WORLD HEALTH STATISTICS VISUALIZATION DASHBOARD

JUPS Health Visionaries

Authors: Umaima Khurshid Ahmad, Shruti Shah, Pinki Sharma, Jiali Ruan

Group Liaison Name : Umaima Khurshid Ahmad

Contents

Introduction of Data and Problem Statement

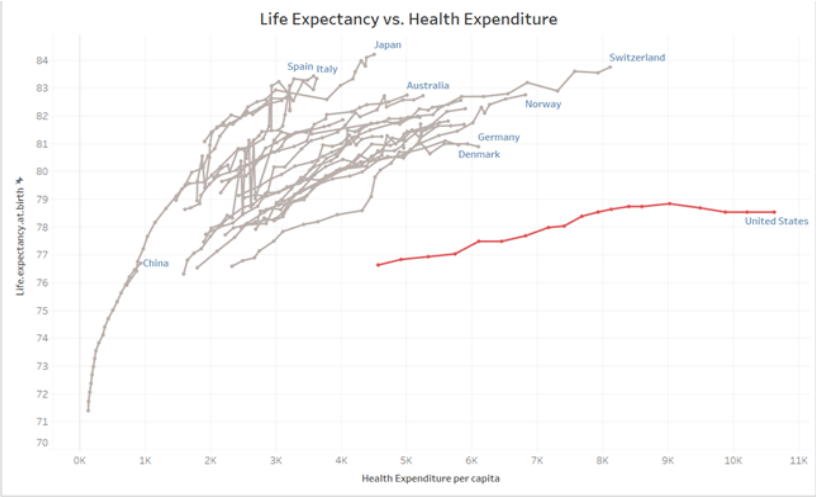
Life expectancy is a crucial variable that is mostly used in order to know how healthy a country or a region is. It is mostly used for assessing population health. We have used a dataset from the world bank to visualize life expectancy and its key-dependent variates. We have mainly focused on figuring out developing and developed countries' life expectancy. Our dataset's major setback was that we didn't have a separate column for developed and underdeveloped countries, so we used the internet to find the most developed and developing countries. The developing countries we are using are Angola, Ethiopia, Africa, and Bangladesh. While the developed countries we tend to use are the United States, Germany, India, and China. We also throw light upon the life expectancy across different regions around the world to know the better standing of how life expectancy has impacted across through other variables like health expenditure and mortality rate.

The dataset was taken from the World Health Bank and the dataset contains several important sub topics. Here, we are mainly focusing on life expectancy , mortality and health expenditures. In total we have 56981 rows and 17 columns over the years 1960 to 2020. Only 2 categorical variables which are country name and country code and the remaining all variables are numeric.

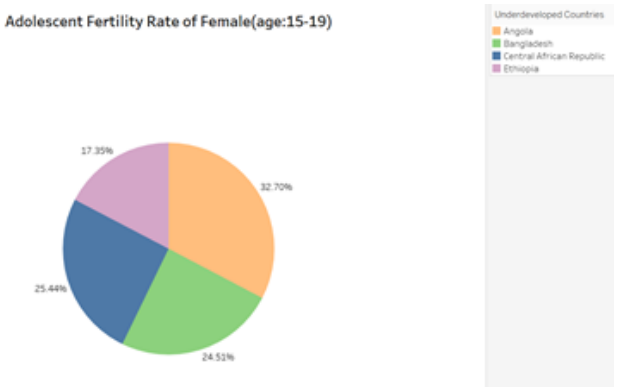
Exploratory Analysis

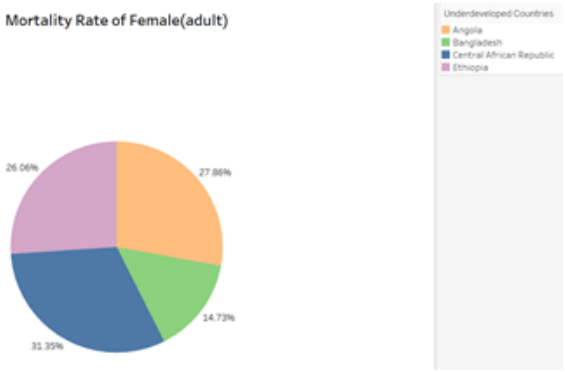
Our dataset mostly consists of numerical values, so we began with comparing few of our data to examine the patterns within our dataset. First, we created a geographical map to check the life expectancy around the world and to see which countries have high life expectancy and where it is very low.

After looking at the analysis, we choose to look at the relationship between health spending and life expectancy and where the US stands amongst other countries. We created a line graph ***Life Expectancy vs. Health Expenditure*** and used a filter on Tableau to display few of the rich countries. It turned out that the life expectancy in the US is lower than other rich countries.

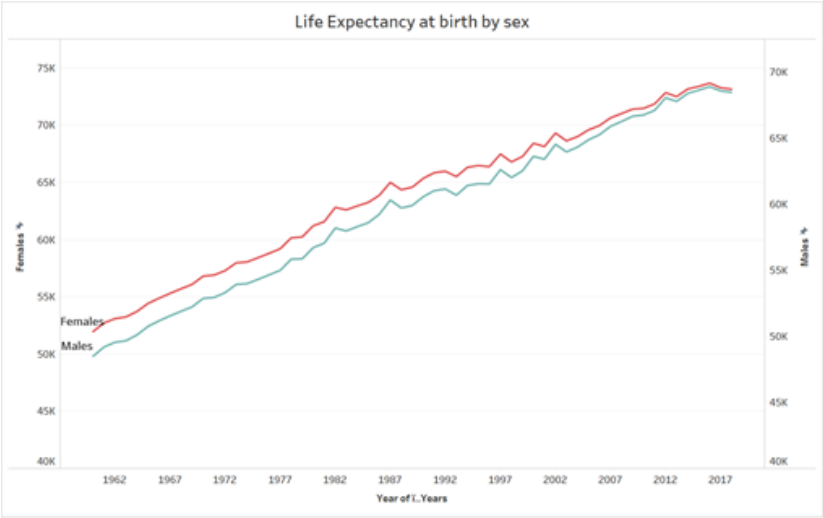


In order to analyze this result in more depth, we chose to look at some of the developed and underdeveloped counties to check how the life expectancy differs in those countries. Our dataset had columns for country names, but it didn't have specific columns indicating developed/underdeveloped countries, so we did some research and manually selected few of the countries in Tableau-filter based on our findings as developed and underdeveloped countries. We created a pie chart to visualize ***the adolescent fertility rate of females(age: 15-19)*** and compared it with a pie chart of ***mortality rate of female(adult)*** in underdeveloped countries.

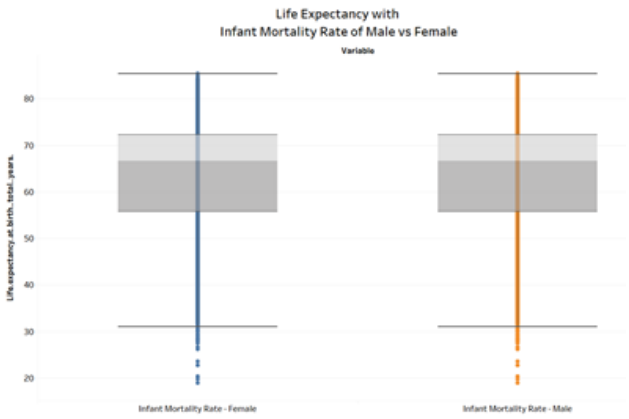


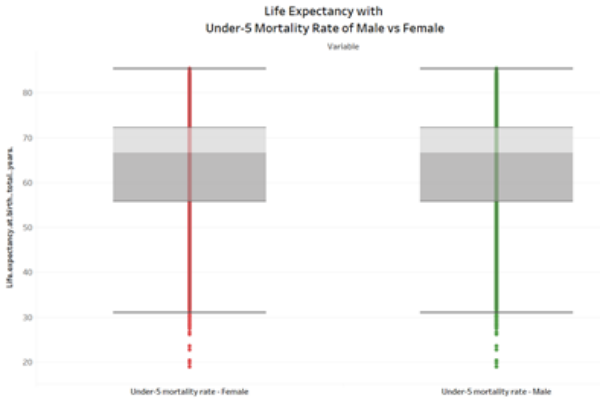


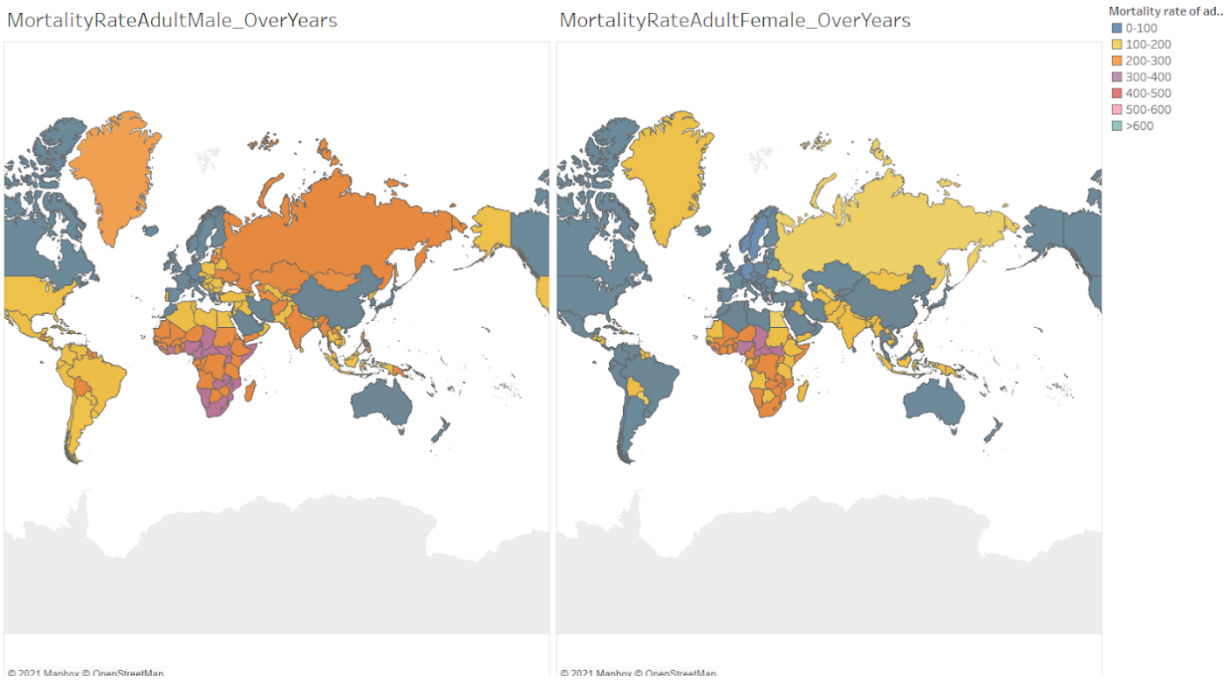
Since our datasets didn't have data for fertility rate of adult females or mortality rate of females from age15-19, it didn't feel right to compare rates of females from different age groups, so we decided not to include this exploratory analysis to our final visualization. However, we ended up making two-line graphs which visualizes how life expectancy changed over the years from 1960 to 2018 and how it is affected by the developed/rich and underdeveloped countries and counting it as part of our final visualization. One of our exploratory analyses includes a line graph comparing ***life expectancy at birth for male and female***.



