Final_Project_Markdown

Ashton Chevallier, Lisa Minas, Olivier Zimmer

August 12, 2017

Dataload

```r
library(dplyr)
library(ggplot2)
library(stargazer)
library(lmtest)
#stargazer(both, )
#Change Directory
raw_data <- read.csv('C:/Users/achevall/Documents/UC Berkeley/w241/Final/BeatuyAd_CausalExperiment/Data,
```

Exploratory Data Analysis

Filters

```r
#First Filter out junk data
filter_data <- function(raw_data) {
  raw_data$RecordedDate <- as.POSIXct(strptime(as.character(raw_data$RecordedDate),
                                               '%Y-%m-%d %H:%M:%S'))
  data <- raw_data %>%
    filter(Finished == 'True',
           Status == 'IP Address',
           Welcome == 'I agree',
           Group != '',
           AudioCheck == 'Pineapple',
           RecordedDate > '2017-07-16 01:00:00')
}

data <- filter_data(raw_data)
```

We created an audio check to test whether people understood English and were paying attention to the survey. We are using the audiocheck to filter out non-compliers. The other filters are to remove junk data: we don't want partial responses, junk IPS, or people that didn't agree with our terms. The date filter is to ensure we are using the correct experiment timing.

Summary Stats

```r
interesting_columns <- c('Race','Age','Gender','Location','Group')
summary(data[,interesting_columns])
```

```
##                                     Race          Age
##    White                            :276    25 - 34:197
##    Asian                            : 56    35 - 44: 70
```

```
## Black or African American          : 27   18 - 24: 56
## Other                              : 12   45 - 54: 31
## American Indian or Alaska Native   :  5   55 - 64: 25
## White,American Indian or Alaska Native:  3   65 - 74: 10
## (Other)                           : 11   (Other): 1
##                                 Gender          Location
##                                    :  0   California  : 64
## {"ImportId":"QID4"}                :  0   Florida     : 36
## Female                             :186   New York    : 34
## Male                               :203   Indiana     : 20
## Other                              :  1   Texas       : 20
## Which gender do you identify with?:  0   Pennsylvania: 18
##                                            (Other)     :198
##               Group
##                     :  0
## {"ImportId":"Group"}:  0
## Control           :195
## Group             :  0
## Treatment         :195
##
##
```

```r
percent_table <- function(column){
  print(head(sort(table(sort(column))/length(column), decreasing = TRUE)))
}

percent_table(data$Race)
```

```
##
##                                White
##                          0.707692308
##                                Asian
##                          0.143589744
##           Black or African American
##                          0.069230769
##                                Other
##                          0.030769231
##      American Indian or Alaska Native
##                          0.012820513
## White,American Indian or Alaska Native
##                          0.007692308
```

We got excellent an split between Treatment and Control, which gives us confidence our randomization
worked. The racial demographic split seems to be fairly representative of the US, but whites and asians seem
to be a bit over sampled. Our location is fairly well balanced but it seems like Texas is under-represented.
Most importantly our gender split is relatively even. We'd prefer more females than males (as the US is
slightly more female), especially because we think females will respond to treatment better. But overall, we
are happy with the sampling and don't see any glaring bias.

Exploratory Graphs

```r
qplot(data$Age, geom = 'bar', fill = 'Count', main = 'Bar plot of Age')
```
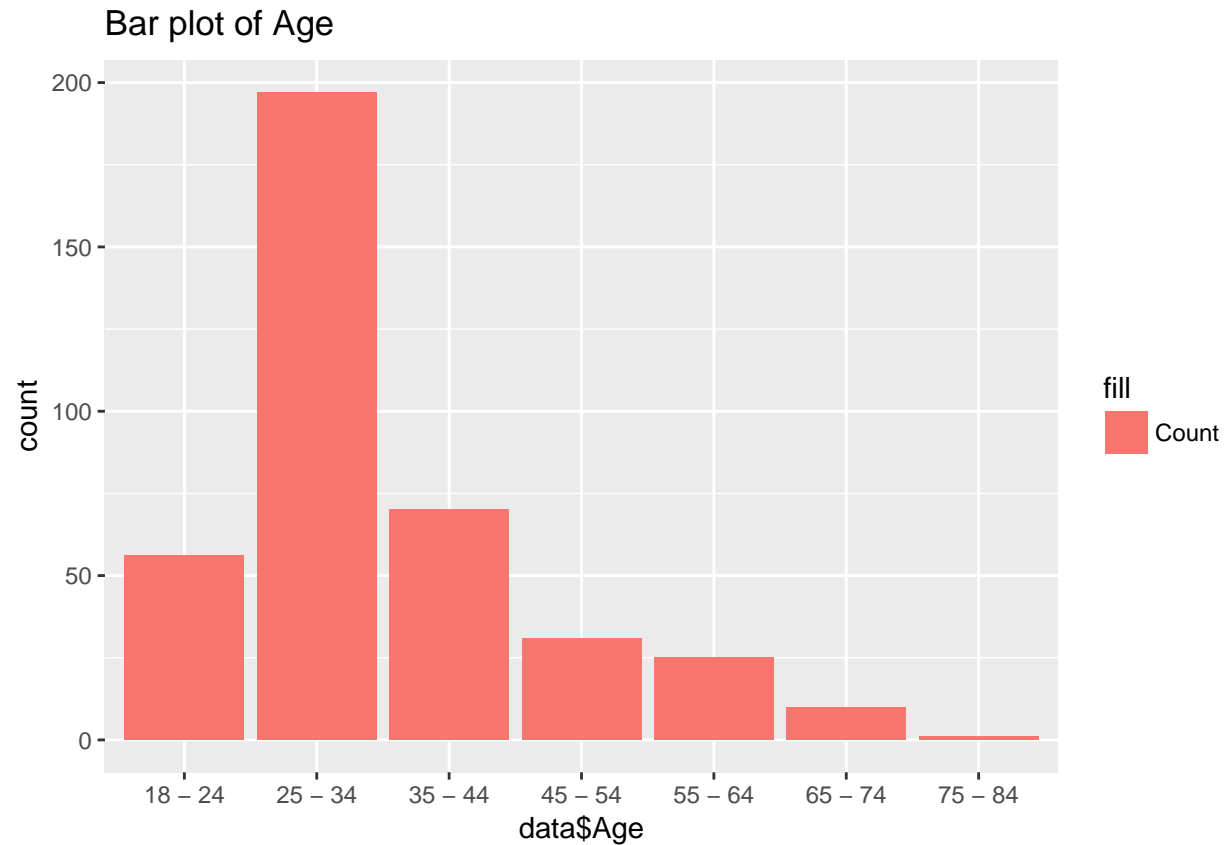
Bar plot of Age

Not a lot of old people on Mechanical Turk, this is expected because I doubt retirees are filling out surveys to pay the bills. But our sample is certainly biased towards the 'millenials'.

```
qplot(data$Location, geom = 'bar', fill = 'r', main = 'Bar Plot of Location') +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  coord_flip()
```

Bar Plot of Location

Nothing too glaringly terrible, but Mechanical Turk seems much more popular in California than other places.

```
#Rotate Race
qplot(data$Race, geom = 'bar', fill= 'Count', main = 'Bar Plot of Race') +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  coord_flip()
```

Bar Plot of Race
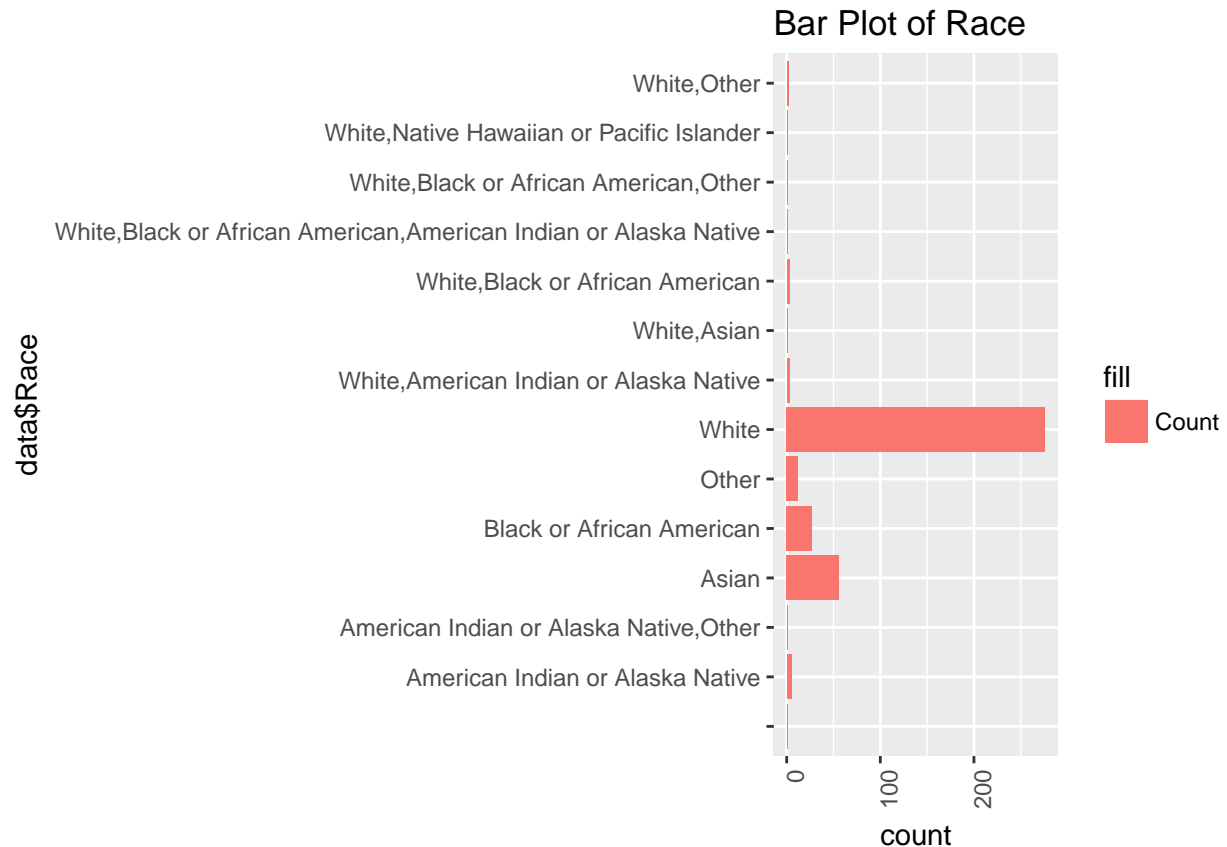
Again, our distribution is fairly representative of the US. Nothing too biased. Although, we might not have enough power make any real causal claims on race.

