Questions:

Does a country's GDP determine its life expectancy?
If GDP does not *determine*, but simply *affects* life expectancy, how strongly does it do so?
What factors might contribute to the strength of GDP's influence on life expectancy?

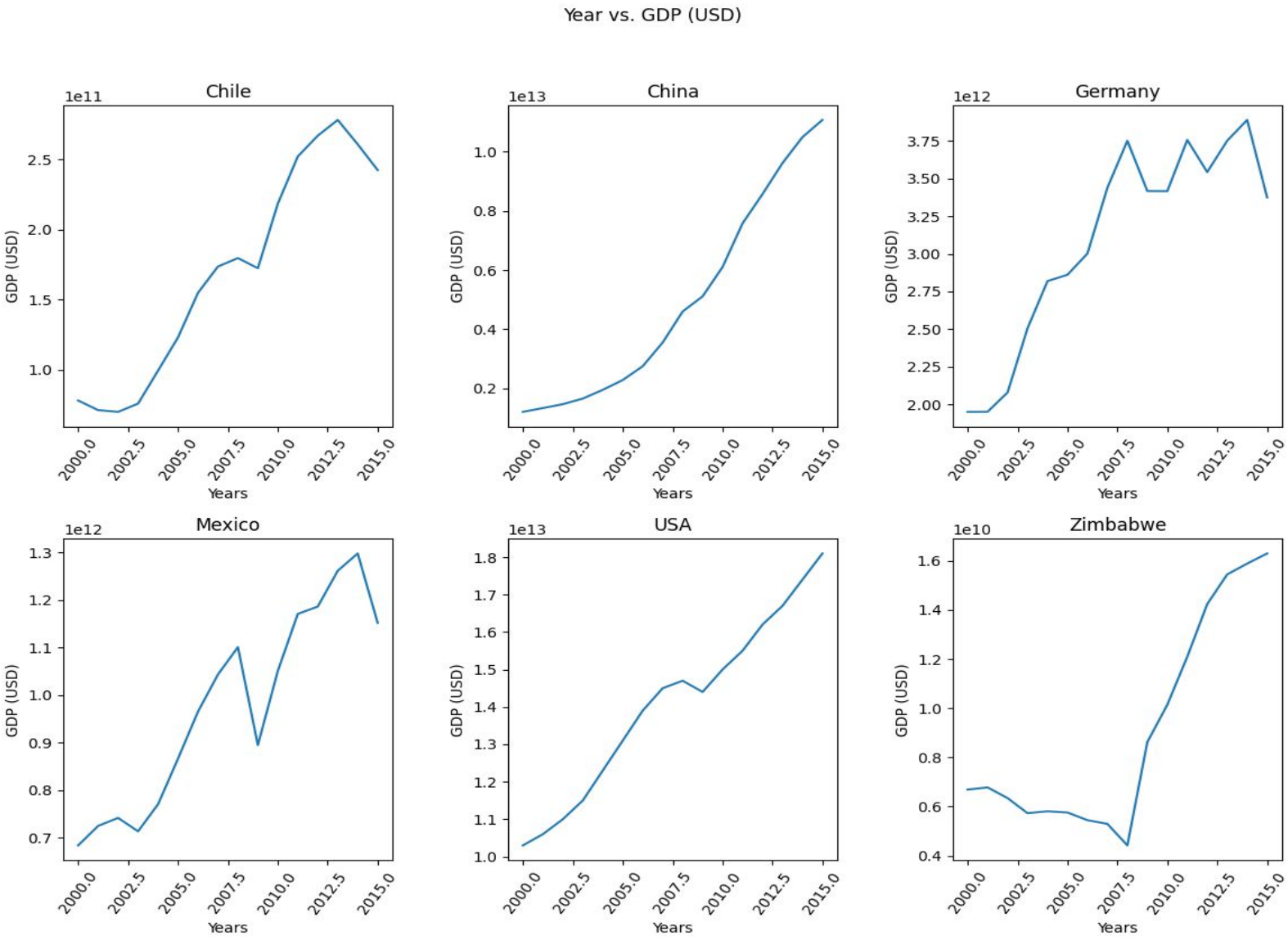
Abstract:

Using Python 3 with Pandas, Matplotlib, and SKLearn libraries, this study concludes that GDP has a strong influence on life expectancy, but is not absolutely determinative. Visual comparisons and linear regression suggest other unseen factors contribute to the variations in the data. Further studies into Germany's practices and into Zimbabwe's history between 2007 and 2015 might yield additional insight into boosting life expectancy.

