Esmé Middaugh
STATS 670, Mini Project #1
2018OCT04

GDP and Life Expectancy since WWII

"Can the increase in life expectancy since WWII be largely explained by increases in GDP per capita?" The simple, unexamined response comes quickly enough; as people have more money to spend they can afford better nutrition, healthcare, and living conditions leading to a longer life expectancy. However, analysis of data on global GDP and life expectancy since the end of WWII show that the the question—and possible explanations—are more nuanced and require deeper investigation. The nature of trivariate data (time, GDP, and life expectancy) coupled with vastly different life expectancy for subsets of our dataset based off of continent, necessitates three prongs of data exploration. First we explore the the relationship between gdp and life expectancy by looking at a single year, 2007. Next, we examine how life expectancy has changed over time by continent. Third, we attempt to explain the relationship between GDP life expectancy over time, guided by the knowledge gleaned from the first two explorations.

Before starting into actually analysis, a brief review of our dataset: we are using the gapminder dataset, a subset of data taken from the World Health Organization containing information on GDP and Life Expectancy from 1952-2007. Specifically, the following variables are available: country, continent (Europe, The Americas, Asia, Africa, and Oceania), year (ranging from 1952 to 2007); lifeExp (average life expectancy by country, pop (country's population) and gdpPercap (Gross Domestic Product per Capita). Now, to the analysis.

GDP and Life Expectancy

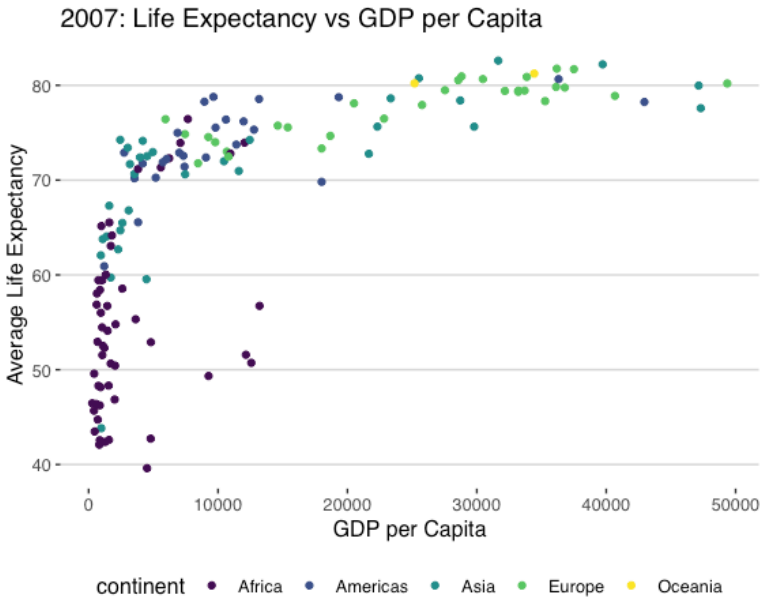

Figure 1

A simple scatter plot with x=gdpPercap and y=lifeExp shows that countries with higher GDP generally have a higher Life Expectancy (Figure One). However, the scatterplot also reveals that the relationship is not linear. As such, fitting a model is difficult; trends cannot be described by a simple linear model, box_cox transformations did not prove helpful (some normalization occurred with tau of 4, but that lacked interpretability).

