

Necklace Maps for COVID-19 Visualization

Mary Abikoye (temiabik@student.ubc.ca), Minglong Li (limnglng@cs.ubc.ca),
Jocelyn Minns (jminns@cs.ubc.ca)

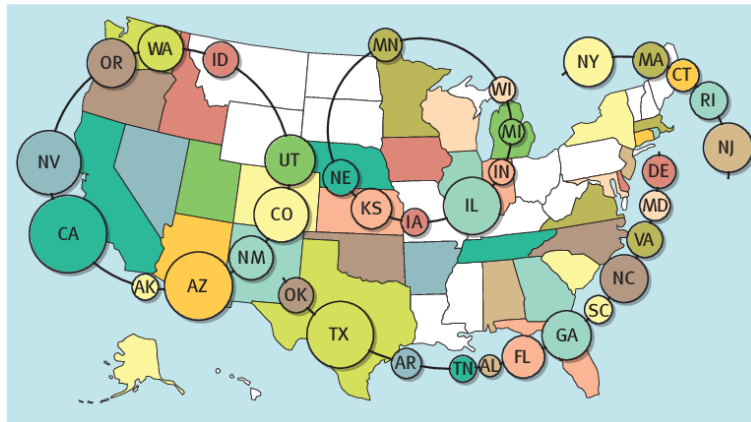


Figure 1

1. Introduction

Necklace maps [8], as shown in figure 1, are a form of visualization that combines elements of proportional symbols maps and boundary labeling. The symbols are placed inside the corresponding intervals. The projection maps each region of the input to a contiguous interval on the necklace in such a way that the interval captures the global location of the region with respect to the necklace.

The advantage of necklace maps over the much more common choropleth maps is that they give more visual weight to the quantitative variable, rather than just the geographic area. This can be preferred when we are dealing with epidemiological data since large geographic areas may have a sparse population, but that large area holds more visual weight than a densely populated city. When users see these choropleth maps, they can easily misunderstand the data to assume the situation is better or worse than the quantitative variable indicates.

Since the rise of the COVID-19 pandemic in early 2020, data about factors in the spread and response to the disease have become commonplace in our daily lives. However, many visualizations of this data continue to favour choropleth maps. For this project we intend to implement necklace maps to visualize three major factors of the COVID-19 pandemic: confirmed COVID-19 cases, vaccination rates and the stringency index.

Our project focuses on examining two research questions. **R1: Does vaccination rate affect COVID-19 confirmed cases?** and **R2: Does stringency index affect COVID-19 confirmed cases?**

2. Related Work

Choropleth maps are a thematic map that is well suited to show quantitative attributes of geographic data [7], however there can be significant drawbacks to relying on this method. When the quantitative attribute is focused on the population, these maps can overemphasize large geographic areas regardless of their sparse population due to their strong visual weight. Different thematic maps have been proposed over the years and while they may each have their own advantages, no clear alternative has been identified [1]. Additionally, choropleth maps can be ill suited for the task of observing changes over time in the data. In a study on animated choropleth maps, the phenomenon of change blindness hindered the user from perceiving changes in the data [3].

Epidemiological data is commonly visualized using choropleth maps, though careful consideration must be given to ensure these maps can be useful to epidemiologists. Variations in classifying the data must be considered to accurately inform those relying on these visualizations [2]. The recent COVID-19 pandemic has seen a surge in visualization approaches [9] including a heavy reliance on choropleth maps, however the limitations of these maps must be considered when creating visualizations to avoid potential misinformation [6].

Necklace maps [8] are a proposed alternative to the choropleth maps. This method combines proportional symbol maps and boundary labeling to create a one-dimensional curve containing scaled symbols corresponding to a quantitative attribute for a particular region. This can help visualize data that is not proportional to region sizes.

3. Data and Task Abstraction

3.1 Datasets

Our World in Data provides publicly available datasets on the COVID-19 pandemic [5]. Specifically, we want to look at three attributes: Confirmed COVID-19 cases, reported vaccination rates and the stringency index.

The confirmed COVID-19 case data gives a daily update of the confirmed case count per million people by country. This data is updated daily for many countries, but reported cases can vary depending on how the area chooses to report their cases on weekends or holidays. To mitigate this reporting inconsistency, we will use the smooth data over the 7-day period.

Vaccination rates provide the reported percentage of people fully vaccinated by country. We will focus on the data of people partially and fully vaccinated.

The stringency index looks at the response policies regarding the COVID-19 pandemic. The Oxford Coronavirus Government Response Tracker (OxCGRT) [6] project calculates this stringency index of countries using 20 indicators divided into 5 categories: Containment and closure policies, Economic policies, Health system policies, Vaccination policies, and Miscellaneous policies. The final stringency index is a number between 0 and 100 where 100 is the strictest.

