**DATA ANALYSIS PROJECT**

**Election Impact Analysis on Petroleum Products**

**(2014-2024)**

**BY V SOORAJ.**

**Introduction:**

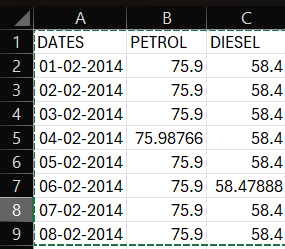
* **Objective**: Analyze the impact of elections (Lok Sabha and Kerala Legislative Assembly) on petrol and diesel prices.
* Elections that are covered during this time intervals are:
  + Lok Sabha: 2014, 2019, 2024
  + Kerala Legislative Assembly: 2016, 2021
* Duration for which the datas are collected are 45 days before and after the election dates.
* Dataset collected are the petrol diesel prices over a period of time.
* The data set collected are of daily basis rates.

**Data Collection:**

* **Sources:**The datas are collected from reputed areas like:
  + **<https://iocl.com/>**
  + **<https://www.bharatpetroleum.in/>**
  + **Local fuel bunks.**
* **Dataset Structure:**
  + **RATES.csv contains the collected dataset in the format of columns as**

**DATES,PETROL,DIESEL..**

* + Sample data is provide below;

****

**Data Preprocessing:**

* **Cleaning the Data:**
  + **Handling missing values or irregularities in price data.**
  + **Changing the date formats which are not in correct format**
* **Transformation:**
* **Aggregating data on a weekly basis .**
* **Weekly basis datas are aggregated and stored onto a new columns named pet(weekly)-petrol rates,die(weekly)-for diesel rates.**

**Data visualization:**

**For Data visualization we have used diffenet softwares and python libraries.The python libraries that used are matplotlib,sklearn,seaborn,pandas etc…**

**The softwares that used are excel and powe BI.**

**Dashboard Design:**

**For Dashboard designing we have used python,The libraries like Dash have a major role in creating the dashboard like appearance,plotly for chart creation and also pandas for some basic aggregate operations on the raw data.**

**Project processes:**

**The Project was to find the complete data set of 10 years from 2014 to 2024 during the election time.The elections covered during this time interval for both lok sabha and Kerala legislative assembly were:**

* **2014 lok sabha election**
* **2016 kerala legislative election**
* **2019 lok sabha election**
* **2021 kerala legislative election**
* **2024 lok sabha election**

**Dates which the elections are conducted In kerala:**

* **10th april-2014 election**
* **16th may-2016 election**
* **23rd april-2019 election**
* **6th april-2021 election**
* **26th april-2024 election**

