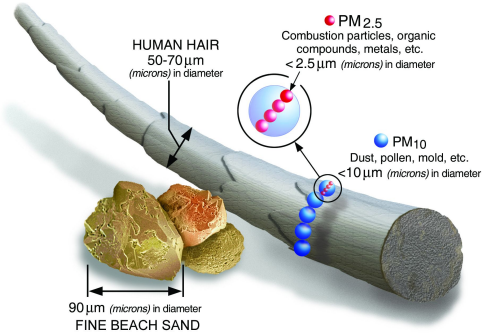
Exploring Particulate Matter 2.5

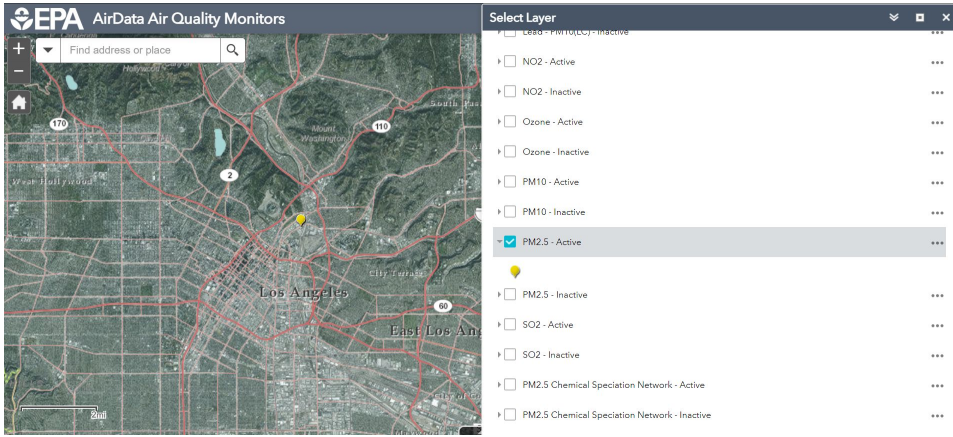
Leon Zha

What Is PM2.5 And Why Does It Matter?



- PM stands for particulate matter
  - mixture of solid particles and liquid droplets found in the air.
- PM2.5 : fine inhalable particles
  - diameters <= 2.5 micrometers
  - 30 times smaller than the width of human hair
- Short-term health effects: eye, nose, throat and lung irritation, coughing, sneezing, runny nose and shortness of breath.
- Long term exposure associated with increased rates of chronic bronchitis, reduced lung function and increased mortality from lung cancer and heart disease.

A Look At Our Data Source



- Data sourced from the EPA
  - https://epa.maps.arcgis.com/apps/webappviewer/index.html?id=5f239fd3e72f424f98ef3d5def547eb5&extent=-146.2334,13.1913,-46.3896,56.5319
- Los Angeles-North Main Street Station Monitor
  - Looking at PM2.5 Chemical Speciation Network - Active
  - Using daily data from 2016-2020

Three Guiding Questions

Are there any particular days with high levels of PM2.5 in 2016?

What does the chemical makeup of PM2.5 in 2016 look like?

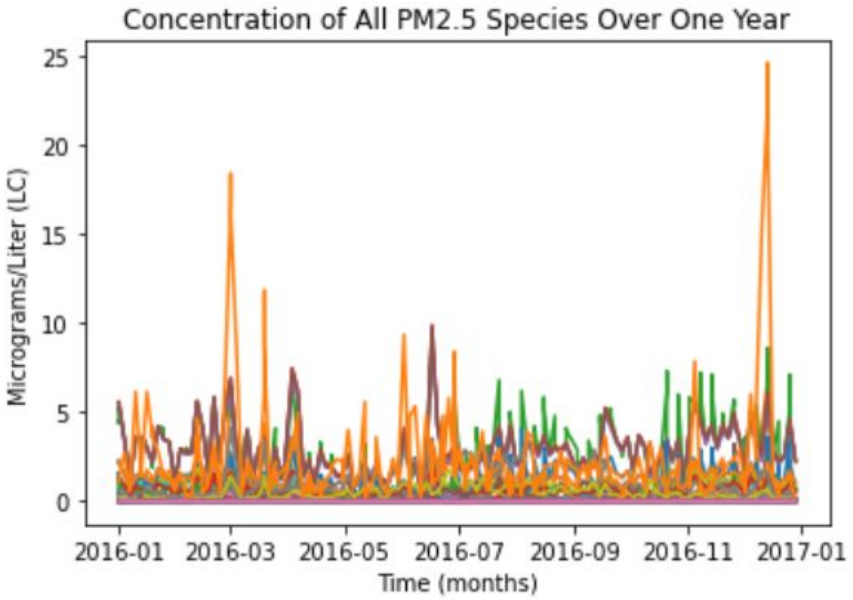
What kind of trends do we see in PM2.5 from year to year?

