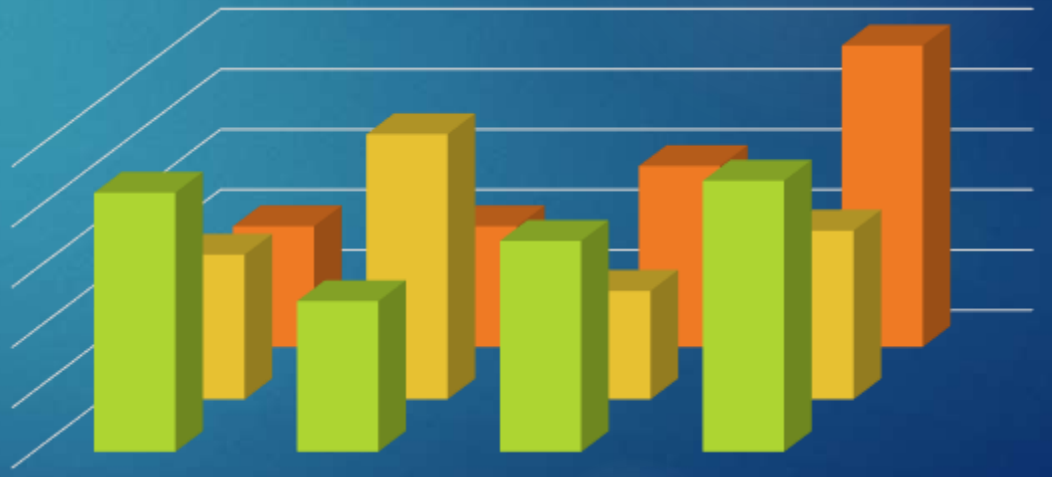
Customer Insights from Cuisine Ratings Dataset

LEVERAGING EDA FOR IMPROVED RESTAURANT PERFORMANCE



Introduction

- **Objective**: Understand customer preferences, behavior, and ratings using data analysis.

- **Dataset Overview**:
  - Contains customer demographics, activity patterns, budget, and ratings.
  - Categorical and numerical features analyzed.

GUVI | (3) Post | DMART-P | Machine | ChatGPT | EDA.ipynb | Free Goo | Downloa | vertopal.c

colab.research.google.com/drive/14hEGiA7TNqLtCucWens5ZP8o8p2gTM-7#scrollTo=QEGrvxeLI_66

EDA.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Share

+ Code + Text

RAM Disk

Gemini

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

0s

```
[6] import pandas as pd
from google.colab import files
upload=files.upload()
```

22s

Choose Files Cuisine_rating.csv

- **Cuisine_rating.csv**(text/csv) - 17828 bytes, last modified: 12/26/2024 - 100% done

```
Saving Cuisine_rating.csv to Cuisine_rating.csv
```

```
[7] df=pd.read_csv('Cuisine_rating.csv')
df
```

0s

	User ID	Area code	Location	Gender	YOB	Marital Status	Activity	Budget	Cuisines	Alcohol	Smoker	Food Rating	Service Rating	Overall Rating	Often A S
0	1	153	Upper East Side,NY	Female	2006	Single	Professional	3	Japanese	Never	Never	5	4	4.5	No
1	2	123	St. George,NY	Female	1991	Married	Student	3	Indian	Never	Socially	1	1	1.0	No

✓ 17s completed at 3:28 PM

16:10

```
df.describe()
```

	User ID	Area code	YOB	Budget	Food Rating	Service Rating	Overall Rating
count	200.000000	200.000000	200.000000	200.000000	200.000000	200.000000	200.000000
mean	100.500000	141.060000	1984.830000	3.815000	3.220000	3.230000	3.225000
std	57.879185	26.130257	16.809339	1.056578	1.411226	1.526022	1.079445
min	1.000000	101.000000	1955.000000	1.000000	1.000000	1.000000	1.000000
25%	50.750000	123.000000	1971.000000	3.000000	2.000000	2.000000	2.500000
50%	100.500000	135.000000	1987.000000	4.000000	3.000000	3.000000	3.000000
75%	150.250000	158.000000	2000.000000	5.000000	5.000000	5.000000	4.000000
max	200.000000	199.000000	2009.000000	5.000000	5.000000	5.000000	5.000000

```
[10] df.shape
```

```
(200, 15)
```

```
[11] df.head()
```

```
df.isnull().sum()
```

	0
User ID	0
Area code	0
Location	0
Gender	0
YOB	0
Marital Status	0
Activity	0
Budget	0
Cuisines	0
Alcohol	0
Smoker	0
Food Rating	0
Service Rating	0

[12]

Service Rating	0
Overall Rating	0
Often A S	0

dtype: int64

```
df.columns.values
```

```
array(['User ID', 'Area code', 'Location', 'Gender', 'YOB',
       'Marital Status', 'Activity', 'Budget', 'Cuisines', 'Alcohol ',
       'Smoker', 'Food Rating', 'Service Rating', 'Overall Rating',
       'Often A S'], dtype=object)
```

[14]

```
#catorical column - Location,Gender,Marital Status,Activity,Cuisines,Alcohol,Smoker,Often A S
#Numerical column - User ID,Area code,YOB,Budget,Food Rating,Service Rating,Overall Rating
```

[15]

```
fig = plt.figure(figsize=(6,6))
df['Gender'].value_counts().plot.pie(autopct='%1.2f%%')
```

```
<Axes: ylabel='count'>
```

