**Term Project Final**

**WHO statistics on life expectancy**



Submitted by:

**Steffi George - 300358150**

**Hazel Ibasco - 300366644**

**CSIS 3860 SEC 071**

Professor : **Rupa Manabala**

**Part A - PowerBI**

**Dataset Chosen**

**WHO statistics on life expectancy:**This dataset contains information collected by the World Health Organization and the United Nations to track factors that affect life expectancy. The data contains 2938 rows and 22 columns. The columns include: country, year, developing status, adult mortality, life expectancy, infant deaths, alcohol consumption per capita, country’s expenditure on health, immunization coverage, BMI, deaths under 5-years-old, deaths due to HIV/AIDS, GDP, population, body condition, income information and education.

Then we came across another dataset **Life Expectancy (WHO) Fixed** which had fixed the issues of some missing values and had the column regions.

Each column represents a different variable or feature that could potentially influence life expectancy. Here's a brief description of each column:

* Country: The name of the country or region.
* year: The year in which the data was collected.
* Status: Indicates whether the country is classified as "developing" or "developed."
* Adult mortality: A measure of mortality among adults in the country.
* Life expectancy: The average number of years a person is expected to live.
* infant deaths: The number of infant deaths in the country.
* Alcohol : The average alcohol consumption per person in the country.
* Percentage expenditure: The country's healthcare expenditure.
* BMI: Average Body Mass Index in the country.
* deaths under 5-years old: The number of deaths in children under 5 years old.
* deaths due to HIV/AIDS: The number of deaths attributed to HIV/AIDS.
* GDP: Gross Domestic Product, a measure of the country's economic output.
* population: The total population of the country.
* Schooling: Data related to education levels in the country.
* Thinness 1-19 years and Thinness 5-9 years
* Income composition of resources:provides an insights into a country's economic structure

**Analysis Questions**

* Which country has the highest life expectancy in the year 2015?
* Which country has the highest under 5 years old death rate ?
* Which region has the highest infant mortality rate ?
* Which are the top 5 developed countries with least under 5 deaths due to HIV?
* Which country has the highest expenditure on healthcare?

**Transformation strategies**

First, we connected the dataset to our Power BI workbook and conducted some transformations such as changing data types, changing column names, adding conditional columns to modify null values, and replacing null values with fill up and fill down.Then we added the second dataset, where we updated the names of the column and the types, and then we merged both tables to add the region from table 2 to table 1.

The merged query was then renamed to “updated life expectancy”. Then we changed the type of region to text,reordered the columns, sorted the country and year.

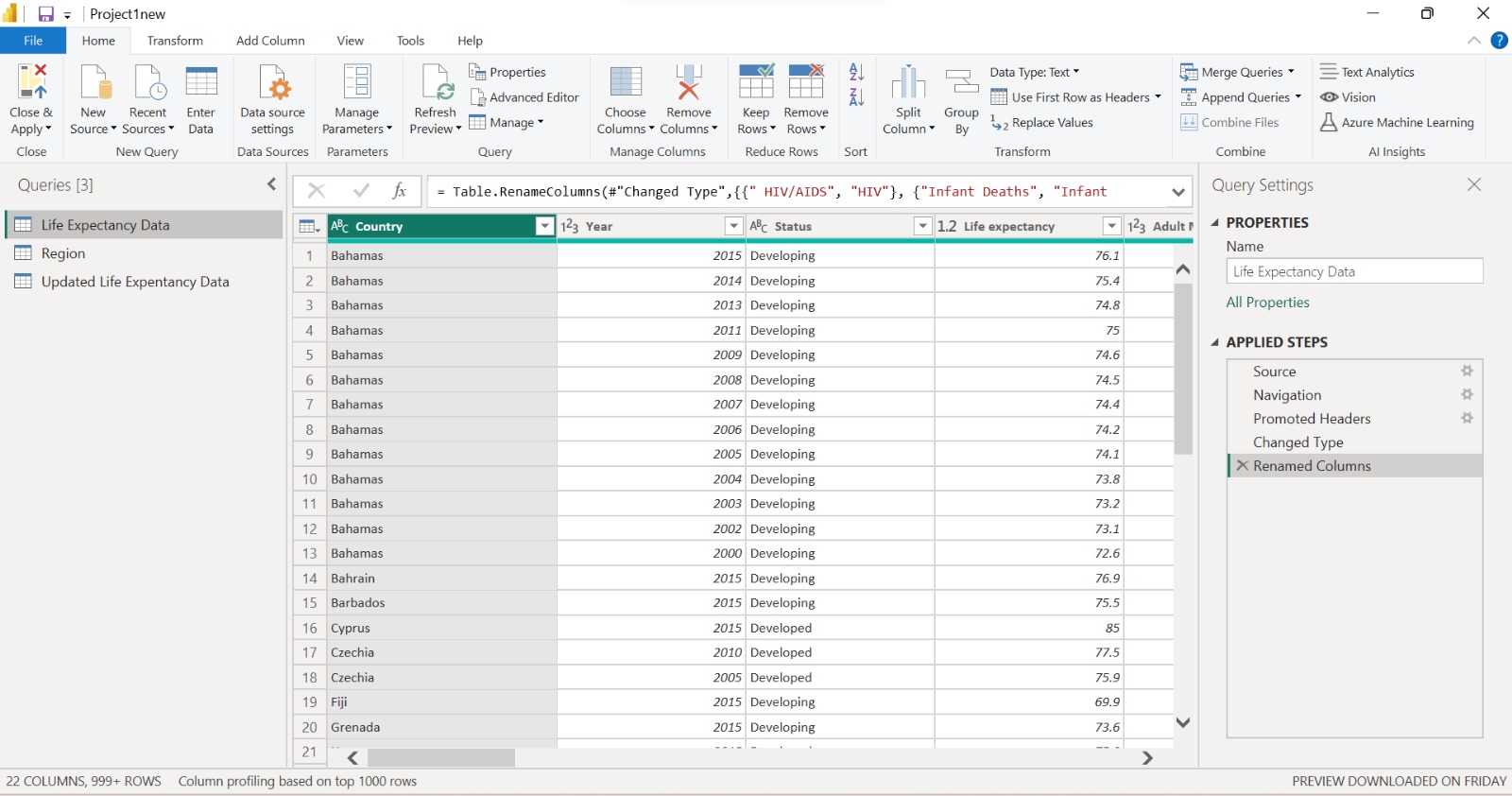


Table 1

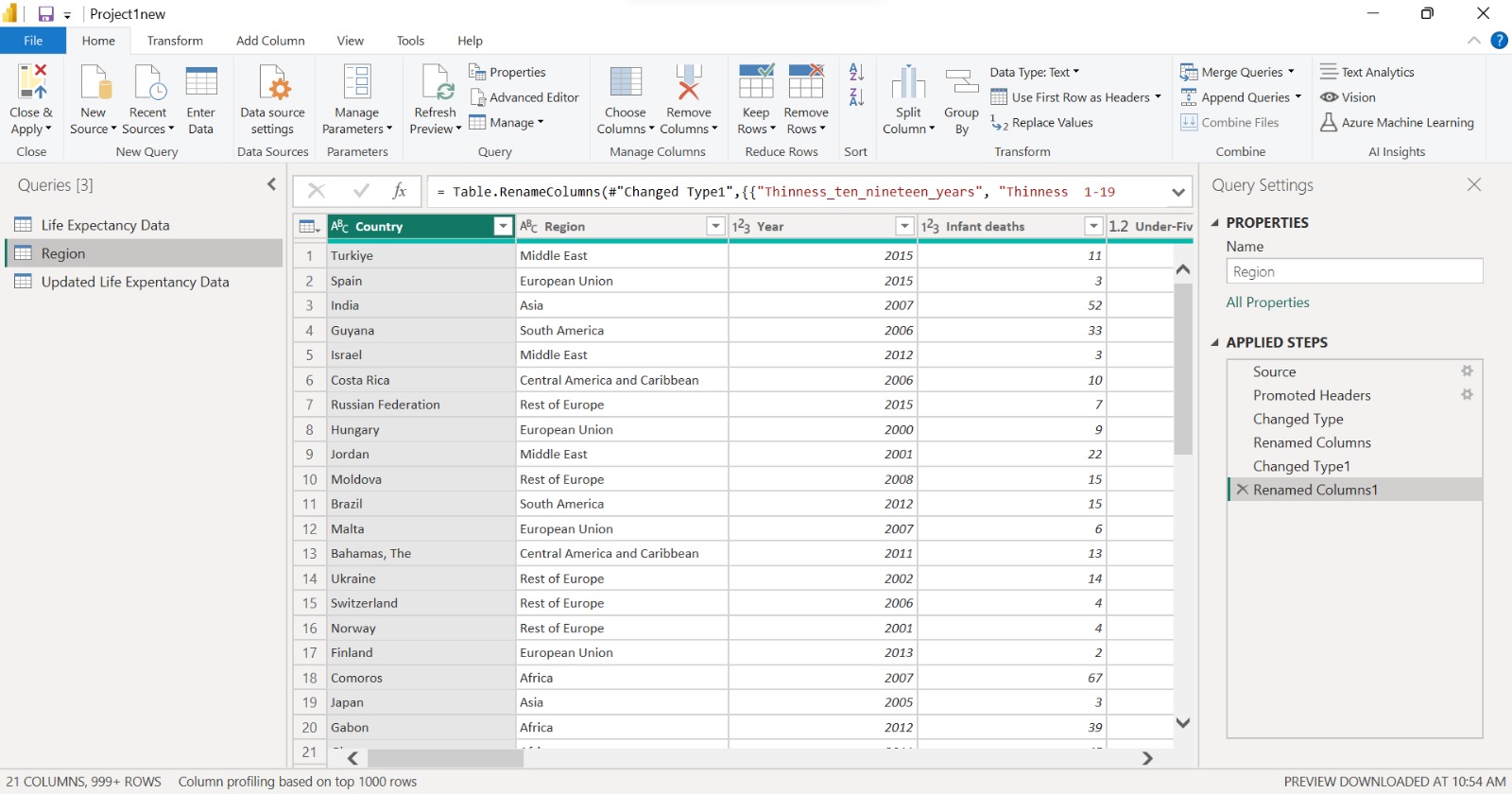
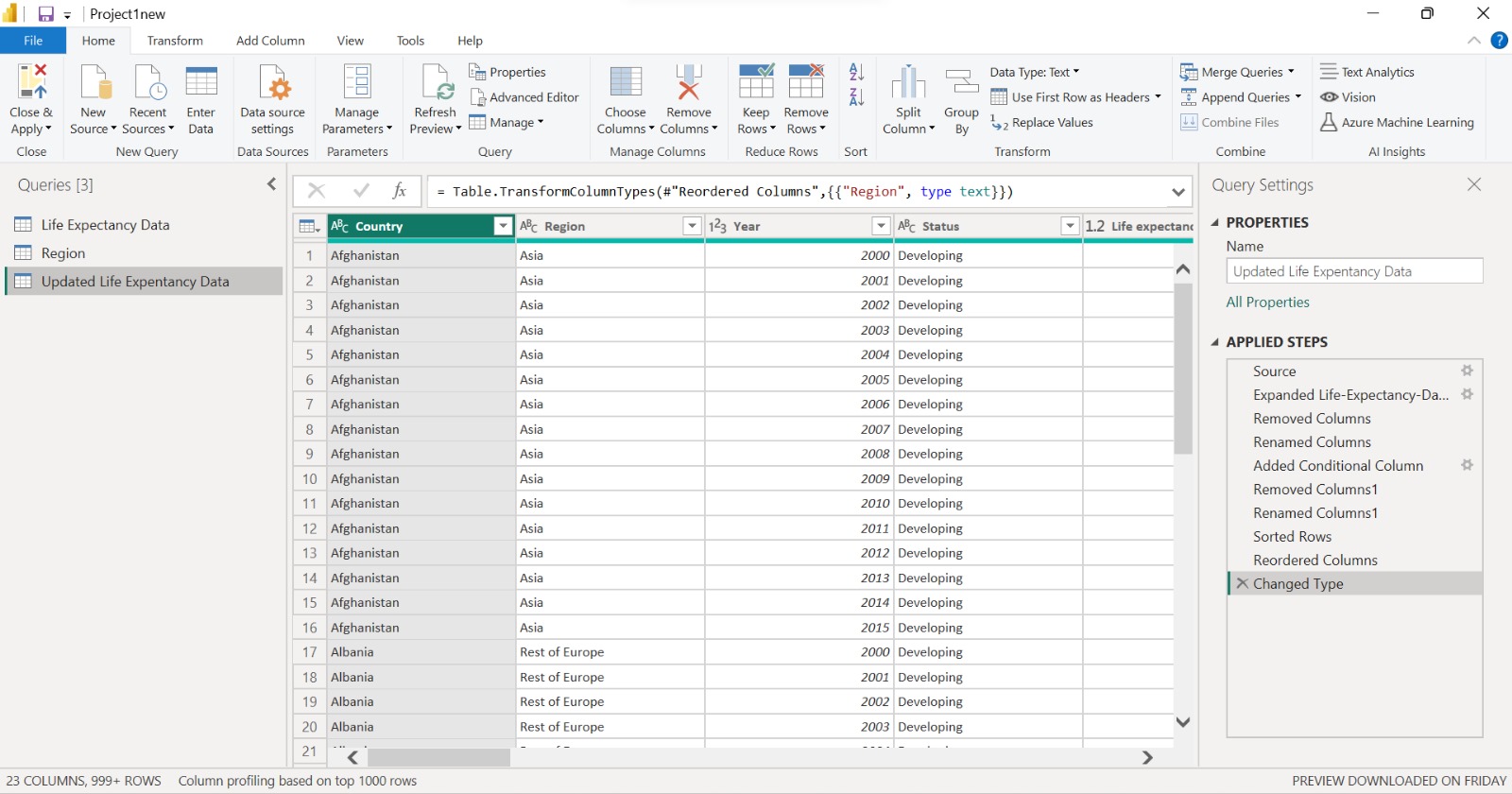
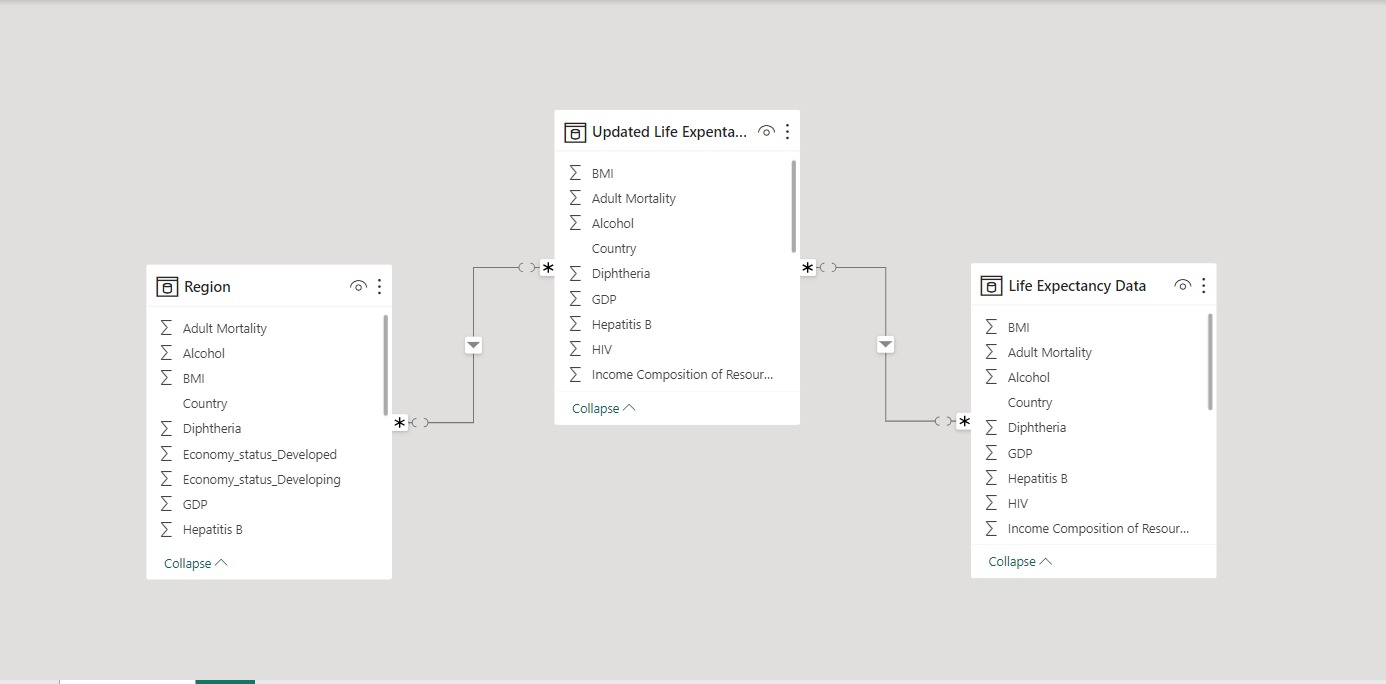


