

Breath of Cancer

By: Alyss, Aruna, Daniel , and Michael

What did we look at:

- National AQI Trends
- Trends between Air Quality Index (AQI) and Population
- Relationship between AQI and Cancer Incidence

What is AQI?

- An index for reporting daily air quality
 - How clean or polluted your air is.
- Provides insight on health effects people may experience within a few hours or days after breathing polluted air.

Air Quality Index (AQI) Values	Levels of Health Concern	Colors
<i>When the AQI is in this range:</i>	<i>..air quality conditions are:</i>	<i>...as symbolized by this color:</i>
0 to 50	Good	Green
51 to 100	Moderate	Yellow
101 to 150	Unhealthy for Sensitive Groups	Orange
151 to 200	Unhealthy	Red
201 to 300	Very Unhealthy	Purple
301 to 500	Hazardous	Maroon

History of AQI in the United States

Prior to 1976:

- 55 cities used 14 different indices
- Each had their own cautionary messages
- Confusing!

In 1976:

- U.S. Clean Air Act required U.S. Environmental Protection Agency (EPA) to establish a national air quality index
- EPA established the Pollutant Standards Index (PSI)

From 1976 to 1998: EPA and U.S. cities used the PSI that covered:

- CO, CO₂, SO₂, NO₂, O₃

In 1999: EPA revised the Index

- Changed name to AQI
- Removed CO₂ as a pollutant
- Revisions to O₃ and PM standards
- New advantages: easy to communicate color-coded categories to public, uniformity among cities

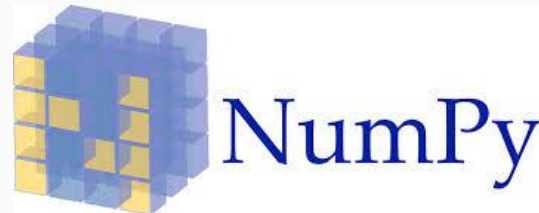
Why is looking at AQI important?

- There can be severe health risks from prolonged exposure to polluted air including:
 - Cardiovascular and respiratory aggravation
 - Increased risk of lung disease (including cancer)
 - Shortened life span



How did we look at our data?

- Pandas
- Numpy
- Matplotlib
- Plotly

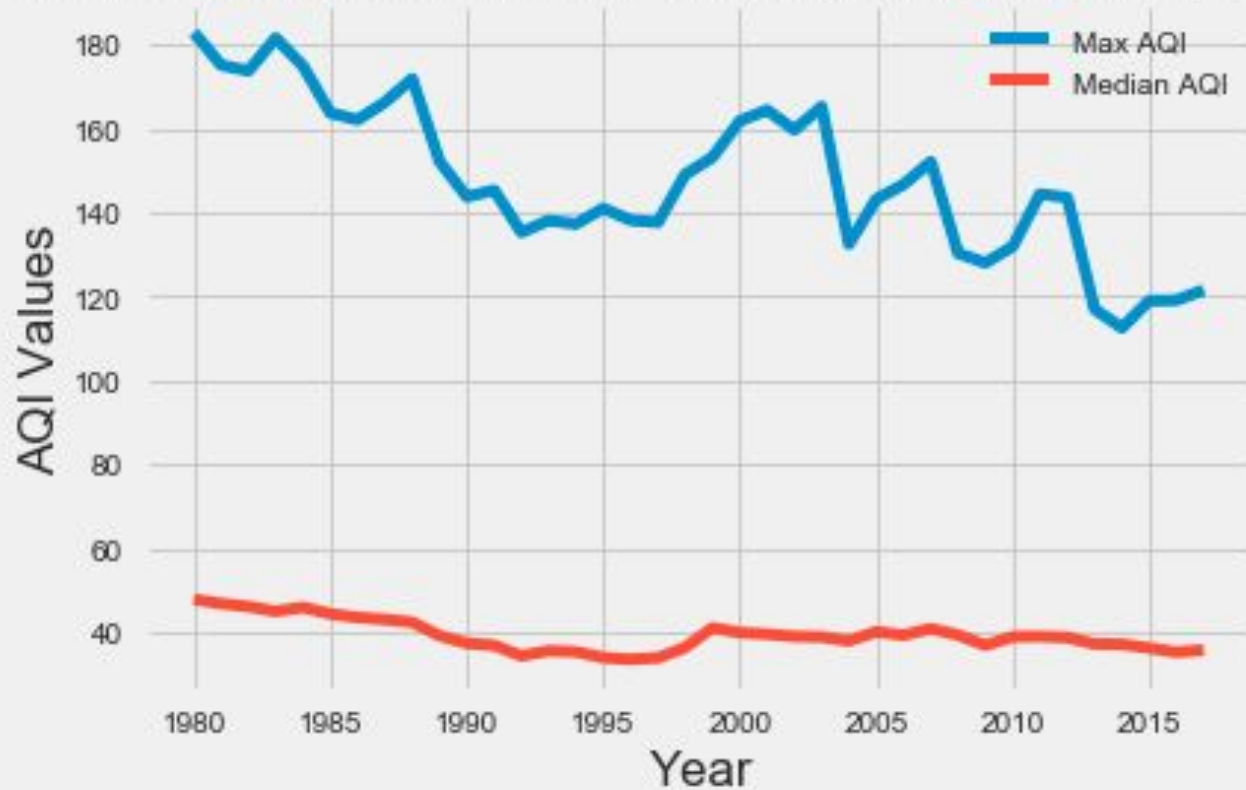


National AQI Trends

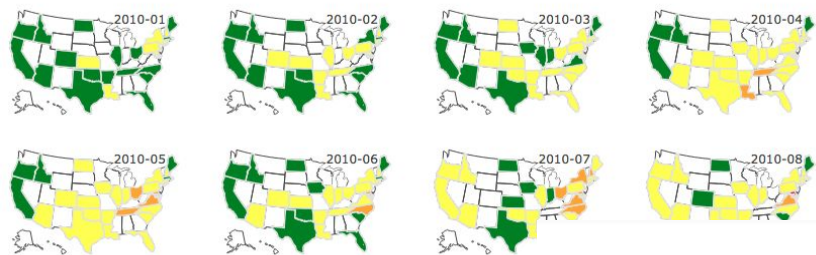
United States: Air Pollution AQI from 2000-2016



