

Michael Rediron
ID: 10138189
Due Date: April 14, 2021
Email: mdrediro@ucalgary.ca

CPSC 583 Final Project Report

1. Introduction

For my project I choose to do a visualization on educational attainment of the Non-Indigenous population, and Indigenous population in Canada for the years 2018-2019. My goal of the visualization was to illustrate the need to help encourage the Indigenous population to pursue higher education. My visualization allows the user to view and compare educational attainment data for each of the population groups in different regions of Canada. The user can click on a region and it will display the data for that region for both population groups in a dialog box for the years 2018/2019. The user can also compare the data of the different regions using the graph boxes. The graph boxes allow the user the ability to control the variables they want to compare such as population type, graph type, year, and region. I felt by allowing the user to have multiple views it would give a better overall picture and help illustrate the concept of the visualization.

In this report it demonstrates the start to finish process I took to create my final visualization. It includes the data description, design process, and final implemented visualization. For the data description it discusses the context of the data, pros/cons of the data, and why I decided to use this data for my visualization. For the design process it illustrates the process in which I took to come up with my final design. It includes initial data sketches and prototype variations which helped encourage the direction I wanted to take and the pro/cons of visualizing the data in that context. It also includes the process of taking my ideas and changing them into a final visualization which I can then improve on. The final implemented visualization discusses the interactions that were added to better portray the data to viewers of the visualization. It also discusses why I ended up choosing this final design and why I feel the final visualization conveys the data in a clear and concise manner that is beneficial to the overall concept.

In conclusion I am happy with my overall design and I feel it conveys the concept of the visualization in an adequate manner. By allowing the viewer to view and compare the data of different population characteristics it clearly illustrates that the Indigenous population is behind the rest of the population of Canada when it comes to pursuing education. This demonstrates the need to help encourage the Indigenous population to pursue education which will overall improve their quality of living. University has improved my life greatly and I hope in the future more Indigenous people receive the benefits from education that I have.

2. Data Description

2.1 Description

The dataset I submitted represents the educational attainment of the total population, non-indigenous population, and indigenous population in Canada and provinces from 2018-2019. The five columns topics I will be using data from is Year ,GEO, Educational attainment level, population characteristics, and value. Year indicates the year that the data was collected. GEO indicates the location of the data and can be Canada as a whole or the provinces/territories. Educational attainment level indicates the highest level of education attained which can be represented by less than high school, high school, trades, college, university and all levels. Population characteristic indicates if the population is off-reserve aboriginal , non-aboriginal population and total population. Value indicates the percentage that has that educational attainment. The dataset source is Statistics Canada and I could not find any information about how the data was collected. I think this dataset will be interesting to work with because it highlights the fact that the educational attainment of the indigenous population is not as strong as the non-indigenous population.

2.2 Pros/Cons Discussion

One pro of this dataset is that it highlights an important social issue that affects that indigenous population and the need to improve in that area. One other pro of using this dataset is that it uses all 3 types all 3 types of data nominal, ordinal and quantitative. One con of this data is that it only has 5 columns and the year column is limited to two years. One other con is that it does not include the on-reserve indigenous population and if this information was taken into account I believe the educational value for the indigenous population would be much lower.

2.3 Data Set Decision:

I want to use this dataset because it highlights an important social issue that affects that indigenous population. This is important to me because I am a indigenous student and I believe it is important to encourage indigenous youth to pursue education. One other reason I should be able to use this dataset is because the data is consistent and has no missing values. This dataset also uses all 3 types of data we learned about in class.

3. Design Process

3.1 Sketchable Data Subsets:

My first subset is the Educational Attainment of Off-reserve Aboriginals in Canada 2019. The tradeoff between sketchability and how it represents the full data set is that it leans strongly towards sketchability. When I was drawing my initial sketches I realized I should have included the Non-aboriginal population also because my sketches were kinda limited. The subset did cover all the educational attainment options though but it would have been better if the Non-aboriginal population was included.

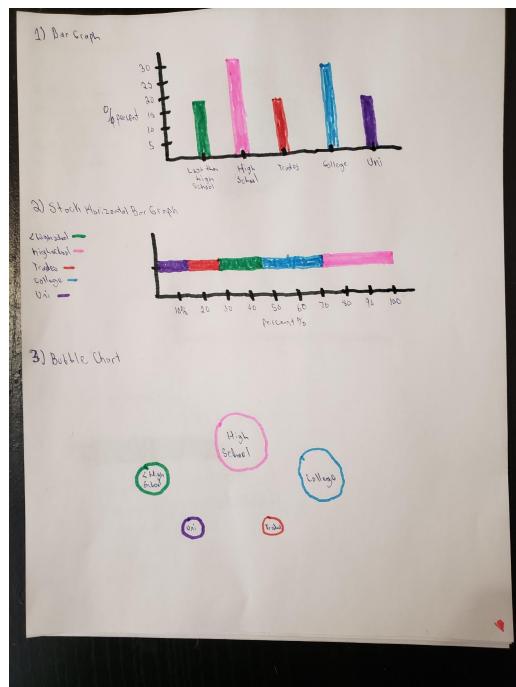
My second subset is the University Attainment of Off-reserve Aboriginals and Non-aboriginals in Alberta, British Columbia, Saskatchewan, and Manitoba in 2019. This subset had much more of an equal trade off in terms of sketchability and the representation of the data set as a whole. This subset limited the educational attainment options to just university and a subset of the regions represented. By making this subset broader it allowed me to give a better representation of the full data set and it also gave me more options when creating the sketches.

