

EPFL

Personality Traits Across the World

COM-480: Data Visualization Process Book

EPFL, Spring 2021

Group name: "Slackers"

Furkan Karakaş
Alexandre Luster
Barış Sevilmiş



Path

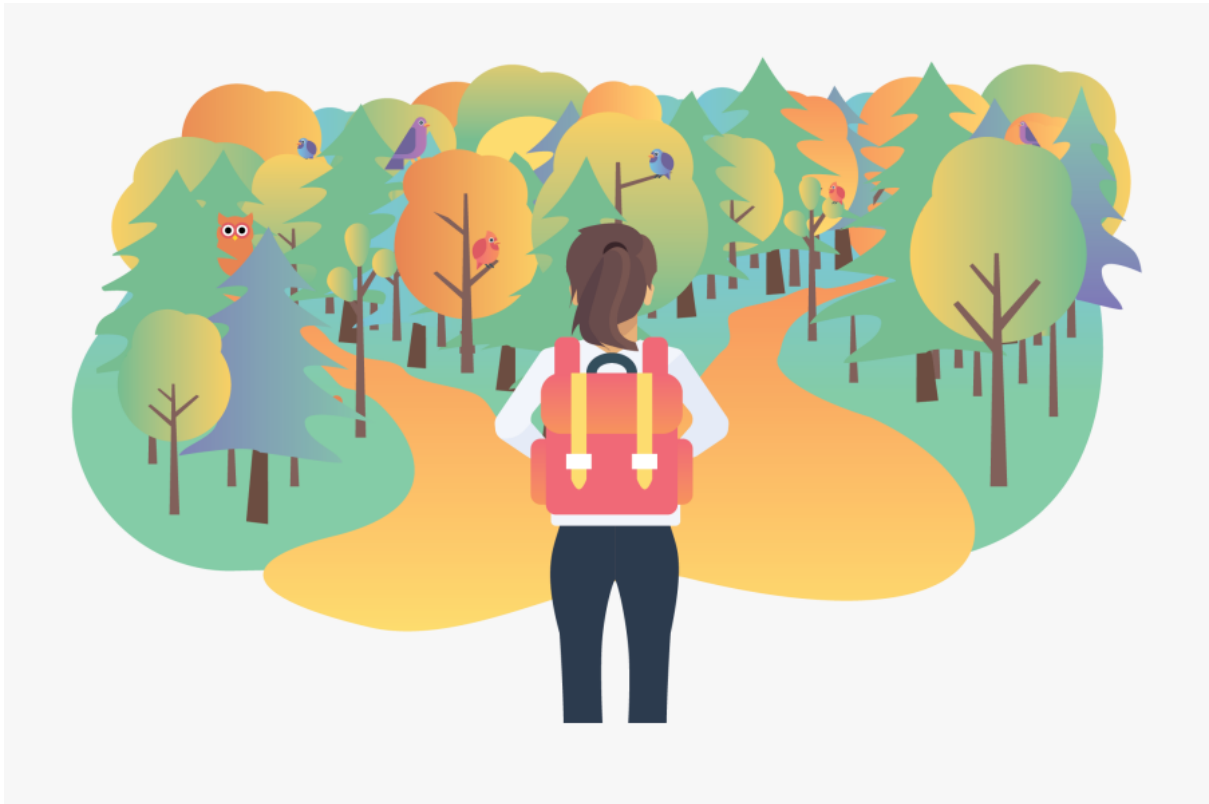


Figure 1: Alternative paths

There exist many alternative paths to design and implement a website, in which visualization of a specific dataset can take place. HTML, CSS and JavaScript are bread and butter in a web development project. Since the aim of this course is to be accustomed with HTML, CSS and JavaScript, we did not want to use heavy back-end or front-end frameworks such as Node.js, React.js, Angular.js etc. Other than that, there exist many JavaScript libraries to choose from, and some of the choices are crucial depending on the data.

D3 library was among the first of our choices, since it does provide very useful world map visualizations. Our map choice was based on a few criteria, which are mentioned in the Challenges & Design Decisions section. Shortly, our priority was to demonstrate the whole content of the map at once, whereas the map would have a visually appealing sight. **D3-legend** library is used as well, where the legend is added to the map. With the legend, users could see the meaning of the corresponding colors, in which data visualization becomes more self explanatory.

