



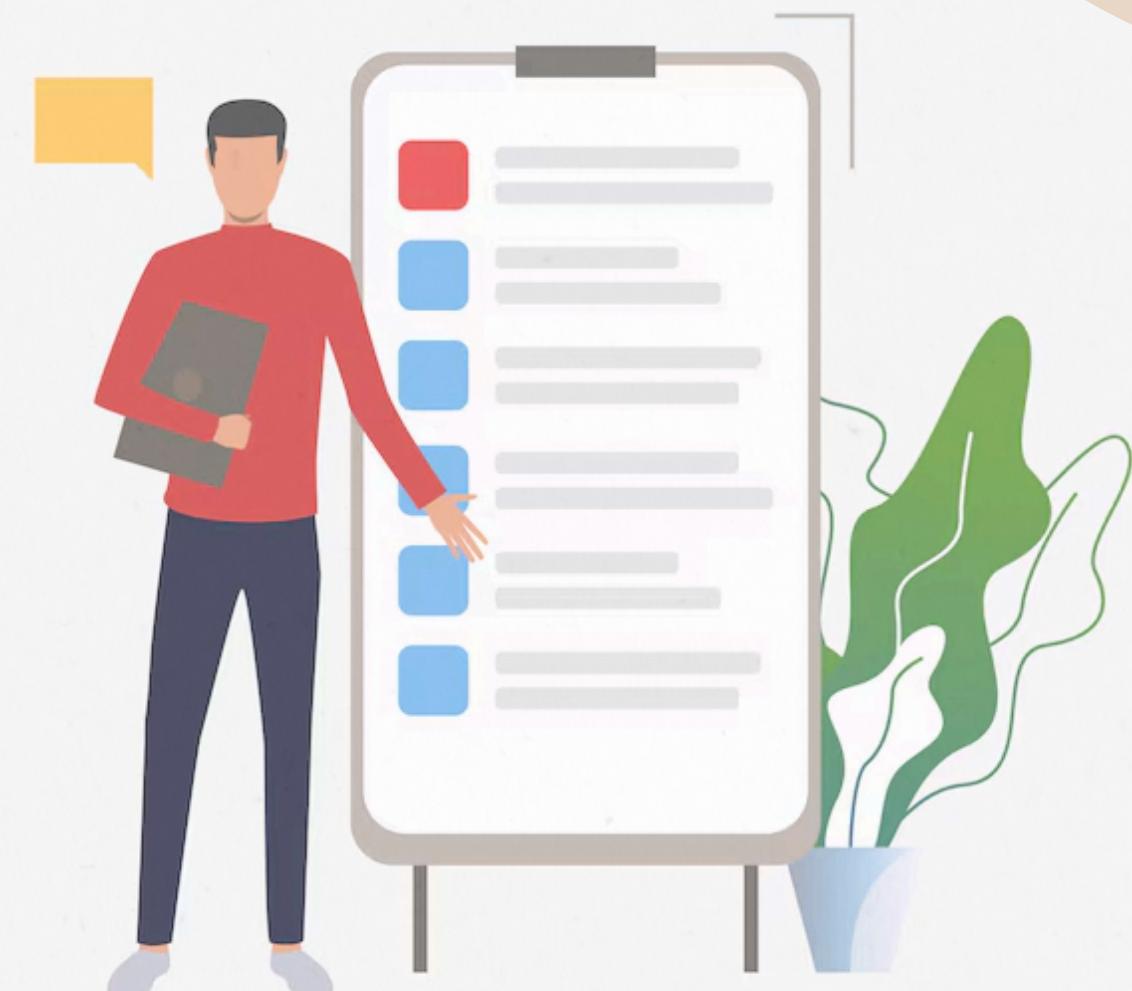
Exploratory Data Analysis (EDA) for Real Estate Pricing

