NC School of Science and Mathematics– Team #7412
Durham, North Carolina
Coach: Dan Teague
Students: Simon Marlan, Vinit Ranjan, Katherine Yang,
Sunwoo Yim, Kelly Zhang

M3 Challenge Finalist, $5,000 Team Prize

Summary - #7412

Car-sharing has recently become an increasingly popular way to commute within the United States. It remove the burdens of car ownership and reduces pollution, traffic, parking limits, and other issues that arise in populations where sole car ownership is predominant. As a result, new companies like Zipcar along with automakers such as Ford Motor Co have begun investing heavily in the car-sharing industry, looking for ways to expand their markets using different car-sharing plans.

We built a mathematical model to characterize current US drivers based on the miles they drive per day and their time spent driving per day. We found that most drivers had a medium daily distance and a medium daily driving time, though many had a high daily distance and medium daily driving time or a medium daily distance and high daily driving time. Interestingly, very few drivers had a low daily distance and low daily driving time. Two numerical ways of determining the nine categories of low/medium/high daily distance (in miles) and low/medium/high daily driving time (in minutes) were used. The final chosen method relies on the idea that time is dependent on distance to determine the percentage of all US drivers who belong to each category. This model does this intrinsically without actually finding the values bounding each of the nine categories. An additional, more rigorous strategy using statistics and probabilities was also found to calculate the percentages in each category but would require data collection that was not available during this investigation. Splitting the two considerations (miles and time) into three separate categories by finding intersections, we can find the values of the boundaries for each of these (low, medium, and high) and find percentages.

Given the four different car-sharing options—round-trip, one-way floating plan, one-way station plan, and fractional ownership—we then created a computational agent-based model to simulate the car-sharing preference for populations of city drivers. We determined that the fractional ownership would be the method of car-sharing most used by US drivers, but companies would gain the most profit by implementing the one-way car-sharing station model, which would have the most participation from consumers in any given city. Out of the four cities—Poughkeepsie, NY; Richmond, VA; Riverside, CA; and Knoxville, TN—Knoxville was the best option for implementing the car-sharing plan of one-way floating. There was little preference between the other three cities, which were not statistically significant when the best car-sharing option was considered individually for each city.

As the technology necessary for self-driving cars and cars with alternative fuel or renewable energy improves, we modified our model with these changes, and our results suggest that when companies switch to these more modern vehicles, they also switch to the one-way floating model, and again implement these changes at Knoxville for the greatest estimated profit.

Our model considers current characterization of US drivers and considers multiple aspects of driving populations and their interactions with the car-sharing industry.

Sharing is Car-ing
Modeling Car-Sharing

Team #7412

February 28, 2016

Table of Contents

Introduction

Background

The United States has recently seen an increase in car-sharing plans. Across the nation, US drivers are opting to rent cars or invest in fractional ownership of a private car rather than own a car. In addition to decreases in traffic, pollution, and fuel usages, car-sharing allows drivers to avoid dealing with the costs and issues of being the sole owner of a car. In response to this increasingly popular practice, companies want to investigate cities to determine which locations would garner more participation, depending on the type of car-sharing plan available, and which locations would be best for a company to expand its car-sharing services to [1,2].

Restatement of the Problem

Given the increasing popularity of car-sharing programs, we have responded to the needs of US drivers and car-sharing business to develop a model to solve the following problems:

1. What percentage of current US drivers have low, medium, and high daily distance and daily driving times? The distance in miles driven per day and the amount of time spent driving each day are two of the main factors used when drivers consider using car-sharing.
2. Consider the four car-sharing business models: round-trip car-sharing, one-way floating car-sharing, one-way station car-sharing, and fractional ownership [2]. First, determine which car-sharing option would garner the most participation in a given city. Next, rank the cities to find which is most prospective for a car-sharing company to expand to—if a company wanted to establish car-sharing, which of the four cities of Poughkeepsie, NY; Richmond, VA; Riverside, CA; and Knoxville, TN, would return the greatest profit for the company?
3. If a company wanted to establish car-sharing involving self-driving cars and green-energy powered cars, which of the four cities of Poughkeepsie, NY; Richmond, VA; Riverside, CA; and Knoxville, TN, would return the greatest profit for the company? [1,3]

1. Who's Driving?

We were tasked with developing a model of driving considerations that involves two of the main factors related to a driver's decisions about car-sharing: the amount of time using the car and miles driven per day. This model was used to determine the percentage of current drivers in the US for each combination of three categories—low, medium, and high—for the two metrics, resulting in nine different characterizations of car usage.

Assumptions, Decisions, and Justifications

- **The data reported in the 2015 AAA American Driving Survey is representative of current US drivers.** We specifically assumed that the average data reported for each age range surveyed is representative of US drivers within that age range. The Survey used dual-frame sampling (landline and cell phone) as well as established screening and phone-based survey procedures, callbacks, and other methodology intended to reduce bias skewing of the survey population and responses in any way [4].

- **In the relationship between driving time and miles driven, the time taken is dependent on the miles.** A driver travels with a set destination. The time spent driving to their destination is dependent on the physical distance, as well as other factors such as traffic and infrastructure design.

Methods

Using data from the AAA's American Driving Survey, published April 2015, we determined the percentage of US drivers in each characterization. For example, we found the percentage of drivers that spend a low amount of time using the car and drive a medium amount of miles per day. For the remainder of our solution to question 1, we refer to daily distance and daily driving time as "considerations," the groups of low, medium, and high as "categories," and the combinations of the groups as "characterizations."

