

Home Page - Select or create a notebook | LinearRegression - Jupyter Notebook | +

localhost:8888/notebooks/LinearRegression.ipynb

jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

In [1]: `import pandas as pd
df=pd.read_csv("ML0.csv")`

In [2]: `df`

Out[2]:

	Area	Price
0	2600	550000
1	3000	565000
2	3200	610000
3	3600	680000
4	4000	725000

In [4]: `import matplotlib.pyplot as plt
from sklearn import linear_model
%matplotlib inline
plt.xlabel("Area")
plt.ylabel("Price")

plt.scatter(df.Area,df.Price,marker="*")`

Out[4]: `<matplotlib.collections.PathCollection at 0x1343b2f8>`

725000
700000
675000

Type here to search

11:36 AM 9/22/2020

Home Page - Select or create a notebook x LinearRegression - Jupyter Notebook x +

localhost:8888/notebooks/LinearRegression.ipynb

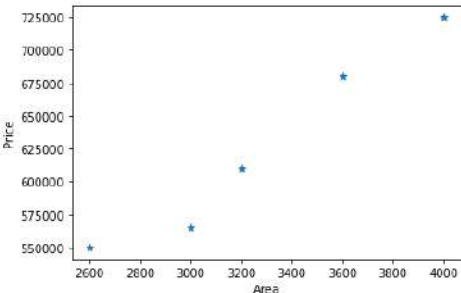
jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

In [4]:

```
import matplotlib.pyplot as plt
from sklearn import linear_model
%matplotlib inline
plt.xlabel("Area")
plt.ylabel("Price")
plt.scatter(df.Area, df.Price, marker="*")
```

Out[4]: <matplotlib.collections.PathCollection at 0x1343b2f8>



A scatter plot showing the relationship between Area (x-axis) and Price (y-axis). The x-axis ranges from 2600 to 4000 with major ticks every 200 units. The y-axis ranges from 550,000 to 725,000 with major ticks every 25,000 units. There are five data points plotted as blue asterisks. The points show a positive correlation between Area and Price.

Area	Price
2600	550000
3000	565000
3200	610000
3600	680000
4000	720000

In [5]:

```
reg=linear_model.LinearRegression()
reg.fit(df[["Area"]], df.Price)
```

Out[5]: LinearRegression()

Type here to search

11:39 AM 9/22/2020

Home Page - Select or create a notebook | LinearRegression - Jupyter Notebook | +

localhost:8888/notebooks/LinearRegression.ipynb

jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

In [15]: `reg.predict([[3300]])`
Out[15]: `array([628715.75342456])`


In [16]: `reg.coef_`
Out[16]: `array([135.78767123])`

In [17]: `reg.intercept_`
Out[17]: `180616.43835616432`

In [18]: `reg.predict([[4000]])`
Out[18]: `array([723767.12328767])`

In [24]: `import matplotlib.pyplot as plt`
`from sklearn import linear_model`
`%matplotlib inline`
`plt.xlabel("Area")`
`plt.ylabel("Price")`

`plt.scatter(df.Area, df.Price, marker="*")`
`plt.plot(df.Area, reg.predict(df[["Area"]]))`
Out[24]: `[<matplotlib.lines.Line2D at 0x15914a8>]`



Windows taskbar: Type here to search | 11:39 AM 9/22/2020

Home Page - Select or create a notebook | LinearRegression - Jupyter Notebook

localhost:8888/notebooks/LinearRegression.ipynb

Jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

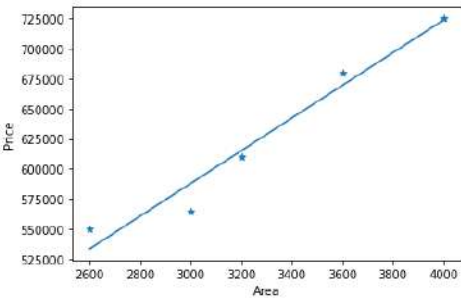
File Edit View Insert Cell Kernel Help Trusted Python 3

In [24]:

```
import matplotlib.pyplot as plt
from sklearn import linear_model
%matplotlib inline
plt.xlabel("Area")
plt.ylabel("Price")

plt.scatter(df.Area, df.Price, marker="*")
plt.plot(df.Area, reg.predict(df[["Area"]]))
```

