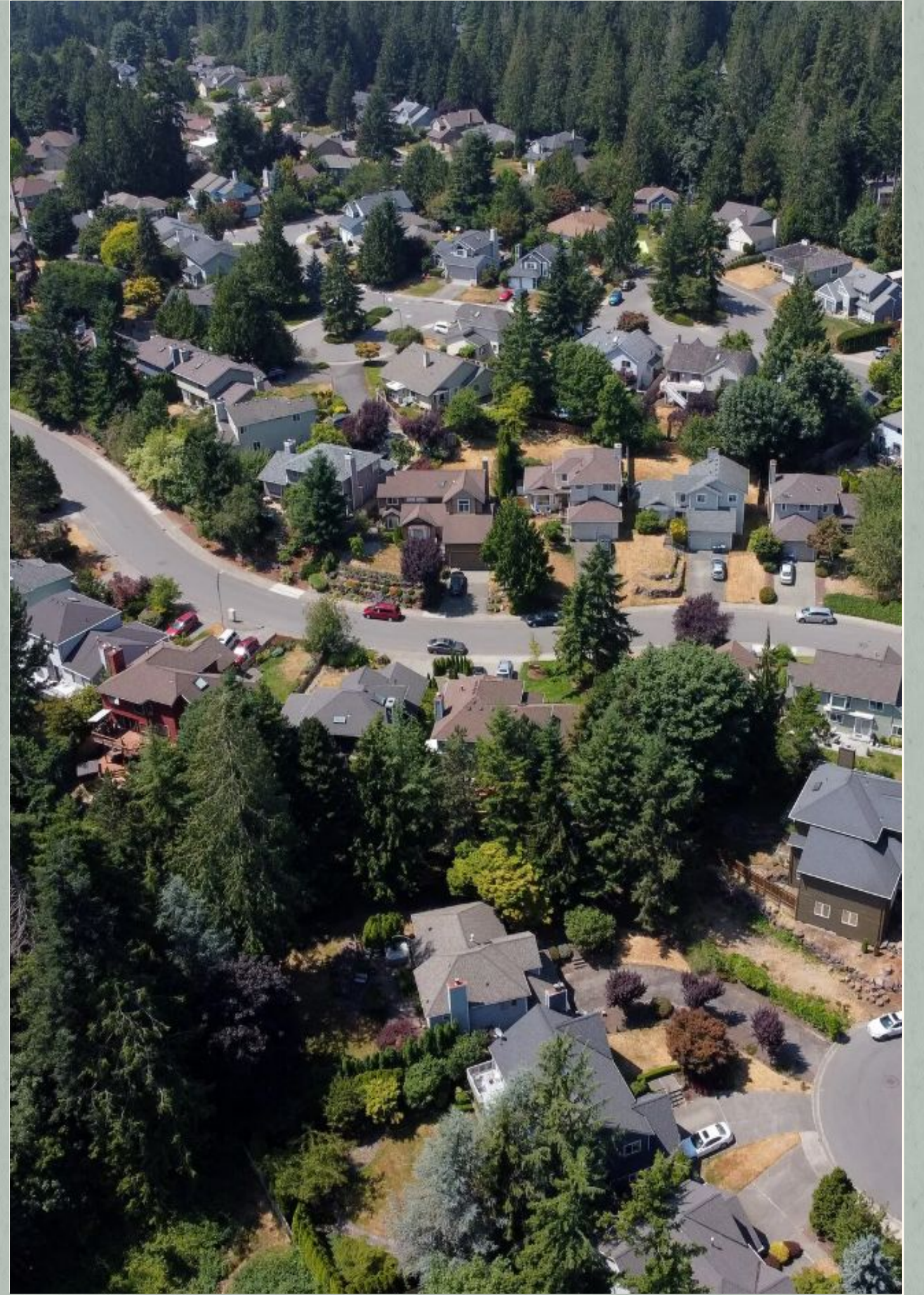


KING COUNTY REAL ESTATE

A Presentation by: Albane Colmenares

09/04/2023



AGENDA

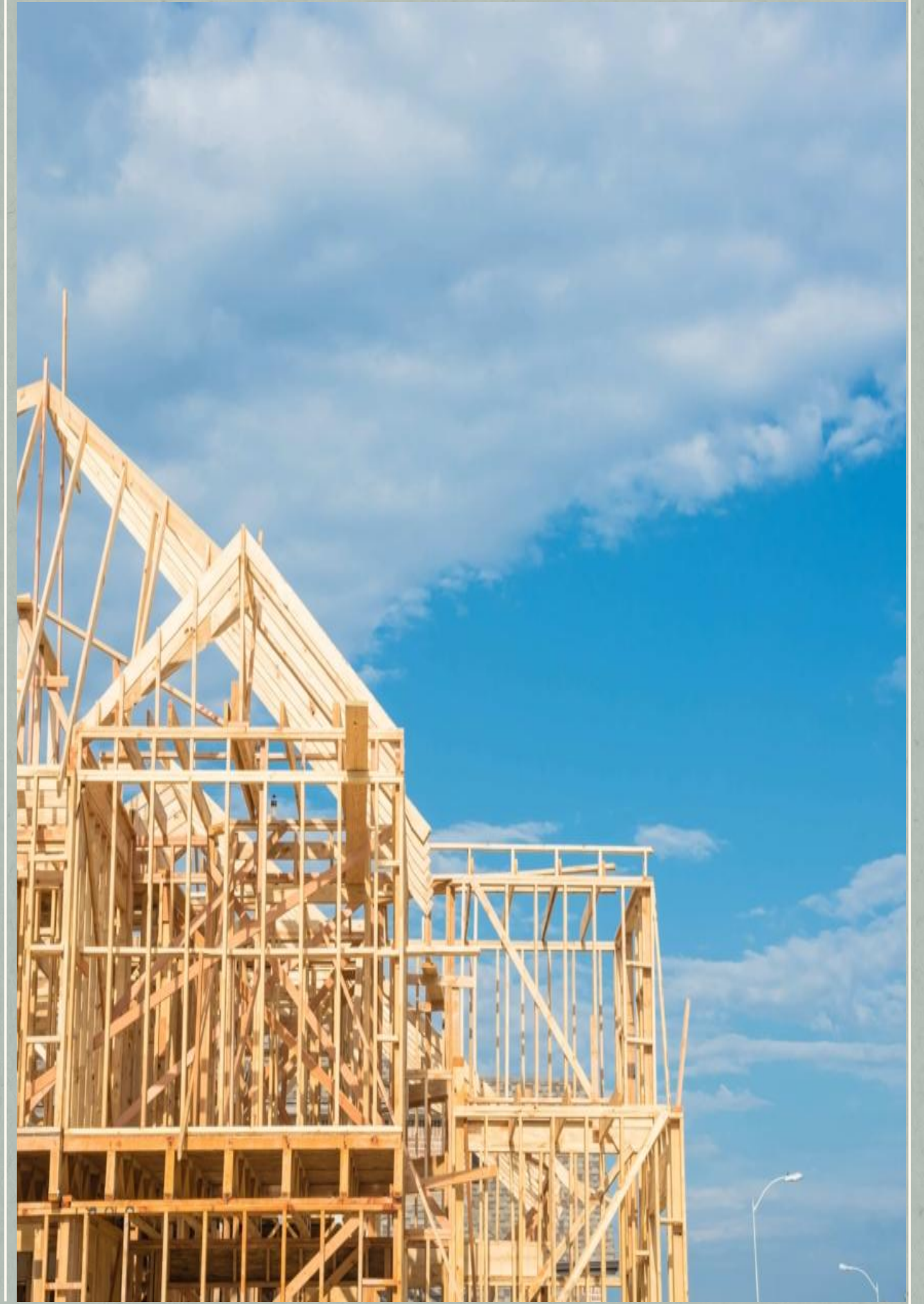
1 BUSINESS UNDERSTANDING