Table 2



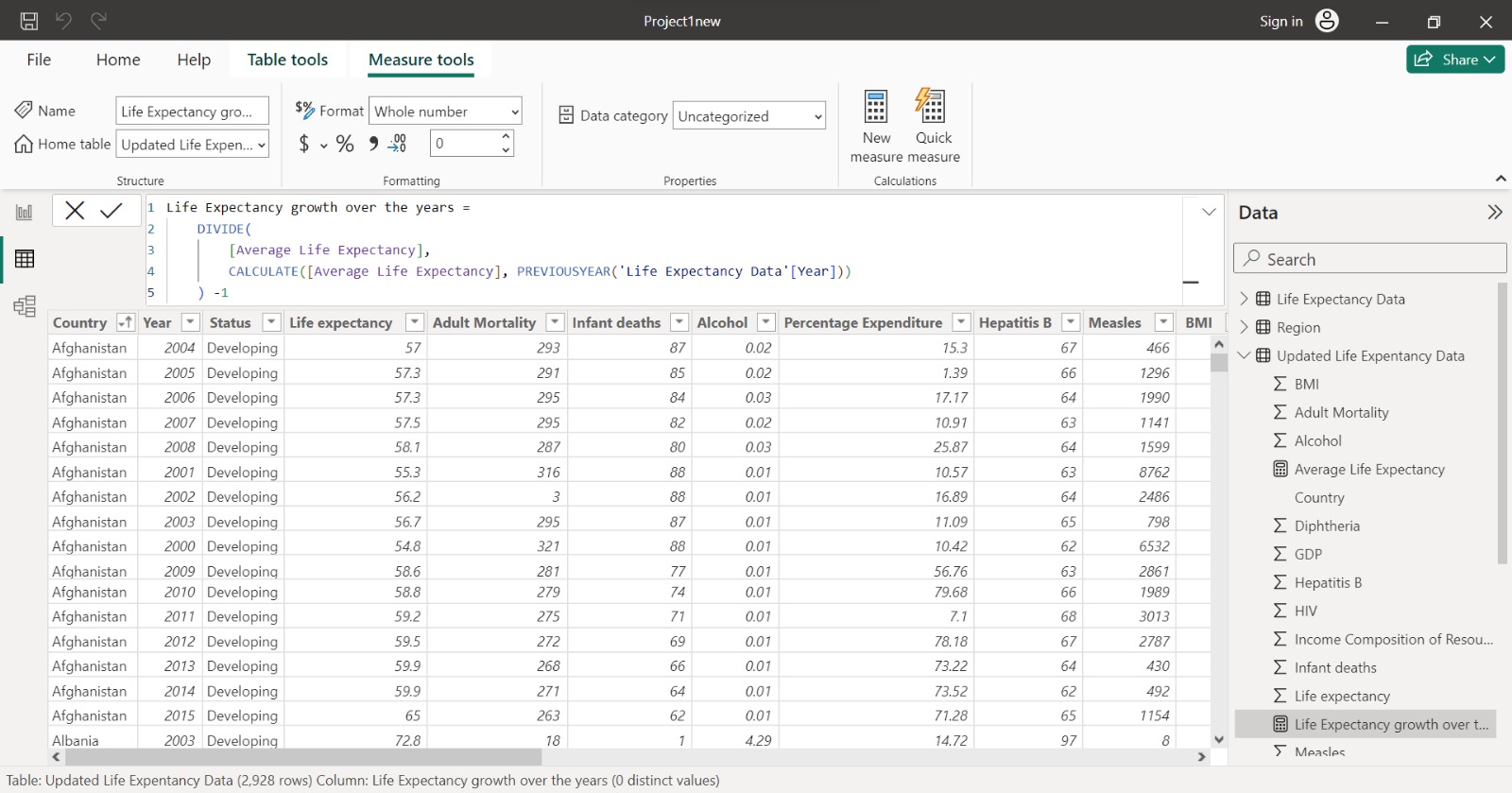
Merged Table

**Data Model**

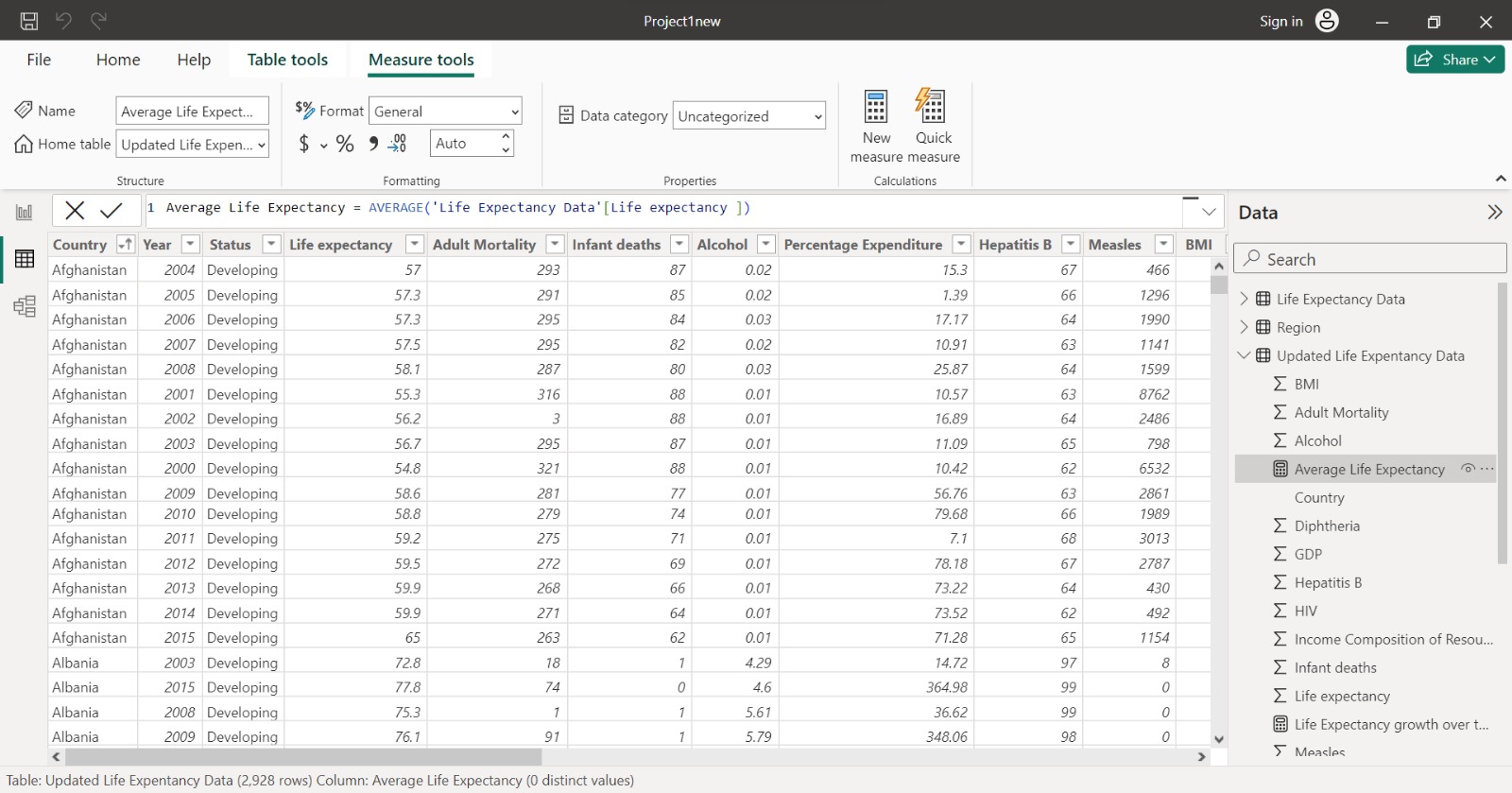


The data model shows a many to many relationship between the region and updated life expectancy using region as a primary key,and also created a relationship between updated life expectancy data and life expectancy data using country as a primary key

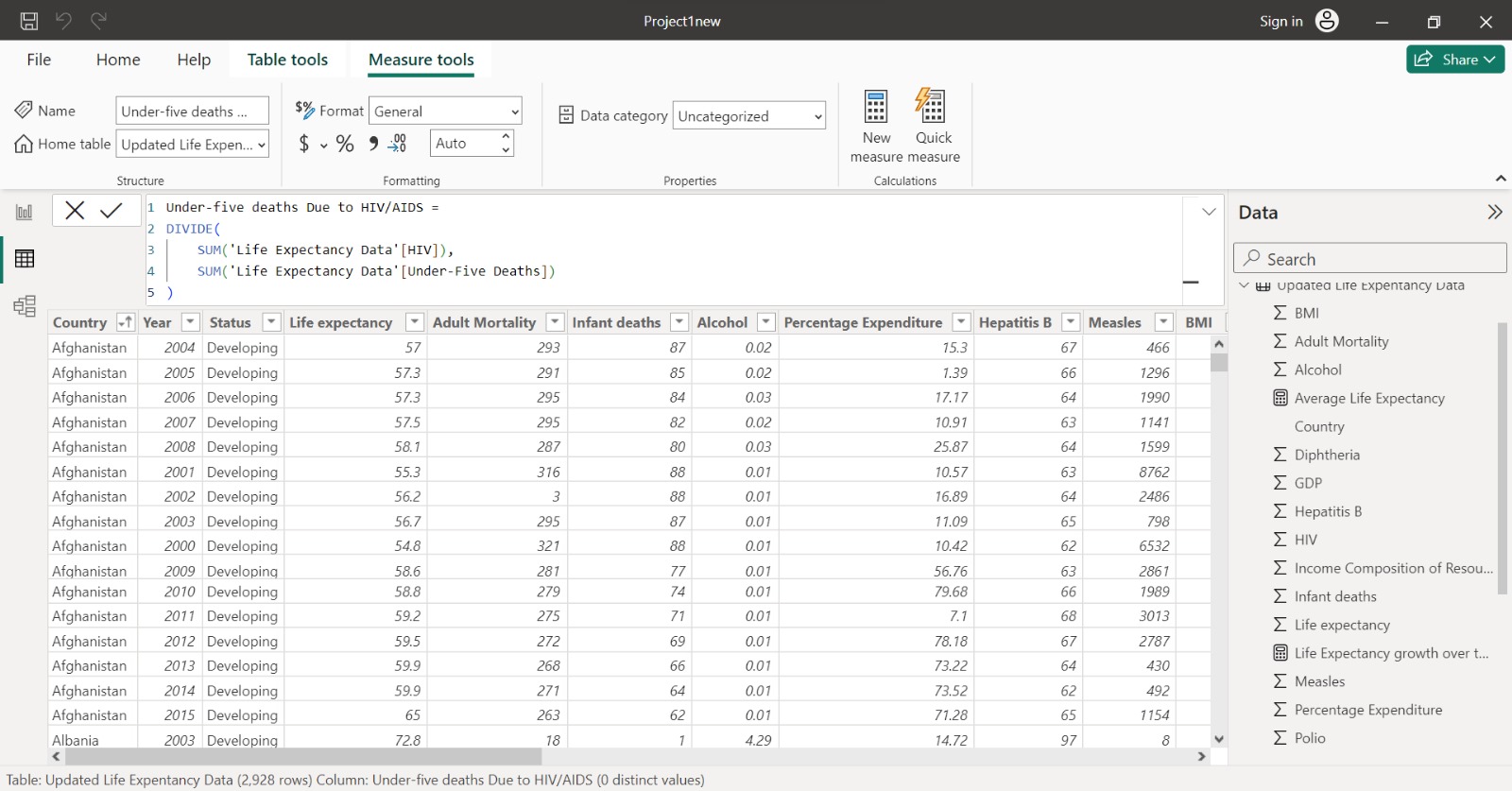
**DAX Commands**

****

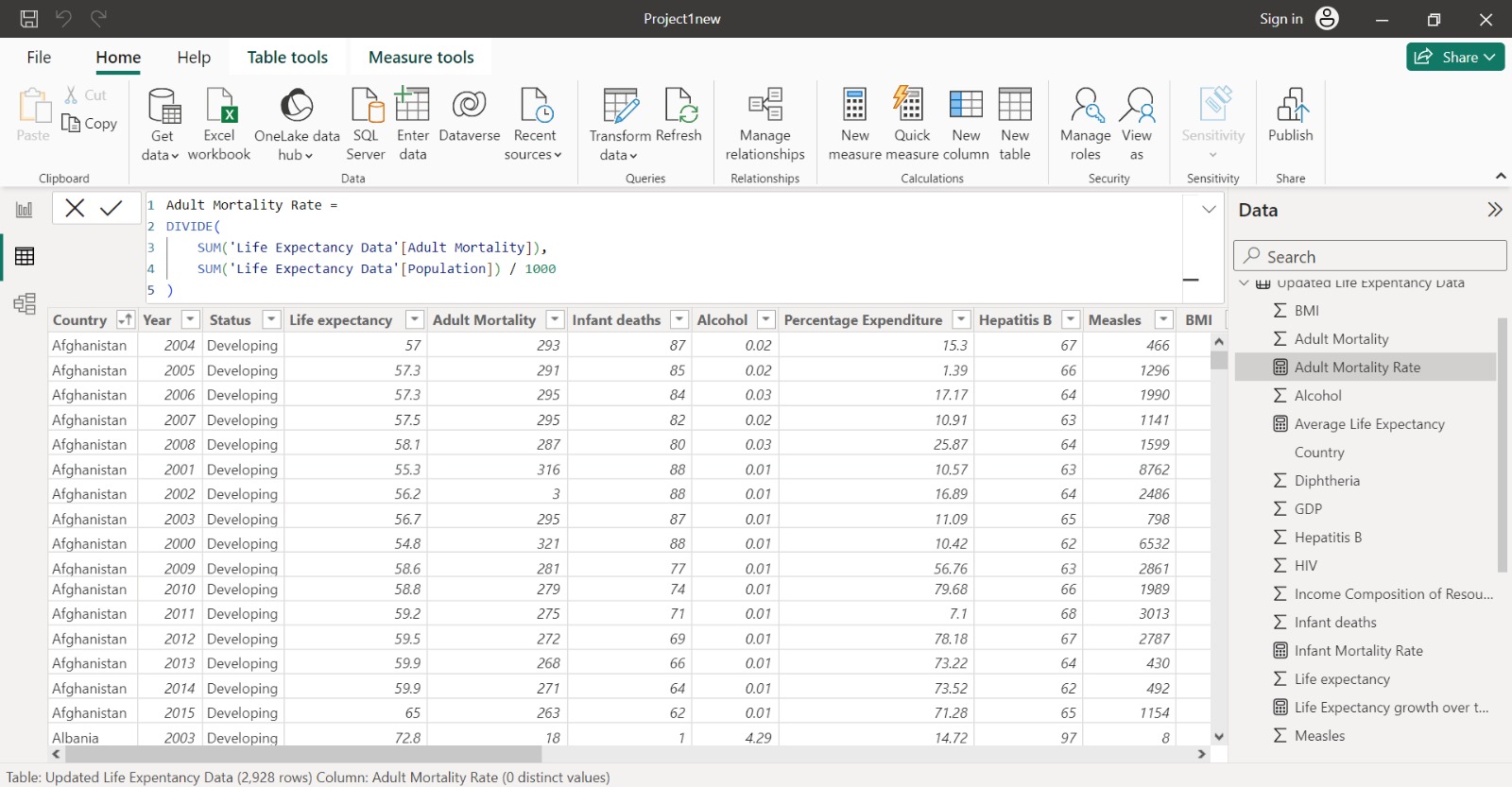
1. Created a DAX command for calculating the life expectancy over the years