**Unveiling the Dynamics of House Valuation
in a Dynamic Market**

- By Deewankar Sharma

Agenda

- Problem Statement
- Problem Overview
- Data Acquisition and Cleaning
- Univariate Analysis
- Multivariate Analysis
- Feature Engineering
- Size Impact Analysis
- Historical Pricing Trends
- Customer Preferences and Amenities
- Conclusion



Problem Statement



In the dynamic landscape of the residential real estate market, determining an optimal and competitive price for a house is a multifaceted challenge. As a key member of the analytics team in a leading real estate company, your task is to conduct a comprehensive analysis to identify and understand the myriad variables that significantly influence house prices. By leveraging advanced data analytics techniques and visualization tools, your goal is to uncover patterns, correlations, and trends within the dataset, enabling the company to make informed decisions and strategically position properties for better business opportunities.

Project Overview



The primary goal of this project is to analyze housing prices using a dataset containing various features related to real estate properties. By examining factors such as size, amenities, historical trends, and feature engineering, we aim to uncover valuable insights that can aid in understanding and predicting housing prices.

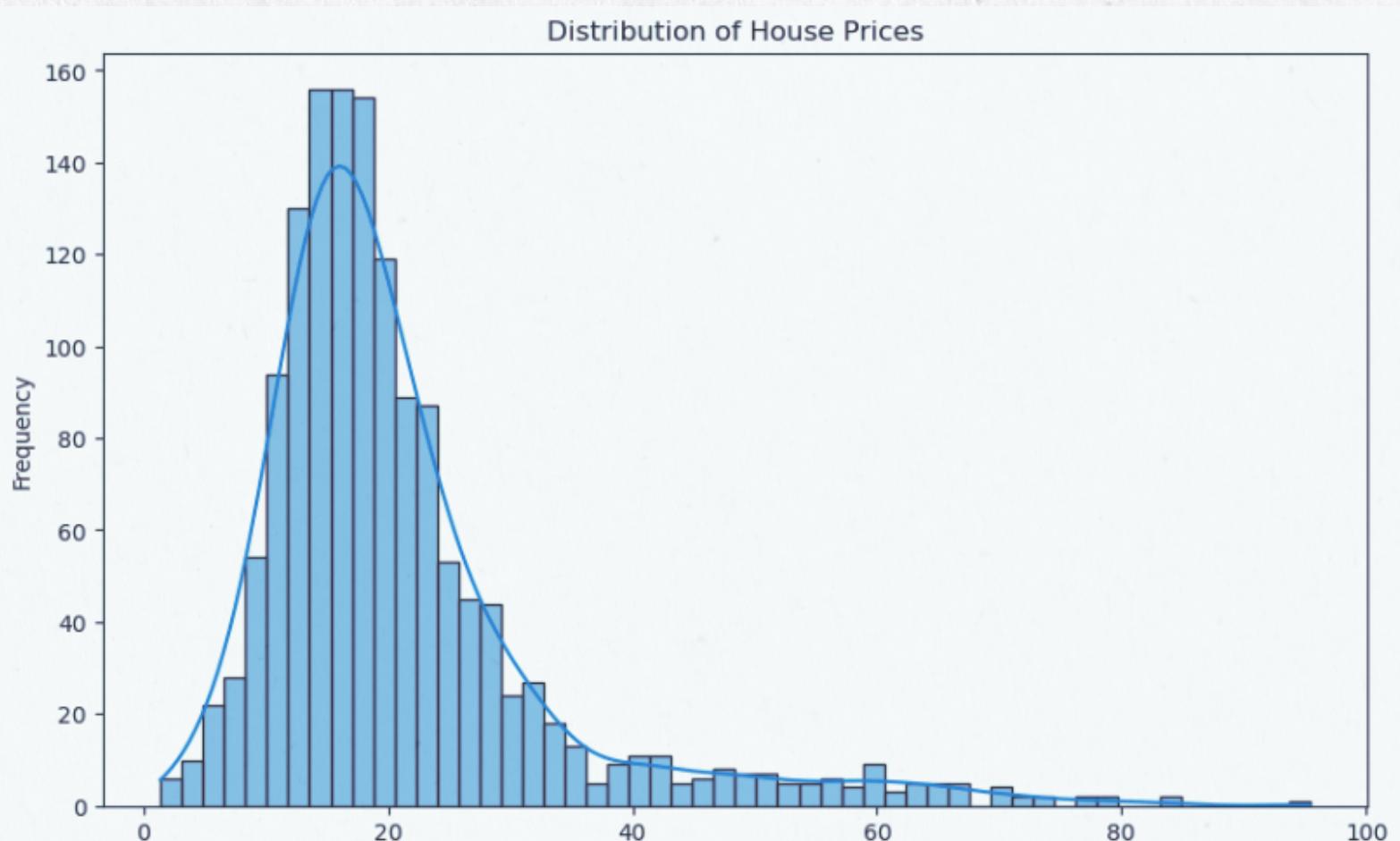
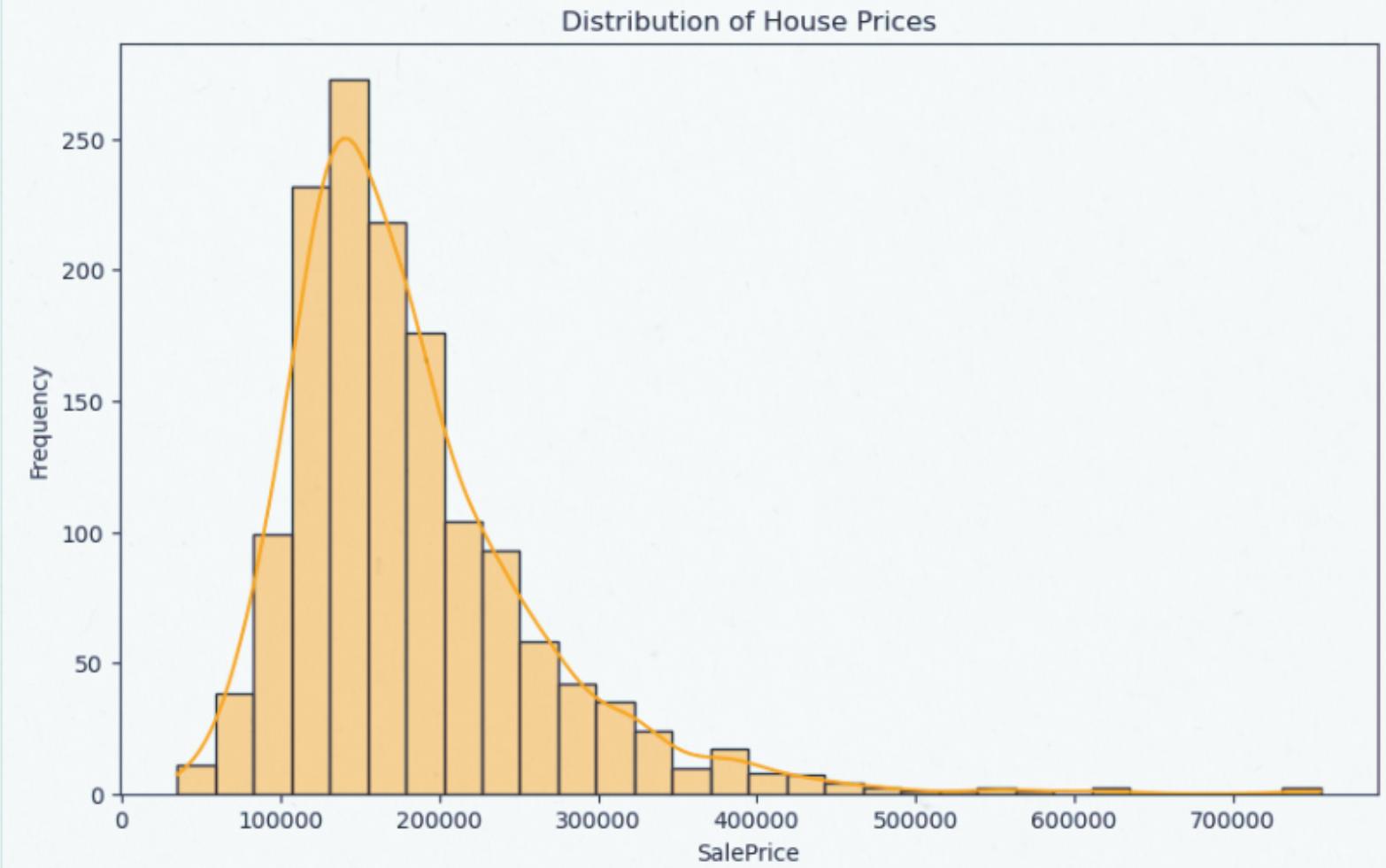
Data Acquisition and Cleaning

