CAPSTONE PROJECT:

ENERGY CONSUMPTION FORECASTING

LIONEL LWAMBA

PROBLEM STATEMENT

- **Challenge:** Accurately predicting future energy consumption is very important for utilities, and could also be important for households
- **Impact:**

a. Matches power supply to household or consumers demand

b. Predicting use helps utilities avoid waste, potentially lowering your electricity bill.

c. By predicting peak usage, utilities can prevent outages and ensure reliable electricity

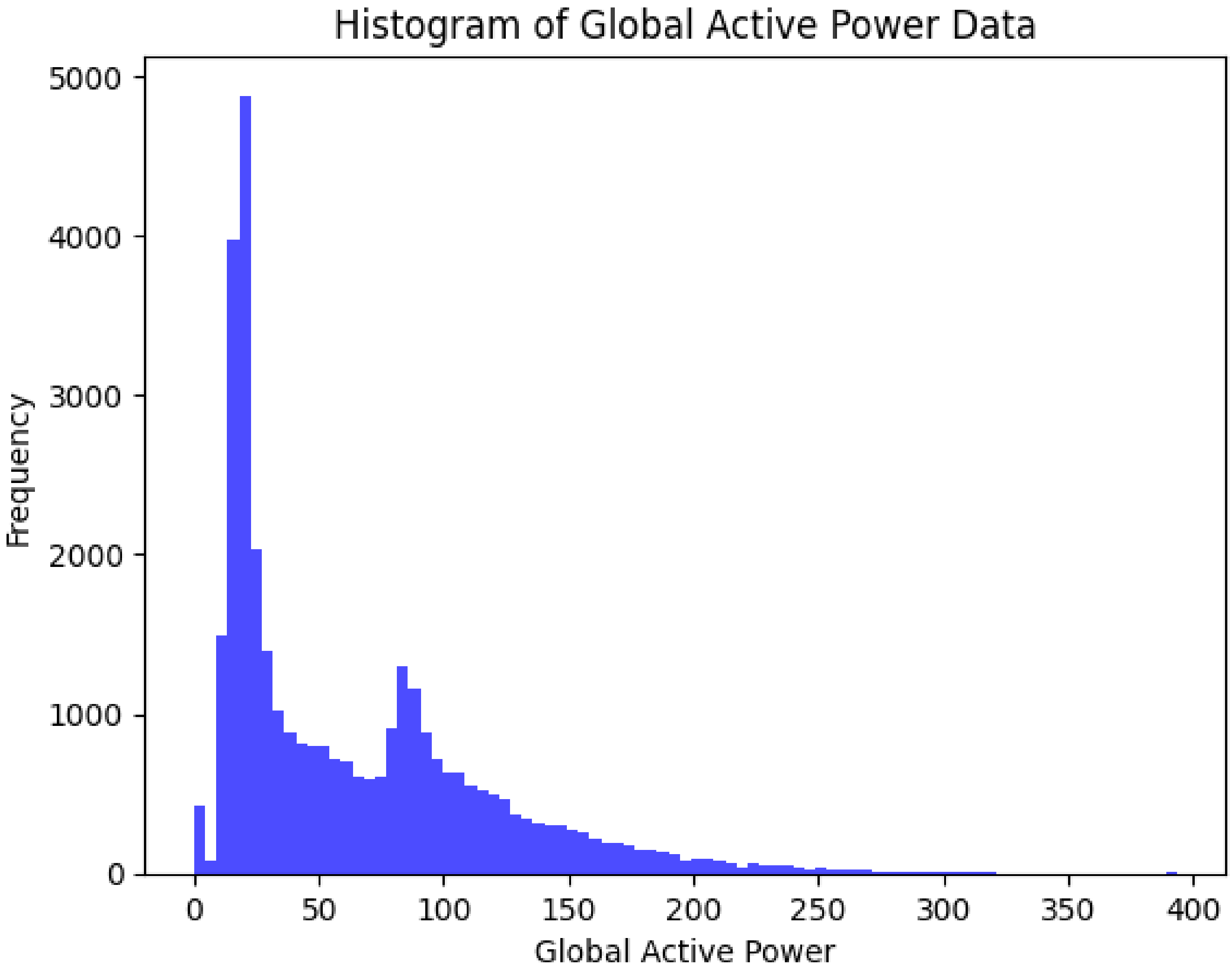
DATA SOURCE & CLEANING

- Data Source: "Individual household electric power consumption" dataset downloaded from the UCI Machine Learning Repository.
- Data gathered in a city neared Paris, France between December 2006 and November 2010 (47 months)
- Converted data types to numerical and DateTime
- 1.25% missing data that was dropped
- Resampled data from minutes to hourly
- Featured Engineering for further analysis

