World Debt Dataset

Lets start by importing some libs

```
import pandas as pd
import numpy as np

import plotly.express as px
from plotly.subplots import make_subplots
import plotly.graph_objects as go
import matplotlib.pyplot as plt

import warnings
warnings.filterwarnings('ignore')
import math
!pip install bar-chart-race
!pip install ffmpeg-python
import bar_chart_race as bcr


pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)
```

```
Requirement already satisfied: bar-chart-race in /opt/python/envs/default/lib/
Requirement already satisfied: pandas>=0.24 in /opt/python/envs/default/lib/py
Requirement already satisfied: matplotlib>=3.1 in /opt/python/envs/default/lib
Requirement already satisfied: kiwisolver>=1.0.1 in /opt/python/envs/default/l
Requirement already satisfied: numpy>=1.17 in /opt/python/envs/default/lib/pyt
Requirement already satisfied: python-dateutil>=2.7 in /opt/python/envs/defaul
Requirement already satisfied: cycler>=0.10 in /opt/python/envs/default/lib/py
Requirement already satisfied: pillow>=6.2.0 in /opt/python/envs/default/lib/p
Requirement already satisfied: fonttools>=4.22.0 in /opt/python/envs/default/l
Requirement already satisfied: pyparsing>=2.2.1 in /opt/python/envs/default/li
Requirement already satisfied: packaging>=20.0 in /opt/python/envs/default/lib
Requirement already satisfied: pytz>=2017.3 in /opt/python/envs/default/lib/py
Requirement already satisfied: six>=1.5 in /opt/python/envs/default/lib/python

[notice] A new release of pip is available: 23.0.1 -> 23.1.2
[notice] To update, run: pip install --upgrade pip
Requirement already satisfied: ffmpeg-python in /opt/python/envs/default/lib/p
Requirement already satisfied: future in /opt/python/envs/default/lib/python3.

[notice] A new release of pip is available: 23.0.1 -> 23.1.2
```

Import dataset

```
debtdata = pd.read_csv('/data/notebook_files/Historical Public Debt Database.csv
```

Now lets see our data

```
debtdata
```

	Country Name	Country Code	Indicator Name	Indicator Code	Attribute	1800	1801	1802	1803	1804	...	20
0	Portugal	182	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	68
1	Qatar	453	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	8.9
2	Romania	968	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	12
3	Russia	922	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	8.0
4	Rwanda	714	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	26
...	...	...	...	...	...	...	...	...	...	...	...	...
186	Tonga	866	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	Na
187	Micronesia, Fed. States of	868	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	25
188	Mongolia	948	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	40
189	Afghanistan	512	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	20
190	Puerto Rico	359	Debt to GDP Ratio	GGXWDG_GDP	Value	NaN	NaN	NaN	NaN	NaN	...	36

191 rows × 222 columns

```
debtdata.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 191 entries, 0 to 190
Columns: 222 entries, Country Name to Unnamed: 221
dtypes: float64(217), int64(1), object(4)
memory usage: 331.4+ KB
```

Cleaning Data

```
# Lets remove some column not needed
debtdata =debtdata.drop(["Indicator Name","Indicator Code","Attribute","Unnamed:
```

```
#Flattening the time series data
years = list(debtdata.columns[2:].values)
debtdata = pd.melt(debtdata,id_vars =['Country Name', 'Country Code'],
                value_vars =years,
               var_name ='Year', value_name ='Debt to GDP Ratio')
```

```
debtdata.dtypes
```

```
Country Name          object
Country Code           int64
Year                  object
Debt to GDP Ratio    float64
dtype: object
```

Reshaping the Dataset

```
debtdata["Year"] = debtdata["Year"].astype(str)
```

```
debtdata_pivot = pd.pivot_table(debtdata, values='Debt to GDP Ratio', index=['Ye
                                columns=['Country Name'], aggfunc='sum', fill_value=0)
debtdata_pivot
```

