

NAME - Hamza

```
In [ ]: pip install wbgapi

In [1]: import pandas as pd
import wbgapi as wb
import matplotlib.pyplot as plt

In [2]: #Pandas function to read the dataset values from World Indicator data
text=pd.read_csv(r"C:\\Hamza ADS2\\World_Bank_Indicators.csv", low_memory=False)

In [3]: #Printing the initial rows and columns
text.head(6)
```

Out[3]:

	economy	AUS	AUS.1	AUS.2	AUS.3	AUS.4	AUS.5	AUS.6	AU
0	series	EG.ELC.NGAS.ZS	EG.ELC.RNWX.KH	EN.ATM.CO2E.PC	EN.ATM.GHGT.KT.CE	NE.IMP.GNFS.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.FE.ZS	SP.POP.TC
1	YR1961	0.0	303000000.0	NaN	NaN	14.9985775248933	19683055213.3498	NaN	1048300
2	YR1962	0.0122204570450935	295000000.0	NaN	NaN	12.6089156220136	19922723709.262	NaN	1074200
3	YR1963	0.0182575038340758	301000000.0	NaN	NaN	13.809598086622	21539926083.548	NaN	1095000
4	YR1964	0.0230066390586998	289000000.0	NaN	NaN	13.7398833051007	23801097547.3177	NaN	1116700
5	YR1965	0.0209198768715818	296000000.0	NaN	NaN	15.2403535244665	25977153096.6514	NaN	1138800

6 rows × 81 columns

```
In [4]: #Index set function
text1=text.set_index('economy')

In [5]: #Printing the initial rows and columns in transpose format
text1.T.head(8)
```

Out[5]:

economy	series	YR1961	YR1962	YR1963	YR1964	YR1965	YR1966
AUS	EG.ELC.NGAS.ZS	0.0	0.0122204570450935	0.0182575038340758	0.0230066390586998	0.0209198768715818	0.022237046920169
AUS.1	EG.ELC.RNWX.KH	303000000.0	295000000.0	301000000.0	289000000.0	296000000.0	284000000.0
AUS.2	EN.ATM.CO2E.PC	NaN	NaN	NaN	NaN	NaN	NaN
AUS.3	EN.ATM.GHGT.KT.CE	NaN	NaN	NaN	NaN	NaN	NaN
AUS.4	NE.IMP.GNFS.ZS	14.9985775248933	12.6089156220136	13.809598086622	13.7398833051007	15.2403535244665	15.1035472626615
AUS.5	NY.GDP.MKTP.CD	19683055213.3498	19922723709.262	21539926083.548	23801097547.3177	25977153096.6514	27309889125.322
AUS.6	SL.UEM.1524.FE.ZS	NaN	NaN	NaN	NaN	NaN	NaN
AUS.7	SP.POP.TOTL	10483000.0	10742000.0	10950000.0	11167000.0	11388000.0	11651000.0

8 rows × 61 columns

```
In [40]: #Forming the economic and climate indicators
cntry_nmes = ['CHN','ARG','BGD','BGR','PAK','FRA','JPN','LUX']
ecn_vrbl = ['SP.POP.TOTL','NE.DAB.TOTL.ZS','NY.GDP.MKTP.CD','SL.UEM.1524.NE.ZS']
clm_vrbl=['EN.ATM.PM25.MC.T1.ZS','EG.ELC.RNWX.KH','EN.ATM.GHGT.KT.CE','EN.ATM.CO2E.GF.KT']
ECONMY = wb.data.DataFrame(ecn_vrbl, cntry_nmes, mrv=7)
CLMT = wb.data.DataFrame(clm_vrbl, cntry_nmes, mrv=7)

#SP.POP.TOTL: Current population of a country
#NE.DAB.TOTL.ZS: Current expenditure of a country
#NY.GDP.MKTP.CD: Current GDP as USD of a country
#SL.UEM.1524.NE.ZS: Current youth unemployment of country
#EN.ATM.PM25.MC.T1.ZS: Value of PM2.5 pollution, exceeding WHO Interim Target-1
#EG.ELC.RNWX.KH: Production of energy from Renewable sources
#EN.ATM.GHGT.KT.CE: Emissions of Greenshouse gases
#EN.ATM.CO2E.GF.KT: Emission of CO2 using gaseous fuel
```

```
In [41]: # Economic indicators over the years
ECONMY.columns = [r.replace('YR','') for r in ECONMY.columns]
ECONMY=ECONMY.stack().unstack(level=1)
ECONMY.index.names = ['Country_Name', 'Year']
ECONMY.columns
ECONMY.fillna(0)
ECONMY.head(8)
```