Feature	Type	Description
DateTime	datetime	The timestamp of the recorded data
Global_active_power	float	The total active power consumed by the household (in kilowatts)
Year	integer	The year when the data was recorded
Quarter	integer	The quarter of the year when the data was recorded (1 = Jan-Mar, 4 = Oct-Dec)
Month	integer	The month when the data was recorded (1 = January, 12 = December)
Day	integer	The day of the month when the data was recorded (1 - 31)
Time_of_day	integer	The hour of the day when the data was recorded (0 - 23)
Day_of_week	integer	The day of the week when the data was recorded (1 = Monday, 7 = Sunday)
Season	integer	The season when the data was recorded (1 = Spring, 2 = Summer, 3 = Fall, 4 = Winter)

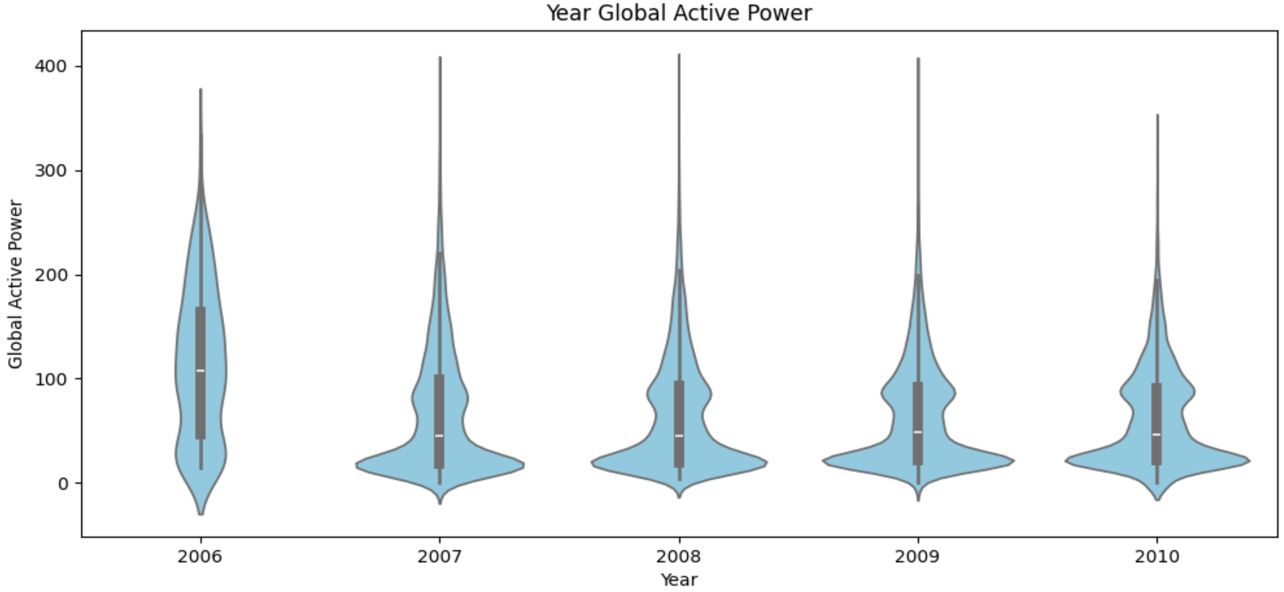
EXPLORATORY DATA ANALYSIS (EDA)

