Visualizing Mobility and COVID-19

Lily Bryant, Frank Yu, James Yoo

{labryant,frankyu,yoo}@cs.ubc.ca

The University of British Columbia

1 Introduction

The COVID-19 Pandemic is without a doubt the largest event in recent memory to have disrupted the lives of billions of people on a global scale. Many across the globe have changed their daily routines as a response to this event; within our society, people have changed how they work, play, and otherwise spend their time. Advances in technology and the ubiquity of mobile devices have accelerated the collection of data on a massive scale, especially with respect to how the mobility of individuals has changed across the course of the pandemic.

Although the scale of mobility data collected throughout the pandemic is massive, it requires an expensive amount of pre-processing and domain-specific modelling that makes interpretation difficult for a layperson. In order for this data to be useful, it needs to be presented in a way that explains trends in mobility and presents the data in a way that is accessible to a layperson. Tasks such as comparing the changes in mobility across regions, or observing a change in a behaviour over time during the pandemic may be of interest to a layperson.

This work aims to create an *interactive explainer* that leverages a number of visualizations that help to present mobility data collected during the COVID-19 Pandemic in a way that makes trends easily observable. Our interactive explainer will be focused on visualizing changes in mobility across Canada, with a particular focus on the province of British Columbia.

2 Background

Mobility is defined as the study that describes how individual humans move within a network or system [8]. Various infection mitigation measures introduced throughout the pandemic have led to marked changes in the mobility of individuals in their communities. The reduction of activities in community locations that can be categorized under retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential have been the focus of efforts to mitigate community transmission of COVID-19. Our work will explore the trends in mobility across these categories.

Taking a focused approach to data from Canada, and British Columbia in particular, also affords us the option of framing our visualizations in the context of government-mandated shutdowns. A specific example could be visualizing how mobility changed with respect to the phases of British Columbia's Restart Plan [6].

3 Related Work

There have been a large number of visualizations performed with COVID-19 data. A seminal example is the COVID-19 Dashboard by the Center for Systems Science and Engineering at Johns Hopkins University [3]. We explore two visualizations that interactively display COVID-19 data in Canada (3.1), and display mobility data across a number of categories (3.2). Related work in visualizing mobility that is orthogonal to COVID-19 is explored in (3.3). Additionally, we explore the viability of using a map that encodes Bayesian surprise and expectation (3.4). We also review an article about the dangers of irresponsible COVID-19 visualizations and discuss how our project will avoid them (3.5).

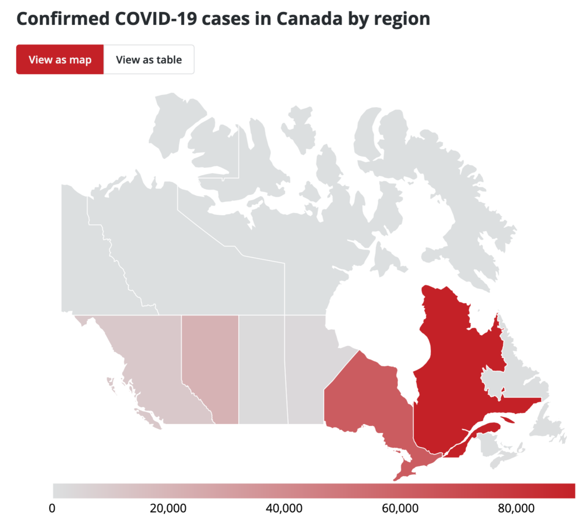

Fig. 1. A choropleth map from the CBC Coronavirus Tracker

3.1 CBC Coronavirus Tracker

Fig. 1 shows a choropleth map from the CBC Coronavirus Tracker [1]. The tracker exists in the form of a web page that contains a number of visualizations. It enables end-users to perform tasks such as viewing the number of new cases, cumulative cases, and death tolls in Canadian provinces across a range of months. This style of presentation applied to mobility data could be something we explore in our project.

3.2 Google COVID-19 Mobility Reports

Fig.2 is an example of a sample PDF that users can generate from the Google COVID-19 Mobility Reports [2]. It shows a small multiple of line charts that visualize the changes from baseline mobility levels across a number of activity categories.

