CHAIR OF TRANSPORTATION SYSTEMS ENGINEERING

MASTER'S THESIS

Factors Affecting the Shift to Autonomous Vehicles: A Safety Perspective Survey in Munich

Kunwar Muhammad

Supervised by:

Dr. Christos Katrakazas

Univ.-Prof.Dr. Constantinos Antoniou

Date of Submission: 28th September 2018

Abstract

According to the advancements in the field of automation, the current human-driven vehicles will soon be replaced with the self-driven vehicles. The autonomous cars are predicted to be successful on roads with respect to safety as they will eradicate road accidents caused by the carelessness of the drivers. Before the introduction of highly autonomous/fully autonomous cars on the roads, it is the need of the time to evaluate the factors that will persuade the people to buy autonomous cars once they are available in the market. This research intends to find the factors with respect to safety of autonomous cars that will affect the shift of mode from conventional vehicles to autonomous vehicles for the city of Munich. For this purpose, an online survey questionnaire was designed and distributed among the residents of the city. Ordinal logistic regression models and Factor analysis models based on the important variables selected from the questionnaire were developed and later interpreted. The results indicate that age, gender and employment are the demographic factors that influence the decision of the respondents. The preference of the mode of automation and drunk driving are the other common factors from both models that impact the decisions of the people to shift towards autonomous cars. In addition, the safety of autonomous cars and the likelihood of elimination of accidents with their introduction were also calculated and interpreted with the help of the model output.

Acknowledgements

I would first like to thank my supervisor Dr. Christos Katrakazas for his tremendous support and invaluable advice throughout my thesis. His excellent supervision made the research easier for me and his knowledge helped me to go through all the difficult phases of the thesis with no trouble.

I am very thankful to Prof. Dr. Constantinos Antoniou, head of the Chair of Transportation Systems Engineering for his assistance and expert opinions during my thesis. His profound knowledge helped me in designing the survey questionnaire which was an integral part of the research.

I would want to express my gratitude to my work place team leader Kornelia Huber for translating the questionnaire in German for me which was a big help in achieving my aim.

I would also like to thank all the friends, colleagues and classmates for their constant support during the last six months and to the respondents of my survey questionnaire, without them this research study was not possible.

Finally, I would dedicate my thesis to my parents who have always been a continuous support during my life and without their love, care and prayers this was not possible.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AV	Autonomous Vehicles
AMOD	Autonomous Mobility on Demand
ADAS	Advance Driver Assistance System
EV	Electric Vehicle
GPS	Global Positioning System
PAV	Privately owned Autonomous Vehicle
SAV	Shared Autonomous Vehicle
SAE	Society of Automotive Engineers
FAD	Fully Automated Driving
V2V	Vehicle to Vehicle
V2I	Vehicle to Infra-structure
SLAM	Simultaneous Localization and Mapping
VMT	Vehicle Miles Travelled
CLD	Casual Loop Diagrams
TAM	Technology Acceptance Model
CFA	Confirmatory Factor Analysis
EFA	Exploratory Factor Analysis
EDA	Exploratory Data Analysis
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
PCA	Principal Component Analysis
Var	Variance
VKT	Vehicle Kilometers Travelled
St. Error	Standard Error
ITS	Intelligent Transportation System
LIDAR	Light Detection and Ranging
NHTSA	National Highway Traffic Safety Administration

Chapter 1: Introduction

1.1 Background and motivation

Nearly 1.3 million people die in road accidents each year and with an average of 3,287 deaths happening per day and around 20-50 million people getting injured or disabled due to road accidents per year around the globe *(ASIRT 2018)*. In the United States alone, 29 people die in car accidents every day implying that in USA a life is lost in road accidents every 50 minutes *(NHTSA 2017)*. Road traffic accidents are ranked as the 9th leading cause of death, and they also account for 2.2% of all deaths around the world *(ASIRT 2018)*. Over-speeding of vehicle is one of the major reasons behind accidents and is known to have caused 27% of motor vehicle crash deaths in 2016 in the United States. Furthermore, it has been a significant factor in more than a quarter of crash deaths since 2007 *(IIHS 2017)*. Including over-speeding of vehicles, most of the accidents occur from the decisions or rather negligence of the driver. Given the sheer amount of loss of life, resulting from negligence of the human in command of the vehicle, elimination of the command of the vehicle using autonomous driving is considered as the next major step towards safer road transport. Most of the major car manufacturers are embracing this idea and have started moving towards the development of autonomous vehicles (AVs) with the idea that less human involvement will overall reduce road accidents, deaths and other injuries caused by the negligence of the drivers. The increase in the safety of transportation by road is the most prominent factor that justifies the introduction of self-driven cars *(Reeve 2015)*.

The evolution of the automated safety systems is envisioned to decrease the number of road accidents-based casualties and injuries *(NHTSA 2017)*. In 2018, many new cars have built-in capability to assist drivers and warn them, such as lane-change warning reverse assistant for letting the driver know of cars and obstacles, automatic braking or slowing of vehicle to avoid collision with a car or some other object that is detected to be too close*(NHTSA 2017)*. These new technologies, which combine high-tech car sensors, such as cameras, radar systems, LiDAR systems and state of the art software helps the cars to recognize safety risks and either timely warn the driver or take automated actions to avoid car accidents *(NHTSA 2017).*

According to NHTSA, there have been five eras of safety pertaining to motor vehicles till now *(NHTSA 2017).*The first era began from 1950 to 2000, in which a few safety features were introduced in the vehicles for the convenience of drivers. It includes features like cruise control, seat belts and antilock brakes. The second era from 2000 to 2010 brought more advanced safety features in the vehicles like electronic stability control, blind spot detection, forward collision warning and lane departure warning. The third era between 2010 and 2016 gave birth to advanced driver assistance features like rearview video systems, automatic emergency braking, pedestrian automatic emergency braking, rear automatic emergency braking, rear cross traffic alert and lane centering assist. The fourth era starts from the year 2016 in which partially automated safety features are introduced in the cars which include lane keeping assist, adaptive cruise control, traffic jam assists and self-parking. Some of these features are already in function in a few high-end cars. The fifth era will continue from the year 2025 onwards, in which it is expected that cars will be driven without assistance of a human on highways with infrastructure to support fully automated driving *(NHTSA 2017).*

The previous paragraphs, described the positive aspects of the self-driven cars. On the other hand, other aspects such as technical challenges faced by the self-driving cards (safety, security, privacy, trust, transparency, reliability, responsibility, accountability, quality assurance process) are required to study as well. Out of all these challenges, safety remains the first and fundamental requirement *(Holstein 2018)*. There are questions regarding the testing of the safety of the self-driven cars as to how a self-driven car will be considered safe or otherwise. There are global standard testing procedures for current set of human-driven vehicles; however there are no standard and proven testing procedures for self-driven cars and this is expected to evolve in the years to come *(Holstein 2018)*.

Before testing self-driven car in real world scenarios, it will be a priority for car manufacturers and developers to know about the user demands and their perspectives regarding these vehicles. Moreover, it would be essential to identify which characteristics users would like to be found in these cars in order to shift to self-driven cars from standard nowadays vehicles.

1.2 Problem statement

Autonomous driving is the topic of utmost importance in the automotive industry these days. Automotive companies are working and testing on this technology with aim to have an edge in the future automotive market. However, with new developments made in this sector, there is a dire need to create public awareness regarding this new technology to investigate the factors that will influence users' future decisions on shifting from conventional vehicles to self-driven cars. According to *Fagnant and Kockelman,(2015) s*afety is one of the factors that will make the autonomous cars better than the conventional cars *(Fagnant and Kockelman 2015)* while according to *Noy et al ,(2018)* autonomous vehicles might not be as safer as they appear and a mere software failure would be enough to cause a big road accident *(Noy et al. 2018).*

The dominant questions are: i) How much safer would autonomous vehicles be as compared to the conventional human driven vehicles and ii) would automated cars be safe enough for public to switch to autonomous vehicles once they are widely available in the market?

The expectations of zero deaths with autonomous vehicles are unreal but self-driven cars will drive safer than a middle-aged driver*(Sivak 2015)*. Moreover, according to the same study there is a possibility that road safety might get worse during the transition phase, when there will be mixed traffic on the road with both conventional and self-driven cars. There is still a need of a strategy which identifies challenges and future work required for the safety approval of the autonomous. This is a challenging task considering the multi-disciplines that are associated with the performance of self-driven cars which include safety engineering, hardware reliability tests, software validation, robotics, security, vehicle testing, human-vehicle interaction, social acceptance and a legal framework *(Koopman and Wagner 2017).*

As mentioned by *(Van Brummelen et al. 2018)*, the three major areas which need further attention with respect to autonomous driving are:

- Reduction of uncertainty in perception
- Reduction in cost of perception systems
- Operating safety for sensors and algorithms

Additionally, for the safety approval process, there is a need of identifying a standard set of safety parameters that would be considered while validating a self-driven car for its safety. Currently, there is a lack of research in this area. There is no standard process of updating the existing practices used by conventional cars and adopting it for self-driven cars. Also, there has not been a concrete development around the creation of a design and validation process that validates safety issues of an autonomous car, which is acceptable to the general public and auto manufacturers with respect to cost, risk and ethics as described by *(Koopman and Wagner 2017)*.

Additionally, there are many other challenges that need to be addressed before self-driven cars become widely accepted in the market. Some of them as mentioned by *(Holstein 2018)* are:

- The questions around driving license for a self-driven car.
- The difficulty in developing global quality assurance or validation process for source codes that drives a self-driven car.
- High cost of state of the art hardware and software systems, like high tech laser systems, localization, mapping and computer vision based.
- The hard task for the car manufacturer to find an economical alternative hardware like using cameras instead of radar.
- The streamlined transfer of the control of the vehicle to the human in driving seat in case the software cannot handle the situation.
- Developing provision for human to override the autonomous software in case the software behaves differently than intended, and many other legal, ethical and technological challenges.

These challenges accumulate to make safety an important aspect of the self-driven car technology that needs to be addressed as one wrong decision might prove to be very detrimental and result in a loss of human lives and assets. One of the methods to understand expectations in the market about safety and other matters, is to research and learn from existing experience (experience of conventional car users)*(Holstein 2018)*.

As a result, a survey which correlates users' perspective with regards to the safety of autonomous vehicles, is likely to help car manufacturers better understand the challenges, work towards solving the problems and meet expectations of the users.

One example of such research is a study carried out in the UK where 1000 participants were surveyed to find out their perceptions regarding autonomous vehicles which were highly dependent on the road user type, gender and age *(Hulse et al. 2018)*. There was another research performed in the UK itself which involved 450 participants and enquired about the willingness of people to buy autonomous cars in future. This research mainly discussed the impact of primary factors like cost, safety and legalization while leaving behind the secondary factors like fuel efficiency, productivity of the user and other environmental impacts caused by the car *(V.Casely 2013)*. Similarly, in Canada, a survey performed on 162 Tesla drivers to understand their experiences and attitudes with the summon and auto-pilot systems and find out the level of trust the users developed with these semi-autonomous systems. This research, however, lacked a longitudinal survey approach as the opinion about the technology was taken after quite some time of the experiences and it might have changed in the meantime *(Dikmen and Burns 2017)*.

1.3 Aim and objectives

The field of autonomous driving is an emerging field and plenty of research work has already been done and is still in progress. The research made in the field as reviewed by the literature lacks a safety-based approach which points out the different parameters that will influence the mode shift.

This thesis focuses on the parameters to understand safety perspective of the autonomous driving sector. It aims at finding the important factors that affect the willingness of the people to shift from their conventional vehicles to autonomous cars, with considerations to the safety perspective for the city of Munich.

The aim of this thesis is to investigate user perspective with regards to the safety of autonomous vehicles in the city of Munich. In order to fulfill the aim of the thesis, the following objectives need to be fulfilled:

- Review of the literature with regards to autonomous vehicles, their safety and the user perspective with regards to safety
- The development of a survey that could reach out to the inhabitants of the city of Munich to collect their opinions about the safety of self-driven cars.
- Identification of factors influencing the shift to autonomous vehicles using statistical analysis.
- Evaluation of the results of the survey and the statistical analysis for drawing useful conclusion for research and businesses.

1.4 Research Overview

The research framework for this research has been developed to proceed in an organized manner. After making an introduction to the problem with this first chapter, an in-depth study of existing literature (Chapter No:2 Literature Review) is presented. After the review of the literature the methodology being developed to perform the research in the (Chapter No 3: Methodology) is explained. This is followed by the analysis part of the research, which demonstrates the method on how the research was analyzed with the specific methods used in the (Chapter No 4: Data Analysis). Finally, the last chapter (Chapter No 5: Conclusions) of this research includes the conclusions found out after the analysis of the data and gives suggestions about future research in this field.

Chapter 2: Literature Review

2.1 Introduction

This chapter gives an overview of autonomous vehicles including their challenges, opportunities and future implications for transportation policies. One of the features of the autonomous cars is their intelligent behaviour which is expected to improve road safety and efficiency. As a study suggests that autonomous cars could save up to 50 minutes each day that were previously spent in driving a car *(NHTSA 2017)*.This literature review discusses the different perspectives of self-driven cars with respect to their intelligence, responsibilities, drawbacks and the different challenges faced by them in the future. Furthermore, it discusses at length the safety of autonomous vehicles and safety of autonomous vehicles with regards to Users perspective. Lastly, it describes the existing gap in the literature concerning each feature discussed below in the review of the literature. The structure of the study is displayed in a reverse pyramid style on purpose in the figure 2.1 below which shows that the review starts from a general overview of autonomous driving which contains the potential benefits and hurdles associated with self-driven cars and then the literature is reviewed with respect to safety and specially safety with respect to User-perspective which is the aim of the study:

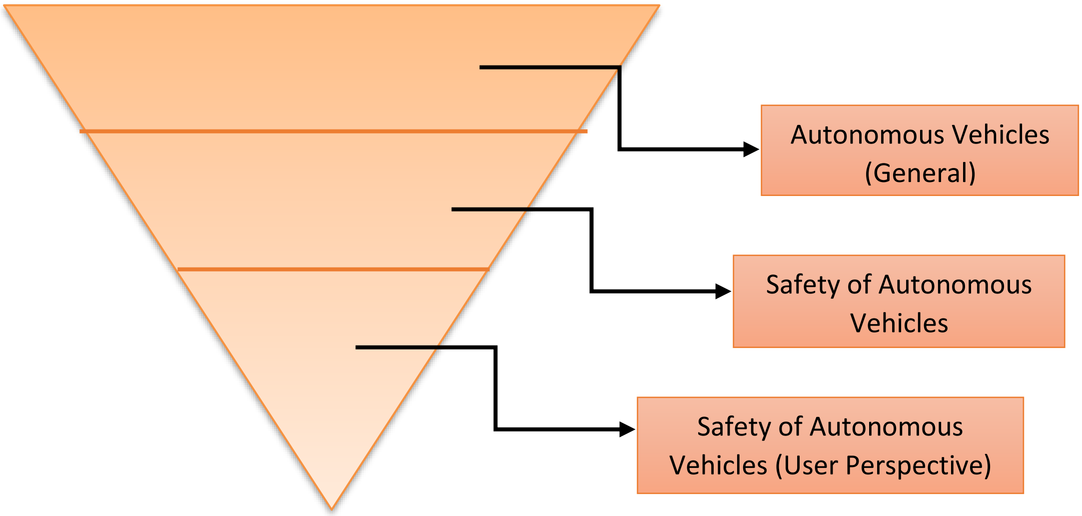

Figure 2.1: Structure of the study

2.2 Autonomous vehicles: Benefits and limitations

Autonomous cars are cars in which the operation of a car occurs without direct driver input to control the steering of the vehicle, giving acceleration and applying brakes plus they are designed in such a way that the driver is not expected to monitor the roadway regularly whilst he travels in self-driving mode*(NHTSA 2017)*.With the development of technology, autonomous cars could be a real game changer in the market. 54% off the total population worldwide currently lives in cities, and they are responsible for 64% of total kilometres travelled as they make 10 billion trips per day and with this rate in the future, the percentage of population living in the cities would jump to 66% by 2050 while the total kilometres travelled will also increase three times. The autonomous cars could increase the VKTs but could decrease the total kilometres travelled per user per day with a single vehicle used by different household members *(Thomopoulos and Givoni 2015).*

According to *Bagloee* and *Fagnant*, there are some potential benefits and future hurdles linked with the introduction of autonomous vehicles. The autonomous cars have certain characteristics which make them better than the conventional cars *(Bagloee et al. 2016).* These characteristics include; safety, overcoming traffic congestion, travel behaviour impacts, parking prices etc *(Fagnant and Kockelman 2015)* while AVs will improve the deficiencies in an EV with respect to charging infrastructure and travel range anxiety. There are almost more than a million road fatalities every year which makes safety an important characteristic while discussing autonomous cars because of less human intervention and the connected technology being able to monitor one's journey with the help of GPS and other navigation devices *(Thomopoulos and Givoni 2015).* There are certain other factors which make the functionality of autonomous vehicles better than the normal vehicles. These factors include safety management regarding crash rate, road congestion, taxi and car ownership *(Bagloee et* *al. 2016*). Autonomous cars are further likely to enhance private car mobility with the autonomous navigation helping them to move around the cities while they could also offer an opportunity of de-privatizing the car use. Such self-driven cars would allow drivers to driver everywhere around the world when they are not used to of the local driving customs like the differences between left-hand drive and a right-hand drive or driving with manual gears or automatic gears etc. Moreover, with the use of shared autonomous cars the car fleet within the city will decrease and ultimately having a small car fleet will decrease the demand for parking *(Thomopoulos and Givoni 2015)*. Autonomous cars will also lead to the increase in the accessibility and the people with disabilities which prevents them from driving a car or people having no driving licenses would be able to ride a car as a passenger or a driver *(Thomopoulos and Givoni 2015).* Furthermore, a study was carried out for the region of Stuttgart with a presence of a large autonomous mobility on demand (AMOD) service with a world without private cars. 45% of all vehicle movements and 20% of all vehicle kilometres travelled could be saved and 85% of all vehicles in the Stuttgart region might not be needed with the mobility on demand service *(Heilig et al. 2017).*

Talking about the hurdles that the autonomous vehicles will face, *Fagnant* makes a point about vehicle costs, autonomous vehicles licensing, litigation and liability, security and privacy concerning the introduction of the self-driven cars *(Fagnant and Kockelman 2015).* Cost is a key element which plays an important role for people to shift from normal vehicles to autonomous vehicles. There is a risk of autonomous cars and shared autonomous vehicles to be provided at a high cost which will affect the purchase of the AV and the households with lower income groups will rely on car travel. If AVs are not parked 96% of their time, then it might increase the congestion or energy use as they will be travelling more kilometres. A major threat mentioned by *(Nikolas,2015)* is the rebirth of sprawled development and its inter connected impacts due to the emergence of autonomous vehicles on dedicated lanes. It would highly affect non-motorized transport and aggravate social separation between those who are able to find and use AVs and those people who will not be able to use them *(Thomopoulos and Givoni 2015).*

However, the predictions related to the emergence of autonomous vehicles could change once they are on the road and the factors affecting their movement might not be the same as travel demand, vehicle miles travelled, and other factors largely depend on how people become acclimatized to the new road atmosphere and road infrastructure.

In short, every new technology has some pros and cons. The introduction of AVs could give birth to some hurdles while on the other side it could solve many issues.

- AVs could increase the VKTs while it could reduce the total kilometres travelled.
- The short comings of EVs with respect to charging stations and travel range could be improved with the introduction of AVs.
- The use of GPS and navigation with V2V and V2I technologies could make roads safer with AVs.
- AVs could decrease the demand for parking with the fleet moving around the cities all the time.
- AVs could increase the VKTs which might increase road congestion and energy use.
- Due to high costs, the households with less household income could rely on conventional vehicles.

2.3 Autonomous Vehicles and User Perspective

Autonomous vehicles are a new technology in the automotive market, and the people are not yet used to of it. The absence of knowledge for this new technology, leads to difficulties in understanding the safety benefits of AVs. This was displayed in a study-based research carried out for 721 participants living in Israel and North America, which aimed to observe their motivation to choose, own and use autonomous vehicles. 44% of the people voted in favour of regular vehicles while 32% voted for PAVs (Privately owned autonomous vehicle) and 24% for SAVs (Shared autonomous vehicles). Only 75% of all people voted in favour of SAVs even if the service is completely free. The research only focused on commuter-based survey and doesn't discuss the important variable of safety *(Haboucha et al. 2017)*. Another research was carried out by *Hill, (2017)* in UK to identify the public attitude with respect to the latest advancement in the field of autonomous driving. In this survey 2,175 people were surveyed using an online questionnaire, and the purpose was to know their interest in cars, self-driving cars, purchasing of a new vehicle and overall connected vehicle technology *(Lewis Hill 2017)*. The majority of the participants (94%) had heard of the driver assistance technologies systems that operate automatically such as cruise control and automatic parking. 76% of the people said that they know a little about these technologies while 60% of all the people were interested in having driver assistance technologies in their vehicle. The purchasing decision of buying a vehicle was still driven mostly by price, running costs and reliability instead of connected driving and driver assistance. A drawback of the research is that it excludes the offline population, those without internet and weighted the known population to overcome the non-response bias *(Lewis Hill 2017).*

2.4 Safety of Autonomous Vehicles

Safety has been an important area of concern in designing autonomous vehicle *(Reschka 2016).* Hence many previous studies have been conducted to check the feasibility of AVs about its safety for the users and non-motorized users (pedestrians and bicyclists). The interaction between vulnerable road users and autonomous cars has been of paramount importance as the autonomous vehicles become more and more prominent on the roads. For example, if a pedestrian with his family is trying to cross the road, the autonomous vehicle should have all the knowledge of this sort of social interactions and should act and communicate accordingly which will help the vulnerable road users to have the trust of the technology *(Saleh et al. 2017)*. In one of the online surveys conducted on 255 individuals by *Brar and Caulfield (2017)*, to find out the effect of autonomous vehicles on the safety of the pedestrians; it was found out that pedestrians were very careful about the introduction of autonomous vehicle and showed high level of concern. However, the research data gathered through the survey mainly focuses on students (227 out of the 255 individuals were students and aged between 16-24 years) which would give a very biased response in results *(Brar and Caulfield 2017).*

Along with vulnerable road users safety, there is a huge task for the different areas associated with the performance of self-driven cars like safety engineering, hardware reliability tests, software validation, robotics, security, testing, human-vehicle interaction, social acceptance and a legal framework to implement safety of the self-driven cars *(Koopman and Wagner 2017)*. *Koopman and Wagner, (2017)* also give the definition of safety which varies for autonomous vehicles as according to him safety means "*at least correctly implementing vehicle-level behaviours such as obeying traffic-laws and dealing with non-routine road hazards such as downed power lines and flooding*". There is also a big question mark on when a fleet of fully autonomous vehicles will be deployed on the roads which are safe enough for the humans to totally leave the driving seat. Along with that, there is a strong need of the creation of a design and validation process that fulfils all the safety issues and is also accepted by the public with respect to cost, risk and ethics *(Koopman and Wagner 2017)*.

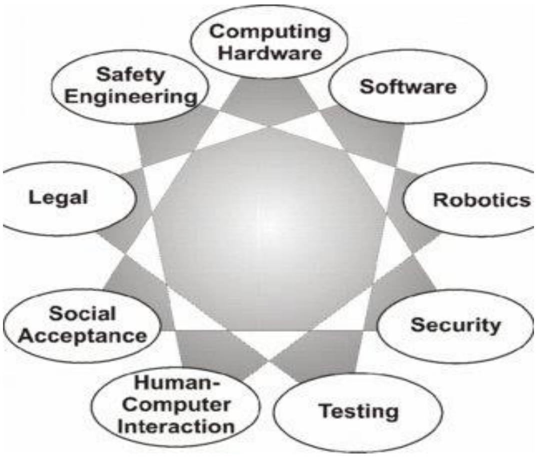

Figure 2.2: Different areas co-ordination, inter-disciplinary approach to ensure safety of AV *(Koopman and Wagner 2017)*

In designing autonomous vehicles, emphasis should be on the safety while the vehicle will be in motion with the different levels of automation. A research associated with the field of safety science talks about the safety features linked with the level of automation of an autonomous vehicle and further tries to put an argument where the conventional vehicles could be safer than the autonomous vehicles *(Noy et al. 2018*). *Noy et al, (2018)* indicate the lack of clear guidance from the SAE (Society of Automotive Engineers) regarding the establishment of design requirements of autonomous vehicles which could help designers in making design decisions. It also doesn't guide regulators plus the driving public to know about the autonomous system and the limits of the system. Most importantly, it doesn't inform automotive industry and consumers about the training needs. He later criticizes about the crashes being wholly eradicated from the road with the AV which he thinks is not true. The research further says that the AV might not be as safer as they appear, and a mere software failure would be enough to cause a big road accident. The paper also describes the fact that driver still would be involved despite of autonomous driving being a self-driven car apart from fully automated driving. It includes certain tasks for the driver if the car is not fully autonomous. The tasks that the driver might have to tackle become more challenging under critical considerations once AV will be on the road. The driver has to allocate tasks that are too complex or uncertain for the automatic reliability. Another task is deskilling, it is when drivers spend more time due to disengagement of automation while the system is under the control of the vehicle. Furthermore, cognition is another task, which is the monitoring of the highly complex system without pre-requisite diagnostic skills. Lastly, it includes the control of the vehicle which is actually driving the autonomous vehicle despite of the vehicle being autonomous. The institutions that could help to gain public trust in autonomous vehicles are manufacturers, government and academia *(Noy et al. 2018)*.

In another study; the author identifies the problems associated with the road safety with the self-driven cars. *Sivak, (2015)* indicates that the road safety might get worse in the transition period in which there will be mix traffic on the road with the conventional cars and the autonomous cars. Although, he predicts that the expectations of zero deaths with an AV are unreal and the self-driven cars will drive safer than an experienced middle-aged driver *(Sivak 2015)*.

