MSIS 5673

Data Visualization Project Abstract

**- Sai Teja Sagi**

# **Data Description**

The Energy Statistics Database contains comprehensive energy statistics on the production, trade, conversion and final consumption of primary and secondary; conventional and non-conventional; and new and renewable sources of energy.

This dataset was kindly published by the United Nations Statistics Division on the UNData site. The data has 1.04 million observations and 6 features each describing the:

Country – Details the country name that has had the transaction done

Commodity Transaction – Indicates the type of energy – renewable or non-renewable

Year – Year in which the transaction happened

Unit – Unit of the commodity transaction i.e., Metric Ton, KJ, etc.

Quantity – Quantity of the transaction

Category – Type of the energy product i.e., coal, wind energy, natural gas, etc.

I have only used the energy details corresponding to renewable types of energy such as wind, solar, hydro, nuclear, geothermal, etc.

# **Analysis Goal**

I have always been curious about the growth of renewable energy in the world. I wish to see how the growth has happened across different countries in the last 25 years. From the analysis, I wish to see which country has been the proponent in advancing different kinds of renewables.

# **Initial Review**

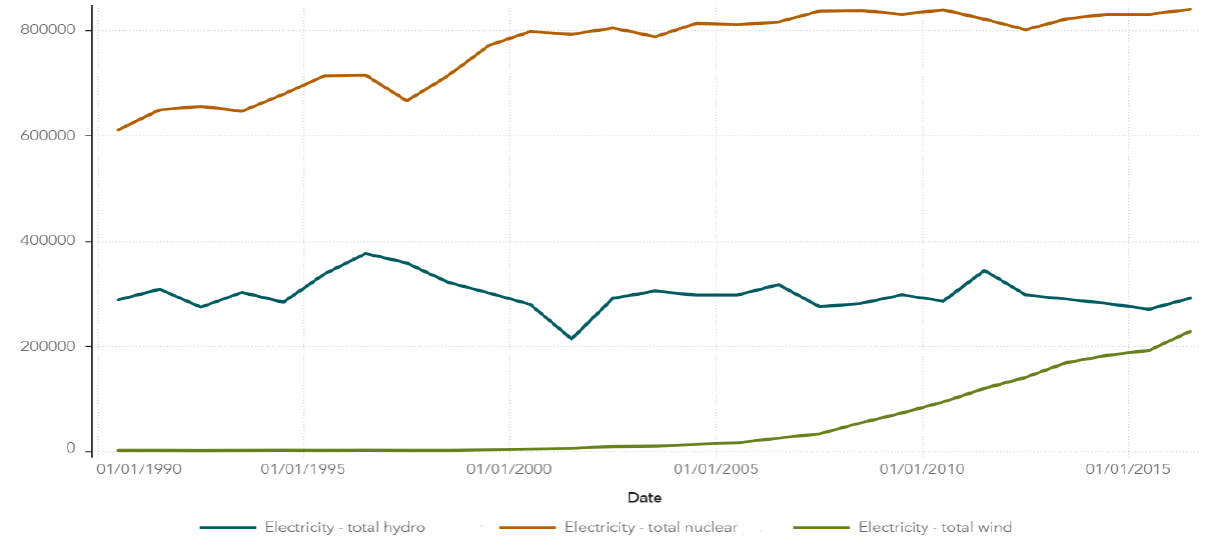
1) We can see from the analysis that the USA is among the countries that have the most significant growth rates in renewable energy usage although its overall capacity is lower in certain cases.

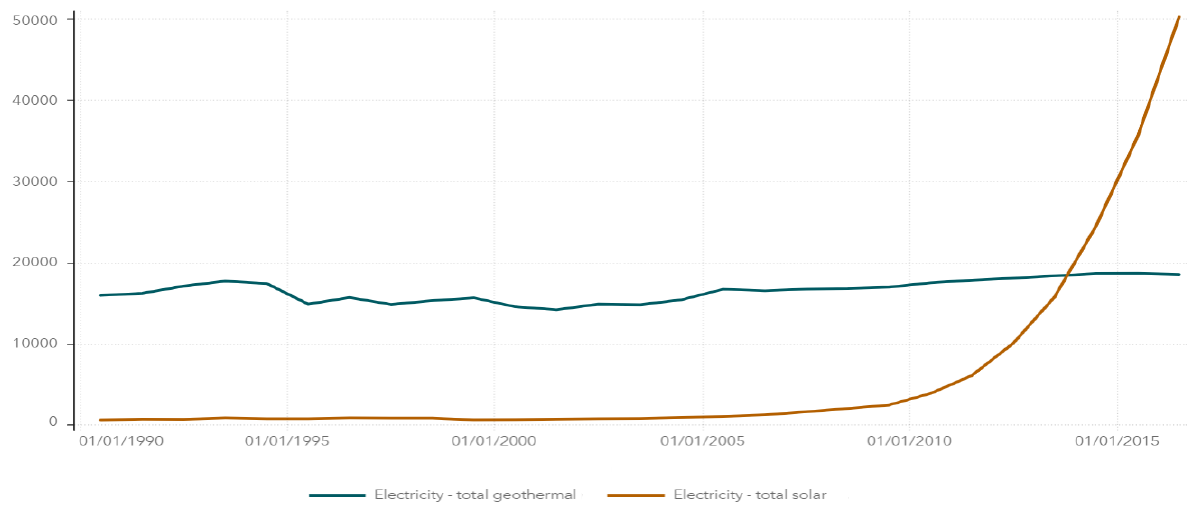
2) China has the largest increase in hydropower generation.

3) Developing countries are more active than developed countries when it comes to increasing their respective capacities of renewable energies.

# **Screenshots and descriptions of future graphs**

I will plot several kinds of charts like line charts, bar graphs, etc. showing the usage, increase in usage and percent increase in usage for different countries. I will use butterfly charts and Pareto charts to see the changes across time. I will build dashboards showing a bigger picture in a single slide.





# **Challenges in the data**

The original dataset is so huge and has a lot of extra information. I had to use a software called SAS to subset only those rows where there are details for the renewable energy transactions. There are a lot of levels in the commodity\_transactions variable and getting a holistic view of the data was difficult. After a deeper understanding, it became easy for me to see the picture.

# **Reviewer Feedback**

**Name** - Jaya Bhatia

**Feedback** – Jaya was a classmate in Tableau class, and I believe she has excellent data viz skills. So, I asked her to review my work. While I have not solidly finished the project, Jaya kind of liked the idea. She advised me to keep the data in long format and remove any aggregations at row level. She asked me to try to put maps in the project because they look visually appealing when used properly. Finally, she has promised to give honest feedback once I show her the final project.