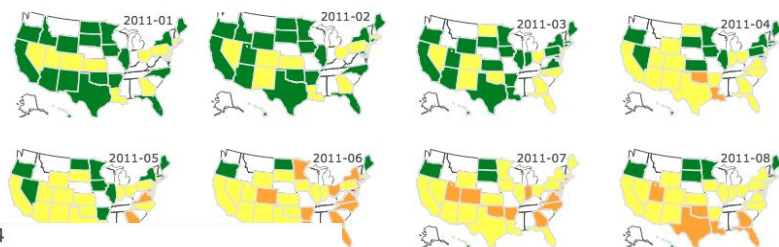
United States: Average Air Quality Index from 1980-2018



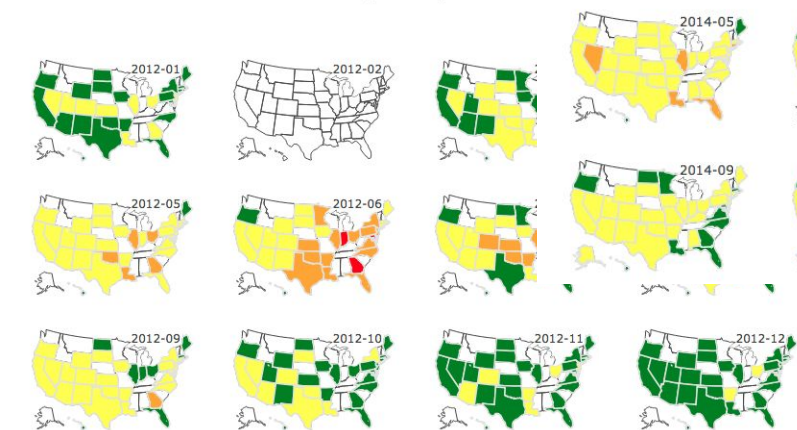
Monthly Max AQI in 2010



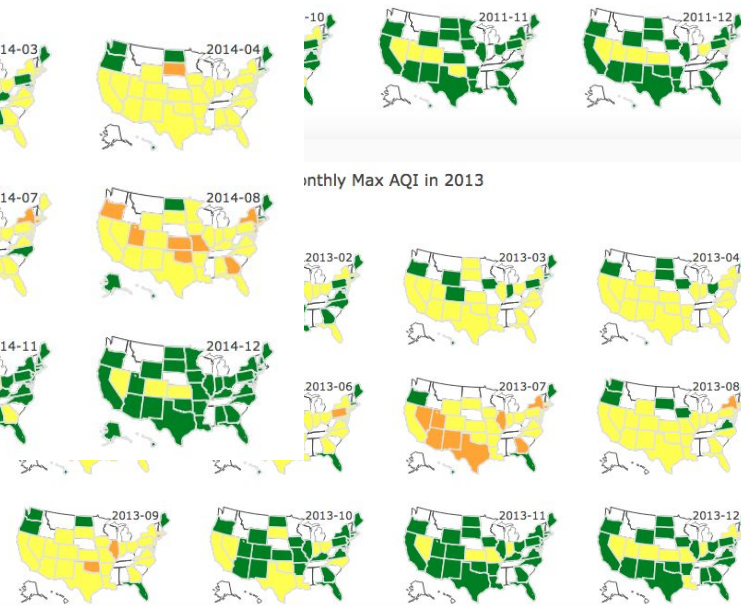
Monthly Max AQI in 2014



Monthly Max AQI in 2012



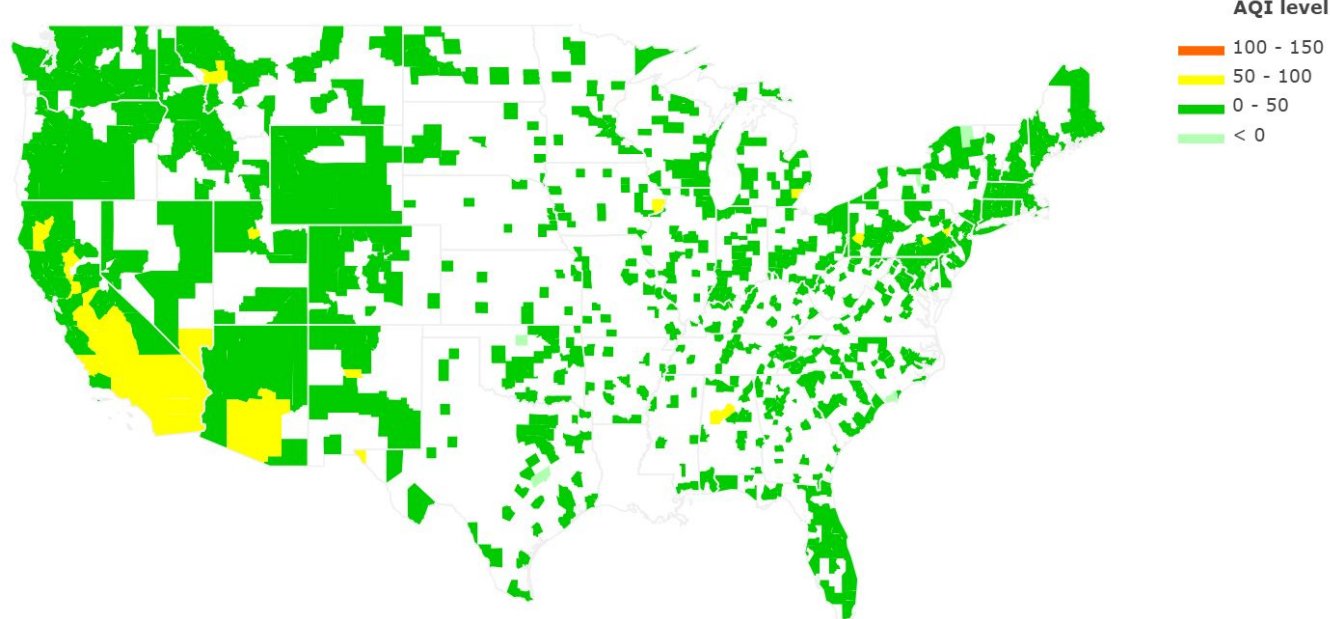
Monthly Max AQI in 2013



Seasonal Trends of Max AQI: 2010-2014

How did we do last year?

USA by Median Air Quality Index (AQI) [2017]



Levels of Health Concern

..air quality conditions are:

Good

Moderate

Unhealthy for Sensitive Groups

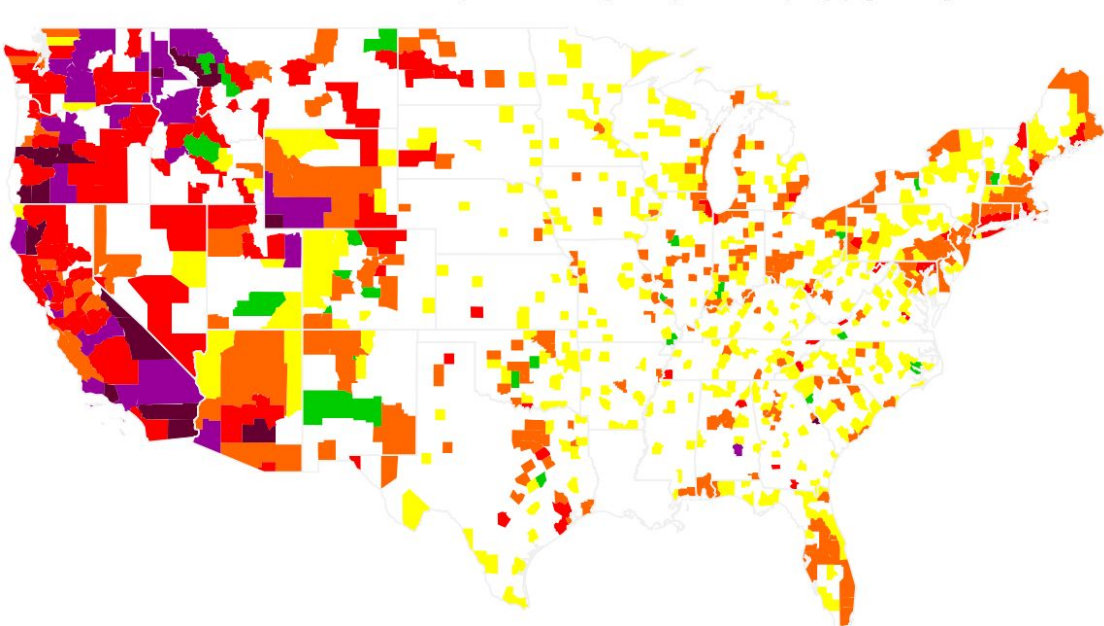
