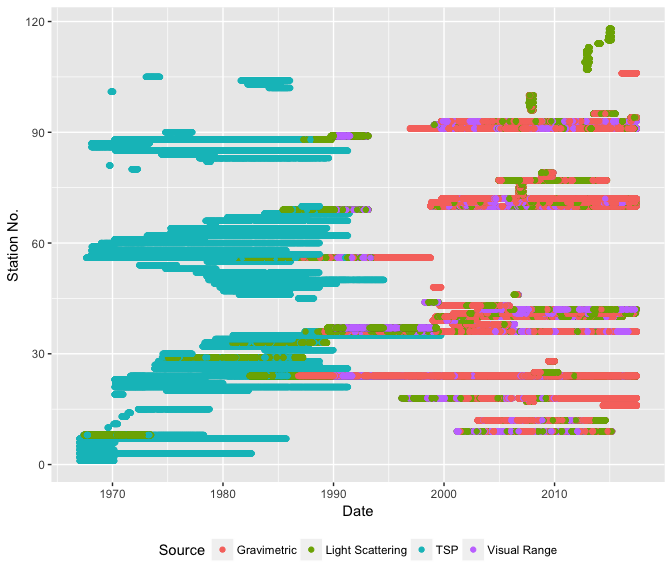
PM Data Exploration

Magali Blanco

10/9/17

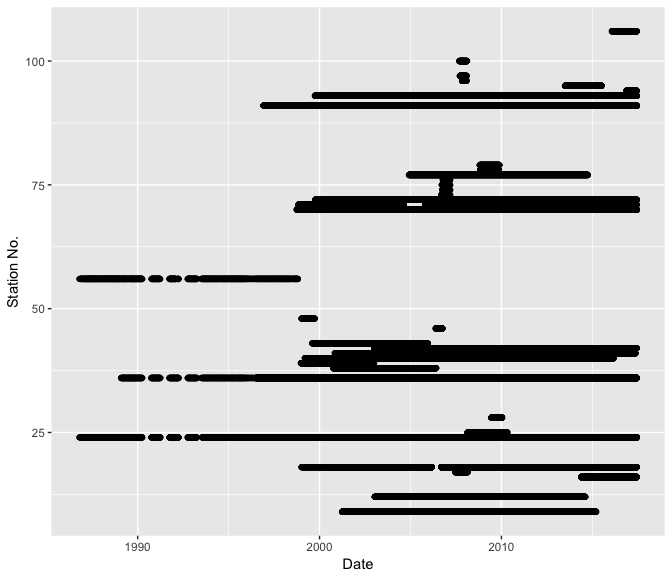
# Data Availability

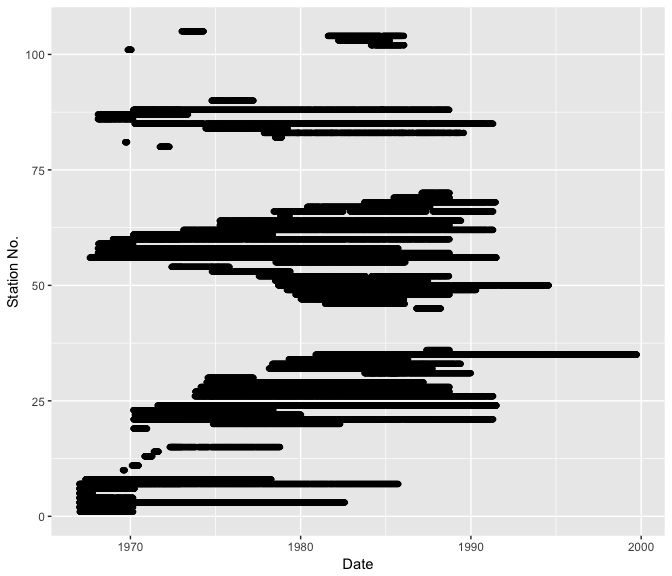
Overall data availability:  
\* Gravimetric PM2.5  
\* TSP  
\* Light Scattering  
\* Visual Range

