

# Property Tax Prediction for Real Estate Professionals

Vaishali Patelia

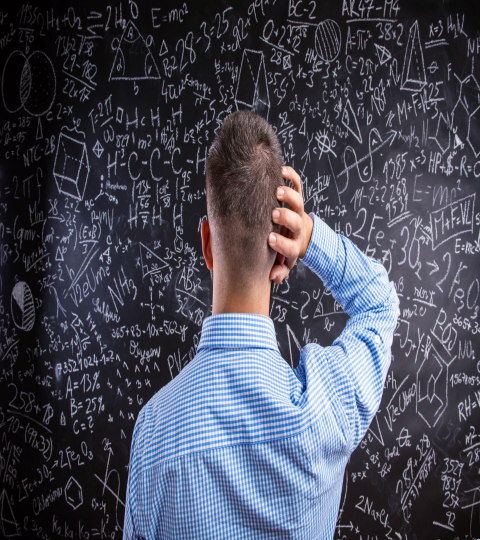
# Outline

2

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



# Introduction<sup>3</sup>



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)  
methods:

Pd.qcut

Pd.cut

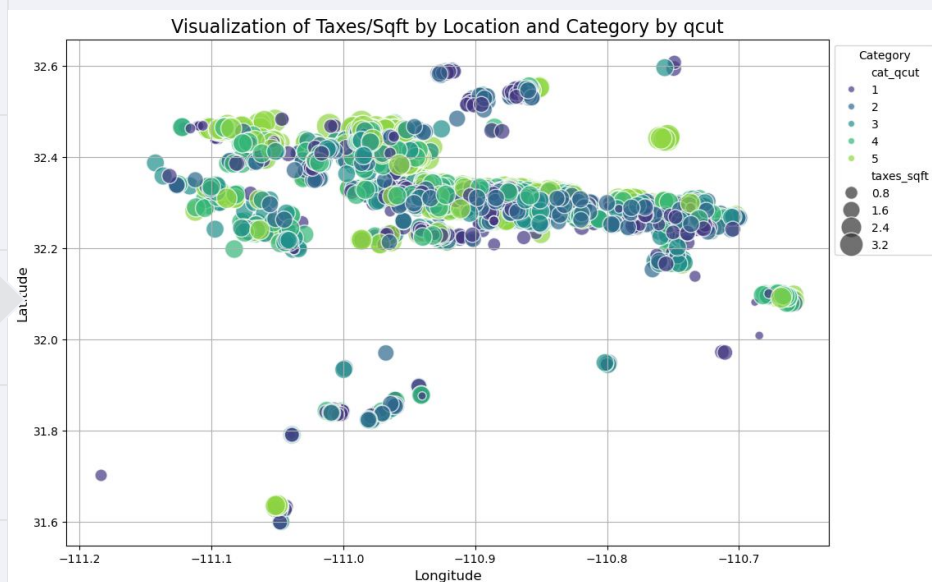
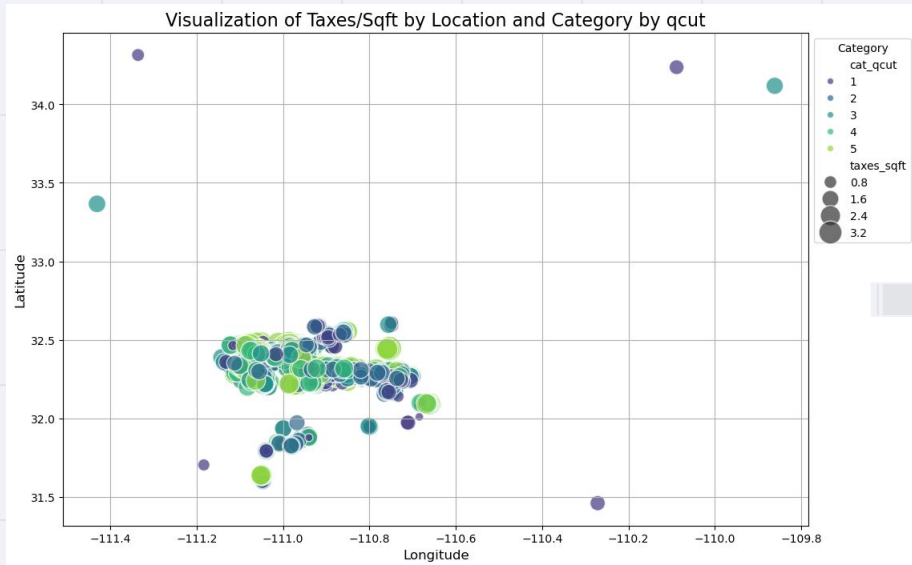
Pd.cut (linespace)

