

FORECASTING RENT

Presentation by Group 6
(BUAN 6356.004)





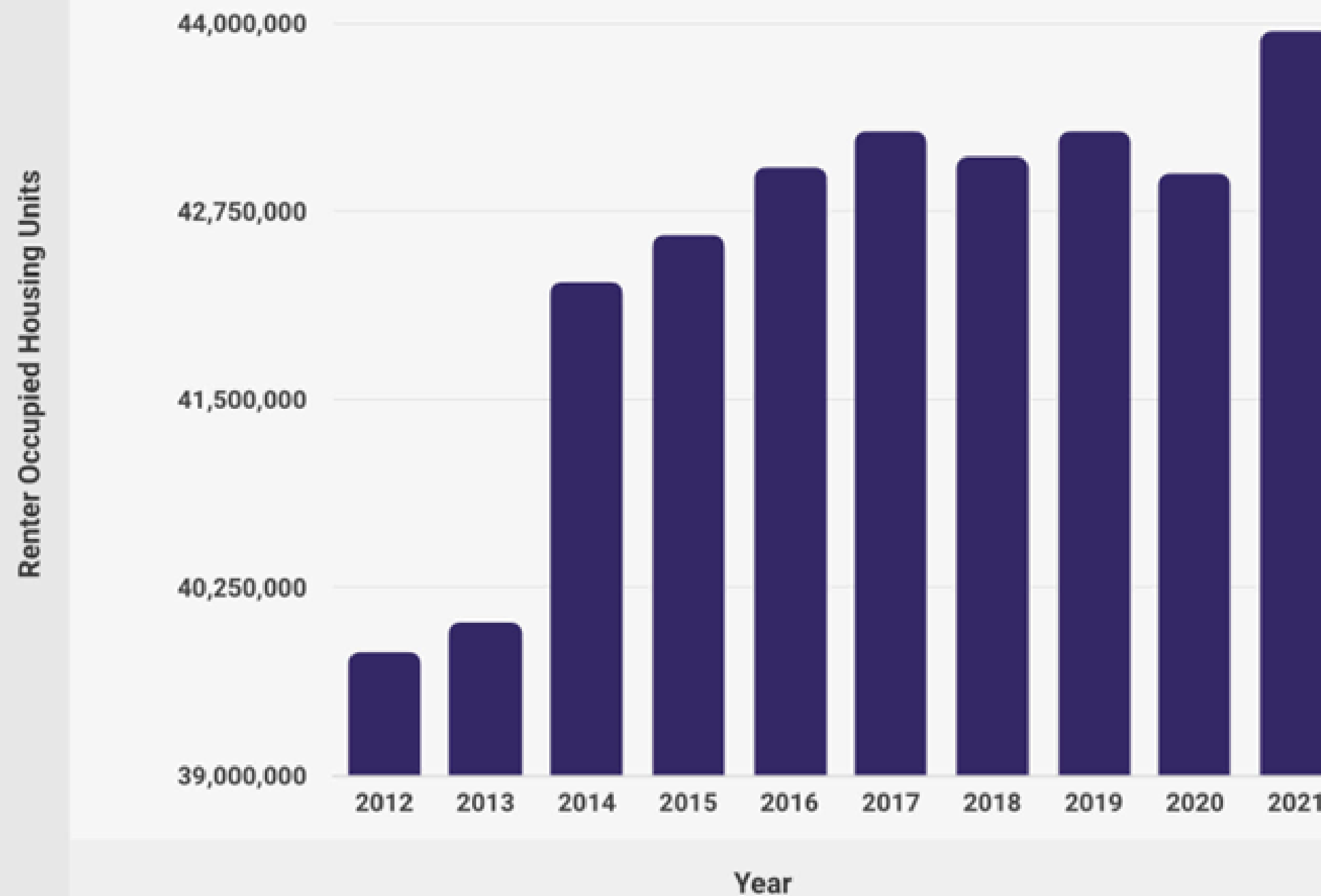
OBJECTIVE

The aim of this project is to use BI techniques like clustering, regression and neural networks to analyze rental apartment data to predict price optimisation for market segmentation to help formulate more effective business strategies.




