Final_Project_Markdown

Ashton Chevallier, Lisa Minas, Olivier

August 12, 2017

Dataload

```
library(dplyr)
library(ggplot2)
library(stargazer)

#Change Directory
raw_data <- read.csv('/Users/ozimmer/GoogleDrive/berkeley/w241/BeatuyAd_CausalExperiment/Data/BeautyAds
```

Exploratory Data Analysis

Filters

```
#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

```
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
##
##
```

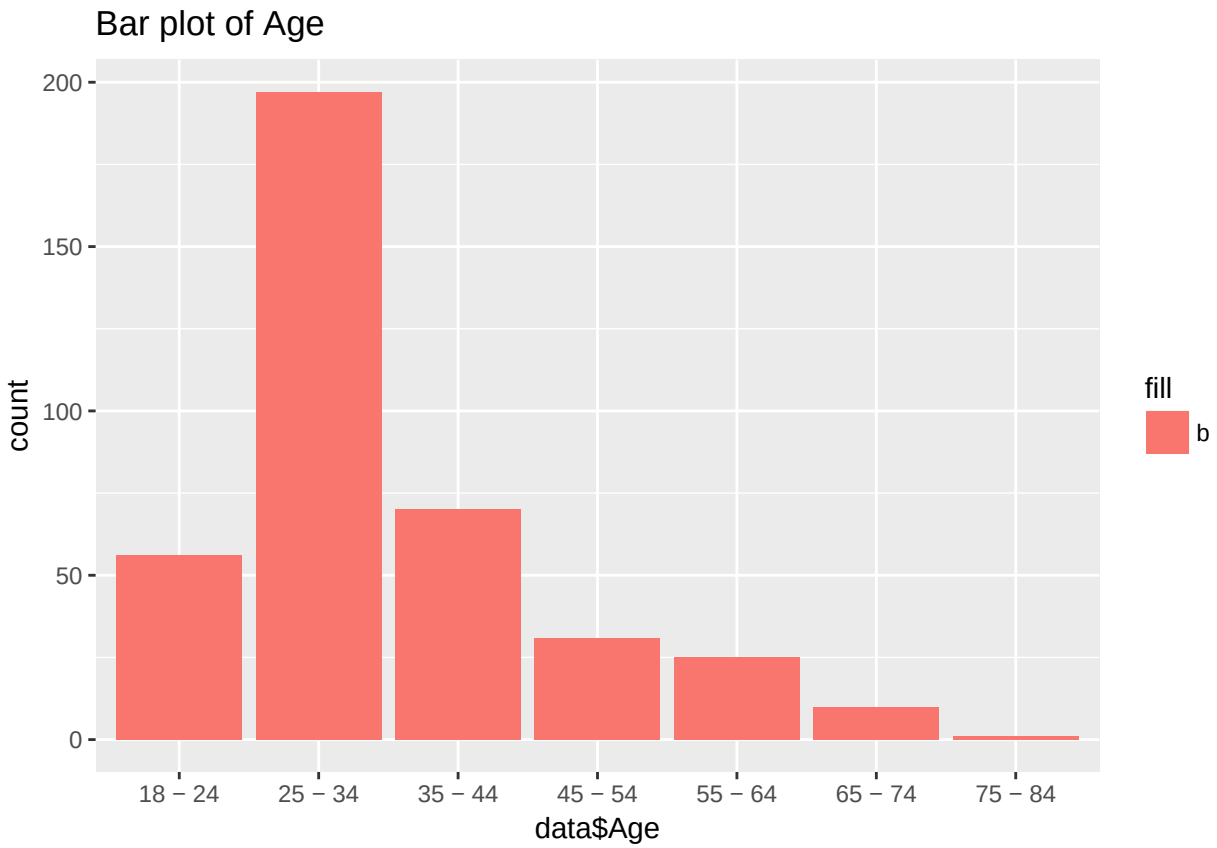
```
percent_table <- function(column){
  print(head(sort(table(sort(column))/length(column), decreasing = TRUE)))
}

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.

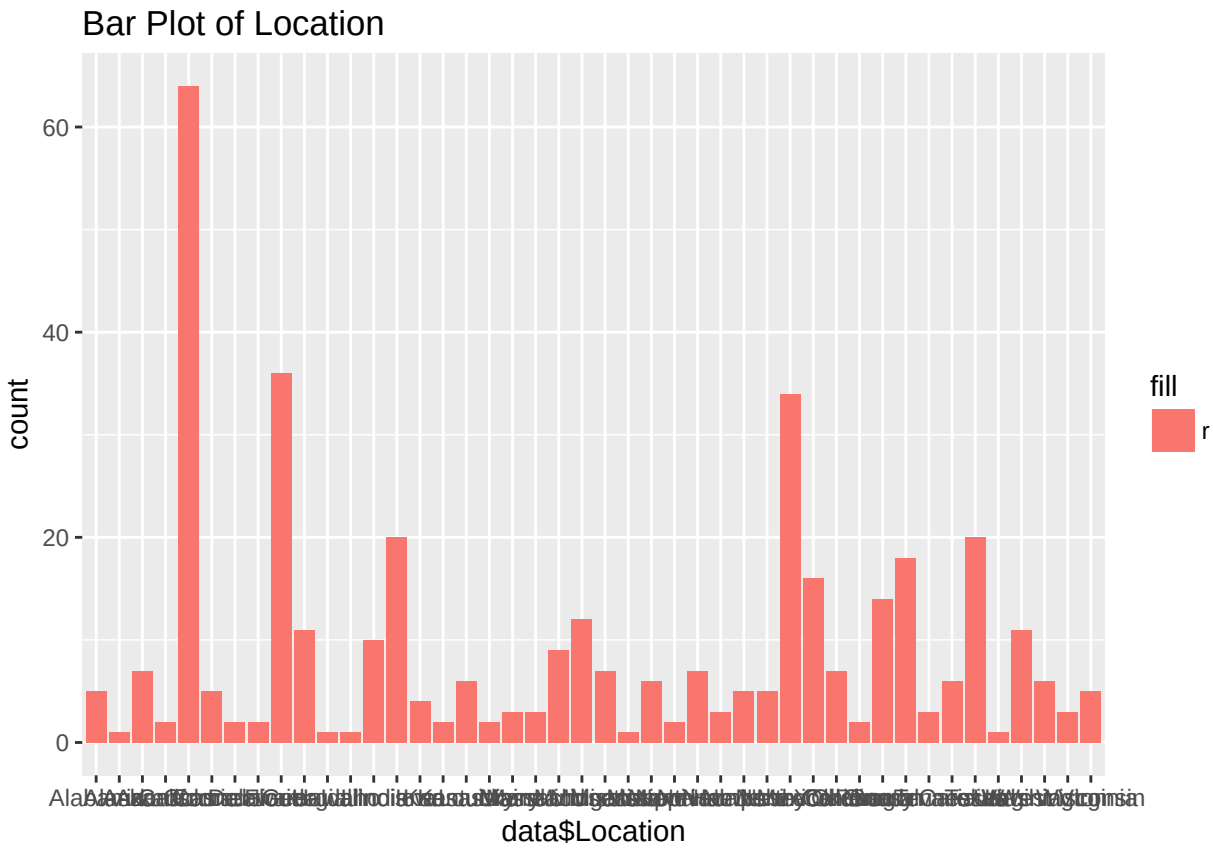
Exploratory Graphs

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



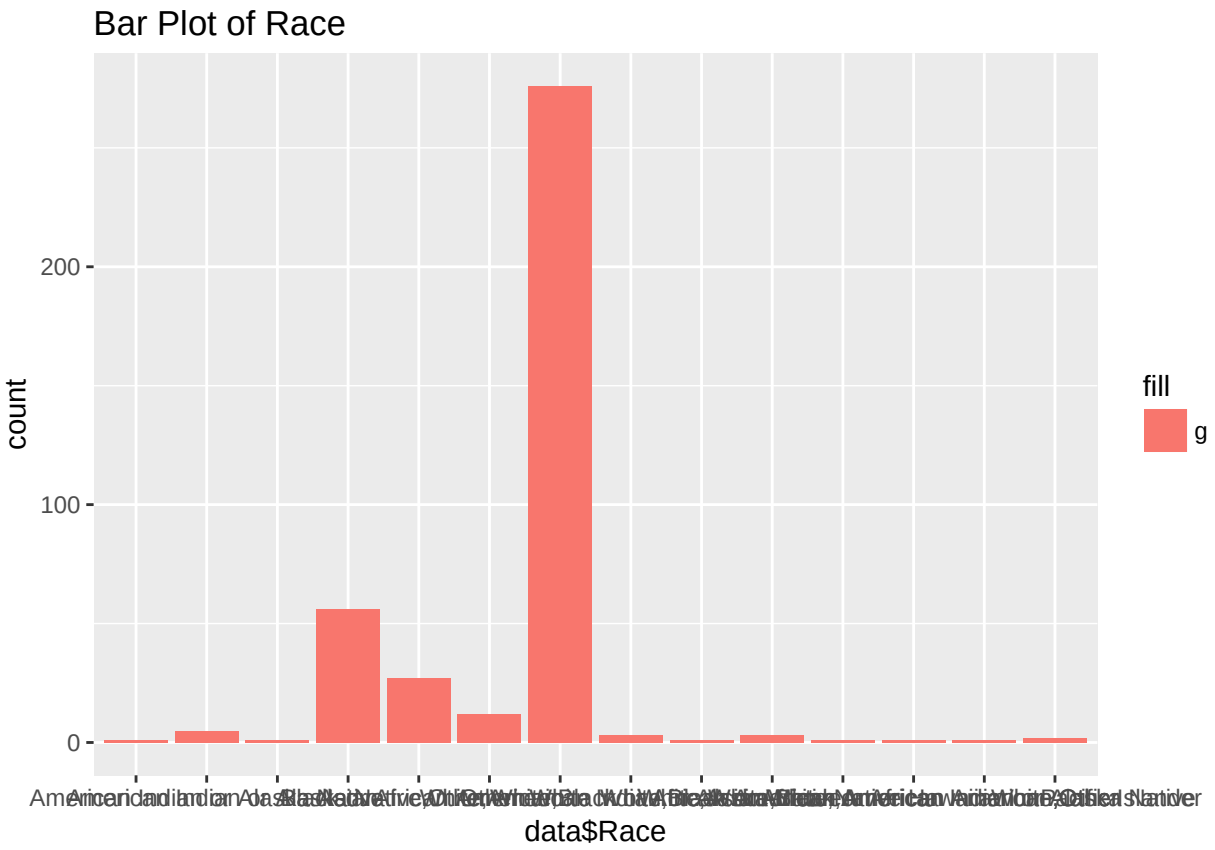
Not a lot of old people on Mechanical Turk, this is expected because I doubt retirees are filling out surveys to pay the bills. But our sample is certainly biased towards the 'millenials'.

```
#INSERT CODE TO ROTATE LOCATIONS
qplot(data$Location, geom = 'bar', fill = 'r', main = '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= 'g', main = '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
```

```
##  {"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
```

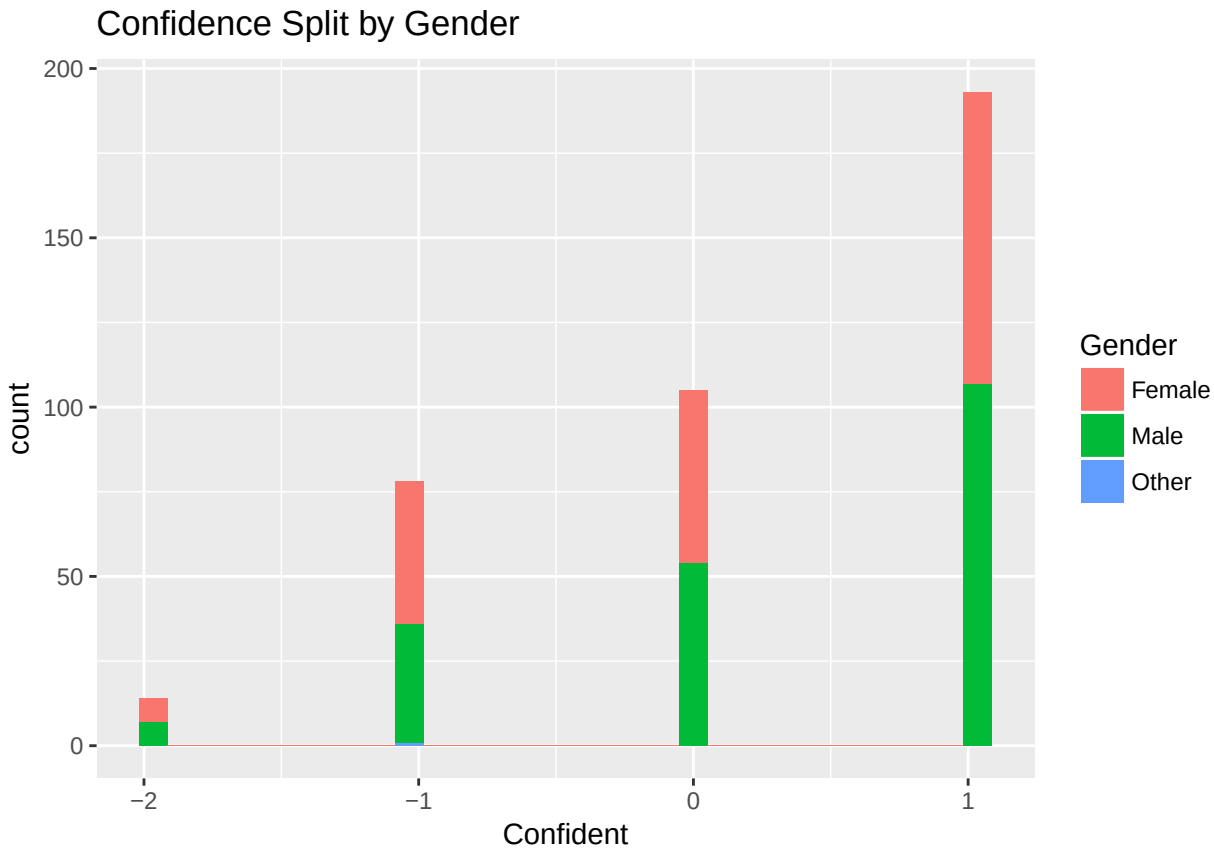
Finally, the relating to the model seem to half a real split in the data. This is probably due to the treatment.

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.

