



Linear Regression Model

**Real Estate
Price Predictor**

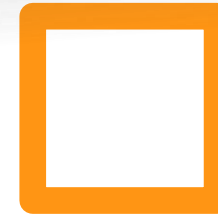


Contents

- ☐ **Dataset**
- ☐ **Feature Selection**
- ☐ **Linear Regression Model**
- ☐ **Onward Deployment**

Dataset

- ❑ Actual transaction prices of 66,000 private residential properties in Singapore from 2017 to 2020 from the Singapore Urban Redevelopment Authority (URA)
- ❑ Dataset contains about 20 dependent parameters such as residential project name, address, locality, planning region, tenure, property type, completion date, etc



Feature Selection

Reduced set of four dependent parameters most likely to impact transaction prices, are selected for ease of onward deployment of model to the Maya Properties chatbot

- ❑ **Locality:** Comprises about 30 districts in Singapore such as Orchard, Tampines, Changi etc
- ❑ **Property Type:** Bungalow, apartment, condominium, terrace, & semi-detached
- ❑ **Area Type:** Land titled or Strata
- ❑ **Tenure:** Freehold or Leasehold



Sample Data Set

```
df1.head()
```

	UnitPrice	PlanArea	Tenure	PropertyType	AreaType
0	1,512	Hougang	Freehold	Condominium	Strata
1	1,047	Yishun	Leasehold	Condominium	Strata
2	2,013	Bedok	Leasehold	Apartment	Strata
3	1,630	Toa Payoh	Leasehold	Condominium	Strata
4	1,306	Clementi	Leasehold	Condominium	Strata

Train Linear Regression Model

```
import sklearn
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size =0.2, random_state=7)
```

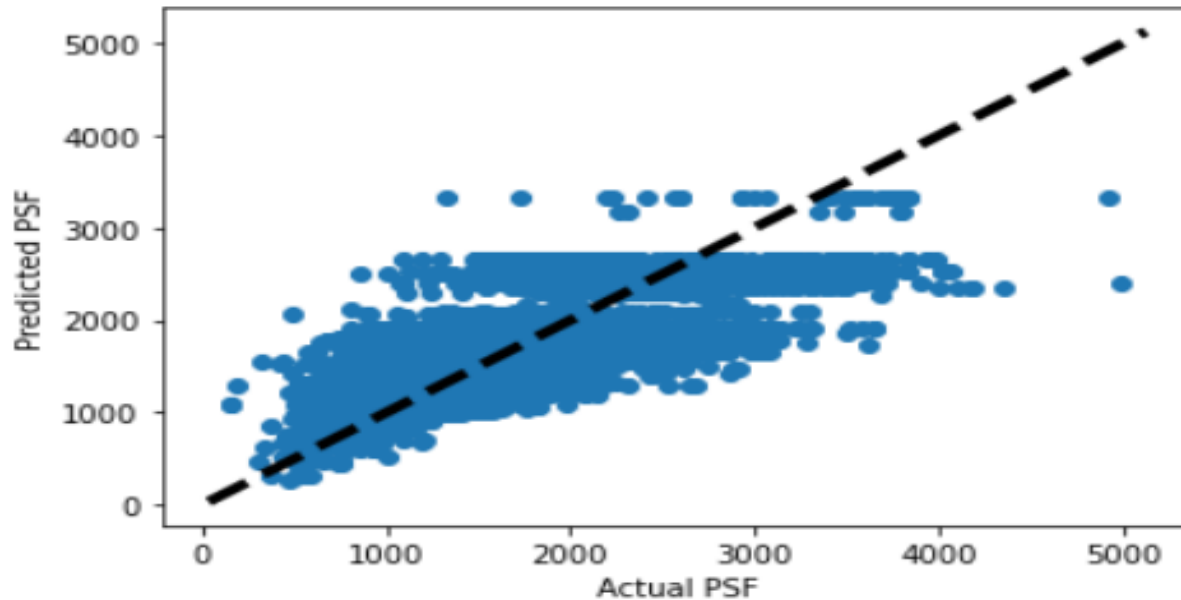
```
model = LinearRegression()
model.fit(X_train, y_train)
y_predict = model.predict(X_test)
```



Model Evaluation

```
# importing r2_score module
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error
# predicting the accuracy score
score=r2_score(y_test,y_predict)
score
```

0.6051216417932486



**R-square Coefficient
0.61**

Predicted vs Actual Prices

y=x line represents perfect fit

Onward Deployment

Model coefficients and Y-intercept output as JSON file for onward deployment to Maya Properties chatbot coded in Javascript in Dialogflow chatbot development

Save as JSON File (Default file name is data.json)

```
# Save using sklearn_export  
from sklearn_export import Export  
export = Export(model)  
result = export.to_json()
```

