**Business Understanding**

The selling price of a home is a critical factor in the real estate market, as it directly influences the financial outcome for both buyers and sellers. For homeowners, the selling price represents the return on their investment and can significantly impact their financial well-being. Potential buyers, on the other hand, rely on the selling price to make informed decisions about purchasing a property within their budget and assessing its value relative to similar properties in the market.

House price is influenced by a multitude of factors, which can be broadly categorized into three main categories: property-specific factors, market factors, and external factors. Property-specific factors encompass attributes such as location, size, condition, amenities, architectural style, and age of the property. Market factors include supply and demand dynamics, interest rates, mortgage availability, and prevailing economic conditions. External factors can range from neighborhood characteristics, such as school quality and crime rates, to broader influences like government policies, infrastructure development, and demographic trends.

Understanding the factors that influence the selling price of residential properties is of paramount importance to various stakeholders involved in the real estate industry. Real estate agents need this knowledge to provide accurate pricing recommendations and effective marketing strategies for their clients. Homeowners can benefit from understanding these factors to make informed decisions when pricing their properties. Investors and developers can leverage this knowledge to identify promising investment opportunities and maximize their returns.

**Problem Statement**

Real estate market is highly volatile, influenced by economic conditions, housing demand, and external factors. As such setting inappropriate prices and making uninformed decisions of when to sell a house can be counterproductive. Research is essential to understand market trends and identify the best time to sell a home for maximizing its selling price. Analyzing property characteristics such as location, size, amenities, condition, and recent market trends through research aids in setting an appropriate selling price. By understanding how different home characteristics impact selling prices, the agency can help home owners mitigate the risk of setting inappropriate prices or making poor investment decisions.

**Objectives**

1. To understand factors that affect price of a house

2. To develop a model that can predict housing price based on various features.

3. To make recommendations on how home owners can optimize selling price of a house

**Modelling**

* Baseline model is with variable with the highest correlation with price which was sqft\_living

***Model evaluation approach***

* Adjusted R-squared to show variation explained
* Durbin-Watson test to detect autocorrelation in residuals, ensuring the validity of regression analysis.
* Durbin-Watson test value between 1.5 and 2.5 indicates no significant autocorrelation

**Model comparison approach**

* Adjusted R-squared to show variation explained. Higher Adjusted R-squared indicates better model
* AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) to compare goodness-of-fit and complexity of the models. Lower AIC and BIC values indicate a better model

**Baseline model interpretation**

* With an F-statistic p-value below 0.05, the overall model is statistically significant
* The results show that sqft\_living explains 49.9% variation in price of houses.
* The model coefficients (const and sqft\_living) are both statistically significant, with t-statistic p-values below 0.05
* For 0 sqft\_living, the estimated price is estimated to be -43990
* An increase of 1 square foot in the living area leads to a price increase of approximately $286.

***Evaluating Model Performance using Durbin-Watson*.**

* The model has a Durbin-Watson of 1.983 which shows that there is no significant autocorrelation.

***Model Comparison***

* Second model had lower value AIC (434,900) and BIC (435,100) compared to model with sqft\_living which has AIC (438,200) and BIC (438,200). As such second model is a better fit.
* Second model explains 59.5% of the total variation in price. This is an improvement from the baseline model which explained 49.9% of the variation.

**Second Model Interpretation**

* The model is statistically significant since the F-statistic p-value is less than 0.05
* Most of the explanatory variables are statistically significant apart from condition\_Average, condition\_Fair, condition\_Good, condition\_Very Good and sqft\_basement which did not have a statistically significant linear relationship with price
* When all independent values are zero the price of a house is $228,800
* An increase in one bedroom reduces price of house by $ 40,120
* An increase in one square footage of living space increases price of house by $266.
* An increase in one square footage of the lot leads to a reduction of price by $0.41
* An increase in one floor increases price by $35,700
* An increase of grade by one unit reduces price by $22,000
* An increase in one unit of the square footage of the basement increases price by $2.32,
* An increase in one year for years since renovated leads to an increase of $2,618 in price
* Having a house with waterfront increases price by $494,900 compared to one without waterfront
* A house with an average view compared to that with no view leads to an increase of an $123,700 in price.
* A house with an excellent view compared to that with no view leads to an increase of an $326,900 in price.
* A house with a fair view compared to that with no view leads to an increase of an $164,800 in price.
* A house with a good view compared to that with no view leads to an increase of an $176,000 in price.
* Selling a house in fall reduces price by $26,040 compared to selling in spring
* Selling a house in summer reduces price by $26,780 compared to selling in spring
* Selling a house in winter reduces price by $32,670 compared to selling in spring

**Conclusions**

* Being on waterfront has highest positive effect followed by having an excellent, good, fair and average view respectively.
* Number of floors, years since renovated and square footage of living space have moderate positive effect
* Bedrooms have the highest negative effect on price followed by selling in winter, summer, fall and grade of house

**Limitations**

* The dataset does not include other factors such as location-specific characteristics, neighborhood amenities which determine price of houses
* The data is collected over a period of time which encompasses different economic situations, changing market conditions, economic trends, and policy changes which were affect house prices but were not captured in the dataset.

**Recommendations**

* Put up houses for sale in spring as it fetches higher prices
* Renovate houses as it increases value of the property
* Increase the square footage of the living space as it increases value
* Do not increase grade of the house as it reduces value of property
* Do not increase number of bedrooms as it reduces price of the house