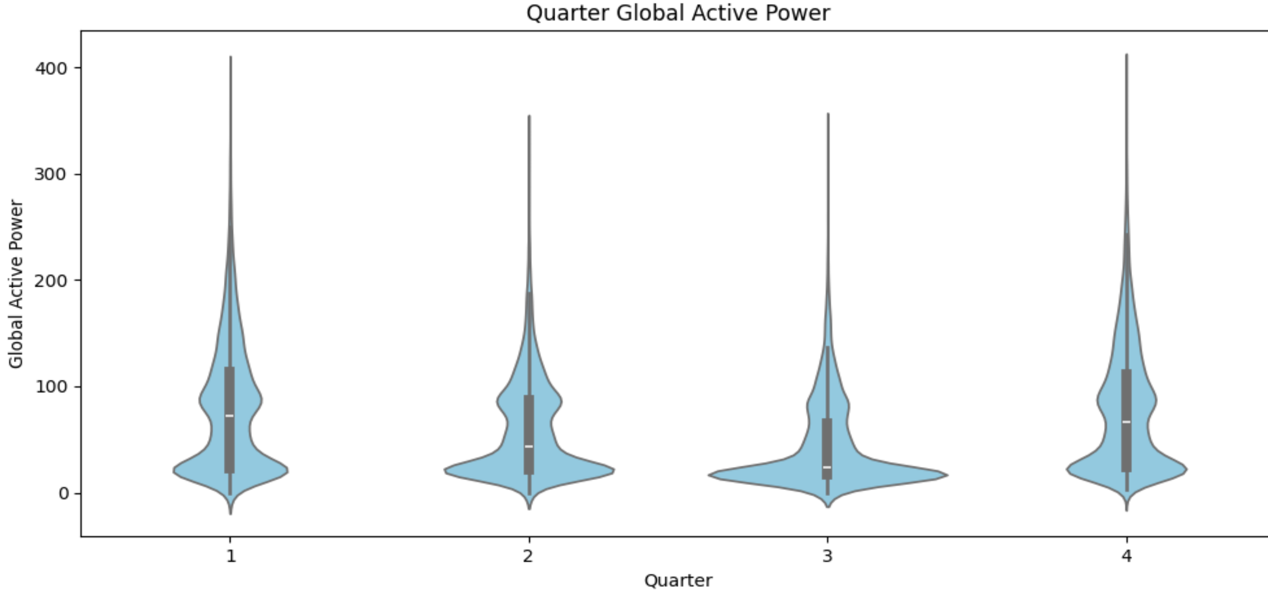
- Global active power is the total amount of electrical power consumed over a given period of time.
- Global active power is not normally distributed

