# OUTCOMES FROM DATASET

1. **Dataset Shape:** The dataset has 21,613 rows and 20 columns. Each row represents a real estate property, and each column corresponds to an attribute such as location, price, bedrooms, etc.

2. **Number of Cells:** The dataset contains a total of 432,260 cells or entries.

3. **Distribution of Property Prices:** The histogram of property prices shows that most properties have lower prices, with a peak around the lower price values. Higher-priced properties are less common.

Distribution of Property Prices
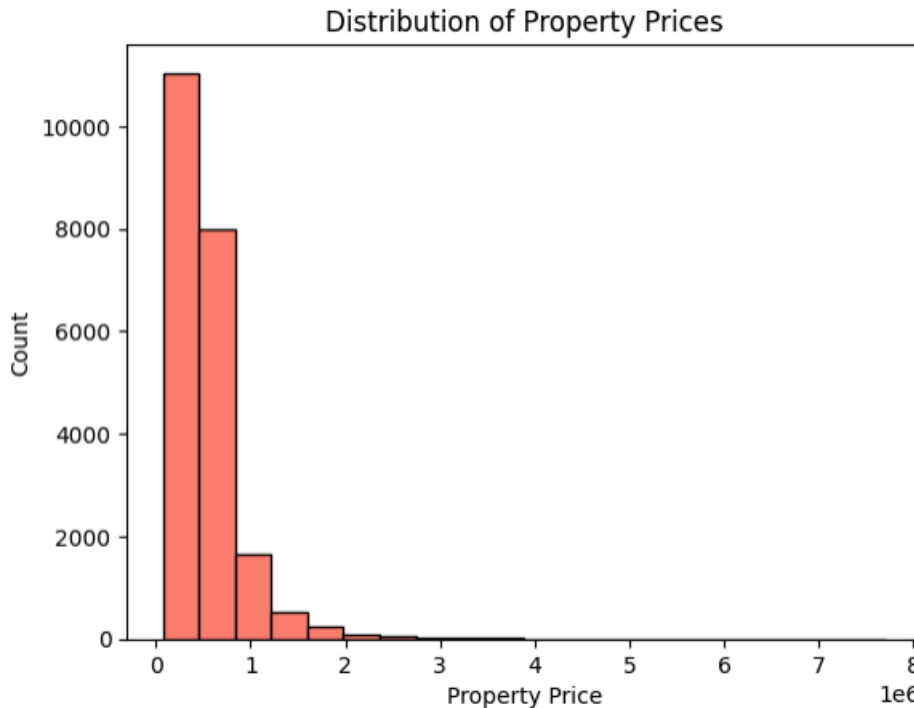
4. **Average Price by Number of Bedrooms:** The average price for different numbers of bedrooms is as follows:

    - 1-bedroom properties: ₹467,253

    - 2-bedroom properties: ₹401,373

    - 3-bedroom properties: ₹466,232

    - 4-bedroom properties: ₹635,420

    - 5-bedroom properties: ₹786,600

    - 6-bedroom properties: ₹825,521

    Observations: As the number of bedrooms increases, the average price tends to rise. However, there are exceptions, such as 1-bedroom properties having higher prices than 2-bedroom properties. Other factors like location and amenities may also influence property prices.