Although the data provided is rich, there are a number of pitfalls. These visualizations do not enable a pairwise or one-to-many comparison between regions for which data was collected. There is also the added complexity of users needing to explicitly download these PDF visualizations on per-region basis. Our project plans to remedy this by providing visualizations for which regional data is displayed on-demand, enabling pairwise and one-to-many comparisons without the user needing to collate an ensemble of downloaded visualizations themselves.

3.3 Mobility Sensing

The concept of using mobile sensing devices to collect data for visualizing mobility is not new [10]. Smartphones and other mobile devices are pervasive enough such that they are able to provide insights into

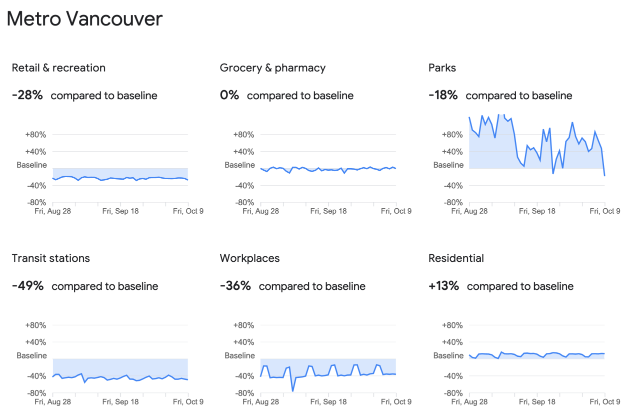

Fig. 2. A small multiple composed of line charts tracking changes in mobility for activity categories in Metro Vancouver

the mobility of individuals who use and interact with them. They also provide an unobtrusive way to collect data on a massive scale [10]. Using data collected from sensors embedded in mobile devices, it was possible to generate detailed line graphs of the number of vehicles that passed through an area at a certain time. This approach of taking data collected from mobile sensing devices and creating visualizations is something that we can explore in visualizing how COVID-19 has changed mobility.

3.4 Surprise Maps

While displaying raw data on a choropleth map can be helpful, it has limitations when it comes to focusing attention on regions of with surprising data. In [7], the authors introduce the *Surprise Map*, a visualization for displaying geographical data that addresses the issue of surprising regional data being hidden from view. In this visualization, the authors use a new form of derived data that shows how surprising/expected a particular value is. This visualization is something that we should consider when designing a choropleth map.

3.5 Responsible COVID-19 Visualizations

Misinformation about COVID-19 is incredibly widespread. Visualizations are widely used to disseminate information that may be incorrect or even harmful. For example, data scientists and statisticians have been producing their own models and and visualizations that attempt to draw conclusions about the spread of the virus [9]. There has also been a marked focus on visualizing case numbers without providing additional context.

In creating responsible visualizations, we will not be visualizing or generating any models of how COVID-19 progresses in a region. We do not have the necessary background in epidemiology to do so, and proceeding without further domain knowledge in this space would be unethical. We will also not be focusing on the visualization of case numbers for our visualizations. Instead, our visualizations will be centered around data relating to the mobility of individuals in a regional basis.

4 Data

We plan to use the COVID-19 Community Mobility Reports dataset that is provided by Google. This is a static dataset that is downloaded from Google. The time period we are currently considering is from February 14, 2020 to October 17, 2020. We plan to update the end of this range until our final submission date. The dataset consists of 3077 item with a total of 8 attributes. The following six attributes represent the measured changes from baseline mobility for each location category encoded as an integer percentage:

- Retail & Recreation
- Grocery & Pharmacy
- Parks
- Transit stations
- Workplaces
- Residential

Canadian provinces and territories are categorical attributes that exist in this dataset. Additionally, our dataset demonstrates time-varying semantics with a ISO 8601 date being one of the 8 attributes.

5 Task Abstractions & Usage Scenarios

For our project, we will create an article that can be viewed by a layperson. We do not assume that our intended audience has specialized knowledge in epidemiology or are intimately familiar with interpreting complex visualizations. We expect our audience to participate in tasks which resemble some form of exploratory data analysis (EDA). Some tasks may include locating which months had the greatest change in mobility in provinces across Canada, or observing how the mobility associated with an activity changed over time.