OVERVIEW

Number of Renter Occupied Units





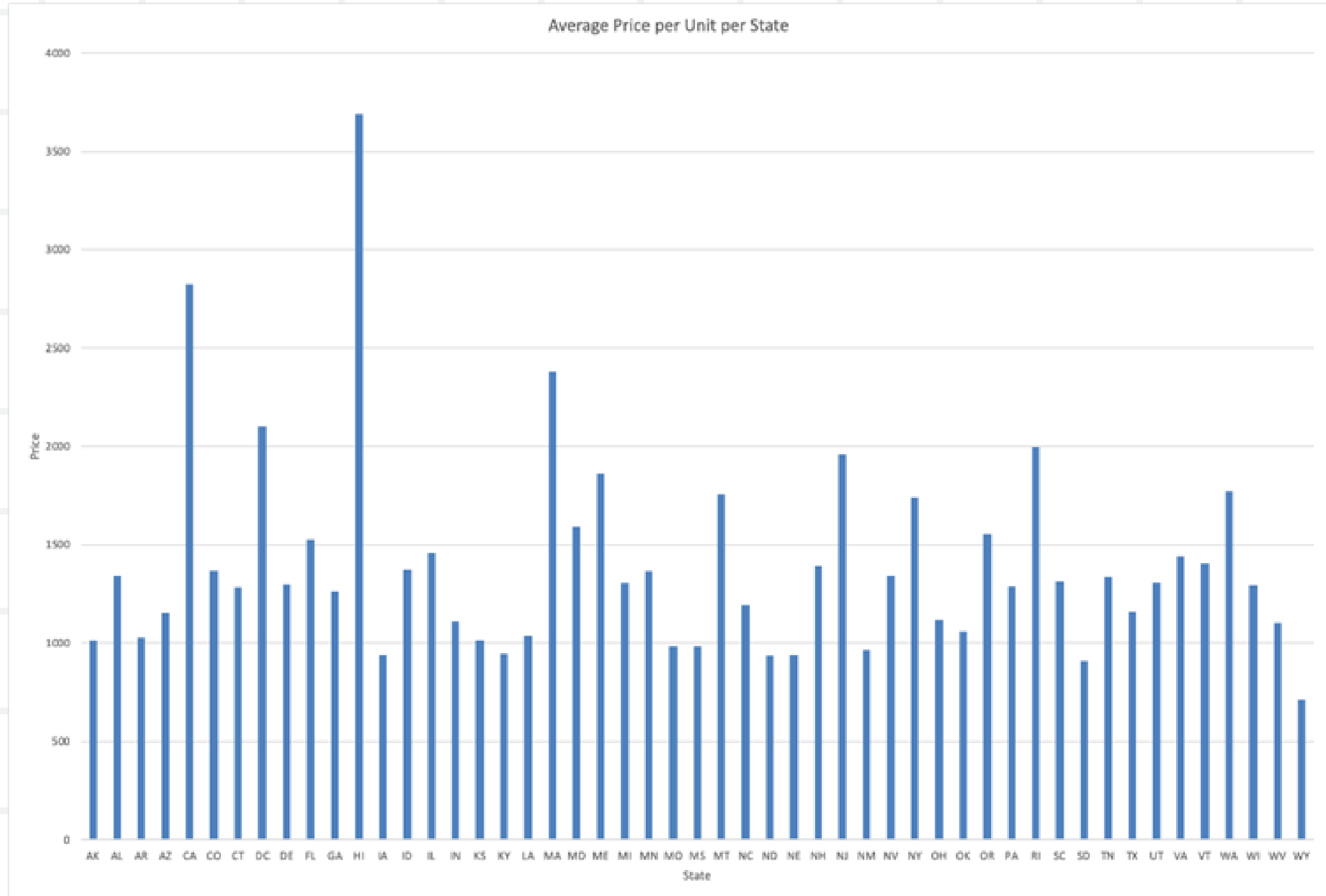
INSIGHTS

- Clustering of data of apartments based on price and square feet as well as on state the apartment is located in and source of the apartment listing.
 - Visualization based on various attribute combinations to better understand the data and gain deeper insights.
 - Regression based on square_feet, bathrooms, bedrooms, state
 - Neural Networks based on square_feet, bathrooms, bedrooms
- 

ATTRIBUTE INFORMATION

- The dataset consists of 22 attributes: 12 numeric, and 10 categorical
- Title and body contain the details and specifications of the apartment; address, city, state, latitude and longitude represent the location of the apartment
- Amenities contains the various facilities provided in the apartment and has pets shows if and what pets are allowed
- No. of Bathrooms and Bedrooms contains the room specifics of the apartment
- price and price display show the cost of the apartment and fee represents any additional fees to be paid
- square feet represents the total area of the apartment
- source shows the source of the apartment listing (website)

