Phase 2 Project Presentation

Linear Regression Modeling *of Real Estate Sale Prices*

// FLATIRON SCHOOL

Title Slide

Name of the Project

Phase 2: Machine Learning Project

Outline

- **Business Problem**
- **Key Takeaways**
  - Key 1 (visualization)
  - Key 2 (visualization)
  - Key 3 (visualization)
- **Model Performance**
  - RMSE, R^2, MAE
- **Summary**
  - Recommendation, Next Steps, Future Analyses
- **Appendix**

Overview

Background / Business Problem

Audience

Brief description of the business problem, the stakeholder, the environment, etc.

Business Question

Describe the problem being solved and the analytical approach used to find a solution

Executive Summary / Key Takeaways

Approach & Solution

Summarize the modeling approach and key findings (*The key points your audience needs to walk away*)

- First Key Finding (predictor #1)
- Second Key Finding (predictor #2)
- Third Key Finding (predictor #3)

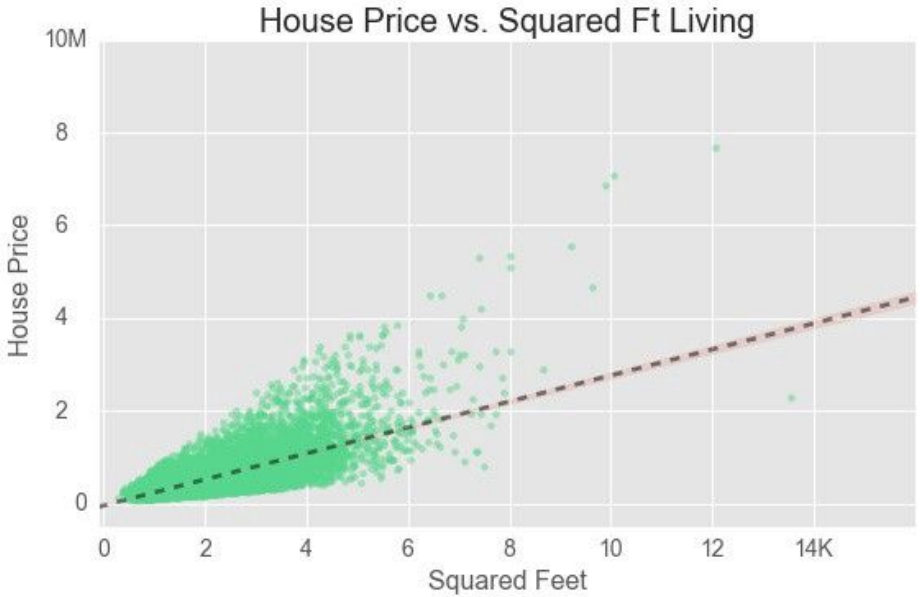
Repeat main takeaway / recommendation in one sentence.

