

THE COMPLETE MACHINE LEARNING EXTERNSHIP:

WEEKLY ML CODE CLUB: 001

DATA WRANGLING

In this project you will first "wrangle" some data from [Gapminder](#), a Swedish non-profit co-founded by Hans Rosling. "Gapminder produces free teaching resources making the world understandable based on reliable statistics."

- [Cell phones \(total\), by country and year](#)
- [Population \(total\), by country and year](#)
- [Geo country codes](#)

These two links have everything you need to successfully complete the first part of this project.

- [Pandas documentation: Working with Text Data](#) (one question)
- [Pandas Cheat Sheet](#) (everything else)

Part One: Load Data

- Make sure you import all necessary python Libraries.
- Declare a variable for each of the three(3) csv files and read the data.

Part Two: Join Data

- Join the 'Cell phones' and 'population' dataframe (With an inner join on 'Country' and 'year') [Helpful link for merging](#)
- The resulting data frame should be in shape (11640, 4).



- In the 'Geo country codes' rename the columns{'country' to 'geo', and 'name' to 'Country'} ([Helpful link for renaming a dataframe column](#))
- Then, select the 'geo' and 'Country' columns from the 'geo_country' dataframe, and join with your 'merged population and cell phone' data. ([Helpful link for merging](#))
- The resulting dataframe's shape should be: (11640, 5)

Part Three: Make Features

- Calculate the number of cell phones per person, and add this column onto your dataframe. (Hint: Might have to divide the total number of cell phones by the total number of population)
- Assert: If you've calculated correctly you'll get '**1.230769**' cell phones per person in the United States in 2017.
- Modify the 'geo' column to make the 'geocodes' uppercase instead of lowercase. ([Helpful Link](#))

Part Four: Process Data

- Use the '.describe()' method to Generate descriptive statistics and describe your dataframe's numeric columns, and then its non-numeric columns. ([Helpful link for implementing describe method](#))
- Assert: (You'll see the Year period range from '1960' to '2019', and there are '194' unique countries represented.)
- In 2017, what were the top 5 countries with the most cell phones total? (Your list of countries should have these totals below :)



Country	Cell phones total
??	1,474,097,000
??	1,170,000,000
??	435,000,000
??	400,000,000
??	227,000,000

