1. Project goal

To interpret the relationship between the number of electric vehicles and air pollution in the U.S, and explore valuable patterns in the electric vehicle population

2. Description

Over the last few years, the government and organizations have been spending their time and money to advocate environmental awareness. With the goal of protecting the earth, the initiative to ban single-use plastic products has been taken in several states of the U.S, and people have been trying to be more cautious on their purchases of the products. One of the ways to help reducing air pollutant emission is to own electric cars rather than typical petrol cars that are known to emit 4.6 tons of carbon dioxide per year

The study is aimed to answer how the increase in the number of electric vehicles actually affects the air environment. Also, the major project goal is to not only study the relationship between two different factors but also to uncover hidden patterns across the sets of data by analyzing them with data science approaches learned in classes. The study will be conducted with recently updated data in 2021 or 2020.

The expected result of the project will be the inverse relationship between the number of electric vehicles and the severity of air pollution in the U.S. The major concern from the expected result will be that the data might give a weak correlation between the two factors since electric vehicles take a small sector of the vehicle population as of now and the ongoing pandemic has altered many systems such as production in factories. Hence, some of the anticipated challenges will be figuring out the pure relationship between the two factors not reflected by other possible factors such as pandemic and natural disasters.

Data science approaches such as predictive approaches and descriptive approaches will be heavily used to explore the potential patterns in characteristics of the electric vehicle population which can be marketable business insight for companies that sell electric vehicles. Exploration of the unseen patterns will be done without any hypothesis, and the data will be analyzed through the data mining approaches to figure out the valuable patterns amongst attributes of the data.

3. Data for the project :

Mainly for the interpretation of the relationship between the number of vehicles and pollution

1. Daily Air Pollution Data - India & USA (Dataset updated Jun 12, 2021)
2. Electric Vehicle by Year updated (updated Sep 17, 2021)
3. Electric Vehicle Population Data (updated Sep 21, 2021)

Mainly for exploration of unseen patterns in the electric vehicle population

4. Electric car sales 2020 (updated Sep 17, 2021)
5. Dataset of electric passenger cars with their specifications (updated Feb 9, 2021)
6. Electric Vehicles per County in NY (updated Oct 2, 2021)