STAT 158 Midterm

SID: 3035924449

Casey Lei

Question 1: Design

Units: The 96 call center workers

Conditions: There are 6 treatment conditions here, each of which is a combination of two factors:

1. The three kinds of photos an employee could receive in their information packet: photo of woman running, collage of achievement-related photos, or no photo
2. Presence of conscious goal setting: employees are either consciously encouraged to raise $1200 or just encouraged to do their best

Response: The number of dollars raised by each employee at the end of a 3 hour shift.

Potential outcomes:

- The potential outcomes are every possible outcome that a subject would have had under each of the 6 treatment combinations. Only one of these is realized (the outcome of the treatment combination the subject actually receives) but the table of potential outcomes would have 6 outcomes for each subject. For subject i, let $Z_i = 1, Z_i = 2, \ldots, Z_i = 6$ indicate which treatment combination the subject received. $Y_i(j,k) = Y_i(1), \ldots, Y_i(6)$ would represent the amount of money i would have raised under each treatment combination (note that treatments 1 through 6 are combinations of some level j of factor 1 and level k of factor 2).
- Because each of the 96 subjects receives one of 6 possible treatment combinations, there are

  6 treatments x 96 subjects = 576 potential outcomes

Reliability: The response appears to be reliable. We can accurately measure the amount of money raised by each employee by summing the donations and unless the donation was noted incorrectly, there should not be any measurement error. One issue however is that even if we reran the experiment under the exact same conditions, we cannot guarantee that the donors would still donate the same amount of money, which could give us different results that do not necessarily reflect the performance of the employees.

Validity: The goal of this experiment to examine how conscious and unconscious goals affect employee performance. While it might be true that more consciously motivated employees can successfully convince more people to donate money, ultimately the amount of money that is donated depends on how much donors can/want to give. This can be influenced by factors irrelevant to our study too, such as the state of the economy or the potential donor's affinity for the university. Moreover, soliciting donations is a very specific task so the results of this experiment may not accurately generalize to other types of work, such as those with more solitary work or collaboration within a team.

Question 2: Procedure

Randomization:

1. In a spreadsheet, write out the 6 treatment combinations in a column.
2. Because we have 96 subjects in a balanced experiment, duplicate the column of treatment combinations 16 times so there are 96 rows in the column.
3. Write out the names of the workers in a column so that each name has a treatment combination on the same row.
4. Randomize the orders of both the worker names and the treatment combinations within their respective columns using the randomize range function in the spreadsheet.

5. Give each worker the treatment on the same row as their name after the randomization.

Other steps:

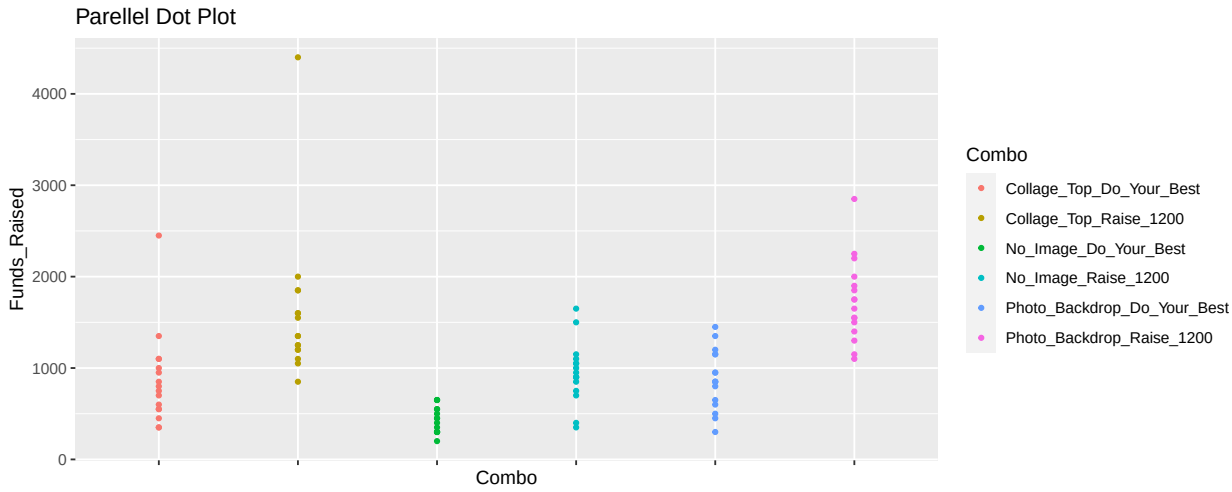
1. I would ensure all subjects were blinded, i.e. they did not know they were being assigned treatments or that other subjects received packets and/or goal-setting different from the ones they received. Knowledge of the experiment may skew subjects to behave differently from how they do naturally because they are aware they are being studied, they want to prove a treatment does/does not have effect, etc.
2. I would have the same person give all employees both forms of conscious goal setting (i.e. same person tells all the employees either to raise $1200 or to do their best). This would remove confounding in how the goal was given; specifically, we are worried that different people may urge employees to raise $1200 or to do their best more/less enthusiastically than others, which could impact the employees' motivation. By having the same person speak to all the employees, we ensure that the treatment is standardized.

