**What does 2.5 bathrooms mean?**

**Bathrooms = 1, 1.25, 1.5, 1.75.**

In the US it means there are two “full” bathrooms and one “half” bathroom.

A full bathroom has a sink, toilet, and either a tub/shower combo or separate tub and shower.

A three-quarters bath has sink, toilet, and either a tub OR a shower, but not both.

A half-bath has a sink and a toilet - no tub or shower. These are often called “powder rooms” since a lady would often say “I need to powder my nose” instead of “I need to pee”.

A quarter-bath has only a toilet. These are uncommon now, but usually found in garages or workshops so Dad didn’t have to come back into the house all dirty/greasy just to use the bathroom.

**Floors = 1,1.5**

One and a half story home plans place all main living areas and usually also the master suite on the first floor. Additional rooms and/or loft space for children, visitors, and flex use are placed upstairs to help provide more square footage without expanding the size of the home's footprint.

**About this file**

1. Executive Summary

Online property companies offer valuations of houses using machine learning techniques. The aim of this report is to predict the house sales in King County, Washington State, USA using Multiple Linear Regression (MLR). The dataset consisted of historic data of houses sold between May 2019 to May 2020.

1. Problem Statement

 “Predict the sales of houses in King County with an accuracy of at least 75-80% and understand which factors are responsible for higher property value - $650K and above.”

1. Dataset Description
The dataset consists of house prices from King County an area in the US State of Washington, this data also covers Seattle. The dataset was obtained from Kaggle*. This data was published/released under CC0*: Public Domain. Unfortunately, the user has not indicated the source of the data. Please find the citation and database description in the Glossary and Bibliography.
♣ The dataset consisted of 21 variables and 21613 observations.
Due to the “Grade” variable which is based on a grading system specific to King County, it seems likely that the data originated from an official source. On the other hand, the existence of the “View” variable relates to the number of viewings the property received. This suggests that the data originated from a real estate agent/company.
In either case, there is no reason to call into question the general accuracy of the data because values such as prices, area etc. do not appear random. Nevertheless, some caution is appropriate before applying models trained on this data to more general cases.

Variables Description Data Type
id a notation for a house Numeric
date Date house was sold String
price Price is prediction target Numeric
bedrooms Number of Bedrooms/House Numeric
bathrooms Number of bathrooms/bedrooms Numeric
sqft\_living square footage of the home Numeric
sqft\_lot square footage of the lot Numeric
floors Total floors (levels) in house Numeric
waterfront House which has a view to a waterfront Numeric