

STA304 Problem Set 2

Names of your Group Members

October 19, 2020

Title of your Report

Name(s) of Author(s)

Date

1. Abstract

Here is where you give a brief (one paragraph overview of your entire paper). This should include some background/introduction, some methodology, results and conclusions.

2. Introduction

The General Social Survey - Family 2017 (GSS-Cycle 31) is a sample survey began on February 2nd and completed on November 30th in 2017. The program aims to capture social trends and provide data on current social issues. The survey mainly explores some characteristics of the modern Canadian family and how diverse they are. It collects a large amount of information for the selected respondents and some data about the family members of the respondents from the ten provinces in Canada.

Research Inspiration

The novel “The Meritocracy Trap” by Daniel Markovits crystallizes the problem, phenomenon, and representations of the American meritocracy. In the 1960s, the American economy was operated by lands, capital, machines, and manpower. However, in today’s society, the one and only one valuable resource is the brain of the elites, and families become the new “factory” who are responsible to output the most valuable resource in this era. One of the consequences of this phenomena is that marriage across different education levels becomes less and less under meritocracy. Parents in a household who have receive a college education are more likely to provide kids a stabilized family relationship and better education in the long run. In 2010, 25% of the newlyweds both had at least a college diploma.

Research objective

Given a wide set of features of the modern Canadian family, we want to investigate the characteristics that come along with households where parents have the same education and households where parents have a different education. We will be developing an accurate predictive logistic regression for parents with the same education to explore characteristics of such households. We can, therefore, identify the household formation under meritocracy in Canadian society.

3. Data

The data is collected through a survey, and its target population, sampling frame and the sampling techniques are explained below. Its target population contains all persons over the age of 15 in Canada, to the exclusion of the residents of the Yukon, Northwest Territories and Nunavut, and full-time institutional residents. The survey frame was created based on the telephone numbers accessible to Statistics Canada obtained from multiple sources, and The Address Register (AR), which is designed for connecting all telephone numbers with the address associated with them. The combination of the two methods is to guarantee that all households are covered with telephone numbers. The households which do not possess telephones are excluded from the survey population. The ten provinces were separated into strata for stratified random sampling. There are 27 strata in total. Every sampling unit was allotted to a stratum. A simple random sampling without replacement of the sampling units was performed to randomly pick households for interview. A respondent was randomly selected from the selected households. There were 20,602 respondents for the 2017 survey.

Features and Selected Variables

The table below lists all selected predictors that are considered affecting the result of this observation. It could be found that the original dataset involves a bunch of categorical variables. Thus, in general, nearly all of those variables are categorical variables indicating specific groups of respondents, except the numerical variable illustrating the age of respondent's first kid. In advance, since 13 variables, out of 81 of those, are chosen to indicate the relationship between our response variable and each of those predictors, the processing of creating a reduced model would be easier and the result might be more accurate than working on a full model. However, it can be a weakness that too many categorical variables restrict the interpretation of data generated from the model, and might not be able to distinguish degrees of response.

Table 1: List of Selected Variables and Descriptions

VARIABLE.NAME	TYPE	CHARACTERISTIC	DETAILED.CATEGORIES
has_same_education	Categorical	whether the partners have the same education	True, False
age_first_child	Numerical	Age of the first kid	Numeric values among 0 and 60
total_children	Categorical	Number of children	Integers among 0 and 7
feelings_life	Categorical	Rate of feelings about life	Integers among 0 and 10
place_birth	Categorical	Originally born in Canada	Born in Canada, Born outside canada, NA, Don't Know
_canada			
province	Categorical	Regional location of province	Quebec, Manitoa, Ontario, Alberta, Nova Scotia, British Columbia, Saskatchewan, Prince Edward Island, New Brunswick, Newfoundland and Labrador
education	Categorical	Education background	High school diploma or a high school equivalency certificate, Trade certificate or diploma, Bachelor's degree (e.g. B.A., B.Sc., LL.B.), College, CEGEP or other non-university certificate or diploma, Less than high school diploma or its equivalent, University certificate or diploma below the bachelor's level, University certificate, diploma or degree above the bachelor, NA

VARIABLE.NAME	TYPE	CHARACTERISTIC	DETAILED.CATEGORIES
hh_type	Categorical	Household type	Low-rise apartment (less than 5 stories), Single detached house, Other, High-rise apartment (5 or more stories), Don't know
partner_birth _province	Categorical	Partner's regional birth (province)	Quebec, Manitoa, Ontario, Alberta, Nova Scotia, British Columbia, Saskatchewan, Prince Edward Island, New Brunswick, Newfoundland and Labrador, Yukon / Northwest Territories / Nunavut
average_hours _worked	Categorical	Average weekly work time in hours	30.0 to 40.0 hours, 50.1 hours and more, NA, 0.1 to 29.9 hours, Don't know, 40.1 to 50.0 hours, 0 hour
self Rated _health	Categorical	Health level rated by self	Excellent, Good, Very good, Poor, Fair, Don't know, NA
self Rated _mental_health	Categorical	Mental health level rated by self	Excellent, Good, Very good, Poor, Fair, Don't know, NA
income_family	Categorical	Family income level	\$25,000 to \$49,999, \$75,000 to \$99,999, \$100,000 to \$ 124,999, \$50,000 to \$74,999, Less than \$25,000, \$125,000 and more
income_respondent	Categorical	Respondent income level	\$25,000 to \$49,999, Less than \$25,000, \$50,000 to \$74,999, \$125,000 and more, \$75,000 to \$99,999, \$100,000 to \$ 124,999

Data Cleaning

Our final dataset contains 5100 observations and each of them contains characteristics of modern Canadian household. Data is splitted into a training set and a testing set where each of them contains 50% (randomly sampled) of the original data.

