People's Preferences in the Adoption of New Mobility Technologies:

A Case Study of Puget Sound Region

By

Saud Bashar Abdulmunem Al Fadhili

B.S., University of Technology- Baghdad, 2017

THESIS

Submitted as partial fulfillment of the requirements

For the degree of Master of Science in Civil Engineering

In the Graduate College of the

University of Illinois at Chicago, 2020

Defense Committee:

Abolfazl Moahammadian, Chair and Advisor, Civil and Materials Engineering

Sybil Derrible, Civil and Materials Engineering

Ramin Shabanpour, Civil and Materials Engineering

TABLE OF CONTENTS

:

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AV	Autonomous Vehicles
SAE	Society of Automotive Engineers
TNC	Transportation Network Companies
NHTSA	National Highway Traffic Safety Administration
FHWA	Federal Highway Administration
ACC	Adaptive Cruise Control
LIDAR	Light Detection And Ranging
VMT	Vehicle Miles Traveled
FAD	Fully Automated Driving
PTA	Public Transport Authority
EMTA	European Metropolitan Transport Authorities
BRT	Bus Rapid Transit
SAV	Shared Autonomous Vehicles
DRS	Dynamic Ridesharing
WTP	Willingness To Pay

Summary

This is a study where U.S cities are classified according to their demographics, transportation modes and the number of shared bikes and vehicles using hierarchical clustering and principle component analysis. An MNL model is developed using the Puget Sound Regional Council Household Travel Survey in addition to other land use and trip characteristics. After that, several Probit models were developed to test the factors affecting people's interests and concerns regarding AV technology using the same survey.

It was found that US cities fall into four clusters with the majority of US cities fell into one cluster because of their suburban planning and their auto-centric transportation systems. The Multinomial Logit model explores the mode choice in Seattle, an outstanding city and is known for its multimodality. It was found what previous literature points out to the factors affecting the mode choice, which are personal, household, trip, and land use variable. In the end, an order probit model was developed to model the factors affecting the level of interests and concerns towards Autonomous vehicles.

1. INTRODUCTION

1.1 Introduction

Transportation has been at the center of human activities since the emergence of urban communities. In recent years, the transportation industry has been experiencing a remarkable transformation as a result of the advent of new mobility technologies. Planners and policy-makers study the people's transportation preferences for many reasons including but not limited to designing policies towards more efficient, safe, and sustainable systems. In this process, spatial characteristics of the network plays a pivotal role. That is why it would be of great importance to classify different cities into relatively homogenous clusters and then focus on mode usage patterns, social perceptions of mobility options, and best applicable policies in these urban centers especially with the rapid changes that are happening in the transportation spectrum. Transportation trends are shifting in US cities and it is important to shed light on the differences of these cities to better understand people's travel behavior and preferences.

In the U.S., automobile mass production lines were first developed by Henry Ford in 1913. Since then, the low cost of auto ownership made it possible for almost any household to purchase their own vehicle and it quickly became part of the American identity. In addition to that, the introduction of the interstate highways and suburbanization in U.S. cities made automobile an integral part of the transportation network and the American identity. This in turn made officials and planners center their focus on road expansions in lieu of promoting public transit and active transportation. However, in recent decades, due to the limitations in land and budget resources together with the ever-increasing travel demand patterns, planners have shifted their focus on management of travel demand rather than accommodating more supply to meet the demand.

The transportation industry is evolving with the rise of many new business models (e.g., the sharing economy), mobility technologies (e.g., autonomous vehicles and electric vehicles), and other new mobility trends (e.g., micro-mobility including shared scooters and bikes). These changes stem from the evolution of mobility and that is evident in the emerging business models empowered by cell phone technology such as Transportation Network Companies (e.g., Uber, Lyft, etc.) and car sharing companies (e.g., Zipcar, etc.) and the advanced Artificial Intelligence technologies that makes fully automated driving possible. Many companies such as Waymo, Tesla, General Motors, and Volvo have invested significant resources to build autonomous vehicles (AVs) and several pilot projects are already conducted in different cities in the U.S. and other countries.

AV technology has the potential to reshape the transportation system once it is mass produced and reach a sufficient market penetration which relies on many factors. These factors depend on the technology itself and people's behavior towards this technology (i.e., how people perceive benefits and disadvantages of AVs and how much they are willing to pay for the technology). Some of the potential changes are averting traffic accidents, providing mobility to the elderly and people with special needs or conditions which would not allow them to drive a conventual vehicle. Also, the fact that the vehicle is now human independent can change the way we perceive transportation from a privately-owned service to an on-demand service which would affect the roads capacity, infrastructure investments, parking needs, land use and trucking.

Since public spending, policy making, and people's choice of current mobility option and perception of emerging mobility technologies are all related. Similar urban areas tend to have similar mobility patterns (dense cities focused on public transit such as New York city and Chicago versus sparse urban areas that are heavily reliant on personal vehicles such as Dallas, and Los Angeles). TNCs, and Carsharing services among many other mobility options are causing rapid changes to transportation demand. In addition to that, the emergence of AVs has created a new spectrum of complex questions that need to be answered first relating to privacy, piracy, and legal liability about AVs and people's interests, and concerns towards this technology among other emerging trends in transportation.

1.2 <u>Objectives of this study</u>

There are many new disruptive technologies (such as autonomous and connected vehicles) and business models (such as shared mobility) that are emerging besides the classical transportation means such as the private vehicles and public transit. AVs as one of the game changers in the transportation industry will be on the market, changing the role of the driver to be a passenger, and with that comes a full range of changes in safety, car ownership, and travel behavior. Several companies such as Waymo, Volvo, and Nissan have been developing AV technologies for the past 10 years and promised their commercially available AV to be in the market within the next few years. Yet no one can say for sure how the demand of this technology will be in the future. The current study is designed to shed light on this area by:

- Classifying U.S. cities according to their mobility shares;
- Developing mode choice models in one of the resulting clusters (Puget Sound Region in Seattle);
- Developing ordered models to disclose the factors affecting the level of interests and concerns for AVs.

By addressing these questions, many important policies can be derived. Also, it is important for AV manufacturers to have an empirical evidence of what are the factors affecting people's adoption behavior about this technology.

1.3 Terminology :

1.3.1 Transportation Network Company (TNC)

According to the state of Texas (House Bill [HB] 1733, 84th Regular Session, codified as new Chapter 1954, Insurance Code) a TNC is defined as “a corporation, partnership, sole proprietorship, or other entity operating in this state that uses a digital network to connect a transportation network company rider to a transportation network company driver for a prearranged ride.” In the State of Washington Transportation network companies (TNCs) include companies such as Uber and Lyft that use a digital network or software application to connect passengers to drivers to provide prearranged rides (Crombie, 2016). In general, TNC definition typically include the following elements:

- The use of a digital platform or software application.
- A prearranged ride between the rider and the passenger(s)
- The use of personal cars by transportation providers(Goodin, 2016).

1.3.2 Car Sharing:

“Carsharing is defined by its environmental and social purpose, rather than business and financial objectives. Carsharing is a service designed for local users in support of community transit and environmental goals. Its mission, vision and values lead to actions aimed at decreasing individual car ownership, reducing vehicle miles traveled, improving urban land use and development, providing affordable access to vehicles for all constituencies – including those less able to afford car ownership - as well as motivating residents to walk, cycle and take public transportation, and decreasing dependence on fossil fuels while reducing the emission of greenhouse gases.”(Car Sharing Association, 2011)

1.3.3 <u>Autonomous Vehicles:</u>

"Any vehicle equipped with driving automation technologies (as defined in SAE J3016). This term can refer to a vehicle fitted with any form of driving automation. (SAE Level 1-5) " (Transportation, 2018)

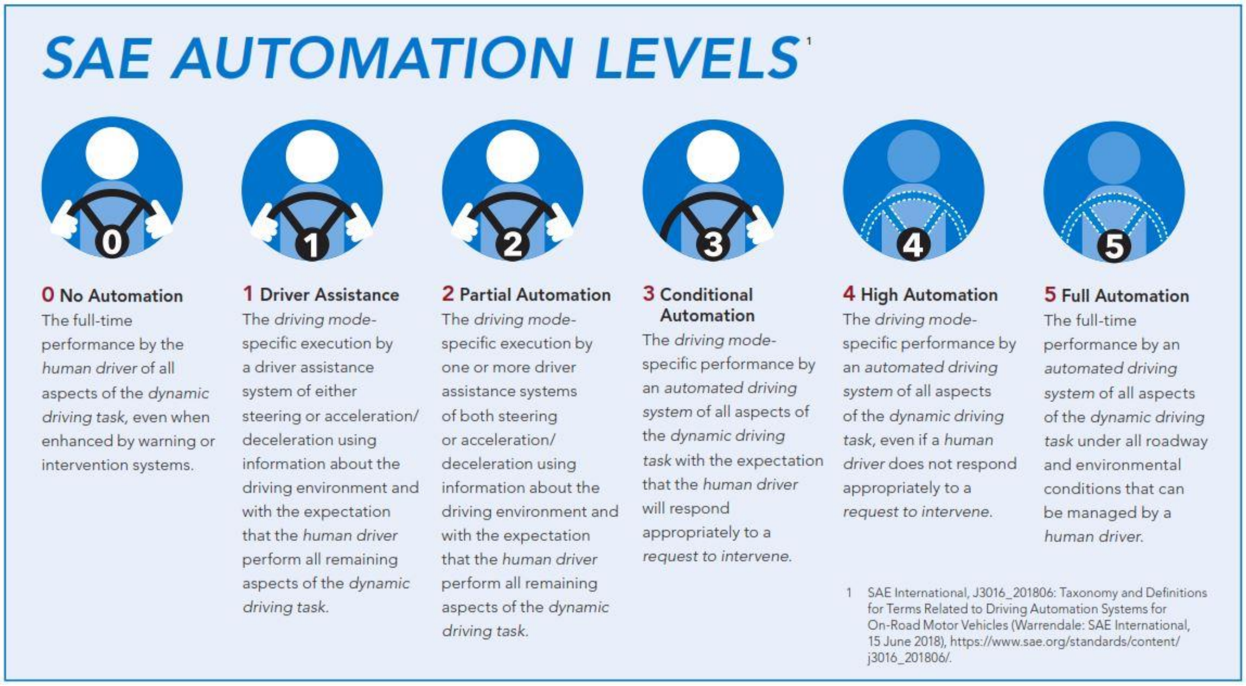

Figure 1: Society of Automobile Engineers Automation Levels (Transportation, 2018)

Currently there are up to level 4 automation in the market ("Driverless Cars: Where They Stand Now | Automobile Magazine - Automobile," 2019) but what is meant by autonomous vehicle in this thesis is level 5 Fully automated Autonomous Vehicles where no human input is required.

1.4 <u>Possible effects of AV Technology</u>

There are many perceived benefits to the transportation system because of the differences between AVs and human-driven vehicles. Some of these differences include the possibility of driving unlicensed “drivers”, being programed not to break the law and the immediate response to traffic impulses. However, the extent of these benefits are not precisely known because of the limited information on the potential market penetration of AVs (Fagnant & Kockelman, 2015).

Safety is sure to increase because of eliminating human error factor. This would allow of 40% fatal traffic accidents reduction in the United States (Administration, 2008). Assuming that technical errors and malfunctions are reduced to minimum, almost 90% of crashes can be prevented (Administration, 2008). Human error is an integral part of any accident even if the cause of crash is attributed to external factors such as the vehicle, roadway, or environment because the inattention or distraction of humans contribute to the crash or the severity of the resulted damage. The potential benefit is vast. With statistics showing over 30 thousands persons die annually in the US due to automobile collisions (USDOT, 2012), 2.2 million of which results in injury (Highway Traffic Safety Administration & Department of Transportation, 2013). In 2010, the associated cost with traffic accidents in deaths and crashes totaled 242 $billion(Blincoe, Miller, Zaloshnja, & Lawrence, 2014). All of these losses could be alleviated if AV technology was to be fully deployed and reach a considerable market penetration.

Researchers are looking into reducing congestion and fuel consumption of AVs by harmonizing vehicles’ acceleration and baking patterns. This is anticipated to allow for smoother braking and speed adjustments resulting into fuel savings, less brake wear and reduction in traffic-destabilization shockwave propagation. Adding to that, it is expected that AVs will be more efficient since it is expected that AVs will be able to navigate with shorter gaps between the vehicles, be able to form platoons and would make more efficient route choices. Adaptive Cruise Control (ACC) is among these features that is currently being integrated into automobiles and drivers are already realizing the many benefits that AVs could bring to mobility although they're not fully developed, yet. Many of the benefits of AVs will not depend on the AVs themselves rather on the cooperative abilities through V2V (Vehicle to Vehicle) communications. This communication technology is assumingly to be the standard especially since the NHTSA intends to mandate V2X (Vehicle to Everything) capabilities with all light duty vehicles (National Highway Traffic Safety Adminstration, 2014) and V2I (Vehicle to Infrastructure) communication. Congestion would be reduced if the safety benefits of AVs are realized even without V2X communication. According to FHWA estimates, traffic accidents cause 25% of congestion, half of which is estimated to be crashes (Systematics, 2004).

There will be notable impacts to travel behavior because of AVs' potential impacts such as increased safety and reduced congestion for example, people too young to drive, the elderly and the disabled will be able to use this technology and thus generating new transportation demand. Parking demand also could change as AVs can self-park in less expensive areas. In general, higher VMT will probably be the result of these changes. This will create many problems in the transportation network such as higher emission rates, greater gasoline consumption, oil dependence and higher obesity rates (Fagnant & Kockelman, 2015).

It is assumed that VMT per AV will be higher than that of non-AV depending on AVs' market penetration rate (i.e., 20% higher at 10% market penetration, and 10% higher at 90% market penetration rate). This indicates a suppressed demand for early adapters. Fagnant and Kockelman's preliminary agent-based modeling simulation (Fagnant, Kockelman, & Bansal, 2015) highlight this idea. A fleet of shared autonomous vehicles (SAVs) who makes 56,000 trips a day, travels 8.7% of its mileage unoccupied, and up to 4.5% when ride sharing was allowed. It was observed that when demand rose by a factor of 5 and ridesharing was permitted, less than 1% net VMT reductions were realized. The analysis of different simulation scenarios proposed that each SAV could serve the travel demand of 10 household owned private vehicles (if all the trips were set to take place a geofence of 12 -24 mile radius).

1.5 <u>Potential Barriers to AV Technology:</u>

Although AV technology can present many opportunities and benefit, there are many challenges to this industry. Still, no one can accurately guarantee the speed and nature of such transition, since such transitions will likely depend on the product properties such as its cost. In addition to state and federal licensing requirements and the security, and privacy concerns associated with AVs. Special research should be put to study these risks in order to full deployment of AVs into transportation systems. The following discussion outlines several barriers that AV technology currently faces.

One barrier to full market penetration of AV technology is the cost of AV platforms. New software and hardware are required for each vehicle. Some hardware devices like the Light Detection and Ranging (LIDAR) systems can have a varying cost between (30,000$-85,000$) in addition to the costs of other equipment necessary for the operations of Autonomous Vehicles (Nick Shchetko, 2014). In order for AVs to be affordable in the future they need not to have a LIDAR system or LIDAR's price should fall dramatically (Nick Shchetko, 2014). In general, research suggests that Americans will be more willing to buy an AV if its price falls to what a conventional vehicle's cost. J.D Power and associates survey (J.D. Power and Associates, 2012) found that the percentage of people who responded that they would "Definitely" or "Probably" purchase an automated vehicle in their next purchase has fallen from 37% to 20% after assuming an increase of additional 3000$ purchase price of

AVs estimated by Volvo senior engineer Erik Coelingh (Economist, 2012). Early-sales' costs will likely be much higher for early adopters.

After overcoming AV certification by state agencies that are in charge of this process come other issues regarding insurance and liability. One of the issues is persuading insurance providers of the technology's ability to function properly in all driving environments. Even if autonomous driving technology was perfected, there are still many scenarios in which crash is unavoidable. For example in case a deer or a human has jumped on the car while driving and the different decisions that needs to be taken in such cases. Different issues arise with the emergence of AVs. First, there is the question of who is accountable for a crash since the car is equipped with sensors, cameras and algorithms. In addition to that there will always the philosophical questions of prioritizing injury of the passengers versus avoiding crash with other obstacles during the trip, and whether the owners should be allowed to make adjustments to AVs or not. Implementation of AVs could be delayed because of the increased cost of AVs. This increase of cost is attributed to the high standards that are expected from AVs to keep safety level the highest.

Cybersecurity remains one of the top concerns of transportation policymakers, auto manufacturers, and future AV drivers. There are many groups that may target AVs and intelligent transportation systems such as computer hackers and terrorist organizations. These attacks might cause collateral damages to the safety of the passengers and to the functionality of the transportation network. Since Each AV represent an access point to the whole system, it might be infeasible to secure the entire system of national transportation. Hickey (2012) the vice president of a software security firm Venezuela has stated that espionage (i.e., to break into a system to gather information) is a more common goal to current cyber-attacks than sabotage(i.e., to actively disrupt a system's common operations).

Sabotage operations would require designing a more complex attack than it would be to gather information. There is also the threat of a security breach, although attacking a complicated system such as AV is a great challenge

The U.S has a demonstrated ability in maintaining and securing large critical national infrastructure systems (e.g., power grids, and air traffic control systems). The frameworks and recommendations developed by The National Institute of Standards and Technology (NIST) are to improve critical infrastructure cyber security that will be implemented in the protocols of the connected and autonomous vehicles.V2V an and V2I protocols have been developed with security implemented in the initial development phase (National Highway Traffic Safety Administration, 2011). All of this should make it more difficult for an attack to be pulled off while also limiting the damage that can be done.

Consumer Watchdog, which is a “California-based consumer education and advocacy organization” has previously raised privacy concerns during a round of AV-enabling legislation (Brandon, 2012).With data sharing at the center of Autonomous and Connected vehicles such concerns are likely to grow. As (Fagnant, 2014) put it: “This gives rise to five questions about data: Who should own or control the vehicle’s data? What types of data will be stored? With whom will these data sets be shared? In what ways will such data be made available? And for what ends will they be used?” As for the crash data, it will probably be available to AVs manufacturers and suppliers since they will likely be accountable for damages caused by the crash. However, privacy concerns arise because no one wants their data to be used against them in court. During this time of the US data ownership and control remains undefined (Martin Kaste, 2013). There are concerns about providing travel data to a centralized and governmental controlled system and the possibility of using such data to monitor individuals. Although some type of monitoring is already occurring through roadside Bluetooth sensors and cell phone towers triangulation, continually monitoring travel data using AVs could bring this phenomenon to a whole new level. On the other hand, these datasets would help traffic engineers and designers in planning better transportation networks and potentially switch from gas tax to Vehicle Miles Traveled, where the owner of the vehicle is taxed based on the number of miles traveled rather than on the gas consumed. It will also help pull up better congestion pricing schemes and that are related to the area and the time of the day. In general, responsible distribution of AV data to all stake holders could provide tremendous benefits to the riders and to the transportation network and mobility now and in the future.

In this thesis, first, a cluster analysis of 50 U.S. cities is executed based on mobility shares of different modes and other variables related to each city. After that, a Multinomial Logit Model (MNL) mode choice model is developed for the Puget Sound Region in Seattle using Biogeme (Bierlaire, 2018) based on the Puget Sound Regional Council (PSRC) 2017 Household Travel Survey. Different attributes for non-chosen modes during each trip were derived using Google API. Land-use variables were also incorporated from the Environmental Protection Agency. After that, multiple Ordered Probit Models were developed using NLOGIT5 to test the effect of personal and household attributes as well as land-use variables on the level of interest and concern towards Autonomous Vehicles (AVs).

