# King County House Analysis



#### Presented By: Group 3.1

- Evah Wangari
- Dorine Langat
- Grace Nyongesa
- Maureen Anduuru
- Stella Kitur
- Vivian Adhiambo

### **Business Understanding**

#### **Business Overview**

- Real estate is one of the most important sectors of any economy.
- Understanding the key drivers of housing prices can provide valuable insights for both buyers and sellers in the market.
- In this project, we analyze a data set of house sales in a northwestern county to identify the factors that influence housing prices in the area.

### Challenges

- 1. The agency currently lacks a reliable method for predicting the impact of home value. Therefore, the agency is unable to provide accurate advice to homeowners about the potential return on investment.
- 2. Houses that stay longer in the market before being sold depreciate in value over time
- 3. Housing affordability: some houses were over valued while others were undervalued as compared to the normal market rate.

### Proposed Solutions

- 1. Develop a reliable method for predicting the impact of home improvements on the estimated value of a home.
- 2. To prevent houses from depreciating in value due to longer time in the market, the real estate agency can work with homeowners to set realistic prices based on market trends, property condition, and location. The agency can also help homeowners to prepare their homes for sale by renovating the property.
- 3. Provide homeowners with a list of home improvements that can increase the value of their homes, and offer advice on which improvements will provide the greatest return on investment.
- 4. Conduct market research and analysis to determine the appropriate value for each home based on its location, size, and other factors.

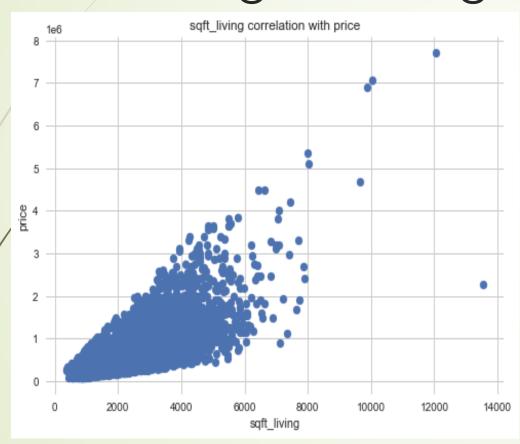
### Problem Statement

The real estate agency is unable to **accurately** identify pricing factors, analyse trends in house prices, and identify undervalued properties, resulting in unreliable information for buyers and sellers.

## Objectives

- To identify features influencing the pricing.
- To analyse trends in house prices over time (time series analysis) and predict future prices.
- To identify undervalued properties (outlier detection) and recommend better pricing strategies

# Correlation between price and Square footage of living space



There is a strong positive correlation between square footage of living space and price of the home.

## House prices according to months

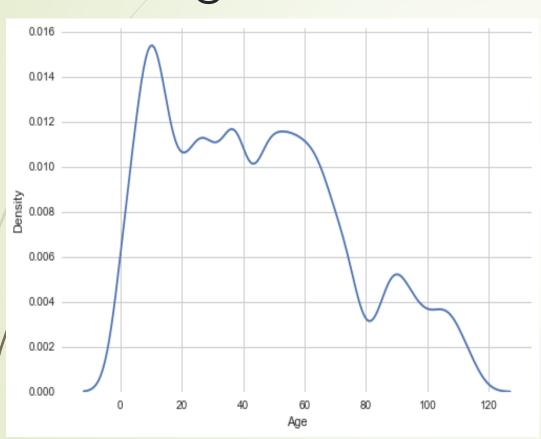


There appears to be a decrease in price especially in the month of March 2015, however, after that there seems to be a sharp increase.