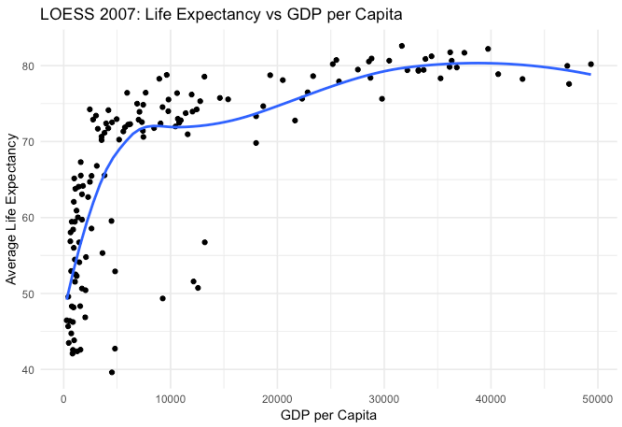

Figure 2

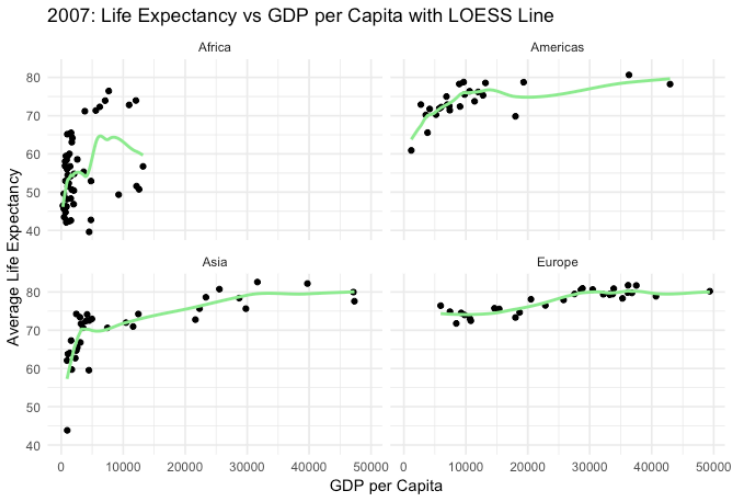

Figure 3

While an initial LOESS model looked like it may be a good fit (Figure 2), examining the residuals (Appendix, Figure 1) showed that there was distinct structure remaining, particularly for African countries. Because of this, it seems prudent to examine the 2007 life expectancy vs. GDP by continent. Faceting out by continent (with Oceania taken out due to lack of data) it becomes clear why fitting the LOESS line was not easy; Africa has a much different distribution than Asia, Europe, or America, with differences not explainable by simple additive or multiplicative shift (even among African countries there appears to be two distinctly different distributions, see Figure Three). While there are differences in distribution between the Americas, Asia, and Europe, there distributions seems similar enough to analyze together. As such, I fitted LOESS line to the remaining continents; we will analyze Africa more closely in our investigation of life expectancy over time by continent.

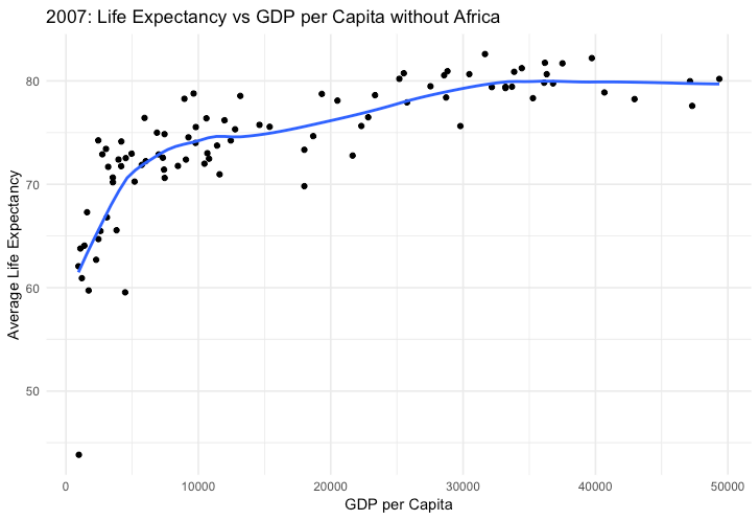

Figure 4

With Africa taken out of our data (Figure 4), two items are apparent. First, that it is much easier to fit a model. A LOESS curve provides a good fit (see Appendix Figure 1 for graph of residuals). Second, while there is an overall increase in Life Expectancy as GDP rises, it is not as steep of a slope as it appeared when Africa was included. There is a sharp increase in Life Expectancy as GDP rises from roughly $0 to $10,000 (it should be noted that the data here comes mostly from Asian and American countries), a slower increase from ~$10,000 to $30,000 and after that a mostly flat relationship. From this model we can generalize that positive relationship between GDP and life expectancy, with the most impact being for increases near

the beginning of the GDP scale, and that as countries get wealthier the relationship between GDP per capita and life expectancy starts to level off.

