**IBM NAANMUDHALVAN**

**PHASE 3**

**DOMAIN – ELECTRICITY PRICE PREDICTION**

**INTRODUCTION:**

Electricity is a basic human need and definitely one of the most important factors of societal progress. In recent decades however, electricity has entered the market as a tradeable commodity and the power industry of many countries has been **deregulated**. In Spain, the Electric Power Act 54/1997 exposed all of the stakeholders to **high amounts of uncertainty** as the price of electricity is determined by countless factors and also, due to the fact that electricity cannot be stored in large quantities efficiently . With the emergence of this new market, the need for reliable forecasting methods at all scales (hourly, daily, long-term, etc.) has also emerged and has become a large area of research.

**DATASET:**

Electricity price prediction

**ABOUT:**

* Loading a dataset
* Preprocessing dataset
* Data cleaning
* Data transformation
* Data reduction

**PROGRAM:**

**LOAD THE DATASET:**

import pandas as pd

a=pd.read\_csv('Electricity.csv')

a

38014 rows × 18 columns

a.head(5)

a.tail(5)

a.isnull()

a.shape

(38014, 18)

a.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 38014 entries, 0 to 38013

Data columns (total 18 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 DateTime 38014 non-null object

1 Holiday 38014 non-null object

2 HolidayFlag 38014 non-null int64

3 DayOfWeek 38014 non-null int64

4 WeekOfYear 38014 non-null int64

5 Day 38014 non-null int64

6 Month 38014 non-null int64

7 Year 38014 non-null int64

8 PeriodOfDay 38014 non-null int64

9 ForecastWindProduction 38014 non-null object

10 SystemLoadEA 38014 non-null object

11 SMPEA 38014 non-null object

12 ORKTemperature 38014 non-null object

13 ORKWindspeed 38014 non-null object

14 CO2Intensity 38014 non-null object

15 ActualWindProduction 38014 non-null object

16 SystemLoadEP2 38014 non-null object

17 SMPEP2 38014 non-null object

dtypes: int64(7), object(11)

memory usage: 5.2+ MB

a.nunique()

DateTime 38014

Holiday 15

HolidayFlag 2

DayOfWeek 7

WeekOfYear 52

Day 31

Month 12

Year 3

PeriodOfDay 48

ForecastWindProduction 29312

SystemLoadEA 36166

SMPEA 8661

ORKTemperature 32

ORKWindspeed 53

CO2Intensity 25115

ActualWindProduction 2940

SystemLoadEP2 36171

SMPEP2 9277

dtype: int64

a.isnull().sum()

DateTime 0

Holiday 0

HolidayFlag 0

DayOfWeek 0

WeekOfYear 0

Day 0

Month 0

Year 0

PeriodOfDay 0

ForecastWindProduction 0

SystemLoadEA 0

SMPEA 0

ORKTemperature 0

ORKWindspeed 0

CO2Intensity 0

ActualWindProduction 0

SystemLoadEP2 0

SMPEP2 0

dtype: int64

a.duplicated().any()

False

df=a['Year'].head(10)

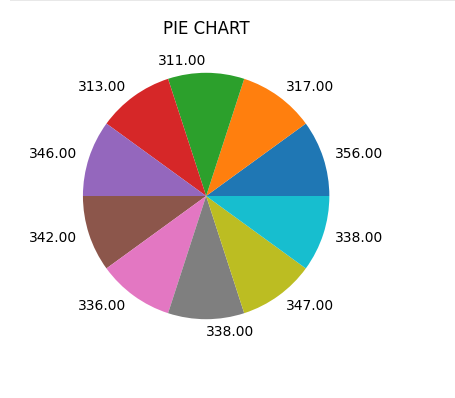
df1=a['ActualWindProduction'].head(10)

fig = plt.figure(figsize =(4,4))

plt.pie(df, labels= df1)

plt.title("PIE CHART")

plt.show()



import numpy as np

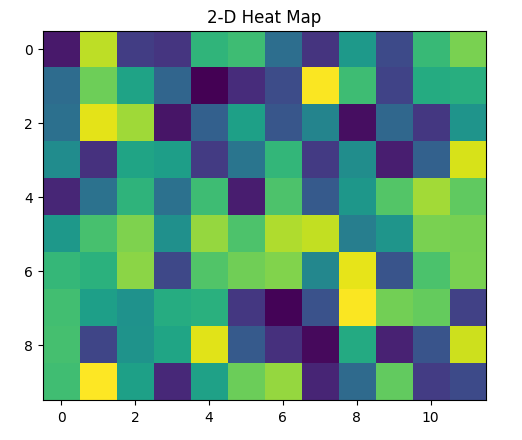
import matplotlib.pyplot as plt

a = np.random.random(( 10, 12 ))

plt.imshow( a )

plt.title( "2-D Heat Map" )

plt.show()



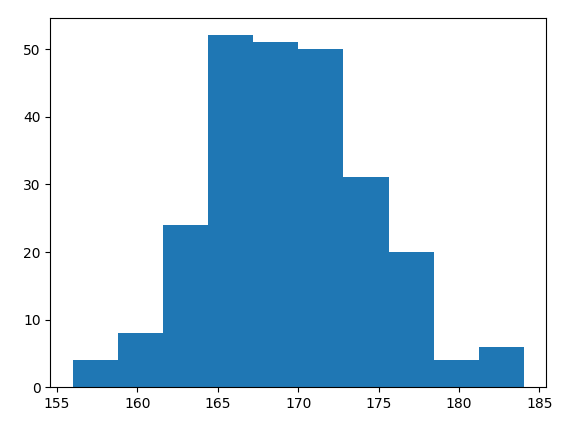
import matplotlib.pyplot as plt

import numpy as np

x = np.random.normal(170, 5, 250)

plt.hist(x)

plt.show()



**CONCLUSION:**

The dataset has been preprocessed and visualized successfully in the given dataset.

**NAME :KHAWSIK R**

**COLLEGE CODE : 4204**

**REGISTER NO :420421106025**