**Major Findings: Pollution, Health, and Economics in 70 Select US Cities**

**Questions:**

**Is there a correlation between temperature change and pollution?**With a very weak trendline, the data do not show any clear correlation between average temperature change from 1990 - 2014 and air pollution in PM 2.5 in μg/m3. One interesting lesson from the data was that of the 70 cities that we looked at, 16, or 22.8%, were at or above the EPA’s recommended National Ambient Air Quality Primary Standard of 12 μg/m3. This means that those cities exceeded standards set by the EPA to provide public health protection, including protecting the health of "sensitive" populations such as asthmatics, children, and the elderly.

**Is there a correlation between temperature change and pulmonary disease?**

In addition to looking at how temperature change and air pollution are related, we wanted to see if there was a correlation between average temperature change and pulmonary disease. To visualize this, we plotted average temperature change from 1990 - 2014 against two different metrics of pulmonary disease as measured by the U.S. Centers for Disease Control (CDC): asthma percentage and pulmonary disease percentage. We found there to be a medium to strong negative correlation between and increase in the average temperature between 1990 - 2014 and the percentage of Americans with asthma and pulmonary disease in 2014. Admittedly, this was unexpected, as intuitively we expected greater increases in temperature to correlate positively with increases in pulmonary diseases.

**What is the correlation between air quality and health measures such as life expectancy and heart disease?**

According to the charts plotted comparing air pollution (PM2.5 particulate matter) from 2014 to heart disease and cancer percentages from 2017 CDC data, there is not much of a trend at all. The data points are scattered with no sense of an association with air quality.

As far as air pollution in 2014 vs life expectancy, it seems the data is also pretty scattered across the board, though there does seem to be a line that decreases with life expectancy as the pollution values increase. Other than that, there is not a concrete trend between pollution and life expectancy.

**Is there a correlation between air quality and health measures such as pulmonary (lung) disease and asthma?**

Our scatterplots and regression analysis did show somewhat of a trendline (but only .08 r-squared) for both asthma and pulmonary disease prevalence, drawing from 2017 CDC data and the 2014 Urban Centres database pollution data. Outlier cities such as Detroit (very negative health outcomes, moderate pollution) and San Francisco (good health outcomes, moderate pollution.) Although it’s beyond the scope of our analysis, poverty rate is likely a strong additional influence on health outcomes, in addition to pollution.

Looking at the heat maps, we can see that they are almost identical when comparing particulate matter in the air to asthma percentages. This makes sense, as the higher the concentration of particulate matter in the air, one would assume that the percentage of people with asthma would be higher as well.

**Is there a correlation between the economic rank of a city in the US and the air pollution in the US in 2014? What about the poverty rate in the US from 2011-2018 and the air pollution in the US in 2014?**

There was no correlation between the economic rank of cities (according to the Milkin Institute, 2018) in the US and the air pollution in the US in 2014. The trendline/regression line was horizontal in the scatter plot of the data, which means that there is no trend between the economic rank and pollution in 2014.

There was a slightly positive correlation between the poverty rate in the US from 2011-2018 US Census American Community Survey and the pollution in the US in 2014. In the scatterplot of the data, the higher the poverty rate, the higher the pollution was in that city in 2014.

**Data Sources:**

US Census API data from American Community Survey 2011-2018 1-Year Data -- https://www.census.gov/data/developers/data-sets/acs-1year.html

CDC 500 Cities health data 2017 -- downloaded and merged via CSV -- https://www.cdc.gov/500cities/index.htm

CSV data of PM2.5 particulate pollution, CO2, and temperature data from the European Commission's global Urban Centres Database - part of Global Human Settlement data project -- https://ghsl.jrc.ec.europa.eu/ucdb2018Overview.php

CDC CSV = separate life expectancy download, 2010-2015 data -- https://healthdata.gov/dataset/us-life-expectancy-birth-state-and-census-tract-2010-2015

Source for "Economic Rank" -- 2018 Milkin Institute Best Performing Cities Index --- http://www.best-cities.org/2018/best-performing-cities-report-2018.pdf