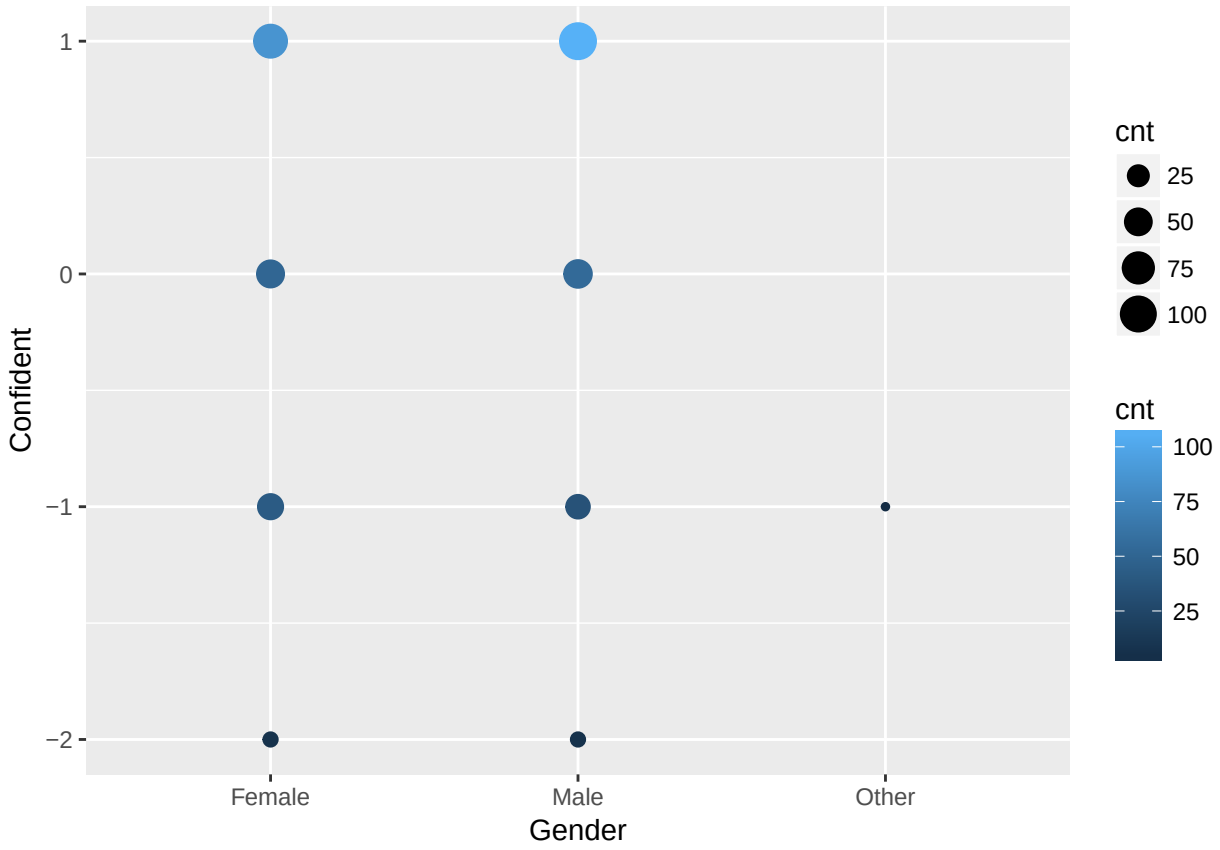
```
recode <- function(field){
  out <- 0
  if(field == 'Agree'){
    out <- 1
  } else if(field == 'Strongly agree'){
    out <- 2
  } else if(field == 'Disagree'){
    out <- -1
  } else if(field == 'Strongly disagree'){
    out <- -2
  }
  out <- as.numeric(out)
  return(out)
}


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)
```

```
qplot(data = data, x = Confident, fill = Gender, main = 'Confidence Split by Gender' )
```

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




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




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

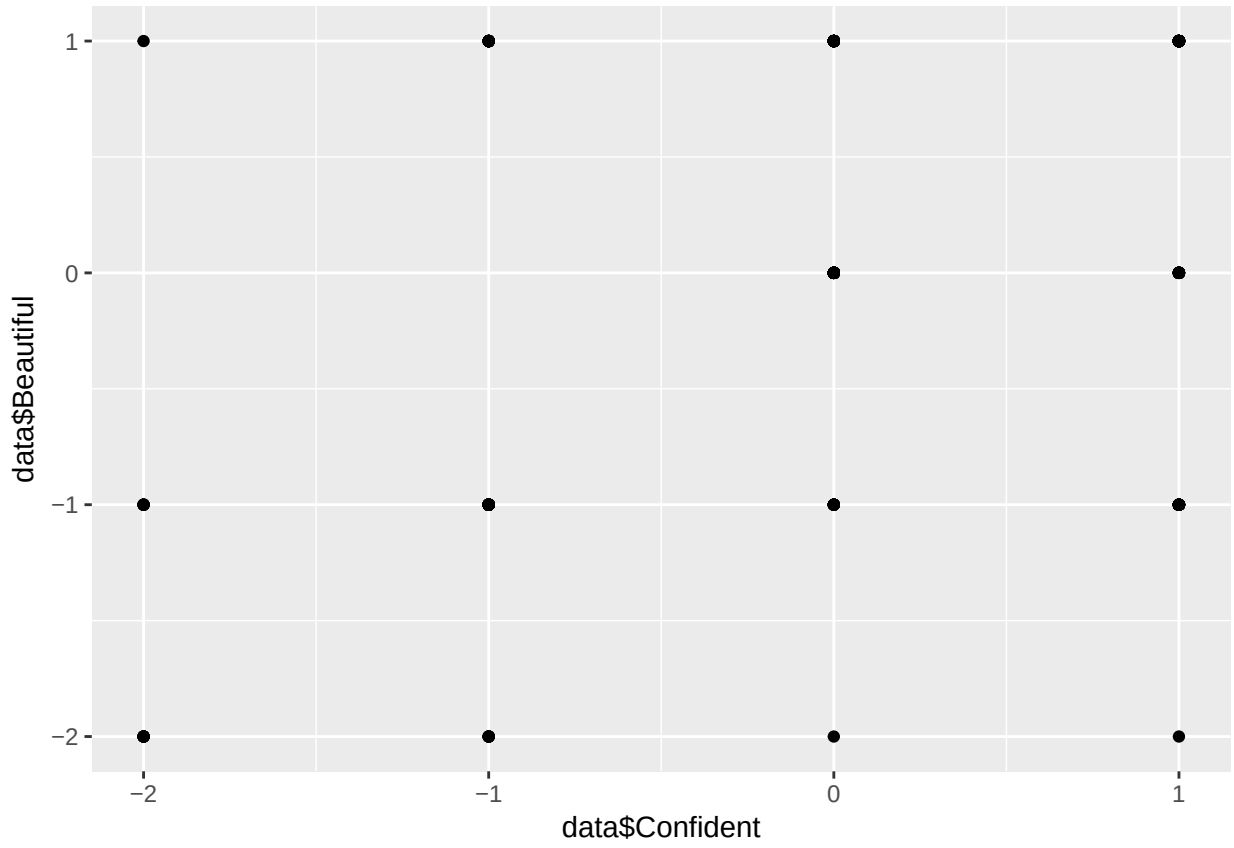
```
## [1] "Male Conf 0.285714285714286"
```

```
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.

```
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
```

It looks like our Confidence and Self Beauty are strongly correlated, which is not surprising.

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

```
##
##  Pearson's product-moment correlation
##
```

```
## data:  data$Relate and data$Beautiful
## t = 5.2758, df = 388, p-value = 2.203e-07
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1636207 0.3490539
## sample estimates:
##       cor
## 0.2587193
```

** Note (Olivier): I would stick to descriptive statistics and not cover here correlation, as at this point this could be influenced by the experiment**

The correlation between relating to the model and beauty is not nearly as strong as the correlation between beauty and confidence.

Randomization Inference

```
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 0.4567"
```

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

```
## [1] "Pval Confident 0.3732"
```

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

```
## [1] "Pval Relate 0.0155"
```

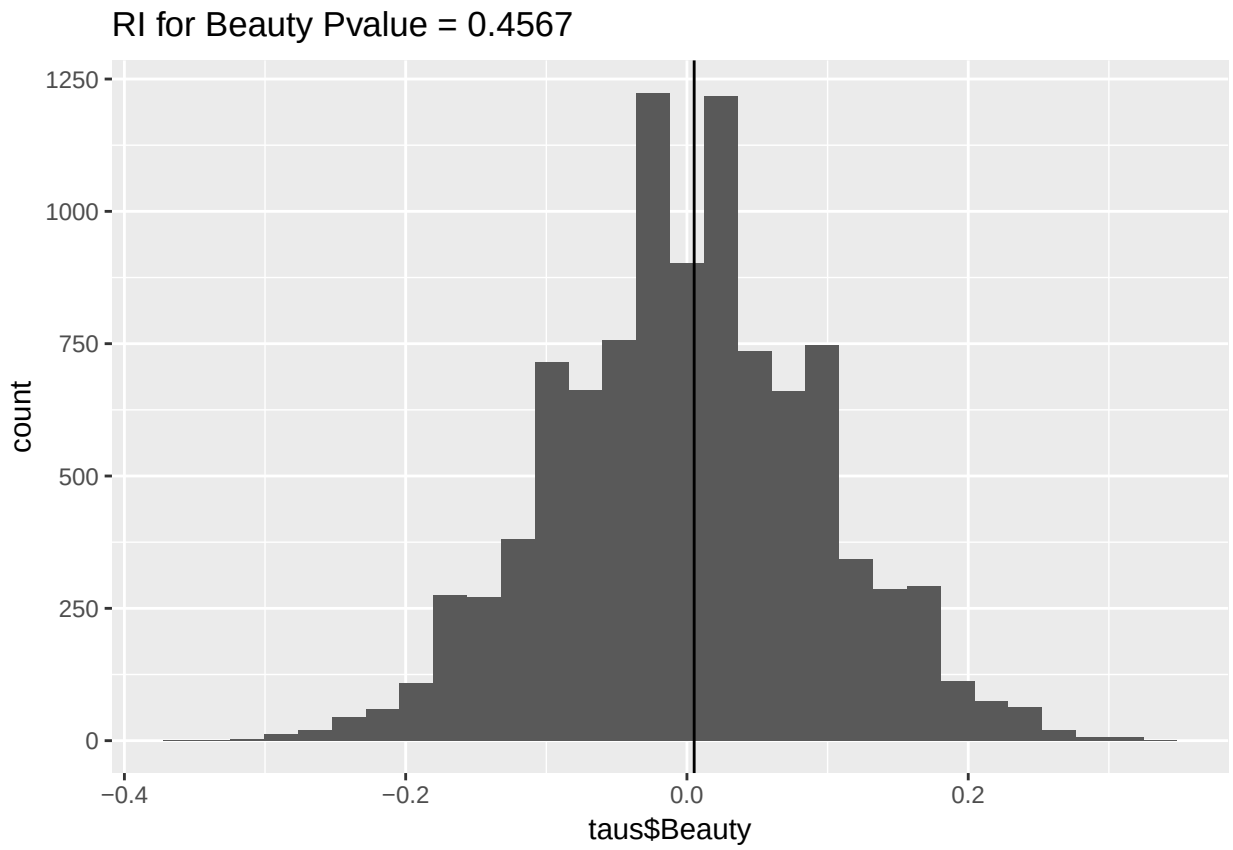
```
print(paste('Pval Important', p_importance))
```