Unhealthy

Very Unhealthy

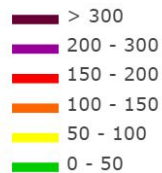
Hazardous

How did we do last year? (cont.)

USA by Max Air Quality Index (AQI) [2017]



AQI level



Levels of Health Concern

..air quality conditions are:

Good

Moderate

Unhealthy for Sensitive Groups

Unhealthy

Very Unhealthy

Hazardous

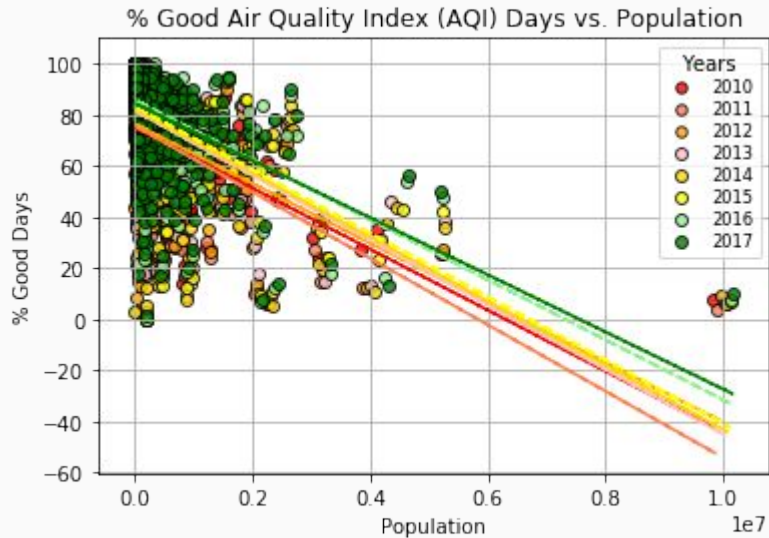
What did we see?

- Since 1980, the quality of the air in the United States is increasing!
 - Recall that lower AQI = better air
- Air quality is seasonal!
 - Better in the winter
- West coast has really bad max AQI values last year.
 - Why?