2 MODELING & REGRESSION RESULTS

3 RECOMMENDATIONS & NEXT STEPS



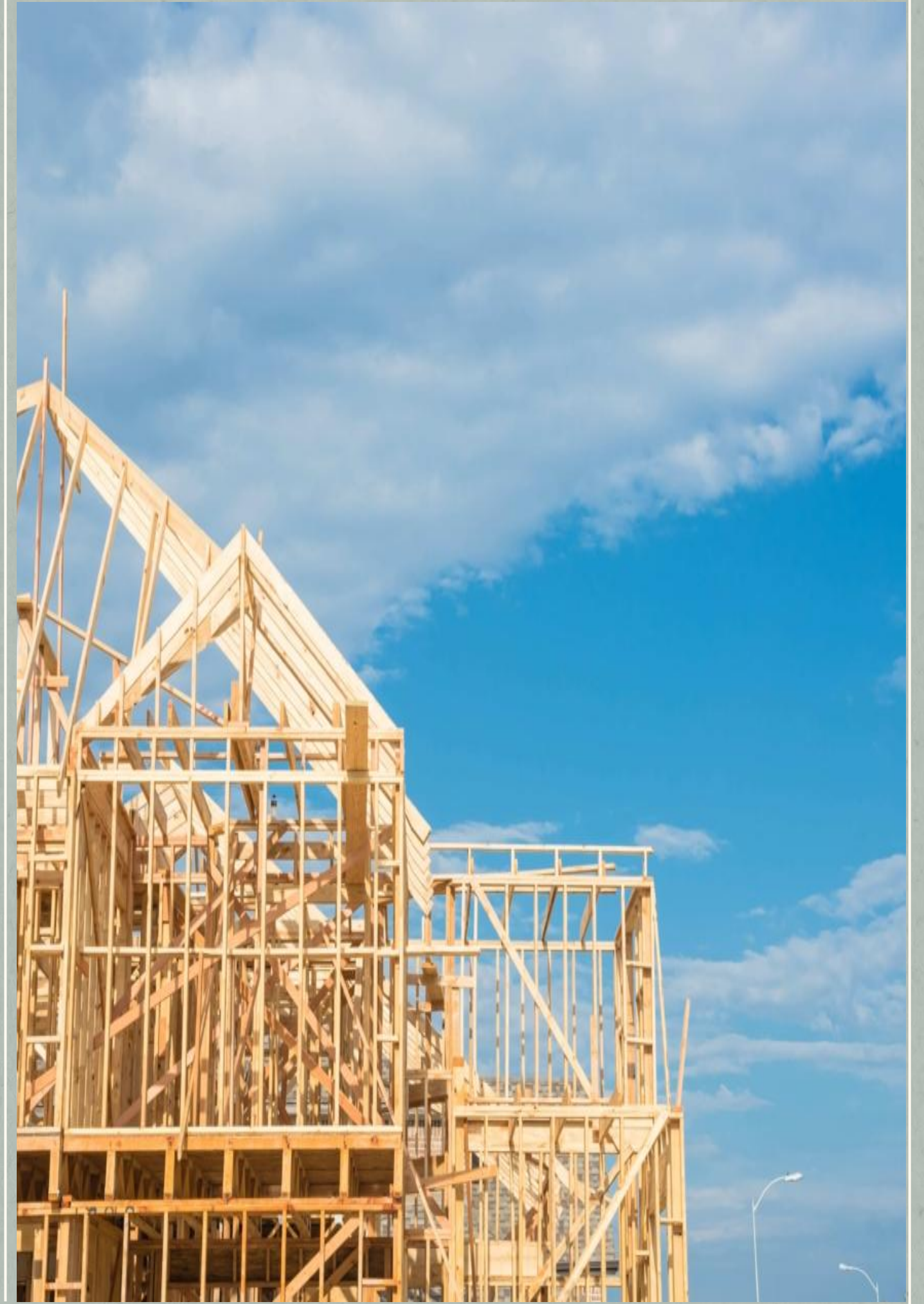
BUSINESS UNDERSTANDING



RENOVATIONS: HOW WILL THEY IMPACT YOUR SELLING PRICE?



