

- -> notebooks from this lecture: <https://github.com/ine-rmotr-curriculum/freecodecamp-intro-to-pandas>
- -> conditional selection, boolean selection / arrays
- -> filtering data
- -> we want to select all the countries which the population is > 70 million
- -> this is with a data frame, rather than a series
- -> this returns a boolean series -> so the companies with > 70 million people return true
- -> we are matching the objects in a series with the objects in a dataframe
- -> we are first extracting the rows which we want
- -> **dropping information**
 - -> `df.drop('index', 'index_2')`
 - -> you can also drop columns
 - -> you can also do this with `axis=...` as an argument, where ... is a number
- -> **broadcasting operations**
 - -> you can tell it to subtract certain values from everything in the dataframe
 - -> there are two different independent objects
 - -> when you combine them with an operation, the columns are aligned
- -> **modifying dataframes**
 - -> the operations are immutable -> they're not changing the original dataframe
 - -> they're just returning data from it
 - -> for example, `df.drop`
 - -> **to add a new column**
 - -> NaN <- not a number, blank
 - -> set a column equal to an array, `df['this']=array`
 - -> `df <-` to print the dataframe
 - -> `df['Language']='English' <-` To replace a row with another one
 - -> you can deliberately change the index of a column in a dataframe
 - -> in this example, he's doing it with multiple columns at once
 - -> these operations are immutable -> the original dataframe isn't being changed
- -> **question**

What will the following code print out?

```
import pandas as pd

certificates_earned = pd.DataFrame({
    'Certificates': [8, 2, 5, 6],
    'Time (in months)': [16, 5, 9, 12]
})
names = ['Tom', 'Kris', 'Ahmad', 'Beau']

certificates_earned.index = names
longest_streak = pd.Series([13, 11, 9, 7], index=names)
certificates_earned['Longest streak'] = longest_streak

print(certificates_earned)
```

Tom 13
 Kris 11
 Ahmad 9
 Beau 7
 Name: Longest streak, dtype: int64

	Certificates	Time (in months)	Longest streak <- This one
Tom	8	16	13
Kris	2	5	11
Ahmad	5	9	9
Beau	6	12	7

	Certificates	Longest streak
Tom	8	13
Kris	2	11
Ahmad	5	9
Beau	6	7