

ORF 467 Final Project: West Region

Kevin Wu, Anna Erkalova, Mike Wagner, and Bulut Cakmak with assistance from files

produced by Kyle Marocchini

Professor Kornhauser

ORF 467 Fall 2017

Table of Contents

Executive Summary:3

Assessment of Population Served Based on analysis of NN Files:.....4

Assessment of Trips Originating in the Region based on NationWide Trip Data:.....8

Assessment of RideShare Potential:.....13

Assessment of Start-of-Day Demand and End-of-Day Supply:17

Assessment of Business Case for your Regional aTaxi Mobility-for-All System:19

Recommendations for Further Analysis:.....21

Appendix:23

 Alaska:23

 Arizona:26

 California:30

 Hawaii:.....33

 Idaho:37

 Nevada:40

 Oregon:44

 Utah:.....47

 Washington:51

Executive Summary:

The West Region includes Alaska, Arizona, California, Hawaii, Idaho, Nevada, Oregon, Utah, and Washington, and includes metropolitan areas such as Los Angeles, San Diego, San Jose, San Francisco, Seattle, and Phoenix. Of the nine states that were part of our analysis, the majority of trips originate from California. For the West, the busiest times of day are evening rush, evening, and morning peak, in that order, and the greatest AVO occurs during the morning peak and evening rush. Major metropolises tend to have higher AVO in their regions, but there are also a few scarcely scattered rural regions with very high AVOs as well.

The Start of Day Demand and End of Day Supply distributions for 1x1 pixels proved to be too computationally difficult to be graphed and visualized in an interactive plot by the computers available to the team. The Start of Day Demand and End of Day Supply distributions for both the 5x5 SuperPixels and 10x10 SuperDuperPixels show that SoD Demands tend to be toward the center of cities while EoD Supplies tend to be further from high population areas. The amount of active taxis during the day was greatest during the morning peak and evening rush.

For the Business Case, ride-share services are the only revenue source. Expenditures include vehicle acquisition and maintenance, energy acquisition and maintenance, and parking land rental.

All data analysis of the original trip files was done using programs in Java written by the team. All interactive plots were created using Bokeh and Python. Html files for each of the interactive plots as well as a supplementary spreadsheet containing all data output from the team's analysis are included with the submission of this project.

Assessment of Population Served Based on analysis of NN Files:

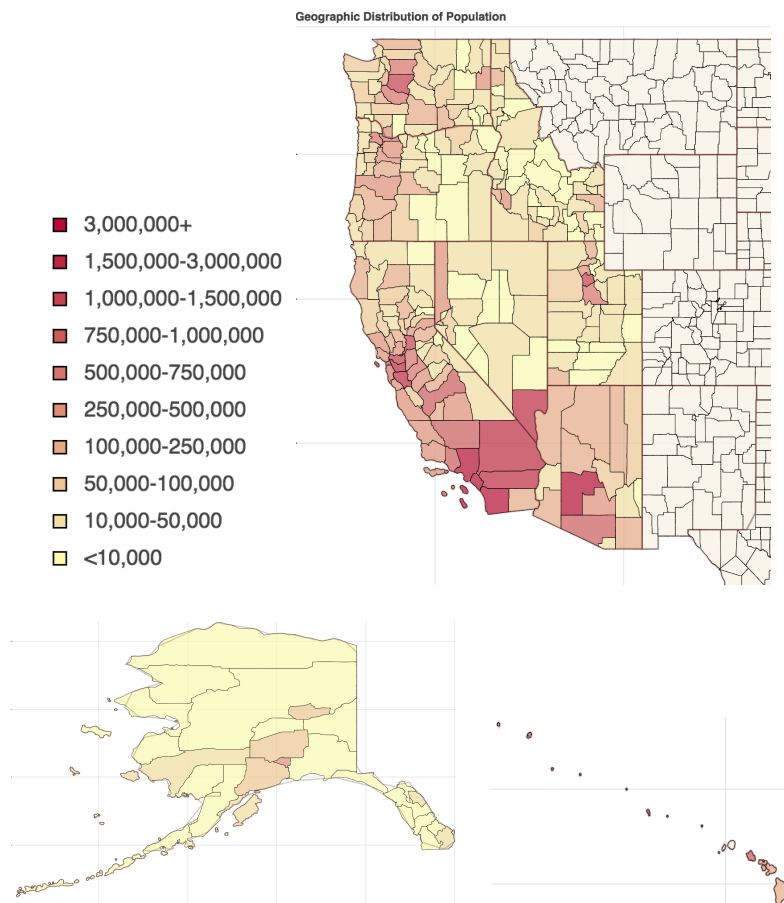
Methodology:

Each row in the NN Files represented a person who visited zero to seven nodes. The PersonTrips was determined by calculating $\max(0, \text{nodes} - 1)$ for each row. Population, PersonTrips, and trip distances were recorded on both a by county and entire region basis, and the frequency of trip lengths was also calculated. The total population is 61,546,382 people with 195,404,528 PersonTrips and an average distance length of 21.543.

Assessment:

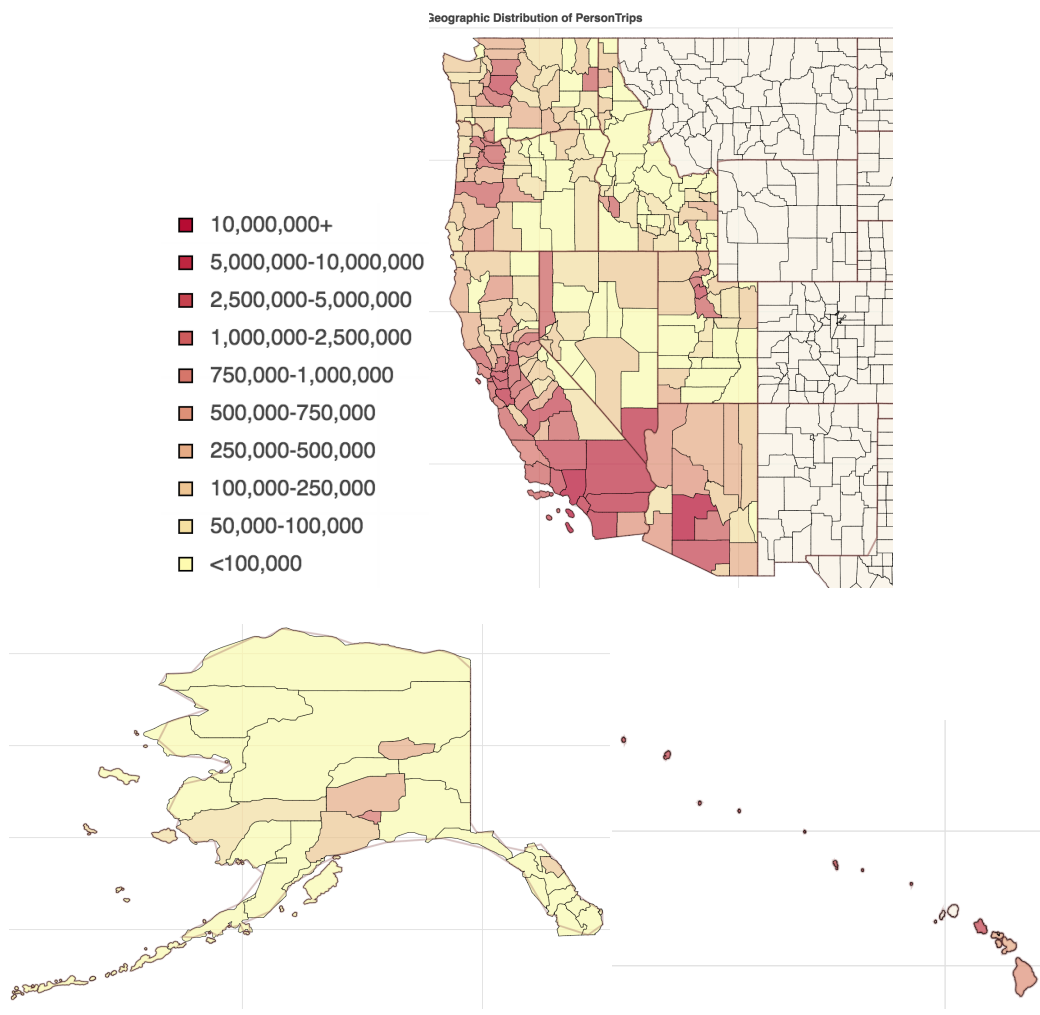
Areas with high populations include the Southern California, Las Vegas, and Phoenix regions.

Figure 1: Geographic distribution of population, according to Nationwide NN (Person) Files. Interactive graph (/ORF 467 Final Project Interactive Plots /Population.html)



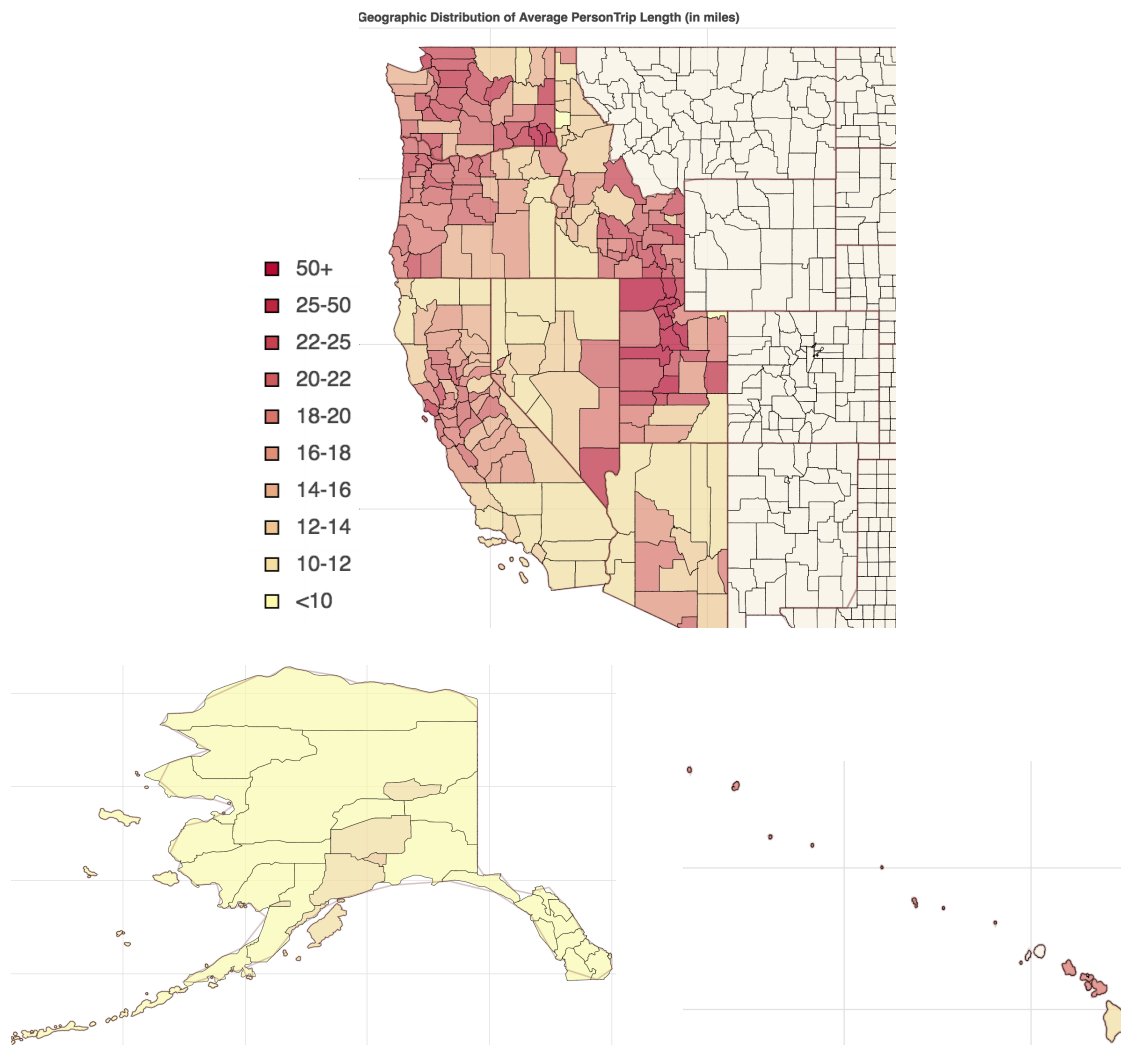
Similar to above, the Southern California, Las Vegas, and Phoenix areas again contain the largest amounts of PersonTrips.

Figure 2: Geographic distribution of PersonTrips, according to Nationwide NN (Person) Files. Interactive graph (/ORF 467 Final Project Interactive Plots /PersonTrips.html)



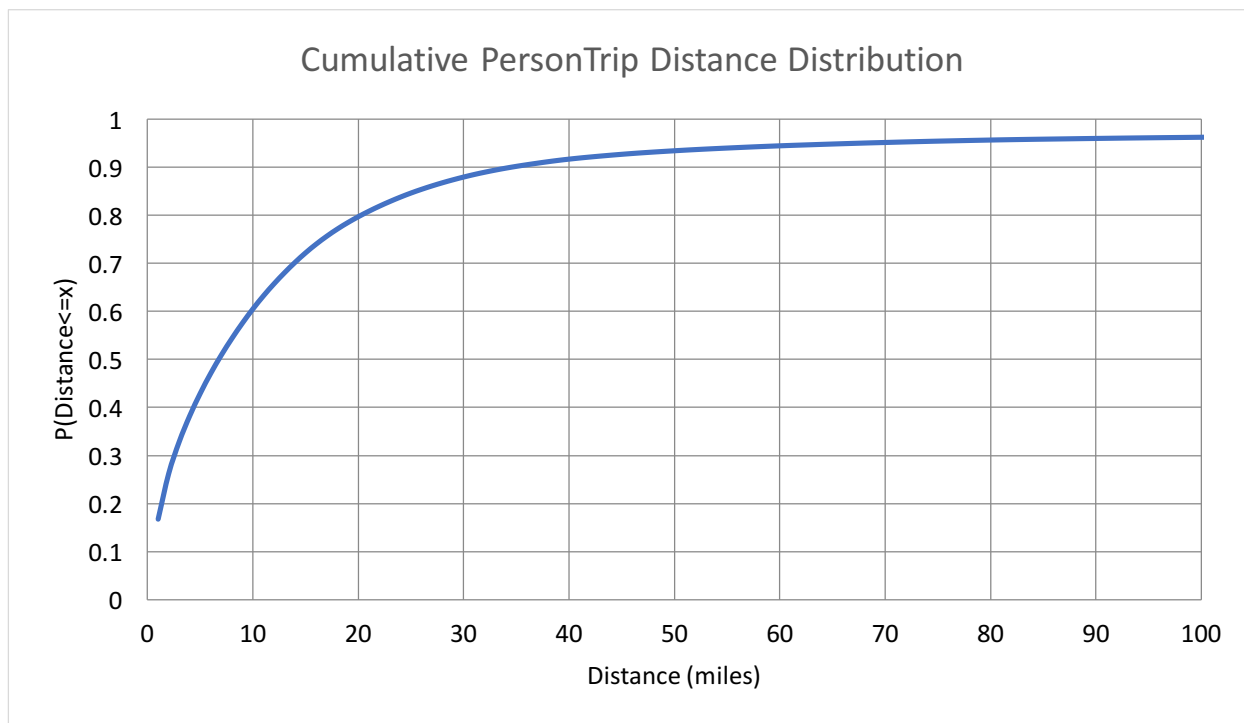
Regions with higher average trip distances include areas in Utah, Idaho, Washington, and Oregon. The overall state of Alaska features low average trip distances.

Figure 3: Geographic distribution of Average PersonTrip Length, according to Nationwide NN (Person) Files. Interactive graph (/ORF 467 Final Project Interactive Plots / Average_PersonTrip_Distance.html)



Over half of all PersonTrips (approximately 60%) have a gcd of 10 or less and approximately 80% of PersonTrips have a gcd of 20 or less.

Figure 4: Cumulative distribution of trip lengths for all trips made by West residents, according to Nationwide NN (Person) Files



Assessment of Trips Originating in the Region based on NationWide Trip Data:

Methodology:

Every trip was assigned a trip type based on the distance of the trip: type zero for trips less than 0.25 gcd, one for trips between 0.25 and 2 gcd, two for trips between 2 and 10 gcd, and three for trips greater than 10 gcd. Each trip was assigned a drop off zone and minimum wait time dependent on the trip type and destination location. Type zero trips were removed from further analysis as there is no opportunity for ride-share. Type one trips were assigned drop-off zones that are the same as the original pixel location of the destinations and minimum wait times of 300 seconds. Type two trips were assigned drop-off zones that are three-by-three-pixel regions of the original destinations and minimum wait times of 420 seconds. Type three trips were assigned drop-off zones that are five-by-five-pixel regions of the original destinations and minimum wait times of 600 seconds.

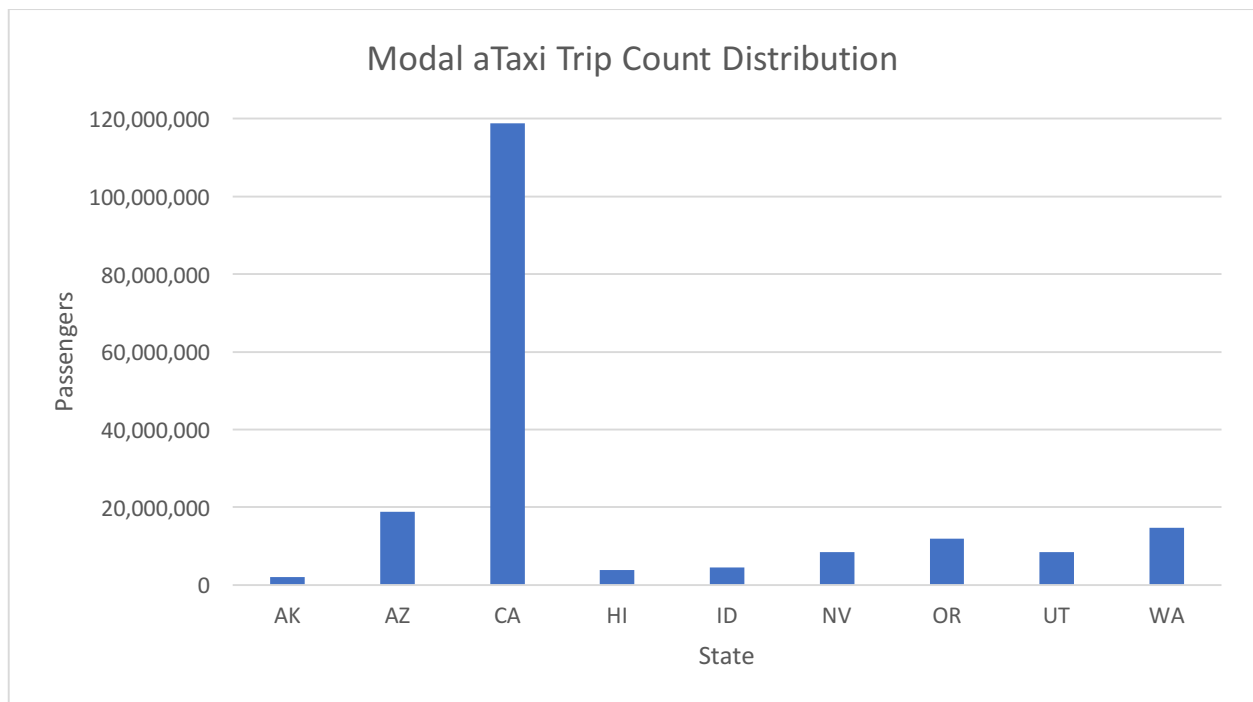
The trips were then sorted by starting pixel, drop-off zone pixel, and departure time. To determine ride-share opportunity, the rows in the data files were scanned to determine if each proceeding trip started from the same starting location, ended in the same drop-off zone, and started at a time less than the previous trip's departure time plus minimum wait time. If these conditions were met, then the proceeding trip could ride share with the previous trip and would leave at the same time as the previous trip; if not, the proceeding trip would begin as a new taxi with departure time equal to departure time plus minimum wait time. The number of passengers for every departed taxi was recorded as well as the sum of the trip distances of all the passengers. The vehicle distance traveled for every taxi was estimated to be the largest passenger trip distance. The average vehicle occupancy could then be calculated by dividing total passenger distance by total vehicle distance.

The arrival time for each departed taxi was estimated by assuming that each vehicle traveled on average thirty gcd per hour. A start-of-day demand and end-of-day supply distribution by pixel was then calculated by aggregating the net supply and demand for every pixel over the day with the assumption that each taxi could only carry a maximum of five passengers. The active fleet by minute distribution was calculated by recording the time duration during which each taxi was active to complete the trip and then summing the total amount of taxis active for each minute.

Population and Trip Distribution Assessment:

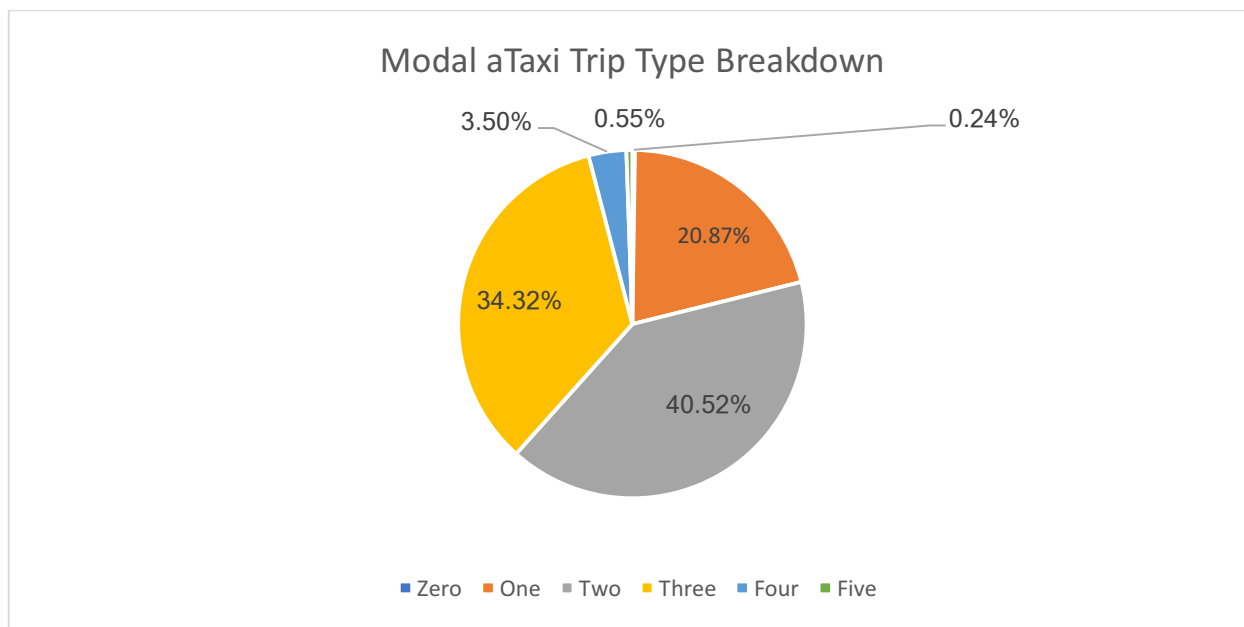
Of the 191,609,413 total modal aTaxi trips, the overwhelming majority are from California.

Figure 5: Count of aTaxi trips per state, according to Nationwide Modal aTaxi PersonTrips Files



Most trips are of type one, two, or three with only a very small percent of type zero, four, or five.

Figure 6: Percentages of aTaxi trip types, according to Nationwide Modal aTaxi PersonTrips Files



The following four charts suggest that the distribution of trips is bimodal, with a larger peak during evening rush hours and a smaller peak during morning rush hours.

Figure 7: Percentages of aTaxi trips that occur during each time-of-day, according to Nationwide Modal aTaxi PersonTrips Files

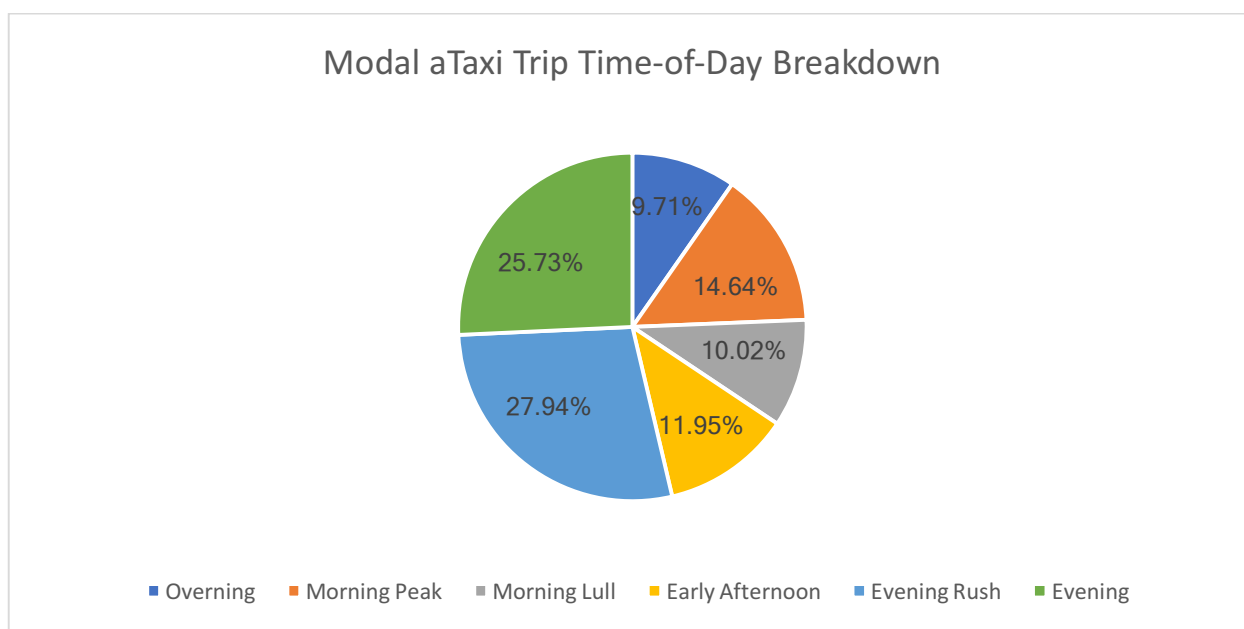


Figure 8: Percentage of aTaxi trips that occur during every hour of the day (where Hour 1 is equal to 12:00 am), according to Nationwide Modal aTaxi PersonTrips Files

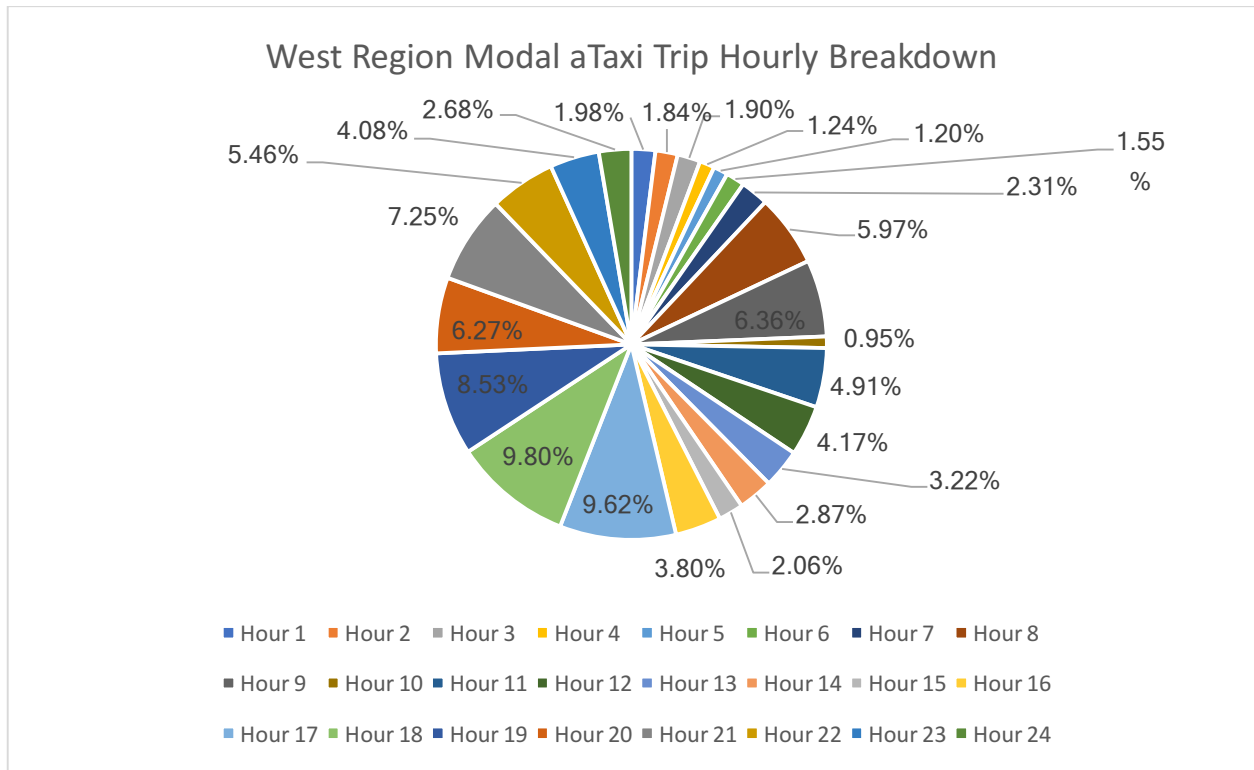


Figure 9: Number of aTaxi trips by time-of-day, broken down by state, according to Nationwide Modal aTaxi PersonTrips Files

