

- -> notebooks from this lecture: <https://github.com/ine-rmotr-curriculum/freecodecamp-intro-to-pandas>
- -> **pandas**
 - -> this is a library for data analysis
 - -> you get the data from different sources
 - -> then you process the data (combine it, merge it)
 - -> then you visualise it with Pandas <- creating reports, statistical analysis
 - -> this is an older module
- -> how data structures from pandas are processed
- -> reading files, grouping data
- -> **the first data structure which pandas has -> the pandas series**
 - -> this looks like an excel file
 - -> all of them fit on top of numpy
 - -> an order / sequence of elements
 - -> this is an example which has a series of lists -> population data
 - -> it is returning the array type we have
- -> the data is backed by a numpy array
- -> you can select elements as you would in a regular list
- -> the series has an index -> this is similar to a list
- -> you don't see the list, but there is an index there
- -> in the pandas series, each list has an associated value -> it's a lot more explicit than a numpy array
- -> ordered sequences of elements
- -> **you can arbitrarily change the index of the series**
 - -> elements have a default index, but you can change it
 - -> you refer to the numbers by their indices
 - -> in this case the indices are the names of countries -> it looks like a Python dictionary
 - -> dictionaries in Python aren't ordered
 - -> names and labels
 - -> you can pass the index to return the value of the element which has that index

• -> **question**

What will the following code print out?

```
import pandas as pd
```

```
certificates_earned = pd.Series(
    [8, 2, 5, 6],
    index=['Tom', 'Kris', 'Ahmad', 'Beau']
)
```

```
print(certificates_earned)
```

- options
 - Tom 8 Kris 2 Ahmad 5 Beau 6 dtype: int64 <- This one
 - Kris 2 Ahmad 5 Beau 6 Tom 8 dtype: int64
 - Tom 8 Kris 2 Ahmad 5 Beau 6 Name: certificates_earned dtype: int64