

MEASURE ENERGY CONSUMPTION

USING MACHINE LEARNING

Project title: Measure Energy Consumption

Phase 3: Development part 1

Topic: *Start building the measure energy consumption model by loading and pre-processing the dataset*

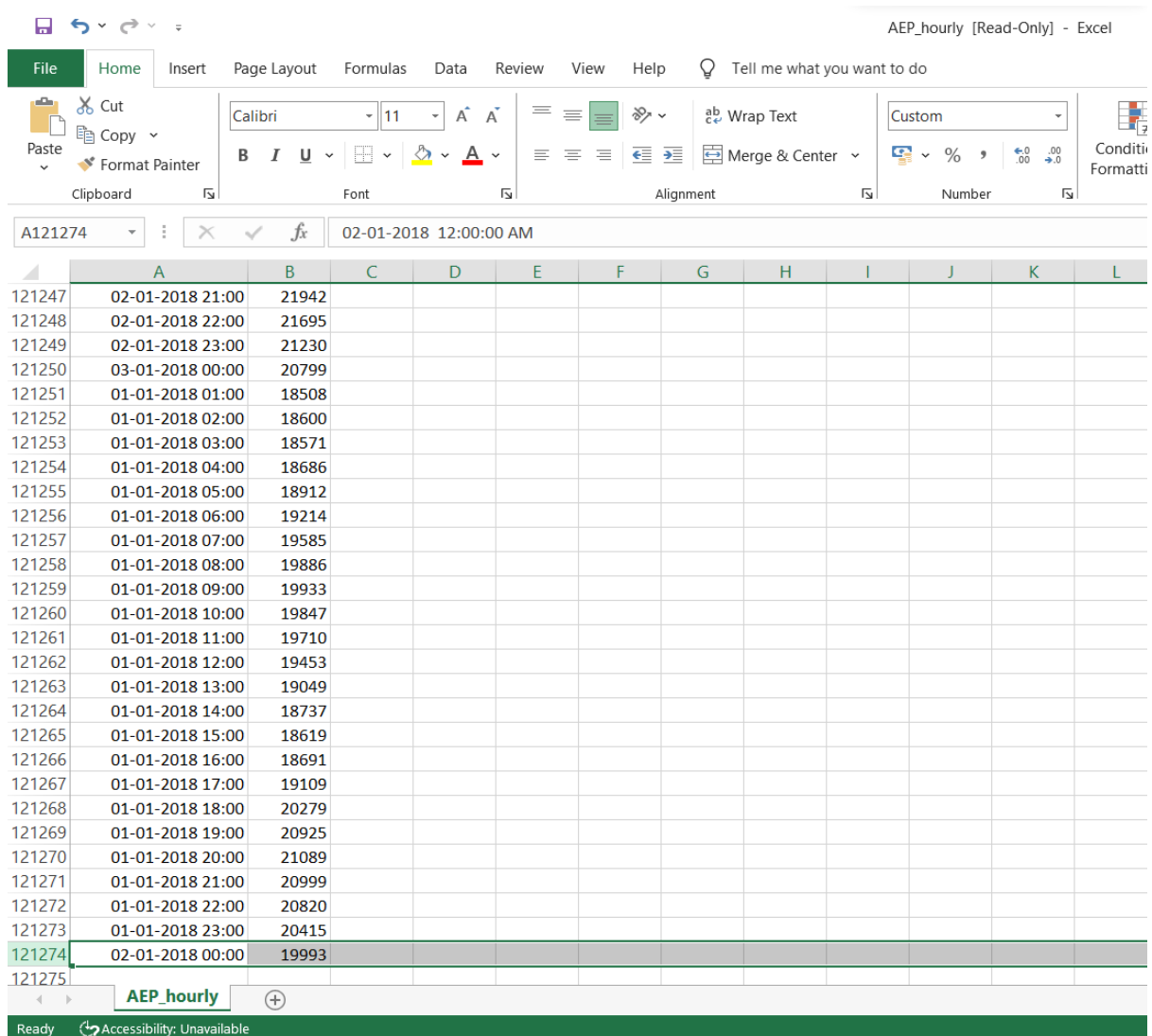
Measure energy consumption

Introduction:

Measuring energy consumption is a critical practice in today's world, as our reliance on energy sources continues to grow, and the environmental impact of our energy usage becomes increasingly apparent. Understanding and monitoring energy consumption is essential for a variety of reasons, including reducing costs, conserving resources, and mitigating the effects of climate change. This introduction will delve into the significance of measuring energy consumption and the various methods and tools used to do so.

Energy consumption measurement plays a pivotal role in our quest for sustainability and efficiency. It provides insights into how we use energy in our homes, businesses, industries, and transportation systems. By quantifying energy usage, we can identify areas where energy is wasted, make informed decisions to reduce consumption, and ultimately lower our carbon footprint.

Given data set:



	A	B	C	D	E	F	G	H	I	J	K	L
121247	02-01-2018 21:00	21942										
121248	02-01-2018 22:00	21695										
121249	02-01-2018 23:00	21230										
121250	03-01-2018 00:00	20799										
121251	01-01-2018 01:00	18508										
121252	01-01-2018 02:00	18600										
121253	01-01-2018 03:00	18571										
121254	01-01-2018 04:00	18686										
121255	01-01-2018 05:00	18912										
121256	01-01-2018 06:00	19214										
121257	01-01-2018 07:00	19585										
121258	01-01-2018 08:00	19886										
121259	01-01-2018 09:00	19933										
121260	01-01-2018 10:00	19847										
121261	01-01-2018 11:00	19710										
121262	01-01-2018 12:00	19453										
121263	01-01-2018 13:00	19049										
121264	01-01-2018 14:00	18737										
121265	01-01-2018 15:00	18619										
121266	01-01-2018 16:00	18691										
121267	01-01-2018 17:00	19109										
121268	01-01-2018 18:00	20279										
121269	01-01-2018 19:00	20925										
121270	01-01-2018 20:00	21089										
121271	01-01-2018 21:00	20999										
121272	01-01-2018 22:00	20820										
121273	01-01-2018 23:00	20415										
121274	02-01-2018 00:00	19993										
121275												

Necessary step to follow:

1.Import Libraries:

Start by importing the necessary libraries.

Program:

```
#import the libraries
```

```
import pandas as pd
```

```
import numpy as np
```

```
import matplotlib.pyplot as plt

import seaborn as sns

pd.options.display.float_format = '{:.5f}'.format

pd.options.display.max_rows = 12

filepath = '../input/hourly-energy-consumption/PJME_hourly.csv'

df = pd.read_csv(filepath)

print("Now, you're ready for step one")
```

Explore the data:

```
# turn data to datetime

df = df.set_index('Datetime')

df.index = pd.to_datetime(df.index)

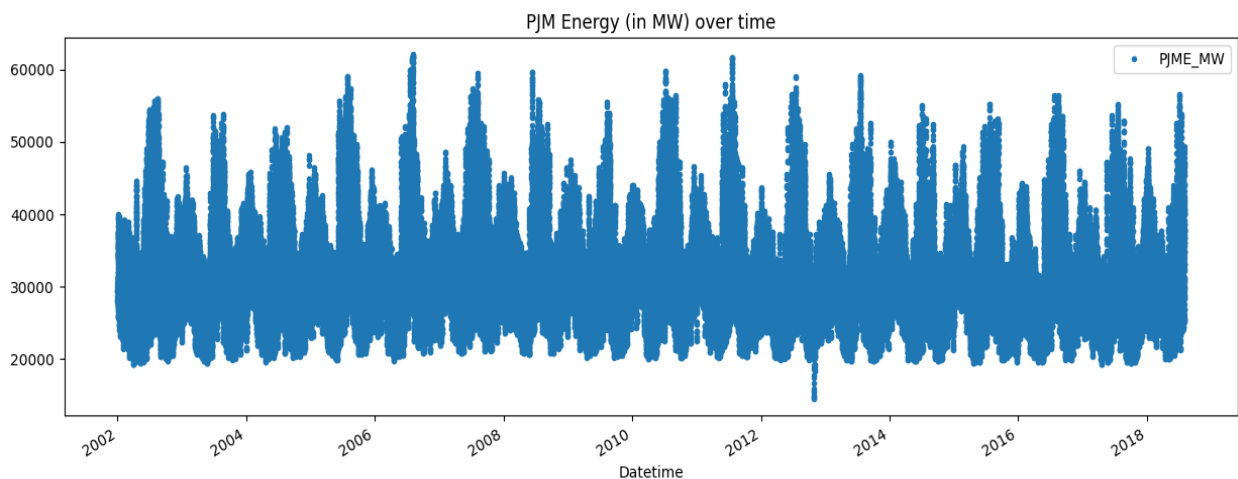
df.plot(style='.',

        figsize=(15, 5),

        title='PJM Energy (in MW) over time')

plt.show()
```

