Melbourne Housing Market

By: Furqan Akhtar

Project Brief

Melbourne’s housing market has finally cooled off.

Provide a general overview of market.

Where's the expensive and in-expensive side of town?

Which area to look for to buy a 2-bedroom unit within city centre with a budget of 0.8$ million?

Additionally, could you find a trend or give a prediction?

Contents

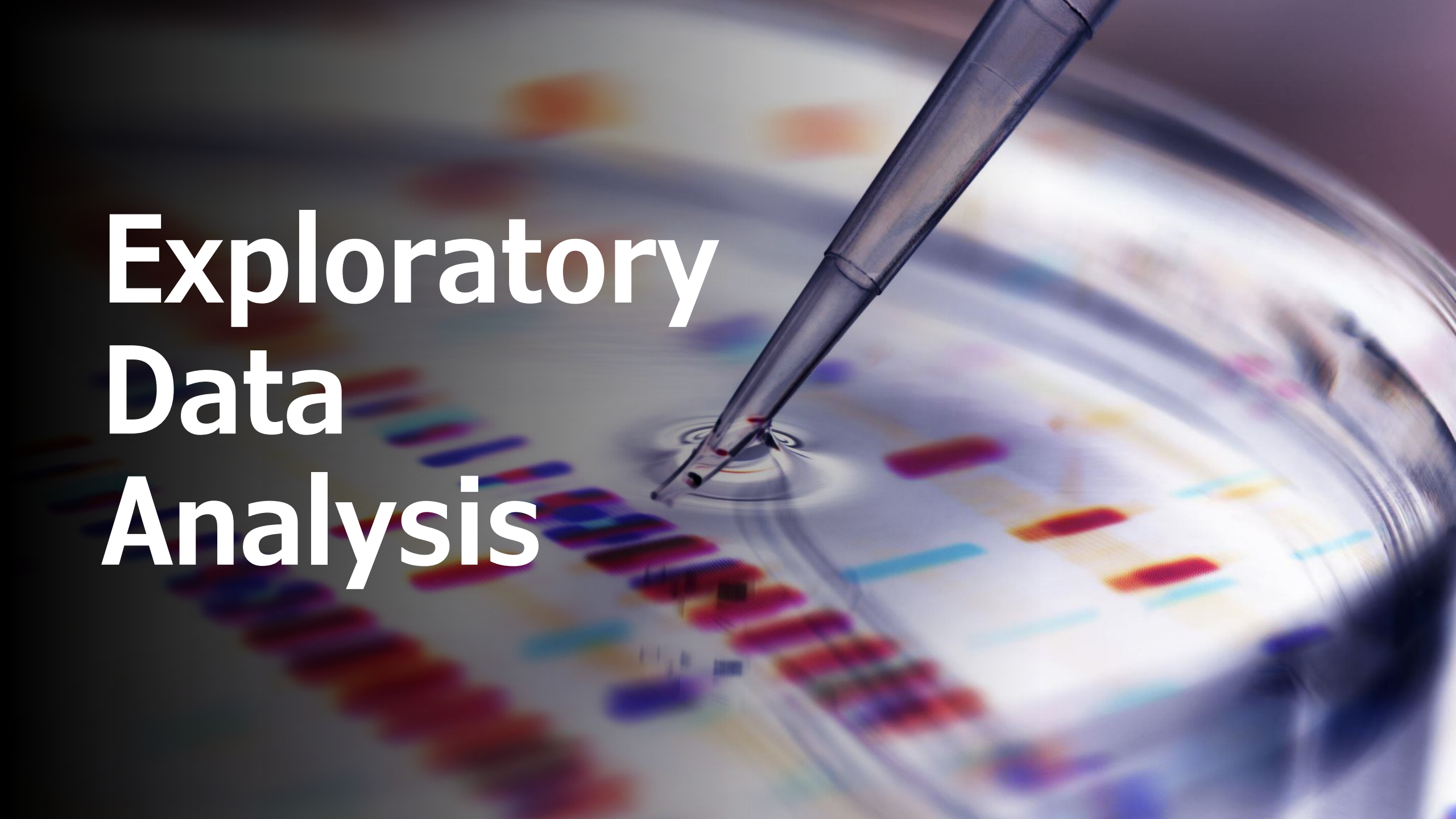
Exploratory Data Analysis

Data Cleaning

Market Analysis

Machine Learning Analysis

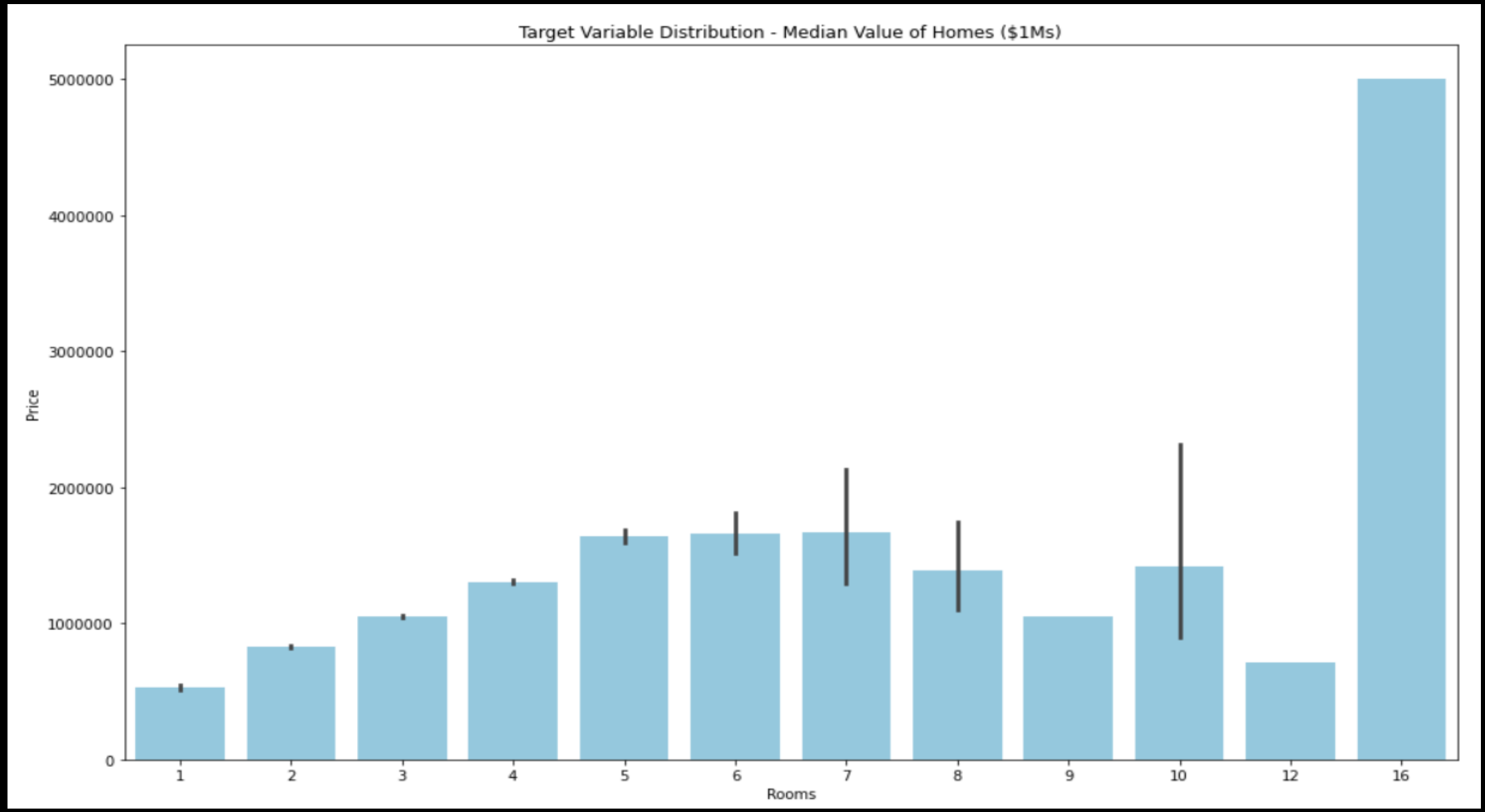
Conclusion


Exploratory
Data
Analysis

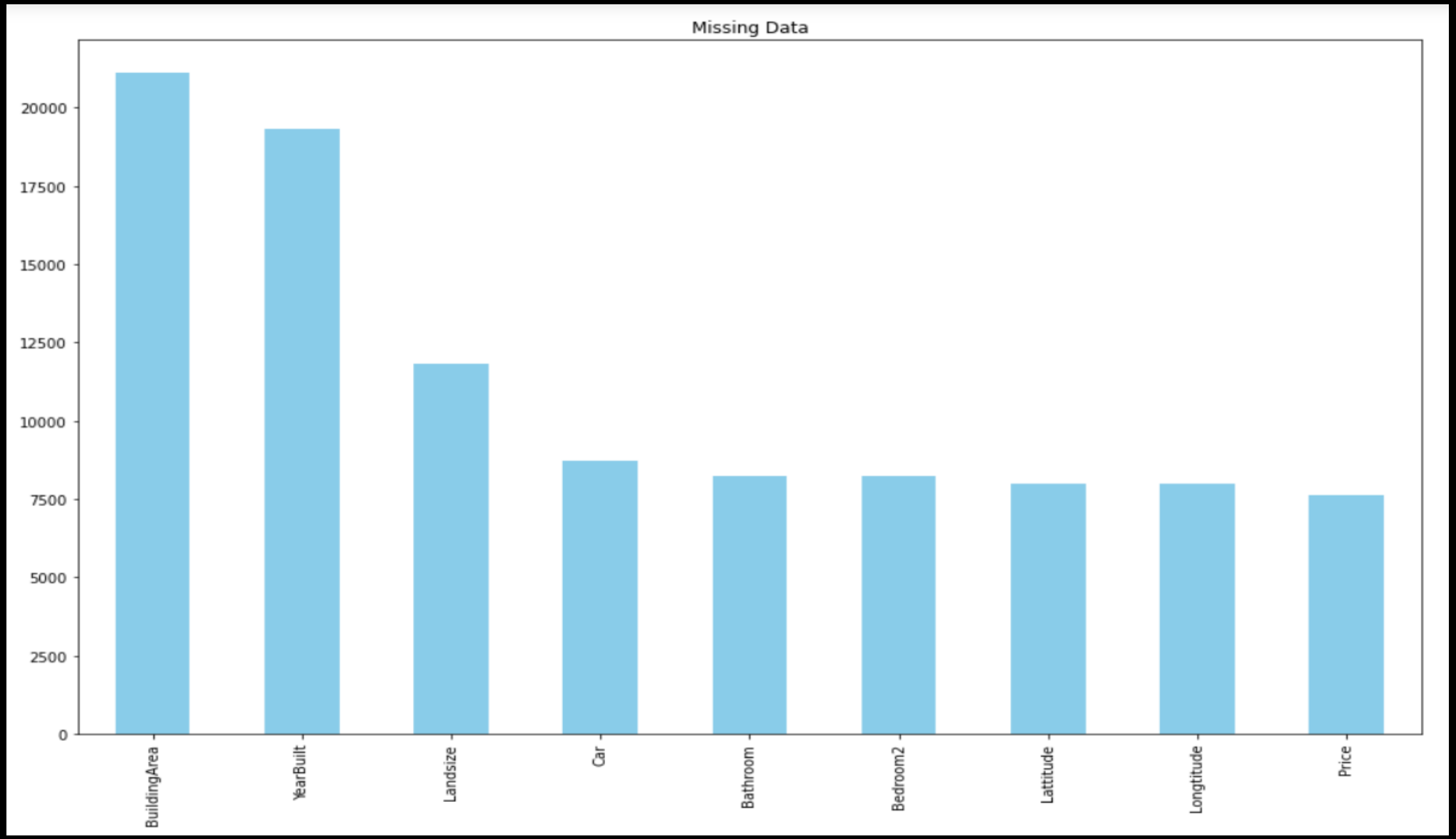
Data Dictionary

	Column Name	Column Type	Brief Description
0	Suburb	string	Suburb
1	Address	string	Address
2	Rooms	int64	Number of rooms
3	Type	string	h = house,cottage,villa, semi,terrace, u = unit, duplex, t = townhouse, dev site = development site, o res = other residential
4	Price	float64	Price in Australian dollars
5	Method	string	S = property sold, SP = property sold prior, PI = property passed in, PN = sold prior not disclosed, SN = sold not disclosed, NB = no bid, VB = vendor bid, W = withdrawn prior to auction, SA = sold after auction, SS = sold after auction price not disclosed, N/A = price or highest bid not available
6	SellerG	string	Real estate agents name
7	Date	datetime64	Date sold
8	Distance	float64	Distance from Cetral Business District (CBD) in Kilometres
9	Postcode	float64	Post code
10	Bedroom2	float64	Number of bedrooms
11	Bathroom	float64	Number of bathrooms
12	Car	float64	Number of parking slots
13	Landsize	float64	Land Size in metres
14	BuildingArea	float64	Building size in metres
15	YearBuilt	float64	Year the house was built
16	CouncilArea	string	Governing council for the area
17	Lattitude	float64	Latitude
18	Longtitude	float64	Longitude
19	Regionname	string	General region (West, North West, North, North east …etc)
20	Propertycount	float64	Number of property

