```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

from sklearn.linear_model import LinearRegression

#read data

data = pd.read_csv('D:\datasetgdp.csv')

data.head()
```

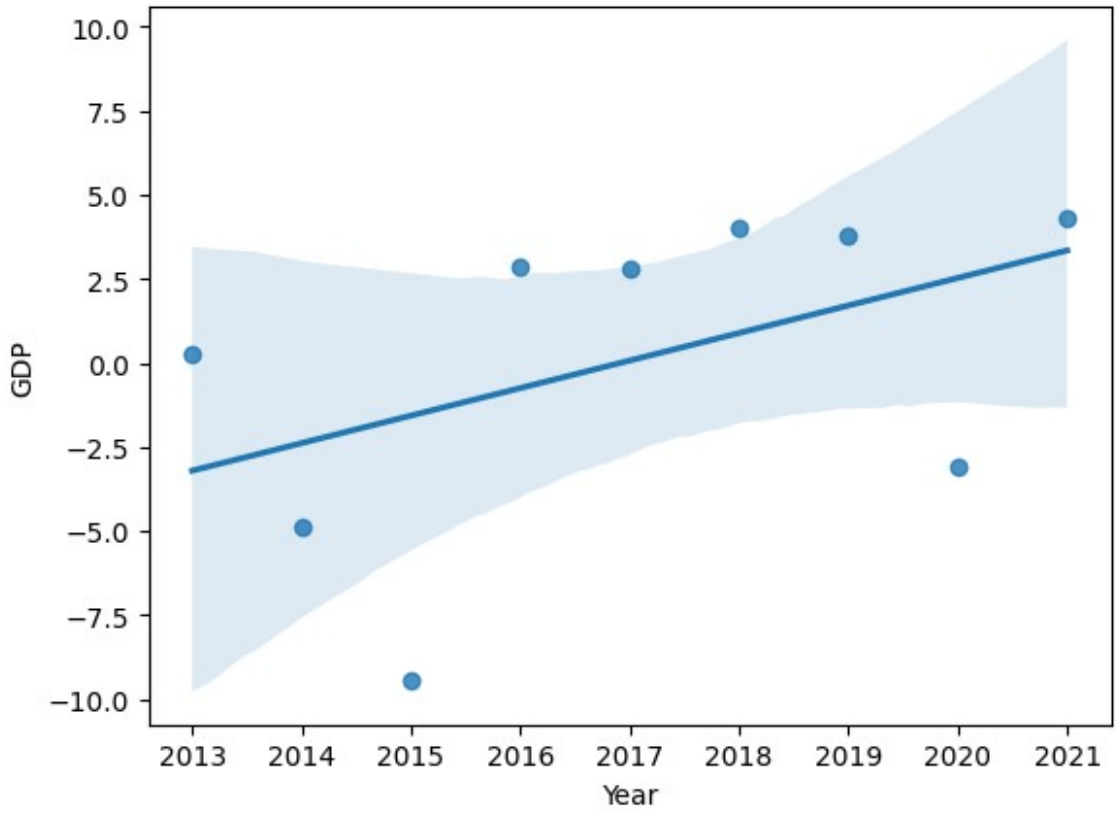
```
   Year       GDP
0  2013  0.273496
1  2014 -4.875250
2  2015 -9.443832
3  2016  2.854803
4  2017  2.810645
```

```
#view data

data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9 entries, 0 to 8
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   Year    9 non-null      int64  
 1   GDP     9 non-null      float64
dtypes: float64(1), int64(1)
memory usage: 276.0 bytes
```

```
#visualize data distribution

sns.regplot(x='Year', y='GDP', data=data)
```

```
<Axes: xlabel='Year', ylabel='GDP'>
```




```
#regression

#create instance

lri = LinearRegression()

#train model

lri.fit(data[['Year']].values, data.GDP)

LinearRegression()

lri.predict([[2022]])

array([4.16663934])
```