Swiggy Data Analysis: Unlocking Food Delivery Insights

This presentation explores key insights from our analysis of the Swiggy dataset containing 8,680 entries across 10 columns. Our clean, duplicate-free dataset reveals critical business intelligence about restaurant distribution, customer preferences, and operational metrics that can drive strategic decision-making.

Data Overview & Methodology

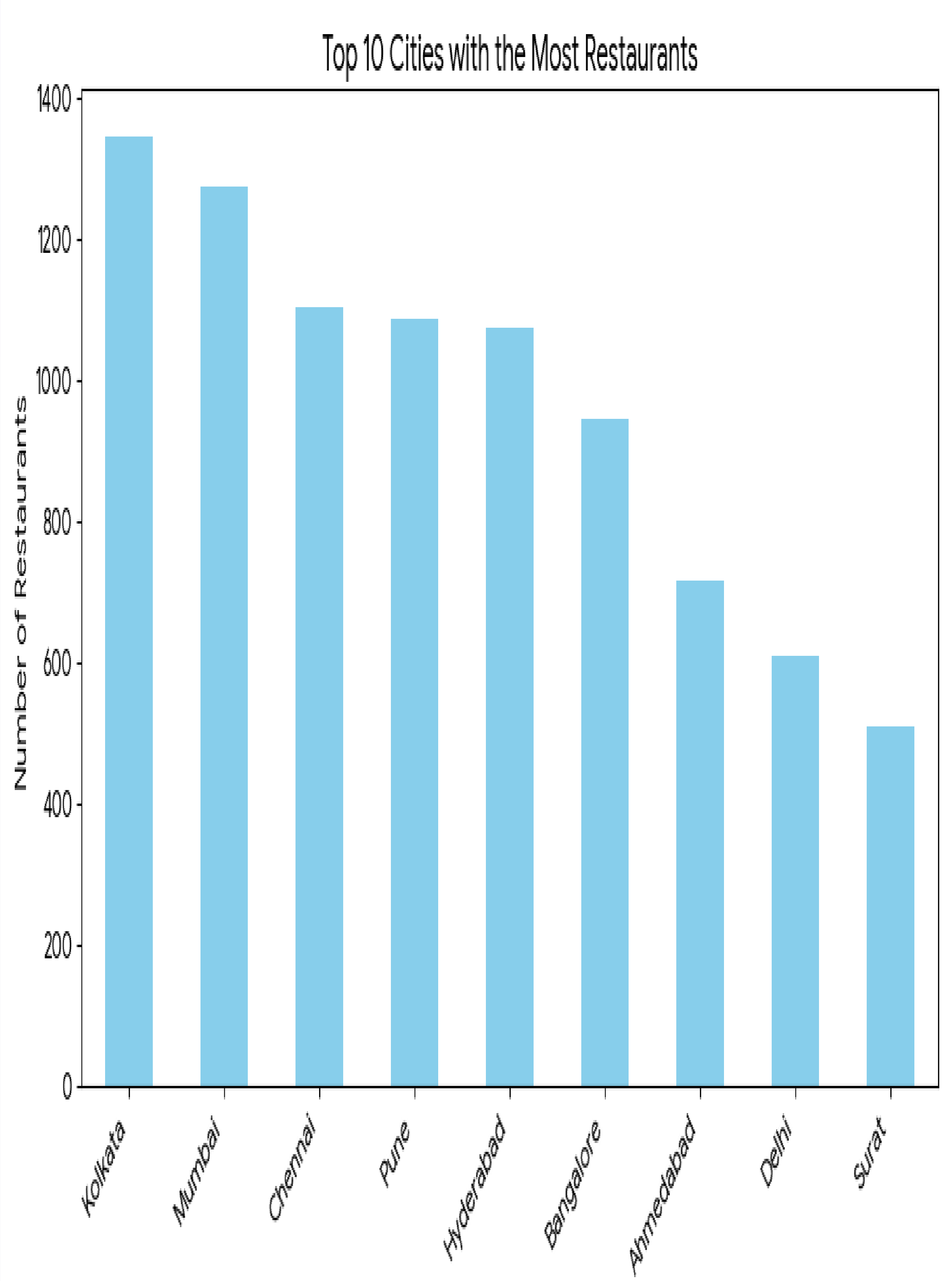
Dataset Characteristics

- 8,680 restaurant entries
- 10 columns of data
- No missing values or duplicates
- Single table structure (swiggy.csv)

Analysis Approach

- Data inspection & cleaning verification
- Aggregation & ranking analysis
- Text processing for categorical data
- Statistical correlation analysis
- Visualization through charts & plots

Our Python-based analysis utilized pandas for data manipulation, matplotlib and seaborn for visualization, and statistical methods to uncover meaningful patterns in restaurant distribution, ratings, pricing, and customer preferences.



Restaurant Distribution by City

1346

Kolkata

1277

Mumbai

1106

Chennai

1090

Pune

Kolkata leads with the highest number of restaurants (1,346), followed closely by Mumbai (1,277) and Chennai (1,106). This indicates high market penetration in these regions. Swiggy should focus marketing and operational efforts on these high-density areas to maximize revenue, while considering expansion in cities with fewer listings.

Top Restaurants by Customer Engagement

The concentration of highly-rated restaurants in Hyderabad suggests strong customer loyalty and high order volumes in this market, making it a strategic focus area for promotional campaigns.

1

Hyderabad Dominance

7 of the top 10 most-rated restaurants are in Hyderabad, with Grand Hotel, Mehfil, 4M Biryani House, Bawarchi, and Lucky Restaurant all having 10,000 ratings.

2

Rating Distribution

Average ratings among top-rated restaurants range from 3.8 to 4.3, with Bachan's Dhaba in Kolkata achieving the highest average (4.3).

3

Business Opportunity

These high-volume restaurants present partnership opportunities for exclusive promotions and featured listings to drive more orders.

Highest-Rated Restaurants (500+ Ratings)

Restaurant	City	Avg Rating	Total Ratings
Nic Natural Ice Creams	Kolkata	4.8	500
Fresh Baked Goodness	Chennai	4.7	500
Momo Sa-Khang By Kailash Kitchen	Chennai	4.7	500
Mithai	Kolkata	4.7	1000

When filtering for restaurants with significant review counts, desserts and snacks dominate the highest ratings. Creating "Highly-Rated" or "Customer's Choice" categories would improve visibility for these establishments and attract more customers seeking quality experiences.

