**Life Expectancy: Exploring the Influences of Diverse Factors**

ISM-6419 – DATA VISUALIZATION FOR STORYTELLING

FINAL PROJECT

**SUBMITTED BY**-

PAWAN JETHWANI

U27212618

[*pawanjethwani@usf.edu*](mailto:pawanjethwani@usf.edu)

**Introduction**:

The Life Expectancy of humans throughout the world is an important aspect in understanding the overall well-being not only of the individual but also of the entire nation the individual belongs to. Life Expectancy can vary in different parts of the globe due to various factors which include and are not limited to health diseases, infant deaths, Population, drug overdose or even GDP of the nation.

In our comprehensive exploration of life expectancy, we embark on a journey to delve into some complex interactions among socio-economic factors, healthcare variables and broader demographic indicators to understand their impacts on life expectancy, if any.

In this project, we will be analyzing various factors that may affect Life Expectancy around the world. We will be focusing on multiple aspects that include Death rates due to substance use disorders, some disease parameters such as HIV/AIDS, Hepatitis A, Measles, Diphtheria and Polio. We will also consider impacts of some Socio-Economic factors like Schooling, Infant Mortality, Population etc. We will also analyze some trends to see how Life Expectancy varies in different countries over time.

This research is important because it clarifies why people live different lengths of time. Through examining variables such as GDP, health, and other pertinent information like the effects of drug usage, we can determine the elements that influence life expectancy. Government agencies, medical professionals, and others working to enhance public health may find this information useful. We may choose more wisely where to concentrate our efforts and resources to support individuals in leading longer and healthier lives by understanding the variables that are linked to life expectancy.

This analysis will be based on a comprehensive dataset containing real time data derived from multiple sources to maintain the authenticity and deliver a better understanding of the impacts due to the factors mentioned above.

We will be trying to answer the below research questions which are formulated as:

**Research Questions**:

1. How are Life Expectancy and Death Rate due to substance use disorder related?

2. What is the relation between various diseases and Life Expectancy?

3. How does Schooling, Infant Death and Population impact Life Expectancy?

4. Does GDP have a significant impact on Life Expectancy?

**Methodology:**

Gathering relevant Data:

For this project I collected data from multiple sources which I found relevant to our exploration. Below are the websites the data was sourced from:

<https://ourworldindata.org/grapher/life-expectancy-un-vs-gdp-per-capita-wb>

<https://www.who.int/data/gho/data/indicators/indicator-details/GHO/life-expectancy-at-birth-(years)>

<https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who?select=Life+Expectancy+Data.csv>

<https://ourworldindata.org/grapher/death-rates-substance-disorders>

Data Compilation and Cleaning:

Once the data was identified, we cleaned it by removing some irrelevant information that came along and also some missing information. We also did a thorough check to ensure that it contains all the needed information for our analysis.

Data Merging:

Next step involved merging the selected datasets into a comprehensive repository using the Data Visualization tool Tableau which we also used for our exploration.

Data Analysis:

Here we performed detailed analysis of the comprehensive dataset to understand the variable and also to identify trends and patterns in the collected data.

Data Visualization:

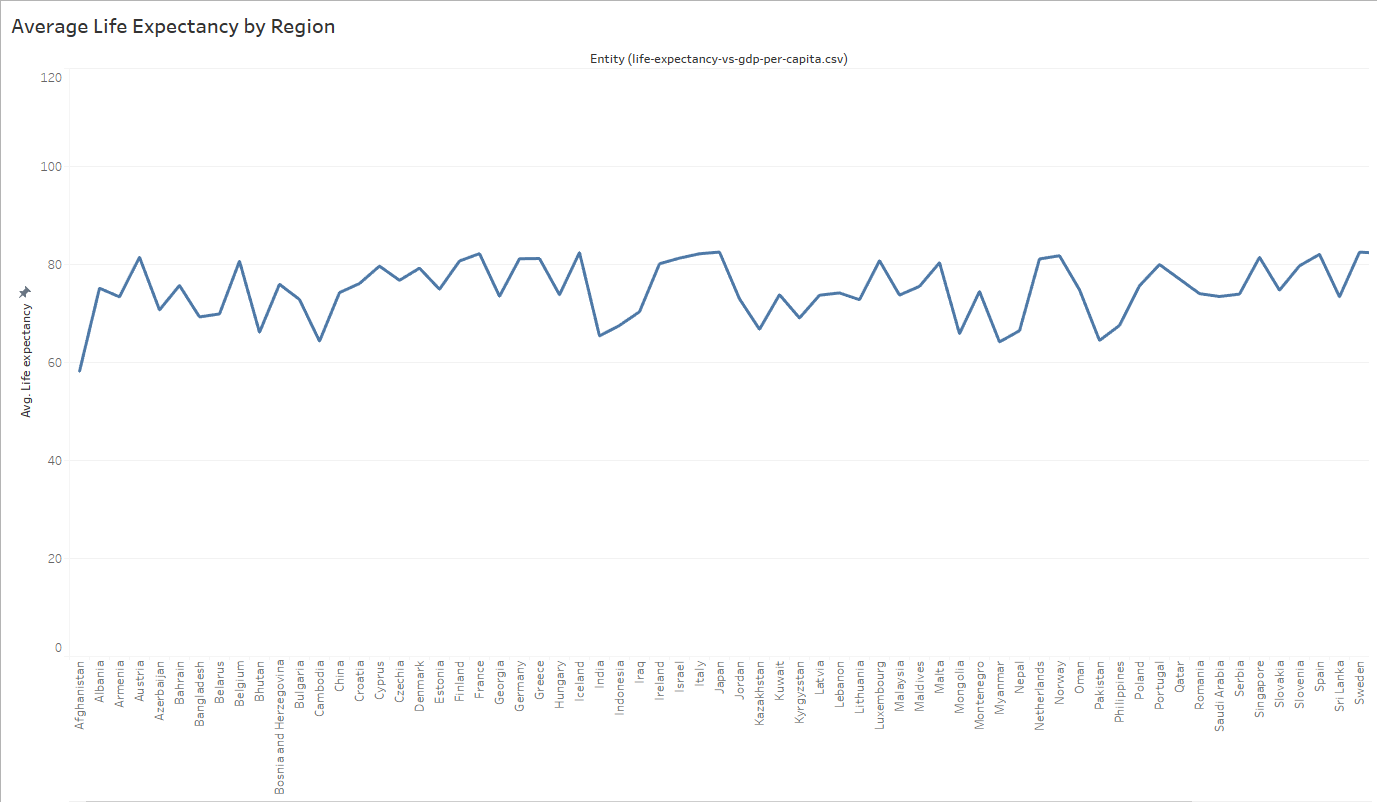
Visualize the finalized data using the data analysis tool Tableau. We chose Tableau for its ease of use and its ability to create interactive and colorful visualizations by providing options to create interactive charts and graphs that enable us to communicate our analysis based on the research questions effectively.

Conclusion:

Again, we used Tableau to come up with interactive dashboards to summarize the findings of our analysis of the various factors impacting Life Expectancy.

