KING COUNTY HOUSING SALES ANALYSIS



MARTIN NGUGI

INTRODUCTION

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.

Business Understanding

The real estate agency helps homeowners buy or sell homes. One of the key services they provide is advice to homeowners about how home renovations can increase the estimated value of their homes. The agency is interested in developing a model that can predict the estimated value of a home after renovations, based on the type and cost of the renovations.

Business Problem

The real estate agency needs to provide accurate advice to homeowners about how home renovations can increase the estimated value of their homes, and by what amount. However, the

agency currently lacks a reliable method for predicting the impact of specific home renovations on

home value. As a result, the agency is unable to provide accurate advice to homeowners about the

potential return on investment for different renovation projects.

The project objectives we aim to solve include:

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

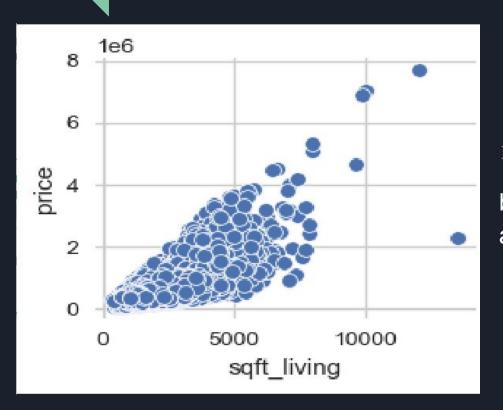
PROBLEM STATEMENT

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 In Accordance To Months

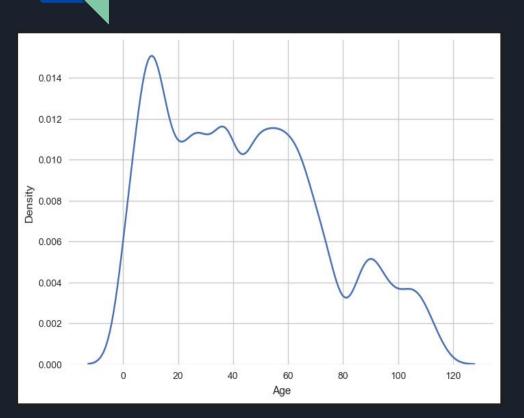


There appears to be a decrease in price especially in the month of March 2015, however, after that there seems to be a sharp increase.

Relationship of Price With Property Features

	Correlation Matrix of Numerical Features																
id	1	-0.015	0.0078	0.0056	-0.0083	-0.14	0.021	-0.0092	-0.00031	0.024	-0.01	-0.0074	-0.005	0.02	0.0009	-0.14	-0.0039
price	-0.015		0.31	0.53		0.085	0.26	0.61	0.32	0.049	0.12	-0.05	0.31	0.021	0.58	0.079	0.89
bedrooms	0.0078	0.31	4.	0.51	0.57	0.025	0.18	0.47	0.3	0.15	0.017	-0.15	-0.0059	0.13	0.39	0.025	0.34
bathrooms	0.0056	0.53	0.51	1	0.75	0.08	0.51		0.28	0.5	0.047		0.03	0.22	0.57	0.082	0.55
sqft_living	-0.0083		0.57	0.75	- 1	0.17	0.36	0.88	0.43	0.31	0.05		0.058	0.24	0.76	0.18	
sqft_lot	-0.14	0.085	0.025	0.08	0.17	1	-0.0099	0.17	0.017	0.052	0.0021	-0.13	-0.084	0.23	0.15		0.096
floors	0.021	0.26	0.18	0.51	0.36	-0.0099	1	0.53		0.49	-0.00072	-0.058	0.058	0.13	0.28	-0.014	0.31
sqft_above	-0.0092	0.61	0.47		0.88	0.17	0.53	- 1	-0.052	0.42	0.019		0.009	0.34	0.73	0.19	0.61
sqft_basement	-0.00031	0.32	0.3	0.28	0.43	0.017		-0.052	1	-0.13	0.068	0.074	0.1	-0.14	0.2	0.019	0.31
yr_built	0.024	0.049	0.15	0.5	0.31	0.052	0.49	0.42	-0.13	1	-0.22	-0.34	-0.14	0.41	0.32	0.07	0.078
yr_renovated	-0.01	0.12	0.017	0.047	0.05	0.0021	-0.00072	0.019	0.068	-0.22	11	0.068	0.032	-0.071	-0.0057	0.0025	0.11
zipcode	-0.0074	-0.05	-0.15			-0.13	-0.058		0.074		0.068	1	0.27	-0.56	-0.28	-0.15	-0.035
lat	-0.005	0.31	-0.0059	0.03	0.058	-0.084	0.058	0.009	0.1	-0.14	0.032	0.27	1	-0.13	0.047	-0.083	0.45
long	0.02	0.021	0.13	0.22	0.24	0.23	0.13	0.34	-0.14	0.41	-0.071	-0.56	-0.13	1	0.34	0.25	0.053
sqft_living15	0.0009	0.58	0.39	0.57	0.76	0.15	0.28	0.73	0.2	0.32	-0.0057	-0.28	0.047	0.34	1	0.18	0.62
sqft_lot15	-0.14	0.079	0.025	0.082	0.18	0.72	-0.014	0.19	0.019	0.07	0.0025	-0.15	-0.083	0.25	0.18	1	0.09
log(price)	-0.0039	0.89	0.34	0.55	0.7	0.096	0.31	0.61	0.31	0.078	0.11	-0.035	0.45	0.053	0.62	0.09	1
	Đ	price	bedrooms	bathrooms	sqf_living	sqft_lot	floors	sqft_above	sqft_basement	yr_built	yr_renovated	zipcode	at	long	sqft_living15	sqft_lot15	log(price)

Density Plot Representing Age of Houses



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

CONCLUSIONS

- a. Square footage of living space in the home: an additional square footage increases the price by \$199.09
- b. Waterfront: the presence of a waterfront has an associated increase in price of \$70,000
- c. Condition of the house: houses in good conditions have an associated increase in price of \$35,650 compared to houses with average condition.
- d. For every additional year in the age of a house, there is an associated decrease in price of \$626.09
 - e. Some of the overvalued properties were found to be older than 50 years of age
- f. 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

- 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.

Thank You

MARTIN NGUGI

EMAIL: nwainaina76@gmail.com

GITHUB: https://github.com/martin-ngugi