Average Price by Number of Bedrooms

5. **Waterfront View Impact on Property Value:** Non-waterfront properties have an average price of ₹531,564, while waterfront properties have an average price of ₹1,661,876. Having a waterfront view significantly increases the property value.
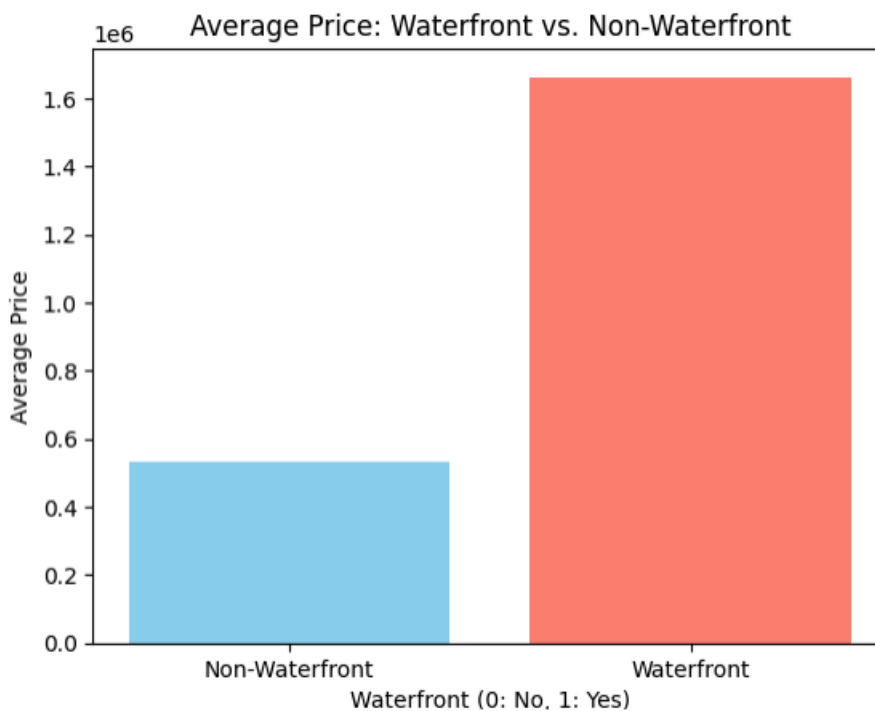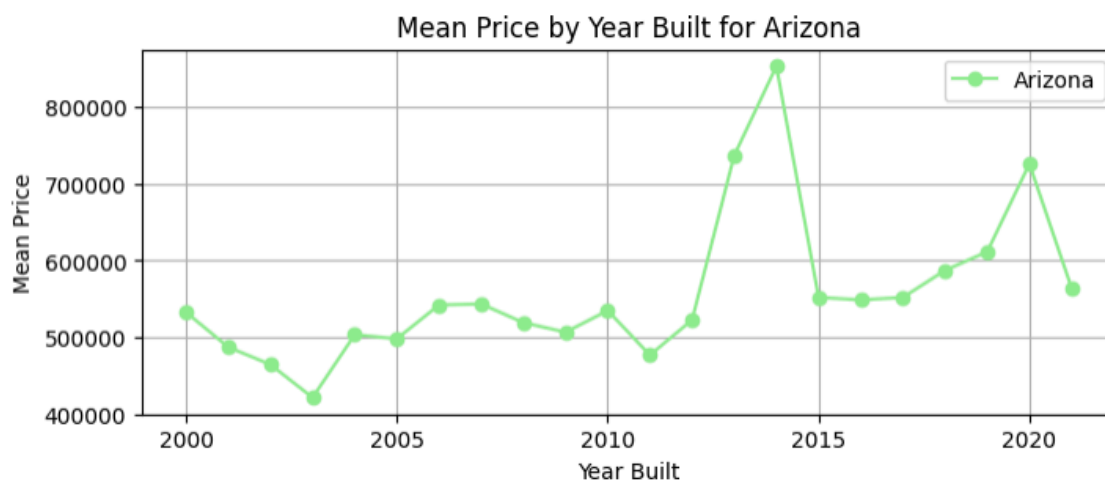
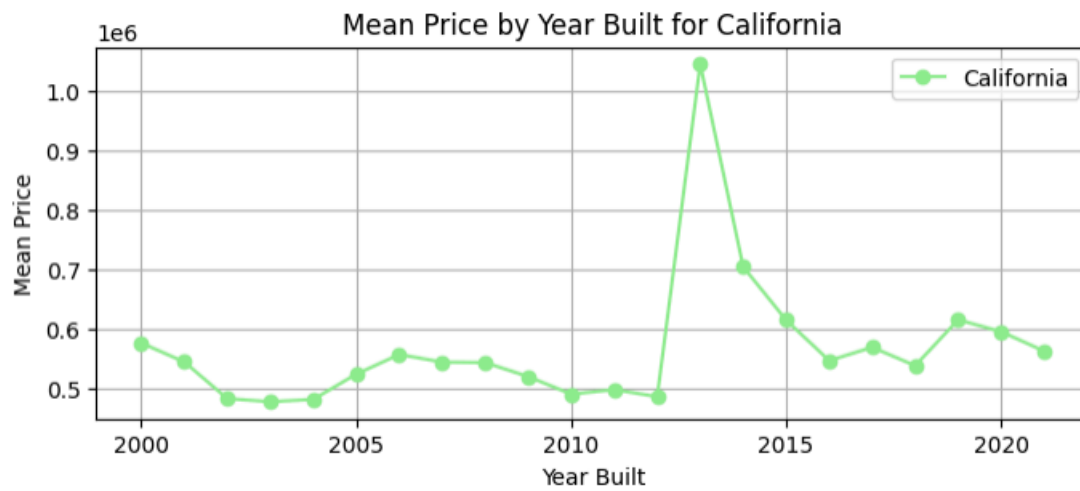Average Price: Waterfront vs. Non-Waterfront

