The Effects of Branded Coffee on Preceived Taste Satisfaction

Amy Zhang, Chris John, Jenna Farac, Simran Gill

2024-11-27

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

```
## Warning: package 'data.table' was built under R version 4.4.1
```

```
library(gridExtra)
```

```
## Warning: package 'gridExtra' was built under R version 4.4.1
```

```
library(stargazer)
```

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

```
##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
library(ggplot2)
library(lmtest)
```

```
## Warning: package 'lmtest' was built under R version 4.4.1
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 4.4.1
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:data.table':
##
##     yearmon, yearqtr
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
library(dplyr)
```

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

## The following object is masked from 'package:gridExtra':
##
##     combine

## The following objects are masked from 'package:data.table':
##
##     between, first, last

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

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

```r
library(car)
```

```
## Warning: package 'car' was built under R version 4.4.1

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.4.1

##
## Attaching package: 'car'

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

```r
set.seed(123)

control <- read.csv("Coffee Survey Control (Responses) - Form Responses 1.csv")
treatment <- read.csv("Coffee Survey Group II (Responses) - Form Responses 1.csv")

# rename column names for control
colnames(control) <- c('timestamp', 'good_and_gather_score', 'chameleon_score', 'age', 'gender','how_of
# rename column names for treatment
colnames(treatment) <- c('timestamp', 'name','good_and_gather_score', 'chameleon_score', 'age', 'gender
# reorder column names for treatment
treatment <- treatment[, c('timestamp', 'good_and_gather_score', 'chameleon_score', 'age', 'gender','hou

#--- Control ----

control$treatment <- 0
control$age <- as.integer(control$age)
```

```
## Warning: NAs introduced by coercion
```

```
# re-labeling gender
control <- control %>%
  mutate(gender = case_when(
    gender == "F" ~ "Female",
    gender == "M" ~ "Male",
    TRUE ~ "Unknown"
  ))

# removing rows where age is null
control<- control %>%
  filter(!is.na(age))

#--- Treatment ----

treatment$treatment <- 1
treatment$age <- as.integer(treatment$age)

# re-labeling gender
treatment <- treatment %>%
  mutate(gender = case_when(
    gender == "F" ~ "Female",
    gender == "M" ~ "Male",
    TRUE ~ "Unknown"
  ))

# removing rows where age is null
treatment<- treatment %>%
  filter(!is.na(age))
```

Balancing Control and Treatment

There are more participants in Control than in Treatment groups. To help create balance between the two groups, will perform random sampling to match the size of the treatment group.

```
print("Before Random Sampling:")
```

```
## [1] "Before Random Sampling:"
```

```
cat("Control size:", nrow(control))
```

```
## Control size: 51
```

```
cat("\nTreatment size:", nrow(treatment))
```

```
##
## Treatment size: 39
```

```r
# selecting smaller group size
#n_control <- nrow(control)
#min_size <- min(n_control, nrow(treatment))

#random sampling the control group
#control <- control[sample(1:n_control, min_size), ]

# combined data
d <- rbind(control, treatment)
table(d$treatment)
```

```
##
##  0  1
## 51 39
```

Organizing the rest of the data from dataset "d"

```r
# creating age groups
d$age_group <- cut(d$age,
                   breaks = c(0, 20, 30, 40, 50, Inf),
                   labels = c("Under 20", "20-30", "31-40", "41-50", "Over 50"),
                   right = FALSE)


# Convert how_often_drink_coffee to integer by factoring
d$how_often_drink_coffee <- factor(d$how_often_drink_coffee,
                                   levels = c("Never",
                                              "Occasionally (up to 1 time a week)",
                                              "Sometimes (a few times a week)",
                                              "Often (almost every day)",
                                              "Every day"))

# yes/no flag for if the participant is aware of the coffee brand at all
d$chameleon_awareness_flag <- ifelse(d$chameleon_awareness == "No", 0, 1)
d$good_and_gather_awareness_flag <- ifelse(d$good_and_gather_awareness == "No", 0, 1)

cat("\nNumber of Rows after cleaning:",nrow(d),"\n")
```

```
##
## Number of Rows after cleaning: 90
```

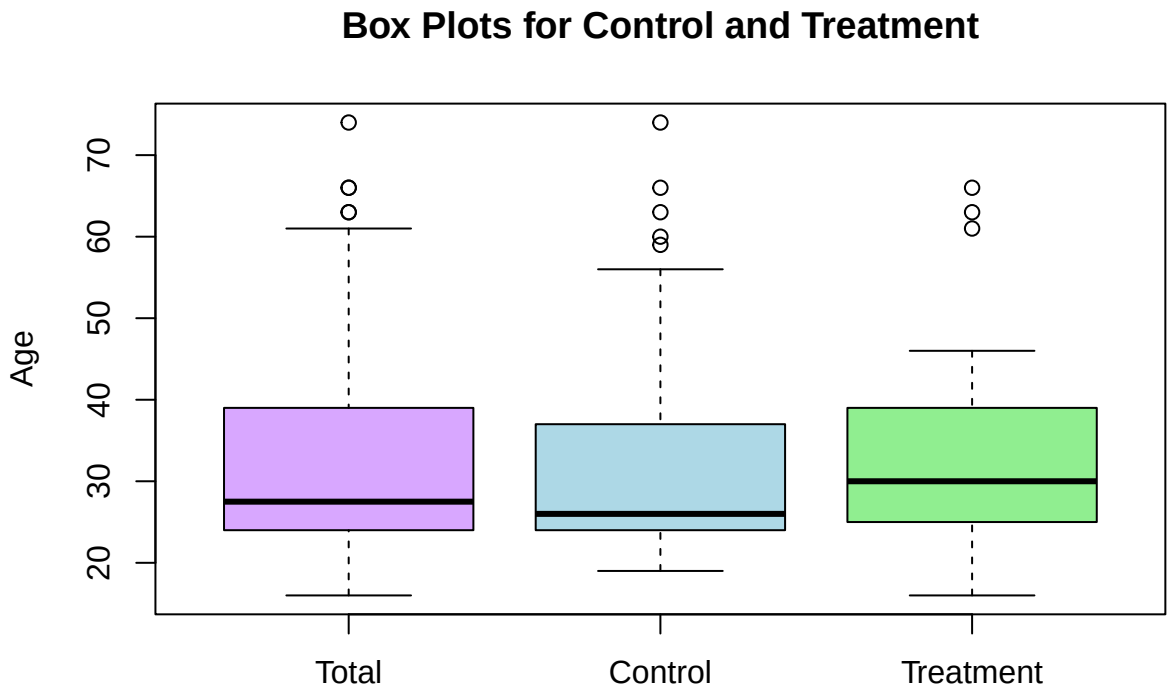
```r
str(d)
```

```
## 'data.frame':    90 obs. of  16 variables:
##  $ timestamp                    : chr  "11/11/2024 9:36:42" "11/11/2024 9:38:27" "11/11/2024 9:41:58
##  $ good_and_gather_score        : int  3 3 1 3 1 5 4 5 3 5 ...
##  $ chameleon_score              : int  5 5 5 5 3 4 2 3 5 5 ...
##  $ age                          : int  34 21 27 23 24 35 24 24 23 24 ...
##  $ gender                       : chr  "Male" "Male" "Female" "Female" ...
##  $ how_often_drink_coffee       : Factor w/ 5 levels "Never","Occasionally (up to 1 time a week)",..
```

```
##  $ hot_or_cold                   : chr  "Hot Coffee" "Hot Coffee" "Cold Coffee" "Hot Coffee" ...
##  $ sweet_or_not_sweet            : chr  "Not Sweet" "Sweet" "Not Sweet" "Not Sweet" ...
##  $ good_and_gather_awareness     : chr  "No" "Yes, Neutral" "Yes, Positive" "Yes, Positive" ...
##  $ chameleon_awareness           : chr  "No" "No" "Yes, Positive" "No" ...
##  $ medical_condition             : chr  "No" "No" "No" "No" ...
##  $ name                          : chr  "Kavin" "Arya Desai" "Liz Ren" "Halah Biviji" ...
##  $ treatment                     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ age_group                     : Factor w/ 5 levels "Under 20","20-30",..: 3 2 2 2 2 3 2 2 2 2 ...
##  $ chameleon_awareness_flag      : num  0 0 1 0 1 1 0 0 1 0 ...
##  $ good_and_gather_awareness_flag: num  0 1 1 1 0 1 0 0 1 1 ...
```

Exploratory Data Analysis

```
# box plot for age by treatment and control

boxplot(d$age, control$age, treatment$age,
        names = c("Total", "Control", "Treatment"),
        main = "Box Plots for Control and Treatment",
        ylab = "Age",
        col = c("#D8A7FF", "lightblue", "lightgreen"),
        border = "black")
```



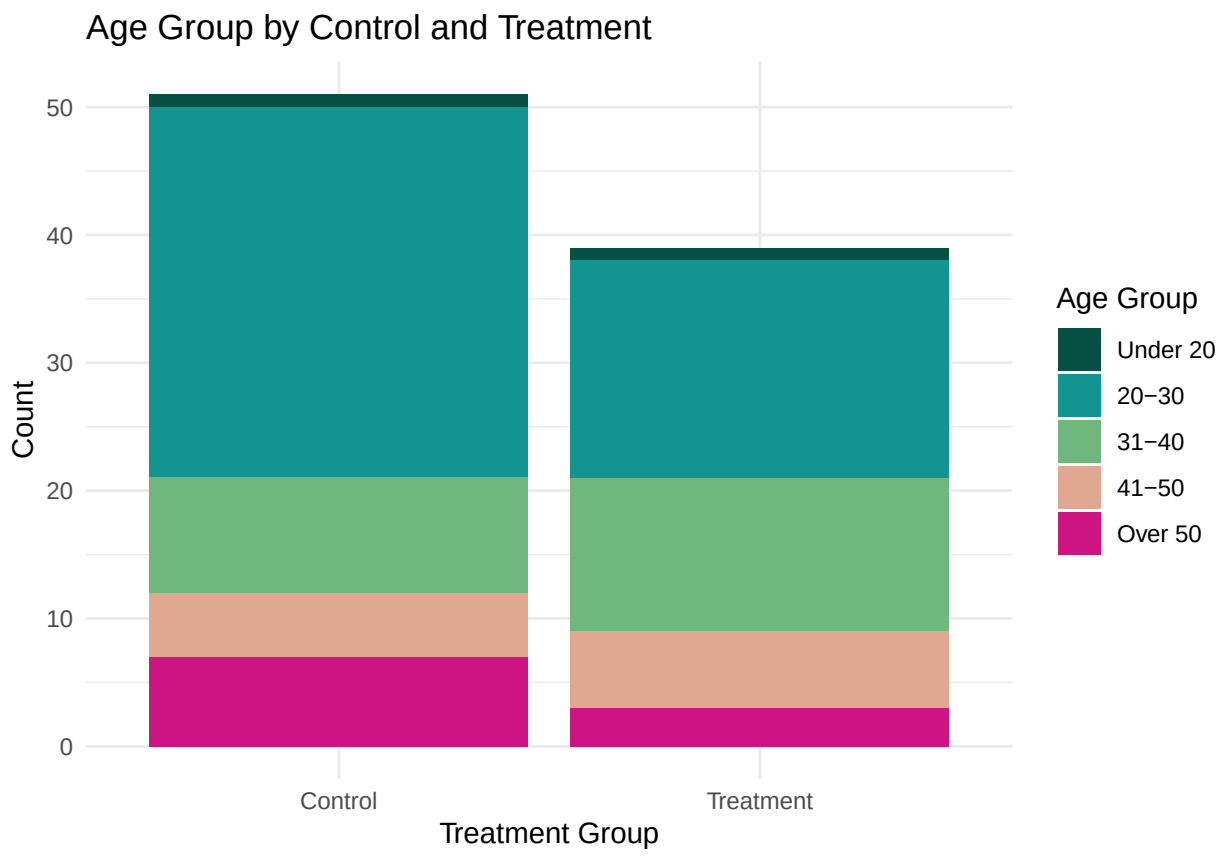
```r
cat("Number of Rows for Treatment Group:", sum(d$treatment == 1))
```

```
## Number of Rows for Treatment Group: 39
```

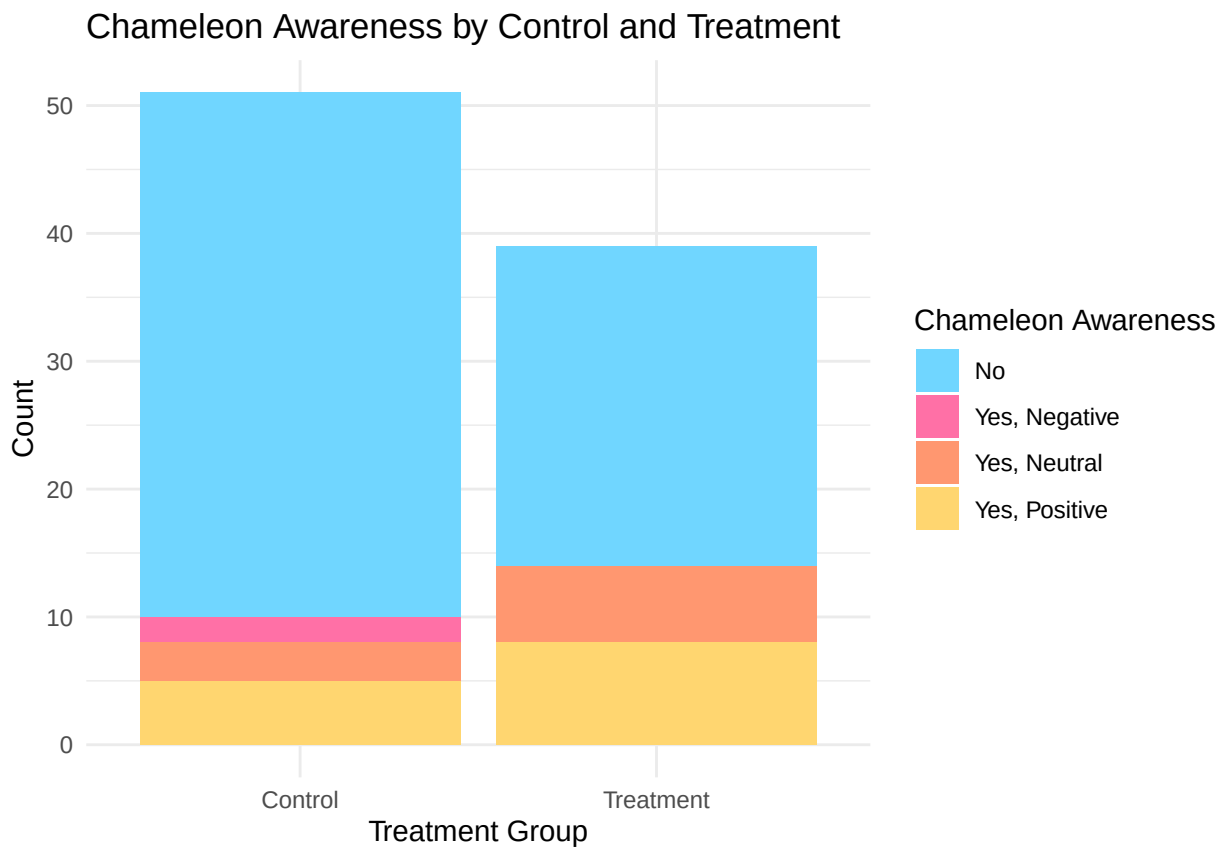
```r
cat("\nNumber of Rows for Control Group:", sum(d$treatment == 0))
```