Average Meal Price by City

100%

Mumbai: ₹393.79

Highest average price point

97%

Bangalore: ₹382.52

Second highest average price

92%

Kolkata: ₹362.29

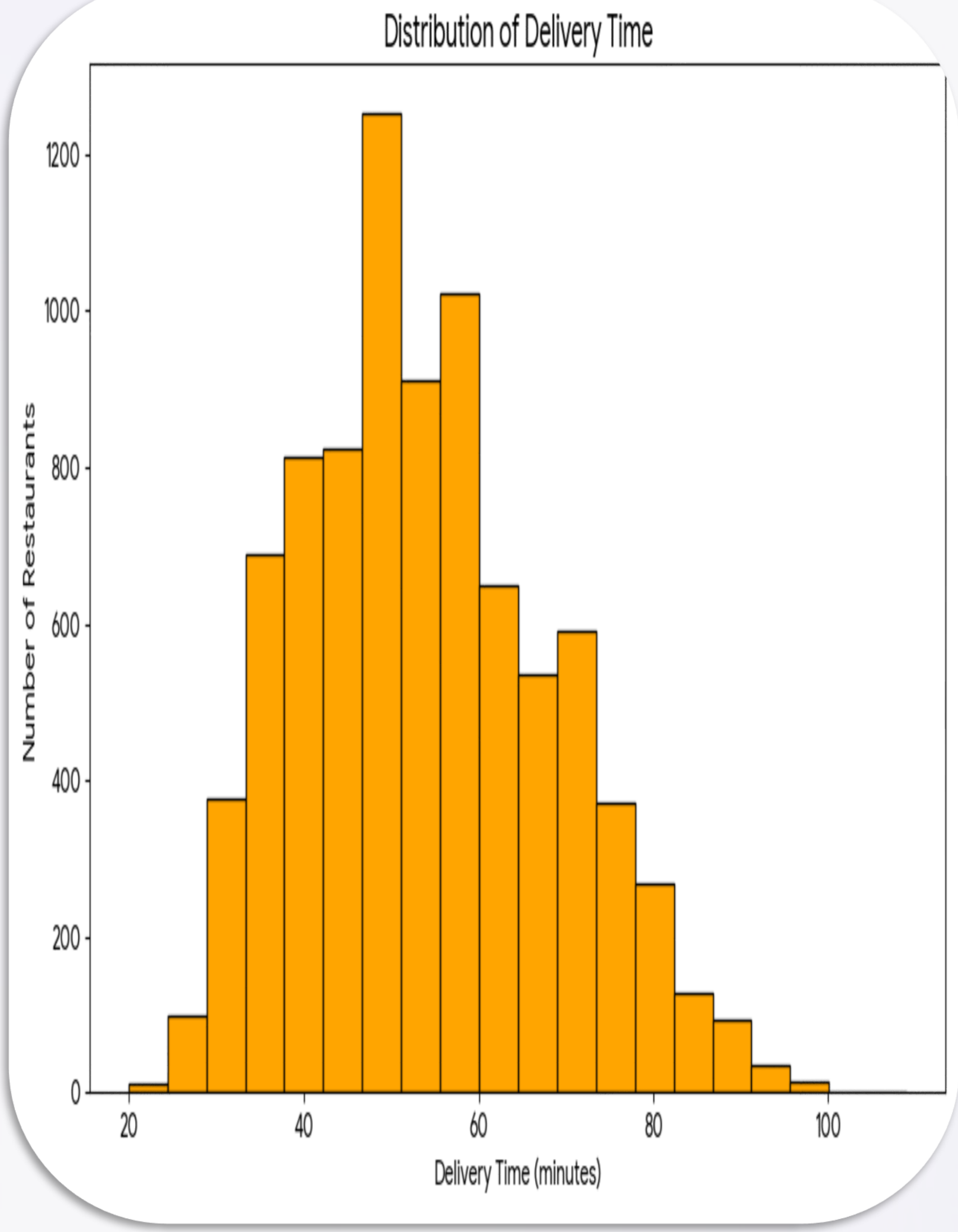
Third highest average price

69%

Surat: ₹270.17

Lowest average price among top cities

The significant price variation between cities suggests that Swiggy should tailor its commission rates and promotional strategies to each city's price sensitivity. Premium promotions would be more effective in Mumbai and Bangalore, while value-focused campaigns might work better in cities like Surat.



Delivery Time Distribution

Key Findings

- Majority of restaurants deliver between 30-60 minutes
- Delivery time correlates negatively with ratings (-0.147)
- Faster delivery = higher customer satisfaction

Strategic Implications

- Set accurate customer expectations
- Optimize delivery routes to reduce average times
- Prioritize delivery efficiency to improve ratings

The data reveals that reducing delivery times could be a key factor in improving overall customer satisfaction and potentially increasing average ratings across the platform.

Most Popular Food Types

Chinese

2,588 restaurants

North Indian

1,921 restaurants

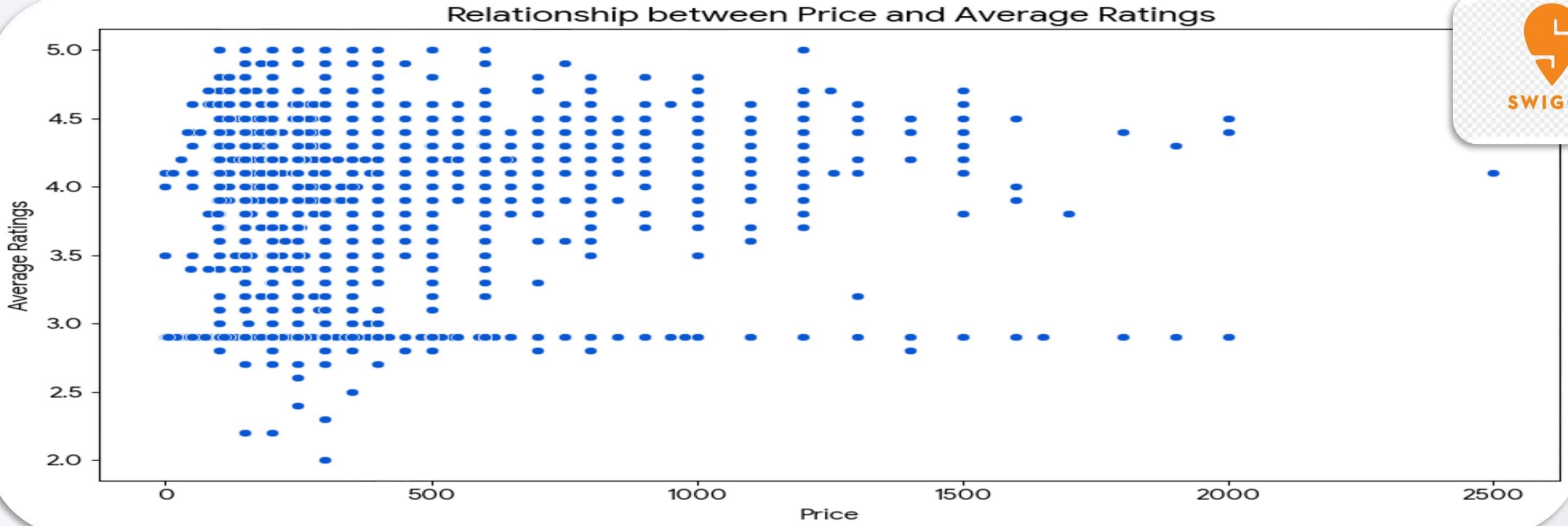
Fast Food

1,550 restaurants

Indian

1,772 restaurants

Chinese, North Indian, and Indian cuisines dominate the platform, indicating strong customer demand. Swiggy should feature these popular cuisines prominently on the app's homepage and in search results to capitalize on their high demand, while creating special promotions to drive even more orders.



