

PHASE II PROJECT

PROBLEM STATEMENT

THROUGH THIS ANALYSIS, WE AIM TO UNCOVER MEANINGFUL INSIGHTS INTO REAL-ESTATE INVESTMENT, IDENTIFY PATTERNS, AND DERIVE ACTIONABLE RECOMMENDATIONS.



THE TEAM



Biwott Gideon



Silas Rugut



Sammy Macharia



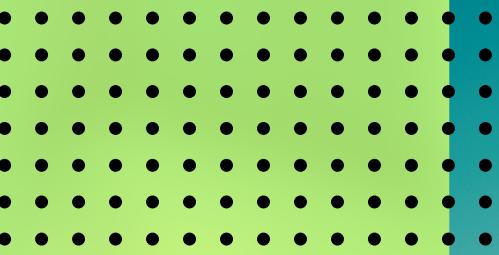
Baker Otieno



Mitch Anyul



Naeema Suleiman Sheikh



Key Business Questions

By the end of this project, we intended to answer these four main questions and more others;

1. Which feature has the strongest impact on sale price?
2. How much more valuable is a house for every additional square foot living space?
3. Can we develop a model to predict potential price growth for different renovations?
4. How much can the estimated sale price increase with individual renovations or combinations?



Business Understanding



This project aims to help real-estate agents achieve optimal Investment Decisions, Satisfaction, and to some extent good returns on investments (ROI). The Team will be working closely with the Real Estate Agency, to help provide added value to its clients (homeowners) by offering personalized advice on home renovations or in the Real Estate Investment. This strengthens the agency's relationship with clients and positions it as a trusted advisor in the competitive real estate market.



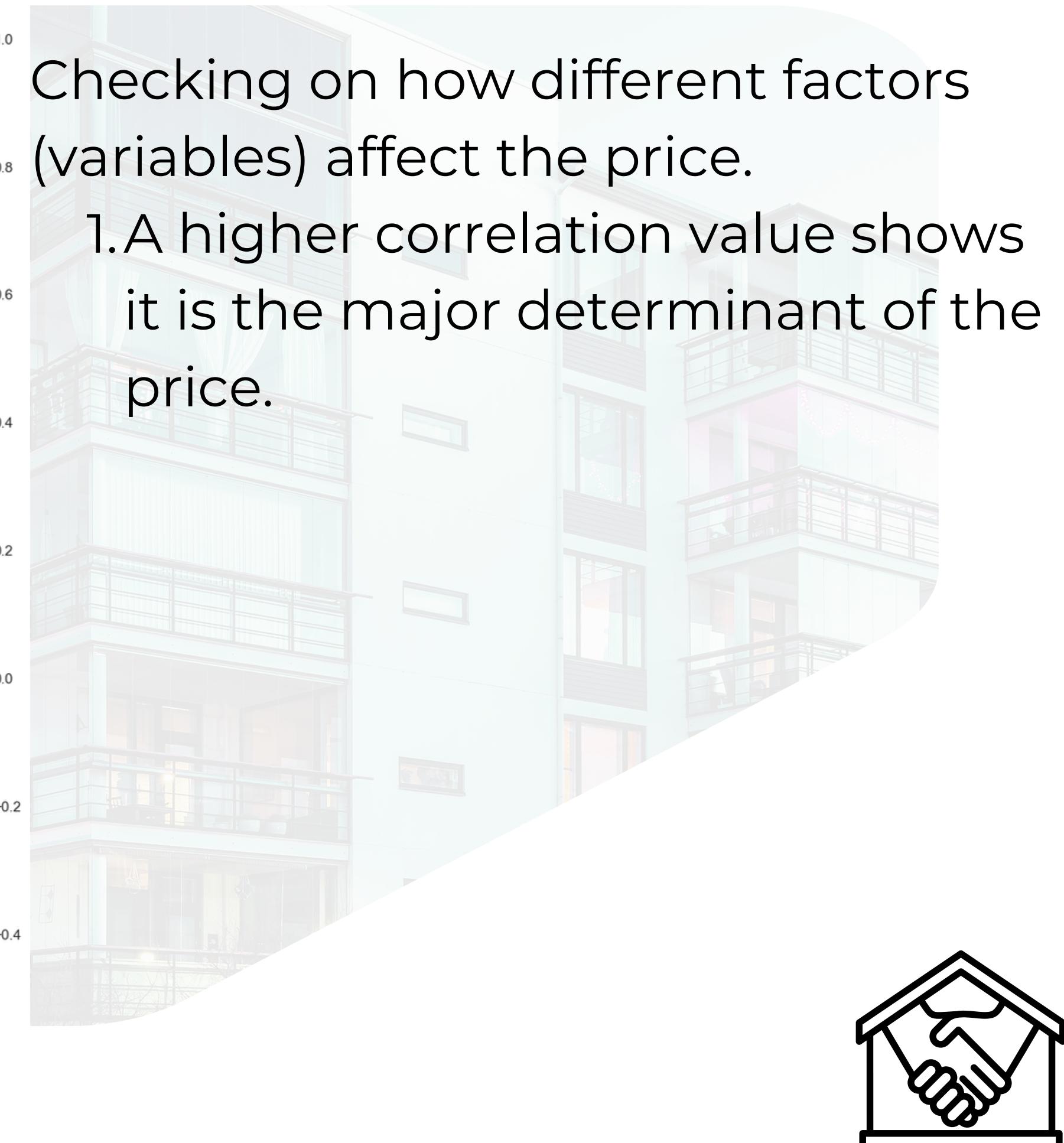
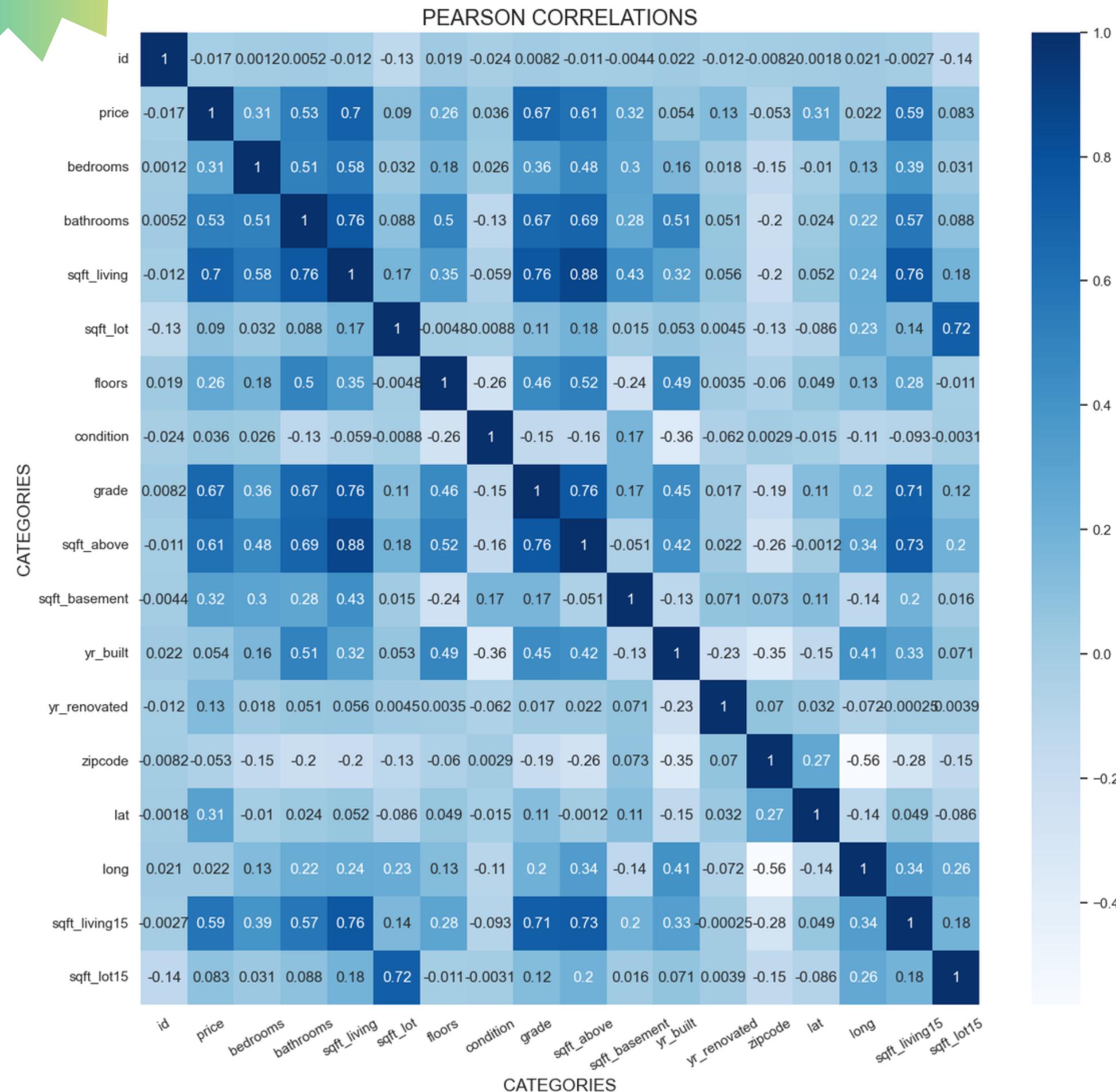
Data Understanding