3.2 Design Direction

Sketchable Dataset (Set 1): Educational Attainment of Off-reserve Aboriginals in Canada 2019

2019 Canada	2016A000 Less than high school	Off-reserve Aboriginal	Percent	239 units	0 v105881611.2.2	18
2019 Canada	2016A000 High school	Off-reserve Aboriginal	Percent	239 units	0 v105881611.3.2	29
2019 Canada	2016A000 Trades	Off-reserve Aboriginal	Percent	239 units	0 v105881611.4.2	14
2019 Canada	2016A000 College	Off-reserve Aboriginal	Percent	239 units	0 v105881621.5.2	26
2019 Canada	2016A000 University	Off-reserve Aboriginal	Percent	239 units	0 v105881621.6.2	14

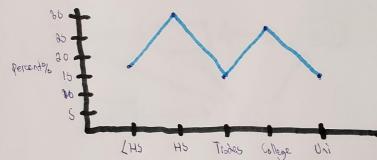
Initial 10 Sketches:



4) Pie Graph



5) Line Chart

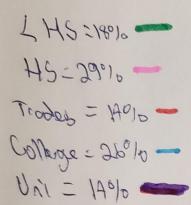


6) Item value representation

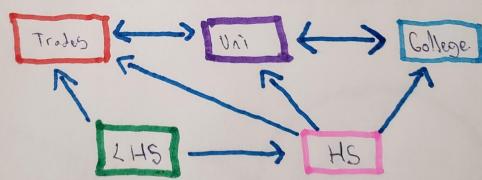
Triangle A = 5 Square B = 1

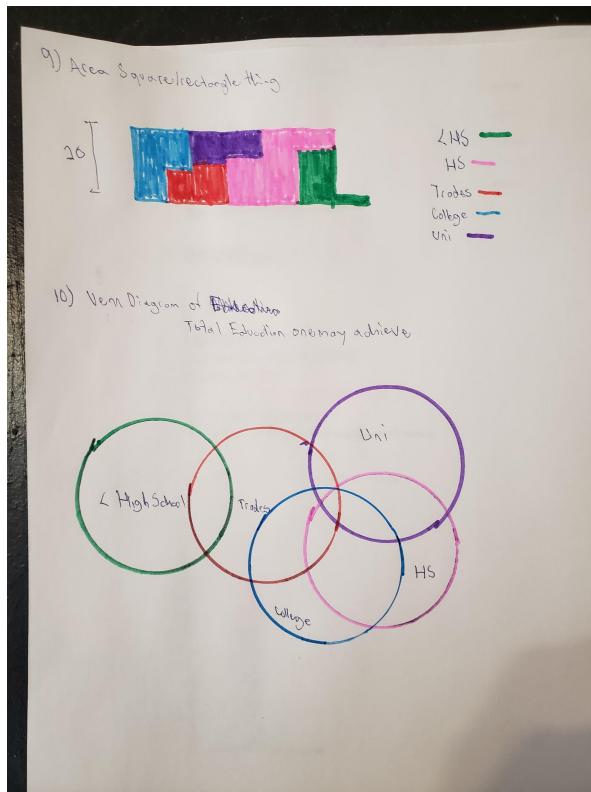
LHS HS = △△□□□
Highschool = △△△△□□□□
Trades = △△□□□□
College = △△△△□
Uni = △△□□□□