**Analysis:**

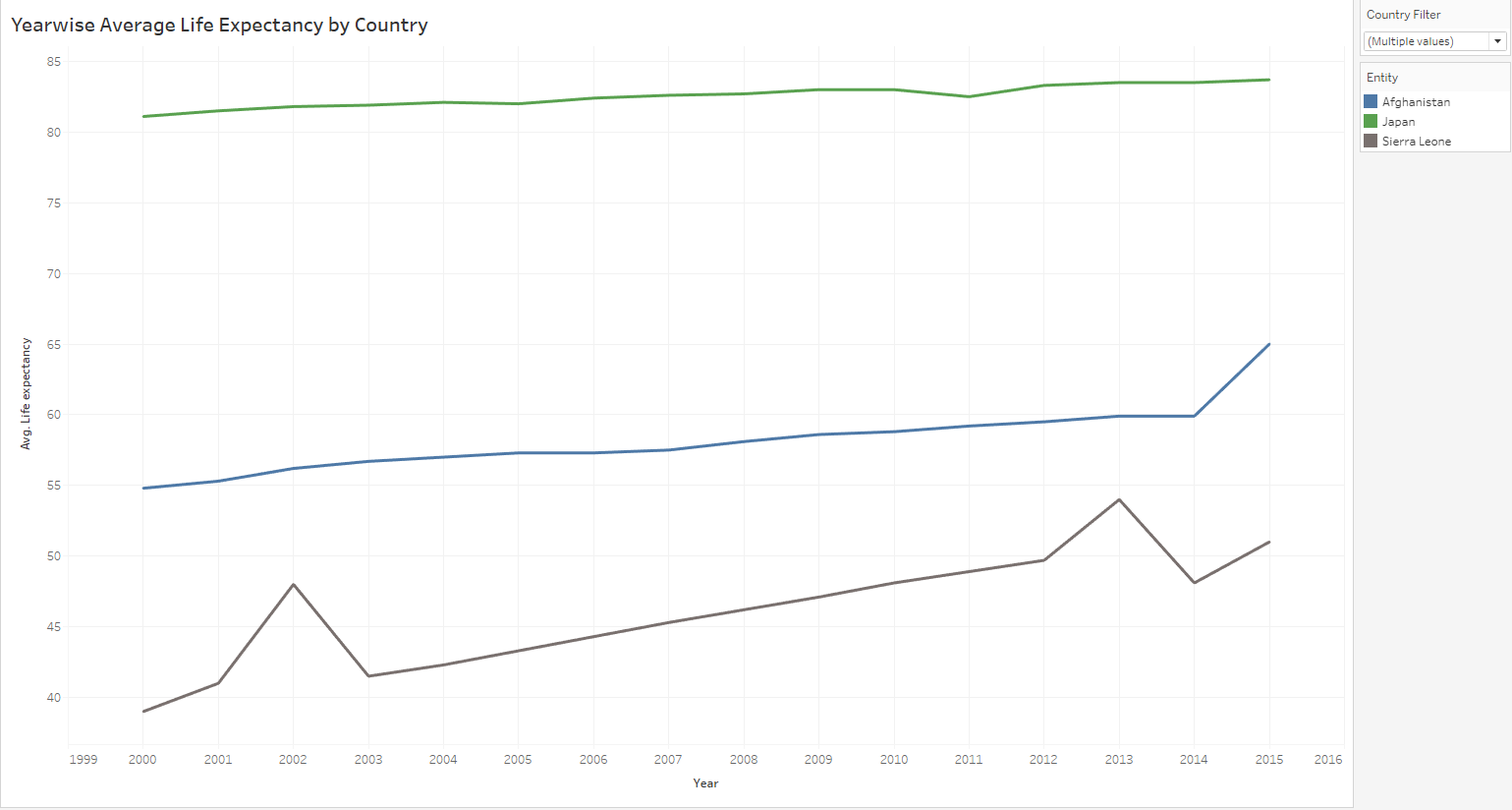
As mentioned earlier, we sourced our data into the visualization tool – Tableau. And here is a preliminary view of Average Life Expectancy in Years by region. Below representation shows a selection of countries in Asia and Europe. Our visual provides a comprehensive coverage of all the regions but for a clear understanding here we selected Asia and Europe at random.



As we can see from the above visual, the life expectancy in the countries belonging to Asia and Europe range from a little under 60 years to a little over 80 years on average. Among the two regions Japan has the highest average of 82.59 years and Afghanistan has the lowest average of 58.19 years.

On digging deeper, I found that among all the continents the lowest Average Life Expectancy of 46.81 years belongs to the country Sierra Leone and Japan still tops the world in the average.

Let’s now take a look at the Year wise Average of Life Expectancy for Japan, Afghanistan and Sierra Leone.



The above visual explains our preliminary analysis of the highest and lowest averages for Japan and Sierra Leone as seen earlier.

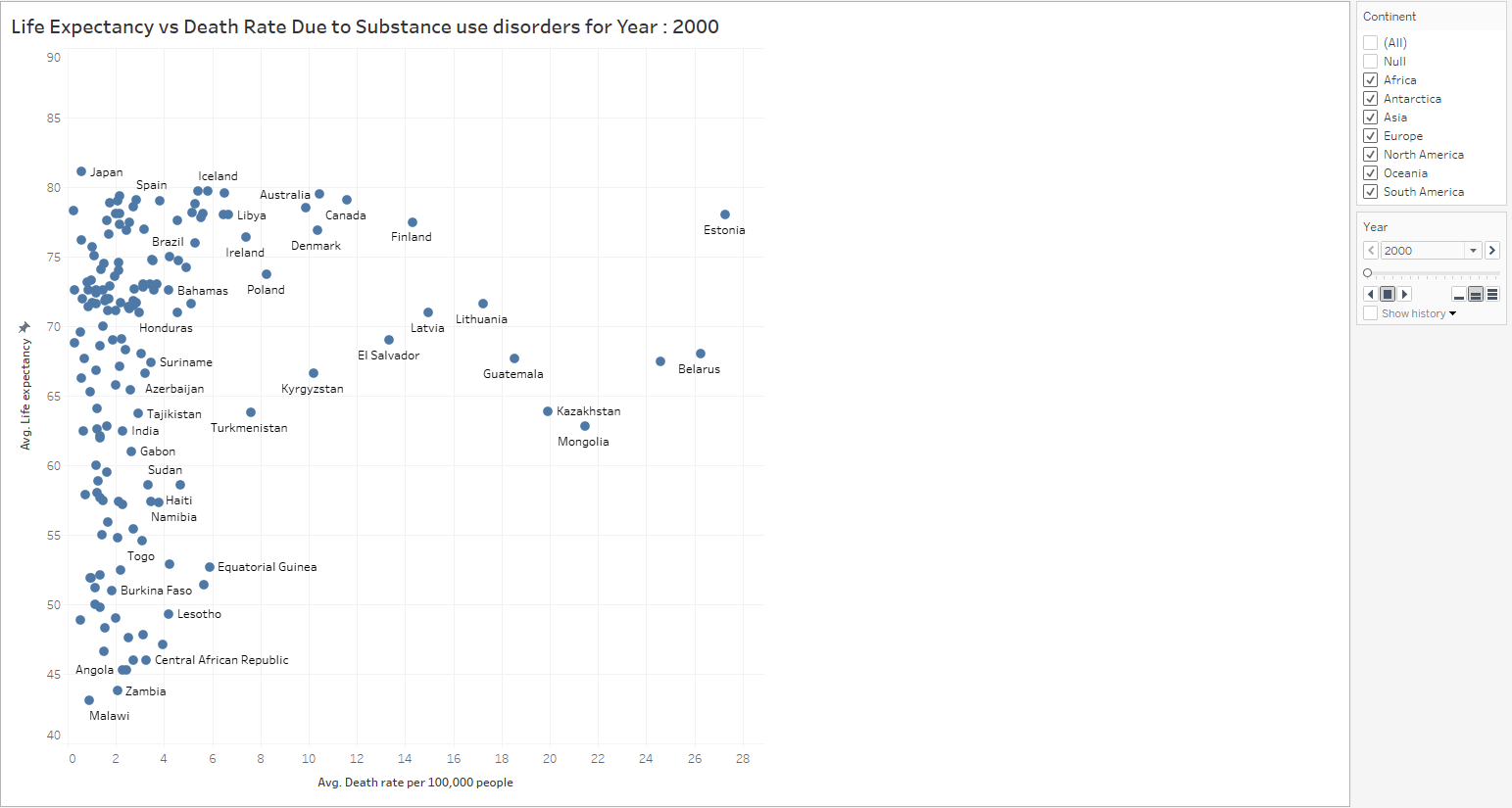
The Averages we see above span between years 2000 to 2015 and we see that Japan ranges from 81.10 in 2010 to 83.70 in 2015.

On the other hand, the Life Expectancy average in Sierra Leone ranges from 39.00 in the year 2000 to 51.00 in 2015.

The above visual also presents the country Afghanistan and explains the yearly averages in a similar fashion. Likewise, we have provided the ability to select the country of choice to analyze similar trends.

Let’s now delve into our research questions to understand the factors that may have caused the variations in the Life Expectancy as we saw in our preliminary analysis.