6. **Average Property Price by Year of Construction:** The average property price varies across different years of construction. In 2015, the average price was high, while in 2002, it was relatively low. These trends in property prices over the years provide insights into market dynamics and economic factors.
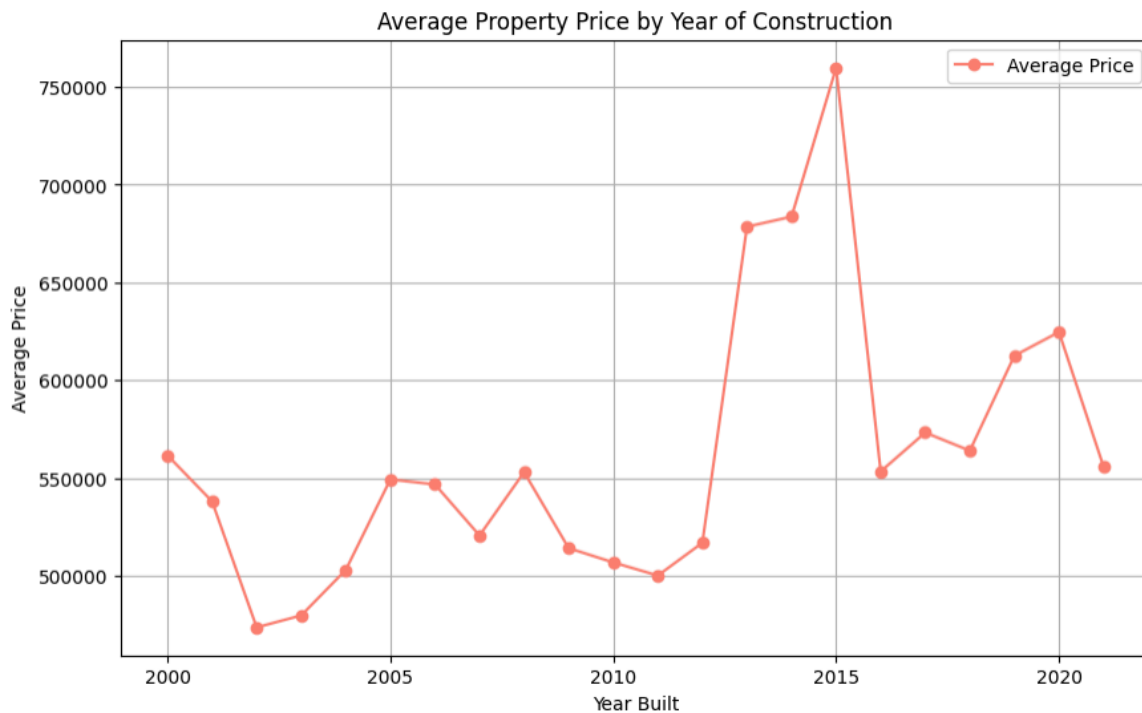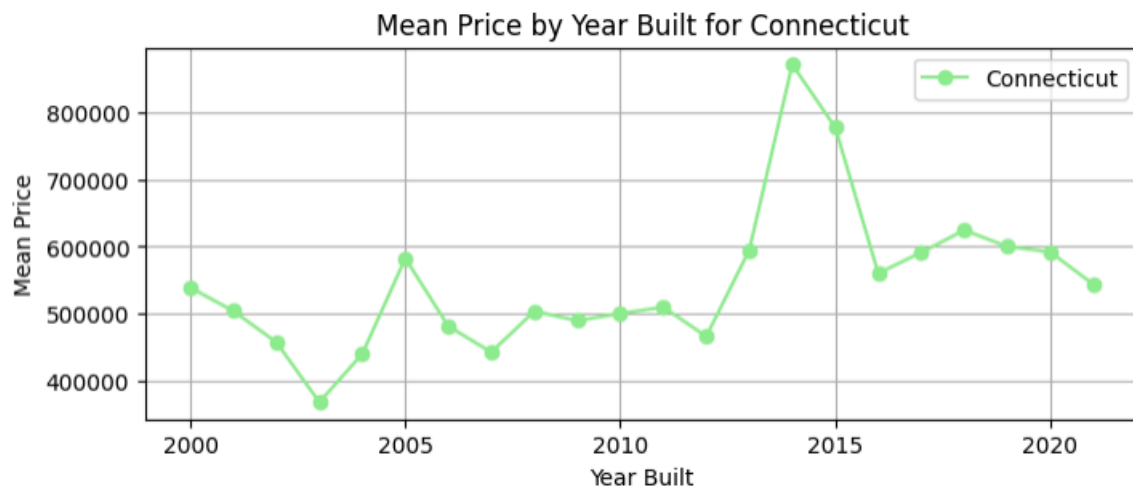