7) Donut Graph

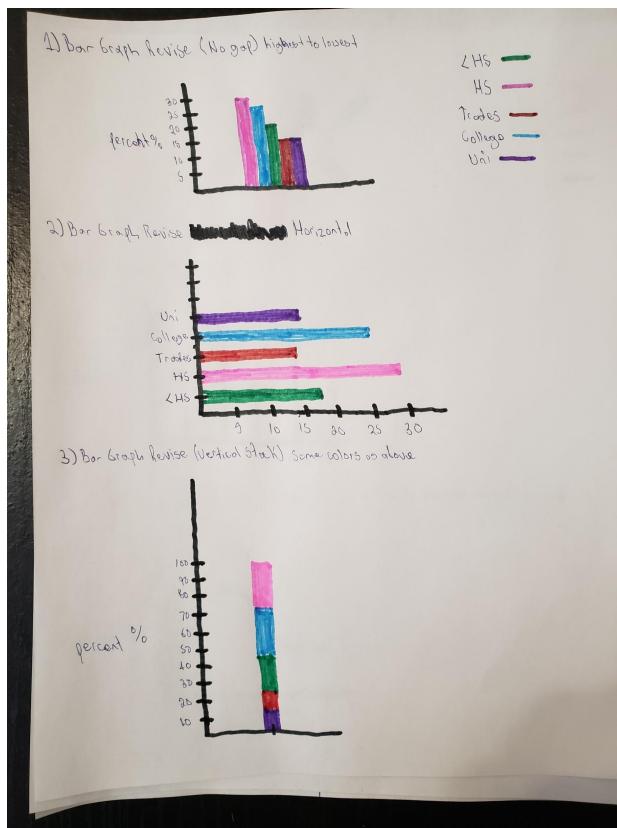


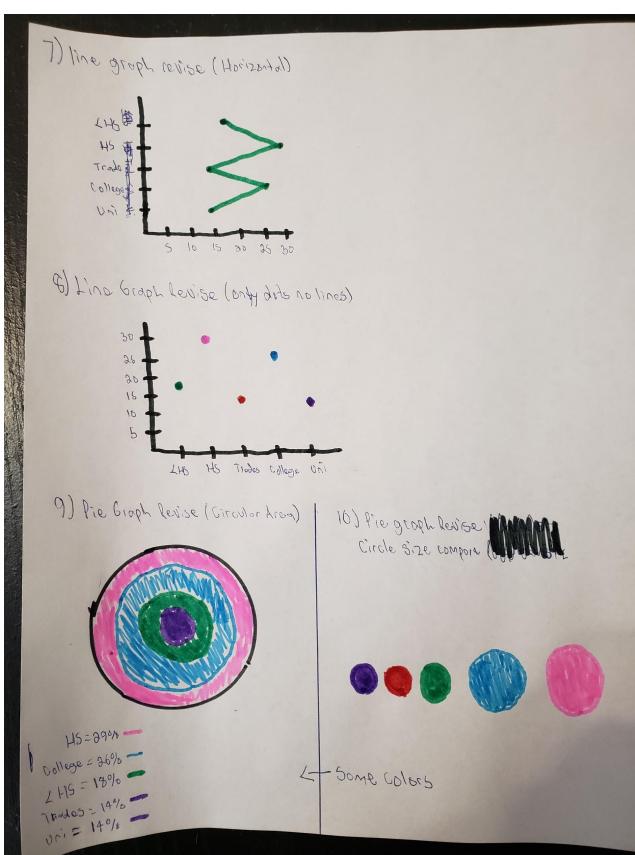
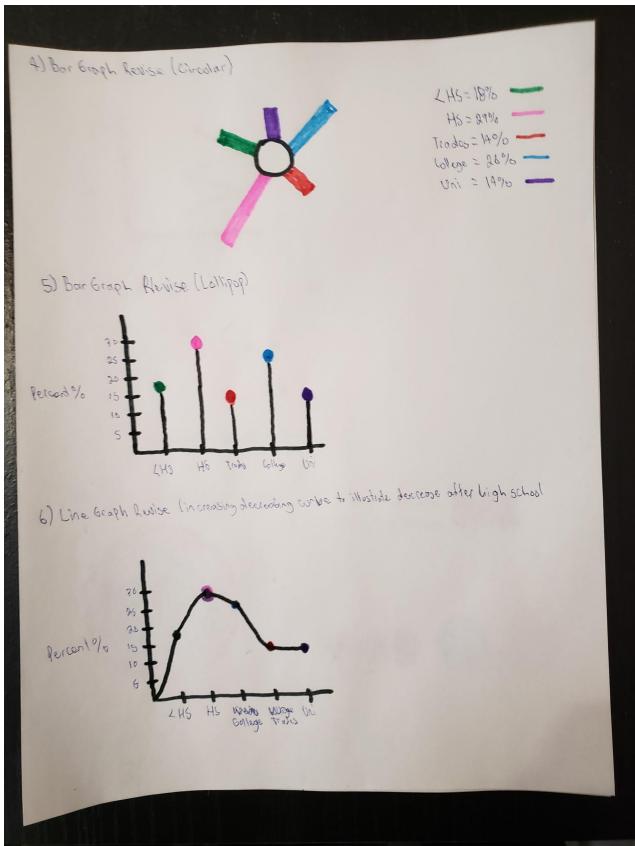
8) Paths to Education Chart





Revised 10 Sketches:

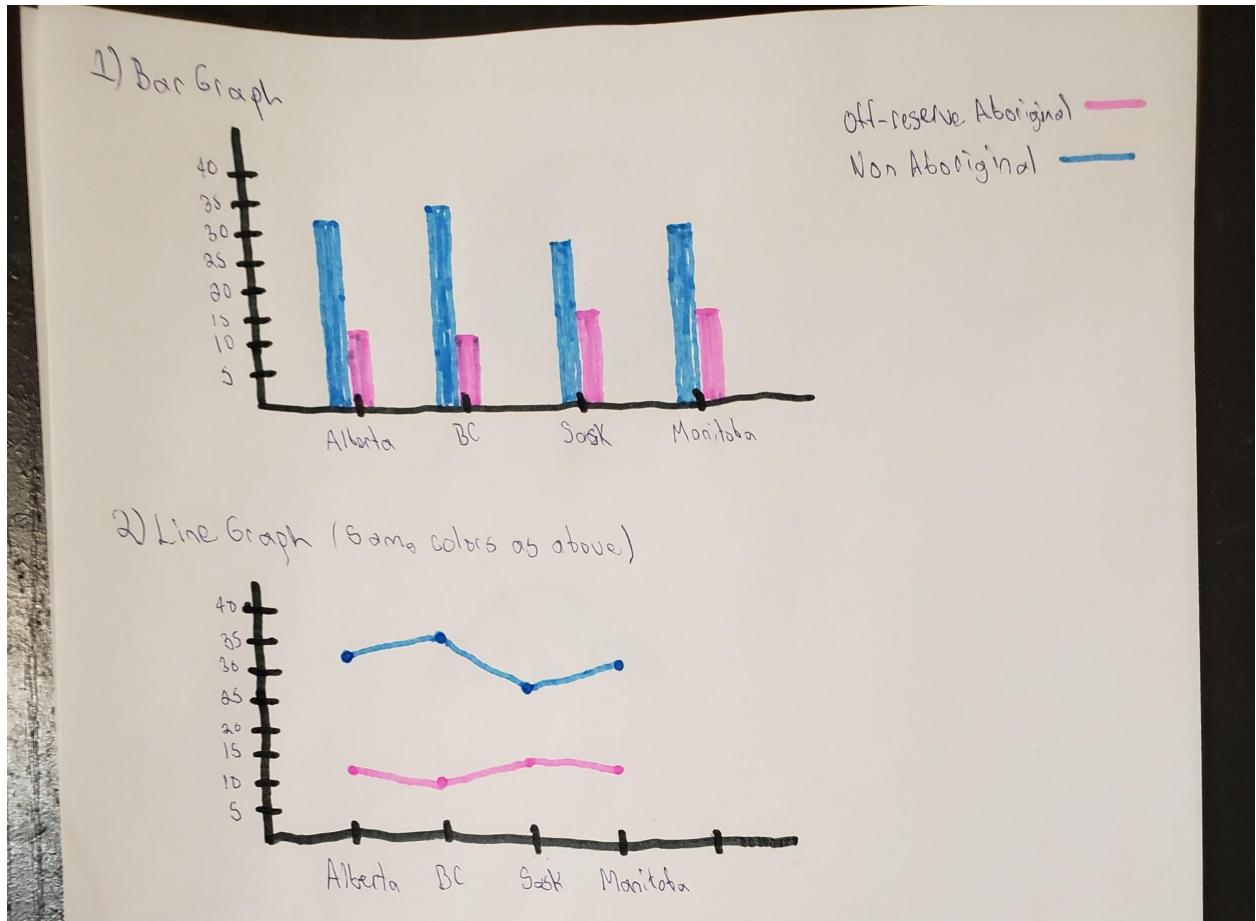


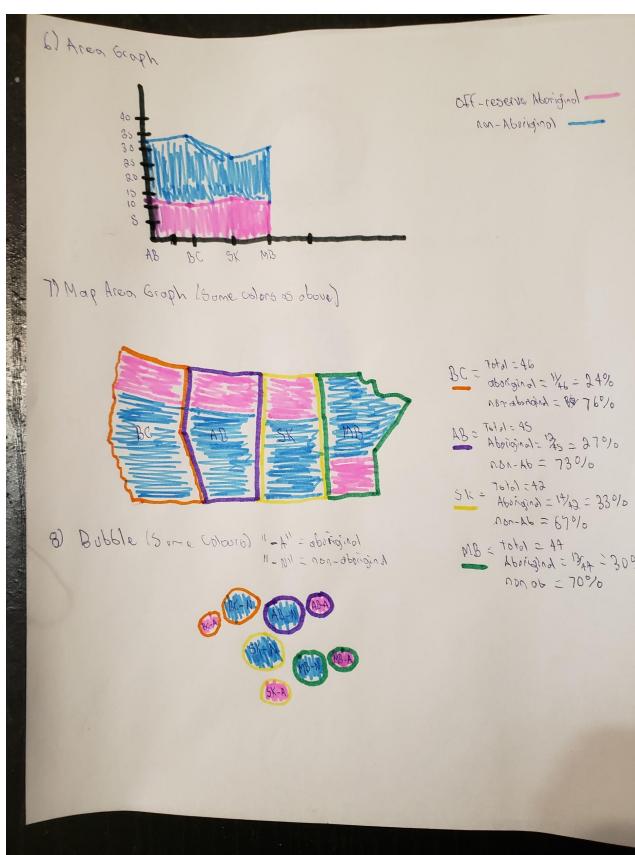
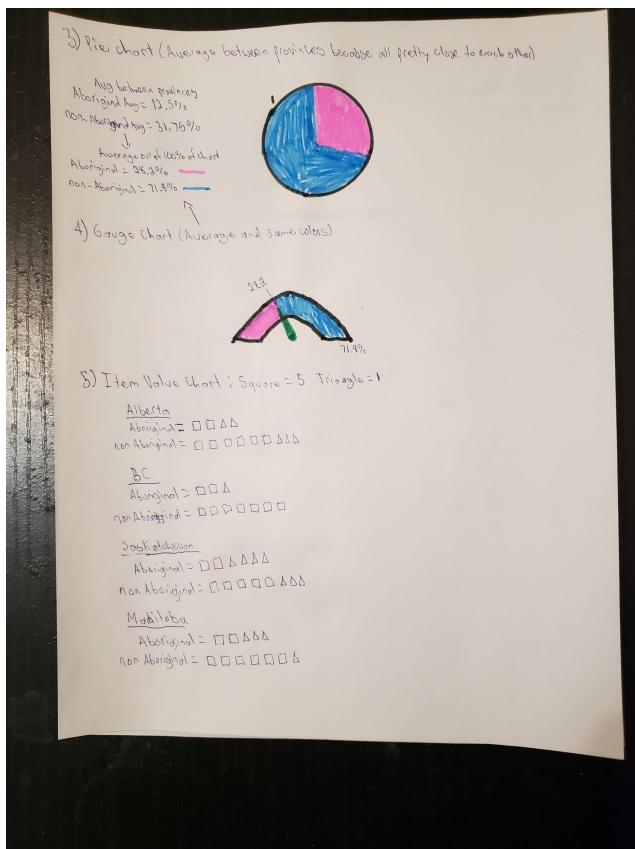


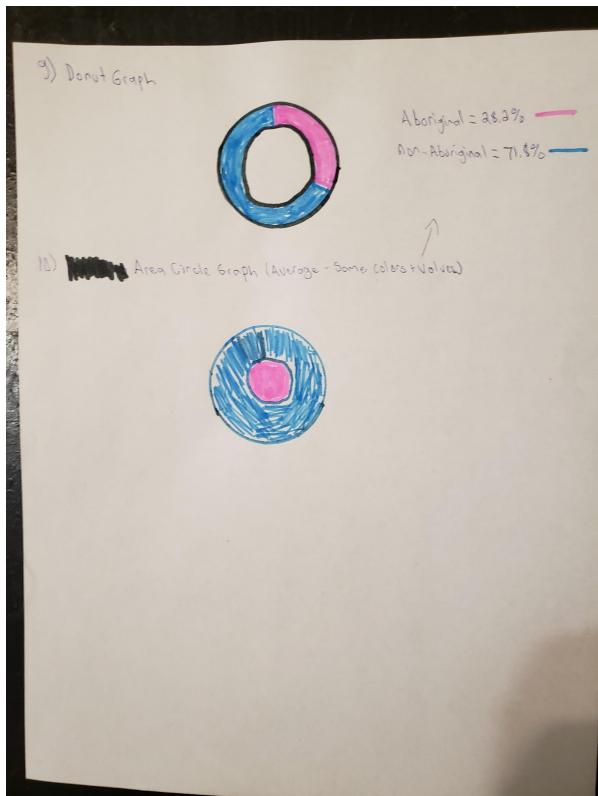
Sketchable Dataset (Set 2): University Attainment of Off-reserve Aboriginals and Non-aboriginals in Alberta, British Columbia, Saskatchewan, and Manitoba in 2019

2019 Alberta	2016A000; University	Off-reserve Aboriginal Percent	239 units	0 v1058817€ 10.6.2	12
2019 Alberta	2016A000; University	Non-aboriginal population Percent	239 units	0 v1058817€ 10.6.3	33
2019 British Col	2016A000; University	Off-reserve Aboriginal Percent	239 units	0 v1058818€ 11.6.2	11
2019 British Col	2016A000; University	Non-aboriginal population Percent	239 units	0 v1058818€ 11.6.3	35
2019 Saskatche	2016A000; University	Off-reserve Aboriginal Percent	239 units	0 v1058817€ 9.6.2	14
2019 Saskatche	2016A000; University	Non-aboriginal population Percent	239 units	0 v1058817€ 9.6.3	28
2019 Manitoba	2016A000; University	Off-reserve Aboriginal Percent	239 units	0 v1058817€ 8.6.2	13
2019 Manitoba	2016A000; University	Non-aboriginal population Percent	239 units	0 v1058817€ 8.6.3	31

