# House Sales in King County, USA

## Data Sourcing

## This data is publicly available open-source data. It was downloaded from Kaggle.com.

## [House Sales in King County, USA | Kaggle](https://www.kaggle.com/datasets/harlfoxem/housesalesprediction)

## Data Content:

* This dataset contains house sale prices for King County, which includes Seattle. It includes homes sold between May 2014 and May 2015.

## Data Relevance

## This data is relevant and crucial to answer the questions of this project.

## Why I chose this data.

* I chose this data as I gave not done retail data.

Data Cleaning

* deleted longitude and latitude to protect location and address
* Price column: convert format to 2 decimal places
* Deleted duethis information is not important to data analysissqft\_living15 - The square footage of interior housing living space for the nearest 15 neighbors  
  sqft\_lot15 - The square footage of the land lots of the nearest 15 neighbors.
* Convert dates to dd-mm-yy and remove time
* Check for missing data
* Check for duplicate data
* Check for mixed data type

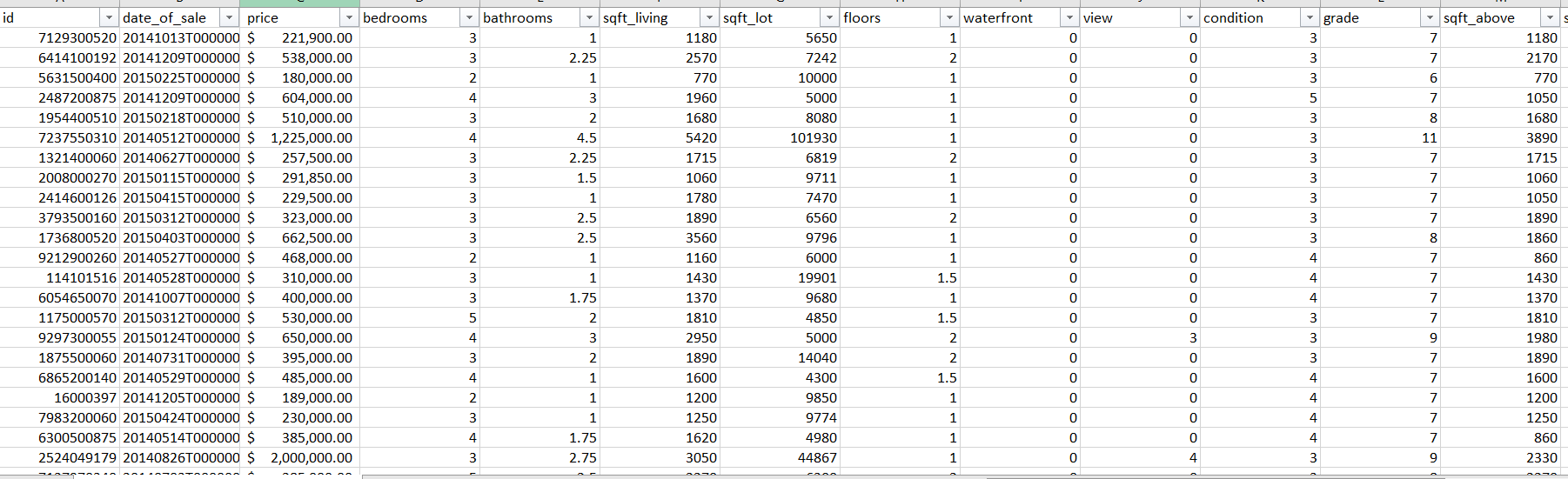
Meanings of columns

* sqft\_lot - Square footage of the land space
* view - An index from 0 to 4 of how good the view of the property was
* condition - An index from 1 to 5 on the condition of the apartment
* grade - An index from 1 to 13, where 1-3 falls short of building construction and design, 7 has an average level of construction and design, and 11-13 have a high quality level of construction and design.

**Consider limitations and ethics**

* Some of the columns have an index that will need to be reported in the data so anyone looking at it can understand the numbers represent.
* The data has zip code and longitude and latitude, deleted longitude and latitude to protect location and address.

Data Profile



Table

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

**Define questions to explore**

* Characteristics of the most expensive houses.
* Does the year of renovation have an impact on sales price, ex. Renovated in 2015 vs 1950.
* Of the following waterfront, view, condition, and grade have the highest impact on sale price.