United Nations Human Development Index in Sub Saharan Africa: An Exploratory Analysis

The Human Development Index (HDI) is a statistic composite index of life expectancy, education, and per capita income indicators. The index number is a geometric average of life expectancy at birth, education index (which is a arithmetic average of expected years of schooling and mean years of schooling), and gross national income per capita. The original goal of HDI was to create a metric that would shift the development of economies towards people-centered policies. The HDI is viewed as a proxy for the quality of life in a country, and rapid change in HDI over time can be an indicator that a country is entering a new stage of growth and development.

Sub Saharan Africa contains some of the world's fastest growing economies, resulting in some of the largest increases in HDI over the past decade.

The datasets I chose to look at to better explore HDI in depth included, HDI over time, Life Expectancy (LE) over time, Education Index (EI) over time, Gross National Income per capita (GNI) over time, and Foreign Direct Investment (FDI) over time. LE, EI, and GNI are direct components of HDI, but FDI is an external factor I wanted to observe. Prior to exploring the data, I was most interested in observing the strength in correlations between HDI and the rest of the datasets.

I chose to focus exclusively on data from Sub Saharan Africa over the past decade.

Change in HDI over time by country
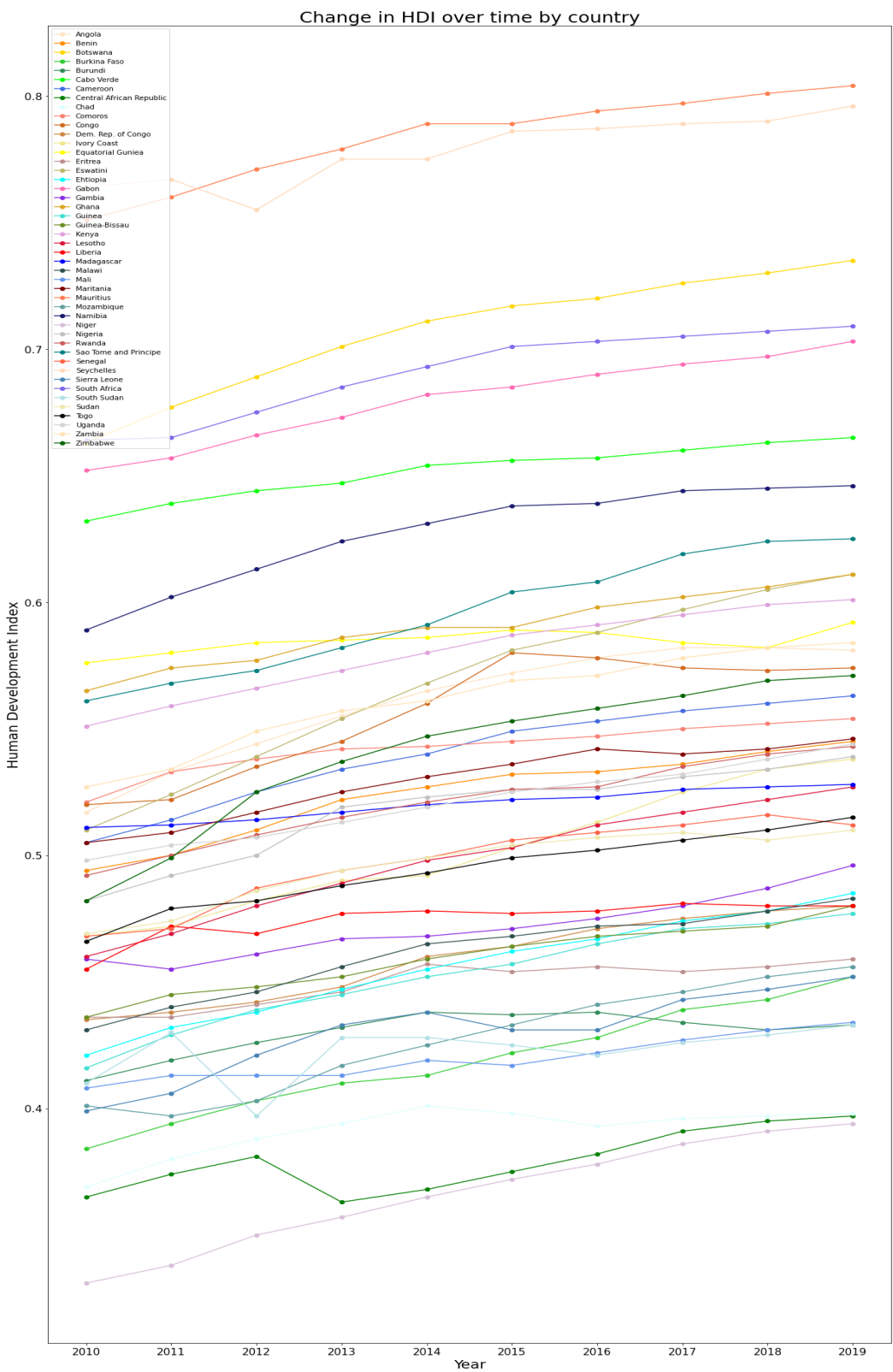
The first graph (pictured above) was meant to show the change in HDI over the past decade for each individual country. Countries with lines at the top have a higher HDI on average. When observing this graph, I became more interested in the lines that had the steepest slope, indicating the countries with the greatest change in HDI over the past decade. I decided to look into who the biggest growers were.