```
##
## Number of Rows for Control Group: 51
```

```r
ggplot(d, aes(x = as.factor(treatment), fill = age_group)) +
  geom_bar(position = "stack") +
  labs(
    title = "Age Group by Control and Treatment",
    x = "Treatment Group",
    y = "Count"
  ) +
  scale_x_discrete(
    labels = c("0" = "Control", "1" = "Treatment")
  ) +
  scale_fill_manual(
    values = c("#065143", "#129490", "#70B77E", "#E0A890", "#CE1483"),
    name = "Age Group"
  ) +
  theme_minimal()
```

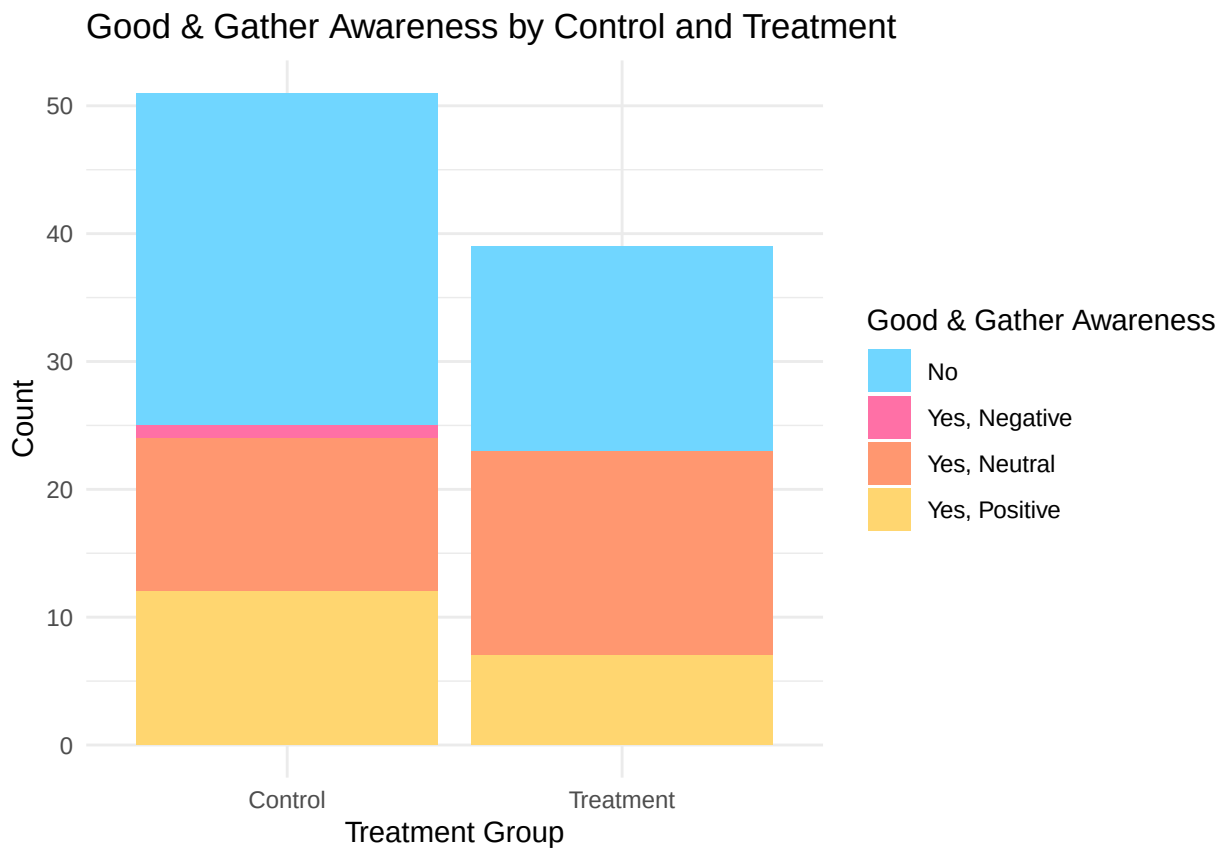


```r
ggplot(d, aes(x = as.factor(treatment), fill = chameleon_awareness)) +
  geom_bar(position = "stack") +
  labs(
    title = "Chameleon Awareness by Control and Treatment",
    x = "Treatment Group",
    y = "Count"
  ) +
  scale_x_discrete(
    labels = c("0" = "Control", "1" = "Treatment")
  ) +
  scale_fill_manual(
    values = c("#70D6FF", "#FF70A6", "#FF9770", "#FFD670"),
    name = "Chameleon Awareness"
  ) +
  theme_minimal()
```



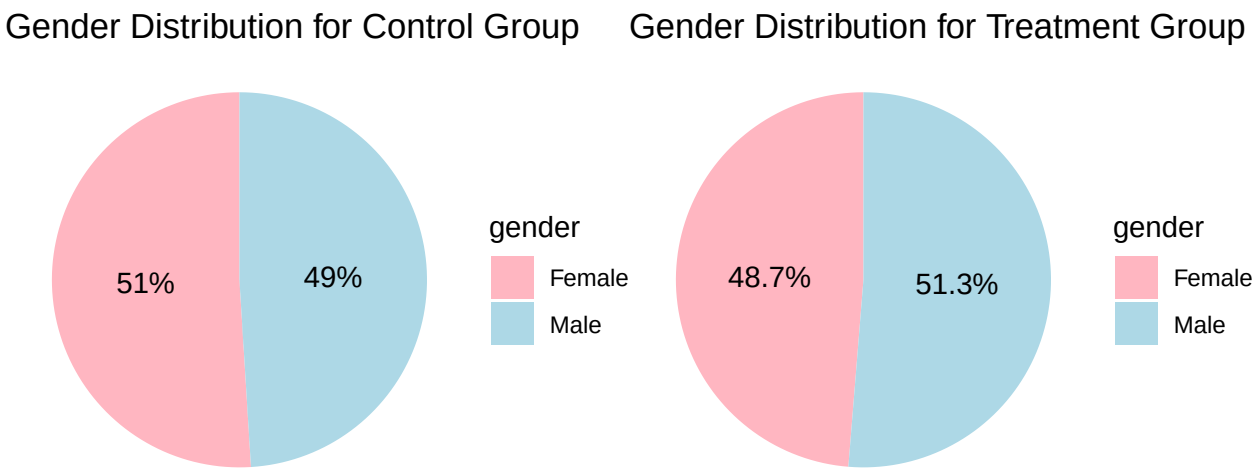