Variables of Interest

```
personal_views <- c('Personal_Views_Confident',
                    'Personal_Views_Beautiful',
                    'Personal_Views_Beauty_Importance',
                    'Personal_Views_Relate_To_Model')

summary(data[,personal_views])
```

```
##         Personal_Views_Confident          Personal_Views_Beautiful
##  Agree                  :193     Agree                  :199
##  Strongly Agree         :105     Disagree               : 98
##  Disagree               : 78     Strongly Agree         : 77
##  Strongly disagree      : 14     Strongly disagree      : 15
##                         :  0                            :  1
##  {"ImportId":"QID26_1"}:  0      {"ImportId":"QID26_2"}:  0
##  (Other)                :  0     (Other)                :  0
##       Personal_Views_Beauty_Importance
##  Agree                  :189
##  Disagree               :113
##  Strongly Agree         : 64
##  Strongly disagree      : 24
##                         :  0
```

5

```
## {"ImportId":"QID26_3"}:  0
## (Other)              :  0
##       Personal_Views_Relate_To_Model
## Disagree              :151
## Agree                 :126
## Strongly disagree     : 69
## Strongly Agree        : 43
##                       :  1
## {"ImportId":"QID26_4"}:  0
## (Other)              :  0
```

```
print('Confidence')
```

```
## [1] "Confidence"
```

```
percent_table(data$Personal_Views_Confident)
```

```
##
##              Agree         Strongly Agree              Disagree
##         0.49487179             0.26923077            0.20000000
##     Strongly disagree                     {"ImportId":"QID26_1"}
##         0.03589744             0.00000000            0.00000000
```

Our group seems to be a very confident bunch. With well over 2/3rds being confident, it seems unlikely we'll find any discerning difference between treatment and control.

```
print('Beauty')
```

```
## [1] "Beauty"
```

```
percent_table(data$Personal_Views_Beautiful)
```

```
##
##              Agree              Disagree         Strongly Agree
##         0.510256410           0.251282051            0.197435897
##     Strongly disagree                     {"ImportId":"QID26_2"}
##         0.038461538           0.002564103            0.000000000
```

Considering how confident our Turk Users are, it's not surprising that they are also fairly positive in their views on their own beauty. There doesn't seem a lot of disagreement on their self beauty.

```
print('Beauty Importance')
```

```
## [1] "Beauty Importance"
```

```
percent_table(data$Personal_Views_Beauty_Importance)
```

```
##
##              Agree              Disagree         Strongly Agree
##         0.48461538            0.28974359            0.16410256
##     Strongly disagree                     {"ImportId":"QID26_3"}
##         0.06153846            0.00000000            0.00000000
```

Similar to before, our Turks have a consistent view of beauty, with only about a 3rd disagreeing.

```
print('Relate to Model')
```

```
## [1] "Relate to Model"
```

```
percent_table(data$Personal_Views_Relate_To_Model)
```

```
## 
##            Disagree                    Agree      Strongly disagree
##          0.387179487            0.323076923            0.176923077
##       Strongly Agree                               {"ImportId":"QID26_4"}
##          0.110256410            0.002564103            0.000000000
```

Finally, the relating to the model seem to half a real split in the data. This is probably due to the treatment.

We'll recode all of our responses to a logical numeric value. 1 will be agree, 2 for strongly agree, -1 for disagree, -2 for strongly disagree.

```r
recode <- function(field){
  out <- 0
  if(field == 'Agree'){
    out <- 1
  } else if(field == 'Strongly agree'){
    out <- 2
  } else if(field == 'Disagree'){
    out <- -1
  } else if(field == 'Strongly disagree'){
    out <- -2
  }
  out <- as.numeric(out)
  return(out)
}


data$Beautiful <- sapply(data$Personal_Views_Beautiful, FUN =  recode)
data$Confident <- sapply(data$Personal_Views_Confident, FUN =  recode)
data$Importance <- sapply(data$Personal_Views_Beauty_Importance, FUN =  recode)
data$Relate <- sapply(data$Personal_Views_Relate_To_Model, FUN =  recode)
```
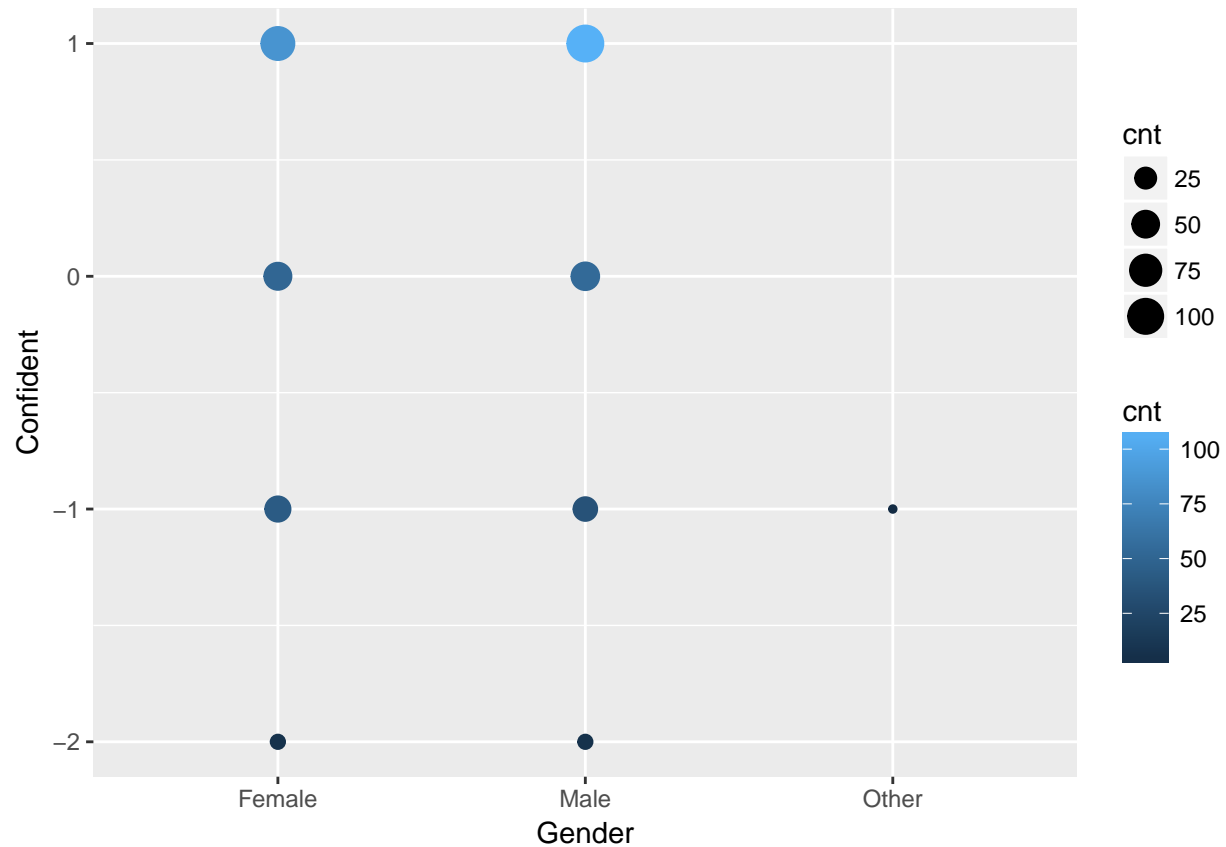
```r
qplot(data = data, x = Confident, fill = Gender, main = 'Confidence Split by Gender' )
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Confidence Split by Gender
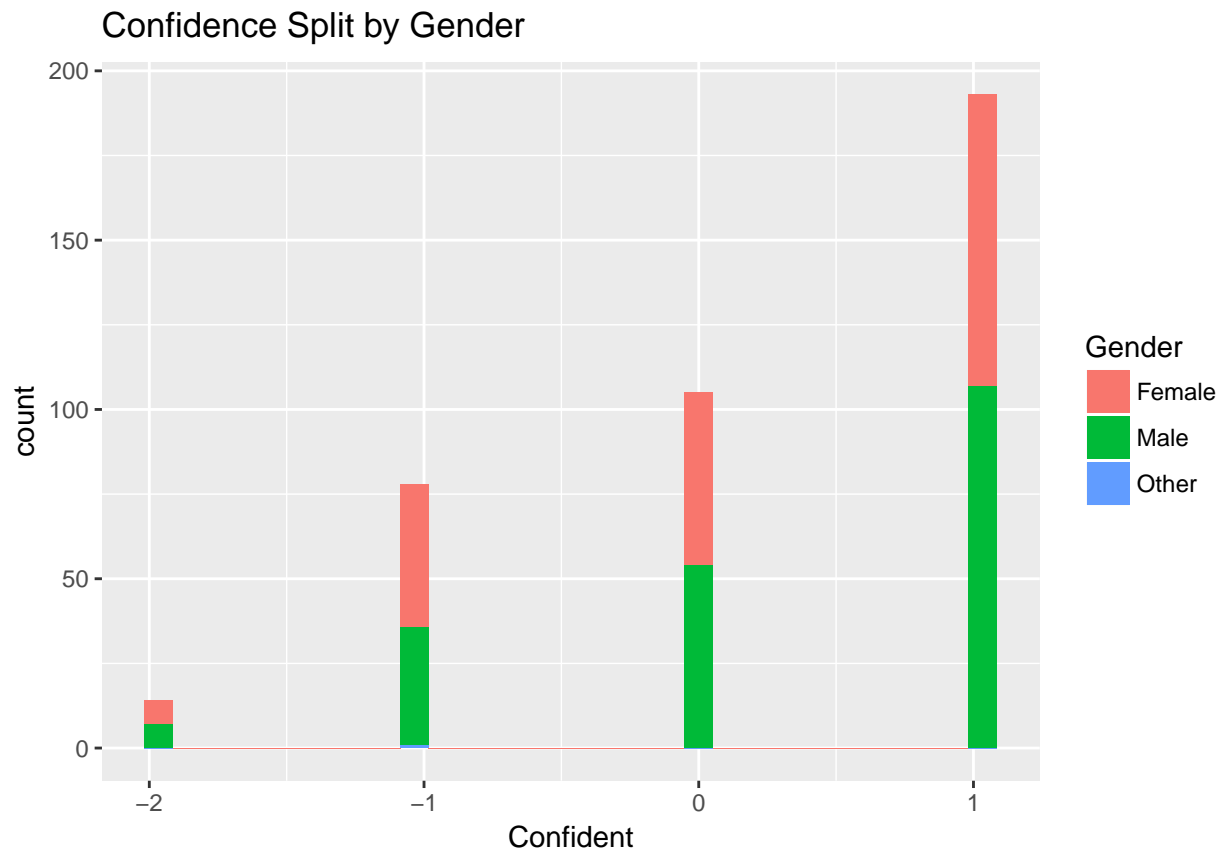

