Causal HW

2022-10-02

Homework 1

Causal Inference via Econometrics and Experimentation

Rithwik Sivadasan (sivad001) and Mohammed Pahadwala (pahad002)

R Markdown

Loading the Required Packages

```
library(readxl)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.1.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(stringr)
```

```
## Warning: package 'stringr' was built under R version 4.1.3
```

```
library(readxl)
library(mltools)
```

```
## Warning: package 'mltools' was built under R version 4.1.3
```

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 4.1.3
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 4.1.3
```

```
## Loading required package: lattice
```

```
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 4.1.3
```

```
## corrplot 0.92 loaded
```

```r
library(tidytext)
```

```
## Warning: package 'tidytext' was built under R version 4.1.3
```

```r
library(stringr)
library(textdata)
```

```
## Warning: package 'textdata' was built under R version 4.1.3
```

```r
library(gmodels)
```

```
## Warning: package 'gmodels' was built under R version 4.1.3
```

Loading the data For Q1

```r
reddit_data <- read.csv('C:/Users/mdphd19/Documents/Carlsons/Fall/Causal/Homework/Assignment 1/data_Q1.
summary(reddit_data)
```

```
##     user_id            treated     intervention_post_score      tenure
##  Min.   :   1.0   Min.   :0.0   Min.   :   3.00          Min.   :   0.0
##  1st Qu.: 453.2   1st Qu.:0.0   1st Qu.:   6.00          1st Qu.:  23.0
##  Median : 905.5   Median :0.5   Median :  14.00          Median : 253.5
##  Mean   : 905.5   Mean   :0.5   Mean   :  45.05          Mean   : 549.9
##  3rd Qu.:1357.8   3rd Qu.:1.0   3rd Qu.:  36.75          3rd Qu.: 849.5
##  Max.   :1810.0   Max.   :1.0   Max.   :1762.00          Max.   :4579.0
##   premium_user     comment_karma      link_karma      num_post_before
##  Min.   :0.0000   Min.   :   -98   Min.   :    1.0   Min.   : 0.000
##  1st Qu.:0.0000   1st Qu.:     0   1st Qu.:   14.0   1st Qu.: 0.000
##  Median :0.0000   Median :    51   Median :  114.0   Median : 0.000
##  Mean   :0.0221   Mean   :  2925   Mean   : 1509.6   Mean   : 1.597
##  3rd Qu.:0.0000   3rd Qu.:  1001   3rd Qu.:  957.5   3rd Qu.: 2.000
##  Max.   :1.0000   Max.   :160591   Max.   :59079.0   Max.   :45.000
##  num_post_before_ss num_sub_before   first_timer         posted
##  Min.   :0.0000     Min.   : 0.00   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.: 0.00   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000     Median : 0.00   Median :1.0000   Median :1.0000
##  Mean   :0.2889     Mean   : 1.03   Mean   :0.5398   Mean   :0.5917
##  3rd Qu.:0.0000     3rd Qu.: 1.00   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :8.0000     Max.   :21.00   Max.   :1.0000   Max.   :1.0000
##    posted_ss        posted_oths        n_posts          l_n_posts
##  Min.   :0.0000   Min.   :0.0000   Min.   :  0.000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:  0.000   1st Qu.:0.0000
##  Median :0.0000   Median :0.0000   Median :  1.000   Median :0.6931
##  Mean   :0.3354   Mean   :0.4182   Mean   :  2.791   Mean   :0.8271
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:  3.000   3rd Qu.:1.3863
##  Max.   :1.0000   Max.   :1.0000   Max.   :182.000   Max.   :5.2095
##    n_posts_ss       l_n_posts_ss      n_posts_oths         length
##  Min.   : 0.000   Min.   :0.0000   Min.   :  0.000   Min.   :   0.000
##  1st Qu.: 0.000   1st Qu.:0.0000   1st Qu.:  0.000   1st Qu.:   0.000
##  Median : 0.000   Median :0.0000   Median :  0.000   Median :   8.667
##  Mean   : 0.684   Mean   :0.3361   Mean   :  2.107   Mean   : 467.388
##  3rd Qu.: 1.000   3rd Qu.:0.6931   3rd Qu.:  2.000   3rd Qu.: 559.932
##  Max.   :12.000   Max.   :2.5649   Max.   :176.000   Max.   :6498.500
##   l_length_ss     l_length_oths
##  Min.   :0.000   Min.   :0.000
##  1st Qu.:0.000   1st Qu.:0.000
```

```
##  Median :0.000    Median :0.000
##  Mean   :2.351    Mean   :1.478
##  3rd Qu.:6.523    3rd Qu.:3.163
##  Max.   :8.834    Max.   :8.731
```

Question 1:

Platforms use various methods to stimulate user's content creation. This includes paying users for reviews and providing awards and badges to users. Reddit is one of the largest platforms for creating and sharing content. On Reddit, users can recognize other contributions by providing gold to each other. However, does getting Reddit gold actually increase the receiver's content generation?

To find out, researchers gave 905 random users reddit gold. Data is included for a similar number of users in the control group who did not receive gold during the time of the experiment. Import the data and examine:

  a) If the control and treatment groups are similar across tenure, premium_user, and num_posts_before metrics.

  b) Does getting reddit gold increase likelihood that the user will post (use the posted metric as the dependent variable and treated as the independent variable)? Use a simple linear model (not a logit) for the analysis.

  c) What sorts of users are more likely to increase their contribution? (use the tenure and the first_timer variables)

  d) Is the SUTVA assumption likely to be violated in the experiment?

#Solution

a)

Checking if the treatment and control groups are same across tenure, premium_user, and num_posts_before

```
t.test(tenure~treated,reddit_data)
```

```
##
##  Welch Two Sample t-test
##
## data:  tenure by treated
## t = 1.373, df = 1789.6, p-value = 0.1699
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -19.09774 108.23144
## sample estimates:
## mean in group 0 mean in group 1
##        572.1680        527.6011
```

```
t.test(premium_user~treated,reddit_data)
```

```
##
##  Welch Two Sample t-test
##
## data:  premium_user by treated
## t = 0.95906, df = 1769.9, p-value = 0.3377
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.006928414  0.020188082
## sample estimates:
```

```
## mean in group 0 mean in group 1
##      0.02541436      0.01878453
```

```
t.test(num_post_before~treated,reddit_data)
```

```
##
##  Welch Two Sample t-test
##
## data:  num_post_before by treated
## t = 0.56253, df = 1796.1, p-value = 0.5738
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.2307971  0.4164325
## sample estimates:
## mean in group 0 mean in group 1
##        1.643094        1.550276
```

We see that the p-value for tenure test is 0.1699, which is well above our threshold of 0.05. Hence, we have enough evidence to say that the treatment and control groups have similar tenures.

We see that the p-value for premium_user test is 0.3377, which is well above our threshold of 0.05. Hence, we have enough evidence to say that the treatment and control groups have similar number of premium users.

We see that the p-value for num_posts_before test is 0.5738, which is well above our threshold of 0.05. Hence, we have enough evidence to say that the treatment and control groups have similar number of posts before this experiment.

b)

Testing relationship between getting a gold and posting on the

```
lm_q1_1 = lm(posted~treated,reddit_data)
summary(lm_q1_1)
```

```
##
## Call:
## lm(formula = posted ~ treated, data = reddit_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6232 -0.5602  0.3768  0.4398  0.4398
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.56022    0.01631   34.34   <2e-16 ***
## treated      0.06298    0.02307    2.73   0.0064 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4908 on 1808 degrees of freedom
## Multiple R-squared:  0.004105,   Adjusted R-squared:  0.003554
## F-statistic: 7.452 on 1 and 1808 DF,  p-value: 0.006396
```

Yes, getting a reddit gold does incease the likelihood that a user will post. The coefficient from the linear regression model comes out to be 0.0628. This means that the likelihood of the user posting increases by 0.0628 by getting a reddit gold. The results are statistically significant as denoted by the p-value of the coefficient.

c)

Checking if tenure of the user or if the user is a first_timer increases the chance of posting.

```
lm_q1_2_1 = lm(posted~treated+tenure,reddit_data)
summary(lm_q1_2_1)
```

```
##
## Call:
## lm(formula = posted ~ treated + tenure, data = reddit_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6365 -0.5740  0.3645  0.4254  0.5410
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.747e-01  1.890e-02  30.399  < 2e-16 ***
## treated      6.186e-02  2.308e-02   2.681  0.00741 **
## tenure      -2.526e-05  1.671e-05  -1.512  0.13077
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4906 on 1807 degrees of freedom
## Multiple R-squared:  0.005363,   Adjusted R-squared:  0.004262
## F-statistic: 4.872 on 2 and 1807 DF,  p-value: 0.007762
```

```
lm_q1_2_2 = lm(posted~treated+first_timer,reddit_data)
summary(lm_q1_2_2)
```

```
##
## Call:
## lm(formula = posted ~ treated + first_timer, data = reddit_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6673 -0.5875  0.3327  0.4125  0.4775
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.60225    0.02031  29.650  < 2e-16 ***
## treated      0.06501    0.02301   2.825 0.004775 **
## first_timer -0.07973    0.02308  -3.454 0.000565 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4893 on 1807 degrees of freedom
## Multiple R-squared:  0.01064,    Adjusted R-squared:  0.009543
## F-statistic: 9.715 on 2 and 1807 DF,  p-value: 6.361e-05
```

From the above models, we can say that if the user is a first timer, it will not increase the chance of contribution. The coefficient gained is less than zero (negative) hence suggesting that if the user is first timer, there is less likelihood of it posting on the platform. We cannot say for sure about tenure as the probability value large and the results are not statistically significant. There is not enough evidence that there is a relationship between tenure and posting on reddit.

d)

Considering economic value of getting a reddit gold, you can use some premium features of reddit like /rlounge which is premium only access. Moreover, a reddit gold user can get wide audience by using the features available exclusively for the gold member. Thus, to reach out to a wider audience a user can employ more than one person to maintain the account. The users that are employed by the gold user could be the part of control group thus effecting the activities in the control group. We could say that SUTVA assumption is violated in the experiment.

Question 2

a)Use a t-test to see if there is a statistical difference in the pre-period between schools in the treatment (bal = 1) and control (bal = 0). This will check if randomization has been done correctly. To do this, calculate the average normalized test score(norm) for the pre period (pre = 1) for math (test_type = 0). Is there a statistical difference between students who got the Balsakhi program and did not get the program? Perform the same test for language (test_type = 1)

filtering data to pre period for math and language

Loading the data For Q2

```
data_Q2 <- read.csv('C:/Users/mdphd19/Documents/Carlsons/Fall/Causal/Homework/Assignment 1/data_Q2.csv')
summary(data_Q2)
```

```
##    studentid            div                std            schoolid
## Min.   :      91   Length:91782       Min.   :3.000   Min.   :101.0
## 1st Qu.:  13203    Class :character   1st Qu.:3.000   1st Qu.:210.0
## Median :12041583   Mode  :character   Median :3.000   Median :331.0
## Mean   :20386646                      Mean   :3.497   Mean   :354.4
## 3rd Qu.:34231143                      3rd Qu.:4.000   3rd Qu.:502.0
## Max.   :64041343                      Max.   :4.000   Max.   :640.0
##
##      divid       researchgroup       sessiond         precheated
## Min.   :31010   Min.   : 1.000    Length:91782       Length:91782
## 1st Qu.:33310   1st Qu.: 4.000    Class :character   Class :character
## Median :36395   Median : 6.000    Mode  :character   Mode  :character
## Mean   :38509   Mean   : 5.912
## 3rd Qu.:43310   3rd Qu.: 8.000
## Max.   :46400   Max.   :11.000
##                 NA's   :1026
##   prehelped           prelater            pre_tot       prepapersubtotal
## Length:91782       Length:91782       Min.   : 0.00    Min.   :100
## Class :character   Class :character   1st Qu.:14.00    1st Qu.:100
## Mode  :character   Mode  :character   Median :28.00    Median :100
##                                       Mean   :31.93    Mean   :100
##                                       3rd Qu.:48.00    3rd Qu.:100
##                                       Max.   :94.00    Max.   :100
##
##   midcheated          midhelped           midlater           mid_tot
## Length:91782       Length:91782       Length:91782       Min.   :  0.0
## Class :character   Class :character   Class :character   1st Qu.: 25.0
## Mode  :character   Mode  :character   Mode  :character   Median : 45.0
##                                                          Mean   : 45.4
##                                                          3rd Qu.: 65.0
##                                                          Max.   :238.0
```

```
##                                                            NA's   :10458
## midpapersubtotal postcheated       posthelped        postlater
## Min.   :100      Length:91782      Length:91782      Length:91782
## 1st Qu.:100      Class :character  Class :character  Class :character
## Median :100      Mode  :character  Mode  :character  Mode  :character
## Mean   :100
## 3rd Qu.:100
## Max.   :100
## NA's   :10458
##    post_tot      postpapersubtotal attritprepost     attritpremid
## Min.   : 0.00    Min.   :100       Min.   :0.0000   Min.   :0.0000
## 1st Qu.:21.00    1st Qu.:100       1st Qu.:0.0000   1st Qu.:0.0000
## Median :40.00    Median :100       Median :0.0000   Median :0.0000
## Mean   :42.19    Mean   :100       Mean   :0.1738   Mean   :0.1139
## 3rd Qu.:62.00    3rd Qu.:100       3rd Qu.:0.0000   3rd Qu.:0.0000
## Max.   :99.00    Max.   :100       Max.   :1.0000   Max.   :1.0000
## NA's   :15948    NA's   :15948
## attritmidpost      pre_verb         pre_math         mid_verb
## Min.   :0.0000   Min.   : 0.00    Min.   : 0.00    Min.   : 0.00
## 1st Qu.:0.0000   1st Qu.: 6.00    1st Qu.: 7.00    1st Qu.:12.00
## Median :0.0000   Median :11.00    Median :16.00    Median :21.00
## Mean   :0.1188   Mean   :14.31    Mean   :17.62    Mean   :22.49
## 3rd Qu.:0.0000   3rd Qu.:21.00    3rd Qu.:27.00    3rd Qu.:33.00
## Max.   :1.0000   Max.   :49.00    Max.   :50.00    Max.   :94.00
##                                                    NA's   :10782
##    mid_math        post_verb        post_math           bal              male
## Min.   :  0.00   Min.   : 0.00    Min.   : 0.0     Min.   :0.0000   Min.   :0
## 1st Qu.: 13.00   1st Qu.:11.00    1st Qu.: 9.0     1st Qu.:0.0000   1st Qu.:0
## Median : 23.00   Median :20.00    Median :20.0     Median :0.0000   Median :0
## Mean   : 22.95   Mean   :21.59    Mean   :20.6     Mean   :0.4893   Mean   :0
## 3rd Qu.: 33.00   3rd Qu.:32.00    3rd Qu.:31.0     3rd Qu.:1.0000   3rd Qu.:0
## Max.   :144.00   Max.   :50.00    Max.   :50.0     Max.   :1.0000   Max.   :0
## NA's   :10458    NA's   :15948    NA's   :15948
##     rank            bot20            third           numstud
## Min.   :  1.00   Min.   :0.0000   Min.   :1.000    Min.   :  7.0
## 1st Qu.: 14.00   1st Qu.:0.0000   1st Qu.:1.000    1st Qu.: 44.0
## Median : 28.00   Median :0.0000   Median :2.000    Median : 62.0
## Mean   : 33.05   Mean   :0.3695   Mean   :1.995    Mean   : 65.1
## 3rd Qu.: 46.50   3rd Qu.:1.0000   3rd Qu.:3.000    3rd Qu.: 82.0
## Max.   :143.00   Max.   :1.0000   Max.   :3.000    Max.   :143.0
##
## thirdinschool    bigschool        numthird            pre
## Min.   :1.000    Min.   :0.00     Min.   : 2.00    Min.   :0.0000
## 1st Qu.:1.000    1st Qu.:0.00     1st Qu.:15.00    1st Qu.:0.0000
## Median :2.000    Median :1.00     Median :21.00    Median :0.0000
## Mean   :2.011    Mean   :0.66     Mean   :21.74    Mean   :0.3333
## 3rd Qu.:3.000    3rd Qu.:1.00     3rd Qu.:27.00    3rd Qu.:1.0000
## Max.   :3.000    Max.   :1.00     Max.   :50.00    Max.   :1.0000
##
##      mid              post           test_type        test
## Min.   :0.0000   Min.   :0.0000    Min.   :0    Min.   :  0.00
## 1st Qu.:0.0000   1st Qu.:0.0000    1st Qu.:0    1st Qu.: 11.00
## Median :0.0000   Median :0.0000    Median :1    Median : 23.00
## Mean   :0.3333   Mean   :0.3333    Mean   :1    Mean   : 26.31
```

```
##   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:2   3rd Qu.: 37.00
##   Max.   :1.0000   Max.   :1.0000   Max.   :2   Max.   :238.00
##                                                 NA's   :8838
##      bal_mid          bal_post          mean            sd
##   Min.   :0.0000   Min.   :0.0000   Min.   :10.83   Min.   : 9.633
##   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:12.83   1st Qu.:10.734
##   Median :0.0000   Median :0.0000   Median :17.62   Median :11.317
##   Mean   :0.1631   Mean   :0.1631   Mean   :21.24   Mean   :14.003
##   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:23.66   3rd Qu.:19.772
##   Max.   :1.0000   Max.   :1.0000   Max.   :40.19   Max.   :21.025
##
##       norm           classmean         classsd         classnorm
##   Min.   :-1.954   Min.   : 4.50   Min.   : 4.87   Min.   :-2.68
##   1st Qu.:-0.580   1st Qu.:23.41   1st Qu.:16.96   1st Qu.:-0.78
##   Median : 0.222   Median :31.72   Median :19.35   Median :-0.16
##   Mean   : 0.362   Mean   :31.93   Mean   :18.88   Mean   : 0.00
##   3rd Qu.: 1.228   3rd Qu.:40.23   3rd Qu.:21.06   3rd Qu.: 0.68
##   Max.   :11.591   Max.   :57.09   Max.   :29.38   Max.   : 4.68
##   NA's   :8838     NA's   :81584   NA's   :81584   NA's   :81584
```

```
pre_math_data = filter(data_Q2, data_Q2$pre == 1 & data_Q2$test_type == 0)


pre_lang_data = filter(data_Q2, data_Q2$pre == 1 & data_Q2$test_type == 1)
```

running t test on the filtered data sets for Math

```
t.test(pre_math_data$norm~pre_math_data$bal)
```

```
##
##  Welch Two Sample t-test
##
## data:  pre_math_data$norm by pre_math_data$bal
## t = 0.34151, df = 10159, p-value = 0.7327
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.03235140  0.04600232
## sample estimates:
## mean in group 0 mean in group 1
##   -6.854839e-09   -6.825465e-03
```

Average norm value for both treatment and control group are more or less similarly divided given norm value ranges between -1.9 and 3.1 and both control and treatment has average norm value of -6.854839e-09 and -6.825465e-03 respectively implying randomization has been done correctly. The p-value for this test is 0.7327, which is well above our threshold of 0.05.We hence have enough evidence to say that the treatment and control groups have similar average normalized maths test scores for pre period.

```
t.test(pre_lang_data$norm~pre_lang_data$bal)
```

```
##
##  Welch Two Sample t-test
##
## data:  pre_lang_data$norm by pre_lang_data$bal
## t = -1.2176, df = 10140, p-value = 0.2234
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.06397876  0.01495078
## sample estimates:
```

```
## mean in group 0 mean in group 1
##   -1.313364e-08    2.451397e-02
```

Average norm value for both treatment and control group are more or less similarly divided given norm value ranges between -1.614 and 3.962 and both control and treatment has average norm value of -1.313364e-08 and 2.451397e-02 respectively .The p-value for this test is 0.2234, which is above our threshold of 0.05. We hence have enough evidence to say that the treatment and control groups have similar average normalized language test scores for pre period.

b) Calculate the average test scores for the post period (post = 1) for math for treatment and control. Is there a statistical difference between students in the two groups of schools? Use a t-test model to test the increase. Perform the same analysis for language test scores.

filtering the data into post period for Math and Language seperately

```
post_math_data = filter(data_Q2, data_Q2$post == 1 & data_Q2$test_type == 0)

post_lang_data = filter(data_Q2, data_Q2$post == 1 & data_Q2$test_type == 1)
```

running t test on the filtered data sets for Math

```
t.test(post_math_data$test~post_math_data$bal)
```

```
##
##   Welch Two Sample t-test
##
## data:  post_math_data$test by post_math_data$bal
## t = -5.807, df = 8391.7, p-value = 6.591e-09
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -2.257751 -1.118161
## sample estimates:
## mean in group 0 mean in group 1
##        19.78144        21.46939
```

Average test score value for treatment group after experiment is 21.4639 and for control it is 19.7814. the t test returned a p-value of 6.591e-09, which is lower than the threshold of 0.05. We can hence conclude that there us a difference in the mean test scores for maths between treatment and control groups #post Balsakhi Experiment

```
t.test(post_lang_data$test~post_lang_data$bal)
```

```
##
##   Welch Two Sample t-test
##
## data:  post_lang_data$test by post_lang_data$bal
## t = -3.773, df = 8407.6, p-value = 0.0001624
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -1.5450260 -0.4885151
## sample estimates:
## mean in group 0 mean in group 1
##        21.09880        22.11557
```

Average test score value for treatment group after experiment is 22.11557 and for control it is 21.09880. the t test returned a p-value of 0.0001624, which is lower than the threshold of 0.05. We can hence conclude that there is a difference in the mean test scores for language between treatment and control groups post Balsakhi Experiment

#(c)Can you conclude if the Balsakhi program increase test scores in reading and mathematics?

Running a linear model for maths to check if we can conclude that Balsakhi program increased the test scores post experiment .

```
summary(lm(test ~ bal, data = post_math_data))
```

```
##
## Call:
## lm(formula = test ~ bal, data = post_math_data)
##
## Residuals:
##      Min      1Q  Median      3Q     Max
## -21.4694 -11.4694  -0.4694  10.5306  30.2186
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  19.7814     0.2024  97.753  < 2e-16 ***
## bal           1.6880     0.2907   5.807 6.58e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.33 on 8424 degrees of freedom
##   (1772 observations deleted due to missingness)
## Multiple R-squared:  0.003987,   Adjusted R-squared:  0.003869
## F-statistic: 33.72 on 1 and 8424 DF,  p-value: 6.585e-09
```

The value of Coefficient for bal in the above linear model is 1.6880 indicating an increase in the test scores for math for the treatment group compared to control group. the P-value from this test is 6.585e-09, which is lower than our threshold of 0.05. Hence we can conclude that the Balsakhi program increases the test scores for Mathematics

Running a linear model for language to check if we can conclude that Balsakhi program increased the test scores post experiment

```
summary(lm(test ~ bal, data = post_lang_data))
```

```
##
## Call:
## lm(formula = test ~ bal, data = post_lang_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -22.116 -10.116  -2.099  10.884  28.901
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  21.0988     0.1877 112.399  < 2e-16 ***
## bal           1.0168     0.2696   3.771 0.000164 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.37 on 8424 degrees of freedom
##   (1772 observations deleted due to missingness)
## Multiple R-squared:  0.001685,   Adjusted R-squared:  0.001567
## F-statistic: 14.22 on 1 and 8424 DF,  p-value: 0.0001637
```

The value of Coefficient for bal in the above linear model is 1.0168 indicating an increase in the test scores for language for the treatment group compared to control group. the P-value from this test is 0.0001637, which is lower than our threshold of 0.05. Hence we can conclude that the Balsakhi program increases the test scores for languages

#(d)Is the SUTVA assumption likely to be violated in the experiment?

SUTVA assumptions state that: 1. Treatment can not affect behavior of control group. 2. External factors do not affect the behavior of treatment or control group. In this experiment, it can be violated if in case students in the treatment group interacts with students from the control group, i.e maybe tutoring them Additionally, the test score of students in both the treatment and control groups can be impacted due to external factors, such as intervention from parents and teachers or students cheating in the exams