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

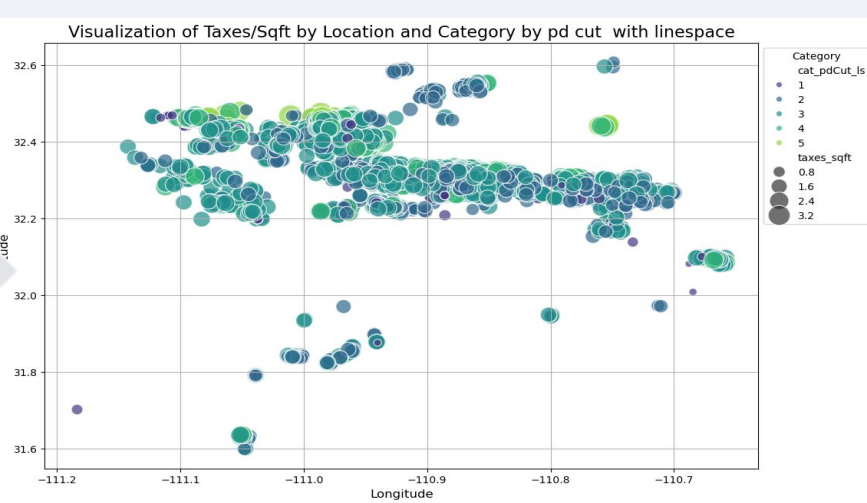
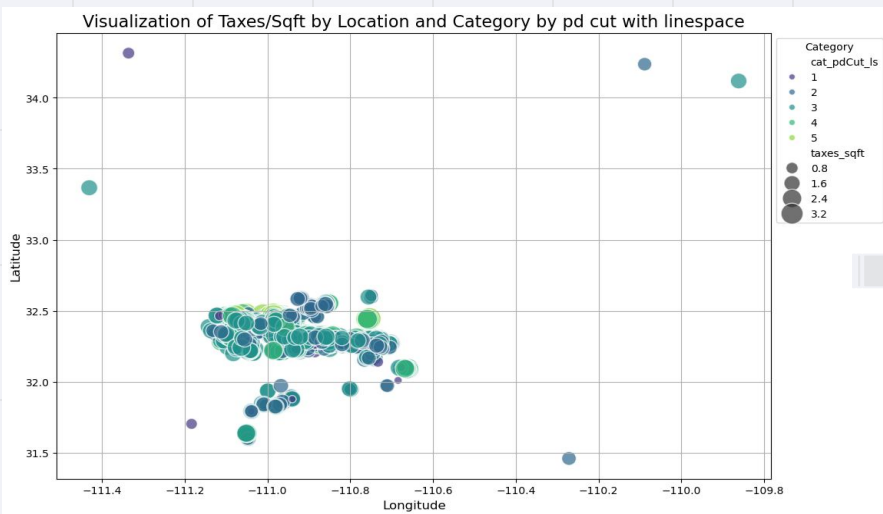
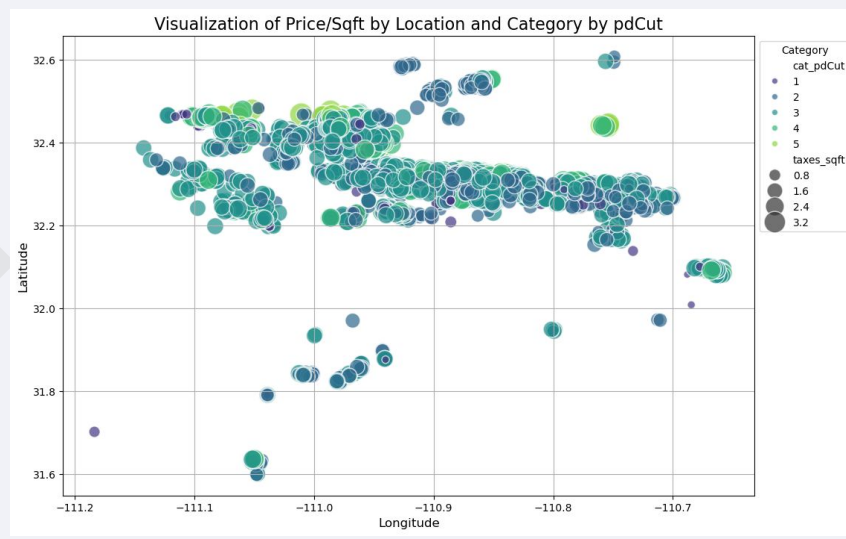
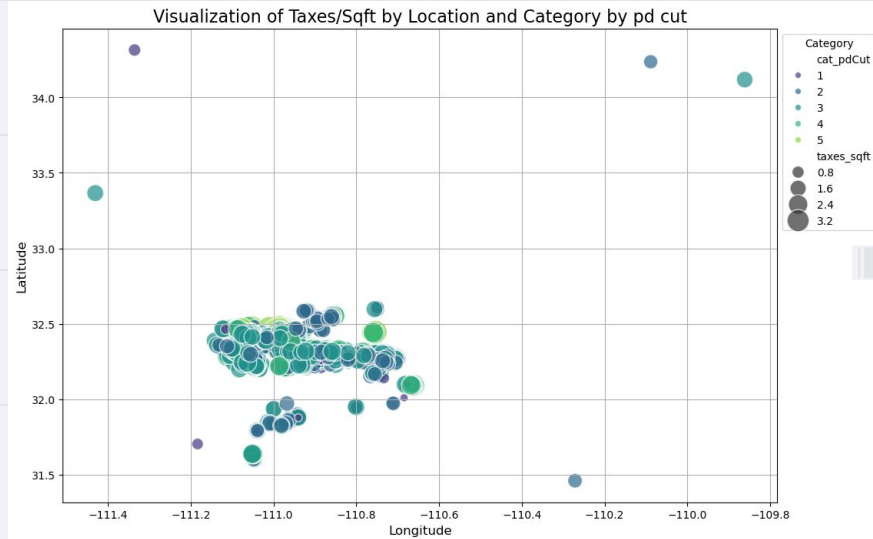
	taxes_sqft	price_zone	bedrooms	bathrooms	lot_acres	year_built
0	1.722875	13.124431	4	5.0	1.33	1986
1	1.696623	12.827988	4	4.0	1.17	1994
2	1.906972	13.760933	4	3.0	1.30	1993
3	2.817765	13.713572	4	5.0	1.23	2004
4	0.922989	13.124343	3	4.0	1.71	2017
...	...	...	...	...	...	...
3679	1.458667	6.132175	3	3.0	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







