Data Description:

With a strong interest in historical work-living conditions, our team starts off exploring how working hours change over time in different countries. We got data of working hours [1] and world map json data from an open source repository [2]. Since a large amount of working hours data was missing in the early years, we only kept data from 1950 to 2017. Because we wanted to make a map visualization, we combined countries-110m.topo.json with working-hours.csv in jupyter notebook.

We further came up with whether working longer hours generates more wealth for a country, specifically GDP per capita. So we joined additional GDP per capita data[3] to each country in our original dataset matched by their unique country codes. In our exploratory data analysis, we noticed interesting and compelling patterns within different continents regarding how their working hours and GDP per capita are distributed. We added continents information by creating a country_to_continents map and mapped each country to its continent.

We end up with the following variables:

1) Year

2) Country

3) Annual Working Hours

4) GDP per Capita

5) Population

6) Continents (South America, Oceania, Europe, Asia, North America, Africa)

We also reformat the data to make our visualization more user-friendly and relatable for the viewers. We noticed that in common practice, working hours are counted weekly rather than annually, so we divided annual working hours in our raw data by 52 weeks to reformat it into weekly working hours. To support the yearly presentation of working hours and GDP per capita, we grouped the data by year and averaged working hours and GDP per capita grouped by continent under each year. In addition, we altered column names to make them easily identifiable in our code.

Design Rationale

[1] Data set: https://www.rug.nl/ggdc/productivity/pwt/?lang=en
[2] Data source: https://github.com/topojson/world-atlas
[3] Data source: https://www.rug.nl/ggdc/productivity/pwt/

The primary focus of our visualizations would be to see how working hours change over time and whether the varying lengths of work contribute to different levels of wealth. In the following paragraphs, we detail our design choices to justify our choices relating to accessibility, usability, clarity, and visual appeal.

For our first graph, we made a map visualization to directly present how working hours changes over time around the world. A map is effective in pairing up the working hours information with their geographical contexts. The mark is countries on the map. The channels are longitudinal position, latitudinal position, and color hue. We vary the regions by color hue according to different lengths of working hours. For this, we chose a quantile scale for working hours. The quantile scale splits the input domain, which is each country's working hours in our case, into intervals. The result is roughly the same number of values falling into each of the intervals. This is what we want because we want to show which quantile each country falls into. One of our concerns with this graph was missing data in many parts of the world, especially Africa. To tackle this, we added a "no-data" category and set it in grey to minimize the disturbance of these areas, so that the viewers will only focus on the others areas with vibrant colors. We also labeled it in the legend to avoid confusing the viewers. Finally, we included a legend on the bottom to provide clarity on which interval the color falls in to ensure precise understanding for graph users, as well as increase readability.

For our second graph, we choose a scatter plot to display working hours versus GDP per capita because we want to directly present distinguishable clusterings to the viewers. We used linear scale to plot the weekly working hours and GDP per capita. Our marks are the circles, and our channels are color hue, size of the circle, horizontal alignment and vertical alignment. We varied the color hue of the circles by continents; this provides a direct and eye-catching way of differentiating how working hours and wealth are distributed across continents. The users can recognize the clusterings and distinguishable patterns quickly upon seeing the visualization. Due to the large amount of overlaps, we added an opacity of 0.7. This allowed all of the marks to hold equal value and importance on the graph. To incorporate population information in the scatter plot, we set the radius of the circles to be varied by population. First, we choose a linear scale for the population to be set as the radius. However, this has made only a few countries with extremely large populations (ex. China and India) standed out while all the others are very tiny. We also experimented on using a quantile scale but that fails to display the difference between very large and moderately large countries and the graph was not as information-rich as we had hoped. Finally we choose a linear scale on the square root of the population which better differentiate countries with large, moderately large and small population while not overpowering the extremely large ones. We realized this has made population vary by circle area. In this case, we feel it is not a bad practice because we do not need the viewers to capture the exact value of the area; it is only meant to provide a general idea of population variance within countries on the graphs. Additionally, we added labeled the circles with their country names to increase readability and add context to the graph. One of our major concerns with the label was overlap due to similar x and y-values. To combat this, we selectively display 2-6 larger industrial countries within each continents. This decision does not loses too much information as the current visible countries already provide a salient context. We also adjusted the labels of the circles close to the right edge of the graph to be displayed on the left, so that

the texts are fully displayed. Finally, we included a legend on the right-hand side to provide clarity and precise understanding for graph users, as well as an axis and a legend to increase readability.

