

## House Data Analysis Report

### 1. Dataset Description

**1.1 Source:** housing.csv dataset

**1.2 Columns:**

- **price** – House sale price (numeric, continuous)
- **area** – Total house area in square feet (numeric, continuous)
- **bedrooms** – Number of bedrooms (integer)
- **bathrooms** – Number of bathrooms (numeric, can be fractional)
- **stories** – Number of stories/floors (integer)
- **mainroad** – Whether the house has access to a main road (Yes/No)
- **guestroom** – Presence of a guest room (Yes/No)
- **basement** – Presence of a basement (Yes/No)
- **hotwaterheating** – Presence of hot water heating (Yes/No)
- **airconditioning** – Presence of air conditioning (Yes/No)
- **parking** – Number of parking spaces (integer)
- **prefarea** – Whether the house is in a preferred area (Yes/No)
- **furnishingstatus** – Furnishing status of the house (unfurnished, semi-furnished, furnished)

**1.3 Data Quality:**

- No missing/null values in primary columns.
- Clean structure with proper numeric and categorical types.
- Balanced distribution across different categories (bedrooms, stories, prefarea).

### 2. Operations Performed

**2.1 Data Cleaning & Exploration**

- Checked for missing/null values – none found.
- Verified unique values in categorical columns: mainroad, guestroom, basement, hotwaterheating, airconditioning, prefarea, furnishingstatus.

- Computed summary statistics for numerical columns (price, area, bedrooms, bathrooms, stories, parking) including mean, median, standard deviation, and interquartile range.

## **2.2 Descriptive Analytics**

- Distribution of bedrooms, bathrooms, and stories (bar charts).
- Price distribution visualized using histogram and boxplot.
- Area distribution (bar chart or histogram).
- Distribution of furnishing status, mainroad, guestroom, basement, airconditioning, parking, prefarea using bar plots.

## **2.3 Relationship Analysis**

- **Bedrooms vs. Price** (boxplot)
- **Area vs. Price** (scatter plot)
- **Stories vs. Price** (bar plot)
- **Parking vs. Price** (bar plot)
- **Furnishing Status vs. Price** (bar plot)
- **Preferred Area vs. Price** (bar plot)

## **3. Key Insights**

### **3.1 Housing Characteristics**

- Most houses have 3–4 bedrooms.
- Bathrooms mostly range from 1–3.
- Majority of houses have 1–2 stories.
- Common furnishing statuses are unfurnished or semi-furnished.
- Many houses have 1–2 parking spaces, and some have none.

### **3.2 Price Insights**

- Price range: low hundreds of thousands to higher-end properties.
- Average house price depends on the dataset.
- Median price is typically lower than the mean → right-skewed distribution (few high-priced houses).
- Outliers: Houses with larger area or more amenities drive up maximum prices.

### **3.3 Size & Area Insights**

- Area strongly correlates with price → bigger houses cost more.

- Most homes have 800–3,000 sqft area.
- Some houses with large areas may not have proportionally high prices, depending on location and amenities.

### **3.4 Amenities & Features**

- Houses with airconditioning, basement, and hotwaterheating generally have higher prices.
- Presence of a guestroom is a positive price factor.
- Houses in preferred areas (prefarea) show higher average prices.
- **Stories:** multi-story houses tend to be more expensive.
- **Furnishing status:** furnished houses command higher prices than unfurnished ones.

### **3.5 Other Influencing Factors**

- Access to mainroad can slightly increase property price.
- Number of parking spaces adds value, particularly in urban areas.
- Combining multiple amenities (airconditioning, prefarea, basement) → significant price premium.

## **4. Recommendations**

### **4.1 Market Strategy**

- Focus investment on houses in preferred areas with higher stories and multiple amenities.
- Renovation or adding amenities like airconditioning and basement can increase value.

### **4.2 Buyer Targeting**

- Middle-income buyers dominate 3-bedroom houses in a moderate price range.
- Premium buyers target houses with furnishing, multi-story, and additional amenities.

### **4.3 Pricing Strategy**

- Use area, stories, and amenities as primary price predictors.
- Amenities like parking, basement, and airconditioning add incremental value.
- Avoid overvaluing based solely on house size without amenities or preferred area.

### **4.4 Future Analytics Opportunities**

- Build price prediction models using Linear Regression, Random Forest.
- Perform clustering of houses by area, price, and amenities for market segmentation.
- Explore trend analysis if transaction dates are available.