Air Quality Measurement and Prediction

What is air pollution? <https://resphealth.org/clean-air/understanding-air-pollution/>

What are the three most harmful greenhouse gases?

In general, fluorinated gases are the most potent and longest lasting type of greenhouse gases emitted by human activities. There are four main categories of fluorinated gases—hydrofluorocarbons (HFCs), perfluorocarbons (PFCs), sulfur hexafluoride (SF6), and nitrogen trifluoride (NF3).

<https://www.epa.gov/ghgemissions/overview-greenhouse-gases>

For people the deadliest gas in the earth's atmosphere is Carbon Dioxide - Atmospheric levels of carbon dioxide—the most dangerous and prevalent greenhouse gas—are at the highest levels ever recorded. Greenhouse gas levels are so high primarily because humans have released them into the air by burning fossil fuels. May 13, 2019

<https://www.nationalgeographic.com/environment/article/greenhouse-gases>

The following studies give background information on air quality: See the Research folder in Git Rep

DurhamCountyGreenhouseGasE.pdf

Conformity Analysis and Determination Report.pdf

TriangleRegionTransportation(1).pdf

Global impact: deaths that were attributable to the joint effects of household and ambient air pollution were:

* chronic respiratory diseases that included COPD [chronic obstructive pulmonary disease],
* chronic airway diseases,
* asthma

From: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4740163/>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4740163/>

<https://www.americashealthrankings.org/>

this is all background information to show that tracking air quality is important for everyone’s quality of life especially in areas that have high concentrations of airborne particulate matter.

**Plots and visualizations:**

**Suggested comparison – Rural vs Urban**

**Urban – California, Texas, New York, Florida**

**Rural – Mississippi, Vermont, Alaska, Wyoming**

**Dates: 2010 to 2020**

**Suggested Data Sets – for air quality coming from Daily Data from the EPA**

**If using weather data – use the free weather api**

**With location and date information – we could use weather api to get temperature and humidity to correlate to air quality data – this plots nicely and is easy to see the relationship**

**See:** [Air Pollution 101 | Kaggle](https://www.kaggle.com/mhajabri/air-pollution-101/notebook#Table-of-contents) Example

**For machine learning – get a comparison of the AQS api data with weather data to predict future air quality**

Data sets for visualizations, and machine learning

Samples of the data sets are in the Git Repo:

<https://www.kaggle.com/sogun3/uspollution> us information from 2009 to 2016 history – Too large to add to Git Repo – see the Kaggle link

[Download Daily Data | US EPA](https://www.epa.gov/outdoor-air-quality-data/download-daily-data) – data set downloads by pollutant to 2021 by area – gives a spreadsheet – Look at the folder ~/DataSets/Daily Data EPA for the monitoring data

<https://java.epa.gov/castnet/clearsession.do> - monitoring datasets - allows many different data reports – see ~/DataSets/Castnet Data/

[CASTNET (epa.gov)](https://java.epa.gov/castnet/epa_jsp/prepackageddata.jsp) – prepackaged data sets -use hourly gas and ozone sets

[AQS API | AirData | US EPA](https://aqs.epa.gov/aqsweb/documents/data_api.html) – API download – for live update – for machine learning

visualizations – examples - reference

<https://www.airnow.gov/> - another resource – only interactive data

[National Air Quality: Status and Trends of Key Air Pollutants | US EPA](https://www.epa.gov/air-trends) – for reference and visualizations

[Particulate Matter (PM) Basics | US EPA](https://www.epa.gov/pm-pollution/particulate-matter-pm-basics)

<https://www.epa.gov/pm-pollution/technical-data-and-reports-particulate-matter-pm-measurements-and-sip-status> -

For Reference - Monitoring sites

<https://www3.epa.gov/ttn/amtic/amticpm.html> Ambient Monitoring Technology site – for reference

<https://www.epa.gov/amtic> Ambient monitoring - with all of the methodology and reporting by EPA

[AMTIC - Ambient Air Monitoring Assessments | US EPA](https://www.epa.gov/amtic/amtic-ambient-air-monitoring-assessments) – Gaseous Monitoring Dashboards

<https://sti-r-shiny.shinyapps.io/CO_Dashboard/> Carbon Monoxide

<https://sti-r-shiny.shinyapps.io/NO2_dashboard/> Nitrogen Dioxide

<https://sti-r-shiny.shinyapps.io/NO_dashboard/> Nitric Oxide

<https://sti-r-shiny.shinyapps.io/ozone_dashboard/> Ozone

<https://sti-r-shiny.shinyapps.io/SO2_dashboard/> Sulfur Dioxide

[PM2.5 Dashboard (shinyapps.io)](https://sti-r-shiny.shinyapps.io/QVA_Dashboard/) Particulate matter dashboard

<https://www.epa.gov/castnet> - monitoring sites

References on data gathering from monitoring sources

<https://www.epa.gov/amtic/air-monitoring-methods> - methodology for data gathering

<https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants>

<https://www.epa.gov/amtic/compendium-methods-determination-inorganic-compounds-ambient-air>

<https://www.epa.gov/amtic/compendium-methods-determination-toxic-organic-compounds-ambient-air>

**Machine Learning modeling: Ideas for machine learning using our data sets**

[**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>