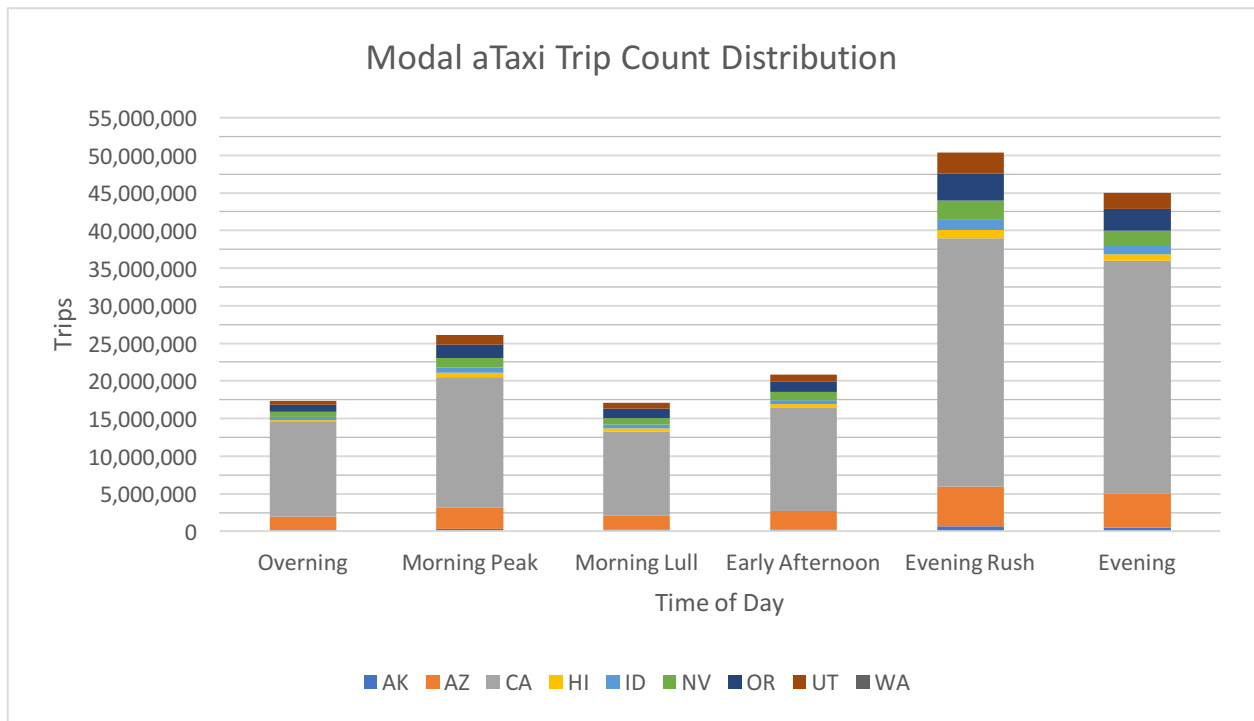
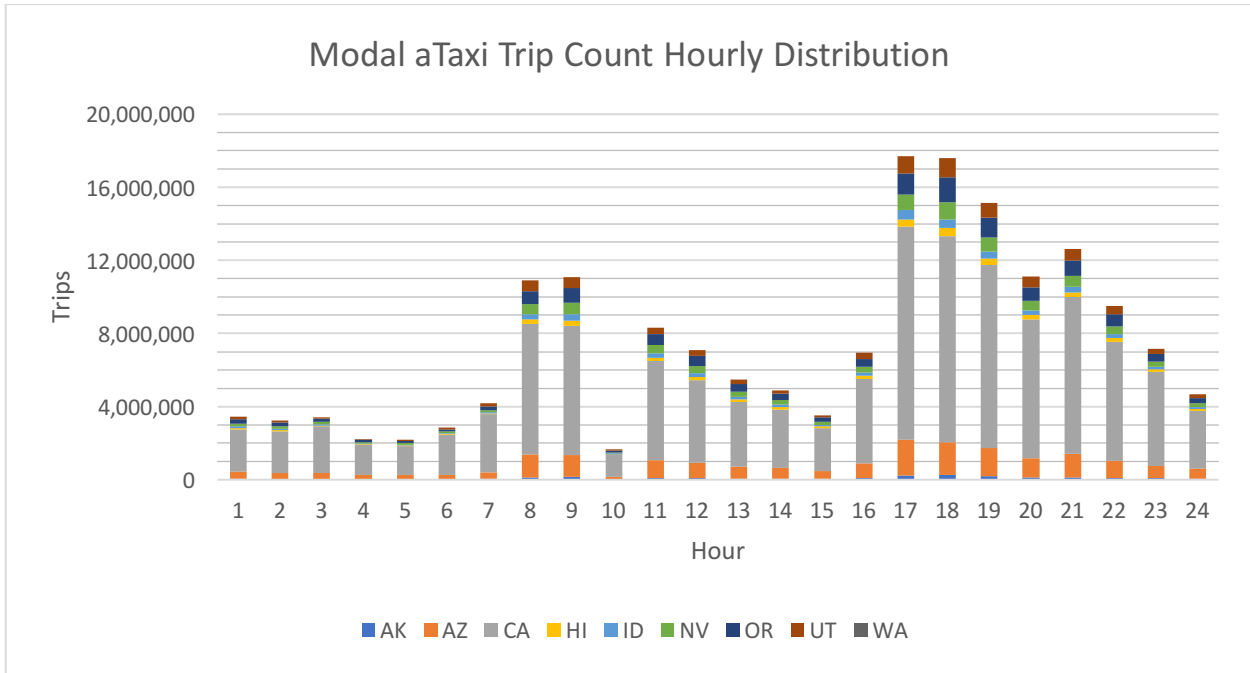
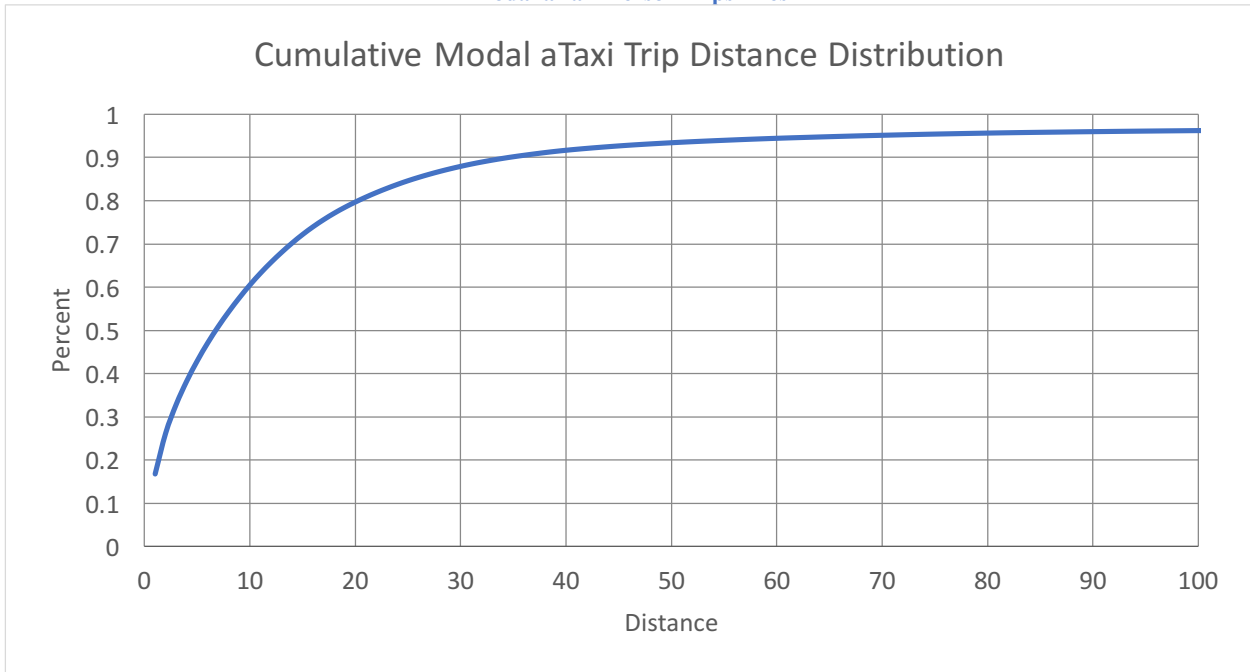


Figure 10: Number of aTaxi trips by hour, broken down by state, according to Nationwide Modal aTaxi PersonTrips Files



The cumulative distance distribution for Modal aTaxi trips is similar to that for all PersonTrips, with approximately 60% of all trips being 10 gcd or less and approximately 80% 20 gcd or less.

Figure 11: Cumulative distribution of trip lengths for all aTaxi trips made by West residents, according to Nationwide Modal aTaxi PersonTrips Files



Assessment of RideShare Potential:

AVO calculations were made with unlimited aTaxi capacity. Total passenger and vehicle distances traveled are 2,568,011,110 and 1,676,471,358 gcd, respectively (1.532 region AVO).

Figure 12: Number of departing aTaxis by average vehicle occupancy (unlimited capacity), according to Nationwide Modal aTaxi PersonTrips Files

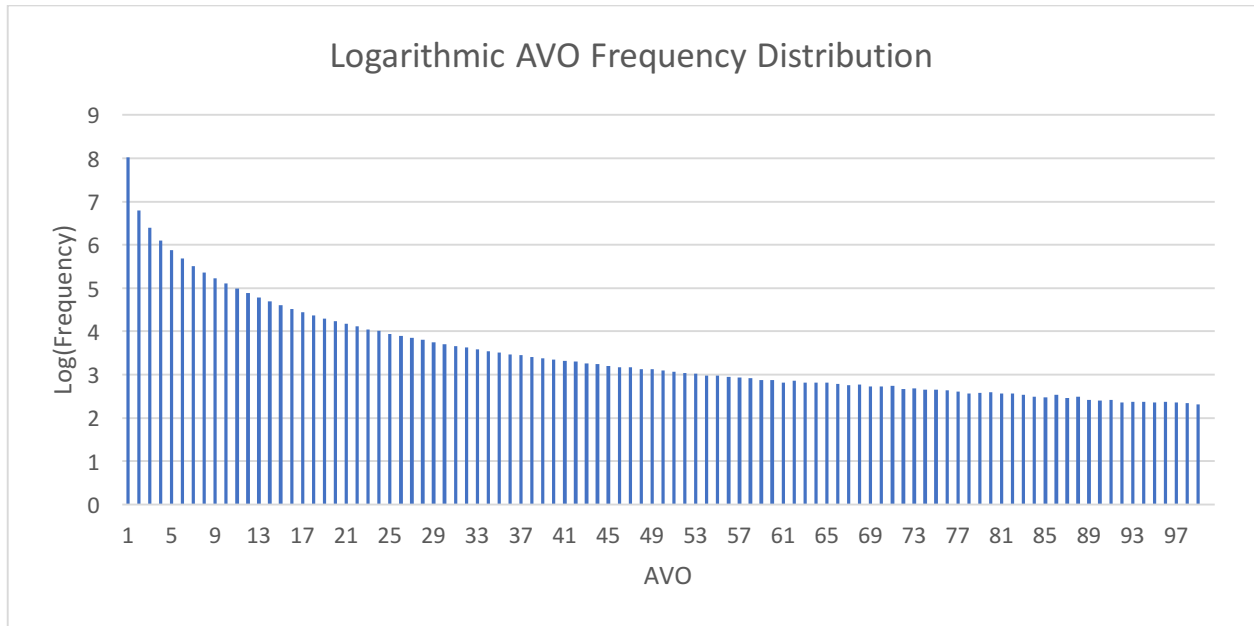
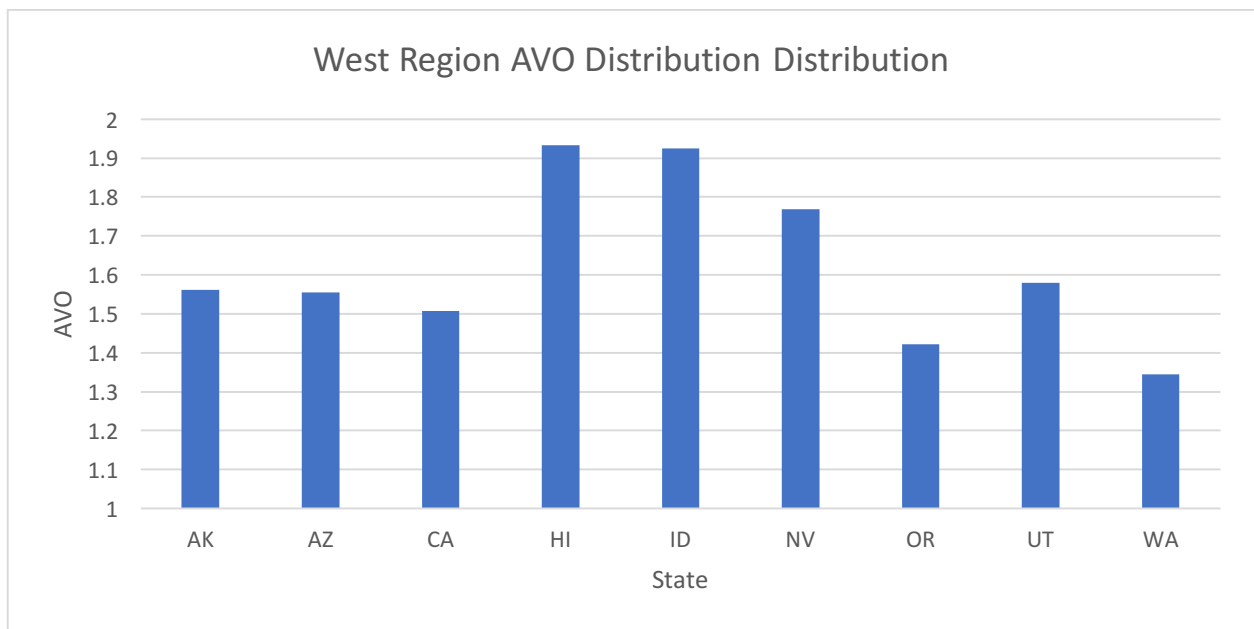
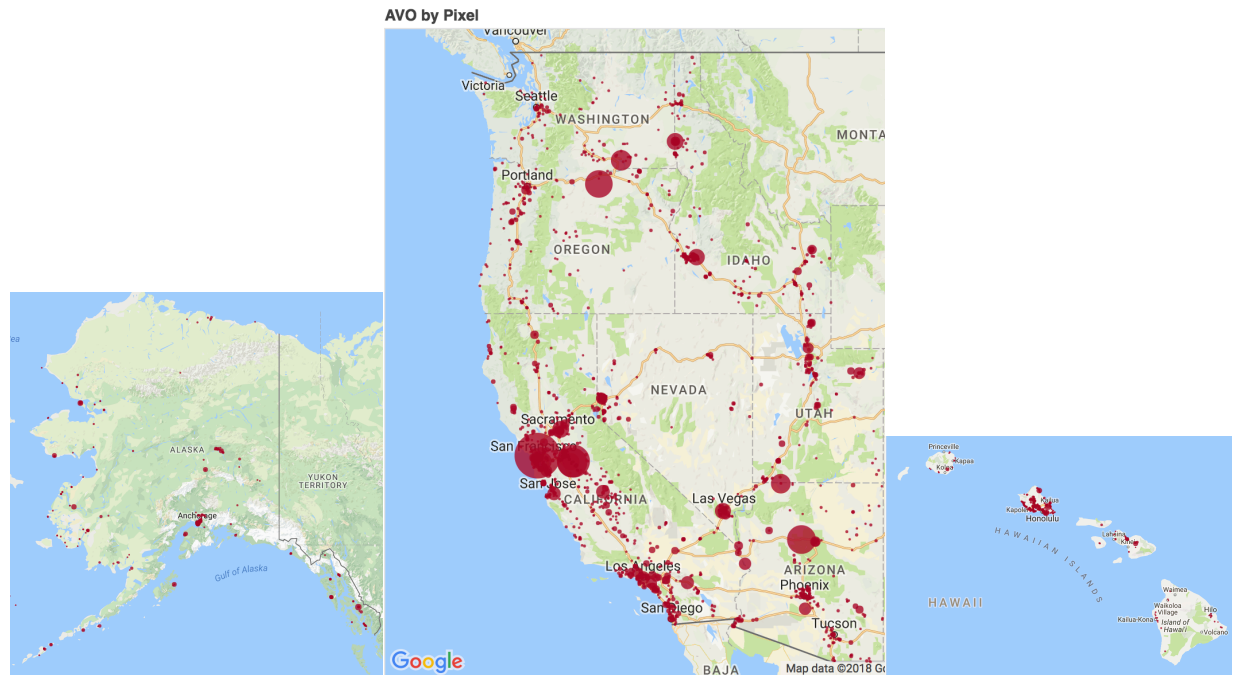


Figure 13: Average vehicle occupancy by state, according to Nationwide Modal aTaxi PersonTrips Files



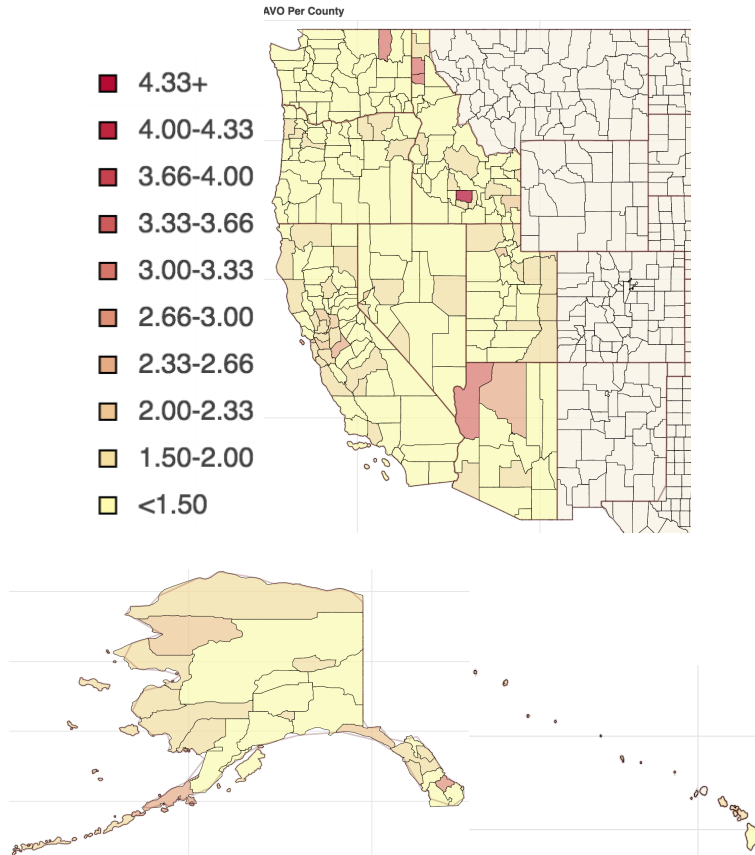
The following plots show the AVO distribution across the United States for pixels with AVO of two or more. The radius of each circle is equal to half the pixel AVO. Large metropolis areas tend to have clusters of pixels with high AVOs, although there are a few pixels in rural areas with unusually high AVOs.

Figure 14: AVO by Pixel, according to Nationwide Modal aTaxi PersonTrips Files. Excludes AVOs of less than 2, radius is proportional to the AVO. Interactive graph (/ORF 467 Final Project Interactive Plots /AVO_Pixel.html)



The county AVO distribution shows that metropolitan counties tend to have higher AVOs. In addition, there are several counties in rural areas with very high AVOs.

Figure 15: AVO by County, according to Nationwide Modal aTaxi PersonTrips Files. Interactive graph (/ORF 467 Final Project Interactive Plots /AVO_County.html)



Of the counties with the top ten AVOs, only one of them encompasses a large metropolitan area.

Figure 16: Chart listing the counties with the top 10 largest AVOs

FIPS	County	AVO
16065	Lincoln, ID	5.642
04015	Mohave, AZ	3.269
16011	Bingham, ID	3.116
53021	Franklin, WA	3.104
16057	Latah, ID	3.085
02275	02275, AK	2.864
06075	San Francisco, CA	2.520
02013	Aleutians East, AK	2.444
06099	Stanislaus, CA	2.395
04005	Coconino, AZ	2.333

The following two charts that show the distribution of AVO throughout the day suggest that the distribution of AVO is also bimodal, with peaks during morning peak and evening rush.

Figure 17: AVO by time-of-day, according to Nationwide Modal aTaxi PersonTrips Files

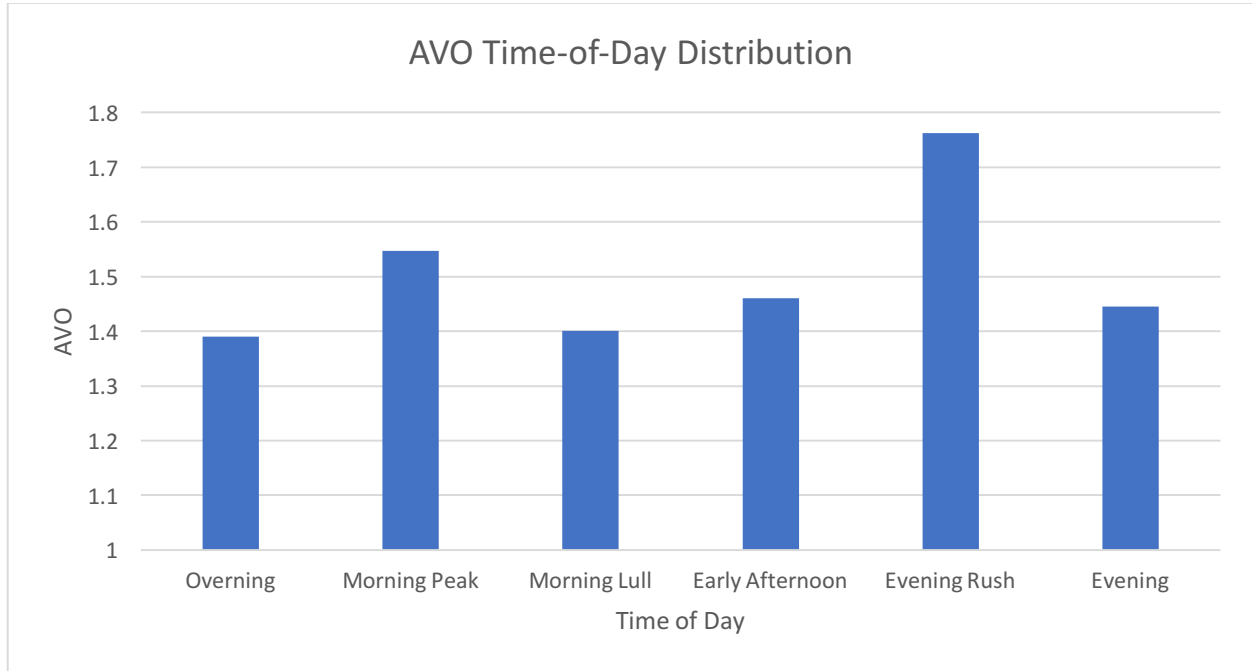
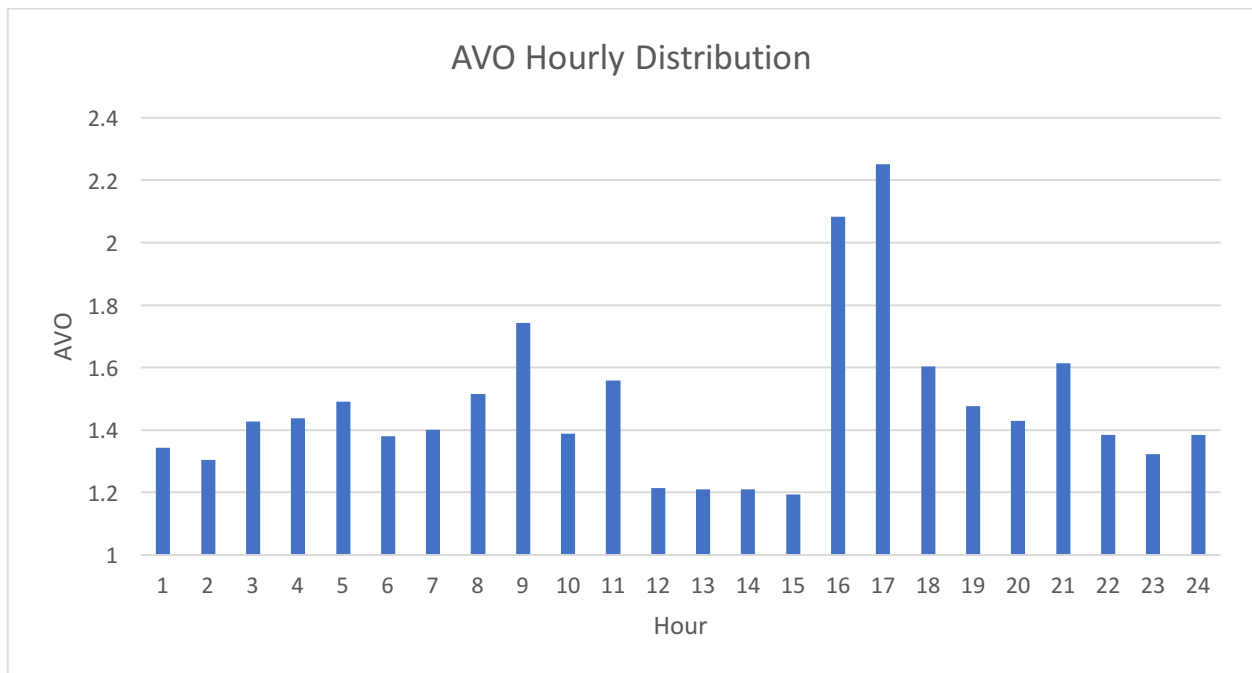


Figure 18: AVO by hour, according to Nationwide Modal aTaxi PersonTrips Files



Assessment of Start-of-Day Demand and End-of-Day Supply:

For all Supply/Demand analysis, aTaxi have a five-passenger capacity. Graphs for 1x1 Pixels were unable to be created due to computational limits of the computers available to the team. Graphs for 5x5 and 10x10 pixels are included below, and the raw data for 1x1 pixels are included in the supplementary spreadsheet. The radius for the circles on these graphs are equal to $\text{Log}_4(\text{Number of Taxis}) / 2$ with red circles indicating demand and green circles supply.

Figure 19: Start of Day Demand and End of Day Supply by SuperPixel, based on Nationwide Modal aTaxi PersonTrips Files. Radius is proportion to the log of the number of aTaxis.

[Interactive graph \(/ORF 467 Final Project Interactive Plots / SoD_Supply_EoD_Demand_5x5.html\)](#)

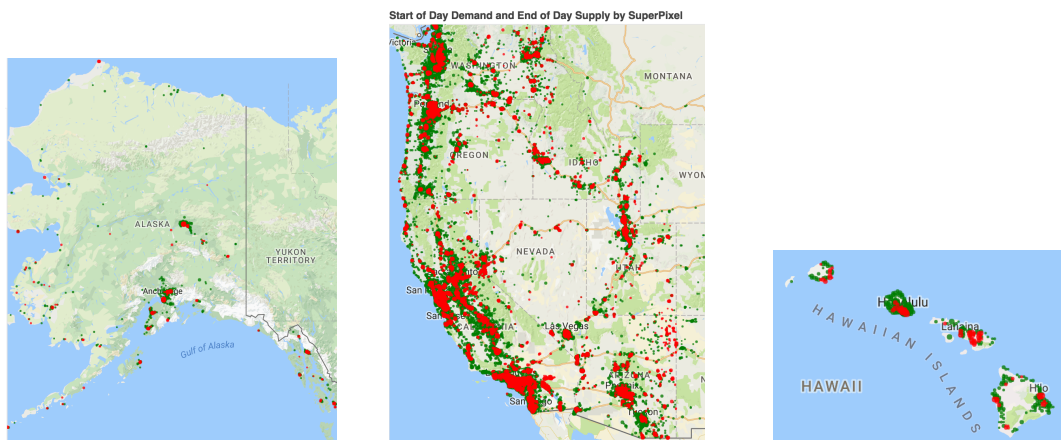
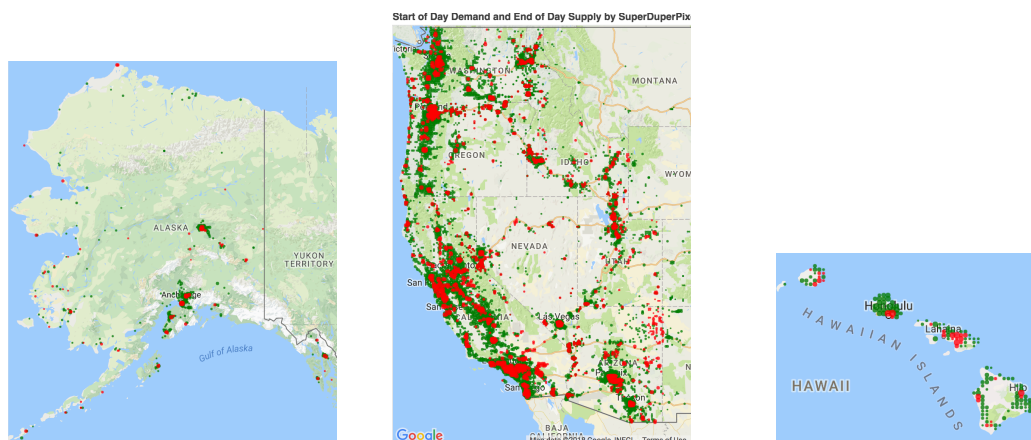


Figure 20: Start of Day Demand and End of Day Supply by SuperDuperPixel, based on Nationwide Modal aTaxi PersonTrips Files. Radius is proportion to the log of the number of aTaxis.