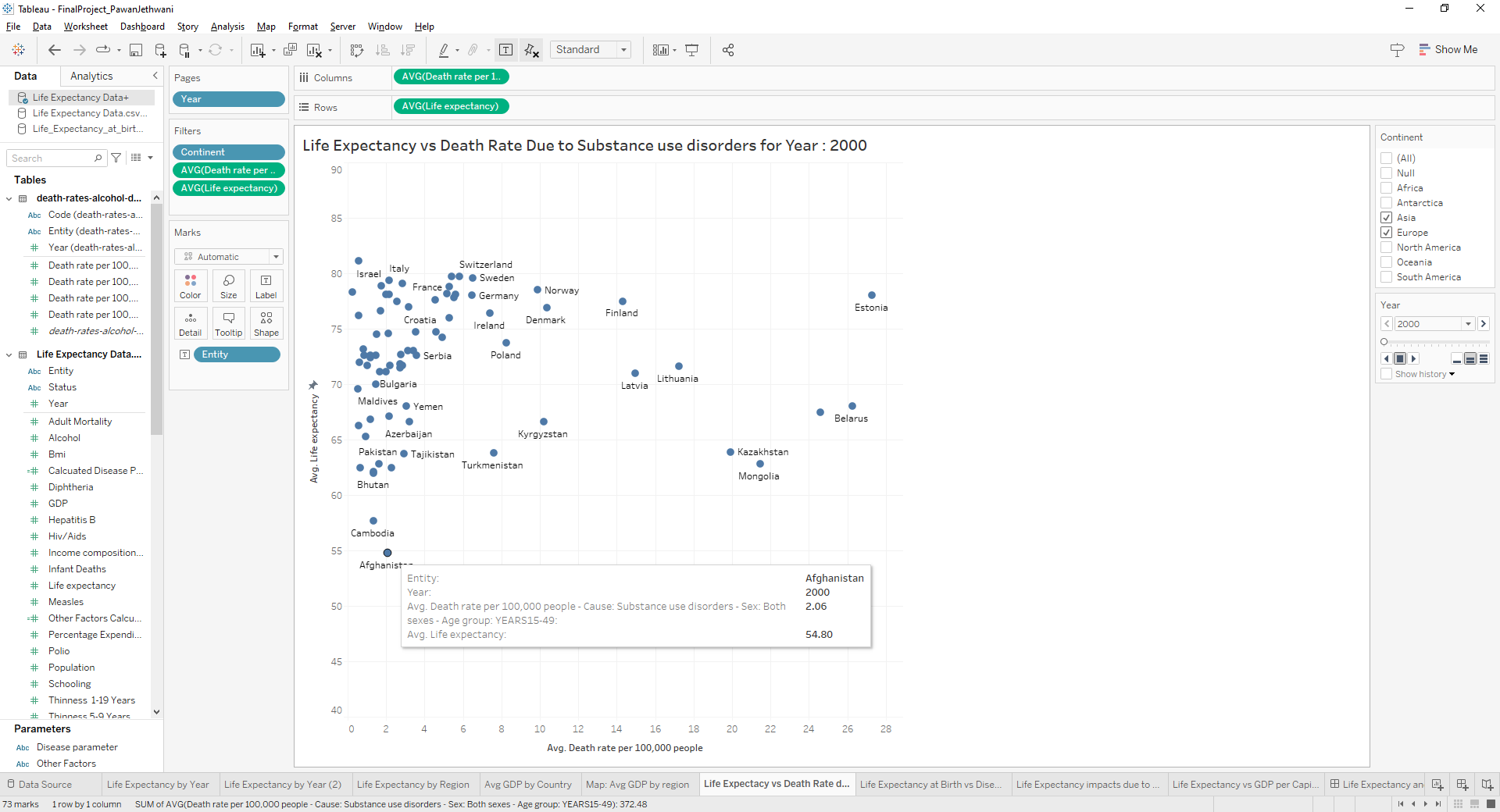
**1). How are Life Expectancy and Death Rate due to substance use disorder related?**



The above visual is a year wise representation of Life Expectancy for all the regions and how it is impacted by the Death rate due to substance abuse disorder.

Here we also analyzed trends to understand how Life Expectancy varies by year as compared to average deaths due to substance use disorders. We have used a page animation to represent this analysis and below are a couple of references for the country Afghanistan for the years 2000 and 2015.

Let’s look at the below representations to understand the same:

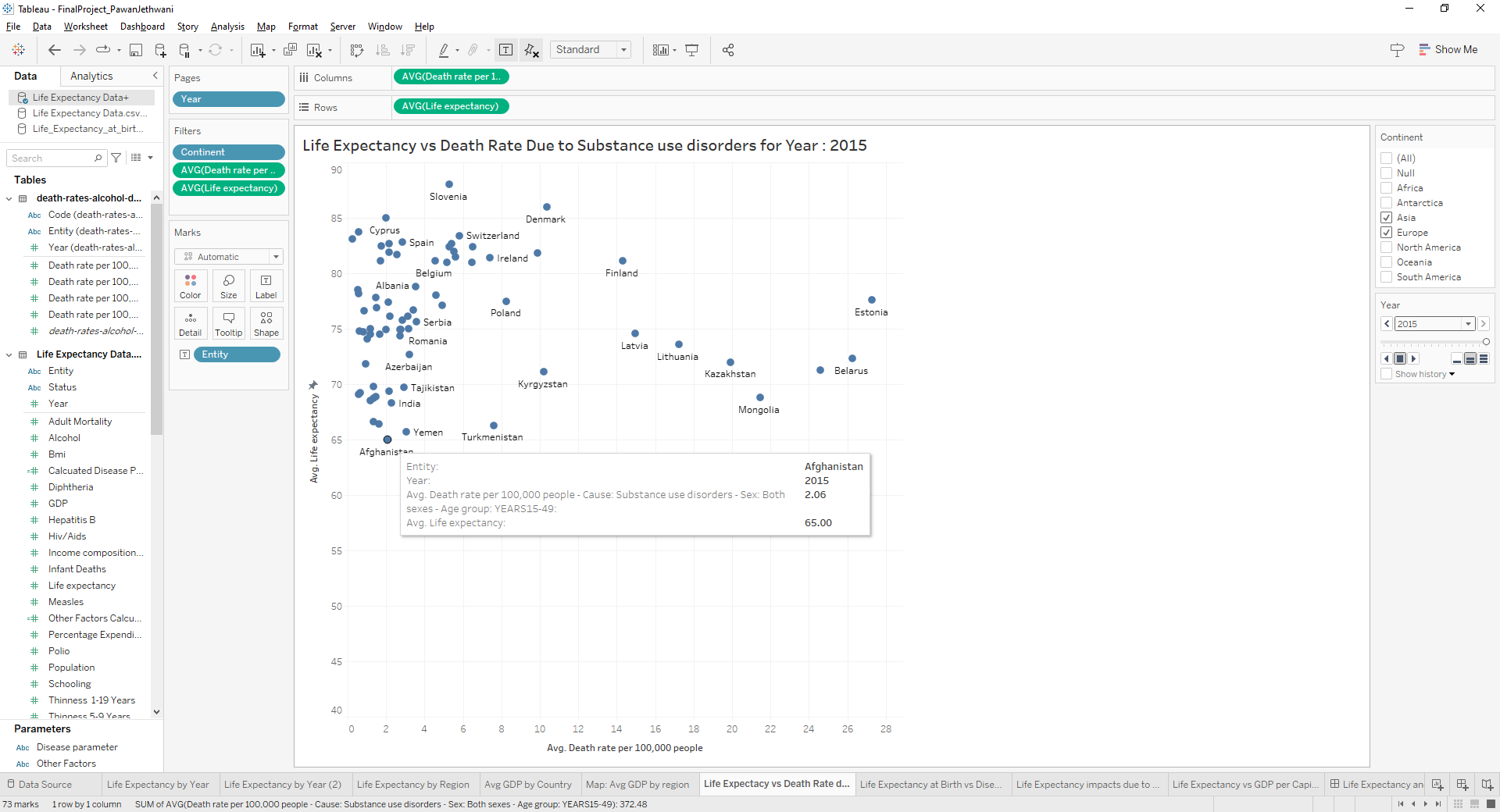


As we see above, for the year 2000 in Afghanistan, we saw an average of 2.06 deaths per 100,000 and the average life expectancy then was 54.80 years.

Similar verifications were done for different regions as well. Let’s look at a few examples:

For Canada (representation not shown) which falls under the region of North America, we see an average of 11.58 deaths per 100,000 and their average life expectancy that year was 79.10.