Out[24]: [matplotlib.lines.Line2D at 0x15914a8]



Area	Price
2600	540000
3000	560000
3200	610000
3600	680000
4000	720000

In [26]: d=pd.read_csv("ML1.csv")

In [27]: d

Type here to search

11:36 AM 9/22/2020

Home Page - Select or create a notebook | LinearRegression - Jupyter Notebook | +

localhost:8888/notebooks/LinearRegression.ipynb

jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

In [29]: `p=reg.predict(d[["Area"]])`

In [30]: `p`

Out[30]: `array([452191.78082192, 494557.53424658, 643923.97260274, 657095.37671233,
683030.82191781, 723767.12328767])`

In [31]: `d["Price"]=p`

In [32]: `d`

Out[32]:

	Area	Price
0	2000	452191.780822
1	2312	494557.534247
2	3412	643923.972603
3	3509	657095.376712
4	3700	683030.821918
5	4000	723767.123288

In [34]: `d.to_csv("ML2.csv",index=False)`

In [35]: `d`

Out[35]:

	Area	Price
--	------	-------

Type here to search

11:37 AM 9/22/2020

Home Page - Select or create a notebook | LinearRegression - Jupyter Notebook | +

localhost:8888/notebooks/LinearRegression.ipynb

jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

In [39]: `ex=pd.read_csv("ML3.csv")`

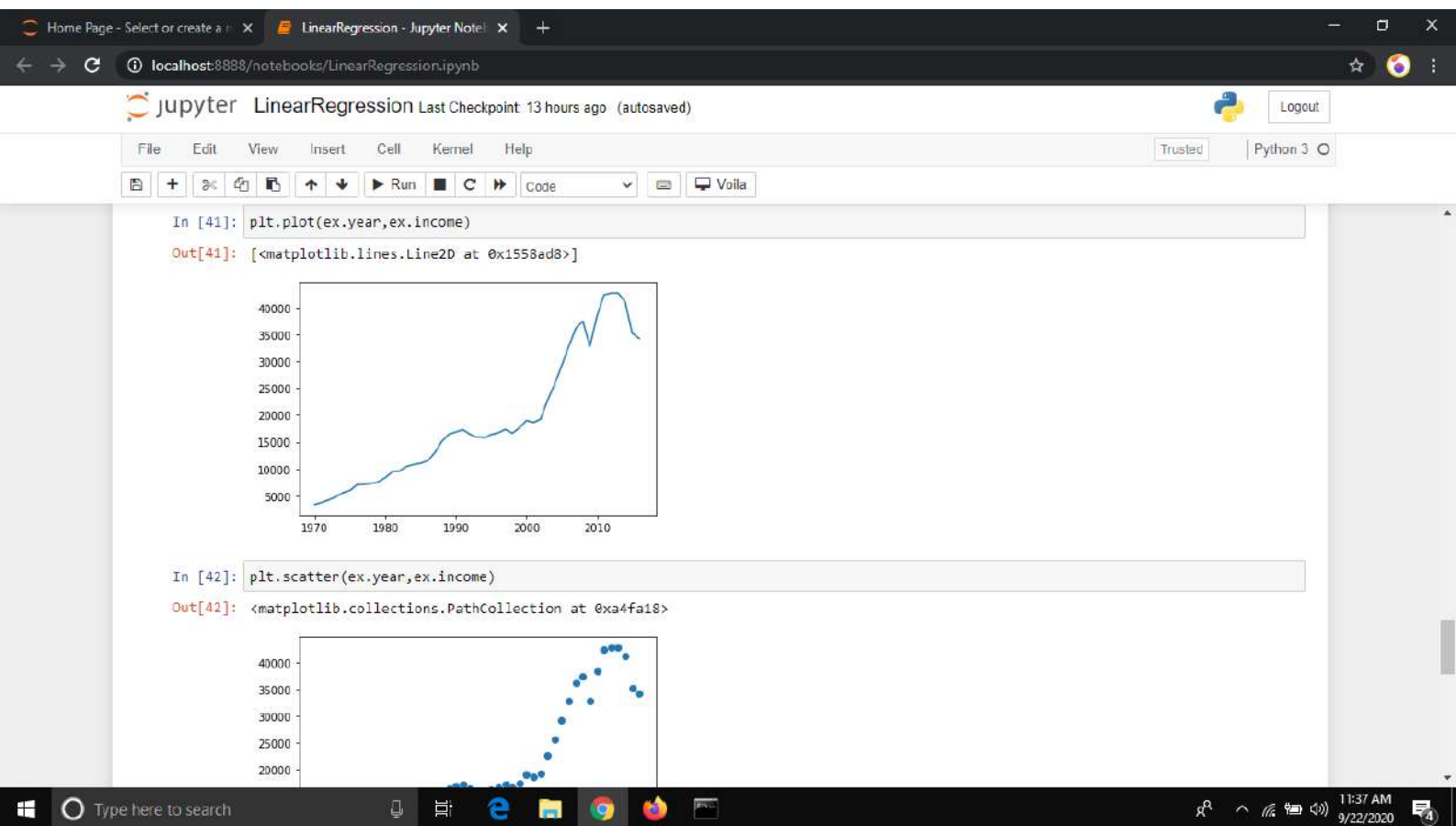
In [40]: `ex`

Out[40]:

	year	income
0	1970	3399.299037
1	1971	3768.297935
2	1972	4251.175484
3	1973	4804.463248
4	1974	5576.514583
5	1975	5998.144346
6	1976	7062.131392
7	1977	7100.126170
8	1978	7247.967035
9	1979	7602.912681
10	1980	8355.968120
11	1981	9434.390652
12	1982	9619.438377
13	1983	10416.536590
14	1984	10790.328720
15	1985	11018.955850
16	1986	11100.001520

Type here to search

11:37 AM 9/22/2020