2. LITERATURE REVIEW

Evolution of cities and transportation systems are among the topics that have sparked the interests of many scientists and engineers. As illustrated in Figure 2, it is estimated that 68% of the world population will be residing in cities by 2050 (UNDESA, 2018). With this wave of urbanization comes a whole range or challenges that need to be studied and are related to livability, mobility, and sustainability in the cities that we live in. Roughly 90% of US population will reside in cities and thus there should be a focus on classifying and examining urban centers and mobility preferences especially when we are witnessing a transition with the emergence of new transportation technologies.

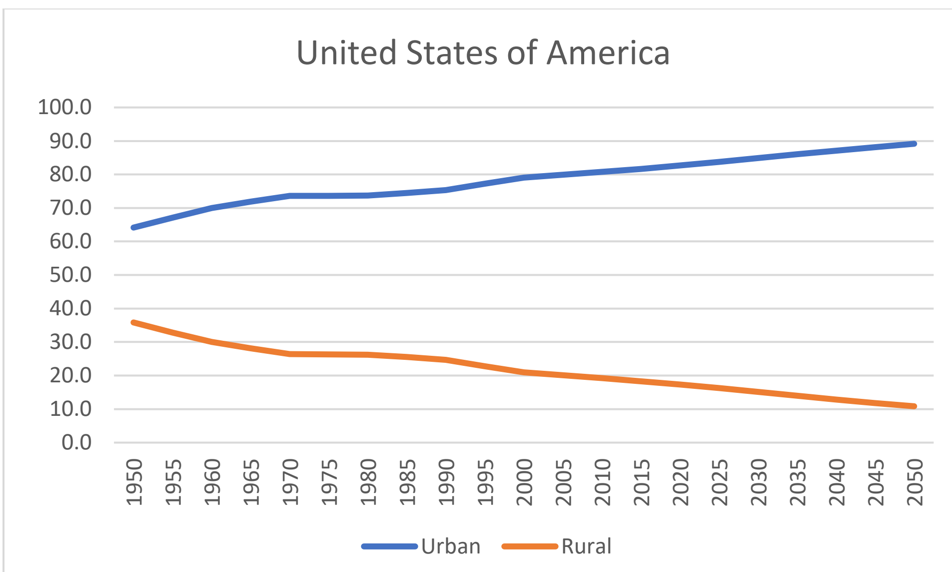

Figure 2: Percentage of Rural and Urban Population living in the United States of America (United Nations, 2018)

(Alonso, Monzón, & Cascajo, 2015) Clustered 18 European cities according to three indicators that they developed from datasets mainly obtained from The Metropolitan Mobility Observatory(MMO, 2019), which is a platform comprised of 24 Public Transport Authority (PTA) in the main Spanish cities. The European Metropolitan Transport Authorities (EMTA, 2019). The indicators constructed were the three pillars of Sustainability- Economic, Social, and Environmental- agreed upon by the United Nations (UN, 2019). The factors that constructed these three indicators were chosen based on previous literature and on definition of these three pillars (e.g., Coverage ratio of public transport represented one of the Economic dimensions based on (Newman & Kenworth, 1999) and its sustainability aspect is efficiency in operation). The composed indices(CIs) (sustainability measures in each city) are based on the comparison among the selected cities (i.e., the value for each city depends on the average performance of the rest) After that, Pearson correlation coefficients were analyzed between sustainability scores of cities and their characteristics (e.g., size, wealth, modal share, etc.) to identify the characteristics that contribute the most to achieving sustainable urban transport. Finally, the cities are clustered based on their Composite Indicators as classification variables. Four clusters appeared (environmentally efficient, socially friendly, economically competitive, and least sustainable). In conclusion, the richest and largest cities are usually the most sustainable in terms of transport systems. CIs can also refer to certain transport policies that could improve sustainable transport deficiencies in other cities, such as increasing the share of public transport and avoiding urban sprawl.

(Kyung, 2010) Clustered cities in Seoul according to their characteristics (e.g., topology of metropolitan cities, population, land, etc.) using principle component analysis (PCA). In the end, the metropolitan areas fell into five clusters. Urban characteristics spread out in the form of concentric circles around the city center of Seoul, where industry is concentrated. This study aimed to help metropolitan regions to understand where it stands from other Korean cities according to its characteristics in a way that it could be used to establish policies and plans for these cities.

In general, modal split is established to be one of the important characteristics in city classifications since modal split is a product of the characteristics of cities. For example, (Toop & Miller, 2014) devised a framework to classify Canadian cities. He concluded that midsize cities in Canada are uniquely “automobile-centric”, and since more than one third of Canadians live in such cities, it’s important to develop transportation management strategies among other policy tools that are tailored for such cities. If classifying cities is important and modal split plays such a crucial role in defining them, it’s as important to develop transportation models for these cities.

(Frank S. Koppelman; Chandra Bhat, 2006) published a self-instructing course on mode choice modeling. After explaining the theory and mathematics that lie at the core of MNL modeling, they began developing this model on the San Francisco Bay Area. They developed two models, one for work trips and the second for shop/other mode choice. After that, a nested Logit model (NL) was developed to circumnavigate the problem of the independence of irrelevant alternative (IIA). In their exhaustive course many models where developed, discussed, and improved upon and several examples for probability estimation for different scenarios were also given.

(Badoe, 2002) proposed that in urban areas households with more than one worker should be modeled as one unit instead of modeling each person as one decision maker. She argued that that is a sound reasoning for at least two reasons. First, the survey is done on a household level, and secondly, most household in urban areas have more than one worker that would in the long run choose one mode that would lower the aggregate disutility of the household. She developed two MNL models on the individual and the household levels, for household that has 2 individuals living in them from the 1986 Transportation Tomorrow’s Survey (TTS) that was conducted in the Greater Toronto Area, Ontario, Canada. The choice set for the model is based on two modes, auto-drive mode and transit all way mode including various Level of Service variables Socio-economic variables and alternative specific variables that were either available or generated using EMME/2. The results were consistent with expectations, that is the variables indicating the Level of Service (e.g., travel time, travel cost, etc.) had a negative sign while the socioeconomics variables had a positive sign. The magnitude was also consistent with expectation (e.g., in-vehicle travel time had consistently less impact than the access time or wait time for transit service) for both the individual and household models with comparable values of these variables. However, the model fit for the household model was better using different metrics (comparing the direct loglikelihood, comparing the likelihood based on marginal probabilities, and finally, an aggregate prediction test on the estimation data).

While numerous studies focuses on the working age population, (Kim & Ulfarsson, 2004) computed a multinomial logit model for travel mode choice for the elderly using the Puget Sound Regional Council household travel survey. They studied the effect of personal attribute (sex, age, availability of a driver's license, etc.), household characteristics (size, income, etc.), neighborhood features (population density, gross rent in a block group, etc.), trip characteristics (number of trips per day, time of the trip, etc.), and the activity purpose (family, errands, shopping, etc.). The modes that were taken into consideration include Private Car/ Truck, Car-/Vanpool, Bus, Paratransit, and Walk. He found that analogous to previous literature, automobile is the dominant mode choice. Walking is a preferable option to public transit as a travel mode. Seniors made 90% of their trips using personal vehicles although almost a third of them had a transit pass. And although older people may drive not out of necessity but to express personal freedom (Benekohal, Michaels, Shim, & Resende, 1994), the model confirms that as they age they become more dependent on public transportation and walking. The least favorable mode is transit and specifically the bus. Public transit routes and schedules are designed for the working aged population that are often inadequate for senior citizens (Carp, 1988). The retired elderly makes most of the trips during off-peak periods where the waiting time is maximum. The estimation results shows the possibility of some measures to provide incentives to increase the share mode for public transit. (Fujii & Kitamura, 2003) found that a one-month free bus pass increased bus use while decreasing the automobile use at least in a short-term period. The elderly who have been living in their current residence for more than 4 years are more likely to use carpools or vanpools, one explanation can be that the social network formed in the neighborhood over the years influences the travel mode. Even though the primary support for providing travel needs for seniors is a task for family members, neighboring friends participating in the same activities provide transportation or those who don't drive. (Burkhardt, 1999). The topic of the residential environment of the neighborhood over the individual travel choices is disputed in the transportation domain. However, the effect of aggregate behavior of a group of individuals or households (i.e., their lifestyle) on the behavior of an individual or household within the same neighborhood or adjacent neighborhood (i.e., the neighborhood effect) has been widely studied in the social science literature (Haurin, Dietz, & Weinberg, 2005). Population density had a negative correlation with single-occupant automobile use. Accessibility to activity centers and the availability of the transit are the main cause for that. The estimation results show that for seniors living in a neighborhood with high residential relocation rates, using public transit s less likely. Also, the workers' driving population increases the tendency of passenger car use when income and population density are controlled for. The crowdedness of public transport in urban dense areas may discourage elderly of opting for that mode of transportation. An indicative of the neighborhood effect is the positive correlation of the percentage of workers in the neighborhood who commute using a private vehicle on elderlies using the same mode when other neighborhood characteristics (e.g., densities.) are controlled.

Finally, (Irfan, Khurshid, Khurshid, Ali, & Khattak, 2018) used a Multinomial Logit Model for stated and revealed preferences about work-trip mode choice to study the effect of simultaneously introducing a new mode of transportation (i.e., Bus Rapid Transit (BRT)) and a congestion pricing policy in order to alleviate the congestion. The collected data was through a questionnaire to residents living in a 2 km buffer zone from the proposed BRT so that they are realistically able to access the new mode. A test of elasticities and change of probabilities when introducing the new suggested BRT and congestion pricing was calculated. Most of the variables were inelastic. Public transport, however, was elastic to both in-vehicle travel time and out-of-vehicle travel time with out-of-vehicle travel time having the heavier effect. This means that increasing the frequency of buses would lead to more use of BRT system. Also, the toll price was elastic in the car mode. Cross elasticity of car toll for BRT was comparatively higher than that with other modes which translates to increase in demand for BRT mode of transportation.

While focusing on people's preferences towards various travel modes are important, it would be also essential to account for new mobility technologies such as autonomous and connected vehicles and shared mobility. Such emerging technologies are considered as a new trend in the transportation spectrum and an increasing number of papers and reports have been published on examining people's stated preferences for adopting such technologies. The analysis methods are broadly based on descriptive analysis or behavioral analysis of people's response to these technologies.

A famous study examining the public opinions of AVs in the U.S, U.K, and Australia (Schoettle, Brandon; Sivak, 2014) was executed through a public survey done in three countries to account for the technological progress in AV technology and to address the concerns about such technology. The survey was designed to ask the respondents about the perceived benefits (with answers ranking: very likely, somewhat likely, somewhat unlikely, and very unlikely) and also their concerns about AV technology with answers (very concerned, moderately concerned, slightly concerned, and not at all concerned). The respondents were also asked about their concerns for different scenarios with similar answers (very concerned, moderately concerned, slightly concerned, and not at all concerned). For example: “Riding a Vehicle with no driver controls available” or “Self-Driving vehicles moving by themselves from one location to another while unoccupied” etc. Another important question concerned the willingness to pay for such technology. Overall, the results for U.S respondents indicates a prior familiarity with the concept of self-driving vehicles and were more likely to have a “very positive” of this technology, but they were also more likely to be “very concerned” than their foreign counterparts about the common concerns that comes with this technology (e.g., legal liability, data privacy, system performance and etc.).In general, more than one third of the sample (35.9%) responded “very concerned” about riding in Level 4 vehicles. In all countries, respondents have reported high concerns about Level 3 and Level 4 automation with greater concern about Level 4 full automation than Level 3, although there is a greater potential safety-risk during the transition if necessary, back to human drivers in Level 3 automation. However, there were similar percentages in concerns about riding self-driving vehicles with 87.3% being “very/moderately/slightly/ concerned” about Level 3 and 87.9% giving similar responses about Level 4 automation. Answers to the question” how extra time would be spent” reveal concerns towards this technology since 41% of the respondents said they would watch the road while 22.4% said that they would prefer not to ride in Level 4 self-driving vehicle. In the US the concern was greatest (92.8%) expressed their concerns that self-driving vehicles would not drive as well as human drivers. In addition to that, there was a higher concern with the highest frequency in the U.S (95.6%) of people about self-driving vehicles’ confusion in unexpected situations. This paper shed lights on the general perception and concerns about the AV technology. It has been found that there was a significant level of familiarity with AV technology with positive initial opinions, high expectations of the benefits but highly concerned as well about actually using this technology. That feeling was generated by perceived issues in security, traffic safety, and relative performance of the self-driving vehicles. Responded has also expressed their concerns about vehicles without driver control, unoccupied vehicle that are moving within the system, and self-driving commercial vehicles, busses and taxis. Most respondents expressed their desire in owning the AV technology while most of them were unwilling to pay extra for this technology. Females were comparatively more concerned than males regarding the utilization and benefits of AVs. Correspondents from the U.S expressed the greatest concerns among the three countries.

(Payre, Cestac, & Delhomme, 2014) focused on Fully Automated Driving (FAD) with a questionnaire answered by French drivers with their stated preferences on FAD on highways, in traffic congestion and for automatic parking. A priori acceptability refers to an evaluation of a technology before having any interactions with it. The survey included 421 drivers and was constituted of six parts. The first part was a description of a fully automated car, and information about demographics. In the second part, the criteria associated with FAD were explained to participants. They have been informed of their responsibility for the driving and they were presented with the first set of the questionnaire. The scale presented for the different choices was 1 referring to “I don’t agree at all” to 7 referring to “I totally agree”. In the third part seven questions were presented. The objective was to measure a priori acceptability. The fourth section contained the driving internality and driving externality. The fifth section was filled with questions about the intention to use automated driving, the willingness to pay for such technology, and road and environment preferences and finally the participants’’ attitude towards automated driving (i.e., whether it’s pleasant or unpleasant or whether it’s useful or useless). The three items in attitude were merged into one dimension. In the sixth section, they filled the DRSS. In the last section, they answered socio-demographic questions: gender, age, year of driving license acquirement, kilometers driven last week, and extra hours estimated to learn automated driving. After acquiring the data from the survey, a descriptive analysis and hierarchical linear regression was used to get a measure of acceptability. In general there was a positive attitudes towards the technology although participants haven't used such car with variations in acceptance according to the environment.(Payre et al., 2014)

(Bansal, Kockelman, & Singh, 2016) have assessed public opinions of and interests in new vehicle technologies in Austin Texas. They examined the willingness to pay for full automation (level4) that is much higher than partial automation (level 3). Their paper develops an ordered Probit model to estimate the impact of demographic, built environment and travel characteristics on willingness to pay. The main motivation for this study was the cost of crashes. The methodology that was used consisted of surveying a group of people that can be random, transportation experts, tech savvy, or belonging to some specific demographic-socioeconomics class. The participants are then introduced to the concepts being studied first and then the questionnaire has different questions about socioeconomics, current trends or travel behaviors the respondents have. After that, questionnaire presented to the participants about the proposed mode (Autonomous Vehicles (AVs), Shared Autonomous Vehicles (SAVs), and Shared Autonomous Vehicles with Dynamic ridesharing (DRS). The questions examine things such as willingness to pay, the diffusion of innovation, adaptation behavior etc. Their finding was that the average willingness to pay (WTP) for level 4 AVs was more than twice as higher for Level 3 (7,253$ vs. 3,300$) with more than 80% interested in Level 4 AV ownership. They have also discovered that apparently for half of the population, the adoption rate was dependent on that of their social networks while the idea of SAVs was not appealing to 80% of them. If the cost was under 100$ for adding a connectivity to the car, 75% of the respondents expressed their interest in adding connectivity to their current vehicles. The biggest concern related to AVs was equipment and system failure while adaptation to a smart vehicle seemed to be of the least concern. The highest perceived benefits by the respondents was fewer crashes while less congestion was the least likely benefit for AVs. While riding an AV, the highest picked activities were looking out the window and talking with friends. More interest in a higher WTP was associated with high-income tech savvy males living in urban areas who had previous crash experience with less dependence on friends' adoption rates. It's possible that these individuals appreciate and evaluate the safety benefits of smart technologies. Respondents also showed an interest to moving closer to central Austin where they can reap higher benefits of SAVs because of higher density in that area while licensed seniors expressed less interests in this technology driven by their concerns about having to learn how to use CAVs and SAVs. Driving population was found to be more adaptable towards using AVs with less dependence on their friends' adoption rate and a higher WTP for level 4 automation and connectivity than level 3 or using SAVs. Costing 3$ per mile. It's likely that long distance traveler can enjoy the safety benefit of Level 4 automation while also being capable of doing other work in route (such as reading, working or talking to friends) while this is not possible with Level 3 automation. As have prior studies found (e.g.,(Celsor & Millard-Ball, 2007)), SAVs utilization is more interesting to people living in denser urban areas. Reasons for that might be the inconvenient parking facilities, and lower vehicle ownership rates (Bansal & Kockelman, 2017).

(Krueger, Rashidi, & Rose, 2016) Looked at stated preferences survey of Shared Autonomous Vehicles (SAVs) by mean of an online survey completed by 345 residents of major metropolitan areas of Australia collecting their characteristics, socioeconomics and then a stated preference survey about SAV. The survey was comprised of three stages. The first stage was to specify a reference trip made by the person. The second stage was presenting the instruction and the last was to state the preferences for five tasks. They were required to choose one out of three mobility choices, two of which were hypothetical. The choices were Shared Autonomous Vehicles, Shared Autonomous Vehicles with ridesharing and the last option was public transit only. The attributes of the hypothetical modes included travel cost, time and waiting time. The value of these attributes was estimated. The results obtained from the data obtained specifies the coefficient of each variable, also the conclusion of the importance of Value of Time (VOT) and also the results about modal changes (e.g., a driver would switch to SAV with DRS while a passenger would opt for DRS.A person who uses Public Transit would switch to SAV without DRS) Also, it’s unlikely that people would choose SAV for medical or dental appointment (Krueger et al., 2016) .

(J. Zmud, Sener, & Wagner, 2016) shed light on the acceptance and the intent to use Autonomous Vehicles technology by trying to answer several questions about the effect of self-driving cars to travel demand(i.e., the likelihood that people will use this technology, the factors that affect the acceptance rate and intent to use and also what is the appeal of this technology to people) and to transportation network (i.e., the effect on traffic and congestion in the future). The data gathering consisted of a two-step process. First, an online survey of 556 Austin Metropolitan area residents were conducted followed by interviews with 44 participants were there was probing to the specific reasons, and drivers of people’s choices. The intent to use is defined as The population was split in half with 50% intent to use the technology (36% Enthusiasts (Extremely Likely) 14% Pragmatists (Somewhat Likely)), and the other 50% with no intent to use (18% Rejecters (Extremely Unlikely) and 32% Traditionalists (Somewhat Likely)).

The reasons that surfaced during the interviews fell under seven categories:

1. Safer than Human Drivers.
2. Stress Relief during the trip.
3. Mobility enabler for senior citizens.
4. Frees up the focus and time to do other productive tasks during the trip.
5. The trust in the testing of this technology.
6. Comparability to Public transit experience.

7. Attraction of new technology.

The demographic variables (e.g., age) for the population were not strongly related to the intent to use. Younger people (less than 30 years old) as well as people older than 65 years old were evenly split with their intent to use. In the cohort (30-45 years old), the percentage were slightly skewed towards the intent to use (53% likely), whereas (55%) of persons 46-65 years old were unlikely to use. All the people (n=11) with a revealed-restrictive disability was likely to use. There was a difference in the intent to use gender wise. Males were more likely to use than females, and the percentage of male enthusiasts (18%) was higher than females (11%). As for the household income, those of lower household income (less than 25,000$) were unlikely to use (56%), on the contrary, household with incomes (25,000$-50,000$) were more likely to use (54%). Educational level of the survey participants was not associated with the intent to use unlike having children in the household. Although more than one third of the 20 sampled households with more than two children were enthusiasts, having children in the household was indicative of less likelihood of intent to use AVs than households without children (51 and 45%, respectively).

Reasons with the most frequency for lowering the likelihood of using a self-driving vehicle for everyday use were:

- Lack of trust in the technology (41%).
- Safety (24%).
- Cost (22%).

