**Analysing Ad Budgets for different media channels**

DESCRIPTION

**Problem:**

The given dataset contains ad budgets for different media channels and the corresponding ad sales of XYZ firm. Evaluate the dataset to:

* Find the features or media channels used by the firm
* Create a model  to predict the sales outcome
* Split as training and testing datasets for the model
* Calculate the Mean Square Error (MSE)

Table

Description automatically generated

* Find the features or media channels used by the firm

Graphical user interface, text, application

Description automatically generated

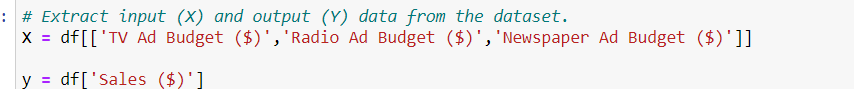
Graphical user interface, text, application

Description automatically generated

The correlation matrix suggests that there is a strong correlation between Sales and TV Ad Budget. The values is continuous. So here we use Multiple linear Regression to predict the Sales based on TV , Radio and Newspaper Ad Budgets.

Chart, scatter chart

Description automatically generated



Since we are using Multiple Linear Regression, X will be the independent variables - TV , Radio and Newspaper Ad Budgets. Y will be the dependent or response variable – Sales.

Graphical user interface, text, application, email

Description automatically generated

Data is randomly split into train and test datas. Train data is 80% of the total data which is 160 records and test data is 40 records.

Graphical user interface, text, application

Description automatically generated

Here we have the intercept and coefficients needed for the linear regression equation y = mx + c where m is coefficient and c is y-intercept.

Text

Description automatically generated

Above, we have the prediction sales values.

Text

Description automatically generated

The root mean square error for the test data is 1.316.

Graphical user interface, text, application

Description automatically generated

The R-Square value is 0.89. Any value closer to 1 suggests a good model. Here our R-square value of 0.89 suggests that our model is close to accurate in predicting the sales. You can see from the below chart that the model does a nice job of predicting values.

Chart, line chart

Description automatically generated

Chart, scatter chart

Description automatically generated