Data Analysis Report on Customer Tipping Behavior

Marketing Analytics Report

Anshul Mantri
Electrical Engineering
IIT Gandhinagar
22110143

Shreel Chawla
Electrical Engineering
IIT Gandhinagar
22110243

Shrinivas Kulkarni
Civil Engineering
IIT Gandhinagar
22110126

Paras Yogi
Civil Engineering
IIT Gandhinagar
22110180

Pavani Khale
Chemical Engineering
IIT Gandhinagar
22110191

***Abstract*—This report analyzes a restaurant tips dataset containing information on bill amounts, tips, payment methods, group size, day of the week, and server identity. The purpose of the study is to explore the key factors that influence tipping behavior and to identify patterns across different servers, days, and payment types. The analysis highlights how variables such as bill size, credit card usage, and group size affect both the amount and percentage of tips. Along with these findings, the report also provides guidance on designing marketing dashboards, selecting appropriate KPIs, and presenting results in a clear format for decision-makers.**

I. Introduction

Tipping is an important aspect of the restaurant business because it reflects both customer satisfaction and service quality. For managers, understanding tipping behavior can provide valuable insights into staff performance, customer preferences, and overall dining experiences. The dataset used in this report includes detailed information such as bill amounts, tips, method of payment, number of guests, the day of the week, and the server responsible for each table.

The main goal of this study is to examine what influences how much customers tip. Specifically, the analysis looks at whether larger bills lead to higher tips, if customers paying with credit cards tend to give more, and how tipping behavior changes across different days and servers. By identifying these patterns, the report helps explain the factors that encourage higher tipping and points out areas where service or management strategies could improve.

In addition to analyzing customer tipping behavior, the report also connects the findings to practical applications. It provides guidance on how restaurants can use these insights to design effective marketing dashboards, track key performance indicators (KPIs), and present results in a clear way for decision-makers. This ensures that the study is not only about data analysis but also about turning insights into useful strategies for restaurant growth and better customer service.

II. Data Description

The dataset used for this analysis provides detailed information about customer bills and tipping behavior in a restaurant setting. Each record captures key aspects of a dining transaction, allowing us to explore both the financial and behavioral side of customer service.

The main columns in the dataset are:

- **Bill** – The total amount spent on the meal, which serves as the base for understanding tipping behavior.

- **Tip** – The amount of money left by the customer as a gratuity, reflecting their satisfaction and generosity.
- **Credit** – Indicates whether the payment was made using a credit card or not. This helps compare tipping patterns between cash and card payments.
- **Guests** – The number of people in the dining group, which can influence both the total bill and the size of the tip.
- **Day** – The day of the week when the transaction occurred, coded as f (Friday), th (Thursday), w (Wednesday), t (Tuesday), and m (Monday). This allows for the study of day-wise trends.
- **Server** – Identifies the server (A, B, or C) who attended the table, making it possible to compare performance across staff.
- **PctTip** – The percentage of the bill given as a tip, which standardizes tips relative to bill size and helps compare across transactions.

The dataset spans multiple weekdays and covers a variety of group sizes, ranging from solo diners to larger parties. Tip percentages vary greatly, from as little as 6% to as high as 40%, showing both modest and highly generous customer behaviors. This wide variation makes the dataset particularly useful for examining what drives tipping differences, how servers perform, and how context such as payment type or group size impacts customer generosity.

III. Methodology

To carry out the analysis of this dataset, a step-by-step approach was followed to ensure that the findings are accurate, meaningful, and easy to interpret. The process included the following key stages:

- **Data Cleaning** - The dataset was first reviewed to check for errors or inconsistencies. This included verifying that day labels were coded correctly, ensuring that all numerical values such as bills, tips, and tip percentages were valid, and confirming that there were no missing or duplicate records that could affect the analysis.
- **Descriptive Analysis** - Basic statistical measures such as averages, minimums, maximums, and ranges were calculated for bill amounts, tip values, and tip percentages. This helped in understanding the overall spread of the data and identifying general patterns in tipping behavior.
- **Comparisons** - The dataset was then examined more closely by comparing tipping behavior across different categories. This included analyzing differences between servers, variations in tipping based on payment method (cash vs. credit), and how group size influenced both the total tips and the percentage of tips.
- **Visualization Suggestions** - To make the data easier to interpret, suitable chart types were identified. Scatter plots were suggested to show the relationship between bill size and tip amount, bar charts to compare the performance of different servers, and line charts to highlight tipping trends across different days of the week.
- **Interpretation** - Finally, the statistical results and visual patterns were translated into meaningful insights. This step connected the data to real-world behavior, showing how customer tipping habits are influenced by various factors and what that means for restaurant management and decision-making.

