Project Proposal

Section 1. Topic and Dataset

1.1. What is the data about?

1.1.1. Describe the data cases and data dimensions in the dataset

The dataset includes “crowdsource questionnaire information on how COVID-19 has impacted Canadians' experiences of discrimination, sense of belonging, trust in institutions and access to health care services.” The data cases are PUMFIDs (Public use microdata file identifiers) which identify each participant who took part in the questionnaire. The data dimensions are Sense of belonging (Local community, Town or city, Province or territory, Canada - born outside of Canada, Country of origin - exclude Canada, People with the same race or skin colour, People with the same ethnic/cultural background, People with the same religion, People who speak the same first language, Online communities most active in), Immigration status, Province of residence, Discrimination - 2yrs before COVID-19 and Since COVID-19 began (Language, Accent, Physical Appearance, Age, Sex, sexual orient or gender ident/express, Indigenous identity, Ethnicity/culture, Race/skin colour, Religion, Physical/mental disability), Discrimination situation - 2yrs before COVID-19 and Since COVID-19 began (Store/bank/rest, School, On the internet, Work, Housing, Police, Courts, Border into Can, Social gatherings, Public areas, Public transit), Difficulty/Problem accessing service(s) (Help for mental health, Cost, Getting an appointment, Contacting physician/nurse to get info, Waited too long to get health care svc, Not available at time required, Medical treatment, Appointment for rehabilitative care, Emergency services/urgent care, Transportation problems), Identifies as a person with a disability, and Emotional/psychological/mental health conditions.

1.1.2. Describe the size of the dataset

The dataset has 36674 data cases with 127 dimensions. We are not planning to use all of those.

1.1.3. Describe the nature of the dataset

Each data case is the set of answers to questions asked from an individual. Every item has a primary key to itself. It’s a relational dataset but probably not in an efficient way.

1.2. How did you collect the data?

The dataset is publicly available at www.statcan.gc.ca. The data can be downloaded directly in a CSV format. It will not require any processing and an API will not be needed.

1.3. Why does the data matter to you?

As an international student, I had been planning to come to Canada last year but the pandemic put a long pause on everything. I personally was affected by the pandemic and embarrassed in situations because I was not privileged to get vaccinated when the time was right. I was stopped at borders and spent a great deal of money and time on quarantine as I experienced many delays. I was wondering how this pandemic affected people living in Canada who are presumed to be different due to their

appearance, ethnicity, etc. and how they were discriminated against or felt left out. Also I was wondering how Canada provided healthcare services to the people who needed them the most.

Section 2: Data Questions

What are interesting data questions?

	Question about the data	Why is this question interesting to you?	Why is visualization well suited to answering this question?
Question 1	Did the pandemic fuel discrimination toward ethnically-diverse groups when applying for/seeking housing, crossing the border of Canada, and when present at public areas, public transit, social gatherings, store/bank/rest, on the internet, etc.?	This question can provide insight into the effect of the pandemic on people's behaviours toward ethnically diverse groups. In addition, the most common discrimination situations could be examined. With many reports of hate speech, physical attacks, racism, xenophobia, and hatred towards minorities and wrong accusations of them spreading the virus, the intensity of discrimination in various provinces can be examined.	It's very important to see trends and patterns of discrimination intensity in a detailed map of all the 10 provinces in Canada. Also, a before/after visualization gives a clearer overview of the changes 2 yrs. before the pandemic vs. during the pandemic.
Question 2	Which discrimination factors were most apparent before and during the pandemic and how did the pandemic affect the intensity of these acts?	It's fascinating to see the most common and least common discrimination factors before and during the pandemic. Comparing the effect of physical discrimination factors (Skin colour and physical appearance) and auditory discrimination factors (Language and accent) would be interesting to visualize as well.	Visualization is suited because it's possible to use scales to show the difference between the number of people who were discriminated against on two timelines based on different factors. Area can be used to approximate the total number of people who were discriminated against as well.

