STA305 Assignment 2

William Wang, 1004278818

Contents

Part 1 – A Published Experiment

(10 marks) Using the virtual Library Guide under Assignments Section in the class website, identify a recent research article in your field of interest (for example, Agriculture, Biology, Health, Psychology) in which at least one real scientific experiment was studied. The article must have been published by at least one University of Toronto author. Further, the year of publication must be between 2000 and 2020, and end with the last digit of your student number. Note, STA1004 students are allowed to write about their current research (or proposal). Based on your article, answer the following

1. What was your selected field of interest?

The selected field of interest was Education and Computers, where the experiment looks at how computer game design can affect learning.

2. Write a proper reference for the article, including the author(s), title, journal, year of publication, volume and page indices. Provide a link to the article or a soft copy of the article.

The reference for the article can be seen below.

Gauthier, A., & Jenkinson, J. (2018). Designing productively negative experiences with serious game mechanics: Qualitative analysis of game-play and game design in a randomized trial. *Computers & Education*, 127, 66–89. doi: 10.1016/j.compedu.2018.08.017

The exact DOI link is http://dx.doi.org/10.1016/j.compedu.2018.08.017

Alternatively, another link that isn't a DOI is shown below:

https://www-sciencedirect-com.myaccess.library.utoronto.ca/science/article/pii/S0360131518302203

A soft copy is also found in the Appendix

3. Which UofT department or institute was the UofT author affiliated with?

The authors of the experiment are from the University of Toronto Biology department, specifically Biomedical Communications, and the University of Toronto's Institute of Medical Sciences.

4. Which database from the Library Guide did you use?

The database used to find the article was Scopus.

5. What was the design of the experiment?

The experiment was a completely randomized design where students were randomly distributed between two of the treatment groups, the computer-based interactive simulation group or the serious game group (i.e. a game that isn't solely for entertainment). The authors designed 2 treatments, a serious game, *MolWorlds*, which only differed from the other treatment, a simulation dubbed *MolSandbox*, by some additional game mechanics to *MolWorlds*, which isolates any changes to the outcome as a result of the added mechanics. The students then played their assigned treatment and the experimenters would look at how the participants interacted in each of their scenarios, comparing the two treatment group results, analyzing recorded gameplay and drawing conclusions.

6. Briefly describe the experimental units used in the study.

A total of 40 students were assigned to be part of one of the two treatment groups, with 20 being randomly allocated towards each of our two different groups. These students were undergraduates, in their first, second, or third years, studying biology at the University of Toronto Mississauga Campus. Of the 40 students who were given the treatment, 15 were first year, 13 second year and 12 third year, and these were randomly assigned to the two groups. There was no difference in gaming habits, school achievement or biology course engagement levels between the individuals. Thus, our experimental units are undergraduate biology students, specifically first-, second- and third-year students, from the University of Toronto Mississauga campus.

7. What was the response variable(s) of the study?

This experiment was conducted to investigate how the design of games could affect learning through productive negativity, i.e. learning from failure, a common process in both education and gaming. One of the response variables in the study was a count of the number of correct concepts that students were able to demonstrate for each treatment. There was also a count for the number of instances in which the experimenters saw productive negativity had occurred. Thus, the response variables for this study were counts of the number of times the experimenters saw productive negativity and conceptual knowledge being retained.

8. Name the factors which were explored in the experiment.

The factor that was explored was the overall design of the game, more specifically specific gameplay mechanics/features were the factors being explored, which the authors were trying to determine the effect of. Thus, this factor had 2 levels, whether the additional mechanics were present and when they were not.

9. How many observations were recorded in the study?

There were 40 observations in the study, corresponding to each of our 40 students in the treatment groups.

10. Identify at least one statistical method used to analyze the data

The authors conducted a Chi-Square Test to test the goodness of fit in their analysis of the data. They also applied Bonferroni's method to account for multiple comparisons in their data analysis as well. Aside from the statistical methods that were taught in the course, they also performed two-sided Mann-Whitney U tests and Wilcoxon signed-rank tests.

Part 2 – My Factorial Experiment

Description of the Design

When ever I am at home, I often use rubber bands to seal bags/wrappers or to keep things closed or together. However, I notice that often times my rubber bands seemingly break in different scenarios, often depending on the temperature it was used at, the size of the rubber band or the brand I am using. Thus, I was curious on whether I could more reliably determine when these rubber bands will break, and to do so I decided to run an experiment. The experiment conducted will examine the stretchability of rubber under various factors, whereby we will test the break point of rubber bands in varying conditions. The design of such experiment will be a 2^3 factorial design with a total of 8 different possible scenarios, to mimic the possibilities in my own home use. For the experiment, I will measure the length at which the rubber band would break at each of the given combinations of possibilities, which would be my response variable. Since I expected that there may be large variability, I decided to also replicate my experiment 2 times to be able to better estimate the variance.

