Swiggy-Copy1

August 29, 2025

1 Consumer Behavior & Delivery Analytics: A Swiggy Dataset Exploration

1.1 Swiggy Company Insights

Swiggy is one of India's largest online food delivery platforms, known for its hyperlocal logistics network and customer-centric approach. Founded in 2014, it has expanded beyond food into groceries, alcohol delivery, and even concierge services through **Swiggy Genie**.

1.1.1 Business Model

- **Marketplace Aggregator**: Connects restaurants with customers via a centralized app.
- **Revenue Streams**:
  - Delivery fees
  - Restaurant commissions
  - In-app advertising
  - Subscription model (**Swiggy One**)
- **Tech-Driven Logistics**:
  - Real-time traffic and weather data
  - Demand forecasting
  - Route optimization for faster deliveries

This business model supports Swiggy's ability to deliver food efficiently while maintaining high customer satisfaction. The dataset used in this project reflects key operational metrics that align with Swiggy's strategic goals — including delivery time, ratings, pricing, and cuisine preferences.

1.2 Dataset Insights

This dataset provides a structured view of restaurant listings on Swiggy, capturing key attributes that influence customer experience and operational efficiency.

1.2.1 Column Descriptions

- **ID**: Unique identifier for each restaurant or entry

- **Area**: Locality or neighborhood where the restaurant is located

- **City**: Name of the city
- **Restaurant**: Name of the restaurant
- **Price**: Average price of an order (in)
- **Avg_Ratings**: Average customer rating (e.g., 4.2 out of 5)
- **Total_Ratings**: Total number of ratings received
- **Food_Type**: Cuisine or category (e.g., Chinese, South Indian)
- **Address**: Full restaurant address
- **Delivery_Time (min)**: Estimated delivery time in minutes

---

These features enable exploratory data analysis (EDA), customer behavior modeling, and performance benchmarking across cities, cuisines, and delivery metrics.

---

2 Problem Statements – Swiggy Dataset Analysis

The objective is to analyze Swiggy's restaurant dataset to uncover patterns, customer preferences, and operational trends for better decision-making.

2.0.1 Data Cleaning

1. Handle missing values in numeric & categorical columns.
2. Remove duplicates and fix data types.
3. Standardize inconsistent values (City, Food type).
4. Detect and treat outliers in Price, Ratings, Delivery time.

2.0.2 Exploratory Data Analysis (EDA)

5. **Average Delivery Time by City** – Compare delivery performance across cities.
6. **Distribution of Ratings by City (Boxplot)** – Identify variability, medians & outliers.
7. **Top 10 Restaurants by Count** – Find most frequently listed / ordered restaurants.

8. **Price Distribution (Histogram)** – Understand price range & spending trends.
9. **Highest-Rated Area in Each City** – Spot local hotspots with top customer satisfaction.
10. **Average Ratings per City** – Benchmark customer experience across cities.
11. **Top 3 Cuisines per City** – Reveal most popular food types region-wise.
12. **Unique Restaurants per City** – Measure restaurant variety available in each city.
13. **City-Level Summary (Price, Ratings, Delivery time)** – Compare overall performance.
14. **City & Area Aggregation (with cuisines & restaurants)** – View combined trends across numeric & non-numeric data.
15. **Correlation Heatmap** – Explore relationships between price, ratings, delivery time, etc.

2.0.3 Business Insights

16. Which cities deliver **fastest vs slowest**?
17. Which cities/areas have **highest vs lowest customer ratings**?
18. Which cuisines dominate in each region?
19. Do **higher prices correlate with better ratings**?
20. Does **delivery time impact ratings**?
21. Which areas/cities should Swiggy **target for promotions or growth**?
22. Which restaurants are **most popular** among customers?

---

[2]:
```python
#core libraries

import numpy as np
import pandas as pd

#visualization libraries

import matplotlib.pyplot as plt
import seaborn as sns
```

[3]:
```python
#load dataset
```

```
df= pd.read_csv(r"C:\Users\Vrishikaa\Downloads\archive (2)\swiggy.csv")
df.head(2)
```

[3]:
```
    ID         Area       City     Restaurant  Price  Avg ratings  \
0  211  Koramangala  Bangalore    Tandoor Hut  300.0          4.4
1  221  Koramangala  Bangalore  Tunday Kababi  300.0          4.1

   Total ratings                                  Food type    Address  \
0            100  Biryani,Chinese,North Indian,South Indian  5Th Block
1            100                           Mughlai,Lucknowi  5Th Block

   Delivery time
0             59
1             56
```

[4]:
```
# to remove duplicates

df.duplicated().sum()
df.drop_duplicates(inplace=True)
```

[5]:
```
#drop duplicates from the column

df = df.drop_duplicates()
```

[6]:
```
#to check if duplicates are droped or not

df.head(2)
```

[6]:
```
    ID         Area       City     Restaurant  Price  Avg ratings  \
0  211  Koramangala  Bangalore    Tandoor Hut  300.0          4.4
1  221  Koramangala  Bangalore  Tunday Kababi  300.0          4.1

   Total ratings                                  Food type    Address  \
0            100  Biryani,Chinese,North Indian,South Indian  5Th Block
1            100                           Mughlai,Lucknowi  5Th Block

   Delivery time
0             59
1             56
```

[7]:
```
#need to get the number of rows and columns

df.shape
```

[7]:
```
(8680, 10)
```

[8]:
```
#information from the data-set

df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8680 entries, 0 to 8679
Data columns (total 10 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   ID             8680 non-null   int64
 1   Area           8680 non-null   object
 2   City           8680 non-null   object
 3   Restaurant     8680 non-null   object
 4   Price          8680 non-null   float64
 5   Avg ratings    8680 non-null   float64
 6   Total ratings  8680 non-null   int64
 7   Food type      8680 non-null   object
 8   Address        8680 non-null   object
 9   Delivery time  8680 non-null   int64
dtypes: float64(2), int64(3), object(5)
memory usage: 678.3+ KB
```

[9]:
```
#to get the idea of all columns by mathematical insights

df.describe()
```

[9]:

	ID	Price	Avg ratings	Total ratings	Delivery time
count	8680.000000	8680.000000	8680.000000	8680.000000	8680.000000
mean	244812.071429	348.444470	3.655104	156.634793	53.967051
std	158671.617188	230.940074	0.647629	391.448014	14.292335
min	211.000000	0.000000	2.000000	20.000000	20.000000
25%	72664.000000	200.000000	2.900000	50.000000	44.000000
50%	283442.000000	300.000000	3.900000	80.000000	53.000000
75%	393425.250000	400.000000	4.200000	100.000000	64.000000
max	466928.000000	2500.000000	5.000000	10000.000000	109.000000

[10]:
```
#to find the missing values

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

[10]:
```
ID               0
Area             0
City             0
Restaurant       0
Price            0
Avg ratings      0
Total ratings    0
Food type        0
```

```
Address          0
Delivery time    0
dtype: int64
```

[11]:
```
#drop column

df.drop('Address',axis=1,inplace=True)
```

[12]:
```
df.head(1)
```

[12]:
```
    ID         Area       City   Restaurant  Price  Avg ratings  \
0  211  Koramangala  Bangalore  Tandoor Hut  300.0          4.4

   Total ratings                               Food type  Delivery time
0            100  Biryani,Chinese,North Indian,South Indian             59
```

[13]:
```
df.set_index(['City'],inplace=True)
df
```

[13]:
```
               ID                  Area               Restaurant  Price  \
City
Bangalore     211           Koramangala              Tandoor Hut  300.0
Bangalore     221           Koramangala            Tunday Kababi  300.0
Bangalore     246             Jogupalya                  Kim Lee  650.0
Bangalore     248           Indiranagar        New Punjabi Hotel  250.0
Bangalore     249           Indiranagar                      Nh8  350.0
…             …                     …                     …     …
Ahmedabad  464626  Panjarapole Cross Road               Malt Pizza  500.0
Delhi      465835                Rohini  Jay Mata Ji Home Kitchen  200.0
Delhi      465872                Rohini       Chinese Kitchen King  150.0
Delhi      465990                Rohini     Shree Ram Paratha Wala  150.0
Ahmedabad  466488            Navrangpura             Sassy Street  250.0

           Avg ratings  Total ratings  \
City
Bangalore          4.4            100
Bangalore          4.1            100
Bangalore          4.4            100
Bangalore          3.9            500
Bangalore          4.0             50
…                  …            …
Ahmedabad          2.9             80
Delhi              2.9             80
Delhi              2.9             80
Delhi              2.9             80
Ahmedabad          2.9             80
```