2. Created a DAX command for calculating the average life expectancy



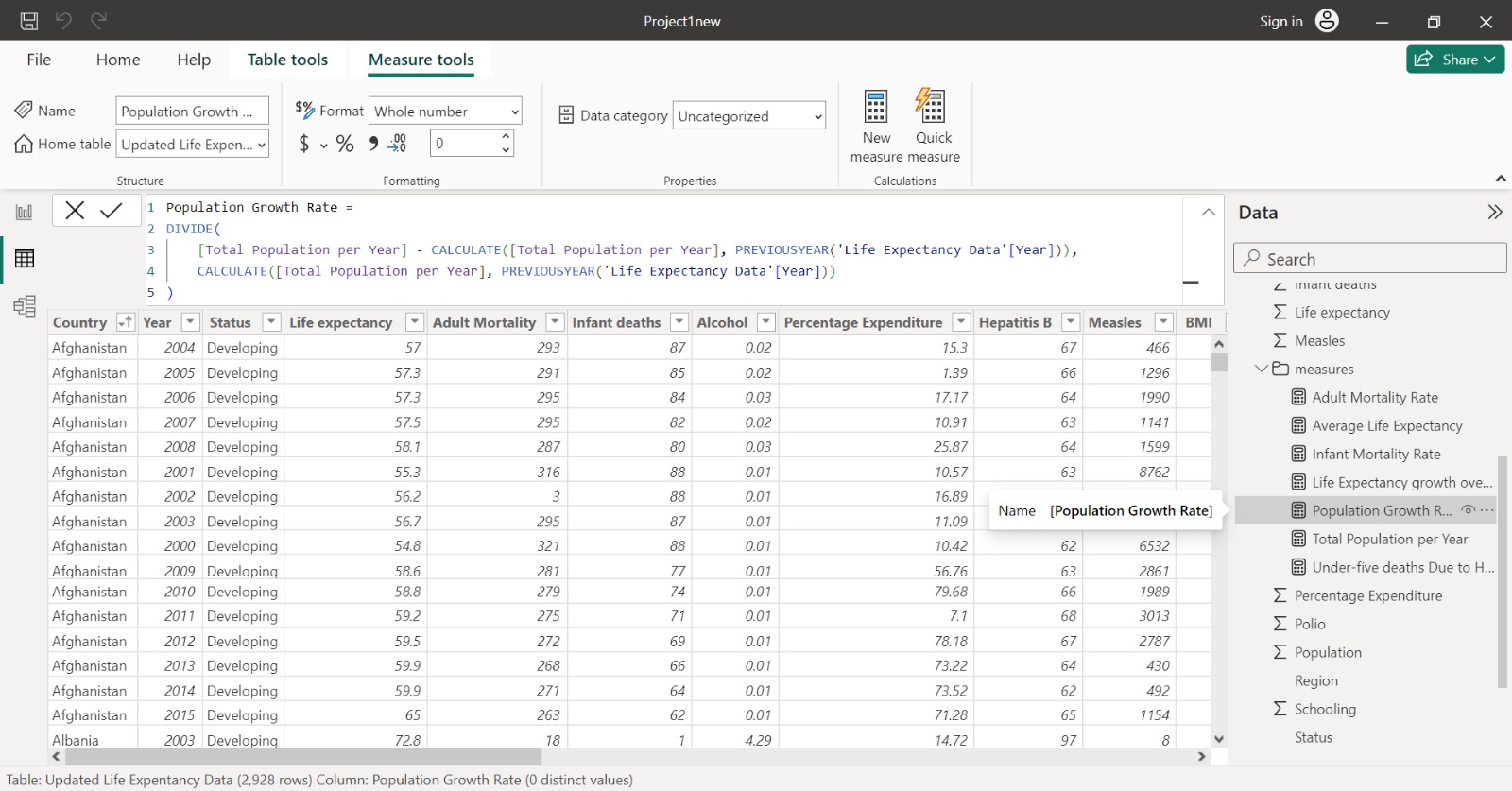
3.Created a DAX command to calculate the ratio of under 5 deaths due to HIV/AIDS



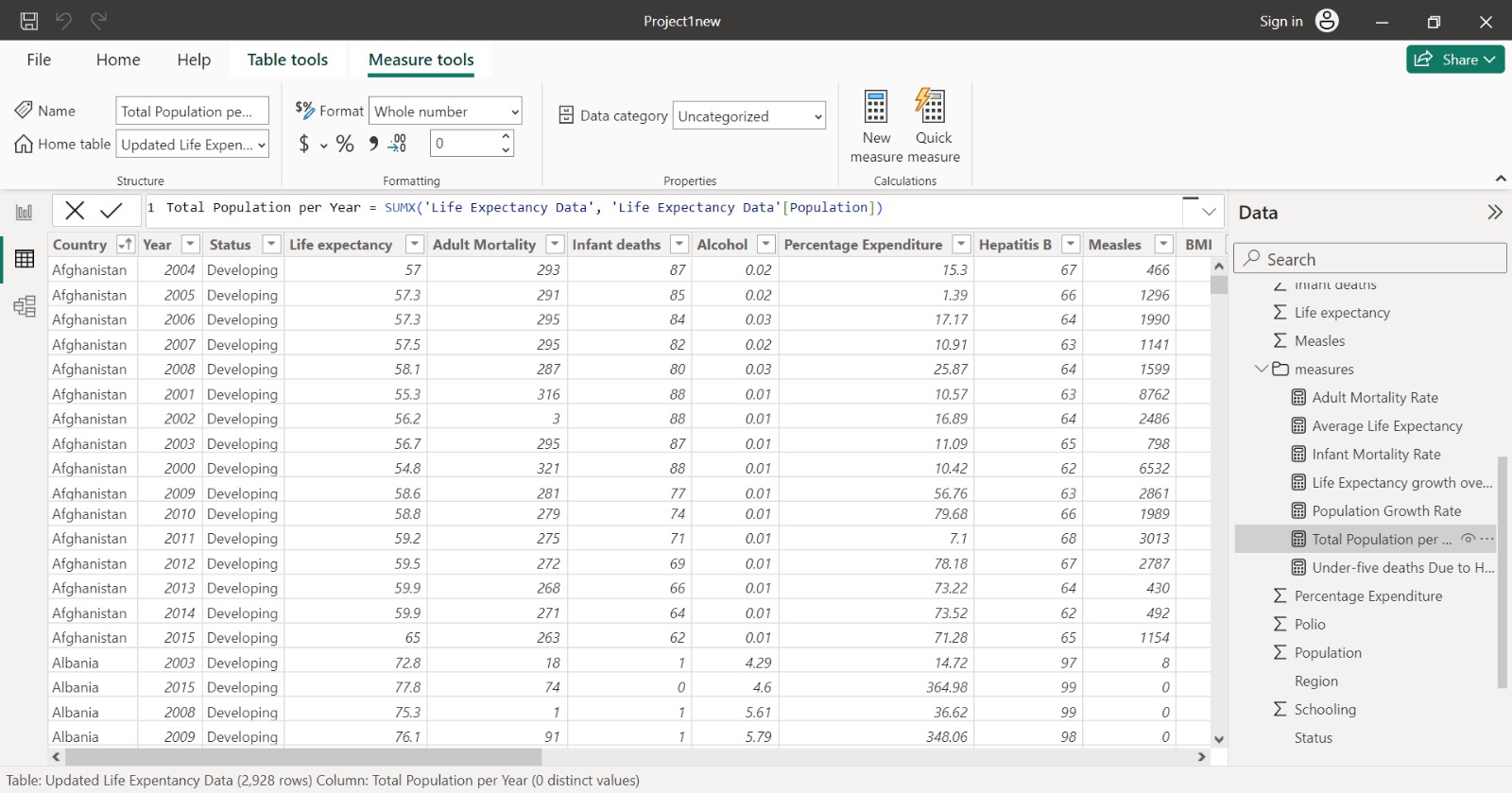
4.Created a DAX command to calculate the adult mortality rate with respect to the population



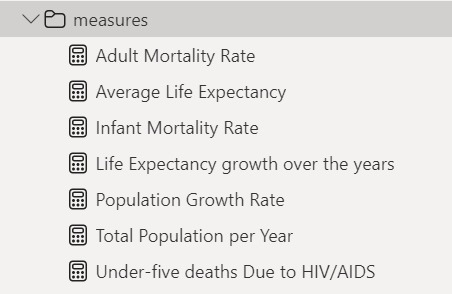
5.Created a DAX command to calculate the infant mortality rate with respect to the population



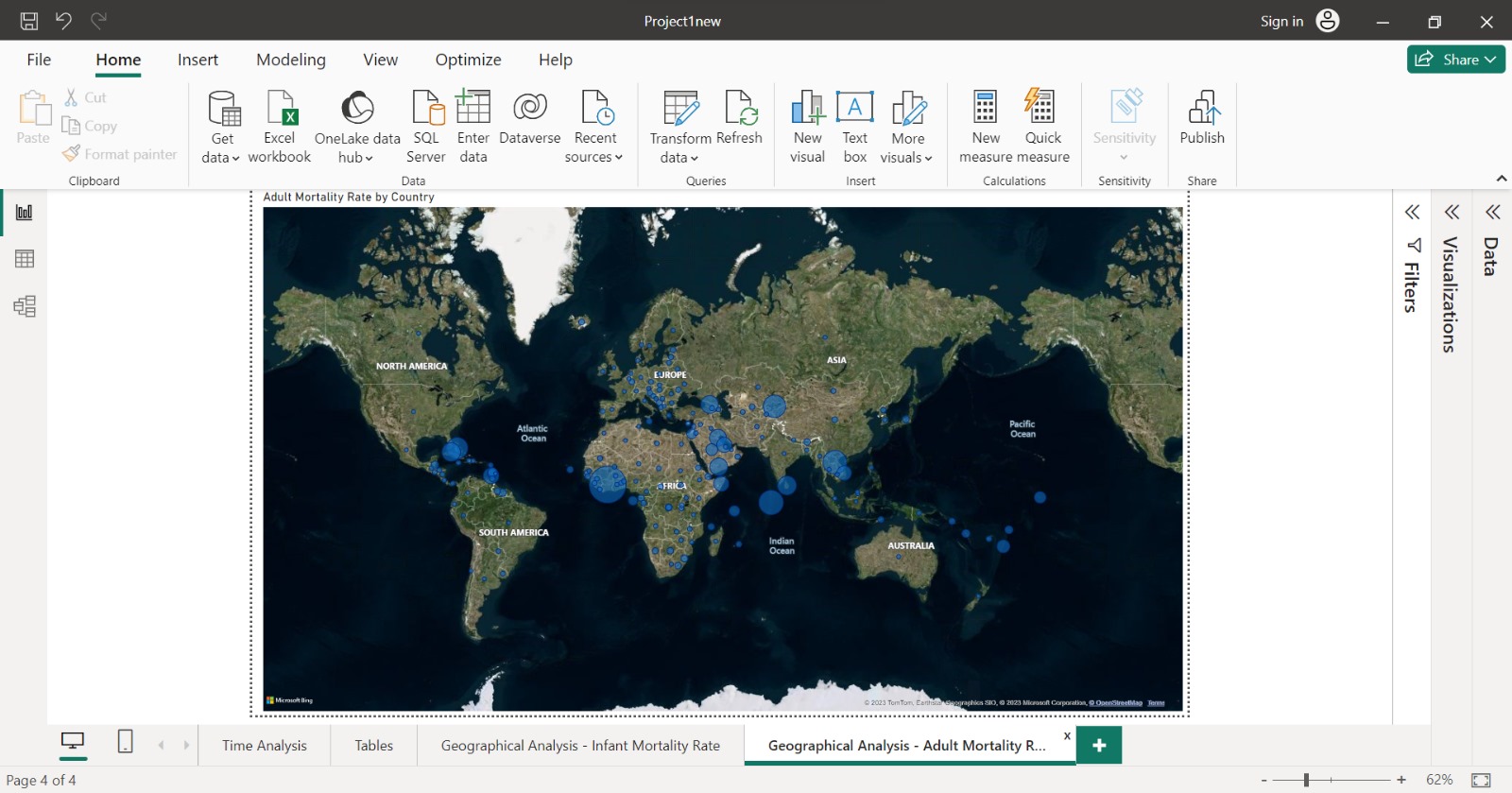
6.Created a DAX command to calculate the population growth rate over the year



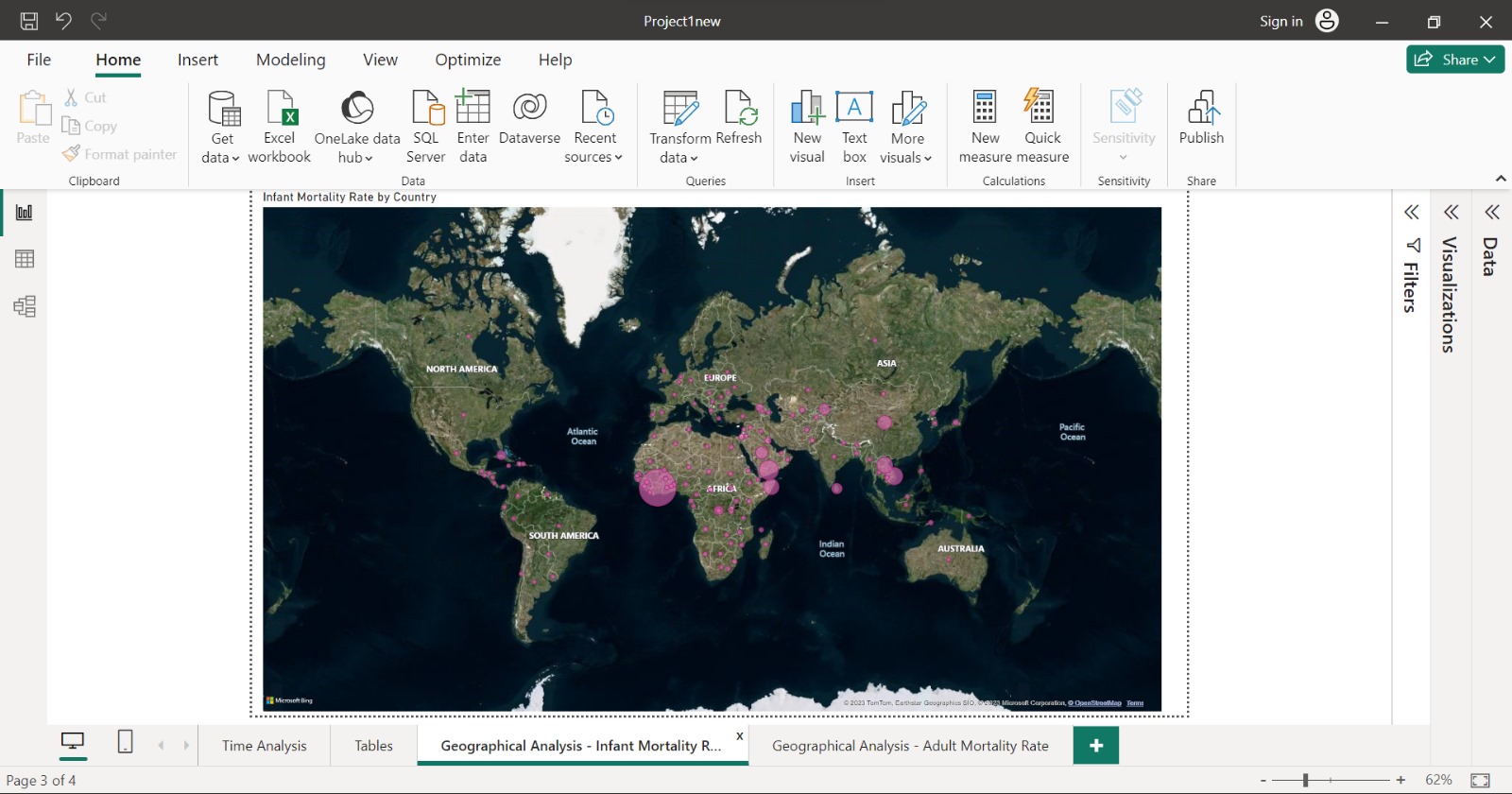
7.Created a DAX command to calculate the total population per year



**Geographic Analysis**

****

Adult mortality rates across the countries



Infant Mortality rates across the countries

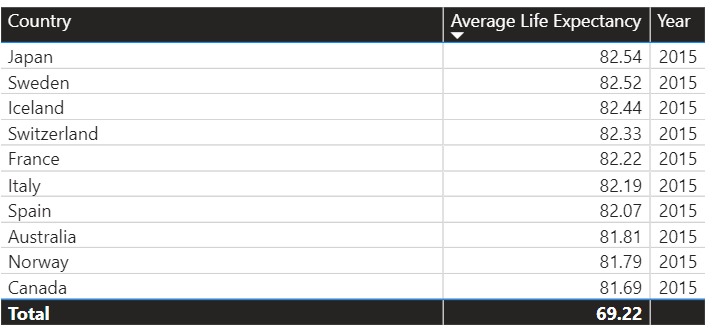
**Time Series Analysis**

****

Created a time series analysis between total adult mortality and total infant deaths by year

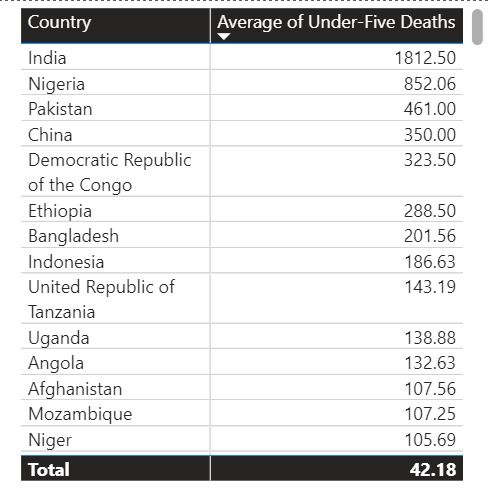
**Answering Analysis Questions**

* Which country has the highest life expectancy in the year 2015?



