MULTI LINEAR REGRESSION

RAZA BAKIR SHAH

PGD DATA SCIENCES & AI (BATCH 6)

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from pandas.core.common import random_state
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
```

In [2]:
```
# Load the training data
data = pd.read_csv('multiple_linear_regression_dataset.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	age	experience	income
0	25	1	30450
1	30	3	35670
2	47	2	31580
3	32	5	40130
4	43	10	47830

In [4]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   age         20 non-null     int64
 1   experience  20 non-null     int64
 2   income      20 non-null     int64
dtypes: int64(3)
memory usage: 612.0 bytes
```

In [5]:
```
data.shape
```

Out[5]: (20, 3)

In [6]:
```
# Define the independent variables (X) and the dependent variable (y)
X = data[['age', 'experience']]
y = data['income']
```

In [7]:
```
X.head()
```

Out[7]:

	age	experience
0	25	1
1	30	3
2	47	2
3	32	5
4	43	10

In [8]:
```
y.head()
```

Out[8]:
```
0    30450
1    35670
2    31580
3    40130
4    47830
Name: income, dtype: int64
```

In [9]:
```
# Create and fit the multiple linear regression model
model = LinearRegression()
model.fit(X, y)
```

Out[9]:
```
LinearRegression
LinearRegression()
```

In [10]:
```
# Retrieve the coefficients and intercept
coefficients = model.coef_
intercept = model.intercept_
```

In [11]:
```
# Print Our Results
print("Intercept:", intercept)
print("Coefficients:", coefficients)
```

```
Intercept: 31261.68985410128
Coefficients: [ -99.19535546 2162.40419192]
```

Make Predictions

In [12]:
```
# Find price of income for age 30 and experience 5 years
predicted_income = model.predict([[30, 5]])
print(predicted_income)
```

```
[39097.85014989]
C:\Users\razab\anaconda3\Lib\site-packages\sklearn\base.py:493: UserWarning: X do
es not have valid feature names, but LinearRegression was fitted with feature nam
es
  warnings.warn(
```

In [13]:
```
# Y = m1 * X1 + m2 * X2 + b (m1,m2,m3 is coefficient and b is intercept)
```

```
-99.19535546 * 30 + 2162.40419192 * 5 + 31261.68985410128
```

Out[13]: 39097.85014990128

In [ ]: