

▼ PROJECT 1

Step 1: Understanding business problem

Problem Statement: A person should develop a software where we enter area in sq.ft and it should return price

```
#Python libraries for Data Analysis
# Pandas for Analysis
#Matplotlib for Visualization
# Scikitlrean for ML
```

▼ Data Collection and Analysis

```
import pandas as pd

df = pd.read_csv('/content/Metupalayam -P1 Dataset.csv')
print(df)

   area  price
0  1000  800000
1  1200  964000
2  1500 1200000
3  1400 1100000
4  2000 1600000

print(type(df))

<class 'pandas.core.frame.DataFrame'>
```

▼ Data Visualization

```
from matplotlib import pyplot as plt

df['area']

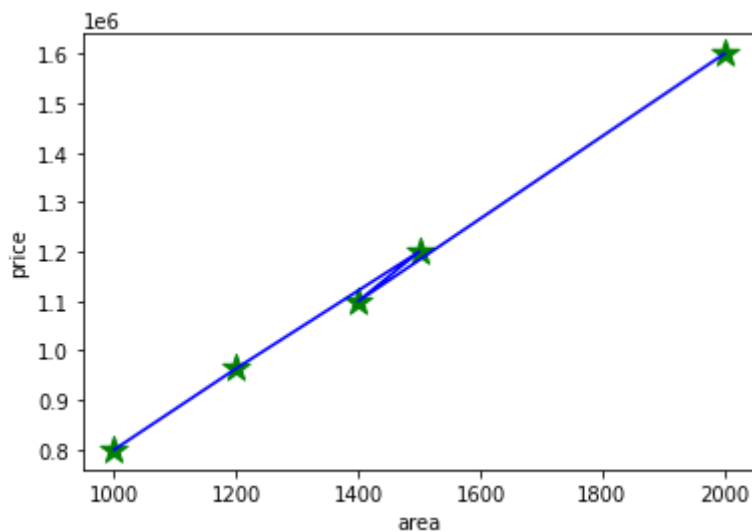
0    1000
1    1200
2    1500
3    1400
4    2000
Name: area, dtype: int64
```

```
X = df.drop('price',axis=1)
print(X)
y = df['price']
print(y)
```

```
      area
0    1000
1    1200
2    1500
3    1400
4    2000
0      800000
1      964000
2     1200000
3     1100000
4     1600000
Name: price, dtype: int64
```

```
plt.xlabel("area")
plt.ylabel("price")
plt.scatter(X,y,color='green',marker='*',s=200)
plt.plot(X,y,color="blue")
```

[<matplotlib.lines.Line2D at 0x7f896367cd50>]



▼ Build ML Model

```
from sklearn import linear_model
#importing library scikitlearn and linear_model algorithm
```

```
#Creating a reference object to that algorithm
reg = linear_model.LinearRegression()
```

```
#Train the Machine -ML
```

```
# using FIT - helps to learn the machine
reg.fit(X,y)
```

```
LinearRegression()
```

```
#Predict the output or test the ML
```

```
reg.predict([[1000]])
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/base.py:451: UserWarning: X does not
  "X does not have valid feature names, but"
array([797154.92957746])
```



```
reg.score(X,y)
```

```
0.9989964816659621
```

▼ How ML works Just to know how it works

Linear Regression Formula

$y = mX + b$

```
m = reg.coef_
m
```

```
array([799.15492958])
```

```
b = reg.intercept_
b
```

```
-2000.0000000002328
```

```
inX = 1000
```

```
y = m*inX+b
y
```

```
array([797154.92957746])
```

▼ Download the Model

dw reg object contains coefficient and intercept value

To dw variables use Pickle library

```
import pickle
```

going to dw ML model that file is python obj as Binary format. Syntax : with
`open('filename',wb)as f:pickle.dump(reg,f)`

```
with open('Metupalayammodel','wb') as f:  
    pickle.dump(reg,f)  
    #Check folder and find modelfile
```

✓ 0s completed at 7:57 PM

