**Electricity Demand Time Series Forecasting**

**Project Overview**

This project focuses on analyzing and forecasting electricity demand in Great Britain using time series data provided by **National Grid ESO**. It can be accessed from: [Kaggle.](https://www.kaggle.com/datasets/albertovidalrod/electricity-consumption-uk-20092022/data) The dataset includes electricity consumption records from **2009** onward, updated **48 times per day** (every 30 minutes). The data was aggregated to **daily frequency** by taking the **maximum consumption value per day**.

The primary goal was to forecast maximum consumption in a day. This was done by performing **exploratory data analysis (EDA)**, assessing stationarity, and by developing accurate time series models for forecasting electricity demand.

**Dataset**

* **Source**: National Grid ESO
* **Frequency**: Half-hourly (converted to daily)
* **Time Range**: From 2009 to 2024
* **Data Preprocessing**:
  + Aggregated data to daily frequency (maximum consumption per day)
  + Interpolated missing data points

**Project Workflow**

**1. Exploratory Data Analysis (EDA)**

* Analyzed trends, seasonality, and outliers in the time series data.
* Interpolated missing data points for consistency.

**2. Statistical Analysis**

* **Classical Decomposition**: Decomposed time series into trend, seasonality, and residual components.
* **Pymannkendall Test**: Checked for monotonic trends in the data.
* **Kruskal-Wallis Test**: Tested for seasonality in the data.

**3. Stationarity Analysis**

* **ADF Test** (Augmented Dickey-Fuller): Checked for stationarity.
* **KPSS Test**: Confirmed stationarity results.
* **ACF and PACF Plots**: Analyzed autocorrelations to identify seasonality and lag patterns.

**4. Time Series Forecasting**

* Built and compared the following models:
  + **ARIMA**: Autoregressive Integrated Moving Average
  + **SARIMA**: Seasonal ARIMA
* Achieved the **best performance** with:
  + **SARIMA(2,0,5)(5,1,2,7)**
* Model evaluation metric: **Mean Absolute Percentage Error (MAPE)** of **3.5%**.

**Results**

The **SARIMA model** effectively forecasted electricity demand with a MAPE of **3.5%**, showcasing its ability to capture both seasonal and non-seasonal patterns.

**Technologies Used**

* **Python**: Core programming language
* **Pandas**: Data preprocessing
* **NumPy**: Numerical operations
* **Matplotlib & Seaborn**: Data visualization
* **Statsmodels**: Time series analysis and modeling
* **Pymannkendall**: Trend analysis

**Future Work**

* Incorporating external features like temperature, day type (weekend/weekday), or holidays for improved accuracy.
* Exploring advanced models such as **Prophet** or **LSTM** for forecasting.

**Conclusion**

This project demonstrates a comprehensive approach to time series forecasting, from EDA to building robust SARIMA models. The achieved MAPE of **3.5%** reflects the accuracy and reliability of the model in predicting electricity demand.

**Contact**

For any doubts, feel free to reach out.