In [28]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn import linear_model
```

In [57]:
```
df=pd.read_csv("canada_per_capita_income.csv")
```

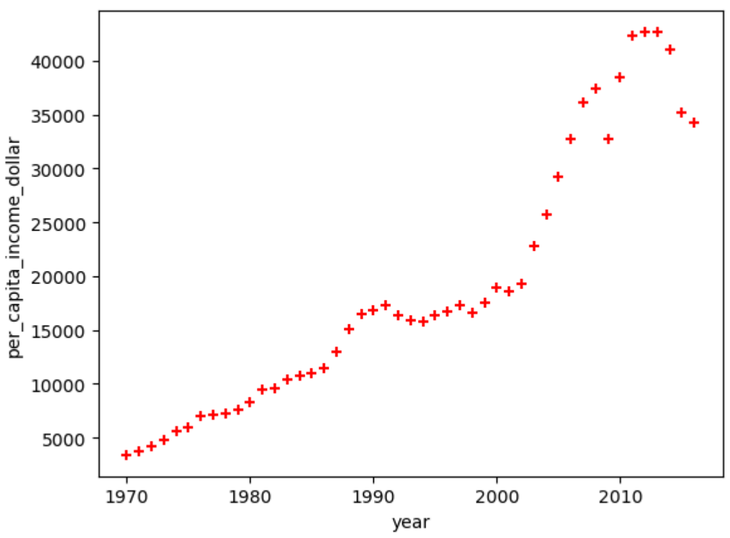
In [58]:
```
df.head()
```

Out[58]:

	year	per_capita_income_dollar
0	1970	3399.299037
1	1971	3768.297935
2	1972	4251.175484
3	1973	4804.463248
4	1974	5576.514583

In [62]:
```
plt.xlabel('year')
plt.ylabel('per_capita_income_dollar')
plt.scatter(df.year,df.per_capita_income_dollar,color='red',marker='+')
```

Out[62]:
```
<matplotlib.collections.PathCollection at 0x21ced6b5e50>
```



In [68]:
```
# Create linear regression object
reg = linear_model.LinearRegression()
reg.fit(df[['year']],df.per_capita_income_dollar)
```

Out[68]:
```
LinearRegression
LinearRegression()
```

In [79]:
```
#predict data for next few years
df2=pd.read_csv('canada_per_capita_income_new.csv")
```

In [80]:
```
df2.head()
```

Out[80]:

	year
0	2017
1	2018
2	2019
3	2020
4	2021

In [83]:
```
df2['Income_per_Capita1']=reg.predict(df2)
```

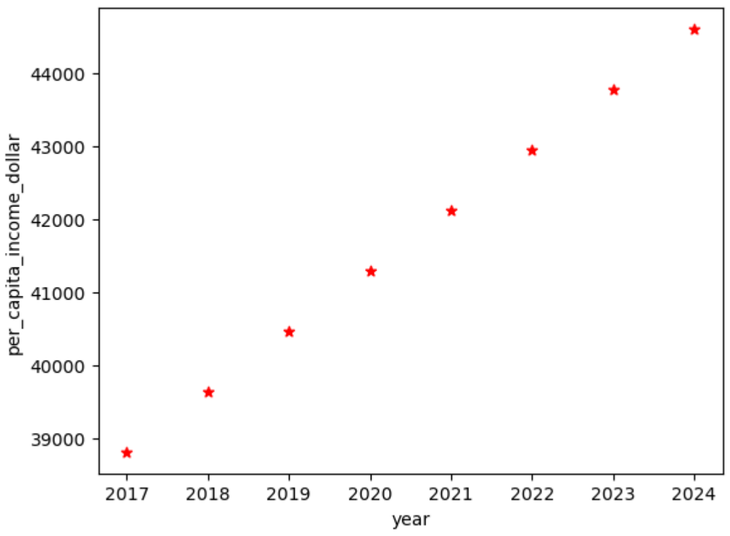
In [85]:
```
df2
```

Out[85]:

	year	Income_per_Capita1
0	2017	38803.298869
1	2018	39631.763944
2	2019	40460.229019
3	2020	41288.694094
4	2021	42117.159170
5	2022	42945.624245
6	2023	43774.089320
7	2024	44602.554395

In [89]:
```
plt.xlabel('year')
plt.ylabel('per_capita_income_dollar')
plt.scatter(df2.year,df2.Income_per_Capita1,color='red',marker='*')
```

Out[89]:
```
<matplotlib.collections.PathCollection at 0x21ced769e50>
```



In [86]:
```
df2.to_csv("prediction.csv")
```