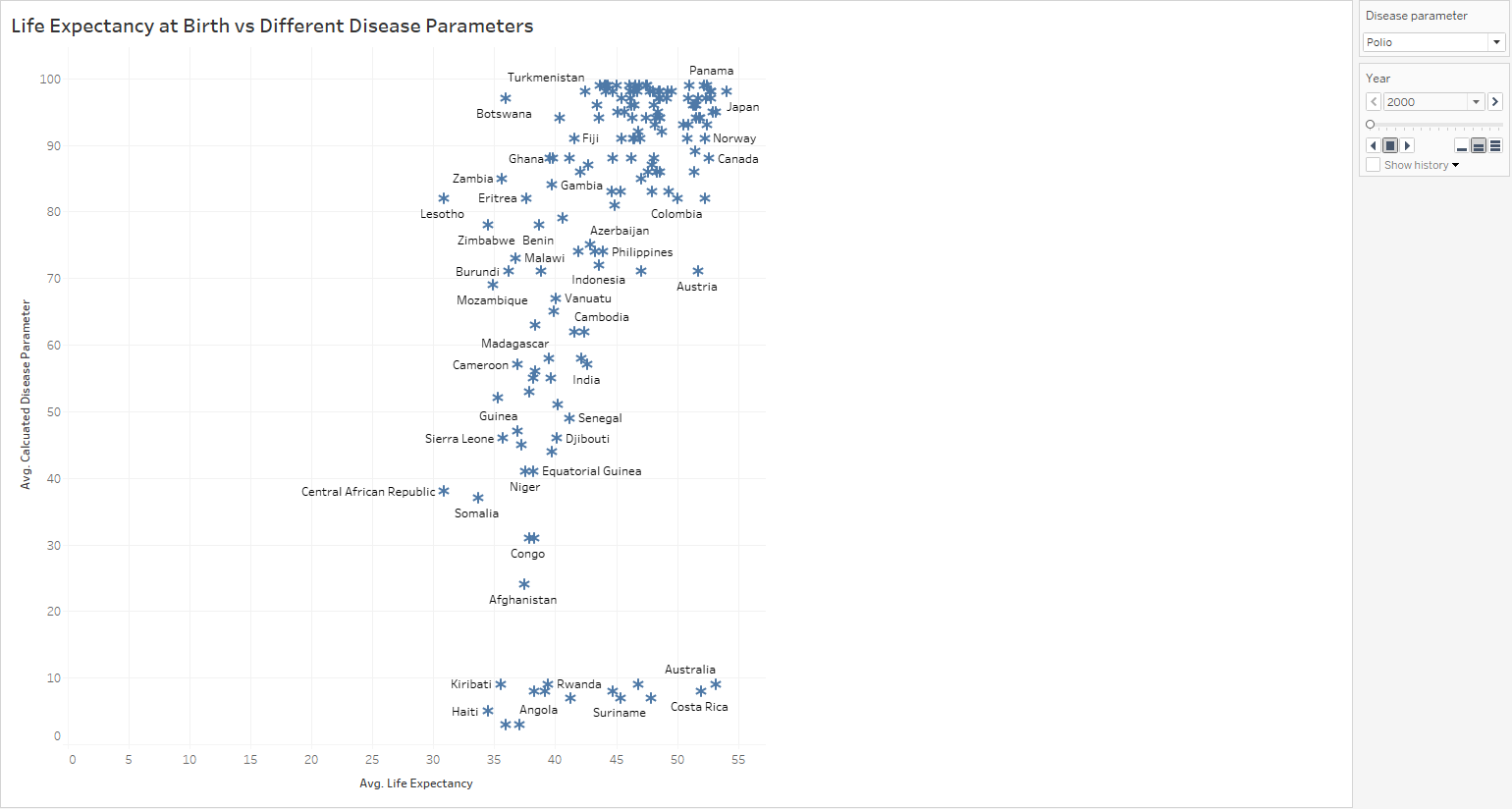
For Brazil (representation not shown) which falls under South America, we see an average of 4.21 deaths per 100,000 and their average life expectancy that year was 75 years.