These three reasons encompassed the majority of the reasons of deterring from intending to use the self-driving vehicles. There were also other individualistic related to personality reasons such as the liking of driving, or the desire for vehicle control. After using the Car Technology Acceptance Model (CTAM) the researchers tested different variables related to socioeconomics and demographic and after isolating the significant variables a regression was implemented and the results obtained indicates that there are several groups of people with higher likelihood of using self-driving vehicles such as people with driving prohibitive physical condition, the people who believe that using a self-driving vehicle would reduce crashes on the transportation network, tech-savvy individuals (i.e., people who use smartphones, text messaging, social media and transportation applications), individuals who are not that concerned with data privacy, also other things were an indication of using self-driving vehicles like the belief that it would be fun to use, think being skillful at it would be important and finally people who are surrounded by others that are likely to be using self-driving vehicles. First this study analyzed the effect of self-driving vehicles on travel behavior (i.e., The effect on auto ownership, the tendency to use shared vehicles instead of buying a private car, choices of mode of travel, and the effect on the amount of VMT). It was observed that the factors that affect the intent to use self-driving vehicles fall under seven categories:

1. Relative higher safety compared to the human counterpart.
2. Stress relief during the trip.
3. The trust in the adequacy of testing of such devices.
4. Being comparable to transit experience
5. Appeal of new technology.
6. Mobility enablers for the elderly senior.
7. Productivity enabler during the trip.

There were also many concerns about it but most of the concerns were about the technology adequacy. Safety, and data privacy were also mentioned as concerns towards using the self-driving vehicles. The main factor affecting people's intent to own was convenience, but there are also other factors such as liability, legislation, vehicle size, brand, and most importantly the cost of this vehicles. There was a question about the willingness to pay for this technology compared to the average price for a new car in 2014 (32,000$) and by far the most frequent answer was "a slight amount" compared to "zero" or "a great amount." It should be noted that categorical option were used instead of discrete numbers because the technology itself is new and not yet available in the market.

As for people's interest in Shared Self-Driving Vehicles, many of residents of Austin speculated that this will be the primary business model for self-driving vehicles. However, there was more interest in owning a self-driving vehicle than in using a shared self-driving vehicle (e.g., Zipcar, Car2go, or Uber taxi) (59% and 41% respectively). The variables didn't seem to have any effect are age and income. people's interest in car sharing. There were categories behind people's rationale: gaining experience and cost. The willingness to pay also was examined in the survey and the choice "slight amount" was the most frequent in comparison to "zero" or "great amount" when people were asked about their willingness to pay for such service above 10$ per hour.

There were questions about the effect of owning self-driving vehicles on household auto ownership. 61% reported that owning a self-driving vehicle would have no effect on household number of vehicles excluding the possibility of sharing the same vehicle with other household members, whereas 23% indicated they would reduce the number of vehicles owned. 16% said that the number of vehicles would increase.

The impact the self-driving vehicles would have on VMT is still unclear. Some speculate that VMT per capita would increase since people can live in places farther from their work or study locations since they will free the time used while driving the car to get to these places. Also, the car might end up making zero occupant trips around the city to drive different members of the same household which would increase the VMT. Others think that the VMT would decrease thanks to the car-sharing programs as has been found for conventional car sharing programs. (Handy & Boarnet, 2014) . An online survey was conducted that concluded that 16% of the respondents indicated that they have driven less than 5,000 miles in 2014, 35% indicated driving (5,000-10,000), 35% reported driving (10,000-15,000) miles and 15% reported driving more than 15,000 miles in 2014. Almost two thirds of the respondents to the survey and to the interviews (66%) said that their VMT would not change even if they obtained a self-driving vehicle. Some speculated an increase in their VMT since they would be taking more leisure trips using their vehicle. The few that reported expected lower VMT attributed it to the efficiency of car sharing (J. P. Zmud & Sener, 2017).

The participants were also asked about the impact of self-driving vehicles to where they live. 80% said that they wouldn't change their residence even if they had a self-driving vehicle.2 respondents pointed out that they would move farther for lower housing prices while two pointed out that they would move closer to get the full benefits of car sharing programs. The 44 respondents were also asked about the impact self-driving vehicles could make to long distance travel. 57% said that they make occasion inter-city trips in Texas, and 43% said that they frequently make them.45% said that they would change the mode of travel if self-driving vehicles were available today but not the frequency of their trips. Also, 42% would change the mode for out of Texas trips while others said the fuel efficiency is a key factor to that decision.

3.CHARACTERIZING CITIES ACCORDING TO MODE SHARES

3.1 <u>Introduction:</u>

There has been a wide increase in shared mobility services in Cities around the world and in the U.S (movmi, 2019).For instance, Seattle’s pilot project with Bike-sharing that has started in 2017 (Lioyd, 2018) other examples include free float carsharing, several ride hailing programs and other mobility options such as electric scooters. 52 cities across the US are classified in this chapter. Their population, number of shared bikes and cars, land area and commute modal splits of each city is used. The assumption is that through Hierarchical cluster analysis method, cities that share similar characteristics would fall into similar clusters. Thus, cities can see where they stand and how they can get to an inclusive transportation system beyond private vehicle ownership following the lead of other U.S cities that at a more advanced stage in this process. Figure 1 depicts the cities that were in this study.

Figure 3: US Cities Analyzed

There are regional differences among cities in the U.S that affect the transportation trends in these cities. Dallas, Texas is an example of automotive-centric long-term planning. Dallas is more suburban than urban, in addition to having one of the cheapest gasoline price in the United States (EIA, 2019). These facts made it easier and more efficient for commuters to use their private vehicles instead of using the public transit service operating in Dallas (Dallas Area Rapid Transit). This is shown in the relative high number of vehicle ownership per household in Dallas 1.59 relative to other denser cities such as Chicago (1.12) or New York City (0.63) (Governing.com, 2016). Cities like Dallas need to follow the footsteps of more multimodal oriented cities that are less dependent on private vehicle ownership in order to develop a more sustainable transportation system.

San Francisco is another notable example. The presence of Silicon Valley near the city made it a field of experiments for all the tech start-ups that want to enter into the domain of data-driven transportation companies enabled by the availability of smartphones and GPS technology such as Uber, bike-sharing companies carsharing companies and more. The funding and operation of public transportation is also significant with the Bay Area Rapid Transit (BART) and the Municipal Rail (MUNI) and Caltrans that connects the city to its suburban areas and the Silicon Valley. All of this in addition to the smaller area of San Francisco compared to Dallas made the percentage of transit riders much higher than it is in Dallas. The following chart is derived from the data used in the analysis that highlights the differences in mode split among a sample of the cities that were under the analysis.

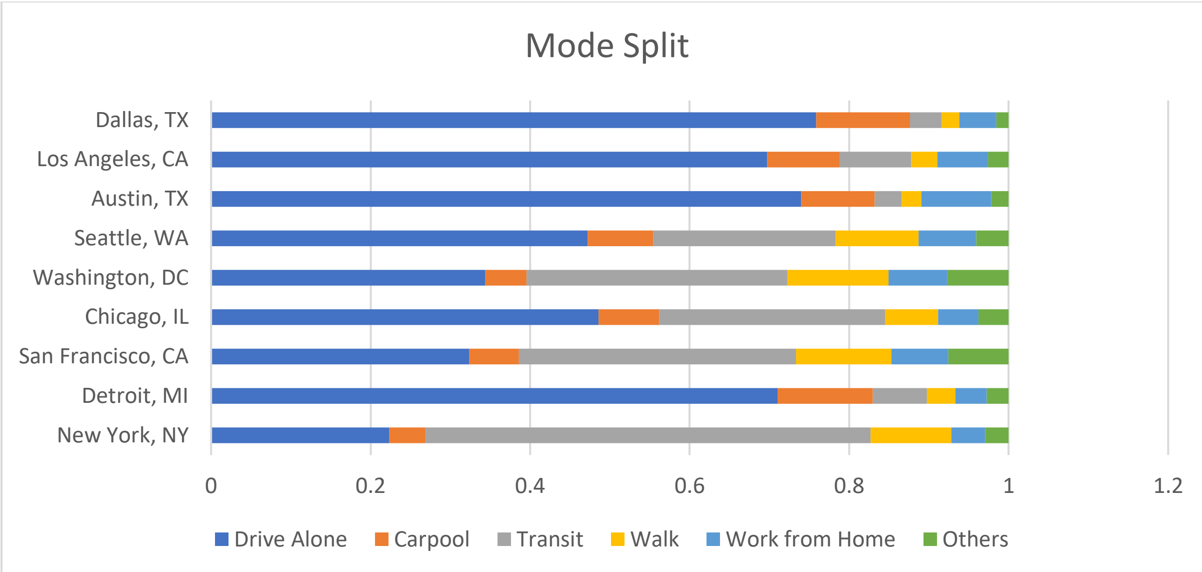

Fig 4: Mode Split

3.2 Methodology:

3.2.1 Data Preparation:

The data used for this analysis comes from two sources:

The first source is the Shared-Use Mobility Center (SUMC) Mobility on Demand (MOD) Learning Center. It aggregates the number of bike-sharing and carsharing vehicles alongside American Community Survey Commuting mode-split data for cities that represent the center of a Core-Based Statistical Areas (CBSAs), which are defined by the Office of Management and Budget as “geographic locations neighboring urban areas of at least 10,000 people and/or are socioeconomically tied to the urban center by commuting. “

The second source is the American Community Survey (ACS (2012-2016)) census data. This data was chosen because it is the same dataset that is used at SUMC Mobility on Demand Learning Center. This dataset contains the percentages of each mode used in commuting in cities across the U.S, also the total number of vehicles and households of each city analyzed. A Vehicle per Household variable was thus developed.

All the variables in this analysis were normalized so that each variable varies between 0 and 1.

3.2.2 Techniques Used:

I- Principle Component Analysis (PCA):

PCA is a dimensionality reduction technique. It has many uses in Statistics and machine Learning especially for datasets that have a linear nature where it's used for compression, or redundancy removal. It is best known for face recognition (Kaur & Himanshi, 2015).

Principle Components can be obtained from any dataset with a linear nature. After obtaining the dataset, the average is subtracted from each dimension. After that the Covariance matrix is calculated. The step that follows is to calculate the eigenvectors and the eigenvalues of the covariance matrix. From this the notion of dimensionality reduction comes into play. The eigenvectors with the highest eigenvalues are the Principle Components of the dataset. After viewing the eigenvalues of the components, we can safely ignore the lowest components without a significant loss of the data. The new dataset is comprised by multiplying the matrix with the eigenvectors in the columns transposed that are arranged by importance (significance) by the mean-adjusted data transposed (Smith, 2002).

II- Hierarchical Clustering:

Cluster analysis is defined as a method of separation of a set of data objects into subsets (i.e., clusters) such that objects in a cluster are like one another than objects in another cluster. It's a widely used technique in Biology, Business Intelligence, Security and many other domains of research (Han, Kamber, & Pei, 2012). The Hierarchical method ass its name suggests groups data objects into a hierarchy or a “tree” of clusters. There are two kinds of Hierarchical methods: Agglomerative and divisive. While the first starts by grouping the data objects closest to each other’s into clusters, the latter starts with all of the data objects as a single cluster before separating them.(i.e., Bottom-up approach for the first, Top-down for the second). The separation process depends on the distance (i.e., Linkage Measures) among these data objects. The four most widely used measures for clusters are as follows, where |p-p’| is the distance between two objects or point, p and ṕ; m_i is the mean for cluster C_i; and n_i is the number of objects in C_i:

$$dist_{min}(C_i, C_j) = min_{p \in C_i, ṕ \in C_j}\{|p - ṕ|\} \text{ …….. 1}$$

$$dist_{max}(C_i, C_j) = max_{p \in C_i, ṕ \in C_j}\{|p - ṕ|\} \text{ …….. 2}$$

$$dist_{mean}(C_i, C_j) = |m_i - m_j| \text{ …….. 3}$$

$$dist_{avg}(C_i, C_j) = \frac{1}{n_i n_j} \sum_{p \in C_i, ṕ \in C_j} |p - ṕ| \text{ …….. 4}$$

3.3 <u>Results:</u>

The following variables were used in the PCA:

1- Percent Drive Alone

2- Percent Carpool

3- Percent Transit

4- Percent Walk

5- Percent Taxi, Motorbike, and Other

6- Percent Work from Home

7- Vehicle Per Household (Normalized)

8- Population Density (Normalized)

9- Number of Shared Cars Per 10,000 inhabitants (Normalized)

10- Number of Shared Bikes Per 10,000 inhabitants (Normalized).

Using 6 principal components, the results obtained were as follow:

Mean

[0.68162 0.09193 0.09633 0.04698 0.02922 0.05392 0.68692 0.19942 0.11038 0.09135]

Principal Components Results

```
[[-0.4128  -0.03206  0.31237  0.08901  0.03296  0.01053 -0.30815  0.43338  0.6105   0.25766]
 [ 0.21509  0.0091  -0.22187 -0.03878  0.00688  0.02959  0.34679 -0.48195  0.59455  0.44268]
 [ 0.05805  0.00708  0.00379  0.00346 -0.04895 -0.02343 -0.21731  0.06258 -0.47194  0.84838]
 [-0.00453 -0.01683  0.08122 -0.08266 -0.01873  0.04152  0.80404  0.56289 -0.07815  0.12145]
 [ 0.6653   0.06053 -0.35345 -0.24258 -0.09461 -0.03518 -0.29291  0.47885  0.20637 -0.04544]
 [-0.37004  0.37064 -0.71345  0.28998  0.1045   0.31837 -0.00834  0.13926 -0.02706  0.01159]]
```

Percentage variance explained by components

[0.60536 0.22689 0.10932 0.03991 0.01044 0.00443]

This shows that the 1st principal component explains over 60% of the total variance of the dataset. The second explains 22.7% of the total variance, and the third explains 10.9% of the total variance in data. The most important variables for the principal component analysis as explained by the projections of those first three components, by order of importance, are the population Density-normalized, number of shared cars (as in total vehicles from programs such as car2go, zipcar, etc), the percentage of commuters driving alone, and the percentage of commuters using transit respectively.

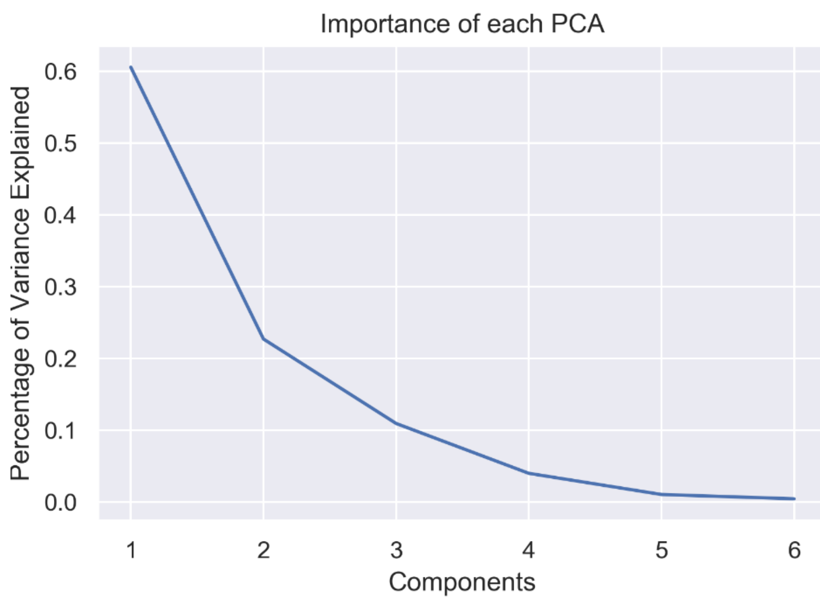

Fig 5: The Importance of Each Principle Component

After applying several hierarchical clustering options (different numbers of 3,4, and 5 clusters) to these three Principle Components of the different U.S cities categorized the following clusters emerged with the highest Silhouette Coefficient of 0.789:

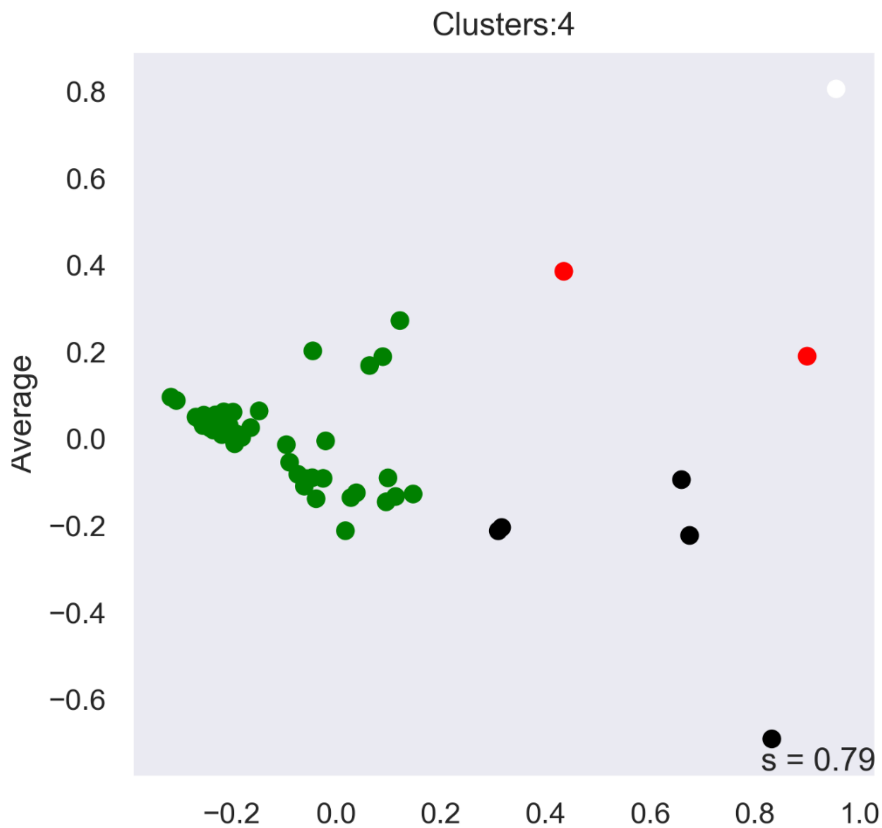

Fig 6: Produced Cities' Clusters

Cluster 1 (Black): ['New York, NY', 'Chicago, IL', 'Philadelphia, PA', 'Boston, MA', 'San Francisco, CA']

Cluster 2 (Red): ['Washington, DC', 'Portland, OR']

Cluster 3 (Green): ['Los Angeles, CA', 'Dallas, TX', 'Houston, TX', 'Miami, FL', 'Atlanta, GA', 'Phoenix, AZ', 'Detroit, MI', 'Minneapolis, MN', 'San Diego, CA', 'Tampa, FL', 'Denver, CO', 'Baltimore, MD', 'St. Louis, MO', 'Charlotte, NC', 'Orlando, FL', 'San Antonio, TX', 'Pittsburgh, PA', 'Sacramento, CA', 'Las Vegas, NV', 'Cincinnati, OH', 'Kansas City, MO', 'Austin, TX', 'Columbus, OH', 'Cleveland, OH', 'Indianapolis city (balance), IN', 'San Jose, CA', 'Nashville-Davidson metropolitan government (balance), TN', 'Virginia Beach, VA', 'Providence, RI', 'Milwaukee, WI', 'Jacksonville, FL', 'Oklahoma City, OK', 'Memphis, TN', 'Raleigh, NC', 'Louisville/Jefferson County metro government (balance), KY', 'New Orleans, LA', 'Hartford, CT','Salt Lake City, UT', 'Birmingham, AL', 'Buffalo, NY', 'Rochester, NY', 'Tucson, AZ', 'Tulsa, OK','Urban Honolulu, HI']

