Insight Analysis Report

Apartment Rental Market in Colombo

Team Outliers – DXP07

DataXplore

15 February 2025

Table of Content

1. Abstract

The demand for rental properties in Colombo is increasing, prompting a data-driven approach to understand the rental trends. This study explores key factors influencing rental prices, such as apartment size, number of bedrooms, proximity to the city centre, furnishing status, and building type. Data pre-processing, exploratory data analysis (EDA), and statistical analysis are conducted to derive meaningful insights.

2. Methodology and Data Analysis

2.1 Data Overview

The dataset consists of apartment listings from various neighborhoods in Colombo, with the following features:

Number of Records: 252

Number of Features: 9

Features:

Apartment_ID: Unique identifier for each apartment listing

Neighborhood: Name of the neighbourhoods

Rental_Price: Monthly rental price

Size_in_Sqft: Size of the apartment in square feet

Distance_to_City_Center_KM: Distance to the Colombo Fort Station

Bedrooms: Number of bedrooms

Bathrooms: Number of bathrooms

Furnished: Whether the apartment is furnished

Building_Type: Type of the building

2.2 Data Cleaning

To ensure data quality and reliability, the following cleaning steps were performed:

- **Duplicate Removal:**

Two duplicate rows were identified and removed to prevent bias in analysis.

- **Inconsistent Formatting:**

  Categorical variables (Furnished and Building Type) had inconsistent casing (e.g., "Furnished", "furnished") These were converted to lowercase for uniformity.

- **Feature Removal:**

  - Apartment_ID was removed as it was a unique identifier and did not provide any significant analytical value.

- **Feature Encoding:**

  Label Encoding: Applied to Furnished, converting "Furnished" to 1 and "Unfurnished" to 0.

  One-Hot Encoding was implemented for the "Building_Type" and "Neighborhood" columns. Since these features contain multiple categorical values with no inherent order, One-Hot Encoding was chosen to prevent misleading ordinal relationships.

2.3 Handling Missing Values

Bedrooms Column:

Bedrooms was stored as an object data type instead of an integer. This issue was corrected. Missing values in Bedrooms were handled using **mode imputation** (most frequent value replacement).

2.4 Data Visualization

Univariate Analysis

- **Distribution of Apartment Sizes**
  - Apartments of nearly 600 sqft were the most frequent in the dataset.
  - A right-skewed distribution indicated that most apartments were smaller in size, with fewer larger apartments.

- **Rental Price Distribution & Outliers**
  - A box plot revealed the presence of outliers in Rental_Price.
  - Winsorization (Capping Method) was applied to handle extreme values while preserving meaningful variations.