Country Name	Afghanistan	Albania	Algeria	Angola	Anguilla	Antigua and Barbuda	Argentina	Armenia
Year								
1800	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1801	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1802	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1803	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1804	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
...	...	...	...	...	...	...	...	...
2011	7.136649	59.412625	9.489045	33.800868	29.205448	92.430731	38.062868	35.691128
2012	6.625249	62.144467	9.491093	29.486661	30.941384	87.147489	39.433889	36.476851
2013	6.748725	70.391063	7.744179	32.874126	30.422149	94.263091	42.195379	37.958656
2014	6.380300	72.038010	7.955684	40.660897	24.112630	100.422126	43.586674	41.447232
2015	6.246007	73.321326	9.060291	64.238741	22.147033	104.357530	52.130934	46.893048

216 rows × 191 columns

Lets work on the data a see some values

```
# Find the country with the highest and lowest debt-to-GDP ratio for each year
max_debtdata = debtdata.loc[debtdata.groupby('Year')['Debt to GDP Ratio'].idxmax
min_debtdata = debtdata.loc[debtdata.groupby('Year')['Debt to GDP Ratio'].idxmin
```

```
max_debtdata
```

	Country Name	Country Code	Year	Debt to GDP Ratio
138	United Kingdom	112	1800	176.840000
329	United Kingdom	112	1801	177.470000
520	United Kingdom	112	1802	188.990000
711	United Kingdom	112	1803	190.670000
902	United Kingdom	112	1804	188.400000
...	...	...	...	...
40413	Japan	158	2011	231.629355
40604	Japan	158	2012	238.014557
40795	Japan	158	2013	244.477406
40986	Japan	158	2014	249.113739
41177	Japan	158	2015	247.975864

216 rows × 4 columns

```
min_debtdata
```

	Country Name	Country Code	Year	Debt to GDP Ratio
139	United States	111	1800	18.082789
330	United States	111	1801	15.719697
521	United States	111	1802	16.469388
712	United States	111	1803	15.799180
903	United States	111	1804	16.179775
...	...	...	...	...
40389	Hong Kong SAR	532	2011	0.581567
40580	Hong Kong SAR	532	2012	0.552267
40771	Hong Kong SAR	532	2013	0.526190
40962	Hong Kong SAR	532	2014	0.066424
41153	Hong Kong SAR	532	2015	0.062575

216 rows × 4 columns

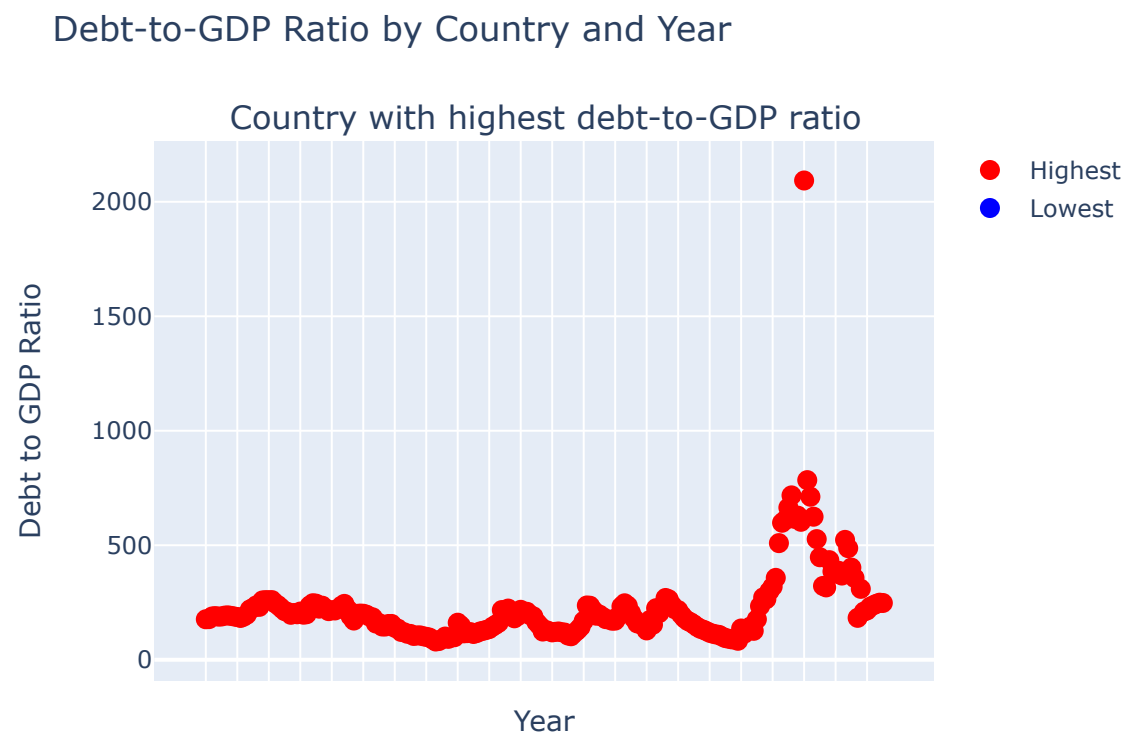
Creating some visualizations of our data

