Analyzing Voter Demographics in Yukon through Surveying and Stratified Random Sampling

Zakir Chaudry, Ankhee Paul, Shupeng Chen, Yi Peng Chang

10/07/2020

Executive Summary

Yukon, Canada's westernmost and one of the least populated provinces, has long been neglected by political parties for a variety of reasons. However, Yukon's importance in nature conservation and mining development is self-evident. In this paper, we will explore Yukon's voter data set from a survey conducted by our team. With the support of the Yukon Liberal Party, we successfully gathered comprehensive data of the Yukon population which include demographics, political preferences, income levels, and educational backgrounds. We divided voters into groups based on their educational background, where they live and which policy they support and then analyzed which party they will vote for. We found some very useful results, for instance, people who live in Pelly-Nisutlin prefer the Liberal Party. The New Democratic Party's supporters care more about the environment and so on. However, we recognize that the results of the current analysis represent a preliminary observation of the limitations inherent in such a large record. In addition to the limitations of using large survey data sets discussed earlier, this study is also constrained by budget. For example, most of our data are derived from densely populated urban areas, thus ignoring the universality of the data. So, our future work will focus on understanding and measuring the sparsely populated suburbs of Yukon. In conclusion, with reliable data sources and our team's efficient data analysis methods, our findings are sure to provide a valuable reference for the Yukon Liberal party in the 2020 federal election.

Introduction

Throughout this report, we approach the task of analyzing the demographics of voters in Yukon with the aim to provide supporting information to the Yukon Liberal Party to aid policy decisions and re-election efforts. Initially, we discuss the survey methodology, and the path that took us to our final choice of method, that being phone surveys. Moving past that, we talk about the survey design itself and the reasons towards certain decisions. We then provide our findings and our interpretation of the results. Finally, we discuss potential weaknesses and biases and future opportunities.

Benefits for the Party

As victors of the previous elections, it is in the party's best interests to conduct this survey, in order to gage their chances of being re-elected. Our use of stratified random sampling from the population, based on electoral districts allows the party to recognize potential voters, and favorable districts. It also helps them to identify voters' concerns about relevant issues. The survey and the graphs depicted in our report give a clear comparison of the voter demographics, based mostly on districts. As a result, the party can compare statistics, identify relevant information - developing and improving their campaigning strategy accordingly, thereby, ensuring their possibility of being re-elected.

Survey Methodology

The population we focused on was eligible voters in Yukon. While there may be some merit to finding the opinions of Non-Citizens and those who will be eligible to vote in the election following, we found that in terms of the immediate helpfulness only those that will influence the upcoming election are people of interest. There were quite a few avenues through which to pursue how the survey would be done. Initially, we wanted to have surveys done in person at grocery stores. Upon completion of the survey, the respondent would have been entered into a raffle for a free laptop. We chose a laptop as a relatively expensive reward and one that appeals to the largest number of people without being something too basic, such as pure cash. A big issue with this methodology is the fact that not everyone attends a grocery store, and even if someone from a household comes, that doesn't necessarily mean they speak for every eligible voter that lives with them. This limits the frame to eligible voters that happen to go to grocery stores. However, the idea of a reward was appealing, so that maintained. Another method was phone surveys. Phone surveys could reach a person regardless of location. The next problem came in the form of what the frame would be. One option was simply generating every single local number and trying to call each one. This would be quite expensive, especially as many of these numbers wouldn't belong to anyone. However, the benefit of this would be a much wider range than if we had simply used census data, as we reach the unlisted numbers. Unlisted numbers make up a huge number of potential respondents than would be otherwise missed. For example, according to an article from the CBC, in 2017 Saskatchewan was home to more than 600,000 unlisted numbers, compared to their population at the time of about 1.2 Million. As time goes on, unlisted numbers will only grow more and more in number, and ignoring them would not only miss a huge number of people, but also the younger demographics who are much more likely to have an unlisted number. As such, we found this to be the best form of survey methodology. In terms of sampling, we chose Randomized Stratified Sampling with respect to electoral districts, as we felt this was the best way to represent the population, choosing roughly 10% of each district. Simple Random Sampling Without Replacement was also an option, but we wanted the districts to be a key point, so it was necessary that we had enough info from each. We also spread on online version of the survey as most people would rather use this instead of a phone survey. A potential issue that could, and most likely will, arise is Non-Response. However, in both having a reward for completing the survey, so that respondents would be more motivated to participate, and casting a very wide net, we believe that this a reasonable approach to mitigate this problem.

Survey Design and Motivation

The purpose of this survey is to categorize the citizens of Yukon by their income levels, educational background, and socio-political leanings. The categories can then be associated with the different districts of Yukon to prepare and inform the Yukon liberal party to come up with appropriate campaigning strategies.

The survey is hosted on Google Forms, making it easily accessible for almost everyone that has a device with internet access. It is short and contains mostly multiple-choice questions allowing easy completion. This survey is also accessible by phone, with personnel working for the Yukon Liberal party will conduct the survey over phone and fill in the Google Form on behalf of the respondent.

This survey discourages multiple submissions by requiring a valid email. The first 3 questions are to determine if the data obtained from this person is relevant for the upcoming election. If the person is under 18 or not a Canadian citizen, this respondent is not eligible to vote. However, their feedback/information can still be useful for further analysis of district backgrounds. The next 2 questions will ask about their sexual identity and preference, this determines the social aspect of the respondent. Then we ask about income and education, which allows further sorting and categorizing of respondents. The rest of the questions place the respondent on a political spectrum and obtain their political standings. In a few simple and basic questions, we acquired a respondent's social, economic, and political backgrounds.