The AAA surveyed 3319 US drivers above the age of 16, with demographic ranges of 16-19, 20-29, 30-49, 50-64, 65-74, and 75+ years of age. Within the survey, the AAA statistically analyzed each demographic and indicated when an age range was "significantly different than the overall estimate, at the 95% confidence level" [4]. We assumed that the US drivers surveyed within an age range were representative of all US drivers within that age range. The values of average daily number of miles driven, average daily duration of driving trips and number of total US drivers are reported for each age range (Table 1).

Age Range	Average daily distance (miles)/person	Average daily driving time (min)/person	Number of US drivers
Overall	29.2	46	214092472
16-19	19.7	28	8429221
20-29	30.7	49	36289088
30-49	35.5	54	73424022
50-64	29.6	47	57439783
65-74	23.2	39	23832010
75+	19.5	36	14616177

Table 1: Demographic data used from AAA American Driving Survey. Blue-colored boxes are significantly lower than the overall estimate at the 95% confidence interval. Red-colored boxes are significantly higher than the overall estimate.

We combined the age ranges in blue that showed significantly lower overall estimates to be the category of low daily distance per person. The age ranges in white that were not significantly different from the overall estimate were combined to be the category of medium daily distance per person, and the age ranges in red that were significantly higher than the overall estimate were combined to be the category of high daily distance per person. These same criteria were used to separate the survey data into categories of low, medium, and high daily driving time.

The fraction of the population that is in each of the categories for miles and time is then found by dividing the total numbers of US drivers in the respectively colored age ranges where blue is low, white is medium and red is high by the total US driving population. Fractions for miles and for time both sum to 1. Categories are shown in Table 2.

Categories	Considerations	
	Distance in miles (*A*)	Time in min (*B*)
Low	0.2190	0.2190
Medium	0.4378	0.4378
High	0.3430	0.3430

Table 2: The fraction of the population listed for each consideration, for each category.

It should be noted that every age range that was significant lower in average daily distance per person was also significantly lower in average daily driving time per person, and the same holds for each age range when the data's significance is categorized into low, medium and high. This pattern supports the idea that the time spent driving depends on the distance driven, each day. The drivers have destinations with set driving distances; the physical distance does not vary, while the time taken is the varying component. Thus, the pattern present in the data for daily driving time and daily distance per person is indicative of this logical real-life occurrence.

Driving Consideration Model

Iteration 1

In the first iteration of this driving consideration model, we isolated the two factors considered and calculated each of the nine categories' percentages (P) by multiplying the fractions of the constituent mile and time categories. The low, medium, and high categories will be referred to by subscripts of 1, 2, and 3, respectively, for both the mile and time considerations. In other words, we multiplied the fraction of total US drivers in a category of daily distance (A) by the fraction of total US drivers in a category of daily driving time (B). Dividing this newly obtained fraction by the total ($\sum_{i=1}^{3}\sum_{j=1}^{3} P_{i,j}$) and multiplying by 100% results in the percentage of total US drivers with *A* daily distance and *B* daily driving time. Note that $\sum_{i=1}^{3} A_i = 1$ and $\sum_{i=1}^{3} B_j = 1$ (see Methods)

$$\sum_{i=1}^{3} A_i * \sum_{j=1}^{3} B_j = (A_1 + A_2 + A_3)(B_1 + B_2 + B_3) = \sum_{i=1}^{3}\sum_{j=1}^{3} P_{i,j} = 1$$

$$P_{i,j} = A_i * B_j * 100\% \text{ for } 1 \leq i,j \leq 3$$

Iteration 1 result percentages of US drivers for each of the nine categories are shown in Table 3:

		Daily Driving Time (min)		
		Low	Medium	High
Daily Distance (miles)	Low	4.79%	9.59%	7.51%
	Medium	9.59%	19.17%	15.01%
	High	7.51%	15.01%	11.76%

Table 3: Results of first iteration of the driving consideration model.

The driving consideration model returned values for the percentage of drivers within each category of low, medium, and high daily mileage and daily driving time. Instances where the constituent percentages for daily distance and daily driving time were the same and multiplied to be the same percentage result in the percentages that have the same value (such as the characterizations low daily distance, medium daily driving time and medium daily distance, low daily driving time, which both have percentages of 9.59%). Most US drivers had a medium daily distance and a medium daily driving time, though many had a high daily distance and medium daily driving time or a medium daily distance and high daily driving time. Interestingly, the smallest percentage of drivers had a low daily distance and low daily driving time.

The percentages for each of the nine categories from Iteration 1 assume that in Table 2, the *A* daily distance percentages can be distributed evenly among the daily driving time percentages and vice versa with the *B* daily driving time percentages. However, this is not always true. For example, 22% of drivers travel a low number of miles every day, we cannot then conclude that 22% of the drivers with low daily driving times must also drive a low number of miles, 22% of the drivers with medium times must drive a low number of miles, and 22% of the drivers with high times must drive a low number of miles. Rather, it makes sense that drivers who drive a low number of miles thus spend less time driving, and so the percentages must be weighted differently.

<u>Iteration 2</u>

For the second iteration of this driving consideration model, we included more information about the relationship between daily distance and daily driving time by considering how time is dependent on distance (miles). To do this, we first demonstrated that the two quantities were linearly correlated with randomly distributed residuals.