IV. Data Analysis

1) What are the mean, median, mode, minimum, and maximum values for Bill, Tip, and PctTip?

	Mean	Median	Mode	Min	Max
Bill	22.729045	20.22	18.62	1.66	70.51
Tip	3.849299	3.35	2.00	0.25	15.00
PctTip	16.619108	16.20	14.90	6.70	42.20

Bills mostly average around $25 but still vary widely. Tips mostly center around $3-$4 , with mostly people tipping in the range of

15-20%.The highest recorded percentage tip is 42.2%, while the lowest is only 6.7%.

2) What is the distribution of Bill and Tip amounts? Are they normally distributed or skewed?

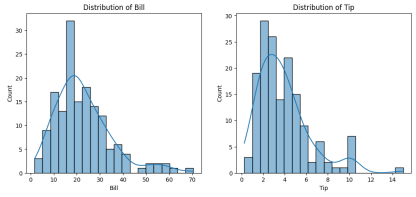


Fig. 1. Distribution of bill and tip

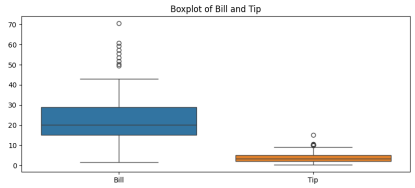


Fig. 2. Boxplot of Bill and Tip

Bill: Right-skewed (most bills are moderate, but a few very large bills pull the average upward). Tip: Also right-skewed (most tips are small, but some large tips occur for big bills) Both Bill and Tip amounts are not normally distributed. Instead, they are positively skewed, meaning most customers spend moderately, while a few spend a lot and tip big. .

3) What is the average percentage tip across all transactions? Does it vary a lot? (Mean, median, and standard deviation of PctTip)

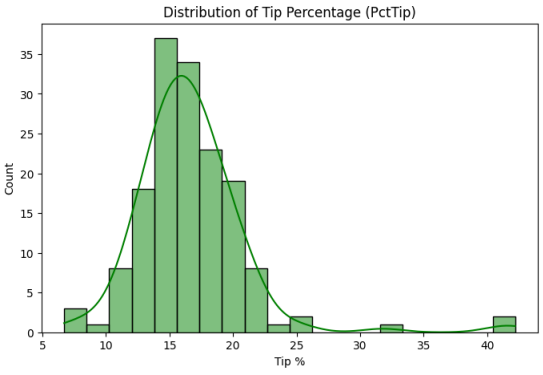


Fig. 3. Distribution of Tip percentage(PctTip)

Average Tip %: 16.62 Median Tip %: 16.20 Standard Deviation: 4.39.
On average, people tip about 16–17% of the bill. The standard deviation shows some variation: most people tip in the 15–20% range, but there are occasional very high or very low tips. Compared to Bill and Tip, PctTip is less skewed, meaning tipping percentages are relatively consistent.

4) Who tips better on average — small groups or large groups?

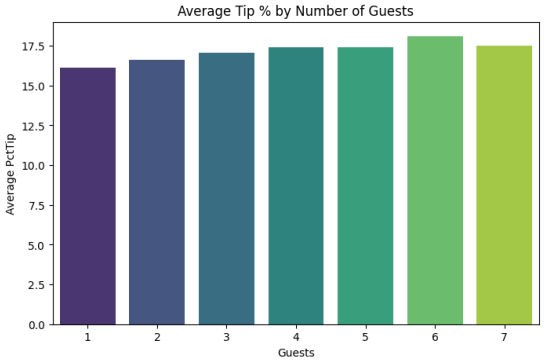


Fig. 4. Average tip percentage by number of guests

Smaller groups (1–2 guests) tend to tip better (higher percentage). As the group size increases, the tip percentage decreases. Likely because in large groups, the total bill is already high, and customers leave a fixed tip rather than scaling up perfectly.

5) Is there a significant positive relationship between the amount of the bill and the tip given by customers?

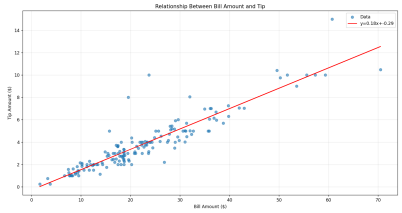


Fig. 5. Tip amount vs. Bill amount scatter plot

The analysis shows a strong positive relationship between the bill amount and the tip given by customers, with a correlation

coefficient of 0.92. This means that as the bill size increases, the tip amount also tends to increase in a nearly linear fashion. The scatter plot, along with the regression line, visually confirms this trend—larger bills are generally associated with larger tips.