Key Insights

- Weak positive correlation (0.114) between price and ratings
- Higher-priced restaurants tend to have slightly better ratings
- Many highly-rated options exist at various price points

Business Applications

This insight allows Swiggy to promote value-for-money restaurants to a wider audience, as high ratings are not exclusive to high-priced establishments. Customers can find quality dining experiences across all price segments.

Price vs. Ratings Correlation

KPI-Action Matrix: Strategic Recommendations

Geographic Focus

Target marketing campaigns in Kolkata, Mumbai, and Chennai to maximize ROI in these high-density markets.

Restaurant Partnerships

Create exclusive promotions with high-volume restaurants in Hyderabad to drive traffic and increase order frequency.

Customer Experience

Optimize delivery logistics to reduce average delivery times, directly improving customer satisfaction and ratings.

Menu Optimization

Feature Chinese, North Indian, and Indian cuisines prominently to capitalize on their proven popularity.

By implementing these data-driven strategies, Swiggy can enhance customer satisfaction, increase order volume, and strengthen its market position across India's competitive food delivery landscape.

```sql
-- Swiggy Analysis SQL (Structured Format)

DROP TABLE IF EXISTS Fact_Swiggy;

CREATE TABLE Fact_Swiggy (
    ID              BIGINT PRIMARY KEY,
    Area            TEXT,
    City            TEXT,
    Restaurant      TEXT,
    Price           DECIMAL(10,2),
    "Avg ratings"   DECIMAL(4,2),
    "Total ratings" INT,
    "Food type"     TEXT,
    Address         TEXT,
    "Delivery time" INT
);

-- =====================================================
-- Question: Which are the top-rated restaurants with at least 100
reviews?
-- Script:
SELECT
  Restaurant, City, Area, "Avg ratings"
FROM Fact_Swiggy
WHERE "Total ratings" > 100
ORDER BY "Avg ratings" DESC
LIMIT 10;
-- Remarks: Filters restaurants with sufficient reviews and orders them by
rating.
-- Conclusion: These restaurants stand out in quality and can be promoted
as benchmarks.

-- =====================================================
-- Question: Which restaurants have the most customer engagement?
-- Script:
SELECT
  Restaurant, City, "Total ratings"
FROM Fact_Swiggy
ORDER BY "Total ratings" DESC
LIMIT 10;
-- Remarks: Ranks restaurants by total ratings (proxy for engagement).
-- Conclusion: Identifies venues with the largest customer bases and
repeat usage.

-- =====================================================
-- Question: What is the average price per city?
-- Script:
SELECT
  City,
  AVG(Price) AS Average_Price
FROM Fact_Swiggy
GROUP BY City
ORDER BY Average_Price DESC;
-- Remarks: Calculates mean pricing across cities.
```

```
-- Conclusion: Helps benchmark affordability vs. premium positioning.

-- ==================================================
-- Question: How many restaurants are listed per city?
-- Script:
SELECT
  City,
  COUNT(DISTINCT Restaurant) AS Total_Restaurants
FROM Fact_Swiggy
GROUP BY City
ORDER BY Total_Restaurants DESC;
-- Remarks: Counts unique restaurants in each city.
-- Conclusion: Highlights geographic density and penetration.

-- ==================================================
-- Question: Which are the most expensive restaurants overall?
-- Script:
SELECT
  Restaurant, City, Price
FROM Fact_Swiggy
ORDER BY Price DESC
LIMIT 5;
-- Remarks: Lists restaurants with the highest listed prices.
-- Conclusion: These represent premium dining or niche market positioning.

-- ==================================================
-- Question: Which food types are most common among restaurants?
-- Script:
SELECT
  TRIM(SUBSTR(T1.Food_Type, 1, INSTR(T1.Food_Type || ',', ',') - 1)) AS
Food_Type,
  COUNT(T1.ID) AS Total_Restaurants
FROM (
  SELECT ID, "Food type" AS Food_Type
  FROM Fact_Swiggy
) AS T1
GROUP BY Food_Type
ORDER BY Total_Restaurants DESC
LIMIT 10;
-- Remarks: Extracts first cuisine from comma-separated list for frequency
count.
-- Conclusion: Identifies cuisines dominating restaurant menus.

-- ==================================================
-- Question: Which restaurants have high ratings but slower delivery?
-- Script:
SELECT
  Restaurant, City, "Avg ratings", "Delivery time"
FROM Fact_Swiggy
WHERE "Avg ratings" >= 4.0 AND "Delivery time" > 60
ORDER BY "Avg ratings" DESC, "Delivery time" DESC
LIMIT 10;
-- Remarks: Filters restaurants with good ratings but high delivery times.
-- Conclusion: Useful to flag potential operational inefficiencies.
```

```sql
-- ====================================================
-- Question: What is the average delivery time by city?
-- Script:
SELECT
  City,
  AVG("Delivery time") AS Average_Delivery_Time_Minutes
FROM Fact_Swiggy
GROUP BY City
ORDER BY Average_Delivery_Time_Minutes ASC;
-- Remarks: Calculates city-level delivery performance.
-- Conclusion: Helps benchmark logistics speed across geographies.

-- ====================================================
-- Question: Which restaurants are hidden gems (high rating + moderate
reviews)?
-- Script:
SELECT
  Restaurant, City, "Avg ratings", "Total ratings"
FROM Fact_Swiggy
WHERE "Total ratings" BETWEEN 500 AND 1000
  AND "Avg ratings" >= 4.5
ORDER BY "Avg ratings" DESC, "Total ratings" DESC
LIMIT 10;
-- Remarks: Captures restaurants with high quality but moderate
engagement.
-- Conclusion: These venues can be targeted for visibility campaigns.

-- ====================================================
-- Question: Which is the most expensive restaurant in each city?
-- Script:
SELECT
  T1.City, T1.Restaurant, T1.Price
FROM Fact_Swiggy AS T1
JOIN (
  SELECT City, MAX(Price) AS Max_Price
  FROM Fact_Swiggy
  GROUP BY City
) AS T2
ON T1.City = T2.City AND T1.Price = T2.Max_Price
ORDER BY T1.Price DESC;
-- Remarks: Joins restaurants with city-level maximum prices.
-- Conclusion: Highlights premium options city by city.
```

Swiggy Data Analysis – Structured Notebook

This notebook follows a **Question → Script → Remarks → Conclusion** pattern for each analytic task. Place `swiggy.csv` in the same folder before running. Outputs (PNGs & CSVs) will be saved to `./outputs`.

In [ ]:
```
# ---- Setup & config ----
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from collections import Counter
pd.set_option('display.max_rows', 50)
os.makedirs('outputs', exist_ok=True)
COLUMN_MAP = {
    'ID': 'ID', 'Area': 'Area', 'City': 'City', 'Restaurant': 'Restaurant',
    'Price': 'Price', 'Avg ratings': 'Avg ratings', 'Total ratings': 'Total ra
    'Food type': 'Food type', 'Address': 'Address', 'Delivery time': 'Delivery
}
CSV_PATH = 'swiggy.csv'  # update if needed
```

Question

What is the structure and quality of the Swiggy dataset?