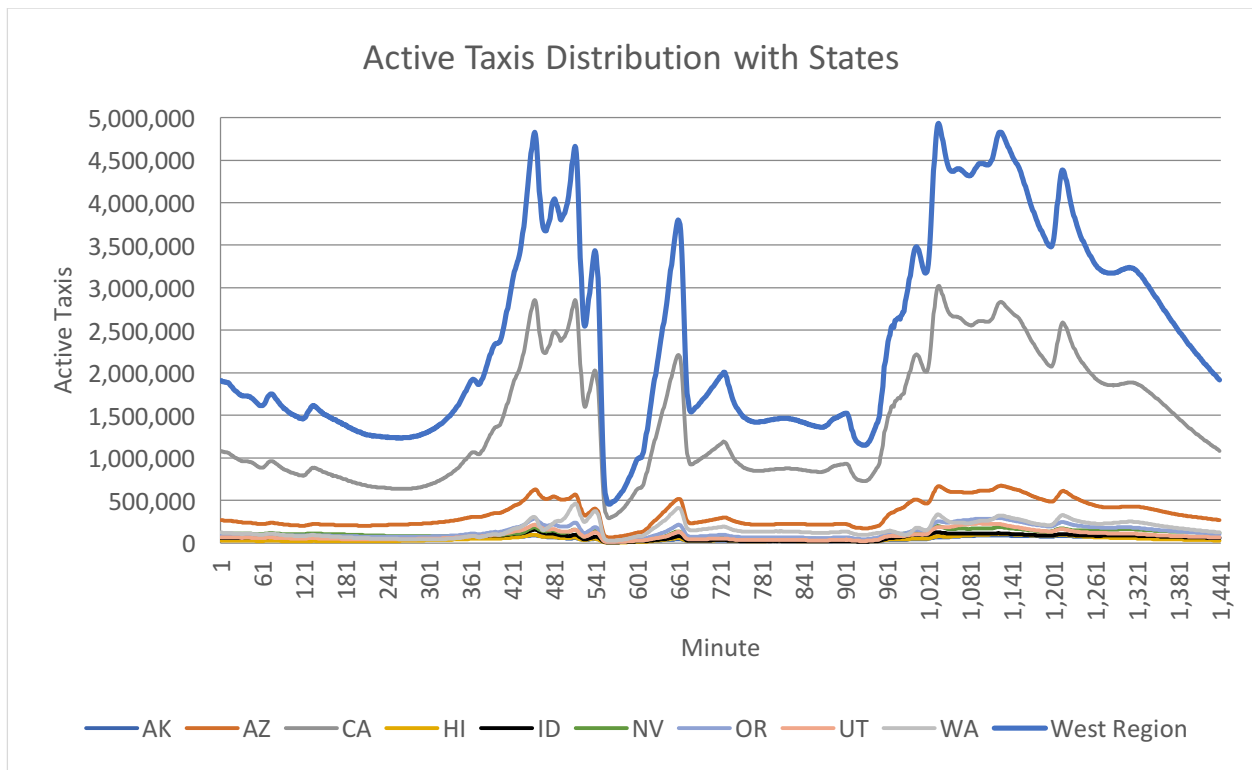
[Interactive graph \(/ORF 467 Final Project Interactive Plots / SoD_Supply_EoD_Demand_10x10.html\)](#)



4,838,588 taxis need to be repositioned at the end of the day using 1x1 Pixels (2,617,637 for 5x5 and 2,010,458 taxis for 10x10). Because only a very small percentage of trips end up in states outside the region, there is not a strong need to collaborate with neighboring regions.

The maximum amount of active taxis during the day occurs during the early evening at 4,930,881 active taxis; this amount is also the minimum fleet size needed. The chart below includes the aggregate region active taxi distribution in addition to a by state breakdown.

Figure 21: Number of Active aTaxis throughout the day, according to Nationwide Modal aTaxi PersonTrips Files. Chart is not cumulative.



Assessment of Business Case for your Regional aTaxi Mobility-for-All System:

Expenditures:

For the business case, one unit of gcd will be assumed to be the equivalent of one mile. ~5,700,000 taxis must be purchased in order to acquire 1.15x the minimum fleet size needed. Approximately ~2,600,000,000 passenger miles and ~1,700,000,000 vehicle miles are traveled each day, which averages to ~300 miles per vehicle per day (~110,000 miles per vehicle per year). Assuming that taxis can be repositioned on a 10x10 pixel system, ~2,000,000 taxis must be repositioned per day. Assuming that repositioning distance is 30% the total daily distance incurred by the repositioned taxi, the total average miles traveled per vehicle per year increases to ~120,000 miles, which will be rounded to 125,000 miles when calculating vehicle replacement and maintenance costs. Assuming that these cars have ~150,000 miles life expectancies (thus requiring 5/6 of the fleet to be replaced every year) and can be purchased for \$60,000 per vehicle, total vehicle acquisition expenditures add to \$342,000,000,000 per year (or \$285,000,000,000 for if only 5/6 of the fleet is repurchased every year). Assuming that these vehicles require maintenance costs of \$2,000 every ~25,000 miles and an annual insurance cost of \$1,000 per vehicle, other annual vehicle expenditures combine to ~\$74,100,000,000.

Other expenditures include those for energy and parking. Assuming that an acre of land can be leased for \$350 per year and that each acre provides 150 parking spots, \$15,000,000 must be spent annually on land rental. In addition, assuming that each vehicle has mileage efficiency of 4 miles per kWh and that all cars need to be recharged once per day, 425,000,000 kWh needs to be produced per day. Assuming that solar energy can be installed for a one-time fee of \$5 per Wh and then produced for \$0.10 per kWh, energy expenditures total to a one-time expense of \$2,125,000,000,000 (to be amortized over ten years) plus an additional \$42,500,000 annually.

Revenue:

The pricing model for ride-sharing services will be to charge each customer \$0.65 per mile traveled. Assuming that each vehicle averages 30 miles per hour, this averages to \$19.50 per hour of service. Because passengers using ride-sharing services would be able to regain productivity previously lost to driving and because the average American's hourly wage is ~\$25, the pricing model becomes a win-win for customers because each customer essentially receives a net \$5.50 per hour of value through the ride-share service. Because there is ~940,000,000,000 miles traveled per year, annual revenue is projected to be \$611,400,000,000.

Other potential sources of revenue include integrating mail and package delivery as well as food and other consumable delivery (similar to UberEats/GrubHub) to the passenger ride-share service as well as the sale of consumer data. These are not included in the current analysis.

Annual Balance Sheet:

Assets (\$MM):		Liabilities (\$MM):	
Ride-Share Revenue:	611,400	Vehicle acquisition (5/6 of fleet / year):	285,000
		Vehicle maintenance and insurance:	74,100
		Solar installation (paid over 10 years)	212,500
		Energy generation:	43
		Parking land rental:	15
Total:	611,400	Total:	571,658
		Net Worth:	39,742

Recommendations for Further Analysis:

This analysis was conducted for only Modal aTaxi trips and thus did not use any trip rerouting to rail stations or airports. Nevertheless, there were still trips in the data files which by distance could have been rerouted to a transit station. Further analysis that includes trip rerouting for trips originating in the region to transit stations and for trips arriving from transit stations to their final destinations can be conducted to more accurately illustrate Start of Day Demand and End of Day Supply distributions, and such analysis can then include trips for different modals.

In addition, current analysis estimates vehicle distance as the greatest passenger trip distance, and further analysis can also be done to improve this estimate by accounting for the actual drop off sequence for passengers based on the passengers' final destinations. Another area of consideration is opening ride-sharing opportunities to passengers that not only share starting positions and destinations but also to passengers whose starting positions and destinations lie on the taxi's current route.

More analysis can also be conducted for taxi repositioning. The current estimate of each taxi's repositioning costs is 30% of the vehicle's daily distance traveled, and this can be improved through optimization models that calculate the minimum total distance traveled by the entire fleet to reposition EoD Supply positions to SoD Demand positions. Additionally, more analysis can be done to determine when and how many additional repositions should be made during the day to minimize overall cost.

For the business case, more analysis can be conducted to develop a better pricing model by measuring how much of the market size the ride-sharing can actually capture based on the pricing strategy. More analysis can also be made to better estimate cost of land rental as well as energy generation capabilities based on the geography.

Team Responsibilities:

Kevin Wu – Lead on coding component and write-up

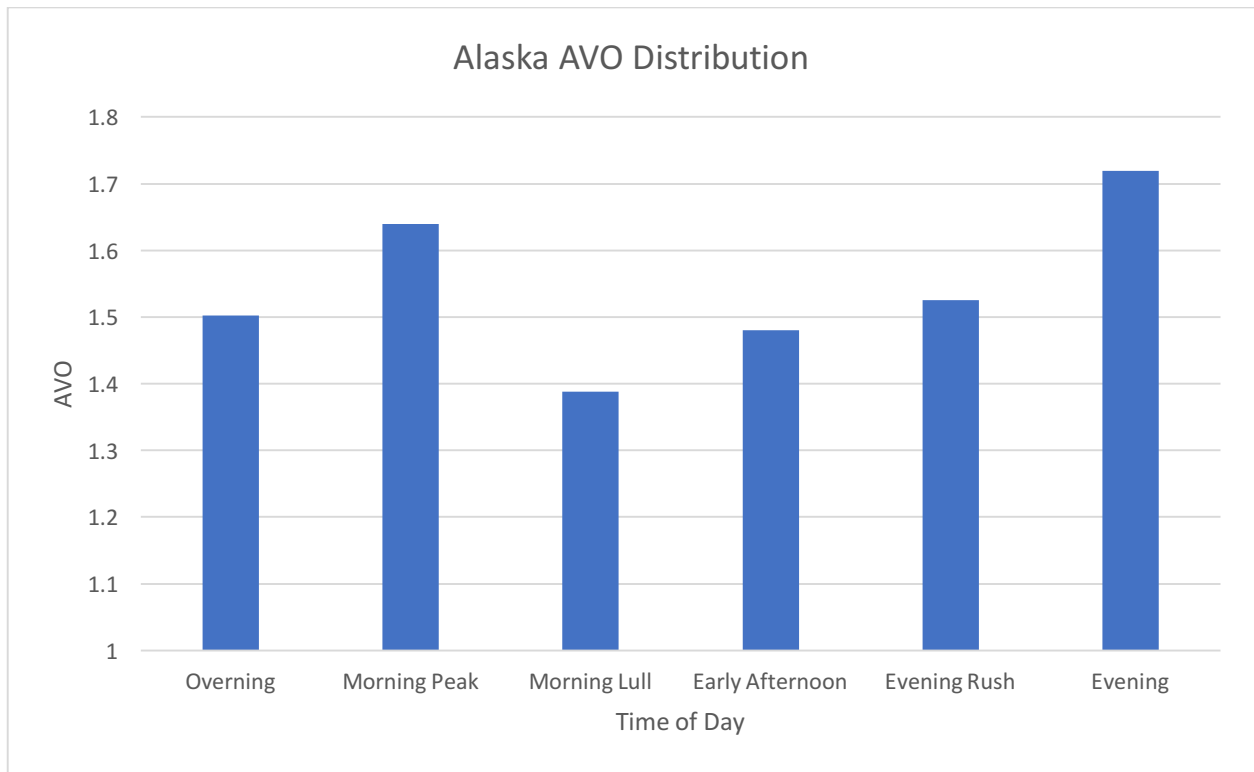
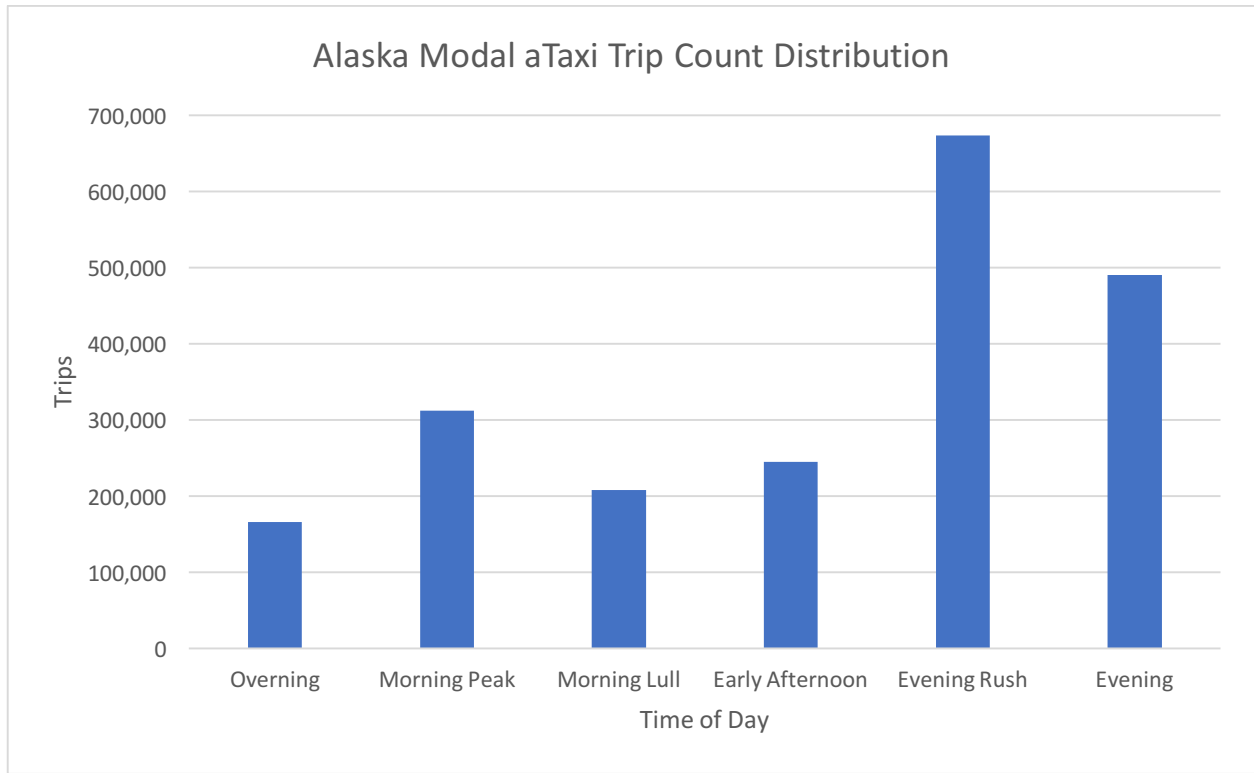
Anna Erkalova – Lead on data visualization and chart creation

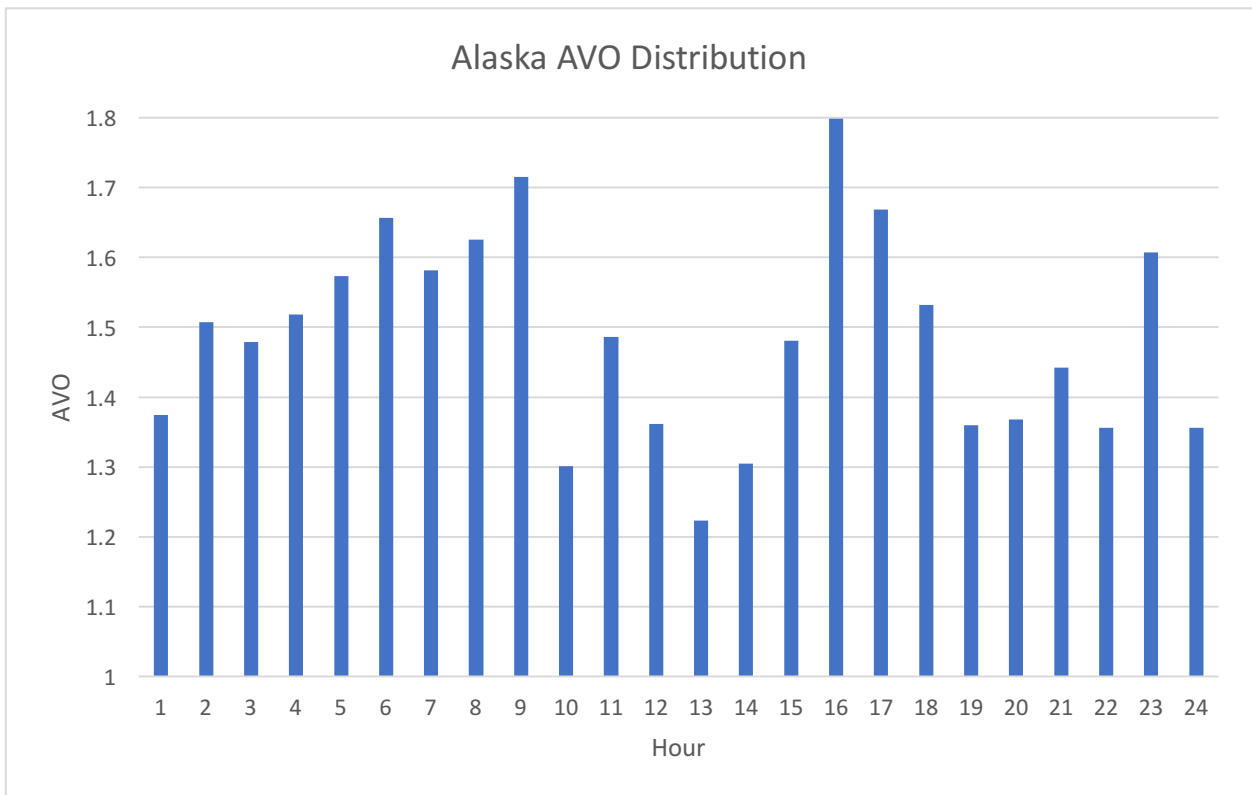
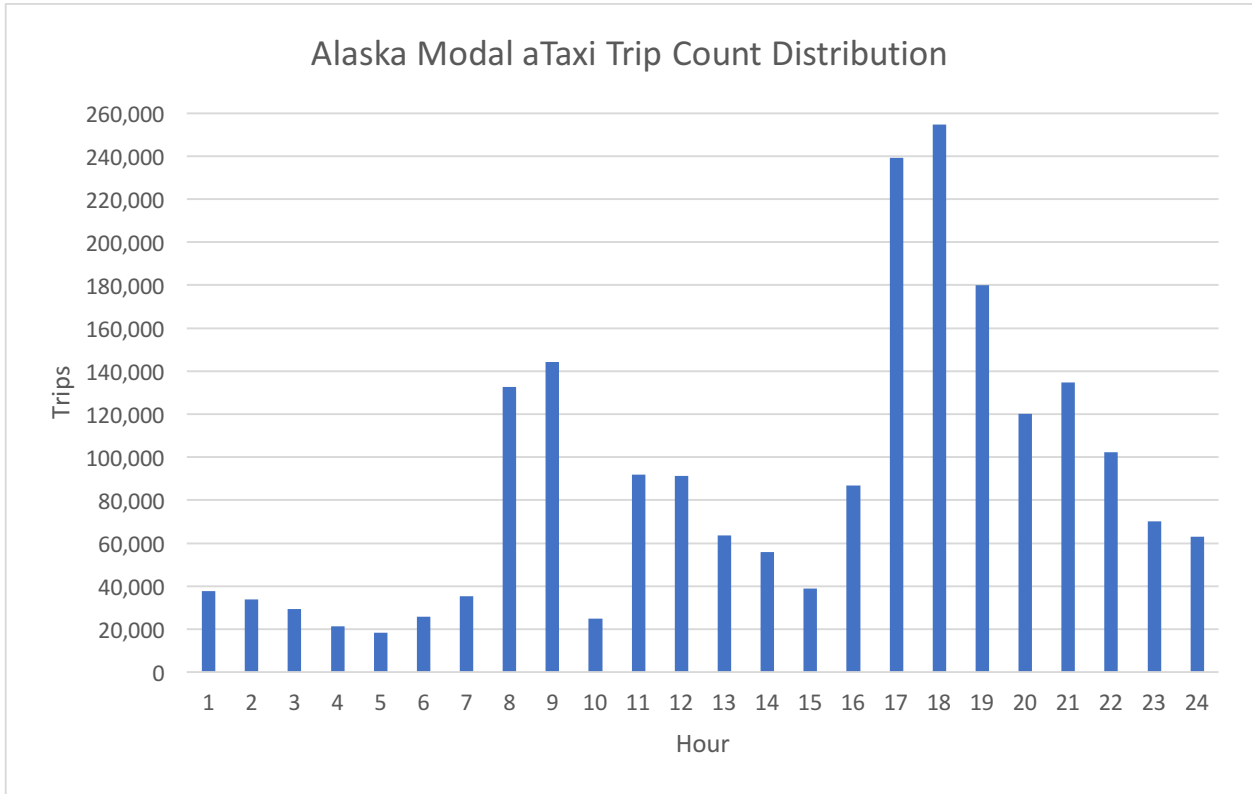
Mike Wagner – Assisted with write-up and business case

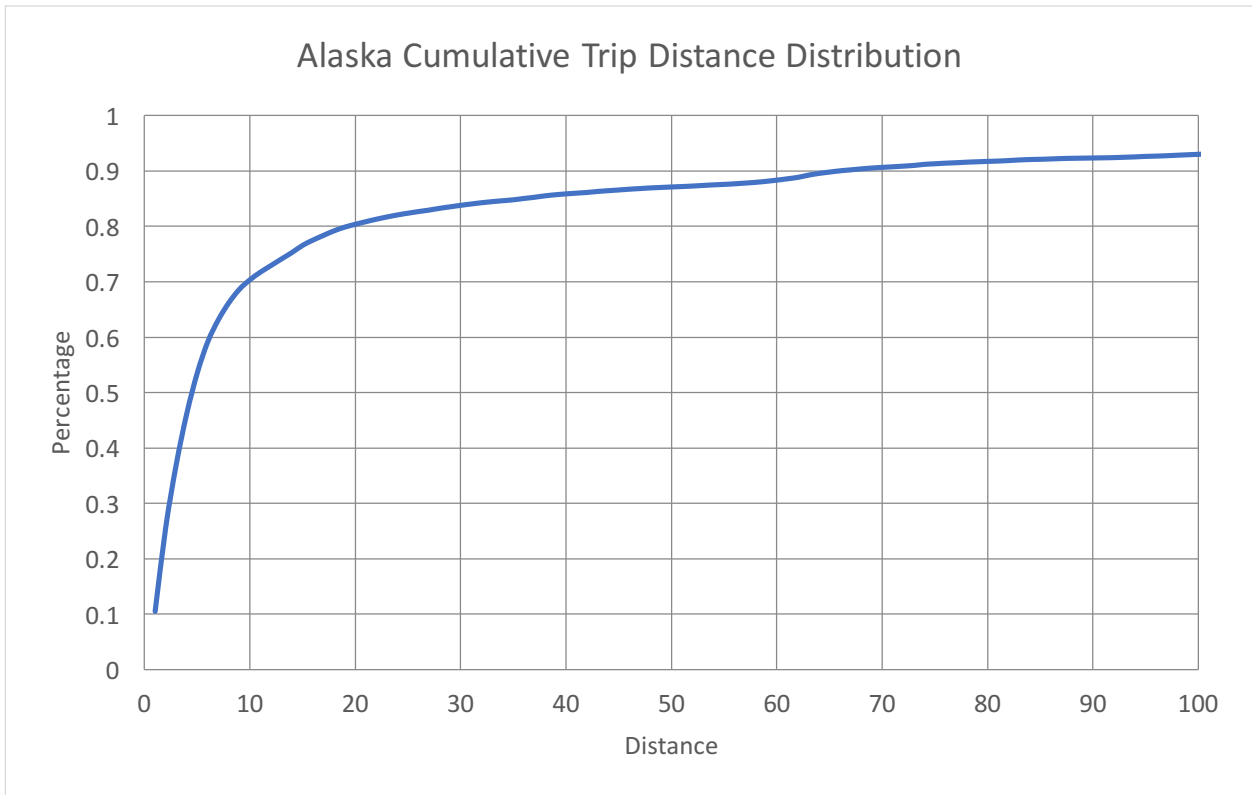
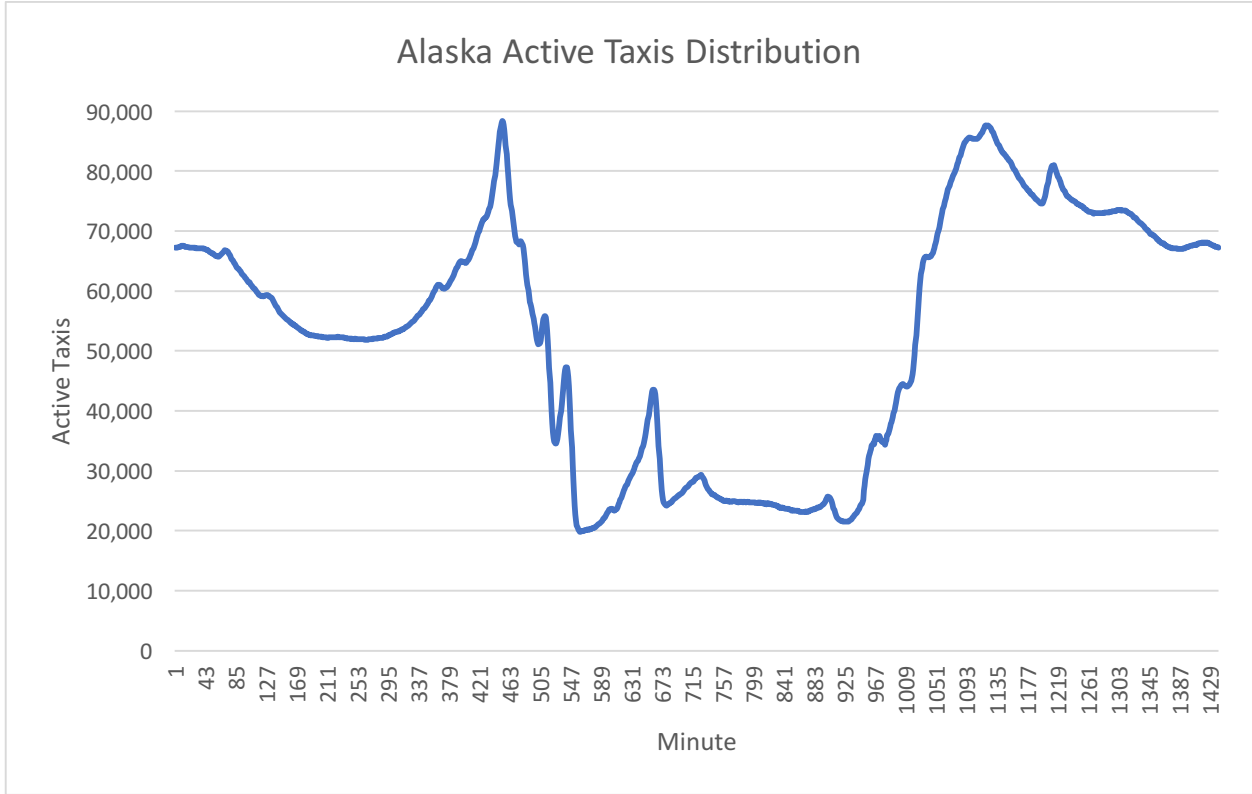
Bulut Cakmak – Assisted with coding component and data visualization/chart creation

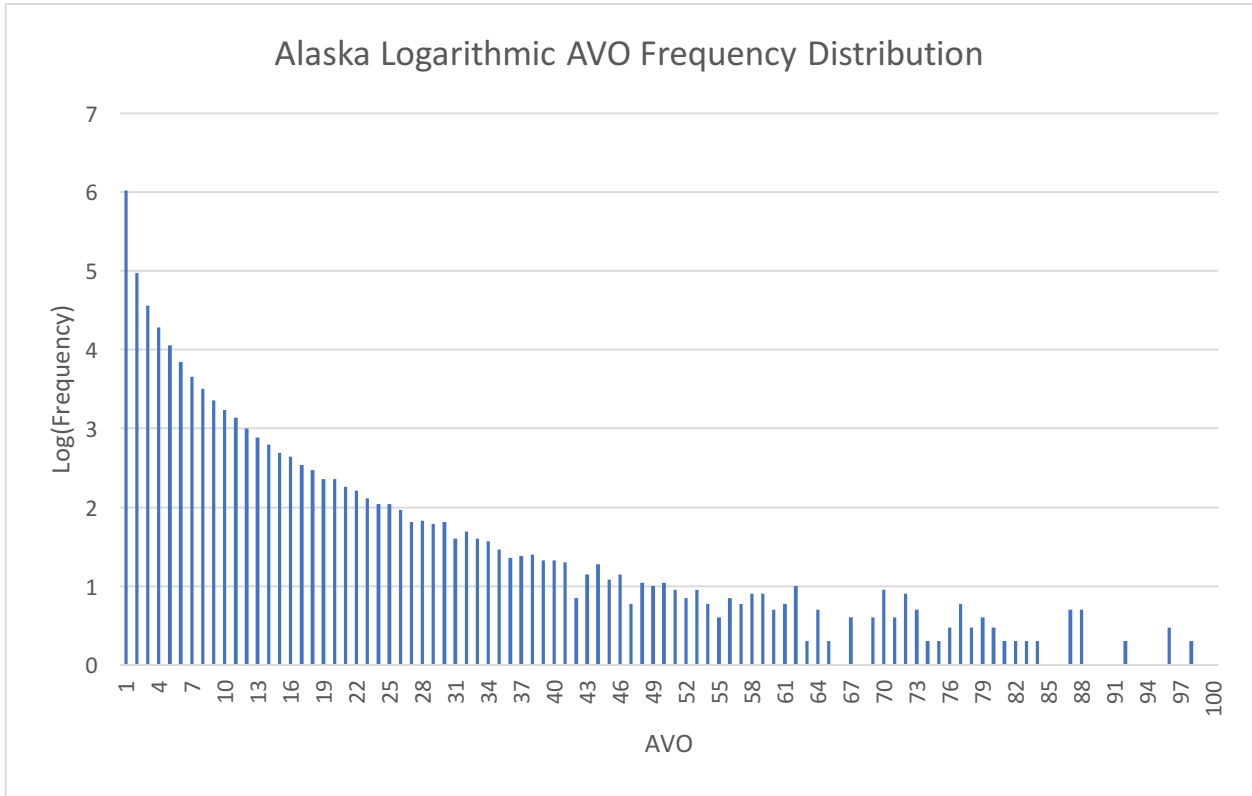
Appendix:

Alaska:

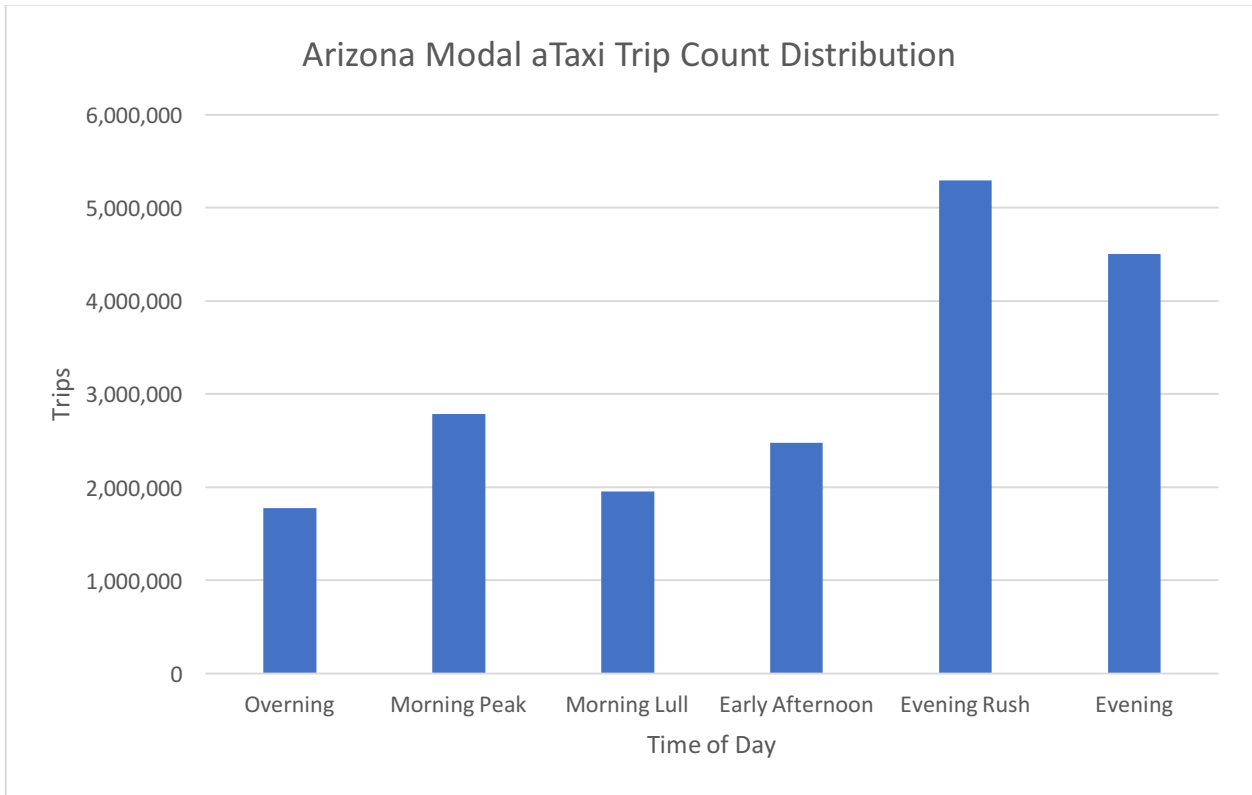


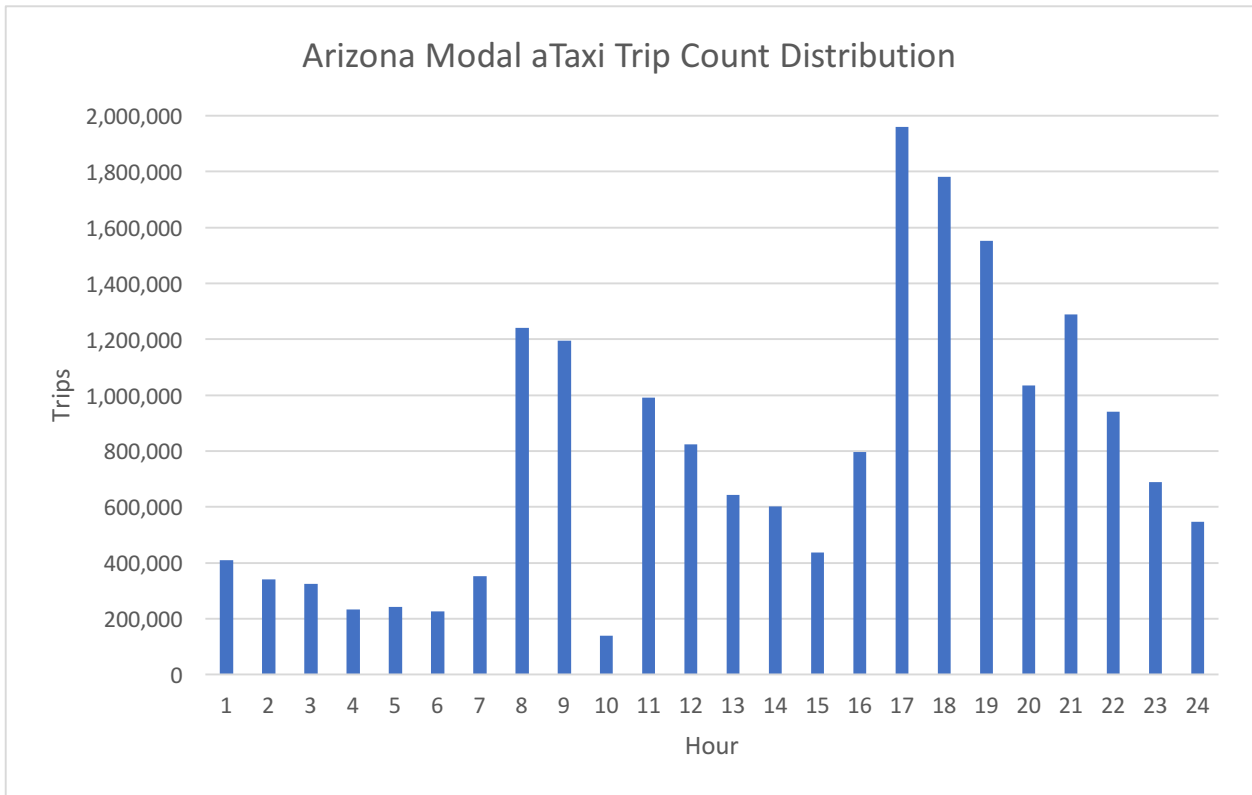
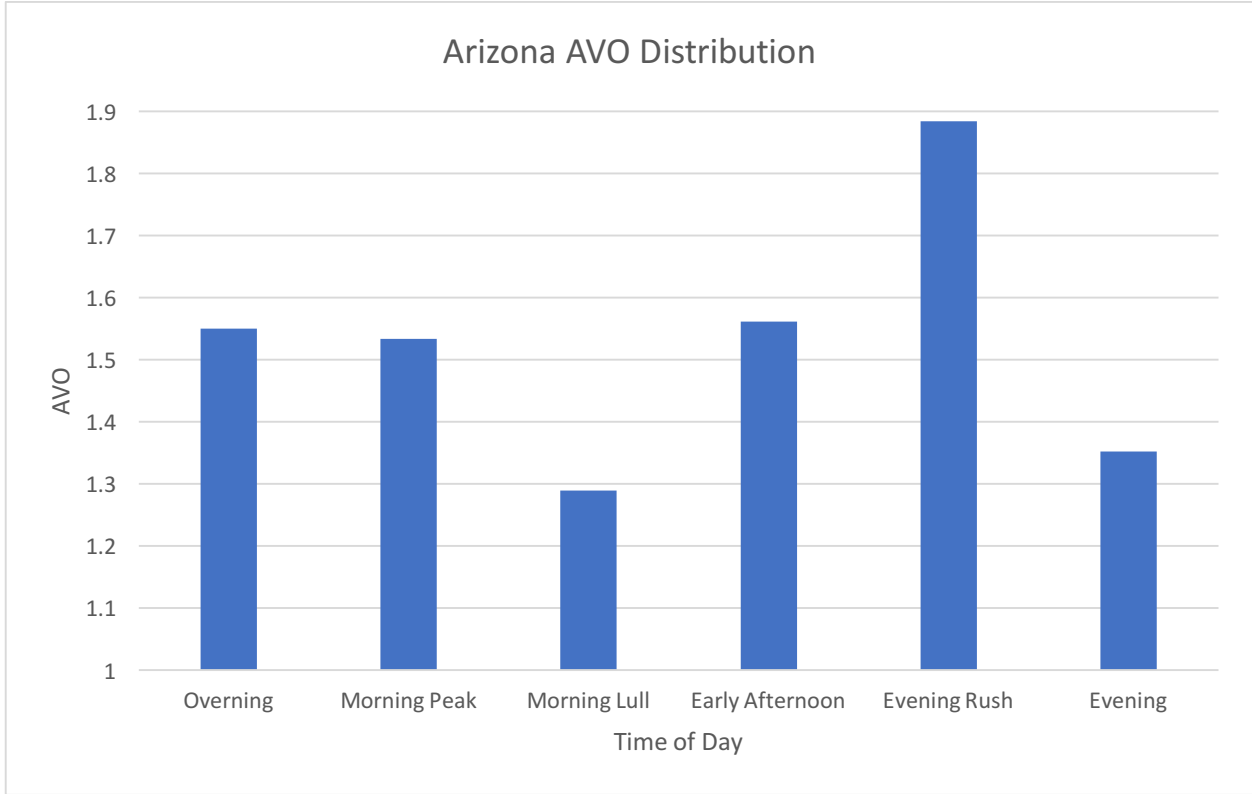


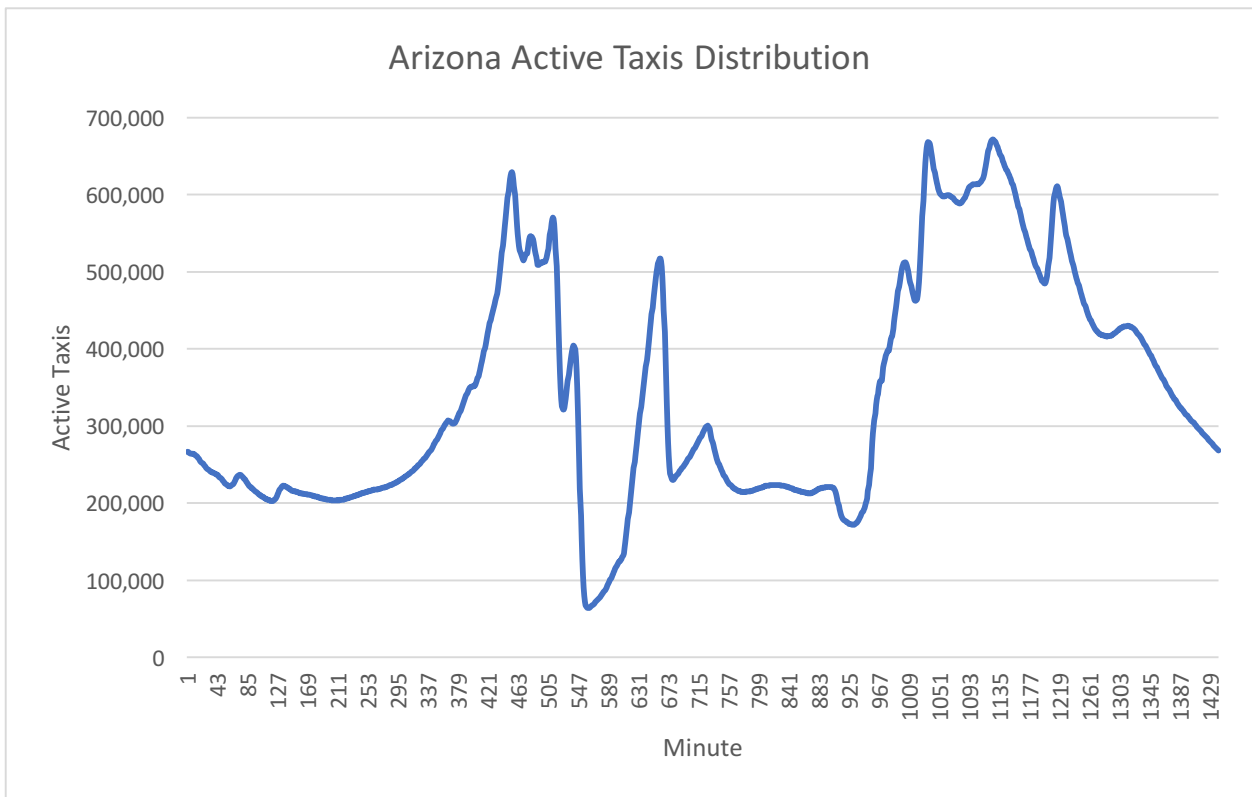
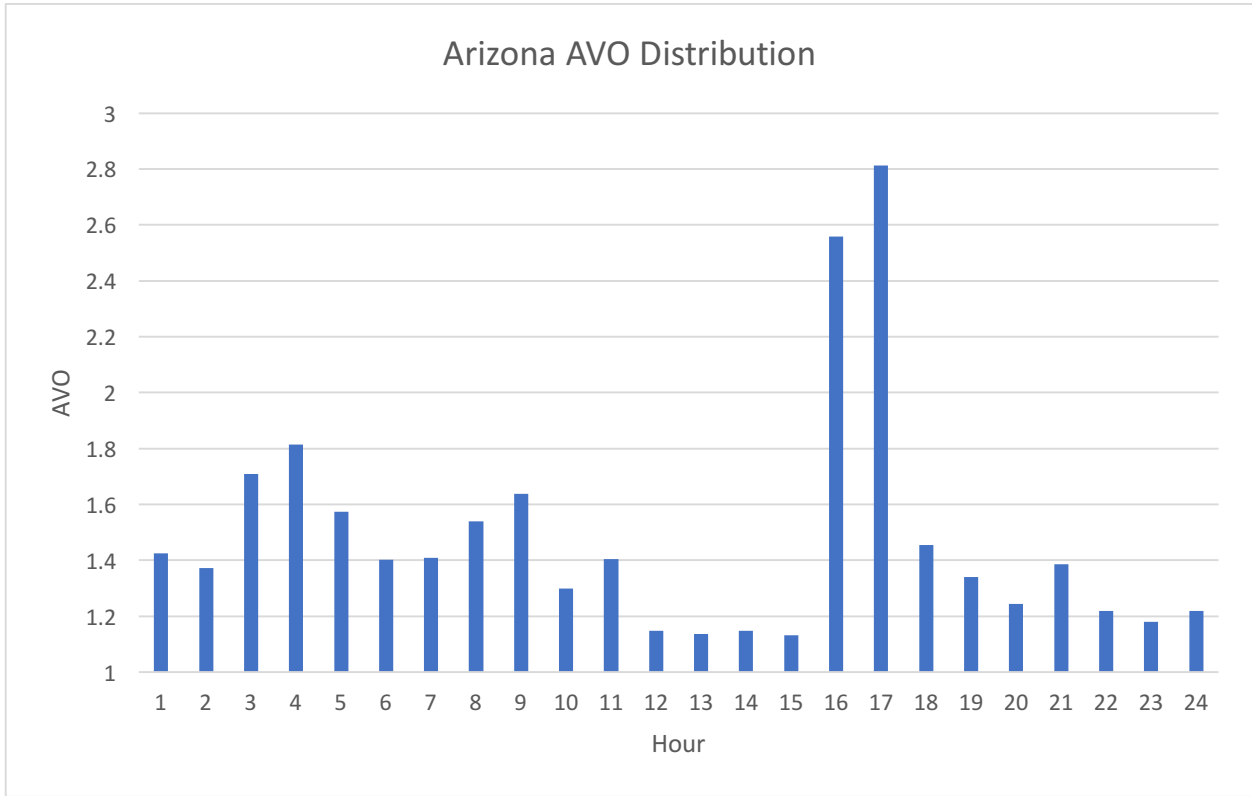


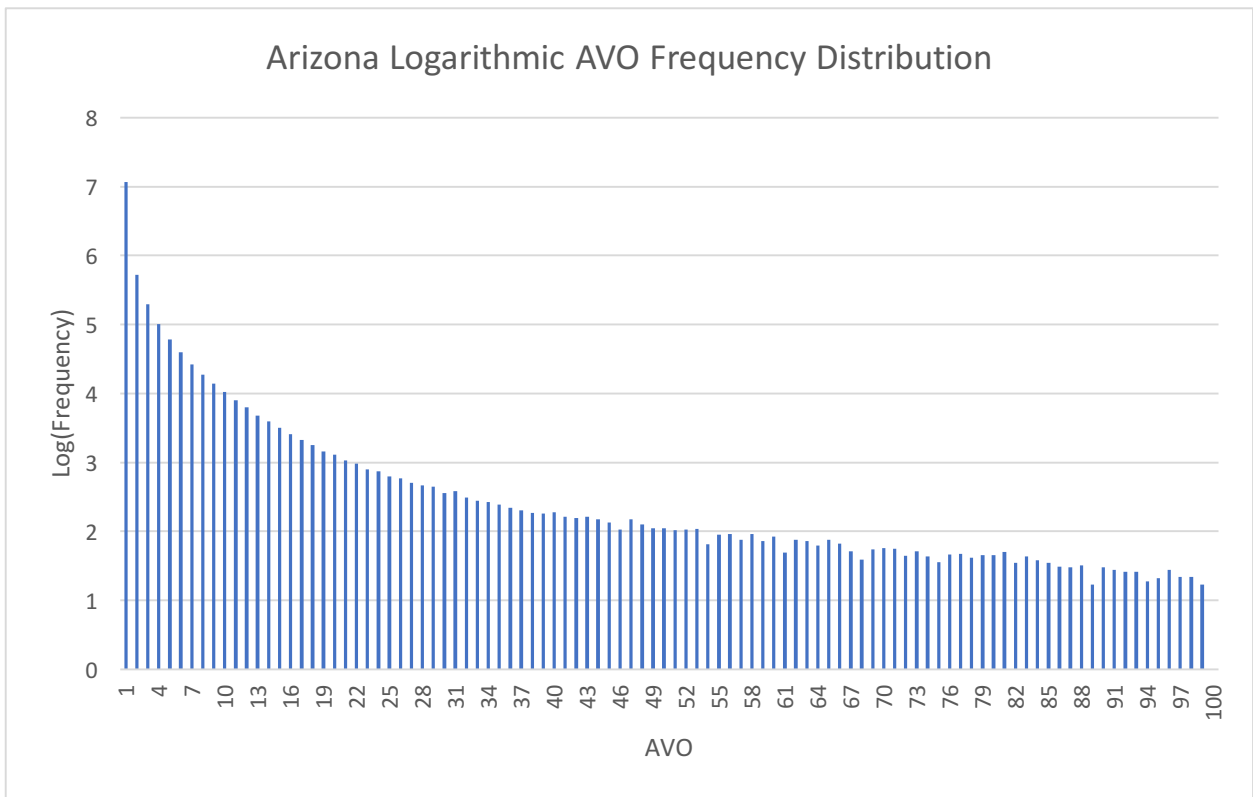
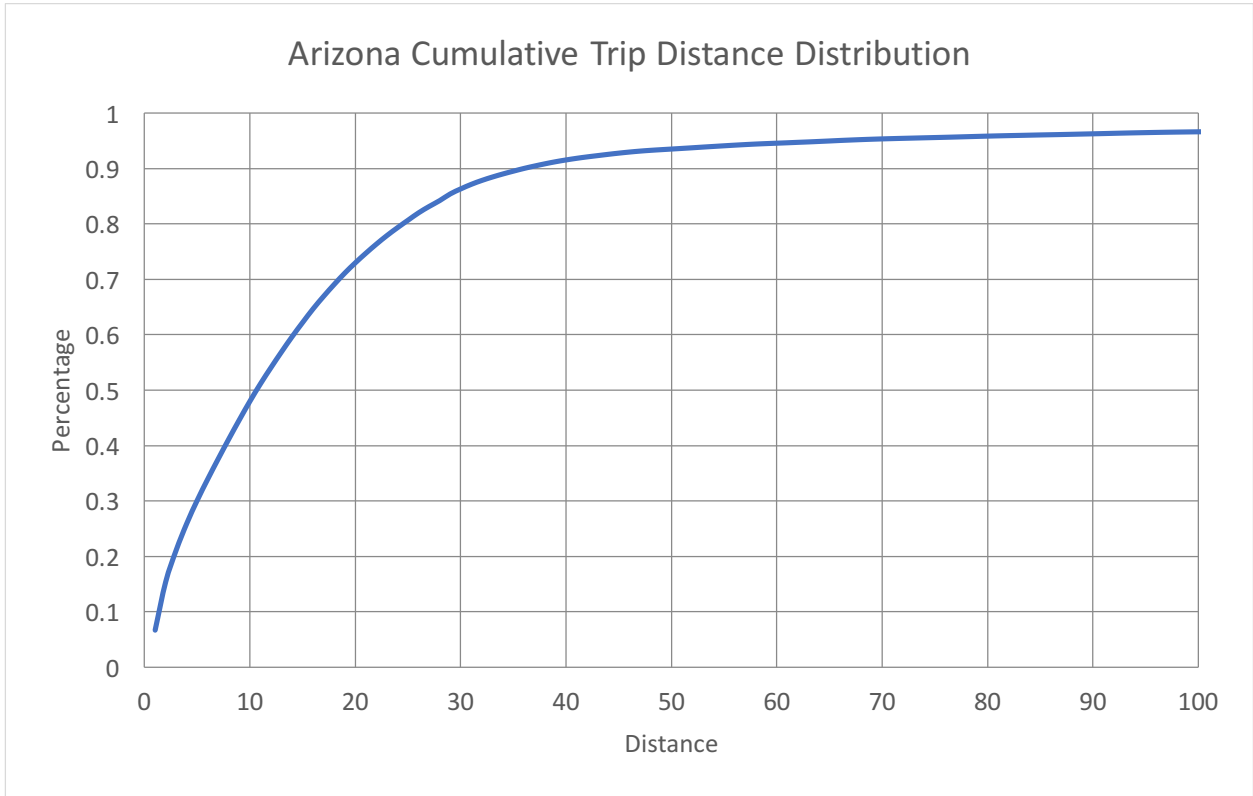


Arizona:

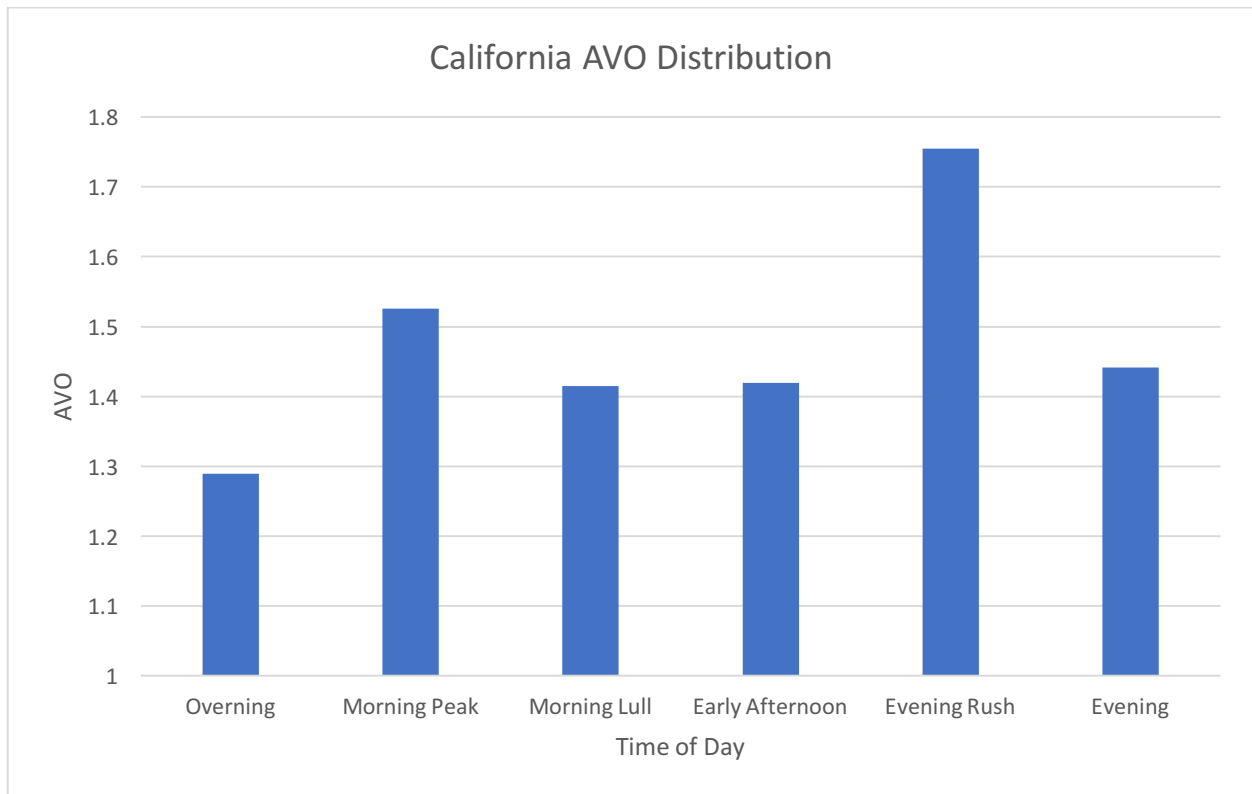
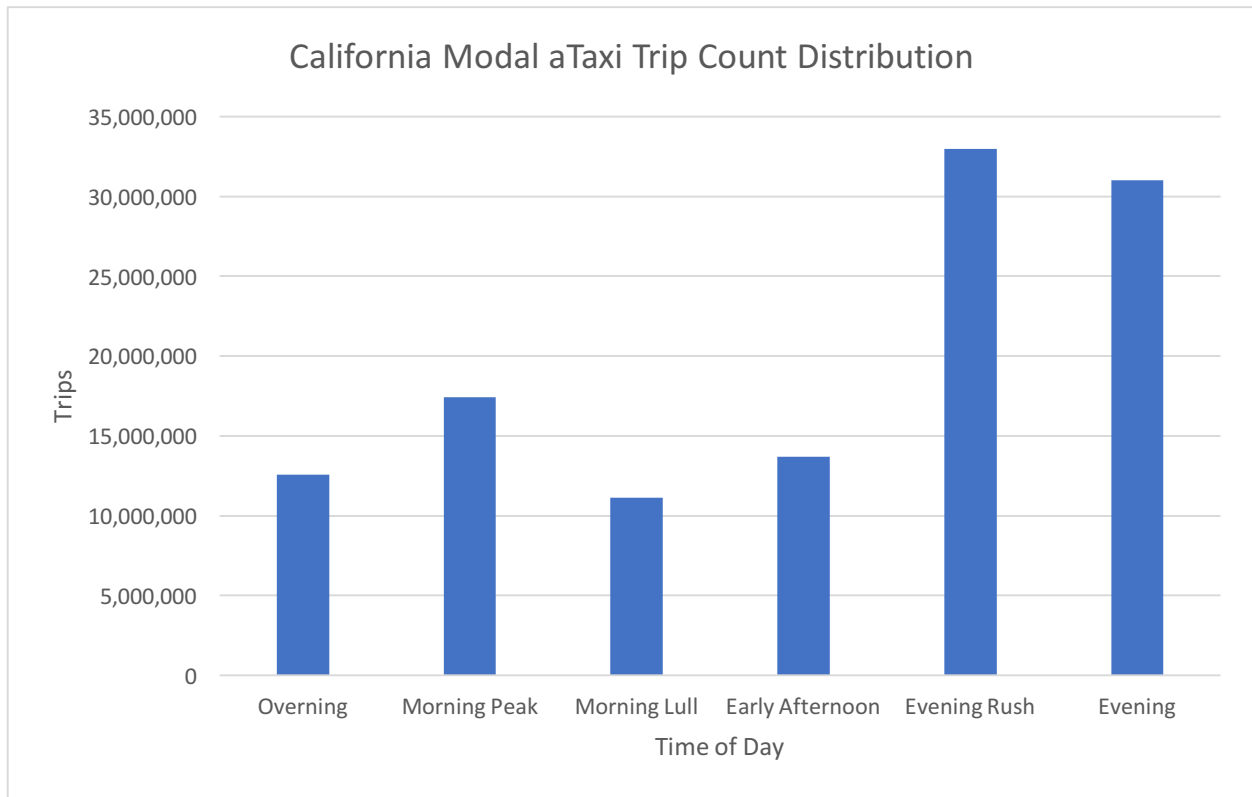


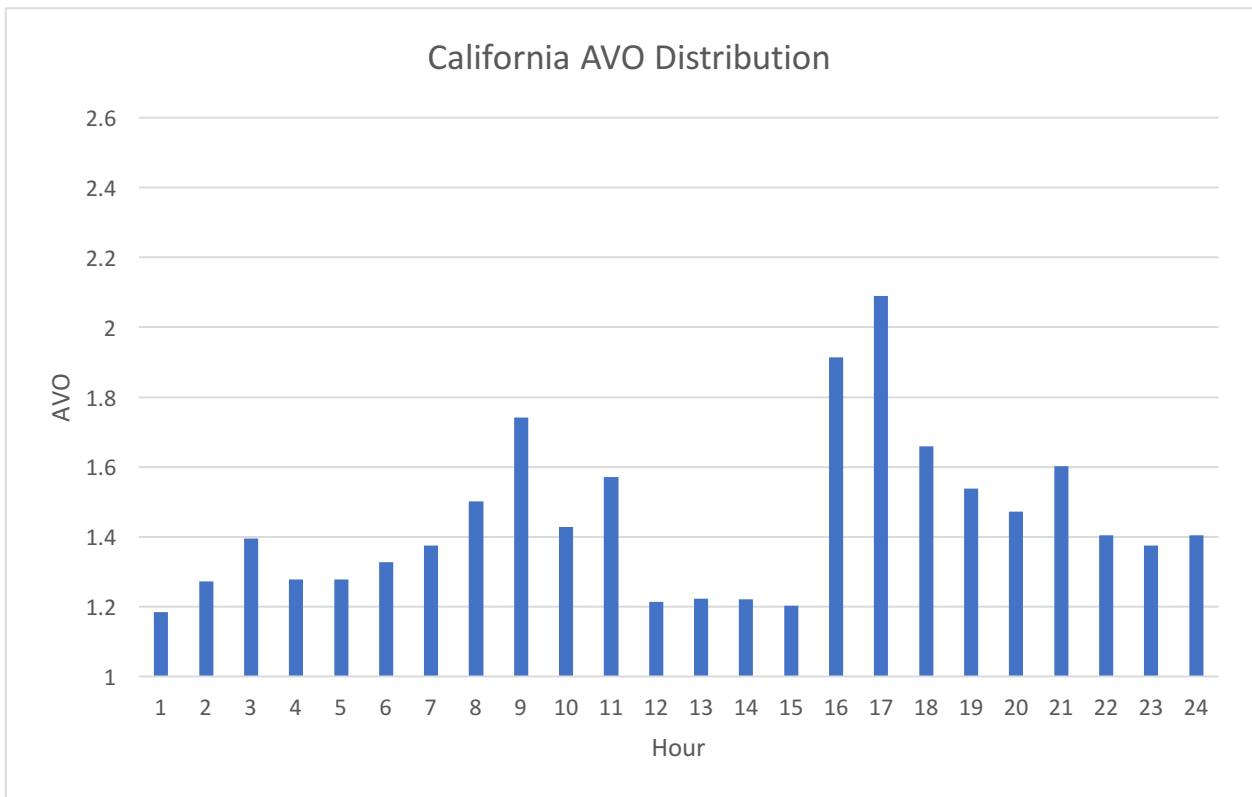
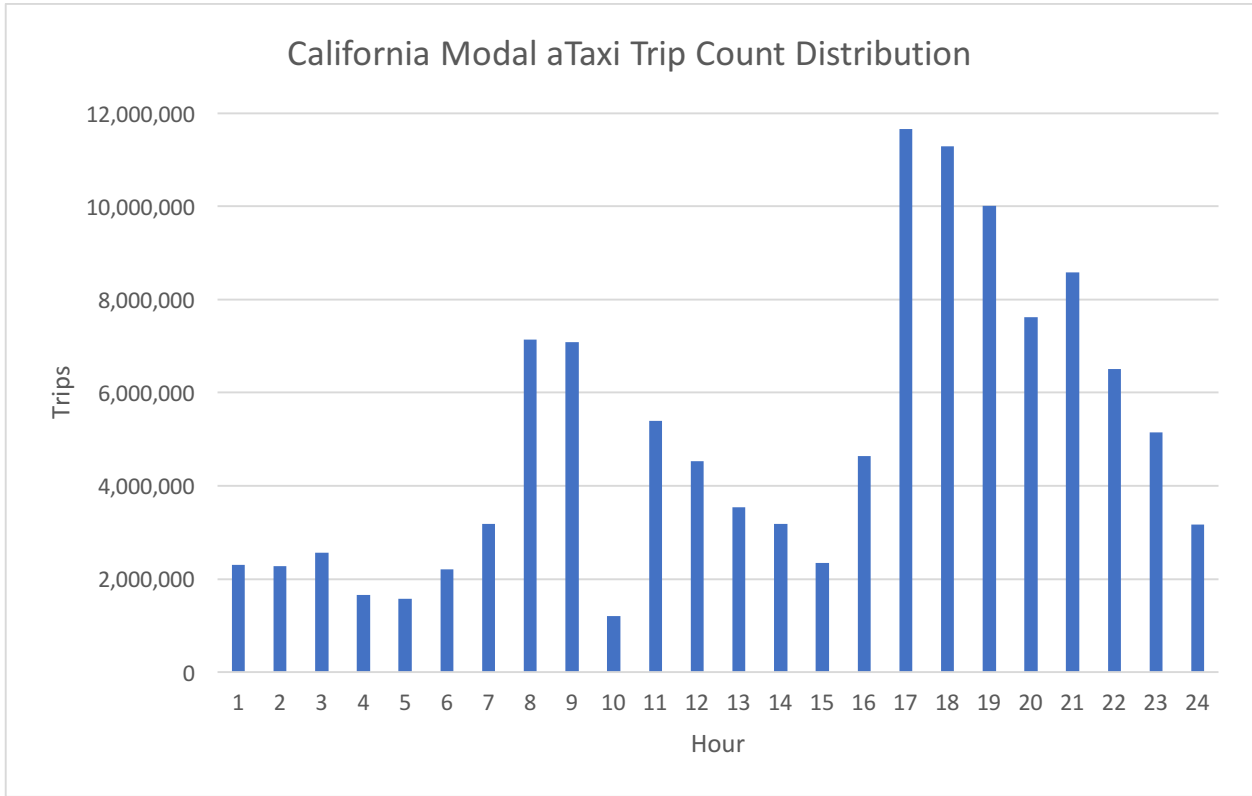