Output:

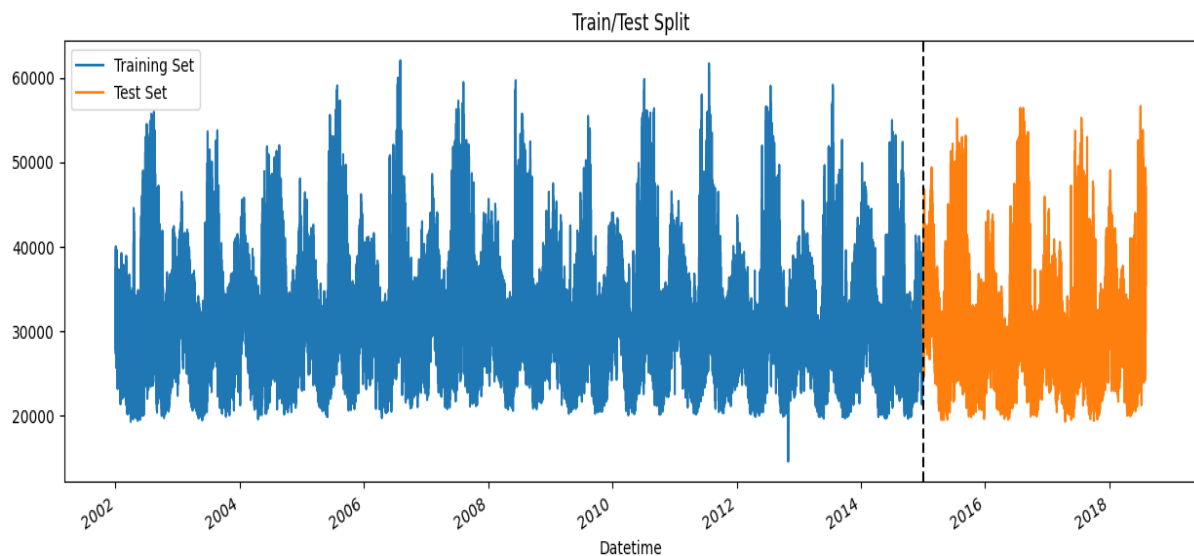


Split the data:

```
train = df.loc[df.index < '01-01-2015']
```

```
test = df.loc[df.index >= '01-01-2015']
```

Output:



Features Engineering:

```
def create_features(df):
```

```
    df = df.copy()
```

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

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

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

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

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

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

```
    df['dayofmonth'] = df.index.day
```