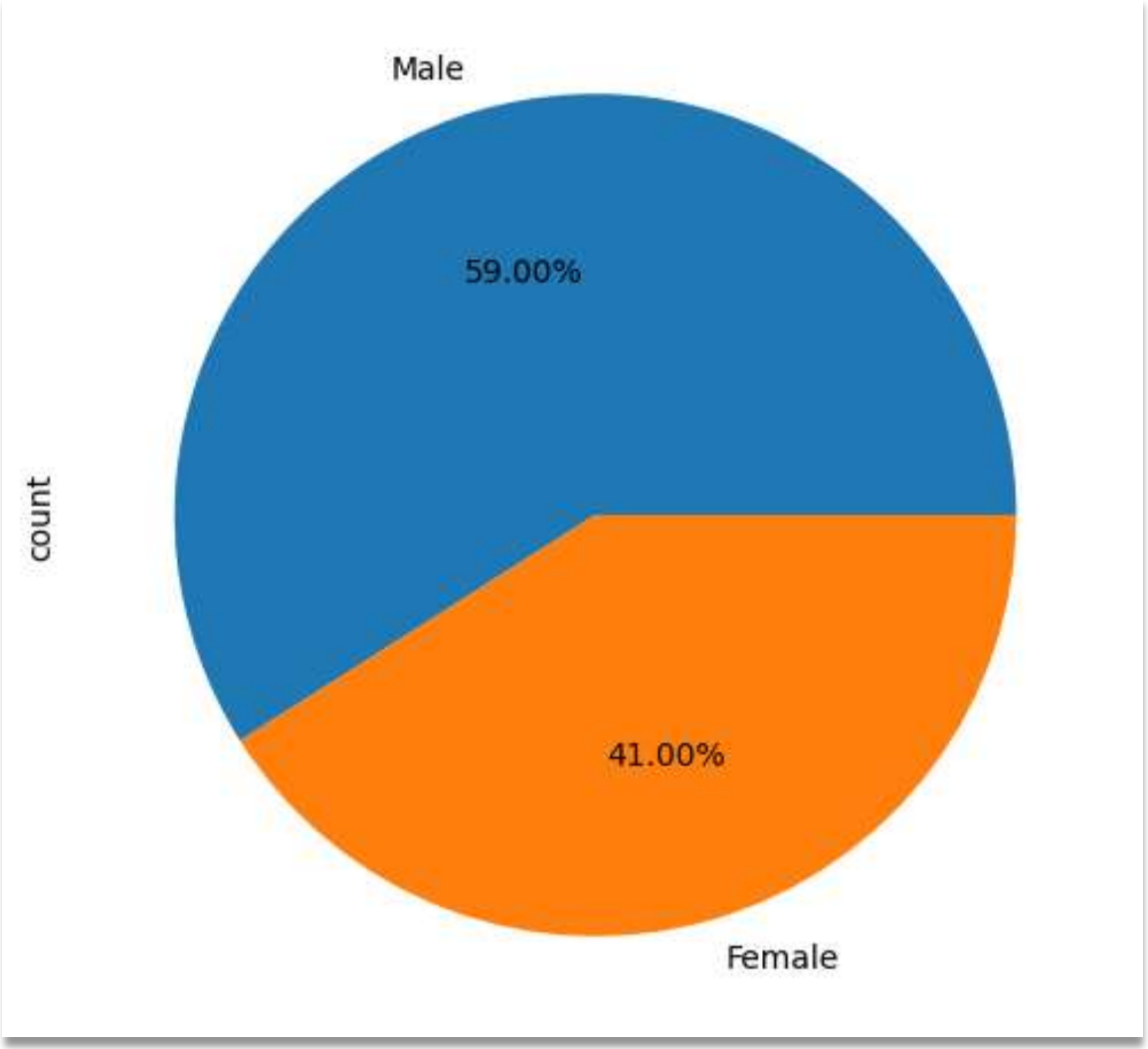
Visualization: Pie chart showing Gender split

Male Majority:

•Males represent 59% of the total dataset, indicating a higher proportion of male customers compared to females.

•***Female Customers:***

•Females account for 41% of the dataset, which is significant but less compared to males.



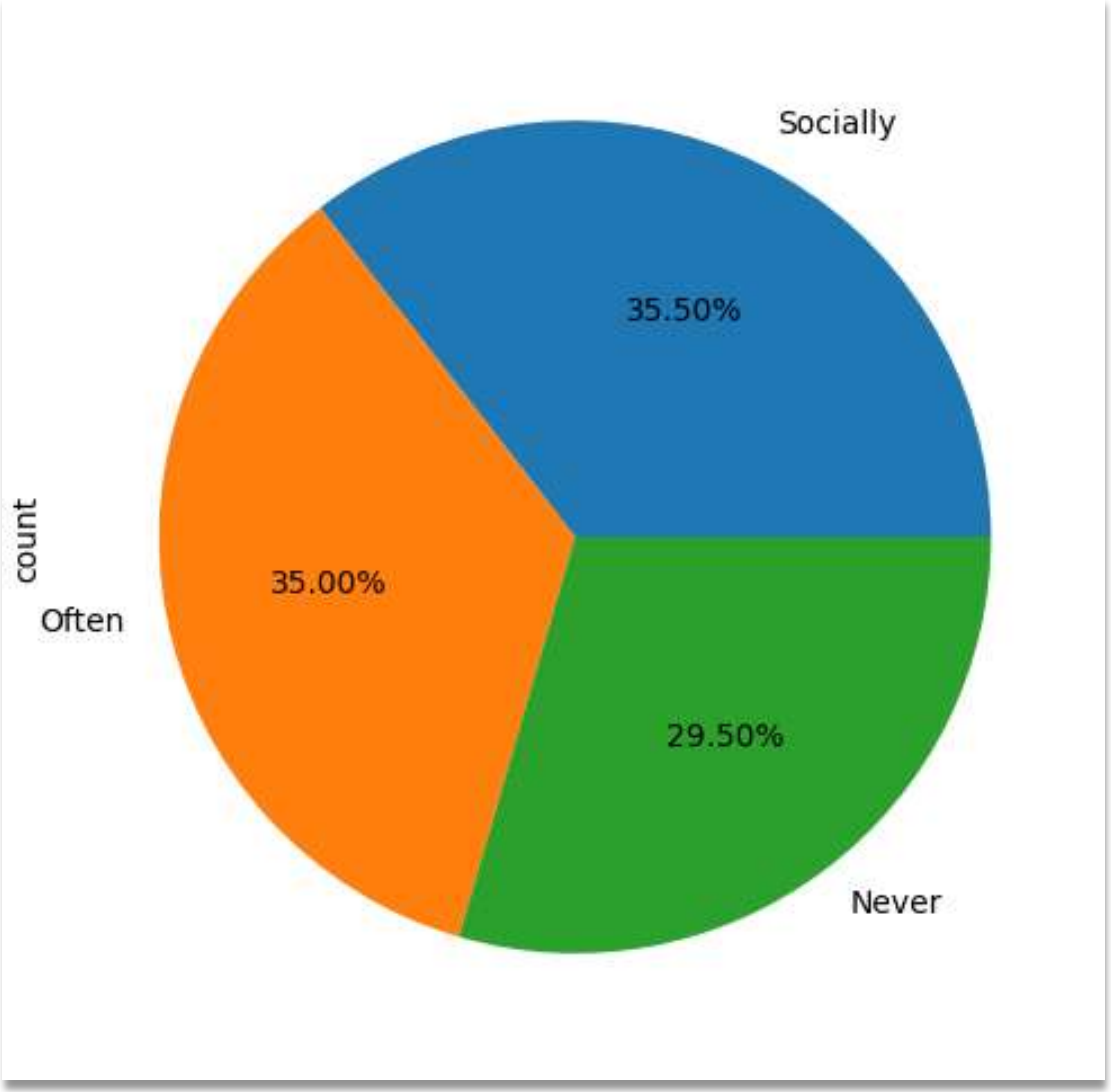
Visualization: Pie chart showing Smoker vs. Non-Smoker distribution.

Key Insight:

Balanced Distribution:
Smoking behavior among customers is fairly evenly distributed across three categories:

- **Socially**: 35.5%
- **Often**: 35.0%
- **Never**: 29.5%

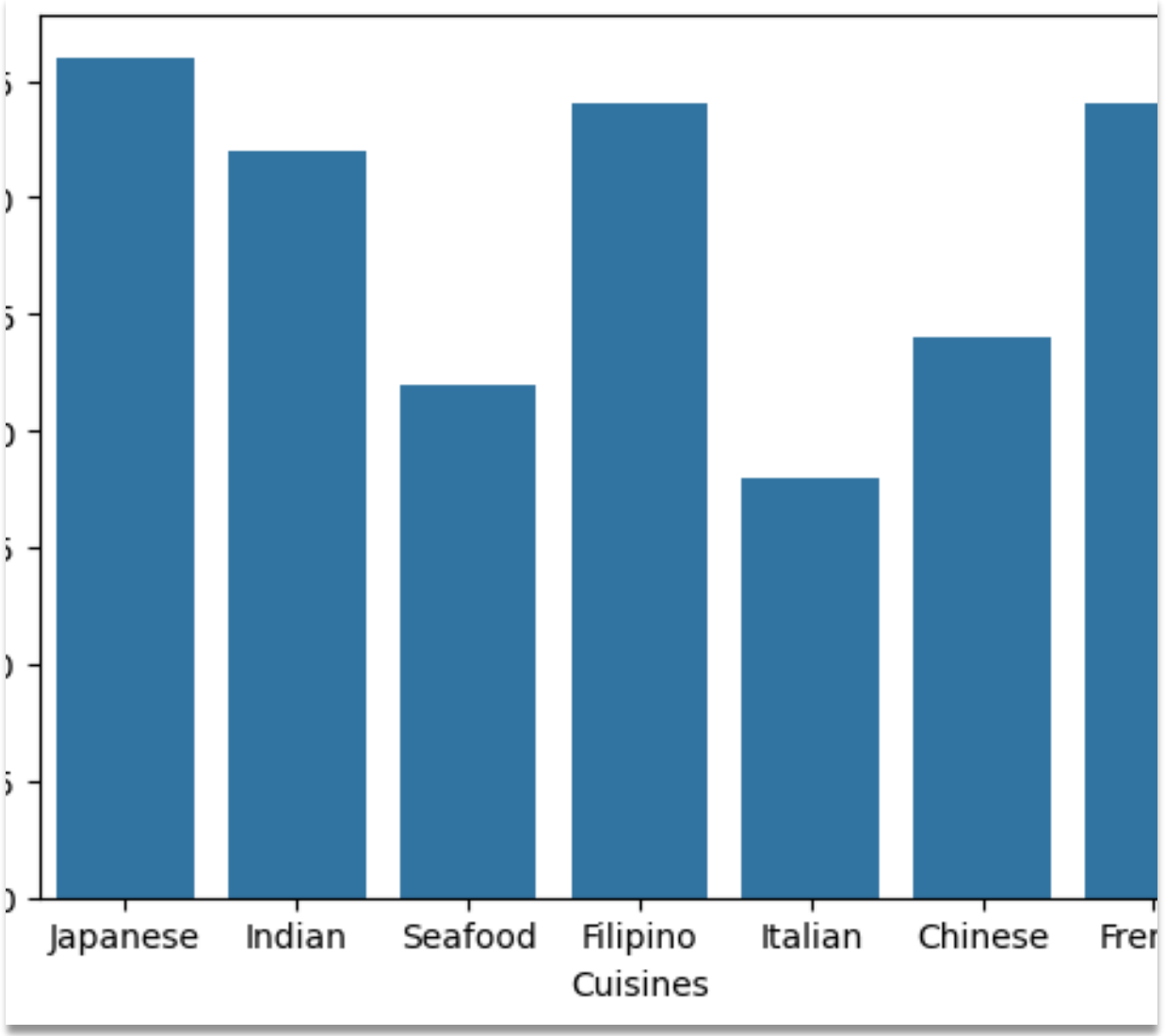
Implications for Business:
A significant portion of customers (70.5%) engage in smoking either socially or often, suggesting the importance of offering designated smoking areas or accommodating smokers in some way.



Key Insight: Top preferred cuisines.

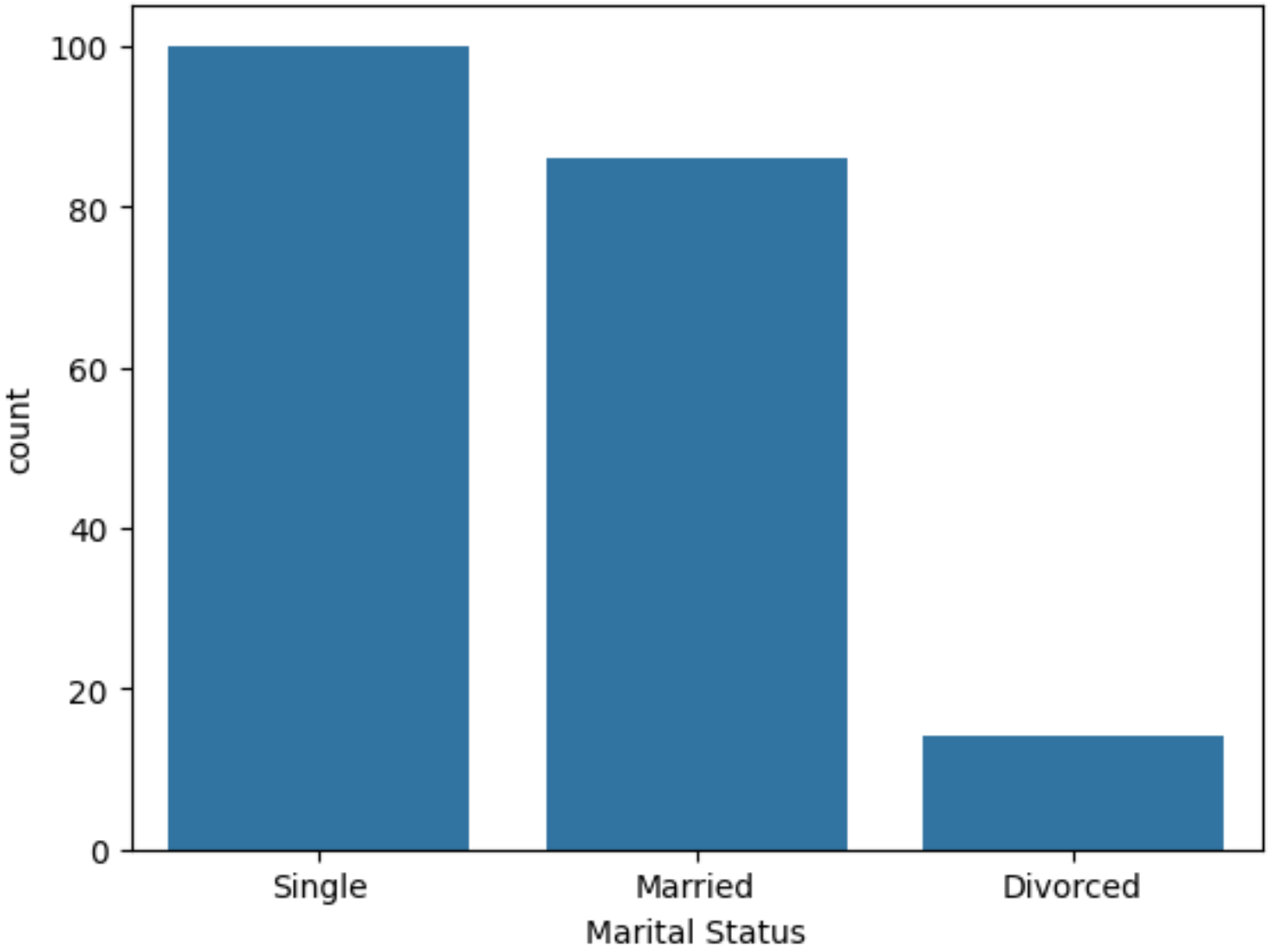
Japanese Cuisine is the most popular: It has the highest number of occurrences.

- **Indian Cuisine is second most popular.**
- **Seafood, Filipino, and Chinese cuisines have a moderate number of occurrences.**
- **Italian and French cuisines have fewer occurrences compared to the others.**



key insights from the provided bar chart:

- **Single individuals are the most represented group.** They have the highest count among the three marital statuses.
- **Married individuals come in second.** The count for married individuals is lower than that of single individuals.
- **Divorced individuals have the lowest representation.** They have the smallest count among the three groups.



