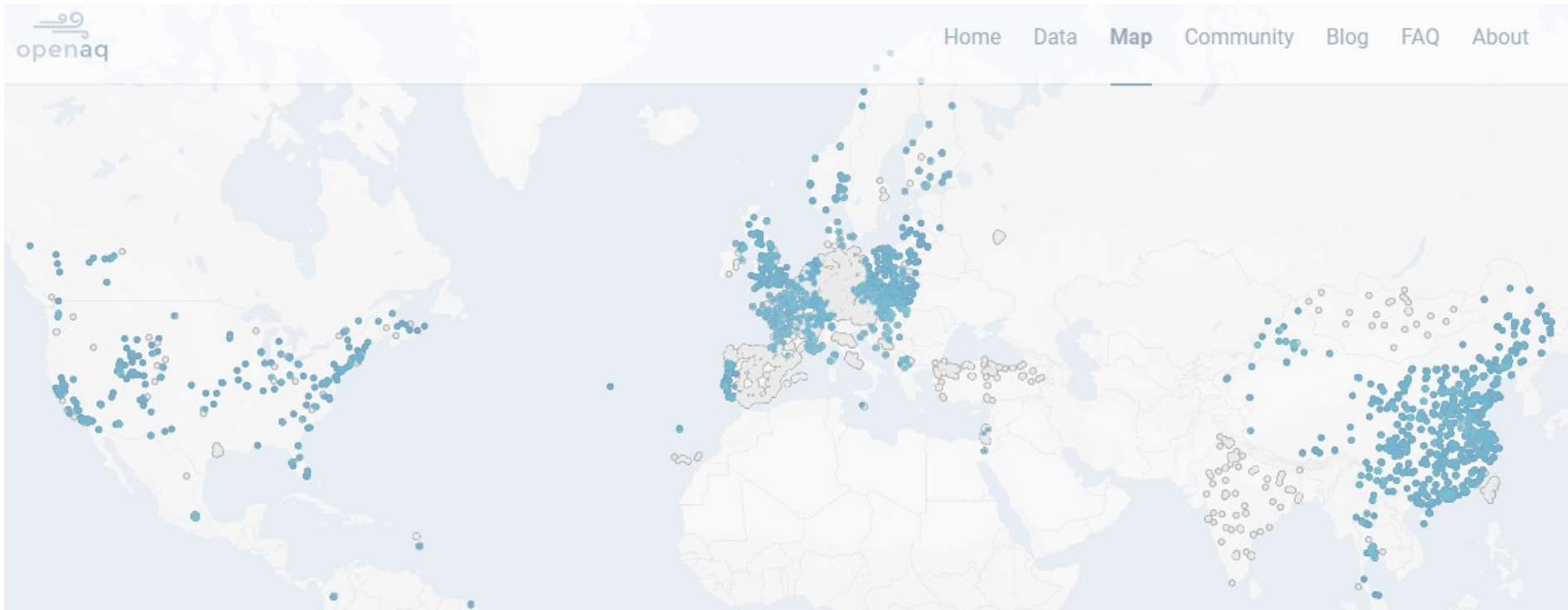# Open AQ Data Study

Jonathan Stumpf
EMSE 6992 Fall 2018

Dataset Background - Open AQ

Data Gathering

OpenAQ posts their data on Amazon Web Services.