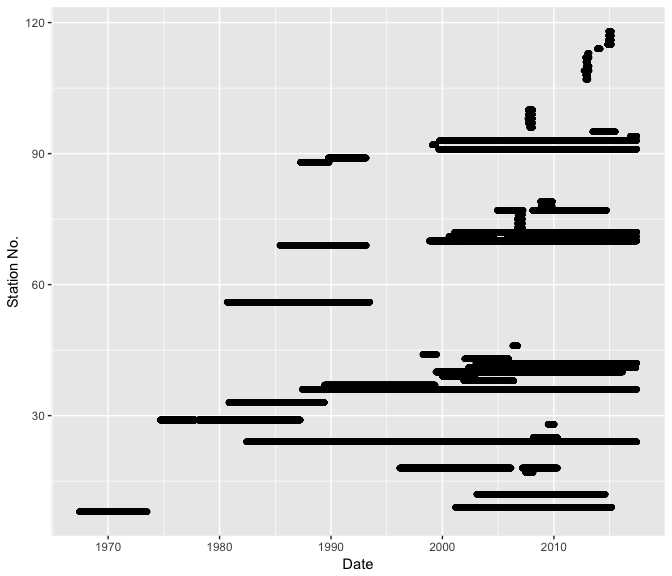


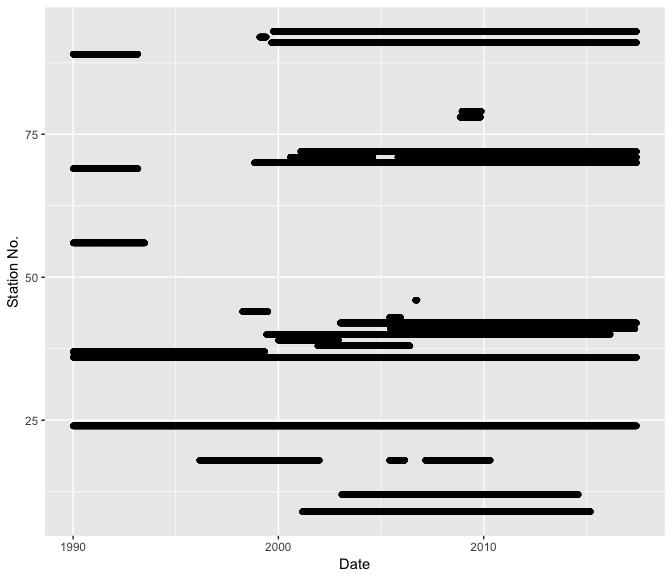
Overall Data Availability

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Start | End | Yrs | Smpl | Avg Smpl/Yr | Avg Smpl/Dy | No. Stations | No. QMUs |
| 1967-01-01 | 2017-06-12 | 50.5 | 545097 | 10794 | 30 | 118 | 38 |