MODELING



MODELING

1. Simple Linear Regression

- A. Highest correlation with price
- B. Initial simple linear regression
- C. Removal of outliers
- D. New simple linear regression

2. Multiple Linear Regression

- A. Categorical feature
- B. Condition & square footage of living area

3. Multiple Linear Regression

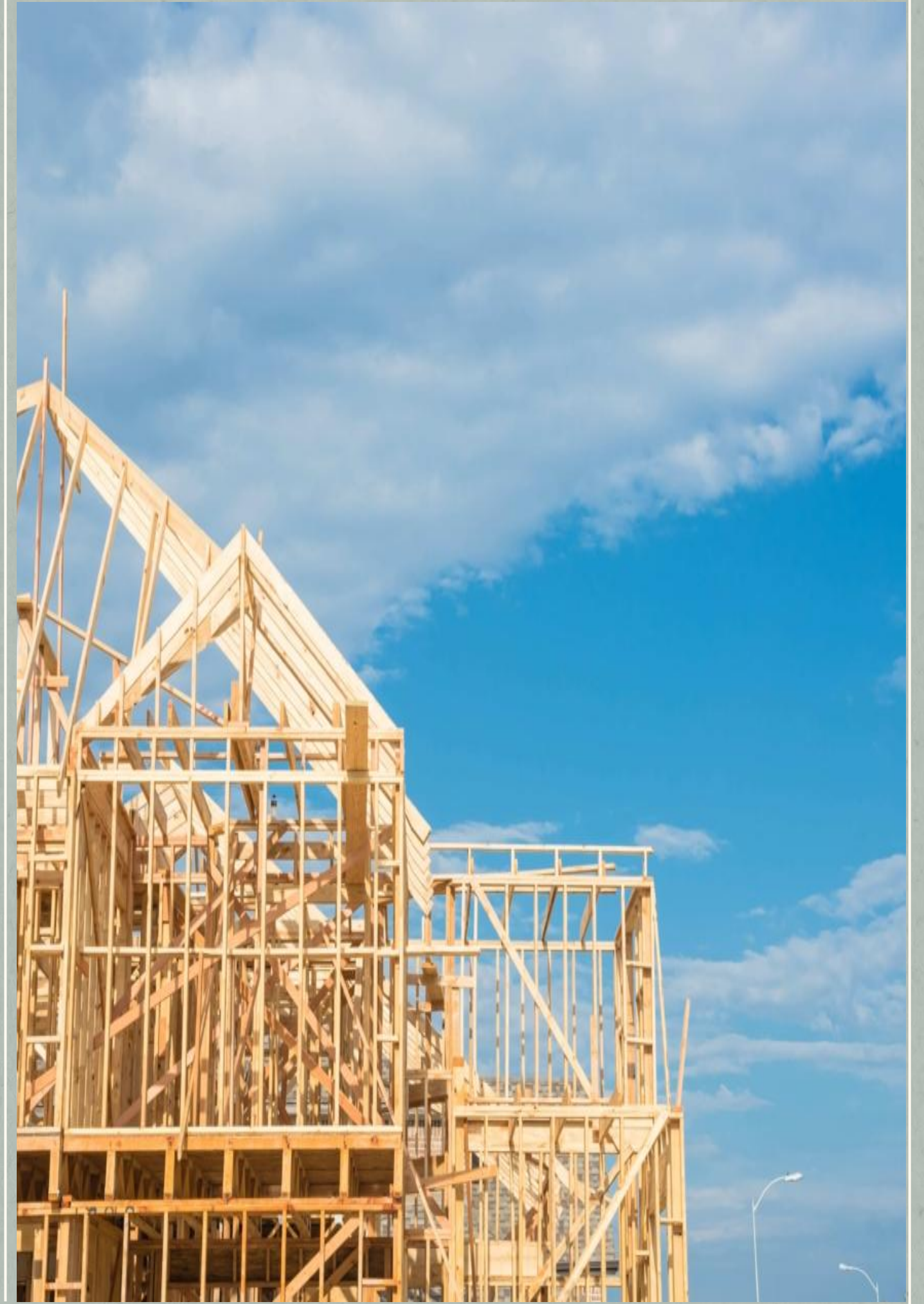
- A. Numeric features
- B. Part 1: all variables

