



# House Price Prediction in King County



Multiple Linear Regression Model

# What We'll Discuss

## TOPIC OUTLINE

- **Introduction**
- **Business Problems**
- **Data**
- **Approach/Methodology**
- **Analysis**
- **Results**
- **Conclusion**
- **Next Step**

# HOUSES FOR SALE



# Introduction



This project focuses on predicting the house and investigating house sales in the King County area.

# Business Problems

- Build a model to predict house price
- Analysis on providing prediction data to the real estate agents for homeowners who want to buy homes

# Data

This project uses the 2014-2015 King County House Sales dataset, which contains around 21k house sale data.

# APPROACH/METHODOLOGY

1

2

3

4

5

**OBTAIN**

**SCRUB**

**EXPLORE**

**MODEL**

**INTERPRET**

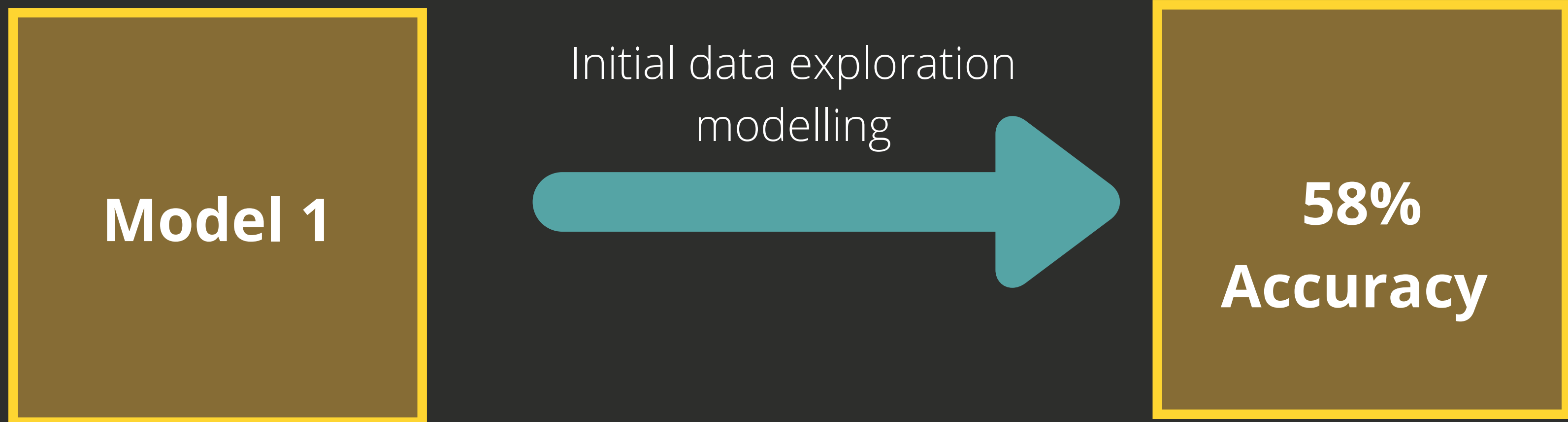
# Analysis



# House Features



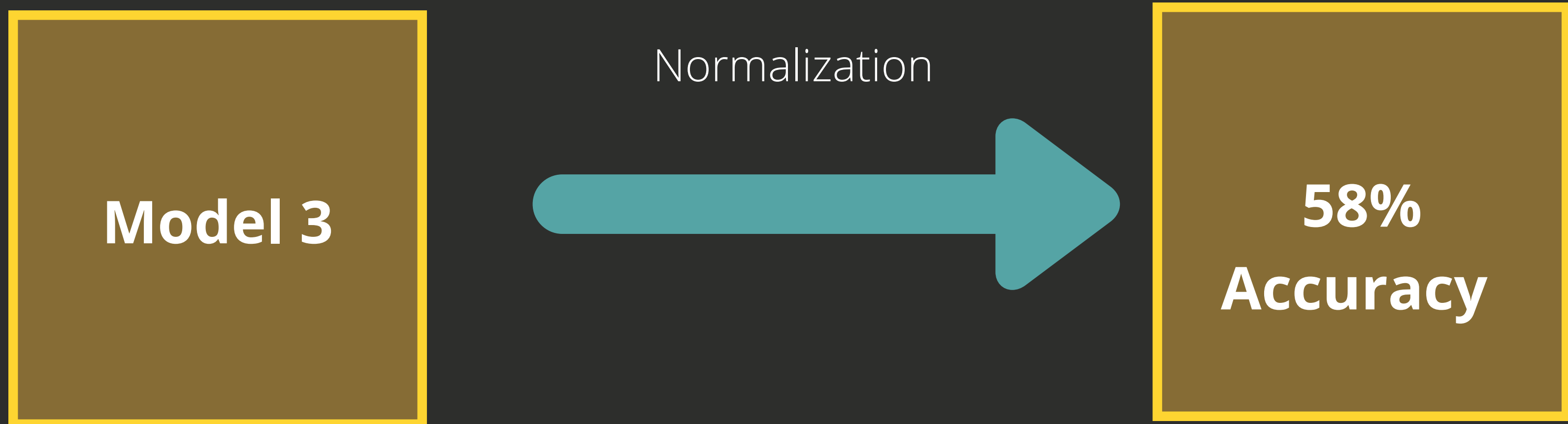