```r
conf_analysis <- data %>% group_by(Gender,Confident) %>% summarize(cnt = n())
qplot(data = conf_analysis, x = Gender, y = Confident, size = cnt, color = cnt)
```

```r
print(paste('Male Conf', mean(data$Confident[data$Gender=='Male'])))
```
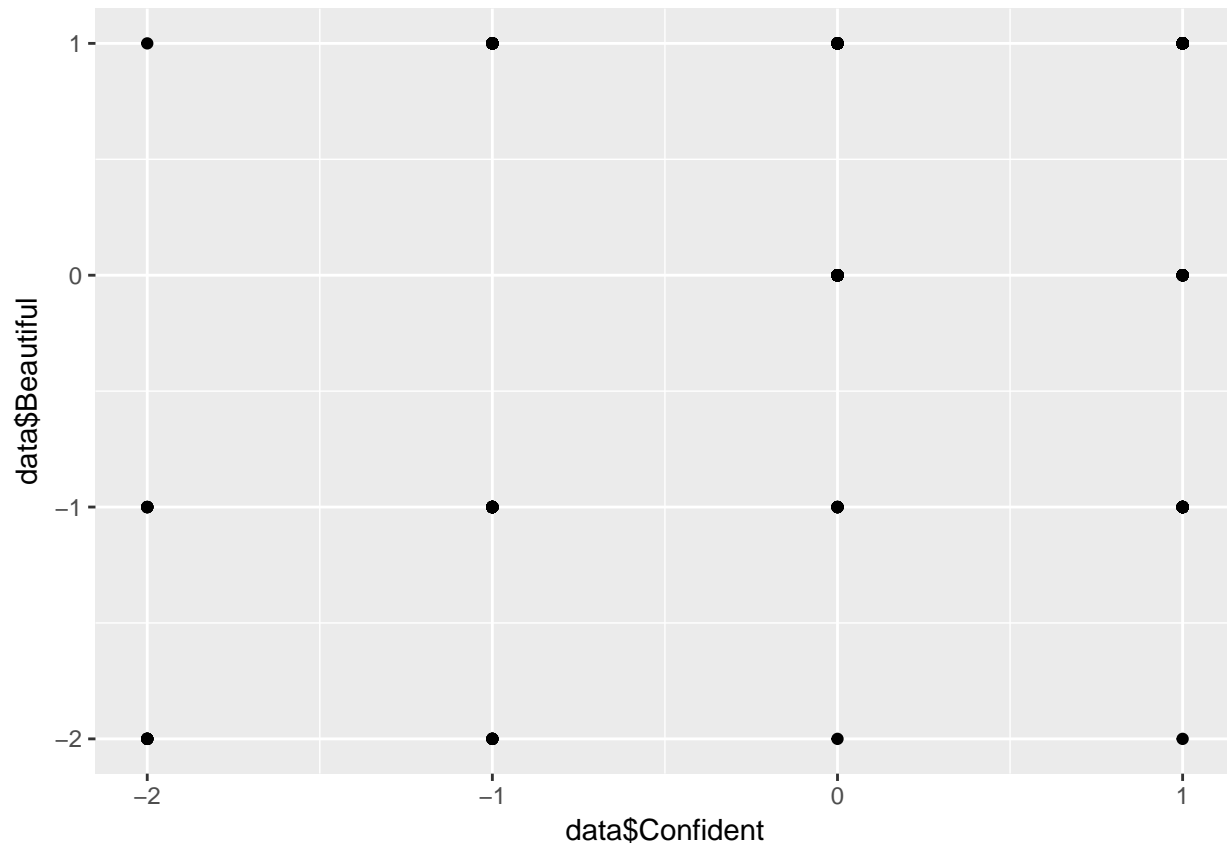
```
## [1] "Male Conf 0.285714285714286"
```

```r
print(paste('Female Conf', mean(data$Confident[data$Gender=='Female'])))
```

```
## [1] "Female Conf 0.161290322580645"
```

All of our turks a fairly confident, but the male population seems to be a little bit more confident than the ladies.

```r
qplot(data$Confident, data$Beautiful)
```

```
cor.test(data$Confident,data$Beautiful)
```

```
##
##  Pearson's product-moment correlation
##
## data:  data$Confident and data$Beautiful
## t = 9.6531, df = 388, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3563296 0.5167800
## sample estimates:
##       cor
## 0.4400606
```

```
chisq.test(data$Confident, data$Beautiful)
```

```
##
##  Pearson's Chi-squared test
##
## data:  data$Confident and data$Beautiful
## X-squared = 258.3, df = 9, p-value < 2.2e-16
```

Randomization Validation

