Project 2 - Final Report

Team members: Zhi Lin (zl846), Tianchen Wang (tw537), Czar Carson (ckc65), Rishabh Prakash (rp429)

Visualizations Overview

Description of The Data

Country GDP Dataset: GDP Among World
World Map TopoJSON: World Atlas TopoJSON

We accessed and downloaded our GDP dataset from Kaggle.com. This dataset included both a JSON and CSV file containing economic data for 184 countries worldwide; through this dataset, we had access to each country's land area, region, population rank, % of world population, GDP, GDP per capita, and more. The dataset was well organized and formatted, requiring very little cleaning. We used the **GDP2.csv** file within our project.

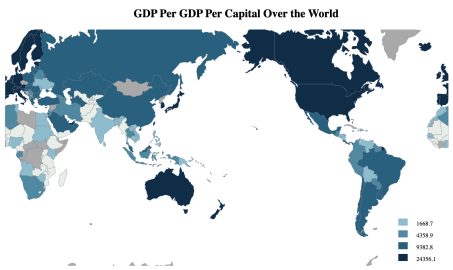
We accessed and downloaded our world map topoJSON through GitHub. We imported only the **countries-110m.json** into the project, which is a world map on a 1:110m scale; we chose this size because it fits nicely on a laptop screen.

For each graph, we used smaller subsets of the dataset containing only the necessary data to create each visualization. For the world map, we utilized both the topoJSON world map and the

GDP dataset; from the GDP dataset, we used the county and GDP per capita. For the scatterplot, we used the country, land area, and GDP per capita from the GDP dataset. For our bar chart, we used the country name, GDP per capita, population rank, growth rate, land area, and density from the GDP dataset. Within our collision graph, we created a smaller dataset containing only the country, region (renamed to the continent), GDP (IMF), and GDP (UN) from the GDP dataset. And for our drop-down circle, we used the length of the GDP dataset to determine how many countries were included.

Overview of Visual and Interactive Design Rationale

Plot 1 Map with Zoom Function:



- This plot shows the 2020 GDP per capita worldwide. This plot answered the question: How is GDP distributed in the world?

➢ *Visual Design*

- **Marks**: Polygon.
- **Visual Channel 1**: Varying the color hue of the Polygons.
- **Visual Channel 2**: Varying the polygons' horizontally aligned position and vertically aligned length.
- **Visual Design Reasoning**: Our dataset includes GDP for countries' information, therefore using a map is the most effective way to show the viewer distribution of GDP worldwide.

➢ *Interaction Design*

- **Interaction**: Zoom in and zoom out the map.
- **Interaction Design Reasoning**:

Zoom in and zoom out: Since we have many countries and some are smaller in size, letting the viewer zoom in and out of the map allows the viewer to see each country on the map more clearly.

Plot 2 Drop-Down Circle:

- This plot shows how many countries we have in the dataset.

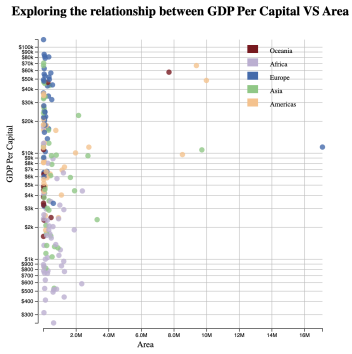
➢ *Visual Design*

- **Marks:** Circle.
- **Visual Channel 1:** Varying the color hue of the circle.
- **Visual Channel 2:** Varying the horizontally aligned position and vertically aligned length of the circle.
- **Visual Design Reasoning**: The numbers in the circles make the data more visible and easy for the viewer to see and add interest.

➢ *Interaction Design*

- **Interaction**: The circle will drop if the viewer clicks the replay button.
- **Interaction Design Reasoning**: By asking the viewers to guess the number of countries, and get the answer by clicking the button, the interaction process can make the viewers feel more fun and engaged.

Plot 3 Scatter Plot:



- This plot shows the relationship between GDP per capita and the country's area.