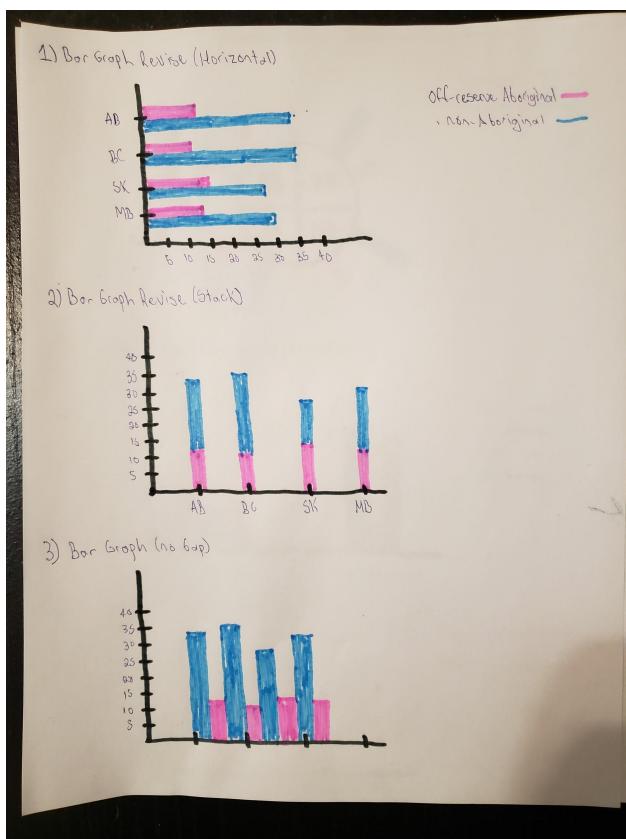
Initial 10 Sketches:



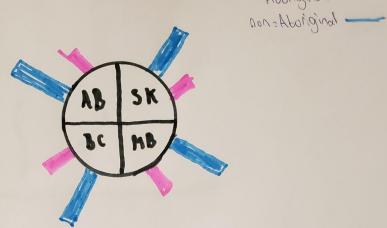




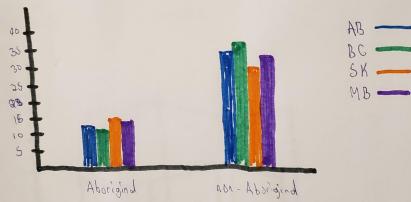
Revised 10 Sketches:



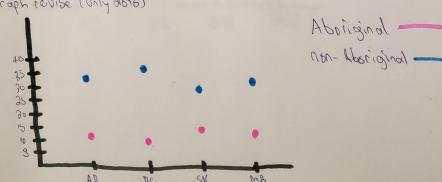
4) Bar Graph (Revise) Circle



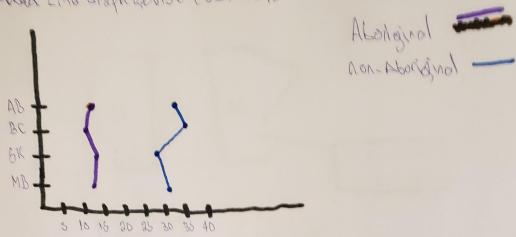
5) Bar Graph Revise (Aboriginal & non-Aboriginal split up)



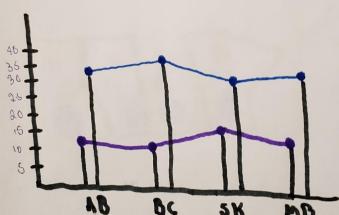
6) Line graph revise (only dots)

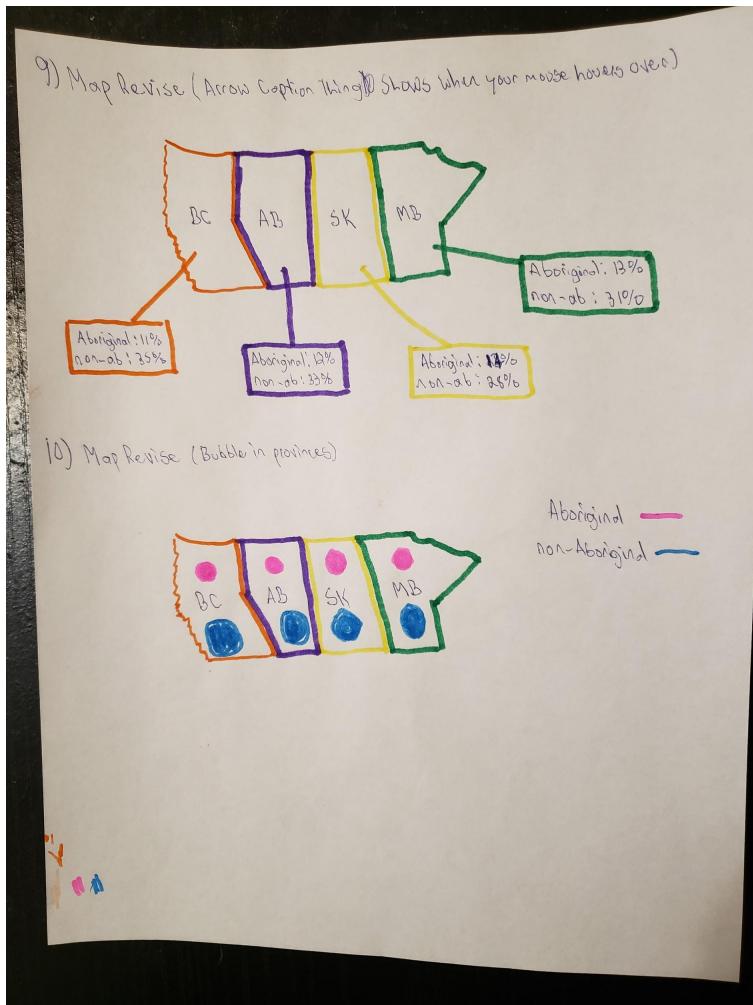


7) Vertical Line Graph Revise (Vertical!)