```
## [1] "Pval Important 0.6774"
```

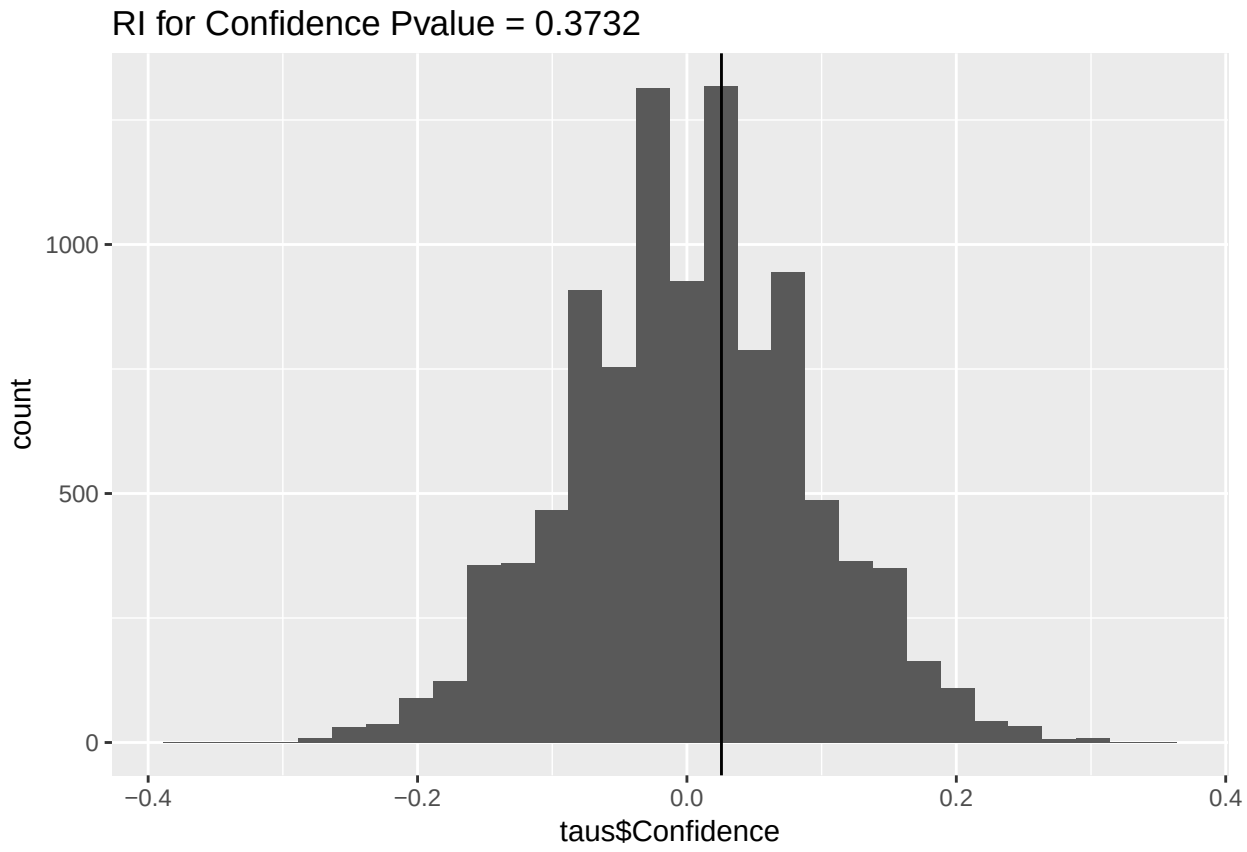
Our only variable that seems to have a real ATE is the relation.

RI Plots

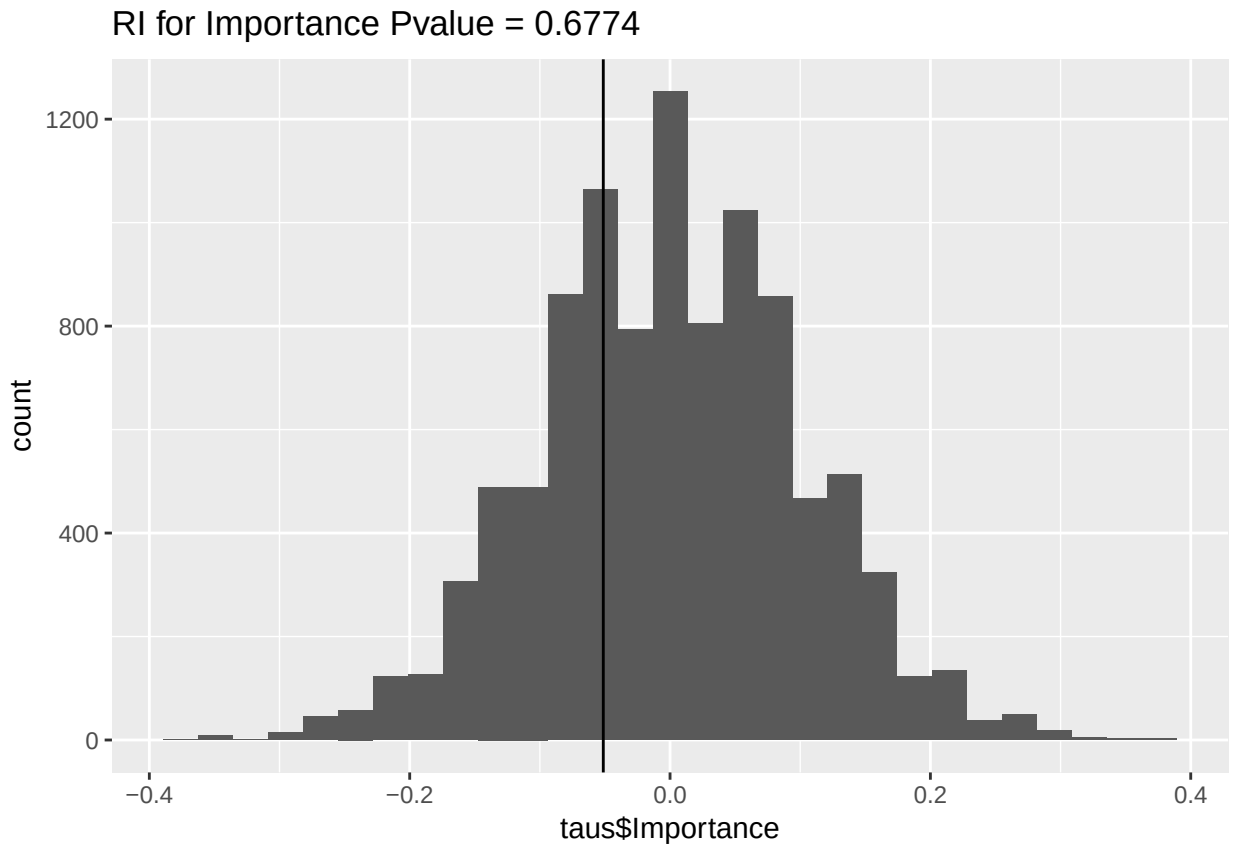
```
qplot(taus$Beauty, bins = 30, main = paste('RI for Beauty Pvalue =', p_beauty)) + geom_vline(xintercept
```



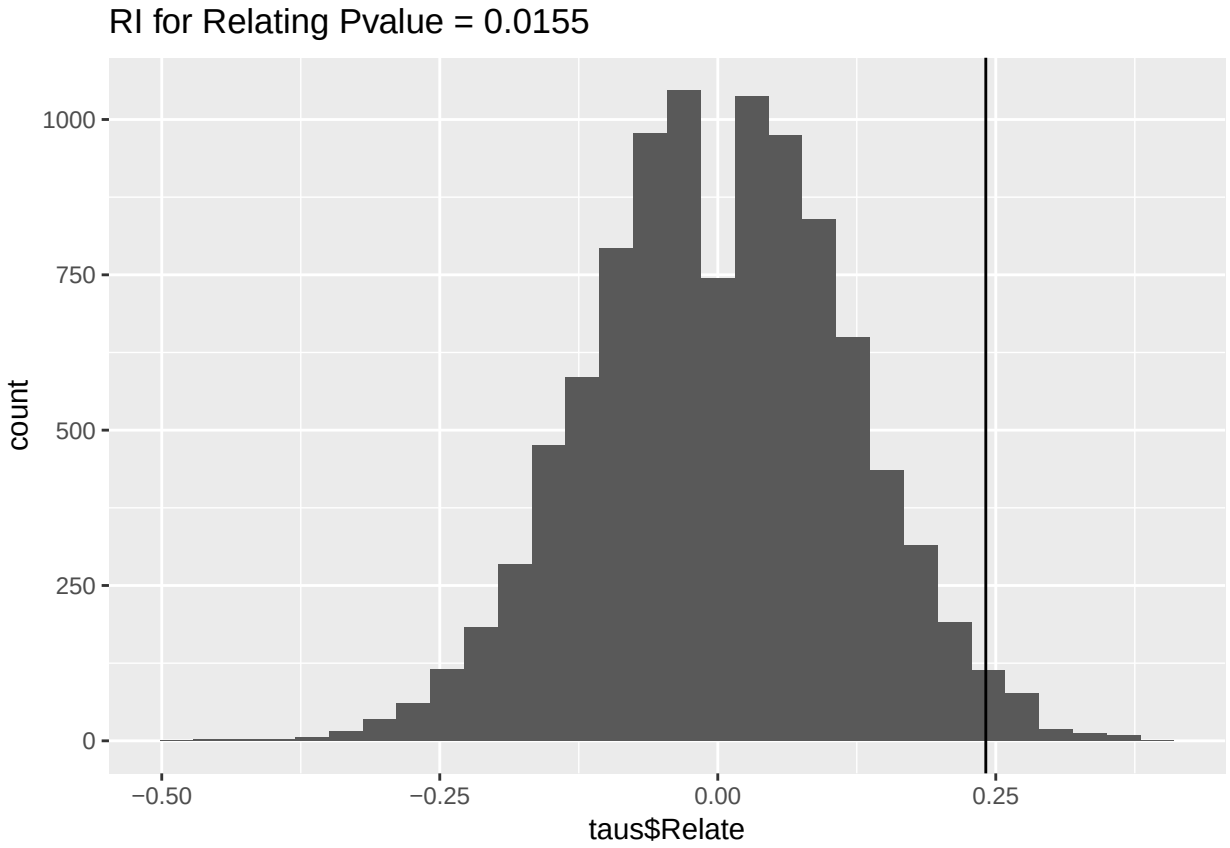

