DATA AND
ARTIFICIAL INTELLIGENCE

Business Analytics with Excel Certification Training

simplilearn

DATA AND
ARTIFICIAL INTELLIGENCE

Predicting Restaurant Tips Using Predictive Analytics on Excel

simplilearn

Objectives

- To predict restaurant tips given input values with the mathematical equation for predicting the value of the tip

Prerequisites

- Data cleaning and preparation
- Data analysis using statistics
- Formatting in excel

Industry Relevance

- **Data cleaning and preparation:** It is the process of cleaning and transforming raw data before processing and analysis.
- **Data analysis using statistics:** In statistical analysis, large amounts of data are collected, analyzed, and converted into valuable information by identifying patterns and trends.
- **Formatting in excel:** It helps to make data look more interesting and descriptive.

Problem Statement

Use Excel to build a model to predict restaurant tips given input values with the mathematical equation for predicting the value of the tips

Dataset Description

Variable	-	Description
• 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

Tasks to Perform

Perform the below tasks on the dataset provided using Excel:

1. Use the restaurant tips file for the analytics using Excel
2. Find out if there are any missing values and clean the data
    - Remove rows with at least one empty cell; choose all data using Control+A and in the home tab, click on Find & Select and go to Special.
    - Check for duplicates: Choose the dataset using Control+A; In the Data tab, click on Remove duplicates
3. Find the features that are independent and dependent
4. Identify which predictive problem is needed

Tasks to Perform

Perform the below tasks on the dataset provided using Excel:

5. Encode the categorical variables to numeric values using IF conditions
    - For each independent numeric value, find its correlation coefficient with respect to the tip. The correlation coefficient is derived from the formula CORREL(array1, array2) where array1 is all the independent variables and array2 is the tip variable.

6. Build an appropriate model with the dataset
7. Calculate the predicted and actual tips values
8. Calculate the RMSE (Root Mean Square Error) of the model. RMSE is the root of the mean of square errors

Project Outcome

- The aim of the project is to predict restaurant tips given input values with the mathematical equation for predicting the value of the tips.
- You will be able to understand the restaurant tips based on what is the customer's age and gender, and what is the time of the day.

Submission Process

1. Complete the project in the Simplilearn lab
2. Complete each task listed in the problem statement
3. Take screenshots of the results for each question and the corresponding code
4. Save it as a document and submit using the assessment tab
5. Tap the "Submit" button (this will present you with three choices)
6. Attach three files and then click "Submit“

Note: Be sure to include screenshots of the output

DATA AND
ARTIFICIAL INTELLIGENCE

Thank You

simplilearn