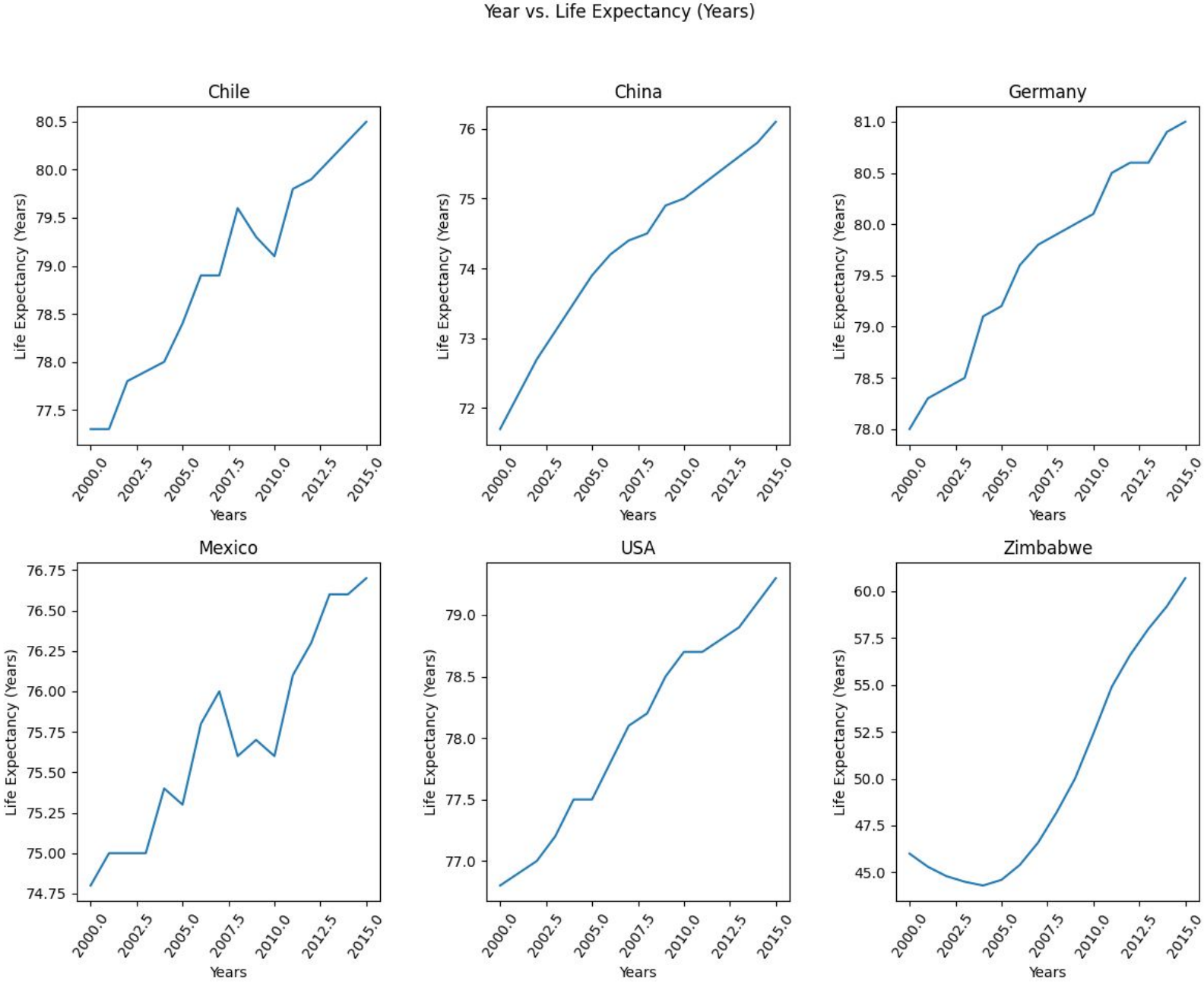
Details:

After importing our data and writing the appropriate functions needed for analysis, we can now get to our visualizations. Let's start by getting a sense of each country's GDP and Life Expectancy.

