offer solutions for private-transport (owning a vehicle or rent it). The number of mobility providers in each city is rapidly growing for each city, offering often the same public/private duplicate service. Nowadays, a fundamental need for these companies is to establish their assessment in the market and gain the biggest share of it since these modern sharing-services are quite new mobility services. The major evidence is portrait in the bike-sharing systems (chapter 2) where they are strongly competing each other by overwhelming the cities with extra offer in order to push out the competitors (Haas, 2017). These recent form of shared mobility defined in the chapter 2 has still an undergoing research in terms of what the outcome will be for the society. Therefore, the companies need to better understand these concepts demanded by further mobility improvements that are arising in the upcoming future in the autonomous vehicles. The next level of automation (Level 3) that car industry is facing on the horizon of 2025 will deal with the automatic pilot (Yole, 2015). This revolution will probably cut down driver/pilot costs making cheaper and economically more accessible the new forms of transportation such as ride-sharing and car-sharing.

A further business-research need is to create an integration of data-driven strategies according to the level of detail of the information in possession of; gather self-knowledge through the data generated through the new technologies. The importance is fundamental to combine methods and tools for the planning, scheduling and operational transport procedures with new algorithms that can extract meaningful understandings and visions that are hard for human to analyse and infer. (Ma and Antoniou, 2018). Lastly, as granted the benefits of these techniques the ultimate need of this research is to adapt modelling and utilisation of learning methods according to the level of detail of the data at the disposal of the practitioners and to the detail of the forecast.

1.3 Objectives

The objectives of this research aim to improve the data-driven methodologies in the sharing-mobility businesses, in specific it has considered the car-sharing systems. The case study has dealt with the scenario of car-sharing in Munich by the courtesy of DriveNow, the car-sharing company from BMW Group, who has shared the data (more details about the data-set described in chapter 3) for the rentals of 2016 with the Chair of Transportation Systems of the Technical University of Munich. The research explore the forecast analysis providing an holistic understanding of the context as complete as possible. The thesis starts broad with the data analysis and narrows down to the forecast task. The study wants to uphold the research on shared mobility, evaluates the data-driven methodology with cutting-edge techniques for the forecast of the vehicles rentals and compare with traditional methods (described in chapter 4 and in chapter 5) in a time series scenario using a Machine Learning approach. Accurate forecast, adds extremely importance to sharing mobility services, the more precise is the prediction the more tailored is the supply which generates a better revenue for the provider. Moreover, a proper forecast enhance the transport system as a whole when the network of sharing vehicles reflects as best the mobility demand, the reliability as a mode of transport increases as well. Therefore, the users can rely as principal transportation systems so the system contributes to resolve mobility problems of the cities. Lastly the literature reviewed and to the best of the author knowledge has investigated on the usage of Neural Networks model for forecasting car-sharing services. However, the literature present a gap of using this technique in this field. According to Ferrero et al. (2018), who have introduced a taxonomy about car-sharing, classifying 137 different papers and covering the last fifteen years of studies on this, Neural Networks are a tool not largely used in the prediction of car-sharing demand. Hence, the following statement introduces the research question of this study that want to fill the gap:

1. **Can Neural Network, deep learning technique be used under time series scenario?**

2. **In what circumstances Neural Network be used and how better is it than traditional time series forecasting?**

1.4 Report Structure

This report has been structured according to the following research framework showed in Figure 1.1: The top blocks of the diagram (Need - Objectives and Literature review) correspond with Chapter 1 and Chapter 2 of the thesis work. The first has emphasised the reasons why are important to research into sharing-mobility as a potential solution for current and future transportation issues. Then, thanks to the availability of data, the need to integrate data-driven strategies into the planning and operational sphere of new mobility services. The second has dealt with a broad research on the stat-of-art of shared mobility, narrowing down to the current situation of car-sharing on the national and municipal level. The last part of literature review focuses on research on car-sharing systems, it has identified four major areas where practitioners are focused Deployment, Optimisation, Schedule and Demand forecast. This segmentation reflects the four dominant research topics and principal operational challenges for the car-sharing providers.

The central part of the diagram corresponds to Chapter 3 only. The section reports a large data analysis, data handling and data visualisation. The information have been processed and reshaped accordingly to best practice of data-analysis. The process familiarise with the case study, reveals insights about the DriveNow in Munich and pave the way to the data structure for the data to feed the modelling section.

The bottom of the diagram matches the Chapter 4, Chapter 5 and Chapter 6. The Chapter 4 introduces the concept of Neural Network and describes the methodology applied to the case study. Chapter 5 includes the results of the prediction for the business zone of DriveNow in Munich, a second data analysis is described with the new outcomes of the forecast. Moreover, Chapter 5 evaluates the quality of the results benchmarks it with a different time-series model. At last Chapter 6, answers to the research questions and provides recommendation for future works.

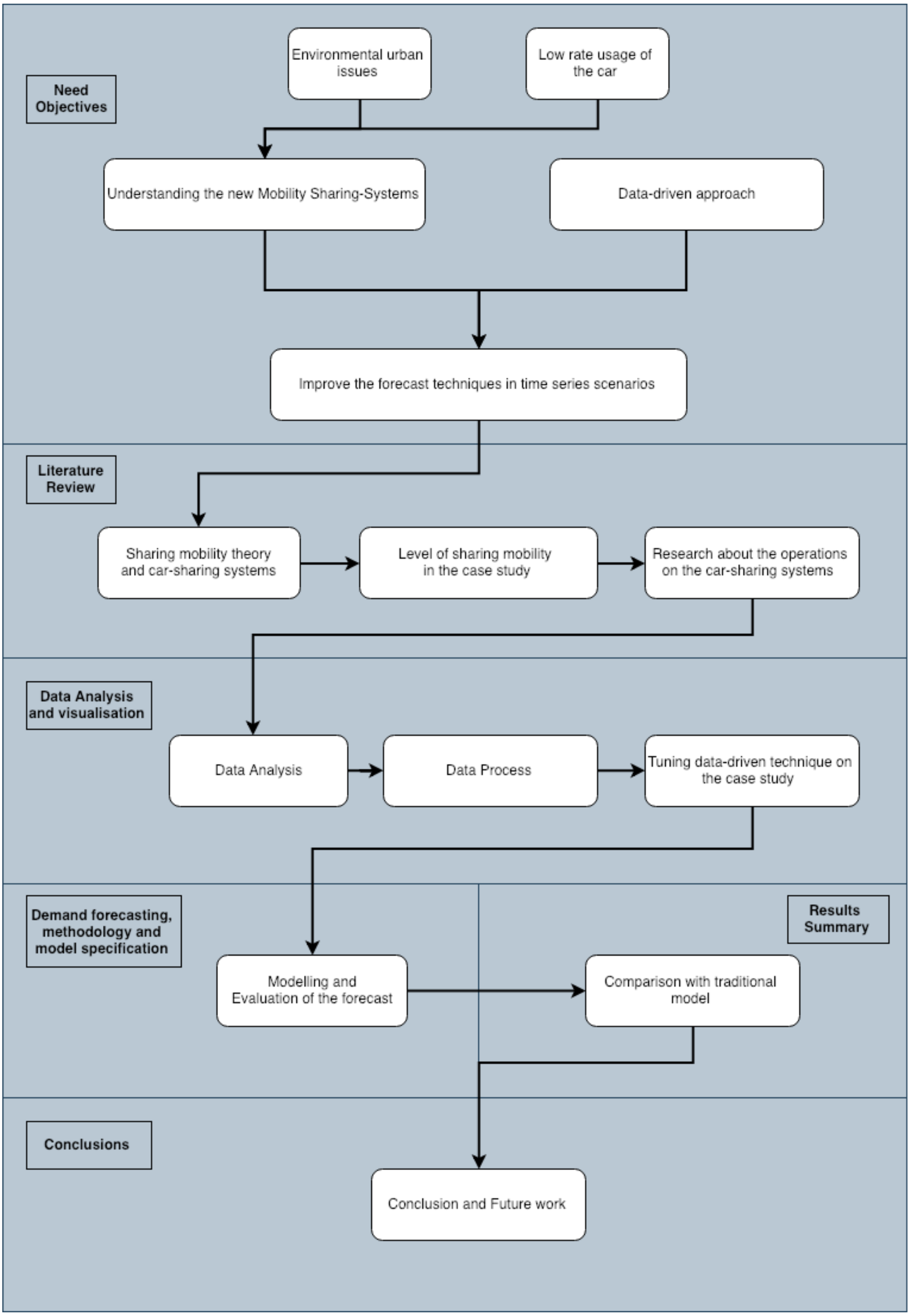


Figure 1.1: Research framework.

CHAPTER 2

Literature review

2.1 Share Mobility and the Car's Triple Revolution

2.1.1 Car's Triple Revolution

The automotive industry is renewing itself under a triple revolution, three major innovations are changing this sector at the same time: a broad electrification of the vehicle engines (EV), driver-less vehicles as known as Autonomous Vehicles (AV) and the introduction of sharing mobility systems. This transformation redesigns the driver model, how the vehicle is controlled and the perception of ownership and the fuelling systems of the cars. The growth of electric vehicles sales has reached 1.6 million this year compared with the number back in 2012 was only 122 thousand (Mckerracher, 2018). The reflection of the market growth is coherent to the 79% drop on the cost of Li-Ion batteries in 6 years (Mckerracher, 2018). Moreover, the predictions say that the cost to make EV compared to the traditional internal combustion engine will be equalised in the 2030. By 2040 the 55% of all new car sales and 33% of the global fleet will be electric. (Mckerracher, 2018).
At the same time, narrowing emission limits in order to increase the urban air quality (such as the possibility of diesel ban in Europe), is another leading factor that require significant electrification of the vehicle fleet. Accordingly, car industries are pushing out an extensive number of EVs models and this number will double up in five years time from 155 at the end of 2017 to 289 different models by 2022 (Mckerracher, 2018).

The next level of innovation is the autonomous driving. Autonomous Vehicles (AV) are supposed to replace drivers within the next twenty years. These are six principal levels of automation that characterise cars (Yole, 2015) :

1. "Level 0, no Autonomy" the driver must command the vehicle in every aspect.

2. "Level 1", launched the Active Cruise Control and the Lane Departure Warning System.

3. "Level 2", introduced until the 2012 the Parking Assistant, and Lane Keep Assistant technology.

4. ”Level 3”, corresponds to the today level it will last until 2022 and it deals with Automatic emergency braking system, Drive monitoring and Traffic Jam assistant. The car can drive without hands on the steering wheel.

5. ”Level 4”, will implement the sensor system, the car fully ”see” by itself.

6. ”Level 5”, the full automation, the car can be autonomously driven in every possible scenario.

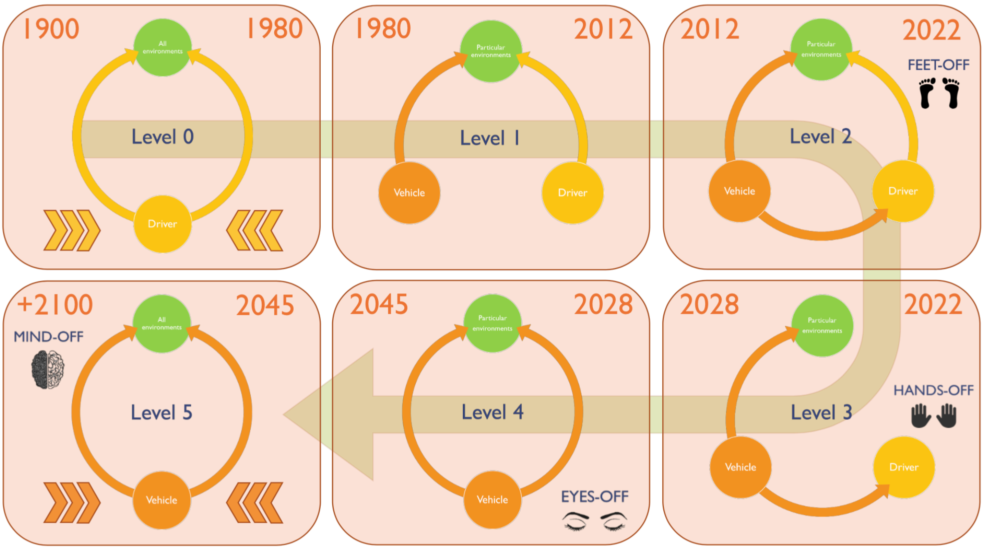


Figure 2.1: Level of automation for AV (credits to Yole.fr).

The last axis of the triple car’s revolution is represented by the degree of ownership of the vehicle. Sharing systems are becoming more and more widespread and popular, share mobility is evolving, innovating the concept of ownership a car (Shaheen et al., 2010). The combination of the three degrees of innovations will generate a completely new concept of mobility: the Autonomous Mobility-on-Demand (AMoD), where cars driven by themselves that do not belong to a single owner, providing mobility services on-demand (Iglesias et al., 2017).

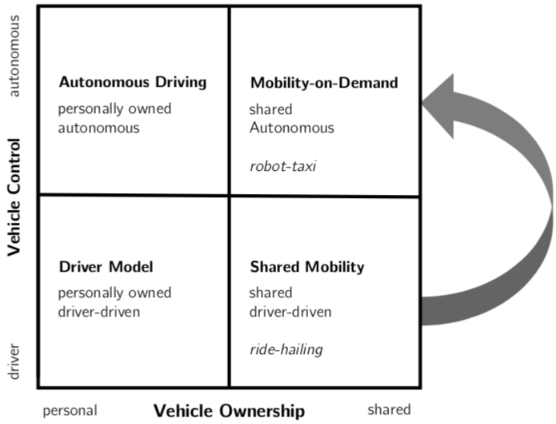


Figure 2.2: From Michelberger (2018), "Future States of mobility" .

2.1.2 Share Mobility

The rise of Share Economy (or Sharing Economy) has implemented perfectly the paradigm that emphasise the access rather than the ownership. This economic system creates services through a central provider that empowers the owners of goods or services to be connected and "share" satisfying others users' needs(Martin, 2016). Although, it has several definitions and a big discourse among the researchers, the Share Economy has its rudiments in the accessibility of the service or goods through a platform. These systems have a structure where users share, swap, barter, trade, gift and rent to get the equal benefits of a normal ownership thanks to the current development of technologies and easiness to access the web.(Martin, 2016) Inside this "umbrella term" several new business models can be found under this definition: *the Rental economy*, based on a company's rental schemes of goods that generally users have a low rate of utilisation, it avoids the burden of the ownership (Car-Sharing). *Peer-to-peer economy*, model of economy similar to Rental but offers the good or service directly by the owner(AirBnB). *On-demand economy*, platforms-based that let the access to services (Uber, Blablacar or TaskRabbit). *Time banking and local exchange trading system*, lack of money in the transactions, focus on barter as a surrogate payment or time spent as value of service (TimeRepublik). *Open source*, can be considered as the pioneer of the Share Economy, generally the software/code is available to anybody and to any purpose. The development of such a service/good encourage the open-collaborative creation of it(Linux). *Social lending and crowdfunding*, platforms where people can help to reach the financial assets necessary for the development of a new idea between people that are keen on the project (Kick starter) (Selloni, 2017).

Share Economy models have given birth to the movement of Sharing Mobility, the transportation strategy that implement the usage of a shared vehicle in order to fulfil a mobility demand (Shaheen et al., 2015). Likewise the other businesses, the transportation field is turning the focus towards the concept of "Mobility as a Service" (MaaS), that changes from a personally-owned system to travel-need one in order to better accommodate the personal desire.
This approach does rely on a multitude mode of transport and according to Willing et al. (2017), nowadays, timetables information of transportation services are more accessible thanks to the progress in IT and software solutions so that intermodal solutions are more attracting and simpler for users. MaaS enables users to get a convenient access to different modal choice under one single provider, an integrated transport systems that let the passenger to choose when and how to get from A to B in a optimum way. The service included in the MaaS involves 1) on-demand vehicles, 2) car-sharing services , 3) bike sharing services and 4) taxi services. (Martin, 2016)

2.2 Types of Sharing Vehicles

The configuration of the rental of a sharing vehicle is divided between Station-Based (SB), Free-Floating (FF) or a mix of SB and FF and they represent the business model which the operator wants to offer the service to the end user. The Station-Based setting, force the user to start and end the trip in a defined area such as a specific car park or a very restricted geographic area. Inside this classification, there are the "one way trip", the end of the rental can be done at the destination station and the round trip category where the start and the end must be at the same zone. The Free-Floating scheme instead lets the user the "freedom" to move, start and end the trip in the whole business zone, a set of areas where the provider has decided to offer the service according to different parameters (Ferrero et al., 2018). The type of sharing vehicles are (Shaheen et al., 2015):

- Car-Sharing
- Bike-Sharing
- Scooter-Sharing
- Ride-Sharing
- Alternative transit services; shuttles
- Courier network services: P2P delivery services
- On-Demand ride services: Ride-sourcing, Ride-splitting, Ride-pooling

The car-sharing systems are differentiated by the owner of the fleet: Private/Public owner and Peer-to-Peer(P2P). The first, uses a personal fleet from an public or private

organisation and let the vehicle be rented for a short-term with the benefits of having a private vehicle without paying the cost of ownership (Shaheen and Cohen, 2012). The P2P instead, uses the people personal vehicles as the object of share. In the next section, it has described specifically this type of Share Mobility. The Bike and Scooter Sharing, has the same approach as the Car-Sharing in terms of Ownership-Access it differs only whether the services are FF or SB. The Ride-Sharing, offers the possibility of share a place in the car in a common trip (Durán Rodas and Constantinos Antoniou Emmanouil Chaniotakis, 2017) such as Carpooling.
Shuttles connect the main network of public service by sharing the vehicle. Courier network services are the service that set up the delivery through private owned vehicles. (Durán Rodas and Constantinos Antoniou Emmanouil Chaniotakis, 2017) Last but not least, On-Demand ride services (or Ride-Hailing or Real-time Ride-Sharing), such as Uber and Lyft belong to group, where the user specifies where and when use the service and it takes them directly to their destinations in a professional driving setting. (Amey et al., 2011).

2.3 Shared Mobility in Munich

2.3.1 The overview

Munich is the capital of the Free State of Bavaria, it is the third largest city (after Berlin and Hamburg) in Germany with a population reaching nearly 1.5 million people, it is the most densely populated with 4,668 people per square kilometre Bunderinstitut für Bau- (2015). The public transport network is composed by 95 km of subway, 75 km of tramway, 456 km of bus lines and 442 km of commuter train (S-Bahn) lines connects the city with the metropolitan region (Montserrat et al., 2017). The rate of car ownership is 432 vehicles/1000 inhabitants lower than the national average of 555 vehicles/1000 inhabitants (Priester and Wulfhorst, 2014) (Eurostat, 2016). The population of the city has grown by 16,5% in the last fifteen years and it has stressed particularly the real estate market and the transportation network (Montserrat et al., 2017).
In Munich, the major operators offering sharing mobility services are summarised in the Table 2.1: The city hosts a broad spectrum of share mobility systems, both FF and SB of: 6 motor vehicles-sharing and 5 bike-sharing. Regarding Ride-sharing and Ride-pooling services 5 companies operate in Munich with many others are planning to get into the market soon (Lyft Team, 2018). The ownership of the fleet of sharing vehicles are both private and public. It bears mentioning that Munich's public mobility companies are looking with very interest in this new market. Indeed, each shared mobility category has its share in the market run by a public authority provider. For example, "Flinkster", "Call-a-bike", and "CleverShuttle" are the SB car-sharing, FF bike-sharing and Ride-pooling providers owned by Deutsche Bahn (DB), the German railway operator. Moreover, the council authority has its own share mobility systems:

Table 2.1: List of share mobility services in the Bavarian Capital.

Name	System	Trip	Owner	Vehicles	Price min
DriveNow	FFCS	One-way	BMW	750	0,30-0,36 €
Car2go	FFCS	One-way	Daimler	550	0,26-0,33 €
Flinkster	SBCS	Return	DB	135	not applicable
Oply	SBCS	Return	Oply	200	not applicable
StattAuto	SBCS	Return	StattAuto	450	not applicable
Emmy	FFSS	One-way	Emmy & GreenCity	-	0,19 €
MVG Rad	FF-SBBS	One-way	MVG	3200	0,08 €
Call a Bike	FFBS	One-way	DB	1200	0,03 € (based on 30 mins)
OBike	FFBS	One-way	OBike	6800	0,03 €
Donkey Republic	SBBS	One-way	Donkey Republic	-	0,10€ the first 2 hrs
Uber	RP	-	Uber	-	0,30-0,50 € (plus km plus extra fee)
Clever Shuttle	RP	-	DB	-	not applicable
IsarTiger	RS	-	MVG	-	not applicable
Filnc	RP	-	Daimler	-	not applicable

”StattAuto”, ”MVG Rad” and the most recent ”IsarTiger” are the SB car-sharing, FF-SB bike-sharing and Dynamic On-Demand Ride-sharing Services.

2.3.2 The Multimodal Mobility Station

The Multimodal Mobility Station is a project run by City council, the Munich city utilities and the city transport provider, where it offers sustainable mobility transport modes renouncing the own car (MVG, 2015). The station is located in ”Münchener Freiheit” in the north of the Scwhabing neighbourhood on the main road Leopoldstraße and offers access to: 1) Metro (Line U3 and U6), 2) Tram (Line 23), 3) Bus (Line 53, 54, 59 and night bus N43 and N44), 4) Car-sharing with dedicated parking spots (StattAuto, DriveNow, Car2go), 5) Bike-sharing (MVG Rad station), 6) Taxi station.

