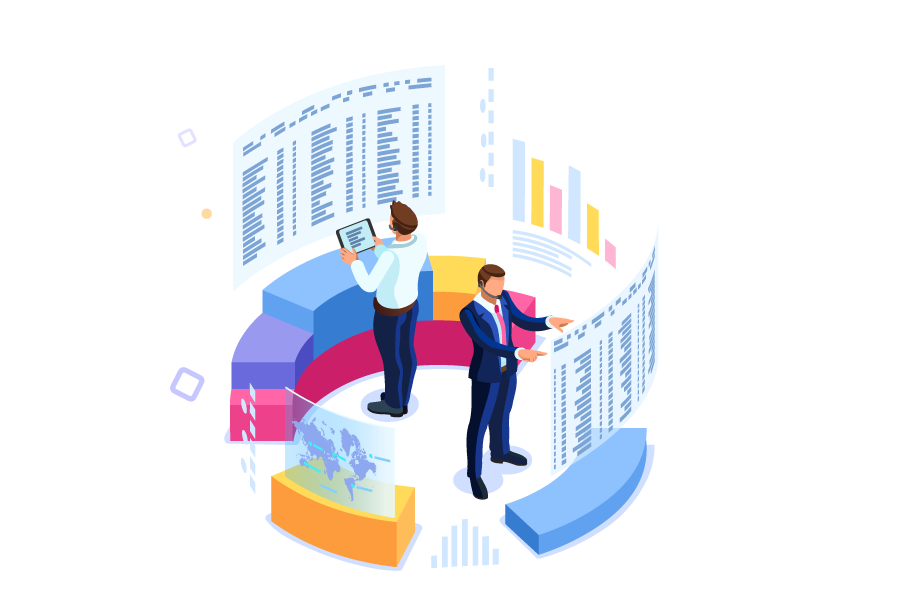
**DAT405: Introduction to data science and AI**

Module1

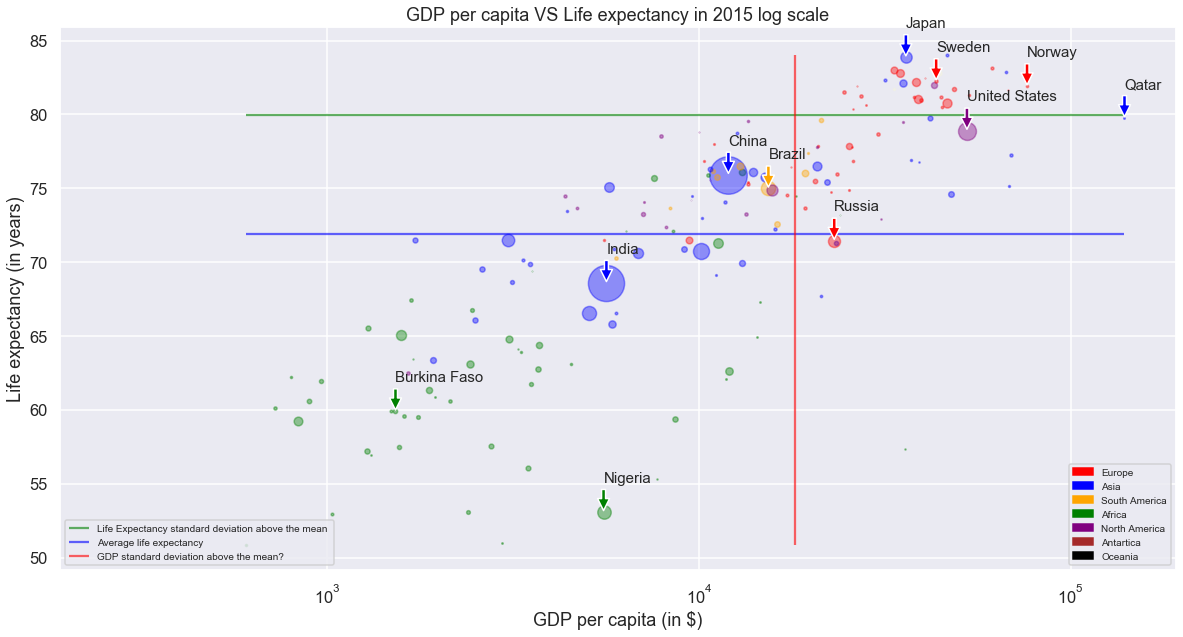
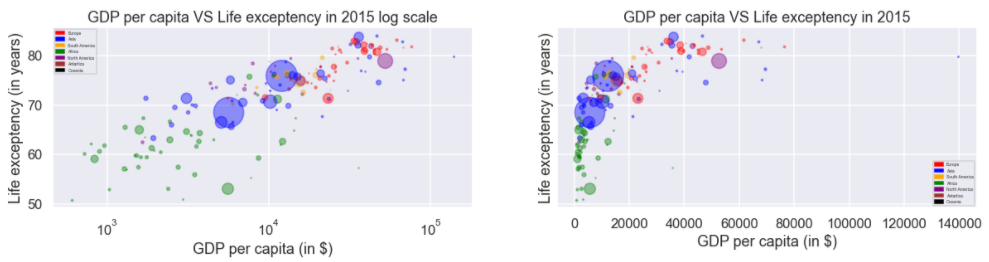