Average Price Against Number of Bedrooms



Sum of Null Values



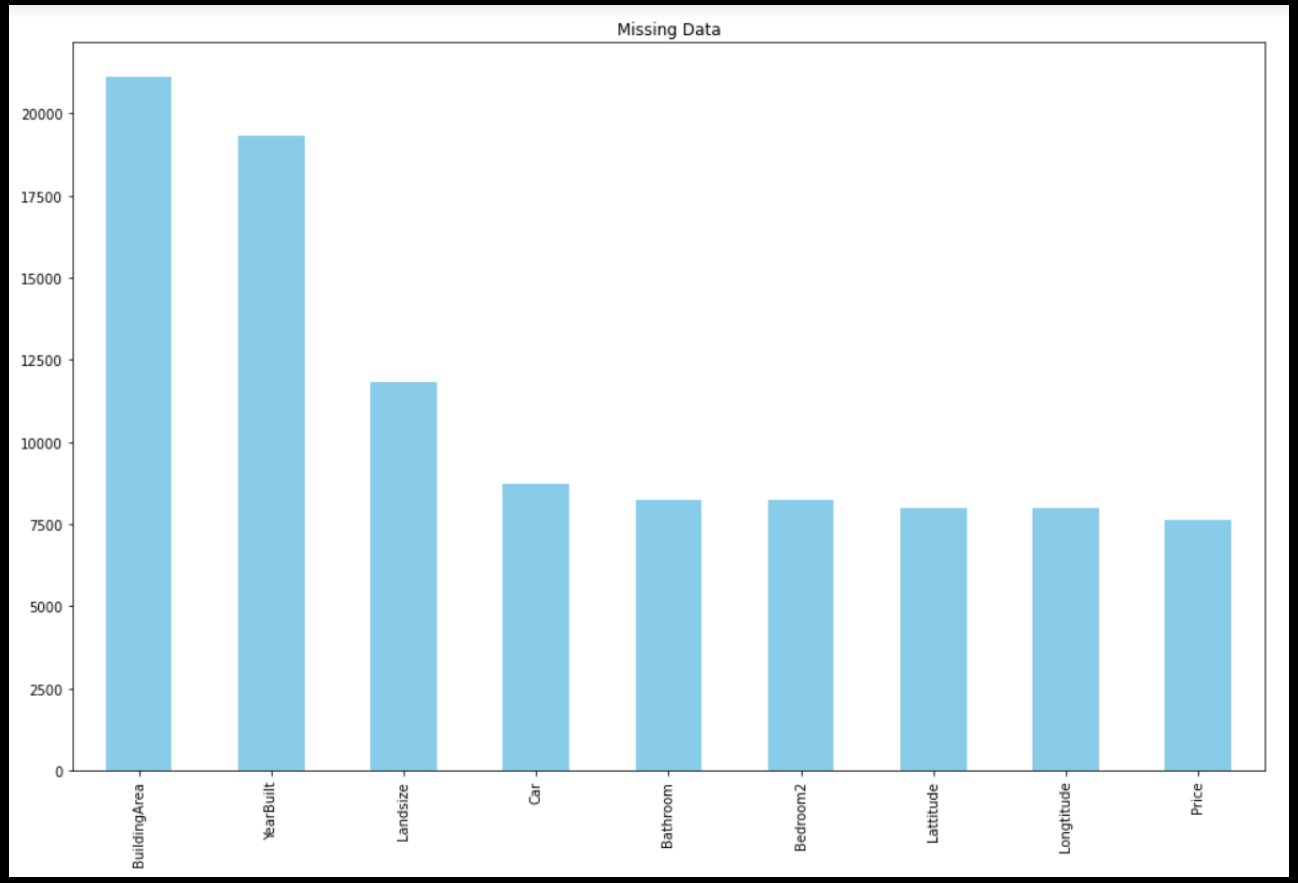
Data
Cleaning

Missing values

Removed columns with more than 10k null values

Removed a duplicate column

Removed a column irrelevant to analysis



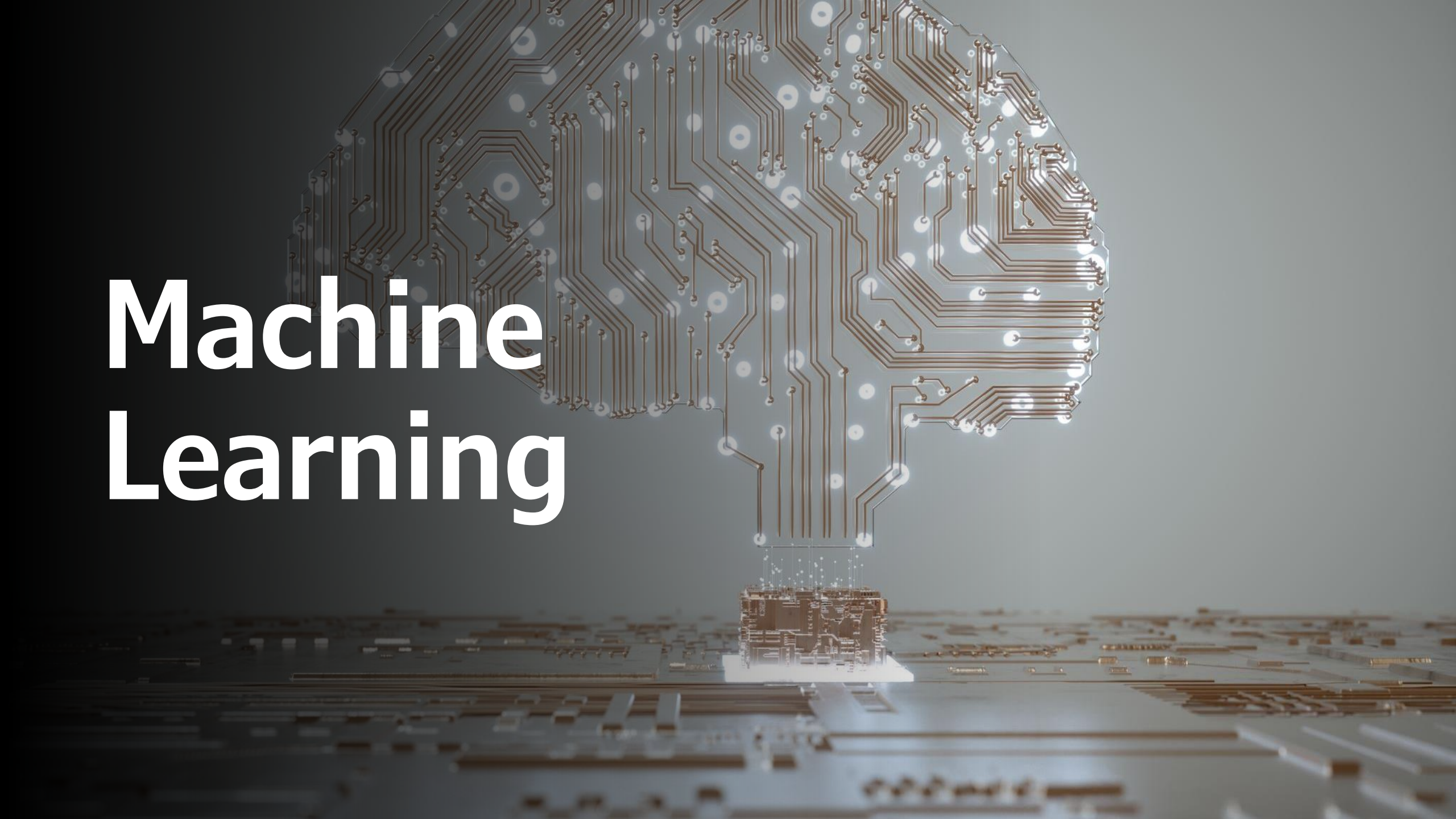
Coded Values

Replaced coded values

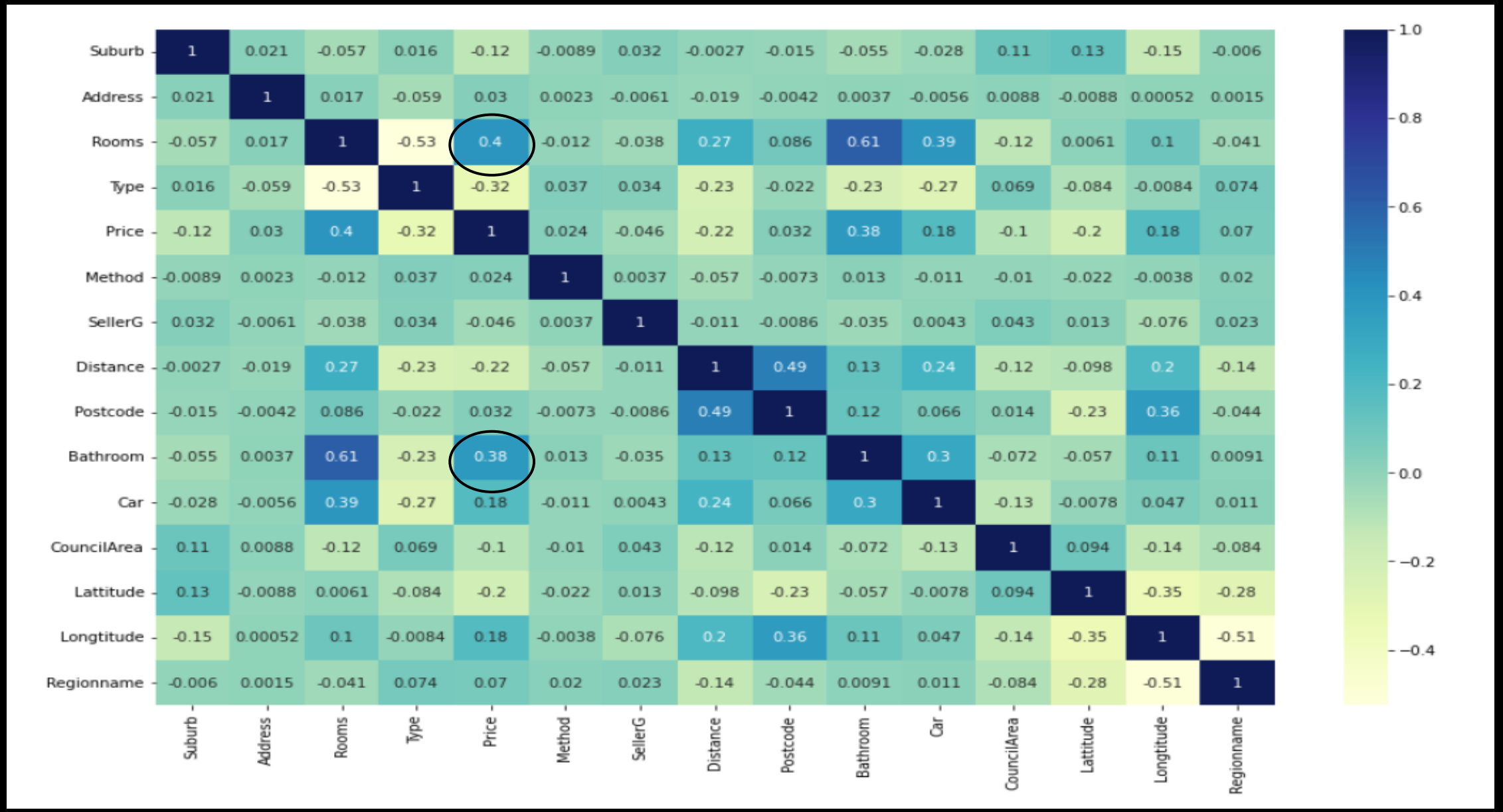
Dropped rows of data where price was not disclosed