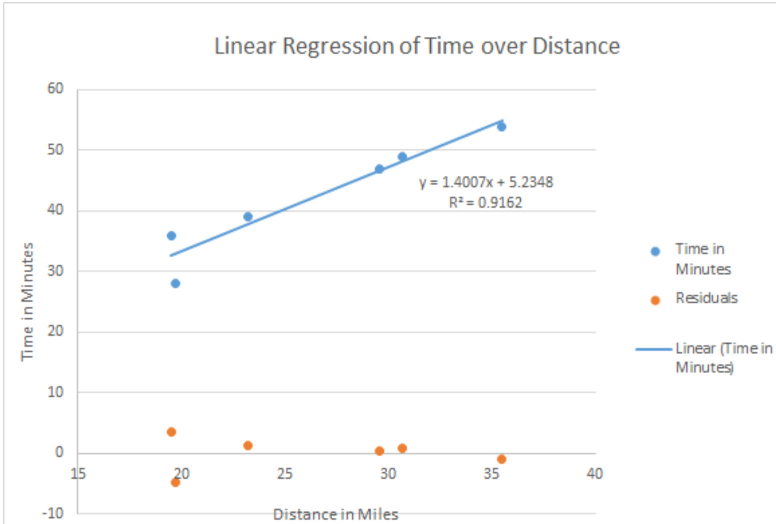

Figure 1: Linear regression of time over distance with residuals.

We added this relationship by using the same mile fractions from Table 2 and dividing the time fractions by a "pseudo-speed" (*W*) for the low, medium, and high categories.

Using Table 1, the "pseudo-speed" was found by dividing the weighted average of miles by the weighted average of time for a specific category (low, medium and high). The weight depends on the number of people in each age range. For example for the low category, the weighted mile average would be calculated by taking the average miles/person in the 16-19, 65-74, and 75+ age ranges multiplied by the corresponding number of people in that specific category and divided by the total number of people in the low mile category. The values from this calculation are shown in Table 4.

Categories	Considerations		"Pseudo-speed" (mi/min)
	Weighted Miles Average	Weighted Time Average (min)	
Low (k=1)	21.42	36.09	0.5935
Medium (k=2)	30.03	47.77	0.6286
High (k=3)	35.5	54	0.6574

Table 4: The "pseudo-speed" weighting factors for time.

Using these calculated values for the "pseudo-speed," we found new weighted values for the daily driving time ($B_{weighted}$) based on the corresponding fractional values for the daily distance (A):

$$B_{weighted,k} = \frac{\frac{A_k}{W_k}}{\sum_{k=1}^{3} \frac{A_k}{W_k}} \text{ for } 1 \leq k \leq 3$$

Categories	Considerations	
	Miles	Time (min)
Low	0.2190	0.2324
Medium	0.4378	0.4389
High	0.3430	0.3287

Table 5: Second Iteration. The fraction of the population listed for each consideration.

The same multiplication method was used as the first iteration ($Pi{,}j = A_i * B_{weighted,\,j} * 100\,\%$ for $1 \leq i, j \leq 3$) to find the percent of the population in each of the nine categories.

The second iteration results in these percentages of US drivers for each of the nine categories:

		Daily Driving Time (min)		
		Low	Medium	High
Daily Distance (miles)	Low	5.09%	9.61%	7.20%
	Medium	10.18%	19.21%	14.39%
	High	7.97%	15.05%	11.27%

Table 6: Results of the second iteration of the driving consideration model.

<u>Additional Conceptual Analysis for Driving Consideration Model</u>
Ideally, we would have daily distance and daily driving time data for each US driver and use these to categorize the drivers into the nine categories. However, with the time limitation, this data collection could not be collected and does not already exist in a readily available form. Because we were limited by the scope of our data, we decided to complete our numerical modeling with the method described in Iteration 2 above.

Given access to additional data containing more individualized data on each driver, we could use normal distribution graphs (bell curves) to calculate the boundaries between low, medium, and high categories for both the daily distance and daily driving time.

Using the collected data, the normality of the distances in miles and the time in minutes can be confirmed. Using these normal distributions, the values at the boundaries for the categories (low, medium, and high) can we be calculated.

It is known that any linear combination of independent random variables that are normally distributed will also be normally distributed. By demonstrating the normality of the overall daily miles and times, we can work backward to create three normal distributions that are linearly combined to create the original overall distribution.

Let us focus on the overall time normal distribution with a known mean (μ_0) and standard deviation (σ_0) (seen in Figure 2 (blue)). We would make the three constituent normal curves (α, β, γ) (seen in Figure 2 (red)) be centered around the means, μ_1, μ_2, μ_3, respectively, where $\mu_1 < \mu_2 < \mu_3$ with the union of $\alpha, \beta,$ and γbeing an area of 1 ($\alpha \cup \beta \cup \gamma = 1$). Let the corresponding standard deviations be $\sigma_1, \sigma_2, \sigma_3$. Therefore, the overall normal distribution can be described by

$$\mu_0 = a\,\mu_1 + b\,\mu_2 + c\,\mu_3 \text{ for some constants a, b and c}$$

$$\sigma_0 = \sqrt{a^2\sigma_1{}^2 + b^2\sigma_2{}^2 + c^2\sigma_3{}^2}$$

The boundaries between the low and medium categories and between the medium and high categories can be found by calculating the value at the intersection of the $\alpha$and βnormal distributions (x), the value at the intersection of the $\beta$and γnormal distributions (y), and the value at the intersection of the α and γnormal distributions (z).

Using normal cumulative distribution function, an equation for the union of α, β, and $\gamma$is

$$normalcdf(-1*10^{99}, x, \mu_1,\sigma_1) + normalcdf(x, y, \mu_2,\sigma_2) \\ + normalcdf(y,1*10^{99}, \mu_3,\sigma_3) = 1$$

$$normalcdf(-1*10^{99},x,\mu_2,\sigma_2) + normalcdf(x,1*10^{99}, \mu_1,\sigma_1) \\ + normalcdf(-1*10^{99},y, \mu_3,\sigma_3) + normalcdf(y,1*10^{99}, \mu_2,\sigma_2) \\ - normalcdf(-1*10^{99},z, \mu_3,\sigma_3) - normalcdf(z,1*10^{99}, \mu_1,\sigma_1) \ = 2$$

Using our limited data, we could approximate $\mu_0, \mu_1, \mu_2, \mu_3$, and $\sigma_0$with the values in Table 7. With more data, $\sigma_1, \sigma_2, \sigma_3$can be found, and $a, b, c, x, y,$ and $z$can be found.

