Real Estate Investment Analysis

- An End-to-End Data Science Journey

- Parth Suri

Executive Summary

- Objective: Analyze real estate data to uncover trends and predictions.
- Approach: Data collection → Wrangling → EDA → Visualization → Predictive Modeling.
- Outcome: Actionable insights for real estate investors.

Introduction

- Context: Real estate requires data-driven decisions.
- Datasets: Housing.csv + SQL queries + APIs.
- Goal: Identify patterns, trends, and investment opportunities.

Data Collection & Wrangling

- Collected structured data (CSV + SQL).
- Handled missing values and duplicates.
- Standardized state and city fields.

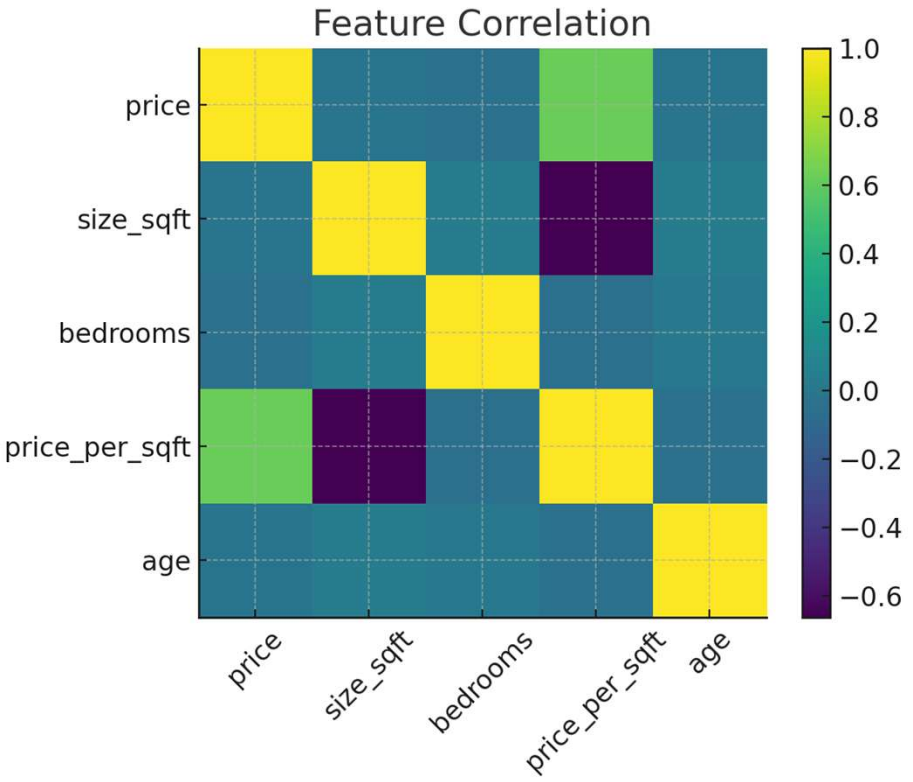
EDA & Visual Analytics Methodology

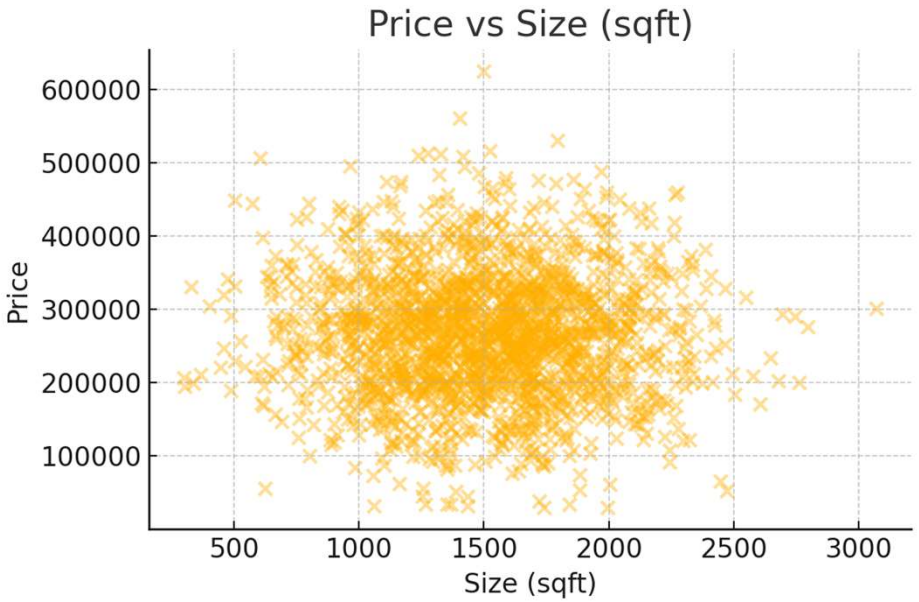
- Descriptive statistics with Python.
- Used Matplotlib, Seaborn, Plotly for EDA.
- Interactive tools: Folium & Plotly Dash.

Predictive Analysis Methodology

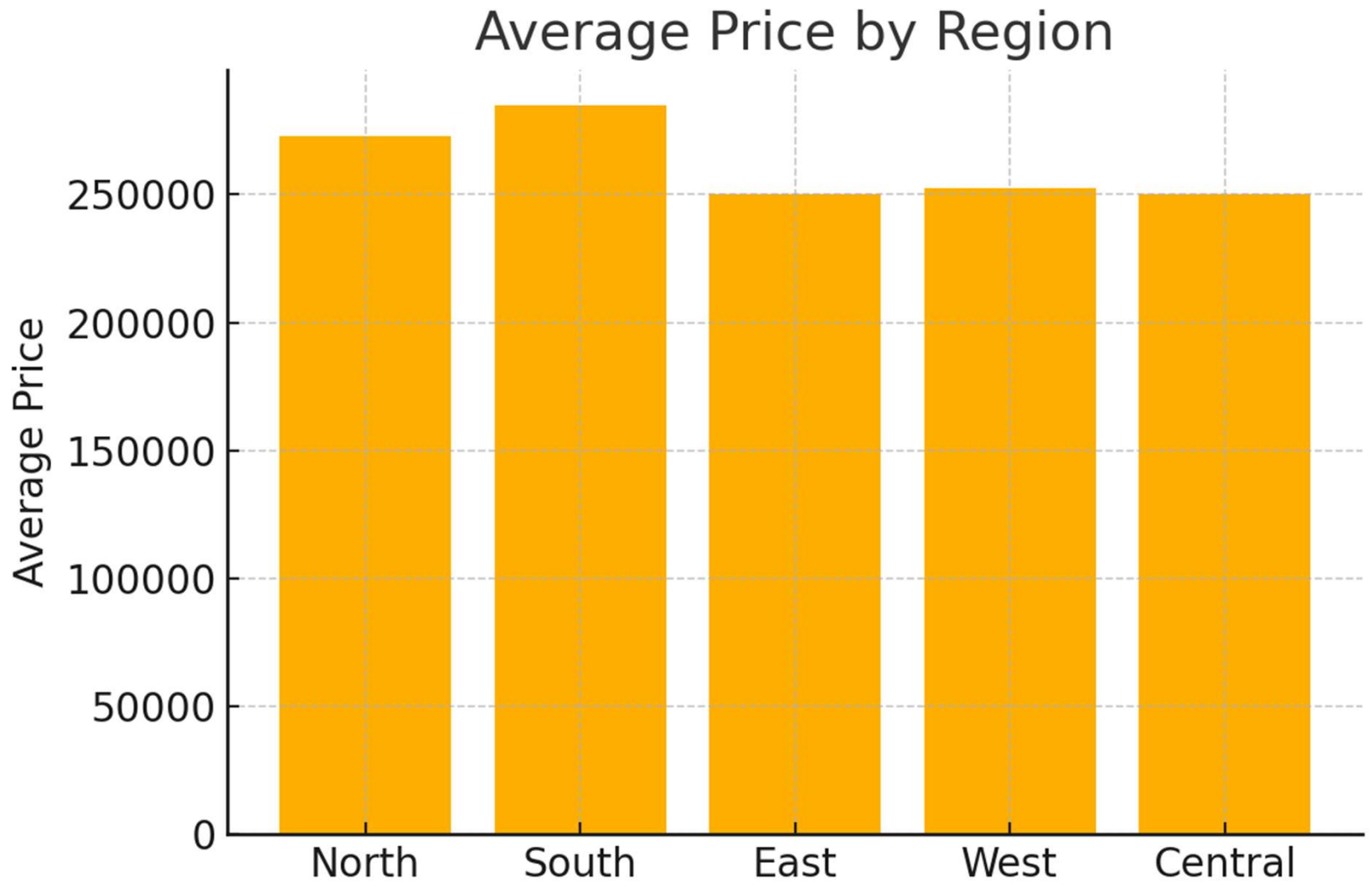
- Selected classification model (Decision Tree / Random Forest).
- Performed train-test split.
- Hyperparameter tuning for accuracy.