```
qplot(taus$Confidence, bins = 30, main = paste('RI for Confidence Pvalue =',p_confident)) + geom_vline(
```




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




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



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
```

```
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
```

```
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
```

```
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
```

```
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
```

```
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
```

```
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
```

```
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
```

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

```
##
## \begin{table}[!htbp] \centering
##   \caption{}
```

```
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{\textit{Dependent variable:}} \\
## \cline{2-5}
## \\[-1.8ex] & Beautiful & Confident & Importance & Relate \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  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) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Observations & 390 & 390 & 390 & 390 \\
## R$^{2}$ & 0.027 & 0.010 & 0.020 & 0.058 \\
## Adjusted R$^{2}$ & 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$^{***}$ \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

Image analysis

Preparing the data

```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

```
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(data[[column]] ~ data[['Group']])
  print(column)
  print(summary(l1))
  }
}
```

```
## [1] "Passion_i_identify_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3529 -0.3529 -0.2895  0.6471  0.7105
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.35294    0.03415  10.335   <2e-16 ***
## data[["Group"]]Treatment -0.06347    0.04811  -1.319    0.188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.467 on 375 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  0.00462,    Adjusted R-squared:  0.001966
## F-statistic: 1.741 on 1 and 375 DF,  p-value: 0.1879
##
## [1] "Passion_i_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3936 -0.3936 -0.2872  0.6064  0.7128
##
```

```
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.39362    0.03443  11.432   <2e-16 ***
## data[["Group"]]Treatment -0.10638    0.04869  -2.185   0.0295 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4721 on 374 degrees of freedom
##   (14 observations deleted due to missingness)
## Multiple R-squared:  0.0126, Adjusted R-squared:  0.009961
## F-statistic: 4.773 on 1 and 374 DF,  p-value: 0.02953
##
## [1] "Passion_o_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6064 -0.5895  0.3936  0.4105  0.4105
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.60638    0.03585  16.914   <2e-16 ***
## data[["Group"]]Treatment -0.01691    0.05057  -0.334    0.738
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4916 on 376 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.0002973,  Adjusted R-squared:  -0.002361
## F-statistic: 0.1118 on 1 and 376 DF,  p-value: 0.7383
##
## [1] "Passion_validate_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98936  0.01064  0.01064  0.02632  0.02632
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.989362   0.009842 100.525   <2e-16 ***
## data[["Group"]]Treatment -0.015677   0.013882  -1.129    0.259
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1349 on 376 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.003381,   Adjusted R-squared:  0.00073
## F-statistic: 1.275 on 1 and 376 DF,  p-value: 0.2595
##
```

```
## [1] "Coffee_i_identify_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1848 -0.1848 -0.1482 -0.1482  0.8518
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.18478    0.02749   6.723 6.74e-11 ***
## data[["Group"]]Treatment -0.03663    0.03861  -0.949    0.343
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3728 on 371 degrees of freedom
##   (17 observations deleted due to missingness)
## Multiple R-squared:  0.002421,   Adjusted R-squared:  -0.0002684
## F-statistic: 0.9002 on 1 and 371 DF,  p-value: 0.3433
##
## [1] "Coffee_i_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3027 -0.3027 -0.2526  0.6973  0.7474
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.30270    0.03295   9.187   <2e-16 ***
## data[["Group"]]Treatment -0.05007    0.04629  -1.082     0.28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4482 on 373 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  0.003127,   Adjusted R-squared:  0.0004542
## F-statistic:  1.17 on 1 and 373 DF,  p-value: 0.2801
##
## [1] "Coffee_o_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4842 -0.4842 -0.4216  0.5158  0.5784
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.42162    0.03663  11.512   <2e-16 ***
```

```
## data[["Group"]]Treatment  0.06259    0.05145   1.216    0.225
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4982 on 373 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  0.003951,   Adjusted R-squared:  0.001281
## F-statistic:  1.48 on 1 and 373 DF,  p-value: 0.2246
##
## [1] "Coffee_validate_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98919  0.01081  0.01081  0.02105  0.02105
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.989189   0.009242 107.031   <2e-16 ***
## data[["Group"]]Treatment -0.010242   0.012984  -0.789    0.431
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1257 on 373 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  0.001665,   Adjusted R-squared:  -0.001011
## F-statistic: 0.6222 on 1 and 373 DF,  p-value: 0.4307
##
## [1] "Couple_i_identify_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.2234 -0.2234 -0.2064 -0.2064  0.7936
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.20635    0.02995   6.890 2.37e-11 ***
## data[["Group"]]Treatment  0.01706    0.04241   0.402    0.688
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4117 on 375 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  0.0004311,  Adjusted R-squared:  -0.002234
## F-statistic: 0.1617 on 1 and 375 DF,  p-value: 0.6878
##
## [1] "Couple_i_prefer_"
##
## Call:
```

```
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3191 -0.3191 -0.3175  0.6808  0.6825
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.317460   0.033973   9.344   <2e-16 ***
## data[["Group"]]Treatment 0.001689   0.048110   0.035    0.972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4671 on 375 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  3.285e-06,  Adjusted R-squared:  -0.002663
## F-statistic: 0.001232 on 1 and 375 DF,  p-value: 0.972
##
## [1] "Couple_o_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5904 -0.5106  0.4096  0.4894  0.4894
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.51064    0.03626  14.083   <2e-16 ***
## data[["Group"]]Treatment  0.07979    0.05128   1.556    0.121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4972 on 374 degrees of freedom
##   (14 observations deleted due to missingness)
## Multiple R-squared:  0.006432,   Adjusted R-squared:  0.003775
## F-statistic: 2.421 on 1 and 374 DF,  p-value: 0.1206
##
## [1] "Couple_validate_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98930  0.01070  0.01070  0.02646  0.02646
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.973545   0.009842  98.922   <2e-16 ***
## data[["Group"]]Treatment 0.015760   0.013955   1.129    0.259
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.1353 on 374 degrees of freedom