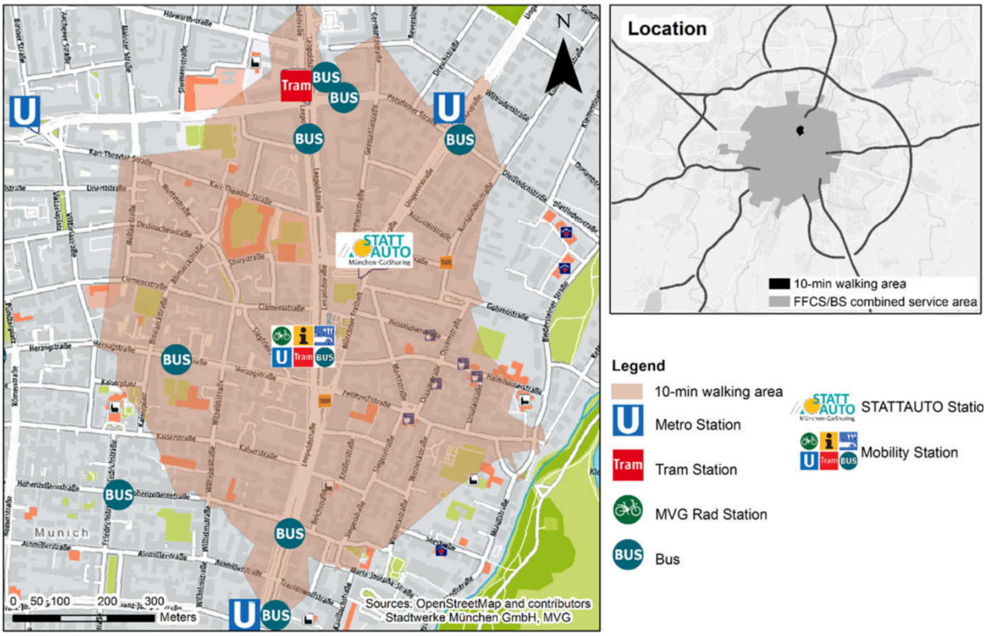


Figure 2.3: Multimodal Mobility Station map from Montserrat et al. (2017).

Montserrat et al. (2017) have evaluated through a survey the impact of the Multimodal Mobility Station. The study has revealed that utilisation of Multimodal Mobility Stations increases the integration of public transport and intermodal solution between FFCS and BS and viceversa and 75% of the interviewed people agree on that this model helps to avoid the usage of private car (Montserrat et al., 2017).

2.4 Car-Sharing

The Car-Sharing is the short-term rental of a car that lets the end user to access to a private or public car fleet. The costumer is generally required to pay a subscription fee to sign up to the car club and the cost of the trips are charged by: a) type of vehicle, b) kilometre, c) time, d) special geographical locations (special price for special locations such as airports and big centres of demand outside the business zones) or by a mix of them (Shaheen et al., 2014). The idea of car-sharing let user benefit nearly as same as private automobile without the burden of an owned personal car Shaheen and Cohen (2012). The cost of FF systems is mainly fixed on the price per minute. Within this price, the service includes a quote of variable costs such as fuel, price per kilometre and the share of fixed costs for the operator like maintenance, re-balancing, insurance and parking. The calculation of the cost of SB systems instead is split between hourly, fuel consumption, and cost per kilometre [1] and the minimum rental time is 1 hour. On the top of hourly cost, SB systems add kilometre rate-cost and fuel consumption separately. Both systems charge different prices in accordance of model selected. Car-sharing operates under geographical zone called "business zone" is common that this area represents urban areas with high density (Shaheen et al., 1998). The definition of business areas for the car-sharing providers differs between the share-mobility structure. The business areas for the SB systems are constrained on origin of the trip. The "return" SB systems force to return the vehicle in station, the same where the start of the rental has begun. These areas are spread in normal parking spot around the business zone or specific car park designed for the service (Bundesverband CarSharing eV, 2018). FF systems, in contrast, consider the business zone as a whole, so rentals can be started and ended in each point within the continuity of the boundaries. However, the business area can include very circumscribed, limited zones called "Satellites". These small portion of areas are strategically points of interest that gather high demand but are detached by the continuity of the city confines. A well-known example are Airports, usually they are located very far away from the city, but are crucial points to be served by any mean of transportation. Another case of "satellites" can be found in dense business districts with high concentration of companies or university campuses dislocated from the city centre.

The freedom to drop the car at any location of the FF systems introduce one of major operational task of every vehicles in the system: *the re-balancing problem*. FF systems inside the business area during the day generate sub-areas called *hot-spots* and *cold-spots*. Hot-spots are the area with high concentration of departures and arrivals where the intensity of the flow of vehicles that are arriving and departing is high. In contrast, cold-spots are representing areas concentrate high amount arrivals and low departures. Therefore the operator must take actions to move part of the fleet and relocate in areas with future concentration of demand (Schmöller et al., 2015). An alternative strategy for operators is represented by creating attractive incentives to user such as discounts on the cost per minute for the specific car in the specific time at the specific location in order to "move" the car from the Cold-spot. The same technique is also used for the operations of refuelling cars that have low level of fuel/power. The re-balancing problem is not present (for return systems, the one-way SB still have it) in SB systems because each "station" has its own balance allocated beforehand by the operator. Indeed, the system of deployment of SB car-sharing focus on relative small area and the balance is kept fixed due to the constraint of returning the vehicle where the trip has started. However, both systems must take into account the maintenance tasks of the fleet and the refuelling of the vehicles.
Both systems present advantages and disadvantages for users and operators according to different business models. For users where high degree of freedom is required, in the FFCS this asset is compensated by generally more expensive service compared

[1] Some SB provider have costs include an initial limit of kilometre included and charge the rate above limit (e.g 6€ per hour and 150km included)

to the hourly rate of SBCS and the parking space is not guaranteed which can cause more time spent in the vehicle (and higher cost). For operators the FFCS is a more attractive business model, but requires high density area for the deployment and has high cost of re-balancing since each car require a person that relocates it accordingly to demand. SBCS are cheaper to implement, do not require very high density areas for the operators, but limits the user to make a return trip. Therefore, the flexibility of the trip is one of the most important driven force for the transport modal choice. Where the purpose of the trip is commuting to workplace or one-way trip, FF represents more suitable systems, but regarding leisure trips, SB can offer more economical solutions (Kopp et al., 2015).

2.4.1 Technology and Operational setting

In order to access to shared vehicles the following tools can be used: with card-RFID, with smart-phone application (both android and iOS) or with hot-line. Every vehicle is equipped with a RFID reader on the windscreen with 3 colours LED which communicates the availability and status of the rent. To open the car start the rental and end it, it is only necessary to tap-in/tap-out the card on the reader and wait for the confirmation signals and insert the personal password or code in the car-panel. Through the smart-phone the geographical position of the user is gathered from the phone and the app displays real-time information on availability and damage status of the car. The user can access to several information: the model of the car, the fuel status and the cost per minute/hour. In addition, the smart-phone's application can display offers and filters where the user can set up different parameters for the rental such as which models of the car to display in the area (Guijarro, 2015). Moreover, most of the applications have a "radar" feature, this function sets up the search at an input zone and given time within a time-frame of a week ahead, then, when the car is available the application sends the notification to the user's smart-phone. Before start the rental within the application, is also possible to check the physical status of the vehicle whether external damages are already recorded or to be reported. The application let the user to unlock and lock the car and to select the destination beforehand and it will be automatically set up on the vehicle's navigator. When the rental is over, the central system verifies that the vehicle is inside the business zone and gives the confirmation signal of the end of the rental.

2.4.2 History of Car-Sharing and current figures in Germany

The first evidence of car-sharing system is in the 1948 in Zurich, Switzerland, when the small community of "Selbstfahrergemeinschaft" started to share their vehicles in order to help people that cannot afford the price of car. The project failed because of too many requests and overbooking in the process of rent a car was too complicated (Shaheen et al., 1998). Right now, the two oldest and largest car-sharing organisations that still are operating are: Mobility Car-sharing Switzerland (since 1987) and Stat-

tAuto Berlin (Since 1988, Now Greenwheels). Another example of early Car-Sharing system was done by Lufthansa in the 1993. It had implemented an automatic rental systems for its employees between the airports of Frankfurt and Munich. Indeed, the German Airline had developed a computer system that released the key and takes into account the kilometre and fuel consumption of the cars, and saved over 20 million of dollars (Shaheen et al., 1998). In the more recent history, in 2008 the city of Ulm in Germany introduced the first fleet from the provider Car2go launching the first Free-Floating service that is known currently. In Germany in 2018 there are more than 2 million people that use car-sharing systems (Bundesverband CarSharing eV, 2018). The German Car-sharing association (Bundesverband CarSharing) has calculated that regarding the Station Based car-sharing services gained in 2017 80.000 new customers reaching 535.000 overall (increase of 17,6 %). The Free-floating figure increased by 315.000 in the 2017 (increase of 25%). The number of shared cars in Germany counts 10.050 for SBCS in 161 companies and 7.900 for FFCS in 4 companies (Bundesverband CarSharing eV, 2018). The German car-sharing services are covering 677 cities and municipalities (80 more than 2016) most of them provide SBCS while FFCS services operate in 12 cities in Germany. This insight show how the density of population and number of inhabitants are key factor for the deployment of FFCS services. The FFCS is the system that is growing at faster pace and mostly popular. Although the number of vehicles are not tremendously increasing (only 100 more vehicles among 12 cities between the 2016 and 2017) in the last year, the popularity among this service is boosting. Therefore, the ratio vehicle-user has got an interesting growth. For the year 2017 the proportion vehicle-customer has grown from 173 to 215 persons-per-vehicle. The SBCS has a lower ratio with 53 persons-per-vehicle (Bundesverband CarSharing eV, 2018).
Another intriguing aspect of the Car-sharing systems in Germany is that the 10,3% of the CS fleet is composed by electric vehicles, 100% above the national rate (0,1%) (Bundesverband CarSharing eV, 2018). In the rank of the German cities of with more car-sharing vehicles per 1.000 inhabitants is summarised in Table 2.2 and Figure 2.4

Table 2.2: Car-sharing Vehicles per 1.000 inhabitants.

City	**Car-sharing vehicles**
Karlsruhe	2,71
Stuttgart	1,47
Freiburg	1,41
Köln	1,27
Heidelberg	1,27
Munich	1,26
Göttingen	1,25
Frankfurt a. M.	1,16
Tübingen	1,10
Hamburg	0,94

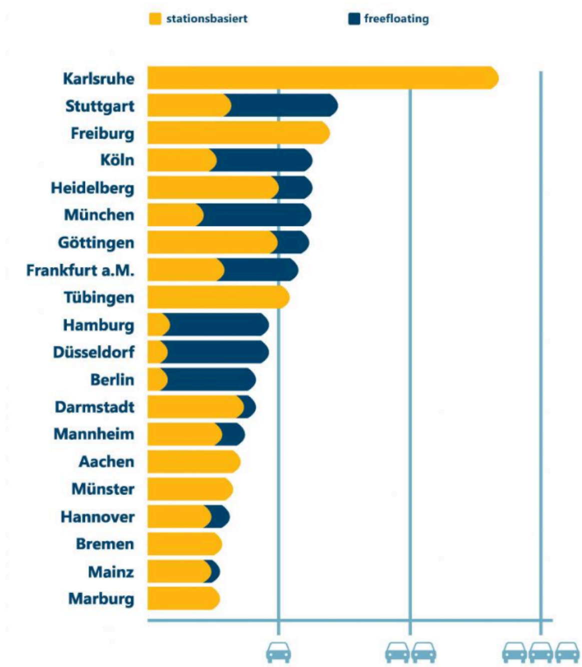


Figure 2.4: Car-sharing vehicles per 1.000 inhabitants according to system (yellow SB, blue FF).

2.5 Studies and Research on Car-Sharing

2.5.1 Impacts of Car-sharing

This section analyses the impact of car-sharing found in the literature, dividing them in social, environmental and economic. The major aim of car-sharing is to reduce the car ownership (Firnkorn and Müller, 2011). The reduction of vehicle ownership leads to positive impacts such as efficient use of road and infrastructure, reduction of the vehicle kilometres per person and economical saving for the users (Shaheen and Cohen, 2012). In addition, Giesel and Nobis (2016) highlight that SBCS systems have a bigger impact (15%) in the reduction of the vehicle ownership rather than FFCS (7%). This generates broad positive impacts: the decrements of vehicle generated emission, with less car ownership less traffic congestions and more availability for parking with a better land use. Many studies revealed that car-sharing users are generally male young adults between 25-40 years of age, high-educated, with low-average income due to the young age and living in small households without children

(Shaheen and Cohen (2012) and Prieto et al. (2017)). Although the usage of share-mobility systems generate fewer vehicle kilometre travelled by mono-use cars, the side effect, on the other hand, could incentive car purchase for non-car owners (Giesel and Nobis, 2016). There have been in-depth studies on the effect of SB car-sharing systems. To begin with, Giesel and Nobis (2016), has summarised the positive effects on car-ownership by SB systems. Bundesverband CarSharing eV (2018) calculated the break-even point for who drive less than 10.000 kilometres per year.

2.5.2 Modelling Car-Sharing

This section focuses on the research on modelling car-sharing systems. Before going into the detail of the literature it bears mentioning the major tasks - challenges for car-sharing operators that are focus of the research as well:

Optimisation Car-sharing providers must optimise their resources. The object of the optimisation in car-sharing providers is to minimise the trips on loss. One well-known challenge for FF systems is to re-balance the fleet in order to have a good distribution of the vehicles accordingly to the demand. The vehicles are then relocated and moved from "cold-spots" avoiding to have them idly parked in areas that have less demand Boldrini and Bruno (2017). Moreover, other operations must be placed on the vehicles, refuelling for example and general maintenance. The object of the research is to create optimised path in order to minimise the time and distance when the fuelling service take place. Another important object of the optimisation is to model the best size of the fleet that generates the maximum revenue.

Factors affecting deployment This branch of research focuses into the analysis of the factors that affect the usage of vehicle-sharing systems. Hence, the operator can decide to open or extend the area in the specific location according to specific factors or to deploy more vehicles in the existing one.

Demand forecasting A prerequisite for any FF system, it brings about better operations and therefore the whole system performs better. Moreover, in a high competitive mobility market, offering the vehicle in the shortest time possible with the best accuracy generate competitive advantage for mobility providers. Furthermore this aspect is getting even more importance at the dawn of the autonomous vehicles coming into the market in the next 10 years.

2.5.2.1 Optimisation

Jorge et al. (2015a) have developed an integer programming model that generate one-way systems from SBCS operators. Hence, mathematical model was developed to assess which route to select in a SBCS. The model resolved that high demand point can be a factor of implementing both systems for special destination. The Airport represents one potential market integration for SBCS. The object of the optimisation,

takes into account potential revenue and loss of shifting the system into a FFCS for the Boston Airport. The results confirmed the validity of the model as a matter of fact the SBCS operator Zip-Car has later implemented into the business strategy.
Kumar and Bierlaire (2012) present a mathematical model to recognise the best number of stations and its best location in Nice, France. The key drivers that impact the presence of car-sharing were: High income and high education, presence of shopping mall, presence of hotels and high population density. On the other hand the factors that negatively affected the optimisation model were: distance and higher share to car usage to reach workplace.
Further object of optimisation of relocation, Jorge et al. (2015b) dealt with the high costs of re-balancing. They have proposed a dynamic pricing policies. The result is enhances profitability for operators, moreover the article shows the benefits of the optimisation of re-balancing made by user with dynamic pricing generates a better fleet balance for the operators.

2.5.2.2 Factors

Lorimier and El-geneidy (2013), have conducted the study for the factors that affect the vehicle usage for CS systems. Size of the station, and high availability of cars play a fundamental factor for the CS usage. Moreover, the seasonal factor affected the bookings of the cars in Montreal with summer season scored better outcomes in terms of CS usage. The intermodality with metro station is also a factor that decreases the availability of the SBCS vehicles and increases the station usage. Demographic factors are fundamental for the deployment of the car-sharing systems (Durán Rodas and Constantinos Antoniou Emmanouil Chaniotakis, 2017). Willing et al. (2017) have proved that in the research considering FFCS with the city of Amsterdam that point of interests such as restaurants, bookstores, banks bust stations are positive influencing the usage of CS. Another important factor that must be taken into account is the price or limited of parking as well as sports events and conferences (Giesel and Nobis, 2016).

2.5.3 Demand Forecasting

Xu (2007) uses an Evolutionary Neural Network to prove that this algorithm is possible to be used as a tool for forecast in car-sharing services. The previous researches use regression analysis method or agent-based simulation models. The demand forecasting has been used to profile certain type of user (Jorge and Correia, 2013). Therefore, the literature review does not offer that much on the demand forecasting of CS itself. Many different methods have been applied to this case but mostly to understand the factors and the locations that affect the most of the systems (Li et al., 2018).

CHAPTER 3

Data analysis and visualisation

Chapter 3 introduces the first step of the method of this research: the data analysis and visualisation of the object of the case study. First, the processes of the data handling before start the analysis, then the spatio-temporal analysis on figures and visualisation. Moreover the analysis has combined the use of different tools: the majority of the coding has been done through R programming language and RStudio environment. Hence, the data have been also processed with Tableau Desktop 2018 for the realisation of the maps, the clustering and the charts. Lastly, for the sake of the reading flow of this section, this chapter will not include the entirety of the charts produced in the analysis, general repetition of trend and figure state among the following sections. Some figures and tables are included in the appendix but the completeness of the work has uploaded in the overall hand-in file.

3.1 Structure of the data

The case study analyses the DriveNow data-set for the rentals of car-sharing service in the city of Munich during the year of 2016. The structure of the raw data has time and space information. The initial data-set were composed by two files. 1) Spreadsheets containing , the values of departures divided in two sheets: the first one from January 2016 to June 2016 and the second from July 2016 to December 2016. 2) the second containing arrivals of 2016 (with the same sheet configuration). The time information of the rentals were displayed both on the columns and rows; the columns were containing the days of the year in "day-month-year" and the rows were displaying the hour in a day in "HH" format. Regarding the space information in the data-set, the first column were composed by the code for the parking zone division by the city council of Munich (**?**). The representation of the first data-set has been exampled in the Table 3.1

. The values information are then, daily-hourly-area-based (time-space) for the

Table 3.1: Sample of first configuration of the data-set (fiction data).

		01/01/2016			02/01/2016		
Start Area	**Begin Hour**	**Average Distance (km)**	**Average Time (min)**	**Number of Starts**	**Average Distance (km)**	**Average Time (min)**	**Number of Starts**
750	00:00:00	12,3	97,3477615	2	NA	NA	NA
	01:00:00	NA	NA	NA	NA	NA	NA
	02:00:00	NA	NA	NA	NA	NA	NA
	03:00:00	NA	NA	NA	8,907865081	62,28888204	6
	04:00:00	8,466167123	1,948563089	1	15,21449844	86,33329888	2
	05:00:00	5,644222925	10,15270954	2	14,92642325	9,188465285	5
	06:00:00	0,007386206	13,82481071	3	NA	NA	NA
	...	...	...	...	...	...	...
	23:00:00	14,52950857	80,88558836	2	14,52950857	42,87858515	5

year 2016. The information provided in the data-set take into account:

- Average minutes spent in the trip
- Average kilometres spent in the trip
- Number of departures and arrivals

Each information is composed by a geographical representation point given by the code of the area, the hour and the day where the rental occurred and a triplet of columns that describe the average of the duration and the distance of the rental. The data-set did not provide a legend for the interpretation of the values. Therefore, the values of average kilometres and average time are not described and does not explain how the average is calculated . Hence, these information have been not taken into account for the analysis. Further clarification on these values will let new studies on the duration and the distance of the trips. The structure of the data analysed represents the time series of the rentals of the car-sharing vehicles highlighted in Figure 3.1. A time series is a time-ordered discrete sequence of data points equally spaced. It is important to mention that the departures and arrivals of the data-set **are not linked**. The information given from DriveNow does not represent the trips generated in network. The starts value point does not belong to an end point, but are solely the number of vehicles that leave or arrive in the specific area.

3.1.1 Data handling

Before introduce the process of data-handling it bears mentioning the theory of Data Science used. The general knowledge about data-science tasks tells that between the 60-80% of the data analysis is spent on cleaning and preparing the data (DASU and JOHNSON, 2003). The data preparation does not represent only an initial task, is a continuous process due to new different necessities that come across during the analysis (Wickham, 2017). The activities of data cleaning are: 1) outlier checking, observations have an anomalous distance from other values. 2) date parse and 3)

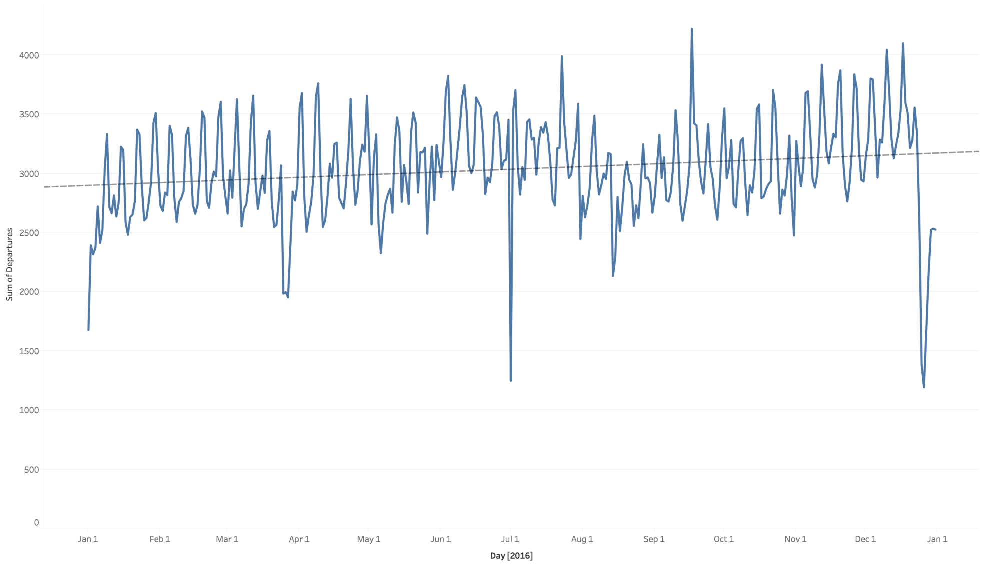
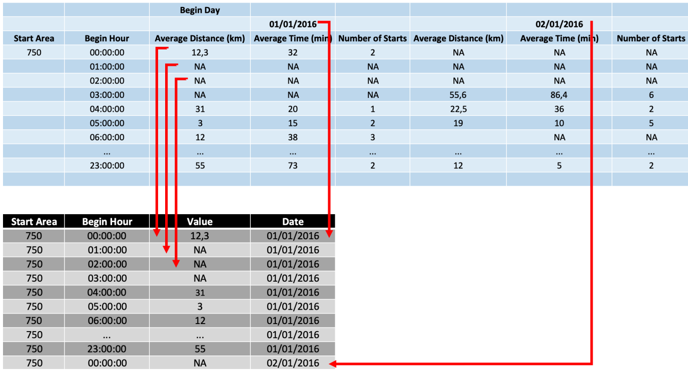


Figure 3.1: Time series of departures of DriveNow.

analysis of the missing values. The definition of ”Tidy Data” is a standard method to create data-set linking the layout of the structure with its meaning (Wickham, 2017). The principles of the tidy are then (Codd, 1990):

- Each variable forms a column
- Each observation forms a row
- Each type of observational unit forms a table

The process of data-tidying of the DriveNow data-set has the scope of creates univocal set of columns that contains the same standard of a Tidy structure. Each variable (distance, duration, number of starts/ends) represents a new single data-set: To generate the new structure of the data has been used the programming R. Having tidy data-sets benefits the whole process of the data analysis, each information is simply ”index-able” with simple command of sub-setting. Furthermore, the new layout pave the way on the structure of the time-series. There is a continuous flow between observation within the structure.
The data-set has been also processed to detect outliers. One particular example of outlier is in the area called ”(leer)” (”Empty” in German). The area presents strange values when compared to the distribution of the rentals in other zones. According to the evidences showed in Figure 3.3 it has decided to remove it for the spatial and



		Begin Day					
			01/01/2016			02/01/2016	
Start Area	Begin Hour	Average Distance (km)	Average Time (min)	Number of Starts	Average Distance (km)	Average Time (min)	Number of Starts
750	00:00:00	12,3	32	2	NA	NA	NA
	01:00:00	NA	NA	NA	NA	NA	NA
	02:00:00	NA	NA	NA	NA	NA	NA
	03:00:00	NA	NA	NA	55,6	86,4	6
	04:00:00	31	20	1	22,5	36	2
	05:00:00	3	15	2	19	10	5
	06:00:00	12	38	3		NA	NA
	...	...	...	...	...	...	...
	23:00:00	55	73	2	12	5	2

Start Area	Begin Hour	Value	Date
750	00:00:00	12,3	01/01/2016
750	01:00:00	NA	01/01/2016
750	02:00:00	NA	01/01/2016
750	03:00:00	NA	01/01/2016
750	04:00:00	31	01/01/2016
750	05:00:00	3	01/01/2016
750	06:00:00	12	01/01/2016
750	...	...	01/01/2016
750	23:00:00	55	01/01/2016
750	00:00:00	NA	02/01/2016

Figure 3.2: Process of data-tidying.

time analysis. Further inquiries has been asked to DriveNow about this area. These values embrace all the rentals outside the business area , however, the distribution of the values are concentrated almost exclusively at midnight in the first 6 months and under the suggestion of DriveNow it has been excluded. (DriveNow, 2018) (Further evidences are reported in the Figure A.2).