6) How does tip percentage vary with the total bill size?

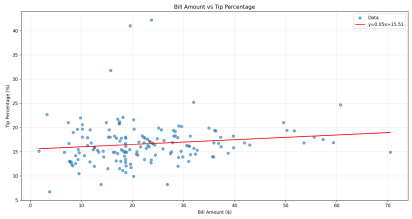


Fig. 6. Tip Percentage vs. Bill amount scatter plot

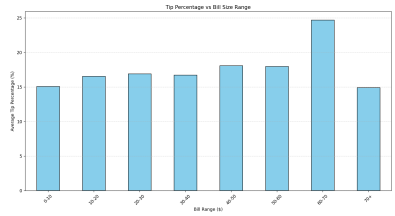


Fig. 7. Average tip percentage across different bill size ranges

The scatter plot of bill amount versus tip percentage, along with the regression line, shows a very weak upward trend, supported by the correlation coefficient of 0.14. This indicates that tip percentage does not strongly depend on the size of the bill. In other words, while customers leave larger absolute tips for higher bills, the proportion of the bill they tip tends to remain fairly constant, with only minimal variation.

This observation is further supported by the bar plot of average tip percentage across different bill ranges. While the 60–70 dollar range shows a slightly higher average tip percentage compared to other groups, the overall pattern remains flat, indicating no consistent trend. Together, these results suggest that customers generally tip similar percentages regardless of the total bill size.

7) How does group size influence tipping behavior, both in terms of absolute tip amounts and tip percentages?

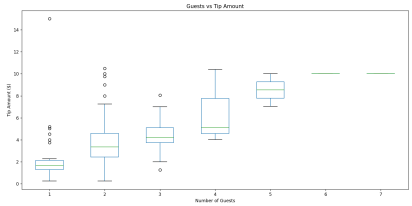


Fig. 8. Distribution of tip amount across different group sizes.

For tip amount vs number of guests, the steady increase in median tip amounts with larger groups suggests that larger groups tend to leave higher tip amounts overall. This is intuitive since bills for larger groups are generally higher.

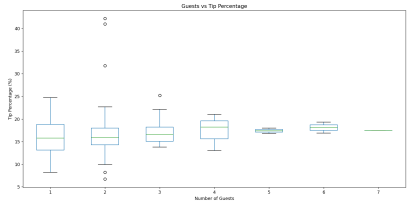


Fig. 9. Distribution of tip percentages across different group sizes.

However, for tip percentage vs number of guests, the overlapping boxes and lack of a positive trend indicate that larger groups do not consistently leave higher tip percentages. The tip percentage remains fairly stable or variable across group sizes without a clear increasing pattern.

Therefore, larger groups tend to leave higher absolute tip amounts because their bills are larger, but they do not tend to leave higher tip percentages of the bill. Tip generosity as a proportion of the bill size does not increase with party size based on your plot observations. This aligns with common restaurant tipping behavior where tip percentage usually stays within a certain range regardless of party size,

while total tip grows with the bill total and number of guests.

8) What is the distribution of payment modes (credit vs cash) in terms of total bills and total tips?

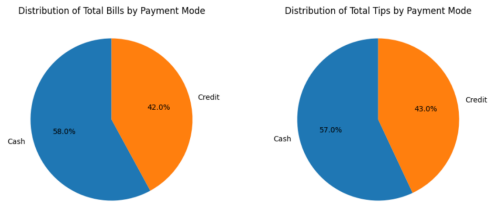


Fig. 10. Distribution of Payment Modes

The pie charts show the distribution of payment modes (credit vs cash) for both total bills and total tips. A majority of customers pay in cash (58% of bills), while credit payments account for 42%. The distribution of tips closely mirrors this, with 57% of tips coming from cash transactions and 43% from credit.

9) If we categorize customers into based on their tipping percentage, which category do most customers fall into, and how does this distribution look?

