**Question of Interest**

We are interested in the house price in the U.S, and we decide to build a model that predicts house price based on several attributes of the house.

**Data Sources**

The dataset is from a Kaggle, the link is posted here: <https://www.kaggle.com/harlfoxem/housesalesprediction>

This dataset collects 21,613 records of house sales between May 2014 and May 2015 for King County, which located in US state of Washington. King County includes Seattle city, which is the biggest city in the state.

**Variables will be used**

Dependent variable:

|  |  |
| --- | --- |
| Variable | Description |
| Price | Price of each home sold, with a range from $75,000 to $7,700,000 |

Independent variables:

|  |  |
| --- | --- |
| Variable | Description |
| date | Date of the home sale |
| bedrooms | The number of bedrooms |
| bathrooms | The number of bathrooms |
| sqft\_living | Square footage of the apartments interior living space |
| sqft\_lot | Square footage of the land space |
| floors | The number of floors |
| waterfront | A dummy variable for whether the apartment was overlooking the waterfront or not |
| 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 |
| sqft\_above | The square footage of the interior housing space that is above ground level |
| sqft\_basement | The square footage of the interior housing space that is below ground level |
| yr\_built | The year the house was initially built |
| yr\_renovated | The year of the house’s last renovation |
| lat | Latitude |
| long | Longitude |
| sqft\_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 |

**Expected Findings**

* Observe upward trend of house price over time
* Square footage would impact the house price significantly
* The higher index of view, condition or grade may lead to higher house price