

oooo

JOSHUA LIN

REAL ESTATE
INVESTMENT
CONSULTING

HIGH-IMPACT VARIABLES THAT PREDICT HOUSING PRICES

oooo

TABLE OF CONTENTS

- Data Science Problem
- Why?
- Methodology
- Insights
- Recommendations



oooo

PROBLEM STATEMENT



**As qualities of a
house change,
how does home
price change?**



WHY?

oooo

Use cases of an ML model in new construction investment:

- Identifying pricing inefficiencies

An accurate model can predict the "real" value of a property

- Identifying features that contribute to home price

Insights can prove useful when considering prospective investments

- Ensuring property value retention

By identifying high-impact features that a prospective investment may or may not have

METHODOLOGY

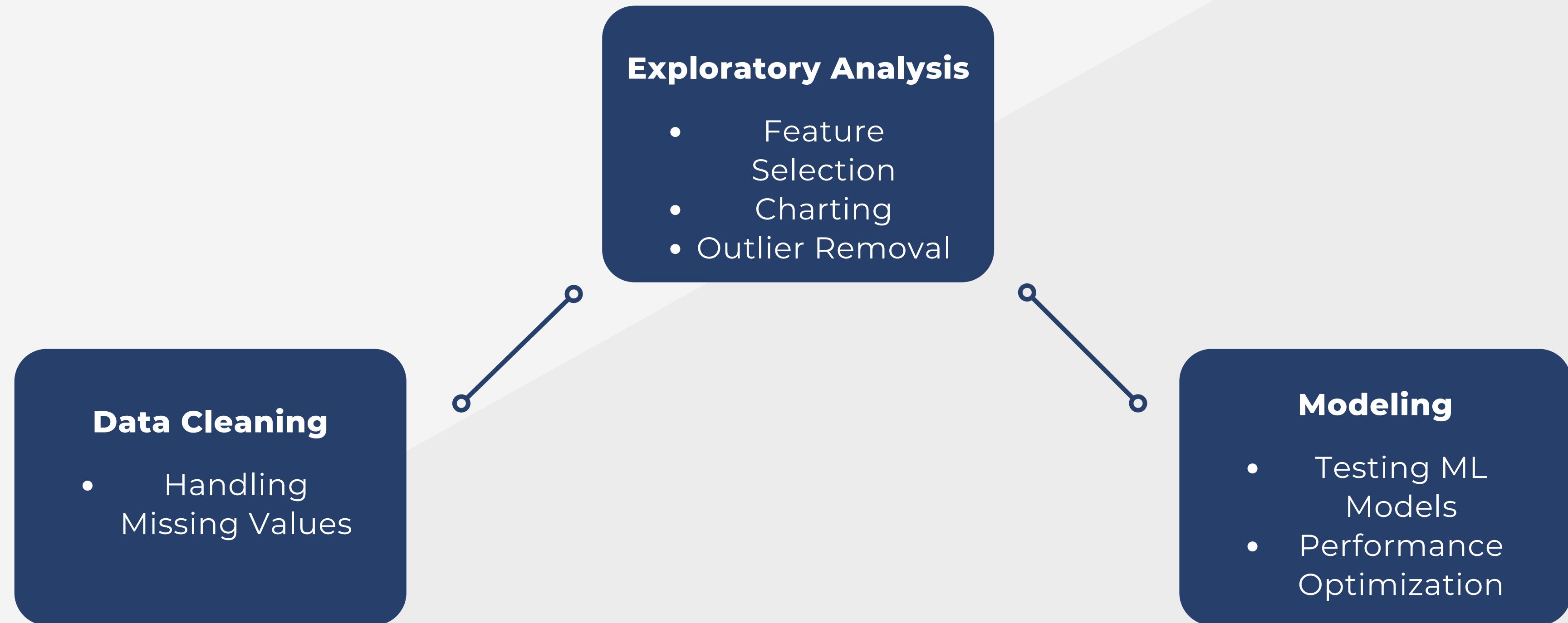
Data Cleaning

- Handling Missing Values

METHODOLOGY



METHODOLOGY



OOOO

INSIGHTS

The variables that have the strongest effect on home price as identified by our model include:

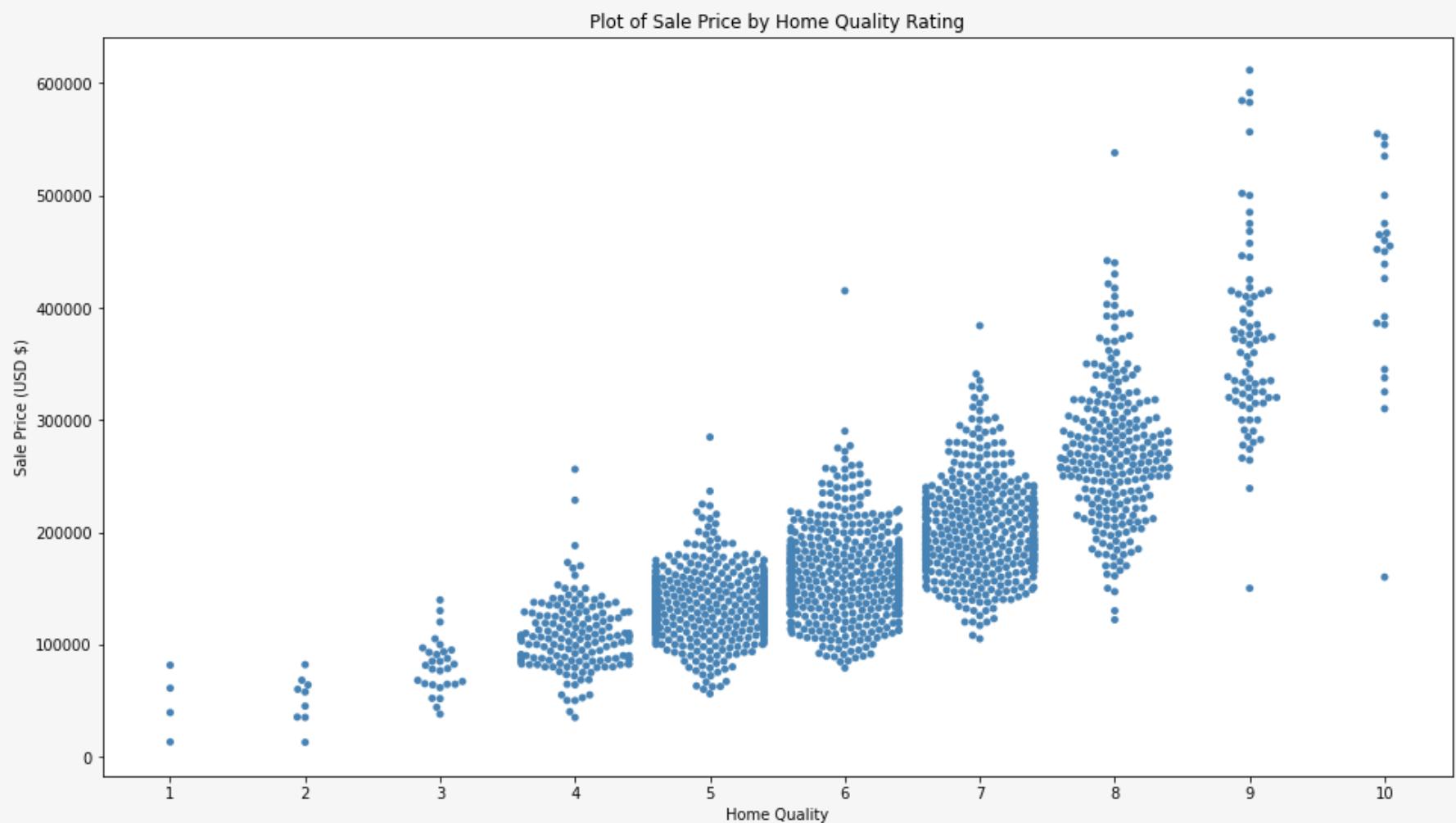
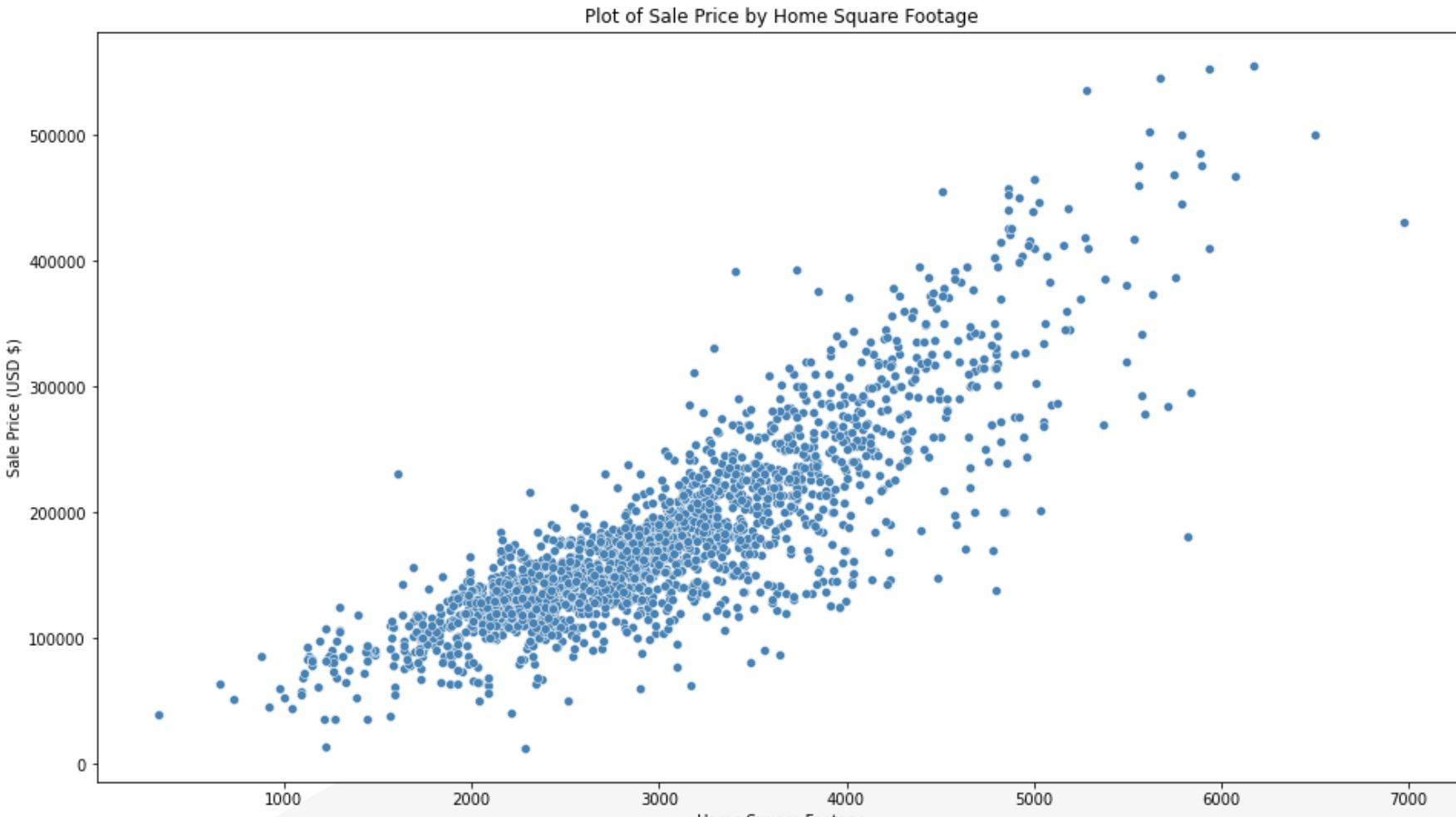
1. Home square footage
2. Home quality
3. Neighborhood
4. Home type

For every 1 unit increase in home square footage, home price increases by 17%.

For every 1 unit increase in home quality rating, home price increase by 10%.

