Group 4 phase 2 project

Calvin Omwega
Collins Biwott
Mercy Kiragu
Ingavi Kilavuka

Business overview

A real estate agency aims to provide detailed pricing models to their clients showing how different features impact home sale prices. By understanding these impacts, the agency can advise sellers on how to enhance their property to maximize sale price and assist buyers in evaluating potential homes

Challenges

Market fluctuations due to shaky micro and macroeconomic conditions

Increasing competition from other agencies requires forward-thinking innovation

Regulatory changes like zoning, taxes, and housing laws can affect the market, necessitating adaptation for compliance and risk management.

Solutions

Collaborate for better data, tools, and expertise in real estate valuation and predictive modeling.

Use advanced data analytics, like machine learning, to analyze past sales, spot trends, and predict property values more accurately.

Create personalized pricing models considering property features, neighborhood traits, market demand, and buyer preferences for accurate pricing recommendations.

Problem statement

Guide sellers to maximize sale prices by advising on feature enhancements and using pricing model insights to inform renovation, upgrade, and staging decisions.

Empower buyers with insights on how property features affect sale prices, helping them prioritize preferences and negotiate effectively based on perceived value.

Elevate agency competitiveness with advanced pricing models, providing detailed insights into home sale factors, attracting sellers aiming for higher returns and buyers seeking expert guidance.

Data understanding

The King County House sale dataset contains information regarding houses sold during the one-year period ranging from May 2014 to May 2015.

TARGET/DEPENDENT VARIABLE:

price – the price of each home sold

PREDICTORS/INDEPENDENT VARIABLES:

- Id
- date
- bedrooms
- bathrooms
- sqft_living
- sqft_lot,
- floors
- sqft_above
- Waterfront
- View
- condition
- sqft_basement
- yr_built
- yr_renovated
- zipcode
- lat
- long
- sqft_living15
- sqft_lot15

Data analysis

Data cleaning

After loading the data, we began cleaning. Checking for missing, fixing missing values checking for duplicated values

Visual explorations

This was done to understand the data, identify errors, and to guide the modeling process.

Modelling

Modeling included simple linear regression to predict prices against other IV'S followed by multiple linear regression.

Conslusions

01

Key Predictors of House Prices: Several key predictors: Condition of the property, View from the property, Latitude, and longitude coordinates, Square footage of living space, Number of bedrooms, Number of bathrooms, Number of floors

02

Impact of Predictors: Analysis indicates significant impacts on house prices from factors including property condition, views, size, bedrooms, bathrooms, location, and number of floors

03

Model Performance: Our predictive models have limitations. Variables needing log transformation for regression assumptions require similar preprocessing for new data. The model's applicability to other counties may be limited due to regional price variations, and outlier removal may affect extreme value predictions.

Recommendations

Feature Enhancement: Encourage sellers to improve critical property features that significantly influence house prices, such as condition, views, and layout.

Property Marketing**: Market properties by highlighting critical features like condition, views, and square footage to attract buyers and support pricing.

Buyer Guidance: Help buyers

Price Guidance: Provide sellers with data-driven price guidance based on predictive models and specific property characteristics

Thank you