One safety perspective of the autonomous vehicles could be described with the help of a U-shaped curve which shows that there is a high risk of young and elderly drivers as compared to middle aged drivers to be involved in fatal accidents the more they drive. The fatality rates with autonomous cars will be independent of the age which is shown by the flat line in the cases 1-4. When the technological and operational cases were realized, the actual safety rate (red dotted line) of AVs would increase as the fatality risk could be slightly reduced (cases 3-4), it could be significantly reduced (case 2) and finally could even eliminate (case 1) for all users irrespective of the age *(Thomopoulos and Givoni 2015).*

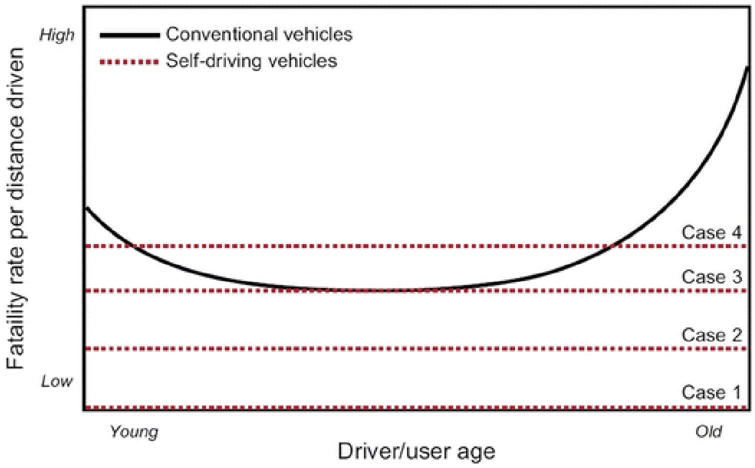

Figure 2.3: Fatality rates per distance driven using conventional vehicles and self-driven vehicles as a function of driver/user age *(Thomopoulos and Givoni 2015)*

Advanced driver assistance systems (ADAS) such as adaptive cruise control, forward collision warning, lane departure etc. have been introduced in most of the vehicles these years and are no more part of just the upper-class vehicles. The goal of these systems is improving driving safety and comfort by making the driver relax by performing routine tasks efficiently *(Hasenjäger and Wersing 2017)*. There are also studies conducted for the purpose of improvement of the road safety through warnings being generated to the driver in case of potential risk. In case he doesn't react then the system could itself perform autonomous actions. These are some of the tasks associated with the introduction of advance driver assistance systems (ADAS). The main aim behind the personalization of ADAS is to increase its usability and make it user friendly in context with the human-vehicle contact as it will play an essential role in improving the safety of the driver assistance systems *(Hasenjäger and Wersing 2017)*. There are previous researches conducted on the development and execution of an integrated advanced driver assistance system (ADAS) for the rural and intercity environment. There is one existing research carried out mainly for single carriage ways due to the intensity of accidents on the inter-urban roads being higher than that on the highways as the inter-urban roads have less structured environment, with more vulnerable road users and the infra-structure is not properly prepared in terms of safety. The proposed remedy to overcome the safety issues is the development of ADAS which is based on advanced perception techniques, vehicle automation and communications between vehicles (V2V) and vehicle to infrastructure (V2I) technology. ADAS performs real time detection of obstacles and classifies them and more importantly it identifies the potential risks *(Jiménez et al. 2016)*. The research however, gives no clue about how these systems are going to work with highways or motorways (dual-carriage ways) and how is it will deal with fully autonomous cars.

The safety of autonomous vehicles is an on-going research and plenty of studies have been performed to identify the potential hazards with respect to safety but none of them have currently indicated the factors that affect the safety of autonomous vehicles and remedies to overcome them.

2.5 Comfort

According to *Bellem et al, (2018)*, "*comfort is associated with a feeling of well-being, an attribution of positive valence towards the eliciting entity, the absence of discomfort and uneasiness".* A close relationship is seen between comfort and trust with respect to highly automated driving and the acceptance of automated vehicles. Therefore, comfort can be one of the factors deeply associated with technology adoption and can be considered as a barrier. As far as the manual driving is concerned, the passenger comfort largely depends upon the driver's driving style, vehicle safety equipment, safety systems and seats while as compared to the autonomous driving the only option that could affect the comfort is by manipulating driving style because the other factors will not change with a higher automation level *(Bellem et al. 2018)*. A simulator study having a sample size of 72 was performed to identify driving strategies based upon passenger comfort. Three variations were configured in the simulation. They include lane changes, accelerations and decelerations to find out the comfort level with respect to each manoeuvre. The research found out that other than self-reported driving style, no other personality trait affects the preferences for automated driving styles. Therefore, automated driving style was recommended because it was perceived comfortable by the participants of the study and not dependent on their personality *(Bellem et al. 2018).* The drawback of the simulation-based studies is always the change in behaviour of the user with the change in environment and could be different when the tests are performed in real conditions.

2.6 Determinants of shifting to self-driven vehicles

To determine the likelihood of people adopting the autonomous vehicles as their mode of transportation in the future, a two-stage data collection survey method was employed by *Zmud et al., (2016).* The impact on the amount of travel, mode choice, car ownership and other similar behaviours were also put into consideration by the researcher. The first stage of the method consisted of an online survey which was created and conducted with 556 residents of metropolitan Austin to know about their intent to use the AV in future and after the survey, qualitative interviews were conducted with the people (205) who showed the intent to use to determine the impact on their travel behaviour. The research concluded that most of the respondents (59%) would rather want to own self-driving cars because of the convenience than just use one, like Car2Go or Uber taxi (41%). The determinants found by the research might change in future with the actual interaction of the users and the autonomous vehicles *(Zmud et al. 2016).*

To analyse the priori acceptability, intentional traits, personal traits and attitudes using a fully automated vehicle, a study was carried out using an online questionnaire. In the study, 421 French drivers were surveyed. 68.1% of the sample priori accepted fully automated driving (FAD). The main predictors to influence the acceptability were attitudes, contextual acceptability and interest in impaired driving, driving related sensation and gender *(Payre et al. 2014).* One of the gaps of the study is that the researcher chose only male respondents for semi-conductive interviews while the people to which the questionnaire was mailed were chosen by the researchers themselves which could be the ones belonging to the same technical institute and more inclined towards FAD and therefore could lead to more biased results.

2.7 Perception about autonomous cars

The introduction and realization of any new technology leads to the perception of the people about it. Perception of people about any new technology is very important for the manufacturers and the new users as public attitudes. The fully automated self-driving car with the expectations of improving road safety are close to becoming a reality and therefore, all the attention is towards the appraisal of public perception about these automated vehicles *(Hulse et al. 2018).* Public attitude towards self-driven cars is of massive importance as the public opinion matters in the future demand for the technology, the governing policies and the future investment in infrastructure *(Howard 2013).*

The perceptions of autonomous vehicles can be observed in *Hulse et al.(2018)*. In this study the authors, surveyed 925 participants using an online survey in the United Kingdom. The results threw light particularly on the perceived risk of collision and injury for the passengers and pedestrians. Moreover, according to the survey results gender and age were the two demographic factors that were of most importance with respect to safety indicators analysis of perceiving autonomous cars with male participants and younger participants were more likely to be travelling in an autonomous car than females or other age groups *(Hulse et al. 2018)*. However, the research lacks a multi-perspective approach towards the relationship between road users and autonomous driving.

In another study by *Van Brummelen et al, (2018)*, perceptions about the different sensor technologies which will be used by autonomous vehicles was demonstrated. They point out three areas which need further attention. These are reduction of uncertainty in perception, reduction in cost of perception systems and operating safety for sensors and algorithms (Van Brummelen *et al.* 2018). Further, they explain the different localization and mapping techniques to be used by the autonomous cars and how those algorithms will work. They also explain the stages of navigation methods used by the AV which include environmental perception, localization and mapping, path planning, decision-making, and vehicle control *(Van Brummelen et al. 2018)*. The study shows that the current perception systems in automation level 1 to 3 will increase vehicle safety but there is still room for improvement before the emergence of fully automated vehicles. Efficiency of computational methods and algorithms in perception is not discussed by the paper which is an important part of perception systems and the paper lacks deep rooting of knowledge to be more helpful for the people working in the field *(Van Brummelen et al. 2018).*

To explore more about public perception regarding autonomous vehicles, some of the existing researches are discussed below.

A case study in 2015 was done by *Kyriakidis et al. (2014)* to study the public opinion about the autonomous cars and how much they are willing to buy them. He used a sample size of 5000 respondents in his case study. The research took place in 109 countries with at least 25 respondents from 40 countries. It considered all levels of autonomous driving and not any specific level which many previous studies have done. The survey was constructed using a 63-question online survey. 33% people found autonomous driving as enjoyable while the majority voted in favour of manual driving. The results show that the respondents were most concerned about software hacking, misuse and data privacy issues with the autonomous cars *(Kyriakidis et al. 2015).*

Another study identified the perception of the people and their attitudes regarding the self-driven cars by surveying 107 people in California. They were asked about the best and least attractive thing about self-driving cars and how do they envision this technology and furthermore their individual demographics, existing travel behaviour and relationship to cars and technology *(Howard 2013)*. Out of the 107 responses, 46% respondents believe that self-driven vehicles should be run with normal traffic while 38% people said that there should be separate dedicated lanes and 11% had no opinion.

Additionally, one online survey on around 1000 participants, stratified by age, income and gender was carried out to find the potential impact of autonomous driving by *Fraedrich et al., (2016).* The study depicts that the acceptance of autonomous vehicles is most strongly impacted by the participants' present level of interest and knowledge of self-driven cars as well as their eagerness to abdicate certain driving functions. The socio-demographics features which were thought to have a strong impact on attitudes and behaviour did not impact the results of the study. The respondents currently expect that there won't be much impact on their personal mobility behaviour (Fraedrich *et al.* 2016)

An additional research, with 450 participants was carried out to identify the willingness of people to use self-driven cars in future by *Casely, (2013)*. The research was categorized into primary and secondary factors influencing the acceptance of self-driven cars. The primary factors were safety, cost and legal structure while the secondary one involved productivity, efficiency and environmental impacts *(V.Casely 2013)*. Although the research was carried out mainly for the primary factors, the results showed that the secondary factors were equally playing an important role. The research suggested that the technology and the laws must improve and develop further for the public to show willingness to shift to autonomous cars *(V.Casely 2013)*.

In an effort to find out the factors affecting the user acceptance of self-driven cars and the future implications for the government and automotive manufacturers; 35 interviews were conducted by *Kaan, (2017)*. The interviews had open ended questions and took an audience between the ages of 18-40 years based upon the research results and improved knowledge in this sector *(Kaan 2017)*. The research showed that many people didn't know much about the technology and it was difficult for them to form an opinion about the technology. The research further explains that there is lack of awareness among the people regarding the benefits of self-driven cars which is an important area for the government and the automotive industry to focus on *(Kaan 2017)*.

The trust of users on driverless cars is one of the factors which is explored by *Kaur and Rampersad, (2018)*. They identify the factors involved in developing the trust in driverless cars with an online survey engaging 101 members from the faculty and student section of the university in which the participants were asked 10 questions each having various sub-questions about the autonomous car adoption. The research identified key factors influencing autonomous vehicles adoption like performance expectancy, reliability, security, privacy and trust. The research also found out the situations in which people are more likely to adopt the autonomous cars which include closed environments, finding a car park, public transport with a companion and on highways where drivers can take full control of the vehicle. The study area was confined due to the university own research as it plans to introduce an autonomous car that shall pick up staff and students from the train and bus stations to the campus (Kaur and Rampersad 2018).

Finally, an online survey conducted with 165 Tesla drivers having ages from 16-65 years was carried out by *Dikmen, (2017)* in which they were asked about their experiences with two advanced driver assistance systems Auto-pilot and Summon. The study finds out that the trust in these systems was very much co-related with attitude and behaviour factors like frequency of use, self-rated knowledge about the systems and ease of learning *(Dikmen and Burns 2017)*.

2.8 Summary of literature review of general aspects of autonomous vehicles

This section 2.8 gives an overview of all the literature studied with respect to the general characteristics of autonomous vehicles. It includes the research that has been performed without any experiments on users and just discusses the future prospects of autonomous vehicles.

Table 2.1 summarizes the review of literature with regards to autonomous vehicles in general.

Table 2.1: General research with respect to autonomous vehicles

Author: (Bagloee *et al.* 2016) **Field of Study**: Civil, Infrastructure, Transportation Engineering and Business Studies **Research area**: Future Opportunities and hurdles associated with Autonomous Vehicles and Vehicle Navigation Model for Autonomous Driving **Overview**: It describes the function of autonomous vehicles on a "Sense-Plan-Act" technique. It explains the reduction of existing negative externalities with their introduction. The impact of AV on road congestion in a positive way. A fleet of Autonomous Vehicles can resolve the practical implications of EV including travel range anxiety, access to charging infrastructure and charging time management. AV could remove the parking spaces in the city-centre; stimulate sub-urban growth and more buildings. **Limitation**: The research foresees the challenges and opportunities regarding AV as previously studied in the literature and don’t do any test in real life to support all the claims. The logics are there but there is lack of practical implications to support the logics.
Author: (Noy *et al.* 2018) **Field of Study**: Safety Science **Research area**: Challenges which will be faced by the autonomous vehicles with respect to safety **Overview**: The research argues about the accidents termed as human-error might be caused by other factors on the road and with error-free technology there won’t be any crashes. The paper also focuses on autonomous driving and socio-technical systems as to how information is travelled throughout the system. The critical role of institutions like manufacturers, government and academia has the responsibility to influence public trust in AV **Limitation**: The research doesn’t give any solid mitigation on how to overcome these blind spots based on stats or history if they have already been applied in some other field facing the same problems. It is too early to say and with just two examples (Google car, Tesla Car) that AV without the involvement of driver will lead to unnecessary losses and injuries.
Author: (Van Brummelen *et al.* 2018) **Field of Study**: Transportation Research (Emerging Technologies)

Research area: An overview of the current autonomous vehicles perception system and identification of the areas of AV perception that require further research

Overview: The study focuses on three main areas which need improvement regarding AVs. They include reduction of uncertainty in perception, reduction in cost of perception systems and operating safety for sensors and algorithms. The paper also explains the localization and mapping techniques used by AV from the initial algorithms to SLAM (Simultaneous Localization and Mapping) used by the Tesla Model S.

Limitation: The research left large gap in literature as it didn't tell about the efficiency of computational methods and algorithms in Perception. The study is very knowledgeable for beginners but needs deep rooting into the algorithms and perception sensors to make it more advanced as it points out the areas which needs further research in this field but lacks practical implications.

Author: (Fagnant and Kockelman 2015)

Field of Study: Civil, Architecture and Environmental Engineering

Research area: An introduction to the transportation professionals and policy makers regarding the AV technology, hurdles, Potential benefits and their impacts.

Overview: The research paper demonstrates the expected benefits of the autonomous vehicles when they will come on roads concerning Safety which makes them safer to travel. With AVs, there will be less expected crashes, Congestion. Traffic operations will also get better through V2I (Vehicle to Infrastructure) and V2V (Vehicle to vehicle) communication. Travel Behaviour Impacts would predict that the number of trips will increase (VMT more travelled) as AV will provide mobility to everyone. Freight transport will get better with increasing productivity; fewer travel times and lower costs through reduced labour.

Limitation: The research is a thoroughly predictive research which could greatly influence the outcomes as predicted when autonomous vehicles are on the road. There is uncertainty in the research which lacks the research as in how the travel demand patterns will change, management of intersections & ITS infrastructure, change in VMT and vehicle emissions and future market penetration rates after the introduction of autonomous vehicles.

Author: (Sivak 2015)

Field of Study: Traffic Safety

Research area: Issues related to self-driving cars and road safety have been discussed

Overview: The research focuses on the safety related issues with the introduction of autonomous cars. It finds out that the expectation of zero fatalities with autonomous driving is not realistic. A self- driving vehicle would perform safer than a middle-aged driver. During the transition period from conventional vehicles to self-driven cars, the safety might get worse.

Limitation: The research doesn't perform any survey or interview to back up its conclusions. It uses the stats from other researches instead which gives just a vague idea about how it is all going to change with the introduction of self-driven vehicles and how this transition phase needs to be tackled when there would be conventional vehicles and self-driven vehicles both on the road.

Author: (Ross and Guhathakurta 2017)

Field of Study: Quality Growth, Regional Development and GIS

Research area: Identify and quantify the impacts of autonomous vehicles on energy through development of scenarios

Overview: The research paper focuses on the energy sector with respect to the levels of autonomous vehicles development. Scenarios were developed between partial automation, full automation vs shared autonomous vehicles and personal autonomous vehicles. The impacts on the energy consumption in 2017 with case scenarios were forecasted. Full automation is likely to consume more energy because it allows vehicle to travel faster and induce travel demand.

Limitation: The research failed to provide a detail overview of the scenario when there is partial automation with shared vehicle domination and what would be the impact on the energy consumption. As the energy consumption on the highway for trucks is more (25% in total) so they should be separately analysed.

Author: (Gruel and M. Stanford 2016)

Field of Study: Transport Planning

Research area: Assessing the long-term effects of autonomous vehicles on travel behaviour and overall transportation system

Overview: The research paper builds a transportation model to describe the effects of the vehicle automation at the system level using case scenarios. The scenarios taken into consideration are that the technology change, but the behaviour doesn't, the new technology induces new behaviour and lastly, the new technology might drive new ownership models with the change in ownership from private vehicles to shared vehicles. The base model used for all the scenario modifications is the CLD (Causal Loop Diagrams) of traffic and congestion introduced by John Sterman.

Limitation: The research didn't discuss the potential of shared AVs to be used as first and last mile connectivity with trunk lines which could help in improving public transport. Further, there is still room for discussion about ride sharing in shared AV's that could save several trips and decrease congestion.

Author: (Jiménez *et al.* 2016)

Field of Study: Traffic Safety

Research area: Developing an integrated driving assistance system for inter-urban environments for safety improvement

Overview: The research paper demonstrates a collision avoidance system which includes obstacle detection system with the help of artificial vision and 3D laser sensor technology PC (Point Cloud) and wireless communications modules for single carriage way to improve safety on the roads of

Spain. The advanced driver assistance system developed provides warnings to the driver if it detects any danger and can even take control of the vehicle to perform evasive maneuvers. **Limitation:** The research is only conducted to be used for single carriage ways and not for dual carriage way motor ways. The tests performed by the research are for partial autonomous driving and there is still space for fully autonomous driving.
Author: (Seif and Hu 2016) **Field of Study**: Big Data and Engineering Perspective **Research area**: In-depth Analysis of the necessary technologies for automated driving in future cities **Overview**: The research ponders over the technology that is going to be used by the newly introduced autonomous cars in a few years. The need for the HD maps for fully autonomous cars with an accuracy of +-10cm which provides exact location, speed and direction of the car, the current traffic situation and the movement of other road users. **Limitation:** The paper just reveals a background of the technology with future update and nothing substantial has been discussed. The technology details are missing.
Author: (Heilig *et al.* 2017) **Field of Study**: Transport Modelling **Research area**: Study of the mode choice behaviour after introduction of AV in Stuttgart region using multi-agent simulation travel demand model called mobiTopp. **Overview**: The research shows that 85% of the total vehicles in the Stuttgart region might be replaced by the introduction of 100% self-driving vehicles specifically AMOD (Autonomous Mobility on Demand) service in which 4 persons share a ride and there is no private vehicle on the road. The research further calculates that 45% of all vehicle movements and 20% of all vehicle trips could be saved in that manner. **Limitation:** The study lacks the information about vehicle usage regarding spatial effects which needs to be further analysed. The assumption of having 100% AV (AMOD) and no private vehicles on the road is a bit unrealistic for the future scenario.
Author: (Wei *et al.* 2013) **Field of Study**: Intelligent vehicles **Research area**: Development of a highly integrated autonomous vehicle **Overview**: The research introduces a platform to develop an autonomous vehicle and the different characteristics and functions of autonomous driving. The different hardware and software of an autonomous vehicle and how they function.

Limitation: The different road tests in different conditions might tell more about the performance statistics of the autonomous vehicle and vehicle's intelligence should be made better for the vehicle to know surrounding vehicles movements and perform actions accordingly.
Author: (Koopman and Wagner 2017) **Field of Study**: Traffic Safety **Research area**: Challenges for the autonomous vehicles fleets **Overview**: The paper signifies the importance of safety verification of autonomous vehicles with respect to different disciplines like safety engineering, hardware reliability, software validation, robotics, security testing, human-computer interaction, social acceptance and a viable legal framework. These all areas add up to determine the successful movement of fully autonomous vehicles in future **Limitation:** The research points out the main challenges regarding cross disciplinary challenges but fails to establish the link between these all disciplines and there is still a need to update the practices that have been already accepted and create a design and validation process that explains all safety concerns.
Author: (Hasenjäger and Wersing 2017) **Field of Study**: Intelligent Transportation Systems **Research area**: Reviewing of personalization approaches for ADAS systems **Overview**: The research paper stresses upon the implicit personalization of the Advance Driver Assistance Systems (ADAS) from the human-machine interaction point of view and to make the technology more acceptable and useful for the people. **Limitation:** The research tells the factors and functions on which the personalization is based upon but doesn't perform any real-time study to show how this implicit personalization affects the human-vehicle relationship.
Author: (Urooj *et al.* 2018) **Field of Study**: Computer Science and IT Department **Research area**: Systematic Literature Review on User Interfaces of Autonomous cars **Overview**: The research paper shows the in-depth literature review which involves the tasks like development of review protocol, identification and selection of primary studies, data extraction and reporting of the results. **Limitation:** This systematic review is least important for someone formulating the rules and regulations regarding autonomous cars as the strength of evidence of the answers of the research questions are very low.
Author: (Thomopoulos and Givoni 2015)

Field of Study: Transport Mobility

Research area: Exploration of the likely and the desirable outcomes of the autonomous vehicles

Overview: The research paper extensively explains the likely benefits and threats with respect to the different impact groups with the dawn of autonomous vehicles on the road and their effects on the low carbon mobility in future.

Limitation: The paper purely depends on the existing literature and doesn't do any further study/survey to find out the benefits and threats which will be caused by the AV when they are on the road.

Author: (Saleh *et al.* 2017)

Field of Study: Transport Modelling

Research area: Understanding between Vulnerable Road Users and Automated Vehicles

Overview: The research paper proposes a computational framework that models trust between vulnerable road users and autonomous vehicles based on a shared intent understanding.

Limitation: A simulation case study might have explained this modelling process in a better way than without it.

2.9 Summary of safety of autonomous vehicles and user perspective

In this section 2.9, the studies that have been performed with respect to the safety of autonomous vehicles has been discussed and it includes the literature in which research experiments and surveys are conducted with regards to the safety of self-driven cars.

Table 2.2 summarizes the literature reviewed with respect to the safety of autonomous vehicles and the user perspective.

Table 2.2: Safety of autonomous vehicles and User Perspective

Author: (Haboucha *et al.* 2017)

Field of Study: Transportation and Geo-Information Engineering

Study area: Israel & North America

Overview: 721 participants asked to state their preference for future mode of transport for daily commutes based on certain attitudinal variables (PRO-AV, Environmental Concern, Technology Interest, Public Transit Attitude, Enjoy Driving) getting 44% in favour of regular vehicles, 32% for PAVs and 24% for SAVs.

Limitation: Only focused on the commuter trips. Safety being one of the most important parameters has not been discussed at all.

Method: State Preference Survey was created to gather individual preferences data of 721 participants. The characteristics are quantified through random utility models like logit kernel model and nested model. Afterwards SPSS and AMOS are used to reduce attitudinal variables through factor analysis.
Author: (Bellem *et al.* 2018) **Field of Study**: Psychology **Study area**: Germany **Overview**: 72 participants were tested with a simulator-based study to observe the maneuvers (different driving styles) with respect to the level of comfort in an automated vehicle. The maneuvers selected were Acceleration, Lane change and deceleration. A pre-drive and post -drive questionnaire was also filled by the participants. **Limitation:** The results gathered in a simulator could be different when performed in real case scenario. The perception of drivers regarding trust will be different if this test was performed on real road ways. **Method:** Simulator based study performed on 72 participants who were paid 33$ each to perform the task. Three types of maneuvers with three different variations were observed for each participant. The participants were asked to fill a pre- and post-simulation questionnaire to gather the opinions. The data analysis took place using the BTL model with the EBA package in R to get the final results.
Author: (Brar and Caulfield 2017) **Field of Study**: Civil, Structural, Environmental Engineering **Study area**: Ireland **Overview**: 255 individuals filled a survey questionnaire about how safe they would feel when autonomous vehicles would be on the road. 87.8% of total population had heard of AV before the survey. The research focuses mainly on safety with respect to vulnerable road users and their interaction with Autonomous Vehicles. The topics which the questionnaire covers are Demographics, Opinions about AV, Expected Benefits and Expected Concerns about AV. **Limitation:** The survey focuses mainly on the student sector with having 227 off the 255 individuals in between the ages of 16-24 which might give biased results when its descriptive analysis is performed in terms of a new technology. The popular mode of transport that most of them used is Public Transport (116/255) which again might give biased results with respect to the sample size. **Method:** The research uses an online platform to create a questionnaire which was forwarded mostly to the Engineering students of the college. 255 respondents filled the questionnaire having 16 questions to answer which were broken into different topics. The views were later analysed based on a descriptive analysis and rating scale was chosen from 1 to 3 to calculate the Mean and Standard Deviation of the sample.