A Quick Look At Data From 2016

- 36265 rows
- 34 columns
- Discard data that doesn’t change
  - Ex. state name, longitude/latitude, county code
- Discard other non-relevant data
  - AQI, daily criteria indicator, first maximum hour
- End up with 67 unique particulate species (rows) and three columns (Parameter Name, Arithmetic Mean, and Date)

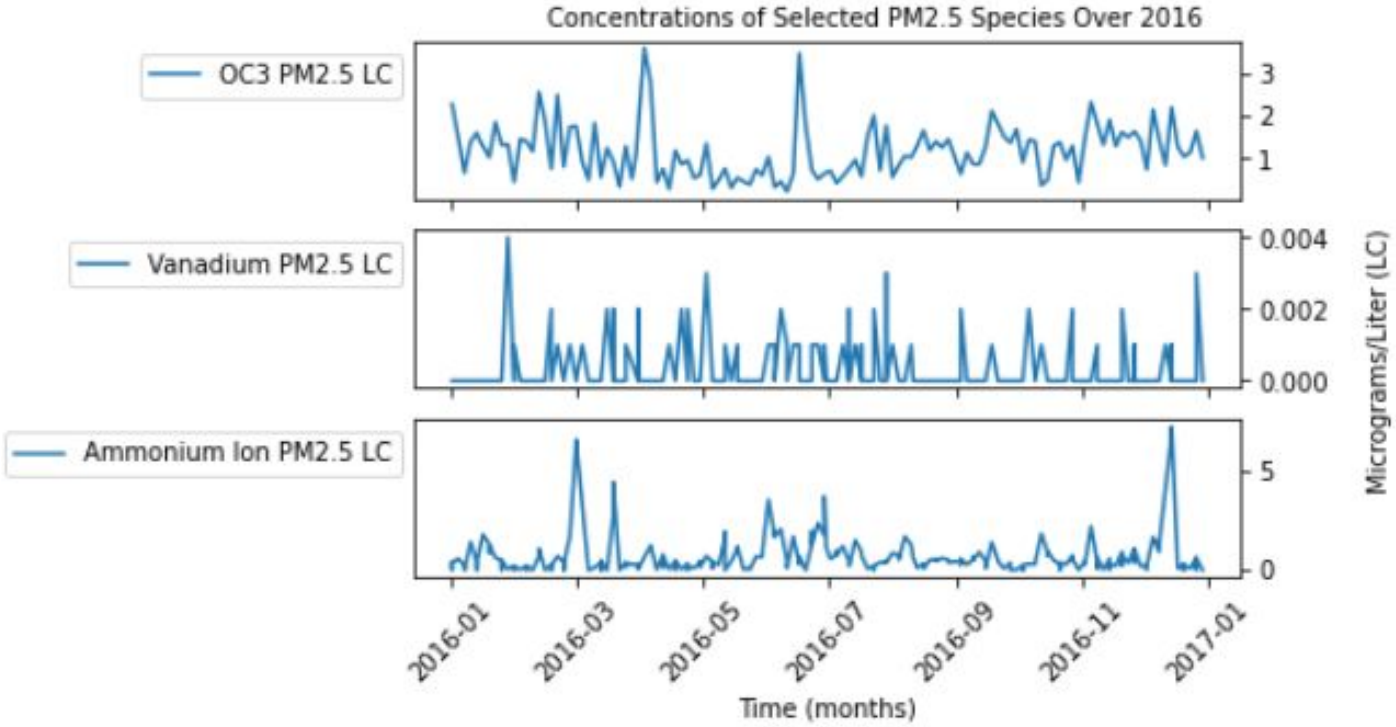
Let’s Look At The Trend For All Species Over 2016

- Very noisy, as expected because of the sheer amount of chemical species we are looking at
- No particular days stands out except for the two big peaks