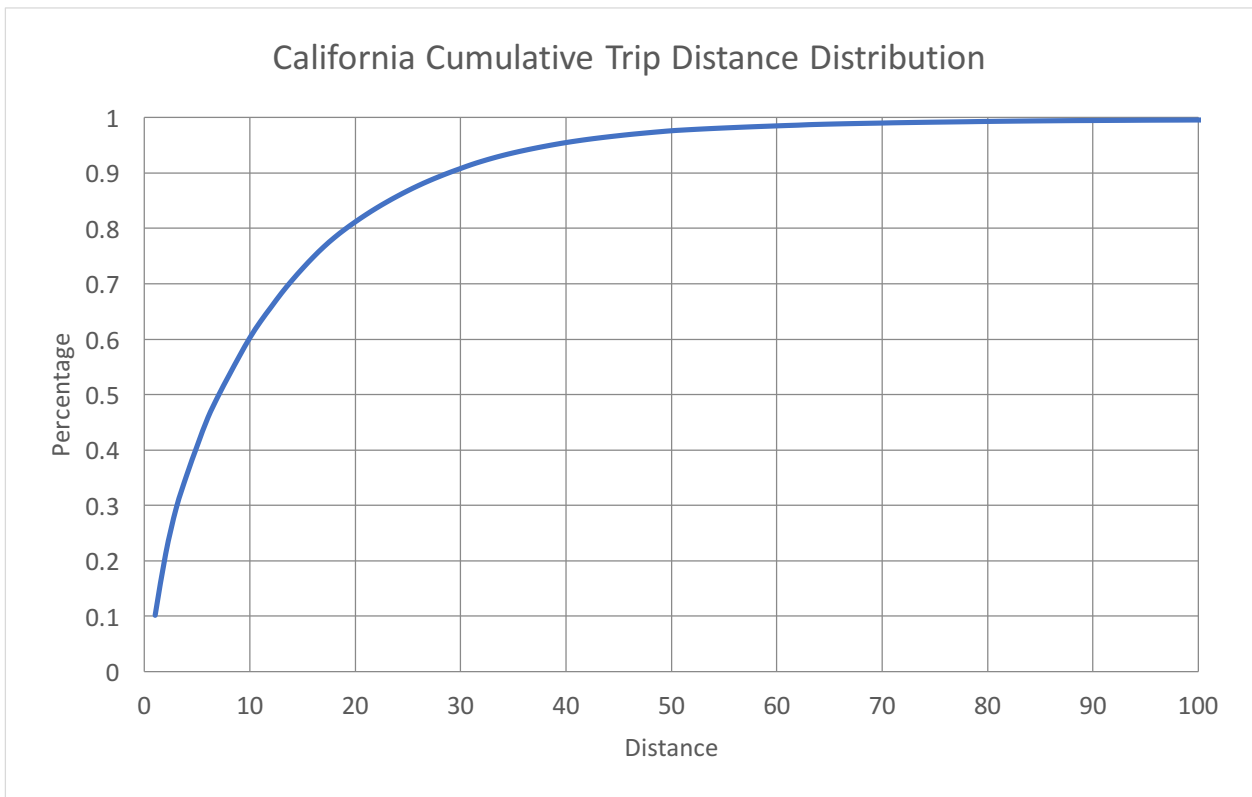
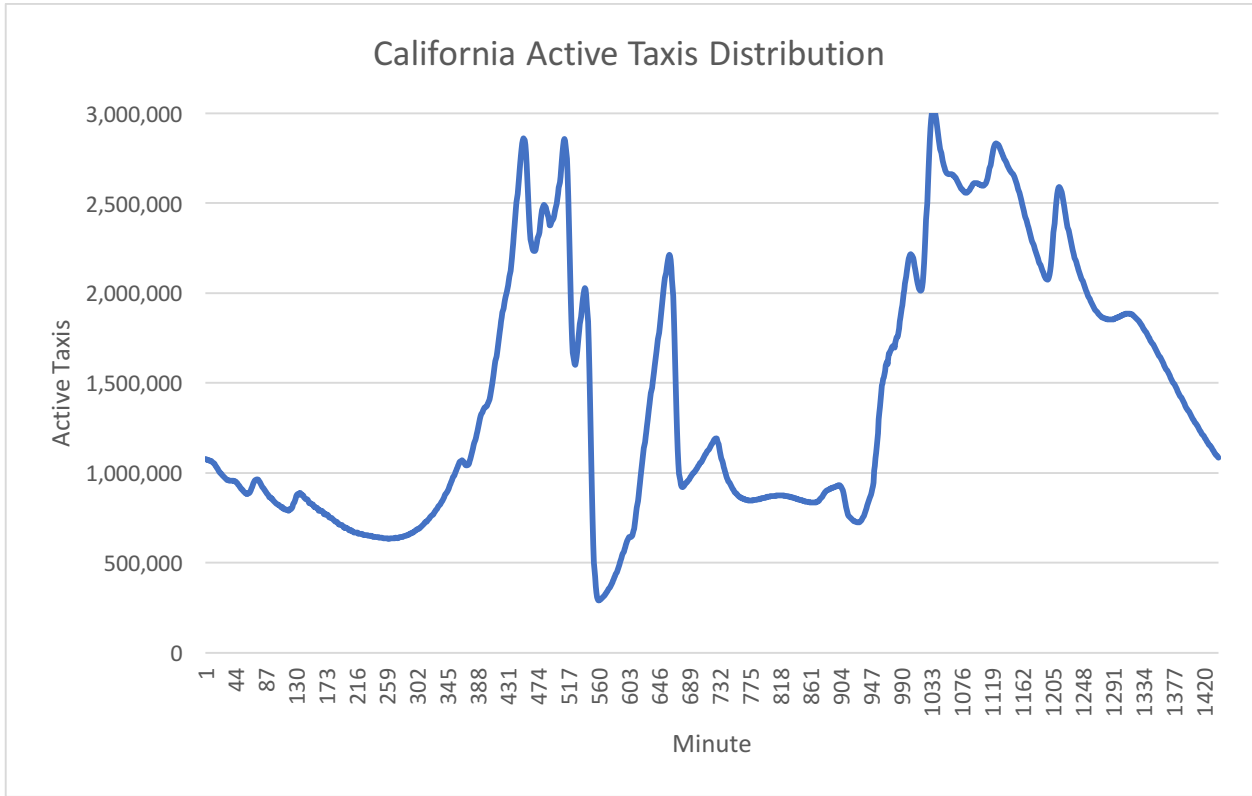


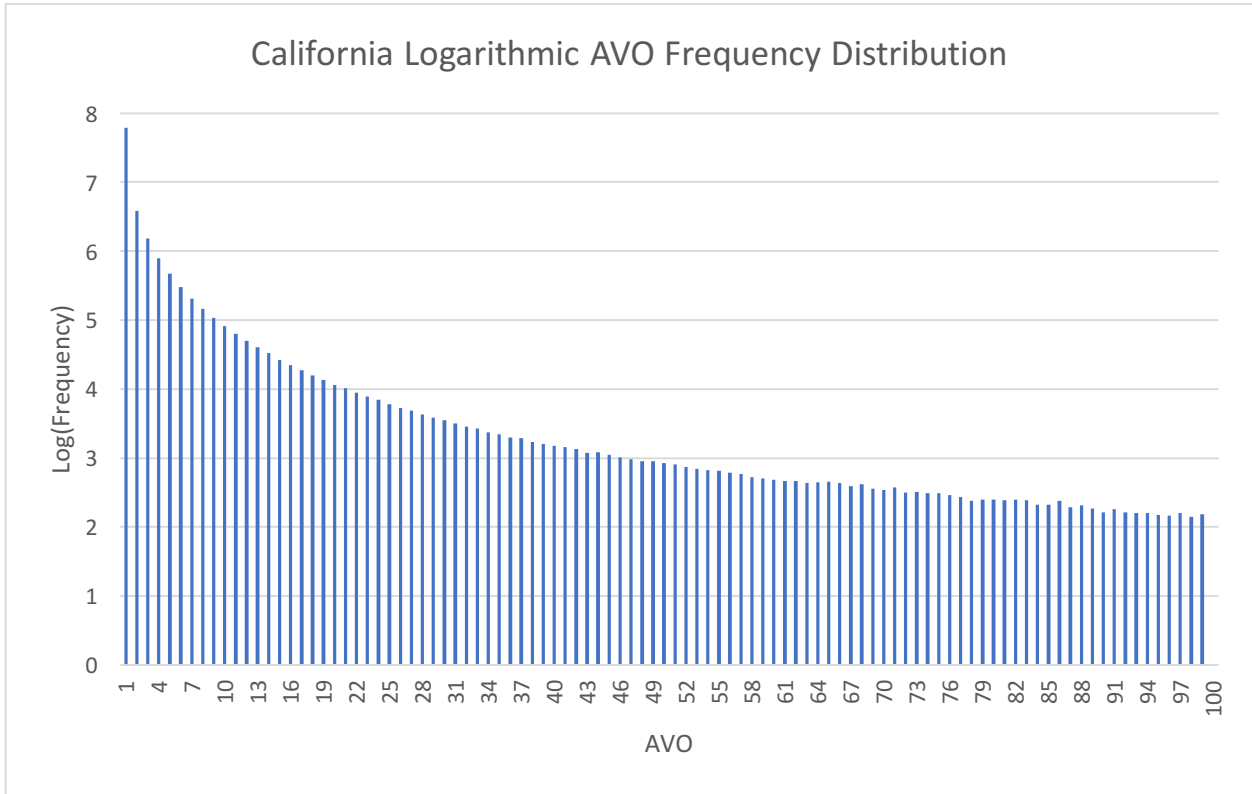


California:

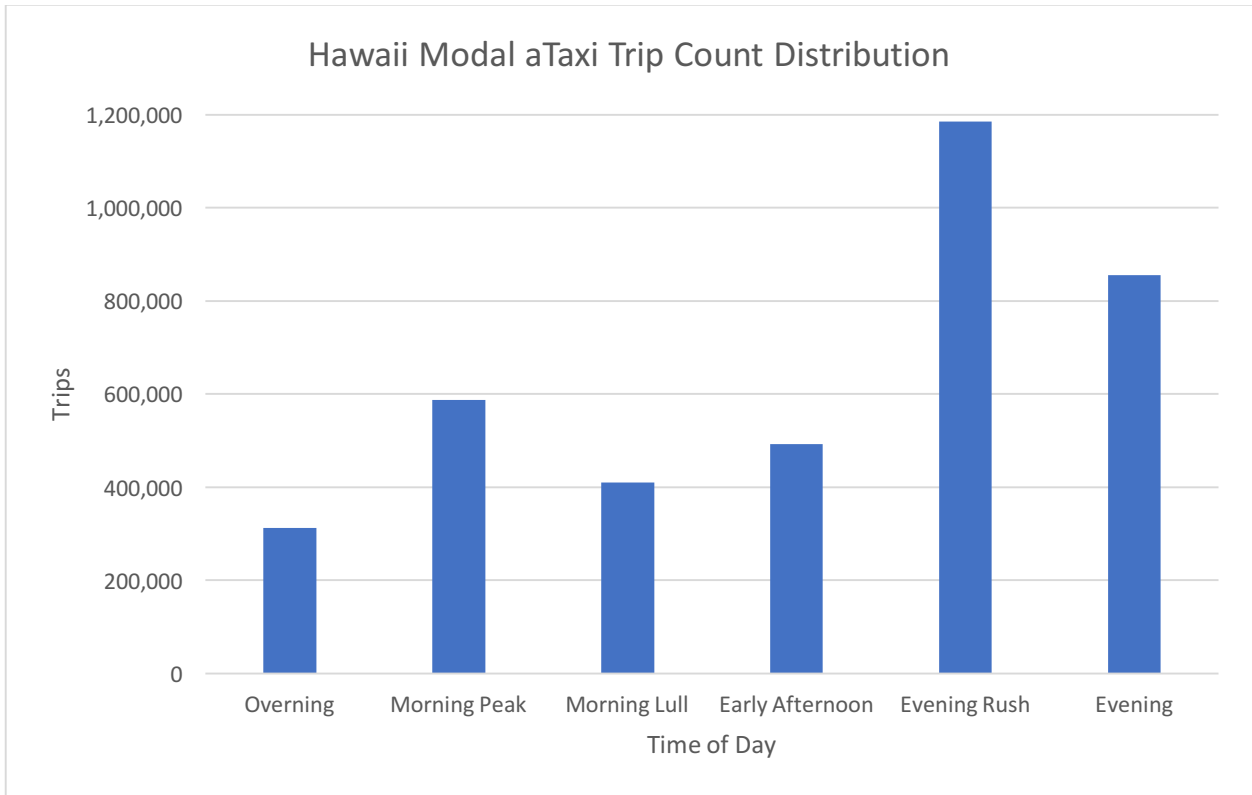


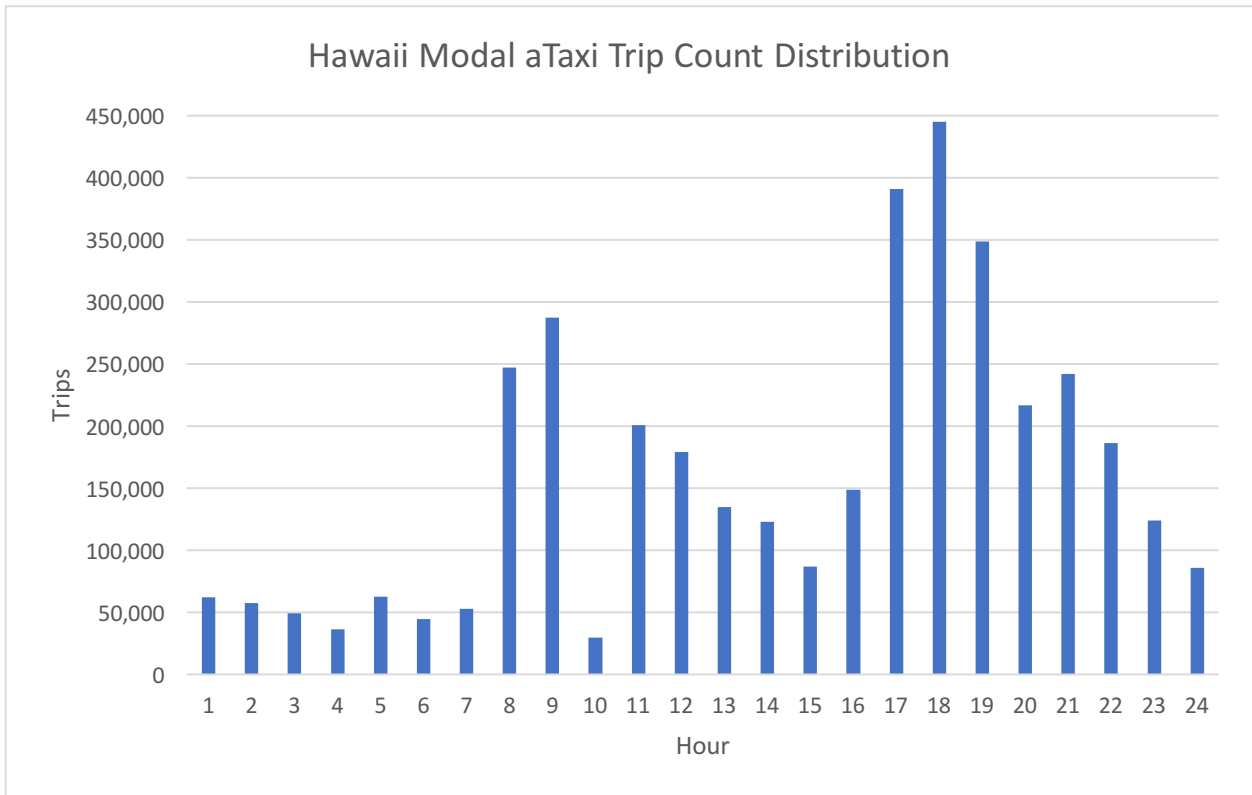
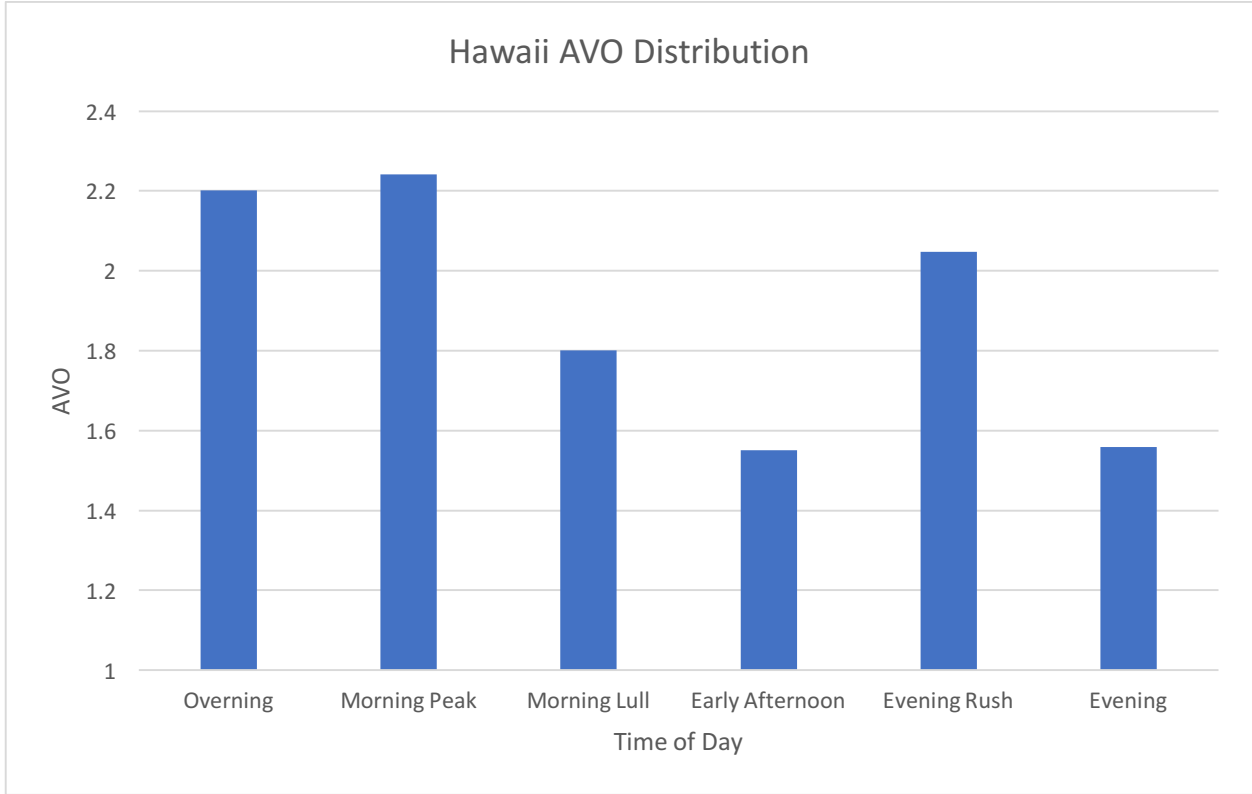


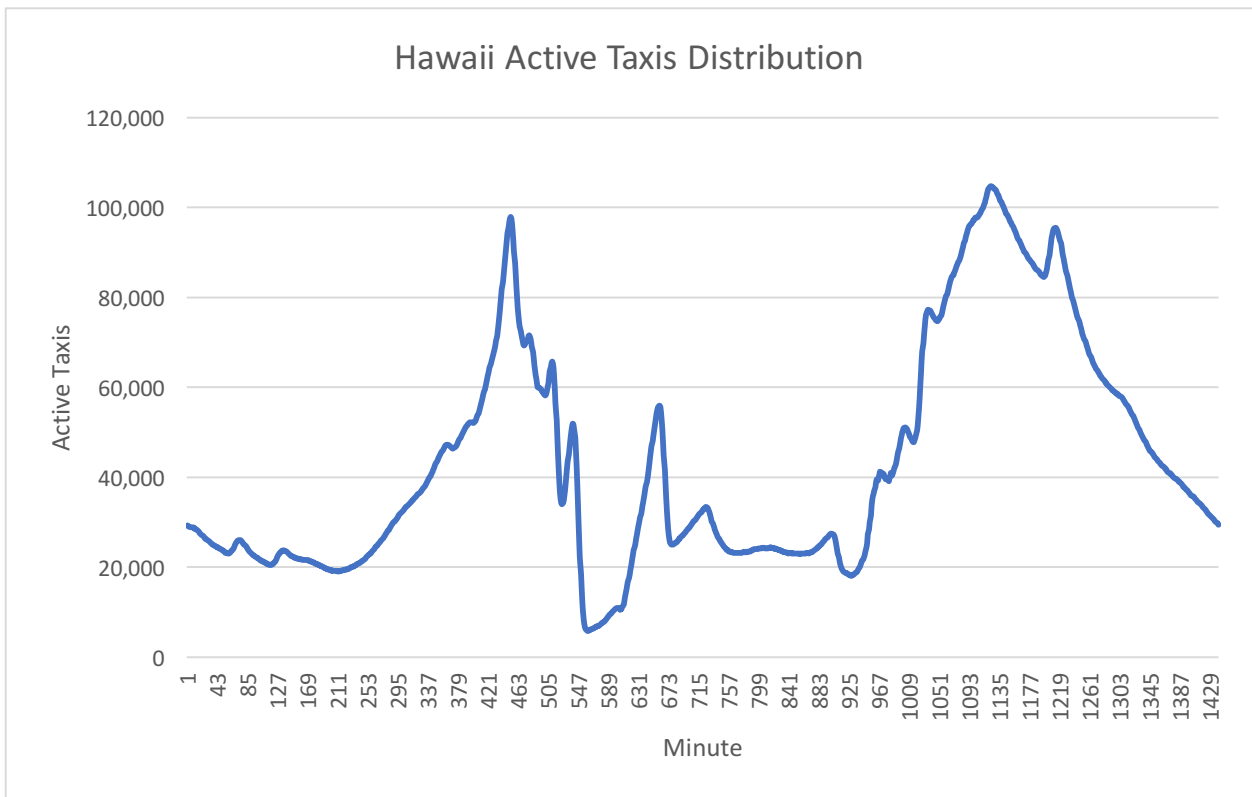
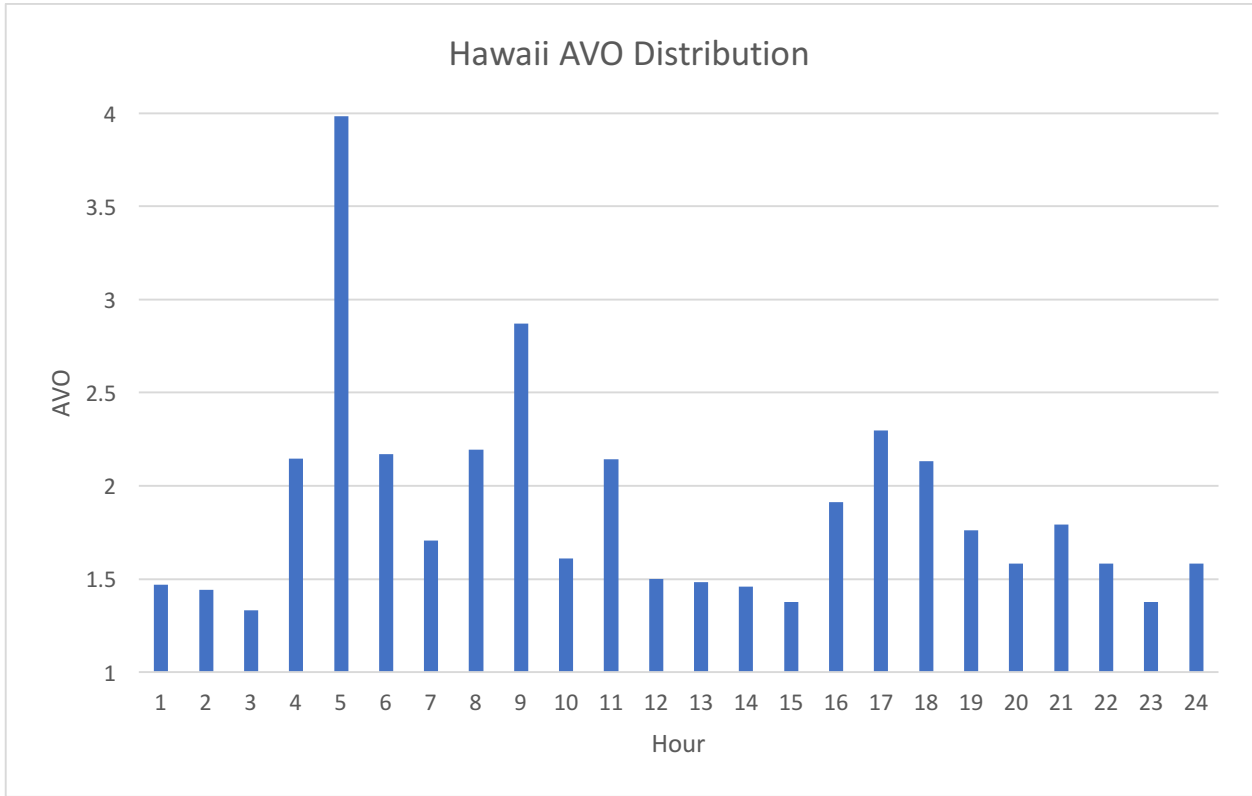