INITIAL ANALYSIS



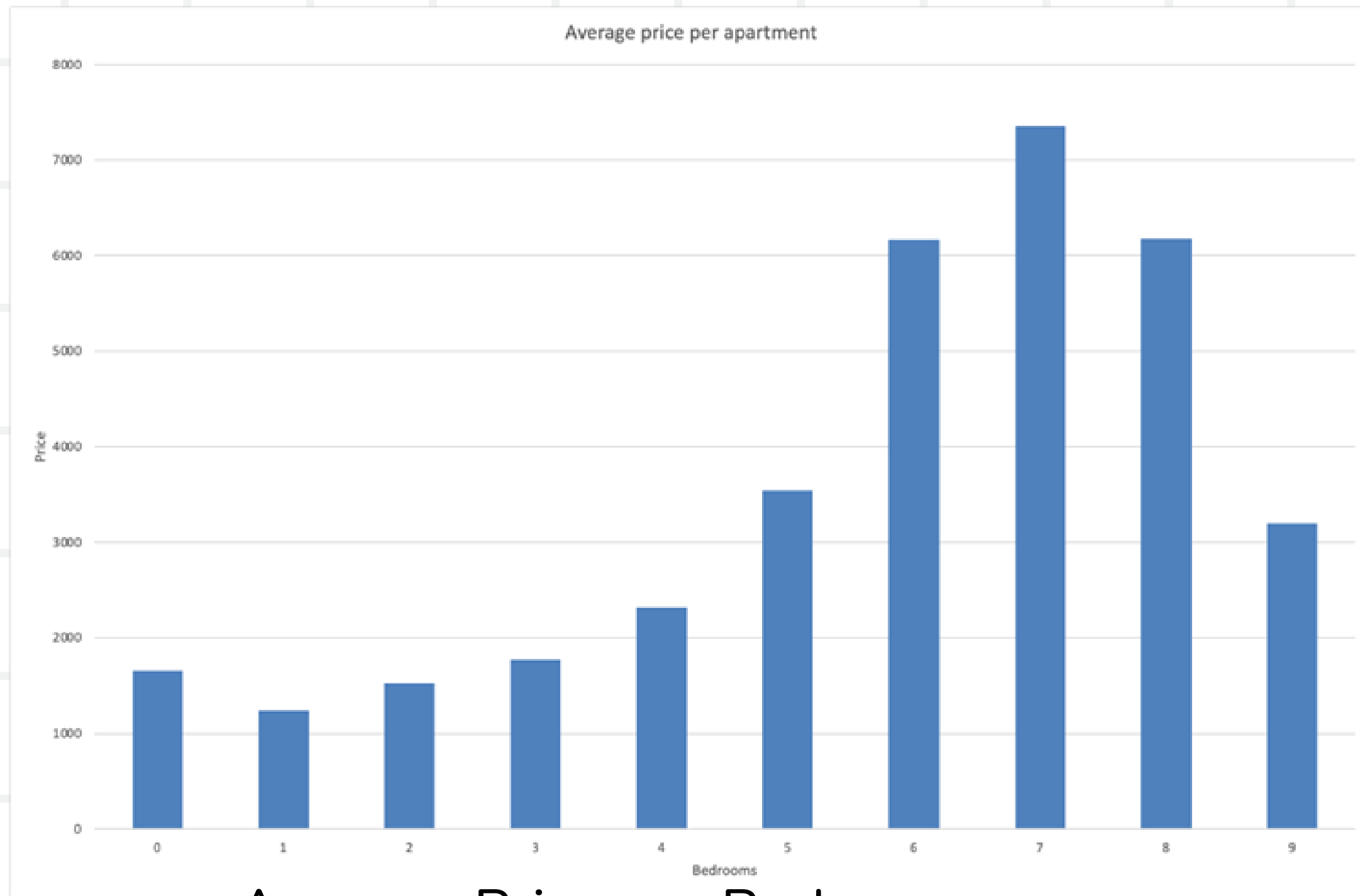
INITIAL ANALYSIS



Top 10 States with Highest Average Price per Unit

1	Hawaii	6	New Jersey
2	California	7	Maine
3	Massachusetts	8	Montana
4	Washington	9	New York
5	Rhode Island	10	Maryland

INITIAL ANALYSIS



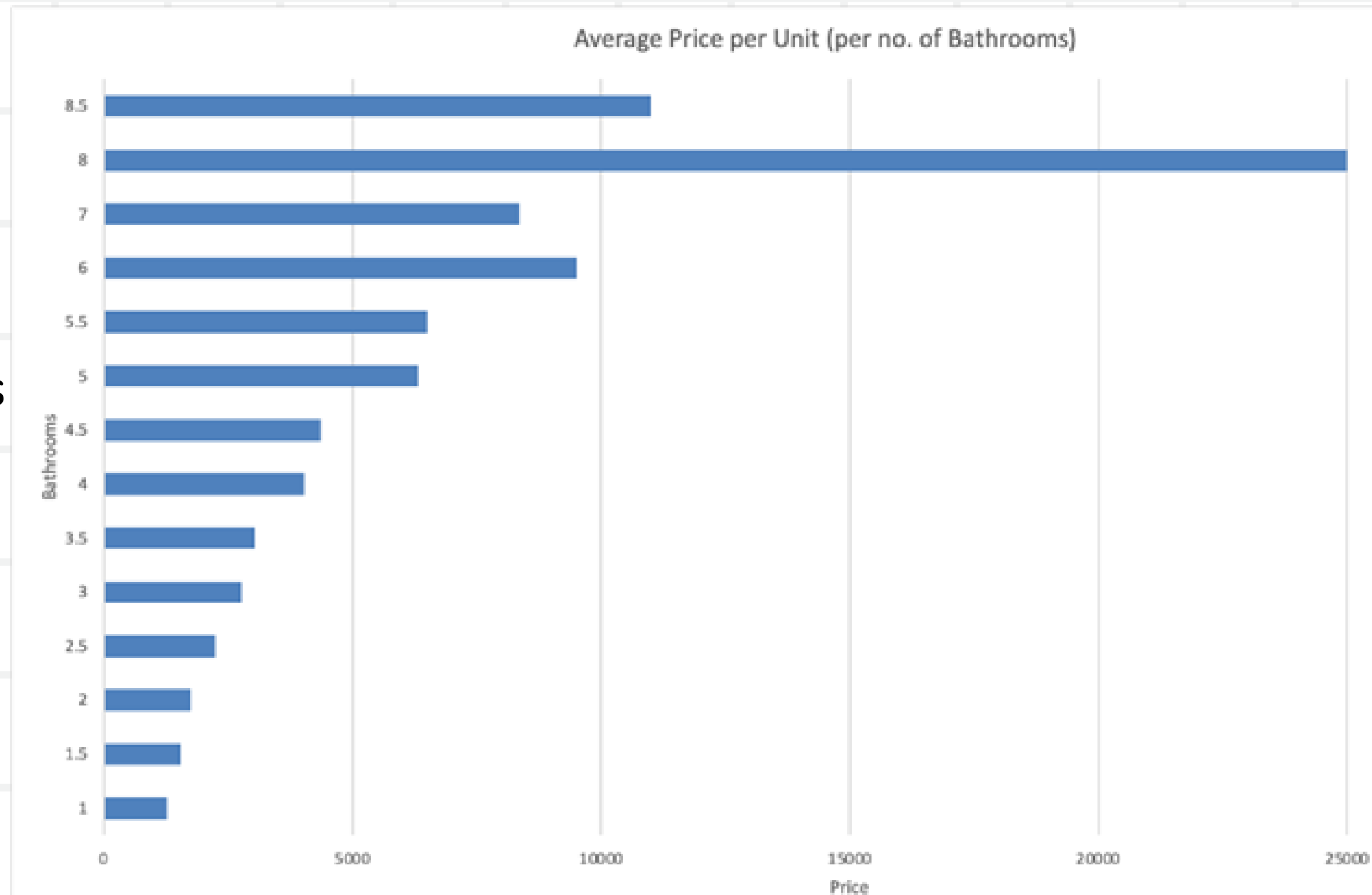
Average Price per Bedrooms

- Apartments with **2** bedrooms have the **lowest** cost of around **\$1200**
- Apartments with **7** bedrooms have the **highest** cost of **>\$7000**
