

# Mortality



Ethel Perez  
Lauren Ennesser  
Kirsten Casey

**What socioeconomic and environmental factors affect mortality? By how much?**

# The Data

- Our data came from three major sources:
  - Centers for Disease Control and Prevention (CDC) – mortality and population numbers
  - United States Census Bureau – poverty numbers
  - Environmental Protection Agency (EPA) – air quality data
- Data were broken down by county into three 5-year intervals from 2002–2016
- CDC data also contains basic demographic information:
  - Race
  - Sex
  - Hispanic origin
  - Age group

County	Age Group Code	Gender Code	Race	Hispanic Origin	Deaths	Population	State	Unreliable	FIPS	5_Year_Avg_Poverty_Estimate	Interval	CO 2nd Max 8-hr	NO2 Mean 1-hr	Ozone 4th Max 8-hr	SO2 Mean 1-hr	PM2.5 Weighted Mean 24-hr	Lead Max 3-Mo Avg
Autauga County, AL	25-34	F	White	Not Hispanic or Latino	10.0	11763.0	AL	True	01001	5090.6	2002-2006	NaN	NaN	NaN	NaN	NaN	NaN
Autauga County, AL	35-44	F	White	Not Hispanic or Latino	25.0	16511.0	AL	False	01001	5090.6	2002-2006	NaN	NaN	NaN	NaN	NaN	NaN
Baldwin County, AL	15-19	F	White	Not Hispanic or Latino	17.0	20477.0	AL	True	01003	16183.6	2002-2006	NaN	NaN	10.000000	NaN	-1.944444	NaN
Baldwin County, AL	20-24	F	White	Not Hispanic or Latino	14.0	16460.0	AL	True	01003	16183.6	2002-2006	NaN	NaN	10.000000	NaN	-1.944444	NaN
Baldwin County, AL	25-34	F	White	Not Hispanic or Latino	32.0	37370.0	AL	False	01003	16183.6	2002-2006	NaN	NaN	10.000000	NaN	-1.944444	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Rock County, WI	65-74	F	Black or African American	Not Hispanic or Latino	22.0	969.0	WI	False	55105	19918.8	2007-2011	NaN	NaN	-1.714286	NaN	NaN	NaN
Rock County, WI	75-84	F	Black or African American	Not Hispanic or Latino	29.0	584.0	WI	False	55105	19918.8	2007-2011	NaN	NaN	-1.714286	NaN	NaN	NaN
Rock County, WI	85+	F	Black or African American	Not Hispanic or Latino	38.0	261.0	WI	False	55105	19918.8	2007-2011	NaN	NaN	-1.714286	NaN	NaN	NaN
Waukesha County, WI	75-84	F	Black or African American	Not Hispanic or Latino	13.0	191.0	WI	True	55133	18804.0	2007-2011	NaN	NaN	-7.142857	NaN	5.000000	NaN
Waukesha County, WI	85+	F	Black or African American	Not Hispanic or Latino	11.0	46.0	WI	True	55133	18804.0	2007-2011	NaN	NaN	-7.142857	NaN	5.000000	NaN
401389 rows x 39 columns																	

```

1 population_parameters = {'Race': 'White', 'Gender Code': 'F', 'Age Group Code': '85+', 'Interval': '2012-2016'}
2 df = flatten_df(mort, population_parameters)
3 df