Replaced null values with mean where applicable

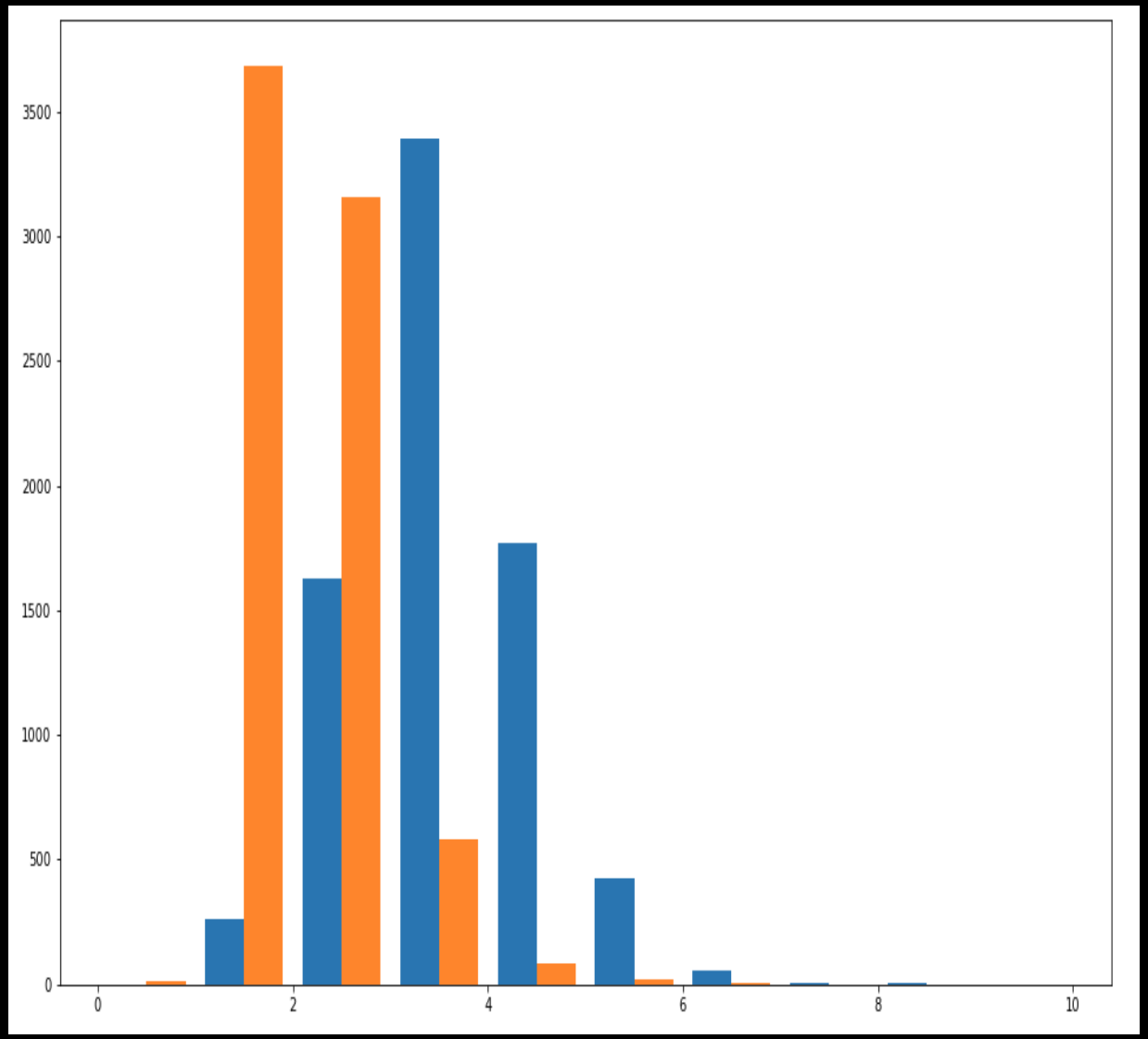
Market Overview

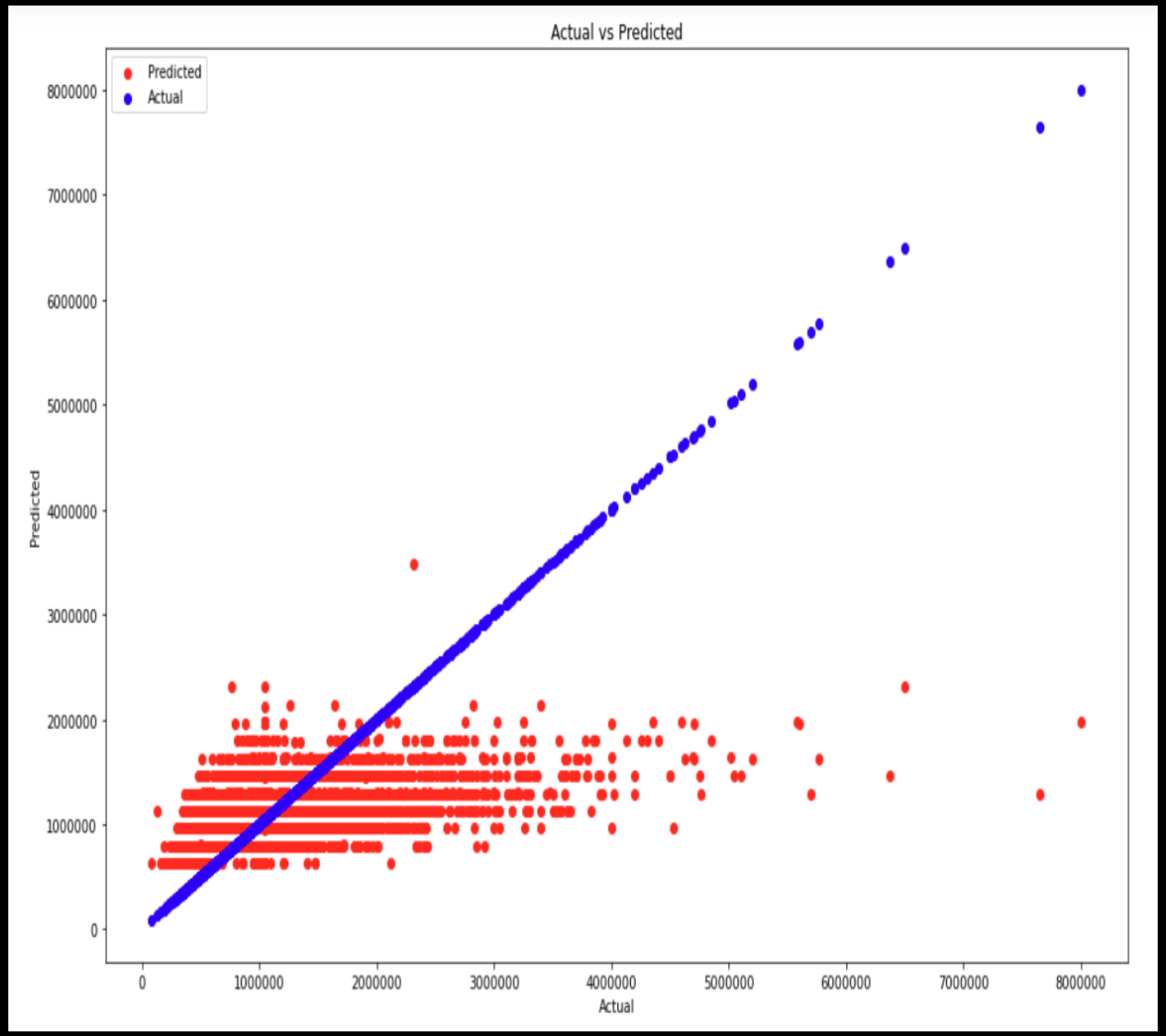

Machine
Learning

Heatmap


Suburb 1 0.021 -0.057 0.016 -0.12 -0.0089 0.032 -0.0027 -0.015 -0.055 -0.028 0.11 0.13 -0.15 -0.006
Address 0.021 1 0.017 -0.059 0.03 0.0023 -0.0061 -0.019 -0.0042 0.0037 -0.0056 0.0088 -0.0088 0.00052 0.0015
Rooms -0.057 0.017 1 -0.53 0.4 -0.012 -0.038 0.27 0.086 0.61 0.39 -0.12 0.0061 0.1 -0.041
Type 0.016 -0.059 -0.53 1 -0.32 0.037 0.034 -0.23 -0.022 -0.23 -0.27 0.069 -0.084 -0.0084 0.074
Price -0.12 0.03 0.4 -0.32 1 0.024 -0.046 -0.22 0.032 0.38 0.18 -0.1 -0.2 0.18 0.07
Method -0.0089 0.0023 -0.012 0.037 0.024 1 0.0037 -0.057 -0.0073 0.013 -0.011 -0.01 -0.022 -0.0038 0.02
SellerG 0.032 -0.0061 -0.038 0.034 -0.046 0.0037 1 -0.011 -0.0086 -0.035 0.0043 0.043 0.013 -0.076 0.023
Distance -0.0027 -0.019 0.27 -0.23 -0.22 -0.057 -0.011 1 0.49 0.13 0.24 -0.12 -0.098 0.2 -0.14
Postcode -0.015 -0.0042 0.086 -0.022 0.032 -0.0073 -0.0086 0.49 1 0.12 0.066 0.014 -0.23 0.36 -0.044
Bathroom -0.055 0.0037 0.61 -0.23 0.38 0.013 -0.035 0.13 0.12 1 0.3 -0.072 -0.057 0.11 0.0091
Car -0.028 -0.0056 0.39 -0.27 0.18 -0.011 0.0043 0.24 0.066 0.3 1 -0.13 -0.0078 0.047 0.011
CouncilArea 0.11 0.0088 -0.12 0.069 -0.1 -0.01 0.043 -0.12 0.014 -0.072 -0.13 1 0.094 -0.14 -0.084
Lattitude 0.13 -0.0088 0.0061 -0.084 -0.2 -0.022 0.013 -0.098 -0.23 -0.057 -0.0078 0.094 1 -0.35 -0.28
Longtitude -0.15 0.00052 0.1 -0.0084 0.18 -0.0038 -0.076 0.2 0.36 0.11 0.047 -0.14 -0.35 1 -0.51
Regionname -0.006 0.0015 -0.041 0.074 0.07 0.02 0.023 -0.14 -0.044 0.0091 0.011 -0.084 -0.28 -0.51 1
Suburb Address Rooms Type Price Method SellerG Distance Postcode Bathroom Car CouncilArea Lattitude Longtitude Regionname
1.0 0.8 0.6 0.4 0.2 0.0 −0.2 −0.4