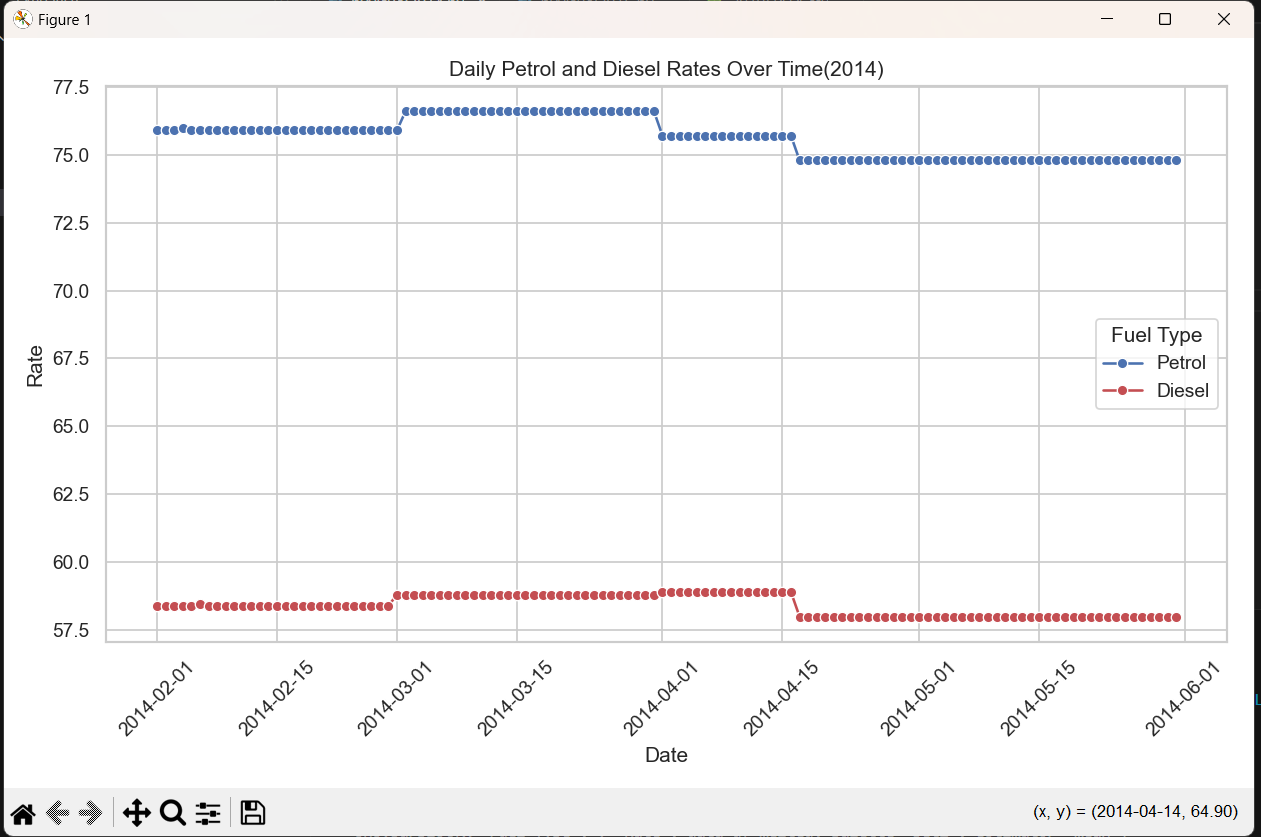
**The next step was to get the details of the price that is both before and after the election dates for both the fuels petrol and diesel.For getting more accuracy we have taken almost 60 days rates before an after the election**

**For taking these points in mind we are meant to take the details of the following month data’s for each day.**

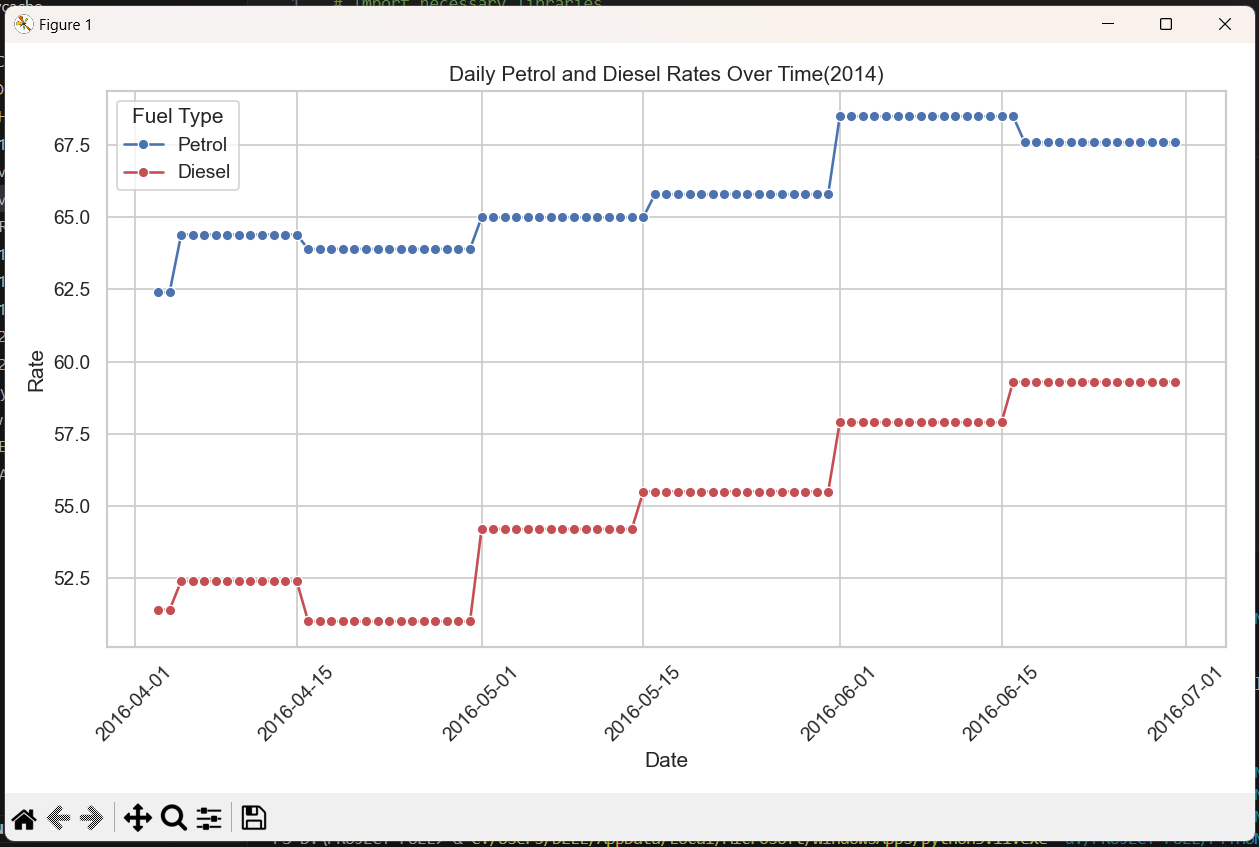
**Ie.2014(February,March,April,May),2016(April,May,June),2019(March,April,May,June),**

