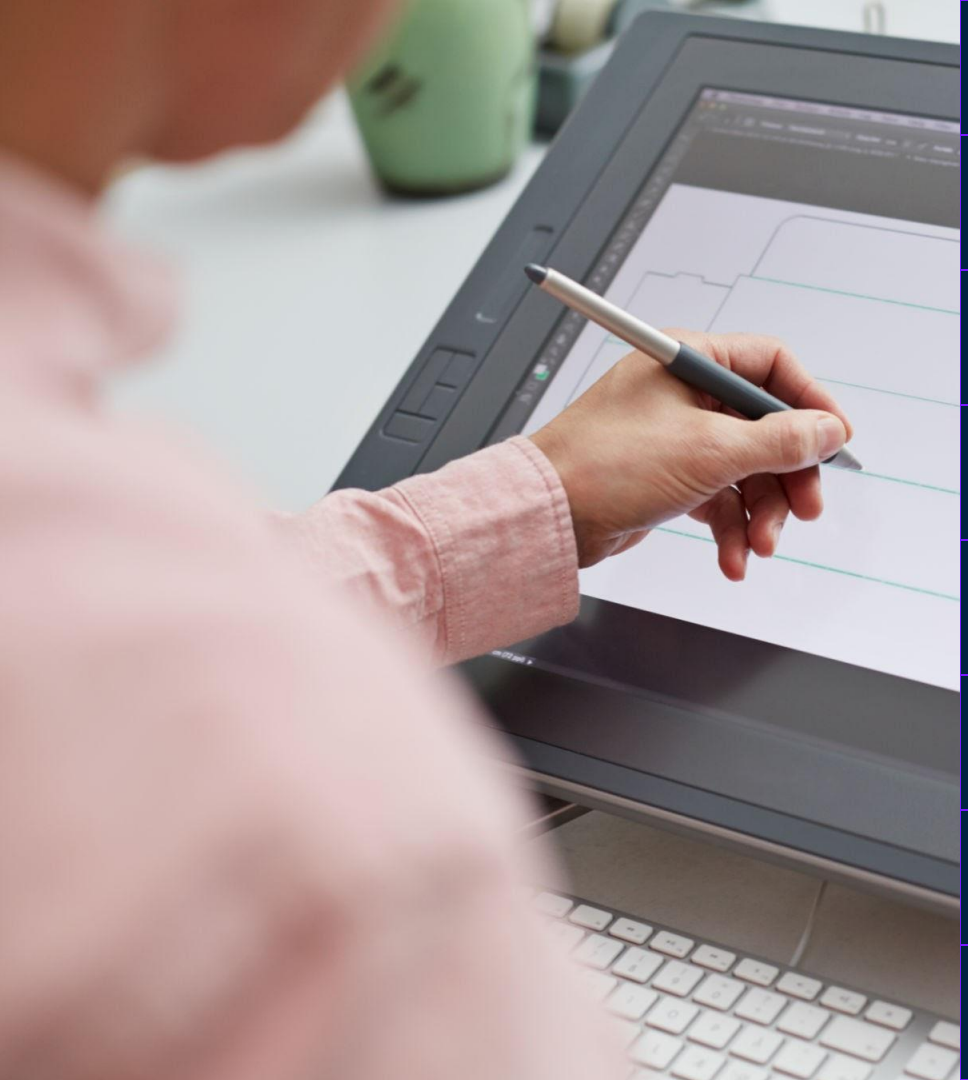


Property Tax Analysis for Real Estate Professionals

Vaishali Patelia

Outline

- Introduction - Use Case
- Data Analysis & Visualization
- Implementation
  - KNN Regression Model
  - MV Regression Model
- Result Analysis
- Conclusion

Introduction

Estimating property taxes is complicated due to varied attributes, geographic differences, and tax regulations.

Inaccurate estimates affect valuations, investments, and financial planning for stakeholders.

Provide precise property tax estimates.

Assist real estate professionals with actionable insights for smarter decision-making.

