

Real Estate Success Blueprint:

Insights into
Seasonality of Sales
&
Property Features



Business Problem

- Seasonal Sales Analysis

Explore the impact of seasons on property sales within the real estate market and quantify this effects. Understanding seasonal variations is crucial in understanding how they affect the industry.

- Property Price Prediction

Evaluate how property characteristics have an affect on their prices. This can aid in accurately predicting property prices.



Data Understanding

This project utilizes House Sales dataset of King County, Washington.

Key columns and their descriptions:

- Price - Sale price (prediction target)
- Date - Date house was sold
- yr_renovated - Year when house was renovated
- bedrooms - Number of bedrooms
- bathrooms - Number of bathrooms
- sqft_living - Square footage of living space in the home
- sqft_lot - Square footage of the lot
- floors - Number of floors (levels) in house
- grade - Overall grade of the house. Related to the construction and design of the house
- condition - How good the overall condition of the house is. Related to maintenance of house.

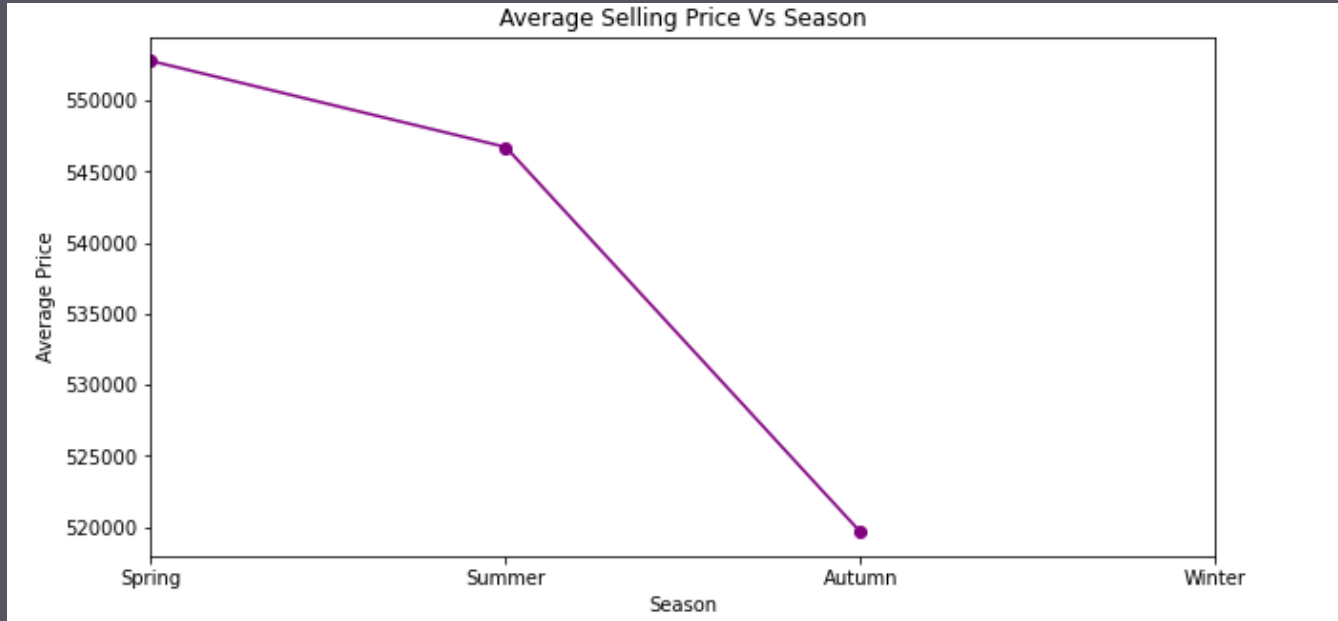


Data Cleaning

- Handled Missing Data.
- Dealing with Outliers
- Addressing Data Integrity.
- Encoding Categorical Variables.