Here is a figure showing GDP over time for each country. These plots are designed to show the shape and magnitude of each country's GDP. Since the scope of GDP data is different on whole orders of magnitude, comparing them all on the same scale would not give a meaningful visual. Here, however, we can see such trends in the country's wealth as Zimbabwe's crash then renewal, Mexico's mid-range dip, Germany's rise and steadying out, and so on.

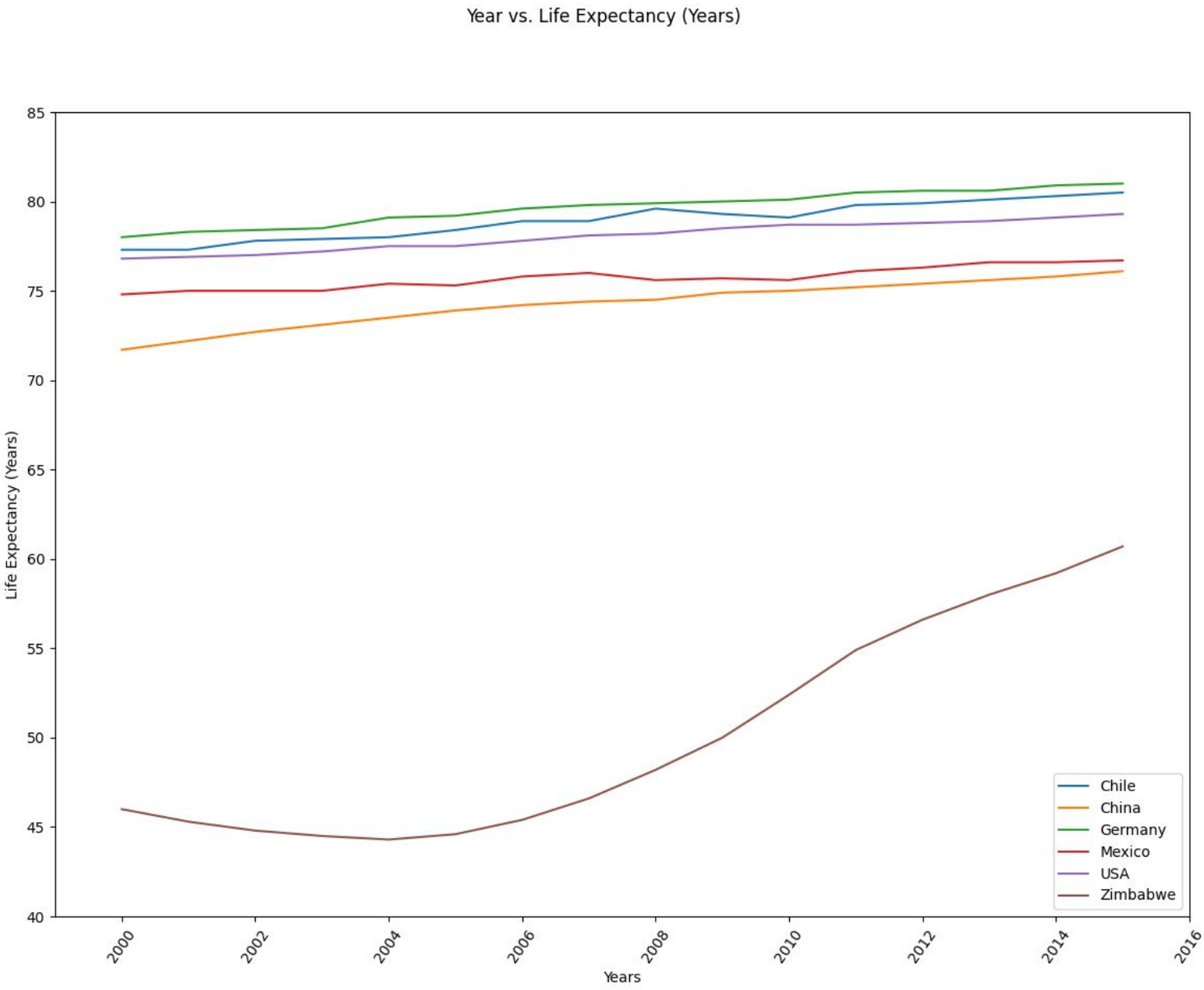


Next we might look at each country's life expectancy over time. Since we are investigating whether (and how much) GDP affects life expectancy, we want to keep in mind how similar these graphs look to the previous. If there is a correlation between the two, we would expect the graphs to look similar.