```
qplot(data = data, x = Gender, geom = 'bar', fill = Group)
```
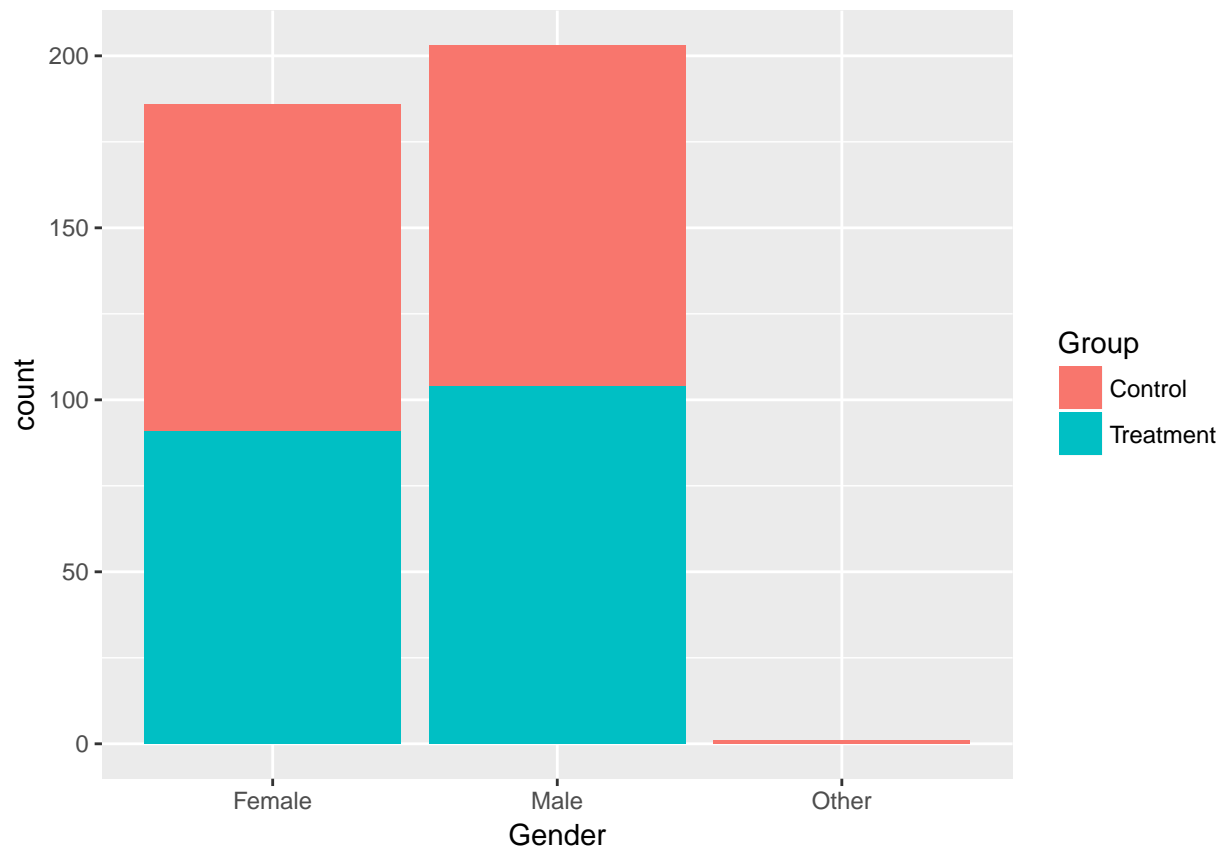
```
table(data$Gender)
```

```
##
##                                  {"ImportId":"QID4"}
##                0                                    0
##           Female                                 Male
##              186                                  203
##            Other Which gender do you identify with?
##                1                                    0
```

```
table(data$Group)
```

```
##
##                {"ImportId":"Group"}           Control
##                0                    0              195
##            Group          Treatment
##                0                195
```

It looks like our randomization worked very well. We got an even split on treatment and control and an almost even split on Male and Females. We have little reason to beleive the randomization didn't work properly.

Randomization Inference

To check to see whether we just got lucky, or actually found a treatment effect before diving into linear models, we'll run a randomization inference check on the ATE between treatment and control.

```r
treat <- filter(data, Group == 'Treatment')
control <- filter(data, Group == 'Control')

ate_Beautiful <- mean(treat$Beautiful) - mean(control$Beautiful)
ate_Confident <- mean(treat$Confident) - mean(control$Confident)
ate_Importance <- mean(treat$Importance) - mean(control$Importance)
ate_Relate <- mean(treat$Relate) - mean(control$Relate)

#
n <- 1e4
copy <- data
taus <- data.frame(matrix(NA, nrow = n, ncol = 4))
names(taus) <- c('Beauty','Confidence','Importance','Relate')
for(i in 1:n){
  copy$Group <- sample(data$Group)
  treat <- filter(copy, Group == 'Treatment')
  control <- filter(copy, Group == 'Control')

  taus[i,1] <- mean(treat$Beautiful) - mean(control$Beautiful)
  taus[i,2] <- mean(treat$Confident) - mean(control$Confident)
  taus[i,3] <- mean(treat$Importance) - mean(control$Importance)
  taus[i,4] <- mean(treat$Relate) - mean(control$Relate)
}

p_beauty <- sum(taus$Beauty > ate_Beautiful)/n
p_confident <- sum(taus$Confidence > ate_Confident)/n
p_relate <- sum(taus$Relate > ate_Relate)/n
p_importance <- sum(taus$Importance > ate_Importance)/n
print(paste('Pval Beauty',p_beauty))
```

```
## [1] "Pval Beauty 0.4523"
```

```r
print(paste('Pval Confident', p_confident))
```

```
## [1] "Pval Confident 0.3727"
```

```r
print(paste('Pval Relate', p_relate))
```

```
## [1] "Pval Relate 0.0153"
```

```r
print(paste('Pval Important', p_importance))
```

```
## [1] "Pval Important 0.6832"
```

Our only variable that seems to have a real a statistically significant ATE is the relating to the model. As noted before, our group of Turks already had a a pretty high level of self beauty and confidence, so we didn't expect to see it change much with treatment.

RI Plots

```r
qplot(taus$Beauty, bins = 30, main = paste('RI for Beauty Pvalue =', p_beauty)) + geom_vline(xintercept
```

RI for Beauty Pvalue = 0.4523
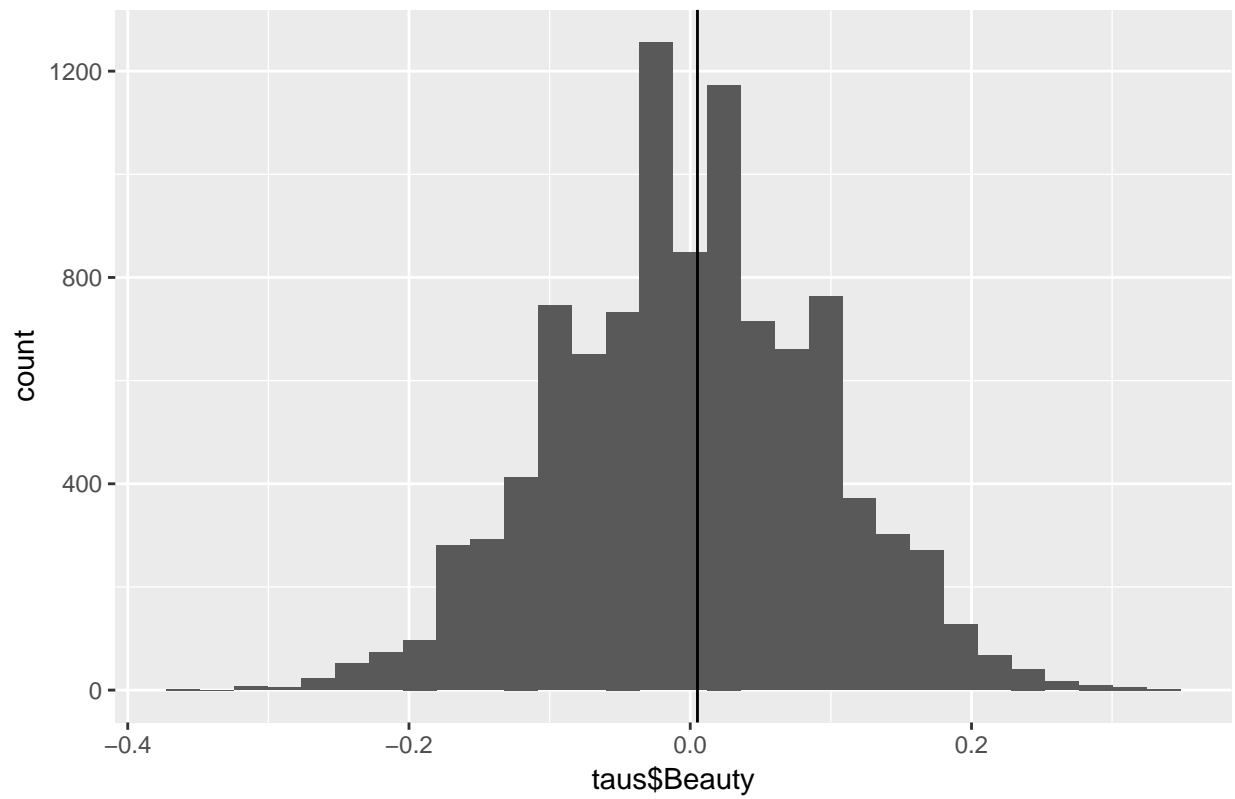