# KNN Regression Model

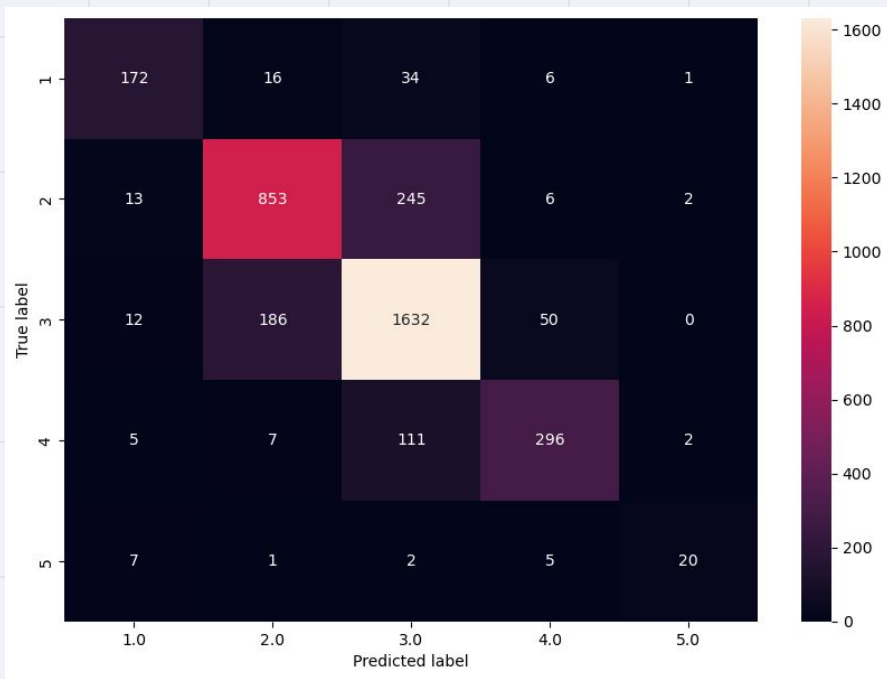
7

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

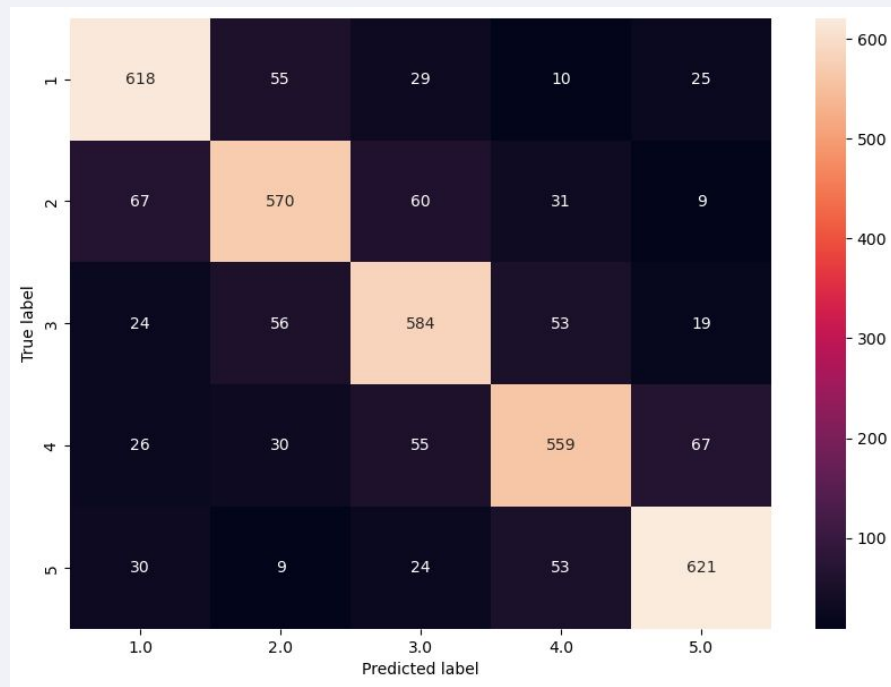
# Confusion Matrix

8

Using pd.cut method



Using pd.qcut method



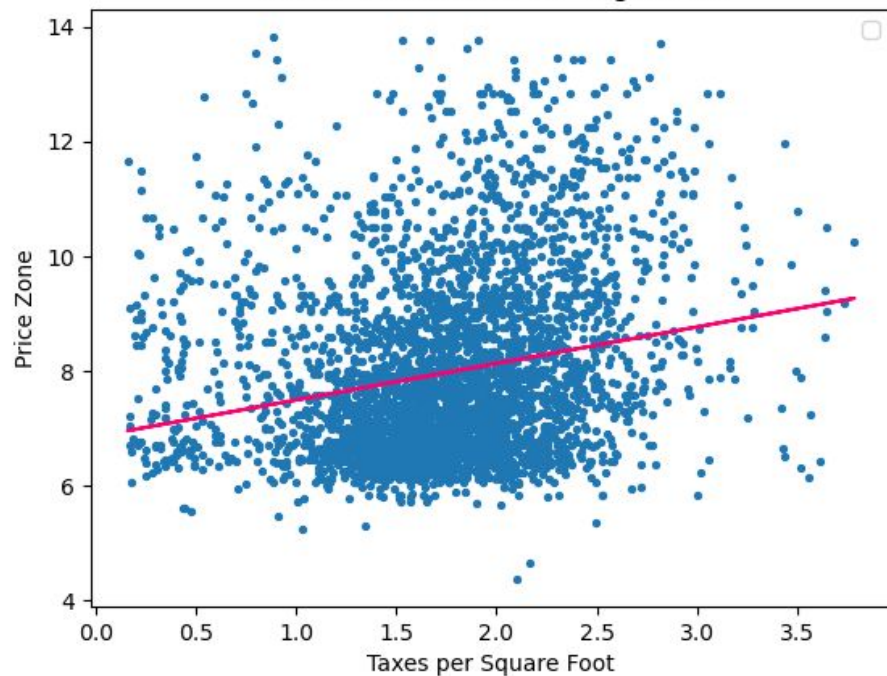
8



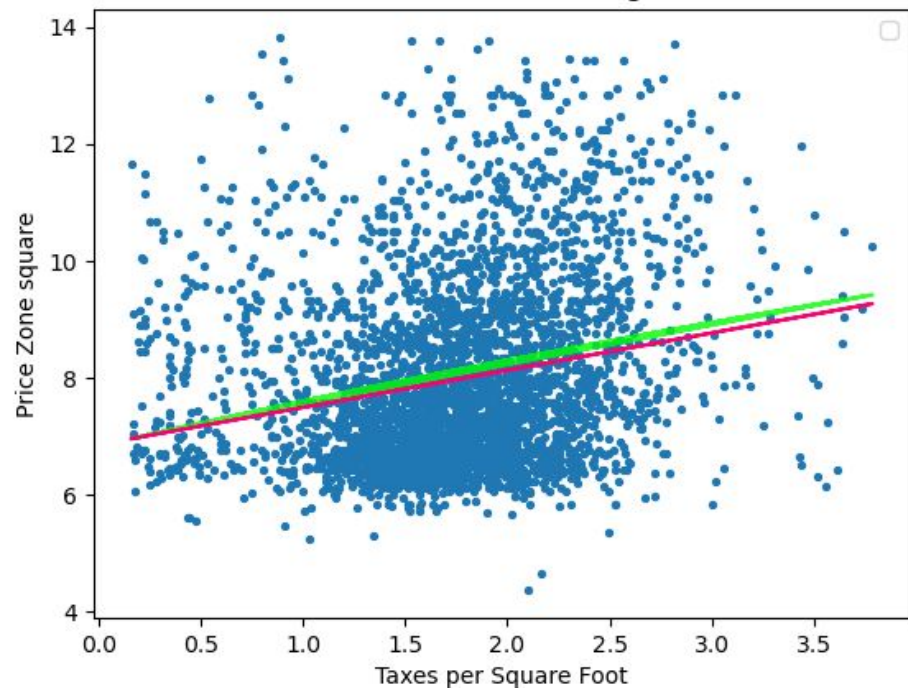
# Simple Linear Regression

9

Scatter Plot with Linear Regression



Scatter Plot with Linear Regression



