Khan Needa - 6118019
Mirajkar Shamim - 6118027
Mohd Umair Sayed - 6118028

EXPERIMENT NO:- 10

Aim: To implement Multiple Linear Regression using python.

Multiple linear regression refers to a statistical technique that is used to predict the outcome of a variable based on the value of two or more variables. It is sometimes known simply as multiple regression, and it is an extension of linear regression. The variable that we want to predict is known as the dependent variable, while the variables we use to predict the value of the dependent variable are known as independent or explanatory variables.

Formula and Calculation of Multiple Linear Regression:

$$Yi = \beta 0 + \beta 1 xi1 + \beta 2 xi2 + ... + \beta p xip + \epsilon$$

where for $i=n$ observations:
yi =dependent variable
xi=explanatory variables
$\beta 0$=y-intercept (constant term)
βp=slope coefficients for each explanatory variable
ϵ=the model's error term (also known as the residuals)

Implementation:

Importing libraries :

In [31]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

Importing dataset :

In [32]:
```
df = pd.read_csv("WeatherDataM.csv")
df.head()
```

Out[32]:

	Temperature (C)	Wind Speed (km/h)	Pressure (millibars)	Humidity
0	9.472222	14.1197	1015.13	0.89
1	9.355556	14.2646	1015.63	0.86
2	9.377778	3.9284	1015.94	0.89
3	8.288889	14.1036	1016.41	0.83
4	8.755556	11.0446	1016.51	0.83

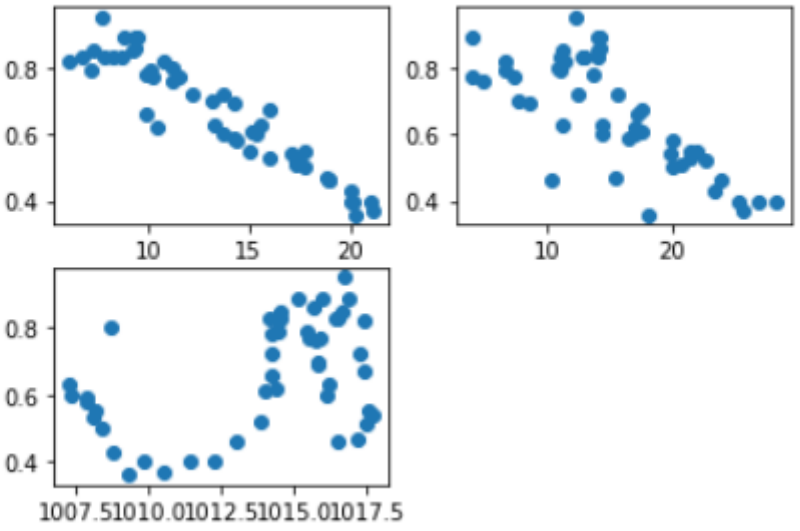
In [33]:
```
df.columns
```

Out[33]:
```
Index(['Temperature (C)', 'Wind Speed (km/h)', 'Pressure (millibars)',
       'Humidity'],
      dtype='object')
```

In [34]:
```
weather_features = ['Temperature (C)', 'Wind Speed (km/h)', 'Pressure (millibars)']
X = df[weather_features]
y = df.Humidity
```

In [35]:
```
plt.subplot(2,2,1)
plt.scatter(X['Temperature (C)'],y)
plt.subplot(2,2,2)
plt.scatter(X['Wind Speed (km/h)'],y)
plt.subplot(2,2,3)
plt.scatter(X['Pressure (millibars)'],y)
```

Out[35]:
```
<matplotlib.collections.PathCollection at 0x9ac31a8>
```



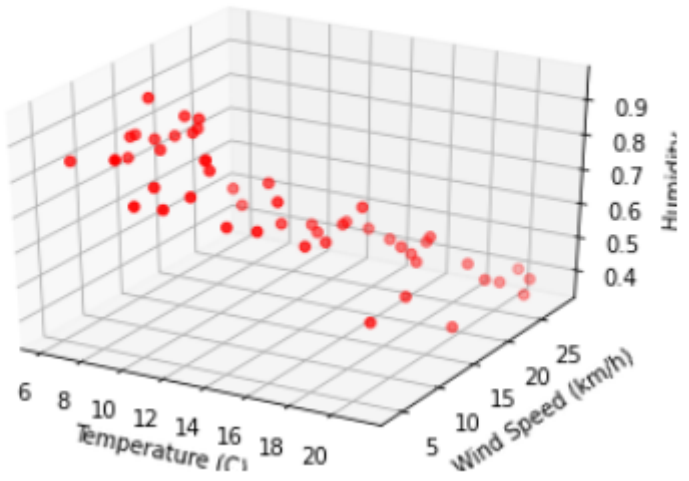
In [36]:
```
X = X.drop("Pressure (millibars)", 1)
X.head()
```

Out[36]:

	Temperature (C)	Wind Speed (km/h)
0	9.472222	14.1197
1	9.355556	14.2646
2	9.377778	3.9284
3	8.288889	14.1036
4	8.755556	11.0446

In [37]:
```
from mpl_toolkits.mplot3d import Axes3D

fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

x1 = X["Temperature (C)"]
x2 = X["Wind Speed (km/h)"]

ax.scatter(x1, x2, y, c='r', marker='o')

# Set axis labels
ax.set_xlabel('Temperature (C)')
ax.set_ylabel('Wind Speed (km/h)')
ax.set_zlabel('Humidity')
```

Out[37]: Text(0.5, 0, 'Humidity')



In [38]:
```
from sklearn.linear_model import LinearRegression

mlr_model = LinearRegression()
mlr_model.fit(X, y)
```

Out[38]:
```
LinearRegression()
```

In [39]:
```
y_pred = mlr_model.predict([[15, 21]])
y_pred
```

Out[39]:
```
array([0.58721315])
```

So a temperature of 15 °C and Wind speed of 21 km/h expects to give us a Humidity of 0.587.

Conclusion:

Hence we have successfully implemented Multiple Linear Regression using Python