```
                                                        Food type  Delivery time
City
Bangalore          Biryani,Chinese,North Indian,South Indian             59
Bangalore                                    Mughlai,Lucknowi             56
Bangalore                                             Chinese             50
Bangalore               North Indian,Punjabi,Tandoor,Chinese             57
Bangalore  Rajasthani,Gujarati,North Indian,Snacks,Desser…             63
…                                                         …            …
Ahmedabad                                              Pizzas             40
Delhi                                            South Indian             28
Delhi                                   Chinese,Snacks,Tandoor             58
Delhi                               North Indian,Indian,Snacks             28
Ahmedabad                                 Chaat,Snacks,Chinese             44

[8680 rows x 8 columns]
```

[14]:
```
print(df.columns.tolist())
```

```
['ID', 'Area', 'Restaurant', 'Price', 'Avg ratings', 'Total ratings', 'Food
type', 'Delivery time']
```

[15]:
```
df
```

[15]:
```
               ID                  Area               Restaurant  Price  \
City
Bangalore     211           Koramangala              Tandoor Hut  300.0
Bangalore     221           Koramangala            Tunday Kababi  300.0
Bangalore     246             Jogupalya                  Kim Lee  650.0
Bangalore     248           Indiranagar        New Punjabi Hotel  250.0
Bangalore     249           Indiranagar                      Nh8  350.0
…              …                     …                        …      …
Ahmedabad  464626  Panjarapole Cross Road               Malt Pizza  500.0
Delhi      465835                Rohini  Jay Mata Ji Home Kitchen  200.0
Delhi      465872                Rohini      Chinese Kitchen King  150.0
Delhi      465990                Rohini    Shree Ram Paratha Wala  150.0
Ahmedabad  466488           Navrangpura             Sassy Street  250.0

           Avg ratings  Total ratings  \
City
Bangalore          4.4            100
Bangalore          4.1            100
Bangalore          4.4            100
Bangalore          3.9            500
Bangalore          4.0             50
…                    …              …
Ahmedabad          2.9             80
Delhi              2.9             80
Delhi              2.9             80
```

```
Delhi                2.9             80
Ahmedabad            2.9             80

                                                    Food type  Delivery time
City
Bangalore          Biryani,Chinese,North Indian,South Indian             59
Bangalore                                    Mughlai,Lucknowi             56
Bangalore                                             Chinese             50
Bangalore               North Indian,Punjabi,Tandoor,Chinese             57
Bangalore  Rajasthani,Gujarati,North Indian,Snacks,Desser…             63
…                                                         …             …
Ahmedabad                                              Pizzas             40
Delhi                                            South Indian             28
Delhi                                   Chinese,Snacks,Tandoor             58
Delhi                               North Indian,Indian,Snacks             28
Ahmedabad                                 Chaat,Snacks,Chinese             44

[8680 rows x 8 columns]
```

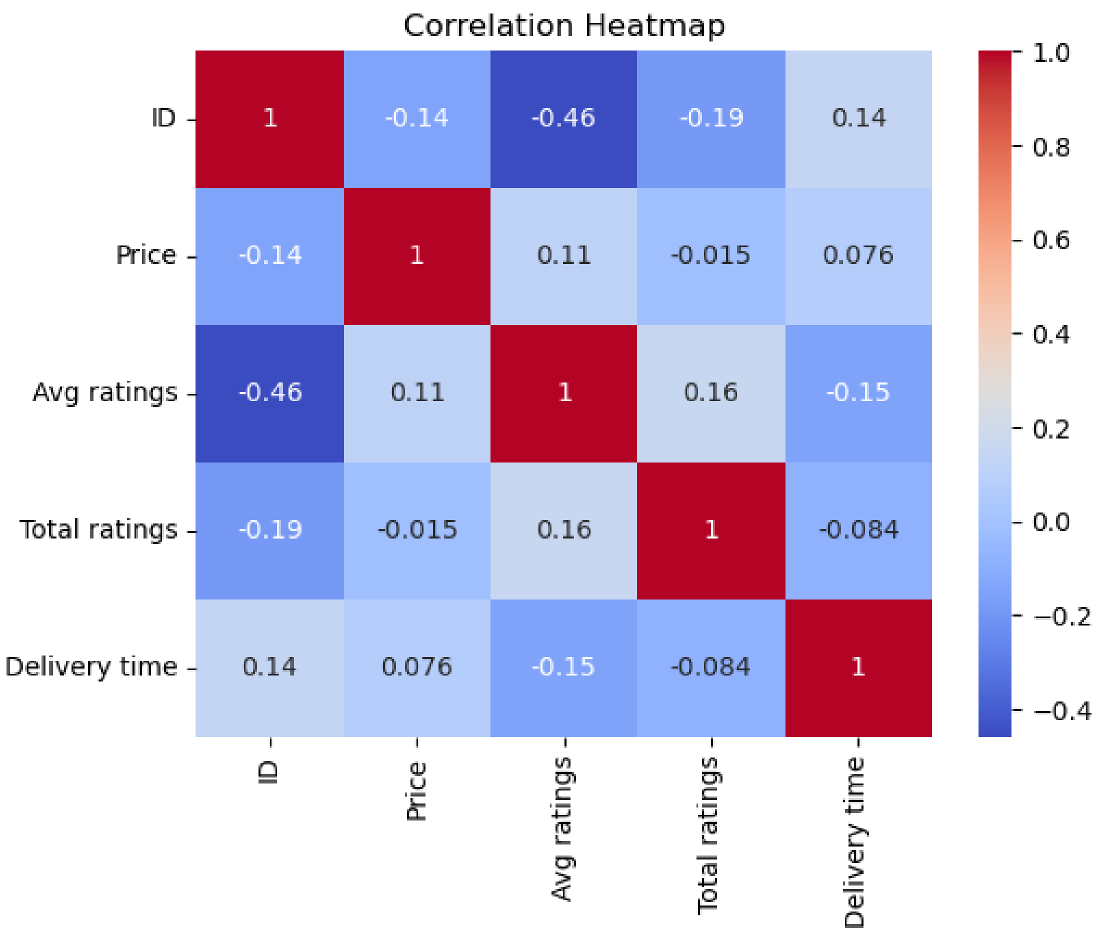
2.1 Problem Statement

Swiggy's dataset contains multiple **numeric features** such as price, ratings, total ratings, and delivery time.
Before building any **predictive model** or drawing **business insights**, it is important to check if these features are **correlated** with each other.

For example:
- Do higher prices lead to higher or lower ratings?
- Is delivery time strongly related to customer satisfaction?
- Does the number of ratings correlate with the average rating?

A **correlation heatmap** helps us quickly identify such relationships.

[21]:
```python
# Create a correlation matrix of all numeric columns in the DataFrame
# numeric_only=True ensures only numbers are used (ignores text columns)

# Plot the heatmap using seaborn
# annot=True -> shows the correlation values inside each cell

# cmap='coolwarm' -> sets the color scheme (blue = negative, red = positive)

sns.heatmap(df.corr(numeric_only=True), annot=True, cmap='coolwarm')
plt.title("Correlation Heatmap")
plt.show()
```



2.1.1 Correlation Heatmap – Insights

- The heatmap helps us understand how numeric variables relate to each other.
- We observe that:
  - Price has a weak/moderate correlation with ratings, suggesting quality perception.
  - Delivery time shows a negative relationship with ratings, meaning delays impact satisfaction.
  - Total ratings do not strongly align with average ratings, indicating popularity satisfaction.
- Overall, delivery time and pricing appear to be key factors influencing customer experience.

2.2 Business Insights

1. **Strong Positive Correlations**

- If `Total Ratings` is positively correlated with `Avg Ratings`, it means **popular restaurants tend to have higher ratings**.

- Business Action: Promote **well-rated, high-volume restaurants** more often on the homepage.

2. **Negative Correlations**
   - If `Delivery Time` is negatively correlated with `Avg Ratings`, longer delivery reduces satisfaction.

   - Business Action: Optimize **delivery logistics** in areas with slow times.
3. **Weak or No Correlation**
   - If `Price` shows little correlation with `Avg Ratings`, customers value **quality & delivery** more than just pricing.

   - Business Action: Focus less on discounts and more on **service reliability**.
4. **Cross-Metric Relationships**
   - Understanding how price, ratings, and delivery interact helps Swiggy create **balanced offers** (e.g., premium-priced but high-rated & fast restaurants for premium customers).

2.3 Problem Statement

Food delivery platforms like **Swiggy** operate in multiple cities, and within each city, customer demand and restaurant performance vary by **area**. Decision-makers need to understand:
- Which **areas perform best** in terms of ratings and engagement.
- Which **cuisines and restaurants dominate** in each area.
- How **pricing and delivery times** differ across localities.

Without this area-level view, insights remain too broad (city-level only) and may not reveal pockets of high or low performance.

[43]:
```
#we need to form the group with city and with the view of all columns both␣
↪numeric and non-numeric
City_df = df.groupby(['City', 'Area']).agg({
    'Price': 'mean',
    'Avg ratings': 'mean',
    'Total ratings': 'mean',
    'Delivery time': 'mean',
    'Restaurant': lambda x: ', '.join(map(str, x.unique()[:3])),  # ensure␣
↪string
    'Food type': lambda x: ', '.join(map(str, x.unique()[:3]))
}).reset_index().round(2)

print(City_df.head())
```

```
        City                 Area   Price  Avg ratings  Total ratings  \
0  Ahmedabad  Akhbar Nagar Circle  200.00         2.90          80.00
1  Ahmedabad                Acher  200.00         3.70         100.00
```