Author: (Lewis Hill 2017) **Field of Study**: Mobility, Safety, Economy and Environment **Study area**: United Kingdom **Overview**: 2,175 respondents aged between 18 to 75 years filled the online questionnaire gathering information about their attitude towards technology (76% agreed that it makes life better). Then it asked their attitude towards cars (59% agreed that their life style depends on cars). Attitude towards purchasing a new vehicle and their interest in connected vehicle technology. **Limitation:** The research excludes the offline population, those without internet and weighted the known population to counteract non-response bias. The participants are only samples of the total population in the research and the results might vary if each person in UK aged 16-75 years was surveyed. **Method:** The methodology of research used is making a survey questionnaire online and getting it filled by 2,175 participants asking 17 questions about technology, cars usage, popularity, prevalence and policy of connected driving technology. Social grade definitions were used to distinguish between the manual and non-manual occupations. Sampling tolerances were added to the percentages at or near the levels.
Author: (Kaan 2017) **Field of Study**: Transport and Logistics **Study area**: Netherlands **Overview**: 35 interviews with open ended questions were conducted with an audience between the ages 18-40 years to determine which factors play an important role in the user acceptance of autonomous vehicles. The conclusions from the interviews helped in finding the second part of research which are the implications of improved knowledge for national government and automotive manufacturers. **Limitation:** The gap in the literature is that different audience might yield different results while the open-ended questions might not enlighten some factors which would otherwise be there. A sample size of 35 is relatively small for a hot topic like autonomous vehicles and might not be credible enough for the results **Method:** The research was conducted using Interviews with Open ended questions in two rounds with a sample size of around 35 and then the factors mentioned in the interviews were used to code the interviews using MS Excel to find out the implications of the knowledge for the government and automotive manufacturers.
Author: (Zmud *et al.* 2016) **Field of Study:** Transport Planning **Study area**: United States

Overview: The research uses two stages of data collection. 556 people were asked using an online survey about their intent to use the autonomous vehicles in future. Interviews were held with the 205 people that showed intent of use to know the impact on their travel behaviour and what factors would influence their travel. **Limitation:** The determinants might change with the passage of time as soon as autonomous vehicles become available. A larger population (>556 people) with less people having full or part time jobs (<64%) might produce different results because household income and age are an important factor in determining the future intent of use of AV. **Method:** The research uses an online questionnaire and gets it filled by 556 respondents of the Austin area. There were instructions and some video links before every question section and after the questionnaire, interviews were conducted of the people by offering them monetary value to know how their intent of using the AV will impact their travel behaviour in future. The research uses a Technology Acceptance Model (TAM) to quantify the qualitative factors.
Author: (Hulse *et al.* 2018) **Field of Study:** Safety Engineering **Study area**: United Kingdom **Overview**: 1000 participants were surveyed to find out their perceptions regarding Autonomous Vehicles. The study found out that the autonomous vehicles are somewhat less risky mode of transport, but it depends on the road user type, gender and age. **Limitation:** The research lacks a multi-perspective approach to find out the interaction between autonomous vehicles and users. Other demographic factors than age, gender and road-user type could be selected. The conclusion is a bit short and doesn't tell the mitigations regarding the result. **Method:** A seven-point ordinal scale was used with a six-item instrument for the participants to rate their opinions regarding risk behaviour of road users. Seven points starting from extremely unlikely to extremely likely and the six-item instrument included different conditions on the road with respect to risk. The survey results were then analysed using SPSS software and CSGLM command which performs linear regression as well as analysis of variance and co-variance.
Author: (Payre *et al.* 2014) **Field of Study:** Transport Planning and Mobility **Study area**: France **Overview**: 421 French drivers (age range 19-73 years) filled an online questionnaire. The research performs an analysis using a priori acceptability method to find out the attitudes, personality traits and intentions to use a Fully Automated Vehicle. 68.1% of the sample priori accepted Fully Automated Driving (FAD). **Limitation:** The sample chosen was biased as only males were chosen for the semi-conductive interviews. The people couldn't be seen while filling the questionnaires which could give false results. Questionnaire was mailed to a specific group of people which could be inclined more

towards technology and innovations and might disturb the output. The results would be totally different if the study was carried out for a Level 4 Automation as it doesn't require driver's assistance at all. **Method:** The methods used by the researcher were semi-directive interviews and the questionnaires. The results of the pilot study were analysed using exploratory factor analysis and the factors were chosen using principal component analysis and later the descriptive Analysis was done using Linear Regression Analysis on R. Hypothesis testing was conducted after the two pilot methods were deployed and using the hypothesis testing, results were derived using a questionnaire which had six sections to elaborate the purpose of the study to the users and explained each and everything.
Author: (Kyriakidis *et al.* 2015) **Field of Study:** Bio-Mechanical Engineering **Study area**: Netherlands **Overview**: 5000 respondents from 109 countries were surveyed using an online questionnaire to find out the user acceptance, concerns and willingness to buy Self-Driving Cars. 33% people found autonomous driving as enjoyable while majority voted in favour of manual driving. The results show that respondents were most concerned about software hacking, misuse and data privacy of autonomous vehicles. **Limitation:** The research uses a small sample size for 40 countries which has at least only 25 respondents which could give unreliable statistics for those countries. The very low-income countries (African countries) were not included in the research which could have easily affected the results. The average age of sample size is 30 years which could give results in one dimension rather than a spread-out result because usually for the older people it is difficult to accept the new technology. **Method:** A 63 Questions questionnaire was prepared online, and 5000 responses were collected using a reward of 0.30$ for filling the questionnaire. The instructions about the questionnaire and general information were all given at the start of the questionnaire. There was no country restriction so many responses could be collected. After getting responses from a huge number of people, they were filtered out and statistics were performed on about 4886 responses. The Analyses were done by calculating Spearman correlation coefficients (criteria for statistical significance at some confidence interval) for each response at individual level and at country level to get the number of opinions.
Author: (V.Casely 2013) **Field of Study:** Computer Science and Robotics Engineering **Study area**: United Kingdom **Overview**: The study carried out involved 450 participants and it demonstrates the willingness of people as in which to what extent they are ready to buy autonomous cars and what factors would

they most likely be considering once autonomous cars are there in the market. The research covers the primary factors (cost, safety and legalization) extensively. Those factors were tested for all of gender, age, demographics etc. The research was carried out using a survey and not interviews which allowed to have a big sample size and reliable results. **Limitation:** The research didn't focus much on the secondary factors (fuel efficiency, Productivity of the user, Environmental impacts of the car) influencing people's decision of shifting to autonomous vehicles which were more important as far as the results of the study are discussed. The sample size includes a large age group of young people and is not equally distributed among different age groups holding different opinions. **Method:** The research uses an online survey method to gather data regarding the factors influencing people's decision to buy autonomous cars in future. For taking the opinions of the new young drivers there were some questionnaires sent to a local school after taking special permission from the school. Monetary gift was 0.50$ for each response on the websites. The results were later analysed using a statistical software which was SPSS.
Author: (Kaur and Rampersad 2018) **Field of Study:** Engineering and technology Management **Study area**: Australia **Overview**: 101 students of Tonsley were surveyed using an online survey questionnaire. The research study investigates the key factors which affects the adoption of self-driving cars. The study area was confined because the University will start self-driven cars inside the University to help staff and students with drop offs from the bus and train stations. The research identified key factors influencing AV adoption like performance expectancy, reliability, security, privacy and trust. **Limitation:** The research was conducted in close environment (inside a university campus) which might have had a huge effect on the results if conducted for other close environments. The research has unequal demographic variables like (males 37% and 62% females) which could affect the outcomes too. **Method:** The research was conducted using an online survey in which 101 students and staff of the University participated and was asked 10 questions with various sub questions. The survey was conducted anonymously. After questionnaire results were quantitatively and descriptively analysed, there was hypothesis validity testing and then statistical analysis was performed by using confirmatory factor analysis (CFA) with SPSS and AMOS software's.
Author: (Dikmen and Burns 2017) **Field of Study:** System Design Engineering **Study area**: Canada **Overview**: 162 Tesla drivers (16 to 65 years age and older) were surveyed using an online questionnaire to ask about their experiences and attitudes with Summon and Auto-Pilot Systems

to value the trust they had in those systems. 62.4% people reported that they experienced an un usual behaviour using Auto-Pilot and 21.2% using Summon. **Limitation:** The research was not done immediately after the incidents and in the meantime the user perspective might have changed which could be covered by a longitudinal survey for a few years in a row to identify the fluctuations in the level of trust. The factors on which the trust in AV is made need to be discussed as well. **Method:** The research was conducted using an online questionnaire which reached to 162 Tesla drivers but only 121 fully completed the survey. There were 5-point Likert scales in which Drivers were asked to rate their trust about using Auto-Pilot and Summon. Different variables were analysed using the Mean and Standard Deviations and the trust for Auto-Pilot and Summon were separately analysed for Initial and Current trust before and after incident.
Author: (Howard 2013) **Field of Study:** City and Regional Planning **Study area**: United States **Overview:** In this research study, 107 participants were surveyed to get responses of their attitude towards the self-driven cars. The survey had all basic questions from demographic perspective to individual travel mode choice behaviour which were designed to get the opinions of different households in Berkeley California. **Limitation:** The research was a bit biased because of the survey distribution method along with the self-selection of participants and the place to conduct the survey. **Method:** The research was conducted through a survey which was filled by a107 regular visitors of the public science museum. The visitors had to fill a survey, watch a 10-minute video and then give an interview. The sample size was analysed using a logit model.
Author: (Fraedrich et al. 2016) **Field of Study:** Transport Planning **Study area:** Germany **Overview:** In this research study, 1000 respondents were surveyed through an online survey which mainly focused on the attitude and approaches of the people regarding autonomous driving in Germany. **Limitation:** The research has more female participants compared to the actual German population (56%) and it contains more single living people with a lower educational level as compared to the German population so there is un-equality in demographics. **Method:** The research was conducted through online survey questionnaire and 1000 respondents were found to have completed the survey. After the survey, interviews were carried out of 250 participants and key indicators were compared to the corresponding shares of the Mobility survey

for Germany in 2008. The factors were analysed using a Boolean Variable DAS. For co-relation of variables Pearson's chi-square test was used.

2.10 Conclusions drawn from Literature Review or Limitations of current literature

It was found out through the above literature that no study fully focuses on the safety factors with respect to the user perspectives to shift from conventional vehicles to choose autonomous vehicles. The studies which have been already done in the field lack a multi-perspective approach. They also lack an impartial demographic survey, longitudinal survey with respect to user experiences with the field of automation. Productivity of the User, Environmental Impacts, and fuel efficiency is one of the few factors which play an important role in people's willingness to choose self-driven cars and haven't been studied in detail.

There is a need of a purely practical approach which is impartial with respect to the participants involved in age, gender and location which is a blend of the people working, not working etc., has a large sample size (considerable to the city population) and basically covers the factors that affect the willingness of the people to shift from their conventional vehicles to autonomous cars considering the safety perspective and is fully concluded with the suggestions and mitigations on how to overcome it.

Chapter 3: Methodology

3.1 Introduction

This chapter will describe all the necessary methodological steps in order to fulfil the aim and objectives of this project. More specifically, the development of the questionnaire survey, as well as the analysis methods are going to be identified, so as to associate the user perspective of safety with a potential shift to autonomous vehicles.

The methodology adopted to perform this research work is to construct a questionnaire survey comprised of different sections keeping in mind the aim of the study which will help in collecting the data for the research. The questionnaire has been chosen as a source to get the data as it is the easiest and the most effective way available for this kind of research. Moreover, more awareness among the respondents could be recorded using this method in little period of time as compared to using any other method like telephonic interviews or face to face interviews *(Hollier et al. 2017)*.The questionnaire is developed after reviewing the aim of the study and will be helpful in analysing the data to reach certain conclusions and it is distributed as an online survey amongst different age groups having different employment statuses and focus on every citizen of the Munich city using social media and other electronic marketing methods. The sample comprises of working students, full time employees or old age people who rarely drive because of their disabilities or do not like the idea of driving the car manually. The safety perspective has been chosen as the basis of doing this survey as there is much concern regarding this particular feature with respect to autonomous driving and not much has been studied in this respect. Munich, being one of the hubs of automotive sector is already doing testing and research on the autonomous driving and hence it would be interesting to see the response of its residents on how much they would consider to shifting towards an autonomous car in near future. The response of the respondents will be recorded and analysed further to reach the conclusions.

It is pursued to identify the factors that the people of Munich will consider if they shift from their normal conventional vehicles to autonomous vehicles and what percentage of people would like to shift to autonomous driving.

3.2 Flowchart of methodology

The methodology can be overviewed from the flow-chart diagram in (Figure 3.1) that depicts the different stages of the work and the steps needed to get the results at the end. It includes the construction of the questionnaire which involves the demographics section (questions related to the person), transport section (questions related to general driving experience), safety section (questions related to general road safety) and safety section concerning autonomous vehicles (questions related to safety aspects of autonomous cars).After the construction of the survey, it is distributed online through social media and specially Facebook groups in Munich. The questionnaire is kept online for one month to collect as much data as possible. In this time period 272 responses are gathered. After the collection of the data, it is analysed using different statistical models like Ordinal Logistic Regression and Factor Analysis. After the analysis and interpretation of results from those models, conclusions are drawn.

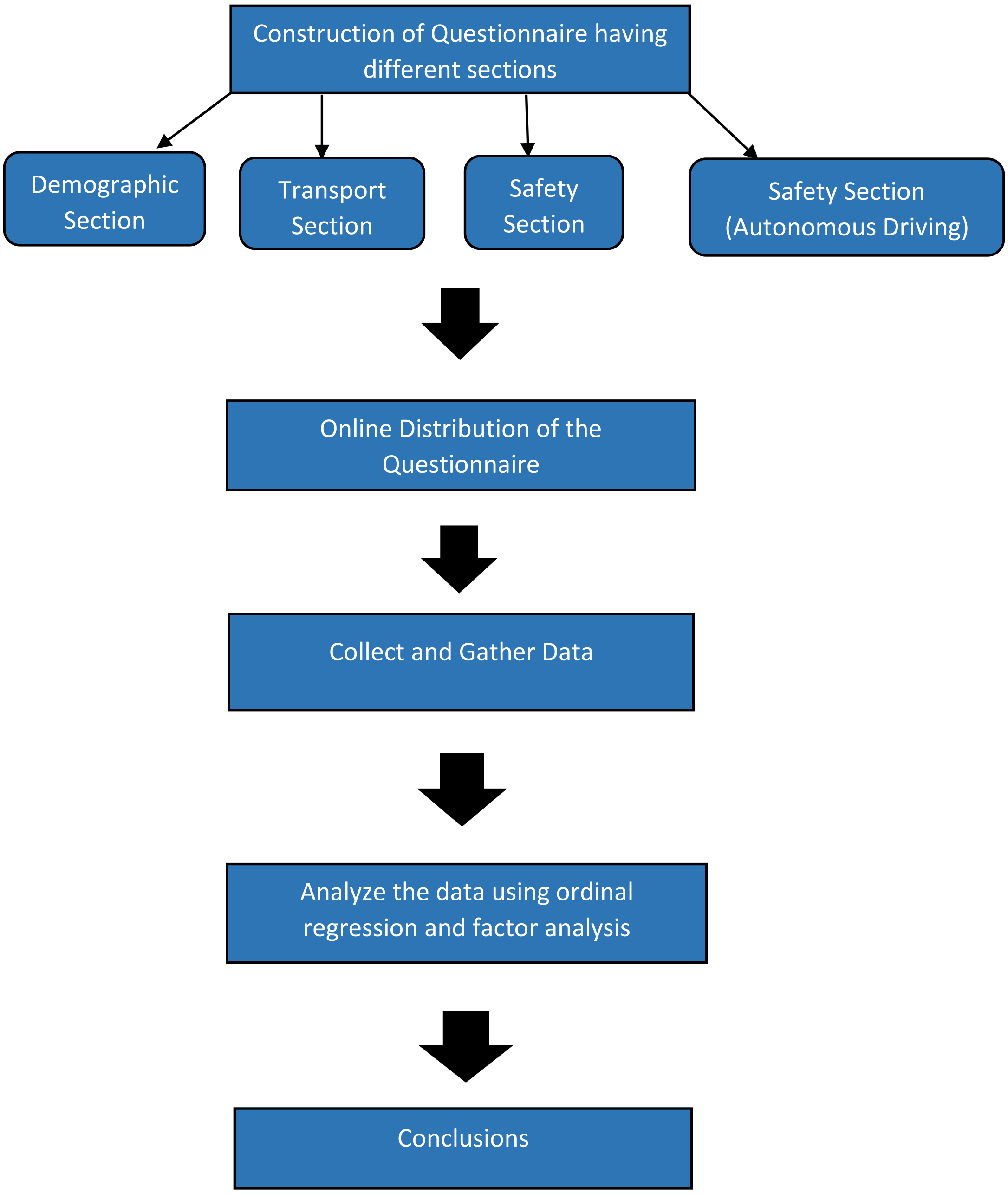

Figure 3.1: Flowchart representing the methodology of the study

3.3 Steps to create general online survey questionnaires

The first and foremost step is to figure out the aim and objective which is purpose of the research. Once the objective is set up, it is needed to work backwards to look out for the data and information that is required for the objective, avoiding questions which induce biasness in the survey. The language is kept simple and understandable to be able to get more survey questionnaires filled out by the participants. Later the feedback is collected using pilot surveying. The data rolling from the survey helps getting near to the aim *(SurveyMonkey)*.

3.4 Methods to develop a research questionnaire

To develop a research questionnaire, at first a list of objectives is written to determine the questions that the survey will use which might be fixed-response/ structured questions (specified set of possible answers) or open-ended/ non-structured questions (scope of the answer has no limits). Then the method of administration is decided to collect the data which could either be face to face interviews, web-based online surveys, telephonic Interviews or written questionnaires. The wording of the questions is selected effectively. The things which should be taken care of in the wordings of questionnaire are comprehension (simple language and brief sentences), answerability (answers can be selected without doing more research), double-barrelled questions (asking one thing at a time and not doubling up the questions) and the response options shouldn't overlap. After that, the survey questionnaire should be organized in a logical way which means that if there is a range of topics, then questions should be bundled by grouping the sensitive questions with neutral and should be asked towards the end of the survey. The questionnaire should be easy to navigate and read which means that a large clear font should be used. There should be enough space for open ended questions, and enough space between consecutive questions to see where one ended and the next begins. Page numbers should be used to make it easy for the reader and answer space should be provided directly under the question. The participants should be provided with the sufficient information they will need to complete the survey effectively. The purpose of the questionnaire should be properly explained. The instructions should be given to fill up the survey. The question format should be explained (i.e. multiple choice, rating scale, etc.) and an example for how to appropriately answer the question should be provided alongside and the total number of questions and the time required to fill the questionnaire should be mentioned in the start of the survey. In the end the revision of the questionnaire is necessary which means if certain questions are being skipped consistently then they need to be reworded. Changing of the layout due to insufficient space below the question and a multiple-choice format could be chosen depending on the range of results not being fulfilled with a yes/no *(ScienceBuddies).*

There are three stages before which the data analysis of a questionnaire can be carried out. These are: survey sampling, questionnaire design and survey administration. Survey sampling includes the selection of the target audience and defining the population with an appropriate sample having a relevant sample size *(Park 2017)*. The selected population, should be specifically based on the topic of the research question, aim of the study and the unit of analysis. For example, the questionnaire used in the study, the potential population includes the residents of the city of Munich. After choosing the population the next step is to choose an appropriate sampling method keeping in mind that the sample should be representative of the population *(Park 2017)*. There are different sampling methods that are used in research like stratified sampling, cluster sampling, systematic sampling, simple random sampling and multistage sampling. *(Etikan 2017)*. The sampling method adopted for this research is random sampling because it can exclude the arbitrariness of the research and manipulation in sampling plus it is easy to use and increases the accuracy of representation at the same time for a large population size *(DePersio 2018)*. The observation number in a random sample is found out by the effect size, the level of significance and the specific data analysis method *(Park 2017)*. The second stage is the questionnaire design which is a survey method to ask other people somethings that are still unknown or insufficiently researched. This method of data collection will not help if the majority of the population have no idea of the subject of the research and will choose "I don't know" or "I can't answer" for most of the questions*(Park 2017).*The respondents personal details like telephone number, address, signature and other things that can help in identifying the respondent should never be asked as it will help in tracing the person and make the survey overall very biased plus the respondents might also not feel good about it due to their leak of personal information *(Park 2017).* The sensitive questions like the age or income of the respondents should not be asked directly because the respondent might feel very uncomfortable with it and instead should be asked within a range like Age 40-50 years or Income 70,000 Euros to 90,000 Euros *(Park 2017).*The respondents become indifferent at the ending part of the questionnaire due to time, fatigue and stress of filling a long questionnaire, hence the closing questions should be formed in such a way that potentially sensitive questions should be left for the ending part of it so that it stops the respondents from leaving them before the collection of important information *(Crawford 1997).*

The last stage of the gist of the questionnaire is the administration of the survey. There are different methods on how to distribute it amongst the targeted audience and the advantages and disadvantages of each method should be considered before finalizing a method *(Park 2017).*

These were the basic steps needed to be in mind when developing a research questionnaire. A questionnaire for a survey should be as unbiased as possible. The basic requirement of the survey is to provide unbiased data so that as realistic as possible analysis could be performed on the answers gathered through the questionnaire. The questions in the questionnaire should also be unbiased when asked and should have options that are as close as practically possible. Most importantly the questionnaire should find the desired target audience to get the desired results. The questions wordings and its content should be such that it doesn't force the respondents to leave the questionnaire in the middle somewhere un-attempted.

There are many methods that have been used to evaluate the attitudes or how the people feel about something that they have already experienced or something that will be a part of the everyday life in future. For these type of studies, Likert-type scales are used in the options of the questions being asked in the questionnaires. Using Likert-Scale, the respondents indicate how much inclined they are towards successive statements which mostly are related to the state of agreeing or disagreeing to some extent (*Tamara van et al. 2007)*. Likert-type scale produces more information than ranking scale or creating paired comparisons. It is better than the other type of scales because of its ability to record an individual's absolute strength of the attitude *(Tamara van et al. 2007)*.

3.5 Questionnaire

The survey questionnaire for the thesis is compiled keeping in view the principles explained in section 3.4 to make a research survey questionnaire. The questionnaire similar to the outline of the literature consists of four sections having total of 42 questions. The first section consists of the questions that are related more to the general transport sector so that the people who don't drive at all or have no idea about driving a car could be filtered out on the spot and the data gathered could be as clean as possible. The second section is made up of the questions that are relevant to the road safety perspective which includes questions about general safety of the passengers, alcohol impaired driving and involvement in other road accidents. The third section of the questionnaire is directly related to the safety with respect to autonomous driving and how people perceive it. It contains the questions regarding general point of view about the autonomous cars and then going further into how autonomous cars would influence the safety on the road with all the driver's assistance tools and steering less cars. The last section of the questionnaire discusses the demographics of the respondents so that it is easy to do the data analysis. It consists of the questions related to gender, education, marital status and income etc. For the analysis part, only a few questions from the total 42 questions of the questionnaire are chosen based on the importance of the questions with respect to the direct relevance with the topic of the thesis and how they could be analysed. The full questionnaire is attached with the appendix of this thesis.

3.6 Data Analysis

The last part of the methodology is the analysis of the data being gathered through the survey to accumulate all the research results and declare the findings. Every methodology needs a specific data analysis technique relevant to the aim and study of the research. The data analysis depends on the type of research being performed. The qualitative data analysis is different to that of the quantitative data. The data analysis techniques depend on which scale is being used in the questionnaire. Either it is nominal, ordinal, interval, ratio, etc.

According to Vosloo (2014) , data analysis is a process that gives order, structure and meaning to a raw data collection procedure *(Vosloo 2014).* The goals of data analysis according to *(Atlas.ti 2011) are:*

1) To give meaning to the data collected
2) To find out the patterns and relationships amongst the data collected
3) To find out general information about the topic of research which the researcher was not aware of before doing research

There are mainly three preliminary stages of Data Analysis *(SSC 2001)*

- Exploratory Data Analysis
- Deriving the main findings
- Archiving

These will be described in the sections that are following.

3.6.1 Exploratory Data Analysis

Exploratory Data Analysis (EDA) is the stage of data analysis in which the researcher gets an overview of the data files before compilation, so as to have an idea of what is being gathered and stored. It helps the researcher in such a way that it gives an idea if there is any need for some additional information, additional questions or if the data collected thus far is sufficient to get to the next stage of analysis. Alternatively, the researcher should stop collecting the data because the results are depicted with the number of observations/surveys already been done. The conclusions made from EDA cannot be accepted as the final results, but this stage helps cleaning the data set and the variations in the data become prominent. If the exploratory data analysis is done in an ideal manner, then the researcher could end up having sufficient confidence to have a clean data set and that version is the final one with all the further analysis being done with the help of that single file of data. In this phase, the form of analysis could be made decisive and the plans for further analysis like coding or programming which could be done side by side with the data collection if required could be made and last but not the least data entry and checking is also the part of the EDA. A complete EDA makes the latter stages of the data analysis relatively simple, steady, ordered and without any arguments *(SSC 2001).*

3.6.2 Deriving the main findings

This is the second stage of the analysis where the further analysis is done on that clean final version of the data set which was cleaned in the exploratory data analysis. This helps the researcher in keeping the analysis files as consistent as possible with one another without any variations or limitations in the data. As the name says, this stage produces the main findings interpretations and recommendations for the researchers to let them start utilizing the results of the study. There could be some extra tasks involved in this stage that could take more time for the unforeseen circumstances which the researcher didn't expect as being planned earlier. Moreover, if the data is altered because of some data entry error which was not found in the earlier stage then help from the software can be taken which can keep a record of the analysis, and it is easy to make changes to the previous data *(SSC 2001).*

3.6.3 Archiving

This is the last stage of this data analysis. It involves the collection of the data and keeping it as a record on a CD, Hard Drive or any storage device and it includes all the stuff used to gather information. Archiving is necessary to keep a good management system of the surveys. It helps in organizing, authenticating, satisfying the sponsor, linking different studies together and obtaining different results after altering administrative boundaries in the research process *(SSC 2001).*

3.7 Qualitative data analysis

The first step in the analysis of qualitative data is to organize the data. The organizing will be different for different research strategies and data collection methods. After the data is organized, the second step in analysing the data is description. In this stage, the different aspects of the study are described by the researcher including the individuals being examined or the reason behind any activity performed including the opinion of the participants and the effects of activities on them. The final phase is the process of interpretation which is the stage in which there is description of the findings and the answers to the why questions and finding the patterns to create an analytic framework *(Vosloo 2014).*

According to *(Vosloo 2014)* there are some general guidelines that are used by the researchers for reading descriptive data. These include:

- Patterns and trends should be noted
- Clustering of the events, people, places if they possess similar characteristics
- Metaphors should be made which help in reducing data and making patterns to help connect data with theory
- Counting frequency of happening of recurrent events
- Establish similarities and differences between and within data sets
- Splitting variables to have more detailed descriptions
- Linking specific data to general concepts and categories
- Attempting to discover the factors involved in the process under study
- Relationship between variables should be studied using matrix display
- Finding intervening variables to establish the presence and effects of variables intervening between observed variables
- Developing logical relationships to understand the trends and patterns

3.8 Performing Data Analysis

Data Analysis is performed using a suitable analysis method using a software or coding. The suitability depends on the type of scale used in the questionnaire, the aim of the study and the comfort of the researcher.

For the current thesis topic, the 5-ert Likert scale is used because it offers a range of options from one extreme to another. It is normally used to measure opinions and attitudes. This current survey is about Users perspective and the measure of how strongly they agree or disagree to a certain opinion *(SurveyMonkey)*. Hence, this scale is used. More specifically the categories used for each question were chosen to be:

1. Very unsafe/unlikely/less/low
2. Rather unsafe/unlikely/less/low
3. Moderate
4. Rather safe/likely/high/high
5. Very safe/likely/high/high

To perform data analysis on the Likert scales used in survey research there have been many methods used already as regression analysis performed by *(Hulse et al. 2018)* to interpret the survey.

Few of the methods used are described below:

1) Cronbach's Alpha
2) Factor Analysis
3) Pearson Product Moment
4) Cluster Analysis
5) Regression Analysis

3.8.1 Interpreting Cronbach's Alpha

Cronbach's Alpha is a measure of internal consistency which is also mentioned as reliability. It is a very common technique used for the data analysis when there are multiple Likert scale questions in a survey/questionnaire. It interprets the scale being used in the questionnaire and determines if the scale is reliable *(Goforth 2015).*

In mathematical terms, Cronbach's alpha could be defined as follows: *(Goforth 2015)*

$$\propto = \frac{k \times \bar{c}}{\bar{v} + (k-1)\bar{c}} \qquad (3.1)$$

where:

k indicates the number of scale items

$\bar{c}$ indicates the average of all covariances between items

$\bar{v}$ indicates the average variance of each item

$\propto$ indicates the Cronbach's Alpha

Cronbach's alpha is computed by correlating the score of each scale item with the total score for each observation which is usually the individual survey respondent and then it is compared with the variance for all individual item scores *(Goforth 2015)*. Cronbach's alpha reliability co-efficient ranges from 0 to 1, mostly between 0.5 and 0.8. However, there is no lower limit of the co-efficient. Cronbach's alpha cannot be used to determine the reliability estimates for single items. Moreover, dimensionality cannot be measured using Cronbach's alpha. Hence a high co-efficient α could be produced for the scales which have same length and variance and with different more than one underlying dimensions. *(Goforth 2015)*

3.8.2 Factor Analysis

Factor Analysis is another measure that is more often used to analyse the questionnaires. It measures the ability and traits of the questions and to ensure that those questions that have been asked in the questionnaire are related to the construct that was intended to measure. The purpose of the factor analysis is to summarize the data to understand the relationships and patterns to interpret them in an easy manner. Mostly, it is used to regroup the variables making a limited group of clusters with them based on shared variance *(Gie Yong and Pearce 2013).*

Factor Analysis is used when the data set contains a long list of variables, and that long list of variables is then grouped which are called factors. This analysis groups similar variables into descriptive categories. This method is very beneficial for studies that have hundreds of variables that take too much time to get the analysis done and are unimportant as well. So, with factor analysis the similar variables are grouped and analysed easily efficiently using a statistical software which i.e. SPSS or R *(Gie Yong and Pearce 2013).*

Factor Analysis could be performed on the data if it has univariate and multivariate normality and is free of univariate and multivariate outliers. It is assumed that there is a linear relationship between the factors and variables when making correlations. A factor is only considered as a factor when it has three or more variables and if there are two or less than two variables then they should be massively correlated with one another ($r>0.70$) and shouldn't correlate with other variables. The sample size recommended for factor analysis should be of at least 300 participants. The variables on which factor analysis is going to be applied upon should have at least 5 to 10 observations and the respondents to variables ratio should be 10:1. Large sample data reduces the errors in the data and Exploratory Factor Analysis works better with large sample sizes. A smaller sample size ($n>150$) could be used if there is a huge number of high factor loading scores (>0.80). The variable contribution to the factor is generally measured by a term called factor loading for a variable. The variables should have a strong correlation r value which should be 0.30 or greater because anything lower than this value would indicate a weak relationship between the variables. A heterogeneous sample should be used instead of using a homogenous sample as homogenous samples affect the variance and factor loadings in a negative way. The sample size should be considered taking care of the missing values in the data. Squared Multiple Correlation (SMC) values for singularity and multicollinearity should be checked in the data set *(Gie Yong and Pearce 2013).*

Factor Analysis could be performed with variables using a correlation matrix or a covariance matrix. It is performed using a correlation matrix when the variables cannot be compared like items from different scales while it could be performed using a covariance matrix in which variables are identical and the items are from the same scales *(Gie Yong and Pearce 2013).* Correlation matrices are in general easier to interpret than Covariance matrices. In general, Factor Analysis is computed, and calculations are made by using matrix algebra. The main statistical tool used by factor analysis is the correlation co-efficient which explains the relationship among two variables.

