According to our data, it can be found that the value of Life expectancy is sequence. In addition to that, we can find that there are 21 attributes that can be used to study the relationship between life expectancy. These 21 factors are differentiated. We can simply separate it into two categories: one is discrete, such as Country, Year and Gender: the choice of these attributes is limited, so we can consider directly visualizing the relationship between each optional value of these attributes and life expectancy. The other category is sequence attribute, such as clean fuels and cooking technologies, GDP, Mortality caused by road traffic injury, etc. For this type of attribute, it means that we should visualize the relationship between two sequential variables. These two categories need to be discussed separately.

Consider first the visualization of discrete data. The statistics of discrete attributes is also different: Country contains about 200 optional values; and gender undoubtedly has only two optional values. There are 20 optional values ​​for Year in our data: from 2000 to 2019. The number of different optional values ​​makes the idea of visualization different.

First consider drawing the relationship between Country and Life expectancy. We once learned in session 6 how to draw visualizations of drinking water conditions in each country. Here we are deeply inspired and directly linked the Life expectancy to each country's section on the world map. Since Life Expectancy is a continuous value here, we thought of using a Colorscaler() to reflect Life Expectancy as the color of the layout.

The specific technical realization is D3, which we learned in the course. We downloaded the geojson data of 248 countries around the world from the website https://geojson-maps.ash.ms/. The reason we did not use the geojson data provided in session 6 is that its resolution is not very satisfatory. We want to show the borders of each country more clearly. Then, we load the dataset and bind the data for each country (all Years and all genders) with the corresponding geojson. Then, for color, we choose the color schema of d3.interpolateGnBu().

Finally got the world map image. (attach pictures)

Similarly, although our main visualization goal is to explore the relationship between Life expectancy and other factors, we found that on this world map, we can also compare different countries with attribute values ​​​​of other parameters in this picture, such as the heat map of GDP and Suicide rate in different countries. These are also features worth exploring. We will discuss this specifically when visualizing continuous attribute values later.

Second, we want to visualize the discrete data Year and Gender. For the year, what we want to see is the relationship between the value of each parameter as the year changes. With the increase of the years, (of course, the advancement of technology), we all know that there will be a more or less increasing trend in life expectancy, and we want to show this trend more clearly. Our data can be classified by country and the data differences between different countries may be more volatile. For example, in the past two decades, the Life expectancy range of Chinese men is (70, 74), while that of Namibia men is (50, 60). Putting the pictures of the two together in the past two decades will make people feel confused, since this is more biased towards highlighting the gap between the two countries' data than the relationship between Life Expectancy and age, which we have shown in the map to some degree.

Therefore, we choose to separate each country first, and first visualize the development trend of Life expectancy in a single country over time. Line charts are very effective for showing trends. Of course, we can't just show only the trend in one country. We need to show trends in all countries likewise, but it's not realistic to draw these lines on the same canvas: we have 248 countries, so we need 248 lines! Therefore, we bind the canvas of the line chart to the world map and used an event binding to the country: whenever an image is clicked, the line chart of the corresponding country will be displayed.

For the last discrete attribute, we also show the relationship between Gender and Life expectancy. Notice that there are only two options for Gender here, male and female. To be specific, we need to show the distribution difference and summary statistics between the two genders. Therefore, we thought of the Box Plot used in session 4. We draw all data points, box plots of male data and female data on the same canvas, and we can see the difference between the two. Attach the picture.

Of course, for men and women in a single country, the development trend of their life expectancy is also very different. Considering this, we have added two different lines for men and women in the above-mentioned line chart to show that within a country, typically the difference between genders (not only the value, but also the trend). Of course, if we want to display the trend of a single gender, we can also choose to hid another, because we have added a selection box at the top of the page to make the selection.

Now we need to deal with the visualization of continuous values. Here we take one of the Basic sanitation services as an example. The value range of Basic sanitation services is (0, 100).

We want to establish a connection between two sequential values. Going back to the line chart we created earlier, it shows the trend of life expectancy with age. Inspired by this, if we also draw the development trend of Basic sanitation services in this graph, we can realize the comparison between the two data trends. Therefore, we consider adding another y-coordinate axis on the right to reflect this.

It is worth noting that here, we did not use the method taught in session 5 to use D3 to draw this line chart. Because in actual operation, we feel that it may be difficult to draw multiple line charts and multiple coordinate axes on the same canvas with the theoretical d3. Therefore, we apply an integrated API Echarts to make the visualization. The figure drawn by Echarts would be smoother and more beautiful than that drawn by d3.

Finally, thinking back to the world map and line chart we have drawn, it could be found that we have been comparing data within a country and comparing the development trends of various factors in a single country. Our data is much more than that: we have data for every country in the entire world. We should not only show a single country, but also visualize the development trend of the whole world, since the data of a country may be more or less accidental. Therefore, it is essential to show all data for all countries.

Back to the question of how to present data for all countries. The data to be displayed are: year, gender, life expectancy and Basic sanitation services. There are four channels of data that need to be displayed. Among them, life expectancy and Basic sanitation services are continuous parameters, so we thought of scatter plot an used these two as the main coordinate axes. For gender, we can show it by two different colors. As for the year, we thought of another implementation method: using animation to show the trend. We can first show the life expectancy with Basic sanitation services for all countries in 2000, grouped by gender. Then animate it over the years to show the trend. With this animation, we can see that the collection of points for the entire world has a somewhat slow-moving trend over the years: exactly what we want. Of course, if we want to see the details of a specific year or a certain gender, we can also stop the animation observation. Attached pictures.

There is also a selection box in the top where we can choose different attributes to show.