**Add a rationale for your design decisions. How did you choose your particular visual encodings and interaction techniques? What alternatives did you consider and how did you arrive at your ultimate choices?**

Before all the visualization work started, we first settled down our research question: How is medical resources related with people’s life expectancy? With the OECD data available from <https://www.oecd-ilibrary.org/social-issues-migration-health/data/oecd-health-statistics/oecd-health-data-health-care-resources_data-00541-en> , and the World Bank population data (

<https://data.worldbank.org/indicator/SP.POP.TOTL>), we were able to start our visualization.

As a starting point, we would like to include 2 variables related to medical resources, as well as the target variable, life expectancy. Along with the year and country variables, the total number of variables came to 5. To be able to incorporate all this information into our visualization, we considered bar plots and heatmaps, but decided that the example we saw in class regarding population, life expectancy, and fertility might be a good practice to follow. The interactive scatter plot is good because on both axes we could include two quantitative variables, and use additional encoding such as size or shape of each individual point in the scatter plot, to represent other variables. Moreover, the interaction added one more variable to consideration. These could easily make up to 5 dimensions.

Eventually, we decided that color might be the easiest way for us to distinguish between different countries, so we encoded color. We also considered the fact that position in plots provides information most effectively. So we put the two variables that we were most concerned about into x and y axes respectively, namely the total number of health employment and life expectancy. The size of the points was used to represent the number of hospital beds. All these variables are encoded to be the average over 10 years of data for each country. The number of beds and the number of health employment are calculated to indicate medical resources for each thousand people. Since we only had 10 years of data, using year as the interaction variable might not reveal the trend very well. Therefore, we used country as the interaction variable, and in the main plot we differentiated between European OECD countries and non-European OECD countries. In this way, we can compare across countries, as well as across OECD countries in different continents. We also used tooltips to help clarify the detailed information contained in each point.

Additionally, using the interaction variable “country,” we sought to include more specific information about each individual country in 2010-2019. There are three subplots we included, first being a stacked bar chart detailing the total number (including categories of professions) of health professionals. The second subplot includes the selected country’s life expectancy along with total health employment. The third subplot includes life expectancy along with total hospital beds. Note that we included the raw values in these subplots, without any calculations. This aims to add additional information to the viewer regarding the medical resources in a given country across those 10 years. We later added color components: we maintained the same colors for life expectancy lines in two subplots; we kept different bar plot colors to indicate differences for variables.

The initial idea for the interaction between main plots and the subplots below was through a click. Namely, if we click on a specific country, we would like to see it light up and the subplots change with respect to the click. But we decided not to do it this way and used the traditional dropdown menu because it is impossible in Altair to dynamically display the selected country name on the plot. More importantly, we will not be able to investigate medical information about a country by its name.

Moreover, we added the main title to indicate our topic. Thereby, we have created and implemented our visualization.

**Describe how the work was split among the team members. Include a commentary on the development process, including answers to the following questions: Roughly how much time did you spend developing your application (in person-hours)? What aspects took the most time?**

We began our meeting for the assignment by 26th April, Tuesday and stuck to a frequency of roughly three times per week. For each meeting, we spent around 4 hours discussing, working and coding together. Thus, we 4 partners spent nearly 25 hours together on this assignment. We followed such a strategy for the reason that our output for the assignment should be an interactive visualization and is hard to be finished in a parallel manner. As we sat together in person, we could split our work into small parts for each member. In this way, we efficiently cooperated with each other. We four worked in the same Colaboratory. Though we experimented separately, we would only save the effective part sequentially to avoid the problem in version control.

Though we would always work together and contribute in every aspect, we did have different emphasis on each part. Bohan and Wenjin found the dataset we finally utilized. All the four members contributed in data cleaning and manipulation. Bohan, Dongyang and Peihong contributed in implementing and polishing the visualization. Dongyang and Wenjin sought help from the professor and TAs and finished the writeup. Bohan and Peihong focused on the aesthetics of the project. Dongyang and Wenjin provided support for team meetings, such as scheduling rooms and taking notes.

In the first two meetings, we confirmed our dataset to use, questions to solve and what visual encoding, visualization tool and interaction techniques to leverage. There was a problem that during the first meeting, we planned to mainly implement a scatter plot with size encoding and time line interactive bar involved. Yet on that Thursday’s course, our TAs shared the coding for this visualization through D3. Though we would leverage Altair, we still thought a visualization highly similar to the chart introduced on the course may not be that good. In this way, we dropped our work from the first meeting and tried to come up with different visualization methods. For the next two meetings, we finished our exploratory analysis of the dataset and managed to get the visualization done following our original plan. During this process, we met quite a few problems in coding and publishing our product. We turned to our professor and TAs and got help to come over most of those boulders smoothly. Though there were still some questions to do with the tool Altair left unsolved, we managed to achieve our visualization with other methods. We then held a final meeting to further polish our visualization and finish our write-up for the assignment.

During the whole process, coming up with a good design of the visualization and the coding part did cause trouble to us. Due to our lack of experience in implementing visualization, we easily got lost when designing the chart. Though we had a clear research question, we failed to always utilize our question to lead our design of visualization, thus we easily walked into the trap of becoming focused only on what kind of techniques to use. These difficulties took us most of the time, but did give us great lessons for moving forward. For example, we managed to prioritize tasks in designing a chart, from the data manipulation, to the main plots, and interaction, and finally the details such as color selection. For the final project, we will further explore other visualization techniques and apply some more theories we learned in class.