To understand ***classical factor analysis*** using mathematical model, let's suppose '**p**' indicates the number of variables (**X_1, X_2,……….X_p)** while '**m**' indicates the number of underlying factors (**F_1, F_2,………….F_m**). **X_j** is the variable which is represented in latent factors. The model infers the presence of '**m**' underlying factors where each observed variable is a linear function of these factors along with a residual variate *(Gie Yong and Pearce 2013).*

This model could replicate the maximum correlations *(Gie Yong and Pearce 2013)*

$$Xj = aj1F1 + aj2F2 + \dots\dots\dots\dots ajmFm + ej \quad (\mathbf{3.2})$$

Where j= 1, 2,………. p and a_{j1}, a_{j2},………a_{jm} are the factor loadings. The equation explains that a_{j1} represents the factor loading of j^{th} variable on the 1^{st} factor. The factor **e_j** represents a unique factor *(Gie Yong and Pearce 2013).*

The factor loadings are helpful as they tell about the contribution of the variable to the factor. The greater the value of the factor loadings the more it would have contributed to the factor. The limitation of this method could be that naming the factors could be problematic and some factor names might not reflect the exact nature of the variables *(Gie Yong and Pearce 2013).*

Factor analysis method has been used quite a lot in the studies that have already been done in the field of autonomous driving. *(Kaur and Rampersad 2018)* uses factor analysis to perform the statistical analysis with SPSS and AMOS software to quantitatively analyse the survey questionnaire to know the key factors in adoption to self-driving cars. *(Payre et al. 2014)* also did the initial analysis using exploratory factor analysis and the factors were chosen using principal component analysis and later the descriptive analysis was done using linear regression analysis on R to do a priori study to find the attitudes, personality traits and intentions to use a Fully Automated Vehicle.*(Haboucha et al. 2017)* also conducted a study in the same background which wanted to know the preference of the people for future mode of transport for daily commutes based on certain attitudinal variables. The characteristics are quantified through random utility models like logit kernel model and nested model. Afterwards SPSS and AMOS are used to reduce attitudinal variables through factor analysis.

3.8.3 Pearson Product Moment Correlation

The Pearson product moment correlation co-efficient is used to measure the strength of linear relationship between two variables and it is normally denoted by r *(Chee 2015).*

$$r = \frac{n\,(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}} \quad (\mathbf{3.3})$$

Where

r =Pearson's correlation coefficient

n= number of paired scores

X= score of first variable

Y= score of second variable

XY= the product of X and Y paired scores

The values taken by the co-efficient r range from +1 to -1 while a value of 0 indicates that there is no relationship between the two variables. The strength of the relationship between two variables will be judged by the Pearson Correlation Coefficient r, which will be near to +1 or -1 dependent on the relationship whether positive or negative. If you get a value of +1 or -1 it means that all the data points are included on the line of best fit and there are no data points to show any variation from this line of best fit *(Chee 2015).* Assumptions that are made with Pearson's correlation are that the variable type should be either interval or ratio measurements having normal distribution amongst them and associated with each other linearly. The outliers must be removed completely or minimized to a limit. Data should have a certain level of homoscedasticity which means that the error term (noise or random disturbance in the relationship between two variables) is the same across all values of independent variable *(Chee 2015)*

Pearson Product Moment correlation is only used for linear relationships and no meaningful relationship could be established if the correlation coefficients are 0 *(Chee 2015)*.

3.8.4 Cluster Analysis

Cluster Analysis also called as segmentation analysis or taxonomy analysis is an exploratory analysis which tends to pick out the structures within the data. It is an attempt to point out cases having similar or homogenous groups if the grouping was not previously done *(Tan et al. 2005)*. The quality of the clustering largely depends on the similarity within the groups them and the distinction between the groups. The greater the similarity within them and the greater the difference between them, the better will be the Clustering *(Tan et al. 2005)*. This analysis doesn't distinguish between independent or dependent variables due to its characteristic of being exploratory *(Cornish 2007)*. SPSS software can handle different cluster analysis methods having binary, ordinal, nominal and scale (ratio or interval) data. Usually the cluster analysis method is used in combination with other data analysis methods like discriminant analysis etc *(Norušis 2012)*.

In general, there are three different procedures which can be used to cluster data: *(Norušis 2012)*

- Hierarchical cluster analysis (when the research has a small data set with an increase in the number of clusters)
- K-means cluster analysis (when the research has a moderately sized data file and it is already known how many clusters you want to analyse the data)
- Two-step-cluster analysis (when the research has a large data file with a mix of continuous and categorical variables)

Clustering is commonly referred to as an entire collection of clusters. There are many different types of clustering's. For example: Hierarchical (nested) clustering, Partitional (unnested) clustering, Exclusive clustering, Overlapping clustering, fuzzy clustering, Complete clustering and Partial clustering *(Tan et al. 2005)*. **Hierarchical** clustering allows clusters to have sub-clusters. It is a set of nested clusters which are formed in an organized manner as a tree. **Partitional** clustering allows the set of data objects to be further divided into non-overlapping subsets (clusters) in such a way that every data object is in exactly one subset. **Exclusive** clustering permits each object to be assigned to only one single cluster. **Overlapping** clusterings are those in which an object can belong to more than one group at the same time. This type of clustering is mostly used when an object is in the middle of two clusters and could very easily be assigned to any of those clusters. It is also called non-exclusive clustering. In **Fuzzy** clusterings each object belongs to every cluster with a membership weight which is between 0 and 1. Zero means that is absolutely doesn't belong and 1 means that it absolutely belongs. In **Complete** clusterings every object is assigned to a cluster while in **Partial** clusterings every object is not assigned to a cluster *(Tan et al. 2005).*

Cluster Analysis has been one of the data analysis methods being used in the transportation sector. Like *(Yang et al. 2014)* uses K-means cluster analysis to classify the near crash cases into different risk levels using braking process features which are maximum deceleration, average deceleration and percentage reduction in the vehicle kinetic energy to find out the relationship among the braking and triggering factors with the driving risk level while *(Li et al. 2018)* uses k- means clustering to develop an unsupervised classifier to group naturalistic driving encounters into distinguishable clusters by combining an auto-encoder.

3.8.5 Regression Analysis

Regression Analysis is method of analysis which is used to find out the association between a dependent variable (target) and independent variable (predictor). This method is used to find out the causal effect relationship between the variables, for forecasting purposes, time series modelling etc. *(Ray 2015)*. Regression analysis has many benefits. It shows the significant relationships between dependent variable and independent variable. It can also help in depicting the strength of impact of multiple independent variables on a dependent variable. It also permits to make a comparison of the effects of variables measured on different scales *(Ray 2015)*.

The regression techniques are further divided and driven by three metrics which are the number of independent variables, type of dependent variables and the shape of regression line *(Ray 2015).*

1. Linear Regression
2. Logistic Regression
3. Polynomial Regression

3.8.5.1 Linear Regression

In this type of regression, the dependent variable is continuous while the independent variable can be continuous or discrete and the regression line is linear in nature. Basically, linear regression builds a relationship between the dependent variable (Y) and one or more independent variables (X) using a straight line which is commonly known as the regression line *(Ray 2015).*

It is represented by an equation *(Ray 2015)*

$$Y = a + b * X + e \qquad (3.4)$$

where,

a = intercept

b= slope of line

e= error term

The equation is used to predict the value of target variable based on given predictor variable.

Linear Regression can be further divided into two types. The first type is simple linear regression. In simple linear regression there is only one independent variable. The second type is multiple linear regression. In multiple linear regression there are more than one independent variables *(Ray 2015).*

Linear regression is very sensitive to outliers and can affect the regression line and disturbing the forecasted values. Multiple regression can be affected by multicollinearity, autocorrelation and heteroskedasticity *(Ray 2015)*.

Linear regression has been used by *(Hulse et al. 2018)* . In this study, a regression analysis was performed along with analysis of variance and co-variance. The purpose was to analyse survey results with regards to the perception of people regarding AVs. *(Payre et al. 2014)* conducted his survey questionnaire studies to do priori study in order to find out the attitudes, personality traits and intentions to use a fully automated vehicle and descriptively analysed the survey using linear regression analysis on R Studio.

4 Logistic Regression

This type of regression analysis is used to find the probability of event = Success and event = Failure. It is normally used when the dependent variable (Y) is binary which means either 0 or 1, True or False or Yes or No *(Ray 2015).*

The value of Y ranges from 0 to 1 and can be shown with the following equation: *(Ray 2015)*

$$odds = \frac{p}{(1-p)} = \frac{probability\ of\ event\ occurence}{probability\ of\ not\ event\ occurence} \qquad (\mathbf{3.5})$$

$$\ln(odds) = \ln\left(\frac{p}{1-p}\right) \qquad (\mathbf{3.6})$$

$$logit(p) = \ln\left(\frac{p}{(1-p)}\right) = b0 + b1X1 + b2X2 + b3X3 \ldots\ldots\ldots + bkXk \qquad (\mathbf{3.7})$$

Logistic Regression needs large sample sizes because estimates of maximum likelihood are less powerful at low sample sizes than ordinary least square. The independent variables should not be correlated *(Ray 2015).*

5 Polynomial Regression

It is the type of regression in which the power of independent variable is more than 1 *(Ray 2015).*

$$\boldsymbol{y = a + b * x\hat{}2} \qquad (\mathbf{3.8})$$

In polynomial regression the best fit line is a curve that fit into data points rather than a straight line as in linear regression.

Chapter 4: Data Analysis

4.1 Introduction

In this chapter, the process of data analysis for the development of mathematical model concerning the likelihood of people shifting from normal vehicles to autonomous vehicles, the safety of autonomous vehicles and the probability of autonomous cars to eliminate traffic accidents will be discussed. The chapter is basically divided into three sections. The first section discusses the descriptive analysis of the questionnaire and the information provided through the online survey via graphical representation of different variables. The second section describes the methodological approach which was adopted to perform the analysis and the reason for adopting that approach. The last section depicts the output given by the model and the different statistical methods used to obtain different results as well as their interpretation.

4.2 Descriptive Analysis

The survey questionnaire was distributed online among the residents of Munich via social media, particularly Facebook groups in Munich. The questionnaire was kept online on the website for one month and it was filled by 272 respondents till the end of the month, but 72 respondents left it uncompleted, so it was decided to not use this data. Therefore, the final 200 usable questionnaires data were taken into the system and coded into Microsoft Excel. The main purpose of the descriptive analysis is the profile of the respondents and the number of the responses of the questions selected as variables.

Age

In this study there are a total of 200 respondents out of which 64% belong to the age group of (18-31 years) old and 33.5% belong to the age group of 31-50 years while just 1.5% belong to the age group of 51-70 years and the rest didn't prefer to answer or greater than 70 years old.

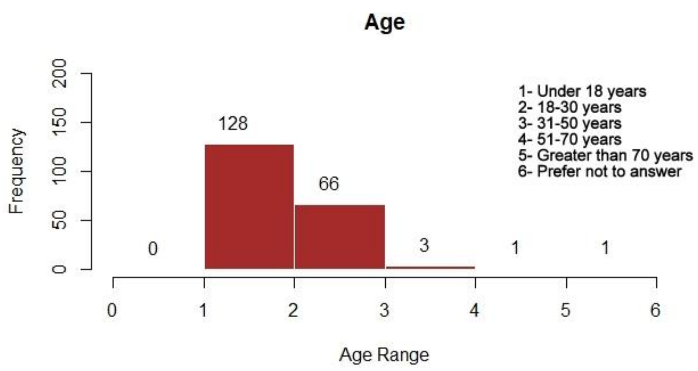

Figure 4.1: Age Group

Gender

The study carried out consisted of 54% male respondents and 43% female respondents out of the total 200 respondents while 2% didn't prefer to answer. Figure 4.2 shows the gender characteristics.

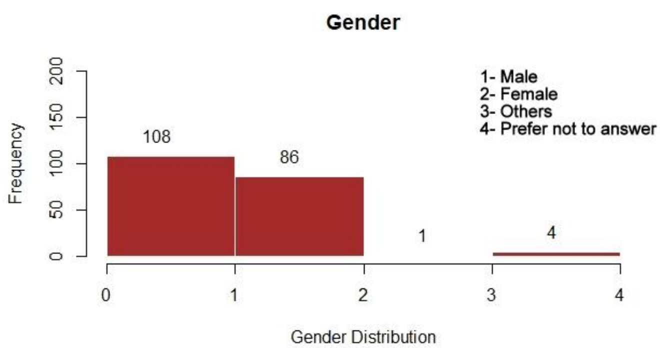

Figure 4.2: Gender Characteristics

Employment status

It was found through the study that majority of the participants (61.5%) were full time workers while 20% participants out of 200 were Students and the rest of the participants were divided among part time employees, public servants, self-employed, homemaker and others as shown in the figure 4.3 below:

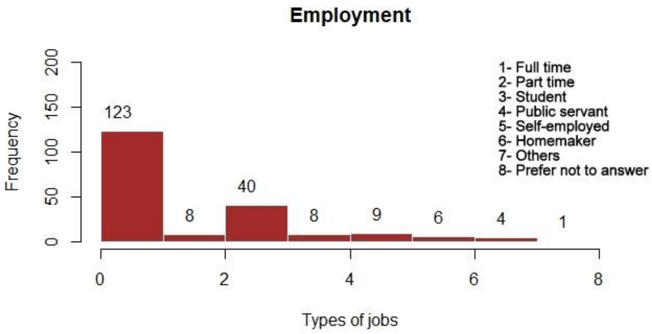

Figure 4.3: Employment status

Mode of transportation

The mode choice of the respondents also varied but most of the participants (61%) use public transportation for their most trips during the day while the second most popular mode choice was the use of a private car (20.5%). 11.5% of the participants use bicycle. The rest of the 7% people use other mode of transport most frequently in a day. This can be depicted in figure 4.4 shown below:

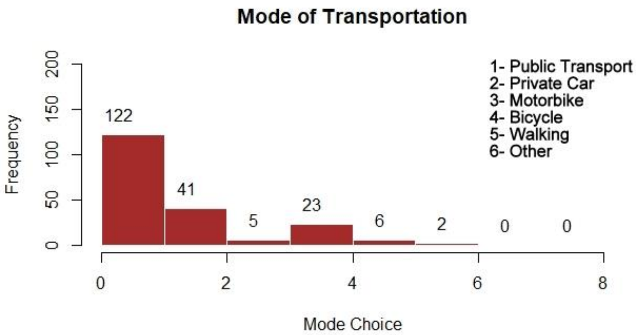

Figure 4.4: Mode of transport

Frequency of driving

Frequency of driving shows that how often the respondents drive a car. It was found that only 19% people drive the car every day while 22.5% drive frequently which means 2-3 days in a week and most of the people (39%) seldom drive which means 2-3 days in a month while 19% respondents never drive a car. The graph 4.5 below shows the results:

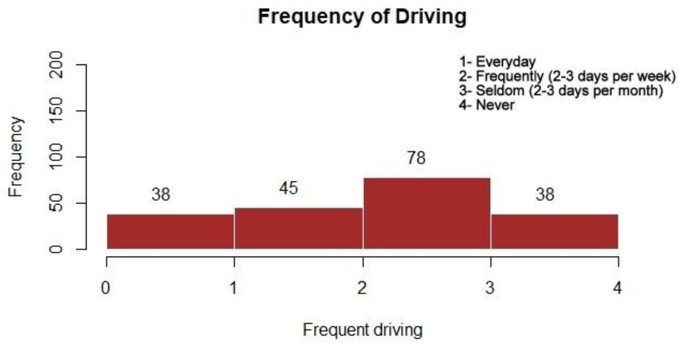

Figure 4.5: Frequency of Driving

Alcohol impaired driving

It was found through the study that the majority of the respondents (87%) declare that they have never driven the vehicle under the influence of alcohol while the rest of them (8.5%) have seldom driven. This is presented in Fig 4.6 below:

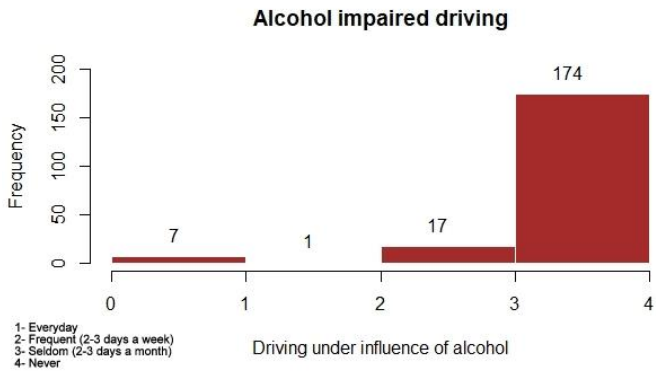

Figure 4.6: Alcohol influenced driving

Performing other tasks while driving

The study shows that most of the respondents (57%) do not perform other tasks like eating food, writing texts while driving, rest of the respondents (43%) do perform these other tasks.

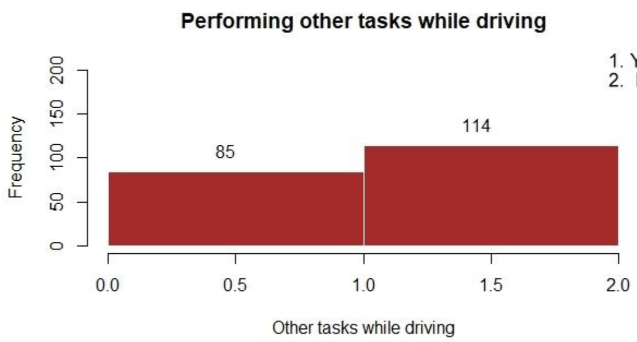

Figure 4.7: Performing other tasks while driving

Preference of automation

The survey questionnaire shows that majority of the respondents (26.5%) prefer partially automated cars closely followed by highly automated (25%) and fully automated (23.5%) while the minorities selected driver assistance (17.5%) and non-autonomous vehicles (7%). The graph 4.8 chart explains more clearly:

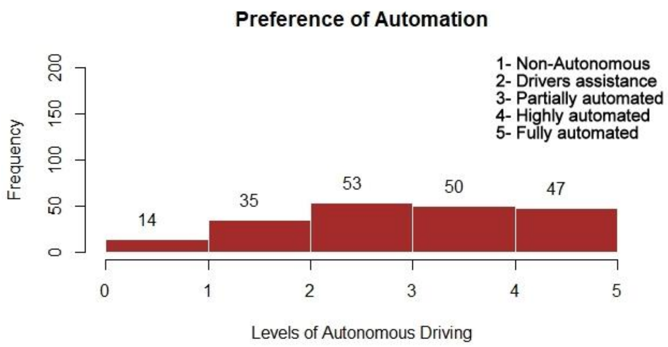

Figure 4.8: Autonomous cars level preference

4.3 Demographic variables distribution

The main demographic variables used for the analysis part are "*age*", "*gender*" and "*employment*". The behaviour of these variables could be seen with respect to the dependent variables to get an image of the distribution of the demographic factors on the different Likert scales.

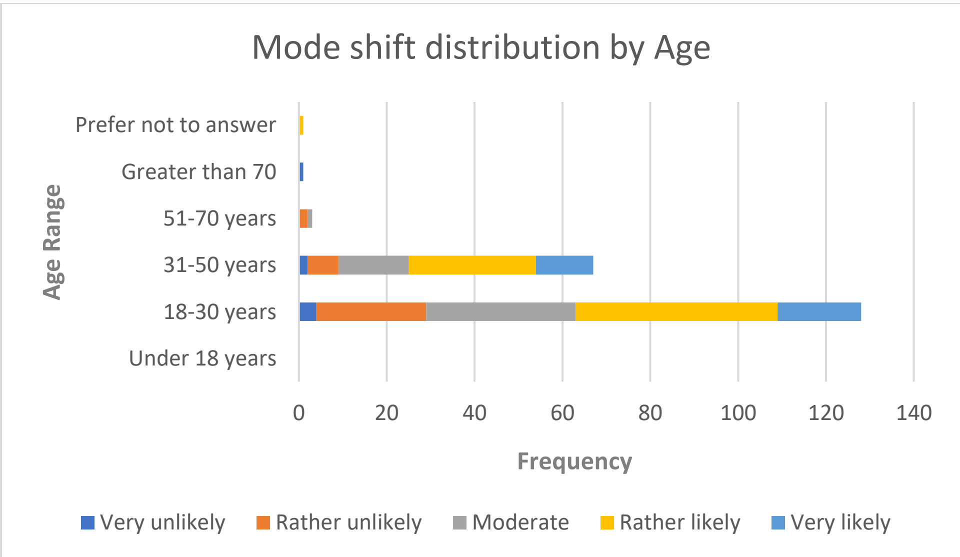

Figure 4.9: Graph between mode shift and age

The distribution shown in Figure 4.9 depicts that majority of the people lying in the age range of 18-30 years are more likely to shift to autonomous vehicles.

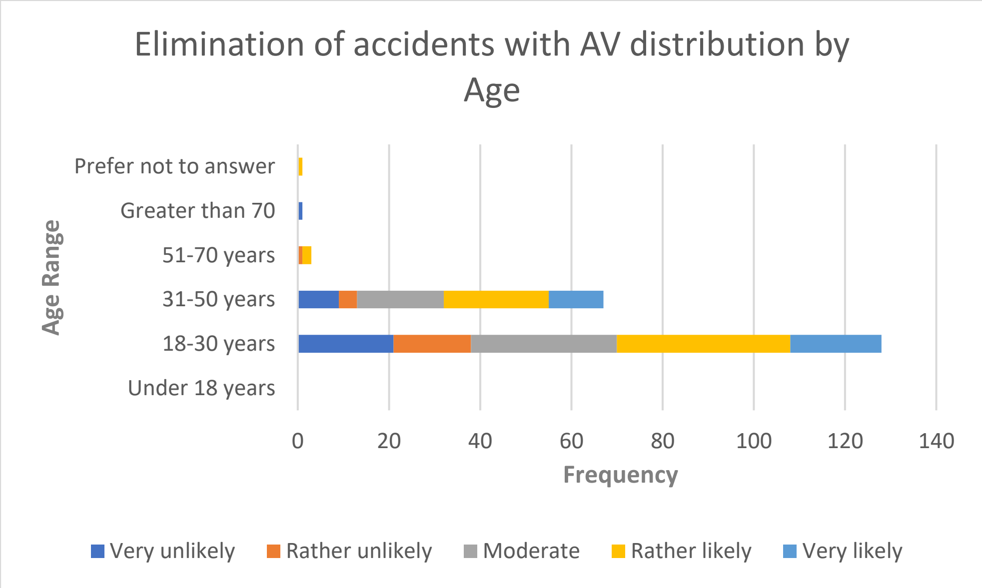

Figure 4.10: Graph between elimination of accidents with AV and Age

In graph shown in Figure 4.10 plotted between the probabilities of elimination of accidents with autonomous vehicles with age shows that more people in the age range of 18-30 years are most likely to think that autonomous vehicles will eliminate traffic accidents.

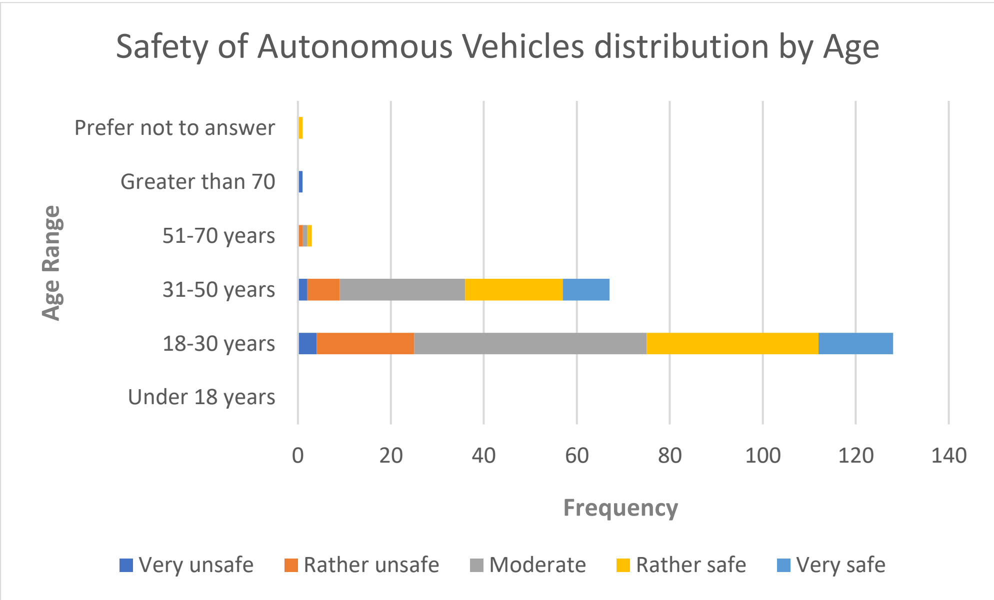

Figure 4.11: Graph between safety of AV and Age

The graph in Figure 4.11 shows that the likelihood of autonomous cars being safer than the human driven vehicles is perceived more by the age range of 18-30 years.

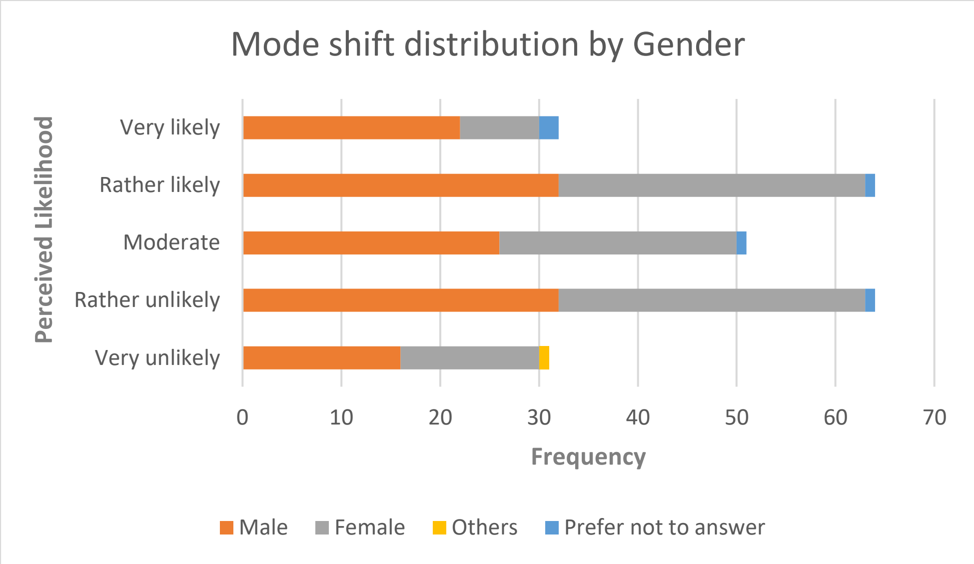

Figure 4.12: Graph between Mode shift and Gender

The graph shown in Figure 4.12 depicts the gender preference to mode shift with respect to shifting from normal vehicles to autonomous vehicles.