```
qplot(taus$Confidence, bins = 30, main = paste('RI for Confidence Pvalue =',p_confident)) + geom_vline(
```

RI for Confidence Pvalue = 0.3727
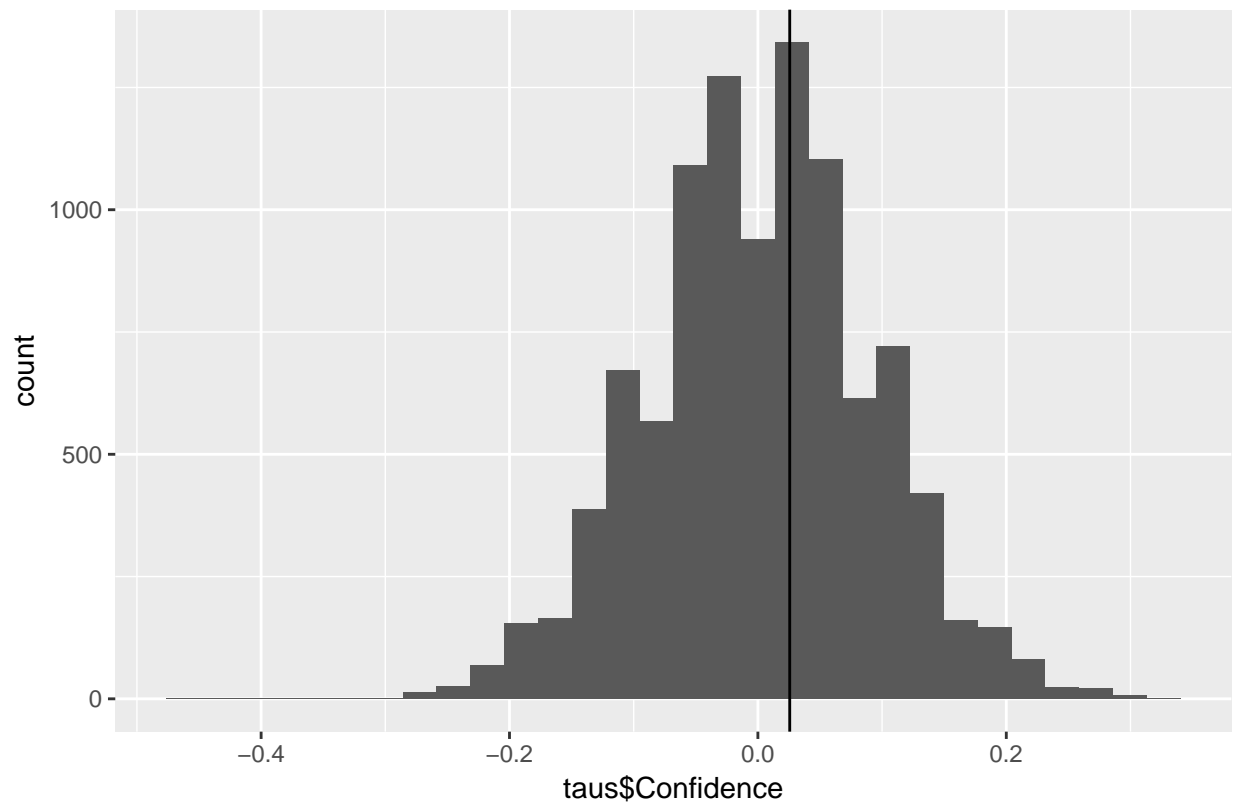
```r
qplot(taus$Importance, bins = 30, main = paste('RI for Importance Pvalue =',p_importance)) + geom_vline
```

RI for Importance Pvalue = 0.6832
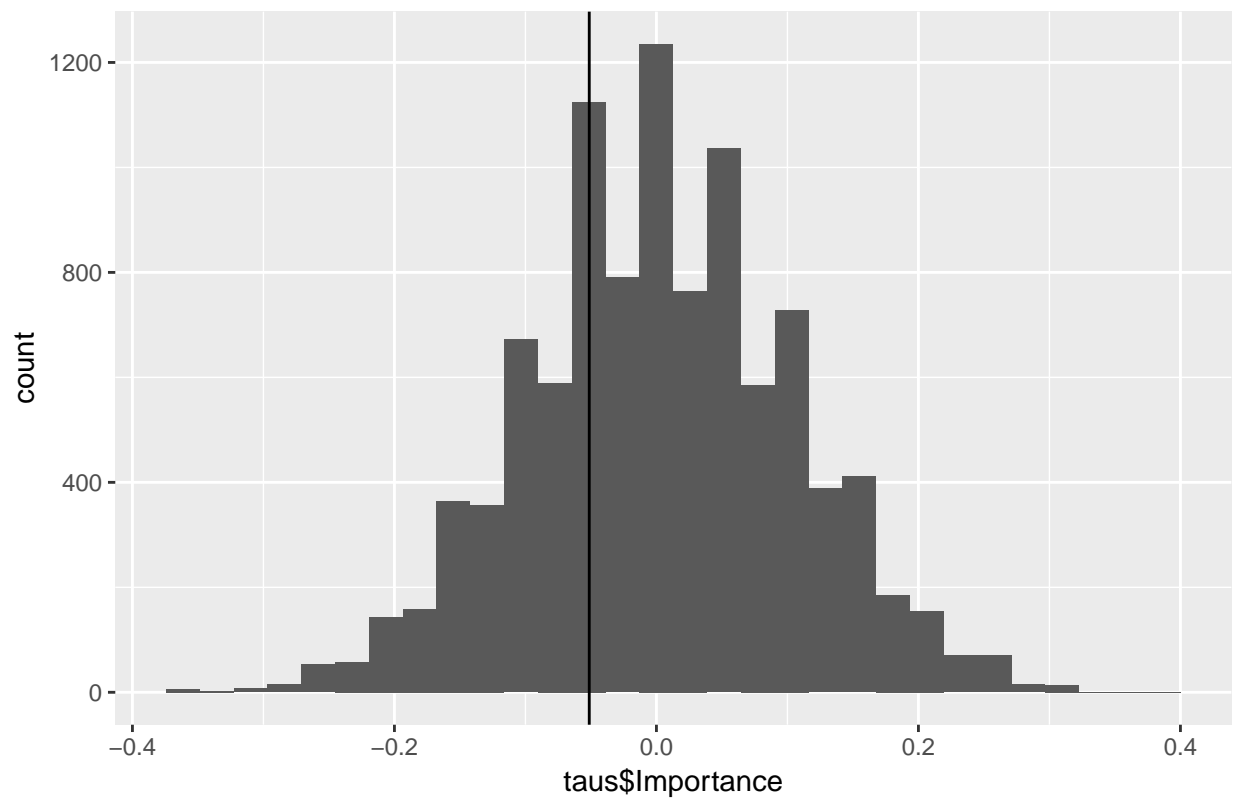
```r
qplot(taus$Relate, bins = 30, main = paste('RI for Relating Pvalue =',p_relate)) + geom_vline(xintercep
```

RI for Relating Pvalue = 0.0153
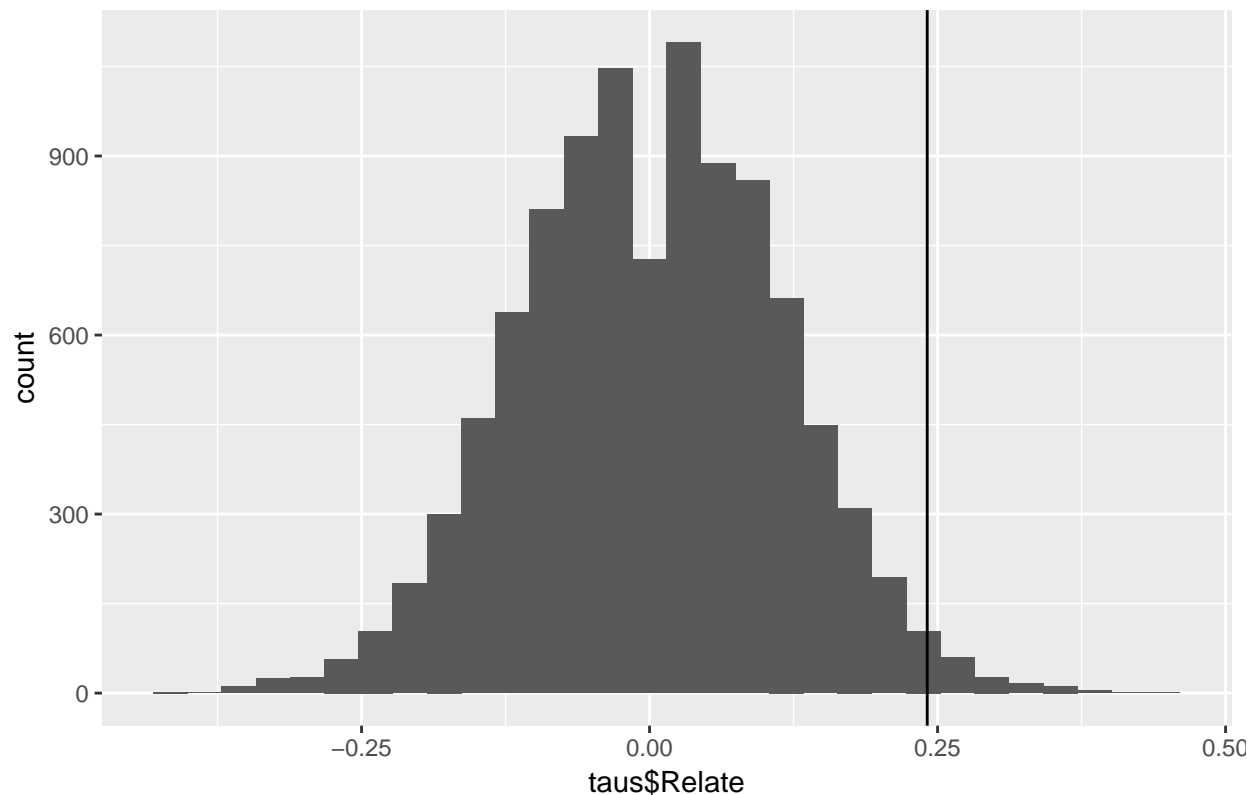
The distributions are fairly normal and centered at zero. We can be fairly confident in our treatment effect on Relating to the model, but not any of the other items.

Linear Models

```
m_beauty <- lm(data = data, Beautiful ~ Group)
summary(m_beauty)
```

```
##
## Call:
## lm(formula = Beautiful ~ Group, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.1846 -1.1795  0.8154  0.8205  0.8205
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.179487   0.067436   2.662   0.0081 **
## GroupTreatment 0.005128   0.095369   0.054   0.9571
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9417 on 388 degrees of freedom
```

```
## Multiple R-squared:  7.452e-06,   Adjusted R-squared:  -0.00257
## F-statistic: 0.002891 on 1 and 388 DF,   p-value: 0.9571
```

The linear model version of statistical inference gives similar results to the RI, but with even larger pvalue. We're guessing the high levels of self beauty didn't leave much room for the treatment to have an effect.

