Level 1 Data Analysis Report

Cognifyz Technologies Internship Program

Date: October 2025

Intern Name: Shazia Afraz

Organization: Cognifyz Technologies Internship

Duration: 28 September 2025 – 28 October 2025

Program: Data Analysis Internship With Python

Introduction

The **Level 1 Data Analysis Report** was created as part of the Cognifyz Technologies internship program to explore and interpret restaurant dataset patterns through structured, data-driven analysis. The primary aim of this stage was to identify foundational insights that define the market characteristics, service distributions, and consumer preferences within the restaurant industry.

The dataset used contains information such as cuisines offered, city-wise restaurant counts, price ranges, online delivery services, ratings, and other operational parameters. This report focuses on **four analytical tasks**, each tied to a specific business objective.

Task 1 – Top Cuisines

Objective:

To identify the top three most common cuisines offered in restaurants, quantify their relative popularity, and provide actionable market insights for menu planning and culinary business strategy.

Code Explanation :

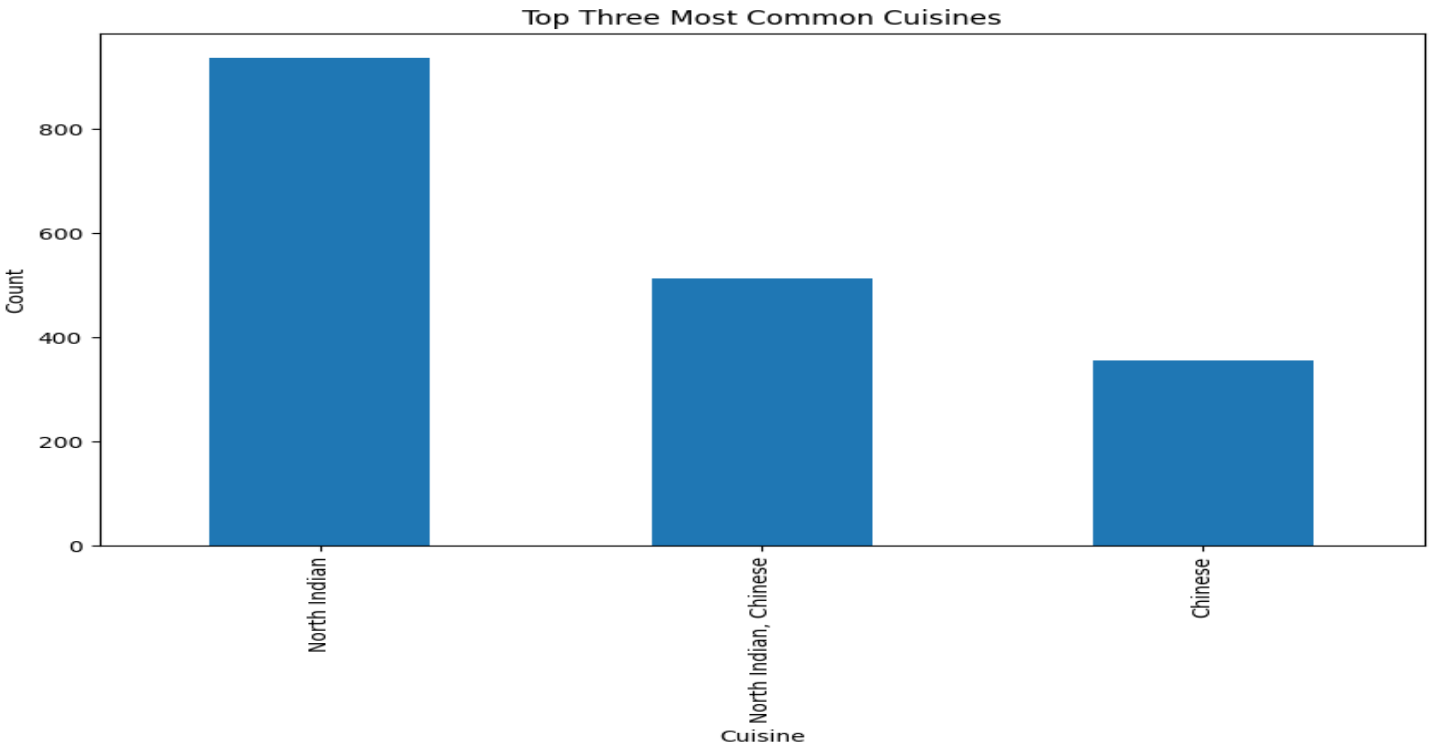
The code calculates the frequency of each cuisine by using the value_counts() method on the 'Cuisines' column. It then selects the top three cuisines using the head(3) function. The resulting counts are printed for review. For visualization, a bar chart is created using matplotlib.pyplot to display the counts of these top cuisines, illustrating their relative popularity.

```
# Get the top three most common cuisines task 1 level 1
cuisine_counts = df['Cuisines'].value_counts()
top_three = cuisine_counts.head(3)
# Plot the top three cuisines
plt.figure(figsize=(10, 6))
top_three.plot(kind='bar')
plt.title('Top Three Most Common Cuisines')
plt.xlabel('Cuisine')
plt.ylabel('Count')
plt.show()
```