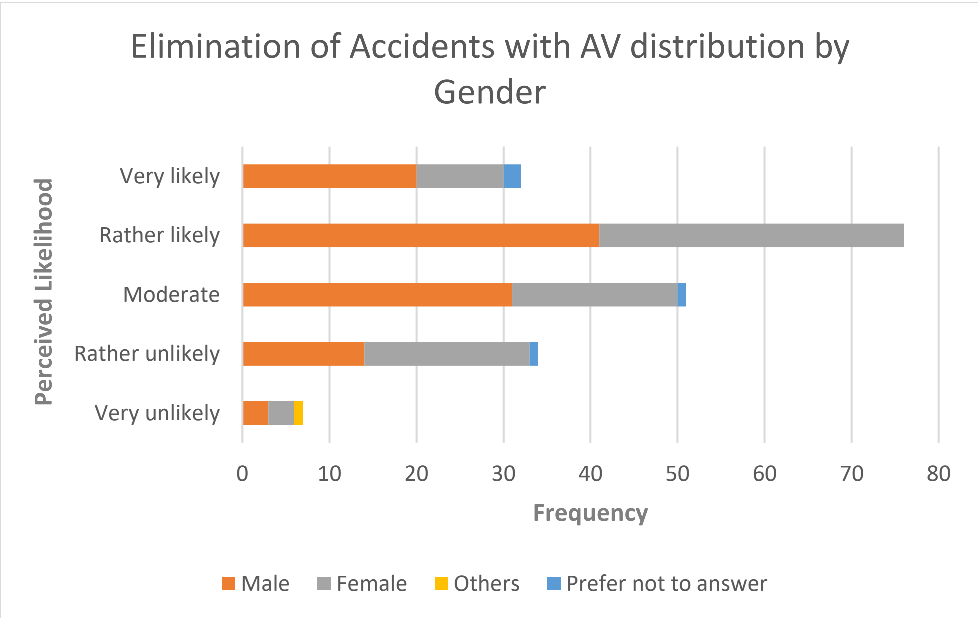

Figure 4.13: Graph between elimination of accidents with AV and Gender

The graph in Figure 4.13 shows the male gender are very likely to perceive the elimination of accidents with the introduction of autonomous cars as compared to female gender.

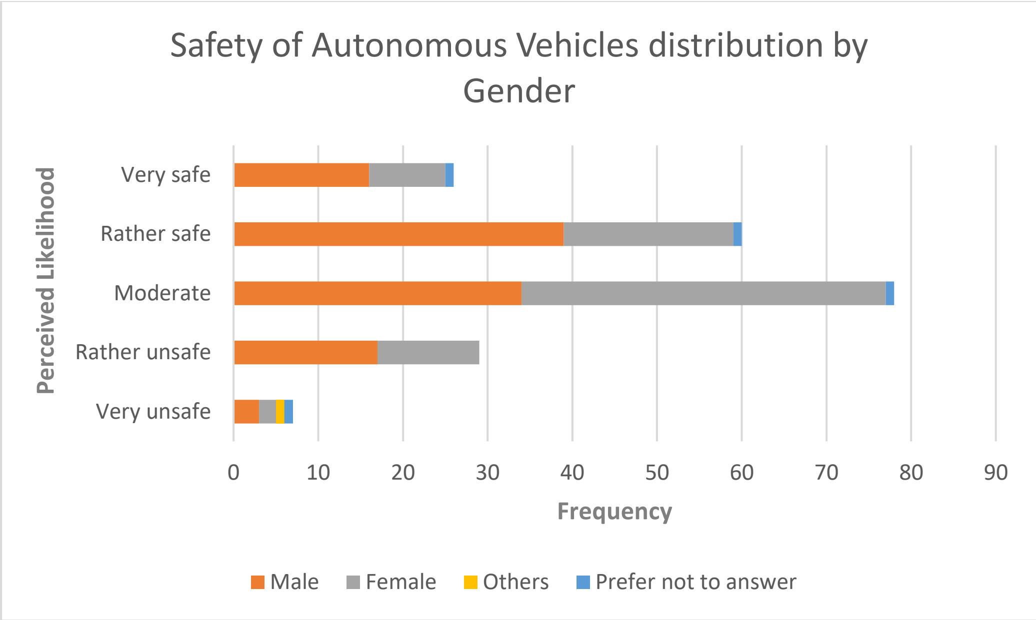

Figure 4.14: Graph between Safety of autonomous vehicles and Gender

The graph shown in Figure 4.14 indicates that majority of the female population choose moderate safety of autonomous vehicles with respect to human driven cars while the male population perceive the autonomous cars rather safer than normal cars.

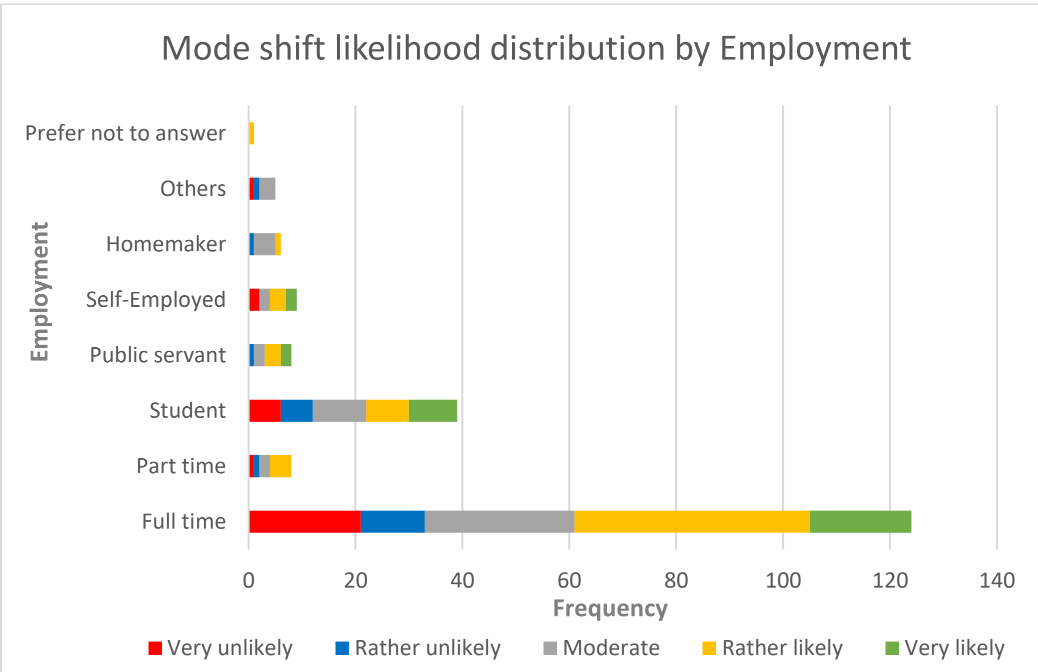

Figure 4.15: Graph between Mode shift and Employment

The graph in Figure 4.15 shows that the people working full-time and students are rather likely to shift towards autonomous vehicles than the people of other professions.

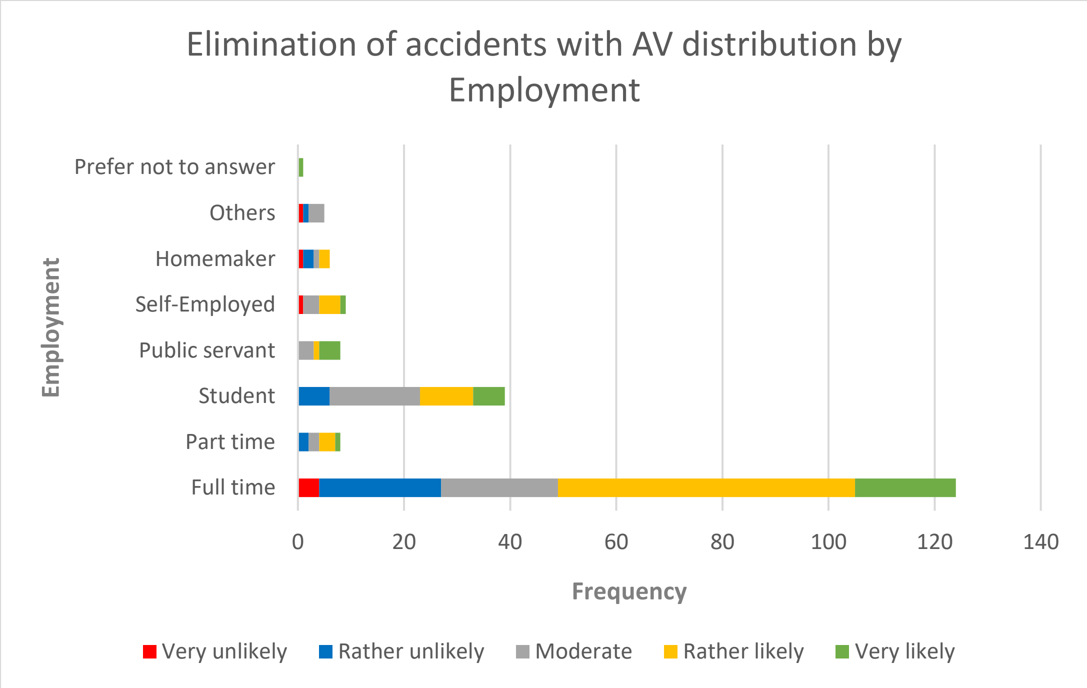

Figure 4.16: Graph between Elimination of accidents with AV and Employment

The graph in Figure 4.16 indicates that the full-time employees and students perceive the elimination of accidents with autonomous cars moderately and rather likely.

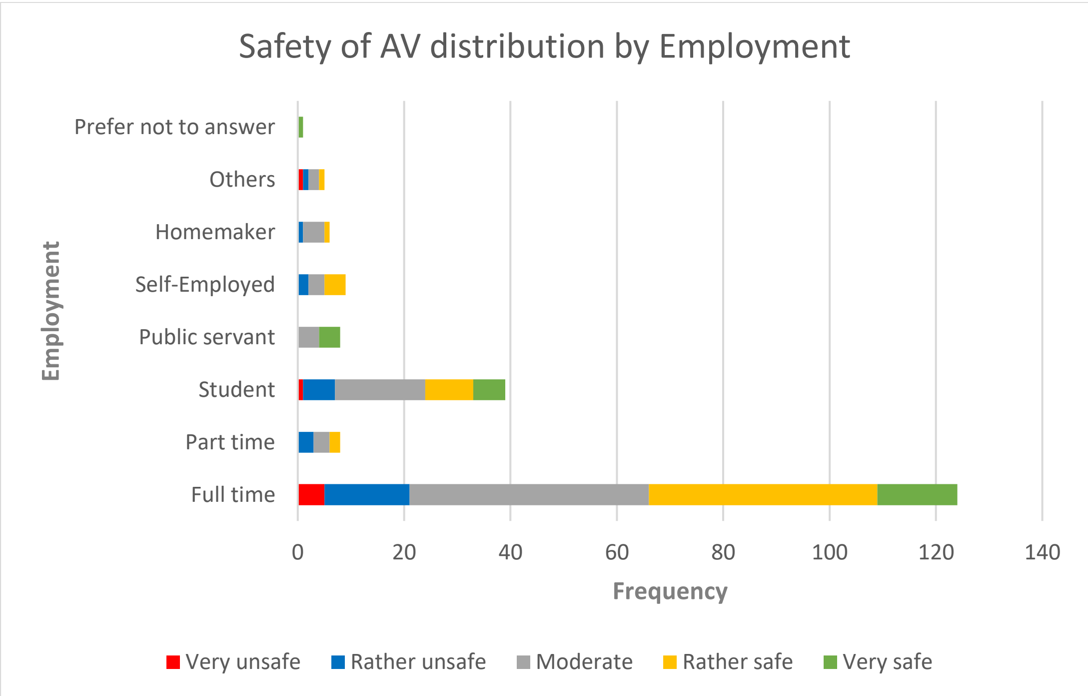

Figure 4.17: Graph between safety of autonomous vehicles and Employment

The graph in figure 4.17 shows that the people working full-time and students perceive the autonomous cars moderate and rather safer as compared to human driven cars.

4.4 Methodological Approach

The purpose of data analysis is to discuss the methods used to organize and analyze the raw data gathered through a survey or a questionnaire to get some validated results with the help of a model constructed through a statistical tool or software to draw some conclusions. For qualitative data, data analysis is a vital part of the research as it links the methodology and the conclusions which summarize the research framework.

The survey questionnaire was made up of 42 questions. Hence, 42 question surveys were distributed among the residents of Munich. 18 variables were chosen from 42 questions. 15 out of those 18 variables were selected as independent variables and the remaining 3 out of 18 variables were chosen as dependent variables. The variables were sorted out based on the type of questions and the influence on the research question in groups of dependent and independent variables.

Dependent variables consisted of the variables which were directly related to the research question and the outcome. The independent variables were the variables which had some influence on the dependent variables and are also called as the predictive variables as they predict the outcome variables. The number of dependent variables and independent variables vary with respect to the type of research and research methodology. Some research works are complete just by using one dependent variable and some use more than one to completely analyze the research question. The predictive variables which were chosen to are:

- Age
- Gender
- Employment
- Mode of transportation
- Frequency of driving
- Involvement in car accident
- Safety in driving yourself
- Other tasks while driving
- Frequency of other tasks
- Vehicle safety being passenger
- Alcohol impaired driving
- Disability preventing from driving
- Safety due to driver assistance systems
- Preference of automation
- Characteristics of AV

Table 4.1 below shows the distribution of different variables into independent and dependent categories.

Table 4.1: List of dependent and independent variables

Independent Variables (Predictive)	Dependent Variables (Outcome)
Age	Likelihood of shifting to Autonomous vehicles
Gender	Perceived safety of autonomous vehicles
Employment status	Perceived likelihood of autonomous cars to eliminate traffic accidents
Primary mode of transportation	
Frequency of driving a car	
Involvement in car accident	
Safety in driving a car yourself	
Other tasks while driving	
Frequency of other tasks performed during driving	
Safety in a vehicle being a passenger	
Alcohol impaired driving	
Disability prevented from driving	
Safety due to Driver assistance systems	
Preference of automation	
Characteristics of Autonomous vehicles	

The data analysis that took place involved 15 statistical analyses having done with two statistical analysis techniques in R Studio. The first technique used was **Ordinal Regression Analysis** which was performed on each dependent variable thrice having all independent variables in the first attempt, then restricting the variables through p-value and t-value tests and the last attempt was made through feature selection method in which only the most statistically significant variables were kept using Random Forest Method. The most suitable model will be the one having a lower AIC value in R Studio. The second technique used was the **Factor Analysis method** which was performed on each output variable twice, once with retaining factors as 4 and one with retaining factors as 5. The factors are retained with the rule of the thumb in which the factors are increased one by one and the process stops when all non-trivial amounts of variance are accounted for *(Brown 2001)* So, the factors that explain large part of the variance are kept in that order *(Rahn 2018)*. Moreover, a Scree plot is performed to note the number of factors retained and it also gives the output graph as four factors to be retained. The predictive variables in this method were chosen based on the questions in the questionnaire having ordinal scales and no demographic questions as they solely depend on a respondent's personal information and vary from person to person and will not co-relate with the other variables.

The main aim of using Ordinal Regression Analysis method for analyzing the data is to calculate a statistically significant model for each outcome variable as to identify the variables that influence (more positively or negatively) the shift from conventional vehicles to autonomous vehicles. The purpose of using Factor analysis as the second method for analysis is to calculate the minimum number of factors needed to produce statistical significant model for each outcome variable. Additionally, it also gives the variables influencing the results.

To target the goal of the research which focuses on the likelihood of people shifting to autonomous vehicles, there were three variables that are considered as dependent variables which include "*mode.shift*", "*Elimination.of.accidents*" and "*Safety.of.AV*" for the people to shift to autonomous vehicles and their perception regarding the elimination of accidents after the introduction of self-driven cars and the safety perspective of autonomous vehicles. All these variables consist of ordinal scales having a scale of 1-5 with 1 as "*very unlikely*" or "*very unsafe*" and 5 as "*very likely*" or "*very safe*".

The final output of the models consists of many trials during which different combinations of the selected variables are tried and tested to get the most significant results. The models are tested and assessed based on statistical tests (t-tests, p-tests, etc...) and on the logical explanation of the results. The selection of the variables is carried out based on their influence on the outcome. Moreover, if the variables strongly statistically co-relate with each other than those were taken into the model.

The aim of the whole process of analyzing the data was the cleaning of variables and testing of variables to get as much statistically significant results as possible which could be presented as the outcome of the model.

4.5 Selected Questions

Following are the few 18 questions which were selected from the total 42 questions in the questionnaire to perform the data analysis. Question Number 1 – 15 are all independent variables for the model and Question Number 16, 17 and 18 are the three questions that are dependent variables.

Q1: Age

Q2: Gender

Q3: Employment status

Q4: Primary mode of transportation

Q5: Frequency of driving a car

Q6: Involvement in any car accident with injury

Q7: Safety while driving the vehicle yourself

Q8: Performing other tasks while driving

Q9: Frequency of performing other tasks while driving

Q10: Safety inside a vehicle being a passenger

Q11: Driving a vehicle under the influence of alcohol

Q12: Disability preventing from manually driving a vehicle

Q13: Safety due to driver assistance systems

Q14: Preference of type of automation

Q15: Characteristics of autonomous vehicles while buying

Q16: Mode shift to autonomous vehicles

Q17: Safety of autonomous cars as compared to human-driven vehicles

Q18: Likelihood of elimination of traffic accidents with autonomous vehicles

4.6 Correlation Matrix

After selecting the independent and dependent variables from the questionnaire it was necessary to see the correlation among all the variables so that the variables which are strongly correlated with each other could be neglected. Therefore, a correlation matrix was created using R Studio to detect the correlated variables as shown below in Figure 4.18.

The visualization of the correlation matrix is a better way of sorting out the variables. It shows that there are no independent variables which are strongly correlated with each other. Hence, all the selected variables were chosen to get output from the model.

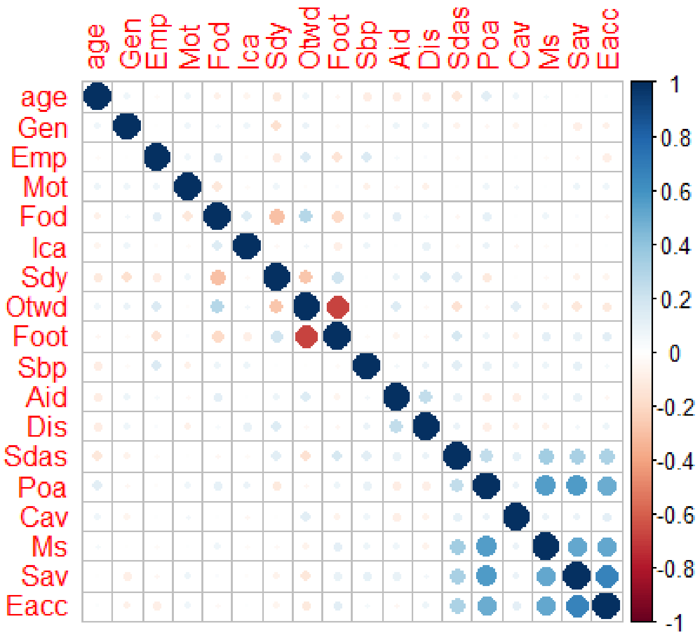

Figure 4.18: Correlation Matrix

4.7 Output of the Models

This section contains the output given by the software R Studio which was performed to analyze the data. There are total 15 model outputs generated with the help of R Studio to compare the results and select the most statistically significant model as the outcome.

Dependent Variables

The dependent variables that have been selected from the questionnaire are as follows:

1) Mode Shift
2) Elimination of Accidents
3) Safety of Autonomous Vehicles

1. Mode shift

The mode shift was calculated through Ordinal Logistic Regression by using the following three techniques:

I. Model 1 (Through all the independent variables selected from the questionnaire)
II. Model 2 (Screening of the variables by p values and t values)
III. Model 3 (Screening of the variables by feature selection method)

I. Model 1 (using all the selected independent variables)

In this technique the ordinal regression was applied on all the selected independent variables from the questionnaire and the effect on the mode shift to autonomous vehicles was taken under consideration for all of them. Results are presented in Table 4.2.

Table 4.2: Mode shift using all selected independent variables

Coefficients:	Value	Std. Error	t value	p value
age. L	-24.514	1.369	-17.906	0.00
age. Q	19.253	1.226	15.704	0.00
age. C	48.450	0.455	106.390	0.00
age^4	37.397	0.798	46.842	0.00
gender. L	1.945	0.894	2.176	0.029
gender. Q	0.770	0.484	1.593	0.111
employment. L	0.695	1.195	0.581	0.561
employment. Q	1.125	1.139	0.988	0.323
employment. C	2.676	1.008	2.655	0.007
employment^4	2.539	1.001	2.514	0.011
employment^5	0.028	0.896	0.032	0.975
employment^6	-1.087	0.762	-1.427	0.154
employment^7	-0.402	0.750	-0.536	0.592
mode.of.transportation. L	0.019	0.984	0.019	0.985
mode.of.transportation. Q	-0.290	0.999	-0.290	0.771
mode.of.transportation. C	-0.920	0.916	-1.003	0.315
mode.of.transportation^4	-1.107	0.780	-1.420	0.155
mode.of.transportation^5	-0.255	0.736	-0.347	0.729
frequency.of.driving	0.268	0.267	1.006	0.314
involvement.in.car.accident. L	-0.364	0.343	-1.059	0.29
safety.in.driving.yourself. L	1.568	0.819	1.916	0.055
safety.in.driving.yourself. Q	-1.183	0.696	-1.700	0.089
safety.in.driving.yourself. C	0.001	0.681	0.001	0.999
safety.in.driving.yourself^4	0.517	0.548	0.942	0.346
other.tasks.while.driving. L	-0.240	0.324	-0.739	0.46
frequency.of.other.tasks. L	0.792	0.756	1.047	0.295
frequency.of.other.tasks. Q	0.090	0.649	0.138	0.890
frequency.of.other.tasks. C	-0.466	0.543	-0.859	0.390
frequency.of.other.tasks^4	-0.479	0.458	-1.045	0.296
vehicle.safety.being.passenger. L	-0.809	0.640	-1.263	0.206
vehicle.safety.being.passenger. Q	0.112	0.581	0.193	0.847
vehicle.safety.being.passenger. C	-0.167	0.517	-0.322	0.747
vehicle.safety.being.passenger^4	-0.612	0.401	-1.528	0.127

alcohol.impaired.driving. L	-0.549	0.885	-0.621	0.535
alcohol.impaired.driving. Q	-0.723	1.381	-0.524	0.600
alcohol.impaired.driving. C	2.092	1.724	1.213	0.225
disability.preventing. L	-0.81	0.704	-1.15	0.249
safety.due.to.driver.assistance. L	1.15	0.810	1.378	0.168
safety.due.to.driver.assistance. Q	-0.16	0.680	-0.24	0.814
safety.due.to.driver.assistance. C	0.740	0.558	1.326	0.185
safety.due.to.driver.assistance^4	-0.550	0.386	-1.425	0.154
preference.of.automation. L	3.608	0.572	6.310	0.00
preference.of.automation. Q	0.105	0.466	0.225	0.822
preference.of.automation. C	-0.826	0.399	-2.074	0.038
preference.of.automation^4	0.216	0.302	0.717	0.474
characteristics.of.AV. L	-0.008	0.569	-0.015	0.988
characteristics.of.AV. Q	1.135	0.704	1.612	0.107
characteristics.of.AV. C	-0.148	0.524	-0.282	0.778
characteristics.of.AV^4	-0.648	0.517	-1.254	0.210
characteristics.of.AV^5	0.152	0.731	0.208	0.835
Intercepts:				
	Value	**Std. Error**	**t value**	**p value**
1= Very unlikely|2= Rather unlikely	11.615	1.145	10.144	0.00
2= Rather unlikely|3= Moderately	12.905	1.125	11.467	0.00
3= Moderately|4= Rather likely	14.771	1.115	13.244	0.00
4= Rather likely|5= Very likely	17.164	1.126	15.240	0.00
Residual Deviance: 466.900				
AIC: 574.900 **BIC: 753.009**				

The interpretation of the output could be described as the effect of each independent variable on the dependent variable with the help of the values of co-efficients, standard error, t-values and p-values.

The independent variables and specific characteristics of respondents are selected based on the metrics of each independent variable and these metrics show that their effect on the dependent variable is statistically significant or not. The co-efficient values of the variables are used to determine how changes in the independent variables are associated with changes in the dependent variable and the statistical significance of these co-efficients is judged by the p-values and t-values *(Frost 2017)*. The p-value is the probability value and it indicates the probability of observing the difference if there is no difference. The conventional and arbitrary threshold for p-value to determine the statistical significance of a variable has been taken as ($p<0.05$) *(Dahiru 2008)*. The values of the independent variables lying in this interval are considered statistically significant for this study.

The second test statistic to figure out the statistical significance of the variables is the t-test statistic. The t test statistic is used to compare the two means or averages to know if they are different from each other and it also tells the statistical significance of the difference between those variables *(Stephanie 2018)*. The arbitrary threshold for the t-value to find out the statistical significance of a variable for this study has been taken as ($t>1.96$). The values of the independent variables lying in this interval have been taken as statistically significant.

From the co-efficient values, it could also be interpreted that how much a specific independent variable category affects the dependent variable. The proportional odds model relates to cumulative probabilities of an ordered categorical response. The analysis suggests that for every category of predictor, the odds of being below the given level of response are multiplied by exponential^(co-efficient value).

Moreover, the sign of the regression co-efficient also explains if there is a positive correlation or negative correlation between the independent variable and the dependent variable. A positive co-efficient illustrates that the mean of the dependent variable tends to increase as the value of the independent variable increases while negative co-efficient shows that the dependent variable decreases with the increase in the independent variable. The co-efficient value demonstrates the change in mean of the dependent variable given a one-unit shift in the independent variable while the other variables in the model remain constant *(Frost 2017).*

The last factor that could be interpreted from the output of the model is the AIC value which is the Akaike's Information Criteria statistic for model selection and the BIC value which is the Bayesian Information Criterion. Both AIC and BIC are based on maximum likelihood estimates of the model parameters. The AIC and BIC values are used to make comparison between models to select the best fitted model.

Considering the above-mentioned parameters, the output of the model 1 shows that the independent variables of "*age", "gender", "employment"* and *"preference of automation"* influence the dependent variable "*mode shift"* more than the other variables because of the strong statistical significance of these variables towards the dependent variable. The most statistical ordinal scales are interpreted for all the above-mentioned variables.

For model 1, with one unit increase in the age group having the largest t-value (>1.96) and lowest p-value (<0.05) , the predicted likelihood of observing the mode shift to autonomous vehicles "very likely" as compared to the mode shift to autonomous vehicles "very unlikely, unlikely, moderately and likely" change by the factor of exp(-48.450)= 9×10^{-22} which is approx.0.99.

