```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import linear_model
from google.colab import files
uploaded = files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving canada_per_capita_income.csv to canada_per_capita_income.csv
```

```
import warnings
warnings.filterwarnings("ignore")
```

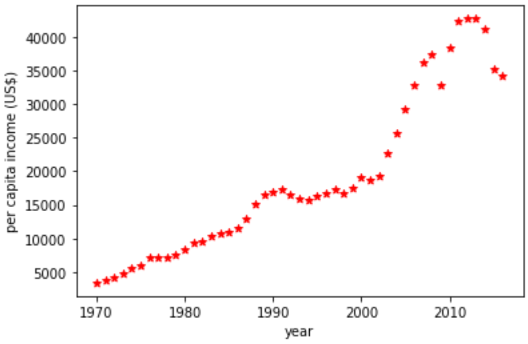
```
df=pd.read_csv("canada_per_capita_income.csv")
```

```
df
```

	year	per capita income (US$)
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

```
%matplotlib inline
plt.xlabel("year")
plt.ylabel("per capita income (US$)")
plt.scatter(df["year"],df["per capita income (US$)"],color="red",marker="*")

plt.show()
```




```
lr = linear_model.LinearRegression()
lr.fit(df[["year"]],df[["per capita income (US$)"]])
```

```
LinearRegression()
```

23 1993 15875.586730

```
lr.predict([[2020]])
```

```
array([[41288.69409442]])
```

y=mx+b

m is the slope or gradient

b is the y-intercept

price= m*area+b

price is dependent variable

area is independent variable

```
lr.coef_
#slope
```

```
array([[828.46507522]])
```

```
lr.intercept_
#y-intercept
```

```
array([-1632210.75785546])
```

```
828.46507522*2020+-1632210.75785546
```

```
41288.69408894004
```

```
data=[(2020),(2021),(2022),(2023),(2024),(2025),(2026),(2027),(2028),(2029),(2030),(2031),(2032),(2033)]
d=pd.DataFrame(data=data,columns=["year"])
d
```

	year
0	2020
1	2021
2	2022
3	2023
4	2024
5	2025
6	2026
7	2027
8	2028

```
lr.predict(d)
```

```
array([[41288.69409442],
       [42117.15916964],
       [42945.62424486],
       [43774.08932009],
       [44602.55439531],
       [45431.01947053],
       [46259.48454575],
       [47087.94962098],
       [47916.4146962 ],
       [48744.87977142],
       [49573.34484664],
       [50401.80992187],
       [51230.27499709],
       [52058.74007231]])
```

```
d.rename(columns={"predict price":"predict price (in US$)"},inplace=True)
```

```
d
```

	year	predict price (in US$)
0	2020	41288.694094
1	2021	42117.159170
2	2022	42945.624245
3	2023	43774.089320
4	2024	44602.554395
5	2025	45431.019471
6	2026	46259.484546
7	2027	47087.949621
8	2028	47916.414696
9	2029	48744.879771
10	2030	49573.344847
11	2031	50401.809922
12	2032	51230.274997
13	2033	52058.740072

```
df
```

	year	per capita income (US$)
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	11482.891530
17	1987	12974.806620

```
d.to_csv("canada future house price prediction.csv")
```

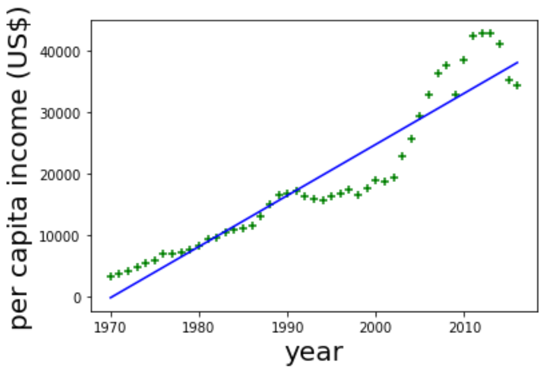
19 1989 16426.725480

```
d.to_csv("canada future house price prediction(1).csv",index=False)
```

21 1991 17266.097690

```
%matplotlib inline
plt.xlabel("year",fontsize=20)
plt.ylabel("per capita income (US$)",fontsize=20)
plt.scatter(df.year,df["per capita income (US$)"],color="green",marker="+")
plt.plot(df.year,lr.predict(df[["year"]]),color="blue")
```

```
[<matplotlib.lines.Line2D at 0x7f611c6d53d0>]
```



Colab paid products - Cancel contracts here