Property tax depend on Property value, Tax Rate, Location, Property type

Data Analysis

Dataset - cleaned_house_data

Property tax depend on

Property value - sold price, sq_ft, bed, bath, lot_acres

Location - zip, lat,long

New column: taxes_sqft , price_zone

Divide taxes_sqft into 5 category (Band of Tax) by using following method:

Pd.qcut
Pd.cut
Pd.cut (linespace)

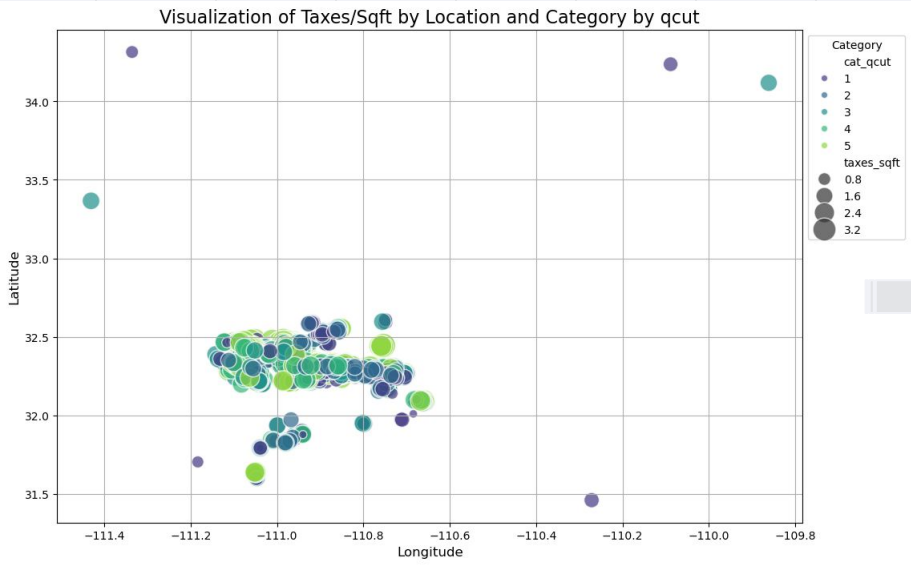
```
]: df[['taxes_sqft','price_zone','bedrooms','bathrooms','lot_acres','year_built']]
```

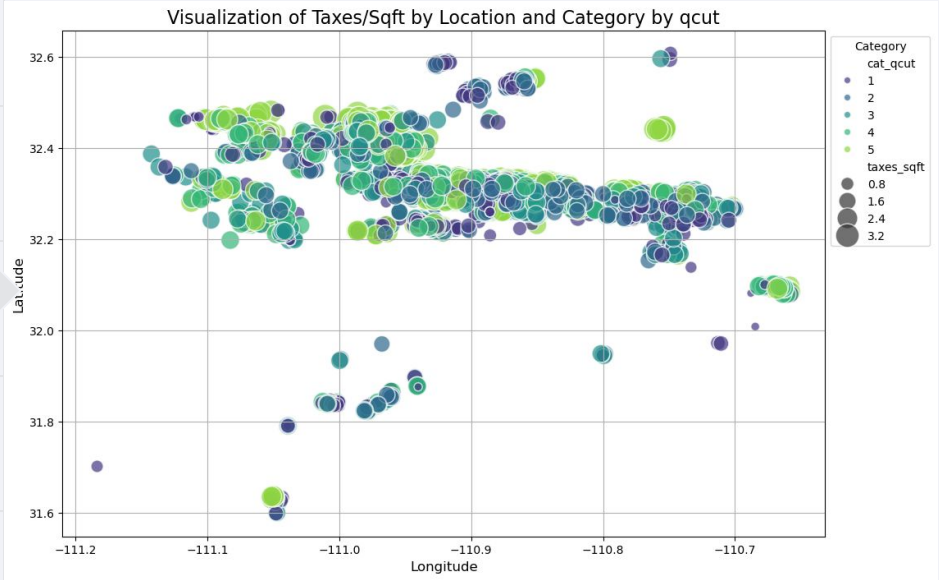
]:

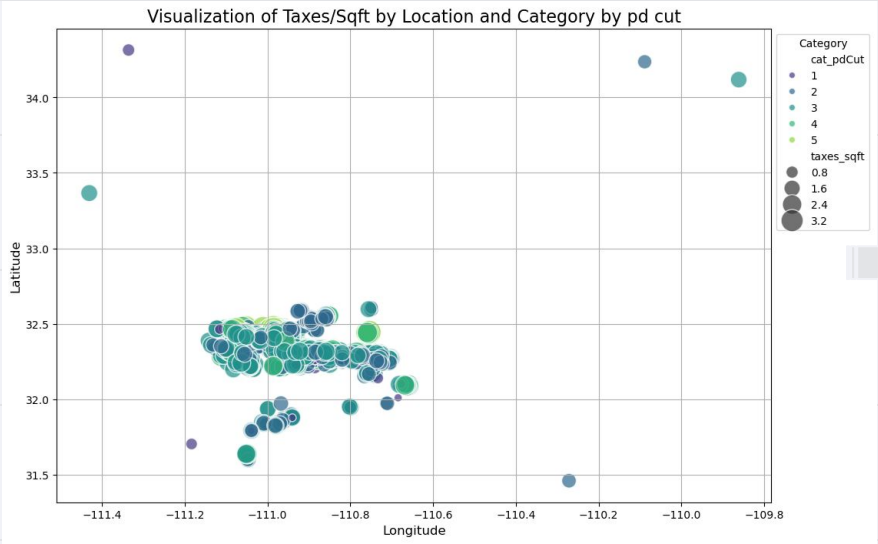
	taxes_sqft	price_zone	bedrooms	bathrooms	lot_acres	year_built
0	1.722875	13.124431	4	5.0	1.33	1986
1	1.696623	12.827988	4	4.0	1.17	1994
2	1.906972	13.760933	4	3.0	1.30	1993
3	2.817765	13.713572	4	5.0	1.23	2004
4	0.922989	13.124343	3	4.0	1.71	2017
...	...	...	...	...	...	...
3679	1.458667	6.132175	3	3.0	3.01	2007
3680	1.624141	6.588921	4	3.0	0.83	1986
3681	2.095916	6.241396	3	2.0	0.18	2002
3682	2.080246	6.413994	4	3.0	1.42	1990
3683	1.563622	6.414368	4	4.0	1.01	2009