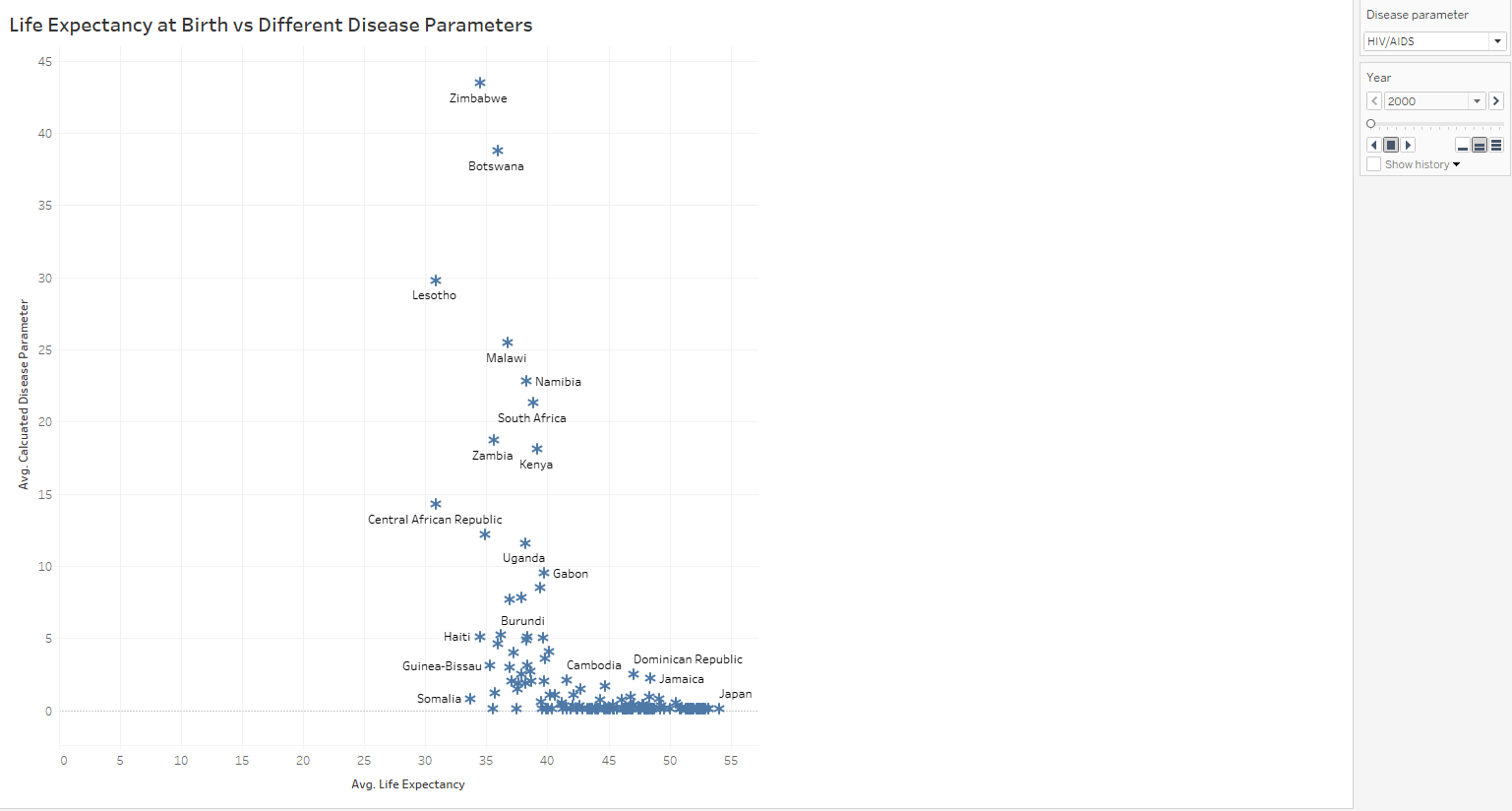


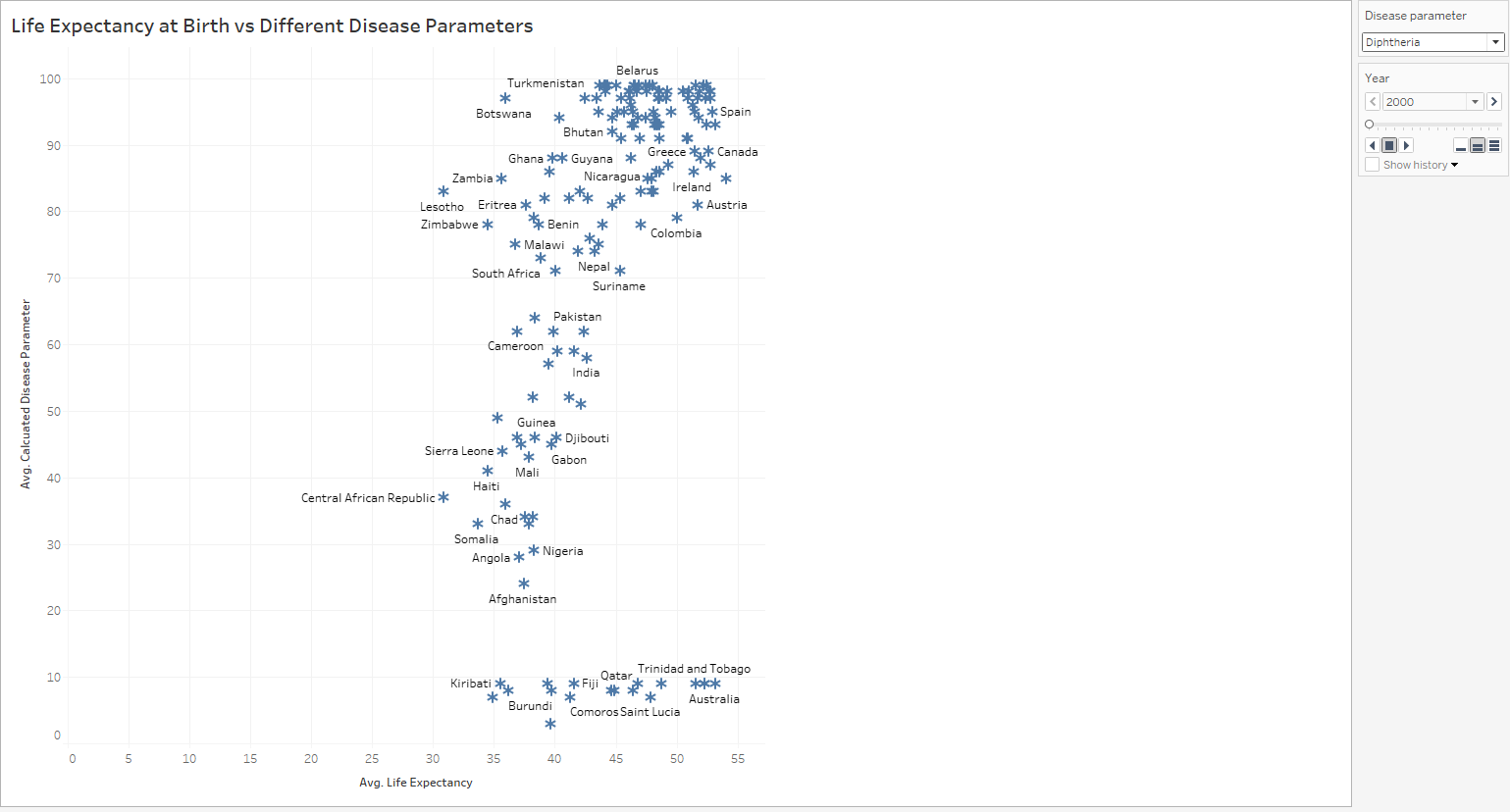
As per the above representation for the year 2015, we again see the same death rate per 100,000, however the average life expectancy increased to 65 years in Afghanistan.

We further analyzed the same for multiple countries and concluded that the average death rate over the years remained almost the same and hence this analysis did not provide any conclusive evidence of dependency on the death rate due to substance use which led us to move on to our next question towards our goal to identify other potential factors impacting Life Expectancy.