3.2 Tasks

The main tasks we hope to accomplish will involve visualizing the datasets listed above. Some tasks we would like to achieve with our implementation of the necklace map algorithm are:

- Discovering trends by comparing symbols on different necklaces. e.g. Is there a noticeable pattern in the case number necklace and the stringency index necklace?
- Look up a specific country's data. e.g. What are the current COVID-19 case numbers in Canada?
- Compare different geographic areas: e.g. How does the vaccination rate necklace of Europe compare to the vaccination rate necklace of North America?

4. Scope

In this section we define the relationship between our work and the original necklace map paper [8]. In addition to generating a visually pleasant necklace of glyphs representing quantitative data, our work extends the capabilities of the necklace map mostly in three aspects: interactive environment, dynamic necklace generation and multiple-necklace display.

Interactive environment The original necklace map is a static visualization, however, we would like to present an interactive map to let the user to freely explore any set of countries with available data. Also, since for each country we have three sets of quantitative data (case numbers, vaccination rate and policy stringency index), the user should also be able to select any or all three datasets to visualize.

Dynamic necklace generation Speckmann and Verbeek [8] generated the necklace manually first and then placed the glyphs onto calculated spots. However, in our case, the size and position of the necklace will be determined automatically based on the selected countries. The countries selected will be divided into multiple clusters based on quantity selected and proximity. For each cluster, the corresponding necklace(s) will be generated.

Multiple-necklace display In order to visualize a maximum of three quantities per country at the same time, we plan to instantiate one necklace per quantity. Therefore, if the user selects to visualize all three datasets, there will be three necklaces at the same time, as opposed to one fixed necklace in [8].

Below we describe the ideal final program and how the user would interact with the tool. It is possible that we cannot realize all the features described below given the time constraints, nonetheless we aim to have the users explore the dataset with the following steps:

- Upon starting the software, a map of all countries in the world appears, all in grey to indicate unselected status. This map should allow zooming in and out, in addition to translational movement.
- The user would then click to select the countries and what datasets they want to visualize. For each dataset they select, a necklace visualizing the data of interest would pop up in concentric circles.
- All selected countries are color coded corresponding to its symbol color on the necklace. If the user wants to unselect certain countries, they would simply click those countries again.

Some additional features might be added if time allows, such as the ability to select countries by searching for it with text input.

5. Solution and Current Status

In this section we discuss our implementation decisions and details, as well as the current status of the project. We break down our approach to the problem into the following aspects.

Language and Packages In order to implement the Necklace map algorithm, it is vital that we select the appropriate programming language and packages. After experimenting with some of the recommended options, we decided to go with python due to our familiarity with it. To build an interactive visualization, we are using the recommended Altair package. In addition, Geopandas is utilized for its easy access to geographical information.

Necklace Generation Automatic generation of necklaces based on user selection is something that was not included in the necklace paper. We devised an algorithm for this purpose as follow:

1. Divide selected countries into a few clusters. The idea is that each cluster will have its own necklace(s). Currently the countries are clustered based on their continents, however, this means that countries far apart in large continents could be clustered together, resulting in undesirable necklaces. In the next iteration, we plan to use a distance-based clustering method to combat this effect.
2. For each cluster, find the convex hull encompassing all countries within the cluster. This step can be accomplished with geographical data provided by Geopandas.
3. Based on the convex hull and the geometric center of the convex hull, we can determine the appropriate radius and center of the necklace. In the case where the user selects to visualize multiple datasets, the output of this step would be a center shared by all necklaces and the radii of the necklaces (up to three radii because one necklace per

dataset). The radii should be equally spaced-out with the smallest one being at least the same as the radius of the convex hull.

As of now, we finished with step 2 and are still fine-tuning some parameters in step 3. An example of the resulting convex hull is shown in figure 2.



Figure 2

Marker Generation In addition to necklace generation, marker generation also requires a lot of effort. We further divide this task into two stages: marker instantiation and placement. The idea is to figure out how to create markers proportional to a given dataset first and then place the markers correctly on the necklace.

Currently, we have been successful with the first stage. An example of visualizing countries with above average COVID-19 case numbers in Africa as a proportional symbol map can be found in figure 3. In that example, symbols are placed at the centroid of the polygon of each country. We plan to move forward concurrently with necklace generation by assuming a fixed necklace for this task.