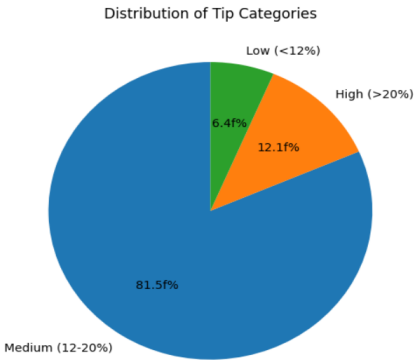


Fig. 11. Distribution of Tip Categories

The pie chart shows the distribution of customers across three tipping categories. The majority of customers (81.5%) fall into the medium tipper range (12–20%), while high tippers (> 20%) make up 12.1% of the total. Only a small proportion (6.4%) are low tippers (< 12%).

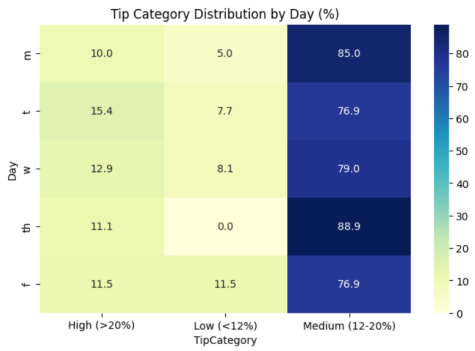


Fig. 12. Tip Category Distribution by Day

The heatmap shows how tipping categories (low, medium, high) are distributed across different days of the week. Medium tippers (12–20%) dominate every day, with shares ranging from 77% to 89%. High tippers (> 20%) appear consistently across all days, varying between 10–15%, with Tuesday having the highest proportion of high tippers (15.4%). Low tippers (< 12%) remain the smallest group overall, with Friday showing the highest proportion (11.5%) and Thursday showing none.

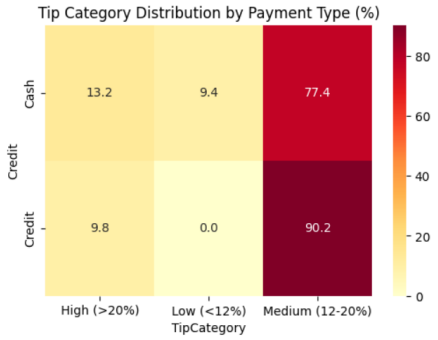


Fig. 13. Tip Category Distribution by Payment Type

Among customers who paid with cash, the majority of tips (77.4%) fell into the medium category (12–20%), while 13.2% were classified as high tips (greater than 20%) and 9.4% as low tips (less than 12%). In contrast, for customers who paid by credit, tipping behavior was even more concentrated in the medium range, with 90.2% of tips falling between 12–20%. High tips accounted for 9.8% of credit payments, whereas no customers in this group left tips below 12%.

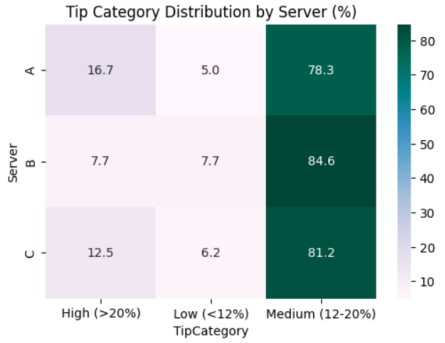


Fig. 14. Tip Category Distribution by Server

The heatmap illustrates how tips are distributed among different servers (A, B, and C) across three tip categories: High (¿20%), Low (¡12%), and Medium (12-20%). Server B has the highest percentage of Medium tips (84.6%) and a balanced Low tip percentage (7.7%). Server A has the highest percentage of High tips (16.7%) but also the lowest percentage of Low tips (5.0%). Server C falls in between with 12.5% High tips and 6.2% Low tips, with the majority of tips being Medium (81.2%).

10) Which days contribute the most to overall revenue and overall tips?

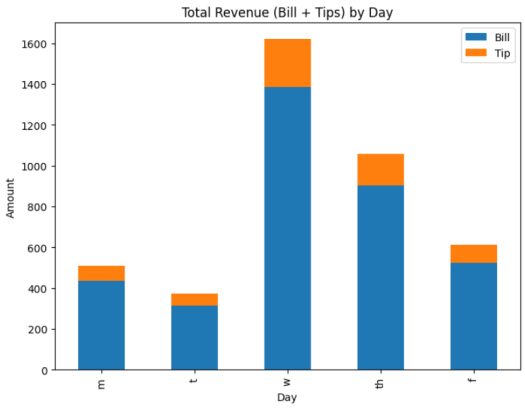


Fig. 15. Revenue With Respect To Day

The stacked bar chart displays the contribution of each day to overall revenue (bills + tips). Among all days, Wednesday shows the highest total revenue, with a significant share of both bills and tips. Friday is the second-highest contributor, followed by smaller but steady contributions on other days. Tips generally scale with the size of the bills across all days, indicating proportional tipping behavior.

11) Which day of the week brings in the highest average bill, highest tips, and highest tip percentage?

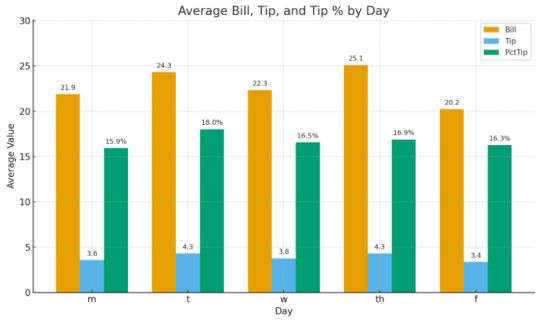


Fig. 16. Average Bill, Tip, and TIP % by Day

The grouped bar chart compares the average bill amount, tip, and tip percentage across weekdays from Monday to Friday. The results indicate that Thursday records the highest average bill, making it the strongest day for overall revenue, while Tuesday stands out for having both high average tips in dollar value and the highest tip percentage, showing that customers are most generous relative to their spending on this day. In contrast, Friday consistently shows the lowest averages across all three measures, making it the least favorable day for both revenue and tips. Overall, Tuesday and Thursday emerge as the most profitable and rewarding days, but for different reasons Thursday for higher customer spending and Tuesday for tipping behavior.

12) Is there a significant difference in average tip percentage between different days of the week?
H_0 (Null hypothesis) : Mean tip percentage is the same for all days.
H_1 (Alternative hypothesis) : At least one day has a different mean tip percentage.

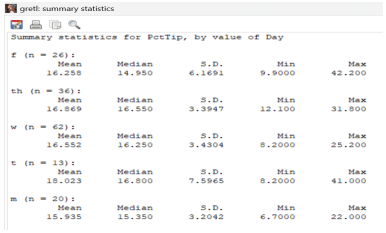


Fig. 17. Statistical data of Percentage Tip across days

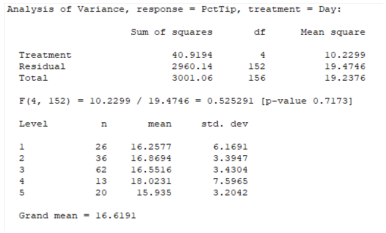


Fig. 18. ANOVA - PercentTip across different days

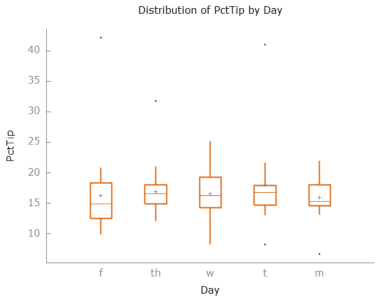


Fig. 19. Box plot showing variation in tip percentage by day

From the test results we get the ,
F- value = 0.5252
P-value = 0.7173

Since P-value is greater than 0.05, we fail to reject the null hypothesis. Thus there is no statistically significant difference in the average tip percentage between different days of week.

From this data, we can infer that the tip percentage is consistent across all days and the day of the week does not influence the tipping behaviour.

13) Do different servers receive significantly different tips on average?
$\mathbf{H}_0$: Mean tip percentage is the same for all servers.
$\mathbf{H}_1$: At least one server has a different mean tip percentage.