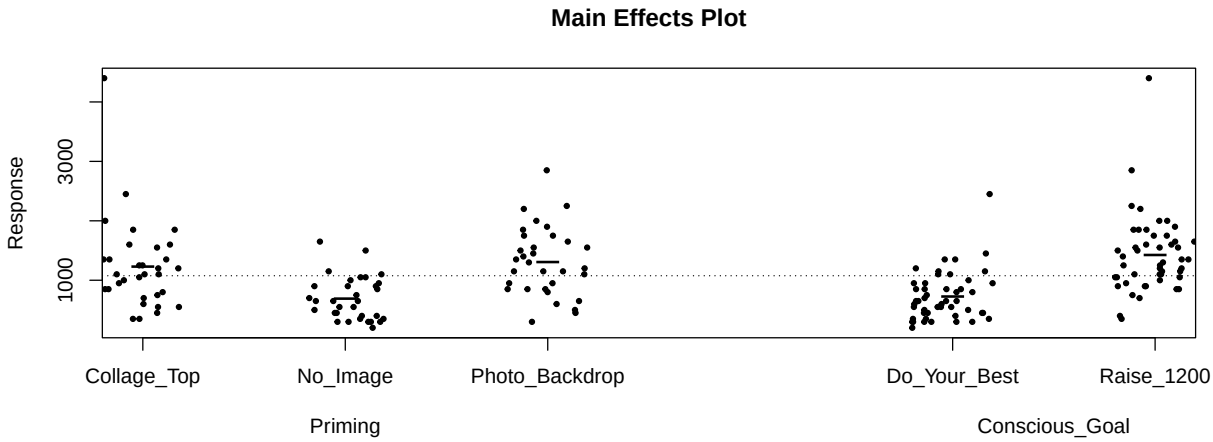
Question 3: EDA

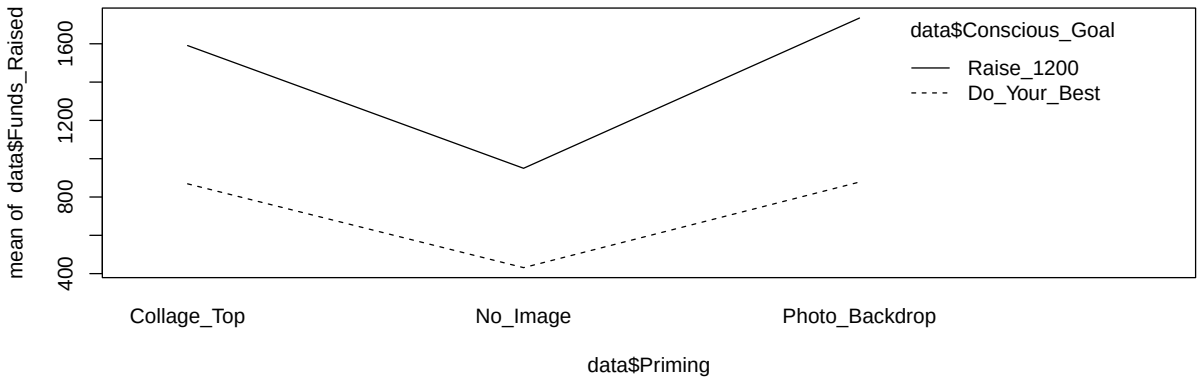
We read in the data and create a new column that combines the existing factor columns (Primed and Conscious_Goal) into a column of treatment combinations (Combo). Then, we create the parallel dot plot, main effects plot, and interaction plot. We also get the treatment group means.