Total change in HDI from 2010 to 2019 by country
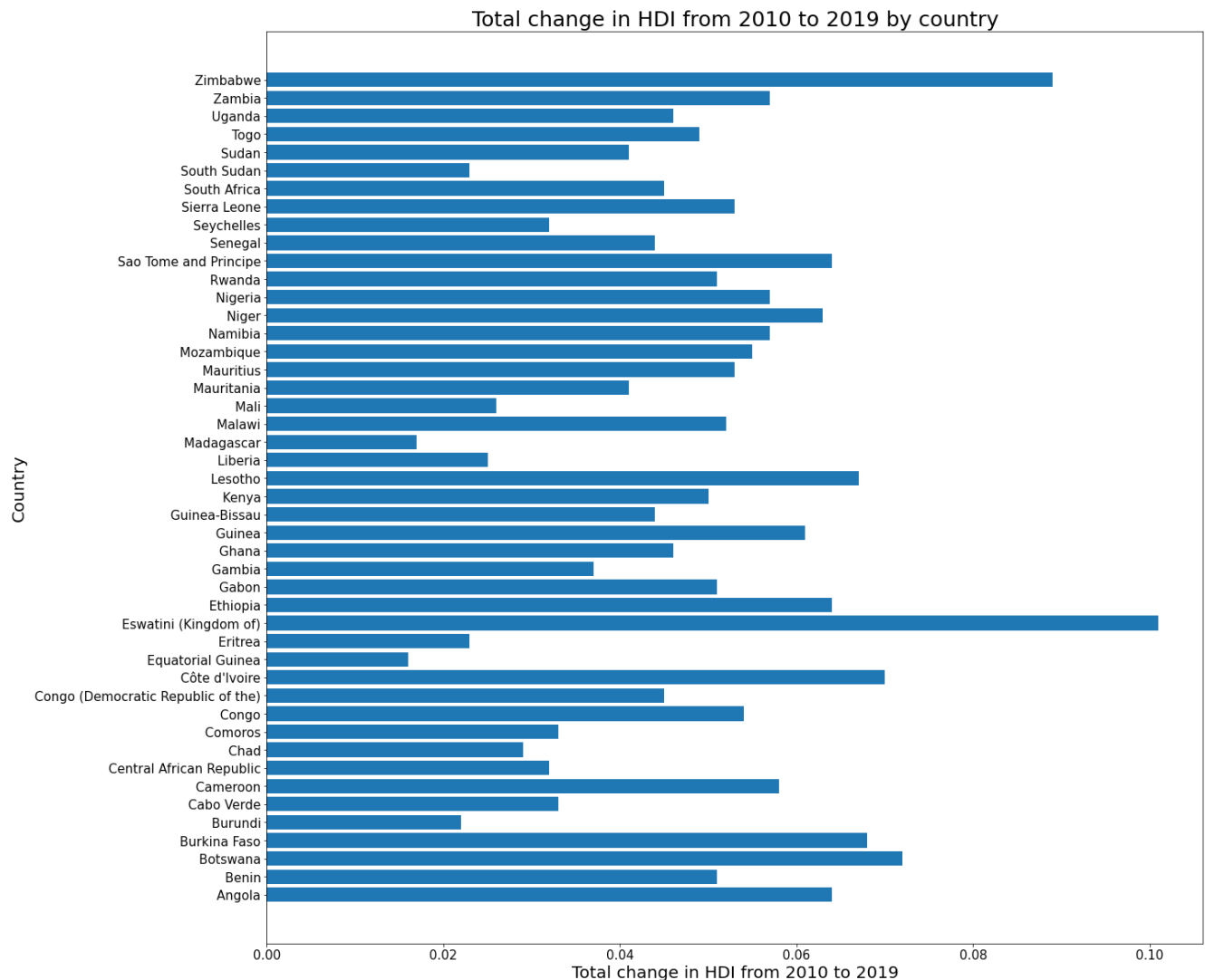
This graph clearly shows that Eswatini, Zimbabwe, and Botswana were the largest gainers in HDI over the past decade. This graph also shows that there was a large discrepancy in HDI growth throughout Sub Saharan Africa. This gives me the idea that there are vastly different growth strategies and stages occurring in the region

Sub-Saharan Africa Average Gross National Income Per Capita Over Time
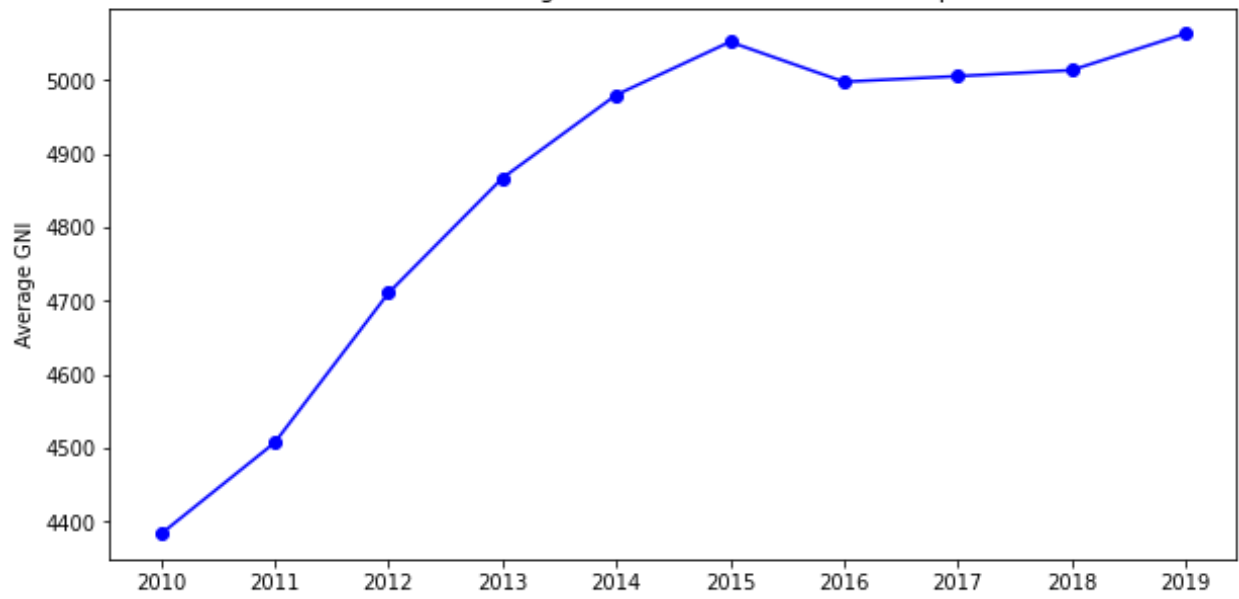
Sub-Saharan Africa Average Life Expectancy at Birth Over Time

