```
import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt
```

```
from google.colab import files
uploaded = files.upload()
```

Choose Files housing_area_price.csv

- **housing_area_price.csv**(text/csv) - 77 bytes, last modified: 3/17/2025 - 100% done

```
Saving housing_area_price.csv to housing_area_price.csv
```

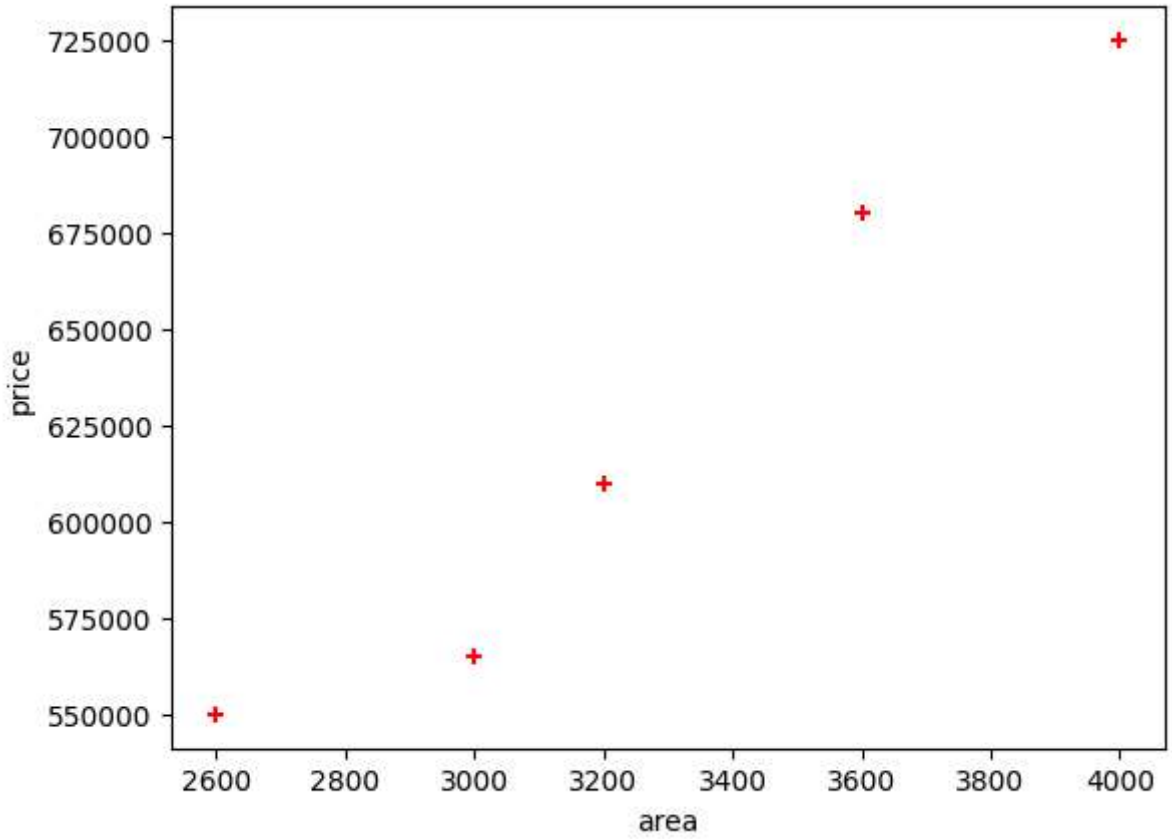
```
import pandas as pd

df = pd.read_csv('housing_area_price.csv')
df.head()
```

| | area | price |
|---|---|---|
| 0 | 2600 | 550000 |
| 1 | 3000 | 565000 |
| 2 | 3200 | 610000 |
| 3 | 3600 | 680000 |
| 4 | 4000 | 725000 |

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red',marker='+')
```

```
<matplotlib.collections.PathCollection at 0x79e2468ba810>
```




```
new_df = df.drop('price',axis='columns')
new_df

price = df.price
price
```

| | price |
|---|---|
| **0** | 550000 |
| **1** | 565000 |
| **2** | 610000 |
| **3** | 680000 |
| **4** | 725000 |

**dtype:** int64

```
reg = linear_model.LinearRegression()
reg.fit(new_df,price)
```

```
reg.predict([[3300]])

reg.coef_

reg.intercept_
```

```
/usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X
  warnings.warn(
180616.43835616432
```

```
reg = linear_model.LinearRegression()
reg.fit(new_df,price)
reg.predict([[3300]])

reg.coef_

reg.intercept_
```

```
/usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X
  warnings.warn(
180616.43835616432
```

```
"""Y = m * X + b (m is coefficient and b is intercept)"""

3300*135.78767123 + 180616.43835616432
```

```
628715.7534151643
```

```
"""(1) Predict price of a home with area = 5000 sqr ft"""

reg.predict([[5000]])
```

```
/usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X
  warnings.warn(
array([859554.79452055])
```