Cluster 4 (White): ['Seattle, WA']

Next, a dendrogram of the data groups was produced as follows:

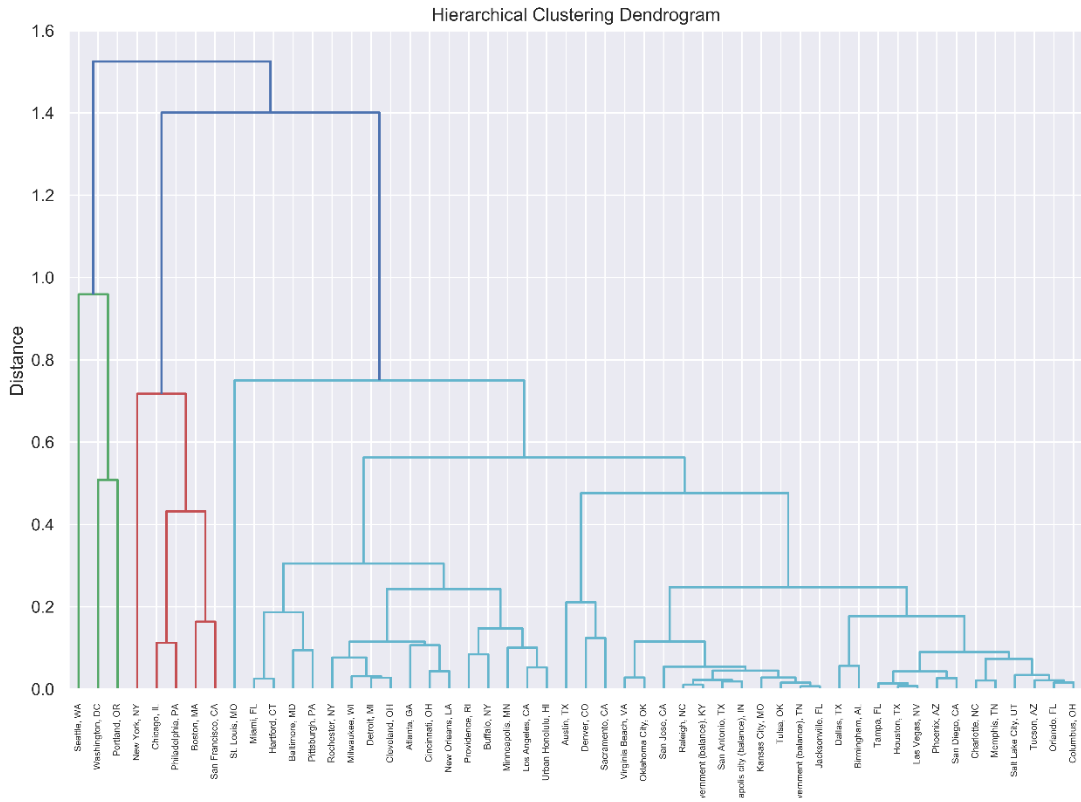

Fig 7: Cities' Dendrogram

3.4 <u>Discussion of The Results:</u>

Cities in the United States have an international reputation of sprawl. Understanding the produced clusters is an exploration of the effect of sprawl on modal share especially that now several cities are trying to be less dependent on private vehicles by introducing pilot projects for new mobility options such as bikes, scooters, carsharing and ride hailing companies.

New York City, Chicago, Philadelphia, and San Francisco ended up in the same cluster. All these cities are dense historical urban centers that has multimodal transportation system within the city. All of this has led them to end up in the same cluster.

One can notice from the results that Seattle has stood alone. This is because it has been taking an active approach towards promoting other modes besides the private vehicle. It's the only city to have transit ridership grown in the age of TNC companies (i.e., Uber, Lyft, etc.) (Graehler, Mucci, & Erhardt, 2018).In 2017, there was a 60% transit ridership growth in Seattle (Levy, 2017) .

Seattle's shared bikes pilot project has been a model of success. The project has shown success in terms of ridership with nearly half a million rides within the pilot period in addition to that the bike operating area covers the entire city with 0.01 collision per one million rides.(Seattle Department of Transportation, 2018). The city of Seattle also hosts two sparate one-way, point to point carsharing programs, BMW's ReachNow and Dailmer's car2go. Seattle provides an approach that can be followed by different cities as a step towards achieving sustainable transportation.

Portland ended up in the same cluster as Washington, DC although the latter is more than double the population density as Portland. This can be explained by Portland's attitude towards shared mobility in its rather successful effort to reduce private vehicle ownership. Portland is one of the first cities to host two separate one-way, point to point carsharing programs (ReachNow and car2go.) Portlans also has the single highest percentage of (bicycle/taxi/motorcycle) in the dataset. All of this has contributed to a lower private vehicle ownership within the city that made it comparable to other dense cities and ended up in the same cluster as Washington, D.C.

The largest cluster, including 44 cities that share similar modal characteristics. Not necessarily all these cities are suburban or sparse yet all of them share a relatively high percentage of “Drive Alone” compared to other modes. Miami is the city with the highest density of over 12,000 person/mile^{2}, yet the over two thirds of commuters drive alone to to work. It’s yet to adopt shared mobility options and that’s evident of the numbers of shared cars that the city has (22 shared cars compared to 647 in Boston, and 695 in Chicago the cities with the closest population densities) and since during the analysis it was shown that in addition to population density, shared cars, and “drive-alone” percentage had the most significant effect on the first principle component we can see how these cities ended up in different clusters. The second highest city in density in the third group was Providence with 60% drive-alone and 0 carshares within the city. Los Angeles is also known for the wide use of private vehicles having arguably the worst traffic jams in the country with 70% of the commuters drive-alone to their work. LA, and other cities in this group can benefit from this analysis to study policies, and planning of other cities in their road towards sustainable transportation.

4.MODE CHOICE MODELING

Introduction:

The primary purpose of this chapter is to introduce and discuss transportation mode shares in the Puget Sound Region in Seattle focusing on six modes (i.e., walk, bike, private vehicle, transit, TNC, and carsharing). The main source of data for this analysis is the Household travel survey that was conducted and published in 2017 by the Puget Sound Regional Council (PSRC). In the field of travel demand modeling, mode choice analysis is arguably one of the most important factors to control since changing to other modes of transportation can increase the efficiency of roadway used in inter and intra-city transport. (Frank S. Koppelman; Chandra Bhat, 2006). Most importantly, understanding the factors that affect mode choice is important for examining and implementing transportation demand management (TDM) policies by metropolitan areas in an effort to decrease traffic.

For over half a century, numerous studies have been conducted using discrete choice modeling to analyze transportation choices. And one of the most popular models is the Multinomial Logit (MNL) model. The Multinomial Logit model was first presented by (McFadden, 1974), where he viewed transportation demand as a utility maximization problem that is affected by personal and population behavior. He showed the success of this model using a sample from California before and after the introduction of new transportation mode, the Bay Area Rapid Transit. After that, many researchers followed the same methodology in travel demand modeling (Frank S. Koppelman; Chandra Bhat, 2006), (L. Frank, Bradley, Kavage, Chapman, & Lawton, 2008), (Koppelman, 1983) among many others.

In this chapter, a Multinomial Logit Model is developed to reveal the significant coefficients effecting the choice of each travel alternative. These coefficients are related to the individual person's characteristics taking the trip, the household characteristics that he/she belongs to, the land-use variables related to the origin and the destination of each individual trip, and the attributes of each alternative mode. After that, a conclusion is drawn from the results found about how each of these characteristics affect mode choice.

4.1 Methodology:

4.1.1 Data Preparation:

The data was obtained by combining several publicly available datasets including Puget Sound Regional Council (PSRC) Household Travel Survey (for person, household and trip attributes), Google API (for estimating the travel time and cost of non-chosen modes for each trip), and the United States Environmental Protection Agency (EPA) dataset (for land-use variables).

The person, household and trip attributes were obtained from the PSRC Household Travel Survey that was conducted and published in 2017. Travel modes with similar characteristics were aggregated by the analyst and other modes with limited number of observations were discarded. All the observations with missing data were discarded as well. The missing data came in many forms: some observations did not have block group coordinates and thus the land-use variables could not be matched to the origin or the destination locations. Some had the location coordinated but the Google API could not fetch some or all the information about the trip time, cost for all the transit mode and most of taxi trips and thus was eliminated from the analysis. Also, the focus of this study was on age groups 18 and above. That is mainly because children don't direct the demand for transportation and their travel behavior is mostly influenced by their parents (Zwerts, Allaert, Janssens, Wets, & Witlox, 2010). Also, children depend on adults for their mobility especially since they don't have the ability to obtain a driver's license (Copperman & Bhat, 2007). The modes considered in this study include walk, bike, personal vehicle, transit (which comprised all bus and rail public transportation modes in the original dataset), TNC and Carsharing. The land-use variables were obtained from EPA's public dataset for the Puget Sound Region with the caveat that the most recent survey was done in 2010. After that, there was an estimation procedure for the travel time and the travel cost for each alternative. The origin and destination block groups were provided in the original dataset but the exact longitude and latitudes of the block group numbers were not provided so the original dataset was merged with another dataset about the geography of each block groups that was also provided by the Puget Sound Regional Council. Google Distance Matrix API was implemented to extract the travel time from each origin and destination coordinated (i.e., latitude and longitude) by means of a python code that would request the data using the provided Google API key (attached in the appendix). It was possible to obtain the travel time for four modes using the Google API (walk, bike, driving, and transit). The carsharing mode was assumed to have the same travel time as the driving mode and it's only different from it with the cost. The time of the TNC mode was similar to the driving mode in addition to 3 minutes (which is the average wait time in the outer boroughs of New York, which was the only reference obtained from the web (Mosendz & Sender, 2014) about Uber wait times). As for the cost function, the cost for the active modes (i.e., walking and biking) is zero, and the cost for driving mode is derived from the latest AAA brochure (AAA, 2018). The parking costs should have been added if there was a reference to the average of the parking time, also there are many obstacles to know whether someone would have parked and paid or whether it's a free parking which would make inconsistencies in the price range. This is the reason why the parking costs were disregarded in this analysis. The current carsharing business models (e.g., Zip Car) is comprised of a recurrent subscription (weekly, monthly, or annually) with varying costs depending on the program that the person/household has subscribed to. Besides the subscription fees there is the fee for sharing the car that is based on the time that the car is shared with minimum time of 30 minutes and in 30 minutes increments (the person can choose to take the car for 30,60,90 minutes and so on) without paying parking, maintenance or fuel cost (Soper, 2016). The carsharing cost was derived from the cost of sharing the car (Zipcar) for a minimum of 30 minutes and in increments of 30 minute. Carshare cost in the city of Seattle according to Zipcar's official website is 7$/hr. The TNC costs were derived from a data table provided by the Intelligent Economist Website about Uber prices across multiple cities inside the U.S. For any city within the dataset there is the minimum fare and an estimated rate per minute or per mile in dollars (Agarwal, 2018). The fare was calculated based on the per minute value to account for any delays calculated by the Google API that are expected to happen during rush hours. The rate was $0.38 US per minute multiplied by the driving time, and the minimum fare ($5.45US) was put instead of any fare that was lower than that amount.

In the end, the final table had several types of variables; the person, household and trip attributes. The land use variables for the origin and destination locations of each trip and an estimate of the time and cost each mode that were of interest. All observations that had missing datapoints were removed from the dataset, and in the end there were around 23,000 datapoints comprising trips that were made using the modes of interest, and after that the modeling process started to choose which variables had significant effect on the mode choice by means of the Multinomial Logit Model using Biogeme Software, which is the most efficient although there wasn't a clear methodology to extract the IIA among the chosen modes.

Table 1: Mode Share

	Mode	Count	Percentage
1	Walk	2589	12.6%
2	Bike	644	3.1%
3	Car	13888	67.8%
4	Transit	2936	14.3%
5	TNC	98	0.5%
6	Carshare	337	1.6%
	Total	20492	100%

Table 2: Description of Explanatory variables used in the MNL Model

Variable	Description	Valid Values	Average	Std. Dev.
Beta_AGE1	Age Group (18-34) years	0 no 1 yes	0.69	0.46
Beta_AGE2	Age Group (34-54) years	0 no 1 yes	0.22	0.41
Beta_AGE3	Age Group (54 and older)	0 no 1 yes	0.08	0.27
BETA_FEM	Female	0 no 1 yes	0.51	0.49
BETA_EDUCATION1	Level of Education until Highschool	0 no 1 yes	0.03	0.18
BETA_EDUCATION2	Level of Education (Highschool-some college)	0 no 1 yes	0.15	0.35
BETA_EDUCATION3	Level of Education (vocational school and higher)	0 no 1 yes	0.81	0.39
BETA_LICENSE	The person has a Driving License	0 no 1 yes	0.95	0.22
BETA_HHSIZ	Household Size	1 1 person 2 2 people 3 3 people 4 ≥ 4 people	2.06	0.95
BETA_HHAVEHICLES	Household Vehicles divided by the number of workers	0 no available vehicles 0.5 0.5 available vehicle 1 1 available vehicle 2 ≥ 2 available vehicles	0.97	0.69
BETA_HHINCOM1	Household Income Under 25,000$	0 no 1 yes	0.12	0.32
BETA_HHINCOM2	Household Income 25000$-99,999$	0 no 1 yes	0.64	0.47
BETA_HHINCOM3	Household Income 99,999$ and higher	0 no 1 yes	0.23	0.42
BETA_HHCARSHARE	Household Participates in a carshare program	0 no 1 yes	0.24	0.42
BETA_HHOFFPARK	Off-street parking spaces at residence	0 0 (no spaces available) 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8	2.87	3.31

		9 9 10 ≥ 10		
BETA_HHSTREETPARKPERMIT	On-street parking availability at/near residence (Permit is required)	0 no 1 yes	0.125	0.33
BETA_OD1A	Origin location: Gross Residential density Household Units/ Acres on unprotected land	Numerical	10.26	11.69
BETA_OD2A_JPHH	Origin Location Jobs Per Households	Numerical	23.26	128.11
BETA_OD4A	Origin Location Distance from population weighted centroid to the nearest transit stop (meters)	Numerical	254.80	203
BETA_OD5DRI	Origin Location Regional Centrality Index – Transit: Census Block Group (CBG) D5drscore (proportional Accessibility of Regional Destination: Transit) relative to max CBSA (Metropolitan Area) D5dr score	Numerical	0.30	0.23
BETA_DD1A	Destination location: Gross Residential density Household Units/ Acres on unprotected land	Numerical	10.30	11.90
BETA_DD2A_JPHH	Destination Location Jobs Per Households	Numerical	23.00	127.31
BETA_DD5AE	Destination Location Working age population within 45 minutes auto travel time, time-decay (network travel time) weighted	Numerical	232560.91	50281
BETA_DD5DRI	Destination Location Regional Centrality Index – Transit: Census Block Group (CBG) D5drscore (proportional Accessibility of Regional Destination :Transit) relative to max CBSA (Metropolitan Area) D5dr score	Numerical	0.30	0.23
BETA_WALKTIME	Walk Time (in minutes)	Numerical	92.15	87.88
BETA_BIKETIME	Bike Time (in minutes)	Numerical	32.77	29.30
BETA_DRIVETIME	Drive Time (in minutes)	Numerical	12.17	6.78
BETA_TRANSITTIME	Transit Time (in minutes)	Numerical	38.16	31.26
BETA_CARSHARETIME	Carshare Time (in minutes)	Numerical	12.17	6.78
BETA_TNCTIME	TNC Time (in minutes)	Numerical	15.1	6.78
BETA_DRIVECOST	Drive Cost (in US Dollars)	Numerical	3.78	3.86
BETA_TNCCOST	TNC Cost (in US Dollars)	Numerical	8.87	5.34

4.1.2 The Multinomial Logit Model (MNL):

Mode choice behavior is discrete in its nature and individuals change their mode choices depending on different attributes including the attributes of each mode such as the time and cost in a way that maximizes their utility, thus a change of attributes of a specific mode would change the preference for that mode. “This is called discrete-choice analysis or the theory of random utility maximization , and the original models are called multinomial logit models” (Daniel, 2002).

The Multinomial Logit Model (MNL) started as a binary choice where the logistic distribution is used but then was generalized to multiple alternatives. Many developments where made and there is a wide array of models that belongs to the Logit family (e.g., binary logit, Multinomial Logit, Nested Logit, etc.). These models have been widely used in transportation demand modeling since the conception of the Multinomial Logit Model (McFadden, 1974) such as (Train, 1978), (Train, 1980), (Kim & Ulfarsson, 2004), (L. Frank et al., 2008),(Kim & Ulfarsson, 2004) and many others. The Multinomial Logit Model is derived assuming the error terms of the utility functions are independent and identical having a Gumbel Distribution (I,e Type I extreme value) (Ben-Akiva & Bierlaire, 1999). That is ε_{in} for all i,n is distributed as :

$$F(\varepsilon) = \exp\,[-e^{-\mu(e-\eta)}],\ \mu > 0 \qquad -\mu(e-\eta)$$

$$F(\varepsilon) = \mu\, e^{-\mu(e-\eta)} \qquad \exp\,[-e^{-\mu(\varepsilon-\eta)}]$$

Where η is a location parameter and μ is a strictly positive scale parameter. The mean of this distribution is

$$\eta + \gamma/\mu$$

Where

$$\Upsilon = \lim_{k\to\infty} \sum_{i=1}^{k} \frac{1}{i} - \ln(k) \cong 0.5772$$

Is the Euler constant. The variance of the distribution is

$$\pi^2 / 6\mu^2$$

The probablity that a given individual n chooses alternative I within the choice set C_n is given by

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_{j\in C_n} e^{\mu V_{jn}}}$$

An important property of the Multinomial Logit Model that is considered a limiation is the assumption that the independence from Irrelevan Alternatives (IIA) that has been highlighted many times in the literature with the famous Blue Bus/Red Bus paradox.(Frank S. Koppelman; Chandra Bhat, 2006). This property means that the ratio of the probabilities of any two alternatives is not affected by the choice set.

