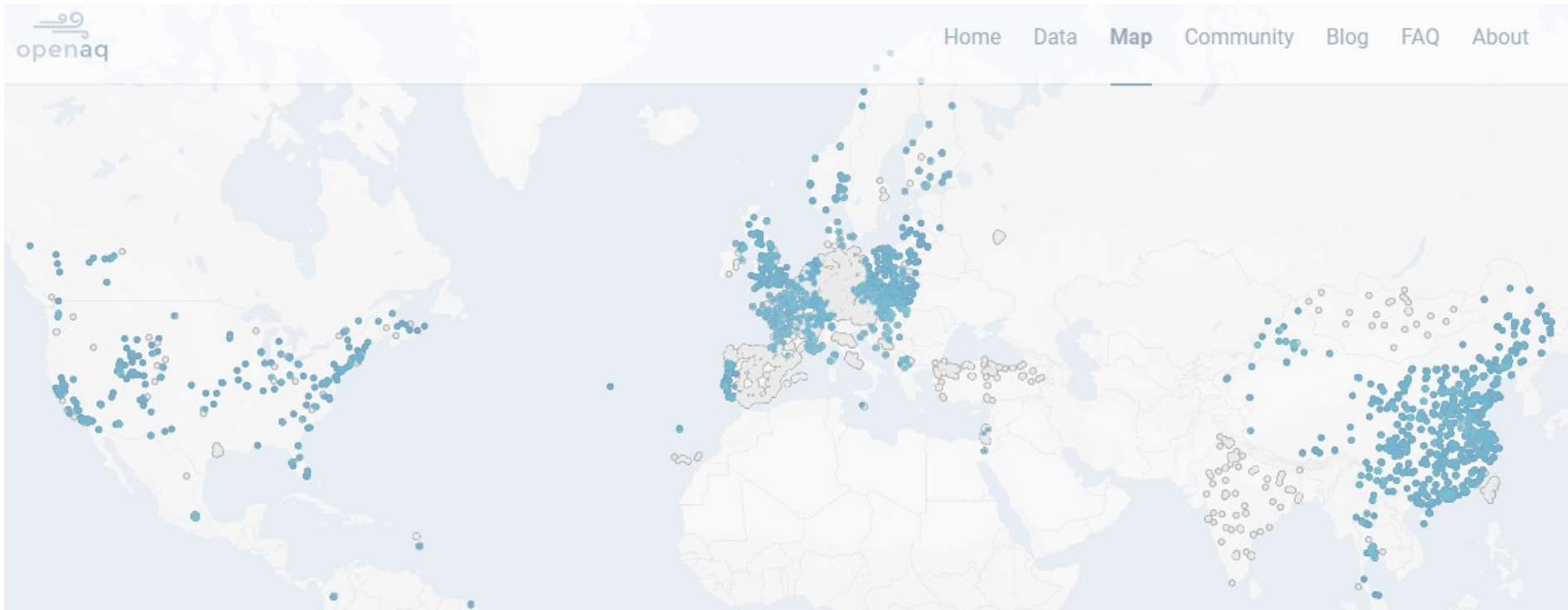




# Open AQ Data Study

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

Query results can be saved off as CSVs...