7. **Yearly Insights by Location:** The average property price by year of construction varies for different locations. For example, in Arizona, the average price was high in 2014 and low in 2003. Similar trends can be observed for California, Connecticut, Florida, Georgia, and Hawaii.
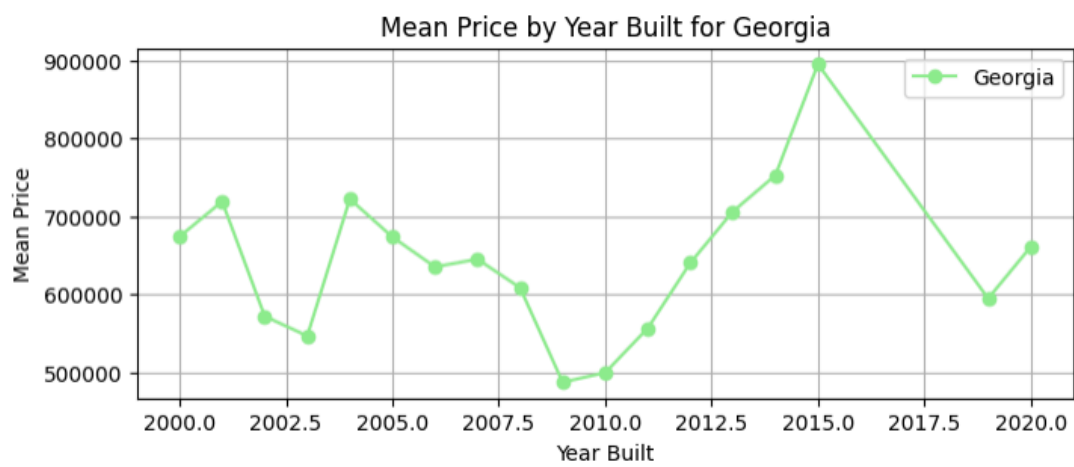
Mean Price by Year Built for California
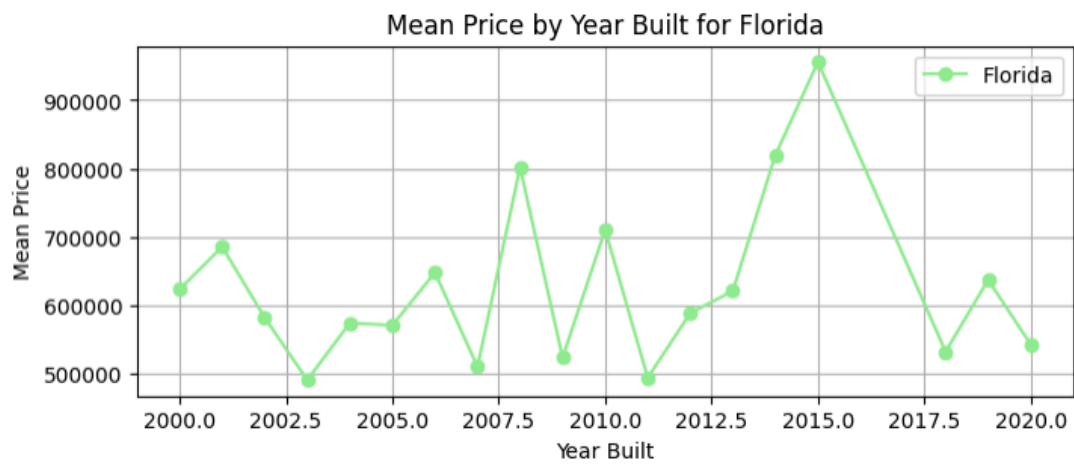
<Figure size 640x480 with 0 Axes>