➢ ***Visual Design***

- **Marks**: Circle.
- **Visual Channel 1**: Varying the color hue of the circles.
- **Visual Channel 2**: Varying the horizontally aligned position and vertically aligned length of the circles.
- **Transformations**: We use a log scale for GDP per capita, which is the y-axis, since the distribution of the GDP per capita is not even.

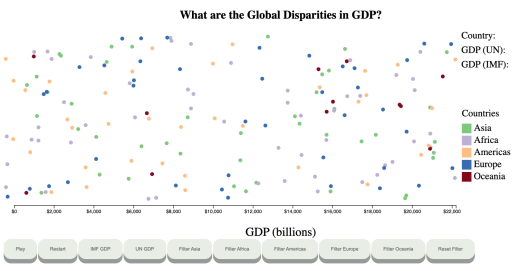
- **Visual Design Reasoning:** We chose the circle as the mark because the circle is more common and is a more recognizable mark; secondly, we used the color hue to distinguish the countries, and we adjusted the opacity to allow the viewer to see the overlapping data points.

➢ ***Interaction Design***

- **Interaction**: When the viewer hovers the mouse over the circle, the circle size will be a larger circle with a darker color; all other circle colors will be lighter, and there will be a black stroke to show the circle; also, when viewer hover the mouse to the datapoint, a text label which includes country and region information will show up in the top of the graph.
- **Interaction Design Reasoning:**

Hover: Since this scatterplot has many points overlapping, we chose to lower the opacity value to make the points more transparent so that viewers can see the different data even though data points overlap. When viewers hover over the circle, the circle will be larger in size, darker in color, and with a black stroke, so viewers can see more clearly the dots they selected. A label with country and region information will be displayed when viewers hover over the circle. This allows viewers to see which country the corresponding data point belongs to.

Plot 4 Collision Chart:



- This plot shows the GDP disparities between countries worldwide and within each continent.

➢ ***Visual Design***

- **Marks**: Circles
- **Visual Channel 1**: Changing the circles' horizontally aligned position

- **Visual Channel 2**: Varying the color hue of the circles
- **Visual Design Reasoning**: Using circles as a mark provided a quick and effective way to represent and differentiate between each country represented within the data. We use varying color hues to further help the viewer identify which continent each country belonged to, we felt this was important as it highlights the global trends in GDP disparities by allowing the viewer to see which continents are represented in the highest and lowest GDPs. We chose to identify GDP on the x-axis as it felt most familiar when utilizing a single axis.

One design tradeoff we made was not using a log scale for our data. Using a log would help distribute the points along the x-axis more evenly, helping the user make better distinctions between the countries with lower GDPs. Because there is such a big disparity between the highest and mid-level GDPs, many of the points get pushed to the side (meaning we may have a country with a $500 billion GDP right next to a country with a $50 million GDP). But we believed that this encapsulated our point, there is a huge disparity in the distribution of global wealth. That there exists a select few countries with insurmountable wealth.

➢ *Interaction Design*

- **Interactions**:

Hover: When the viewer hovers their mouse over each circle, the circle slightly increases in size to help indicate which circle is selected; To the left of the chart, the corresponding country name, continent, IMF GDP, and UN GDP are shown for the selected circle.

Collision*:* The viewer can utilize the Play button to run the collision simulation and the Restart button to reset each circle to a randomized position.

Filters & Axis*:* The viewer can utilize the IMF GDP and UN GDP to change the positioning of each circle corresponding to their IMF or UN GDP valuation (The GDPs are slightly different for each country); The buttons will also dynamically shift the GDP bottom axis to be representative of the new maximum and minimum values. The Filter buttons allow the viewer to filter the dataset by continent, allowing the viewer to only see the data for those corresponding countries; selecting a filter will also dynamically shift the GDP bottom axis to represent the new dataset range.

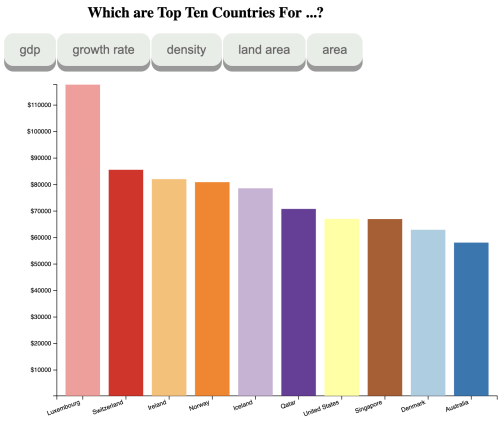
- **Interaction Design Reasoning**:

Hover: Because our data is represented by circles, we took special care to ensure users can easily identify each country.