Life Expectancy over Time by Continent

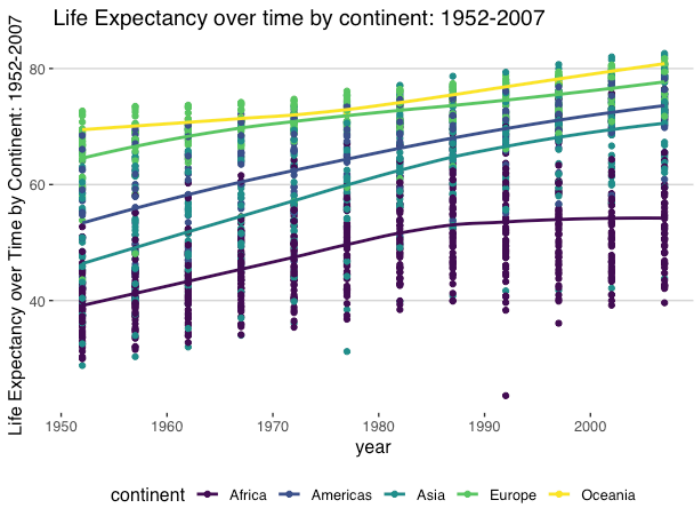

Figure 5

As a first look at how life expectancy has changed over time, I tried plotting by color, with a smoother over the top to help show trends. While this graph was severely over plotted, it also revealed specific outliers which I investigated more in depth. For each of the outliers at 1977, 1992 and 2002. there was a genocide in Cambodia, genocide in Rwanda, and civil war in Sierra Leone, respectively.

Looking instead at the weighted mean of the continents over time with a smoother for good value provides a clearer picture. Before overall analysis, a brief note on another meaningful outlier: the sudden decrease in Asia's life expectancy in 1962. After a little digging, the answer is apparent; from 1958 to 1962 Mao's 'Great Leap Forward' resulted in a famine costing 45 million lives, enough to bring down the life expectancy for the entire continent for that year (China's life expectancy went from 50 years in 1957 to 44 in 1962).

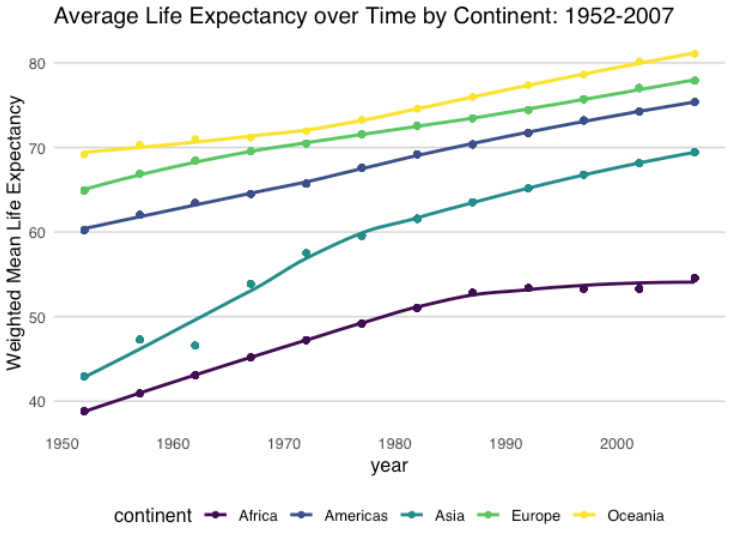

Figure 6

For all continents, life expectancy has gone up, and for the period from 1952 to the 1985 the gap in life expectancy was narrowing for all countries, with Asia's increasing particularly fast up till the late 1970s. Asia's growth begins to slow around 1980, though still closing the gap between itself and Europe, America, and Oceania, whose life expectancy is nearly linear over the entire timer period. Among Europe, Oceania, and the America's there has been a narrowing in the gap between their average life expectancy and while everyones 'position' has stayed the same, The Americas and Asia are now much closer to where Europe and Oceania are, with both surpassing Oceania's 1952 life Expectancy. Based off of our analysis of GDP and life expectancy in a single year, this is

expected; at a certain GDP level there stops being large increases in life expectancy as such a high life expectancy has been reached that it can't reach much higher. As continents get closer to this level it makes sense that the rate of increase in their average life expectancy begins to slow.

