



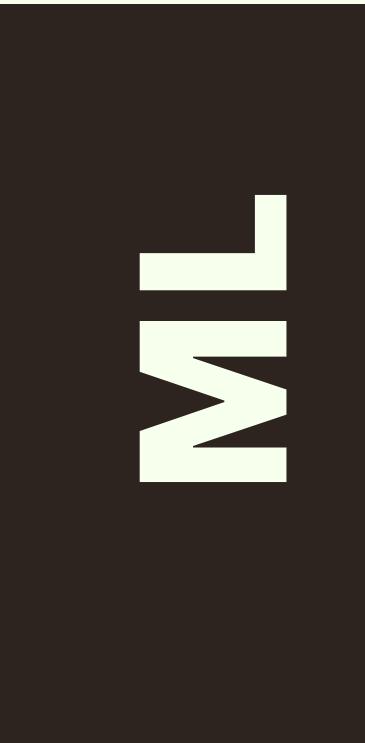
# RealEstate Insight Engine



PREDICTIVE PRICING, SMART RECOMMENDATIONS, AND MARKET ANALYTICS

Group 2

Caroline Bordin, Erjola Lluka,  
Jihye Yoon, Sughra Shadab,  
Supreet Ahuja



# Data Sourcing and Data Cleaning

The dataset was sourced from

[Kaggle.com](https://www.kaggle.com/datasets/polartech/500000-us-homes-data-for-sale-properties)

(<https://www.kaggle.com/datasets/polartech/500000-us-homes-data-for-sale-properties>)

**All the Data Analytics Projects or tasks start from getting the Data Cleaning**

- Removed null values and unnecessary columns, Handled data types, and converted units for consistency.
- Dropped rows with specific criteria (e.g., null values in 'price\_per\_unit' and 'bathroom\_number').
- Removing Outliers, using statistical methods.

## Sample of the cleaned dataset

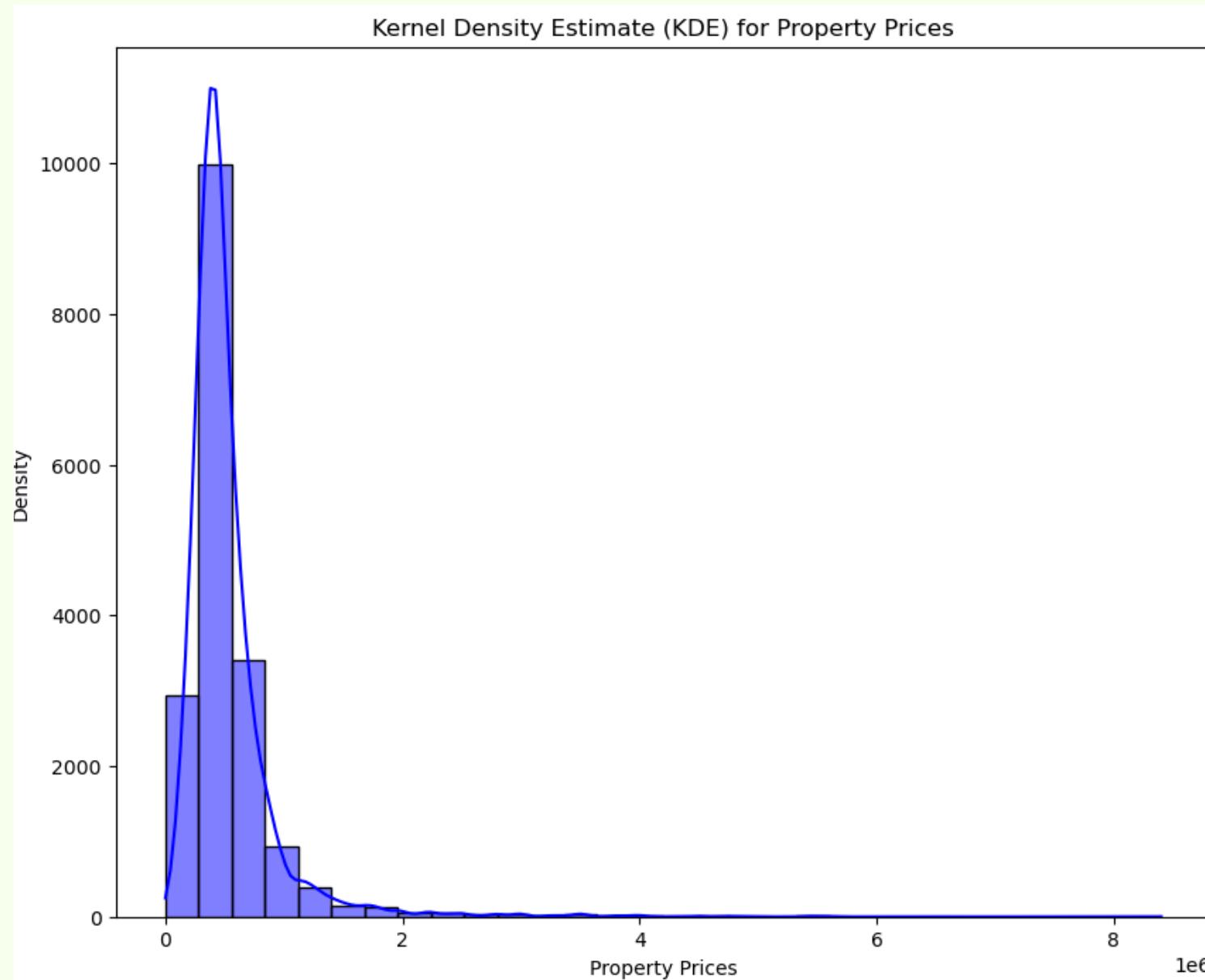
| property_id | address         | city          | state | latitude  | longitude  | price  | bedroom_number | bathroom_number | living_space |
|-------------|-----------------|---------------|-------|-----------|------------|--------|----------------|-----------------|--------------|
| 8352055     | t, Bullhead Cit | Bullhead City | AZ    | 35.11426  | -114.61839 | 54900  | 2              | 1               | 688          |
| 64954116    | r UNIT 33, Bu   | Bullhead City | AZ    | 35.1227   | -114.58975 | 379900 | 3              | 3               | 1649         |
| 71677227    | Dr, Bullhead    | Bullhead City | AZ    | 35.1181   | -114.62062 | 549900 | 3              | 4               | 3250         |
| 8351548     | t, Bullhead Cit | Bullhead City | AZ    | 35.114693 | -114.61625 | 169900 | 3              | 2               | 1157         |
| 8352036     | t, Bullhead Cit | Bullhead City | AZ    | 35.1147   | -114.62122 | 169900 | 2              | 2               | 1440         |





# Visualizing Relationships

## Kernal Density Estimate (KDE) curve



Y - axis showing the density or frequency of prices

X-axis showing the price range

matplotlib

python™



Linear Model



- The curve in the shape of the Bell curve indicates that the distribution of the property prices is approximately normal. or we say that the prices are symmetrically distributed around the MEAN and prices are close to average
- The peak of the Bell curve represents the central tendency.
- A normal distribution indicates that there is a relatively stable and predictable pattern in the property pricing



# PRICE PREDICTION MODEL



This part of the project aims to predict  
Property prices using  
**Supervised Machine Learning**

by using the model - [RandomForestRegressor](#)

- Our model considered 4 basic features to predict property prices:
  1. Bedroom Number
  2. Bathroom Number
  3. Living Space
  4. City Name





# Price Prediction



*The sample of the code where 4 inputs of clients are entered*

```
# Using the model to predict 'price' by Client's input  
client_data = pd.DataFrame({'bedroom_number': [4], 'bathroom_number': [3], 'living_space': [2500], 'city': ['Chino Valley']})
```

*And below we can see the predicted price for '4 Beds', '3 Baths' with approx '2500sqft.' of 'living area' in the city 'Chino Valley' from the code with R2Square and Mean Asolute error*