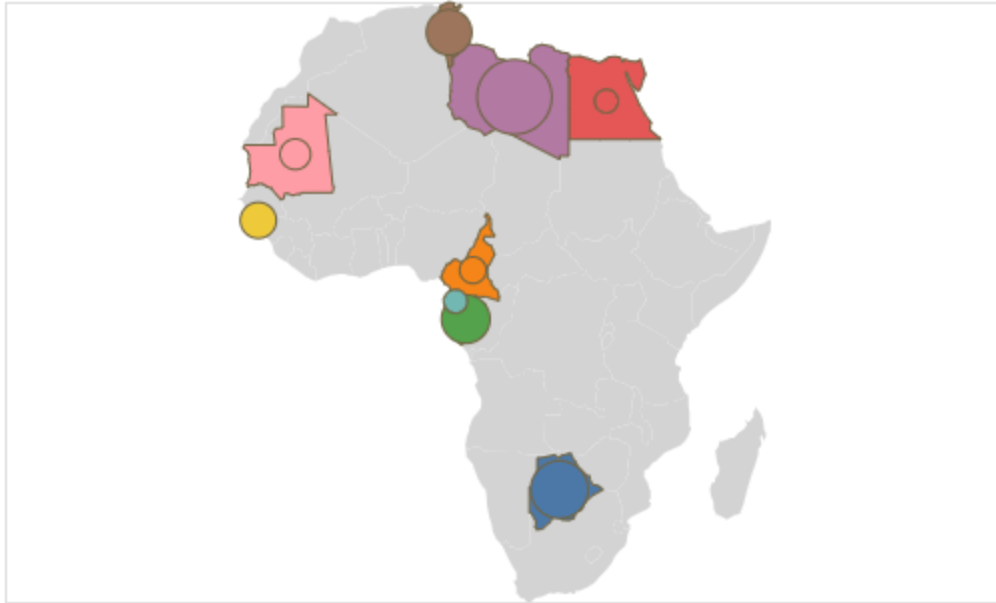


Figure 3

User Interaction This is the view the users have before the decision of chives are made. The view changes and displays different images or charts based on the user's choices. We are considering

1. Manipulating views like change in view over time. This is for the choice of duration of time needed for the maps. Select for the drop-down options. Zoom is used when there is a drop-down of a decided country or countries to view.
2. Facet
3. Reducing items and attributes

Decision is still being made on how the view should look like. For now we are focusing more on the Necklace maps.

We are also familiarizing ourselves with Altair so that we can build the interface. We started by implementing an interactive single-selection tool. By hovering over a single country, a textbox with all three quantities associated with the country will appear by the cursor, as illustrated in figure 4.