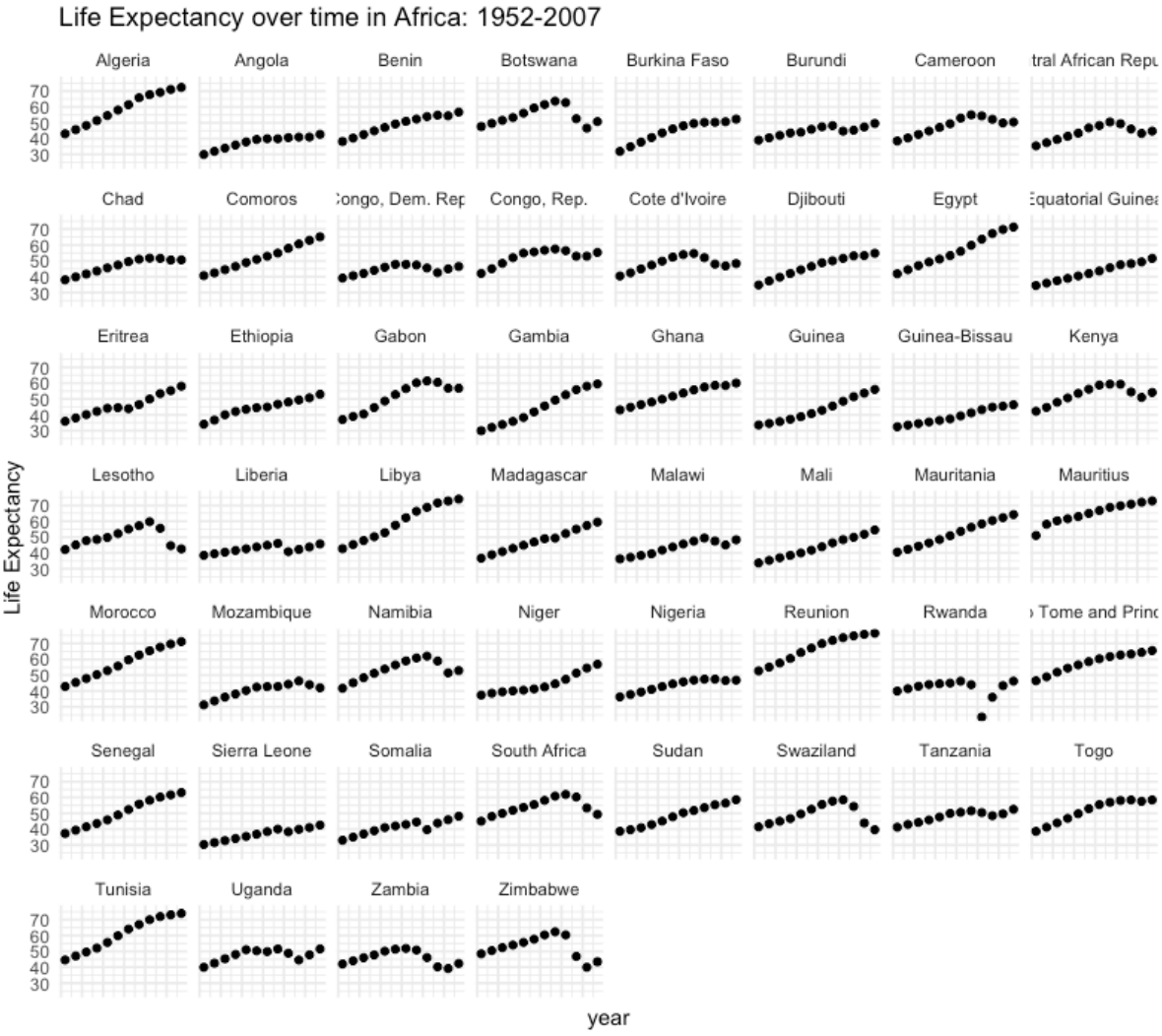

Figure 7

What is worrisome however, is that in 1985 the increase in Africa's life expectancy starts to level off entirely, at only around 54 years. In alignment with our findings from the earlier attempt at fitting a LOESS model, when faceting out by country there appears to be two distinct tracts among African countries; those whose life expectancy has continued to rise, and those for whom it has fallen (see Figure 7) since the 80s. Comparing the list of countries with the most severe decrease in life expectancy to a list of countries with the highest rates of AIDS in Africa it becomes all too apparent why those countries have had such a steep decrease in life expectancy, and why Africa's GDP vs Life Expectancy model had two distinct clusters. Countries experiencing the worst of the AIDS epidemic lack a positive relationship between GDP and life expectancy.

GDP and Life Expectancy over Time

To analyze the relationship between GDP and life expectancy over time I created two coplots, both graphing life expectancy against the log of GDP per capita (log taken to help with the wide spread). The first (Figure 8) explores life expectancy and GDP per capita given continent (color indicates different time frames), and is useful for analyzing how time has effected the relationship within a particular continent . The second (Figure 9) compares the relationship of Life Expectancy and GDP given time and provides a more in depth look at the differences between continents.