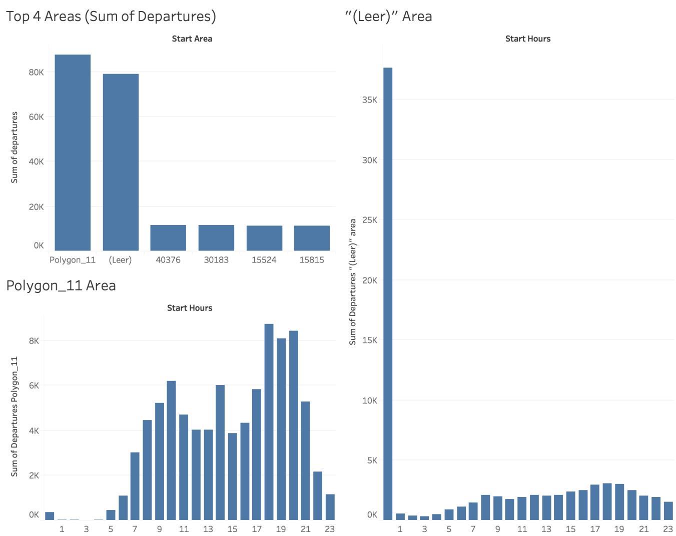


Figure 3.3: Sum of Departures of 2016; Top Left: Top 4 areas, Bottom Left: sum departures based on hour, Right sum of departures based on hour..

3.1.2 Definition of zones

The zones where the rental information are allocated are related to the distribution of the parking zones for the city council of Munich. These zones are defined how the municipality describes the parking regulation for each micro area of each neighbourhood (Par, 2018). On the top of these area definition, the information collected are related to the ”satellites” of DriveNow as well. Hence, it has overlapped the map of Munich with two additional layers: the distribution of the parking zone and the business zone of DriveNow in Munich. The outcome of this process has generated the map in Figure 3.4. The representation of the business zone of DriveNow in Munich is highly fragmented in 274 areas. On the other hand this disposition can generate high level of detail about the rentals of shared vehicles. In Figure 3.4 it has presented the representational centre for each point (on the right side of the figure). Moreover, on the left map, it has highlighted with red colour the zones where the link between the name of the areas of the data-set and the areas of the spatial file is missing. There are twelve points that cannot be associated between the spatial file and the data-set of the rentals. The reason is that the name of the zone between the two files must be the same otherwise the relationship cannot be verified. The missing zones are mainly the satellites and one peripheral zone of the entire network.

3.1.3 DriveNow in Munich

DriveNow is the FF car-sharing services owned by the BMW group, in Munich and offers the possibility to drive BMW and MINI models for the price generally in between 0,33 and 0,36 per minute. The fleet is composed by electric (BMW Model i3) and combustion engine vehicles. The car-sharing provider operates in Munich with 750 cars in 12 different models. This is the largest and the most variety fleet among any other FF or SB competitors. DriveNow vehicles can be rented through smart-phone app or by using an RFID card tapping it on the windscreen sensor . The registration costs between the 4,99 and 29(according to the promotions period during the year) and includes from 15 to 30 free minutes. This sub-section introduces the figures for DriveNow in Munich during the 2016. The car-sharing service has recorded 1.110.239 departures and 1.109.558 arrivals. Theoretically speaking, the number of departures should match the number of arrivals, because in car-sharing systems each arrival generate a new departure. However, this definition is not possible to see in the case study because the number of arrivals and departures does not take into account the extra movement of the vehicles such as re-balancing and maintenance. Furthermore, the number does not include the ”leer” area (explained in section 3.1) hence is not possible to have the ”perfect match” between arrivals and departures.

$$\Delta_{trips} = +681 \text{more departures}$$

The next analysis has taken into account the efficiency rate of utilisation of the vehicles in the DriveNow network based on the following assumption:

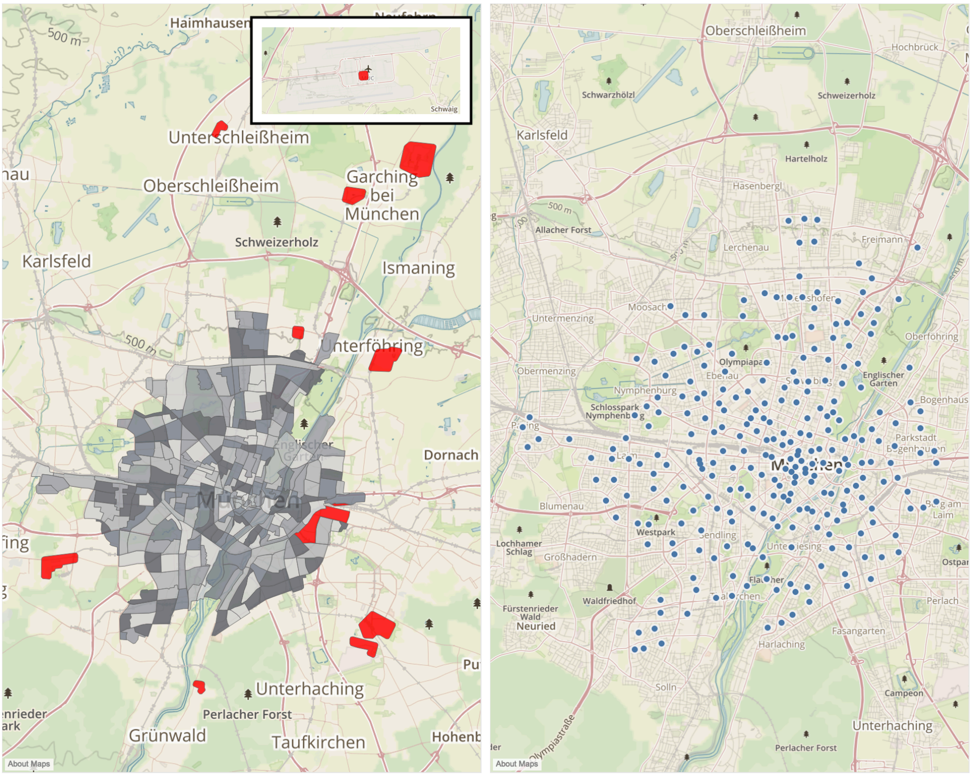


Figure 3.4: Definition of the zones for the study, red colour areas where missing geographical correlation between shape file and data-set.

The sum of number of departures and number of arrivals at the same instant given for a given day for all the areas, it corresponds to the number of active vehicles in the same instant.

$$\eta_{NetworkMunich} = \frac{\sum_{Areas} S_{time} + A_{time}}{NV_{fleet}} \tag{3.1}$$

The assumption made is within a tolerance of $\pm = 15 minutes$ the vehicles departing and arriving represent the total vehicles moving. In the literature it has described the number of vehicles per fleet of share mobility operators in Munich. The number of vehicles in the fleet of DriveNow in Munich is then known (750 vehicles Table 2.1) as a matter of fact the difference between the moving vehicles and the number in the fleet represents the not used cars in the specific time instant. The ratio of the DriveNow Munich efficiency is summarised in the appendix. The general average utilisation of the fleet over the year is of 33,7 % with hourly peaks of even 107,3% at 07:00 PM on 03/07/2016 (so additional vehicles were added into the fleet) and the least utilisation rate is 1,5 % at 03:00 on 26/12/2016. The day with the best average utilisation of the fleet is on 17/09/2016 with the 46,4 % of rate efficiency (highest 96,1% at 08:00 PM and lowest 8,9 % at 04:00). Of course, is clear that the trend of the booking have very high peaks in the rush time and very little utilisation in the night reflecting the same trend of transport demand the temporal trend analysis has been discussed in the subsection 3.2.1.

	Hour																							
Date	X0	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23
21/10/2016	17.2	12.0	6.4	5.2	8.1	14.3	22.1	34.4	47.2	47.3	37.9	34.3	36.4	41.3	34.3	41.2	44.0	53.5	51.6	34.0	42.3	44.8	40.4	29.2
22/10/2016	27.9	23.5	10.9	10.4	10.4	9.2	12.3	17.5	28.7	37.6	38.8	43.1	43.2	52.3	44.5	51.7	52.7	61.6	73.3	91.6	85.7	63.9	48.4	42.4
23/10/2016	36.4	26.7	19.2	17.5	11.2	9.6	10.0	11.6	17.6	28.8	32.7	37.1	48.7	53.6	53.3	52.0	50.9	58.3	70.3	82.0	86.0	59.3	44.1	38.1
24/10/2016	13.9	6.0	3.2	2.7	7.5	19.3	22.8	31.1	48.1	43.7	34.9	29.6	42.3	46.1	48.4	56.4	56.1	58.0	58.7	57.2	46.5	33.1	34.0	26.8
25/10/2016	11.1	6.4	4.9	2.9	4.7	14.1	20.9	34.3	50.9	41.6	33.9	29.5	26.7	26.9	28.1	30.4	35.2	45.7	53.3	58.9	52.3	40.3	30.8	24.5
26/10/2016	15.6	7.9	4.3	2.0	6.4	14.0	16.8	27.1	34.5	41.3	33.3	26.7	29.7	35.3	33.9	32.7	40.9	52.0	64.0	67.6	58.0	51.9	40.0	28.3
27/10/2016	17.5	8.8	4.8	2.9	7.3	17.5	16.3	31.2	32.8	37.3	33.1	30.8	28.1	33.2	28.8	36.9	34.1	44.9	59.6	67.1	57.6	39.9	39.5	35.2
28/10/2016	17.5	7.1	6.8	5.1	12.1	19.1	21.6	32.8	46.8	40.0	37.2	33.7	30.9	34.8	29.6	31.9	36.9	40.1	54.4	64.9	62.0	54.3	48.8	29.7
29/10/2016	24.7	22.5	10.5	8.8	15.5	18.0	14.9	21.1	22.5	32.7	36.5	33.9	34.3	41.6	40.5	45.6	48.4	54.3	58.3	74.5	70.7	56.5	52.9	41.2
30/10/2016	29.1	22.8	24.7	8.3	6.7	5.9	8.8	13.2	19.7	23.2	29.9	33.1	33.1	42.7	41.5	38.5	39.6	41.2	52.8	60.4	54.9	44.4	37.3	37.3
31/10/2016	11.5	7.1	3.5	2.5	4.9	10.1	16.3	23.7	34.8	29.9	28.0	31.5	33.3	36.3	35.2	39.3	41.2	47.9	46.3	44.5	37.5	39.5	32.0	21.5
01/11/2016	26.8	21.7	11.6	9.9	9.1	6.9	10.0	12.7	17.2	28.9	32.8	35.2	31.7	34.8	39.1	42.1	44.9	55.3	74.1	92.0	78.4	65.3	53.6	36.1
02/11/2016	15.6	7.5	4.8	3.6	9.1	18.5	26.1	37.2	43.2	39.5	38.7	37.6	36.9	42.0	51.6	49.6	55.6	55.1	59.5	54.9	42.7	31.2	30.4	28.7
03/11/2016	14.9	8.1	5.3	1.9	4.7	12.0	21.1	31.5	48.5	41.1	32.1	28.4	38.0	35.1	34.3	43.3	45.7	45.5	57.7	64.1	56.3	42.1	36.3	23.7
04/11/2016	20.1	10.0	5.6	4.0	6.8	13.9	21.7	34.9	43.5	42.3	34.8	35.3	29.6	36.3	32.4	36.1	41.1	51.7	62.5	66.8	59.9	49.5	42.1	30.3
05/11/2016	27.9	21.2	11.2	10.1	8.0	9.7	12.8	19.1	27.6	41.1	44.9	46.3	35.7	41.5	47.2	46.1	55.9	61.6	72.5	97.3	82.1	59.2	48.8	38.5
06/11/2016	30.9	27.1	12.5	11.1	7.9	6.3	9.3	10.3	19.7	34.7	37.6	44.3	48.7	56.8	59.6	54.7	53.9	64.8	78.5	90.4	80.3	63.1	51.2	41.2
07/11/2016	12.0	5.6	3.7	3.9	9.6	23.9	28.0	34.3	49.2	45.7	37.7	34.1	48.8	56.3	59.6	59.9	60.5	60.5	69.6	55.6	45.2	33.2	31.6	26.5
08/11/2016	15.2	2.4	0.0	1.2	5.3	14.0	20.8	29.3	35.9	42.8	34.1	33.2	37.9	37.7	36.7	36.1	45.6	51.9	67.2	66.4	52.8	47.1	42.4	28.1
09/11/2016	15.1	8.1	4.1	4.1	5.5	15.6	22.3	32.4	41.2	42.7	33.3	27.6	31.2	34.4	35.7	36.3	42.8	48.4	59.7	64.3	54.0	47.6	38.0	25.6
10/11/2016	13.6	8.8	4.3	2.7	5.7	11.3	22.8	38.4	49.6	42.5	32.4	30.3	33.6	33.6	32.3	36.5	43.6	54.1	62.4	71.9	53.6	45.7	39.5	24.1
11/11/2016	16.7	10.7	6.3	2.8	7.6	16.9	24.1	43.3	50.0	49.1	38.4	38.8	35.3	38.0	43.5	44.3	46.9	50.3	62.9	82.4	56.5	49.1	41.2	31.5
12/11/2016	32.1	22.1	10.7	9.1	11.2	9.5	14.5	21.3	32.0	39.1	44.4	49.5	46.0	45.1	48.3	59.6	63.5	60.8	83.9	96.4	87.3	57.9	52.3	41.1
13/11/2016	34.3	25.3	18.8	12.5	9.2	6.7	9.6	15.1	18.3	27.3	32.8	41.2	46.9	50.9	56.0	53.5	55.9	65.6	73.5	91.7	70.9	53.9	45.7	40.3
14/11/2016	10.9	4.7	3.2	2.1	8.0	19.9	27.1	32.5	42.0	43.9	40.9	32.7	44.5	52.3	61.7	57.2	60.4	65.2	62.3	54.9	42.1	33.9	32.5	18.3
15/11/2016	13.2	7.2	3.6	2.7	6.1	16.8	23.6	33.5	47.5	45.2	37.1	32.4	34.7	34.5	31.6	38.3	46.7	57.5	63.2	70.8	61.2	51.6	36.1	24.7
16/11/2016	14.9	8.5	4.3	3.9	4.9	13.2	26.9	41.2	52.3	58.7	41.3	34.3	33.5	35.2	36.0	40.7	42.5	44.8	65.2	73.6	64.5	49.5	41.2	27.3
17/11/2016	16.8	11.1	4.9	2.8	6.7	12.8	25.9	38.3	47.9	48.7	39.3	35.7	37.7	36.1	45.9	42.7	44.4	49.2	68.9	81.9	66.8	45.9	45.9	32.4
18/11/2016	18.3	7.9	4.5	3.9	6.8	13.3	24.8	35.1	47.5	50.5	38.4	38.3	35.9	36.3	38.3	41.7	44.4	50.9	64.3	71.2	69.7	52.4	47.9	34.8
19/11/2016	29.6	26.0	14.7	9.1	7.7	12.7	9.2	20.0	31.9	41.1	48.5	47.9	40.4	43.7	42.8	50.5	52.5	61.3	73.9	94.4	79.6	63.7	51.2	40.9
20/11/2016	38.1	31.5	18.8	12.4	4.3	4.8	8.8	11.6	18.3	31.3	42.3	45.6	56.8	60.3	64.8	66.0	67.7	78.8	91.9	80.9	70.4	56.7	46.5	32.7

Figure 3.5: Example of the matrix of efficiency of the fleet.

3.2 Statistical analysis

This section goes into the deep analysis for the temporal and spatial components of the data-set. The temporal analysis takes into account the pace of the departure according to hour and days of the year of 2016 without the constraint of the geographical information described in the previous section. The spatial analysis combines the result of the temporal analysis and examines the trends in the city of Munich on a geographic basis.

3.2.1 Temporal trend

The figures of the trips of the cars of DriveNow in the 2016 are are summarised in Table 3.2. The general month trend shows that the number of the trips during summer season is higher than the rest of the year until the break in August (Summer Vacation). The principal reason of a drop of 9 % between July and August is probably caused by the summer vacation period. Moreover, Table 3.3 lists the overall number of rentals according to the day of the week. The results are in accordance on what described by Schmöller et al. (2015) for the cities of Munich and Berlin in 2013. The author explained and analysed the two German cities in space and time. The results of this analysis agree that the weekends (and Friday) have registered higher demand than the rest of the other days of the week. This result links the leisure activities that generally take place in the weekends and Friday evening with the usage of the car-sharing FF services. Thus, the box-plot in Figure A.1 displays the use of DriveNow service

Table 3.2: Figures of Number of departures and arrivals in for every month of 2016.

Month	**Departures**	**Arrivals**	**Previous Month +/-**
January	86.399	98.013	-
February	86.821	95.735	0,488
March	87.868	96.189	1,206
April	91.704	100.984	4,366
May	92.887	101.973	1,290
June	98.554	110.516	6,101
July	97.651	101.594	-0,916
August	88.836	92.072	-9,027
September	91.547	94.978	3,052
October	93.957	97.563	2,633
November	98.005	101.998	4,308
December	96.010	99.962	-2,036
Overall	**1.110.239**	**1.191.577**	**11**

Table 3.3: Departures on day of the week.

Day of the week	Departures
Monday	148.429
Tuesday	147.693
Wednesday	149.650
Thursday	155.212
Friday	167.281
Saturday	182.611
Sunday	159.363

according to the days of the week. Both the box-plot and Figure 3.6 summarise the trend of departures[1] in the hourly-daily-weekly-monthly basis and completely highlights the different trend of the car-sharing bookings in time. The peak of the demand is concentrated in the late afternoon-evening hours. However, the pace of the lines is different; working days are characterised by two major peak periods, the first between 07:00 and 10:00 and the second in the evening from 17:00 to 20:00. On the other hand, the weekend days behave in a different way: there is a smooth increase

[1] The arrivals have same trend and it has plotted them in the appendix

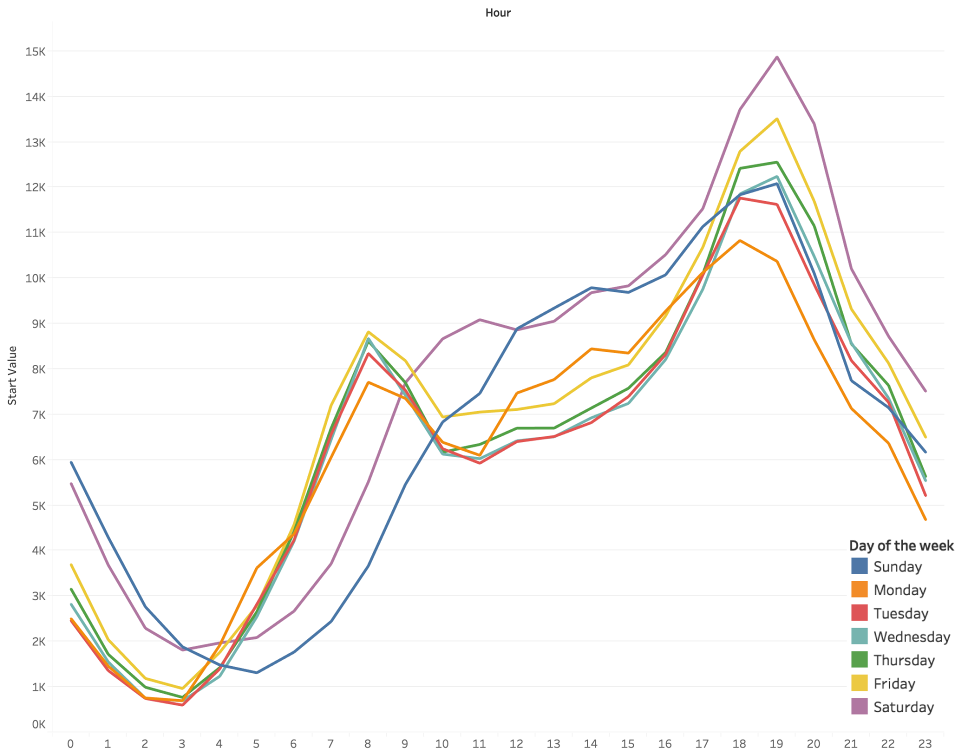


Figure 3.6: Departures by day-hour of the week.

of the bookings without any drop in the afternoon hours. The growth of the rentals continue until the peak period between 16:00 and 22:00 on Saturday and 18:00 and 21:00 on Sunday. The chart Figure 3.1 plots the departures[2] for every day of the year 2016. This figure helps to understand the general trend of the rentals and to find out if the time series contains high seasonality. The trend line shows that there is a slightly increase from the previous year, but the bookings are not affected by seasonality the general trend is constant. Table 3.4 has recapitulated the values of the linear trend model. Speaking of the pace of the bookings for every single day; the biggest drop is occurred on the 26th of December with 1.194 departures due to Christmas. On the other hand, the day with the highest number of bookings is Saturday the 17th of September 2016. This result is interesting because it corresponds with the starting day of the ”Oktoberfest 2016” (see Figure 3.7). The world-famous Beer-fest has its tradition to start every year on the 3rd Saturday of September, therefore this

[2]only departures not relevant to show the arrivals as well, the both charts are posted in the appendix

Table 3.4: Linear trend model.

