Simon Eckerstorfer   11911424   WS2021/22

COVID-19 Dashboard - Report

Find my deployed web-app [here](#).

Motivation

The task at hand was creating a Dashboard with visualizations for a COVID-19 dataset. The dashboard is intended to be used by a public health expert and a government official with certain use cases. The former focuses on the COVID vaccine and its effects on infection and mortality rates, while the latter is interested in the correlation between infection rates and ICU workload on one hand, and certain risk factors (e.g. cardiovascular death rate) on the other.
Both users need a quick overview of the data and means to inspect certain aspects in more detail.
The used dataset contains a huge amount of different parameters and gets updated very frequently. However, for certain aspects (like ICU workload), the data gets thinner and is non-existent for many countries. Thus, the decision was made to focus on Europe, where data is available for almost all countries.

I decided to use my dashboard design 1 from A4, which contains a Scatter Plot, a Stacked Bar Chart, a bivariate Choropleth Map of Europe and a Bar Chart. The first two visualizations focus on the task of the public health expert, while the other two are aimed to fulfill tasks of the government official. However, the tasks also overlap and all visualizations are of interest for both users (e.g. infection rate helps both users).
The design was chosen as it has superior encoding of the essential aspects in comparison with dashboard 2. It carries more information, while also delivering the information in a more user-friendly and intuitive way.

Prototyping / Design Process

Prototyping my design from A4 resulted in a web-app that enables users to explore COVID data for European countries in an intuitive way. The focus lays on vaccination/infection/mortality rates as well as ICU workload and diverse risk factors for an infection. By default, a user views data for all countries at once, giving a general overview on the Scatter Plot and the Choropleth Map. In accordance with the principle of overview and detail, a user can select a country, for which data on vaccination rates and risk factors appears on the 2 Bar Charts, respectively. Another level of detail can be

entered by hovering the mouse over data points - detailed tooltips are implemented for every chart, displaying exact data and additional information.

Additionally, the user can also filter the data on the scatterplot by clicking the legend - this might be useful to focus either on infections or on deaths in detail, if the other aspect is not interesting for a user.

An aspect where the dashboard shines is spotting extreme values. It is very easy to tell which countries have the lowest/highest vaccination rates etc with the Scatter Plot. I discovered an interesting detail in that context: Gibraltar has a vaccination rate of 123%, as many non-residents were vaccinated in the country.
The bivariate color scheme of the Choropleth Map also lets users spot extremes. It is less concise (as only 9 colors are available), but the data points can be linked to the countries quicker on the map view. Also, bigger-scale trends can be spotted, like lower infection rates in the East at the time of writing this report. Central and West Europe seems to be a hotspot for infection right now.

The Stacked Bar Chart on the top-right enables users to compare vaccination rates of selected countries with an European Average, but also a selected comparison country. It also lets users inspect the rate of people with one, two or three vaccine doses respectively. This is a point where very wide variations can be spotted between the countries: For example, Portugal has nearly 50% of people with booster shots, while only 3.7% received their third shot in Bosnia and Herzegovina.
Users can also inspect the rate of unvaccinated people via the tooltips.

Staying with the comparison between Portugal and Bosnia and Herzegovina, inspecting the risk factor Bar Chart for each country respectively, users can spot that both countries are among the highest in Europe when it comes to Diabetes prevalence. The tooltips tell us that near-to 10% of the population of both countries have diabetes. However, Bosnia has a 50% bigger share of their people smoking, and nearly triple the cardiovascular death rate.
In general, the risk factors are normalized with the biggest observed value - meaning the percentages shown are always relative to the extreme values.
This was an aspect where I struggled most while implementing. I tried various approaches to show the risk factors on a common scale, while also keeping the absolute values in mind. I landed on the above-mentioned approach of normalizing the different risk values by dividing with the biggest observed and showing them on a common scale from 0 to 1. By hovering over the data, users can inspect the concrete values for each bar instead of the normalized factor. A change compared to the design of A4 was ditching the percent value labels above the Bars. I had them on my first

web-prototype, but most of the people I showed it to were confused by that. They thought the percentage label (e.g. above "smokers") meant that that percentage of the population smokes. This, however, is a misinterpretation, as the percentage only shows the correlation to the biggest observed factor. By hiding the percentage label, users are forced to look at the scale before arriving at a conclusion - showing them that the value is a normalized factor and giving them the ability to observe concrete values with mentioned tooltips.

A different problem I had with the risk factors was the value of the Human Development Index. While all other values theoretically increase risk of an infection when they are higher (diabetes, smoking, …), a higher HDI should decrease that risk. To keep the mental model of higher bars - more risk, I inverted the HDI, so higher bars mean lower HDI and those (theoretically) higher risk of infection.