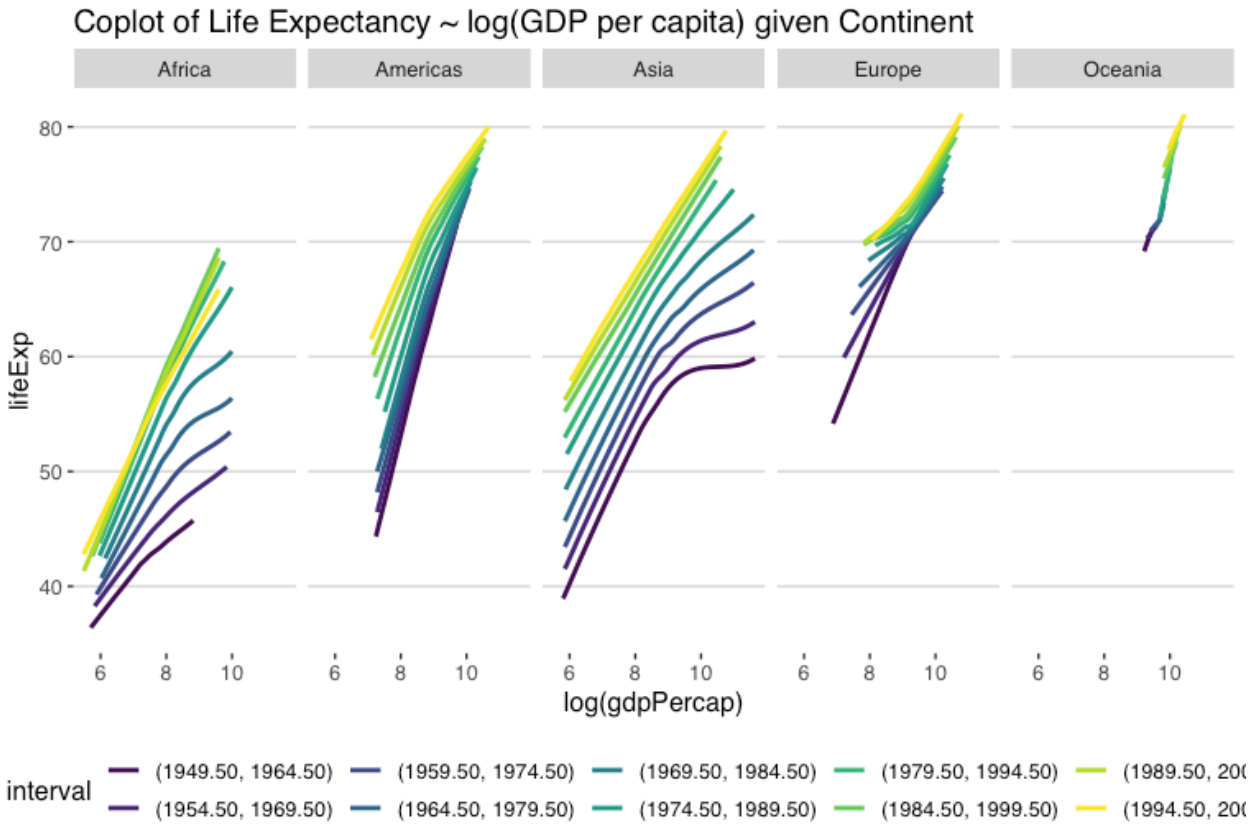

Figure 8

For each continent, the relationship between GDP and life expectancy has been slightly different. While for all there definitely appears to be an additive shift (the model for the most recent years is higher than for previous years for all continents), each has its own specific characteristics:

Africa: As discussed earlier, the relationship between GDP and life Expectancy is different in Africa than in the other countries. Perhaps unsurprisingly, it also has a more

complicated relationship with time; the slope and shape of the distribution changes over time, indicating that there is an interaction, in addition to an additive shift.
The Americas: Mostly an additive shift, although there also appears to be some leveling off of the slope in recent years as a higher life expectancy is reached.
Asia: The changing slope over time show that there is an interaction between GDP, Life Expectancy, and time; over time, the slope has gone from leveling off to a more linear shape.
Europe: There seems to be an interaction, as over time Europe's slope has gotten less steep. Because of this there seems to have been a convergence, with the range of Europe's life Expectancy getting much smaller and higher.
Oceania: While difficult to analyze (and see) as it has so few data points, it appears to have a mostly additive shift with time, nothing else.

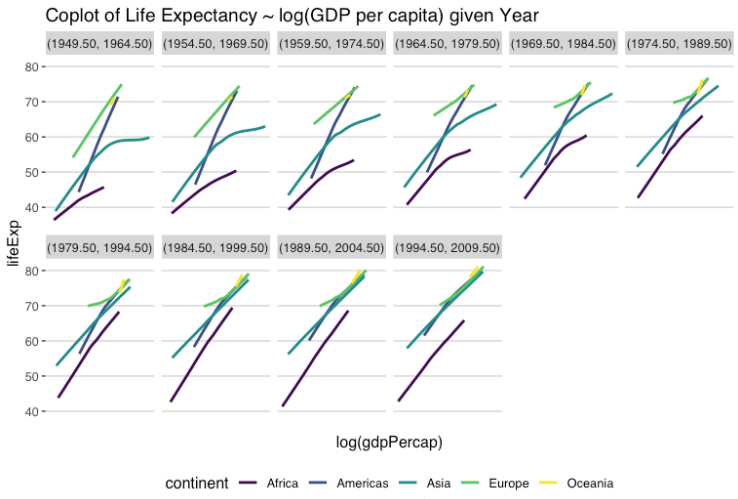

Figure 9

Figure 9 demonstrates the convergence in life Expectancy over time. Going from the upper left to the bottom right, there is a clear pattern of the models getting closer and closer together.

There is a time effect on life expectancy in addition to a GDP effect; from Figure 9 it is clear that for all continents there is both an additive effect and an interaction; over the different time intervals the intercept increases and the shape of the distribution changes for each continent.

Changes in life expectancy can by no means be entirely explained by changes in GDP per capita alone. While at a broad level increased GDP per capita correlates to higher life expectancy, life expectancy also seems highly dependent on continent, where within that continent it is located (for example, with the AIDS crisis in Africa), and also time. At the very least time seems to have had an additive shift effect, indicating that perhaps something else (better nutrition information, health practices, etc.) related to time has also been contributing. Time has had an additive effect, but for many continents there has also been an interaction with the relationship between GDP and life expectancy changing over time. Through exploration of GDP and Life Expectancy in a single year, life expectancy over time by continent, and the relationship of GDP and life expectancy over time it is safe to say that while high GDP and high life expectancy are certainly related, the increase in life expectancy since WWII cannot wholly be explained by increases in GDP.