Each respondent also has a correlated district if they decide to enter the raffle and answer the last question. The respondent's information can be used to construct overviews of each district's social, political, and economic situations. The Yukon Liberal party should use this information to assess their strategies for individual districts.
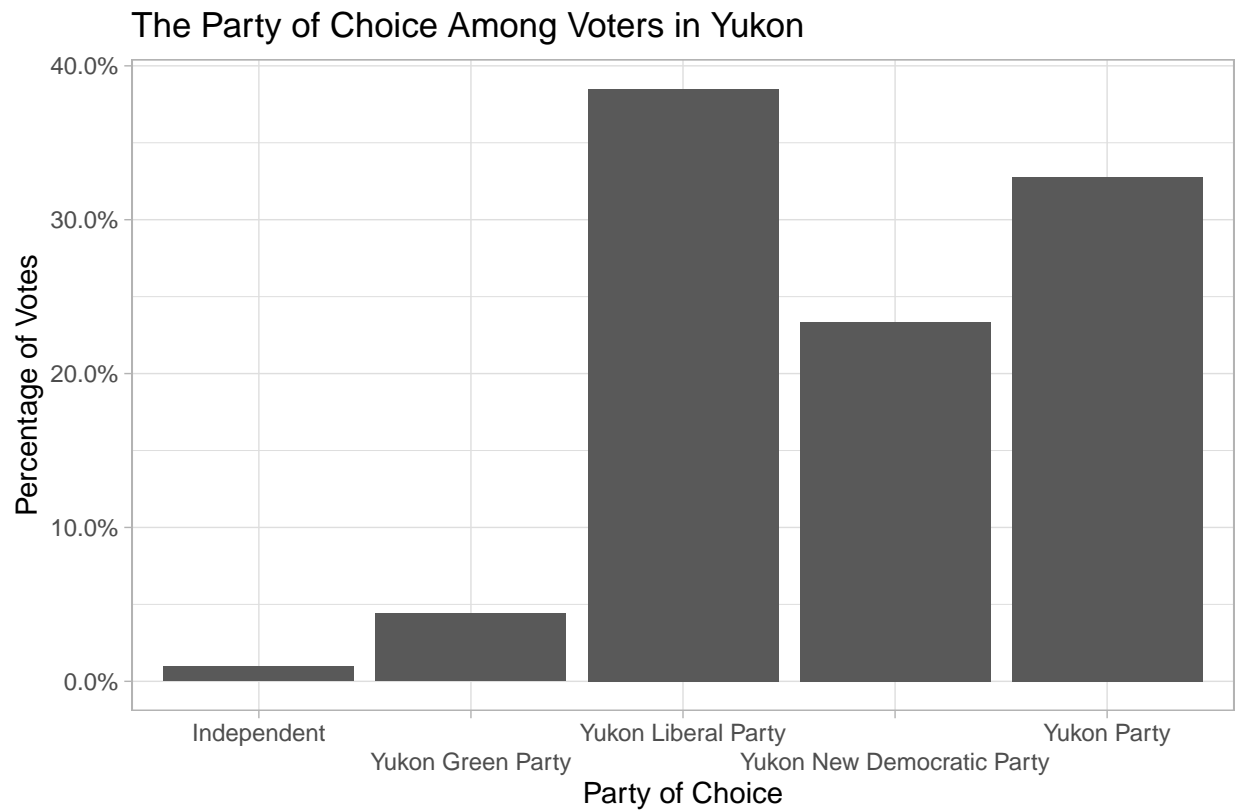
Findings

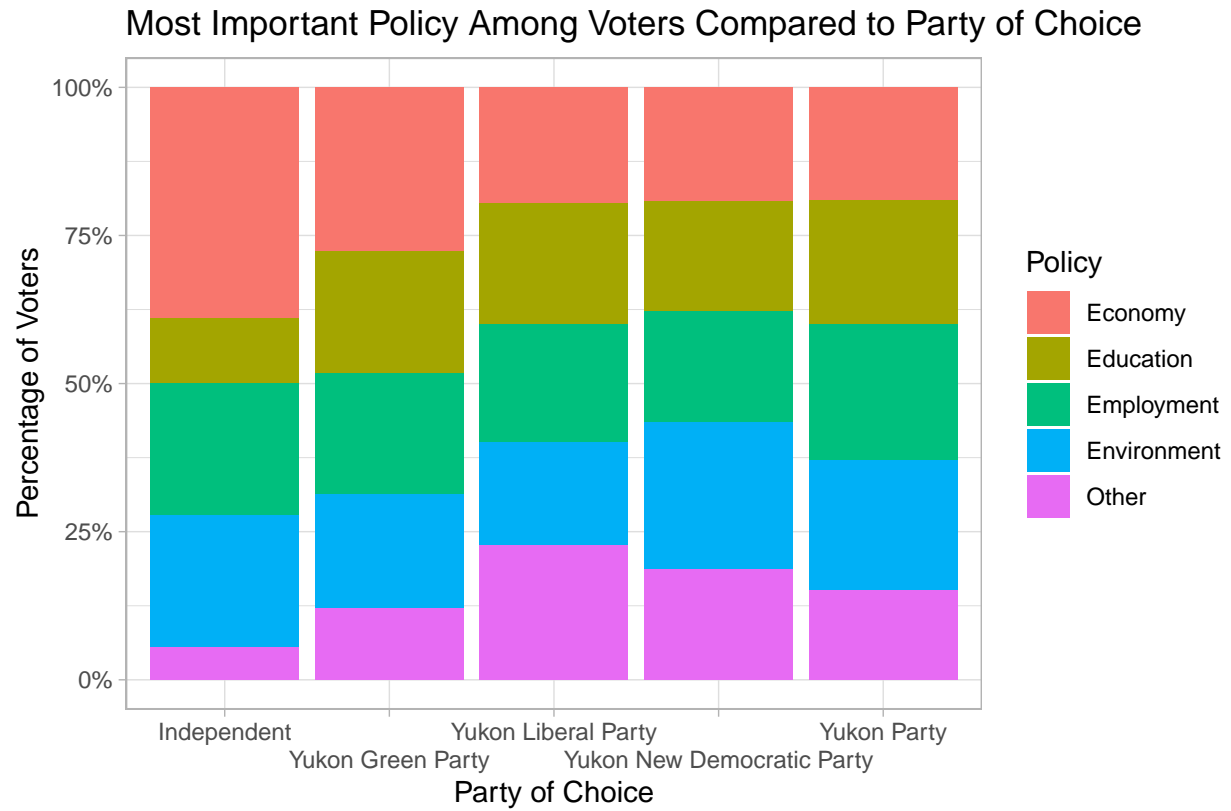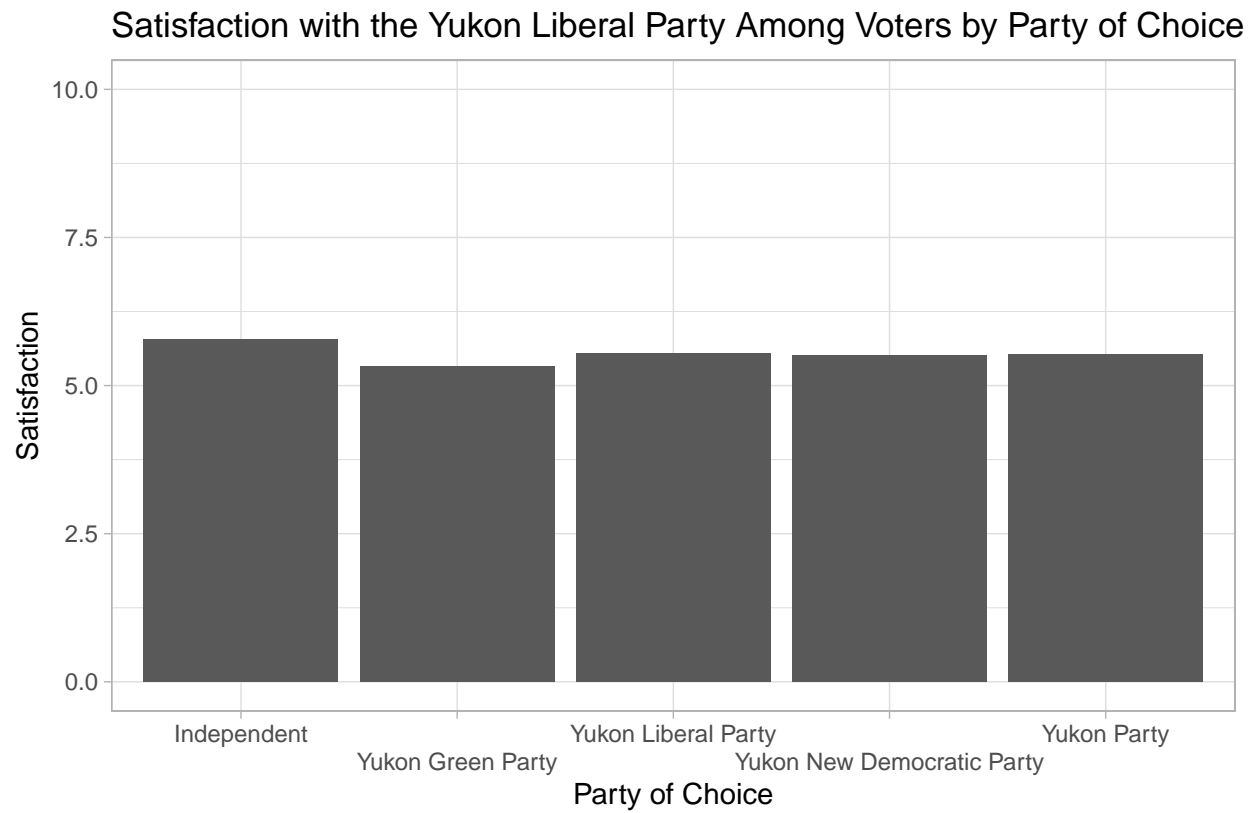## The Party of Choice Among Voters in Yukon