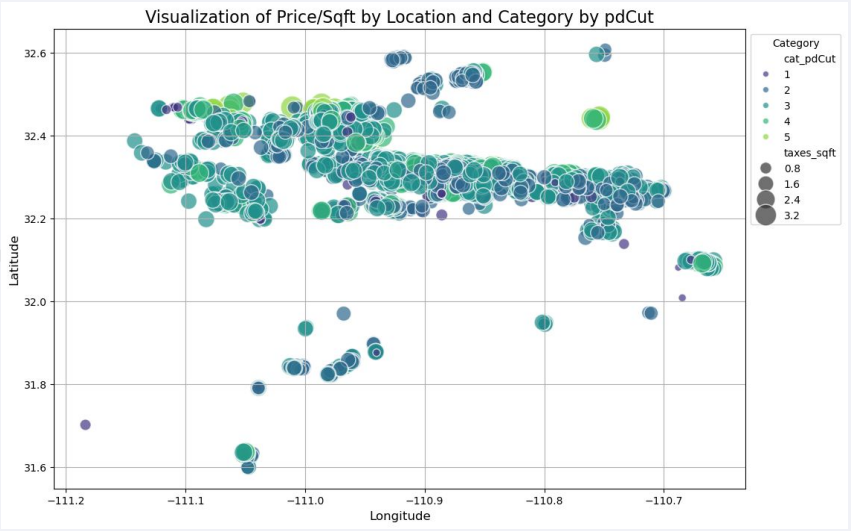
Data Visualization

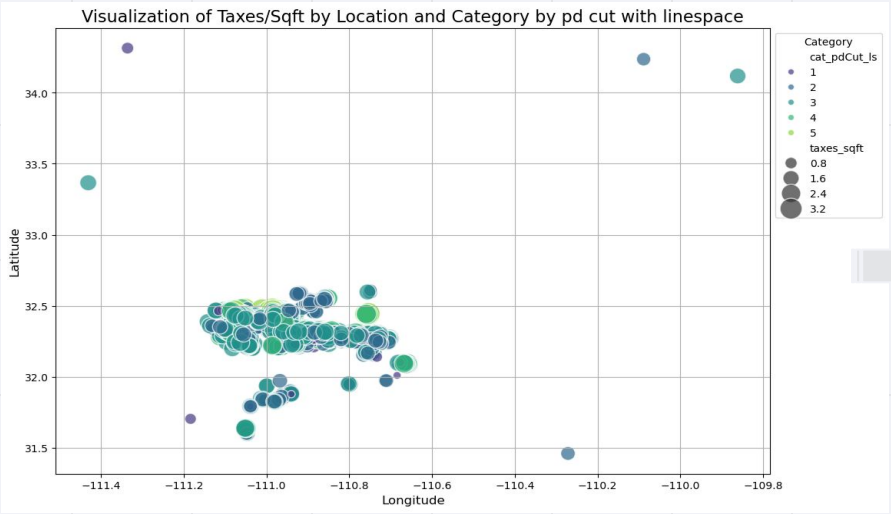
Filtering a DataFrame : using Mix max Method

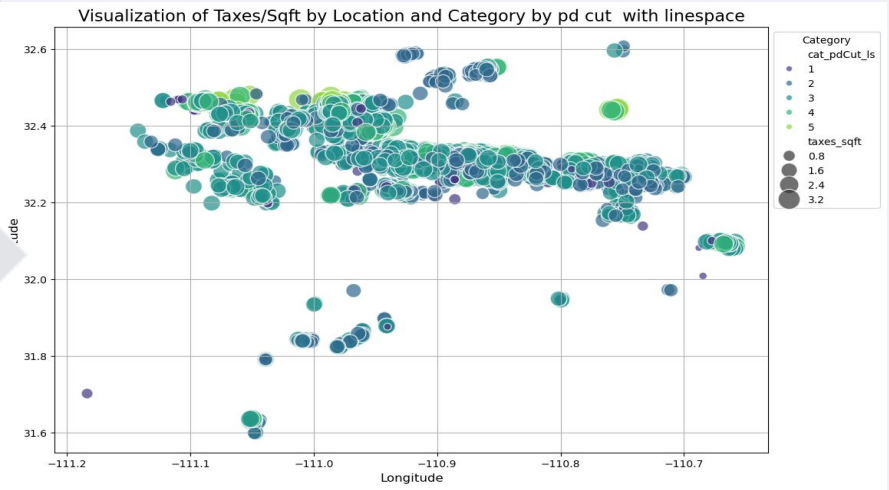





Visualization of Taxes/Sqft by Location and Category by pd cut
Category
cat_pdCut
1
2
3
4
5
taxes_sqft
0.8
1.6
2.4
3.2
Latitude
Longitude
34.0
33.5
33.0
32.5
32.0
31.5
−111.4
−111.2
−111.0
−110.8
−110.6
−110.4
−110.2
−110.0
−109.8


Visualization of Price/Sqft by Location and Category by pdCut
Category
cat_pdCut
1
2
3
4
5
taxes_sqft
0.8
1.6
2.4
3.2
Latitude
Longitude
32.6
32.4
32.2
32.0
31.8
31.6
−111.2
−111.1
−111.0
−110.9
−110.8
−110.7


Visualization of Taxes/Sqft by Location and Category by pd cut with linespace
Category
cat_pdCut_ls
1
2
3
4
5
taxes_sqft
0.8
1.6
2.4
3.2
Latitude
Longitude
34.0
33.5
33.0
32.5
32.0
31.5
−111.4
−111.2
−111.0
−110.8
−110.6
−110.4
−110.2
−110.0
−109.8