✓ 0.0s

Results:

- North Indian cuisine is the most prevalent, with the highest count among restaurants.
- North Indian & Chinese fusion cuisine is the second most common, reflecting demand for diverse menu options.
- Chinese cuisine ranks third, maintaining strong presence across dining establishments.
- These cuisines significantly outnumber other types, signifying leading consumer preferences in the market.



Visualization:

The bar chart titled "Top Three Most Common Cuisines" displays the count of restaurants for each cuisine type:

- North Indian stands tallest, followed by North Indian, Chinese, and Chinese.
- The count axis shows the precise number of restaurants per cuisine, while the cuisine names are labelled along the x-axis for clear comparison.
- This visual format quickly highlights the popularity gaps and ranking among the top three.

Business Implications:

- **Menu design:** Restaurants can increase competitiveness by featuring North Indian and Chinese dishes, catering to dominant customer preferences.

- **Fusion strategy:** Including popular fusion cuisines (like North Indian, Chinese) can further broaden appeal and attract a larger customer base.
- **Market positioning:** New entrants should consider the prevalence of these cuisines for both differentiation and alignment with established demand trends.
- **Consumer targeting**: Marketing campaigns focused on these cuisine types are likely to reach the largest audience and maximize promotional impact.

Task 2 – City Analysis

Objective:

To analyze the distribution of restaurants across India's top 10 cities and identify urban clusters with the highest food service activity, providing actionable insights for restaurant businesses, investors, and planners based on city-wise restaurant density.

Explain Code :

This code analyzes the dataset to find which city has the most restaurants:

- The code calculates which city has the highest number of restaurants by counting the occurrence of each city in the 'City' column, finding the one with the maximum count, and then printing the result with the city name and its restaurant count.
- It then prepares a bar chart showing the top 10 cities with the most restaurants using the value counts from the 'City' column, displaying the cities on the x-axis and the restaurant counts on the y-axis.
- The plot is customized with a title, x-label, y-label, and a specific figure size for better visualization before displaying the bar chart.