4. Multiple Linear Regression

- A. Final model



REGRESSION RESULTS

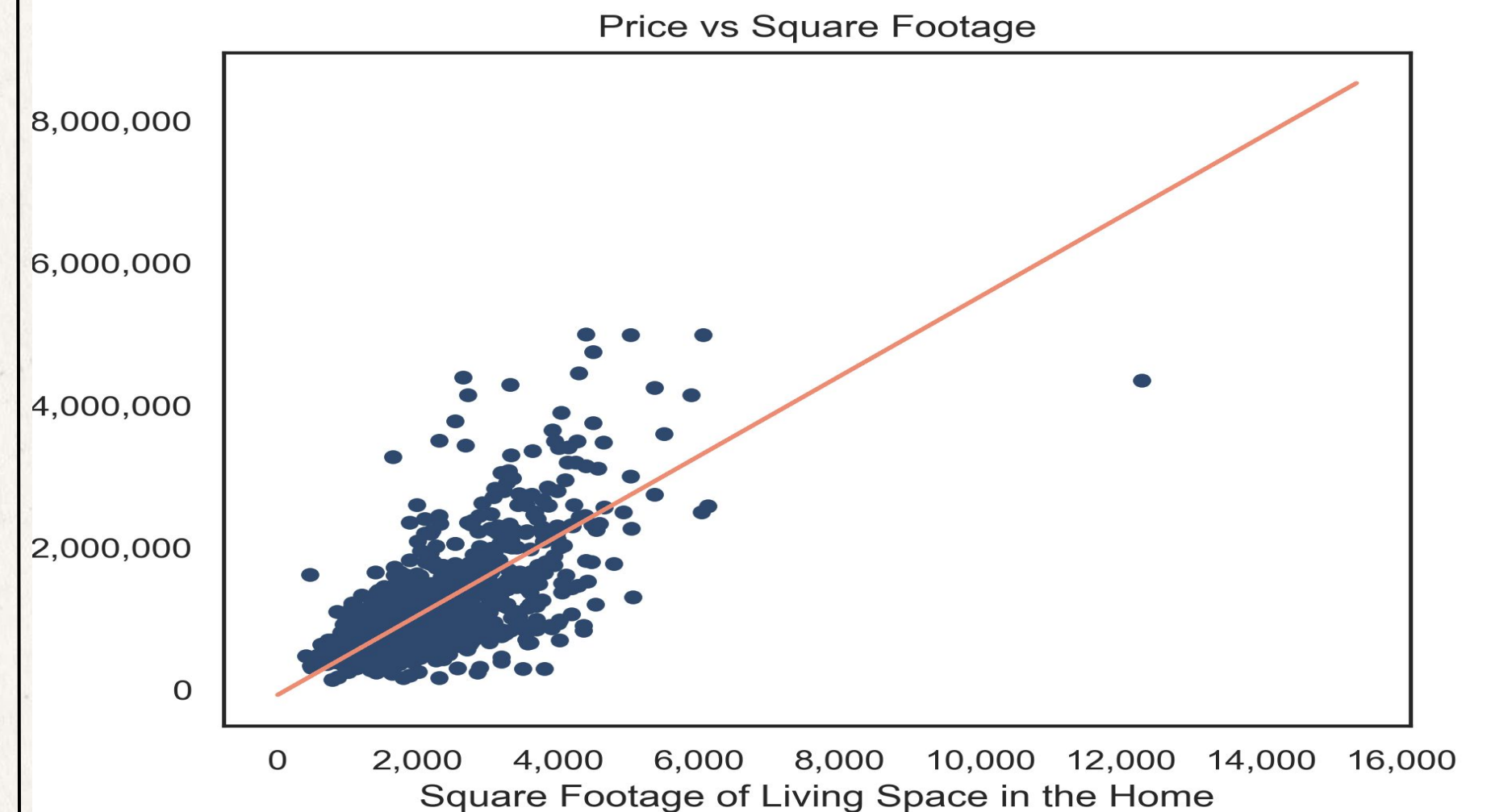


SIMPLE LINEAR REGRESSION

HIGHEST CORRELATION WITH PRICE
Square footage of living space

Results

1. Model Explanatory Power: 42.8%
Adjusted R-squared
2. Price increase by additional square foot: \$475.92
Coefficient & p-values
3. Estimated price for 0 sq. ft house: \$77,330
Constant coefficient or intercept

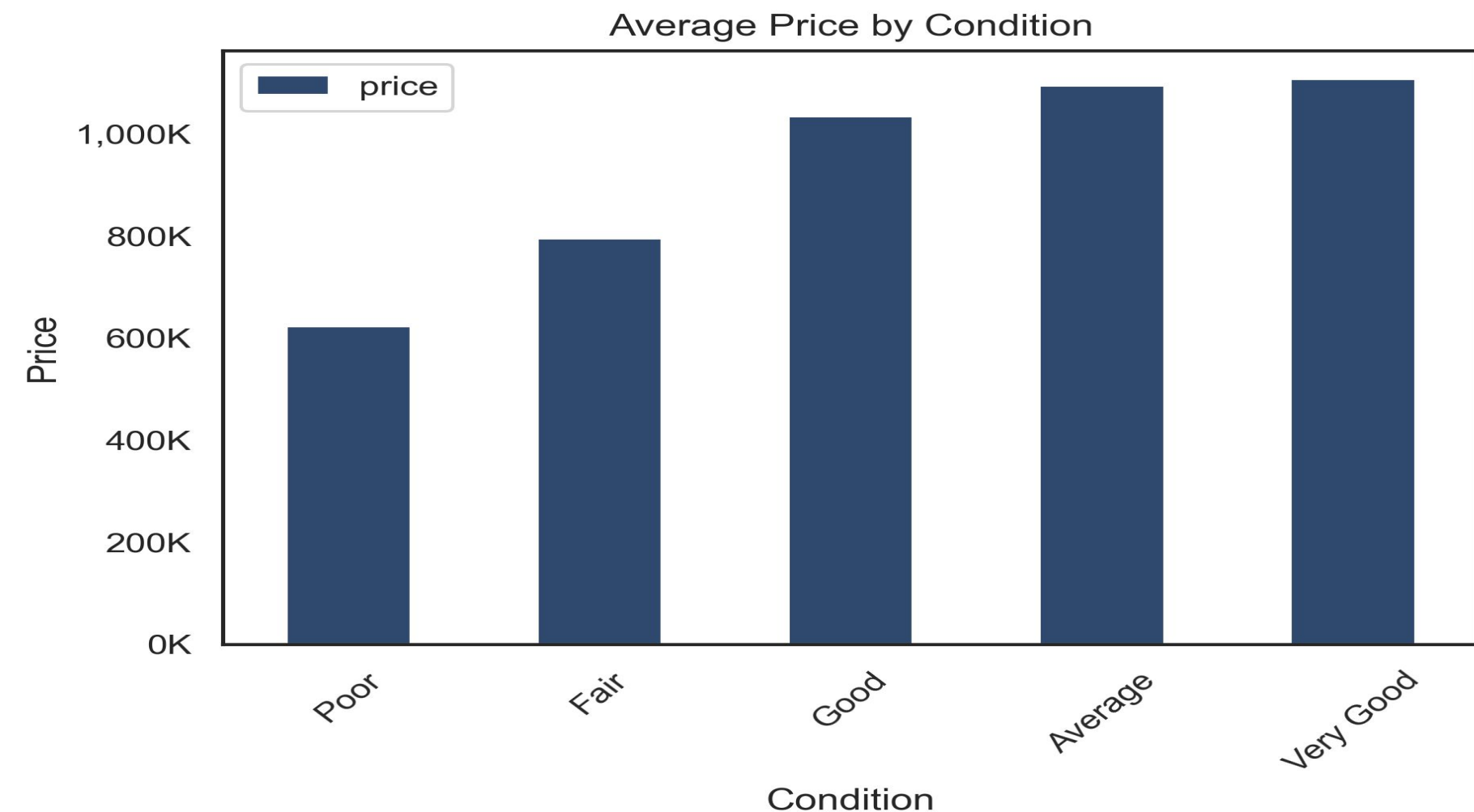


MULTIPLE LINEAR REGRESSION

CATEGORICAL FEATURE
Overall condition of the house

Results

1. Condition explains 42.9% of the price's variance
Adjusted R-squared
2. "Very good" condition generates \$135,700 increase
Price increase by additional square foot: \$477.25
Coefficient & p-values
3. Significant relationship with price
F-statistic: 4478, Prob (F-statistic): 0



