**Measure Energy Consumption**

**Phase-1 Document Submission**

**Project:** Measure Energy Consumption

**Team Members:**

110521106020- Sanjay S

110521106328- Santhosh kumar S

110521106327- Raghul M

**OVERVIEW:**

**The project focuses on addressing the widespread issue of inefficient energy consumption in various sectors, including residential, commercial, and industrial. Inefficient energy usage leads to increased costs, environmental concerns, and a strain on energy resources. The primary goal is to design and implement a system that enables precise measurement, analysis, and optimization of energy consumption to promote sustainability, cost-efficiency, and reduced environmental impact.**

**DESIGN THINKING:**

**Empathize:**

**Understand the needs, concerns, and behaviors of energy consumers.Conduct surveys, interviews, and observe energy consumption patterns.**

**Define:**

**Clearly define the problem statement and the goals of the project.Create user personas to represent different types of energy consumers.**

**Ideate:**

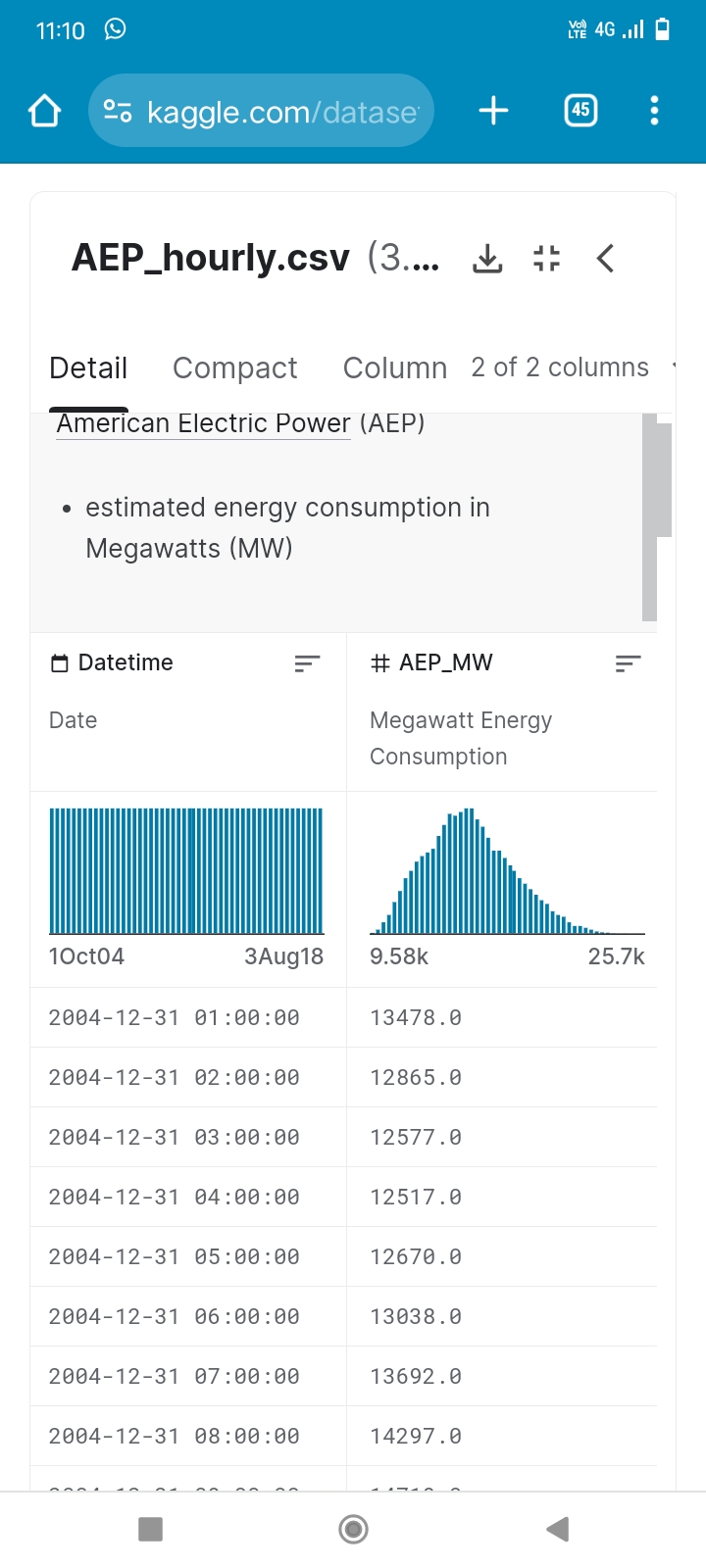
**Brainstorm potential solutions to address energy optimization.Consider technologies like IoT devices, machine learning algorithms, and smart meters.**

**Prototype:**

**Develop a prototype of the energy optimization system.Test it in a controlled environment or with a small group of users.**

**Dataset Link:**

**Link:** [**https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption**](https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption)



Data preprocessing

1. Data Cleaning:

Handle Missing Data: Identify and address missing values in the dataset. Depending on the extent of missing data, you can choose to remove rows with missing values or impute missing values using methods like mean, median, or interpolation.

Remove Duplicates: Check for and remove duplicate records, if any, to ensure data integrity.

Outlier Detection and Handling: Identify outliers in energy consumption measurements and decide whether to remove or transform them based on the context of your analysis

**PYTHON PROGRAMMING:**

from mpl\_toolkits.mplot3d import Axes3D

from sklearn.preprocessing import StandardScaler

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)*

There is 1 csv file in the current version of the dataset:

print(os.listdir('../input'))

['database.sqlite', 'Tweets.csv']

nRowsRead = 1000 *# specify 'None' if want to read whole file*

*# Tweets.csv has 14640 rows in reality, but we are only loading/previewing the first 1000 rows*

df1 = pd.read\_csv('../input/Tweets.csv', delimiter=',', nrows = nRowsRead)

df1.dataframeName = 'Tweets.csv'

nRow, nCol = df1.shape

print(f'There are **{nRow}** rows and **{nCol}** columns')

There are 1000 rows and 15 columns

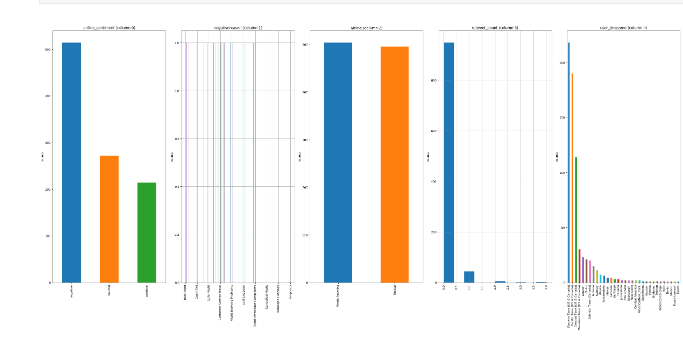
df1.head(5)



Distribution graphs (histogram/bar graph) of sampled columns:

linkcode

plotPerColumnDistribution(df1, 10, 5)



**CONCLUSION:**

In conclusion, sentiment analysis is a powerful tool for modern marketing. By harnessing the capabilities of this technology, businesses can gain deep insights into customer emotions, preferences, and perceptions. These insights enable data-driven decision-making, the optimization of marketing strategies, and ultimately, the enhancement of customer satisfaction and brand success in a dynamic and competitive marketplace.