The gender also influences the mode shift which can be interpreted in a way that with one unit increase in the gender, the predicted odds of observing "very likely" to shift towards autonomous vehicles versus the odds of "very unlikely, unlikely, moderately and likely" shifting to autonomous vehicles change with a factor of exp(-1.945)= 0.14

The employment status of the respondents also depict that with one unit increase in the employment category with the most significant t-values and p-values, the probability of observing the shift towards autonomous vehicles "very likely" against the probability of observing the shift towards autonomous vehicles "very unlikely, unlikely, moderately and likely" alter with a factor of exp(-2.676) = 0.068 and exp(-2.539)=0.078.

The last variable is the preference of automated driving. With a unit increase in this variable, the predicted odds of observing the switch towards autonomous vehicles "very likely or likely" against observing the shift towards autonomous vehicles "very unlikely, unlikely or moderately" change by a factor of exp(-3.608) = 0.027.

II. Model 2 (screening of the variables through p values and t values)

The second technique used to do the ordinal logistic regression was to screen the variables depending on their t-values and p-values from the first model. The same arbitrary threshold used for the first model are used in this model as well with the variables lying in the interval of ($t>1.96$) and ($p<0.05$) are selected from model 1 and this model 2 is run with those variables. The results are presented in Table 4.3.

Table 4.3: Mode shift using the variables selected through p values and t values

Coefficients:	Value	Std. Error	t value	p value
age. L	-6.686	1.136	-5.884	0.000
age. Q	4.519	1.107	4.080	0.000
age. C	13.063	0.392	33.328	0.000
age^4	10.601	0.822	12.888	0.000
employment. L	2.556	0.908	2.804	0.004
preference.of.automation. L	3.725	0.499	7.460	0.000
preference.of.automation. Q	-0.658	0.335	-1.966	0.030
Intercepts:				
	Value	**Std. Error**	**t value**	**p value**
1= Very unlikely|2= Rather unlikely	1.569	0.608	2.579	0.007
2= Rather unlikely|3= Moderately	2.677	0.587	4.556	0.000
3= Moderately|4= Rather likely	4.291	0.573	7.490	0.000
4= Rather likely|5= Very likely	6.327	0.591	10.697	0.000
Residual Deviance: 516.365				
AIC: 554.365 **BIC: 617.034**				

In this model 2 there are three independent variables. As the variables selected in this model are already the ones that are statistically significant. The effect of the values of co-efficients could be interpreted in terms of which ordinal category has the more influence than the other categories.

For model 2, with one unit increase in the most statistically significant age group determined with t-value and p-value, the predicted likelihood of observing the mode shift towards autonomous vehicles "very likely" against the other ordinal categories of "very unlikely, unlikely, moderate and likely" change with a factor of exp (-13.063) = 2.1×10^{-6}.

The second variable is the employment status. With one unit increase in the employment status, the likelihood of observing the mode shift towards autonomous vehicles "very likely or likely" against the mode shift "very unlikely, unlikely or moderately" changes with a factor of exp(-2.556) = 0.077.

The last variable is the preference of automation. The model output shows that with one unit increase in the preference of automation, the probability of the switch to autonomous vehicles "likely or very likely" against the switch to autonomous vehicles "unlikely, very unlikely or moderately" changes with a factor of exp(-3.725) = 0.024.

III. Model 3 (screening of the variables through feature selection method)

The third technique used to obtain results through ordinal regression method was by using a feature selection method. The feature selection method used for this model is random forest method. In this method the variables are selected based on their high values and importance towards dependent variable. The results are presented in table 4.4.

Table 4.4: Mode shift using feature selection method

Coefficients:	**Value**	**Std. Error**	**t value**	**p value**
employment. L	1.967	0.922	2.135	0.032
employment. Q	2.006	0.915	2.194	0.028
preference.of.automation. L	3.432	0.526	6.530	0.000
Intercepts:				
	Value	**Std. Error**	**t value**	**p value**
1= Very unlikely|2= Rather unlikely	-1.748	0.506	-3.454	0.000
2= Rather unlikely|3= Moderately	-0.577	0.485	-1.189	0.234
3= Moderately|4= Rather likely	1.125	0.492	2.288	0.022
4= Rather likely|5= Very likely	3.362	0.535	6.282	0.000
Residual Deviance: 494.232				
AIC: 558.232 **BIC: 645.903**				

The statistically significant independent variables for this model 3 following the designated parameters explained under output of model 1 are "*employment*" and *"preference of automation".*

From table 4.4, the first variable is the employment status. With the one-unit increment in the employment status, the probability of observing the mode shift towards autonomous vehicles "likely or very likely" versus observing the mode shift towards autonomous vehicles "unlikely, very unlikely or moderately" alters with a factor of exp(-2.006)= 0.13.

The next variable is the preference of automation. With a one unit increase in the preference of automation, the predicted odds of observing the mode shift towards autonomous vehicles "very likely or likely" against observing the mode shift towards autonomous vehicles "unlikely, very unlikely or moderately" change with a factor of exp(-3.432) = 0.032.

Only two ordinal categories of the variables *"employment" and "preference of automation"* are considered statistically significant as they lie in the interval decided to interpret the model.

2. Elimination of Accidents

Elimination of accidents with autonomous vehicles is also calculated through ordinal logistic regression with three methods.

I. Model 4 (using all the important independent variables selected from the questionnaire)
II. Model 5 (screening of the variables by p values and t values)
III. Model 6 (screening of the variables by feature selection method)

I. Model 4 (using all the selected independent variables)

In this technique the ordinal regression was applied on all the selected independent variables from the questionnaire and the effect on the elimination of accidents by autonomous vehicles was taken under consideration for all of them. The results are presented in table 4.5 below.

Table 4.5: Elimination of accidents by AV using all the selected independent variables

Coefficients:	Value	Std. Error	t value	p value
age. L	-14.701	1.476	-9.959	0.000
age. Q	11.509	1.380	8.338	0.000
age. C	27.928	0.528	52.866	0.000
age^4	19.911	0.889	22.388	0.000
gender. L	-0.089	0.773	-0.115	0.908
gender. Q	0.276	0.456	0.605	0.545
employment. L	8.311	0.696	11.941	0.000
employment. Q	10.698	0.486	22.033	0.000
employment. C	11.208	0.672	16.667	0.000
employment^4	6.290	0.854	7.366	0.000
employment^5	3.326	0.927	3.590	0.000
employment^6	0.017	0.836	0.020	0.984
employment^7	1.627	0.824	1.975	0.048
mode.of.transportation. L	-0.307	1.126	-0.273	0.785
mode.of.transportation. Q	-0.753	1.083	-0.695	0.487
mode.of.transportation. C	0.243	1.080	0.225	0.822
mode.of.transportation^4	0.413	0.970	0.425	0.671
mode.of.transportation^5	0.253	0.785	0.322	0.748
frequency.of.driving. L	-0.570	0.652	-0.873	0.382
frequency.of.driving. Q	-0.764	0.441	-1.735	0.082
frequency.of.driving. C	-0.545	0.314	-1.734	0.082
involvement.in.car.accident. L	-0.086	0.353	-0.242	0.809
safety.in.driving.yourself. L	-0.991	0.754	-1.315	0.188
safety.in.driving.yourself. Q	-0.503	0.628	-0.800	0.424
safety.in.driving.yourself. C	-0.336	0.665	-0.506	0.613
safety.in.driving.yourself^4	-0.104	0.547	-0.190	0.849
other.tasks.while.driving. L	0.186	0.310	0.599	0.549
frequency.of.other.tasks. L	-0.269	0.766	-0.351	0.725
frequency.of.other.tasks. Q	-1.350	0.649	-2.081	0.037

frequency.of.other.tasks. C	-0.854	0.562	-1.520	0.128
frequency.of.other.tasks^4	0.084	0.470	0.179	0.858
vehicle.safety.being.passenger. L	-0.053	0.657	-0.081	0.935
vehicle.safety.being.passenger. Q	-0.084	0.583	-0.145	0.885
vehicle.safety.being.passenger. C	0.657	0.508	1.295	0.195
vehicle.safety.being.passenger^4	-0.322	0.395	-0.815	0.415
alcohol.impaired.driving. L	-2.635	1.005	-2.623	0.008
alcohol.impaired.driving. Q	-0.541	1.381	-0.391	0.695
alcohol.impaired.driving. C	2.296	1.671	1.374	0.169
disability.preventing.from.driving. L	0.833	0.673	1.237	0.216
safety.due.to.driver.assistance. L	1.601	0.621	2.58	0.009
safety.due.to.driver.assistance. Q	0.118	0.524	0.226	0.821
safety.due.to.driver.assistance. C	0.128	0.500	0.257	0.797
safety.due.to.driver.assistance^4	-0.112	0.403	-0.277	0.782
preference.of.automation. L	2.850	0.520	5.481	0.000
preference.of.automation. Q	1.174	0.447	2.626	0.008
preference.of.automation. C	0.067	0.384	0.174	0.862
preference.of.automation^4	0.212	0.314	0.674	0.5
characteristics.of.AV. L	0.307	0.662	0.464	0.643
characteristics.of.AV. Q	-0.898	0.828	-1.085	0.278
characteristics.of.AV. C	1.012	0.603	1.678	0.093
characteristics.of.AV^4	1.207	0.580	2.080	0.037
characteristics.of.AV^5	-2.220	0.862	-2.575	0.01
Intercepts:				
	Value	**Std. Error**	**t value**	**p value**
1= Very unlikely|2= Rather unlikely	0.879	1.188	0.740	0.460
2= Rather unlikely|3= Moderately	3.476	1.119	3.108	0.001
3= Moderately|4= Rather likely	5.391	1.107	4.870	0.000
4= Rather likely|5= Very likely	8.371	1.119	7.478	0.000
Residual Deviance: 422.796				
AIC: 534.796 **BIC: 719.502**				

In the output of model 4, the statistically significant variables are selected on the same principles as decided in the mode shift model. The change here is the dependent variable which is the elimination of accidents with the introduction of autonomous vehicles.

The statistically significant variables in the model output are *"age"*, *"employment"*, *"frequency of other tasks while driving"*, *"alcohol impaired driving"*, *"safety due to driver assistance systems"*, *"preference of automation"* and *"characteristics of autonomous vehicles"*.

For the age group, with one unit increase in the age group, the predicted odds of observing the elimination of accidents “likely or very likely” with the introduction of self-driven cars against observing the elimination of accidents “unlikely, very unlikely or moderately” with self-driven cars changes with a factor of exp(-27.928)= 7.4×10^{-13} .

For the employment status, the exp(-10.698)= 2.25×10^{-5} which shows that with one unit increase in the most statistically significant (regards to p-value and t-value) employment status, the likelihood of observing the elimination of accidents “likely or very likely” with the dawn of self-driven cars against observing the elimination of accidents “unlikely, very unlikely or moderately” change with a factor of 2.25×10^{-5}.

The next independent statistically significant variable is alcohol impaired driving. The respondents that drive the car under the influence of alcohol have an exp(-2.635) = 0.07 depicting that with one unit increase in the alcohol impaired drivers, predicted odds of observing the elimination of accidents “very likely or likely” after the introduction of self-driven cars versus observing the elimination of accidents “very unlikely, unlikely or moderately” alter with a factor of 0.07.

The variable of safety with driver assistance systems has an exp(-1.601) = 0.20 exhibiting that with one unit increase in this variable, the probability of observing the elimination of accidents “very likely or likely” with self-driven cars against observing the elimination of accidents “very unlikely, unlikely or moderately” with self-driven cars changes with a factor of 0.20.

For the preference of automation the exp(-2.850) = 0.05 shows that with one unit increase in the variable, the predicted odds of observing the elimination of accidents “likely or very likely” with autonomous vehicles versus observing the elimination of accidents with autonomous vehicles “unlikely or very unlikely or moderately” alter with a factor of 0.05.

The last statistically significant independent variable is the characteristics of autonomous vehicles. The people who will consider the characteristics of the autonomous vehicles while buying have an exponential(-2.220) = 0.10 indicating that with one unit increase in the variable, the likelihood of observing the elimination of accidents “likely or very likely” versus observing the elimination of accidents “unlikely, very unlikely or moderately” change with a factor of 0.10.

II. Model 5 (Screening of the variables through p values and t values)

The second technique used to do the ordinal logistic regression was to screen variables depending on their t-values and p-values from the first model. The same arbitrary threshold used for the first model is used in this model as well with the variables lying in the interval of ($t>1.96$) and ($p<0.05$) are selected from model 4 and the model 5 is run with those variables. Results of the model 5 are presented in table 4.6.

Table 4.6: Elimination of accidents by AV by using variables selected through p-values and t-values

Co-efficients:	Value	Std. Error	t value	p value
age. L	-13.200	1.257	-10.504	0.000
age. Q	10.424	1.144	9.115	0.000
age. C	25.362	0.427	59.363	0.000
age^4	18.172	0.749	24.272	0.000
employment. L	7.329	0.629	11.635	0.000

employment. Q	9.429	0.431	21.874	0.000
employment. C	9.782	0.600	16.309	0.000
employment^4	5.748	0.765	7.511	0.000
employment^5	2.461	0.817	3.011	0.002
alcohol.impaired.driving. L	-1.864	0.803	-2.320	0.020
safety.due.to.driver.assistance. L	1.277	0.566	2.256	0.024
preference.of. automation. L	2.607	0.456	5.715	0.000
preference.of.automation. Q	0.809	0.388	2.086	0.037
characteristics.of.AV. L	-1.761	0.771	-2.284	0.022
Intercepts:	**Value**	**Std.Error**	**t value**	**p value**
1= Very unlikely|2= Rather unlikely	1.359	0.841	1.617	0.106
2= Rather unlikely|3= Moderately	3.769	0.765	4.929	0.000
3= Moderately|4= Rather likely	5.513	0.755	7.297	0.000
4= Rather likely|5= Very likely	8.217	0.774	10.616	0.000
Residual Deviance: 452.0437				
AIC: 514.0437 BIC: 616.291				

The most statistically significant variables for model 5 are the *age group*, *employment status of the respondents*, *driving under the influence of alcohol*, *considering the safety due to driver assistance system*, *preference of automation* and the respondents considering *the characteristics of the autonomous cars* while buying in future.

The co-efficient value shows that the most statistically significant age group (with respect to p-value and t-values) have an exp(-25.362) = 9.6×10^{-12} which means that with one unit increase in the age group, the predicted odds of observing the elimination of accidents “very likely or likely” with autonomous cars against observing the elimination of accidents “unlikely or very unlikely or moderately” change with a factor of 9.6×10^{-12}.

The employment status has an exp(-9.429) = 8.03×10^{-5} showing that with one unit increase in the employment status , the likelihood of observing the elimination of accidents “very likely or likely” against observing the elimination of accidents “unlikely, very unlikely or moderately” alters with a factor of 8.03×10^{-5}.

The next statistically significant variable is the alcohol impaired driving. The people who drive under the influence of alcohol have an exp(-1.864) = 0.15 depicting that with one-unit increment in the variable, the predicted odds of observing elimination of accidents “very likely or likely” with autonomous cars against observing elimination of accidents “unlikely, very unlikely or moderately” change with a factor of 0.15.

The safety due to driver assistance systems has an exp(-1277) = 0.28 exhibiting that with one-unit increment, the likelihood of observing elimination of accidents “very likely or likely” after the introduction of autonomous cars against observing elimination of accidents “very unlikely or unlikely” change with a factor of 0.28.

For the preference of automation the exp(-2.607) = 0.07 which shows that with one unit increase in the preference of automation, the probability of observing elimination of accidents "very likely or likely" with self-driven cars versus observing the elimination of accidents "very unlikely, unlikely or moderately" change with a factor of 0.07.

The last statistically significant independent variable is the characteristics of autonomous vehicles. The people who will consider characteristics of the autonomous vehicles while buying have an exp(-1.761) = 0.17 showing that with one-unit increase in the people considering the characteristics while buying autonomous vehicles, the predicted odds of observing the elimination of accidents "Very likely or likely" with autonomous cars against observing the elimination of accidents "Very unlikely or unlikely or moderately" change with a factor of 0.17.

III. Model 6 (Screening of the variables through feature selection method)

The third technique used to obtain results through ordinal regression method was by using a feature selection method. The feature selection method used for this model is random forest method. In this method the variables are selected based on their high values and importance towards dependent variable. The results for model 6 are presented in table 4.7.

Table 4.7: Elimination of accidents with AV by using feature selection method

Co-efficients:	Value	Std. Error	t value	p value
safety.due.to.driver.assistance. L	1.253	0.505	2.481	0.013
preference.of.automation. Q	2.411	0.416	5.791	0.000
Intercepts:	**Value**	**Std.Error**	**t value**	**p value**
1= Very unlikely|2= Rather unlikely	-3.446	0.411	-8.380	0.000
2= Rather unlikely|3= Moderately	-1.334	0.232	-5.741	0.000
3= Moderately|4= Rather likely	0.170	0.219	0.778	0.437
4= Rather likely|5= Very likely	2.551	0.279	9.148	0.000
Residual Deviance: 500.734				
AIC: 524.734 **BIC: 564.314**				

As shown in the output of the model 6, there are only two statistically significant independent variables which are "*safety due to driver assistance systems"* and "*preference of automation".*

The co-efficient values decide about how much they impact the elimination of accidents after the introduction of autonomous vehicles. The safety due to driver assistance system have an exp(-1.253) = 0.28 depicting that with one unit increase in the variable, the predicted odds of observing the elimination of accidents "very likely or likely" with autonomous cars against observing the elimination of accidents "very unlikely, unlikely or moderately" alter with a factor of 0.28.

The second independent variable is the preference of automation. The preference of automation has an exp(-2.411) = 0.08 which shows that with one unit increase in the variable, the probability of observing the elimination of accidents "very likely or likely" with autonomous cars versus observing the elimination of accidents "very unlikely, unlikely or moderately" alter with a factor of 0.08.

3. Safety of Autonomous Vehicles

Safety of autonomous vehicles is also calculated through ordinal logistic regression with three methods.

I. Model 7 (using all the independent variables selected from the questionnaire)
II. Model 8 (screening of the variables by p values and t values)
III. Model 9 (screening of the variables by feature selection method)

I. Model 7 (using all the selected independent variables)

Using this technique, the ordinal regression was applied on all the selected independent variables from the questionnaire and the likelihood of autonomous cars being safer than the human driven vehicles was calculated. The results are presented in table 4.8.

Table 4.8: Safety of autonomous vehicles by using all the selected independent variables

Co-efficients:	Value	Std. Error	t value	p value
age. L	-25.945	1.487	-17.449	0.000
age. Q	19.848	1.340	14.807	0.000
age. C	49.080	0.490	100.117	0.000
age^4	37.245	0.872	42.699	0.000
gender. L	-0.927	0.813	-1.140	0.254
gender. Q	-0.190	0.467	-0.406	0.685
employment. L	8.490	0.737	11.513	0.000
employment. Q	10.117	0.476	21.228	0.000
employment. C	9.507	0.645	14.747	0.000
employment^4	6.110	0.885	6.901	0.000
employment^5	2.092	0.974	2.147	0.031
employment^6	1.047	0.822	1.274	0.203
employment^7	1.912	0.790	2.418	0.015
mode.of.transportation. L	-0.777	1.199	-0.647	0.517
mode.of.transportation. Q	-1.523	1.195	-1.274	0.202
mode.of.transportation. C	-2.389	1.085	-2.200	0.027
mode.of.transportation^4	-0.183	0.897	-0.204	0.838
mode.of.transportation^5	-0.159	0.791	-0.201	0.84
frequency.of.driving. L	-0.599	0.274	-2.182	0.029
involvement.in.car.accident. L	-0.545	0.380	-1.431	0.152
safety.in.driving.yourself. L	-2.386	0.795	-2.999	0.002
safety.in.driving.yourself. Q	0.730	0.676	1.079	0.28
safety.in.driving.yourself. C	-0.347	0.690	-0.503	0.615
safety.in.driving.yourself^4	-0.717	0.562	-1.275	0.202
other.tasks.while.driving. L	-0.721	0.318	-2.262	0.023
frequency.of.other.tasks. L	-0.651	0.777	-0.837	0.402
frequency.of.other.tasks. Q	-0.493	0.653	-0.754	0.451
frequency.of.other.tasks. C	-1.289	0.560	-2.300	0.021
frequency.of.other.tasks^4	-0.300	0.475	-0.632	0.527

vehicle.safety.being.passenger. L	0.817	0.667	1.224	0.221
vehicle.safety.being.passenger. Q	-0.227	0.593	-0.382	0.702
vehicle.safety.being.passenger. C	0.104	0.526	0.199	0.842
vehicle.safety.being.passenger^4	0.179	0.415	0.432	0.666
alcohol.impaired.driving. L	0.558	0.893	0.624	0.532
alcohol.impaired.driving. Q	-1.912	1.421	-1.345	0.178
alcohol.impaired.driving. C	3.072	1.764	1.741	0.081
disability.preventing.from.driving. L	0.710	0.701	1.012	0.311
safety.due.to.driver.assistance. L	2.574	0.745	3.452	0.000
safety.due.to.driver.assistance. Q	-0.876	0.602	-1.454	0.146
safety.due.to.driver.assistance. C	0.435	0.528	0.824	0.41
safety.due.to.driver.assistance^4	0.261	0.406	0.642	0.52
preference.of.automation. L	4.039	0.596	6.768	0.000
preference.of.automation. Q	1.040	0.463	2.242	0.024
preference.of.automation. C	-0.059	0.399	-0.149	0.881
preference.of.automation^4	-0.256	0.321	-0.799	0.424
characteristics.of.AV. L	1.039	0.787	1.320	0.187
characteristics.of.AV. Q	0.933	0.886	1.053	0.292
characteristics.of.AV. C	0.676	0.672	1.004	0.315
characteristics.of.AV^4	0.364	0.588	0.618	0.536
characteristics.of.AV^5	-0.785	0.853	-0.919	0.358
Intercepts:	**Value**	**Std.Error**	**t value**	**p value**
1= Very unsafe|2= Rather unsafe	8.337	1.221	6.826	0.000
2= Rather unsafe|3= Moderate	10.968	1.150	9.538	0.000
3= Moderate|4= Rather safe	14.261	1.152	12.374	0.000
4= Rather safe|5= Very safe	17.301	1.185	14.592	0.000
Residual Deviance: 373.4322				
AIC: 481.4322 **BIC : 659.541**				

As can be seen in table 4.8, the statistically significant variables for model 7 are “*age*”, “*employment*”, “*mode of transportation*”, “*frequency of driving*”, “*safety in driving yourself*”, “*performing other tasks while driving*”, “*frequency of other tasks while driving*”, “*safety due to driver assistance systems*” and “*preference of automation*”.

The age group has an exp(-49.08) = $4.8x10^{-22}$ which shows that with one unit increase in the age group, the predicted odds of observing the autonomous vehicles being “very safe or safe” than the human driven cars against observing the autonomous vehicles being “very unsafe, unsafe or moderately safe” than the human driven cars change by a factor of $4.8x10^{-22}$.

The model shows that with one unit increase in the most statistically significant (with regards to p-value and t-values) employment status, the probability of observing the autonomous cars being "very safe or safe" than the current human driven cars against observing the autonomous cars being "unsafe, very unsafe or moderate" than the normal vehicles change with a factor of exp(-10.11)= 4×10^{-5}.

The mode of transportation has an exp(-2.38) = 0.092 showing that with the increment of one-unit in the mode of frequently used transportation, the likelihood of observing the autonomous cars being "very safe or safe" than the human driven cars against observing the autonomous cars being "very unsafe, unsafe or moderate" than the normal cars alter with a factor of 0.092.

Furthermore, the model 7 output shows that the frequency of driving the car has an exp(-0.599) = 0.54 which means that with one-unit increase in the variable, the predicted odds of observing the autonomous vehicles being "very safe or safe" versus observing the autonomous vehicles being "very unsafe, unsafe or moderate" than the normal vehicles alter with a factor of 0.54.

Likewise, the feeling of safety while driving the car themselves has an exp(-2.38) = 0.092 exhibiting that with one-unit increment in the variable, the predicted odds of observing the autonomous cars being "very safe or safe" than the human driven vehicles against observing the autonomous cars being "very unsafe, unsafe or moderately safe" than the human driven cars change with a factor of 0.092.

The people who perform other tasks while driving have an exp(-0.721) = 0.486 indicating that with one-unit increase in the persons who perform other tasks while driving, the probability of observing the autonomous vehicles being "very safe or safe" versus observing the autonomous vehicles being "very unsafe, unsafe or moderately safe" alter with a factor of 0.486.

With regards to the frequency of other tasks while driving, the output of the model shows that with one unit increase in the people who perform other tasks while driving, the predicted odds of observing the autonomous vehicles being "very safe" than the human driven vehicles against the autonomous vehicles being "very unsafe, unsafe or moderately safe "change by a factor of exp(-0.493) = 0.610.

Observing the safety due to driver assistance system, according to the model output, if there is one-unit increment in this variable, the likelihood of observing the self-driven cars being "very safe or safe" than normal vehicles against observing the self-driven cars being "very unsafe, unsafe or moderately safe" alter with a factor of exp(-2.57) = 0.076.

Finally, it is shown in the preference of automation that with a one-unit increment in the variable, the probability of observing the autonomous cars being "very safe or safe" against observing the self-driven cars being "very unsafe, unsafe or moderately safe" change with a factor of exp(-4.03) = 0.017.

II. Model 8 (Screening of the variables through p values and t values)

The second technique used to do the ordinal logistic regression was to screen the variables depending on their t-values and p-values from the model 7. The same arbitrary threshold used for the first model is used in this model as well with the variables lying in the interval of ($t>1.96$) and ($p<0.05$) are selected from model 7 and the model 8 is run with those variables. The results of the output of model 8 are shown in table 4.9 below:

Table 4.9: Safety of autonomous vehicles by using the variables selected through p-values and t-values

Co-efficients:	Value	Std. Error	t value	p value
age. L	-19.090	1.417	-13.472	0.000
age. Q	15.702	1.305	12.028	0.000
age. C	38.220	0.479	79.695	0.000
age^4	28.939	0.871	33.215	0.000
employment. L	8.564	0.697	12.273	0.000
employment. Q	9.919	0.450	22.018	0.000
employment. C	9.304	0.612	15.184	0.000
employment^4	5.754	0.859	6.695	0.000
employment^5	1.947	0.913	2.131	0.030
employment^6	2.122	0.777	2.730	0.006
safety.in.driving.yourself. L	-2.138	0.705	-3.029	0.002
other.tasks.while.driving. L	-0.631	0.300	-2.100	0.035
frequency.of.other.tasks. L	-1.193	0.523	-2.281	0.022
safety.due.to.driver.assistance. L	2.164	0.677	3.194	0.001
preference.of.automation. L	3.807	0.552	6.891	0.000
preference.of.automation. Q	1.047	0.428	2.445	0.014
Intercepts:	**Value**	**Std.Error**	**t value**	**p value**
1= Very unsafe|2= Rather unsafe	5.685	0.878	6.471	0.000
2= Rather unsafe|3= Moderate	8.173	0.814	10.029	0.000
3= Moderate|4= Rather safe	11.365	0.817	13.910	0.000
4= Rather safe|5= Very safe	14.169	0.868	16.317	0.000
Residual Deviance: 389.1705				
AIC: 469.171				
BIC: 601.103				

The statistically significant independent variables in this model are "*age*", "*employment*", "*safety in driving yourself*", "*other tasks while driving*", "*frequency of other tasks while driving*", *safety due to driver assistance systems*" and "*preference of automation*".

To interpret the co-efficient of the age of people, with one unit increase in the age range, the predicted odds of observing the autonomous vehicles being "very safe or safe" than normal cars against observing the autonomous vehicles being "unsafe, very unsafe or moderately safe" than human driven cars changes with a factor of exp(-38.22) = 2.51×10^{-17}.

