

Series Exercises - due tuesday 09/14 | by Jaya, Clarissa and Charlotte

To get accustomed to Series, let's explore some data on the wealth of 10 randomly selected countries. Data below presents the GDP per capita for these countries in 2008.

Exercise 1

Use the code below to get started:

```
In [23]: import pandas as pd
gdppercap = pd.Series([34605, 34493, 12393, 44200, 10041,
                      58138, 4709, 49284, 10109, 42536],
                      index=['Bahrain', 'Belgium', 'Bulgaria',
                             'Ireland', 'Macedonia', 'Norway',
                             'Paraguay', 'Singapore',
                             'South Africa', 'Switzerland'])
```

Exercise 2

Find the mean, median, minimum and maximum values of GDP per capita in this data.

```
In [24]: f'Mean of GDP per capita of countries is {gdppercap.mean()}'
```

```
Out[24]: 'Mean of GDP per capita of countries is 30050.8'
```

```
In [25]: f'Median of GDP per capita of countries is {gdppercap.median()}'
```

```
Out[25]: 'Median of GDP per capita of countries is 34549.0'
```

```
In [26]: f'Minimum GDP per capita is {gdppercap.min()}'
```

```
Out[26]: 'Minimum GDP per capita is 4709'
```

```
In [27]: f'Maxmimum GDP per capita is {gdppercap.max()}'
```

```
Out[27]: 'Maximum GDP per capita is 58138'
```

Exercise 3

Programmatically, determine which country in our data has the highest income per capita, and which has the lowest income per capita.

(Obviously, this is easier to do by just looking at the data, but that's only because this dataset is very small. With a real dataset, you would need to do it with code, so please write code to accomplish this task.)

Hint: Country names form the index for this Series, so to get country names you'll need to access the index.

```
In [28]: print("Country with the highest GDP per capita is")
         gdppercap[gdppercap == gdppercap.max()]
```

```
Country with the highest GDP per capita is
Out[28]: Norway      58138
         dtype: int64
```

```
In [29]: print("Country with the lowest GDP per capita is")
         gdppercap[gdppercap == gdppercap.min()]
```

```
Country with the lowest GDP per capita is
Out[29]: Paraguay    4709
         dtype: int64
```

Exercise 4

Get Python to print out the names of all the countries that have GDP per capita less than \$20,000.

```
In [30]: print("Countries with GDP per capita below $20,000 are")
         gdppercap[gdppercap < 20000]
```

```
Countries with GDP per capita below $20,000 are
Out[30]: Bulgaria      12393
         Macedonia    10041
         Paraguay      4709
         South Africa   10109
         dtype: int64
```

Exercise 5

Get Python to print out the GDP per capita of Switzerland.

```
In [31]: print("GDP per capita of Switzerland is")
         gdppercap['Switzerland']
```

```
Out[31]: GDP per capita of Switzerland is
         42536
```

Exercise 6

Using this formula, calculate the Gini coefficient for our income data.

HINT: Be careful with 0-indexing! Python counts from 0, but mathematical formulas count from 1!

HINT 2: I'm gonna make you calculate Gini coefficients again later, so maybe you should write a function to do this and make life easier later?

```
In [32]: rank = 0
         num = 0
         den = 0
         for gdp in gdppercap.sort_values():
             rank += 1
             num += rank*gdp
             den += gdp
         gc = ((2*num)/(rank*den)) - ((rank + 1)/rank)
         f'The Gini Coefficient of countries in 2018 is {gc:.2f}'
```

```
Out[32]: 'The Gini Coefficient of countries in 2018 is 0.34'
```

Exercise 7

The result we just generated offers a snap-shot of inequality for this subset of countries. But what are the dynamics of inequality for these countries?

There is an idea in economics called the “convergence hypothesis”, which argues that poorer countries are likely to grow faster, and as a result global inequality is likely to decline. Economists advocating for this hypothesis pointed out that while rich countries had to invent new technologies in order to grow, many poor countries simply had to take advantage of innovations already developed by rich countries.

To test this hypothesis, let’s do a small analysis of the dynamics of income inequality in our sample. Create the following Series in your Python session, which provides the average growth rate of GDP per capita for all the countries in our sample from 2000 to 2018.

```
In [33]: for gdp in gdppercap.sort_values():  
         print(gdp)
```

```
4709  
10041  
10109  
12393  
34493  
34605  
42536  
44200  
49284  
58138
```

```
In [34]: avg_growth = pd.Series([-0.29768835, 0.980299584, 4.52991925,  
                                3.686556736, 2.621416804, 0.775132075,  
                                2.015489468, 3.345793635, 1.349993318,  
                                0.982775018],  
                                index=['Bahrain', 'Belgium', 'Bulgaria',  
                                       'Ireland', 'Macedonia', 'Norway',  
                                       'Paraguay', 'Singapore',  
                                       'South Africa', 'Switzerland']  
                                )
```

Using this data on average growth rates in GDP per capita, and assuming growth rates from 2000 to 2018 continue into the future, estimate what our Gini Coefficient may look like in 2025 (remembering that income in our data is from 2008, so we’re extrapolating ahead 17 years)?

In [35]:

```

gdp2025 = pd.Series([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], index=['Bahrain', 'Belg
                    'Ireland', 'Macedonia', 'Norway',
                    'Paraguay', 'Singapore',
                    'South Africa', 'Switzerland'])
for i in (['Bahrain', 'Belgium', 'Bulgaria', 'Ireland', 'Macedonia', 'Norwa
a = gdp2025[i] * (1+(avg_growth[i]))**17
gdp2025[i] = a
gdp2025

rank = 0
num = 0
den = 0
for gdp in gdp2025.sort_values():
    rank += 1
    num += rank*gdp
    den += gdp
gc = ((2*num)/(rank*den)) - ((rank + 1)/rank)
f'The Gini Coefficient of countries in 2025 is {gc:.2f}'

```

Out[35]: 'The Gini Coefficient of countries in 2025 is 0.85'

Exercise 8

Interpret your result – does it seem to imply that we are seeing convergence or not?

- With these results the null hypothesis can be rejected for now as it seems that **the convergence hypothesis is confirmed** by the estimated output in this analysis: the Gini coefficient increases from 0.34 to 0.85, i.e. on average equality increases over the respective time frame.