Specific Observations:

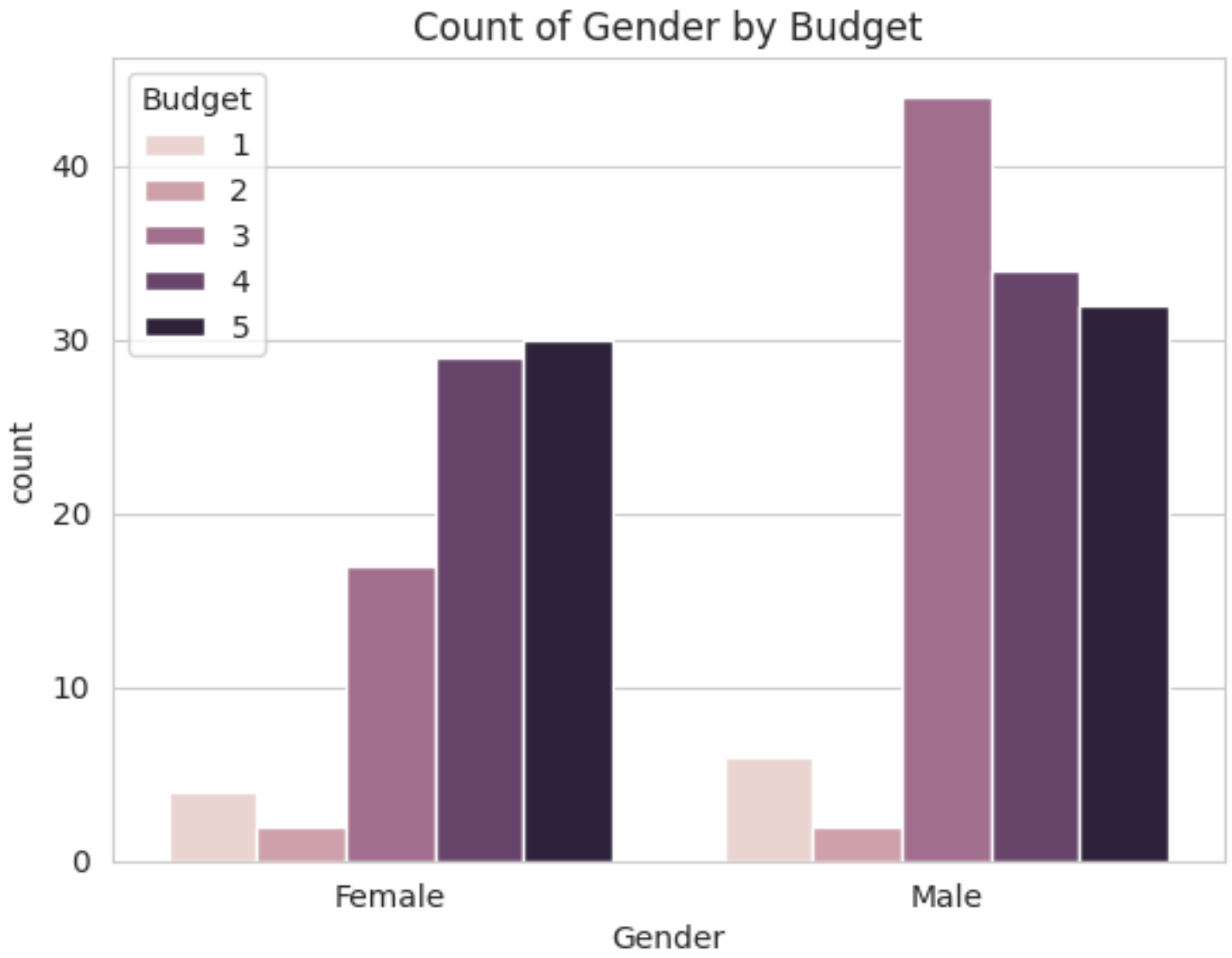
For Budget Level 1: Males have a slightly higher count than females.

For Budget Levels 2 and 3: Males have a significantly higher count than females.

For Budget Level 4: Males have a slightly higher count than females.

For Budget Level 5: Males have a significantly higher count than females.

In summary: The chart suggests that males tend to have higher budgets and are more likely to fall into higher budget levels compared to females.



Specific Observations

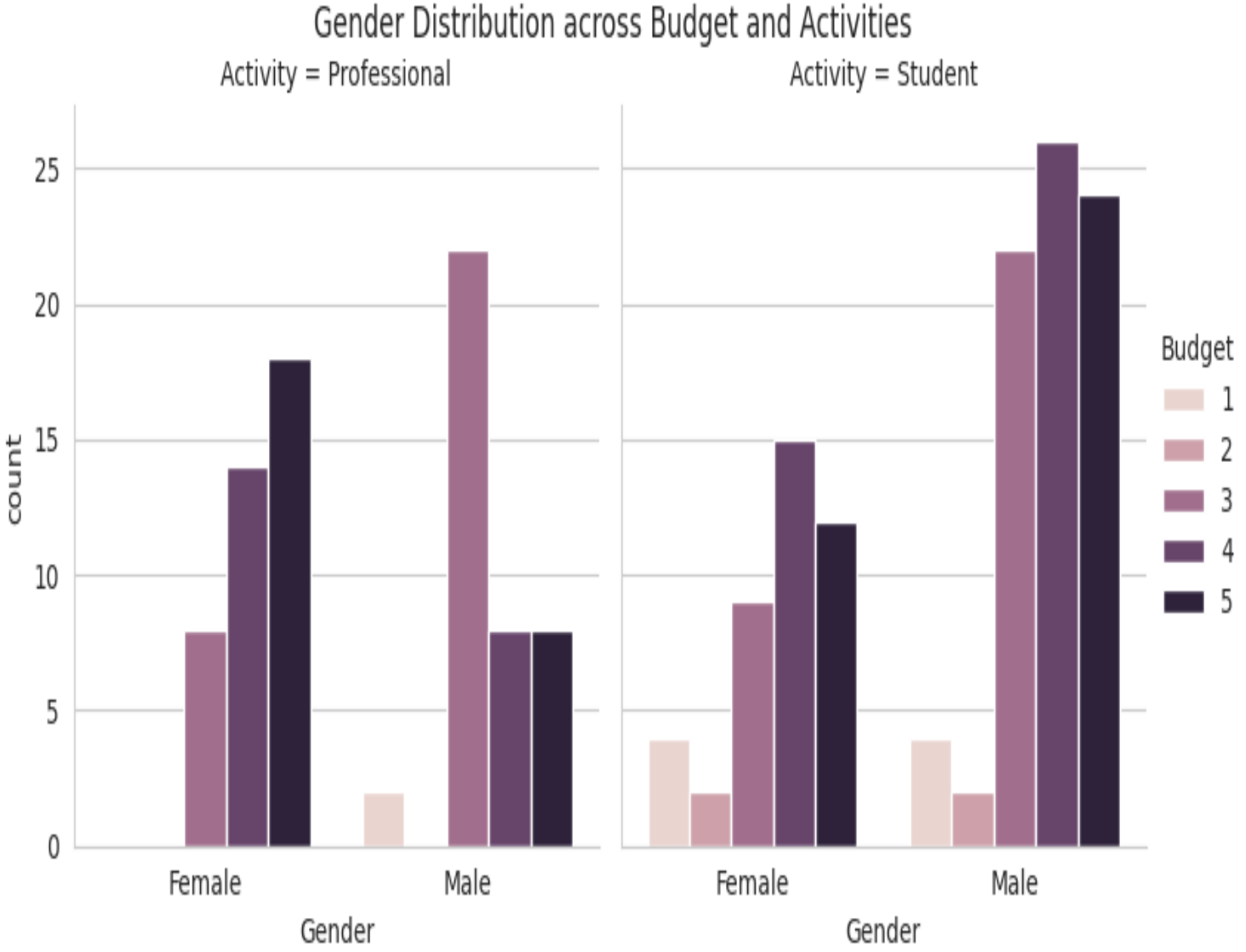
•For Professionals:

- Males have a higher count in all budget levels except for Budget 1.
- Budget Level 4 has the highest count for both genders.