Let’s Look At The Trend For Selected Species Over 2016

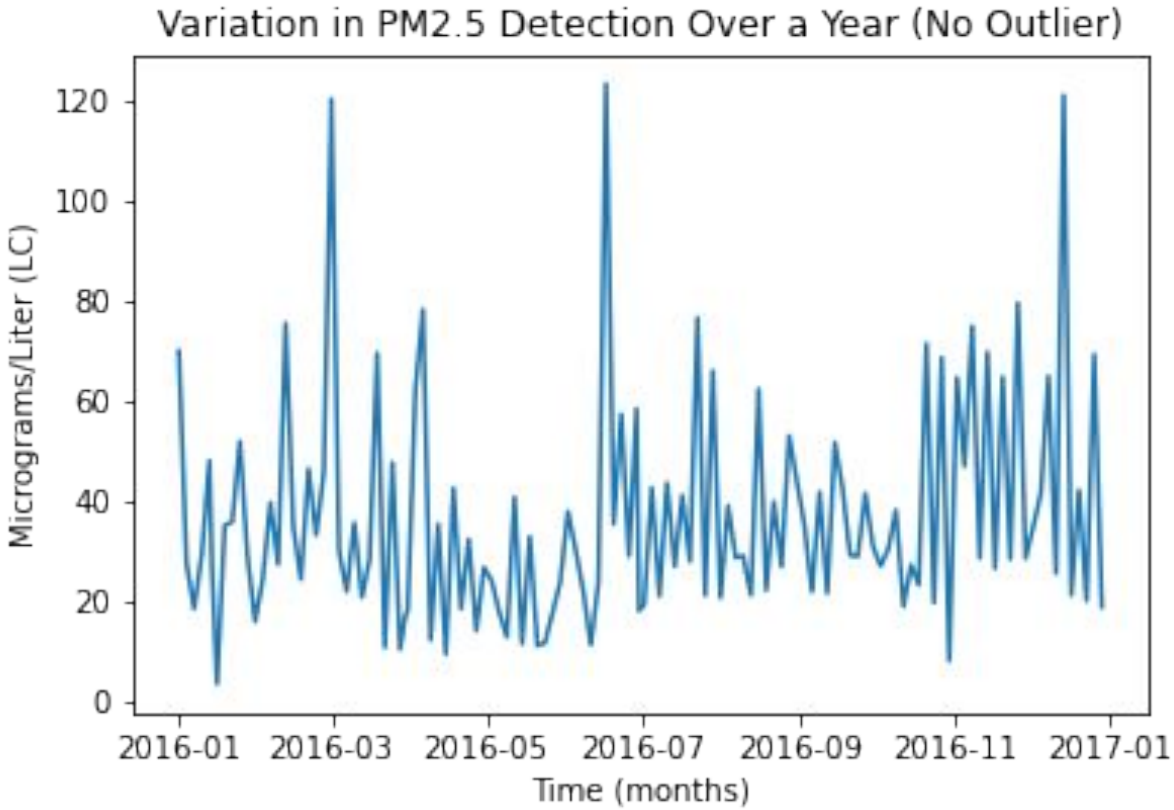
- I pulled out the first three particulate types and examined their trends



Initial Thoughts

- This data is highly variable depending on the time and type of particulate
- Rather than comparing species to species, it will be easier to aggregate all the PM2.5 levels together into one day
  - But first, I should remove the outlier species (the big orange spikes in the first graph)
    - Outlier species ended up being Total Nitrate measurements

- Here is a graph showing total amount of measured PM2.5 per day
- There are 3 clear spikes and I want to check out what days they are