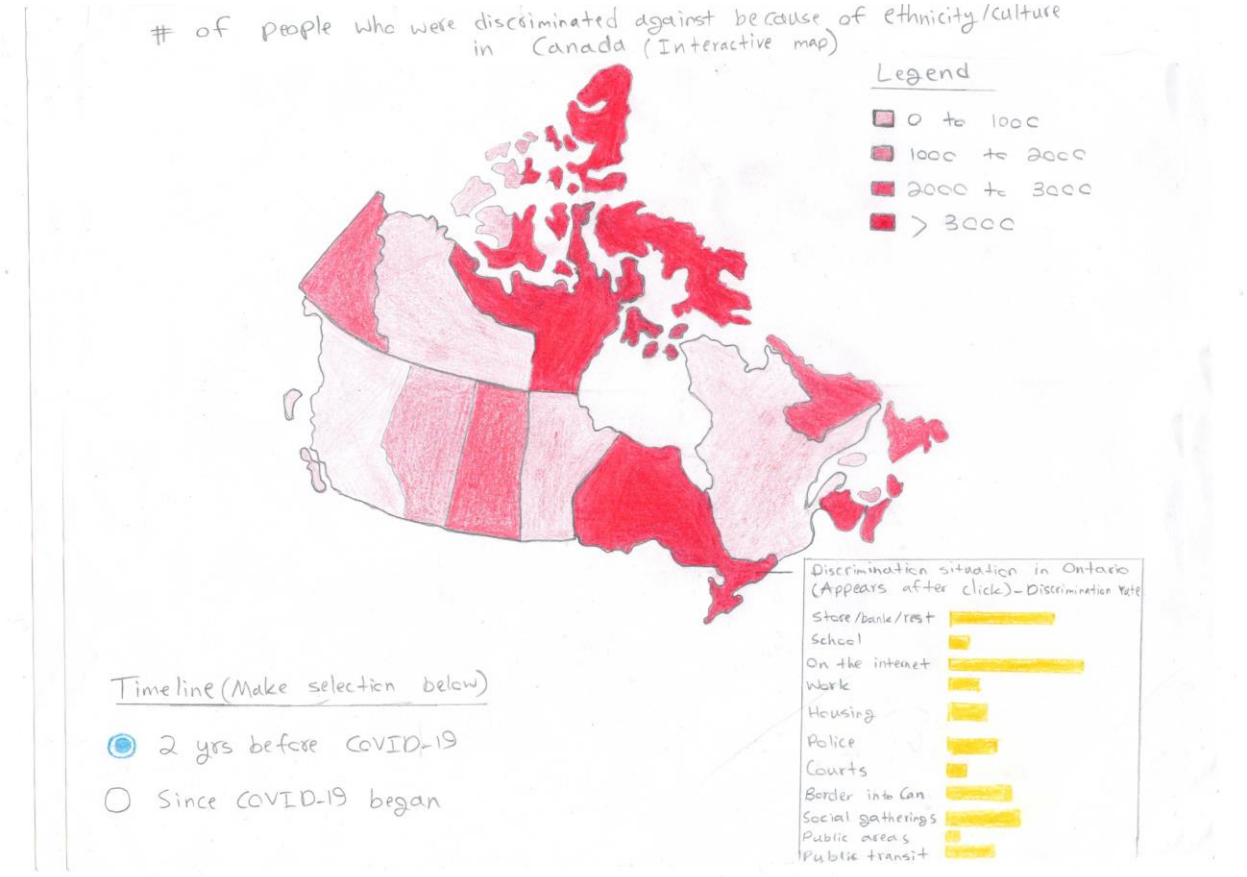
Question 3	For immigrants who were discriminated against because of race or skin colour, ethnic/cultural background, religion, language, and sexual orientation, compare their sense of belonging to different communities they are a part of.	It's important to see how different discrimination factors affect immigrants' sense of belonging to different communities. In addition, the influence on the strength of belonging is fascinating. In times of isolation and loneliness during the pandemic, people's relationships with communities need to be examined very carefully.	Visualization is the only conceivable way to see how discrimination factors affect the strength of connection to different communities for individuals.
Question 4	For two groups of people which are the disabled and people with mental disorders, what kinds of difficulties accessing healthcare services have been more tangible?	These two groups of people tend to be the most vulnerable of people who need serious medical attention. The question of how well Canada could manage providing healthcare to them is a question of altruism.	Visualization can offer a sense of the most difficulties vulnerable groups have encountered accessing healthcare services during a critical time and even compare them.

Section 3: Design Ideation

For each design, provide 1) a sketch; 2) a description of the sketch; and 3) a description of the data questions the sketch can help answer.