•For Students:

- Males have a higher count in all budget levels except for Budget 1.
- Budget Level 5 has the highest count for both genders.

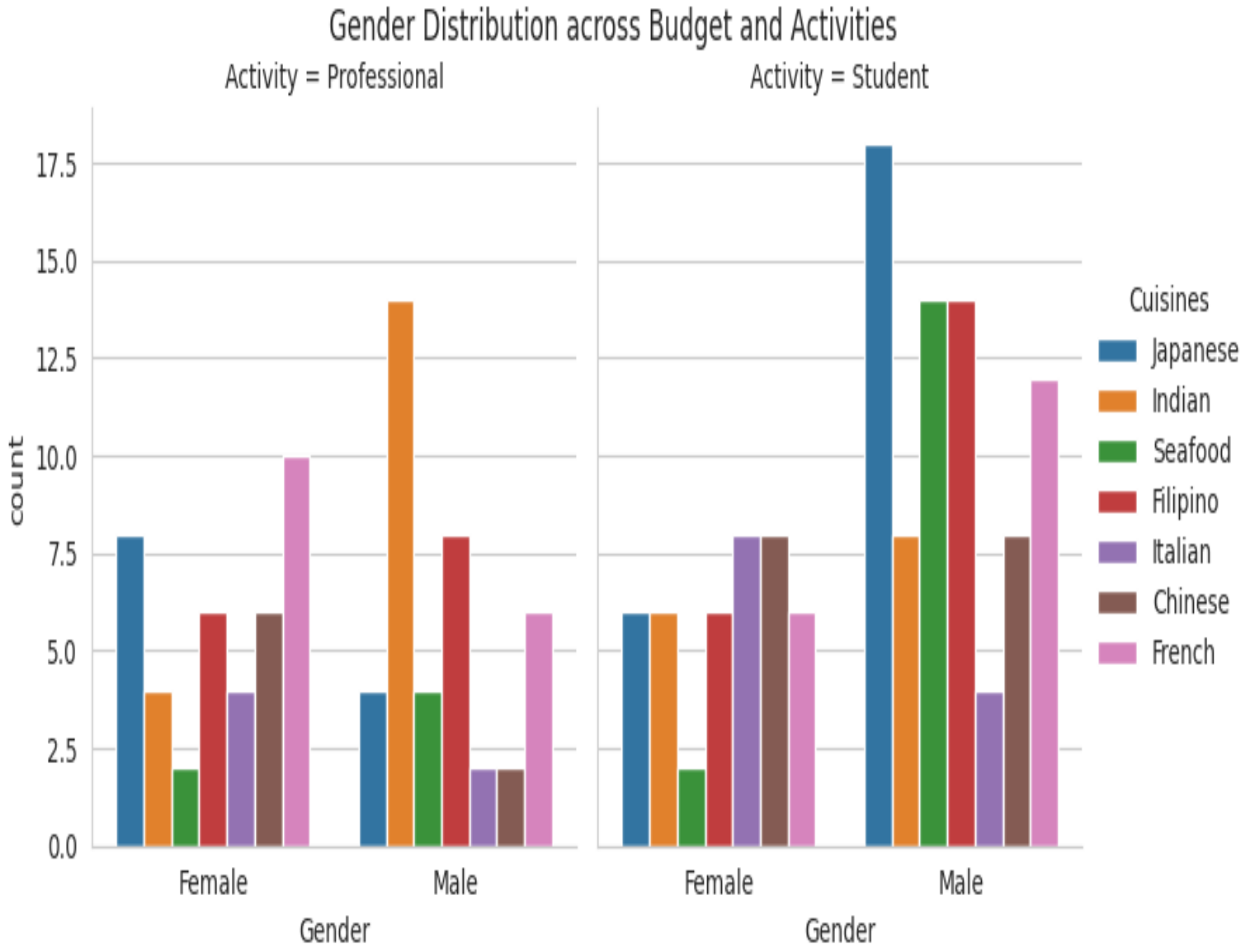
In summary: The chart reveals that males generally have higher budgets across both professional and student activities, with Budget Level 5 being the most common for both groups.



Specific Observations:

- **Professional Activity:** Japanese cuisine is significantly more popular than other cuisines.
- **Student Activity:** While Japanese cuisine still leads, the distribution of preferences for other cuisines is more even.

In summary: The chart reveals that Japanese cuisine is consistently favored across both professional and student activities, while the popularity of other cuisines varies slightly between the two groups.