Key Results

Finding 1

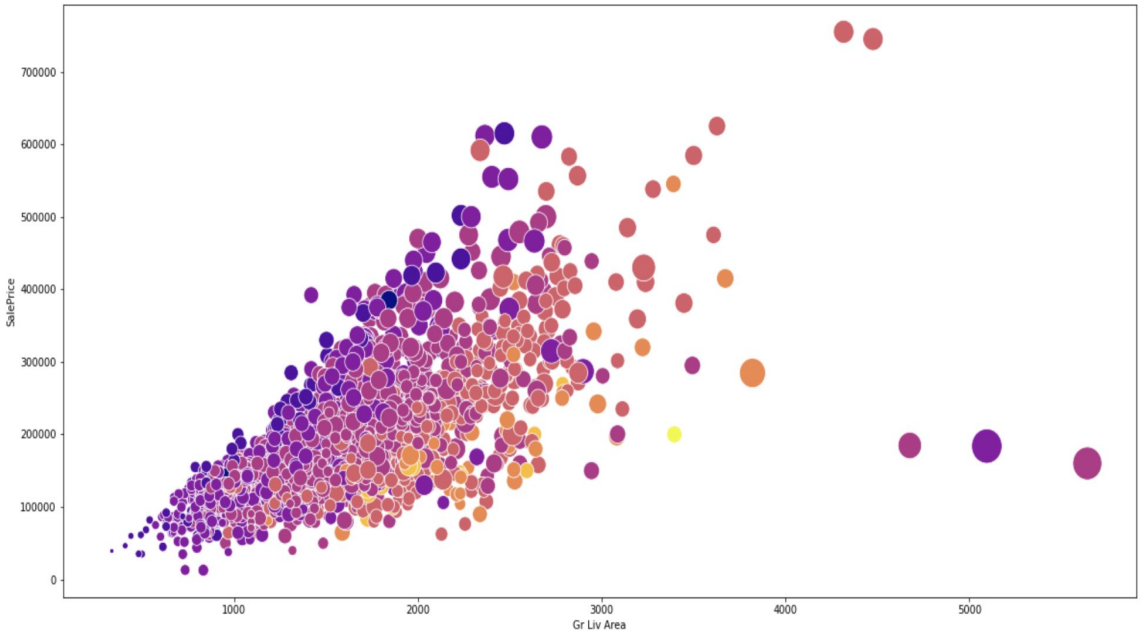
Eg. # of Bedrooms

- **Variable with highest coefficient/predictive power for outcome**



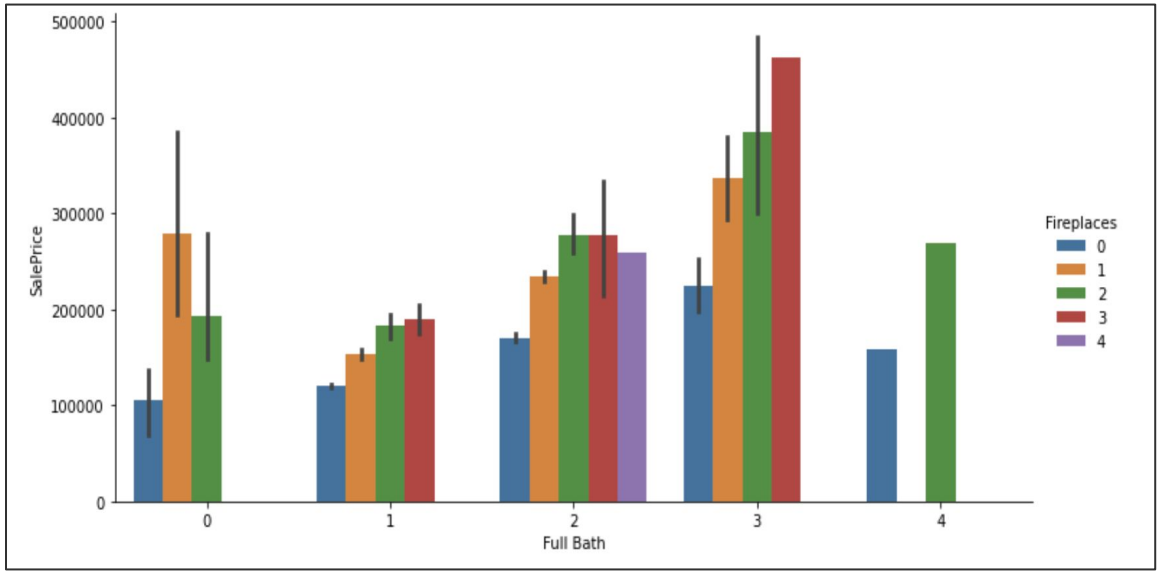
Findings # 2

- **Variable with second highest coefficient power for outcome**



Data Analysis & Data Insights

- **Variable with third highest coefficient power for outcome**



Results

&

Recommendations

Analysis Results

Describe the key results of the Analysis

- **List and strength of the key predictors**
- **Performance of the model**

Result #1

- The regression model explained ___% of the variance….

Result #2

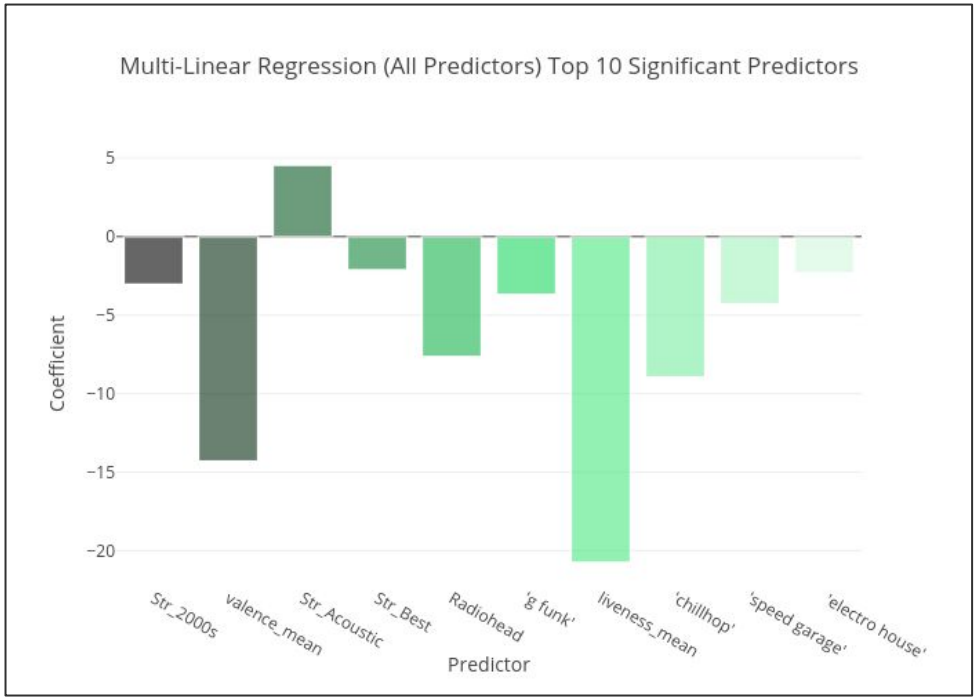
- The strongest predictor of sale price was ___