### Relationship of price with property features

	Correlation Matrix of Numerical Features																		
id	1	-0.017	0.0057	0.0024	-0.011	-0.13	0.02	-0.00033	-0.028	0.01	-0.01	-0.0045	0.022	-0.01	-0.0097	-0.0067	0.022	-0.0029	-0.14
price	-0.017	1	0.31	0.52	0.7	0.087	0.26	0.26	0.0095		0.61	0.32	0.051	0.13	-0.052	0.31	0.022	0.59	0.083
bedrooms	0.0057	0.31	1	0.51	0.57	0.03	0.18	-0.0029	0.0024	0.36	0.47	0.3	0.15	0.019	-0.15	-0.0056	0.13	0.39	0.029
bathrooms	0.0024	0.52	0.51	1	0.75	0.086	0.5	0.063	-0.17	0.66	0.69	0.28	0.51	0.051		0.028	0.22	0.57	0.087
sqft_living	-0.011	0.7	0.57	0.75	1	0.17	0.36	0.11	-0.093	0.75	0.88	0.43	0.32	0.054		0.057	0.24	0.76	0.18
sqft_lot	-0.13	0.087	0.03	0.086	0.17	1	-0.0091	0.024	0.0026	0.1	0.18	0.02	0.049	0.0036	-0.13	-0.084	0.23	0.14	0.72
floors	0.02	0.26	0.18	0.5	0.36	-0.0091	1	0.016	-0.3	0.47	0.53	-0.25	0.49	0.0027	-0.058	0.054	0.13	0.28	-0.013
waterfront	-0.00033	0.26	-0.0029	0.063	0.11	0.024	0.016	1	0.018	0.072	0.073	0.084	-0.022	0.078	0.027	-0.014	-0.038	0.087	0.028
condition	-0.028	0.0095	0.0024	-0.17	-0.093	0.0026	-0.3	0.018	1	-0.19	-0.19	0.16	-0.4	-0.069	-0.0026	-0.033	-0.11	-0.11	0.0053
grade	0.01	0.64	0.36	0.66	0.75	0.1	0.47	0.072	-0.19	1	0.74	0.16	0.46	0.017	-0.19	0.11	0.2	0.71	0.11
sqft_above	-0.01	0.61	0.47	0.69	0.88	0.18	0.53	0.073	-0.19	0.74	1	-0.053	0.42	0.02	-0.26	0.0049	0.34	0.73	0.19
sqft_basement	-0.0045	0.32	0.3	0.28	0.43	0.02	-0.25	0.084	0.16	0.16	-0.053	1	-0.13	0.074	0.072	0.11	-0.14	0.2	0.019
yr_built	0.022	0.051	0.15	0.51	0.32	0.049	0.49	-0.022	-0.4	0.46	0.42	-0.13	1	-0.22	-0.35	-0.15	0.41	0.32	0.071
yr_renovated	-0.01	0.13	0.019	0.051	0.054	0.0036	0.0027	0.078	-0.069	0.017	0.02	0.074	-0.22	1	0.069	0.034	-0.072	-5e-05	0.0024
zipcode	-0.0097	-0.052	-0.15			-0.13	-0.058	0.027	-0.0026	-0.19	-0.26	0.072	-0.35	0.069	1	0.26	-0.56	-0.28	-0.15
lat	-0.0067	0.31	-0.0056	0.028	0.057	-0.084	0.054	-0.014	-0.033	0.11	0.0049	0.11	-0.15	0.034	0.26	1	-0.13	0.049	-0.084
long	0.022	0.022	0.13	0.22	0.24	0.23	0.13	-0.038	-0.11	0.2	0.34	-0.14	0.41	-0.072	-0.56	-0.13	1	0.34	0.26
sqft_living15	-0.0029	0.59	0.39	0.57	0.76	0.14	0.28	0.087	-0.11	0.71	0.73	0.2	0.32	-5e-05	-0.28	0.049	0.34	1	0.18
sqft_lot15	-0.14	0.083	0.029	0.087	0.18	0.72	-0.013	0.028	0.0053	0.11	0.19	0.019	0.071	0.0024	-0.15	-0.084	0.26	0.18	1
	.0	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	condition	grade	sqft_above	sqft_basement	yr_built	y_renovated	zipcode	at	long	sqft_living15	sqft_lot15

# Density plot representing the house ages



Most houses are distributed within 10 to 50 years, while a small portion of the properties are above 90+ years.

### Conclusions

- 1. Some of the features that influence the pricing of houses include:
- Square footage of living space in the home: an additional square footage increases the price by \$199.09
- Waterfront: the presence of a waterfront has an associated increase in price of \$70,000
- Condition of the house: houses in good conditions have an associated increase in price of \$35,650 compared to houses with average condition.

2. For every additional year in the age of a house, there is an associated decrease in price of \$626.09

3.

- Some of the overvalued properties were found to be older than 50 years of age
- The square footage of interior housing living space for the nearest 15 neighbors influences the pricing of houses, in that, an additional square footage leads to an increase in price by \$48.35

### Recommendations

We recommend that:

1. There is a need to do further exploration into other **features** in order to better understand the determinants of house prices.

2. The agency should consider **re-purposing** the old houses and targeting business owners rather than homeowners, this may also be achieved by market research.

3. The agency should consider investing in properties that can increase their profitability, such as properties that have a waterfront.

### Next Steps

- 1. New features can be generated from the existing data to provide **more insights** into the housing market. For example: In-depth information about regions
- 2. Visualise the properties on a map. This would enable our stakeholders to see the affordability of properties per region. Additionally, this would help to also determine the best regions to invest in.
- 3. Revise the models so that it reflects the current market trends so that this will allow for better accuracy in predicting the prices.

GitHub Group Repo

Thank You