```

Define Population Parameters

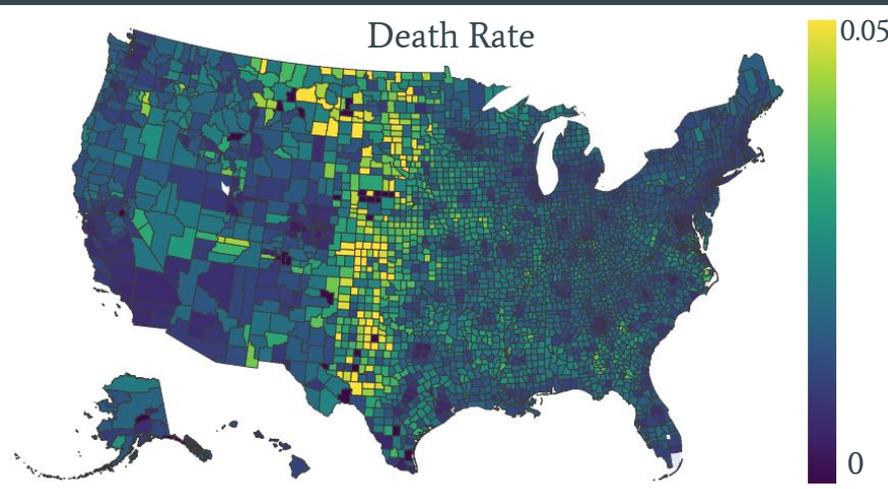
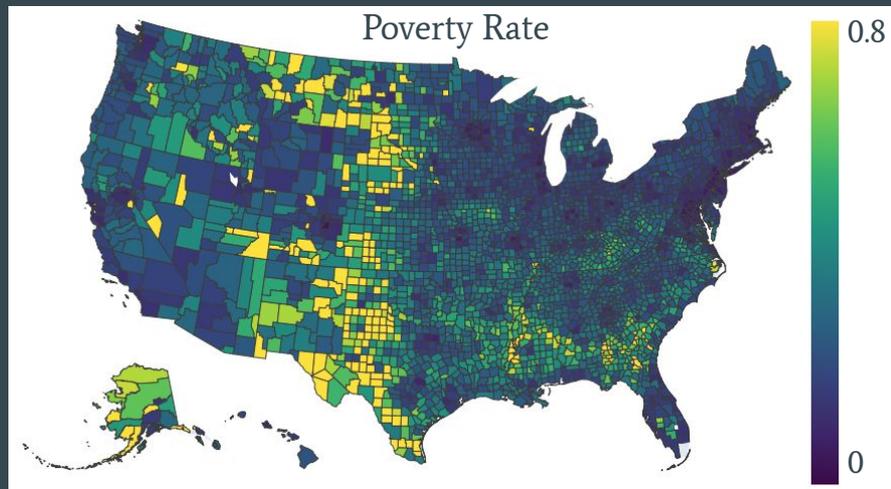
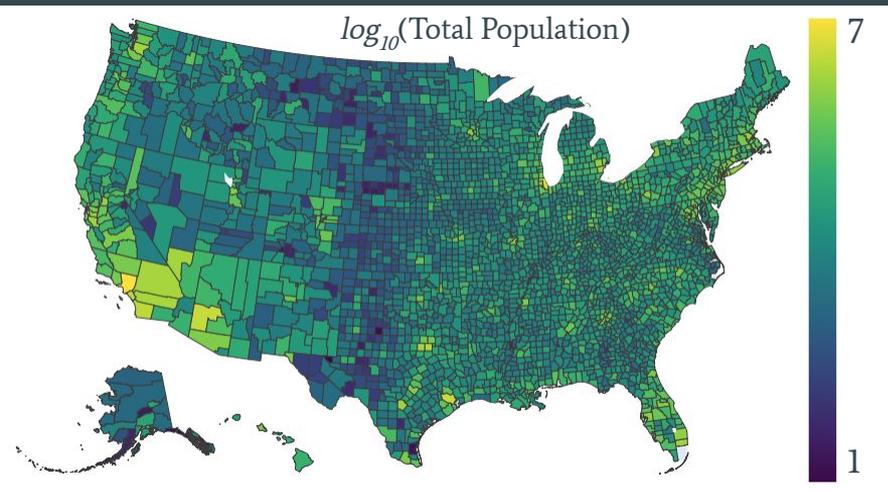
Returns one row for each county

	FIPS	Deaths	Population	County	5_Year_Avg_Poverty_Estimate	Total 5yrAvg County Population	Total 5yrAvg County Deaths	CO 2nd Max 1-hr	CO 2nd Max 8-hr	NO2 98th Percentile 1-hr	NO2 Mean 1-hr	Ozone 2nd Max 1-hr	Ozone 4th Max 8-hr	SO2 99th Percentile 1-hr	SO2 2nd Max 24-hr
0	01001	64.0	365.2	Autauga County, AL	7205.6	34145.0	495.4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	01003	315.8	2156.8	Baldwin County, AL	25849.6	147086.2	1926.4	NaN	NaN	NaN	NaN	-38.333333	-7.428571	NaN	NaN
2	01005	30.8	215.0	Barbour County, AL	6731.2	13496.4	263.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	01007	26.8	173.8	Bibb County, AL	4268.2	12976.2	227.2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	01009	92.2	606.2	Blount County, AL	9114.8	37613.8	591.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3142	56037	42.8	294.6	Sweetwater County, WY	3989.0	25804.6	261.6	-96.571429	-88.888889	-66.6	-92.830189	-38.333333	-5.714286	-51.466667	-94.571429
3143	56039	15.4	153.8	Teton County, WY	1706.4	10321.8	80.6	-96.914286	-94.000000	NaN	NaN	-40.000000	-10.857143	NaN	NaN
3144	56041	13.6	136.0	Uinta County, WY	2291.0	10731.0	126.0	NaN	NaN	-88.0	-96.226415	-40.000000	-9.428571	NaN	NaN
3145	56043	14.8	122.0	Washakie County, WY	988.0	3961.6	76.4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3146	56045	14.0	116.8	Weston County, WY	733.8	2349.2	60.8	NaN	NaN	NaN	NaN	-43.750000	-11.785714	-95.555556	-98.571429

3147 rows x 27 columns

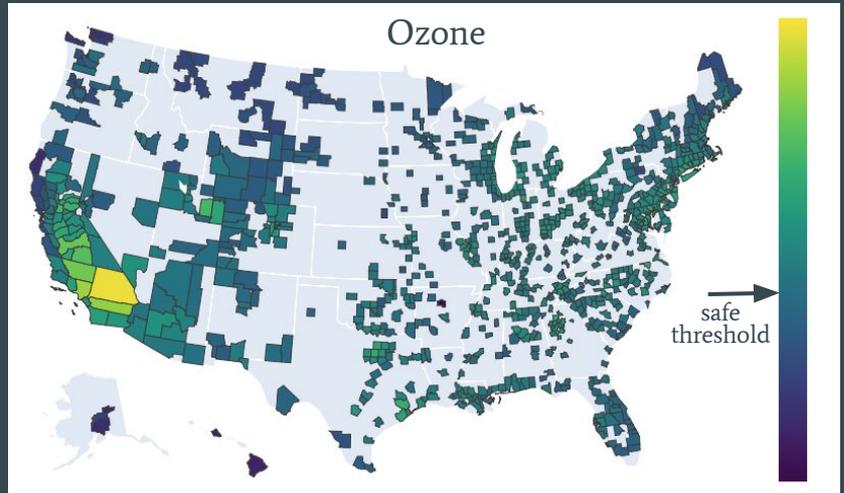
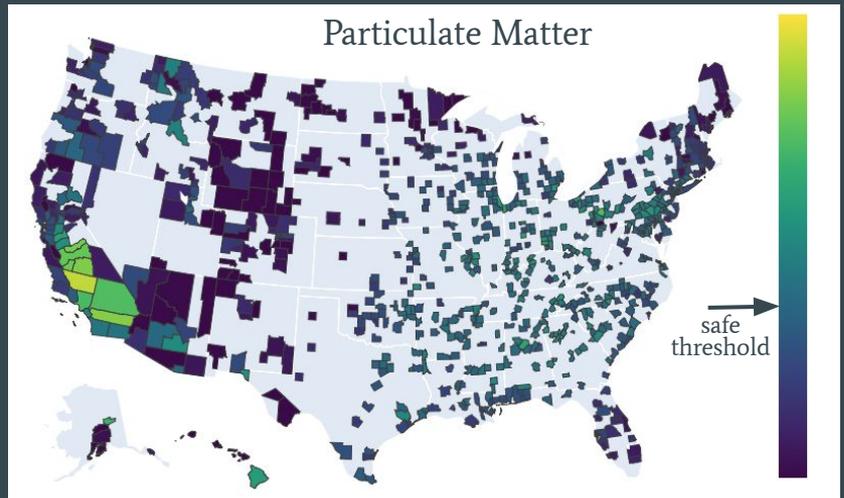
# Geographical Trends

- The correlation between lower population, higher poverty rates, and higher death rates are clearly visible.



# Geographical Trends

- Air quality data includes several pollutants, each with a safety threshold.
- We calculated the mean percent difference over/under the safe threshold.



# Data Challenges

- CDC data is suppressed for numbers of people  $< 10$  in any given population.
- Not every county takes air quality data every year, so many counties are missing.
- These issues are most apparent for:
  - Rural counties
  - Demographic groups with fewer people (i.e. people of color, less common age groups, etc.)
- These issues are compounded when we try to select for populations that are too specific (e.g. Hispanic/Latino Native American women who died between 15-19 years old in the years 2002-2006).
- These data will be more useful for general trends, absolute numbers should be treated with caution.

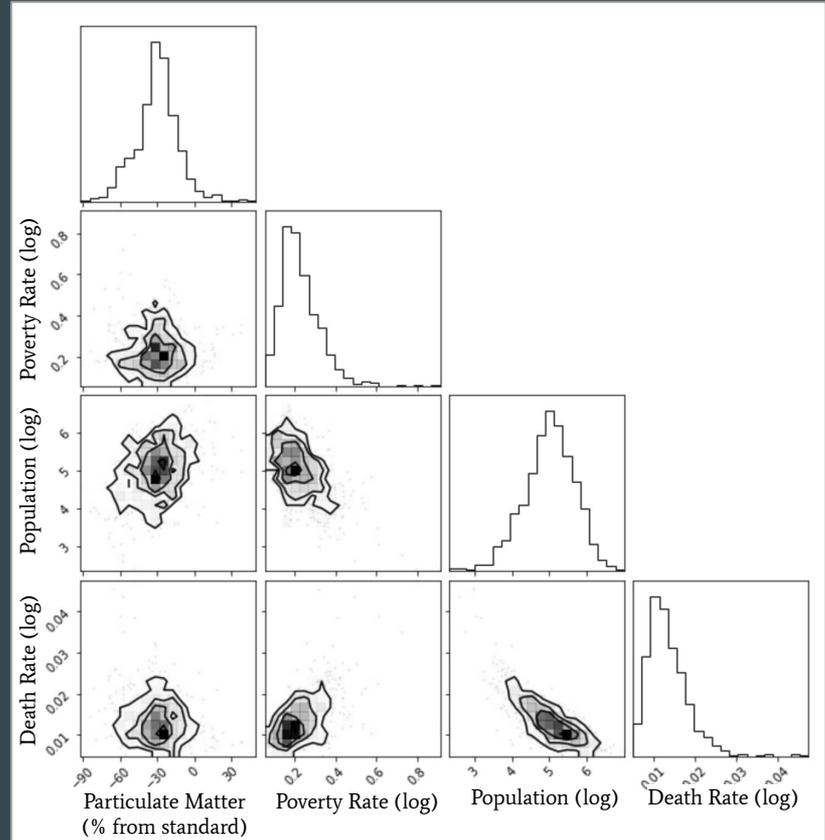
# What can we do with this?

We built a framework to compare a vast array of mortality, population, poverty, demographic, and environmental factors quickly and easily.