- Apartments with **9** bedrooms cost **lesser** than apartments with **5,6,7 & 8** bedrooms
- **Cost** of apartments **peaked** at **7** bedrooms and decreased thereafter

INITIAL ANALYSIS

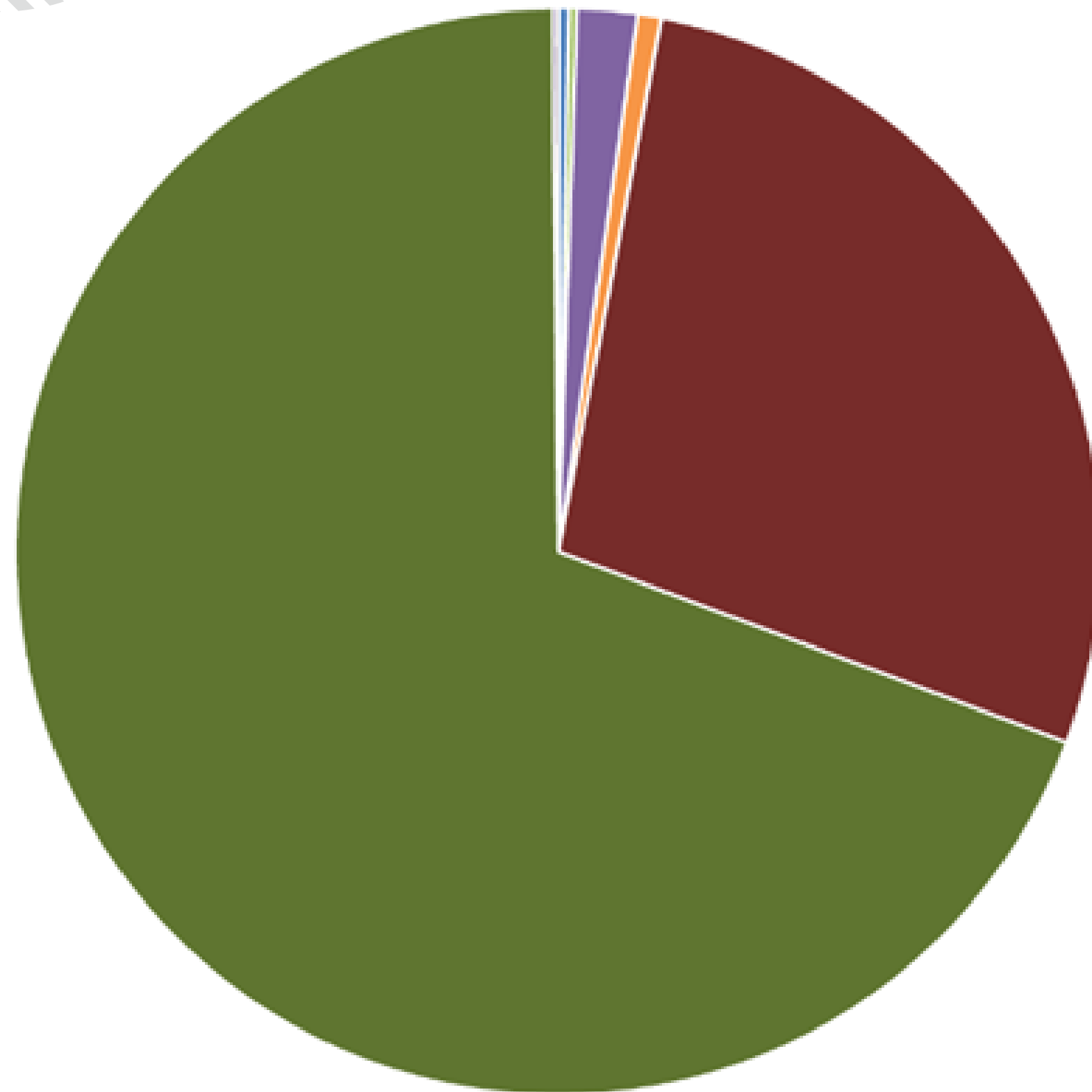


- Apartments with 1 bathroom have the **lowest** cost of around **\$1200**
- Apartments with 8 bathrooms have the **highest** cost of **\$25000**
- **Cost** of apartments **peaked** at 8 bathrooms and decreased thereafter



