David Rheams

Codecademy Practice

I choose the Life Expectancy Project for the final Codecademy project: Data Visualization with Python. The data for the global Gross Domestic Product (GDP) and Life Expectancy was gathered from public repositories at the World Bank and World Health Organization respectively.

The project asks the following questions:

1. How do you think the histories and the cultural values of each country relate to its GDP and life expectancy?
2. What would have helped make the project data more reliable?
3. What were the limitations of the dataset?
4. Which graphs better illustrate different relationships?

We will use the following four graphs to work through these questions. Each country has gone through significant changes and challenges between 2000 and 2015:

- **China** saw a major shift in its economy from a socialist-oriented structure to a capitalist one.
- The **United States** experienced the longest war of its history, if not the deadliest, certainly one of the most expensive.
- **Zimbabwe** emerged from its own years of war and revolution towards a slightly more stable government, a fact that can be seen in the increase in life expectancy.
- **Germany**, the **U.S.**, **Mexico**, and **Chile** were all deeply affected by the Great Recession in 2008, though the German economy does not appear affected when viewed through GDP only.
-

The data from the World Band and the World Health organization is reliable, but very limited. GDP tells one very little about the daily life of a countries citizens. The U.S. had a GDP even during the Great Recession, where many of its citizens struggled to find employment and inequality skyrocketed.

Also, by including only two socio-economic data sets a reader is more likely to find a correlation between the two. If additional data sets were used, a broader, and more accurate picture would emerge. The way it's currently put, the relationship between GDP and Life Expectancy appears direct, when the reality is, of course, much more complex.

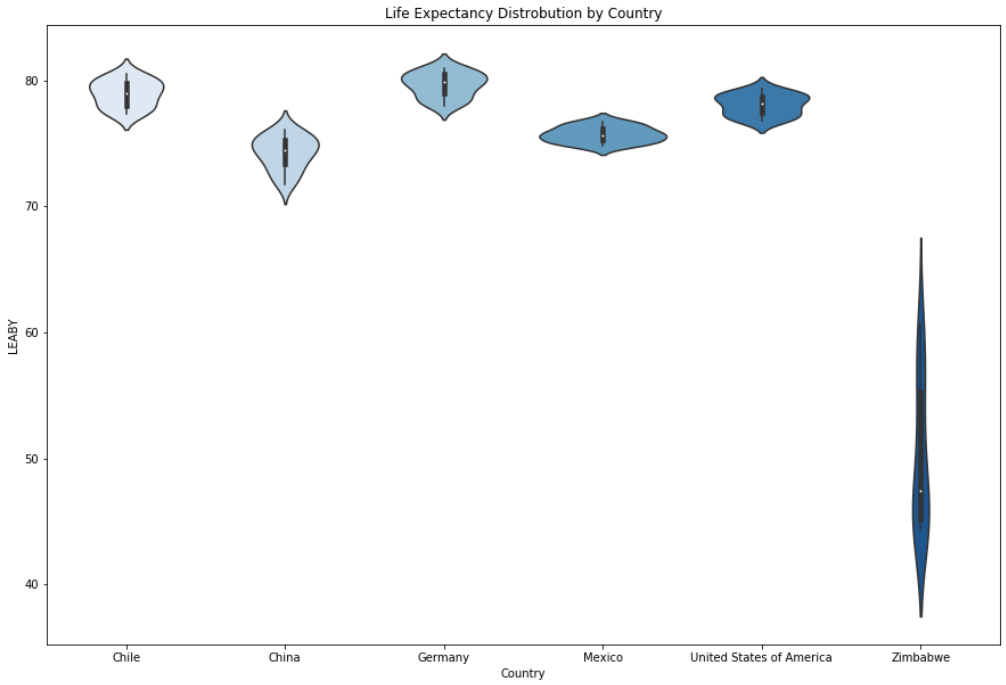

Figure 1: Violin Plot of the Life Expectancy Distribution by Country (2000 – 2015)

The following scatter plot reviews Life Expectancy at Birth through a movement of GDP over the years of this study. The chart is limited because the discrete charts make trends hard to follow. The colors help, but overall, this data is better expressed as a line graph (see figure 3).