**Predicted Price: [516541.8475]**

**R-squared (R2) Score on Training Data: 0.96211636469677**

**Mean Absolute Error on Training Data: 40150.864!**



# Linear Regression

Exploring relationships in Arizona State Real Estate Data

**Sneak peak of the dataset**

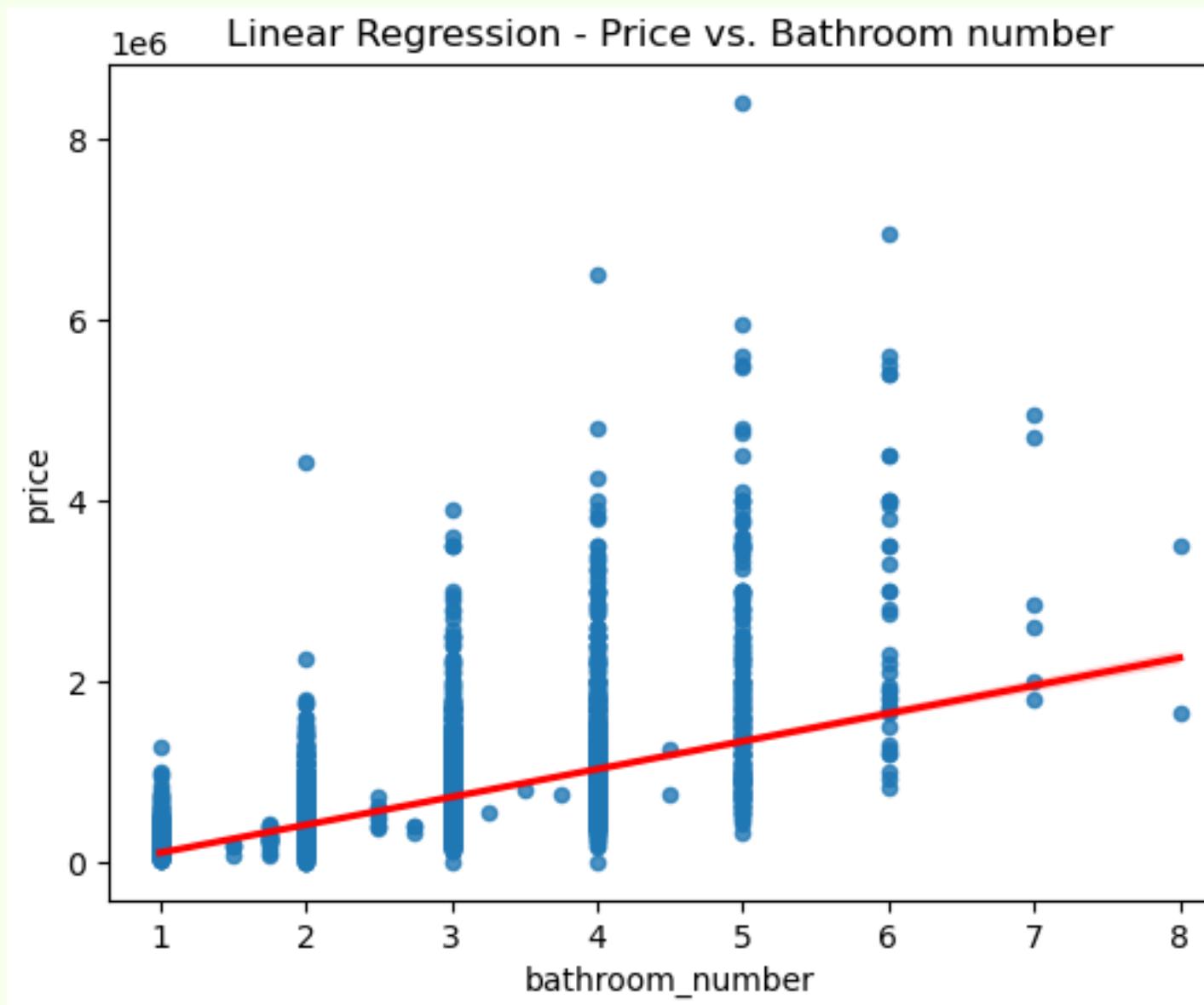
| property_id | address         | city          | state | latitude  | longitude  | price  | bedroom_number | bathroom_number | living_space |
|-------------|-----------------|---------------|-------|-----------|------------|--------|----------------|-----------------|--------------|
| 8352055     | t, Bullhead Cit | Bullhead City | AZ    | 35.11426  | -114.61839 | 54900  | 2              | 1               | 688          |
| 64954116    | r UNIT 33, Bu   | Bullhead City | AZ    | 35.1227   | -114.58975 | 379900 | 3              | 3               | 1649         |
| 71677227    | Dr, Bullhead C  | Bullhead City | AZ    | 35.1181   | -114.62062 | 549900 | 3              | 4               | 3250         |
| 8351548     | t, Bullhead Cit | Bullhead City | AZ    | 35.114693 | -114.61625 | 169900 | 3              | 2               | 1157         |
| 8352036     | t, Bullhead Cit | Bullhead City | AZ    | 35.1147   | -114.62122 | 169900 | 2              | 2               | 1440         |



**Chosen Features: Bathroom number and  
Living Space (sqft.)  
with  
Target variable:Price**



# Visualizing Relationships



Coefficients:

Bathroom\_number: 60836.14706558022

- A positive bathroom\_number coefficient implies an increase in price with additional bathrooms.
- A positive living\_space coefficient implies an increase in price with every unit added to the living space.

Living\_space: 317.9105964419876



# Property Recommendation



## Unsupervised Machine Learning

- 1. Discovering Hidden Patterns and Buyer Preferences**
- 2. Handling Lack of Clear Labels**
- 3. Learning from Limited Information**



Group2- Predictive Pricing, Smart Recommendations, and Market Analytics

# Process

1 GET DATA

2 CREATE A FUNCTION:  
Use SKLEARN.HAVERSINE FORMULA to  
calcute distances.

3 APPLY FILTER



# RESULT



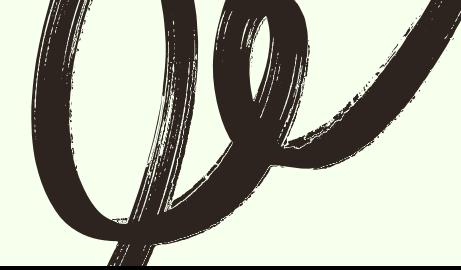
## Make a view History

By putting information of Property and Search

### Similar Nearby Properties:

|   | address  | property_id | latitude       | \               |              |   |
|---|--|-------------|----------------|-----------------|--------------|---|
| 1 | 411 Riverfront Dr L0T 5, Bullhead City, AZ 86442 | 104049502   | 35.1137        |                 |              |   |
|   | longitude  | distance    | bedroom_number | bathroom_number | living_space | \ |
| 1 | -114.63446                                       | 0.0         | 6.0            | 8.0             | 5173.0       |   |
|   | price  |             |                |                 |              |   |
| 1 | 2200000.0  |             |                |                 |              |   |





# WEBSITE: Property Recommendation

Demonstration.

## Property Search

Find properties that match your criteria!

Latitude: 34.766987  
Longitude: -114.49841  
Bedrooms: 4  
Bathrooms: 2  
Price: 189500

Please type the property's information!



Address: 854 Citrus St, Bullhead City, AZ 86442  
Location : Bullhead City  
Latitude :35.114260  
Longitude :-114.618385  
Bedroom\_number : 2  
Bathroom\_number :1  
Price : 54900



Address: 10739 S Paradise Trail Rd, Yucca, AZ 86438  
Location : YUCCA  
Latitude :34.91736  
Longitude :-114.15683  
Bedroom\_number : 2  
Bathroom\_number :2  
Price : 209000



Address: 1825 E Shore Villas Dr UNIT 51, Bullhead City, AZ 86442  
Location : Bullhead City  
Latitude :35.1227  
Longitude :-114.58975  
Bedroom\_number : 2  
Bathroom\_number :3  
Price : 349900



