**Project Report**

**Project Title**

**Energy Consumption Prediction using Regression Analysis**

**Team Leader:**

SHIVANOOR VIGNESH 21R11A0597

**Team Members:**

G . Sandhya Rani 21R11A0569

P. Sreekaree 21R11A6743

PSBN Sriya 21R11A0595

N.Sai Charan 21R11A0588

MD.Akhbar Ali 21R11A0586

**Exploratory Data Analysis**

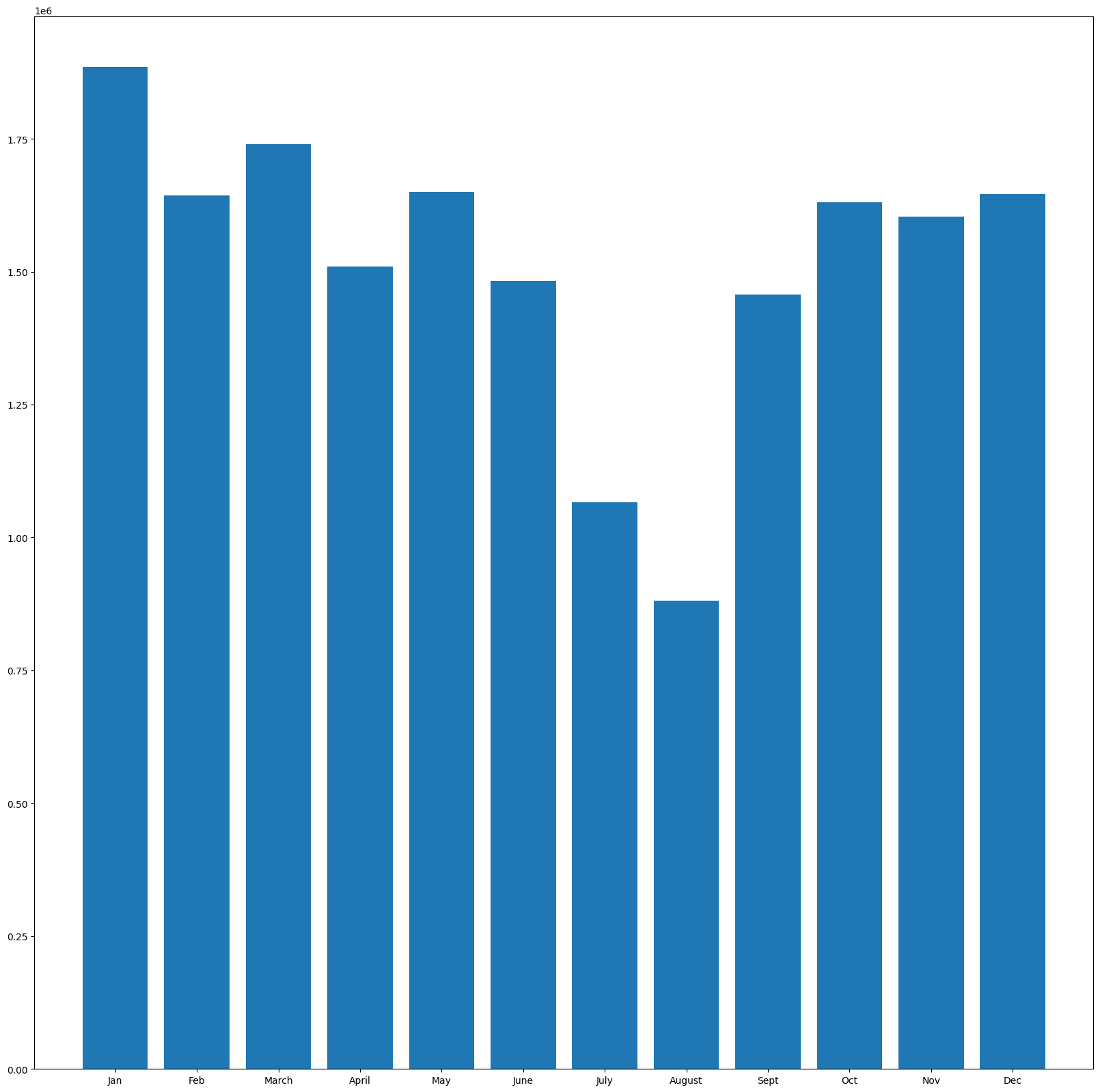
**Import data set**

For importing the dataset and to perform Exploratory Data Analysis we have to import some packages or library which are essential.