X	Priming	Conscious_Goal	Funds_Raised	Combo
1	Photo_Backdrop	Raise_1200	1300	Photo_Backdrop_Raise_1200
2	Photo_Backdrop	Raise_1200	1500	Photo_Backdrop_Raise_1200
3	Photo_Backdrop	Raise_1200	1100	Photo_Backdrop_Raise_1200
4	Photo_Backdrop	Raise_1200	1550	Photo_Backdrop_Raise_1200
5	Photo_Backdrop	Raise_1200	1550	Photo_Backdrop_Raise_1200
6	Photo_Backdrop	Raise_1200	1900	Photo_Backdrop_Raise_1200

```
## # A tibble: 6 x 2
##   Combo                       mean_funds_raised
##   <chr>                                   <dbl>
## 1 Collage_Top_Do_Your_Best                 869.
## 2 Collage_Top_Raise_1200                  1591.
## 3 No_Image_Do_Your_Best                    431.
## 4 No_Image_Raise_1200                      950
## 5 Photo_Backdrop_Do_Your_Best              878.
## 6 Photo_Backdrop_Raise_1200               1734.
```







EDA Comments

Parallel Dot Plot: The ranges of funds raised differs greatly across different treatment combinations, with the smallest range within the employees who received no priming and were told to do their best (No_Image_Do_Your_Best). The employees who received conscious goal setting

(Raise_1200) consistently raised more money than those who did not across all levels of Priming. It seems like No_Image_Do_Your_Best has the lowest mean Funds_Raised of all the treatments and Photo_Backdrop_Raise_1200 has the highest. Collage_Top_Do_Your_Best, Collage_Top_Raise_1200, and Photo_Backdrop_Raise_1200 all have one large outlier.

Main Effects Plot: Collage_Top, Do_Your_Best, and Raise_1200 all have a large outlier. Raise_1200 has a greater mean effect on the amount of funds raised than Do_Your_Best does, and both Collage_Top and Photo_Backdrop have similar means that are greater than the mean effect of No_Image.

Interaction Plot: The interaction plot between Priming and Conscious_Goal looks rather parallel, which implies there may be little interaction between the plots, although we cannot say for sure. It also seems like Raise_1200 has a stronger effect than Do_Your_Best, and Photo_Backdrop and Collage_Top have stronger effects than No_Image.

Question 4: Inference & Contrasts

Randomization Inference

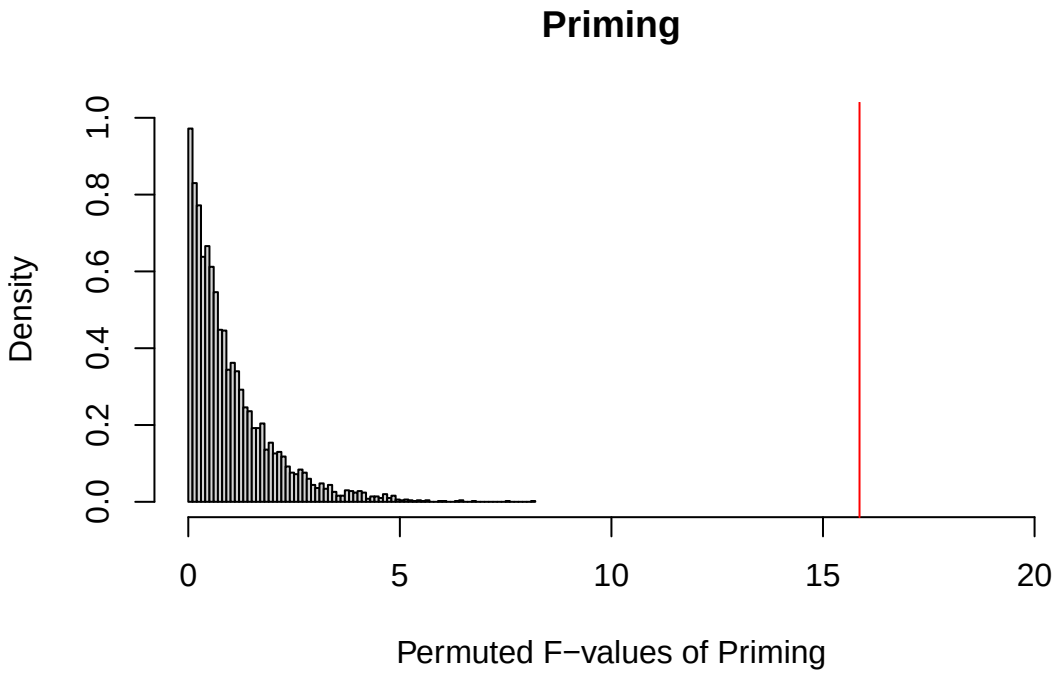
We conduct randomization inference by randomizing the Funds_Raised column across all the treatment combinations (Combo) to create the null distribution. Our **null hypothesis** is that neither priming nor conscious goal setting has an effect on the amount of funds an employee raises, i.e. the mean funds raised across all treatment combinations is equal. Our **alternative hypothesis** is that at least one treatment has an effect on the mean funds raised, i.e. at least one mean treatment's means funds raised is different from the others. We use the F-value for our analysis.

First, we get the observed F-value from our data.