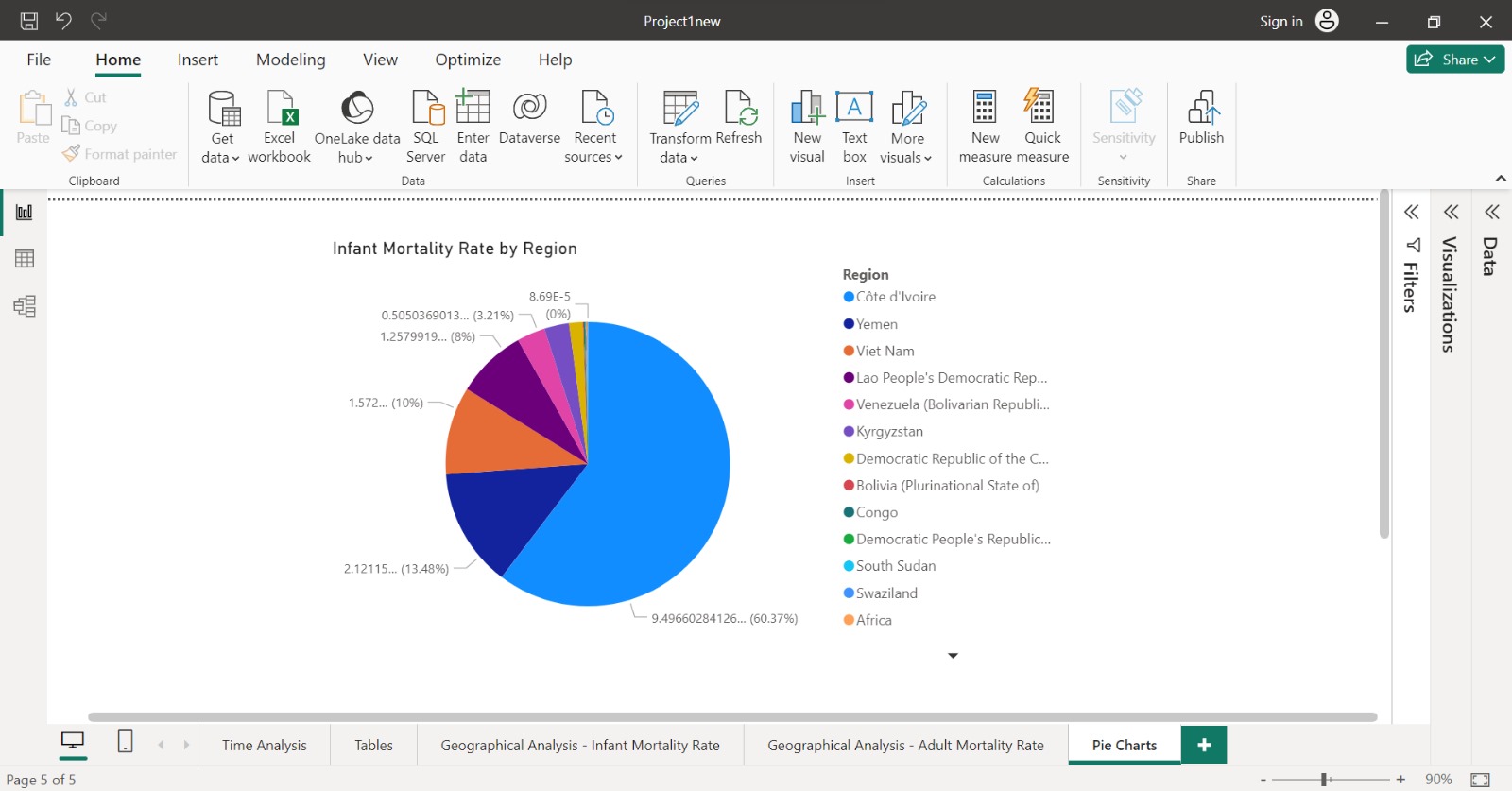
Japan has the highest life expectancy in the year 2015

* Which country has the highest under 5 years old death rate ?



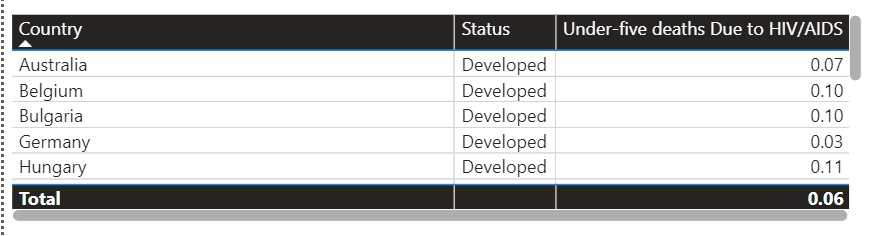
India has the highest under 5 years old death rate

* Which region has the highest infant mortality rate ?

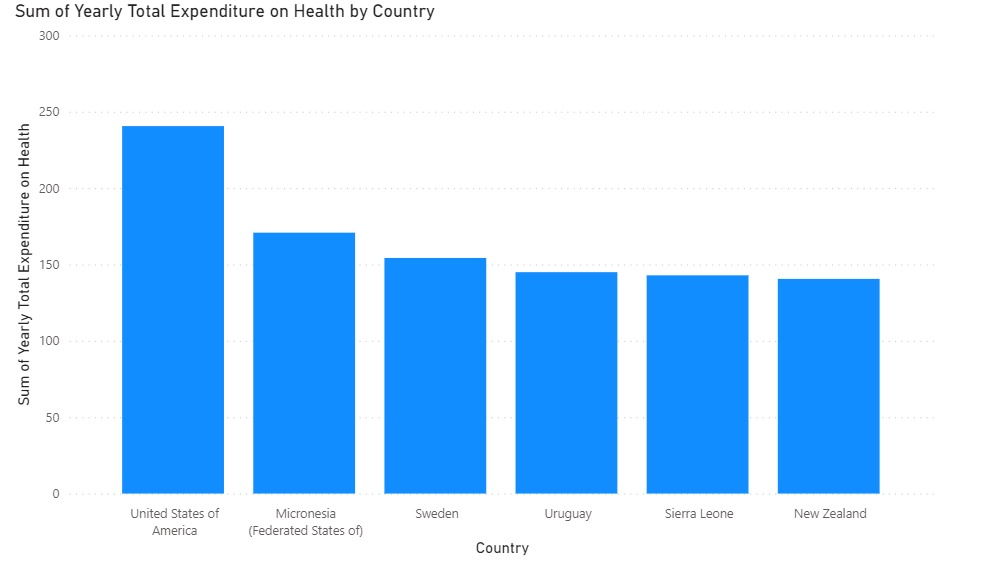


Côte d'Ivoire has the highest infant mortality rate.

* Which are the top 5 developed countries with least under 5 deaths due to HIV?

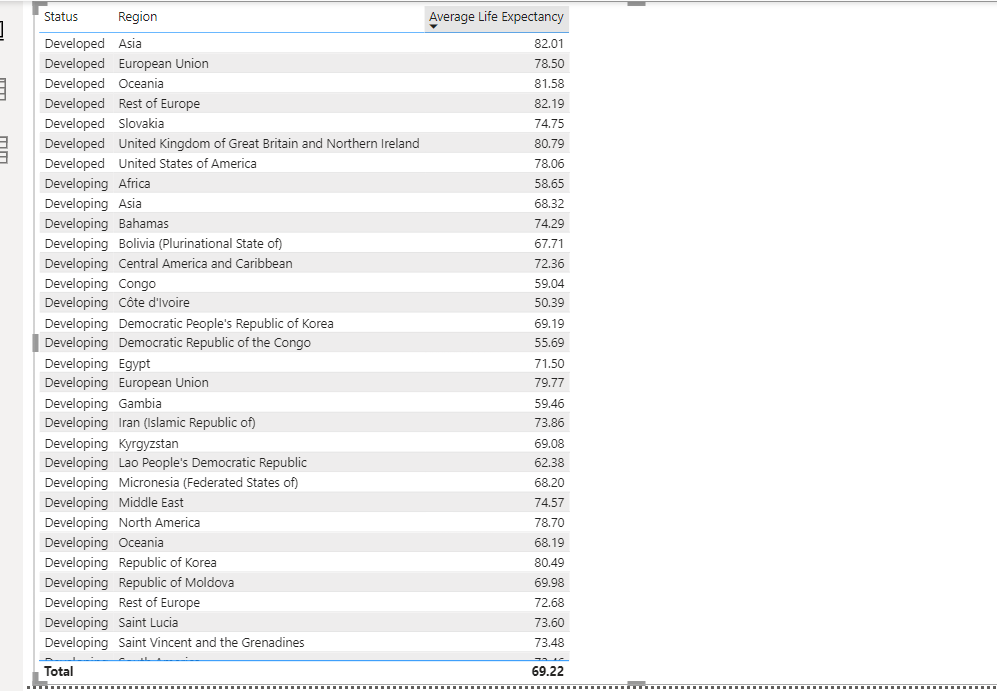


* Which country has the highest expenditure on healthcare?



The United States of America has the highest expenditure on healthcare.

**More Visualizations**

****

This visualization shows the average life expectancy across different regions based on their status

**Team Member Contribution**

PART A

Steffi George - Documentation

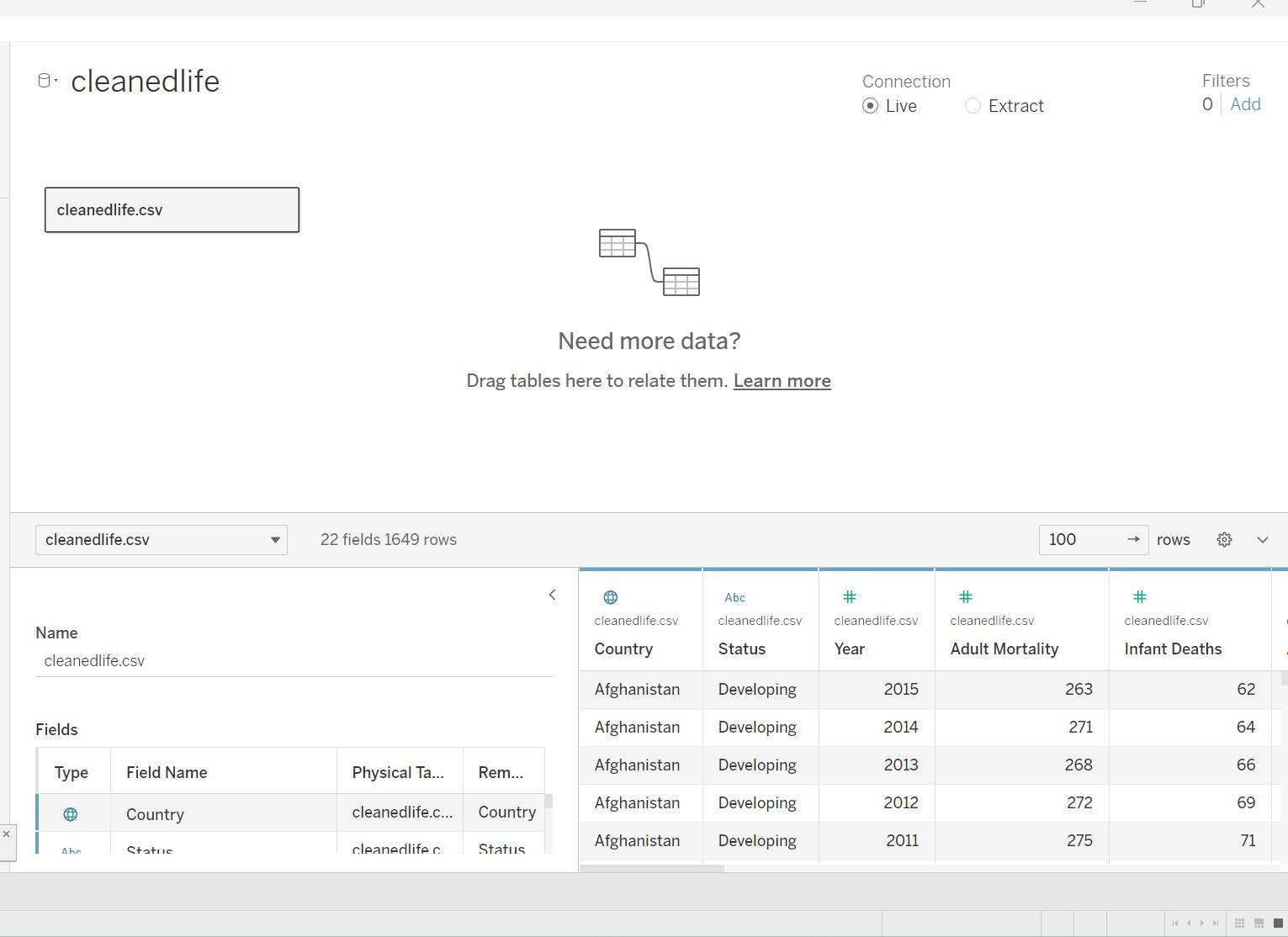
Hazel Ibasco -Power BI Analysis

**Part B - Tableau**

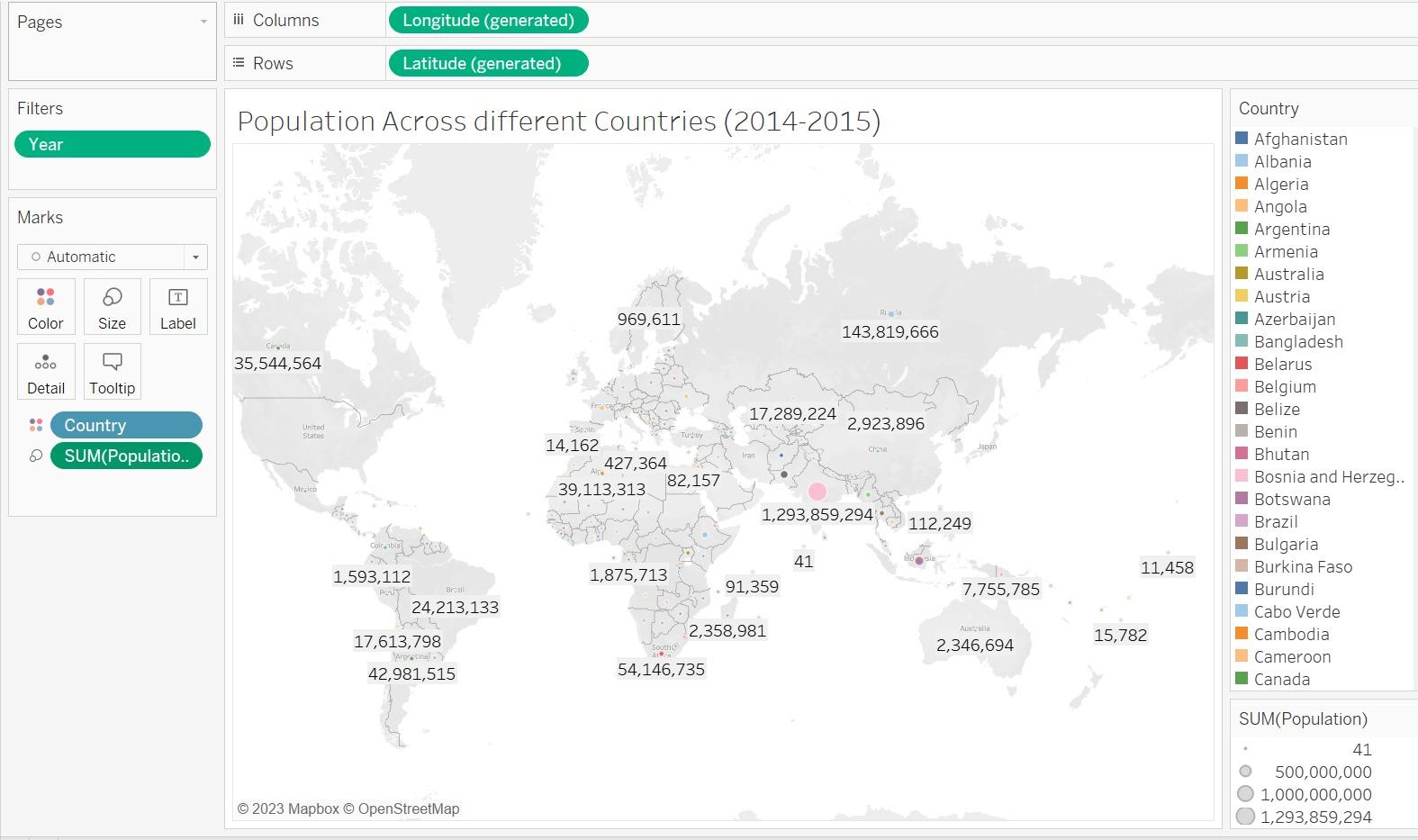
**Analysis Questions**

* Which Country has the highest schooling rate among the 7 countries provided, within the year 2012-2015? Australia, 60.8
* Which Country has the highest life expectancy in the year 2013? Belgium, 87
* Which Country has the highest under 5 death rate? India, 18,500
* Name two countries which have the sum of infant death rate as 0? Austria and Belize
* What is the minimum life expectancy of the developed countries in the year 2014? 73.4

