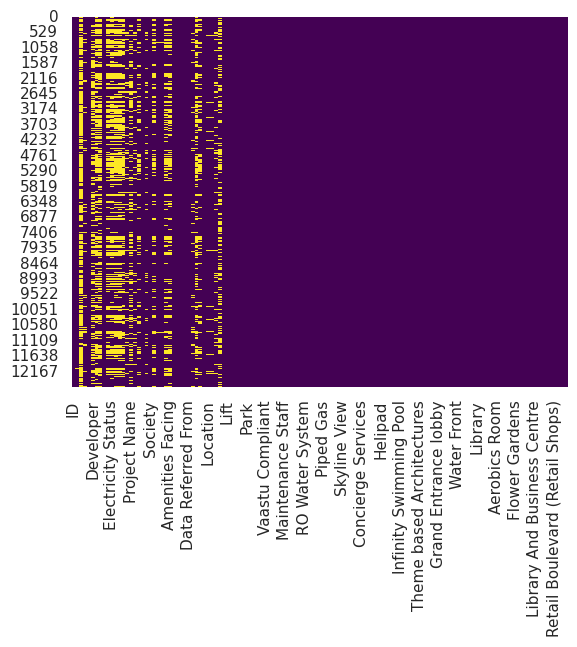
**Assignment 1- Report**

**Data Loading and Initial Cleaning:**

* The code starts by reading an Excel file named 'properties.xlsx' into a pandas DataFrame.
* It then removes several columns that are deemed unnecessary for the analysis or have too many NA values.
* Some column names are renamed for clarity or consistency.
* The code displays information about the DataFrame, including data types and non-null counts.
* A heatmap is created to visualize missing values in the dataset.
* 

Data Transformation:

* For columns containing only 'N' and 'Y' values, these are replaced with 0 and 1 respectively.
* The 'Floor No' column is transformed, mapping specific floor names to numeric values (e.g., 'Ground' to 0, 'Lower Basement' to -1).
* For 'Developer' and 'Project Name' columns, only the first word is kept (ignoring 'The' if it's the first word).
* The 'Possession Status' column is simplified, replacing some values with more general categories.

Price Data Cleaning:

* Rows where the price is listed as 'Call for Price' are removed.
* The 'Price' column is converted from a string format (e.g., "5 Cr" or "50 Lac") to an integer value in rupees.

Date Conversion:

* The 'Availability Starts From' column is converted to datetime format.

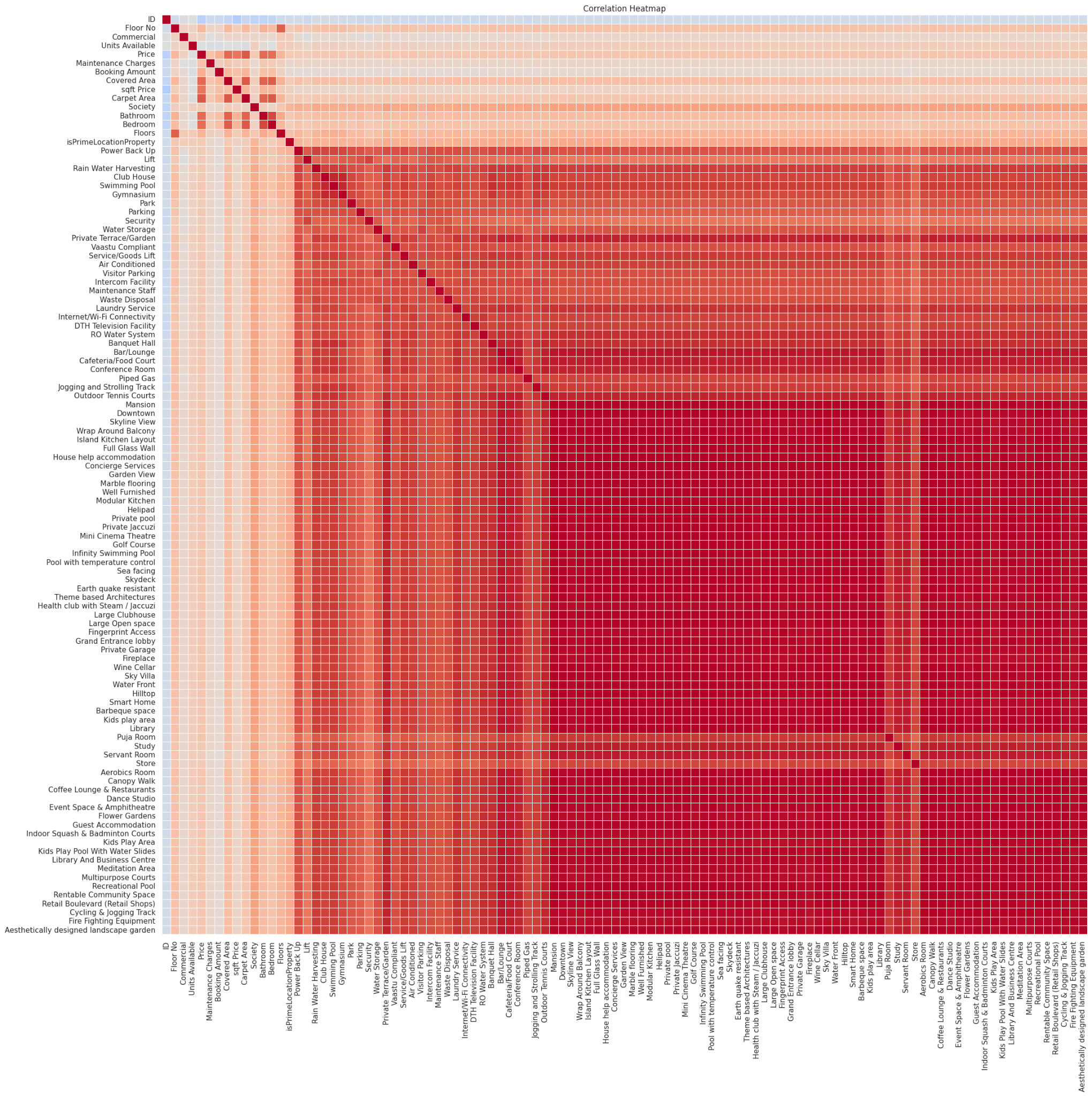
Handling Missing Values:

* Missing values are filled using both forward and backward fill methods.

Outlier Removal:

* The code defines a function to remove outliers using the Z-score method.
* This function is applied to several numeric columns with different threshold values.

Correlation Analysis:

* A correlation matrix is computed for all numeric columns.
* This correlation matrix is visualized using a heatmap.
* 

Data Export:

* The final cleaned DataFrame is saved to an Excel file named 'clean.xlsx'.
* The cleaned data is then read back into a new DataFrame.

**Task-1:**

Distribution of Real Estate Properties by Price Range Across Cities

1. Overview: The analysis categorizes real estate properties into three price ranges (Low, Medium, and High) based on the 33rd and 66th percentiles of the price distribution. This segmentation allows for a comparison of property distributions across different cities, with a focus on Mumbai and Thane, and a separate analysis for other cities.
2. Mumbai and Thane Analysis:

Mumbai:

* Total properties: 8,174
* Low-priced: 1,974 (24.2%)
* Medium-priced: 2,491 (30.5%)
* High-priced: 3,709 (45.3%)

Thane:

* Total properties: 3,989
* Low-priced: 2,029 (50.9%)
* Medium-priced: 1,523 (38.2%)
* High-priced: 437 (10.9%)

Observations: a) Mumbai has a significantly larger real estate market compared to Thane, with more than twice the number of properties. b) Mumbai's market is skewed towards higher-priced properties, with nearly half of its properties in the high price range. c) Thane's market is dominated by low and medium-priced properties, with only about 11% in the high price range.

1. Other Cities Analysis:

* Hyderabad: 8 properties (all low-priced)
* Kalyan: 9 properties (7 low-priced, 2 medium-priced)
* Nagpur: 7 properties (5 low-priced, 2 medium-priced)
* Bhiwandi: 2 properties (all low-priced)
* Palghar: 2 properties (all low-priced)
* Agartala: 1 property (low-priced)
* Gurgaon: 1 property (low-priced)

Observations: a) The data for other cities is limited compared to Mumbai and Thane. b) Most properties in other cities fall into the low price range. c) Only Kalyan and Nagpur have some medium-priced properties. d) No high-priced properties are recorded in cities other than Mumbai and Thane.

1. Conclusions:

a) Market Concentration: The real estate market is heavily concentrated in Mumbai and Thane, with these two cities accounting for the vast majority of properties in the dataset.

b) Price Range Distribution:

* Mumbai offers a diverse range of properties across all price segments, with a slight bias towards high-priced properties.
* Thane presents more affordable options, with a focus on low and medium-priced properties.
* Other cities primarily offer low-priced properties, suggesting they might be more suitable for budget-conscious investors.

c) Investment Opportunities:

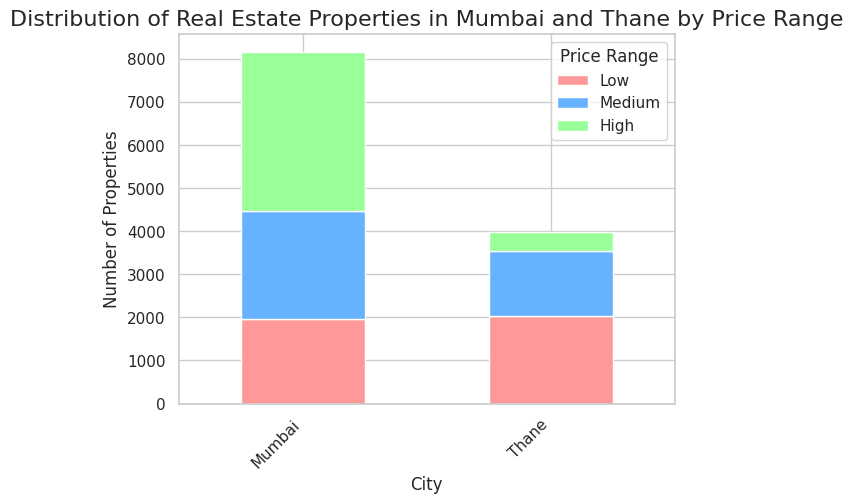
* High-budget investors may find more options in Mumbai.
* Mid-range investors could consider both Mumbai and Thane.
* Budget investors have options across all cities, with Thane offering a large number of low-priced properties.

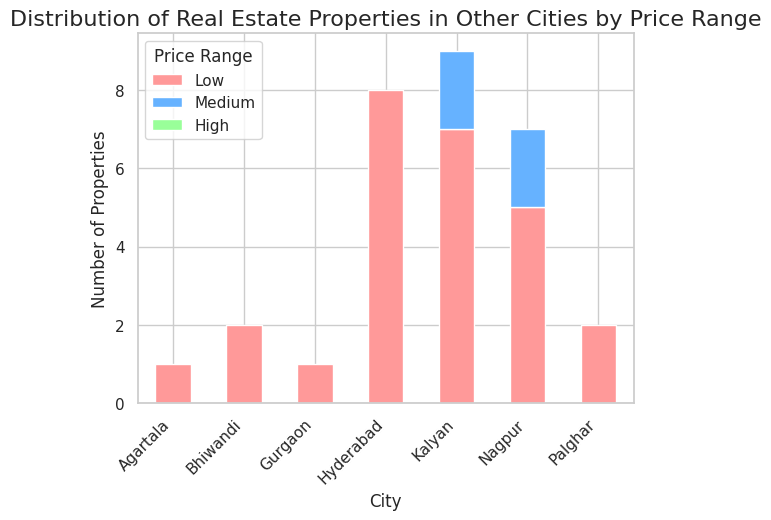
d) Market Maturity: The diversity and volume of properties in Mumbai suggest a more mature and developed real estate market compared to other cities.

e) Data Limitations: The analysis of cities other than Mumbai and Thane is limited due to the small number of properties recorded. This may not provide a comprehensive view of their real estate markets.

1. Recommendations:

* For a comprehensive investment strategy, consider properties in both Mumbai and Thane to balance high-end and affordable options.
* Further investigation into the other cities' real estate markets may be necessary to gain a more accurate picture of their potential.
* The concentration of properties in Mumbai and Thane suggests these areas may offer more liquidity for investors looking to buy or sell properties.





**Task-2:**

Real Estate Investment Summary by Price Range

1. Overview: This analysis covers real estate properties in multiple cities, categorized into Low, Medium, and High price ranges. The data focuses on total amenities offered by different property types across these price ranges in various cities.
2. City-wise Analysis:

a) Mumbai:

* Offers the most diverse range of property types and price ranges.
* Apartments dominate across all price ranges, with high-end apartments offering the most amenities.
* Residential houses show a wide range of amenities, particularly in medium and high price ranges.
* Penthouses and villas are primarily in the high price range with extensive amenities.

b) Thane:

* Second most diverse market after Mumbai.
* Apartments are prevalent across all price ranges, with high-end apartments offering the most amenities.
* Residential houses and villas show a good spread of amenities in medium and high price ranges.
* Builder floor apartments are mainly in the low to medium price range with fewer amenities.

c) Kalyan:

* Limited to apartments only, primarily in the low to medium price range.
* Fewer amenities compared to Mumbai and Thane, even in the medium price range.

d) Palghar:

* Very limited data, showing only low-priced apartments with minimal amenities.

e) Nagpur:

* Focused on apartments in low to medium price ranges.
* Shows a wide variability in amenities, particularly in the low price range.

1. Property Type Analysis:

a) Apartments:

* Most common property type across all cities and price ranges.
* Amenities increase significantly from low to high price ranges.

b) Residential Houses:

* Prominent in Mumbai and Thane, offering a wide range of amenities, especially in medium and high price ranges.

c) Villas:

* Found mainly in Mumbai and Thane in the high price range, offering extensive amenities.

d) Penthouses:

* Exclusive to Mumbai, primarily in the high price range with the most comprehensive amenity packages.

e) Builder Floor Apartments:

* Seen in Thane, mainly in low to medium price ranges with fewer amenities.

1. Amenity Trends:

* Low Price Range: Basic amenities like power backup, lift, parking, security, and water storage are common.
* Medium Price Range: Adds amenities like swimming pools, gyms, and better furnishings to the basic package.
* High Price Range: Offers luxury amenities such as club houses, smart home features, and premium finishes.

1. Investment Insights:

a) Budget Investors:

* Focus on Thane and Kalyan for affordable apartments with basic amenities.
* Some opportunities in low-priced Mumbai apartments, but with fewer amenities.

b) Mid-range Investors:

* Thane offers a good balance of amenities and affordability in medium-priced properties.
* Mumbai provides diverse options in this range with more amenities.

c) Luxury Investors:

* Mumbai is the primary market for high-end properties with extensive amenities.
* Thane also offers some high-end options, particularly in apartments and villas.

1. Presentation Method:

To effectively present this information to investors, I recommend using a combination of:

a) Interactive Dashboard:

* Create an interactive dashboard using tools like Tableau or Power BI.
* Allow investors to filter by city, property type, and price range.
* Include dynamic charts showing amenity distribution and property availability.

b) Heatmap:

* Develop a heatmap showing the concentration of amenities across cities and property types.
* Use color intensity to represent the number of amenities.

c) Comparative Bar Charts:

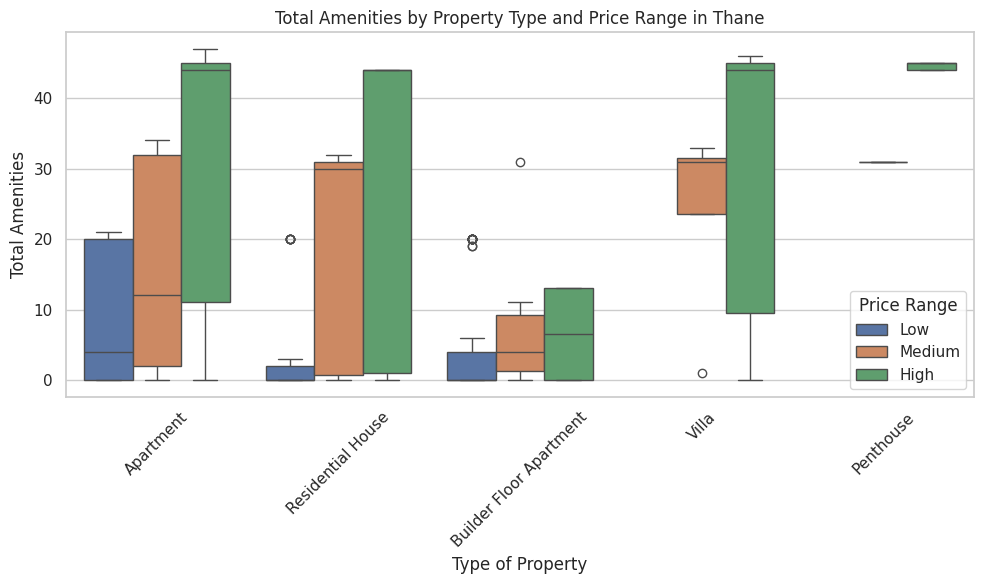
* Display side-by-side bar charts comparing amenities across price ranges for each city.

