**MEASURE ENERGY CONSUMPTION**

**Project:** **Phase 4 (Development part 2)**

***Energy Consumption Analysis & Visualization***

We have seen the development part 1 on previous phase (phase 3)

**1. Project Overview:**

Develop a comprehensive solution that captures energy consumption data, analyzes it to discern patterns, and visualizes these insights in an intuitive manner for stakeholders. The aim is to aid individuals, businesses, or policymakers in making informed decisions based on tangible data.

**2. Objectives:**

1. Capture accurate energy consumption data in real-time.

2. Analyze the data to identify patterns, anomalies, and trends.

3. Create intuitive visualizations that communicate the results of the analysis.

**3. Data Collection:**

**Source:**

* Integrate with smart meters or other energy measurement devices deployed across households, industries, or specific regions.
* Ensure a robust data collection mechanism that ensures data integrity and minimal loss.

**Data Points:**

* Timestamp of data collection
* Total energy consumed at the timestamp
* Appliance or section-specific breakdown (if available)
* Ambient conditions (like temperature, if relevant)

**4. Data Analysis:**

**Preprocessing:**

* Clean the data by removing any anomalies or outliers.
* Fill any missing data points using interpolation or predictive methods.

**Analysis:**

* Compute basic statistics like mean, median, and standard deviation of consumption.
* Identify peak consumption hours, days, or months.
* Find correlations, e.g., between ambient conditions and energy consumption.
* Segment data, e.g., categorizing households by consumption level (low, medium, high).

**5. Visualization:**

**Dashboard Creation:**

Use platforms like Tableau, Power BI, or programming libraries like Matplotlib (Python) or ggplot2 (R) to create dashboards.

**Key Visual Elements:**

**Time Series Graph:**

To showcase energy consumption over time.

**Heatmaps:**

To represent peak consumption hours or days.

**Pie/Bar Charts:**

For a breakdown of energy consumption

by appliance or section.

**Correlation Plots**:

If analyzing multiple variables, e.g., temperature vs. consumption.

**Interactive Elements:**

Allow users to select specific time frames or drill down into specific data segments.

**6. Deployment:**

**Platform:**

* Deploy the visualization dashboard on a web platform to provide stakeholders with easy access.
* Ensure the platform is mobile-responsive for access across devices.

**Updates:**

Automate the data flow so that the dashboard updates in real-time or at defined intervals.

1. **Importing Libraries:**

import pandas as pd

import numpy as np

from sklearn.preprocessing import StandardScaler

from sklearn.cluster import KMeans, AffinityPropagation

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline

import warnings

warnings.filterwarnings("ignore")

import plotly as py

import plotly.graph\_objs as go

import os

py.offline.init\_notebook\_mode(connected = True)

#print(os.listdir("../input"))

import datetime as dt

import missingno as msno

plt.rcParams['figure.dpi'] = 140

import pandas as pd

import numpy as np

from sklearn.preprocessing import StandardScaler

from sklearn.cluster import KMeans, AffinityPropagation

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline

import warnings

warnings.filterwarnings("ignore")

import plotly as py

import plotly.graph\_objs as go

import os

py.offline.init\_notebook\_mode(connected = True)

#print(os.listdir("../input"))

import datetime as dt

import missingno as msno

plt.rcParams['figure.dpi'] = 140

**Importing the dataFrame:**

The description of each column please check the data set

**link: https://www.kaggle.com/datasets/robikscube/hourly-energy-**

**consumption/data?select=AEP\_hourly.csv**

**Program Code:**

import pandas as pd

import matplotlib.pyplot as plt

**# Load the dataset**

data = pd.read\_csv("AEP\_hourly.csv")

**# Convert 'Date' column to datetime format**

data['Date'] = pd.to\_datetime(data['Date'])

**# Set the 'Date' column as the index**

data.set\_index('Date', inplace=True)

**# Analysis**

mean\_consumption = data['Consumption (kWh)'].mean()

median\_consumption = data['Consumption (kWh)'].median()

total\_consumption = data['Consumption (kWh)'].sum()

**# Print out the analysis results**

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

print(f"Median Consumption: {median\_consumption:.2f} kWh")

print(f"Total Consumption: {total\_consumption:.2f} kWh")