Sub-Saharan Africa Average Education Index Over Time

Sub-Saharan Africa Average HDI Over Time
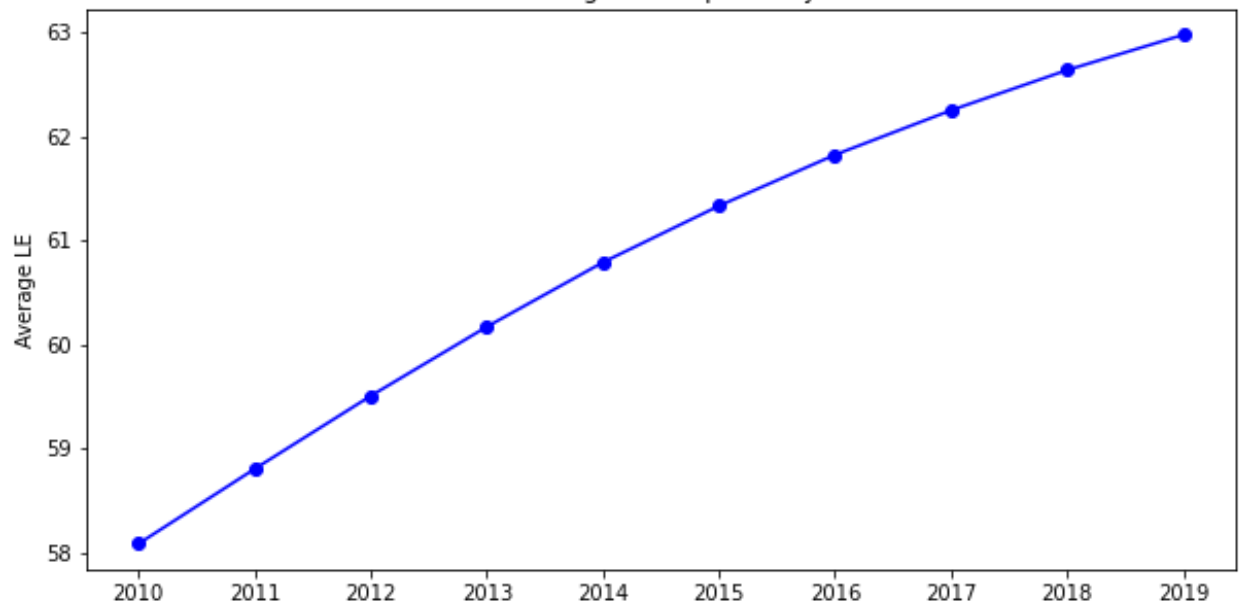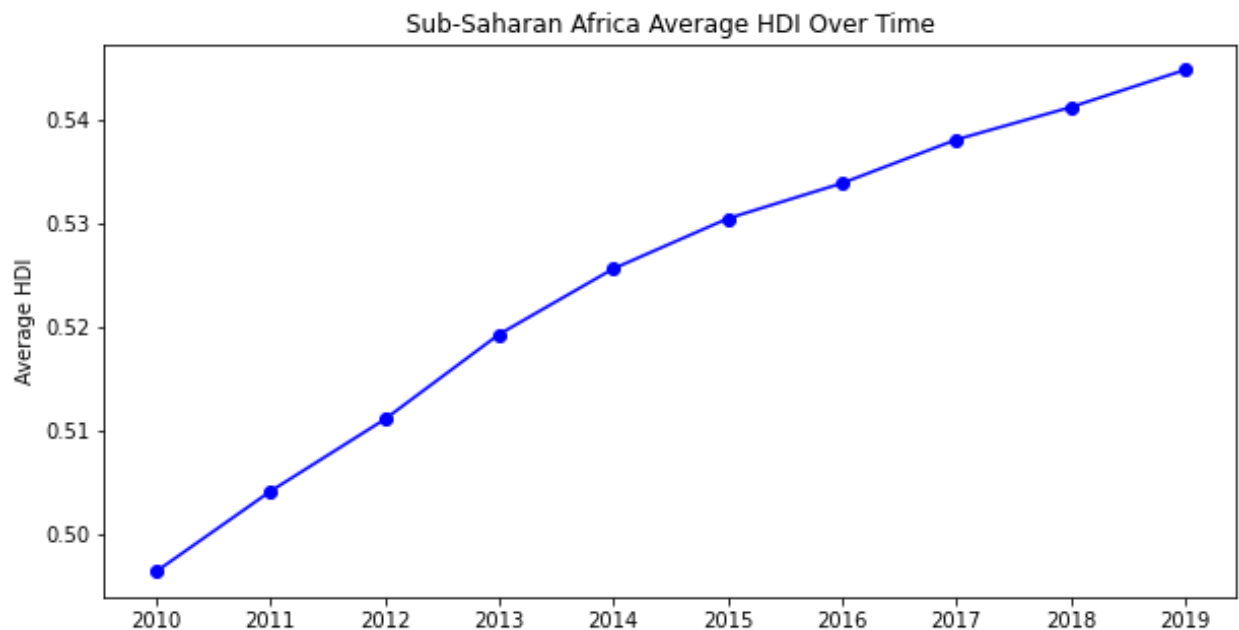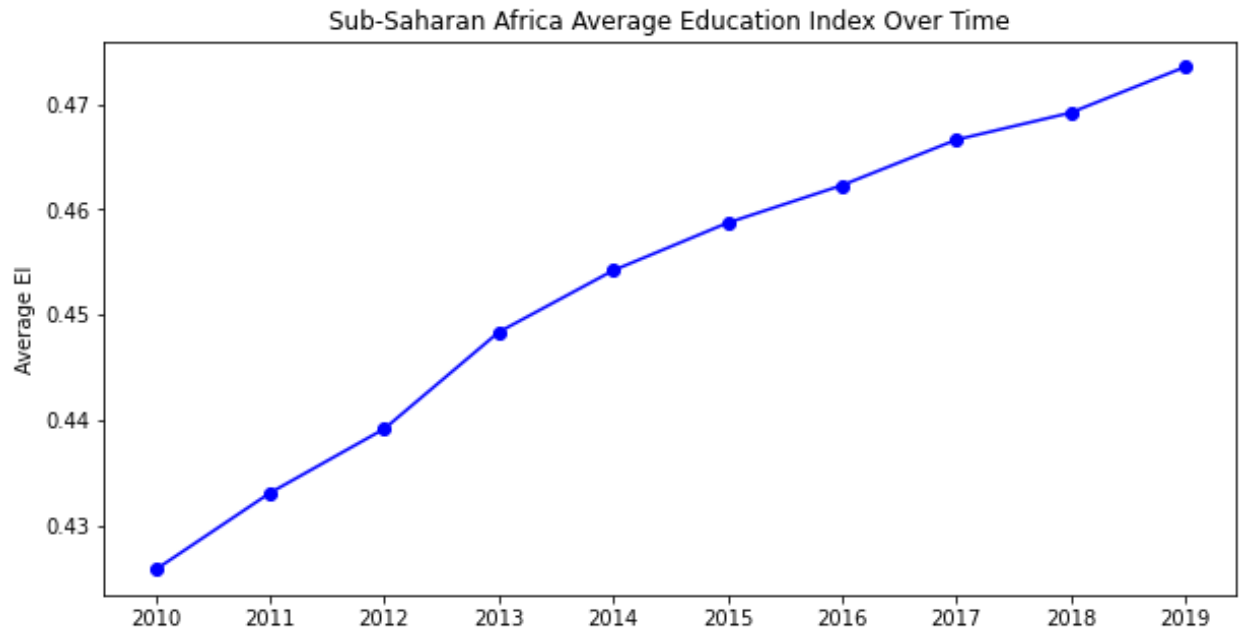
The next four graphs I decided to observe show HDI and each of its components (EI, LE, and GNI) over time for all of Sub Saharan Africa. I got the total averages for each year and projected them over time to observe the shapes of each graph. I wanted to see if there were any unusual trends. HDI sees a steady growth that slightly starts to increase by smaller increments over time. That is true for the components as well (as expected). GNI is the most volatile of the components which gives an indicator that it is most likely responsible for most of the variability in HDI. I also observe that it appears that EI follows HDI most closely. This could indicate that is the best proxy for HDI growth