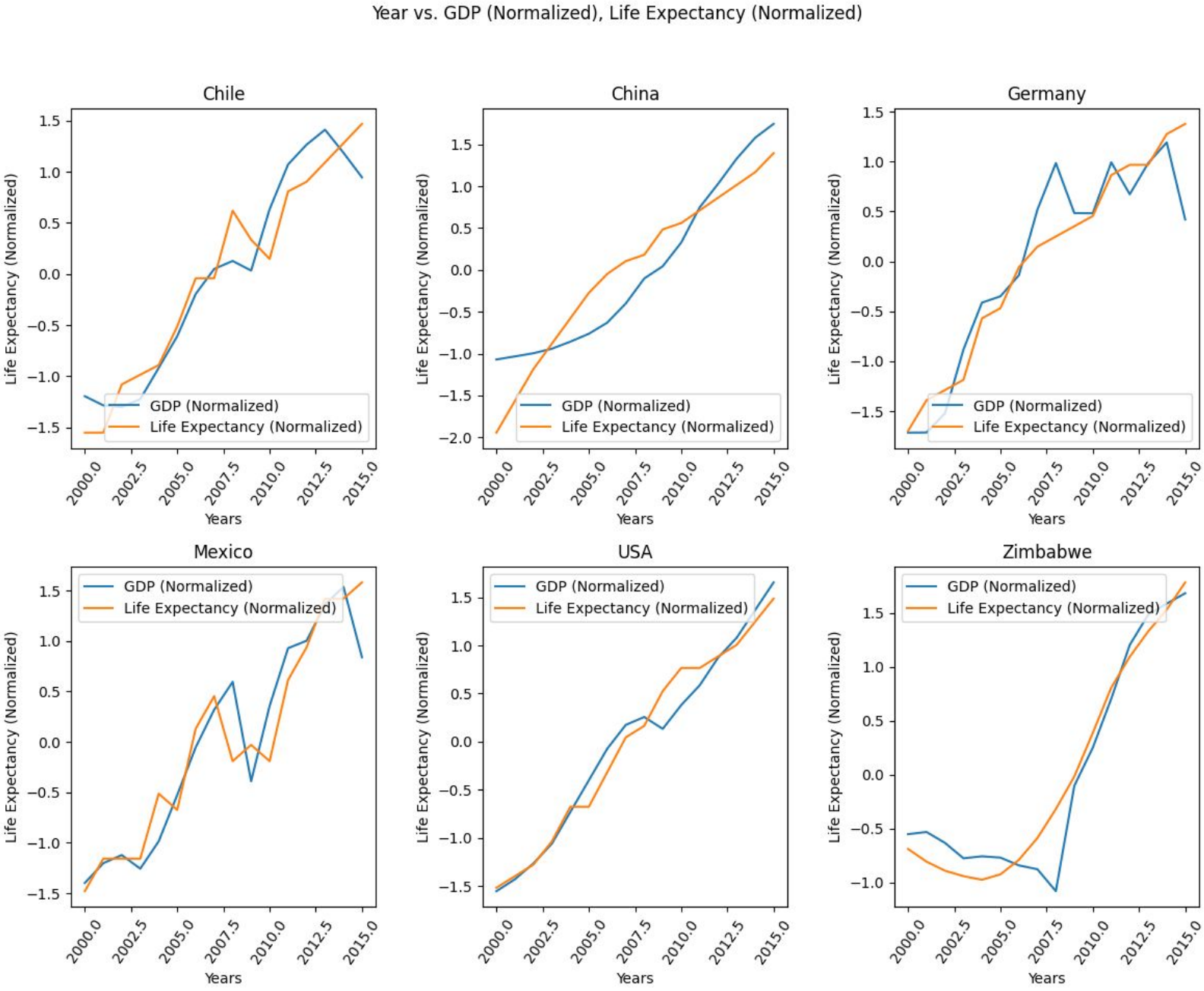
The plots indeed look similar, but not identical. Just as in the GDP graphs, here we can see Zimbabwe's crash and Mexico's dip, but we do not see Germany's life expectancy steadying as did its GDP. Chile and Mexico each showed a dip in GDP in the latest years which is not reflected in their life expectancy. So far, the comparisons suggest that GDP generally affects life expectancy, but is not entirely determinative of it.

Next, we will plot together all the life expectancy data over time to get a visual of how each country compares to the others.