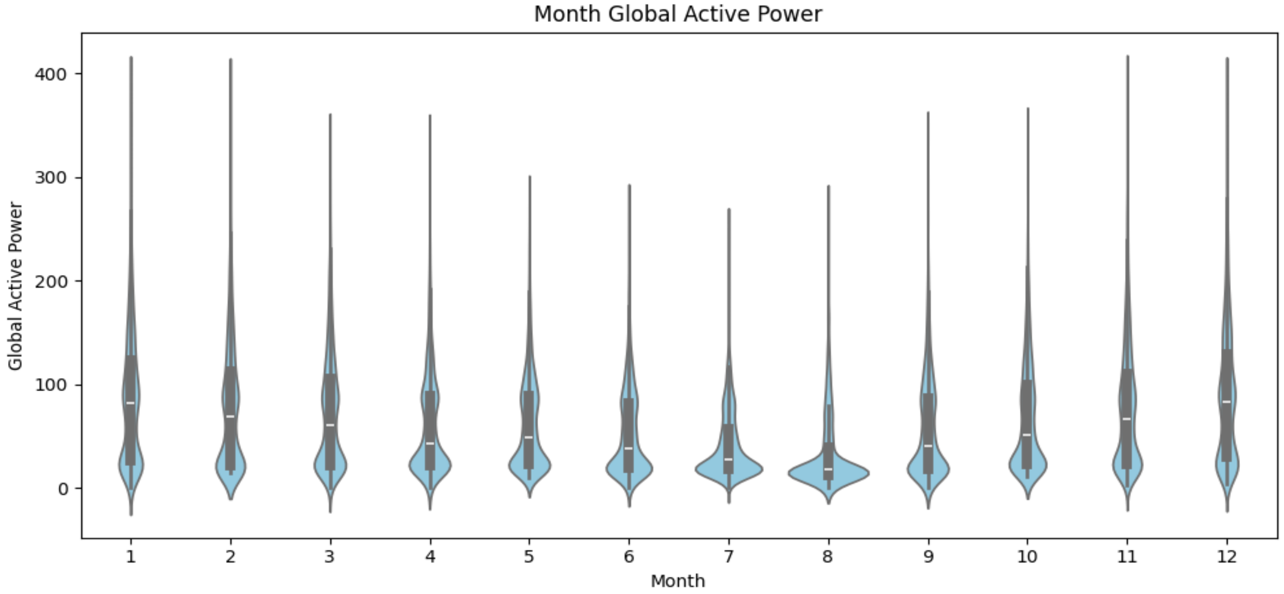


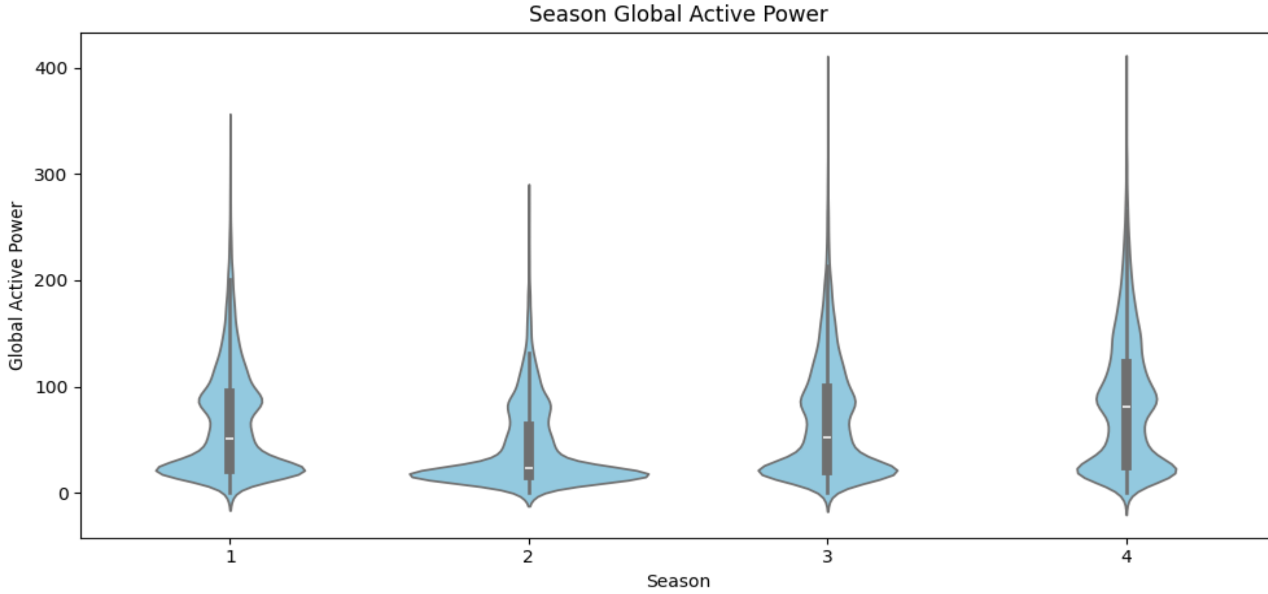
EXPLORATORY DATA ANALYSIS (EDA)