**Loading the dataset to Tableau environment**

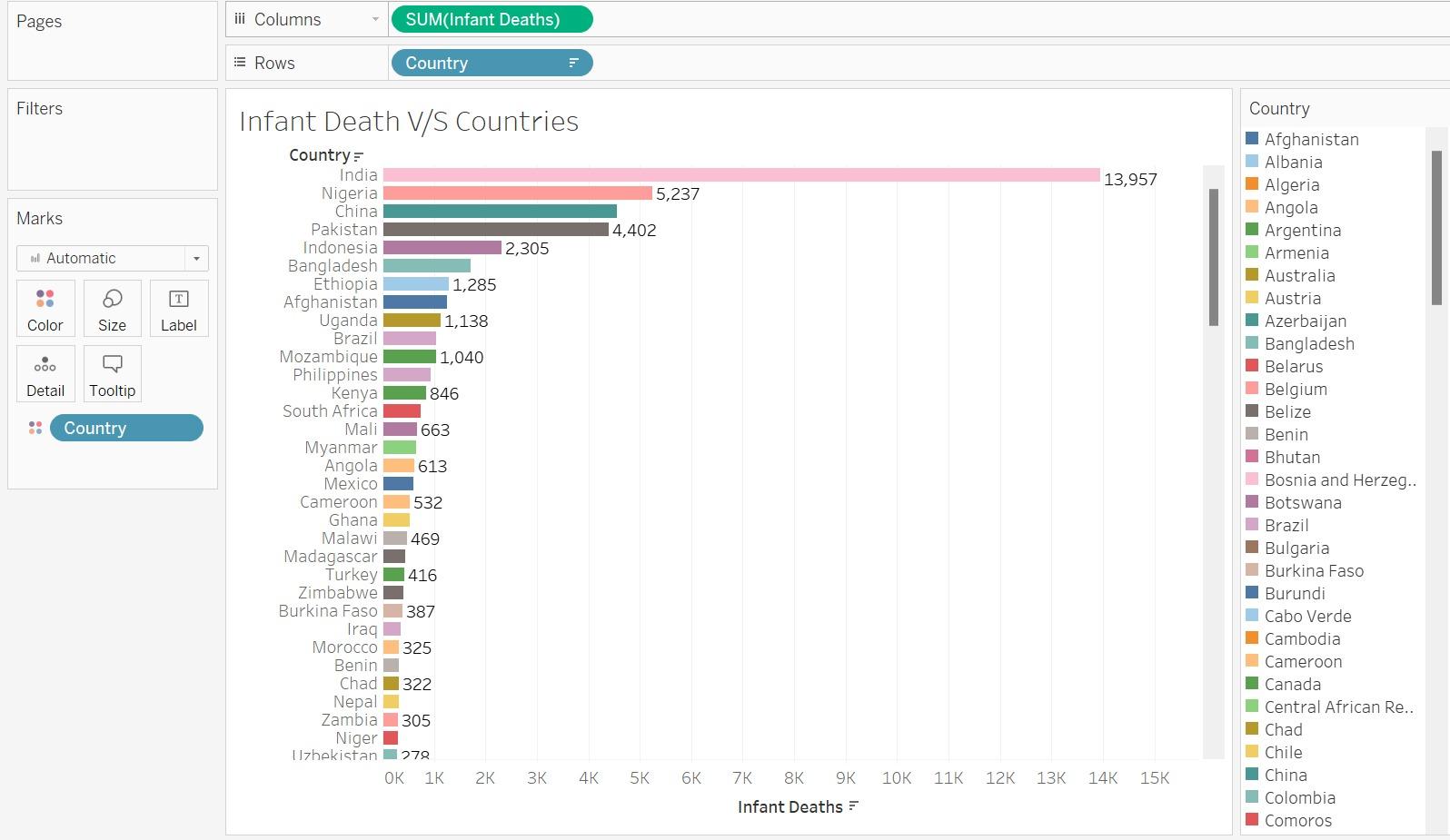


The csv file is named ‘cleanedlife.csv’. After we loaded it, we used the data interpreter for cleaning, and used Live connection.



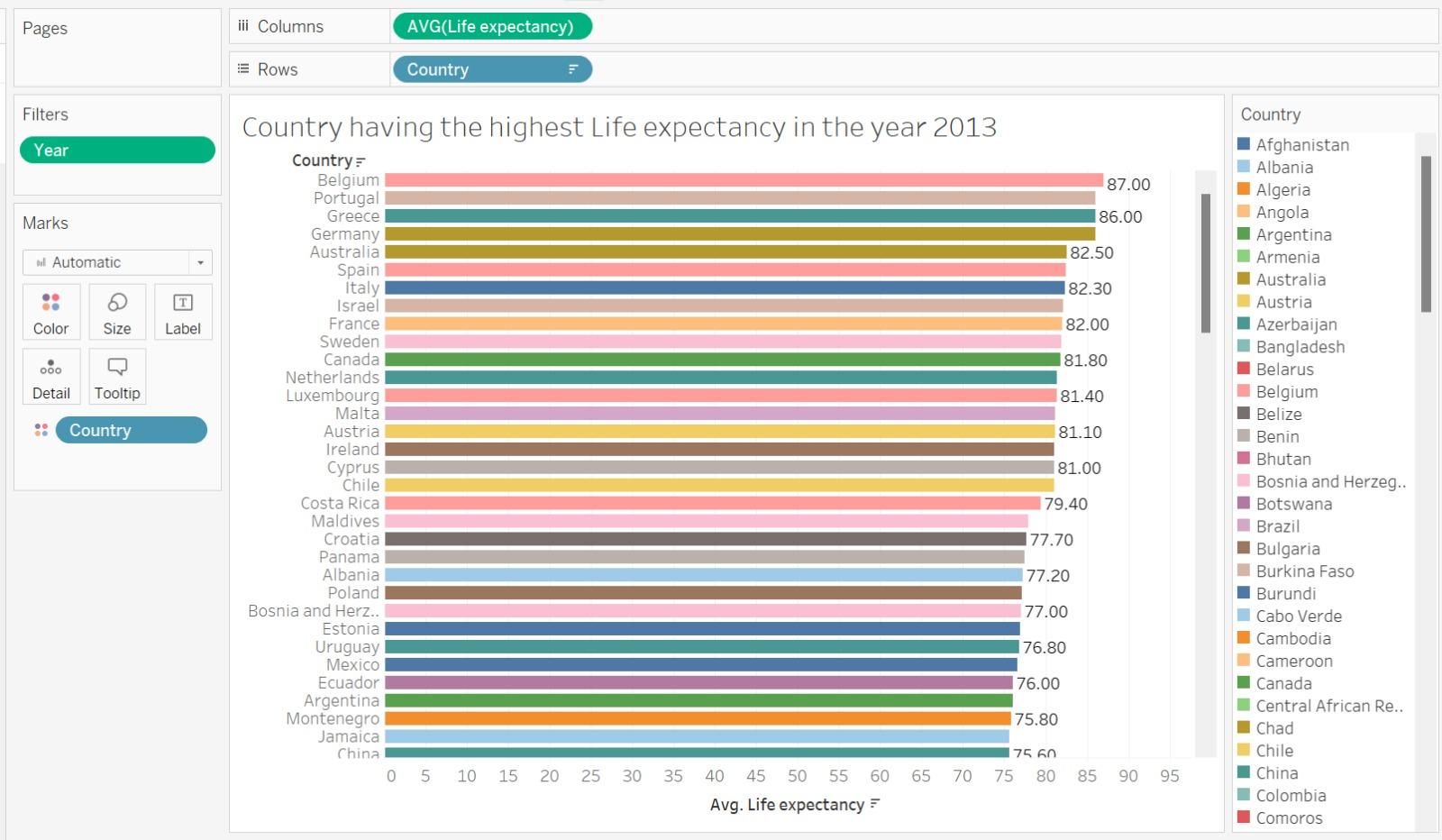
This graph gives us the insight of the population across different countries from the year 2014-2015. The size of the bubble indicates how large or how small the total population is for each country, while the color differentiates the country itself.

**Important illustrations that explain our chosen dataset**



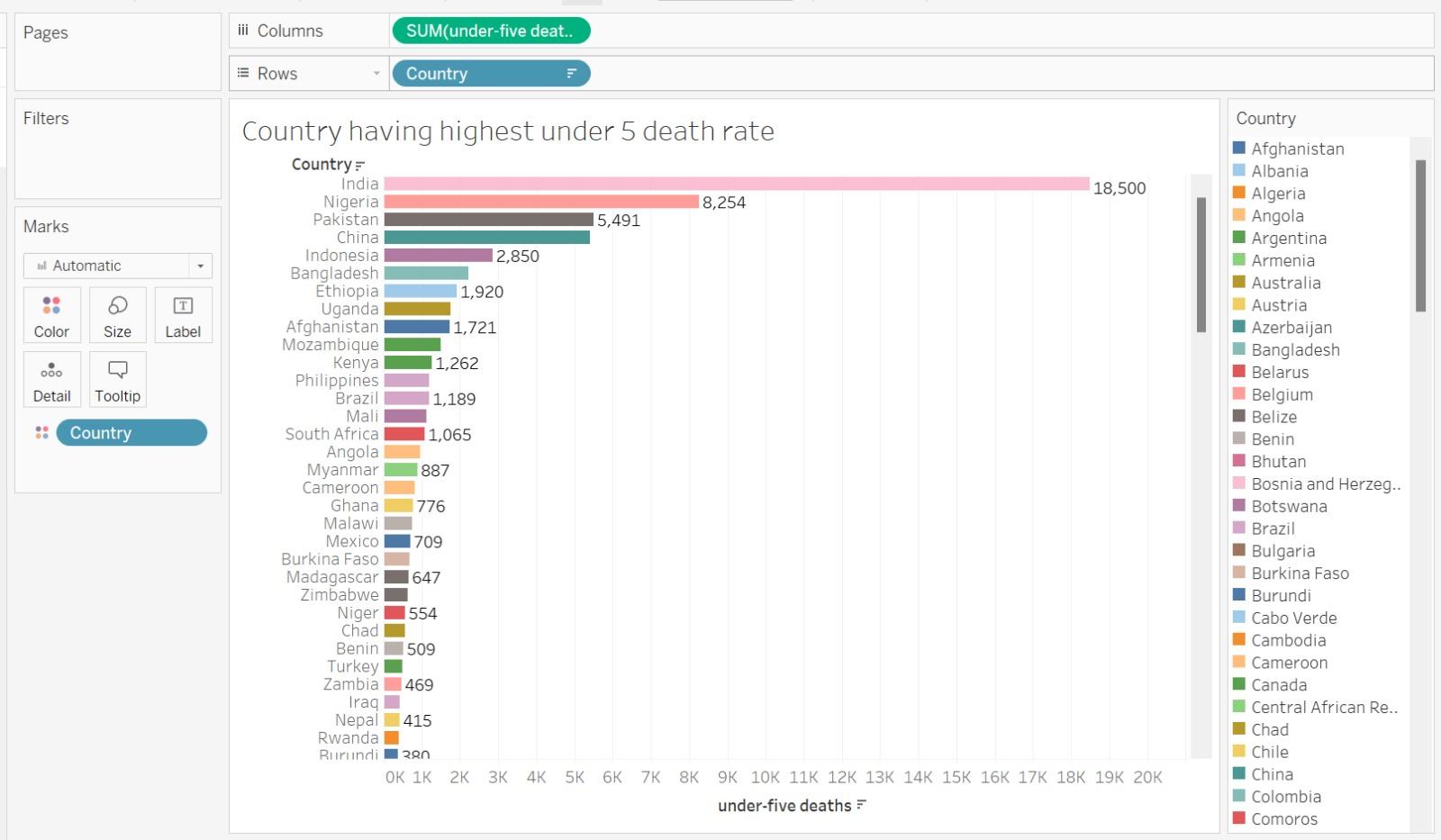
**Infant Death per countries**

This chart shows the top countries with the highest number of infant deaths in the year 2000-2015. It shows that India ranked first with a whopping total of 13,957 deaths of infants, followed by Nigeria with a total of 5,237 deaths.



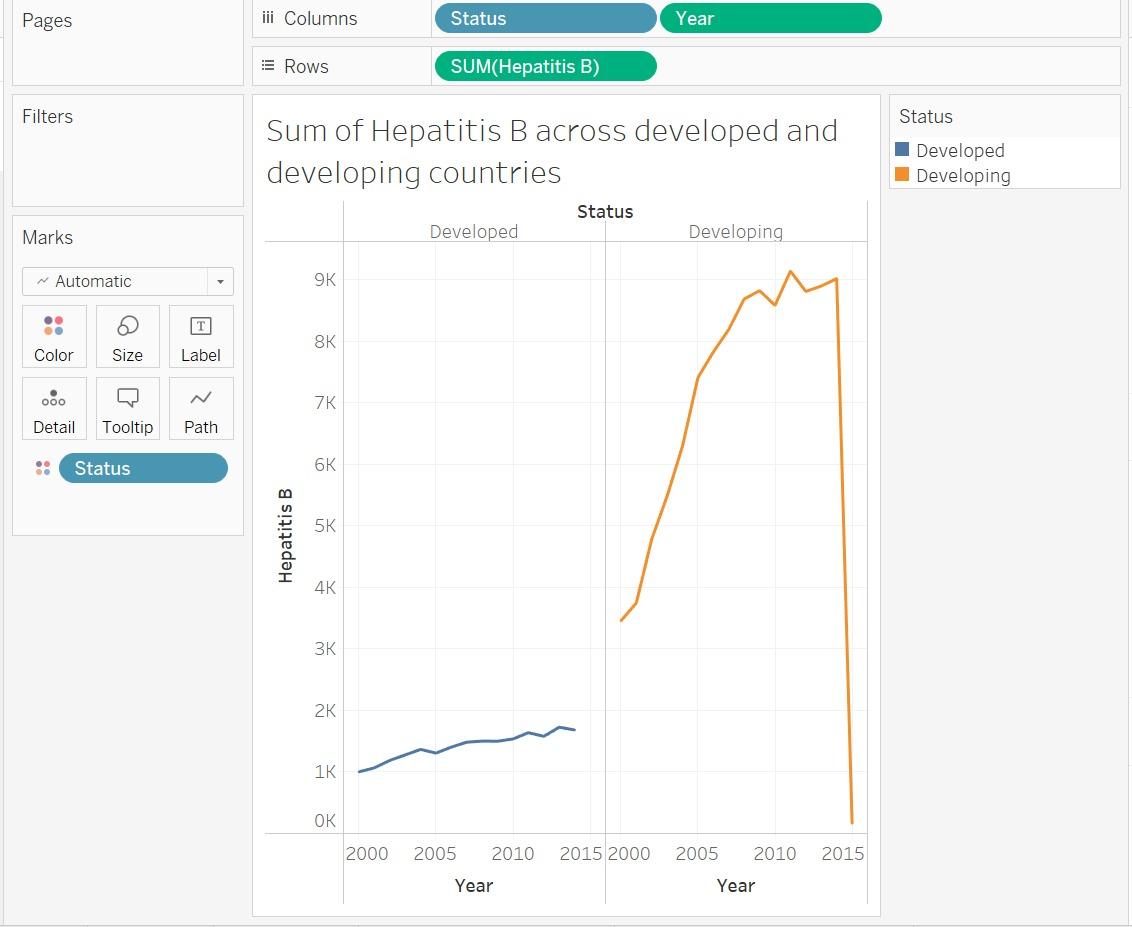
**Countries with the highest average life expectancy in 2013**

The chart indicates that Belgium has the highest average life expectancy in the world for the year 2013, age 87. It also shows that the next three countries ranked after Belgium are also from Europe (Portugal, Greece, and Germany). This answers our analysis question - Which Country has the highest life expectancy in the year 2013?