4. Model

Logistic regression was chosen to fit the relationship between various Canadian household features and our target: whether or not the partners have the same education level. The model was chosen because the target variable is binary, and the logit link is appropriate for a target variable that follows a binomial distribution.

Base Model

We begin by building a base model that includes all the predictors. There are a lot of insignificant predictors and hence, we intended to perform variable selection to explain the variability with a simpler model.

Variable Selection

```
##
## Call: glm(formula = hasSameEducation ~ age_first_child + as.factor(total_children) +
##       as.factor(feelings_life) + as.factor(place_birth_canada) +
##       as.factor(province) + as.factor(education) + as.factor(hh_type) +
```

```

##   as.factor(partner_birth_province) + as.factor(average_hours_worked) +
##   as.factor(selfRatedHealth) + as.factor(selfRatedMentalHealth) +
##   as.factor(income_family) + as.factor(income_respondent),
##   family = binomial, data = train_df)
##
## Coefficients:
##                                     (Intercept)
##                                     -17.770999
##                                     age_first_child
##                                     -0.009654
##   as.factor(total_children)2
##                                     -0.175792
##   as.factor(total_children)3
##                                     0.027056
##   as.factor(total_children)4
##                                     -0.185489
##   as.factor(total_children)5
##                                     0.444429
##   as.factor(total_children)6
##                                     -0.157155
##   as.factor(total_children)7
##                                     0.152825
##   as.factor(feelings_life)3
##                                     -20.739932
##   as.factor(feelings_life)4
##                                     15.735174
##   as.factor(feelings_life)5
##                                     -0.897031
##   as.factor(feelings_life)6
##                                     -0.403344
##   as.factor(feelings_life)7
##                                     -0.170160
##   as.factor(feelings_life)8
##                                     -0.597849
##   as.factor(feelings_life)9
##                                     -0.790773
##   as.factor(feelings_life)10
##                                     -0.670254
##   as.factor(place_birth_canada)Born outside Canada
##                                     0.627877
##   as.factor(place_birth_canada)Don't know
##                                     1.773604
##   as.factor(province)British Columbia
##                                     0.252202
##   as.factor(province)Manitoba
##                                     0.068344
##   as.factor(province)New Brunswick
##                                     0.903245
##   as.factor(province)Newfoundland and Labrador
##                                     0.866185
##   as.factor(province)Nova Scotia
##                                     0.837005
##   as.factor(province)Ontario
##                                     0.963726

```

```

##                                as.factor(province)Prince Edward Island
##                                -0.508870
##                                as.factor(province)Quebec
##                                -0.230198
##                                as.factor(province)Saskatchewan
##                                0.848437
## as.factor(education)College, CEGEP or other non-university certificate or di...
##                                -31.302282
## as.factor(education)High school diploma or a high school equivalency certificate
##                                -30.227783
## as.factor(education)Less than high school diploma or its equivalent
##                                0.242033
## as.factor(education)Trade certificate or diploma
##                                -1.289828
## as.factor(education)University certificate or diploma below the bachelor's level
##                                -0.758986
## as.factor(education)University certificate, diploma or degree above the bach...
##                                -21.535879
## as.factor(hh_type)High-rise apartment (5 or more stories)
##                                -2.514771
## as.factor(hh_type)Low-rise apartment (less than 5 stories)
##                                -1.538622
## as.factor(hh_type)Other
##                                -2.055632
## as.factor(hh_type)Single detached house
##                                -1.710735
## as.factor(partner_birth_province)British Columbia
##                                0.114618
## as.factor(partner_birth_province)Manitoba
##                                -0.022238
## as.factor(partner_birth_province)New Brunswick
##                                -0.287278
## as.factor(partner_birth_province)Newfoundland and Labrador
##                                -0.082147
## as.factor(partner_birth_province)Nova Scotia
##                                -0.418124
## as.factor(partner_birth_province)Ontario
##                                -0.554558
## as.factor(partner_birth_province)Outside Canada
##                                -0.219124
## as.factor(partner_birth_province)Prince Edward Island
##                                0.729108
## as.factor(partner_birth_province)Quebec
##                                0.693048
## as.factor(partner_birth_province)Saskatchewan
##                                -0.604531
## as.factor(partner_birth_province)Yukon / Northwest Territories / Nunavut
##                                0.291787
## as.factor(average_hours_worked)0.1 to 29.9 hours
##                                20.403319
## as.factor(average_hours_worked)30.0 to 40.0 hours
##                                20.470465
## as.factor(average_hours_worked)40.1 to 50.0 hours
##                                20.315810