Individual trend lines:

Panes		**Line**	**Coefficients**				
Row	**Column**	**p-value**	**Term**	**Value**	**StdErr**	**t-value**	**p-value**
Departures	Date	0,0002918	Start Date	0,746401	0,20405	3,65794	0,0002918
			intercept	-28727,8	8682,85	-3,30856	0,0010313

combination (Saturday plus main event in the city) has produced such peak in the number of bookings of car-sharing.

Figure 3.7: Trend daily in 2016 - particular on the Oktoberfest.

3.2.2 Spatial trend

The spatial trend sub-section takes into account the information analysed in the temporal trend and adds geographical facts of the figures of DriveNow in Munich. First of all the chart Figure A.3 shows the rank between the top 11 zones between overall departures and arrivals. The area that gathers the most bookings is defined as Polygon_11. This area is part of the missing geographical information described in the subsection 3.1.2 and plotted in the map at Figure 3.4. As mentioned before in the chapter, the "missing" areas marked in red are representing mostly the satellites. As a matter of fact, it is clear that the Polygon_11 is the representation of the Airport of Munich. Compared to the other red locations at Figure 3.4, the most relevant in terms of demand is the Airport. The other satellites representing detached aggregation points or special points of interests such as Hospitals, Shopping Malls, Theme parks, University Campus (Technical University of Munich) and business districts. The description of the satellites represents a metaphor that summarises some of the influential factors of the use of car-sharing described in the literature review: parking cost included, utilisation of CS vehicles for leisure activities and the generic user profile is male and young.

3.2.3 Strategic Zones

Another space-time analysis performed in this section is the calculation of the hot spots and cold spots in the Munich DriveNow's network. The investigation has dealt with the number of the departures and arrivals for each zone in the business area. Figure A.3 shows three major clusters: the cluster 1 is composed by the areas that have scored a +170 until -881 booking differentials this cluster considers the areas that are in approximately in balance (proximity of 0) and the areas that are defined cold spots where there is a predominance of the arrivals that generate a negative unbalance. Within the cluster 1, the "coldest" spot is dominated by the satellites of the network with Polygon_4 with -881 imbalance and Polygon_7 with -687. However, the mentioned problem of missing geographical information about the exact location of these satellites does not allow to infer where these areas are located in Munich. Regarding the other cold spots, there are: the zone 509522 located in the area called "Am Hart" with -583 imbalance and the zone 62296: the zone includes the west side of the Olympiapark, the "Olympiaeinkaufzentrum" (Shopping Mall). Regarding the cluster 2, this aggregation considers the areas 1773, 1670, 1530, 1255, 1252 and 1071 the major departure areas rather than arrival. These areas are grouped in the nearby of major intersection with the public transport: Central Station (in specific the North Gate), Münchener Freiheit, the multimodal mobility station with concentration of the public transport services (see literature at subsection 2.3.2) and Sendliger Tor intersection with 4 metro lines and tram and busses services. The result of cluster 2 is very interesting because combines very dense areas with the overall concentration of public transport services in Munich. The cluster 3 has more departures than arrivals, but less than Cluster 2 has overall numbers. In this group there are mostly residential areas of the city of Munich. The analysis shows that the areas at the edge of the cluster chart Figure A.3 (which has been re-sized to show the major values of the balance-clusters), there are the areas that have a yearly unbalanced trend. These zones are probably objects of continuous re-balancing strategies and operation for DriveNow. These areas are sectors of the business zones that have principal trip characteristics independently from the time of the day and the day of the year. However, it has analysed in detail the trend of the departures on a daily basis and the Airport (Polygon_11) turned to have very high unbalanced days. For example, between the 22th and 26th of December according to the Christmas holidays, the figure is at -418, which means that half of the fleet is parked at the Airport during these days. This probably represents one of the most critical days in terms of re-balancing problem for DriveNow.

Table 3.5: First 13 Areas with higher concentration of Departure and Arrivals.

Area	Value Start	Value End	Diff. START - END	% DEPARTURES	% ARRIVALS	Cluster
Polygon_11	87555	87401	154	7,891520547	7,877100611	1
40376	11642	11988	-346	1,049318511	1,080430225	1
15524	11437	11334	103	1,03084142	1,021487836	1
30183	11516	11008	508	1,03796186	0,992106767	3
62296	9982	10438	-456	0,899699139	0,940734959	1
39866	9764	9987	-223	0,88005033	0,900088143	1
16526	9813	9614	199	0,884466805	0,866471153	3
15815	11224	9451	1773	1,011643271	0,851780619	2
44749	8857	8860	-3	0,798300468	0,798516166	1
4373	8640	8727	-87	0,778741791	0,786529411	1
Polygon_7	8022	8709	-687	0,723040121	0,784907143	1
Polygon_3	8096	8593	-497	0,729709901	0,77445253	1
Polygon_4	7604	8485	-881	0,685364882	0,764718924	1

Table 3.5 ranks the areas based on the percentage of the overall bookings (the complete table can be found in Table A.1) and highlights in orange the satellites areas. Yet, the Airport (Polygon_11) is the most important area for the business of the DriveNow with the 7,89% of departures and the 7,87% of arrivals of the whole business network. This figure gains even more importance considering the size of other areas are relatively small because of the business zone has been divided in 275 areas. The area has very critical situation during the major holidays period, with concentration of a huge number of vehicles "stuck" in a very remote area that cannot be picked by people that are around. Therefore, the operator must deal with an important plan of re-balancing, with very attractive incentives for customers and compete with public transport and other private companies that serve that route. DriveNow in Munich has to deal with this two sides of the same coin for the Airport. Although the area has a quite good yearly turnaround it exceeds with 154 more departures than arrivals during the 2016. The data does not take into account information on the re-balancing procedures therefore, the figure of departures and arrivals reveals the true amount of trips at each zone.

3.3 Visualisation

This section provides an understanding of the departures and arrivals according to the visualisation of the data on maps. Visualisations have big significance in the data analysis, it displays large number of information that people can immediately understand and gather together at first glance without going into deep on values and parameters. Furthermore, it helps to understand broad trends that cannot be seen in the tables. The section provides the visualisation on map of trend of departures and arrivals of the cars during the day. Lastly, this section considers special circumstances occurred in Munich during the 2016 to see the different outcome on the network of DriveNow.

3.3.1 Map of Departures and Arrivals

Figure 3.8 and Figure 3.9 resume the trend of the trips in the business zone of DriveNow car-sharing vehicles in Munich. The display of the morning departures between 04:00-07:00 and 08:00-11:00 shows that rentals start in the proximity of the ring-road Bundesstraße 2 and 2R (as known as "Mittlerer Ring") with more concentration in the middle-north of Munich. Comparing the same time period with the arrivals figure, the trips made tend to end the rental in the north of the city reaching the borders of the business zone. These are business districts in "Parkstadt Schwabing" with high density of workplaces and offices. Moreover, another principal area of arrivals is represented in the inner city (orange and yellow areas in the centre between the 08:00-11:00 arrival figure), the university district has higher concentration compared to rest of the network. Meanwhile, the departures and arrivals generated between 12:00 - 15:00 are inside the northern area and does not have a clear "from-to" path in the map. The afternoon-evening period from 16:00 - 19:00 the departures are spread all over the network and the concentration of the arrivals is divided by residential areas and city centre. In the evening-night period and the night period, the bookings start from the city centre and university district and end throughout the business zone with peaks in the city centre. It is interesting to mention that the number of arrivals in the first frame correspondingly to 00:00 - 03:00 time period, represents also the time frame where the public transport has the biggest interruption of the service or long waiting time. The metro lines are replaced with low frequency night bus and night tram. It can be argued that the usage of car-sharing in Munich complement the interruption of the public transport.

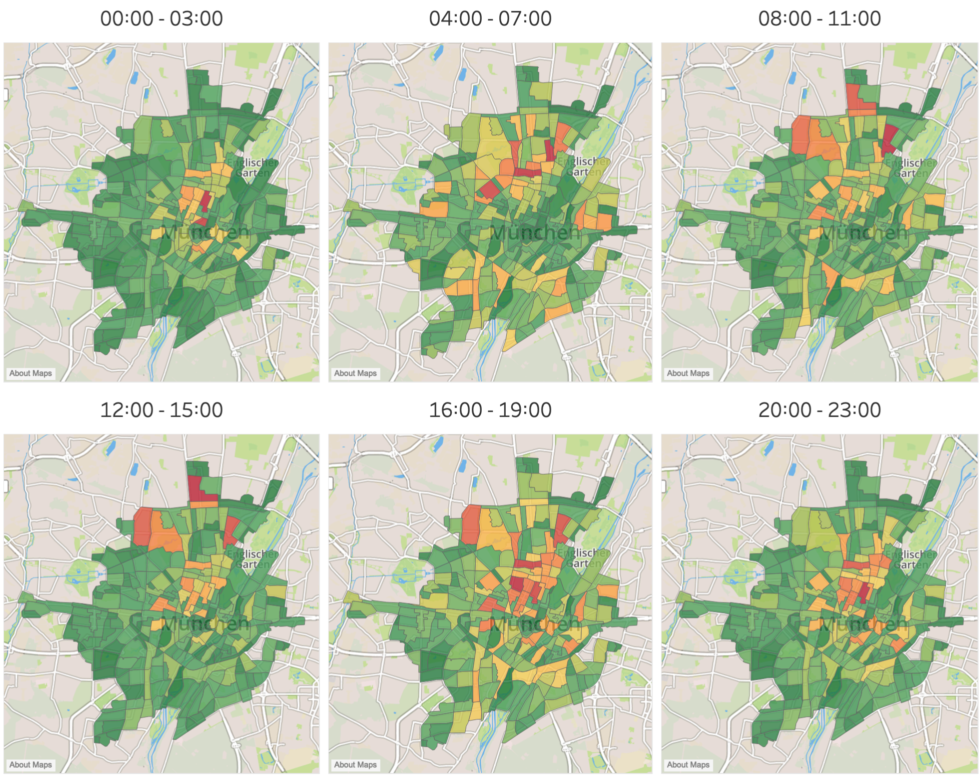


Figure 3.8: Departures on Munich based on time-space.

3.3.2 "Specific circumstances" visualisation

This sub-section implements the tool of map visualisation for the usage of the car-sharing under special circumstances. It has selected two major events that happened in Munich 2016: the Oktoberfest, and the Munich terror shooting on the 22nd of July. That obviously had reflected issues in the traffic management and the shut down of all the public transport. The purpose of the this visualisation analysis is to have a quick understanding of the relationship between events that pressure the transportation system at its most with an high spacial level of detail. Thanks to tools such as Tableau, it is possible to print out results on map that are easy to be analysed by the operator.
The Oktoberfest in 2016 had taken place from the 17th of September until the 3rd of October. The trend in the rentals as described before increases during these events. The map shows the detail on the Theresienwiese (location of the Oktoberfest) area. The maps shows that the users of DriveNow use this mean of transport to get to

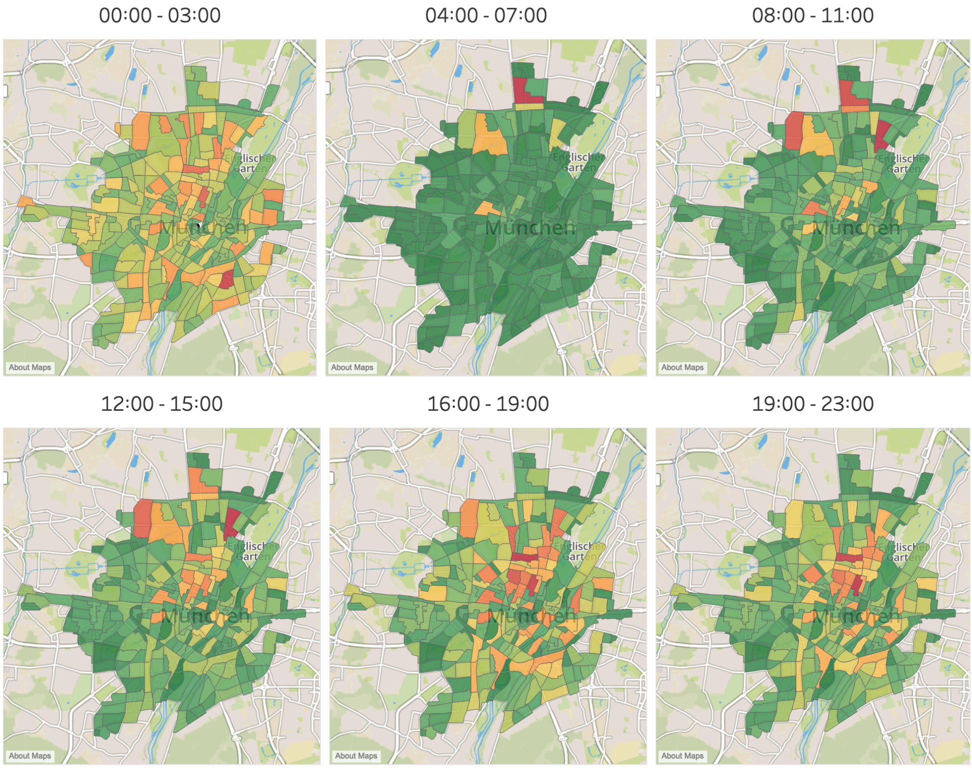


Figure 3.9: Arrivals on Munich based on time-space.

the event and to leave it. The next visualisation considers the Munich terror attack that took place in the proximity of the ”Olympiaeinkaufzentrum”, the shopping mall in the ex-Olympic area in the north-west part of the city on 22/07/2016 at 17:52. It has tried to understand under such circumstances the response of the people in that area. The map shows that from the time of the shootings on-wards people have more used car-sharing services than the average, especially nearby the area of the shootings. Immediately after the shootings the traffic situation in Munich was very critical. Public transports were shut down and that why it can be argued that the car-sharing systems has been utilised as emergency response mean of transport under this circumstance.

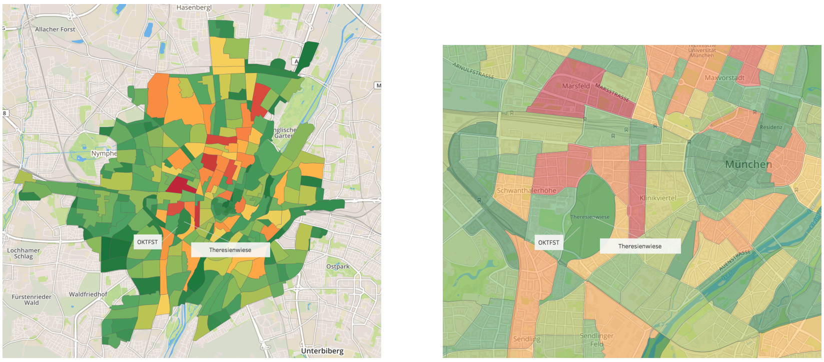


Figure 3.10: Oktoberfest bookings.

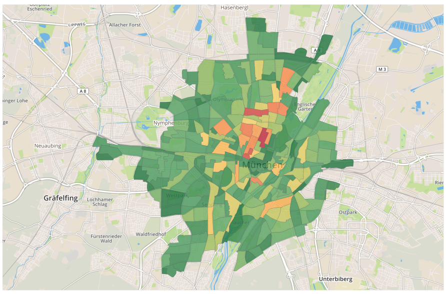


Figure 3.11: Average Departures 18:00-23:00.

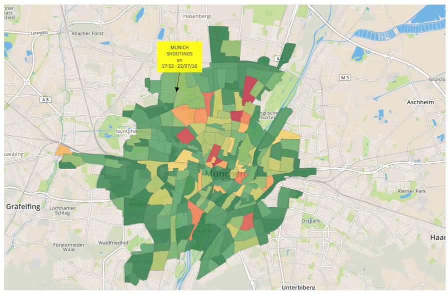


Figure 3.12: Departures on 22/07/16 between 18:00-23:00.

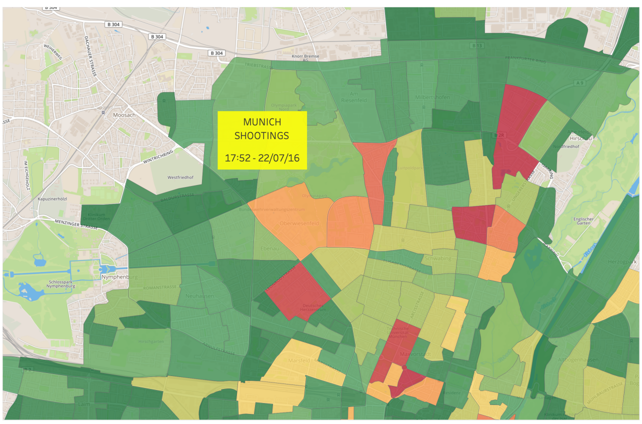


Figure 3.13: Particular in the Olympiaeinkaufzentrum area.

CHAPTER 4

Demand forecast, methodology and model specifications

4.1 Introduction

In this section it has introduced the model utilised for the forecast of the departures and arrivals of the car-sharing vehicles. The case study deals with the network of the car-sharing provider DriveNow in the city of Munich for the year 2016. Lastly, the chapter includes also the variable selection criterion, the description of the aggregation of the data-set in order to reduce the calculation time of the training process.

4.2 Aim of prediction

The prediction performed in this project aims to forecast departures and arrivals within a time frame of one week. The forecast has considered the flows of the rentals in the zones described in chapter 3 within the ”Business Area” of DriveNow[1]. Moreover, the zones have been reclassified due to computational time issues described in Figure 4.5 and then aggregated in macro-areas as the neighbourhoods of the city (within the DriveNow ”Business Area”). Therefore, the forecast value for the macro-areas has been dis-aggregated by the proportion of each micro-area to the major one. The forecast represents one of the tools that the car-sharing providers must use and integrate in their strategical decisions; mobility companies in order to allocate as best as possible their resources (vehicles and personnel), a varied set of tools in order to predict the demand as precise as possible and then optimise their operations and deals. For example this tool can be used for a more accurate prediction of the ”hot-spot” and ”cold-spot” of the operational business zone. This definition implies where there is an unbalance between arrivals and departures of the car in a given time-frame. Generally speaking residential neighbourhoods are very busy areas during the

[1] According with the availability of the geographical information of the areas in the data-set given

morning and evening on working days (due to the work-home trip purpose) but are ”cold spots” during the rest of the off peak time. Therefore, the disequilibrium of the rentals between ”hot and cold spots” generates one of the causes of the re-balancing problem described in the literature in subsection 2.4.1. Here the forecasting task plays a fundamental role when the car-sharing provider must design an optimal strategy between deployment re-balancing and pricing offer (user re-balance) optimising the number of trips on loss. In other words, understanding the future demand of mobility is fundamental for car-sharing services for not missing future revenues cause to an unmet demand and to the re-balancing problem; an overrate or underrate value causes a system disequilibrium that turns into an unsuccessful deployment and potentially into a cost of re-balance.

4.3 Configuration of Models

Neural Networks are part of the vast techniques of Machine Learning, in this section it has analysed the Neural Network theory that stands behind the model implementation at section 4.4. It has summarised the meaning of its parameters, the object of Neural Networks algorithm and the distinction between Multi-layer and single-layer networks.

