Multi Linear Regression Hands-on Training:

Problem Statement:

You are the Data Scientist at a X name company. Some of the companies to have better profit based on some features. You have to analyse the data of all the companies and find the overall predictions value of all companies.

Task to be done:

Part – 1: Data Manipulation:

- Load the data set as name as df (data set).
- Print the first 7 values of the df (data set).
- Print the last 7 values of the df (data set).
- Print the sum of the null values into the df (data set).
- Print the all feature's (Columns) data types to given the data set.
- Check the correlation of given the data set.
- Print the column names.

Part – 2: Data visualization:

- Build a Countplot for number of states occurring the df. Give the x-axis and y-axis names as "State" and "count" in previous one.
- Build a Boxplot for display the distribution between "state" and "profit".
- Build same boxplot distribution between as following as:
  1) State and Marketing Spend (x-axis, y-axis)
  2) State and Administration (x-axis, y-axis)
  3) State and R&D Spend (x-axis, y-axis)
- Build the Histogram for df where set the figure size = 7,7
- Build the pairplot for df where set the figure size = 6,6

- Build the displot for “profit” feature
- Build the heatmap for find out the correlation of given the data set.

Part – 3: Model Train and Test:

- Import the libraries for performing linear regression
- Set the X and y variables for allocating independent and dependent data from df.
- Print the X and y values into the array format only.
- Handle the categorical feature (state) convert to numerical format(Optional) or Just drop the categorical column.

Part – 4: Model Fitting:

- Import the train and test and split function.
- Set the train, test and split and set the test_size = 0.2, random_state = 0.
- Print all train and test shapes.
- Fit the linear model for “model” name instance.
- Print the coefficient and intercept values
- Find the r2_score.
- Print the MAE, MSE, RMSE values.

********** ********** END ********** **********