**Here are some examples...**

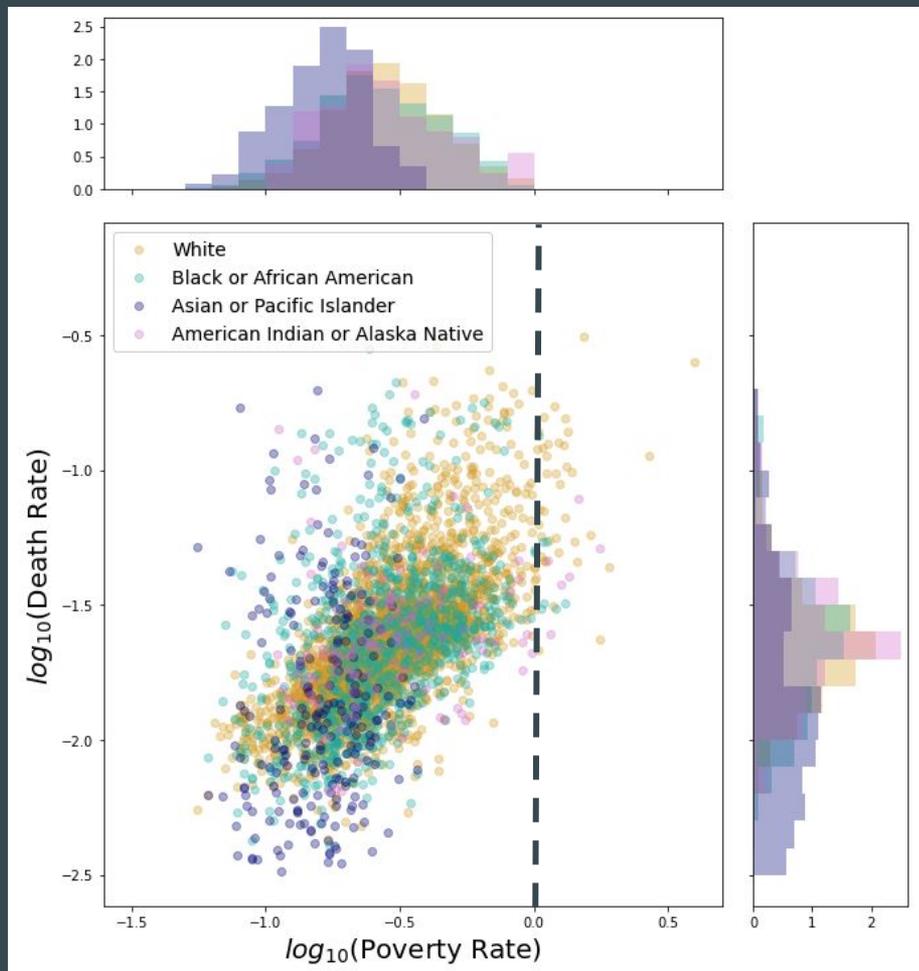
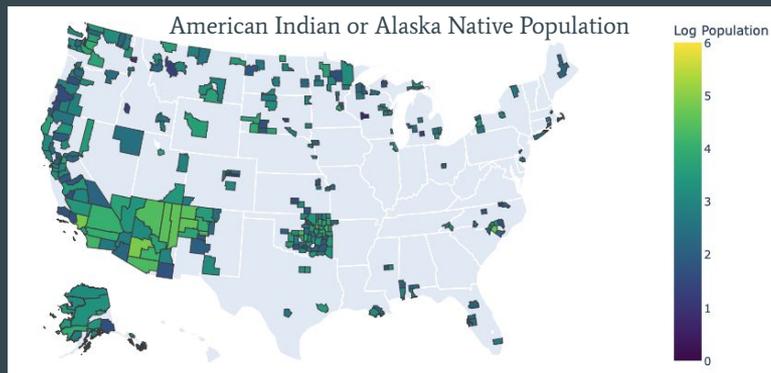
# Basic Modeling

- Poverty and county population are strongly correlated to death rate
- Air quality has little to no correlation with death rates in the available data
- Most air quality measurements are better than the EPA safety threshold, and declined over the 15-years investigated