**2). What is the relation between various diseases and Life Expectancy?**





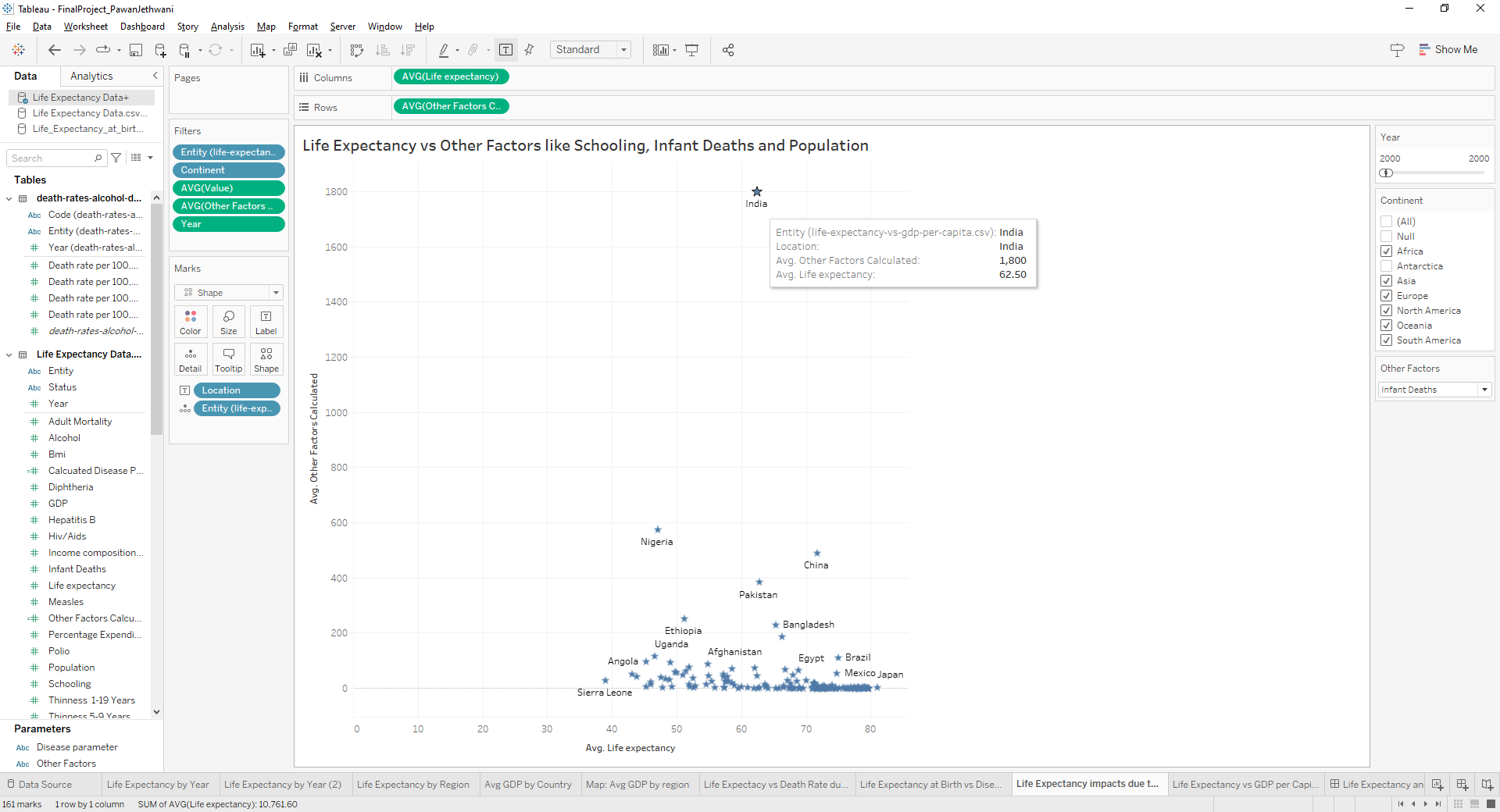


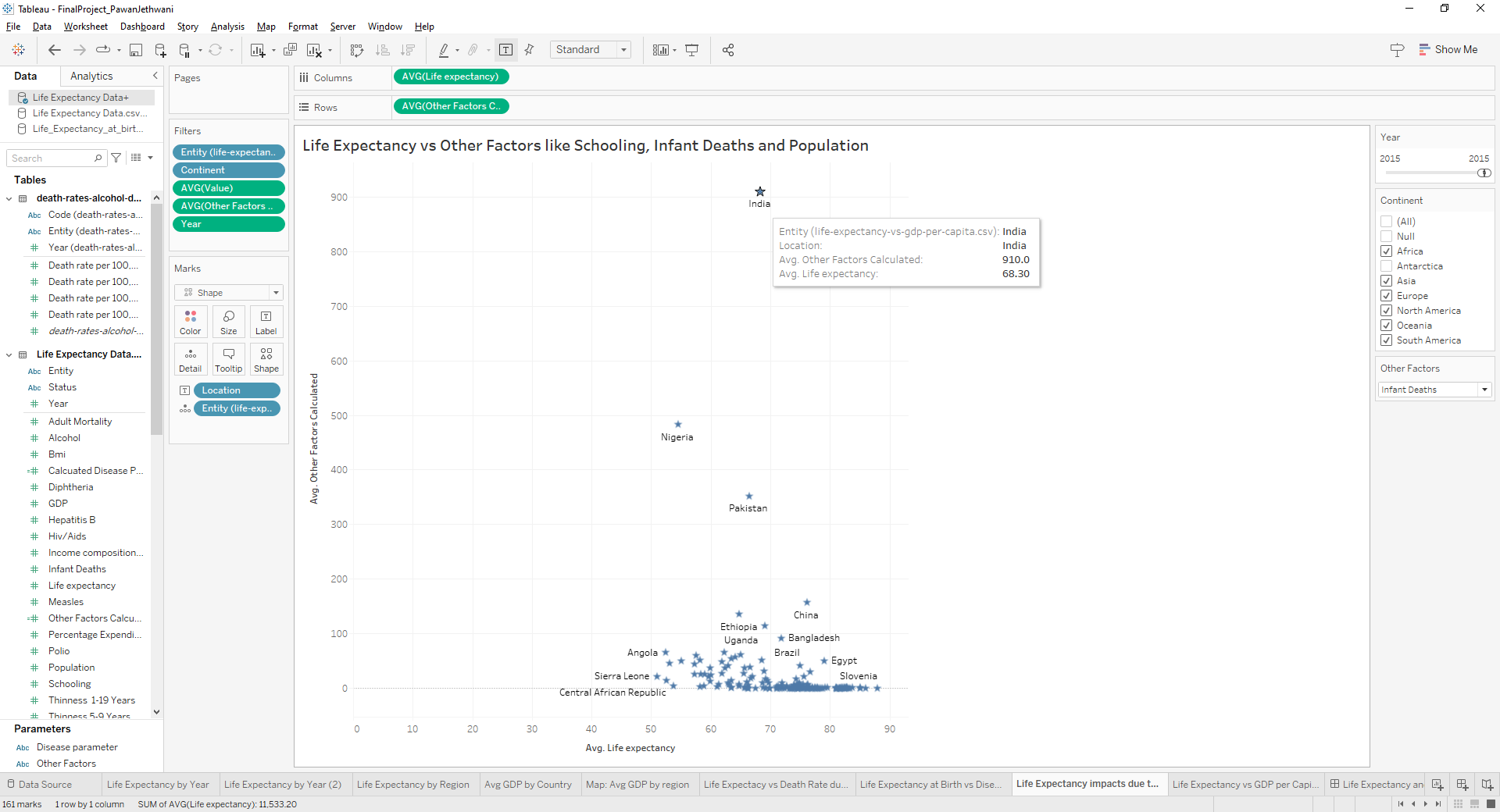
The above three representation reflects the impact of diseases Polio, HIV/AIDS and Diphtheria respectively on Life expectancy for the countries over the years.

For our analysis, I checked the impact of these diseases for the country India and found that over the years 2000 to 2015, India on an average had about 57 deaths due to Polio, 3 deaths due to HIV/AIDS and 58 deaths due to Diphtheria. However, the life expectancy increased from 62.5 years to 68.30 which was a significant improvement. During these years there were significant improvements in the healthcare and India and a vaccination standard was followed for children since birth.

We extended our analysis to examine the potential effects of certain socio-economic factors, as highlighted in the following research question:

**3). How does Schooling, Infant Death and Population impact Life Expectancy?**





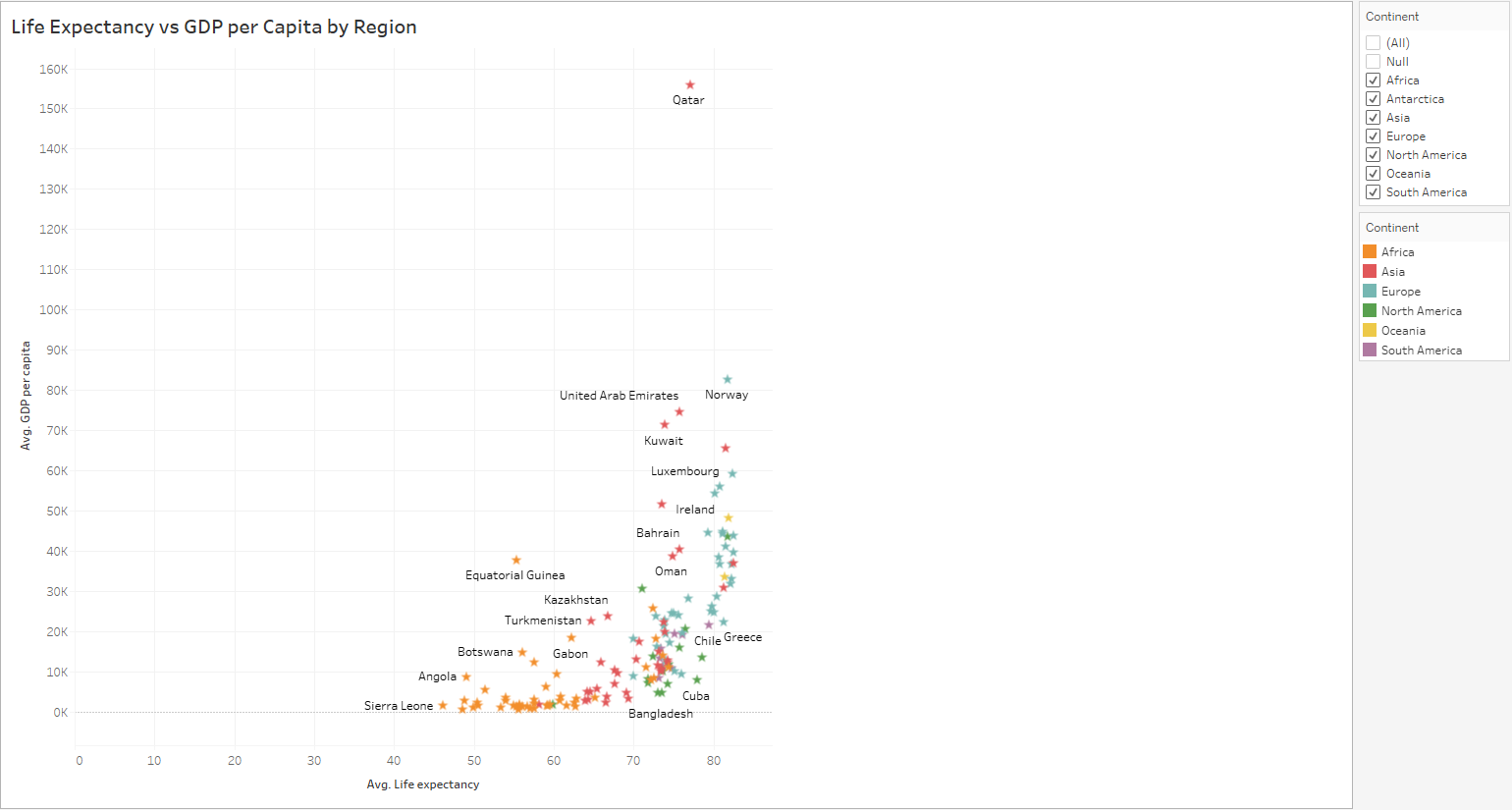
Above representation is a consideration of an important factor Infant Deaths, and shows how reduction in the number of infant deaths which can be another reason of improved healthcare, led to significant improvement in the average life expectancy.

Here, for reporting purposes, we considered an example of a developing nation India which had an average of 1800 infant mortality in the year 2000 and had an average life expectancy of 62.50 years which significantly increased to 68.30 in the year 2015 which as we can see could be a direct impact of the number of infant mortalities which went down to 910. This was about 50% reduction in infant mortality rate.

In the Tableau workbook we have provided an option to select the various other parameters for analysis as well as a selection for region. However, on verifying, out of all the factors under consideration for multiple regions, Infant Deaths was the one which showed significant reduction probably leading to increased life expectancy.