8) Bar/Line Revise (1st steps/line combo)





3.3 Process

Design Direction in the Sketches:

My sketches on a general level are pretty generic and have clear similarities between both sets of sketches. I am not very creative and I am having trouble coming up with unique ways to represent the data, but I feel once I am able to use a more computer oriented approach my ideas will become more unique.

Discussion of Process and Results:

My process of creating the data subsets was to choose something that was easily sketchable and choose something that would not take too much time. The result of this was my sketches becoming limited and generic. For example in the first dataset by only using Off-reserve Aboriginals it limited how I could represent the data in my sketches and it also left the overall motivation of the representation of the data set in an inconsistent state. One way I

can improve on the next representations is to put more in depth thought into what the data represents and then reflect it into visual ideas.

Conclusion:

Overall, creating these sketches was a valuable learning experience and it did bring into view the process it takes to make a good visualization. This process helped me discover the limitations of my data and some of the creative obstacles I will need to overcome. I was also able to think of a few ideas of how to implement the data set. One idea I had was to have a map where when the mouse would hover/click over a region it would display the data selected from a few drop down menus such as year, educational attainment type, and population characteristic.

3.4 Design Direction

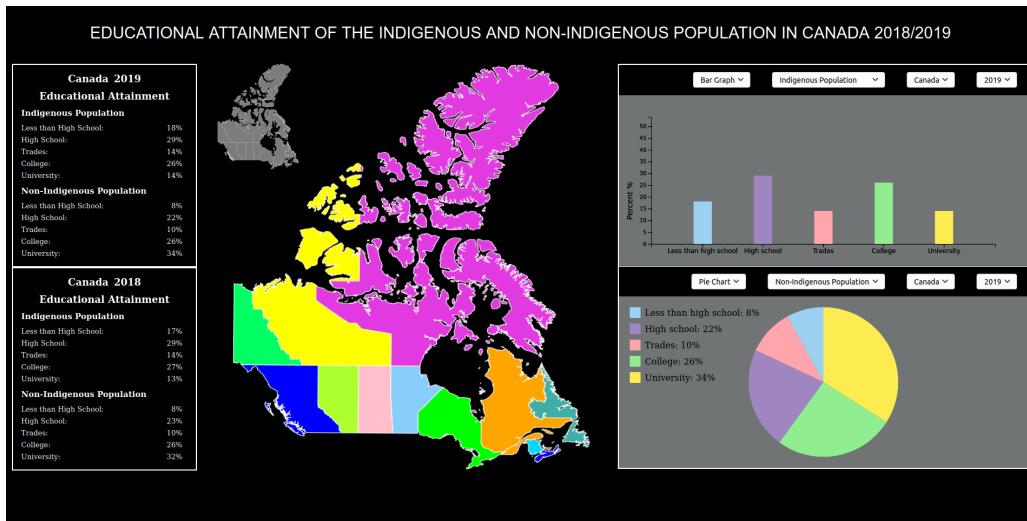
Data Set: Educational Attainment of the Indigenous and Non-Indigenous Population in Canada for 2018/2019

Columns: Year, Location, Education, Population Type, Value (These are renamed from original dataset for easier access during implementation)

General Design Direction:

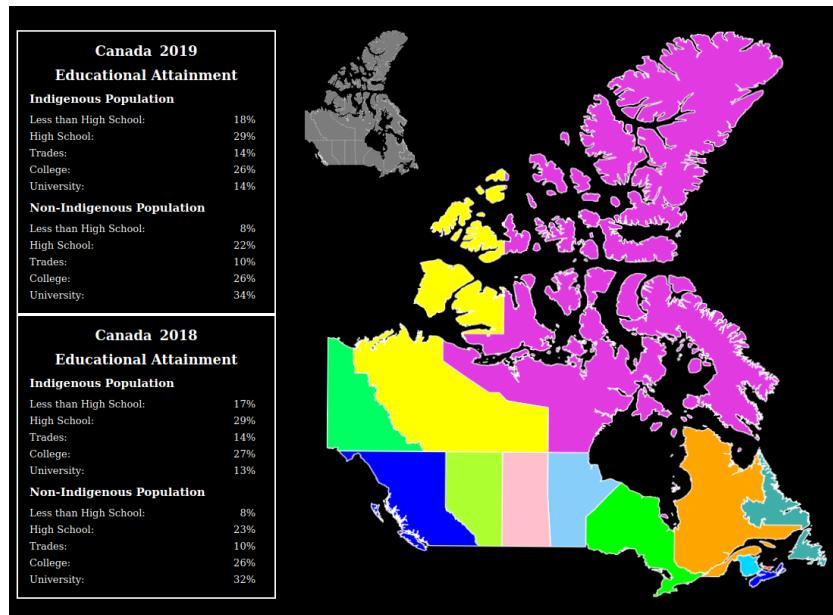
My general design direction is to have a dashboard type of visualization that incorporates a map of Canada. The visualization will have a map of Canada that will provide information on the dataset when clicked on by the user. When the user clicks on a region it gets the data for that region and displays it in a dialog box to the left of the map. There would be a top dialog box and bottom dialog box corresponding to the years 2018 and 2019. The information that would be displayed would be the region name, year, population characteristic and educational attainment. It will also have a section to the right of the map where it displays the information in graph/chart form. By having a section for the data to be displayed in graph/chart form it will allow the user to compare differences in regions, years and population characteristics. It will also allow the data to be viewed as different representations/graphs to better understand what the data is trying to convey. I decided to take the dashboard direction because I like the idea of incorporating a Canadian map and its provinces, but it would be hard to display all of the dataset information just using a map. I feel by going this route it will represent the data effectively and in an organized way.