The factors that will be tested are the temperature at which the rubber is at, the size of the rubber band and the brand of the rubber band. The temperatures we will be testing are a below freezing temperatures (-18 degrees Celsius) and a high temperature (100 degrees Celsius). The rubber bands will all have a thickness of 0.2 mm, and have a width/height of 0.4 mm, however the component of size we will be looking at is the length, and therefore the circumference, of the rubber band. We will have 2 variations of length that we will change to examine the effect of size, medium-sized rubber bands with a 17-20 cm circumference and large-sized rubber bands with 27-30 cm circumferences. The brands that will be tested are Studio brand rubber bands and JAM brand rubber bands. Thus, I tested 3 variables all with 2 factor levels.

The set up of my experiment required me to obtain two bags of rubber bands with assorted sizes of the two specified brands, a tape measure, and places that would simulate my temperatures. Thus, to simulate the temperatures I used a freezer, and a pot of boiling water, to simulate the below freezing and high temperatures respectively. In the case of the freezer, I left the rubber bands it for 24 hours to ensure the temperatures would be retained, then conducted my experiment one band at a time, keeping the rest in the freezer. For the high temperature, I brought a pot of water to a boil then placed the rubber bands in the water for 20 minutes, after, I would remove one band at a time and immediately begin my experiment, while keeping the rest in the boiling water. The procedure of my experiment entails taking the rubber band, placing it on one end of the tape measure and pulling the band until it snaps, where our break point length is where it snaps.

One criticism that may come up in my design that comes to mind is the range in sizes contributing to some inaccuracy, as we are not breaking up the range into smaller more precise sizes, we are assuming that no material/rubber changes for that range of sizes. The different brands may also contribute to some problems, as the rubber/material is not exactly the same between brands, meaning our results may be affected the unknown composition of the rubber. There may also be some problems with the exactness of the response variable measurement. As the band snapping is a release of tension, this would contribute to inaccurate measurements as a human cannot react fast enough to stop themselves from continue pulling at the exact moment it breaks. As a result, I tried to minimize this reaction by being very careful with how much I pulled and pulling at increments. Below is a table of all possible combinations of factors that were tested.

Setup	Brand	Size (cm)	Temperature (°C)
1	Studio	17-20 cm	Freezing (-18°C)
2	Studio	17-20 cm	High (100°C)
3	Studio	27-30 cm	Freezing (-18°C)
4	Studio	27-30 cm	High (100°C)
5	JAM	17-20 cm	Freezing (-18°C)
6	JAM	17-20 cm	High (100°C)
7	JAM	27-30 cm	Freezing (-18°C)
8	JAM	27-30 cm	High (100°C)

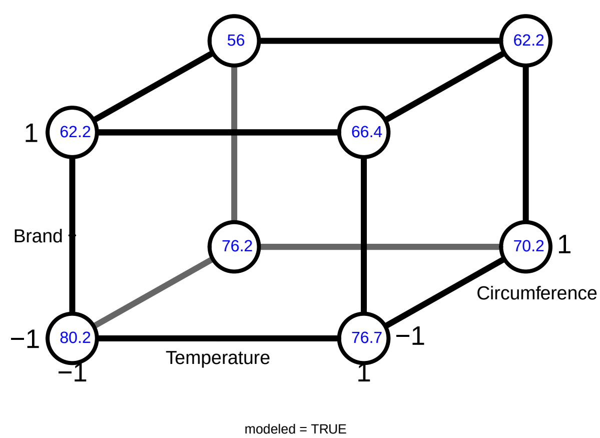
Analysis of the Data

As the first step in my analysis I set each of my factors to a specific value, either 1 or -1. So, I set the freezing temperatures, Studio brand and medium-sized rubber bands as the value of 1 and the other level of the factors would be -1. Since we replicated 2 times, we need to look at the average of the break lengths of the 2 runs with same conditions. Thus, after compiling my 16 observations, I found the average of the break point lengths for each of our set of conditions across the two runs, seen below and in Table 7 of the Appendix.

Table 2: Table of the averages of the duplicate runs

Run	Temperature	Brand	Circumference	Avg_Break_Length
1	1	1	1	62.2
2	-1	1	1	56.0
3	1	1	-1	66.4
4	-1	1	-1	62.2
5	1	-1	1	70.2
6	-1	-1	1	76.2
7	1	-1	-1	76.7
8	-1	-1	-1	80.2

At a glance, I noticed that the break point for Studio brand got lower when temperature increased, while JAM brand had the opposite effect, where it increased as temperature increased, and this seemed to apply no matter the size. Thus, this gave me an idea of what two factors may affect our response.