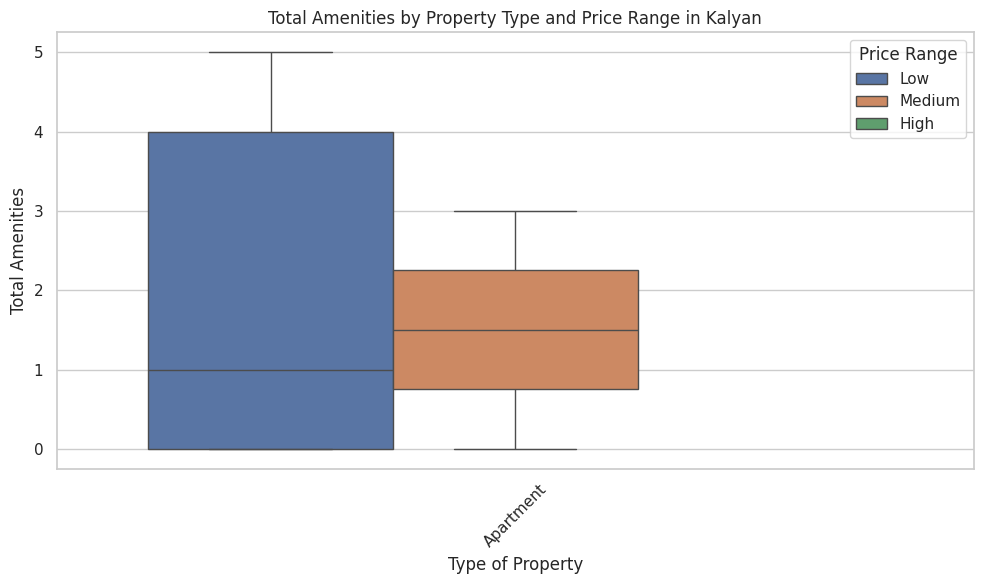
d) Infographics:

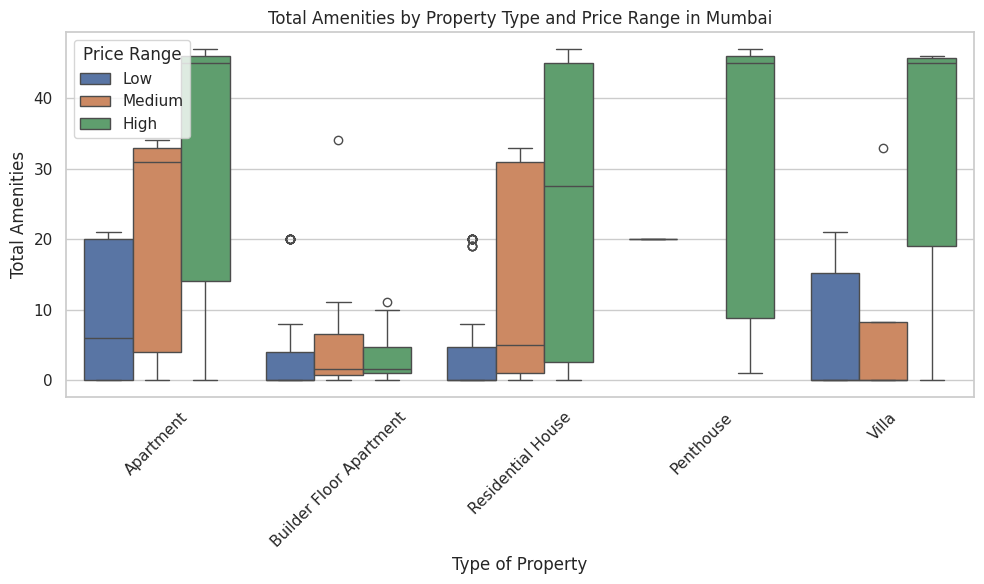
* Create visual summaries for each city, highlighting key property types and their typical amenities.

e) Interactive Property Finder:

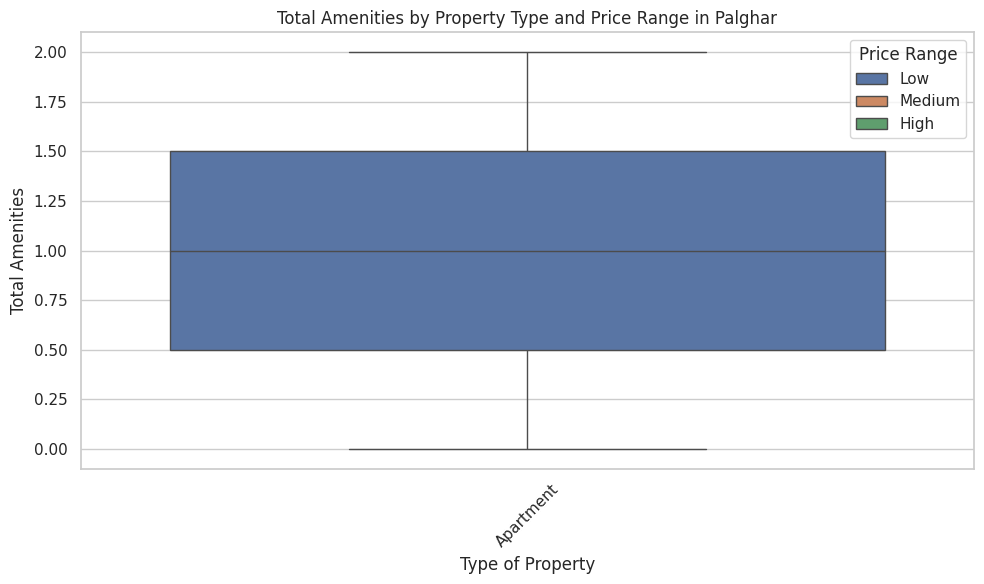
* Develop a tool where investors can input their budget and desired amenities to find matching properties across cities.

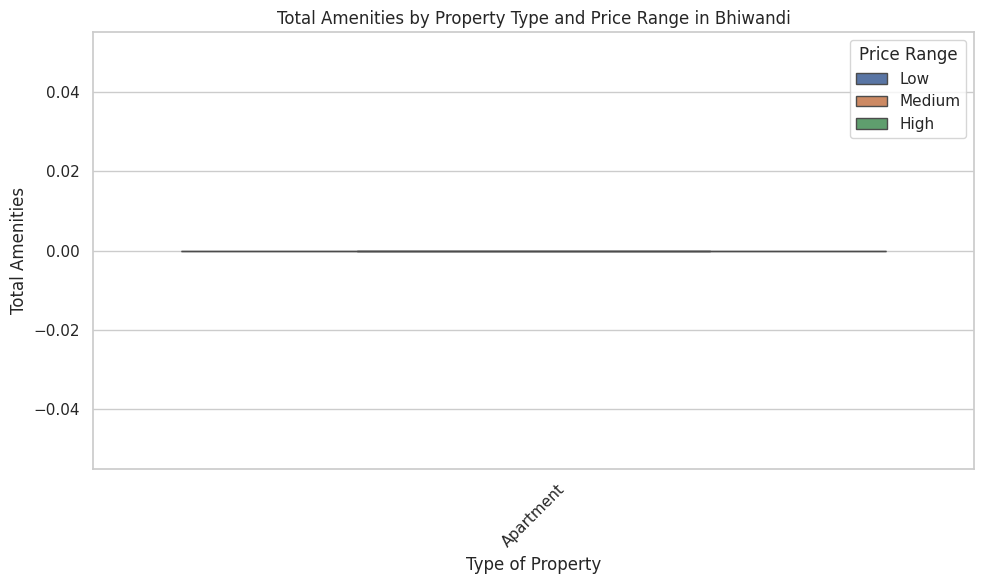




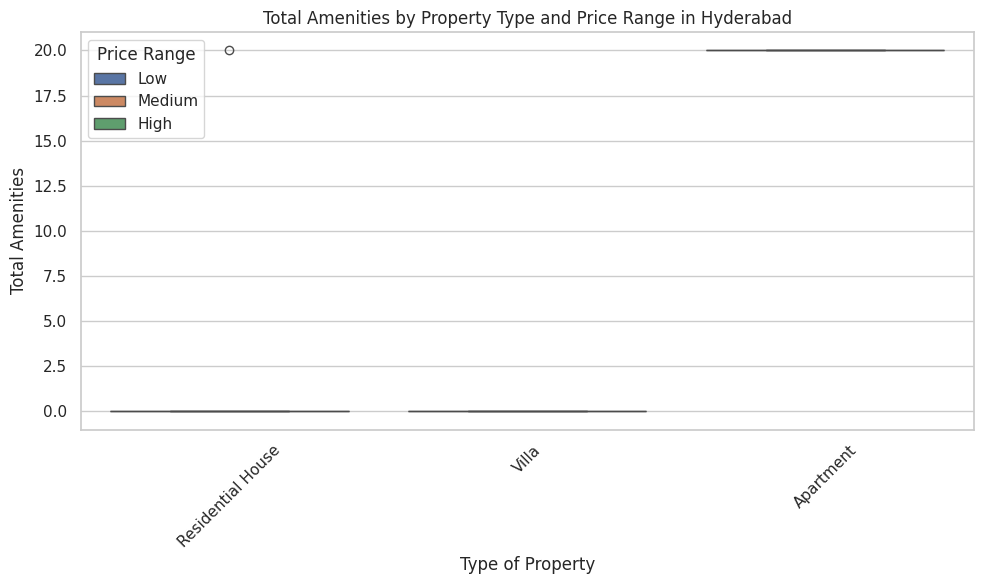


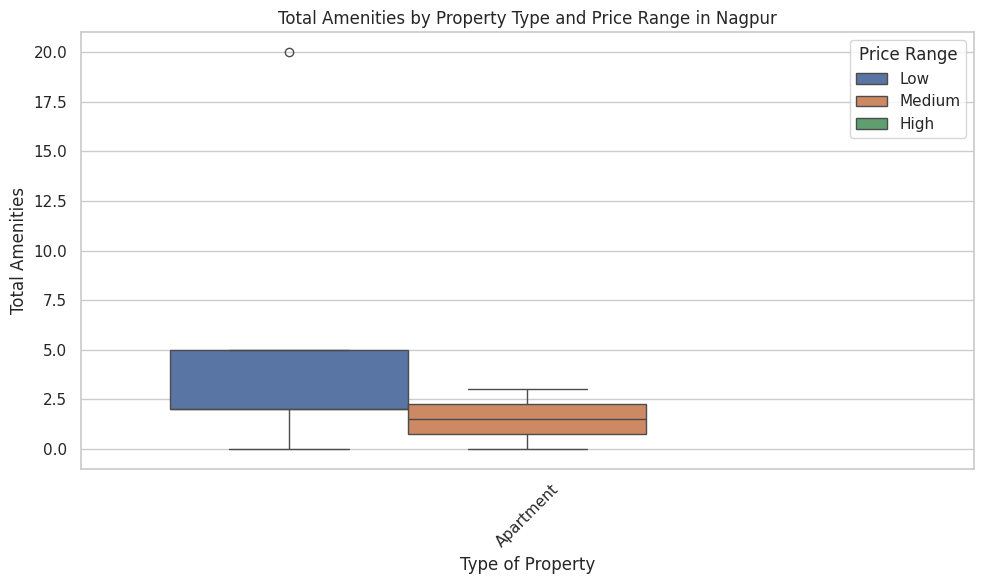












**Task-3:**  
Based on the provided graphs and data for Thane and Mumbai, I'll analyze the differences in property types, sizes (using Carpet Area), and prices between these two cities for investors.

1. Carpet Area Distribution:

Thane:

- Median carpet area: Approximately 500-600 sq ft

- Range: Mostly between 300-1000 sq ft, with some outliers up to 3000 sq ft

- Distribution: Slightly right-skewed, with a concentration of properties in the 400-800 sq ft range

Mumbai:

- Median carpet area: Slightly higher than Thane, around 600-700 sq ft

- Range: Wider than Thane, mostly between 300-1500 sq ft, with outliers up to 3500 sq ft

- Distribution: More spread out, with a significant number of properties in the 700-1200 sq ft range

Key Insight: Mumbai offers a wider range of property sizes, including more larger properties, while Thane has a higher concentration of mid-sized properties.

2. Price Distribution:

Thane:

- Median price: Around 1-1.5 crore INR

- Range: Mostly between 50 lakhs to 2 crores INR, with few outliers up to 5 crores

- Distribution: Relatively tight, with most properties clustered in the 75 lakhs to 1.5 crores range

Mumbai:

- Median price: Significantly higher, around 2-2.5 crores INR

- Range: Much wider, from 1 crore to over 10 crores INR, with some extreme outliers up to 40 crores

- Distribution: Highly right-skewed, with a long tail of high-priced properties

Key Insight: Mumbai properties are generally more expensive than Thane, with a much wider price range and more luxury options.

3. Carpet Area vs. Price Relationship:

- Both cities show a positive correlation between carpet area and price.

- Mumbai's trend line is steeper, indicating a higher price per square foot compared to Thane.

- Mumbai has more high-priced outliers, especially for larger carpet areas.

- Thane's properties cluster more tightly, suggesting more uniform pricing.

Key Insight: While larger properties generally cost more in both cities, the price increase per square foot is more pronounced in Mumbai.

4. Property Types (inferred from data patterns):

Thane:

- More uniformity in property types, likely dominated by apartments

- Fewer extreme outliers in both size and price suggest fewer ultra-luxury or very large properties

Mumbai:

- Greater diversity in property types, including apartments, penthouses, and possibly standalone houses

- The presence of very high-priced, large carpet area properties indicates luxury segments not seen in Thane

Investment Recommendations:

1. Budget-Conscious Investors:

- Focus on Thane for more affordable options

- Look for properties in the 400-800 sq ft range, priced between 75 lakhs to 1.5 crores

- Expect more standardized apartment options

2. Mid-Range Investors:

- Consider both Thane and Mumbai

- In Thane, look for larger apartments (800-1200 sq ft) in the 1.5-2.5 crore range

- In Mumbai, focus on 600-900 sq ft properties in the 2-3 crore range

3. Luxury Investors:

- Primarily focus on Mumbai

- Look for properties above 1500 sq ft, priced at 5 crores and above

- Expect a variety of luxury options including large apartments, penthouses, and possibly standalone houses

4. Investors Seeking Variety:

- Choose Mumbai for a wider range of property types and sizes

- Be prepared for higher prices per square foot compared to Thane

5. Value Investors:

- Explore Thane for potentially undervalued properties, especially in the upper size range (1000-1500 sq ft)

- Look for properties priced below the Mumbai average but offering similar carpet areas

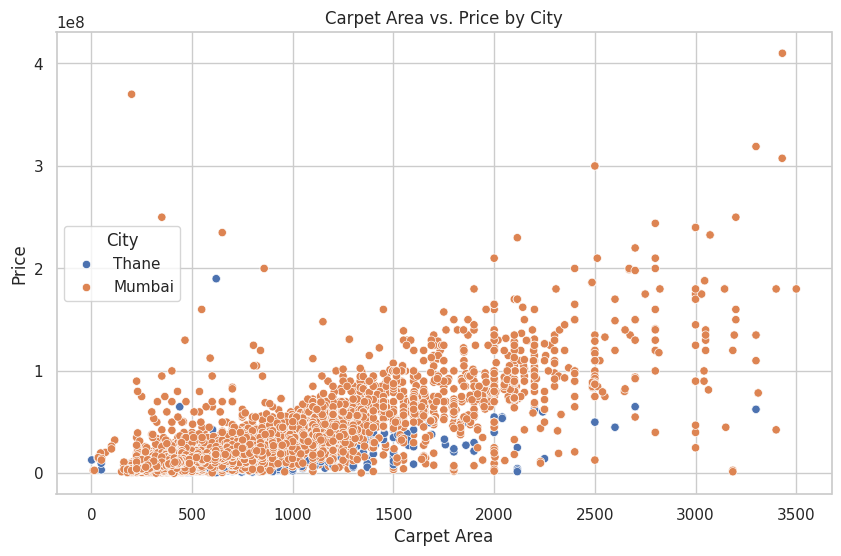
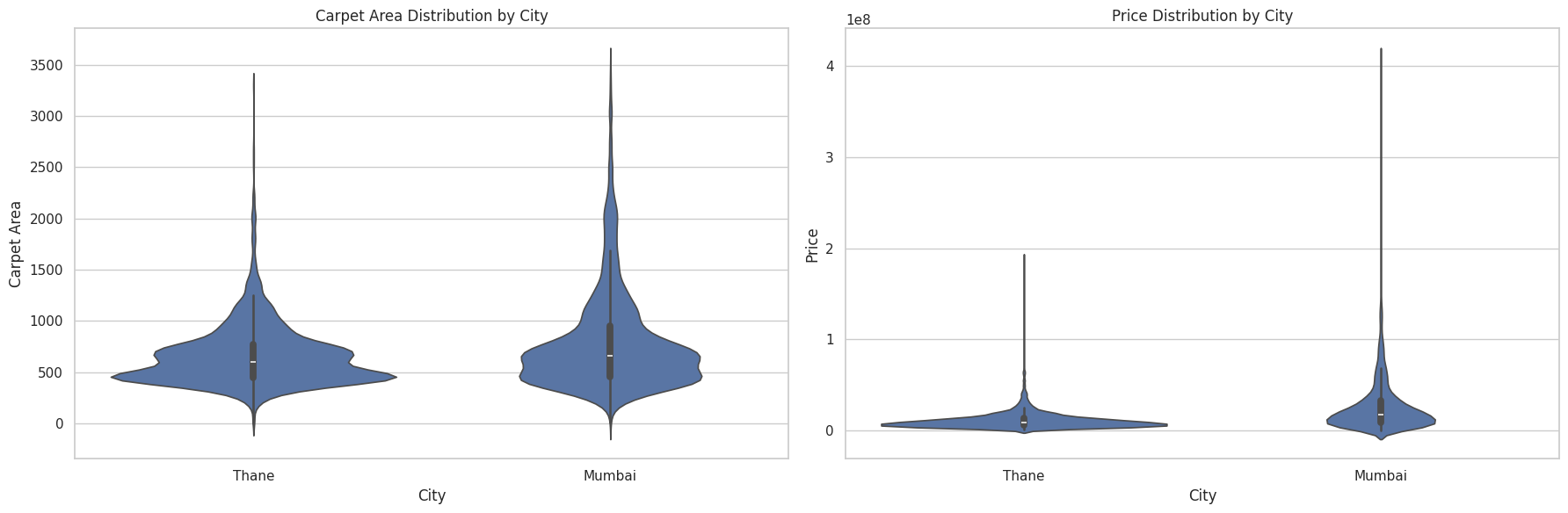
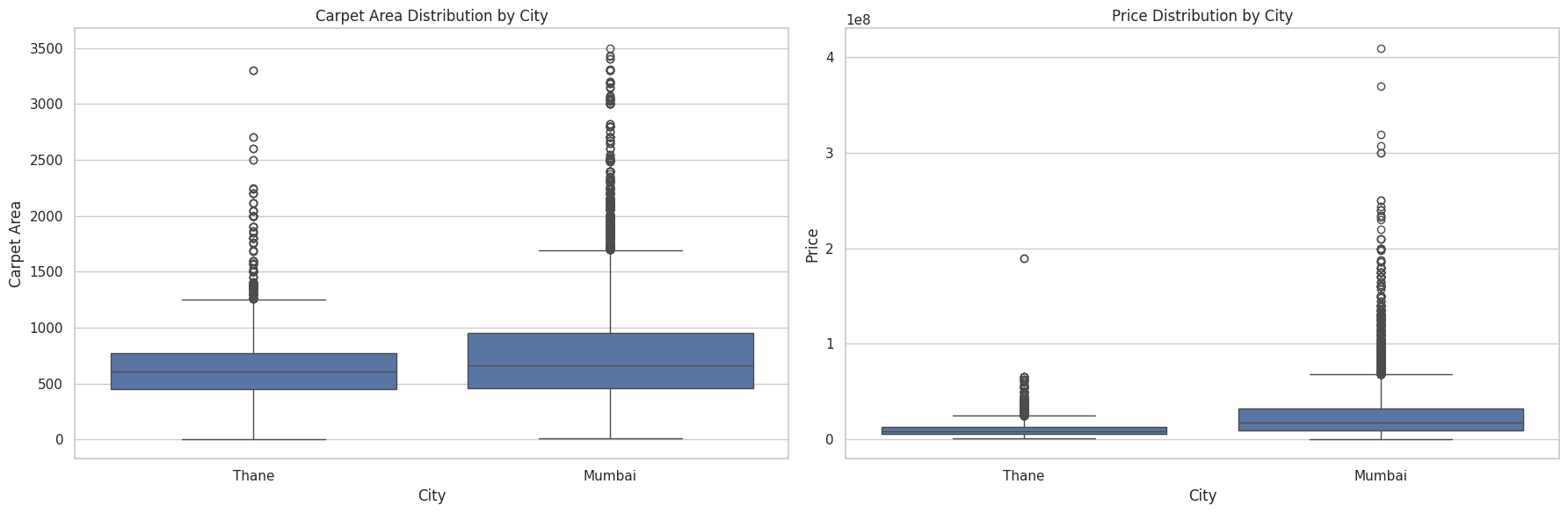
6. Long-term Growth Investors:

- Consider upcoming areas in both cities

- In Thane, look for larger properties that might appreciate as the city develops

- In Mumbai, focus on mid-range properties in developing neighborhoods that might see future price growth

Overall, Thane offers more affordable and standardized options, making it suitable for first-time buyers and budget-conscious investors. Mumbai provides a wider range of choices, including high-end luxury properties, but at significantly higher price points. The choice between the two will depend on the investor's budget, risk appetite, and specific property requirements.



**Task-5:**

Based on the provided code and the resulting visualizations, I'll analyze the Carpet Area per unit of Price (CAPP) across different cities and price ranges to identify properties that offer the best value for investors preferring larger area properties with relatively lower budgets.

1. Overview of CAPP by City and Price Range:

The box plots and violin plots provide a comprehensive view of the CAPP distribution across cities and price ranges. Here's a summary of the findings:

a) Mumbai:

- Low price range: Highest median CAPP, wide distribution

- Medium price range: Lower median CAPP than low range, but still relatively high

- High price range: Lowest median CAPP, narrow distribution

b) Thane:

- Low price range: High median CAPP, slightly lower than Mumbai

- Medium price range: Lower median CAPP than low range, but higher than Mumbai's medium range

- High price range: Lowest median CAPP, but higher than Mumbai's high range

c) Other Cities (Kalyan, Nagpur, Palghar):

- Generally higher CAPP across all price ranges compared to Mumbai and Thane

- Wider distributions, indicating more variability in value

2. Best Value Properties by Price Range:

a) Low Price Range:

- Kalyan offers the highest CAPP, followed closely by Nagpur and Palghar

- Thane provides better value than Mumbai in this range

- Mumbai still has some high CAPP outliers, indicating potential good deals

b) Medium Price Range:

- Kalyan and Nagpur continue to offer high CAPP

- Thane generally provides better value than Mumbai

- Mumbai shows more variability, with some properties offering high CAPP

c) High Price Range:

- Limited data for cities other than Mumbai and Thane

- Thane offers better CAPP than Mumbai in this range

- Mumbai has the lowest CAPP in the high price range, but with some high-value outliers

3. City-specific Insights:

a) Kalyan:

- Consistently high CAPP across all price ranges

- Best option for investors seeking large areas on a budget

- Limited number of properties available

b) Nagpur:

- High CAPP in low and medium price ranges

- Good alternative to Mumbai and Thane for budget-conscious investors

- Limited high-priced options

c) Palghar:

- High CAPP in the low price range

- Limited data for medium and high price ranges

- Potential for good value in budget properties

d) Thane:

- Better overall value than Mumbai across all price ranges

- Particularly good CAPP in the low price range

- More consistent CAPP across price ranges compared to Mumbai

e) Mumbai:

- Lower overall CAPP compared to other cities

- Some high-value outliers in all price ranges

- Widest range of options, but generally at a premium

4. Recommendations for Investors:

a) Budget-conscious investors seeking maximum area:

- Focus on Kalyan, Nagpur, and Palghar in the low price range

- Consider Thane as a more accessible alternative to Mumbai

b) Mid-range investors:

- Explore Kalyan and Nagpur for the best CAPP

- Consider Thane for a balance of good CAPP and urban amenities

c) High-budget investors prioritizing area:

- Look for high-CAPP outliers in Mumbai's high price range

- Consider Thane for consistently better CAPP in the high price range

d) Value hunters:

- Focus on the upper quartile of CAPP in each city and price range

- Pay special attention to outliers in Mumbai, which may represent unique opportunities

e) Long-term investors:

- Consider properties in Kalyan, Nagpur, and Palghar, which may appreciate as these areas develop

- Look for high-CAPP properties in upcoming areas of Thane

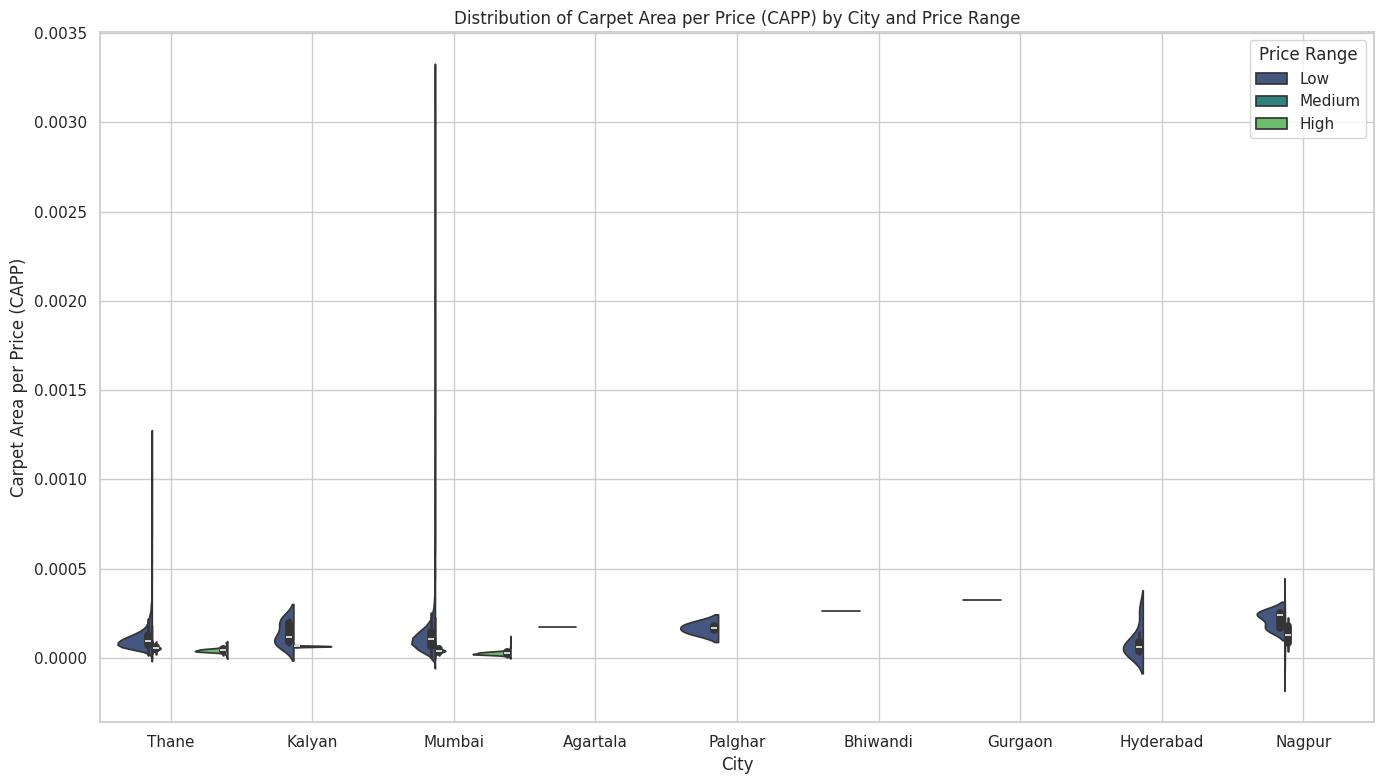
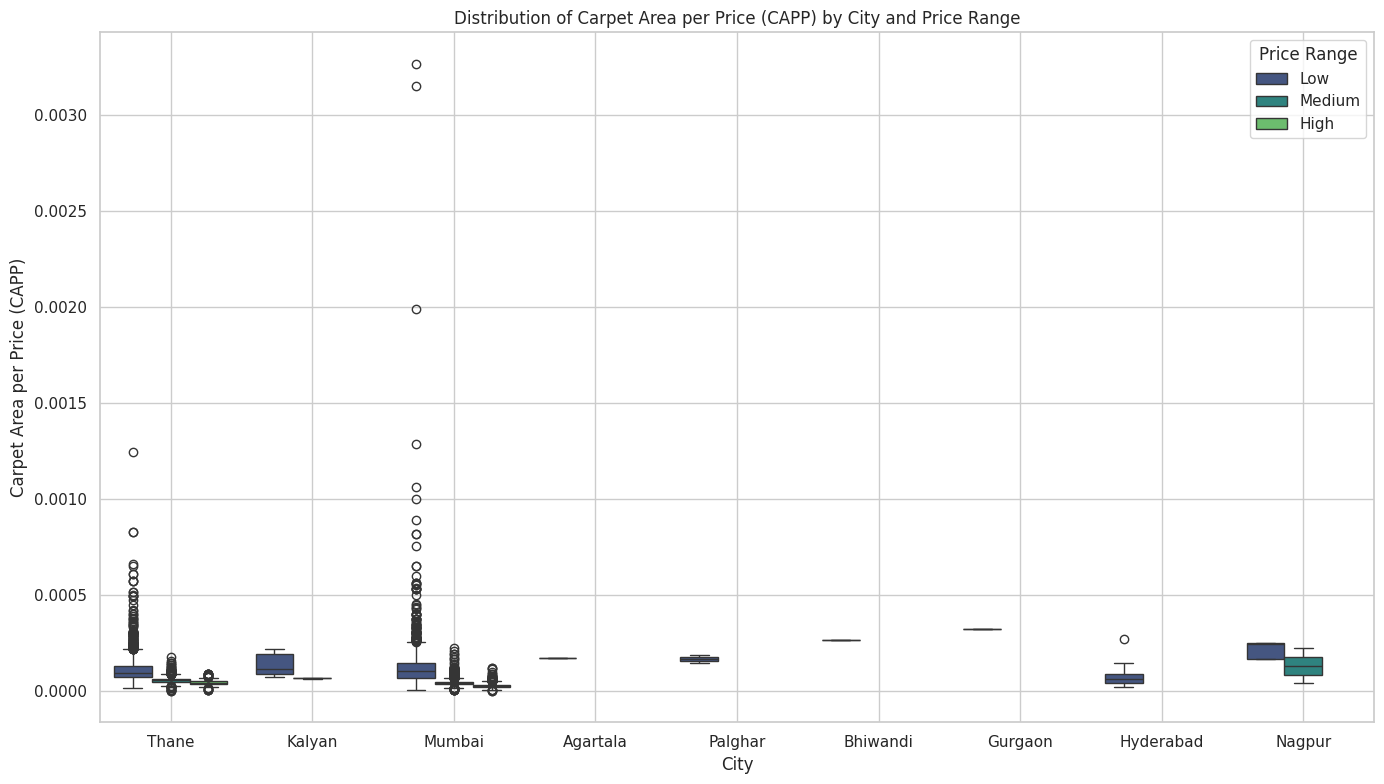
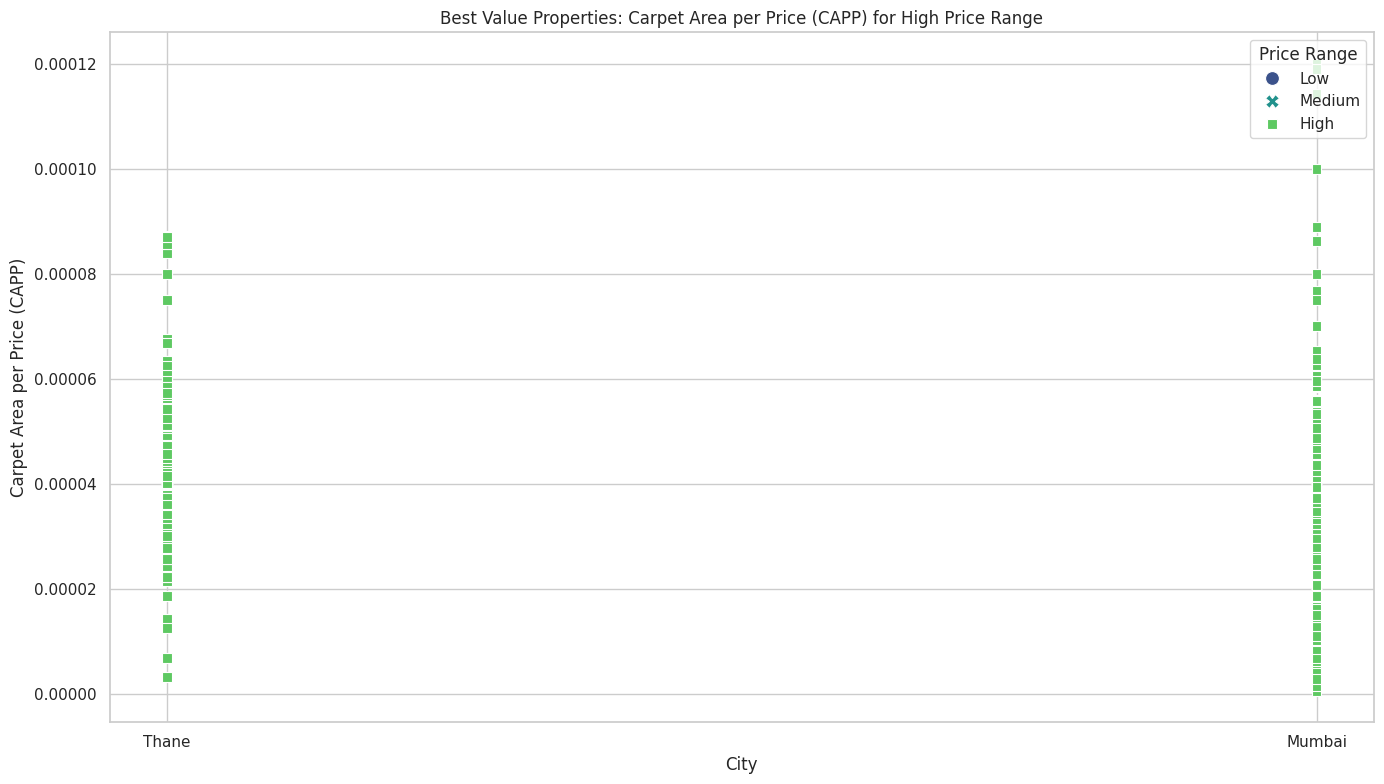
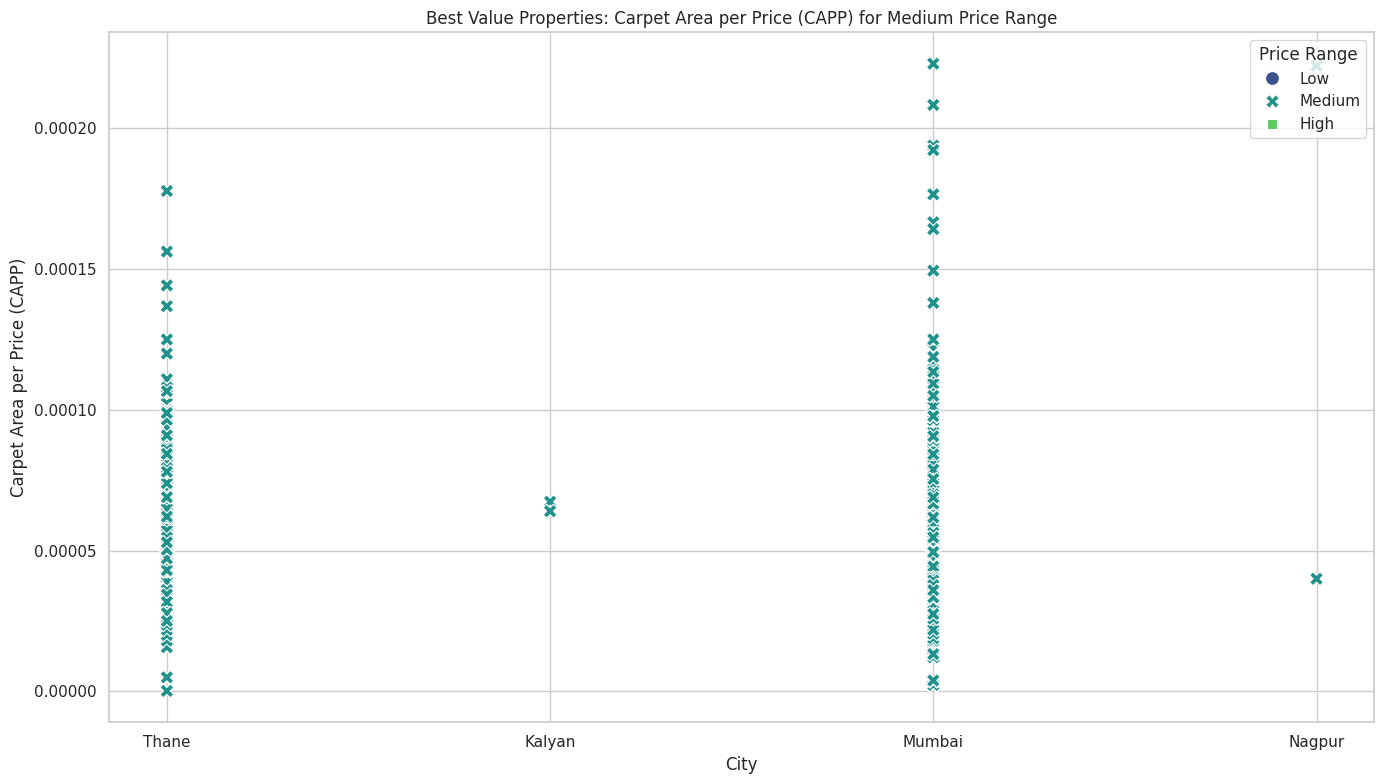
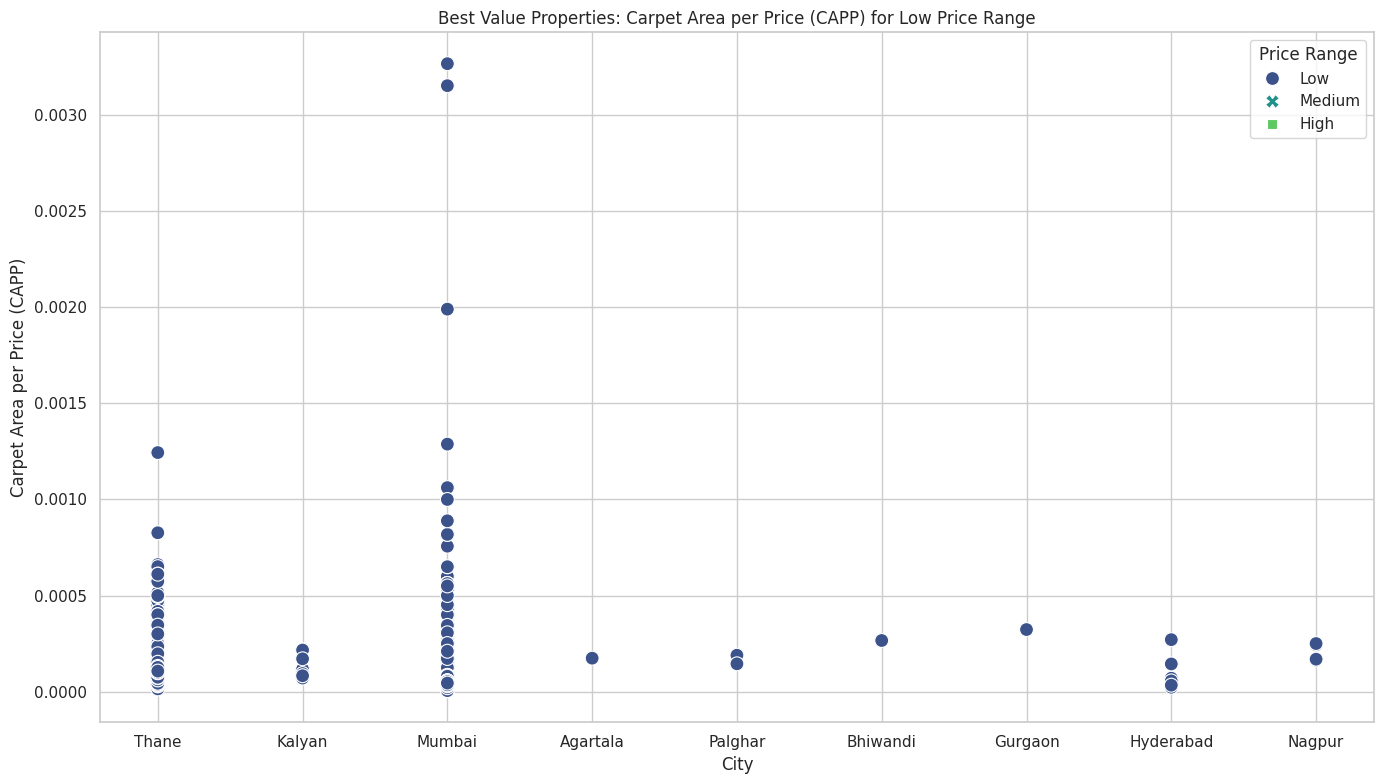
5. Additional Considerations:

- While CAPP is a useful metric, investors should also consider factors like location, amenities, and potential for appreciation.

- The high CAPP in smaller cities may come with trade-offs in terms of urban development and job opportunities.

- Mumbai's lower CAPP is often offset by factors like better infrastructure, job markets, and potential for price appreciation.

In conclusion, investors seeking larger areas on a budget should primarily focus on Kalyan, Nagpur, and Palghar in the low to medium price ranges. Thane offers a good compromise between value and urban amenities across all price ranges. While Mumbai generally offers lower CAPP, it presents unique opportunities in the form of high-value outliers and potential for long-term appreciation.



**Task-4:**  
The goal of this analysis is to compare the average Carpet Area (ACA) of high-budget properties located in prime versus non-prime locations across different cities. The cities analyzed are Kalyan, Mumbai, and Nagpur.

#### **Code Implementation Summary:**

1. **Data Preparation**:
   * The dataset was first grouped by City and Price Range to calculate the average Carpet Area (ACA).
   * This average ACA was then merged back into the main dataset to allow for filtering of high-budget properties.
2. **Segregation**:
   * The dataset was divided into prime and non-prime properties based on the isPrimeLocationProperty flag.
   * Average ACA was calculated separately for prime and non-prime properties within each city.
3. **Visualization**:
   * Separate bar charts were created for each city to visually compare the average ACA between prime and non-prime locations.

#### **Visual Analysis:**

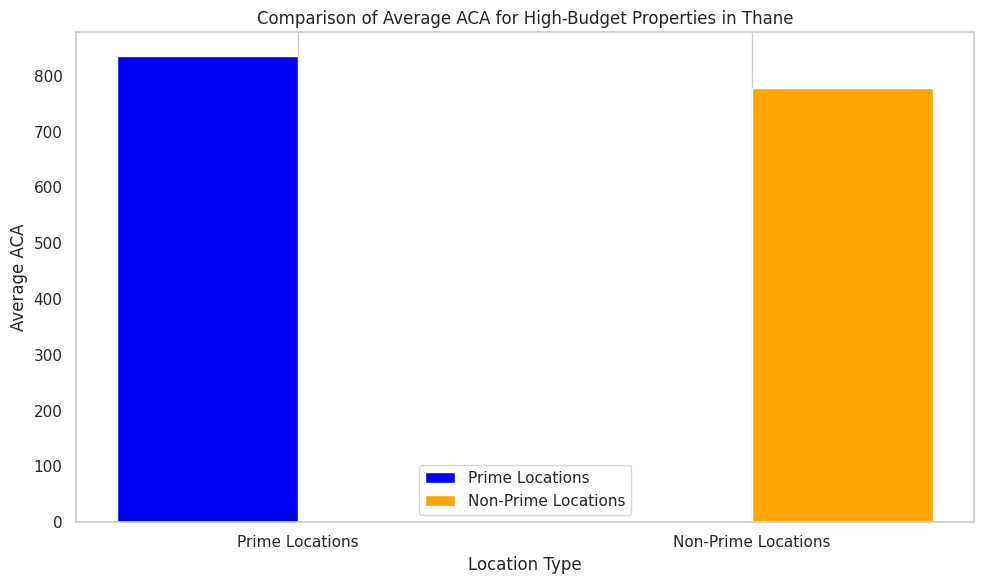
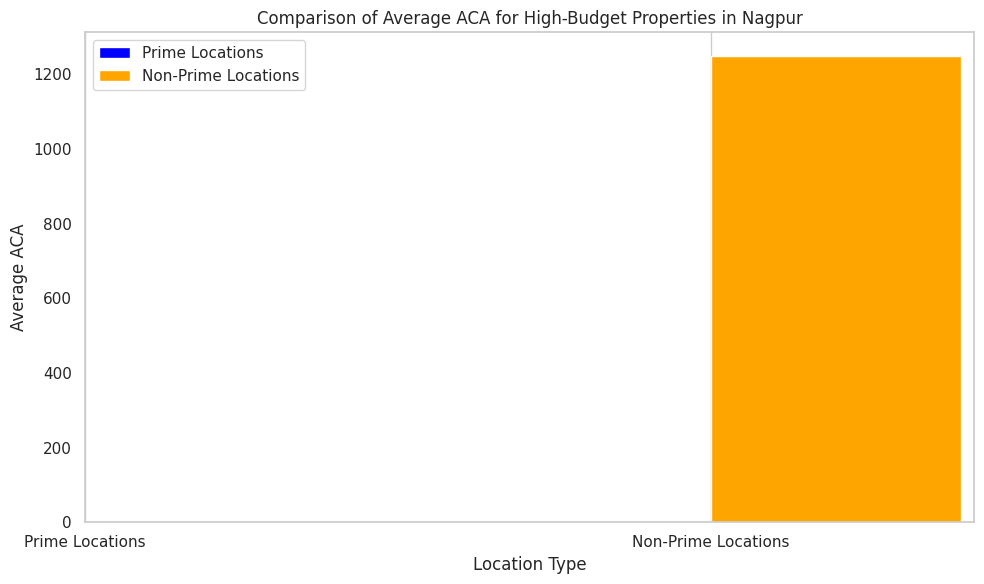
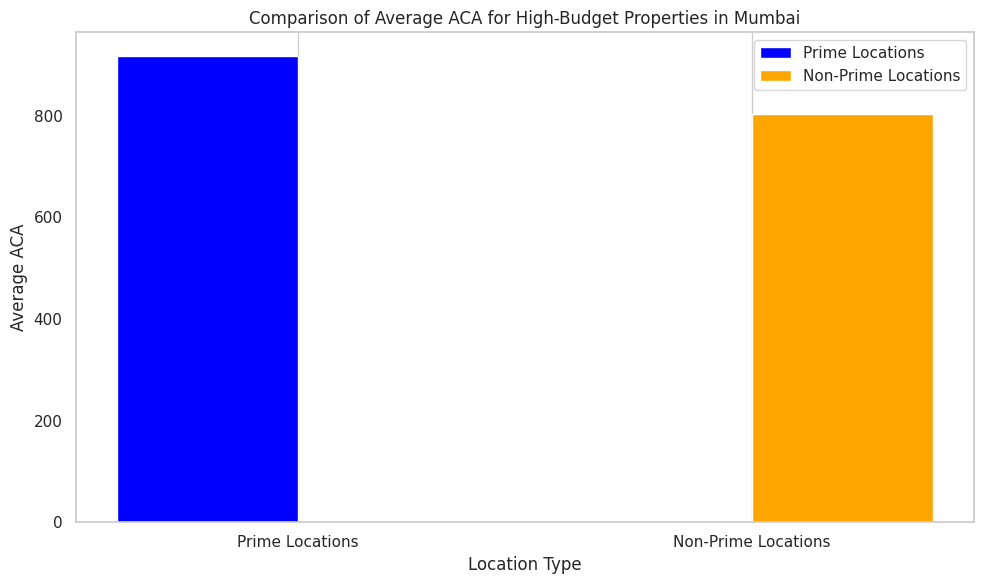
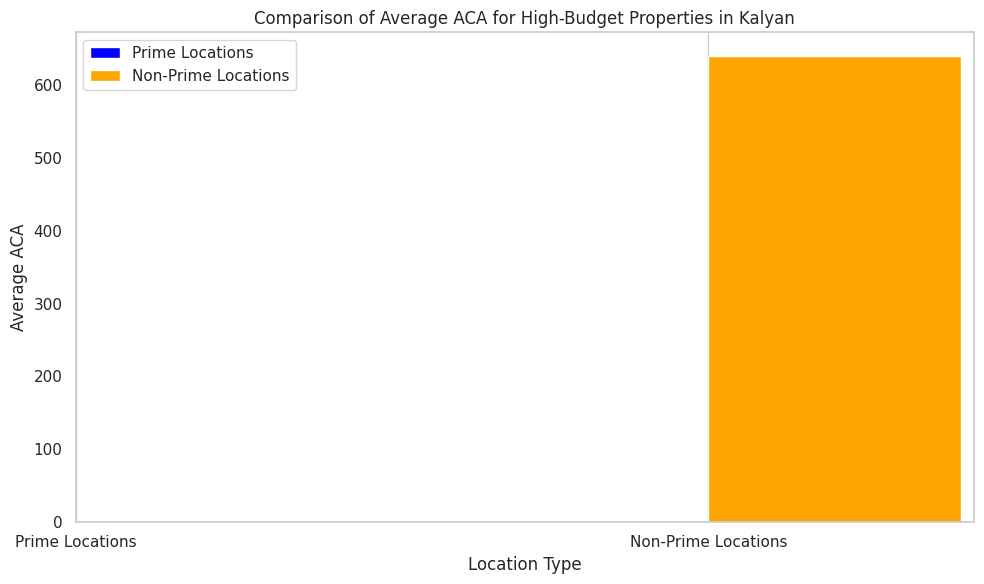
1. **Kalyan**:
   * **Prime Locations**: Data unavailable or insignificant.
   * **Non-Prime Locations**: The average ACA is 640 sq. ft.
   * **Observation**: In Kalyan, only non-prime high-budget properties have a significant ACA, indicating a possible preference for larger carpet areas in non-prime locations or fewer prime high-budget properties available.
2. **Mumbai**:
   * **Prime Locations**: The average ACA is 918 sq. ft.
   * **Non-Prime Locations**: The average ACA is 804 sq. ft.
   * **Observation**: In Mumbai, prime locations have a higher average ACA compared to non-prime locations, reflecting the premium nature of space in prime areas.
3. **Nagpur**:
   * **Prime Locations**: Data unavailable or insignificant.
   * **Non-Prime Locations**: The average ACA is 1250 sq. ft.
   * **Observation**: Like Kalyan, Nagpur's data suggests larger properties in non-prime locations, potentially due to a scarcity of prime high-budget properties.

**4. Thane**:

* + 1. **Prime Locations**:The average ACA is 836sq. ft.
    2. **Non-Prime Locations**: The average ACA is 777 sq. ft.
    3. **Observation**: Like Mumbai, Thanje’s prime locations have a higher average ACA compared to non-prime locations, reflecting the premium nature of space in prime areas.

#### **Conclusion:**

* **Mumbai** and **Thane** are cities where high-budget properties in prime locations have a higher ACA compared to non-prime locations, emphasizing the premium nature of real estate in such areas.
* **Kalyan** and **Nagpur** exhibit a trend where non-prime locations offer larger carpet areas, possibly due to limited availability or lower demand for prime location properties in these cities.
* **City-Specific Trends**: There is considerable variation in ACA between prime and non-prime locations across cities. Mumbai shows a more typical trend of higher ACA in prime locations, whereas Nagpur's non-prime areas offer much larger carpet areas.
* **Investor Implications**: Investors targeting Mumbai and **Thane** may prefer prime locations for more spacious properties, while in cities like Nagpur, non-prime locations may offer better value in terms of carpet area.
* **Missing Data**: For cities like Kalyan and Nagpur, the lack of ACA data in prime locations indicates a need for further investigation or a potential data gap.



