# Business Analytics with Excel Certification Training

Course-End Project Problem Statement

# Course-End Project - 2

# Predicting Restaurant Tips

## Overview

In this project, you must analyze a dataset containing restaurant tips data. You aim to clean the data, identify the independent and dependent variables, and build a predictive model to estimate tip amounts. You will use regression analysis to encode categorical variables into numeric values. The project aims to provide insights that can help improve service and customer satisfaction in the restaurant industry.

## Instructions

- Review the learning materials in BA with Excel
- Carefully read the situation, tasks, actions, and result sections to grasp the assignment fully
- Complete and submit your assignment via the Learning Management System (LMS)
- Follow the provided guidelines closely, ensuring your report includes all required analyses and interpretations

## Situation

You are a data analyst working for a restaurant chain. The management has provided you with a dataset (Restaurant tips dataset.xlsx) containing information about customer tips. The dataset includes customer gender, smoking status, day of visit, and total bill amount. They are looking for actionable insights to improve service and increase customer satisfaction.

## Task

Your task is to analyze this data using Excel and build a predictive model to estimate tip amounts.

## Input Dataset

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

| sex | The 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 |

## Action

- Open the dataset input file: **Restaurant tips dataset.xlsx**
- Check for missing values or duplicates and clean the data
- Convert categorical features (sex, smoker, day, and time) to numeric values using IF statements
- Create a pivot table to sum the tips for males and females separately, then create a pie chart
- Create another pivot table to sum the tips for males and females at various times, and then create a bar chart
- The tip feature is the dependent variable, while the rest are independent variables. Find the correlation between the tips and the other variables.
- The problem is a multiple regression problem. Use the data analysis add-in for regression.
- Use coefficients and intercepts to form the regression equation
- Run the regression equation for all input values to get the predicted y

## Result

Design a tips prediction model that estimates tip amounts based on customer features. This model provides insights for improving customer service and satisfaction by enabling the restaurant chain to make data-driven decisions. Upload the Excel file to the Learning Management System (LMS).