Executive Summary

In recent years, car-sharing has become an increasingly attractive option for consumers. Its strengths align with the major concerns of our time: economic stability, personal autonomy, and environmental protection. Additionally, many Americans seem to be letting go of some previously held attitudes about cars as symbols of status and freedom, and opting instead for this less-costly and more environmentally friendly alternative [22].

Of course, car-sharing services have a long way to go before becoming the dominant mode of transportation. In this paper, our task was to develop models that categorized the driving habits of US residents, to predict the performance of different types of car-sharing services in various cities, and to account for the effect of new technologies, namely environmentally friendly autonomous vehicles, on the usage and adoption of car-sharing.

First, we sought to better evaluate the US population's attitude toward car-sharing. Time of car usage and miles driven per day were identified as main factors that drivers considered when looking at car-sharing services. We identified the primary factors that affect time of car usage and miles driven per day – age, location, and gender [9], and used these factors to determine the percentage of the US driving population that drive a low, medium, and high amount of time and low, medium, and high amount of distance.

To rank the types of car sharing models and their usefulness within each city we multiplied the various factors that affect participation in car sharing, with each factor raised to its correlation coefficient as a method of weighting them. This resulted in a clear ranking of one way floating system as the best method of car sharing and Poughkeepsie, NY as the best tested city within which to create or expand a car-sharing company, with an estimated participation percentage of 4.01%.

Lastly, we factored in the effect of the introduction of environmentally friendly autonomous vehicles by weighting each city by its attitude towards environmental protection. We assessed the city's concerned for the environment by finding the size of the left-leaning population, who have been shown to hold more environmentally-conscious views [21]. This again resulted in Poughkeepsie, NY being ranked as the best city, followed closely by Richmond, VA.

Clearly, car-sharing addresses the concerns of an evolving society, with a greater need for mobility and accessibility of transportation. As the future brings new and unexpected developments, we can only guess at the true success of car-sharing, however we believe our models can provide helpful and accurate guidelines for the decision-makers of today and tomorrow.

Share and (Car) Share Alike

Team 7867

February 27, 2016

Contents

1 Introduction

1.1 Background

While in the past car ownership was seen as a sign of status and independence, today it is often eschewed for more cost effective and convenient modes of transportation. For many, that means car-sharing, which is generally defined as a short-term rental system based off of mileage travelled and/or time used [17].

Even though only a relatively small percentage of US drivers currently use car sharing with any regularity, the market is expected to grow rapidly in the near future [22]. The potential growth of car sharing has enticed companies as large as Ford and General Motors to invest hundreds of millions into developing their own car sharing businesses [5].

1.2 Restatement of the Problem

To address the increased interest in the car-sharing field, and the increased need for accurate, yet often unattainable data, we have developed models to estimate the following:

1. Who's driving? How much and how far do US drivers drive per day?

2. What car sharing business types are best? Which would be most popular with consumers? In which locations would each model excel? What cities would be best to establish a new car sharing business?

3. How would the integration of environmentally friendly autonomous vehicles impact the car sharing market?

2 Who's Driving?

2.1 Analysis of the Problem

We are tasked with splitting the US population into nine sections based on their average daily driving time and mileage. While likely related, time and mileage are not necessarily directly correlated due to differences in the average driving speed of different locations across the United States. To develop our model we made the following assumptions and simplifications:

2.2 Assumptions, Simplifications, and Justifications

- **Assumption:** The primary contributing factors of driving habits of US citizens are gender, location, and age group.
  Justification: Relative to the difference in driving time and distance produced by the above three factors, the difference produced by other factors is relatively insignificant. For instance, while the difference between the average daily driving time of males and females is ten minutes, the difference between white and African American drivers is merely five minutes. Furthermore, gender, location, and age group encompass many of the other factors that surveys have found to cause a difference in daily driving time and distance. For example, although surveys have shown a relationship between time driven and family size, this is encompassed by the location variable, as larger families often move to suburban areas [9].
- **Assumption:** If given a population percentage for a larger age range than necessary (e.g. given percent of population ages 15-19 while looking for the percent of population ages 16-19) then we will assume that the ages are distributed linearly within the age group (e.g. 16-19 is four-fifths of the 15-19 population).
  Justification: The US population will not vary largely between close ages.
- **Simplification:** Calculations are performed as if gender, location, and age group are statistically independent.
  Justification: The lack of obvious and significant correlation between gender, location, and age group allows us to treat them as if they were independent [9].