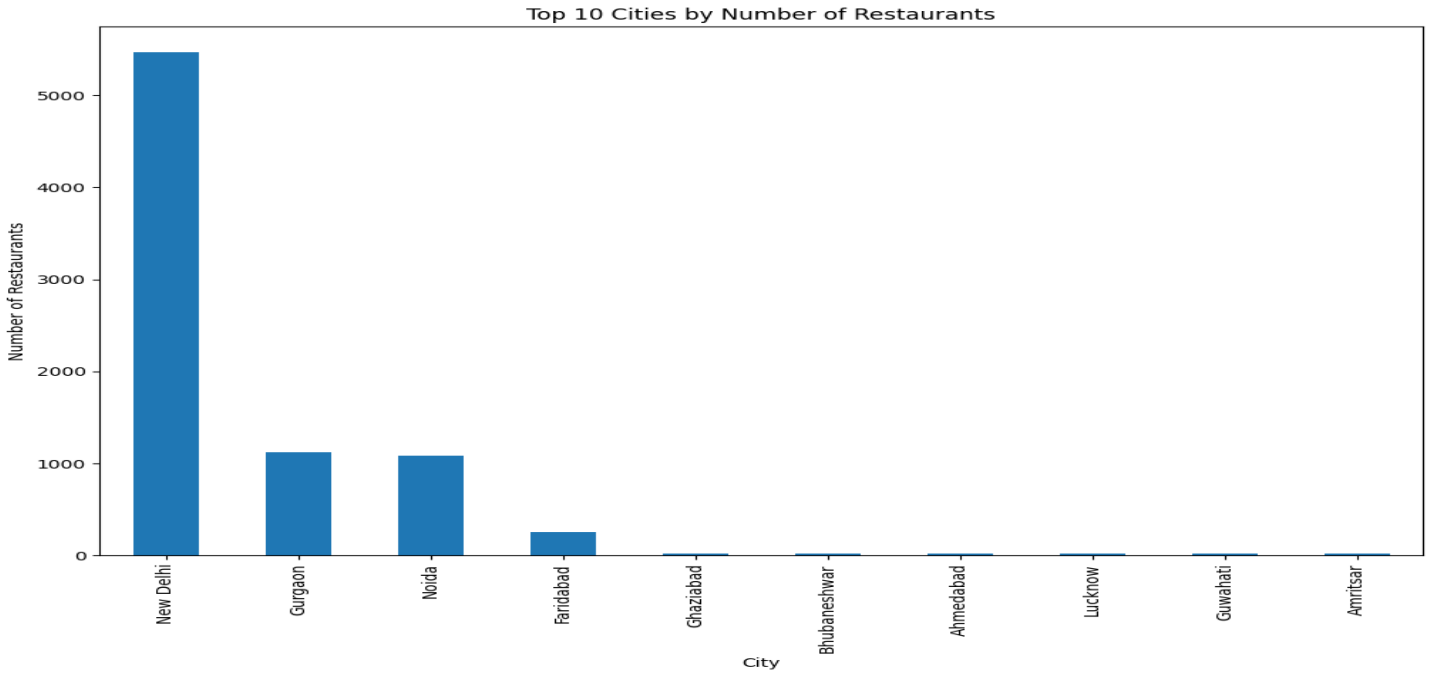
```
# Identify the city with the highest number of restaurants level 1 task 2
city_counts = df['City'].value_counts()
top_city = city_counts.idxmax()
top_count = city_counts.max()
print(f'The city with the highest number of restaurants is {top_city} with {top_count} restaurants.')
# Plot bar chart of top 10 cities
plt.figure(figsize=(12, 8))
city_counts.head(10).plot(kind='bar')
plt.title('Top 10 Cities by Number of Restaurants')
plt.xlabel('City')
plt.ylabel('Number of Restaurants')
plt.show()
```

✓ 0.1s

Results:

- **New Delhi** leads overwhelmingly in restaurant count, with more than 5,000 establishments—far surpassing other cities.
- **Gurgaon** and **Noida** form the secondary tier, each hosting over 1,000 restaurants, highlighting their status as significant food service hubs in the NCR region.

- All other cities (Faridabad, Ghaziabad, Bhubaneswar, Ahmedabad, Lucknow, Guwahati, Amritsar) trail significantly behind, underlining a strong geographic imbalance in restaurant presence and urban food culture.



Visualization:

The bar chart titled **"Top 10 Cities by Number of Restaurants"** visually demonstrates the sharp disparity in restaurant counts:

- New Delhi is the tallest bar, followed by smaller—yet notable—bars for Gurgaon and Noida.
- Remaining cities are represented by much shorter bars, reflecting their comparatively modest restaurant sectors.
- The x-axis lists the cities, while the y-axis indicates the number of restaurants in each, allowing for immediate comparison and ranking.