4.2 Results:

Table 3: MNL Model Results:

Mode	Walk	Bike	Auto	Transit	TNC	Carshare
Alternative Specific Constant	3.63 (0.00)	-1.46 (0.00)		-0.226 (0.36)	-1.61 (0.01)	-5.05 (1.00)
BETA_AGE1	0.235 (0.00)			-0.245 (0.00)	1.25 (0.00)	
BETA_FEM	-0.368 (0.00)	-0.734 (0.00)				
BETA_EDUCATION1			0.557 (0.00)			
BETA_EDUCATION2				0.105 (0.11)		1.32 (0.00)
BETA_LICENSE	-0.849 (0.00)			-1.40 (0.00)	-1.20 (0.00)	
BETA_HHSIZ		0.154 (0.00)			-0.213 (0.00)	

BETA_HHAVEHICLES			1.16 (0.00)	-0.482 (0.00)	-0.584 (0.00)	
BETA_HHINCOME1			1.08 (0.00)			
BETA_HHINCOME2			1.64 (0.00)			
BETA_HHINCOME3			2.00 (0.00)			
BETA_HHCARSHARE		0.275 (0.00)	-0.340 (0.00)			2.66 (0.00)
BETA_HHOFFPARK		-0.0425 (0.00)	0.0258 (0.00)			
BETA_HHSTREETPARKPERMIT			0.234 (0.00)			
BETA_OD1A			0.0102 (0.00)			
BETA_OD2A_JPHH (*e^-4)				5.61 (0.00)		
BETA_OD4A (*e^-005)				1.17 (0.00)		
BETA_OD5DRI			-2.54 (0.00)	0.955 (0.00)	1.55 (0.00)	
BETA_DD1A			0.0106 (0.00)		0.0142 (0.00)	
BETA_DD2A_JPHH (*e^-5)			7.78 (0.00)			
BETA_DD5AE(* e ^ -6)				2.48 (0.00)		
BETA_DD5DRI			-2.27 (0.00)	0.87 (0.00)	1.06 (0.00)	
BETA_WALKTIME	-0.103 (0.00)					
BETA_BIKETIME		-0.0461 (0.00)				
BETA_DRIVETIME			-0.0883 (0.00)			
BETA_TRANSITTIME				-0.0232 (0.00)		
BETA_CARSHARETIME						-0.134 (0.00)
BETA_DRIVECOST			-0.074 (0.00)			
BETA_TNCCOST					-0.284 (0.00)	

4.3 Discussion of The Results:

The base mode for the alternative specific constant(ASC) was private vehicle. The constants shows a high preference towards Walking and the lowest of preferences towards Carshare with ASC is close to zero (the same for private vehicle). The model suggests that being in the youngest age cohort has a positive effect on walking and negative effect on Transit. TNC result was very interesting since it shows a positive correlation with being of that cohort. Being a female has a negative correlation with both active modes (Walk, and Bike) being of a higher educational group has a positive correlation with both Transit and Carshare unlike being of the lower educational group were the positive correlation is with Private Vehicle. Having a license had a negative correlation on all non-auto modes (i.e., Walk, transit, and TNC) with the highest negative correlation to Transit and lowest to Walk mode. The size of the individual household showed a negative correlation with Bike mode but negative correlation with TNC mode. The association of the household income groups showed precisely how does income affect preferences towards private vehicle mode. Where the coefficient increases gradually from 1.16 at the lowest income group to 1.64 for the middle-income households and is highest for the highest income group to 2.0. Participating in a carshare program had a very strong positive correlation with opting or Carsharing mode and a negative correlation with Auto mode which is very promising for diverting transportation demand towards more sustainable means of transportation. The availability of parking (off-street, on-street with a permit) also had a positive correlation with Auto mode. The effects of land-use variables for the origin and destination locations has also significant negative and positive effects on mode preference. (e.g., D5AE had a positive impact on Transit and negative impact on Auto mode, etc.) The time and cost of each mode had a significant negative effects on choosing the mode varying from one mode to another.

5. AV INTERESTS AND CONCERNS

5.1 Introduction:

Autonomous Vehicles (AVs) have been gaining more and more significance and impact over the transportation industry. Since 2012, there are at least 21 states that have considered legislations related to autonomous vehicles. There is an increasing trend in the number of States introducing bills related to AVs from 6 states in 2012 to 15 states in 2015. Further, in 2018, 15 sates enacted 18 AV related bills. (NCSL, 2019)

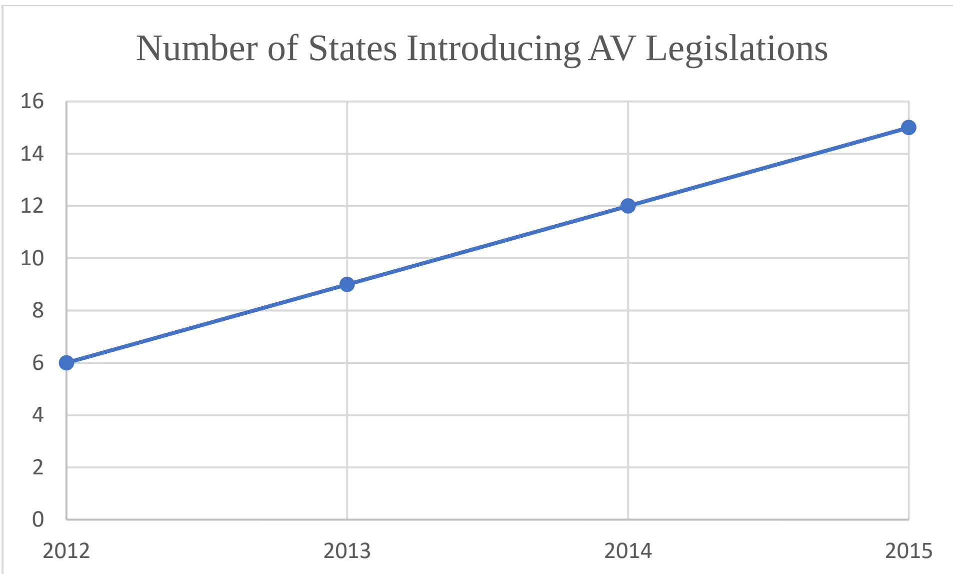

Fig 8: Numbers of States Introducing AV Legislations

These trends indicate the accelerated significance this technology is gaining, and with that comes public opinion adaptability and readiness to embrace this technology - especially with the many perceived benefits for full automated vehicles to be operating. For example, the new accessibility for the elderly and the disabled to independent travels instead of relying on other family members or friends (Fagnant & Kockelman, 2015), (Harper, Hendrickson, Mangones, & Samaras, 2016). In addition to that, if this technology is coupled with connectivity then AVs are expected to lower traffic congestion especially if it had reached a significant level of market penetration (Chandra & Camal, 2016). With that comes many concerns about safety, regulation, security and privacy of this technology (Fagnant & Kockelman, 2015).

Currently, there are not many studies about consumer adoption of AVs, leaving many gaps to fill. Some has done simple descriptive analysis to investigate individuals' demographic characteristics' effects on their opinions of concerning AVs (Kyriakidis, Happee, & De Winter, 2015), (Payre et al., 2014). Other more recent studies took a deeper look into the effects of demographics, socioeconomics and the built environment on people's adoption behavior (Shabanpour, Golshani, Shamshiripour, & Mohammadian, 2018), (Bansal & Kockelman, 2017). Since major gaps exists in the literature further analysis should be made into the preferences of people towards this technology and the factors affecting people's opinion and thus adoption rate for it. It is also pivotal in marketing research to see which market segment to target and also possibly which service or feature the people are more interested in. This is the inspiration of this chapter as statistical modeling is made using NLOGIT5, to test the people's interests in AVs (e.g. Autonomous carshare, AV Ownership, etc.) and their concerns (e.g. legal liability, poor weather performance, etc.) as the rest of the chapter will present the results and the discussion of the statistical models produced.

5.2 Methodology:

5.2.1 Data Preparation:

The data used for this analysis comes from the Puget Sound Regional Council Household Travel Survey that was published in 2017 where they asked participants about their preferences towards AVs. The individual attributes were merged with the household attributes and the land use variables that were obtained from the United States Environmental Protection Agency that were published in 2010 and was queried using the block group code for each household in the survey. After that all the empty choices where omitted from the dataset for each interest or concern. NLOGIT5 Software was used in this analysis after recoding the interests and concerns (0(very interested/concerned)-4(not at all interested/concerned)) in order for the software to work properly and the different variables where tested for their significance.

Since the AV technology is not yet available for the real market stated preferences (SP) choice experiments are used in this study. Stated preferences experiments can be divided into two types: single-alternative selection, where the person is asked to choose the most preferred alternative among a choice set, which is the dominant type (Flynn, Louviere, Peters, & Coast, 2007) and rating/ranking of alternatives. (Hausman & Ruud, 1987)(Shabanpour et al., 2018). Ranked choices are better since they shed light on the preference gradient of a person which shed more light on acceptability rather than the singe choice experiment, but it may produce some complications to the experiments and/ or produce some biases in the model outcomes.(Ben-Akiva, Morikawa, & Shiroishi, 1992)

5.2.2 **Descriptive Statistics:**

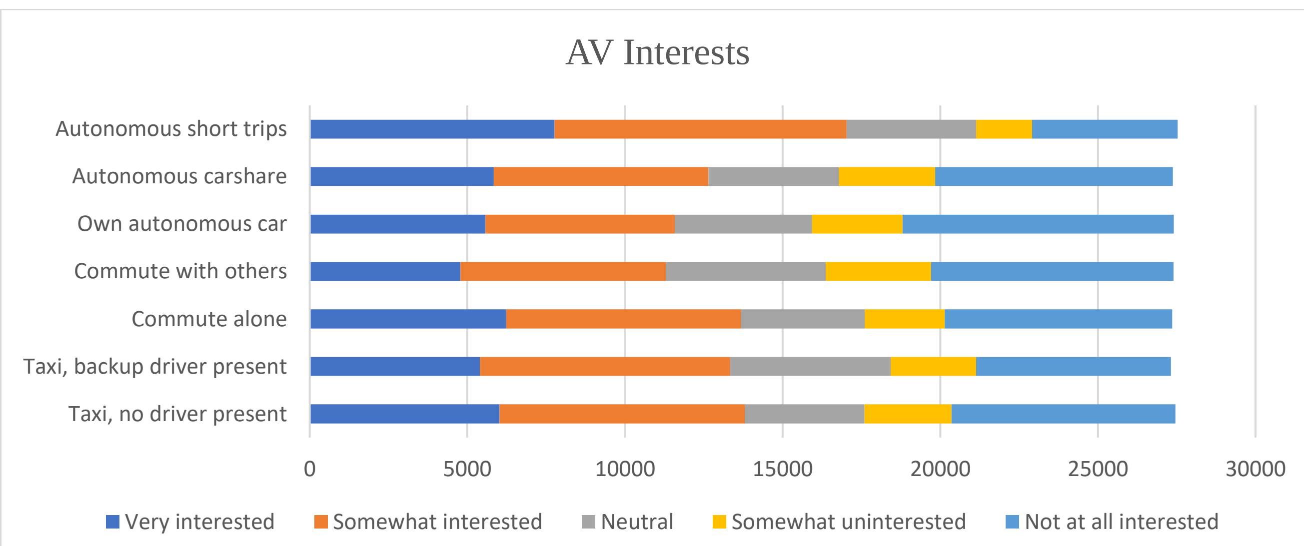

Figure 9: Distribution of Levels of Interests for different AV uses

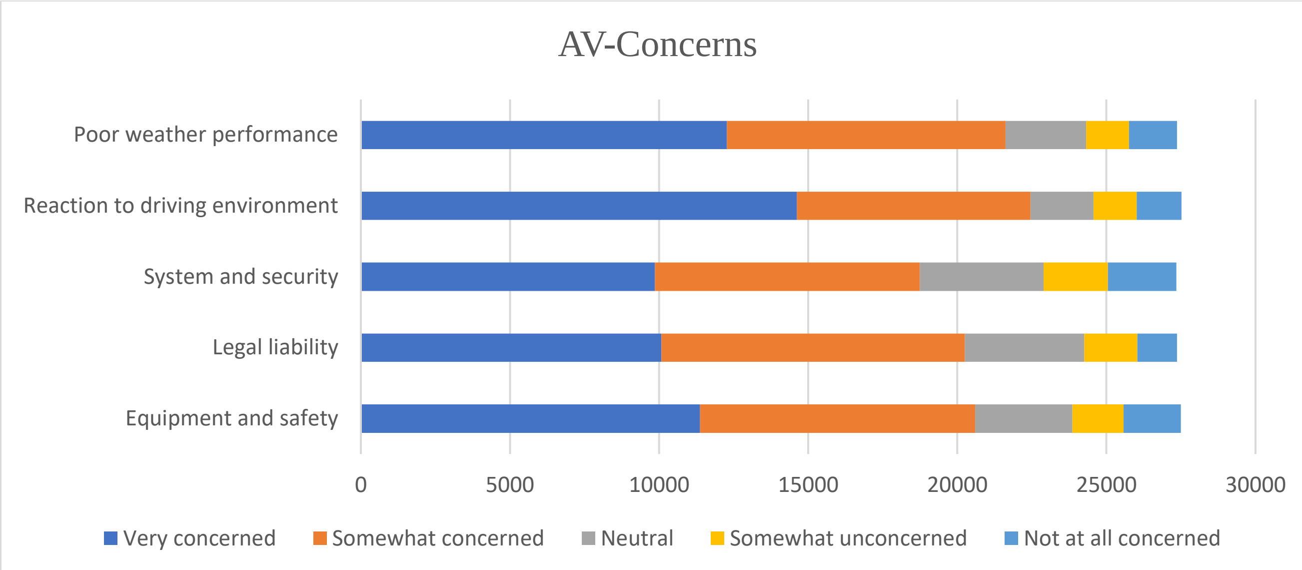

Figure 10: Distribution of Levels of Concerns for AVs

Table 3: Description of Explanatory variables used in the Probit Model

Variable	Description	Valid Values		Average	Std. Dev.
Beta_AGE1	Age Group (18-34) years	0 1	No Yes	0.60	0.48
Beta_AGE2	Age Group (34-54) years	0 1	No Yes	0.61	0.43
Beta_AGE3	Age Group (54 and older)	0 1	No Yes	0.61	0.35
FEMALE	Female	0 1	No Yes	0.47	0.49
BETA_EDUCATION1	Level of Education ≥ Highschool	0 1	No Yes	0.61	0.24
BETA_EDUCATION2	Level of Education (Highschool-some college)	0 1	No Yes	0.61	0.34
BETA_EDUCATION3	Level of Education (vocational school and higher)	0 1	No Yes	0.61	0.4
PTSTUDEN	The person is a part-time student	0 1	No Yes	0.04	0.23

LICENSE2	The person has a Driving License	0 No 1 Yes	0.95	0.21
LEARNING	The person has a Learner's Permit	0 No 1 Yes	0.007	0.08
WORKER	The person is a worker	0 No 1 Yes	0.83	0.44
SMARTPHONE	The person has a smartphone	0 No 1 Yes	0.96	0.21
HHSIZE	Household Size	1 1 person 2 2 people 3 3 people 4 4 people Etc.	2.10	0.99
CITYOFRE	Sample address located in Redmond	0 No 1 Yes	0.18	0.39
CITYOFSE	Sample address located in Seattle urban village	0 No 1 Yes	0.57	0.49
VEHICLE	Number of Household Vehicles	0 No vehicles 1 1 vehicle 2 2 vehicles Etc.	1.45	0.91
NUMWORKE	Number of worker in a Household	0 no worker 1 1 worker 2 2 workers Etc.	1.63	0.59
BETA_HHINCOM1	Household Income Under 25,000$	0 No 1 Yes	0.61	0.34
BETA_HHINCOM2	Household Income 25000$-99,999$	0 No 1 Yes	0.61	0.48
BETA_HHINCOM3	Household Income 99,999$ and higher	0 No 1 Yes	0.60	0.42
CARSHARE	Household Participates in a carshare program	0 No 1 Yes	0.22	0.41
RESTRICT	On-street parking availability at/near residence (Permit is required)	0 No 1 Yes	0.11	0.32
D1A	Household location: Gross Residential density Household Units/ Acres on unprotected land	Numerical	11.81	13.72
D2B_E5MI	Household Location 5-tier employment entropy (denominator set to observed employment types in the CBG)	Numerical	0.62	0.29

D3AMM	Household Location Network density in terms of facility miles of multi-modal links per square mile	Numerical	3.49	3.25
D3APO	Household Location Network density in terms of facility miles of pedestrian-oriented links per square mile	Numerical	19.80	8.15
D5CRI	Household Location Regional Centrality Index – Auto: CBG D5cr score relative to max CBSA D5cr score	Numerical	0.52	0.19
D5CEI	Household Location Regional Centrality Index – Auto: CBG D5ce score*Proportional Accessibility to Regional Destination* relative to max CBSA D5ce score	Numerical	0.69	0.18

5.2.3 <u>**Ordered Probit Model:**</u>

When the dependent variables have a continuous nature that is meaningful yet doesn't represent a linear metric Regression models will fail in the modeling process and would produce heteroskedasticity - modelers use the Ordered Probit model to get around this shortcoming (Jackman, 2000). This model has its origins from Biostatistics (Aitchison, Silvey, Aitchison, & Silvey, 1957) but was used in social, political science and also in Transportation (Pudney & Shields, 2000), and (Abdel-Aty, 2003).

The main theme behind this model is the existence of a latent continuous metrics behind the ordinal response observed by analyst. Thus, a threshold partitioning the continuous metrics into different ordered regions that corresponds to the different ordinal categories. Thus, the latent continuous variable, y* is a linear combination of some predictors, x, plus a disturbance term that has a standard normal distribution:

$$y_i^* = x_i\beta + e_i\,, e_i \sim \;\; N(0,1), \forall\ \mathrm{i} = 1, \dots, \mathrm{N}. \qquad (1)$$

y_i, is the observed ordinal variable, takes on values 0 through m according to the following scheme:

$$y_i = j \iff \mu_{j-1} < \; y_i^* \leq \mu_j,$$

Where j=0…, m and in this analysis I'm defining $\mu_{-1} = \; -\infty$, and $\mu_m = \; +\infty$.

In this model we are concerned with observing how changes in the predictors affect the probability of observing a particular ordinal outcome. For example:

$$P[y_i = 0] \;\; = P[\mu_{-1} < \; y_i^* \leq \mu_0],$$

$$= P[-\infty < \; y_i^* \leq \mu_0],$$

$$= P[y_i^* \leq \mu_0],$$

Substituting from (1),

$$= P[x_i\beta + e_i \; \leq \mu_0],$$

$$= P[e_i \; \leq \mu_0 - x_i\beta],$$

$$= \boldsymbol{\phi}\big(\mu_0 - x_i\beta\big);$$

$$P[y_i = 1] = P[\mu_0 < \; y_i^* \leq \mu_1],$$

$$= P[\mu_0 < x_i\beta + e_i \; \leq \mu_1],$$

$$= P[\mu_0 - x_i\beta < e_i \; \leq \mu_1 - x_i\beta],$$

$$= \boldsymbol{\phi}\big(\mu_1 - x_i\beta\big) - \boldsymbol{\phi}\big(\mu_0 - x_i\beta\big).$$

In a similar manner

$$P[y_i = 2] = \boldsymbol{\phi}(\mu_2 - x_i\beta) - \boldsymbol{\phi}(\mu_1 - x_i\beta)$$

In general, the equation becomes:

$$P[y_i = j] = \boldsymbol{\phi}(\mu_j - x_i\beta) - \boldsymbol{\phi}(\mu_{j-1} - x_i\beta)$$

For j=m (i. e., the highest category) the generic form reduces to

$$P[y_i = m] = \boldsymbol{\phi}(\mu_m - x_i\beta) - \boldsymbol{\phi}(\mu_{m-1} - x_i\beta)$$

$$= 1 - \boldsymbol{\phi}(\mu_{m-1} - x_i\beta)$$

5.3 Model Results:

5.3.1 Interests in different forms of AV Use:

In this section, the probit models for each interests are presented as it was produced by NLOGIT5 with the model specification and the effect and the significance level of each variable. The Level of interests in the survey was converted to (0-4) scale where Zero representing the highest level of interest and 4 representing the lowest.

I- AV Interest 1: Taxi, no driver present

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable              INTEREST
Log likelihood function      -5213.69139
Restricted log likelihood    -5515.97413
Chi squared [   6 d.f.]        604.56548
Significance level                .00000
McFadden Pseudo R-squared       .0548013
Estimation based on N =    3681, K =  10
Inf.Cr.AIC  =  10447.4 AIC/N =     2.838
Model estimated: Feb 19, 2020, 21:20:29
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                Standard            Prob.      95% Confidence
INTEREST|  Coefficient     Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|    1.46230***      .06808    21.48  .0000     1.32886   1.59574
    AGE1|    -.60832***      .03879   -15.68  .0000     -.68435   -.53229
  FEMALE|     .41764***      .03656    11.42  .0000      .34599    .48929
 EDUCAT3|    -.33405***      .04879    -6.85  .0000     -.42967   -.23843
VEHICLE_|     .05703***      .02051     2.78  .0054      .01683    .09723
CARSHARE|    -.29067***      .04726    -6.15  .0000     -.38330   -.19805
   D3AMM|    -.02198***      .00573    -3.84  .0001     -.03321   -.01075
        |Threshold parameters for index
   Mu(1)|     .73667***      .01957    37.64  .0000      .69832    .77503
   Mu(2)|    1.13541***      .02117    53.64  .0000     1.09393   1.17689
   Mu(3)|    1.36810***      .02295    59.60  .0000     1.32312   1.41309
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

The variables showed high significance in their effect on the level of interest towards AVs as Taxis with no backup drivers. Younger age cohort, higher level of education and participation in a carshare program, and increase in multimodality coefficient all increase the level of interest toward this utilization of AV technology unlike being female and household vehicle ownership.