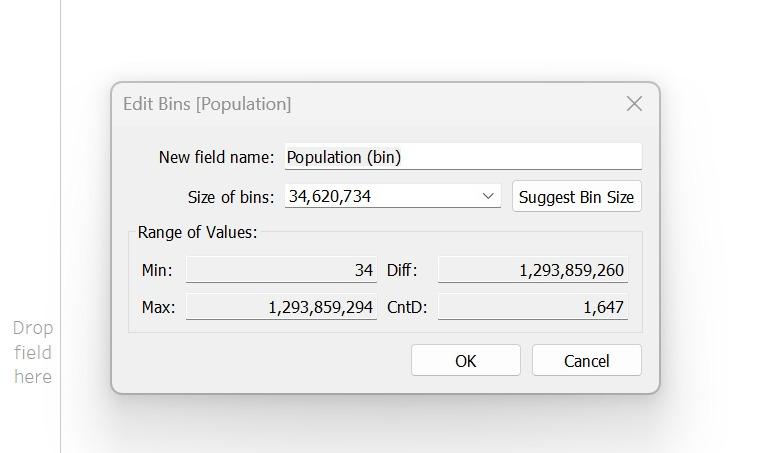
**Countries with the highest ‘Under 5’ death rate**

This chart shows the countries with the highest death rate of children under 5 years old from the year 2000-2015. It shows that the top ranking countries are very similar to the top ranking countries in the previous chart which shows the highest countries with infant deaths. This answers our analysis question - Which Country has the highest under 5 death rate?

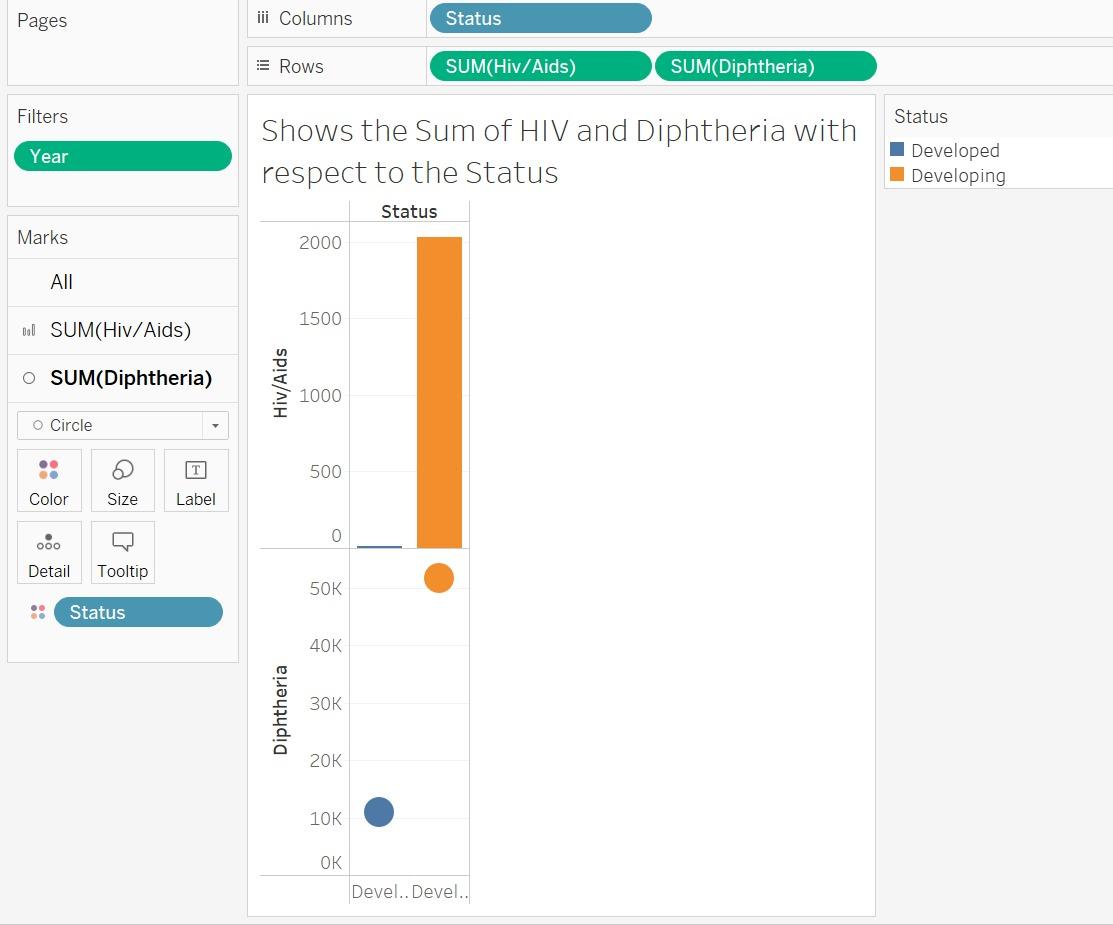
****

**Hepatitis B cases across Developed and Developing countries in 2000-2015**

This line graph shows the difference between the number of cases of Hepatitis B in developed and developing countries. The huge difference of cases between the two categories is very apparent from the year 2000-2015 with the de

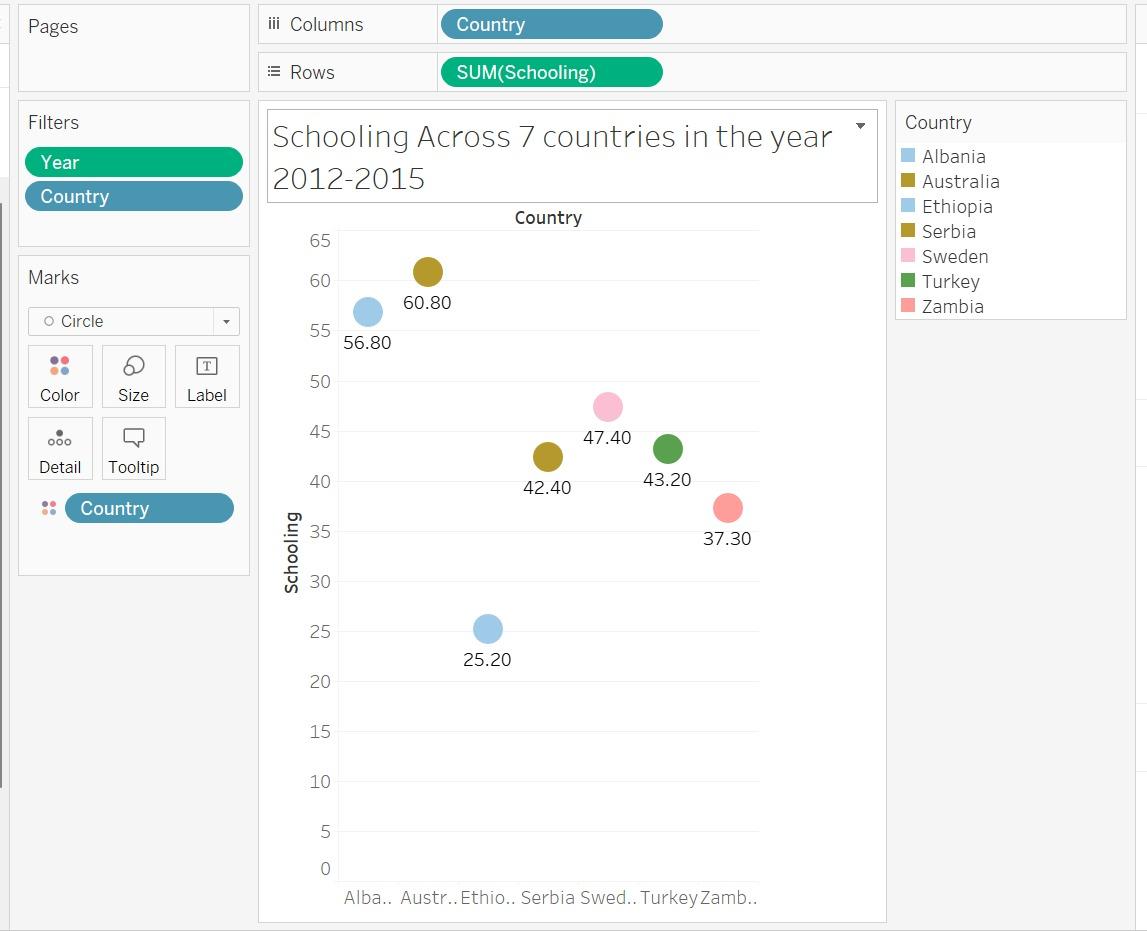
****

We created bins for the population with the difference of 1,293,859,260 billion.

****

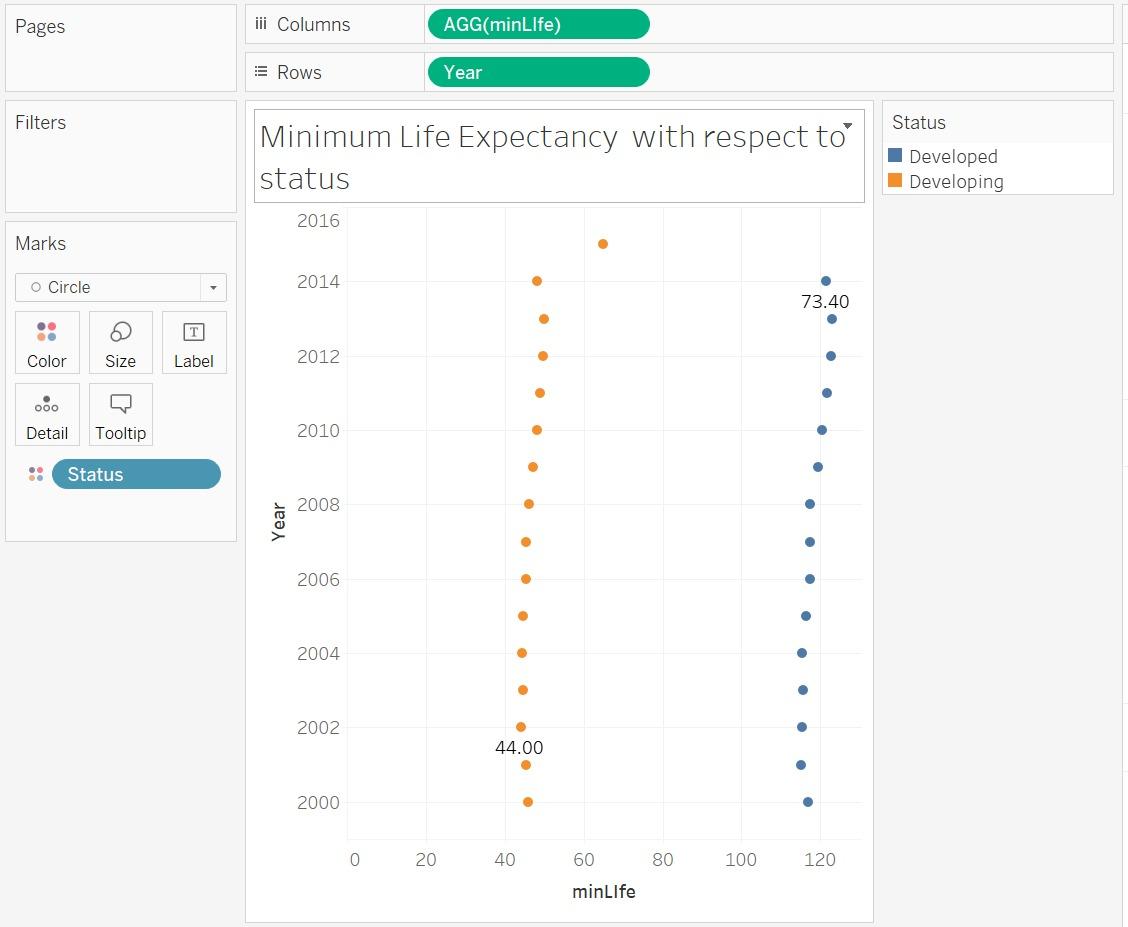
**HIV and Diphtheria cases in developed and developing countries**

In this illustration, it shows a combination of two charts: bar and circle. These charts explain the number of cases of HIV and Diphtheria between developed and developing countries in the year 2007-2015. Again, it shows a huge difference in numbers between the statuses with the developing countries ranking way higher than the developed countries.

****

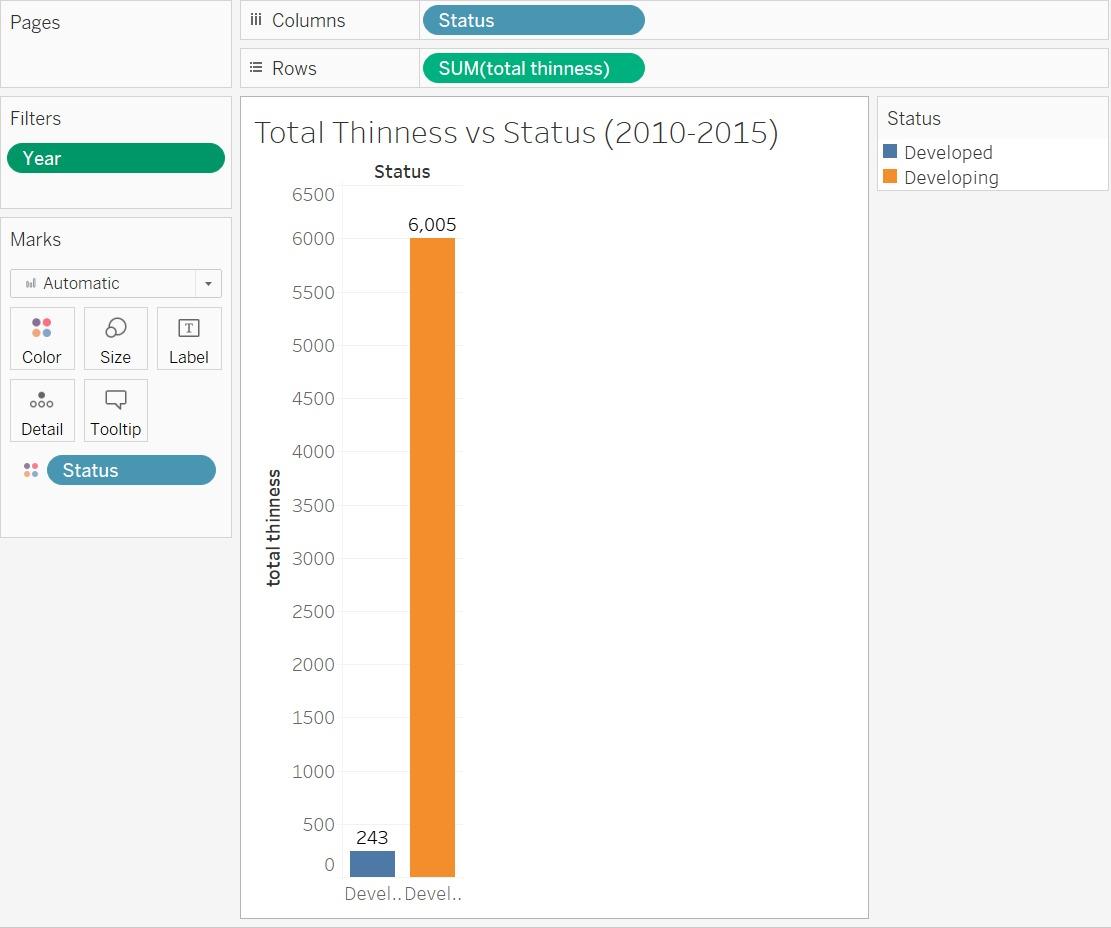
**Schooling across 7 chosen countries in 2012-2015**

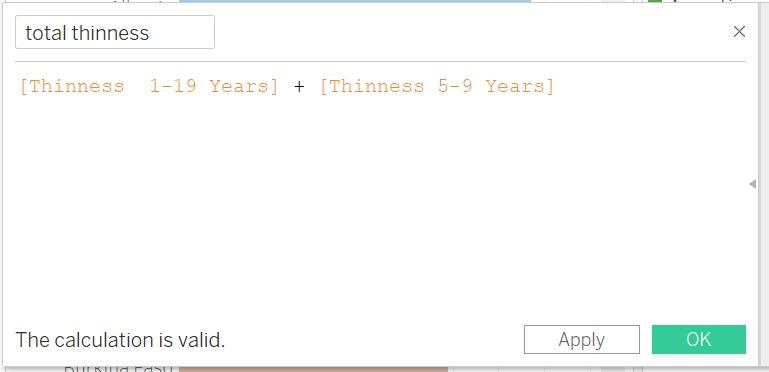
This circle views graph illustrates the average of schooling across 7 chosen countries from both developing and developed countries. As per this chart, Australia, being a developed country, has the highest average schooling compared to a developing country, Ethiopia, which has the lowest average in this chart. This answers our analysis question - Which Country has the highest schooling rate among the 7 countries provided, within the year 2012-2015?