Out[41]:

	series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL
Country_Name	Year				
ARG	2015	101.074922	5.947493e+11	NaN	43131966.0
	2016	101.039698	5.575314e+11	NaN	43590368.0
	2017	102.649034	6.436287e+11	22.639999	44044811.0
	2018	101.889164	5.248197e+11	23.730000	44494502.0
	2019	96.962864	4.528184e+11	25.860001	44938712.0
	2020	93.070816	3.895910e+11	30.170000	45376763.0
	2021	93.456347	4.914927e+11	23.240000	45808747.0
BGD	2015	106.728219	1.950787e+11	NaN	156256287.0

```
In [42]: # Climate indicators over the years
CLMT.columns = [r.replace('YR','') for r in CLMT.columns]
CLMT=CLMT.stack().unstack(level=1)
CLMT.index.names = ['Country_Name', 'Year']
CLMT.columns
CLMT.fillna(0)
CLMT.head(8)
```

Out[42]:

	series	EG.ELC.RNW.X.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1.ZS
Country_Name	Year				
ARG	2013	2.942000e+09	90835.257	364480.010986	0.825134
	2014	2.719000e+09	96691.456	363149.993896	0.177212
	2015	2.752000e+09	98359.941	370000.000000	0.306158
	2016	NaN	102268.963	373329.986572	0.061206
	2017	NaN	NaN	377209.991455	0.061212
	2018	NaN	NaN	375100.006104	NaN
	2019	NaN	NaN	369049.987793	NaN
BGD	2013	1.400000e+08	43369.609	176779.998779	99.999808

```
In [43]: #Forming dataframes for economic and climatic indicators
e1=ECONMY.reset_index()
c1=CLMT.reset_index()
e2=e1.fillna(0)
c2=c1.fillna(0)
```

```
In [44]: #Final dataframe with climate and economic variables
finl = pd.merge(e2, c2)
finl.head(6)
```

Out[44]:

series	Country_Name	Year	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNW.X.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE
0	ARG	2015	101.074922	5.947493e+11	0.000000	43131966.0	2.752000e+09	98359.941	370000.000000
1	ARG	2016	101.039698	5.575314e+11	0.000000	43590368.0	0.000000e+00	102268.963	373329.986572
2	ARG	2017	102.649034	6.436287e+11	22.639999	44044811.0	0.000000e+00	0.000	377209.991455
3	ARG	2018	101.889164	5.248197e+11	23.730000	44494502.0	0.000000e+00	0.000	375100.006104
4	ARG	2019	96.962864	4.528184e+11	25.860001	44938712.0	0.000000e+00	0.000	369049.987793
5	BGD	2015	106.728219	1.950787e+11	0.000000	156256287.0	1.580000e+08	48782.101	192179.992671

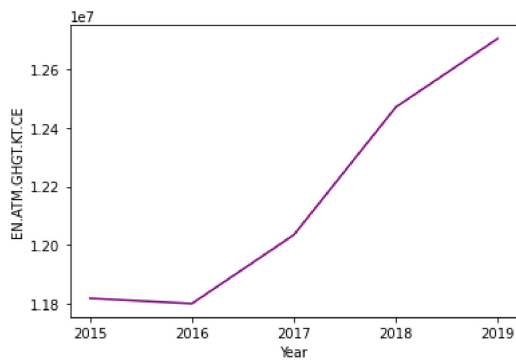
```
In [11]: # Descriptive statistics summary for China
t1=finl[(finl['Country_Name']=='CHN')]
t1.describe()
```