```r
ggplot(d, aes(x = as.factor(treatment), fill = good_and_gather_awareness)) +
  geom_bar(position = "stack") +
  labs(
    title = "Good & Gather Awareness by Control and Treatment",
    x = "Treatment Group",
    y = "Count"
  ) +
  scale_x_discrete(
    labels = c("0" = "Control", "1" = "Treatment")
```

```
  ) +
  scale_fill_manual(
    values = c("#70D6FF", "#FF70A6", "#FF9770", "#FFD670"),
    name = "Good & Gather Awareness"
  ) +
  theme_minimal()
```

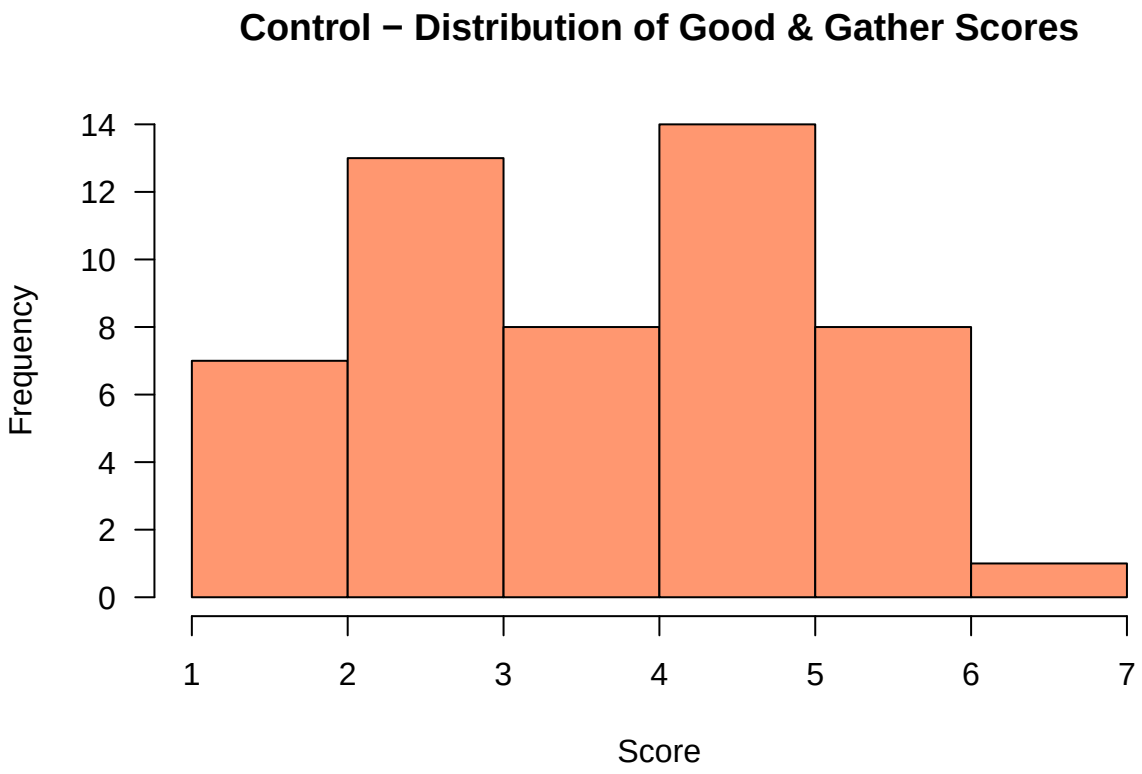



```
### Control Group Gender ###
control_gender_counts <- control %>%
  group_by(gender) %>%
  tally() %>%
  mutate(percentage = n / sum(n) * 100)

control_pie <- ggplot(control_gender_counts, aes(x = "", y = n, fill = gender)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar(theta = "y") +
  geom_text(aes(label = paste0(round(percentage, 1), "%")),
            position = position_stack(vjust = 0.5)) +
  labs(title = "Gender Distribution for Control Group") +
  scale_fill_manual(values = c("lightpink", "lightblue", "purple")) +
  theme_void()

### Treatment Group Gender ###
treatment_gender_counts <- treatment %>%
  group_by(gender) %>%
```

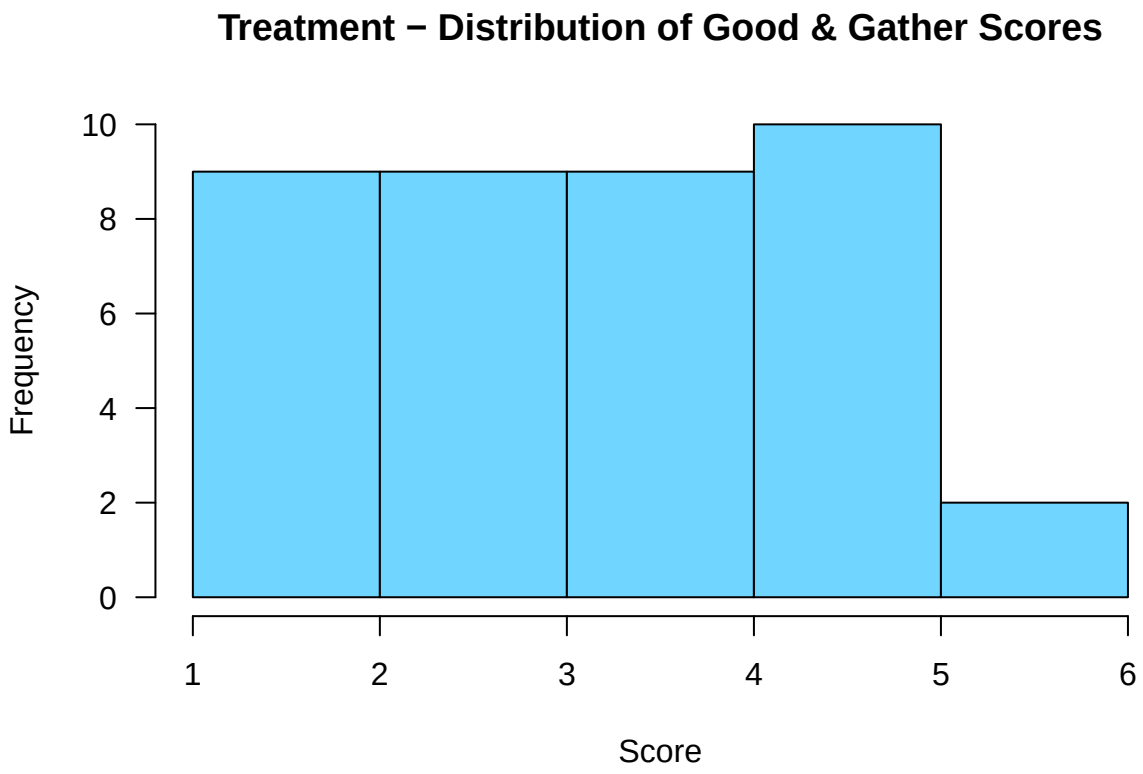
```
  tally() %>%
  mutate(percentage = n / sum(n) * 100)