Importing necessary libraries, loading the dataset, and conducting initial exploratory checks. Additionally, data cleaning is crucial to ensure data quality. This process includes handling missing values, removing duplicates, and dropping unnecessary columns that do not contribute to the analysis. By ensuring the integrity of the data, we lay the foundation for meaningful insights to be derived from further analysis.



Univariate Analysis

It Focuses on exploring individual variables within the dataset to understand their distributions and characteristics. Through visualizations such as histograms and density plots, we gain insights into the distribution of key variables such as house prices and Price Per SqFt. Understanding the distribution of these variables provides essential context for subsequent analyses and helps identify potential outliers or patterns within the data.



Feature Engineering

It involves creating new features or modifying existing ones to enhance the model's predictive ability. In this analysis, we introduce new variables such as Price Per SqFt and Years Since Built. Price Per SqFt normalizes house prices relative to their lot area, providing a standardized metric for price comparisons. Years Since Built captures the age of the property at the time of sale, which can influence its valuation. By incorporating these features, we aim to improve the accuracy and interpretability of our analysis.

```
: # Create a new feature for price per square foot
Housing['Price Per SqFt'] = Housing['SalePrice'] / Housing['LotArea']

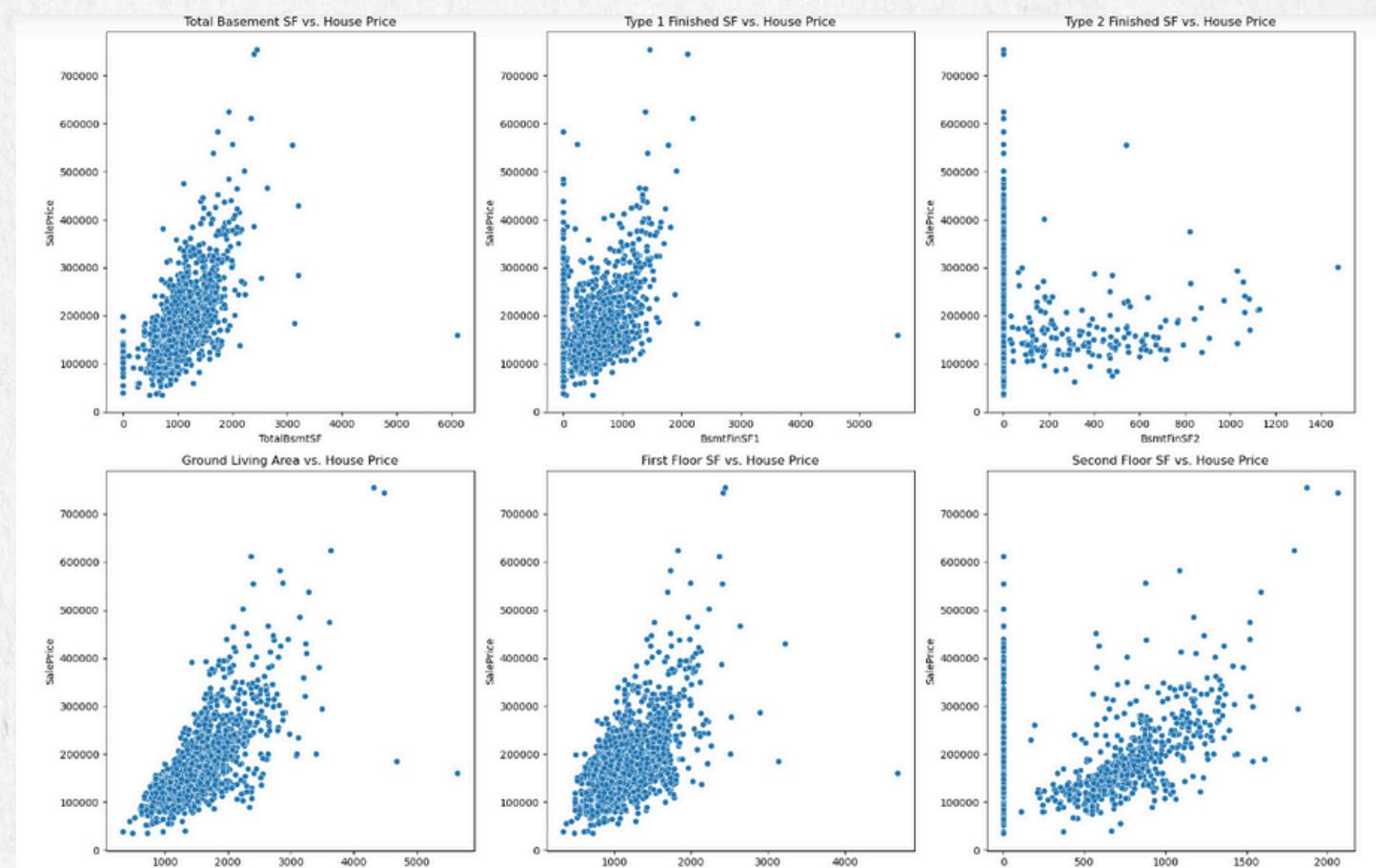
# Create a new feature representing the property's age
current_year = 2024                                # Assuming the current year is 2024
Housing['YrSold'] = current_year - Housing['YearBuilt']

print(Housing['YrSold'])
```