2.3 Design of the Models

The model was designed in two parts. The first part determines which combinations of gender, location, and age are considered "low," "medium," and "high" for both minutes and miles of driving daily. The second part uses census data about the proportions of the US driving population that fit into each category of gender, location, and age to calculate the proportion of US drivers that fit into each classification of driving habits.

2.3.1 Driving Habits of Demographic Classifications

For each demographic group (e.g. male drivers, drivers between ages 16-19), we standardized the miles and minutes driven daily by taking the ratio of the average value for each group and the average of the entire population. In effect, this gives weights each variable relative to the mean value of that variable. As shown below, these equations are:

$$T_s = \frac{Daily\ Duration\ of\ Driving\ Trips\ of\ Demographic\ X\ \ (minutes)}{Daily\ Duration\ of\ Driving\ Trips\ of\ Total\ Population\ (minutes)} \text{and}$$

$$M_s = \frac{Miles\ Driven\ Daily\ of\ Demographic\ X\ \ (miles)}{Miles\ Driven\ Daily\ of\ Total\ Population\ (miles)}.$$

These equations yield the values:

Demographic	T_s	M_s
Male	1.1	1.15
Female	0.89	0.853
Urban Residents	0.93	0.911
Rural Residents	1.1	1.15
Age 16-19	0.61	0.675
Age 20-29	1.1	1.06
Age 30-49	1.2	1.23
Age 50-64	1.0	1.03
Age 65+	0.82	0.719

Using these values, we determined a standardized score for each of the 20 combinations of the 3 demographics by taking the geometric mean of their 3 standardized scores. This was performed twice, once for T_s and once for M_s. As shown below, these equations are:

$$P_T = \sqrt[3]{Ts_1 \times Ts_2 \times Ts_3} \text{ and}$$

$$P_M = \sqrt[3]{Ms_1 \times Ms_2 \times Ms_3}.$$

These equations yield the values:

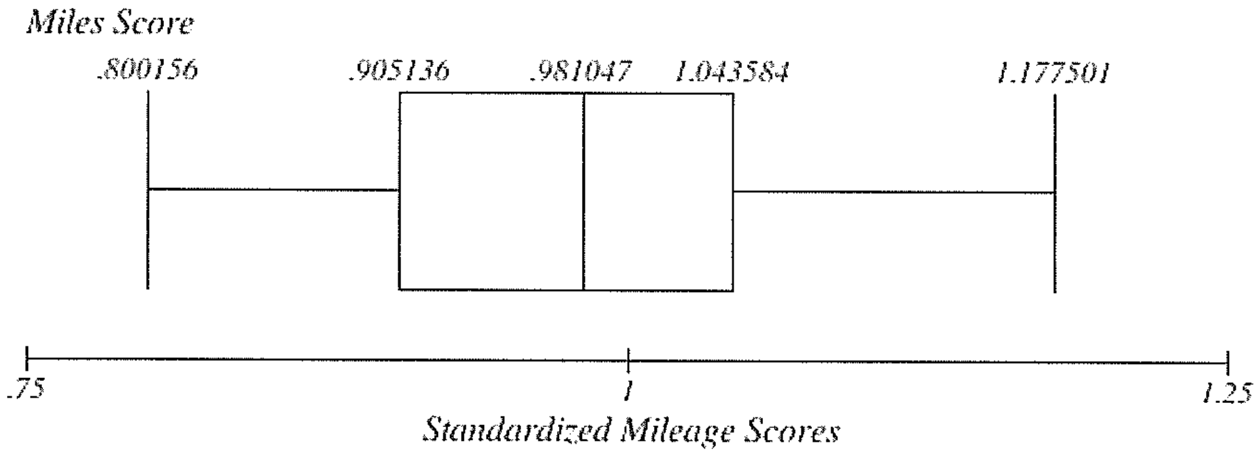
Duration Scores (P_T)