treatment_pie <- ggplot(treatment_gender_counts, aes(x = "", y = n, fill = gender)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar(theta = "y") +
  geom_text(aes(label = paste0(round(percentage, 1), "%")),
            position = position_stack(vjust = 0.5)) +
  labs(title = "Gender Distribution for Treatment Group") +
  scale_fill_manual(values = c("lightpink", "lightblue", "purple")) +
  theme_void()

grid.arrange(control_pie, treatment_pie, ncol = 2)
```



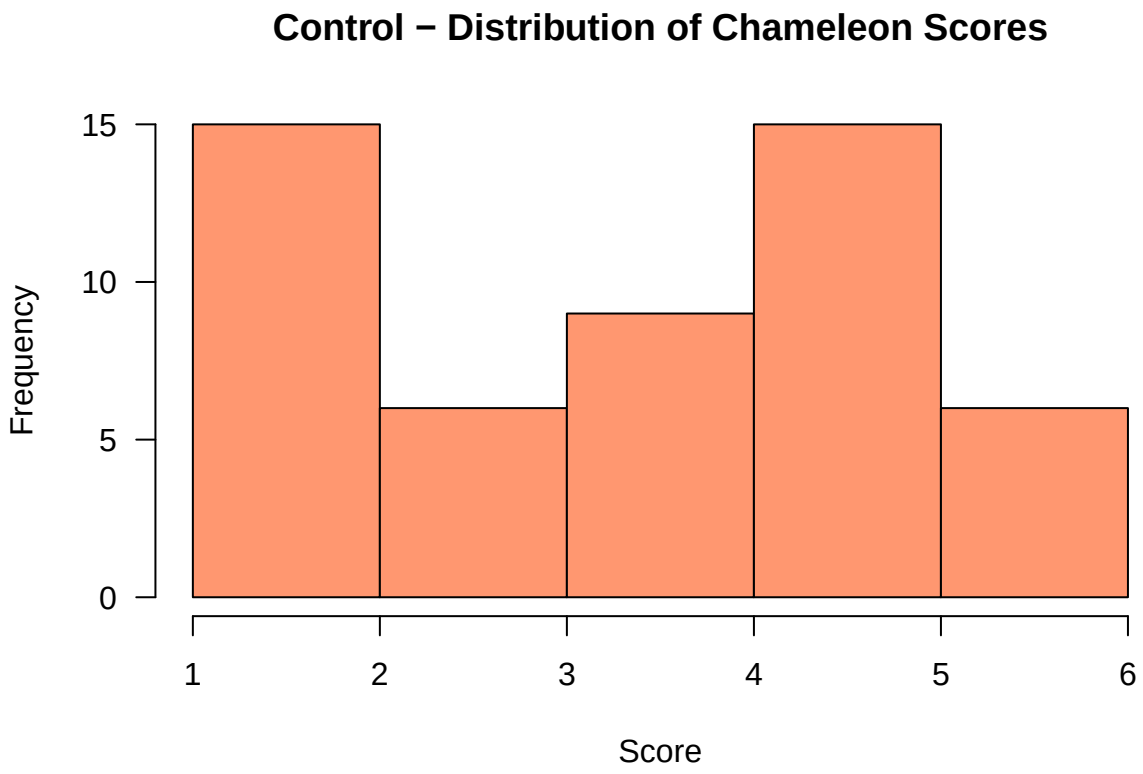

```
c <- d[d$treatment == 0, ]
t <- d[d$treatment == 1, ]

# control good and gather
hist(
  c$good_and_gather_score,
  col = '#FF9770',
  main = "Control - Distribution of Good & Gather Scores",
  ylab = "Frequency",
  xlab = "Score",
  las = 1 )
```



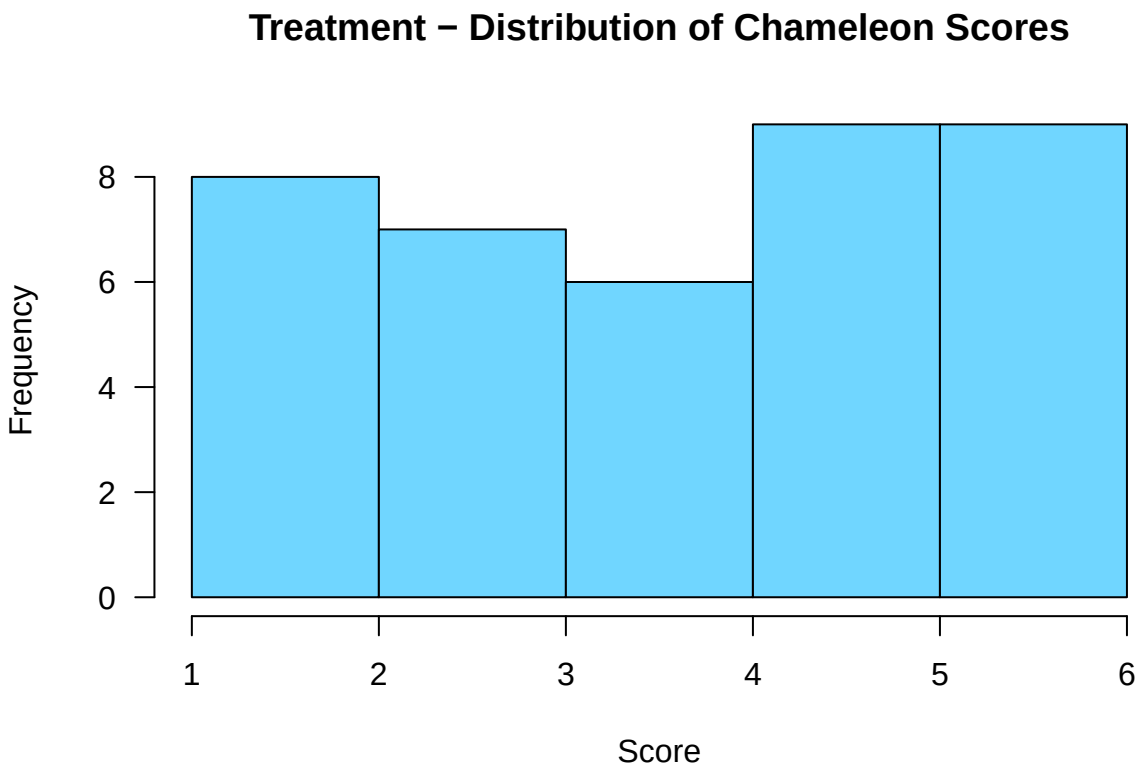

```r
# treatment good and gather
hist(
  t$good_and_gather_score,
  col = '#70D6FF',
  main = "Treatment - Distribution of Good & Gather Scores",
  ylab = "Frequency",
  xlab = "Score",
  las = 1 )
```




```r
# control chameleon
hist(
  c$chameleon_score,
  col = '#FF9770',
  main = "Control - Distribution of Chameleon Scores",
  ylab = "Frequency",
  xlab = "Score",
  las = 1 )
```




```r
# treatment chameleon
hist(
  t$chameleon_score,
  col = '#70D6FF',
  main = "Treatment - Distribution of Chameleon Scores",
  ylab = "Frequency",
  xlab = "Score",
  las = 1 )
```



Simple Average Treatment Effect

```
ate_good_and_gather <- mean(d$good_and_gather_score[d$treatment == 1], na.rm = TRUE) -
                       mean(d$good_and_gather_score[d$treatment == 0], na.rm = TRUE)
cat("ATE Good & Gather:", ate_good_and_gather)
```