4.3.1 Neural Network

Artificial Neural Networks (ANN) or simply Neural Networks are computing systems that take inspiration by the system of biological neural networks in the animal brains. The neural networks contain computation units as well as the biological networks are called *neurons*. These unites are connected each other such as the synaptic connections by *weights* (Bohte, 2018). The network is composed by *inputs*, *outputs* and *intermediate parameters*. The inputs are scaled by weights connecting the computation units and defining the computational function of the cognitive process. The learning process takes place in the selection and change of the weights that link the computation unit. The network is “fed” by training data containing the inputs and the outputs of the objective to be learnt in this way this process “teaches” and provides the assessment to the network (the model). Below the mathematical representation of the linear function at each output node (Aggarwal, 2018). Thus, the formula for a general Neural Network is:

$$Y_k = \phi\{W \cdot X\} = \phi\{\sum_{j=1}^{d} w_j x_j\} \tag{4.1}$$

- X_i, Input layer containing d nodes ($X_1 = [x_1 ... X_d]$) with
- W_j, Weights ($W_j = [w_1 ... w_d]$)

- ϕ, Activation function
- Y_k, output layer

4.3.2 Parameters

The networks are composed by *single-layer* or *multi-layer*. The single-layer network is the earliest and simplest form of Neural Network were called *perceptron* a decision function that takes binary inputs and generates an outputs. (Aggarwal, 2018). The

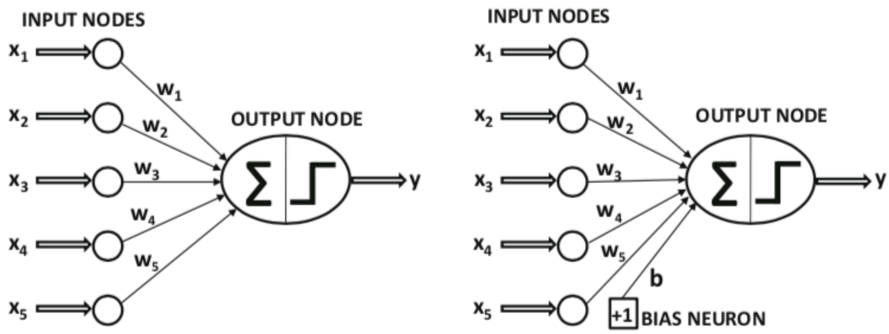


Figure 4.1: The perceptrons elemental structure..

figure Figure 4.1 shows a single input layer that is connected to the output node. The weights $(w_1...w_d)$ contain the features that are multiplied and added at the output node. The bias showed in the right side of the Figure 4.1 represents one of the Neural Network hyper-parameters to help the activation of the neuron. This parameter helps to implement in the model an invariant part of the prediction (Aggarwal, 2018). The goodness of the result is measured by the error on the output that the Neural Network has calculated compared to the real one, hence the objective function of these models is to minimise it, the smaller is the error the closer is the result to the reality (Aggarwal, 2018).

$$\min \sum_{x_i, y_k} (y_k - \hat{y_k})^2 \tag{4.2}$$

The importance of nonlinear activation functions becomes significant when one moves from the single-layered perceptron to the multi-layered architectures. In the book (Aggarwal, 2018) are summarised the principal activation functions for Neural Networks (with v as argument of the Neural Network):

$$\phi = sign(v), \text{ Sign function}$$

$$\phi = \frac{1}{1+e^{-v}}, \text{ Sigmoid function}$$

$$\phi = \frac{e^{2v}-1}{e^{2v}+1}, \text{ Tanh function}$$

$$\phi = \max\{v, 0\}, \text{ Rectified Linear Unit [ReLU]}$$

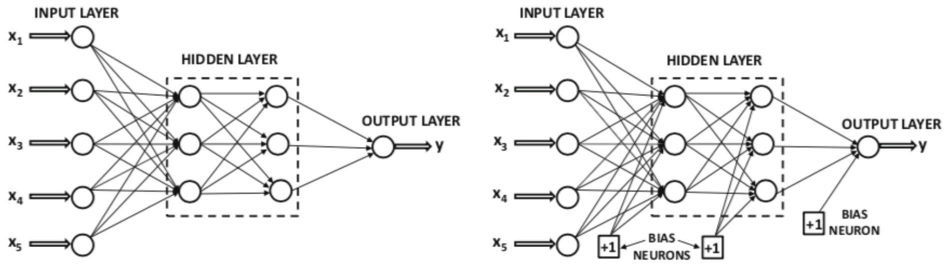
$$\phi = \max\{\min[v, 1], -1\}, \text{ Hard tanh}$$

The sign activation can be used for binary outputs at prediction time. The sigmoid can be used for computations for probabilities. The tanh function can be used when the outputs of the computations are desired to be both positive and negative. Morevore, it is preferable to the sigmoid (which are correlated, by $tanh(v) = 2 \cdot sig(2v) - 1$) is easier and faster to train. The ReLU and hard tanh are recent activation functions used in training of multi-layer Neural Networks and perform better in the training phase of such models.
Multi-layer Neural Networks are the configurations that have more than one computational layer. The process of the transmission of the data until the output layer is processed trough intermediate layers (between input and output) called *hidden layers* then all the neurons in one layer are connected to those of the next layer. Multi-layer networks are named also feed-forward networks, because the structure feed by every hidden layer in advancing in direction from input to the output(Aggarwal, 2018). Figure 4.2 shows the graphic representation of Multi-layers Neural Networks.

To sum up, the following list shows the algorithm summary of a generic Neural Network explained in the class of Statistical Learning at Technical University of Munich on the 17th April 2018 (Ma and Antoniou, 2018).

- "Send the input X_i"
- "Calculate the output Y_k"
- "Given the correct output O_k calculate the error"
- "Adjust the w_ij"
- "Adjust the w_jk"

example.png example.png example.bb example.bb



(a) No bias neurons (b) With bias neurons

Figure 4.2: Multi-layers Neural Networks a) without bias b) with bias.

- ”Repeat with a new example/training input”

4.4 Implementation of Neural Network model for DriveNow

The implementation of the Neural Network method for the ”business zone” of DriveNow is characterised by the three major applications:

- Interpret the Spatial variables
- Interpret the Time variables
- Keep both information in the same model

The data-set has the pair of geographical information and the time information for each day and hour of the year 2016 for the rentals of departure and arrival of the DriveNow car-sharing in Munich. The structure of the data-set reveals the possibility to use the variables in time series way as described in chapter 3.
To use the configuration of time series in a Neural Network is bear mentioning the number of lags where the model must refer to; it has selected an $\alpha = 3$, this value represents the number of time steps that the model analyses in the past in order to forecast the number of departures and arrivals at current time $t = 0$. Therefore, with $\alpha = 3$ the model looks at time series at $t = -1$, $t = -2$ and $t = -3$. The following graph schematise the configuration of the Neural Network as a time series.
Both inputs and outputs of the model are the rentals values in the parking zones, however, the variables have been expanded by the number of lags characterised by $\alpha = 3$. The framework has been replicated for all the areas, 274 unique polygons time three lags in the past with the overall of 822 input variables therefore 274 outputs.

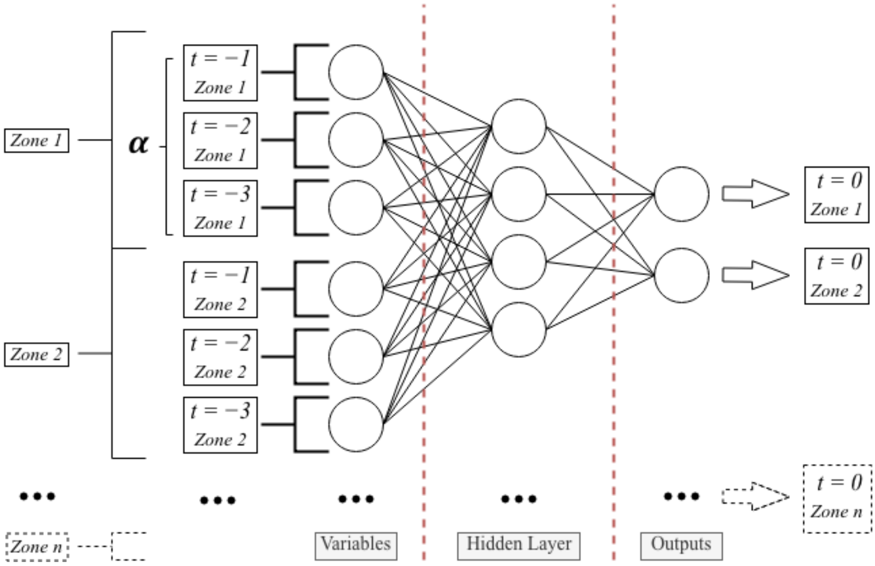


Figure 4.3: Example of spatio-temporal Neural Network.

The following definitions are the values to take into account for the size of the model to be fed into the Neural Network: Then, the number of rows and columns of the matrix for the model are:

$$Rows = p(D - (\gamma_1 + \sum_{\alpha}(\gamma_\alpha - \gamma_{\alpha-1})))$$

$$Columns = z(\alpha + 1)$$

$$Size\ Model = p(D - (\gamma_1 + \sum_{\alpha}(\gamma_\alpha - \gamma_{\alpha-1}))) \ \times \ z(\alpha + 1) \tag{4.3}$$

- $z = number\ of\ zones$
- $\alpha = Number\ of\ lags$
- $D = Number\ of\ days\ in\ the\ time\ series$
- $p = segmentation\ per\ day,\ D$
- $\gamma_\alpha\ value\ of\ the\ steps$

The columns in the model matrix stand for the number of variables and number of outputs where the variables are $V = z\alpha$ and the outputs are $O = z$.

Thank to the data handling performed, the current setting of the data-set can consider the level of detail up to the hourly basis. Indeed, the data-set reveals the total number of rentals in each zone at every hour. The time lags can be represented, therefore, on hourly basis where $\alpha = 3$ and $\gamma_1 = 1$,$\gamma_2 = 2$ and $\gamma_3 = 3$ this stands for observations at time $t = -1$ *hour* , $t = -2$ *hours* and $t = -3$ *hours*. The criteria to assign α and γ_α comes from to keep the balance between the size of the model because each additional α increases the dimension by 274 more inputs and outputs (number of zones in the data-set) and having a time lag of far from each other causes an increases the reduction of the data-set.

4.5 Model Specifications

The flow chart Figure 4.4 describes the steps for the final model specifications, it summarises the procedures in order to set up the parameters of the Neural Network. Firstly, the machines used to run the models are: (1) MacBook Pro late 2013 , with Intel Core i7-4750HQ @ 2 GHz, RAM 8.00 GB (Personal Laptop), (2) Desktop, with processor Intel Core i5-7400 @ 3 GHz, RAM 8.00 GB (Desktop at the Chair of Transportation Systems Engineering, Technical University of Munich)

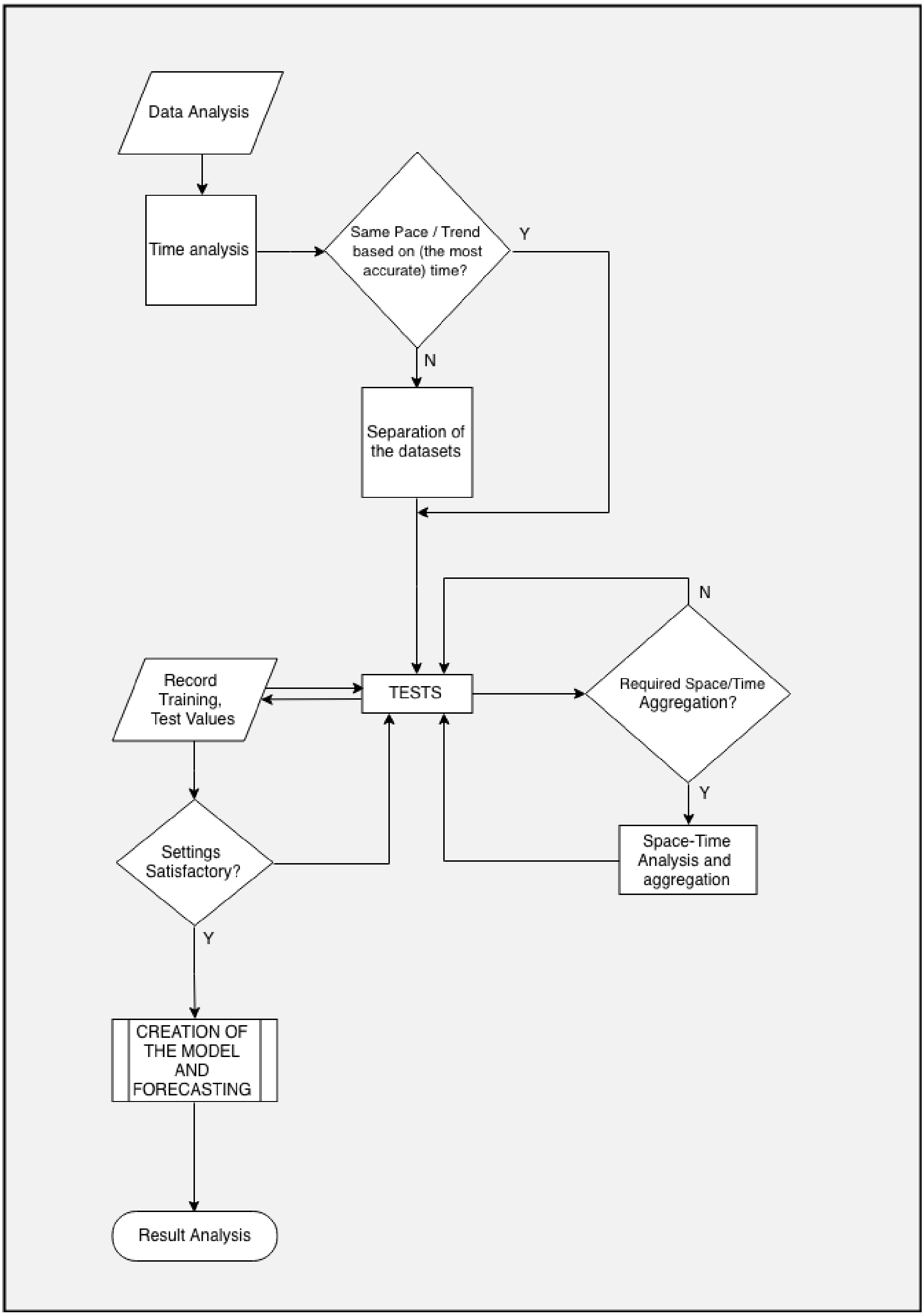


Figure 4.4: Flow chart in model specification.

The input values of this method come from the data handling and data analysis described in chapter 3. The first step is to understand the usage of the cars in different time each day, then, test at the highest level of detail in the given data-set. If it necessary to aggregate the information, the process analyses time and space and performs the aggregation accordingly to the criterion discovered in the analysis.

To begin with, the charts below (Figure 4.5, Figure 4.6) draw the attention to the different pace of the rentals during the day between the working days (Monday to Friday), Saturday and Sunday. Therefore, it has differentiated the model into three singular models: ”working days” model, ”Saturday” model and ”Sunday” model. Although Saturday and Sunday are both not working days the charts highlight the different peak and off peak times between these two days. Indeed, the city case study is Munich and it bears mentioning that during Sundays and holidays most of the shops are closed, so this regulation generates different purpose trip on Saturday and Sunday. The box-plot in Figure 4.6 clearly points out the different values in the different day, in particular during the weekend where Saturday and Sunday have a different volume of rentals compared to the mean working days.

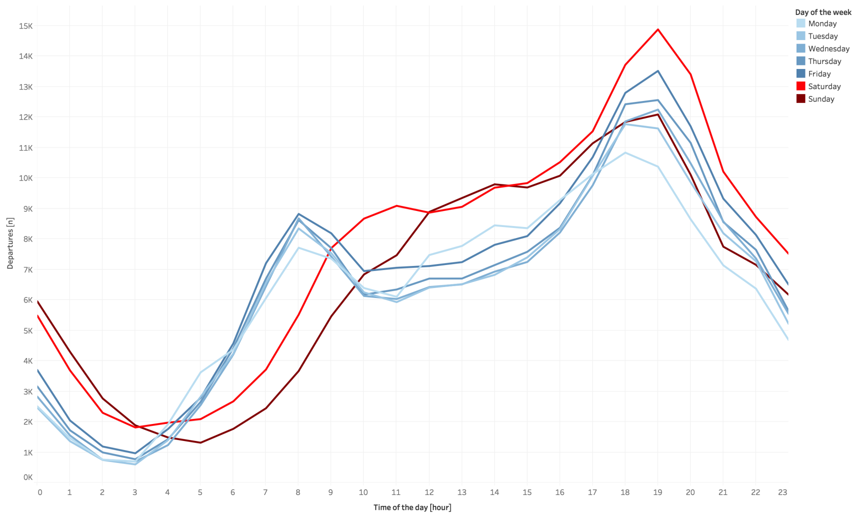


Figure 4.5: Departures according to the time of the day - DriveNow, 2016.

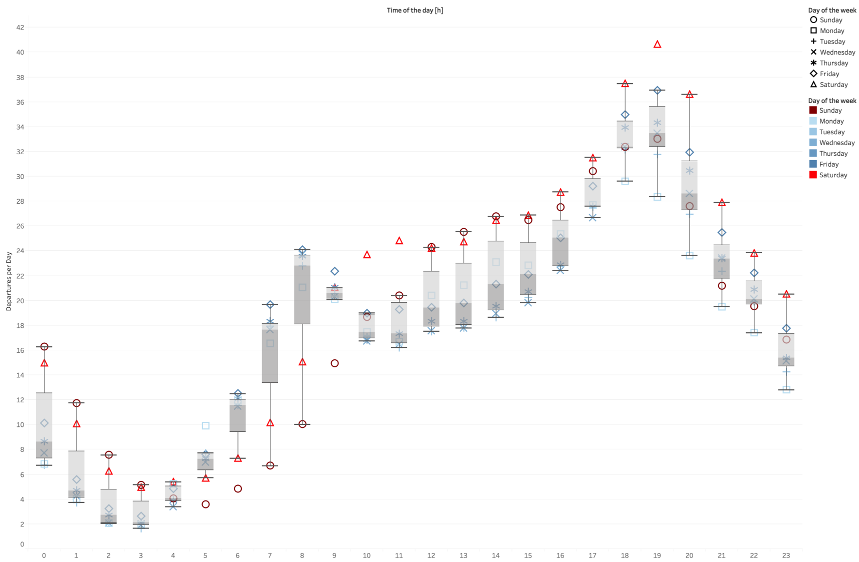


Figure 4.6: Box-plot, Average hourly-daily departures of DriveNow in 2016.

4.5.1 Space and time aggregation

In addition, it has processed the data-set with a space and time aggregation. The reshape of the inputs can be driven either by having a different target for the forecast or by trading off between level of accuracy and computational time. In the first instance, it has aggregated the hourly values of each model as the different time period. The "working days" scenario has two clear peaks time, it has then grouped the departures by "Peak/Off-Peak" for each day; between 07:00 - 10:00 "Peak Morning", 17:00 - 20:00 "Peak Evening", from 21:00 to 06:00 "Off-Peak Evening" and from 11:00 - 16:00 "Off-Peak Morning". Regarding the Saturday scenario, the division of periods has Peak-time from 09:00 to 21:00, and 22:00 - 08:00 Off-Peak and Sunday has 10:00-20:00 of "Peak" and 21:00 - 09:00 of "Off-Peak".
The configuration of the Neural Network takes into account a triplet of nodes for each zone of data-set. The input has 274 overall individual geographical locations that belong to the parking regulation of the city of Munich. The aim of the space aggregation is to give a broader, general understanding of the trend for each neighbourhood therefore, it has created 21 sets corresponding to the geographical borders of each Munich district. It has overlapped the business zone of DriveNow on the neighbourhoods of Munich map, creating a new map with new density for each new macro-zone. The result of the forecast has then considered these aggregated zone in the time period defined. However, in order to keep the same level of detail, the result of the prediction for the district has been allocated to each sub-zone based on the share of departures in the neighbourhood.

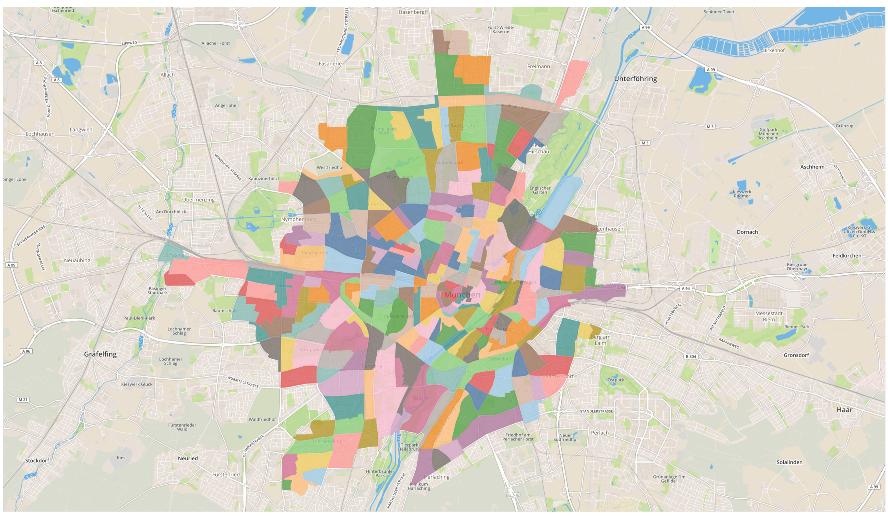


Figure 4.7: Zone Configuration based on parking regulation.

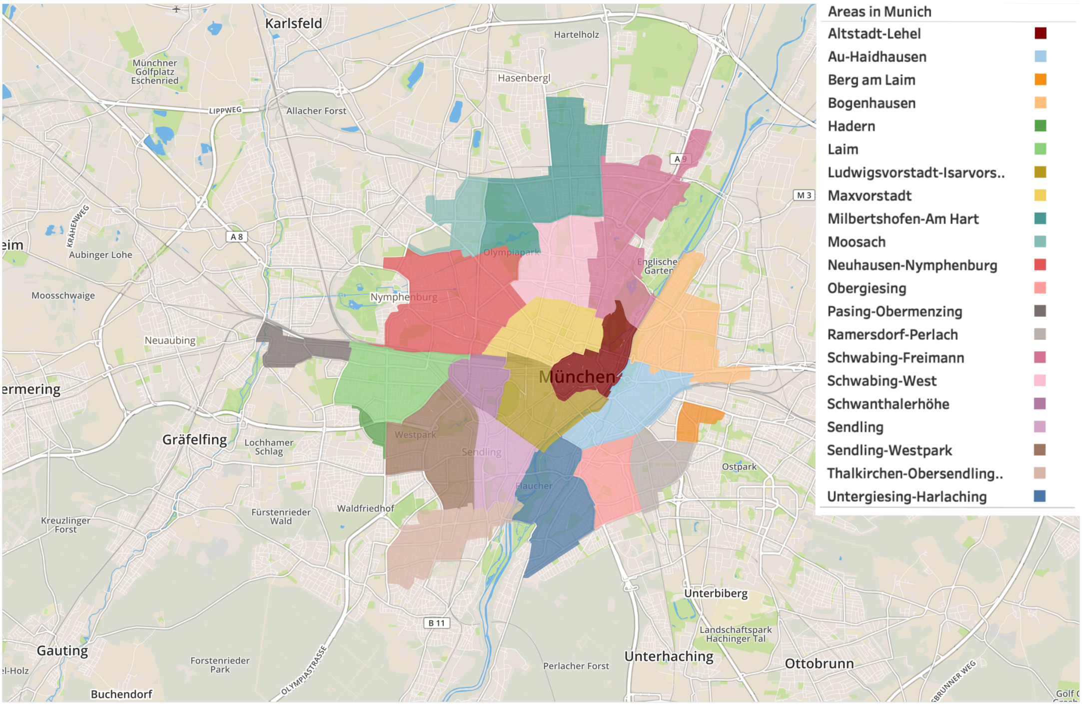


Figure 4.8: New Zone Aggregation based on districts (in German Stadtbezirke).