Collision: We use our collision simulation as an interactive tool as it is a fun way and interesting way to demonstrate the disparity. It's one thing to see the disparity in a static graph, but seeing the points physically move provides an even more powerful visualization.

Filter & Axis: Filtering is our solution to not using a log scale; by allowing the user to see only the countries within a specific continent, they not only see each country easier but they can assess the disparity within each continent. Dynamically shifting the axis makes the chart usable for the user, as it is necessary to have an appropriate scale for our range of data when we filter.

Plot 5 Bar Chart:



- This plot shows the top 10 countries for each category.

➢ ***Visual Design***

- Marks: Rectangles

- Visual Channel 1: Varying the color hue of the rectangles.

- Visual Channel 2: Varying the horizontally aligned position and vertically aligned length of the rectangles.

- Visual Design Reasoning: Barchart is the most straightforward way of visualizing a distribution of a set of data points. In our case, performing a ranking or comparison of values across different subgroups of different attributes of our data. From a bar chart, you can easily see which countries are the highest and how each of the values in the country compares against the other countries.

➢ ***Interaction Design***

- Interactions:

Button Click: The view can click on each button to select the top 10 countries they want to visualize in that category. When a button clicks, a smooth transition will take less than 2 seconds to update the new bar value/height. Also, the bars will get a new randomized set of colors to keep the interaction visually appealing.

Hover: The view can hover over any of the bars; the bar will slightly expand horizontally, and the color will be lightened. On the top of the selected bar, the specific number of the bar will be shown. There will also be a dotted line across the top boundary of the bar.

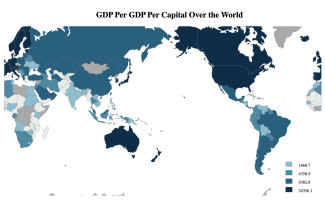
- Interaction Design Reasoning:

Button Click: Different options can give the viewer more freedom to see the category they are interested in exploring. The transitions help convey a mood or tone to viewers and set the stage for the next category. The purpose of the dynamic color scheme is to keep the interaction more fun and visually appealing.

Hover: These interactions can make the selected bar stand out when selected or hovered over so that users can easily confirm that this is the bar they want to see. Not only this makes the interaction more fun, but this interaction's obviousness will also increase the amount of feedback the user receives and therefore increase the readability of the data.

Storytelling

Plot 1 Map:

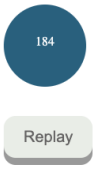


This plot shows the 2020 GDP per capita around the world. This plot answered the question: How is GDP distributed in the world? From

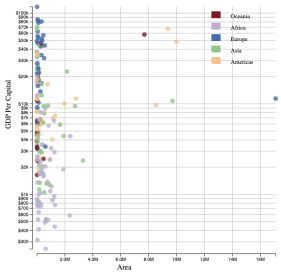
the map on the left, we can see that some countries, such as the United States, Australia, Canada, South Korea, Japan, the United Kingdom, and France, have relatively high GDP per capita, and some countries, such as China, Russia, Brazil are in the middle; it is worth noting that, since we only have 184 countries, there are some countries in the map that do not have corresponding values in the dataset, so these countries are colored dark gray. The viewer can zoom in and out of the map to check the color of each different area.

Plot 2 Drop-Down Circle:

How many countries do we have in this database? There are more than 190 countries worldwide, but only 184 are in the database; we cannot collect data for some countries for various reasons. Accuracy is one of the important criteria for data visualization; we want to be able to tell the viewer exactly how many countries are included in this database. When the viewer refreshes the interface, the circle will fall, and the viewer can also click replay to replay the animation.