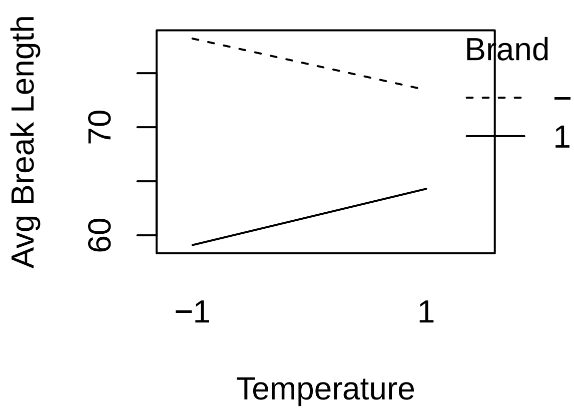
Once I had the averages I was able to create a cube plot of my factors, as well as determine the main effects of my factors of `Temperature`, `Circumference`, and `Brand`, being 0.225, -5.225, and -14.125 respectively. These main effects suggest that there is a negative effect for our factors of `Circumference` and `Brand` on our response variable, and a positive one for `Temperature`. In addition to the main effects, I found the interaction effects, which for the two-factor interactions are 4.975, -0.125, and 0.025, for the interactions between `Temperature` & `Brand`, `Temperature` & `Circumference`, and `Brand` & `Circumference`, respectively. These effects suggest that when that specific interaction occurs, we could see a positive effect on the response for the effects for `Temperature` & `Brand`and `Brand` & `Circumference`, while a negative effect may occur for `Temperature` & `Circumference`. A similar calculation was done for the three-factor interaction. Along with these effects, I found the variances and standard error of the data, and since I did a replicated factorial experiment a Lenth, Half-Normal or Normal Plot would not be appropriate, so, I computed the estimated variance of each of the 16 set of observations, which is seen in Table 6 of the Appendix. Furthermore, the pooled variance estimate with 8 degrees of freedom was found to be 4.205, which we can use to determine the variance and standard error of factorial effects for duplicated runs, which are 1.0512 and 1.0253, both of which are rather small relative to our response. For a better idea of the effects I created a linear model for the factorial design, the code of which can be found in the appendix.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	68.7625	0.5126524	134.1308415	0.0000000
Temperature	0.1000	0.5126524	0.1950639	0.8502059
Brand	-7.0625	0.5126524	-13.7763907	0.0000007
Circumference	-2.6250	0.5126524	-5.1204284	0.0009067
Temperature:Brand	2.5000	0.5126524	4.8765985	0.0012296
Temperature:Circumference	-0.0625	0.5126524	-0.1219150	0.9059732
Brand:Circumference	0.0000	0.5126524	0.0000000	1.0000000
Temperature:Brand:Circumference	0.5375	0.5126524	1.0484687	0.3250624

We can see that of our effects, only the main effects for Brand, and Circumference as well as the interaction effect for Temperature & Brand had p-values significantly less that 0.05, which is the typical significance level used. This suggests that we have evidence that only Brand, and Circumference and the interaction between Temperature & Brand will have a statistically significant effect on our response variable. This is further supported by the estimated 95% confidence intervals of our all factorial effects, shown in the table below.

	2.5 %	97.5 %
(Intercept)	135.160643	139.889357
Temperature	-2.164357	2.564357
Brand	-16.489357	-11.760643
Circumference	-7.614357	-2.885643
Temperature:Brand	2.635643	7.364357
Temperature:Circumference	-2.489357	2.239357
Brand:Circumference	-2.364357	2.364357
Temperature:Brand:Circumference	-1.289357	3.439357

We can see from the table that only the effects previously mentioned have CIs that do not contain 0, aside from the intercept itself, and that Brand and Circumference have a statistically significant negative effects and the interaction between Temperature & Brand has a significant positive effect on our response. We can also see that only the Brand and Temperature interaction plot has lines that are close to intercepting, further supporting the interaction, all three interaction plots can be found in the Appendix.



Thus, it seems that my initial idea was correct, where Temperature and Brand did have some interaction between them which affected the response.