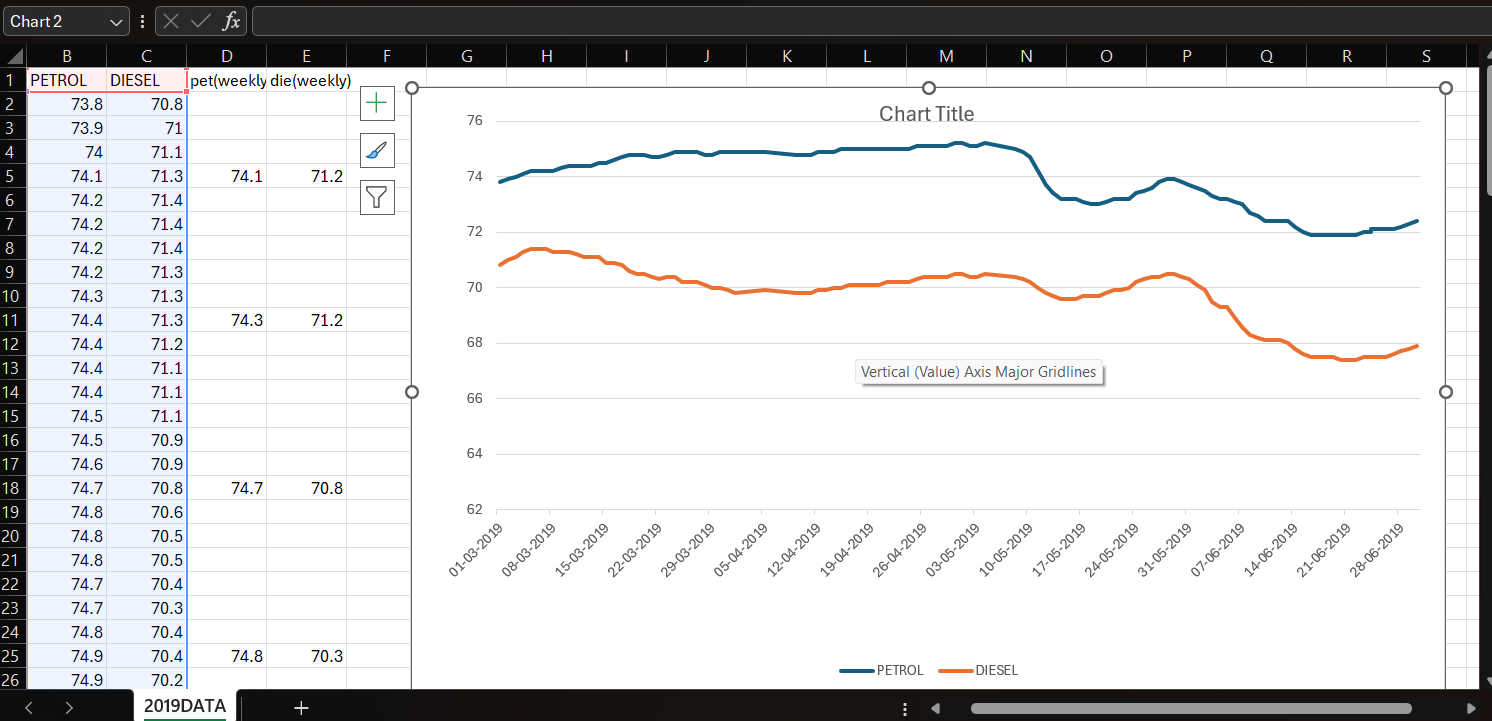
**2021(March,April,May,June),2024(March,April,May,June).**