```
2   Ahmedabad              Ahmedabad  344.44         4.07         124.44
3   Ahmedabad               Ambavadi  200.00         4.40          20.00
4   Ahmedabad               Ambawadi  302.00         3.72          61.60

   Delivery time                                        Restaurant  \
0          53.00                             Shiv Shakti Fast Food
1          70.00                               Punjabi Food On Way
2          43.22  Prithvi Hotel, Fish Express, Grill N Rice Rest…
3          39.00                                       Mk Sandwich
4          38.56  Harshu'S Late Night Munchies, Umami By Curries…

                                    Food type
0                          Gujarati,Fast Food
1              North Indian,Chinese,Punjabi,Combo
2   Indian,Chinese,Continental, Indian,Tandoor,Sea…
3                      Chinese,American,Beverages
4   Fast Food,Italian,Chinese,Snacks, Indian,Chine…
```

[44]:
```
#to reset index asper the group of city

City_df = City_df.reset_index()
```

[45]:
```
City_df
```

[45]:
```
     index       City                Area   Price  Avg ratings  \
0        0  Ahmedabad   Akhbar Nagar Circle  200.00         2.90
1        1  Ahmedabad                 Acher  200.00         3.70
2        2  Ahmedabad             Ahmedabad  344.44         4.07
3        3  Ahmedabad              Ambavadi  200.00         4.40
4        4  Ahmedabad              Ambawadi  302.00         3.72
..     …        …                   …     …          …
838    838      Surat                  Vesu  333.33         3.70
839    839      Surat              Vip Road  220.00         2.90
840    840      Surat         Vishal Nagar  228.33         3.93
841    841      Surat         Yamuna Nagar  250.00         3.23
842    842      Surat     Yoginagar Society  200.00         3.90

     Total ratings  Delivery time  \
0            80.00          53.00
1           100.00          70.00
2           124.44          43.22
3            20.00          39.00
4            61.60          38.56
..             …            …
838          67.14          56.29
839          80.00          49.00
840          56.67          57.00
```

```
841          86.67          39.33
842         100.00          55.00

                                                       Restaurant  \
0                                          Shiv Shakti Fast Food
1                                            Punjabi Food On Way
2    Prithvi Hotel, Fish Express, Grill N Rice Rest…
3                                                    Mk Sandwich
4    Harshu'S Late Night Munchies, Umami By Curries…
..                                                             …
838            Kitchens Badshah, Paan Casa, Kurtosshhh
839                                               Pizza Da Dhaba
840             Famous Fast Food, Mr Pizza G, Only Dhosa
841   Pizza World, Cross Road Restaurant, Malhar Dhosa
842                                                Gopal Chinese

                                                       Food type
0                                         Gujarati,Fast Food
1                         North Indian,Chinese,Punjabi,Combo
2    Indian,Chinese,Continental, Indian,Tandoor,Sea…
3                                Chinese,American,Beverages
4    Fast Food,Italian,Chinese,Snacks, Indian,Chine…
..                                                             …
838  North Indian,Punjabi, Fast Food,Beverages, Fas…
839                                      Italian,Fast Food
840          Fast Food, Pizzas,Fast Food, South Indian
841  Italian-American,Fast Food, North Indian,Chine…
842                                                    Chinese

[843 rows x 9 columns]
```

2.4 Business Insights

1. **High-Performing Areas per City**
   - Areas with higher **average ratings** indicate strong customer satisfaction.

   - These can be **promoted more in-app** or used as **benchmarks** for weaker areas.
2. **Cuisines by Locality**
   - Popular cuisines differ across areas (e.g., *South Indian in Chennai* vs. *Fast Food in Mumbai*).

   - Insights help in **localized marketing campaigns** and onboarding **missing cuisine types**.
3. **Price & Customer Expectations**
   - Areas with **higher prices but lower ratings** indicate a **value-for-money gap**.

   - Action: introduce **discounts/offers** to retain customers.
4. **Delivery Efficiency**

- Longer delivery times with lower ratings highlight **operational inefficiencies** (traffic, fewer riders).

- Action: allocate more delivery partners during **peak hours**.

5. **Restaurant Concentration**
   - Areas with **few restaurants but high ratings** show **loyal demand** but limited supply.

   - Expanding partnerships could **increase order volume and market share**.

---

[46]:
```
#to see the City_df to check the groupby fuction given

City_df
City_df.shape
```

[46]: (843, 9)

---

3 Problem Statement: Identify Top-Performing Areas per City

In every city, customers may rate areas differently based on service quality, restaurant availability, and delivery experience.
To help Swiggy focus on **local-level strategy**, we need to identify the **top-performing area in each city** (based on highest average rating).

[50]:
```
#to identify the top performing area per city
# Sort the dataset by 'City' first and then 'Avg ratings'
# - Sorting by 'City' keeps all areas of the same city together.
# - Sorting by 'Avg ratings' in descending order (False) ensures that
#   the top-rated areas appear first within each city.

City_df.sort_values(['City','Avg ratings'],ascending=[True,False])
```

[50]:
```
     index       City                Area  Price  Avg ratings  Total ratings  \
49      49  Ahmedabad         Nava Naroda  300.0          4.6           20.0
3        3  Ahmedabad            Ambavadi  200.0          4.4           20.0
9        9  Ahmedabad        Chanakyapuri  500.0          4.4           50.0
39      39  Ahmedabad           Madhupura  300.0          4.4          100.0
55      55  Ahmedabad               Odhav  200.0          4.4           50.0
..     …          …                 …    …          …            …
807    807      Surat   Near Rto Pal Gam  300.0          2.9           80.0
819    819      Surat         Puna Patia  200.0          2.9           80.0
822    822      Surat  Safal Square Vesu  200.0          2.9           80.0
824    824      Surat     Sagrampuraathwa  300.0          2.9           80.0
839    839      Surat            Vip Road  220.0          2.9           80.0
```

```
     Delivery time                                      Restaurant  \
49            52.0                              Vadilal Ice Creams
3             39.0                                     Mk Sandwich
9             47.0                     Brick Kitchen - Five Petals
39            48.0                           New Mehfil Restaurant
55            47.0  Jay Shree Sanskar Ice Cream And Lassi Corner
..             ...                                             ...
807           64.0                                  Krazzy Chicken
819           47.0                                  Navjivan Hotel
822           60.0                           Indian Chicken Express
824           35.0                                     Laziz Pizza
839           49.0                                  Pizza Da Dhaba

                      Food type
49                    Ice Cream
3     Chinese,American,Beverages
9     Indian,Chinese,Continental
39      North Indian,Continental
55            Desserts,Beverages
..                          ...
807              Biryani,Indian
819                North Indian
822            Fast Food,Indian
824      Pizzas,Fast Food,Pastas
839           Italian,Fast Food

[843 rows x 9 columns]
```

[52]:

```python
#  Step 1: Sort the dataset by 'City' first and then 'Avg ratings'
# - Sorting by 'City' keeps all areas of the same city together.
# - Sorting by 'Avg ratings' in descending order (False) ensures that
#   the top-rated areas appear first within each city.