```
## ATE Good & Gather: -0.4434389
```

```
ate_chameleon <- mean(d$chameleon_score[d$treatment == 1], na.rm = TRUE) -
                       mean(d$chameleon_score[d$treatment == 0], na.rm = TRUE)
cat("\nATE Chameleon:", ate_chameleon)
```

```
##
## ATE Chameleon: 0.2865762
```

Average Treatment Effect using Linear Regression

```
# Basic Linear regression to estimate ATE
model_gg <- lm(good_and_gather_score ~ treatment, data=d)
ate_regression <- coef(model_gg)["treatment"]
print(ate_regression)
```

```
##  treatment
## -0.4434389
```

```r
summary(model_gg)
```

```
##
## Call:
## lm(formula = good_and_gather_score ~ treatment, data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.05882 -1.05882 -0.05882  0.94118  2.94118
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.0588     0.1992  20.377   <2e-16 ***
## treatment    -0.4434     0.3026  -1.466    0.146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.422 on 88 degrees of freedom
## Multiple R-squared:  0.02383,    Adjusted R-squared:  0.01273
## F-statistic: 2.148 on 1 and 88 DF,  p-value: 0.1463
```

```r
stargazer(model_gg,
          type = "latex",   # Use "html" for HTML output or "latex" for LaTeX
          title = "Baseline Model for Good and Gather Score",
          covariate.labels = c("Treatment"),
          star.cutoffs = c(0.10, 0.05, 0.01, 0.001),   # Significance stars
          report = "vc*p")  # Optional: Save output to a text file
```

```
##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac a
## % Date and time: Wed, Dec 11, 2024 - 9:01:59 PM
## \begin{table}[!htbp] \centering
##   \caption{Baseline Model for Good and Gather Score}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{1}{c}{\textit{Dependent variable:}} \\
## \cline{2-2}
## \\[-1.8ex] & good\_and\_gather\_score \\
## \hline \\[-1.8ex]
##  Treatment & $-$0.443 \\
##   & p = 0.147 \\
##   & \\
##  Constant & 4.059$^{****}$ \\
##   & p = 0.000 \\
##   & \\
## \hline \\[-1.8ex]
## Observations & 90 \\
## R$^{2}$ & 0.024 \\
```

```
## Adjusted R$^{2}$ & 0.013 \\
## Residual Std. Error & 1.422 (df = 88) \\
## F Statistic & 2.148 (df = 1; 88) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{1}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

Table 1: Baseline Model for Good and Gather Score

	Dependent variable:
	good_and_gather_score
Treatment	−0.443
	p = 0.147
Constant	4.059****
	p = 0.000
Observations	90
R^2	0.024
Adjusted R^2	0.013
Residual Std. Error	1.422 (df = 88)
F Statistic	2.148 (df = 1; 88)
Note:	*p<0.1; **p<0.05; ***p<0.01

```r
# Basic Linear regression to estimate ATE
model_c <- lm(chameleon_score ~ treatment, data=d)
ate_regression <- coef(model_c)["treatment"]
print(ate_regression)
```

```
## treatment
## 0.2865762
```

```r
summary(model_c)
```

```
##
## Call:
## lm(formula = chameleon_score ~ treatment, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0513 -1.5863  0.2353  1.2353  2.2353
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.7647     0.2169   17.36   <2e-16 ***
## treatment     0.2866     0.3294    0.87    0.387
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.549 on 88 degrees of freedom
## Multiple R-squared:  0.008525,   Adjusted R-squared:  -0.002741
## F-statistic: 0.7567 on 1 and 88 DF,  p-value: 0.3867
```

ATE Adjusted for Covariates

c('timestamp', 'good_and_gather_score', 'chameleon_score', 'age', 'gender','how_often_drink_coffee', 'hot_or_cold', 'sweet_or_not_sweet', 'good_and_gather_awareness','chameleon_awareness', 'medical_condition', 'name')

```
model_gg_covariates <- lm(good_and_gather_score ~ treatment + log(age) + gender + chameleon_awareness,

ate_with_covariates <- coef(model_gg_covariates)["treatment"]
print(ate_with_covariates)
```

```
##  treatment
## -0.5684904
```

```
summary(model_gg_covariates)
```

```
##
## Call:
## lm(formula = good_and_gather_score ~ treatment + log(age) + gender +
##     chameleon_awareness, data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.10730 -0.82418 -0.00309  0.84285  2.56540
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       -0.5236     1.5124  -0.346  0.73005
## treatment                         -0.5685     0.2870  -1.981  0.05094 .
## log(age)                           1.3032     0.4280   3.045  0.00312 **
## genderMale                         0.3850     0.2764   1.393  0.16736
## chameleon_awarenessYes, Negative  -1.8873     0.9450  -1.997  0.04908 *
## chameleon_awarenessYes, Neutral    0.9324     0.4750   1.963  0.05298 .
## chameleon_awarenessYes, Positive  -0.6172     0.4014  -1.538  0.12792
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.301 on 83 degrees of freedom
## Multiple R-squared:  0.2293, Adjusted R-squared:  0.1736
## F-statistic: 4.116 on 6 and 83 DF,  p-value: 0.00115
```

```
vif(model_gg_covariates)
```

```
##                       GVIF Df GVIF^(1/(2*Df))
```

```
## treatment           1.074927  1        1.036787
## log(age)            1.044599  1        1.022056
## gender              1.015046  1        1.007495
## chameleon_awareness 1.107752  3        1.017202
```

```r
model_gg_covariates_v2 <- lm(good_and_gather_score ~ treatment + gender + log(age) + chameleon_awareness

anova(model_gg_covariates , model_gg_covariates_v2)
```

```
## Analysis of Variance Table
##
## Model 1: good_and_gather_score ~ treatment + log(age) + gender + chameleon_awareness
## Model 2: good_and_gather_score ~ treatment + gender + log(age) + chameleon_awareness +
##     good_and_gather_awareness
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     83 140.57
## 2     80 136.00  3    4.5703 0.8961  0.447
```

Interpretation We tested multiple covariates to see if we can improve the regression model for Good&Gather Score. The main covariates we see has a positive impact is how a participate views the Chameleon brand and age group.

When it comes to the Chameleon, even though the participants has a negative view of Chameleon coffee as a brand, they are still likely to score Good & Gather -1.8285 after treatment is provided. The p-value for Chameleon awareness is 0.0561, which means this variable is marginally significant.