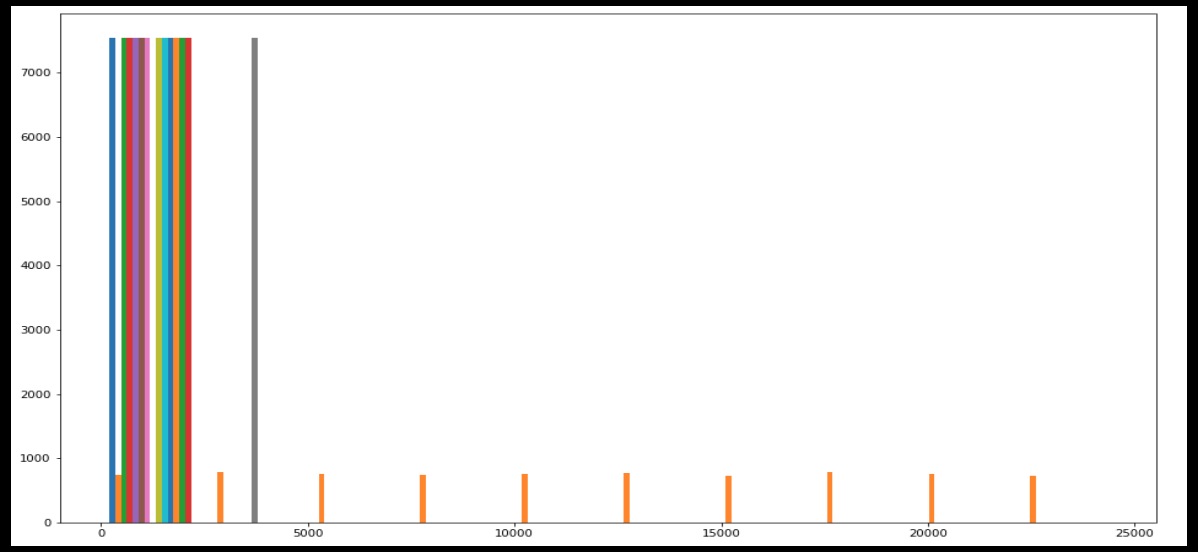
Test 1

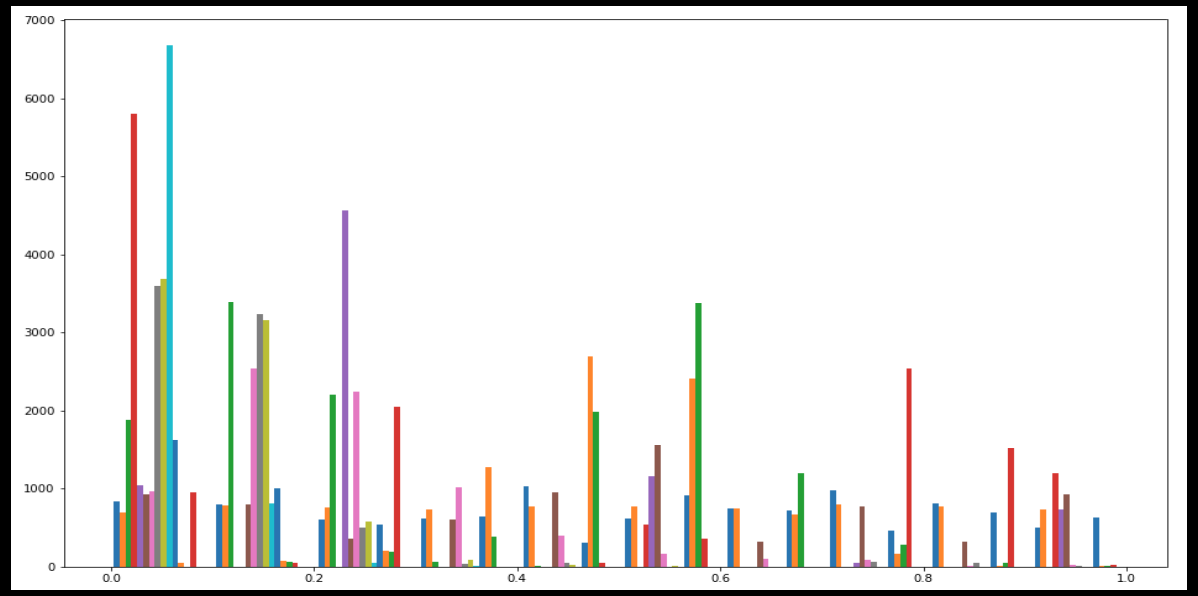


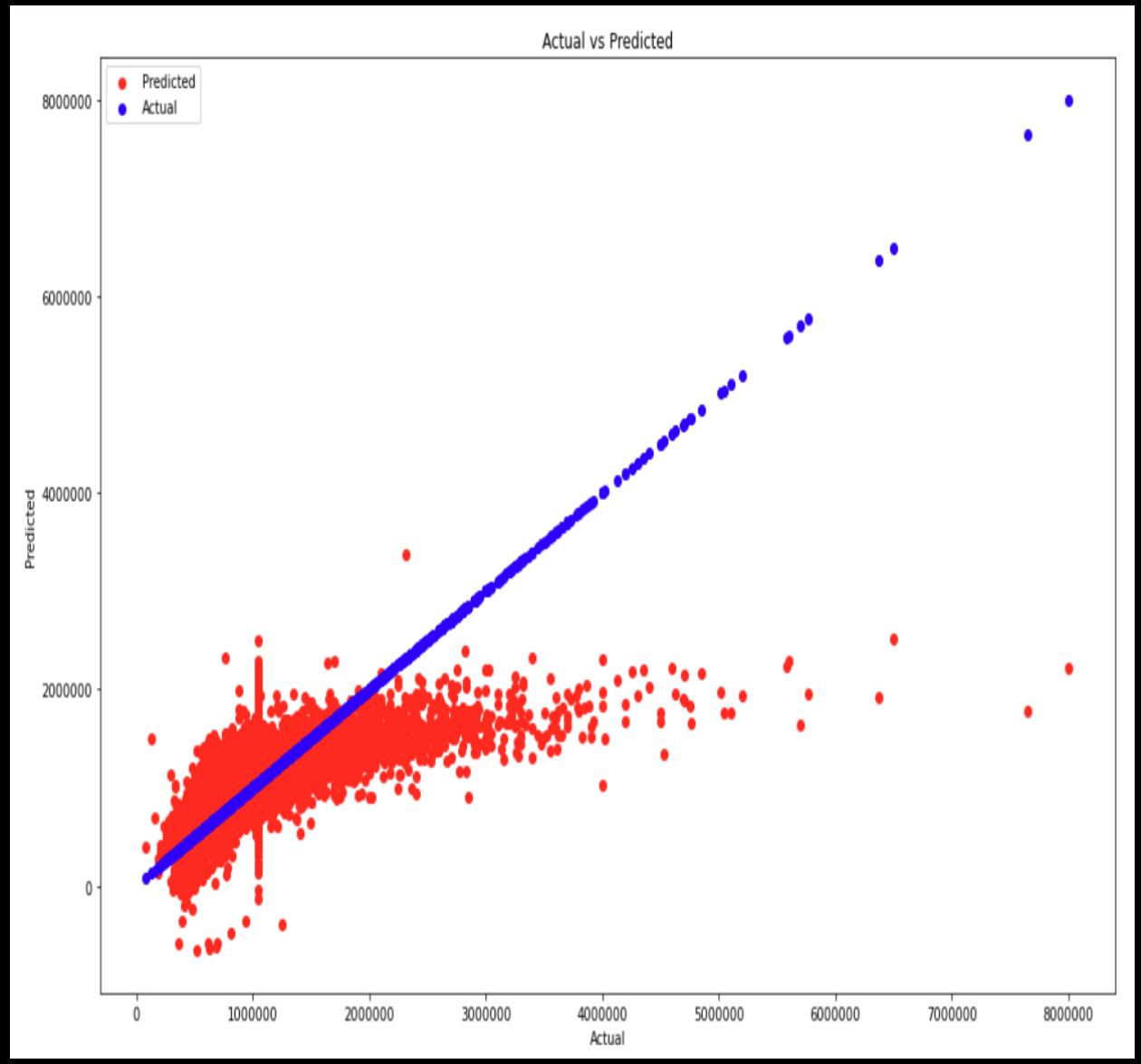


R2 score: 18.98%