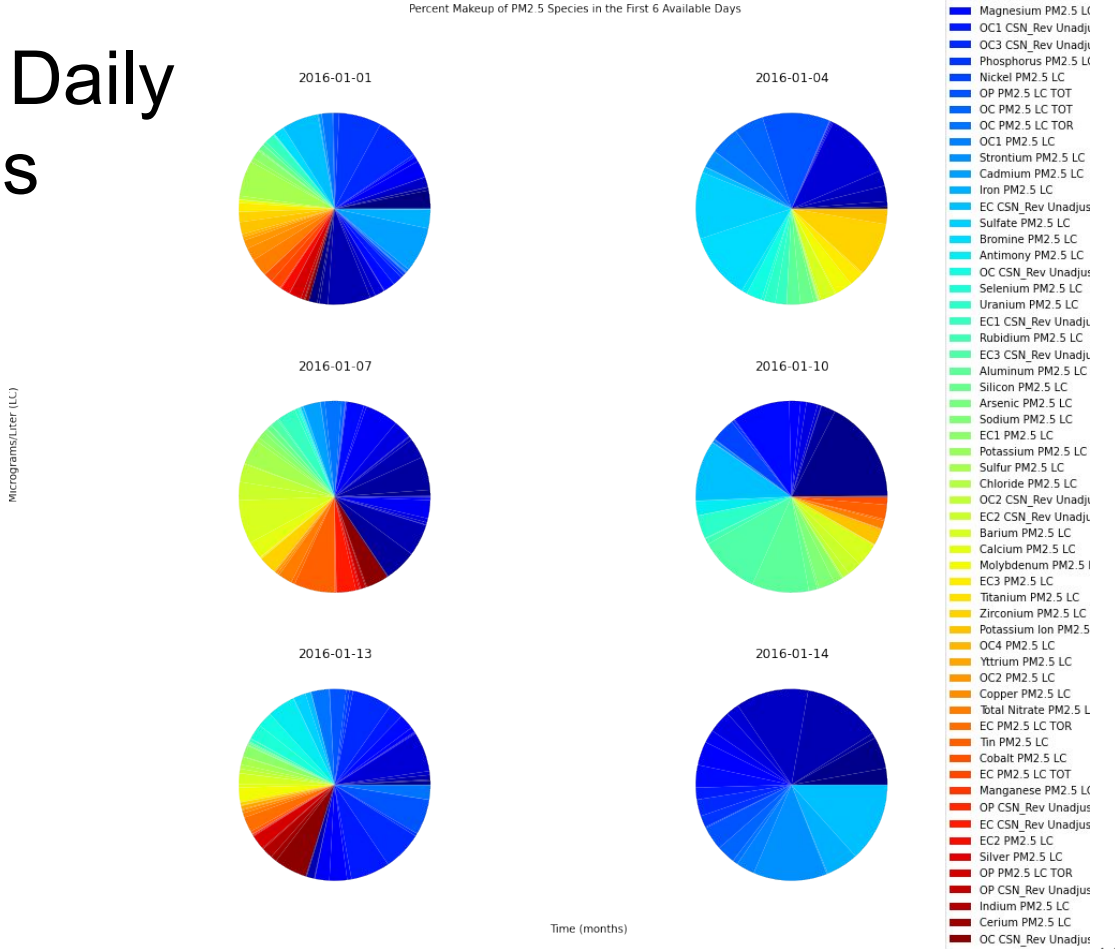


Top 3 Days With The Most PM2.5 Detection

- The three days with highest PM2.5 measurements in 2016 are, in order by PM2.5 measurements
  - 2016-06-17 with an aggregated 123.32181 LC
  - 2016-12-14 with an aggregated 121.04824 LC
  - 2016-03-01 with an aggregated 120.30023 LC
- Unfortunately, none of the three days seem to be special as far as I know
  - Some preliminary research reveals that there don’t seem to be any particularly special events happening on any of those 3 days in LA

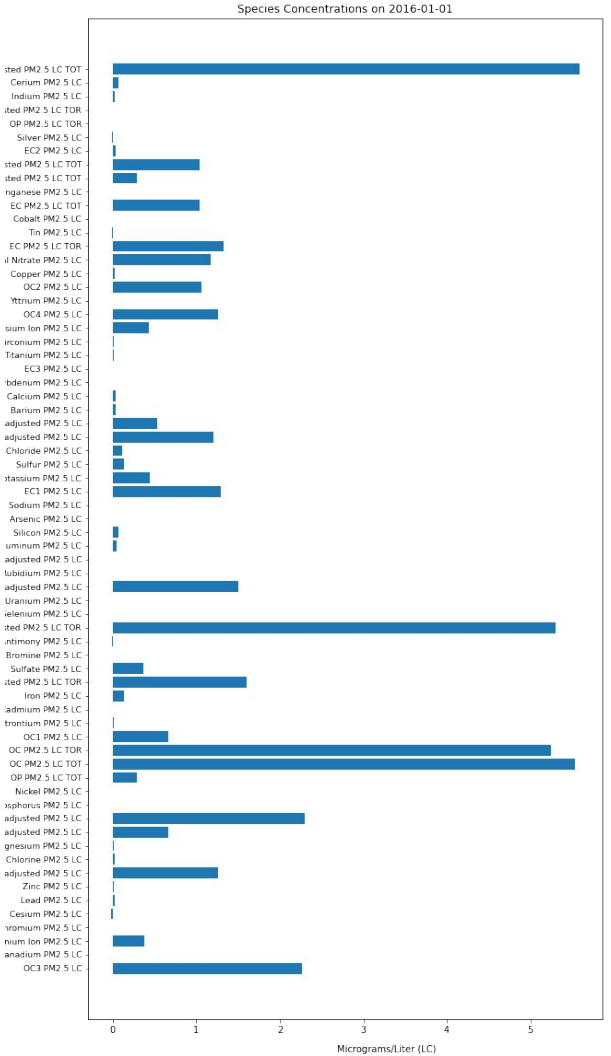
Let's See What the Daily Makeup Of PM2.5 Is

- I took data from the first 6 days and broke it down to see what the daily PM2.5 makeup would look like
- Unfortunately, with 67 parameters adding percentages overcrowds the graph



Putting Some Concrete Numbers Onto PM2.5 Makeup Per Day

- Here I've taken the first day and plotted out how much of each chemical was detected
- Interesting to note that there are a few negative numbers
  - Most likely due to standardization errors

