**World Energy Analysis**

**DSC 530 EDS Final Project**

**Paul Hart**

Energy is currently a hot topic across the world, and it’s a very important topic because it us vital for everyday living. The energy industry is at the top of news headlines over the type of energy that is being used and for its implications on climate change. The world is currently in an energy transition to lower carbon energy. However, as the world makes this transition, low cost and reliable energy is going to be needed to allow the world to operate as new technology is developed and new energy infrastructure is being built. My passion around this critical topic encouraged me to perform an exploratory data analysis (EDA) on world energy. I believe this was a challenging topic because world energy consumption and electricity production are time series data, and the right questions and correct variables had to be used to extract the most information. The most comprehensive data sets I could find on this topic was from a website called our world in data. This source has multiple data sets on energy for the world and specific countries. Using two data sets from this source I combined them to produce one data set that contained information on energy consumption from fossil fuels and green energy, and electricity generation from fossil fuels and green energy. All energy source variables were in kilowatt-hour (kwh), which is a unit of energy. The goal of the project was to determine what energy source is currently used the most in the world, and more specifically what type of fossil fuel source and what clean green energy source is used the most in the world, and how do the top five producing energy countries compare. Also, is there any correlation between energy source variables, and can predictions be made with regression analysis on whether the use of fossil fuels will increase or decrease in the future.

World energy data analysis has shown that the top used energy sources in the world are still from fossil fuels. Coal is the most used energy source in the world for electricity generation, and oil is still the most consumed energy source in the world. The top five energy producing countries in the world are China, United States, Russia, Saudi Arabia, and Canada. Out of these countries the United States and China use the most coal, with China taking the lead in coal use around the year 2018. Since the analysis showed that nuclear energy was the top used green energy source in the world, that source was chosen to reflect green energy use between the top 5 energy producing countries. The two countries that use the most nuclear energy are the United States and Canada. Correlation analysis has shown that the United States is heading in the right direction, and there is a weak to strong negative correlation between United States coal use and world coal use. China is trending in the opposite direction and correlation analysis has shown that coal use in China is increasing, and that there is a strong positive correlation between China coal use and world coal use. China appears to be the largest contributor to coal in the world. I think the world agrees that coal use has to decrease, and in most countries, this is happening, however what about oil? Oil is still a number one energy source in the world for many different industries. The analysis has shown that electricity generation from oil has dropped significantly, but oil consumption has flatlined. Regression analysis has shown that electricity generation from oil will continue to decrease in the future, and oil consumption will actually slightly increase in the future. This has been visualized in the time series data, where oil consumption has flattened but there are still spikes and valleys in the data. The least squares regression analysis has showed that for every year, oil consumption increases by 3.57 kwh. It looks like oil is still in demand and will continue to be needed in the future.

Data source = [https://ourworldindata.org](https://ourworldindata.org/)