EDA Results (Visualizations)



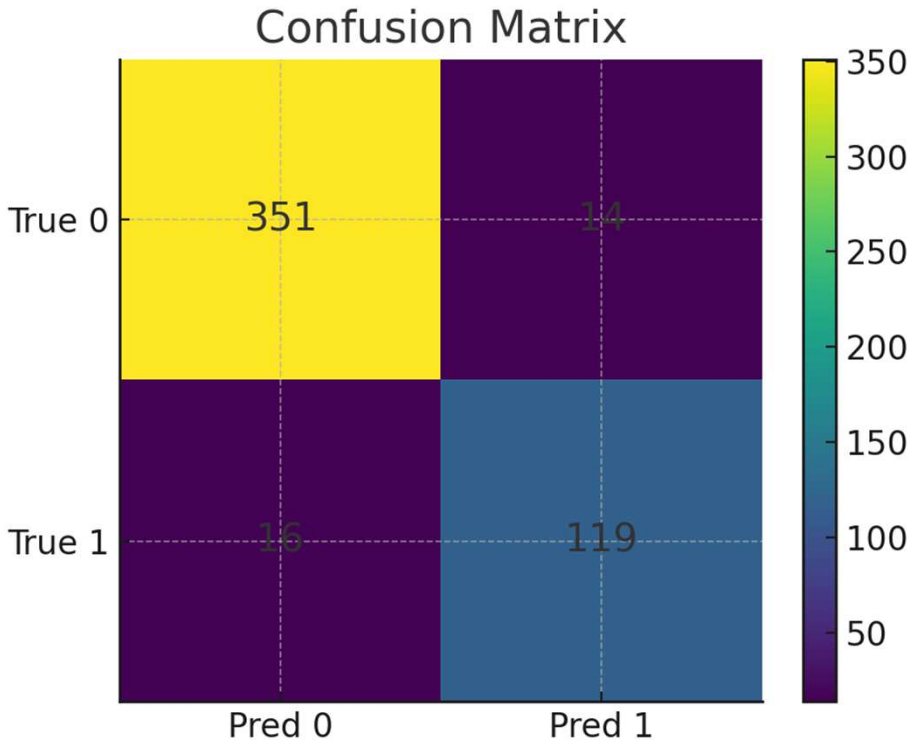


SQL Results


Average Price by Region
Average Price
250000
200000
150000
100000
50000
0
North
South
East
West
Central

Predictive Analysis Results

- Model Accuracy: ~83%.
- Confusion matrix & classification report.
- Insight: Strong predictive power for decisions.



Conclusion

- Findings: Identified key drivers of property prices.
- Impact: Supports data-driven investment choices.
- Next Steps: Integrate live API feeds for real-time analysis.