Conclusions

From my analysis above, I found that only 3 factorial effects would have a statistically significant, the main effects for `Brand` and `Circumference` and the interaction effect of `Brand` and `Temperature.` Our analysis, provided ample support which suggested that a our main effects have a negative effect on our break point length, and the interaction effect has a positive effect. Thus, this means that in the experiment as we changed from JAM brand to Studio brand, i.e. -1 to 1, the break point length of the band decreased by about 14 cm, thus suggesting that JAM brand is significantly more elastic than Studio brand. Furthermore, as the size decreased from large rubber bands to medium, we also saw a decrease of approximately 5 cm in the break point length. This makes sense, as a larger band should be able to stretch farther than one that was smaller in size to it, and our two sizes differed by roughly 5 cm as well which could account for that decrease. We also saw an interesting interaction between `Temperature` and `Brand`, where we found that the effect of temperature on the response seemed to differ between brands. We also found a positive effect on the break point length, suggesting that as the temperature increases, and we change from JAM brand to Studio brand, we would see a positive change to our response variable. However, this effect could be a result of the wide range of temperature I tested, one being below freezing and one at boiling. Thus, I plan to further investigate the effect of temperature at more levels, perhaps at room temperature. Overall, this experiment has shown that if I needed a more elastic band than I should use JAM brand, and/or a larger sized rubber band.

Appendix

Below is the code to install any packages if they are not present on your system.

```
install.packages("tidyverse")
install.packages("FrF2")
```

Below is the code to set up

```
library(tidyverse)
library(FrF2)
library(readxl)
# The Data Can be accessed via https://github.com/William-Wang-github/STA305-A2
STA305_Assignment_2_Control_Data_ <- read_excel("STA305 Assignment 2 Control Data .xlsx")
STA305_Assignment_2_Data_Cleaned_2 <- read_excel("STA305 Assignment 2 Data Cleaned 2.xlsx")
```

Below is the code to create a new data set where the corresponding values for our factors are created, where freezing = 1, hot = -1, Studio = 1, JAM = -1, Medium sizes(17-20 cm) = 1, and Large sizes(27-30 cm) = -1

```
#Creates a new data set where our factors are changed into 1 and -1
Data_Transformed_Factors <- STA305_Assignment_2_Data_Cleaned_2 %>%
  mutate(`Temperature`= ifelse(`Temperature (Celsius)`<0, 1, -1)) %>%
  mutate(`Brand`= ifelse(`Rubber band Brand` == "Studio", 1, -1)) %>%
  mutate(`Circumference`= ifelse(`Circumference (cm)`<20, 1, -1)) %>%
  mutate(`Break_Length`= `Break Point Length at temp (cm)`) %>%
  mutate(`Run`=`Rubber Band Number`) %>%
  select(`Run`, `Temperature`,`Brand`, `Circumference`, `Break_Length`)
knitr::kable(Data_Transformed_Factors, caption = "Table of All 16 Runs of Data")
```

Table 5: Table of All 16 Runs of Data

Run	Temperature	Brand	Circumference	Break_Length
1	1	1	1	62.7
2	-1	1	1	55.6
3	1	1	-1	64.6
4	-1	1	-1	59.3
5	1	-1	1	72.1
6	-1	-1	1	75.5
7	1	-1	-1	77.0
8	-1	-1	-1	80.8
9	1	1	1	61.6
10	-1	1	1	56.4
11	1	1	-1	68.3
12	-1	1	-1	65.1
13	1	-1	1	68.3
14	-1	-1	1	76.9
15	1	-1	-1	76.3
16	-1	-1	-1	79.7

Below is the code to create a data frame which shows the two runs that were completed in separate columns, as well as the difference and estimated variance between the response variable of the two runs with the same set of conditions.

```r
Replicated_df <- Data_Transformed_Factors[1:8,1] %>%
  cbind(Data_Transformed_Factors[9:16,1]) %>%
  cbind(Data_Transformed_Factors[1:8,2:5])%>%
  cbind(Data_Transformed_Factors[9:16,5])
  colnames(data=Replicated_df) <- c("Run1", "Run2", "Temperature", "Brand",
                                    "Circumference", "Break_Length1", "Break_Length2")
Replicated_df <- Replicated_df %>%
  mutate(diff = (Break_Length1-Break_Length2)) %>%
  mutate(estvar = diff^2/2)
knitr::kable(Replicated_df, caption = "Table of all replicate runs and the estimated
             variance for each set of conditions")
```

Table 6: Table of all replicate runs and the estimated variance for each set of conditions

Run1	Run2	Temperature	Brand	Circumference	Break_Length1	Break_Length2	diff	estvar
1	9	1	1	1	62.7	61.6	1.1	0.605
2	10	-1	1	1	55.6	56.4	-0.8	0.320
3	11	1	1	-1	64.6	68.3	-3.7	6.845
4	12	-1	1	-1	59.3	65.1	-5.8	16.820
5	13	1	-1	1	72.1	68.3	3.8	7.220
6	14	-1	-1	1	75.5	76.9	-1.4	0.980
7	15	1	-1	-1	77.0	76.3	0.7	0.245
8	16	-1	-1	-1	80.8	79.7	1.1	0.605