Crossfilter library was used to filter out the data. There are two filtering operations that can take place. Age filtering is the first option, in which users could choose minimum and maximum age limits from the slider. With this optionality, the world map will be colored only according to the age limitations. Sex filtering is the second option, in which male, female or both are the choices. Map will be again colored according to the chosen sex. Both filtering could be used simultaneously.

TopoJSON library was used to encode the topology in our project. This library allows the geometries of countries, lines between them and their coordinates to be stored efficiently in the same file. From the loaded raw data, it provides API such as `topoJSONdata.objects.countries` to efficiently access the countries on the world map.

Challenges & Design Decisions

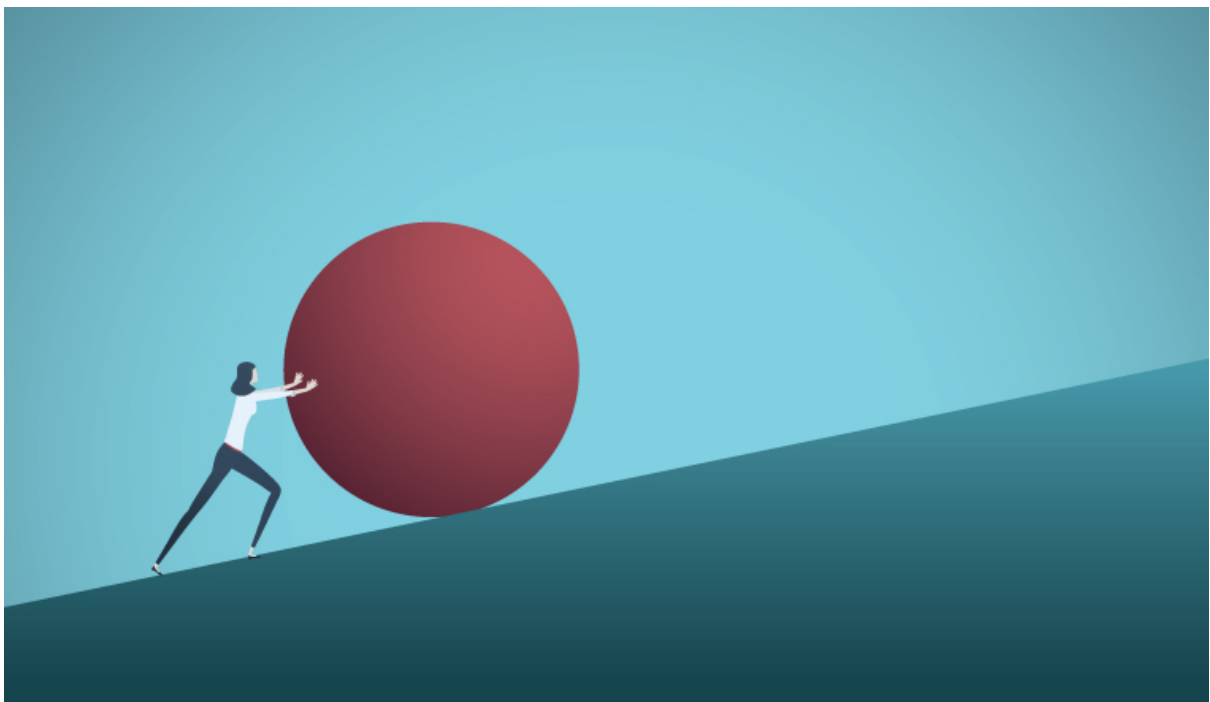


Figure 2: Challenges on the path

There were some of challenges and design decisions that we faced during the implementation:

Normalizing the scores

Instead of having a fixed color scale, we fix a relative color scale based on the current selection. The mean score in the select trait (or mean of means if several traits are

selected) is computed for each country, and these scores are normalized to the range [0,1]. This is important so that differences between countries are more visible. For example, we wanted the definition of "high" openness to be the highest mean score, not the unrealistic score of 1. This is also important because "low" and "high" carries a different meaning for each trait: in the dataset, the mean score for openness is 0.734, but only 0.574 for neuroticism. Hence, it is not really a good idea to put them to the same scale.

