Milestone 1

Julia Cuellar

DSC 540

Final Project-M1

CSV:

Electric vehicle chargepoints

2019 EVCP use Q1 and Q2.csv

Usage data for Electric Vehicle chargepoints in Council car parks in Leeds including how often they are used, duration of charge, electricity consumed and type of vehicle.

Website:

Electric Vehicle Population Data

https://data.wa.gov/Transportation/Electric-Vehicle-Population-Data/f6w7-q2d2/data

This dataset shows the Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) that are currently registered through Washington State Department of Licensing (DOL).

API:

Electric Vehicle Title and Registration Activity

https://data.wa.gov/resource/rpr4-cgyd.csv?electric_vehicle_type=Battery%20Electric%20Vehicle%20(BEV)

This shows records of title activity (transactions recording changes of ownership), and registration activity (transactions authorizing vehicles to be used on Washington public roads).

The relationship I will make between the three is by model.

To accomplish all of the milestones, I will need to keep the relationship between the three data sets in mind. With the relationship between my three data sets being that of model (varies), I would need to remove unnecessary data variables from every data set, not just my CSV data file. I might also need to add information or fill in data by harnessing data mining techniques for my CSV data file. Plus, I could possibly remove duplicate data as well for my CSV data file. A similar approach to handle my website data will occur too. Considering that all three of my data sets can have a CSV file, I could work with my website data in that format as well. (Yet, I probably could not on the account of the requirement for our final project.) Likewise, for my API data, I will take alike steps to manipulate the data. However, for this one in particular, I will probably pull a request in my Python code and exploit the data like so. Although, I could once

again utilize a CSV file to clean and format my data. For the last milestone, once all versions of my data sets have been managed to have a readable format, I would need to merge all three data sets and store within a database. Hopefully, by then, I will know how to merge different data files to be able to store into a grand database of the final data from all three data sets.