	Miles	Time
Overall Average	30.0183	47.3502
Overall Standard Deviation	5.2306	6.8536
Low Average	21.41700618	36.08665714
Medium Average	30.02588795	47.77434173
High Average	35.5	54

Table 7: Approximations that could be used in this conceptual model.

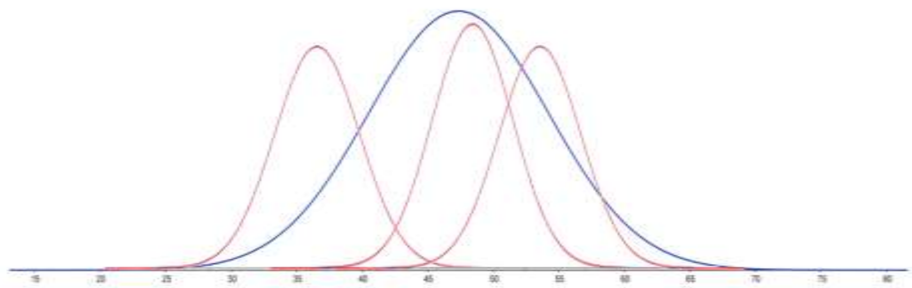

Figure 2: Representation of the overall time distribution with the three constituent normal parts.

Finally by determining the boundaries for each the mile and time between the low and medium categories and between the medium and high categories, we can find a more accurate measure of how much of the high category of time ($j = 3$) is related to the low category of miles ($i = 1$) and every other combination:

$$P_{1,3} = normalcdf(-1*10^{99}, x, \mu_{mile,1}, \sigma_{mile,1}) * normalcdf(-1*10^{99}, x, \mu_{time,3}, \sigma_{time,3}) * 100\%$$

2. Zippity Do or Don’t?

Our second task was to develop a model to study different car-sharing plans in different cities. We first analyzed four different car-sharing business plans [3]:

- **Round-trip car-sharing:** Customers rent vehicles, which they pick up and return to the same point.
- **One-way car-sharing floating model:** Customers rent cars on demand and return them to a defined area near their destination. Usually a “jockey” moves the vehicle to another area.
- **One-way car-sharing station model:** Customers pick up cars at an initial stations and drop them off elsewhere.
- **Fractional ownership:** Multiple owners jointly purchase a car.

The next part of this task was to determine which of the car-sharing plans considered would garner the most participation in each of the given cities: Poughkeepsie, NY; Richmond, VA; Riverside, CA; and Knoxville, TN. Finally, we used our model and estimated the car-sharing company’s profit to determine which city would be the best target for expansion of such a company.

Model Design

We isolated round-trip car-sharing as a different service, satisfying a different customer goal, and analyzed it qualitatively.

Then, we used an agent-based model to compare one-way car-sharing with floating and station plans, as well as fractional ownership, by simulating a city's driving population's behavior through agent-based modeling. Our model essentially simulates the daily driving conditions in a city at any given time. From this, we were able to determine the amount of people that would participate in a specific car-sharing business plan in a specific city. We used participation to determine the most popular of the three business plans for each city. Then we determined which city a car-sharing company should further investigate expanding its service to. By determining the best car-sharing plan and ranking the cities as prospective locations, our model can be used by a company to compare future car-sharing business sites.

Assumptions, Decisions, and Justifications

- **The decisions and conditions of drivers, local infrastructure, worth of dollars, and gas are not affected by sudden events.** These are all conditions that might modify how drivers behave and the system of car-sharing within a given city.We decided that our agent-based model would not include events such as natural disasters or industry booms.
- **The price of gas is simplified to the average cost of gas across the nation.** Gas prices change significantly over time so we decided to use the most recent data available for the United States [5].
- **If a person is considering car-sharing, they do not own their own car.** If a person already owns a car of their own (being the sole owner), it is least expensive and fastest for them to drive directly from point A to point B, as they only pay for the price of gas and not rental fees.
- **Round-trip car-sharing satisfies a different goal.** We decided that round-trip car-sharing is only feasible for people that leave or enter a city from a transport center, like an airport. Thus, round-trip car-sharing is not presented as an option in our agent-based model and is instead studied through a separate analysis using qualitative judgment.
- **Companies interested in the future of car-sharing in cities are interested only in services where cars are lent, not sold.** As these cities investigated are established centers of commerce and development, there are already dealerships and other sources where people may buy cars (both for sole ownership and fractional ownership)
- **Fractional ownership of a car is split between two people**. As there is no reliable data on average private fractional ownership, we decided that a person that participates in fractional ownership would pay half the cost of sole ownership.

Separate Analysis of Round-Trip Car-Sharing in Cities

Through research and comparison of user testimonies for round-trip car-sharing we decided that the nature of round-trip car-sharing makes it a different service, compared to the other three options studied. In any city, people with relatively constant daily driving plans would either want to get from point A to point B, or leave point A, drive around, and return to point A. The first group would more likely rent or pay for floating or station-based one-way car-sharing or fractional ownership because it would be unfeasible to rent a car, drive from the pickup station to work, school, or errands, and drive back to that station.

