

Name: Faisal Zulfiqar

Student ID: 22010044

ASSIGNMENT 1: VISUALISATION

GITHUB LINK: <https://github.com/Faisal-Zulfiqar786/Rework-Assignment-1.git>

DATASET LINK: https://archive.ics.uci.edu/ml/machine-learning-databases/00235/household_power_consumption.zip

DATASET

The "Household Power Consumption" dataset was used, and it contains measurements of one household's electric power usage taken over a nearly four-year period using a one-minute sample rate. The information was gathered in a single home in Sceaux, France, using an electrical monitoring system that consisted of a series of measures made each minute.

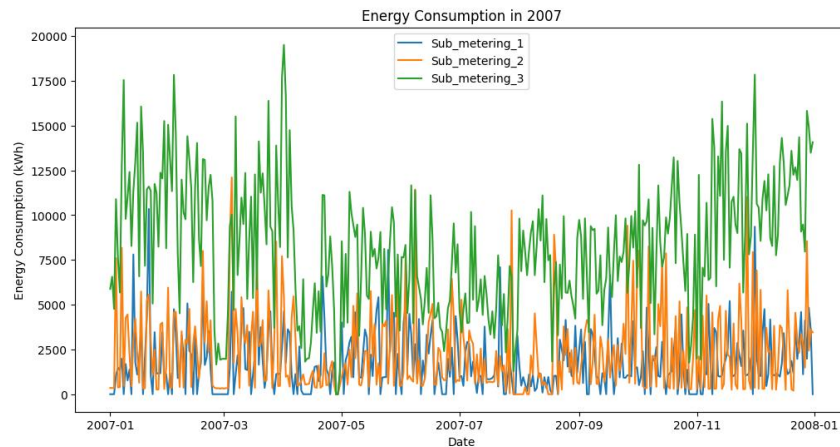
2,075,259 measurements from January 2007 to November 2010 are included in the dataset. Nine attributes are included in the data:

Attribute Name	Description
Date	The day that the power usage was calculated using the dd/mm/yyyy format.
Time	The period when power usage was calculated using the hh:mm:ss format.
Global Active Power	The amount of kilowatts that the household uses for active power overall
Global Reactive Power	Kilowatts represent the amount of reactive electricity that the home uses overall
Voltage	The voltage level measured at the time, expressed in volts
Global Intensity	the total current intensity at the time of measurement, expressed in amperes
Sub-metering 1	The amount of active energy used by the kitchen in kilowatt-hours between the current measurement and the prior one
Sub-metering 2	The amount of energy used actively by the washing room in kilowatt-hours between the current measurement and the prior one
Sub-metering 3	Kilowatt-hours of active energy used by an electric water heater and an air conditioner between the present reading and the prior measurement

VISUALIZATION

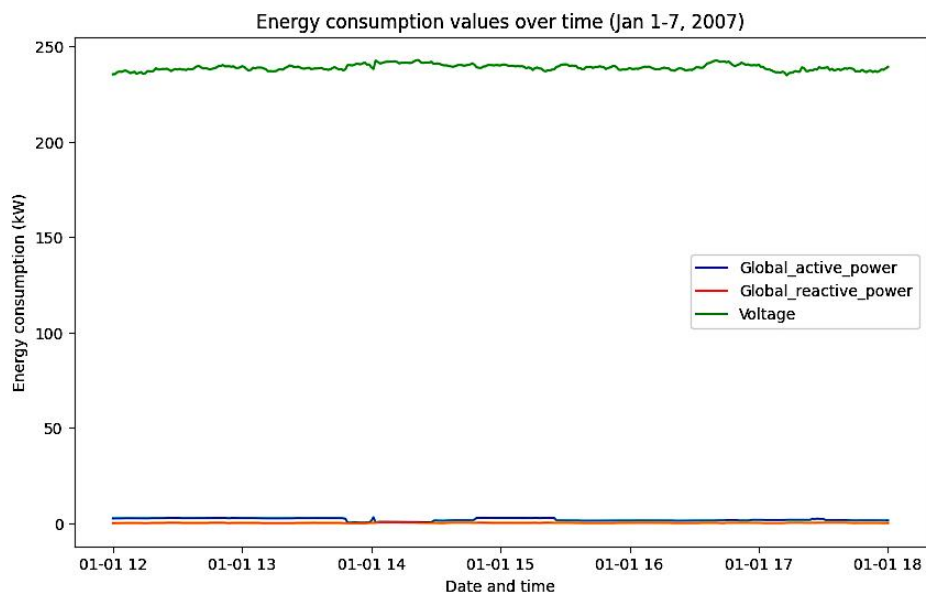
MULTIPLE LINE PLOT

This programme reads a household power consumption dataset and generates a line plot that displays the energy usage for three separate submeters over the year 2007. The code plots the sub-metering data for each day in 2007 and displays the three sub-meters as three separate lines on the plot.



LINE PLOT 1: ENERGY CONSUMPTION FOR THREE DIFFERENT SUB-METERSTHE ADVANTAGE OF THIS PLOT IS THAT IT ENABLES US TO EXAMINE HOW THE ENERGY USAGE FOR EACH OF THE THREE SUBMETERS CHANGES OVER TIME. THIS TYPE OF DATA VISUALISATION ALLOWS US TO SPOT PATTERNS AND TRENDS IN THE DATA THAT MIGHT NOT BE IMMEDIATELY OBVIOUS WHEN EXAMINING THE RAW DATA. ADDITIONALLY, IT IS SIMPLE TO COMPARE THE ANNUAL ENERGY USAGE FOR EACH SUB-METER THANKS TO THE PLOT.

