SEEM 3650 project proposal

1.Members' Name

Yan Tsz Wun 1155160184

Wai Tsz Him 1155159299

Leung Ka Ko 1155159158

2. Description of the problem

Our problem is “What hotspots should we focus on in Hong Kong in order to reduce air pollution?”. We want to analyze sources and trends of air pollution in Hong Kong and predict future air quality to inform policy decisions about emissions regulations, public transportation, and urban planning. Some of the stakeholders are the government, the general public, and the energy industry. Some decisions to be made are identifying pollution hotspots, major sources of emissions, and patterns of air quality deterioration, recommending targeted policy interventions for reducing emissions, improving public health, and raising awareness about air pollution.

3. Description of the dataset

target variable(Y):

Air quality index (AQI) data: Hourly/Daily readings from monitoring stations

target variable:

Pollutant levels: Level of pollutants that will affect the AQI(e.g., PM2.5, PM10, NO2)

Meteorological data: Data of the environment that will affect the AQI (e.g., temperature, humidity, wind speed, wind direction, etc.)

Traffic data: Data of traffic which generates pollutant (e.g., number of vehicles, types, and emissions estimates)

Population density: this will affect the generation of traffic and pollutants (e.g., number of residents per unit area, distribution of vulnerable populations, etc.)

4. Supervised or unsupervised? Regression or classification?

Unsupervised learning and Regression

5. Comments and/or concerns

We want to find out the relation between Air quality index and target variable to get a better understanding of the current state of air pollution in Hong Kong and the factors driving it.And we want to identification of pollution hotspots, major sources of emissions, and vulnerable populations. We hope the model can predict for future AQI, allowing for proactive policy interventions and public health measures.