* import pandas
* import NumPy
* import seaborn

import matplotlib

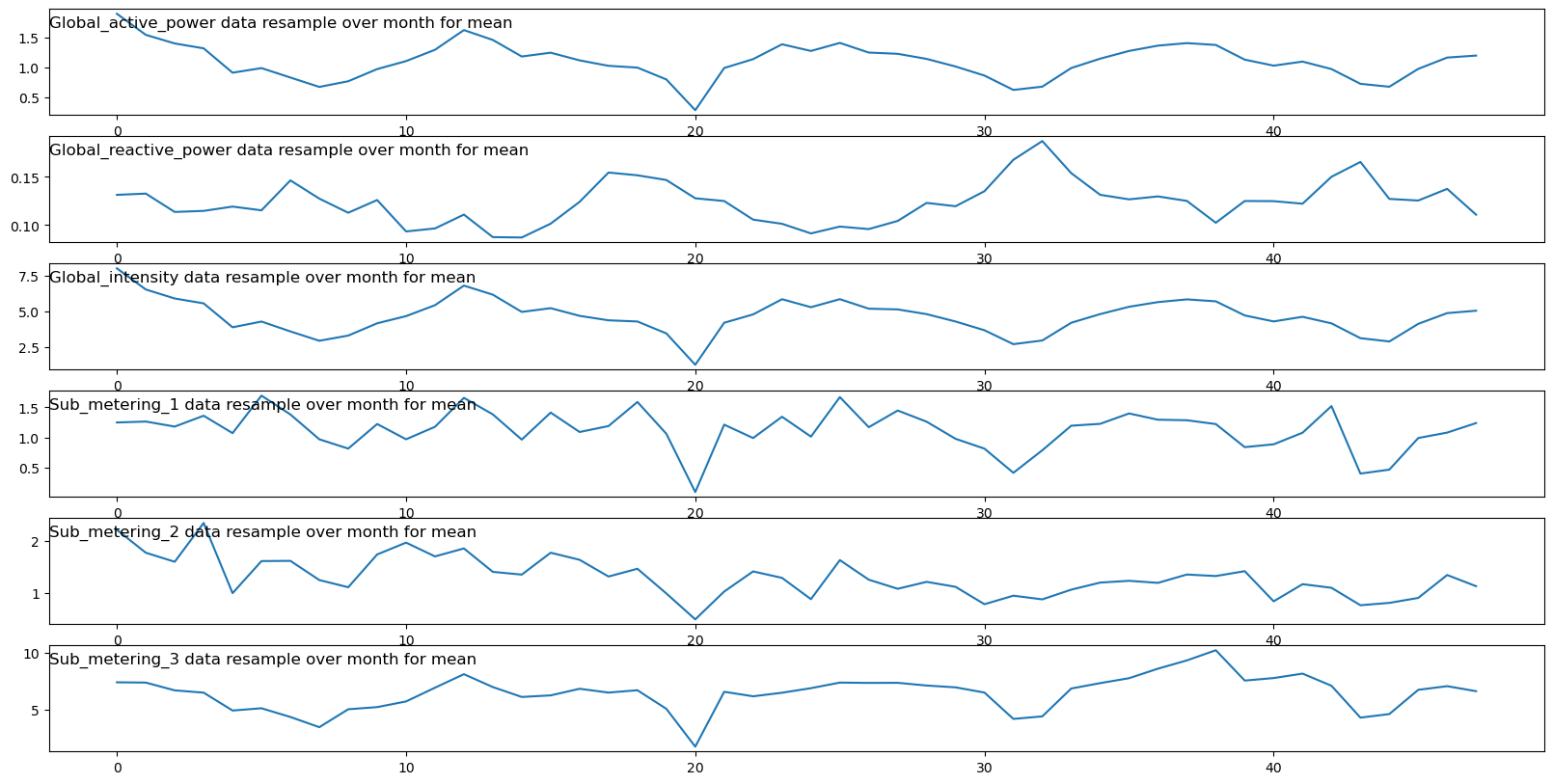
1. **MONTHLY VALUES**



Observations:

1. As we can see in the above graph January month has high usage of the electricity
2. Lowest used electricity was used in august

**Monthly mean values**

****

Observation:

1. The monthly mean values of every attribute, like global active power, are lower during the 20th month and higher at the beginning.