# Baseline Model



# Feature Engineering and Testing



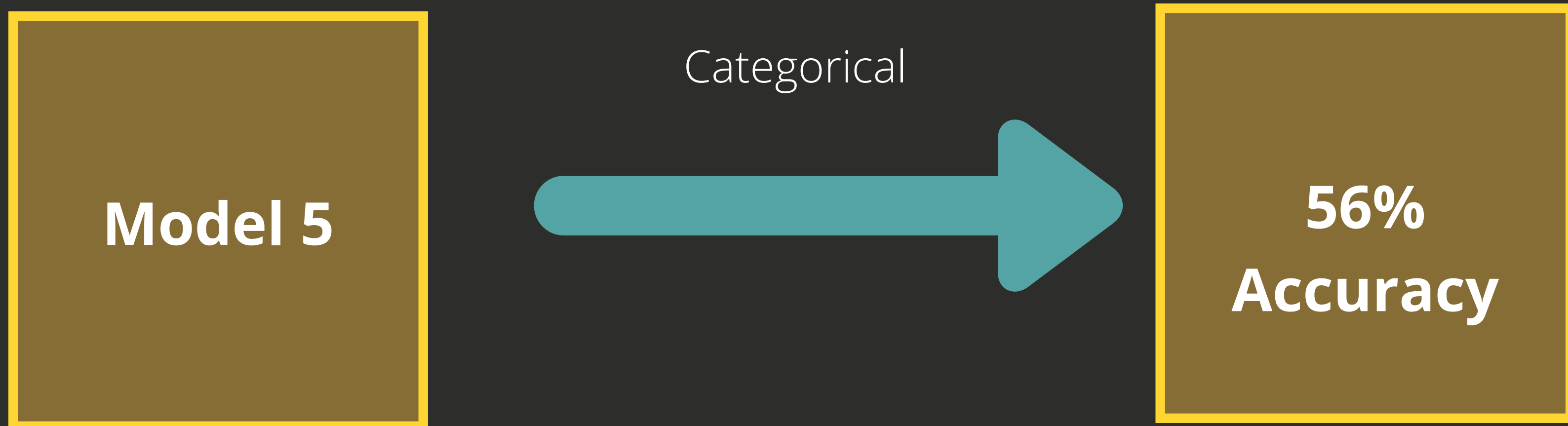
# Applying Log Transformation



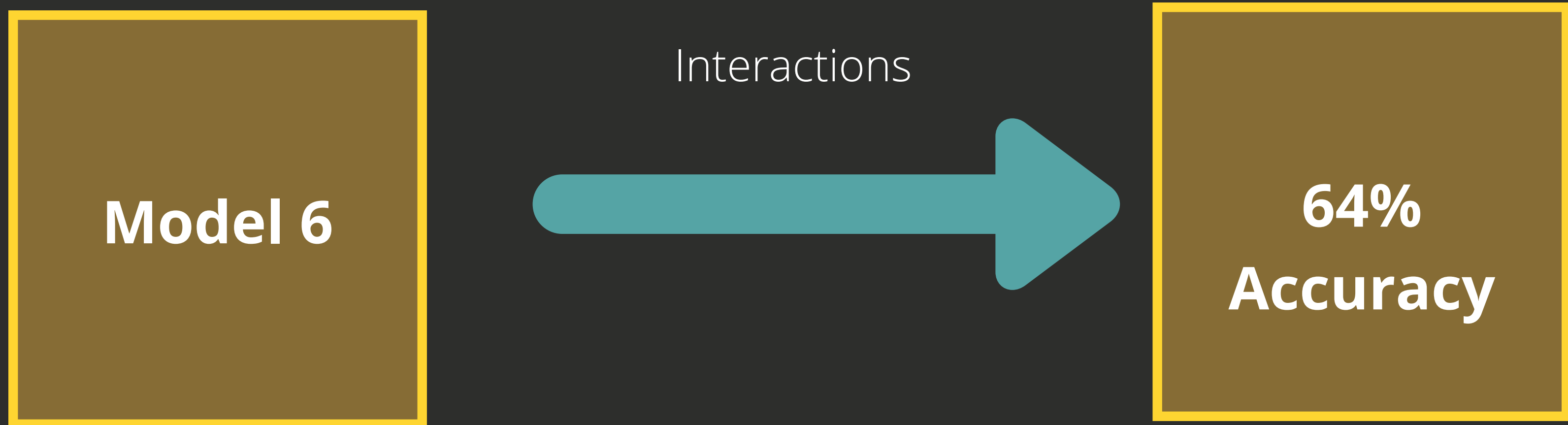
# Multicollinearity Checks



# Applying Categorical Terms



# Applying Interaction Terms

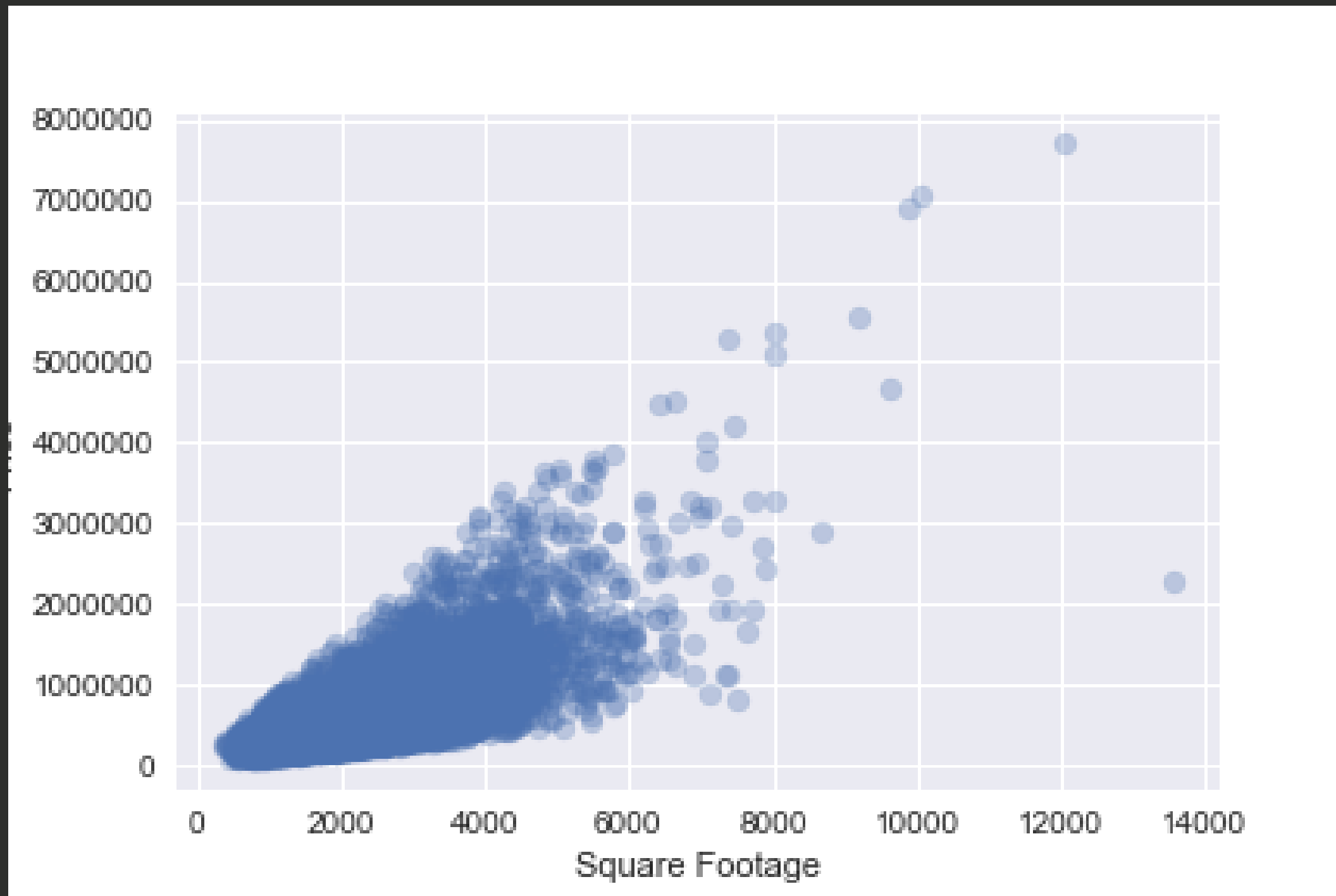


# Final Model

Model	Model 6
Features	9
R2 Score	0.64
RMSE Value	0.31
Square Footage	0.108
Bathroom	0.082