[2]According with the geographical information available in the data-set

4.5.2 Variable selection

The final configuration of the model is represented by 63 inputs and 21 outputs. The variable selection has taken into account the macro-areas of the city of Munich and has aggregated according to space and time period. Each input node represents the number of departure/arrival in a given zone during the time period. The number of lags α has been set equal to 3 and it has selected a continuous interval of γ. The images Figure 4.9,Figure 4.10 summarise final variable-output configuration. The matrix has been also normalised in order to avoid the mathematical instability issue, it has used the Feature Scaling method:

$$N = \frac{x - x_{min}}{x_{max} - x_{min}}$$

The normalisation has standardised the range of the values of the rentals and has contributed to a better computational time for the training process. According to Orr and Müller (1998), having a high alteration of mean value from the zero value

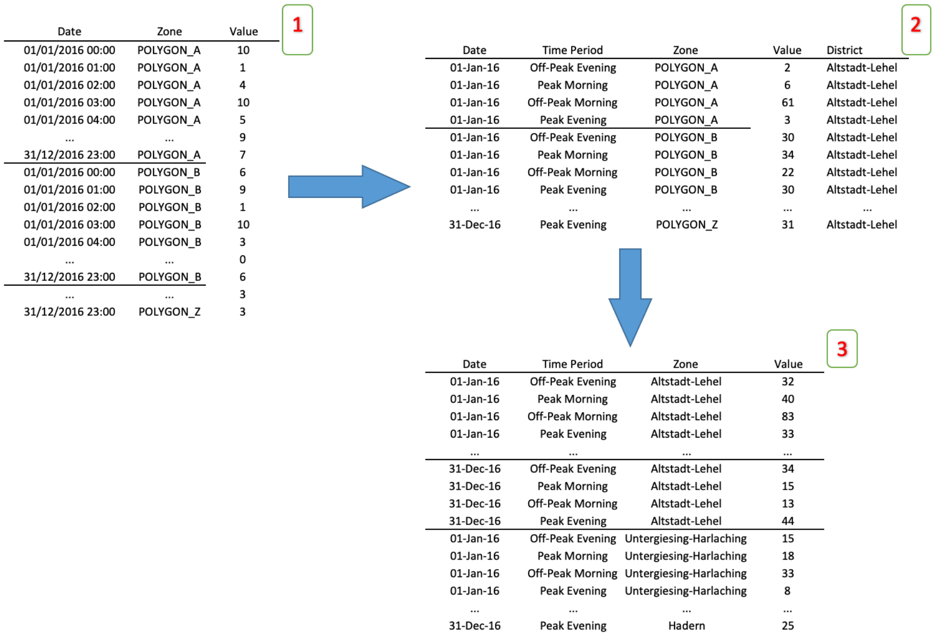


Figure 4.9: Process of space and time aggregation.

slows the algorithm with higher time for converging. Therefore, ”it is good to shift the inputs so that the average over the training set is close to zero”. Table 4.1 shows the impact of this process during the training phase and confirms what described before. Indeed, it has performed different test running on the same machine to select the best configuration for the neural network described in the next subsection. Both tests have the same configuration in terms of inputs-outputs, learning rate and activation function. However, the data-set normalised test has been shut down when the duration time reached ten times higher compared to the normalised test.

Table 4.1: Normalisation effects on computational time.

Area	Dimension Matrix	Step-Max	Variables-Outputs	Normalisation	Training Time [hours]
Macro-Areas	1032x84	$1e^6$	61-23	Yes	1.973
Macro-Areas	1032x84	$1e^6$	61-23	No	19.921

INPUTS										OUTPUTS			
Altstadt-Lehel			Untergiesing-Harlaching			...	Hadern			Altstadt-Lehel	Untergiesing-Harlaching	...	Hadern
V1	V2	V3	V4	V5	V6	...	V61	V62	V63	O1	O2	...	O21
17	4	9	20	8	12	...	14	20	7	12	3	...	6
14	21	19	8	11	4	...	16	9	22	9	19	...	7
19	21	20	16	16	5	...	4	3	19	13	2	...	2
1	13	5	5	14	19	...	4	3	12	19	20	...	2
17	3	5	6	15	0	...	20	20	18	17	0	...	9
21	4	21	2	3	19	...	10	6	3	10	16	...	2
0	17	0	18	21	8	...	20	1	0	10	14	...	8
16	10	8	15	15	19	...	15	20	18	10	11	...	16
17	17	22	3	8	4	...	18	20	0	20	21	...	2
19	3	22	7	5	9	...	5	9	10	16	0	...	5
18	20	4	6	18	16	...	7	2	11	10	19	...	7
12	9	21	3	13	1	...	7	15	2	2	21	...	22
15	21	0	20	1	13	...	6	7	12	8	19	...	14
13	19	20	1	5	2	...	1	14	13	18	9	...	2
13	3	3	2	15	9	...	16	20	10	7	3	...	2
22	2	18	6	13	5	...	1	8	3	8	7	...	2
5	16	3	16	20	5	...	11	20	6	9	15	...	15
11	15	3	3	22	21	...	4	15	8	2	18	...	7
7	1	6	4	12	18	...	0	20	6	21	22	...	19
8	1	20	5	9	12	...	10	7	9	7	12	...	20
8	0	21	9	6	7	...	15	8	16	22	13	...	14
17	10	6	2	18	4	...	1	14	11	22	16	...	18
20	11	20	13	11	10	...	8	16	5	3	12	...	16
19	1	5	1	0	12	...	5	21	17	14	13	...	19
21	20	11	22	3	5	...	4	9	17	19	20	...	16
19	10	8	9	11	20	...	22	16	7	18	8	...	11

Figure 4.10: Configuration of the Input/Output of the Neural Network.

4.5.3 Neural Network settings for the test

This section describes the settings for the Neural Network, it is going into the detail of each parameter used. The tests have been evaluated based on the following criterion: the duration of the training time, compared with the goodness of the model measured by the root squared mean error (RMSE). this is the measure that reflects how close the prediction matches the historical data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2}$$

.
First of all, the model has been run in the R programming language, using the development environment of RStudio. The package utilised specifically for the model is called "neuralnet" Stefan Fritsch, Frauke Guenther. The package lets the user adjusts the settings of activation function for example the number of layers number of neurons and many more (Fritsch et al., 2016). In order to choose the most important settings for the prediction the following tests have been performed:

- Number of layers,
- Number of neurons for each layer
- The "stepmax"

The rest of the other parameters has been left as default. Regarding the test of the the maximum steps for the training of the neural network ("Stepmax") (Fritsch et al.,

2016) it has incremented from ten by the power of ten each run with a configuration of 56 neurons in one single layer (rule of thumb). The model did not converged until reaching the threshold of $1e6$. Therefore, this is the minimum number where the model is stable and provides the minimum computational time, hence it has selected for this number for the other tests performed. The number of neurons and the number of layers that characterise every Neural Network model have been selected through a sensitivity analysis. In order to choose the number of neurons, it has tested the rule

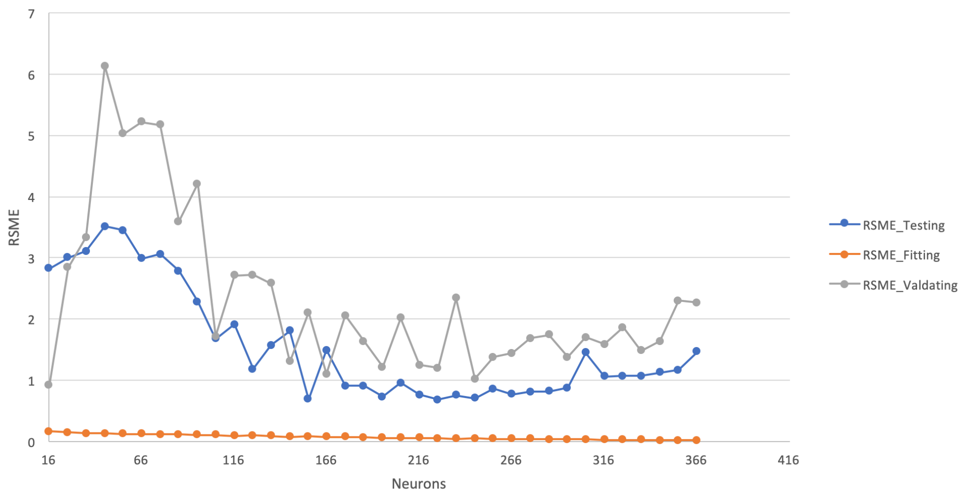


Figure 4.11: Sensitivity Analysis Neurons.

of thumb which suggests that "the number neurons in layer in between the inputs and the outputs of a Neural Network should be equal to the two-third of the sum of inputs and outputs". This would recommend to use a layer with 56 neurons. In order to have the best configuration it has tested it through a sensitivity analysis, it has processed the training phase of the Neural Network with a set of 36 different neuron configurations (from 16 to 366 neurons, increase by 10 neurons per step) and it has recorded the computational time and the quality of the prediction for the training, validation, prediction phases. The result of the test has been plotted as Figure 4.12 and the values are in Table A.2. The rule of thumb (the two-third of inputs of outputs) has not been confirmed, in fact the graph shows that by adding neurons the error is decreasing. Between 116 and 166 neurons the error is tilting until it stabilises. Furthermore, the trend of the error for more than 244 neurons starts to increase, thus adding additional neurons is counter-productive.
The output of the sensitivity analysis suggests to use a 196 neurons configuration. Furthermore it has tested whether creating a multi-layer Neural Network would bring the beneficial results to the forecast. In agreement with Aggarwal (2018) and to the best

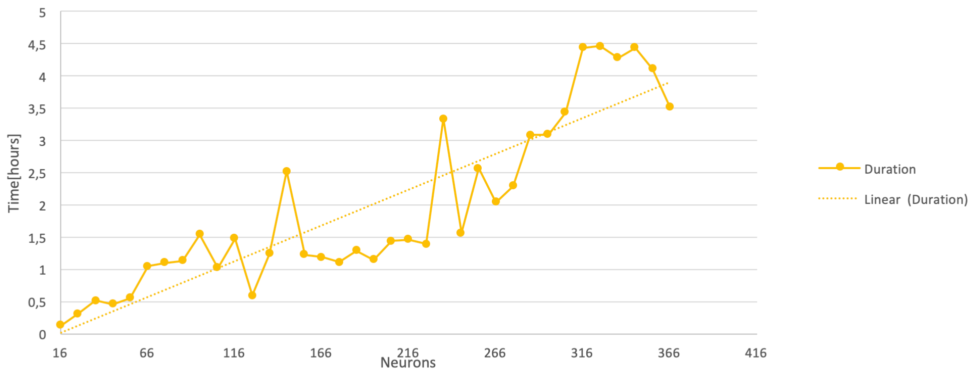


Figure 4.12: Sensitivity Analysis Time.

of the author knowledge the multi-layer configurations are mostly used when problems and patters are for very high complexity such as image recognition. However, it has simply tested whether a feed-forward network brings about some improvements in the case study. Results (Table 4.2) have then proved that for this case study the

Table 4.2: Multi-layers test .

# neurons	Test Duration	RSME Testing	RSME Fitting	RSME Validating
76	02:03	2.052658879	0.119030076	4.403660273
2 Layers x76	03:59	2.055645009	0.112261776	8411680638

multi-layer Neural Networks does not improve the outcome of the forecasting.
In the next step, it has analysed the best activation function configuration. The default function in the configuration of the "neuralnet" package is the *tanh*. According to Aggarwal (2018) activation functions are correlated with the type of problem for example the *ReLU*, $hard - tanh$ and *Sing* functions are more suitable for the image recognition when the object of the output is binary. The *Sigmoid* is better for probability outputs. The *tanh*, therefore, has been found the most suitable for the problem that is trying to resolve. Moreover, this section analyses the phases of training, validation and prediction of the model. To begin with, Figure 4.13 shows how the data-set has been structured for the preparation of the model. The model aims to forecast the last week of the year. Therefore, the part dedicated to the prediction has been split from the rest of the data-set. The remaining records has been again sub-set into Testing(2) and Training(1). The testing share, has the 20% of the overall data-set and the rest 78% is dedicated to the training phase. Furthermore, in Figure 4.13 the red lines represent the different composition of the testing phase for

the K-fold cross-validation. The k-fold cross-validation is one of the model validation technique, that is used to evaluate how the results of the training phase generalise an unseen data-set (James et al., 2000). The method follows the steps:

1. Division of the data-set for training and testing equally sized in k sub-sets
2. Selection of the first $Test_i$ sub-set as Testing set
3. The rest of the data-set is considered as training, $Train_k$
4. Run the training and the testing phase for the $Test_k$ and $Train_k$
5. Repeat the process for all the k divisions

The benefit of this approach is that all the records of the data-set are both used for the testing and training phases. Thus, the cross-validation can detect if in the data-set there are sections that can determine overfitting or selection bias.
It has then performed the k-fold validation with the k = 5. Table 4.3 summarises the result of this process. The results of the process demonstrated that regarding the training phase the data do not affect the evaluation of the error as much as the in the testing and in the validating. The k = 5 represents the bottom part of the data-set and the last part of the year accordingly.

Table 4.3: K-fold cross validation with configuration of 56 neurons.

K	Duration	RMSE_Fitting	RMSE_Testing	RMSE_Validating
1	00:58	0.125141713	3.03016833	2.942765162
2	00:53	0.125020337	3.196910127	3.353348993
3	00:55	0.122507666	3.457813666	4.15626356
4	00:54	0.124029123	3.904821734	3.739422504
5	00:57	0.124624483	3.946262758	3.866927487

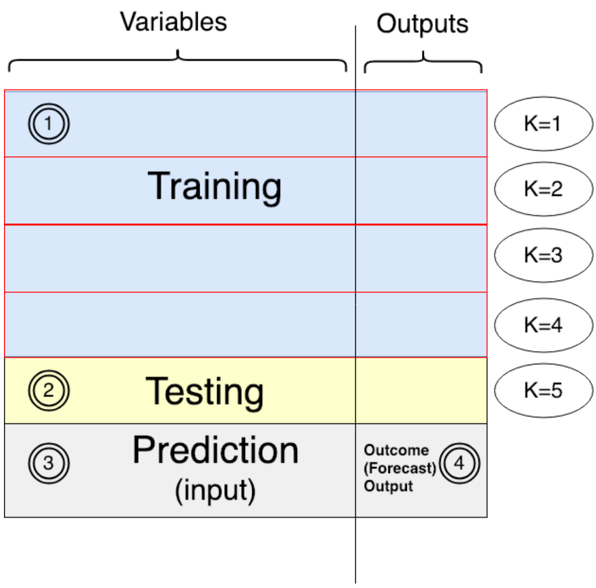


Figure 4.13: Proportion data-set between the training, testing and forecasting.

To sum up the modelling phase has conducted the several tests to implement the model specification, first it has conceptualised the problem into:

- 3 Neural Networks according to the pace of the departures on Working Days, Saturday and Sunday.
- The input and output matrices take into account the aggregation of the zones based on the district distribution The input matrix considers the aggregation of the departures accordingly to peak-off-peak period per day of the week.

Then the tuning the model, into the following model specifications.

- Single-Layer Neural Network with 196 Neurons
- Repetition Step-max at $1e6$
- Activation function $tanh$

CHAPTER 5

Results Summary

This chapter summarises the result obtained by the Neural Network. It has run three different models according to the day-of-the-week scenarios discussed in the previous chapter. Table 5.1 displays the result of the prediction of the model on the working days for the last week of the year 2016. The prediction of the model has configured

Table 5.1: Summary Results Working Days Neural Network model.

Time-Period	Day/AREA	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21
Off-Eve	26/12/2016	4	0	0	3	-11	-12	8	-9	1	0	9	7	20	9	4	-4	-5	-1	-5	9	-9
Peak-Morn	26/12/2016	11	0	-9	18	-2	7	0	25	26	-1	-1	5	6	17	30	20	-12	-2	9	2	6
Off-Morn	26/12/2016	7	10	8	5	2	6	6	15	5	3	16	10	-3	5	17	3	7	2	0	10	2
Peak-Eve	26/12/2016	7	10	-1	6	-2	0	12	16	13	4	6	8	-4	7	18	0	20	11	4	8	-10
Off-Eve	27/12/2016	-22	8	18	-4	-4	-2	12	20	20	9	9	23	-8	-14	-1	12	13	8	-6	1	2
Peak-Morn	27/12/2016	14	3	-2	-1	-1	-1	3	13	11	17	4	6	-2	-2	12	14	8	-5	-4	-8	-8
Off-Morn	27/12/2016	-2	15	4	14	7	21	32	29	28	-8	3	12	5	-2	18	14	-5	-10	4	-5	-2
Peak-Eve	27/12/2016	8	2	8	7	-4	13	-2	4	-3	-10	7	14	-2	2	0	-8	14	-3	-4	12	1
Off-Eve	28/12/2016	-32	-3	14	-5	8	47	11	-15	-1	50	-34	19	-7	19	-47	31	-16	10	-42	-43	71
Peak-Morn	28/12/2016	17	-24	-9	-10	6	12	23	17	-14	24	25	-23	32	7	-19	34	-12	-15	-4	13	12
Off-Morn	28/12/2016	4	15	-2	4	7	1	-3	2	-2	-4	16	4	-1	-11	13	-4	-9	2	21	-1	3
Peak-Eve	28/12/2016	-20	6	5	6	9	8	-27	-6	-2	39	28	-8	-6	-10	31	-3	-9	15	11	10	18
Off-Eve	29/12/2016	15	3	12	33	-3	15	47	2	37	-2	10	33	11	10	-39	37	-13	4	7	38	43
Peak-Morn	29/12/2016	9	-14	-12	9	4	22	42	5	7	9	19	22	-1	29	-11	-1	-4	25	-17	10	10
Off-Morn	29/12/2016	5	-3	5	6	5	-1	23	9	9	0	18	10	-3	5	-7	8	-12	8	16	-5	-1
Peak-Eve	29/12/2016	15	-21	-25	1	-6	-13	51	53	12	-118	-82	82	35	45	-50	-29	-57	-3	3	81	121
Off-Eve	30/12/2016	-20	19	0	5	4	34	18	-19	-20	1	-18	24	1	20	-1	2	13	-7	-38	4	23
Peak-Morn	30/12/2016	11	5	7	4	3	8	15	15	-2	1	19	6	1	4	8	2	3	0	3	-3	0
Off-Morn	30/12/2016	-7	8	3	2	-1	0	16	19	4	-3	22	0	8	4	6	4	0	-2	7	0	0
Peak-Eve	30/12/2016	-11	-10	-41	-11	-10	9	3	11	19	0	-17	31	22	-1	31	15	-22	13	9	10	4

a single-layer Neural Network with 196 neurons in between. The results of the model are in terms of goodness of the prediction (RSME) are:

- Training phase = 0.728
- Testing phase = 0.060
- Validating phase = 1.218

The results are acceptable for the prediction step with 1.218 RSME. Moreover, the results for the second model are in Table 5.2. The Saturday's and Sunday's models

Table 5.2: Summary Results Saturday Neural Network model.

Time-Period	Date/AREA	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21
Off-Peak	31/12/2016	-19	-45	-55	19	49	81	-37	25	-93	41	66	58	-64	55	-48	73	82	-20	-110	-49	-42
Peak	31/12/2016	-68	43	59	-24	53	-83	-28	-81	95	-72	16	-12	2	37	-140	128	30	-114	-128	150	-114

have an error of around 0.005 on the training phase, however the error in the prediction is 4.684 and 5.326. This outcome is probably due to the overfitting of the model.

Therefore, it is not suggested to use this configuration for the demand forecasting of the weekends. The results of the forecasting analysis of this model have turn out that the data-set reshaped for this time-frame is too little if considering the time frame of a single year (52 Saturday and 53 Sundays). Another approach on the level of forecasting can be then taken into account. Keeping the original data-structure of hourly basis and have two degree of forecasting for Car-sharing services, the broad more generic considering the peak-period of the working time and one extremely detailed forecasting for each hour of the weekend. In addition, the results

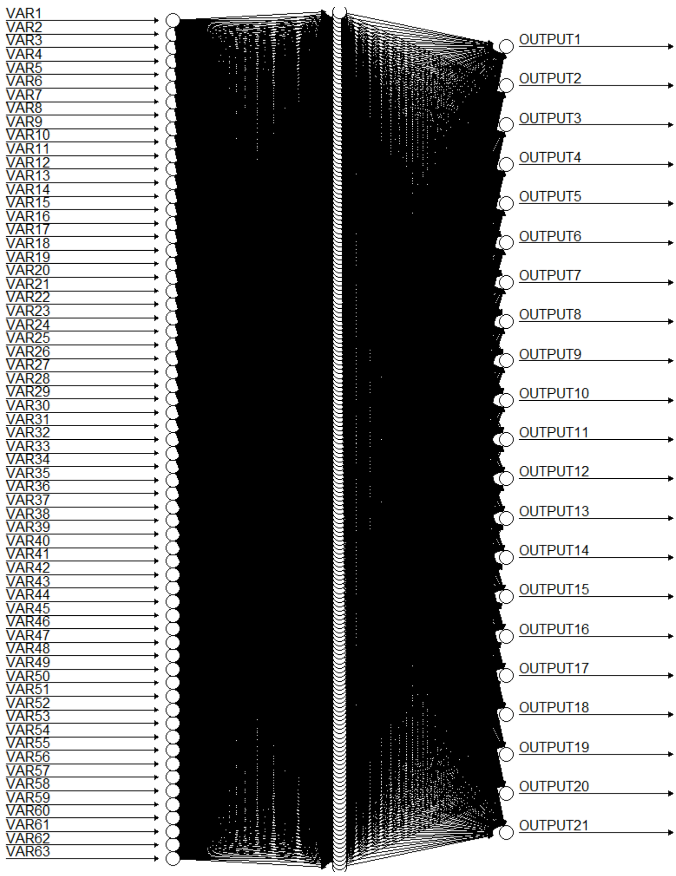


Figure 5.1: Representation of the Model created, larger resolution in Figure A.4.

of the Neural Network have been compared with a traditional time-series forecasting method: ARIMA. The acronym stands for Autoregressive integrated moving average. It has set the ARIMA parameters (p,d,q) in accordance to the same configuration of the Neural Network. Therefore single-Layer Neural Network model created has been compared with an ARIMA(3,0,0). The limitation of the ARIMA model is that it cannot forecast considering the whole business network of DriveNow, because each forecast is independent with the area of bookings. Regarding the precision of the error, the Neural Networks are more precise in the training phase with the tendency to overfitting where the number of the observations are not high enough. The ARIMA model has recorded an RSME in the prediction between 1,381 and 0,197. This results can be compared with the Neural Network results although in some situation the ARIMA performs better, Neural Network has the stability to provide one single error for the entire network. Meanwhile, ARIMA models that can predict singularly either better or worse Neural Network.