**Task-6:**

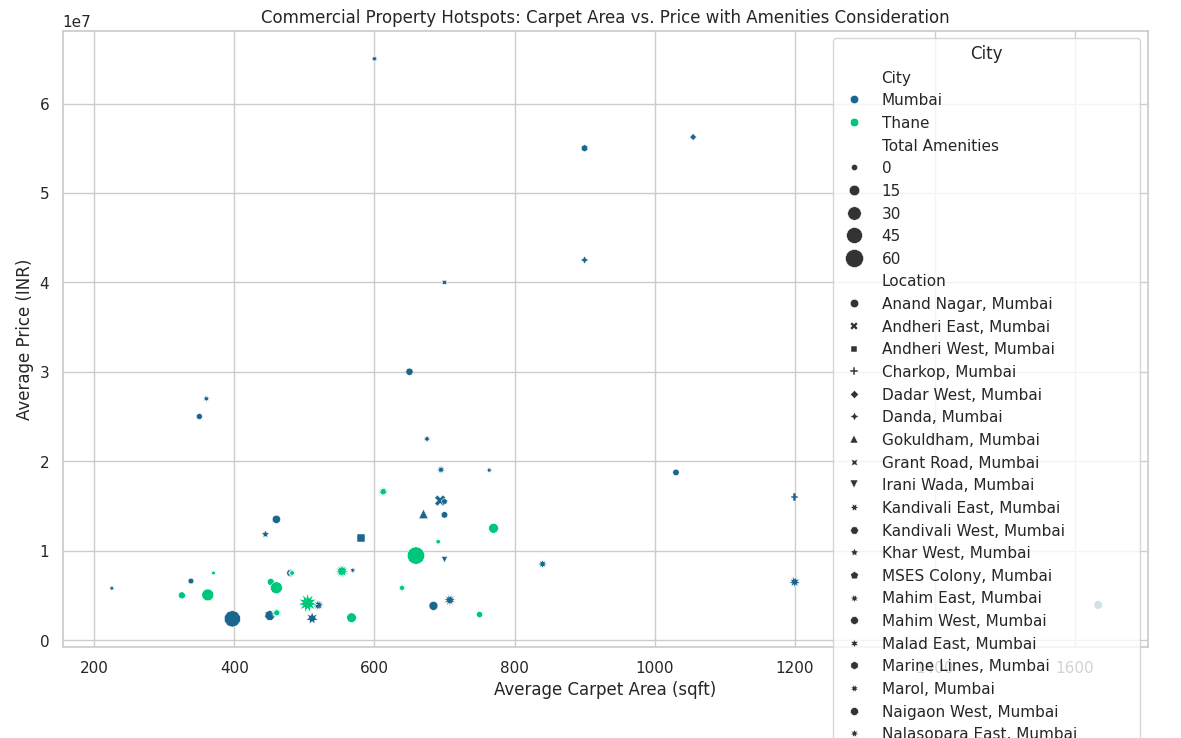
#### **Data and Methodology:**

* **Data Source**: The analysis is based on a dataset containing property prices in Mumbai and Thane, along with the availability of various amenities.
* **Amenities:** Key columns in the dataset include City, Possession Status, Availability Starts From, Price, Type of Property, Carpet Area, and a variety of amenities like 'isPrimeLocationProperty', 'Power Back Up', 'Lift', 'Parking', 'Security', 'Air Conditioned', 'Visitor Parking', 'Intercom Facility', 'Maintenance Staff', 'Cafeteria/Food Court', 'Conference Room'.
* **Approach**: For each city, the average property price was calculated for properties with and without each amenity. The difference between these averages helps gauge the impact of each amenity on property prices.

**Visualization:**

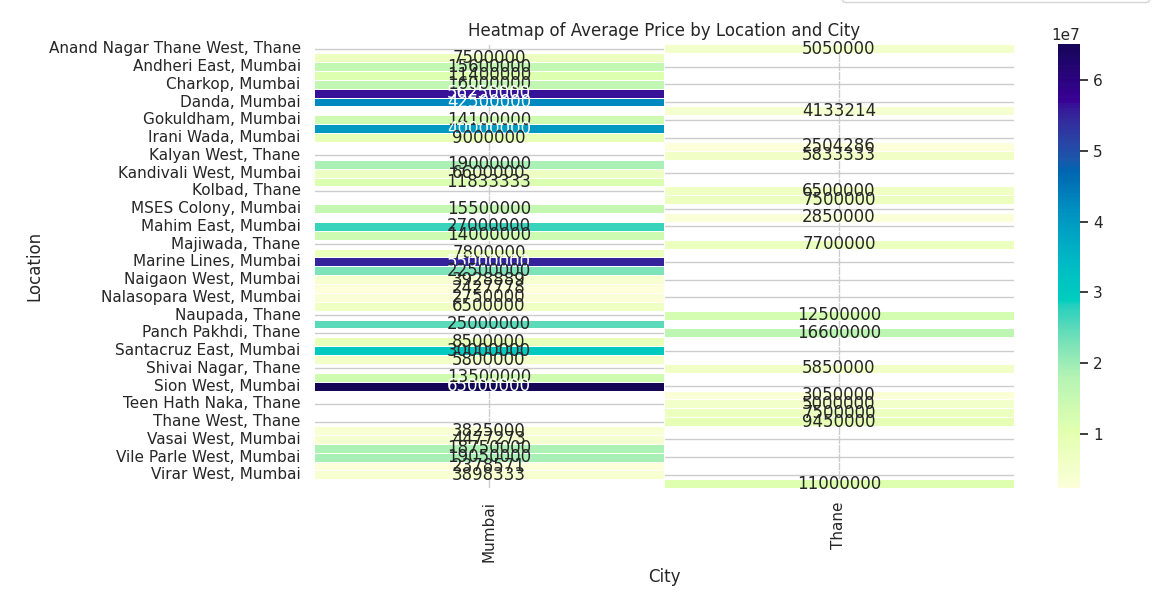
### **Visualization 1: Scatter Plot of Carpet Area vs. Price with Amenities Consideration**

* **Description**: This scatter plot visualizes the relationship between carpet area and price for commercial properties, with points sized by the total number of amenities.
* **Findings**: Larger properties with more amenities tend to have higher prices, indicating that amenities significantly affect property value.



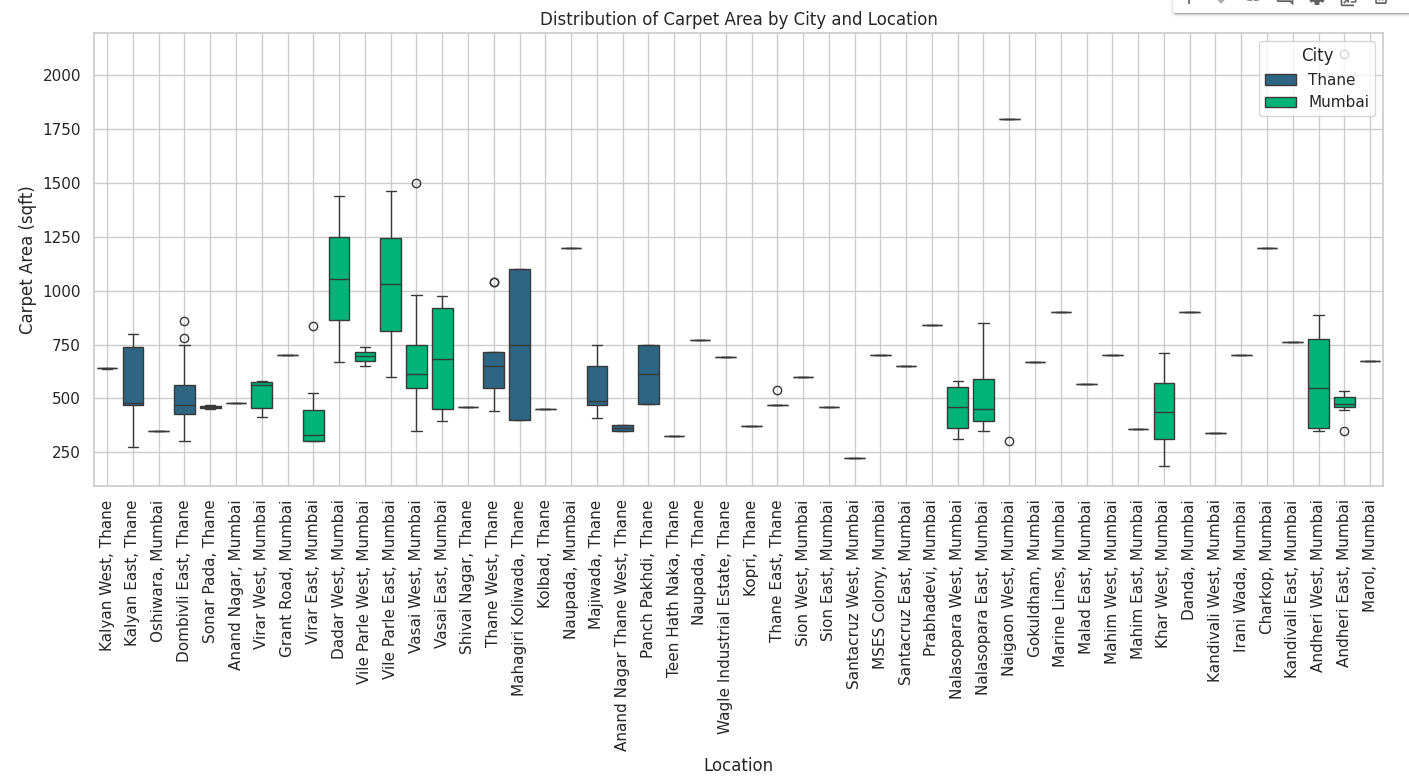
### **Visualization 2: Heatmap of Price by Location and City**

* **Description**: A heatmap showing the average price of commercial properties by location and city.
* **Findings**: Locations with higher average prices are highlighted, and the heatmap helps identify areas with potentially higher property values.



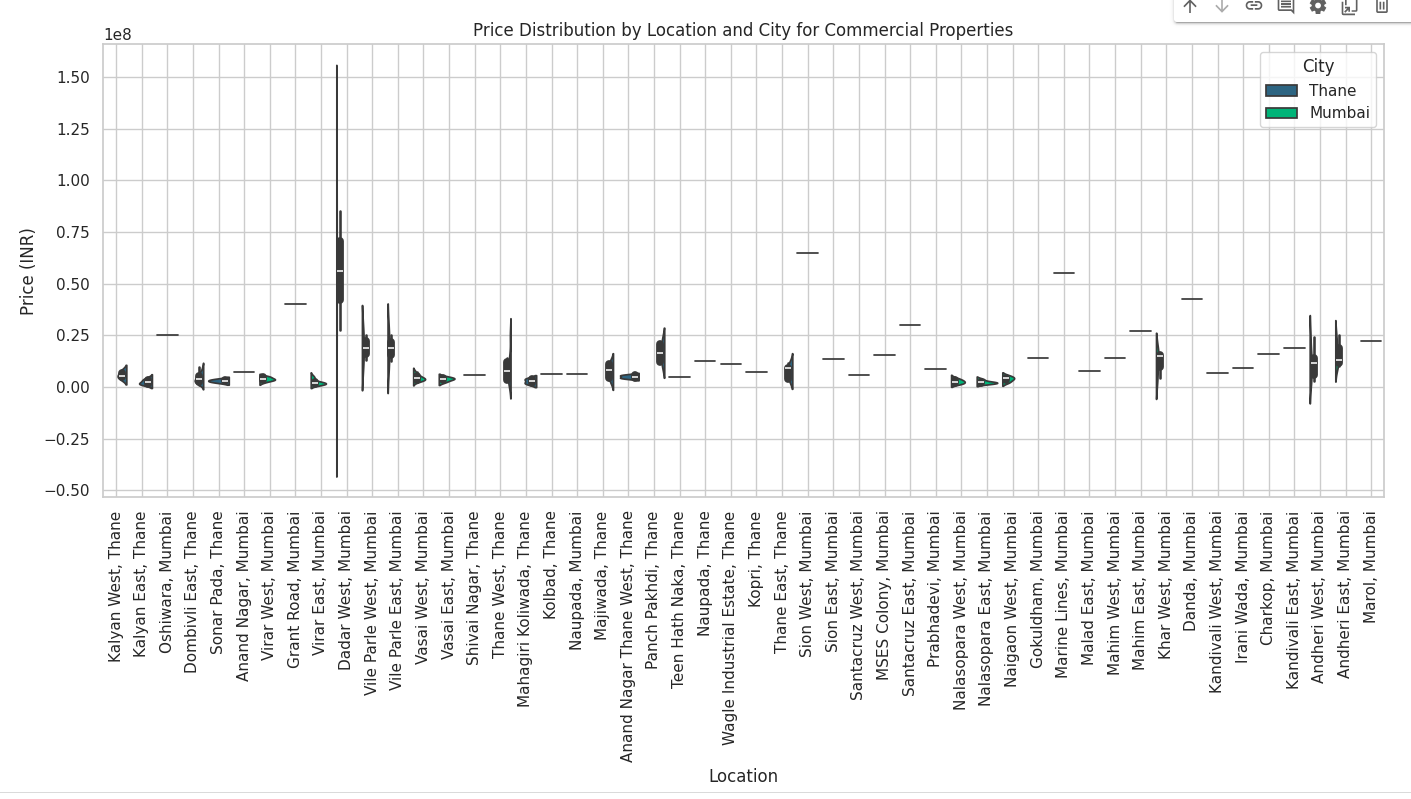
### **Visualization 3: Box Plot of Carpet Area by City and Location**

* **Description**: This box plot shows the distribution of carpet area across different locations for Mumbai and Thane.
* **Findings**: Variations in carpet area are observed across locations and cities, indicating differing property sizes and possibly influencing prices.



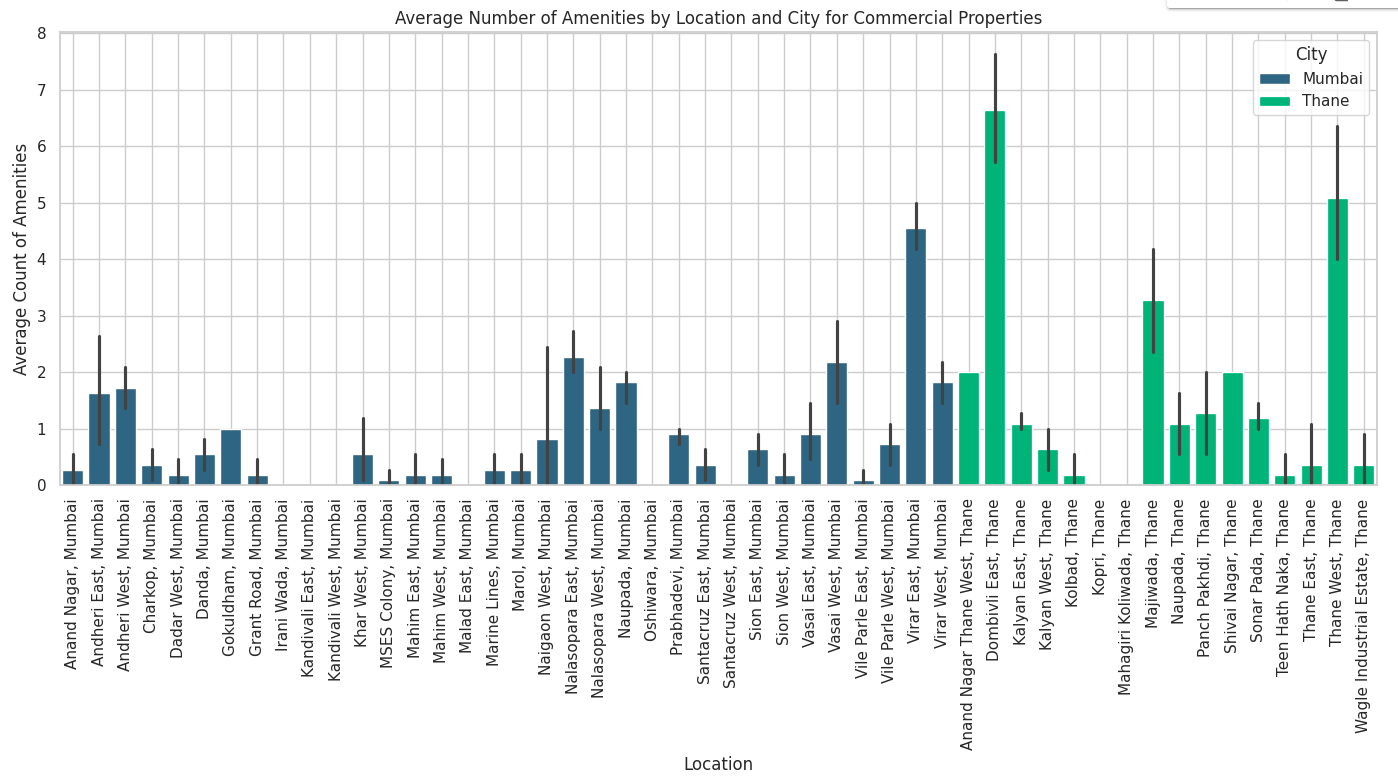
### **Visualization 4: Violin Plot of Price Distribution by City and Location**

* **Description**: A violin plot depicting the distribution of property prices across different locations, with variations based on city.
* **Findings**: The distribution of prices provides insights into pricing trends across different locations and cities, with amenities influencing the price distribution.



### **Visualization 5: Bar Plot of Average Number of Amenities by Location**

* **Description**: This bar plot shows the average number of amenities available at different locations.
* **Findings**: Locations with more amenities typically have higher property prices, suggesting that amenities play a crucial role in determining property values.



#### **Conclusions:**

**Impact of Possession Status on Property Prices**

1. **Price Differences**:
   * 'Ready to Move' properties generally command higher prices compared to 'Under Construction' properties. This reflects the immediate availability and reduced risk for buyers.
   * The difference in price between these types varies between Mumbai and Thane, with Mumbai generally exhibiting higher price differences.

### **Commercial Property Hotspots**

1. **Price and Amenities**:
   * Commercial properties in locations with more amenities tend to have higher average prices.
   * High-value hotspots are identifiable through heatmaps and scatter plots, highlighting areas with significant property investments.
2. **Distribution Analysis**:
   * Box and violin plots reveal variations in carpet area and price distributions across different locations, helping investors identify valuable properties and investment opportunities.