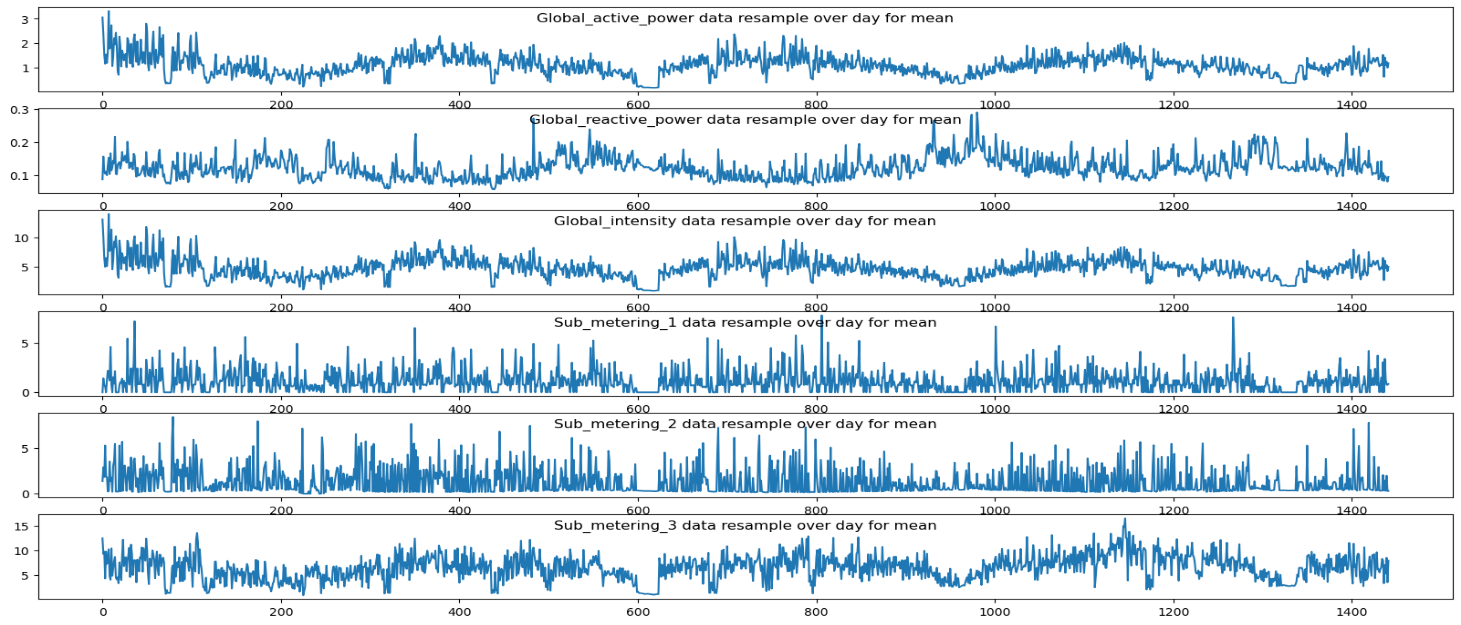
2. The monthly mean values of the column global reactive power are higher after the 30th month.