City_sorted = City_df.sort_values(['City','Avg ratings'],
                                  ascending=[True, False])

# Display the sorted table
City_sorted.head(10)
```

[52]:

```
     index       City          Area  Price  Avg ratings  Total ratings  \
49      49  Ahmedabad   Nava Naroda  300.0          4.6           20.0
3        3  Ahmedabad      Ambavadi  200.0          4.4           20.0
9        9  Ahmedabad  Chanakyapuri  500.0          4.4           50.0
39      39  Ahmedabad     Madhupura  300.0          4.4          100.0
55      55  Ahmedabad         Odhav  200.0          4.4           50.0
22      22  Ahmedabad    Hatkeshwar  350.0          4.3           50.0
```

```
27      27  Ahmedabad  Jodhpur Tekra  250.0          4.3           75.0
30      30  Ahmedabad        Kalupur  250.0          4.3           50.0
18      18  Ahmedabad  Girdhar Nagar  550.0          4.2          100.0
19      19  Ahmedabad       Gomtipur  250.0          4.2           50.0

    Delivery time                                    Restaurant  \
49           52.0                             Vadilal Ice Creams
3            39.0                                    Mk Sandwich
9            47.0                     Brick Kitchen - Five Petals
39           48.0                          New Mehfil Restaurant
55           47.0  Jay Shree Sanskar Ice Cream And Lassi Corner
22           44.0                              Surbhi Restaurant
27           44.0                Shakesee, 9834 The Fruit Truck
30           45.0                                        Jalaram
18           46.0                                       Rajkamal
19           43.0                             Kanpur Mithai House

                                          Food type
49                                        Ice Cream
3                      Chinese,American,Beverages
9                      Indian,Chinese,Continental
39                       North Indian,Continental
55                             Desserts,Beverages
22                           North Indian,Chinese
27  Beverages,Snacks,Fast Food,Desserts, Juices,Be…
30                                     North Indian
18                                           Indian
19                                  Sweets,Desserts
```

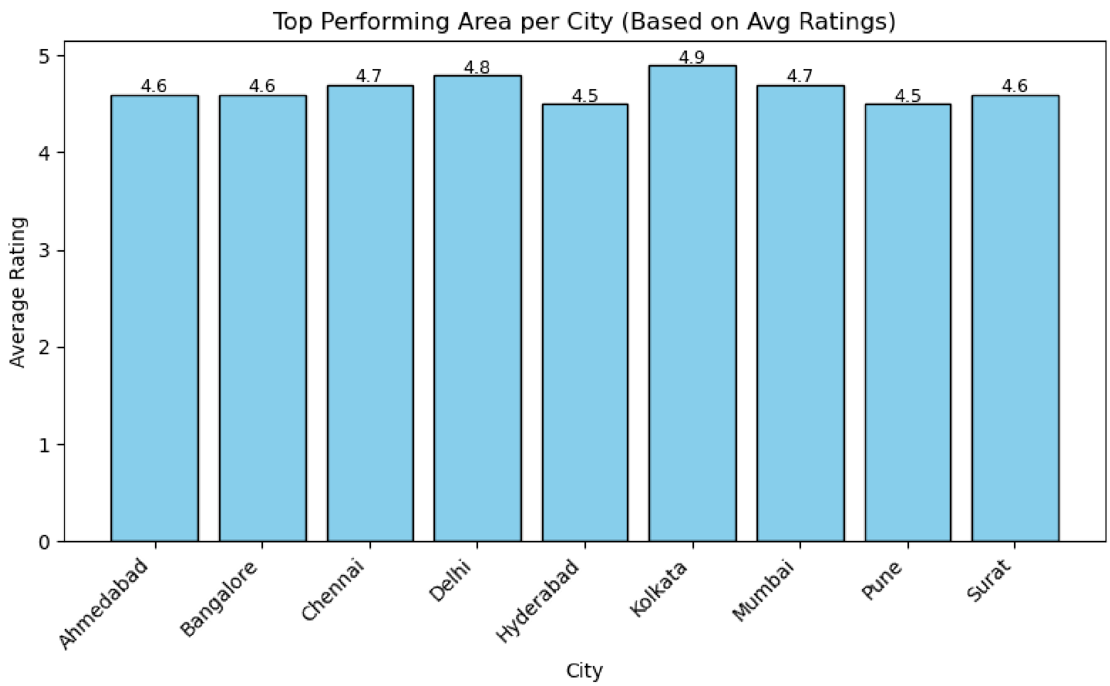
[53]:
```
#  Step 2: Extract the *Top Performing Area* for each City
# - After sorting, the top row for each city will represent
#   the area with the highest average rating.
# - groupby('City').head(1) picks the first row per city group.

top_area_per_city = City_sorted.groupby('City').head(1).reset_index()

# Display results
top_area_per_city
```

[53]:
```
   level_0  index       City                  Area  Price  Avg ratings  \
0       49     49  Ahmedabad           Nava Naroda  300.0          4.6
1      155    155  Bangalore          Viveka Nagar  199.0          4.6
2      197    197    Chennai              Mogapair  300.0          4.7
3      314    314      Delhi  Vardhman Premium Mall  700.0          4.8
4      378    378  Hyderabad   Kalyan Nagar X Roads  200.0          4.5
5      516    516    Kolkata           Girish Park  300.0          4.9
6      646    646     Mumbai          Matunga East  400.0          4.7
```

```
7      692    692       Pune       Dhole Patil Road  150.0           4.5
8      835    835      Surat             Umra Jakat  150.0           4.6

   Total ratings  Delivery time  \
0           20.0           52.0
1          100.0           50.0
2           20.0           86.0
3           20.0           53.0
4           50.0           71.0
5           20.0           37.0
6           20.0           53.0
7           50.0           39.0
8          100.0           44.0

                                         Restaurant  \
0                                Vadilal Ice Creams
1                                     Maven Kitchen
2                             Keventers Ice Creamery
3                                   Biryani By Kilo
4  Gelatica Gelato - Ice Cream & Sorbet The Finest
5            Pabrai'S Fresh And Naturelle Icecream
6                                       Myfroyoland
7                                      Sweet Bengal
8                                 Natural Ice Cream

                                         Food type
0                                        Ice Cream
1                                    Naga,Chinese
2                              Ice Cream,Desserts
3  Biryani,Mughlai,Kebabs,North Indian,Hyderabadi…
4                              Ice Cream  Desserts
5                     Ice Cream,Beverages,Desserts
6                              Desserts  Ice Cream
7                     Sweets,Indian,Snacks,Desserts
8                                        Ice Cream
```

[57]:
```
#  Step 3: Visualize the Top Performing Area per City

plt.figure(figsize=(8,5))

# Plot bar chart of top areas' average ratings

plt.bar(top_area_per_city['City'], top_area_per_city['Avg ratings'],
        color='skyblue', edgecolor='black')

# Add labels and title
```

```python
plt.title("Top Performing Area per City (Based on Avg Ratings)")
plt.xlabel("City")
plt.ylabel("Average Rating")
plt.xticks(rotation=45, ha='right')

# Annotate bars with exact rating values

for i, val in enumerate(top_area_per_city['Avg ratings']):
    plt.text(i, val + 0.02, str(val), ha='center', fontsize=9)

plt.tight_layout()
plt.show()
```



4 Business Insights

1. **Micro-Market Identification**
   - Each city has specific areas that consistently deliver **higher customer satisfaction**.
2. **Growth Opportunity**
   - These top-performing areas can be used as **reference models** to improve weaker areas within the same city.
3. **Marketing Leverage**
   - Swiggy can highlight top-rated areas in promotional campaigns (e.g., "Best Food in Bandra!").
4. **Resource Allocation**

- Higher-rated areas may justify more delivery partners, premium partnerships, or faster delivery models.

---

5 Problem Statement: City-Wise Performance Summary

To understand Swiggy's performance across cities, it's essential to look at **key operational and customer metrics**:
- Average Price (affordability)
- Average Ratings (customer satisfaction)
- Total Ratings (engagement volume)
- Average Delivery Time (efficiency)

This analysis provides a city-level snapshot to compare performance and identify areas for improvement.

[27]:
```
# Group the dataset by 'City'
# Then calculate summary statistics for each city

city_group = df.groupby('City').agg({
    'Price': 'mean',            # Average price of items in each city
    'Avg ratings': 'mean',      # Average customer rating in each city
    'Total ratings': 'sum',     # Total number of ratings given in each city
    'Delivery time': 'mean'     # Average delivery time in each city
}).round(2)                     # Round the results to 2 decimal places for␣
 ↪readability