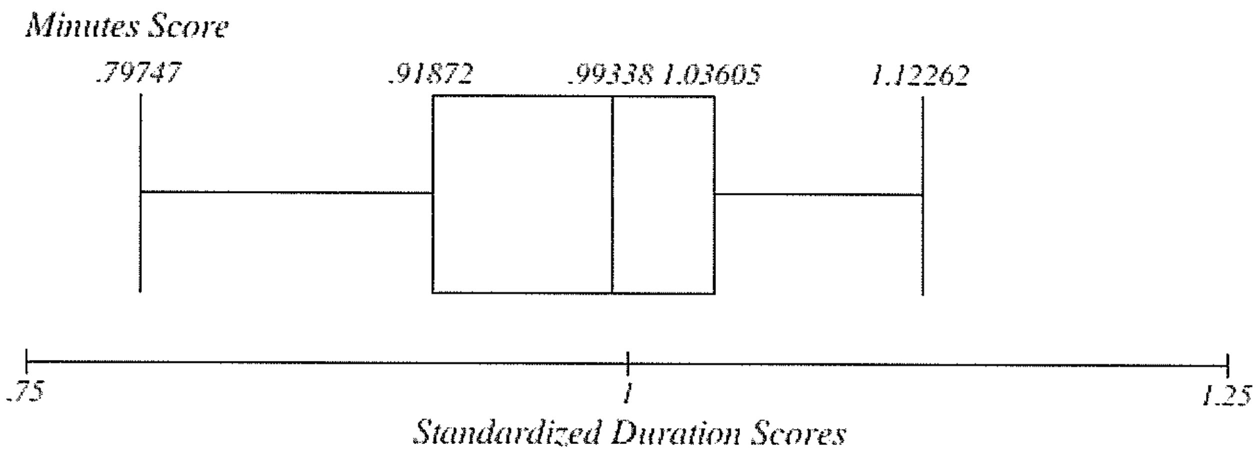
	Ages 16-19	Ages 20-29	Ages 30-49	Ages 50-64	Ages 65+
Urban Male	0.86	1.03	1.07	1.02	0.95
Rural Male	0.90	1.09	1.12	1.07	0.99
Urban Female	0.80	0.96	0.99	0.95	0.88
Rural Female	0.84	1.01	1.04	1.00	0.92

Mile Scores (P_M)

	Ages 16-19	Ages 20-29	Ages 30-49	Ages 50-64	Ages 65+
Urban Male	0.891	1.04	1.09	1.02	0.910
Rural Male	0.963	1.12	1.18	1.11	0.984
Urban Female	0.800	0.931	0.978	0.921	0.817
Rural Female	0.872	1.01	1.07	1.00	0.890

This generated the following box plots





- We assigned the groups whose minute or mile score was lower than or equal to the first quartile of the scores to the category of "low."
- We assigned the groups whose minute or mile score is greater than or equal to the third quartile of the scores to the category of "high."
- The groups whose scores fell within the interquartile range were assigned to the category of "middle"

Here are the final tables for minutes and miles driven daily, with each demographic combination categorized into low (green), medium (white), and high (red).

Duration Scores (P_T)

	Ages 16-19	Ages 20-29	Ages 30-49	Ages 50-64	Ages 65+
Urban Male	Low	Med.	High	Med.	Med.
Rural Male	Low	High	High	High	Med.
Urban Female	Low	Med.	Med.	Med.	Low
Rural Female	Low	Med.	High	Med.	Med.

