

# Predicting Restaurant Tips Using Machine Learning

Leon Motang

November 30, 2025

## Abstract

This study explores the prediction of restaurant tips based on various customer and transaction features using machine learning techniques. Using a dataset of restaurant bills, we analyze the relationship between total bill, party size, day, time, and customer attributes such as sex and smoking status, and the corresponding tip amount. Regression models are trained to estimate expected tips, enabling restaurant management to understand tipping patterns and optimize service.

## 1 Introduction

Predicting tips in restaurants is a valuable problem for understanding customer behavior and optimizing service. This project leverages machine learning techniques to predict the amount of tip a waiter might receive based on the attributes of the transaction.

The dataset and source code are available at the GitHub repository: Waiter Tips Prediction with Machine Learning.

## 2 Dataset and Preprocessing

The dataset contains the following features:

- **Total bill:** Amount of the bill in dollars.
- **Tip:** Amount tipped by the customer.
- **Sex:** Customer gender.
- **Smoker:** Whether the customer is a smoker.
- **Day:** Day of the week.
- **Time:** Lunch or Dinner.
- **Size:** Number of people in the party.

Categorical variables were encoded numerically for modeling:

- Sex: Female = 0, Male = 1

- Smoker: No = 0, Yes = 1
- Day: Thur = 0, Fri = 1, Sat = 2, Sun = 3
- Time: Lunch = 0, Dinner = 1

### 3 Exploratory Data Analysis

Figure 1 shows the relationship between total bill and tip, where bubble size represents party size, and color indicates the day of the week.

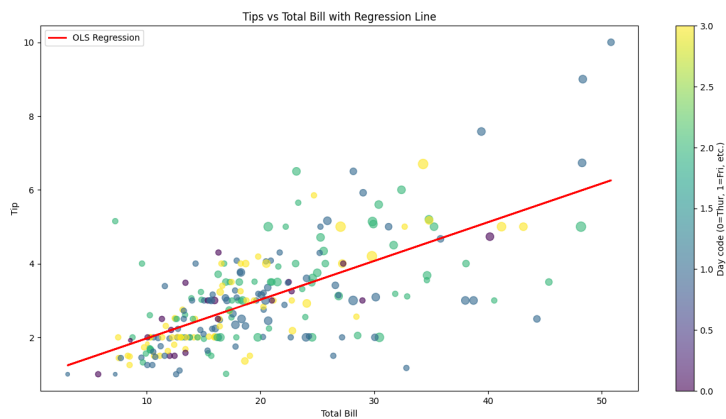


Figure 1: Scatter plot of total bill vs. tip with regression line.

Figure 2 shows the distribution of tips by time of day, highlighting when more tips occur.

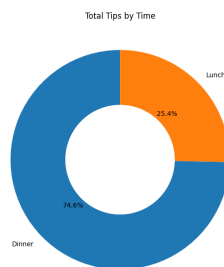


Figure 2: Pie chart showing distribution of tips by time (Lunch/Dinner).

### 4 Modeling

A linear regression model was trained using scikit-learn with the features: total bill, sex, smoker, day, time, and size. The model can predict the tip amount for a new transaction. For example, for a party with a total bill of \$24.50, consisting of 4 diners during dinner on Thursday, with a male non-smoker, the predicted tip is \$3.74.

## 5 Conclusion

This project demonstrates the application of machine learning for predicting restaurant tips. The regression model provides insight into tipping patterns, which can help restaurant staff anticipate earnings and improve service. Future work could explore more advanced models, such as ensemble methods, for higher predictive accuracy.