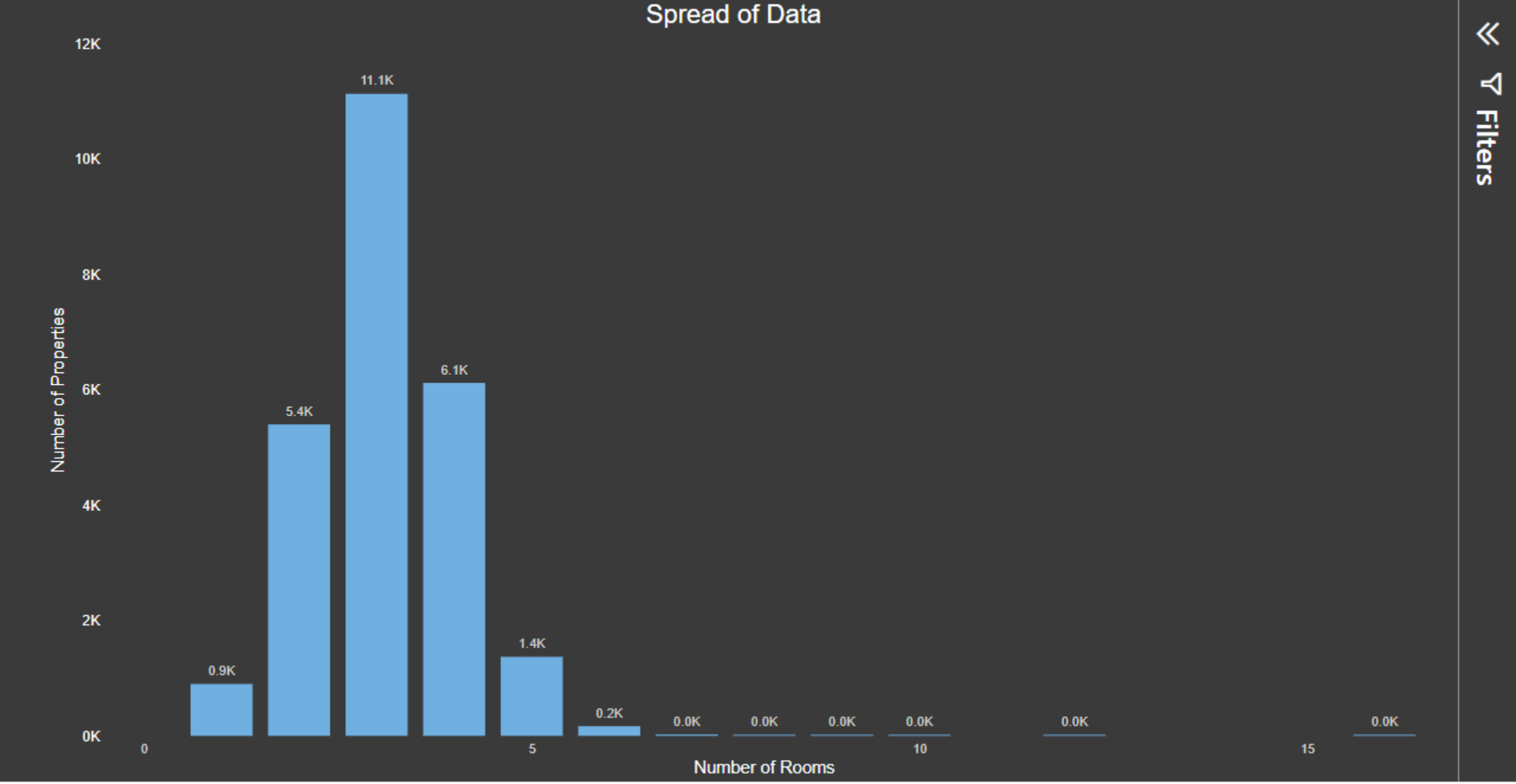
Test 2





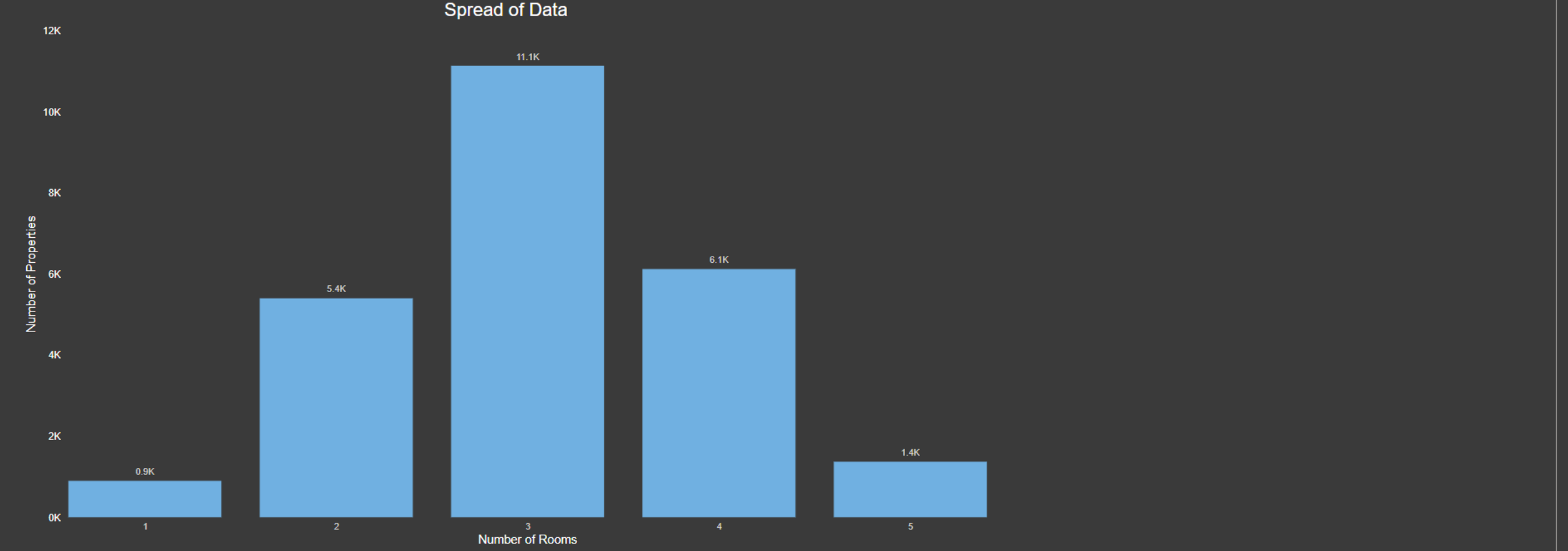


R2 score: 42.86%



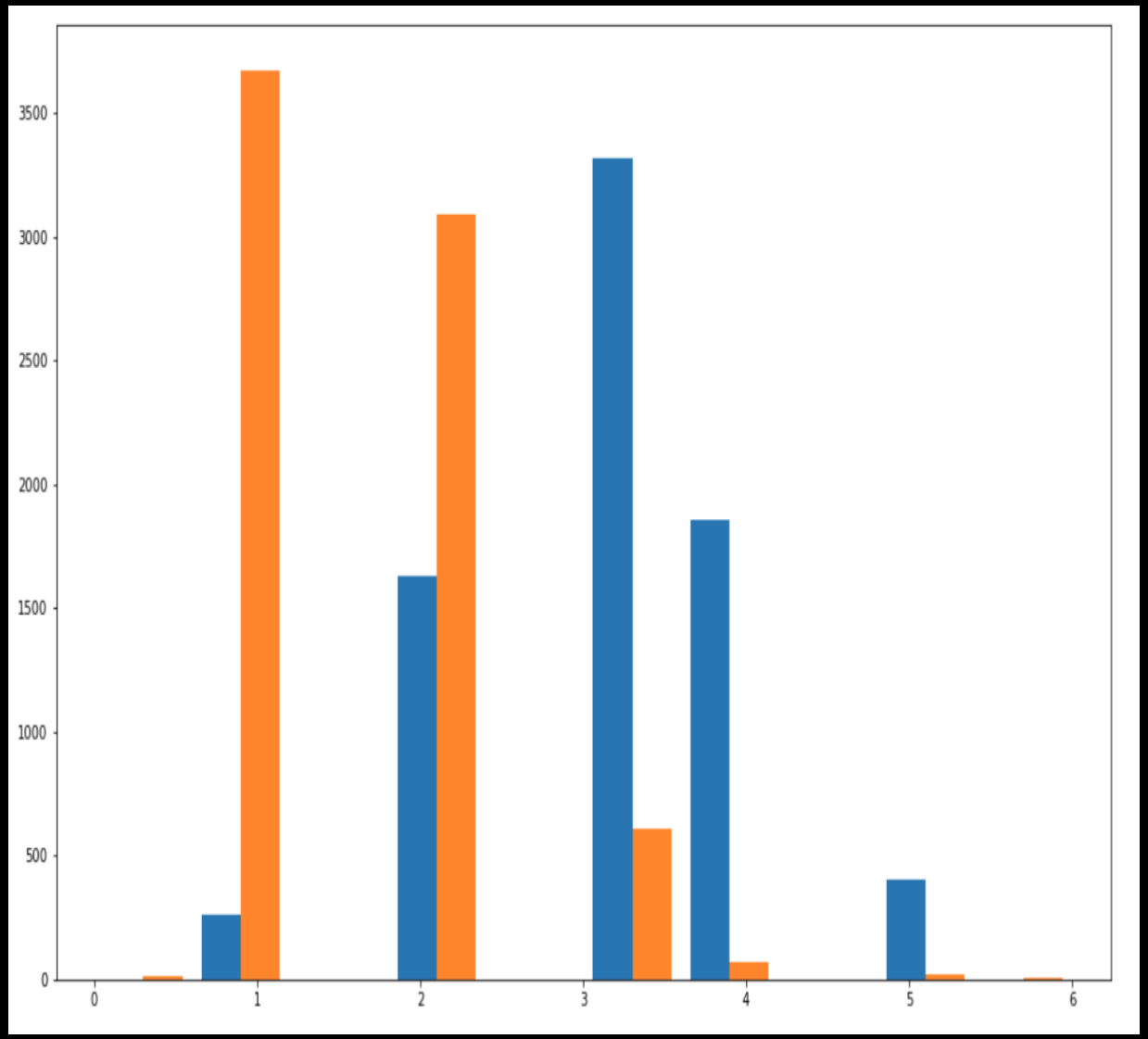
FP, Page 1