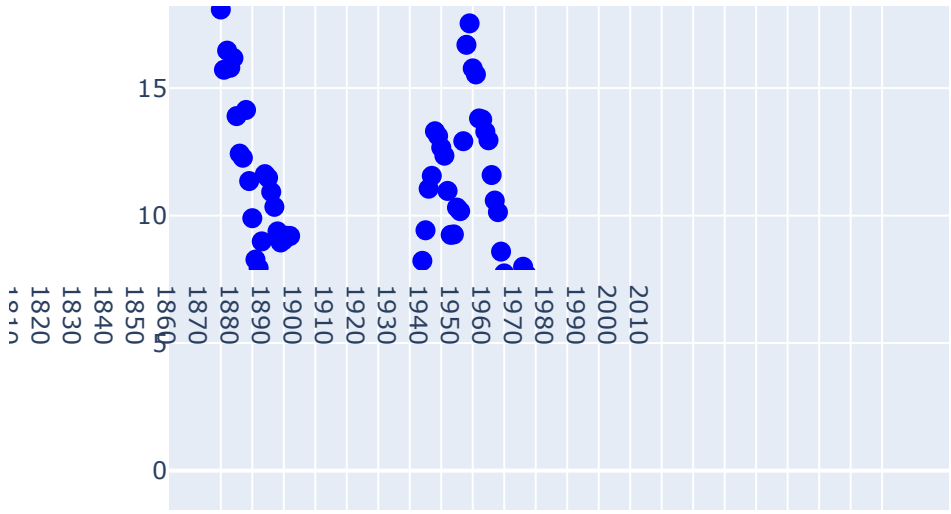
```
# Create the interactive subplots
fig = make_subplots(rows=2, cols=1, shared_xaxes=True,
                    subplot_titles=("Country with highest debt-to-GDP ratio",
                                    "Country with lowest debt-to-GDP ratio"))

# Add the scatter plots for the highest and lowest ratios
fig.add_trace(go.Scatter(x=max_debtdata['Year'], y=max_debtdata['Debt to GDP Rat
                         mode='markers', marker=dict(size=10, color='red'),
                         hovertext=max_debtdata['Country Name'], name='Highest')
fig.add_trace(go.Scatter(x=min_debtdata['Year'], y=min_debtdata['Debt to GDP Rat
                         mode='markers', marker=dict(size=10, color='blue'),
                         hovertext=min_debtdata['Country Name'], name='Lowest'),

# Update the axis and layout properties
fig.update_layout(height=900, title_text="Debt-to-GDP Ratio by Country and Year"
                  xaxis=dict(title="Year"), yaxis=dict(title="Debt to GDP Ratio"

# Show the plot
fig.show()
```





