Analysing Ad Budgets for different media channels

Version : Python 3.7.6

Interface : Jupyter Lab

Summary :

1. Load the dataset
   - The advertisement dataset is loaded into a pandas DataFrame and the shape of the dataset is noted

2. Extract features
   - The features (different ad campaigns) are mapped to x variable and the target(sales) is mapped to a y variable

3. Identify the data type
   - The features are of continuous data type

4. Choose the model
   - Since the features are continuous, a supervised learning algorithm – linear regression is chosen

5. Train and test the model
   - The dataset is now split into training and testing set using the sklearn library's model_selection which provides the train_test_split method.

   - Create the estimator object for linear regression using scikit's class – linear_model

     Python code snippet

     ```
     from sklearn.linear_model import LinearRegression
     # instantiate the estimator
     Linreg = LinearRegression()
     ```

   - Using the fit method, the data_set is fit into the model
   - Using the predict method, the model can be tested on the testing dataset
   - Now the intercept can be seen using intercept_ and coefficient can be seen using the coef_ attributes
6. Strive for accuracy
   - From sklearn.metrics library the mean_squared_error is used to find the MSE for the model
   - The lower the MSE, the better the model

Note:

This assignment has been completed in Jupyter lab launched through the Anaconda Navigator on Desktop.