

Life Expectancy is only Weakly Correlated with a Country's GDP

A Data Science Capstone Project by Justin Papreck

Abstract

Using data acquired from the World Health Organization (WHO) and the World Bank, the relationship between GDP and life expectancy at birth was examined in 6 different countries. The countries spanned 5 continents: Africa, Asia, Europe, North America, and South America. Using these data, it became clear that a country's GDP has little direct correlation with the life expectancy, as the two countries with the lowest GDP had both the highest and lowest life expectancies. While other factors clearly come into play, the relationships, changes, and distributions of GDP and life expectancy over time is worth investigating.

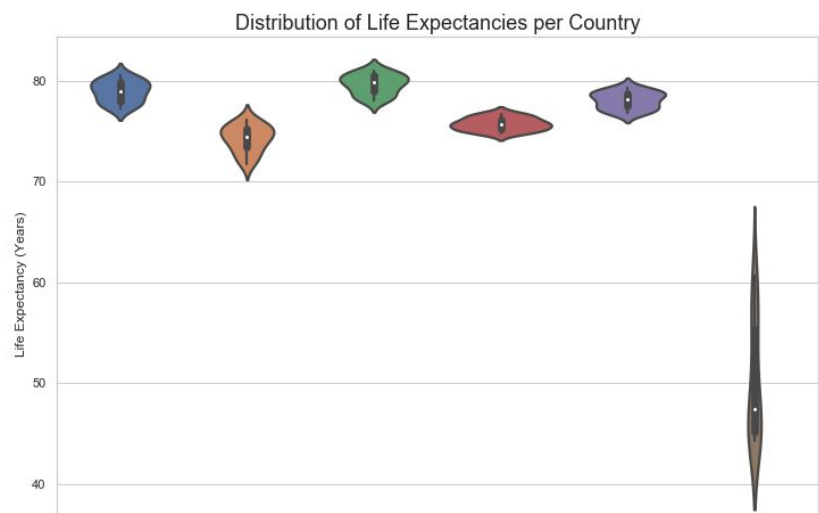
Data

The data acquired from the World Health Organization included entries for 6 countries (Chile, China, Germany, Mexico, United States of America, and Zimbabwe) for 16 consecutive years (2000 to 2015) on just the life expectancy at birth, represented in years. For the same timeframe, the gross domestic product (GDP), a monetary measure of a country's market value of all final goods and services in a given time, represented in US dollars. The data was processed using Python with Pandas with visuals created using Seaborn.

Findings

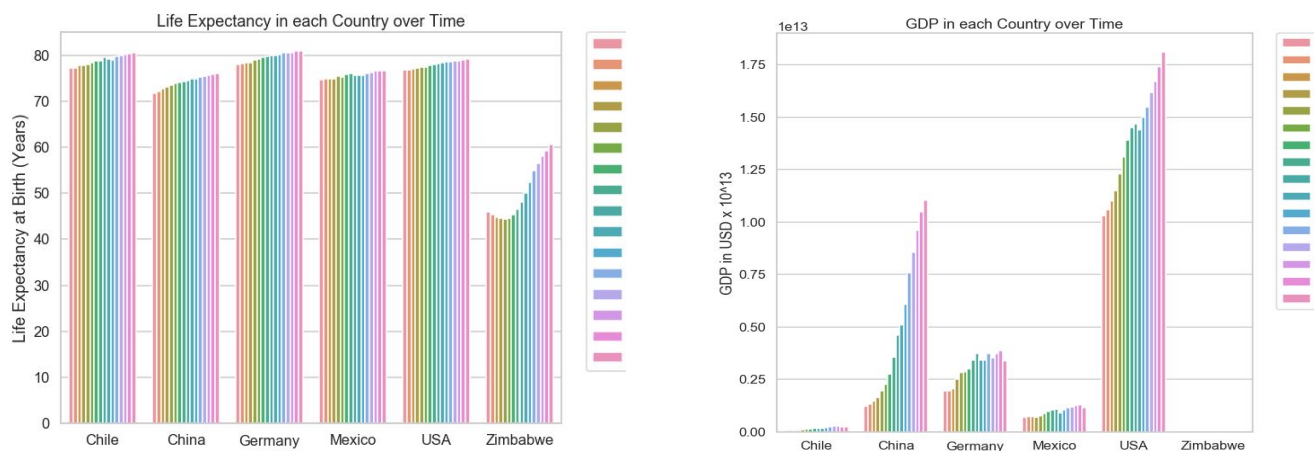
The initial investigation led to questionable conclusions showing that there was absolutely no correlation with a country's GDP and Life Expectancy based on bar plots (not shown) or a violin plot. Examining the distributions of all life expectancies across the 16 year time frame from which the data were collected showed that Zimbabwe had a massive distribution compared to those of the other countries.

