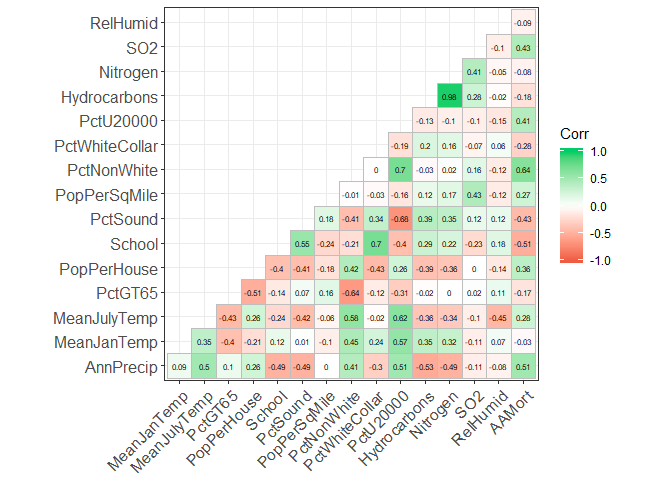
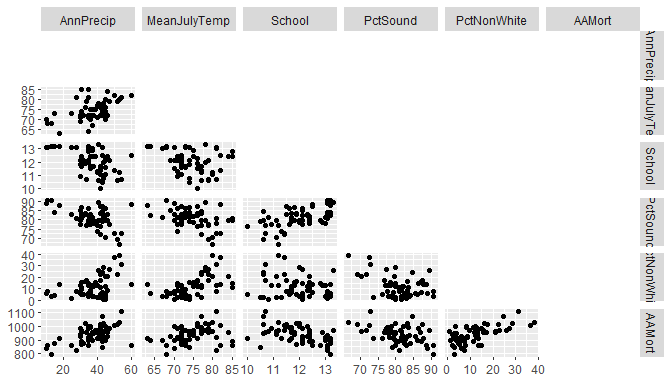
Environmental Impacts

Ethan Kemeny

1. As humans, we want to know how we can live better and longer. This leads to researching what factors contribute to our health. This dataset includes environmental variables from 60 cities and the cities’ age-adjusted mortality rate. Using this data, we can attempt to find the relationships between different factors and how they impact mortality. We can also try to predict mortality for cities with various levels of the different variables.
2. The following graph shows the correlation values between the different variables. The darker the green, the more positive the relation, the more red means more negative relation. White boxes indicate little or no relation.

As we can see there are several strong positive and strong negative relationships between the different variables and mortality (the furthest right column). Some of the stronger positive relations are Percent Non-White and Annual Precipitation. This means as these factors go up, so does mortality generally. Some of the strong negatives include School and Percent Sound. This means as these factors go up, mortality goes down.

The following graph shows a few of the stronger relationship variables. The ones in the green box correspond to mortality.

Here we can see they are fairly linear.

1. The data seems to be mostly linear so we can perform multiple regression on this data. This will allow us to find which factors have the biggest impact on mortality and be able to help determine that relationship. It will also allow us to predict mortality given the certain levels of the different variables of a city.
2. I’m not sure how to find Q-Q interactions and how to deal with variables that show almost no relation to the dependent variable.