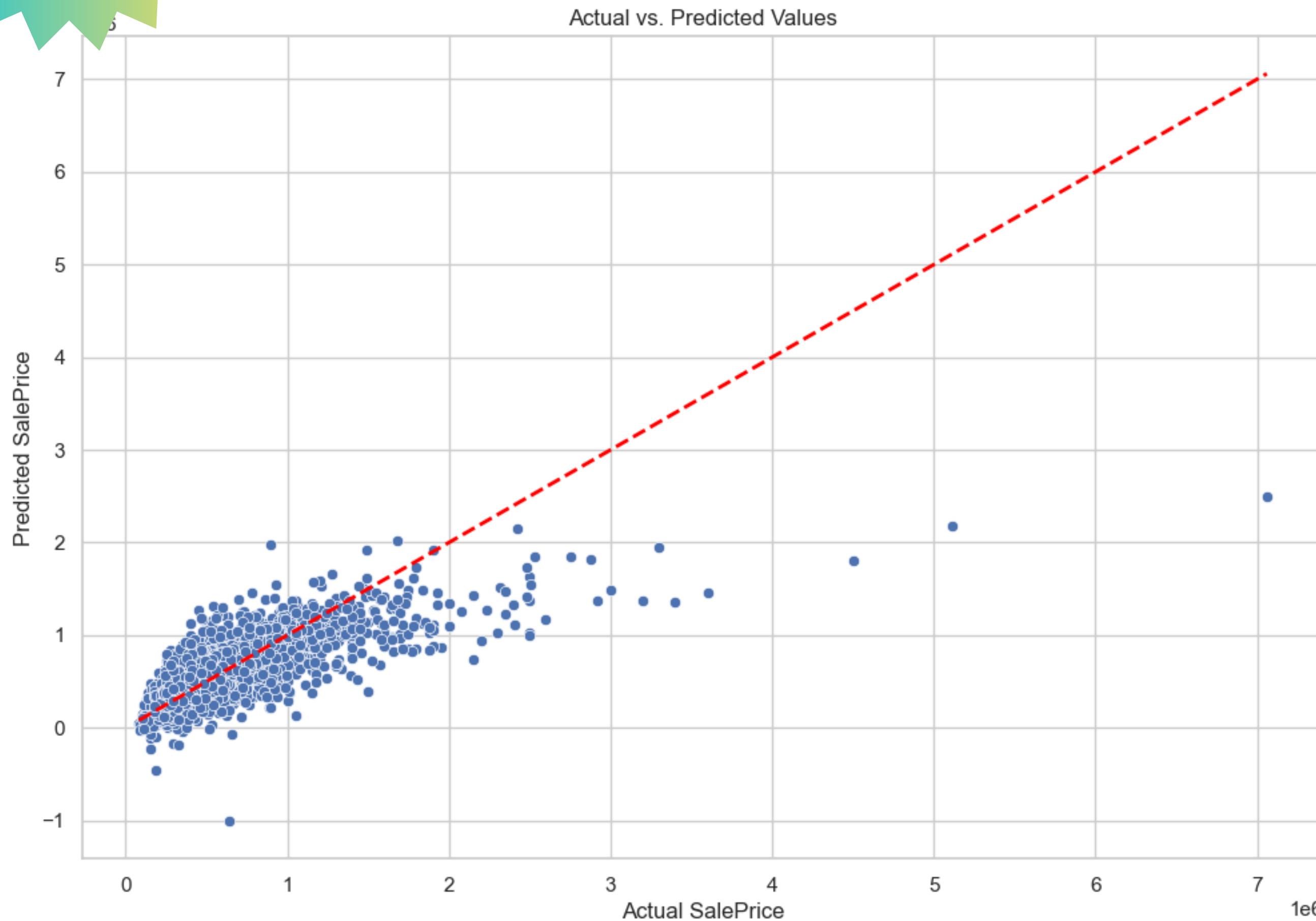
Online property companies offer valuations of houses using machine learning techniques. This report aims to predict house sales in King County, Washington State, USA using Multiple Linear Regression (MLR). The dataset consisted of historical data of houses sold between May 2014 to May 2015. The dataset consisted of 21 variables and 21613 observations.