3. The monthly mean values of global intensity are lower only for the 20th month.

4. The monthly mean values of the column sub-metering 1 are higher except for the 20th month.

5. The monthly mean values of the columns sub-metering 2 and sub-metering 3 are also similar to sub-metering 1.

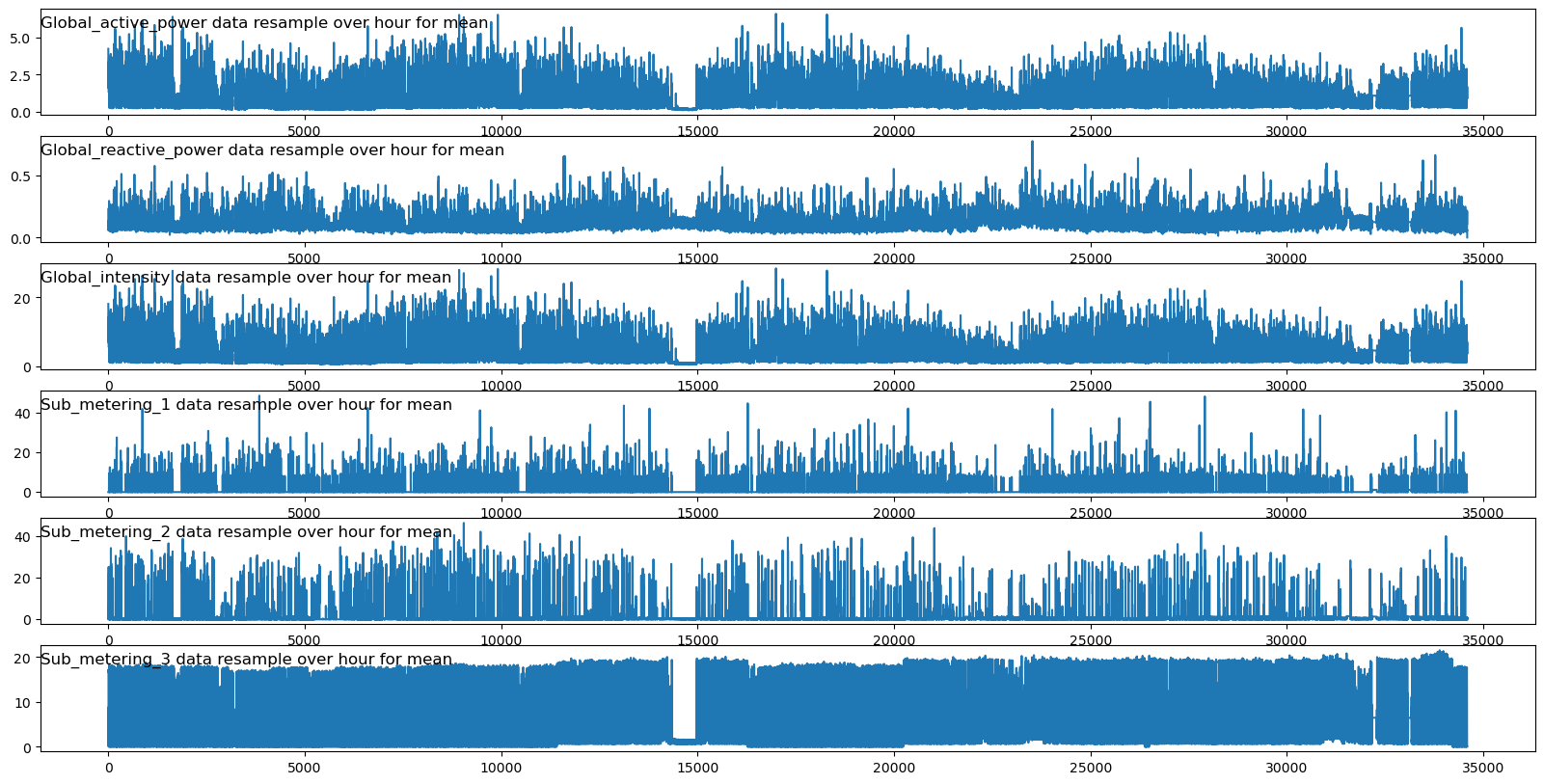
**Daily Readings of Global reactive ,active,intensity,submetering\_1,submetering \_2,submetering\_3**



**Observations:**

1.Showing daily average values for a different variable in a dataset, and revealing patterns and trends over time.

**Hourly mean values:**



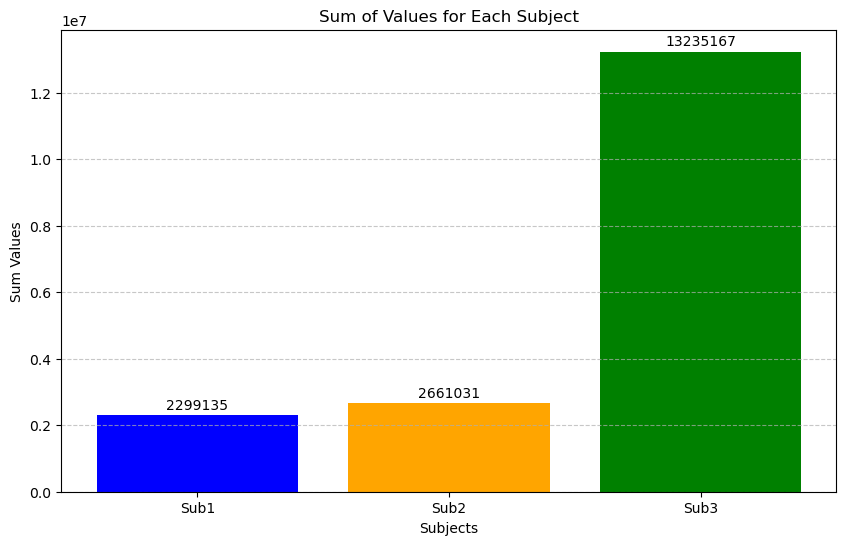
Observation:

1.The multiple line graphs global active powers, global reactive power, global intensity ,submetering\_1,