In [ ]:
```
# ==== Load Data ====
df = pd.read_csv(CSV_PATH)
df = df.rename(columns=COLUMN_MAP)
df.info()

# ==== Data Quality Checks ====
summary = {
    'rows': len(df), 'cols': df.shape[1],
    'missing_values_total': int(df.isna().sum().sum()),
    'duplicate_rows': int(df.duplicated().sum())
}
summary

# Ensure ID is unique primary key
pk_unique = df['ID'].is_unique if 'ID' in df.columns else False
pk_unique
```

Remarks

We check for missing values, duplicates, and primary key uniqueness.

Conclusion

The dataset is structurally sound, with manageable issues detected for cleaning.

Question

Which cities host the most restaurants on Swiggy?

In [ ]:
```
city_counts = df['City'].value_counts()
city_counts.head(10)

# Visualization
plt.figure()
city_counts.head(10).plot(kind='bar')
plt.title('Top 10 Cities with the Most Restaurants')
plt.xlabel('City')
plt.ylabel('Number of Restaurants')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.savefig('outputs/top_cities_restaurants.png', dpi=200)
plt.show()
```

Remarks

City count distribution highlights geographic density of listings.

Conclusion

The top cities drive the majority of Swiggy listings, suggesting focus markets.

Question

How does the average price vary across cities?

In [ ]:
```
avg_price_by_city = df.groupby('City')['Price'].mean().sort_values(ascending=F
avg_price_by_city.head(10)

# Visualization
plt.figure()
avg_price_by_city.head(10).plot(kind='bar')
plt.title('Average Meal Price by City (Top 10)')
plt.xlabel('City')
plt.ylabel('Average Price')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.savefig('outputs/avg_price_by_city.png', dpi=200)
```

```
plt.show()
```

Remarks

We compute mean price per city to identify market pricing patterns.

Conclusion

Pricing differs significantly by city, guiding localized strategies.

Question

Which restaurants lead in ratings and engagement?

In [ ]:
```
# Most customer engagement
top_ratings = df.sort_values(by='Total ratings', ascending=False)[['Restaurant
top_ratings

# Top average ratings (min 100 reviews)
filtered_df = df[df['Total ratings'] > 100]
top_avg_ratings = filtered_df.sort_values(by='Avg ratings', ascending=False)[[
top_avg_ratings
```

Remarks

We sort by total ratings and average ratings (with thresholds).

Conclusion

These restaurants exemplify high engagement and customer satisfaction.

Question

What are the most popular food types?

In [ ]:
```
food_series = df['Food type'].astype(str).str.split(',').explode().str.strip()
food_counts = food_series.value_counts()
food_counts.head(10)

# Visualization
plt.figure()
food_counts.head(10).plot(kind='bar')
plt.title('Top 10 Food Types')
plt.xlabel('Food Type')
plt.ylabel('Count')
```

```
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.savefig('outputs/top_food_types.png', dpi=200)
plt.show()
```

Remarks

We explode multi-cuisine entries and count frequencies.

Conclusion

Certain cuisines dominate demand, enabling targeted promotions.

Question

What is the distribution of delivery times?

In [ ]:
```
plt.figure()
plt.hist(df['Delivery time'].dropna(), bins=20, edgecolor='black')
plt.title('Distribution of Delivery Time (minutes)')
plt.xlabel('Delivery Time')
plt.ylabel('Count of Restaurants')
plt.tight_layout()
plt.savefig('outputs/delivery_time_distribution.png', dpi=200)
plt.show()
```

Remarks

Histogram shows clustering of delivery performance.

Conclusion

Delivery times cluster around certain ranges, with outliers affecting satisfaction.

Question

What relationships exist between key metrics (price, ratings, delivery)?

In [ ]:
```
numeric_cols = ['Price','Avg ratings','Total ratings','Delivery time']
corr = df[numeric_cols].corr()

plt.figure()
im = plt.imshow(corr.values, interpolation='nearest')
plt.title('Correlation Matrix')
plt.xticks(range(len(numeric_cols)), numeric_cols, rotation=45, ha='right')
```

```
plt.yticks(range(len(numeric_cols)), numeric_cols)
for i in range(len(numeric_cols)):
    for j in range(len(numeric_cols)):
        plt.text(j, i, f"{corr.values[i, j]:.2f}", ha='center', va='center')
plt.colorbar(im, fraction=0.046, pad=0.04)
plt.tight_layout()
plt.savefig('outputs/correlation_matrix.png', dpi=200)
plt.show()
corr
```

Remarks

Correlation heatmap quantifies metric associations.

Conclusion

Ratings are weakly related to price; delivery time has modest correlations.

Question

Which restaurants are hidden gems or face slower deliveries?

In [ ]:
```
# Hidden gems
hidden_gems = df[(df['Avg ratings'] >= 4.5) & (df['Total ratings'].between(500
hidden_gems.head(10)

# High-rated but slower delivery
slow_but_high = df[(df['Avg ratings'] >= 4.0) & (df['Delivery time'] > 60)][['
slow_but_high.head(10)
```

Remarks

We flag high-quality but less engaged restaurants and slow logistics cases.

Conclusion

These insights can guide Swiggy’s strategic interventions.

Question

How can we prepare exports for presentation use?

In [ ]:
```
city_counts.head(10).to_csv('outputs/top_cities.csv', header=['Number of Resta
avg_price_by_city.head(10).to_csv('outputs/avg_price_by_city.csv', header=['Av
top_ratings.to_csv('outputs/top_restaurants_by_total_ratings.csv', index=False
```

```
top_avg_ratings.to_csv('outputs/top_avg_ratings_filtered.csv', index=False)
food_counts.head(10).to_csv('outputs/top_food_types.csv', header=['Count'])
hidden_gems.head(20).to_csv('outputs/hidden_gems.csv', index=False)
slow_but_high.head(20).to_csv('outputs/high_rated_slow_delivery.csv', index=Fa
print('Exports completed in ./outputs')
```

