REAL ESTATE PRICE PREDICTION USING MULTIPLE LINEAR REGRESSION IN JMP

Project Overview:

This project was undertaken for a real estate firm aiming to enhance its property valuation process. The client provided a dataset comprising various features of residential properties, including size, location, age, and amenities. The objective was to develop a predictive model using **Multiple Linear Regression in JMP** to estimate property prices accurately and identify key factors influencing pricing.

Problem Statement:

The real estate firm faced challenges in accurately pricing properties due to the subjective nature of traditional appraisal methods. They required a data-driven approach to predict property prices, enabling more consistent and objective valuations. The goal was to build a statistical model that could quantify the impact of various property features on the selling price.

Data Preparation & Cleaning:

- Imported the dataset into JMP and examined the structure for compatibility.
- Handled missing values through appropriate imputation techniques.
- Converted categorical variables, such as property type and location, into dummy variables using JMP's Recode function.
- Detected and addressed outliers by analyzing boxplots and leverage plots.
- Normalized continuous variables to ensure uniform scaling.

Exploratory Data Analysis (EDA):

- Utilized JMP's Graph Builder to create distribution plots and identify patterns.
- Generated correlation matrices to assess relationships between variables.
- Identified multicollinearity issues among predictors, prompting further analysis.(Studocu)

Model Construction:

- Developed a Multiple Linear Regression model using JMP's Fit Model platform, setting the property price as the dependent variable.
- Included independent variables such as square footage, number of bedrooms and bathrooms, age of the property, and location indicators.
- Performed stepwise regression to retain statistically significant predictors.
- Assessed multicollinearity using Variance Inflation Factor (VIF) and addressed issues by removing or combining correlated variables.
- Validated model assumptions through residual plots and lack-of-fit tests.(JMP)

Diagnostics & Interpretation:

- The final model achieved an Adjusted R^2 of 0.78, indicating strong explanatory power.
- Key predictors included square footage, number of bathrooms, and proximity to amenities.
- Interpreted coefficients to understand the impact of each predictor on property price.
- Evaluated confidence intervals and standard errors to ensure statistical validity.

Visualizations Created:

- Developed interactive regression profilers to visualize how changes in predictors affect property price.
- Created scatterplots and bubble charts to illustrate relationships between variables.
- Generated residual histograms and Q-Q plots to assess model diagnostics visually.

Final Report & Deliverables:

- Compiled a comprehensive report using JMP's journal feature, including:
    - EDA summaries
    - Model diagnostics and validation steps
    - Regression tables with coefficient values and p-values
    - Interpretation of results in business terms

- Recommendations for property valuation strategies(Exode HR CA, GitHub, Academia)

Impact:

- The model provided the firm with a reliable tool for property valuation, reducing reliance on subjective assessments.
- Enabled the identification of key factors influencing property prices, informing strategic decisions.
- Improved pricing accuracy led to increased client trust and satisfaction.(GitHub)