Size Impact Analysis

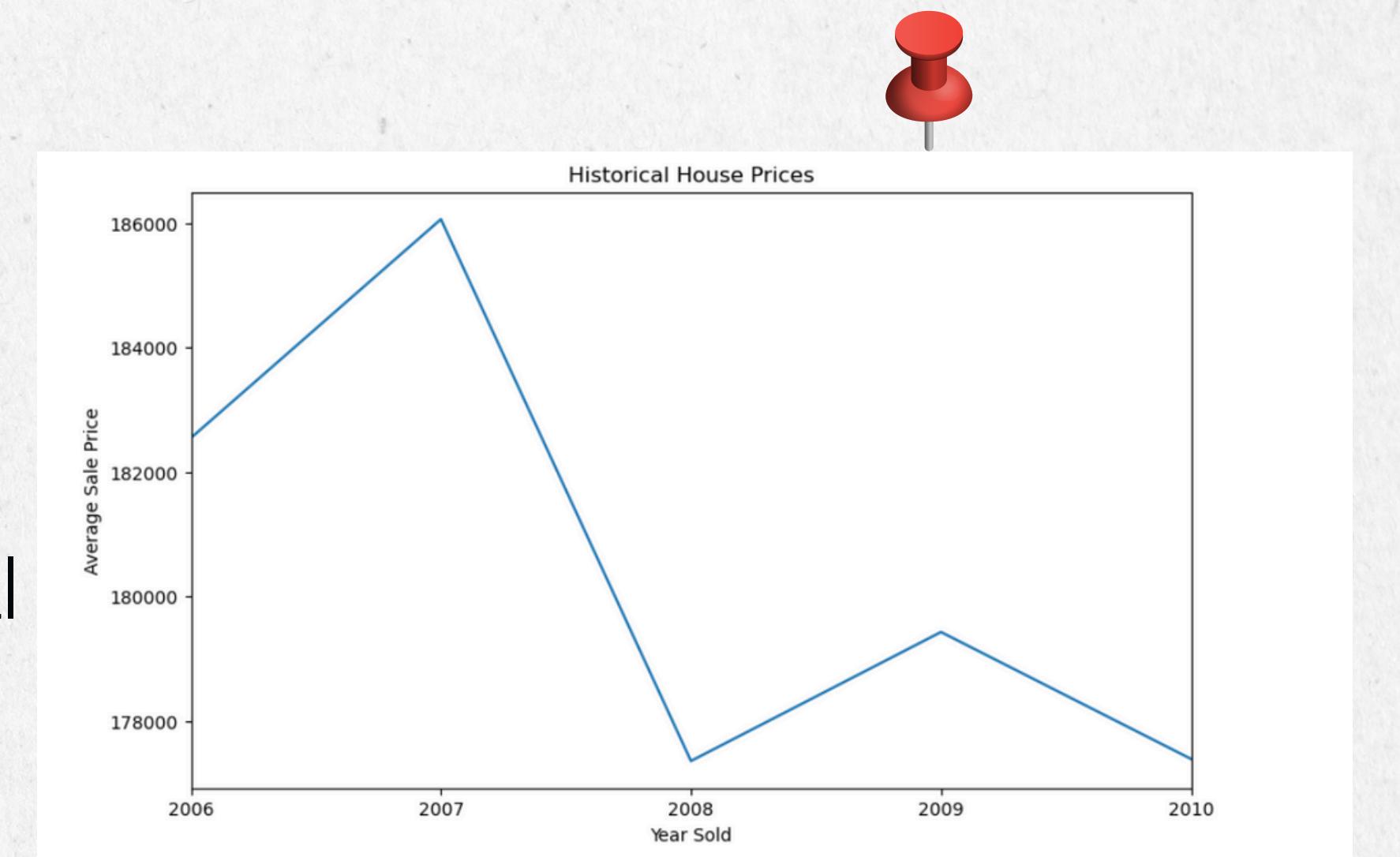
99

The size impact analysis focuses on understanding how size-related features, such as basement area and living area, affect house prices. By analyzing scatterplots, we aim to identify trends and patterns that illustrate the relationship between these features and house prices. This analysis provides insights into the value of additional space within a property and helps quantify its impact on valuation.



