

in 5 to 6 lines

Problem Statement:

Understand the distribution of total ratings to identify the most and least common rating values.

Analysis:

- **80.0 Rating** has the highest frequency, indicating it is the most common rating.
- **100.0 Rating** follows with the second highest frequency.
- **20.0 Rating** has a moderate frequency.
- **50.0 Rating** is the least common among the four.

This suggests a significant clustering around the 80.0 rating, highlighting potential biases or

Understand the distribution of delivery times to identify the most frequent delivery time intervals.

Analysis:

- The highest frequency of delivery times is around 700, occurring at approximately 50 units.
- The distribution is roughly symmetric with a peak around the 50-60 unit range.
- Delivery times taper off towards both ends of the range (20-90 units).
- There is a clear central peak, indicating most deliveries are completed within the 50-60 unit time frame.
- This pattern suggests a consistent delivery performance with some variability.

This insight can help in understanding delivery efficiency and planning improvements.

p

Understand the distribution of orders across different cities to identify the city with the highest and lowest frequency of orders.

Analysis:

- **Mumbai** has the highest frequency of orders, accounting for 15.4% of the total.
- **Kolkata** follows closely with 15.1%.
- **Pune** and **Chennai** also have significant percentages at 13.3% and 12.5%, respectively.
- **Surat** has the lowest frequency, representing 6.2% of the total orders.

This analysis indicates that Mumbai and Kolkata are key hubs for orders, while Surat has the least activity. This can help in strategic planning and resource allocation.

Understand the distribution of average ratings to identify any significant trends or clustering.

Analysis:

- **3.0 Average Rating:** Has the highest frequency, with counts exceeding 3000. This suggests that a large portion of ratings cluster around the 3.0 mark.
- **4.0 to 4.5 Average Rating:** Shows another peak with counts up to approximately 500, indicating a smaller but notable clustering in this range.
- The data points are scattered across the 2.0 to 5.0 average ratings, showing varying counts at different average ratings.
- The notable peaks at 3.0 and 4.0-4.5 suggest these ratings are more common compared to others.

This insight helps in understanding how users are rating on average, highlighting areas of high and low frequency in the rating distribution.

Understand the distribution of price and delivery time densities to identify key trends.

Analysis:

- The **price distribution** (blue) peaks around 600, indicating most items are priced around this value.
- The **delivery time distribution** (orange) peaks near 0, suggesting most deliveries occur almost immediately or very quickly.
- The density plot reveals two distinct distributions, showing that pricing and delivery times follow different patterns.
- There is a notable density at lower prices and faster delivery times, indicating efficiency.
- The higher price ranges have lower densities, reflecting fewer occurrences.

This KDE plot helps visualize and compare the distribution patterns of prices and delivery times effectively.

Understand the distribution of total ratings to identify the most common and significant values.

Analysis:

- The highest peak occurs around **75 total ratings**, indicating this value has the highest density.
- Other notable peaks are around **25**, **50**, and **100 total ratings**, showing these values are also relatively common.
- The density plot demonstrates the spread and frequency of total ratings, with multiple peaks indicating variability in the data.
- The overall distribution suggests that total ratings are not uniformly distributed and tend to cluster around specific values.

This insight helps in identifying key trends and patterns in the distribution of total ratings.

Problem Statement:

Understand the relationship between price and average ratings across different cities.

Analysis:

- The scatter plot shows that **higher prices tend to be associated with higher average ratings**.
- **Mumbai** and **Bangalore** have clusters of high-price, high-rating points.
- **Kolkata** and **Surat** generally show lower prices and average ratings.
- Different color-coded dots represent different cities, providing a clear comparison.
- The overall trend suggests that **price positively correlates with average ratings** across the cities analyzed.

This visualization highlights how pricing strategies might influence customer satisfaction in various urban markets

Problem Statement:

Understand how average ratings correlate with delivery time.

Analysis:

- **Shorter Delivery Times (22 to 40 units):** Generally have higher average ratings, indicating greater customer satisfaction.
- **Mid-range Delivery Times (40 to 70 units):** Show a decline in average ratings, with some fluctuations.
- **Longer Delivery Times (70 to 94 units):** Continue to have lower average ratings, suggesting decreasing satisfaction with increased delivery time.
- **Trend:** Quicker deliveries are associated with higher customer satisfaction.

Problem Statement:

Understand the relationship between delivery time and price across different cities.

Analysis:

- **Delivery Time vs Price:** The scatter plot shows the relationship between delivery time (x-axis) and price (y-axis) across various cities.
- **Color Coding:** Each city is represented by a different color, helping to differentiate data points.
- **Trend:** Higher prices are generally associated with longer delivery times, although there are some variations.

- **Clusters:** Cities like **Mumbai** and **Delhi** show a wide range of prices and delivery times, indicating diverse delivery scenarios.

This analysis can help identify patterns in delivery efficiency and pricing strategies across different urban markets.

Problem Statement:

Understand the variation of average ratings based on the total number of ratings.

Analysis:

- **Data Points:** Total ratings are shown at approximately 20, 50, 80, and 100, with corresponding average ratings.
- **Significant Drop:** Average ratings drop sharply to around 2.8 at 80 total ratings, indicating a dip in user satisfaction.
- **Sharp Increase:** At 100 total ratings, the average rating jumps to around 4.2, showing improved user satisfaction.
- **Trend:** There is a variability in the data, with some confidence intervals depicted by the shaded area.

This graph highlights critical points where average ratings significantly change, providing insights into user satisfaction trends at different rating levels.

Problem Statement:

Understand the average price distribution across various areas.

Analysis:

- **Koramangala** shows the highest average price, nearing 800 units.
- **Jogupalya**, **Indiranagar**, and **Domlur** also have higher average prices, indicating premium locations.
- **Film Nagar** and **Banashankari** fall in the mid-range average prices.
- **Andheri East** and **Powai** also display moderate prices with some variability.
- **Shivajinagar** has the lowest average price, close to 100 units.

This bar chart helps compare average prices by area, identifying the most and least expensive locations.

Problem Statement:

Understand the distribution of prices across various cities to identify trends and outliers.

Analysis:

- **Bangalore:** Shows the widest range of prices with significant outliers, indicating high price variability.
- **Hyderabad:** Has a narrower range, suggesting more consistent pricing.
- **Mumbai:** Displays high median prices with several outliers above the interquartile range (IQR).
- **Surat:** Has the lowest median price and least variability.
- **Pune:** Shows moderate median prices with fewer outliers compared to Bangalore and Mumbai.

This box plot helps compare the price distribution across cities, highlighting variations, trends, and outliers in each city's pricing.

Problem Statement:

Understand the distribution of delivery times across various cities to identify trends and outliers.

Analysis:

- **Bangalore:** Displays a median delivery time around 50 units, with a few higher outliers indicating variability.
- **Hyderabad:** Shows lower median delivery times compared to Bangalore, suggesting more efficient deliveries.
- **Mumbai:** Exhibits a broader range of delivery times, indicating significant variability.
- **Pune:** Has a consistent delivery time distribution, with fewer outliers.
- **Kolkata:** Shows a relatively narrow range of delivery times, with a median slightly higher than Hyderabad.
- **Delhi:** Displays moderate variability in delivery times, similar to Mumbai.
- **Chennai:** Shows a consistent delivery time range with a median close to Hyderabad.
- **Ahmedabad and Surat:** Both have lower median delivery times, indicating efficient deliveries.

This box plot helps compare delivery performance across different cities, highlighting variations in delivery times and identifying cities with more consistent or variable delivery efficiency.