Comparing the current design of my visualization with the paper prototype from A4, only a few major changes can be found. The biggest one is having a "global" focus country that can be selected in multiple ways (Dropdown at the top, clicking on scatter datapoint, clicking on country in map). During the prototyping as a web-app, I found it to be confusing to have different countries selected for the 2 views on top and the 2 views on the bottom respectively, especially with the additional selection for a "comparison country" in the Stacked Bar Chart view. To preserve the mental model of a user, the change to have only a single country in focus at a time seemed to be the most intuitive.

Furthermore, I decided to focus on Europe in the whole dashboard, instead of only showing Europe on the map while having data points for all countries in the other views. This change also improved the perception of the dashboard, as the views seem more linked when showing the same set of countries.

Another point of confusion I got as feedback was the point of time the dashboard focuses on and whether it can be selected. As none of the views focus on time ranges (e.g. no line charts), but only use data for a single point in time, the decision was made to have a static date for which the data is shown. This date is communicated clearly on top of the page. For this assignment, I statically use the most recent dataset from [owid-covid-latest](). However, an enhancement to dynamically pull the most recent data when opening the web-app and displaying the corresponding date would be easy to implement

Implementation Details

A great deal of implementation details were discussed in the previous section, as results of prototyping are closely related to the final implementation.

When it comes to the tasks that are solved, the two views on top are aimed at giving insights about vaccination rates in correlation with the infection and mortality rate respectively. Like mentioned above, different levels of detail can be entered with the different views and the tooltips. The Bar Chart on the bottom right solves the task of giving insight into risk factors, while the Choropleth Map enables users to correlate disease rates with ICU workload.

A technical decision I made during the implementation was the abundant usage of information points encoded with question mark symbols. I feel like they are an appropriate approach to explain certain aspects of the dashboard to new users, while also not overloading the view with information not interesting to trained users.
An example is the explanation of the Cohropleth Map´s legend. The bivariate color scheme could be confusing at first, but is easily understandable with the tooltip-text elaborating on it.
Furthermore, details of the data were also presented in the mentioned way. The exact meanings of the different risk factors are explained via the question mark icon and a tooltip.

Another technical decision I made was using the top-right view as a comparison view exclusively. No single country´s vaccination data is shown, but always a comparison of two. If no comparison country is selected (in addition to the general focus country), the European Average is taken to compare to.

When using the Scatter Plot to discover data, I realised that finding the two points for a country (infection and death rate) was difficult while not selecting it as a focus county. In order to tackle that problem, I created a popout effect for the two correlating points when hovering over one of them. They increase in size and receive a dark outline.

Discussion

During the process of creating the dashboard through assignment 4 and 5, I learned how valuable it is to create concrete tasks for the visualization. With the huge amount of different data points in the covid dataset, it would have been nearly impossible to create a concise dashboard that does not "fizzle out" into too many areas, while not focusing at any of them.
I also see it as central to look at the dataset at hand before coming up with designs, as some aspects were not represented in the data properly (like ICU workload for most countries outside of Europe).
The iterative prototyping and expansion of design space (mostly during assignment 4, but also while implementing the web app) was essential for the visualization as a whole.

If I had stuck with the first ideas, I would not have arrived anywhere near the current outcome. The importance of iterative design was mentioned often during the lecture and I felt it myself during this assignment.

A thing I would do differently next time is asking users for feedback during the design phase. This way, going the extra mile while implementing could have been prevented for aspects like the percentage labels over the Bars for the risk factors. Also, expanding the space of possible visualizations even further would have been a way to enhance the design.

The users are clearly able to solve the tasks that I defined in assignment 4. How that can be done has been elaborated in this report already. In a few words, tasks 1 and 2 are linked to the top views, tasks 3 can be done with the bar chart and task 4 is solved with the map of Europe.

Conclusion

In summary, the dashboard visualises different aspects of a covid dataset in a user-friendly and intuitive way. It enables users to inspect vaccination, infection and mortality rates, as well as showing ICU workload and diverse risk factors for a covid infection.

It focuses on countries in Europe and a single (theoretically most up-to-date) timepoint. Furthermore, a single country can be selected as a focus country. Those three design decisions changed after assignment 4 to enhance the how intuitive the visualization is and its consistency between different views.

Some general improvements have been made during the implementation phase, like the extensive usage of tooltips and explanation texts behind info-icons. Those things should enable new users to comprehend the dataset correctly on one hand, and give an option to see details for experienced users on the other.

I learned the importance of the whole design pipeline first hand. Inspecting the dataset closely, defining user tasks concisely and creating multiple paper prototypes were important cornerstones of the process.

In conclusion, I am proud of the design and visual encodings/interactions I came up with. I learned a lot during the design process and implementation and will keep those things in mind during future projects.