Average Price per Bathrooms

INITIAL ANALYSIS



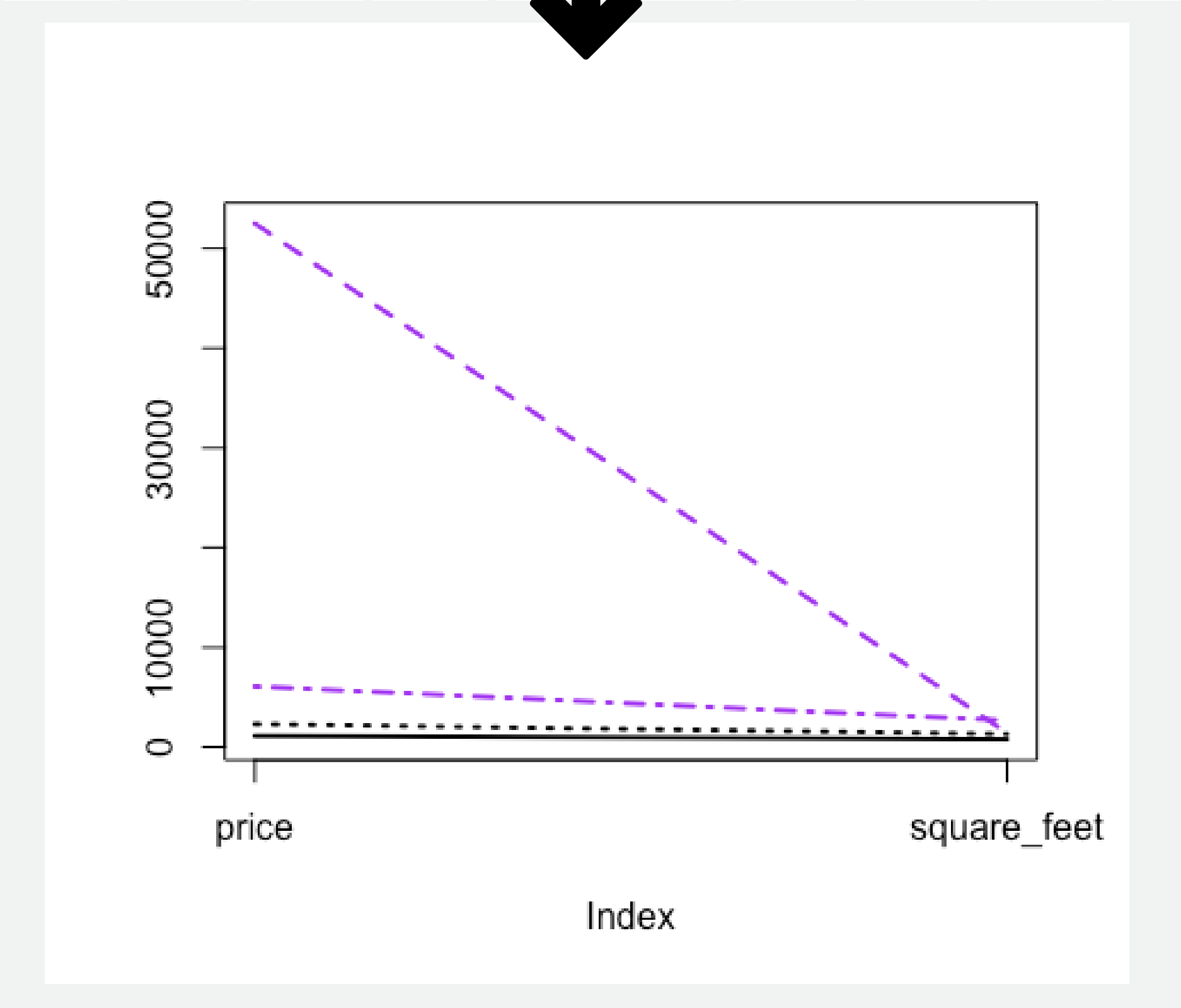
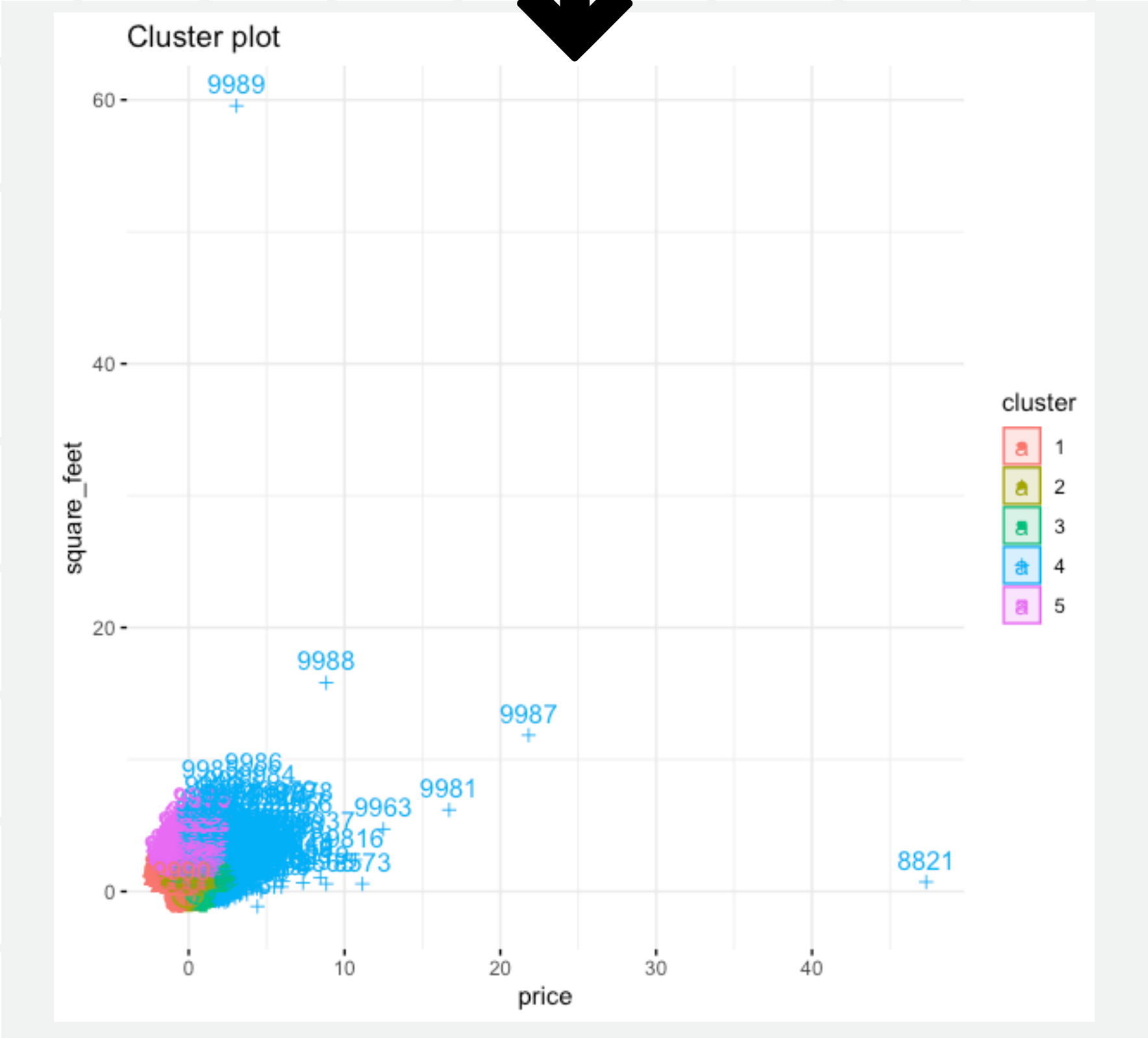
Apartment Listing Websites

