MEASURE ENERGY CONSUMPTION

NAME : LINGESH G

REG NO: 61772221L02

Phase 2 Submission Document



Introduction:

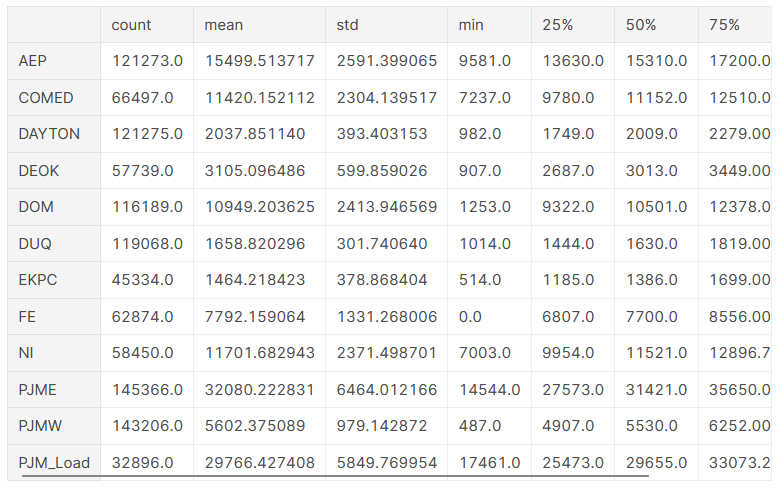
* Measuring energy consumption is a fundamental practice in understanding how energy is utilized in various contexts, whether it's in **homes, businesses, industries, or transportation**.
* The world faces a pressing challenge **in managing its energy resources efficiently and sustainably**. As the global population grows, so does the demand for energy, leading to increased consumption and environmental concerns.
* To address these issues, the integration of artificial intelligence (AI) into energy management has become crucial.
* Measuring energy usage **helps individuals, organizations, and governments effectively manage their energy resources**. It provides insights into where and how energy is being used.
* Accurate measurement enables cost tracking and management, allowing individuals and businesses to identify areas where energy-saving measures can lead to financial savings.
* Monitoring energy consumption is crucial for assessing and reducing environmental impacts. It **allows for tracking carbon emissions and evaluating the ecological footprint** associated with energy use.
* Measuring energy consumption is a key step in improving energy efficiency. It helps identify inefficiencies and opportunities for optimization.

Data Source :

To measure energy consumption accurately, you need reliable and relevant data sources. The choice **of data sources depends on the specific context and goals** of your energy consumption measurement project.

**Dataset Link** :

<https://www.kaggle.com/code/robikscube/starter-hourly-energy-consumption>



Data Preprocessing :

Data pre processing is a crucial step when **working with energy consumption data**, as it helps ensure that the data is clean, structured, and ready for analysis.

Model Develoment :

Developing a model to measure energy consumption **involves building a predictive or descriptive model that can estimate or explain energy** usage based on various factors.

Clearly define the problem you want to solve with the energy consumption model. For instance, you might want to predict energy usage for a building, optimize energy efficiency, or detect anomalies in consumption.

Ensure that you have pre processed and cleaned the energy consumption data as described in the previous answer.

Choose the appropriate modeling approach based on the nature of the problem. Common approaches include

:

* Regression: For predicting energy consumption based on input features.
* Time Series Analysis: If your data is time-dependent, you may need to use time series forecasting techniques.
* Machine Learning: Utilize various machine learning algorithms, such **as linear regression, decision trees, random forests, or neural networks.**
* Anomaly Detection: For identifying unusual consumption patterns.
* Optimization: To find the most energy-efficient strategies.

Model Evaluation:

Choose evaluation metrics that are relevant to your specific energy consumption problem. Common metrics include:

* Mean Absolute Error (MAE): Measures the average absolute difference between actual and predicted values.
* Root Mean Squared Error (RMSE): Penalizes large errors more than MAE and is **commonly used for regression problems**.
* Mean Absolute Percentage Error (MAPE): Measures the percentage **difference between actual and predicted values**.
* R-squared (R^2): Indicates the proportion of variance explained by the model.
* Coefficient of Determination (COD): Similar to R-squared, providing insights into how well the model fits the data.
* F1-score or precision and recall for classification problems (e.g., anomaly detection).