Out[11]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNWX.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.0	5.000000e+00	5.000000e+00	5.000000	5.000000e+00	5.000
mean	98.179512	1.255600e+13	0.0	1.394874e+09	5.677020e+10	154008.132800	1.216615e+07	49.803
std	0.851175	1.483887e+12	0.0	1.123892e+07	1.269420e+11	211271.098125	4.046339e+05	45.519
min	97.225697	1.106155e+13	0.0	1.379860e+09	0.000000e+00	0.000000	1.180083e+07	0.000
25%	97.670114	1.123328e+13	0.0	1.387790e+09	0.000000e+00	0.000000	1.181847e+07	0.000
50%	97.880906	1.231041e+13	0.0	1.396215e+09	0.000000e+00	0.000000	1.203527e+07	81.118
75%	98.851164	1.389482e+13	0.0	1.402760e+09	0.000000e+00	366949.356000	1.247109e+07	81.239
max	99.269679	1.427994e+13	0.0	1.407745e+09	2.838510e+11	403091.308000	1.270509e+07	86.660

The average total expenditure for China is 98.18

```
In [14]: #Line plot visualisation for Emissions of Greenhouse gases in China
plt.plot(t1["Year"], t1["EN.ATM.GHGT.KT.CE"],color="purple")
plt.xlabel("Year")
plt.ylabel("EN.ATM.GHGT.KT.CE")
plt.show()
```



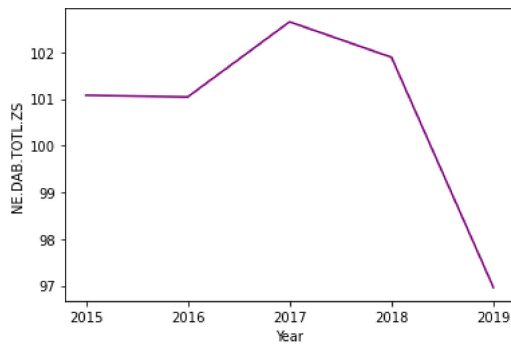
```
In [15]: # Descriptive statistics summary for Argentina
t2=finl[(finl['Country_Name']=='ARG')]
t2.describe()
```

Out[15]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNWX.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000e+00	5.000000e+00	5.00000	5.000000	5.000
mean	100.723136	5.547095e+11	14.446000	4.404007e+07	5.504000e+08	40125.78080	372937.994385	0.085
std	2.204105	7.209881e+10	13.238084	7.143132e+05	1.230732e+09	54961.86715	3421.487899	0.126
min	96.962864	4.528184e+11	0.000000	4.313197e+07	0.000000e+00	0.00000	369049.987793	0.000
25%	101.039698	5.248197e+11	0.000000	4.359037e+07	0.000000e+00	0.00000	370000.000000	0.000
50%	101.074922	5.575314e+11	22.639999	4.404481e+07	0.000000e+00	0.00000	373329.986572	0.061
75%	101.889164	5.947493e+11	23.730000	4.449450e+07	0.000000e+00	98359.94100	375100.006104	0.061
max	102.649034	6.436287e+11	25.860001	4.493871e+07	2.752000e+09	102268.96300	377209.991455	0.306

The average total population of Argentina is lower than China. The average youth unemployment of Argentina is higher than China

```
In [16]: #Line plot visualisation for total expenditure in Argentina
plt.plot(t2["Year"], t2["NE.DAB.TOTL.ZS"],color="purple")
plt.xlabel("Year")
plt.ylabel("NE.DAB.TOTL.ZS")
plt.show()
```



```
In [17]: # Descriptive statistics summary for Bangladesh
t3=finl[(finl['Country_Name']=='BGD')]
t3.describe()
```