Question Exploration and Hypothesis Testing:

After the data exploration portion, I discovered three questions and corresponding hypotheses I wanted to understand better:

Question 1: Which of HDI's components' increase over the past decade has the strongest correlation with HDI's increase over the past decade for the entirety of Sub-Saharan Africa?

Hypothesis 1: From observing the shape of the raw graphs of Education Index, Life Expectancy, and Gross National Income against the graph of HDI, I hypothesize that the education index has the greatest correlation with HDI increase over the past decade. This would tell us that improvements in education have had the strongest relationship with improvements in HDI over the past decade.

Question 2: Are the correlations for each of the decade's top HDI growers with the individual components significantly different from the aggregate or from each other?

Hypothesis 2: I would assume that the correlations will be slightly different, hypothesizing that the top performers were doing something differently than the rest of the pack. I think that the correlations between the top three would be more similar between each other than with the aggregate under the belief that they were doing something similar to achieve higher HDI growth, even if that just means hitting a certain stage in their growth.

Question 3: Are there any factors outside of HDI's components that could have stronger relationships to HDI's % increases for the top 3 countries? Specifically, Foreign Direct Investment?

Hypothesis 3: FDI will not have a better correlation that the strongest relationship with an HDI component for any of the top countries. Otherwise, I would assume that direct foreign investment is a better proxy for HDI than its own internal components.