Result #3

- For every unit increase ***X*** in ___ the average sale price increased by ___

Analysis Results

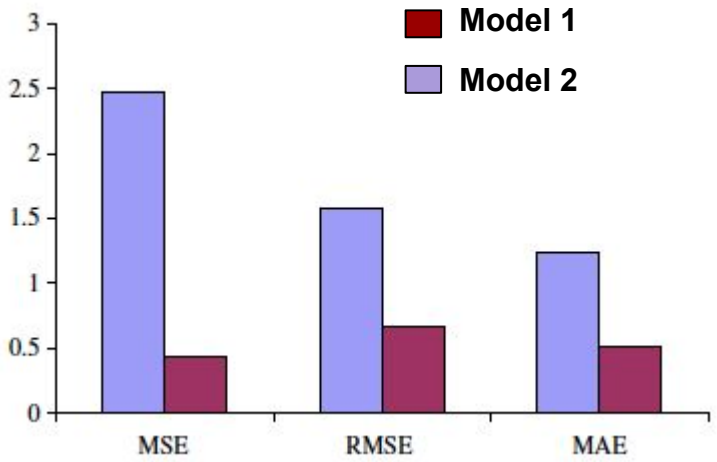
Multiple Linear Regression Model Feature Importance



Analysis Results

Multiple Linear Regression Performance

	MAE	RMSE	R2
ML	2.40	5.17	0.99
MC	7.59	10.57	0.96
RR	8.59	11.22	0.95
PD	5.18	8.22	0.98



Next Steps, Recommendations & Improvements

Lessons learned; possible project improvements; next steps time permitting

- **Source other data sets**
- **Research another data science technique/model**
- **Additional data, new features, other ideas**

1. **Recommendation**

2. **Recommendations**

Appendix
(Not required for Flatiron Presentation)

Example Visualizations

Model selection process:

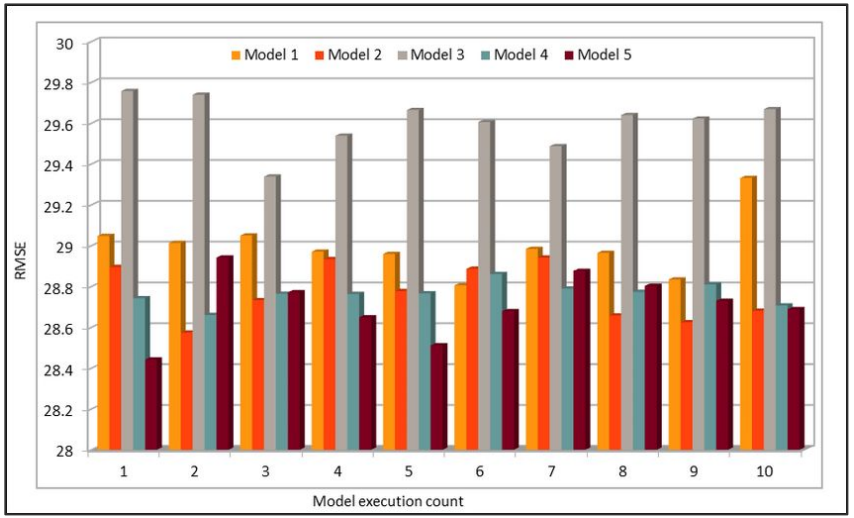
- Regression
- Benchmark Model
- Other models
- Hyper parameter tuning
- Cross validation