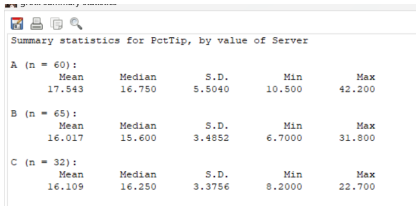


Fig. 20. Statistical data of Percentage Tip across servers

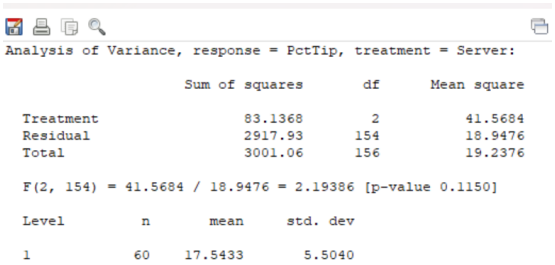


Fig. 21. ANOVA - PercentTip across different servers

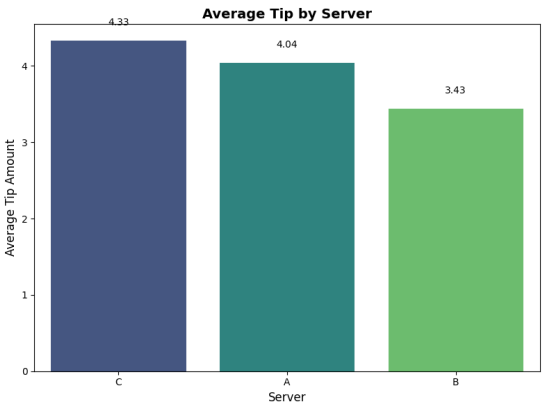


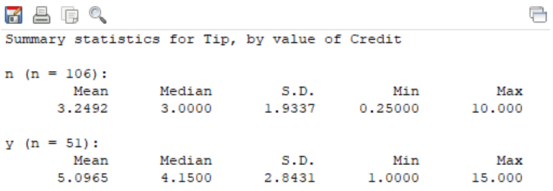
Fig. 22. Bar graph showing tip percentage variation among servers

From the test results we get the ,
F- value = 2.1938
P-value = 0.1150

Since P-value is greater than 0.05, we fail to reject the null hypothesis. Thus there is no statistically significant difference in the average tip percentage received by different servers.

From this data, we can infer that servers relieve similar tip percentages on average and server identity does not significantly affect the tipping behavior.

14) Is the average tip for credit card customers significantly higher than cash customers?
$\mathbf{H}_0$: Average tip for credit and cash customers is the same.
$\mathbf{H}_1$: There is a significant difference in average tips between credit and cash customers.

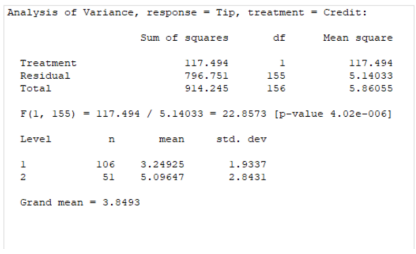


```
Summary statistics for Tip, by value of Credit

n (n = 106):
      Mean      Median        S.D.         Min         Max
    3.2492      3.0000      1.9337     0.25000      10.000

y (n = 51):
      Mean      Median        S.D.         Min         Max
    5.0965      4.1500      2.8431      1.0000      15.000
```

Fig. 23. Statistical data of Tip across payment mode



```
Analysis of Variance, response = Tip, treatment = Credit:

                    Sum of squares       df      Mean square

Treatment                117.494        1          117.494
Residual                 796.751      155          5.14033
Total                    914.245      156          5.86055

F(1, 155) = 117.494 / 5.14033 = 22.8573 [p-value 4.02e-006]

Level         n       mean     std. dev

1           106    3.24925       1.9337
2            51    5.09647       2.8431

Grand mean = 3.8493
```

Fig. 24. ANOVA - Tip across different payment modes

Fig. 25. Box plot showing variation in tip percentage by payment mode

From the test results we get the ,
Mean tip by cash customers = 3.2492
Mean tip by credit customers = 5.0965
F- value = 22.8573
P-value = $4.02*10^{-}6$

Since the P-value is lesser than 0.05, we reject the null hypothesis. Thus there is a statistically significant difference in the average tips given by credit and cash customers.

From this data, we can infer that credit card customers on average tip more compared to cash customers. Thus payment method has a significant impact on the amount tipped.

15) To analyze whether server performance (measured by PctTip) changes across different days of the week and to check for any interaction effect between Server and Day.
$\mathbf{H}_0$: Server performance (average tip) does not vary by day.
$\mathbf{H}_1$: Server performance (average tips) varies by day.

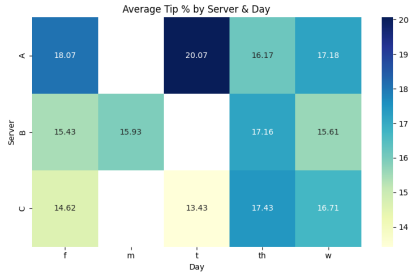


Fig. 26. Heatmap- Avg PercentTip across different day and server

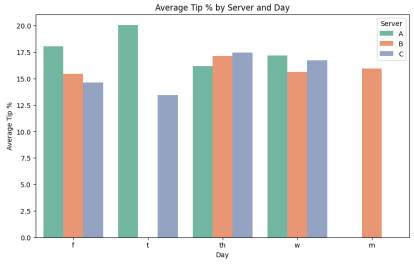


Fig. 27. Bar Graph between server and day for tip percentage

```
----- Two-Way ANOVA Table -----
                      sum_sq     df         F    PR(>F)
C(Server)           0.302114    2.0  0.007989  0.992044
C(Day)             30.375764    4.0  0.401599  0.527263
C(Server):C(Day)  208.201313    8.0  1.376319  0.227948
Residual         2741.841328  145.0       NaN       NaN

----- P-Values from ANOVA -----
C(Server)          0.992044
C(Day)             0.527263
C(Server):C(Day)   0.227948
Residual                NaN
```

