

My current mentor is 75 years old. Another Nobel Prize winning scientist in our department is 86. Both are highly accomplished and continue to shine in their laboratories every day. My mentor says he will not stop working because he knows retirement can contribute to mental decline and shorter life. I tell him if I am half smart and energetic at his age, I will consider my life a great success. They both know that these problems start with Alzheimer's and the associated dementia which currently costs the world about \$1 trillion US. Dementia causes disability, ultimately resulting in physical, psychological, social and economic impact on caregivers, families and society.

Many organizations, including pharmaceutical companies, are cognizant of these facts but they always look for more information. In light of increasing public vigilance against using drugs but preventive methods, better and unique business decisions can be made -for instance a drug company investing on community adult education can be viewed favorably. Smart companies will do this because they recognize that, just as my mentor, higher education lowers Alzheimer's risk, as reported by NIH, Alzheimer's association, WEBMD, Oxford Journals, etc. However, many countries spend more money in the wrong place, with astounding yearly growth. According to Amnesty International, total global military expenditure ballooned from US\$1.14 trillion in 2001 to \$US 1.711 trillion in 2014, which means we are currently spending twice the amount dementia costs us -just to kill each other or not to be killed. This dilemma is certainly a worry or at least of great interest to savvy company CEOs or international leaders who have to make strategic decisions. So for this project, I would like to do a serious but fun project for which I have three purposes:

1. Do something of useful for a large audience
2. Learn something novel by doing something different that I have never done before
3. Demonstrate a few limits, fun surprises and intricacies of data science

My main idea is to use public data to study the relationship between a few factors and their impact on each other -especially on health. I will look at major economic policies, government per capita and GDP expenditure on education, health, social services and military might. I will use resources such as Google public data, nationmaster's, Amnesty international's and Gapminder's data on arms deals, World Bank economic indicator database, World Health Organization's and Alzheimer's organization's data on global impacts of Dementia. For my last purpose, as a fun fact, I will use University of California Irvine Machine Learning Repository for World Flags Data Set and CIA's factbooks. I will look for potentially wrong relationships between some factors, such as independence from some countries and flag colors. As a vision scientist, I am curious about some psychological studies which claim that our color preferences influence our behavior. For instance, other things being equal, do countries with more red color in their flag have a different war history and culture?

For all aims, I will:

1. First create density functions for the distributions of different factors
2. Test if these distributions are parametric or non-parametric (using Kolmogorov-Smirnov for example) and apply appropriate statistical tests to classify them, such as Mann Whitney or ranksum tests
3. Perform tests on interactions by analyzing causal relationships

4. Identify trends in global movements of major economic, education and health indicators.

In my preliminary analysis, I have investigated the proportion of government expenditures for 78 countries covering all continents and economic standings. The expenditures I looked at: military, education and mental health, are quite skewed. That is to say there is more money allocated for military expenses than mental health measures. This fact may not be surprising since mental health should not be the only focus. However, as the top two figures show, countries with lower per capita GDP are lagging behind. Another important observation is that these countries are spending more on education than the richest ones. The size of the circles and right-ward x-axis movement on the top panels indicates larger GDP. Looking at the two bottom panels, it is obvious that military expenditure is much larger than that for mental health (bottom left panel) and education (bottom right panel). The fact that there are more (higher value) points in both panels indicates the skewness towards military expenses. In other words, points lying above the red dotted-line indicate a disproportionate monetary focus on defense expenses. It will be highly interesting to go deeper into the data and include other factors, even fun facts such as independence year, flag colors, life expectancy, higher education completion rates etc. These investigations can give important clues on investment opportunities and humanitarian activities.