Business Implications:

- **Market Saturation & Opportunity**: New Delhi is a highly competitive and saturated market, requiring differentiation, quality, and strong marketing for new entrants and existing businesses to thrive.
- **Emerging Growth Hubs**: Gurgaon and Noida offer both scale and growth for expansion, benefiting from urbanization and spending power, but with less saturation than Delhi.
- **Untapped Potential**: Lesser-represented cities present opportunities for pioneering brands or franchise expansion, as lower supply may signal unmet consumer demand and less competition.

- **Data-driven Site Selection**: Businesses and investors can use city-wise data to inform their site selection, understand competitive landscapes, and optimize resource allocation for maximum ROI.

Task 3 – Price Range Distribution

Objective:

To assess the distribution of restaurants by price range and understand market segmentation based on affordability, revealing key trends in pricing strategy across the industry.

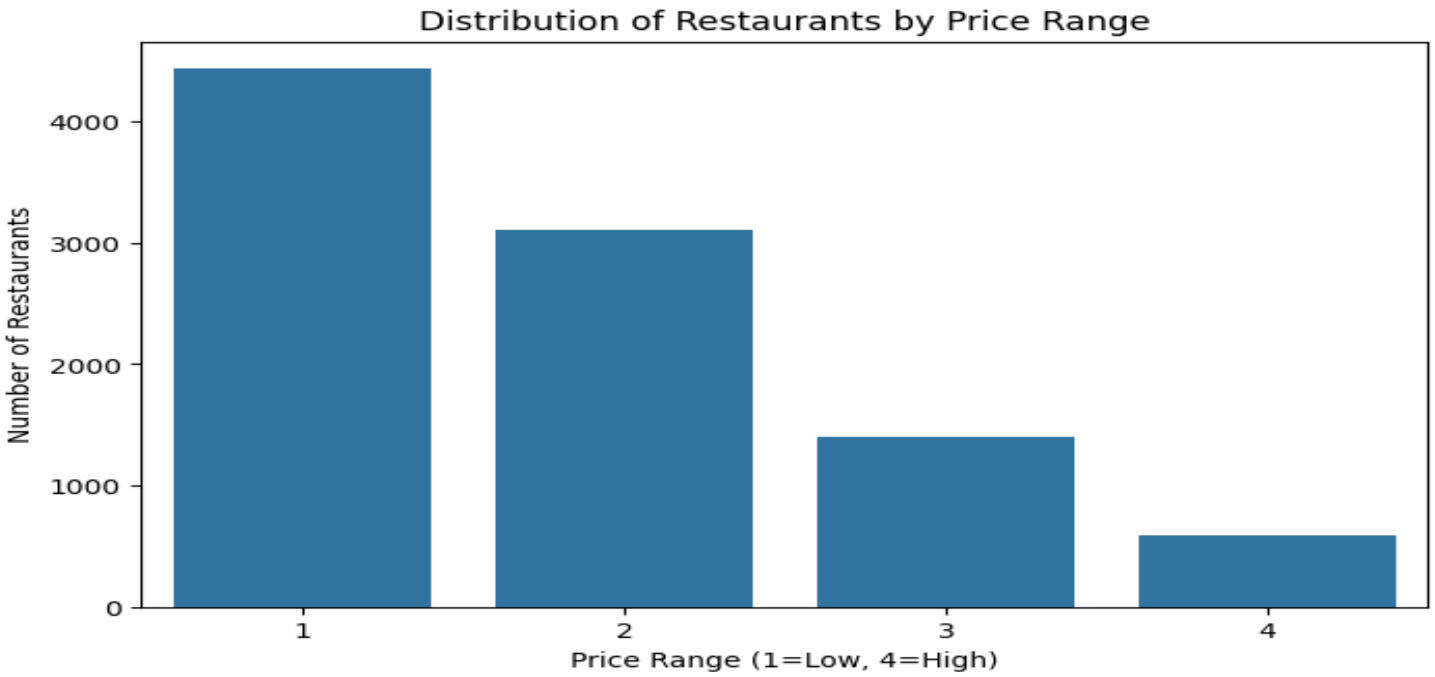
Explain Code

- The code calculates the count of restaurants in each price range from the 'Price range' column, sorts them, and then computes the percentage that each price range represents out of the total number of restaurants.
- It creates a new DataFrame containing the price ranges, their respective restaurant counts, and their percentages, then prints this summary data for review.
- Finally, it visualizes the distribution using a barplot with 'Price Range' on the x-axis and the count on the y-axis, and customizes the chart with labels and a title before displaying it.

```
# Count of restaurants in each price range level 1 task 3
price_counts = df['Price range'].value_counts().sort_index()

# Percentage calculation
price_percent = (price_counts / df.shape[0]) * 100

price_data = pd.DataFrame({
    'Price Range': price_counts.index,
    'Count': price_counts.values,
    'Percentage': price_percent.values
})

print(price_data)

# Visualization: Histogram
plt.figure(figsize=(8,5))