Looking at the other countries' distributions, the only one that looks notably different is that of Mexico, which seems to have a much higher number in the lower half than in the upper half as the other countries seem to show.



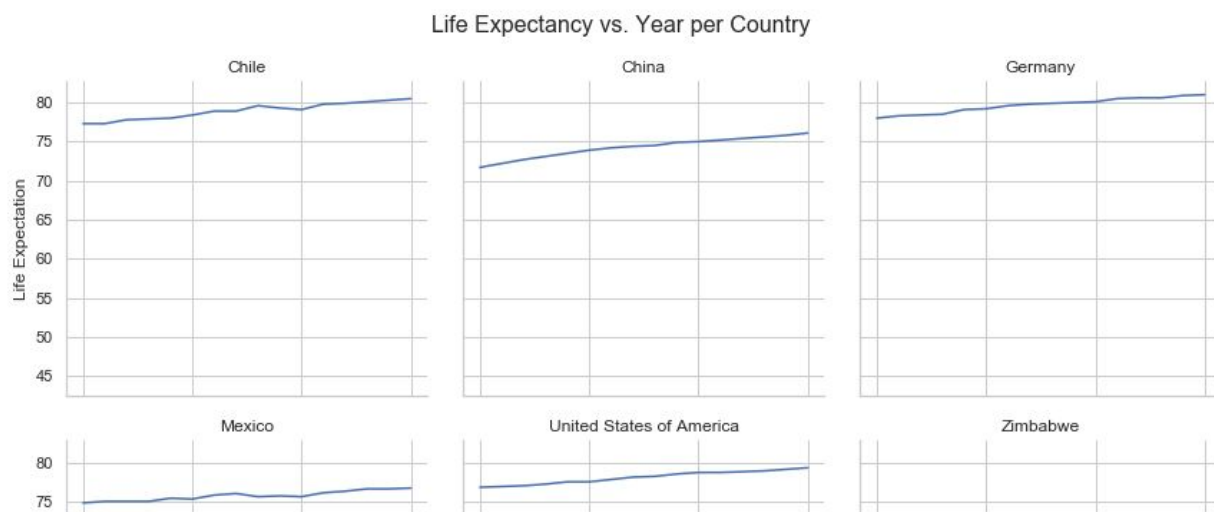
Now that we can see that most of the countries have a relatively narrow distribution compared to that of Zimbabwe, it will be important to consider this when looking at the further data in comparison with GDP. The range of life expectancies for Zimbabwe is between 35 and 68, a value that falls short of the minimum life expectancy for any other countries, all of which fall within the range of 70 to 82.