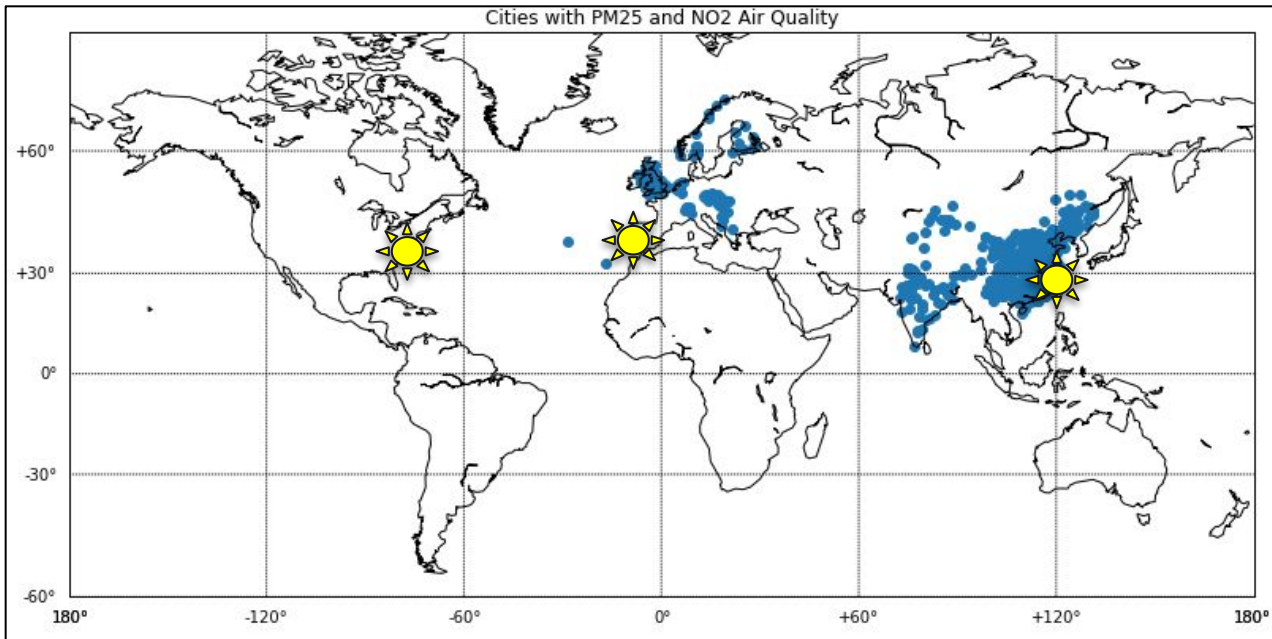
The screenshot displays the Amazon Athena Query Editor interface. The top navigation bar includes 'Athena', 'Query Editor' (active), 'Saved Queries', 'History', 'AWS Glue Data Catalog', 'Settings', 'Tutorial', 'Help', and 'What's new'. On the left, the 'Database' dropdown is set to 'default', and a search bar for 'Filter tables and views...' is present. Below this, 'Tables (1)' are listed, including 'openaq', and 'Views (0)' are shown with a note to create a view from a query. The main editor area shows two queries; 'New query 2' is active and contains the SQL: 

```
1 SELECT *
2 FROM openaq
3 WHERE location='浦东川沙'
```

. Below the query editor are buttons for 'Run query', 'Save as', and 'Create', along with performance metrics: '(Run time: 22.59 seconds, Data scanned: 4.91 GB)'. A 'Format query' button and a 'Clear' button are also visible. The 'Results' section at the bottom shows a table with columns: date, parameter, location, value, unit, and city. The first row of results is: 

	date	parameter	location	value	unit	city
1	{utc=2018-10-17T23:00:00.000Z, local=2018-10-18T07:00:00+08:00}	pm25	浦东川沙	20.0	µg/m³	上海

# Data Cleaning | Down Sampling



Cleaning Data:

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

# Pollutants Analyzed



‘PM2.5 refers to atmospheric particulate matter (PM) described as the mixture of solid particles and liquid droplets in the air. Some examples include dust, ash, sea-spray, and soot from power generation, domestic heating and vehicle engines’ -Department for Environment Food & Rural Affairs