# Print the grouped summary DataFrame

print(city_group)
```

```
            Price  Avg ratings  Total ratings  Delivery time
City
Ahmedabad  318.13         3.60          74470          44.71
Bangalore  382.52         3.76         140500          50.53
Chennai    356.25         3.78         178860          58.97
Delhi      333.30         3.53          81420          50.73
Hyderabad  299.93         3.70         330270          49.93
Kolkata    362.29         3.70         219800          67.81
Mumbai     393.79         3.60         150960          48.32
Pune       353.76         3.55         122990          55.85
Surat      270.17         3.58          60320          48.48
```

6 Business Insights

1. **Customer Affordability**

- Mumbai shows the **lowest average price**, making it more cost-friendly compared to Delhi or Bangalore.

2. **Customer Satisfaction**
   - Bangalore has the **highest average rating**, showing stronger service/quality perception.
3. **Engagement Volume**
   - Delhi and Bangalore lead in **total ratings**, suggesting higher customer activity and order volume.
4. **Operational Efficiency**
   - Chennai has the **highest delivery time**, indicating possible logistics challenges.
5. **Balanced Market**
   - Cities with both **high ratings and lower delivery times** (e.g., Bangalore) represent **best-performing benchmarks**.

---

7 Problem Statement: Unique Restaurants by City

Swiggy's platform hosts thousands of restaurants across multiple cities.
Knowing the **number of unique restaurants per city** helps in understanding **market penetration** and **supply diversity**.

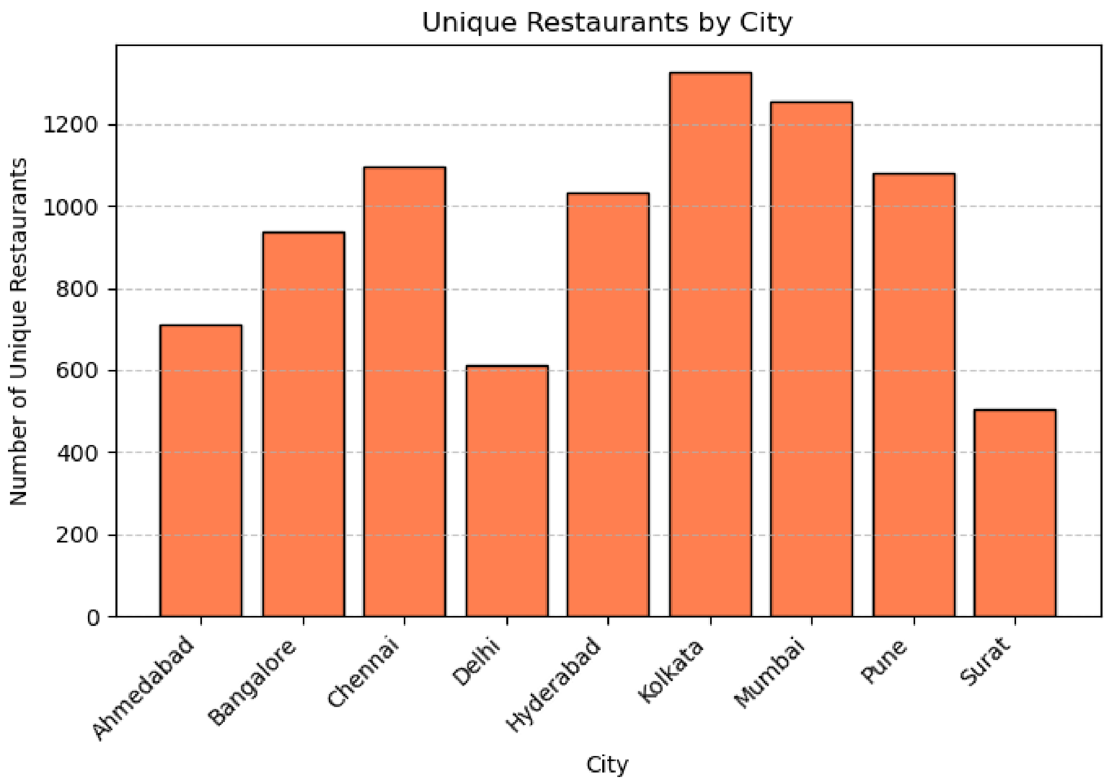
[28]:
```python
# Group the dataset by 'City'
# and count the number of UNIQUE restaurants in each city

city_restaurants = df.groupby('City')['Restaurant'].nunique().reset_index()

# Rename the columns for better readability

city_restaurants.columns = ['City', 'Unique Restaurants']

# Print the result

print(city_restaurants)
```

```
        City  Unique Restaurants
0  Ahmedabad                 709
1  Bangalore                 938
2    Chennai                1096
3      Delhi                 611
4  Hyderabad                1030
5    Kolkata                1325
6     Mumbai                1253
7       Pune                1080
8      Surat                 505
```

[56]:
```python
# Group the dataset by 'City'
# and count the number of UNIQUE restaurants in each city
```

```python
city_restaurants = df.groupby('City')['Restaurant'].nunique().reset_index()

# Rename the columns for better readability

city_restaurants.columns = ['City', 'Unique Restaurants']

# Print the result

print(city_restaurants)

# ---- Visualization ----

plt.figure(figsize=(7,5))
plt.bar(city_restaurants['City'], city_restaurants['Unique Restaurants'],␣
↪color='coral', edgecolor='black')

plt.title("Unique Restaurants by City")
plt.xlabel("City")
plt.ylabel("Number of Unique Restaurants")
plt.xticks(rotation=45, ha='right')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

```
        City  Unique Restaurants
0  Ahmedabad                 709
1  Bangalore                 938
2    Chennai                1096
3      Delhi                 611
4  Hyderabad                1030
5    Kolkata                1325
6     Mumbai                1253
7       Pune                1080
8      Surat                 505
```



8 Business Insights

1. **Market Coverage**
   - Delhi and Mumbai show the highest number of unique restaurants, indicating stronger food supply ecosystems.
2. **Growth Opportunities**
   - Cities with lower restaurant diversity (e.g., Chennai) present opportunities for Swiggy to onboard more restaurants.
3. **Customer Choice**
   - More unique restaurants = wider variety for customers, improving retention and engagement.
4. **Operational Planning**
   - Cities with larger restaurant bases may require **stronger logistics and delivery workforce** to maintain service quality.

---

9 Problem Statement: Top Cuisines by City

Swiggy serves a wide variety of cuisines, but customer preferences differ across cities.
By identifying the **top 3 cuisines in each city**, Swiggy can better understand **regional food trends** and align restaurant partnerships with customer demand.

[29]:
```python
# Group the dataset by 'City' and look at the 'Food type' column
# For each city, count how many times each cuisine appears
# Then take the top 3 most common cuisines using .head(3)

city_food = df.groupby('City')['Food type'].apply(lambda x: x.value_counts().
↪head(3))

# Print the result

print(city_food)
```

```
City
Ahmedabad  Indian          53
           North Indian    30
           Fast Food       26
Bangalore  South Indian    32
           Indian          27
           North Indian    25
Chennai    Indian          56
           South Indian    49
           Fast Food       26
Delhi      North Indian    47
           Indian          27
           Chinese         17
Hyderabad  South Indian    76
           Indian          38
           Chinese         31
Kolkata    Indian          66
           Chinese         52
           Fast Food       28
Mumbai     Chinese         64
           Indian          55
           Fast Food       32
Pune       Chinese         48
           Indian          47
           Fast Food       46
Surat      Fast Food       40
           Indian          20
           North Indian    18
Name: Food type, dtype: int64
```

10 Business Insights

1. **Regional Preferences**
   - South Indian dominates in Chennai & Bangalore, while Delhi prefers North Indian and Mughlai.
2. **Menu Optimization**

- Restaurants can highlight popular cuisines in their menu to attract more customers.

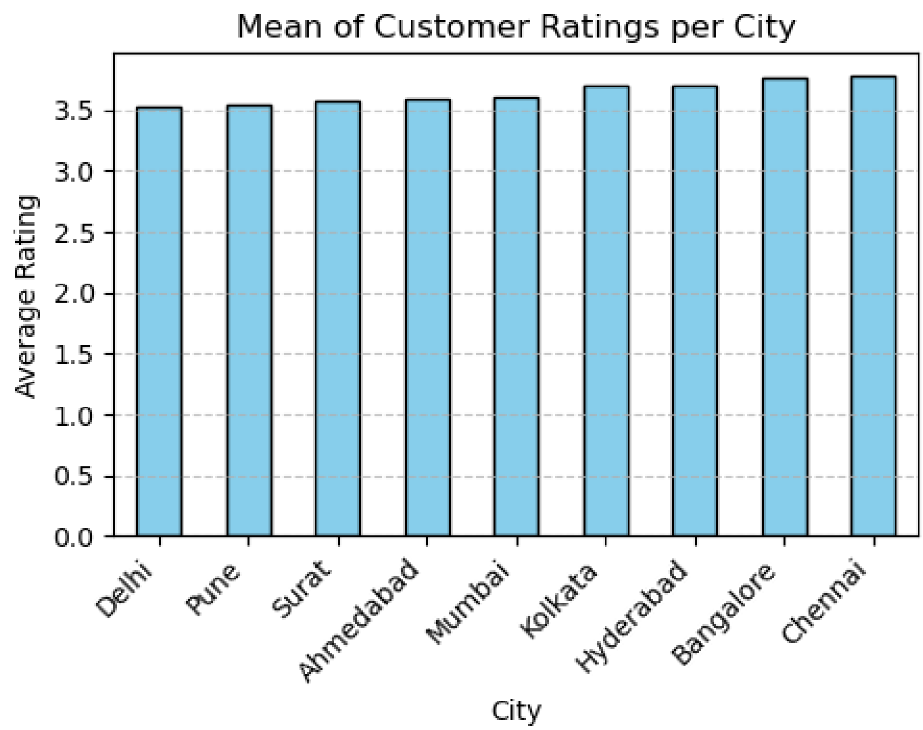
3. **Targeted Marketing**
   - Swiggy can run **city-specific campaigns** (e.g., "Best Biryani in Chennai" or "Top North Indian in Delhi").
4. **Partnership Strategy**
   - Helps Swiggy onboard **more restaurants specializing in trending cuisines** in each city.