Below is the code to calculate the pooled estimated variance, as well as the variance and standard error of a factorial effect for duplicated runs.

```r
# Calculations for pooled variance, and factorial effect var and std dev
pooledestvar <- sum(Replicated_df$estvar)/8
factorial_effvar <- 2*(1/8)*pooledestvar
factorial_effse <- sqrt(factorial_effvar)
```

Below is the code to create a table showing the averages of the two replicated runs.

```r
#Average out Break Points
Avg_Transformed_Factors2 <- Data_Transformed_Factors %>%
  group_by(`Temperature`,`Brand`,`Circumference`)%>%
  mutate(Avg_Break_Length = round(mean(Break_Length),digits = 1)) %>%
  filter(row_number(Avg_Break_Length)==1) %>%
  select(`Run`,`Temperature`,`Brand`,`Circumference`, `Avg_Break_Length`)
knitr::kable(Avg_Transformed_Factors2, caption = "Table of the averages of the duplicate runs")
```

Table 7: Table of the averages of the duplicate runs

Run	Temperature	Brand	Circumference	Avg_Break_Length
1	1	1	1	62.2
2	-1	1	1	56.0
3	1	1	-1	66.4
4	-1	1	-1	62.2
5	1	-1	1	70.2
6	-1	-1	1	76.2
7	1	-1	-1	76.7

Run	Temperature	Brand	Circumference	Avg_Break_Length
8	-1	-1	-1	80.2

Below is the code to create the cube plot for our 2^3 factorial experiment

```
lmavg <- lm(data = Avg_Transformed_Factors2, Avg_Break_Length ~ Temperature*Brand*
                     Circumference)
cubePlot(lmavg, "Temperature","Circumference", "Brand",
         main = "Cube Plot for Rubber Band Experiment")
```