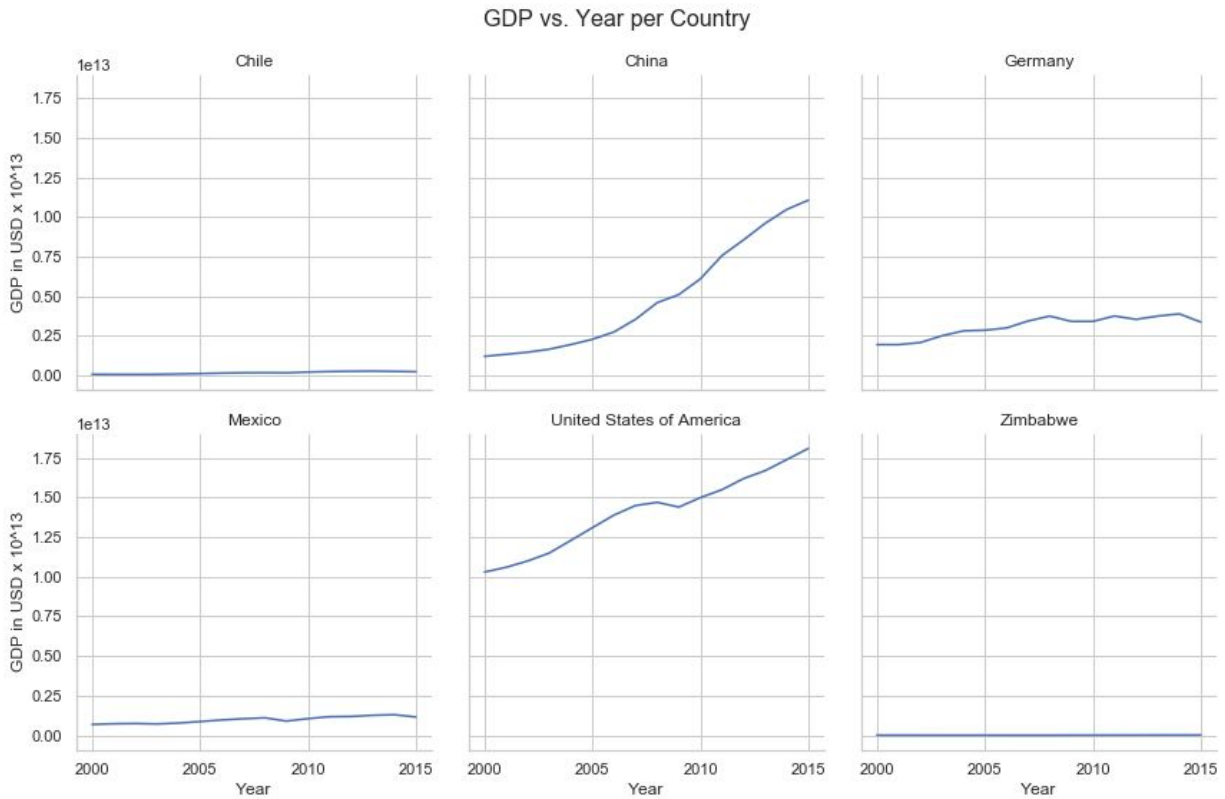
To better visualize the changes in life expectancy over time, I employed a few different visualizations, starting with bar graphs, which I thought gave a very good representation of both the life expectancy and GDP over time, though typically, we don't usually use bar graphs for this purpose, so I followed these up with some more traditional facet grids of line graphs to show each country's change of life expectancy and GDP by year.



As we examine these bar graphs from left to right, we see how the life expectancy and the GDP change each year. By quick glance, it is abundantly clear that the countries Chile and Zimbabwe have such low GDP that they are barely visible if at all on the right graph, yet when comparing the life expectancy from Chile with the other countries we can see it actually has one of the highest life expectancies, while Zimbabwe shows the opposite.