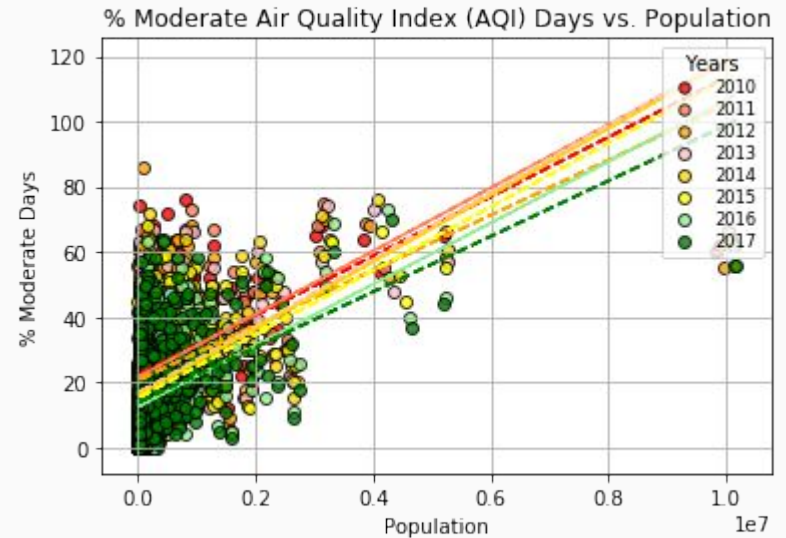


Population & AQI Trends

Effects of Population on AQI (by county)



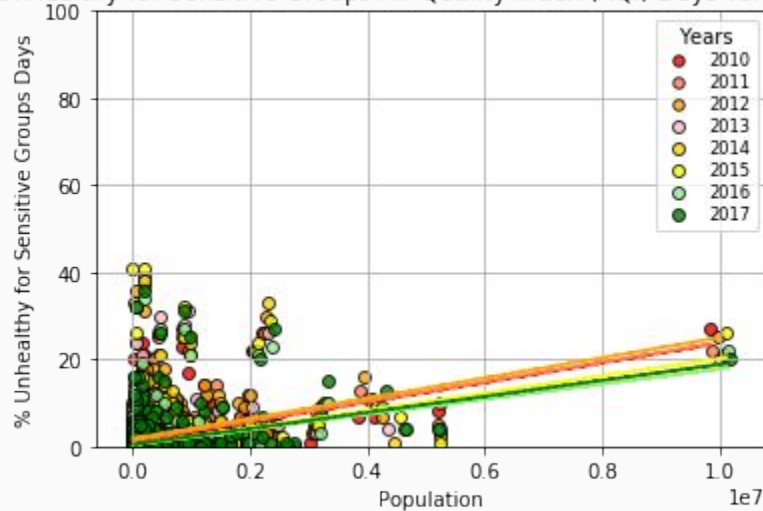
$$r^2 = 0.173$$



$$r^2 = 0.146$$

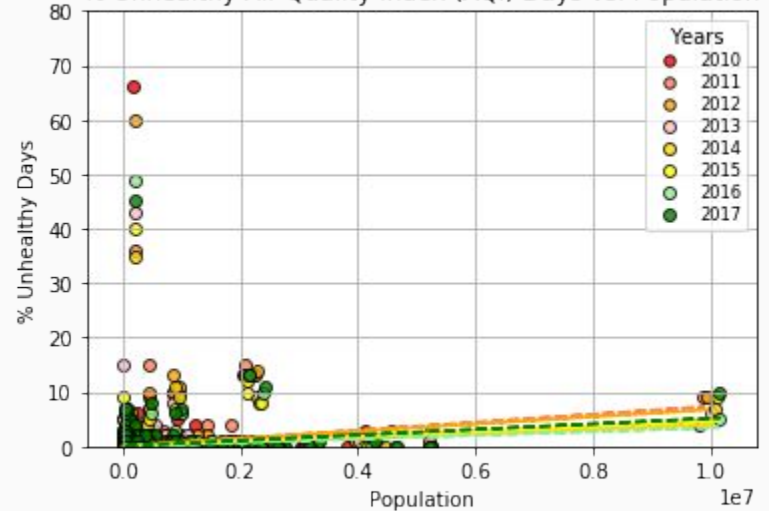
Effects of Population on AQI (by county... continued)

