# Title of Your Report

# Names of your Group Members

## Due Date

## Title of your Report

# Name(s) of Author(s)

## Date

## 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.

#### Introduction

Here is where you should give insight into the setting and introduce the goal of the analysis. Here you can introduce ideas and basic concepts regarding the study setting and the potential model. Again, this is the introduction, so you should be explaining the importance of the work that is ahead and hopefully build some suspense for the reader. You can also highlight what will be included in the subsequent sections.

#### Data

Introduce the data, explain why it was selected. Make sure to comment on important features and highlight any potential drawbacks to the data.

#### **Data Cleaning**

```
## -- Attaching packages ------ tidyverse 1.3.0 --
## v ggplot2 3.3.2
                   v purrr
                           0.3.4
## v tibble 3.0.4
                           1.0.2
                   v dplyr
## v tidvr
          1.1.2
                   v stringr 1.4.0
## v readr
          1.4.0
                  v forcats 0.5.0
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()
                 masks stats::lag()
```

#### Split Data

## Model

Introduce the selected model here. It is expected that you will use some mathematical notation here. If you do please ensure that all notation is explained. You may also want to discuss any special (hypothetical) cases of your model here, as well as any caveats.

#### Base Model

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

## Variable Selection and Regularization

```
##
  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(self_rated_health) + as.factor(self_rated_mental_health) +
##
       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
##	as.lactor(partner_birtn_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
##	as.lactor(average_nours_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:
## Residual Deviance: 1252 AIC: 1396
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
##
##
   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 = 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(self_rated_mental_health)Fair
##
                                                                            -0.295141
##
                                             as.factor(self_rated_mental_health)Good
##
                                                                            -0.593105
##
                                             as.factor(self_rated_mental_health)Poor
##
                                        as.factor(self_rated_mental_health)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(self_rated_health) + as.factor(self_rated_mental_health) +
##
       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
##	as.lactor(province)Newroundrand 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
## ##	as.lactor(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
##
```

## Null Deviance: 2130 ## Residual Deviance: 1252 AIC: 1396 ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred ## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
  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
##
##
   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:
## Residual Deviance: 1346 AIC: 1360
```

#### Forward stepwise AIC/BIC Model

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

#### Backward stepwise AIC Model

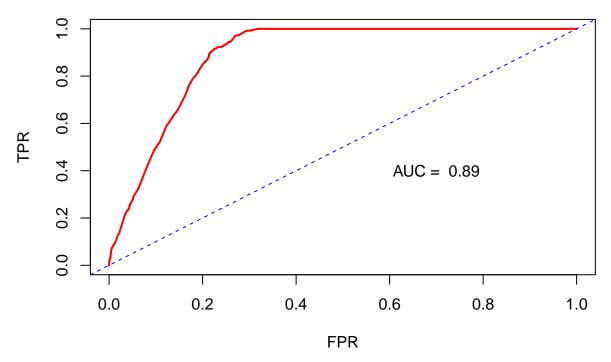
```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

## Backward stepwise BIC Model

## Setting direction: controls < cases

#### LRT

```
## Loading required package: foreign
## Loading required package: survival
## Loading required package: MASS
##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##
##
       select
## Loading required package: nnet
## Attaching package: 'epiDisplay'
## The following object is masked from 'package:ggplot2':
##
       alpha
## Likelihood ratio test for MLE method
## Chi-squared 51 d.f. = 60.55973 , P value = 0.168971
## Likelihood ratio test for MLE method
## Chi-squared 14 d.f. = 33.08574 , P value = 0.002800105
Validation
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Setting levels: control = FALSE, case = TRUE
```

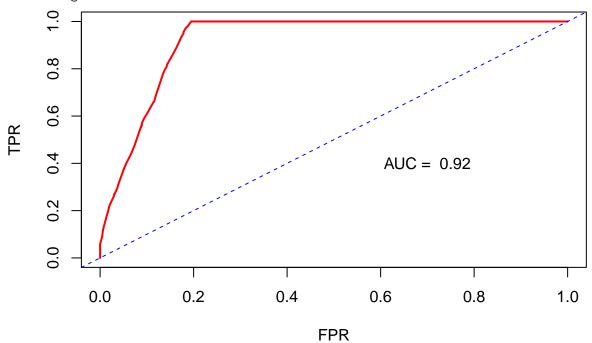


## [1] "AUC: 0.888534183572698"

# Influential Points

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Setting levels: control = FALSE, case = TRUE