BROU BONI Joël

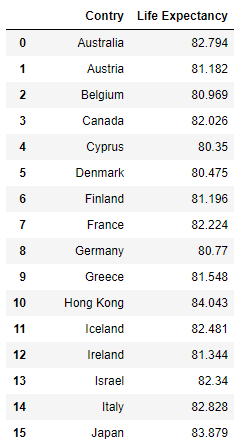
SANOGO Ibrahim Bechir

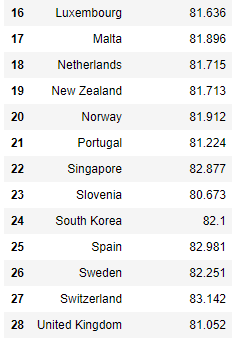
# Part 1

1. After downloading the dataset related to GDP per Capita and life expectancy, we wrote a Python program that draws a scatter plot using this two information for each country. We chose to represent our point cloud for a single year and not for all the years in the dataset. This seemed more coherent to us given the lack of data in the 1900s (some countries did not exist at that time). We focused on 2015 because it was one of the years with the most information. On top of that, the continent to which the countries belonged was also provided, which made our code much easier. It was then possible to group each country by continent and to have a continental vision of GDP per capita according to life expectancy. We also took advantage of the number of inhabitants per country provided, so the size of each point was proportional to it.
2. Most of the time, the greater the GDP per capita the higher the life expectancy. Indeed, the richer the country the better the quality of healthcare for example. It seems that we have the same correlation between the GDP per capita and the life expectancy in our plot. The countries with the lowest GDP are often located in Africa and are located at the bottom left of our plot. This explains why most of the dots on this area are in green. Our result seems reasonable.

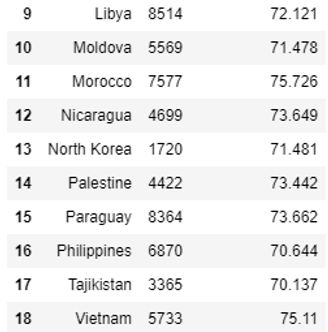


1. As mentioned earlier, there was a lot of data missing for several countries and for different years. We decided to remove from the dataset all the information except those for 2015 to focus on that year. Even for this year, we observed that some information was still missing so we decided to exclude from our study the countries which lacked information about their GDP of the life expectation.
2. There are 29 countries in this case.





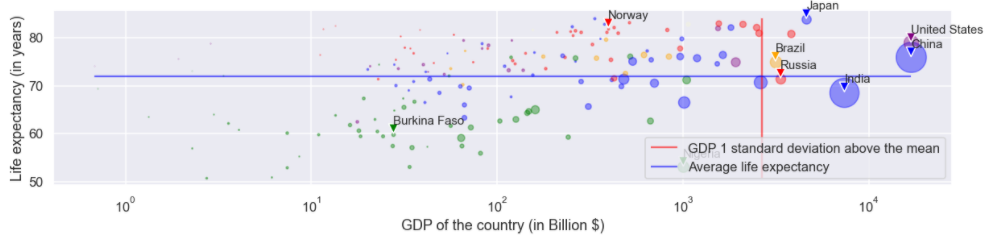
1. The following array represents some countries which have high life expectancy but have low GDP. We arbitrary decided to fix high life expectancy at 70 years old and the GDP per capita at 9000. These values seem reasonable to answer this question.

