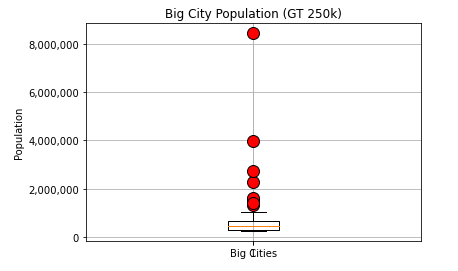
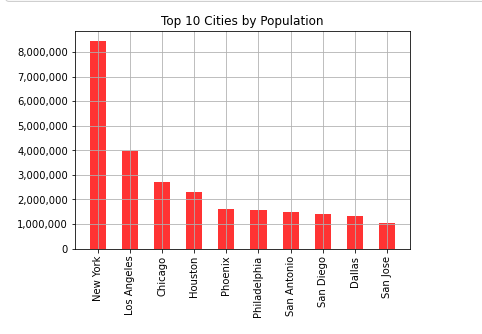
UofM Data Visualization & Analytics Bootcamp

Project 1

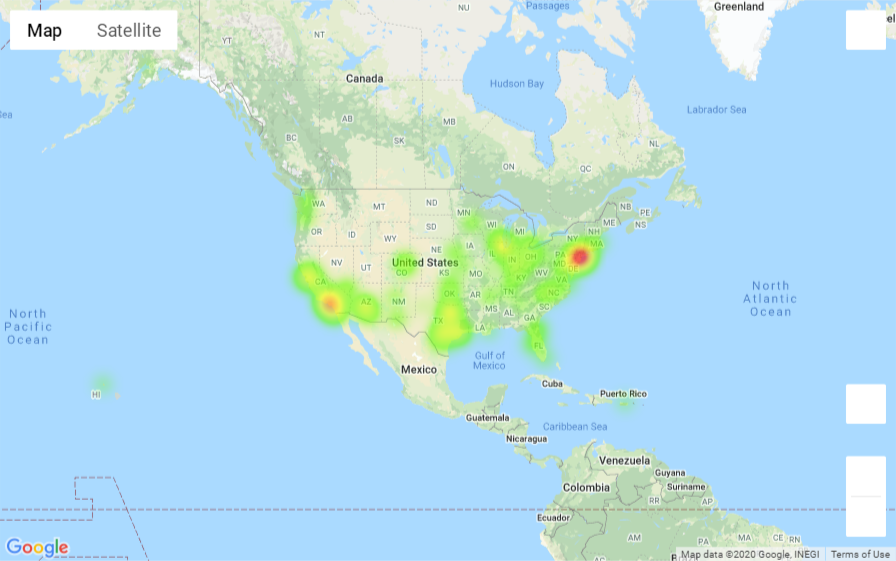
Title: Plug-in Electric Vehicle Market Analysis

Team Members: Andrew Perez, Michael Badinger, Brent Meulebroek, Gary Jeter

1. General: Manufacturers and consumers both have an interest in understanding which are the best cites to own & operate plug-in electric vehicles (PEV). Consumers want confidence in their investments. Manufacturers want to target the best markets. Our analysis identified the top 20 cities in the United States based on three main factors – population, climate and PEV infrastructure.
2. Research Questions: After initial research of optimal conditions for electric cars, the team identified the three most relevant questions to guide the analysis:
   1. Which cites have a large enough population to provide consumer confidence in PEV support and services?
   2. Which cities have the optimal weather conditions for plug-in electric vehicles (PEV)?
   3. Which cities have the best infrastructure to support PEVs?
3. Sources:
   1. PEV Research. A variety of resources were reviewed when researching PEVs. ScienceMag.org was the primary source used to establish city specification limits. Link: <https://www.sciencemag.org/news/2015/02/best-and-worst-places-drive-your-electric-car>
   2. Data Sources.
      1. Open Weather API (<http://api.openweathermap.org>): provided climate data for target US cities.
      2. Open Charge Map API ([*https://openchargemap.org/site/develop/api*](https://openchargemap.org/site/develop/api)*):* provided PEV infrastructure data (charger stations) data for target US cities
      3. Census Python Library API (2018 Data): provided census data for US cities
      4. Google Maps API (maps.googleapis.com/maps/api ): used to generate heat maps for analysis.
4. Analysis.
   1. Census Data. The team limited the analysis to cities with populations greater than 250K. The team used box plots to identify outliers, and a bar chart to visualize the cities by populations.