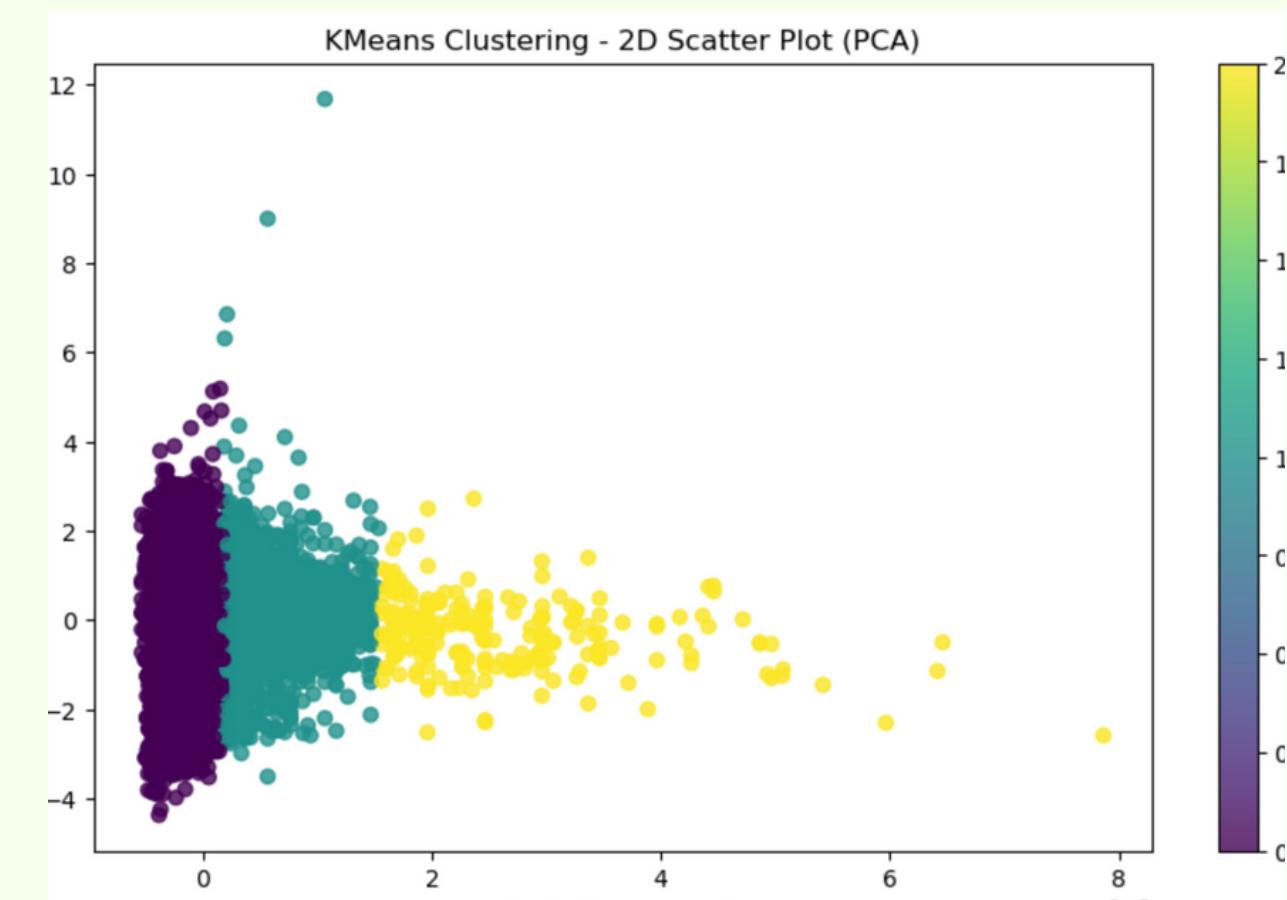
# Model Validation

Kmeans, PCA

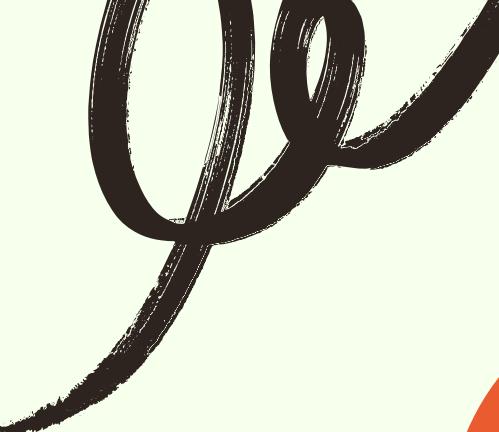
Clustering

Silhouette Score is

0.6647729692800081



Silhouette Score: 0.6647729692800081



# Market Analysis

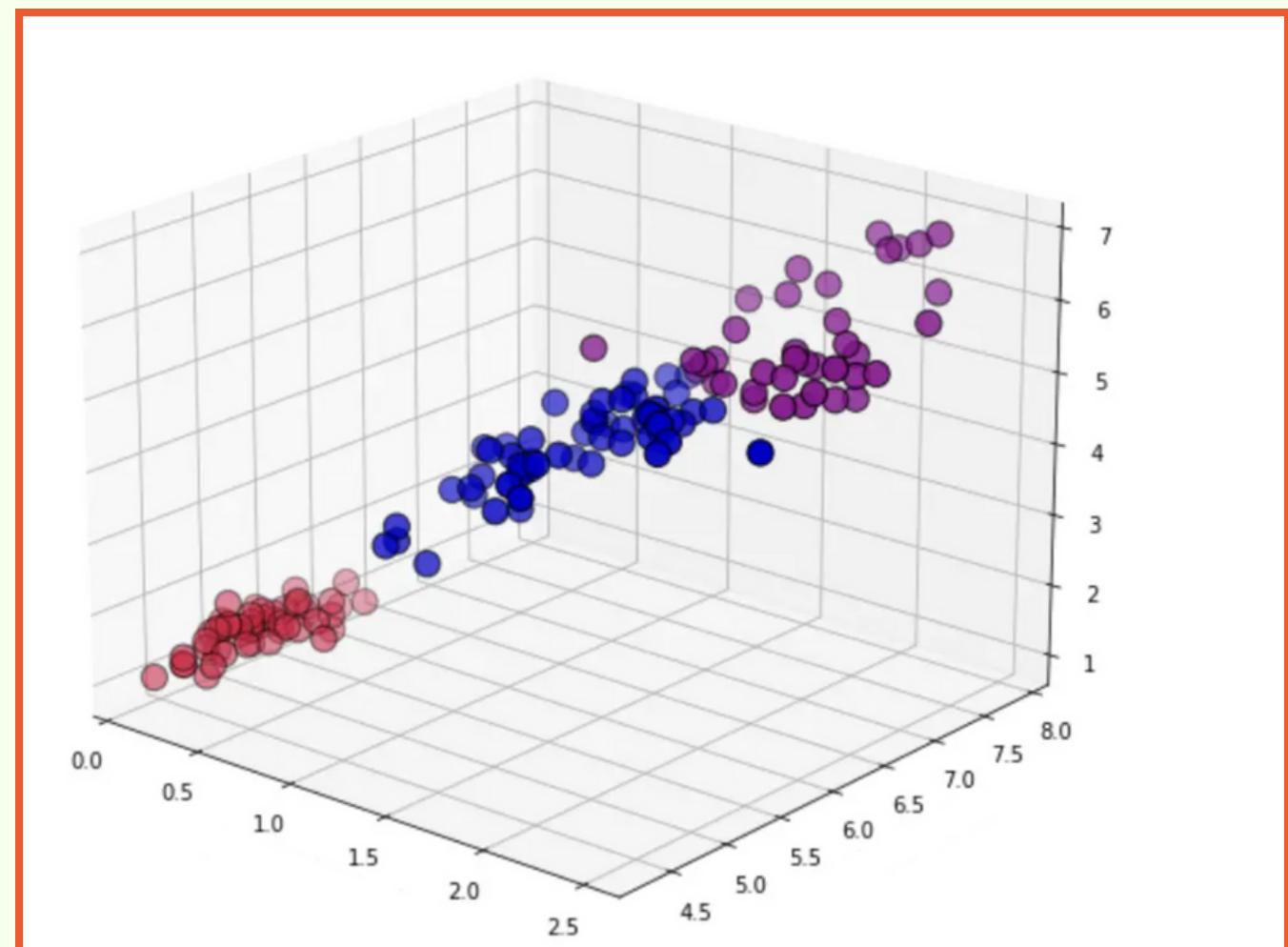