From research and personal firsthand experiences, we determined that the major customer base of round-trip car-sharing businesses are travelers from outside a city. This group of people is most likely to enter a city through an airport, train station, or other long-distance transport center and rent a car at a business location, use it to drive around the city, return the car to the same business location (completing the round-trip), and then leave the city through the transport center. Table 8 shows the data used [6]:

City	Airport Name	Annual passengers	Runways (length in ft)
Poughkeepsie, NY	None	--	--
Richmond, VA	Richmond International Airport	3,500,000	9,003; 6,607; 5,326
Riverside, CA	Riverside Municipal Airport	not available	9,005; 9,000
Knoxville, TN	McGhee Tyson Airport	1,700,000	5,401; 2850

Table 8: Data used in qualitative ranking of four cities for round-trip car-sharing.

Using the data observed, we qualitatively ranked Richmond, VA, as the city that would have the most participation in round-trip car-sharing of the four cities, as it has the most reported annual passengers and three runways, theoretically allowing for the largest group of people traveling from outside the city. We determined the city with the second-most participation to be Riverside, CA, based on the runway size—the larger the runway, the larger the planes that can land are, and the more people arrive at the airport. By this qualitative analysis, the city with the third-most participation would be Knoxville, TN, as there is an airport but its runways are both smaller than those of any other city's airport considered in this ranking. Finally, Poughkeepsie, NY, is left to be the least round-trip car-sharing participation of the four cities, as there is no airport as a large source of outside travelers. In this qualitative analysis, the cities that would garner the most participation in round-trip car-sharing were ranked based on available data for city airports. A company would want to implement round-trip car-sharing facilities according to the following priority: Richmond, VA; Riverside, CA; Knoxville, TN; Poughkeepsie, NY.

Agent-Based Model

To account for variability in cities, data were taken from the Texas A&M Transportation Institute 2015 Urban Scorecard [8] and data from the US Census Bureau [13] to find the total number of commuters, average commute to work, and minutes of delay due to traffic in the four cities, shown in Table 9.

City	Commuters	Avg time to work (min)	Minutes of delay/day
Poughkeepsie, NY	20,500	31.2	6
Richmond, VA	50,300	28.8	6
Riverside, CA	108,100	32.0	10
Knoxville, TN	31,000	21.4	6

Table 9: Data used for agent-based model from 2015 Urban Scorecard and US Census Bureau.

We then simulated the movement of people through cities based on the data we collected by using NetLogo, an open-source agent-based modeling software. This model simulates a single day for the commuters between their place of residence and their commercial workplace. The figure below shows the interface:

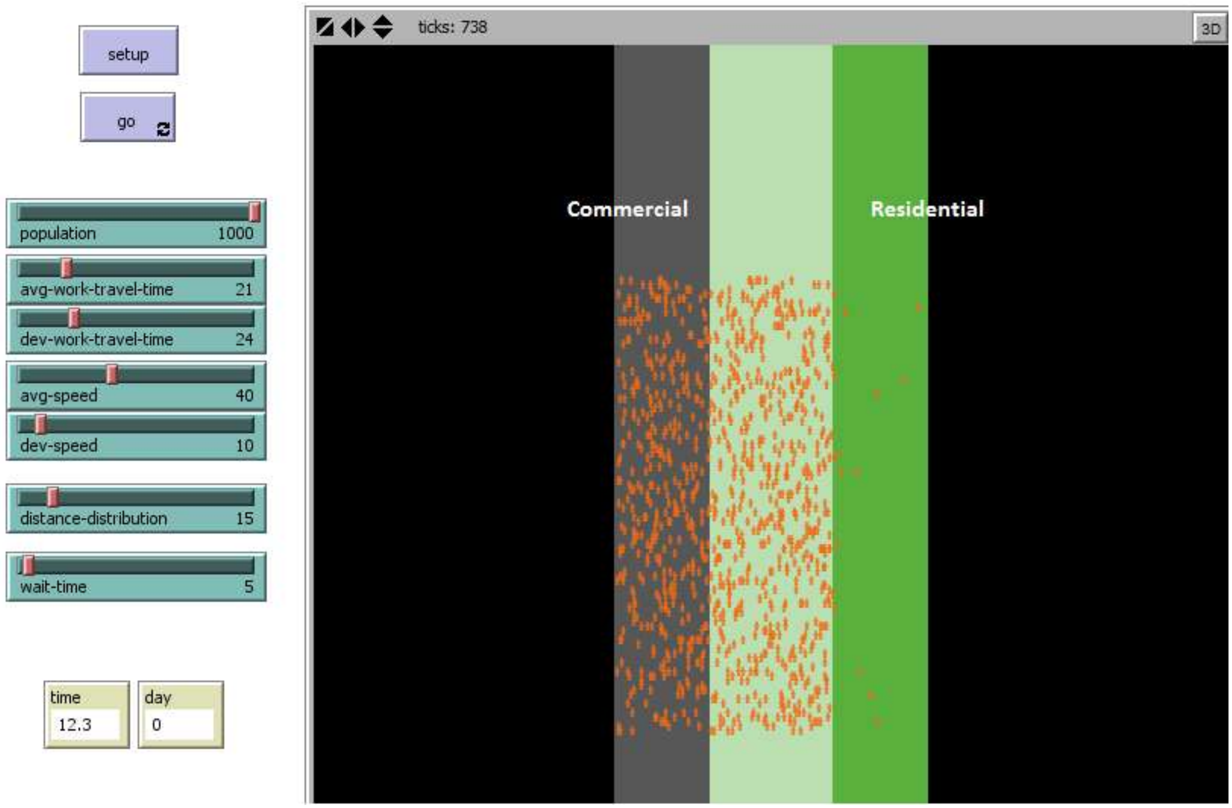

Figure 3: Screenshot of agent-based Netlogo model. The sliders on the left allow for value changes. The window on the right shows people leaving their homes in the residential areas (green) for destinations in the commercial area (grey), with the lighter green being an area of overlap with both residents and commercial development. The physical space between the green and grey sides is representative of time, not distance, and is set by the slider avg-work-travel-time.