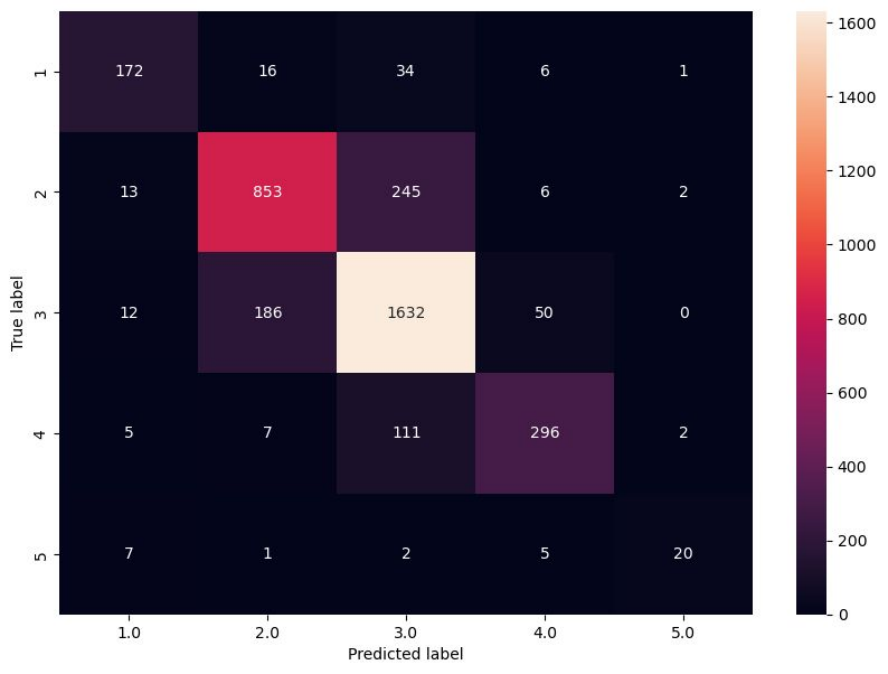

Visualization of Taxes/Sqft by Location and Category by pd cut with linespace
Category
cat_pdCut_ls
1
2
3
4
5
taxes_sqft
0.8
1.6
2.4
3.2
Longitude
32.6
32.4
32.2
32.0
31.8
31.6
−111.2
−111.1
−111.0
−110.9
−110.8
−110.7

KNN Regression Model

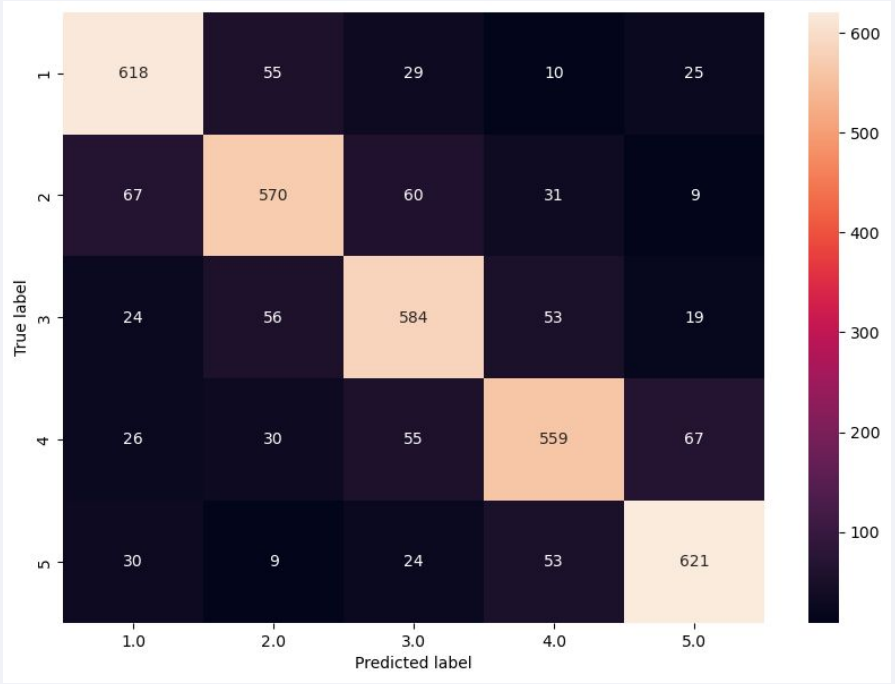
Accuracy	Cat by qcut	Cat by pdCut	Cat by pdCut linespace
dataframe	0.8013	0.8064	0.8064
Filtered dataframe	0.8010	0.8070	0.8070