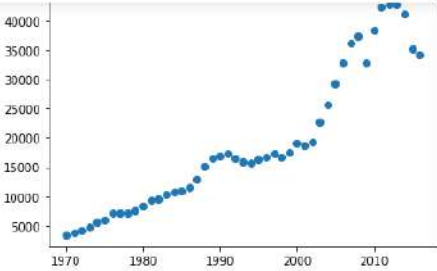
Home Page - Select or create a notebook | LinearRegression - Jupyter Notebook | +

localhost:8888/notebooks/LinearRegression.ipynb

jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

Run Code Voila



```
In [44]: rr=linear_model.LinearRegression()

In [46]: rr.fit(ex[["year"]],ex[["income"]])

Out[46]: LinearRegression()

In [47]: rr.predict([[2020]])

Out[47]: array([41288.69409442])

In [49]: plt.xlabel("Year")
plt.ylabel("Per Capita Income (US$)")
plt.scatter(ex.year,ex.income)
plt.plot(ex.year,rr.predict(ex[["year"]]))

Out[49]: [ <matplotlib.lines.Line2D at 0x4d76628>]
```

Type here to search

11:37 AM 9/22/2020

Home Page - Select or create a notebook

LinearRegression - Jupyter Notebook

+

localhost:8888/notebooks/LinearRegression.ipynb

jupyter LinearRegression Last Checkpoint: 13 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

+

%

+

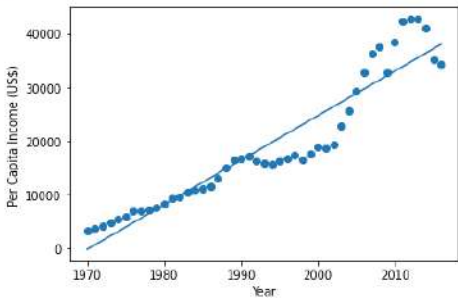
+

Run

Code

Voila

Out[49]: [matplotlib.lines.Line2D at 0x4d75628]



A scatter plot showing Per Capita Income (US\$) on the y-axis (ranging from 0 to 40,000) against Year on the x-axis (ranging from 1970 to 2010). The data points are blue dots, and a solid blue line represents the linear regression fit. The income shows a general upward trend with some fluctuations, particularly a sharp increase around 2000.

Year	Per Capita Income (US\$)
1970	~3000
1980	~8000
1990	~15000
2000	~18000
2010	~35000

In [50]: rr.predict([[2020]])

Out[50]: array([41288.69409442])

In []:

Type here to search

11:37 AM 9/22/2020