Submetering\_2,submetering\_3 have different values in different time period

2.it has submetering \_3 has almost equal in the every hour except 1500th hour

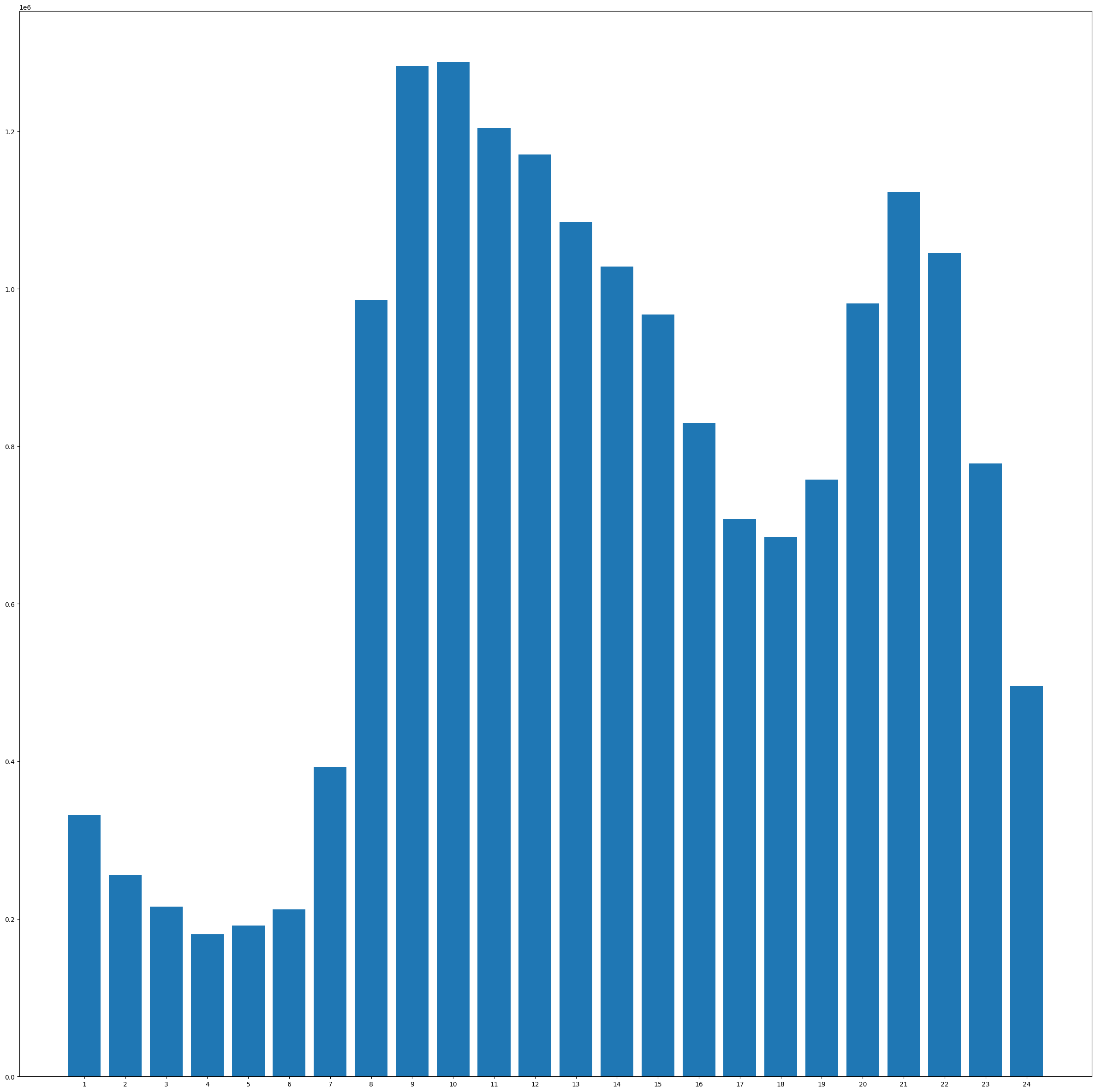
**Sum Values of the each Submetering :**



Observation:

1. From the above graph we can conclude that all the submetering3 has highest
2. While compared to submetering1 lower compared to submetering 1