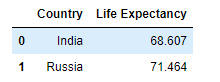


We get countries like Tajikistan, North Korea or Bangladesh which have a high life expectancy despite a low GDP per capita! The result is pertinent.

1. No, every strong economy does not have high life expectancy. This is the case for most for the countries but not for every one of them.

By comparing the life expectancy of a country to the mean life expectancy and the total GDP from this country to this country GDP standard deviation + mean of the GDP of all the countries, we reached the conclusion mentioned above.

We got countries such as India and Russia by this method and this rule should apply to all emerging countries and in particular BRICS (Brazil, Russia, India, China, South Africa) but without China nowadays. It emphasizes the economic inequalities in the same country.

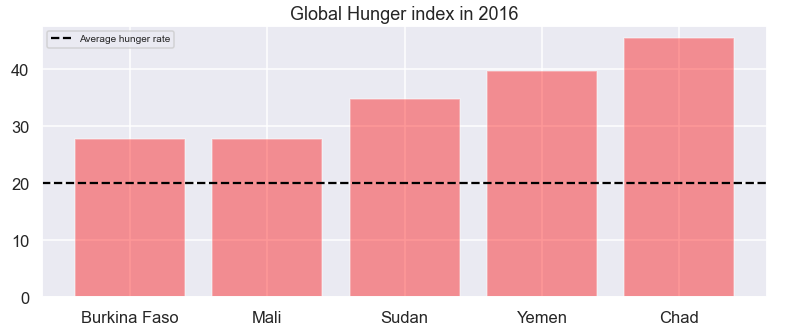
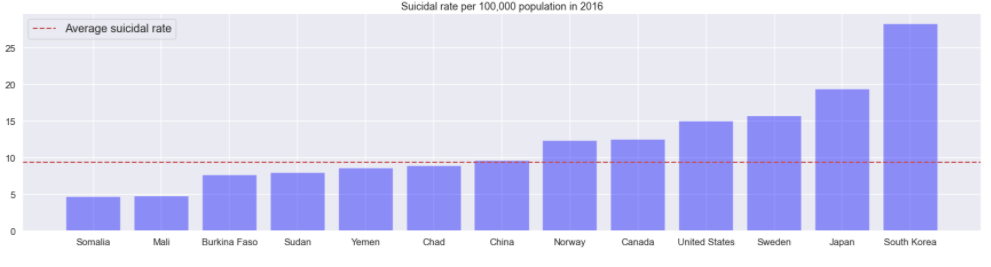


1. GDP per capita is an important indicator of economic performance and a useful unit to make cross-country comparisons of average living standards.

However, it is not a measure of personal income and using it for cross-country comparisons also has some known weaknesses. It does not take into account income distribution in a country. In addition, cross-country comparisons based on the U.S. dollar can be distorted by exchange rate fluctuations and often do not reflect the purchasing power in the countries being compared

Part 2

1. For this part we have chosen to base ourselves on a study carried out a few years ago, saying that it is in the countries where the population is happiest and richest that there are the most deaths by suicide. So the rest of the project, we decided to compare the suicide rate in the world according to the GDP per capita to see if they were correlated. We focused even more on 10 countries. Five that encounter problems of access to water and food above the world average and seven others among the 20 most developed countries in the world. So we have combined 4dataset. One which gave the global hunger rate in the world and the dataset of part one another one which gave the suicide rate in the world and finally the one who is giving the happiest countries in the world from who we choosed the developed countries. We had to clean the 2 new dataset the same way as before and make them correlate with the previous one to get the scatter plot below. To better our comparison we could also had the global



1. This study confirms that it is the countries where the population is supposed to be the happiest that have the highest number of deaths by suicide, which is quite paradoxical. This can be explained, for example, by the much stronger social pressure in these countries, such as South Korea, which is the country most affected by the study.