GoSection8
Home Rentals
Listanza
ListedBuy
Real Estate Agent
RealRentals
rentbits
RentDigs.com
RentLingo
RENTOCULAR
tenantcloud
FALSE

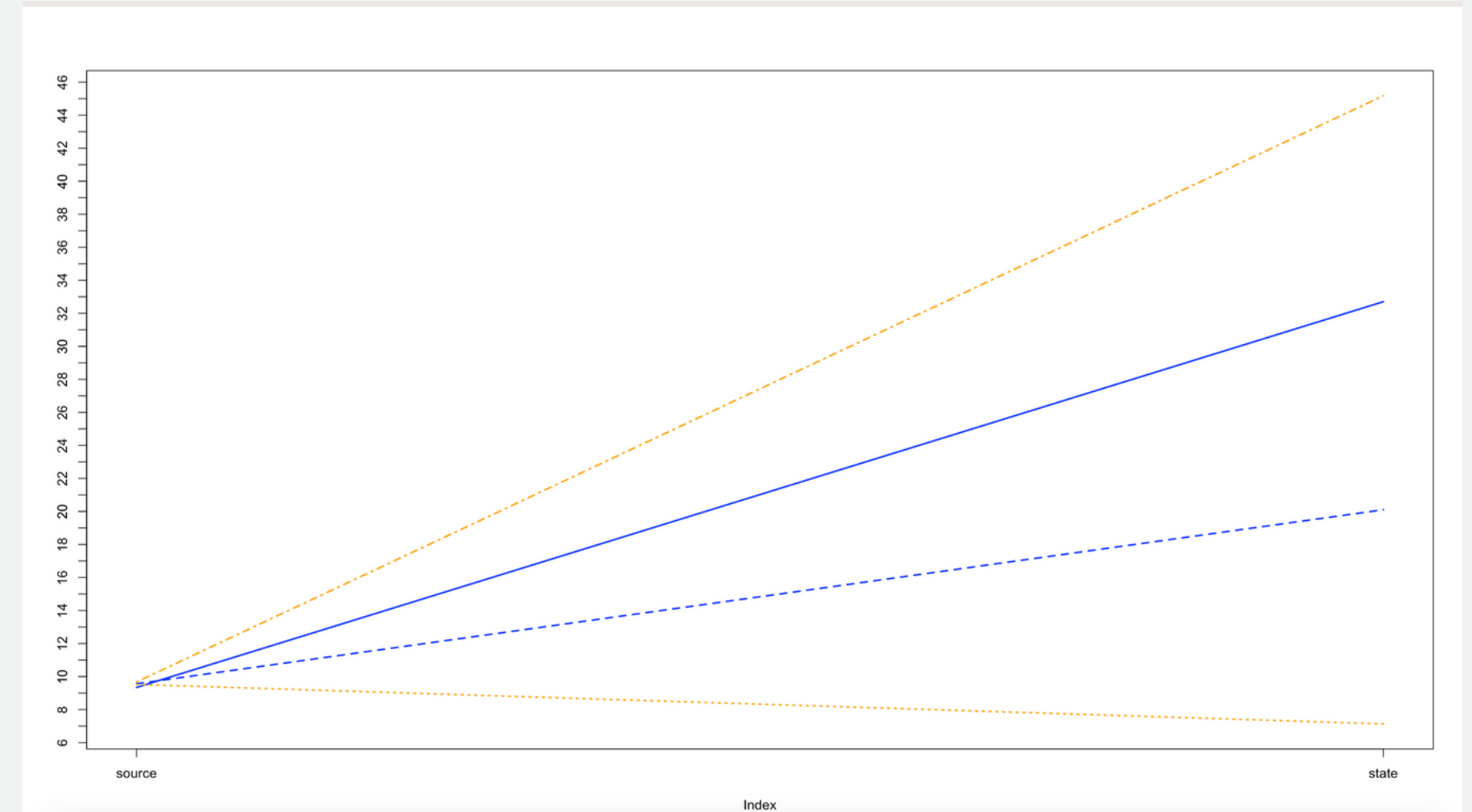
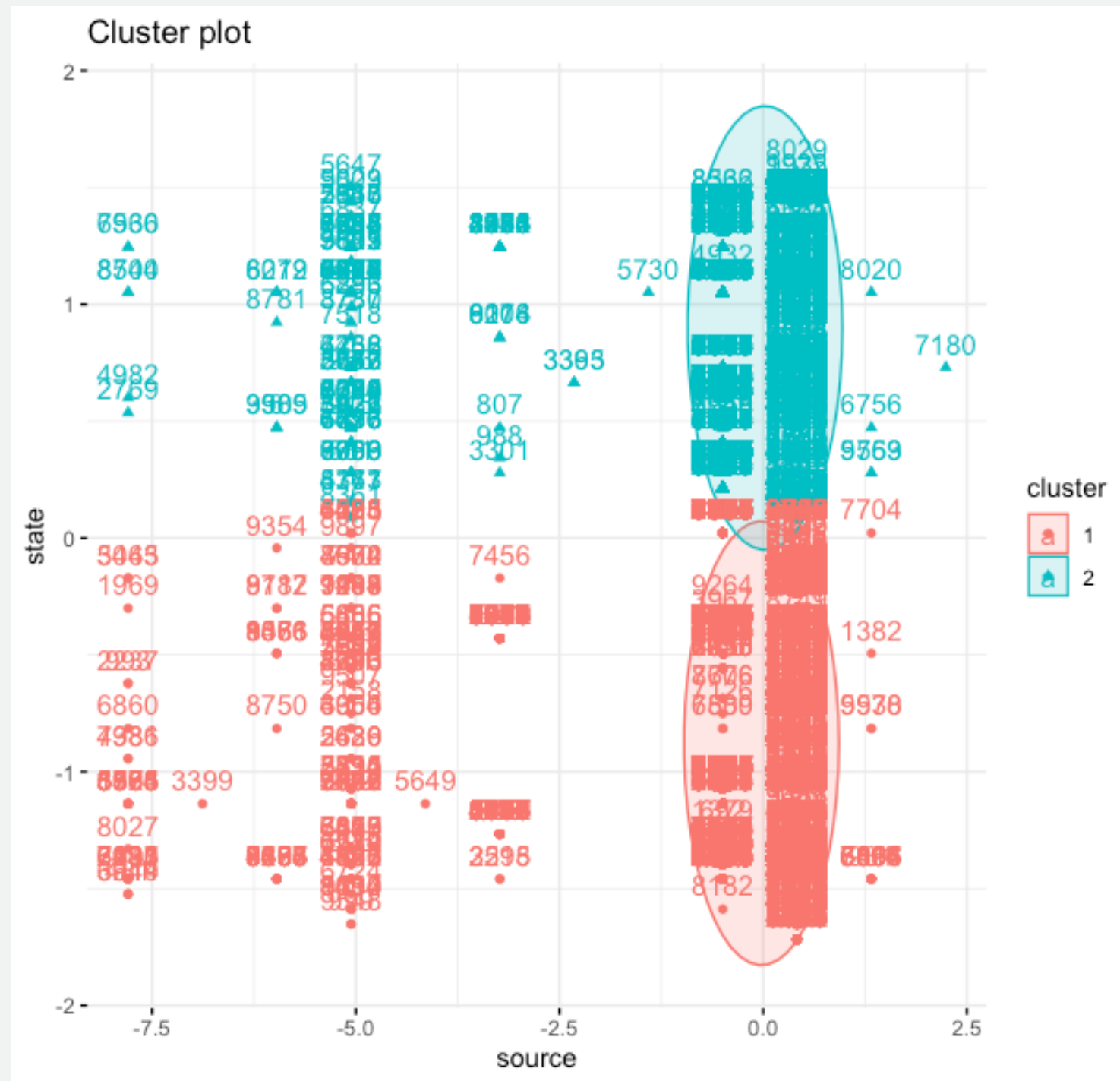
Listing Websites

- **RentLingo** is the **most popular** website for apartment listing with almost **70%** of listings being made on RentLingo
- The **second most popular** website is **RentDigs.com** with around **20%** of listings being made on it.