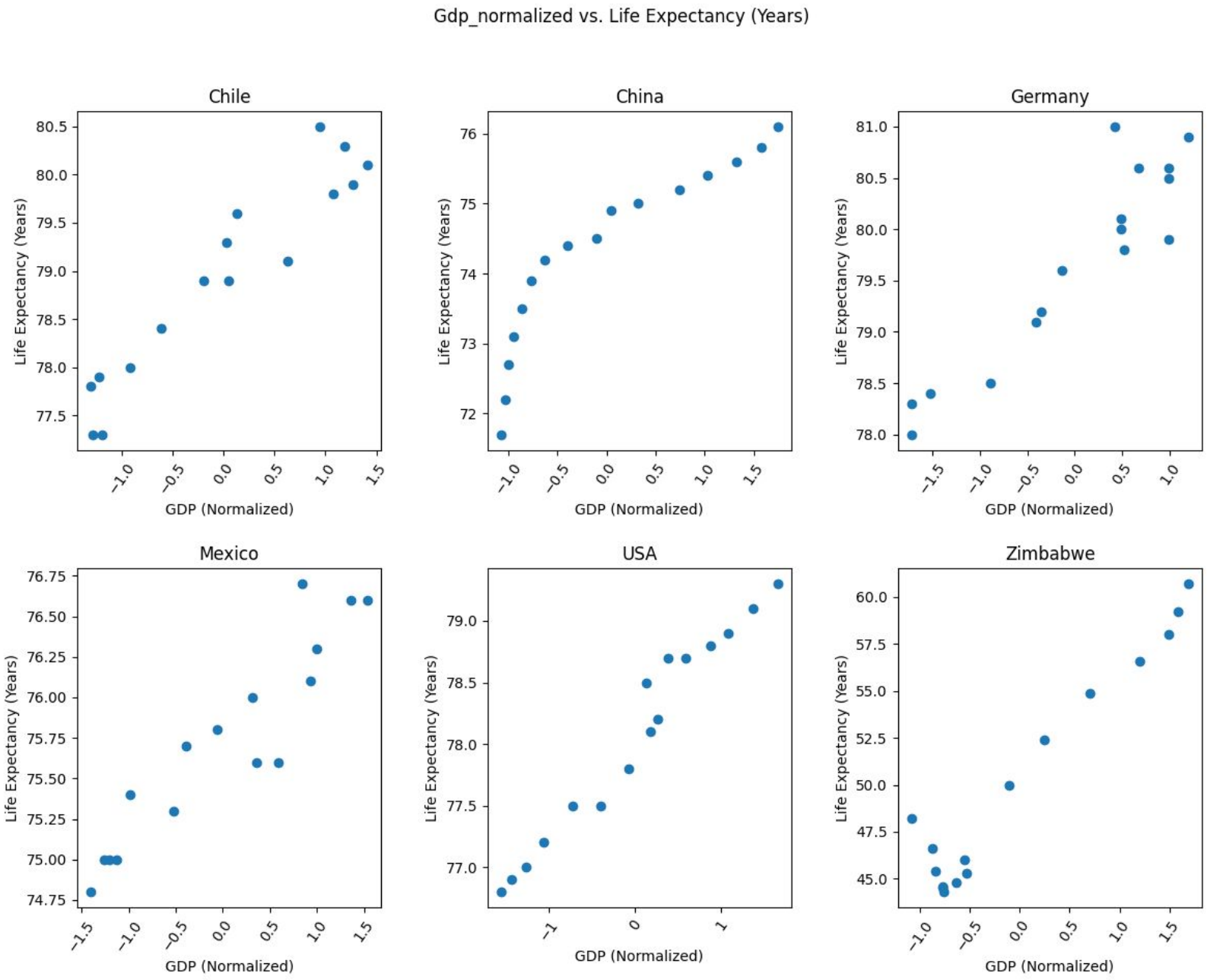
All on the same scale, we can see that most nations in our data have life expectancies in the 70-80 year range, with Germany consistently having the highest. Zimbabwe, on the bottom, has anywhere between 10-25 years below the top grouping. This graph also shows that, while the lowest overall, Zimbabwe showed the greatest *improvement* of life expectancy than any other country.

Now we will plot normalized sets of GDP and life expectancy data for each country side by side to see how well each set of data follows the other.