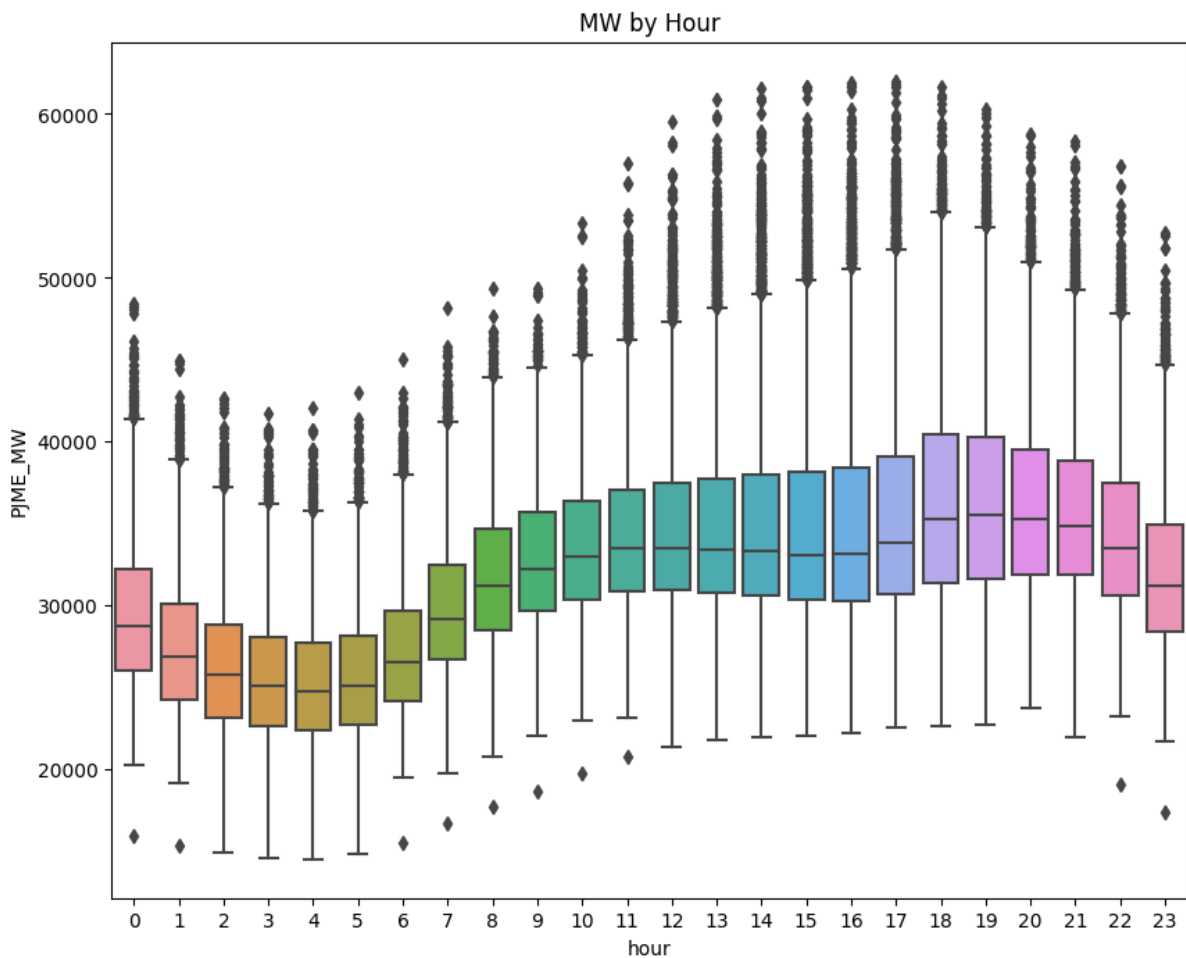
```
    df['weekofyear'] = df.index.isocalendar().week
```

```
    return df
```

```
df = create_features(df)
```

```
fig, ax = plt.subplots(figsize=(10, 8))
sns.boxplot(data=df, x='hour', y='PJME_MW')
ax.set_title('MW by Hour')
plt.show()
```

virtualization:



Modelling and preprocessing :

```
# preprocessing
```

```
train = create_features(train)
```

```
test = create_features(test)
```

```
features = ['dayofyear', 'hour', 'dayofweek', 'quarter', 'month', 'year']
```

```
target = 'PJME_MW'
```

```
X_train = train[features]
```

```
y_train = train[target]
```

```
X_test = test[features]
```

```
y_test = test[target]
```

Building the model:

```
import xgboost as xgb
```

```
from sklearn.metrics import mean_squared_error
```

```
# build the regression model
```

```
reg = xgb.XGBRegressor(base_score=0.5, booster='gbtree',
```

```
                        n_estimators=1000,
```

```
                        early_stopping_rounds=50,
```

```
                        objective='reg:linear',
```

```
                        max_depth=3,
```

```
                        learning_rate=0.01)
```

```
reg.fit(X_train, y_train,
```

```
        eval_set=[(X_train, y_train), (X_test, y_test)],
```

```
        verbose=100)
```

```
fi = pd.DataFrame(data=reg.feature_importances_,
```

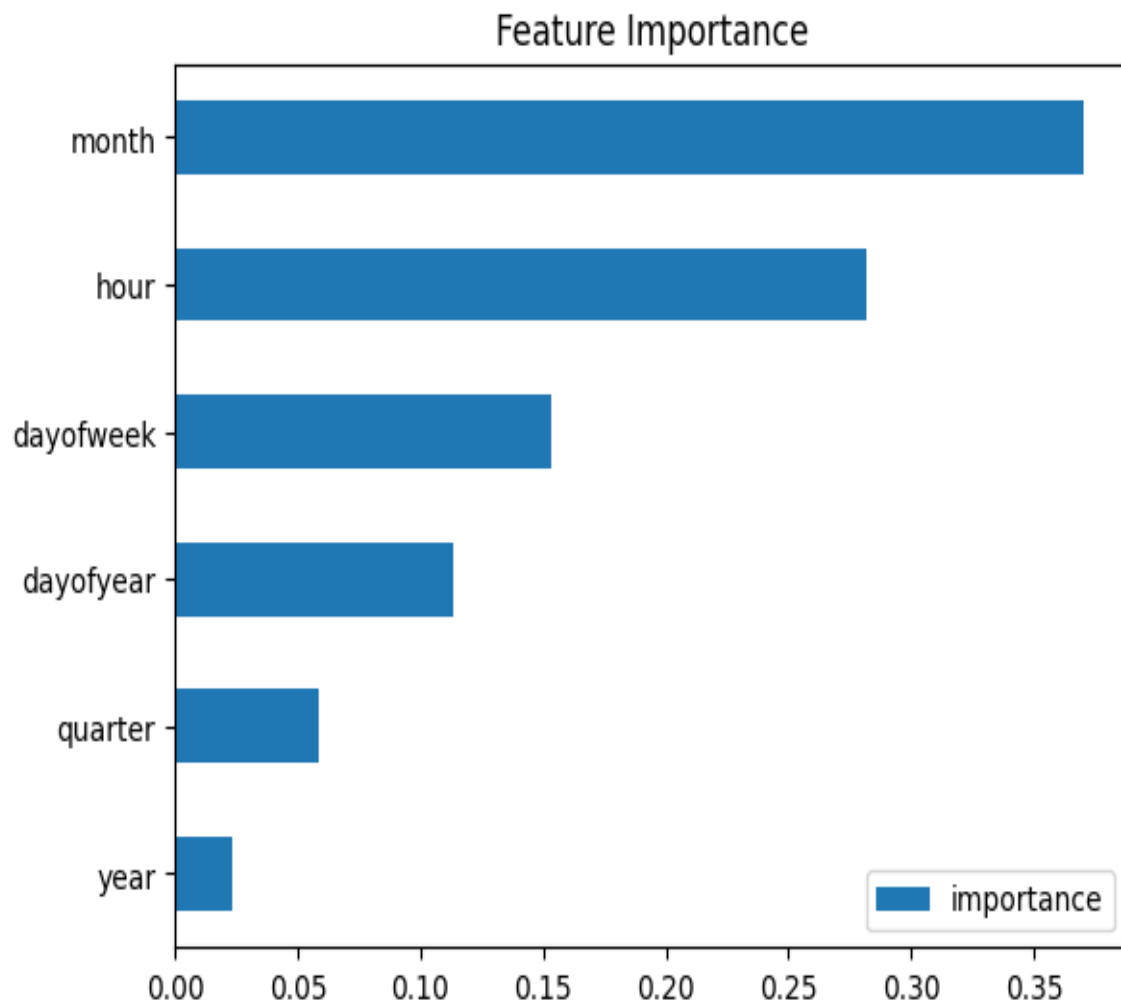
```
                  index=reg.feature_names_in_,
```

```
                  columns=['importance'])
```

```
fi.sort_values('importance').plot(kind='barh', title='Feature  
Importance')
```

```
plt.show()
```