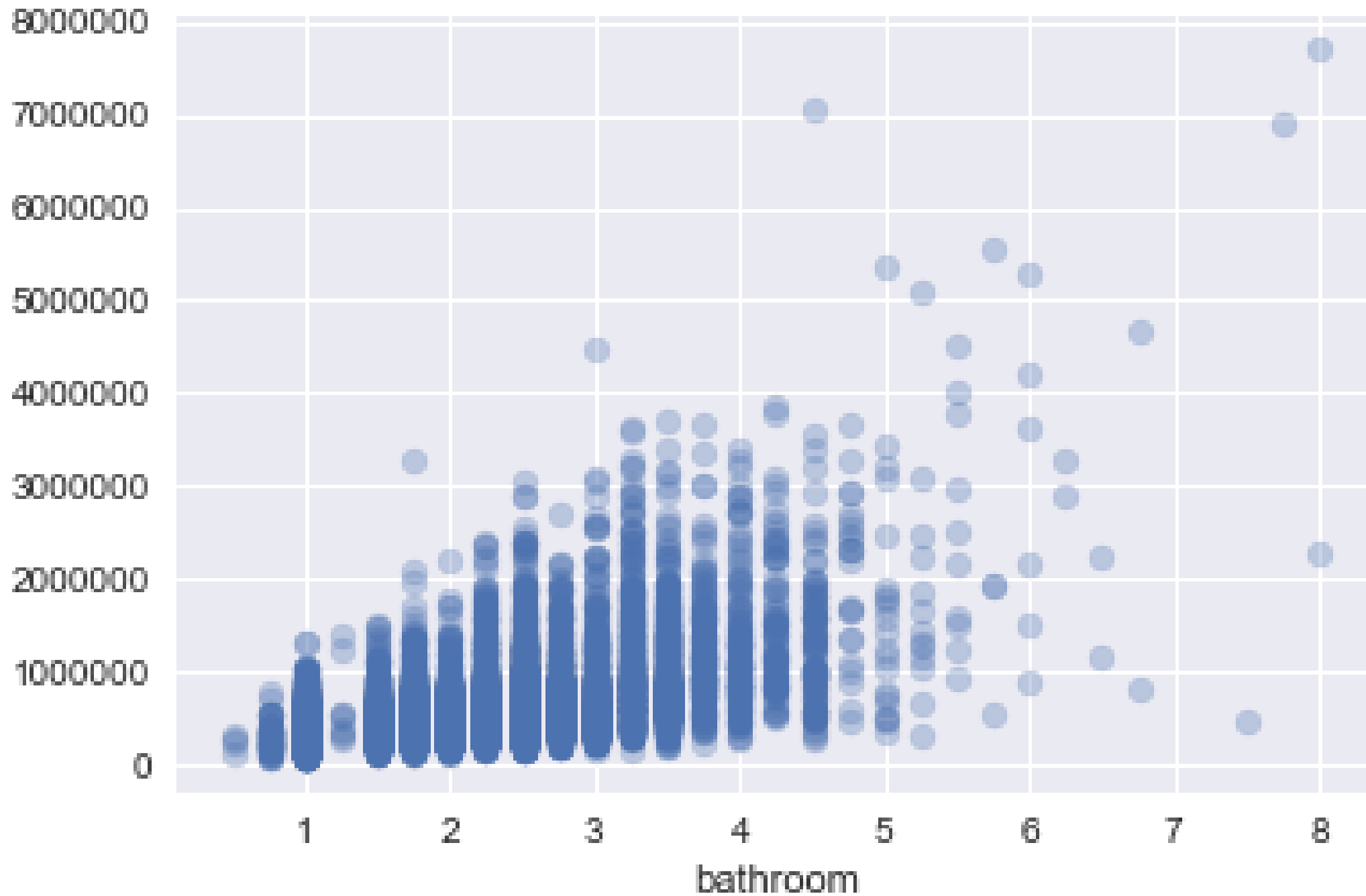


# Square Footage Vs Price



The most distinctive feature is the square footage, where the larger house will be more expensive. It is noticeable that the graph is getting broader as the square foot increases. So for every 1% increase in square footage, price increases by 0.11%

# Bathroom Vs Price



From the graph, it is clear that as the number of bathrooms increases the house price increases. So for every 1% increase in the bathroom, price increase by 0.08%

# Conclusion:

- If you are looking for housing that won't make your bank account fragile, then go for the housing with a minimal bathroom so that you could share them
- It's advisable to shrink on square footage that would make our house purchase a very cost-effective one.

# NEXT STEP

<b>Linear Assumptions</b>	Multicollinearity
<b>Distance from economic hubs</b>	Zip code data
<b>Outliers</b>	Handling outliers

# THANK YOU!

EMAIL:

[janakipurushothamman@gmail.com](mailto:janakipurushothamman@gmail.com)

GITHUB:

[@JanakiGanesh](https://github.com/JanakiGanesh)