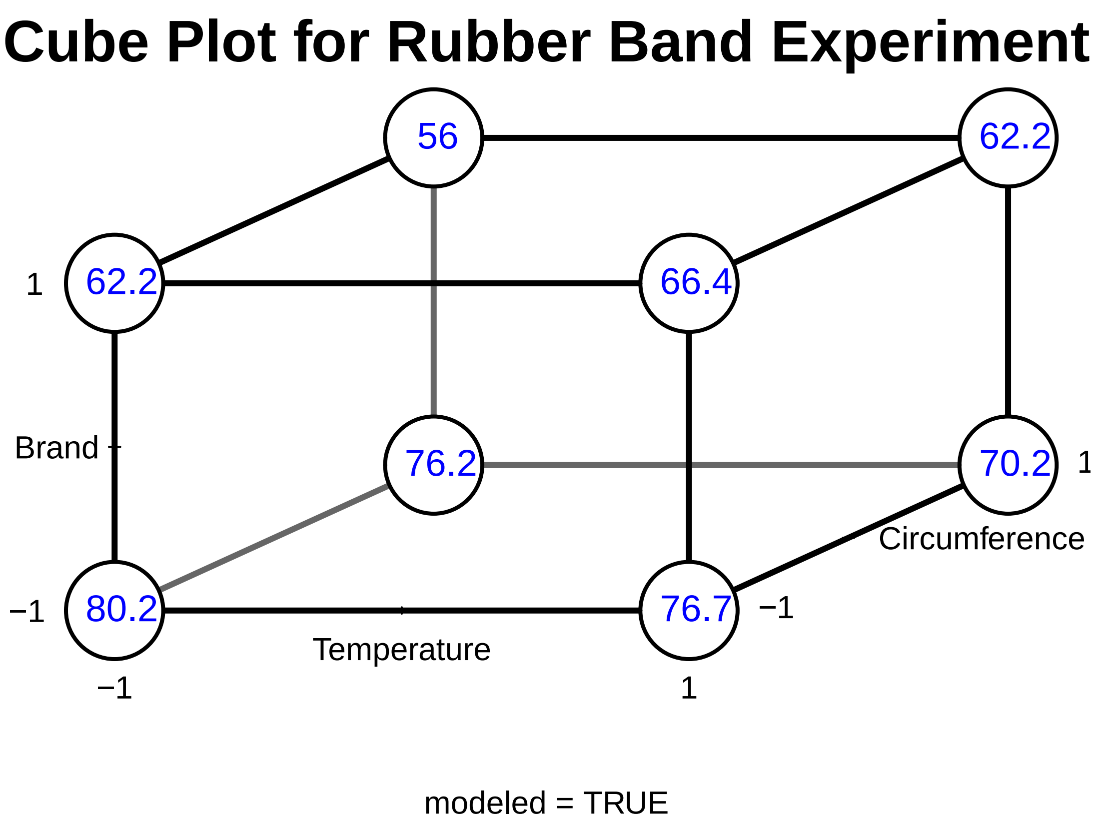


Figure 1: Cube Plot for the three factors

Below is the code which returns a summary of the coefficients of the linear model for the factorial experiment as well as the confidence intervals of said coefficients

```
lmreplicated <- lm(data = Data_Transformed_Factors, Break_Length ~ Temperature*Brand*
                     Circumference)
summary(lmreplicated)
```

```
##
## Call:
## lm.default(formula = Break_Length ~ Temperature * Brand * Circumference,
##     data = Data_Transformed_Factors)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9000 -0.5875  0.0000  0.5875  2.9000
```

```
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  6.876e+01  5.127e-01 134.131 1.07e-14 ***
## Temperature                  1.000e-01  5.127e-01   0.195 0.850206
## Brand                       -7.062e+00  5.127e-01 -13.776 7.44e-07 ***
## Circumference               -2.625e+00  5.127e-01  -5.120 0.000907 ***
## Temperature:Brand            2.500e+00  5.127e-01   4.877 0.001230 **
## Temperature:Circumference   -6.250e-02  5.127e-01  -0.122 0.905973
## Brand:Circumference          4.294e-16  5.127e-01   0.000 1.000000
## Temperature:Brand:Circumference  5.375e-01  5.127e-01   1.048 0.325062
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.051 on 8 degrees of freedom
## Multiple R-squared:  0.9679, Adjusted R-squared:  0.9397
## F-statistic: 34.42 on 7 and 8 DF,  p-value: 2.345e-05
```

```
round(2*lmreplicated$coefficients,3)
```

```
##                     (Intercept)                     Temperature
##                         137.525                           0.200
##                           Brand                   Circumference
##                         -14.125                          -5.250
##               Temperature:Brand       Temperature:Circumference
##                           5.000                          -0.125
##             Brand:Circumference Temperature:Brand:Circumference
##                           0.000                           1.075
```

```
2*confint.lm(lmreplicated)
```

```
##                                     2.5 %     97.5 %
## (Intercept)                    135.160643 139.889357
## Temperature                     -2.164357   2.564357
## Brand                          -16.489357 -11.760643
## Circumference                   -7.614357  -2.885643
## Temperature:Brand                2.635643   7.364357
## Temperature:Circumference       -2.489357   2.239357
## Brand:Circumference             -2.364357   2.364357
## Temperature:Brand:Circumference -1.289357   3.439357
```

Below is the code to calculate the main effect for each of our factors

```
TempMainEffect <-sum(Avg_Transformed_Factors2$Temperature*
                       Avg_Transformed_Factors2$Avg_Break_Length)/4
CircumMainEffect <- sum(Avg_Transformed_Factors2$Circumference*
                          Avg_Transformed_Factors2$Avg_Break_Length)/4
BrandMainEffect <- sum(Avg_Transformed_Factors2$Brand*
                         Avg_Transformed_Factors2$Avg_Break_Length)/4
```

Below is the code to calculate the interaction effects for our factors

```
TempBrandIntEffect <- (sum(Avg_Transformed_Factors2$Temperature[Avg_Transformed_Factors2$Brand
                                                                =="1"]*
                             Avg_Transformed_Factors2$Avg_Break_Length[Avg_Transformed_Factors2$Brand
                                                                      =="1"])/2-
```

```r
                          sum(Avg_Transformed_Factors2$Temperature[Avg_Transformed_Factors2$Brand
                                                                   =="-1"]*
                       Avg_Transformed_Factors2$Avg_Break_Length[Avg_Transformed_Factors2$Brand
                                                                 =="-1"])/2)/2

TempCircumIntEffect <- (sum(Avg_Transformed_Factors2$Temperature[Avg_Transformed_Factors2$Circumference
                                                                  =="1"]*
                             Avg_Transformed_Factors2$Avg_Break_Length[Avg_Transformed_Factors2$
                                                                       Circumference =="1"])/2
                       -sum(Avg_Transformed_Factors2$Temperature[Avg_Transformed_Factors2$
                                                                  Circumference=="-1"]*
                       Avg_Transformed_Factors2$Avg_Break_Length[Avg_Transformed_Factors2$Circumference
                                                                 =="-1"])/2)/2

BrandCircumIntEffect <- (sum(Avg_Transformed_Factors2$Brand[Avg_Transformed_Factors2$Circumference
                                                             =="1"]*
                       Avg_Transformed_Factors2$Avg_Break_Length[Avg_Transformed_Factors2$Circumference
                                                                 =="1"])/2-
                       sum(Avg_Transformed_Factors2$Brand[Avg_Transformed_Factors2$Circumference
                                                          =="-1"]*
                       Avg_Transformed_Factors2$Avg_Break_Length[Avg_Transformed_Factors2$Circumference
                                                                 =="-1"])/2)/2
```