Most Important Policy Among Voters Compared to Party of Choice

Figure 2

Satisfaction with the Yukon Liberal Party Among Voters by Party of Choice

Figure 3

Confidence for the Yukon Liberal Party Among Voters by Party of Choice

Figure 4

Figure 5

Income Among Voters in Yukon by District

Figure 6

Satisfaction with the Yukon Liberal Party Among Voters by District

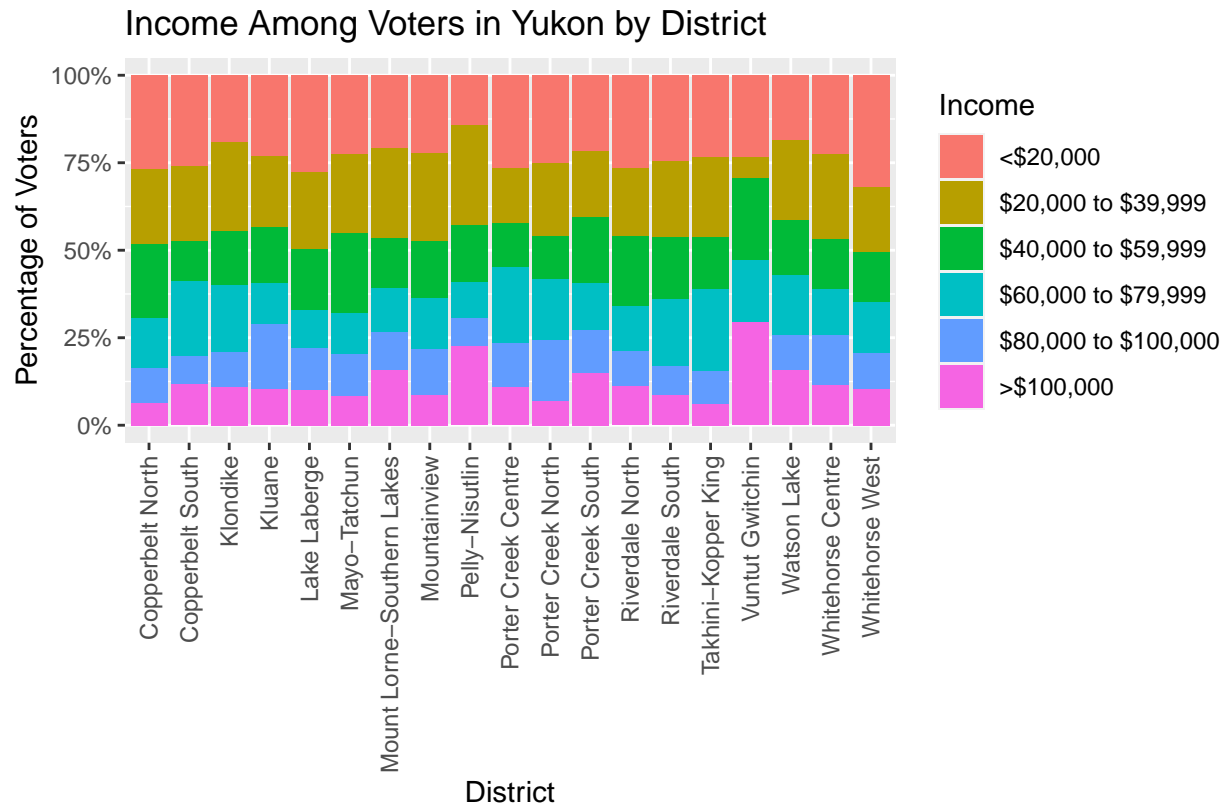Figure 7