Final Design Direction Image:

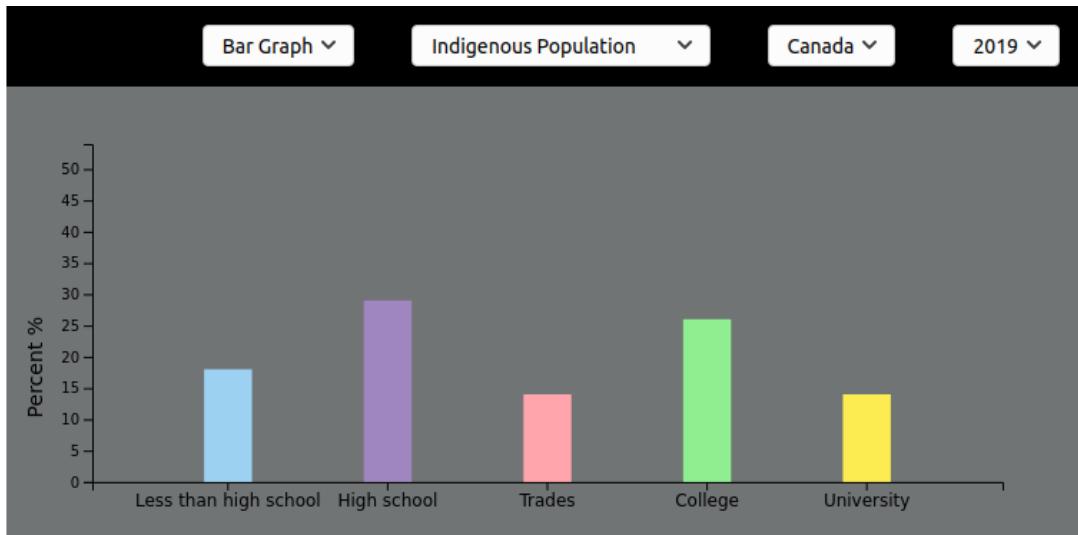


3.5 Prototyping Variations

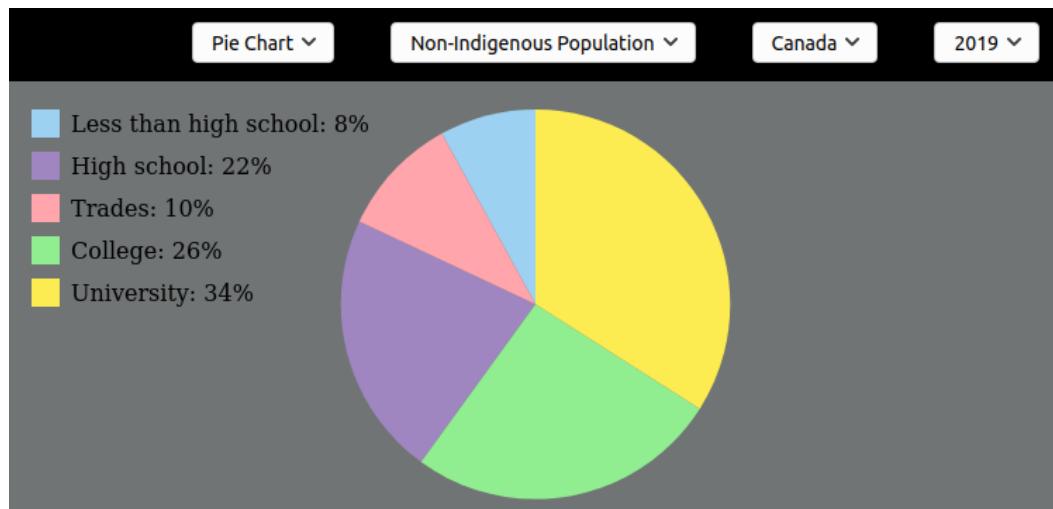
- Map with a Dialog Boxes: This variation would be a map of Canada and provinces that allows the user to click on the region and get the educational attainment data for the Indigenous and Non-Indigenous population for the years 2018 and 2019. The dialog boxes display the percentage for each educational level per year organized by the population type. By representing the data in this form it allows the user the ability to view the data from a clear and complete view. To get the Canada data click on the small Canada map and it defaults to this info when the page is first loaded.



- 2) Bar Graph: The second variation I chose is a bar graph where each bar represents an educational level such as less than high school, high school, trades, college, and university. The user is also able to change the graph type, region, year, and population type using the drop down menus.



- 3) Pie Chart: The third variation I chose was a pie chart where each slice represents an educational attainment level. There is also a legend that is colored coded to each educational level and it also includes the value of each. The user is also able to change the graph type, region, year, and population type using the drop down menus.



Project Links:

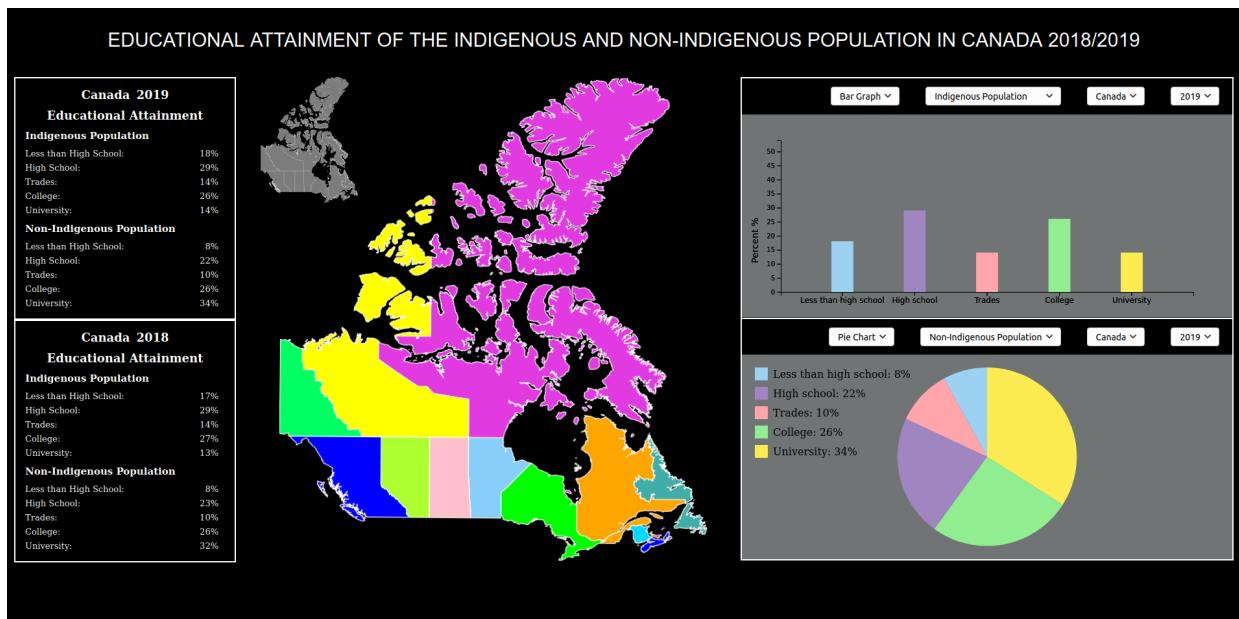
- 1) Live Examples: <https://michaelrediron.github.io/CPSC583-DataVisualization/>
- 2) Code Repo: <https://github.com/MichaelRediron/CPSC583-DataVisualization>

