

Project 2 Final Submission

Sydney Bednar, Noor-E-Jehan Umar, Victoria Eshun, Emily Robinson

Project Topic: “Will Moving to this Country Make You Live Longer?”

Project Idea

For this project, we wanted to assess, through interactive visualization, whether the wealth of a country affects the life expectancy and causes of deaths of its citizens. In particular, we wanted to see whether how much money a country has varies the most common causes of death in that country. To test this, we used a gapminder dataset, with information about different countries' GDPs, populations and life expectancies, and a causes of death dataset that compared the different causes of death in countries' as well as their frequency. We also used a topo json dataset of countries to make our map visualization, and a flags icon dataset to add visuals for each country.

Data Description

- *Cause of Death Dataset:*

This dataset includes historical data of different causes of deaths for all ages around the world. The original dataset can be found [here](#). There are columns to store the name of the country/territory, the country/territory code and the year of the incidents. The other columns consist of different causes of death including multiple diseases, violence, drug abuse, and environmental factors. Each of these columns store the number of people who have died from the given cause during that year in each country.

- *Gapminder Dataset:*

We found three gapminder datasets, one for the life expectancy of different countries' populations, one for GDP, and one for population. Since these were different datasets, our first step was combining all three of them. We did this by using the Pandas and NumPy libraries to perform a left join on the datasets on the country and year column. Furthermore, we used the filter function to filter our countries to only include the most recent year, 2019. We also realized that we needed to convert some data to numeric values to make it usable. For example, population and GDP were both strings, which we converted to numbers. Furthermore, some of

the GDP values were in billions and some were in millions, so we had to remove the current GDP column and find a more suitable one. The raw datasets can be viewed [here](#).

Additional Data:

- *TopoJSON:*

We used a TopoJSON file to make our map visualization. We found one that had information on countries. This can be viewed [here](#).

- *Flag Icons*

We also used another gapminder dataset that had visuals of country flags. We wanted to use this dataset to help with our filter visualizations, where users could click on a flag to filter the results associated with that country. The dataset can be viewed [here](#).

- *Country Codes*

We used a world countries dataset that had country codes associated with each country, and added the respective country code to each country in our TopoJSON file. This made it easier to compare countries in the TopoJSON dataset to our GapminderDataset since they now had a consistent common key. This dataset can be viewed [here](#).

Design Rationale

Map:

For our first visualization, we worked on an interactive map that shows the user differences in GDP, population and life expectancy across countries. We felt that a map was definitely the best visualization for this data because it was divided by country, and we felt that it was an efficient way to separate the different features by showing them as their own map views.

The main channels we used for this visualization were color and area. This is the most standard way to represent different data on maps and would be most intuitive for our users. We used sequential single hue color scales for each map visualization, with darker colors representing higher values and lighter colors representing lower values. We chose to go with sequential scales rather than diverging ones because we felt it would be more intuitive for the user if one indicator (GDP, population, life expectancy) was associated with one single color,

especially if they were switching between them using the buttons. If there was more than one hue for each, it might be confusing for the user and lead them to spend more time figuring out which color means higher values and which one means lower values. We also used area as a visual channel, scaling each country to their relative geographical areas so that the user can easily visualize the relative sizes of the countries.

Another design choice we made was to remove the zoom feature from our map. While we initially had included this feature, upon gathering feedback from other students in the class and through discussion with each other, we realized that there was no reason for the user to need to zoom into the map. This was because the main point of our visualization was to show broad differences in different countries' GDPs, populations, and life expectancies, which was done through the differences in color luminosity. Zooming into a country would not provide the user with any useful additional information and might just make their interactions frustrating if it keeps accidentally zooming in. Lastly, because the user can see more information about a country through our hover interaction, there is no need for more information to show through zooming in.

We also added a color legend for each of the three map visualizations on the top left. We chose to add a legend because it is custom in most colored maps, and we thought adding it on the top left would make sure the user could see the information easily if they needed to, without having to scroll down to look for a legend.