CHAPTER 6

Conclusion

6.1 Conclusions

The research has studied the case-study of DriveNow in Munich in the year 2016. The object of this study has dealt with first spacial-temporal data-analysis, then the tuning and the building of a Machine Learning algorithm that can predict the results of the future bookings in the car-sharing systems. The findings of the data-analysis are that the efficiency of the network is highly subjective with the time in a day, late afternon and evening hours have concentrated the majority of the trips and scores the best efficiency in the fleet. Summer seasons are more profitable for car-sharing providers (Shaheen and Cohen, 2012), it has demonstrated that this is the same case for the DriveNow's Munich Network. Regarding the strategical zones, airport is the most valuable area in the whole network. It has also revealed that satellites in the network are very attractive points for the customers, however, can represent further challenges in terms of re-balancing. The method presented generalises the structure of time-series as a Neural Network, and it applies it to the case study. The results reveal a quite good estimation of the future departures. The research is reinforced by comparing the forecasting model with a traditional time-series forecasting model. Overall the Neural Networks can be a useful tool for the forecasting of time-series. Moreover, biggest competitive advantage of Neural Networks is that it is able to consider the whole business area for car-sharing services as a whole system. The forecast then, takes into account an additional data that is not given: the relationships between each area in the network. On the other hand, classical ARIMA must be done by each area and the information of the interconnection in the business area are lost. Regarding the goodness of the forecast Neural Networks tend to overfit the training phase, when the number of the observations are not high enough. In the research then the models of Saturdays and Sundays are one of this example. Therefore the following statements are the answers to the research question:

1. **Yes, Neural Networks have a high performance in the forecasting of time-series however the method must be contextualised within the disposal of high amount of observation.**

2. **Neural Network can preform better when considering the car-sharing system as a whole, meanwhile traditional time-series forecasting must take into account each sequence within the same area.**

6.2 Recommendations and future works

This study has mainly dealt with the bookings of DriveNow considering only the bookings. Recent inquires to the car-sharing operator has revealed that the values of AVG km and AVG min can be linked with the number of the departures. Whereas the triplet of Distance, Time and Number of bookings contains a "1" under the column "Number of Starts" the trip can be considered as an historical trip. It has calculated the amount this characteristic: the 69,8% of the data-set is composed by a triplet that contains historical trip. Therefore the major recommendation for the continue of this study is to utilise this highly valuable information that can develop several future works that research into travel behaviour analysis.

Bibliography

(2018). Parking Redistribution Muenchen. Technical report.

Aggarwal, C. C. (2018). *Neural Networks and Deep Learning*.

Amey, A., Attanucci, J., and Mishalani, R. (2011). Real-Time Ridesharing: Opportunities and Challenges in Using Mobile Phone Technology to Improve Rideshare Services. *Transportation Research Record: Journal of the Transportation Research Board*, 2217(1):103–110.

Bohte, S. (2018). *Artificial neural networks as models of neural information processing*.

Boldrini, C. and Bruno, R. (2017). Stackable vs Autonomous Cars for Shared Mobility Systems : a Preliminary Performance Evaluation.

Bunderinstitut für Bau-, S. u. R. (2015). Laufende Stadtbeobachtung Raumabgrenzungen. Technical report.

Bundesverband CarSharing eV (2018). Current figures and data for CarSharing in Germany _ bcs Bundesverband CarSharing eV.

Codd, E. (1990). The relational model for database management.

Cohen, A. and Shaheen, S. (2016). Planning for shared mobility. Technical Report 583.

Cohen, A. P., Shaheen, S. A., and McKenzie, R. (2008). Carsharing : A Guide for Local Planners.

DASU, T. and JOHNSON, T. (2003). *Exploratory Data Mining and Data Cleaning*.

Delbosc, A., Currie, G., Delbosc, A., and Currie, G. (2013). Causes of Youth Licensing Decline : A Synthesis of Evidence Causes of Youth Licensing Decline : A Synthesis of Evidence. 1647.

DriveNow (2018). Inquieries DriveNow comunications.pdf.

Durán Rodas, D. and Constantinos Antoniou Emmanouil Chaniotakis, U.-P. (2017). Identification of spatio-temporal factors affecting arrivals and departures of shared vehicles. Methodology for comparing multiple cities on a local level using open-source data.

Eurostat (2016). Passenger cars in the EU. Technical Report April.

Ferrero, F., Perboli, G., Rosano, M., and Vesco, A. (2018). Car-sharing services: An annotated review. *Sustainable Cities and Society*, 37(September 2017):501–518.

Firnkorn, J. and Müller, M. (2011). What will be the environmental effects of new free-floating car-sharing systems? The case of car2go in Ulm. *Ecological Economics*, 70(8):1519–1528.

Fritsch, S., Guenther, F., Suling, M., and M. Mueller, S. (2016). Training of neural networks. Package 'neuralnet'. The R project for statistical computing. page 13.

Giesel, F. and Nobis, C. (2016). The Impact of Carsharing on Car Ownership in German Cities. In *Transportation Research Procedia*, volume 19.

Guijarro, F. (2015). Analysis of the Integration of Carsharing Interoperability among Operators in the City of Munich. pages 1–89.

Haas, B. (2017). Chinese bike share graveyard a monument to industry ' s " arrogance " not enough demand.

Iglesias, R., Rossi, F., Wang, K., Hallac, D., and Sep, R. O. (2017). Data-Driven Model Predictive Control of Autonomous Mobility-on-Demand Systems.

International Energy Agency (2017). Tracking Clean Energy Progress 2017. Technical report.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2000). *An introduction to Statistical Learning*, volume 7.

Jorge, D., Barnhart, C., Homem, G., and Correia, D. A. (2015a). Assessing the viability of enabling a round-trip carsharing system to accept one-way trips : Application to Logan Airport in Boston. *Transportation Research Part C*, 56:359–372.

Jorge, D. and Correia, G. (2013). Carsharing systems demand estimation and defined operations: A literature review. *European Journal of Transport and Infrastructure Research*, 13(3):201–220.

Jorge, D., Molnar, G., Homem, G., and Correia, D. A. (2015b). Trip pricing of one-way station-based carsharing networks with zone and time of day price variations. *Transportation Research Part B*, 81:461–482.

Khan, M. and Machemehl, R. (2017). Seeing Cities Through Big Data.

Kopp, J., Gerike, R., and Axhausen, K. W. (2015). Do sharing people behave differently ? An empirical of free-floating car-sharing members. *Transportation*, pages 449–469.

Kumar, V. P. and Bierlaire, M. (2012). Optimizing Locations for a Vehicle Sharing System. pages 1–30.

Li, Q., Liao, F., Timmermans, H. J., Huang, H., and Zhou, J. (2018). Incorporating free-floating car-sharing into an activity-based dynamic user equilibrium model: A demand-side model. *Transportation Research Part B: Methodological*, 107:102–123.

Lorimier, A. D. and El-geneidy, A. M. (2013). Understanding the Factors Affecting Vehicle Usage and Availability in Carsharing Networks : A Case Study of Communauto Carsharing System from A Case Study of Communauto Carsharing System. 8318.

Lyft Team, M. (2018). Lyft Welcomes Munich Team.

Ma, T. and Antoniou, C. (2018). Statistical Learning (University Course), on the 17th April 2018 at Technical University of Munich.

Martin, C. J. (2016). The sharing economy: A pathway to sustainability or a nightmarish form of neoliberal capitalism? *Ecological Economics*, 121:149–159.

Mckerracher, B. C. (2018). Global sales Electric Vehicle outlook Outlook : 2018. page 2018.

Michelberger, P. (2018). *An End-to-End Machine Learning Approach for Predicting Passenger Demand of a Ride-hailing Service*. PhD thesis.

Montserrat, M., Pfertner, M., Hema Sharanya, R., Martin, S., and Gebhard, W. (2017). Impacts of a multimodal mobility service on travel behavior and preferences : user insights from Munich ' s first Mobility Station. *Transportation*, 44(6):1325–1342.

MVG (2015). Die Mobilitätsstation an der Münchner Freiheit , The mobility station at Münchner Freiheit. (September).

Orr, G. B. and Müller, K.-R. (1998). *Neural Networks: Tricks of the Trade, this book is an outgrowth of a 1996 NIPS workshop.*

Priester, R. and Wulfhorst, G. (2014). A generic code of urban mobility : how can cities drive future sustainable development ? *Transportation Research Procedia*, 4(89):90–102.

Prieto, M., Baltas, G., and Stan, V. (2017). Car sharing adoption intention in urban areas: What are the key sociodemographic drivers? *Transportation Research Part A: Policy and Practice*, 101:218–227.

Schmöller, S., Weikl, S., Müller, J., and Bogenberger, K. (2015). Empirical analysis of free-floating carsharing usage: The munich and berlin case. *Transportation Research Part C: Emerging Technologies*.

Selloni, D. (2017). *Codesign for the Public Interest*.

Shaheen, S., Chan, N., Bansal, A., and Cohen, A. (2015). Shared Mobility a Sustainablity and Technology Workshop: Definition, Industry Development and Early Understanding. *University of California Berkeley Transportation Sustainability Research Center*, page 30.

Shaheen, S., Rodier, C., Murray, G., Cohen, A., and Martin, E. (2010). Carsharing and Public Parking Policies: Assessing Benefits, Costs, and Best Practices in North America. page 56.

Shaheen, S., Sperling, D., and Wagner, C. (1998). Carsharing in Europe and North America: Past, Present, and Future. *Transportation Quarterly*, 52(3):35–52.

Shaheen, S. A. and Cohen, A. P. (2012). Carsharing and Personal Vehicle Services: Worldwide Market Developments and Emerging Trends. *International Journal of Sustainable Transportation*, 7(1):5–34.

Shaheen, S. A., Martin, E. W., Cohen, A. P., Chan, N. D., Pogodzinsk, M., Martin, E. W., and Chan, N. D. (2014). Public Bikesharing in North America During a Period of Rapid Expansion : Understanding Business Models , Industry Trends & User Impacts , MTI Report 12-29 Public Bikesharing in North America During a Period of Rapid Expansion : Understanding Business Mode. *Mineta Transportation Institute Report 12-29*, pages 2–3.

United Nations Department of Economic and Social Affairs (2018). World Urbanisation Prospects: The 2018 Revision. Technical report.

Wickham, H. (2017). Data Science with R by Garrett Grolemund. pages 1–30.

Willing, C., Brandt, T., and Neumann, D. (2017). Intermodal Mobility. *Business and Information Systems Engineering*, 59(3):173–179.

Xu, J.-X. (2007). A new Evolutionary Neural Network for forecasting net flow of a car sharing system. pages 1670–1676.

Yole (2015). From Technologies to Markets Sensors and Data Management for Autonomous Vehicles Table of Contents. Technical report.

APPENDIX A

Appendix

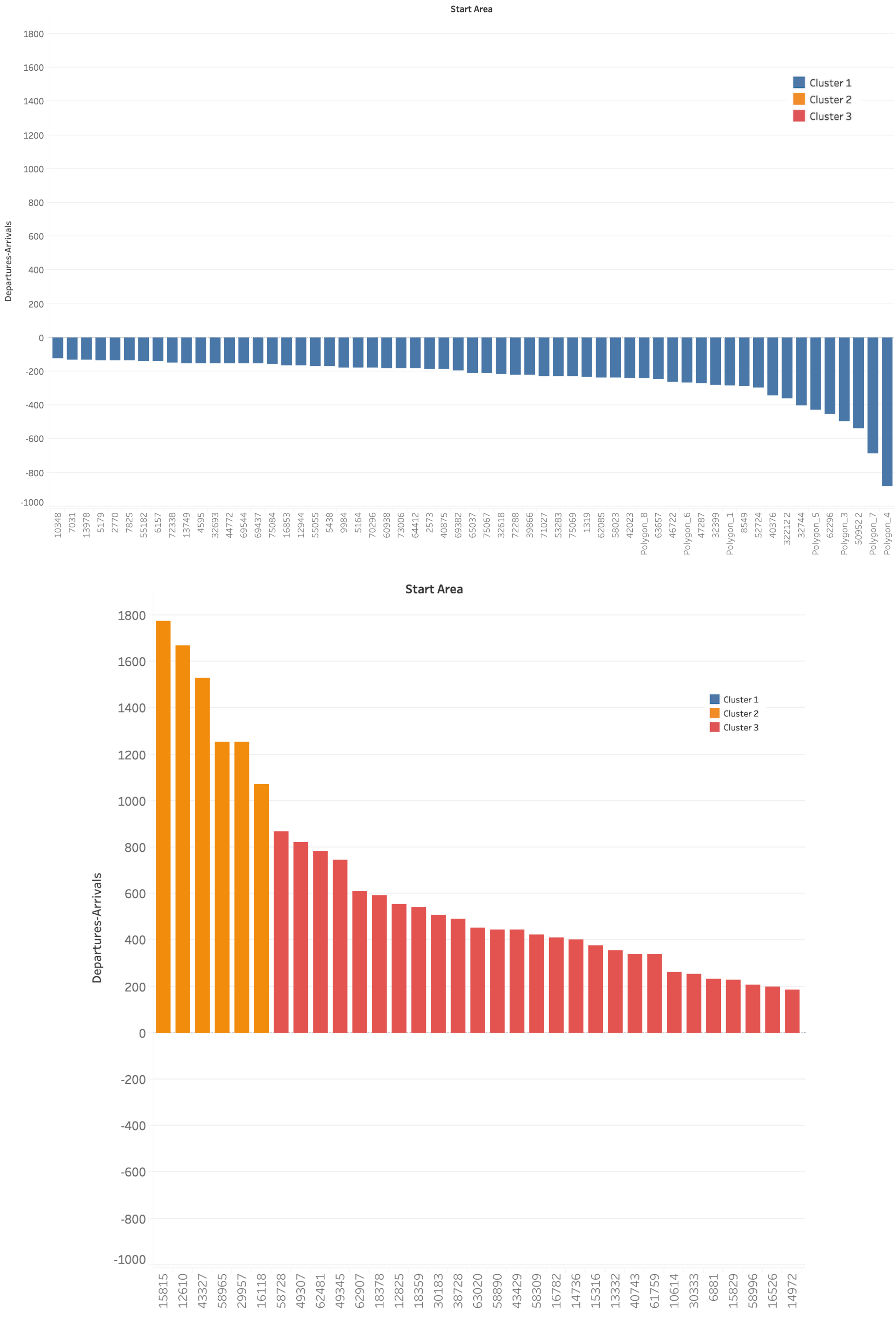


Figure A.3: Tails of the chart for the Balance of the areas.

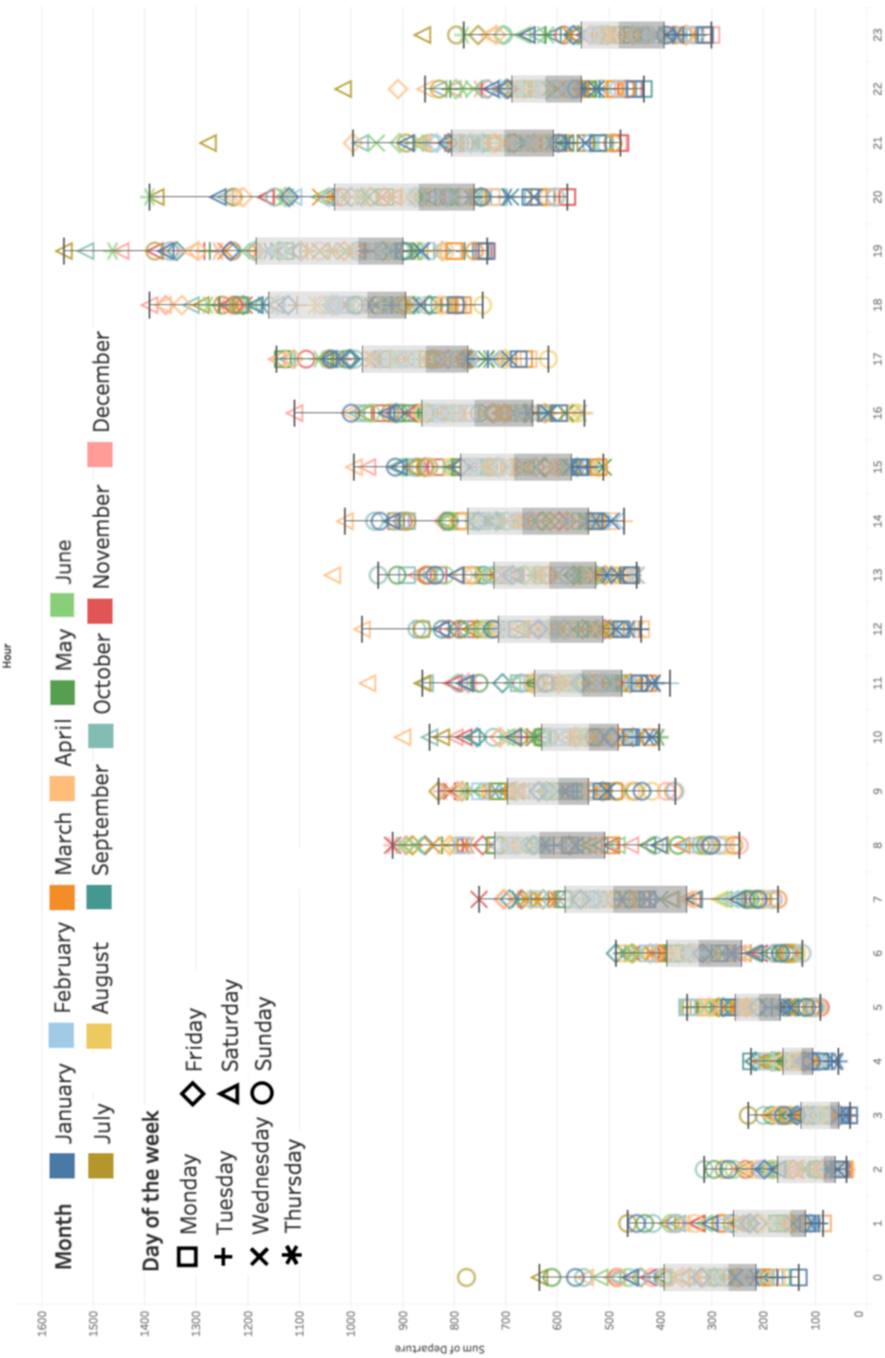


Figure A.1: Box-plot departures-hourly-monthly-weekly .

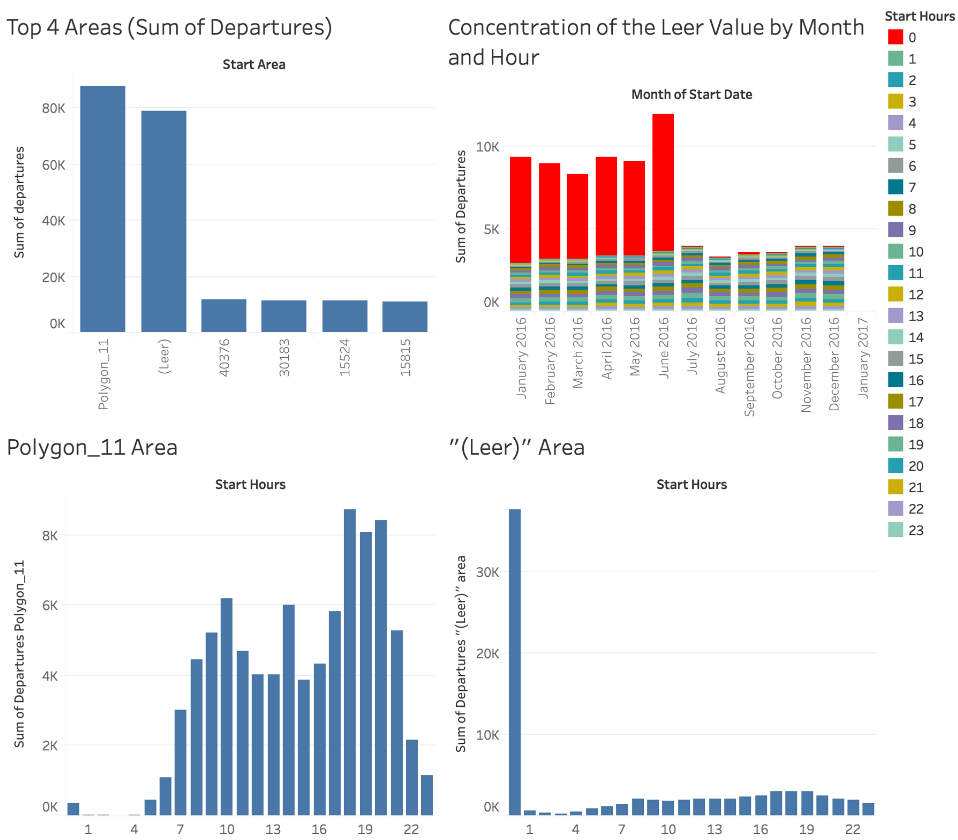


Figure A.2: Leer values outliers.