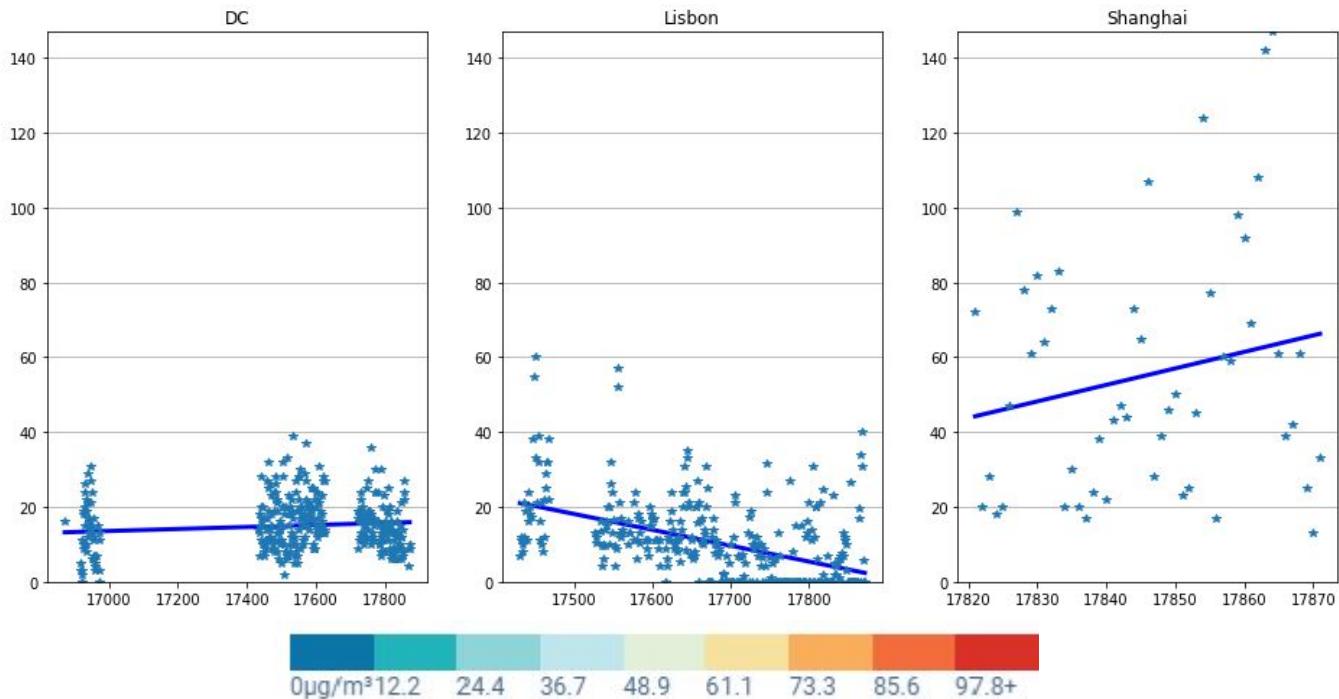
‘CO - Carbon Monoxide

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

# PM<sub>2.5</sub>



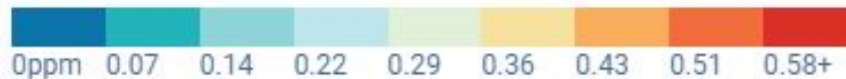
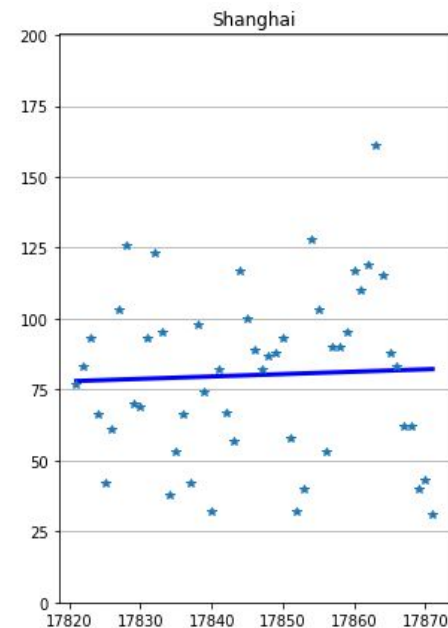
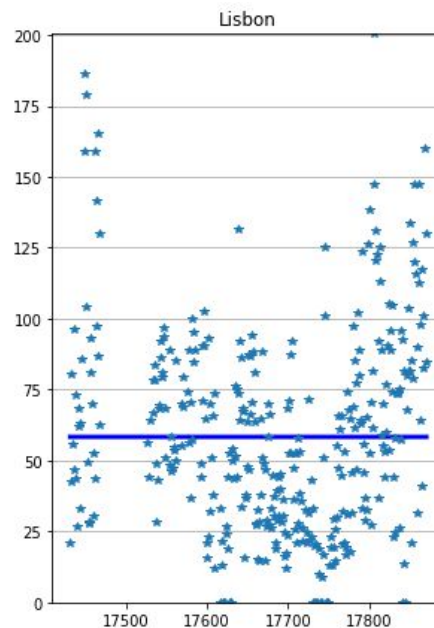
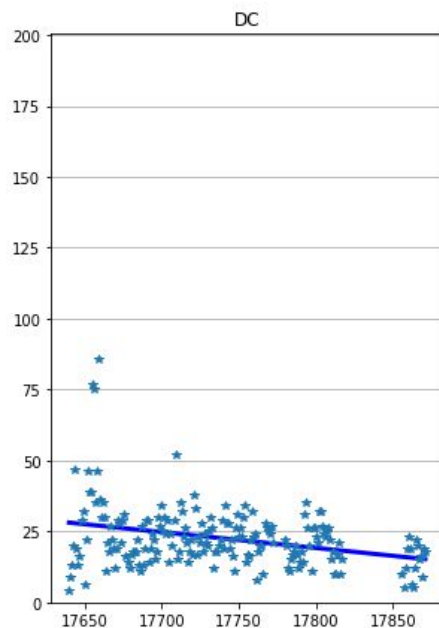
DC	Lisbon	Shanghai
15.0 (14.0)	10.3 (9.5)	55.3 (47.0)



# NO<sub>2</sub>



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)