3.1. Sketch 1

3.1.1. Image of the sketch



3.1.2. Description of the sketch and design rationale

This sketch is a map of Canada with a detailed separation of all the 10 provinces that displays the number of people that were discriminated against in each province because of ethnicity/culture. This is meant to be an interactive map, with the audience having the choice to select a timeline for the map (2 yrs. before COVID-19 vs. Since COVID-19 began) and select a province to view the discriminated rate in various circumstances. Area is main visual mark that is used along with colour value (# of people who were discriminated against), length/1D size (Discrimination rate in various situations), and colour hue. The map gives a good overview of patterns and trends in two different timelines and also provides more detail for each province by selection.

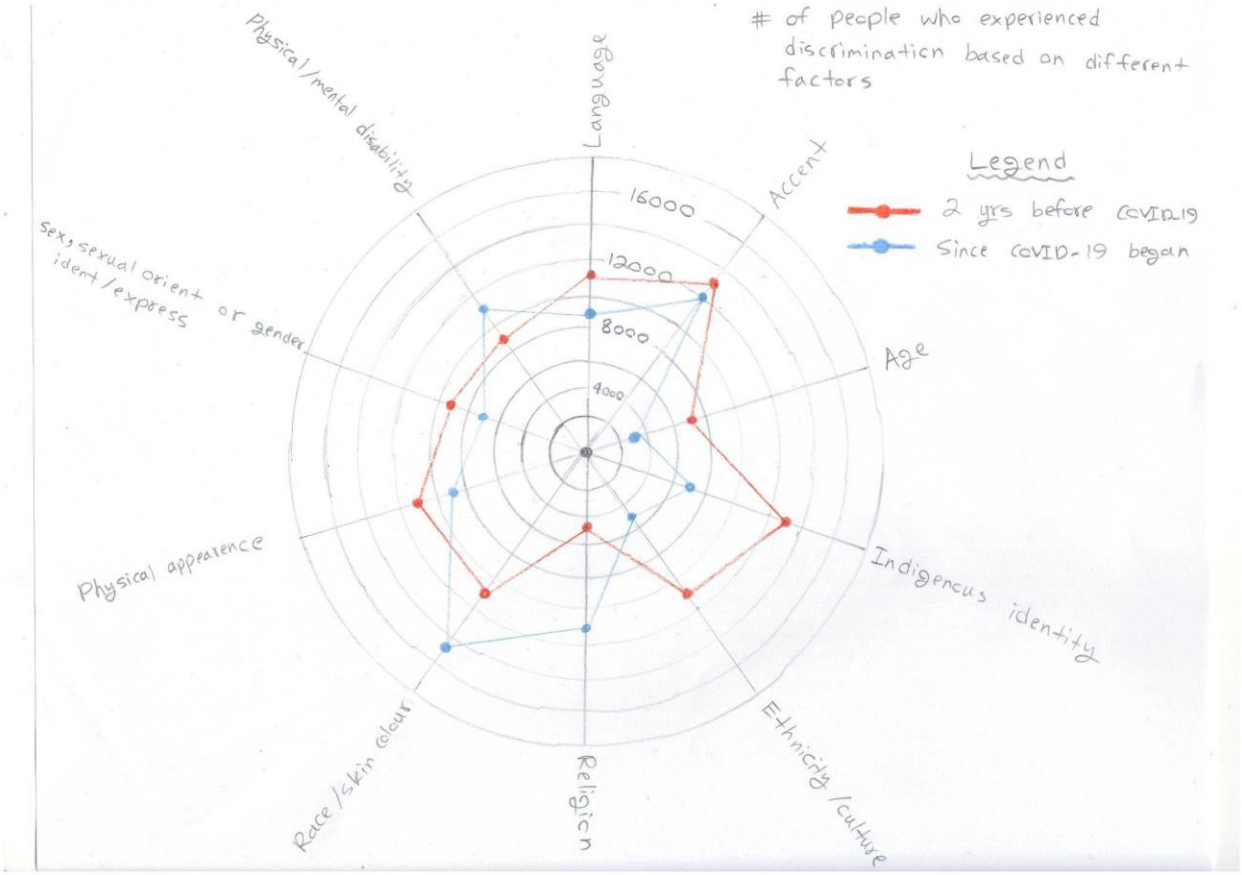
3.1.3. Which data questions the sketch can help answer, and why

This sketch can answer the following data question: “Did the pandemic fuel racially-motivated discrimination toward diverse ethnicity when applying for/seeking housing, crossing the border of

Canada, and when present at public areas, public transit, social gatherings, store/bank/rest, on the internet, etc.“