To support these types of tasks, we will author a Distill [4] article, which can be categorized as a type of interactive explainer. This format is particularly well-suited for our audience for a number of reasons. They enable the presentation of data with explicit explanations that may guide and help the audience in understanding it, even if they are not domain-experts. Additionally, they enable the inclusion of interactive diagrams which allow the audience to explore the data in the context of the explanations provided in the article.

In (5.1), we describe a use-case in which users *consume* data presented in way that describes changes in mobility from a bird's-eye view for the whole of Canada; in (5.2), we discuss a use-case in which users filter specific regions for which to view mobility data, and are able to vary the amount of data displayed with respect to time; in (5.3), we detail a use-case where users make direct comparisons between the provincial and national level; and, in (5.4), we describe a use-case where users compare overall data trends with outliers and determine their potential cause based on contextual information.

5.1 Comparing geographical data by region

Since this project is centered around creating a interactive explainer, we anticipate that one use-case that users may engage in is observing the differences in mobility between different regions of Canada. This would fall into the task abstraction of users that want to *consume* data. For this task, we think that users will want to see how each category of mobility is affected in different regions and if there are any regional trends. This usage scenario is mostly supported in the visualization we see in *Fig. 1* (though with different data); however, we also want the user to be able to select the data they want to have visualized which is not featured in *Fig. 1*.

5.2 Comparing trends over time

A use-case that we have considered for users of our interactive explainer is comparing trends in regions over time. For example, a user in one province may want to see how mobility has changed relative to neighbouring provinces. Although this task appears similar to the one described in (5.1), we specify that users could want to engage in observing how the data changes with respect to time. This would fit into the task abstraction of users who wish to *consume* data that has already been processed by a tool. We can anticipate that this is a task that users may use to discover what effects their changes in mobility, relative to other regions, have had on the incidence of cases in their region. This usage scenario is currently not supported with the Google COVID-19 Mobility Reports (*Fig. 2*), as it is limited to displaying data for only one region at a time.

5.3 Comparing provincial versus national trends

In a similar vein to both (5.1) and (5.2), users may also want to make a comparison between provincial values for British Columbia with averaged totals for Canada overall over time. Users could *consume* filtered data for both area selections and make inferences about the differences in mobility between regions. This may be of particular interest to users who have knowledge of COVID-19 case numbers or varied quarantine measures and wish to evaluate their impact. This task is not *directly* supported in either the Google COVID-19 Mobility Report small multiples (*Fig. 2*) or a simple choropleth design like *Fig. 1* due to the lack of data aggregation.

5.4 Comparing outliers against general trends

This dataset does not follow smooth trends due to a significant number of outlier data points, including holidays, cyclical behaviours in the general population like typical work weeks, and the effects of Google's data collection method. Therefore, we anticipate users will wish to *identify* outliers in the data and determine the causes of specific spikes and troughs to add context. Because there are a substantial number of factors that could culminate in an outlier, we expect that users will want to be provided contextual information rather than needing to rely on their external knowledge. From an ethical standpoint, users should also be alerted when information is missing so they do not make unsubstantiated assumptions.

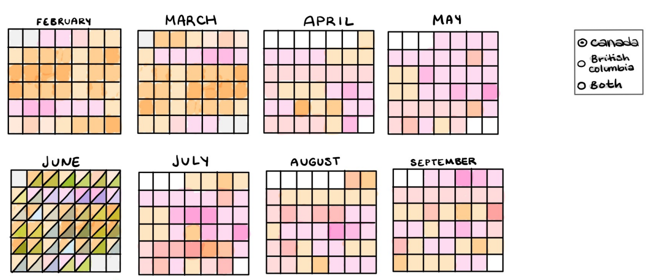

Fig. 3. A calendar heatmap visualization representing changes in mobility for an arbitrary attribute in BC and Canada

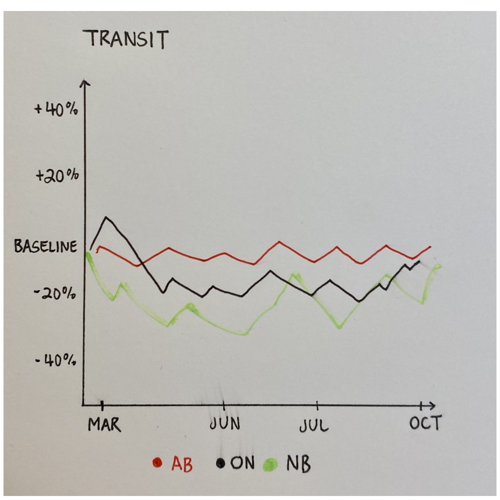

Fig. 4. A line chart for changes in mobility in the Transit category for AB, ON, and NB