Appendix

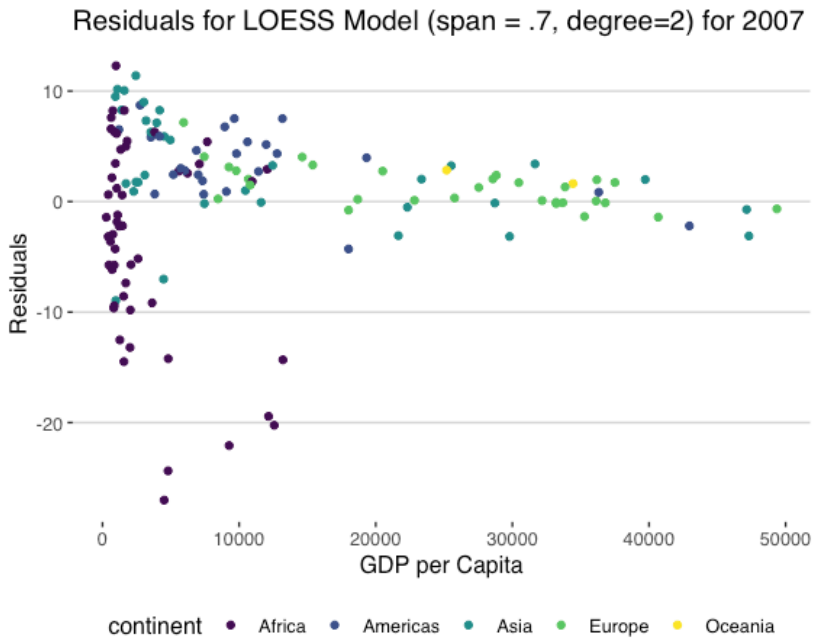

Figure 10

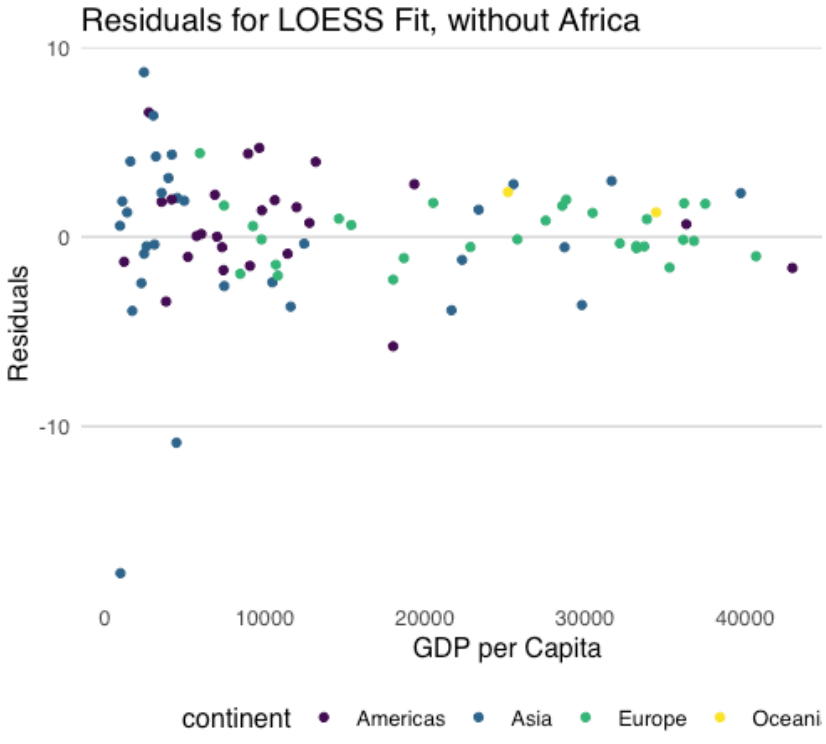

Figure 11

Sources:
http://endgenocide.org/learn/past-genocides/the-cambodian-genocide/
https://www.history.com/topics/africa/rwandan-genocide
https://www.hrw.org/news/2012/04/11/armed-conflict-sierra-leone
https://www.nytimes.com/2010/12/16/opinion/16iht-eddikotter16.html
https://en.wikipedia.org/wiki/List_of_countries_by_HIV/AIDS_adult_prevalence_rate