**Team Magnolia:**

1. Topic:
   1. Topic Proposal:
   2. Github group: <https://github.com/HemaG39635070/GWU-Team_Magnolia>
   3. Dataset: <https://www.kaggle.com/datasets/amirhosseinzinati/electric-vehicle-population-data>
   4. Scope: Price Estimation: Pricing data is available, predict the price of EVs based on attributes like make, model, and battery capacity.
   5. Columns:

**1. VIN (1-10)**

: The 1st 10 characters of each vehicle's Vehicle Identification Number (VIN).

**2. County**

: This is the geographic region of a state that a vehicle's owner is listed to reside within. Vehicles registered in Washington state may be located in other states.

**3. City**

: The city in which the registered owner resides.

**4. State**

: This is the geographic region of the country associated with the record. These addresses may be located in other states.

**5. Postal Code**

: The 5 digit zip code in which the registered owner resides.

**6. Model Year**

: The model year of the vehicle, determined by decoding the Vehicle Identification Number (VIN).

**7. Make**

: The manufacturer of the vehicle, determined by decoding the Vehicle Identification Number (VIN).

**8. Model**

: The model of the vehicle, determined by decoding the Vehicle Identification Number (VIN).

**9. Electric Vehicle Type**

: This distinguishes the vehicle as all electric or a plug-in hybrid.

**10. Clean Alternative Fuel Vehicle (CAFV) Eligibility**

: This categorizes vehicle as Clean Alternative Fuel Vehicles (CAFVs) based on the fuel requirement and electric-only range requirement in House Bill 2042 as passed in the 2019 legislative session.

**11. Electric Range**

: Describes how far a vehicle can travel purely on its electric charge.

**12. Base MSRP**

: This is the lowest Manufacturer's Suggested Retail Price (MSRP) for any trim level of the model in question.

**13. Legislative District**

: The specific section of Washington State that the vehicle's owner resides in, as represented in the state legislature.

**14. DOL Vehicle ID**

: Unique number assigned to each vehicle by Department of Licensing for identification purposes.

**15. Vehicle Location**

: The center of the ZIP Code for the registered vehicle.

**16. Electric Utility**

: This is the electric power retail service territories serving the address of the registered vehicle. All ownership types for areas in Washington are included: federal, investor owned, municipal, political subdivision, and cooperative. If the address for the registered vehicle falls into an area with overlapping electric power retail service territories then a single pipe | delimits utilities of same TYPE and a double pipe || delimits utilities of different types. We combined vehicle address and Homeland Infrastructure Foundation Level Database

[HIFLD](https://gii.dhs.gov/HIFLD)

Retail\_Service\_Territories feature layer using a geographic information system to assign values for this field. Blanks occur for vehicles with addresses outside of Washington or for addresses falling into areas in Washington not containing a mapped electric power retail service territory in the source data.

**17. 2020 Census Tract**

: The census tract identifier is a combination of the state, county, and census tract codes as assigned by the United States Census Bureau in the 2020 census, also known as Geographic Identifier (GEOID). More information can be found here:

<https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13>

<https://www.census.gov/programs-surveys/geography/guidance/geo-identifiers.html>

* 1. EDA’s to be done:
* Age group <-> satisfaction demographics
* Understanding the satisfaction rates based on the Loyality of customers
* Type of travellers, class and Satisfaction rates.
* Arrival and departure delays impact on satisfaction
* Boxplots of Most important reason for satisfaction
  1. Modelling:
* Test - train split
* Logistic regression or Random Forest
* Model testing