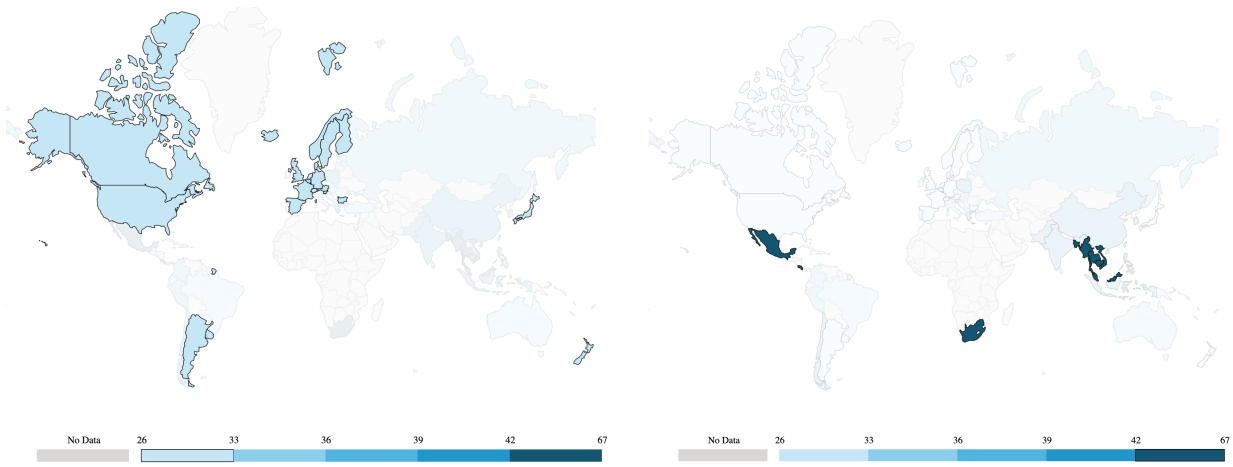
For our third and fourth graphs, we chose line graphs because we want to directly convey linear trends of how working hours and GDP per capita change over years across continents. We used a linear scale to plot years to weekly working hours for the third graph and linear scale to plot years to GDP per capita for the fourth graph. The marks consist of six different colors of lines and circles when hovering over. The channels are vertical alignment, horizontal alignment, and color hue. We varied the color hue of the lines and circles by continents; this provides a direct and eye-catching way of differentiating each continent. We utilized the same color scheme for each of the continents from the second graph to create a seamless transition. Since they share the same context, we did not feel the need to include an extra legend for these graphs as the information can be carried over from graph two.

Interactive Elements:

THE WORLD MAP

Our motivation for the world map is to give an intuitive understanding of the relationship between country location and working hours, and how it has changed over years. The world map contains three interactive elements. First, users can hover their mouse onto a country. The country will be highlighted and more information will be displayed in a box next to the country, also we modify the code in order to better display the box in a suitable location. The second interactive element is that when users hover over the legend with a specific range, the corresponding countries will be highlighted, and others will be dimmed so that the viewers could only focus on the countries within the selected range. This improved accessibility and allowed the viewers to directly capture patterns among different parts of the world.

For example, the graph belows shows that in 2009, countries with the lowest working hours are clustered in North America and Western Europe; and countries with the longest working hours are clustered in the southeast Asia and Mexico.

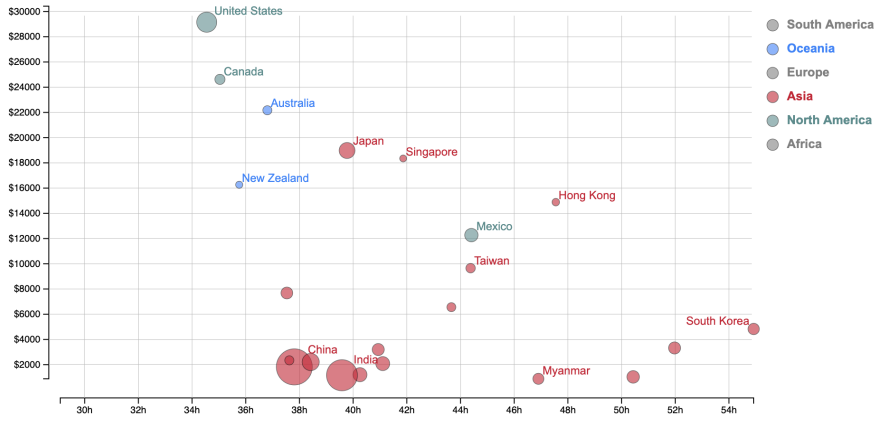


When hovering over, the quantile being selected is highlighted. This improves usability as it implies to the viewers that the legends are hoverable. The third interaction is that users can change data of the year by dragging the slider bar on the bottom. Through dragging the slider, the user will be able to see the change of working hours in every country over the years. Additionally, we added a year bubble that moved along with the cursor so that the viewers could be updated with the current year. We set the default year to 1980 when the viewers load the data. We chose 1980 because it is in the middle part of the slider and allows enough spaces for the viewers to slide in either direction.