Gravimetric data availability  


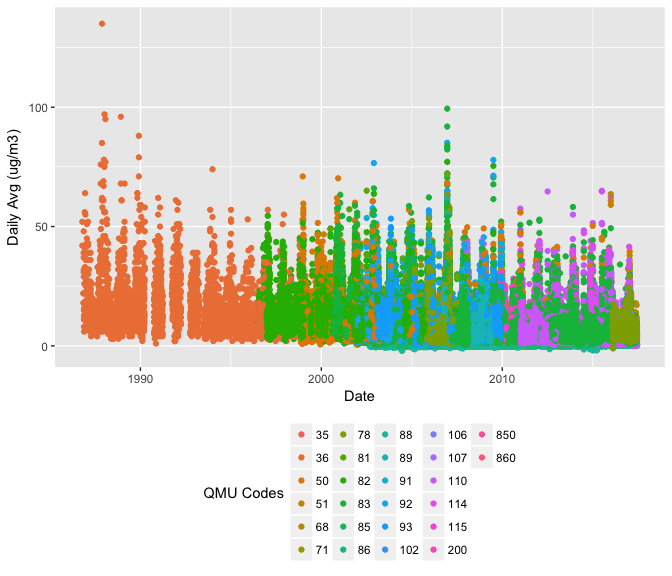
TSP data availability  


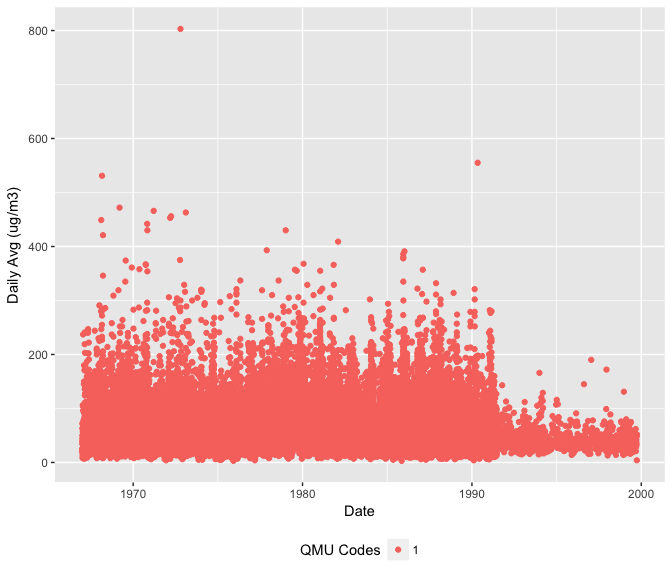
Light scattering data availability  


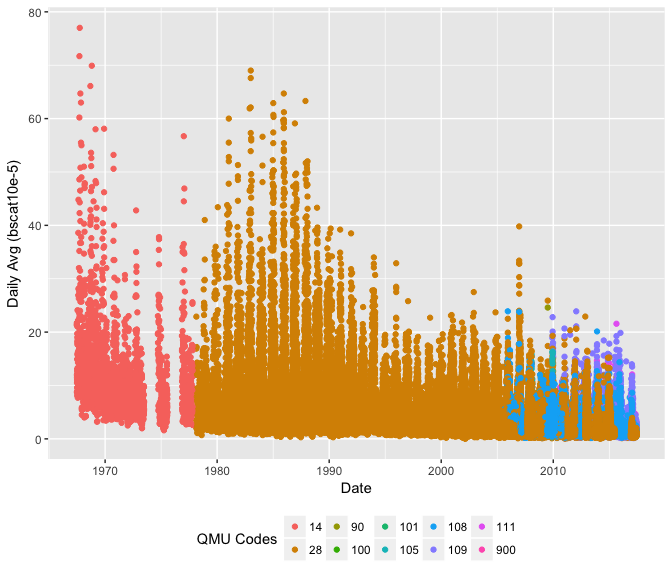
Visual range data availability  


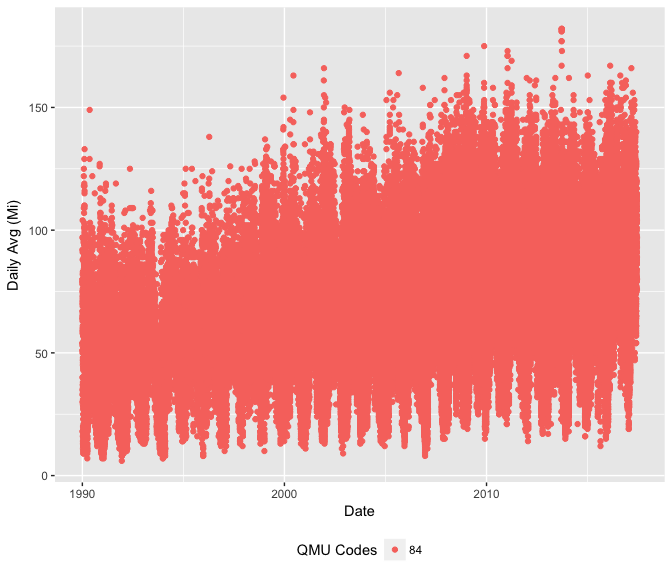
# PM Over Time by Source

Are they continuous?  
Do they increase/decrease over time? (make plots for time ranges?)  
Are there temporal trends?  
Are changes in methods associated with changes in PM?  
Zoom in to see temporal patterns??

Gravimetric Samples of PM2.5 Over Time  


TSP Over Time  


Light Scattering Over Time  


Visual Range Over Time  


Data Availability by Source

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Source | Start | End | YRS | Smpl | Avg Smpl/Yr | Avg Smpl/Dy | Smpl Freq | No. Stations | No. QMUs |
| Gravimetric | 1986-10-17 | 2017-06-12 | 30.7 | 258177 | 8410 | 23 | 1 | 36 | 26 |
| TSP | 1967-01-01 | 1999-09-30 | 32.8 | 43734 | 1333 | 4 | 3 | 71 | 1 |
| Light Scattering | 1967-06-01 | 2017-06-12 | 50.1 | 127109 | 2537 | 7 | 1 | 56 | 10 |