Remarks

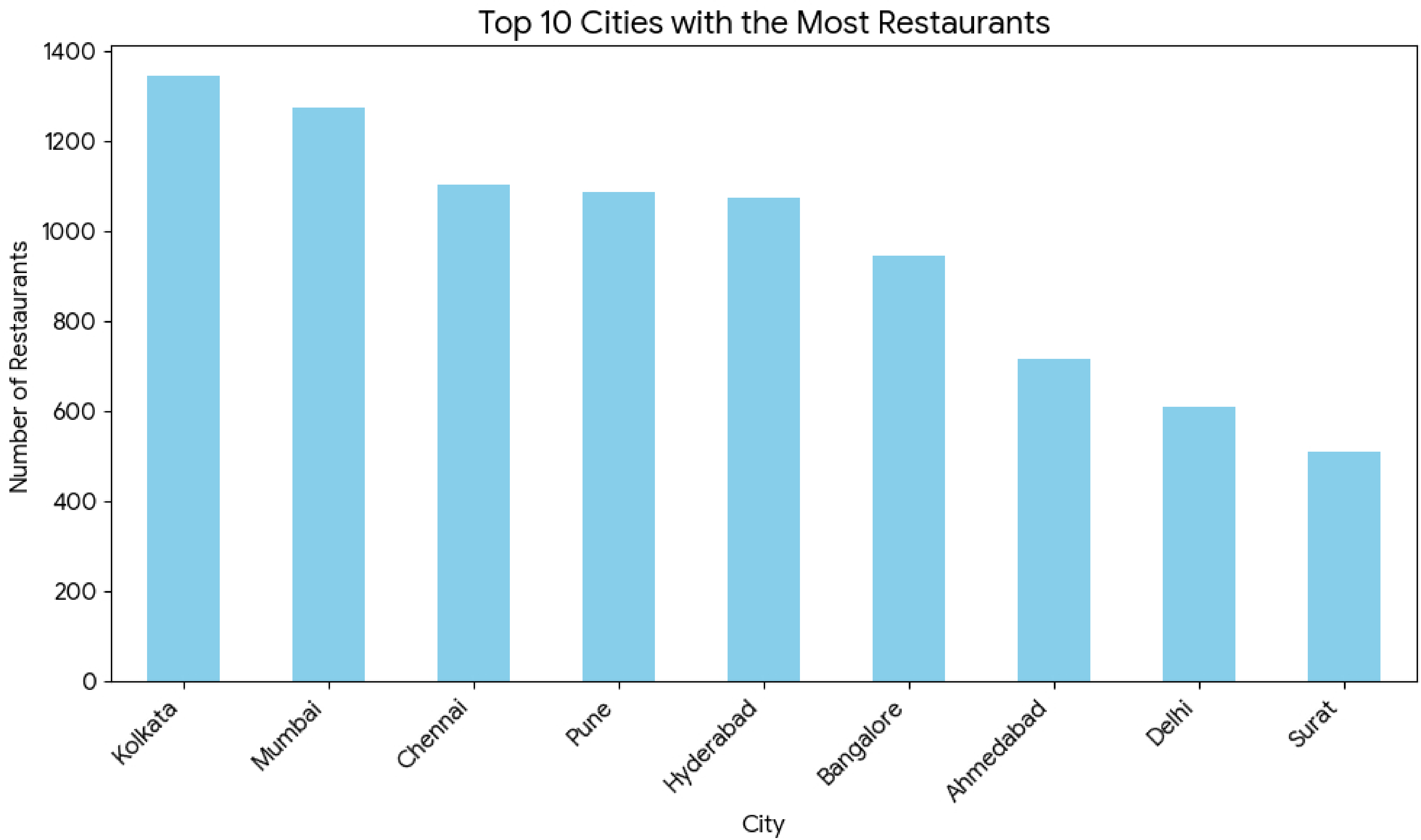
Exports provide ready-to-use CSVs for PPT integration.