Another reason why we wanted to use **relative** scores rather than **absolute** ones is that in the world map, the mean scores of different countries are usually close to each other. Hence, that would make every country look the same with a similar color.



Figure 3: World map design

Map selection

There are various map choices to be used, where 3D options are available as well. Our choice was to use a simple but convenient 2D map. Although a 3D map may provide a decent visualization, it would require the user to interact with the map only through rotation, which may decrease the user satisfaction. In addition, the 3D map does not provide a total appearance of every world country, which does cause an issue in visualization as the purpose of the project would be unfulfilled.

Traits

Traits in this project are absolute and users would have to know about these traits for the project to become interesting. Instead of putting a link for users to read direct descriptions, or to provide long explanations directly, pages with corresponding trait images in background are ensured. These pages do contain some famous sayings about these traits as well as short dictionary definitions to enable users to understand the basics of these traits. Famous quotes may evoke interest in users towards the subject in addition to the descriptive backgrounds.

Sample questions

Lastly, some sample questions were provided at the last page for users to understand the dataset at a better level. This part is quite simplistic, as it contains a positive and negative impact question for each trait making up to 10 questions each with 5 answers. Instead of providing a background as in traits, a simple white background is utilized. Each question is completed with a corresponding image to form a visual understanding of the topic as well.

Changes from the First Milestone

In the first milestone, we mostly did data visualization in the Jupyter notebook in order to gain insightful information from the dataset. We were not sure which aspects of the data to include in our final product since we do not have a large dataset as the number of respondents in specific countries other than Canada, the United States etc. is relatively small.

At first glance, we had in mind to investigate the personality traits separately in the map. Later, we decided to combine different personality traits and we wanted to investigate the traits jointly. In Figure 1, you can see the initial layout of the map that we wanted to use.