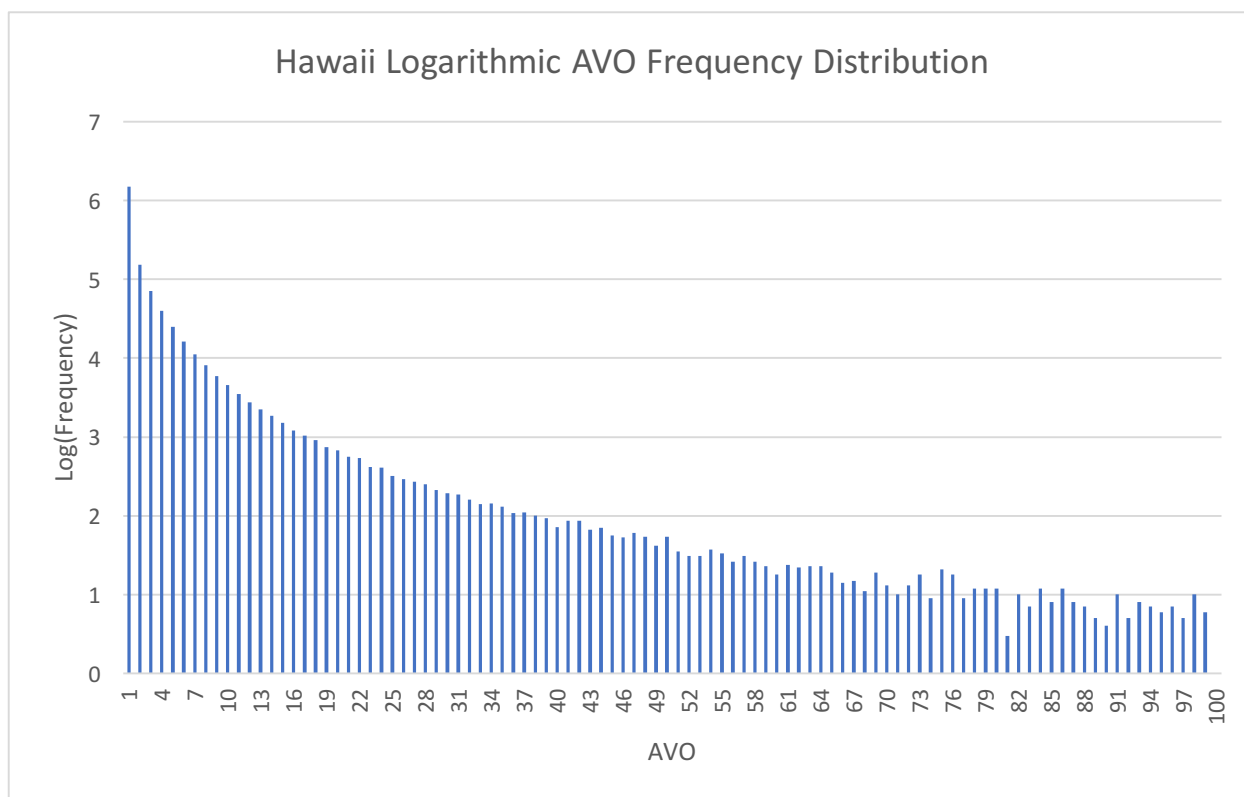
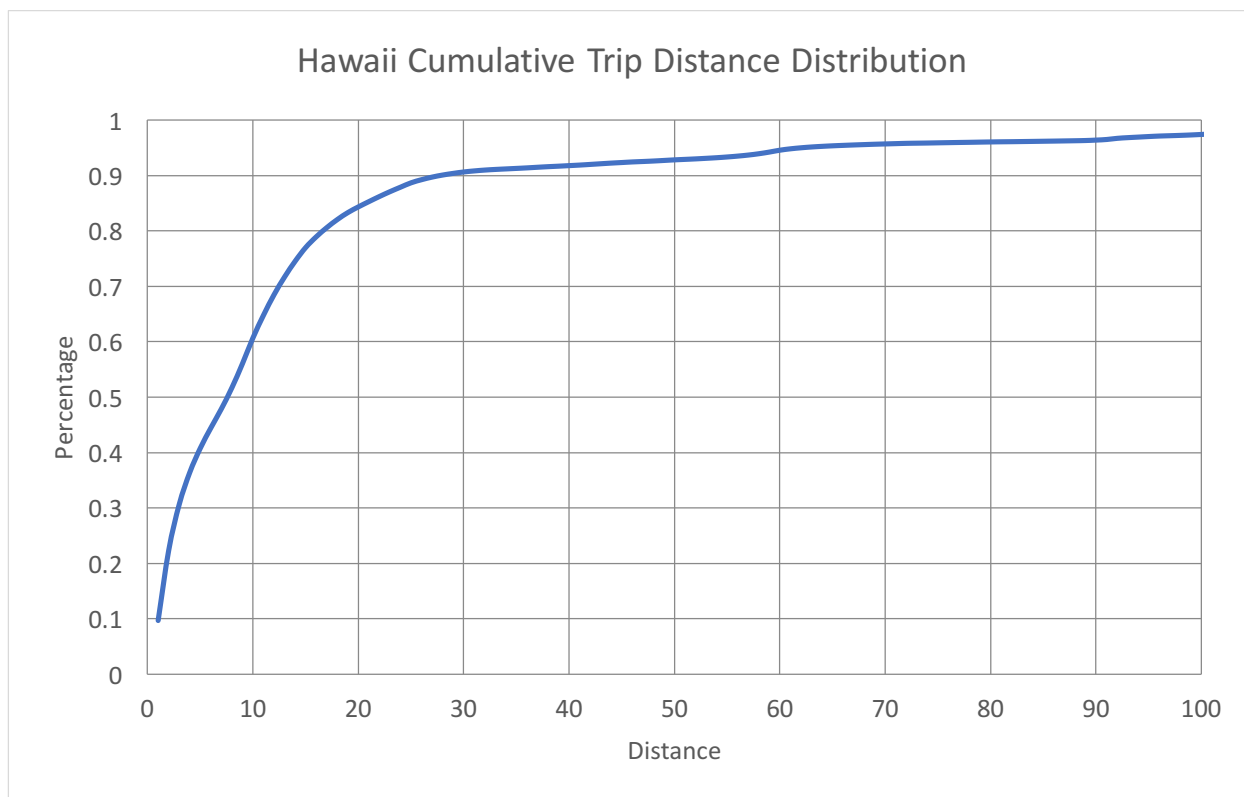


Hawaii:

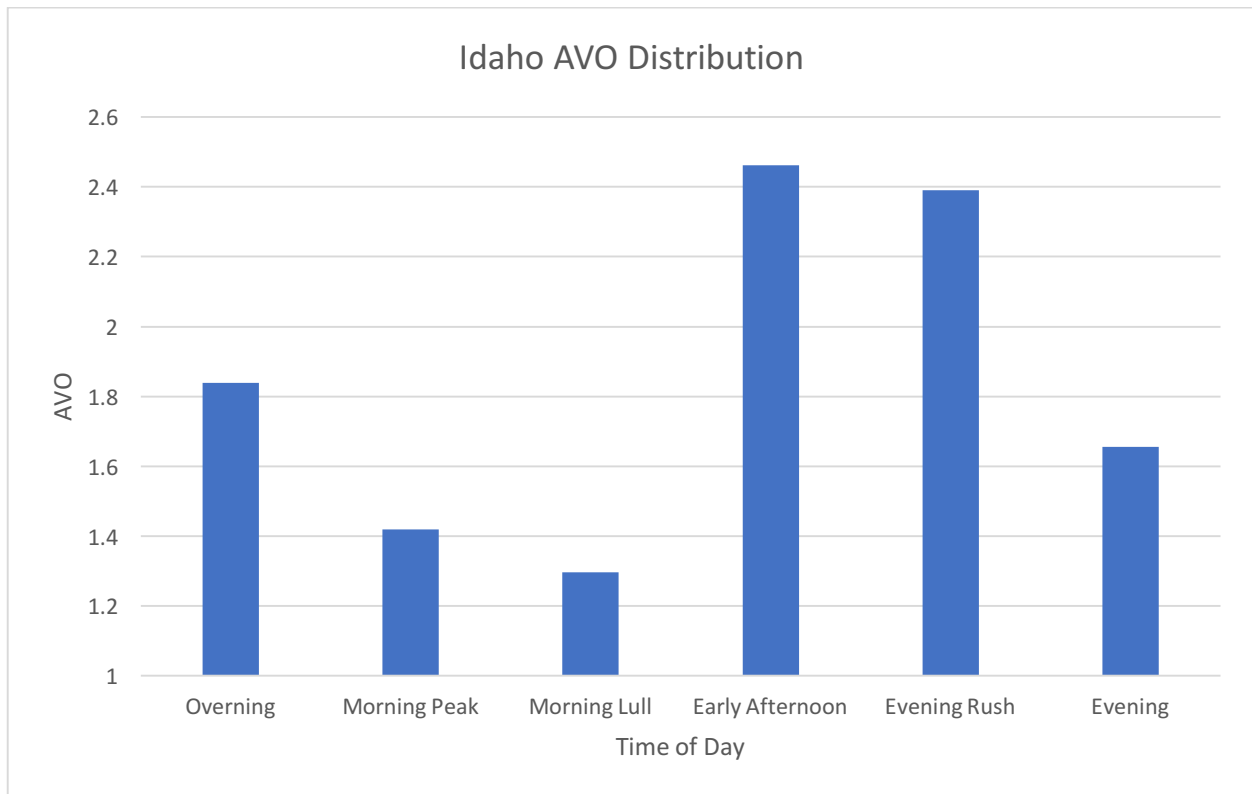
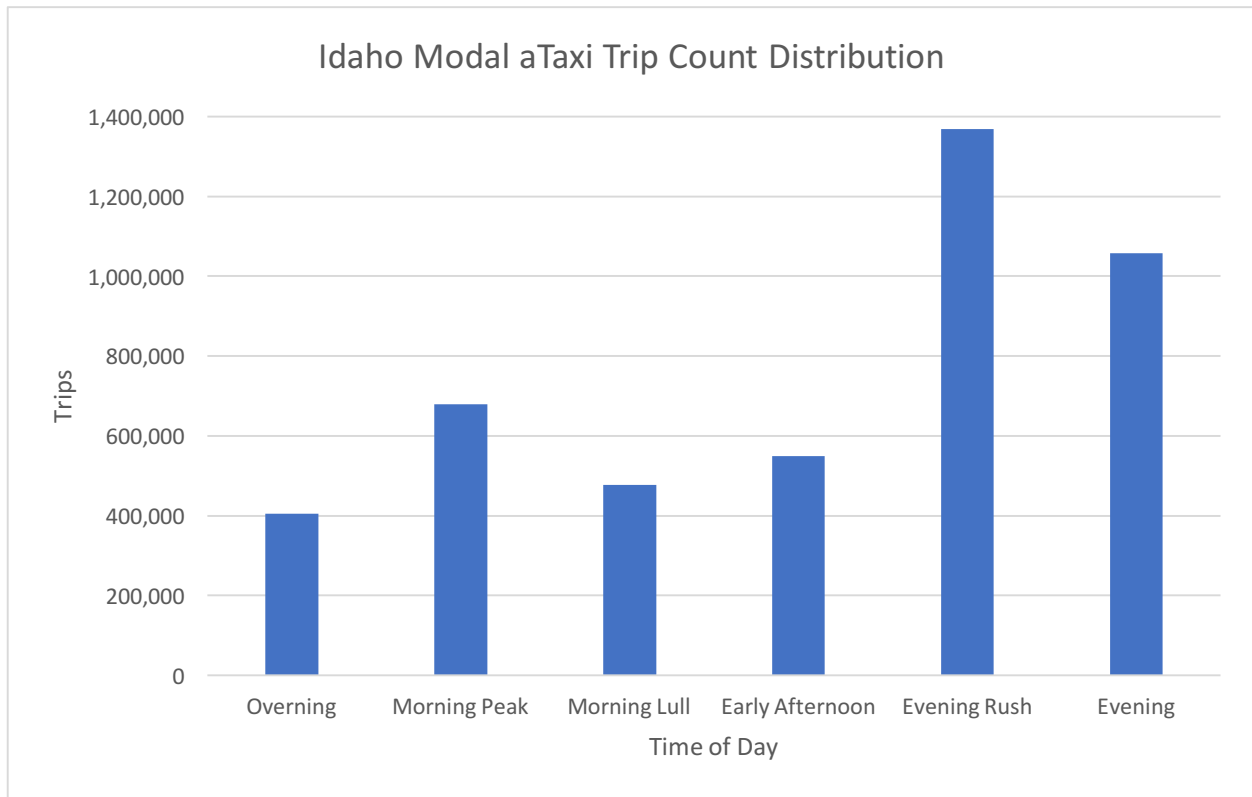


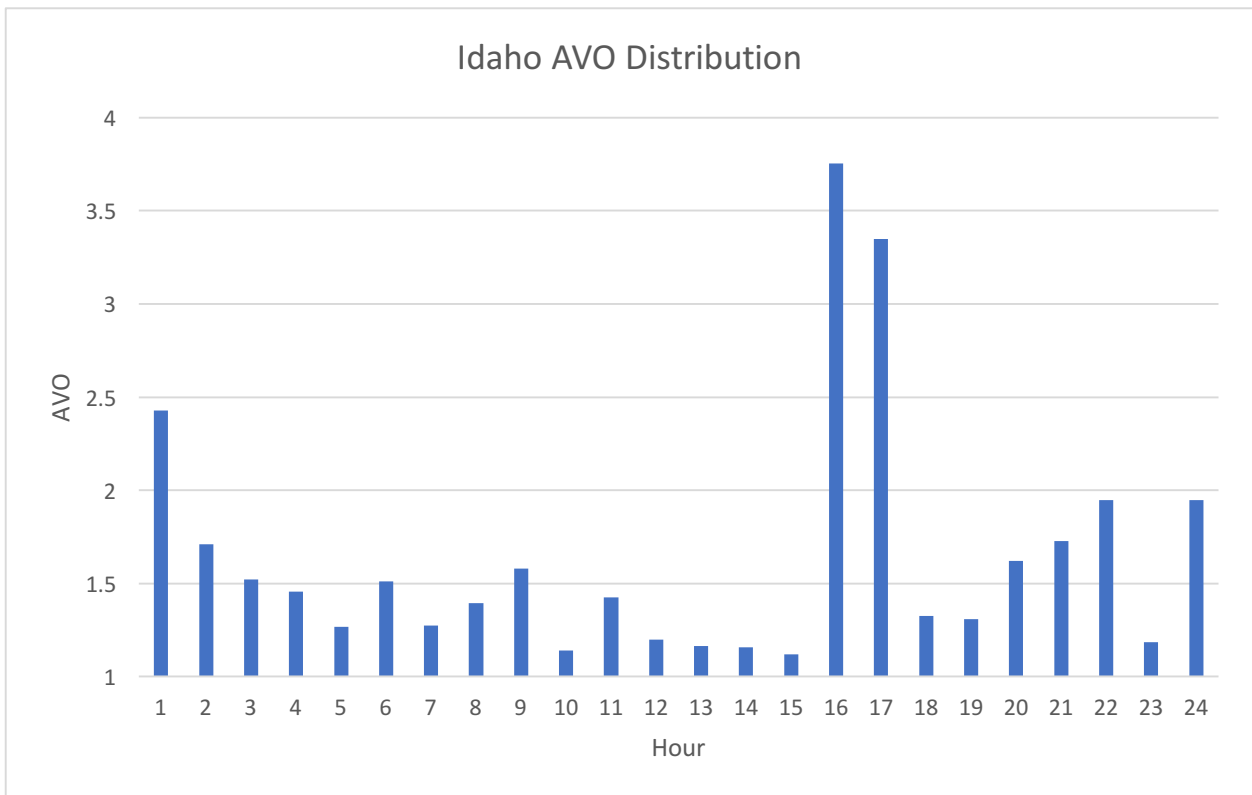
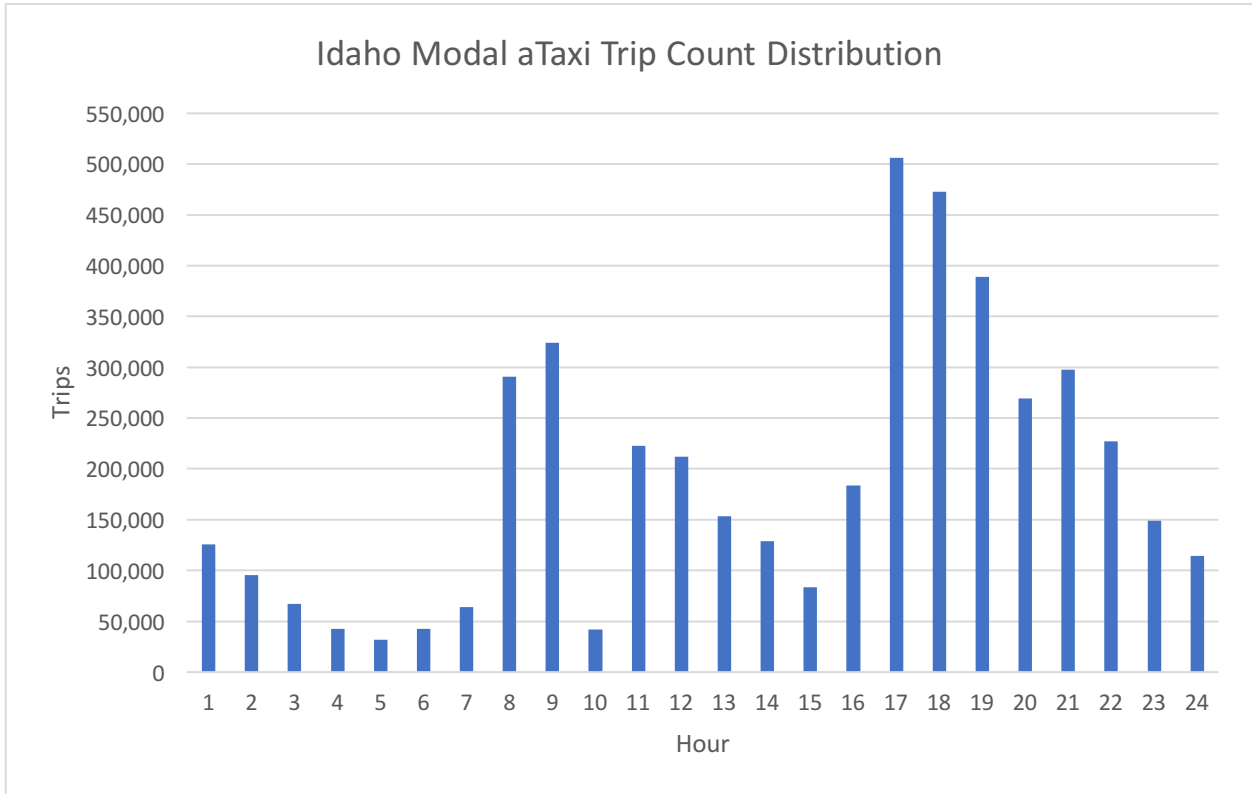


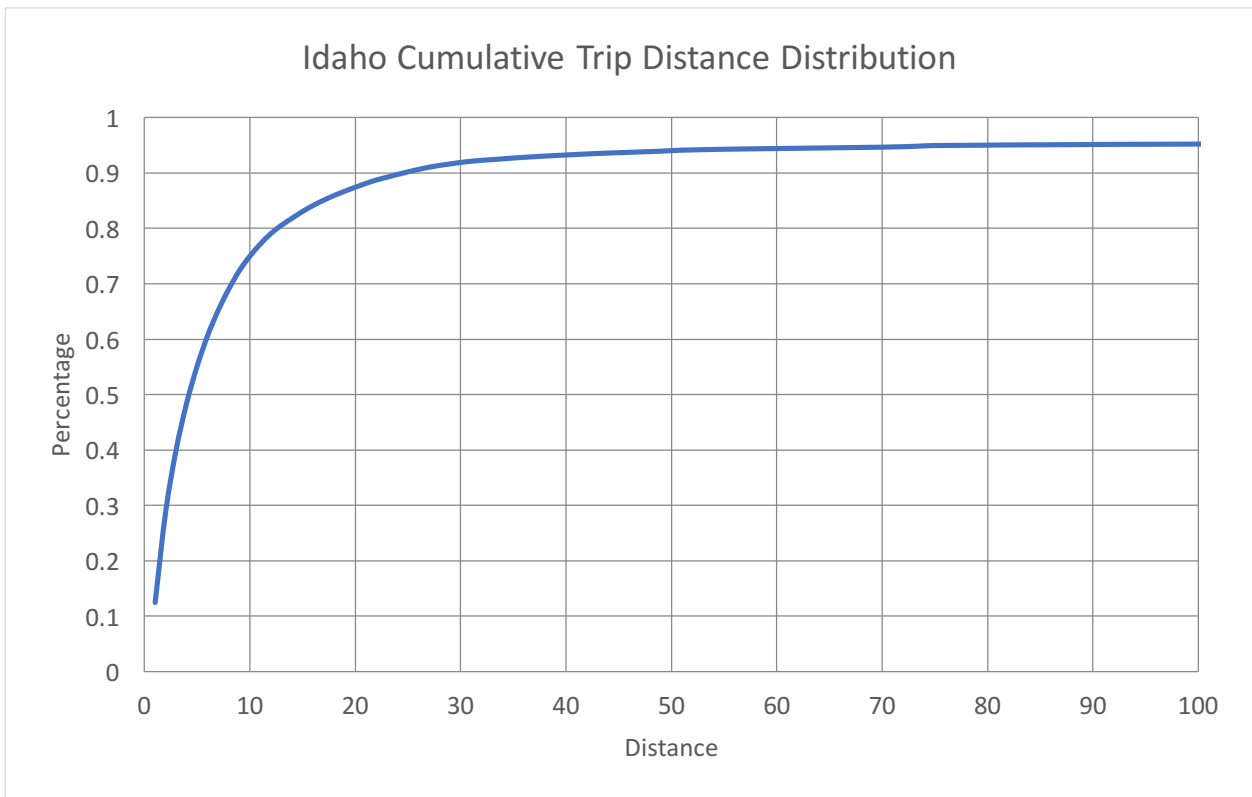
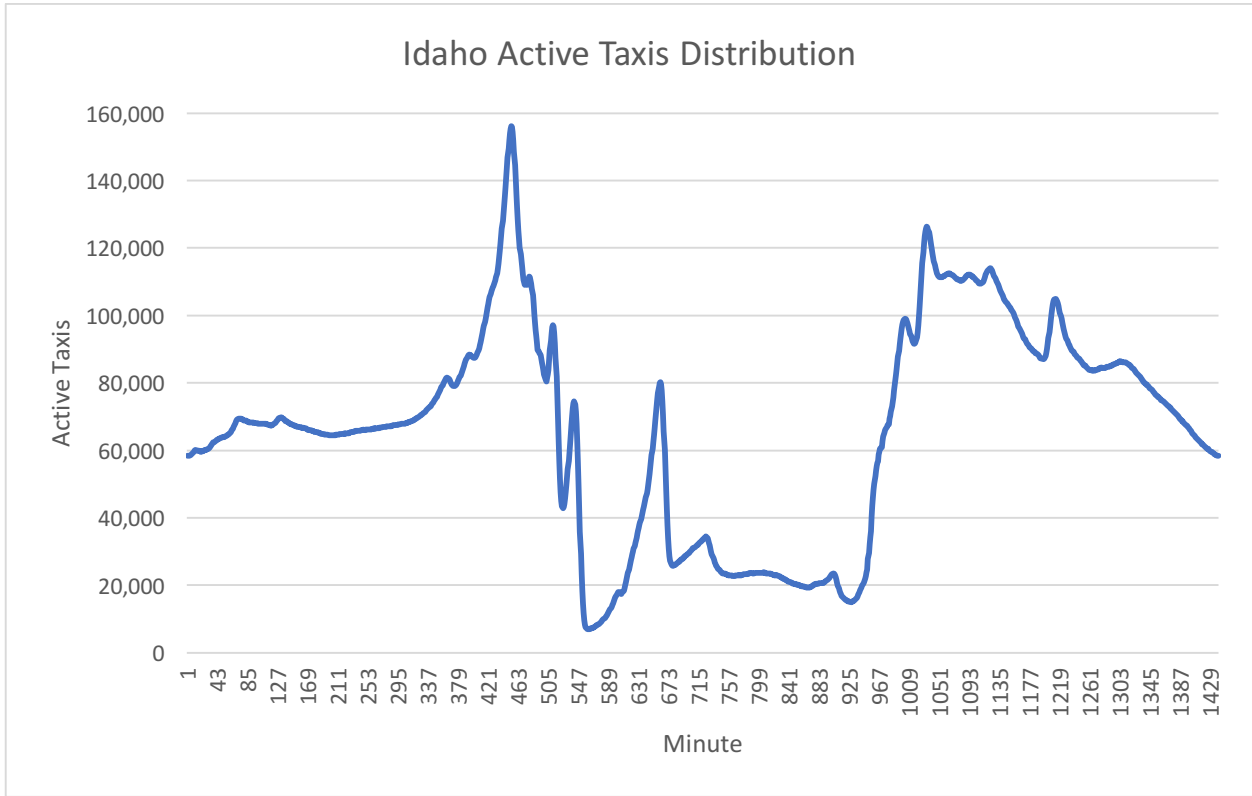


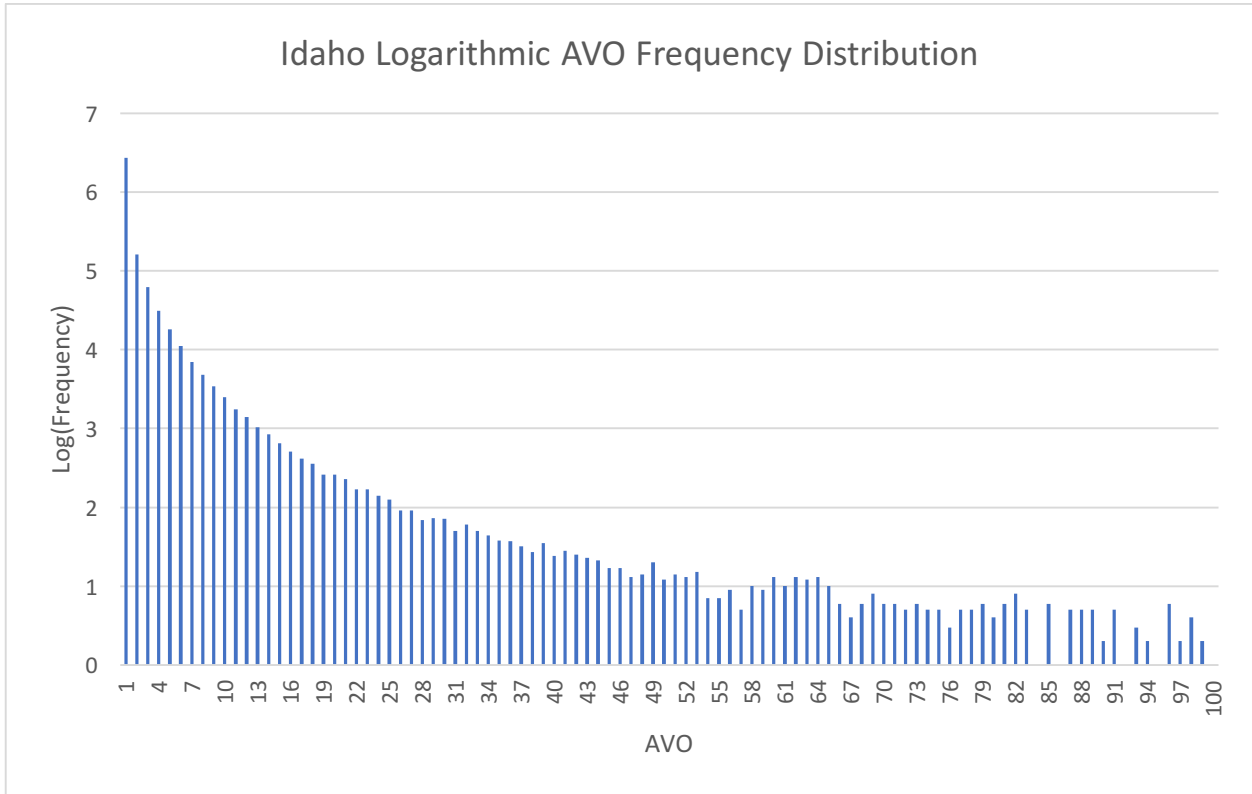


Idaho:

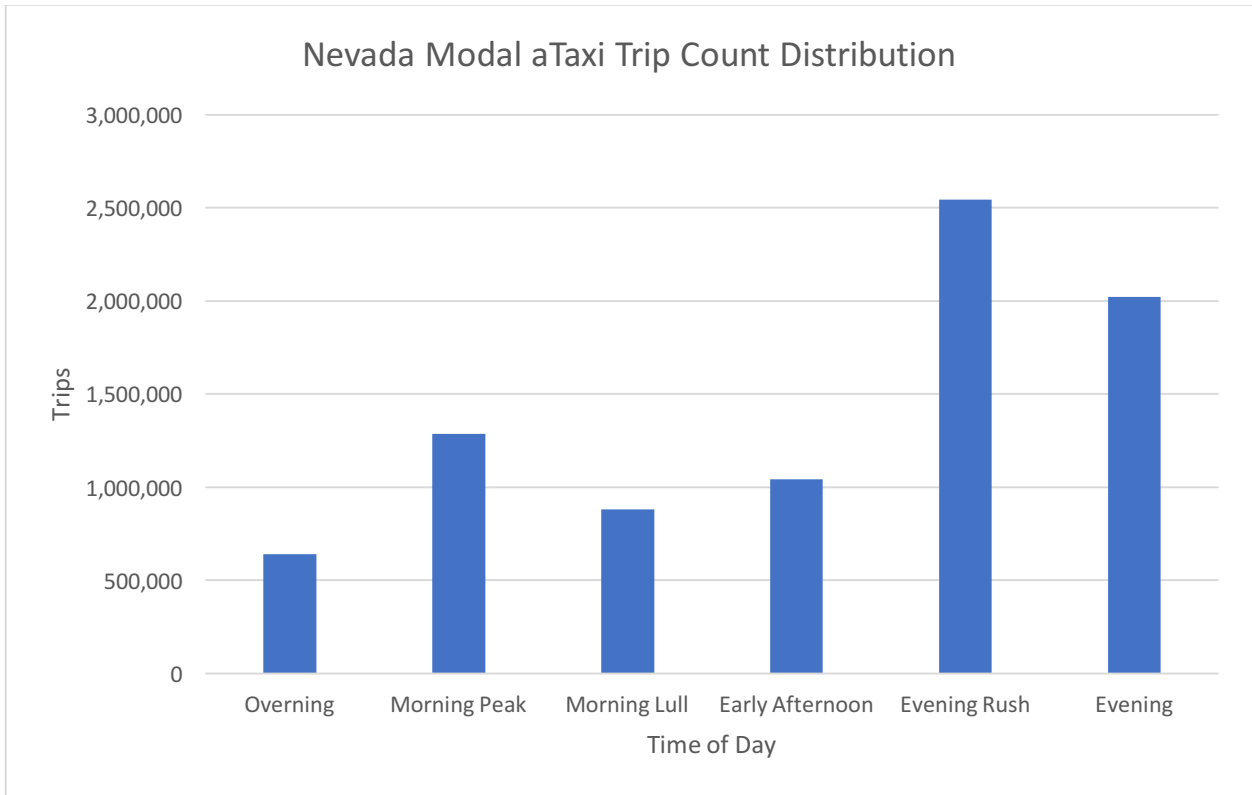


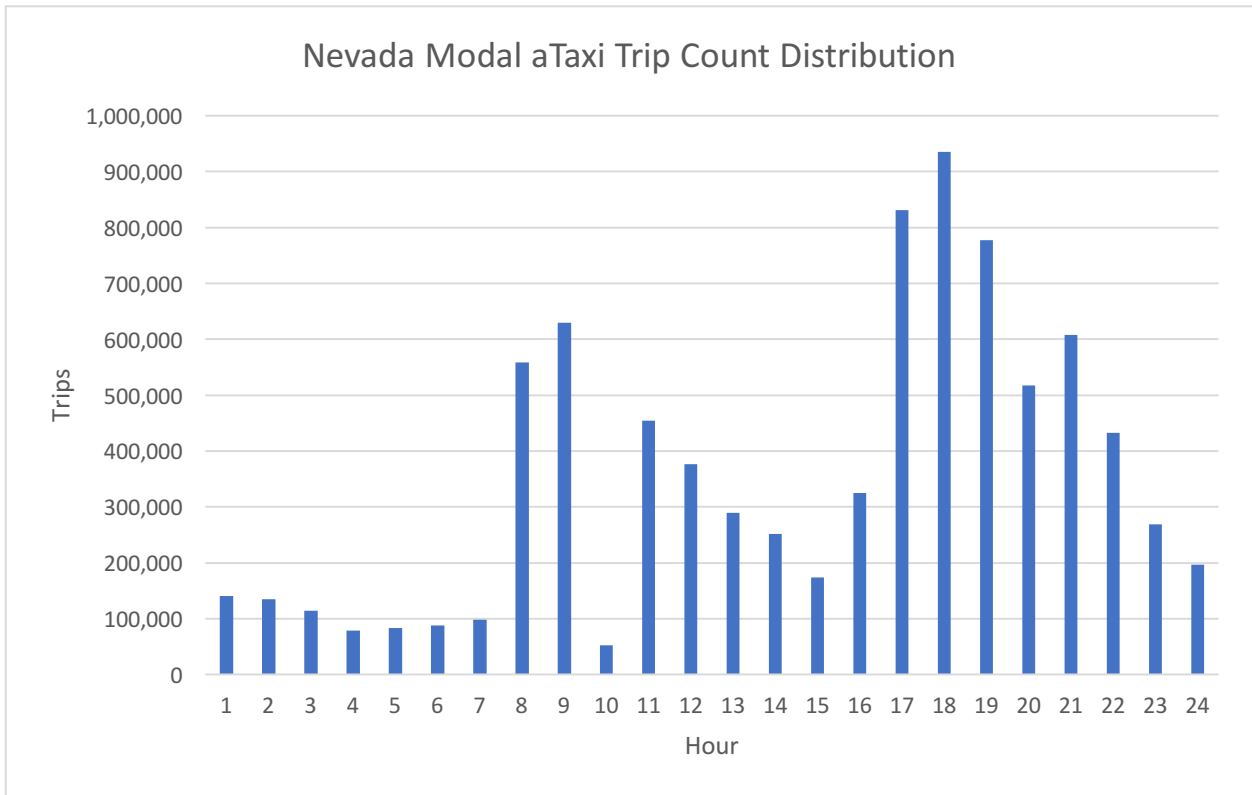
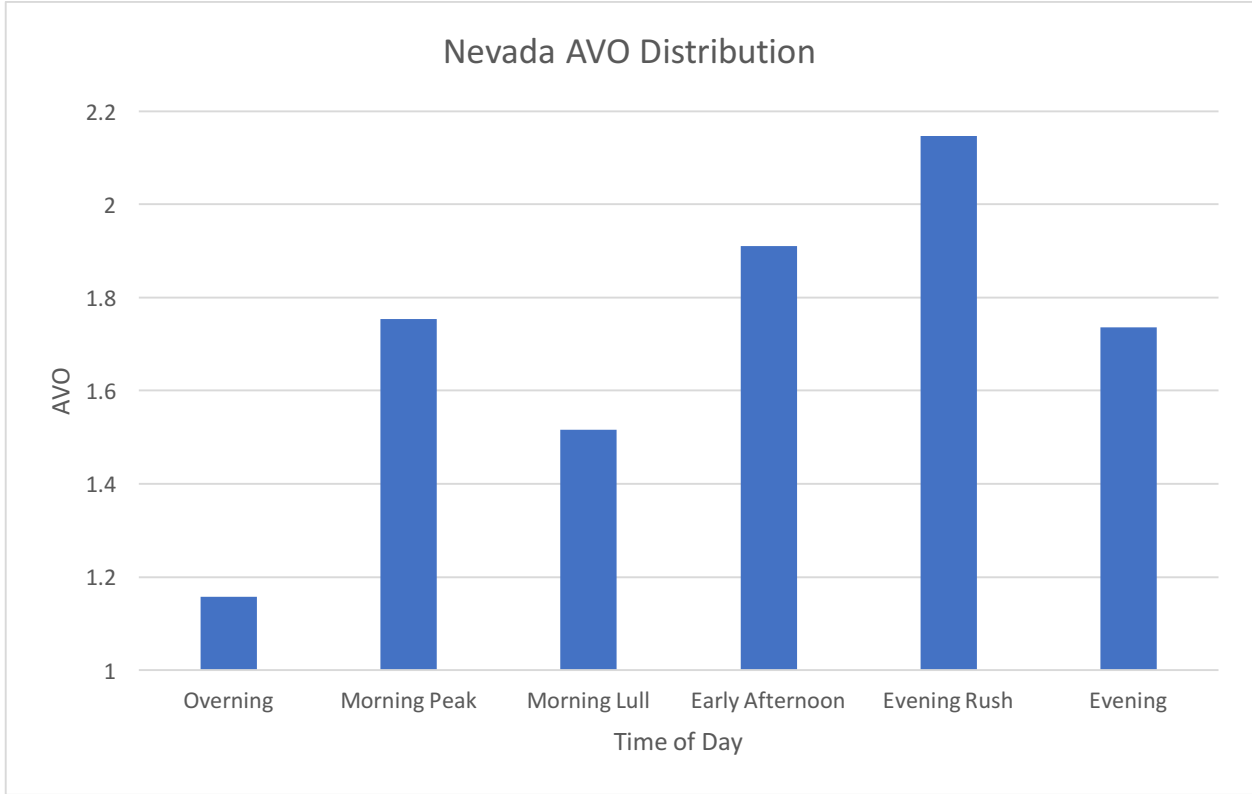