MULTIPLE LINEAR REGRESSION

PART 1: NUMERIC FEATURES

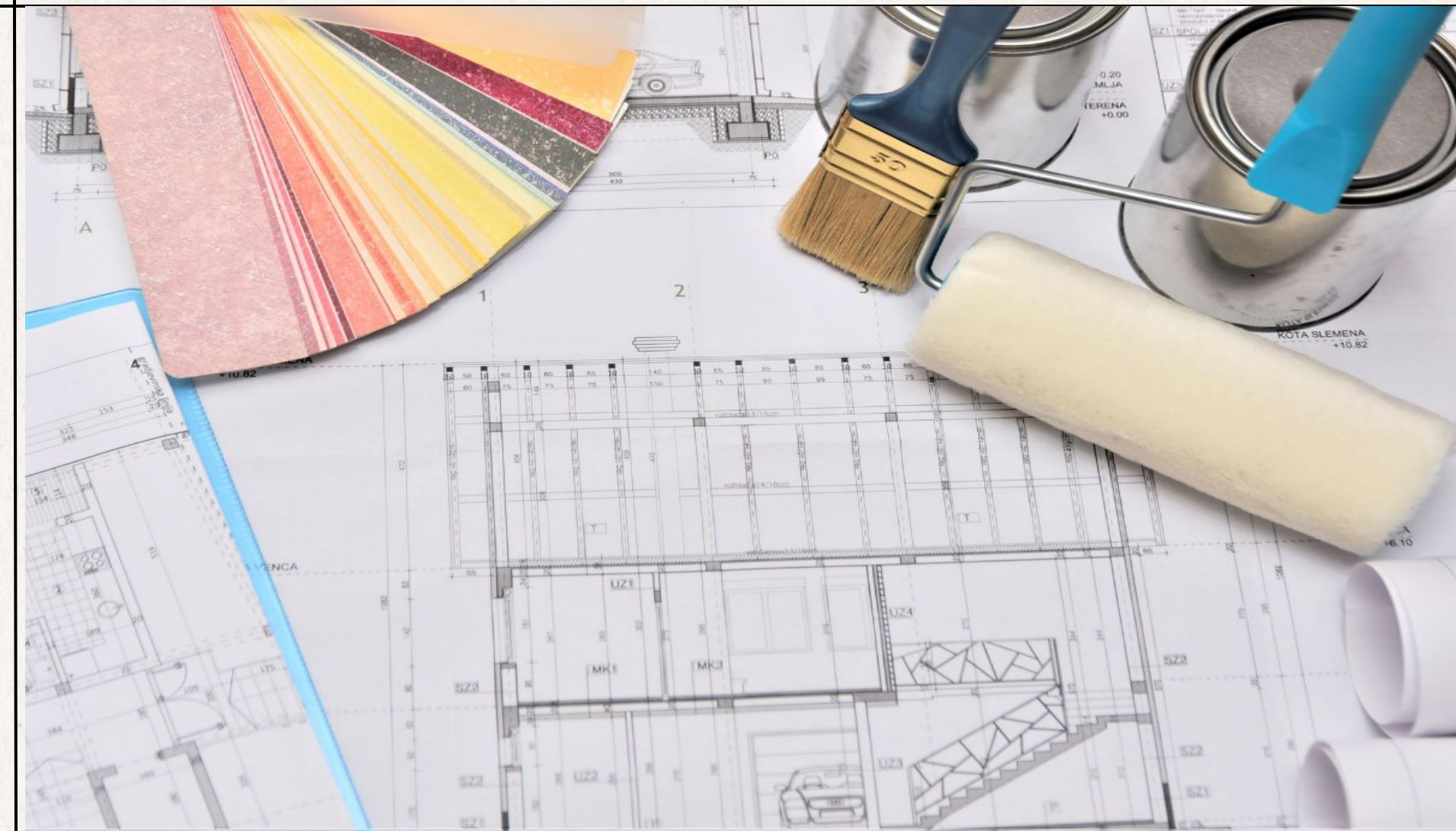
- Square footage living area
- Square footage patio
- Square footage above
- Square footage basement
- Square footage garage
- Square footage lot
- Year built
- Floors
- Bedrooms
- Bathrooms

Results

1. 46.4% of variance in price explained by numeric features

Adjusted R-squared

2. Features associated with an decrease in price:
bedrooms, square footage of garage, year built
 3. No significant impact on price: square footage of lot
Coef: 0.0019 and p-value (0.97) greater than 5%
- Coefficient & p-values*