Program:

**MEASURE ENERGY CONSUMPTION**

**import pandas as pd**

**import matplotlib.pyplot as plt**

**df = pd.read\_csv('your\_dataset.csv')**

**df['Timestamp'] = pd.to\_datetime(df['Timestamp'])**

**df.set\_index('Timestamp', inplace=True)**

**print(df.head())**

**# Calculate basic statistics**

**mean\_consumption = df['EnergyConsumption (kWh)'].mean()**

**max\_consumption = df['EnergyConsumption (kWh)'].max()**

**min\_consumption = df['EnergyConsumption (kWh)'].min()**

**print(f"Mean Energy Consumption: {mean\_consumption:.2f} kWh")**

**print(f"Maximum Energy Consumption: {max\_consumption:.2f} kWh")**

**print(f"Minimum Energy Consumption: {min\_consumption:.2f} kWh")**

**# Plot hourly energy consumption**

**plt.figure(figsize=(12, 6))**

**plt.plot(df.index, df['EnergyConsumption (kWh)'], linestyle='-', marker='o', markersize=4)**

**plt.title('Hourly Energy Consumption')**

**plt.xlabel('Time')**

**plt.ylabel('Energy Consumption (kWh)')**

**plt.grid(True)**

**plt.xticks(rotation=45)**

**plt.tight\_layout()**

**plt.show()**

**Note Book :**

**In [1] :**

import matplotlib.pyplot as plt # plotting

import numpy as np # linear algebra

import os # accessing directory structure

import pandas as pd # data processing, CSV file I/O (e.g. pd.read\_csv)