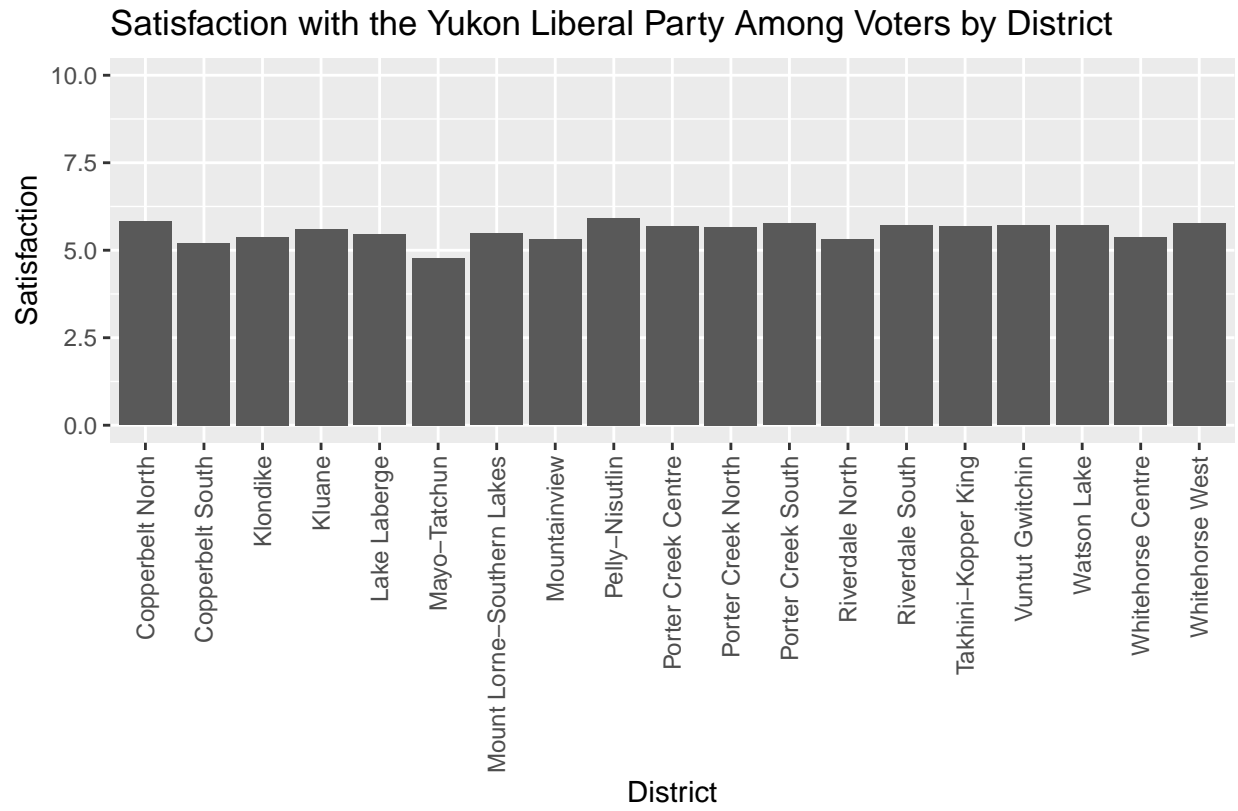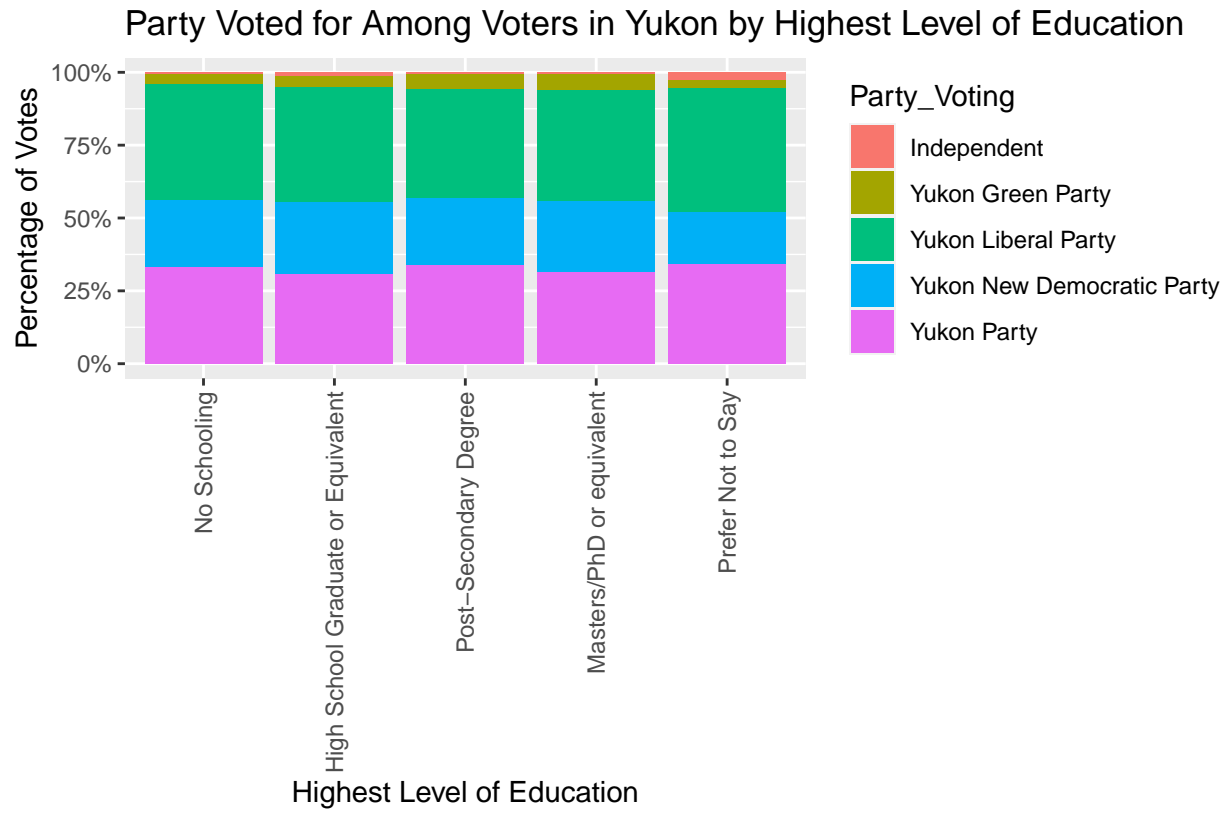Party Voted for Among Voters in Yukon by Highest Level of Education
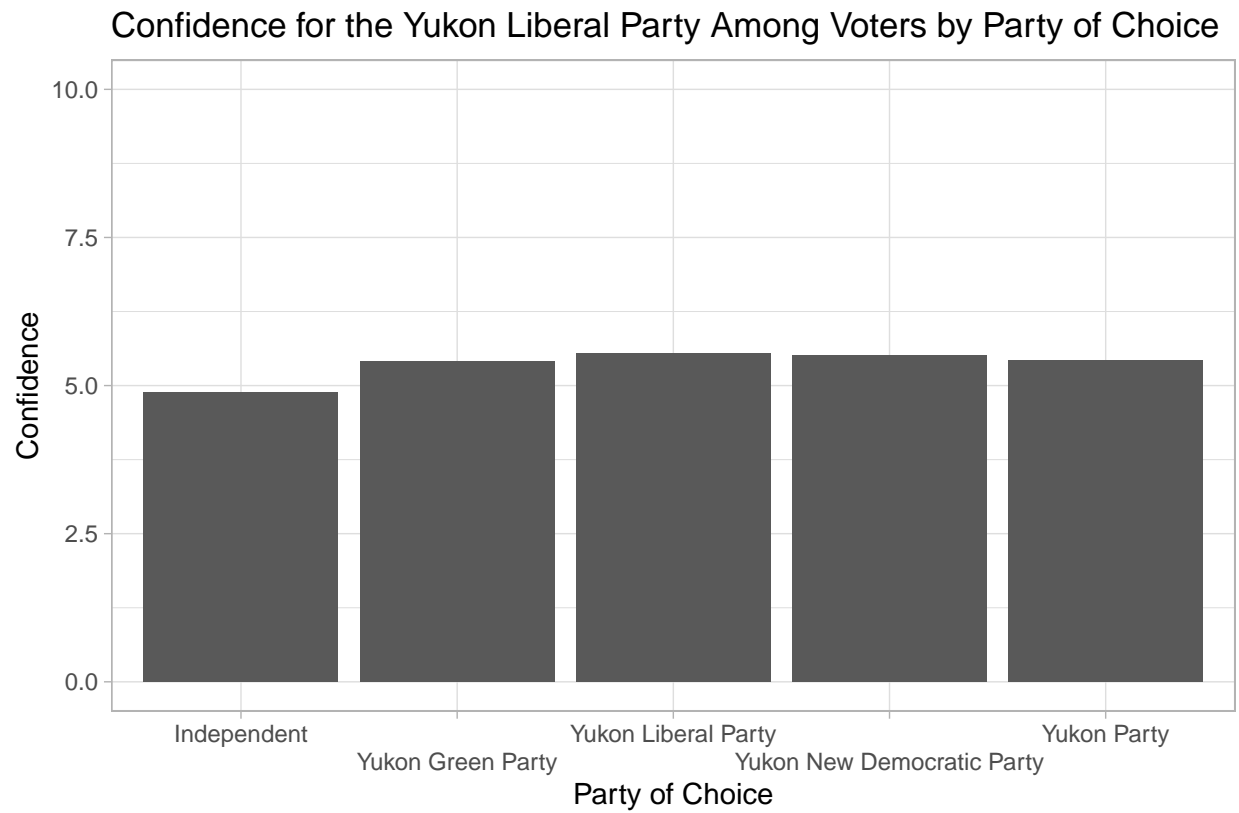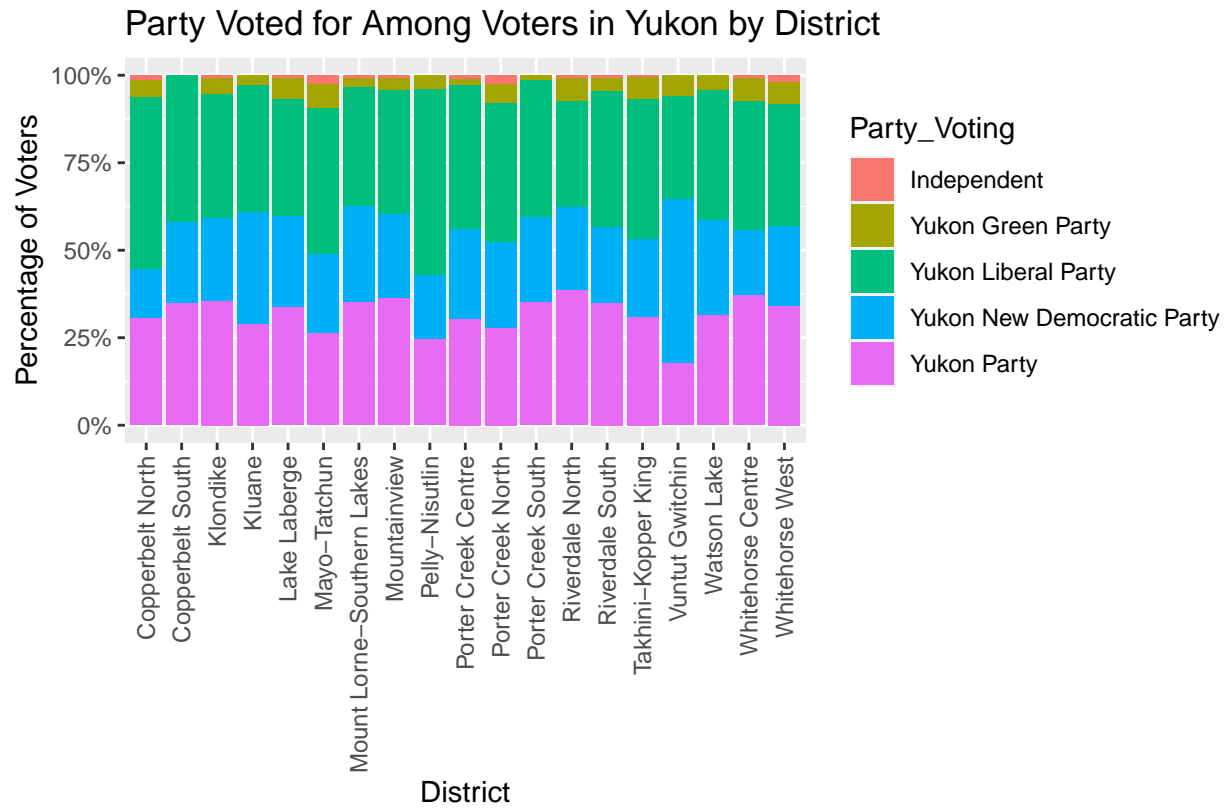
Figure 8