If we just look at the graph, it seems like there isn't much difference between male and female's life expectancy, but if we look at the numbers it turns out that females live more than men. We choose to visualize this in our exploratory and final visualization using two geographic maps where one displays ***the mortality rate of adult male*** and the other displays ***the mortality rate of adult female***. Our initial analysis also includes a box plot showing the Infant and Under-5 mortality of male and female, but it didn't look appealing and the data didn't have much of a difference between infant and under-5 mortality.

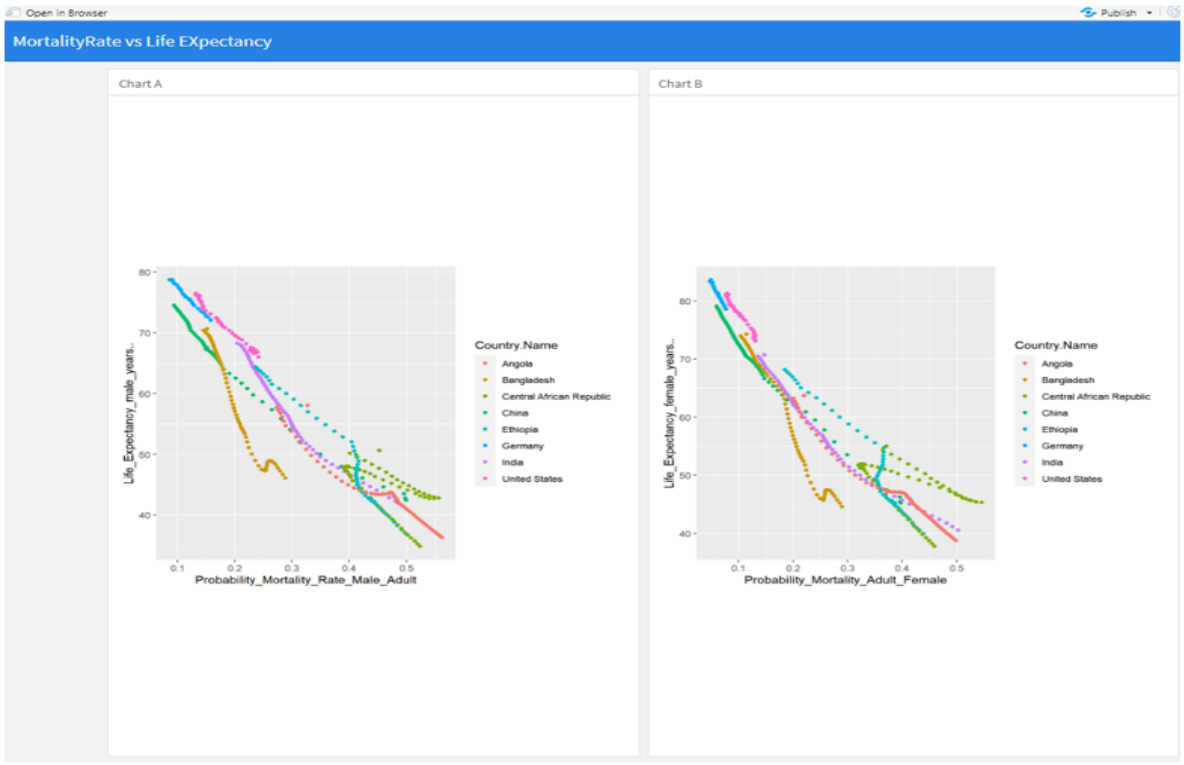






In the above graph, looking at the legend showing different colors for different ranges, it is clear that over the years the mortality rate of adult female is less compared to male.

Furthermore, to initialize the relation between life expectancy and mortality we created a scatterplot showing ***life expectancy vs mortality of adult male/female*** over the years for developed and under developed countries in flex board. In the graph below, we see there is an inverse relationship between life expectancy and mortality.



After going through all exploratory visual analysis, we were clear that we want to talk about life expectancy and how it has changed over the years from 1960 to 2018 and how developed and underdeveloped countries affect life expectancy

and how it links with health expenditure. Also, we decided to look at the mortality rates of male and female across the world and also find the relation between life expectancy and mortality. Our finale visualization includes :

1. Animated plot showing the relationship of life expectancy and health expenditure
2. Heatmap of life expectancy and health expenditure around the world for the 20th century
3. Heatmap of life expectancy of developed and underdeveloped countries from 1960 – 2018
4. Shiny app that shows total life expectancy depending on what year user chooses between developed and underdeveloped countries
5. Tableau Dashboard to visualize how life expectancy has changed over the years through line graphs, treemaps and geographical maps that shows life expectancy in 1900s vs 2000s
6. Animated geographical maps to show the mortality rate of male vs female around the world.
7. Shiny app that shows probability of mortality versus life expectancy for adult male and female over the period of years.

Visualizations:

Animated Plot: Figure 1 (Animation won't work in PDF)

As current health expenditure per capita increases for each country, the life expectancy also increases in a positive relationship in the plot. We used color, shape, size, motion and area for our graph mapping, and the color bubbles are countries with their populations. We learned the animation plot from the class by using the GDP and life expectancy, and people also used animation plots for health expenditure per capita and life expectancy, so we decided to follow the path. Since the health expenditure on our dataset is for the current international $, so we should be fine for not adjusting to individual country dollar $ and inflation. We also search gapminder dataset at R, and it doesn't have the Health Expenditure data in it, so we can't join to that dataset, and it's too late to search outside data and create the graph again. The legend can't display with the plot together, and it's also very clutter to display over 100 countries in one legend, and we think we just want to show the positive relationship or global trending between the health spending and life expectancy globally, and we don't want to know the specific country value in the plot.

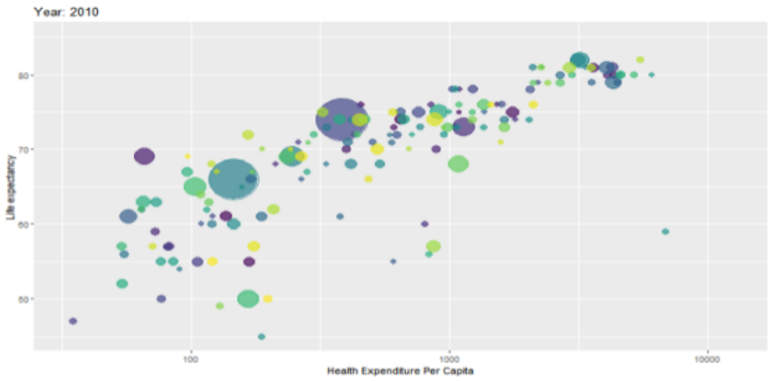
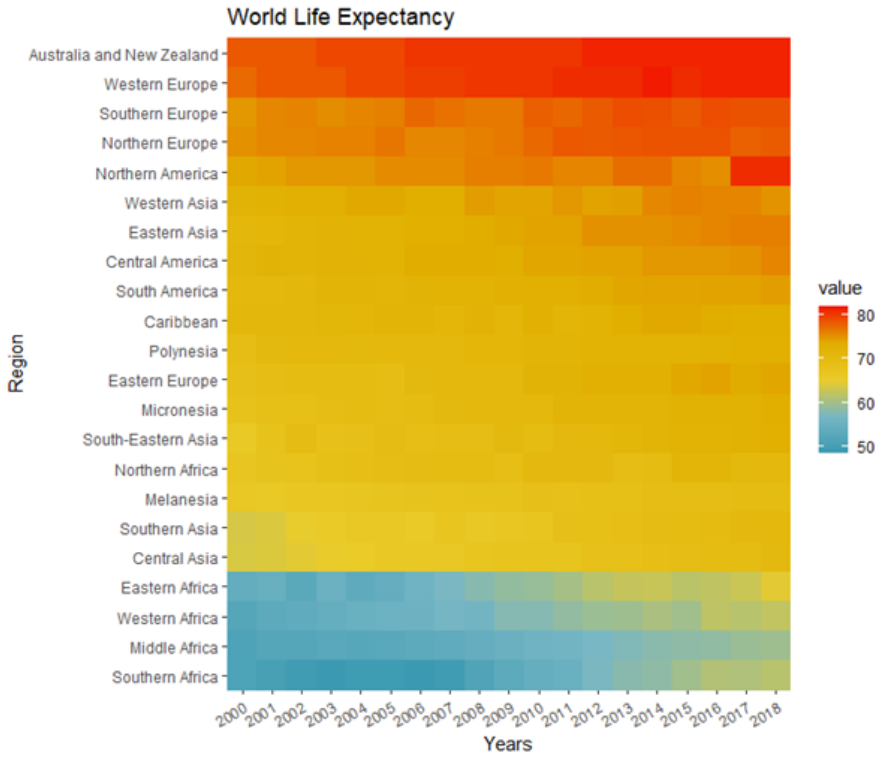