##   (14 observations deleted due to missingness)
## Multiple R-squared:  0.003398,   Adjusted R-squared:  0.0007338
## F-statistic: 1.275 on 1 and 374 DF,  p-value: 0.2595
##
## [1] "Work_i_identify_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5027 -0.4628 -0.4628  0.4974  0.5372
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.46277    0.03651  12.674   <2e-16 ***
## data[["Group"]]Treatment  0.03988    0.05157   0.773     0.44
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5006 on 375 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  0.001592,   Adjusted R-squared:  -0.00107
## F-statistic: 0.598 on 1 and 375 DF,  p-value: 0.4398
##
## [1] "Work_i_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6738 -0.6649  0.3262  0.3351  0.3351
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.664894   0.034402  19.327   <2e-16 ***
## data[["Group"]]Treatment 0.008903   0.048716   0.183    0.855
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4717 on 373 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  8.954e-05,  Adjusted R-squared:  -0.002591
## F-statistic: 0.0334 on 1 and 373 DF,  p-value: 0.8551
##
## [1] "Work_o_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -0.7884  0.2116  0.2116  0.2394  0.2394
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.76064    0.03054  24.905   <2e-16 ***
## data[["Group"]]Treatment   0.02772    0.04314   0.643    0.521
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4188 on 375 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  0.0011, Adjusted R-squared:  -0.001564
## F-statistic: 0.413 on 1 and 375 DF,  p-value: 0.5208
##
## [1] "Work_validate_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98413  0.01587  0.01587  0.02128  0.02128
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.978723   0.009870  99.166   <2e-16 ***
## data[["Group"]]Treatment 0.005404   0.013939   0.388    0.698
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1353 on 375 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  0.0004006,  Adjusted R-squared:  -0.002265
## F-statistic: 0.1503 on 1 and 375 DF,  p-value: 0.6985
##
## [1] "Fit_i_identify_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4032 -0.4032 -0.3723  0.5968  0.6277
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.40323    0.03580  11.262   <2e-16 ***
## data[["Group"]]Treatment -0.03089    0.05050  -0.612    0.541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4883 on 372 degrees of freedom
##   (16 observations deleted due to missingness)
```

```
## Multiple R-squared:  0.001005,   Adjusted R-squared:  -0.001681
## F-statistic: 0.3741 on 1 and 372 DF,  p-value: 0.5412
##
## [1] "Fit_i_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3936 -0.3936 -0.3817  0.6064  0.6183
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.38172    0.03582  10.657   <2e-16 ***
## data[["Group"]]Treatment  0.01190    0.05052   0.235    0.814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4885 on 372 degrees of freedom
##   (16 observations deleted due to missingness)
## Multiple R-squared:  0.000149,   Adjusted R-squared:  -0.002539
## F-statistic: 0.05545 on 1 and 372 DF,  p-value: 0.814
##
## [1] "Fit_o_prefer_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5957 -0.5806  0.4043  0.4194  0.4194
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.58065    0.03618  16.049   <2e-16 ***
## data[["Group"]]Treatment  0.01510    0.05103   0.296    0.767
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4934 on 372 degrees of freedom
##   (16 observations deleted due to missingness)
## Multiple R-squared:  0.0002353,  Adjusted R-squared:  -0.002452
## F-statistic: 0.08756 on 1 and 372 DF,  p-value: 0.7675
##
## [1] "Fit_validate_"
##
## Call:
## lm(formula = data[[column]] ~ data[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98925  0.00000  0.00000  0.01075  0.01075
##
```

```
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.989247   0.005347 184.997   <2e-16 ***
## data[["Group"]]Treatment 0.010753   0.007542   1.426    0.155
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07293 on 372 degrees of freedom
##   (16 observations deleted due to missingness)
## Multiple R-squared:  0.005434,   Adjusted R-squared:  0.002761
## F-statistic: 2.033 on 1 and 372 DF,  p-value: 0.1548
```

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

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

```
## [1] "_i_identify_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   0.000   1.000   1.536   2.000   5.000      30
##
## Call:
## lm(formula = data[[question_average]] ~ data[["Group"]] + data$Gender)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7431 -1.3108 -0.3108  0.6892  3.6892
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                1.39162    0.12294  11.320   <2e-16 ***
## data[["Group"]]Treatment -0.08085    0.13884  -0.582    0.561
## data$GenderMale            0.35144    0.13911   2.526    0.012 *
## data$GenderOther         -0.39162    1.32082  -0.296    0.767
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 1.315 on 356 degrees of freedom
##   (30 observations deleted due to missingness)
## Multiple R-squared:  0.01888,    Adjusted R-squared:  0.01061
## F-statistic: 2.284 on 3 and 356 DF,  p-value: 0.07873
##
## [1] "_i_prefer_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   2.000   1.969   3.000   5.000      30
##
## Call:
## lm(formula = data[[question_average]] ~ data[["Group"]] + data$Gender)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3662 -0.6892 -0.2122  0.7878  3.4648
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.6892     0.1158  14.581  < 2e-16 ***
## data[["Group"]]Treatment   -0.1540     0.1313  -1.173    0.242
## data$GenderMale             0.6770     0.1316   5.145 4.44e-07 ***
## data$GenderOther           -0.6892     1.2493  -0.552    0.582
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.244 on 356 degrees of freedom
##   (30 observations deleted due to missingness)
## Multiple R-squared:  0.07362,    Adjusted R-squared:  0.06581
## F-statistic:  9.43 on 3 and 356 DF,  p-value: 5.19e-06
##
## [1] "_o_prefer_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   2.000   3.000   2.981   4.000   5.000      28
##
## Call:
## lm(formula = data[[question_average]] ~ data[["Group"]] + data$Gender)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3316 -0.7471  0.2529  0.8333  2.4177
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 2.5823     0.1298  19.889  < 2e-16 ***
## data[["Group"]]Treatment    0.1648     0.1469   1.122   0.2627
## data$GenderMale             0.5844     0.1472   3.971 8.65e-05 ***
## data$GenderOther            2.4177     1.4017   1.725   0.0854 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.396 on 358 degrees of freedom
##   (28 observations deleted due to missingness)
## Multiple R-squared:  0.05102,    Adjusted R-squared:  0.04306
```

```
## F-statistic: 6.415 on 3 and 358 DF,  p-value: 0.0003047
##
## [1] "_validate_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   4.000   5.000   5.000   4.994   5.000   5.000      28
##
## Call:
## lm(formula = data[[question_average]] ~ data[["Group"]] + data$Gender)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98941 -0.00041  0.00045  0.01059  0.01145
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 4.9995505  0.0069152 722.977   <2e-16 ***
## data[["Group"]]Treatment -0.0110004  0.0078252  -1.406    0.161
## data$GenderMale             0.0008560  0.0078409   0.109    0.913
## data$GenderOther            0.0004495  0.0746514   0.006    0.995
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07433 on 358 degrees of freedom
##   (28 observations deleted due to missingness)
## Multiple R-squared:  0.005528,   Adjusted R-squared:  -0.002806
## F-statistic: 0.6633 on 3 and 358 DF,  p-value: 0.5751
```