Mile Scores (P_M)

	Ages 16-19	Ages 20-29	Ages 30-49	Ages 50-64	Ages 65+
Urban Male	Low	Med.	High	Med.	Med.
Rural Male	Med.	High.	High	High	Med.
Urban Female	Low	Med.	Med.	Med.	Low
Rural Female	Low	Med.	High	Med.	Low

2.3.2 Percentages of US Driving Population in each Category

Using data from the US Census, we determined the percentages of the driving population of the US that fit into each category by multiplying the proportions by gender, age, and locations of the driving population using the following equation:

$$P = (\%\ of\ US\ Driving\ Population\ in\ Age\ Group)\times (\%\ of\ US\ Driving\ Population\ in\ Gender)\times (\%\ of\ US\ Driving\ Population\ in\ Area\ Type\ (Urban\ or\ Rural)).$$

For example, for Urban Males in the age group 20-29:

1. 17.10% of US driving population between ages of 20-29
2. 49.74% of US driving population is male
3. 79.8% of US driving population lives in urban areas
4. $P = 0.1710\times0.4974\times0.7980 = 0.0679$
5. The proportion of drivers who are 20-29 aged males who live in urban areas is .0679.

This yielded the following tables:

Low Rating = Green
Middle Rating = White

High Rating = Red

Population Proportions for D_s

	16-19	20-29	30-49	50-64	65+
Urban Male	0.0175	0.0679	0.143	0.104	0.0635
Rural Male	0.00418	0.0162	0.0343	0.0248	0.0152
Urban Female	0.0180	0.0701	0.148	0.107	0.0656
Rural Female	0.00431	0.0168	0.0354	0.0256	0.0157

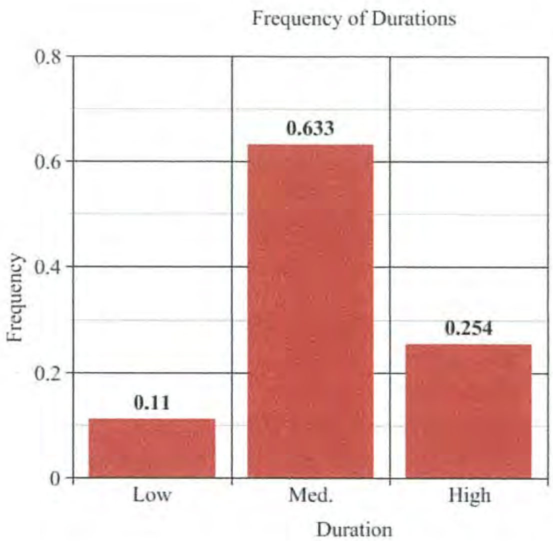
Population Proportions for M_s

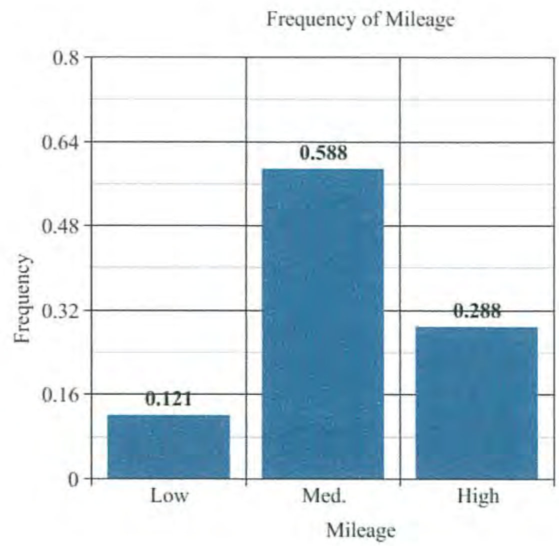
	16-19	20-29	30-49	50-64	65+
Urban Male	0.0175	0.0679	0.143	0.104	0.0635
Rural Male	0.00418	0.0162	0.0343	0.0248	0.0152
Urban Female	0.0180	0.0701	0.148	0.107	0.0656
Rural Female	0.00431	0.0168	0.0354	0.0256	0.0157

To find the proportion of the US population that drives with "low" mileage and minutes we summed the population proportion for each demographic that was considered low. We repeated this process to determine the proportions for medium and high.

Duration	Proportion
Low	0.110
Medium	0.633
High	0.254

Mileage	Proportion
Low	0.121
Medium	0.588
High	0.288





Finally, we multiplied the proportion of drivers in each duration rating by a proportion of drivers in a distance rating to create a two-way table of the time-distance proportion in the US driving population.