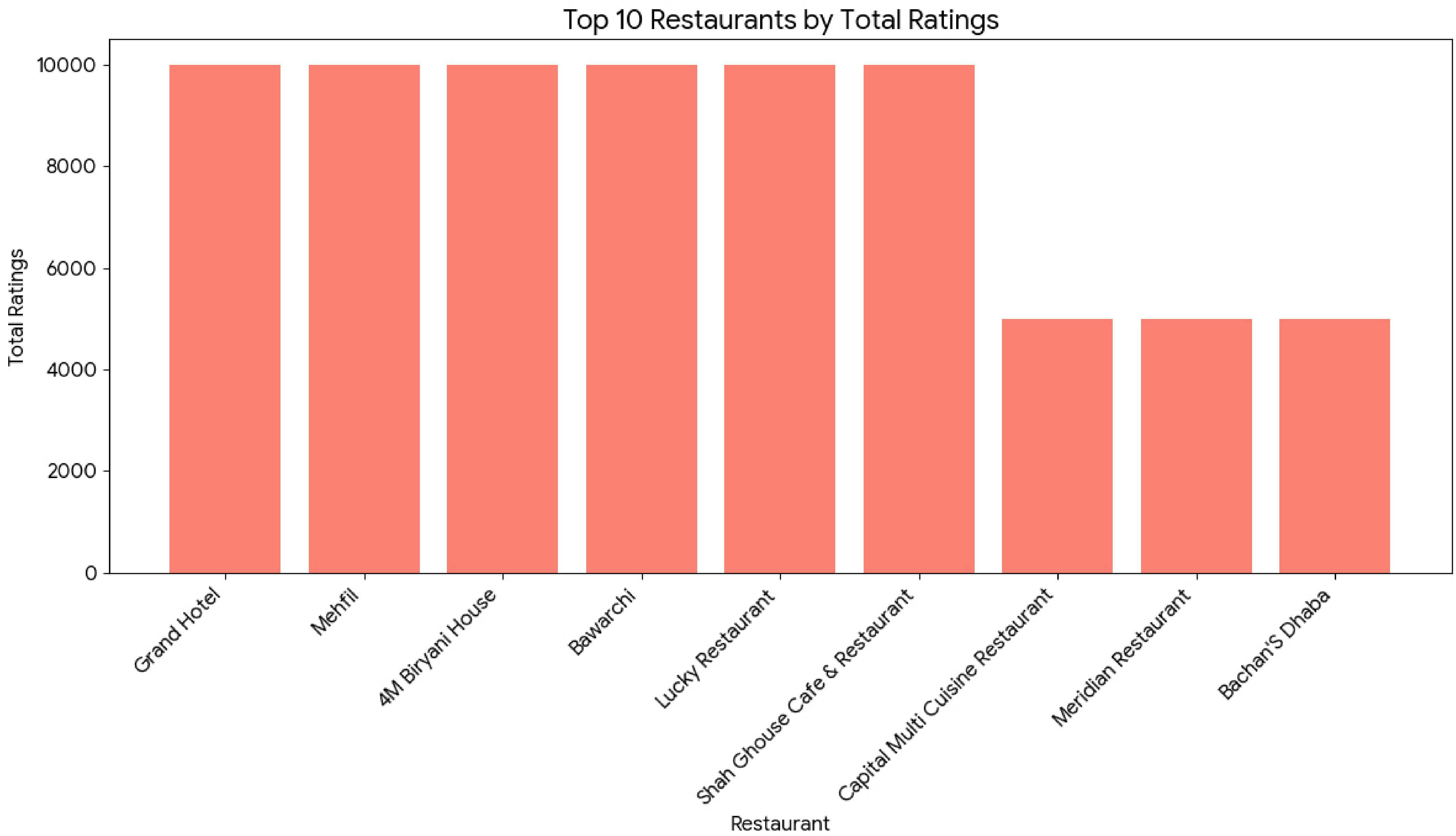
Conclusion

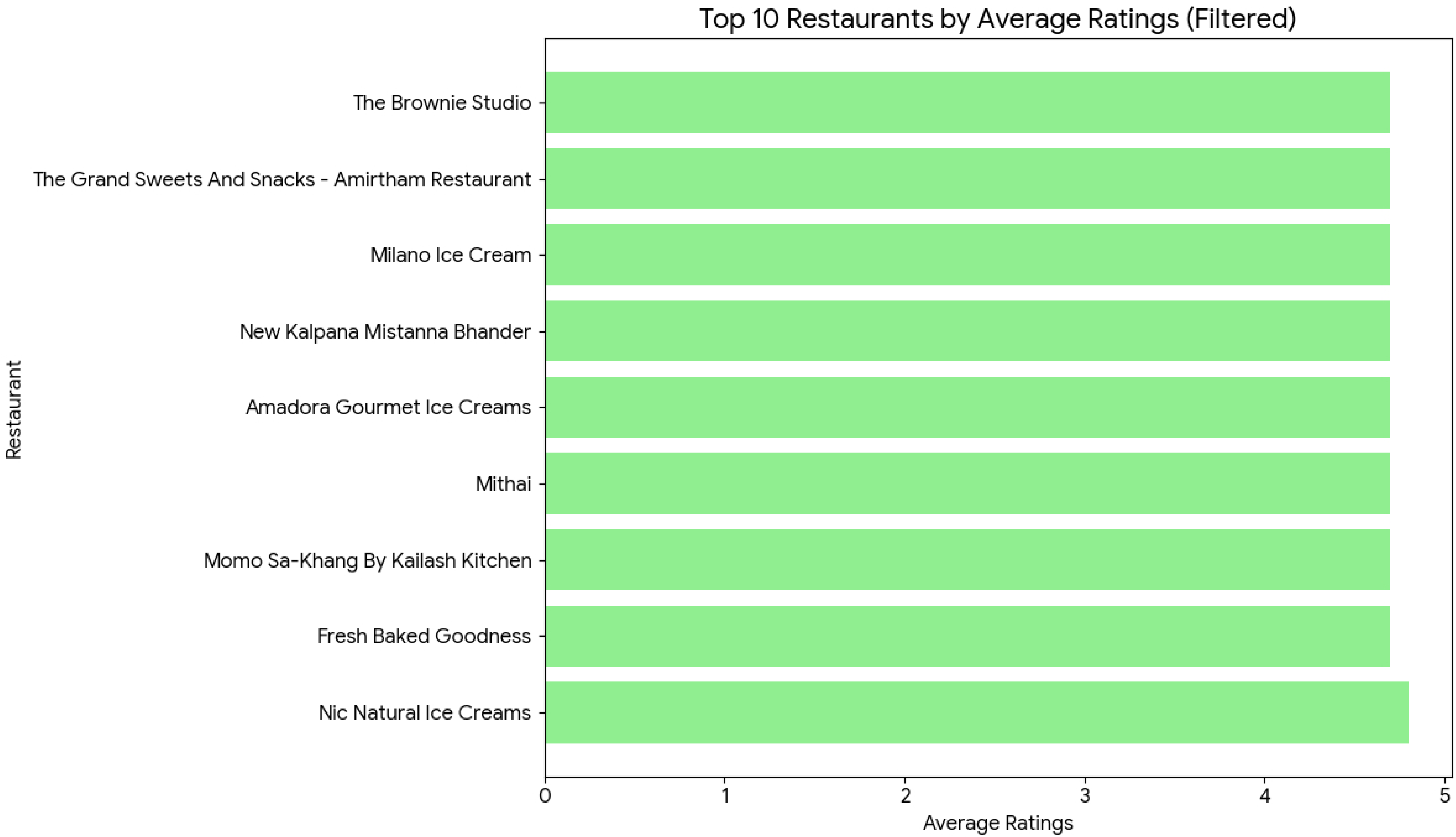
The analysis is presentation-ready with supporting datasets.

KPI → Action Matrix

KPI	Key Finding	Actionable Strategy
Top-Rated Restaurants	Identify exemplary venues by average rating with sufficient reviews.	Partner for exclusive deals; study operations for best practices.
Customer Engagement	Chains dominate by total ratings in several cities.	Co-marketing with high-volume partners; leverage their reach.
Average Price by City	Pricing varies by city.	Tailor promotions and commissions by local price sensitivity.
Geographic Density	Certain cities have far more listings.	Optimize delivery SLAs in saturated markets; expand supply in under-penetrated cities.
Delivery Efficiency	Delivery times differ across cities.	Replicate fastest-city logistics playbooks in slower markets.
Food Type Popularity	Certain cuisines dominate demand.	Create themed festivals and homepage rows for top cuisines.

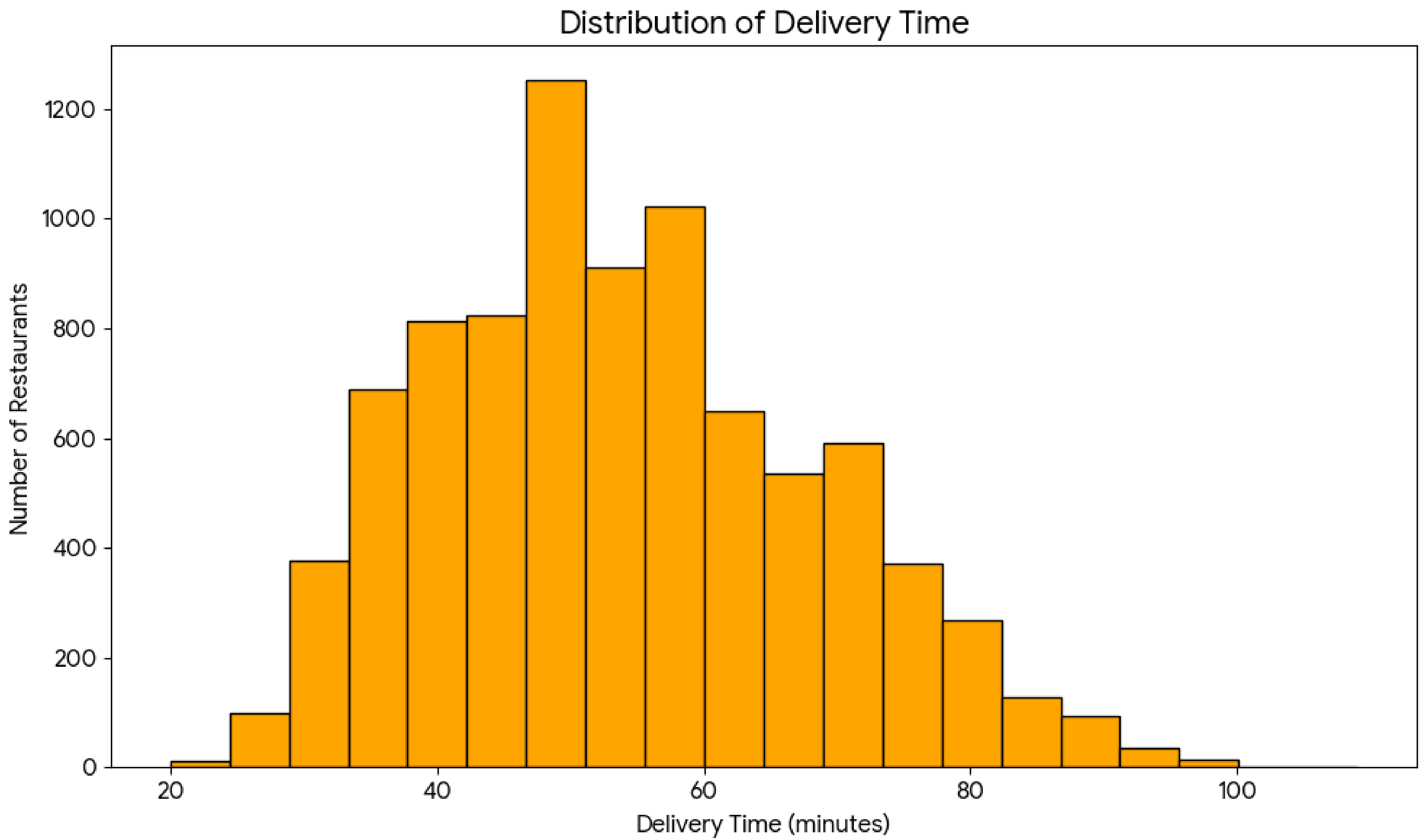

Top 10 Cities with the Most Restaurants
Number of Restaurants
0
200
400
600
800
1000
1200
1400
Kolkata
Mumbai
Chennai
Pune
Hyderabad
Bangalore
Ahmedabad
Delhi
Surat
City