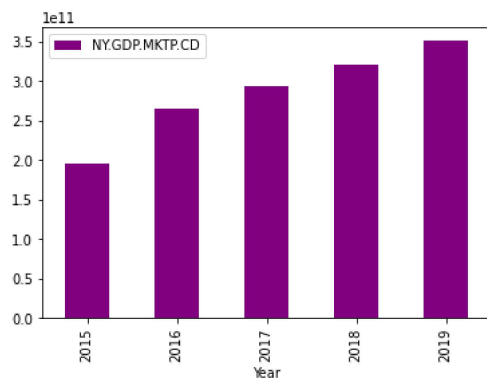
Out[17]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNW.X.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000e+00	5.000000e+00	5.000000	5.000000	5.000
mean	104.855918	2.853374e+11	4.826000	1.596683e+08	3.160000e+07	20475.061200	202632.000732	59.978
std	1.458094	5.971492e+10	6.626521	2.684717e+06	7.065975e+07	28088.184101	9881.383801	54.753
min	102.969504	1.950787e+11	0.000000	1.562563e+08	0.000000e+00	0.000000	192179.992676	0.000
25%	103.881122	2.652362e+11	0.000000	1.579772e+08	0.000000e+00	0.000000	193960.006714	0.000
50%	105.329352	2.937546e+11	0.000000	1.596854e+08	0.000000e+00	0.000000	203080.001831	99.942
75%	105.371393	3.213790e+11	11.370000	1.613767e+08	0.000000e+00	48782.101000	208000.000000	99.952
max	106.728219	3.512384e+11	12.760000	1.630462e+08	1.580000e+08	53593.205000	215940.002441	99.999

The average total electricity production from renewable sources in Bangladesh is lower than Argentina and China. The average total population of Bangladesh is higher than Argentina but lower than China

```
In [21]: #Bar plot visualisation for current market GDP in Bangladesh
t3.plot(x="Year", y="NY.GDP.MKTP.CD", kind="bar",color="purple")
```

Out[21]: <AxesSubplot:xlabel='Year'>



```
In [22]: # Descriptive statistics summary for Bulgaria
t4=finl[(finl['Country_Name']=='BGR')]
t4.describe()
```

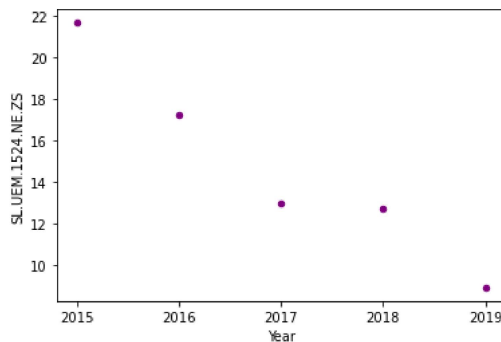
Out[22]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNWX.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000e+00	5.000000e+00	5.000000	5.000000	5.000
mean	96.822546	5.984284e+10	14.676000	7.076512e+06	6.214000e+08	2419.486600	54477.999115	0.065
std	1.571584	7.778530e+09	4.886290	8.020515e+04	1.389493e+09	3313.842559	2419.580624	0.059
min	95.092863	5.078200e+10	8.900000	6.975761e+06	0.000000e+00	0.000000	51279.998779	0.000
25%	95.687406	5.395390e+10	12.680000	7.025037e+06	0.000000e+00	0.000000	53229.999542	0.000
50%	96.771743	5.919945e+10	12.930000	7.075947e+06	0.000000e+00	0.000000	54189.998627	0.108
75%	97.459949	6.636342e+10	17.219999	7.127822e+06	0.000000e+00	5944.207000	56430.000305	0.109
max	99.100768	6.891542e+10	21.650000	7.177991e+06	3.107000e+09	6153.226000	57259.998322	0.109

The average Value of PM2.5 pollution, exceeding WHO Interim Target-1 in Bulgaria is lower than Argentina and Bangladesh. The average total expenditure of Bulgaria is lower than Argentina and Bangladesh.

```
In [29]: # Scatter plot visualisation for youth unemployment in Bulgaria
t4.plot(x="Year", y="SL.UEM.1524.NE.ZS", kind="scatter", color="purple")
```

Out[29]: <AxesSubplot:xlabel='Year', ylabel='SL.UEM.1524.NE.ZS'>



```
In [30]: # Descriptive statistics summary for Pakistan
t5=finl[(finl['Country_Name']=='PAK')]
t5.describe()
```