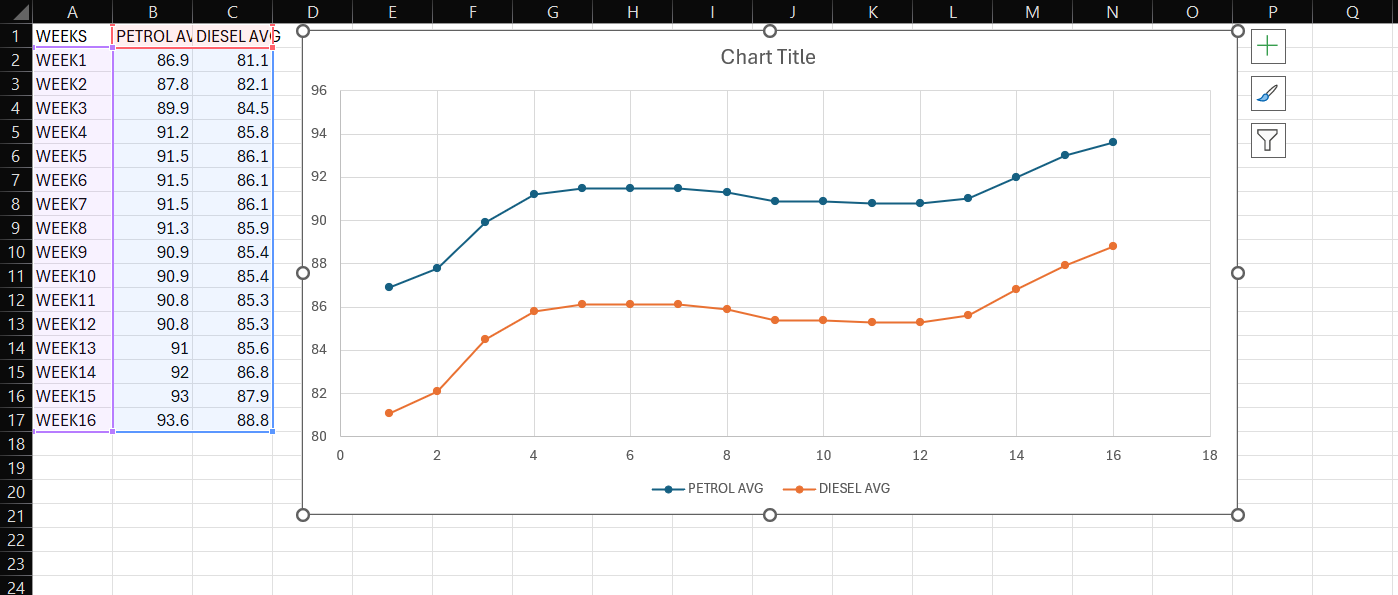
**Visualization of Dataset(yearly basis):**