****

**Minimum Life Expectancy vs Status**

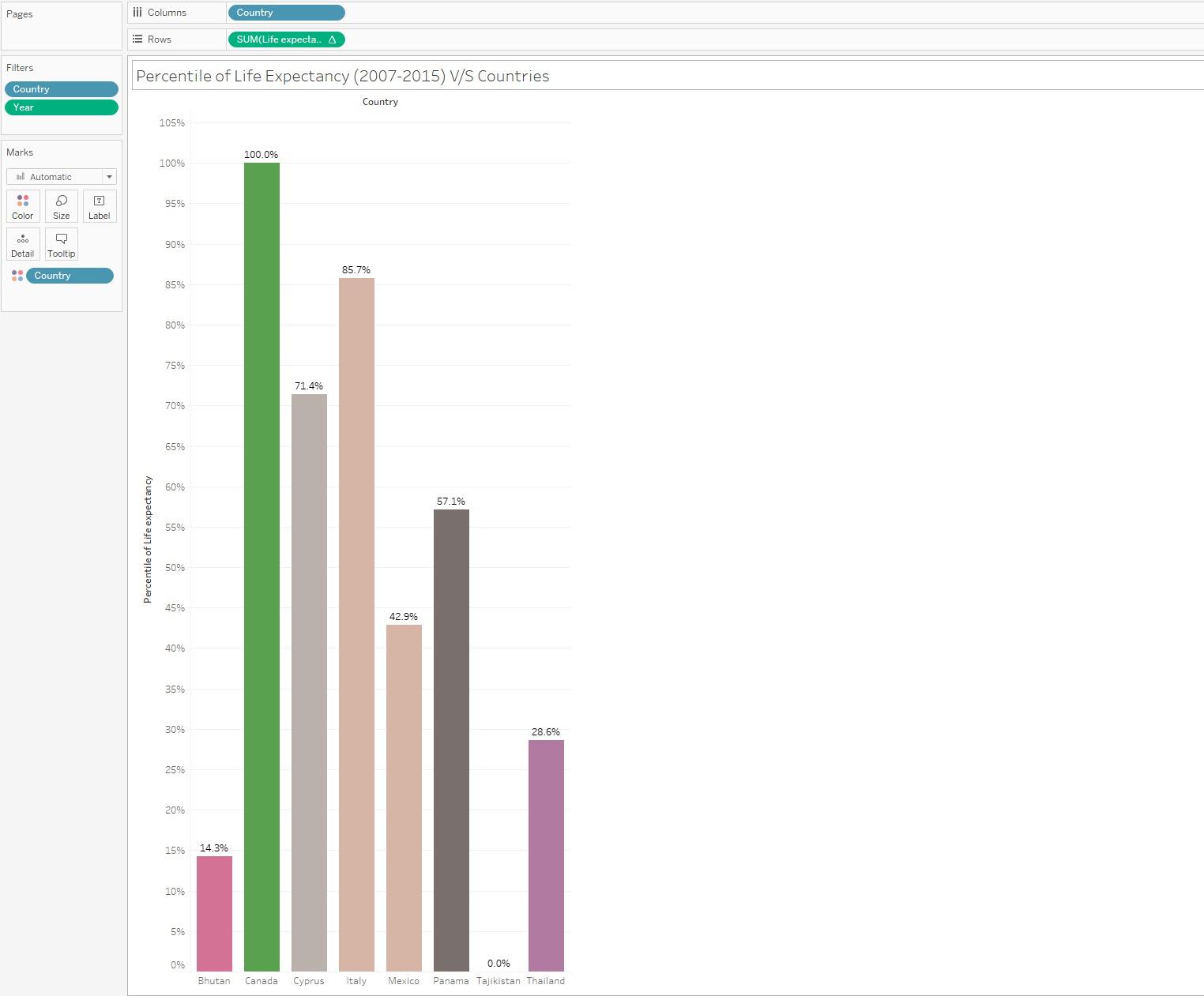
This chart shows the lowest life expectancy between developed and developing countries throughout the years. Before we created the chart, we first created a calculated field which allows us to get the minimum life expectancy in both statuses. The syntax for the calculated field is MIN([Life expectancy]). The field was used to achieve the chart, which leads to the conclusion that the developing countries achieved the all-time low of life expectancy in 2002, which was 44 years old. The developed countries’ minimum life expectancy was 73 years old, which is still pretty high. This answers our analysis question - What is the minimum life expectancy of the developed countries in the year 2014?

****

****

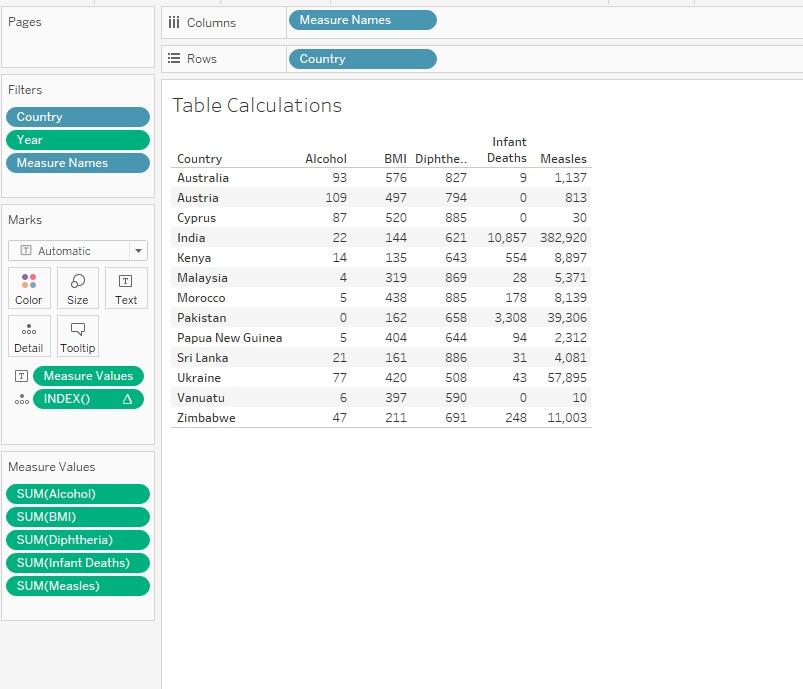
**Total thinness in developed and developing countries**

This bar chart shows the total thinness between developed and developing countries. We have also created a calculated field before creating the graph itself. In the calculated field, the thinness of 1-19 years old and the thinness of 5-9 years old were added together. We applied the field in the chart, thus showing us the large difference between the statuses.

****

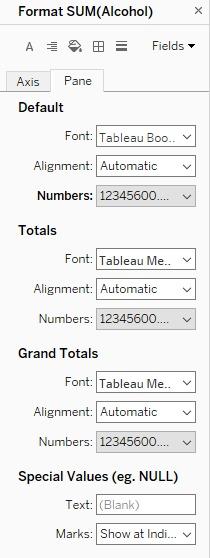
**Percentile of life expectancy across the chosen countries in 2007-2015**

The graph shows the percentile of life expectancy of life expectancy with respect to our chosen 8 countries as you can view in the x axis,and we also filtered the year within the range of 2007-2015. Before achieving the graph, we created a quick table calculation for sum of life expectancy which in term calculates the percentile value.

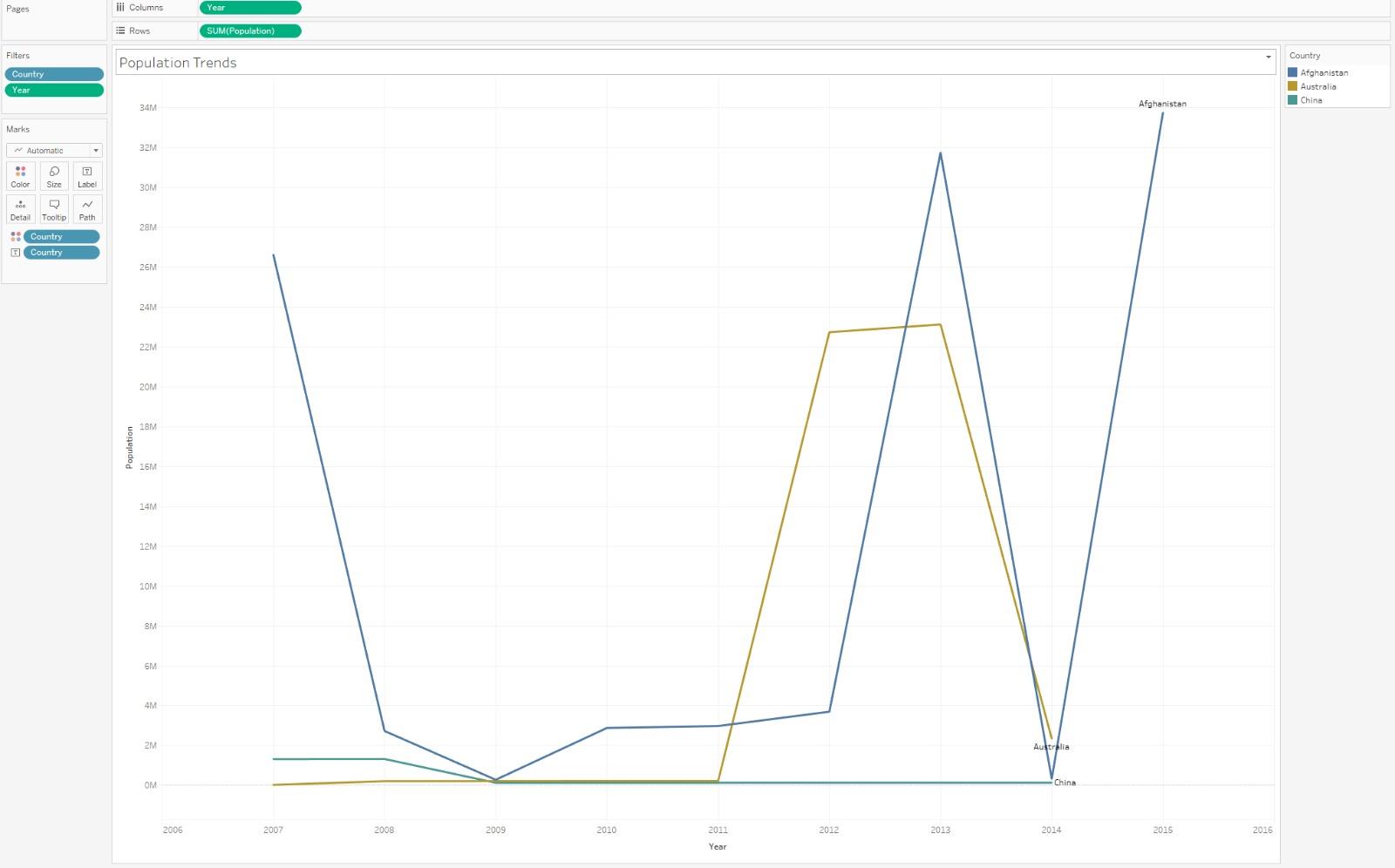
****

**13 countries vs other relevant measures**

This table lists the measures - alcohol, BMI, Diphtheria, infant deaths and measles, of our 13 chosen countries to represent.

****

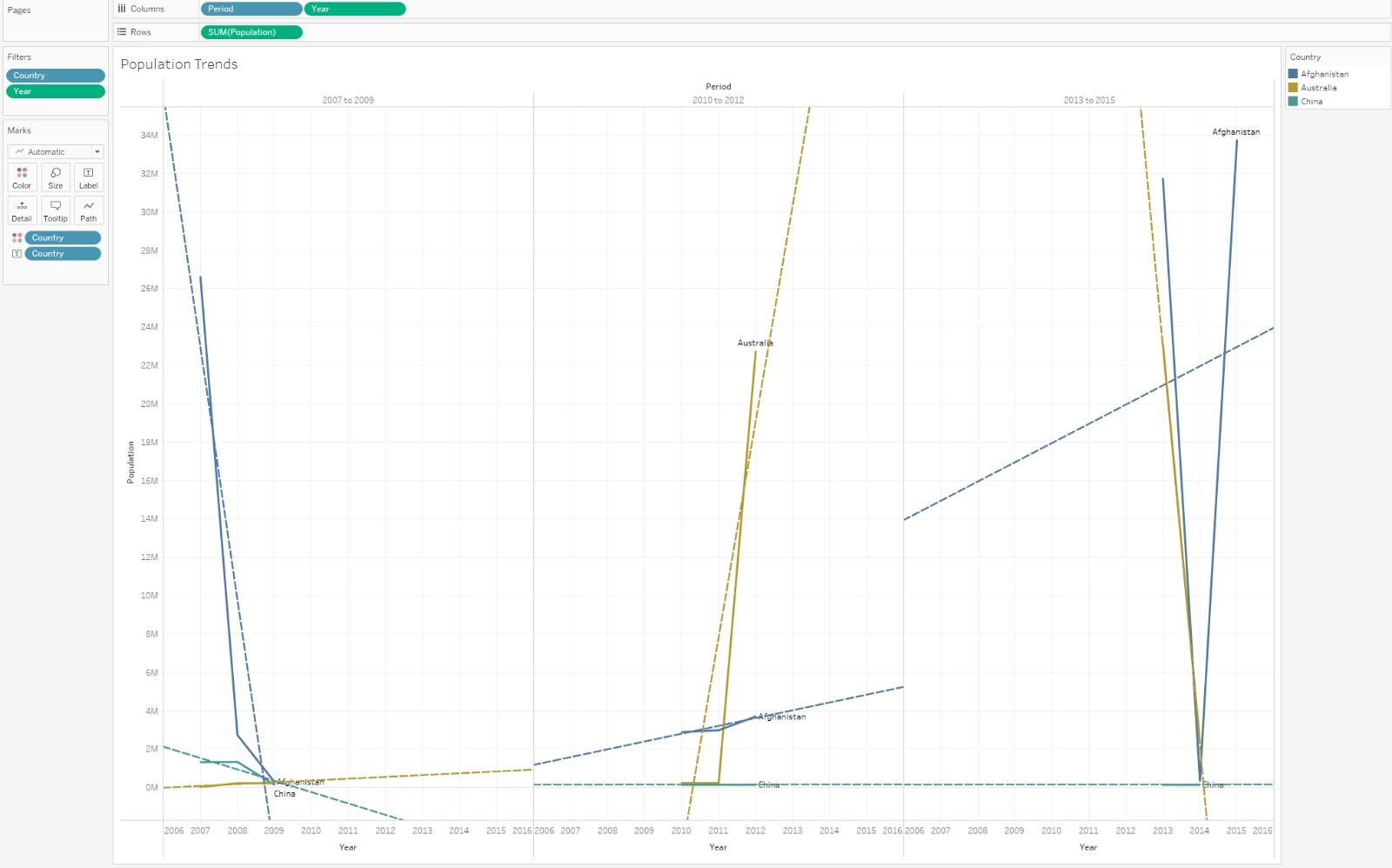
We formatted the Measure values into two decimal places.