THE SCATTER PLOT

We allow the viewers to drag a year soldier at the bottom on the graph to check how working hours versus GDP per capita are distributed around the world in different years. We added a year bubble that moved along with the cursor so that the viewers could be updated with the current year. We also set the initial year to 1980 to ensure consistency with the first graph. When users hover over the legend, we blur out the other continents so that the users can focus only on the selected continent. In addition, all of the country labels that were selectively hidden before will become visible since now the viewers focus on a smaller sample and there should be less clustering issues. The graph below effectively shows a distinguished pattern between Asia and South America countries.

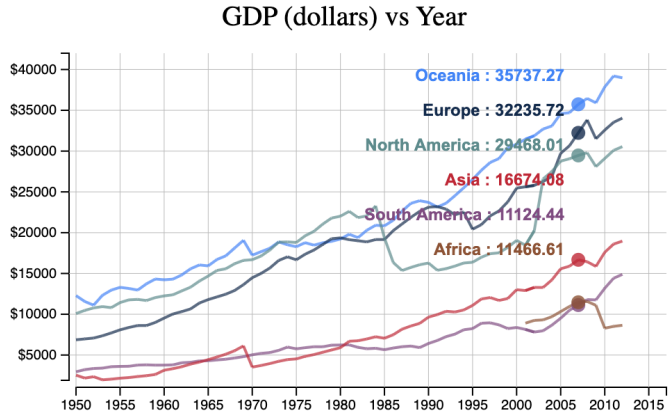
To further strengthen usability, we support viewers to click on the legend to select and deselect single or multiple continents. We style the cursor to become a hand when viewers hover it over legends, so that it is self-explanatory that the legends are clickable. On click, all of the unselected continents grey out so it is easy for the viewers to track the selected items. We also allowed the viewers to deselect the continents by clicking it again. When the last continent is unselected, all of the continents become visible again. Once a continent is selected, it does not blur out when the viewer hovers over the other continents. We've made sure such a clicking function is consistent with what we see on the professional websites so that the users will not likely be confused on the usability. Implementing such functionality is extremely challenging and time consuming. We've experimented with different logic to control the visibility in different conditions and eventually come up with a helper function to control and update the state of the marks.



Like the first graph, we support checking information of individual countries by hovering over. We present factual values like country name, continents, population, working hours and gdp per capita so that the viewers could instantly check data they are interested in.

THE LINE GRAPHS

We made the two line graphs below because the x and y-axis of the scatter plot above changes rapidly as viewers slide the slider. To avoid viewers completely losing track of how working hours or GDP per capita change over years, we add the two line graphs to allow simultaneous tracking of how the two indicators change on a larger scope. Therefore we use the same slider to control the scatter plot above and these two line charts to offer a more data-rich and salient context of how the two indicators change. We also support viewing individual points when the mouse enters the graph. Moving the mouse along the year axis, the circles become an effective tracker of how working hours or GDP per capita of each continent fluctuates. It also makes it easier for the viewers to make comparisons between continents each year.



The Story:

Working is a significant part of our lives. It's something we do almost every day. In today's hustle and bustle world, it's easy to assume that we are all working more than ever. But is that really the case? Since it is so important, exploring closely at how much time we spend working would tell us a lot about our lives, even our societies.

The first map here shows people around the world average working hours per week since 1950. Based on that, we found that in many countries today, people work much less than in the past 60 years. Working less means people being able to enjoy leisure time and spend time becoming more educated. This is substantial progress, but there are large differences between countries, for example people from countries in Asia do not work less, but even work more than before.

For that reason, we want to explore the relationship between working hours and prosperity. When considering such differences in prosperity, a natural question is: who works more, people in richer countries like Norway or in poorer ones like Myanmar? By looking at the second graph (scatter plot), the answer is clear: people who work more hours usually come from poorer countries. There is also an insight: people from more populous countries usually work

more than less populated countries this may be due to Involution. To better understand the relationship between continents and working hours, we introduced two additional graphs. From that, we see that people from Asia usually work more over the years; people from South America, North America, Oceania, and Europe tend to work less over the years. Combined with the fourth graph, we see that people from poorer areas do not mean to work more, for example people from South America tend to work less over the year.

Team Member Breakdown:

Xin Lin

- Attend every meeting and generated the project idea
- Cleaned data set
- Coded the first graph (map)
- Helped debugging and wrote report

Bixin Zhang

- Attend every meeting and generated the project idea
- Searched and cleaned data set
- Coded the first graph (map)
- Sketched the visualization example

Karen Kuang

- Attend every meeting and generated the project idea
- Cleaned the data set
- Coded the second, third, and fourth graph (scatter and plot graphs)
- Wrote report mostly

Overall, we spent the two milestones brainstorming possible topics and visualizations, finding and cleaning datasets, discussing tasks, and sketching design. Initially, we just planned to have only two graphs (world map and scatter plot), we decided that Bixin and Xin both work for the first graph since it may be more difficult than the second graph. While Karen is more passionate with the project idea so she introduced two new graphs which makes our project more interesting and comprehensive. Finally, the first graph took around 25 hours, the second graph took around 20 hours, the third and fourth graph took around 4 hours.