II- AV Interest 2 : Taxi, backup driver present

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                 INT2
Log likelihood function     -5411.20017
Restricted log likelihood   -5643.66826
Chi squared [   7 d.f.]       464.93618
Significance level               .00000
McFadden Pseudo R-squared      .0411910
Estimation based on N =   3668, K =  11
Inf.Cr.AIC  =  10844.4 AIC/N =    2.956
Model estimated: Feb 19, 2020, 21:23:47
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
    INT2|  Coefficient       Error       z     |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|    2.15284***      .08921    24.13  .0000     1.97799   2.32768
    AGE1|    -.49218***      .03952   -12.45  .0000     -.56965   -.41471
  FEMALE|     .18164***      .03591     5.06  .0000      .11127    .25202
 EDUCAT3|    -.31225***      .04895    -6.38  .0000     -.40819   -.21632
SMARTPON|    -.41490***      .07103    -5.84  .0000     -.55411   -.27569
  HHINC1|    -.09208*        .05437    -1.69  .0904     -.19865    .01449
CARSHARE|    -.14572***      .04716    -3.09  .0020     -.23816   -.05328
   D5CRI|    -.50076***      .09182    -5.45  .0000     -.68072   -.32081
        |Threshold parameters for index
   Mu(1)|     .79674***      .01963    40.59  .0000      .75827    .83521
   Mu(2)|    1.27812***      .02121    60.27  .0000     1.23655   1.31968
   Mu(3)|    1.55049***      .02331    66.52  .0000     1.50481   1.59617
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

Variables that have shown significant effect on the level of interest towards AVs as taxis with a backup driver are shown in this model. In addition to the previous factors, owning a smartphone is associated with higher level of interest towards AVs.

III- AV Interest 3: Commute alone

```
--------------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                INT_3
Log likelihood function     -3694.12661
Restricted log likelihood   -3762.13752
Chi squared [   7 d.f.]       136.02182
Significance level               .00000
McFadden Pseudo R-squared      .0180777
Estimation based on N =   2457, K =  11
Inf.Cr.AIC  =   7410.3 AIC/N =    3.016
Model estimated: Feb 19, 2020, 21:55:01
Underlying probabilities based on Normal
--------+-----------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
   INT_3|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Index function for probability
Constant|      .70957***      .04450    15.95  .0000      .62236    .79679
    AGE3|      .60924***      .12570     4.85  .0000      .36287    .85560
  FEMALE|      .35239***      .04418     7.98  .0000      .26580    .43898
 EDUCAT1|      .32165***      .11799     2.73  .0064      .09040    .55290
LEARNING|     -.73993***      .26592    -2.78  .0054    -1.26113   -.21874
  HHINC3|     -.09841*        .05041    -1.95  .0509     -.19722    .00040
CARSHARE|     -.15768***      .05265    -3.00  .0027     -.26087   -.05449
   D3AMM|     -.01667**       .00676    -2.47  .0136     -.02991   -.00343
        |Threshold parameters for index
   Mu(1)|      .67298***      .02220    30.31  .0000      .62946    .71649
   Mu(2)|     1.07240***      .02473    43.36  .0000     1.02393   1.12088
   Mu(3)|     1.30664***      .02708    48.24  .0000     1.25355   1.35972
--------+-----------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------------------
```

This model shows the variables affecting the interest commuting alone in an AV. Having a learner's permit increases the level of interest since learners are in need of driver assistance. Higher level of income, participating in a carshare program and higher coefficient of multimodality all increase the level of interest, while being in the oldest age cohort and being female both have negative effect on the level of interest in owning an AV.

IV- AV Interest 4: Commute with others

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable              AVINT_4
Log likelihood function     -3600.40443
Restricted log likelihood   -3766.29450
Chi squared [   5 d.f.]       331.78015
Significance level               .00000
McFadden Pseudo R-squared      .0440460
Estimation based on N =   2457, K =   9
Inf.Cr.AIC  =   7218.8 AIC/N =    2.938
Model estimated: Feb 19, 2020, 22:05:24
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
 AVINT_4|  Coefficient      Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|    1.37014***     .06792    20.17  .0000     1.23701   1.50326
    AGE1|    -.68918***     .05174   -13.32  .0000     -.79060   -.58777
  FEMALE|     .22621***     .04418     5.12  .0000      .13963    .31280
LEARNING|    -.78713***     .26432    -2.98  .0029    -1.30519   -.26907
CARSHARE|    -.24488***     .05379    -4.55  .0000     -.35031   -.13945
VEHICLE_|     .12581***     .02569     4.90  .0000      .07546    .17615
        |Threshold parameters for index
   Mu(1)|     .73200***     .02384    30.71  .0000      .68528    .77873
   Mu(2)|    1.23963***     .02549    48.63  .0000     1.18967   1.28960
   Mu(3)|    1.53703***     .02802    54.85  .0000     1.48211   1.59195
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

The same variables that are significant in affecting people's interests in previous uses of AVs are the most significant in affecting interest in commuting with others using AVs. Vehicle ownership negatively affects the interest in AVs.

V- AV Interest 5: Own autonomous car

```
--------------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                     INT5
Log likelihood function     -5191.83422
Restricted log likelihood   -5420.04386
Chi squared [   5 d.f.]       456.41928
Significance level               .00000
McFadden Pseudo R-squared      .0421048
Estimation based on N =    3670, K =    9
Inf.Cr.AIC  =  10401.7 AIC/N =    2.834
Model estimated: Feb 19, 2020, 22:22:23
Underlying probabilities based on Normal
--------+-----------------------------------------------------------------------
        |                  Standard            Prob.       95% Confidence
    INT5|  Coefficient       Error       z     |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Index function for probability
Constant|     1.44411***      .05621    25.69  .0000     1.33394   1.55428
    AGE1|    -.66615***       .03857   -17.27  .0000     -.74175   -.59055
  FEMALE|     .35596***       .03683     9.67  .0000      .28378    .42815
 EDUCAT3|    -.20239***       .04969    -4.07  .0000     -.29977   -.10500
  HHINC3|    -.09908**        .04389    -2.26  .0240     -.18510   -.01307
   D3AMM|    -.01450**        .00563    -2.57  .0101     -.02554   -.00345
        |Threshold parameters for index
   Mu(1)|     .61940***       .01860    33.30  .0000      .58295    .65586
   Mu(2)|    1.01906***       .02046    49.80  .0000      .97896   1.05917
   Mu(3)|    1.24705***       .02219    56.20  .0000     1.20356   1.29054
--------+-----------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
```

The same variables are consistent in their effect. This model shows that younger age cohort, higher level of education, higher income and multimodality all effects the level of interest positively. Being a female has the opposite effect.

VI- AV Interest 6: Autonomous carshare

```
--------------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                 INT6
Log likelihood function     -5281.65173
Restricted log likelihood   -5435.00848
Chi squared [   5 d.f.]       306.71351
Significance level               .00000
McFadden Pseudo R-squared      .0282165
Estimation based on N =    3669, K =   9
Inf.Cr.AIC  =  10581.3 AIC/N =    2.884
Model estimated: Feb 19, 2020, 22:34:26
Underlying probabilities based on Normal
--------+-----------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
    INT6|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Index function for probability
Constant|      .91222***      .08088    11.28  .0000      .75370   1.07074
    AGE2|      .45305***      .04292    10.56  .0000      .36893    .53717
  FEMALE|      .34111***      .03649     9.35  .0000      .26960    .41263
PTSTUDEN|    -.36405***      .10683    -3.41  .0007    -.57343   -.15467
 EDUCAT3|    -.44327***      .04913    -9.02  .0000    -.53956   -.34698
LICENSE2|      .18235**       .07290     2.50  .0124      .03947    .32522
        |Threshold parameters for index
   Mu(1)|      .66634***      .01890    35.25  .0000      .62930    .70339
   Mu(2)|     1.07042***      .02050    52.22  .0000     1.03024   1.11060
   Mu(3)|     1.28341***      .02202    58.28  .0000     1.24025   1.32657
--------+-----------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------------------
```

The variables that are shown here have a consistent effect across all interests. All of them align with expectations. For example owning a license has a negative effect on the level of interest in Autonomous carshare.

VII- <u>AV Interest 7: Autonomous short trips</u>

```
--------------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                  INT7
Log likelihood function      -5305.46960
Restricted log likelihood    -5539.47774
Chi squared [   9 d.f.]        468.01628
Significance level                .00000
McFadden Pseudo R-squared       .0422437
Estimation based on N =    3688, K =  13
Inf.Cr.AIC  =  10636.9 AIC/N =    2.884
Model estimated: Feb 19, 2020, 22:48:26
Underlying probabilities based on Normal
--------+-----------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
    INT7|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Index function for probability
Constant|    1.24852***      .06816    18.32  .0000     1.11493   1.38211
    AGE1|    -.57016***      .03837   -14.86  .0000     -.64537   -.49495
  FEMALE|     .23953***      .03611     6.63  .0000      .16875    .31031
PTSTUDEN|    -.22894**       .10744    -2.13  .0331     -.43951   -.01836
 EDUCAT3|    -.33484***      .04826    -6.94  .0000     -.42942   -.24026
  HHSIZE|     .06883***      .01863     3.70  .0002      .03232    .10534
  HHINC3|    -.16719***      .04466    -3.74  .0002     -.25472   -.07967
CARSHARE|    -.18407***      .04703    -3.91  .0001     -.27625   -.09189
RESTRICT|    -.13818**       .05869    -2.35  .0186     -.25321   -.02315
   D3AMM|    -.01659***      .00567    -2.93  .0034     -.02770   -.00548
        |Threshold parameters for index
   Mu(1)|     .78657***      .01938    40.59  .0000      .74859    .82455
   Mu(2)|    1.17937***      .02154    54.76  .0000     1.13716   1.22158
   Mu(3)|    1.37024***      .02327    58.90  .0000     1.32464   1.41584
--------+-----------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------------------
```

The last model tests people's interest in autonomous short trips. All the variables are highly significant. Having a restricted on-street parking increases the level of interest to this feature. Thus, the city can use regulations coupled with this technology so that AVs can be used to solve the first/last mile problem.

5.3.2 AV Concerns:

The same modeling procedure to test the interests in AVs are used again to test people's concerns about them. The ordered dependent variable was coded as follows:

0: Very Concerned

1: Somewhat Concerned

2:Neutral

3: Somewhat Not Concerned

4: Not at All Concerned

I- AV Concern 1: Equipment and safety

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                    CON1
Log likelihood function     -2422.39474
Restricted log likelihood   -2455.69397
Chi squared [   4 d.f.]        66.59847
Significance level               .00000
McFadden Pseudo R-squared      .0135600
Estimation based on N =    1842, K =    8
Inf.Cr.AIC  =    4860.8 AIC/N =     2.639
Model estimated: Feb 16, 2020, 15:13:06
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
    CON1|  Coefficient       Error       z     |z|>Z*        Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|      .63464***      .16763     3.79  .0002       .30609     .96320
    AGE1|      .25595***      .05445     4.70  .0000       .14924     .36267
  FEMALE|     -.26400***      .05183    -5.09  .0000      -.36559    -.16242
 LICENSE|     -.22415**       .10245    -2.19  .0287      -.42495    -.0233
 Educat3|     -.18460***      .06654    -2.77  .0055      -.31502    -.05419
        |Threshold parameters for index
   Mu(1)|      .78840***      .02802    28.14  .0000       .73349     .84331
   Mu(2)|     1.24502***      .03513    35.44  .0000      1.17618    1.31387
   Mu(3)|     1.60550***      .04387    36.60  .0000      1.51952    1.69148
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

In this model, Being of the first category of the three age groups had a positive value. Having a license was also associated with higher level of concerns since drivers are more aware of the risk of driving a vehicle. Being a female and having higher levels of education indicates higher concern towards Equipment and Safety.

II- AV Concern 2: Legal liability

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                 CON2
Log likelihood function     -2422.75709
Restricted log likelihood   -2434.92637
Chi squared [   2 d.f.]        24.33856
Significance level               .00001
McFadden Pseudo R-squared      .0049978
Estimation based on N =    1833, K =   6
Inf.Cr.AIC  =    4857.5 AIC/N =    2.650
Model estimated: Feb 16, 2020, 15:32:23
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
    CON2|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|      .20954***      .05045     4.15  .0000      .11067     .30842
    AGE1|      .16456***      .05287     3.11  .0019      .06094     .26818
  FEMALE|     -.18948***      .05133    -3.69  .0002     -.29009    -.08887
        |Threshold parameters for index
   Mu(1)|      .87738***      .02886    30.40  .0000      .82082     .93394
   Mu(2)|     1.44797***      .03763    38.48  .0000     1.37422    1.52172
   Mu(3)|     1.75305***      .04563    38.42  .0000     1.66361    1.84249
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

Being in the younger population of the sample, unlike being female has an association with lower levels of concerns. Only these two variables were found to have a statistically significant effect on Legal Liability concern.

III- AV Concern 3: System and security

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                  CON3
Log likelihood function      -2549.31107
Restricted log likelihood    -2574.31677
Chi squared [   3 d.f.]         50.01138
Significance level                .00000
McFadden Pseudo R-squared       .0097135
Estimation based on N =    1831, K =   7
Inf.Cr.AIC  =   5112.6 AIC/N =     2.792
Model estimated: Feb 16, 2020, 15:46:01
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
    CON3|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|      .35575***      .04962     7.17  .0000      .25850    .45300
    AGE1|      .14742***      .05224     2.82  .0048      .04504    .24981
  FEMALE|     -.31020***      .05092    -6.09  .0000     -.41000   -.21040
 Educat1|      .24889**       .10806     2.30  .0213      .03709    .46068
        |Threshold parameters for index
   Mu(1)|      .82479***      .02762    29.87  .0000      .77067    .87892
   Mu(2)|     1.34547***      .03437    39.14  .0000     1.27810   1.41284
   Mu(3)|     1.66189***      .04114    40.40  .0000     1.58126   1.74252
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

Being in the lower rank of age and education is associated with lower degrees of concerns. Being a female has the exact opposite effect.

IV- AV Concern 4: Reaction to driving environment

```
----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                 CON4
Log likelihood function     -2234.30324
Restricted log likelihood   -2256.09592
Chi squared [   4 d.f.]        43.58537
Significance level               .00000
McFadden Pseudo R-squared      .0096595
Estimation based on N =   1846, K =   8
Inf.Cr.AIC  =   4484.6 AIC/N =    2.429
Model estimated: Feb 16, 2020, 16:28:25
Underlying probabilities based on Normal
--------+-------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
    CON4|  Coefficient      Error       z    |z|>Z*         Interval
--------+-------------------------------------------------------------------
        |Index function for probability
Constant|     .34595***      .11093     3.12  .0018      .12852    .56338
    AGE3|    -.18289**       .07833    -2.33  .0196     -.33641   -.02936
  FEMALE|    -.24587***      .05327    -4.62  .0000     -.35029   -.14146
 EDUCAT1|     .29637***      .10982     2.70  .0070      .08112    .51162
 LICENSE|    -.30220***      .10772    -2.81  .0050     -.51332   -.09108
        |Threshold parameters for index
   Mu(1)|     .78855***      .02946    26.77  .0000      .73081    .84629
   Mu(2)|    1.18427***      .03706    31.96  .0000     1.11164   1.25690
   Mu(3)|    1.50330***      .04606    32.64  .0000     1.41302   1.59359
--------+-------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
----------------------------------------------------------------------------
```

It is observed that the values coincides with logical sense and previous research in a way that being of older age group, being a female and having a driver license is associated with higher level of concerns unlike having lower educational levels.

V- AV Concern 5: Poor weather performance

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable                   CON5
Log likelihood function       -2374.05776
Restricted log likelihood     -2406.22149
Chi squared [   4 d.f.]          64.32747
Significance level                 .00000
McFadden Pseudo R-squared        .0133669
Estimation based on N =    1838, K =    8
Inf.Cr.AIC  =    4764.1 AIC/N =     2.592
Model estimated: Feb 16, 2020, 16:45:58
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                Standard            Prob.      95% Confidence
    CON5|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability
Constant|     .57453***      .11079     5.19  .0000      .35738    .79168
    AGE1|     .12177**       .05320     2.29  .0221      .01749    .22604
  FEMALE|    -.35951***      .05196    -6.92  .0000     -.46134   -.25768
 LICENSE|    -.30953***      .10122    -3.06  .0022     -.50791   -.11115
CARSHARE|    -.01654**       .00797    -2.08  .0379     -.03216   -.00092
        |Threshold parameters for index
   Mu(1)|     .87041***      .02933    29.67  .0000      .81291    .92790
   Mu(2)|    1.32716***      .03656    36.30  .0000     1.25550   1.39883
   Mu(3)|    1.66565***      .04517    36.88  .0000     1.57712   1.75417
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

The levels of concerns are higher when the person is female,licensed and is a participant of a carshare program. Being of younger age groups is associated with lower levels of concerns.

5.4 **<u>Results Discussions:</u>**

Understanding behavioral aspect of using a new technology is an integral part of the success of that technology. Here, different models were constructed about people's interests and concerns about AVs. Different land use variables had a significant effect about certain interests of AV interests while all of them had an insignificant effect on the concerns. Some variables were consistent in their effect, for example, Age and Female have had a consistent effect of lowering the level of interest and increasing the level of concern in the sense that being in the lower age cohort has yielded lower level of concern unlike being in the oldest age cohort, although it's expected that a portion of the demand for AV technology would come from the elderly population(Bansal et al., 2016). Carshare users were also more interested in general about AVs and less concerned about this technology except for poor weather performance.

The population were more concerned about AV technology than interested. Several factors were consistent in their effects. Coming from a younger age cohort, having higher levels of education and coming from a household with higher income are big drivers to level of interest while being a female had a consistent negative effect.

In general, as AV technology is being developed several model should be incorporated to test the readability of the market and the effect of different policies on the users so that this technology can become an agent of sustainable transport systems in cities in the U.S and across the world. Different theories from psychology, social sciences, and the diffusion of innovation should be implemented. These models as well as others that are in the literature presents a peak into the interests and the concerns of the people towards this technology and what factors affect them.

6. CONCLUSION

In conclusion, cities are the center of human growth that's why special consideration should be given to solving urban problems. Transportation is a pivotal part of any major city. That's why transitioning towards a sustainable system is essential as cities continue to grow. Transportation is a major source of Greenhouse Gases (GHG) emissions in the world. Focusing on the U.S, transportation is the source of almost a third of GHG emissions (US EPA, n.d.). while cities are home to major transportation activities not all cities were created the same. The first chapter tried to classify cities according to their mobility shares, and the number of shared bikes and cars. This is to provide a roadmap for cities that they can follow in order to achieve sustainable urban transport. Four clusters were produced with the majority of U.S cities in one cluster where they are the more auto centric cities such as L.A, Dallas, and Florida. Older, denser cities formed a cluster of their own, such as New York City, and Chicago. Washington and Portland came out as a cluster of their own for their multimodal system and their support for new sharing schemes in transportation. Seattle came out as a cluster of its own. Seattle is the only city that has seen growth in Transit ridership and it also harbored one of the most successful bikeshare pilot programs in the country. An efficient public transportation system and population density is the common factor of the first three clusters while most other cities lied in the forth clusters. (L. D. Frank & Pivo, 1994) found the effect of population, employment density and mixed use-built environment has tested negative for all work and leisure trips with Single Occupancy Vehicle. This has inspired the further analysis of Mode choice in the Puget Sound Region in Seattle and the test of people's interests and concerns in Autonomous Vehicles as it's the nest emergent technology that s expected to have a disruptive effect on the Transportation network.

