TITLE: **Energy Analysis in Python**

AUTHOR: **Giovanni Aratico**

DATE: **12/03/2022**

DESCRIPTION: Provides a brief high level overview of world energy metrics in a Python Shiny Dashboard Application.

QUESTIONS:

- 1.) How does world electricity generation change over time?
- 2.) How does the distribution of electricity generation sources change over time?
- 3.) Can we approximate trends and predict energy usage by source and country over a given period of time?

CHALLENGES:

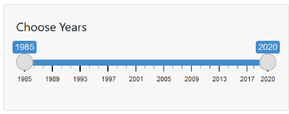
- Not all countries have a full report of certain metrics for certain time periods.
- Some sources like BioFuel did not start reporting until 1985.

ANALYSIS:

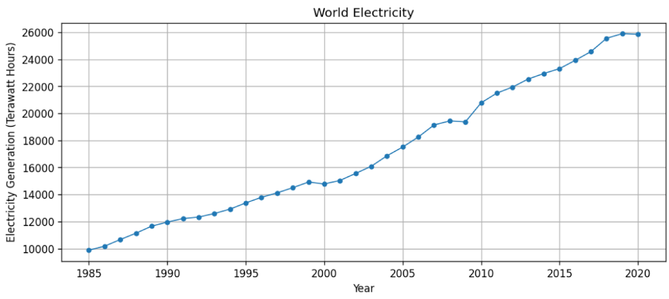
- Data sourced from Kaggle in CSV format.
- CSV load into Python, all wrangling/transformations are done inside each respective method tab.
- Each question is allocated an individual tab for visual interactivity.
- Code was then separated into multiple modules for easier readability.
- For the regression analysis:
  - If there is no data from the selected time period, an error is shown to the user.
  - If the data is only partial, the app will create a model using whatever data is available while also notifying the user of incomplete data.
  - Users can choose between a linear model and a polynomial model to the 2nd degree.
  - The R squared metric is reported on screen for both partial and complete data for models.

ANSWERS:

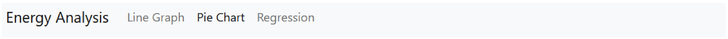
- 1.) World electricity generation has increased over time. This is not unexpected due to increasing populations and developing countries.
  - World electricity increased from approx. 10,000 tW hours to just under 26,000 tW hours.
  - Overall follows a roughly linear trend.

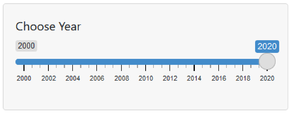


Range:

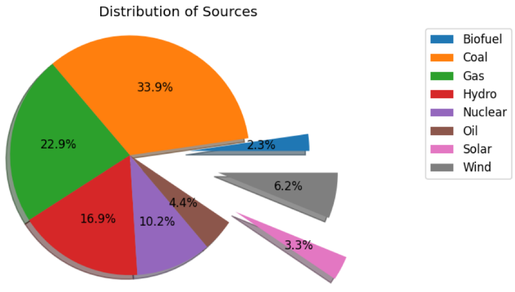


- 2.) Renewable energy sources (Solar, Wind) have increased in usage where fossil fuels (Coal, Oil) have decreased over all. Noteworthy trends include:
  - Solar increases from < 0.1% to 0.1% in the 2000 - 2010 time span, then from 0.1% to 3.3% in the 2010 - 2020 time span. A significant growth.
  - Nuclear decreases from 17.1% to 10.2% in the 2000 - 2020 time span.
  - Oil decreases from 8.0% to 4.4% in the 2000 - 2020 time span.
  - Coal is the largest share for every year in the reported span.

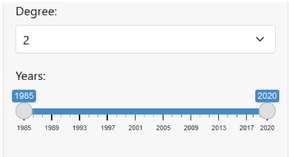




Distribution:

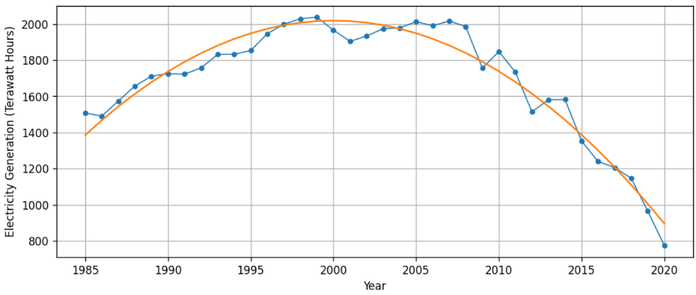


- 3.) Trends shows correlations gravitate closer to linear or single curve when you select larger and more developed geographic regions and using energy sources accounting for a bigger share. When selecting single countries and sources accounting for a smaller share, the trend is more likely to be varied and thus harder to get an accurate model using a 1st or 2nd degree model.



Linear Regression:

```
R SQUARED: 0.2119973178395778 7
```



SOURCES:

- https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption

TECH USED:

- Python
- NumPy
- Pandas
- Shiny for Python
- MatplotLib
- Scikit-Learn

FOLDER STRUCTURE:

- /data/
  - wec.csv
- /img/
  - plot1.png
  - plot2.png
  - plot3.png
- /src/
  - nav_controls.py
  - plot_controls.py
  - wrangle.py
- /app.py
- /Energy.ipynp