## Unsupervised Machine Learning

**Identify trends,  
hotspots, and  
investment opportunities.**

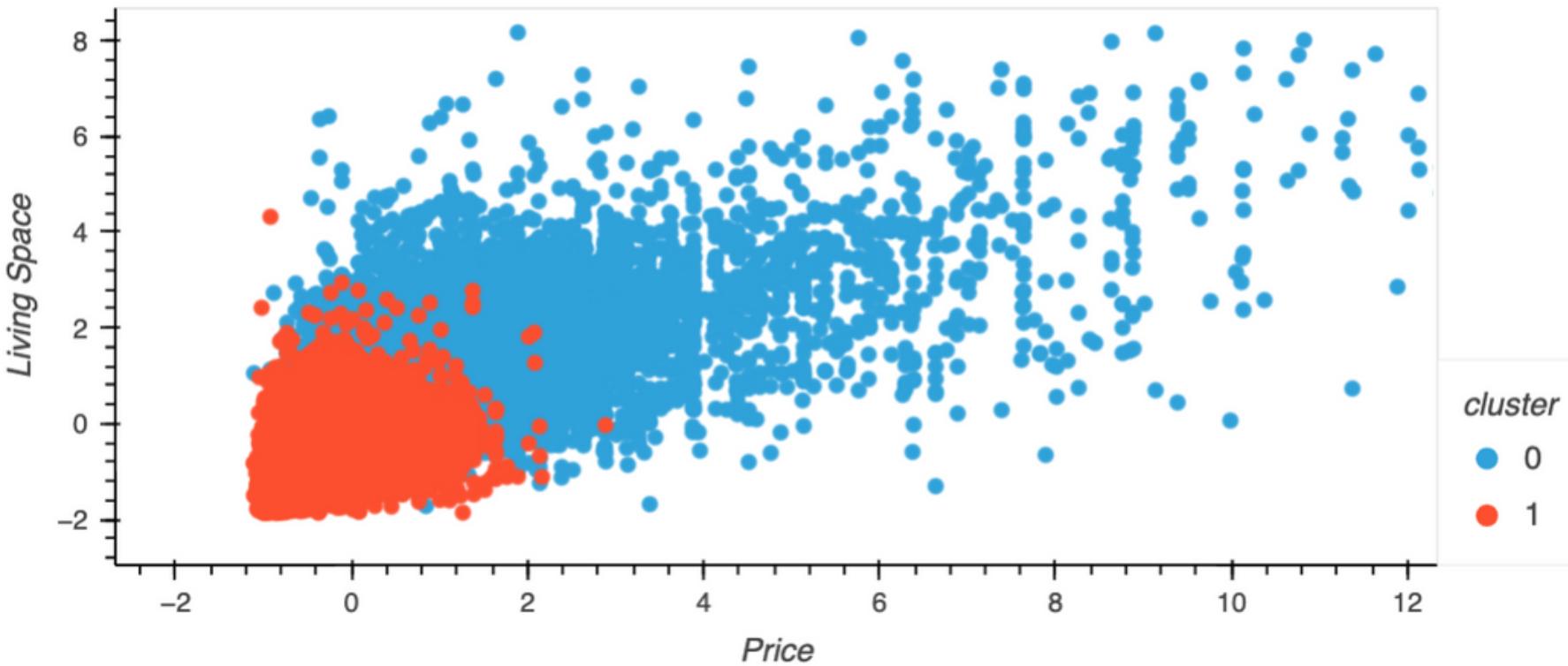
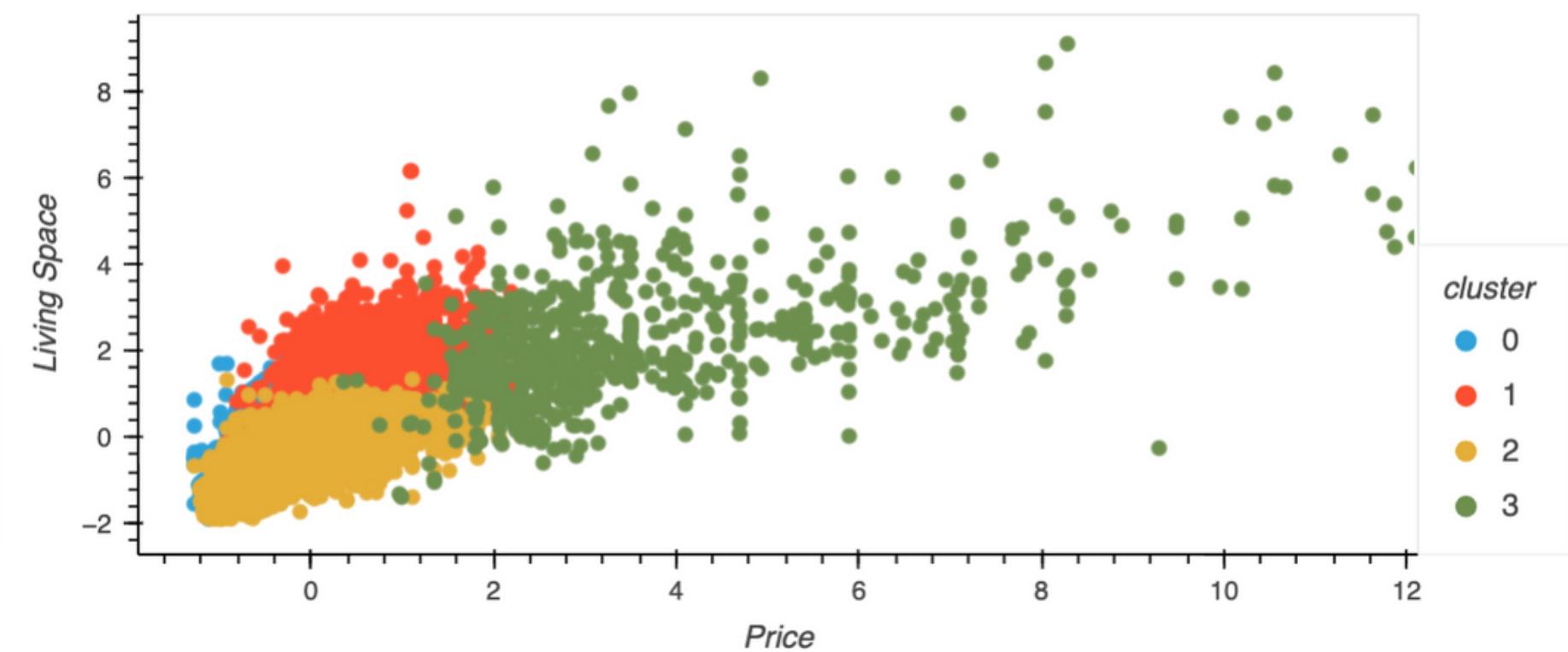
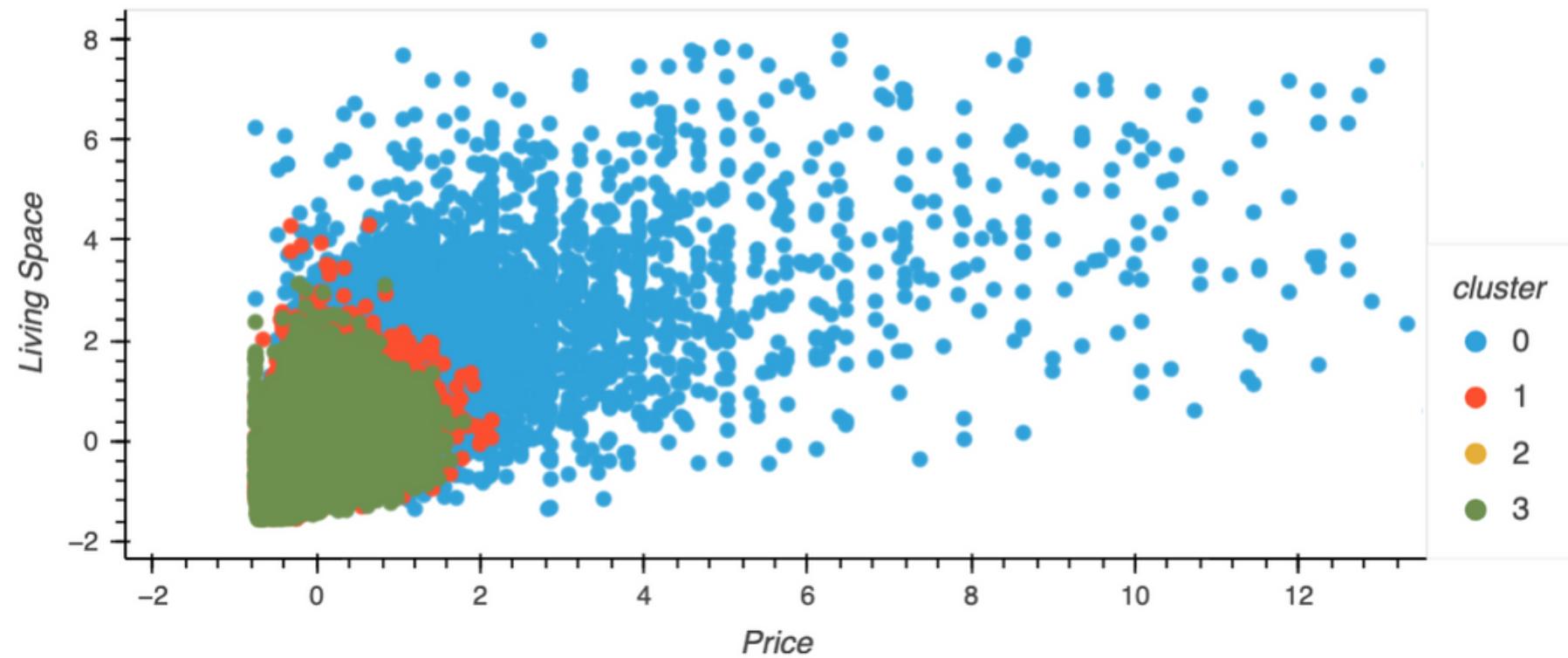
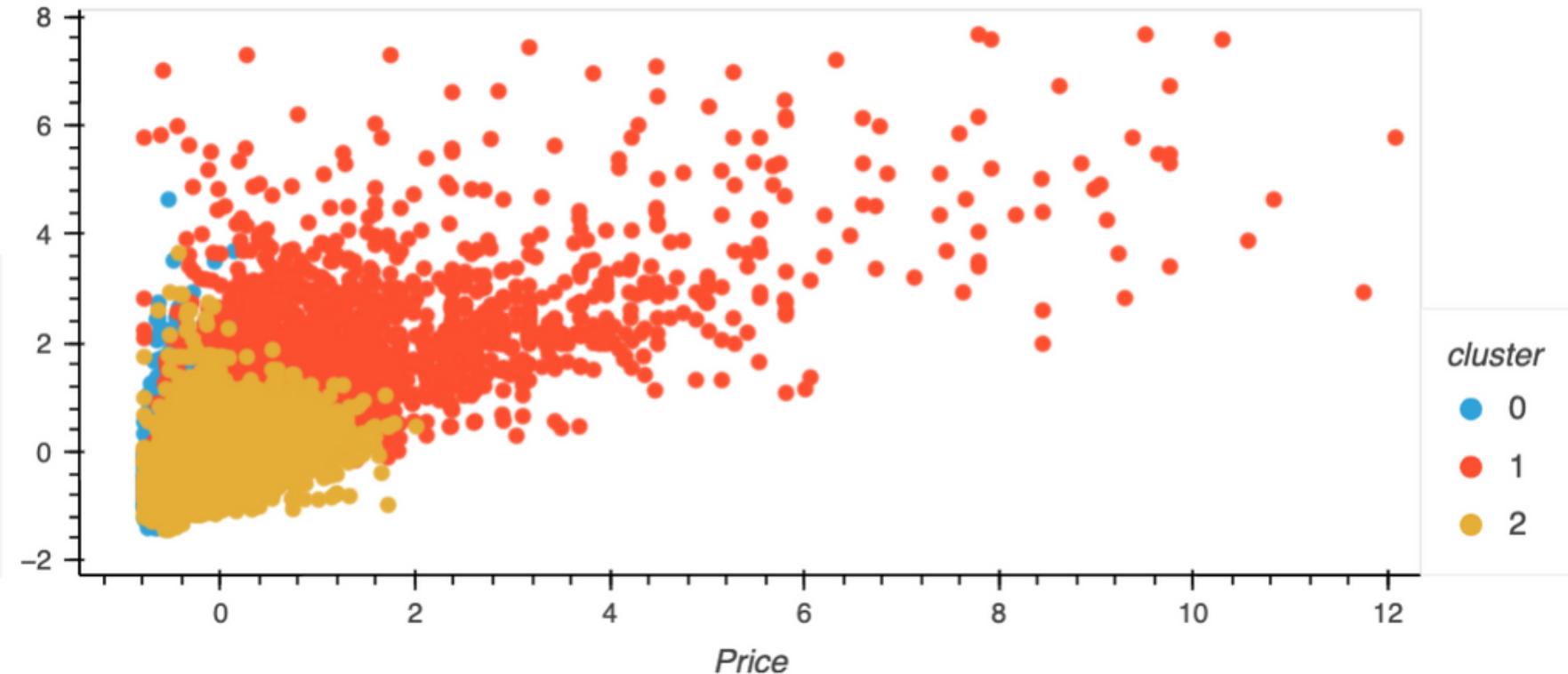
Group2- Predictive Pricing, Smart Recommendations, and Market Analytics