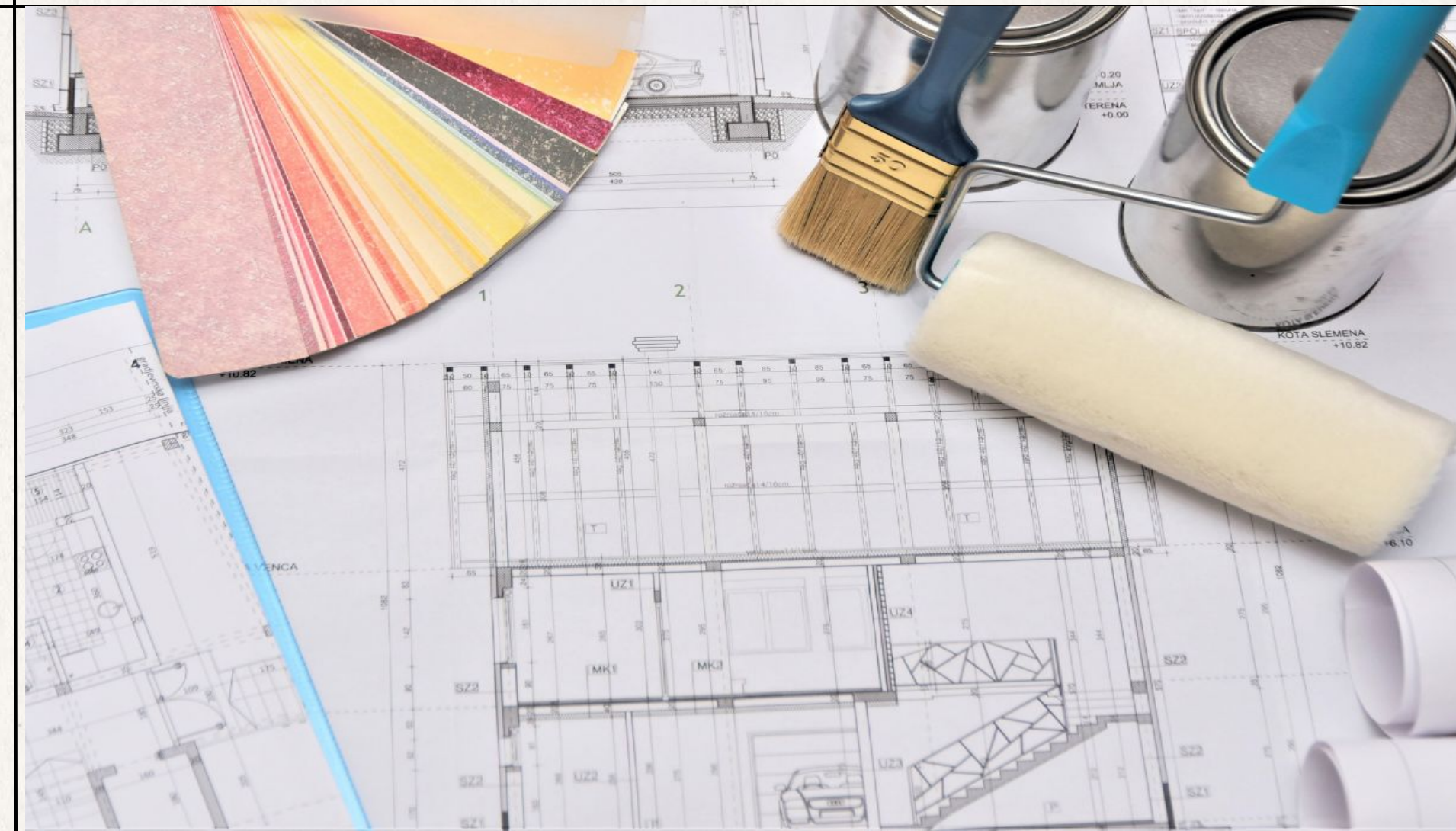
MULTIPLE LINEAR REGRESSION

PART 2: NUMERIC FEATURES

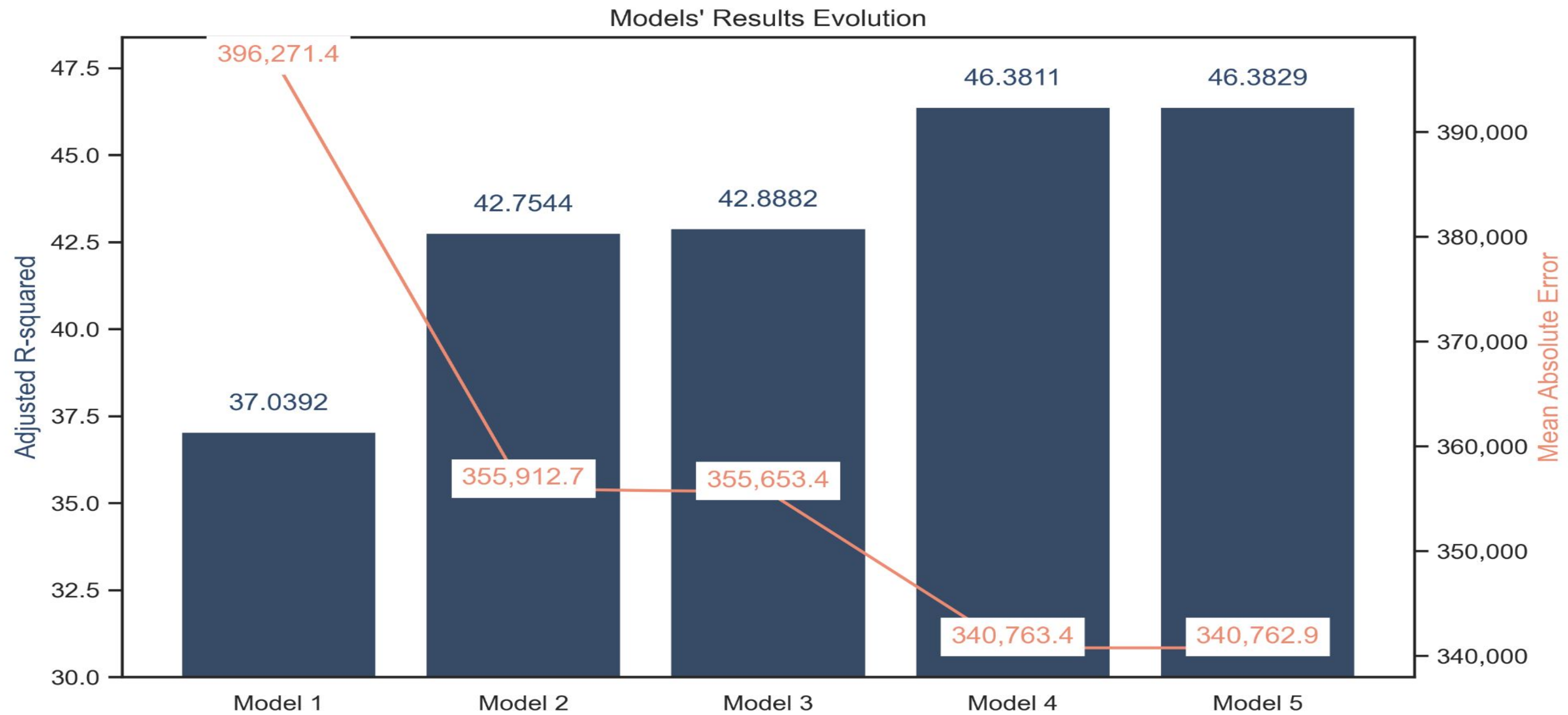
- Square footage living area
- Square footage patio
- Square footage above
- Square footage basement
- Square footage garage
- ~~Square footage lot~~
- Year built
- Floors
- Bedrooms
- Bathrooms