Methods

As shown in Table 8, the model accounted for variability in the cities by individually considering the average commute times to a certain destination.

Since the goal was to find the participation within each of the cities for each of the other three car-sharing plans, we first implemented the results from the driving consideration model into the agent-based model to serve as a basis of comparison. Since the values are averages, we first randomly generated 1,000 individuals and distributed around these average values. Each agent was also given a randomly generated scalar value (S) which was closer to 0 if they preferred to save more time or closer to 1 if they preferred to save more money. This accounts for personal variation in each individual's needs.

To determine the effectiveness of the methods, within each method, the agents calculated the cost (*C*) and time (*T*) taken for their individual commutes. The cost equations are as shown.

For the one-way floating plan, the cost (C_f) is given by the following, where *d* is the distance traveled in miles:

$$C_f = max\{11, 2(1.5 + .16T + .8d)\}$$

This is based on Uber's model of incorporating distance, as the two are closely related. There is a $5.50 minimum fee, a cost of $0.16 per minute, a $1.50 service cost, and a cost of $0.80 per mile [10]. The multiplier of two is incorporated again for the same reasons as the floating station model.

For the one-way station plan, the cost (C_s) is given by the following equation, where *T* is the time in minutes of the commute. The *max* function will take the higher of the two values.

$$C_s = max\left\{18, 2(9*\frac{T}{60})+\frac{7}{30}\right\}$$

This is based on Zipcar's model of charging $8-10 hourly per trip as well as a $7 monthly membership fee (it is divided by 30 as these are daily costs) [9]. Multiplying both by 2 accounts for the trip to the commercial area and back.

For fractional ownership, we found the daily cost to own a car to be $20 per day and $0.11 per mile driven [5]. So the cost (C_p) is given by the following, where

$$C_p = 20 + .11d$$

The time was found from the NetLogo model itself, and the following equation was used:

$$R = S*C+(1-S)*T$$

In this case, *R* becomes the ranking variable that is used as a metric to choose which option. In the model, each individual agent calculated *R* values for the three transportation options. The option that provided a minimum for *R* was chosen by that agent.

Results

First, we ran the model with the three car-sharing strategies implemented. We used multiple trials for the various cities and took averages to find which strategy would be preferred if all options were available. This provided the customer perspective on the situation. From the 1000 agents, we found the number that preferred a certain strategy. We used a 95% confidence interval around the mean to show statistical significance.

Method	Average # of Times Chosen	StDev	CI Low	CI High
Fractional	847	26	839	855
One-way Floating	42	31	32	51
One-way Station	111	14	107	116

Table 10: Agent choice statistical data per 1000 agents for Poughkeepsie, NY.

Method	Average # of Times Chosen	StDev	CI Low	CI High
Fractional	828	21	822	834
One-way Floating	50	34	40	60
One-way Station	122	18	116	127

Table 11: Agent choice statistical data per 1000 agents for Richmond, VA.

Method	Average # of Times Chosen	StDev	CI Low	CI High
Fractional	855	21	849	861
One-way Floating	38	31	28	47
One-way Station	107	14	103	112

Table 12: Agent choice statistical data per 1000 agents for Riverside, CA.

Method	Average # of Times Chosen	StDev	CI Low	CI High
Fractional	761	22	755	767
One-way Floating	90	32	81	100
One-way Station	149	31	140	158

Table 13: Agent choice statistical data per 1000 agents for Knoxville, TN.

The 95% confidence intervals show that, for each city, the agents each find that sharing a vehicle with another agent would be preferred. From the the perspective of a rental company, there is little or no profit resulting from fractional car ownership strategy. This left us with the one-way floating and station plans. For each of the cities, the 95% confidence interval was higher for the one-way station strategy than the one-way floating strategy. As it is most profitable, we decided that a company would proceed by this strategy only.

After finding this, we ran the simulation again to find the profit for a company per 1000 agents assuming the only strategy available to them was the one-way floating car strategy. This formed our metric by which we could rank the cities for the car-sharing company to move forward. Note that our metric used profit per 1000 agents instead of complete profit because having more people would require more cars and rental facilities to be available through this strategy. So, this would increase the company's costs and decrease profit. Keeping the metric as profit per 1000 agents simplifies this issue. The results of the simulation are as shown:

City	Avg profit/1000 ($)	StDev	CI Low	CI High	Rank
Poughkeepsie, NY	1586.60	73.96	1521.77	1651.43	2
Richmond, VA	1469.00	126.96	1357.71	1580.29	3
Riverside, CA	1428.20	124.55	1319.03	1537.37	4
Knoxville, TN	1915.60	99.05	1828.77	2002.42	1

Table 14: The results of the simulation, showing the profitability per 1000 people in each of our test cities. Low and high limits for 95% confidence intervals are shown.

Note that Knoxville has a statistically significant advantage over the other three cities in terms of profit per 1000, so it clearly ranks as the best city. The other three cities have been ranked based on their means, but the overlap in the 95% confidence intervals shows that there is not a statistically significant difference in their means.

3. Road Map to the Future

The modification of the model simply required a modification of the profit and cost to the consumer. We were interested in the available profit per 1000 agents for the company, so we chose not to consider the fractional ownership strategy here for the same reasons as in the previous sections, as it does not contribute to a company's profit. Due to the availability of self-driving cars as a new technology, the two remaining strategies, the one-way floating and one-way station strategies, essentially combine into one strategy. Since the human aspect of the car is removed, there is no more need to leave the car at a station. Thus, the remaining plan is the one-way floating car strategy. This means that the program only needed to run for this one strategy to find the city with the highest profit.