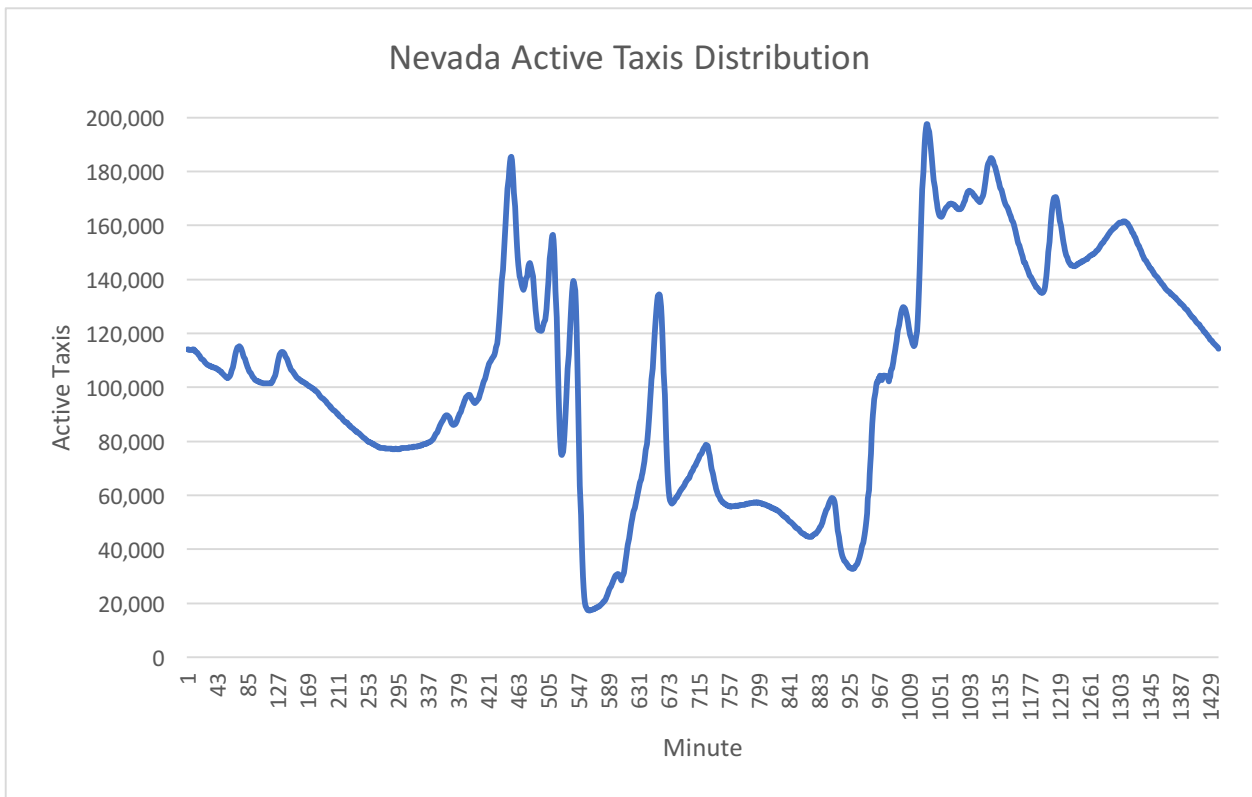
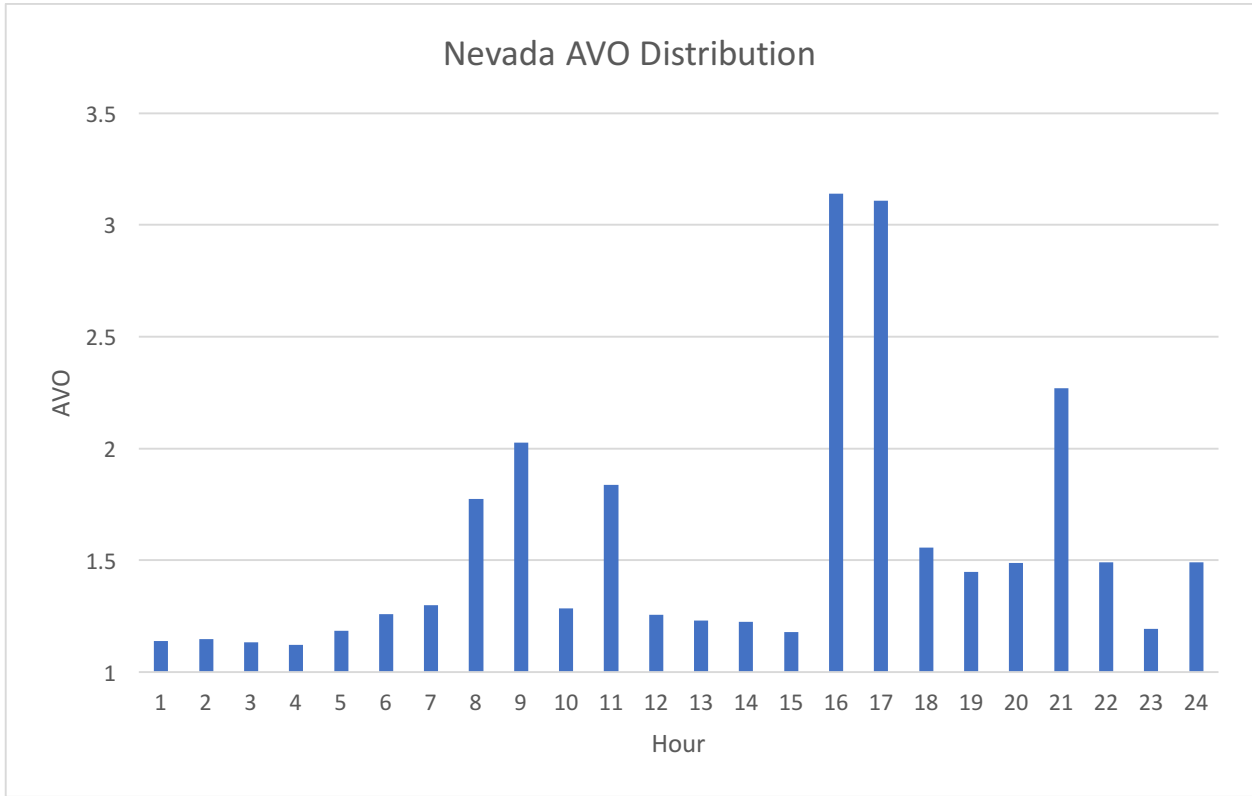


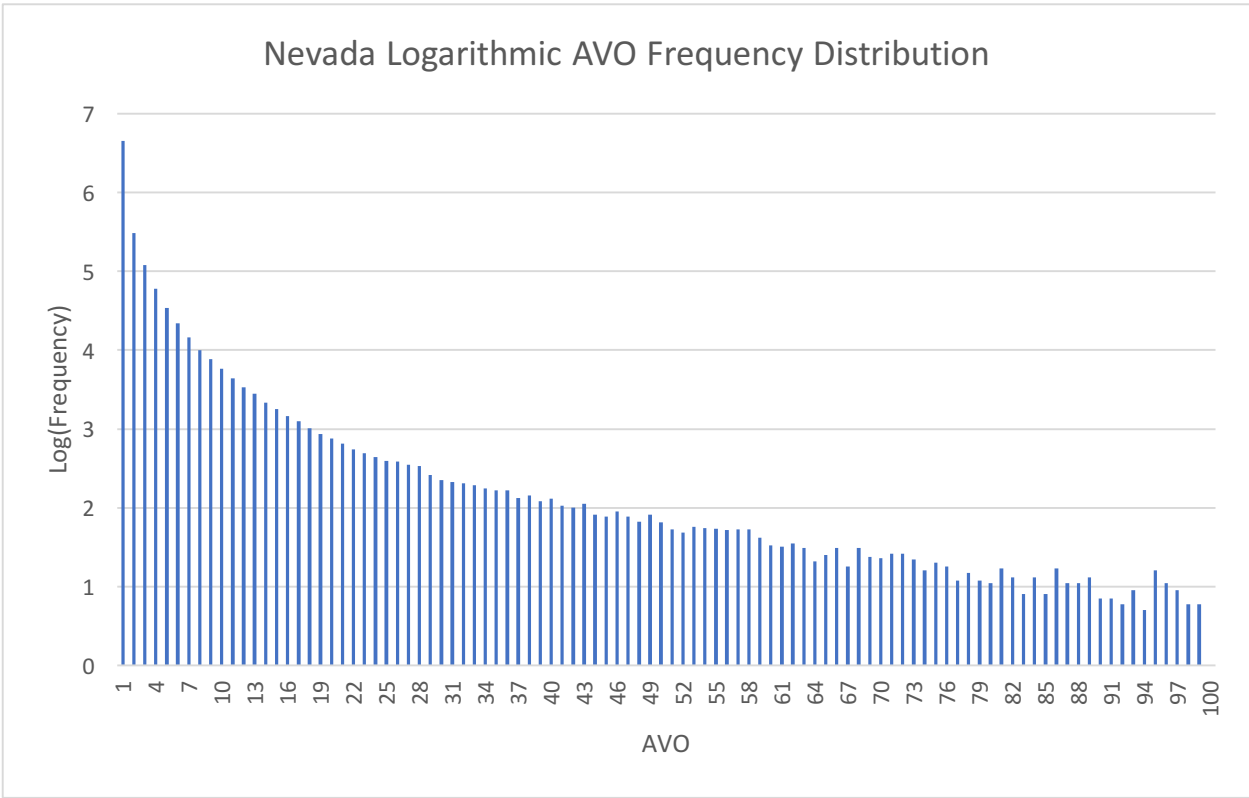
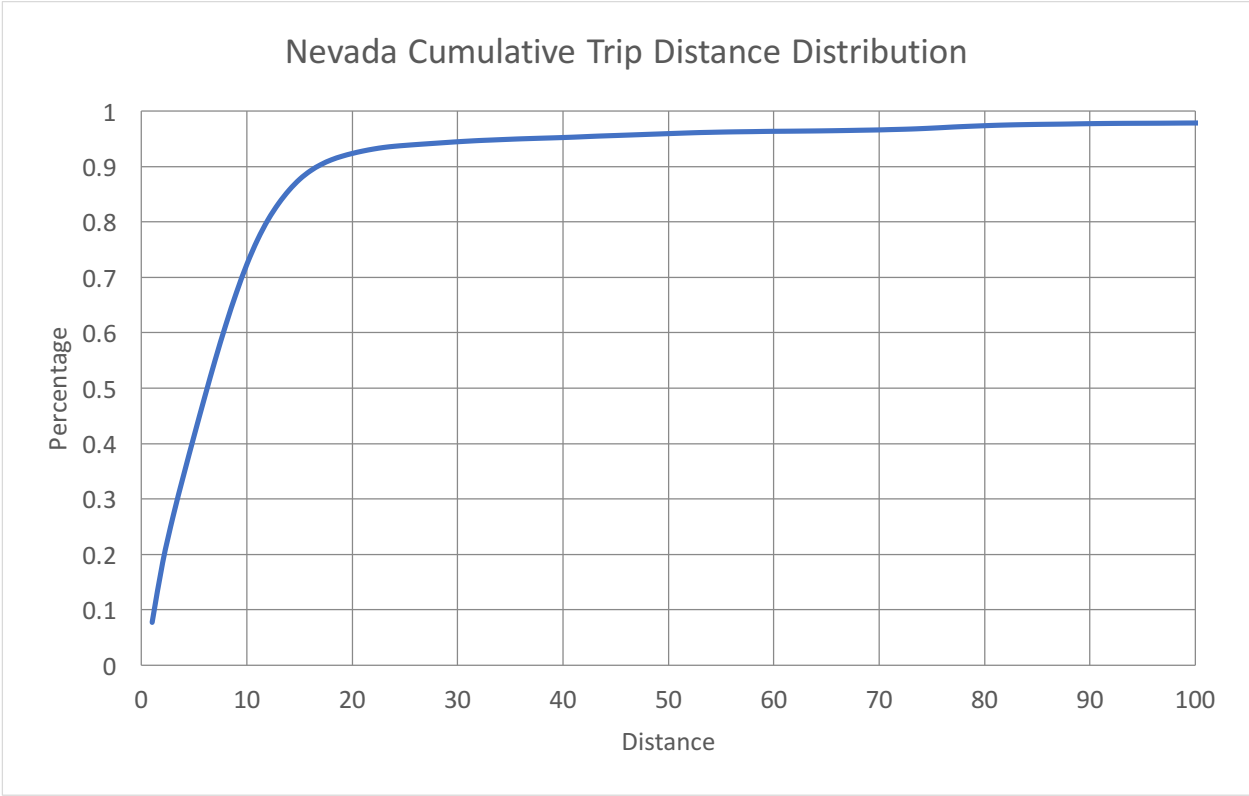


Nevada:

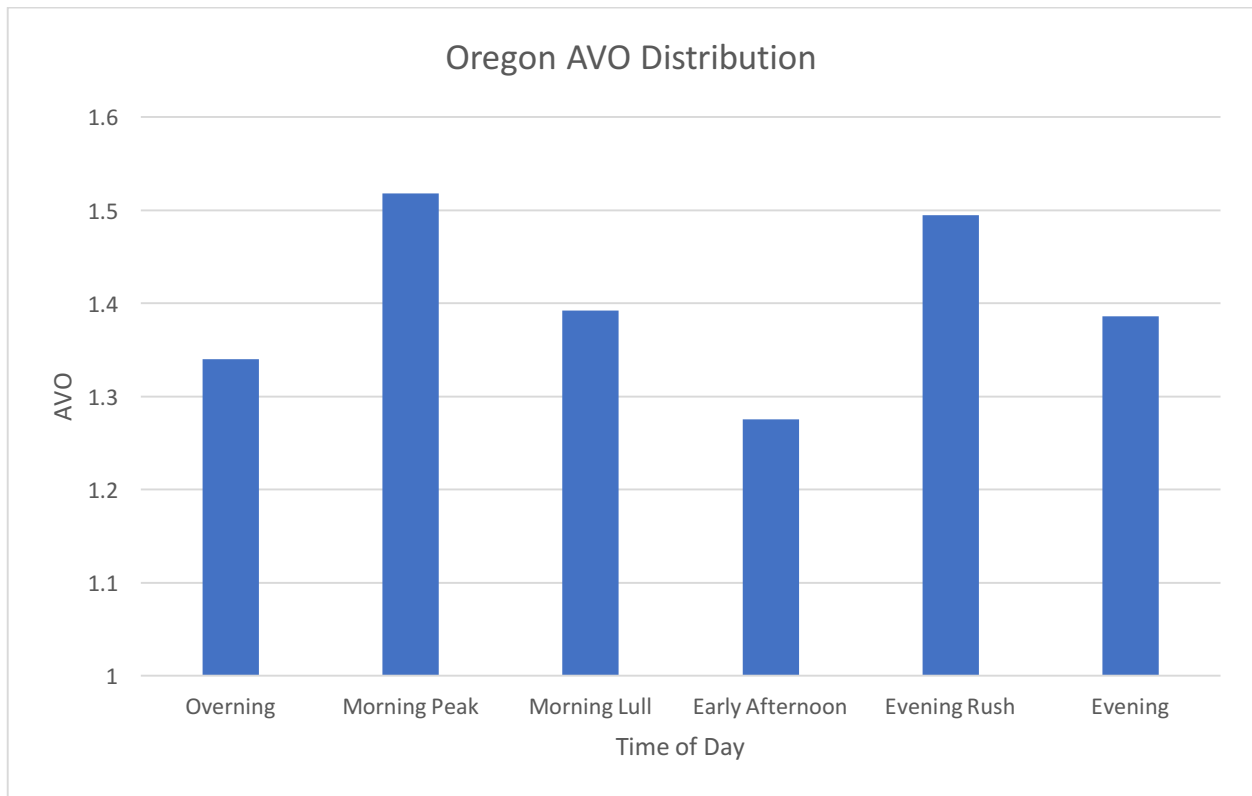
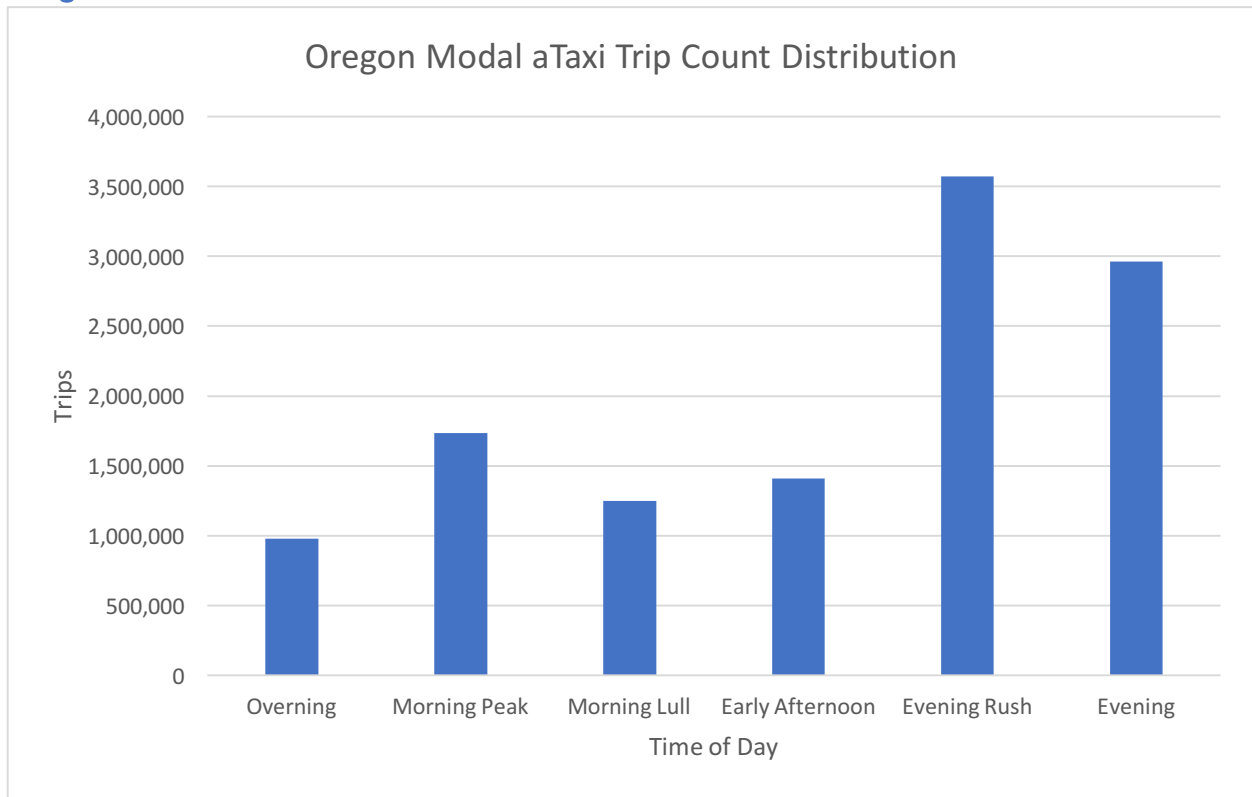


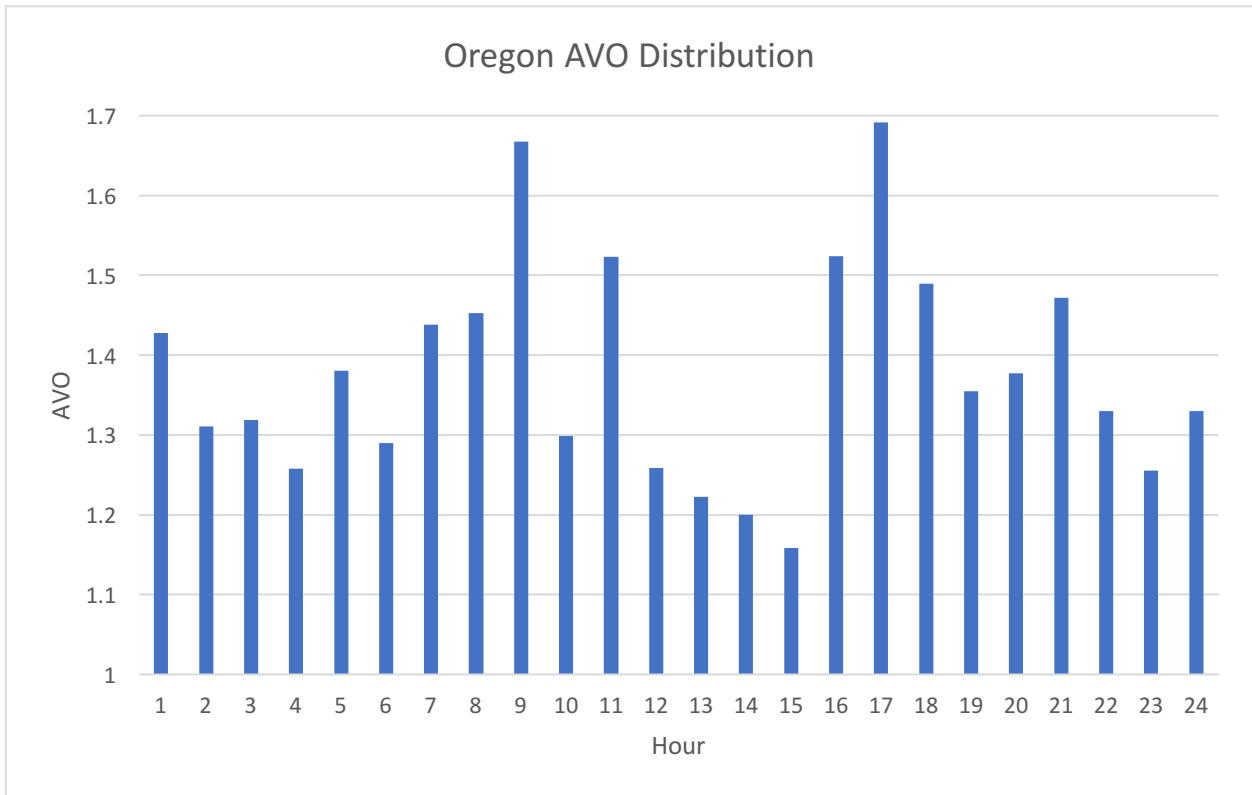
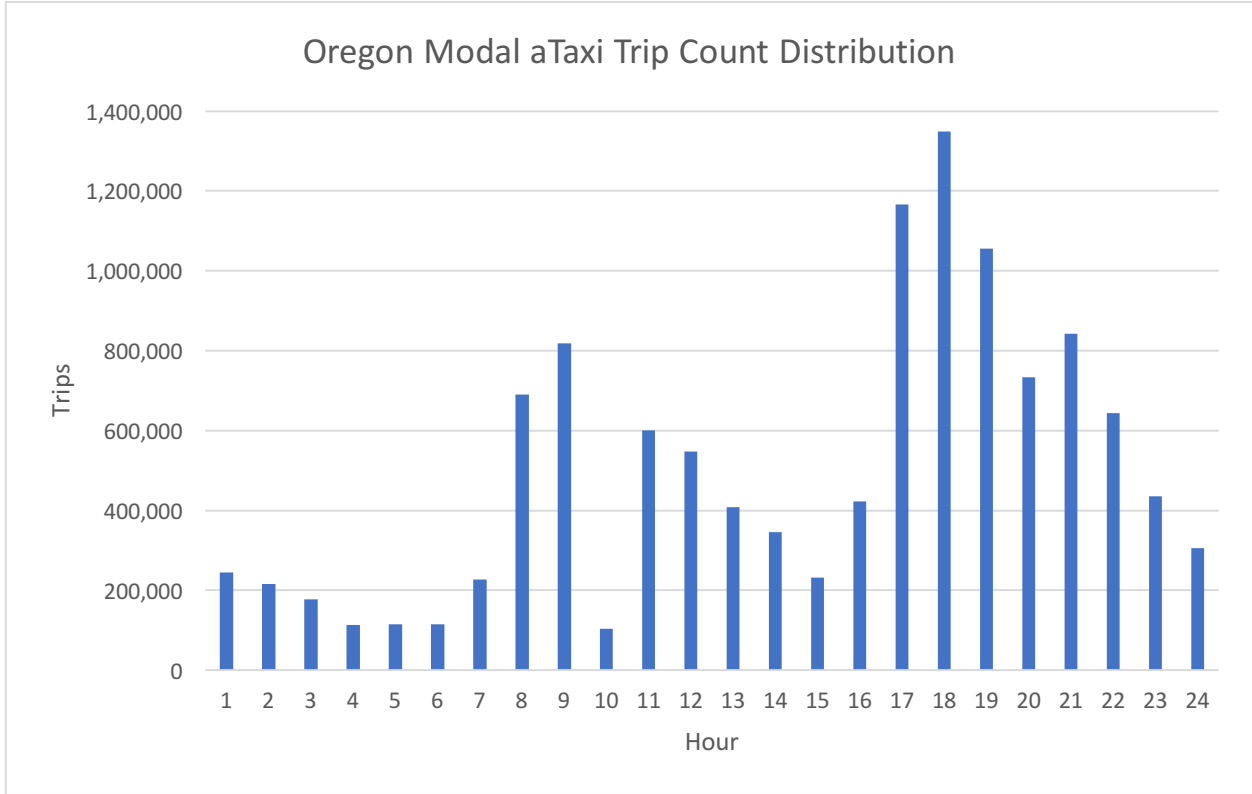