```

```

##               as.factor(average_hours_worked)50.1 hours and more
##                                     20.171525
##               as.factor(average_hours_worked)Don't know
##                                     20.522417
##               as.factor(selfRatedHealth)Excellent
##                                     -1.182859
##               as.factor(selfRatedHealth)Fair
##                                     -1.168308
##               as.factor(selfRatedHealth)Good
##                                     -1.188375
##               as.factor(selfRatedHealth)Poor
##                                     -0.691678
##               as.factor(selfRatedHealth)Very good
##                                     -1.320975
##               as.factor(selfRatedMentalHealth)Fair
##                                     -0.049350
##               as.factor(selfRatedMentalHealth)Good
##                                     -0.628972
##               as.factor(selfRatedMentalHealth)Poor
##                                     0.537989
##               as.factor(selfRatedMentalHealth)Very good
##                                     -0.151457
##               as.factor(income_family)$125,000 and more
##                                     0.418493
##               as.factor(income_family)$25,000 to $49,999
##                                     -0.147494
##               as.factor(income_family)$50,000 to $74,999
##                                     0.012795
##               as.factor(income_family)$75,000 to $99,999
##                                     0.492312
##               as.factor(income_family)Less than $25,000
##                                     -0.057759
##               as.factor(income_respondent)$125,000 and more
##                                     0.444826
##               as.factor(income_respondent)$25,000 to $49,999
##                                     0.211286
##               as.factor(income_respondent)$50,000 to $74,999
##                                     0.191087
##               as.factor(income_respondent)$75,000 to $99,999
##                                     0.122290
##               as.factor(income_respondent)Less than $25,000
##                                     0.268872
##
## Degrees of Freedom: 2549 Total (i.e. Null); 2478 Residual
## Null Deviance: 2130
## Residual Deviance: 1252 AIC: 1396
##
## Call: glm(formula = hasSameEducation ~ as.factor(feelings_life) + as.factor(place_birth_canada) +
##       as.factor(education) + as.factor(selfRatedMentalHealth),
##       family = binomial, data = train_df)
##
## Coefficients:
##                                     (Intercept)

```

```

## -0.551053
## as.factor(feelings_life)3 -19.581579
## as.factor(feelings_life)4 16.512902
## as.factor(feelings_life)5 0.008673
## as.factor(feelings_life)6 0.492938
## as.factor(feelings_life)7 0.711252
## as.factor(feelings_life)8 0.299564
## as.factor(feelings_life)9 0.159265
## as.factor(feelings_life)10 0.247421
## as.factor(place_birth_canada)Born outside Canada 0.395409
## as.factor(place_birth_canada)Don't know 1.532862
## as.factor(education)College, CEGEP or other non-university certificate or di... -30.945224
## as.factor(education)High school diploma or a high school equivalency certificate -31.372697
## as.factor(education)Less than high school diploma or its equivalent -0.049749
## as.factor(education)Trade certificate or diploma -1.355330
## as.factor(education)University certificate or diploma below the bachelor's level -0.842047
## as.factor(education)University certificate, diploma or degree above the bach... -21.305119
## as.factor(selfRatedMentalHealth)Fair -0.295141
## as.factor(selfRatedMentalHealth)Good -0.593105
## as.factor(selfRatedMentalHealth)Poor 0.290096
## as.factor(selfRatedMentalHealth)Very good -0.213357
##
## Degrees of Freedom: 2549 Total (i.e. Null); 2529 Residual
## Null Deviance: 2130
## Residual Deviance: 1313 AIC: 1355
##
## Call: glm(formula = hasSameEducation ~ age_first_child + as.factor(total_children) +
## as.factor(feelings_life) + as.factor(place_birth_canada) +
## as.factor(province) + as.factor(education) + as.factor(hh_type) +
## as.factor(partner_birth_province) + as.factor(average_hours_worked) +
## as.factor(selfRatedHealth) + as.factor(selfRatedMentalHealth) +
## as.factor(income_family) + as.factor(income_respondent),
## family = binomial, data = train_df)

```

```

##
## Coefficients:
##
## (Intercept)
## -17.770999
## age_first_child
## -0.009654
## as.factor(total_children)2
## -0.175792
## as.factor(total_children)3
## 0.027056
## as.factor(total_children)4
## -0.185489
## as.factor(total_children)5
## 0.444429
## as.factor(total_children)6
## -0.157155
## as.factor(total_children)7
## 0.152825
## as.factor(feelings_life)3
## -20.739932
## as.factor(feelings_life)4
## 15.735174
## as.factor(feelings_life)5
## -0.897031
## as.factor(feelings_life)6
## -0.403344
## as.factor(feelings_life)7
## -0.170160
## as.factor(feelings_life)8
## -0.597849
## as.factor(feelings_life)9
## -0.790773
## as.factor(feelings_life)10
## -0.670254
## as.factor(place_birth_canada)Born outside Canada
## 0.627877
## as.factor(place_birth_canada)Don't know
## 1.773604
## as.factor(province)British Columbia
## 0.252202
## as.factor(province)Manitoba
## 0.068344
## as.factor(province)New Brunswick
## 0.903245
## as.factor(province)Newfoundland and Labrador
## 0.866185
## as.factor(province)Nova Scotia
## 0.837005
## as.factor(province)Ontario
## 0.963726
## as.factor(province)Prince Edward Island
## -0.508870
## as.factor(province)Quebec
## -0.230198