import seaborn as sns

plt.style.use('ggplot') # Make it pretty

**In [2] :**

# Data is saved in parquet format so schema is preserved.

**df = pd.read\_parquet('../input/est\_hourly.paruqet')**

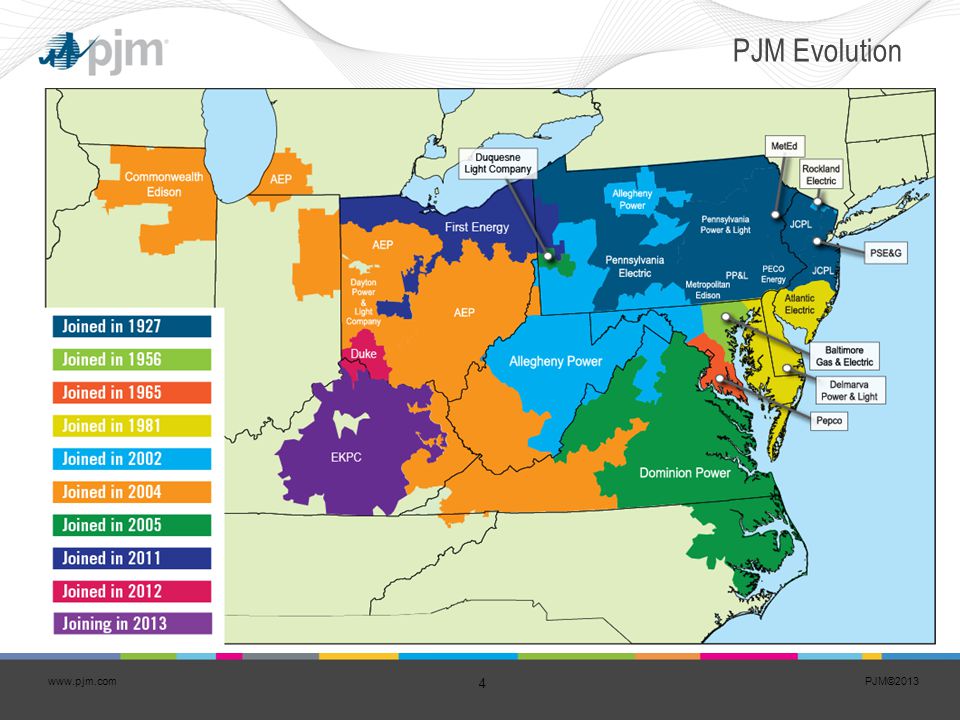
**In [3] :**

**#Show PJM Regions**

**from IPython.display import Image**

**Image(url= "http://slideplayer.com/4238181/14/images/4/PJM+Evolution.jpg")**

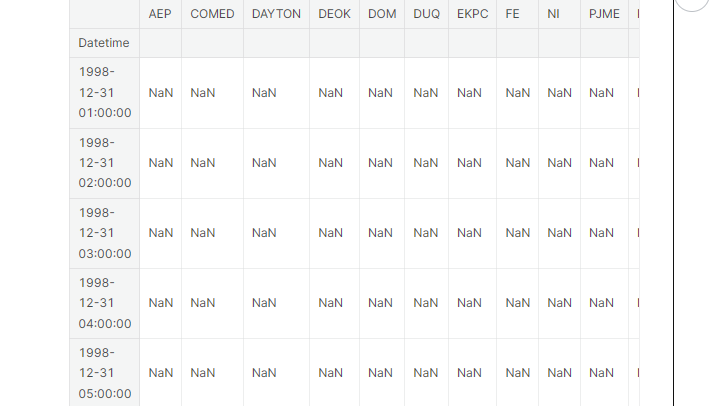
**Out [3] :**



**In [4] :**

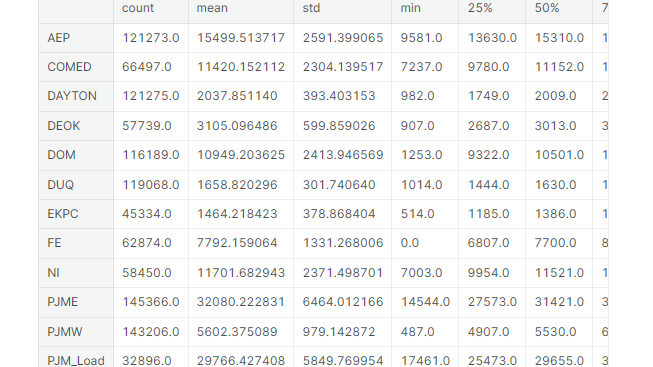
df.head()

out [4] :

****

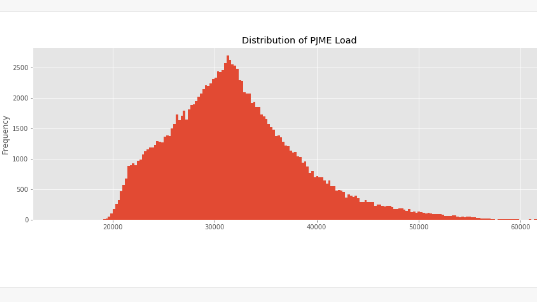
**In [5] :**

df.describe().T

out [5] : 

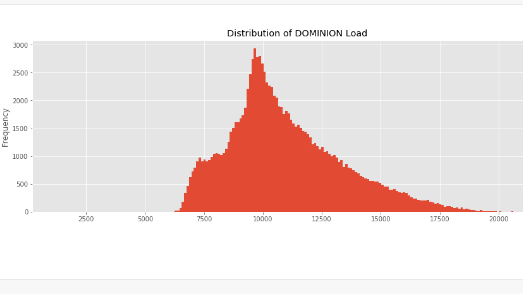
In [6] :

\_ = df['PJME'].plot.hist(figsize=(15, 5), bins=200, title='Distribution of PJME Load')

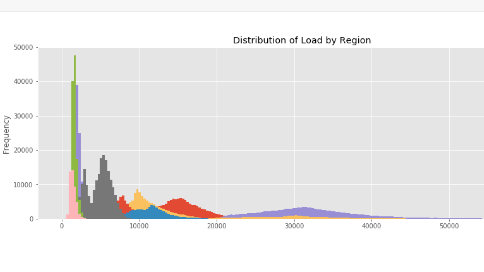


In [7] :

\_ = df['DOM'].plot.hist(figsize=(15, 5), bins=200, title='Distribution of DOMINION Load')

****

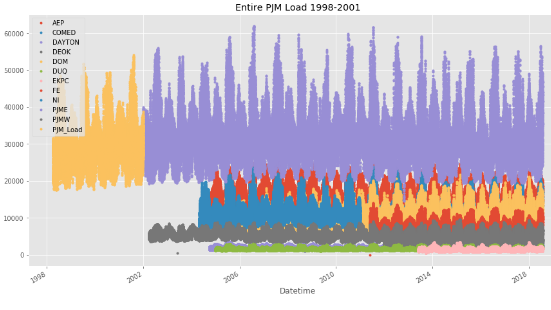
**In [8] : \_**

**= df.plot.hist(figsize=(15, 5), bins=200, title='Distribution of Load by Region') **

Plot Time Series

**In [9] :**

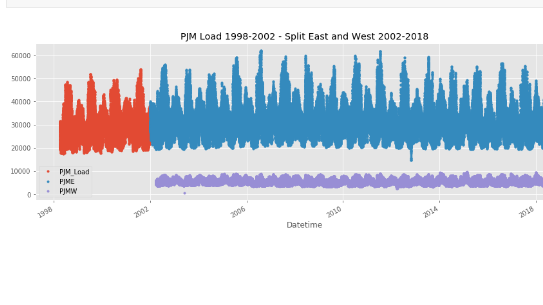
**plot = df.plot(style='.', figsize=(15, 8), title='Entire PJM Load 1998-2001')**