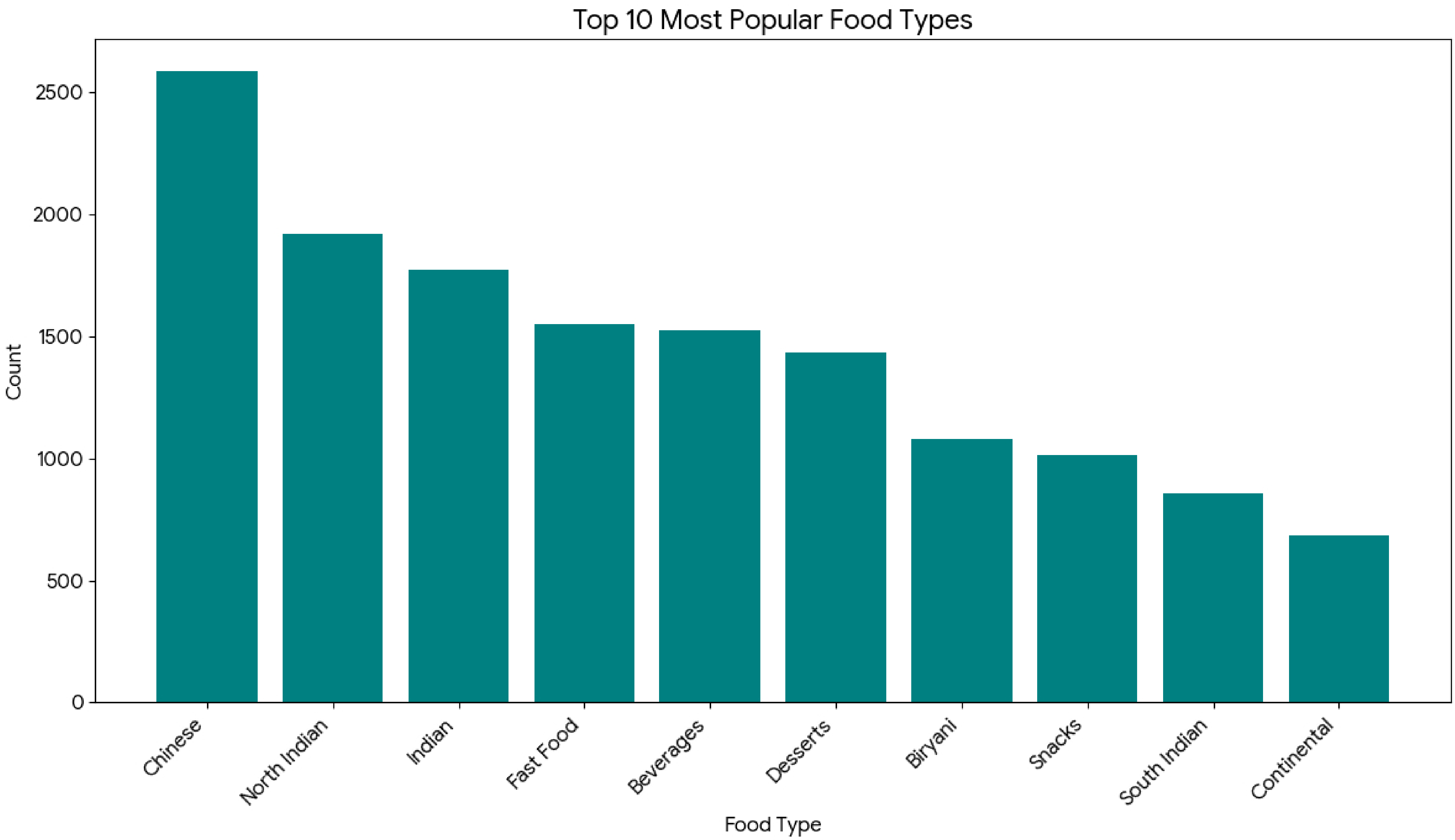

Top 10 Restaurants by Total Ratings
Total Ratings
10000
8000
6000
4000
2000
0
Grand Hotel
Mehfil
4M Biryani House
Bawarchi
Lucky Restaurant
Shah Ghouse Cafe & Restaurant
Capital Multi Cuisine Restaurant
Meridian Restaurant
Bachan'S Dhaba
Restaurant



