```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split


df=pd.read_csv('/content/Lab_3_canada_per_capita_income.csv')


df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 47 entries, 0 to 46
Data columns (total 2 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   year                    47 non-null     int64  
 1   per capita income (US$)  47 non-null     float64
dtypes: float64(1), int64(1)
memory usage: 880.0 bytes
```

```
df.isna().sum()
```

```
year                      0
per capita income (US$)    0
dtype: int64
```

```
df.rename(columns={'per capita income (US$)':'income'},inplace=True)


df.columns
```

```
Index(['year', 'income'], dtype='object')
```

```
X = df['year'].values.reshape(-1,1)
y = df['income'].values


X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)



model = LinearRegression()
```

```
model.fit(X_train, y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
income_2020 = model.predict([[2020]])
```

```
income_2020
```

```
array([41027.67748165])
```

```
model.score(df[['year']],df.income)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:432: UserWarning: X has feature names, but LinearRegression was fitted without f
  warnings.warn(
0.8906178144427537
```