IMPACT OF AMENITIES AND DEMOGRAPHICS ON APARTMENT RENTAL PRICES: A MULTIPLE LINEAR REGRESSION ANALYSIS IN JMP

Project Overview:

This project focused on understanding how different property features and neighborhood demographics affect rental pricing for apartments in an urban area. The client—an investment advisory firm—wanted a statistical model built in **JMP** that could guide their property acquisition and pricing strategies for rental properties across different city districts.

Problem Statement:

The advisory firm needed to make investment decisions about which types of apartments (based on features) yield higher rental returns. However, their current pricing model lacked analytical rigor. They wanted a **JMP-based regression model** to quantify the influence of various features—such as number of rooms, balcony, parking, and local income levels—on monthly rental prices.

Data Cleaning and Preparation:

- Imported structured Excel data into JMP (2500 rows, 15 variables).
- Detected and imputed missing values using median imputation for numerical fields and mode for categorical ones.
- Created dummy variables for categorical fields like "Balcony" (Yes/No), "Parking" (Yes/No), and "Building Type" (e.g., Studio, 1BHK, 2BHK).
- Scaled variables such as distance from nearest metro station, property age, and number of units in building.

Exploratory Data Analysis (EDA):

- Conducted univariate profiling to assess normality, skewness, and presence of outliers.
- Generated scatterplot matrices using JMP's Graph Builder to identify relationships.
- Analyzed rental trends across city zones by income bracket and proximity to commercial hubs.

- Identified potential multicollinearity between features like number of rooms and property size.

Model Development in JMP:

- Used the **Fit Model** platform to construct a multiple linear regression model with monthly rent as the dependent variable.
- Independent variables included: number of rooms, balcony availability, parking, distance to metro, local median income, and building age.
- Performed backward elimination based on p-values and checked for model improvements.
- Ensured multicollinearity control using VIF < 5 for all predictors.
- Validated residual assumptions (normality, constant variance) via plots and goodness-of-fit statistics.

Model Summary and Diagnostics:

- Final model yielded **Adjusted R^2 = 0.81**, showing strong predictive power.
- Statistically significant predictors: local income, metro distance, parking, and balcony presence.
- Presence of a **balcony** increased rent by an average of **₹3,200/month** ($p < 0.01$).
- Proximity to metro had a negative correlation: every 1 km farther reduced rent by approx. ₹1,500.
- No major violations found in residual plots; Cook's Distance identified a few outliers that were flagged but retained.

Visualization Outputs:

- Regression profiler to simulate rent values under different configurations.
- Interactive residual histograms and leverage plots to check model fitness.
- Summary tables with standard errors, p-values, and CI bounds.

Report Deliverables:

- Complete project documented in a JMP Journal file with embedded graphs and text.

- Supplemented with PDF export of findings structured in 5 parts:
    1. Data understanding
    2. Feature analysis
    3. Model construction
    4. Coefficient interpretation
    5. Investment recommendations based on regression outputs

Impact & Result:

- Firm revised its rental pricing matrix and adjusted investment forecasts for neighborhoods with poor metro access.
- Also identified property types (e.g., those with balcony and parking) that command a premium, optimizing asset selection.