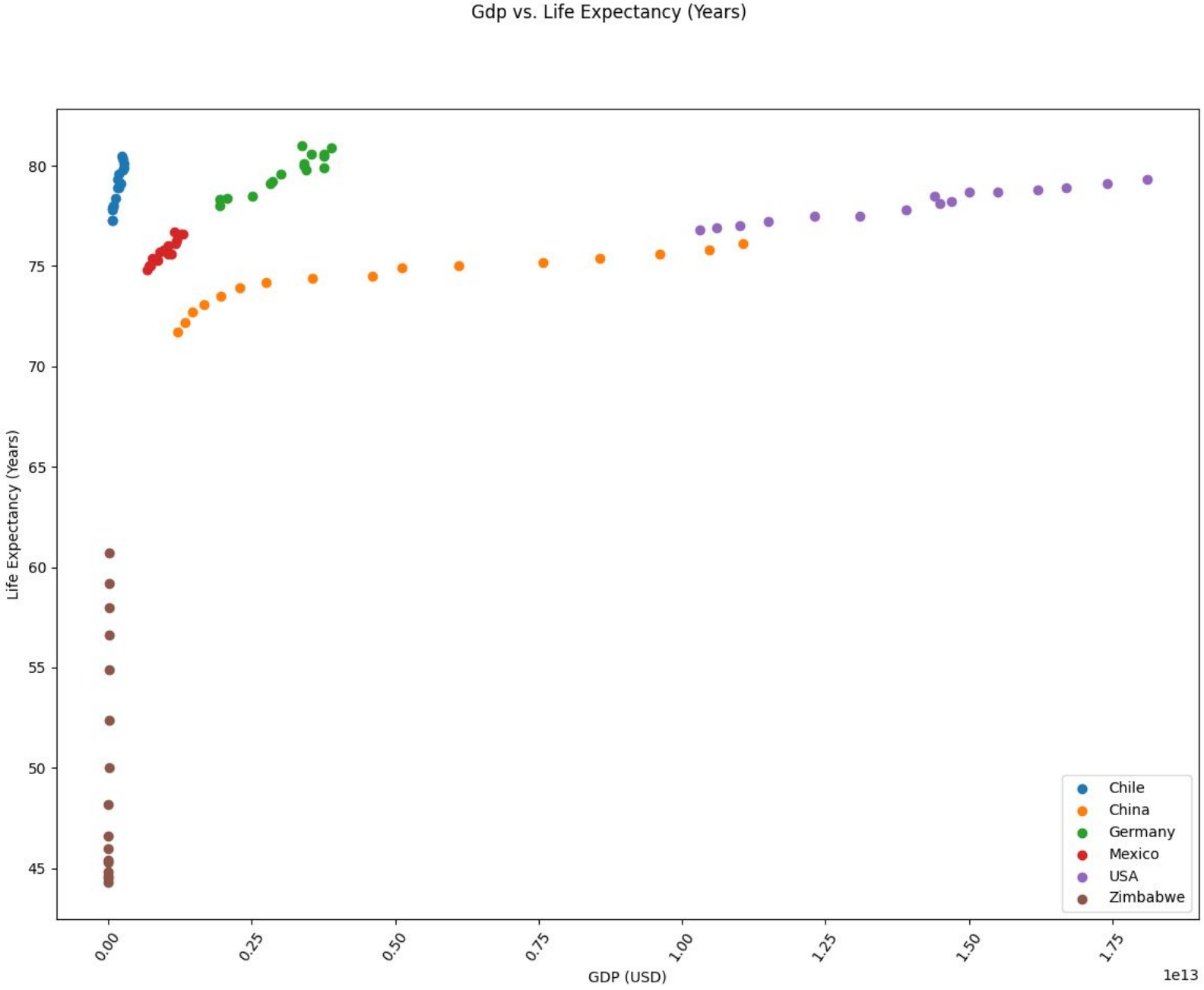


This powerful visual shows how well life expectancy (orange) follows GDP (blue). Chile, China, Germany, and the USA each have a general trend upward that show a positive correlation, but the anomalies in the trends for Zimbabwe and Mexico that are followed by *both lines* show strongly that these two sets of data are related.