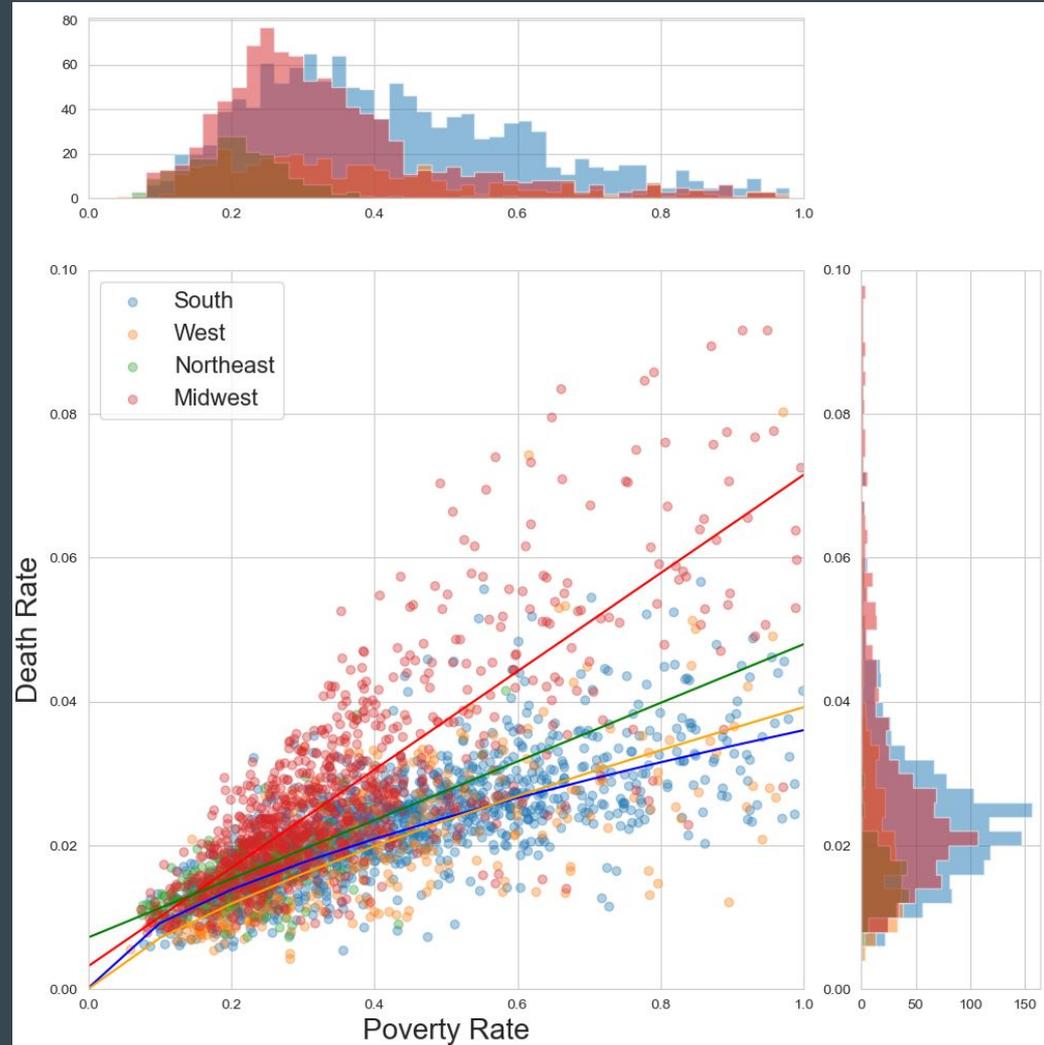
# Demographic Trends - Race

- There is very little reliable data for “Asian or Pacific Islander” and “American Indian or Alaska Native” populations.



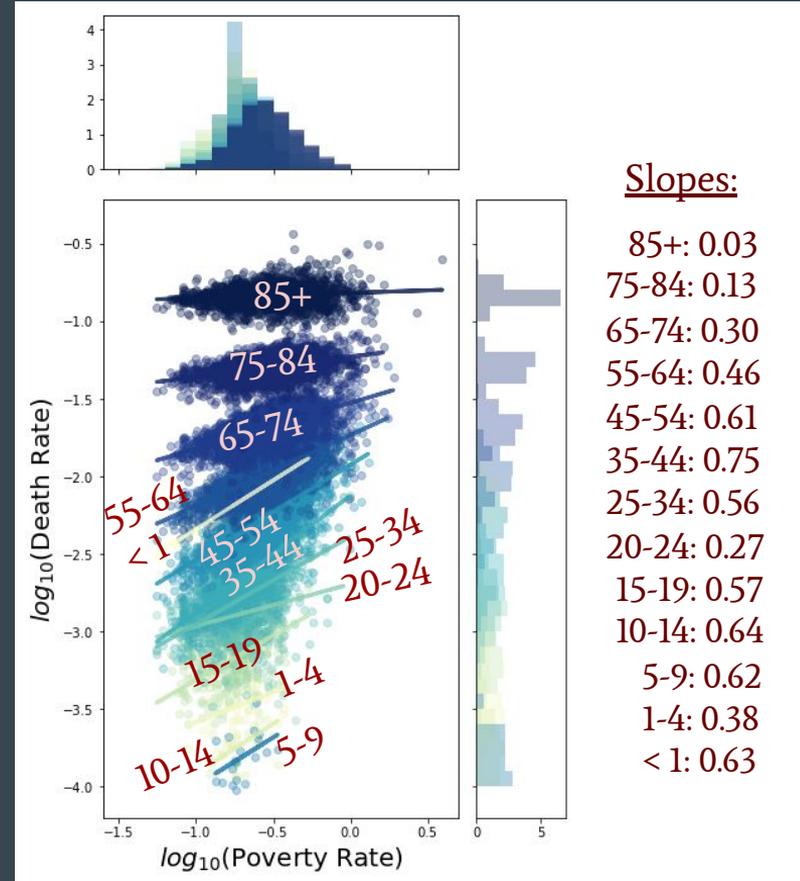
# Trends by US Region

- Correlations suggest that poverty in the Midwest more severely affects the death rate than it does in other regions of the US.



# Demographic Trends - Age Group

- Unsurprisingly, the populations with the highest death rates are the oldest age groups, particularly 85+.
- Children ages 5-9 have the lowest death rates.
- Infants <1 year old have comparatively high death rates.
- The deaths of younger people seem particularly affected by poverty (the slopes are higher).



# Summary

What factors correlate most strongly with higher mortality?

- **Age** – highest mortality for 85+, also comparatively high for infants <1 year old
- **County population** – lower population correlates with higher mortality
- **Poverty** – higher poverty correlates with higher mortality

Other factors...