**Task-7:**

#### **Data and Methodology:**

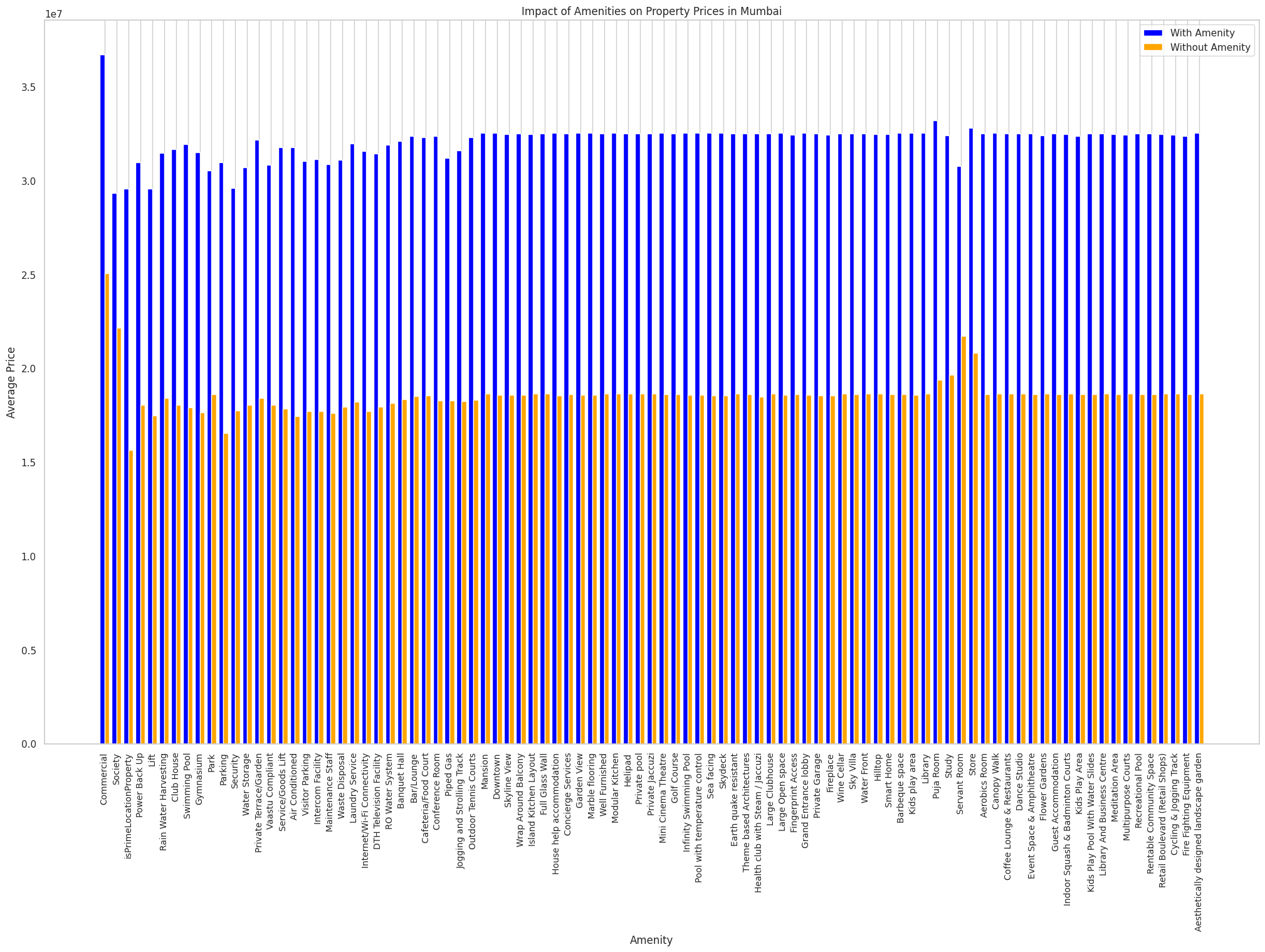
* **Data Source**: The analysis is based on a dataset containing property prices in Mumbai and Thane, along with the availability of various amenities.
* **Amenities**: A comprehensive list of amenities was considered, including common features like parking, swimming pool, and gymnasium, as well as luxury features like a private pool, helipad, and wine cellar.
* **Approach**: For each property, metrics such as average price and carpet area were calculated. Amenities were summed to provide insights into the total amenities available at various locations.

#### **Visualization:**

1. **Mumbai**:
   * The graph shows the average property prices in Mumbai for properties with and without specific amenities.
   * **Key Observations**:
     + Properties with amenities such as commercial access, prime location, and features like a swimming pool, gymnasium, and clubhouse show a significant increase in average prices.
     + High-end amenities like private pools, helipads, and mini cinema theaters also substantially increase property prices.
     + Basic amenities like power backup and parking also show a noticeable price difference, though not as pronounced as the luxury amenities.
2. **Thane**:
   * The graph illustrates the average property prices in Thane, comparing properties with and without each amenity.
   * **Key Observations**:
     + Similar to Mumbai, amenities like prime location, swimming pools, and gymnasiums significantly impact property prices in Thane.
     + The impact of luxury amenities, while present, is less pronounced in Thane compared to Mumbai, possibly due to a difference in market dynamics.
     + Basic amenities like power backup and parking continue to show a positive impact on property prices, but again, the effect is less dramatic than in Mumbai.

#### **Conclusions:**

1. **Impact of Amenities**:
   * In both Mumbai and Thane, amenities generally lead to higher property prices. The extent of this impact varies depending on the type of amenity.
   * Luxury amenities add substantial value in Mumbai, while their impact is relatively muted in Thane, possibly due to differences in buyer preferences and market segmentation.
2. **Differences Between Cities**:
   * Mumbai shows a more pronounced increase in property prices for both basic and luxury amenities, indicating a higher premium placed on these features.
   * Thane, while still showing positive impacts from amenities, reflects a market where the presence of such features is not as critical to property valuation as in Mumbai.
3. **Recommendations for Investors**:
   * For investors in Mumbai, properties with luxury amenities could offer higher returns due to their significant impact on prices.
   * In Thane, focusing on properties with essential amenities like parking, security, and gymnasiums may offer a better value proposition.



**Task-8:**  
**Introduction:**

Investors are interested in understanding how the 'Possession Status' (whether a property is 'Ready to Move' or 'Under Construction') and the 'Availability Starts From' dates affect property prices. The goal is to determine if there are significant price differences between 'Ready to Move' and 'Under Construction' properties, and how these differences vary between the cities of Mumbai and Thane.

