

## HOUSING DATASET

Source:

<https://www.kaggle.com/harlfoxem/housesalesprediction/downloads/housesalesprediction.zip/1>

### Dataset description:

Variable	Description
price	house price in US dollars (prediction target)
bedrooms	number of Bedrooms/House
bathrooms	number of bathrooms/House
sqft_living	square footage of the home
sqft_lot	square footage of the lot
floors	total floors (levels) in house
waterfront	house which has a view to a waterfront
view	has been viewed
condition	how good the condition is ( Overall )
grade	overall grade given to the housing unit, based on King County grading system
sqft_above	square footage of house apart from basement
sqft_basement	square footage of the basement
yr_built	built Year
yr_renovated	year when house was renovated
zipcode	zip
lat	latitude coordinate
long	longitude coordinate
sqft_living15	living room area in 2015(implies-- some renovations). This might or might not have affected the lotsize area
sqft_lot15	lotSize area in 2015(implies-- some renovations)
date	year house was sold
years_built	Age of the house at the moment of sale in years
years_renov	How many years passed after a renovation before the sale

The dataset has been modified – years\_built and years\_renov were created and added to the dataset. Years\_built = date – yr\_built. Years\_renov = 0 if yr\_renovated = 0; otherwise, Years\_renov = date – yr\_renovated. The ID variable has been removed.