Strength of an association between price and other variables



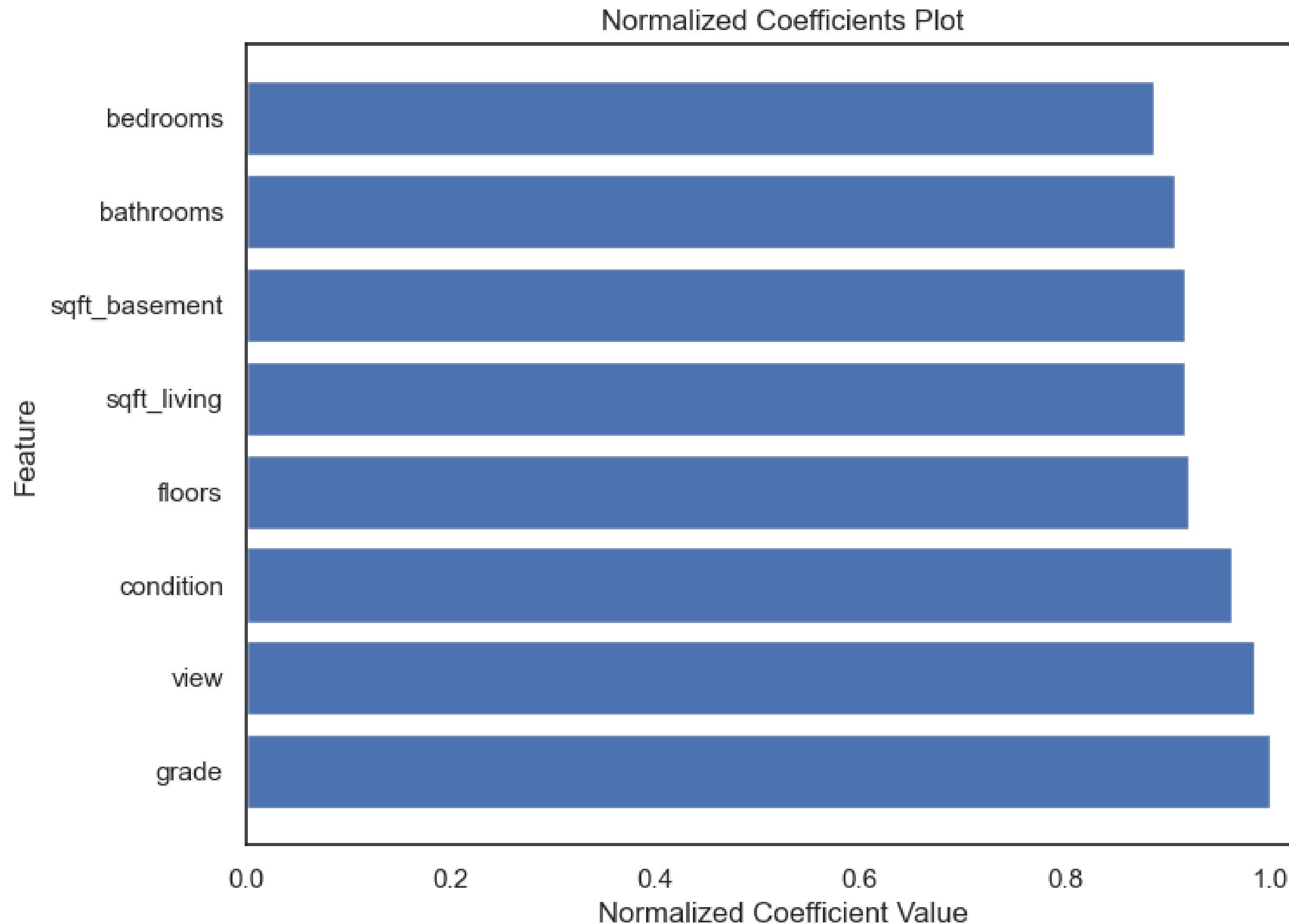
Data Understanding



The points cluster around the diagonal line indicating good predictive performance.



Price Against Other Independent Variables



Top Three Factors Affecting Price

1. Grade

2. View

3. Condition

Bottom three Factors Affecting Price.

1. Bedrooms

2. Bathrooms

3. Floors



Conclusion

1. The MLR result exhibited a statistical significance at a 95% confidence level ($p=0.00$) to the effect of; sqft_living, bedrooms, floors, sqft_basement, condition, grade, view, and bathrooms on price. The independent variables explain 59.3% of the price of the houses.
2. The factors ranked from the one with the highest impact on prices to the least are; House grade, View, Condition, Floors, sqft_living, sqft_basement, and bathrooms. The least impactful variable on the price is bedrooms.



Recommendations

1. Focus on Key Property Features: Our stakeholders should focus on key property features such as grade, view, and condition that significantly influence housing prices.
2. Upgrade Property Grades: Real estate developers and agencies should focus on improving the overall grade of properties, as it has a strong positive correlation with housing prices.
3. Enhance Views: Properties with better views tend to have higher prices. Real estate developers and agencies should consider and enhance the views of their properties.



Next Steps

1. Improve on the model by researching/adding other variables which would improve the model to explain the remaining 40.7%.
2. Explore additional features or transformations of existing features that could enhance the model's predictive power.
3. Revisit the data to identify and address outliers or influential points that may be impacting the model's performance.
4. Validate the model on new data to assess how well the model performs on unseen data



GROUP 1