- For example:
    - Low minutes proportion × Low mileage proportion = Low minutes and low mileage proportion
    - $0.110 \times 0.121 = 0.013 = 1.3\%$

	Low Duration	**Medium Duration**	**High Duration**
Low Mileage	1.3%	7.7%	3.1%
Medium Mileage	6.4%	37%	15%
High Mileage	3.2%	18%	7.3%

2.5 Testing and Verification

Sensitivity Analysis

If the proportions of people of a certain gender, age group, or location are modified to reflect the demographics of a different population, the model produces driving habit proportions as shown in the below analyses.

- By increasing the size of age group 30-49 by 10% or by 0.0361, the proportions of drivers in each category change at most by 21% in the *low duration, low distance* category and by 12.5% and lower in the other 8 categories.
- By increasing the proportion of male drivers by 10% or by 0.04955, the proportions of drivers in each category change at most by 15.6% in the *high duration, high distance* category and by 10.1% and lower in the other 8 categories.
- By increasing the proportion of urban drivers by 10% or by 0.0798, the proportions of drivers in each category change at most by 14.8% in the *high duration, high distance* category and by 9.8% and lower in the other 8 categories.

3 Zippity Do or Don't?

3.1 Analysis of the Problem

We were tasked with determining which of the four listed car-sharing business options would lead to the greatest customer participation in a given city. We then had to rank each city by participation. In order to do so, we had to come up with a model that took into consideration such characteristics as motor-vehicle density, car ownership levels, and individuals who walked to work. To do so, we also made the following assumptions and simplifications:

3.2 Assumptions, Simplifications, and Justifications

- **Assumption:** The probability that a customer will use a car sharing station is 1 at a distance of 0 and 0 at a distance of ½ miles.
  Justification: Since the Transit Cooperative Research Program has shown that customers are willing to walk a maximum of ½ of a mile to a car sharing station, if someone is ½ miles or more away from a station, the likelihood that they will use the carsharing station is very low. If they are a distance of 0 away from the car-sharing station (as in, at a car-sharing station) they will likely use the station. [2]
- **Simplification:** Factors with unknown correlations will be assumed to have a correlation of 0.5.
  Justification: Given an unknown correlation, the safest correlation to use would be the average of the two extremes (.5).
- **Simplification:** Although we were asked to rank four types of car-sharing (round trip, one-way station, one-way floating, and fractional ownership), we simply ranked one-wa station and one-way floating.
  Justification: A one-way station is an objectively better form of the round trip model for the consumer, meaning it will always perform better than the round trip model, rendering

it obsolete. We discounted fractional ownership as it generally only succeeds in small-scales in rural communities [2].

3.3 Ranking Different Car Sharing Models

In general, customers prefer one way floating, which has shown much greater growth in actual markets. Our analysis supports this conclusion.

- Surveys [2] have shown that consumers are willing to walk a maximum of roughly ½ of a mile to a car sharing station. This applies both to round-trip and one way station programs. We therefore assumed the probability that a customer will use a car-sharing station is 1 at a distance of 0 and 0 at a distance of ½ miles. We used the following logistic growth equation to model the average distance that a station based car sharing user would be from a station:

$$\int_0^{\frac{1}{2}} \frac{1}{1+e^{(20x-4)}}\,dx = \frac{1}{4}$$

- We also considered a more simplistic geometric argument to examine the reasonability of our result: if a station based car sharing company wants to cover all potential users in a continuous area, and no location can be more than ½ mile from a station, and if a person is randomly placed in the area, they ought to be roughly half that distance, ¼ mile, from the nearest station. With the many other factors involved, we went with ¼ as our estimate.

- Companies offering one way station based services often stress that they have an advantage in areas where parking is difficult. A one-way station based car sharing system would only be more convenient if most of the time no parking spot could be found within the distance it takes to drive to the nearest station. The average street parking spot in the United States is 18 feet in length [24] which, if spaces were lined up in a row with no gaps for driveways or intersections, would give 73.3 parking spaces per ¼ mile on one side of the street, 146.7 on both sides. Because parking spaces are not continuous, we used a conservative estimate of 70 parking spaces per ¼ mile. The odds of there being at least one open street parking space among 70 is

$$1 - n^{70}$$

where n is the probability that any one parking space is filled. Thus, a station based service is more convenient for n such that

$$1 - n^{70} < \frac{1}{2}$$

which gives

$$n > .9901\ldots$$

Thus, even if 99% of street parking spaces are full, a one way floating system is still usually more convenient for the customer, which is consistent with actual consumer preferences [23].