- **Bedroom Count Analysis**
  - A majority of the apartments had **one bedroom**, followed by two-bedroom apartments.
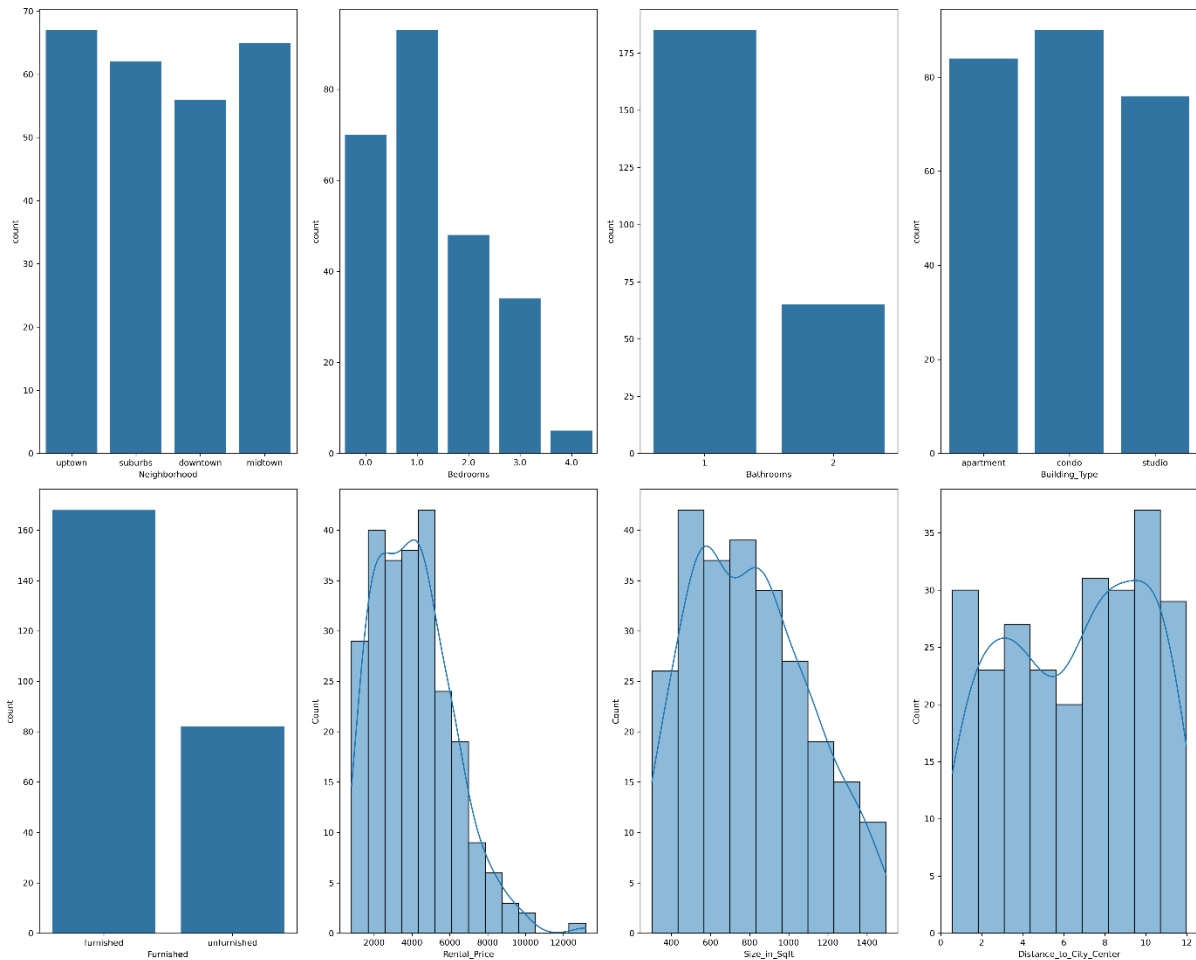

Figure 1

Bivariate Analysis

- Rental Price vs. Bedrooms:
  - Apartments with four bedrooms had the highest rental prices.
- Rental Price vs. Building Type:
  - Apartments and condos had significantly higher rental prices than studios.
- Rental Price vs. Distance to City Center:
  - A negative relationship was observed, where apartments closer to the city center had higher rental prices.

Reasoning about the colour Choice

All graphs were designed and visualized using a blue color scheme to ensure consistency, visual clarity, and ease of interpretation. Blue is a neutral, professional, and widely recognized color in data visualization, improving readability and contrast against white/light backgrounds.