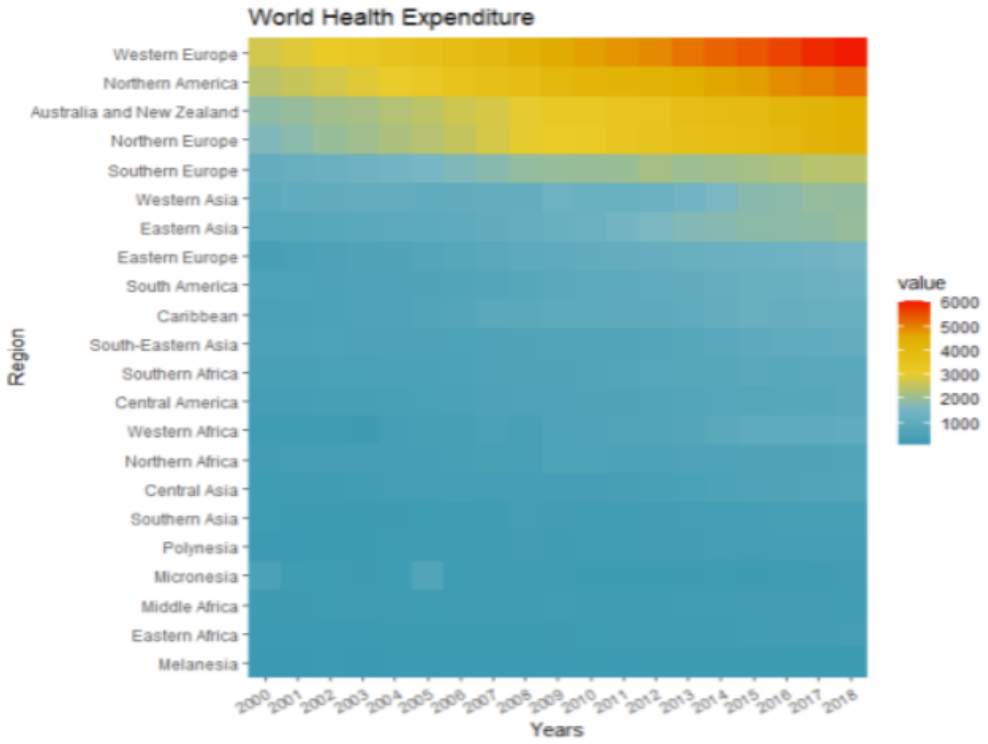
Figure 1 : Life Expectancy in relation to Health Expenditure

Heatmaps: Figure 2 and Figure 3

We plan to further look at the life expectancy and health expenditure around the world through plot heatmaps between these two variables alongside years. We mainly used color, shape and area in our mapping, and two contrast colors of

blue and red are used to represent two ends of values for life expectancy and health spending, and the yellow color works nicely to represent the middle value. We would like to see the global impact, so we chose regions instead of countries to be able to display in the heatmaps. We chose red color for life expectancy greater than 80 because there were only a few regions that can reach that point, and we want those regions to stand out. The heatmaps don't look good with just a few continents and also can’t dig into the problem regions easily where people may want to perform research or provide interventions for. Furthermore, South America is very different from North America with life expectancy and health spending, and we should not combine them in the map. We have provided sorting in the heatmap to group the higher and lower values together





Umaima

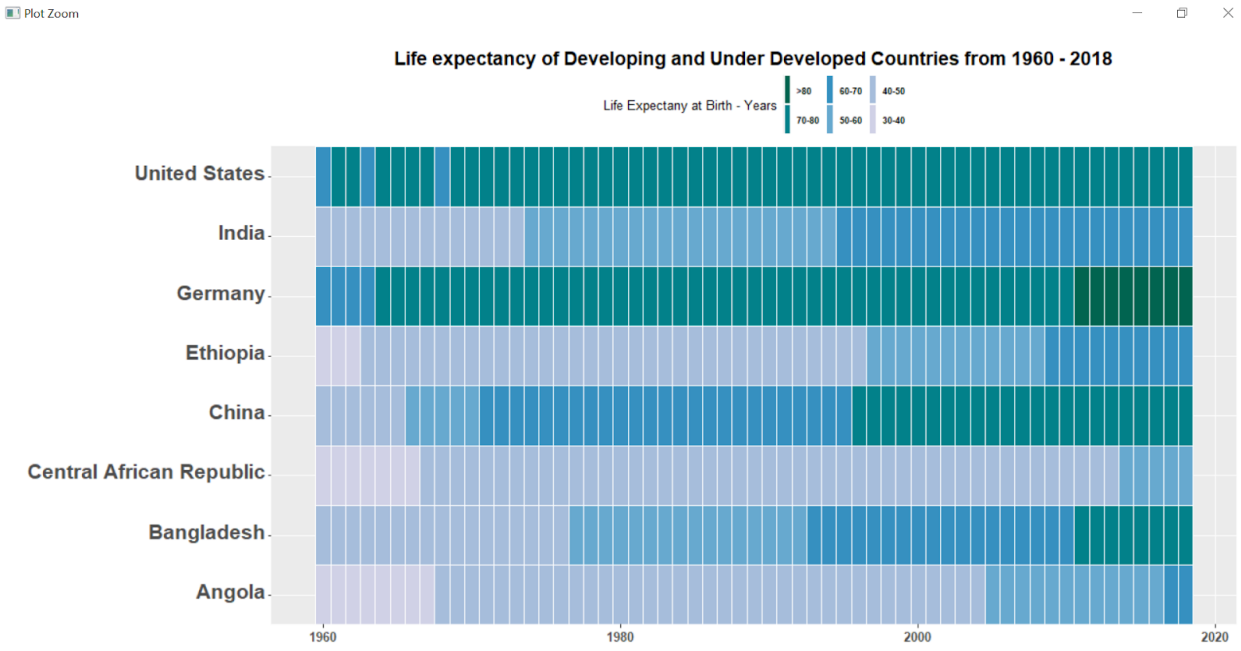


Figure x.1 shows a heatmap for developed and underdeveloped countries.

Figure x.1 shows a heatmap to glance at the life expectancy at birth of the developed and underdeveloped countries. The life expectancy variable is continuous; hence, binning was done to create segmentation (as parts) of the variable to fit the heatmap better. The heat map was chosen to visually see the trend change of life expectancy years amongst all the countries listed. This visualization also identifies the least and the most developed countries by just looking at the color-changing patterns. This color scheme from light to dark was chosen so that it's easier for the user to know even without actually seeing the legend values that lighter colors show less life expectancy while darker shades show more life expectancy. We can see an increasing trend of life expectancy in the country named Angola. Even the shades are still

lighter compared to the developing countries like the United States and Germany. Bangladesh also shows an increasing trend in life expectancy. Overall, with the passage of time and facilities, we can see that the life expectancy is increasing gradually in every country. However, the developing countries still have a long way to reach 70-80 years of life expectancy. The legend box was moved under the heading while also adjusting the size and type-face of country names and year to be visible to the audience.

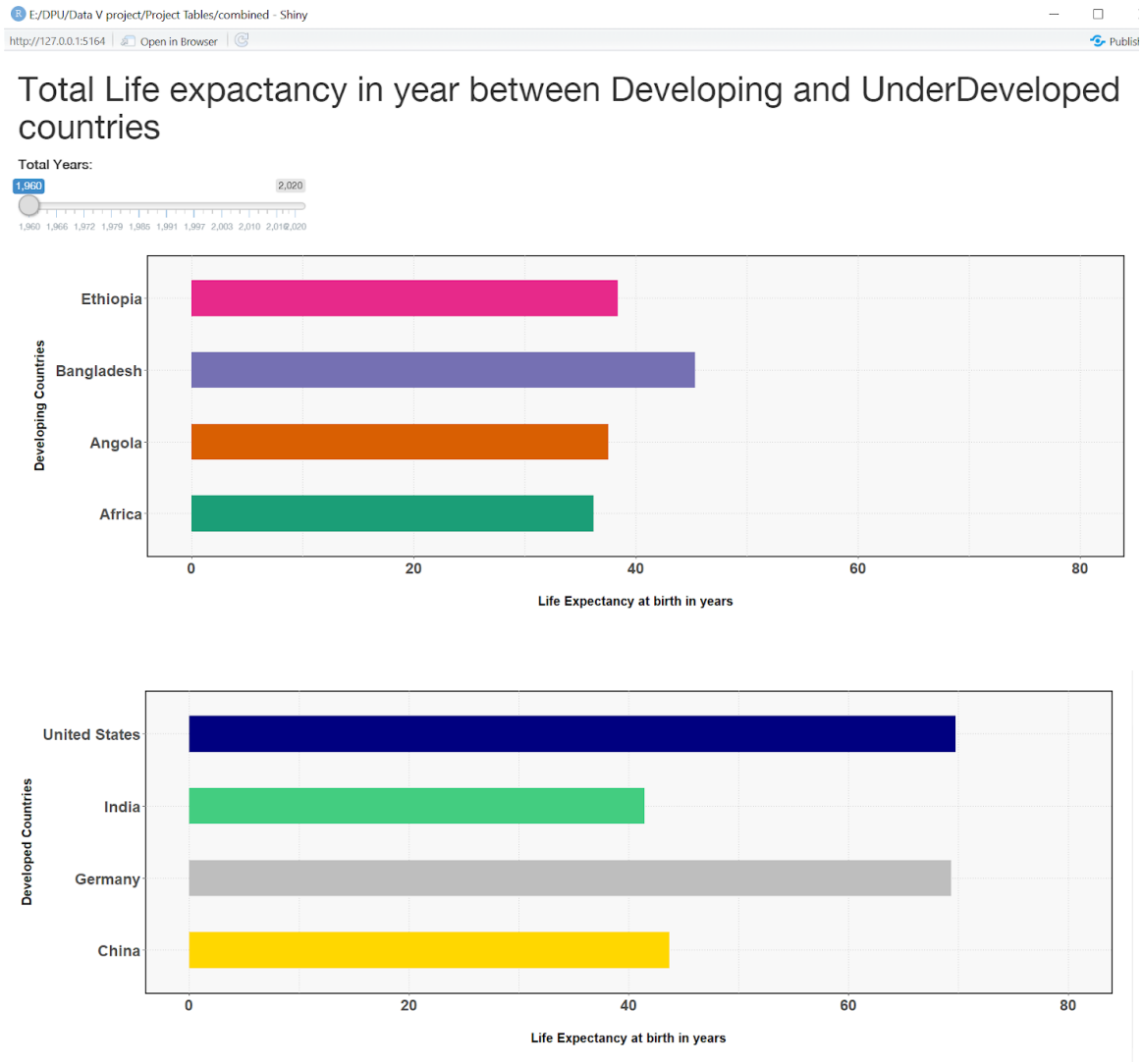
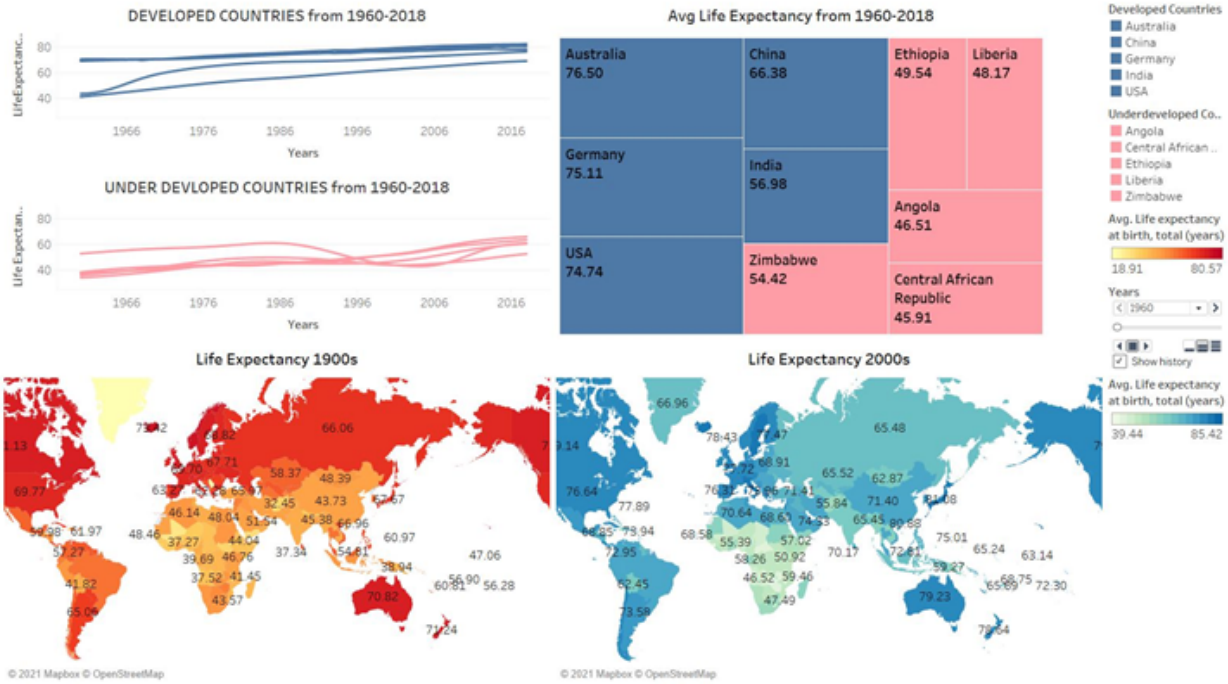