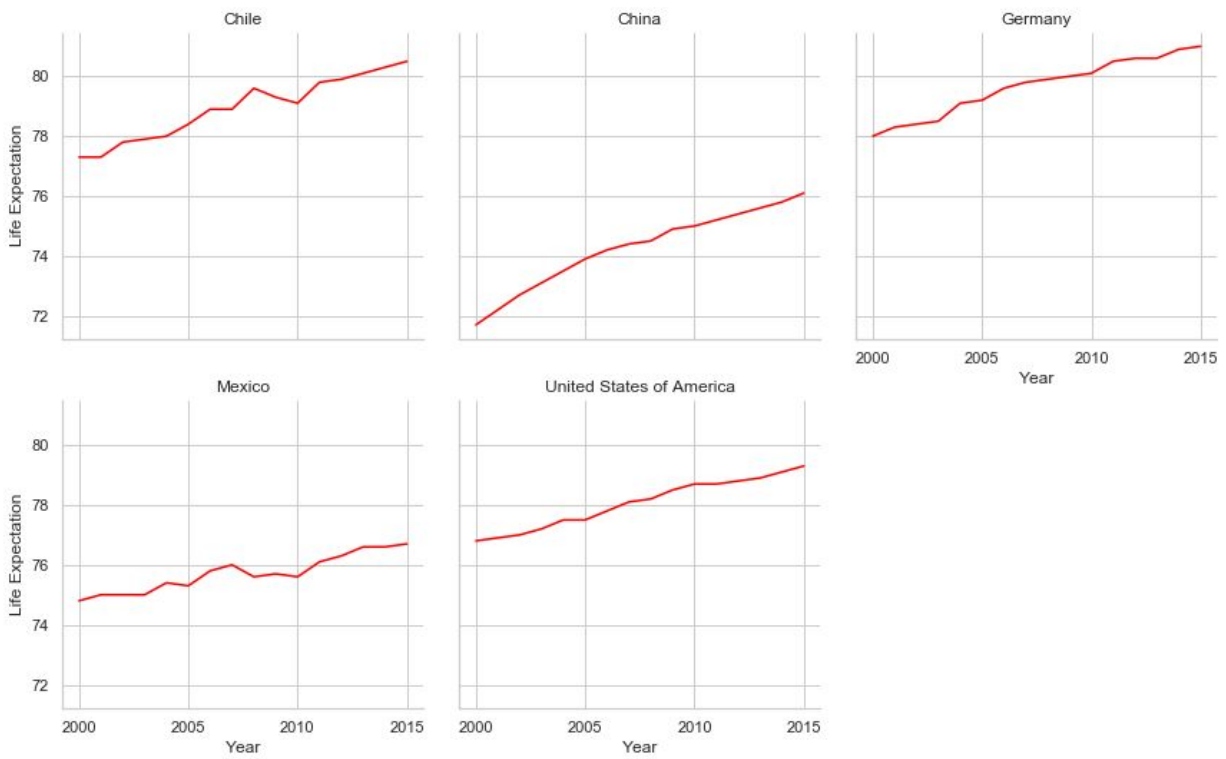
But looking past this, there is a notable trend in that while the GDP of each country increases, the life expectancy also tends in the positive direction. This is further shown in the following facet grids of the same data using line graphs.