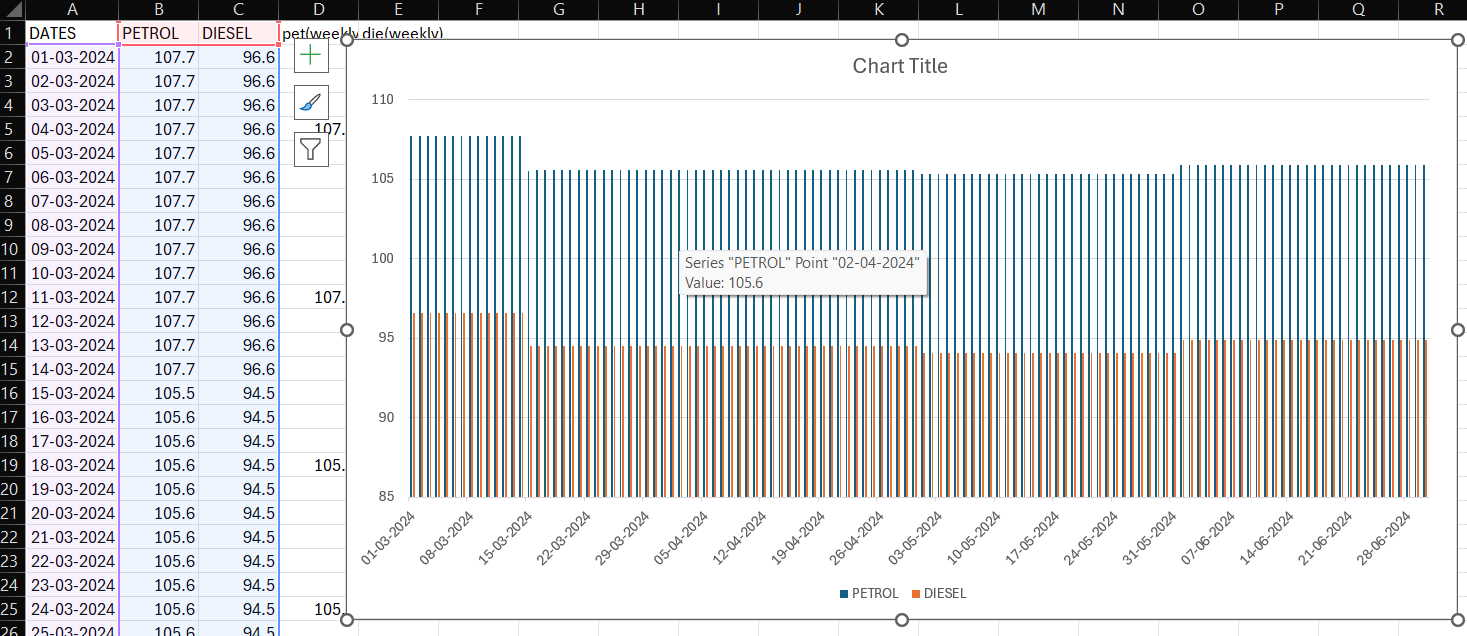
**2014:**

**2016:**

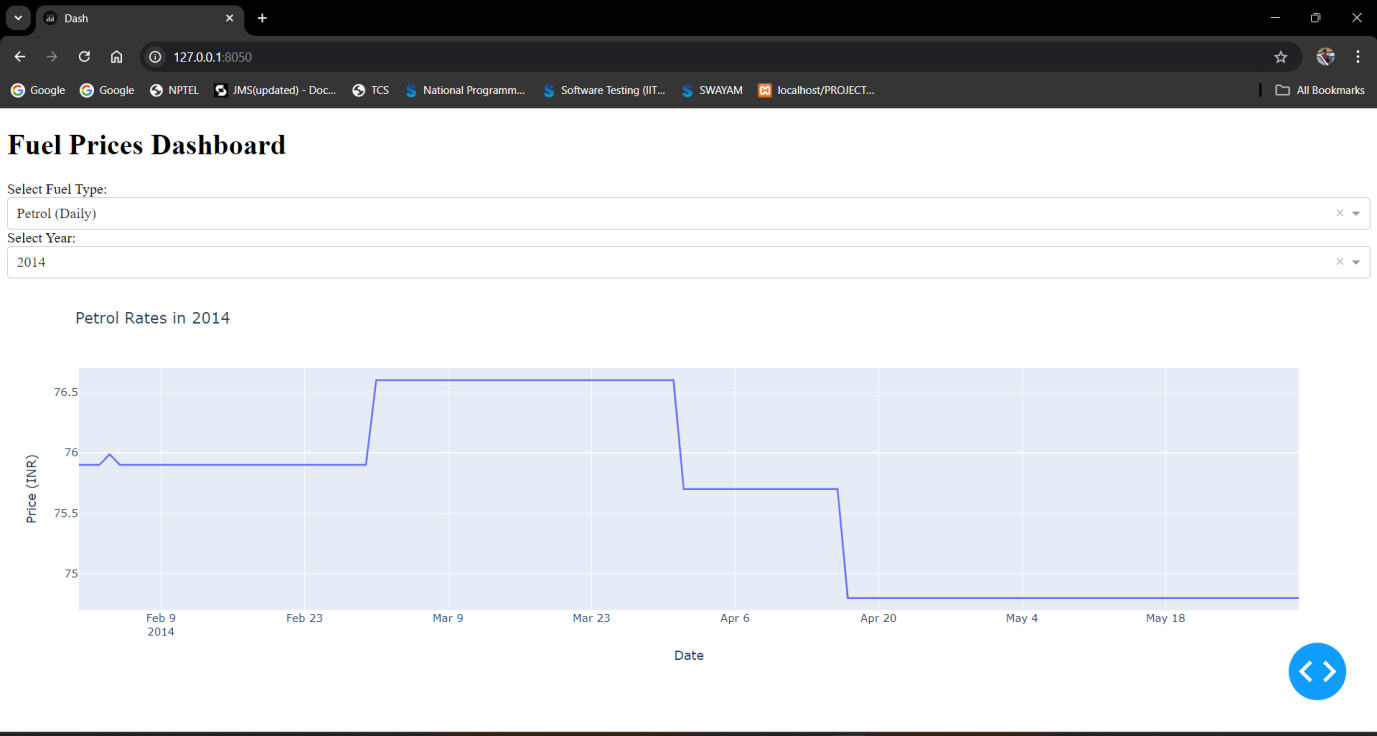
****

**2019:**

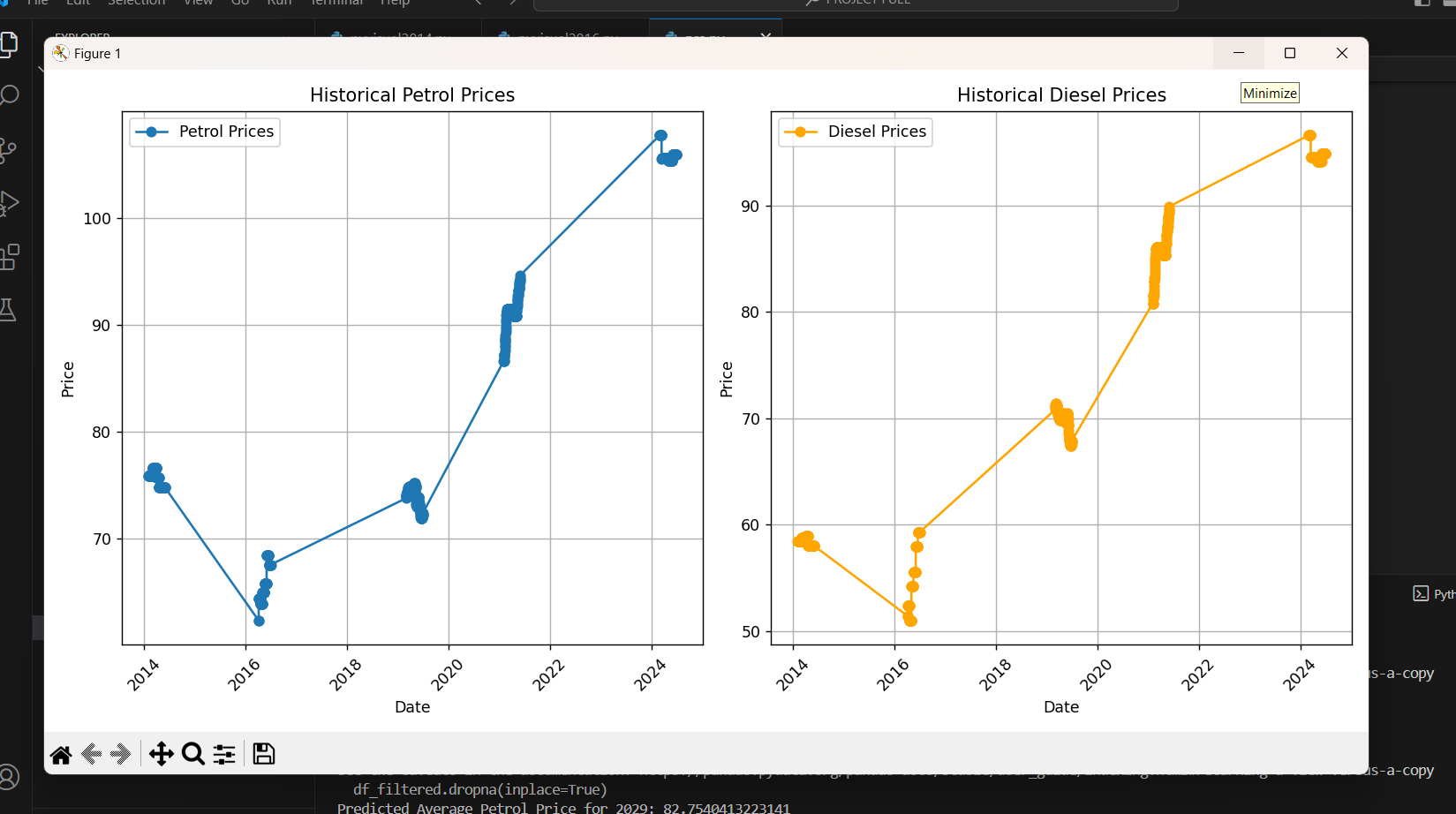
**2021(weekly basis):**

**2024:**

**Dashboard prepared using python libraries:**

****

**Charts containing the integrated values from 2014-2024**

****

**Predictive Modeling:**

Objective: Predict future price trends of petroleum products.

Approach:

Describe the machine learning models used (e.g., linear regression, random forest).

Show performance metrics and compare models.

**Model used:**Random Forest Algorithm

Why Random Forest Algorithm?

Random Forest is an effective choice for your dataset, particularly for predicting fuel prices, due to its ability to handle complex, non-linear relationships. Fuel prices often fluctuate based on various factors such as global markets,elections, leading to patterns that may not follow simple trends. Random Forest excels in capturing these non-linear dynamics by constructing multiple decision trees and aggregating their