Mean Price by Year Built for Connecticut

<Figure size 640x480 with 0 Axes>

Mean Price by Year Built for Florida

<Figure size 640x480 with 0 Axes>

Mean Price by Year Built for Georgia

<Figure size 640x480 with 0 Axes>

Mean Price by Year Built for Hawaii
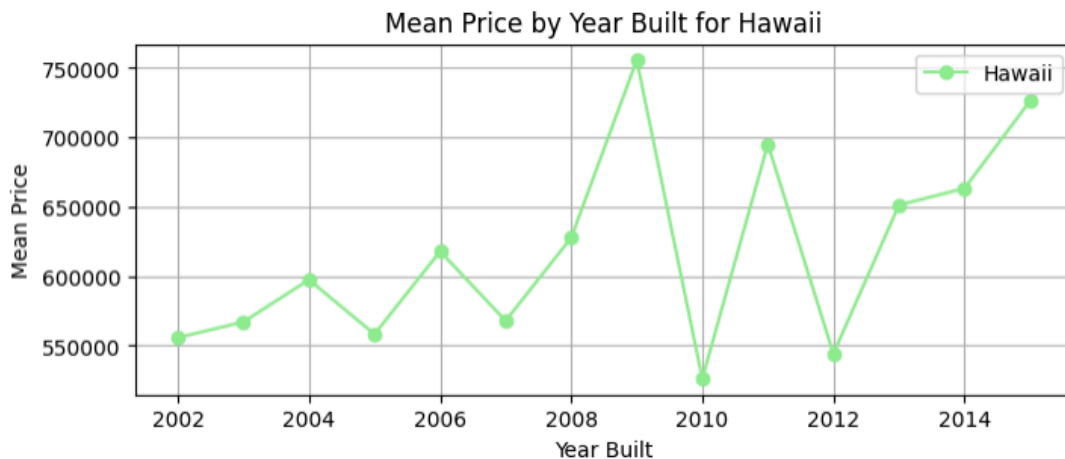
- In year 2014, Average Property Price by Year of Construction is high for city Arizonz and low in year 2003.
- In year 2013, Average Property Price by Year of Construction is high for city California and low in year 2003.
- In year 2014, Average Property Price by Year of Construction is high for city Connecticut and low in year 2003.
- In year 2015, Average Property Price by Year of Construction is high for city Florida and low in year 2003.
- In year 2015, Average Property Price by Year of Construction is high for city Georgia and low in 2009.
- In year 2009, Average Property Price by Year of Construction is high for city Hawaii and low in year 2010.

