House Sales in King County, WA

Jonathan Fava
Data Analysis

Executive Summary

The dataset contains residential real estate sales in King County, Washington USA, which includes the city of Seattle, from May 2014 to May 2015. King County has compiled this dataset and published it for public usage with no copyright. The purpose of this analysis to educate potential home buyers on the housing market in Kings County as well as predict future values of real estate in this area.

Data Source

The data is open source and available on Kaggle
https://www.kaggle.com/datasets/harlfoxem/housesalesprediction

Data Profile

The data set has 23 unique columns and 21613 home sales(rows)

Data Limitations and Ethics

The data only consists of real estate sales from 2014 to 2015, so any analysis will reflect changes in value will be based on these years only and not a longer historical period.

Data Cleaning and Consistency Checks

- Removed excess numbers from date column values
- Created year and month columns
- Checked for missing and duplicate data
- Checked for mixed data types

Column Details

Variable	Description
id	Unique id for home
date	Date of sale
Year	Year of sale
month	Month of sale
price	Price of sale
bedrooms	Number of bedrooms
bathrooms	Number of bathrooms
sqft_living	Interior livable square footage
sqft_loft	Land square footage
floors	Number of floors
waterfront	Variable representing if there is a view of a body of water
view	An index from 0 to 4 representing how nice of a view the home has
condition	An index from 1 to 5 of the condition of the property
grade	Index of 1 to 13 of how well the property was designed
sqft_above	Square footage of the home above ground level
sqft_basement	Square footage of the home below ground level
yr_built	Year the home was built
yr_renovated	Year the home was last renovated
zipcode	Zip code of the home
lat	Latitude of the home
long	Longitude of the home
sqft_living15	Interior livable square footage of the 15 nearest properties
sqft_lot15	Land square footage of the 15 nearest properties

Questions

1. What are the statistics of homes sold in each zip code? Where are homes the most and least valuable?
2. What characteristics do the highest and lowest selling homes have?
3. Are waterfront homes significantly more valuable?
4. How does "grade" effect sale price?