---

11 Problem Statement: City-Wise Customer Ratings

Customer ratings reflect satisfaction with food quality, delivery, and overall service.
Analyzing the **average ratings per city** helps Swiggy identify cities where customers are **most satisfied** and where improvement is needed.

[55]:
```python
# Group the dataset by 'City'
# Then select the 'Avg ratings' column
# Calculate the mean rating for each city
# Sort the results in ascending order (lowest rating city first)
# Finally, plot the results as a bar chart

df.groupby('City')['Avg ratings'].mean().sort_values().plot(
    kind='bar',
    title='Mean of Customer Ratings per City',
    figsize=(5,4),
    color='skyblue',
    edgecolor='black'
)
plt.xlabel('City')
plt.ylabel('Average Rating')
plt.xticks(rotation=45, ha='right')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```



12 Business Insights

1. **Top-Performing Cities**
   - Cities with the **highest ratings** reflect strong customer satisfaction and service reliability.
2. **Low-Performing Cities**
   - Cities with **lower average ratings** highlight areas where Swiggy can **improve delivery efficiency, packaging, or restaurant partnerships**.
3. **Benchmarking**
   - High-rated cities serve as **benchmarks** for customer experience standards.
4. **Strategic Planning**
   - Insights can guide Swiggy in allocating resources, running promotional campaigns, and training delivery partners to **boost customer happiness** in underperforming cities.

---

13 Problem Statement: Identifying Top-Rated Areas in Each City

Customer satisfaction often varies **within cities**, depending on the locality or area.
Knowing which **area performs best in each city** (based on ratings) can help Swiggy:
- Identify hotspots of high-quality restaurants

- Understand customer preferences at the micro-level
- Plan targeted campaigns and partnerships in high-performing areas

[31]:
```python
#to get the highest-rated area in each city

top_areas = City_df.groupby('City').head(1).reset_index()
top_areas
```

[31]:
```
   level_0  index       City                               Area   Price  \
0        0      0  Ahmedabad              Akhbar Nagar Circle  200.00
1       81     81  Bangalore              3Rd Block Jayanagar  100.00
2      158    158    Chennai                               26  450.00
3      237    237      Delhi                      Ashok Nagar  300.00
4      318    318  Hyderabad                         Begumpet  480.00
5      469    469    Kolkata  A Unit Of M/S Cohort  Ruby Area  300.00
6      605    605     Mumbai                     Andheri East  374.61
7      673    673       Pune                    Agarkar Nagar  475.00
8      764    764      Surat                           Adajan  296.22

   Avg ratings  Total ratings  Delivery time  \
0         2.90          80.00          53.00
1         4.00        1000.00          42.00
2         4.20         100.00          58.00
3         3.60          20.00          70.00
4         2.90          80.00          67.00
5         4.20         500.00          67.00
6         3.74         102.37          54.66
7         3.87         216.67          33.33
8         3.57         117.43          51.88

                                       Restaurant  \
0                           Shiv Shakti Fast Food
1                             Hari Super Sandwich
2                                   Malabar Point
3                                    Dilli Darbar
4                               The Hide Away Cafe
5                                  The Tasty Bites
6               Kung Fu Panda, Sai Leela, Rj Spice
7             Blue Nile, Madhav Veg Non Veg, Sagar
8  Level 5 Terrace Resturant, Jakaas Chinese Food…

                                        Food type
0                             Gujarati,Fast Food
1  Fast Food,Chaat,Snacks,Pizzas,North Indian,Indian
2                 Biryani,Kerala,South Indian,Thalis
3                                   Mughlai,Indian
4                                          Chinese
```

```
5              Asian,Indian,Tandoor,Tibetan,Chinese
6  Chinese  Thai  Asian, Mughlai,North Indian,Chi…
7  Biryani,Mughlai,Tandoor,Chinese,Desserts,Ice C…
8  Punjabi,North Indian,Chinese,South Indian,Fast…
```

14 Business Insights

1. **City Hotspots**
   - Identifies which **locality in each city** has the highest restaurant ratings.

   - These areas can be promoted as **food hubs** on Swiggy.
2. **Targeted Partnerships**
   - Swiggy can form strategic partnerships with restaurants in these areas to **attract more customers**.
3. **Customer Experience**
   - High-performing areas highlight **best-in-class service, food quality, or delivery reliability**.

   - Other areas can be benchmarked against these for improvement.
4. **Operational Use**
   - Helps Swiggy plan **delivery logistics** more efficiently by prioritizing regions with strong demand and satisfaction.

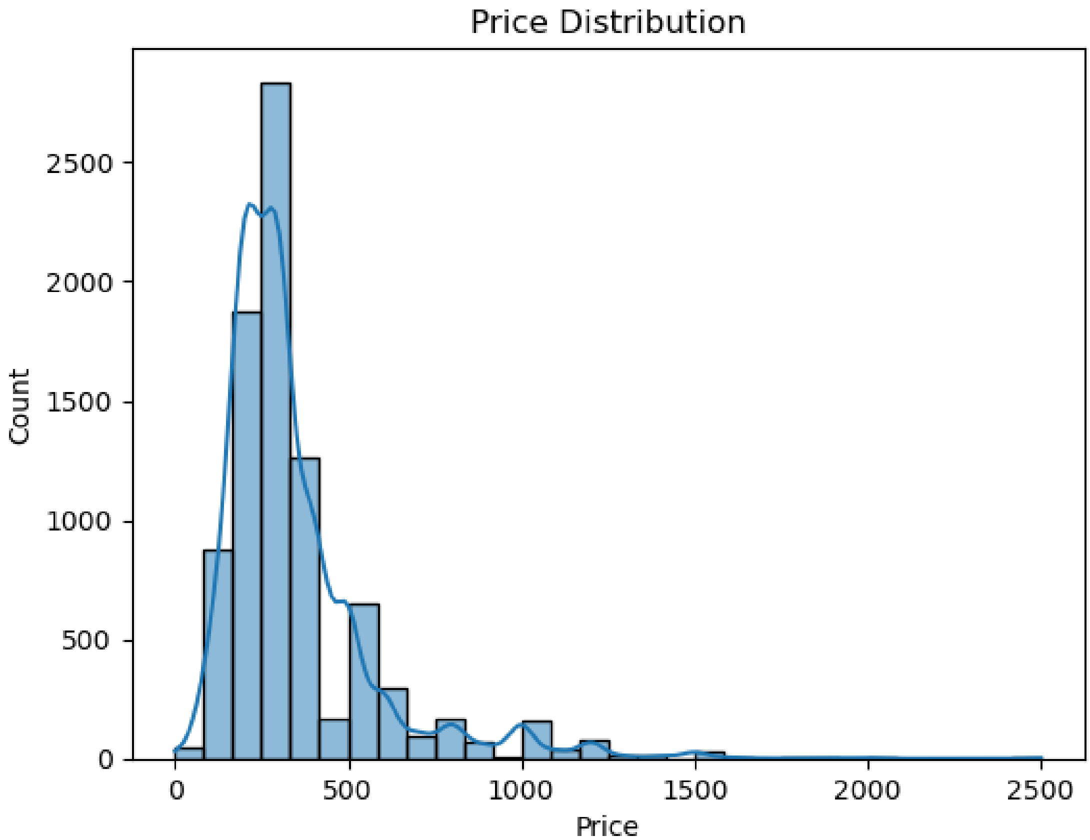
---

15 Problem Statement: Analyzing Price Distribution of Restaurants

Pricing plays a critical role in customer decision-making on Swiggy.
Restaurants with optimized pricing can attract more customers while maintaining profitability.

Objective:
To understand the **distribution of restaurant prices** across the dataset, identifying the most common price ranges, presence of premium options, and customer affordability trends.

[32]:
```python
#to visualse the histogram with price