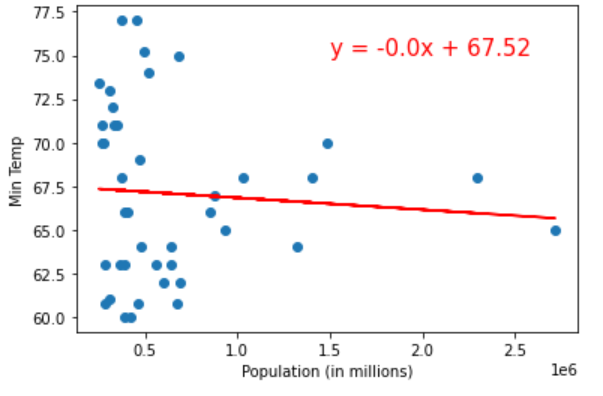
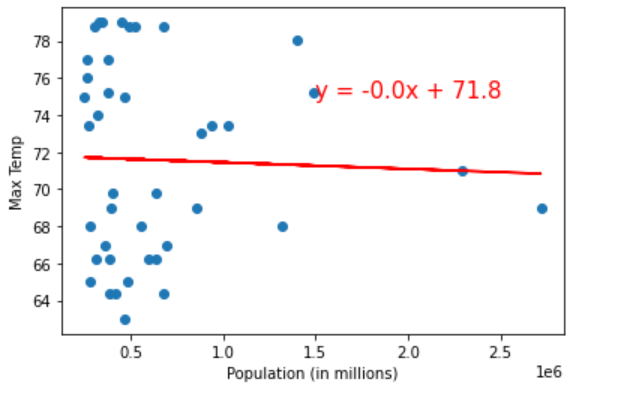


The six outliers were not removed since the larger populations are attractive markets for PEV manufacturers. This analysis limited the population of optimal cities to 84 based on populations. The heatmap below provides a visual of population. The concentration of cities in Southern California including Los Angeles and the spread of cities in Eastern Texas are locations of interest.

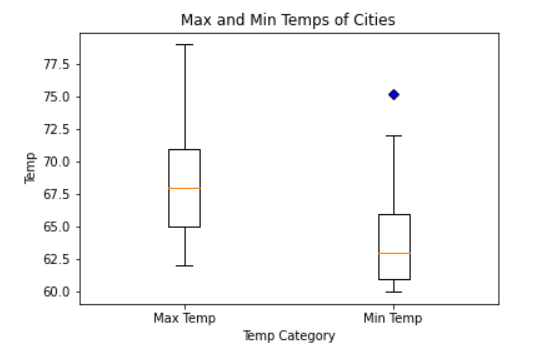


* 1. Weather Data. Using the Open Weather API, we pulled weather data for the 84 cities identified in the census analysis. Through research the team learned that a recommended temperature range is between 40 degrees and 115 degrees Fahrenheit. The API provided maximum and minimum average temperatures for each city.

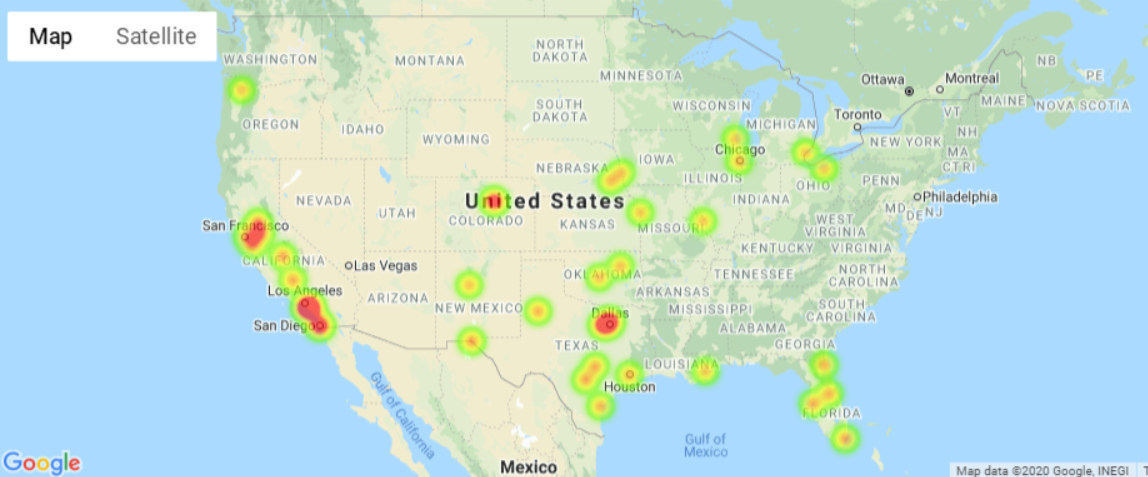
To confirm the hypothesis that population and temperature are not correlated, we performed a regression analysis between population and max temperature and min temperature. Using linear regression, the data showed that both slopes were zero and the R-Squared value for each were essentially zero (Min Temp R-Sq = 0.001; Max Temp R-Sq = 0.005). This allowed us to confidently analyze weather completely independently of population.



Using box plots, we visually confirmed that there were no outliers in temperature among the 84 cities. No cities were outside the 1.5 times the interquartile range, and the temperature data appeared to be normal. Further research discovered that operating environments between 60 - 80 degrees Fahrenheit provides the longest range and best range and battery life for PEVs. We decided to maintain the recommended temperature range (40-115 degrees) since the open weather API pulls current temperature and to account for seasonal & daily variations in temperature.

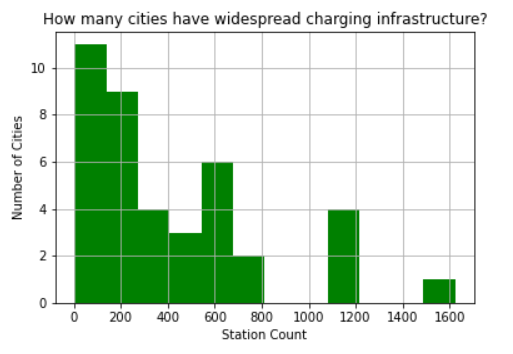


The weather analysis identified 70 cities. The heat map below reflects the 70 cities weighted by temperature. This map highlights the concentration of cities in Californa, Texas and Florida as locations of interest.



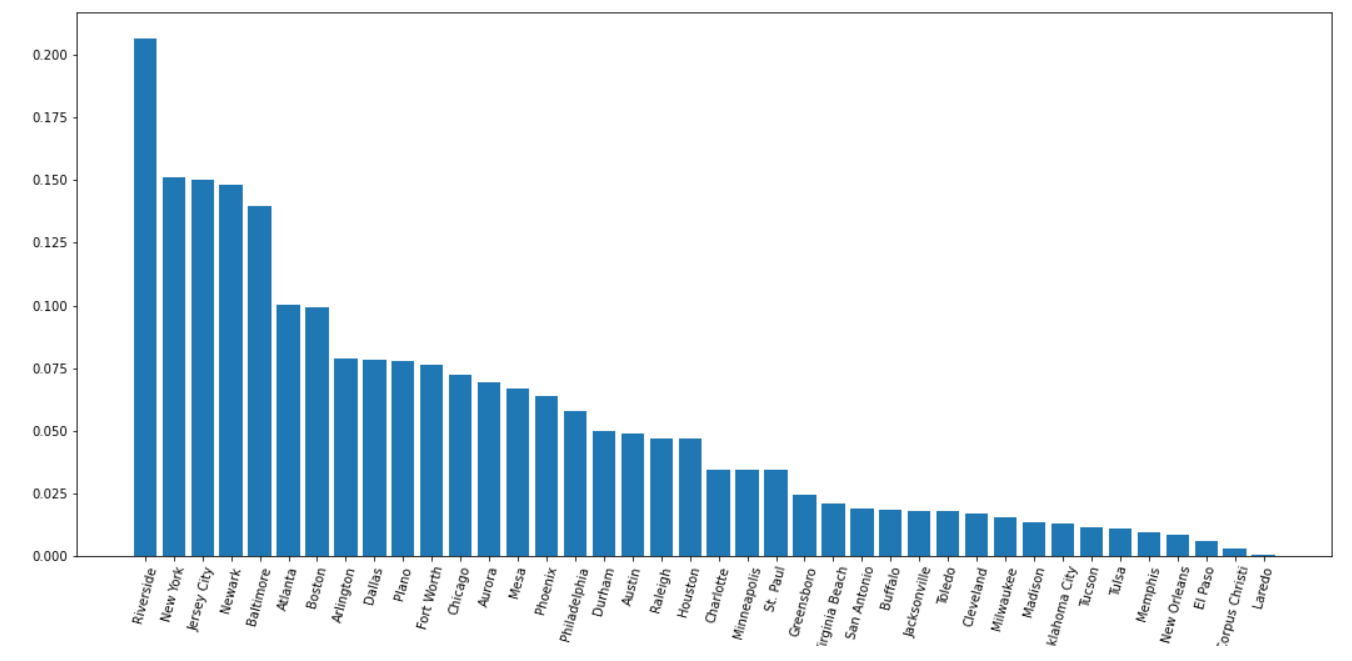
* 1. Infrastruture Data. A high quantity of charging stations that are readily available is critical infrastruture that make a city PEV friendly. We used Open Charge Map API to pull data that provided quantity of charging stations per city.

