

# Categorizing Sold\_Price

By Jacob Alamparambil

# Problem

1. Inconsistency
2. Lack of Segmentation
3. Inefficiency
4. Missed Insights

# Why Categorization Matters

- Creates a shared language across teams
- Enables market trend analysis
- Supports pricing, planning, and forecasting
- Simplifies communication internally & externally

# Steps

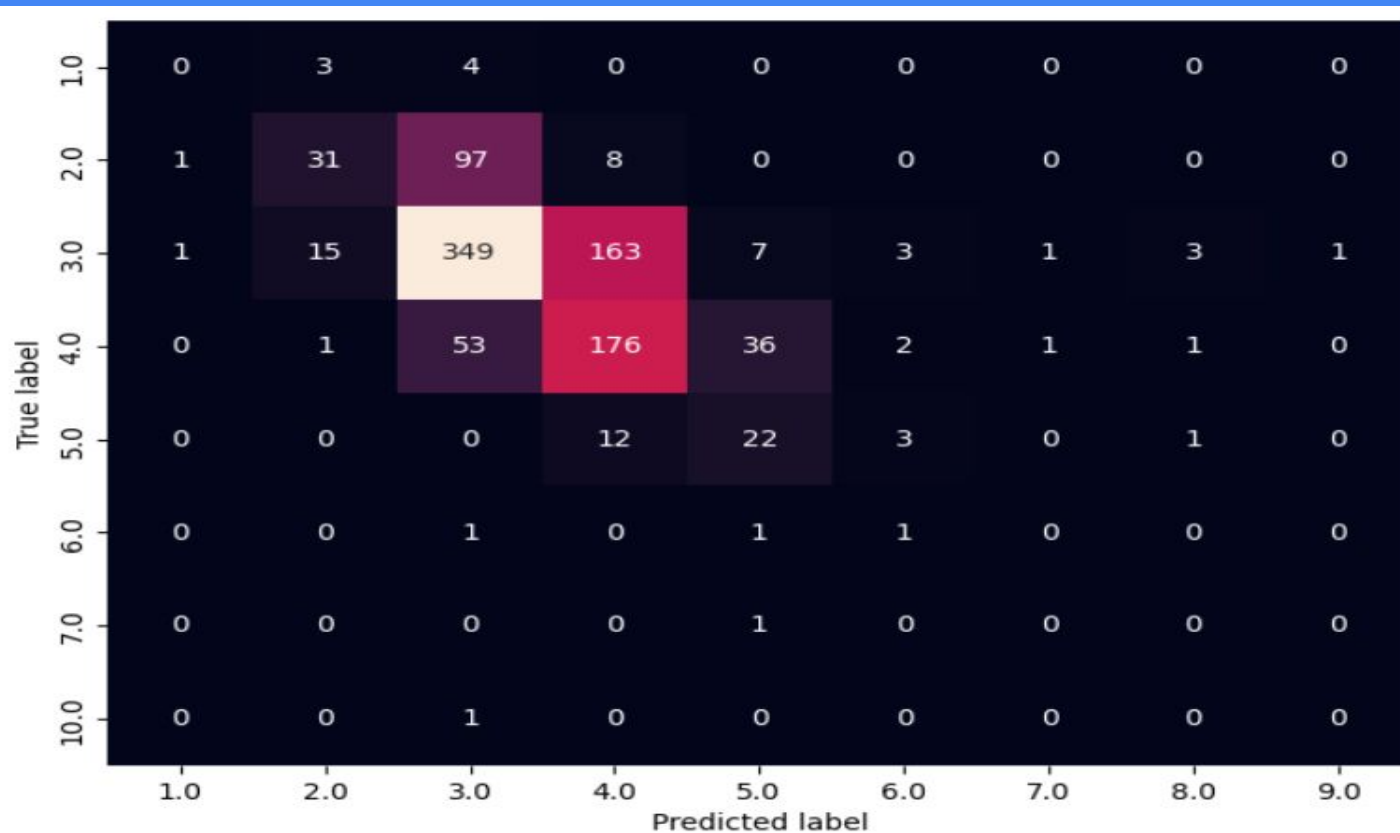
- Preprocessing cleaned data
- Feature Selection
- Normalization
- Split dataset
- Run model

# Sample categorization

- Divide price/sqft into 24 categories

```
Drive already mounted a
count      5000.00000
mean       3.75460
std        1.38606
min        0.00000
25%        3.00000
50%        4.00000
75%        4.00000
max        24.00000
```

# Actual vs Predicted Comparison



# Results

```
Name: price/sqrft, Length: 5000, dtype: float64  
Model Accuracy: 57.90%
```

# Conclusion

- Consistent categorization unlocks insights
- Works with rule based logic
- Supports every team across the business
- Foundation for smarter real estate analytics