```
m_confident <- lm(data = data, Confident ~ Group)
summary(m_confident)
```

```
##
## Call:
## lm(formula = Confident ~ Group, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2359 -0.2359 -0.2103  0.7641  0.7897
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.21026    0.06375   3.298  0.00106 **
## GroupTreatment  0.02564    0.09016   0.284  0.77626
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8903 on 388 degrees of freedom
## Multiple R-squared:  0.0002084,   Adjusted R-squared:  -0.002368
## F-statistic: 0.08088 on 1 and 388 DF,   p-value: 0.7763
```

Similar to the beauty inference. It doesn't look like there was much room for a treatment to cause a noticable change.

```
m_importance <- lm(data = data, Importance ~ Group)
summary(m_importance)
```

```
##
## Call:
## lm(formula = Importance ~ Group, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.09744 -1.04615 -0.04615  0.90256  0.95385
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.09744    0.07232   1.347    0.179
## GroupTreatment -0.05128    0.10228  -0.501    0.616
##
## Residual standard error: 1.01 on 388 degrees of freedom
## Multiple R-squared:  0.0006475,   Adjusted R-squared:  -0.001928
## F-statistic: 0.2514 on 1 and 388 DF,   p-value: 0.6164
```

No surprises here. It looks like the video doesn't really change what you think about the importance of beauty.

```
m_relate <- lm(data = data, Relate ~ Group)
summary(m_relate)
```

```
##
```

```
## Call:
## lm(formula = Relate ~ Group, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7026 -0.7026 -0.4615  1.2974  1.5385
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.53846    0.07958  -6.766  4.9e-11 ***
## GroupTreatment  0.24103    0.11255   2.142   0.0329 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.111 on 388 degrees of freedom
## Multiple R-squared:  0.01168,    Adjusted R-squared:  0.009134
## F-statistic: 4.586 on 1 and 388 DF,  p-value: 0.03286
```

Our linear model confirms what we saw with the RI. Remember, we recoded Agree to 1 point and Strongly agree to 2 points, so the treatment effectively yields a .25 point increase in how much they person can relate to the model. It's not straightforward what the coefficient means, but effectively people relate more to the non-stereotypical model. We'll explore how the other co-variates effect the ATE.

```r
stargazer(m_beauty, m_confident, m_importance, m_relate, header=F, type = 'text')
```

```
##
## ==============================================================================
##                                    Dependent variable:
##                      -----------------------------------------------
##                      Beautiful Confident Importance  Relate
##                         (1)       (2)       (3)       (4)
## ------------------------------------------------------------------------------
## GroupTreatment          0.005     0.026    -0.051    0.241**
##                        (0.095)   (0.090)   (0.102)   (0.113)
##
## Constant               0.179***  0.210***   0.097    -0.538***
##                        (0.067)   (0.064)   (0.072)   (0.080)
##
## ------------------------------------------------------------------------------
## Observations            390       390       390       390
## R2                     0.00001   0.0002    0.001     0.012
## Adjusted R2            -0.003    -0.002    -0.002     0.009
## Residual Std. Error (df = 388)  0.942     0.890     1.010     1.111
## F Statistic (df = 1; 388)       0.003     0.081     0.251     4.586**
## ==============================================================================
## Note:                                      *p<0.1; **p<0.05; ***p<0.01
```

Covariate Analysis

```r
m_beauty <- lm(data = data, Beautiful ~ Group + Gender)
summary(m_beauty)
```

```
##
```

```
## Call:
## lm(formula = Beautiful ~ Group + Gender, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3016 -1.0832  0.6984  0.9157  0.9168
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.301615   0.082447   3.658 0.000289 ***
## GroupTreatment -0.001103   0.094453  -0.012 0.990689
## GenderMale     -0.217306   0.094543  -2.298 0.022068 *
## GenderOther    -2.301615   0.934846  -2.462 0.014252 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9312 on 386 degrees of freedom
## Multiple R-squared:  0.0272, Adjusted R-squared:  0.01964
## F-statistic: 3.597 on 3 and 386 DF,  p-value: 0.01373
```

```
data %>% group_by(Group, Gender) %>% summarize(avg = mean(Beautiful))
```

```
## Source: local data frame [5 x 3]
## Groups: Group [?]
##
##      Group Gender         avg
##      <fctr> <fctr>       <dbl>
## 1   Control Female  0.33684211
## 2   Control   Male  0.05050505
## 3   Control  Other -2.00000000
## 4 Treatment Female  0.26373626
## 5 Treatment   Male  0.11538462
```

It looks like our treatment still isn't significant with the inclusion of gender as a co-variate, but interestingly enough women have a much higher rate of beauty than men do.

```
m_confident <- lm(data = data, Confident ~ Group + Gender)
summary(m_confident)
```

```
##
## Call:
## lm(formula = Confident ~ Group + Gender, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2938 -0.2938 -0.1532  0.7228  0.8468
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.15319    0.07865   1.948   0.0521 .
## GroupTreatment   0.01655    0.09010   0.184   0.8544
## GenderMale       0.12404    0.09018   1.375   0.1698
## GenderOther     -1.15319    0.89174  -1.293   0.1967
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.8883 on 386 degrees of freedom
## Multiple R-squared:  0.009848,   Adjusted R-squared:  0.002152
## F-statistic:  1.28 on 3 and 386 DF,  p-value: 0.281
```

```r
data %>% group_by(Group, Gender) %>% summarise(avg = mean(Confident))
```

```
## Source: local data frame [5 x 3]
## Groups: Group [?]
##
##       Group Gender         avg
##      <fctr> <fctr>       <dbl>
## 1   Control Female  0.25263158
## 2   Control   Male  0.18181818
## 3   Control  Other -1.00000000
## 4 Treatment Female  0.06593407
## 5 Treatment   Male  0.38461538
```

Again, adding the gender covariate didn't improve the model much.

```r
m_importance <- lm(data = data, Importance ~ Group + Gender)
summary(m_importance)
```

```
##
## Call:
## lm(formula = Importance ~ Group + Gender, data = data)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -2.20039 -1.01983  0.03811  0.85755  1.03811
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.20039    0.08878   2.257   0.0246 *
## GroupTreatment -0.05794    0.10171  -0.570   0.5693
## GenderMale     -0.18056    0.10181  -1.773   0.0769 .
## GenderOther    -2.20039    1.00670  -2.186   0.0294 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.003 on 386 degrees of freedom
## Multiple R-squared:  0.0198, Adjusted R-squared:  0.01218
## F-statistic: 2.599 on 3 and 386 DF,  p-value: 0.05194
```

```r
data %>% group_by(Group, Gender) %>% summarise(avg = mean(Importance))
```

```
## Source: local data frame [5 x 3]
## Groups: Group [?]
##
##       Group Gender         avg
##      <fctr> <fctr>       <dbl>
## 1   Control Female  0.31578947
## 2   Control   Male -0.09090909
## 3   Control  Other -2.00000000
## 4 Treatment Female  0.02197802
## 5 Treatment   Male  0.06730769
```

Treatment effect remains insignificant with additional co-variates gender covariate. Somewhat surprisingly,

men don't place as high importance on beauty.

```
m_relate <- lm(data = data, Relate ~ Group + Gender)
summary(m_relate)
```

```
##
## Call:
## lm(formula = Relate ~ Group + Gender, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9472 -0.7032 -0.2444  1.0528  1.7556
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -0.2968     0.0963  -3.082   0.0022 **
## GroupTreatment   0.2440     0.1103   2.212   0.0275 *
## GenderMale      -0.4587     0.1104  -4.154 4.02e-05 ***
## GenderOther     -1.7032     1.0919  -1.560   0.1196
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.088 on 386 degrees of freedom
## Multiple R-squared:  0.05822,    Adjusted R-squared:  0.0509
## F-statistic: 7.954 on 3 and 386 DF,  p-value: 3.7e-05
```

```
data %>% group_by(Group, Gender) %>% summarise(avg = mean(Relate))
```

```
## Source: local data frame [5 x 3]
## Groups: Group [?]
##
##        Group Gender         avg
##        <fctr> <fctr>       <dbl>
## 1    Control Female -0.35789474
## 2    Control   Male -0.69696970
## 3    Control  Other -2.00000000
## 4 Treatment Female  0.01098901
## 5 Treatment   Male -0.56730769
```

With the gender co-variate our treatment remains statistically significant. Unsurprisingly, the effect is a lot stronger for women than men.

Covariate Summary

```
stargazer(m_beauty, m_confident, m_importance, m_relate, header=F, type = 'text')
```