8. **Basement Influence:** Properties with a basement have a higher average price (₹622,374) compared to properties without a basement (₹486,884). Having a basement tends to increase the property value, as basements are often considered valuable additional space.

9. **Property Price vs. Square Footage of Living Space:** There is a strong positive relationship between the square footage of living space (sqft_living) and property price. As the square footage of living space increases, the property price tends to rise. The Pearson correlation coefficient of approximately 0.7 confirms this relationship. The distribution of sqft_living is skewed to the right, indicating that more properties have smaller living spaces.
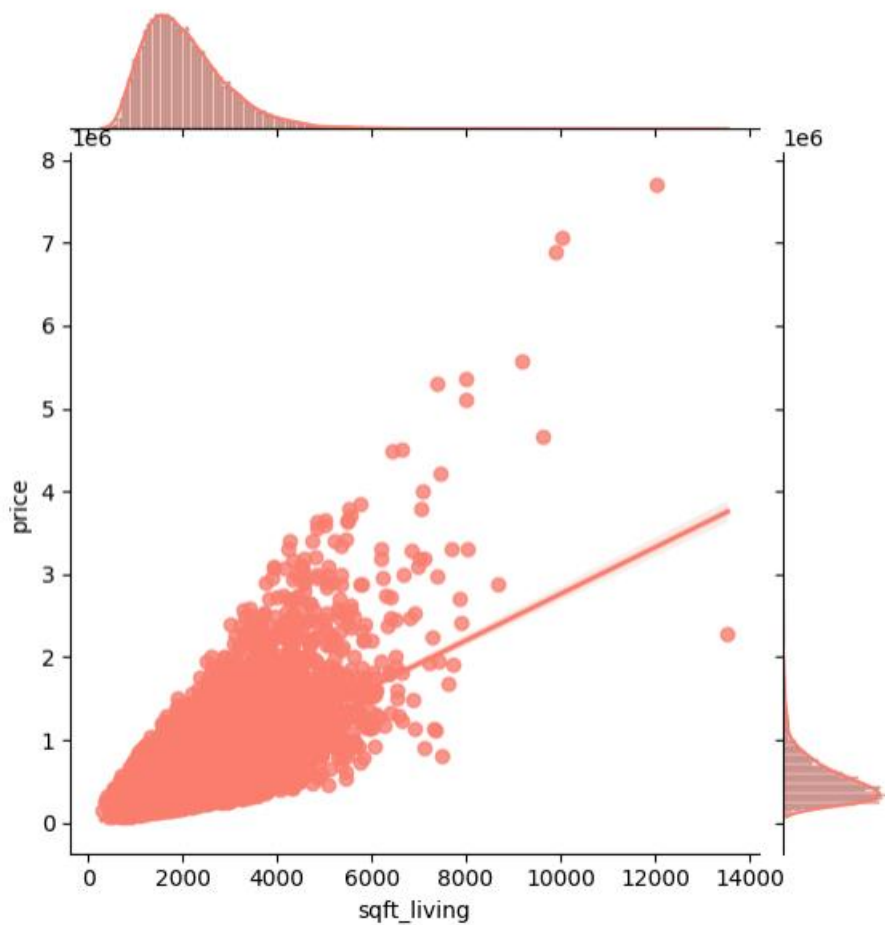
10. Spearman Correlation: The Spearman correlation coefficients and their significance for the given variables are as follows:

   - Price vs. Bedrooms: Moderate positive correlation (correlation coefficient ≈ 0.34), significant relationship (p-value ≈ 0.0).