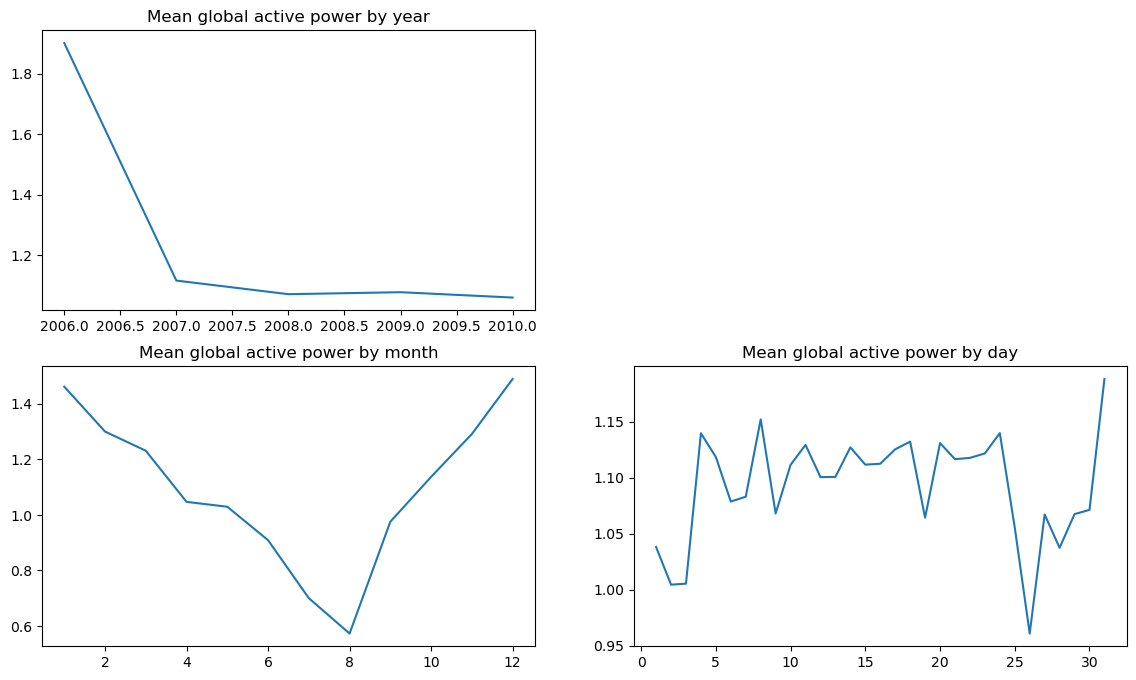
**Hourly consumption**



Observation:

1. From the above graph we can conclude that most of the consumption during the 10Th hour and most of the consumption was during 9th , 10th hour only may be because in every house at time of 9 every house complete their work and watch movies in their house

**Mean values of Global active Power By day, hour, Month:**

****

**Observation:**

1.There is sudden down fall occurred in the year of 2007 and it continued from 2007 to 2010

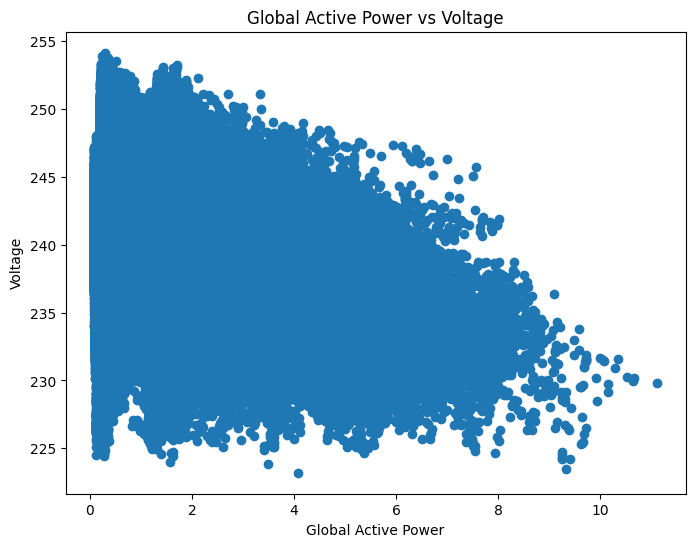
2. There is sudden fall occurred when during 8th month and its gradually increased from then to

12th month

3.There is sudden fall in the days between end of the month of day 25th and took almost 5 days to

Increase

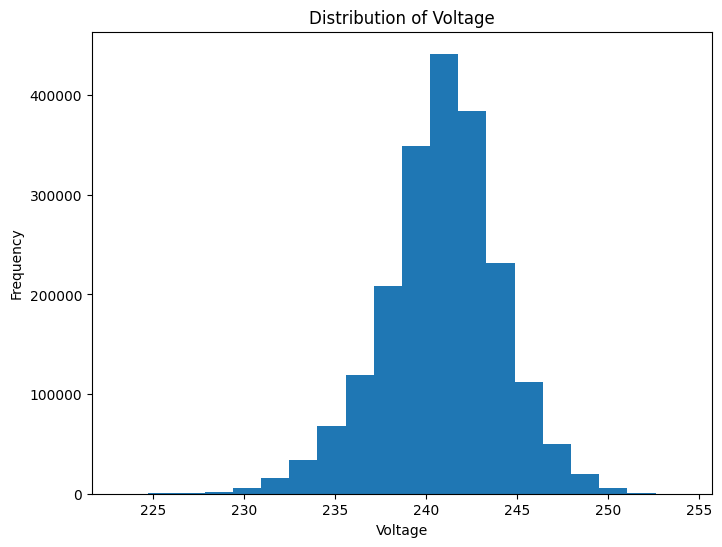
**Global Active power vs Voltage**



**Observation**

1. From the above graph, we can observe that when the global active power is between 0 and 2, there is high voltage present.
2. We can see that when the global active power reaches 8-10, low voltage is produced