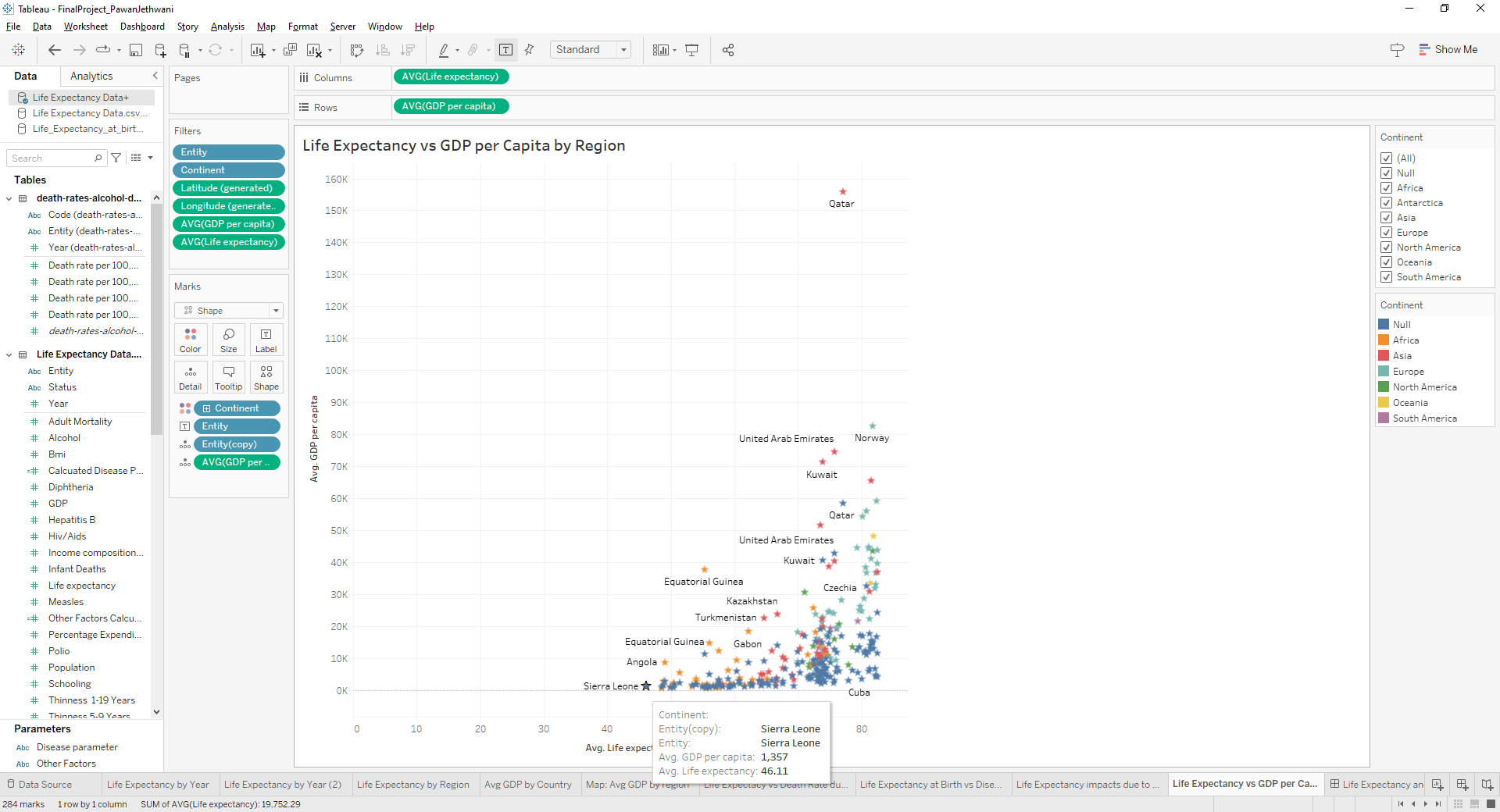
We further extended our research to verify the below:

**4). Does GDP have a significant impact on Life Expectancy?**



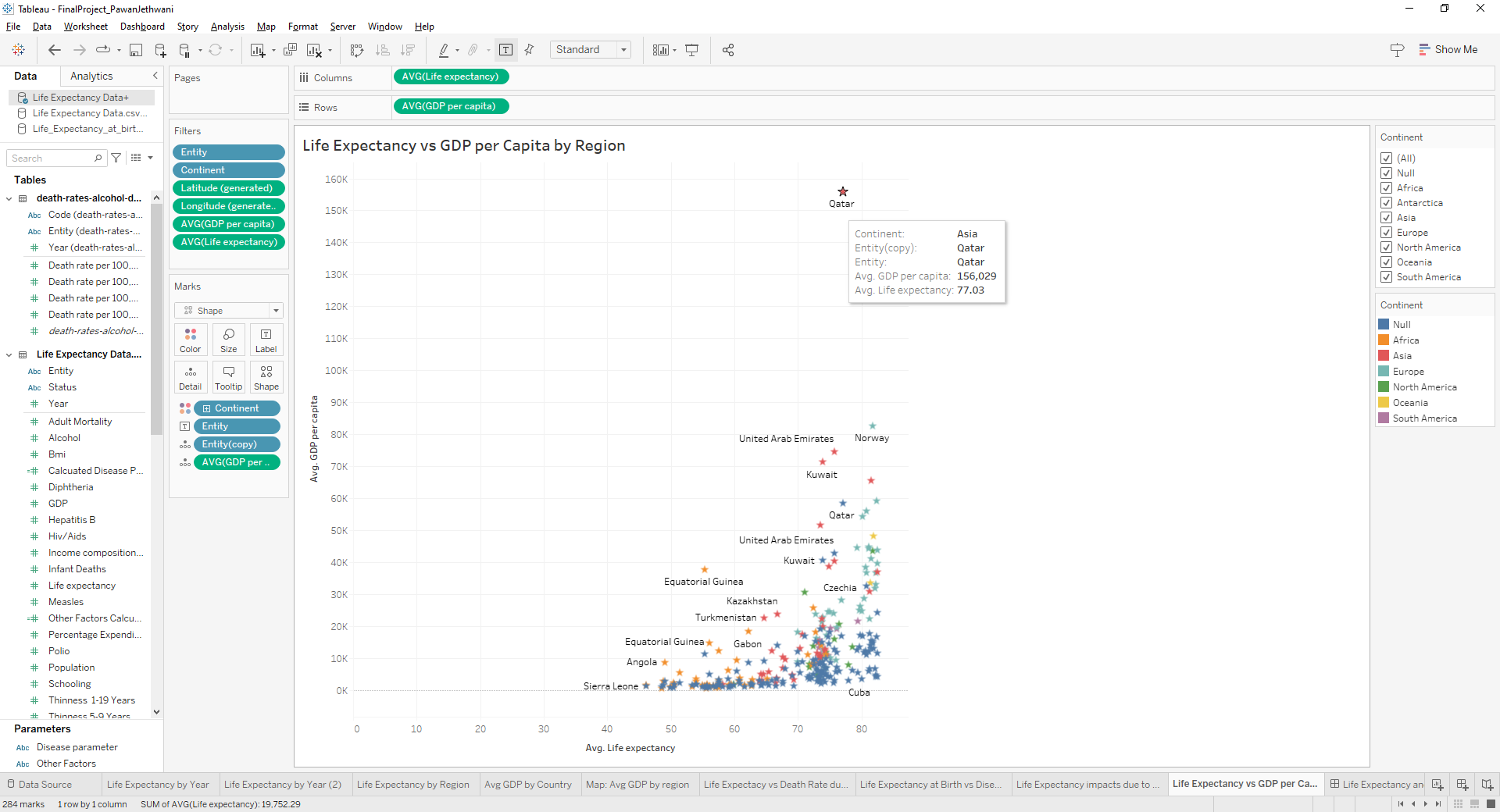
The above representation reflects a comparison of GDP per capita to that of average Life Expectancy by Region.

To analyze the impact of GDP on Life Expectancy we considered the examples of Sierra Leone and Qatar that are at the bottom and the top of the visual as seen respectively.



The above selection highlights the relation between Avg GDP and Avg Life Expectancy for the country Sierra Leone. Here we see that this country has an average GDP per capita of 1357 and a life expectancy average of 46.11.