We also wanted to test if adding Good&Gather brand awareness as a variable to model has an significant effect to the model. From the ANOVA test we can see that the p-value is 0.5600 which is greater than 0.05. This indicated Good&Gather brand awareness has no statistically significant impact on scoring the coffee.

```r
stargazer(model_gg, model_gg_covariates,
          type = "latex",    # Use "html" for HTML output or "latex" for LaTeX
          title = "Regression Results for Good and Gather Score",
          covariate.labels = c("Treatment", "log(Age)", "Gender",
                               "Chameleon Awareness (Negative)",
                               "Chameleon Awareness (Neutral)",
                               "Chameleon Awareness (Positive)"),
          star.cutoffs = c(0.10, 0.05, 0.01, 0.001),    # Significance stars
          report = "vc*p")  # Optional: Save output to a text file
```

```
##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac a
## % Date and time: Wed, Dec 11, 2024 - 9:02:00 PM
## \begin{table}[!htbp] \centering
##   \caption{Regression Results for Good and Gather Score}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{\textit{Dependent variable:}} \\
## \cline{2-3}
## \\[-1.8ex] & \multicolumn{2}{c}{good\_and\_gather\_score} \\
## \\[-1.8ex] & (1) & (2)\\
```

```
## \hline \\[-1.8ex]
##  Treatment & $-$0.443 & $-$0.568$^{*}$ \\
##   & p = 0.147 & p = 0.051 \\
##   & & \\
##  log(Age) &  & 1.303$^{***}$ \\
##   &  & p = 0.004 \\
##   & & \\
##  Gender &  & 0.385 \\
##   &  & p = 0.168 \\
##   & & \\
##  Chameleon Awareness (Negative) &  & $-$1.887$^{**}$ \\
##   &  & p = 0.050 \\
##   & & \\
##  Chameleon Awareness (Neutral) &  & 0.932$^{*}$ \\
##   &  & p = 0.053 \\
##   & & \\
##  Chameleon Awareness (Positive) &  & $-$0.617 \\
##   &  & p = 0.128 \\
##   & & \\
##  Constant & 4.059$^{****}$ & $-$0.524 \\
##   & p = 0.000 & p = 0.731 \\
##   & & \\
## \hline \\[-1.8ex]
## Observations & 90 & 90 \\
## R$^{2}$ & 0.024 & 0.229 \\
## Adjusted R$^{2}$ & 0.013 & 0.174 \\
## Residual Std. Error & 1.422 (df = 88) & 1.301 (df = 83) \\
## F Statistic & 2.148 (df = 1; 88) & 4.116$^{***}$ (df = 6; 83) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```r
model_c_covariates <- lm(chameleon_score ~ treatment + log(age) + gender + chameleon_awareness ,  data =

ate_with_covariates <- coef(model_c_covariates)["treatment"]
print(ate_with_covariates)
```

```
##  treatment
## 0.01849654
```

```r
summary(model_c_covariates)
```

```
##
## Call:
## lm(formula = chameleon_score ~ treatment + log(age) + gender +
##     chameleon_awareness, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.7951 -1.0407  0.1036  1.2510  2.4980
```

Table 2: Regression Results for Good and Gather Score

	Dependent variable:	
	good_and_gather_score	
	(1)	(2)
Treatment	−0.443	−0.568*
	p = 0.147	p = 0.051
log(Age)		1.303***
		p = 0.004
Gender		0.385
		p = 0.168
Chameleon Awareness (Negative)		−1.887**
		p = 0.050
Chameleon Awareness (Neutral)		0.932*
		p = 0.053
Chameleon Awareness (Positive)		−0.617
		p = 0.128
Constant	4.059****	−0.524
	p = 0.000	p = 0.731
Observations	90	90
R^2	0.024	0.229
Adjusted R^2	0.013	0.174
Residual Std. Error	1.422 (df = 88)	1.301 (df = 83)
F Statistic	2.148 (df = 1; 88)	4.116*** (df = 6; 83)
Note:		*p<0.1; **p<0.05; ***p<0.01

```
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      -0.7601     1.6373  -0.464 0.643693
## treatment                         0.0185     0.3107   0.060 0.952674
## log(age)                          1.1962     0.4634   2.582 0.011589 *
## genderMale                        0.2760     0.2992   0.922 0.359081
## chameleon_awarenessYes, Negative  1.2500     1.0230   1.222 0.225208
## chameleon_awarenessYes, Neutral   1.1921     0.5142   2.319 0.022881 *
## chameleon_awarenessYes, Positive  1.6751     0.4345   3.855 0.000227 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.409 on 83 degrees of freedom
## Multiple R-squared:  0.2261, Adjusted R-squared:  0.1702
## F-statistic: 4.043 on 6 and 83 DF,  p-value: 0.001329
```

```r
stargazer(model_c,
          type = "latex",   # Use "html" for HTML output or "latex" for LaTeX
          title = "Baseline Model for Chameleon Score",
          covariate.labels = c("Treatment"),
          star.cutoffs = c(0.10, 0.05, 0.01, 0.001),   # Significance stars
          report = "vc*p")  # Optional: Save output to a text file
```

```
##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac a
## % Date and time: Wed, Dec 11, 2024 - 9:02:00 PM
## \begin{table}[!htbp] \centering
##   \caption{Baseline Model for Chameleon Score}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{1}{c}{\textit{Dependent variable:}} \\
## \cline{2-2}
## \\[-1.8ex] & chameleon\_score \\
## \hline \\[-1.8ex]
##  Treatment & 0.287 \\
##   & p = 0.387 \\
##   & \\
##  Constant & 3.765$^{****}$ \\
##   & p = 0.000 \\
##   & \\
## \hline \\[-1.8ex]
## Observations & 90 \\
## R$^{2}$ & 0.009 \\
## Adjusted R$^{2}$ & $-$0.003 \\
## Residual Std. Error & 1.549 (df = 88) \\
## F Statistic & 0.757 (df = 1; 88) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{1}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

Table 3: Baseline Model for Chameleon Score

	Dependent variable:
	chameleon_score
Treatment	0.287
	p = 0.387
Constant	3.765****
	p = 0.000
Observations	90
R^2	0.009
Adjusted R^2	−0.003
Residual Std. Error	1.549 (df = 88)
F Statistic	0.757 (df = 1; 88)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
model_c_covariates_v2 <- lm(chameleon_score ~ treatment + log(age) + gender + chameleon_awareness + good

anova(model_c_covariates , model_c_covariates_v2)
```

```
## Analysis of Variance Table
##
## Model 1: chameleon_score ~ treatment + log(age) + gender + chameleon_awareness
## Model 2: chameleon_score ~ treatment + log(age) + gender + chameleon_awareness +
##     good_and_gather_awareness
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     83 164.74
## 2     80 156.86  3    7.8875 1.3409  0.267
```

```
stargazer(model_c, model_c_covariates,
          type = "latex",    # Use "html" for HTML output or "latex" for LaTeX
          title = "Regression Results for Chameleon Score",
          covariate.labels = c("Treatment", "log(Age)", "Gender",
                               "Chameleon Awareness (Negative)",
                               "Chameleon Awareness (Neutral)",
                               "Chameleon Awareness (Positive)"),  # Manually add all factor levels
          star.cutoffs = c(0.10, 0.05, 0.01, 0.001),
          report = "vc*p")  # Optional: Save output to a text file
```

```
##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac a
## % Date and time: Wed, Dec 11, 2024 - 9:02:00 PM
## \begin{table}[!htbp] \centering
##   \caption{Regression Results for Chameleon Score}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{\textit{Dependent variable:}} \\
```

```
## \cline{2-3}
## \\[-1.8ex] & \multicolumn{2}{c}{chameleon\_score} \\
## \\[-1.8ex] & (1) & (2)\\
## \hline \\[-1.8ex]
##  Treatment & 0.287 & 0.018 \\
##   & p = 0.387 & p = 0.953 \\
##   & & \\
##  log(Age) &  & 1.196$^{**}$ \\
##   &  & p = 0.012 \\
##   & & \\
##  Gender &  & 0.276 \\
##   &  & p = 0.360 \\
##   & & \\
##  Chameleon Awareness (Negative) &  & 1.250 \\
##   &  & p = 0.226 \\
##   & & \\
##  Chameleon Awareness (Neutral) &  & 1.192$^{**}$ \\
##   &  & p = 0.023 \\
##   & & \\
##  Chameleon Awareness (Positive) &  & 1.675$^{****}$ \\
##   &  & p = 0.0003 \\
##   & & \\
##  Constant & 3.765$^{****}$ & $-$0.760 \\
##   & p = 0.000 & p = 0.644 \\
##   & & \\
## \hline \\[-1.8ex]
## Observations & 90 & 90 \\
## R$^{2}$ & 0.009 & 0.226 \\
## Adjusted R$^{2}$ & $-$0.003 & 0.170 \\
## Residual Std. Error & 1.549 (df = 88) & 1.409 (df = 83) \\
## F Statistic & 0.757 (df = 1; 88) & 4.043$^{***}$ (df = 6; 83) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 4.4.1
```

```
library(dplyr)

reshaped_data <- d %>%
  pivot_longer(
    cols = c(good_and_gather_score, chameleon_score), # Columns to pivot
    names_to = "brand",                               # New column for brand
    values_to = "score"                               # New column for scores
  ) %>%
  mutate(brand = case_when(
    brand == "good_and_gather_score" ~ "Good and Gather",
    brand == "chameleon_score" ~ "Chameleon"
```