****

# **Plotting Regions**

In [10] :

JME','PJMW']] \ .plot(style='.', figsize=(15, 5), title='PJM Load 1998-2002 - Split East and West 2002-2018')

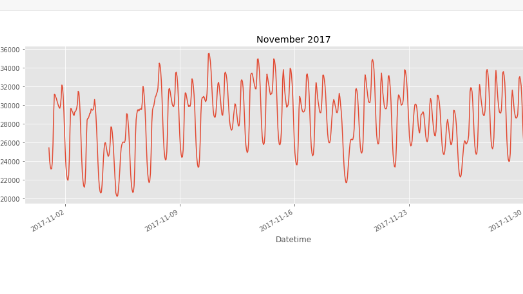


## Summer Demand vs Winter Demand

Note the dips mid-day in the winter months. Conversely in summer months the daily load is more bell shaped. This is due to high mid-day energy consumtion by air conditioning. In winter months people tend to use less energy mid-day.

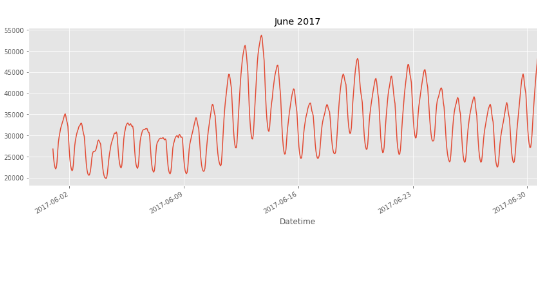
In [11] :

\_ = df[['PJM\_Load','P\_ = df['PJME'].loc[(df['PJME'].index >= '2017-11-01') & (df['PJME'].index < '2017-12-01')] .plot(figsize=(15, 5), title = 'November 2017')



In [12] :

\_ = df['PJME'].loc[(df['PJME'].index >= '2017-06-01') &(df['PJME'].index < '2017-07-01')] \.plot(figsize=(15, 5), title = 'June 2017')



Create Time Series Features

In [13] :

df['dow'] = df.index.dayofweek

df['doy'] = df.index.dayofyear

df['year'] = df.index.year

df['month'] = df.index.month

df['quarter'] = df.index.quarter

df['hour'] = df.index.hour

df['weekday'] = df.index.weekday\_name

df['woy'] = df.index.weekofyear

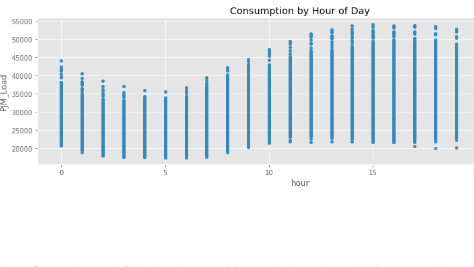
df['dom'] = df.index.day # Day of Month

df['date'] = df.index.date

In [14] :

\_ = df[['PJM\_Load','hour']].plot(x='hour’, y='PJM\_Load', kind='scatter',

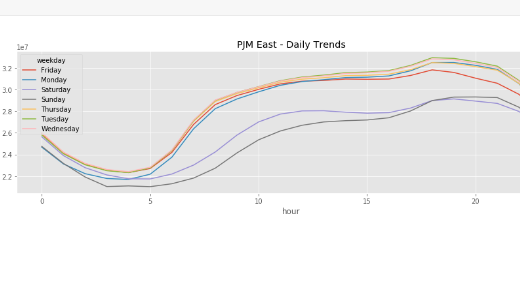
figsize=(14,4), title='Consumption by Hour of Day')

****

Note Saturday and Sunday demand is much less than during a work week. This is also true for holidays.

In [15] :

\_ = df.pivot\_table(index=df['hour'], columns='weekday', values='PJME', aggfunc='sum').plot(figsize=(15,4), title='PJM East - Daily Trends')

