Climate Change

Stephen Smitshoek

College of Science & Tech, Bellevue University

DSC530-T302: Data Exploration and Analysis

Professor Catherine Williams

August 10, 2022

Do humans have a direct impact on climate change? The global temperature is increasing, the ice caps are melting, and the sea levels are rising, and all of this is a direct result of humans. These are all facts that are told to us by scientists all the time, but can we use historical data to fact check this and see the results for ourselves?

Four different sources of data from Kaggle were used, which included, global average temperatures, sea levels, CO2 levels in the atmosphere and CO2 emissions by humans. After the data was cleaned a PDF plot was drawn which showed very clearly that the average yearly temperature has been significantly higher post 1950 than it was pre-1950. This clearly indicated that there is a change in the climate but not what is causing it. The check the human contribution on different aspects of climate change correlations were fit to the data with the following results.

Temperature vs CO2 Emissions 0.70 p < 0.001

CO2 Emissions Vs CO2 Levels 0.99 p < 0.001

Sea Level Vs CO2 Levels 0.99 p < 0.001

While these correlations are very high, they do not imply causation, however when paired with the scientific communities’ findings it becomes very hard to argue against.

One variable that could have used some more depth was the CO2 emissions. This variable was reported as an emission number for every country in each year since 1751, however there were several countries that had no reported emissions during the early centuries. In the analysis, these were assumed to be zero, but it could be that the numbers were not accurately recorded. Additionally, it was unclear if the CO2 emissions were strictly CO2 or were a CO2 equivalency. CO2 is known to be a green house gas, but it is not the only one and so other green house gases are sometimes reported as a CO2 equivalent number. If the emissions reported were strictly CO2 than there are many other green house gases that are not being considered.

There were two aspects to this EDA that could have used more insight. The first was to fit a non-linear curve to the temperature over time. While an attempt was made to do this the curve never quite fit right. It is possible that a separate curve needed to be fit to the second region in the model (post 1990). If this curve fitting had worked than predicting the future rise in temperature would have been valuable.