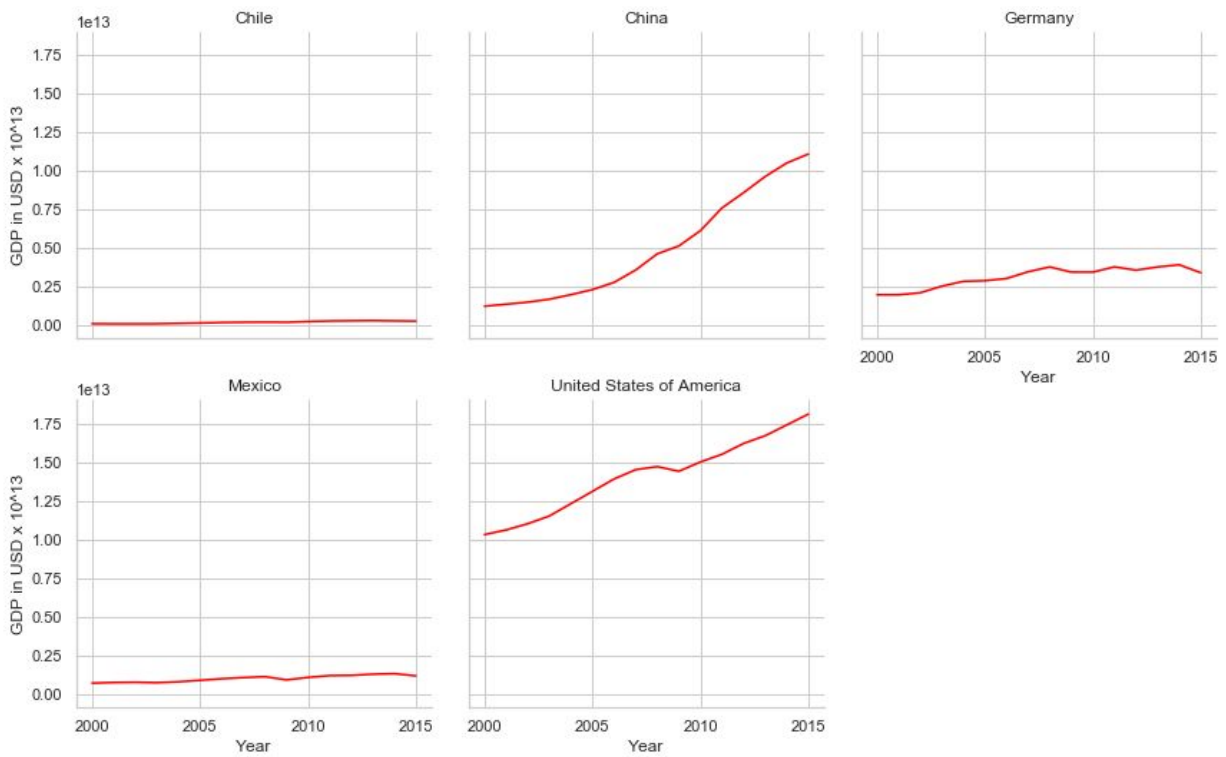


Before continuing with a more multidimensional representation of this data, I wanted to look closer at the 5 countries which had much closer life expectancies to examine the trends without the scaling effect from including Zimbabwe. Here it becomes a little more evident that the countries with the lowest GDP had more fluctuation in life expectancies each year. The dip in the life expectancy in Chile didn't correlate with the seemingly stable GDP. Mexico did have a fluctuation in GDP around the same time as there was a drop in life expectancy, though the USA had a major stall in the GDP from 2008 to 2012, yet the life expectancy continued to increase. Furthermore, the country with the greatest increase in GDP was China, yet China still lags behind the other countries in life expectancy.

