Business Analytics with Excel

Predicting restaurant tip using predictive analytics on Excel

-Amulya Billemane M S

Description

Use excel to predict restaurant tips.

Description:

The dataset in file ***Restaurant tips dataset.xlsx*** contains tips data for different customers. The following are the features in the dataset:

Sex	Gender of the customer
Smoker	Indicates if the customer is a smoker or not
Day	Day of the restaurant visit
Time	Indicates whether the tip was for lunch or dinner
Size	Number of members dining
Total Bill	Bill amount in USD
Tip	Tip amount in USD

The following project tasks are required to be performed in excel:

- Use the restaurant tips file for the analytics using Excel
- Find out if there are any missing values and clean the data
- Find the features that are independent and dependent
- Identify which predictive problem is needed.
- Encode the categorical variables to numeric values using IF conditions
- Build an appropriate model with the dataset.
- Calculate the predicted and actual tips values.
- Calculate the RMSE(Root Mean Square Error) of the model. RMSE is the root of the mean of square errors.

Tools required: Microsoft Excel, Data Analysis Add-in.

Expected Deliverables: Model to predict restaurant tips given input values with the mathematical equation for predicting the tips value.

In this project:

- I utilized the restaurant tips file to conduct comprehensive analytics using Excel.
- The initial steps involved thorough scrutiny of the data to identify and rectify any missing values.
- Following this, I meticulously classified the independent variables as Sex, Smoker, Day, Time, Size, and Total Bill, with Tip identified as the dependent variable.
- The categorical values were systematically converted into numerical equivalents using the IFS function.
- Multiple regression analysis was then employed to derive a data model as evidenced in the Excel files.
- Subsequently, the predicted tip was computed utilizing the predictive analysis multiple regression formula, enabling an assessment of the disparity between the actual and predicted tip.
- Subsequently, the RMSE (Root Mean Square Error) value was calculated to gauge the model's accuracy.
- Finally, the tip values were visually represented in a graphical format.