Glasses For Letter Grade Impression

BA830 - Fall 2019

team 7

12/13/2019

Team 7 Members: Kunpeng Huang, Yoki Liu, Lyufan Pan, Yunlei Zhou, Jiayuan Zou, Sherry Zuo

Key Elements of Project Writeup When you write up your project paper, there is no predefined format. However, these factors must be included: a. Your causal question of interest.
b. Your experimental design and what data you collected.
c. Why you chose the experiment design.
d. Summary statistics about the data.
e. Treatment effects.
f. Limitations of your study.
Make sure the paper is well-written and that there are figures and tables that are easy to understand.

```
library(data.table)
library(tidyverse)
```

```
## -- Attaching packages ----------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts -------------------- tidyverse_conflicts() --
## x dplyr::between()   masks data.table::between()
## x dplyr::filter()    masks stats::filter()
## x dplyr::first()     masks data.table::first()
## x dplyr::lag()       masks stats::lag()
## x dplyr::last()      masks data.table::last()
## x purrr::transpose() masks data.table::transpose()
```

```
library(lfe)
```

```
## Loading required package: Matrix

##
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:data.table':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday,
##     week, yday, year
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
library(stargazer)
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(naniar)
library(ggmap)
```

```
## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.
```

```
## Please cite ggmap if you use it! See citation("ggmap") for details.
```

```
smart <- fread("survey.csv")
```

```
#skimr::skim(smart)
#View(smart)
dim(smart)
```

```
## [1] 228  43
```

```
head(smart, 2)
```

```
##                                                    StartDate
## 1:                                                Start Date
## 2: {""ImportId"":""startDate"",""timeZone"":""America/Denver""}
##                                                      EndDate
## 1:                                                  End Date
## 2: {""ImportId"":""endDate"",""timeZone"":""America/Denver""}
##                      Status                    IPAddress
## 1:             Response Type                   IP Address
## 2: {""ImportId"":""status""} {""ImportId"":""ipAddress""}
##                      Progress       Duration (in seconds)
## 1:                   Progress       Duration (in seconds)
## 2: {""ImportId"":""progress""} {""ImportId"":""duration""}
##                      Finished
## 1:                   Finished
## 2: {""ImportId"":""finished""}
##                                                  RecordedDate
## 1:                                               Recorded Date
## 2: {""ImportId"":""recordedDate"",""timeZone"":""America/Denver""}
##                     ResponseId                  RecipientLastName
## 1:                  Response ID                 Recipient Last Name
## 2: {""ImportId"":""_recordId""} {""ImportId"":""recipientLastName""}
##                      RecipientFirstName                    RecipientEmail
## 1:                  Recipient First Name                   Recipient Email
## 2: {""ImportId"":""recipientFirstName""} {""ImportId"":""recipientEmail""}
##                          ExternalReference
## 1:                  External Data Reference
## 2: {""ImportId"":""externalDataReference""}
##                      LocationLatitude
## 1:                   Location Latitude
## 2: {""ImportId"":""locationLatitude""}
##                      LocationLongitude
## 1:                   Location Longitude
## 2: {""ImportId"":""locationLongitude""}
##                      DistributionChannel                    UserLanguage
## 1:                   Distribution Channel                   User Language
## 2: {""ImportId"":""distributionChannel""} {""ImportId"":""userLanguage""}
##                        Q2.1_First Click
## 1:                 Timing - First Click
## 2: {""ImportId"":""QID11_FIRST_CLICK""}
##                        Q2.1_Last Click
## 1:                  Timing - Last Click
## 2: {""ImportId"":""QID11_LAST_CLICK""}
##                        Q2.1_Page Submit
## 1:                 Timing - Page Submit
## 2: {""ImportId"":""QID11_PAGE_SUBMIT""}
##                        Q2.1_Click Count
## 1:                 Timing - Click Count
## 2: {""ImportId"":""QID11_CLICK_COUNT""}
##                                    Q2.3
## 1: Please guess the year of this person.
## 2:              {""ImportId"":""QID12""}
```

```
##                                     Q2.4
## 1: Please guess this person's personality.
## 2:                {""ImportId"":""QID13""}
##                                           Q2.5
## 1: Please guess the letter grade of this person.
## 2:                       {""ImportId"":""QID14""}
##                         Q3.1_First Click
## 1:                  Timing - First Click
## 2: {""ImportId"":""QID41_FIRST_CLICK""}
##                         Q3.1_Last Click
## 1:                  Timing - Last Click
## 2: {""ImportId"":""QID41_LAST_CLICK""}
##                         Q3.1_Page Submit
## 1:                  Timing - Page Submit
## 2: {""ImportId"":""QID41_PAGE_SUBMIT""}
##                         Q3.1_Click Count
## 1:                  Timing - Click Count
## 2: {""ImportId"":""QID41_CLICK_COUNT""}
##                                     Q3.3
## 1: Please guess the year of this person.
## 2:              {""ImportId"":""QID43""}
##                                      Q3.4
## 1: Please guess this person's personality.
## 2:                {""ImportId"":""QID44""}
##                                           Q3.5
## 1: Please guess the letter grade of this person.
## 2:                       {""ImportId"":""QID45""}
##                         Q4.1_First Click
## 1:                  Timing - First Click
## 2: {""ImportId"":""QID36_FIRST_CLICK""}
##                         Q4.1_Last Click
## 1:                  Timing - Last Click
## 2: {""ImportId"":""QID36_LAST_CLICK""}
##                         Q4.1_Page Submit
## 1:                  Timing - Page Submit
## 2: {""ImportId"":""QID36_PAGE_SUBMIT""}
##                         Q4.1_Click Count
## 1:                  Timing - Click Count
## 2: {""ImportId"":""QID36_CLICK_COUNT""}
##                                     Q4.3
## 1: Please guess the year of this person.
## 2:              {""ImportId"":""QID38""}
##                                      Q4.4
## 1: Please guess this person's personality.
## 2:                {""ImportId"":""QID39""}
##                                           Q4.5                       Q5.1
## 1: Please guess the letter grade of this person.     What is your gender?
## 2:                       {""ImportId"":""QID40""} {""ImportId"":""QID46""}
##                         Q5.2                     Q5.3
## 1:      What year are your? What's your personality?
## 2: {""ImportId"":""QID33""} {""ImportId"":""QID34""}
##                          Q5.4
## 1: What's your letter grade?
## 2:  {""ImportId"":""QID35""}
```

```
##                                                       Q48
## 1: Can you guess the purpose of our experiment? (optional)
## 2:                             {""ImportId"":""QID48_TEXT""}
```

```
smart=smart%>%select(-Status, -DistributionChannel, -RecipientEmail, -RecipientFirstName, -RecipientLas
```

```
#only keep finished records
smart=smart%>%filter(Finished=="True")%>%filter(Progress==100)
```

```
smart=smart%>%select(-Progress, -Finished)
```

```
length(unique(smart$ResponseId))
```

```
## [1] 226
```

```
length(unique(smart$IPAddress))
```

```
## [1] 211
```

```
#some people take more than once, we only want to keep each person once.
smart=smart%>%distinct(IPAddress, .keep_all = TRUE)
```

```
# Add treatment/control dummy variable:
smart2 = smart %>%
  mutate(Treatment = ifelse(`Q4.1_First Click`=="", FALSE, TRUE))

# Variable selection:
smart_cleaned = smart2 %>%
  select(-c(`Q2.1_First Click`:`Q2.1_Click Count`, `Q3.1_First Click`:`Q3.1_Click Count`,
            `Q4.1_First Click`:`Q4.1_Click Count`, Q48, IPAddress, StartDate, EndDate, ResponseId)) %>%
  select(1:8,
         class_baseline=Q2.3, personality_baseline=Q2.4, grade_baseline=Q2.5,
         gender_p=Q5.1, class_p=Q5.2, personality_p=Q5.3, grade_p=Q5.4,
         everything())
```

```
control = smart_cleaned %>%
  filter(Treatment == F) %>%
  select(-c(Q4.3, Q4.4, Q4.5), class=Q3.3, personality=Q3.4, grade=Q3.5)

treatment = smart_cleaned %>%
  filter(Treatment == T) %>%
  select(-c(Q3.3, Q3.4, Q3.5), class = Q4.3, personality=Q4.4, grade=Q4.5)

data = rbind(control, treatment)
#write_csv(data, "data.csv")
```

```
data2<-read.csv("data.csv", na.strings = c("", "NA"))
```

```r
data2 = na.omit(data2)
```

```r
#write_csv(data2, "cleaned_data.csv")
```

```r
dim(data2)
```

```
## [1] 205  15
```

```r
# Convert GPA
data2$grade = ifelse(data2$grade == "A(3.75-4.0)", 9,
                     ifelse(data2$grade == "A-(3.5-3.75)", 8,
                            ifelse(data2$grade == "B+(3.25-3.5)", 7,
                                   ifelse(data2$grade == "B(3.0-3.25)", 6,
                                          ifelse(data2$grade == "B-(2.75-3.0)", 5,
                                                 ifelse(data2$grade == "C+(2.5-2.75)", 4,
                                                        ifelse(data2$grade == "C(2.25-2.5)", 3,
                                                               ifelse(data2$grade == "C-(2.0-2.25)", 2,
                                                                      ifelse(data2$grade == "Prefer not
                                                               ))))))))
table(data2$grade)
```

```
##
##  2  3  4  5  6  7  8  9
##  1  1  3  7 18 40 76 59
```

```r
# Convert GPA_P
data2$grade_p = ifelse(data2$grade_p == "A(3.75-4.0)", 9,
                     ifelse(data2$grade_p == "A-(3.5-3.75)", 8,
                            ifelse(data2$grade_p == "B+(3.25-3.5)", 7,
                                   ifelse(data2$grade_p == "B(3.0-3.25)", 6,
                                          ifelse(data2$grade_p == "B-(2.75-3.0)", 5,
                                                 ifelse(data2$grade_p == "C+(2.5-2.75)", 4,
                                                        ifelse(data2$grade_p == "C(2.25-2.5)", 3,
                                                               ifelse(data2$grade_p == "C-(2.0-2.25)", 1
                                                                      ifelse(data2$grade_p == "Prefer no
                                                                             1)))))))))
table(data2$grade_p)
```

```
##
##  3  4  5  6  7  8  9
##  2  3  5 30 33 79 42
```

```r
# Convert GPA_baseline
data2$grade_baseline = ifelse(data2$grade_baseline == "A(3.75-4.0)", 9,
                     ifelse(data2$grade_baseline == "A-(3.5-3.75)", 8,
                            ifelse(data2$grade_baseline == "B+(3.25-3.5)", 7,
                                   ifelse(data2$grade_baseline == "B(3.0-3.25)", 6,
                                          ifelse(data2$grade_baseline == "B-(2.75-3.0)", 5,
                                                 ifelse(data2$grade_baseline == "C+(2.5-2.75)", 4,
                                                        ifelse(data2$grade_baseline == "C(2.25-2.5)", 3
                                                               ifelse(data2$grade_baseline == "C-(2.0-2
```

```
                                                                    ifelse(data2$grade_baseline == "P
                                                                           1)))))))))
table(data2$grade_baseline)
```

```
##
##  1  4  5  6  7  8  9
##  1  1  2 19 55 96 31
```

```
data2$gender_p = ifelse(data2$gender_p == "Male", 1,
                 ifelse(data2$gender_p == "Female", 0, NA))

table(data2$gender_p)
```

```
##
##   0   1
## 130  67
```

```
data2$personality_baseline = ifelse(data2$personality_baseline == "Logical", 1,
                             ifelse(data2$personality_baseline == "Emotional", 0, NA))

table(data2$personality_baseline)
```

```
##
##   0   1
##  79 126
```

```
data2$personality_p = ifelse(data2$personality_p == "Logical", 1,
                             ifelse(data2$personality_p == "Emotional", 0, NA))

table(data2$personality_p)
```

```
##
##  0  1
## 69 84
```

```
data2$personality = ifelse(data2$personality == "Logical", 1,
                           ifelse(data2$personality == "Emotional", 0, NA))

table(data2$personality)
```

```
##
##   0   1
##  53 152
```

```
#save final cleaned version of data
#write_csv(data2, "data2.csv")
```

```
analysis<-fread("data2.csv")
```

```
summary(analysis)
```

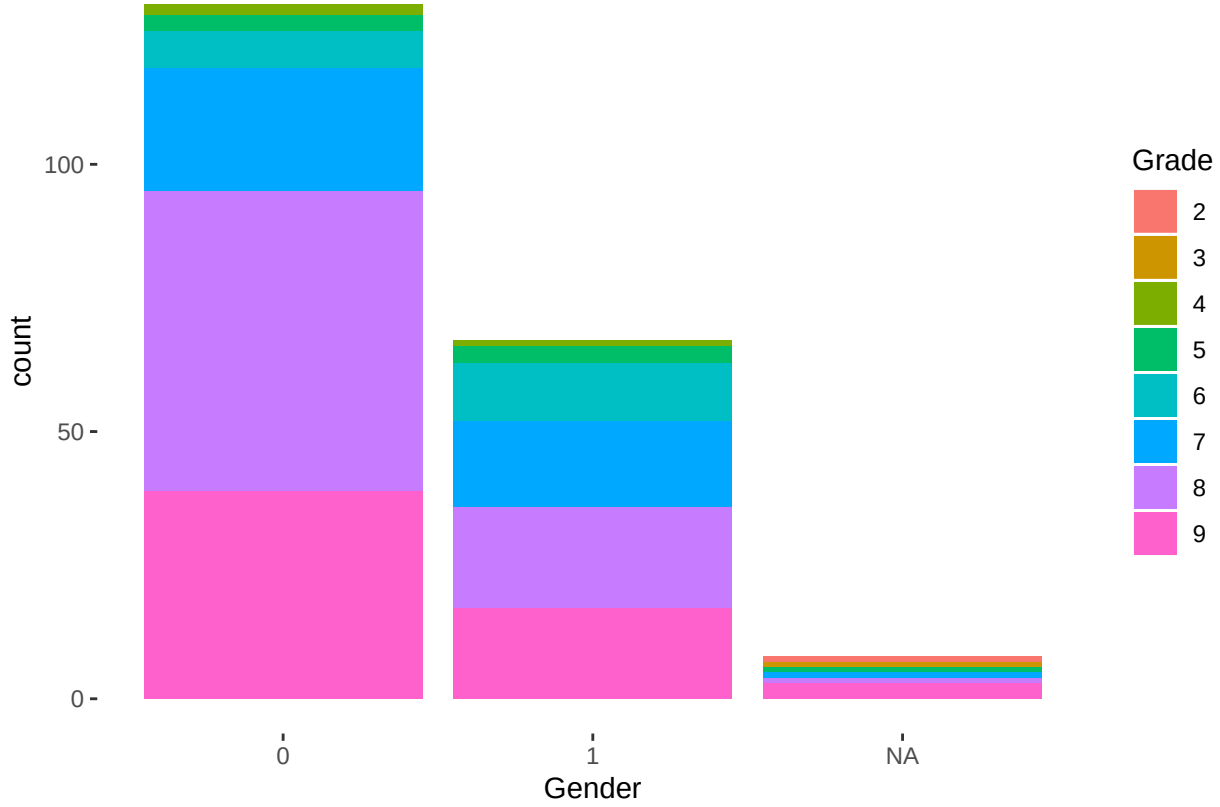
```
##  Duration..in.seconds. RecordedDate       LocationLatitude
##  Min.   :   35.00      Length:205         Min.   :-27.473
##  1st Qu.:   62.00      Class :character   1st Qu.: 34.065
##  Median :   82.00      Mode  :character   Median : 42.037
##  Mean   :  312.92                         Mean   : 38.942
##  3rd Qu.:  116.00                         3rd Qu.: 42.347
##  Max.   :17896.00                         Max.   : 53.796
##
##  LocationLongitude   class_baseline     personality_baseline
##  Min.   :-158.0419   Length:205         Min.   :0.00000
##  1st Qu.: -83.3713   Class :character   1st Qu.:0.00000
##  Median : -71.0975   Mode  :character   Median :1.00000
##  Mean   : -27.1948                      Mean   :0.61463
##  3rd Qu.:   2.3387                      3rd Qu.:1.00000
##  Max.   : 153.0215                      Max.   :1.00000
##
##  grade_baseline      class              gender_p        class_p
##  Min.   :1.0000   Length:205         Min.   :0.0000   Length:205
##  1st Qu.:7.0000   Class :character   1st Qu.:0.0000   Class :character
##  Median :8.0000   Mode  :character   Median :0.0000   Mode  :character
##  Mean   :7.6146                      Mean   :0.3401
##  3rd Qu.:8.0000                      3rd Qu.:1.0000
##  Max.   :9.0000                      Max.   :1.0000
##                                      NA's   :8
##  personality_p        grade_p        personality          grade
##  Min.   :0.00000   Min.   :3.0000   Min.   :0.00000   Min.   :2.0000
##  1st Qu.:0.00000   1st Qu.:7.0000   1st Qu.:0.00000   1st Qu.:7.0000
##  Median :1.00000   Median :8.0000   Median :1.00000   Median :8.0000
##  Mean   :0.54902   Mean   :7.5464   Mean   :0.74146   Mean   :7.7024
##  3rd Qu.:1.00000   3rd Qu.:8.0000   3rd Qu.:1.00000   3rd Qu.:9.0000
##  Max.   :1.00000   Max.   :9.0000   Max.   :1.00000   Max.   :9.0000
##  NA's   :52        NA's   :11
##  Treatment
##  Mode :logical
##  FALSE:102
##  TRUE :103
##
##
##
##
```

```
names(analysis)
```

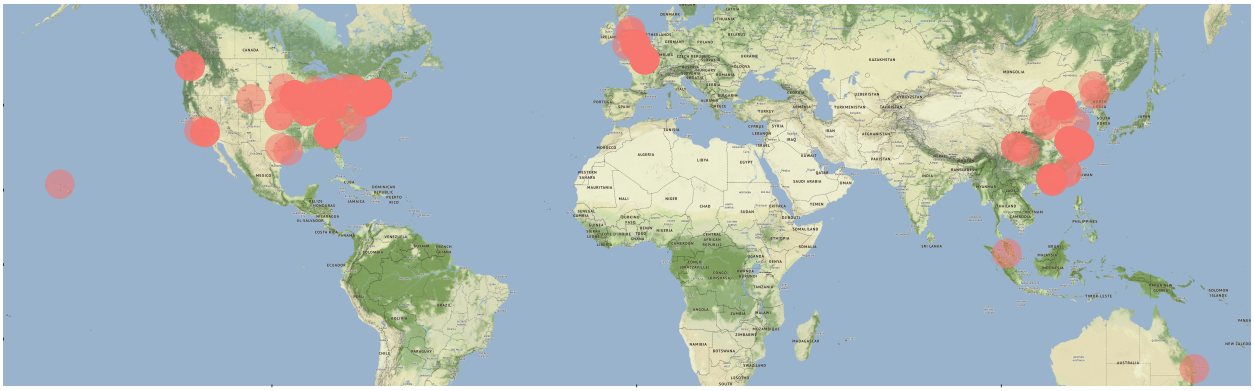
```
##  [1] "Duration..in.seconds." "RecordedDate"
##  [3] "LocationLatitude"      "LocationLongitude"
##  [5] "class_baseline"        "personality_baseline"
##  [7] "grade_baseline"        "class"
##  [9] "gender_p"              "class_p"
```

```
## [11] "personality_p"          "grade_p"
## [13] "personality"            "grade"
## [15] "Treatment"
```

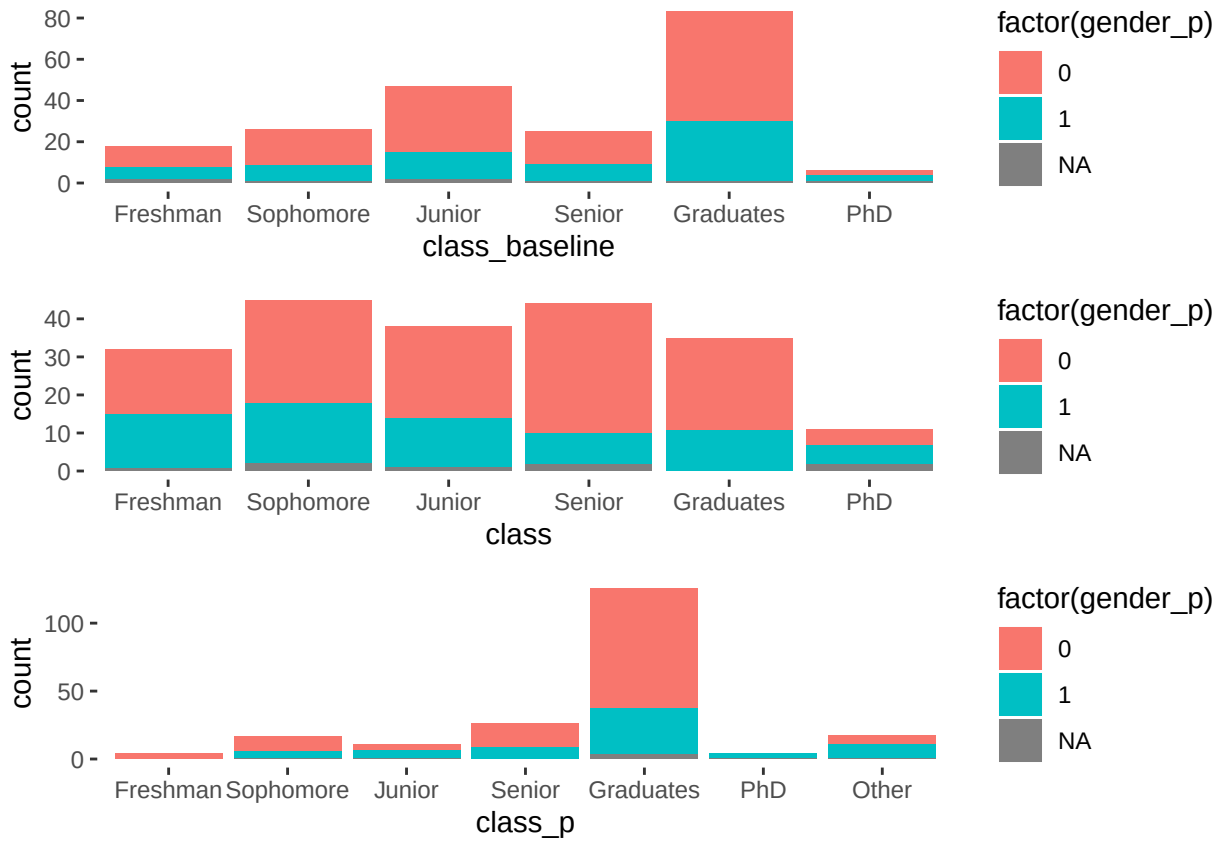
```
ggplot(data=analysis, aes(x=factor(gender_p), fill=factor(grade)))+
  geom_bar()+
  labs(x="Gender", fill="Grade")+
  theme(panel.background = element_rect(fill="white"))
```



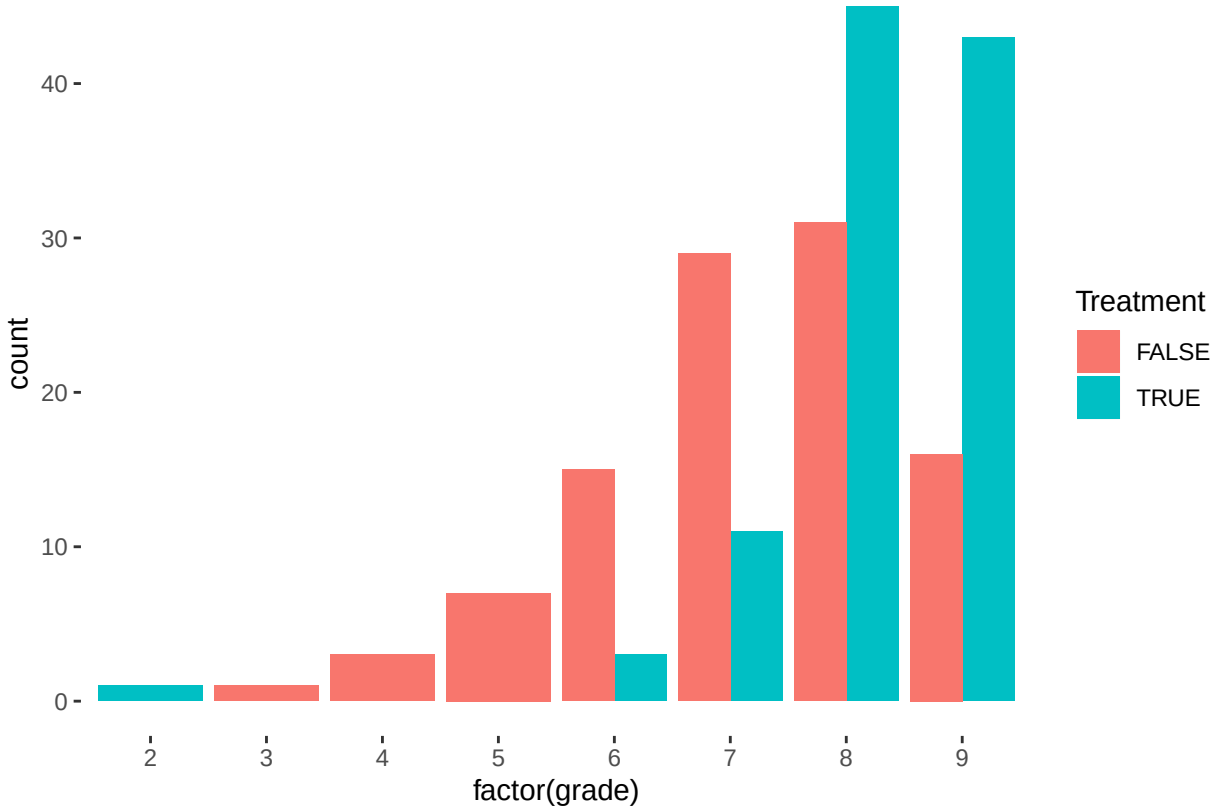

```
qmplot(x = LocationLongitude, y = LocationLatitude, data = analysis, colour = 'blue', alpha=0.4,  size =
  theme(legend.position="")
```



```
analysis$class_baseline<-factor(analysis$class_baseline, levels=c("Freshman", "Sophomore", "Junior", "Se
p1<-ggplot(analysis, aes(x=class_baseline, fill=factor(gender_p)))+
  geom_bar()+
  theme(panel.background = element_rect(fill="white"))
analysis$class<-factor(analysis$class, levels=c("Freshman", "Sophomore", "Junior", "Senior", "Graduates
p2<-ggplot(analysis, aes(x=class, fill=factor(gender_p)))+
  geom_bar()+
  theme(panel.background = element_rect(fill="white"))
analysis$class_p<-factor(analysis$class_p, levels=c("Freshman", "Sophomore", "Junior", "Senior", "Gradua
p3<-ggplot(analysis, aes(x=class_p, fill=factor(gender_p)))+
  geom_bar()+
  theme(panel.background = element_rect(fill="white"))
grid.arrange(p1,p2,p3, nrow=3, ncol=1)
```




```
ggplot(analysis, aes(x=factor(grade), fill=Treatment))+
  geom_bar(position = "dodge")+
  theme(panel.background = element_rect(fill="white"))
```



##Randomization Check/Balance Check

```
summary(lm(grade_baseline~Treatment, analysis))
```

```
##
## Call:
## lm(formula = grade_baseline ~ Treatment, data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.62136 -0.62136  0.37864  0.39216  1.39216
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.607843   0.101792 74.7392   <2e-16 ***
## TreatmentTRUE 0.013516   0.143606  0.0941   0.9251
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.028 on 203 degrees of freedom
## Multiple R-squared:  4.3636e-05, Adjusted R-squared:  -0.0048823
## F-statistic: 0.0088585 on 1 and 203 DF,  p-value: 0.92511
```

Since the p-value is 0.9251 which is larger than 0.05, the R-square is -0.0048823 which is too small, so we fail to reject the null hypothesis, thus we think the ATE is 0, so we have 95% confidence that the glasses treatment on grade taken before the experiment started does not have effect. So this results indicate that the validity of the study's conclusion is high.

##Causal Effect

```
summary(lm(grade~Treatment, data=analysis))
```

```
##
## Call:
## lm(formula = grade ~ Treatment, data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.19417 -0.20588 -0.19417  0.80583  1.79412
##
## Coefficients:
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept)    7.20588    0.11520  62.551 < 2.2e-16 ***
## TreatmentTRUE  0.98829    0.16252   6.081 5.838e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.1635 on 203 degrees of freedom
## Multiple R-squared:  0.15409,    Adjusted R-squared:  0.14992
## F-statistic: 36.978 on 1 and 203 DF,  p-value: 5.8378e-09
```

```
lowerci<-0.98829-1.96*0.16252
lowerci
```

```
## [1] 0.6697508
```

```
upperci<-0.98829+1.96*0.16252
upperci
```

```
## [1] 1.3068292
```

The ATE of treatment on grade is **0.98829**, the 95% confidence interval is [**0.6697508, 1.3068292**].

##Fixed Effect Check ###Grade_Baseline

```
summary(felm(grade~grade_baseline, data=analysis))
```

```
##
## Call:
##    felm(formula = grade ~ grade_baseline, data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.31903 -0.87395  0.12605  0.68097  2.46126
##
## Coefficients:
##                Estimate Std. Error t value  Pr(>|t|)
## (Intercept)    4.313387   0.618596  6.9729 4.278e-11 ***
## grade_baseline 0.445071   0.080514  5.5278 9.889e-08 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.1794 on 203 degrees of freedom
## Multiple R-squared(full model): 0.13083   Adjusted R-squared: 0.12655
## Multiple R-squared(proj model): 0.13083   Adjusted R-squared: 0.12655
## F-statistic(full model):30.557 on 1 and 203 DF, p-value: 9.8894e-08
## F-statistic(proj model): 30.557 on 1 and 203 DF, p-value: 9.8894e-08
```

Since the p-value is less than 0.05, so we reject the null hypothesis, so we have 95% confidence that the baseline has effect for the grade, that's what we want to do baseline during our survey. Thus, it's a fixed effect. ###Personal Gender

```
with(analysis, table(gender_p, grade))
```

```
##         grade
## gender_p  2  3  4  5  6  7  8  9
##        0  0  0  2  3  7 23 56 39
##        1  0  0  1  3 11 16 19 17
```

```
male_grade<-analysis[gender_p==1, grade]
table(male_grade)
```

```
## male_grade
##  4  5  6  7  8  9
##  1  3 11 16 19 17
```

```
female_grade<-analysis[gender_p==0, grade]
table(female_grade)
```

```
## female_grade
##  4  5  6  7  8  9
##  2  3  7 23 56 39
```

```
t.test(male_grade, female_grade)
```

```
##
##  Welch Two Sample t-test
##
## data:  male_grade and female_grade
## t = -2.1912, df = 116.782, p-value = 0.030423
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.746452065 -0.037704077
## sample estimates:
## mean of x mean of y
## 7.4925373 7.8846154
```

Since the p-value is 0.030423 which smaller than 0.05, so we reject the null hypothesis, so we have 95% confidence that the true difference in means is not equal to 0, so which means it's not 50/50 split by gender in each grade level, so the experiment didn't achieve its desired randomizaion, so we need to consider gender as fixed effect.
###Personal Personality

```r
with(analysis, table(personality_p, grade))
```

```
##              grade
## personality_p  2  3  4  5  6  7  8  9
##             0  0  0  1  2  5 14 21 26
##             1  1  1  1  3  9 20 30 19
```

```r
logical_grade<-analysis[personality_p==1, grade]
table(logical_grade)
```

```
## logical_grade
##  2  3  4  5  6  7  8  9
##  1  1  1  3  9 20 30 19
```

```r
emotional_grade<-analysis[personality_p==0, grade]
table(emotional_grade)
```

```
## emotional_grade
##  4  5  6  7  8  9
##  1  2  5 14 21 26
```

```r
t.test(logical_grade, emotional_grade)
```

```
##
##  Welch Two Sample t-test
##
## data:  logical_grade and emotional_grade
## t = -1.92402, df = 150.802, p-value = 0.056235
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.802585092  0.010659626
## sample estimates:
## mean of x mean of y
## 7.4880952 7.8840580
```

Since the p-value is 0.056235 which larger than 0.05, so we fail to reject the null hypothesis, so we have 95% confidence that the true difference in means is not equal to 0, so which means it's 50/50 split by personality in each grade level, so the experiment achieve its desired randomizaion, so we do need to consider personality as fixed effect. ###Class Since most of our survey participants are around Graduates, but their classes guessing of pictures are randomly, so we do not think it's a fixed effect here.

```r
summary(felm(grade~Treatment | grade_baseline + gender_p, data=analysis))
```

```
##
## Call:
##    felm(formula = grade ~ Treatment | grade_baseline + gender_p,      data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.86147 -0.57468  0.13280  0.55207  1.85032
```

```
##
## Coefficients:
##               Estimate Std. Error t value  Pr(>|t|)
## TreatmentTRUE  1.00573    0.13825   7.275 8.977e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.95269 on 189 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared(full model): 0.33205   Adjusted R-squared: 0.30731
## Multiple R-squared(proj model): 0.21877   Adjusted R-squared: 0.18983
## F-statistic(full model):13.422 on 7 and 189 DF, p-value: 5.028e-14
## F-statistic(proj model): 52.925 on 1 and 189 DF, p-value: 8.9765e-12
```