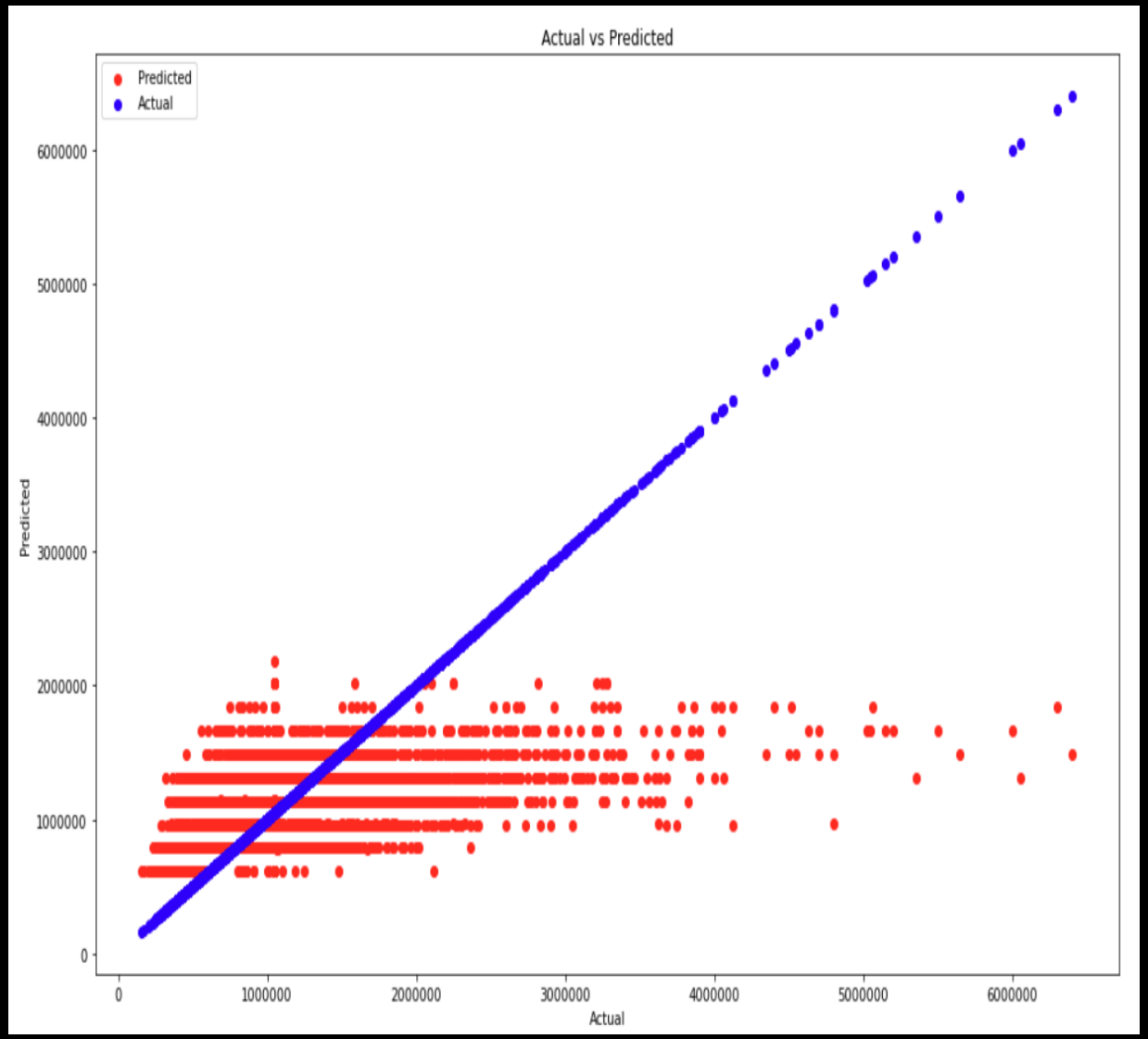
Data updated on 12/2/22, 10:04 AM


Spread of Data
12K
10K
8K
6K
4K
2K
0K
Number of Properties
0.9K
5.4K
11.1K
6.1K
1.4K
1
2
3
4
5
Number of Rooms