It’s notable the Alternative Specific Constant was positive for the mode walk when the base mode was private vehicle, which is a very good sign considering that Seattle itself is the 8 most walkable city in the U.S. (Soper, 2013). Different variables were significant that aligned with previous literature. Socioeconomics, Demographic, land use and alternative specific variables were significant and affected the utility of each mode differently. In the future, these variables can be used to determine the effect on the probability of choosing alternative mode so that congestion can be lowered by diverting demands towards public transportation in accompanying with other modes of transportation such as TNC, or Carshare.

In the future, it is very foreseeable to see Autonomous Vehicles driving people to and from public transit station or using carsharing schemes so that vehicle occupancy is higher. Either way, ordered models were developed to test people’s interests and concerns. Special attention should be focused on females and older population as these characteristics have consistently showed less interest and more concerns in AVs.

There have been many limitations to this study mainly related to data availability. The newest data obtained from the Environmental Protection Agency was released in 2010 and it is expected that there has been demographic changes to the region examined. Also, the limitation of data availability regarding the transportation modes in the Puget Sound Region (e.g., parking fees, etc.) that would describe more accurately the cost associated with each modes during the analysis.

To sum up, this can be taken as a primary look into the Puget Sound Region mode choice and AV interests and concerns. They can be used as policy tools to see the effect of each policy on the travel demand and on interest and concerns in AV technology. It can also be the basis for a more developed models to be implemented, for example tour based models (L. Frank et al., 2008).

Bibliography:

AAA. (2018). *Your Driving Costs: How Much Are You Really Paying to Drive*.

Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using ordered probit models. *Journal of Safety Research*, *34*(5), 597–603. https://doi.org/10.1016/j.jsr.2003.05.009

Administration, N. H. T. S. (2008). *National Motor Vehicle Crash Causation Survey. U.S. Department of Transportation, Report DOTHS 811 059.*

Agarwal, P. (2018). UberX Pricing by City [Raw Data] | Intelligent Economist. Retrieved October 20, 2019, from https://www.intelligenteconomist.com/uberx-pricing-data/

Aitchison, A. J., Silvey, S. D., Aitchison, B. Y. J., & Silvey, S. D. (1957). *Biometrika Trust The Generalization of Probit Analysis to the Case of Multiple Responses Published by : Oxford University Press on behalf of Biometrika Trust Stable URL : http://www.jstor.org/stable/2333245 Accessed : 18-08-2016 18 : 18 UTC Your use of th*. *44*(1), 131–140.

Alonso, A., Monzón, A., & Cascajo, R. (2015). Comparative analysis of passenger transport sustainability in European cities. *Ecological Indicators*, *48*, 578–592. https://doi.org/10.1016/j.ecolind.2014.09.022

Badoe, D. A. (2002). Modelling work-trip mode choice decisions in two-worker households. *Transportation Planning and Technology*, *25*(1), 49–73. https://doi.org/10.1080/03081060290032051

Bansal, P., & Kockelman, K. M. (2017). Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. *Transportation Research Part A: Policy and Practice*, *95*, 49–63. https://doi.org/10.1016/j.tra.2016.10.013

Bansal, P., Kockelman, K. M., & Singh, A. (2016). ASSESSING PUBLIC OPINIONS OF AND INTEREST IN NEW VEHICLE TECHNOLOGIES: AN AUSTIN PERSPECTIVE Prateek. *Transportation Research Part C*, *263*(1), 258–264. https://doi.org/10.1016/j.ejor.2017.04.057

Ben-Akiva, M., & Bierlaire, M. (1999). *Discrete Choice Methods and their Applications to Short Term Travel Decisions*. 5–33. https://doi.org/10.1007/978-1-4615-5203-1_2

Ben-Akiva, M., Morikawa, T., & Shiroishi, F. (1992). Analysis of the reliability of preference ranking data. *Journal of Business Research*, *24*(2), 149–164. https://doi.org/10.1016/0148-2963(92)90058-J

Benekohal, R. F., Michaels, R. M., Shim, E., & Resende, P. T. V. (1994). Effects of aging on older drivers' travel characteristics. *Transportation Research Record*, (1438), 91–98.

Bierlaire, M. (2018). PandasBiogeme: a short introduction. *Technical Report TRANSP-OR 181219. Transport and Mobility Laboratory, ENAC, EPFL.* Retrieved from https://wilpf.org/wilpf/who-we-are/

Blincoe, L., Miller, T. R., Zaloshnja, E., & Lawrence, B. A. (2014). The economic and societal impact of motor vehicle crashes, 2010. *Motor Vehicle Crashes: Economic and Societal Impact, In Depth*, *2015*(May), 1–285. https://doi.org/10.1016/j.annemergmed.2015.06.011

Brandon, J. (2012). Privacy concerns raised over California "robot car" legislation | Fox News. Retrieved December 23, 2019, from https://www.foxnews.com/auto/privacy-concerns-raised-over-california-robot-car-legislation

Burkhardt, J. E. (1999). Mobility changes: Their nature, effects, and meaning for elders who reduce or cease driving. *Transportation Research Record*, (1671), 11–18. https://doi.org/10.3141/1671-03

Car Sharing Association. (2011). *CarSharing Association Code of Ethics and Standards of Practice*.

2011(January), 1–15. Retrieved from http://www.carsharing.org/research/CSA_CodeofEthics.pdf

Carp, F. M. (1988). Significance of Mobility for the Well-Being of the Elderly. *Transportation Research Board Special Report*, *2*(218).

Celsor, C., & Millard-Ball, A. (2007). Where does carsharing work? Using geographic information systems to assess market potential. *Transportation Research Record,* (1992), 61–69. https://doi.org/10.3141/1992-08

Chandra, S., & Camal, F. (2016). A Simulation-based Evaluation of Connected Vehicle Technology for Emissions and Fuel Consumption. *Procedia Engineering*, *145*, 296–303. https://doi.org/10.1016/j.proeng.2016.04.077

Copperman, R. B., & Bhat, C. R. (2007). Exploratory analysis of children's daily time-use and activity patterns: Child Development Supplement to U.S. Panel Study of Income Dynamics. *Transportation Research Record*, (2021), 36–44. https://doi.org/10.3141/2021-05

Crombie, M. B. (2016). *Regulation of Transportation Network Companies*.

Daniel, L. (2002). *The Path to Discrete-Choice Models*.

Driverless Cars: Where They Stand Now | Automobile Magazine - Automobile. (2019). Retrieved December 23, 2019, from https://www.automobilemag.com/news/autonomy-levels-cars-driverless/

Economist. (2012). Look, no hands, Self-driving cars are just around the corner. *TechnologyQuarterly*.

EIA. (2019). Gasoline and Diesel Fuel Update. Retrieved November 20, 2019, from https://www.eia.gov/petroleum/gasdiesel/

EMTA. (2019). EMTA : European Metropolitan Transport Authorities. Retrieved November 16, 2019, from https://www.emta.com/?lang=en

Fagnant, D. J. (2014). *The Future of Fully Automated Vehicles: Opportunities for Vehicle- and Ride-Sharing, with Cost and Emissions Savings*. 163.

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181. https://doi.org/10.1016/j.tra.2015.04.003

Fagnant, D. J., Kockelman, K. M., & Bansal, P. (2015). Operations of shared autonomous vehicle fleet for Austin, Texas, market. *Transportation Research Record*, *2536*, 98–106. https://doi.org/10.3141/2536-12

Flynn, T. N., Louviere, J. J., Peters, T. J., & Coast, J. (2007). Best-worst scaling: What it can do for health care research and how to do it. *Journal of Health Economics*, *26*(1), 171–189. https://doi.org/10.1016/j.jhealeco.2006.04.002

Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation*, *35*(1), 37–54. https://doi.org/10.1007/s11116-007-9136-6

Frank, L. D., & Pivo, G. (1994). Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transportation Research Record*, 44–44.

Frank S. Koppelman; Chandra Bhat. (2006). *A Self Instructing Course in Mode Choice ModelinMultinomial and Nested Logit Models*. 1–249. Retrieved from papers2://publication/uuid/F79D2407-5BA7-4570-9C49-C6A4EC9C1B75

Fujii, S., & Kitamura, R. (2003). What does a one-month free bus ticket do to habitual drivers? An experimental analysis of habit and attitude change. *Transportation*, *30*(1), 81–95.

https://doi.org/10.1023/A:1021234607980

Goodin, G. (2016). *Transportation Network Companies Testimony of*.

Governing.com. (2016). Vehicle Ownership in U.S. Cities Data and Map. Retrieved November 20, 2019, from https://www.governing.com/gov-data/car-ownership-numbers-of-vehicles-by-city-map.html

Graehler, M., Mucci, R. A., & Erhardt, G. D. (2018). Understanding the Recent Transit Ridership Decline in Major US Cities: Service Cuts or Emerging Modes? *98th Annual Meeting of the Transportation Research Board*, (January), 1–19. Retrieved from http://usa.streetsblog.org/wp-content/uploads/sites/5/2019/01/19-04931-Transit-Trends.pdf

Han, J., Kamber, M., & Pei, J. (2012). 10 – Cluster Analysis: Basic Concepts and Methods. In *Data Mining: Concepts and Techniques*. https://doi.org/10.1016/B978-0-12-381479-1.00010-1

Handy, S., & Boarnet, M. G. (2014). *Impacts of Gas Price on Passenger Vehicle Use and Greenhouse Gas Emissions*. Policy Brief. Retrieved from http://www.arb.ca.gov/cc/sb375/policies/hwycapacity/highway_capacity_brief.pdf

Harper, C. D., Hendrickson, C. T., Mangones, S., & Samaras, C. (2016). Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions. *Transportation Research Part C: Emerging Technologies*, *72*, 1–9. https://doi.org/10.1016/j.trc.2016.09.003

Haurin, D. R., Dietz, R. D., & Weinberg, B. A. (2005). The Impact of Neighborhood Homeownership Rates: A Review of the Theoretical and Empirical Literature. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.303398

Hausman, J. A., & Ruud, P. A. (1987). Specifying and testing econometric models for rank-ordered data. *Journal of Econometrics*, *34*(1–2), 83–104. https://doi.org/10.1016/0304-4076(87)90068-6

Highway Traffic Safety Administration, N., & Department of Transportation, U. (2013). *Traffic Safety Facts 2013 Data Key Findings People Killed and Injured, and Fatality and Injury Rates*. (July), 1–12. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812169

Irfan, M., Khurshid, A. N., Khurshid, M. B., Ali, Y., & Khattak, A. (2018). Policy implications of work-trip mode choice using econometric modeling. *Journal of Transportation Engineering Part A: Systems*, *144*(8), 1–10. https://doi.org/10.1061/JTEPBS.0000158

J.D. Power and Associates. (2012). 2012 U.S. Automotive Emerging Technologies Study. Retrieved November 15, 2019, from https://www.jdpower.com/business/press-releases/2012-us-automotive-emerging-technologies-study

Jackman, S. (2000). *Models for Ordered Outcomes Political Science 200C Spring 2000 Simon Jackman*. 1–20. Retrieved from https://web.stanford.edu/class/polisci203/ordered.pdf

Kaur, R., & Himanshi, E. (2015). Face recognition using Principal Component Analysis. *Souvenir of the 2015 IEEE International Advance Computing Conference, IACC 2015*, 585–589. https://doi.org/10.1109/IADCC.2015.7154774

Kim, S., & Ulfarsson, G. F. (2004). Travel mode choice of the elderly: Effects of personal, household, neighborhood, and trip characteristics. *Transportation Research Record*, (1894), 117–126. https://doi.org/10.3141/1894-13

Koppelman, F. S. (1983). Predicting transit ridership in response to transit service changes. *Journal of Transportation Engineering*, *109*(4), 548–564. https://doi.org/10.1061/(ASCE)0733-

947X(1983)109:4(548)

Krueger, R., Rashidi, T. H., & Rose, J. M. (2016). Preferences for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *69*, 343–355. https://doi.org/10.1016/j.trc.2016.06.015

Kyriakidis, M., Happee, R., & De Winter, J. C. F. (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. https://doi.org/10.1016/j.trf.2015.04.014

Kyung, M. (2010). Charaterization of Cities in Seoul Metropolitan Area by Cluster Analysis. *Journal of the Korean Society for Geo-Spatial Information Science*, *18*(1), 83–88.

Levy, N. (2017). Seattle-area transit ridership growing faster than other big U.S. cities, with more options on the way – GeekWire. Retrieved November 22, 2019, from https://www.geekwire.com/2017/seattle-area-transit-ridership-growing-faster-big-u-s-cities-options-way/

Lioyd, S. A. (2018). How did Seattle’s bike-share pilot go? - Curbed Seattle. Retrieved November 19, 2019, from https://seattle.curbed.com/2018/6/19/17479724/seattle-dockless-bike-share-data

Martin Kaste. (2013). Yes, Your New Car Has A “Black Box.” Where’s The Off Switch? : All Tech Considered : NPR. Retrieved November 15, 2019, from https://www.npr.org/sections/alltechconsidered/2013/03/20/174827589/yes-your-new-car-has-a-black-box-wheres-the-off-switch

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303–328. https://doi.org/10.1016/0047-2727(74)90003-6

MMO. (2019). Observatorio de la Movilidad [OMM]. Retrieved November 16, 2019, from http://www.observatoriomovilidad.es/

Mosendz, P., & Sender, H. (2014). *EXCLUSIVE: Here's How Long It Takes to Get an Uber in U.S. Cities*. p. 1. Retrieved from https://www.newsweek.com/exclusive-heres-how-long-it-takes-get-uber-across-us-cities-289133

movmi. (2019, August 23). What Will Shared Mobility Look Like in 2020? Retrieved November 19, 2019, from Current Sustainable/Renewable Energy Reports website: http://movmi.net/shared-mobility-2020/

National Highway Traffic Safety Administration. (2011). USDOT Connected Vehicle Research Program: Vehicle-to-Vehicle Safety Application Research Plan. *U.S. Department of Transportation: National Highway Traffic Safety Administration*, (October), 1–13. https://doi.org/DOT HS 811 373

National Highway Traffic Safety Adminstration. (2014). U.S. Department of Transportation Releases Request for Comment (RFC) on Vehicle-to-Everything (V2X) Communications | NHTSA. Retrieved November 15, 2019, from https://www.nhtsa.gov/press-releases/us-department-transportation-releases-request-comment-rfc-vehicle-everything-v2x

NCSL. (2019). Autonomous Vehicles | Self-Driving Vehicles Enacted Legislation. Retrieved November 25, 2019, from http://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx

Newman, P., & Kenworth, J. (1999). Sustainability and Cities: Overcoming Automobile Dependence - Peter Newman, Jeffrey Kenworthy - Google Books. Retrieved November 16, 2019, from Book website:

https://books.google.com/books?id=pjatbiavDZYC&printsec=copyright&source=gbs_pub_info_r#v=onepage&q&f=false

Nick Shchetko. (2014). Laser Eyes Pose Price Hurdle for Driverless Cars - WSJ. Retrieved November 15, 2019, from https://www.wsj.com/articles/laser-eyes-pose-price-hurdle-for-driverless-cars-1405969441

Payre, W., Cestac, J., & Delhomme, P. (2014). Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(PB), 252–263. https://doi.org/10.1016/j.trf.2014.04.009

Pudney, S., & Shields, M. (2000). Gender, race, pay and promotion in the British nursing profession: Estimation of a generalized ordered probit model. *Journal of Applied Econometrics*, *15*(4), 367–399. https://doi.org/10.1002/1099-1255(200007/08)15:4<367::AID-JAE563>3.0.CO;2-Z

Schoettle, Brandon; Sivak, M. (2014). A survey of public opinion about autonomous and self-driving vehicles in the US, UK and Australia. *UMTRI, Transportation Research Institute*, (July), 1–38.

Seattle Department of Transportation. (2018). *2017 Free-Floating Bike Share Pilot Evaluation Report*. (August).

Shabanpour, R., Golshani, N., Shamshiripour, A., & Mohammadian, A. (Kouros). (2018). Eliciting preferences for adoption of fully automated vehicles using best-worst analysis. *Transportation Research Part C: Emerging Technologies*. https://doi.org/10.1016/j.trc.2018.06.014

Smith, L. I. (2002). A tutorial on Principal Components Analysis. *Parasite Epidemiology and Control*. https://doi.org/10.1016/j.parepi.2017.04.003

Soper, T. (2013). America’s most walkable cities: No. 1 New York; No. 2 San Francisco; No. 8 Seattle –

GeekWire. Retrieved November 28, 2019, from https://www.geekwire.com/2013/walk-score-2014/

Soper, T. (2016). Zipcar brings one-way car-sharing service to Seattle, including drop-off at new airport lot – GeekWire. Retrieved October 22, 2019, from https://www.geekwire.com/2016/zipcar-brings-one-way-car-sharing-service-seattle-including-drop-off-new-airport-lot/

Systematics, C. (2004, July 19). *Traffic Congestion and Reliability: Linking Solutions to Problems*. Retrieved from https://rosap.ntl.bts.gov/view/dot/20655

Toop, E., & Miller, E. J. (2014). *What Is a Midsize City? A Transportation Policy-Based Framework for Classifying Cities*. Retrieved from https://trid.trb.org/view/1287977

Train, K. (1978). A validation test of a disaggregate mode choice model. *Transportation Research*, *12*(3), 167–174. https://doi.org/10.1016/0041-1647(78)90120-X

Train, K. (1980). A Structured Logit Model of Auto Ownership and Mode Choice. *The Review of Economic Studies*, *47*(2), 357. https://doi.org/10.2307/2296997

Transportation, U. . D. of. (2018). *Preparing for the Future of Transportation: Automated Vehicles 3.0*. Retrieved from http://www.ghbook.ir/index.php?name=فرهنگ و رسانه های نوین&option=com_dbook&task=readonline&book_id=13650&page=73&chkhashk=ED9C9491B4&Itemid=218&lang=fa&tmpl=component

UN. (2019). Sustainable Development | UNITED NATIONS ECONOMIC and SOCIAL COUNCIL. Retrieved November 16, 2019, from https://www.un.org/ecosoc/en/sustainable-development

UNDESA. (2018). 68% of the world population projected to live in urban areas by 2050, says UN | UN DESA | United Nations Department of Economic and Social Affairs. Retrieved November 16, 2019, from https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

United Nations, P. D. (2018). World Urbanization Prospects - Population Division - United Nations. Retrieved November 16, 2019, from https://population.un.org/wup/Country-Profiles/

US EPA, O. (n.d.). *Sources of Greenhouse Gas Emissions*. Retrieved from https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

USDOT. (2012). *Traffic Safety Facts 2012 Data Truck*. (January).

Zmud, J. P., & Sener, I. N. (2017). Towards an Understanding of the Travel Behavior Impact of Autonomous Vehicles. *Transportation Research Procedia*, *25*, 2500–2519. https://doi.org/10.1016/j.trpro.2017.05.281

Zmud, J., Sener, I. N., & Wagner, J. (2016). *Consumer Acceptance and Travel Behavior Impacts of Automated Vehicles Final Report, Texas A{&}M Transportation Institute PRC 15-49 F*.

