Political Orientation of Canadian University Students

Alen Mitrovski, Harry Hwang, Xiaoyan Yang

10/8/2020

Executive Summary

Petit Poll, a Canadian polling company, has been asked by the Liberal party of Canada to survey university students across the county in order to better understand their current political party orientation. We designed a 16-question survey that would highlight key areas of interest for the political party. The questions focused on areas of potential improvement such as tuition rates, the job market and purchasing power. With the results of this survey, the Liberal party should have a clear indication of what they need to implement to gain student votes in the following election campaign.

Through the Neyman Allocation theorem, the sample size of the study was calculated. This theorem produced the ideal sample size for each given stratam considering their relative weight to the total sample size survey. The target population of this study is Canadian university students. Students were selected from the University of Toronto, University of Alberta, University of British Columbia, and McGill University to form the sample frame. The sampled population in our study is represented by the 1000 students from each of the four Canadian universities that completes the study survey. The stratified sampling method was utilized to conduct this survey. The survey was first sent, via email, to half of the calculated student sample size for each university. If the minimum desired respondent level was not reached, a second set of surveys would be sent to the other half of the sample university students. Through this procedure we collected 1000 student responses per university.

The data used in this study was simulated to most accurately depict the current university population in Canada. We focused our results on three key questions: "Should the government lower tuition costs?", "How confident are you to purchase a home in your current province in 10 years?", "Are you satisfied with the state of the job market in your current province?". We graphed the results of these specific responses and compared between the universities to examine if there is a difference in opinion depending on their geographical location. We found that all university students, from each surveyed university, agree that tuition costs should be lowered. Additionally, we found that students from McGill university, located in Quebec, had very low confidence in purchasing a property in 10 years and low satisfaction with the current job market in their province. In contrast, University of Alberta students, located in Alberta, has very high confidence in purchasing property in 10 years and high satisfaction with the job market in their province. These statistics seem to correlate with the general high median income in Alberta and low housing index.

This data is important as it indicates key areas that the Liberals can look to improve upon to gain student votes before the next election. Quebec is the 2nd largest province in Canada, in terms of population, accounting for 22% of the total Canadian population (Population Estimates, 2020). The government should focus on the concerns presented by McGill university students in terms of finding work in the provincial job market and purchasing reasonably priced property. This new generation of voters could be critical in 3 years when the next election takes place. In addition, there is agreeance among all university students (regardless of location) that tuition costs should be lowered. The Liberals can look to implement new programs, in conjunction with the universities, to provide students with a reasonably priced education that they will be satisfied with. Considering the consensus opinion on this national student issue, the government would be wise to focus their resources on gaining student support on this issue.

One limitations of this paper is the possibility of dishonest answers. A method to eliminate this issue was not

implemented in this study. Another limitation is the potential sampling bias considering we haven't further stratified the populations to account for male vs female or ethical discrepancies. In the future, researchers should conduct a follow up study that accounts for these differences in opinions among students. Additionally, implementing strategies that can determine whether a student is answering honestly will allow researchers to discard any invalid data. Future follow-up studies could look to gather data from the maritime provinces and territories which were not covered in this study. Furthermore, researchers can aim to perform regression analysis on the survey responses in this study to find correlating relationships between the survey responses of various questions.

Introduction

Elections are conducted in a manner which grants citizens the right to vote and choose between the political parties. In the 2019 Canadian federal election, the Liberal party won the election by securing a minority government despite the Conservatives winning the popular vote (CBC, 2019). The Liberal party is now looking to gain a better understanding of the political orientation of the Canadian student population so that they can plan for the following election campaign where they will look to regain a majority government. The following study was conducted on behalf of the Liberal party of Canada to better understand the political orientation of university students in Canada. The data used in this study was simulated to depict an accurate representation of the student population in the country. In order to collect data, surveys were sent out, through email, to students from the University of Toronto, University of Alberta, University of British Columbia, and McGill University. These universities are located in the four provinces (Ontario, Quebec, Alberta and British Columbia) that make up 86% of the total Canadian population and thus, sufficiently represent all university students in Canada (Population Estimates, 2020).

On a national scale, we will be examining the political allegiances of students and their perspectives on key political topics. On a provincial scale, we will be comparing the results of certain factors between the different university's students in order to understand the geographical differences of their responses. In this paper, we will discuss the following: 1. The survey methodology 2. The results of the surveys collected 3. The relevance of the data 4. The limitations of this study 5. The next steps for future researchers

Full code and data supporting this analysis is available at: https://github.com/yang-xiaoyan/Political-Orientation-of-Canadian-University-Students-simulated-

Survey Methodology

First, it is important to define the parameters of this student political party orientation study. A target population is defined as the set of all units covered by the main objective of the study (Wu, 2020). In this study, the target population is Canadian university students. A sampling frame is defined as the list of units in the survey population (Australian Bureau). For the purpose of this study, students were selected from the University of Toronto, University of Alberta, University of British Columbia, and McGill University to form the sample frame. A sampled population, also known as a study population, is defined as the population represented by the survey sample (Wu, 2020). The sampled population in our study is represented by the 1000 students from each of the four Canadian universities that completes the study survey.

In this study, the stratified sampling method was utilized in order to best represent the overall Canadian university student population. Stratified sampling is a sampling method that divides a population into X number of subpopulations (Wu, 2020). In this study, the target population of Canadian university students was stratified into four subpopulations. Each subpopulation represented one of four selected universities (University of Toronto, University of Alberta, University of British Columbia, and McGill University). These universities are located in the four provinces (Ontario, Quebec, Alberta and British Columbia) that make up 86% of the total Canadian population and thus, provide an accurate representation of the target population (Population Estimates, 2020).

As per Wu and Thompson (2020), there are four reasons to utilize stratified sampling – Administrative Convenience, Estimation of Subpopulation Parameters, Efficiency Considerations and More Balanced Samples (Wu, 2020). In terms of administrative convenience, our study was designed to survey four major universities that are representative of the provinces that make up 86% of the Canadian population. In our case, it was much more convenient to stratify our target population in this manner as it allowed us to remain organized throughout and work more efficiently with a smaller dataset. In terms of estimation of subpopulation parameters, stratifying our study has allowed us not only to extrapolate and make estimations on a national scale, but also to examine properties of our subpopulations (Students at each selected university). As Wu and Thompson write, efficiency considerations are a benefit of stratified sampling as it provides more accurate estimates and efficient inferences (Wu, 2020). As it pertains to more balanced samples, designing our study in this manner has allowed us to ensure that we obtain a relatively proportionate representation of Canadian university students as we are obtaining an equal amount of responses from each of the 4 selected universities. For example, if this study was not stratified, we may have received a large amount of responses from the University of Toronto but very few from the University of Alberta. As a result, our data would not be an accurate representation of students in Canada as we are incorporating geographically biased results.

In order to calculate our sample size for the study, we referred to the Neyman allocation theorem under stratified simple random sampling (Appendix A). This theorem aims to produce the ideal sample size for a given stratam considering their relative weight to the total sample size. For our purposes, historical data was utilized in the calculation of the sample mean and variance. As shown in Wu and Thomson, the Neyman allocation theorem is demonstrated below:

Define:

Let 1. N_h denote the stratum population size 2. n_h denote the stratum sample size proportional to population size 3. h denote the number of stratum in naturals from 1,2,... to H 4. $\bar{y_h} = \frac{1}{n_h}\sum_{i\in\mathcal{S}_\langle} y_h i$ denote stratum sample mean under stratified sampling 5. $\bar{yst} = \sum_{h=1}^{H} W_h \bar{y_h}$ denote the stratified sample mean estimator under stratified simple random sampling 6. $E(\bar{y}st = \mu_y$ denote the population mean, and the stratified sample mean $\bar{y}st$ is anunbiased estimator of the population mean 7. $y_h i$ denote as the value of the study variable y for unit i in stratum h,i=1,2....,N_h 8. $s_y^2 h = \frac{1}{n_h - 1}\sum_{i\in\mathcal{S}_\langle}(y_h i - \bar{y}_h)^2$ denote as stratum sample variance 9. $v(\bar{y}st) = \sum_{h=1}^{H} W_h^2(1 - \frac{n_h}{N_h})(\frac{s_y^2 h}{n_h})$ denote as an unbiased variance estimator for $\bar{y}st$ 10. $W_h = N_h/N$ denote as the stratum weight, which is the relative size of the stratum within the overall population. 11. $\sum_{h=1}^{H} N_h = N, \sum_{h=1}^{H} W_h = 1$ 12. $n_h = \frac{n}{N}N_h = nW_h, h = 1, 2, ..., H$

Through the utilization of this theorem, we calculated the following sample size for each university: - University of Toronto (2706 students) - University of British Columbia (2113 students) - University of Alberta (1749 students) - McGill University (1430 students)

In order to reach our desired respondents, we (1) obtained consent from the four universities to collect data from their students via survey. As per our request, we were (2) given the email addresses of randomly selected, and currently registered, students from each university (5412 students for UofT, 4225 students for UofBC, 3498 students for UofA and 2860 students for McGill). Our initial step was (3) sending our survey of 16 carefully selected questions to half of these students via email. This was performed for each of the select universities. A one-week deadline was implemented. If a minimum of 1000 responses were not collected from each university during this initial period, then (4) the other half of students would be sent the survey via email. Another one-week deadline was implemented for this sample of students. If the target of 1000 responses per university was still not achieved, then (5) the responses that were collected up to that point would be used in the study analysis. The time span of this survey data collection was 4 weeks.

The estimated cost of this study was dependent on the number of respondents. We rewarded each successful respondent with a $2.00 digital coupon redeemable at a recognizable Canadian coffee shop. This coupon was attached to the congratulatory page of a completed survey. The maximum completed responses after the first round of survey dissemination was 8000 for the total of the four universities. In addition, the desired minimum survey response was 1000 per university and 4000 for the total four universities. Therefore, estimated range of cost was between $8000 and $16 000. Labor cost was excluded as we are paid employees by Petit Poll.

In order to protect respondents' privacy, non-disclosure clauses were signed with the select universities and non-disclosure declarations were provided in survey emails. As per the ethical and legal policies of Petit Poll, the company will never disclose any personal information to a third party without consent, or request, and each person who participated in our survey is protected by Canadian Privacy Act.

Results

```
survey_size <- 4000
```

```
# U of Alberta
set.seed(441)
uofa <- tibble(
  age = rpois(n = survey_size_alb, 20),
  citizen = sample(x = c("yes","no"), size = survey_size_alb,
                   replace = TRUE, prob = c(0.05, 0.95)),
  tuition = sample(x=c(1:5), size = survey_size_alb, replace = TRUE,
          prob= c(0.0,0.04,0.16,0.2,0.6
          )),
  hours = rnorm(n = survey_size_alb, mean = 11.5, sd = 1),
  wage = sample(x = c("yes","no"), size = survey_size_alb,
                   replace = TRUE, prob = c(0.25, 0.75)),
  income = sample(x=c(1:5), size = survey_size_alb, replace = TRUE,
          prob= c(0.15,0.15,0.35,0.2,0.15
          )),
  hospital = sample(x = c("less","more","same"), size = survey_size_alb,
                   replace = TRUE, prob = c(0.5, 0.1, 0.4)),
  purchase = sample(x=c(1:5), size = survey_size_alb, replace = TRUE,
          prob= c(0.05,0.1,0.2,0.27,0.38
          )),
  future = sample(x=c(1:5), size = survey_size_alb, replace = TRUE,
          prob= c(0.42,0.31,0.07,0.08,0.02
          )),
  taxes = sample(x = c("yes","no"), size = survey_size_alb,
                   replace = TRUE, prob = c(0.25, 0.75)),
  vote = sample(x = c("yes","no"), size = survey_size_alb,
                   replace = TRUE, prob = c(0.4, 0.60)),
  campus = sample(x=c(1:5), size = survey_size_alb, replace = TRUE,
          prob= c(0.1,0.15,0.15,0.25,0.35
          )),
  job = sample(x=c(1:5), size = survey_size_alb, replace = TRUE,
          prob= c(0.1,0.1,0.2,0.2,0.4
          )),
  news = sample(x = c(1:5), size = survey_size_alb,
                  replace = TRUE, prob = c(0.15, 0.23, 0.26, 0.23, 0.15)),
  interest = sample(x = c("yes","no"), size = survey_size_alb,
                   replace = TRUE, prob = c(0.35, 0.65)),
  party =  sample(x = c("Liberal","Conservative", "NDP","Green","Bloc Québécois","No Response"), size =
                  replace = TRUE, prob = c(0.13,0.7,0.1,0.03,0.067,0.0)),
)
```

```
# UBC

set.seed(541)
```

```r
uofb <- tibble(
  age = rpois(n = survey_size_ubc, 19.5),
  citizen = sample(x = c("yes","no"), size = survey_size_ubc,
                   replace = TRUE, prob = c(0.25, 0.75)),
  tuition = sample(x=c(1:5), size = survey_size_ubc, replace = TRUE,
          prob= c(0.0,0.035,0.165,0.15,0.65
          )),
  hours = rnorm(n = survey_size_ubc, mean = 10.5, sd = 1),
  wage = sample(x = c("yes","no"), size = survey_size_ubc,
                   replace = TRUE, prob = c(0.2, 0.8)),
  income = sample(x=c(1:5), size = survey_size_ubc, replace = TRUE,
          prob= c(0.05,0.017,0.3,0.23,0.25
          )),
  hospital = sample(x = c("less","more","same"), size = survey_size_ubc,
                   replace = TRUE, prob = c(0.45, 0.2, 0.35)),
  purchase = sample(x=c(1:5), size = survey_size_ubc, replace = TRUE,
          prob= c(0.2,0.25,0.35,0.13,0.07
          )),
  future = sample(x=c(1:5), size = survey_size_ubc, replace = TRUE,
          prob= c(0.45,0.27,0.15,0.08,0.05
          )),
  taxes = sample(x = c("yes","no"), size = survey_size_ubc,
                   replace = TRUE, prob = c(0.25, 0.75)),
  vote = sample(x = c("yes","no"), size = survey_size_ubc,
                   replace = TRUE, prob = c(0.46, 0.54)),
  campus = sample(x=c(1:5), size = survey_size_ubc, replace = TRUE,
          prob= c(0.10,0.10,0.1,0.25,0.45
          )),
  job = sample(x=c(1:5), size = survey_size_ubc, replace = TRUE,
          prob= c(0.22,0.25,0.33,0.13,0.07
          )),
  news = sample(x = c(1:5), size = survey_size_ubc,
                  replace = TRUE, prob = c(0.05, 0.28, 0.33, 0.29, 0.05)),
  interest = sample(x = c("yes","no"), size = survey_size_ubc,
                   replace = TRUE, prob = c(0.32, 0.68)),
  party =  sample(x = c("Liberal","Conservative", "NDP","Green","Bloc Québécois","No Response"), size =
                  replace = TRUE, prob = c(0.24, 0.34,0.22,0.1,0.02,0.08)),
)
```

```r
# Mcgill
set.seed(641)
mcgill <- tibble(
  age = rpois(n = survey_size_mcg, 21),
  citizen = sample(x = c("yes","no"), size = survey_size_mcg,
                   replace = TRUE, prob = c(0.3, 0.7)),
  tuition = sample(x=c(1:5), size = survey_size_mcg, replace = TRUE,
          prob= c(0.0,0.04,0.06,0.3,0.6
          )),
  hours = rnorm(n = survey_size_mcg, mean = 12, sd = 1),
  wage = sample(x = c("yes","no"), size = survey_size_mcg,
                   replace = TRUE, prob = c(0.15, 0.85)),
  income = sample(x=c(1:5), size = survey_size_mcg, replace = TRUE,
          prob= c(0.1,0.15,0.2,0.3,0.25
          )),
```

```
  hospital = sample(x = c("less","more","same"), size = survey_size_mcg,
                    replace = TRUE, prob = c(0.7, 0.1, 0.2)),
  purchase = sample(x=c(1:5), size = survey_size_mcg, replace = TRUE,
          prob= c(0.5,0.25,0.15,0.15,0.05
          )),
  future = sample(x=c(1:5), size = survey_size_mcg, replace = TRUE,
          prob= c(0.4,0.4,0.1,0.05,0.05
          )),
  taxes = sample(x = c("yes","no"), size = survey_size_mcg,
                    replace = TRUE, prob = c(0.1, 0.9)),
  vote = sample(x = c("yes","no"), size = survey_size_mcg,
                    replace = TRUE, prob = c(0.36, 0.64)),
  campus = sample(x=c(1:5), size = survey_size_mcg, replace = TRUE,
          prob= c(0.1,0.1,0.2,0.35,0.25
          )),
  job = sample(x=c(1:5), size = survey_size_mcg, replace = TRUE,
          prob= c(0.37,0.28,0.15,0.13,0.07
          )),
  news = sample(x = c(1:5), size = survey_size_mcg,
                  replace = TRUE, prob = c(0.15, 0.18, 0.34, 0.18, 0.15)),
  interest = sample(x = c("yes","no"), size = survey_size_mcg,
                    replace = TRUE, prob = c(0.27, 0.73)),
  party =  sample(x = c("Liberal","Conservative", "NDP","Green","Bloc Québécois","No Response"), size =
                  replace = TRUE, prob = c(0.34, 0.16,0.1,0.04,0.32,0.04)),
)
```

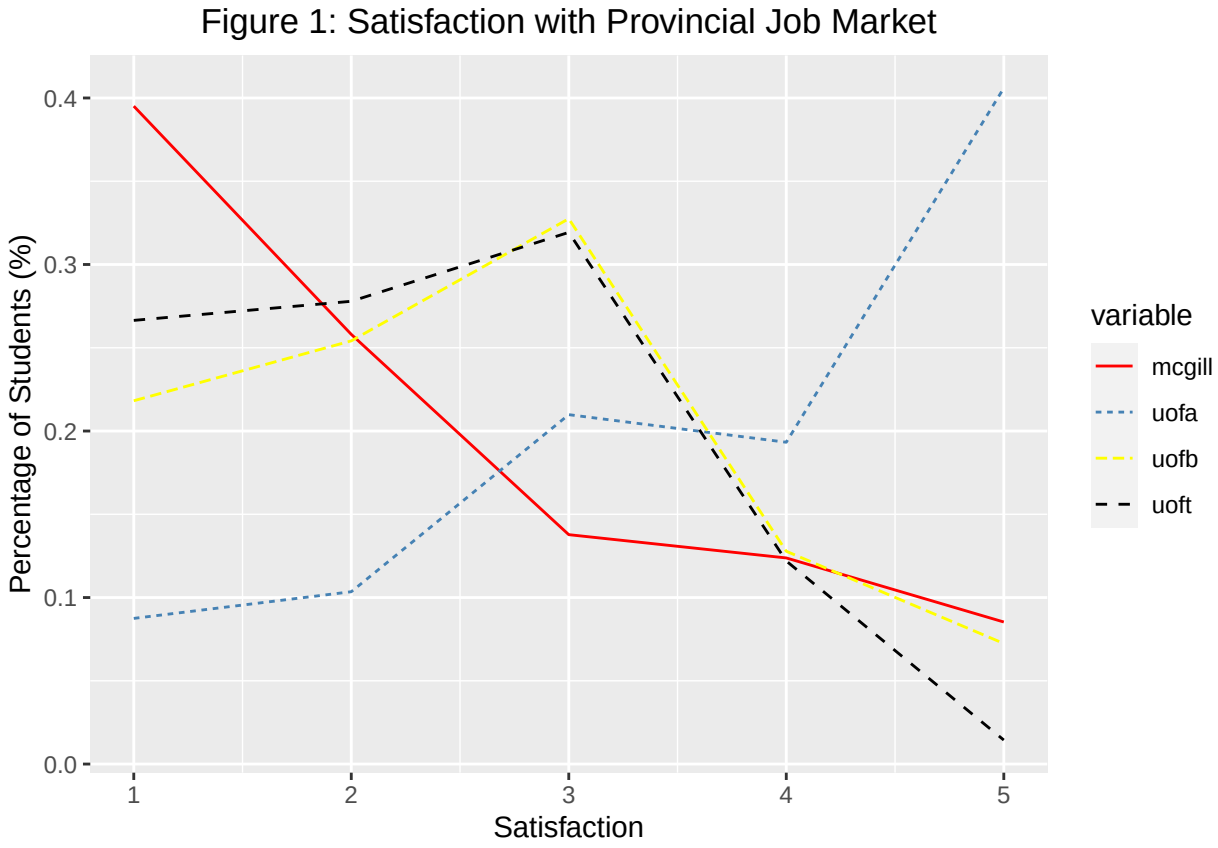
In Figure 1, the satisfaction of university students' provincial job market is depicted for the four selected universities based on a five-point scale. Notably, approximately 40% of McGill university students responded with 1 out of 5 satisfaction with the Quebec job market. While only 8% responded with 5 out of 5 satisfaction with the job market. In contrast, approximately 40% of University of Alberta students responded that their satisfaction with the Quebec job market is 5 out of 5. The highest proportion of University of Toronto and University of British Columbia students responded with 3 out of 5 satisfaction with their provincial job market (approximately 32% respondents for each university).

```
library(dplyr)

# Q13 job market

uoft_job=data.frame(cbind(table(uoft$job)), prop.table(table(uoft$job)))
colnames(uoft_job) <- c("uoft job count", "jobScale", "uoft")

uofa_job=data.frame(cbind(table(uofa$job)), prop.table(table(uofa$job)))
colnames(uofa_job) <- c("uofa job count", "jobScale", "uofa")

uofb_job=data.frame(cbind(table(uofb$job)), prop.table(table(uofb$job)))
colnames(uofb_job) <- c("uofb job count", "jobScale", "uofb")

mcg_job=data.frame(cbind(table(mcgill$job)), prop.table(table(mcgill$job)))
colnames(mcg_job) <- c("mcg job count", "jobScale", "mcgill")

temp1 <- merge(x=uoft_job, y=uofa_job, by ="jobScale",all=TRUE)
temp2 <- merge(x=uofb_job, y=mcg_job, by = "jobScale", all=TRUE)

sum_job <- merge(x=temp1, y=temp2, by = "jobScale", all=TRUE)
```

```r
sum_job[is.na(sum_job)] <- 0

sum_job$jobScale <- as.numeric(as.character(sum_job$jobScale))

df2 <- sum_job %>% select(jobScale, uoft, uofa, uofb, mcgill) %>%
  gather(key="variable", value="value", -jobScale)
head(df2)
```

```
##   jobScale variable      value
## 1        1     uoft 0.26644494
## 2        2     uoft 0.27790096
## 3        3     uoft 0.31929047
## 4        4     uoft 0.12195122
## 5        5     uoft 0.01441242
## 6        1     uofa 0.08747856
```

```r
#

ggplot(df2, aes(x=jobScale, y=value)) +
  geom_line(aes(color=variable, linetype=variable)) +
  scale_color_manual(values=c("red", "steelblue", "yellow", "black"))+
  labs(x= "Satisfaction", y = "Percentage of Students (%)") +
  ggtitle("Figure 1: Satisfaction with Provincial Job Market") +
  theme(plot.title=element_text(hjust=0.5))
```



In Figure 2, the confidence level of university students to purchase a home within 10 years is depicted for the four selected universities based on a five-point scale. Approximately 50% of McGill students responded with 1 out of 5 confidence of purchasing a home. Only 4% of McGill students responded with 5 out of 5 confidence

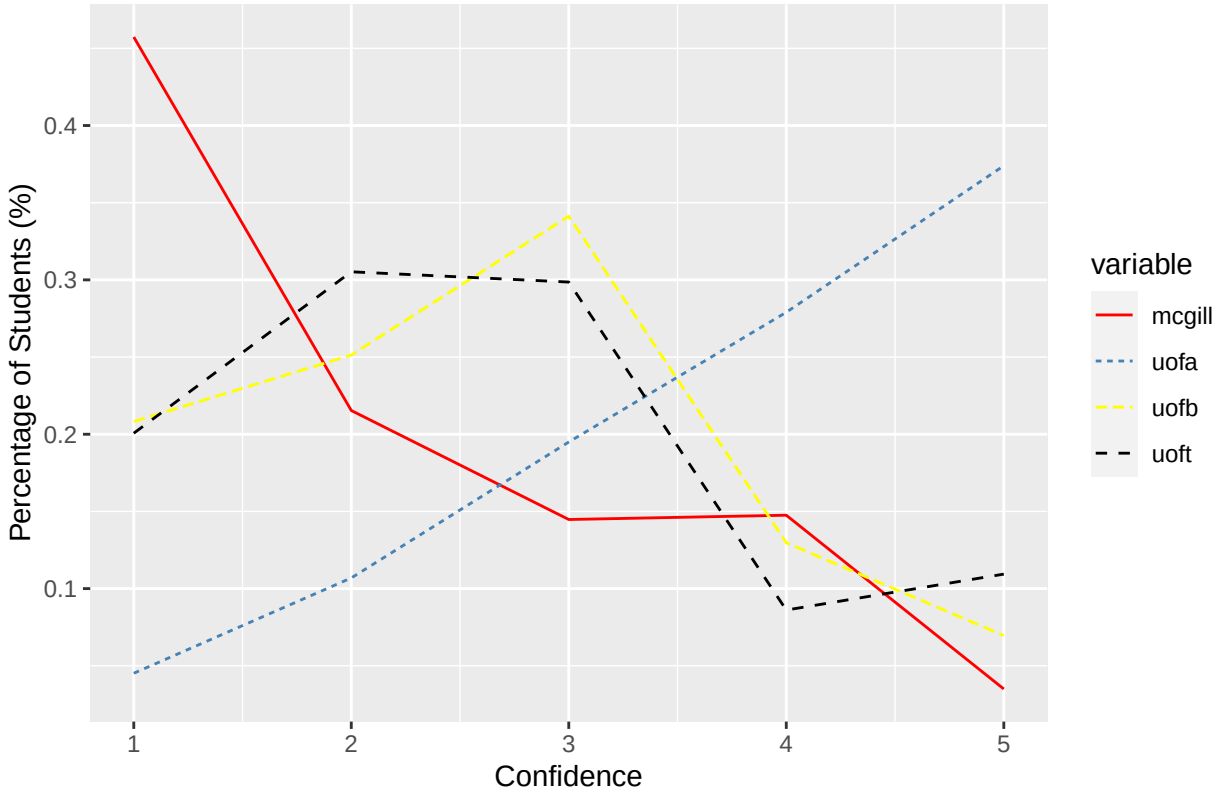
in purchasing a home in Quebec. Alternatively, approximately 38% of University of Alberta students have 5 out of 5 confidence that they will be able to purchase a home in Alberta within 10 years. Only 5% of University of Alberta students responded with 1 out of 5 confidence in purchasing a home in their province. Approximately 60% of University of Toronto students responded with either 2, or 3, out of 5 confidence in purchasing a home. Similarly, approximately 58% of University of British Columbia students responded with either 2, or 3, out of 5 confidence in purchasing a home.

```
library(dplyr)

# Q8, purchase

uoft_purchase=data.frame(cbind(table(uoft$purchase)), prop.table(table(uoft$purchase)))
colnames(uoft_purchase) <- c("uoft purchase count", "PurchaseScale", "uoft")

uofa_purchase=data.frame(cbind(table(uofa$purchase)), prop.table(table(uofa$purchase)))
colnames(uofa_purchase) <- c("uofa purchase count", "PurchaseScale", "uofa")

uofb_purchase=data.frame(cbind(table(uofb$purchase)), prop.table(table(uofb$purchase)))
colnames(uofb_purchase) <- c("uofb purchase count", "PurchaseScale", "uofb")

mcg_purchase=data.frame(cbind(table(mcgill$purchase)), prop.table(table(mcgill$purchase)))
colnames(mcg_purchase) <- c("mcg purchase count", "PurchaseScale", "mcgill")

temp1 <- merge(x=uoft_purchase, y=uofa_purchase, by ="PurchaseScale",all=TRUE)
temp2 <- merge(x=uofb_purchase, y=mcg_purchase, by = "PurchaseScale", all=TRUE)

sum_purchase <- merge(x=temp1, y=temp2, by = "PurchaseScale", all=TRUE)
sum_purchase[is.na(sum_purchase)] <- 0

sum_purchase$PurchaseScale <- as.numeric(as.character(sum_purchase$PurchaseScale))

df1 <- sum_purchase %>% select(PurchaseScale, uoft, uofa, uofb, mcgill) %>%
  gather(key="variable", value="value", -PurchaseScale)
head(df1)
```

```
##   PurchaseScale variable      value
## 1             1     uoft 0.20066519
## 2             2     uoft 0.30524760
## 3             3     uoft 0.29859571
## 4             4     uoft 0.08610495
## 5             5     uoft 0.10938655
## 6             1     uofa 0.04516867
```

```
#

ggplot(df1, aes(x=PurchaseScale, y=value)) +
  geom_line(aes(color=variable, linetype=variable)) +
  scale_color_manual(values=c("red", "steelblue", "yellow", "black"))+
  labs(x= "Confidence", y = "Percentage of Students (%)") +
  ggtitle("Figure 2: Confidence in Purchasing Home") +
  theme(plot.title=element_text(hjust=0.5))
```

Figure 2: Confidence in Purchasing Home



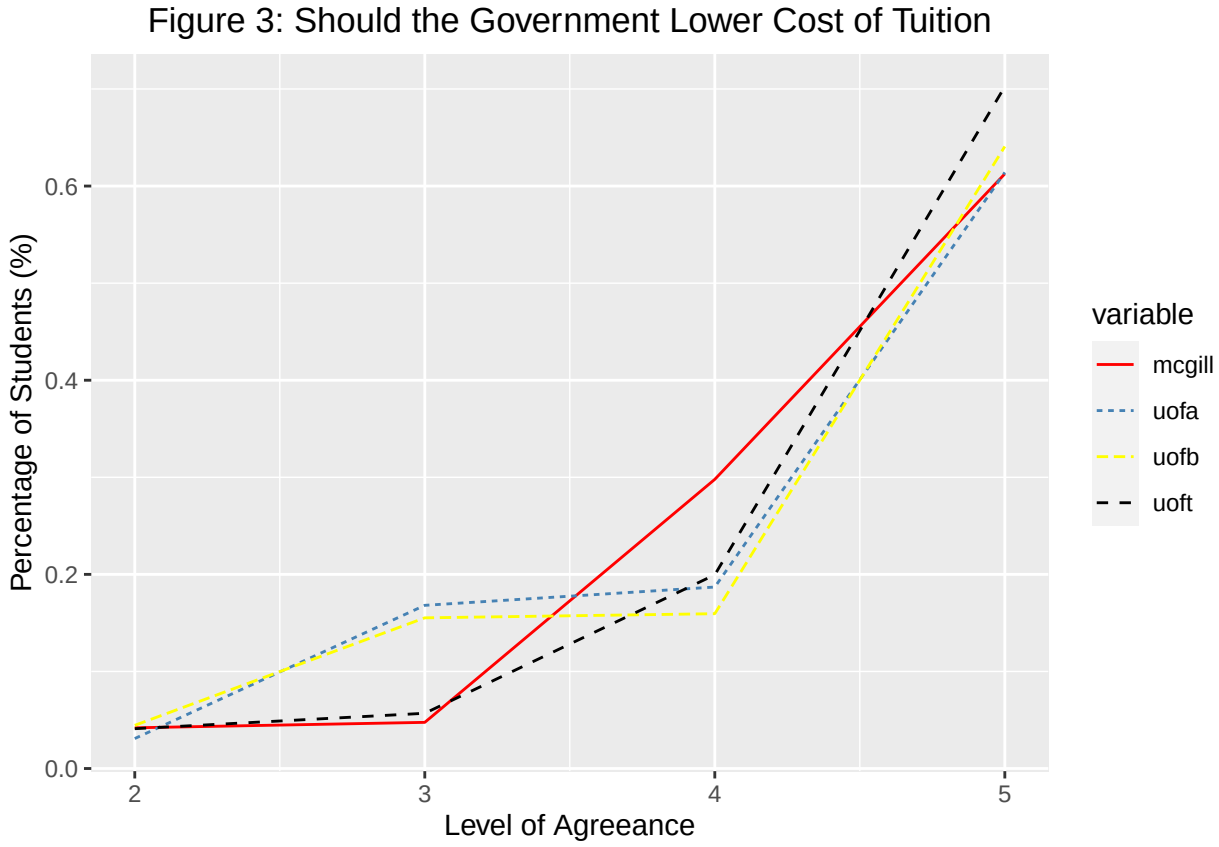
In Figure 3, university students were asked whether the government should lower tuition costs with answers varying from one to five in terms of agreeance. 60% of students in each of the four universities responded with 5 out of 5 agreeance that the government should lower the cost of tuition. No student in any of the universities responded with 1 out of 5 agreeance for this question.

```
library(dplyr)
################################################################################
# Tuititon

uoft_tuition=data.frame(cbind(table(uoft$tuition)), prop.table(table(uoft$tuition)))
colnames(uoft_tuition) <- c("uoft tuition count", "TuitionScale", "uoft")

uofa_tuition=data.frame(cbind(table(uofa$tuition)), prop.table(table(uofa$tuition)))
colnames(uofa_tuition) <- c("uofa tuition count", "TuitionScale", "uofa")

uofb_tuition=data.frame(cbind(table(uofb$tuition)), prop.table(table(uofb$tuition)))
colnames(uofb_tuition) <- c("uofb tuition count", "TuitionScale", "uofb")

mcg_tuition=data.frame(cbind(table(mcgill$tuition)), prop.table(table(mcgill$tuition)))
colnames(mcg_tuition) <- c("mcg tuition count", "TuitionScale", "mcgill")

temp1 <- merge(x=uoft_tuition, y=uofa_tuition, by ="TuitionScale",all=TRUE)
temp2 <- merge(x=uofb_tuition, y=mcg_tuition, by = "TuitionScale", all=TRUE)

sum_tuition <- merge(x=temp1, y=temp2, by = "TuitionScale", all=TRUE)
sum_tuition[is.na(sum_tuition)] <- 0
sum_tuition$TuitionScale <- as.numeric(as.character(sum_tuition$TuitionScale))
```

```
df <- sum_tuition %>% select(TuitionScale, uoft, uofa, uofb, mcgill) %>%
  gather(key="variable", value="value", -TuitionScale)
head(df)
```

```
##   TuitionScale variable      value
## 1            2     uoft 0.04101996
## 2            3     uoft 0.05691057
## 3            4     uoft 0.19955654
## 4            5     uoft 0.70251293
## 5            2     uofa 0.03087479
## 6            3     uofa 0.16809605
```

```
#

ggplot(df, aes(x=TuitionScale, y=value)) +
  geom_line(aes(color=variable, linetype=variable)) +
  scale_color_manual(values=c("red", "steelblue", "yellow", "black"))+
  labs(x= "Level of Agreeance", y = "Percentage of Students (%)") +
  ggtitle("Figure 3: Should the Government Lower Cost of Tuition") +
  theme(plot.title=element_text(hjust=0.5))
```



Discussion

Elections give citizens leverage to shape policy of political parties, while also providing the interface of connecting government policy to the public's desired expectations. As a result, Canadian political parties are continually interested in finding new avenues of reaching their potential voters. The purpose of this particular

survey was to gain a better understanding of university students' political views and how they contrast in different provinces of Canada.

According to a 2017 Statistics Canada report, Alberta has the highest median income in Canada at $70,300 per year (Canadian Income, 2019). Ontario ($62,700), British Columbia ($62,100) and Quebec ($52,400) also have varied median income per year. This variation could serve as an explanation as to why our resulting survey data in Figure 1 differs between the four universities. Considering the economic prosperity that we see in median income statistics for Albertans, it is reasonably understandable that we see high satisfaction among University of Alberta students in terms of the job market in their province (Figure 1). Likewise, this correlation between median income levels and satisfaction with the provincial job markets can explain why students of McGill university (Located in Quebec) report unsatisfaction with the Quebec job market (Figure 1). Quebec is among the lowest reported median income per year in Canada, only ahead of Nova Scotia (Canadian Income, 2019). Furthermore, University of Toronto and University of British Columbia students are both relatively neutral in terms of satisfaction with their provincial job markets (Figure 1).

As reported by Statistics Canada, the housing index varies greatly in certain province with indices of 97.2 (Alberta), 105.4 (Ontario), 105.5 (British Columbia) and 108.8 (Quebec) (New Housing, 2020). These indices could help explain the varying results in Figures 2. As seen in Figure 2d, McGill students are not confident that they will be able to purchase a property in 10 years. The relatively high price of housing in the province of Quebec (108.8 index) could explain the McGill students' lack of confidence. On the other end of the spectrum, University of Alberta students are confident in purchasing property within 10 years (Figure 2). This seems to be attributable to the very low housing index in Alberta (New Housing, 2020). Overall, based on the data collected by Statistics Canada and our simulated data, there seems to be a positively correlated relationship between housing index and confidence in buying property for university students in Canada.

If the Liberal party of Canada wishes to regain a majority government in the following election, they will need to address the some of the issues encountered by university students. An overwhelming majority of students surveyed from each university believe the government should lower tuition rates (Figure 3). Considering this consensus opinion among Canadian university students, the government would be wise to create a plan with universities to remove certain incidental fees and lower the cost of individual courses. Additionally, the Quebec is the 2nd largest province in Canada, in terms of population, accounting for 22% of the total Canadian population (Population Estimates, 2020). The Liberal party could look to improve their standing within the province by improving the key issues that students are facing within that province. McGill University students are not satisfied with the provincial job market (Figure 1) and lack confidence in purchasing property in Quebec (Figure 2). By creating government programs that recruit students after graduation and offering internship programs within the Quebec universities, students from the province will have the opportunity to gain valuable work experience that can aid their ability to purchase property in a few years.

Limitations

1. There are several students who were unwilling to answer certain questions when presented the opportunity to answer with "No Response". This could be mitigated by altering questions in such a way that the respondent feels comfortable answering while also providing the researchers the appropriate data needed to conduct their study.
2. There is a possibility that not all survey responses are honest or conducted with full attention to the survey. Students could be answering randomly to obtain a free coffee coupon in the end. This could be mitigated by adding survey questions that directly oppose other questions in the survey. This ensures that respondents are answering with awareness to the survey questions. For example, one survey question could read "Are you in favor of basic income" while another question follows-up later on in the survey with "Do you oppose basic income". These questions require different answers (ie. Yes vs No) and would allow researchers to discard invalid data.

Future Steps

1. Future researchers can look towards creating a study that increases the number of stratified subpopulations for the Canadian student population. For example, additional subpopulations for a Maritime and/or Territories university could provide an even better representation of the entire student population as their data was left out of this survey.
2. Future researchers could also follow-up this study with research that investigates why students orient themselves towards certain political parties. They can survey for factors such as geography, education level, program of study, gender, ethnicity.

Appendix A

As noted in Wu and Thomspon, Confidence Intervals are calculated using the following formula:

3.1.4 Confidence Intervals

Confidence intervals for the population mean μ_y based on the stratified sample mean can be constructed as

$(\bar{yst}\text{-}Z_{\alpha}/_2 \cdot \sqrt{v(\bar{y}st)}, \bar{yst}+Z_{\alpha}/_2 \cdot \sqrt{v(\bar{y}st)})$

Theorem 3.2 Under stratified simple random sampling, the Neyman allocation, which minimizes $V(\bar{yst})$ subject to $\sum_{h=1}^{H} n_h = n$, is given by

$n_h = \frac{nW_h\sigma_y h}{\sum_{k=1}^{H} W_k\sigma_y k} = \frac{nN_h\sigma_y h}{\sum_{k=1}^{H} N_k\sigma_y k}, h = 1, 2, ..., H$

The Confidence Interval Calculation for our survey is presented in below:

```
## [1]  7.000866 15.391952
```

Appendix B Survey & Link

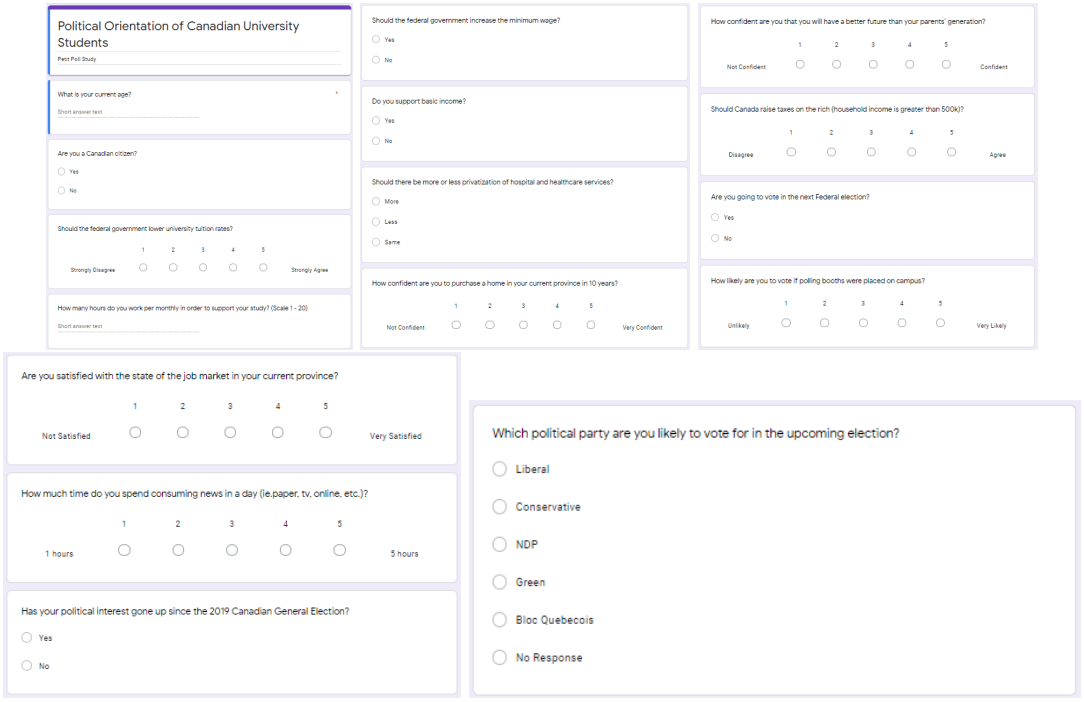


Figure 1: Figure 4. Survey Questions

A link to the study survey is located at the following: https://docs.google.com/forms/d/e/1FAIpQLSf2nTbfJ-9MljrOg6ykcMsZStgj-R3hnUNrvqapgSk5uG7nsg/viewform?usp=sf_link

References

1. Australian Bureau of Statistics. Frames and Population. www.abs.gov.au/websitedbs/d3310114.nsf/home/Basic+Survey++Frames+&+Population#:~:text=The%20frame%20refers%20to%20the,etc)%20in%20the%20survey%20population.&tex
2. "Canadian Income Survey, 2017." Statistics Canada, 26 Feb. 2019, www150.statcan.gc.ca/n1/daily-quotidien/190226/t003b-eng.htm
3. Hadley Wickham, Jim Hester and Winston Chang (2020). devtools: Tools to Make Developing R Packages Easier. https://devtools.r-lib.org/,https://github.com/r-lib/devtools.
4. JJ Allaire and Yihui Xie and Jonathan McPherson and Javier Luraschi and Kevin Ushey and Aron Atkins and Hadley Wickham and Joe Cheng and Winston Chang and Richard Iannone (2020). rmarkdown: Dynamic Documents for R. R package version 2.3. URL https://rmarkdown.rstudio.com.
5. Kirill Müller (2017). here: A Simpler Way to Find Your Files. R package version 0.1. https://CRAN.R-project.org/package=here
6. "Liberals Take Losses but Win Enough in Quebec and Ontario to Form Minority Government | CBC News." CBCnews, CBC/Radio Canada. 22 Oct. 2019. www.cbc.ca/news/politics/federal-election-results-2019-cbc-leaders-1.5329485.
7. "Population Estimates, Quarterly." Population Estimates, Quarterly, Government of Canada, Statistics Canada, 29 Sept. 2020, www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901.
8. "New housing price index, monthly". 2020. Statistics Canada. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810020501.
9. R Core Team (2020). R: A language and environment for statistical computing. R, Foundation for Statistical Computing, Vienna, Austria. URL, https://www.R-project.org/.
10. Wu, Changbao, And Mary E. Thompson . Sampling Theory and Practice. Springer Nature, 2020.