3.6 Process of Implementing Representation & Presentation

First I decided on the page layout and where I was going to put each variation. Then I used html and css to create each section. The sections are the dialog box section on the left side, the map section in the center, and the graph section on the right. For the left section I then used html to display the data categories that would be filled in by a click event. Then I found a free Canada SVG with provinces and pasted the svg into the middle section. I then created a fillDataBox function using javascript that will fill the data box when given a region as input. After that was functional I created an event listener that would fill the data box every time a click event happened on the map. I then moved on to the graph section where I made a function that would create a bar graph and a pie chart. Both of those functions used d3 in order to create. After those were working I had an event listener that would trigger if the selection box had changed and then it would recreate the top or bottom graph accordingly. I also made it so the bar graphs would activate when a region was clicked and I added a little bit of animation so the visualization fades in when the page loads.

I knew my dataset had a lot of information to represent, so I knew beforehand I wanted to combine all 3 of my variations to come up with a semi final design. This is why I decided on a dashboard visualization which allows me to represent that data in multiple ways and it allows the data to be viewed as a whole. The implementation itself allowed me to represent the data in ways that I couldn't have done with sketches and overall I am happy with the result.

3.7 Final Static Design (Already had implemented in part 3 so its the same)



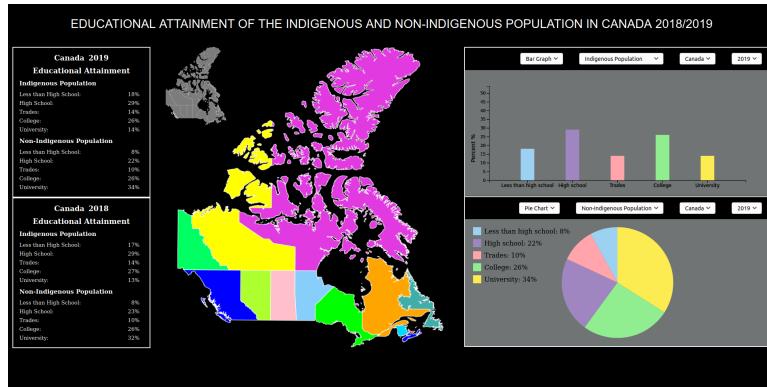
3.8 Prototyping Interactions

- A) Click on map region: When the user clicks on a province or full Canada map (top left) it will display the educational attainment data for both population groups in the dialog box located on the left side of the screen. It will display the data for 2019 in the top box and the data for the 2018 in the bottom box. Also when a region is clicked the data will also be displayed in bar graph form where the top bar graph defaults to the Indigenous population and the bottom bar graph defaults to the Non-Indigenous population. Both bar graphs default to the year 2019, so if the 2018 data needs to be viewed it needs to be done manually.
- B) Drop down menus to control graphs: The user can use the drop down menus to control all variables of the dataset and it allows the user the ability to control the graph type. The variables the user can control are the population groups, regions, and years.

4. Final Implementation Visualization

Project Links:

- 1) Live Examples: <https://michaelrediron.github.io/CPSC583-DataVisualization/>
- 2) Code Repo: <https://github.com/MichaelRediron/CPSC583-DataVisualization>



My visualization allows the viewer to visualize the dataset from a holistic perspective. By using multiple views it allows the viewer to easily explore and compare the data. For example by using a map view it allows the user to easily click on the region they're interested in which will then display the data clearly in the dialog box. Also by allowing the user the ability to control the dataset variable in the graph section it gives them the ability to compare different scenarios when analyzing the data. For example the user can compare educational attainment of the Indigenous population data for Alberta and Manitoba. Or if they want to view the data of one region they can select that region on both graphs and change the years or population groups to align with what they are trying to compare. Using the visualization the viewer will be able to

easily understand the concept the visualization is trying to convey which is that the educational attainment of the Indigenous population could be improved.

4.1 Process Reflection

My dataset had a lot of information to display, so I knew it would be hard to display it all using just one representation. This is why I decided to use multiple views on my final visualization which allowed me the ability to better portray the dataset as a whole and convey the message of the visualization in a complete way. By adding interactions it helped give a better overall experience and it allowed the visualization to convey information that a static visualization could not. Overall I am happy with the interactions that I implemented and I feel they are very beneficial to the final design.

5. Discussion

To reflect on my final design I feel that it accurately represents the data and the message that I was going for. If I had more time I would have added more graph options to the graph sections which would allow the user to view the data from even more perspectives. One more feature I would have liked to add is hover tooltips for the graphs and map region. Also I would have liked to make the visualization responsive that way it would remain the same on all devices and browsers, so if it is a bit off on your device that may be why. Some advantages of the visualization are that it allows the user to control the variable when comparing data and it gives multiple views which provides a better perspective to the viewer. I learned a lot about the design process of creating a good visualization and how to represent data that is beneficial and clear to the viewer. I now understand that there is a lot more to the process than it seems when creating a good visualization.

6. Conclusion

One of my main achievements that I had during this project is that I feel that I now have more of a grasp of the overall design process when creating a visualization, which will be very beneficial to my future. A lot of the things I have learned will be applicable to projects I will have in the future such as creating a website. One other main achievement I had was that I feel that I was able to give a good representation of the message I was trying to portray when creating the visualization. Overall this project was a valuable learning experience and I am thankful for the skills I have learned.