Lastly, we chose to display information for each country's gdp, population and life expectancy in an additional tooltip that shows upon hover. We did this because this would make it easier for the user to get all the information of a country in one spot if they wanted, without having to switch between the maps. We kept these in close proximity to the country being selected, and we also included a thicker outline stroke around the country border upon hovering.

Bar Chart:

For our second visualization, we created an interactive and dynamic bar chart that changes according to the country selected. The bars show the top five causes of death in the country and how many people died from that cause. We chose a bar chart to present this data because it was comparing different amounts of a value. We chose a horizontal bar chart with causes of death on the y axis and number of deaths on the x axis because we were comparing nominal values (causes of death) and felt they were more emphasized when they were on the y axis.

We used a `scaleLog()` for the x axis because the data for the number of deaths was pretty similar in value to each other and was becoming very clustered when we used a linear scale.

The log scale made it easier to see smaller differences between the number of deaths by the different causes.

For the bar chart, we used color as well as varying horizontal position on an aligned scale as our main visual channels. We created a `scaleOrdinal()` color scale for the causes of death, so each cause would have a unique color; we felt that it made sense to vary hue because we were dealing with categorical data. This would ensure that the user would be able to use color to differentiate the different causes from each other, as well as tell when different countries have the same top cause of death. The varying horizontal position lets the user see differences in the frequency of the different causes of death.

We decided to use the flags of each country along with the names to allow users to filter the countries because this would make it easier for the user to know which country they would be looking at. Users can switch between top 5 gdp and bottom 5 gdp countries using a dropdown menu. We felt this made sense because it made sure the screen was not too cluttered with all ten flags present at once.

Interactive Elements and Rationale

Map:

- *Buttons*

Above the map, the user is presented with three buttons labeled 'GDP', 'Population' and 'Life Expectancy'. We used buttons to indicate to the user that these different names were clickable. When the user clicks on a button, the associated colored map for that value is shown. For example, if the user clicks on GDP, the map is colored using the sequential scale we used for GDP. The map changes according to the id of the button that is clicked.

Furthermore, when a button is clicked, a box shadow appears behind it, its font-weight turns bold and its background color changes to the same color as the fill of the map associated with. The other buttons remain the same as they initially were. This helps give the user feedback for which button has been clicked on. When a user switches to a different button, the other buttons return to normal.

Lastly, when a user clicks on a specific button, a customized legend for that particular indicator (GDP, population or life expectancy) is shown. This helps the user see that darker colors are associated with higher values and lighter with lower values.

- *Hover*

When users hover on a specific country on the map, a tooltip appears, giving them more information about: 1. The name of the country 2. The GDP 3. The life expectancy and 4. The population. We added this feature because we got feedback from other students that they wanted a better way to compare GDP, population and life expectancy between countries. While the user can see the difference in color luminosity to tell how high or low these values are on a map and can switch between them using the buttons, they might forget which country they were trying to compare or not even know what country it is.

Therefore, we added this additional information so the user just has to hover on the country to get all the information at once. They can also get more specific information in numbers for the GDP, life expectancy and population this way. However, if they want a bigger, more general picture, they can just look at the map as a whole.

To make it discoverable, we added text at the top, telling users that they could hover to get more information. Furthermore, the tooltip that appears is close to the country that is being hovered on, indicating to the user which area it is associated with.

Upon hovering over a country, its outline would appear thicker - making the user's selection more distinguishable. This also helps the user realize which country they are selecting at a given time.

Bar Chart:

- *Drop-Down*

Above the bar chart, the user can choose to view a list of the top 5 countries or bottom 5 countries in terms of GDP. We included this drop down menu to consolidate the view and to ensure that the user could view all elements of the chart without scrolling. The arrow on the far right of the menu signals to the user that it affords clicking and choosing an option from the subsequent list. When the user chooses an option, the flags and labels below it update to show the chosen set of countries.

- *Hover Over Flags Filter*

Below the drop down, the user can hover over each flag to view the top 5 causes of death in that country. When a flag is moused over, the associated label is bolded, the flag appears with a shadow under it, and the chart updates its axes and bars to provide feedback to the user that their selection has changed. We decided to show flags for a visual cue to users, and we wanted the bar chart to be neat, so we chose to include a filter by country to reduce the amount of data shown on the chart at one time.