2.5 Statistical Analysis

Correlation Analysis

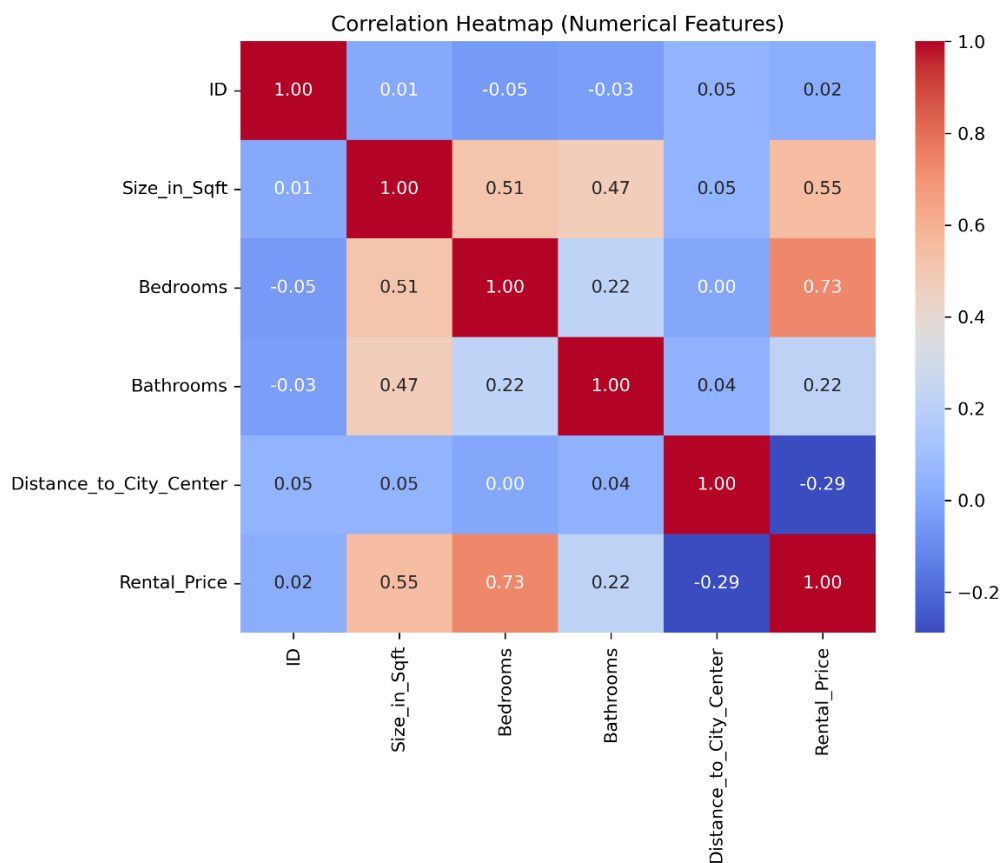Some of the key findings from correlation analysis are as follows

Figure 2

Feature Pair	Correlation Coefficient	Interpretation
Bedrooms & Rental Price	0.73	Strong Positive Correlation
Size in Sqft & Rental Price	0.55	Moderate Positive Correlation

Distance to City Center & Rental Price	-0.29	Weak Negative Correlation
Bedrooms & Size in Sqft	0.51	Moderate Positive Correlation

Table 1

3. Results

Rental prices in Colombo are primarily influenced by:

- **Number of bedrooms** (higher bedrooms = higher price).
- **Apartment size** (larger apartments have higher rents).
- **Proximity to the city center** (closer apartments are more expensive)
- **Size per Squarefeet** (Higher area = higher prices)

Furnished apartments are more common and tend to have higher rental prices.

Apartments and condos have significantly higher rental prices than studios.

4. Discussion

4.1 Limitations

- **Handling of Outliers:**

  Winsorization was used to cap extreme rental prices, but alternative approaches like logarithmic transformations could have been tested.

- **Lack of Time Series Data**

  The dataset does not include timestamps, making it difficult to analyze rental price fluctuations over time.

- **Potential Bias in Neighborhood Representation:**

  Some neighborhoods have a small number of listings, affecting generalizability.

4.2 Further Exploration

- **Predictive Modelling**
  A machine learning model (such as regression models or tree-based algorithms) could be built to predict rental prices.

- **Clustering analysis**

Neighborhoods could be clustered based on rental prices and apartment features to identify distinct market segments.

- **Feature Engineering**
  Additional features like amenities, security, public transport access, and seasonal trends could improve predictions.

4.3 Potential Use Cases

Use Case	Description
Real Estate Investment	Property investors can use insights to identify high-demand rental properties.
Price Benchmarking	Landlords can set competitive rental prices based on market trends.
Tenant Decision-Making	Renters can choose apartments based on their budget and preferred features.
Urban Planning	Authorities can use data to identify areas with rising demand for housing.

Table 2

Conclusion

In conclusion, our analysis of the apartment rental market in Colombo highlights key factors influencing rental prices, including apartment size, number of bedrooms, and proximity to the city center. Larger apartments, especially those with four bedrooms, tend to have higher rental prices, while condos and apartments are priced higher than studios. The negative correlation between rental price and distance from the city center suggests that apartments closer to Colombo Fort Station are more expensive. Additionally, furnished apartments are more common and may slightly impact rental rates. These insights provide valuable guidance for property companies to develop competitive pricing strategies and better understand market demand trends.