| Visual Range | 1990-01-01 | 2017-06-12 | 27.5 | 87832 | 3194 | 9 | 1 | 25 | 1 |

# Missing Data

????

Percent NAs in Datasets

|  |  |
| --- | --- |
| Source | NAs |
| Overall | 5% |
| Gravimetric | 7% |
| TSP | 0% |
| Visual Range | 4% |
| Light Scattering | 5% |

How often does each source sample PM? Daily? Every 3 days? 7 days?

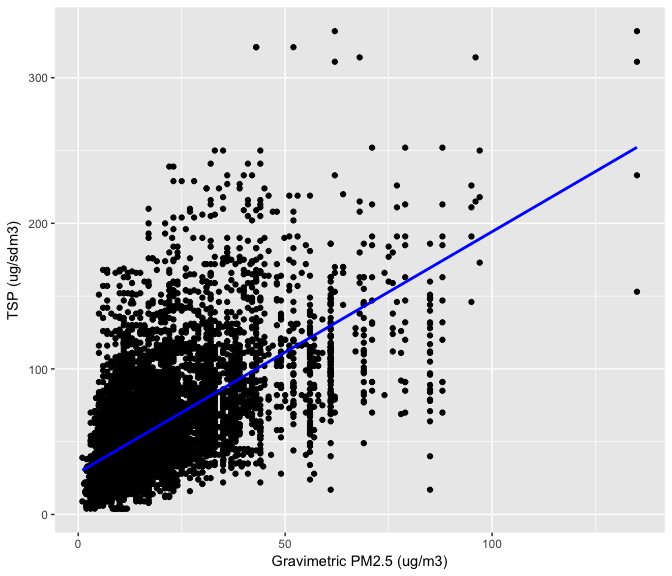
How much data should be here? for each station or mean # days/station: (observed days) / (unique stations) x (1 / x days sampling frequency)

??? What percent of data is missing, by year?

# Correlate Gravimetric reference data to ?? TSP ??

???by ObservedDate & StationCode

Compare to RH values. Are there RH ranges that are associated with poor gravimetric-TSP correlation?

Gravimetric PM2.5 vs TSP 

Gravimetric PM2.5 and TSP Linear Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 28.692503 | 0.4905033 | 58.49604 | 0 |
| Grav.Avg | 1.655792 | 0.0218249 | 75.86705 | 0 |

# Summary of concentration levels per method/instrument

How do these values compare to the method/instrument's detection limits?

Gravimetric: PM2.5