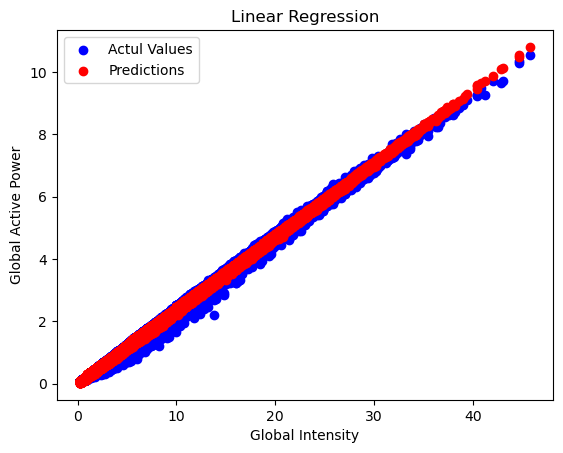
**Distribution Of Voltage:**



Observation:

1. From the above histogram provides insights into the distribution of voltage values. It can help identify the most common voltage levels and any outliers in the data.

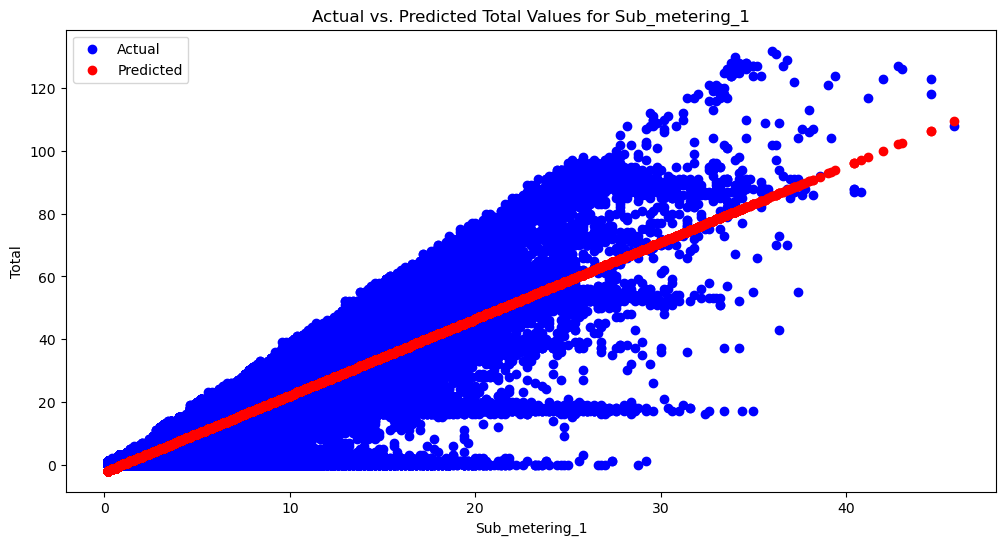
**Prediction Vs Actual Values**



Observation:

1.we can see that graph of predicted values of the global active power and global intensity which is compared to actual values of it