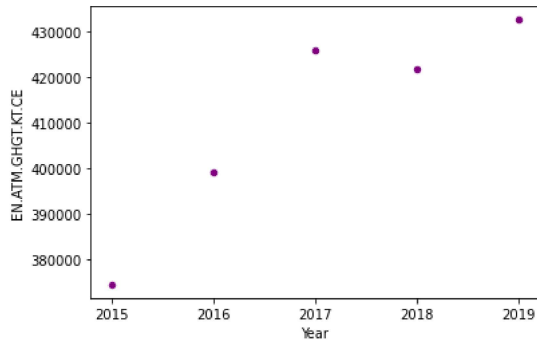
Out[30]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNWX.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000e+00	5.000000e+00	5.000000	5.000000	5.000
mean	108.657822	3.200859e+11	4.804000	2.079516e+08	1.680000e+08	30091.402000	410732.000732	59.760
std	1.766379	3.224969e+10	4.508773	6.779055e+06	3.756594e+08	41204.715396	23823.170520	54.553
min	106.445849	2.705561e+11	0.000000	1.994270e+08	0.000000e+00	0.000000	374510.009766	0.000
25%	107.228145	3.136299e+11	0.000000	2.036314e+08	0.000000e+00	0.000000	399109.985352	0.000
50%	109.028811	3.209095e+11	6.610000	2.079062e+08	0.000000e+00	0.000000	421619.995117	99.549
75%	110.123849	3.392056e+11	7.850000	2.122283e+08	0.000000e+00	74982.816000	425920.013428	99.614
max	110.462457	3.561282e+11	9.560000	2.165653e+08	8.400000e+08	75474.194000	432500.000000	99.640

The average total expenditure of Pakistan is higher than Bulgaria and Bangladesh. The average greenhouse gas emissions of Pakistan is higher than Bulgaria and Bangladesh.

```
In [31]: #Scatter plot visualisation for greenhouse gas emission in Pakistan
t5.plot(x="Year", y="EN.ATM.GHGT.KT.CE", kind="scatter", color="purple")
```

```
Out[31]: <AxesSubplot:xlabel='Year', ylabel='EN.ATM.GHGT.KT.CE'>
```



```
In [32]: # Descriptive statistics summary for France
t6=finl[(finl['Country_Name']=='FRA')]
t6.describe()
```

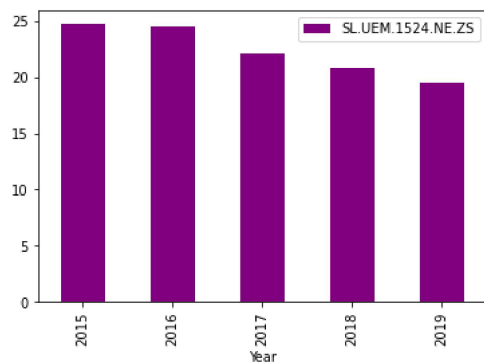
```
Out[32]:
```

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNWX.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000e+00	5.000000e+00	5.0	5.000000	
mean	100.840816	2.605426e+12	22.322000	6.690825e+07	6.983400e+09	0.0	427955.999756	
std	0.236295	1.540671e+11	2.265694	2.815854e+05	1.561536e+10	0.0	9328.185050	
min	100.566449	2.439189e+12	19.530001	6.654827e+07	0.000000e+00	0.0	414040.008545	
25%	100.605068	2.472964e+12	20.760000	6.672410e+07	0.000000e+00	0.0	422739.990234	
50%	100.957362	2.595151e+12	22.150000	6.691802e+07	0.000000e+00	0.0	433220.001221	
75%	101.010618	2.728870e+12	24.469999	6.710193e+07	0.000000e+00	0.0	433600.006104	
max	101.064581	2.790957e+12	24.700001	6.724893e+07	3.491700e+10	0.0	436179.992676	

The average youth unemployment in France is higher than Pakistan and Bulgaria. The average electricity production from renewable source in France is higher than Pakistan and Bulgaria

```
In [37]: #Bar plot visualisation for youth unemployment in France
t6.plot(x="Year", y="SL.UEM.1524.NE.ZS", kind="bar", color="purple")
```

```
Out[37]: <AxesSubplot:xlabel='Year'>
```



```
In [45]: # Descriptive statistics summary for Japan
t7=finl[(finl['Country_Name']=='JPN')]
t7.describe()
```