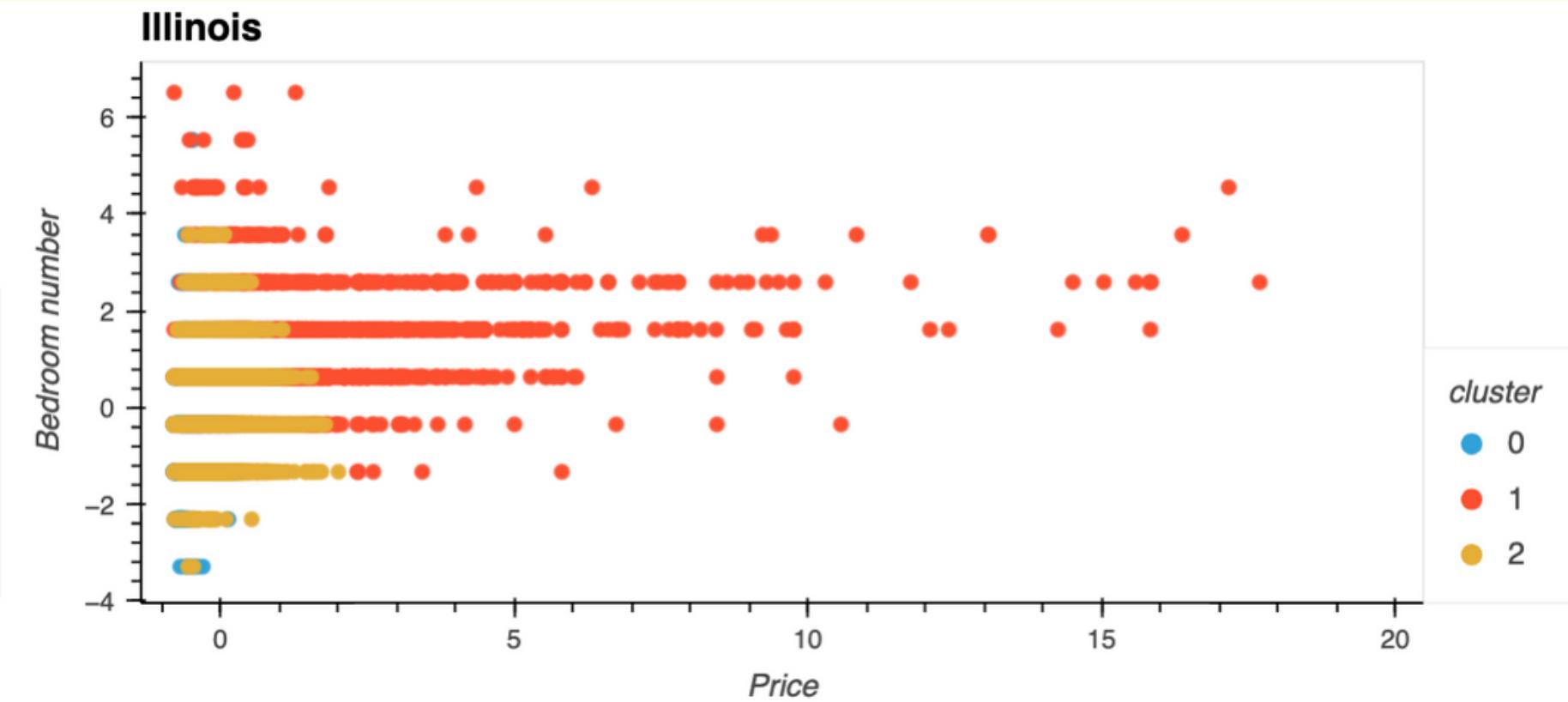
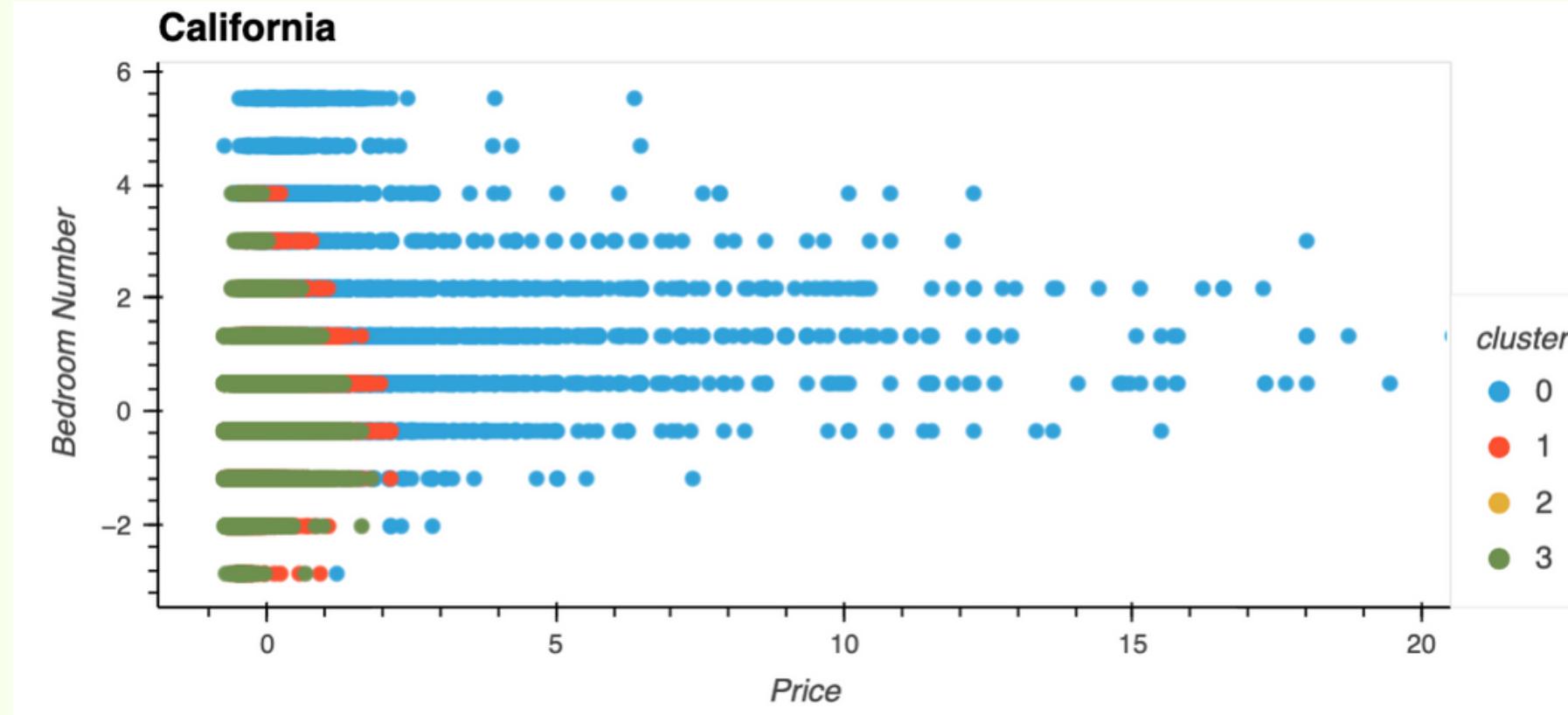
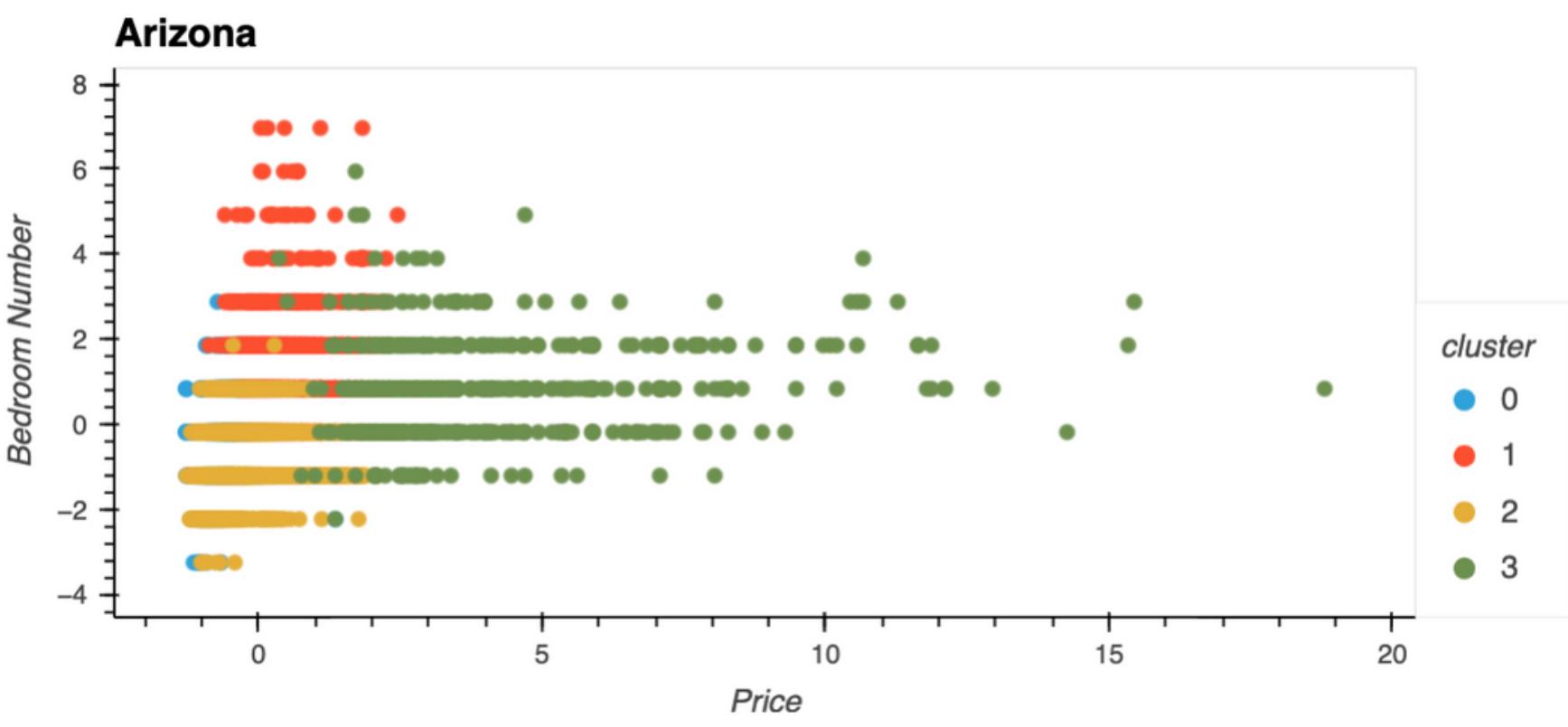
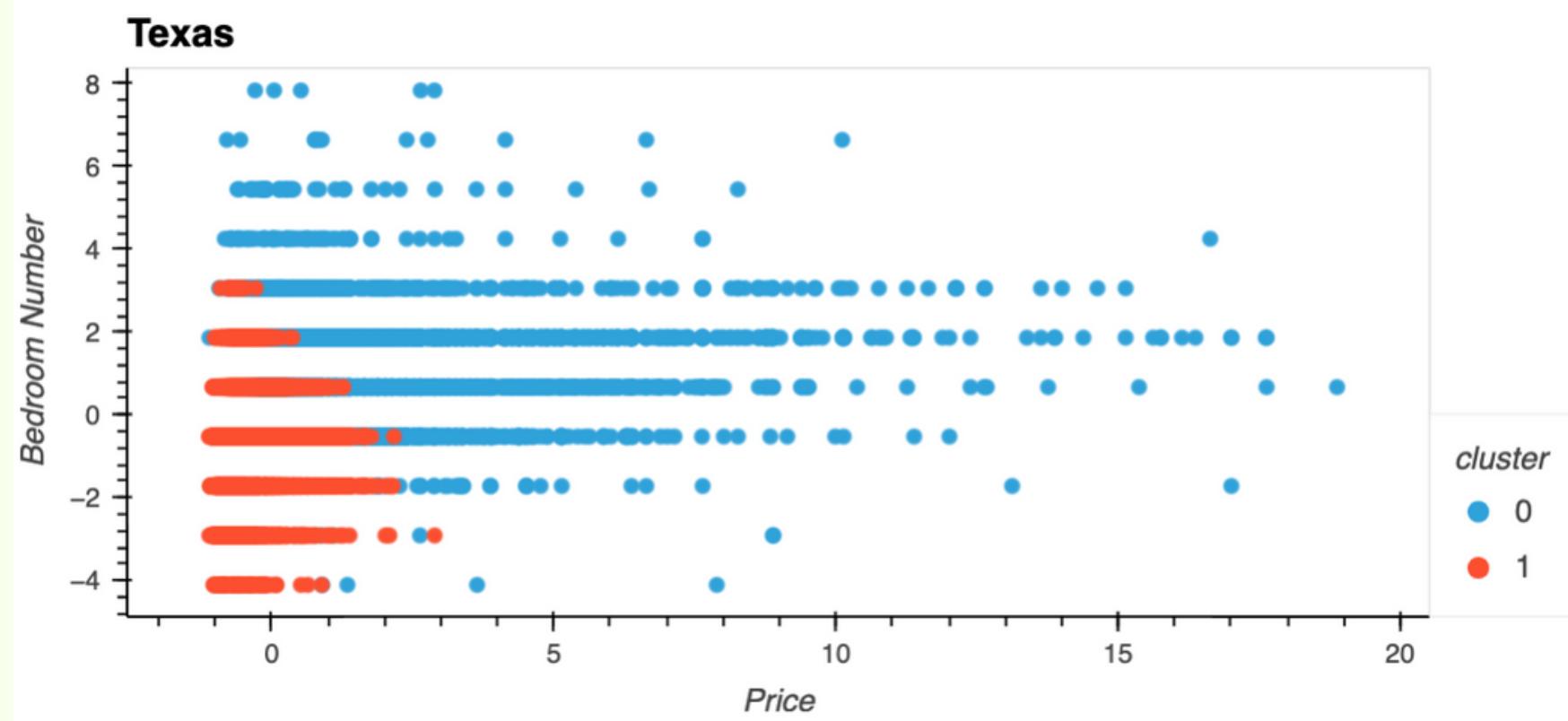
# Process