3.4 Ranking Cities

We used several different factors in considering which cities were most likely to be able to host successful car sharing programs. Using data from the TCRB report sponsored by the Federal Transit Authority [2] we discovered that the demographic statistics most correlated with success were the percent of households with no vehicles, the percent of individuals walking to work, and the percentage of one person households, with positive correlations, and the number of vehicles per household, with a negative correlation. In addition, we decided to include the proportion of individuals falling with demographics that we considered "low, low" from section 2 of this report, giving those a weight equivalent to a correlation of 0.5, as we were unsure of the real correlation. We used the correlations of each variable to weight them, giving our demographic factor *d* for each city to be

$$d = \prod_{i=1}^{n} \quad (a_i)^{c_i} \ ,$$

where a_i is the proportion of each city falling within each demographic (e.g proportion of one person households) and c_i is the correlation of that statistic with car-sharing program success. We also took into account the amount of college students in the vicinity of each city, as while college students are not counted in census data for a city they are large users of car-share programs, with participation rates of 10-15% [13]. Thus, to get an estimate of the percentage of each city's residents, including college students, who would participate in a car-sharing program we used the following formula:

$$\frac{100(d(P)+.10(C))}{(C+P)} \ ,$$

where *d* is the demographic factor from the previous equation, P is the city population, and C is the student population of nearby universities. In calculating for each of the four given cities, we chose to exclude populations of community colleges, as their students are predominately from the immediate area, and thus would already be counted in the demographic statistics.

This gave the following predictions for each of the four cities:

Poughkeepsie, NY	4.01%
Knoxville, TN	2.51%
Richmond, VA	2.16%
Riverside, CA	1.52%

Poughkeepsie's numbers are bolstered by its high density of housing units per acre, its high percentage of residents who walk to work, and the several large college populations nearby, including Vassar and Marist, relative to its small population. On the other hand, Riverside has a high number of vehicles per household and a low number of residents who walk to work, giving it a lower score.

3.5 Assessing the Reasonability of the Model

We also tested data for Traverse City, MI in our model. Traverse City is notable in that an attempt at introducing a car-sharing program failed [2], so we should expect our model, if it is accurate, to give it a very low score. Our model indeed gave a value of only 0.53%, much lower than any of the given cities, implying that our model is consistent with the failure of this program. Our model seems to indicate that this is the result of a low housing density, the lack of a large nearby college population, and many of the same factors as Riverside.

4 Road Map to the Future

4.1 Analysis of the Problem

We are tasked with determining how the introduction of environmentally friendly autonomous vehicles would affect our previous rankings of cities and car sharing business types. To do so, we made the following assumptions and simplification:

4.2 Assumptions, Simplifications, and Justifications

- **Assumption:** The population proportion that cares deeply about environmental issues is roughly the same as the population with left-leaning views.
  Justification: Views on environment and politics and environment are closely linked, with those holding more left-leaning views generally caring more deeply about the environment [21].
- **Assumption:** Existing car sharing business types would become irrelevant and do not need to be ranked.
  Justification: The existing car sharing models are predicated on the idea that cars must either be returned directly to a station by the customer themselves (round trip station and one way station), or else they must be parked somewhere in the streets (one way floating). As an automated car can simply drive itself to a station after the customer has left, it would negate the need for customers using the one way floating model to park the car, which is its primary downside. Therefore, all autonomous car-sharing companies would take on a combined one way floating and one way station business model.
- **Simplification:** Factors with unknown positive correlations will be assumed to have a correlation of 0.5.
  Justification: Given an unknown positive correlation, the safest correlation to use would be the average of the two extremes (.5).