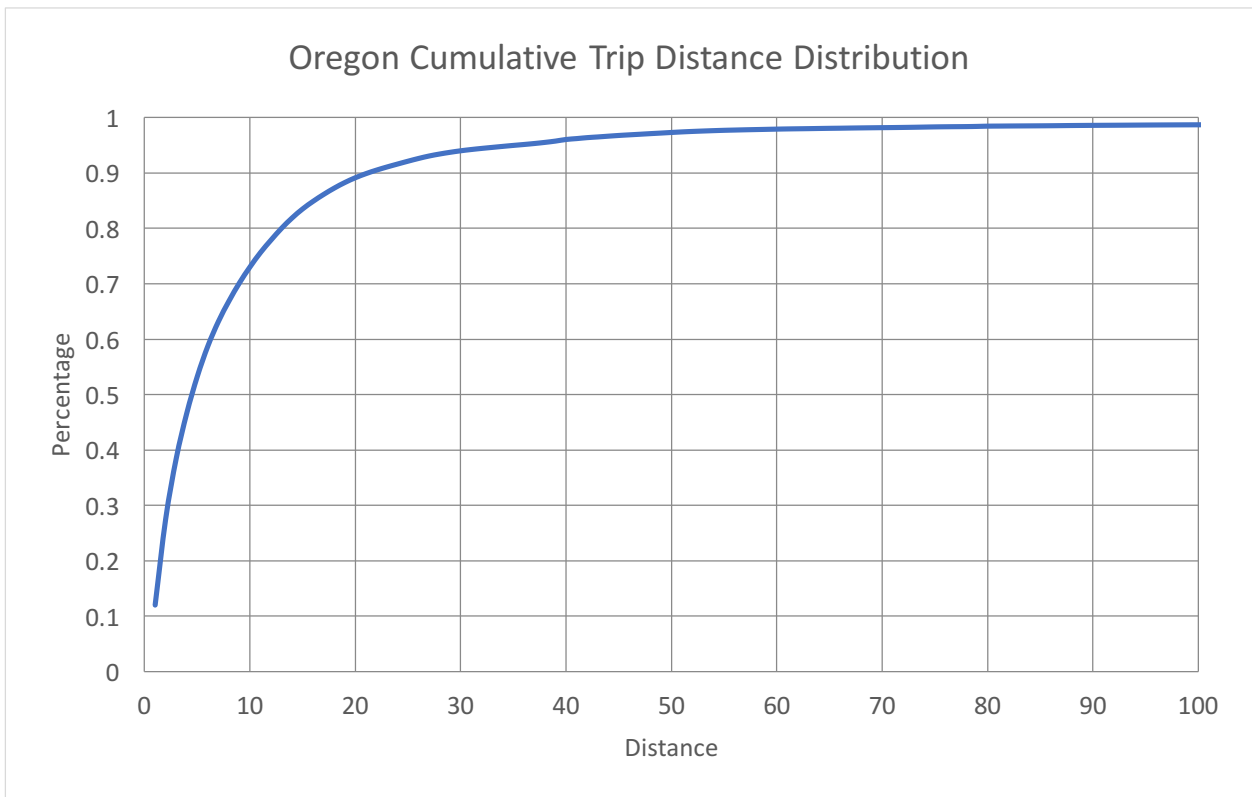
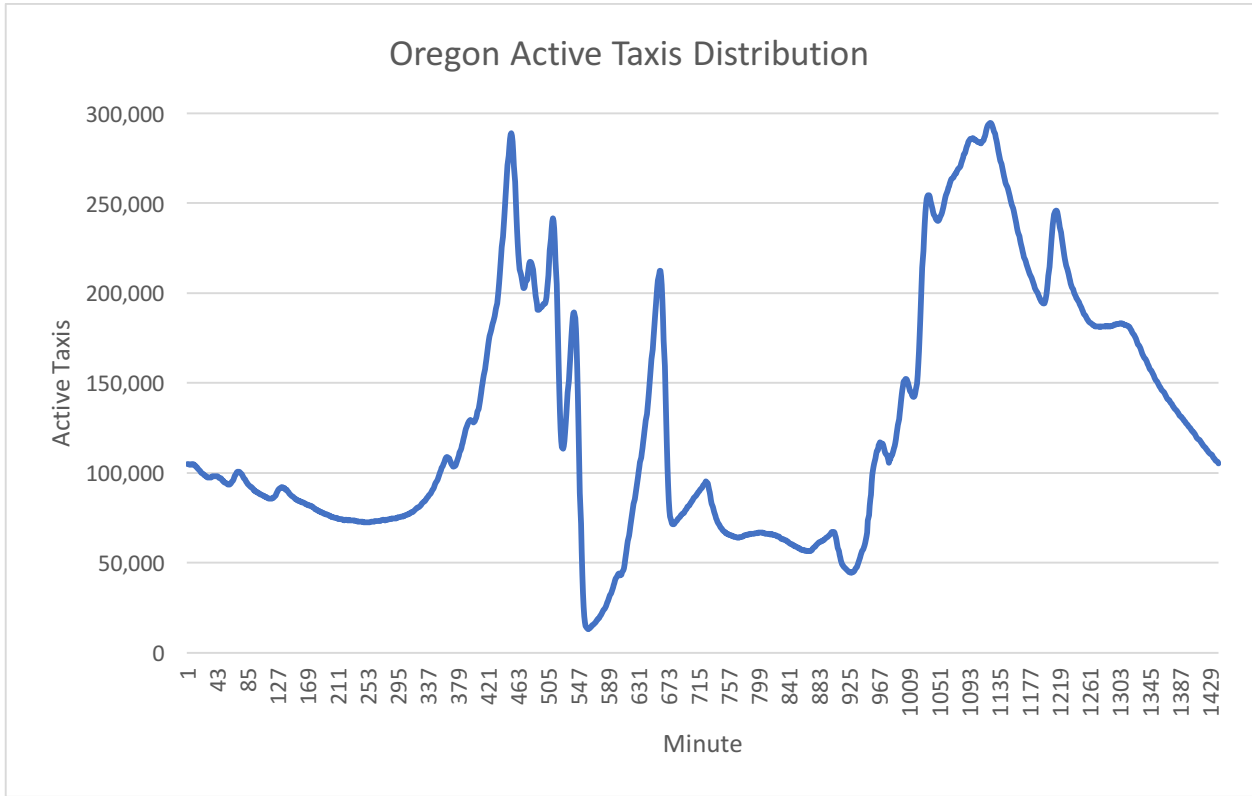


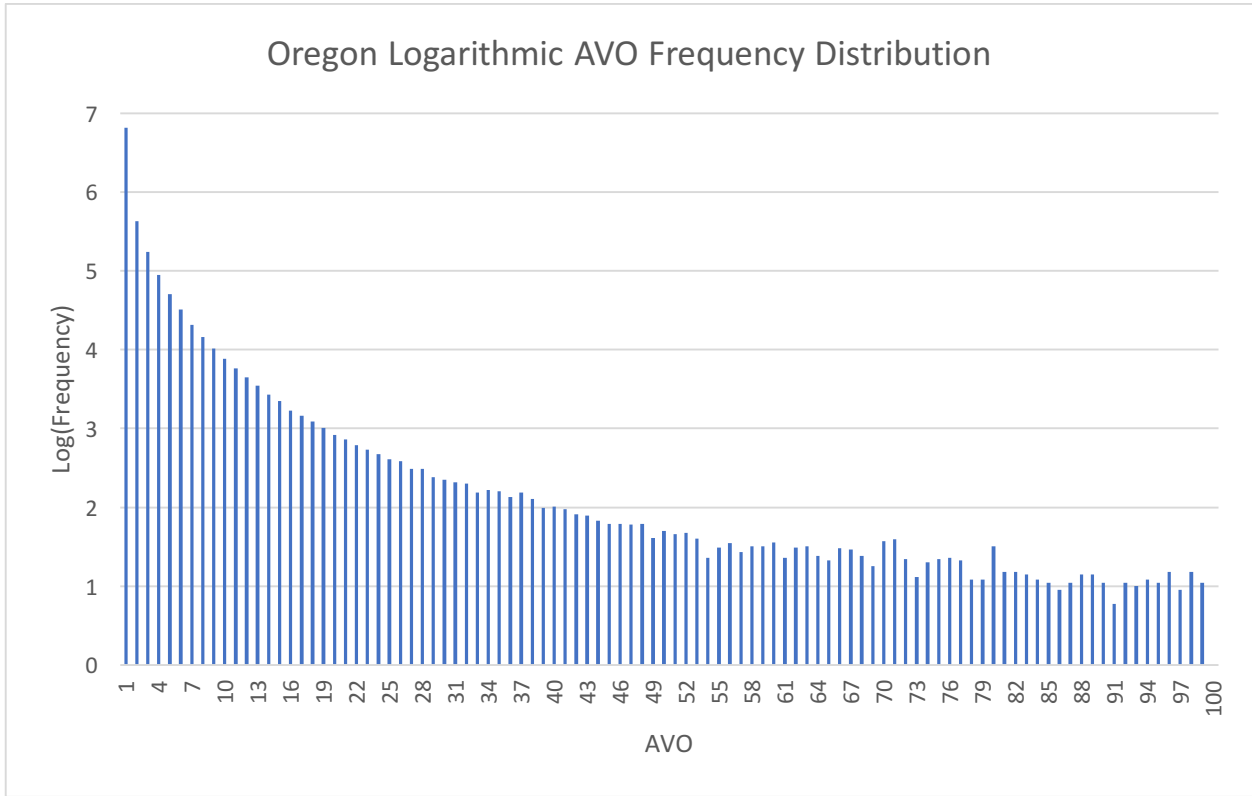


Oregon:

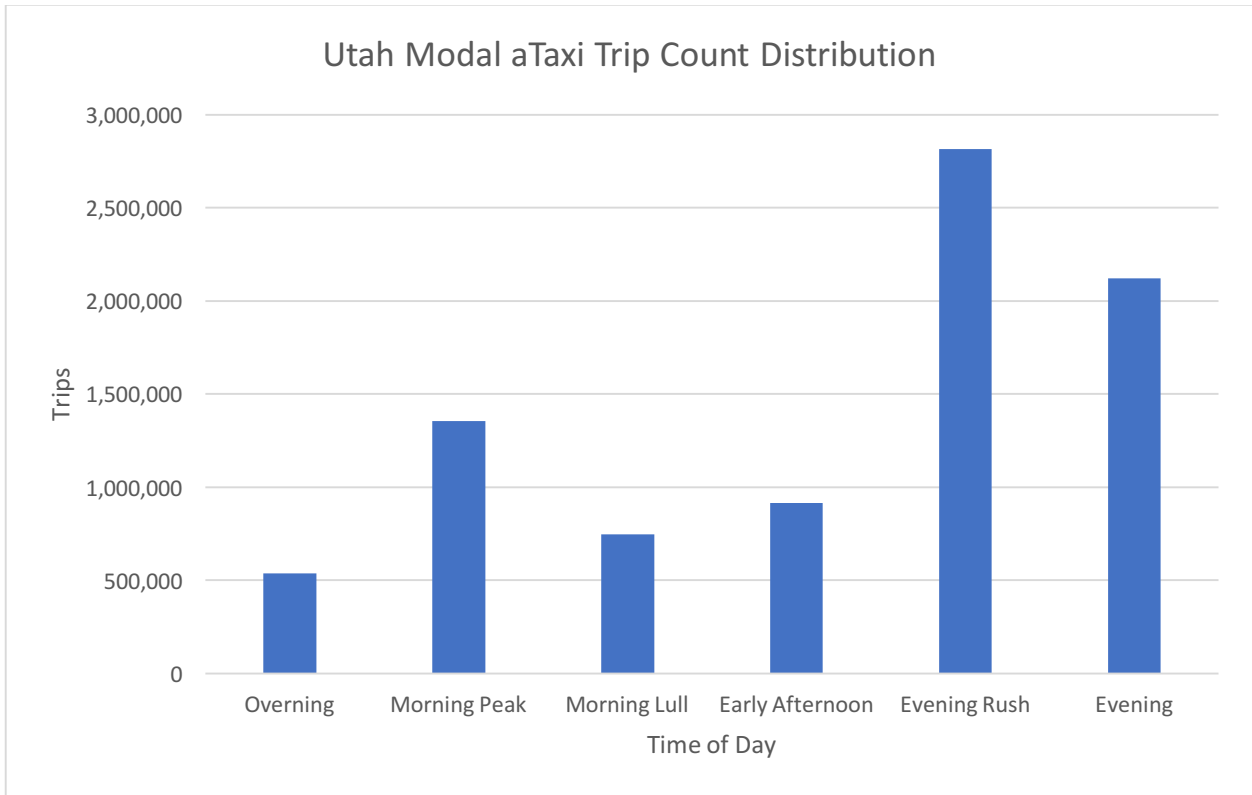


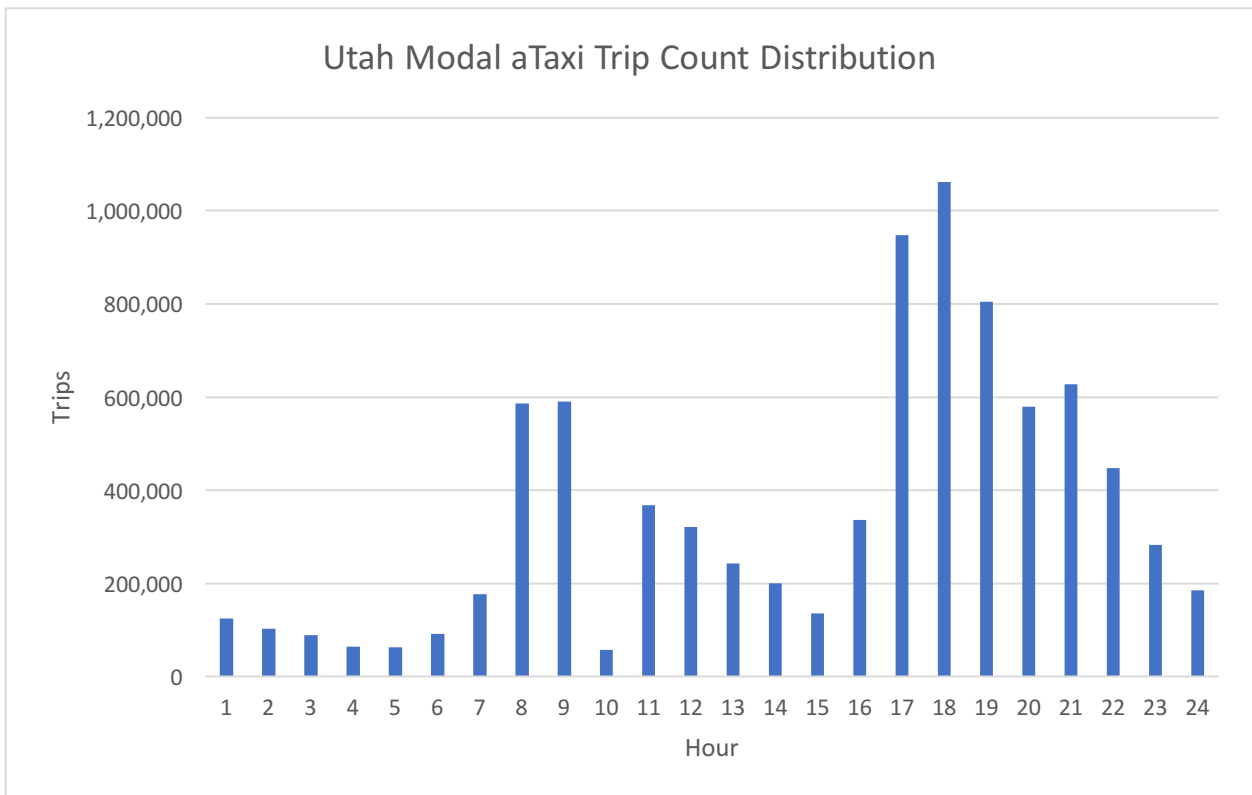
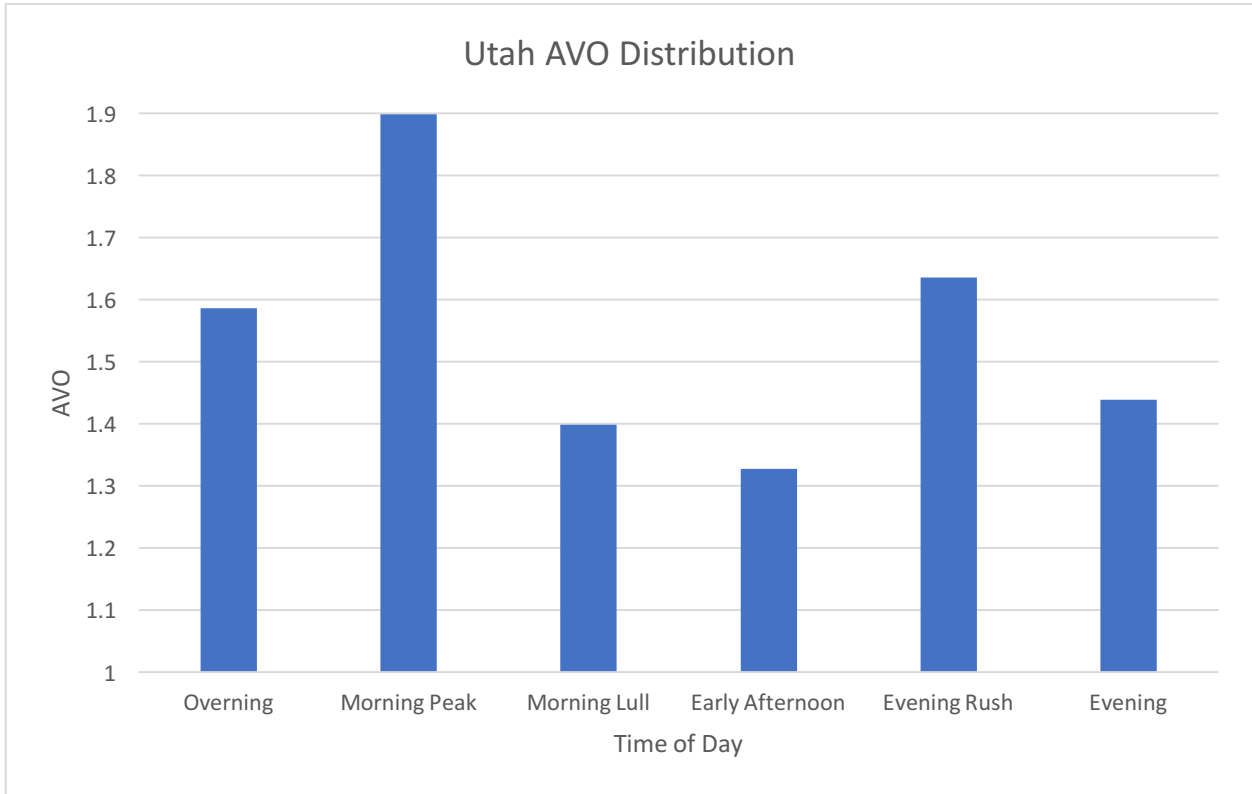


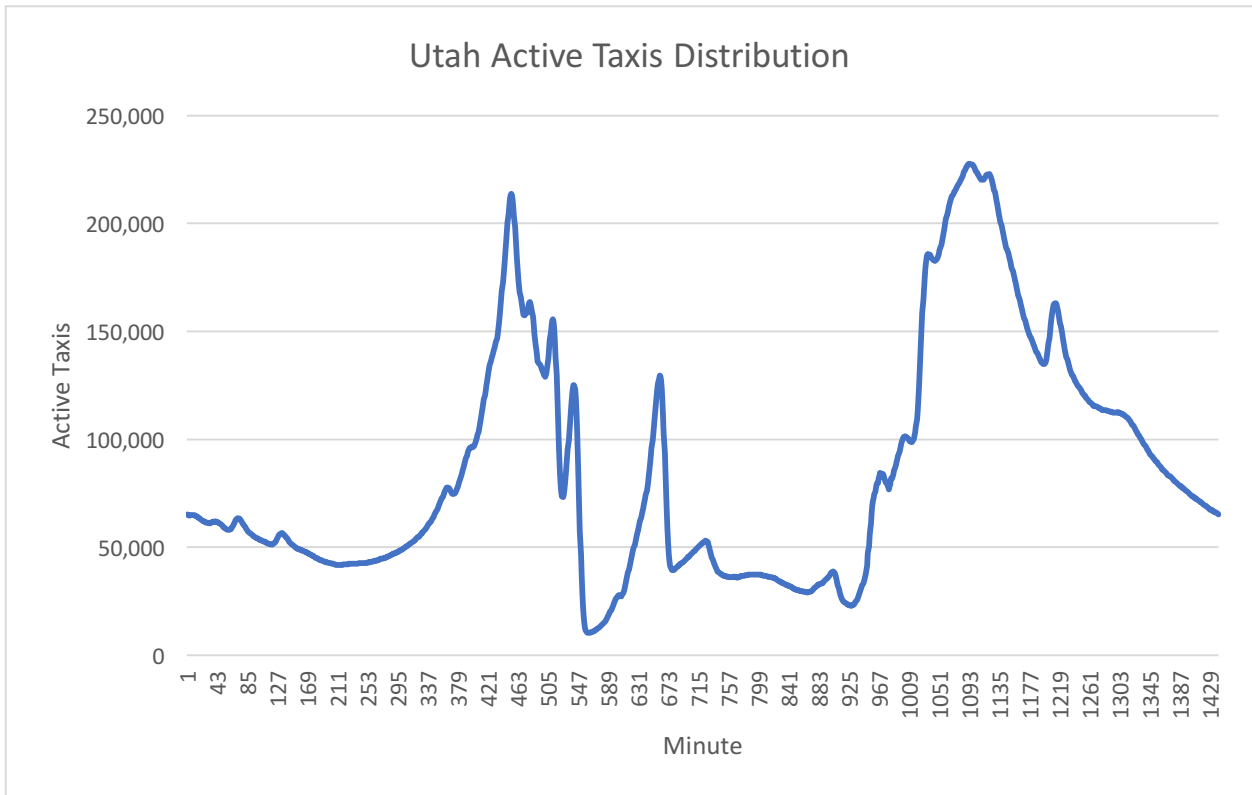
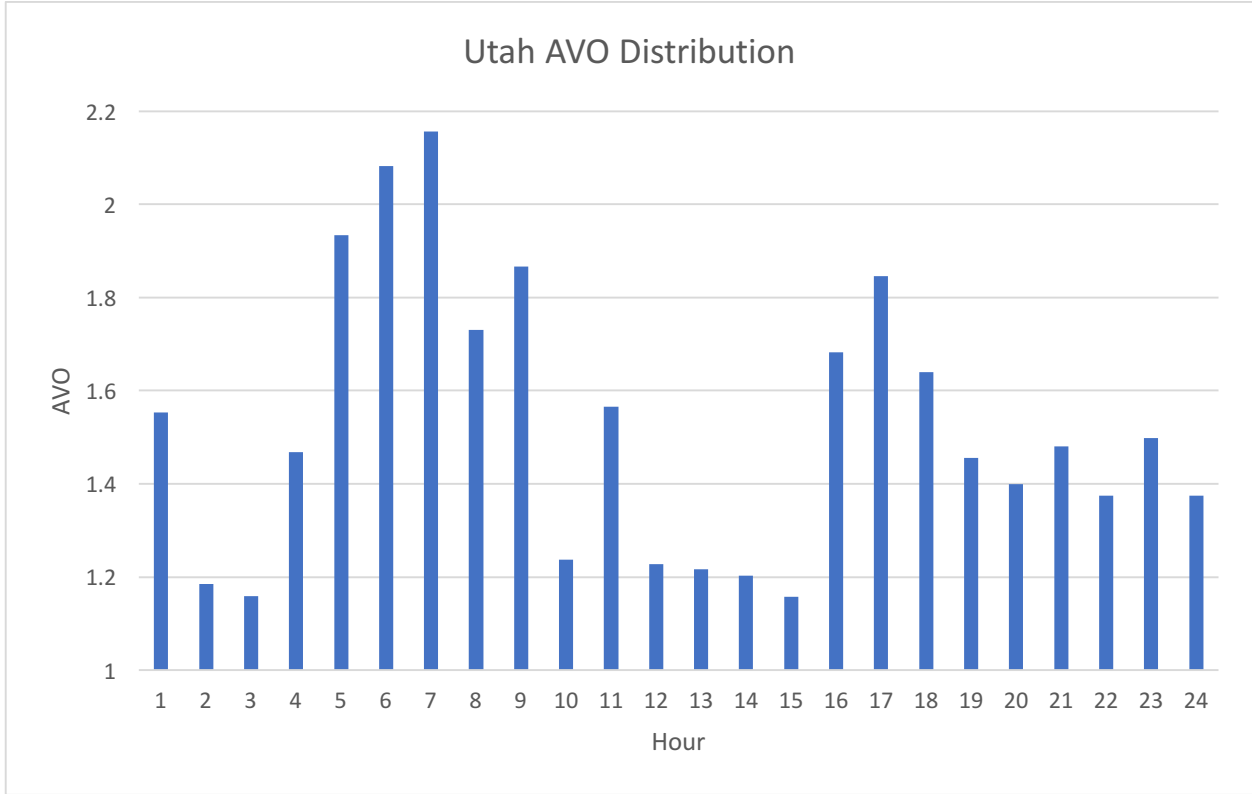


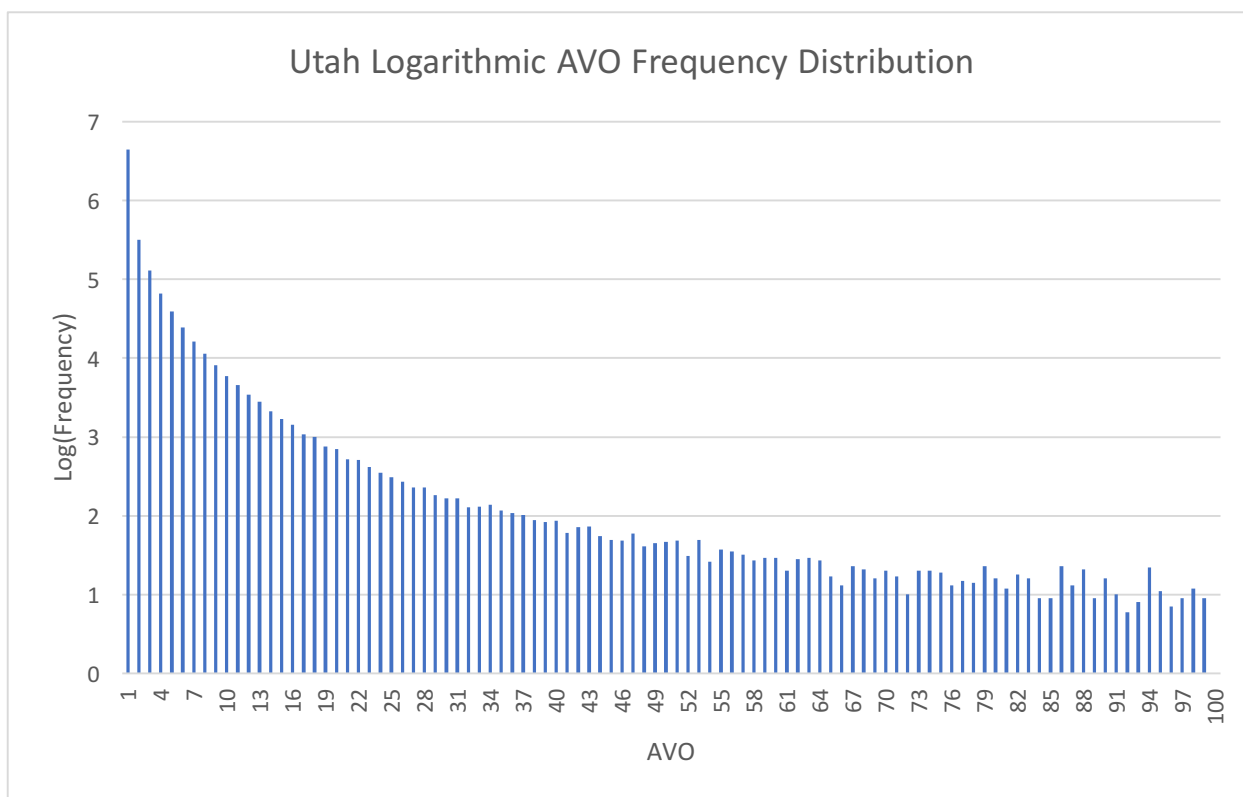
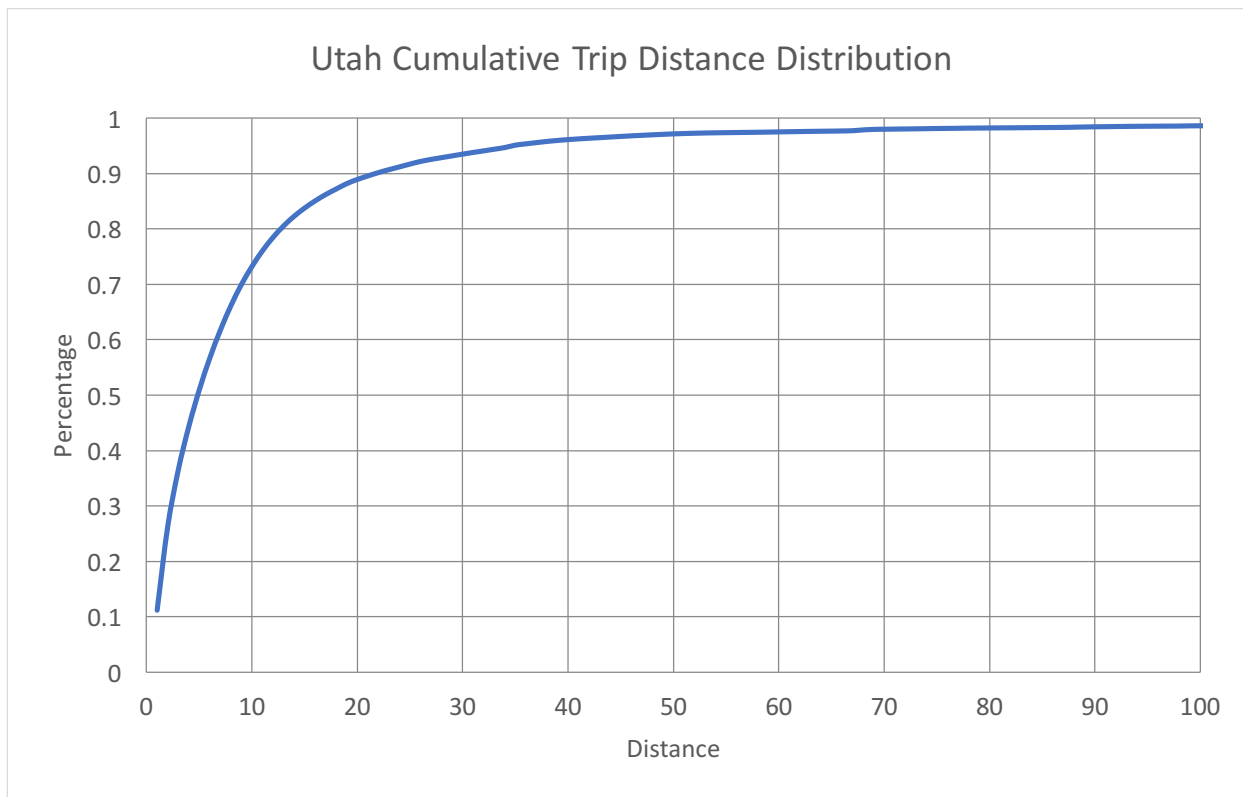


Utah:









Washington:

