

Analyzing the differences of the 2019 Canadian Federal Election If Everyone Had Voted

Zijun Ye (1005414507)

21/12/2020

Abstract

In this report, two datasets (CES & GSS) were used to analyze the differences if everyone has voted in the 2019 Canadian Election. The method used is logistic regression model and post-stratification. Canadian citizens that have different demographical characteristics such as sex, household income, the importance of religious belief, education, mother tongue, province of living, etc. would have different intentions of voting. It was found that the estimated percentage of votes for the Conservative Party was lower than the actual result.

Keywords

Logistic Regression Model, Post-stratification, 2019 Canadian Election, Percentage votes for the Conservative Party

Introduction

The 2019 Canadian federal election was held on October 21, 2019. The Liberal Party led by the incumbent Prime Minister Justin Trudeau has won this election. For all the Canadian citizens who are eligible and register to vote, there is approximately 67 percent effectively participating in voting [1].

This report summarizes the statistical model and analysis results associated with the differences of that if everyone who is a registered voter takes part in this election. The purpose of this report is to investigate what the most important difference comparing to the actual result is. The main reasons that influence their intention to vote and would cause such differences such as citizens in different regions, their household condition, etc. are all important points needed to be taking into consideration.

In this report, the model is built using the general linear regression and post-stratification (MRP) method to estimate the percentage of votes for the Conservative Party base on the 2019 Canadian Election Study (CES) Data available on the ces-ec.ca website [2] and the 2017 General Social Survey: Families Cycle 31 (GSS) Data provided by Statistics Canada [3]. The model selected and dataset used is specified in detail in the next several sections

Methodology

Data

This report uses two main datasets, the first one is the 2019 Canadian Election Study (CES) Data which are used for building the general linear regression model to estimate whether a Canadian citizen is willing to vote for the Conservative party. This CES Data contains 620 variables and 37822 pre-election survey interviews. The second one is the 2017 General Social Survey: Families Cycle 31 (GSS) Data which the observations would be regrouping into cells that have the same demographical characteristics and experience the post-stratified process to estimate the final percentage of votes for the Conservative Party. The cleaning up process of the GSS Data was based on the R codes provided by STA304 Professor Rohan Alexander and it contains 461 variables and 20602 observations.

Initially, the variables selection are based on finding the common variables in both CES and GSS dataset. Those variables that could reflect the living condition and the basic demographic information of the citizens

are primarily been selected. After selecting and cleaning up the variables, which includes the uniform process of modifying the variables' names and values for both datasets. Removing those observations who are not Canadian citizens. Most of the variables' categories have been reduced and variables are all transformed into either binary or categorical which would make further interpretation of the model coefficients easier and make each partitioning featured demographic cell of people more clear. The following Table 1&2 shows all the information with the selected reformatted variables and there are 13 variables including one outcome response variable "vote_conservative", 6 binary variables, and 6 categorical variables. All the variables in table 1&2 would be used for the next step of modeling.

Table 1: Original Variables Information

Variables	Description	Type
vote_conservative	Whether a citizen would vote for the Conservative Party	Binary
sex	is measured as either male or female	Binary
religion_importance	How religion is important for a citizen	Categorical
employment	Whether a citizen is employed	Binary
have_children	Whether a citizen have at least a child	Binary
born_canada	Whether a citizen was born in Canada	Binary
household_size	How many people are living together in a household	Categorical
province	Which province is a citizen currently living in	Categorical
household_income	How much household income made in a year	Categorical
age_group	The age of a citizen in 2019	Categorical
marital	The martial status of a citizen	Binary
education	The educational level of a citizen	Binary
mothertongue	The mother tongue of a citizen (can be more then one language)	Categorical

Table 2: Original Variables Information (continuous)

Variables	Number_of_Categories	Values
vote_conservative	2	1 (vote for the Conservative), 0 (vote for the others party)
sex	2	Female, Male
religion_importance	4	Not at all important, Not very important, Somewhat important, Very important
employment	2	Employed, Unemployed
have_children	2	Yes, No
born_canada	2	Yes, No
household_size	4	One person household, Two person household, Three person household, Four or more person household
province	13	Alberta, British Columbia, Manitoba etc. (All the provinces in Canada)
household_income	6	Less than \$25,000, \$25,000 to \$49,999, \$50,000 to \$74,999, \$75,000 to \$99,999, \$100,000 to \$ 124,999, \$125,000 and more
age_group	4	18-30, 30-55, 55-70, Above 70
marital	2	Married, Not Married
education	2	Less than highschool diploma or equivalent, More than highschool diploma
mothertongue	4	English, French, Multiple, Others

Model

The method of the General Linear Model (GLM) would be used to analyze the CES dataset. The logistic regression is chosen since the outcome variable is whether a citizen would vote for the Conservative Party. The assumption of this model which the predictors must be independent of each other. In order to satisfy the assumption of the GLM model, removing the variables that would have collinearity with others. The mathematic notation of the original model can be express as,

$$\begin{aligned} \log\left(\frac{P}{1-P}\right) = & \beta_0 + \beta_1 \textit{sex} + \beta_2 \textit{religion_importance} \\ & + \beta_3 \textit{employment} + \beta_4 \textit{have_children} + \beta_5 \textit{born_canada} \\ & + \beta_6 \textit{household_size} + \beta_7 \textit{province} + \beta_8 \textit{household_income} \\ & + \beta_9 \textit{age_group} + \beta_{10} \textit{marital} + \beta_{11} \textit{education} + \beta_{12} \textit{mothertongue} \end{aligned}$$

Model 1. Original Full Model

where P indicate the probability of a Canadian citizen would vote for the Conservative Party. The coefficients β_0 is the intercept term and other β_i 's are the estimated coefficient of each predictor in the model.

The next step is using the variance inflation factor (VIF) to detect the magnitude of multicollinearity. If the magnitude of VIF is 1 means that there is no correlation between a predictor and the remaining predictor variables. The general rule of thumb is that VIFs exceeding 4 warrant further investigation.

The further variable selection from these 12 predictors involves the stepwise selection method by Bayesian Information Criterion (BIC). Since there are 2^{12} possible models, the basic idea for such a method is to iteratively remove the non-important predictors or add useful predictors from the full model according to some criterion and choose the best model. The model after being stepwise BIC selection will become the final model.

The model's estimated coefficients are required to be exponential which could be interpreted as the odds ratio of the interest, and the 95% confidence intervals are used to assess the significance of the estimates. The AUC-ROC curve would also be introduced to check the final model's goodness of fit. This curve tells how much a model is capable of distinguishing between classes. The higher the AUC, the better the model is at predicting the true value. All of the variables, model selection, and model diagnostic would be done by the software programming language R.

Post-stratification

After successfully building the model, the GSS dataset would be used as the census data and the observation would be partitioned into different demographic cells (stratum) according to the values of each model predictors. Then the estimation of the response variable would be done by using the logistic regression model and be calculated according to each cell's estimation and weights. The next thing is to aggregate the cell-level estimates up to a population-level estimate by weighting each cell by its relative proportion in the population. The formula for the post-stratification calculation can be expressed as:

$$\hat{y}^{PS} = \frac{\sum N_j \hat{y}_j}{\sum N_j}$$

where \hat{y}_j is the estimate in each cell and N_j is the population size of the j^{th} cell based off demographics. \hat{y}^{PS} indicates the post-stratification prediction over the total population.

Results

As mention above, the information of result coefficient summary of the original full model shown in the appendix (1) have encountered the “NA” issue of the predictor “household_size”, which describes the number of people living together in a household, the likely reason might be that there exists another predictor which have a correlation with the “household_size”. After several trials by removing different predictors, it was finally found that there is the correlation between “religion_important” and “household_size”, so only one is kept in the model between them for doing the next steps of analysis. The predictor “religion_important” remains since the religious belief is important for a person to vote. Therefore the second version of the full model can be shown as:

$$\begin{aligned} \log\left(\frac{P}{1-P}\right) = & \beta_0 + \beta_1 sex + \beta_2 religion_importance \\ & + \beta_3 employment + \beta_4 have_children + \beta_5 born_canada \\ & + \beta_6 province + \beta_7 household_income + \beta_8 age_group \\ & + \beta_9 marital + \beta_{10} education + \beta_{11} mothertongue \end{aligned}$$

Model 2. Original Full Model 2

The VIF in Table 4 in the appendix (2) shows that all the values are under 4 which indicates there is no multicollinearity between predictors in Model 2. Among all predictors of Model 2 in the coefficients summary in the appendix (3), only some of them are significant regression coefficients. Therefore, applying the stepwise selection method is necessary to determine which of them has a significant effect on prediction. After evaluation on the stepwise BIC method, it provides a simpler with fewer predictors and each has significant regression coefficients. This BIC model is chosen to be the final model in which the model expression and the odds ratio coefficient summary table are demonstrated below,

$$\begin{aligned} \log\left(\frac{P}{1-P}\right) = & \beta_0 + \beta_1 sex + \beta_2 religion_importance \\ & + \beta_3 have_children + \beta_4 province + \beta_5 household_income \\ & + \beta_6 marital + \beta_7 education + \beta_8 mothertongue \end{aligned}$$

Model 3. Final Model

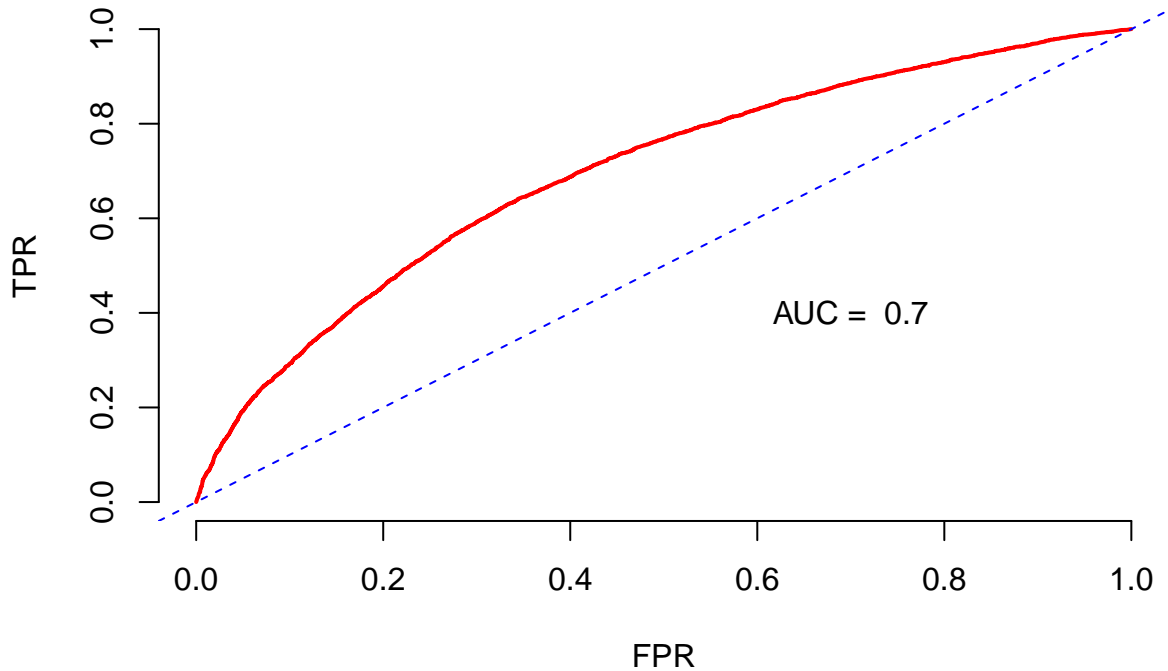
Table 3: Odds Ratio and 95% Confidence Interval of the Coefficients

	est	2.5	97.5
(Intercept)	0.8688	0.7229	1.0442
sexMale	1.6093	1.4999	1.7266
religion_importanceNot very important	1.2707	1.1365	1.4208
religion_importanceSomewhat important	1.4935	1.3442	1.6595
religion_importanceVery important	1.8977	1.6938	2.1260
have_childrenYes	1.2467	1.1485	1.3532
provinceBritish Columbia	0.3372	0.2935	0.3875
provinceManitoba	0.4728	0.3970	0.5632
provinceNew Brunswick	0.3386	0.2668	0.4296
provinceNewfoundland and Labrador	0.2155	0.1633	0.2844
provinceNorthwest Territories	0.1697	0.0379	0.7591
provinceNova Scotia	0.1856	0.1450	0.2376
provinceNunavut	0.3731	0.1233	1.1291
provinceOntario	0.3129	0.2815	0.3479

	est	2.5	97.5
provincePrince Edward Island	0.0857	0.0388	0.1896
provinceQuebec	0.2049	0.1748	0.2402
provinceSaskatchewan	0.7372	0.6129	0.8866
provinceYukon	0.1981	0.0553	0.7099
household_income\$125,000 and more	1.1063	0.9756	1.2545
household_income\$25,000 to \$49,999	0.8144	0.7127	0.9306
household_income\$50,000 to \$74,999	0.9556	0.8430	1.0833
household_income\$75,000 to \$99,999	0.9351	0.8228	1.0627
household_incomeLess than \$25,000	0.6181	0.5269	0.7251
maritalNot Married	0.8134	0.7505	0.8816
educationMore than highschool diploma	0.7687	0.7116	0.8304
mothertongueFrench	0.7132	0.6100	0.8339
mothertongueMultiple	0.7700	0.7047	0.8415
mothertongueOthers	0.9984	0.8526	1.1691

As shown in Model 3 above, the remained 8 predictors are “sex, religion_importance, have_children, province, household_income, marital, education, and mothertongue”. This model is telling us these predictors have the most influence on the intention of people to vote for the 2019 election. Table 3 indicates the natural scale of the odds ratio estimated coefficients of each predictor and their lower and upper bound of the 95% confidence intervals. If the confidence interval does not contain 1, it indicates this predictor is significant. The following figure of the ROC curve shows that the area under the curve (AUC) is 70% which means that this logistic regression can distinguish whether a Canadian citizen would vote for the Conservative Party 70 percent of the time.

Figure 1. ROC curve



Before starting the post-stratification, the census data would suppose to be partitioning into 19969 cells. However, there are only 10 provinces in the census data and more than half cells do not contain any observations. Therefore, only 5516 cells have formed at the end. It's worth noting that there is one cell involving, which contains the most numbers of observations, 61 observations, and their characteristics are:

female, having a very important religious belief, having no child, living in Ontario, having yearly household income at the level of \$25,000 to \$49,999, married, having high school or less educational level and English native speaker.

Then, the final model that estimating the Conservative Party support would be used in each cell to calculate the estimated probability to vote for the Conservative Party. According to the post-stratification formula, the result of estimation of that the proportion of voters in favor of voting for the Conservative Party to be 0.2762. This is based on our post-stratification analysis of the proportion of voters in favor of the Conservative Party modeled by a logistic regression model, which accounted for the 8 variables “sex, religion_importance, have_children, province, household_income, marital, education, and mothertongue”.

Discussion

The following analysis is about the final model mention in the last section with 4 categorical variables and 4 binary variables. The interpretation of these estimated coefficients in the natural scale of predictors would be, in general, the odds ratio between this citizen in this specific predictor group and a citizen that in the reference group, controlling other predictors have the reference group value. For example, “sexMale” which refers to a male citizen have an estimated value of 1.6093, which means that the odds ratio of voting for the Conservative Party between a male citizen and a female citizen is 1.6093, which can be interpreted as a male would be more likely to vote for the Conservative Party than a female, controlling other characteristics be the same. Also, the 95% confidence interval of “sexMale” does not include the value of 1, which indicates this predictor is significant for the model. For the predictor “religion_importance”, we can see that as the religious belief of a citizen becoming more important, the probability of voting for the Conservative Party becomes higher. One more interesting finding is that the majority of the provinces such as Alberta, Manitoba, British Columbia, Saskatchewan, etc. have voted for the Conservative Party according to the data provided by the Statista [4]. From our model coefficient table, we can notice that the predictor “province” has Alberta as the reference group, and no other provinces would have a higher odds ratio comparing to it. Whereas, Manitoba, British Columbia, Saskatchewan, etc. have an odds ratio nearly closed to 1 comparing to Alberta, which means people who live in these provinces would also have more intention to vote for the Conservative Party as expected. There are more interesting findings that can be discovered from this table.

The sample data is drawn from the 2019 Canadian Election Study (CES) Data provided by the ces-ec.ca website. The census data is drawn from the 2017 General Social Survey: Families Cycle 31 (GSS) Data provided by Statistics Canada. There are originally 12 predictor variables selected in both datasets, and the stepwise BIC method is used to select the best simple logistic models for predicting the probability of voting for the Conservative Party in this election. Eventually, the actual percentage of votes for the Conservative Party in the 2019 Canadian Election was 33.34% according to the data provided by Elections Canada [1]. The census data are partitioned into corresponding cells according to the 8 predictors models and post-stratification analyses are performed and have the result that if everyone votes in this election, the estimated percentage that votes for the Conservative Party is 27.62% which is lower than the actual one. However, although this research has utilized recent survey data and census data, there might exist some potential weaknesses problem during the entire analysis. The following subsection will discuss the weaknesses and raise some notices for further research.

Weakness

First of all, the 2019 Canadian Election Study (CES) Data contains two main parts of the survey, the Campaign Period Survey (CPS) and the Post-Election Survey (PES). All the variables selected for the model are from the CPS part since there are too many missing values in the PES part. The most important variable, vote choice from the CPS part, is lack of reliability because what people actually vote after the CPS might be different from what they have answered. Also, the number of common variables in both CES Data and GSS Data are quite a few, the answer bank for both data may not share the exact information as well. Therefore, there exists much bias when doing the cleanup process. Besides, the GSS was done in

2017, Although the variable ‘age’ has been increased by 2 to accommodate the CES, it is not practical to modify other variables that would potentially influence the accuracy of the predictions. The GSS Data as the census contains a small number of samples that could not represent the population at all. There is no sample coming from Northwest Territories, Nunavut, and Yukon provinces. The problem of less amount of samples would also cause absent cells in the post-stratification step, the census data only establish 5516 out of 19969 cells.

The model has some weaknesses either. The model used in this report is the logistic regression by considering all variables as a fixed effect. However, some variables would be varied as the time elapsing, those variables usually would be treated as the random effect which has their own distribution. Thus, if holding every predictor constant, the change in probability of the outcome variable over different values of the predictor of interest is only true when all other predictors are held constant or in the same group which is quite narrowing. In other words, the ordinary logistic regression could not cover more variance than the mixed-effect regression model. Last but not least, we only estimate the percentage votes for the Conservative Party which as a result is lower than the actual percentage votes. However, we don’t have enough evidence to show that such a difference in the percentage change would change the result of which party would win the election if everyone has voted.

Next Steps

After discussing the weaknesses, any subsequent work should be considered, which may help further research to be more successful. As mentioned before, the PES part of the CES should be completed since people would have different thoughts when the election is holding and after the election was held. The PES data would be more reliable for improving the accuracy of the model prediction. The GSS should be held every year and such a survey should cover as many people as it can to make it more representative of the entire population. Also, the method of these surveys could be improved such as avoiding imperfect sampling population and non-response problems. The survey organization should increase the volume of the survey in some efficient way such as introducing various ways for collection and launch some incentive policies to gain more attention for the survey. Moreover, both dataset contains more useful and significant variables, but this report only analyzes some of them due to technical and time constraints. Therefore, more variables should be used to build a more effective and more practical model with mixed effects that can control more of the unpredictable variances. Moreover, other major parties should also be modeled and check the difference that if everyone has voted to compare to the actual result.

Reference

- [1] Elections Canada. (2019). FORTY-THIRD GENERAL ELECTION 2019 Official Voting Results. <https://www.elections.ca/res/rep/off/ovr2019app/home.html#3>.
- [2] Stephenson, Laura B; Harell, Allison; Rubenson, Daniel; Loewen, Peter John, 2020, “2019 Canadian Election Study - Online Survey”, <https://doi.org/10.7910/DVN/DUS88V>, Harvard Dataverse, V1
- [3] General social survey on Family (cycle 31), 2017 Retrieved Dec 01,2020 from: <http://dc.chass.utoronto.ca>
- [4] Statista. (2019). “Preliminary share of total votes for each party in the Canadian federal election held on October 21, 2019 (as of October 22, 2019), by province*”. Retrieved Dec 19, 2020 from <https://www.statista.com/statistics/1062264/canada-election-share-of-popular-vote-per-party-by-province/>.

Appendix

(1) R Summary of the Original Model

```
##
## Call:
## glm(formula = vote_conservative ~ age_group + sex + religion_importance +
##       employment + have_children + born_canada + province + household_income +
##       marital + education + mothertongue + household_size, family = binomial,
##       data = election)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7198  -0.8114  -0.6162   1.0450   2.4821
##
## Coefficients: (3 not defined because of singularities)
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.185280   0.120003  -1.544  0.12260
## age_group30-55    0.090761   0.064767   1.401  0.16111
## age_group55-70    0.117734   0.069047   1.705  0.08817 .
## age_groupAbove 70  0.235596   0.084459   2.789  0.00528 **
## sexMale          0.459884   0.036537  12.587 < 2e-16 ***
## religion_importanceNot very important 0.236184   0.057056   4.140 3.48e-05 ***
## religion_importanceSomewhat important 0.395465   0.053819   7.348 2.01e-13 ***
## religion_importanceVery important    0.637133   0.058129  10.961 < 2e-16 ***
## employmentUnemployed -0.072193   0.043268  -1.669  0.09521 .
## have_childrenYes    0.197708   0.043415   4.554 5.27e-06 ***
## born_canadaYes     -0.006551   0.055159  -0.119  0.90546
## provinceBritish Columbia -1.097036   0.071021 -15.447 < 2e-16 ***
## provinceManitoba     -0.746871   0.089243  -8.369 < 2e-16 ***
## provinceNew Brunswick -1.082922   0.121641  -8.903 < 2e-16 ***
## provinceNewfoundland and Labrador -1.528890   0.141568 -10.800 < 2e-16 ***
## provinceNorthwest Territories -1.798283   0.764537  -2.352  0.01867 *
## provinceNova Scotia  -1.687125   0.125959 -13.394 < 2e-16 ***
## provinceNunavut     -0.974352   0.566065  -1.721  0.08520 .
## provinceOntario     -1.164058   0.054103 -21.516 < 2e-16 ***
## provincePrince Edward Island -2.463889   0.405102  -6.082 1.19e-09 ***
## provinceQuebec      -1.591685   0.081136 -19.617 < 2e-16 ***
## provinceSaskatchewan -0.309753   0.094305  -3.285  0.00102 **
## provinceYukon       -1.635805   0.651812  -2.510  0.01209 *
## household_income$125,000 and more    0.098574   0.064220   1.535  0.12480
## household_income$25,000 to $49,999 -0.202834   0.068840  -2.946  0.00321 **
## household_income$50,000 to $74,999 -0.045996   0.064291  -0.715  0.47434
## household_income$75,000 to $99,999 -0.066834   0.065331  -1.023  0.30631
## household_incomeLess than $25,000 -0.463116   0.082867  -5.589 2.29e-08 ***
## maritalNot Married -0.201811   0.041384  -4.877 1.08e-06 ***
## educationMore than highschool diploma -0.268093   0.039608  -6.769 1.30e-11 ***
## mothertongueFrench -0.335368   0.079774  -4.204 2.62e-05 ***
## mothertongueMultiple -0.250705   0.046455  -5.397 6.79e-08 ***
## mothertongueOthers  -0.008530   0.087075  -0.098  0.92196
## household_sizeOne person household    NA         NA         NA         NA
## household_sizeThree person household    NA         NA         NA         NA
## household_sizeTwo person household    NA         NA         NA         NA
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 21008  on 17623  degrees of freedom
## Residual deviance: 19222  on 17591  degrees of freedom
## AIC: 19288
##
## Number of Fisher Scoring iterations: 4
```

(2) VIF Table

Table 4: Variance Inflation Factor

	VIF
age_group	1.6499
sex	1.0722
religion_importance	1.0894
employment	1.4564
have_children	1.2779
born_canada	1.2405
province	2.1681
household_income	1.3453
marital	1.3259
education	1.1088
mothertongue	2.5960

(3) Model 2 Coefficients Table

Table 5: Model 2 Coefficients Table

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.1852803	0.1200026	-1.5439685	0.1225960
age_group30-55	0.0907614	0.0647671	1.4013500	0.1611094
age_group55-70	0.1177341	0.0690473	1.7051222	0.0881716
age_groupAbove 70	0.2355963	0.0844595	2.7894590	0.0052796
sexMale	0.4598843	0.0365367	12.5869185	0.0000000
religion_importanceNot very important	0.2361843	0.0570557	4.1395431	0.0000348
religion_importanceSomewhat important	0.3954649	0.0538188	7.3480817	0.0000000
religion_importanceVery important	0.6371327	0.0581291	10.9606417	0.0000000
employmentUnemployed	-0.0721933	0.0432676	-1.6685280	0.0952110
have_childrenYes	0.1977081	0.0434153	4.5538805	0.0000053
born_canadaYes	-0.0065509	0.0551594	-0.1187636	0.9054626
provinceBritish Columbia	-1.0970361	0.0710211	-15.4466124	0.0000000
provinceManitoba	-0.7468708	0.0892434	-8.3689234	0.0000000
provinceNew Brunswick	-1.0829224	0.1216405	-8.9026467	0.0000000
provinceNewfoundland and Labrador	-1.5288901	0.1415684	-10.7996596	0.0000000
provinceNorthwest Territories	-1.7982825	0.7645366	-2.3521208	0.0186667
provinceNova Scotia	-1.6871253	0.1259595	-13.3941926	0.0000000

	Estimate	Std. Error	z value	Pr(> z)
provinceNunavut	-0.9743518	0.5660653	-1.7212711	0.0852016
provinceOntario	-1.1640585	0.0541032	-21.5155235	0.0000000
provincePrince Edward Island	-2.4638885	0.4051020	-6.0821432	0.0000000
provinceQuebec	-1.5916849	0.0811363	-19.6174141	0.0000000
provinceSaskatchewan	-0.3097535	0.0943049	-3.2845961	0.0010213
provinceYukon	-1.6358053	0.6518116	-2.5096290	0.0120858
household_income\$125,000 and more	0.0985736	0.0642204	1.5349252	0.1248022
household_income\$25,000 to \$49,999	-0.2028344	0.0688404	-2.9464425	0.0032145
household_income\$50,000 to \$74,999	-0.0459962	0.0642913	-0.7154339	0.4743410
household_income\$75,000 to \$99,999	-0.0668340	0.0653313	-1.0230009	0.3063074
household_incomeLess than \$25,000	-0.4631160	0.0828668	-5.5886792	0.0000000
maritalNot Married	-0.2018112	0.0413842	-4.8765248	0.0000011
educationMore than highschool diploma	-0.2680926	0.0396083	-6.7686006	0.0000000
mothertongueFrench	-0.3353678	0.0797740	-4.2039733	0.0000262
mothertongueMultiple	-0.2507045	0.0464554	-5.3966745	0.0000001
mothertongueOthers	-0.0085304	0.0870748	-0.0979658	0.9219594

(4) Estimated Percentage of Votes for the Conservative Party

$$\frac{\text{alp_predict}}{0.2762}$$