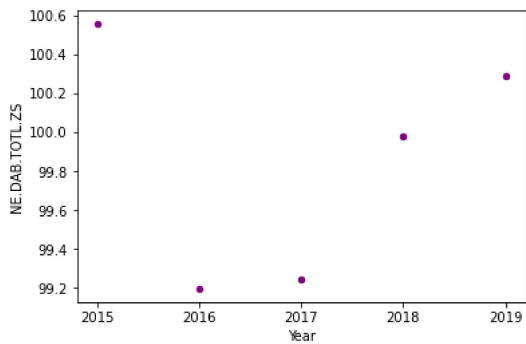
Out[45]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNW.X.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000e+00	5.000000e+00	5.000000	5.000000e+00	
mean	99.851532	4.908120e+12	4.558000	1.269266e+08	1.605840e+10	92486.873800	1.229972e+06	
std	0.613748	2.680149e+11	0.761919	2.061026e+05	3.590767e+10	126648.570206	4.364137e+04	
min	99.194327	4.444931e+12	3.670000	1.266330e+08	0.000000e+00	0.000000	1.166510e+06	
25%	99.240335	4.930837e+12	3.930000	1.268110e+08	0.000000e+00	0.000000	1.204370e+06	
50%	99.980029	5.003678e+12	4.610000	1.269720e+08	0.000000e+00	0.000000	1.246640e+06	
75%	100.289082	5.037835e+12	5.110000	1.270760e+08	0.000000e+00	229517.530000	1.261870e+06	
max	100.553889	5.123318e+12	5.470000	1.271410e+08	8.029200e+10	232916.839000	1.270470e+06	

The average current GDP of Japan is higher than France and Pakistan. The total youth unemployment in Japan is lower than than France and Pakistan.

```
In [49]: # Scatter plot visualisation for total expenditure in Japan
t7.plot(x="Year", y="NE.DAB.TOTL.ZS", kind="scatter",color="purple")
```

Out[49]: <AxesSubplot:xlabel='Year', ylabel='NE.DAB.TOTL.ZS'>



```
In [50]: # Descriptive statistics summary for Luxembourg
t8=finl[(finl['Country_Name']=='LUX')]
t8.describe()
```

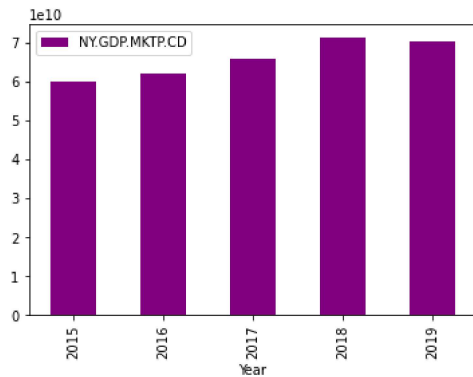
Out[50]:

series	NE.DAB.TOTL.ZS	NY.GDP.MKTP.CD	SL.UEM.1524.NE.ZS	SP.POP.TOTL	EG.ELC.RNW.X.KH	EN.ATM.CO2E.GF.KT	EN.ATM.GHGT.KT.CE	EN.ATM.PM25.MC.T1
count	5.000000	5.000000e+00	5.000000	5.000000	5.000000e+00	5.000000	5.000000	
mean	67.686598	6.589628e+10	16.556000	595181.000000	6.640000e+07	676.92820	10152.000046	
std	1.093820	4.873976e+09	1.819898	20049.72945	1.484749e+08	928.16254	248.937547	
min	66.240142	6.007158e+10	14.180000	569604.000000	0.000000e+00	0.000000	9850.000381	
25%	67.292733	6.221689e+10	15.380000	582014.000000	0.000000e+00	0.000000	10000.000000	
50%	67.461078	6.571218e+10	16.990000	596336.000000	0.000000e+00	0.000000	10119.999886	
75%	68.298985	7.019572e+10	17.350000	607950.000000	0.000000e+00	1624.48100	10310.000420	
max	69.140052	7.128502e+10	18.879999	620001.000000	3.320000e+08	1760.16000	10479.999542	

The average total population of Luxembourg is lower than Japan and France. The average value of electricity from renewable sources in Luxembourg is lower than Japan and France.