Neighborhood example:

A 16% increase in home price if the home is in Crawford vs. Sawyer





RISK EXPOSURE

R2 score:

84%

Root Mean Squared Error:

\$23,638



TAKEAWAYS

Detect pricing inefficiencies

When prices in the market do not indicate actual property value

Understand the qualities that affect home price

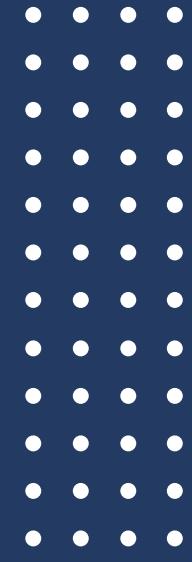
Home Square Footage
Home Quality
Neighborhood
Home Type (Concept)

Ensuring investment value retention

Understanding qualities that your investments have

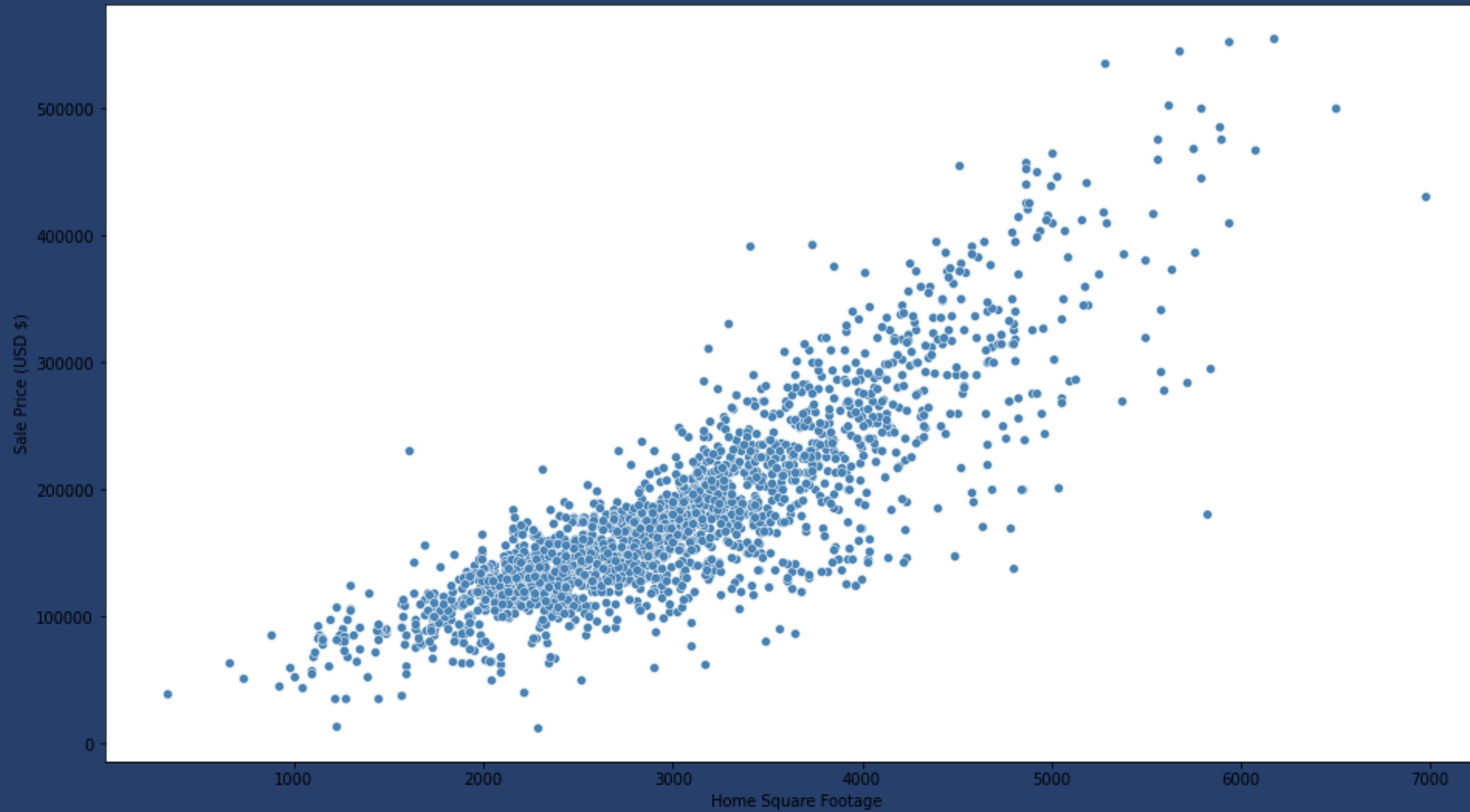


QUESTIONS

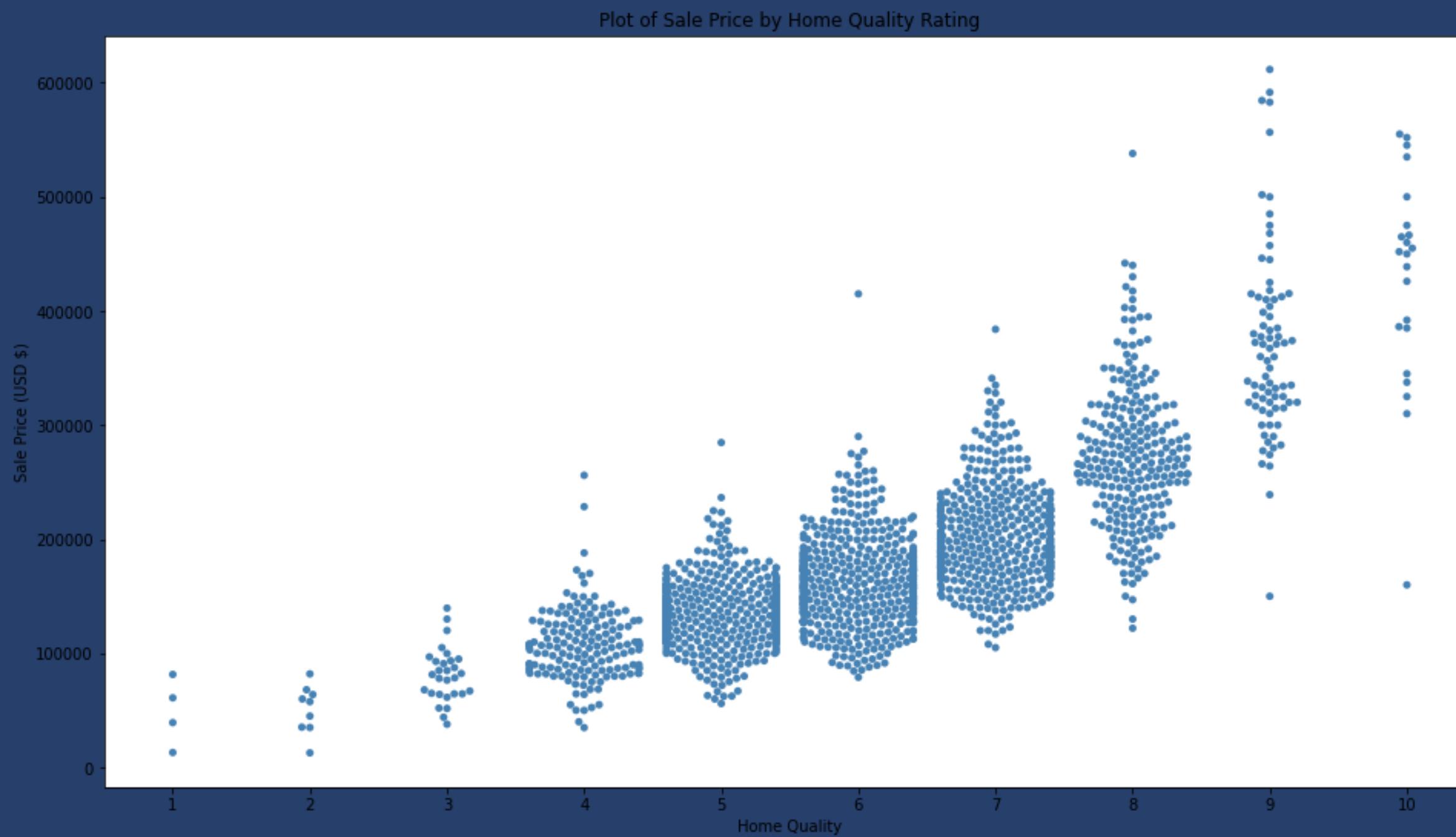


APPENDIX

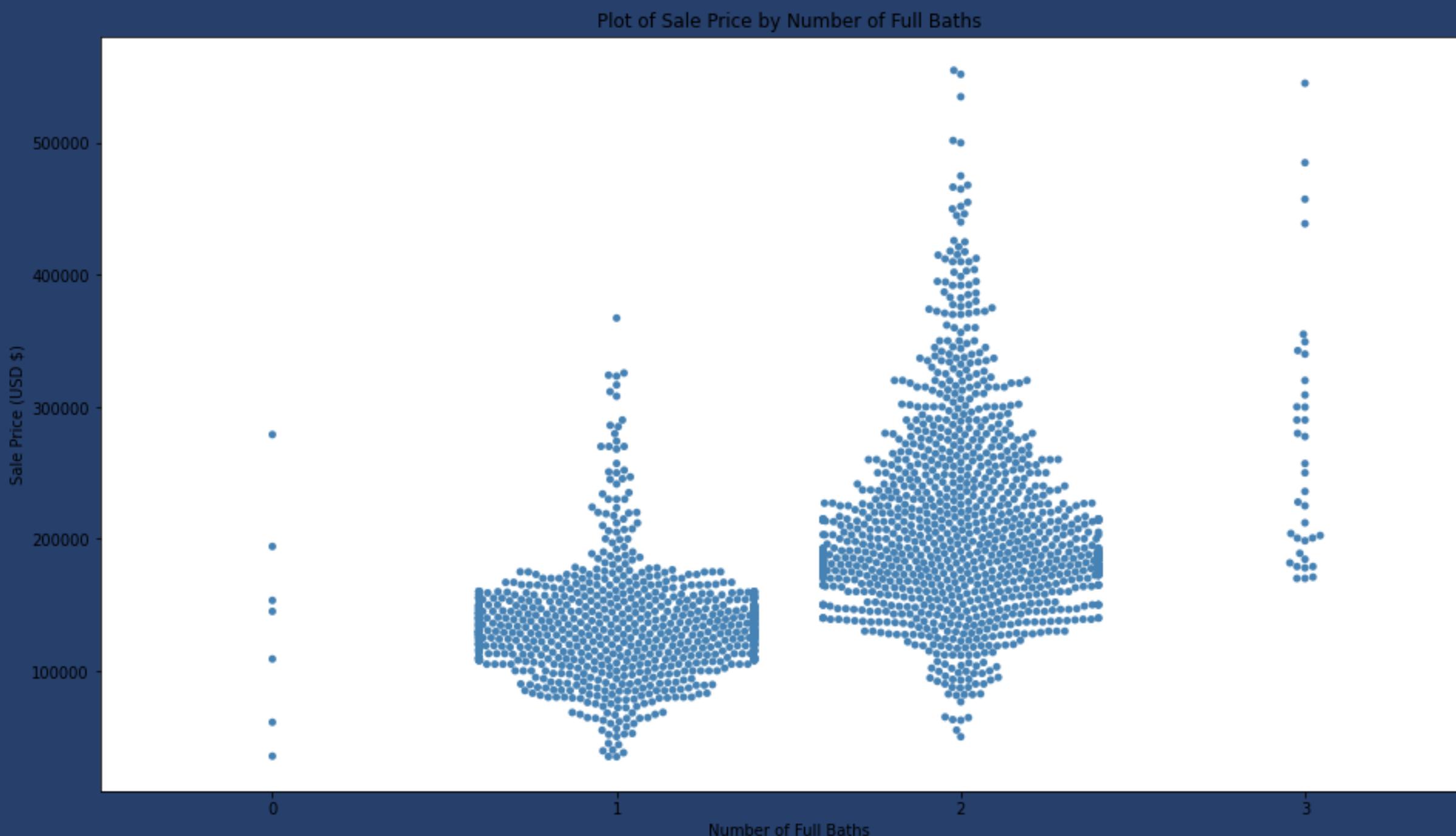
Plot of Sale Price by Home Square Footage