- Race also plays a role, but given the limited data we have it is difficult to accurately quantify its effect.
- The data for air quality is also extremely limited and we were not able to identify trends with mortality.

# Recommendations and Next Steps

If you want to reduce mortality:

- *Invest in programs that pull people out of poverty.* Across the board, regardless of demographics, poverty is associated with higher mortality.

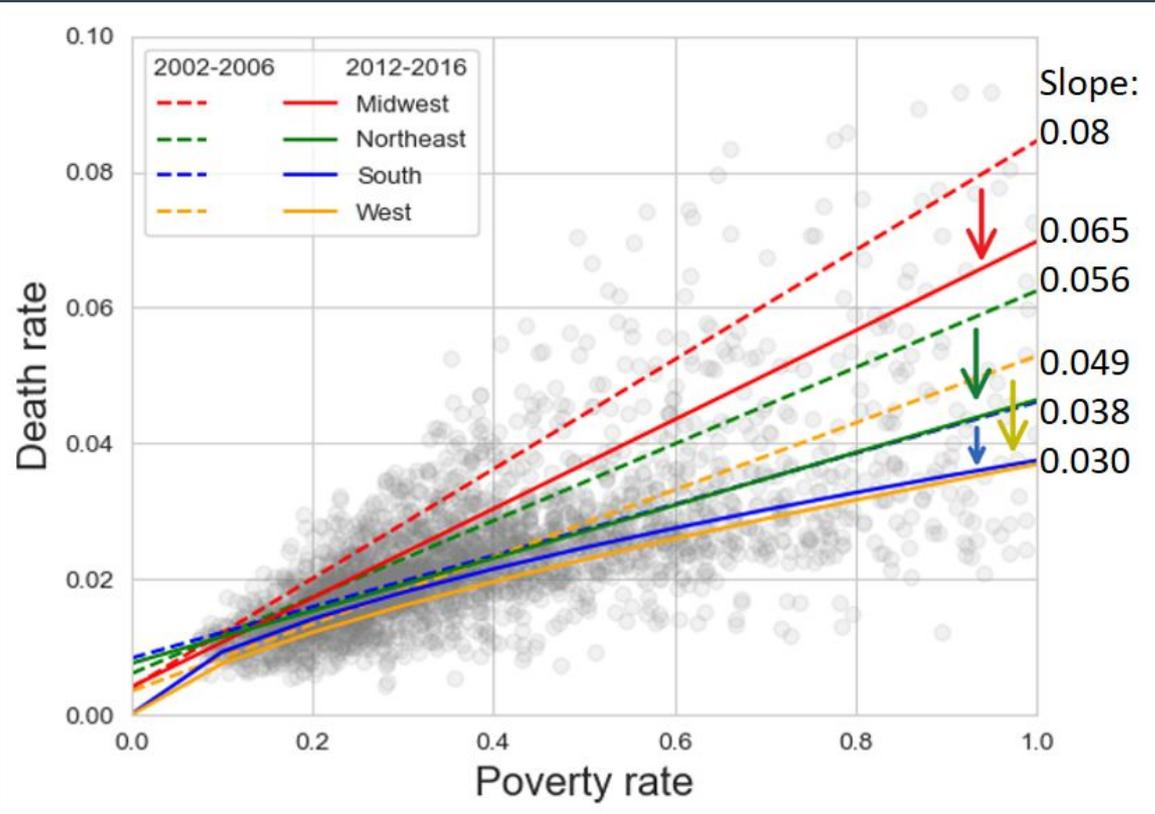
What else could we do with this data?

- Further exploit geographical connections (clusters of counties).
- Use predictive modeling to estimate data values for missing counties.
- Incorporate data from 2017-present to compare the effect of COVID-19.