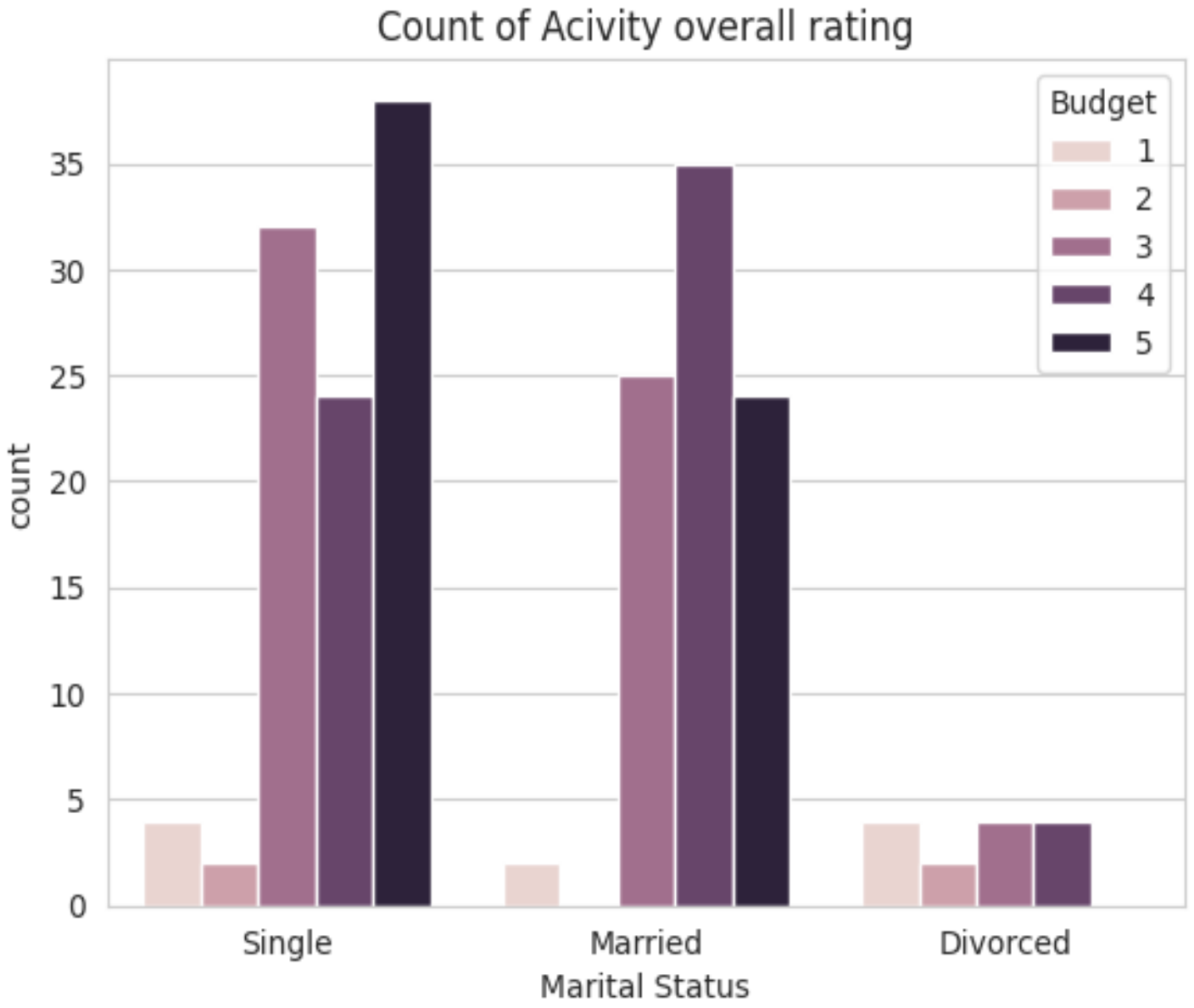
Specific Observations:

Single individuals tend to have higher budgets. This is evident from the higher bars for Single individuals across all budget levels.

Married individuals also have a significant presence in higher budget levels.

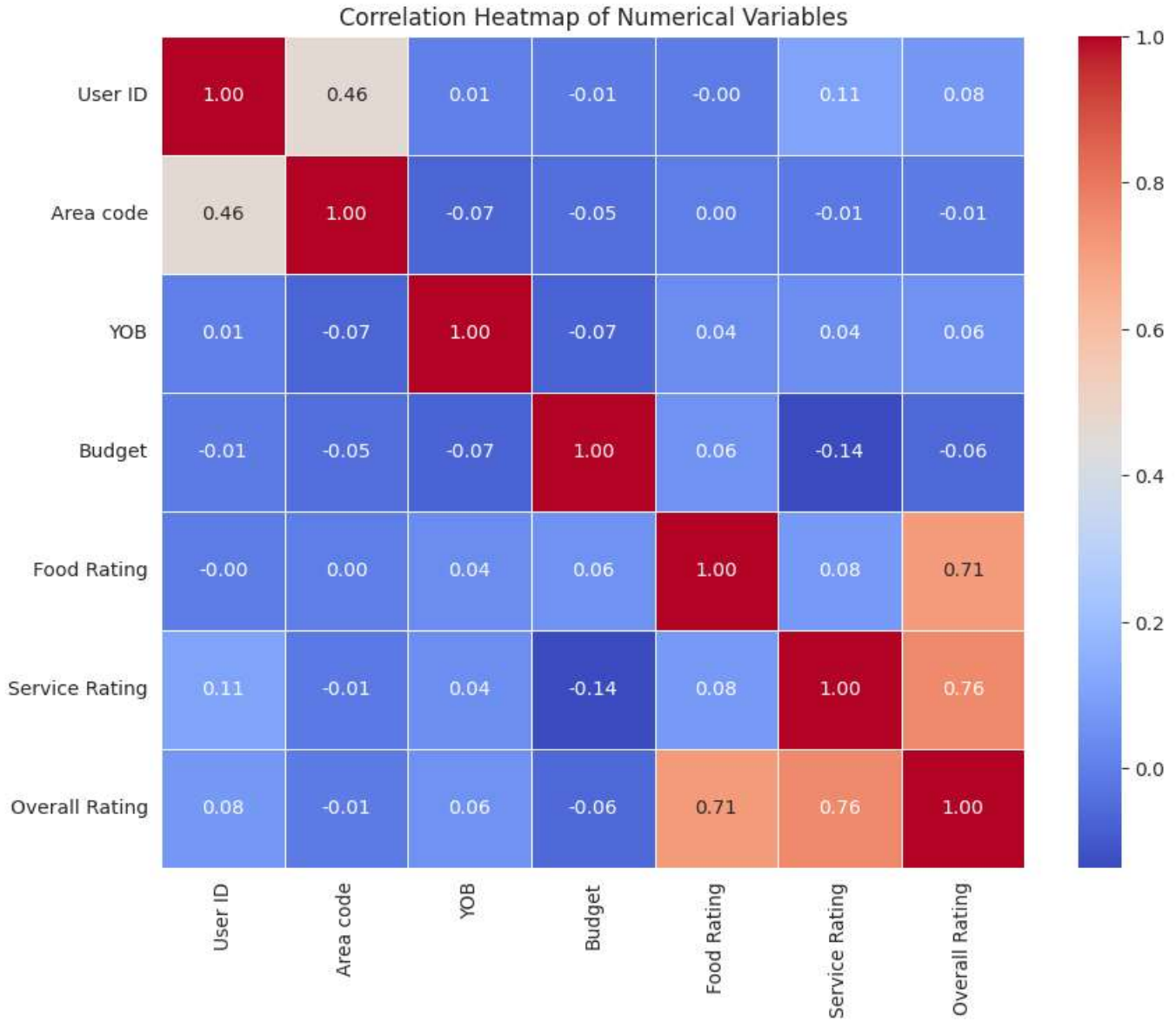
Divorced individuals are more concentrated in lower budget levels.

In summary: The chart suggests that Single individuals are more likely to fall into higher budget levels compared to Married and Divorced individuals



Heatmap

The correlation heatmap reveals that food and service quality strongly influence overall customer satisfaction. Customers tend to rate food and service similarly. Conversely, factors like age (YOB) and budget have little impact on ratings. To maintain customer satisfaction, businesses should prioritize high standards in food and service. Additionally, exploring customer segmentation and investigating other potential factors affecting satisfaction, such as dining experience and ambiance, could provide further insights.



Heatmap of Gender, Budget, and Activity

Target specific segments: The heatmap suggests that tailoring marketing or promotional strategies to different gender and activity groups could be effective. For example, targeting males with higher budgets for professional activities or females with higher budgets for student activities.

Investigate budget drivers: Understanding the factors that influence budget decisions for different groups could inform pricing strategies and product offerings.

