**Project Summary for EDA on Raw Dataset**

* **Dataset Overview**:
  + Processed a dataset with 5000 observations and 16 features focused on property characteristics.
  + Based on overview and statistics used pandas, matplotlib and seaborn libraries for plotting, graphs for better estimation for better prospects.
* **Data Cleaning**:
  + Performed data type analysis and statistical description for better clarity.
  + Identified and handled missing values by dropping rows, reducing the dataset to 4370 rows.
  + Checked for duplicates; none were found.
* **Data Type Adjustment**:
  + Updated data types for specific columns to improve calculation efficiency:
    - Bathrooms: Converted from float to int.
    - Garage: Converted from float to int.
    - Fireplaces: Converted from float to int.
* **Data Visualization**:
  + Created a scatter plot to analyze the relationship between sold\_price and zipcode for identifying high-demand areas and outliers.
  + Generated a box plot to examine property trends based on construction year (year\_built) and sales (sold\_price).
* **Market Insight**:
  + Explored properties with the largest and smallest carpet areas for investment or selling strategies.
* **Output**:
  + Saved the cleaned dataset in CSV format for further analysis.