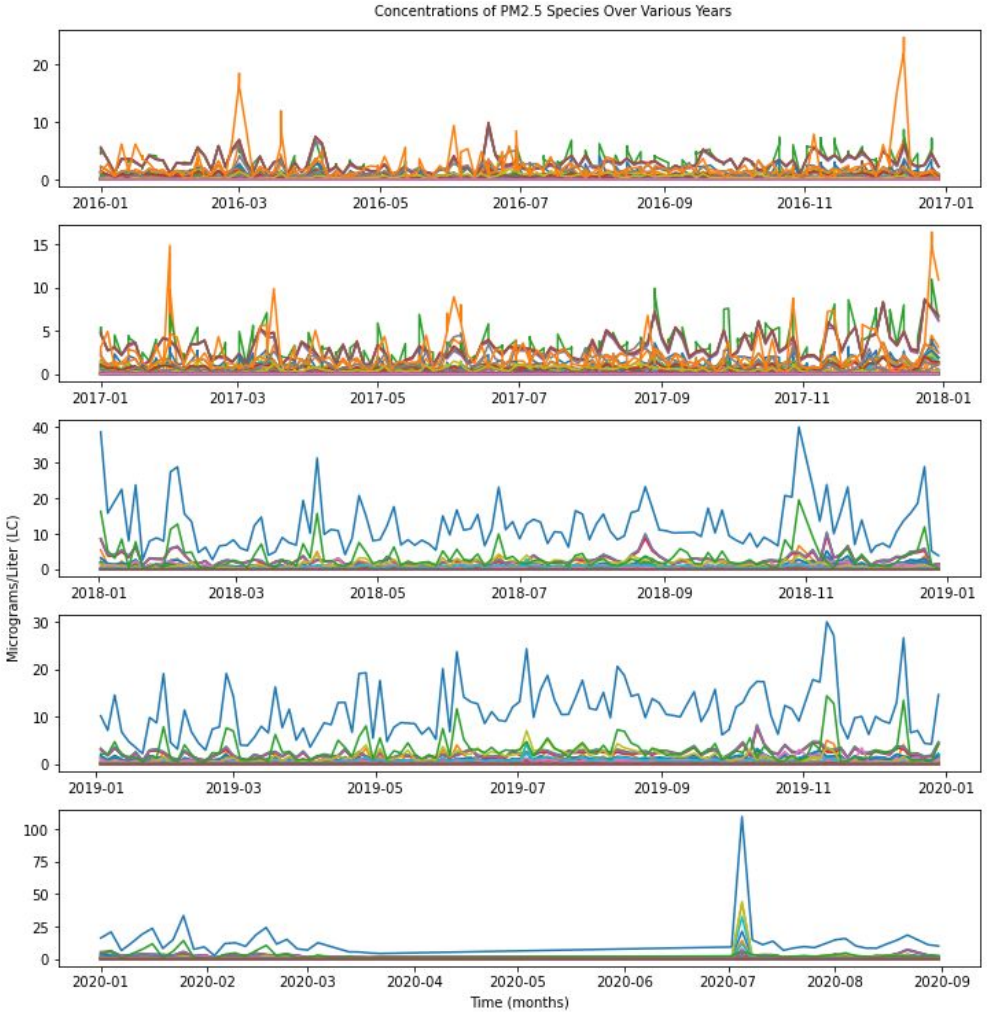


PM2.5 Daily Makeup Findings

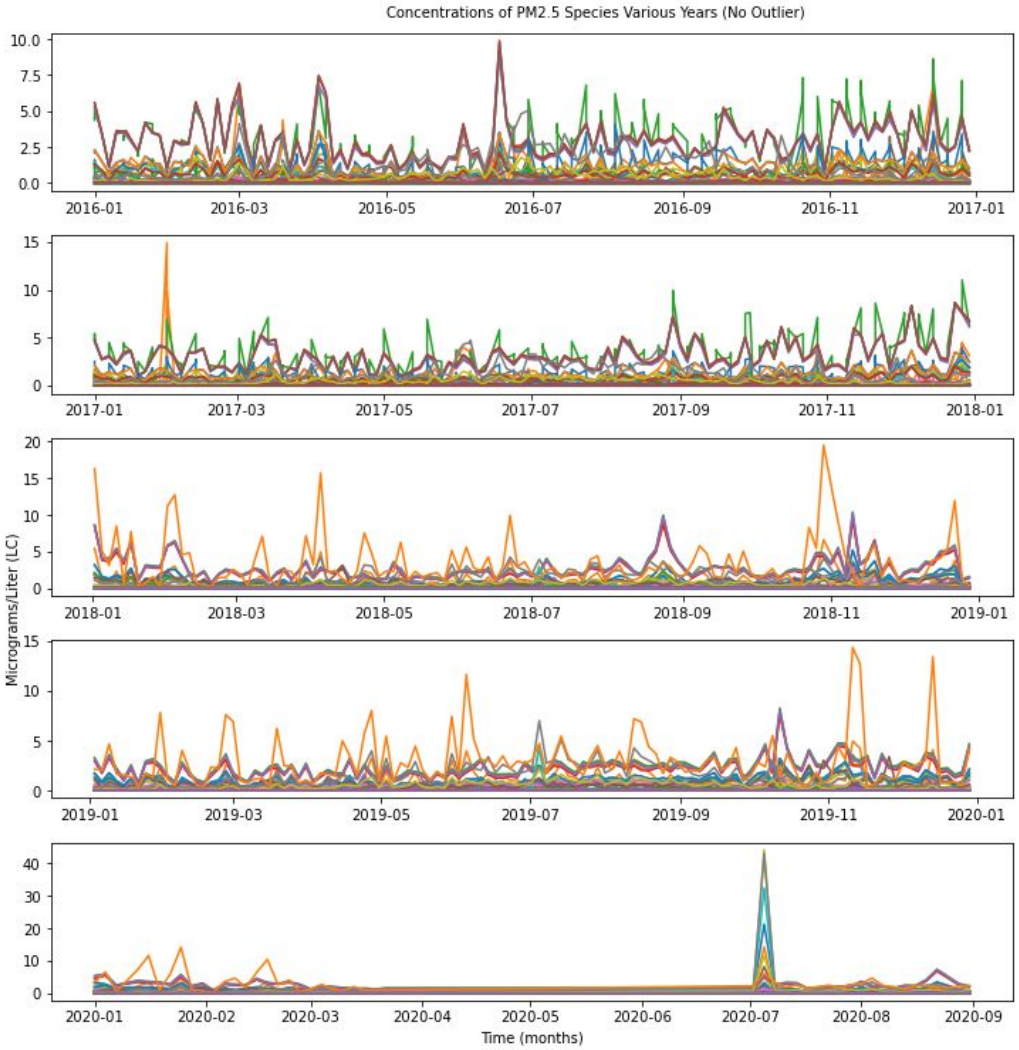
- Keep in mind that nitrate, as an outlier, was removed
- After nitrate, various types of carbon such as OC1, OC2, EC, etc. seem to be the most common PM2.5 species
  - OC stands for Organic Carbon, EC stands for Elemental Carbon
- Variation by day is also a lot
  - The 6th day, for example, has much fewer species that make up PM2.5 compared to the 3rd
- Check out the attached notebook for a more detailed versions of all the graphs, but especially the last 2