Below is the code to create the two-way interaction plots between our factors

```r
interaction.plot(Avg_Transformed_Factors2$Temperature, Avg_Transformed_Factors2$Circumference,
                 Avg_Transformed_Factors2$Avg_Break_Length, type = 'l', xlab = "Temperature",
                 ylab = "Avg Break Length", trace.label = "Circumference")
```

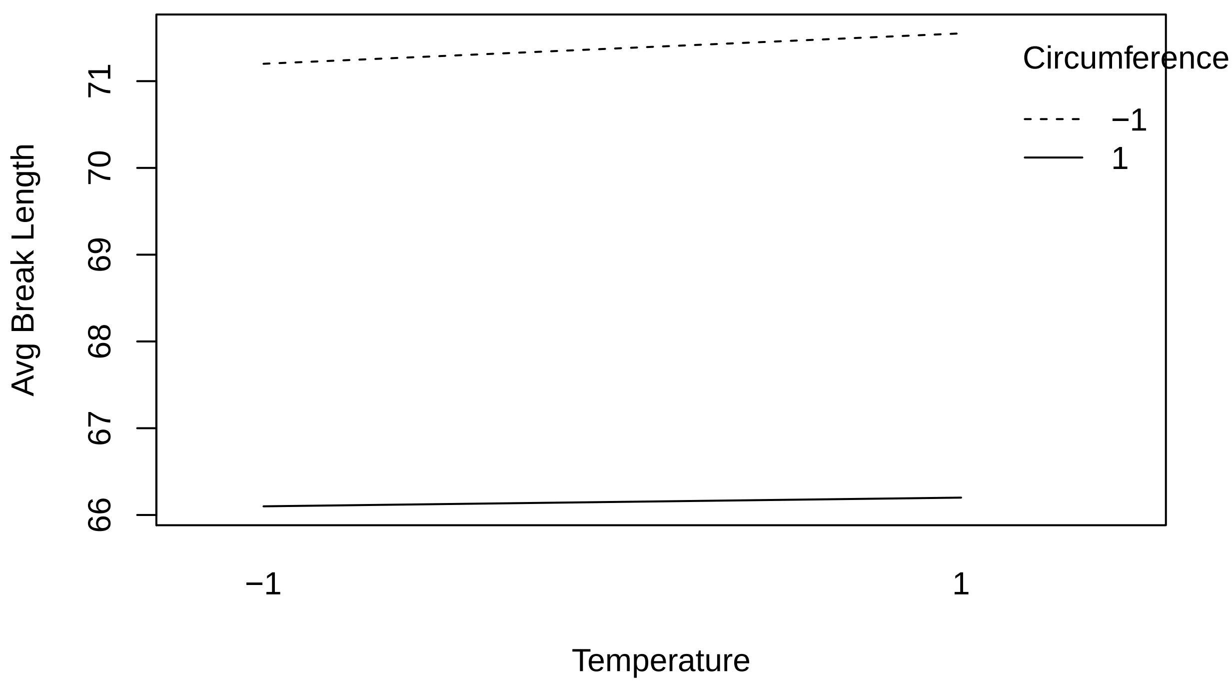


Figure 2: Interaction Plot of Temperature and Circumference

```r
interaction.plot(Avg_Transformed_Factors2$Temperature, Avg_Transformed_Factors2$Brand,
                 Avg_Transformed_Factors2$Avg_Break_Length, type = 'l', xlab = "Temperature",
                 ylab = "Avg Break Length", trace.label = "Brand")
```