Trend of Average Sales Vs Season



Spring has the highest average sales

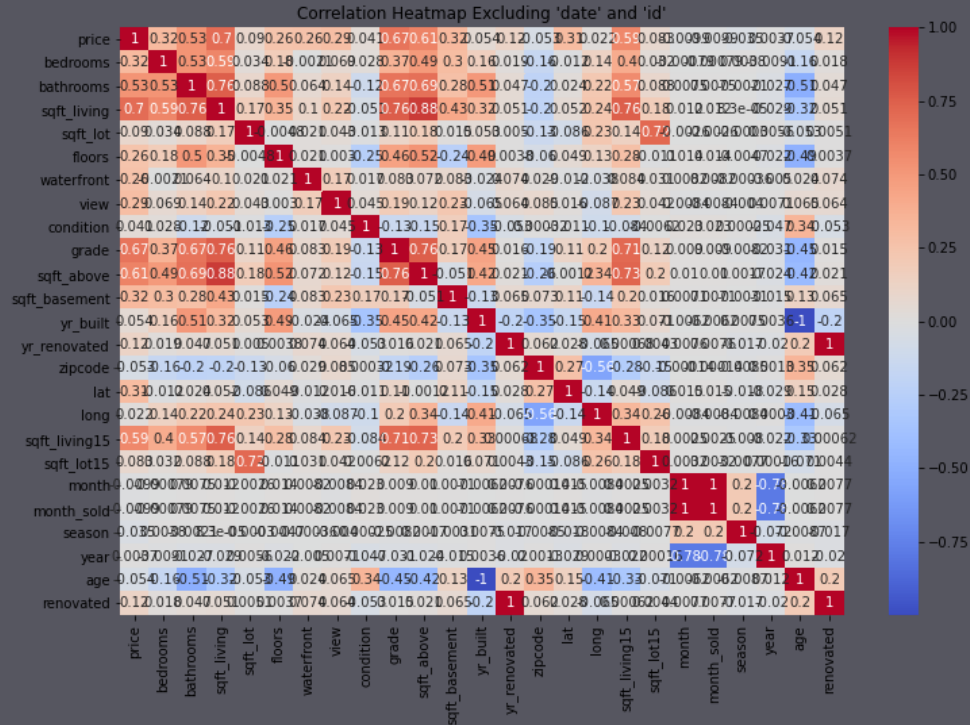


Sales Trend Per Season

season		price		
		sum	mean	median
1	Spring	3.603038e+09	552782.763271	465000.0
2	Summer	3.459641e+09	546719.464286	455000.0
0	Autumn	2.686134e+09	531276.474881	443725.0
3	Winter	1.919972e+09	519613.645467	430000.0



Correlation Heatmap



Sqft_living is the most correlated | Age least correlated

Modelling

Model 1

Features used: 'bedrooms', 'bathrooms', 'sqft_living', 'sqft_lot' and 'floors'

Model 2:

Features used: 'bedrooms', 'bathrooms', 'sqft_living', 'sqft_lot', 'floors', 'grade', 'condition'

Model 3:

'bedrooms', 'bathrooms', 'sqft_living', 'sqft_lot', 'floors', 'grade', 'condition', 'renovated'

Model 4:

'sqft_living', 'bedrooms', 'sqft_lot', 'grade', 'condition', 'renovated'



Interpretation of Models

- Model 3 performs consistently well with an RMSE of 241808 which is the lowest out of all the other models.
- This is the average prediction error of our model meaning that on average our predictions make an error of 241,808 dollars in terms of property prices.
- The model however has some room for improvement because an error of 241,807 is relatively high.
- The standard deviation, which represents the variability of scores across the data sets is 18716
- The 4th model did not improve even after using variables with price increase. However it was not so far off from the third model.



Regression Results

- The average prediction error on our test data of our model is 242,102.21 meaning that with every prediction there is an error of approximately 242,102.21 dollars in property prices. This can be interpreted as, on average our price predictions deviate from the actual values by 241,102.21
- The standard deviation, which represents the variability of scores across the data sets, is low at 11348.26. A lower std indicates that the model is performing consistently. Standard deviation shows how close predictions are to the actual values across different subsets of data.



Challenges

- Outlier Handling:
- Feature Selection:
- Model Evaluation:
- Feature Engineering:
- Interpretability:
- Seasonal Variations:
- Categorical Data Conversion:
- Identifying Impactful Features:
- Model Selection:



Business Recommendations

- Align property listings and promotions with peak sales seasons.
- Implement dynamic pricing strategies to adapt property prices in real time based on market conditions, seasons and demand changes.
- Stay updated and adaptable with market trends.
- Gathering feedback from customers to understand their preferences and use these feedback to refine strategies.
- Making use of insights from Real Estate Professionals and Market Experts
- Tailoring pricing recommendations to individual properties based on their unique features and market demand using data driven insights.



Future Steps

- Data Enrichment:
- Geographic Analysis:
- Real-time Data Updates:
- Market Segmentation:
- Additional Data Sources:
- Competitor Price Monitoring:
- Customer Feedback and Insights:



Thank You!

