Aluko_Week 2: Data Science Practicum - Project Proposal

1. **Name, Contact info (e.g., email/phone):** Olumide Aluko (olumidelk@gmail.com/ 281-857-7607).
2. **Title of the project:** Forecasting Electricity Demand with Time Series
3. **High-level description of the project: what question or problem are you addressing?** A time series with electricity demand (Mega Watts) for the state of Victoria (Australia) from 2011-12-31 to 2014-12-31 is available. Demand for electricity in Australia has been in the spotlight for the general population due to the recently increasing price. Still, forecasts of the electricity demand have been expected to decrease due to various factors. The project aims to generate a forecasting model capable of predicting the next day's energy demand at the hourly level by accurately predicting monthly electricity demand. The proposed project design will be achieved using a time series forecasting with scikit-learn regressors
4. **What type of data science task is it?**
    - prediction using time series with scikit-learn regressors
    - data visualization
5. **Data: Brief description of data. How big do you expect the data will be? Is the amount of your data too big or too small? How long do you expect to collect the data if you're web-scraping or collecting data?** This dataset is for operational demand, which is the demand met by local scheduled generating units, semi-scheduled generating units, non-scheduled intermittent generating units of aggregate capacity more significant than 30 MW, and by generation imports to the region. The operational demand excludes the demand met by non-scheduled non-intermittent generating units, non-scheduled intermittent generating units of aggregate capacity smaller than 30 MW, exempt generation (e.g., rooftop solar, gas tri-generation, tiny wind farms, etc.), and demand of local scheduled loads. It also excludes giant industrial users (such as mines or smelters). I plan to download the GitHub dataset, whose sizes are 2,803 KB.
6. **How will you analyze the data? What machine learning methods do you plan to use, and what business intelligence aspect do you plan on incorporating?** I plan to analyze the dataset with Time Series and using Jupiter Notebook Python 3.
7. **Describe any anticipated difficulties and problems. Discuss how you may overcome the problems:** I anticipated difficulties from updates, required libraries, error codes,

Hyperparameter optimization, cleaning, etc. And I intend to research journals, stack flow, and any other material source that would guide the successful run of the dataset.

8. **Suggest a timeline for the project. This should be a weekly breakdown of what you plan on doing each week:**

- Week 2 – Project proposal research & submission
- Week 2 – Data collection
- Week 3 – Data cleaning
- Week 4 - Data exploration
- Week 5 - EDA
- Week 6 - Build models
- Week 7 - Hyperparameter optimization
- Result visualization & conclusion