```



```

##                                     as.factor(province)Saskatchewan
##                                     0.848437
## as.factor(education)College, CEGEP or other non-university certificate or di...
##                                     -31.302282
## as.factor(education)High school diploma or a high school equivalency certificate
##                                     -30.227783
## as.factor(education)Less than high school diploma or its equivalent
##                                     0.242033
## as.factor(education)Trade certificate or diploma
##                                     -1.289828
## as.factor(education)University certificate or diploma below the bachelor's level
##                                     -0.758986
## as.factor(education)University certificate, diploma or degree above the bach...
##                                     -21.535879
## as.factor(hh_type)High-rise apartment (5 or more stories)
##                                     -2.514771
## as.factor(hh_type)Low-rise apartment (less than 5 stories)
##                                     -1.538622
## as.factor(hh_type)Other
##                                     -2.055632
## as.factor(hh_type)Single detached house
##                                     -1.710735
## as.factor(partner_birth_province)British Columbia
##                                     0.114618
## as.factor(partner_birth_province)Manitoba
##                                     -0.022238
## as.factor(partner_birth_province)New Brunswick
##                                     -0.287278
## as.factor(partner_birth_province)Newfoundland and Labrador
##                                     -0.082147
## as.factor(partner_birth_province)Nova Scotia
##                                     -0.418124
## as.factor(partner_birth_province)Ontario
##                                     -0.554558
## as.factor(partner_birth_province)Outside Canada
##                                     -0.219124
## as.factor(partner_birth_province)Prince Edward Island
##                                     0.729108
## as.factor(partner_birth_province)Quebec
##                                     0.693048
## as.factor(partner_birth_province)Saskatchewan
##                                     -0.604531
## as.factor(partner_birth_province)Yukon / Northwest Territories / Nunavut
##                                     0.291787
## as.factor(average_hours_worked)0.1 to 29.9 hours
##                                     20.403319
## as.factor(average_hours_worked)30.0 to 40.0 hours
##                                     20.470465
## as.factor(average_hours_worked)40.1 to 50.0 hours
##                                     20.315810
## as.factor(average_hours_worked)50.1 hours and more
##                                     20.171525
## as.factor(average_hours_worked)Don't know
##                                     20.522417

```

```

##          as.factor(self_rated_health)Excellent
##                                     -1.182859
##          as.factor(self_rated_health)Fair
##                                     -1.168308
##          as.factor(self_rated_health)Good
##                                     -1.188375
##          as.factor(self_rated_health)Poor
##                                     -0.691678
##          as.factor(self_rated_health)Very good
##                                     -1.320975
##          as.factor(self_rated_mental_health)Fair
##                                     -0.049350
##          as.factor(self_rated_mental_health)Good
##                                     -0.628972
##          as.factor(self_rated_mental_health)Poor
##                                     0.537989
##          as.factor(self_rated_mental_health)Very good
##                                     -0.151457
##          as.factor(income_family)$125,000 and more
##                                     0.418493
##          as.factor(income_family)$25,000 to $49,999
##                                     -0.147494
##          as.factor(income_family)$50,000 to $74,999
##                                     0.012795
##          as.factor(income_family)$75,000 to $99,999
##                                     0.492312
##          as.factor(income_family)Less than $25,000
##                                     -0.057759
##          as.factor(income_respondent)$125,000 and more
##                                     0.444826
##          as.factor(income_respondent)$25,000 to $49,999
##                                     0.211286
##          as.factor(income_respondent)$50,000 to $74,999
##                                     0.191087
##          as.factor(income_respondent)$75,000 to $99,999
##                                     0.122290
##          as.factor(income_respondent)Less than $25,000
##                                     0.268872
##
## Degrees of Freedom: 2549 Total (i.e. Null);  2478 Residual
## Null Deviance:      2130
## Residual Deviance: 1252  AIC: 1396
##
## Call:  glm(formula = hasSameEducation ~ as.factor(education), family = binomial,
##           data = train_df)
##
## Coefficients:
##                                     (Intercept)
##                                     -0.3263
##  as.factor(education)College, CEGEP or other non-university certificate or di...
##                                     -20.2398
##  as.factor(education)High school diploma or a high school equivalency certificate
##                                     -20.2398

```

```
##          as.factor(education)Less than high school diploma or its equivalent
##                                                    -0.1592
##          as.factor(education)Trade certificate or diploma
##                                                    -1.4983
## as.factor(education)University certificate or diploma below the bachelor's level
##                                                    -0.8322
##  as.factor(education)University certificate, diploma or degree above the bach...
##                                                    -20.2398
##
## Degrees of Freedom: 2549 Total (i.e. Null);  2543 Residual
## Null Deviance:      2130
## Residual Deviance: 1346  AIC: 1360
```

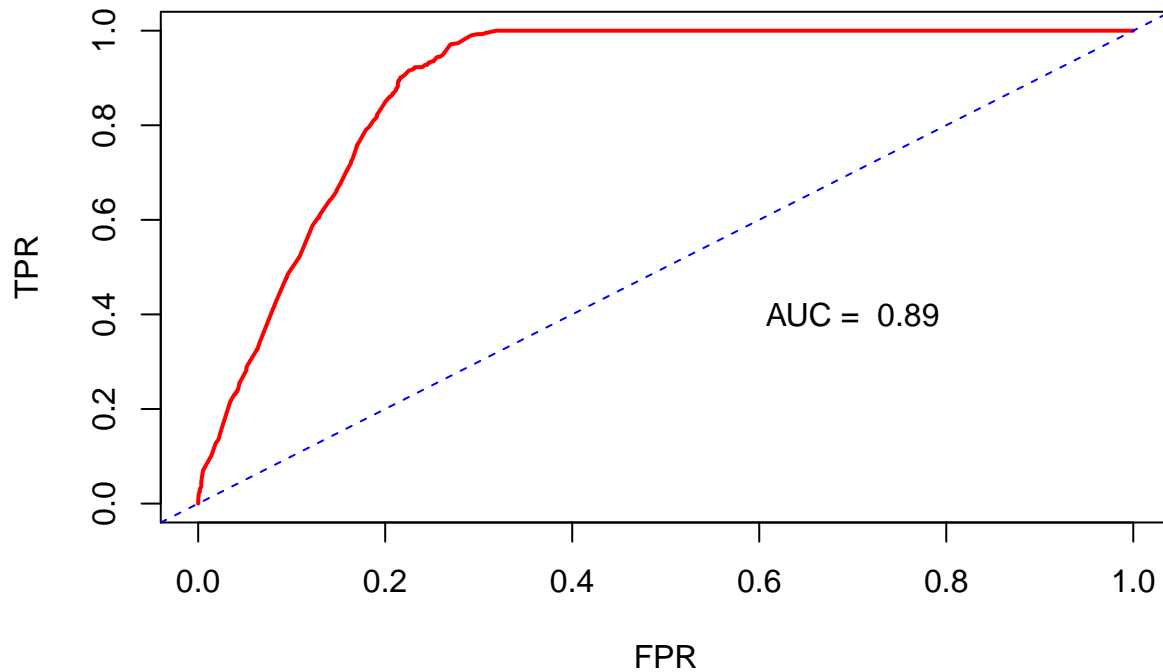
We performed variable selection through forward and backward AIC/BIC selection. Note that we have obtained the same model through the forward stepwise AIC method and the forward stepwise BIC method. To decide which model has a better fit, we performed the likelihood ratio test (forward stepwise AIC model as base model) and measured their prediction power through AUC value. We concluded that the backward stepwise AIC model performs the best in terms of Goodness of Fit and it is simple enough to provide good predictions.

Candidate.Model	Number.of.Predictors	Likelihood.Ratio.Test	AUC
Forward Stepwise AIC	13	1	0.91
Backward Stepwise AIC	4	0.1690	0.90
Backward Stepwise BIC	1	0.0028	0.88

Validation

We validated our model through its prediction accuracy in the test set. We observed that our model has an AUC value of 0.89, which is close to its AUC value with the training set (0.9). Therefore, our model generalizes well and provides good predictions.

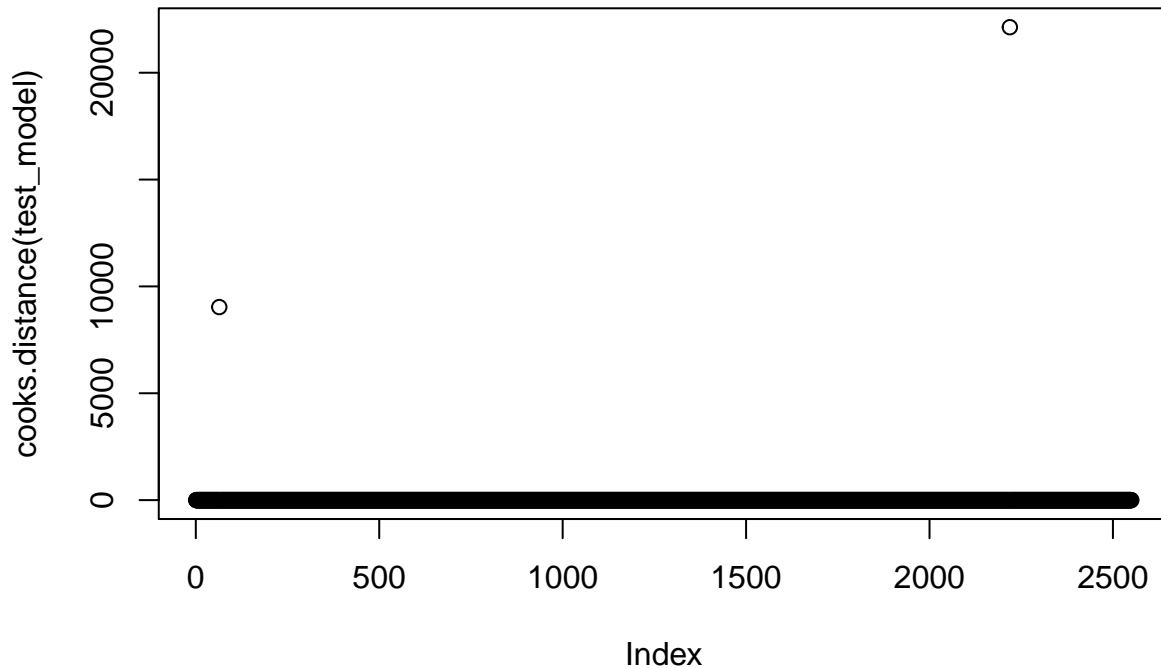
```
## Setting levels: control = FALSE, case = TRUE
## Setting direction: controls < cases
```



```
## [1] "AUC: 0.888534183572698"
```

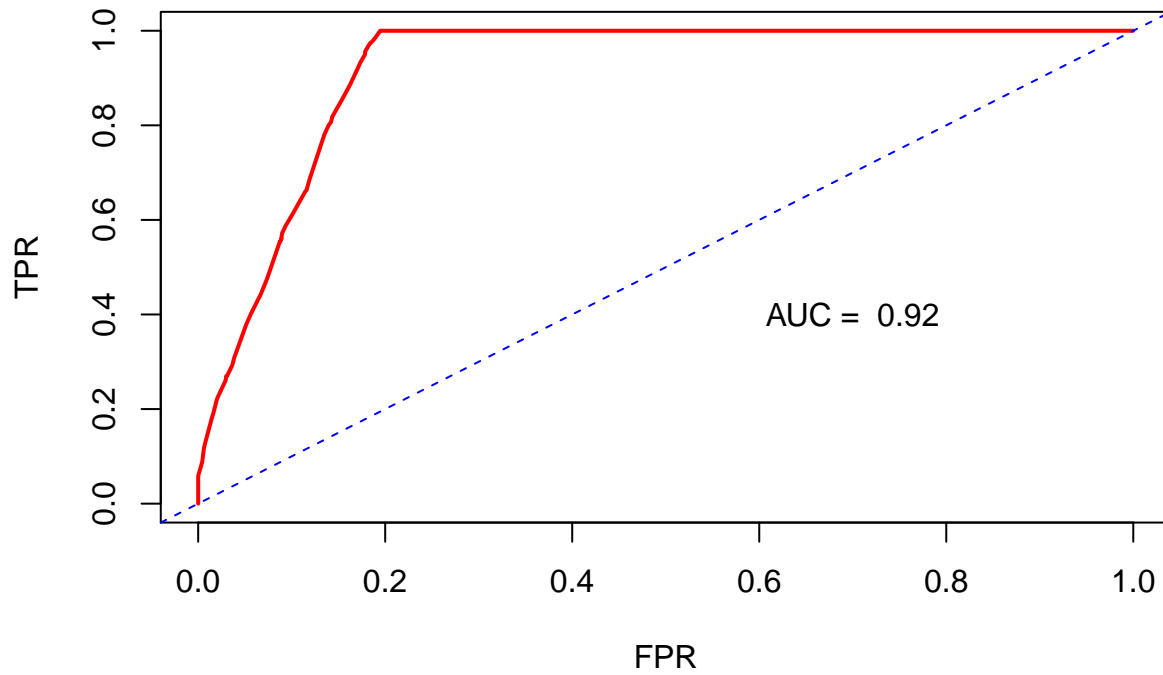
Influential Points

We identified influential points through Cook's Distance to assess observations' influence over the model fit. We observed that there are two observations that have significant effect on the model fit, hence we removed them and validate the model again. We observe that the model performs significantly better after removing the influential points.



```
## Setting levels: control = FALSE, case = TRUE
```

```
## Setting direction: controls < cases
```

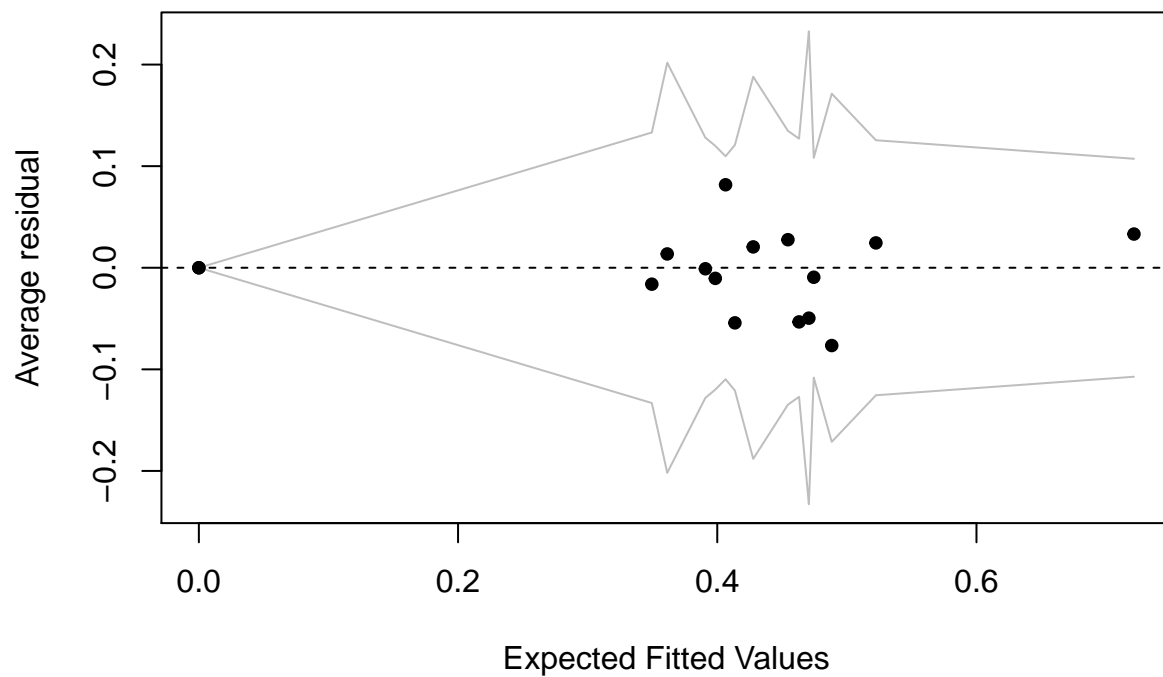


```
## [1] "AUC: 0.91884240808972"
```

Diagnostics

In this section, we will verify the model assumptions.