Second data set

Load, Clean & Recode the data

```
raw_data2 <- read.csv('/Users/ozimmer/GoogleDrive/berkeley/w241/BeatuyAd_CausalExperiment/Data/BeautyAds

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
```

```r
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')
for (column in columns_to_analyse){
  l1 <- lm(data2[[column]] ~ data2[['Group']] + data2[['Gender']])
  print(column)
  print(summary(l1))
  #print(coef(summary(l1))[2,])
}
```

```
## [1] "Beautiful"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]] + data2[["Gender"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2509 -1.2117  0.7491  0.7509  0.7883
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.213388   0.089946   2.372   0.0181 *
## data2[["Group"]]Treatment -0.001733   0.095144  -0.018   0.9855
## data2[["Gender"]]Male       0.037485   0.097259   0.385   0.7001
## data2[["Gender"]]Other      0.786612   0.960269   0.819   0.4132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.956 on 401 degrees of freedom
## Multiple R-squared:  0.00196,    Adjusted R-squared:  -0.005507
## F-statistic: 0.2625 on 3 and 401 DF,  p-value: 0.8524
##
## [1] "Confident"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]] + data2[["Gender"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3749 -0.3749  0.6251  0.6373  0.8063
```

```
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.19375    0.08330   2.326   0.0205 *
## data2[["Group"]]Treatment   0.01219    0.08812   0.138   0.8900
## data2[["Gender"]]Male       0.16895    0.09008   1.876   0.0614 .
## data2[["Gender"]]Other     -1.19375    0.88935  -1.342   0.1803
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8854 on 401 degrees of freedom
## Multiple R-squared:  0.01399,    Adjusted R-squared:  0.006615
## F-statistic: 1.897 on 3 and 401 DF,  p-value: 0.1295
##
## [1] "Importance"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]] + data2[["Gender"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.1840 -1.1004  0.8160  0.8211  0.8996
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.183984   0.091142   2.019   0.0442 *
## data2[["Group"]]Treatment -0.078506   0.096409  -0.814   0.4160
## data2[["Gender"]]Male     -0.005065   0.098552  -0.051   0.9590
## data2[["Gender"]]Other     0.816016   0.973034   0.839   0.4022
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9688 on 401 degrees of freedom
## Multiple R-squared:  0.003601,   Adjusted R-squared:  -0.003853
## F-statistic: 0.4831 on 3 and 401 DF,  p-value: 0.6942
##
## [1] "Relate"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]] + data2[["Gender"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8579 -0.5988 -0.3004  1.1421  1.6996
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 -0.4404     0.1032  -4.269 2.46e-05 ***
## data2[["Group"]]Treatment    0.2983     0.1091   2.734  0.00654 **
## data2[["Gender"]]Male       -0.2592     0.1116  -2.323  0.02066 *
## data2[["Gender"]]Other      -1.5596     1.1014  -1.416  0.15756
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.097 on 401 degrees of freedom
## Multiple R-squared:  0.03601,    Adjusted R-squared:  0.0288
## F-statistic: 4.994 on 3 and 401 DF,  p-value: 0.002066
```

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']])
  print(column)
  print(summary(l1))
  }
}
```

```
## [1] "Passion_i_identify_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3632 -0.3632 -0.3112  0.6368  0.6888
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.36316    0.03432  10.580   <2e-16 ***
## data2[["Group"]]Treatment -0.05193    0.04817  -1.078    0.282
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4731 on 384 degrees of freedom
##   (19 observations deleted due to missingness)
## Multiple R-squared:  0.003018,   Adjusted R-squared:  0.0004217
## F-statistic: 1.162 on 1 and 384 DF,  p-value: 0.2816
##
## [1] "Passion_i_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3790 -0.3790 -0.3214  0.6210  0.6786
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.37895    0.03462  10.945   <2e-16 ***
## data2[["Group"]]Treatment -0.05752    0.04859  -1.184    0.237
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4773 on 384 degrees of freedom
##   (19 observations deleted due to missingness)
## Multiple R-squared:  0.003636,   Adjusted R-squared:  0.001041
## F-statistic: 1.401 on 1 and 384 DF,  p-value: 0.2372
##
## [1] "Passion_o_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6842 -0.6020  0.3158  0.3980  0.3980
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.68421    0.03473  19.700   <2e-16 ***
## data2[["Group"]]Treatment -0.08217    0.04874  -1.686   0.0926 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4787 on 384 degrees of freedom
##   (19 observations deleted due to missingness)
## Multiple R-squared:  0.007347,   Adjusted R-squared:  0.004762
## F-statistic: 2.842 on 1 and 384 DF,  p-value: 0.09264
##
## [1] "Passion_validate_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98947  0.01053  0.01292  0.01531  0.01531
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.989474   0.008223 120.334   <2e-16 ***
## data2[["Group"]]Treatment -0.004780   0.011539  -0.414    0.679
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1133 on 384 degrees of freedom
##   (19 observations deleted due to missingness)
## Multiple R-squared:  0.0004466,  Adjusted R-squared:  -0.002156
## F-statistic: 0.1716 on 1 and 384 DF,  p-value: 0.6789
##
## [1] "Coffee_i_identify_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.2064 -0.2064 -0.1846 -0.1846  0.8154
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.20635    0.02890   7.140 4.75e-12 ***
## data2[["Group"]]Treatment -0.02173    0.04056  -0.536    0.592
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3973 on 382 degrees of freedom
##   (21 observations deleted due to missingness)
## Multiple R-squared:  0.0007512,  Adjusted R-squared:  -0.001865
## F-statistic: 0.2872 on 1 and 382 DF,  p-value: 0.5924
##
## [1] "Coffee_i_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3231 -0.3231 -0.2910  0.6769  0.7090
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.29101    0.03363   8.654   <2e-16 ***
## data2[["Group"]]Treatment  0.03207    0.04719   0.680    0.497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4623 on 382 degrees of freedom
##   (21 observations deleted due to missingness)
## Multiple R-squared:  0.001208,   Adjusted R-squared:  -0.001407
## F-statistic: 0.4619 on 1 and 382 DF,  p-value: 0.4971
##
## [1] "Coffee_o_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4815 -0.4815 -0.4615  0.5185  0.5385
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.48148    0.03640   13.23   <2e-16 ***
## data2[["Group"]]Treatment -0.01994    0.05108   -0.39    0.696
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5004 on 382 degrees of freedom
```

```
##   (21 observations deleted due to missingness)
## Multiple R-squared:  0.0003989,  Adjusted R-squared:  -0.002218
## F-statistic: 0.1525 on 1 and 382 DF,  p-value: 0.6964
##
## [1] "Coffee_validate_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.97436  0.02564  0.02564  0.02646  0.02646
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.973545   0.011615   83.82   <2e-16 ***
## data2[["Group"]]Treatment 0.000814   0.016299    0.05     0.96
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1597 on 382 degrees of freedom
##   (21 observations deleted due to missingness)
## Multiple R-squared:  6.529e-06,  Adjusted R-squared:  -0.002611
## F-statistic: 0.002494 on 1 and 382 DF,  p-value: 0.9602
##
## [1] "Couple_i_identify_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.2783 -0.2783 -0.2423  0.7217  0.7577
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.27835    0.03156   8.820   <2e-16 ***
## data2[["Group"]]Treatment -0.03608    0.04463  -0.808    0.419
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4396 on 386 degrees of freedom
##   (17 observations deleted due to missingness)
## Multiple R-squared:  0.00169,    Adjusted R-squared:  -0.0008959
## F-statistic: 0.6536 on 1 and 386 DF,  p-value: 0.4193
##
## [1] "Couple_i_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3454 -0.3454 -0.3402  0.6546  0.6598
```