Visualization of the evaluation metrics

- Bar chart of error terms
- Pie chart of the coefficients
- Line chart of the

Model Performance Results

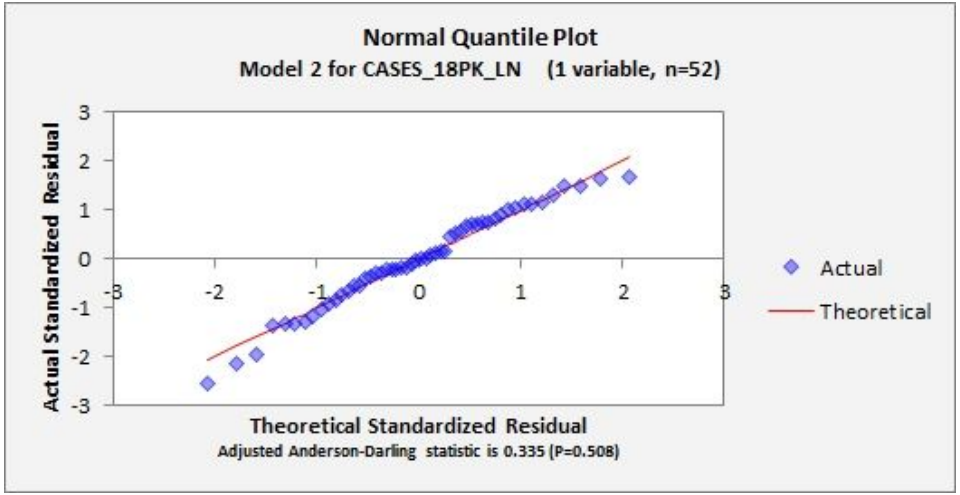
- Final MAE, RMSE, Accuracy, F1 Score, AUC, ROC
- Confusion matrix
- MAE, RMSE
- Accuracy, F1 Score
- AUC, ROC



Assumptions

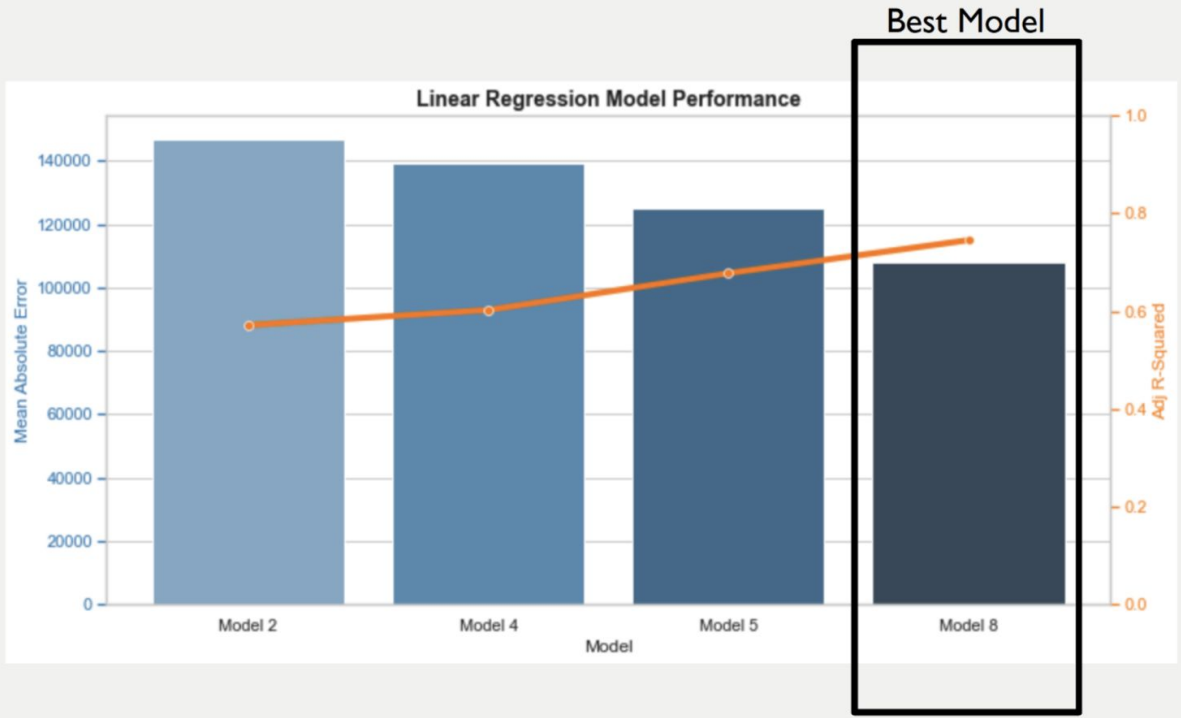
State any assumptions used for the analysis and obtaining results or the results of the tests of the assumptions of a linear regression

1. Results of the test of linearity
2. Results of the test of multicollinearity
3. Results of the test of heteroscedasticity
4. Results of the test of normality



MODEL PREDICTIONS

Model Summary



Model 8 performed best with 74% of variance explained

- 0.745 Adjusted R-squared
- $108K Mean Absolute Error
- $182K Root Mean Sq. Error
- 0.00 P-Value

Variables in Model 8

- Target: Price
- Predictors (10):
  - Sq. Ft. Living Space
  - Sq. Ft. Per Bedroom
  - View
  - Grade
  - Condition
  - Waterfront
  - Basement Indicator
  - Renovation Indicator
  - Miles From Seattle
  - Zip code Median Price-Per-Sq. Ft.

Reference