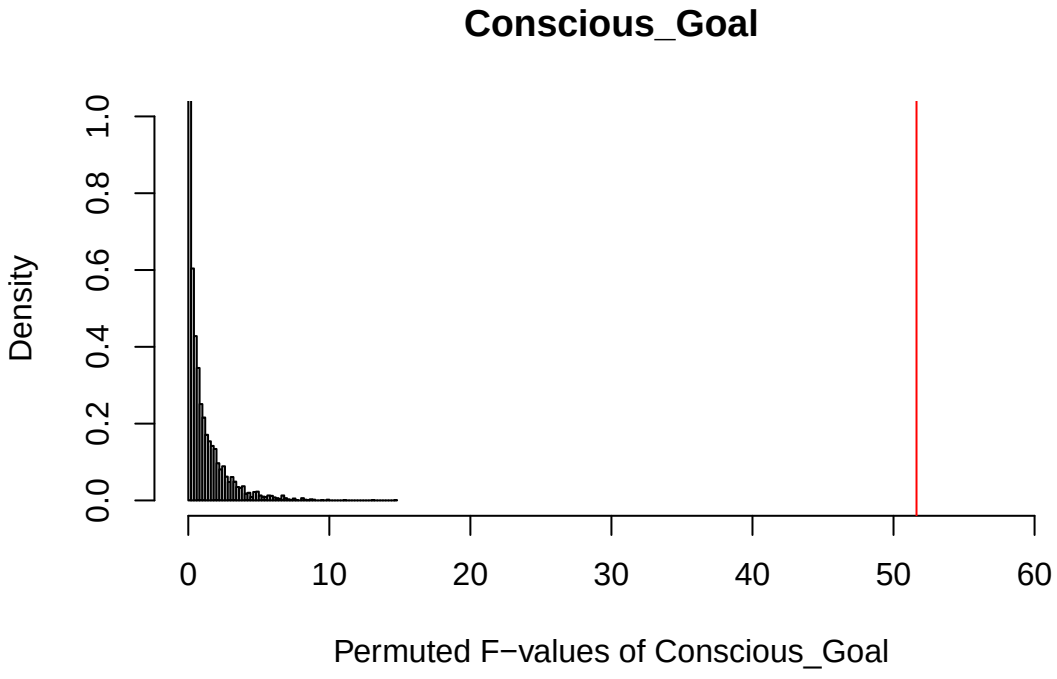
```
# ANOVA
aov.rslt <- aov(Funds_Raised ~ Priming * Conscious_Goal, data = data)

# Get F-values
obs_F <- summary(aov.rslt)[[1]]$`F value`
summary(aov.rslt)
```

```
##                         Df   Sum Sq  Mean Sq F value   Pr(>F)
## Priming                  2  7204740  3602370  15.863 1.26e-06 ***
## Conscious_Goal           1 11725026 11725026  51.631 1.88e-10 ***
## Priming:Conscious_Goal   2   461927   230964   1.017    0.366
## Residuals               90 20438281   227092
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Next, we conduct randomization inference and create null distribution by running 5000 permutations of the randomization of values in Funds_Raised across the individual factors and the treatment combinations.

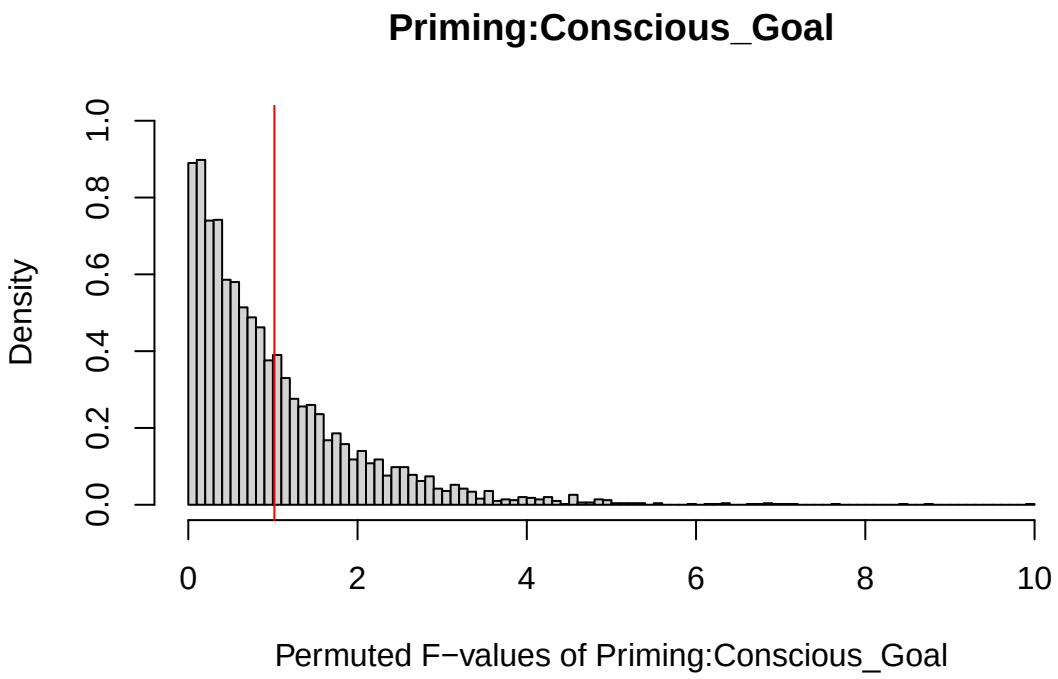