- Data transformed using feature engineering

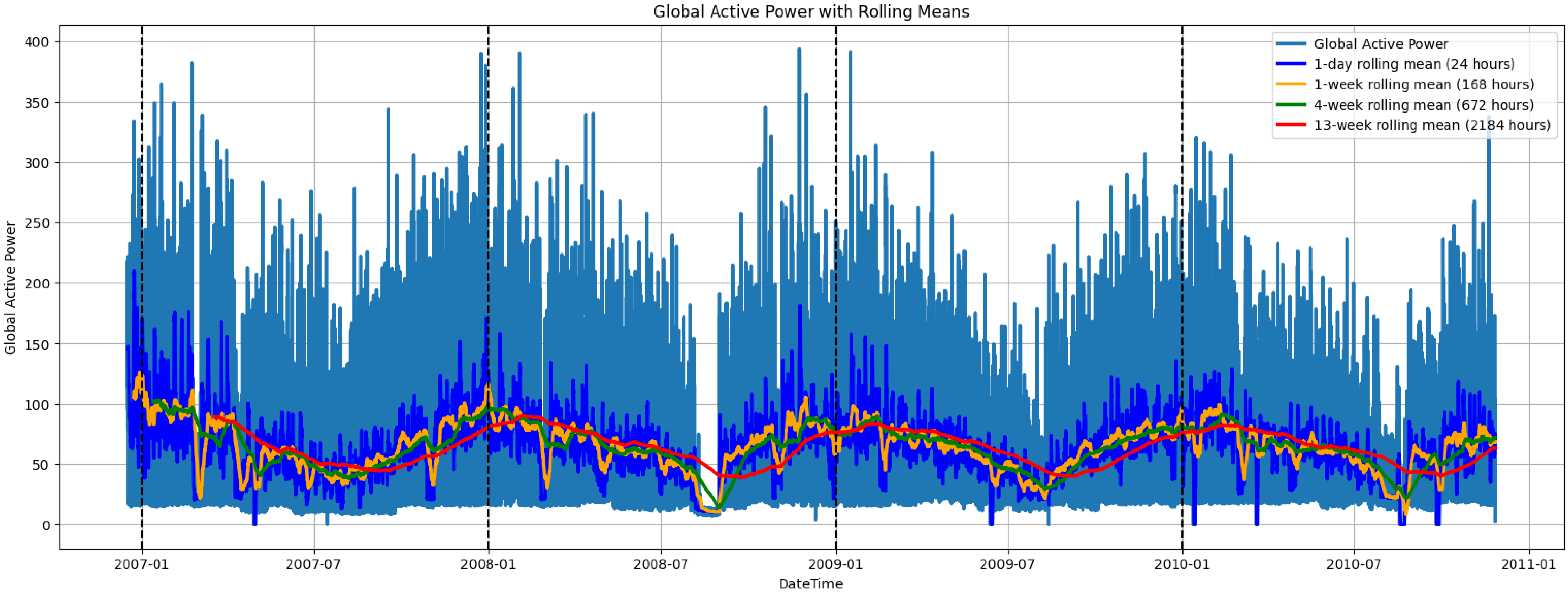






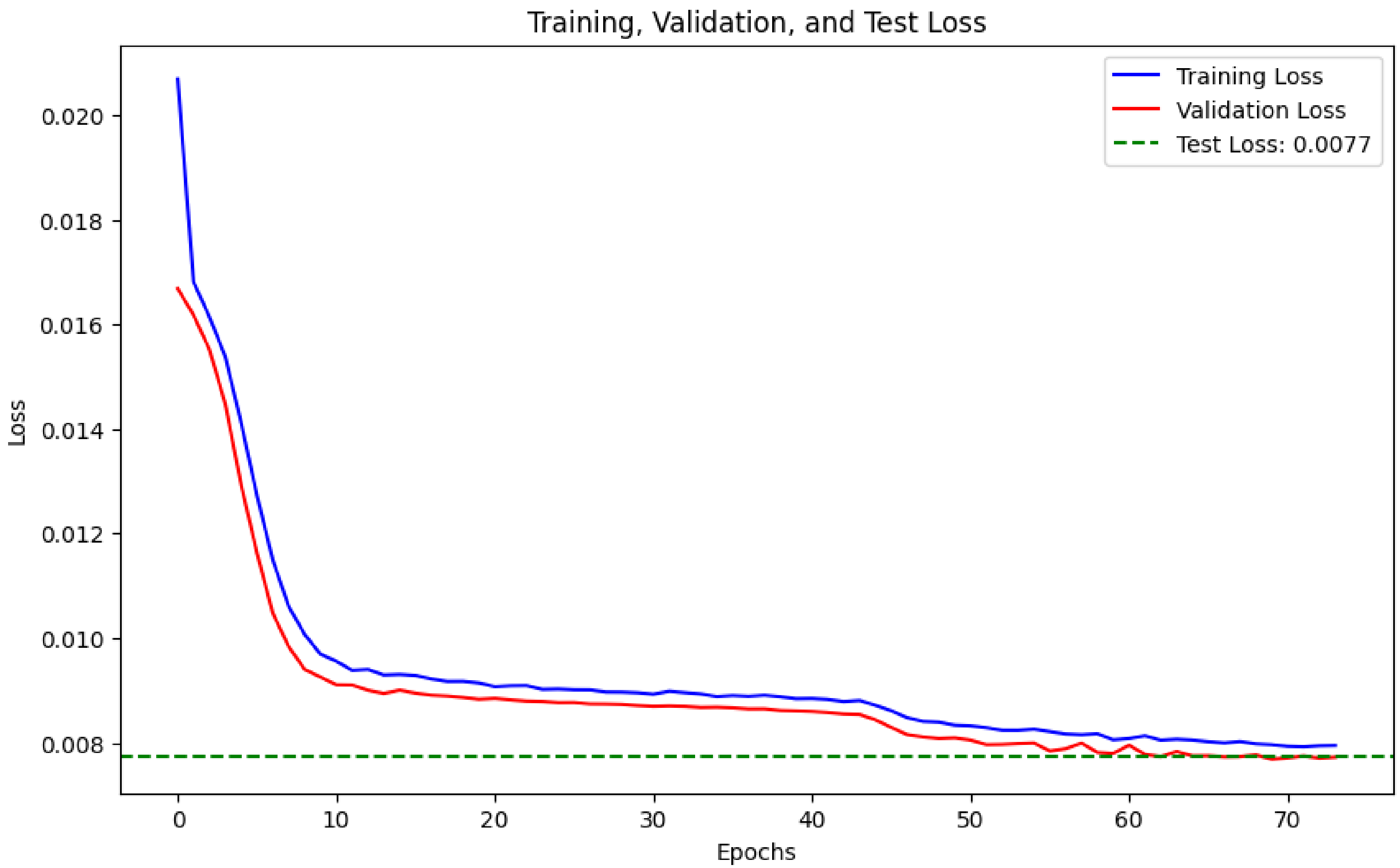


ACTIVE POWER CONSUMPTION WITH ROLLING MEANS