Binned residual plot



5. Results

```
##
## Call:
## glm(formula = hasSameEducation ~ as.factor(feelings_life) + as.factor(place_birth_canada) +
##      as.factor(education) + as.factor(self Rated mental health),
##      family = binomial, data = test_df_removed)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.273    0.000    0.000    0.000    1.461
##
## Coefficients:
##
## (Intercept)
## as.factor(feelings_life)4
## as.factor(feelings_life)5
## as.factor(feelings_life)6
## as.factor(feelings_life)7
## as.factor(feelings_life)8
## as.factor(feelings_life)9
## as.factor(feelings_life)10
## as.factor(place_birth_canada)Born outside Canada
## as.factor(place_birth_canada)Don't know
## as.factor(education)College, CEGEP or other non-university certificate or di...
## as.factor(education)High school diploma or a high school equivalency certificate
## as.factor(education)Less than high school diploma or its equivalent
## as.factor(education)Trade certificate or diploma
## as.factor(education)University certificate or diploma below the bachelor's level
## as.factor(education)University certificate, diploma or degree above the bach...
## as.factor(self Rated mental health)Excellent
## as.factor(self Rated mental health)Fair
## as.factor(self Rated mental health)Good
## as.factor(self Rated mental health)Poor
## as.factor(self Rated mental health)Very good
##
## (Intercept)
## as.factor(feelings_life)4
## as.factor(feelings_life)5
## as.factor(feelings_life)6
## as.factor(feelings_life)7
## as.factor(feelings_life)8
## as.factor(feelings_life)9
## as.factor(feelings_life)10
## as.factor(place_birth_canada)Born outside Canada
## as.factor(place_birth_canada)Don't know
## as.factor(education)College, CEGEP or other non-university certificate or di...
## as.factor(education)High school diploma or a high school equivalency certificate
## as.factor(education)Less than high school diploma or its equivalent
## as.factor(education)Trade certificate or diploma
## as.factor(education)University certificate or diploma below the bachelor's level
## as.factor(education)University certificate, diploma or degree above the bach...
## as.factor(self Rated mental health)Excellent
## as.factor(self Rated mental health)Fair
```

	Estimate	Std. Error
(Intercept)	9.7743	31155.7152
as.factor(feelings_life)4	-14.5907	18659.7365
as.factor(feelings_life)5	-16.3560	11187.1143
as.factor(feelings_life)6	2.5743	10750.8321
as.factor(feelings_life)7	2.8484	10716.8498
as.factor(feelings_life)8	2.8338	10716.8498
as.factor(feelings_life)9	2.9043	10716.8498
as.factor(feelings_life)10	2.8822	10716.8498
as.factor(place_birth_canada)Born outside Canada	0.3331	0.1694
as.factor(place_birth_canada)Don't know	17.3950	1567.6617
as.factor(education)College, CEGEP or other non-university certificate or di...	-34.5778	1166.2708
as.factor(education)High school diploma or a high school equivalency certificate	-33.9275	1299.7609
as.factor(education)Less than high school diploma or its equivalent	-0.1959	0.1938
as.factor(education)Trade certificate or diploma	-34.8317	2009.5174
as.factor(education)University certificate or diploma below the bachelor's level	-33.9536	2808.3085
as.factor(education)University certificate, diploma or degree above the bach...	-34.5480	1483.6883
as.factor(self Rated mental health)Excellent	-13.0364	29255.6521
as.factor(self Rated mental health)Fair	4.1717	29249.9686
as.factor(self Rated mental health)Good	-13.0586	
as.factor(self Rated mental health)Poor	-13.7192	
as.factor(self Rated mental health)Very good	-12.7905	

```

## as.factor(self Rated Mental Health)Good 29255.6521
## as.factor(self Rated Mental Health)Poor 32233.9823
## as.factor(self Rated Mental Health)Very good 29255.6521
## z value
## (Intercept) 0.000
## as.factor(feelings Life)4 -0.001
## as.factor(feelings Life)5 -0.001
## as.factor(feelings Life)6 0.000
## as.factor(feelings Life)7 0.000
## as.factor(feelings Life)8 0.000
## as.factor(feelings Life)9 0.000
## as.factor(feelings Life)10 0.000
## as.factor(place Birth Canada)Born outside Canada 1.966
## as.factor(place Birth Canada)Don't know 0.011
## as.factor(education)College, CEGEP or other non-university certificate or di... -0.030
## as.factor(education)High school diploma or a high school equivalency certificate -0.026
## as.factor(education)Less than high school diploma or its equivalent -1.011
## as.factor(education)Trade certificate or diploma -0.017
## as.factor(education)University certificate or diploma below the bachelor's level -0.012
## as.factor(education)University certificate, diploma or degree above the bach... -0.023
## as.factor(self Rated Mental Health)Excellent 0.000
## as.factor(self Rated Mental Health)Fair 0.000
## as.factor(self Rated Mental Health)Good 0.000
## as.factor(self Rated Mental Health)Poor 0.000
## as.factor(self Rated Mental Health)Very good 0.000
## Pr(>|z|)
## (Intercept) 0.9997
## as.factor(feelings Life)4 0.9994
## as.factor(feelings Life)5 0.9988
## as.factor(feelings Life)6 0.9998
## as.factor(feelings Life)7 0.9998
## as.factor(feelings Life)8 0.9998
## as.factor(feelings Life)9 0.9998
## as.factor(feelings Life)10 0.9998
## as.factor(place Birth Canada)Born outside Canada 0.0493
## as.factor(place Birth Canada)Don't know 0.9911
## as.factor(education)College, CEGEP or other non-university certificate or di... 0.9763
## as.factor(education)High school diploma or a high school equivalency certificate 0.9792
## as.factor(education)Less than high school diploma or its equivalent 0.3121
## as.factor(education)Trade certificate or diploma 0.9862
## as.factor(education)University certificate or diploma below the bachelor's level 0.9904
## as.factor(education)University certificate, diploma or degree above the bach... 0.9814
## as.factor(self Rated Mental Health)Excellent 0.9996
## as.factor(self Rated Mental Health)Fair 0.9999
## as.factor(self Rated Mental Health)Good 0.9996
## as.factor(self Rated Mental Health)Poor 0.9997
## as.factor(self Rated Mental Health)Very good 0.9997
##
## (Intercept)
## as.factor(feelings Life)4
## as.factor(feelings Life)5
## as.factor(feelings Life)6
## as.factor(feelings Life)7
## as.factor(feelings Life)8