```
##
## ==========================================================================
##                                    Dependent variable:
##                      --------------------------------------------
##                      Beautiful Confident Importance  Relate
##                         (1)       (2)       (3)       (4)
## --------------------------------------------------------------------------
## GroupTreatment         -0.001     0.017     -0.058   0.244**
```

```
##                                        (0.094)   (0.090)   (0.102)    (0.110)
##
## GenderMale                             -0.217**   0.124    -0.181*   -0.459***
##                                        (0.095)   (0.090)   (0.102)    (0.110)
##
## GenderOther                            -2.302**  -1.153    -2.200**   -1.703
##                                        (0.935)   (0.892)   (1.007)    (1.092)
##
## Constant                               0.302***   0.153*    0.200**  -0.297***
##                                        (0.082)   (0.079)   (0.089)    (0.096)
##
## -----------------------------------------------------------------------
## Observations                             390       390       390        390
## R2                                       0.027     0.010     0.020      0.058
## Adjusted R2                              0.020     0.002     0.012      0.051
## Residual Std. Error (df = 386)           0.931     0.888     1.003      1.088
## F Statistic (df = 3; 386)                3.597**   1.280     2.599*     7.954***
## =======================================================================
## Note:                                            *p<0.1; **p<0.05; ***p<0.01
```

Gender did make a pretty big significance in almost every model, even when the treatment was insignificant. Since we only suspected the treatment effect to really have a effect on gender we aren't going fishing for other covariate significance.

Image analysis

Preparing the data

Essentially we are counting a 1 if the response matched whether you were in treatment or control.

```r
# Combine and recode randomization 1 & 2 for the images
images <- c('Passion', 'Coffee', 'Couple', 'Work', 'Fit')
questions <- c('_i_identify_', '_i_prefer_', '_o_prefer_', '_validate_')
randomization <- c('1', '2')

# Correct column names misspellings
data <- dplyr::rename(data, Coffee_validate_1 = Coffe_validate_1,
          Fit_i_identify_1 = FIt_i_identify_1,
          Work_i_identify_2 = work_i_identify_2)

# Recode image output to one column
for (image in images){
  for (question in questions){
    column1 <- paste(image, question, '1', sep = "")
    data[[column1]] <- ifelse(data[[column1]] == 'Ad 1', 2,
                        ifelse(data[[column1]] == 'Ad 2', 1, 0))
    column2 <- paste(image, question, '2', sep = "")
    data[[column2]] <- ifelse(data[[column2]] == 'Ad 2', 2,
                        ifelse(data[[column2]] == 'Ad 1', 1, 0))
    new_column <- paste(image, question, sep ="")
    data[[new_column]] <- data[[column1]] + data[[column2]] - 1
    data[[new_column]] <- ifelse(data[[new_column]] == -1, NA, data[[new_column]])
  }
```

```
}
```

Getting the ATE for IMAGES questions

```r
images <- c('Passion', 'Coffee', 'Couple', 'Work', 'Fit')
questions <- c('_i_identify_', '_i_prefer_', '_o_prefer_', '_validate_')
models <- list(20)
i <- 1
for (image in images){
  for (question in questions){
    column <- paste(image, question, sep = "")
    l1 <- lm(data[[column]] ~ data[['Group']])
    models <- list(models, l1)
    i <- i + 1
  print(column)

  print(coeftest(l1))
  }
}
```

```
## [1] "Passion_i_identify_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.352941   0.034151 10.3348   <2e-16 ***
## data[["Group"]]Treatment -0.063467   0.048105 -1.3193   0.1879
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Passion_i_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.393617   0.034433 11.4316  < 2e-16 ***
## data[["Group"]]Treatment -0.106383   0.048695 -2.1847  0.02953 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Passion_o_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.606383   0.035850 16.9143   <2e-16 ***
## data[["Group"]]Treatment -0.016909   0.050566 -0.3344   0.7383
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Passion_validate_"
##
```

```
## t test of coefficients:
##
##                             Estimate Std. Error  t value Pr(>|t|)
## (Intercept)                 0.989362   0.009842 100.5248   <2e-16 ***
## data[["Group"]]Treatment -0.015678   0.013882  -1.1293   0.2595
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Coffee_i_identify_"
##
## t test of coefficients:
##
##                             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)                 0.184783   0.027485  6.7230 6.739e-11 ***
## data[["Group"]]Treatment -0.036634   0.038612 -0.9488    0.3433
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Coffee_i_prefer_"
##
## t test of coefficients:
##
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.302703   0.032951  9.1865   <2e-16 ***
## data[["Group"]]Treatment -0.050071   0.046292 -1.0816   0.2801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Coffee_o_prefer_"
##
## t test of coefficients:
##
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.421622   0.036626 11.5116   <2e-16 ***
## data[["Group"]]Treatment 0.062589   0.051455  1.2164   0.2246
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Coffee_validate_"
##
## t test of coefficients:
##
##                              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)                 0.9891892  0.0092421 107.0307   <2e-16 ***
## data[["Group"]]Treatment -0.0102418  0.0129840  -0.7888   0.4307
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Couple_i_identify_"
##
## t test of coefficients:
##
##                             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)                 0.206349   0.029949  6.8901 2.367e-11 ***
```

```
## data[["Group"]]Treatment 0.017055   0.042410  0.4021     0.6878
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Couple_i_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.3174603  0.0339734  9.3444   <2e-16 ***
## data[["Group"]]Treatment 0.0016886  0.0481095  0.0351    0.972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Couple_o_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.510638   0.036259  14.083   <2e-16 ***
## data[["Group"]]Treatment 0.079787   0.051278   1.556    0.1206
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Couple_validate_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.9735450  0.0098415 98.9222   <2e-16 ***
## data[["Group"]]Treatment 0.0157598  0.0139552  1.1293    0.2595
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Work_i_identify_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.462766   0.036513  12.6742   <2e-16 ***
## data[["Group"]]Treatment 0.039880   0.051568   0.7733    0.4398
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Work_i_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.6648936  0.0344017 19.3274   <2e-16 ***
## data[["Group"]]Treatment 0.0089032  0.0487163  0.1828    0.8551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## [1] "Work_o_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.760638   0.030542 24.9047   <2e-16 ***
## data[["Group"]]Treatment 0.027721   0.043136  0.6427   0.5208
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Work_validate_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.9787234  0.0098696 99.1659   <2e-16 ***
## data[["Group"]]Treatment 0.0054036  0.0139392  0.3877   0.6985
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Fit_i_identify_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.403226   0.035803 11.2623   <2e-16 ***
## data[["Group"]]Treatment -0.030885   0.050498 -0.6116   0.5412
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Fit_i_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.381720   0.035818 10.6571   <2e-16 ***
## data[["Group"]]Treatment 0.011897   0.050520  0.2355    0.814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Fit_o_prefer_"
##
## t test of coefficients:
##
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.580645   0.036179 16.0492   <2e-16 ***
## data[["Group"]]Treatment 0.015100   0.051029  0.2959   0.7675
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## [1] "Fit_validate_"
##
## t test of coefficients:
##
```

```
##                         Estimate Std. Error  t value Pr(>|t|)
## (Intercept)            0.9892473  0.0053474 184.9973   <2e-16 ***
## data[["Group"]]Treatment 0.0107527  0.0075422   1.4257   0.1548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
stargazer(models, type = 'text')
```

```
##
## ==
## 20
## --
##
## =============================================================================
##
##                          ----------------------------------------------------
##
##                              (1)              (2)              (3)              (4)
## ---------------------------------------------------------------------------------
## data[["Group"]]Treatment    -0.063          -0.106**          -0.017          -0.016
##                             (0.048)          (0.049)          (0.051)          (0.014)
##
## Constant                    0.353***         0.394***         0.606***         0.989***
##                             (0.034)          (0.034)          (0.036)          (0.010)
##
## ---------------------------------------------------------------------------------
## Observations                  377              376              378              378
## R2                           0.005            0.013            0.0003           0.003
## Adjusted R2                  0.002            0.010            -0.002           0.001
## Residual Std. Error    0.467 (df = 375)  0.472 (df = 374)  0.492 (df = 376)  0.135 (df =
## F Statistic           1.741 (df = 1; 375) 4.773** (df = 1; 374) 0.112 (df = 1; 376) 1.275 (df = 1
## =============================================================================
## Note:
```

Only 1 question out of the 20 had any real significance. Essentially, for the passion photo, showing treatment led to preferring the non-stereotypical image. At an alpha rate of .05 we know that we could have simply gotten this result from a random selection at least 1 out of 20 times. To make sure we there is a real effect here, we'll need to run the experiment again, otherwise we might be fishing for significance.

Taking the total and average by questions for all images

```r
questions <- c('_i_identify_', '_i_prefer_', '_o_prefer_', '_validate_')
images <- c('Passion', 'Coffee', 'Couple', 'Work', 'Fit')

for (question in questions){
  question_average <- paste(question, 'average', sep = "")
  data[[question_average]] <- rep(0, nrow(data))
  for (image in images){
    column <- paste(image, question, sep = "")
    data[[question_average]] <- data[[question_average]] + data[[column]]
  }
}
```