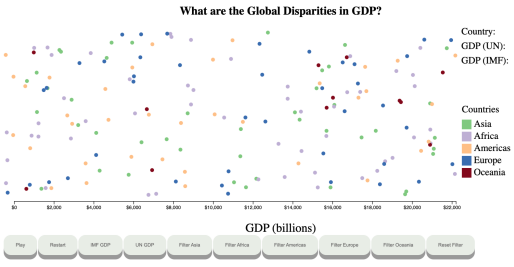


Plot 3 Scatter Plot:



The scatterplot on the left shows the relationship between the area of the country and the GDP per Capita. The x-axis of this scatterplot is the area, and the y-axis is GDP per Capita. This graph shows no strong linear relationship between area and GDP per capita. But we can still find some interesting facts from the figure: the GDP of countries with a large land area is often not low, but the relationship of GDP of countries with a small land area seems irrelevant. The viewer can place the mouse on the point, and the region and country will be displayed above the picture; a legend on the upper right shows different colors for different regions.

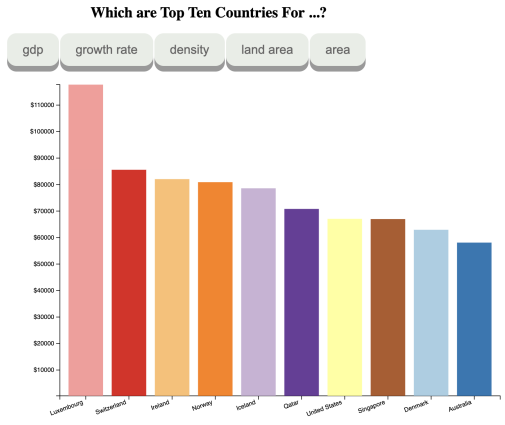
Plot 4 Collison Chart:



This chart shows the disparities in GDP globally. With this chart, we wanted to explore the wealth that exists worldwide and answer the question: what does the global distribution of wealth look like? Our graph shows how much bigger the United States GDP is compared to the rest of the world, only challenged by China. Within our graph, we can see that only around a dozen countries crack the $1 trillion threshold, while the rest of the world is beneath; this demonstrates that most of the wealth is centered among a select few countries. When we filter by continent, we see this same wealth disparity continue. The biggest disparities exist in Asia, America, and Oceania, all having one country vastly more

wealthy than the rest. Europe and Africa's disparities aren't as extreme, but still very much present. From our graphs, we see that extreme wealth disparity exists both globally and continentally, and above all else, America is absurdly rich.

Plot 5 Bar Chart:



This bar chart shows the top 10 countries in these categories: GDP per capita, growth rate, density, land area, and area. While most of the above graphs give an overview of the distribution of all countries of a specific category/attribute worldwide, this visualization allows the viewer to explore five different categories/attributes and highlights only the first ten highest data points. Therefore, this bar chart is complementary to the other graphs because it shows different aspects of our dataset in a more detailed and focused manner.

We have some interesting findings from this bar chart. The trends of the top 10 countries in each category are very different. Singapore has the highest density and is a lot higher even than the ranked-2 country. On the other hand, the growth rates of the top countries are very similar. The reason might be that a country is a large entity, so it is relatively stable. It is hard for a country to dramatically change its growth rate over time.

Conclusion

In this project, we learned the relationship between the distribution of GDP and other factors in different countries.

From the map, we can see the distribution of GDP by country; from the scatter plot, we can see that the relationship between land area and GDP is weak, especially for countries with small areas; from the Collison chart, we see the vast wealth disparities that exist globally.

Contribution

Zhi Lin	Czar Carson	Tianchen Wang	Rishabh Prakash
1. Visualization Design 2. Barchart visualization implementation 3. Code merge	1. Visualization Design 2. Collison chart visualization implementation 3. Code merge	1. Map visualization implementation 2. Scatterplot visualization implementation 3. Drop-down circle	

4. Final polishing 5. Documentation	4. Final polishing 5. Documentation	implementation 4. Code merge 5. Final polishing 6. Documentation	

Time	Part
8 Hour	Map
6 Hour	Scatter plot / Drop-Down Circle
9 hours	Collision Chart
7 hours	Bar Chart
6 hours	Documentation / Final Report