Figure x.2 shows a shiny application of life expectancy of specific years

The visualization is a shiny application comparing the developed and underdeveloped countries' life expectancy at birth. The comparison is through a bar graph, and the year is selected from the slider bar. After iterative drafting, the visualization year selection was changed into a sidebar rather than a drop-down menu. This was done because there are a lot of years, and from a user experience point of view, it should be easier for the user to see all the bars at a glance. Also, the major change is between the comparison of the 19th century with the 20th century due to which we felt a need to put a slider bar that is easily draggable as compared to a long list choosing a value for the 20th century that is at the end on the long list. Other than that, as we have learned in class about proximity, this visualization shows a direct comparison between the similar boxes(developed bar graph and developing countries bar graph). The default limit axis for the life expectancy for years of the bar developing countries box was initially up to 40, and the developed countries were till 60 years. If we are comparing two visualizations, their axis should be the same. Hence both axes were transformed to be at the same limit of 80 years. The first box's width was different because of the length of one of the country names, "Central African Republic''; hence it was changed to "Africa". The overall theme of the bar graph was changed; the background and grid lines were minimized. Our main focus is to deeply analyze the life expectancy of

countries and how it changed over years. Hence, the audience can easily choose a specific year and analyze the rate of change in life expectancy between the countries.



The above visualization shows an interactive dashboard to show how life expectancy has changed over the years from 1960-2018 worldwide and how developed and underdeveloped countries impact life expectancy. The variables we have used are life expectancy at birth, years and country names, since we didn't have specific columns indicating developed and underdeveloped countries, we got some reference from secure websites and from our geographical maps and based on that we selected few countries as developed and underdeveloped using Filter in Tableau. We decided to go with line graphs because we wanted to see the relationship between life expectancy and the countries. As we can see, life expectancy is low in underdeveloped countries in comparison to developed countries. The treemap shows the average life expectancy in those countries and if we look at the number, we can tell how much difference there is in life expectancy between developed and underdeveloped countries. Before, we had picked different colors for each county and sequential palette for the treemap, but we changed the colors after getting feedback on our presentation. We went with blue color for developed countries and light pink for the underdeveloped countries and made the axis similar on both the line graphs so that it's easier to convey that underdeveloped countries have lower life expectancy. Since we were using filters in the dashboard, the colors instinctively changed in the treemap. The life expectancy was mapped to size so if we look at the treemap, we can see that the box size is big for developed countries and small for underdeveloped countries, also the colors on treemap is making it obvious that we don't need to look at the legends. In order to glance over how life expectancy has changed over the years, we chose to go with two geographical maps where one map shows life expectancy from 1960 to 1999 and the other maps shows life expectancy from 2000 to 2018. For the two maps, we created two custom sequential palettes to display the highs and the lows of the life expectancy around the world. If we just look at the maps, it doesn't seem like there isn't much difference in life expectancy, but if we look at the numbers we can say that life expectancy has increased in underdeveloped countries. Overall, the interactive dashboard visualizes life expectancy precisely through these graphs.

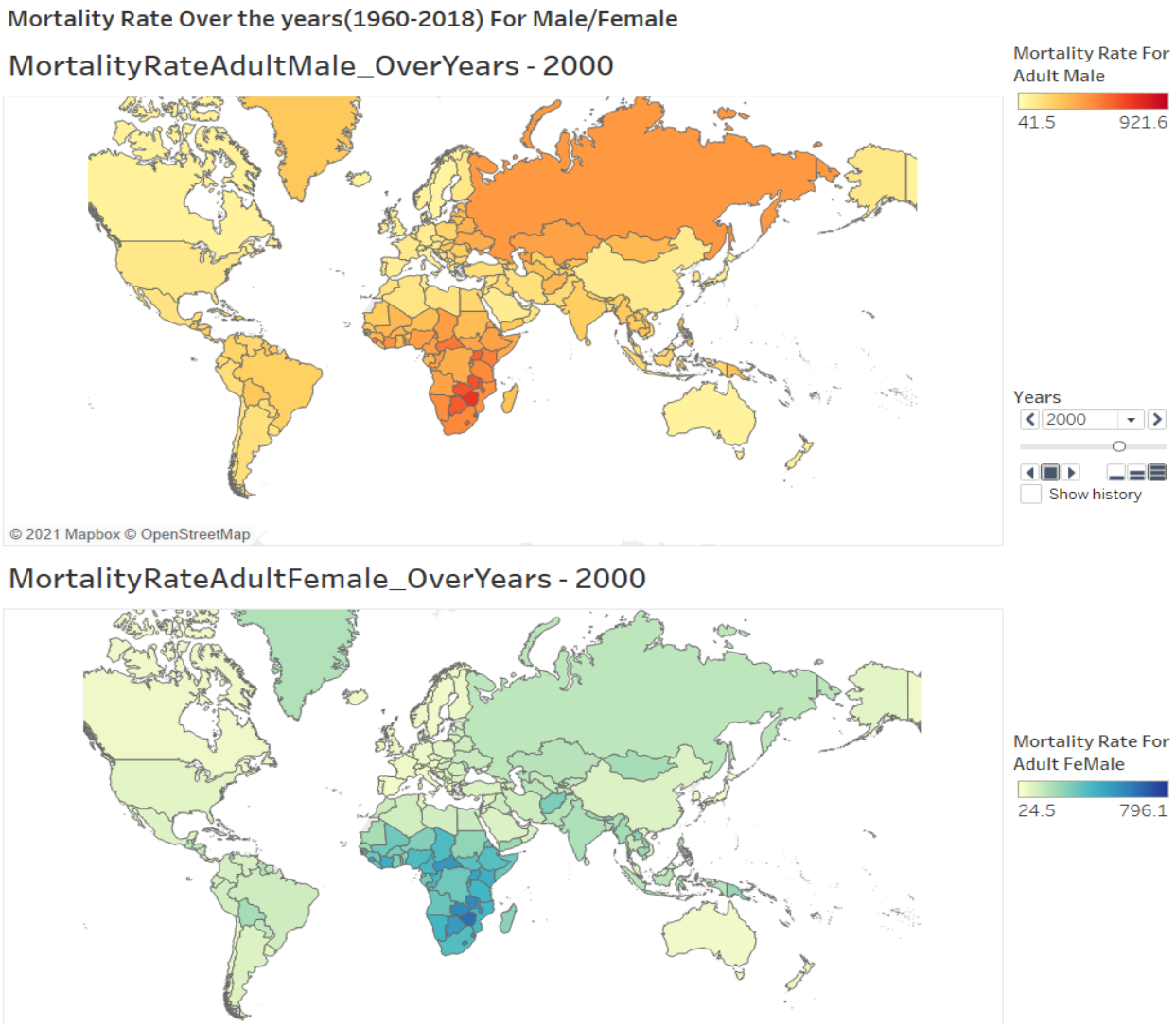


Figure 1a: map showing mortality rate of adult male and female of each country of specific year

The above figure 1a is visualizations comparing geographical maps between the mortality rate of adult Male and Female over the years from 1960 to 2018. All the variables that we used to create the map after cleaning the dataset were years, mortality rate of adult female and male and country name. We used a colorbrewer app to create our own color palette to show different saturation of a hue in a continuous fashion with light color depicting less mortality and darker the color, higher the mortality. Colorbrewer app gave more flexibility in choosing color vs the default color Tableau provides. Thinking about the concept of data, audience and message, we created a dashboard showing animation of mortality rates of adult Male and Female over different years of the world in tableau. Here, we used male and female mortality rate comparison map, in order to see whether over the time mortality rate is increasing or decreasing for each developed and underdeveloped country.