```

```
## as.factor(feelings_life)9
## as.factor(feelings_life)10
## as.factor(place_birth_canada)Born outside Canada *
## as.factor(place_birth_canada)Don't know
## as.factor(education)College, CEGEP or other non-university certificate or di...
## as.factor(education)High school diploma or a high school equivalency certificate
## as.factor(education)Less than high school diploma or its equivalent
## as.factor(education)Trade certificate or diploma
## as.factor(education)University certificate or diploma below the bachelor's level
## as.factor(education)University certificate, diploma or degree above the bach...
## as.factor(self_rated_mental_health)Excellent
## as.factor(self_rated_mental_health)Fair
## as.factor(self_rated_mental_health)Good
## as.factor(self_rated_mental_health)Poor
## as.factor(self_rated_mental_health)Very good
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1976.15  on 2446  degrees of freedom
## Residual deviance:  993.97  on 2426  degrees of freedom
## AIC: 1036
##
## Number of Fisher Scoring iterations: 20
```

Here you will include all results. This includes descriptive statistics, graphs, figures, tables, and model results. Please ensure that everything is well formatted and in a report style. You must also provide an explanation of the results in this section. You can overflow to an Appendix if needed.

Please ensure that everything is well labeled. So if you have multiple histograms and plots, calling them Figure 1, 2, 3, etc. and referencing them as Figure 1, Figure 2, etc. in your report will be expected. The reader should not get lost in a sea of information. Make sure to have the results be clean, well formatted and digestible.

6. Discussion

Addressing general social survey, questionnaire is one of the ways to collect data from every potential target population. Since the online self-completed questionnaire is purposed to indicate some basic personal information and household information, and the other income data are collected through some administrative files, the questions in the questionnaire could be more reliable that reduce the burden and improve accuracy of data. Furthermore, the questionnaire increases initiative while it is an easy way to collect every respondents' data. On the other hand, due to some situations that there could be some respondents not able to complete the questionnaire, the response rate of the questionnaire might never be hundred-percent fulfilled as expected. Equivalently, the result interpreted from the data might not represent the Canadian residents' information accurately, because not all of the target population are included in the dataset. In terms of this questionnaire of general social survey, some questions are ambiguous to answer for some special groups and there are some categories that are not representative of some respondents' situations. According to the overview of the general social survey report in 2017, the questions are not mandatory to respond to, which leads to many NA entries in the data set.

7. Weaknesses

Here we discuss weaknesses of the study, data, analysis, etc. You can also discuss areas for improvement.

8. Next Steps

Here you discuss subsequent work to be done after this report. This can include next steps in terms of statistical analysis (perhaps there is a more efficient algorithm available, or perhaps there is a caveat in the data that would allow for some new technique). Future steps should also be specified in terms of the study setting (eg. including a follow-up survey on something, or a subsequent study that would complement the conclusions of your report).

9. References

1. “Pandoc User’s Guide.” Pandoc, pandoc.org/MANUAL.html#multiline-tables.
2. Sonderegger, Derek L. “A Sufficient Introduction to R.” Derek Sonderegger, Ph.D., 3 Oct. 2017, dereksonderegger.github.io/570L/15-rmarkdown-tricks.html.
3. Yihui Xie, Christophe Dervieux. “R Markdown Cookbook.” 10.1 The Function Knitr::Kable(), 21 Sept. 2020, bookdown.org/yihui/rmarkdown-cookbook/kable.html.
4. Comprehensive Reference to the Information Available from the General Social Survey (GSS), Government of Canada, Statistics Canada, 20 Feb. 2019, www150.statcan.gc.ca/n1/pub/89f0115x/89f0115x2019001-eng.htm.
5. General Social Survey C31 MAIN SURVEY- FAMILY, https://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=assembleInstr&lang=en&Item_Id=335815