4.3 Design of the Model

Given that we already had the score for each city, it was relatively simple to take into account the effect of the introduction of environmentally friendly autonomous cars. We simply took the percentage of citizens with left-leaning views raised to the one-half (their correlation) and multiplied it by the previous score of the city, thereby weighting each city by its environmental views. This resulted in the following rankings of each city:

1. Poughkeepsie, NY (1.99)
2. Richmond, VA (1.89)
3. Knoxville, TN (1.15)
4. Riverside, CA (.56)

5 Strengths and Weaknesses

5.1 Strengths

- Our models incorporate complete census data describing the whole population accurately as opposed to sample data which could misrepresent the population.

- Several factors were considered in the development of each model, ensuring that they respond to small changes in environment accordingly and better approximate real-life scenarios.

5.2 Weaknesses

- Our model assumes the independence of proportions whose independence is not guaranteed, especially in the assumption of independence between driving duration and driving mileage.
- Our model for the impact of electric and self-driving vehicles on the use of car-sharing assumes that the main factor is environmental consciousness, as indicated by political affiliation. This is not an ideal indicator because political affiliation is only weakly associated with attitudes regarding conservation.

6 Further Work

Given additional time, we would have liked to explore the effects of self-driving vehicles on the car-sharing market. It would have been very interesting to trace the relationship between autonomous cars and the need for sharing stations. Although much of the technology needed for this vision to become a reality has yet to be developed, it is safe to say that a car sharing business based on self-driving cars would dramatically increase fleet efficiency and customer satisfaction. Also, we would have liked to have been able to collect more reliable data on the usage of Zip Cars across the United States. This would have allowed us to better determine the accuracy of our predictions for Section 3.

7 Conclusion

In this report, we developed solutions to categorize the United States population by their average driving time and driving distance per day, to rank the effectiveness of various car sharing methods in various cities, and to take into account the effect of future technological developments such as autonomous cars and environmentally friendly cars. As a result of our models, we were able to determine that the one way floating model of car sharing outperforms other types of car sharing in every scenario. We also determined that Poughkeepsie, NY would likely be the best tested city to open up a new car sharing company, as expected participation could reach as high as 4.01%. Finally, we concluded that if environmentally friendly self-driving cars were to be introduced, Poughkeepsie, NY would remain the best city to open up a new car sharing company, but would be followed closely by Richmond, VA.

8 References

[1] http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf

[2] http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_108.pdf

[3] http://richmondva.areaconnect.com/statistics.htm

[4] http://nhts.ornl.gov/tables09/FatCat.aspx

[5] http://www.businessinsider.com/r-ford-rd-chief-says-automaker-wants-to-develop-ride-hailing-services-2015-12

[6] http://nhts.ornl.gov/tables09/ae/work/Job44289.html

[7] https://www.fhwa.dot.gov/policyinformation/statistics/2010/pdf/dl20.pdf

[8] https://www.census.gov/prod/2011pubs/acs-15.pdf

[9] https://www.aaafoundation.org/sites/default/files/2015AmericanDrivingSurveyRerport.pdf

[10] http://www.autorentalnews.com/channel/rental-operations/article/story/2015/03/carsharing-state-of-the-market-and-growth-potential.aspx

[11] https://riversideca.gov/planning/pdf/demographics/City_OK2010.pdf

[12] http://www.ucr.edu/

[13] http://www.autorentalnews.com/article/story/2011/11/car-sharing-grading-the-college-campus-market/page/1.aspx

[14] http://www.calbaptist.edu/

[15] https://lasierra.edu/

[16] http://www.ucr.edu/

[17] http://tsrc.berkeley.edu/carsharing

[18] http://city-stats.org/

[19] http://www.brookings.edu/~/media/Files/rc/papers/2011/0818_transportation_tomer/0818_transportation_tomer.pdf

[20] http://www.rita.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/index.html

[21] http://phys.org/news/2015-07-political-views-strongly-linked-attitudes.html

[22] http://www.autoblog.com/2013/09/17/12m-carsharing-users-predicted-by-2020-today-its -2 - 3m/

[23] http://sf.streetsblog.org/2010/03/29/san-francisco-first-city-in-the-nation-to-count-its-parking-spaces/

[24] http://www.danbury-ct.gov/filestorage/21015/21087/21123/23014/Page8-1112.pdf