CLUSTERING



CLUSTERING



REGRESSION

DEPENDANT VARIABLE

Price

INDEPENDANT VARIABLES

regular: square_feet (***), square_feet3,
bathrooms, bedrooms, state
interaction: square_feet×bathrooms,
square_feet3×bedrooms

REGRESSION

CLEANING

- Remove missing records
- Replace missing value with mean
- Replace 'null' string with 'None' (pets_allowed) or 'Unavailable' (address)
- Format 'bedrooms' and 'bathrooms' as numeric columns
- Convert all rent periods to 'monthly'

PRE-PROCESSING

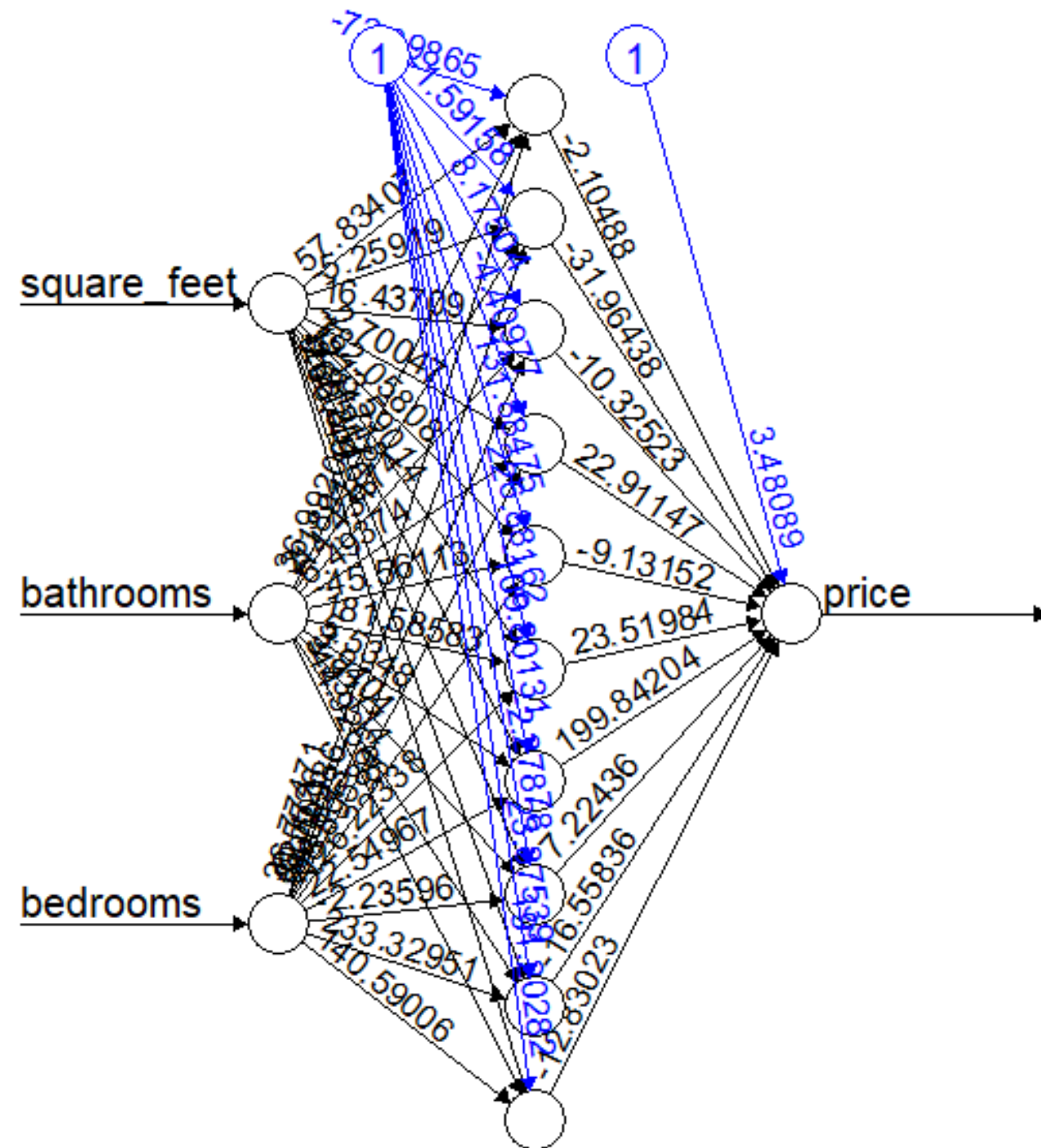
- Remove outliers using Cooks'D

REGRESSION

SUMMARY

- Residual standard error: 438.1 on 9724 degrees of freedom
- Multiple R-squared: 0.636
- Adjusted R-squared: 0.6339
- F-statistic: 308.9 on 55 and 9724 DF
- p-value: $< 2.2e-16$

NEURAL NETWORKS



DEPENDANT VARIABLE

Price

INDEPENDANT VARIABLE

square_feet
bathrooms
bedrooms

NEURAL NETWORKS

SUMMARY

- Mean Absolute Error (MAE): 0.5217533
- Mean Squared Error (MSE): 0.5585854
- Root Mean Squared Error (RMSE): 0.7473857
- Relative Root Mean Squared Error (RMSE): 74.73857
- R-Squared: 0.1536861

MODEL PERFORMANCES

CHOICE

In our case, **regression** performed better than **neural network**.


POINTERS

- Include categorical variables in the neural network
- Try different combinations of hidden layers and learning rates



CONCLUSION

In summary, this project utilized Business Intelligence techniques, including clustering, regression, and neural networks, to analyze rental apartment data and address the primary goal of predicting price optimization and market segmentation. Additionally, it provided valuable insights for more effective business strategies. Through clustering, distinct market segments were identified, while regression analysis illuminated factors influencing rental prices. The incorporation of neural networks showcased a commitment to innovative predictive modeling. The project outcomes contribute essential information for strategic decision-making in the dynamic rental housing market.



REFERENCES



🔍 DATA SOURCE

<https://archive.ics.uci.edu/dataset/555/apartment+for+rent+classified>

THANK YOU

Group 6