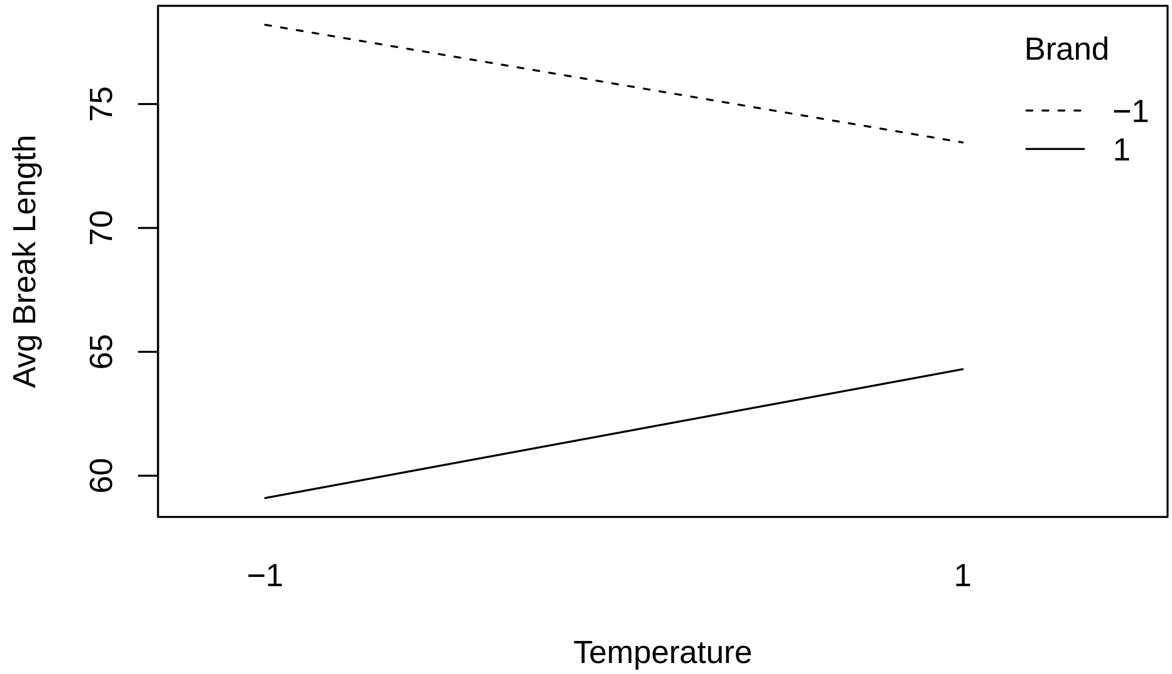


Figure 3: Interaction Plot of Temperature and Brand

```r
interaction.plot(Avg_Transformed_Factors2$Brand, Avg_Transformed_Factors2$Circumference,
                 Avg_Transformed_Factors2$Avg_Break_Length, type = 'l', xlab = "Brand",
                 ylab = "Avg Break Length", trace.label = "Circumference")
```

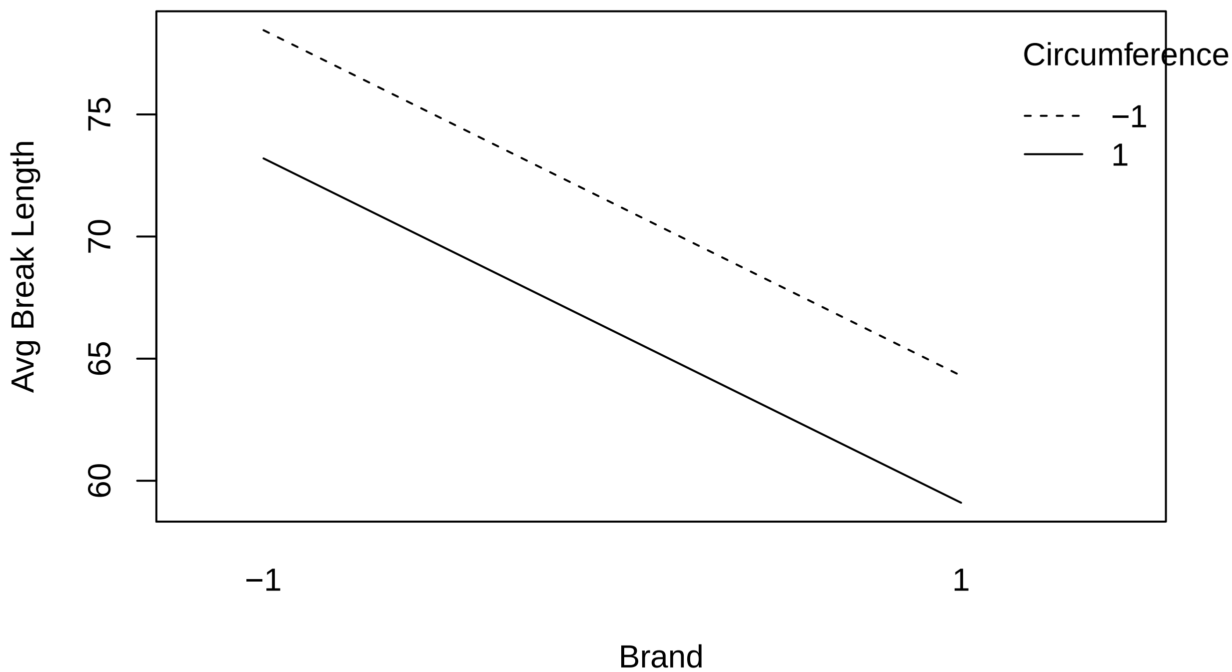


Figure 4: Interaction Plot of Brand and Circumference