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
Polygon_11	87555	87401	154	7,891520547	7,877100611	1
40376	11642	11988	-346	1,049318511	1,080430225	1
30183	11516	11008	508	1,03796186	0,992106767	3
15524	11437	11334	103	1,03084142	1,021487836	1
15815	11224	9451	1773	1,011643271	0,851780619	2
62296	9982	10438	-456	0,899699139	0,940734959	1
16526	9813	9614	199	0,884466805	0,866471153	3
39866	9764	9987	-223	0,88005033	0,900088143	1
44749	8857	8860	-3	0,798300468	0,798516166	1
4373	8640	8727	-87	0,778741791	0,786529411	1
62907	8635	8026	609	0,77829113	0,7233511	3
29957	8465	7213	1252	0,762968665	0,65007868	2
14736	8452	8051	401	0,761796947	0,72560425	3
62481	8433	7652	781	0,760084436	0,689643984	3
58996	8421	8215	206	0,75900285	0,74038491	3
63020	8372	7919	453	0,754586375	0,713707621	3
18378	8307	7714	593	0,748727785	0,695231795	3
58023	8170	8410	-240	0,73637968	0,757959476	1
58728	8130	7262	868	0,732774394	0,654494853	3
15437	8122	8141	-19	0,732053337	0,733715588	1
Polygon_3	8096	8593	-497	0,729709901	0,77445253	1
Polygon_7	8022	8709	-687	0,723040121	0,784907143	1
64529	7975	8101	-126	0,71880391	0,730110549	1
18359	7814	7274	540	0,704292634	0,655576365	3
43429	7645	7203	442	0,6890603	0,64917742	3
Polygon_4	7604	8485	-881	0,685364882	0,764718924	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
32744	7586	7988	-402	0,683742503	0,719926313	1
58890	7345	6902	443	0,662020655	0,622049501	3
58085	7298	7381	-83	0,657784444	0,665219844	1
53283	7295	7526	-231	0,657514047	0,678288111	1
16782	7293	6884	409	0,657333783	0,620427233	3
Polygon_5	7149	7579	-430	0,644354753	0,683064788	1
32693	7098	7251	-153	0,639758013	0,653503467	1
49345	7085	6340	745	0,638586295	0,571398701	3
58965	7046	5791	1255	0,635071141	0,521919539	2
11729	7000	7028	-28	0,630925062	0,633405374	1
15901	6935	6991	-56	0,625066472	0,630070713	1
75084	6694	6851	-157	0,603344624	0,617453076	1
10709	6679	6798	-119	0,601992642	0,612676399	1
12944	6674	6842	-168	0,601541981	0,616641942	1
15316	6659	6282	377	0,600189999	0,566171394	3
49307	6503	5682	821	0,586129383	0,512095808	3
14816	6498	6359	139	0,585678722	0,573111095	1
38728	6453	5961	492	0,581622775	0,537240955	3
33055	6419	6382	37	0,578558282	0,575183992	1
61759	6296	5960	336	0,567472027	0,537150829	3
11061	6273	6365	-92	0,565398988	0,573651851	1
33079	6257	6335	-78	0,563956874	0,570948071	1
10814	6215	6280	-65	0,560171323	0,565991142	1
30154	6137	6113	24	0,553141015	0,550940104	1
5043	6094	6169	-75	0,549265333	0,555987159	1
12610	6015	4345	1670	0,542144893	0,391597375	2

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	**Value Start**	**Value End**	**Diff. START - END**	**% OF DEPARTURES**	**% OF ARRIVALS**	**Cluster**
13332	5808	5455	353	0,523487537	0,49163721	3
71027	5802	6030	-228	0,522946745	0,543459648	1
16239	5747	5576	171	0,517989476	0,502542454	1
58744	5716	5756	-40	0,515195379	0,51876513	1
18045	5683	5653	30	0,512221018	0,509482154	1
6398	5634	5729	-95	0,507804543	0,516331728	1
6157	5577	5717	-140	0,50266701	0,515250217	1
62951	5568	5532	36	0,501855821	0,498576911	1
44772	5451	5605	-154	0,491310359	0,505156107	1
63048	5370	5496	-126	0,484009655	0,495332376	1
17991	5320	5443	-123	0,479503047	0,490555699	1
11133	5314	5438	-124	0,478962254	0,490105069	1
43516	5294	5400	-106	0,477159611	0,486680282	1
27534	5257	5323	-66	0,473824722	0,479740581	1
55055	5224	5396	-172	0,470850361	0,486319778	1
4595	5217	5369	-152	0,470219436	0,483886376	1
12825	5204	4649	555	0,469047718	0,418995672	3
72288	5188	5410	-222	0,467605603	0,487581541	1
41077	5184	5244	-60	0,467245075	0,472620629	1
30333	5175	4922	253	0,466433885	0,443600064	3
40875	5159	5346	-187	0,464991771	0,481813479	1
43327	5119	3589	1530	0,461386485	0,323462135	2
5438	4976	5148	-172	0,448497587	0,463968535	1
48360	4928	4774	154	0,444171244	0,430261419	1
5467	4909	5007	-98	0,442458733	0,451260772	1
9984	4853	5030	-177	0,437411332	0,45333367	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
42411	4826	4778	48	0,434977764	0,430621923	1
58508	4760	4798	-38	0,429029042	0,432424443	1
43560	4696	4724	-28	0,423260585	0,425755121	1
32916	4691	4754	-63	0,422809924	0,4284589	1
67353	4645	4726	-81	0,418663845	0,425935372	1
16269	4571	4478	93	0,411994066	0,40358413	1
10061	4523	4575	-52	0,407667722	0,41232635	1
16543	4519	4498	21	0,407307194	0,405386649	1
53415	4479	4530	-51	0,403701908	0,408270681	1
15372	4462	4479	-17	0,402169661	0,403674256	1
63657	4458	4703	-245	0,401809133	0,423862475	1
32618	4382	4600	-218	0,394959089	0,414579499	1
63262	4369	4459	-90	0,393787371	0,401871736	1
8311	4345	4430	-85	0,391624199	0,399258083	1
75065	4298	4306	-8	0,387387988	0,388082462	1
75083	4295	4365	-70	0,387117592	0,393399894	1
5716	4239	4274	-35	0,382070191	0,38519843	1
40743	4166	3828	338	0,375490544	0,345002244	3
16118	4138	3067	1071	0,372966844	0,276416375	2
69351	4109	4142	-33	0,370353012	0,373301801	1
47287	4094	4368	-274	0,369001029	0,393670272	1
64412	4054	4239	-185	0,365395743	0,382044021	1
Polygon_1	4051	4338	-287	0,365125347	0,390966493	1
27558	4010	4124	-114	0,361429929	0,371679534	1
60938	4006	4190	-184	0,3610694	0,377627848	1
5164	4002	4180	-178	0,360708871	0,376726588	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
8549	4002	4291	-289	0,360708871	0,386730572	1
58179	3982	3984	-2	0,358906228	0,359061897	1
16853	3977	4145	-168	0,358455568	0,373572179	1
69544	3946	4101	-155	0,355661471	0,369606636	1
55182	3908	4047	-139	0,352236449	0,364739833	1
59445	3905	3939	-34	0,351966053	0,355006228	1
46722	3893	4157	-264	0,350884467	0,374653691	1
37202	3874	3916	-42	0,349171956	0,35293333	1
5179	3847	3982	-135	0,346738388	0,358881645	1
9636	3815	3865	-50	0,343854159	0,348336905	1
32399	3757	4038	-281	0,338626494	0,3639287	1
6821	3736	3793	-57	0,336733719	0,341847835	1
61425	3714	3733	-19	0,334750812	0,336440276	1
58635	3683	3736	-53	0,331956715	0,336710654	1
43692	3608	3704	-96	0,325196804	0,333826623	1
75067	3608	3823	-215	0,325196804	0,344551614	1
1129	3601	3637	-36	0,324565878	0,327788182	1
15607	3595	3649	-54	0,324025086	0,328869694	1
52724	3575	3874	-299	0,322222443	0,349148039	1
36608	3571	3586	-15	0,321861914	0,323191757	1
59578	3571	3651	-80	0,321861914	0,329049946	1
4586	3564	3560	4	0,321230989	0,320848482	1
12585	3557	3594	-37	0,320600064	0,323912765	1
10348	3544	3670	-126	0,319428346	0,33076234	1
33237	3540	3549	-9	0,319067817	0,319857096	1
13749	3528	3680	-152	0,317986231	0,331663599	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
29561	3485	3532	-47	0,314110549	0,318324955	1
58309	3467	3043	424	0,31248817	0,274253351	3
5935	3426	3488	-62	0,308792752	0,314359412	1
11647	3408	3445	-37	0,307170373	0,310483995	1
47722	3399	3519	-120	0,306359184	0,317153317	1
5706	3391	3426	-35	0,305638127	0,308771601	1
75046	3287	3392	-105	0,296264383	0,305707318	1
12986	3283	3323	-40	0,295903854	0,299488625	1
75064	3246	3128	118	0,292568965	0,281914059	1
7031	3217	3349	-132	0,289955132	0,301831901	1
59517	3210	3250	-40	0,289324207	0,292909429	1
59709	3129	3159	-30	0,282023503	0,284707965	1
58566	3120	2998	122	0,281212313	0,270197682	1
75063	3114	3137	-23	0,280671521	0,282725193	1
62085	3106	3343	-237	0,279950463	0,301291145	1
24293	3036	3145	-109	0,273641213	0,283446201	1
Polygon_6	3005	3273	-268	0,270847116	0,294982326	1
49028	2974	2999	-25	0,268053019	0,270287808	1
6881	2939	2708	231	0,264898394	0,244061149	3
14972	2936	2749	187	0,264627998	0,247756314	3
2492	2934	3049	-115	0,264447733	0,274794107	1
69437	2929	3084	-155	0,263997073	0,277948516	1
5211	2919	2939	-20	0,263095751	0,26488025	1
13349	2903	2917	-14	0,261653637	0,262897478	1
2573	2903	3089	-186	0,261653637	0,278399146	1
75066	2891	2943	-52	0,260572051	0,265240754	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
73006	2878	3063	-185	0,259400333	0,276055871	1
10582	2876	2911	-35	0,259220068	0,262356722	1
13978	2867	3001	-134	0,258408879	0,27046806	1
40052	2863	2895	-32	0,25804835	0,260914707	1
5824	2859	2914	-55	0,257687822	0,2626271	1
61824	2816	2827	-11	0,253812139	0,25478614	1
42023	2767	3010	-243	0,249395664	0,271279194	1
67402	2764	2838	-74	0,249125267	0,255777526	1
7825	2764	2900	-136	0,249125267	0,261365336	1
56033	2753	2828	-75	0,248133814	0,254876266	1
70296	2722	2902	-180	0,245339717	0,261545588	1
39992	2707	2739	-32	0,243987735	0,246855054	1
3587	2678	2679	-1	0,241373902	0,241447495	1
5978	2658	2711	-53	0,239571259	0,244331527	1
50381	2643	2699	-56	0,238219277	0,243250015	1
49163	2624	2465	159	0,236506766	0,222160536	1
69382	2612	2809	-197	0,23542518	0,253163872	1
Polygon_2	2584	2676	-92	0,23290148	0,241177117	1
7103	2538	2643	-105	0,228755401	0,23820296	1
52820	2511	2501	10	0,226321833	0,225405071	1
31541	2506	2621	-115	0,225871172	0,236220189	1
43351	2475	2448	27	0,223077076	0,220628394	1
6571	2470	2531	-61	0,222626415	0,228108851	1
36922	2447	2345	102	0,220553375	0,211345419	1
42889	2440	2272	168	0,21992245	0,204766222	1
12656	2421	2488	-67	0,218209939	0,224233433	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
1319	2407	2642	-235	0,216948089	0,238112834	1
16392	2394	2401	-7	0,215776371	0,216392473	1
59033	2384	2401	-17	0,21487505	0,216392473	1
61315	2374	2489	-115	0,213973728	0,224323559	1
15066	2348	2364	-16	0,211630292	0,213057812	1
63413	2324	2423	-99	0,209467121	0,218375245	1
15829	2297	2069	228	0,207033553	0,186470649	3
7971	2285	2365	-80	0,205951967	0,213147938	1
29687	2247	2264	-17	0,202526945	0,204045214	1
58877	2237	2261	-24	0,201625623	0,203774836	1
67324	2204	2295	-91	0,198651262	0,20683912	1
56168	2173	2299	-126	0,195857166	0,207199624	1
29458	2166	2200	-34	0,195226241	0,198277152	1
10614	2154	1893	261	0,194144655	0,170608477	3
58699	2084	2137	-53	0,187835404	0,192599215	1
63365	2062	2105	-43	0,185852497	0,189715184	1
47659	2043	2123	-80	0,184139986	0,191337451	1
42901	2009	1985	24	0,181075493	0,178900067	1
38589	1976	2050	-74	0,178101132	0,184758255	1
12206	1942	2061	-119	0,175036639	0,185749641	1
12305	1921	2019	-98	0,173143864	0,18196435	1
11788	1889	1891	-2	0,170259635	0,170428225	1
59615	1878	1937	-59	0,169268181	0,17457402	1
14860	1847	1880	-33	0,166474084	0,169436839	1
41973	1836	1899	-63	0,165482631	0,171149232	1
75069	1835	2067	-232	0,165392498	0,186290397	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
45373	1793	1893	-100	0,161606948	0,170608477	1
34799	1761	1754	7	0,158722719	0,158080966	1
Polygon_8	1755	1999	-244	0,158181926	0,18016183	1
1032	1730	1830	-100	0,155928623	0,16493054	1
65037	1713	1924	-211	0,154396376	0,173402382	1
10583	1700	1744	-44	0,153224658	0,157179706	1
62095	1670	1795	-125	0,150520693	0,161776131	1
750	1630	1679	-49	0,146915407	0,151321517	1
44066	1612	1619	-7	0,145293029	0,145913959	1
38448	1584	1601	-17	0,142769328	0,144291691	1
48976	1527	1547	-20	0,137631796	0,139424888	1
7706	1525	1645	-120	0,137451531	0,148257234	1
72338	1516	1667	-151	0,136640342	0,150240005	1
27604	1514	1628	-114	0,136460078	0,146725092	1
2770	1499	1634	-135	0,135108095	0,147265848	1
10907	1447	1446	1	0,130421224	0,130322164	1
36362	1396	1459	-63	0,125824484	0,131493802	1
16164	1383	1389	-6	0,124652766	0,125184984	1
19383	1332	1441	-109	0,120056026	0,129871534	1
10853	1314	1324	-10	0,118433647	0,119326795	1
61283	1301	1372	-71	0,117261929	0,123652842	1
42746	1298	1301	-3	0,116991533	0,117253897	1
32238	1232	1253	-21	0,111042811	0,112927851	1
42925	1214	1183	31	0,109420432	0,106619032	1
58611	1161	1173	-12	0,104643428	0,105717772	1
8677	1140	1163	-23	0,102750653	0,104816513	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
40062	1101	1123	-22	0,099235499	0,101211473	1
1537	1065	1159	-94	0,095990742	0,104456009	1
45235	1064	1126	-62	0,095900609	0,101481851	1
13145	1061	1108	-47	0,095630213	0,099859584	1
59527	1043	948	95	0,094007834	0,085439427	1
27509	1043	1075	-32	0,094007834	0,096885426	1
40428	1020	1043	-23	0,091934795	0,094001395	1
29435	1008	1025	-17	0,090853209	0,092379128	1
37565	1000	942	58	0,090132152	0,084898671	1
75070	960	1068	-108	0,086526866	0,096254545	1
12409	944	913	31	0,085084751	0,082285018	1
67455	942	951	-9	0,084904487	0,085709805	1
42308	912	992	-80	0,082200522	0,08940497	1
14073	907	919	-12	0,081749862	0,082825774	1
16570	903	934	-31	0,081389333	0,084177664	1
12963	894	907	-13	0,080578144	0,081744262	1
15212	873	881	-8	0,078685368	0,079400987	1
31615	861	908	-47	0,077603783	0,081834388	1
57970	844	856	-12	0,076071536	0,077147837	1
42013	843	925	-82	0,075981404	0,08336653	1
46595	826	880	-54	0,074449157	0,079310861	1
13314	824	870	-46	0,074268893	0,078409601	1
50952 2	800	1338	-538	0,072105721	0,120588559	1
6622	785	810	-25	0,070753739	0,073002042	1
32212 2	753	1114	-361	0,06786951	0,10040034	1
6428	748	748	0	0,06741885	0,067414232	1

Table A.1: Figures of departures and arrivals, classification of the areas.

Area	Value Start	Value End	Diff. START - END	% OF DEPARTURES	% OF ARRIVALS	Cluster
71562	715	779	-64	0,064444489	0,070208137	1
16098	708	713	-5	0,063813563	0,064259822	1
35263	706	787	-81	0,063633299	0,070929145	1
56186	673	711	-38	0,060658938	0,06407957	1
40558	661	660	1	0,059577352	0,059483146	1
33854	611	734	-123	0,055070745	0,066152468	1
14905	355	354	1	0,031996914	0,031904596	1
17284	304	312	-8	0,027400174	0,028119305	1
71152	293	318	-25	0,02640872	0,028660061	1
11920	237	252	-15	0,02136132	0,022711746	1
11873	187	199	-12	0,016854712	0,01793507	1
57931	146	147	-1	0,013159294	0,013248519	1
58047	41	44	-3	0,003695418	0,003965543	1
13314 2	3	4	-1	0,000270396	0,000360504	1

Table A.2: Sensitivity Analysis number of neurons Neural Network with single layer.

# neurons	Duration	RSME Testing	RSME Fitting	RSME Validating
16	0:13	2,824384993	0,159571629	0,916231582
26	0:31	3,00433195	0,144471909	2,840559726
36	0:51	3,104441515	0,13437413	3,336853732
46	0:46	3,510557238	0,129008587	6,128985117
56	0:55	3,447865374	0,125169029	5,02270423
66	1:05	2,98210642	0,119599859	5,218493231
76	1:10	3,061952083	0,116754004	5,169953375
86	1:13	2,785969828	0,113279964	3,592643735
96	1:54	2,27964138	0,109970029	4,203677574
106	1:02	1,671978932	0,104672282	1,713790608
116	1:48	1,911989278	0,090721488	2,710296491
126	0:59	1,181313343	0,095628452	2,717862745
136	1:24	1,567625494	0,088664996	2,577917206
146	2:52	1,808734795	0,07530136	1,306853216
156	1:23	0,688033702	0,078503471	2,10729912
166	1:19	1,491608331	0,07313563	1,098059671
176	1:11	0,905641361	0,072746219	2,057878465
186	1:29	0,908144353	0,068679159	1,63403514
196	1:51	0,727703441	0,060183734	1,217934196
206	1:44	0,956015603	0,060196588	2,014903628
216	1:46	0,762051641	0,056355169	1,247643792
226	1:39	0,68173301	0,052503391	1,198946432
236	3:37	0,755621182	0,045060431	2,342115816
246	1:56	0,702972712	0,048112393	1,019663514
256	2:57	0,858618491	0,044107031	1,373082918
266	2:04	0,772023959	0,042704672	1,444428564
276	2:29	0,813449651	0,040564308	1,687804102
286	3:08	0,822039698	0,036037807	1,739581764
296	3:09	0,873469807	0,030567897	1,372284139
306	3:44	1,45015563	0,029785755	1,697409556
316	4:44	1,057344342	0,024892355	1,58597069
326	4:46	1,066111556	0,025132431	1,864051029
336	4:28	1,06795847	0,022856252	1,482953748
346	4:44	1,123000621	0,021598957	1,637240439
356	4:11	1,169883332	0,017649774	2,302565076
366	3:52	1,462726761	0,018651422	2,266786043

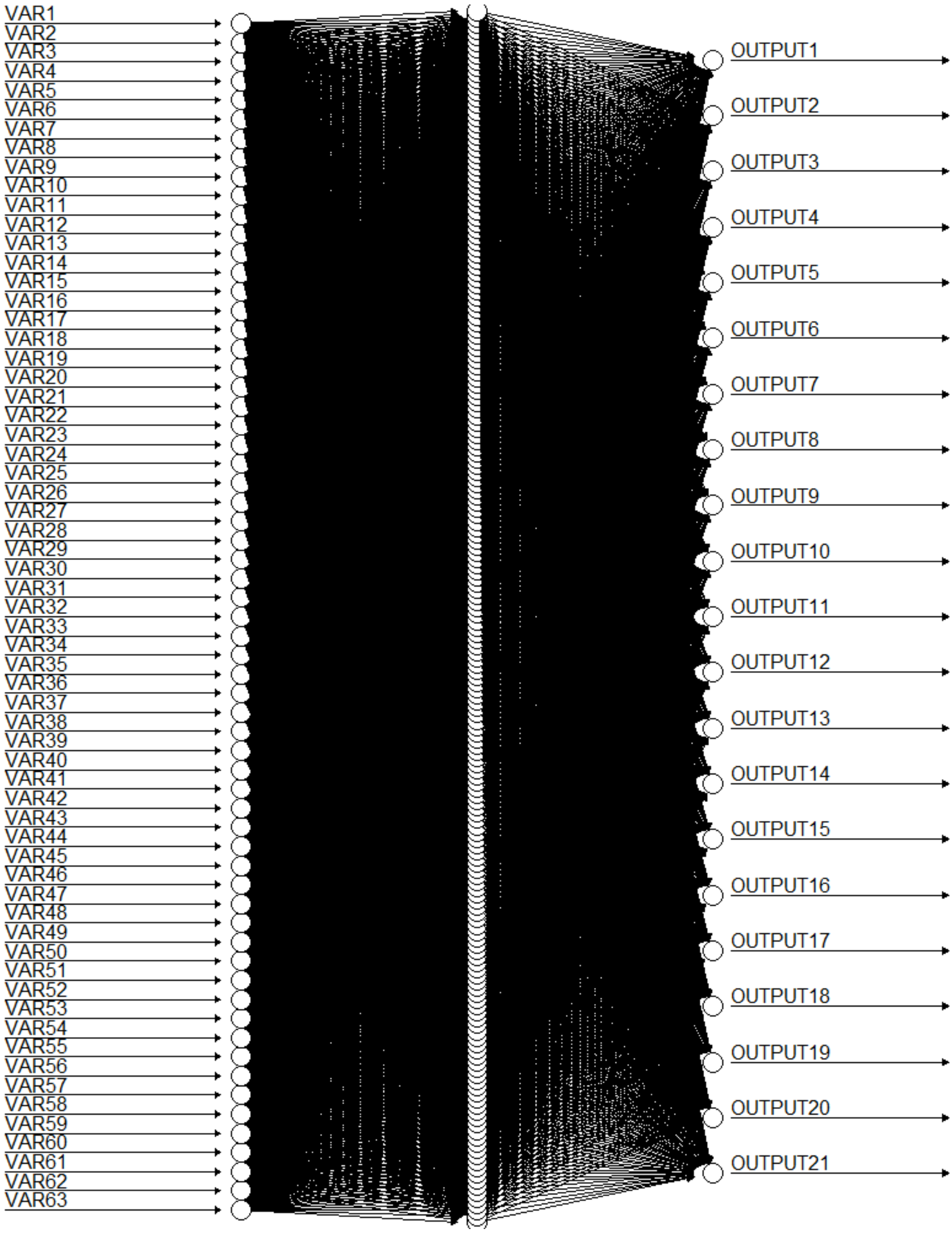


Figure A.4: Representation of the Model created.