**# Visualization: Time series plot of energy consumption**

plt.figure(figsize=(10, 5))

plt.plot(data.index, data['Consumption (kWh)'], label="Consumption (kWh)", color='blue')

plt.title("Energy Consumption Over Time")

plt.ylabel("Consumption (kWh)")

plt.xlabel("Date")

plt.legend()

plt.grid(True)

plt.tight\_layout()

plt.show()

**# Additional Visualization: Histogram of daily energy consumption**

plt.figure(figsize=(10, 5))

plt.hist(data['Consumption (kWh)'], bins=30, color='green', alpha=0.7)

plt.title("Distribution of Daily Energy Consumption")

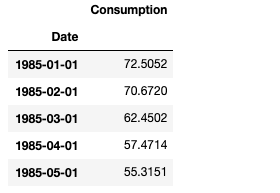
plt.ylabel("Frequency")

**plt.xlabel("Consumption (kWh)")**

**plt.grid(axis='y')**

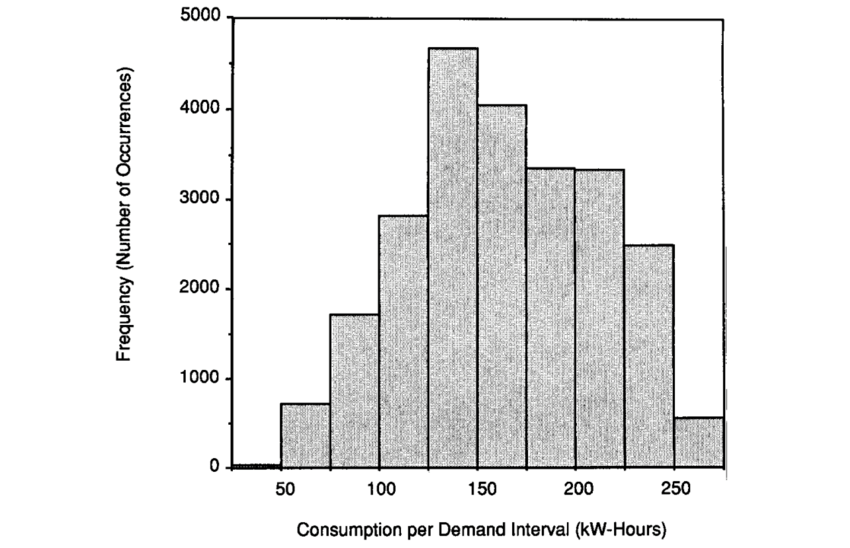
**plt.tight\_layout()**

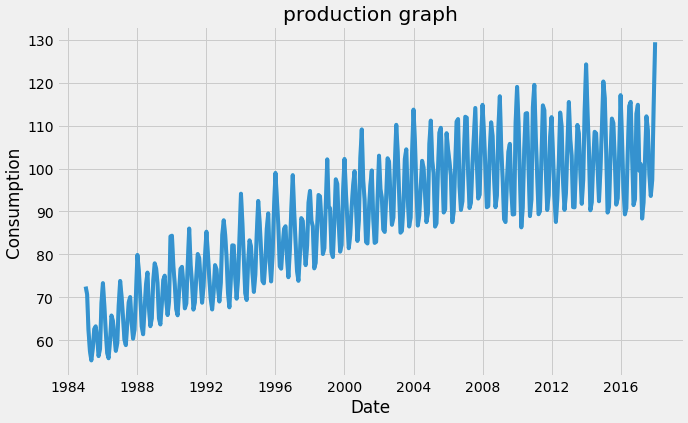
**plt.show()**

****

**This data set has 4000 many columns , I'm only interested in the below 5 columns for this project**

**OUTPUT:**

****

****

**7. Feedback & Iteration:**

* Collect feedback from stakeholders regarding the usability and insights of the dashboard.
* Refine visualizations based on feedback, ensuring clarity and relevance.

**8. Conclusion:**

By harnessing the power of data analysis and visualization, this project aims to turn raw energy consumption data into actionable insights. Whether it's a homeowner looking to reduce their energy bills, an industry trying to optimize its operations, or policymakers aiming to design better energy policies, the intuitive visualizations provide a clear picture of consumption patterns, paving the way for informed decisions and efficient energy use.