results, offering robust predictions,which can provide insights into the primary drivers of price changes, such as election periods or economic factors. By using Random Forest, you're also reducing the risk of overfitting, as it averages the predictions of many trees, making it a powerful tool for forecasting future fuel prices.

**Prediction Results:**

The prediction we conducted aims to forecast the average petrol and diesel prices in Kerala for the year 2029, based on historical data surrounding major election years, including 2014, 2016, 2019, 2021, and 2024. Utilizing a Random Forest model, we captured the underlying patterns in fuel price fluctuations, which are influenced by various factors such as elections, market dynamics, and government policies. The model predicted that the average petrol price in 2029 would be approximately ₹82.75, while the average diesel price would be around ₹76.40. These predictions, however, seem lower than expected given the observed gradual increase in prices over the past years

**CONCLUSION**:

· **Pre-Election** : Analysis of the data shows a noticeable trend of fuel prices, both petrol and diesel, gradually decreasing in the weeks leading up to election dates.

· **Fuel Price Around Election Dates**: On or just before the election dates, there is a further drop in fuel prices. This pattern appears consistent across multiple election years, strategy to maintain lower fuel costs during the election period to avoid public discomfort

· **Post-Election** : After the election, fuel prices tend to rise gradually, reversing the pre-election declines. This post-election price increase seems to follow a pattern of price normalization or recovery,

· **2029 Prices(model predicted value)**: Using a Random Forest model, predictions for the 2029 election cycle suggest an average petrol price of ₹82.75 and diesel price of ₹76.40. While these values appear lower than expected given historical trends, this anomaly highlights the need for incorporating additional economic factors or market behaviors In future predictions.