## Setting direction: controls < cases



## [1] "AUC: 0.91884240808972"

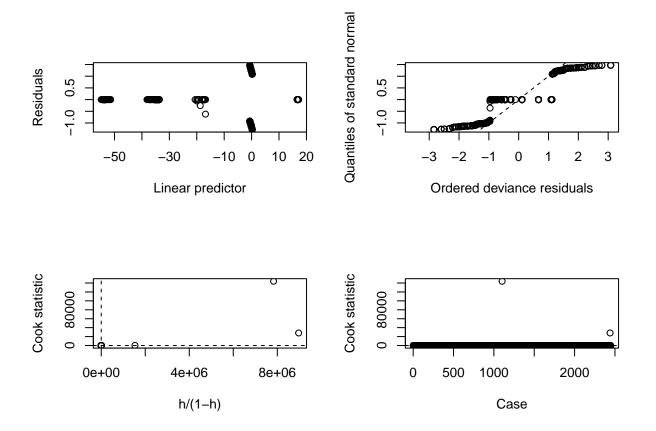
```
##
## Call:
  glm(formula = hasSameEducation ~ as.factor(feelings life) + as.factor(place birth canada) +
##
       as.factor(education) + as.factor(self_rated_mental_health),
       family = binomial, data = test_df_removed)
##
## Deviance Residuals:
##
      Min
               10 Median
                                30
                                       Max
## -1.273
            0.000
                    0.000
                            0.000
                                     1.461
##
  Coefficients:
##
                                                                                        Estimate
## (Intercept)
                                                                                          9.7743
## as.factor(feelings_life)4
                                                                                        -14.5907
## as.factor(feelings_life)5
                                                                                        -16.3560
## as.factor(feelings_life)6
                                                                                          2.5743
## as.factor(feelings_life)7
                                                                                          2.8484
## as.factor(feelings life)8
                                                                                          2.8338
## as.factor(feelings_life)9
                                                                                          2.9043
## as.factor(feelings life)10
                                                                                          2.8822
## as.factor(place_birth_canada)Born outside Canada
                                                                                          0.3331
## as.factor(place_birth_canada)Don't know
                                                                                         17.3950
## as.factor(education)College, CEGEP or other non-university certificate or di...
                                                                                        -34.5778
## as.factor(education)High school diploma or a high school equivalency certificate
                                                                                        -33.9275
## as.factor(education)Less than high school diploma or its equivalent
                                                                                         -0.1959
## as.factor(education)Trade certificate or diploma
                                                                                        -34.8317
## as.factor(education)University certificate or diploma below the bachelor's level
                                                                                        -33.9536
## as.factor(education)University certificate, diploma or degree above the bach...
                                                                                        -34.5480
## as.factor(self_rated_mental_health)Excellent
                                                                                        -13.0364
## as.factor(self_rated_mental_health)Fair
                                                                                          4.1717
## 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
##
                                                                                      Std. Error
## (Intercept)
                                                                                      31155.7152
## as.factor(feelings life)4
                                                                                      18659.7365
## as.factor(feelings life)5
                                                                                      11187.1143
## as.factor(feelings_life)6
                                                                                      10750.8321
## as.factor(feelings life)7
                                                                                      10716.8498
## as.factor(feelings_life)8
                                                                                      10716.8498
## as.factor(feelings life)9
                                                                                      10716.8498
## as.factor(feelings life)10
                                                                                      10716.8498
## as.factor(place birth canada)Born outside Canada
                                                                                          0.1694
## as.factor(place_birth_canada)Don't know
                                                                                       1567.6617
## as.factor(education)College, CEGEP or other non-university certificate or di...
                                                                                       1166.2708
## as.factor(education)High school diploma or a high school equivalency certificate
                                                                                       1299.7609
## as.factor(education)Less than high school diploma or its equivalent
                                                                                          0.1938
## as.factor(education)Trade certificate or diploma
                                                                                       2009.5174
## as.factor(education)University certificate or diploma below the bachelor's level
                                                                                       2808.3085
## as.factor(education)University certificate, diploma or degree above the bach...
                                                                                       1483.6883
## as.factor(self_rated_mental_health)Excellent
                                                                                      29255.6521
## as.factor(self_rated_mental_health)Fair
                                                                                      29249.9686
## 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
                                                                                        0.000
## (Intercept)
## 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
                                                                                        0.000
## as.factor(self_rated_mental_health)Excellent
## 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 '**' 0.05 '.' 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
Diagnostics
##
## Attaching package: 'boot'
## The following object is masked from 'package:survival':
##
```

##

aml



## Results

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 labelled. 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.

# Discussion

Here you will discuss conclusions drawn from the results and comment on how it relates to the original goal of the study (which was specified in the Introduction).

# Weaknesses

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

# **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).

## References