When comparing between two countries, the United States of America(developed) and Angola; we found the mortality rate of adult male decreased in the 19th century; however, for female groups there was no substantial decrease. In the 2000s century we see the mortality rate of adult females decreased for both developed and underdeveloped countries in comparison to Males.

In further, to understand the relation between life expectancy and mortality in developed and under developed countries we created a scatterplot using a shiny app.

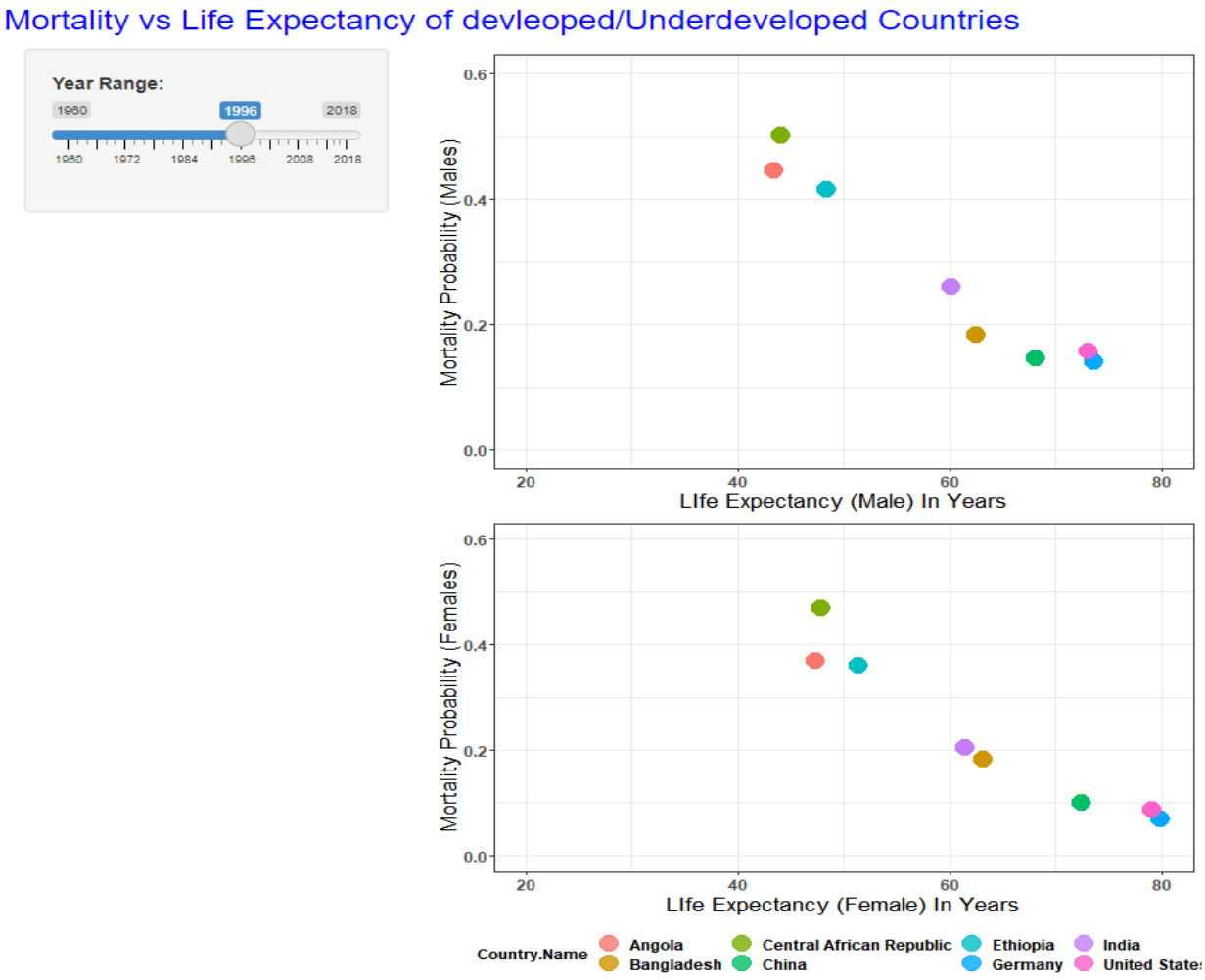


figure 1b: Scatter plot showing life expectancy vs mortality for both developed and underdeveloped country of specific year controlled through slider.

The above figure 1b is the visualization created using a shiny application. We are trying to show the relationship between probability of mortality versus life expectancy of developed and underdeveloped countries for both Male and Female over the years(1960-2018). The comparison is through scatter plots with same color coding for countries for both Male and Female and main focus being to find out the correlation between mortality and Life expectancy for both Male and Female over the years for some countries. We have taken 4 developed countries (USA ,Germany, China and India) and 4 underdeveloped countries (Bangladesh, Angola, Central African Republic and Ethiopia) using the filter option and removed all the null values in order to clean our dataset. After iteratively drafting the visualization, we came up with a slider for years in a shiny app. We have made sure the axes are the same scale for both graphs to get proximity as changing the years was changing the scale of the axes so had to fix the scale. The background theme and axis labels and size are modified for easy readability. The size of scatter plots are also changed and if you see legend position and key size is also changed. legend is only shown for one graph vs both the graph and positioned at bottom horizontally to fit the graph. The values in legend are made bold and bigger in size. Year slider changes the plots each time for a different year and gives flexibility for users to better understand the relationship over the years.

We see there is an inverse relationship between Life expectancy and Mortality and over the years Female had better Life expectancy than Male.Both developed and underdeveloped countries over the years have a mortality rate decreasing which is a good sign.

Analysis and Discussion

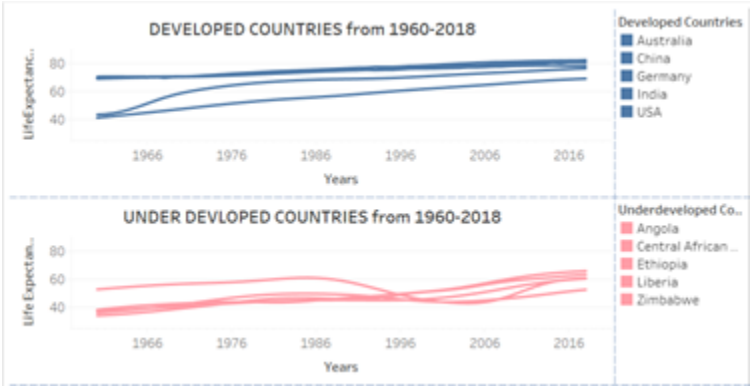
We have used heatmaps, animation plots/graphs and dashboards in R programming, Tableau and Shiny. We have the following interesting findings and conclusion from our data and visualization: Over the years, life expectancy has gradually increased for almost every country around the world, However if we see the under developed countries like Angola and Central African Republic, the life expectancy is still low in the 19th century and in the mid of the 20th century but its still increasing but on a very low rate. If we compare it with the developed countries, the life expectancy increase is still low. By looking at the visualizations, we can also come to a conclusion that overall expectancy of females is more as compared to the life expectancy of males. We also identified several factors that drive improvements in life expectancy, and they are as current health expenditure per capita increases for each country, the life expectancy also increases in a positive relationship. The continuing increase of life expectancy is also due to decline in mortality rate. Since 2000, the mortality rate of adult females decreased for both developed and under developed countries. Even in the developed countries like the USA, we have examined that there are differences in life expectancy between the male and female.We also analyzed through draft visualizations that the educational variables were not having any correlations with the life expectancy variables. We can still further examine through a larger dataset and different feature set for the dependency of life expectancy.

Appendix

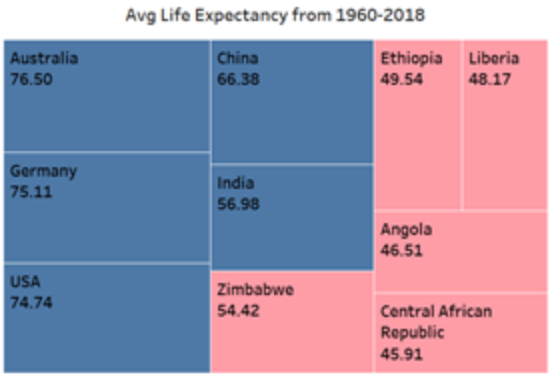
Note: All the codes are attached with the submission folder

1. Individual Report of Shruti Shah

We were working with different dataset at the beginning in which I did data cleaning, but the dataset didn't have accurate data, so we decided to go with this dataset. I created a few exploratory visualizations mentioned earlier in the report such as the pie charts for the fertility and mortality rates for females, geographical map to get basic idea of life expectancy around the world, line graph showing life expectancy at birth for male vs female. I have stayed connected with the teammates throughout and participated in the weekly group meetings.
As a team, we planned that it would be a good idea to highlight how life expectancy has changed over the years. I was in charge of working with developed and underdeveloped variables and to see how it impacts life expectancy. At first, I created a simple map to look at life expectancy across the world. Our datasets didn't have specific columns for developed and underdeveloped countries, so I looked at the map to check in which countries the life expectancy is very low and very high. I also did some research(reference link is at the end) and based on my findings, I used filters in Tableau and selected few countries as developed and underdeveloped countries. I created two-line graphs : one visualizing life expectancy at birth over the years in developed countries and other visualizing life expectancy in underdeveloped countries.