Highest Level of Education Among Voters in Yukon by District

Figure 9

Most Important Policy Among Voters in Yukon by District

Figure 10

Gender Among Voters in Yukon by District
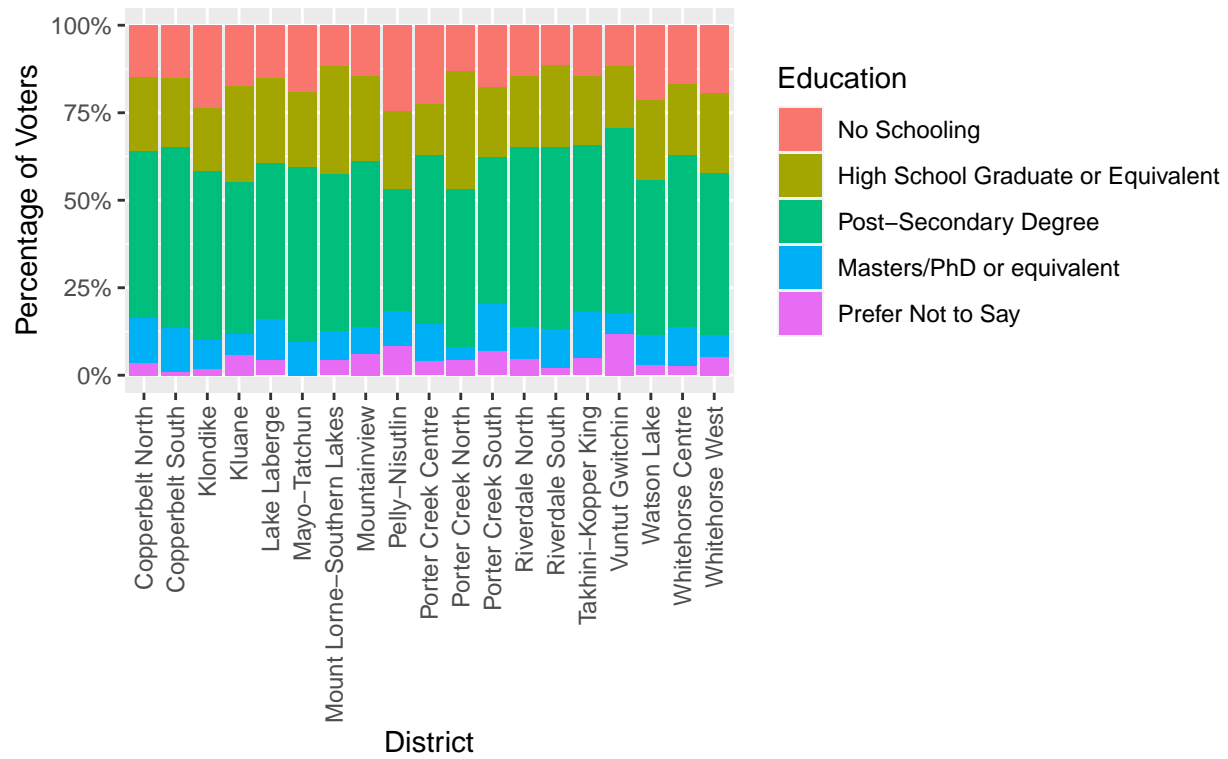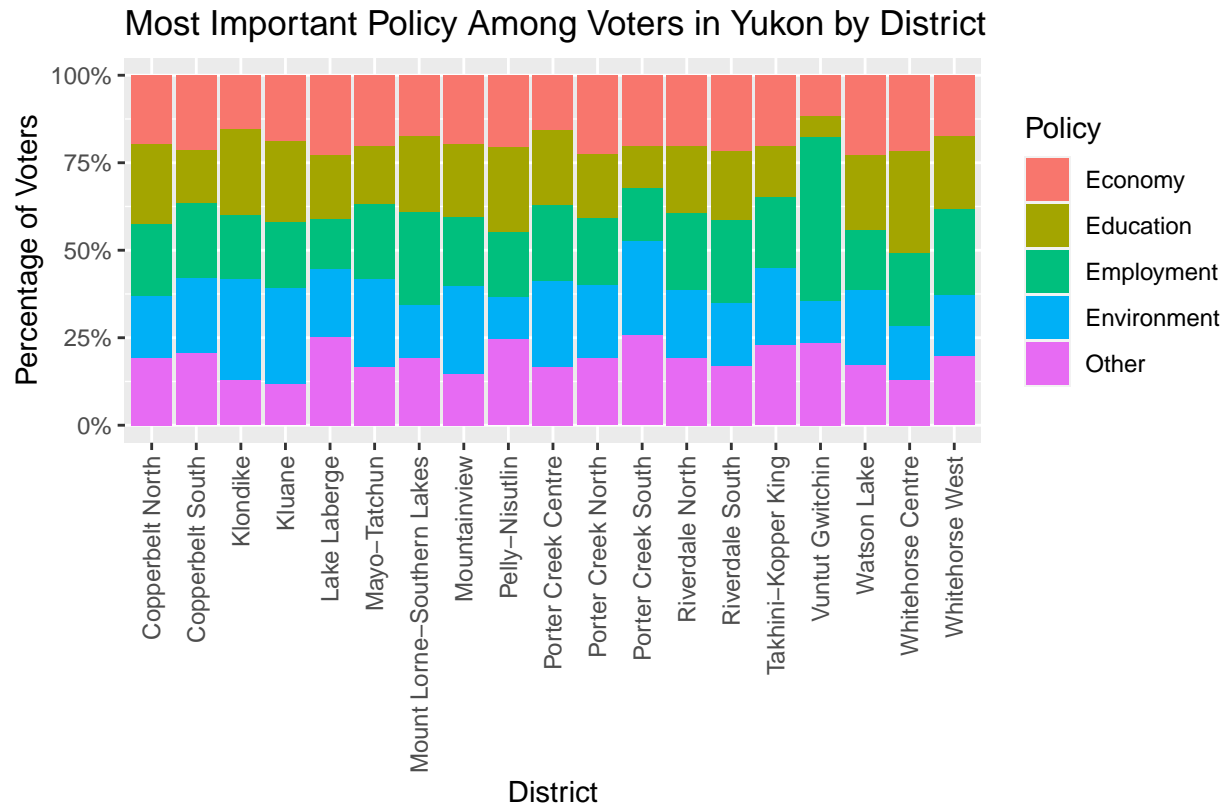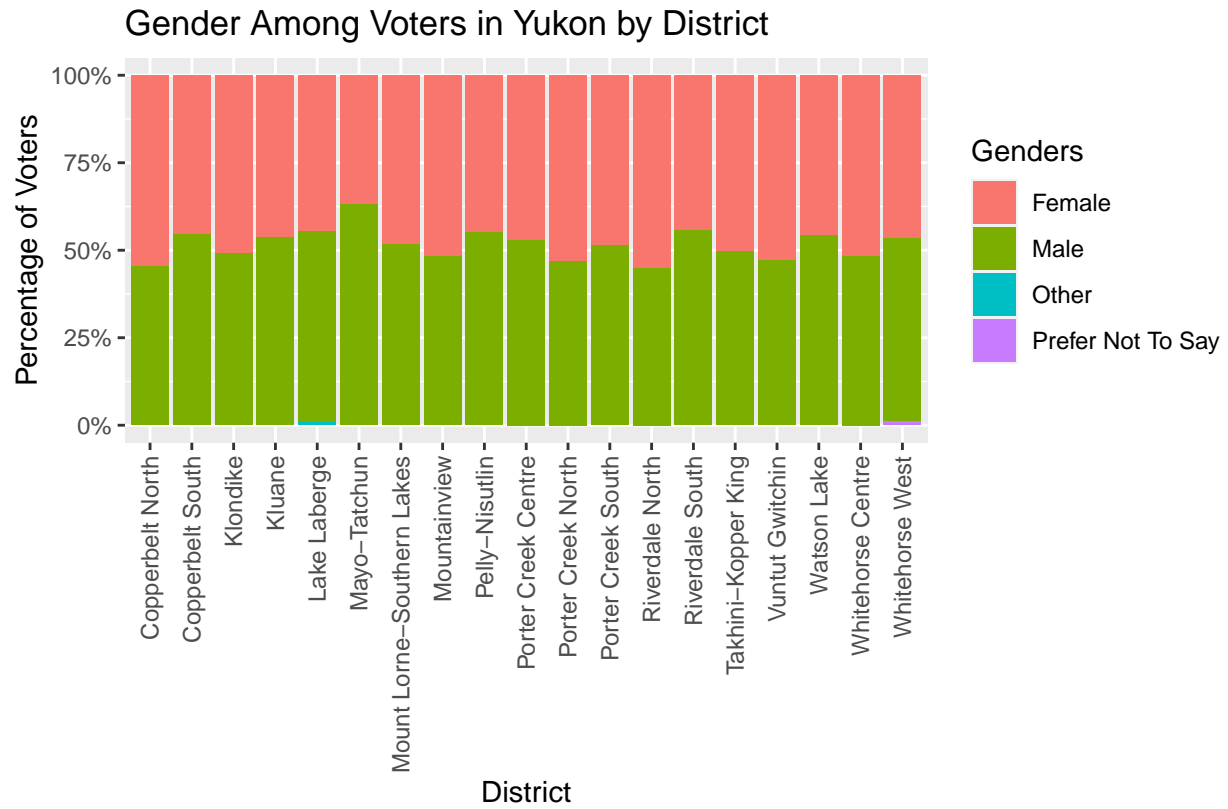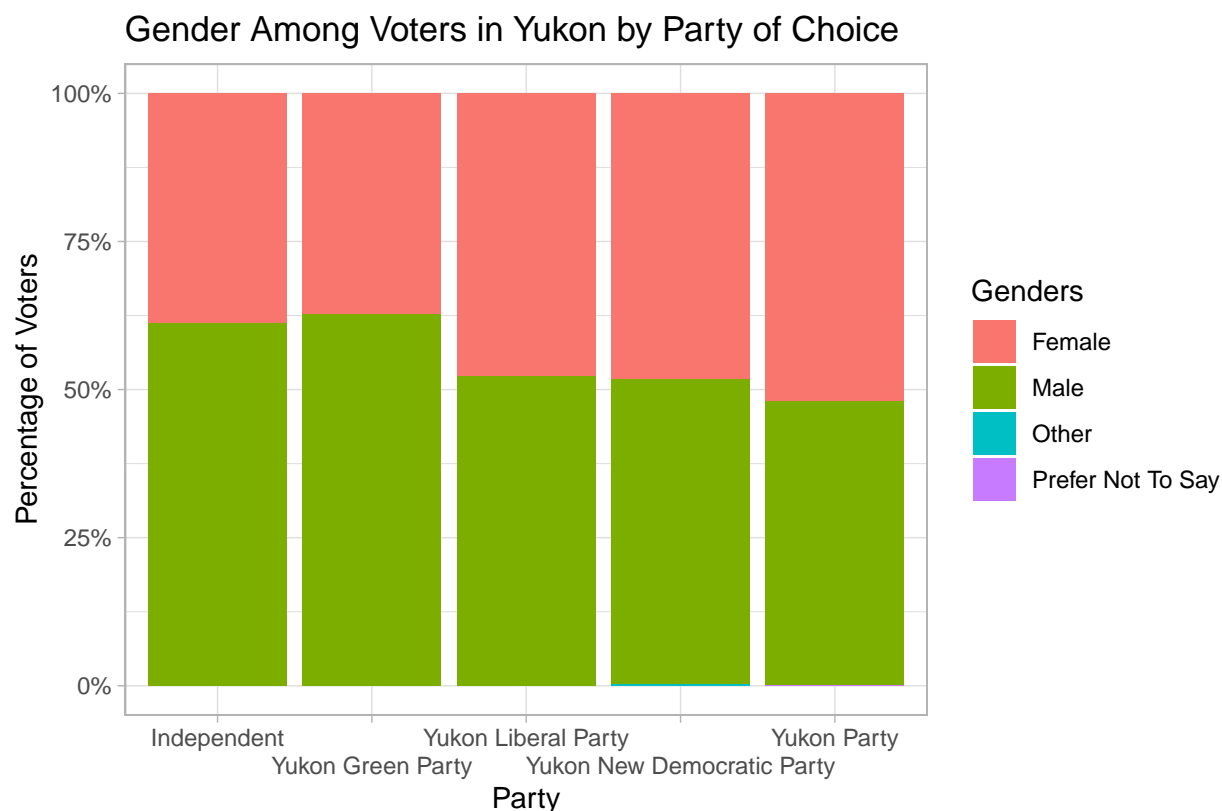