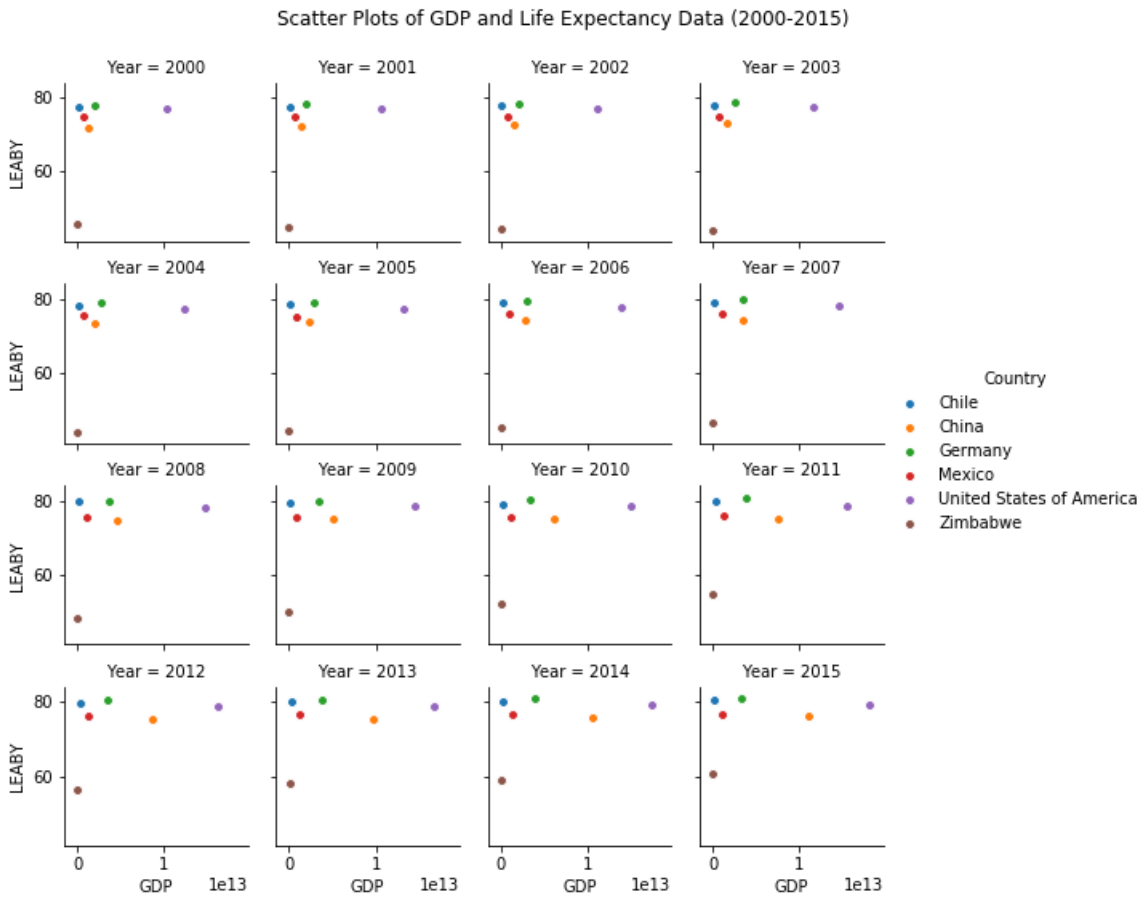

Figure 2: Scatter Plots of GDP and Life Expectancy Data (2000-2015)

The two line graphs below (Figures 3 and 4) are better representations of Life Expectancy and GDP. First, they do not falsely correlate data, they are simply a chart with one data set. Second, it is much easier to follow by country rather than by year.