We have a lot of countries, i'll find the G20 countries and do some some data visualization

```
G20 = debtdata[debtdata["Country Name"].isin(["Argentina","Australia","Brazil","
                                              "Euro area","France","Germany","In
                                              "Italy","Japan","Korea","Mexico","
                                              ,"South Africa","Turkey","United K
```

```
G20_pivot = pd.pivot_table(G20, values='Debt to GDP Ratio', index=['Year'],
                           columns=['Country Name'], aggfunc='sum', fill_value=0
```

```
G20_pivot
```

Country Name	Argentina	Australia	Brazil	Canada	China	Euro area	France	Germany
Year								
1800	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1801	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1802	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1803	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1804	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
...	...	...	...	...	...	...	...	...
2011	38.062868	24.200510	61.241677	81.513469	33.086210	86.668889	85.209228	78.310693
2012	39.433889	27.798238	62.310147	84.839928	34.021460	91.250438	89.581390	79.515999
2013	42.195379	30.829111	60.364423	86.079855	36.928375	93.344643	92.395436	77.057504
2014	43.586674	34.322830	63.314056	86.201314	39.830932	94.254036	95.342725	74.479724
2015	52.130934	37.637749	73.697095	91.499318	42.919283	92.490987	96.141149	70.988156

216 rows × 20 columns

```
# Summary statistics for Debt-to-GDP Ratio by year
G20.groupby('Year')['Debt to GDP Ratio'].describe()

# Summary statistics for Debt-to-GDP Ratio by country
G20.groupby('Country Name')['Debt to GDP Ratio'].describe()
```

	count	mean	std	min	25%	50%	75%	max
Country Name								
Argentina	117.0	50.791834	27.185182	9.347420	33.667569	46.623333	60.603656	152.111761
Australia	111.0	51.674653	35.113359	9.682676	23.213111	37.637749	66.106920	136.486486
Brazil	109.0	51.761209	27.222013	10.642757	30.823111	47.889400	68.457200	120.740741
Canada	146.0	62.138196	27.705148	18.445006	40.921172	59.587769	79.817956	155.466577
China	32.0	20.162956	12.421483	0.971136	6.659085	22.248240	27.509040	42.919283
Euro area	21.0	75.463102	10.238533	64.903590	68.032888	70.567387	84.062493	94.254036
France	112.0	73.863105	50.210686	14.388017	32.802367	66.041870	97.248742	237.034600
Germany	113.0	36.657211	17.774360	4.169661	20.104092	38.178008	43.767375	81.003775
India	65.0	50.079912	19.566994	24.999195	35.993487	40.400501	69.642845	84.243037
Indonesia	40.0	38.019346	18.568973	16.430100	24.808902	32.178867	42.613470	95.893500
Italy	155.0	84.927293	32.022708	24.213945	60.275516	90.834497	106.727154	159.720009
Japan	140.0	60.881925	59.479670	4.438238	22.811377	37.899869	69.980696	249.113739
Korea	56.0	17.037302	8.936470	2.631236	9.987567	16.782733	20.865775	37.888716
Mexico	93.0	28.549708	18.469075	0.072204	14.610449	28.117591	42.611257	78.144800
Russia	53.0	53.257457	26.765157	7.425429	37.472145	55.764579	74.200426	116.000000
Saudi Arabia	25.0	49.884826	37.021115	1.567105	12.055995	56.838427	81.564231	102.991788
South Africa	102.0	56.528960	23.583256	23.198200	38.200898	46.932259	71.907162	124.587983
Turkey	65.0	35.154797	12.072866	18.263554	27.644993	33.317822	39.162600	77.935973
United Kingdom	216.0	114.613486	67.309219	27.271153	47.030578	107.546903	176.045292	269.797597
United States	213.0	30.343693	28.494901	0.021858	7.434861	20.155968	47.848521	121.197526

```
fig = px.scatter(G20, x='Year', y='Debt to GDP Ratio', color='Country Name', hov
```