Filters

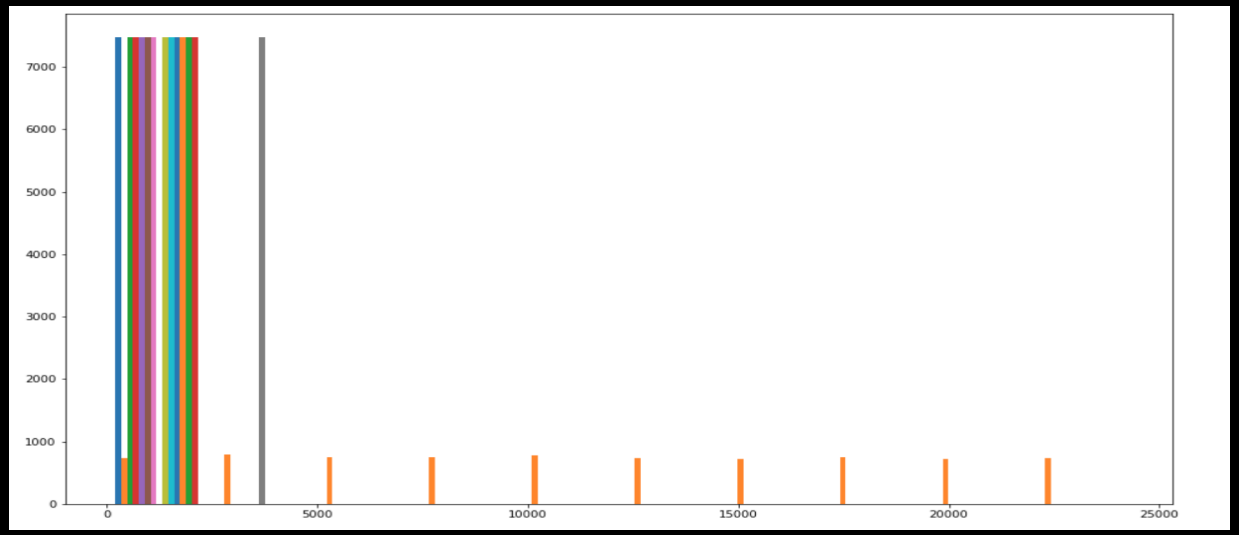
Test 3

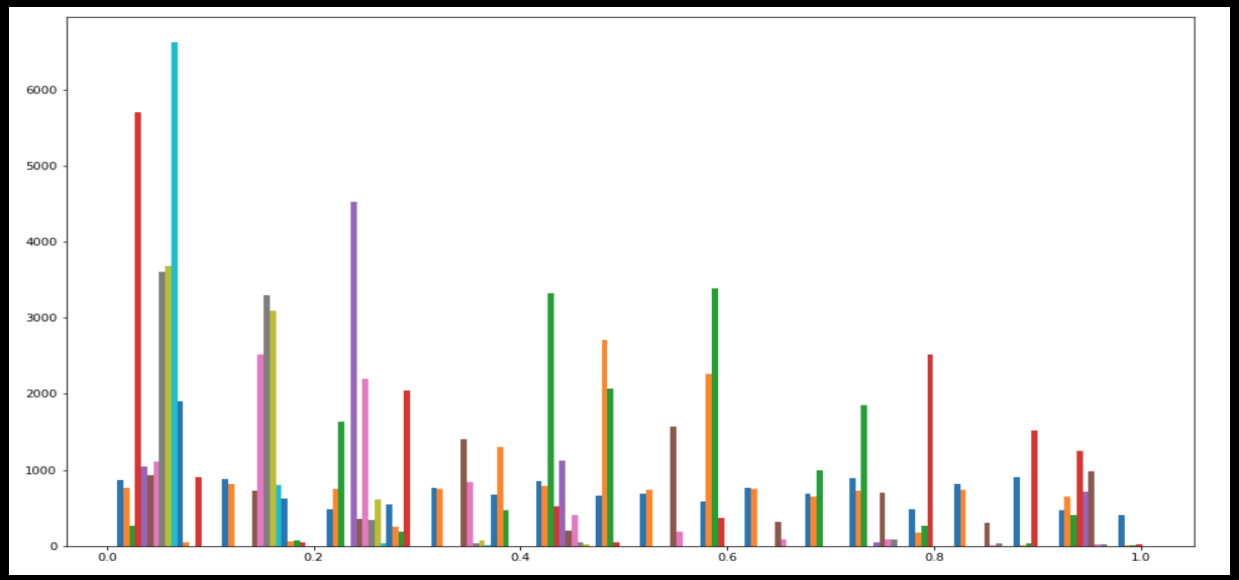


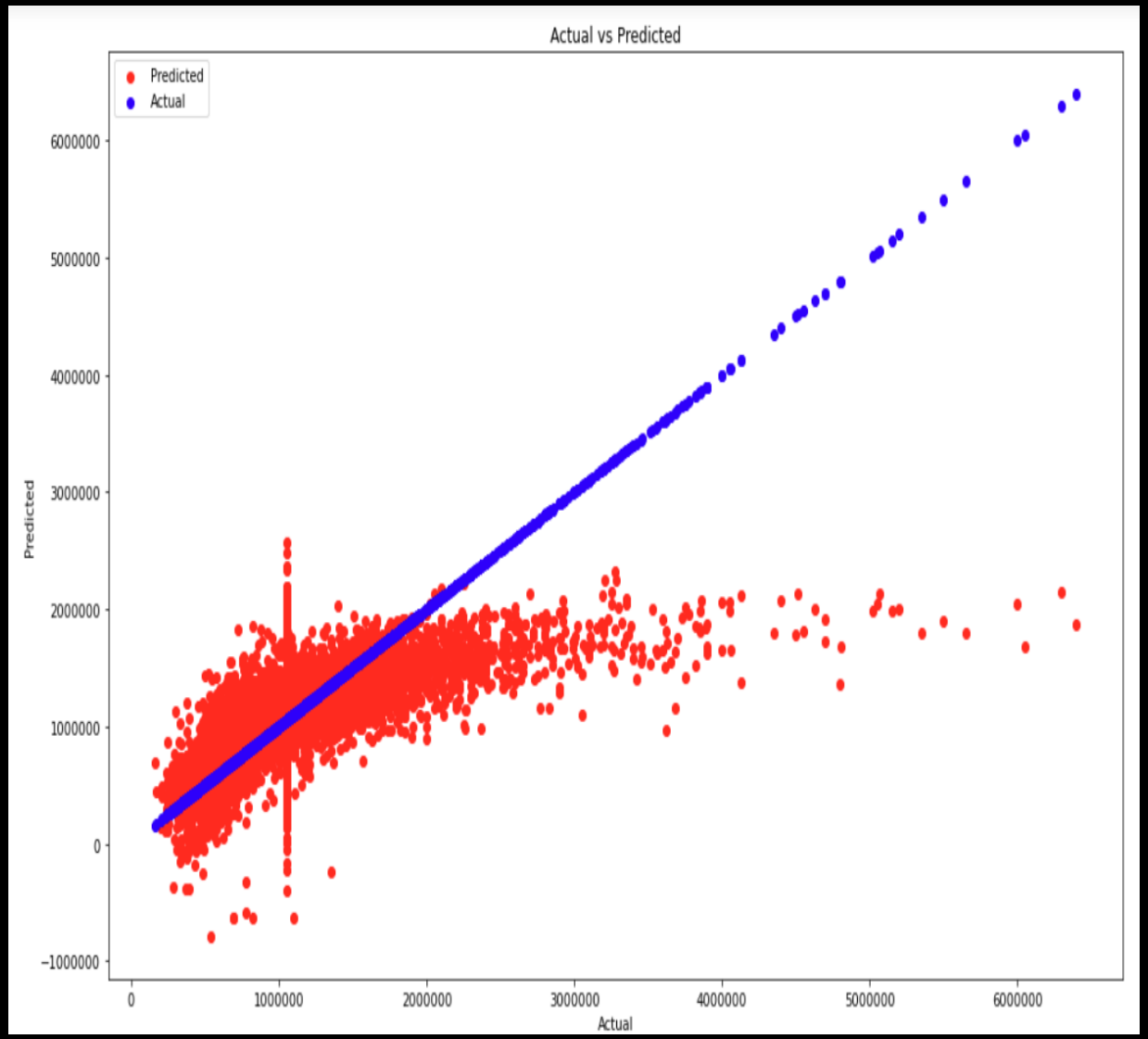


R2 score: 19.71%

Test 4







R2 score: 43.83%

Comparison

Including all bedrooms	
Number of iterations	**R^2 Score**
1st (Only 2 columns)	18.98%
2nd (All columns)	42.86%

Removed houses with more than 5 bedrooms	
Number of iterations	**R^2 Score**
3rd (Only 2 columns)	19.71%
4th (All columns)	43.83%

Conclusion

- Provided an overview of Melbourne’s real estate market
- **Most expensive** area of the city: **Boroondara** City Council
- **Cheapest** area of the city: **Moorabool** Shire Council

Where should they buy a 2-bedroom unit?

- **Melbourne** City Council
- **Moreland** City Council

Further work

- ARIMA Model
- Lasso & Ridge regression
- Analyse with up to date data

Thank you for Listening

Any Questions