****

**Population trend of 3 countries from the year 2007-2015**

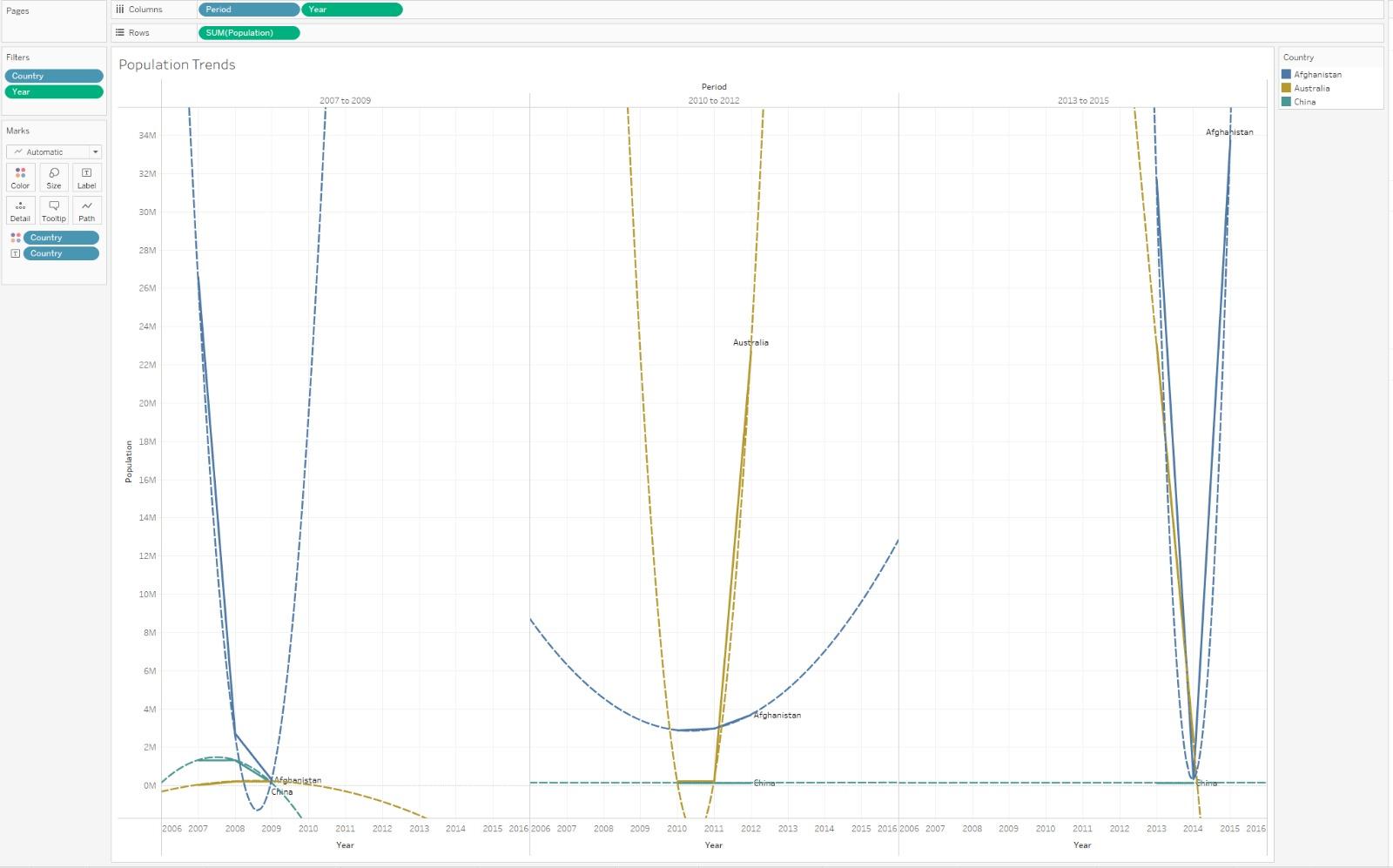
This line chart shows the population trends of our 3 chosen countries (Australia, Afghanistan, and China) in the year 2007-2015.





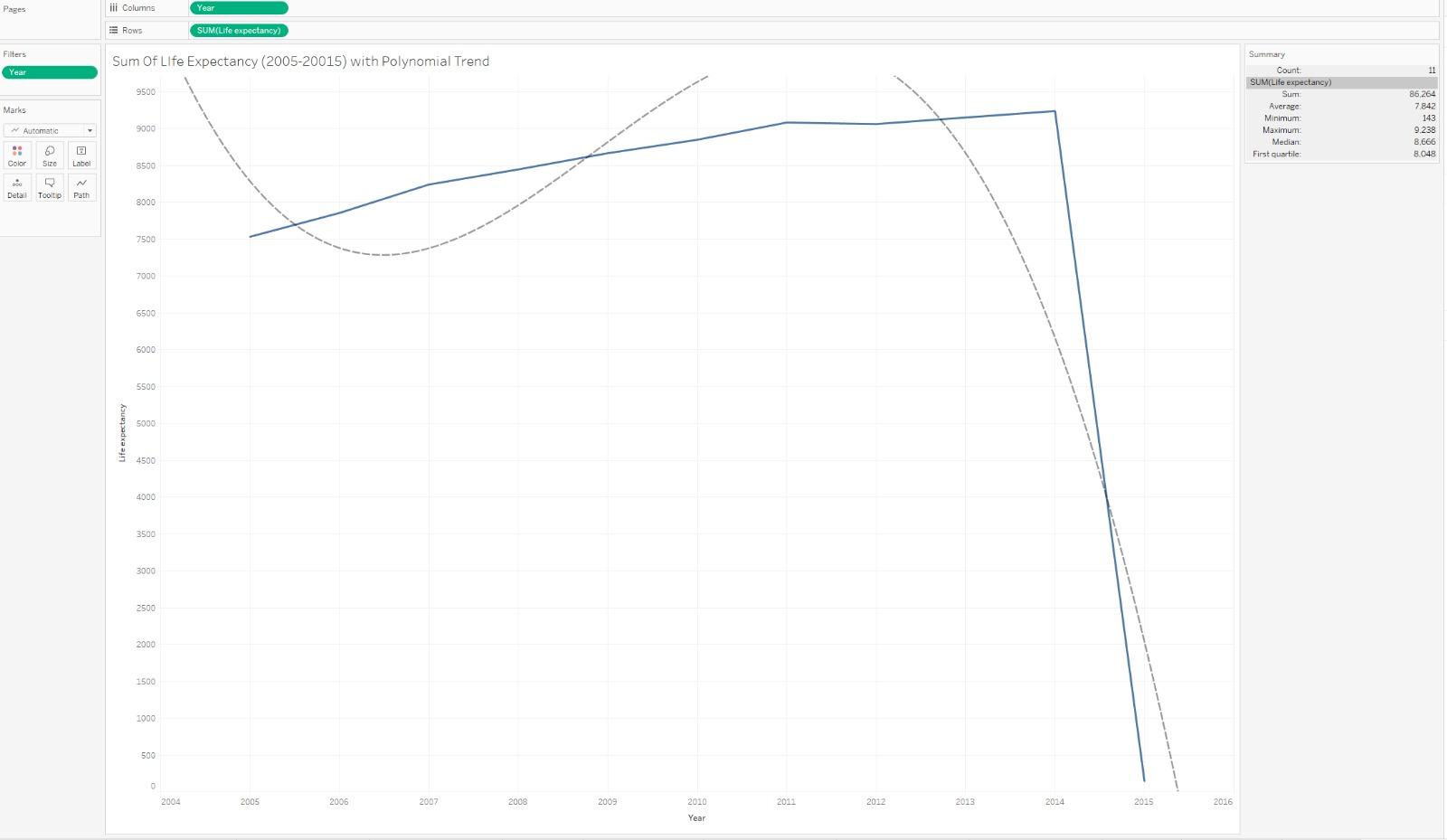
**Population trends through the years**

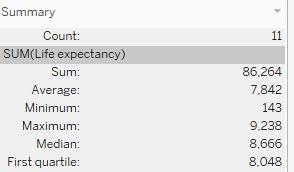
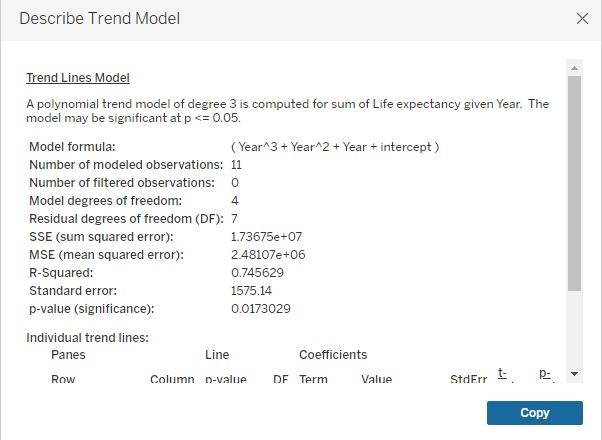
This graph explains the population trends for the same 3 chosen countries in different years (2007-2009, 2010-2012, 2013-2015).



**Population trends through the years with polynomial trend**

This graph explains the population trends with the *polynomial curve* for the same 3 chosen countries in different years (2007-2009, 2010-2012, 2013-2015).



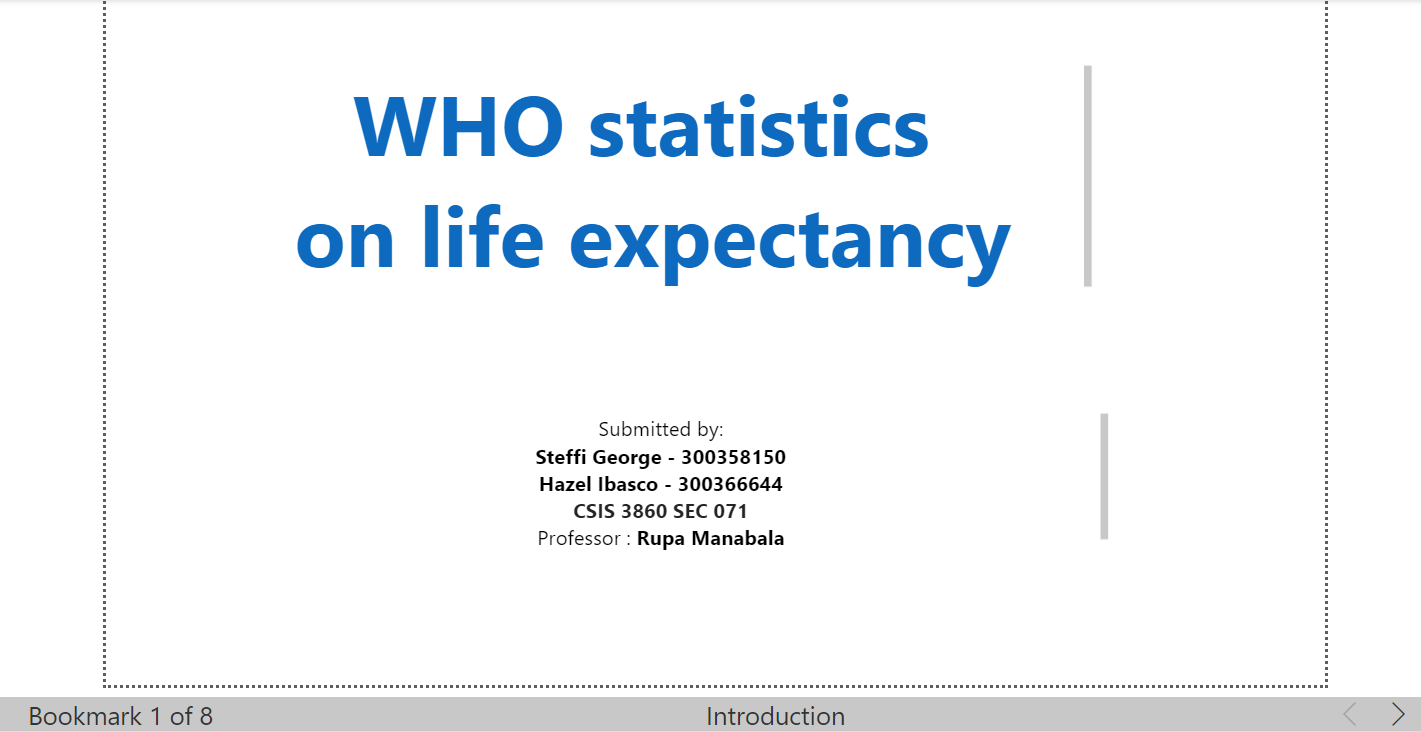


**Life Expectancy in 2007-2015 with Polynomial trend**

The line graph above shows the polynomial trend across all countries in the year 2007-2015. The trend model shows the information about the polynomial trend model. Summary is provided as well, with the sum, average, minimum, maximum, median and first quartile.

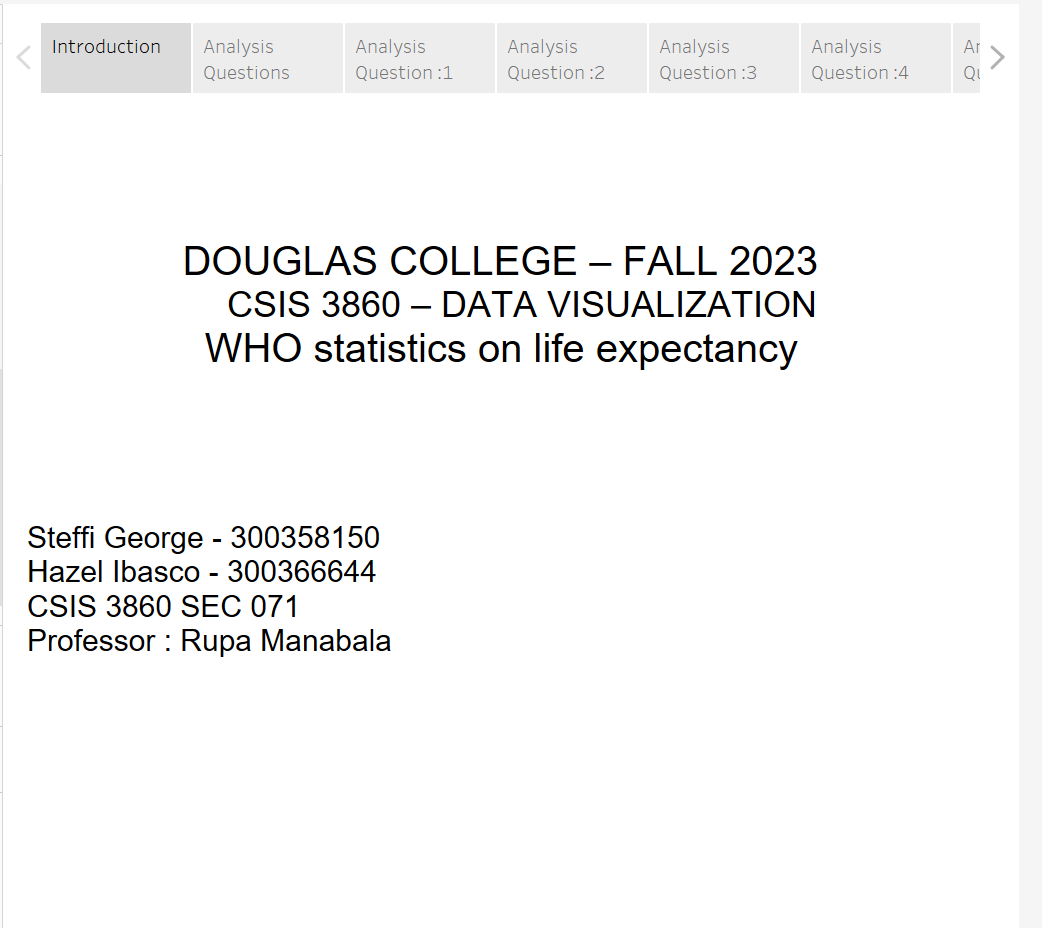
**Storytelling using PowerBI**

Created a powerBI story with 8 pages which will give a brief description of the life expectancy analysis.



**Storytelling using TABLEAU**

Created a tableau story with 10 pages which will give a brief description of the life expectancy analysis.



**Team Member Contribution**

PART B

Steffi George - Tableau

Hazel Ibasco - Documentation

**Reference**

* Life Expectancy (WHO) Fixed. (2023, March 30). Kaggle. <https://www.kaggle.com/datasets/lashagoch/life-expectancy-who-updated/discussion>
* *Life Expectancy (WHO)*. (2018, February 10). Kaggle. <https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who>
* *10 Open Datasets For Linear Regression*. (2021, January 1). TELUS International. <https://www.telusinternational.com/insights/ai-data/article/10-open-datasets-for-linear-regression>