% Unhealthy for Sensitive Groups Air Quality Index (AQI) Days vs. Population



$r^2 = 0.111$

% Unhealthy Air Quality Index (AQI) Days vs. Population



$r^2 = 0.026$

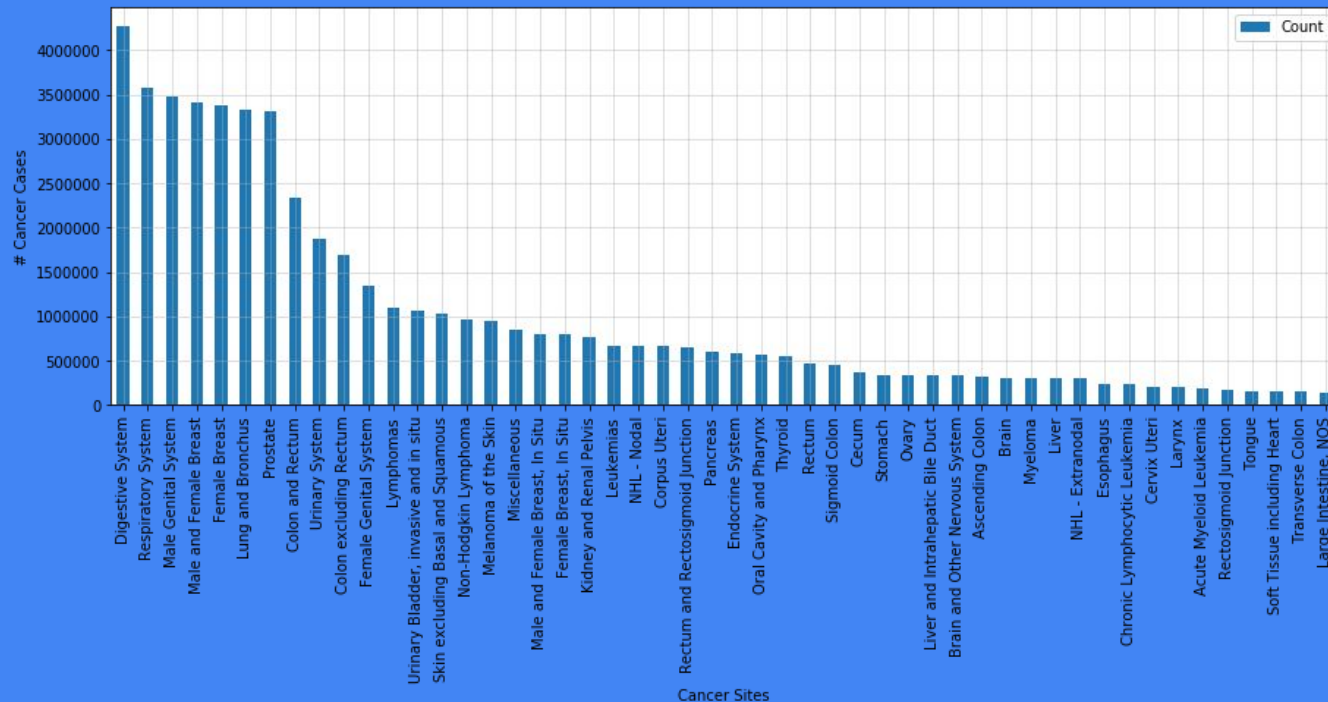
What did we see?

- As population increases, the less likely a county is to have AQI values that are considered “Good”.
 - Moderate days increase.
 - Unhealthy days increase.

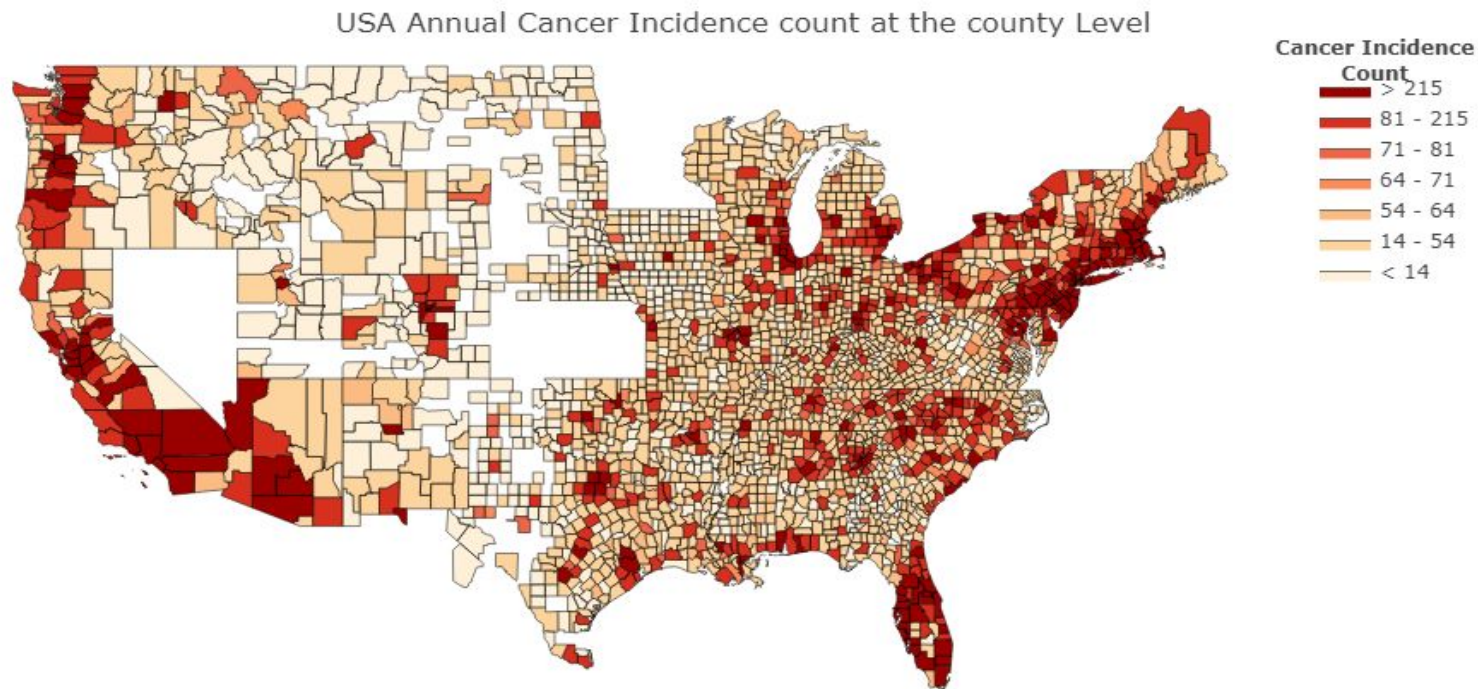


Cancer...