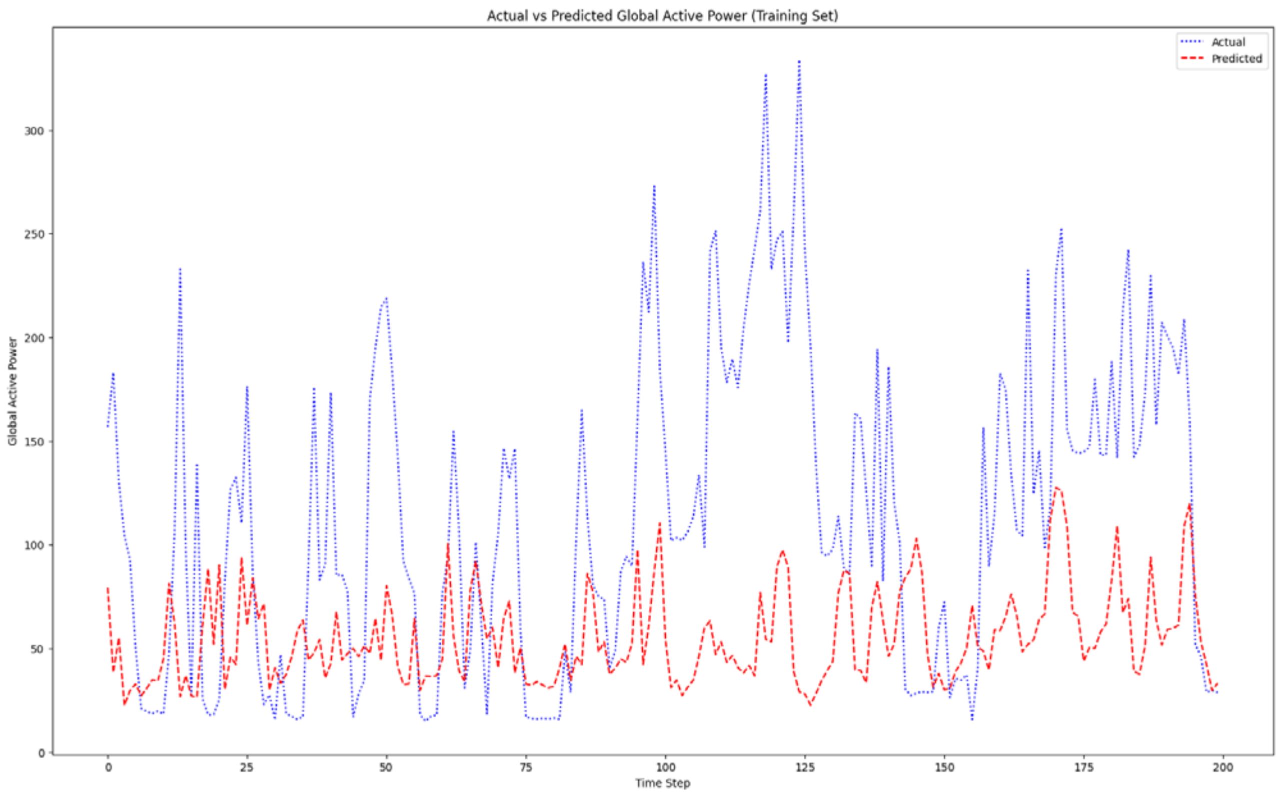


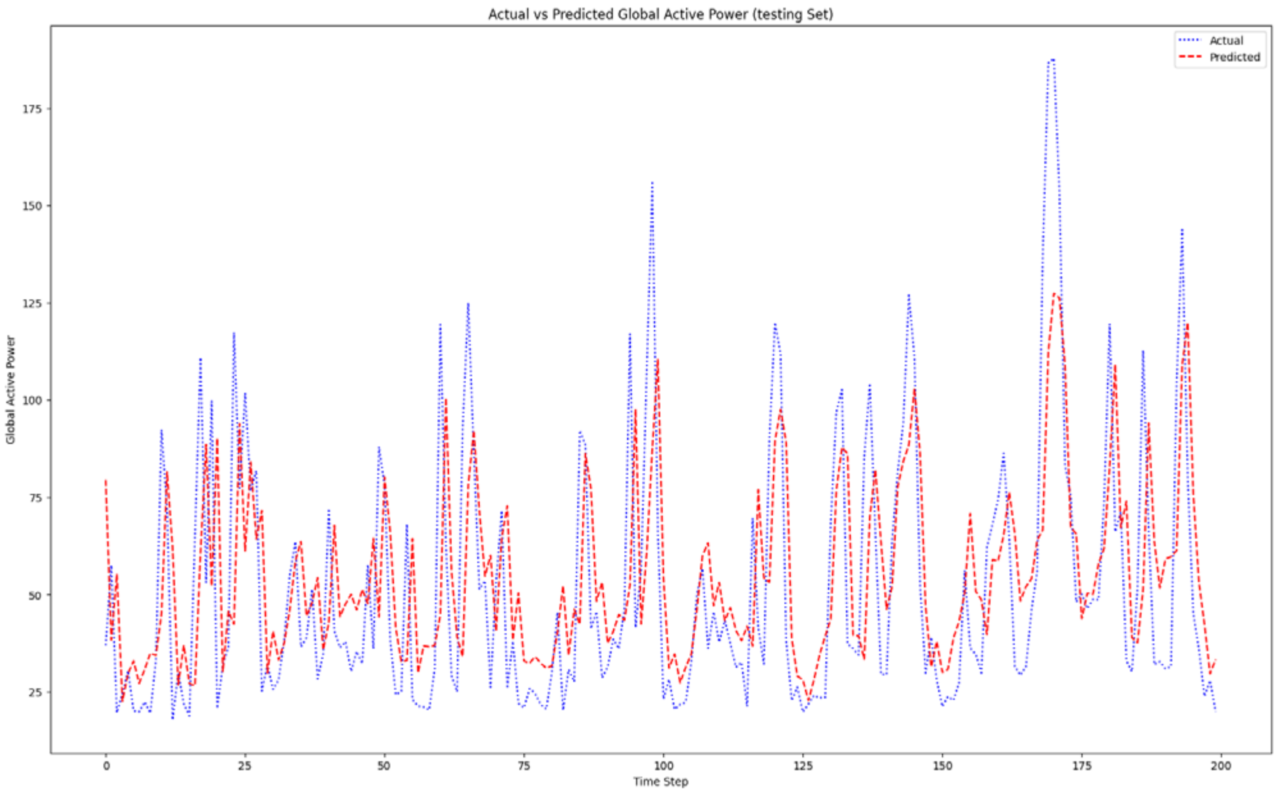
LSTM MODEL DEVELOPMENT

- Train LSTM architectures using historical energy consumption data
- optimize with techniques like dropout and early stopping.
- Evaluate model performance (MAE, MSE) on training and testing data