The cost was modified as such: while the minimum cost was kept the same, the \$1.50 service cost was removed. Thus, the new equation for cost (C_s^*) becomes

$$C_s^* = max\{11, 2(0.16T + 0.8d)\}$$

So the simulations were run to find the profit per 1000 agents. The results are as shown:

City	Avg profit/1000 ($)	StDev	CI Low	CI High	Rank
Poughkeepsie, NY	304.80	57.71	254.21	355.39	3
Richmond, VA	309.20	68.19	249.43	368.97	2
Riverside, CA	236.80	60.71	183.59	290.01	4
Knoxville, TN	431.60	46.12	391.17	472.03	1

Table 15: The results of the simulation, showing the profitability per 1000 people in each of our test cities for the future scenario. Low and high limits for 95% confidence intervals are shown.

The results were similar to those in the previous section. Knoxville was still the best city for the company in terms of profit. The relative ranks shifted slightly, with Richmond being pushed to second place; however, the means of these two profit values are quite close, and the confidence intervals overlap almost completely. The confidence interval from Riverside also overlaps these two, so there is no statistical significance between these three means.

Model Evaluation

Sensitivity Analysis

It is important that our agent-based model shows tolerance to a small variance in its inputs. If our model is robust, the outputs will not vary dramatically given small differences in inputs. To test our model's sensitivity, we thus intentionally introduce small amounts of error and compare the resulting output to our original solution.

First, we varied the time the drivers spent going to work from 20 to 30 minutes. As expected, we found that increasing the average driving time decreased the popularity and profit earned from the options, but the relative rankings of the option and city popularity remained the same.

We also tried iterating the time spent waiting for a vehicle from 10 to 30 minutes. Interestingly, we found that the number of users using the car-sharing options or the profit earned did not change significantly. This makes sense because at a certain threshold of wait time, users would simply stop waiting for a vehicle and pursue other options such as fractional ownership.

Strengths

- The driving consideration model is based on real-world data gleaned from current sources. Our model thus accurately represents the current US driver situation in the four cities and can be used to predict actual changes in behavior due to the car-sharing implementations.
- The qualitative analysis of round-trip car-sharing provides insight for the viability of round-trip car-sharing by considering realistic and logical information available for the cities' airports, a main source of outside travelers that utilize round-trip rental services.
- The use of agent-based modeling allows for individual decisions to contribute to the overall outcome of the model. We chose agent-based modeling as it is an excellent tool to simulate real-life complex systems. Rather than a deterministic model, our model allows for variations in behavior to create different outcomes which can be analyzed over many iterations.
- The agent based model contains stochastic elements, simulating the random nature of the car drivers and the owners. Factors such as travel distance, wait time for rentals, and an individual's preference for saving time over money are all inherently random, which is accounted for in the NetLogo code for our agent-based model.
- Using time instead of distance within the agent-based model increasing the flexibility of the model as it does not account for the specific layout of the city but rather the time people take when driving within a city. This makes the model more flexible and easily applicable to other cities with more data available.
- Our agent-based model simulates the behavior of cities' driving population in a scenario where a company has introduced car-sharing, the city has "recovered" from any temporary spikes in behavior, and the entire system has returned to "normal" behavior. In other words, our model shows a simulation of future daily preference and profit for the people in a city and a car-sharing company, respectively.

Weaknesses

- Our driving consideration model determined the low, medium, and high categories based on the 95% confidence intervals as given by our data. However, we could not access all the data the source used, and so we do not know the exact boundary lines of these three categories for both the daily distance and daily driving time.
- The driving consideration model is based on average data, which is skewed right, and not raw data from transportation sources, which is not available. The AAA reports that the average driving time is 46 minutes, while the median is 22 minutes. The daily driving miles are also similarly right skewed [4].
- Specific city layout and traffic-related factors that vary by location or season are not accounted for, as our agent-based model “spaces” agents using time instead of distance. For example, Riverside, CA, is considered part of the greater Los Angeles area [7]. These conditions as well as the preexistence and physical location of car-sharing facilities and parking availability may cause differences in people's decisions regarding car-sharing that are not accounted for in our model: Riverside, CA, may be much more populated during certain tourist or business conditions or centers of commerce, or Poughkeepsie, NY, may be snowed in.
- Our agent-based model was based on the typical workday, where the majority of drivers work 9am-5pm jobs. This means that the vast majority of commutes happen in the hours before 9am and after 5pm. It’s possible that the residents of a city would drive during hours between 9am and 5pm or at night, but the model does not take this into account. In addition, driving patterns may completely change on weekends as people go to different locations.
- Our agent-based model does not account for cities’ immediate response to the introduction of car-sharing businesses. We created a simulation of the future conditions and did not model a gradual shift or change in people’s choice of methods over time. Companies would not use our model to predict immediate profits but rather long-term gain from predicted future uses, given a city and car-sharing plan.

Conclusion and Future Work

Our solution models the behavior of US drivers and simulates different car-sharing plans in the context of different cities being considered by car-sharing companies.

First, we characterized the driving population in the US by creating a model that determines the percentage of current US drivers that have low, medium, and high driving time per day and driving miles per day. These are two of the main factors that people consider when making decisions regarding car-sharing. Our mathematical model uses data reported from nationwide surveys and returned percentages for each of nine combined categories: low, medium, and high for both daily driving time and daily driving miles. In the future, we could collect data about these two considerations and find more accurate numbers. With a more accurate characterization of the percentage of the population into these nine categories, these percentages could be used in conjunction with the data from the second agent-based model to find the tendencies of people in each of the nine categories.