**Thank you.**

**Extra Slides**

# Poverty reduction over the 15-year period



All regions show a reduction in the poverty rates in the 15-year examined

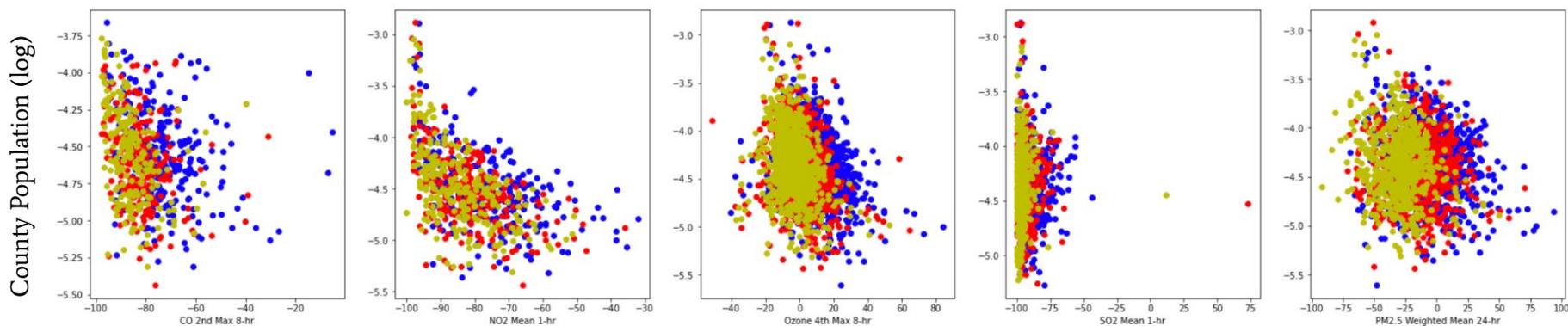
# 15 year trends for pollution

Reduction in air pollutants over 15-year period

Blue: 2002-2006

Red: 2007-2011

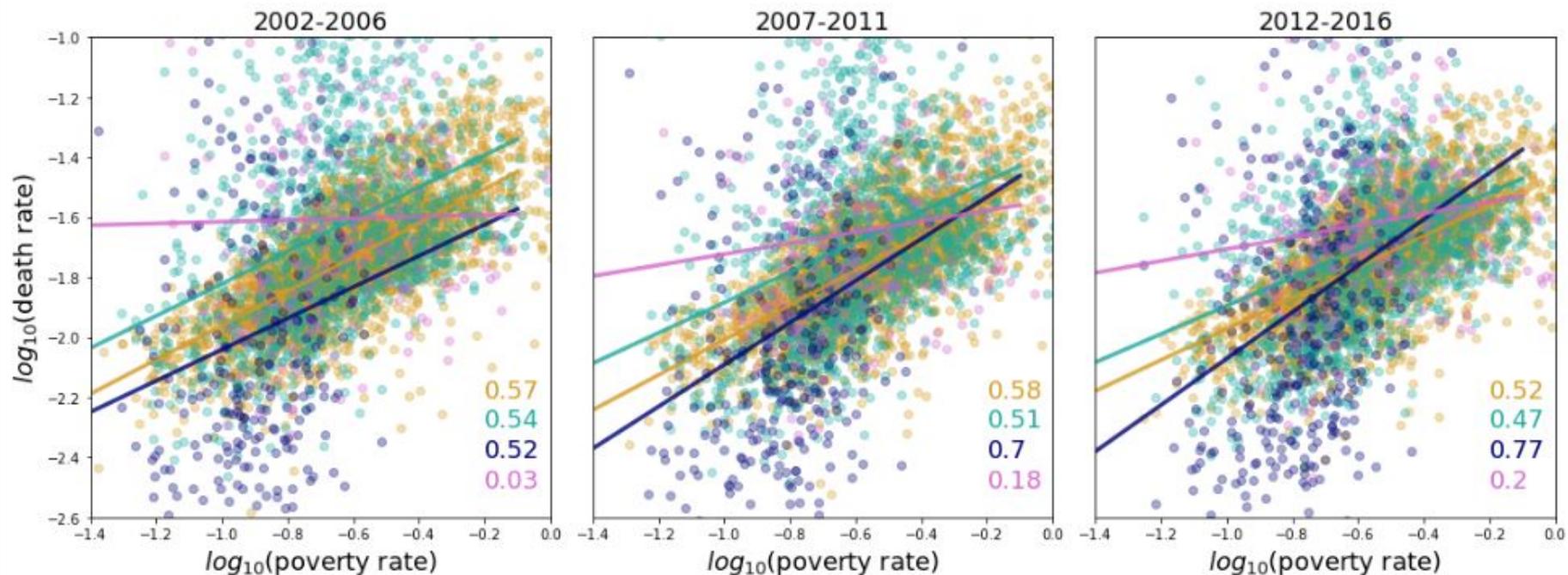
Yellow: 2012-2016



# Demographic Trends - Race

Including 'unreliable' data:

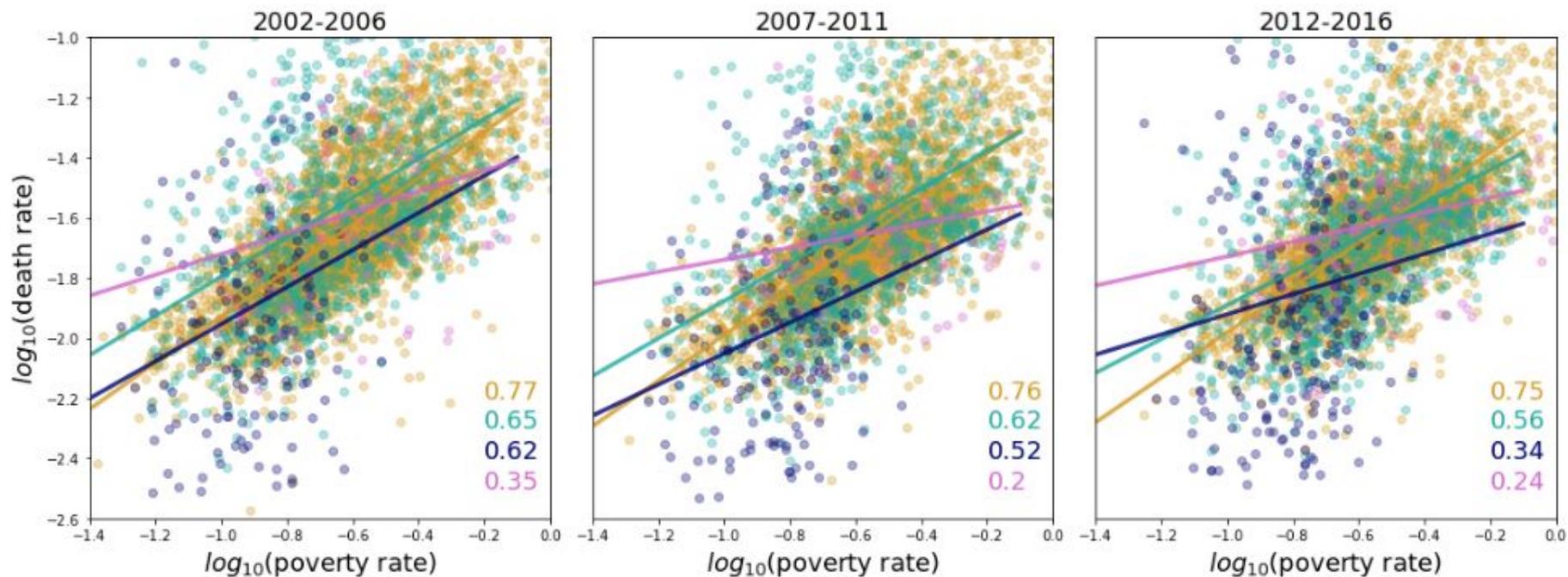
- White
- Black or African American
- Asian or Pacific Islander
- American Indian or Alaska Native