3.2. Sketch 2

3.2.1. Image of the sketch



3.2.2. Description of the sketch and design rationale

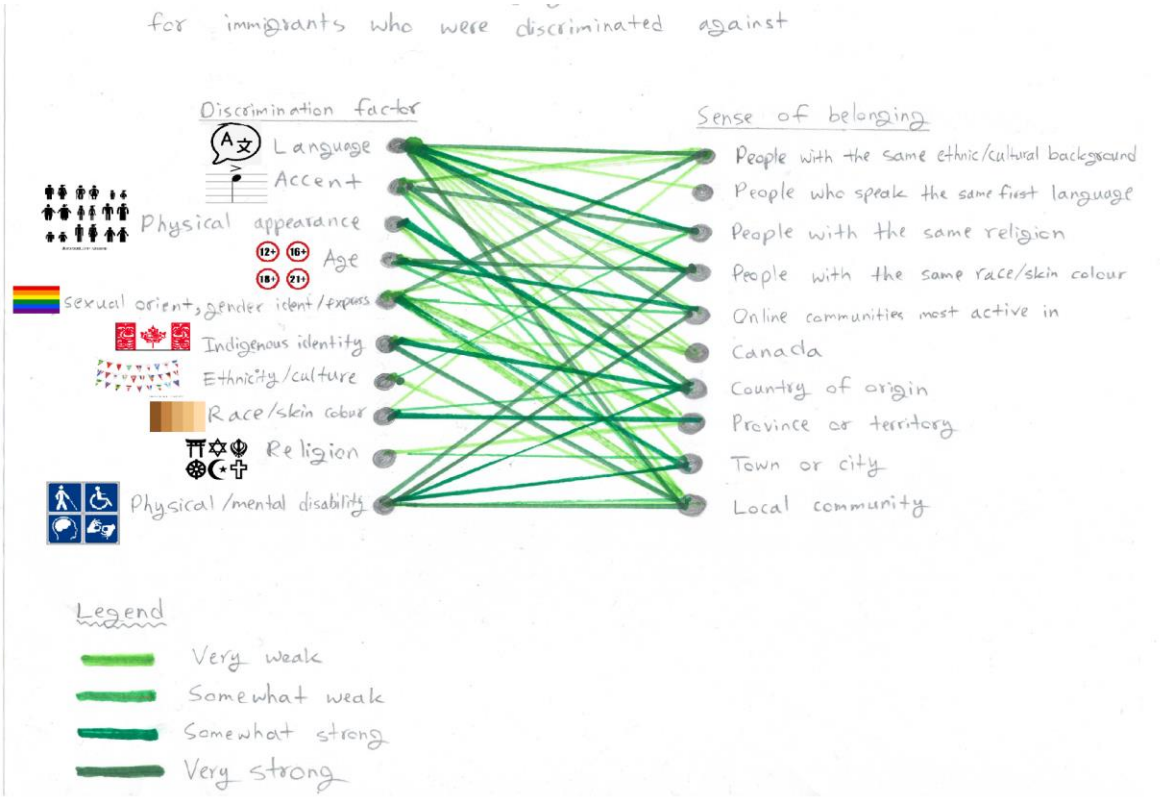
This sketch provides a visualization of the number of people that were discriminated against according to different factors 2 yrs. before COVID-19 and since COVID-19 began. All visual marks are used: points to indicate the location of the factor on the scale, lines to connect discrimination factors on the same timeline, and area to indicate the total number of people who were discriminated against. Colour hue is used to differentiate between two different timelines and the position of points on the scale indicates the number of people who were discriminated against because of the factor on the aligned scale. Each line that connects to the centre of the circle indicates a discrimination factor, and position is used to align those factors accordingly.

3.2.3. Which data questions the sketch can help answer, and why

This sketch can answer the following data question: “Which discrimination factors were most apparent before and during the pandemic?“

