Description of the Linear Regression Model Linear regression is a statistical method that models the relationship between a dependent variable and one or more independent variables by fitting a linear equation to observed data. The simplest form is a linear equation for a line (y = mx + b), where:

where: y is the dependent variable,

x is the independent variable,

m is the slope of the line,

b is the y-intercept.

y=β 0 +β1x1 +β2x2 +...+βnxn +ϵ β 0 , β 1 ,...,βn are coefficients, x 1 , x 2 ,...,xn are the features,
ϵ is the error term. The goal is to find the values of the coefficients that minimize the error in prediction.

The dataset will include two features: inflation rate and employment rate. The target variable will be the Gross Regional Domestic Product (GRDP) growth rate. I'll create 20 data rows with some random values that could realistically represent these economic indicators.

In [2]:
```
import pandas as pd
import numpy as np

# Set random seed for reproducibility
np.random.seed(42)

# Generate synthetic data
data = {
    'Inflation Rate (%)': np.random.normal(2, 0.5, 20),  # Average inflation rate a
    'Employment Rate (%)': np.random.uniform(90, 100, 20),  # Employment rate betwe
    'GRDP Growth Rate (%)': np.random.normal(3, 1, 20)  # Average GRDP growth aroun
}

# Create DataFrame
df = pd.DataFrame(data)

# Save to CSV (optional, for use in Jupyter)
df.to_csv('economic_data.csv', index=False)

# Display the first few rows of the DataFrame
print(df.head())
```

```
   Inflation Rate (%)  Employment Rate (%)  GRDP Growth Rate (%)
0            2.248357            94.560700              3.822545
1            1.930868            97.851760              1.779156
2            2.323844            91.996738              3.208864
3            2.761515            95.142344              1.040330
4            1.882923            95.924146              1.671814
```

In [ ]:
```
create a linear regression model
```

In [5]:
```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error

# Generate synthetic data
np.random.seed(42)  # Seed for reproducibility
data = {
    'Inflation Rate (%)': np.random.normal(2, 0.5, 20),  # Simulating a normal dist
    'Employment Rate (%)': np.random.uniform(90, 100, 20),  # Uniform distribution
    'GRDP Growth Rate (%)': np.random.normal(3, 1, 20)  # Simulating GRDP growth ra
}
df = pd.DataFrame(data)

# Splitting data into training and test sets
X = df[['Inflation Rate (%)', 'Employment Rate (%)']]
y = df['GRDP Growth Rate (%)']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_sta

# Model training
model = LinearRegression()
model.fit(X_train, y_train)

# Predicting the test set results
y_pred = model.predict(X_test)

# Evaluating the model
mse = mean_squared_error(y_test, y_pred)

# Print the coefficients of the model
print("Coefficients of the model:")
print("Coefficient for Inflation Rate (%):", model.coef_[0])
print("Coefficient for Employment Rate (%):", model.coef_[1])

# Print mean squared error
print(f"Mean Squared Error: {mse}")

# Optionally, print a few predictions alongside actual values
comparison = pd.DataFrame({'Actual GRDP Growth Rate (%)': y_test, 'Predicted GRDP G
print(comparison.head())
```

```
Coefficients of the model:
Coefficient for Inflation Rate (%): -0.5050473859029889
Coefficient for Employment Rate (%): -0.0661619218644957
Mean Squared Error: 1.191187100337155
    Actual GRDP Growth Rate (%)  Predicted GRDP Growth Rate (%)
0                      3.822545                        2.597627
17                     2.614918                        2.617831
15                     1.236960                        2.875582
1                      1.779156                        2.540232
```

In a linear regression model, the coefficients represent the relationship between each independent variable and the dependent variable. They quantify the magnitude and

direction of the impact that each predictor variable has on the target variable. Here's a breakdown of how to interpret these coefficients:

Understanding the Coefficients Coefficient for Inflation Rate (%):

This coefficient tells us how much the GRDP Growth Rate (%) is expected to change for every one percent increase in the Inflation Rate, assuming all other variables remain constant. If this coefficient is positive, it indicates that as inflation increases, the GRDP growth rate tends to increase. Conversely, a negative coefficient would suggest that higher inflation is associated with a lower GRDP growth rate. Coefficient for Employment Rate (%):

Similarly, this coefficient measures the change in the GRDP Growth Rate (%) associated with a one percent increase in the Employment Rate, while holding other factors constant. A positive coefficient indicates that an increase in the employment rate is associated with an increase in the GRDP growth rate. If the coefficient were negative, it would suggest that higher employment rates are linked to lower GRDP growth rates. Practical Example