A histogram of the number of cites by charging station count showed a skewed distribution with the bulk of the cities gaving 800 station or less.

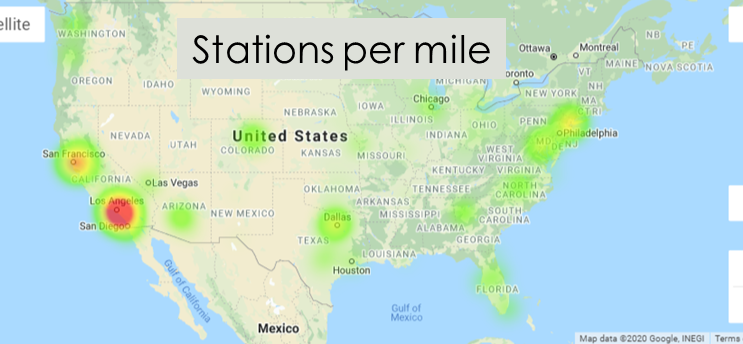


A city’s footprint is where the majority of the city’s population lives, works, worships and plays. The radius of this footprint is 50 miles and is used to calcualte the square miles within a cities footprint (7,854 square miles - area of a cirle). We used the footprint to calcualte the charging station density.

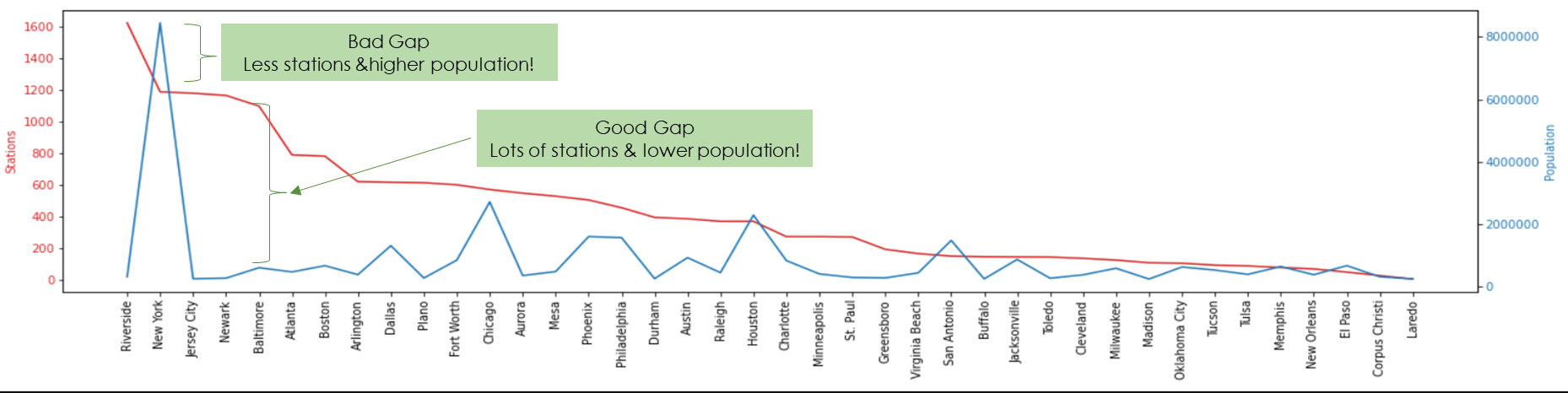
Using a bar chart of charging station density.



Using a heatmap to display the desity of charging stations, Southern California once again bcomes a location of interest.