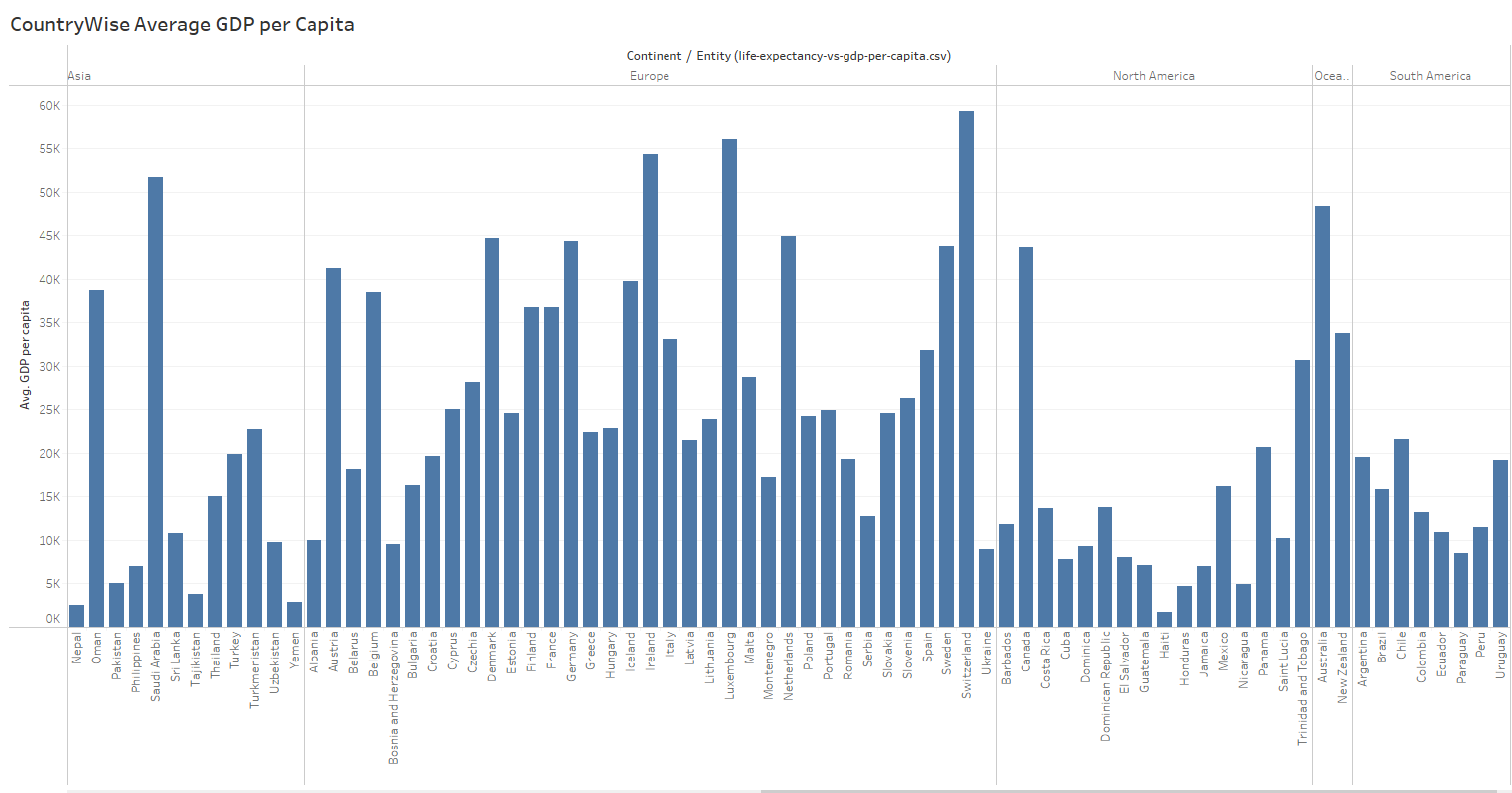
Let’s now take a look at Qatar which stands at the top of the chart as seen.

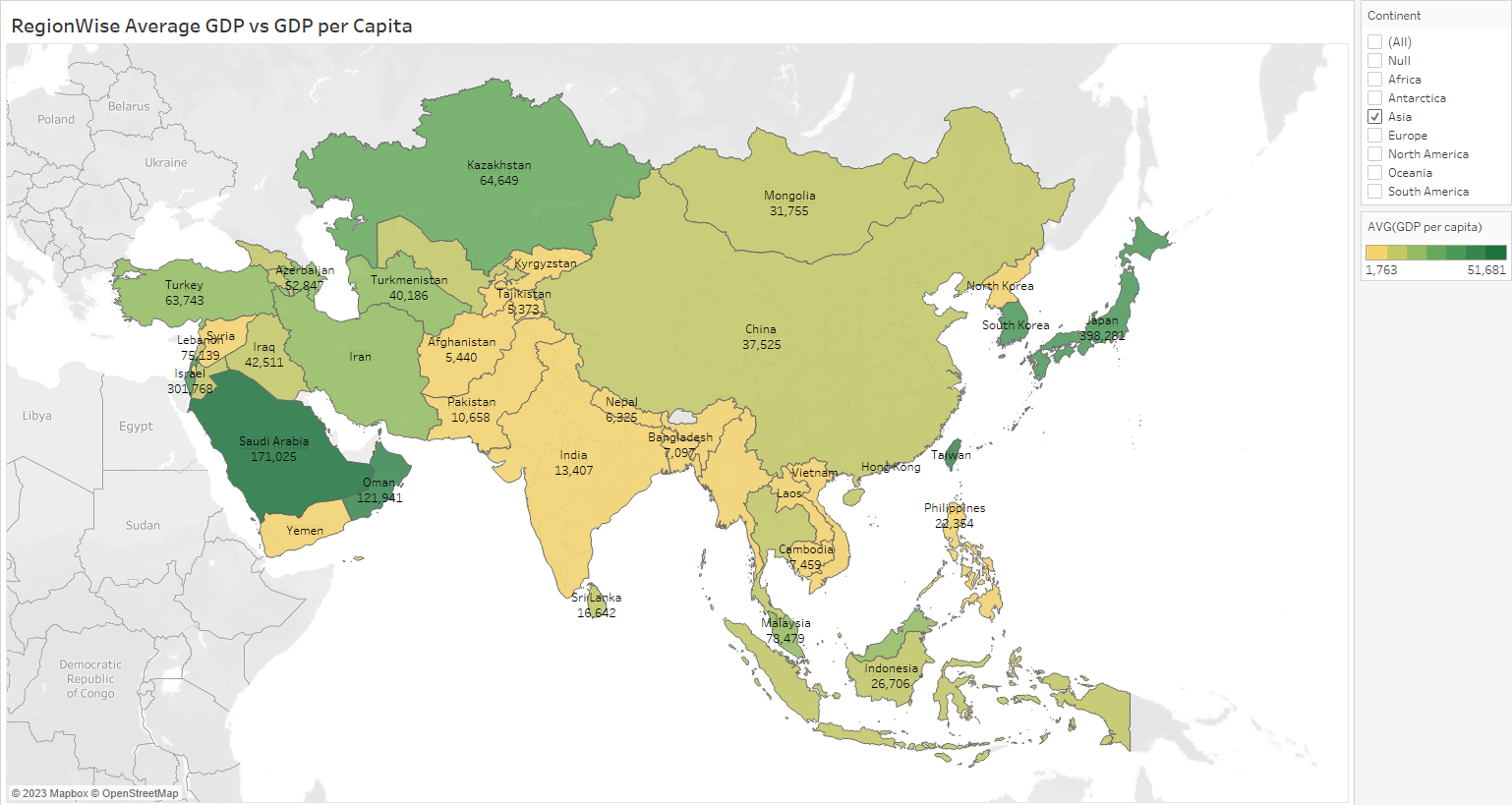


Qatar as seen has an Average GDP of 156,029 per capita and we see a significantly higher Average Life Expectancy of 77.03 as compared to that of Sierra Leone.

This comparison helps us understand that GDP in addition to some other factors as seen earlier, does play a significant role in increasing the Life expectancy of a nation.

In addition to the above research questions, I also tried to gain further insights through the GDPs of nations. Below are a few representations that were taken into consideration.





The above two representations show country wise and Region wise GDP comparisons which were also compared to Average Life Expectancies in the analysis which we saw earlier.

**Conclusion:**

The results of my analysis, in summary, highlight the important influence of several factors on life expectancy. Interestingly, a greater GDP per capita is found to be a significant factor in longer life expectancies, demonstrating the complex relationship between economic growth and general well-being.

Our results also demonstrate the vital impact that healthcare interventions play, especially when it comes to lowering newborn mortality rates and combating common illnesses like polio, HIV/AIDS and diphtheria to name a few. There is a real chance that average life expectancy may rise as countries work to improve socioeconomic conditions and healthcare facilities. These findings highlight the significance of comprehensive strategies that include targeted disease prevention, healthcare infrastructure, and economic prosperity in promoting longer and healthier lives for people everywhere.

In addition to the analysis above, I am keen towards continuing my research and identifying more potential factors that may help various organizations towards taking steps for better and improved life expectancies in the coming years. To move towards this approach, I have formulated few more research questions that I plan to investigate on. Here are the questions:

**Further Research Questions:**

Can we identify regions where life expectancy has shown notable improvement or decline?

How did Covid impact life expectancy in different regions?

How does Environmental sustainability impact on life expectancy?

Can advancement in education Impact life expectancy?

Can lifestyle changes such as Healthy eating, Exercise improves life expectancy?