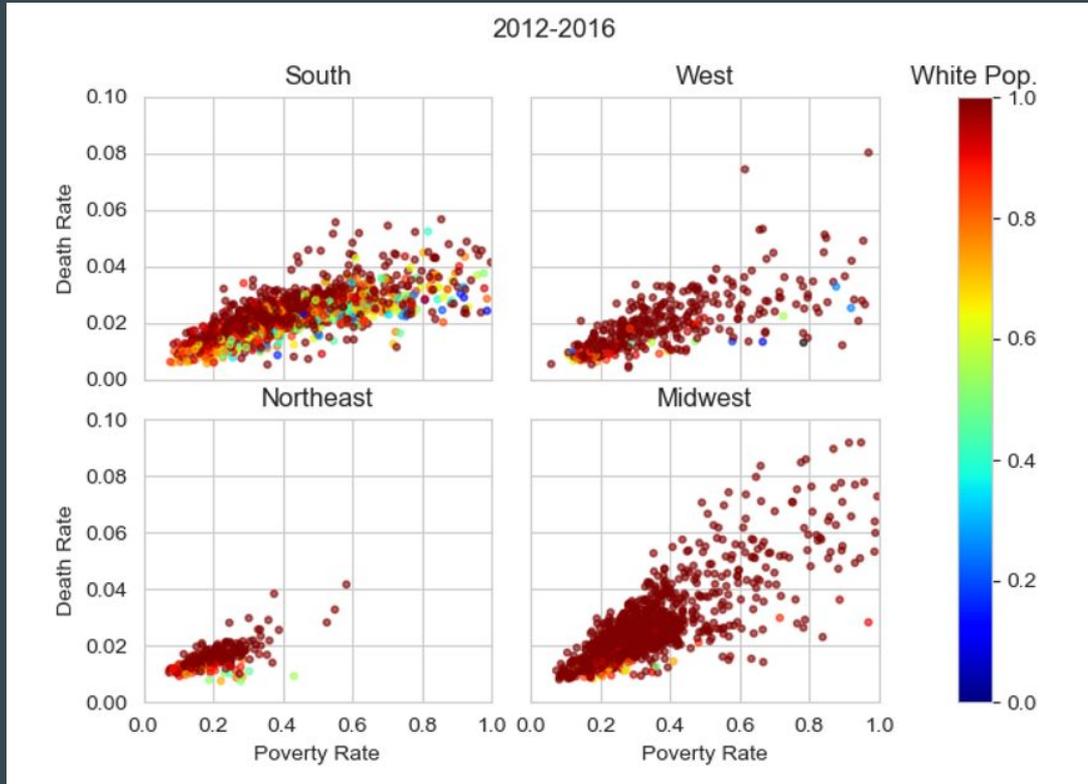
# Demographic Trends - Race

Excluding 'unreliable' data:

- White
- Black or African American
- Asian or Pacific Islander
- American Indian or Alaska Native



# Demographic and Regional Trends - Race



- The Northeast shows the lowest death rates and poverty rates.
- The data does not indicate a clear relationship between race, death rates, and poverty rates.
- The South has the most counties with high populations of non-white people.

# Case Study: St. Bernard Parish, LA

- Shows an 4% increase in  $\text{SO}_2$  over the 15 year interval
  - Still under standard, but one of few counties that shows an increase
- Major industry is oil refining (a known producer of  $\text{SO}_2$ )
- Hit by Hurricane Katrina in 2005, suffered major flooding and oil spills
- Nearly 50% population loss between the 2002-2006 interval and 2011-2016 interval (41 thousand - 21 thousand)
- Poverty rate has more than doubled (20% - 44%)
- No significant change in death rate over the three intervals (1.6% -1.5%)



# Case Study: LA County

- Home of “The Most Polluted City in the Country”
- Since the 1940s has been working to improve air quality
  - Californians can no longer keep crops warm by burning garbage
- 2002-2007 average population: ~9.5 million
- 2.5% increase in population between 2002 - 2016
- Death rate has remained about the same (0.06%)
- Poverty rate has increased slightly (2%)
- Particulate Matter decreased significantly (from 70% over standard to 9%)
- Ozone decreased from 73% to 41% over standard