The model predicts that with one-unit increment in the employment status of people, the likelihood of observing the autonomous vehicles being "very safe or safe" than human driven vehicles against observing the autonomous vehicles being "very unsafe, unsafe or moderately safe" than human driven cars alter with a factor of exp(-9.91) = 4.96×10^{-5}.

Furthermore, with one-unit increment in the feeling of safety while driving the car themselves, the probability of observing the autonomous vehicles being "very safe or safe" than human driven vehicles against observing the autonomous vehicles being "very unsafe, unsafe or moderately safe" than human driven cars change with a factor of exp(-2.13) = 0.11.

While observing the respondents who perform other tasks while driving. It is seen that if there is a unit increase in the people that perform other tasks while driving, the likelihood of observing the autonomous cars being "very safe or safe" than human driven cars against observing the autonomous cars being "very unsafe, unsafe, moderate" than human driven cars alter with a factor of exp(-0.63) = 0.53.

Likewise, with a one-unit increase in the frequency of performing other tasks while driving, the predicted odds of observing the autonomous cars being "very safe or safe" than the normal cars against observing the autonomous cars being "unsafe, very unsafe or moderate" change with a factor of exp(-1.19) = 0.304.

With respect to the safety with driver assistance system, according to the model output, if there is one-unit increment in this variable, the likelihood of observing the self-driven cars being "very safe or safe" than normal vehicles against observing the self-driven cars being "very unsafe, unsafe or moderately safe" alter with a factor of exp(-2.16) = 0.11.

In the last, with one-unit increase in the preference of automation, the likelihood of observing the autonomous cars being "very safe or safe" versus observing the autonomous cars being "very unsafe, unsafe or moderate" changes with a factor of exp(-3.807) = 0.022.

III. Model 9 (Screening of the variables through feature selection method)

The third technique used to obtain results through ordinal regression method was by using a feature selection method. The feature selection method used for this model is random forest method. In this method the variables are selected based on their high values and importance towards dependent variable. The results of this model are displayed in table 4.10 below:

Table 4.10: Safety of autonomous vehicles using the variables selected through feature selection method

Co-efficients:	Value	Std. Error	t value	p value
safety.due.to.driver.assistance. L	1.833	0.610	3.003	0.002
preference.of.automation. L	3.041	0.460	6.602	0.000
preference.of.automation. Q	0.982	0.365	2.692	0.007
Intercepts:	**Value**	**Std.Error**	**t value**	**p value**
1= Very unsafe|2= Rather unsafe	-3.593	0.421	-8.534	0.000
2= Rather unsafe|3= Moderate	-1.456	0.268	-5.414	0.000
3= Moderate|4= Rather safe	1.175	0.264	4.445	0.000
4= Rather safe|5= Very safe	3.463	0.338	10.242	0.000
Residual Deviance: 451.622				
AIC: 475.622 **BIC: 515.202**				

The statistically significant variables for model 10 are *"safety due to driver assistance system"* and *"preference of automation".* The interpretation of these variables could be done in a similar style as the previous models. The first variable could be interpreted in a way that if there is a one unit increase in the safety of driver assistance systems, the predicted odds of observing the autonomous cars being "very safe or safe" than human driven vehicles against observing the autonomous cars being "very unsafe, unsafe or moderate" change with a factor of exp(-1.83) = 0.16.

The last variable is the preference of automation. With the increment of one-unit in the preference of automation, the probability of observing the autonomous cars being "very unsafe or safe" than human driven vehicles against observing the autonomous cars being "very unsafe, unsafe or moderate" change with a factor of exp(-3.04) = 0.047.

4.8 Comparison of models

The comparison of models is based on the AIC and BIC values for all the models of a single dependent variable. There is a stronger evidence of choosing one model over the other if there is a large difference in either AIC or BIC values. The lower the values, the better is the model. The difference between the AIC and BIC is that AIC doesn't penalize the number of parameters as strongly as BIC does. As the sample size N approaches to infinity, the model with the least AIC value will be the model closest to the true model because of the smallest Kullback-Liebler divergence *(dmartin 2014).*

Table 4.11: Comparison of models

Dependent Variable	Method	AIC Value
Mode shift	Using all selected variables	574.900
	Screening of variables using p and t values	554.365
	Screening of variables using feature selection method	558.238
Elimination of accidents	Using all selected variables	534.796
	Screening of variables using p and t values	514.044
	Screening of variables using feature selection method	524.734
Safety of autonomous vehicles	Using all selected variables	481.432
	Screening of variables using p and t values	469.171
	Screening of variables using feature selection method	475.622

The AIC values of the model shows that the model made up with the screening of the variables using the p-values and t-values is closest to the true model in case of all three dependent variables. The next model which comes closest to the true model is the model made up with the variables selected with the feature selection method in which random tree method was used. The least preferred model for this study would be the model made up with all selected variables.

4.9 Factor Analysis

The second analysis method used for this study is the factor analysis method. In this method, the ordinal variables selected from the questionnaire are quantified into their underlying interval versions under the aim of maximizing the variance which is demonstrated by the selected number of principal components extracted from the interval data. This is called categorical principal component analysis or non-linear factor analysis *(ttnphns 2017)*. There is a limit on the number of factors that can be specified while performing factor analysis. If the data set has v variables and f factors, then $(v-f)^2$ must be greater than (v+f). For this study, the number of variables is 9, the number of factors that can be specified is 4 so $(9-4)^2>(9+4)$. If the number of factors is 5 then it is also feasible as $(9-5)^2>(9+5)$ *(McCaffrey 2017)*. The retention of factors could be based upon a number of rules: i) It could be based upon the eigen values being greater than 1 called as Kaiser Criterion. ii) Another method of finding how many factors to be retained is Cattell's Scree test which investigates the plot of the eigen values for breaks or discontinuities. iii) The third factor retention method is Parallel Analysis *(Hayton et al. 2004)*. iv) The amount of factors could also be retained with the help of thumb rule in which analysis occurs till all non-trivial variance is accounted *(Brown 2001).* For this study, two methods are used for factors retention. One is the thumb rule of calculating factors till all non-trivial variance is explained. Secondly, the eigen values are also plotted versus the components to plot the Scree graph. The results are the same as both methods give four or five factors to retain.

1) Mode Shift

The first dependent variable is the mode shift to autonomous vehicles. For this variable, the factor analysis was performed twice. The number of factors used is directly proportional to the number of variables to be analyzed. For this study, the factor analysis was performed with four factors and with five factors.

I. Model 10 (using factor number 4)
II. Model 11 (using factor number 5)

I. Model 10 (Using factor number 4)

For this model, the factor number was kept at four to get the required factors against the loadings. The results of factor analysis with uniquenesses and loadings are shown in table 4.12 and 4.13 below:

Table 4.12: Mode shift using factor analysis with 4 factors

Uniquenesses:	
mode.of.transportation	frequency.of.driving
0.953	0.52
safety.in.driving.yourself	vehicle.safety.being.passenger
0.776	0.96
alcohol.impaired.driving	safety.due.to.driver.assistance.systems
0.005	0.328
preference.of.automation	characteristics.of.autonomous.vehicles
0.01	0.972
mode.shift	
0.617	

Table 4.13: Loadings of model 10

Loadings:	Factor1	Factor2	Factor3	Factor4
mode.of.transportation				
frequency.of.driving				0.683
safety.in.driving.yourself				-0.44
vehicle.safety.being.passenger				
alcohol.impaired.driving		0.993		
safety.due.to.driver.assistance.systems			0.801	
preference.of.automation	0.972			
characteristics.of.autonomous.vehicles				
mode.shift	0.506		0.354	
	Factor1	Factor2	Factor3	Factor4
SS loadings	1.246	1.025	0.863	0.724
Proportion Var	0.138	0.114	0.096	0.08
Cumulative Var	0.138	0.252	0.348	0.429

The first part of the output of factor analysis produces "Uniquenesses", which ranges from 0 to 1. A high uniqueness value interprets that it doesn't fit neatly into the factors. If uniqueness of a variable is subtracted from 1, it gives a quantity called "communality". The communality is the proportion of variance of the *i*th variable contributed by the m factors *(Ford 2016).*

The ''Loadings" range from -1 to 1. A value is high if it is near to 1. There is no value for certain variables in the "Factor" column which means that their loadings were less than 0.1 or very close to 0 and R doesn't print those loadings. The loadings less than 0.3 are ignored. Loadings are the correlations with the unobserved factors *(Ford 2016).*

Variances are technically associated with the loadings and considered a part of them. These values summarize the factors and the usual more interested part of them is the "Cumulative Var" which ranges from 0 to 1 and tells the total proportion of variance. "Proportion Var" is the term which shows the proportion of variance explained by each factor. The "SS loadings" row depicts the sum of squared loadings and is used to find out the value of a factor. A factor is to be kept if its "SS loading" is greater than 1 *(Ford 2016).* The p-values at the end of the output is the probability that the source data perfectly fits the number of factors specified, so larger values are better *(McCaffrey 2017).*

If model 10 is interpreted according to the above-mentioned principles, "*the preference of automation"* has high large positive value of 0.972 with "Factor1" which means that this independent variable is strongly correlated with Factor1, "*alcohol impaired driving"* has high positive large value of 0.993 for "Factor2" , "*safety due to driver assistance systems"* is strongly correlated with "Factor3" with a value of 0.801 while "*frequency of driving"* has high positive value with "Factor4" with a factor of 0.683.

The SS loadings for model 10 show that the values for the first two factors "Factor1" and "Factor2" are greater than 1, while"Factor3" and "Factor 4" have a proportion variance of 9.6% and 8% respectively so all these factors are significant and are kept.

II. Model 11 (Using factor number 5)

For this model 11, the factor number was changed to five to see the change in the loadings. The results are presented in table 4.14 and 4.15 respectively below:

Table 4.14: Factor analysis of mode shift with 5 factors

Uniquenesses:	
mode.of.transportation	frequency.of.driving
0.005	0.712
safety.in.driving.yourself	vehicle.safety.being.passenger
0.623	0.867
alcohol.impaired.driving	safety.due.to.driver.assistance.systems
0.005	0.697
preference.of.automation	characteristics.of.autonomous.vehicles
0.548	0.946
mode.shift	
0.226	

Table 4.15: Loadings with factor analysis of model 11

Loadings:					
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
mode.of.transportation			**0.991**		
frequency.of.driving				-0.516	
safety.in.driving.yourself				**0.604**	
vehicle.safety.being.passenger					0.349
alcohol.impaired.driving		**0.997**			
safety.due.to.driver.assistance.system	0.354				0.36
preference.of.automation	**0.616**				
characteristics.of.autonomous.vehicles					
mode.shift	**0.873**				
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
SS loadings	**1.279**	**1.045**	**0.999**	0.706	0.34
Proportion Var	**0.142**	**0.116**	**0.111**	0.078	0.038
Cumulative Var	0.142	0.258	0.369	0.448	0.486

In this model, "*mode shift*" and "*preference of automation*" are highly correlated with "Factor1" with a positive value of 0.873 and 0.616 respectively while "*alcohol impaired driving*" has a high positive value of 0.997 for "Factor2". The "*mode of transportation*" is strongly correlated with "Factor3" and has a factor of 0.991. "*Safety in driving yourself*" has a correlation with "Factor4" with a value of 0.604 while there is no correlation for "Factor5".

The SS loadings for the model 11 show that Factor 1 and Factor 2 are significant. Factor 3 could also be considered significant with a value just near to 1. Factor 4 and Factor 5 have both the values of SS loadings below 1 as well as their proportion variance is 7.8% and 3.8% respectively so they are considered insignificant.

2) Elimination of Accidents

Elimination of accidents is also analyzed using factor analysis. It was performed twice likewise for this variable as well.

I. Model 12 (using factor number 4)
II. Model 13 (Using factor number 5)

I. Model 12 (using factor number 4)

In this model 12, the factor analysis was performed by using a factor number 4. The results of the model are displayed in the table 4.16 and 4.17 below:

Table 4.16: Elimination of accidents with AV by using factor analysis with 4 factors

Uniquenesses:	
mode.of.transportation	frequency.of.driving
0.942	0.571
safety.in.driving.yourself	vehicle.safety.being.passenger
0.697	0.904
alcohol.impaired.driving	safety.due.to.driver.assistance.systems
0.005	0.63
preference.of.automation	characteristics.of.autonomous.vehicles
0.567	0.963
Elimination.of.accidents	
0.398	

Table 4.17: Loadings of factor analysis of model 12

Loadings:				
	Factor1	**Factor2**	**Factor3**	**Factor4**
mode.of.transportation				
frequency.of.driving			**-0.611**	
safety.in.driving.yourself			0.535	
vehicle.safety.being.passenger				
alcohol.impaired.driving		**0.988**		
safety.due.to.driver.assistance.systems	0.407			0.383
preference.of.automation	**0.634**			
characteristics.of.autonomous.vehicles				
Elimination.of.accidents	**0.775**			

	Factor1	**Factor2**	**Factor3**	**Factor4**
SS loadings	**1.213**	**1.007**	**0.753**	0.351
Proportion Var	**0.135**	**0.112**	**0.084**	0.039
Cumulative Var	0.135	0.247	0.33	0.369

In model 12 with four factors, *"elimination of accidents"* and *"preference of automation"* have correlation with "Factor1" and have values of 0.775 and 0.634 respectively. *"Alcohol impaired driving"* has a high positive value of 0.988 for "Factor2" while *"frequency of driving"* has a negative high value of 0.611 for "Factor3" while there is no such correlation for "Factor4".

The SS loadings could be interpreted in a way that the "Factor1" and "Factor2" have values greater than 1 and a proportion variance of 13.5% and 11% while "Factor3" has a proportion variance of 8.4%. All three factors are considered significant.

II. Model 13 (using factor number 5)

In this model, the factor number is changed from 4 to 5 to see the impact of five factors. The results are depicted in table 4.18 and 4.19 below:

Table 4.18: Elimination of accidents with AV by using factor analysis with 5 factors

Uniquenesses:	
mode.of.transportation	frequency.of.driving
0.005	0.65
safety.in.driving.yourself	vehicle.safety.being.passenger
0.661	0.912
alcohol.impaired.driving	safety.due.to.driver.assistance.systems
0.005	0.639
preference.of.automation	characteristics.of.autonomous.vehicles
0.629	0.955
Elimination.of.accidents	
0.131	

Table 4.19: Loadings of factor analysis for model 13

Loadings:					
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
mode.of. transportation			**0.988**		
frequency.of.driving				-0.571	
safety.in.driving.yourself				0.561	
vehicle.safety.being.passenger					
alcohol.impaired.driving		**0.994**			
safety.due.to.driver.assistance.systems					0.499
preference.of.automation	0.493				
characteristics.of.autonomous.vehicles					

Elimination.of.accidents	**0.925**				
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
SS loadings	**1.196**	**1.028**	**0.993**	**0.715**	0.481
Proportion Var	**0.133**	**0.114**	**0.11**	**0.079**	0.053
Cumulative Var	0.133	0.247	0.357	0.437	0.49

In this model 13, *"elimination of accidents"* have a high value of 0.925 with "Factor1", *"alcohol impaired driving"* have a high value of 0.994 with "Factor2" , *"mode of transportation"* has a very high positive value for "Factor3" while there are no such strong correlations for "Factor4" and "Factor5" respectively.

The SS loadings show that "Factor1" and "Factor2" are significant factors having values greater than 1 while "Factor3" could also be considered significant with a value near to 1. "Factor4" has a proportion variance of approximately 8% so it is considered significant as well. "Factor5" are insignificant.

3) Safety of Autonomous Vehicles

Safety of autonomous vehicles is also analyzed by factor analysis using factor 4 and factor 5.

I. Model 14 (using factor number 4)
II. Model 15 (using factor number 5)

I. <u>Model 14 (using factor number 4)</u>

In this model, four factors are used to interpret the loadings and recognize significant factors. The results are presented in table 4.20 and 4.21 below:

Table 4.30: Safety of autonomous vehicles by factor analysis with 4 factors

Uniquenesses:	
mode.of.transportation	frequency.of.driving
0.926	0.582
safety.in.driving.yourself	vehicle.safety.being.passenger
0.703	0.909
alcohol.impaired.driving	safety.due.to.driver.assistance.systems
0.005	0.649
preference.of.automation	characteristics.of.autonomous.vehicles
0.49	0.965
Safety.of.AV	
0.31	

Table 4.21: Loadings with factor analysis for model 14

Loadings:				
	Factor1	**Factor2**	**Factor3**	**Factor4**
mode.of.transportation				
frequency.of.driving			**-0.632**	

safety.in.driving.yourself			0.508	
vehicle.safety.being.passenger				
alcohol.impaired.driving		**0.994**		
safety.due.to.driver.assistance.systems				0.494
preference.of.automation	**0.667**			
characteristics.of.autonomous.vehicles				
Safety.of.AV	**0.803**			
	Factor1	**Factor2**	**Factor3**	**Factor4**
SS loadings	**1.206**	**1.036**	**0.742**	0.477
Proportion Var	**0.134**	**0.115**	**0.082**	0.053
Cumulative Var	0.134	0.249	0.331	0.385

The interpretation of the model could be done as "*preference of automation*" and "*safety of autonomous vehicles*" have positive high values of 0.667 and 0.803 respectively and therefore correlate with "Factor1". "*Alcohol impaired driving*" has a high value of 0.994 which means that it correlates with "Factor2". "*Frequency of driving*" correlates with "Factor3" having a negative value of 0.632. There are no correlations for "Factor4".

The SS loadings show that "Factor1" and "Factor2" have high significant values of greater than 1 with proportion variances of 13.4% and 11.5% respectively which should be kept. "Factor3" has a proportion variance of 8.2% which is considered significant for this study as well.

II. Model 15 (using factor number 5)

In this model, five factors are used to interpret the loadings and recognize significant factors against the four factors being used in the previous model. The results of factor analysis are demonstrated in table 4.22 and 4.23 respectively.

Table 4.22: Factor analysis performed on safety of autonomous vehicles with 5 factors

Uniquenesses:	
mode.of.transportation	frequency.of.driving
0.9	0.624
safety.in.driving.yourself	vehicle.safety.being.passenger
0.67	0.896
alcohol.impaired.driving	safety.due.to.driver.assistance.systems
0.005	0.693
preference.of.automation	characteristics.of.autonomous.vehicles
0.487	0.005
Safety.of.AV	
0.313	

Table 4.23: Loadings of factor analysis for model 15

Loadings:					
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
mode.of. transportation					
frequency.of.driving				-0.557	
safety.in.driving.yourself				0.566	
vehicle.safety.being.passenger					0.292
alcohol.impaired.driving		**0.992**			
safety.due.to.driver.assistance.systems	0.396				
preference.of.automation	**0.693**				
characteristics.of.autonomous.vehicles			**0.994**		
Safety.of.AV	**0.823**				
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
SS loadings	**1.348**	**1.024**	**1.002**	**0.724**	0.308
Proportion Var	**0.15**	**0.114**	**0.111**	**0.08**	0.034
Cumulative Var	0.15	0.264	0.375	0.455	0.49

For this model, the "*preference of automation"* and "*safety of autonomous vehicles"* have high positive values for "Factor1" which means that these two factors are strongly correlated with Factor1. Similarly, "*alcohol impaired driving"* has a high loading value for "Factor2", "*characteristics of autonomous vehicles"* is strongly correlated with "Factor3" while no such strong correlations with "Factor4" and "Factor5" are to be found.

The SS loadings show that "Factor1", "Factor2", "Factor3" and "Factor4" are all considered significant in this scenario and are kept with proportion variances of 15%, 11.4%, 11.1% and 8% respectively.

4.10 Comparison between Ordinal logistic regression and Factor analysis

As the analysis was performed with two techniques and both giving different outputs, the factors that affect the dependent variables could be selected on a comparison between the two techniques for all three dependent variables.

Table 4.24: Comparison of the models

Dependent Variable	Ordinal Logistic Regression	Factor Analysis
Mode shift	age, employment and preference of automation	preference of automation, alcohol impaired driving, frequency of driving and safety due to driver assistance systems
Elimination of accidents	age, employment, alcohol impaired drivers, safety due to driver assistance	

	systems, preference of automation and characteristics of autonomous vehicles	alcohol impaired driving, mode of transportation and frequency of driving
Safety of autonomous cars	age, employment, safety in driving the car yourself, performing other tasks while driving, frequency of performing other tasks, safety due to driver assistance systems and preference of automation	preference of automation, alcohol impaired driving, safety in driving the car yourself, characteristics of autonomous vehicles

As shown in table 4.24, the demographic factors were not taken into consideration for factor analysis, apart from those factors if the other common factors are compared for each dependent variable, it can be seen that preference of automation and safety in driving the car yourself are the significant variables in the mode shift towards autonomous driving and the safety of autonomous cars in comparison to human-driven vehicles while the variable of alcohol impaired driving plays an important role for the people to perceive the elimination of accidents with autonomous driving for both ordinal logistic regression and factor analysis.

Chapter 5: Conclusions

5.1 Summary

This section gives a brief background of the research indicating the different developments being made in the field autonomous driving, the purpose of conducting this research along with the methods deployed in conducting the research.

In this research, factors affecting the shift to autonomous vehicles were found out by an online survey questionnaire for the inhabitants of the city of Munich. The purpose of the research was to find out the awareness among the masses regarding autonomous vehicles, and the extent of the desire of the people to shift towards autonomous vehicles in the future as the field of autonomous driving is entirely new and still developing, and the models are being tested in the market. There has been a lot of research which is being carried out in the field of autonomous driving, but it lacks a more in-depth analysis of the safety parameters and the factors in general which would play a vital role for the people to choose self-driven cars over human driven cars in future.

Firstly, the literature is reviewed keeping in mind the aim of the study. The focus is on safety with respect to autonomous vehicles. The gap in the literature which is found helped in developing the research methodology. The research is mostly qualitative and is performed by creating a survey questionnaire which includes the questions related to the general perspective of driving a car moving towards autonomous vehicles and the safety of autonomous vehicles. After getting the required data from the questionnaire, it is analyzed through R Studio using two analysis techniques which are Ordinal Logistic Regression Analysis and Factor Analysis. These two techniques are applied due to their suitability and relevance to the nature and amount of data gathered. R Studio is used to perform the data analysis using fifteen independent variables and three dependent variables for Ordinal logistic regression while using nine factor loadings for Factor Analysis. There are fifteen models in total for the analysis. The output of the models is interpreted and compared at the end to choose the model very near to the true model and which factors it points out. The results show that the output given by the regression model made with the p-values and t-values is statically significant and close to the true model with respect to the AIC values.

5.2 Significant Findings

This section explains the main findings of the research with different constructed models and their future implications.

The significant findings of the research depict that that p-values and t-values play an important role in finding the most statistically significant model. The model chosen at the end based on the comparisons with regards to BIC, AIC was the one created with the variables that were screened based on p-values and t-values.

The factors that influence the mode shift towards autonomous vehicles are "*age", "employment"* and "*preference of automation"* while the factors that influence the perception of people regarding the elimination of accidents with the introduction of autonomous cars are "*age", "employment", "alcohol impaired drivers", "safety due to drivers assistance systems", "preference of automation"* and "*characteristics of autonomous vehicles".* With regards to the anticipation of people about safety of autonomous vehicles as compared to the human-driven vehicles, the factors that affect the safety of autonomous vehicles are "*age", "employment", "safety in driving the car yourself", "performing other tasks while driving", "frequency of performing other tasks", "safety due to driver assistance systems"* and "*preference of automation".*

For factor analysis, there were two models for each dependent variable. The demographic features were not selected amongst the variables selected for factor analysis as they vary from person to person. The models were created by using four and five factors as they are retained by the factor retention methods. The factors which affect the mode choice are "*preference of automation", "alcohol impaired driving*", "*safety due to driver assistance systems"* and "*frequency of driving"* as these loadings are statistically correlated with more factors and show more proportion of variance . The factors that affect the perception of people about elimination of accidents with the introduction of self-driven cars are "*alcohol impaired driving", "frequency of driving" and* "*mode of transportation".* The factors selected from the output of the model for the safety of autonomous vehicles with respect to introduction of autonomous cars are "*preference of automation", "alcohol impaired driving", "safety in driving the car yourself*" and "*characteristics of autonomous vehicles".*

It can be concluded from the model results that there are certain human behaviors and specific vehicle characteristics which play a vital role in this transition from standard vehicles to self-driven cars while there are demographic factors as well which vary a lot but have an impact on the mode shift. The factors point out towards the upcoming challenges and hurdles that the autonomous cars will have to face for people to leave their human-driven vehicles and shift to the self-driven vehicles. It will be a significant job for the manufacturers and legislators to deal with all the technical parameters and safety hazards which people are careful about while shifting to autonomous cars.

5.3 Limitations and Future Research

This section explains the limitations related to the methodological approach and the possible sample bias for conducting this research and the future research that can be carried out to make this thesis better.

One of the limitations is the relatively small sample size as compared to the city population and better results could be obtained if the sample size is greater. Another limitation is the possible sample bias and if the research is performed by controlling individual characteristics then it might yield better results. If there had been more time, then a large sample size could have given more reliable results and the comparison could be made with more statistical models.

For the future research and recommendations, the following changes could be made:

1. The research was carried out for the city of Munich and its inhabitants and the same research could be performed in another city or region that might yield different results based on the approach, thinking and privacy concerns of the people.
2. The factors which affect the mode shift towards autonomous cars could be further practically implemented in the field of self-driven cars while making policies and general awareness among the people.
3. Furthermore, analytical models other than ordinal logistic and factor analysis could be used to analyze the factors which are not part of this research study due to time limitations.

Bibliography

ASIRT (2018). "Road Safety Facts." Retrieved 15th July, 2018.

Atlas.ti (2011). "Data Analysis – Analyzing Data in Qualitative Research." Retrieved 18.06, 2018, from https://atlasti.com/analyzing-data-data-analysis/.

Bagloee, S. A., et al. (2016). "Autonomous vehicles: challenges, opportunities, and future implications for transportation policies." Journal of Modern Transportation **24**(4): 284-303.

Bellem, H., et al. (2018). "Comfort in automated driving: An analysis of preferences for different automated driving styles and their dependence on personality traits." Transportation Research Part F: Traffic Psychology and Behaviour **55**: 90-100.

Brar, J. S. and B. Caulfield (2017). Impact of autonomous vehicles on pedestrians' safety. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC).

Brown, J. D. (2001). "What is an eigenvalue?": 15-19.

Chee, J. (2015). Pearson's Product-Moment Correlation: Sample Analysis.

Cornish (2007). "Cluster Analysis."

Crawford, I. M. (1997). Marketing Research and Information Systems. Rome, FAO Regional Office for Africa.

Dahiru (2008). "P – VALUE, A TRUE TEST OF STATISTICAL SIGNIFICANCE? A CAUTIONARY NOTE."

DePersio, G. (2018). "What are the advantages of using a simple random sample to study a larger population?". Retrieved 16th June, 2018, from https://www.investopedia.com/ask/answers/042915/what-are-advantages-using-simple-random-sample-study-larger-population.asp.

Dikmen, M. and C. Burns (2017). Trust in autonomous vehicles: The case of Tesla Autopilot and Summon. 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC).

dmartin (2014). "AIC guidelines in model selection." Retrieved 15th September, 2018, from https://stats.stackexchange.com/q/81431.

Etikan, B. (2017). "Sampling and Sampling Methods." **5**(6).

Fagnant, D. J. and K. Kockelman (2015). "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations." Transportation Research Part A: Policy and Practice **77**: 167-181.