Another way to visualize this is to make a scatter plot with GDP against life expectancy for each country. In such plots, correlated data would show a positive linear trend. Let's have a look:



All of these plots show a positive linear relationship, so we can conclude from these visuals that GDP and life expectancy are correlated. At this point, however, we might ask: *how* correlated? Is GDP *determinative* of life expectancy, such that a certain GDP will (by some natural human tendency) always yield a certain life expectancy? For example, does reaching $1 billion in GDP *always* ensure a life expectancy of at least 80 years (absolute determination), or does such a GDP merely support a higher life expectancy with other factors significantly accounting for variances (strong correlation)? To answer these questions, it will be helpful to make a scatter plot of GDP vs life expectancy for all nations together.



This graph is telling. It seems that achieving a specific GDP does not always ensure a particular life expectancy. This is very evident from the fact that the data does not all follow a single curve. Rather, we see different life expectancies for comparable GDPs. For instance, China had comparable GDPs to both Mexico (~$ 2 billion) and Germany (~ $3 billion), while its life expectancy was around five years lower in each case! Furthermore, the USA achieved a far higher GDP than most other nations, and still had lower life expectancies than Chile and Germany. This is especially surprising with Chile, whose life expectancies are similar to that of the USA, but with GDPs around 100 times smaller!

Clearly GDP strongly affects, but is not entirely determinative of, the country's life expectancy. This is further reinforced by performing linear regression on each country's data, and examining the coefficients produced. Taking normalized GDP as the feature (x values, independent variable), and life expectancy as the label (y value, dependent variable), the code produces the following output:

Country: Chile -- R^2: 0.89843 -- coef: 1.06362
Country: China -- R^2: 0.80527 -- coef: 1.23000
Country: Germany -- R^2: 0.98116 -- coef: 0.92202
Country: Mexico -- R^2: 0.94617 -- coef: 0.56338
Country: USA -- R^2: 0.92683 -- coef: 0.80684
Country: Zimbabwe -- R^2: 0.94671 -- coef: 5.66278

Such high R^2 values show that the life expectancy data is well accounted-for by the GDP values. Generally (except for China), 90% or more of the variance in the life expectancy data is explained by the GDP. This is indeed a strong correlation! The coefficients, however, show that the *strength* of GDP's influence on each country's life expectancies are different. For example, Zimbabwe was far more affected by its GDP (5.67) than was Mexico (0.56). In other words, the *growth* of GDP in Zimbabwe boosted its life expectancies far more than did Mexico's. In seven years, Zimbabwe added around $10 billion to its GDP and about 13 years to its life expectancy. The USA, on the other hand, added around *$30 trillion* to its GDP and only one year to its life expectancy in those same seven years.

Why would that be?

One reason could be that the relationship between GDP and life expectancy might be less linear and more asymptotic, such that increases in GDP at lower values would yield higher gains in life expectancy, which eventually approach a limit (~ 80 years) at increasingly higher GDPs. We would need more data to test this hypothesis, namely data at lower GDPs, and data that fills the gap between Zimbabwe's values and those of the higher group.

Another reason could be other factors not included in our data. What developments other than GDP did Zimbabwe make in those years? Beyond the GDP generated by a country within itself, such outside factors as relief from war, changes in government and religion, food and financial relief, and medical technology gifted from other nations might boost life expectancy. If such factors did contribute to life expectancy, and GDP was similarly given space to grow as a result, then we may be giving too much credit to GDP.

Conclusions:

We may conclude from this study that GDP certainly has a strong effect on life expectancy, but is not the only determining factor.

To gain further insight into boosting life expectancy, we might look at the differences between Germany's and the USA's lifestyles, private enterprises, and social programs. Given that Germany's GDP is lower than the USA's while boasting the top life expectancy, such a study might impart wisdom for optimizing GDP utilization to yield higher life expectancy results.

We might also look at what other factors besides GDP were present in Zimbabwe between 2004 and 2015 that boost its life expectancy so greatly. Such insights could be of use

to other countries in a similar position. Knowing these factors could improve our knowledge, and thereby help us better ‘engineer’ higher life expectancy for ourselves and others.

For further study, we might look at why 80-81 seems to be the limit on average life expectancy. Since this point (~80 years old) is achieved by the lower GDP ranges (e.g. Chile) all the way up to the higher (e.g. USA), money does not seem to be the limiting factor. So what *is* the limiting factor for this age? Is this a *natural* limit? As in, despite our resources, humans just will not live much longer than this? Or is this a technological limit? That is, despite our wealth, and supposing our bodies can be pressed further, does the world’s present medical technology only support around 81 years?