6 Proposed Solution

We plan to curate an article on the changes in mobility trends due to the COVID-19 Pandemic across Canada using Observable [5], a platform which enables the interleaving of rich D3.js visualizations within a text document. Details about the visualizations we provide are listed in the following subsections.

6.1 Choropleth Map of Canada

This visualization focuses on showing an overview of how mobility is currently being affected in each province and territory in Canada. *Fig. 1* shows a rough idea of what this visualization will look like using a similar dataset; however, our design will also include a slider so the user can see how mobility changes over time. Because this figure is designed to be an overview, we limit the scope of the data that will be visualized in two ways. First, we will only be showing the average mobility of all categories as opposed to individual ones. Second, we also limit the time frame to only the current month. We reserve more detailed information for visualizations that come later in the article. Some of the limitations of choropleth maps that we have to be take into consideration are that it assumes a uniform distribution of data over the whole area and that we have to select the color scale such that differences between regions are differentiable from one another.

6.2 Provincial/National Comparison Calendar Heatmap

Fig. 3 displays a calendar visualization mock-up with heatmap-style colour coding. For a given mobility attribute, we will encode the daily average difference from pre-pandemic levels according to a pre-defined shade range. Users will have the ability to display values for both British Columbia and Canada, either independently or simultaneously. If the latter option is selected, the visualization will switch to two contrasting colour palettes as shown in the bottom-left corner of *Fig. 3*. We plan to include three iterations of this calendar design within our article to highlight changes in individual mobility attributes. In particular, they will focus on each of retail & recreation, grocery & pharmacy, and parks.

In addition to the basic calendar layout, we plan to overlay information about relevant provincial and federal holidays as well as government-imposed lockdown measures. After preliminary review of the data, these external factors notably affected mobility values and thus should be presented to the user.

6.3 Provincial Mobility Small Multiples

This visualization specifically addresses one of the shortcomings we found in the Google Community Mobility Report PDFs. We plan to create a small multiple of line charts that visualizes changes in mobility for the categories of Transit, Workplaces, and Residential. *Fig. 4* shows a mock-up of one of the line charts that could be embedded in the small multiple.

This visualization would enable the user to compare the mobility trends of numerous provinces, as opposed to just being limited to one region, and without the need to re-download a PDF each time.

7 Work in Progress

All work described in this section is currently available at:

`https://jyoo980.github.io/`

Please note that due to the nature of continuous integration and deployment, minor periods of unavailability may occur between new deployments.

7.1 Choropleth Map of Canada

Our initial design for this choropleth map has not changed significantly since the initial proposal. A sample image of the current state of the choropleth map visualization can be seen in *Fig 5*. In this visualization we have a map that is divided into provinces and visualizes the percent change for each of the six mobility categories. This is done through the use of the radial button selectors. Furthermore, we currently provide a slider for the user to select a specific date to view data.

One key component of this visualization that is remaining is to integrate a working legend into the visualization. At the same time, we need to determine whether it is best to use a legend color range that is static and is determined ahead of time based on all available data within that category, or a dynamic one in which we determine the range based on the data that is currently being visualized

Going forward, we have discussed the quantity and variety of data we wish to display with this visualization. Presently, we have decided on only visualizing an average of each mobility category for the past 30 days. However, this number may change depending on the data updates and how much information is "lost" due to the averaging.

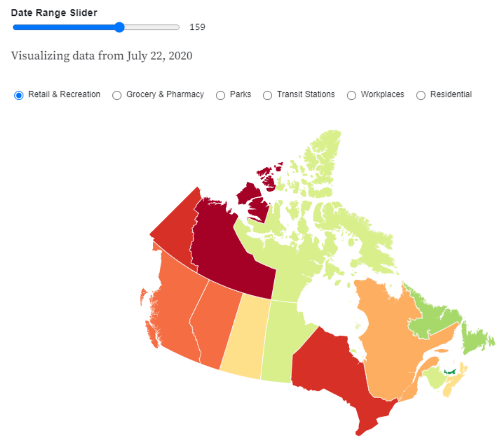

Fig. 5. A choropleth map of Canada showing the percentage change in mobility for the retail and recreation.

