Justin Madsen

DSC550 / Brett Werner

Project Milestone 1

I am looking into how deaths due to air pollution have changed throughout the years. I am curious to see how pollution related deaths compare to other hygiene-related deaths, and if changes to travel (world events or war time in certain areas) show any relevant changes to these death rates. I'm hoping that I can find a downward trend of deaths throughout the years. If not, then a focus on what can be down to turn trends around in that region may be a good way forward after the project. The data provided after the appropriate model is fit may also provide a good idea of when previously high death rate countries will get close to zero death rates. Ideally, this data would then be used to direct future efforts in providing additional hygiene precautions to at-risk countries in order to promote healthier ways of living.

Subsetting the data by country might be easier accomplished in R, however Jupyter will allow for easier graphs through for loops to demonstrate deaths over time per country. A linear regression model may end up being the best way forward, however determining which algorithm to use will depend on how the variables stack. Having a variable with total population of the countries in their respective years would provide the ability to determine the odds of someone dying by the various means in the dataset.