Now Let's Expand Out The Findings

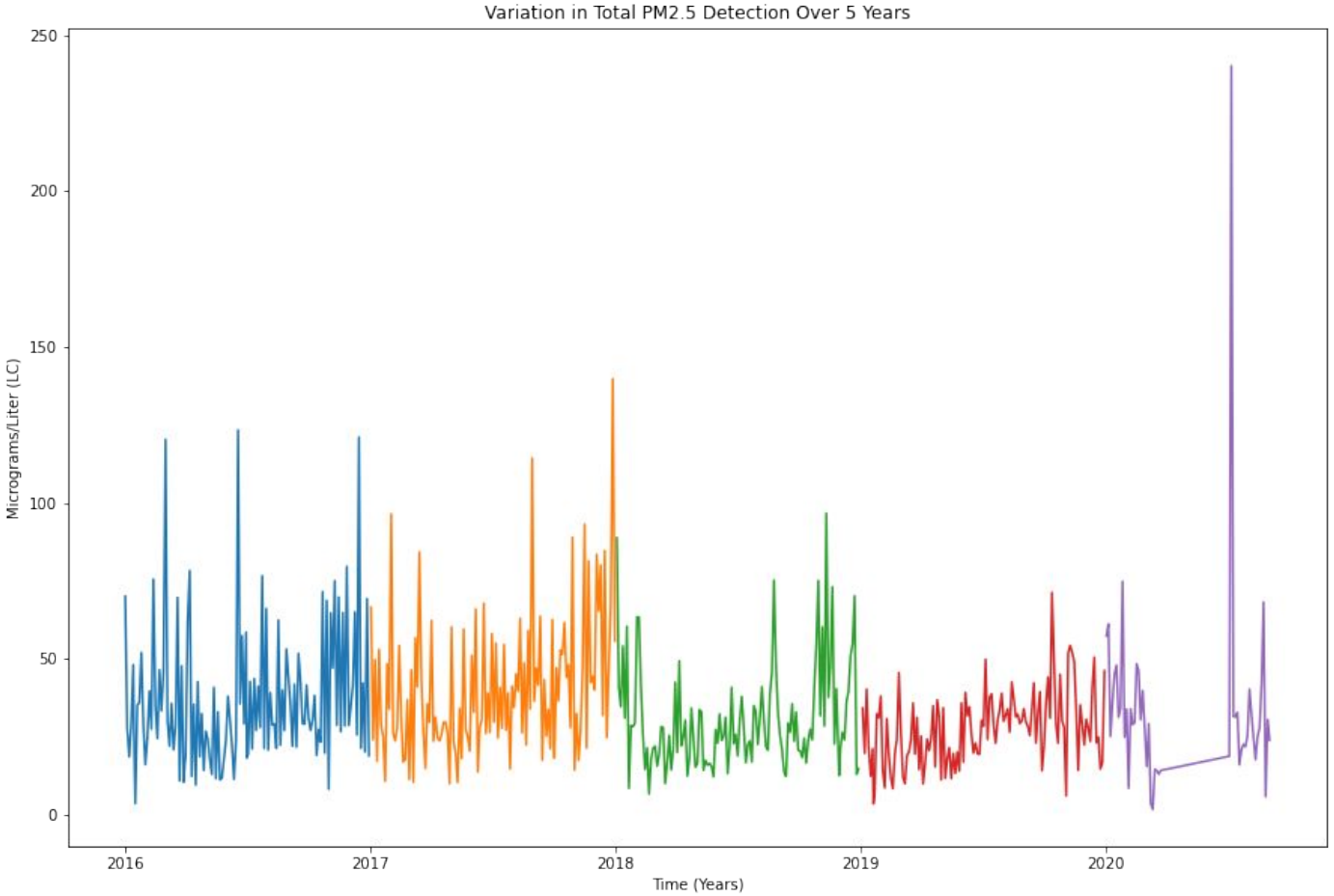
- Here is a graph showing each PM2.5 species' concentrations over 5 years
- There are clear outliers in each graph