Life Expectancy vs. Year per Country

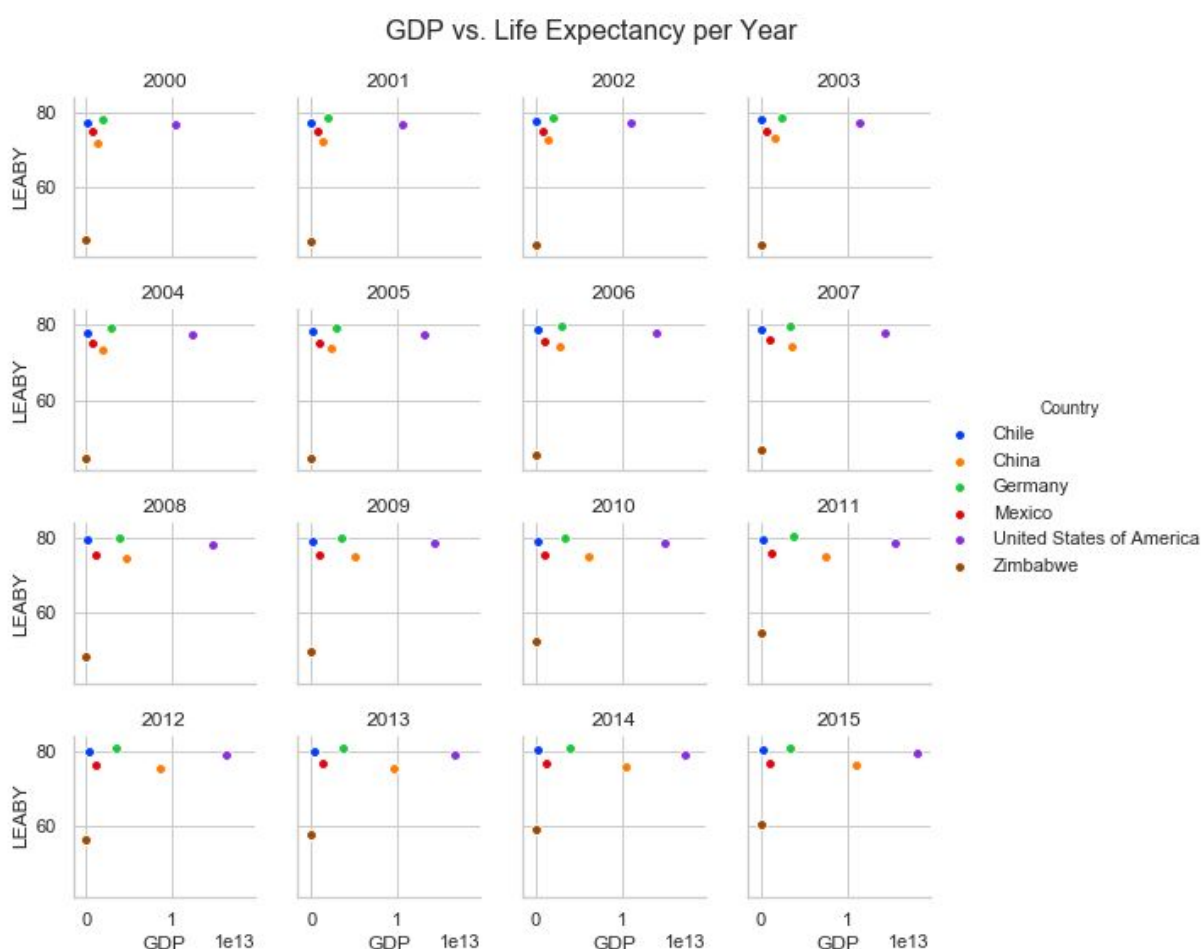


GDP vs. Year per Country



Finally, using a facet grid scatter plot, I can show the distribution of the life expectancies (life

expectancy at birth year: LEABY) with GDP for each of the countries in each of the years. There is a cluster in the upper left of each of these containing Chile, Germany, and Mexico. Initially China also lies within this cluster, though over time, the GDP of China increases, thus the data point moves to the right over time. Along the same concept, the data point for Zimbabwe moves downward and then upward as the life expectancy changes, though the point remains on the far left as the country's GDP does not change. The USA is substantially different from the cluster, as the GDP of the USA is significantly higher than the other, yet vertically it is on the same level, as the life expectancy is no different from the other countries excluding Zimbabwe.



Analysis

The first indicator that life expectancy is not directly correlated with the GDP is the difference between the two countries with the lowest GDP: Chile and Zimbabwe. While Zimbabwe shares both the lowest GDP and the lowest life expectancy, Chile has the second lowest GDP but

actually touts the highest life expectancy. This is highly suggestive that there is something else coming into play. Looking further into Zimbabwe, the life expectancy actually decreased between 2000 and 2005. **Why is Zimbabwe the only country that has a decrease in life expectancy? And why did it rebound in an almost exponential fashion?**

During the early 2000s, Zimbabwe was in a bout of political turmoil. There were periods of economic depression as the president assumed control of the government and used policy to punish his opponents, despite the country-wide ramifications. With hundreds of thousands of people without jobs, over 100,000 people died from malnutrition and starvation. After some political changes, the people of Zimbabwe were able to again find jobs and afford the necessities for survival. It was following these changes that the life expectancy rebounded.

In this case, what must be considered is how low the country's GDP is compared to the rest of the world, but further data would be needed on the distribution of wealth in the country, as it seems most of the wealth was held by very few people while the rest of the country was starving. Even though Chile has the next lowest GDP, it's current GDP is 20 times that of Zimbabwe. Even though the GDP isn't directly correlated with the life expectancy, if the GDP of a country is too low, the difference in the economy changes from having the resources to sustain the population to not having the resources to sustain the people, which massively impacts the life expectancies of the population.

Another major difference that can be seen in the data is the massive economic growth of China (by GDP) in this time period. In the late 1990s and early 2000s, China changed its economic policy and redirected its focus to globalized trade. Over the 16 year period for which this data examined, China showed a dramatic increase in GDP, but the change in life expectancy was underwhelming. In fact, even though China now has the second highest GDP, it has the second lowest life expectancy - suggesting that the economic changes that have been made are bringing a lot of money into the country, but the wealth isn't being put back to maintain the health of the people.

Other considerations that were lacking from these data were the populations, population density, number of doctors or health service providers available to the people, number of people or percent of population with access to clean water and sanitation.