```
## Pvalue permutation of: Priming
## [1] 0
```




```
## Pvalue permutation of: Conscious_Goal
```

```
## [1] 0
```




```
## Pvalue permutation of: Priming:Conscious_Goal
## [1] 0.3646
```

We have significant p-values for the individual factors of Priming and Conscious_Goal but not for the interaction between them. The interaction has a p-value of 0.3646, so it is not significant at the $\alpha = 0.05$ level. The randomization inference yielded p-values of 0 for Priming and Conscious_Goal, but this does not mean that the actual p-value is 0, only that by chance none of the F-values we got in this randomization inference were greater than the observed F-values for those two factors. Thus, we reject the null and find that there is significant treatment effect for both Priming and Conscious_Goal at the $\alpha = 0.05$ level, but we fail to reject for the interaction Priming:Conscious_Goal.

Contrasts

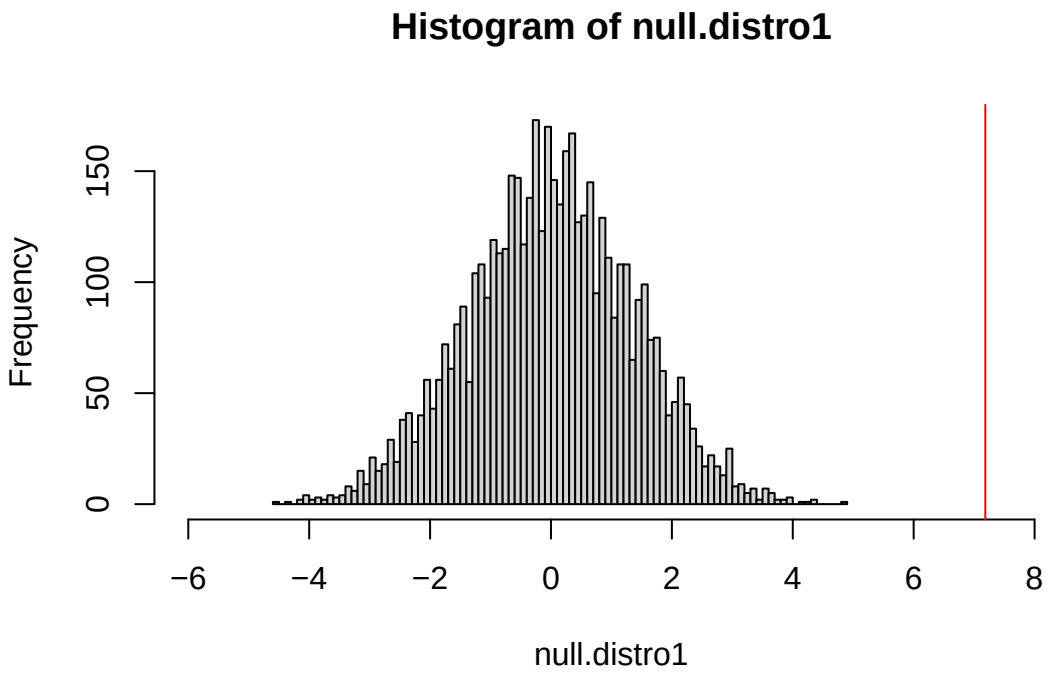
1. Raise_1200 vs. Do_Your_Best: This contrast would show us whether providing a very specific work goal or a vague encouragement impacts the amount of money employees raise more, which is directly relevant to the research question of conscious goal setting.
2. No_Image vs. Image (Photo_Backdrop or Collage_Top): This contrast would show us whether providing an employee with no image vs. showing them an encouraging image impacts the amount of funds they raise, which is directly relevant to the research question of whether nonconscious goal setting impacts worker performance.
3. Collage_Top vs. No_Image: Collage_Top is most directly achievement-related photo in our Priming levels, and in our EDA it seemed to yield higher Funds_Raised than other levels. This contrast would give more insight on whether this specific level has a significant impact on Funds_Raised compared to the "control" treatment of No_Image.

Multiple Testing: Because we are conducting three tests instead of one, we use the Bonferroni correction to correct for multiple testing and divide $\alpha = 0.05/3$ for each and check the 0.0167 quantile for significance in each test. This maintains the familywide error rate of 0.05 and controls for Type I errors.

Contrast 1: Raise_1200 vs. Do_Your_Best

Null hypothesis: The mean funds raised by employees told to raise $1200 is equal to the mean funds raised by employees told to do their best.

Alternative: The mean funds raised by employees told to raise $1200 is higher than the mean funds raised by employees told to do their best.



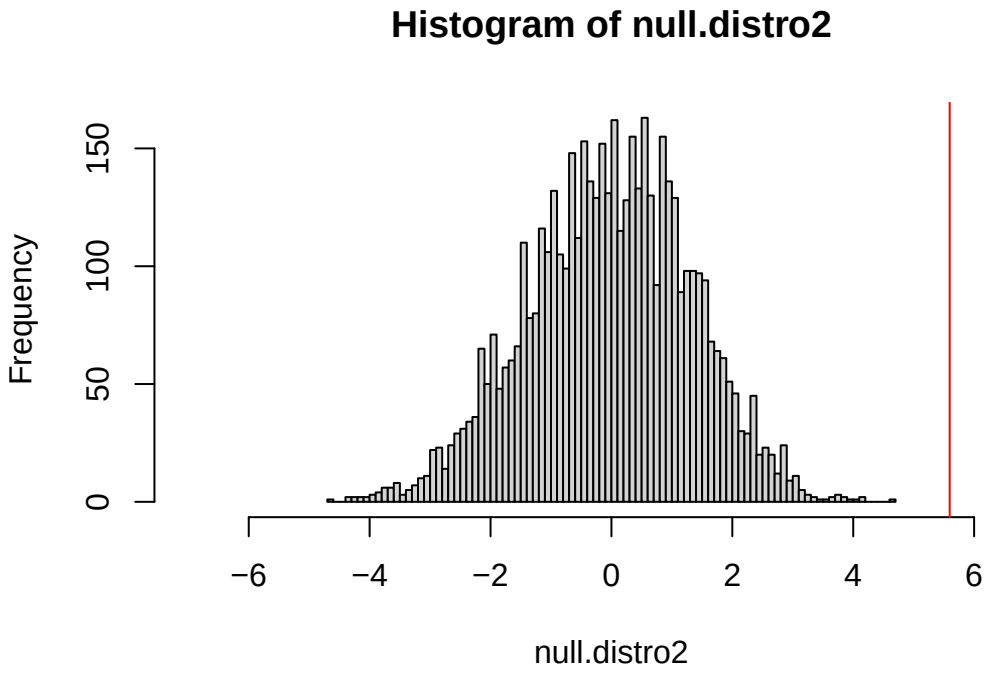