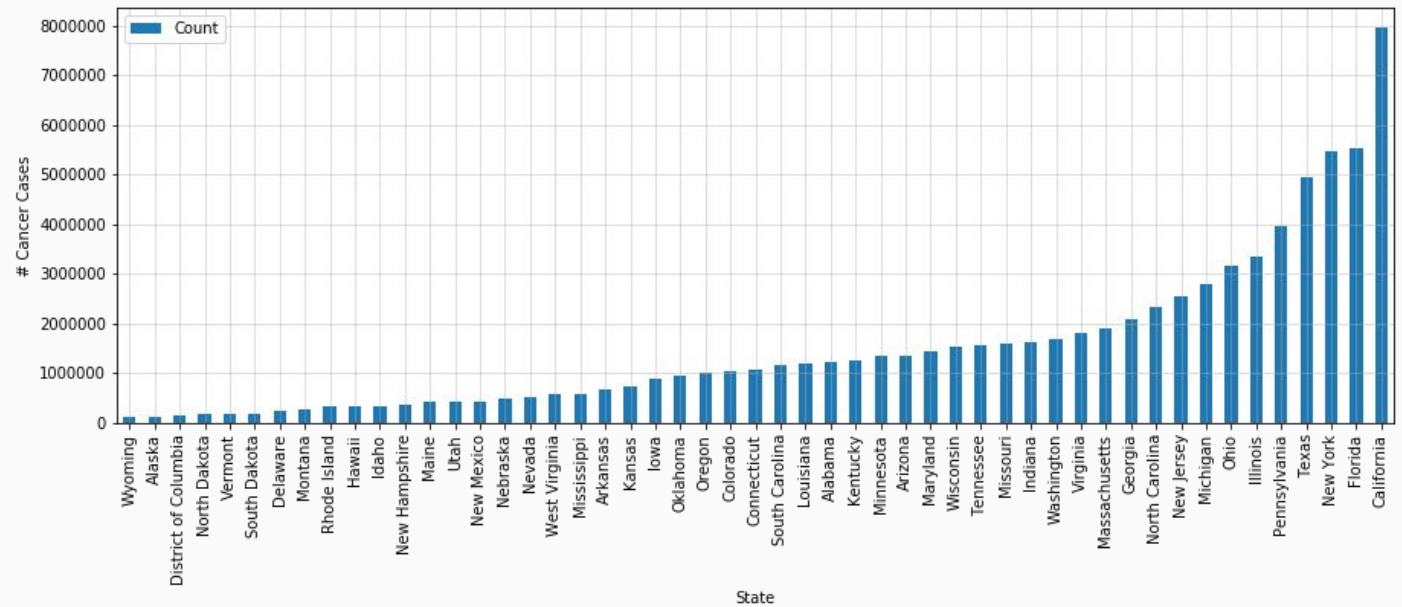
50 Most Common Cancer Types between 1998 to 2014



Cancer Incidences (by county)

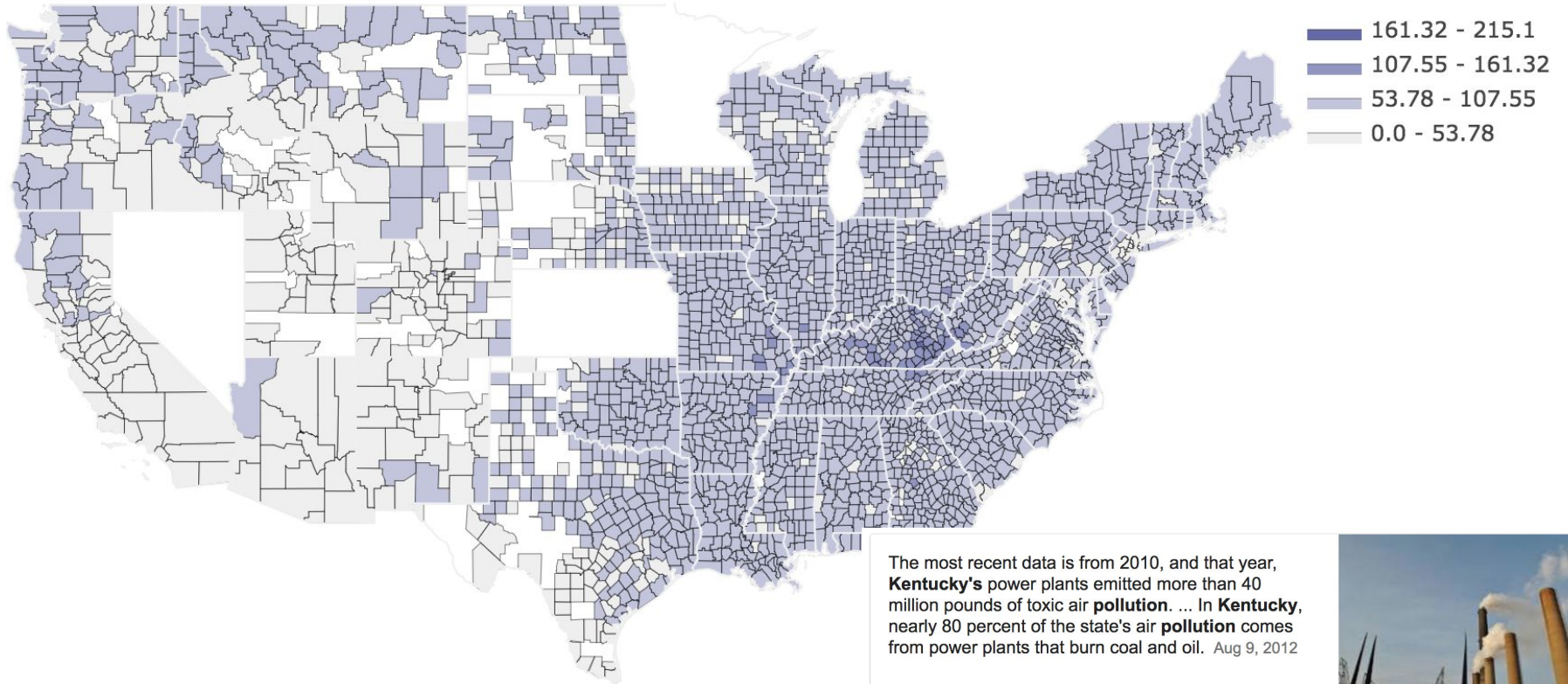


Total Cancer Counts by State from 1998 to 2014



Age-Adjusted Cancer Incidence Rate by County (2011-2015)