Results

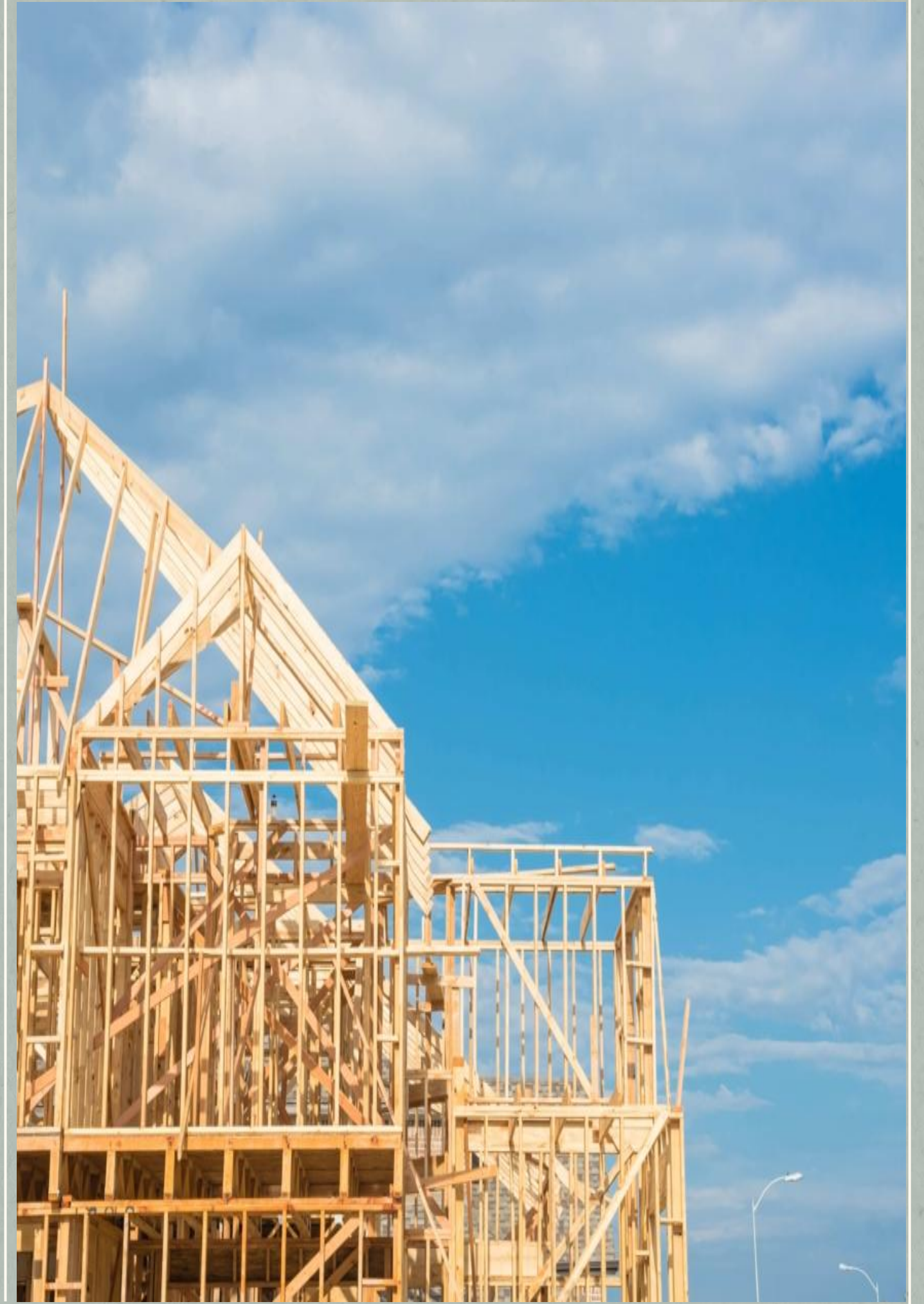
1. 46.4% of variance in price explained by numeric features
Adjusted R-squared
2. Significant impact on price: F-statistic increased (2,866)
3. Features associated with highest increase in price:
 - a. bathrooms: \$116,200
 - b. square footage living area: \$303.3 → \$30,330 for 100
 - c. floors: \$33,900
4. Lowest association with price increase:
 - a. square footage of garage and basement



MODELS' RESULTS



RECOMMENDATIONS

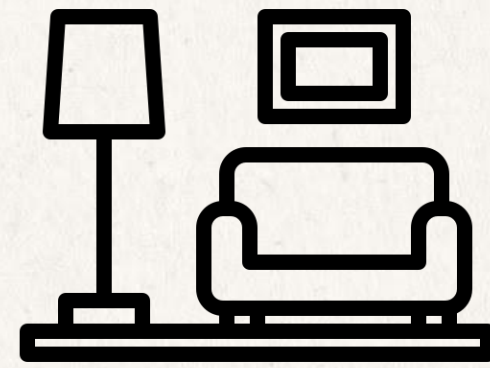


RECOMMENDATIONS



Very Good Condition

Associated with an increase of \$135,700 in price



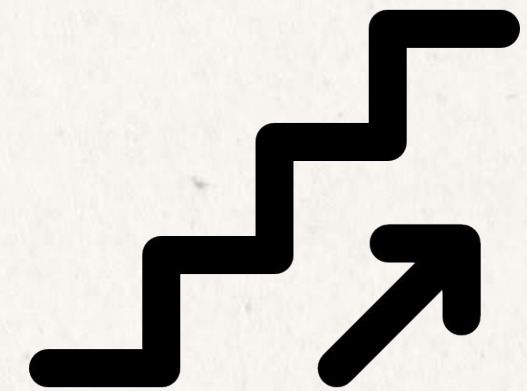
Living Area Square Footage

Associated with an increase of \$303.26 in price per square foot.