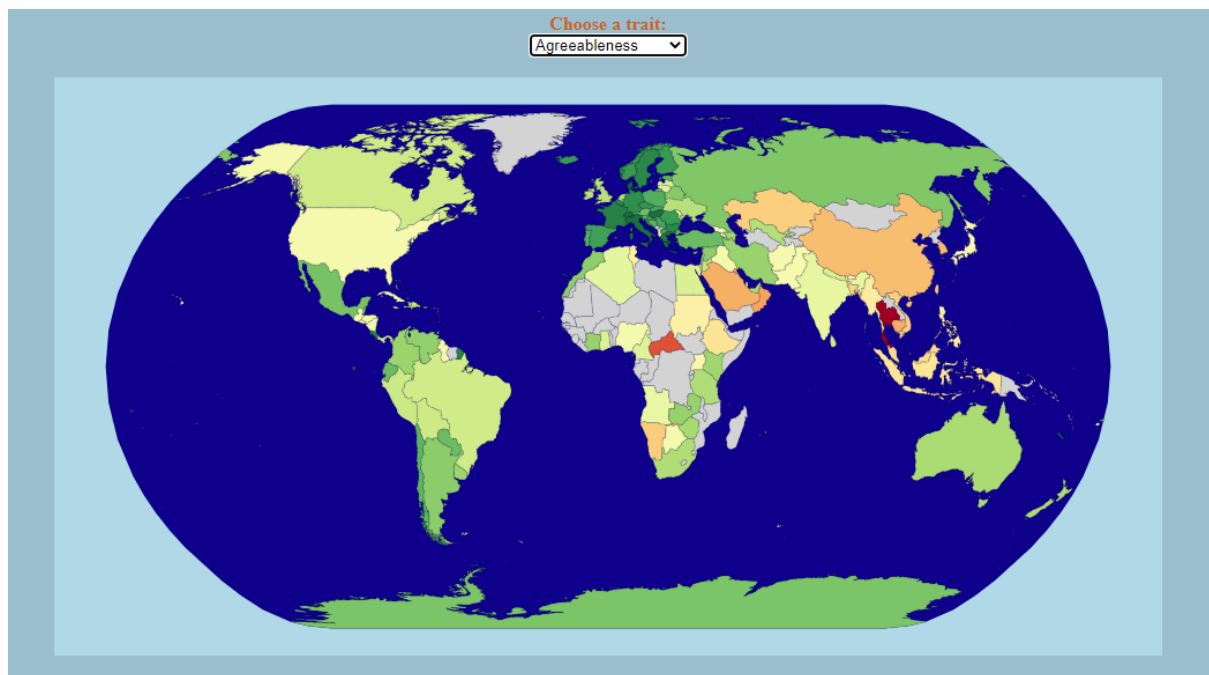


Figure 4: Initial layout of the website

In the first milestone, we wanted to include histograms to visualize the distribution of the score of a specific trait within a country. We added this histogram for discrete bins for selected combinations of traits. However, as we mentioned above, this distribution will not be very useful for the countries with few participants since the number of respondents are quite low. Other than this histogram, we decided to include bar charts where we plot the average scores of the traits with respect to four age bins "teenagers (younger than 20)", "young adults (between 20 and 39)", "middle-aged adults (between 40 and 59)" and "seniors (60 and over)". We also had in mind to put the histogram with an arrow right next to the country. However, we decided to include it in its own section so as not to crowd the main section too much.

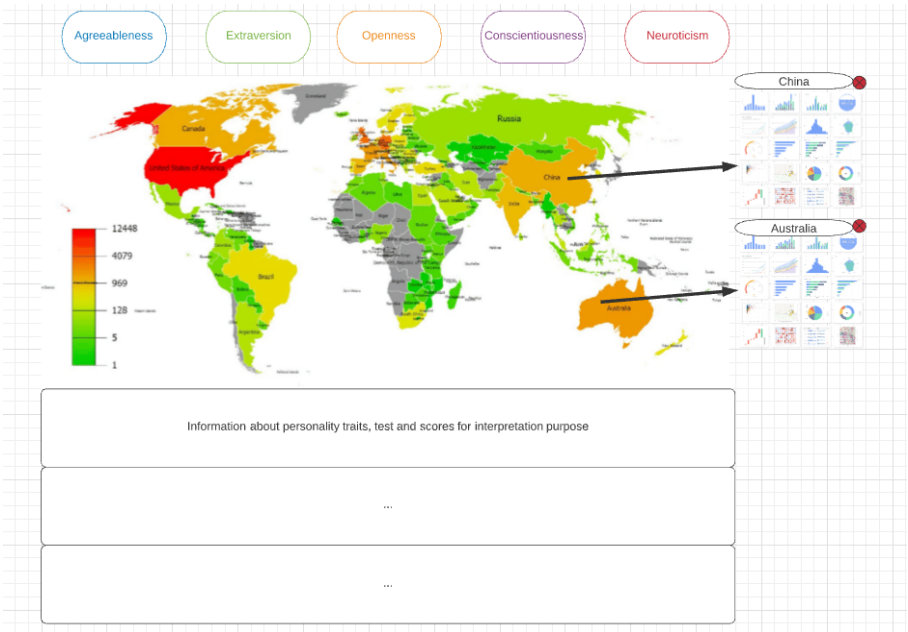


Figure 5: The initial design idea of the project

Peer Assessment



Figure 6: Peer Review

We were constantly in communication with each other during the implementation of the project. Whenever we faced some problems on the implementation side or we were not sure which API of the D3.js framework we were supposed to use in the website, we were able to consult each other so that other team members could point out useful documentation or tutorials on the web how the work could be done. In terms of design, we laid out different ideas on the table which would look better and more appealing to the user of the end product. We discussed several possible solutions, and we chose the ones that would appeal the most to the end product users. In terms of implementation, we had four major components in the website:

1. integration of the world map to the raw CSV data in D3.js framework
2. implementation of the bar chart in the D3.js framework
3. making the website more appealing to the end users by writing Cascading Style Sheets (CSS)
4. capturing the events such as click, change, mouse hover, etc. of the HTML elements in JavaScript and defining the appropriate logic
5. preparing the process book and a short video

Alexandre was responsible for the first two points, and **Bariş** mostly took care of doing the third part, whereas **Furkan** was focusing on the fourth part above. Nevertheless, these tasks cannot be easily separated from each other. For example, Furkan also worked on correcting minor mistakes in the world map and bar charts. In general, there was involvement from every team member in every part of the project, and we were able to manage this process effectively.

For the fifth part in the list of the tasks, everyone worked on preparing the final process book and our team member **Furkan** recorded the short, two-minute-long video.