After creating an AWS account, one can use a tool called Athena to run queries

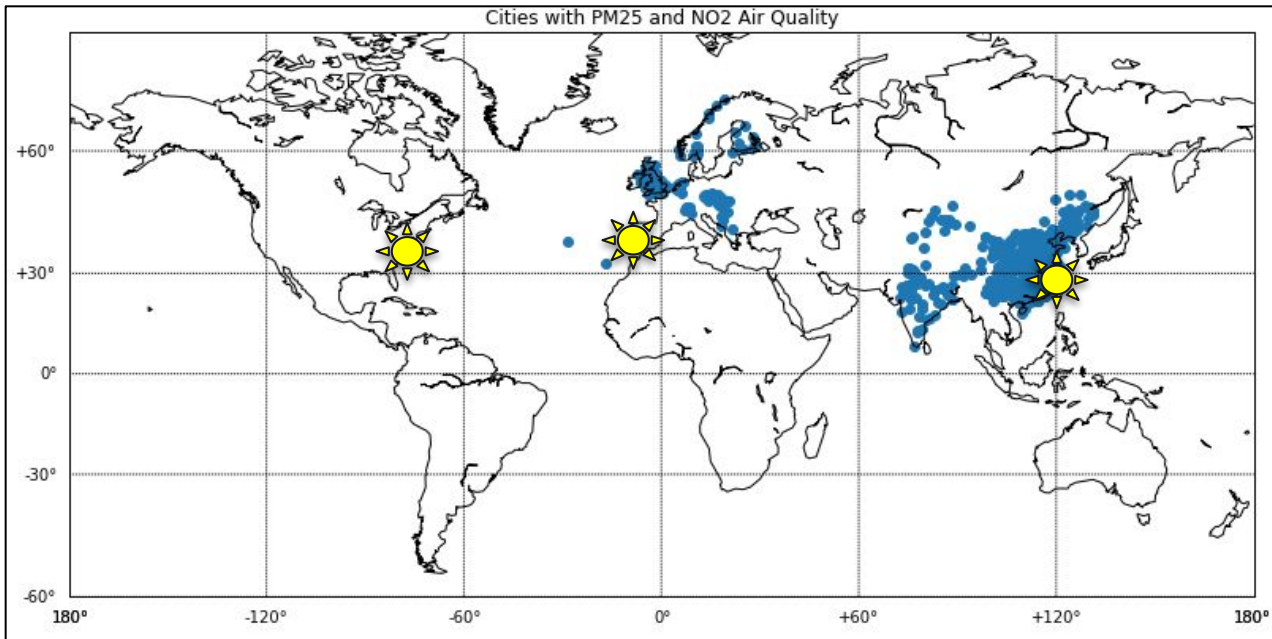Querie results can be saved off as CSVs...

Data Cleaning | Down Sampling

Cities with PM25 and NO2 Air Quality

Cleaning Data:

- Nested structures within pandas' 'read_csv'
- Unit conversion for alignment
- Outlier elimination

Pollutants Analyzed

'PM2.5 refers to atmospheric particlulate matter (PM) described as the mixture of solid particles and liquid droplets in the air. Some exampls include dust, ash, sea-spray, and soot from power generation, domestic heating and vehicle engines' -Department for Environment Food & Rurual Affairs
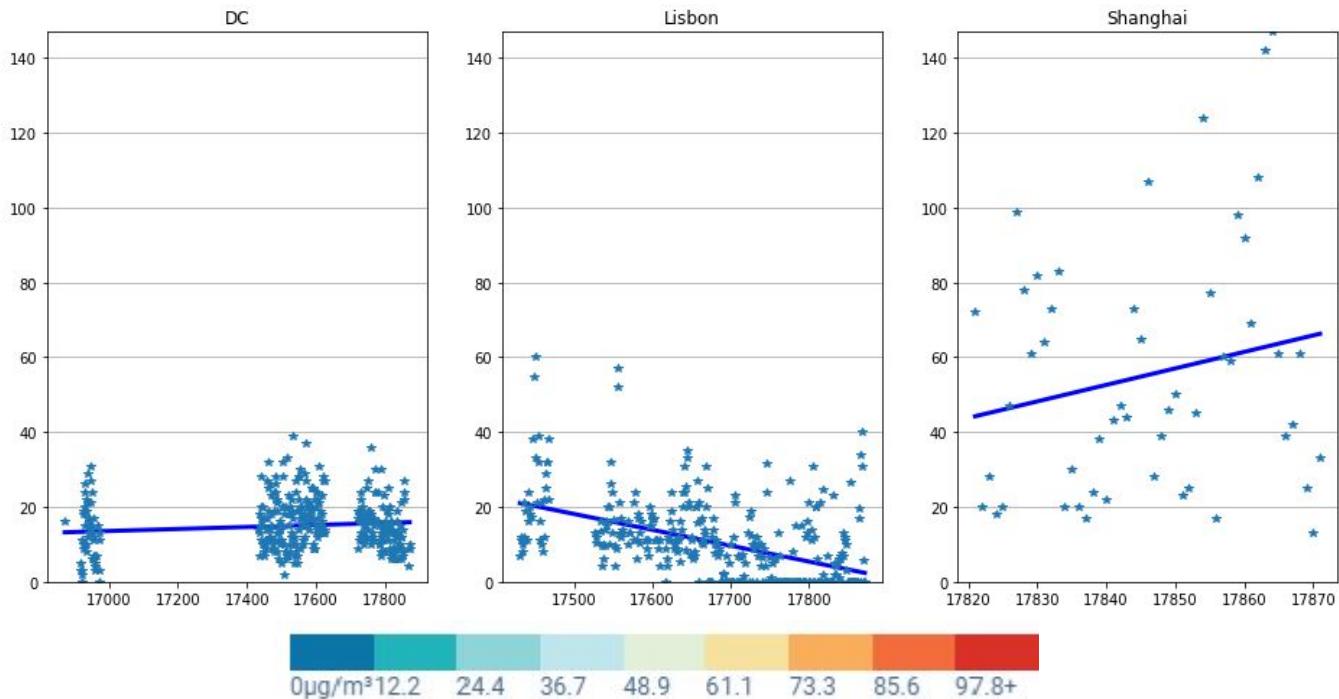
'CO - Carbon Monoxide