LSTM MODEL ACTUAL VS PREDICTION

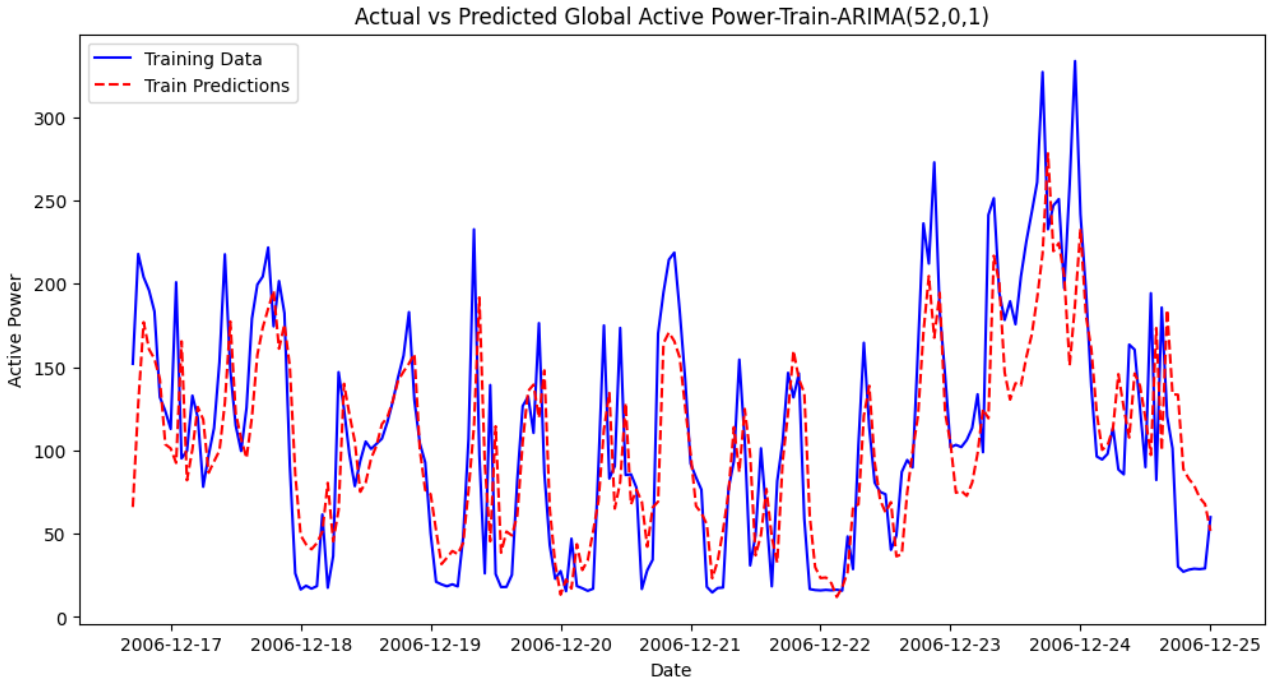


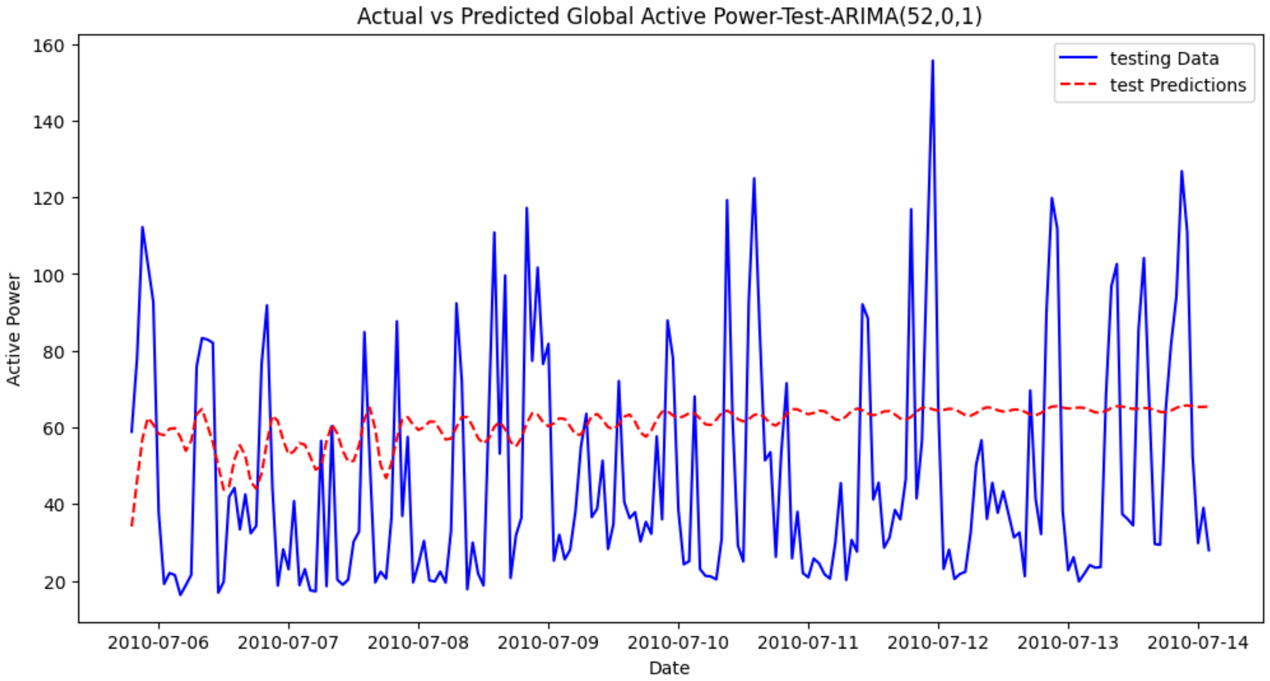


MODEL COMPARISON

- **LSTM Model:**
- MAE-Train: 25.26, & Test: 20.90
- MSE-Train: 1274.81, & Test: 856.86
- Consistent performance on training and test data suggests good generalization.
- **ARIMA (2,0,2):**
- Training MSE: 1445.72
- Test MSE: 2155.50
- Higher test error indicates overfitting to training data.
- **ARIMA (52,0,1):**
- Training MSE: 1253.20
- Test MSE: 2141.75
- Slightly better than (2,0,2) but still shows overfitting.

ARIMA MODEL(52,0,1) ACTUAL VS PREDICTION





CONCLUSION & RECOMMENDATIONS

- LSTM model outperforms ARIMA in forecasting energy consumption.
- Continuously update and refine the model with additional data sources for improved accuracy
- Larger timeframe evaluation

THANK YOU

References:

- **Individual Household Electric Power Consumption**. UCI Machine Learning Repository. Hebrail, Georges & Berard, Alice. (2012). Individual Household Electric Power Consumption. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C58K54