# Title: Housing Prices and Property Attributes

Niels Metselaar – s1119113

Rostyslav Redchyts – s1113227

## Aim:

Housing prices are crucial for both residents and property owners. This study aims to identify which attributes of a property significantly determine its price and to precisely predict these prices based on the investigated attributes. The primary goal is to conduct a statistical analysis to answer questions such as, “What possible amenities should one do without to afford a cheaper house?” or “How much will I get for my home given certain characteristics?”

## Statement of Hypotheses:

**H0 (Null Hypothesis):** The number of bathrooms and bedrooms, living space square footage, lot area, basement presence, total area above ground, number of floors in the house, view quality, house condition, grade of the building’s construction and design, and the years since it was built do not predict housing prices.

**HA (Alternative Hypothesis):** The number of bathrooms and bedrooms, living space square footage, lot area, basement presence, total area above ground, number of floors in the house, view quality, house condition, grade of the building’s construction and design, or the years since it was built can predict housing prices.

## Data:

To address these questions, we will use data from approximately 22,000 properties (before cleaning) from the real estate markets in King County, Washington State, USA, for the years 2014-2015. This dataset provides detailed information about various attributes of the properties along with their total prices, offering an excellent opportunity for statistical analysis.

*The data can be found at the following URL:* [*House Sales in King County, USA*](https://www.kaggle.com/datasets/harlfoxem/housesalesprediction/data)

## Population:

*Observational Units*: The observational units are individual real estate properties (houses) in the King County area.

We will attempt to generalize our findings to the population of all housing properties in the USA for the years 2014-2015. The total number of housing units in the USA in 2014 was approximately 134.32 million, including apartments, which are not covered in our dataset. Therefore, the exact number of houses is difficult to estimate.

*Response Variable:*

1. Price of the property (USD)

*Explanatory Variables:*

1. Number of bathrooms (bathrooms
2. Number of bedrooms (bedrooms)
3. Square footage of living spaces (sqft\_living)
4. Square footage of the lot (sqft\_lot)
5. Square footage of the basement (sqft\_basement)
6. Square footage above ground level (sqft\_above)
7. Number of floors (floors)
8. View quality (scale of 0-4) (view)
9. Condition of the house (scale of 1-5) (condition)
10. Year built (yr\_built)
11. Year renovated (yr\_renovated)
12. Grade of building’s construction and design (scale of 1-13) (grade)