Confusion Matrix

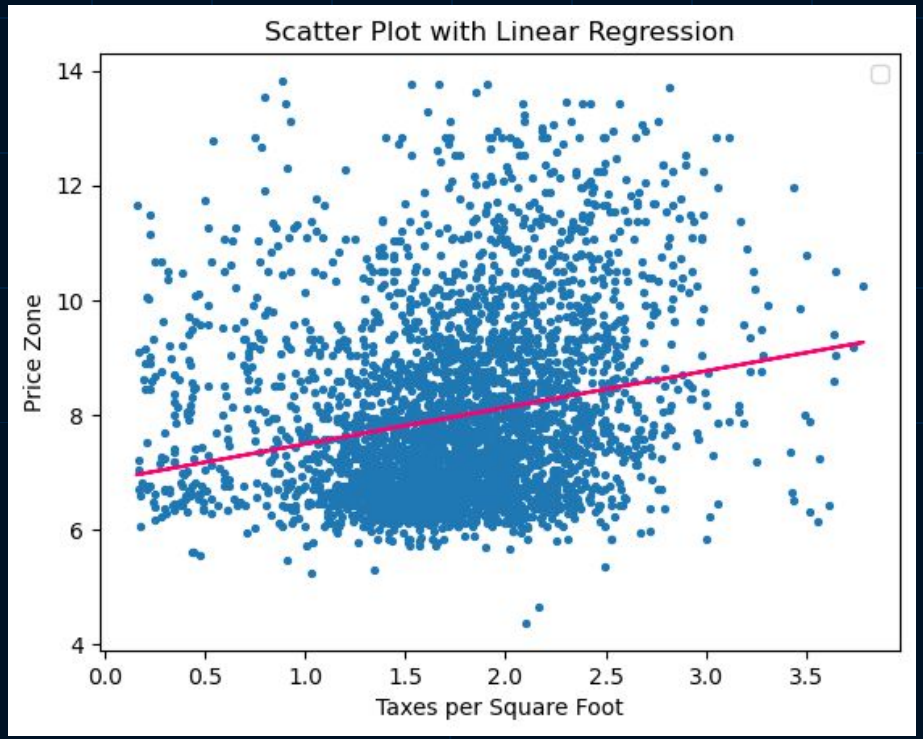
Using pd.cut method

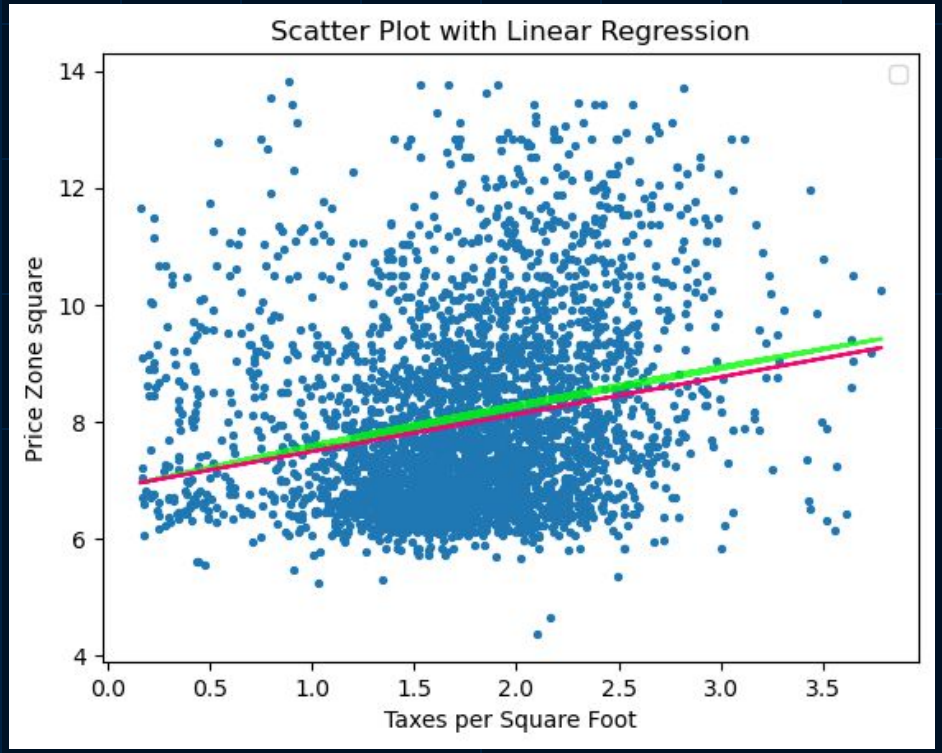


Using pd.qcut method



Simple Linear Regression



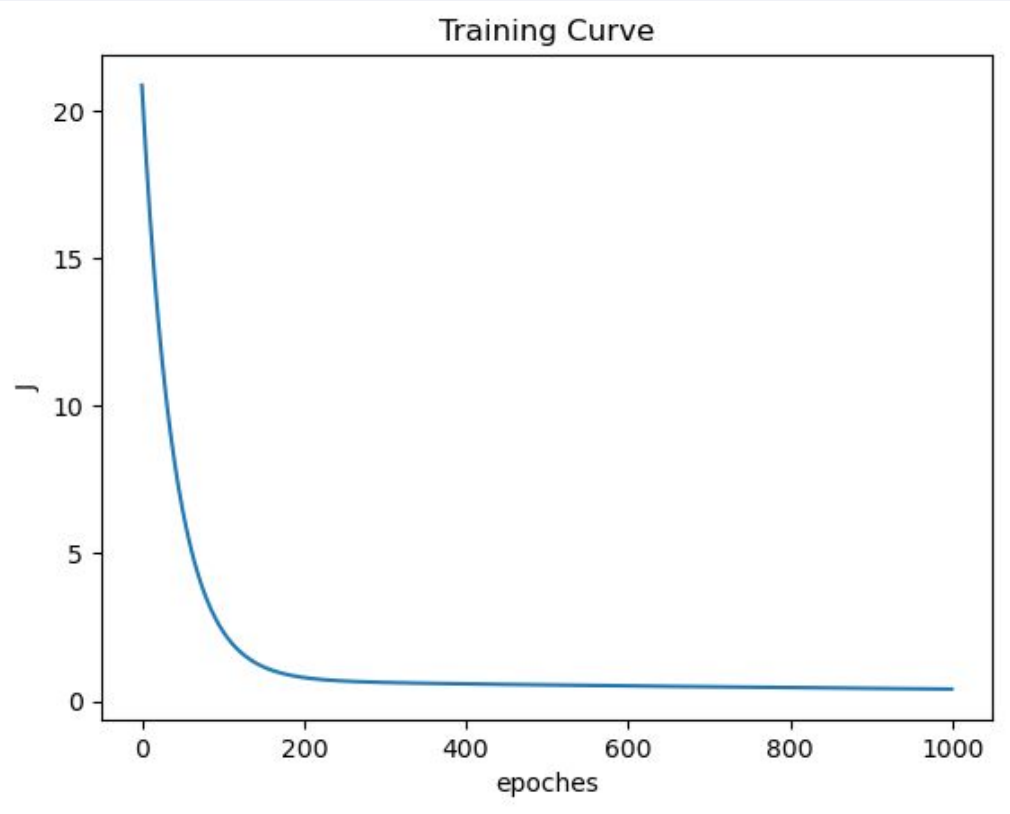


MV Regression Model

Dependent variable
X = price_zone bedrooms
bathrooms lot_acres
year_built

Independent variable
y = taxes_sqft

Model optimized with hyperparameters (eta=1e-8, epochs=1e3)



Result Analysis

Predicted Values
taxes_sqft
2.57164559
1.92747161
1.31464421
2.52677025
1.82934706

Actual Values

	taxes_sqft	price_zone	bedrooms	bathrooms	lot_acres	year_built
0	1.722875	13.124431	4	5.0	1.33	1986
1	1.696623	12.827988	4	4.0	1.17	1994
2	1.906972	13.760933	4	3.0	1.30	1993
3	2.817765	13.713572	4	5.0	1.23	2004
4	0.922989	13.124343	3	4.0	1.71	2017

Conclusion

Property taxes strongly depend on features like property value (sold_price, sq_ft, etc.) and location (zip, lat, long).

The categorization and regression modeling provide real estate professionals with actionable insights for accurate tax predictions

Thank You!