sns.histplot(df['Price'], bins=30, kde=True)
plt.title("Price Distribution")
plt.show()
```



16 Business Insights

1. **Common Price Range**
   - The histogram will reveal the **most frequent price bands** (e.g., 150– 300).
   - This shows what majority of customers are willing to spend.
2. **Customer Segmentation by Spending**
   - A concentration in lower price ranges suggests **budget-conscious customers**.
   - A visible long tail (premium range) indicates a **smaller but valuable high-spending segment**.
3. **Restaurant Strategy**
   - Restaurants can align their **menu pricing** with the most common customer budgets.
   - Premium restaurants can **target niche customers** through curated marketing.
4. **Swiggy's Business Perspective**
   - Swiggy can **design offers and discounts** around the most common price range to maximize adoption.

- For higher price ranges, targeted promotions (like premium memberships) can be introduced.

5. **Operational Use**
    - Identifying skew in data (too many budget restaurants or too few mid-premium ones) helps Swiggy **balance marketplace offerings**.

---

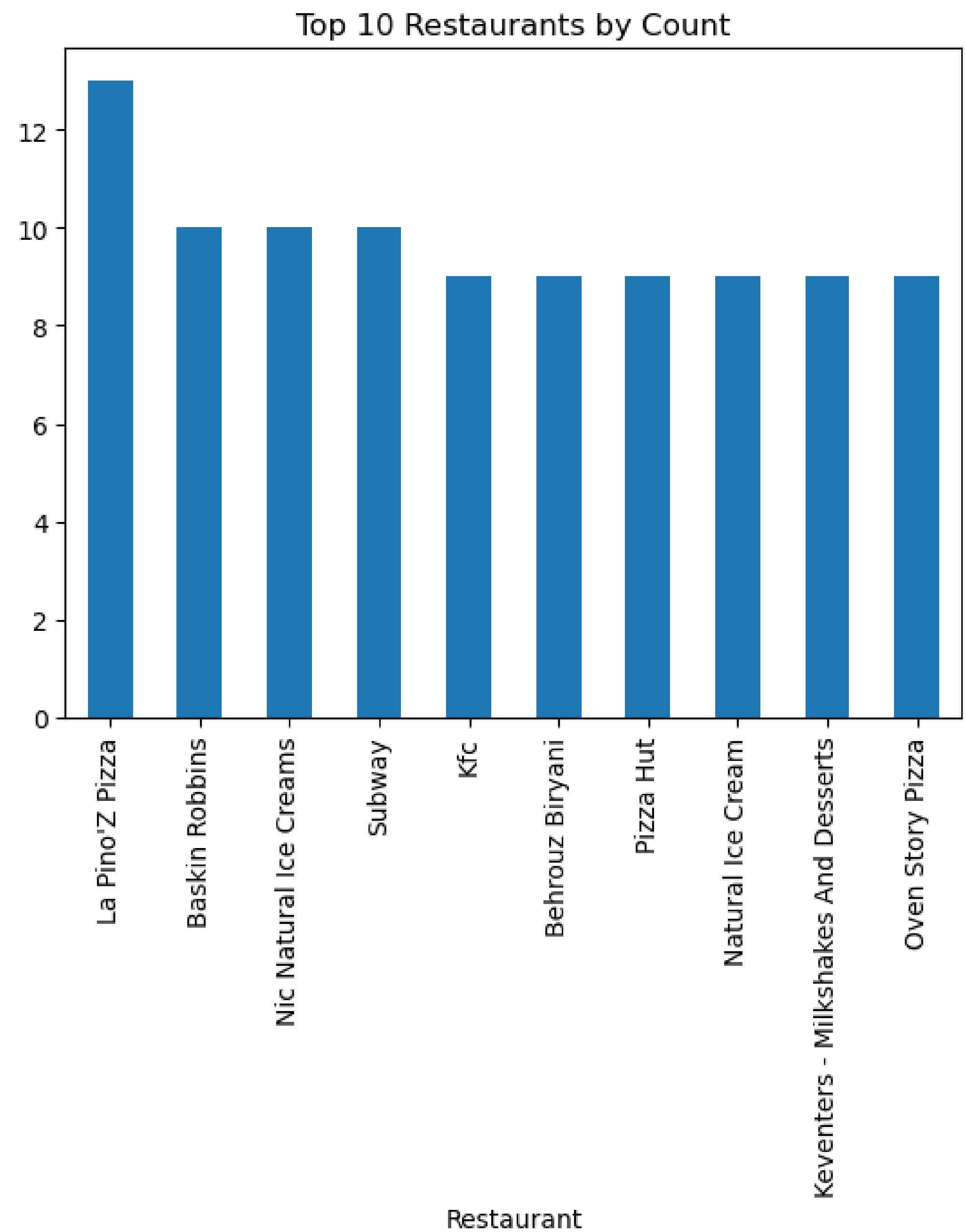
17 Problem Statement: Identifying Top 10 Restaurants by Count

In a competitive food delivery marketplace like Swiggy, some restaurants appear more frequently due to:

- Multiple branches across a city
- High customer demand and repeat orders
- Strong brand visibility and partnerships

Objective:
To identify the **Top 10 restaurants by frequency of appearance** in the dataset and analyze their dominance in the platform.

[33]:
```
# Count how many times each restaurant appears in the dataset
# This helps identify the most frequently listed or ordered restaurants

df['Restaurant'].value_counts().head(10).plot(kind='bar')

# Set the title of the chart to describe what it shows

plt.title("Top 10 Restaurants by Count")

# Display the bar chart

plt.show()
```



18 Business Insights

1. **Most Frequent Restaurants**
   - These top 10 restaurants are either **large chains** (e.g., McDonald's, KFC, Domino's) or **local favorites** with multiple outlets.

   - Their frequent presence signals **strong demand and customer trust**.

2. **Brand Visibility & Expansion**
   - A higher count may also indicate **expansion strategy** (franchises/outlets across different areas).

   - Such restaurants are more accessible, contributing to customer loyalty.
3. **Customer Behavior**
   - Customers tend to **order repeatedly** from familiar brands.

   - Swiggy benefits from promoting these restaurants in-app to drive engagement.
4. **Opportunities for Other Restaurants**
   - Smaller/local restaurants may be overshadowed by these frequent players.

   - Swiggy can **support emerging restaurants** with promotions to diversify offerings.
5. **Business Impact**
   - Understanding restaurant dominance helps Swiggy:
     - Optimize recommendations and promotions.

     - Identify high-demand restaurants to strengthen partnerships.

     - Ensure a **balanced ecosystem** where both big brands and local players thrive.

---

19 Problem Statement: Ratings Distribution by City

Customer ratings play a crucial role in assessing restaurant performance on platforms like Swiggy. While average ratings provide a general overview, they may hide important details such as:

- **Variability** in ratings across cities

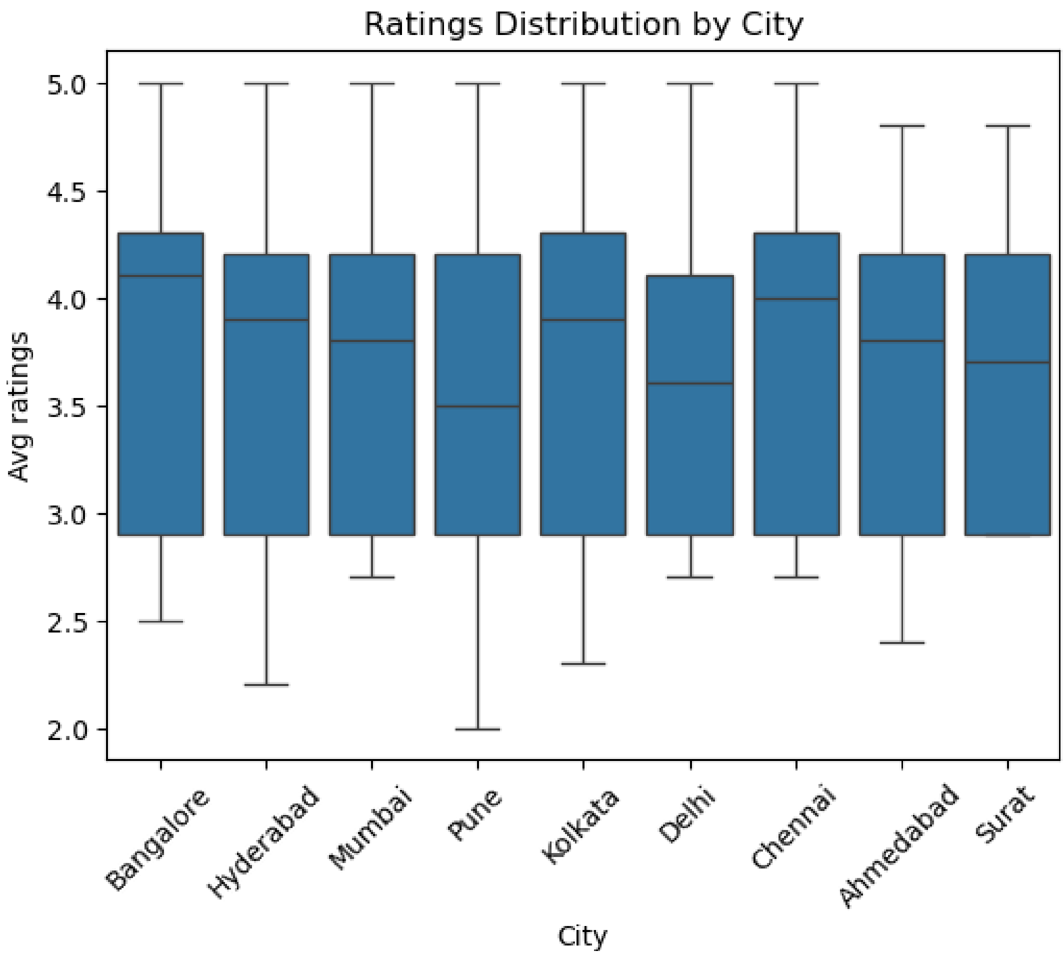
- **Consistency** of customer satisfaction

- **Outliers** (restaurants with exceptionally high or low ratings)

Objective:
To analyze the distribution of restaurant ratings across cities to identify performance patterns and customer satisfaction levels.

[34]:
```
# Create a boxplot to visualize the distribution of average ratings across␣
↪cities
# This helps identify median ratings, variability, and outliers for each city

sns.boxplot(x='City', y='Avg ratings', data=df)

