

Data Source

Data Set: [House Sales in King County, USA](#)

Summary:

- This data set contains house sale prices for King County in Washington and includes homes sold during May 2014 through 2015.

Reason for using:

- I found this data set interesting to be used for learning and predicting house sales. There is ample data on date sold, price, zip code, number of beds and baths, and much more useful information needed for real estate and analytical purposes.

Data Profile

- [House Sales Data Profile](#)

Limitations & Ethical Considerations

Limitations:

- One entry showed a house having 33 bedrooms with 1.75 bathrooms. This value was changed to 3 bedrooms.
- There were 10 entries that listed zero bathrooms. These entries were removed since it is irregular to have houses without bathrooms.
- It was mentioned on Kaggle and after checking myself, there were inconsistencies with zip codes having similar longitudes and latitudes as

other zip codes. When addressing geographical location, longitude and latitudes were used since the zip codes were not reliable. The zip code column will be kept for now and may be excluded later if not useful in the data set.

Ethical Considerations:

- Each home is assigned to a unique id. Address, name of homeowner, and any PII are not included in this data set.
- **Measurement bias** - Although there is no duplicate data, as well as no missing data, there are still some levels of human error when collecting the data. For example, having a house with 33 bedrooms or zero bathrooms is unrealistic, thus inaccurate.
- **Exclusion bias** - It is possible that not all suitable homes were accounted for during the time of data collection. I can imagine some homeowners of specific race/ethnicities being excluded from the process or other homeowners requesting to be omitted from the data.

Key Questions

1. What is the relationship between price and square feet?
2. What is the relationship between price and number of bathrooms?
3. Is there a trend between price and location?
4. Is there a trend between price and view?
5. Is there a trend between price and waterfront?
6. Is there a trend between price and if the house was renovated?
7. Can home value be determined over time?