Predict Selling Price

By Jacob Alamparambil

Overview

- Feature Selection
- Split dataset
- KNN vs Linear Regression
- Conclusion

Feature Selection(KNN)

Columns X:

- Latitude
- Longitude

Column Y:

- Price/Square Feet

Feature Selection(Linear Regression)

Columns X:

- Bathrooms
- Sqrt_ft
- Price/Sqrt_ft

Column Y:

- Sold_price

Linear Regression results

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content
                            OLS Regression Results
==============================================================================
Dep. Variable:             sold_price   R-squared:                       0.932
Model:                            OLS   Adj. R-squared:                  0.885
Method:                 Least Squares   F-statistic:                     19.84
Date:                Mon, 17 Nov 2025   Prob (F-statistic):               0.00
Time:                        21:46:14   Log-Likelihood:                -50970.
No. Observations:                4000   AIC:                         1.052e+05
Df Residuals:                    2366   BIC:                         1.155e+05
Df Model:                        1633
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

KNN results

••• Drive already mounted at /content/drive; to attempt to forcibly remount, call dr
74.52880952380953