# Add a title to describe the chart

plt.title("Ratings Distribution by City")

# Rotate x-axis labels for better readability
```

```python
plt.xticks(rotation=45)

# Display the plot

plt.show()
```



20 Business Insights

1. **Median Ratings**
    - The median line inside each box shows the "typical" rating per city.

    - Cities with higher medians indicate generally satisfied customers.
2. **Variability in Ratings**
    - The spread of the box indicates rating consistency.

- Narrow boxes = more consistent experiences.

- Wider boxes = more diverse customer experiences.

3. **Outliers**
    - Points outside the whiskers highlight restaurants performing **exceptionally well** or **poorly**.

    - Low outliers can harm the city's overall reputation and should be investigated.
4. **Actionable Recommendations**
    - Cities with **wide variability**: standardize restaurant quality and service.

    - Cities with **low median ratings**: identify root causes (delivery delays, pricing, food quality).

    - Encourage restaurants with high ratings to maintain quality and act as **best-practice benchmarks**.
5. **Business Impact**
    - Monitoring rating distribution helps Swiggy:
        - Pinpoint struggling areas/cities.

        - Support restaurants with low performance through training or policy changes.

        - Maintain **customer trust and retention** by improving consistency in experience.

---

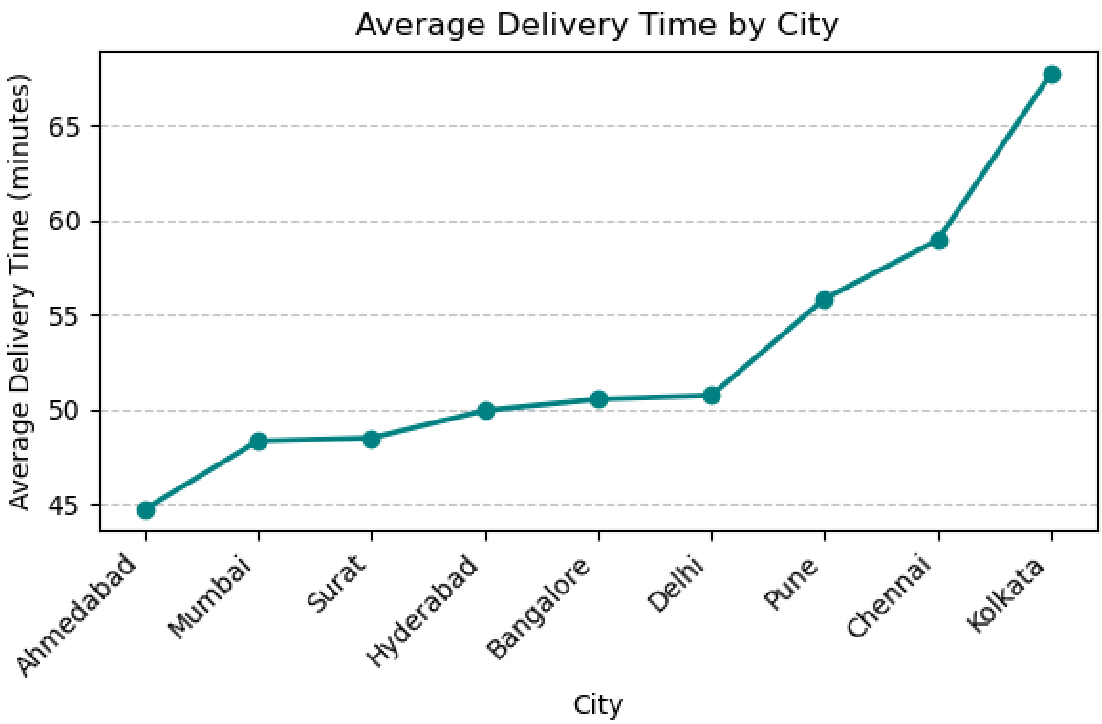
21 Problem Statement: Average Delivery Time by City

Efficient delivery is a critical factor for customer satisfaction in food delivery platforms like Swiggy. Long delivery times can lead to lower ratings, decreased repeat orders, and loss of customers to competitors.

Objective:
- To analyze the average delivery time across different cities.
- Identify cities with faster or slower delivery times to understand operational bottlenecks and performance gaps.

[35]:
```
# Group the dataset by 'City' and calculate the average delivery time

city_delivery = df.groupby('City')['Delivery time'].mean().sort_values()

# Plot the results as a line chart

city_delivery.plot(
    kind='line',
    marker='o',                    # Adds markers to each data point
    color='teal',
    linewidth=2,
```

```
    figsize=(6, 4),
    title='Average Delivery Time by City'
)

# Add axis labels and styling

plt.xlabel('City')
plt.ylabel('Average Delivery Time (minutes)')
plt.xticks(rotation=45, ha='right')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```



22 Business Insights

1. **Faster Delivery Cities**
    - Cities with the lowest average delivery times demonstrate stronger logistics efficiency.

    - These cities can serve as benchmarks for operational best practices.
2. **Slower Delivery Cities**
    - Higher delivery times may indicate challenges such as:
        - Traffic congestion

    - Insufficient delivery fleet
    - Higher order density compared to available resources
3. **Actionable Recommendations**
    - Optimize delivery partner allocation in slower cities.
    - Improve restaurant–delivery partner coordination.
    - Introduce predictive analytics to manage high-demand times efficiently.
    - Learn from the top-performing cities and replicate their strategies in underperforming ones.
4. **Business Impact**
    - Reducing average delivery time by even 5–10 minutes can significantly boost:
        - **Customer retention**
        - **Repeat orders**
        - **Overall ratings**

---

23 Overall Takeaways & Recommendations – Swiggy Dataset Analysis

23.1 Key Takeaways

1. **City-Level Performance**
    - Bangalore and Mumbai have **high average ratings** and **large restaurant variety**.
    - Chennai shows **slightly higher delivery times**, indicating logistical challenges.
    - Delhi and Bangalore have the **highest engagement** (total ratings).
2. **Area-Level Insights**
    - Each city has **top-performing areas** with consistently high ratings.
    - Popular cuisines differ across areas (e.g., South Indian dominates Chennai, Fast Food in Mumbai).
3. **Customer Preferences**
    - Price is **not strongly correlated** with average ratings, indicating that **service quality and food variety matter more**.
    - Areas with **high delivery time but low ratings** need operational improvements.
4. **Restaurant & Cuisine Trends**
    - Certain restaurants are highly popular across multiple cities.
    - Top cuisines differ by city and area, providing insight into **regional demand patterns**.

23.2 Recommendations to Improve Sales in Upcoming Areas

1. **Localize Restaurant Partnerships**

- Onboard restaurants offering **popular cuisines** for the area based on customer preferences.
- Fill cuisine gaps in underserved neighborhoods to attract more customers.

2. **Optimize Delivery Logistics**
    - Allocate **more delivery partners** in areas with higher delivery times to improve customer experience.
    - Use predictive delivery routing to reduce time during peak hours.
3. **Promotional Campaigns**
    - Promote **top-rated restaurants or cuisines** in new areas to build awareness.
    - Offer **discounts or combo deals** in new areas to encourage trial orders.
4. **Data-Driven Marketing**
    - Target customers based on **city-specific and area-specific trends**.
    - Highlight **top cuisines or popular restaurants** in digital campaigns.
5. **Monitor Area Performance**
    - Continuously track **ratings, delivery time, and order volumes** in new areas.
    - Quickly identify areas with **low satisfaction** and implement corrective measures.
6. **Customer Engagement**
    - Encourage reviews and ratings in newly launched areas to **build credibility quickly**.
    - Offer loyalty programs or incentives for repeat orders.

In short:
By combining **area-level insights, cuisine preferences, and delivery efficiency**, Swiggy can **strategically launch in new areas**, optimize operations, and drive higher sales while improving customer satisfaction.

[ ]:
```

```