Question 1:

Average Component % Change Over Time for All of Sub-Saharan Africa
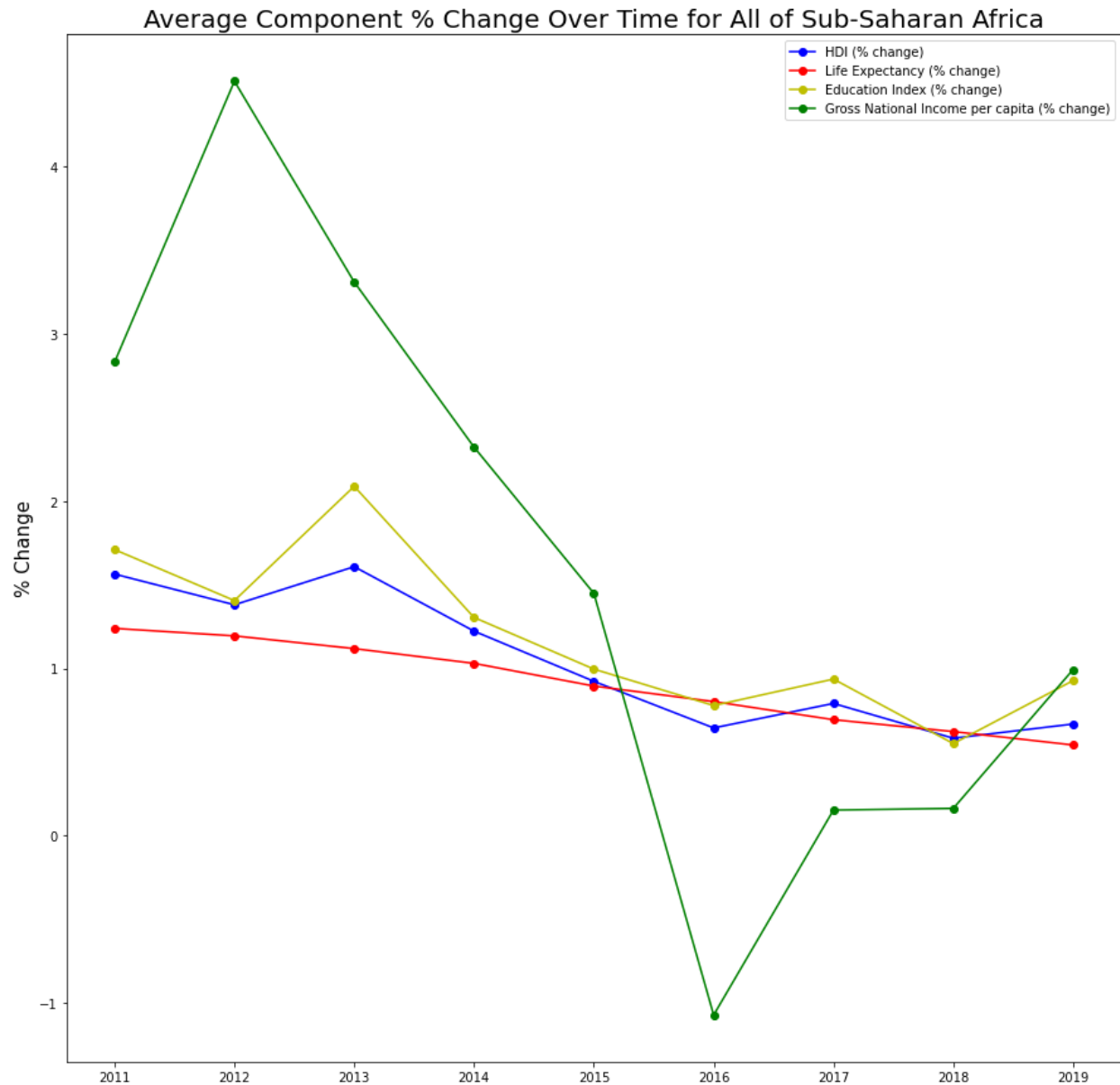
To answer question 1, I wanted to visualize the % change of HDI, LE, EI, and GNI on the same graph to better visualize the correlations between the individual components. The resulting graph shows the Life Expectancy and Education Index appear to move much closer with HDI than GNI does. To back up the visualization with data, I found the correlations of each component with HDI:

Education Index vs. HDI = 0.9584908743042713

Life Expectancy vs. HDI = 0.9320688268139777

Gross National Income vs. HDI = 0.8761916686066716

The correlations back up the visualization. The correlation proves my hypothesis to be correct. Education Index's % change over time has the strongest relationship with HDI's % change over time. It is complicated to deduce the logical reasoning behind that strong correlation because correlation does not equal causation. In the least, it shows that EI is the best proxy overall for HDI growth. This could mean that growth in education is the most reliable way to increase HDI over time.

Question 2:

Average Component % Change Over Time for Zimbabwe

Average Component % Change Over Time for Eswatini
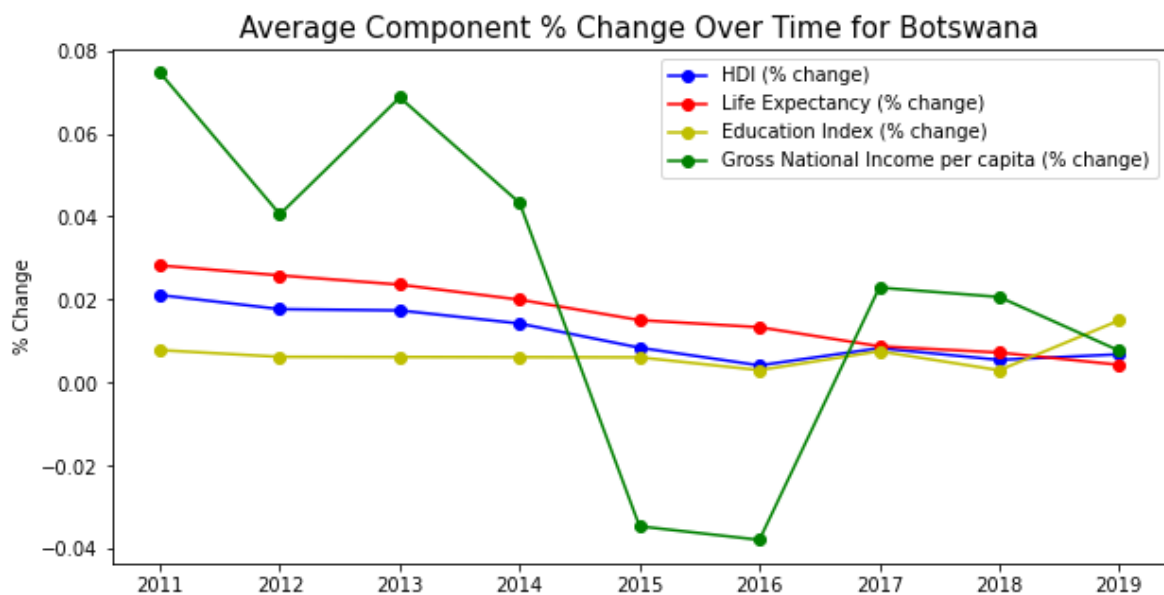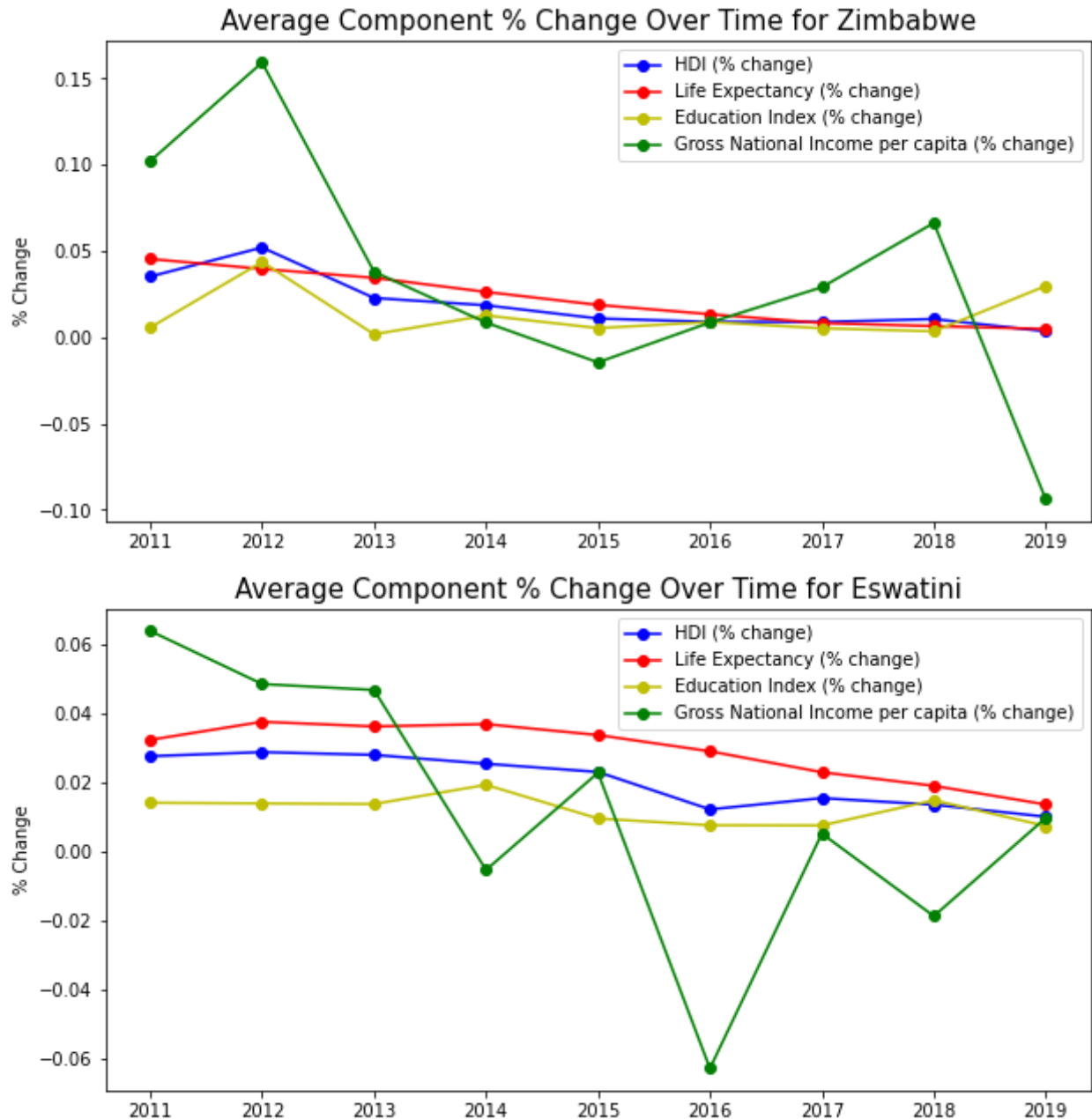
The three graphs above detail the same exploration of questions 1 but instead of the entirety of Sub Saharan Africa, I am only looking at the three largest growers of the past decade. I want to see if the is a large difference in the same correlations. From the visualization, the trend of GNI being the most volatile seems to hold, and LE and EI follow HDI much closer; however, there are clearly different trends in % growth for each country despite HDI following a similar pattern. Here are the correlations that correspond to give better insight into the graph:

Eswatini: EI vs. HDI = 0.6396033803699559

Eswatini: LE vs. HDI = 0.8819529991028152

Eswatini: GNI vs. HDI = 0.7463892582644187

Zimbabwe: EI vs. HDI = 0.4789898631105538

Zimbabwe: LE vs. HDI = 0.8740695691891619

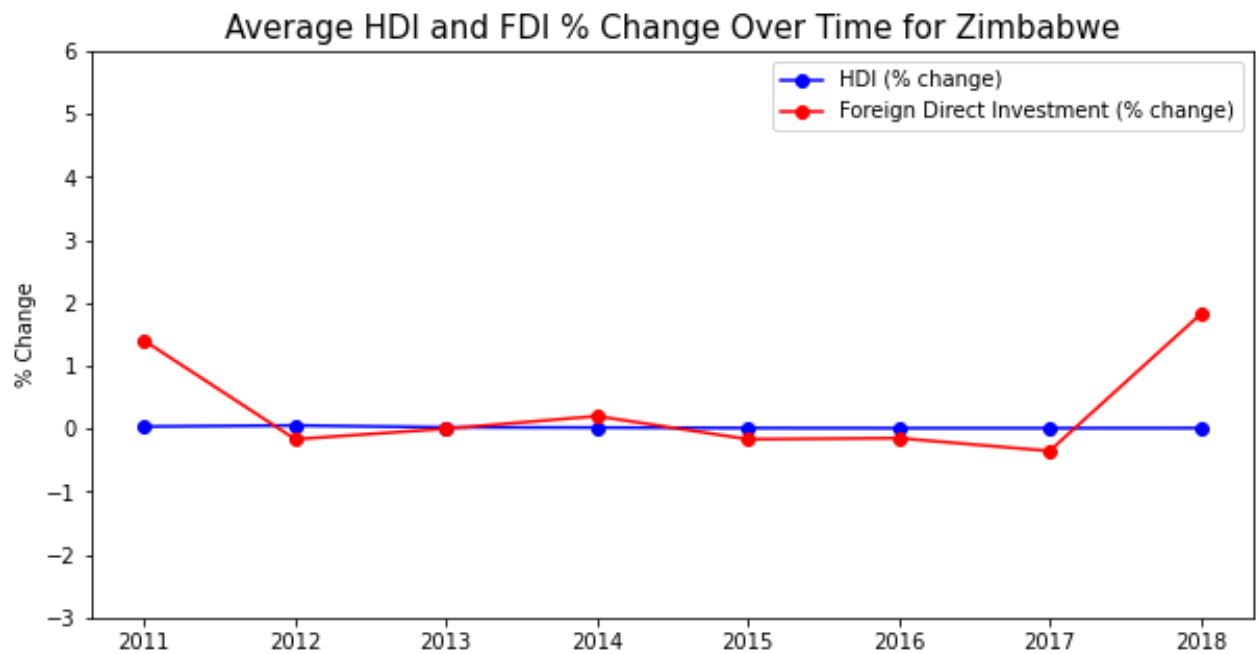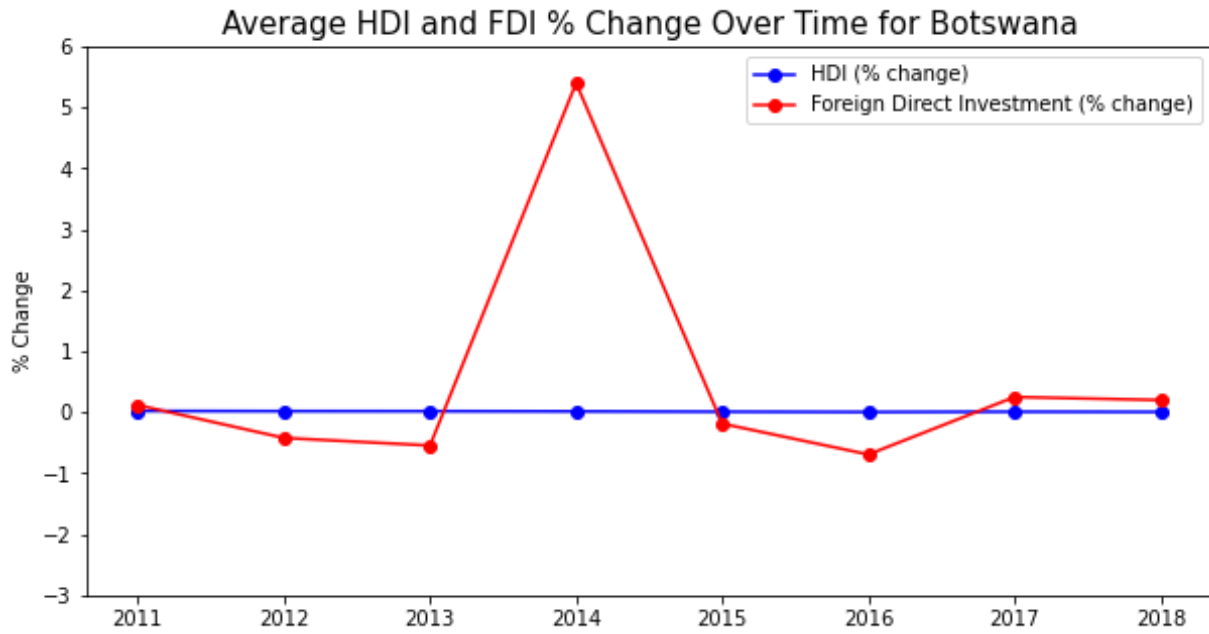Zimbabwe: GNI vs. HDI = 0.8551928453940894
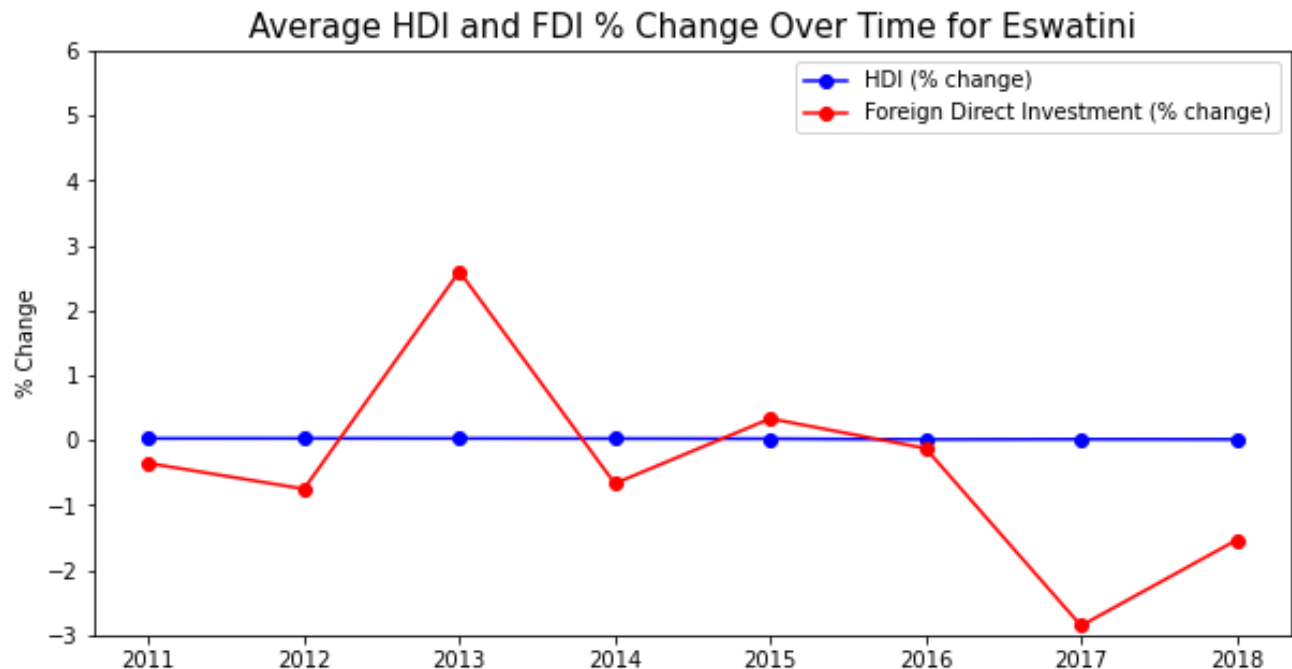

