|  |  |  |
| --- | --- | --- |
| Each City will have all 6 pollutants merged into a CSV | There will be 3 city CSV files for 2016, 2018, 20202 | The columns and rows will be dropped |
| Each Pollutant will have the monitoring data for the six cities | There will be 3 pollutant CSV files for 2016, 2018, 2020 | The columns and rows will be dropped |
| **The data files can be merged or segmented more based on what visualizations are being used** |  |  |

Data sets for charts – the data will be merged as follows:

Data sets for machine learning – the data will be merged as follows:

|  |  |  |
| --- | --- | --- |
| Air Quality predictions in a specific city over the years 2016, 2018, 2020 – initial training may only take one year into account before trying to run all three years together to see what the training results are. | Predictor Variables are:  Location  Date  Lead reading  Carbon Monoxide reading  Ozone reading  Nitrogen Dioxide reading  Sulfur Dioxide reading  Particulate Matter 10 reading [150 ug/m3] means  [150 micrograms/cubic meter]  [added to the set if time permits]  Seasonal date  Location  [taken from weather api if historical data can be captured]  Humidity  Temperature | Prediction: Air Quality will change based on the number of pollutants and their monitored volumes. A subset of this prediction is that 2022 air quality should show more pollutants present after everyone goes back to work. |

Weather data - <https://www.ncei.noaa.gov/products/land-based-station/comparative-climatic-data>

Air Quality Values - <https://www.airnow.gov/?city=Cary&state=NC&country=USA>

Standards for each pollutant - <https://www.epa.gov/naaqs> <https://www.epa.gov/criteria-air-pollutants/naaqs-table>

Background - <https://gispub.epa.gov/air/trendsreport/2020/#home>

Machine Learning – model data research

[**https://www.kaggle.com/guidosalimbeni/air-pollution-analysis-and-prediction**](https://www.kaggle.com/guidosalimbeni/air-pollution-analysis-and-prediction)

<https://www.kaggle.com/sharmamanali/air-quality-index-analysis-ml-visualisation#1.2-Importing-Libraries-->

<https://www.kaggle.com/mhajabri/air-pollution-101>

<https://www.kaggle.com/anbarivan/indian-air-quality-analysis-prediction-using-ml> - aqi FORMULAS FOR CHECK

Path to data

<https://www3.epa.gov/airquality/montring.html>

<https://www.epa.gov/aqs> - documentation

<https://aqs.epa.gov/aqsweb/documents/AQS_Data_Dictionary.html> - online reference

<https://aqs.epa.gov/aqsweb/airdata/download_files.html#Meta> – data files for machine learning [includes the weather data files]