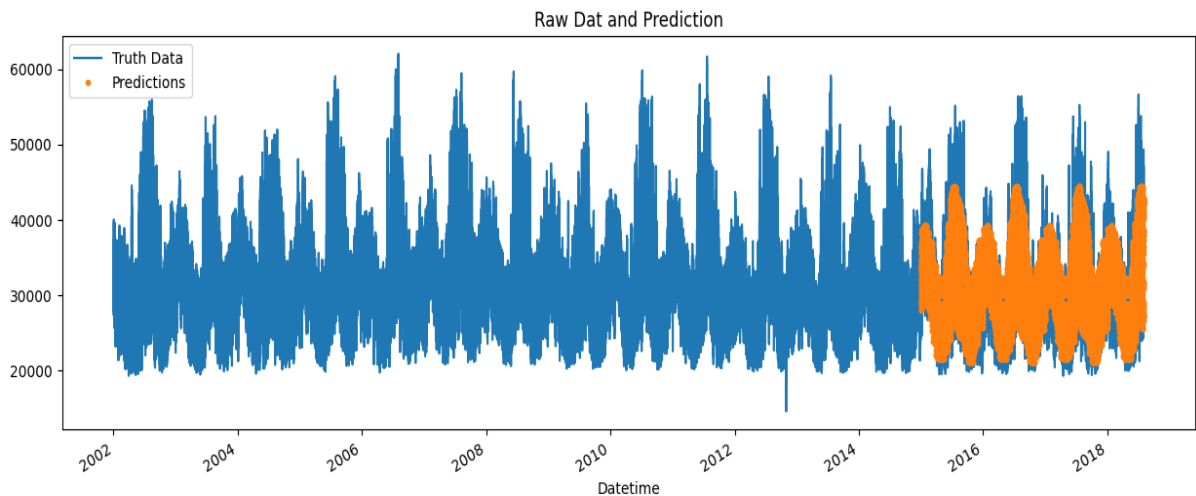
Output:



Forecasting on test data:

```
test['prediction'] = reg.predict(X_test)
df = df.merge(test[['prediction']], how='left', left_index=True,
right_index=True)
ax = df[['PJME_MW']].plot(figsize=(15, 5))
df['prediction'].plot(ax=ax, style='.')
plt.legend(['Truth Data', 'Predictions'])
ax.set_title('Raw Dat and Prediction')
plt.show()
```

Output:



Conclusion:

In conclusion, measuring energy consumption is not just a matter of tracking numbers; it's a pivotal practice that has far-reaching implications for our planet, our wallets, and our overall well-being. Whether at the individual, industrial, or governmental level, the act of quantifying and monitoring energy usage holds immense value.

By carefully measuring energy consumption, we can pinpoint inefficiencies, make informed decisions to reduce energy waste, and ultimately save money while reducing our impact on the environment. The information collected through these measurements empowers us to set and achieve energy-saving goals, contributing to a more sustainable and responsible world.

In the context of global climate change, energy consumption measurements are crucial in the fight against greenhouse gas emissions. They help us track our progress toward reducing our carbon footprint and implementing effective energy-efficient policies.