```
## [1] "P-value:  0"
```

For Contrast 1, we get a p-value of approximately 0 (only because this is randomization inference, not model-based) which is less than our Bonferroni corrected α of 0.0167. Thus, we reject the null.

Contrast 2: Image (Photo_Backdrop or Collage_Top) vs. No_Image

Null hypothesis: The mean funds raised by employees who received either a photo backdrop or a collage is equal to the mean funds raised by employees who received no image.

Alternative hypothesis: The mean funds raised by employees who received either a photo backdrop or a collage is greater than the mean funds raised by employees who received no image.



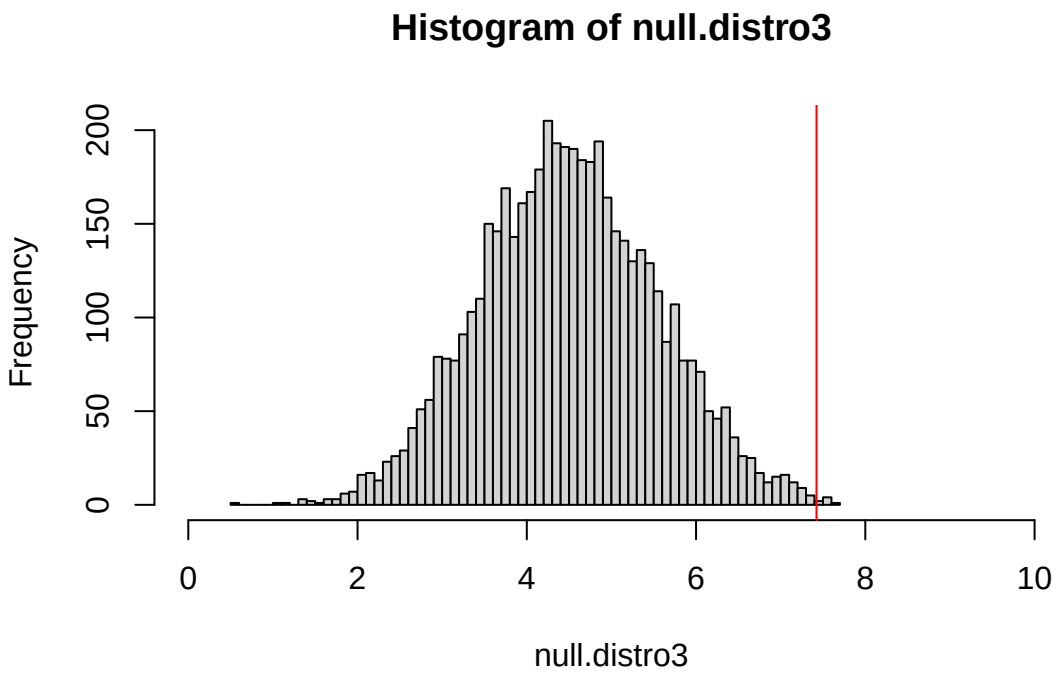