7.2 Provincial/National Comparison Calendar Heatmap

Following the initial calendar heatmap sketch shown in *Fig. 3*, we have implemented a similar first draft design as shown in *Fig. 6*. Currently, the calendar features working drop-down menus for selecting the displayed attribute (parks, retail & recreation, and grocery & pharmacy) and dataset (British Columbia or Canada) as well as a dynamic legend.

Similarly to (7.1), one item left to finalize is the visualization's legend, which gives concrete values to the displayed colour scale. Presently, the calendar dynamically updates to the respective value ranges as different attribute and dataset selections are chosen. Although this is helpful for viewing one selection, it increases the complexity of comparing between selections as the same shade may not represent the same percentage value from baseline. In contrast, fixing the legend values may reduce precision within shades for individual views. Ultimately, we plan to introduce a static legend with value limits representing the highest and lowest values in the unfiltered dataset.

Two notable features currently missing from our draft design include important date highlighting, such as government quarantine restriction changes and holidays, and the ability to select both the British Columbia and Canada datasets simultaneously for a given attribute. In the last phase of development, we will explore ways to include key date information without detracting from the user's ability to view overall trends. Additionally, we hope to include the split daily value model in some capacity, but will survey viable options available in the context of the complexity of the visualization. We also plan to incorporate some other minor visual changes, such as including a tooltip for displaying concrete values for a given day. As well, we may look into splitting the months into two rows.

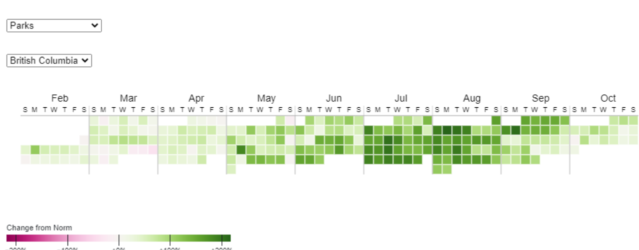

Fig. 6. A calendar heatmap displaying daily percentage mobility change for park attendence in British Columbia.

7.3 Provincial Mobility Line Charts

Our initial design for this visualization was to be a collection of small multiples consisting of line charts of the type in *Fig. 7*. A mock-up of this visualization quickly revealed that this would lead to a number of issues. One issue was the fact that displaying line charts for the categories of Transit, Workplaces, and Residential in small multiples would quickly lead to an increase in visual occlusion for the user.

We also found that using small multiples was not conducive to a user task that would be common in this domain. Specifically, we found it unlikely that users were interested in comparing changes in mobility across categories. Rather, we found that user would likely want to compare changes in mobility between provinces and regions *within* a category.

The issue of the ease of user-interaction was also demonstrated in how difficult it was to create granular controls for small multiples. Creating a single widget that would enable the user to toggle the range of dates and provincial selections did not provide the finer-grained level of control that we initially sought out to create. Creating a smaller set of widget for each line chart in the small multiple was found to be confusing and visually overwhelming.

Consequently, we made the choice to split the small multiple into a number of larger line charts. *Fig. 7* shows a line chart for the Transit category. A tooltip was also added to enable users to explore the data with a finer level of detail than what could be afforded with a non-interactive line chart. The strength of this design choice was especially evident in mitigating some of the effects of visual occlusion present when data from multiple provinces and regions were selected to be displayed.

8 Milestones and Schedule

Table 1 outlines a tentative schedule for the implementation of our project from initial implementation to final paper submission. Estimating person-hours on tasks is an incredibly error-prone process. Within *Table 2* is a listing of milestones we have currently reached with our visualizations. It details which group member was assigned which visualization component, actual completion date, and the actual amount of time that each milestone took to complete. An equal amount of the workload was distributed across each group member for these milestones.