Then, we created an agent-based model to simulate the daily behavior of a city with different car-sharing plans available. We considered the four given cities: Poughkeepsie, NY; Richmond, VA; Riverside, CA; and Knoxville, TN. For each of the four, people most preferred using one-way car-sharing with stations where they picked up and dropped off cars at established stations. For the company, Knoxville would be the most profitable (per 1000 people) to introduce this. The flexibility of our agent-based model also allows it to be applied to other cities.

Finally, we considered a scenario where a car-sharing company rents out renewable energy cars with emerging technology for self-driving cars. The renewable energy part essentially removes the cost of gas and replaces it with electricity cost [5]. This drastically decreases the cost for the companies and potentially increases profit margins. The self-driving cars also will remove many service fees within costs, so consumers can pay even less than what they are paying now. The cost for the companies will also decrease as a result. However, our model showed that, even with these changes, Knoxville would be the most profitable (per 1000 people).

Our agent-based model indicated that implementing one-way car-sharing with stations is preferable for a company using gas-powered cars, while a floating plan is preferable when using self-driving and renewable energy-powered cars. The one-way car-sharing station plan has been highly successful in real-world implementation in the case of Zipcar car-sharing [11]. This plan results in profits for car-sharing business and reduces the “last mile” issue. “Last mile” is a problem encountered in transportation, essentially indicating the trouble of transporting people from their destination or source location to the beginning or end of a public transport line [20]. For example, a subway may run through a city, but employees working in office buildings far away from a station would not be likely to make use of cheaper public transport. Instead, the stations used in this one-way car-sharing plan can continue to be implemented in cities as portions of a parking lot or street parking slots [12], allowing for flexibility of location and increasing driver participation in car-sharing programs. The model we developed can be used to estimate a city’s driving population’s behavior when given car-sharing options as well as assess different cities as prospective new locations or expansion sites for the future of car-sharing plans.

References

[1] Bergen, Mark, and Kara Swisher. "GM Invests $500 Million in Lyft and Strikes Strategic Autonomous Car Alliance." *Recode*. Vox Media, Inc., 04 Jan. 2016. Web. 28 Feb. 2016.

[2] Helper, Lauren. "Zipcar, Google and Why the Carsharing Wars Are Just Beginning." *GreenBiz*. GreenBiz Group, 14 July 2015. Web. 28 Feb. 2016.

[3] Reuters, Matthew DeBord and. "Ford Is Betting $4.5 Billion on Radically Transforming the Company and Its Mission." *Business Insider*. Business Insider, Inc, 11 Dec. 2015. Web. 28 Feb. 2016.

[4] Triplett, Tim, Robert Santos, and Sandra Rosenbloom. *American Driving Survey*. Rep. N.p.: AAA Foundation for Traffic Safety, 2015. Online.

[5] Stepp, Erin. "Annual Cost to Own and Operate a Vehicle Falls to $8,698, Finds AAA | AAA NewsRoom." *AAA NewsRoom*. AAA, 27 Apr. 2015. Web. 28 Feb. 2016.

[6] "AirportIQ 5010." *AirportIQ 5010*. Aeronautical Information Services, n.d. Web. 28 Feb. 2016.

[7a, b, c] "Richmond International Airport, Riverside Municipal Airport, McGhee Tyson Airport." *Wikipedia*. Wikimedia Foundation, n.d. Web. 28 Feb. 2016.

[8] Lomax, Tim. "Annual Urban Mobility Scorecard." *Urban Mobility Information*. N.p., 2015. Web. 28 Feb. 2016.

[9] "Winter, Meet Wanderlust." *Car-sharing: An Alternative to Car Rental with Zipcar*. Zipcar.com, n.d. Web. 28 Feb. 2016.

[10] "How Much Does Uber Cost? Uber Fare Estimator." *RideSharingDriver*. N.p., 02 Dec. 2014. Web. 28 Feb. 2016.

[11] Krantz, Matt. "Can Zipcar's Stock Success Continue? - USATODAY.com." *USATODAY.COM*. N.p., 6 May 2011. Web. 28 Feb. 2016.

[12] "Boston Metro." *Car-Sharing and Hourly Car Rental- Enterprise CarShare*, n.d. Web. 28 Feb. 2016.

[13] "United States - Average Commute Time by State." *United States - Average Commute Time by State*. U. S. Census Bureau, n.d. Web. 28 Feb. 2016.

[14] Jarosz, Beth, and Rachel T. Cortes. "In U.S., New Data Show Longer, More Sedentary Commutes." *In U.S., New Data Show Longer, More Sedentary Commutes*. Population Reference Bureau, n.d. Web. 28 Feb. 2016.

[15] "Fastfleet." *Fastfleet*. Zipcar, n.d. Web. 28 Feb. 2016.

[16] "Best car-sharing reviews 2016 - Reviews.com." *Reviews.com*. N.p., n.d. Web. 28 Feb. 2016.

[17] Loizos, Connie. "Can Fractional Car Ownership Work?" *TechCrunch*. Techcrunch, n.d. Web. 28 Feb. 2016.

[18] "Solar Installs in Maine, New Hampshire, Massachusetts." *ReVision Energy*. N.p., n.d. Web. 28 Feb. 2016.

[19] Fitzpatrick, Alex. "General Motors Is Launching a New Car-Sharing Service." *Time*. Time, n.d. Web. 28 Feb. 2016.

[20] "Last Mile Definition | Investopedia." *Investopedia*. 2016, Investopedia, LLC., 30 Aug. 2006. Web. 28 Feb. 2016.

Software used:

Netlogo 5.3.1

Geogebra

Microsoft Excel

Google Drive