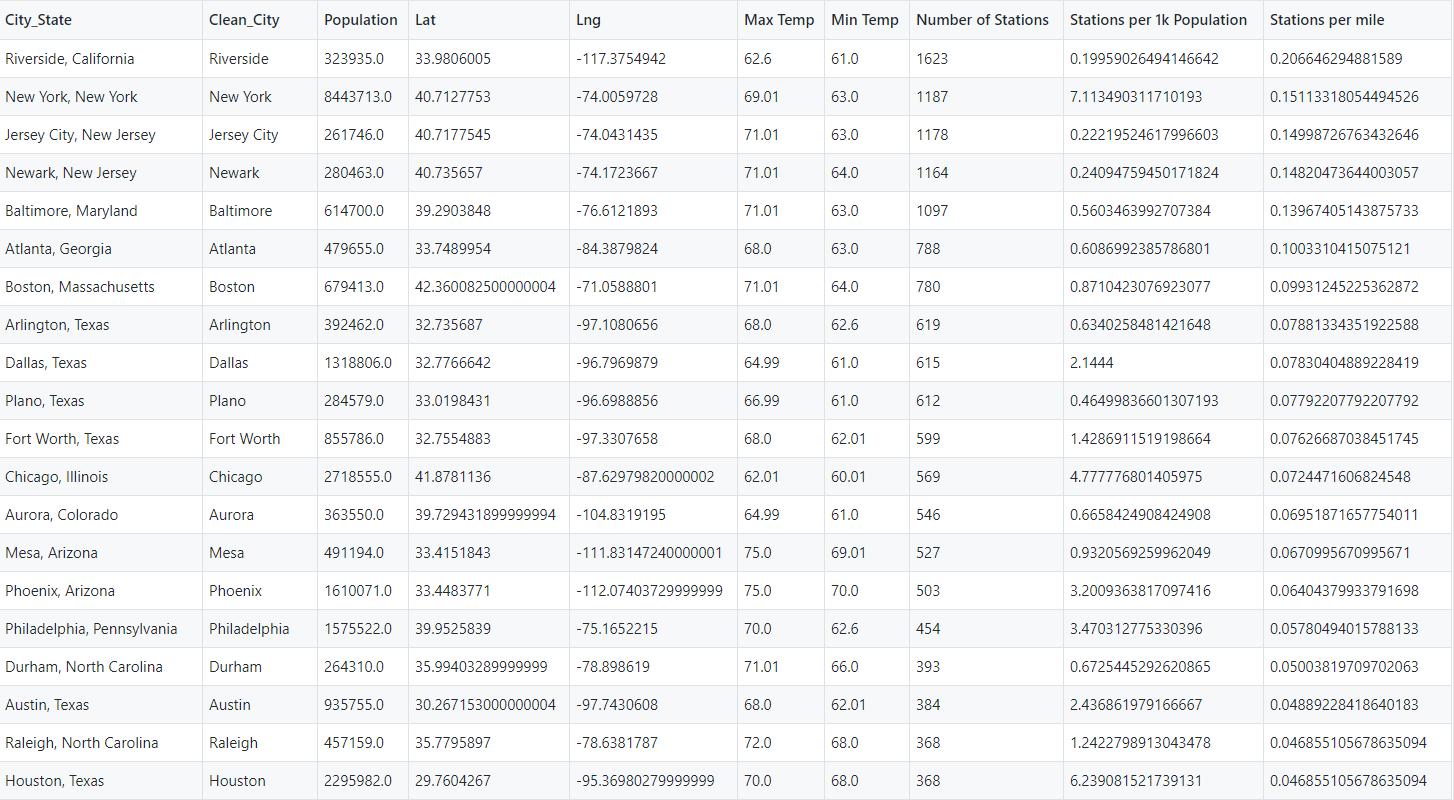


In addition to density, availability of charging stations is a factor in making a city in desireabile for PEVs. The graphical analysis below shows population and count of charging stations. A high quantity of charging stations with a lower populations is more desirable then a low quantity of stations with a high populations.



This analysis provided us the top 20 cities within the United States to sell, own and operate a plug-in electric vehicle.

1. Results. Considering population, weather and infrastruture, the top 20 cites based on our analysis is in the below table.



1. Retrospective: There were a lot of great lessons learned within this project which we will carry over to future projects. Access to data limited analysis opportunities. The team wanted to include Federal and State tax incentives as an additional screening criteria. Since the data did not align effectively with cities, we did not include it. The Census API data is messy but includes a wealth of information. Although it was time consuming to get the information, it contributed significantly to the analysis. GitHub was good for collaboration, but there were a number of growing pains. For example the team intially did not use a common Kernel which resulted in issues. We all expereinced challenges with merging branches. Our intitial query to the Open Charge API crashed PC was too general and resulted in the program to crash due to it search through 500 MB JSON. A big lesson learned was that time of day when running the script changed the results significantly. The Open Weather API uses current temperature. Future analysis we would use historical temperatures.