Fig. 28. ANOVA-server performance and day interaction

From the ANOVA, we observed that the p values are greater than 0.05 thus we fail to reject the null hypothesis.

We can say server performance does not significantly change by day. No strong interaction effect is observed between server and day. In the graphs too some small variations are observed on some days. However these are not statistically significant.

This helps the restaurant/business understand whether certain servers perform better on specific days and if customer tipping behavior depends on the day and server together.

16) Does tipping behavior (PctTip) differ by payment mode (credit vs cash) and group size, and is there an interaction between them?
$\mathbf{H}_0$: No difference in average tip between credit and cash and also across guest count thus no interaction.
$\mathbf{H}_1$: At least one difference exists.

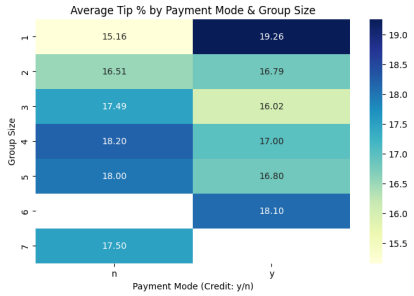


Fig. 29. Heatmap- Avg PercentTip across different payment mode and group size

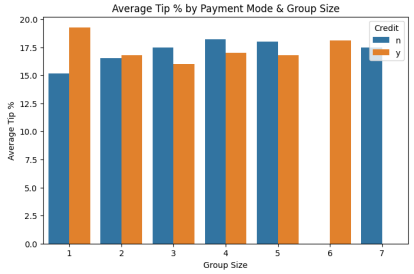


Fig. 30. Bar Graph between Group size and payment mode for tip percentage

```
----- Two-Way ANOVA Table -----
                      sum_sq     df         F    PR(>F)
C(Credit)                NaN    1.0       NaN       NaN
C(Guests)          12.969991    6.0  0.108867  0.990223
C(Credit):C(Guests)  112.406177    6.0  0.943513  0.454851
Residual           2879.113917  145.0       NaN       NaN

----- P-Values from ANOVA -----
C(Credit)                NaN
C(Guests)           0.990223
C(Credit):C(Guests)    0.454851
Residual                 NaN
Name: PR(>F), dtype: float64
```

Fig. 31. ANOVA- payment mode and group size interaction

From the ANOVA, we observed that the p values are less than 0.05 thus we fail to reject the null hypothesis.

We can say that payment mode or group size does not affect tip percentage. There is no significant interaction between payment mode and group size. Tipping behavior remains consistent regardless of payment method and group size. Also the reason for the NAN values can be very little variance in tip percentage within credit or cash groups.

This analysis helps us understand whether larger groups tend to tip differently when paying by credit card versus cash. It provides insights into customer behavior and can guide billing strategies.

V. Insights and Recommendations

Our analysis of the restaurant's tipping data reveals several key trends that have direct, practical implications. By understanding what drives customer tipping, management can make informed decisions to boost both revenue and staff morale. This section discusses the main findings from the data and provides recommendations based on each insight.

Insight 1: Bigger Bills Mean Bigger Tips

The most straightforward finding is that the total bill amount is the main driver of the tip amount. We saw a strong, positive correlation (0.92) here, which basically means that when a customer's bill is higher, their tip is almost always higher too. It's a very reliable pattern.

- **Recommendation:** The restaurant's clearest path to increasing tips for its staff is to focus on increasing the average bill size.

- **Actionable Steps:** Servers can be trained on effective upselling techniques, like recommending appetizers or premium dishes. The restaurant could also create special "set menus" or combos that encourage customers to spend a bit more. These efforts to raise the average check will directly boost server income.

Insight 2: People Tip by Habit, Not by Bill Size

Interestingly, while the dollar value of tips goes up with the bill, the tip *percentage* mostly stays the same. The correlation between the bill size and the tip percentage was very weak (0.14). This tells us that people don't really change their tipping rate, even when their bill is high. Most customers (81.5%) are "medium tippers," consistently leaving between 12% and 20%.