# MV Regression Model

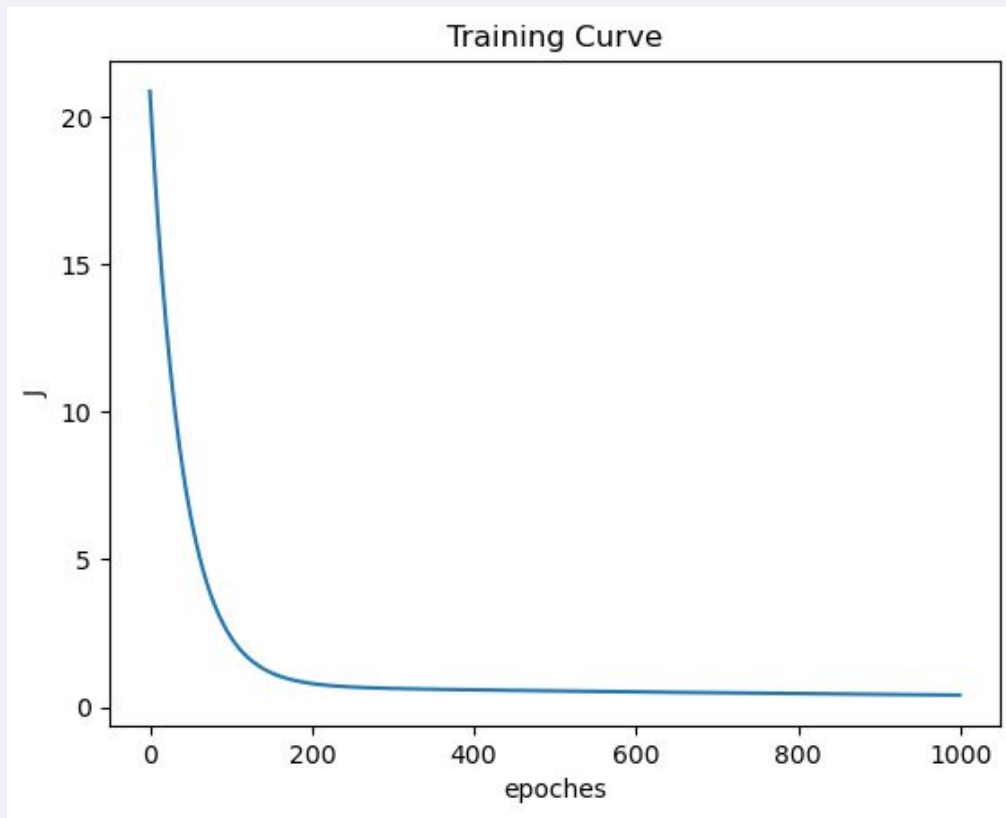
Dependent variable

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	Actual Values						
taxes_sqft		taxes_sqft	price_zone	bedrooms	bathrooms	lot_acres	year_built
2.57164559	0	1.722875	13.124431	4	5.0	1.33	1986
1.92747161	1	1.696623	12.827988	4	4.0	1.17	1994
1.31464421	2	1.906972	13.760933	4	3.0	1.30	1993
2.52677025	3	2.817765	13.713572	4	5.0	1.23	2004
1.82934706	4	0.922989	13.124343	3	4.0	1.71	2017

Predicted Tax = Predicted taxes\_sqft \* sqft

RESULTS OF SKLEARN MODEL TEST

MAE: 0.5738599599999998

MSE: 0.20307177062649126

RMSE: 0.450634852875908

R-Squared: -0.508975605126756

# Conclusion

Property taxes strongly depend on features like property value (**sold\_price**, **sq\_ft**) and location (**zip**, **lat**, **long**).

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



Thank  
You!