##Limitation

```
summary(felm(grade~Treatment+personality | grade_baseline + gender_p, data=analysis))
```

```
##
## Call:
##    felm(formula = grade ~ Treatment + personality | grade_baseline +      gender_p, data = analysis)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -2.951298 -0.672204 -0.009534  0.578417  1.757184
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## TreatmentTRUE  0.66267    0.12955  5.1153 7.69e-07 ***
## personality    1.11184    0.14692  7.5676 1.65e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8363 on 188 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared(full model): 0.48801   Adjusted R-squared: 0.46622
## Multiple R-squared(proj model): 0.40118   Adjusted R-squared: 0.3757
## F-statistic(full model):22.399 on 8 and 188 DF, p-value: < 2.22e-16
## F-statistic(proj model): 62.975 on 2 and 188 DF, p-value: < 2.22e-16
```

```
summary(felm(grade~Treatment+class | grade_baseline + gender_p, data=analysis))
```

```
##
## Call:
##    felm(formula = grade ~ Treatment + class | grade_baseline + gender_p,      data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.88656 -0.55989  0.15496  0.58122  1.92812
##
## Coefficients:
##               Estimate Std. Error t value  Pr(>|t|)
```

```
## TreatmentTRUE   0.87703    0.13716  6.3943 1.293e-09 ***
## classSophomore  0.51845    0.22005  2.3561  0.019521 *
## classJunior     0.59990    0.22725  2.6399  0.009006 **
## classSenior     0.54328    0.22590  2.4049  0.017167 *
## classGraduates  0.96262    0.23305  4.1305 5.487e-05 ***
## classPhD        0.98737    0.35738  2.7628  0.006312 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9185 on 184 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared(full model): 0.39555   Adjusted R-squared: 0.35613
## Multiple R-squared(proj model): 0.29304   Adjusted R-squared: 0.24693
## F-statistic(full model):10.034 on 12 and 184 DF, p-value: 5.213e-15
## F-statistic(proj model): 12.711 on 6 and 184 DF, p-value: 5.5137e-12
```

```
summary(felm(grade~Treatment*personality | grade_baseline + gender_p, data=analysis))
```

```
##
## Call:
##    felm(formula = grade ~ Treatment * personality | grade_baseline +      gender_p, data = analysis)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -3.008968 -0.632629  0.052621  0.596440  1.889850
##
## Coefficients:
##                            Estimate Std. Error t value  Pr(>|t|)
## TreatmentTRUE               1.15764    0.28782  4.0221 8.363e-05 ***
## personality                 1.29341    0.17379  7.4425 3.495e-12 ***
## TreatmentTRUE:personality -0.61382    0.31930 -1.9224   0.05608 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.83036 on 187 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared(full model): 0.49793   Adjusted R-squared: 0.47377
## Multiple R-squared(proj model): 0.41278   Adjusted R-squared: 0.38452
## F-statistic(full model):20.606 on 9 and 187 DF, p-value: < 2.22e-16
## F-statistic(proj model): 43.817 on 3 and 187 DF, p-value: < 2.22e-16
```

```
summary(felm(grade~Treatment*class| grade_baseline + gender_p, data=analysis))
```

```
##
## Call:
##    felm(formula = grade ~ Treatment * class | grade_baseline + gender_p,      data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.83331 -0.54880  0.11490  0.53303  1.82092
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
```

```
## TreatmentTRUE                 0.566200   0.361649  1.5656 0.119208
## classSophomore                0.364933   0.293266  1.2444 0.214989
## classJunior                   0.366625   0.297905  1.2307 0.220057
## classSenior                   0.351909   0.287017  1.2261 0.221775
## classGraduates                1.024126   0.347598  2.9463 0.003644 **
## classPhD                      1.304569   0.611522  2.1333 0.034261 *
## TreatmentTRUE:classSophomore  0.388238   0.458820  0.8462 0.398590
## TreatmentTRUE:classJunior     0.566018   0.474129  1.1938 0.234134
## TreatmentTRUE:classSenior     0.498612   0.463692  1.0753 0.283684
## TreatmentTRUE:classGraduates  0.062503   0.498265  0.1254 0.900316
## TreatmentTRUE:classPhD       -0.281937   0.784445 -0.3594 0.719713
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9233 on 179 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared(full model): 0.40581   Adjusted R-squared: 0.34938
## Multiple R-squared(proj model): 0.30504   Adjusted R-squared: 0.23904
## F-statistic(full model):7.1912 on 17 and 179 DF, p-value: 3.3542e-13
## F-statistic(proj model): 7.1426 on 11 and 179 DF, p-value: 5.1172e-10
```

```
summary(felm(grade~personality*class| grade_baseline + gender_p, data=analysis))
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite

## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite

##
## Call:
##    felm(formula = grade ~ personality * class | grade_baseline +      gender_p, data = analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.27503 -0.44463  0.19984  0.56847  1.68049
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## personality                  1.49500    0.31884  4.6889 5.41e-06 ***
## classSophomore               0.79088    0.35229  2.2450  0.02599 *
## classJunior                  0.66482    0.36401  1.8264  0.06945 .
## classSenior                  0.44114    0.36193  1.2189  0.22449
## classGraduates               1.07569    0.49846  2.1580  0.03225 *
## classPhD                     0.68842    0.36221  1.9006  0.05895 .
## personality:classSophomore -0.53948    0.44093 -1.2235  0.22274
## personality:classJunior    -0.30930    0.45960 -0.6730  0.50183
## personality:classSenior    -0.15549    0.44528 -0.3492  0.72734
## personality:classGraduates -0.38241    0.56946 -0.6715  0.50274
## personality:classPhD             NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.87116 on 180 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared(full model): 0.46807   Adjusted R-squared: 0.42079
## Multiple R-squared(proj model): 0.37786   Adjusted R-squared: 0.32256
## F-statistic(full model):9.8995 on 16 and 180 DF, p-value: < 2.22e-16
## F-statistic(proj model): 9.9387 on 11 and 180 DF, p-value: 5.2209e-14
```

We have one optional question in the end of our survey to ask whether people can detect our purpose or not, and they can take a guess.

```
guess <- read_csv("glasses.csv")
```

```
## Parsed with column specification:
## cols(
##   `Can you guess the purpose of our experiment? (optional)` = col_character()
## )
```

```
names(guess)
```

```
## [1] "Can you guess the purpose of our experiment? (optional)"
```

```
names(guess)<-"purpose_guess"
```