README (CS685A)

Group - 2

Environment setup:

Create a new virtual environment and install all the dependencies mentioned in **requirements.txt.**

Data Source and Extraction:

Downloaded raw data from the CPCB interface (link here) and stored them in the **temp_data/** folder. Processed this raw data to create the datasets used in insights and our model. The processed datasets are stored in the **data/** folder and the script used to process the raw data is **scripts/data_populate.ipynb**.

Execution:

Simply unzip the **data.zip** file and run each of the notebooks present in the **scripts/** folder. The following are the scripts for each of the tasks:

Insights:

- Seasonality of AQI - **AQI seasonality.ipynb**
- Spatial and temporal comparison of AQI across cities - **AQI City Comparison.ipynb**
- Dates for maximum and minimum AQI for each city - **AQI City Stats.ipynb**
- Analysing the AQI for a single day for each hour - **AQI_hour.ipynb**
- Analysing particle concentrations - **AQI_particle_concentration.ipynb**
- Analysing impact of pandemic on AQI - **Pre and Post Lockdown Comparison.ipynb**
- Correlation among particles and with AQI - **correlation.ipynb**

Model:

- Prediction model of AQI based on current particle concentrations - **Prediction.ipynb**

Other Links:

- Project Report
- Project Presentation Slides
- Codebase - https://github.com/nikhilag2711/Project_CS685