sns.barplot(x=price_data['Price Range'], y=price_data['Count'])
plt.title("Distribution of Restaurants by Price Range")
plt.xlabel("Price Range (1=Low, 4=High)")
plt.ylabel("Number of Restaurants")
plt.show()
```

Results:

- Price Range 1 (Lowest) contains the highest number of restaurants, indicating a strong presence of budget-friendly establishments.
- Price Range 2 is also well represented, making up the second largest group and reflecting broad appeal for moderately priced dining.
- Price Range 3 has a moderate number of restaurants, signaling a smaller but visible market for premium casual experiences.
- Price Range 4 (Highest) has the fewest restaurants, highlighting that high-end dining is a minority segment within the market.



Visualization:

The bar chart titled "**Distribution of Restaurants by Price Range**" shows the count of restaurants for each price tier:

- The bars steeply decline from left to right, with Price Range 1 towering above the rest and each subsequent range having fewer restaurants.
- The x-axis categorizes price levels from 1 (low) to 4 (high), while the y-axis quantifies the restaurant count.
- This distribution visually communicates that affordability is the dominant trend.

Business Implications:

- **Market Targeting:** The dominance of lower price ranges shows that most consumers prioritize value and affordability, making these segments ideal for high-volume, cost-competitive business models.
- **Premium Strategy:** The smaller share of higher-priced restaurants suggests exclusivity but also a need for strong differentiation, quality, and experience to justify premium pricing.
- **Expansion Planning:** Entrepreneurs and investors can prioritize mid- and low-range openings for broader reach, while high-end concepts should focus on niche, affluent urban markets.
- **Menu and Promotion:** Restaurants in lower price categories can benefit from frequent promotions and combo deals to attract price-sensitive diners, whereas premium locations can leverage exclusivity and personalized service.