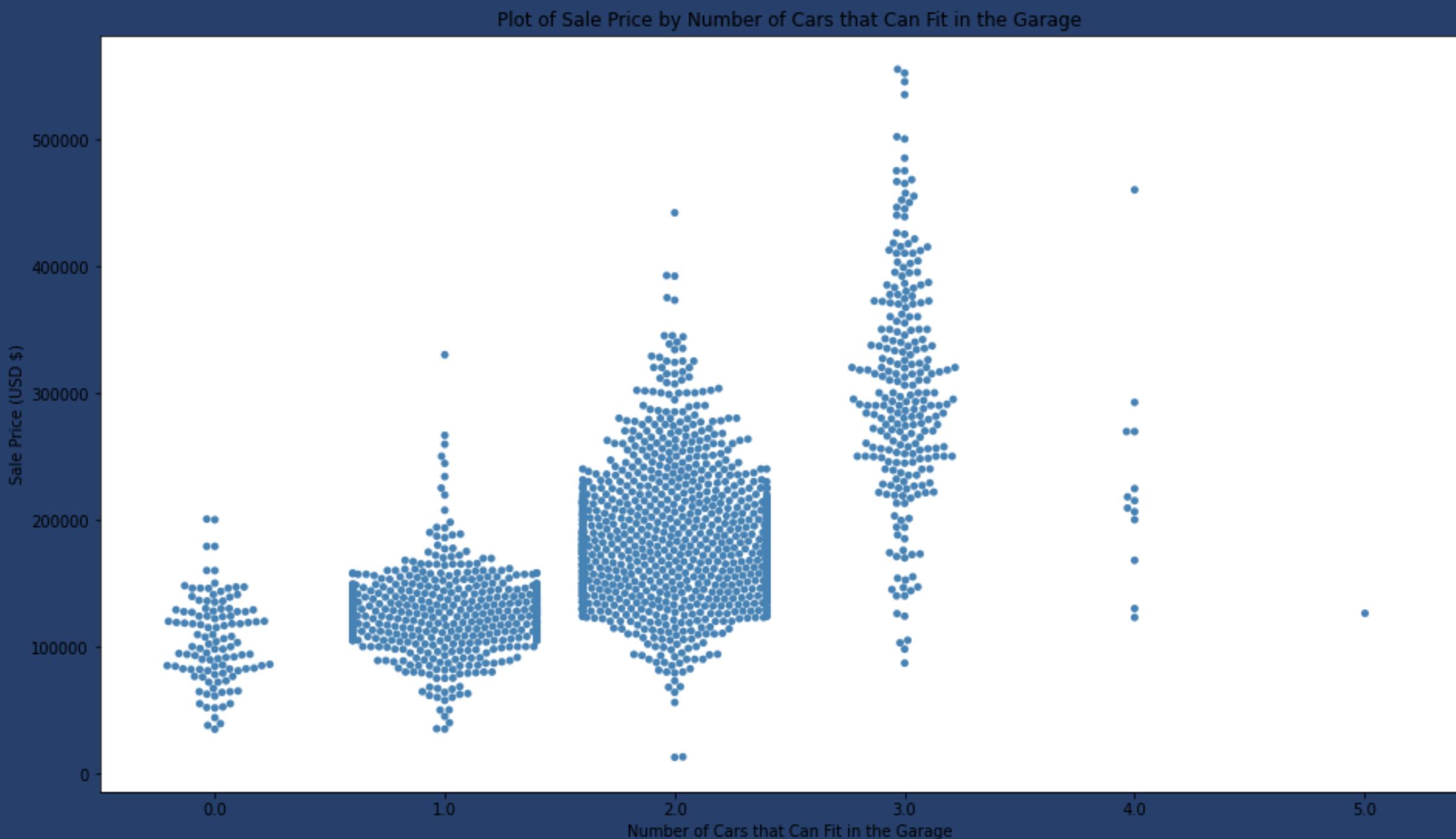
APPENDIX



APPENDIX

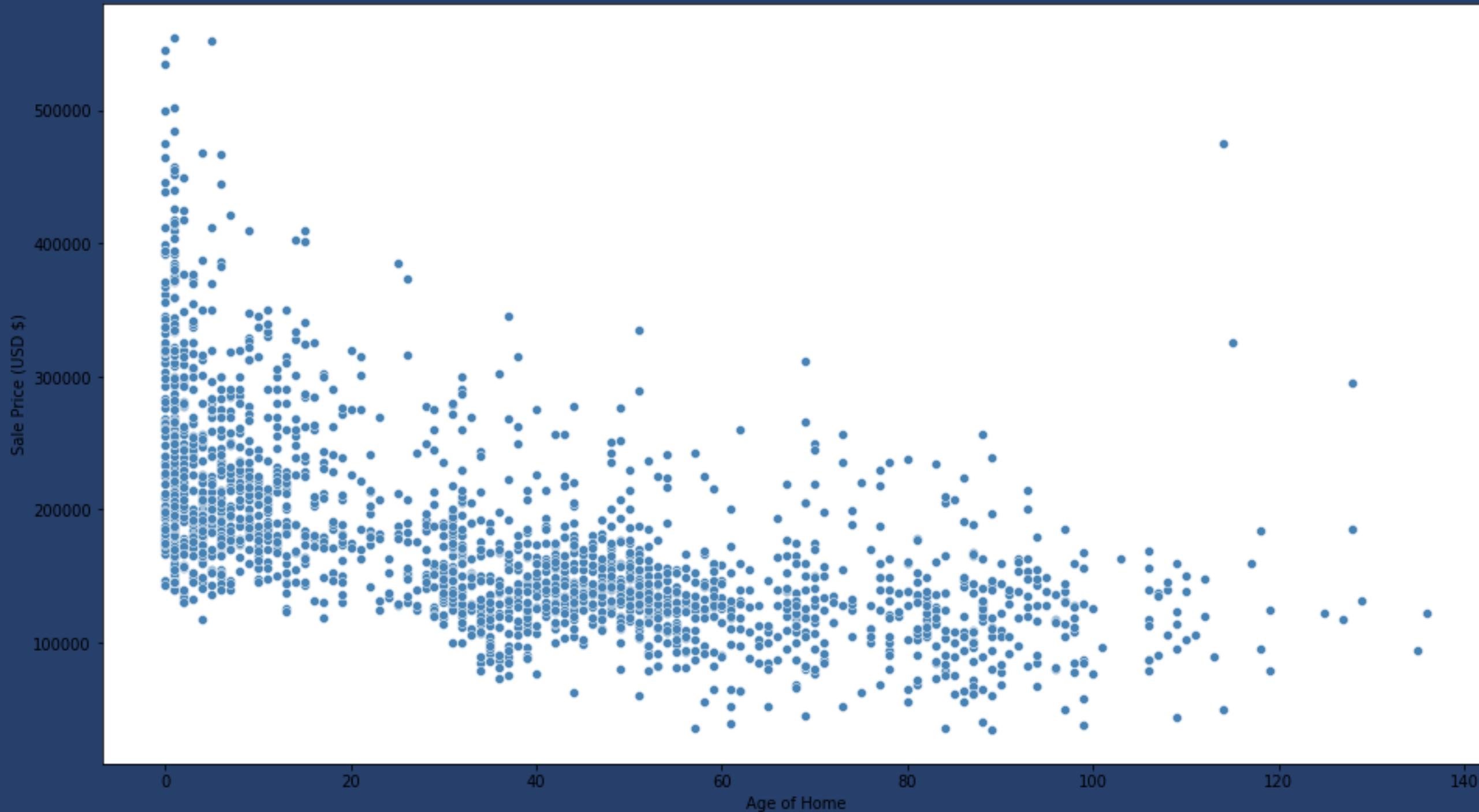


APPENDIX

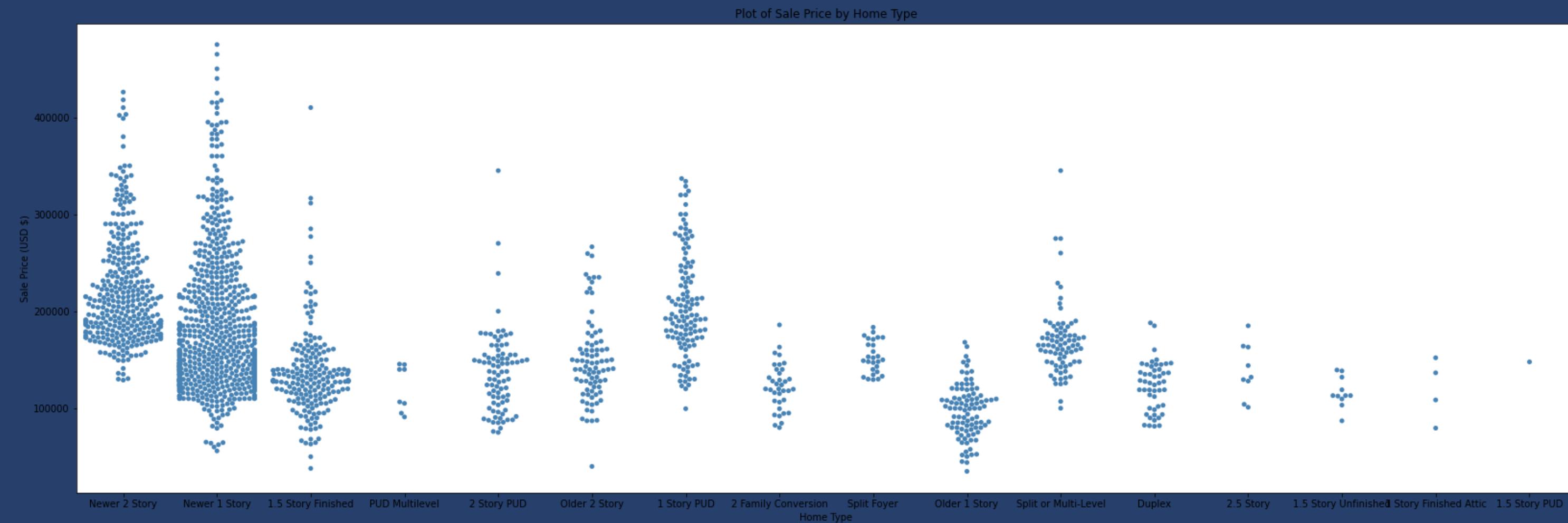


APPENDIX

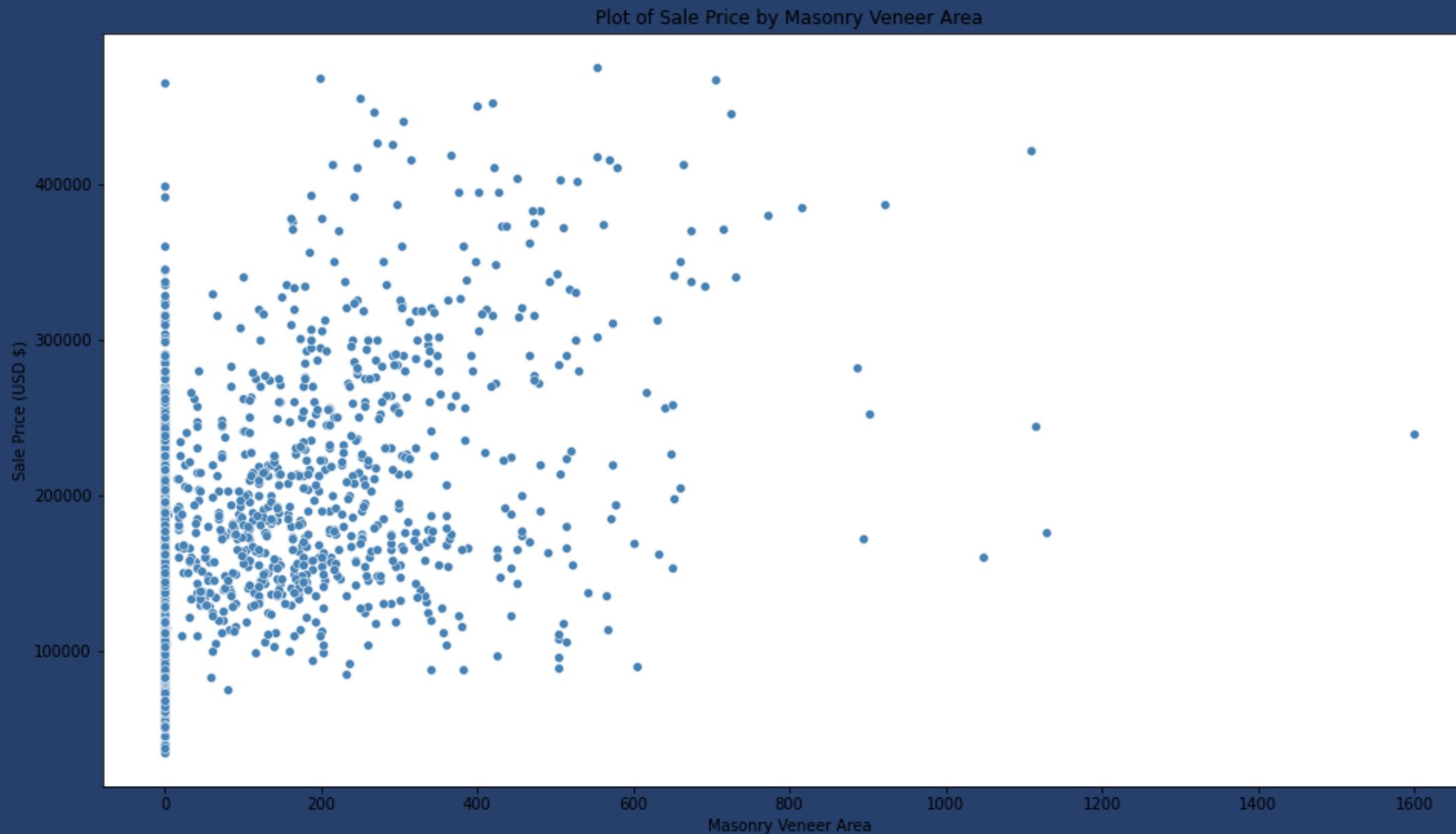
Plot of Sale Price by Age of Home



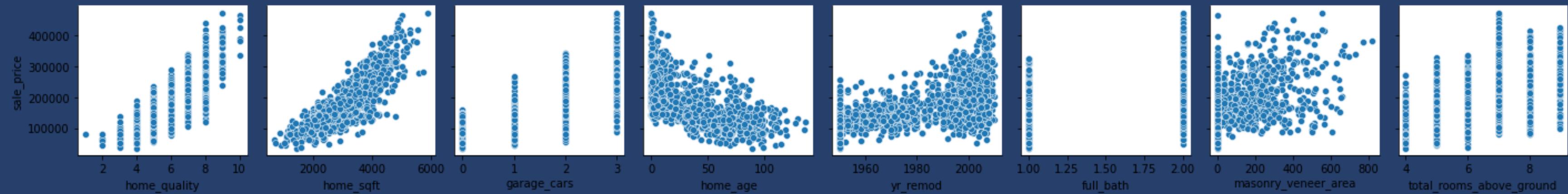
APPENDIX