3.3. Sketch 3

3.3.1. Image of the sketch



3.3.2. Description of the sketch and design rationale

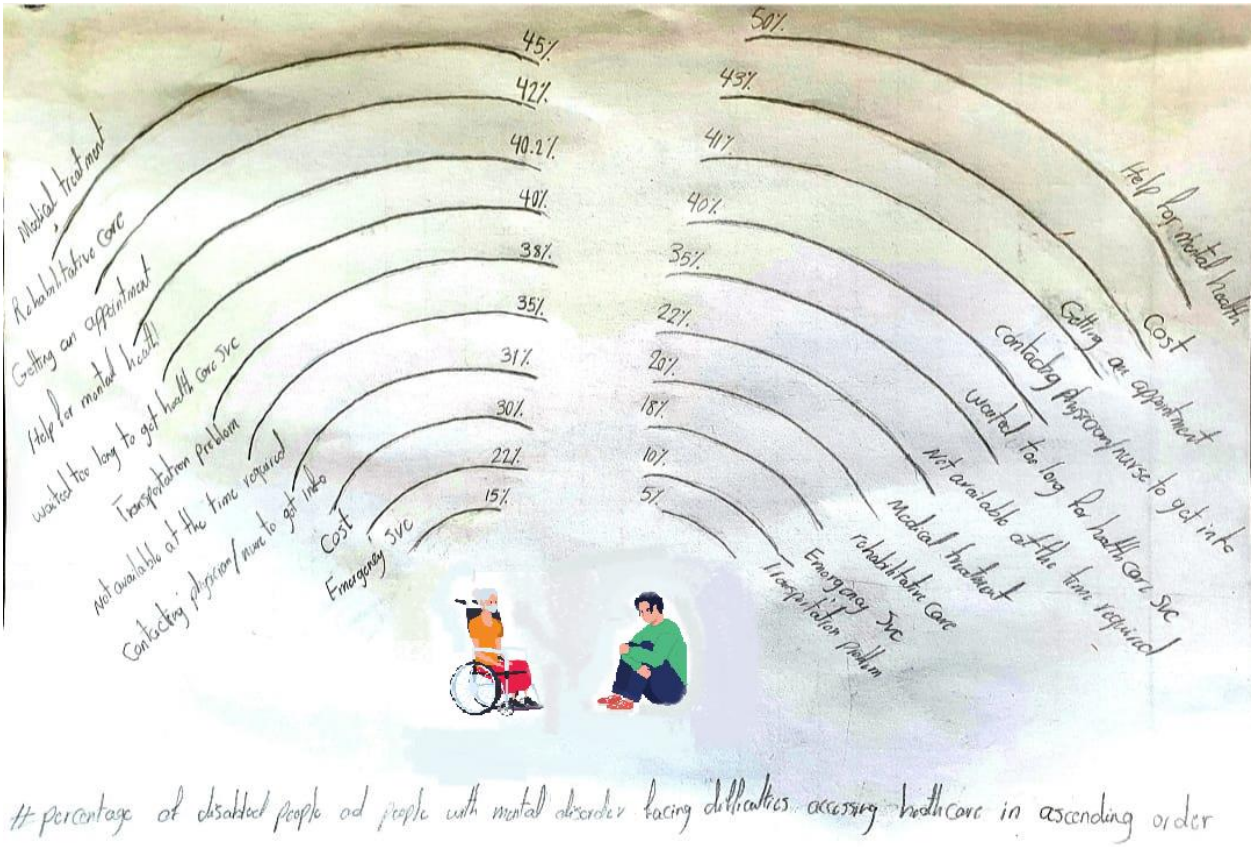
In this sketch, the strength of the sense of belonging of immigrants to different communities are visualized based on how they were discriminated against. Points and lines are used as visual marks in this visualization. Spatial region is used to position similar nominal/categorical attributes. Discrimination factors are aligned on the left and factor for sense of belonging are aligned on the right. Moreover, colour value is used to indicate the strength of belonging to different communities. Each line that connects the two groups corresponds to a single person.

3.3.3. Which data questions the sketch can help answer, and why

This sketch can answer the following data question: “For immigrants who were discriminated against because of race or skin colour, ethnic/cultural background, religion, language, and sexual orientation, compare their sense of belonging to different communities?“

3.4. Sketch 4

3.4.1. Image of the sketch



3.4.2. Description of the sketch and design rationale

This sketch tries to demonstrate how the disabled and the people suffering from mental health issues experienced inconvenience receiving health care services they were in need of during the pandemic. Percentage of disabled people and people with mental disorders facing certain kinds of difficulties accessing healthcare in ascending order. It uses the line as a visual mark. The visual variables are spatial region, shape and curvature.

3.4.3. Which data questions the sketch can help answer, and why

This sketch can answer the following data questions: “What percentage of the disabled people or people with mental health disorders faced each kind of difficulties receiving healthcare services during the pandemic? “, “Order of difficulties these people faced based on the lowest to highest percentage of the population experiencing these difficulties ”