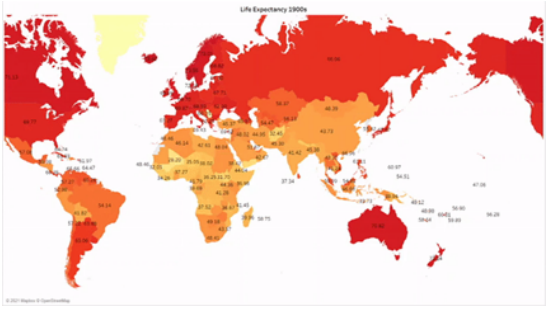


The two graphs shows that life expectancy is low in underdeveloped countries and high in developed countries. To look at the analysis in more detail, I created one treemap which displayed the average life expectancy in both developed and underdeveloped countries. The data in the treemap is divided based on the size of average life expectancy at birth and I used the same colors that I used in line graphs to make it easy to understand and appealing to the audience.

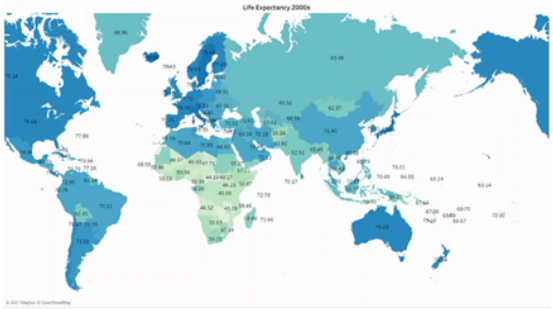


In the treemap, we can see that the developed countries are all grouped together in blue and the box sizes are big whereas the underdeveloped countries are all grouped together in light pink revealing underdeveloped countries. I also created two geographical maps visualizing how life expectancy has changed from the last two centuries. I decided to create two maps because when we look at the data, we found out that the life expectancy has increased over the years, especially in poor/underdeveloped countries. I wanted to use sequential palettes for the maps, but Tableau didn't have many options for it, so I created two custom sequential palettes. I used the ColorBrew website for my reference. Later, I thought it would be interesting to use animation on these maps to show how life expectancy changed each year, so I did animation on both maps in Tableau.

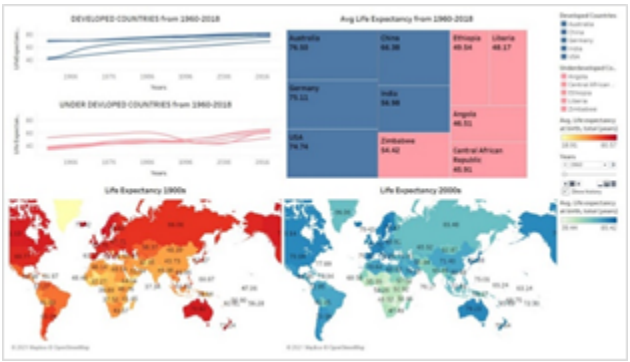
Life Expectancy 1960-1999



Life Expectancy 2000-2018



When we went over all our final visualization, we felt like there would be too many animated visuals since two of my team members also created animation and two members created a shiny app. We planned to put my visuals in Tableau Dashboard and count it all as one visualization since the variables I was working with were developed/underdeveloped countries, and life expectancy. Also, the variables are linked to each other in a way that it has a story to tell. Overall, the final visualization that I worked on was this:

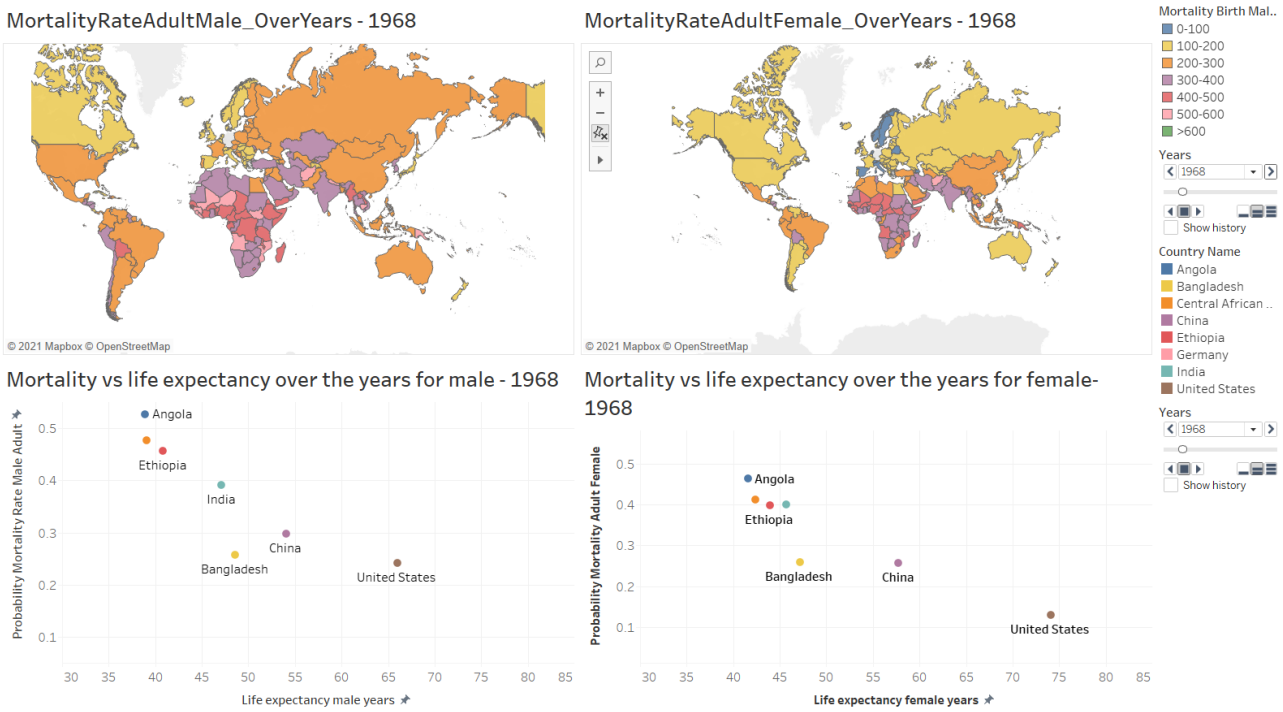


Our project has taught me how the life expectancy links with all the aspects like country's health expenditure, fertility rates, mortality rates, country's status like rich/developed countries will have more life expectancy in comparison to poor/underdeveloped countries. This course has taught me how to do data cleaning and how even a simple exploratory visualization can give us the broad ideas and as we make more visuals/drafts and play with different techniques, we'll know which is the correct way to visualize that data. I had never worked with R before taking this course, so it was a good learning experience.

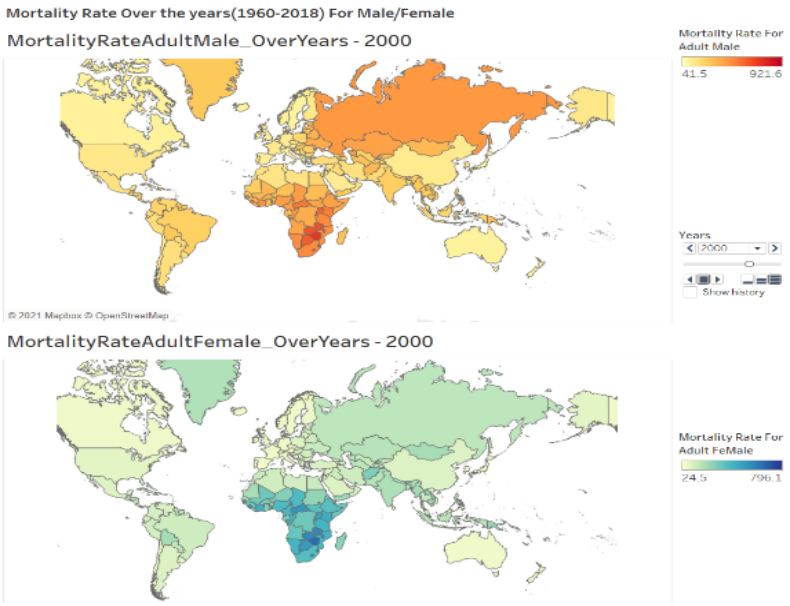
2. Individual report of Pinki Sharma:

I always connected with my teammates on what we should do and where we can improve by regularly participating in the meetings and sharing constructive ideas. As a team, we planned to drill down Life expectancy and Mortality over the years in different parts of the world. I chose to see how mortality rate changed over the years in different parts of the world, especially focusing on Developed and Underdeveloped countries. Our initial dataset did not have developed and underdeveloped columns so we filtered countries after taking reference from the website.. Also, thought of linking Life expectancy and Mortality Rate to give an overall gist of our project. In our dataset we had different features related to mortality rate like infant and Adult male and female mortality, but we did not have a total mortality rate to compare with total Life Expectancy. I took the mortality rate of adult male and females as a reference to compare with life expectancy.

At first, I created a simple world map as an exploratory analysis to see how the mortality rate of male and female was changing over the years in different countries. I categorized mortality rate variable by creating ranges (i.e. 0-100,100-200) to show better in the map with categorical distinct colors. To create an interactive dashboard in tableau I also added the relationship of Life expectancy vs Mortality as shown below.



As per the feedback I received broke up the above dashboard into individual comparison graphs. To give more visual appeal and user readability I created a dashboard in tableau showing visual animation of mortality rate of adult male and female over the years (1960-2018) in a world map. In the above map, I had Mortality Rate as categorical ranges(0-100,200-300) which I changed back to numerical sequential legend as shown below. I created my own sequential color palette using the color brewer app to give more details with color in my final visualization as seen below. Here, I used a dashboard, showing a comparison map of Male and Female mortality rate over the years for all countries. I also added the 'Years' variable as a page attribute for users to change years and compare the mortality. It also accounted for animation.