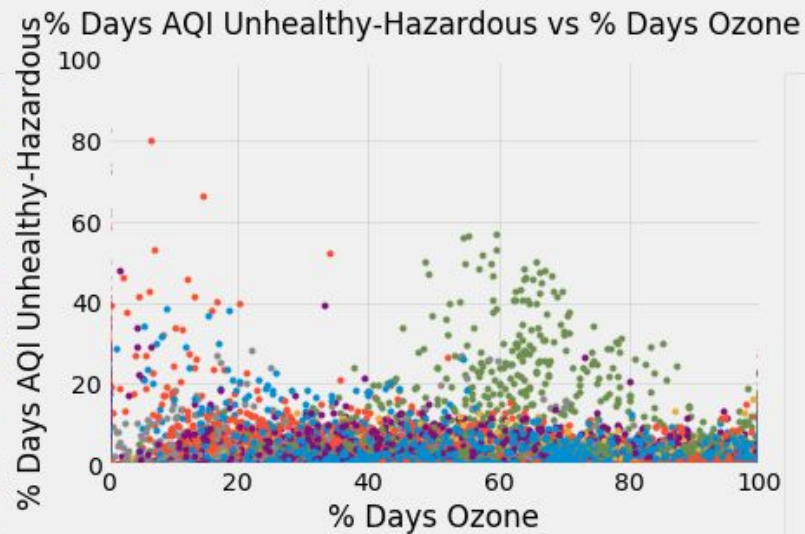
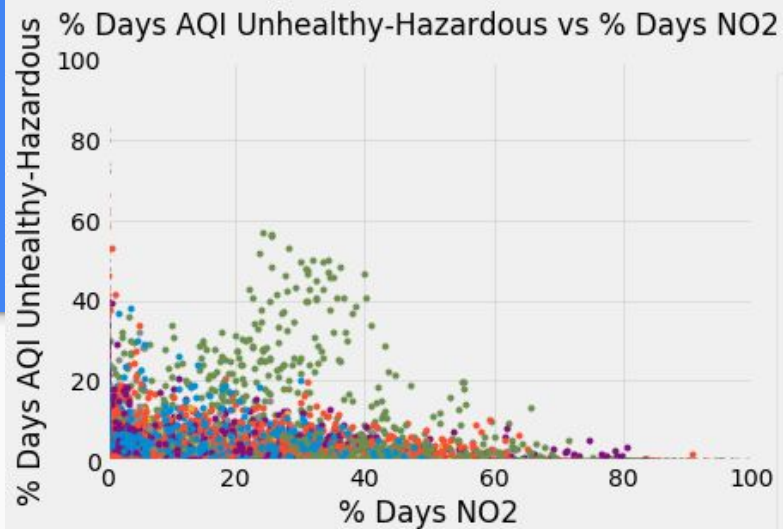
Age-Adjusted Incidence Rate



The most recent data is from 2010, and that year, **Kentucky's** power plants emitted more than 40 million pounds of toxic air **pollution**. ... In **Kentucky**, nearly 80 percent of the state's air **pollution** comes from power plants that burn coal and oil. Aug 9, 2012



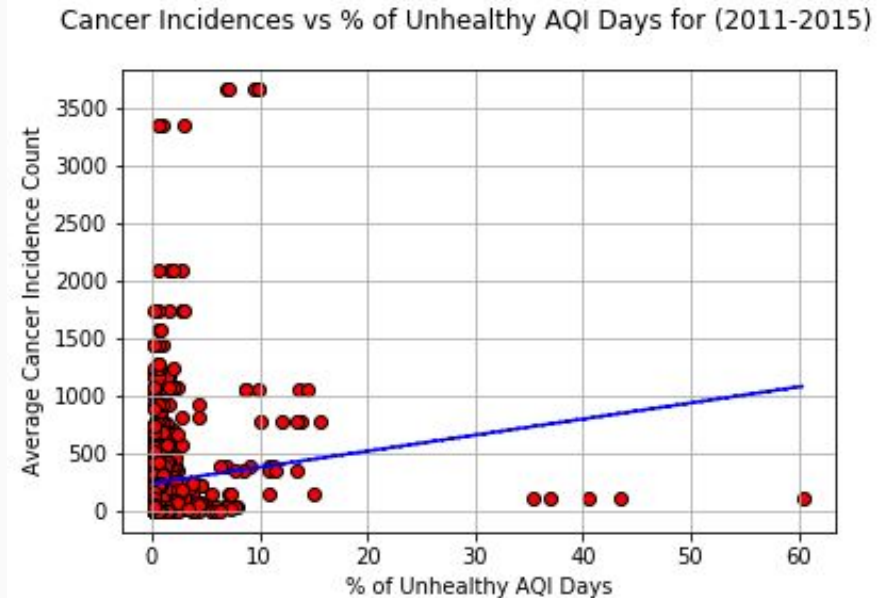
[Kentucky Ranks First in Toxic Air Pollution from Power Plants | 89.3 ...](http://wfpl.org/kentucky-ranks-first-toxic-air-pollution-power-plants/)
wfpl.org/kentucky-ranks-first-toxic-air-pollution-power-plants/



- alabama
- arizona
- arkansas
- california
- colorado
- connecticut
- delaware
- district of columbia
- florida
- georgia
- hawaii
- idaho
- illinois
- indiana
- iowa
- kansas
- kentucky
- louisiana
- maine
- maryland
- massachusetts
- michigan
- minnesota
- mississippi
- missouri
- montana
- nebraska

Cancer Incidence vs AQI

- It's well known that there is a correlation between rising air pollution and rising cancer incidence rates.
- Our data supports this correlation.



```
# Fitted linear regression model
```

```
variables = ['state', 'year', 'cancer_type', 'aqi_90th', 'pcnt_days_co',  
            'pcnt_days_no2', 'pcnt_days_so2', 'pcnt_days_ozone', 'pcnt_days_pm2pt5',  
            'pcnt_days_pm10']
```