**Actual values vs predicted values of Global intensity:**

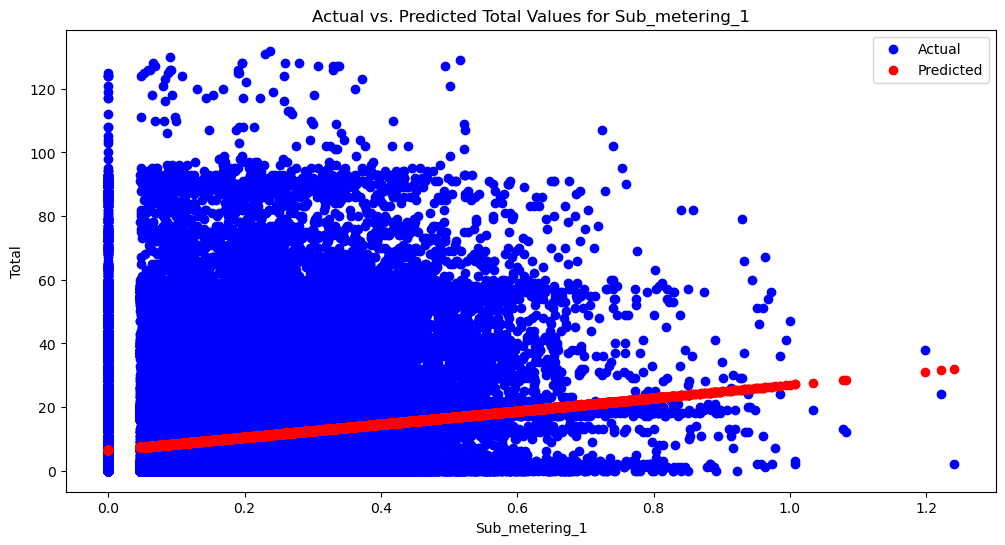


Observation:

1.from the above graph we can say that predicted values are linearly increasing

The global intensity

**Actual values vs global reactive power:**



Observation:

1.From above graph between the actual values and predicted values bule one is actual values and red colour one is predicted

**Voltage Varying Throughout the year**

# 

**Observation:** **Month**

**1 242.154158**

**2 241.459095**

**3 241.520813**

**4 241.040326**

**5 239.008586**

**6 239.930387**

**7 240.004303**

**8 240.204213**

**9 240.464033**

**10 240.777288**

**11 241.042169**

**12 242.748110**

**Observation :**

* The mean value of Voltage is **240.839 in volt**
* From the above mean values the consumption of voltage is less in the month of May comparing to all other months
* From the above mean values the consumption of voltage is very high in the month of December comparing to all other months

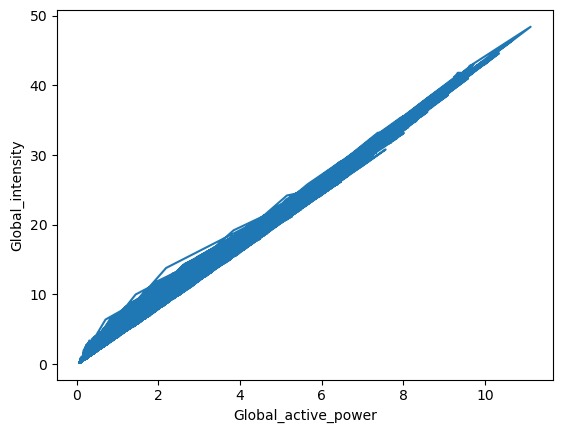
**Heat map for Numeric columns**

# 

**Observation:**

* As we can observe from the above heat map the relation between global active power and global intensity is directly proportional
* Also we can say that global active power and sub\_metering 3 are co-related

**Co-relation between Global Intensity and Global Active Power**



1. **Observation :**

* The Co-relation value between global intensity and global active power is 0.998
* This depicts that these two factors are proportional

**Suggestion and finding**

* Monthly Patterns: The monthly mean values of global active power, global reactive power, voltage, and global intensity show different patterns over the months. January has the highest usage of electricity, while August has the lowest.
* Daily and Hourly Patterns: There are distinct patterns in the daily and hourly readings of global active power, global reactive power, and global intensity. These patterns can provide insights into peak consumption hours and daily trends.
* Submetering Analysis: The submetering data shows that submetering 3 consistently has the highest values, followed by submetering 1 and submetering 2. This suggests that certain appliances or areas in the system are consistently consuming more energy.
* Voltage and Power Relationship: The relationship between global active power and voltage indicates that high voltage is present when the global active power is between 0 and 2, while low voltage is produced when the global active power reaches 8-10.
* Prediction Accuracy: The comparison of actual values with predicted values shows that the model's predictions are generally in line with the actual values, indicating a reasonable level of accuracy.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | | Global reactive power | Global active power | Global Intensity |
| Findings  from  graphs | | 1.The monthly mean values of the column global reactive power are higher after the 30th month. | 1.monthly mean values  2.yearly trends  3.voltage and correlation  4.There is sudden fall occurred when during 8th month and its gradually increased from then to 12th month  5.We can see that when the global active power reaches 8-10, low voltage is produced | 1. The monthly mean values of global intensity are lower only for the 20th month. |
| Solutions/Efficency | 1. Seasonal Adjustments and Infrastructure Review: Consider adjusting energy production and distribution strategies to align with seasonal consumption patterns, while also reviewing infrastructure for efficiency and reliability.  2. Voltage Optimization and Predictive Maintenance: Focus on optimizing voltage regulation systems based on observed voltage-power relationships, and implement predictive maintenance strategies to prevent sudden drops in energy consumption. | | | |

s