- 1 GET DATA
- 2 USE K-MEANS CLUSTERING
- 3 IDENTIFY PATTERNS



**Texas****Arizona****California****Illinois**





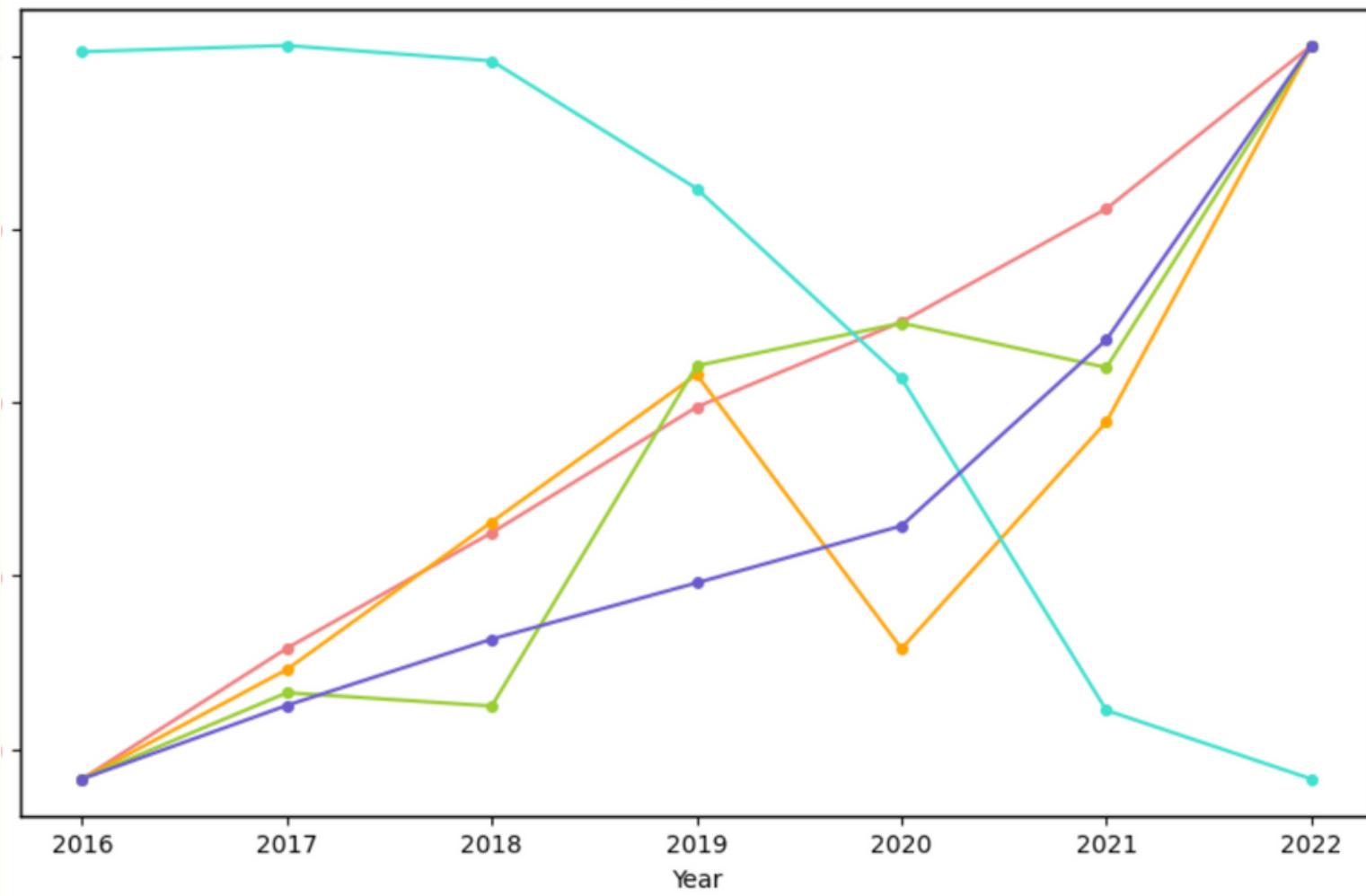
# Investment Opportunities

WHAT IS THE PURPOSE OF THE INVESTMENT?

- 1 BUY AND SELF-USE → **HOUSE PRICE**
- 2 BUY AND LEASE → **HOUSE PRICE /  
INCOME**
- 3 BUY AND SELL → **HOUSE PRICE/  
INCOME/POPULATION**