'NO2- Nitrous Dioxide - primarily gets in the air from the burning of fuel. Emissions from cars, trucks, buses, power plants, off road equipment
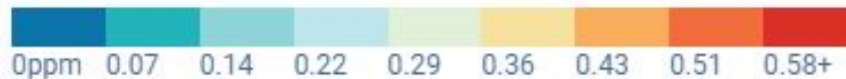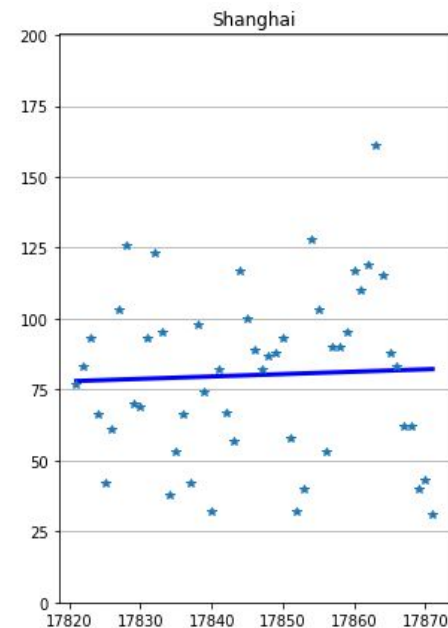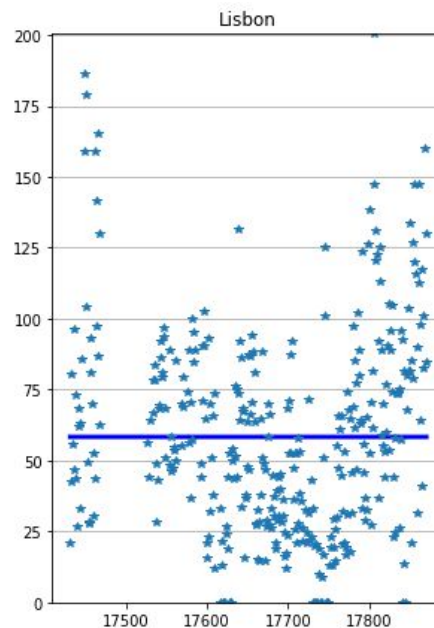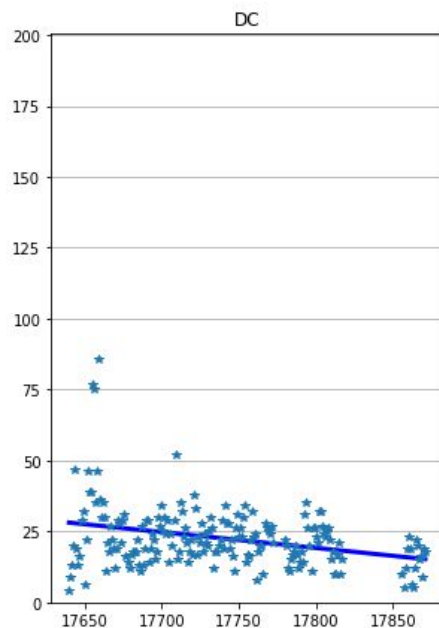
PM2.5

DC	Lisbon	Shanghai
15.0 (14.0)	10.3 (9.5)	55.3 (47.0)

NO2

DC	Lisbon	Shanghai
22.6 (21.0)	58.3 (53.4)	80.1 (83.0)

CO

DC	Lisbon	Shanghai
641.6 (580.0)	556.0 (330.0)	921.6 (900.0)