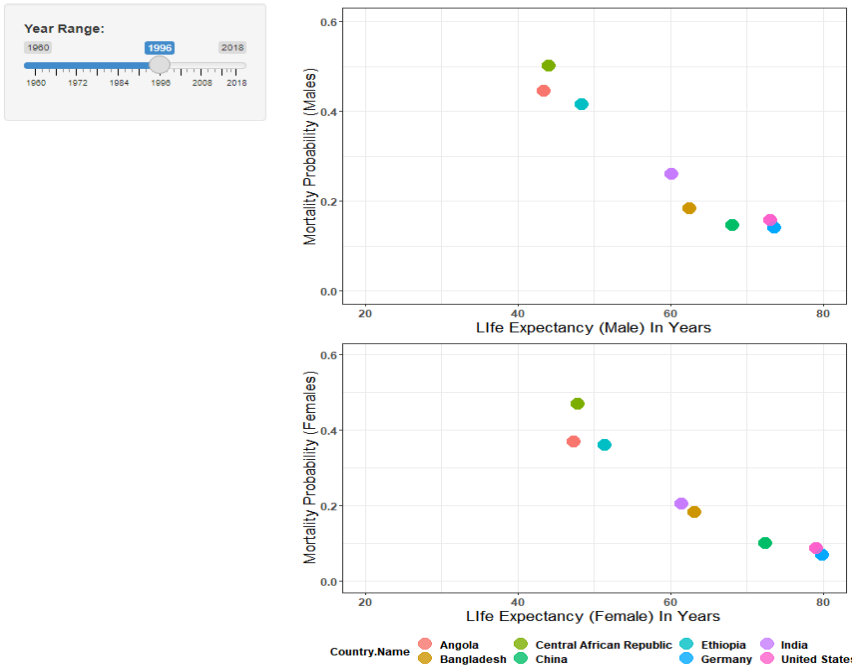


The animation above, shows the mortality rate of adult male being more compared to adult females. Even though the number of mortality rates of adult Females is high, over the period of years 1960-2018 it was decreasing compared to male. The lighter shade of color shows lesser values and darker shade of color shows higher mortality rate. For both Male and Female, Africa had higher mortality and USA, UK, Germany, China had lower mortalities. When comparing between two countries, the United States of America(developed) and Angola(underdeveloped); we found the mortality rate of adult male decreased in the 19th century; however, for female groups there was no substantial decrease. In the 2000s century we see the mortality rate of adult females decreased for both developed and underdeveloped countries in comparison to Males.

Furthermore, I wanted to show how mortality rates of adult male and females are affecting life expectancy in some of the developed and underdeveloped countries in different time periods. It was not easy to compare mortality rate directly with life expectancy, as we had total life expectancy in years and mortality rate in relative to 1000 people. So, to generalize mortality value, I created a new column for Male and Female Mortality by dividing existing mortality columns by 1000. The initial exploratory analysis of the same was done in Flex Board(refer in exploratory analysis part in the group report above). Then as an iterative process, to make an interactive dashboard with animation and world maps, I created the graphs in Tableau. Furthermore, to increase the visual appeal and have better control on the labels, axes, plotting and slider effect for years, I created my final visualization in a shiny app with 'Years' variable as a slider.

Eventually, compared it with life expectancy for male and females by creating a scatter plot in **Shiny App** to see whether there is any relationship between life expectancy and mortality rate for both male and female for developed and underdeveloped countries. In short, I used categorical variables (Country Name) and numerical variables (Life Expectancy for Male/Female, Mortality rate for Adult Male/Female and Years). I have made sure the axes are the same scale for both graphs to get proximity as changing the years was changing the scale of the axes in the below graph so I had to fix the scale. The background theme , axis labels and size are modified for easy readability. The size of scatter plots are also changed and if you see legend position and key size is also changed. legend is only shown for one graph vs both the graphs and positioned at bottom horizontally to fit the graph. The values in legend are made bold and bigger in size. Year slider changes the plots each time for a different year and gives flexibility for users to better understand the relationship of mortality and Life expectancy over the years.

Mortality vs Life Expectancy of devleoped/Underdeveloped Countries



Life expectancy over the years have changed drastically for females rather than men. The above visualization shows Mortality is strongly connected to Life expectancy; when Life expectancy increases, mortality decreases for both developed and underdeveloped countries but drastically for Female Adults than men. In both Male and Female cases, we see a negative trend as mortality decreases, Life expectancy increases. For instance, life Expectancy when compared across the 19th century and 2000s, we see Females have higher Life Expectancy and lower mortality compared to Male for both developed and developing countries. China, Germany, USA ,developed countries have better Life expectancy and lower mortality compared to Africa, Ethiopia and Angola.

In my opinion our visualization will help people to get a general idea that even the developed countries also have differences in life expectancy/mortality for male and females.

Throughout the project I learned deeply about shiny apps and tableaus and also learned how to organize the visualization to depict story by maintaining proper hygiene for visualization.I also learned proximity by maintaining the same axis for both the graphs when comparing with each other as in my case for Male and Female.

I was trying to make a dashboard with all three variables(Life expectancy,Mortality,Health Expenditures) over the years in different countries and show the relationship in one visualization. **If I had more time** would have done this as currently we have broken this into two relationships which are Life expectancy vs Health Expenditure and Mortality vs Life expectancy.

All my codes in R(Shiny App) and Tableau files are shared above with my initial(Pinki Sharma).

3. Individual report of Jenny

We have a great team in terms of organizing group activities, feedback and technical talents. I found the initial WHO dataset from Kaggle and introduced the dataset to the team. I have created the google drive to organize documents, helped with creating and organizing documents, created visualizations and provided help and feedback to the team on time. I always treat my teammates with respect.

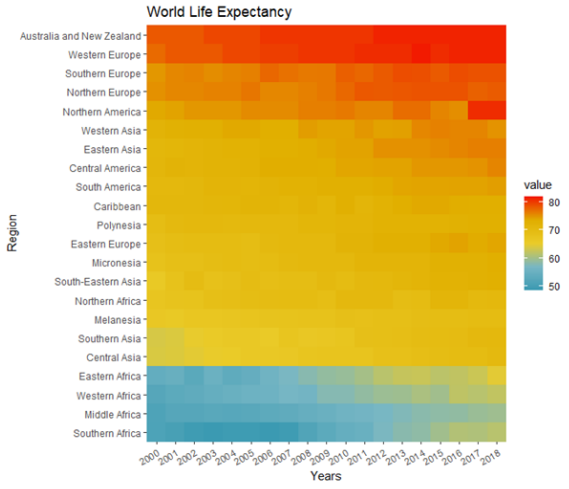
Data correlations and important variables that stand out as key variables that are important for drawing conclusion:

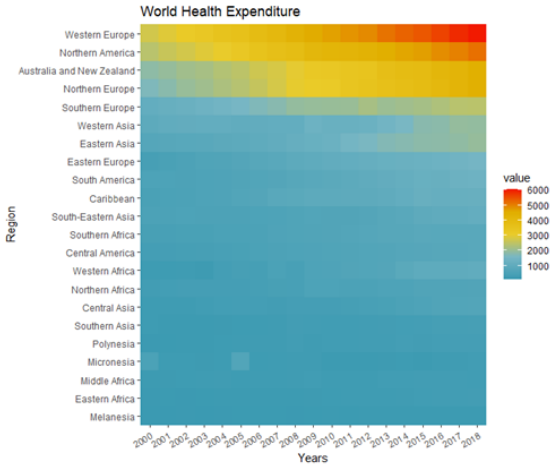
As I reviewed the WHO dataset, I chose two variables and tried to study the correlations between life expectancy and health expenditure since they are important factors in the world health statistic dashboard and help us to draw conclusions. Therefore, I mainly used the life expectancy and health expenditure for visualization within the dataset, and I have the below conclusions after I studied the data and created visualization for them :

- As current health expenditure per capita increases for each country, the life expectancy also increases in a positive relationship in the plot.
- Life expectancy is the key metric for assessing population health.
- The lowest life expectancy is the South and West African with a little over 50 years, but South African has had improvement with the life expectancy since 2009.
- Austrian, New Zealand and European always have longer life expectancy than other countries.
- Global spending on health continued to rise almost every year.
- African region spent less with health per capita in general.
- Austrian and Zealand spent a modern amount on health but with longer life. expectancy which perhaps due to the good environment factor and healthy lifestyles.
- North America has increased health spending graduality since 2012.
- Eastern Asia has also increased health spending graduality since 2011.

To create the rich, deep displays of the data that go beyond the simple line graph for life expectancy, I used the following design criteria for my visualization:

Time Series HeatMaps are great to display life expectancy and health expenditure from year to year as 2000 to 2018, and there are only 3 main colors, and they are highly contrasted with red and blue to represent the two end values for subtleties in color mapping, and the yellow color are for the middle value. In which way, people can identify the long/short life expectancy and high/low health spending easily from heatmaps. Furthermore, two heatmaps are shown side by side, so people realize that the health spending has a positive impact on life expectancy.
We would like to see the global impact, so we chose regions instead of countries to be able to display in the heatmaps. We chose red color for life expectancy greater than 80 because there were only a few regions that can reach that point, and we want those regions to stand out.





The design criteria that went into my visualizations:

I learned the animation plot from the class by using the GDP and life expectancy, and people also used animation plots for health expenditure and life expectancy, so I decided to follow the path. Since the health expenditure in our dataset is for the current international $, so we should be fine for not adjusting to individual country dollar $ and inflation. I also search gapminder dataset at R, and it doesn't have the Health Expenditure data in it, so I can't join to that dataset, and it's too late to search outside data and create the graph again. The legend can't display with the plot together, and it's also very clutter to display over 100 countries in one legend, and I think we just want to show the positive relationship or global trending between the health spending and life expectancy globally, and we don't want to know the specific country value in the plot.