```
## [1] 0
```

For Contrast 2, we get a p-value of approximately 0. This is less than our Bonferroni corrected α of 0.0167. Thus, we reject the null.

Contrast 3: Collage_Top vs. No_Image

Null hypothesis: The mean funds raised by employees who received the collage image is equal to the mean funds raised by employees who received no image.

Alternative hypothesis: The mean funds raised by employees who received the collage image is greater than the mean funds raised by employees who received no image.




```
## [1] 0.0012
```

For Contrast 3, we get a p-value of 0.0012. This is less than our Bonferroni corrected α of 0.0167. Thus, we reject the null.

Question 5: Summary

Overall, we find evidence that both unconscious and conscious goal setting have a statistically significant positive impact on the amount of funds an employee raises. We do not find significant evidence that the interaction between these factors has an effect. We find that being told to raise $1200 has a significantly stronger positive effect on the funds an employee raises compared to being told to do their best—from the EDA, this appears to have a mean effect of $500+. We find that both being shown any image at all and being shown the collage specifically have significant positive effects on the amount of funds raised as well. In all, the employees who received the combination of no image and were told to do their best raised the least mean funds of all the treatment combinations, while those with a photo backdrop who were told to raise $1200 raised the most mean funds.

Strengths of Study: This study randomizes treatment assignments, which reduces biases and potentially allows for causal inference more so than an observational study because we are more confident that our treatment groups are comparable so conclusions from our study can be generalized. We have control over the factors we are choosing to study, unlike a purely observational study. Our sample size is relatively large, with 96 subjects, which better allows us to detect effects because we can duplicate our treatments over multiple subjects and compare the responses and detect interactions.

Weaknesses: The scope of this experiment is rather limited. Phone calls for donations are a very specific type of work that does not resemble other office or blue collar labor and relies on the donors' ability and willingness to donate, which is not within our or the employees' control and is not completely determined by how hardworking the employee is. This experiment also only studies short-term effects, since we collect the

responses at the end of three hours, so we cannot make conclusions on how the factors may impact long-term work performance.

Question 6: Blocking

A couple ways to block could be by the employee's gender or their age. These would be relevant because male and older voices tend to be perceived as more assertive, and older people also tend to be more confident when speaking, all of which could make the employee sound more compelling to potential donors over the phone. By blocking, we can ensure that all the employees within a certain gender or age group receive the treatments an equal amount, so we will not run into issues such as all young employees receiving conscious goal setting but raising less money than older employees because they sound less credible over the phone, despite being motivated to work harder. This would be a feasible design if we have an equal number of employees of each gender or age groups. We cannot remove the attributes of gender or age from a subject, so by blocking on factors like these that are not relevant to the question we are researching (goal setting), we can control for unwanted variation due to these factors. We could also pull the data from the subgroups and analyze them as their own experiments.

Assuming we have equal numbers of only male and female employees:

1. Divide the employees by gender into two blocks, male and female. There should be $96/2 = 48$ people in each block.
2. In a spreadsheet, write the male employee names in a column and the female employee names in another column.
3. Write the 6 unique treatment combinations in two columns, one next to the male names and one next to the female names.
4. Duplicate the treatment combinations 8 times in each column.
5. Randomize the order of names and treatments for all four columns.
6. Assign each name to the treatment combination next to it.