****

## Trends change depending on time of year

## In [16] :

## fig, ax = plt.subplots(figsize=(15,5))

## sns.boxplot(df.loc[df['quarter']==1].hour,df.loc[df['quarter']==1].PJME)ax.set\_title('Hourly Boxplot PJME Q1')

## ax.set\_ylim(0,65000)fig, ax = plt.subplots(figsize=(15,5))

## sns.boxplot(df.loc[df['quarter']==2].hour, df.loc[df['quarter']==2].PJME)ax.set\_title('Hourly Boxplot PJME Q2')

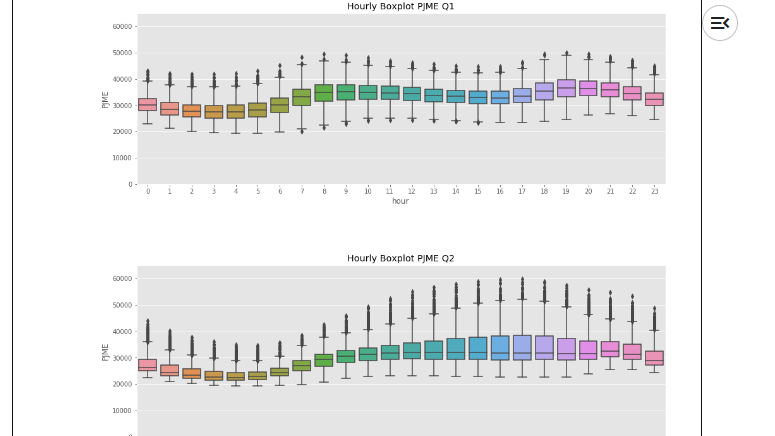
## ax.set\_ylim(0,65000)fig, ax = plt.subplots(figsize=(15,5))

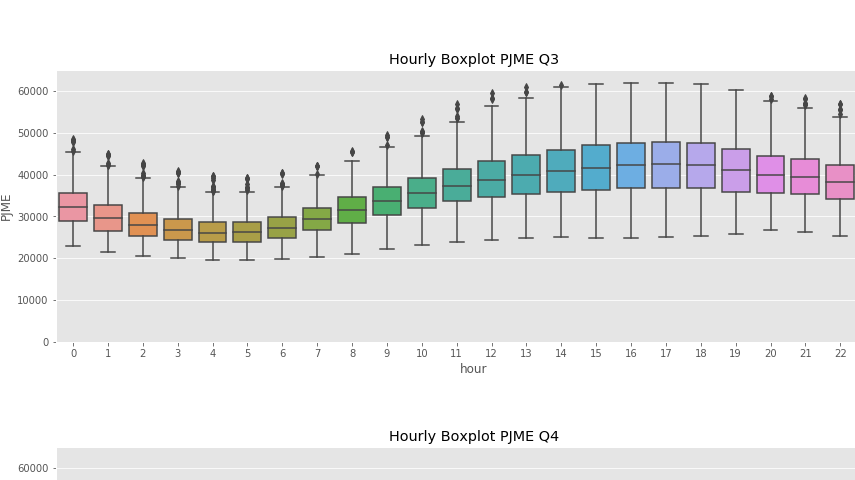
## sns.boxplot(df.loc[df['quarter']==3].hour, df.loc[df['quarter']==3].PJME)ax.set\_title('Hourly Boxplot PJME Q3')

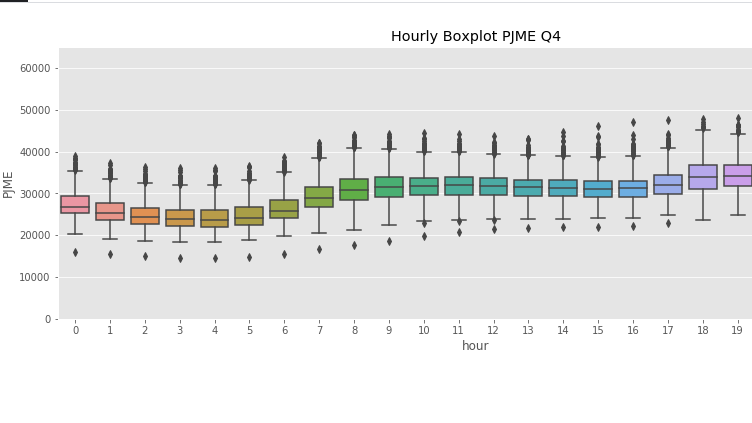
## ax.set\_ylim(0,65000)fig, ax = plt.subplots(figsize=(15,5))

## sns.boxplot(df.loc[df['quarter']==4].hour, df.loc[df['quarter']==4].PJME)

## ax.set\_title('Hourly Boxplot PJME Q4')\_ = ax.set\_ylim(0,65000)

****

****

****

**Conclusion:**

In conclusion, measuring energy consumption is a crucial task with **significant implications for sustainability, cost management**, and environmental impact reduction. It involves collecting, pre processing, and data to gain insights into how energy is used in various settings, from residential homes to industrial facilities.

Accurate measurements of energy consumption start with reliable data. **Data pre processing and cleaning are essential to ensure the data is accurate, complete, and free of errors**.

Building models to **measure energy consumption is essential for forecasting, anomaly detection, and optimizing energy usage**.