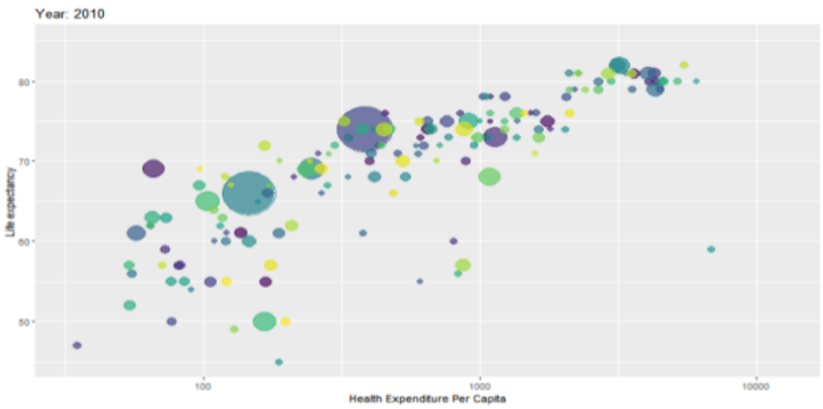


Figure 1 : Life Expectancy in relation to Health Expenditure

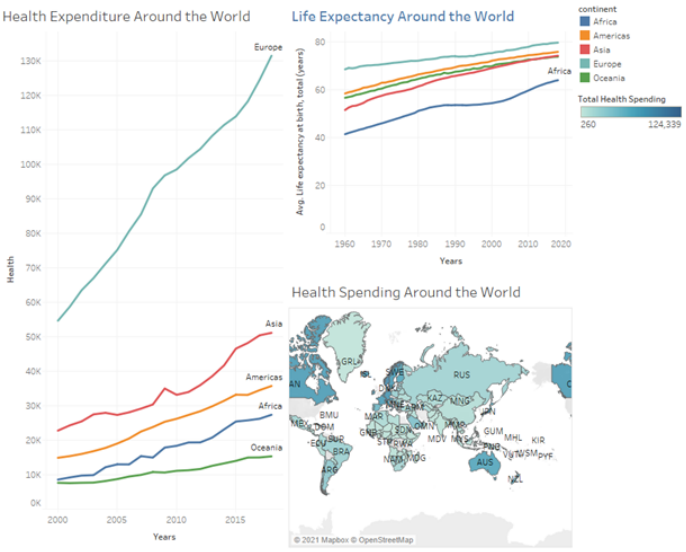
Yes. I used R codes for the animation plot and two heatmaps to create the life expectancy and health expenditure. I provided the R codes below:

Interactivity which helps move a visualization technique into a

deeper/richer domain:

I created an interactive dashboard in Tableau for a deeper and richer research for the relationship of life expectancy and health spending around the world as below:

When we hover at the health expenditure in the region, the other 2 sheets will show the life expectancy of the region and the countries with their health spending to related regions. The purpose is to show the positive relationship with health spending and life expectancy with each region and its associated countries.



Conclusions can be drawn about the data from my visualizations and how might they

help people analyze the data further:

The conclusions I have for my visualization are:

· Over the years, life expectancy has gradually increased for almost every country.

· Global spending on health continued to rise almost every year.

· The lowest life expectancy is the South and West African with a little over 50 years, but South African has had improvement with the life expectancy since 2009.

· Austrian, New Zealand and European always have longer life expectancy than other countries.

· Global spending on health continued to rise almost every year.

· African region spent less with health per capita in general.

· Austrian and Zealand spent a modern amount on health but with longer life. expectancy which perhaps due to the good environment factor and healthy lifestyles.

· North America has increased health spending graduality since 2012.

· Eastern Asia has also increased health spending graduality since 2011.

People will have a general idea for life expectancy and health spending around the world, and they can also perform research and intervention to the problem areas like the Africa region and help them to improve health spending and life expectancy.

If you had more time, where might you have liked to have developed your visualizations further?

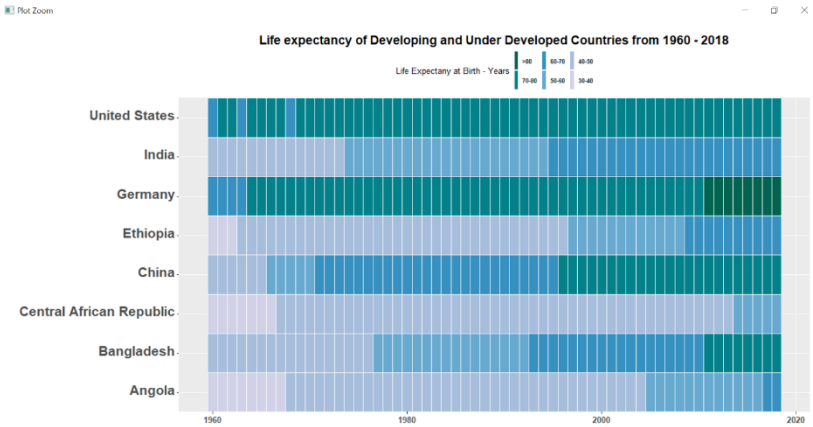
If I have more time with the project, I will try to explore more data such as specific communicable diseases and their impact on life expectancy and health around the world. I will also want to use R or dashboard to graph all data within the WHO dataset.

4. Individual report of Umaima Khurshid Ahmad

Each paragraph refers to the question stated for the individual summary

I used the variables life expectancy in years at birth, country names, and the corresponding years from the dataset. Life expectancy in years at birth was an important variable to use as our main aim in the project in exploring the trend of life expectancy of countries over the years. Country name as being the major variable was further categorized into developed and underdeveloped countries. Our dataset's major setback was that we didn't had a separate column for developed and underdeveloped countries, so we used the internet to find the most developed and developing countries.

I applied two visualization techniques; one was a heatmap showing life expectancy of years at birth between the underdeveloped and developed countries. Another was the Shiny application that shows the comparison between countries' life expectancy by selecting a specific year.

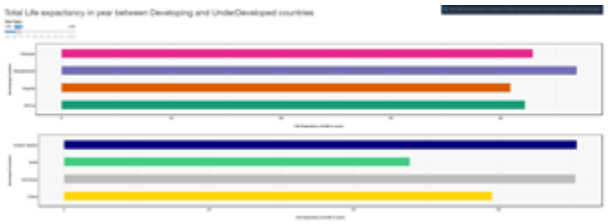
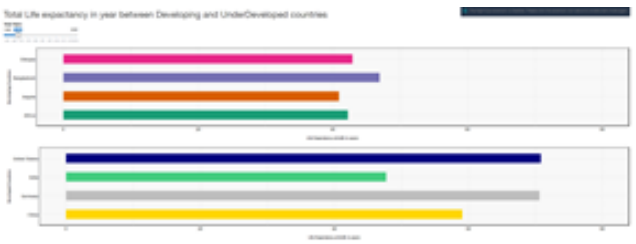


From the heatmap, we can clearly see the trends from the color scheme of the life expectancy of underdeveloped countries is low while countries that are developed countries have a higher color region across all the projected years. The visualization clearly shows why the raising trend life expectancy of some developing countries. On the other hand, the shiny application shows an interactive behavior with the audience, if the audience wants to know the life expectancy of the developed and underdeveloped countries of specific countries, they can easily see the visualization and do some analysis. We can see that the developing countries have a lower life expectancy as compared to the developed countries.

For the heatmap, I transformed the data variable life expectancy in years at birth to mutate it for creating a rich effect. The column had all numerical numbers from 30 to 80. The mutated variable created that transformed all the variables into categories (2d binning). The different categories were from "30","30-40","40-50","50-60","60-70","70-80",">80". With the help of colors, the audience could easily visualize the country's life expectancy they fall into through a specific year. The shiny application was revised from the feedback given by the professor. The audience can compare life expectancy details between specific years. This is an efficient way to analyze how life expectancy was between the developed and the underdeveloped countries in a specific year. After the feedback, The select option was replaced by a slider option, as the year column has many variables from 2960 to 212018. As a user experience, it is easier for the user to visually see all the years and drag to a specific year rather than choose from a drop-down menu.

As learned in this course, data visualization is an iterative process to out crisp and compelling visualization. There were several changes made during the creation of these visualizations. Concerning the heatmap, the audience can get confused with the different labels of years as labels were automatically created for the life expectancy in years. We could easily see in the presentation images as well that the message for life expectancy between the developed and the underdeveloped countries was portrayed wrong as both of the graphs were ending at the same point. Still, if we looked closely, we could have seen the developing countries' x-axis end at 40 years while developed countries ended at 60 years. If we want a message to be given as the life expectancy is constant between the developing and developed countries, we

would have used the default range. Still, our main idea is to inform the audience about the shortfall of life expectancy between the developing and underdeveloped countries; hence the axis was changed and equalized.

Before Adjusting the scale on Year 1970	After Adjusting the scale on Year 1970
	

The color scheme in the shiny application complements the comparison between the countries themselves. Simultaneously, it is also an overall comparison between the developing and developed countries. We have studied a general rule of thumb that the human eye would relate those things together that are boxed across each other.

The segmentation and continuous color scheme in the heatmap show how life expectancy has changed how low to high between the countries. (Lighter color shows low life expectancy while darker colors show higher life expectancy)

I developed both the heatmap and the shiny application using R. Hence coding was done to clean the data and develop the following visualizations. The added advantage of Shiny application is that it has created a more prosperous domain by presenting an interaction with the audience. The audience can now easily visualize and compare different years of life expectancy amongst developed and developing countries.

Overall, I believe that the visualizations I created show two different aspects of the data. Firstly, through the heatmap an overall structure on how life expectancy has a chance over the years by just looking through a naked eye, the audience can visualize the lighter to higher color trend(low to high life expectancy). Secondly, the shiny application enhances the heatmap by projecting a deeper insight into years that the audience wishes to see.

My visualizations demonstrate a pattern and a correlation between the life expectancy of countries shown. We can see the United States has a constant life expectancy till 2018 while Germany is also being declared as a developed country. We see an increase in life expectancy from 2010 to 2018. Furthermore, Angola being an underdeveloped country, we see an increasing trend in life expectancy through different shades of colors, but we can say it is developed as it has less life expectancy. We can easily visualize that the Central African Republic is the most underdeveloped country through the 19th and 20th centuries as its life expectancy is in the bracket of 30 – 50 years while the most developed countries have the bracket of more than 80 years of life expectancy.

I was trying to make a cartogram of the world amongst a small multiple to show how life expectancy has changed overall. However, doing that in R was a little challenging. If I have more time, I would have explored how to make a cartogram in R and side small multiple graphs of years and life expectancy.

References

Code Link: https://drive.google.com/drive/folders/1Mi6Tg5JUzPh8lmi4A5jpcd1qzZAtSjc_?usp=sharing

Developed Link: Top 25 Developed and Developing Countries (investopedia.com)

Underdeveloped Link : UN list of least developed countries | UNCTAD

Dataset Link: https://drive.google.com/drive/folders/1Mi6Tg5JUzPh8lmi4A5jpcd1qzZAtSjc_?usp=sharing

Dataset Source Link: https://databank.worldbank.org/source/world-development-indicators