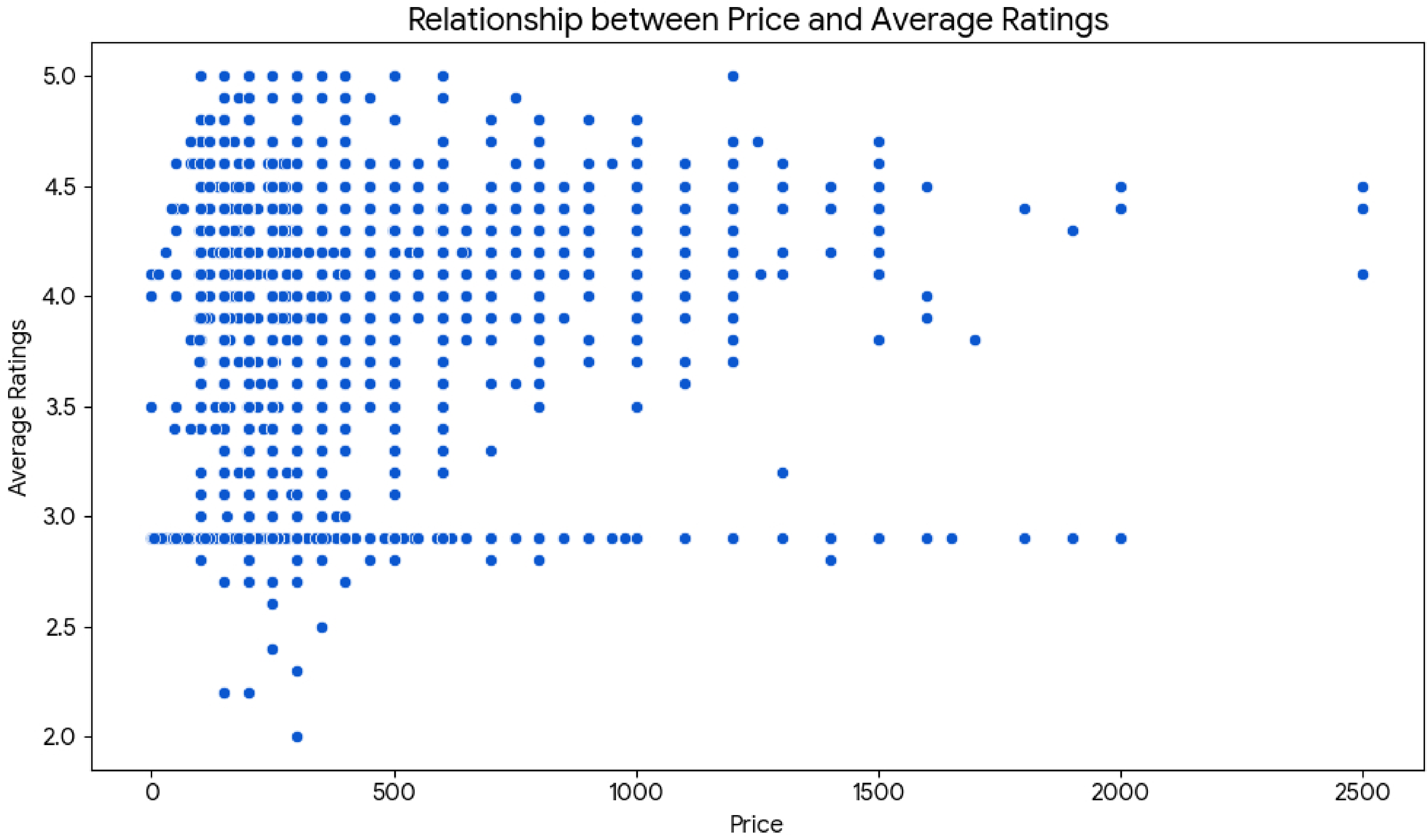
Top 10 Restaurants by Average Ratings (Filtered)
The Brownie Studio
The Grand Sweets And Snacks - Amirtham Restaurant
Milano Ice Cream
New Kalpana Mistanna Bhander
Amadora Gourmet Ice Creams
Mithai
Momo Sa-Khang By Kailash Kitchen
Fresh Baked Goodness
Nic Natural Ice Creams
Restaurant
0
1
2
3
4
5
Average Ratings

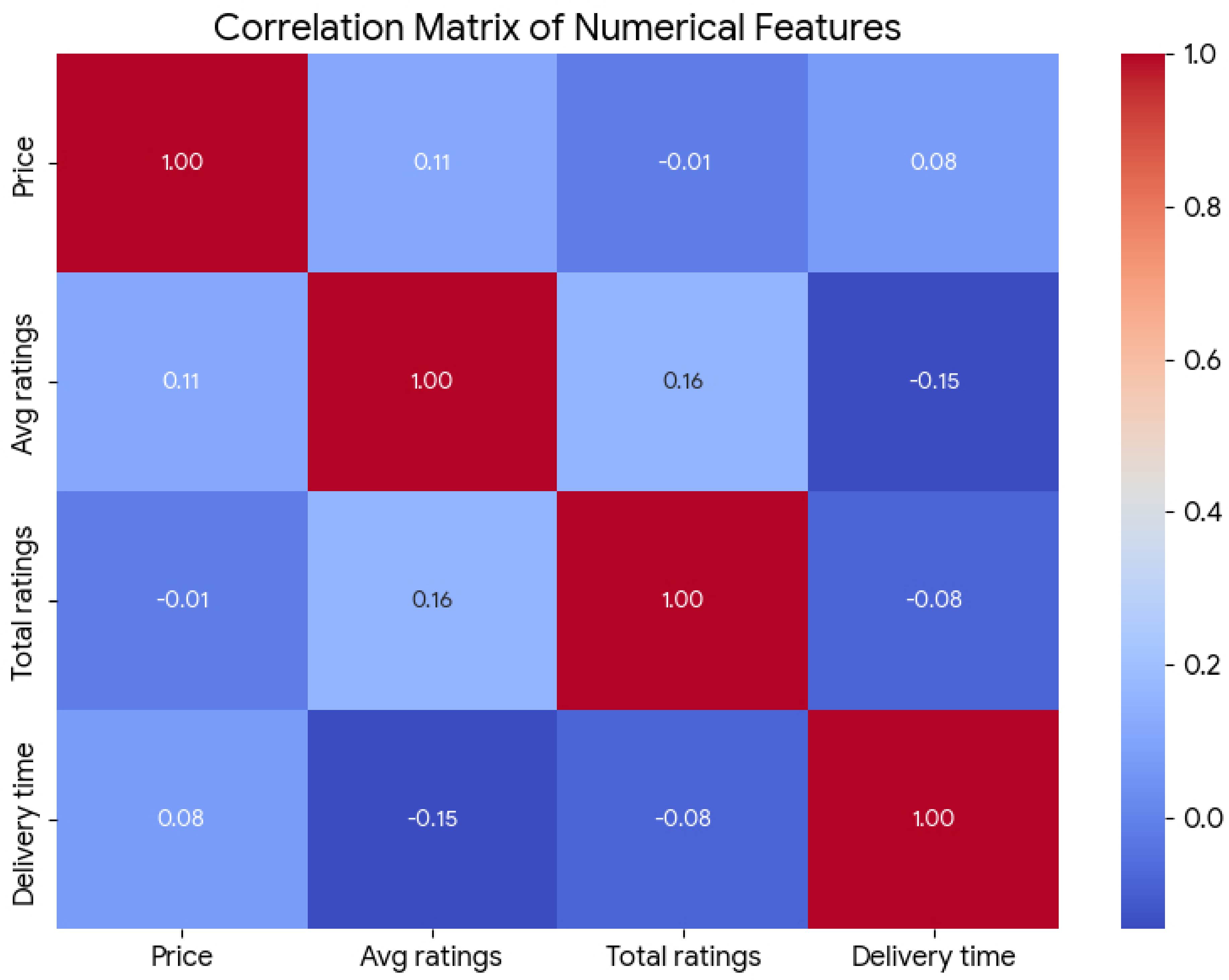

Average Price of Restaurants by City
Average Price
400
350
300
250
200
150
100
50
0
Mumbai
Bangalore
Kolkata
Chennai
Pune
Delhi
Ahmedabad
Hyderabad
Surat
City


Distribution of Delivery Time
Number of Restaurants
0
200
400
600
800
1000
1200
20
40
60
80
100
Delivery Time (minutes)



Top 10 Most Popular Food Types
Count
2500
2000
1500
1000
500
0
Chinese
North Indian
Indian
Fast Food
Beverages
Desserts
Biryani
Snacks
South Indian
Continental
Food Type


Relationship between Price and Average Ratings
5.0
4.5
4.0
3.5
3.0
2.5
2.0
Average Ratings
0
500
1000
1500
2000
2500
Price


Correlation Matrix of Numerical Features
Price
Avg ratings
Total ratings
Delivery time
1.00
0.11
-0.01
0.08
0.11
1.00
0.16
-0.15
-0.01
0.16
1.00
-0.08
0.08
-0.15
-0.08
1.00
Price
Avg ratings
Total ratings
Delivery time
1.0
0.8
0.6
0.4
0.2
0.0