**Project: Restaurant Tips Data Analysis**

# Objective

The goal of this project was to analyze the Restaurant Tips Dataset to understand the factors influencing tipping behavior. Using Excel, the dataset was cleaned, transformed, and analyzed with pivot tables, charts, and regression techniques to extract meaningful insights.

# Steps Performed

## 1. Data Preparation

- Opened the dataset file: Restaurant tips dataset.xlsx.  
- Checked for missing values and duplicate entries, then cleaned the data to ensure accuracy.  
- Converted categorical variables — sex, smoker, day, and time — into numeric values using Excel IF statements, making them usable for correlation and regression analysis.

## 2. Exploratory Analysis

- Created a pivot table to calculate the total tips given by males and females separately.  
 - Visualized the results with a pie chart to compare the proportion of tips by gender.  
- Built another pivot table to calculate tips by gender across different times (Lunch/Dinner).  
 - Represented the results in a bar chart, highlighting tipping differences by gender and time.

## 3. Correlation Analysis

- Considered Tip as the dependent variable and all other features as independent variables.  
- Calculated the correlation coefficients to measure the strength of relationships between tips and other features (bill size, gender, smoking status, day, and time).

## 4. Regression Modeling

- Defined the problem as a multiple regression problem with "Tip" as the output variable.  
- Used the Excel Data Analysis Toolpak to run regression analysis.  
- Extracted coefficients and intercepts from the regression results to construct the regression equation:

Tip (ŷ) = Intercept + b₁x₁ + b₂x₂ + … + bₙxₙ

- Applied the regression equation across the dataset to generate predicted tip values (ŷ).

# Key Learnings

- Learned data cleaning techniques (handling missing values and duplicates).  
- Practiced categorical-to-numeric transformation using Excel IF statements.  
- Strengthened skills in creating pivot tables and visualizations (pie and bar charts).  
- Understood how to measure correlations between dependent and independent variables.  
- Gained hands-on experience with multiple regression analysis in Excel.  
- Developed the ability to form a regression equation and use it for prediction.

# Conclusion

This project provided a complete workflow of data cleaning, visualization, statistical analysis, and prediction modeling using Excel. The analysis highlighted the role of factors such as gender and time on tipping patterns, while regression helped quantify the influence of multiple variables on the tip amount.