   - Price vs. Floors: Moderate positive correlation (correlation coefficient ≈ 0.32), significant relationship (p-value ≈ 0.0).

   - Price vs. View: Weak positive correlation (correlation coefficient ≈ 0.29), significant relationship (p-value ≈ 0.0).

   - Price vs. Grade: Strong positive correlation (correlation coefficient ≈ 0.66), significant relationship (p-value ≈ 0.0).

 These correlations provide insights into how different features impact property prices. However, it's important to note that correlation does not imply causation.

11. **The correlation between the average square footage of the 15 closest neighbors (`sqft_living15`) and the selling price (`price`)** is weak. The scatterplot and Spearman correlation coefficient suggest that there is no clear linear relationship between these variables. Therefore, `sqft_living15` does not seem to be a strong predictor of property prices. Other factors, such as the actual square footage of the property itself (`sqft_living`), may have a more significant impact on price.
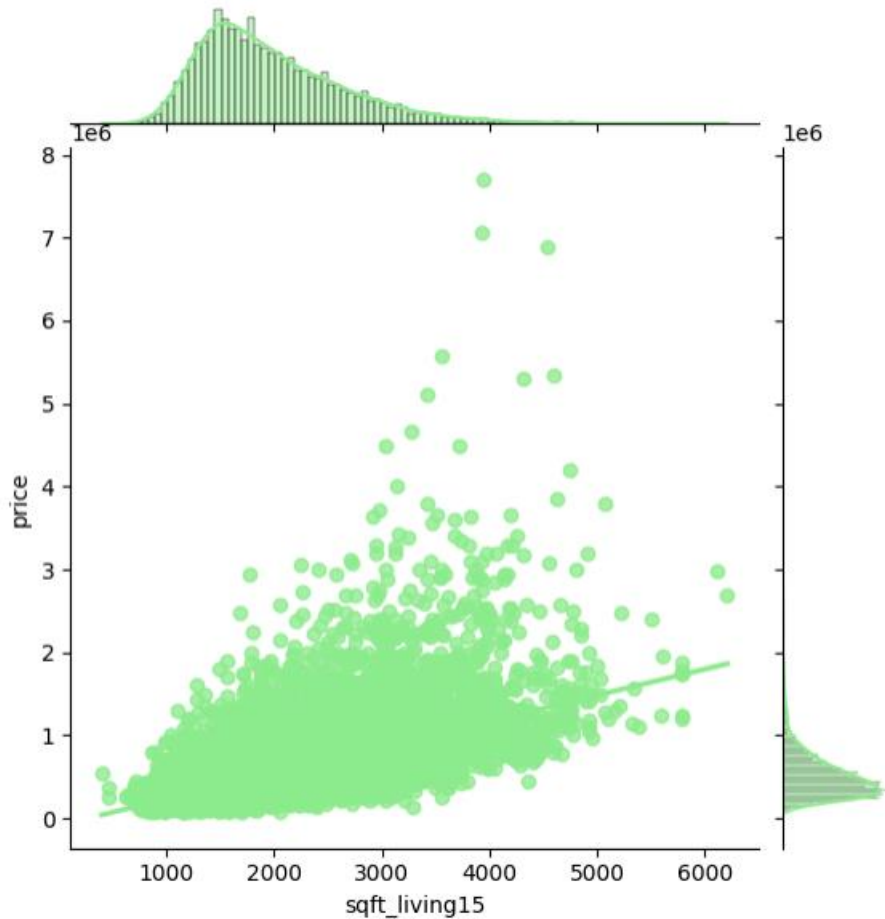
12. The scatterplot shows the distribution of selling prices against the average lot square footage of the 15 closest neighbors. There appears to be a weak or no correlation between these variables, as indicated by the lack of a clear linear trend. The average lot size of nearby properties does not significantly impact the selling prices. Other features, such as the square footage of the living space, may be more influential in determining property prices.