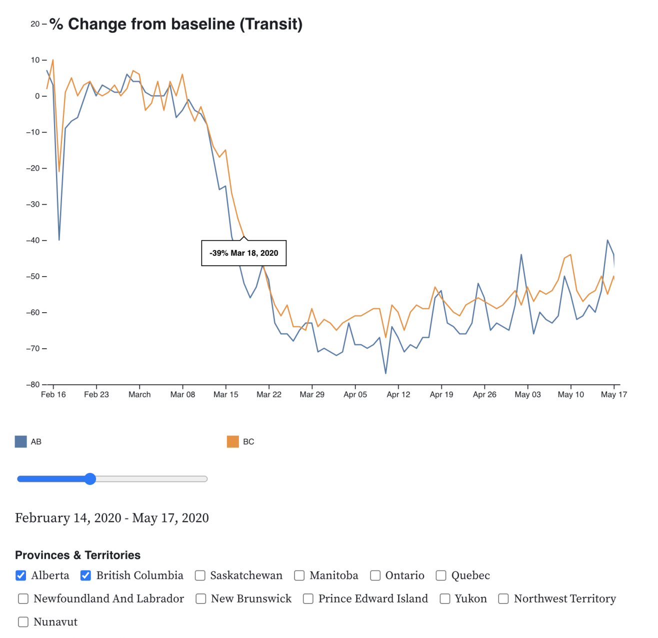

Fig. 7. A line chart for changes in mobility in the Transit category for AB, and BC

References

[1] CBC coronavirus tracker. `https://newsinteractives.cbc.ca/coronavirustracker/`. Accessed 10.15.2020.

Expected Completion	Task	Est. Hours (Total/Person)	Actual Hours	Status
October 21	Write project proposal and divide tasks	18 / 6	18	**Complete**
November 10	Implement first draft of article visualizations	15 / 5	$\sim$ 22	**Complete**
	Write first draft of interactive explainer article content	9 / 3	9	**Complete**
November 17	Update paper to include new changes	9 / 3	12	**Complete**
December 3	Implement final draft of article visualizations	15 / 5	-	**In progress**
	Write final draft of interactive explainer article content	9 / 3	-	**In progress**
December 10	Create content for final presentation and prepare talking points	9 / 3	-	**To do**
December 14	Update paper to include final designs and conclusions	12 / 4	-	**To do**

Table 1. Proposed project timeline

Visualization Component	Assignee	Actual Completion Date (Draft)	Actual Hours Spent (Draft)
Choropleth Map of Canada	Frank Yu	November 8	$\sim$ 8
Provincial/National Comparison Calendar Heatmap	Lily Bryant	November 13	$\sim$ 8
Provincial Mobility Line Charts (Previously Small Multiples)	James Yoo	November 10	$\sim$ 6

Table 2. Breakdown of visualization component ownership for drafts

[2] COVID-19 community mobility reports. `https://www.google.com/covid19/mobility/`. Accessed 10.14.2020.

[3] COVID-19 dashboard. `https://coronavirus.jhu.edu/map.html`. Accessed 10.15.2020.

[4] Distill: latest articles about machine learning. `https://distill.pub/`. Accessed 10.14.2020.

[5] Observable: Use data to think, together. `https://observablehq.com/`. Accessed 10.18.2020.

[6] BC's restart plan. `https://www2.gov.bc.ca/gov/content/safety/emergency-preparedness-response-recovery/covid-19-provincial-support/bc-restart-plan`, Jul 2020. Accessed 10.16.2020.

[7] M. Correll and J. Heer. Surprise! bayesian weighting for de-biasing thematic maps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):651–660, 2017. doi: 10.1109/TVCG.2016.2598618

[8] N. Keyfitz. Individual mobility in a stationary population. *Population Studies*, 27(2):335, July 1973. doi: 10.2307/2173401

[9] A. Makulec. Ten considerations before you create another chart about COVID-19. `https://medium.com/nightingale/ten-considerations-before-you-create-another-chart-about\-covid-19-27d3bd691be8`, Apr 2020.

[10] L. You, F. Zhao, L. Cheah, K. Jeong, C. Zegras, and M. Ben-Akiva. Future mobility sensing: An intelligent mobility data collection and visualization platform. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2653–2658, 2018. doi: 10.1109/ITSC.2018.8569697