#### **Data and Methodology:**

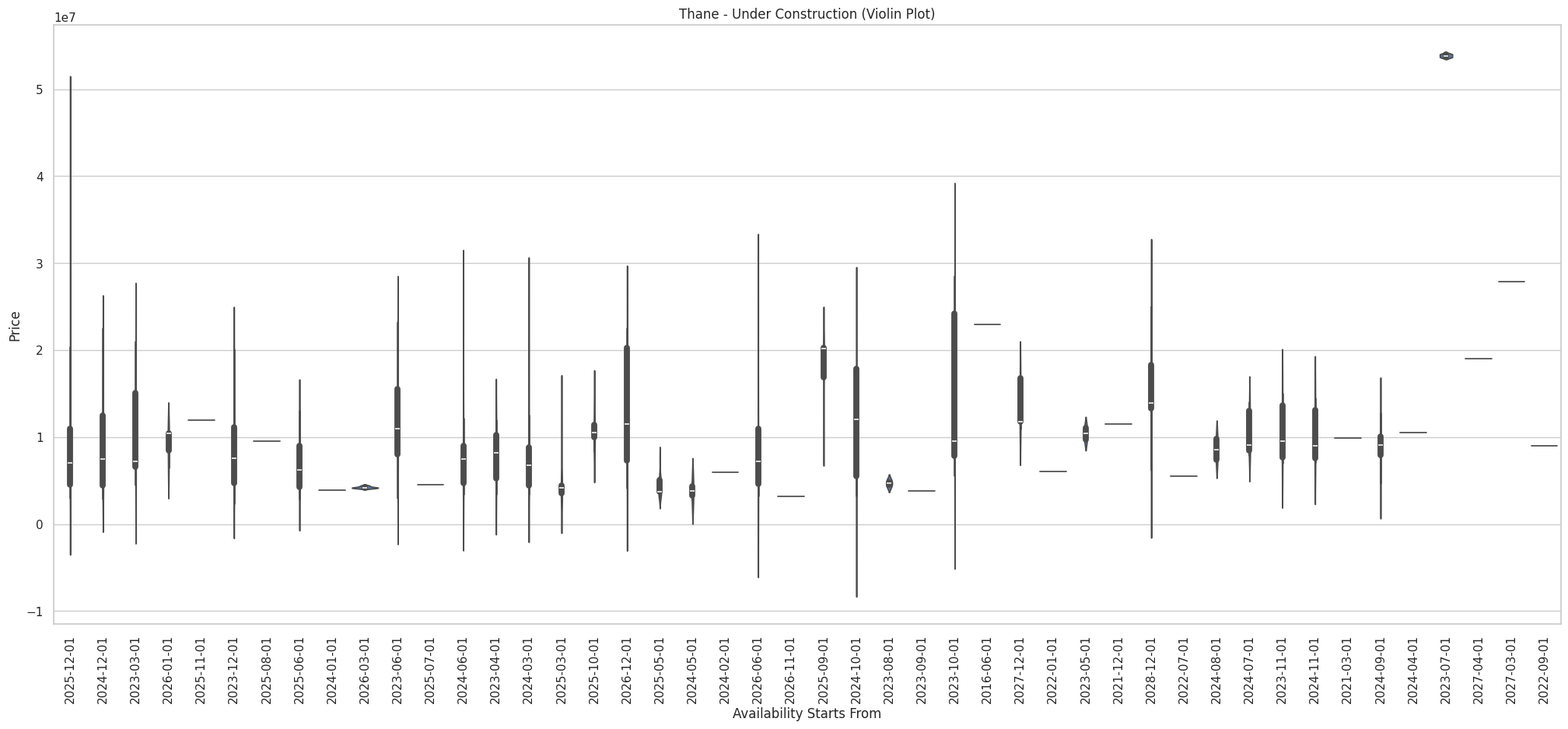
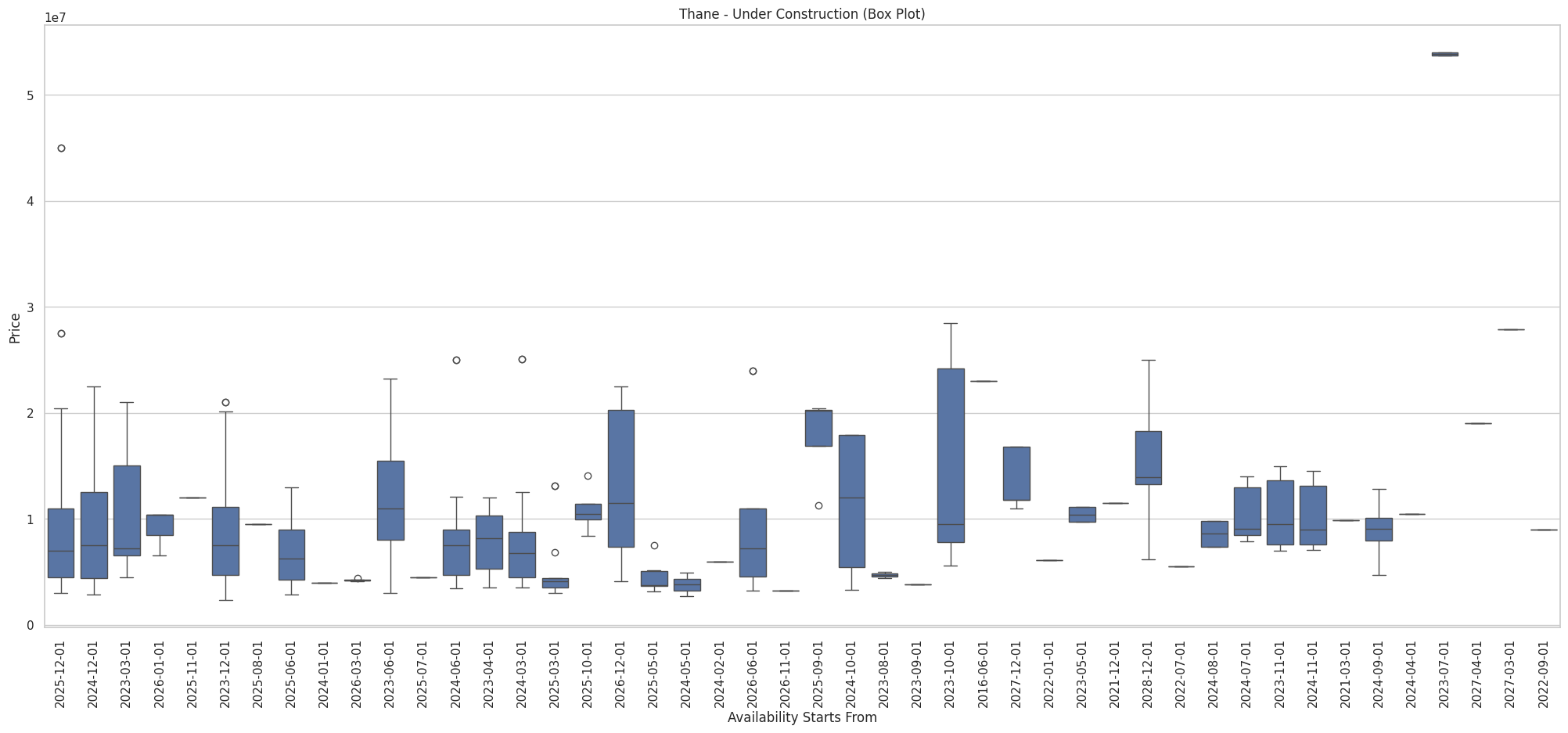
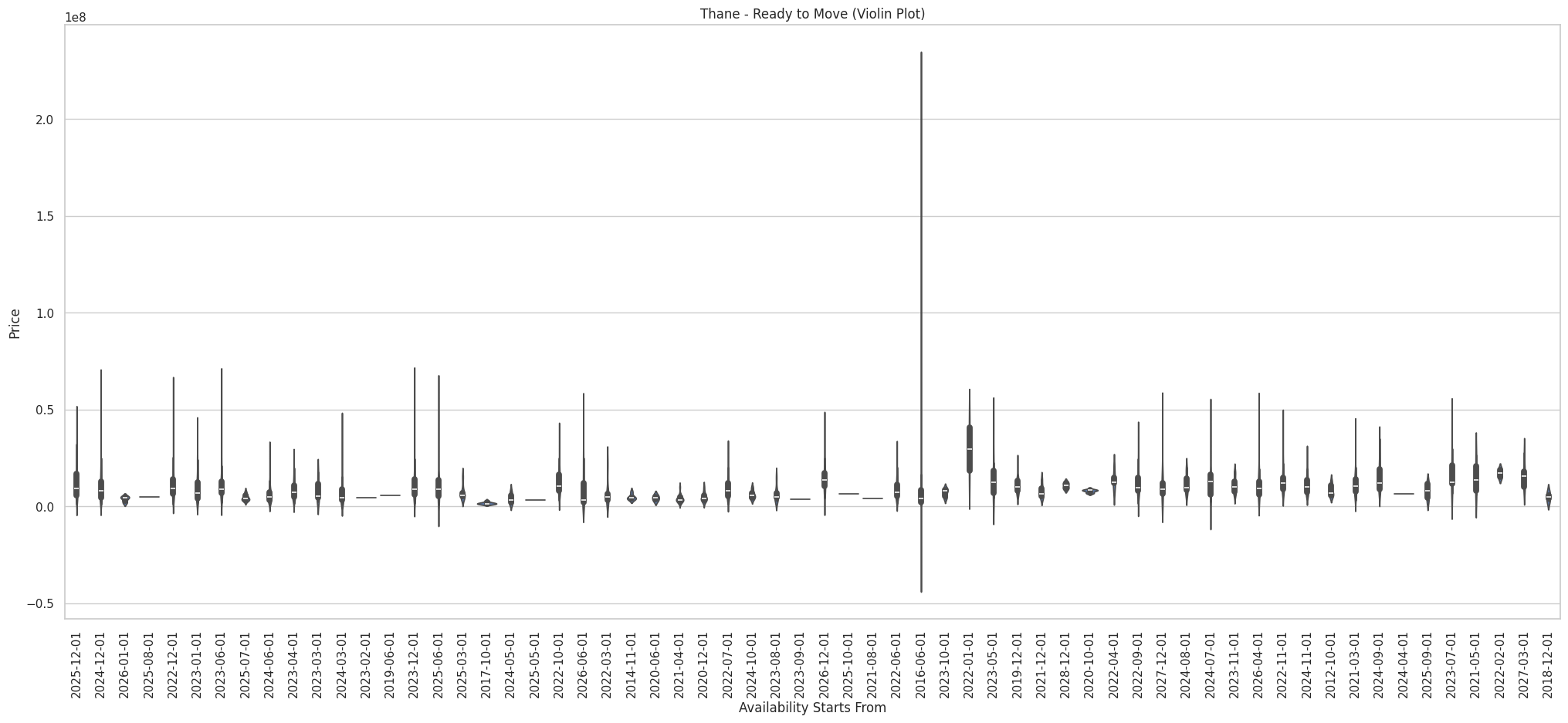
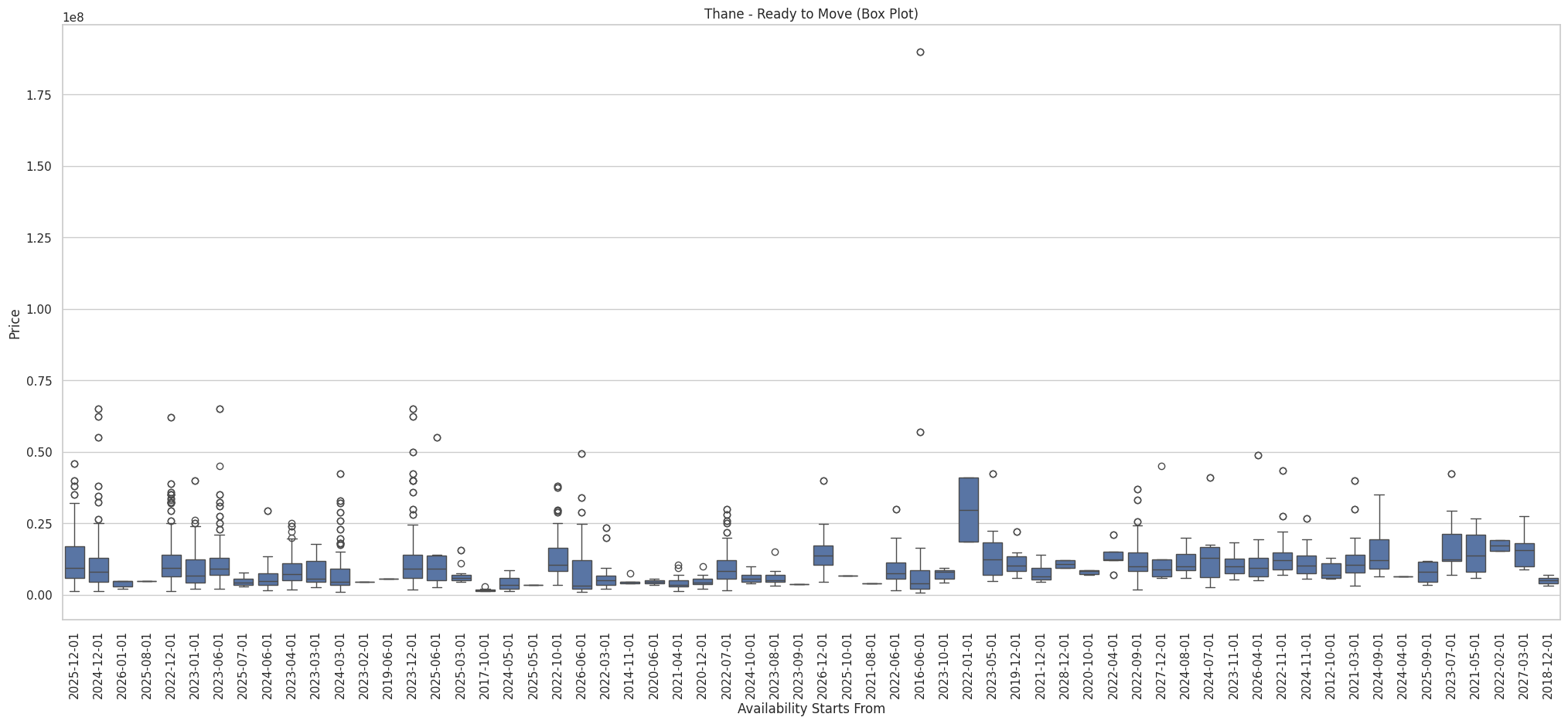
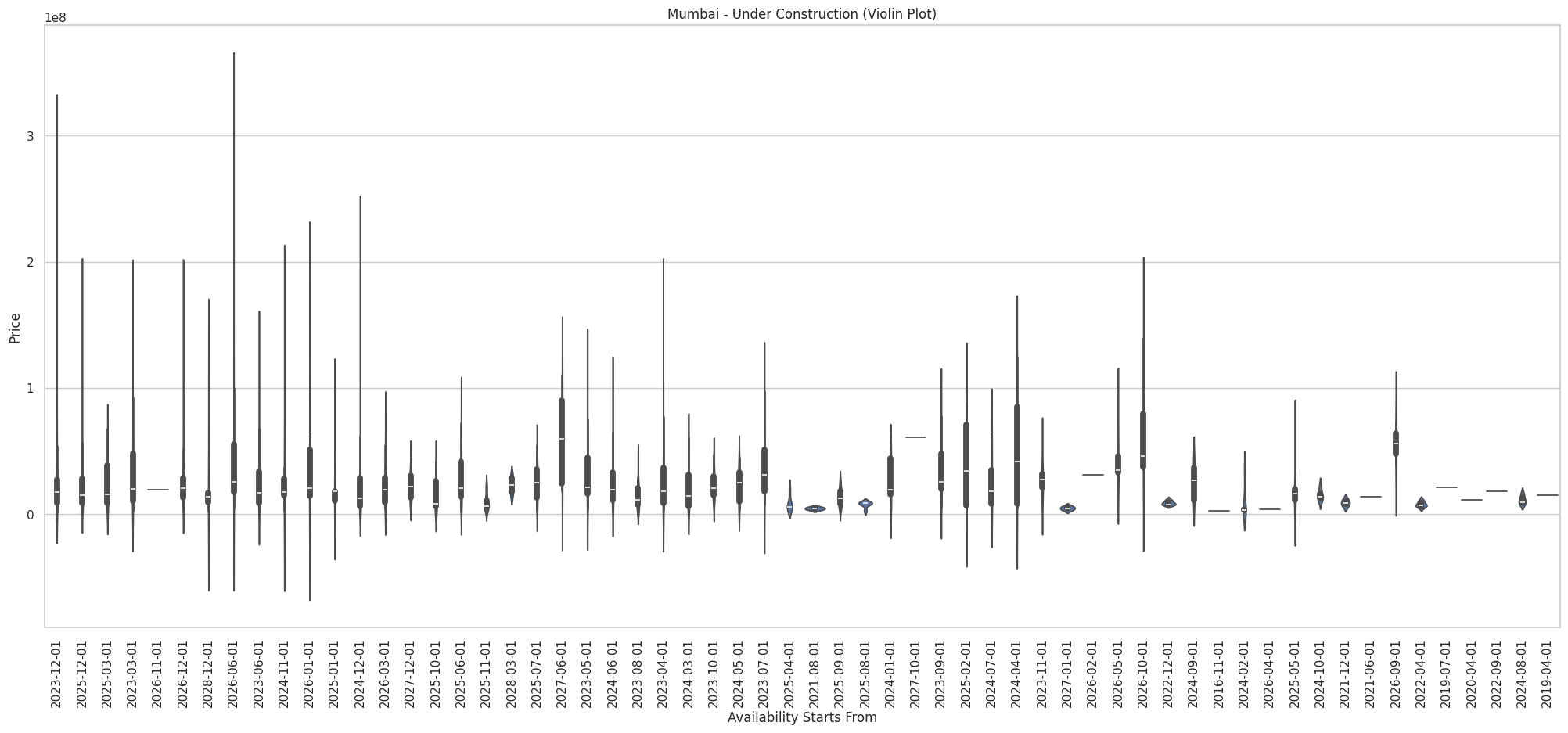
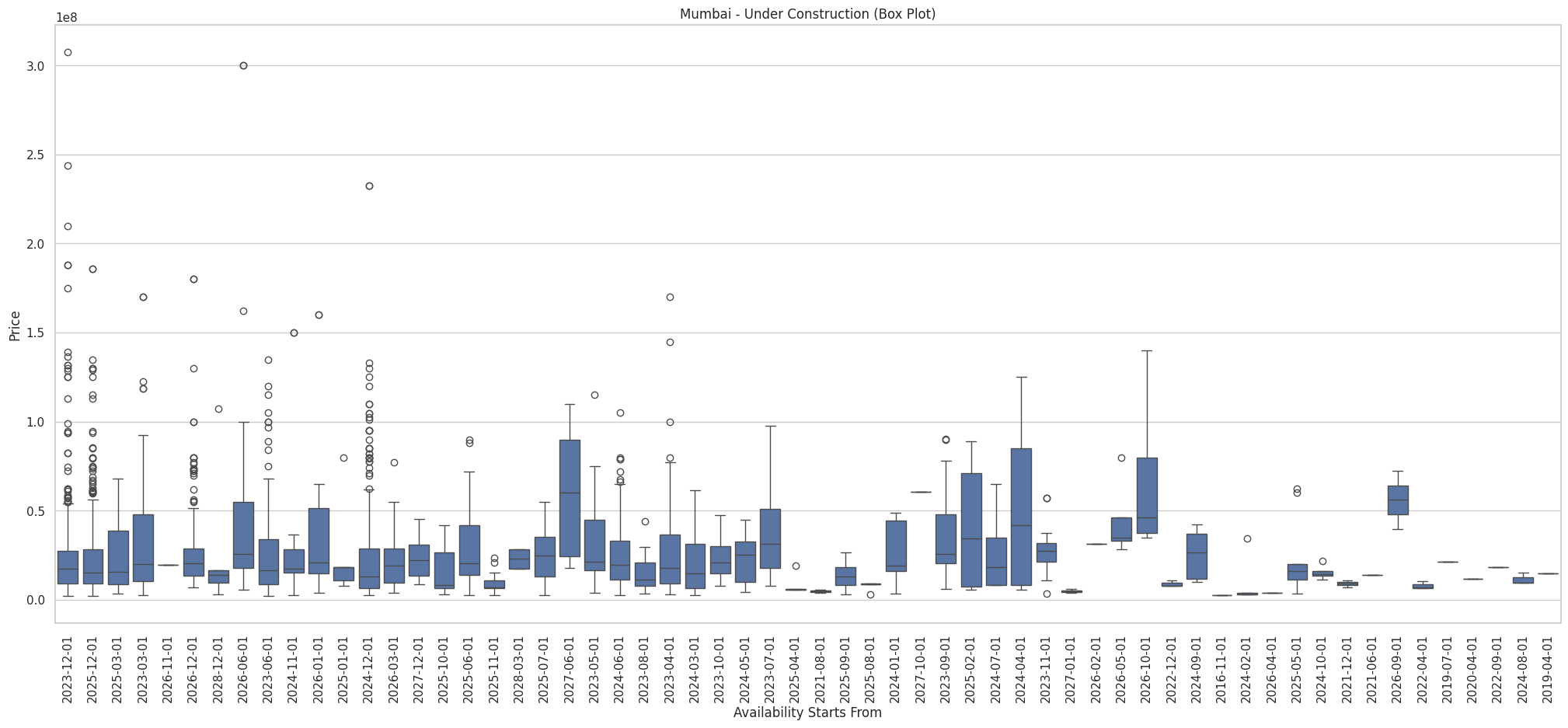
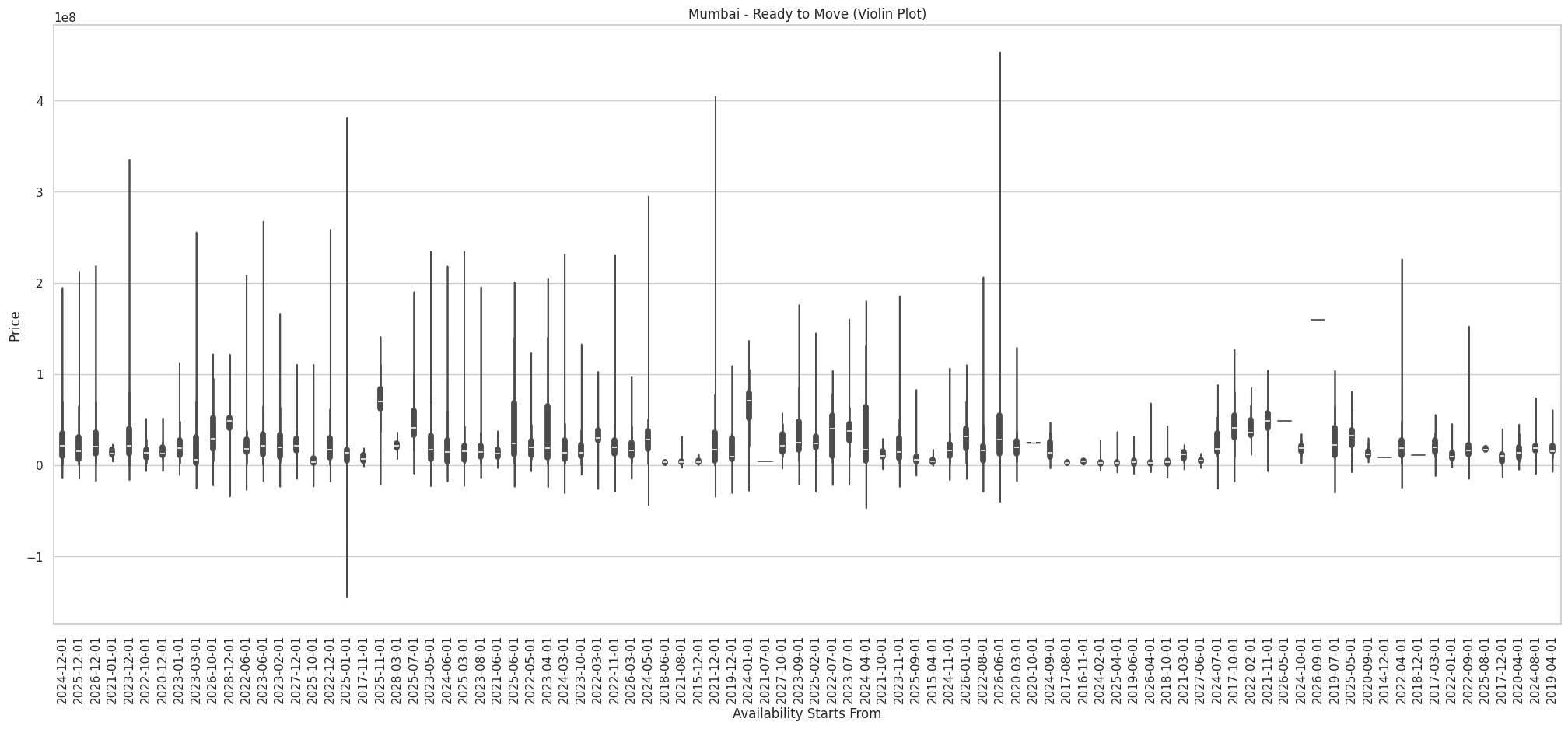
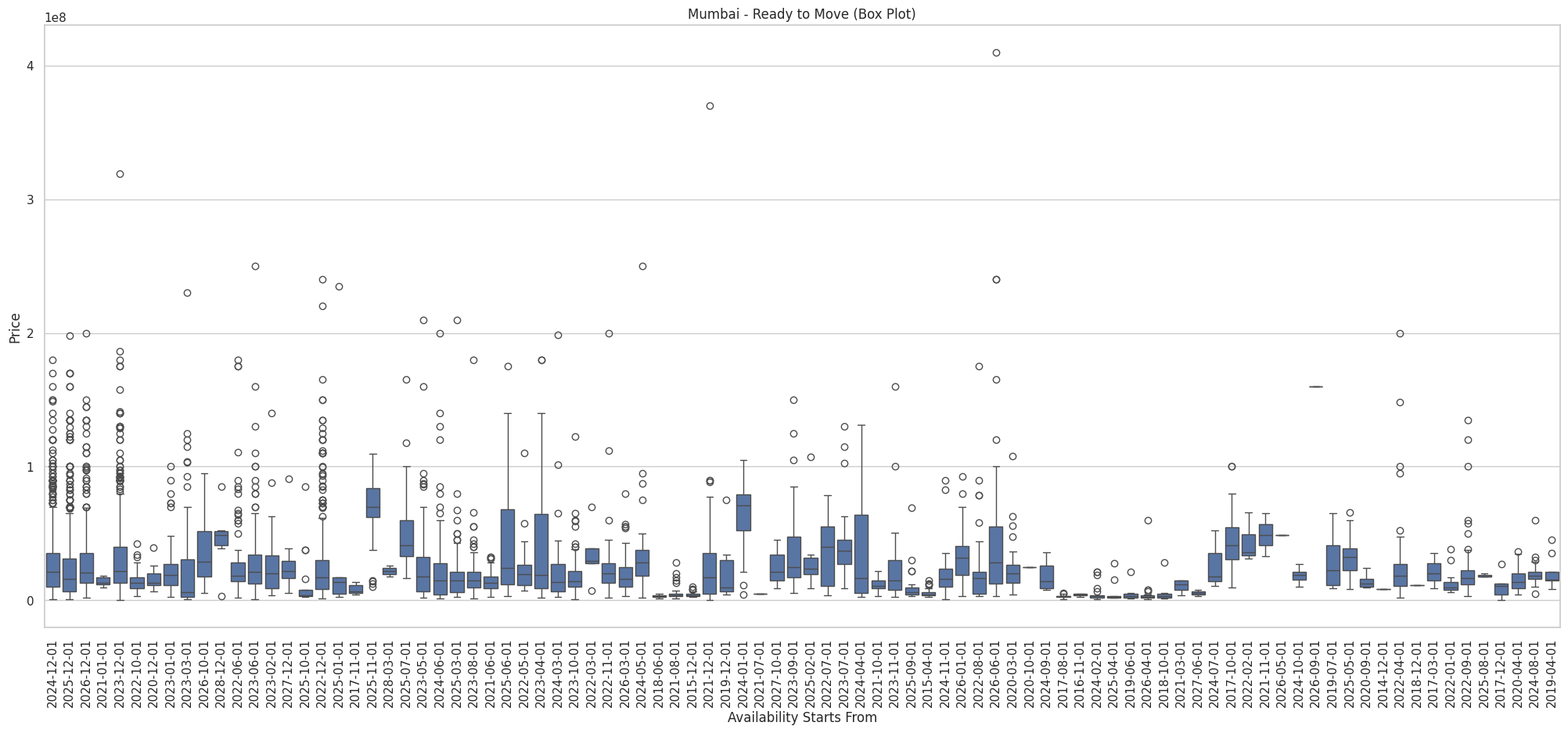
* **Data Source**: The analysis uses a subset of the property dataset focused on Mumbai and Thane, filtered to include only properties that are either 'Ready to Move' or 'Under Construction'.
* **Approach**: The property prices are compared based on the 'Possession Status' and the 'Availability Starts From' dates. Box plots and violin plots are generated to visualize the distribution of property prices for each category in both cities.