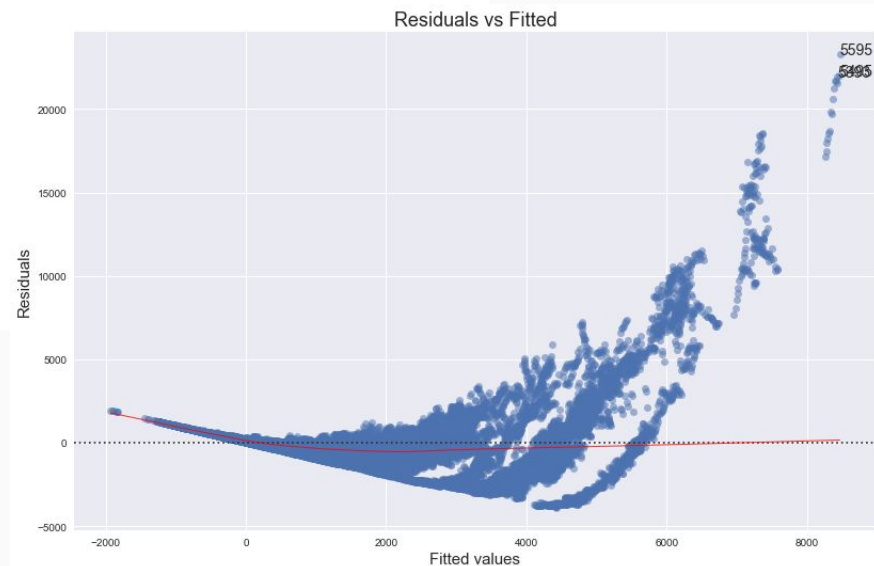
```
model_f = 'cancer_counts ~ {}'.format(' + '.join(variables))
```

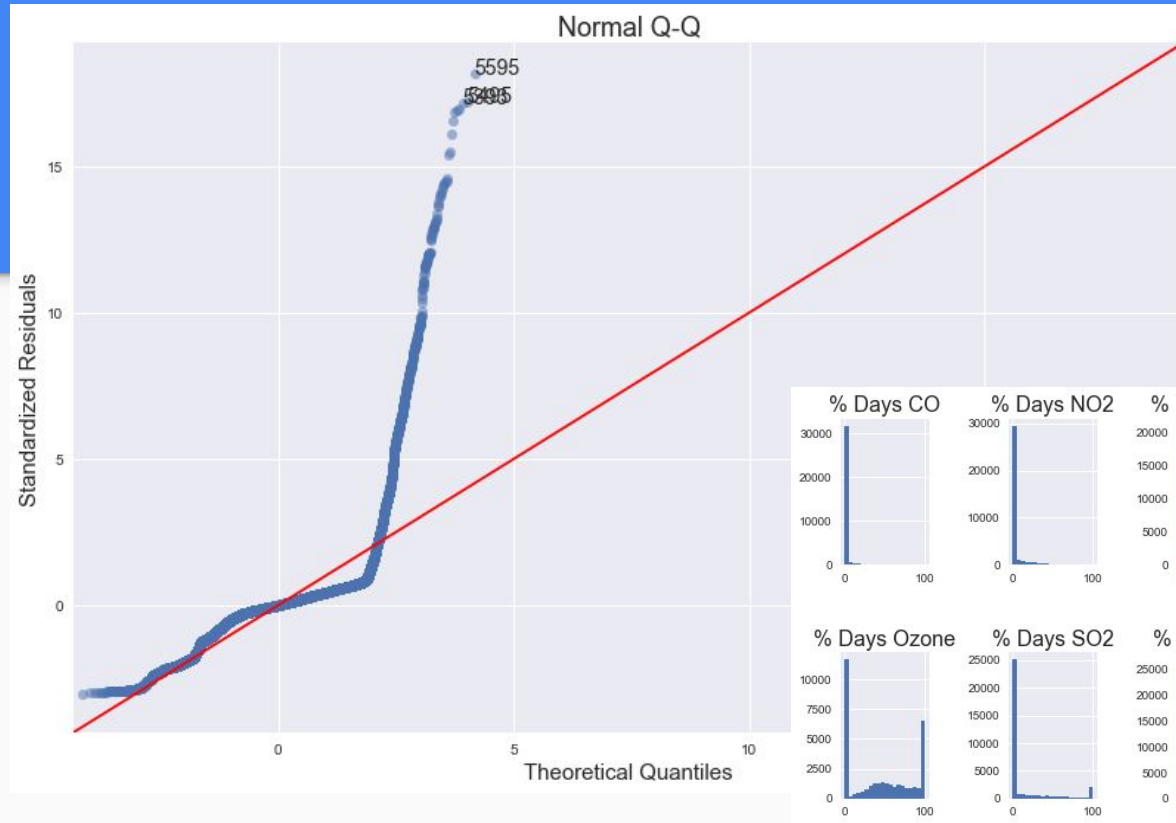
```
model = smf.ols(formula=model_f, data=cancer_types)
```

```
model_fit = model.fit()
```

```
model_fit.summary2()
```

Model:	OLS	Adj. R-squared:	0.575
Dependent Variable:	cancer_counts	AIC:	1131592.3665
Date:	2018-06-13 20:11	BIC:	1133038.7824
No. Observations:	65970	Log-Likelihood:	-5.6564e+05
Df Model:	158	F-statistic:	565.5
Df Residuals:	65811	Prob (F-statistic):	0.00
R-squared:	0.576	Scale:	1.6443e+06





QQ plot shows how well the distribution of residuals fit the normal distribution

Fin.