The Story

The main questions we were trying to explore through our visualizations were:

1. Whether richer countries have higher life expectancies and populations
2. Whether the health of a country varies by how rich it is, i.e whether the main causes of death in the richest countries are similar to each other and different from poorer countries.

Through the map, we discovered that many countries with the highest GDPs in the world (USA, China, India) have some of the highest populations in the world too. Some countries with high GDPs, like Japan and Germany, did not have the highest populations. However, we inferred that this was because these countries are really small in size, so we decided to focus more on the bigger countries mentioned before for analyzing this trend. We concluded that our visualization suggested that there could be a trend between countries with higher populations having higher GDPs too.

When we compared life expectancy to GDP, we found that there was not really a clear trend between countries with higher GDPs and countries with higher life expectancies. While Japan and Germany, which have 2 of the highest five GDPs in the world also had some of the highest life expectancies in the world, the USA, China and India, the leading three countries in terms of GDP, had lower life expectancies in comparison to a lot of other countries. In conclusion, we decided that this meant there was not really a clear correlation between the two. This was surprising because we had assumed richer countries would have better health facilities to help their citizens live longer.

We had given our users this information in the first visualization because we wanted them to understand the basic relationships between life expectancy and GDP of a country, before moving onto the specific causes of deaths in those countries. While looking at life expectancies is one way to assess the health of a country, looking at morbidity provides additional insight. In our bar chart, we wanted to see whether how rich a country is affects the most common causes of death in it. We used the countries with the top five GDPs: China, USA, India, Japan and Germany, and the countries with the bottom five GDPs: Tuvalu, Nauru, Marshall Islands, Kiribati and Palau.

Our barchart showed us that the top five countries did indeed have a lot of overlap for the most common causes of death. In particular, cardiovascular diseases, neoplasms, chronic respiratory diseases, and digestive diseases were common amongst all five of them, with cardiovascular disease and neoplasms being the highest in frequency for all of the countries.

Similarly, we saw a lot of overlap between the bottom 5 countries as well, with all of them having Cardiovascular Diseases, Neoplasms, Diabetes Mellitus and Chronic Respiratory Diseases in common as their top causes of death.

We saw that there was actually a lot of overlap in the top causes of death between countries with the highest and lowest GDPs. While there were some differences in some top causes of death, such as diabetes being very prevalent in the countries with the lowest GDPs but not in those with the highest GDPs, most of the causes were pretty similar. Cardiovascular diseases, in particular, seemed to be the leading cause of death in all of these countries, followed closely by neoplasms. This leads us, and the user interacting with our site, to the conclusion that wealthier, larger countries are still plagued with the same diseases as poorer, smaller countries.

If we had more time for this project and access to more data, we would have considered plotting the same visualization but using the GDP per capita for each country to see if that affects our results in any way. This is because a lot of the results in our visualization could be attributed to the fact that the top countries and the bottom countries had very different population sizes.

Team Contributions

Sydney:

- Helped sketch and brainstorm project ideas
- Filtered data for causes of death and country attributes
- Created initial bar chart axes
- Created flag hover interactions and styling
- Created color legends
- Worked on updating bar chart colors and bars
- Worked on map color scales, tooltip location, and legends updating
- Worked on styling graphs, buttons and filters
- Helped write report

Victoria:

- Helped sketch and brainstorm project ideas
- Cleaned and combined multiple datasets into one gapminder dataset
- Created map color scales
- Worked on map and bar chart hover functions
- Worked on placing and binding tooltip within chart area
- Wrote comments for code and edited descriptions for visualizations

Noor-E-Jehan:

- Helped sketch and brainstorm project ideas

- Worked on updating bar chart colors and bars
- Worked on placing and binding tooltip within chart area
- Worked on updating map color scales when button clicked
- Helped on button styling for map
- Wrote report

Emily:

- Helped sketch and brainstorm project ideas
- Created mesh for map
- Made stroke on hover for map countries