Capstone Final Report

The motivation behind this capstone comes from the ever-growing concern of global warming. Most people have accepted that it is real and getting worse. The larger question in my opinion concerns responsibility: Which nations are the primary cause of it. Global warming is due to carbon gases being trapped in the atmosphere, causing the temperatures in the air and water spaces to warm. It is reasonable to assume that this accelerated warming process started in the mid to late 1800s, corresponding to the beginning of the Industrial Revolution. With this project I look to do two things: 1) Given a set of dates and corresponding temperatures, I will forecast the average temperatures 35 years past the dates in my dataset and 2) Investigate if a set of countries have a larger contribution to global warming based on their rate of increase of their average temperatures.

I hypothesize that the temperatures for the industrialized nations will have risen at a faster rate than the developing nations because the industrialized nations have more factories, burn more coal and clear away more vegetation that would remove the carbon gas. The approach for this project is to investigate the average temperatures for a set of industrialized nations and underdeveloped nations at monthly intervals. Then I will build time series models for each city to forecast temperatures beyond those given by the datasets. Finally, I will compare the rates of change for the average temperatures using a moving average graph.

Findings

I was able to draw surprising conclusions from this project. The first, contrary to my initial thoughts, is that the average temperatures of the developing nations rose at about the same rate as those of underdeveloped nations. For the period from 1870 to 2013, the developing nations of India, Haiti, Senegal, and Cambodia experienced their average temperatures increase by 1.1, 1.6 and 1.4 degrees Celsius respectively, whereas the industrialized nations of Japan and the UK experienced average temperature increases of 1.5 and 1.1 respectively. However, the United States proved to be an outlier as its average temperature increased by 3 degrees Celsius. After seeing these numbers, I brought in the country of China, since it is a very large, industrialized country. Nevertheless, its average temperature increase was about 1.1, right in line with the other countries in this report.

Of the models that I used for this project I found the ARIMA model to give the best results. It required more work at the beginning, but its forecast results were more accurate and precise. Before doing a forecast, I trained an ARIMA model on a partial dataset for each city (approximately 80% of the data) and used that training data to forecast average temperatures for the years corresponding to the test sets. The results for Cambodia, Senegal and Haiti were pretty accurate for the first 15 years, seemed to do slightly worse proceeding forward, was still good for a long-term forecast. However, the results for the United States, India and Japan were amazingly close from the beginning to the end of the forecast period. The next model I tried was the SARIMA model. This choice was motivated by the forecasting results of the Haiti dataset. Since the SARIMA model is geared to accommodate seasonality, it was worth fitting the data from Haiti with that model. The results were not accurate; the top of the forecast values severely overshot the actual values and the bottom of the forecast missed the actual bottom and gradually got worse.

The findings from this project suggest that efforts to control global warming should be a global effort; indeed all nations should do their part to lower carbon footprints. That being said, the project also implies extra care should be taken in the United States as evidenced by the 3 degree-Celsius rise in average temperature from 1870 to 2013.