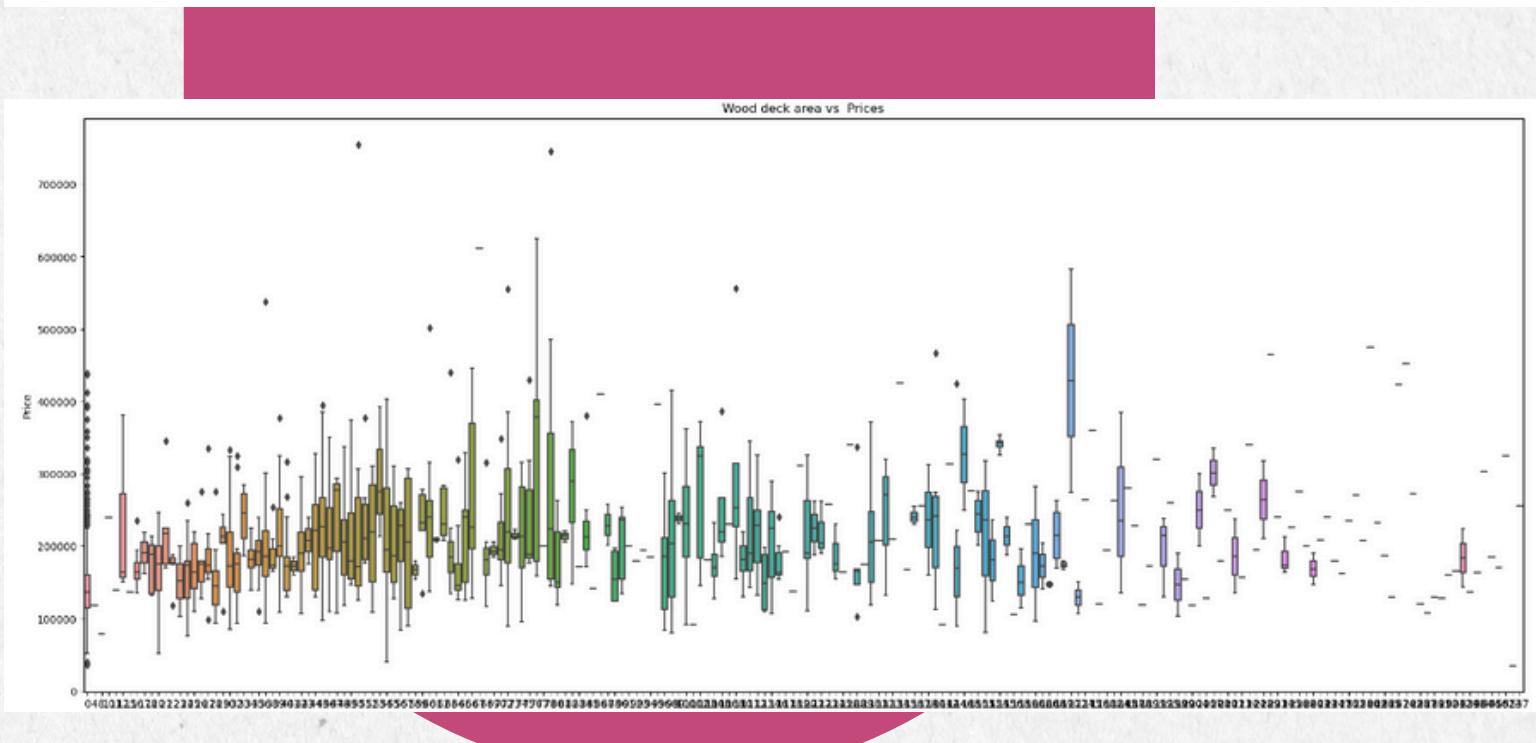
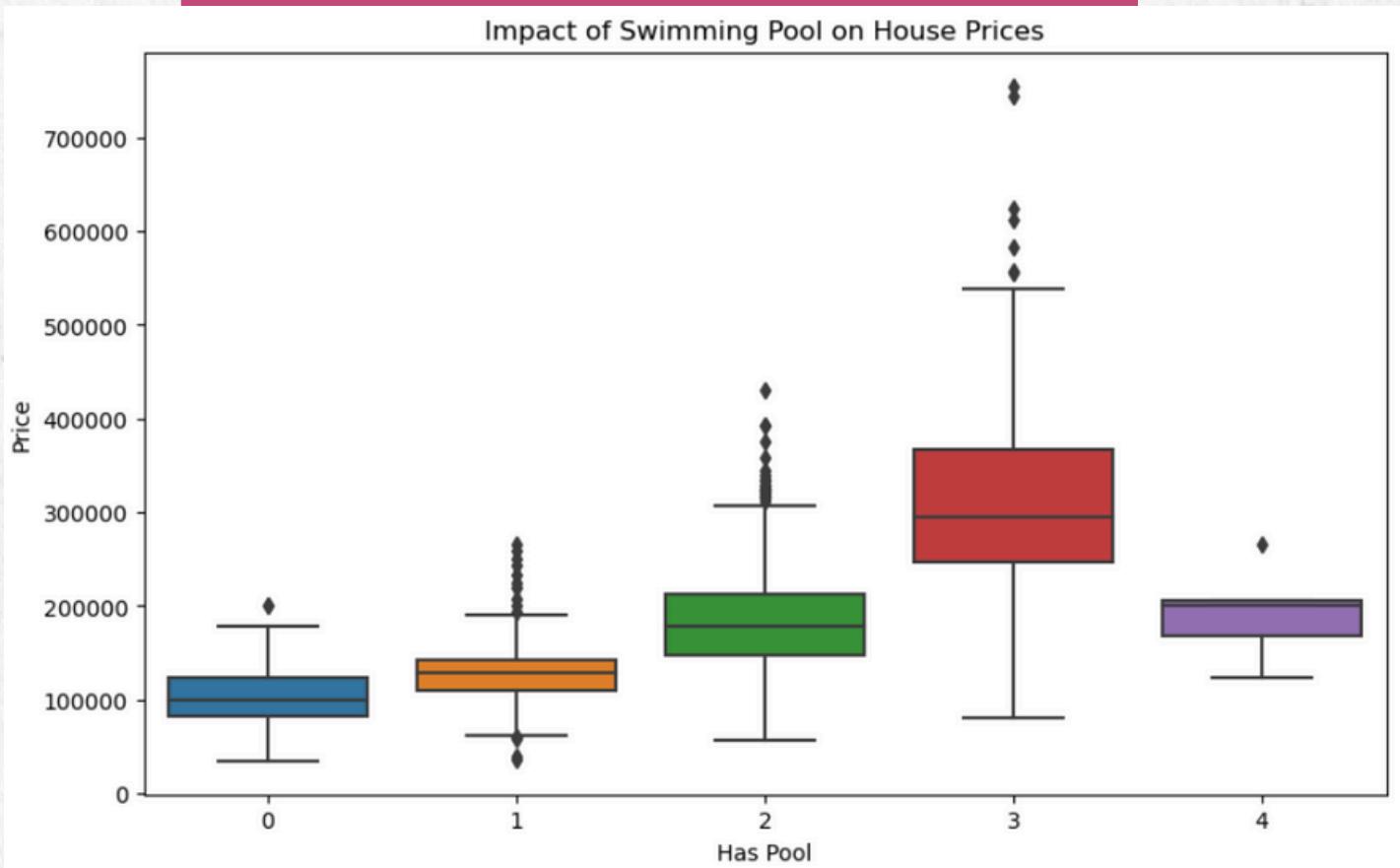
Historical Pricing Trends

Examining historical pricing trends over time provides valuable insights into market behavior and external influences. Through time-series analysis, we can identify patterns, trends, and cyclical fluctuations in house prices. Understanding these historical trends allows us to anticipate future market movements and adapt our pricing strategies accordingly.



Customer Preferences and Amenities

Customer preferences and amenities play a significant role in shaping house prices. By analyzing the impact of amenities such as pools, garages, and outdoor spaces on house prices, we gain insights into buyer preferences and market demand. Additionally, examining categorical features such as heating quality and kitchen condition provides further insights into customer preferences and their influence on valuation.



Conclusions

In conclusion, Exploratory Data Analysis (EDA) is a foundational step in understanding the dynamics of house valuation in a dynamic real estate market. We have uncovered valuable insights into the factors influencing house prices by conducting comprehensive analyses and leveraging advanced analytics techniques. These insights can inform strategic decision-making, optimize pricing strategies, and enhance the competitive positioning of properties within the market. Continued analysis and monitoring of market trends will be essential to adapt to changing market conditions and maintain a competitive edge.