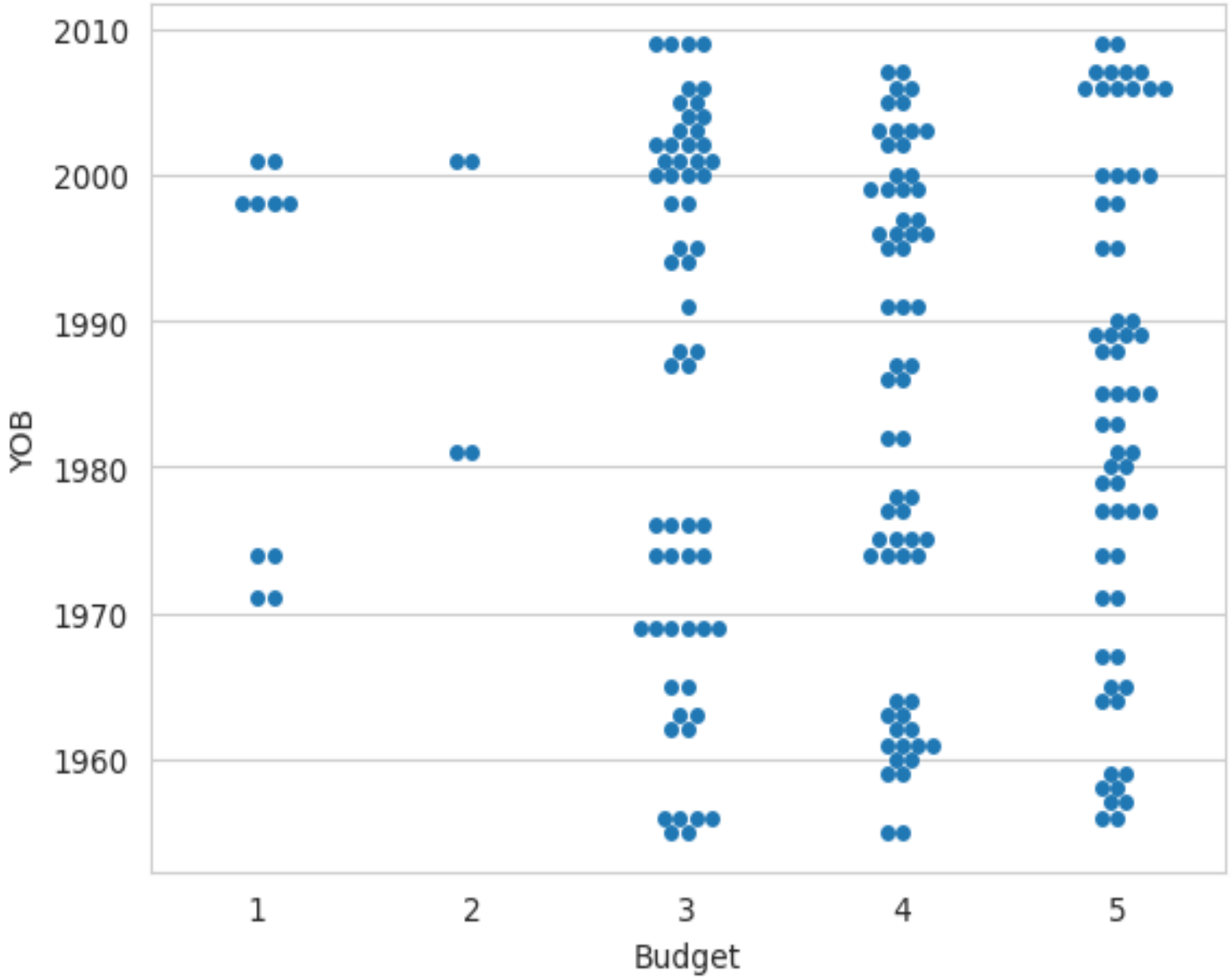
Consider dynamic pricing: Implementing dynamic pricing based on activity and gender could optimize revenue and cater to different budget sensitivities.

Swarmplots for Customer Demographics

Budget Distribution: The distribution of budgets seems to be skewed towards higher values. There are a larger number of observations with higher budget values compared to lower ones.

Year of Birth (YOB) Distribution: The YOB data points are more evenly distributed across the range. There are no apparent clusters or gaps in the data.

Relationship between Budget and YOB: There doesn't appear to be a clear linear relationship between budget and YOB. The data points are scattered without a discernible pattern.



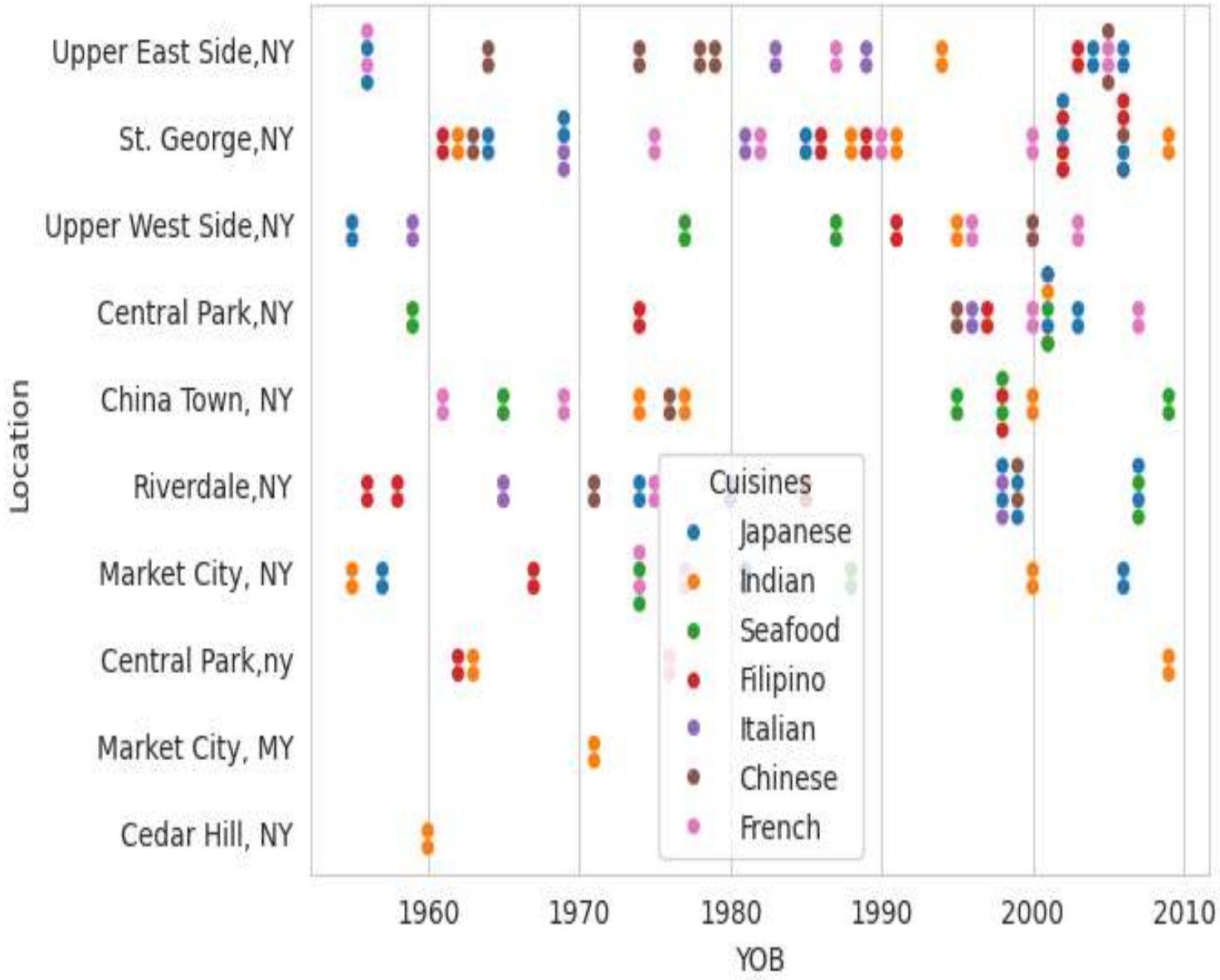
Cuisine Distribution: The plot shows the distribution of various cuisines (Japanese, Indian, Seafood, etc.) across different locations (Upper East Side, St. George, etc.) and years (1960-2010).