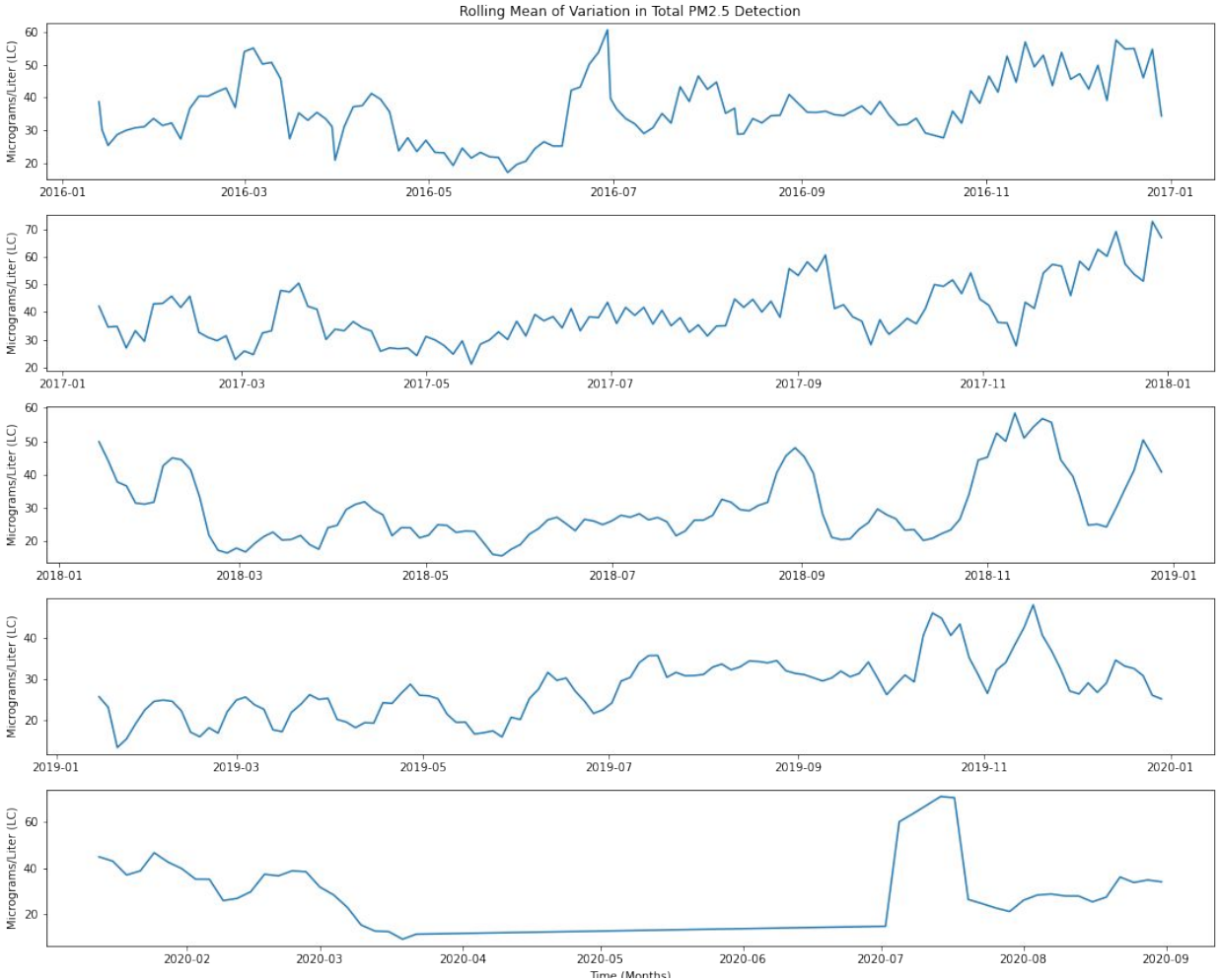
- Here are the plots after I removed outliers from each year (Total Nitrate in the first two years and Reconstructed Mass in the last 3)
- Still a bit noisy, so let's sum the measurements from each day into one number again



- This graph shows the summed the amount of PM2.5 recorded per species every day
- Notice a slight upward trend as the year progresses



- To smooth out the curves I applied a rolling mean to each year by combining every 5 days into one
- Now it's easier to easier to see trends

