The New York State Energy Research and Development Authority (NYSERDA) has data on how many electric vehicles drive through a zip code on average across New York State. There are many factors that affect whether a household or person gets an EV. New England ISO, the regional power systems regulator, published monthly forecast data for EV charging load across eight regions in New England up to 2045. However, we realized that data isn't granular enough to pinpoint which location needs how much infrastructure at a given time. Therefore, first we created models based on the NYSERDA data and projected how many EVs drive through each zip code in New England. From our different models (Linear Regression, Decision Tree, Random Forest, Neural Network), we determined Random Forest Regression with log relation (to account for charging deserts because some zip codes are extremely skewed with high number of EVs) to be our best model (highest r-squared and lowest RMSE values), especially as our training and test metrics were very close together (less than 1% difference). We used cross validation and grid search to hypetune our parameters. Using the model, we were able to get a prediction on how many EVs are on the road for each New England zip code.

Then, for each zip code in New England, we identified their corresponding region and found what proportion of the total EVs in that region each zip code accounts for. We multiplied the forecasted EV load for each region in each time step by the proportion of capacity each zip code uses. This gave us EV charging demand forecasts for each New England zip code. To have a better understanding of how EV load demand changes over time, we created a number of different visualizations. In our time series choropleth map, we visualized how EV load per zip code changes every month over the next 20 years. There are a couple of major patterns. We see every zip code increasing in EV load, particularly regions in more urban areas of Massachusetts and parts of Connecticut. Another thing we see is seasonality as demand wanes in and out through the months, although with overall increase. We also created a difference choropleth map, looking at how much difference there is between current demand and how much demand that zip code needs to prepare for. The map represents how much more EV load they will need to provide using today as a baseline. It accounts for seasonality by taking the maximum additional load between now and that month, so even if demand wanes, EV infrastructure must still satisfy previous higher demands.

After the data visualization, we agreed our initial suspicion that EV demand will not be met by current capacity and growth is valid. Therefore, we projected how many additional charging ports need to be installed in each zipcode and at what time step (month/year). The US Department of Energy publishes a list of all charging stations in the US. We filtered for US New England by zip code and summed across all charging stations in a zip code and recorded the total number of charging ports. After conducting further analysis, we found that projecting the number of charging stations is unreasonable because each station has a different capacity. Instead, we focused on the number of charging ports (if we had more time for the project, a future direction is to distinguish between the different types of chargers–Level 1, Level 2, and fast chargers). We assumed that the current number of charging ports satisfies the current charging load. Then, using a ratio for the number of charging ports and the base load (maximum of 2025 load), we projected how many charging ports will be needed in each zip code for each time step. Then, we subtracted the current number of charging stations and calculated the difference between adjacent columns to create an installation schedule. In our projections, we accounted for a 0.5% annual improvement in technology (so the same number of chargers can charge more cars) and the dependency from neighboring zip codes. We also made the assumption that each zip code's demand can be met by half of the average of its nearing two zip codes. This dependency also helped us account for zip codes that currently have zero charging stations. Additionally, we created two matplotlib plots that show the increment in the number of charging ports in each zip code in five year increments (2025, 2030, 2035, 2040, 2045) on the regular scale and log scale to see the differences more clearly.