- **Recommendation:** It's more effective to focus on great service to protect this stable tipping culture than to try and convince customers to tip a higher percentage.
- **Actionable Steps:** Since most people seem to have a go-to tipping percentage (around 16–17%), the goal should be to provide a consistently positive experience that meets their expectations. This helps avoid low tips and encourages customers to come back, which is better for business in the long run than pushing for an extra percent on the tip.

Insight 3: Credit Card Users are More Generous Tippers

A really interesting find was that customers who pay with a credit card tip significantly more on average ($ 5.10) than those who pay with cash ($ 3.25). The ANOVA test confirmed this is a statistically significant difference. Even more, we saw that no credit card users fell into the "low tipper" category (<12%).

- **Recommendation:** The restaurant should make it as easy as possible for customers to pay by credit card.
- **Actionable Steps:** Having modern, easy-to-use payment terminals is a great start. These devices should also be set up to suggest common tip percentages (like 15%, 18%, 20%). This simple "nudge" makes tipping easier for the customer and seems to encourage more generous, standard amounts.

Insight 4: Large Groups Don't Mean a Better Tip Percentage

It's no surprise that big groups lead to bigger bills and therefore larger tips in total dollars. However, our analysis showed that these groups don't necessarily leave a better tip *percentage*. The rate tends to stay the same or even dip slightly for larger parties.

- **Recommendation:** The restaurant should implement an automatic gratuity policy for large groups.
- **Actionable Steps:** The restaurant can define what it considers a "large group" (for example, 6 or more people) and add a standard service charge of around 18%. This policy should be clearly stated on the menu and website so customers know what to expect. This ensures servers are fairly paid for the extra work that comes with serving a big table.

Insight 5: Service and Tipping are Consistent

Finally, our statistical tests found no significant difference in tip percentages across different days of the week or between the three servers. This points to a very consistent level of service and customer experience.

- **Recommendation:** Management should view this consistency as a major strength and focus on team-wide improvements.
- **Actionable Steps:** Since there are no underperforming days or servers, the restaurant can be confident in its current service standard. Any new training should be aimed at the entire team—like improving menu knowledge or introducing new service standards—rather than singling out individuals. This data shows that the restaurant is operating smoothly and reliably, which is a great foundation to build on.

VI. Limitations

Although this dataset provides useful insights into tipping behavior, there are certain limitations that should be taken into account when interpreting the results:

- **Timeframe of Data** - The dataset only covers a limited period, which means the findings may not fully represent long-term customer behavior. Tipping patterns could change over different seasons, holidays, or years, and this dataset does not capture those variations.
- **Number of Servers** - The data includes only three servers (A, B, and C). While this allows for comparison within the group, it is not enough to generalize about servers in larger restaurants or across different establishments. More servers would provide a stronger basis for evaluating performance differences.
- **Lack of Customer Demographics** - Important information such as customer age, gender, income level, or cultural background is not available. These factors often influence tipping behavior, and without them, the analysis cannot fully explain why some customers tip more generously than others.
- **Missing External Context** - Factors like holidays, special promotions, weather conditions, or restaurant events are not included in the dataset. These external elements can significantly impact customer spending and tipping but cannot be analyzed here.

Because of these limitations, the findings should be seen as indicative rather than absolute. The results are still valuable for understanding general patterns, but they should ideally be supported with additional data for more complete conclusions.

VII. Conclusion

In conclusion, this report analyzed the key factors that drive tipping at this restaurant. We found that the main driver for the *size* of a tip is the total bill—the two have a strong, positive correlation. However, the tip *percentage* is very consistent, hovering around a 16-17% norm, which suggests customer habit plays a bigger role than the bill size in determining the rate.

Our analysis also revealed that payment method matters, as credit card users tip significantly more than cash users. We also noted that while large groups are great for revenue, their tip percentage doesn't scale up with their bill, which can put servers at a disadvantage. On a positive note, the consistency of tipping across all days and servers suggests the restaurant provides a reliable service standard.

From these findings, we recommend a few clear actions: focus on strategies to increase the average check size, encourage credit card use with built-in tip suggestions, and establish an automatic gratuity for large parties.

It's important to acknowledge the study's limitations, like the specific timeframe of the data and lack of customer details. Even so, this analysis provides a solid, data-backed foundation for the restaurant to make strategic decisions that can boost revenue and support its staff.