Table 4: Regression Results for Chameleon Score

	Dependent variable:	
	chameleon_score	
	(1)	(2)
Treatment	0.287	0.018
	p = 0.387	p = 0.953
log(Age)		1.196**
		p = 0.012
Gender		0.276
		p = 0.360
Chameleon Awareness (Negative)		1.250
		p = 0.226
Chameleon Awareness (Neutral)		1.192**
		p = 0.023
Chameleon Awareness (Positive)		1.675****
		p = 0.0003
Constant	3.765****	−0.760
	p = 0.000	p = 0.644
Observations	90	90
R^2	0.009	0.226
Adjusted R^2	−0.003	0.170
Residual Std. Error	1.549 (df = 88)	1.409 (df = 83)
F Statistic	0.757 (df = 1; 88)	4.043*** (df = 6; 83)
Note:		*p<0.1; **p<0.05; ***p<0.01

```
  ))

# View the reshaped data
head(reshaped_data)
```

```
## # A tibble: 6 x 16
##   timestamp     age gender how_often_drink_coffee hot_or_cold sweet_or_not_sweet
##   <chr>       <int> <chr>  <fct>                  <chr>       <chr>
## 1 11/11/2024~    34 Male   Often (almost every d~ Hot Coffee  Not Sweet
## 2 11/11/2024~    34 Male   Often (almost every d~ Hot Coffee  Not Sweet
## 3 11/11/2024~    21 Male   Often (almost every d~ Hot Coffee  Sweet
## 4 11/11/2024~    21 Male   Often (almost every d~ Hot Coffee  Sweet
## 5 11/11/2024~    27 Female Often (almost every d~ Cold Coffee Not Sweet
## 6 11/11/2024~    27 Female Often (almost every d~ Cold Coffee Not Sweet
## # i 10 more variables: good_and_gather_awareness <chr>,
## #   chameleon_awareness <chr>, medical_condition <chr>, name <chr>,
## #   treatment <dbl>, age_group <fct>, chameleon_awareness_flag <dbl>,
## #   good_and_gather_awareness_flag <dbl>, brand <chr>, score <int>
```

```
ate  <- mean(reshaped_data$score[d$treatment == 1], na.rm = TRUE) -
                      mean(reshaped_data$score[d$treatment == 0], na.rm = TRUE)
print(ate)
```

```
## [1] 0.1025641
```

```
ate <- mean(reshaped_data$score[reshaped_data$treatment == 1], na.rm = TRUE) -
       mean(reshaped_data$score[reshaped_data$treatment == 0], na.rm = TRUE)
print(ate)
```

```
## [1] -0.07843137
```

```
model1<- lm(score ~ treatment,  data = reshaped_data)
model2 <- lm(score ~ treatment + log(age) + gender + chameleon_awareness + brand,  data = reshaped_data)
```

```
stargazer(model1, model2,
          type = "latex",   # Use "html" for HTML output or "latex" for LaTeX
          title = "Regression Results",
          covariate.labels = c("Treatment", "log(Age)", "Gender",
                               "Chameleon Awareness (Negative)",
                               "Chameleon Awareness (Neutral)",
                               "Chameleon Awareness (Positive)",
                               "brand Good and Gather"),  # Manually add all factor levels
          star.cutoffs = c(0.10, 0.05, 0.01, 0.001),
          report = "vc*p")  # Optional: Save output to a text file
```

```
##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac a
## % Date and time: Wed, Dec 11, 2024 - 9:02:00 PM
## \begin{table}[!htbp] \centering
##   \caption{Regression Results}
```

```
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{\textit{Dependent variable:}} \\
## \cline{2-3}
## \\[-1.8ex] & \multicolumn{2}{c}{score} \\
## \\[-1.8ex] & (1) & (2)\\
## \hline \\[-1.8ex]
##  Treatment & $-$0.078 & $-$0.275 \\
##   & p = 0.727 & p = 0.217 \\
##   & & \\
##  log(Age) &  & 1.250$^{****}$ \\
##   &  & p = 0.0003 \\
##   & & \\
##  Gender &  & 0.331 \\
##   &  & p = 0.124 \\
##   & & \\
##  Chameleon Awareness (Negative) &  & $-$0.319 \\
##   &  & p = 0.663 \\
##   & & \\
##  Chameleon Awareness (Neutral) &  & 1.062$^{***}$ \\
##   &  & p = 0.005 \\
##   & & \\
##  Chameleon Awareness (Positive) &  & 0.529$^{*}$ \\
##   &  & p = 0.090 \\
##   & & \\
##  brand Good and Gather &  & $-$0.022 \\
##   &  & p = 0.917 \\
##   & & \\
##  Constant & 3.912$^{****}$ & $-$0.631 \\
##   & p = 0.000 & p = 0.592 \\
##   & & \\
## \hline \\[-1.8ex]
## Observations & 180 & 180 \\
## R$^{2}$ & 0.001 & 0.121 \\
## Adjusted R$^{2}$ & $-$0.005 & 0.085 \\
## Residual Std. Error & 1.490 (df = 178) & 1.422 (df = 172) \\
## F Statistic & 0.123 (df = 1; 178) & 3.375$^{***}$ (df = 7; 172) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

Table 5: Regression Results

	Dependent variable:	
	score	
	(1)	(2)
Treatment	−0.078	−0.275
	p = 0.727	p = 0.217
log(Age)		1.250****
		p = 0.0003
Gender		0.331
		p = 0.124
Chameleon Awareness (Negative)		−0.319
		p = 0.663
Chameleon Awareness (Neutral)		1.062***
		p = 0.005
Chameleon Awareness (Positive)		0.529*
		p = 0.090
brand Good and Gather		−0.022
		p = 0.917
Constant	3.912****	−0.631
	p = 0.000	p = 0.592
Observations	180	180
R^2	0.001	0.121
Adjusted R^2	−0.005	0.085
Residual Std. Error	1.490 (df = 178)	1.422 (df = 172)
F Statistic	0.123 (df = 1; 178)	3.375*** (df = 7; 172)
Note:		*p<0.1; **p<0.05; ***p<0.01