import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn import metrics

df=pd.read\_csv('https://raw.githubusercontent.com/kunalPisolkar24/DSBDA\_Lab/master/twelve/sample\_weather.txt',sep='\s+')

# sep-> specifies that data is separated by one or more whitespace characters

print(df.info())

df.columns=['Station\_ID','WMO\_Number','Timestamp','Temperature','Dew\_point','Wind\_speed','pressure','Humidity','Visibility','Cloud\_cover','Wind\_direction','Precipitation','Snow\_depth','Radiation','UV\_Index','Weather\_code','Sunrise\_sunset','Sea\_level\_pressure','Ground\_level\_pressure']

df.info()

df.head()

df.isnull().sum()

import matplotlib.pyplot as plt

avg\_temp=df['Temperature'].mean()

avg\_dew=df['Dew\_point'].mean()

avg\_wind=df['Wind\_speed'].mean()

print(f"Average Temperature : {avg\_temp}")

print(f"Average Dew Point : {avg\_dew}")

print(f"Average Wind Speed : {avg\_wind}")

plt.figure(figsize=(15,5))

#Temperature plot

plt.subplot(1,3,1)

plt.hist(df['Temperature'],bins=20,color='skyblue',edgecolor='black')

#bins means no. of bars in histogram

plt.title('Temeprature Distribution')

plt.xlabel('Temperature')

plt.ylabel('Frequency')

plt.show()

#Dew point plot

plt.subplot(1,3,2)

plt.title('Dew Point Distribution')

plt.hist(df['Dew\_point'],bins=20,color='skyblue',edgecolor='black')

plt.xlabel('Dew point')

plt.ylabel('Frequency')

#Wind Speed plot

plt.subplot(1,3,3)

plt.hist(df['Wind\_speed'],bins=20,color='skyblue',edgecolor='black')

plt.title('Wind Speed Distribution')

plt.xlabel('Wind Speed')

plt.ylabel('Frequency')