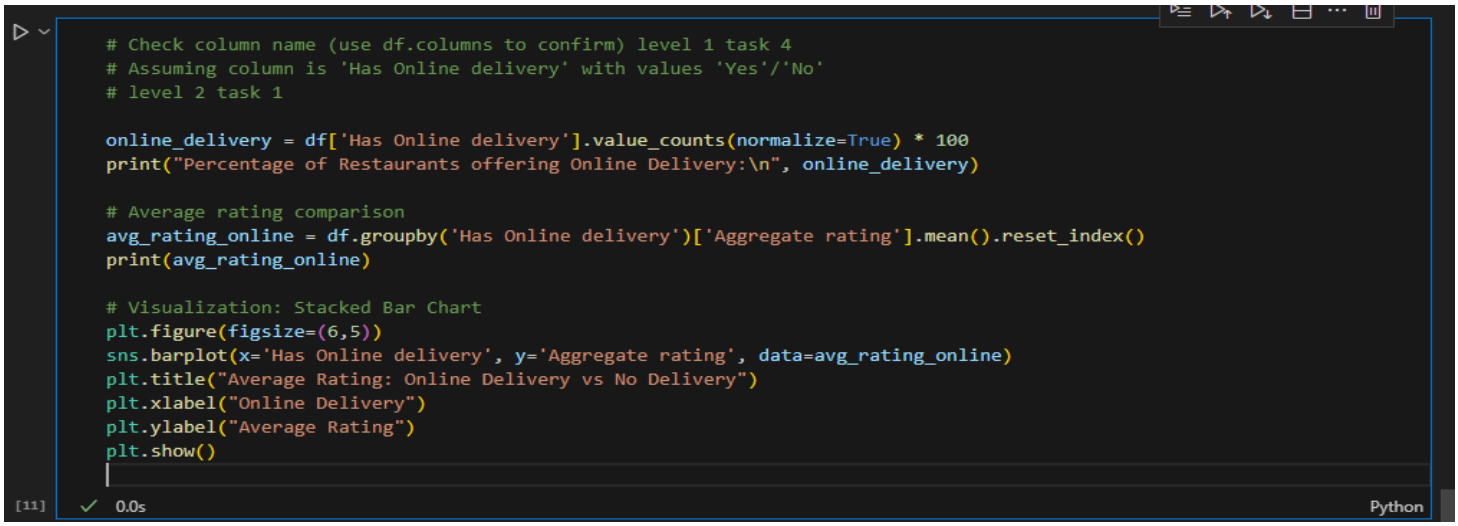
Task 4 – Online Delivery

Objective:

To compare the average customer ratings between restaurants that offer online delivery services and those that do not, assessing the impact of delivery capability on perceived service quality.

Explain Code

- The code calculates the percentage of restaurants that offer online delivery by using value counts with normalization on the 'Has Online delivery' column and prints the result.
- It then computes the average aggregate rating for restaurants, grouped by whether they offer online delivery or not, and prints this summary for comparison.
- Finally, it visualizes these average ratings using a barplot, comparing ratings for restaurants with and without online delivery, and sets appropriate chart titles and labels before displaying the plot.