Zwerts, E., Allaert, G., Janssens, D., Wets, G., & Witlox, F. (2010). How children view their travel behaviour: A case study from Flanders (Belgium). *Journal of Transport Geography*, *18*(6), 702–710. https://doi.org/10.1016/j.jtrangeo.2009.10.002

APPENDIX A : Principle Component Analysis Code:

```
Created on Tue Apr  9 13:44:44 2019

@author: Eng.Saud Bashar
"""

#Libraries needed to run the tool
import numpy as np
np.set_printoptions(suppress=True, precision=5, linewidth=150) #to control what is printed: 'suppress=True' prevents exponential prints of numbers, 'precision=5' allows a max of 5 decimals, 'linewidth'=150 allows 150 characters to be shown in one line (thus not cutting matrices)
import pandas as pd
from sklearn.decomposition import PCA
from sklearn.preprocessing import LabelEncoder #To switch categorical letters to numbers
from mpl_toolkits.mplot3d import Axes3D
import matplotlib.pyplot as plt
import seaborn as sns
sns.set(style='darkgrid')

#Ask for file name and read the file
#file_name = raw_input("Name of file:")
file_name = 'Regional'
data = pd.read_csv(file_name + '.csv', header=0, index_col=0)

#Print number of rows and colums read
print("{0} rows and {1} columns".format(len(data.index), len(data.columns)))
print("")

#Defining all the data X1 and X2, all data X
```

```
X1=data.Pct_drv_aln #Percent Drive Alone

X2=data.Pct_carpool #Percent Carpool

X3=data.Pct_Transit #Percent Transit

X4=data.Pct_walked #Percent Walked

X5=data.Pct_taxi_moto_bike #Percent Taxi Motor Bike

X6=data.Pct_workhome #Percent Worked From Home

X7=data.VPHH

X8=data.Pop_Dens_Norm #Percent Not Drive alone

X9=data.SCP10KN

X10=data.SBP10K

X = np.column_stack((X1,X2,X3,X4,X5,X6,X7,X8,X9,X10))

#X = np.column_stack((X1, X2)) #Use only two variables to illustrate how transformation is done with two variables (with more the distances get distorted in a graph)



#Calculate and show covariance matrix

print("Covariance matrix")

print(np.cov(X, rowvar=0).round(2)) #rowvar=0 means that each column is a variable. Anything else suggest each row is a variable.

print('')

a = np.linalg.eigvals(np.cov(X, rowvar=0))

print(a/a.sum()) #To show that percentage variance explained by components is the eigenvalues

print('')



#Calculate and show correlation coefficients between datasets

print("Correlation Coefficients")

print(np.corrcoef(X, rowvar=0).round(2))

print("")
```

```
#Define the PCA algorithm
ncompo = int(input("Number of components to study:"))
print("")
pca = PCA(n_components=ncompo)

#Find the PCA
pcafit = pca.fit(X) #Use all data points since we are trying to figure out which variables are relevant

print("Mean")
print(pcafit.mean_)
print("")
print("Principal Components Results")
print(pcafit.components_)
print("")
print("Percentage variance explained by components")
print(pcafit.explained_variance_ratio_)
print("")

print(X)
X_new = pca.transform(X)
print(X_new)



#Plot percentage variance explained by components
perc = pcafit.explained_variance_ratio_
perc_x = range(1, len(perc)+1)
plt.plot(perc_x, perc,)
plt.xlabel('Components')
plt.ylabel('Percentage of Variance Explained')
```

```
plt.title('Importance of each PCA')
plt.savefig(file_name + '_pervard', dpi=300)
plt.show()
```

APPENDIX B: Hierarchical Clustering Analysis:

```
# -*- coding: utf-8 -*-
"""
Created on Tue Apr 30 11:56:33 2019

@author: Eng.Saud Bashar
"""

#Libraries needed to run the tool
import numpy as np
import pandas as pd
from sklearn.cluster import AgglomerativeClustering
from sklearn import metrics#$
from scipy.cluster.hierarchy import dendrogram, linkage #for dendrogram specifically
import matplotlib.pyplot as plt
import seaborn as sns



sns.set(style='darkgrid')

#Ask for file name and read the file
#file_name = raw_input("Name of file:")
file_name = 'SM_PCA_Data'
data = pd.read_csv(file_name + '.csv', header=0)
#data = data.set_index('CC')

#Print number of rows and colums read
print("{0} rows and {1} columns".format(len(data.index), len(data.columns)))
print("")
```

```
#Defining all the data X1 and X2, all data X
X1=data.PCA1
X2=data.PCA2
X3=data.PCA3
X = np.column_stack((X1,X2,X3))
fig = plt.figure()
fig.set_size_inches(5,5) #define the size of the figure
clusters = 4
Y_hierarchy = AgglomerativeClustering(linkage='average', n_clusters=4)
Y_hierarchy.fit(X)
Y_hierarchy_labels = Y_hierarchy.labels_
Y_hierarchy_silhouette = metrics.silhouette_score(X, Y_hierarchy_labels, metric='sqeuclidean')
print()
print()
#
print("Silhouette for Hierarchical Clustering: {0}".format(Y_hierarchy_silhouette))
# print("Hierarchical Clustering: {0}".format(Y_hierarchy_labels))
print()
print()
#print the name of each state in each cluster
def ClusterIndicesNumpy(clustNum, labels_): #numpy
  return np.where(labels_ == clustNum)[0]

C0 = ClusterIndicesNumpy(0, Y_hierarchy.labels_)
C1 = ClusterIndicesNumpy(1, Y_hierarchy.labels_)
C2 = ClusterIndicesNumpy(2, Y_hierarchy.labels_)
C3 = ClusterIndicesNumpy(3, Y_hierarchy.labels_)
C4 = ClusterIndicesNumpy(4, Y_hierarchy.labels_)
```

```
State_Array = data.CC

#Define CLuster 1 List
a = []
for x in range(len(C0)):
   a.append(State_Array[C0[x]])
print ('Cluster 1 (Black): ', a)
a.clear()
print()

#Define CLuster 2 List
for x in range(len(C1)):
   a.append(State_Array[C1[x]])
print ('Cluster 2 (Red): ', a)
a.clear()
print()

#Define CLuster 3 List
for x in range(len(C2)):
   a.append(State_Array[C2[x]])
print ('Cluster 3 (Green): ', a)
a.clear()
print()

#Define CLuster 4 List
for x in range(len(C3)):
   a.append(State_Array[C3[x]])
print ('Cluster 4 (White): ', a)
a.clear()
```

```
print()

colormap = np.array(['Black','red','green','white']) #Define colors to use in graph - could use c=Y but colors are too similar when only 2-3 clusters

plt.scatter(X1, X2, c=colormap[Y_hierarchy_labels])

plt.grid(False)

plt.annotate("s = " + str(Y_hierarchy_silhouette.round(2)), xy=(1, 0), xycoords='axes fraction', horizontalalignment='right', verticalalignment='bottom')

plt.title("Clusters:4")

plt.ylabel("Average")

fig.savefig(file_name + '_clustering2.png', dpi=300)

plt.show()
```

```
#Plot Hierarchical clustering results

data = data.set_index('CC')

linkage_types = ['ward', 'average', 'complete']

Z = linkage(X, linkage_types[2])

dendro = plt.figure()

dendro.set_size_inches(12,8)

dendrogram(Z, labels=data.index)

plt.title('Hierarchical Clustering Dendrogram')

plt.xlabel('Index from Dataframe')

plt.ylabel('Distance')

plt.savefig(file_name + '_dendro-cmp3.png', dpi=600)

plt.show()
```

APPENDIX C: Code Used to Extract Travel Information for Non-Chosen Mode :

```
# -*- coding: utf-8 -*-
"""
Created on Mon Sep  9 03:24:17 2019

@author: Eng.Saud Bashar
"""

from urllib import request
import json
from datetime import datetime
from enum import Enum

class Modes(Enum):
  DRIVING = 'driving'
  WALKING = 'walking'
  BICYCLING = 'bicycling'
  TRANSIT = 'transit'

def str_or_enum(value):
  return value.value if isinstance(value, Enum) else value;

def get_destination_info(origin: dict, destination: dict, mode: Modes = 'driving', departure_time: datetime = datetime.now()):
  url = 'https://maps.googleapis.com/maps/api/distancematrix/json?origins={origin}&destinations={destination}&departure_time={departure}&mode={mode}&key={key}'.format(
      origin='{},{}'.format(origin['x'], origin['y']),
      destination='{},{}'.format(destination['x'], destination['y']),
      departure=int(departure_time.timestamp()),
```

```
        mode=str_or_enum(mode),
        key='Google API Key'
      );

  response = request.urlopen(url).read()

  json_result = json.loads(response);
  info = json_result['rows'][0]['elements'][0]

  return info;



# Example of using it
result = get_destination_info(
  {'x': '36.2148335', 'y': '44.0130055'},
  {'x': '36.1643384', 'y': '44.025237'},
  mode=Modes.DRIVING,
  departure_time=datetime.now()
  );

print(result)
```

```
# -*- coding: utf-8 -*-
"""
Created on Tue Sep 10 15:07:04 2019

@author: Eng.Saud Bashar
"""

# -*- coding: utf-8 -*-
```

```
"""
Created on Mon Sep  9 03:24:17 2019

@author: Eng.Saud Bashar
"""

from urllib import request
import json
import numpy as np
import pandas as pd
from datetime import datetime
from enum import Enum
file_name = 'PILOT'

#Create a pandas dataframe from the csv file.
data = pd.read_csv(file_name + '.csv', header=0, index_col=0) #Remove index_col = 0 if rows do not have headers



#Print number of rows and colums ready
print("{0} rows and {1} columns".format(len(data.index), len(data.columns.values)))
print('')
type (data.index)
data.set_index('pyid',inplace=True)
class Modes(Enum):
  DRIVING = 'driving'
  WALKING = 'walking'
  BICYCLING = 'bicycling'
  TRANSIT = 'transit'
```

```
def str_or_enum(value):
  return value.value if isinstance(value, Enum) else value;


def get_destination_info(origin: dict, destination: dict, mode: Modes = 'driving,walking,bicycling,trnsit',
departure_time: datetime = datetime.now()):
  url =
'https://maps.googleapis.com/maps/api/distancematrix/json?origins={origin}&destinations={destination}&depa
rture_time={departure}&mode={mode}&key={key}'.format(
      origin='{},{}'.format(origin['x'], origin['y']),
      destination='{},{}'.format(destination['x'], destination['y']),
      departure=int(departure_time.timestamp()),
      mode=str_or_enum(mode),
      key='Google API Key'
    );

  response = request.urlopen(url).read()

  json_result = json.loads(response);
  info = json_result['rows'][0]['elements'][0]

  return info;



result = get_destination_info(
  {'x': data.O_x_gps, 'y': data.O_y_gps},
  {'x': data.D_x_gps, 'y': data.D_y_gps},
  mode=Modes.WALKING,
  departure_time=datetime.now()
```

```
);

print(result)
```

```
# -*- coding: utf-8 -*-
"""
Created on Tue Sep 10 18:58:28 2019

@author: Eng.Saud Bashar
"""
import urllib
import json
import pandas as pd
import numpy as np
import datetime, time

secret_key = 'Google API Key'

def main():
# Load TAZ records with x and y coordinates attached
taz = pd.read_csv(r'G:\Thesis\TRAVELTIME DERIVATIVES\PILOT.csv')
# Create a Google-formatted coordinates field
taz['g_coord'] = taz['y_gps'].astype('str') + ',' + taz['x_gps'].astype('str')
# set standard departure time for tomorrow at 8 AM
dep_hr = 8
dep_time = datetime.datetime.now()
dep_time = dep_time.replace(hour=dep_hr,day=dep_time.day+1)
dep_time = str(int(time.mktime(dep_time.timetuple())))
# Skims can be auto, transit, bike, or walk
mode = 'auto'
```

```
# Create empty skims to be filled with results

skims = {'auto_8_dist': np.zeros([4000,4000]), 'auto_8_time_ff': np.zeros([4000,4000]),
'auto_8_time_congested': np.zeros([4000,4000])}

# list of TAZ IDs to find data for, max size of 25 per request

taz_lists = [range(1,25)]

#taz_list = [xrange(i,i+25) for i in range(1,4000,25)]

# taz_lists = [xrange(i,i+25) for i in range(1,200,25)]

# taz_lists = [xrange(300,303)]

for taz_list in taz_lists:

# Look up for otaz in taz_list:

results = {}

urlfeed = ""

print (otaz)

origin = taz[taz['tripid'] == otaz]['g_coord'].values[1]

# get list of different destinations

destination = ''

dtaz_list = []

for dtaz in taz_list:

  if otaz != dtaz:    # skip intrazonal trips where otaz==dtaz

destination += taz[taz['tripid'] == dtaz]['g_coord'].values[0] + '|'

                    dtaz_list.append(dtaz)

 #                      remove trailing |

   destination = destination[:-1]

    urlfeed +=
"https://maps.googleapis.com/maps/api/distancematrix/json?origins="+origin+"&destinations="+destination+ \


"&mode="+mode+"&departure_time="+dep_time+"&key="+secret_key+"&units=imperial"


#                      Fetch url and store
```

```
  results[otaz] = json.loads(urllib.urlopen(urlfeed).read())   results[otaz]['dtaz_list'] = taz_list
#                    loop through each origin
#                     for otaz, data in results.iteritems():
#                    loop through each destination
#                   try:
#                    for i in xrange(len(results[otaz]['rows'][0]['elements'])):
#                   dtaz = dtaz_list[i]
#                   dist = results[otaz]['rows'][0]['elements'][i]['distance']['value']
#                   time_ff = results[otaz]['rows'][0]['elements'][i]['duration']['value']
#                   time_cong = results[otaz]['rows'][0]['elements'][i]['duration_in_traffic']['value']    # congested
#                   skims['auto_8_dist'][otaz-1][dtaz-1] = dist*0.000621371    # convert meters to miles
#                   skims['auto_8_time_ff'][otaz-1][dtaz-1] = time_ff/60    # convert seconds to minutes
#                   skims['auto_8_time_congested'][otaz-1][dtaz-1] = time_cong/60    # convert seconds to minutes
#
#                   except:
#                   print ('no values returned')
#
#                    for skimname, data in skims.iteritems():
#                    try:
#                     pd.DataFrame(data).to_csv(skimname+'.csv')
#                   except:
#                    print ('error writing to file')


if __name__ == '__main__':
   main()
```

APPENDIX D : MNL Model Code in Biogeme:

```
// File modechoiceBL.mod


[ModelDescription]

"Simple Multinomial logit choice model"



[Choice]

Mode



[Beta]

// Name        Value     LowerBound     UpperBound  status (0=variable, 1=fixed)

ASC_WALK        0       -10000       10000      0

ASC_BIKE     0       -10000       10000      0

ASC_Transit       0       -10000       10000      0

ASC_Share        0       -10000       10000      0

ASC_TNC        0       -10000       10000      0

BETA_AGE1_WALK       0       -10000       10000      0

BETA_AGE1_TRANSIT       0       -10000       10000      0

BETA_AGE1_TNC       0       -10000       10000      0

BETA_EDUCATION1_AUTO       0       -10000       10000      0

BETA_EDUCATION2_SHARE       0       -10000       10000      0

BETA_EDUCATION2_TRANSIT       0       -10000       10000      0

BETA_HHINCOME1_AUTO      0       -10000       10000      0

BETA_HHINCOME2_AUTO      0       -10000       10000      0

BETA_HHINCOME3_AUTO      0       -10000       10000      0
```

BETA_FEM_WALK 0 -10000 10000 0

BETA_FEM_BIKE 0 -10000 10000 0

BETA_LICENSE_TNC 0 -10000 10000 0

BETA_LICENSE_WALK 0 -10000 10000 0

BETA_LICENSE_TRANSIT 0 -10000 10000 0

BETA_HHSIZ_BIKE 0 -10000 10000 0

BETA_HHSIZ_TNC 0 -10000 10000 0

BETA_HHAVEHICLES 0 -10000 10000 0

BETA_HHAVEHICLES_TNC 0 -10000 10000 0

BETA_HHAVEHICLES_TRANSIT 0 -10000 10000 0

BETA_HHCARSHARE_SHARE 0 -10000 10000 0

BETA_HHCARSHARE_AUTO 0 -10000 10000 0

BETA_HHCARSHARE_BIKE 0 -10000 10000 0

BETA_HHOFFPARK_BIKE 0 -10000 10000 0

BETA_HHOFFPARK_AUTO 0 -10000 10000 0

BETA_HHSTREETPARKPERMIT 0 -10000 10000 0

BETA_OD1A_AUTO 0 -10000 10000 0

BETA_OD2A_JPHH 0 -10000 10000 0

BETA_OD4A 0 -10000 10000 0

BETA_OD5DRI_AUTO 0 -10000 10000 0

BETA_OD5DRI_TRANSIT 0 -10000 10000 0

BETA_OD5DRI_TNC 0 -10000 10000 0

BETA_DD1A_TNC 0 -10000 10000 0

BETA_DD1A_AUTO 0 -10000 10000 0

BETA_DD2A_JPHH 0 -10000 10000 0

BETA_DD5AE_TRANSIT 0 -10000 10000 0

BETA_DD5DRI_AUTO 0 -10000 10000 0

BETA_DD5DRI_TRANSIT 0 -10000 10000 0

BETA_DD5DRI_TNC 0 -10000 10000 0

BETA_WALKTIME 0 -10000 10000 0

BETA_BIKETIME 0 -10000 10000 0

BETA_DRIVETIME 0 -10000 10000 0

BETA_TRANSITTIME 0 -10000 10000 0

BETA_CARSHARETIME 0 -10000 10000 0

BETA_DRIVECOST 0 -10000 10000 0

BETA_TNCCOST 0 -10000 10000 0

[Utilities]

// Id Name Avail linear-in-parameter expression (beta1*x1 + beta2*x2 + ...)

1 Walk one ASC_WALK * one + BETA_FEM_WALK * Female + BETA_WALKTIME * Walktime + BETA_AGE1_WALK * AGE1 + BETA_LICENSE_WALK * License

2 Bike one ASC_BIKE * one + BETA_FEM_BIKE * Female + BETA_BIKETIME * Biketime + BETA_HHSIZ_BIKE * HHsize + BETA_HHCARSHARE_BIKE * HHcarshare + BETA_HHOFFPARK_BIKE * HHoffpark

3 Car one BETA_HHOFFPARK_AUTO * HHoffpark + BETA_HHAVEHICLES * HHavailvehicles + BETA_HHSTREETPARKPERMIT * HHstparkpermit + BETA_OD5DRI_AUTO * OD5DRI + BETA_DD5DRI_AUTO * DD5DRI + BETA_DRIVETIME * Drivetime + BETA_DRIVECOST * Drivecost + BETA_HHCARSHARE_AUTO * HHcarshare + BETA_OD1A_AUTO * OD1A + BETA_DD1A_AUTO * DD1A + BETA_EDUCATION1_AUTO * EDUCATION1 + BETA_HHINCOME1_AUTO * HHINCOME1 + BETA_HHINCOME2_AUTO * HHINCOME2 + BETA_HHINCOME3_AUTO * HHINCOME3

4 Transit one ASC_Transit * one + BETA_LICENSE_TRANSIT * License + BETA_OD4A * OD4A + BETA_OD2A_JPHH * OD2A_JPHH + BETA_DD2A_JPHH * DD2A_JPHH + BETA_TRANSITTIME * Transittime + BETA_HHAVEHICLES_TRANSIT * HHavailvehicles + BETA_OD5DRI_TRANSIT * OD5DRI + BETA_DD5DRI_TRANSIT * DD5DRI + BETA_DD5AE_TRANSIT * DD5AE + BETA_AGE1_TRANSIT * AGE1 + BETA_EDUCATION2_TRANSIT * EDUCATION2

5 Share one ASC_Share * one + BETA_HHCARSHARE_SHARE * HHcarshare + BETA_CARSHARETIME * Carsharetime + BETA_EDUCATION2_SHARE * EDUCATION2

6 TNC one ASC_TNC * one + BETA_HHAVEHICLES_TNC * HHavailvehicles + BETA_TNCCOST * TNCcost + BETA_HHSIZ_TNC * HHsize + BETA_DD1A_TNC * DD1A + BETA_OD5DRI_TNC * OD5DRI + BETA_DD5DRI_TNC * DD5DRI + BETA_LICENSE_TNC * License + BETA_AGE1_TNC * AGE1

[Expressions]

// Define here arithmetic expressions for name that are not directly

// available from the data

one = 1

[Model]

$MNL

// Description of the choice subsets to compute the new

// variable for McFadden's IIA test

// Name list_of_alt

C123 1 2 3

C345 3 4 5