Location Preferences:

- **Upper East Side** appears to have a wide variety of cuisines represented.
- **St. George** and **Upper West Side** have a more limited range of cuisines.
- **Cedar Hill** has only a few cuisines represented.

Temporal Trends:

- It's difficult to discern clear temporal trends due to the scatter in the data points. However, some locations might show a preference for certain cuisines in specific time periods.



Key Insight: Highlight cuisine preferences across different budgets and locations.

Cuisine Distribution: The plot shows the distribution of various cuisines (Japanese, Indian, Seafood, etc.) across different locations (Upper East Side, St. George, etc.) and budgets.

Location Preferences:

•**Upper East Side** appears to have a wide variety of cuisines represented.

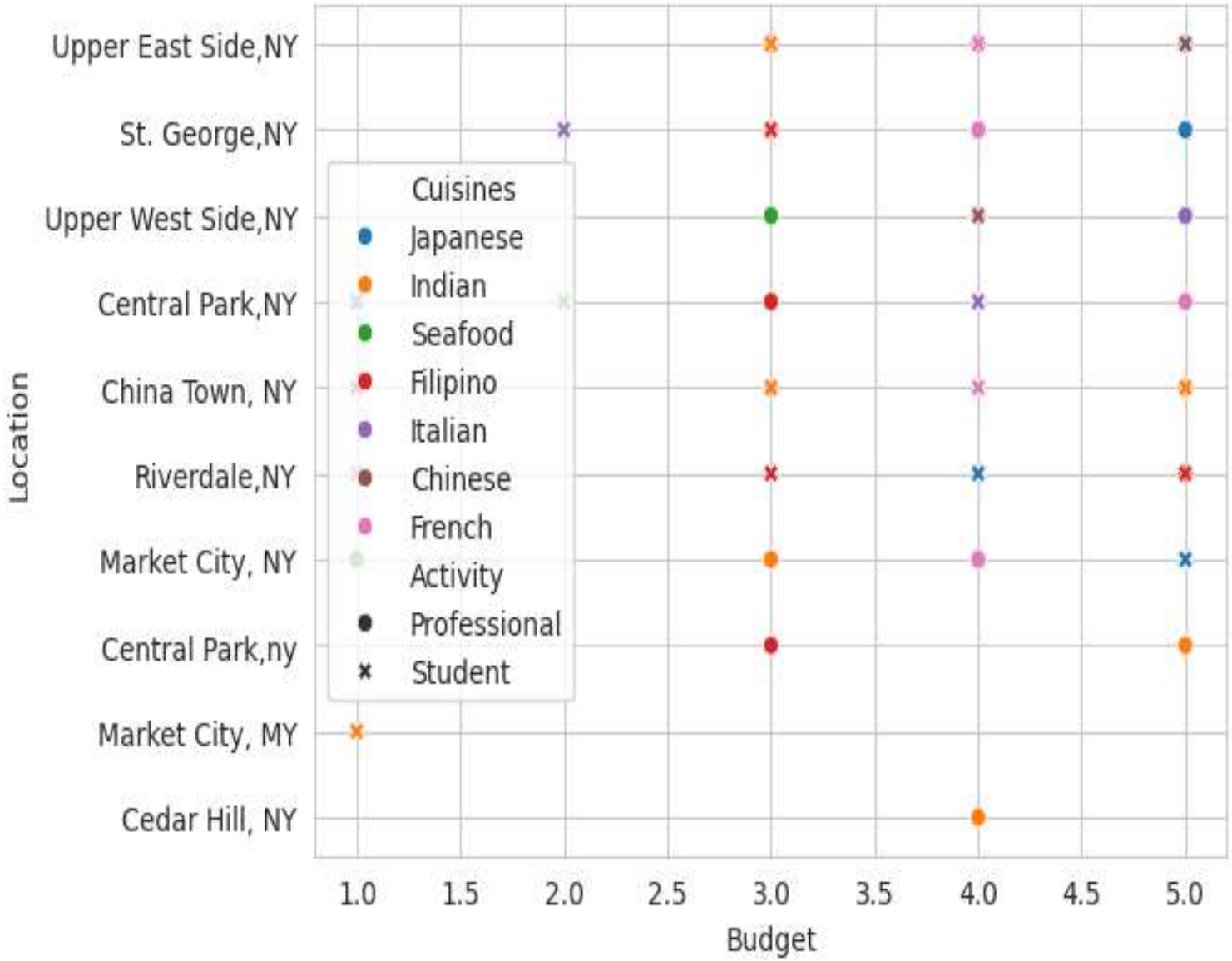
•**St. George** and **Upper West Side** have a more limited range of cuisines.

•**Cedar Hill** has only a few cuisines represented.

Budget Variation:

•There is a wide range of budgets observed across the different cuisines and locations.

•Some locations (like Upper East Side) seem to have cuisines with higher budgets compared to others.



Recommendations for the Client:

1. **Smoking & Alcohol Preferences**
    1. Use insights from smoking and alcohol habits to optimize seating arrangements, bar services, or designated areas.
    2. Promote "non-smoking" areas for families and "social drinking" areas for groups.
2. **Gender-Based Marketing Strategies**
    1. Design specific campaigns targeting male and female customers differently, as their preferences may vary.
    2. Example: Female customers might appreciate health-focused menu options or family dining promotions.
3. **Tailor Offers by Age Group**
    1. Use the "Year of Birth" (YOB) data to identify generational preferences.
    2. For younger customers, focus on trendy cuisines or casual dining.
    3. For older customers, offer more traditional or premium options.
4. **Leverage Seasonal and Regional Insights**
    1. Identify if certain locations or cuisines perform better during specific times of the year and tailor campaigns accordingly.
    2. Promote regional specialties in locations with lower performance to boost interest.
5. **Optimize Marketing for Low-Performing Locations**

- Analyze why certain locations have fewer customers and experiment with localized promotions, discounts, or community events to draw attention.

Recommendations for the Client:

•**Target Popular Locations**
Focus marketing efforts on the top dining locations identified in the dataset. These areas have the highest customer concentration, which could increase footfall and revenue.
•**Promote Top Cuisines**
Highlight the most preferred cuisines (e.g., Italian, Chinese) in promotions, especially in popular locations.
Introduce special offers, combo deals, or loyalty programs to boost customer engagement for these cuisines.
•**Enhance Customer Experience**
Improve food and service quality, as high ratings in these areas correlate with overall customer satisfaction.
Regularly gather feedback to address potential areas of improvement.
Leverage Budget Insights
Offer tailored menus or packages to different budget groups.
For high-budget customers, consider premium dining options, exclusive experiences, or personalized services.
For lower-budget groups, provide affordable deals or discounts to attract more customers.