#### **Visualization and Key Observations:**

1. **Mumbai - Ready to Move**:
   * **Box Plot**: The box plot shows the distribution of property prices for 'Ready to Move' properties in Mumbai, categorized by the 'Availability Starts From' dates.
     + **Key Observations**:
       - Properties available for immediate possession (earlier dates) tend to have higher prices.
       - There is a wider range of prices for 'Ready to Move' properties, indicating a diverse market with various options.
       - Some outliers exist, indicating that certain properties are priced significantly higher than others in the same availability category.
   * **Violin Plot**: The violin plot provides a detailed view of the distribution, showing the density of property prices.
     + **Key Observations**:
       - The price distribution is skewed towards higher values, especially for properties with earlier availability dates.
       - There is a noticeable concentration of properties within certain price ranges, suggesting specific segments within the market.
2. **Mumbai - Under Construction**:
   * **Box Plot**: The box plot displays the price distribution for 'Under Construction' properties in Mumbai.
     + **Key Observations**:
       - Prices for 'Under Construction' properties tend to be lower compared to 'Ready to Move' ones, but with some overlap.
       - Properties with later availability dates (indicating longer construction times) tend to have a broader price range.
       - The presence of outliers suggests that some under-construction properties are priced competitively, possibly due to luxury features or prime locations.
   * **Violin Plot**: The violin plot for 'Under Construction' properties in Mumbai shows the price distribution.
     + **Key Observations**:
       - The distribution is more spread out compared to 'Ready to Move' properties, indicating varying investor confidence and potential for price appreciation.
       - There is less concentration in specific price ranges, showing a more speculative market.
3. **Thane - Ready to Move**:
   * **Box Plot**: The box plot for 'Ready to Move' properties in Thane highlights the price distribution.
     + **Key Observations**:
       - Similar to Mumbai, properties available for immediate possession in Thane tend to have higher prices.
       - The overall price range in Thane is lower compared to Mumbai, reflecting different market dynamics.
       - Outliers are present but less pronounced, suggesting a more stable market.
   * **Violin Plot**: The violin plot for 'Ready to Move' properties in Thane shows the price density.
     + **Key Observations**:
       - The price distribution is relatively compact, with less variation than in Mumbai.
       - The concentration of prices in specific ranges is more pronounced, indicating more homogeneity in the Thane market.
4. **Thane - Under Construction**:
   * **Box Plot**: The box plot for 'Under Construction' properties in Thane shows the price distribution.
     + **Key Observations**:
       - 'Under Construction' properties in Thane exhibit a narrower price range compared to Mumbai, reflecting lower speculative activity.
       - Prices are generally lower, with fewer outliers, indicating a more cautious market.
   * **Violin Plot**: The violin plot for 'Under Construction' properties in Thane.
     + **Key Observations**:
       - The price distribution is relatively uniform, with no significant skew towards higher values.
       - The market appears to be more predictable, with fewer surprises in pricing.

#### **Conclusions:**

1. **Price Differences Between 'Ready to Move' and 'Under Construction' Properties**:
   * **Mumbai**:
     + 'Ready to Move' properties generally command higher prices, reflecting the immediate utility and reduced risk for buyers.
     + 'Under Construction' properties offer a wider price range, indicating potential for future appreciation and investor speculation.
   * **Thane**:
     + The price difference between 'Ready to Move' and 'Under Construction' properties is less pronounced compared to Mumbai, suggesting a more stable market with less speculative activity.
2. **Impact of 'Availability Starts From' Dates**:
   * Properties with earlier availability dates tend to have higher prices, particularly in the 'Ready to Move' category, as they offer immediate possession and reduced uncertainty.
   * The variation in prices is more significant in Mumbai, where market dynamics are more complex, with a mix of luxury and mid-range properties.
3. **Market Dynamics**:
   * **Mumbai**: The market is more diverse, with significant opportunities for both immediate returns ('Ready to Move' properties) and long-term appreciation ('Under Construction' properties).
   * **Thane**: The market is more predictable, with less speculative activity, making it attractive for more risk-averse investors.



**Task-9:**

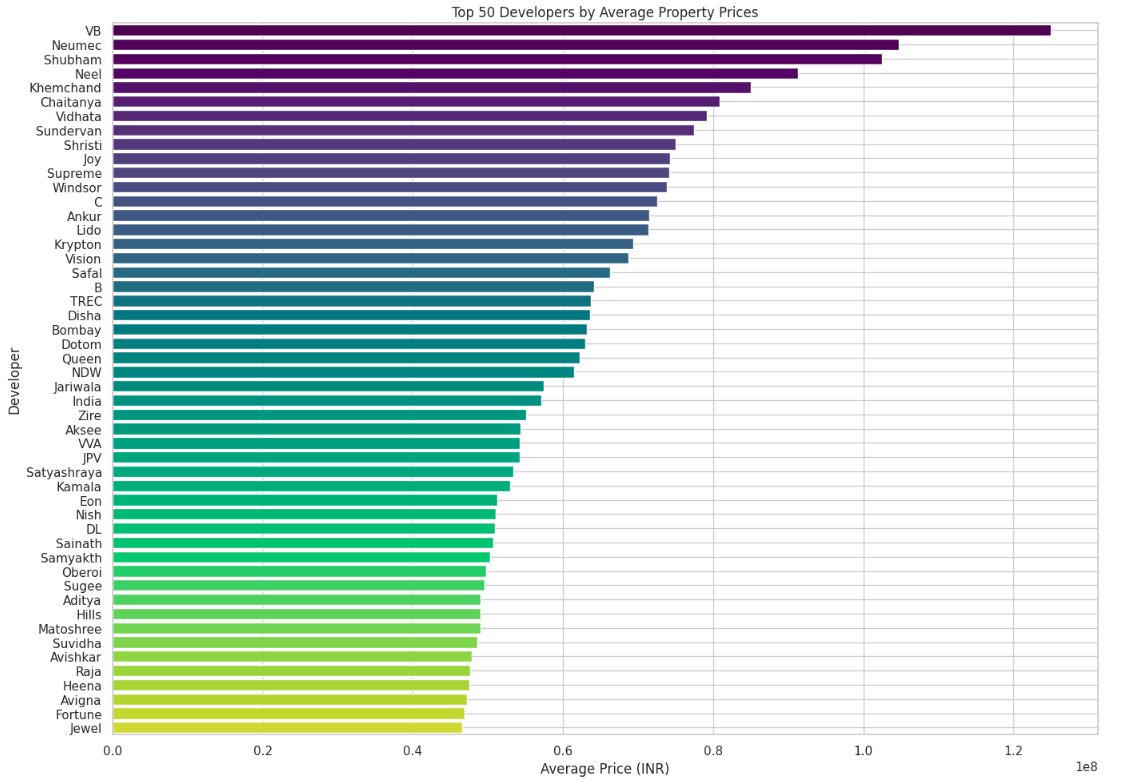
### **Data and Methodology**

#### **Data Preparation**

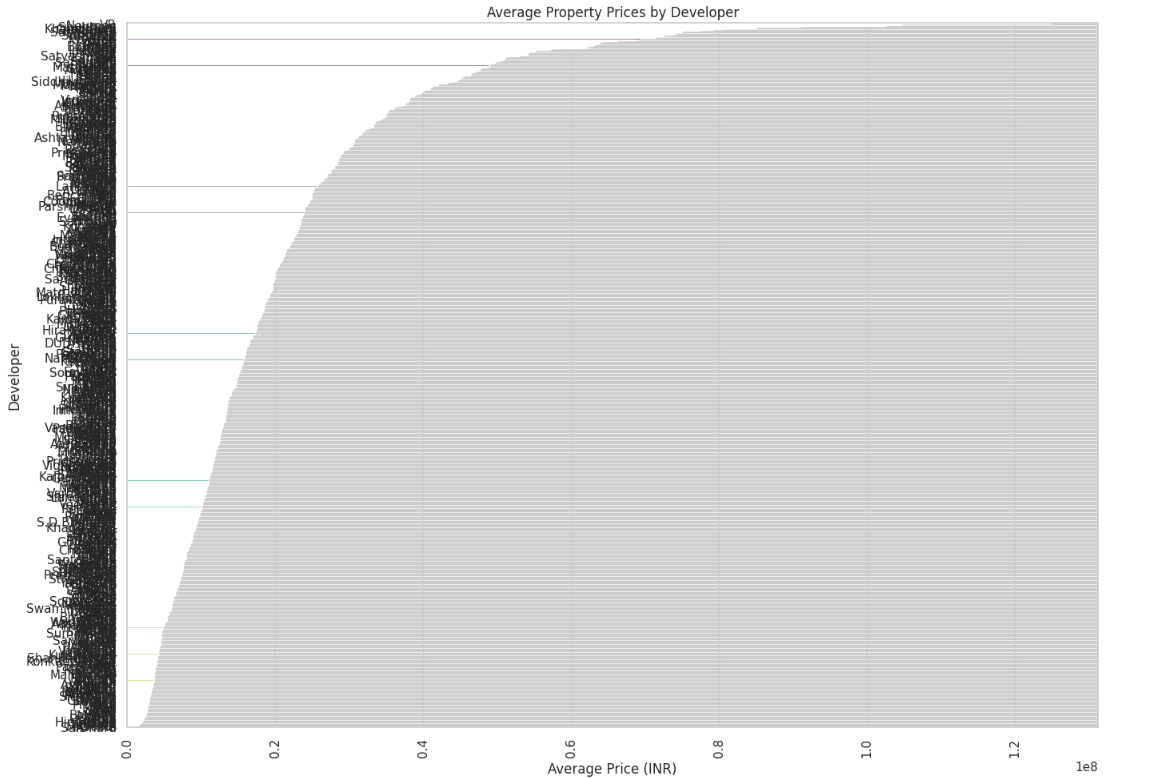
* **Dataset**: We used a dataset containing various property attributes, including Price, Developer, and several numerical features related to property characteristics.
* **Features** **of** **Interest**: We focused on numerical features (excluding Price) and analyzed how these features relate to average property prices for different developers.
* **Grouping and Aggregation:**
* Grouped the data by Developer.
* Calculated the mean values for Price and other numerical features.
* Sorted the developers based on average property prices to identify high-end developers.

**Visualizations:**

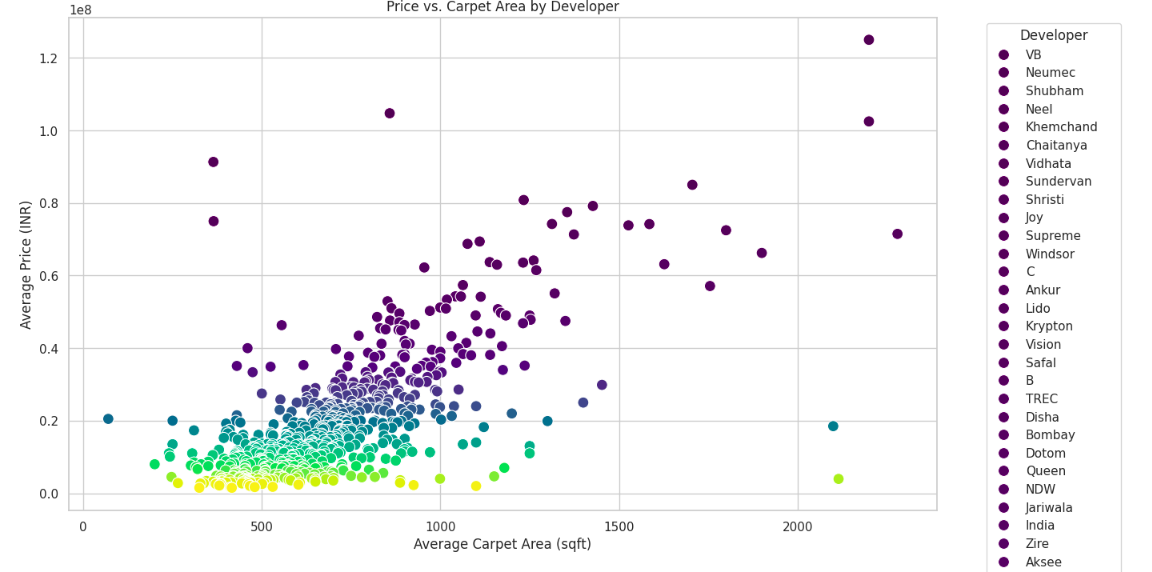
**Bar Plot of Average Property Prices by Developer**

****

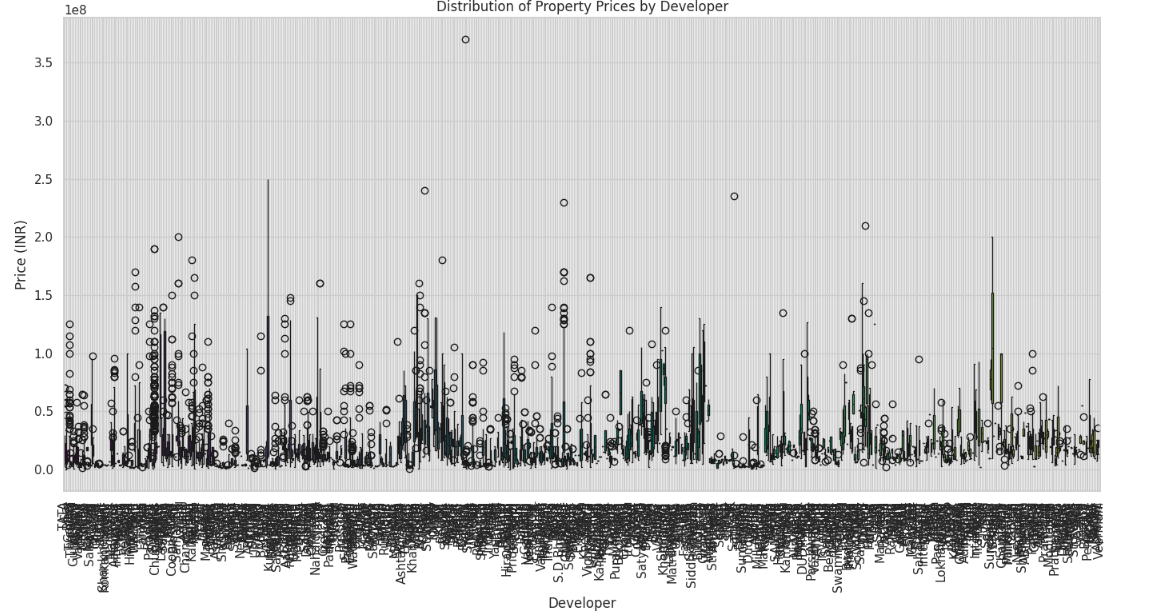
**Bar Plot of Average Property Prices by Developer (All Developers)**

****

**Scatter Plot of Price vs. Carpet Area by Developer**

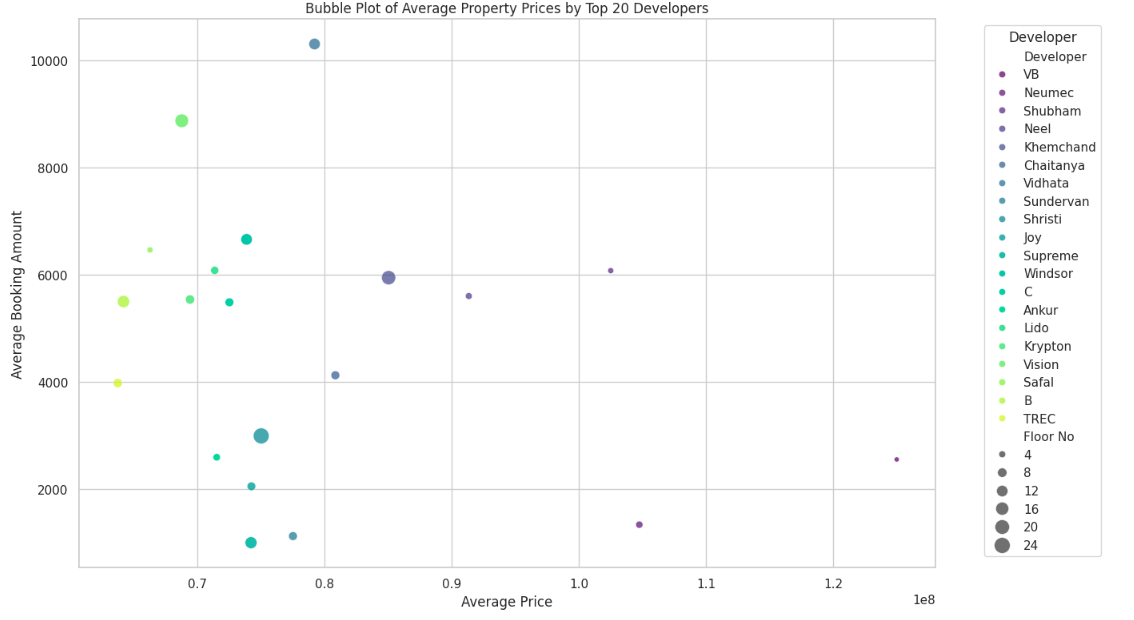
****

**Box Plot of Property Prices by Developer**

****

**Bubble Plot of Average Property Prices by Top 20 Developers**

Here we can change the index of numerical features to get the required plot. Below is hte example for Avg booking amount feature.

****

**Conclusion:**

* **High-End Developers**: The bar plots indicate which developers are associated with higher average property prices, helping identify premium developers.

We can see that VB developers have the highest avg prize.

* **Feature Relationships:** The bubble plot and scatter plot reveal how property prices relate to other features, such as carpet area, and how different developers compare.

From the above graph we can see the relationship between features, developers, average prize

* **Price Distribution:** The box plot provides insights into the variability and range of property prices associated with different developers, highlighting any disparities.