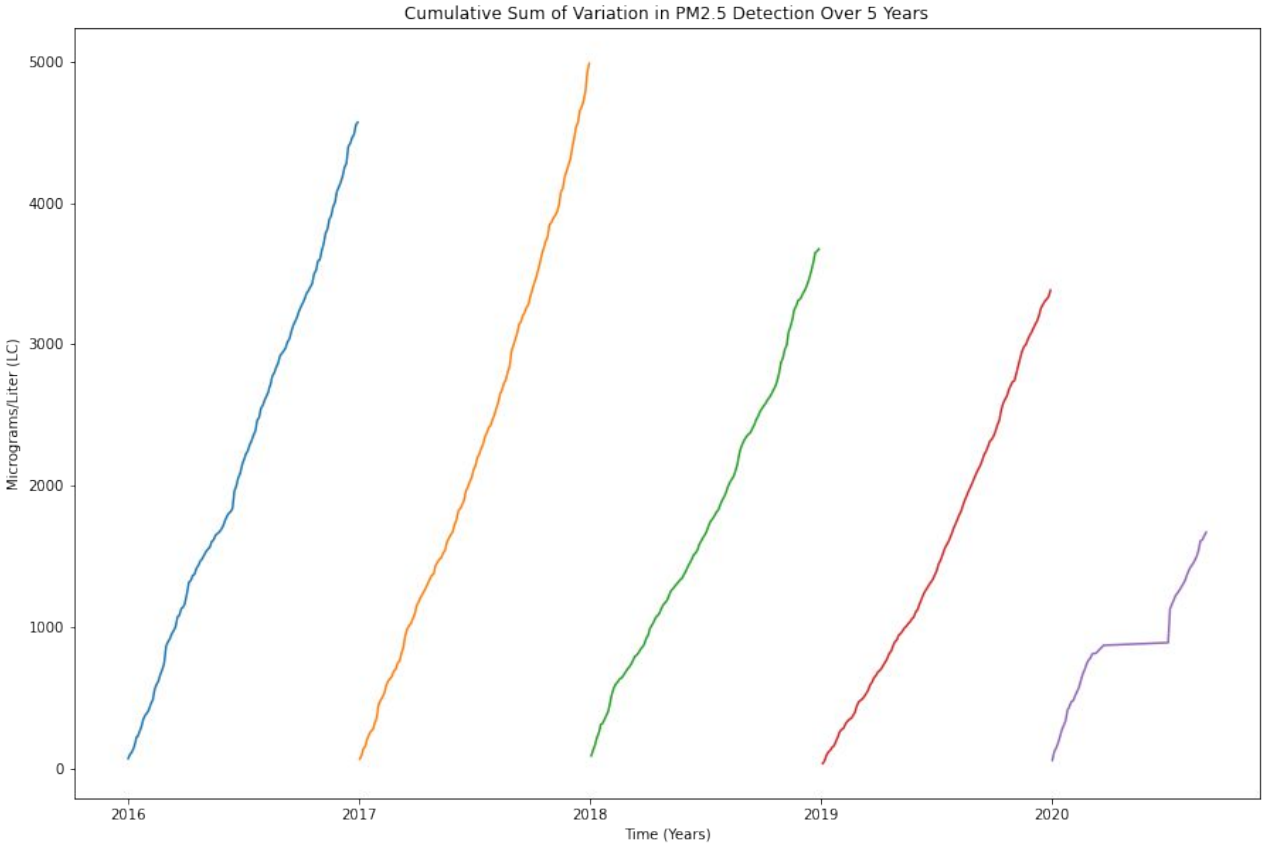


Thoughts On Yearly Trends

- 2020 has much lower levels of PM2.5
  - Probably due to COVID-19
  - The spike in July coincides with activity in support of the BLM movement in LA
- Despite having different y-axes windows, each year seems to follow a pattern where PM2.5 levels rise throughout the year
  - Usually bottoms out near May

Looking at Rate of Growth Per Year

- Finally, I applied a cumulative sum function onto each year and found that the rate of growth is roughly linear



Answers To My Initial Questions:

- Are there any particular days with high levels of PM2.5 in 2016?
  - Yes, they are March 1st, June 17th, and December 14
  - None of the days seem special, so I'm not sure why PM2.5 levels are so high
- What does the chemical makeup of PM2.5 in 2016 look like?
  - The chemical makeup varies a lot day by day
  - Nitrate tops the list, but by so much it is an outlier. Various forms of carbon are the next most common type of PM2.5 species
- What kind of trends do we see in PM2.5 from year to year?
  - Each year aside from 2020 sees lowest PM2.5 measurements around May before steadily rising again throughout the year
  - The rate of growth of PM2.5 each year is linear

Further Questions To Investigate And Next Steps

- Are there certain species that are representative of other species (i.e. increase and decrease together)? How can one go about finding these species?
- In general, is there a way to reduce the dimensions of the data without losing important trends and features?
- Can I compare variation in daily PM2.5 makeup daily throughout the years?
- How do certain species affect the health of humans, animals, and the environment? Arsenic, for example, is probably more dangerous than carbon.
- How do meteorological effects such as rain, temperature, heavy winds, UV radiation, etc.play a role in PM2.5 makeup?
- How do human activities such as travel, celebrations, and protests affect PM2.5 makeup?