Botswana: EI vs. HDI = 0.0745614921205372

Botswana: LE vs. HDI = 0.9104070540236903

Botswana: GNI vs. HDI = 0.8329148439674214

Overall, Life Expectancy is the strongest correlation for each of the top countries, indicating that the most successful HDI growth is proxied by health improvements in the country. Aside from LE, EI appears to be the worst proxy, contradicting the findings for the entirety of Sub Saharan Africa. These findings confirm both of my hypotheses for this question. The top performers are similar in their correlation rankings and different from the rest of Sub Saharan Africa. In periods of high HDI growth, health improvements in the country are the most accurate indicator of HDI growth. An indirect conclusion could be drawn that overall, education is the steadiest improver of HDI, while health improvements are the quickest improver.

Question 3:

Average HDI and FDI % Change Over Time for Botswana

Average HDI and FDI % Change Over Time for Zimbabwe

Average HDI and FDI % Change Over Time for Eswatini

For question 3, I wanted to observe if external factors appeared to have a significant relationship with HDI, possibly even a better correlation than its own components. I was particularly interested in foregin direct investment (FDI), as it is a complicated factor in fragile, emerging economies. In theory, FDI would only go towards benefiting HDI, as it would hit each of its components through correlative investments (i.e. hospitals, infrastructure, schools, etc.). I would expected foreign investment, if done correctly, to at least have a similar correlation with HDI compared to LE, EI and GNI. I wanted to continue to observe the top 3 countries because you would expect foreign investment to play a role. As the visualization shows, FDI is highly volatile and doesn't appear to influence HDI at all. The graph's scale is a bit challening, but even then, % change in HDI hardly moves. Also, each country has vastly different FDI graphs, showing that it is not a reliable contributor.

Here are the additional correlations:

Eswatini: FDI vs. HDI = 0.5093578147192478

Zimbabwe: FDI vs. HDI = 0.03471126514706657

Botswana: FDI vs. HDI = 0.12733385713504144

My hypothesis was proven right in that the HDI components were still more accurate; however, I was very wrong in assuming FDI would have strong correlations with successful countries. Being a complicated factor, FDI's correlations vary in country depending on the investor and the sum of the investment. It is clearly a poor indicator of HDI success.