Bathroom

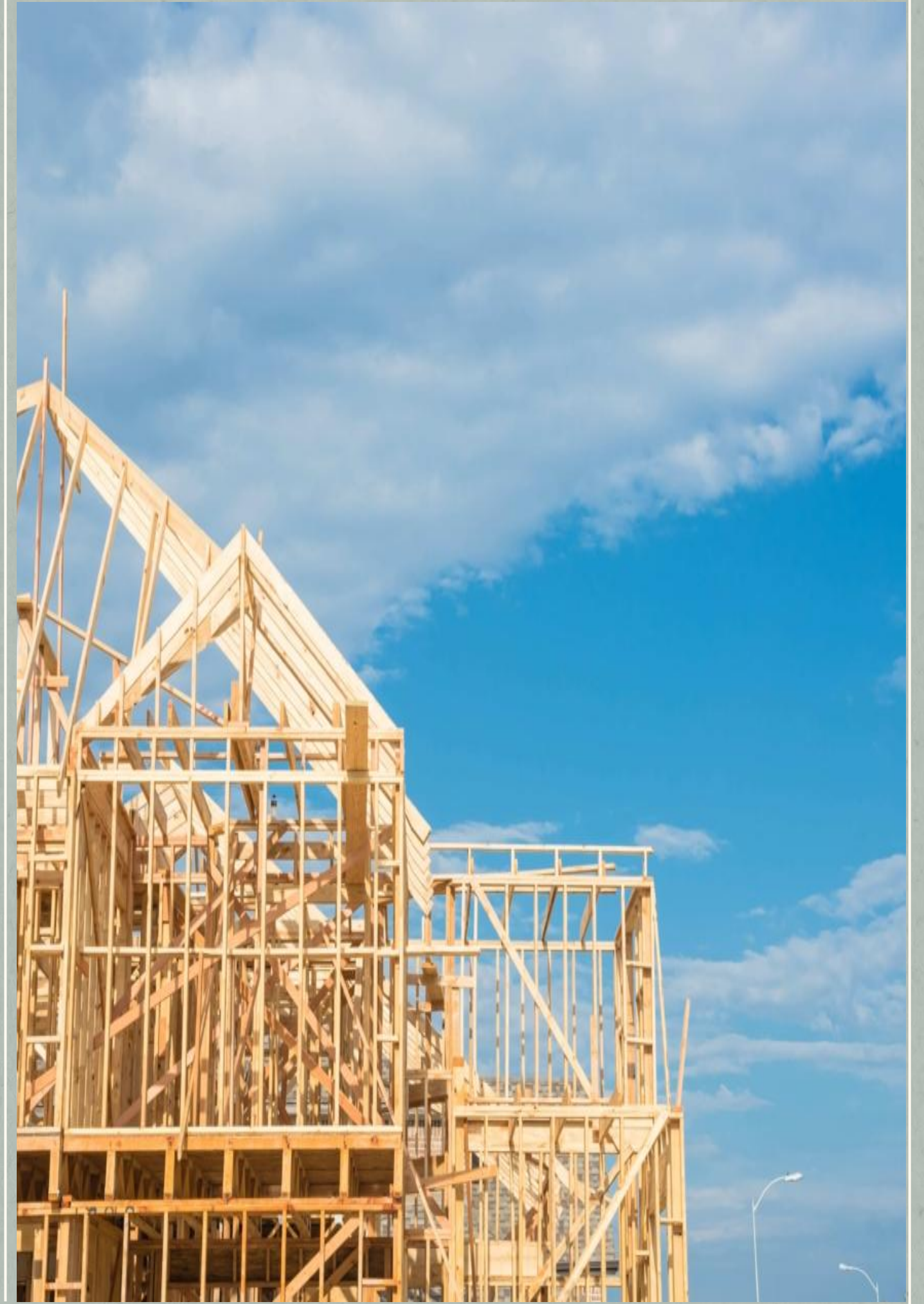
Add one bathroom if your house permits it.
Associated with \$116,200 in price increase



Floor

Add one floor. Consider a mezzanine. Estimated increase in price: \$33,900

LIMITS & NEXT STEPS



LIMITS & NEXT STEPS

LIMITS

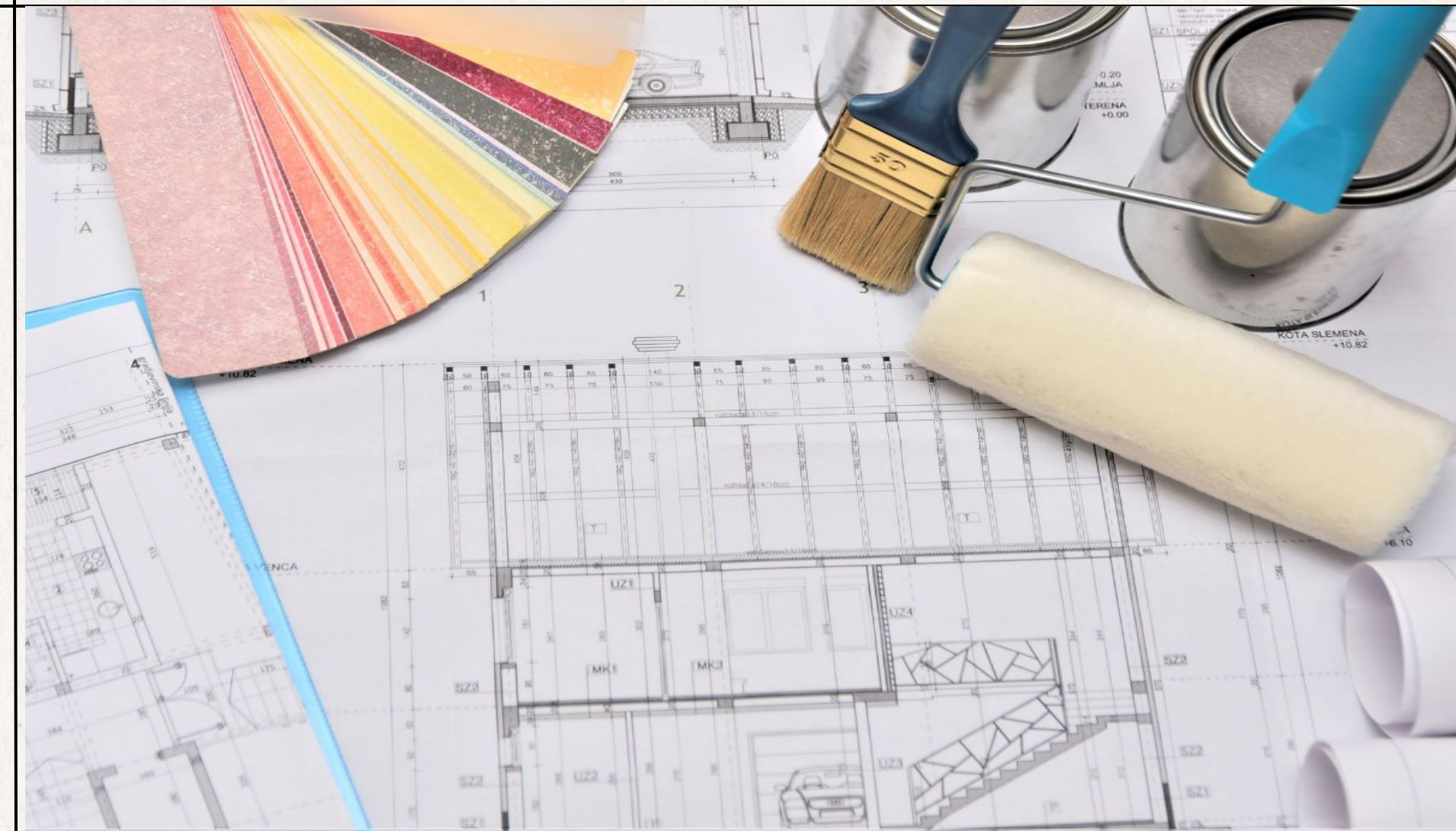
Linear Regression Assumptions Threatened

- Not a normal distribution
- Multicollinearity among independent variables

NEXT STEPS

Linear Regression Assumptions Threatened

- Normalizing distribution
- Scaling data to improve predictions and lower MAE



CONTACT INFORMATION

For more details, contact:

albane.colmenares@gmail.com



THANK YOU