```
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.345361   0.034165  10.109   <2e-16 ***
## data2[["Group"]]Treatment -0.005155   0.048316  -0.107    0.915
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4759 on 386 degrees of freedom
##   (17 observations deleted due to missingness)
## Multiple R-squared:  2.949e-05,  Adjusted R-squared:  -0.002561
## F-statistic: 0.01138 on 1 and 386 DF,  p-value: 0.9151
##
## [1] "Couple_o_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5722 -0.5051  0.4278  0.4949  0.4949
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.57216    0.03580  15.981   <2e-16 ***
## data2[["Group"]]Treatment -0.06701    0.05063  -1.323    0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4987 on 386 degrees of freedom
##   (17 observations deleted due to missingness)
## Multiple R-squared:  0.004517,   Adjusted R-squared:  0.001938
## F-statistic: 1.752 on 1 and 386 DF,  p-value: 0.1865
##
## [1] "Couple_validate_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.97938  0.02062  0.02062  0.04124  0.04124
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.95876    0.01244  77.074   <2e-16 ***
## data2[["Group"]]Treatment  0.02062    0.01759   1.172    0.242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1733 on 386 degrees of freedom
##   (17 observations deleted due to missingness)
## Multiple R-squared:  0.003546,   Adjusted R-squared:  0.0009646
## F-statistic: 1.374 on 1 and 386 DF,  p-value: 0.2419
```

```
##
## [1] "Work_i_identify_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4973 -0.4949 -0.4949  0.5051  0.5051
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.497326   0.036659  13.566   <2e-16 ***
## data2[["Group"]]Treatment -0.002481   0.051373  -0.048    0.962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5013 on 379 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  6.153e-06,  Adjusted R-squared:  -0.002632
## F-statistic: 0.002332 on 1 and 379 DF,  p-value: 0.9615
##
## [1] "Work_i_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6684 -0.6546  0.3316  0.3454  0.3454
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.66845    0.03469  19.267   <2e-16 ***
## data2[["Group"]]Treatment -0.01381    0.04862  -0.284    0.777
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4744 on 379 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  0.0002128,  Adjusted R-squared:  -0.002425
## F-statistic: 0.08068 on 1 and 379 DF,  p-value: 0.7765
##
## [1] "Work_o_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.7594 -0.7268  0.2406  0.2732  0.2732
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)                0.75936    0.03203  23.711   <2e-16 ***
## data2[["Group"]]Treatment -0.03255    0.04488  -0.725    0.469
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4379 on 379 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  0.001386,   Adjusted R-squared:  -0.001249
## F-statistic: 0.5261 on 1 and 379 DF,  p-value: 0.4687
##
## [1] "Work_validate_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98454  0.01546  0.01546  0.01604  0.01604
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.9839572  0.0091282 107.793   <2e-16 ***
## data2[["Group"]]Treatment 0.0005789  0.0127923   0.045    0.964
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1248 on 379 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  5.403e-06,  Adjusted R-squared:  -0.002633
## F-statistic: 0.002048 on 1 and 379 DF,  p-value: 0.9639
##
## [1] "Fit_i_identify_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4787 -0.4787 -0.4691  0.5213  0.5309
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.478723   0.036510  13.112   <2e-16 ***
## data2[["Group"]]Treatment -0.009651   0.051232  -0.188    0.851
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5006 on 380 degrees of freedom
##   (23 observations deleted due to missingness)
## Multiple R-squared:  9.338e-05,  Adjusted R-squared:  -0.002538
## F-statistic: 0.03549 on 1 and 380 DF,  p-value: 0.8507
##
## [1] "Fit_i_prefer_"
##
```

```
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5266 -0.4897  0.4734  0.4734  0.5103
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.52660    0.03653   14.41   <2e-16 ***
## data2[["Group"]]Treatment -0.03691    0.05126   -0.72    0.472
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5009 on 380 degrees of freedom
##   (23 observations deleted due to missingness)
## Multiple R-squared:  0.001362,   Adjusted R-squared:  -0.001266
## F-statistic: 0.5183 on 1 and 380 DF,  p-value: 0.472
##
## [1] "Fit_o_prefer_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6702 -0.6134  0.3298  0.3866  0.3866
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.67021    0.03501  19.144   <2e-16 ***
## data2[["Group"]]Treatment -0.05681    0.04913  -1.156    0.248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.48 on 380 degrees of freedom
##   (23 observations deleted due to missingness)
## Multiple R-squared:  0.003507,   Adjusted R-squared:  0.0008846
## F-statistic: 1.337 on 1 and 380 DF,  p-value: 0.2482
##
## [1] "Fit_validate_"
##
## Call:
## lm(formula = data2[[column]] ~ data2[["Group"]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98454  0.01546  0.01546  0.02128  0.02128
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.978723   0.009805  99.816   <2e-16 ***
## data2[["Group"]]Treatment 0.005813   0.013759   0.422    0.673
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1344 on 380 degrees of freedom
##   (23 observations deleted due to missingness)
## Multiple R-squared:  0.0004694,  Adjusted R-squared:  -0.002161
## F-statistic: 0.1785 on 1 and 380 DF,  p-value: 0.6729
```

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))
}
```

```
## [1] "_i_identify_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   1.000   1.735   3.000   5.000      35
##
## Call:
## lm(formula = data2[[question_average]] ~ data2[["Group"]] + data2$Gender)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9908 -0.8816 -0.2053  1.0092  3.4803
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  1.5197     0.1245  12.211  < 2e-16 ***
## data2[["Group"]]Treatment   -0.1092     0.1340  -0.815 0.415815
## data2$GenderMale             0.4711     0.1360   3.464 0.000594 ***
## data2$GenderOther           -1.5197     1.2928  -1.176 0.240557
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.287 on 366 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.03809,    Adjusted R-squared:  0.0302
## F-statistic: 4.831 on 3 and 366 DF,  p-value: 0.002606
```

```
##
## [1] "_i_prefer_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   2.000   2.168   3.000   5.000      35
##
## Call:
## lm(formula = data2[[question_average]] ~ data2[["Group"]] + data2$Gender)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4788 -0.8699 -0.3750  0.6250  3.2339
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  1.8699     0.1272  14.705  < 2e-16 ***
## data2[["Group"]]Treatment  -0.1039     0.1369  -0.759    0.449
## data2$GenderMale            0.6089     0.1389   4.383 1.53e-05 ***
## data2$GenderOther          -1.8699     1.3209  -1.416    0.158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.315 on 366 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.05794,    Adjusted R-squared:  0.05022
## F-statistic: 7.504 on 3 and 366 DF,  p-value: 6.923e-05
##
## [1] "_o_prefer_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   2.000   3.000   3.041   4.000   5.000      35
##
## Call:
## lm(formula = data2[[question_average]] ~ data2[["Group"]] + data2$Gender)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2934 -1.0120 -0.0263  0.9737  2.2551
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 3.01197    0.12575  23.952   <2e-16 ***
## data2[["Group"]]Treatment -0.26707    0.13539  -1.973   0.0493 *
## data2$GenderMale            0.28144    0.13739   2.049   0.0412 *
## data2$GenderOther          -0.01197    1.30625  -0.009   0.9927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.3 on 366 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.02163,    Adjusted R-squared:  0.01361
## F-statistic: 2.698 on 3 and 366 DF,  p-value: 0.04569
##
## [1] "_validate_average"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   4.000   5.000   5.000   4.986   5.000   5.000      35
```

```
##
## Call:
## lm(formula = data2[[question_average]] ~ data2[["Group"]] + data2$Gender)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.99898  0.00102  0.01138  0.01781  0.02817
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 4.97183    0.01119 444.427   <2e-16 ***
## data2[["Group"]]Treatment   0.01679    0.01204   1.394    0.164
## data2$GenderMale            0.01036    0.01222   0.847    0.397
## data2$GenderOther           0.02817    0.11621   0.242    0.809
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1157 on 366 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.00725,    Adjusted R-squared:  -0.0008878
## F-statistic: 0.8909 on 3 and 366 DF,  p-value: 0.4459
```