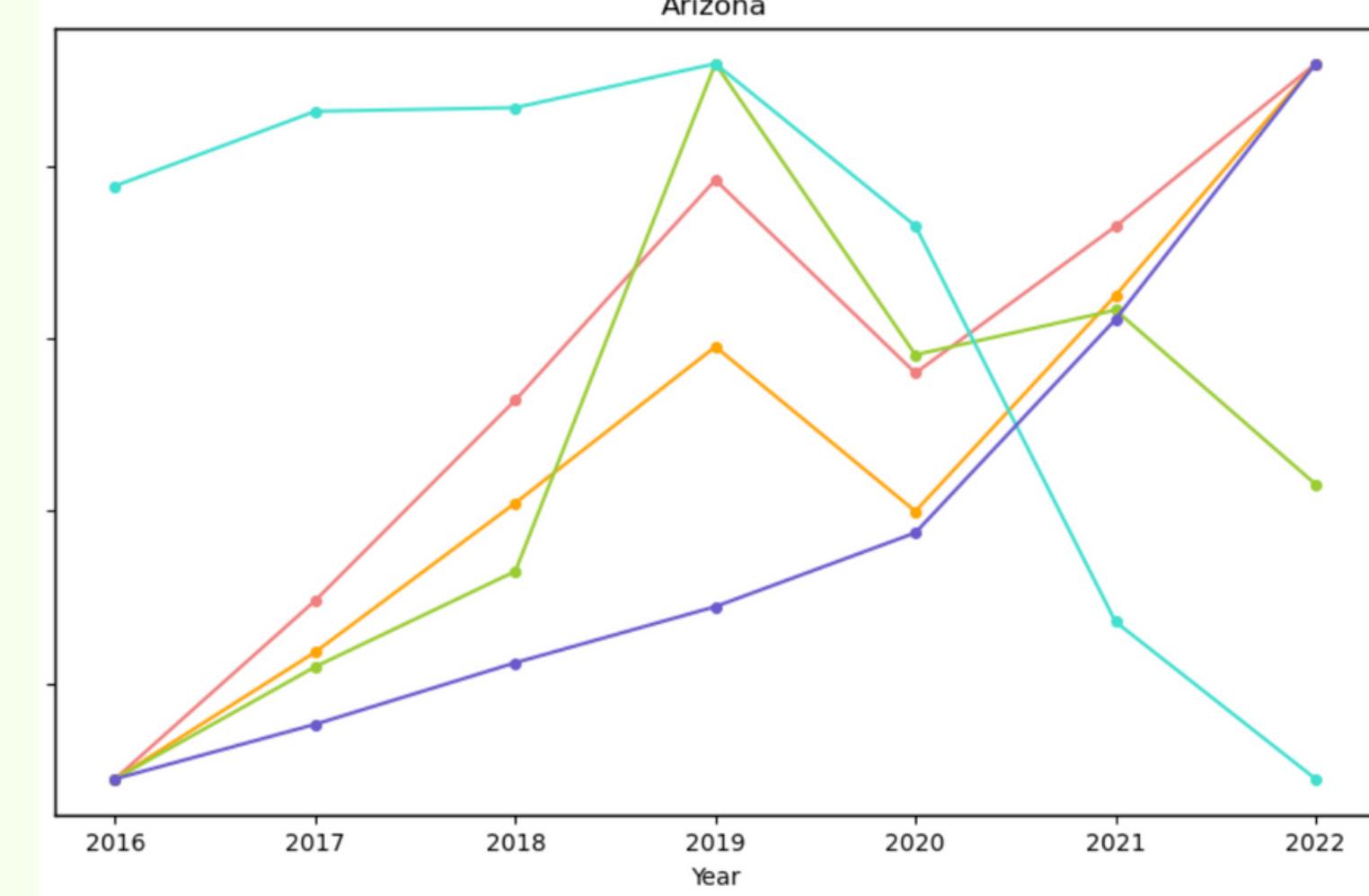


Texas

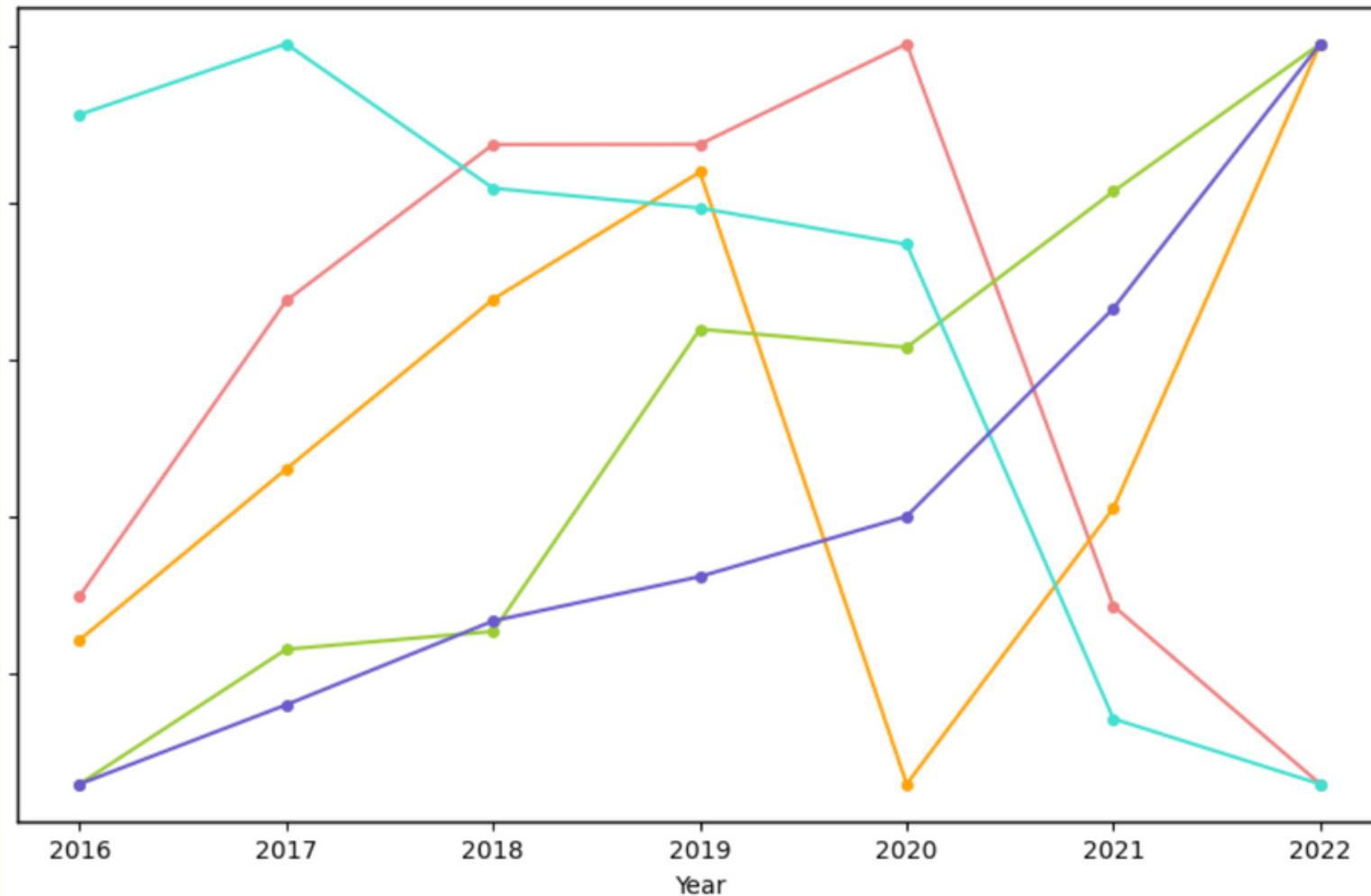


- population
- employment
- income
- home size(sqft)
- house prices

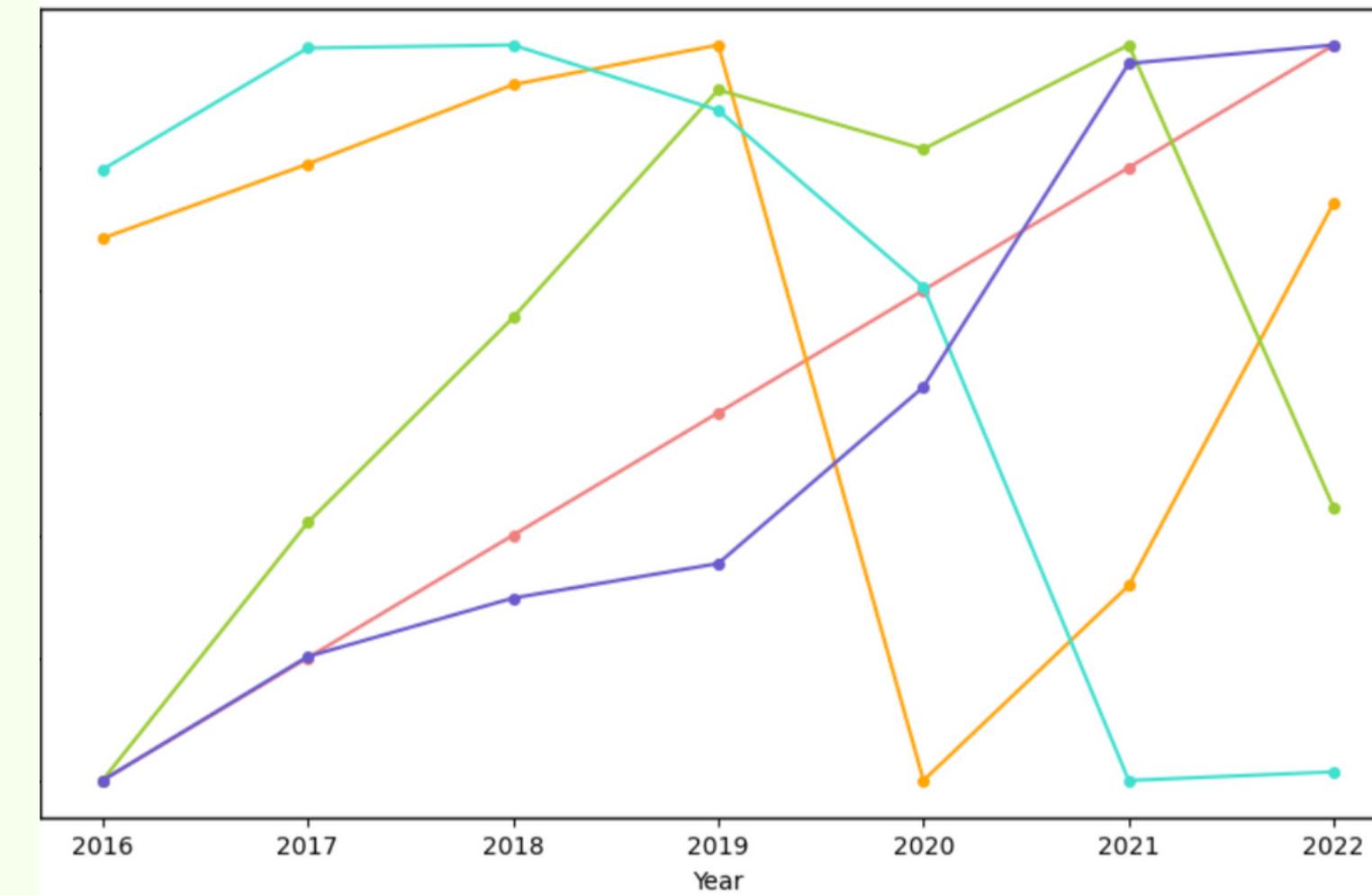
Arizona



California



Illinois

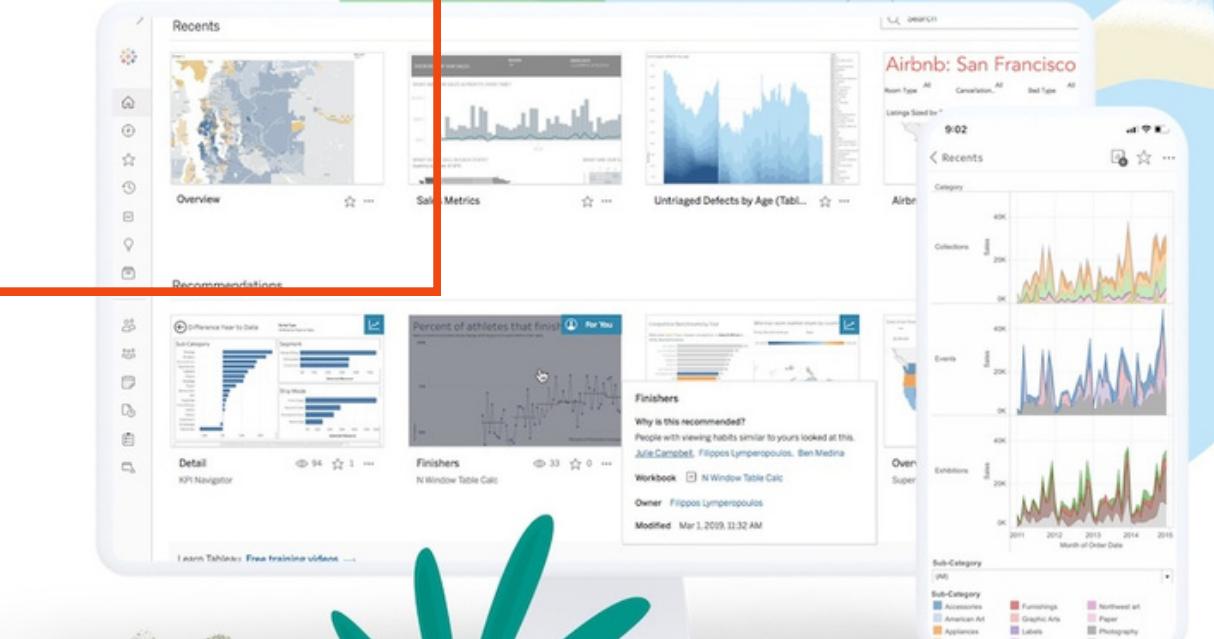




# Visualizing and Analyzing Data



## Tableau



Group2- Predictive Pricing, Smart Recommendations, and Market Analytics

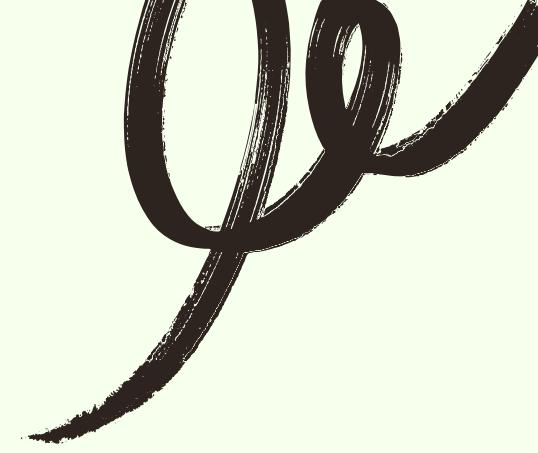


# Link to dashboard

[https://public.tableau.com/app/profile/sughra.shadab/viz/Project4-Dashboard2\\_17007549160420/Dashboard2?publish=yes](https://public.tableau.com/app/profile/sughra.shadab/viz/Project4-Dashboard2_17007549160420/Dashboard2?publish=yes)



- 1. Relationship between Location and Price**
- 2. Property count and Property Types**
- 3. Minimum and Maximum Price of Property**
- 4. Features Affecting selling Prices**



# Thank you!

Group2- Predictive Pricing, Smart Recommendations, and Market Analytics