```
average_lm <- function(data, question){
  question_average <- paste(question, 'average', sep = "")
  #print(question_average)
  question_average <- lm(data[[question_average]] ~ data[['Group']] + data$Gender)
  #print(coeftest(question_average))
  return(question_average)
}

lm_i_identify <- average_lm(data, '_i_identify_')
lm_i_prefer <- average_lm(data, '_i_prefer_')
lm_o_prefer <- average_lm(data, '_o_prefer_')
lm_validate <- average_lm(data, '_validate_')

stargazer(lm_i_identify, lm_i_prefer, lm_o_prefer, lm_validate, header=T, type = 'text',
          column.labels = c('I Identify','I Prefer','Others Prefer','Validate'))
```

```
##
## ===============================================================================================
##                                          Dependent variable:
##                       ------------------------------------------------------------------------
##                                          data[[question_average]]
##                          I Identify        I Prefer        Others Prefer           Val
##                             (1)               (2)               (3)
## -----------------------------------------------------------------------------------------------
## data[["Group"]]Treatment   -0.081            -0.154             0.165               -0
##                            (0.139)           (0.131)           (0.147)              (0
##
## GenderMale                 0.351**           0.677***          0.584***              0
##                            (0.139)           (0.132)           (0.147)              (0
##
## GenderOther                -0.392            -0.689             2.418*              0.0
##                            (1.321)           (1.249)           (1.402)              (0
##
## Constant                   1.392***          1.689***          2.582***            5.00
##                            (0.123)           (0.116)           (0.130)              (0
##
## -----------------------------------------------------------------------------------------------
## Observations                360               360               362
## R2                         0.019             0.074             0.051                 0
## Adjusted R2                0.011             0.066             0.043                -0
## Residual Std. Error   1.315 (df = 356)   1.244 (df = 356)   1.396 (df = 358)    0.074 (
## F Statistic           2.284* (df = 3; 356) 9.430*** (df = 3; 356) 6.415*** (df = 3; 358) 0.663 (d
## ===============================================================================================
## Note:                                                             *p<0.1; **p<0.05
```

Because the original data set was essentially 20 interactions between images and questions, we wanted to summarize the effects to each question. In aggregate, the measurement still makes sense. We are in effect seeing if each survey taker matches up to the stereotypical or non-stereotypical image and seeing if the treatment makes any difference. Even in aggregate none of the questions had any statistically significant treatment effect. Although, interestingly males seemed to match up with their photos of their treatment more than females.

Second data set

Load, Clean & Recode the data

```r
raw_data2 <- read.csv('C:/Users/achevall/Documents/UC Berkeley/w241/Final/BeatuyAd_CausalExperiment/Data

data2 <- filter_data(raw_data2)
data2 <- filter(data2, RecordedDate > '2017-08-10 00:00:00')

data2$Beautiful <- sapply(data2$Personal_Views_Beautiful, FUN =  recode)
data2$Confident <- sapply(data2$Personal_Views_Confident, FUN =  recode)
data2$Importance <- sapply(data2$Personal_Views_Beauty_Importance, FUN =  recode)
data2$Relate <- sapply(data2$Personal_Views_Relate_To_Model, FUN =  recode)

# Combine and recode randomization 1 & 2 for the images
images <- c('Passion', 'Coffee', 'Couple', 'Work', 'Fit')
questions <- c('_i_identify_', '_i_prefer_', '_o_prefer_', '_validate_')
randomization <- c('1', '2')

# Correct column names misspellings
data2 <- dplyr::rename(data2, Coffee_validate_1 = Coffe_validate_1,
           Fit_i_identify_1 = FIt_i_identify_1,
           Work_i_identify_2 = work_i_identify_2)

# Recode image output to one column
for (image in images){
  for (question in questions){
    column1 <- paste(image, question, '1', sep = "")
    data2[[column1]] <- ifelse(data2[[column1]] == 'Ad 1',
                            2, ifelse(data2[[column1]] == 'Ad 2', 1, 0))
    column2 <- paste(image, question, '2', sep = "")
    data2[[column2]] <- ifelse(data2[[column2]] == 'Ad 2',
                            2, ifelse(data2[[column2]] == 'Ad 1', 1, 0))
    new_column <- paste(image, question, sep ="")
    data2[[new_column]] <- data2[[column1]] + data2[[column2]] - 1
    data2[[new_column]] <- ifelse(data2[[new_column]] == -1, NA, data2[[new_column]])
  }
}
```

Text analysis

```r
columns_to_analyse <- c('Beautiful', 'Confident', 'Importance', 'Relate')
models <- list(4)
for (column in columns_to_analyse){
  l1 <- lm(data2[[column]] ~ data2[['Group']] + data2[['Gender']])

  models <- list(models, l1)
}
stargazer(models, header = FALSE, type = 'text', column.labels = columns_to_analyse)

##
## =
```

```
## 4
## -
##
## =========================================================================
##                              Dependent variable:
##                       ----------------------------------------
##                                  data2[[column]]
##                       Beautiful Confident Importance  Relate
##                          (1)       (2)       (3)       (4)
## -------------------------------------------------------------------------
## data2[["Group"]]Treatment   -0.002    0.012    -0.079   0.298***
##                            (0.095)   (0.088)   (0.096)   (0.109)
##
## data2[["Gender"]]Male        0.037    0.169*    -0.005  -0.259**
##                            (0.097)   (0.090)   (0.099)   (0.112)
##
## data2[["Gender"]]Other       0.787   -1.194     0.816   -1.560
##                            (0.960)   (0.889)   (0.973)   (1.101)
##
## Constant                    0.213**   0.194**   0.184**  -0.440***
##                            (0.090)   (0.083)   (0.091)   (0.103)
##
## -------------------------------------------------------------------------
## Observations                  405       405       405       405
## R2                          0.002     0.014     0.004     0.036
## Adjusted R2                -0.006     0.007    -0.004     0.029
## Residual Std. Error (df = 401)  0.956     0.885     0.969     1.097
## F Statistic (df = 3; 401)   0.262     1.897     0.483    4.994***
## =========================================================================
## Note:                              *p<0.1; **p<0.05; ***p<0.01
```

Our second experiment confirms the treatment effect on relating to the models in the non-stereotypical ad.

Image analysis

```
images <- c('Passion', 'Coffee', 'Couple', 'Work', 'Fit')
questions <- c('_i_identify_', '_i_prefer_', '_o_prefer_', '_validate_')

for (image in images){
  for (question in questions){
    column <- paste(image, question, sep = "")
    l1 <- lm(data2[[column]] ~ data2[['Group']])
  }
}
```

Taking the total and average by questions for all images

```
questions <- c('_i_identify_', '_i_prefer_', '_o_prefer_', '_validate_')
images <- c('Passion', 'Coffee', 'Couple', 'Work', 'Fit')

for (question in questions){
```

```
  question_average <- paste(question, 'average', sep = "")
  data2[[question_average]] <- rep(0, nrow(data2))
  for (image in images){
    column <- paste(image, question, sep = "")
    data2[[question_average]] <- data2[[question_average]] + data2[[column]]
  }
}

# for (question in questions){
#   question_average <- paste(question, 'average', sep = "")
#   print(question_average)
#   print(summary(data2[[question_average]]))
#   l1 <- lm(data2[[question_average]] ~ data2[['Group']] + data2$Gender)
#   print(summary(l1))
# }

lm_i_identify_ <- average_lm(data2, '_i_identify_')
lm_i_prefer_ <- average_lm(data2, '_i_prefer_')
lm_o_prefer_ <- average_lm(data2, '_o_prefer_')
lm_validate_ <- average_lm(data2, '_validate_')

stargazer(lm_i_identify_, lm_i_prefer_, lm_o_prefer_, lm_validate_, header=T, type = 'text',
          column.labels = c('I Identify','I Prefer','Others Prefer','Validate'))
```

```
##
## ============================================================================
##                                   Dependent variable:
##                         ----------------------------------------------
##                                   data[[question_average]]
##                         I Identify I Prefer Others Prefer Validate
##                             (1)       (2)       (3)         (4)
## --------------------------------------------------------------------------
## data[["Group"]]Treatment   -0.109    -0.104    -0.267**     0.017
##                            (0.134)   (0.137)   (0.135)     (0.012)
##
## GenderMale                 0.471***  0.609***  0.281**      0.010
##                            (0.136)   (0.139)   (0.137)     (0.012)
##
## GenderOther                -1.520    -1.870    -0.012       0.028
##                            (1.293)   (1.321)   (1.306)     (0.116)
##
## Constant                   1.520***  1.870***  3.012***    4.972***
##                            (0.124)   (0.127)   (0.126)     (0.011)
##
## --------------------------------------------------------------------------
## Observations                 370       370       370         370
## R2                          0.038     0.058     0.022       0.007
## Adjusted R2                 0.030     0.050     0.014       -0.001
## Residual Std. Error (df = 366)  1.287  1.315     1.300       0.116
## F Statistic (df = 3; 366)  4.831***  7.504***  2.698**      0.891
## ============================================================================
## Note:                                      *p<0.1; **p<0.05; ***p<0.01
```

Interestingly, our 2nd experiment did find significance of people matching the image that they thought

others would prefer. It's hard to trust this result, as the 1st experiment performed so differently. However, interestingly we see once again that males tended to match the non-stereotypical ad regardless of treatment.