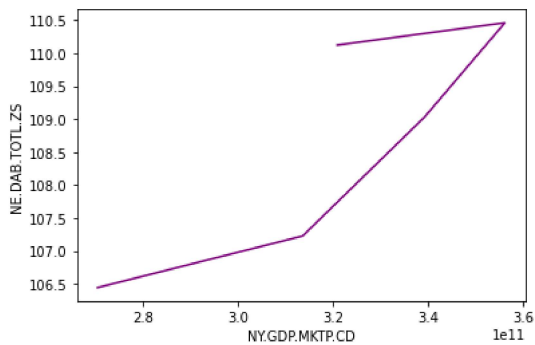
```
In [52]: #Bar plot visualisation for current GDP of Luxembourg
t8.plot(x="Year", y="NY.GDP.MKTP.CD", kind="bar", color="purple")
```

```
Out[52]: <AxesSubplot:xlabel='Year'>
```

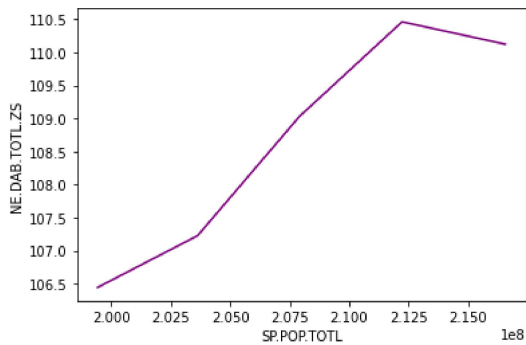


### CORRELATION ANALYSIS - PAKISTAN

```
In [67]: plt.plot(t5["NY.GDP.MKTP.CD"], t5["NE.DAB.TOTL.ZS"], color="purple")
plt.xlabel("NY.GDP.MKTP.CD")
plt.ylabel("NE.DAB.TOTL.ZS")
plt.show()
```



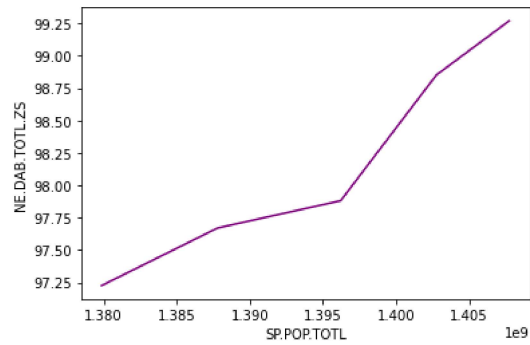
```
In [79]: plt.plot(t5["SP.POP.TOTL"], t5["NE.DAB.TOTL.ZS"], color="purple")
plt.xlabel("SP.POP.TOTL")
plt.ylabel("NE.DAB.TOTL.ZS")
plt.show()
```



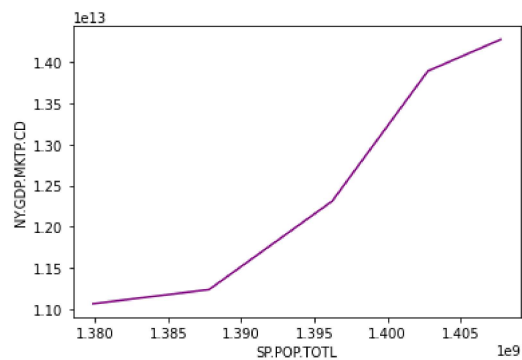
### CORRELATION ANALYSIS - CHINA



```
In [83]: plt.plot(t1["SP.POP.TOTL"], t1["NE.DAB.TOTL.ZS"],color="purple")
plt.xlabel("SP.POP.TOTL")
plt.ylabel("NE.DAB.TOTL.ZS")
plt.show()
```



```
In [85]: plt.plot(t1["SP.POP.TOTL"], t1["NY.GDP.MKTP.CD"],color="purple")
plt.xlabel("SP.POP.TOTL")
plt.ylabel("NY.GDP.MKTP.CD")
plt.show()
```



```
In [ ]:
```