The second line plot shows how the values for several variables, including global active power, global reactive power, and voltage, have changed over time.



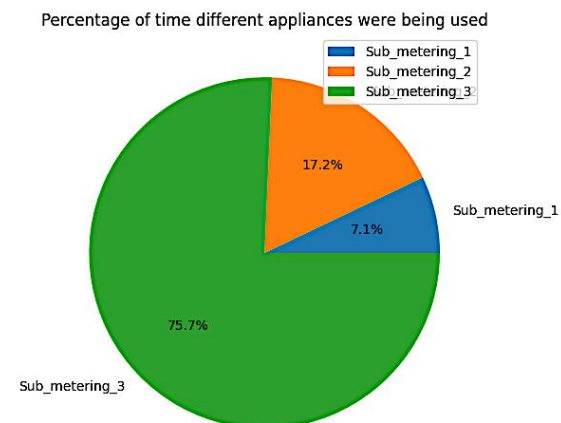
LINE PLOT 2: ENERGY CONSUMPTION VALUES OVER TIME FOR DIFFERENT VARIABLES

Name: Faisal Zulfiqar

Student ID: 22010044

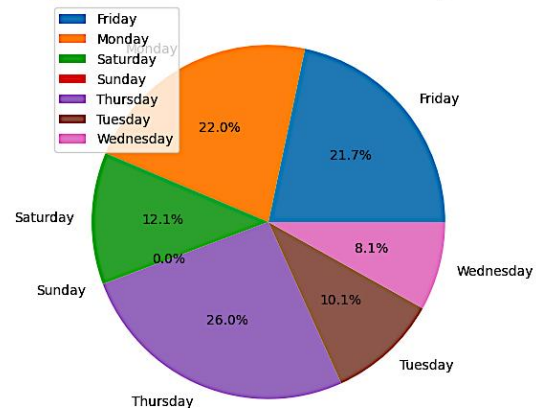
PIE CHART

The code employs matplotlib to generate three pie charts from a dataset. The first graph illustrates the percentage of time that various appliances were used. The graph indicates that the two appliances that used the most electricity during the first week of January 2007 were an electric water heater and an air conditioner.



PIE CHART 1:PERCENTAGE OF TIME DIFFERENT APPLIANCES

Percentage of Time Power was Consumed on Different Days of the Week

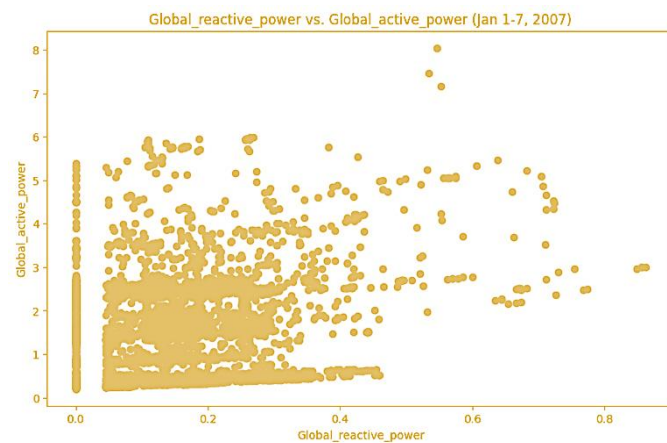
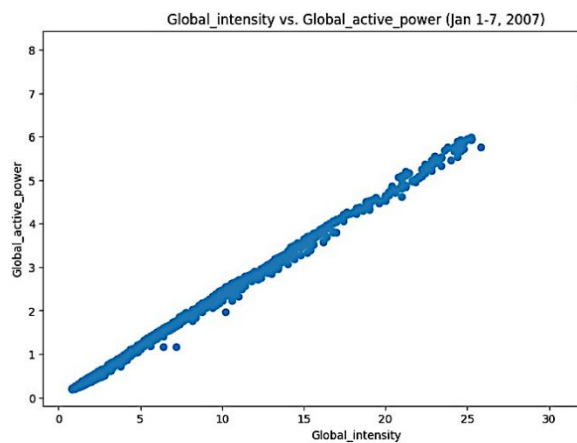


PIE CHART 2: PERCENTAGE OF TIME POWER WAS CONSUMED ON DIFFERENT DAYS OF THE WEEK

The second graph displays the proportion of electricity used during various times of the week. In the first week of January 2007, the day with the highest power use was Thursday (26%), according to this pie chart.

SCATTER PLOT

We may visually check if there is a correlation or relationship between the two variables by plotting the scatter plot of each feature variable against the output variable independently. If there is no clear pattern and the points on the plot are evenly spaced, it is likely that the feature variable and the output variable do not have a strong correlation. On the other hand, it suggests that the feature variable is connected to the output variable if the points are clustered or exhibit a distinct pattern. Building a predictive model may benefit from using this knowledge to determine which features have a substantial effect on the output variable.



SCATTER PLOTS 1