Ford, C. (2016). "Getting Started with Factor Analysis."

Fraedrich, E., et al. (2016). Technical Report: User Perspectives on Autonomous Driving. A Use-Case-Driven Study in Germany.

Fraedrich, E., et al. (2016). "User Perspectives on Autonomous Driving: A Use-Case-Driven Study in Germany."

Frost, J. (2017). "How to Interpret P-values and Coefficients in Regression Analysis."

Gie Yong, A. and S. Pearce (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis.

Goforth, C. (2015). "Using and Interpreting Cronbach's Alpha."

Gruel, W. and J. M. Stanford (2016). Assessing the Long-term Effects of Autonomous Vehicles: A Speculative Approach.

Haboucha, C. J., et al. (2017). "User preferences regarding autonomous vehicles." Transportation Research Part C: Emerging Technologies **78**: 37-49.

Haboucha, C. J., et al. (2017). "User preferences regarding autonomous vehicles."

Hasenjäger, M. and H. Wersing (2017). Personalization in advanced driver assistance systems and autonomous vehicles: A review. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC).

Hayton, J. C., et al. (2004). "Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis." Organizational Research Methods **7**(2): 191-205.

Heilig, M., et al. (2017). "Potentials of Autonomous Vehicles in a Changing Private Transportation System – a Case Study in the Stuttgart Region." Transportation Research Procedia **26**: 13-21.

Hollier, L. P., et al. (2017). "Comparing online and telephone survey results in the context of a skin cancer prevention campaign evaluation." J Public Health (Oxf) **39**(1): 193-201.

Holstein, D.-C., Pelliccione (2018). "Ethical and Social Aspects of Self-Driving Cars."

Howard, D. (2013). "Public Perceptions of Self-Driving Cars: The Case of Berkeley, California."

Hulse, L. M., et al. (2018). "Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age." Safety Science **102**: 1-13.

IIHS (2017). "General statistics."

Jiménez, F., et al. (2016). "Advanced Driver Assistance System for Road Environments to Improve Safety and Efficiency." Transportation Research Procedia **14**: 2245-2254.

Kaan, J. (2017). "User Acceptance of Autonomous Vehicles: Factors & Implications."

Kaur, K. and G. Rampersad (2018). "Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars." Journal of Engineering and Technology Management **48**: 87-96.

Koopman, P. and M. Wagner (2017). "Autonomous Vehicle Safety: An Interdisciplinary Challenge." IEEE Intelligent Transportation Systems Magazine **9**(1): 90-96.

Kyriakidis, M., et al. (2015). "Public opinion on automated driving: Results of an international questionnaire among 5000 respondents." Transportation Research Part F: Traffic Psychology and Behaviour **32**: 127-140.

Lewis Hill, A. C., Joe Wheeler (2017). Intelligent Vehicles: A public attitude survey.

Li, S., et al. (2018). Clustering of Naturalistic Driving Encounters Using Unsupervised Learning.

McCaffrey, J. (2017). "Revealing Secrets with R and Factor Analysis." Retrieved 15th September, 2018, from https://visualstudiomagazine.com/articles/2017/03/01/revealing-secrets-r-factor-statistics.aspx.

NHTSA (2017). "Automated Vehicles for Safety." Retrieved 16th July, 2018, from https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety.

Norušis, M. (2012). Cluster Analysis. IBM SPSS Statistics 19 Statistical Procedures Companion.

Noy, I. Y., et al. (2018). "Automated driving: Safety blind spots." Safety Science **102**: 68-78.

Park, H. M. (2017). "Factor Analysis for Questionnaire Survey Data: Exploratory and Confirmatory Factor Analysis."

Payre, W., et al. (2014). "Intention to use a fully automated car: Attitudes and a priori acceptability." Transportation Research Part F: Traffic Psychology and Behaviour **27**: 252-263.

Rahn, M. (2018). "Factor Analysis: A Short Introduction, Part 4–How many factors should I find?". Retrieved 27th September, 2018, from https://www.theanalysisfactor.com/factor-analysis-how-many-factors/.

Ray, S. (2015). "7 Types of Regression Techniques you should know!". from https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/.

Reeve, M. a. (2015). "Why we should get used to the idea that self- driving cars will sometimes crash."

Reschka, A. (2016). Safety Concept for Autonomous Vehicles. Autonomous Driving: Technical, Legal and Social Aspects. M. Maurer, J. C. Gerdes, B. Lenz and H. Winner. Berlin, Heidelberg, Springer Berlin Heidelberg**:** 473-496.

Ross, C. and S. Guhathakurta (2017). "Autonomous Vehicles and Energy Impacts: A Scenario Analysis." Energy Procedia **143**: 47-52.

Saleh, K., et al. (2017). Towards trusted autonomous vehicles from vulnerable road users perspective. 2017 Annual IEEE International Systems Conference (SysCon).

ScienceBuddies. "Designing a Survey." Retrieved 07th June, 2018, from https://www.sciencebuddies.org/science-fair-projects/references/how-to-design-a-survey.

Seif, H. G. and X. Hu (2016). "Autonomous Driving in the iCity—HD Maps as a Key Challenge of the Automotive Industry." Engineering **2**(2): 159-162.

Sivak, S. (2015). Road Safety with Self-Driving Vehicles.

SSC (2001). "Modern Approaches to the Analysis of Survey Data."

Stephanie (2018). "T Test (Student's T-Test): Definition and Examples." Retrieved September 08, 2018.

SurveyMonkey. "How to design a survey." Retrieved 07th June, 2018, from https://www.surveymonkey.com/mp/survey-guidelines/?ut_source=header.

Tamara van, G., et al. (2007). Data Collection and Analysis. Handbook of Research on Educational Communications and Technology, Routledge.

Tan, P. N., et al. (2005). Cluster Analysis: Basic Concepts and Algorithms.

Thomopoulos, N. and M. Givoni (2015). "The autonomous car—a blessing or a curse for the future of low carbon mobility? An exploration of likely vs. desirable outcomes." European Journal of Futures Research **3**(1): 14.

ttnphns (2017). "Is there Factor analysis or PCA for ordinal or binary data?". Retrieved 15th September, 2018, from https://stats.stackexchange.com/q/215483.

Urooj, S., et al. (2018). Systematic literature review on user interfaces of autonomous cars: Liabilities and responsibilities. 2018 International Conference on Advancements in Computational Sciences (ICACS).

V.Casely, J., Quartulli (2013). "A study of public acceptance of autonomous cars."

Van Brummelen, J., et al. (2018). "Autonomous vehicle perception: The technology of today and tomorrow." Transportation Research Part C: Emerging Technologies **89**: 384-406.

Vosloo, J. J. (2014). "Data Analysis and Interpretation ".

Wei, J., et al. (2013). Towards a viable autonomous driving research platform. 2013 IEEE Intelligent Vehicles Symposium (IV).

Yang, Z., et al. (2014). Driving risk assessment using cluster analysis based on naturalistic driving data. 17th International IEEE Conference on Intelligent Transportation Systems (ITSC).

Zmud, J., et al. (2016). Self-Driving Vehicles Determinants of Adoption and Conditions of Usage.

Appendix: Survey Questionnaire

Introduction:

Autonomous Vehicles are the future part of the road traffic. It is the need of the time to research about this field furthermore and look for the factors that will affect the shift of the people from the normal vehicles to autonomous vehicles. The factors that will play a vital role with respect to the safety are being discussed in this survey to analyse the data that will be useful for further research in this field.

Instructions and Information:

Autonomous driving is the self-driving of a vehicle without the involvement of a human driver. The autonomous car uses V2V (Vehicle to Vehicle) and V2I (Vehicle to Infra-structure) sensors connectivity technology to run throughout the system. According to *(NHTSA 2017)*, there are mainly five levels of automation.

Level 0: No Automation (The driver performs all the tasks)

Level 1: Driver Assistance (The driver controls the vehicle, but some driving assistance functions are built in the design of the vehicle)

Level 2: Partially Automated (The driver remains engaged with the driving task and monitors the surroundings, but the vehicle has automated functions like acceleration and steering etc.)

Level 3: Highly Automated (The vehicle has the capacity to perform all driving functions under certain conditions while the driver has the option to control the vehicle when wanted)

Level 4: Fully Automated (The vehicle has the capacity to perform all driving functions under all conditions while driver has the option to take over the vehicle)

The survey questionnaire below will take approximately 10 minutes to answer. Please take your time. The purpose of the survey is to find out the safety factors that will affect the shift from the manual driving vehicles to autonomous vehicles. The survey is based upon the residents of the city of Munich. Be open minded and unbiased to the questions. Thankyou

Transport Section:

Q.No.	Questions	Preferred Choices

1.	What is your primary mode of transportation for most trips in a day?	1. Public Transportation 2. Private Car 3. Motorbike 4. Bicycle 5. Walking 6. Other
2.	Do you possess a driving license?	1. Yes 2. No
3.	If Yes, which license do you possess?	1. EU 2. International 3. Other
4.	How often do you drive a car? (choose the closest one)	1. Everyday 2. Frequently (2-3 days per week) 3. Seldom (2-3 days per month) 4. Never
5.	As a driver, were you involved in any car accident with injury since you have started driving?	1. Yes 2. No
6.	How would you self-assess your type of driving behaviour?	1. Aggressive 2. Defensive 3. Other
7.	On a scale from 1-5, how do you rate your driving skills?	1. Very Weak 2. Weak 3. Average 4. Above Average 5. Professional

Safety Section:

Q.No.	Questions	Preferred Choices
8.	How safe do you feel while driving the vehicle yourself?	1. Very unsafe 2. Rather unsafe 3. Moderate 4. Rather safe 5. Very safe
9.	How safe does wearing a helmet on a bike or seatbelt in a car make you feel while driving?	1. Very unsafe 2. Rather unsafe 3. Moderate 4. Rather safe 5. Very safe

10.	Do you perform other tasks while driving like talking or texting on the cell phone, eating etc?	1. Yes 2. No
11.	How often do you perform these other tasks while driving?	1. Very seldom 2. Rather Seldom 3. Moderate 4. Rather frequently 5. Very frequently
12.	How much preference do you give to safety while buying a new vehicle?	1. Very Low 2. Low 3. Moderate 4. High 5. Very High
13.	How safe do you feel inside a vehicle being a passenger?	1. Very unsafe 2. Rather unsafe 3. Moderate 4. Rather safe 5. Very safe
14.	How much do you think vulnerable road users (Pedestrians, Bicyclists etc.) affect road safety for car or motorbike drivers?	1. Very less 2. Rather less 3. Moderate 4. Rather high 5. Very high
15.	Has any disability prevented you from manually driving a vehicle?	1. Yes 2. No
16.	Do you know people who drink too much alcohol (as allowed legally) and drive their vehicles afterwards?	1. Very few 2. Rather few 3. Moderate 4. Rather many 5. Too many
17.	How often have you driven a vehicle under the influence of alcohol?	1. Very seldom 2. Rather Seldom 3. Moderate 4. Rather frequently 5. Very frequently
18.	How much do you think that Driver assistance systems (Absolute Braking System, Advanced Cruise Control, Lane assist etc.) offer more road safety?	1. Very less 2. Rather less 3. Moderately 4. Rather high 5. Very high

19.	Have you received a ticket for any violation of any traffic rule or law in the last year including Parking?	1. Yes 2. No
20.	If you have received a ticket for any traffic violation, which violation have you done?	________________
21.	How many times have you been fined for Over Speeding?	1. None 2. 1-2 3. 3-7 4. 7-10 5. >10

Safety Section with respect to Autonomous Driving:

Q. No	Questions	Preferred Choices
22.	How familiar are you with autonomous driving?	1. Very unfamiliar 2. Rather unfamiliar 3. Moderately 4. Rather familiar 5. Very familiar
23.	How much does the idea of Autonomous Vehicles fascinate you?	1. Very less 2. Rather Less 3. Moderate 4. High 5. Very High
24.	If you have to choose, which type of automation would you prefer?	1. Non-Autonomous 2. Driver Assistance 3. Partially Automated 4. Highly Automated 5. Fully Automated
25.	How satisfactory would you feel inside an Autonomous Vehicle?	1. Very unsatisfactory 2. Rather unsatisfactory 3. Moderately 4. Rather satisfactory 5. Very satisfactory
26.	How much do you think Driver assistance functions help in improving road safety?	1. Very less 2. Rather less 3. Moderate 4. Rather high 5. Very high
27.	How safe would you feel travelling in a vehicle without a steering wheel?	1. Very unsafe 2. Rather unsafe 3. Moderate

		4. Rather safe 5. Very safe
28.	Which characteristic of Autonomous Vehicles would you consider while buying an autonomous vehicle?	1. Level of Automation 2. Automatic Braking 3. Automatic Lane Change 4. Comfort of Passenger 5. All of the above 6. Others __________
29.	If you would purchase a new vehicle, what is the maximum amount you would be willing to pay for an autonomous vehicle?	1. Less than 30,000 Euros 2. 30,000 – 50,000 Euros 3. 51,000 – 70,000 Euros 4. 71,000 – 100,000 Euros 5. More than 100,000 Euros
30.	How likely would you be to shift to autonomous vehicle considering the right price and reliability with respect to safety?	1. Very unlikely 2. Rather unlikely 3. Moderately 4. Rather likely 5. Very likely
31.	Would you agree that your vehicle data (location, speed, etc.) will be passed on to vehicles in the area for better networking?	1. Yes 2. No
32.	In the last year in USA it was known that the software of an autonomous car could be hacked. How high do you estimate this risk with autonomous driving?	1. Very unlikely 2. Rather unlikely 3. Moderately 4. Rather likely 5. Very likely
33.	Realizing the Social Dilemma of Autonomous Vehicles, if the vehicle is going to be involved in an accident with a group of pedestrians, should it save the driver or the group of Pedestrians?	1. Driver 2. Pedestrian
34.	How safe do you think autonomous vehicles would be compared to human-driven vehicles?	1. Very unsafe 2. Rather unsafe 3. Moderate 4. Rather safe 5. Very safe
35.	How likely do you think that Autonomous Vehicles will eliminate traffic accidents?	1. Very unlikely 2. Rather unlikely 3. Moderate 4. Rather likely 5. Very likely

Demographics Section:

Q. No	Questions	Preferred Choices
36.	What is your Age?	1. Under 18 2. 18-30 3. 31-50 4. 51-70 5. >70 years 6. Prefer not to answer
37.	Please choose your Gender:	1. Male 2. Female 3. Other 4. Prefer not to answer
38.	What is your Marital Status?	1. Single 2. Married 3. Divorced 4. Widowed 5. Separated 6. Others 7. Prefer not to answer
39.	Do you have Children?	1. Yes 2. No 3. Prefer not to answer
40.	What is your Highest Education qualification?	1. Less than High School Diploma 2. High School Degree 3. College Degree 4. Bachelors Degree 5. Masters Degree 6. Doctorate Degree 7. Prefer not to answer
41.	What is your current Employment Status?	1. Full time 2. Part time 3. Student 4. Retired 5. Self-Employed 6. Homemaker 7. Others ____________ 8. Prefer not to answer
42.	What is your net household income per year?	1. Less than 20,000 Euros. 2. 21,000-60,000 Euros 3. 61,000-80,000 Euros

		4. More than 80,000 Euros 5. Prefer not to answer

Fragebogen

Einleitung:

Autonome Fahrzeuge sind die Zukunft des Straßenverkehrs. Es ist an der Zeit diesen Bereich tiefer zu erforschen und die Faktoren zu ergründen, welche den Wechsel von selbst gesteuerten Autos zu autonomen Fahrzeugen ermöglichen. Der Zweck der Untersuchung ist es Sicherheitsfaktoren aufzudecken, welche den Wechsel von manuell gefahrenen Fahrzeugen zu autonom gesteuerten Fahrzeugen beeinflussen, um Ansatzpunkte für eine zukünftige tiefergehende Analyse zu schaffen.

Anweisungen und Informationen:

Autonomes Fahren ist ein selbstfahrendes Fahrzeug ohne das Einwirken eines menschlichen Fahrers. Das autonome Fahrzeug benutzt V2V (Fahrzeug zu Fahrzeug) und V2I (Fahrzeug zu Infrastruktur) Sensoren-Kommunikationstechnologie um vollumfänglich funktionieren zu können. Es gibt fünf unterschiedliche Stufen des autonomen Fahrens:

Level 0: Kein autonomes Fahren (Der Fahrer übernimmt vollständig)

Level 1: Fahrerassistenz (Der Fahrer kontrolliert das Fahrzeug, aber verschiedene Fahrerassistenzsysteme unterstützen den Fahrer)

Level 2: Teilweise Autonom (Der Fahrer bleibt mit den Fahreraufgaben beschäftigt und überwacht die Umgebung während das Fahrzeug autonom Funktionen wie z.B. Beschleunigung und Steuerung übernimmt)

Level 3: Überwiegend Autonom (Das Fahrzeug kann alle Funktionen zur Bedienung des Fahrzeugs unter bestimmten Bedingungen übernehmen während der Fahrer die Möglichkeit besitzt die Kontrolle zu übernehmen)

Level 4: Voll-Autonom (Das Fahrzeug kann alle Funktionen zur Bedienung des Fahrzeugs unter allen Bedingungen übernehmen während der Fahrer die Möglichkeit besitzt die Kontrolle zu übernehmen)

Die Beantwortung des Fragebogens wird ca. 10 Minuten dauern. Bitte nehmen Sie sich diese Zeit! Der Zweck der Untersuchung ist es Sicherheitsfaktoren aufzudecken, welche den Wechsel von manuell gefahrenen Fahrzeugen zu autonom gesteuerten Fahrzeugen beeinflussen. Die Untersuchung beruht auf Befragungen von Einwohnern Münchens. Gehen Sie aufgeschlossen und vorurteilsfrei an die Fragen heran.

Transportabschnitt:

Frage Nummer	Frage	Antwort

1.	Wie bewegen Sie sich meist im Verlauf eines Tages fort?	1. Öffentliche Verkehrsmittel 2. Pkw 3. Motorrad/Roller 4. Fahrrad 5. Zu Fuß 6. Anderes
2.	Besitzen Sie einen Führerschein?	1. Ja 2. Nein
3.	Falls Ja, welchen Führerschein haben Sie?	1. EU Führerschein 2. Internationaler Führerschein 3. Anderer
4.	Wie oft fahren Sie selbst ein Auto? (Wählen Sie die nächstliegende Antwortmöglichkeit)	1. Täglich 2. Häufig (2-3 Tage pro Woche) 3. Selten (2-3 Tage pro Monat) 4. Nie
5.	Waren Sie als Fahrer in einen Verkehrsunfall involviert, bei dem es Verletzte gab?	1. Ja 2. Nein
6.	Wie würden Sie Ihren Fahrstil beschreiben?	1. Aggressiv 2. Defensiv 3. Anders
7.	Bitte schätzen Sie Ihre Fahrerfertigkeiten auf einer Skala von 1-5 ein:	1. Sehr schwach 2. Schwach 3. Durchschnittlich 4. Überdurchschnittlich 5. Professionell

Sicherheitabschnitt:

Frage Nummer	Frage	Antwort
8.	Wie sicher fühlen Sie sich wenn Sie selbst das Fahrzeug fahren?	1. Sehr unsicher 2. Eher Unsicher 3. Teils/teils 4. Eher sicher 5. Sehr sicher
9.	Wie sicher fühlen Sie sich, falls Sie einen Helm tragen oder einen Sicherheitsgurt anlegen?	1. Sehr unsicher 2. Eher unsicher 3. Teils/teils 4. Eher sicher 5. Sehr sicher

10.	Beschäftigen Sie sich während des Autofahrens mit anderen Tätigkeiten wie z.B. telefonieren, SMS schreiben, essen?	1. Ja 2. Nein
11.	Wie oft üben Sie diese anderen Tätigkeiten während des Fahrens aus?	1. Sehr selten 2. Eher selten 3. Teils/teils 4. Eher häufig 5. Sehr Häufig
12.	Wie wichtig sind Ihnen Sicherheitsfaktoren beim Kauf eines Fahrzeugs?	1. Sehr unwichtig 2. Eher unwichtig 3. Teils/teils 4. Eher wichtig 5. Sehr wichtig
13.	Wie sicher fühlen Sie sich in einem Fahrzeug als Mitfahrer?	1. Sehr unsicher 2. Eher unsicher 3. Teils/teils 4. Eher sicher 5. Sehr sicher
14.	Wie wirken sich ungeschützte Verkehrsteilnehmer (Fußgänger, Fahrradfahrer usw.) auf die Straßensicherheit für Auto oder Motorradfahrer aus?	1. Gar nicht 2. Eher nicht 3. Teils/teils 4. Eher sehr 5. Sehr
15.	Hat Sie eine Behinderung/Beeinträchtigung bisher davon abgehalten selbst ein Fahrzeug zu steuern?	1. Ja 2. Nein
16.	Jeder weiß man soll nicht Alkohol trinken und anschließend ein Fahrzeug steuern. Haben Sie Bekannte, die Alkohol (über der erlaubten Grenze) trinken und anschließend ein Fahrzeug steuern?	1. Sehr wenige 2. Eher wenige 3. Ein paar 4. Mehrere 5. Viele
17.	Wie oft haben Sie ein Fahrzeug unter Alkoholeinfluss gesteuert?	1. Sehr selten 2. Eher selten 3. Teils/teils 4. Eher häufig 5. Sehr häufig
18.	Wirken sich Ihrer Ansicht nach Fahrerassistenzsysteme (Notfallbremsassistent, Automatische Distanzregelung, Spurhalteassistent, usw.) auf die Straßensicherheit aus?	1. Sehr wenig 2. Eher wenig 3. Teils/teils 4. Eher hoch 5. Sehr hoch
19.	Haben Sie innerhalb des letzten Jahres Strafzettel wegen Missachtung der Straßenverkehrsordnung einschließlich Falschparkens erhalten?	1. Ja 2. Nein

20.	Wenn Sie ein Ticket für einen Verkehrsverstoß erhalten haben, welchen Verstoß haben Sie dann begangen?	________________
21.	Wie oft erhielten Sie Strafzettel für Übertretung der Höchstgeschwindigkeit?	1. Keine 2. 1-2 3. 3-7 4. 7-10 5. >10

Sicherheitsabschnitt von autonomen Fahrzeugen:

Frage Nummer	Frage	Antwort
22.	Wie vertraut sind Sie mit autonomen Fahren?	1. Nicht vertraut 2. Eher nicht vertraut 3. Teils/teils 4. Eher vertraut 5. Sehr vertraut
23.	Sind Sie vom autonomen Fahren fasziniert?	1. Überhaupt nicht 2. Eher nicht 3. Teils/teils 4. Eher sehr 5. Sehr
24.	Für welches Level des autonomen Fahrens würden Sie sich entscheiden?	1. Kein autonomes Fahren 2. Fahrerassistenz 3. Teilweise Autonom 4. Überwiegend Autonom 5. Voll-Autonom
25.	Wie zufrieden würden Sie sich innerhalb eines autonomen Fahrzeugs finden?	1. Nicht zufrieden 2. Eher nicht zufrieden 3. Teils/teils 4. Eher zufrieden 5. Sehr zufrieden
26.	Haben Fahrerassistenzsysteme einen Einfluss auf die Straßensicherheit?	1. Sehr wenig 2. Eher wenig 3. Teils/teils 4. Eher sehr 5. Sehr
27.	Wie sicher würden Sie sich in einem Fahrzeug ohne Lenkrad fühlen?	1. Sehr unsicher 2. Eher unsicher 3. Teils/teils 4. Eher sicher 5. Sehr sicher

28.	Welche Charakteristiken eines autonomen Fahrzeugs würden Sie persönlich beim Kauf eines Fahrzeugs berücksichtigen?	1. Level des autonomen Fahrens 2. Automatisches Bremsen 3. Automatisches Spurwechseln 4. Sitzkomfort 5. Alle der genannten 6. Andere
29.	Falls Sie ein neues Fahrzeug kaufen würden, wie viel würden Sie maximal für ein autonomes Fahrzeug ausgeben?	1. Weniger als 30.000 2. 30.000 – 50.000 Euros 3. 51.000 – 70.000 Euros 4. 71.000 – 100.000 Euros 5. Mehr als 100.000
30.	Wie wahrscheinlich wäre für Sie ein Wechsel zu einem autonomen Fahrzeug, falls ein angemessener Preis und Sicherheit vorausgesetzt wären?	1. Sehr unwahrscheinlich 2. Eher unwahrscheinlich 3. Teils/teils 4. Eher wahrscheinlich 5. Sehr wahrscheinlich
31.	Wären Sie damit einverstanden, dass Ihre Fahrzeugdaten (Standort, Geschwindigkeit usw.) an Fahrzeuge in der Umgebung für eine bessere Vernetzung weitergegeben werden?	1. Ja 2. Nein
32.	Im letzten Jahr wurden in den USA Fälle bekannt, dass die Software von autonomen Fahrzeugen gehackt wurde. Wie hoch schätzen Sie dieses Risiko bei autonomen Fahrzeugen ein?	1. Sehr unwahrscheinlich 2. Eher unwahrscheinlich 3. Teils/teils 4. Eher wahrscheinlich 5. Sehr wahrscheinlich
33.	Wie stehen Sie zu folgendem sozialen Dilemma: Ein autonomes Fahrzeug wird in einen Unfall mit einer Gruppe Fußgänger involviert. Soll das autonome Fahrzeug den Fahrer oder die Gruppe Fußgänger schützen?	1. Fahrer 2. Fußgänger
34.	Wie schätzen Sie die Sicherheit von autonomen Fahrzeugen im Vergleich zu von Menschen gesteuerten Fahrzeugen ein?	1. Sehr unsicher 2. Eher unsicher 3. Teils/teils 4. Eher sicher 5. Sehr sicher
35.	Wie wahrscheinlich finden Sie, dass Verkehrsunfälle durch autonome Fahrzeuge verhindert werden könnten?	1. Sehr unwahrscheinlich 2. Eher unwahrscheinlich 3. Teils/teils 4. Eher wahrscheinlich 5. Sehr wahrscheinlich

Demographics Section:

Q. No	Questions	Preferred Choices
36.	Wie alt sind Sie?	1. Jünger als 18 Jahre 2. 18-30 3. 31-50 4. 51-70 5. Älter als 70 Jahre 6. Keine Angabe
37.	Bitte wählen Sie Ihr Geschlecht	1. Herr 2. Frau 3. Anders 4. Keine Angabe
38.	Wie ist Ihr Familienstand?	1. Ledig 2. Verheiratet 3. Geschieden 4. Verwitwet 5. Geschieden 6. Keine Angabe
39.	Haben Sie Kinder?	1. Ja 2. Nein 3. Keine Angabe
40.	Was ist Ihr höchster Bildungsabschluss?	1. Hauptschul-/Mittelschulabschluss 2. Mittlere Reife 3. Abitur 4. Bachelor 5. Master 6. Doktorat 7. Keine Angabe
41.	Ordnen Sie sich bitte einer beruflichen Kategorie zu:	1. Angestellte/-r 2. Arbeiter/-in 3. Auszubildende/-r 4. Beamte/-r 5. Hausfrau/-mann 6. Schüler/-in 7. Selbstständige/-r 8. Student/-in 9. Rentner/-in 10. Sonstiges ______________
42.	Wie hoch ist Ihr Haushalts-Nettoeinkommen?	1. Weniger als 20.000 Euros 2. 21.000 - 40.000 Euros 3. 41.000 - 60.000 Euros 4. 61.000 - 80.000 Euros

		5. Mehr als 80.000 Euros 6. Keine Angabe