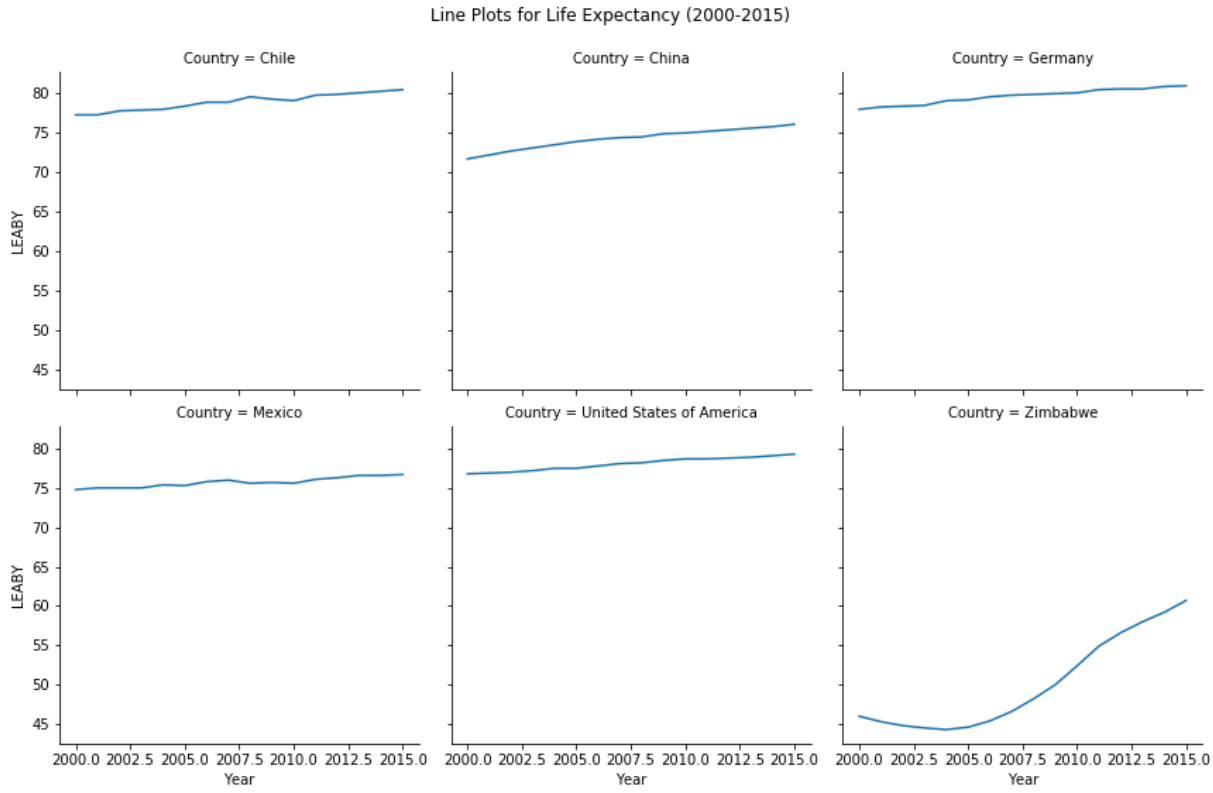

Figure 3: Line Plots for Life Expectancy at Birth (2000 – 2015)

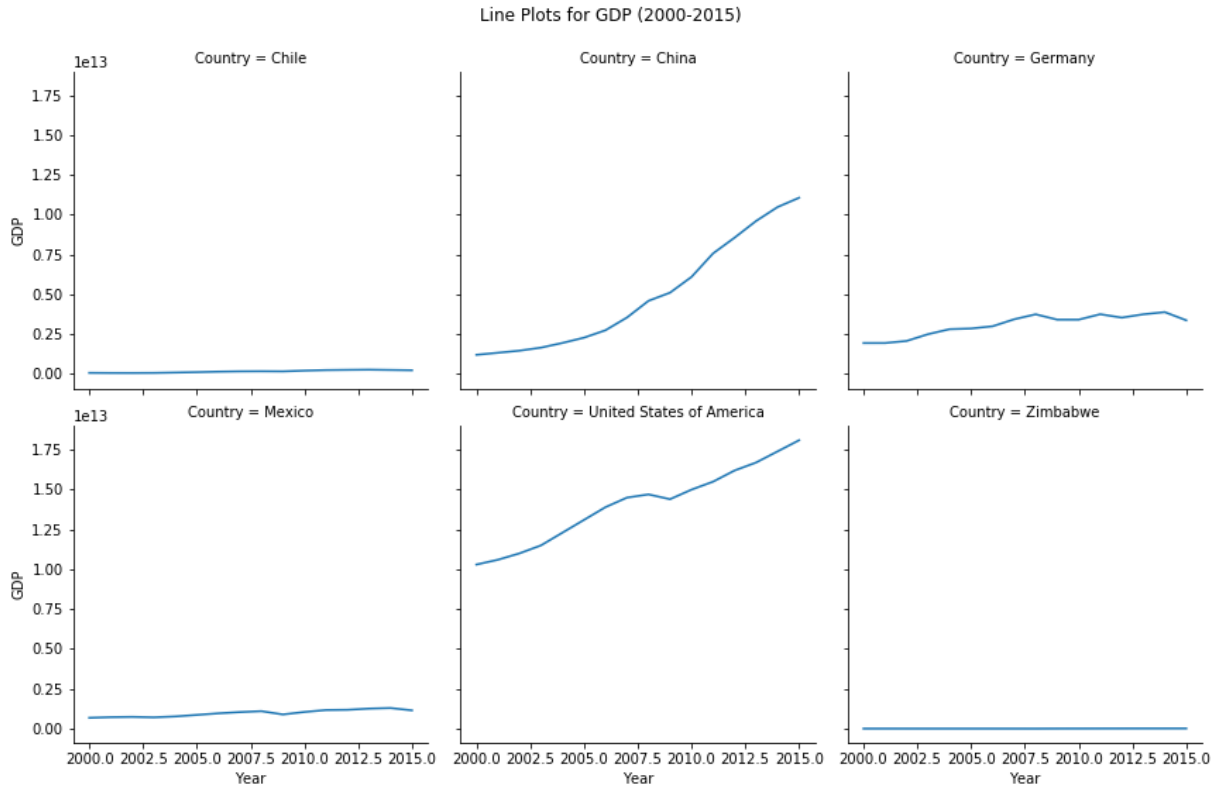

Figure 4: Line Plots for GDP (2000-2015)

For a full output of the project see the following link:

Life Expectancy Project-2