Restaurant Tips

Introduction

predicting restaurant tips using Excel regression analysis. In this project, analyse a dataset containing restaurant tips data. Clean the data, identify the independent and dependent variables, and build a predictive model to estimate tip amounts. use regression analysis to encode categorical variables into numeric values. The dataset includes customer gender, smoking status, day of visit, and total bill amount. And finally use dashboard to present data in a clear, visual format.

Objective

To build a model that estimates tips based on factors like total bill, gender, day, smoker, size and time. Summarize insights from pivot tables and charts, and create a dashboard.

Data Preparation

removing duplicates, and encoding categorical variables. The data headers were formatted in bold, and borders were applied to all the data for clear separation and readability.

Exploratory Data Analysis

Summarize insights from pivot tables and charts. A pivot table was created to examine tip totals by gender. The results show that male customers contributed $485, while female customers contributed $245, confirming that males provided a significantly larger share of total tips. This distribution was visualized using a pie chart, highlighting the dominant proportion of tips coming from male customers.

Another pivot table was used to analyse tips by gender across meal times. During dinner, males tipped $390.96 and females tipped $156.11, showing a large difference. During lunch, the tip values were much closer, with males tipping $95.11 and females tipping $88.40. These values were displayed using a bar chart, clearly showing that dinner contributes the majority of tips and that male customers leave higher tips overall.

The correlation analysis shows that most categorical variables such as gender, smoker status, day, and time have negligible relationships with tip amounts, with correlation values close to

zero. In contrast, bill amount and table size show meaningful positive correlations with tip, at 0.675 and 0.488 respectively, indicating that higher bills and larger parties tend to leave higher tips.

Model Building

The regression model explains 46.9% of the variation in tip amounts, indicating a moderately strong predictive capability. The model is statistically significant (F = 34.72), confirming that the predictors collectively influence tip behaviour. Total Bill coefficient is 0.09437 it means, it is a highest impact variable. As the bill increases, the tip rises proportionally. Group Size coefficient is 0.1739 it means larger groups tend to leave higher tips. Each additional person increases the tip amount. Gender, smoker status, time of day, and day of the week show no significant effect on tips and do not contribute meaningfully to prediction.

Dashboard

Dashboard to present data in a clear, visual format. Total 243 customers, this is the total number of transactions in the dataset. Average total bill is $20 this means on average customers spend $20 per bill. Average tip is $3 that means, the typical tip amount left by customers is $3. Four charts presented in the dashboard, Pie Chart, Bar Chart, Scatter Plot, Column Chart.

Tip totals by gender were visualized using a pie chart, male customers contributed $485, while female customers contributed $245, it highlighting the dominant proportion of tips coming from male customers.

Using a bar chart clearly showing that, during dinner, males tipped $390.96 and females tipped $156.11, showing a large difference. During lunch, the tip values were much closer, with males tipping $95.11 and females tipping $88.40. Dinner contributes the majority of tips and that male customers leave higher tips overall.

The scatter charts each green dot represents a single observation; the predicted tip plotted against the actual tip. X-axis: Likely shows the predicted tip amount and Y-axis: Shows the actual tip amount. Upward trend indicates that as predicted tips increase, actual tips tend to increase.

Tip by Day (Column Chart) chart shows total tips collected on different days. Sunday tip collected $260, Saturday tip collected $270, Thursday tip collected $180, Friday tip collected $60. Weekends (Sat & Sun) bring in the highest tip totals. Friday has the lowest tip total.

Conclusion

The dashboard analysis shows clear patterns in customer tipping behaviour. Male customers consistently contribute a higher portion of total tips compared to female customers, accounting for almost twice the amount. Dinner time generates the largest share of tips, with both genders tipping more during dinner than lunch. Weekend days especially Saturday and Sunday show the highest total tip amounts, indicating peak customer activity and spending during those days.

The average customer spends $20 and leaves an average tip of $3, giving a quick snapshot of typical transaction behaviour. Overall, the insights indicate that tips are strongly influenced by customer spending levels, group size, and the time and day of the visit, with weekends and dinner hours being the most profitable periods.