Figure 11

Figure 12

The main focus of this report is to share information regarding the voters of Yukon to the Yukon Liberal Party. Rather than give advice regarding particular policy decisions or campaigning strategies, we found it best to find and put together as much data as possible from which the Yukon Liberal Party can base their future actions off of. Given the material of the survey previously discussed, there were quite a few different points of interest to share.

Figure 1 dictates a fairly simple relationship regarding who voters plan to vote for. The results are fairly similar to those of the previous election. However, this does mean that while the Yukon Liberal Party has the lead, the other two main contenders are not far behind, so there's no room to rest on laurels at this time without maintaining the strength of the party and improving beyond that to reach a greater audience.

Moving forward, Figure 2 outlines the most important policy for voters based on which party they plan to vote for. As you can see, there's not too much variation between parties. That said, two key observations are that, as a group, Independent voters weigh the economy much heavier than other issues relative to the other parties. The other observation of note is that the percentage of those who plan to vote for the Yukon Party that find the environment to be their primary issue is greater than that of all other parties, whereas the voters who plan to vote for the Yukon Liberal Party tend to not consider it their highest priority, so this may be something to look into.

Figures 3 and 4 describe the satisfaction and confidence that voters have towards the Yukon Liberal Party based on their party of choice respectively. Most people are more satisfied than not across the grid, but not by a significant factor, so this is for sure an area to improve upon. Confidence is the same, with the exception of Independent voters being slightly less confident in the Yukon Liberal Party.

Figures 5, 6, 7, 9, 10, and 11 compare voters in each electoral district to their party of choice, their income, their satisfaction with the Yukon Liberal Party, their highest level of education, their most important policy, and their gender breakdown respectively. These figures serve to aid any campaigning efforts in each of these places, as understanding and appealing to these demographics in each district will surely contribute heavily to re-election efforts.

14

The remaining two plots, Figures 8 and 12, highlight supplemental information that we believed may be important, with Figure 8 describing how voters plan on voting based on their highest level of education and Figure 12 based on their gender.

We are certain that the information obtained from these findings will be of great use during future policy planning and campaigning efforts.

Shortcomings and biases

Assuming that nobody lies on survey for the sake of the laptop, there are multiple weaknesses which might lead to inconclusive results. There is an inevitable problem of missing data as people may choose not to answer specific questions or take the survey, invalidating generalization to the population. Additionally, the method of distribution remains insufficient to reach every single inhabitant of the province, depriving us of valuable information/data. By trying to reach everyone through their phones, there might arise the problem of overlapping information as respondents could possess multiple phone numbers. External factors (or lurking variables) such as the economic conditions caused by the current pandemic, and policies considered personally irrelevant amongst the sample, can lead to biased answers against the party. Moreover, respondents can find the survey tedious due to its length, rendering inaccurate and unfavorable answers.

Code and data supporting this analysis is available at https://github.com/ZakirChaudry/PS2/

Appendix

Simple Random Sampling: A sampling technique where everyone is chosen entirely by chance and each member of the population has an equal chance of being included in the sample.

Stratified Random Sampling: A method of sampling that involves the division of a population into smaller groups, the smaller groups are formed based on members' shared attributes or characteristics, and everyone of each group has an equal chance of being included in the sample.

Link to Survey: https://docs.google.com/forms/d/e/1FAIpQLSeKtBfexOZDvcqszeNDH_dL5x3bkHHV4jNet4fdUS9Iv5-V3w/viewform?usp=sf_link

Screenshots of the Survey:

Yukon Liberal Party General Election

We want to know how you feel about the Yukon liberal party for the upcoming election. Participation in this survey is completely voluntary and the data you provide will be anonymous. Withdrawl from the survey at any time will not penalize you. You can choose to enter the raffle for a 2020 Dell XPS13 at the end of the survey.

* Required

Email address *

Your email

Do you consent for your answers to this survey to be analyzed by data scientists working for the Liberal Party of Yukon *

○ Yes

○ No

Are you a Canadian citizen *

○ Yes

○ No

What is your date of birth *

Date

yyyy-mm-dd

What gender do you associate with *

What gender do you associate with *

○ Male

○ Female

○ Prefer not to say

○ Other: _____

What is your sexual preference *

○ Heterosexual

○ Homosexual

○ Bisexual

○ Prefer not to say

○ Other: _____

What is your level of education *

○ No schooling

○ Secondary education

○ Bachelor

○ Master

○ Doctoral

○ Other: _____

Income *

Income *

○ <$20,000

○ $20,000 - $39,999

○ $40,000 - $59,999

○ $60,000 - $79,000

○ $80,000 - $100,000

○ >$100,000

Are you planning to vote in the upcoming election? *

○ Yes

○ No

○ Unsure

Which policies are you most concerned with *

○ Environment

○ Employment

○ Economic

○ Education

○ Other: _____

How satisfied are you with the liberal party *

	1	2	3	4	5	6	7	8	9	10	
very dissatisfied	○	○	○	○	○	○	○	○	○	○	satisfied

How satisfied are you with the liberal party *

	1	2	3	4	5	6	7	8	9	10	
very dissatisfied	○	○	○	○	○	○	○	○	○	○	satisfied

How confident are you in the liberal party winning *

	1	2	3	4	5	6	7	8	9	10	
not very confident	○	○	○	○	○	○	○	○	○	○	very confident

What are some concerns you have for the liberal party *

Your answer

Which party are you planning to vote for *

○ Yukon Liberal

○ Yukon party

○ Yukon New Democratic

○ Yukon Green

○ Independent

Which party did you previously vote for *

○ Yukon Liberal

○ Yukon party

○ Yukon New Democratic

What are some concerns you have for the liberal party *

Your answer

Which party are you planning to vote for *

○ Yukon Liberal

○ Yukon party

○ Yukon New Democratic

○ Yukon Green

○ Independent

Which party did you previously vote for *

○ Yukon Liberal

○ Yukon party

○ Yukon New Democratic

○ Yukon green

○ Independent

Optional: for a chance to win Dell XPS13. What district do you live in

Your answer

Submit

Page 1 of 1

20

References

Forrest, Maura. 2015. "Yukon Voter Turnout Second Highest in the Country." *Yukon News.* Yukon News. https://www.yukon-news.com/news/yukon-voter-turnout-second-highest-in-the-country/.

Government of Canada, Statistics Canada. 2017. "Same-Sex Couples and Sexual Orientation... By the Numbers." *Statistics Canada.* Government of Canada. https://www.statcan.gc.ca/eng/dai/smr08/2015/smr08__203__2015.

———. 2019. "Census Profile, 2016 Census Yukon [Territory] and Canada [Country]." *Statistics Canada.* Government of Canada. fm?Lang=E&Geo1=PR&Code1=60&Geo2=&Code2=& SearchText=Yukon&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR& amp;GeoCode=60&type=0.

Mahto, Ananda. 2019. *Splitstackshape: Stack and Reshape Datasets After Splitting Concatenated Values.* https://CRAN.R-project.org/package=splitstackshape.

McKee, Lori. 2017. "Report of the Chief Electoral Officer of Yukon on the 2016 General Election." *Elections Yukon.* The Chief Electoral Officer of Yukon. https://electionsyukon.ca/sites/elections/files/english__website__2016__election__report__1.56.55__pm.pdf.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2020. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Dana Seidel. 2020. *Scales: Scale Functions for Visualization.* https://CRAN.R-project.org/package=scales.

Xie, Yihui, Christophe Dervieux, and Emily Riederer. 2020. *R Markdown Cookbook.* Boca Raton, Florida: Chapman; Hall/CRC. https://bookdown.org/yihui/rmarkdown-cookbook.

Zakreski, Dan. 2017. "Sask. Home to More Than 600K Unlisted Numbers." *CBC News.* CBC/Radio Canada. https://www.cbc.ca/news/canada/saskatoon/saskatchewan-phones-411-1.4013540.