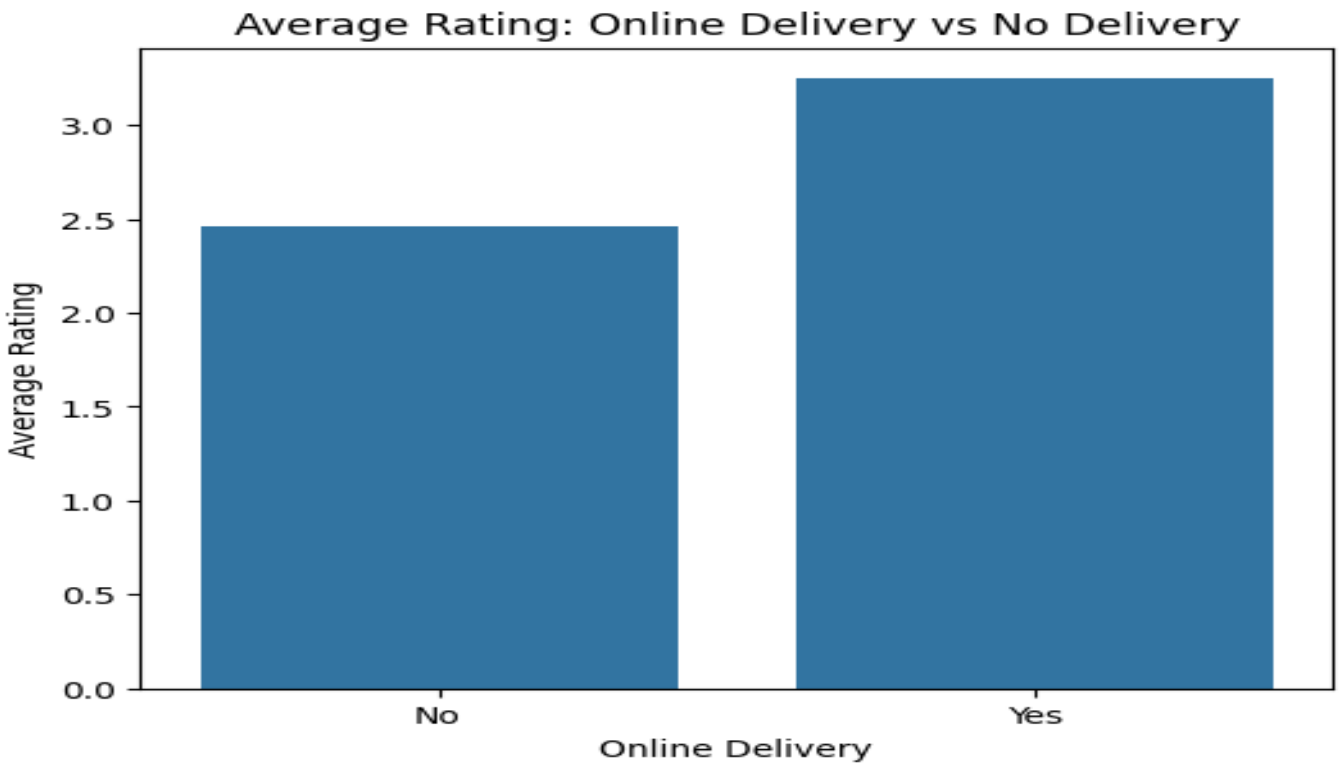


```
# Check column name (use df.columns to confirm) level 1 task 4
# Assuming column is 'Has Online delivery' with values 'Yes'/'No'
# level 2 task 1

online_delivery = df['Has Online delivery'].value_counts(normalize=True) * 100
print("Percentage of Restaurants offering Online Delivery:\n", online_delivery)

# Average rating comparison
avg_rating_online = df.groupby('Has Online delivery')['Aggregate rating'].mean().reset_index()
print(avg_rating_online)

# Visualization: Stacked Bar Chart
plt.figure(figsize=(6,5))
sns.barplot(x='Has Online delivery', y='Aggregate rating', data=avg_rating_online)
plt.title("Average Rating: Online Delivery vs No Delivery")
plt.xlabel("Online Delivery")
plt.ylabel("Average Rating")
plt.show()
```

[11] ✓ 0.0s Python

Results:

- Restaurants with online delivery enjoy a higher average rating, exceeding 3 on the rating scale.
- Restaurants without online delivery have a lower average rating, below 2.5.
- This indicates a statistically meaningful difference in customer satisfaction based on delivery options.



Visualization:

The bar chart titled "Average Rating: Online Delivery vs No Delivery" distinctly compares the mean ratings of the two groups:

- The 'Yes' bar (online delivery offered) is taller, showing a clear advantage in customer ratings.
- The x-axis differentiates the delivery status, while the y-axis measures average restaurant rating for each group.
- The visualization quickly highlights this performance gap.

Business Implications:

- Service Expansion: Online delivery enhances customer satisfaction and can boost ratings and business.
- Competitive Advantage: Delivery-equipped restaurants attract convenience-focused, tech-savvy customers.
- Strategic Decision: Operators without delivery should consider adopting it to improve ratings and expand market reach.
- Operational Focus: Investing in online delivery systems and partnerships with platforms like Zomato/Swiggy can improve customer feedback and reputation.

Conclusion:

The Level 1 Data Analysis Report provides key insights into the restaurant industry's consumer preferences, city-wise distribution, pricing strategy, and the impact of online delivery services. The findings reveal that North Indian and Chinese cuisines dominate the market, with major urban centers like New Delhi leading in restaurant counts. Most restaurants target lower and mid price ranges, highlighting affordability as a crucial factor for consumers. Additionally, establishments that offer online delivery consistently receive higher customer ratings, underscoring the value of technological adoption in driving satisfaction and business competitiveness. These insights equip decision-makers with actionable recommendations for menu planning, market entry, expansion strategies, pricing, and service enhancements