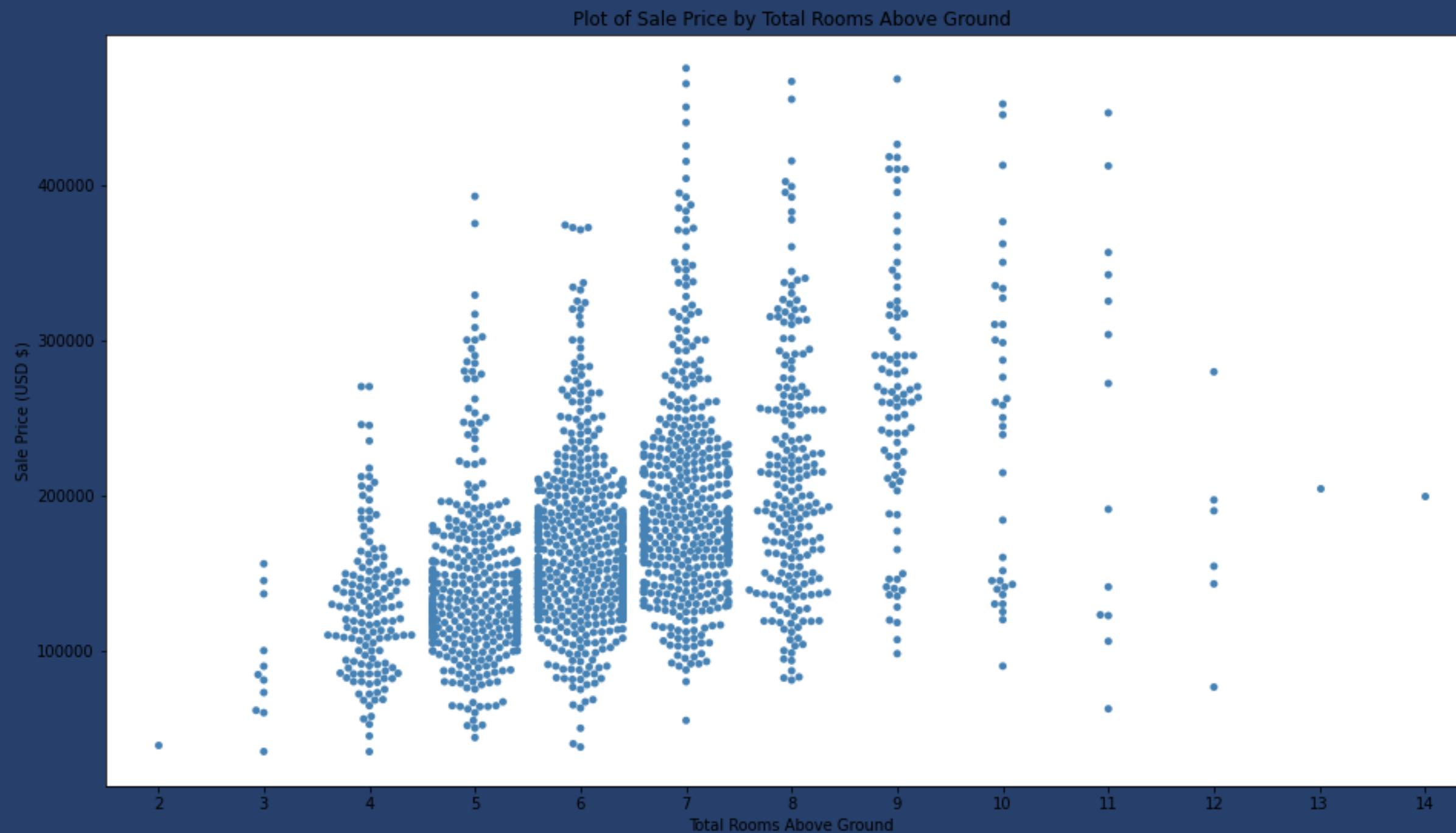
APPENDIX



APPENDIX

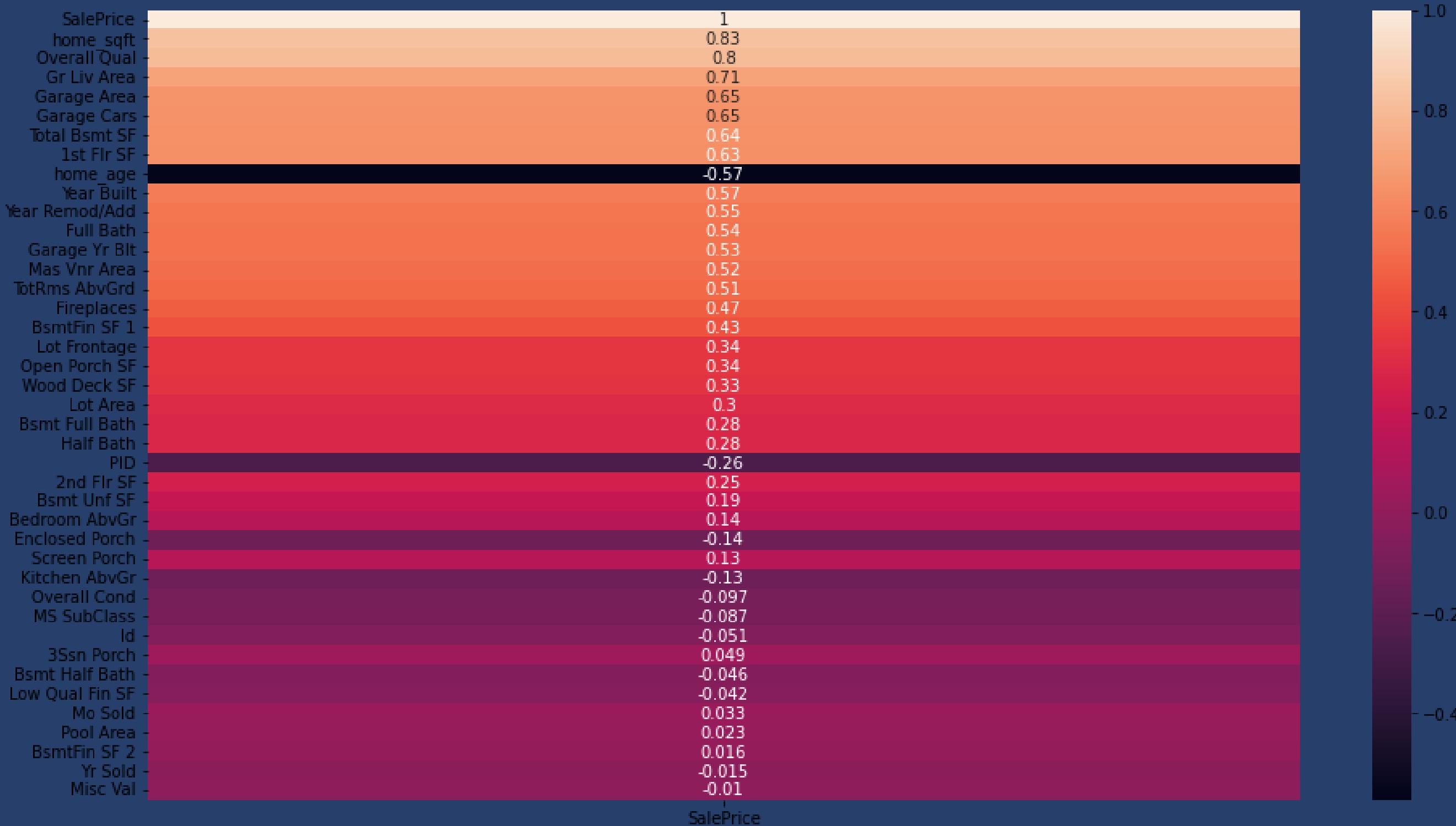


APPENDIX



APPENDIX

Correlation Coefficients of Numeric Values



APPENDIX

Top 10 Strongest Coefficients as Identified by the Model (log)

	Features	Coefficient
1	ss_home_sqft	0.173751
13	ohe_neighborhood_Crawfor	0.163102
11	ohe_neighborhood_ClearCr	0.132807
33	ohe_neighborhood_Veenker	0.120682
43	ohe_home_type_Split Foyer	0.114954
37	ohe_home_type_2 Story PUD	-0.114078
17	ohe_neighborhood_IDOTRR	-0.107370
0	ss_home_quality	0.102906
9	ohe_neighborhood_BrDale	-0.084809
38	ohe_home_type_Duplex	-0.079169

Coefficient translation examples:

Home square footage

Crawford home vs. Sawyer home

ClearCreek home vs. Sawyer home

Veenker home vs. Sawyer home

A Split Foyer home vs. a 2 story home built in 1946 or newer

A 2 story PUD vs. a 2 story home built in 1946 or newer