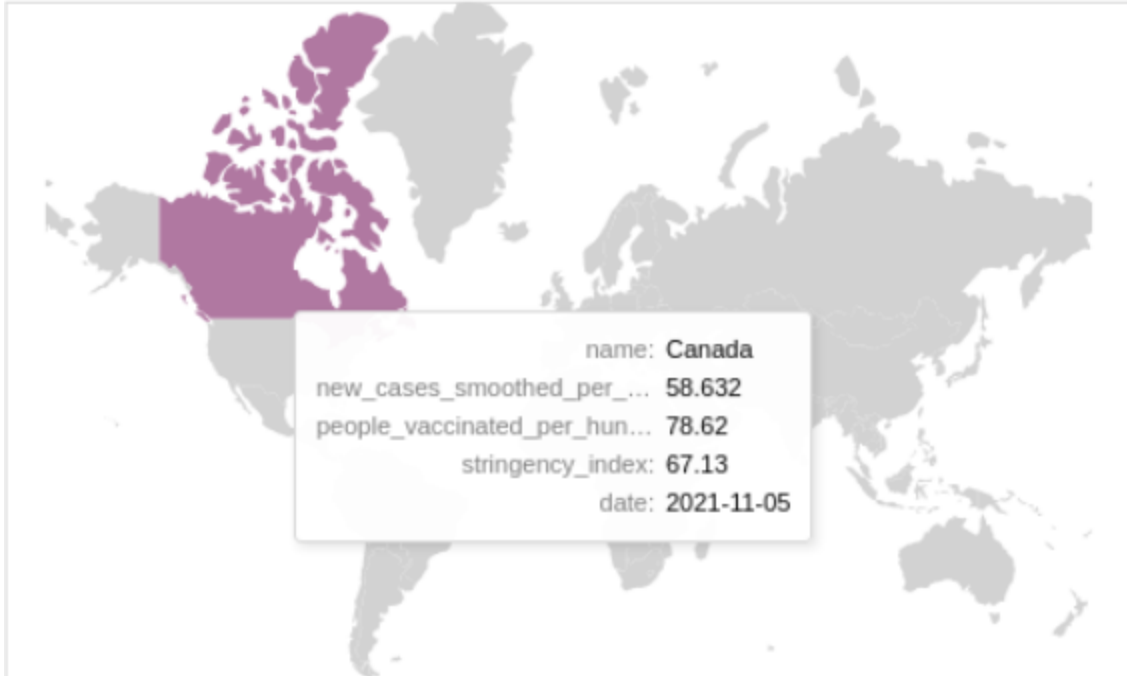


Figure 4

6. Milestones and Timeline

| Weeks | Tasks and Descriptions |
|-------|---|
| W1 | <ul style="list-style-type: none"> ● Write Project Proposal (Oct 21 noon) (all members - 4 hours each) |
| W2 | <ul style="list-style-type: none"> ● Literature Research (all members - 4 hours each) <ul style="list-style-type: none"> ○ Other map visualization techniques ○ Capabilities and learning curves of multiple tools ● Data Exploration (all members - 1 hours each) <ul style="list-style-type: none"> ○ Examine data available. Determine what filtering is required from existing data: https://ourworldindata.org/covid-stringency-index, https://ourworldindata.org/covid-cases, https://ourworldindata.org/covid-vaccinations |
| W3 | <ul style="list-style-type: none"> ● Programming Ramp-Up (all members - 6 hours each) <ul style="list-style-type: none"> ○ Become familiar with chosen programming tools ○ Explore capabilities and limitations of Altair ● Geometric Map Visualization (Jocelyn - 6 hours) <ul style="list-style-type: none"> ○ Visualize the world map ○ Link database with the application ○ Implement choropleth maps to explore compatibility of the geographic data and COVID-19 datasets |

- W4
 - Establish Necklace Algorithm Implementation (Minglong - 6 hours)
 - Determine convex hull for necklace shape
 - Plot proportional symbols given our dataset

 - W5
 - **Write Project Updates (Nov 16 3pm)** (all members - 4 hours each)
 - Discuss feedback
 - Reevaluate scope of the project
 - **Peer Reviews (Nov 17 in class)**

 - W6
 - **Post-update Meetings (Nov 24 & other times that week TBD)**
 - Modification Based on Comments (all members - 1 hours each)
 - Make necessary modifications and adjust plans based on reviews from the class and the comments from the meeting
 - Complete Necklace Algorithm Implementation (all members - 6 hours each)
 - Hard code pre-defined clusters of countries
 - Draw necklaces for each cluster and dataset

 - W7
 - Advanced Features (all members - 6 hours each)
 - Design and implement automatic generation of the necklaces
 - Enable user interaction to select countries and datasets

 - W8
 - Advanced Features Cont. (all members - 6 hours each)
 - Finish up previous features
 - Implement country search bar if time allows
 - Performance Optimization (all members - 2 hours each)
 - Latency optimization if time allows

 - W9
 - **Final Presentation (Dec 15 2-6pm)**
 - **Final Paper (Dec 17 8pm)**
-

7. Reference

- [1] Lonni Besançon, Matthew Cooper, Anders Ynnerman, and Frederic Vernier. An evaluation of visualization methods for population statistics based on choropleth maps. arXiv preprint arXiv:2005.00324, 2020.
- [2] Cynthia A Brewer and Linda Pickle. Evaluation of methods for classifying epidemiological data on choropleth maps in series. *Annals of the Association of American Geographers*, 92(4):662–681, 2002.
- [3] Carolyn Fish, Kirk P Goldsberry, and Sarah Battersby. Change blindness in animated choropleth maps: An empirical study. *Cartography and Geographic Information Science*, 38(4):350–362, 2011.
- [4] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, et al. A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 5(4):529–538, 2021.

- [5] Lucas Rod'es-Guirao Cameron Appel Charlie Giattino Esteban Ortiz- Ospina Joe Hasell Bobbie Macdonald Diana Beltekian Hannah Ritchie, Edouard Mathieu and Max Roser. Coronavirus pandemic (covid-19). Our World in Data, 2020. <https://ourworldindata.org/coronavirus>.
- [6] Carsten Juergens. Trustworthy covid-19 mapping: Geo-spatial data literacy aspects of choropleth maps. *KN-journal of cartography and geographic information*, 70(4):155–161, 2020.
- [7] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014.
- [8] Bettina Speckmann and Kevin Verbeek. Necklace maps. *IEEE Trans. Vis. Comput. Graph.*, 16(6):881–889, 2010.
- [9] Yixuan Zhang, Yifan Sun, Lace Padilla, Sumit Barua, Enrico Bertini, and Andrea G Parker. Mapping the landscape of covid-19 crisis visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–23, 2021.