THE BEST BROWNIES

Injoong Kim, Sam Lee, Henri Ly

INTRODUCTION

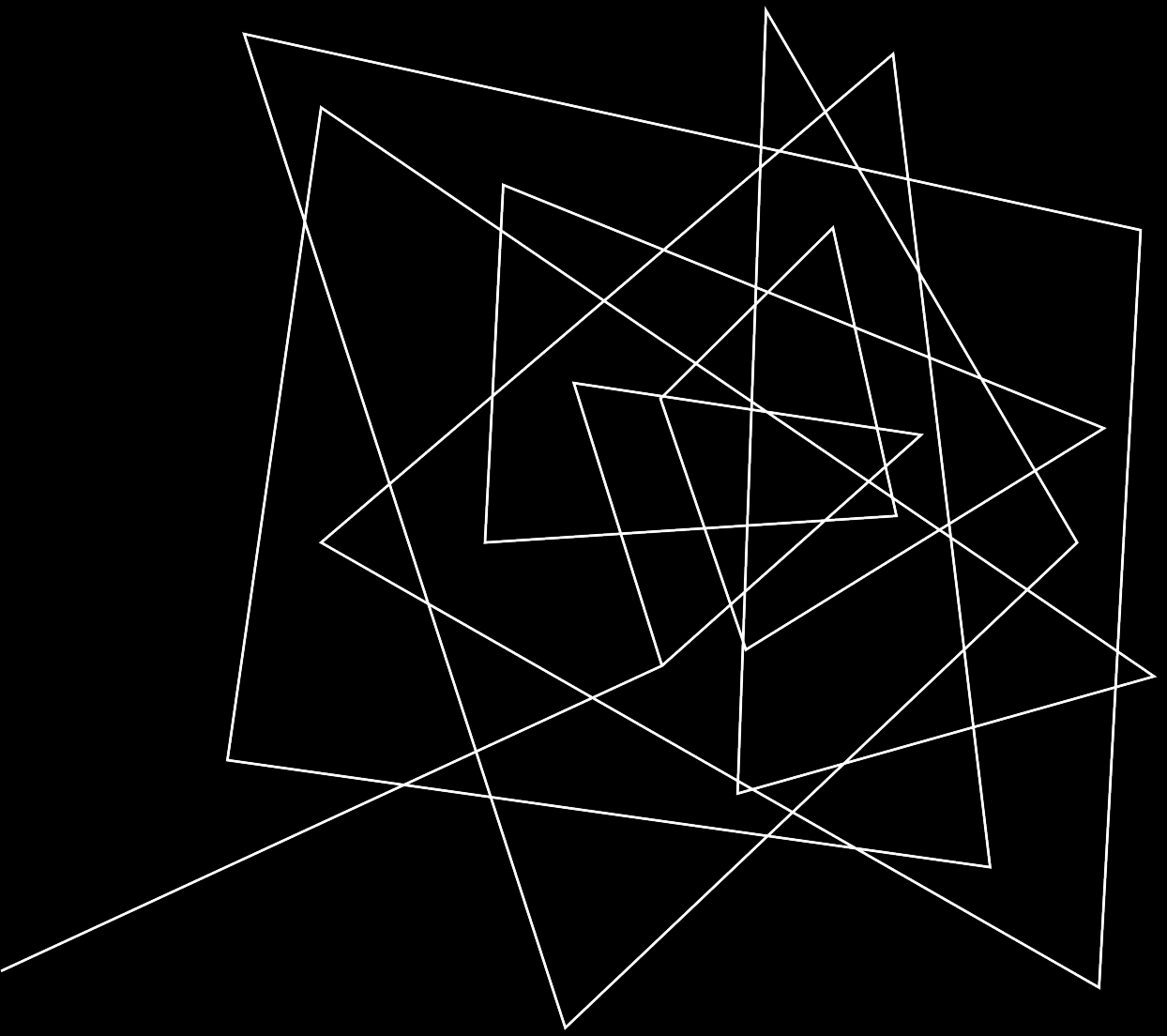
The price of brownie mix can significantly vary, but it remains unclear if it impacts the taste or overall satisfaction of the baked product. For this study, we will examine eight treatments, including male versus female participant, two types of brownie mix, and the time spent in the oven.The aim of this study is to determine the effect of these factors on the taste and overall satisfaction of the brownies.

RESEARCH QUESTIONS

Brownie mix varies wildly in price, but does it really affect taste or satisfaction?

Do people have consistent preferences when it comes to leaving brownies in the oven longer?

Does a person's gender affect which brownie they prefer?



HYPOTHESIS

Null Hypothesis: There is no difference in satisfaction from a brownie with varying levels of gender, time in the oven, or brownie mix used.

Alternative Hypothesis: At least one of the test means for these varying treatments is different.

EXPERIMENTAL DESIGN

- Originally Complete block design: CB[3]
- Three factors
  - Brand - Ghirardelli and Betty Crocker
  - Oven time - Normal and 5 minutes longer
  - Gender
- Switched to BF[2] after factoring out blocking variable (More on this later)
  - Brand
  - Oven time

R Code Used for Randomization

```r
library(tidyverse)

factors = c("Brand", "OvenTime", "Gender")
Brand = c("BettyCrocker", "Ghirardelli")
OvenTime = c("Normal", "Extra")
Gender = c("Male", "Female")
combinations = expand.grid(Brand, OvenTime, Gender)
combinations = combinations %>% setNames(factors)
combinations["ID"] = 1:(length(Brand)*length(OvenTime)*length(Gender))
combinations = combinations[c(4,1:3)]

combinations = combinations %>% as_tibble()
combinations

scale = 10; difference = 1.5

treatments = nrow(combinations)
brownie_scores = c(rep(0, treatments-1), difference)
#We want to detect a different means significantly by 1.5 points
between_var = var(brownie_scores)

within_var = ((scale-(scale/5))/6)^2
#We assume the 99% range is roughly around 8 because
#people will give ratings at least 2 or higher
n = ceiling(power.anova.test(groups=treatments,
                             between.var = between_var,
                             within.var = within_var,
                             sig.level = 0.05,
                             power=0.80)$n)

replications = rep(1:n, max(combinations["ID"]))
order = 1:length(replications)

data = tibble(
  "Order" = order,
  "Person" = rep(1:28, each=4),
  "Treatment" = rep(1:nrow(combinations), n)
)
data = data[sample(1:nrow(data)),] #Randomize the treatments
data = mutate(data,
              Brand=combinations$Brand[data$Treatment],
              OvenTime=combinations$OvenTime[data$Treatment],
              Gender=combinations$Gender[data$Treatment]) %>% arrange(Person)
```

RESULTS (RAW)

1	Order	Person	Treatment	Brand	OvenTime	Gender
2	1	1	1	BettyCrocker	Normal	Male
3	2	1	3	BettyCrocker	Extra	Male
4	3	1	4	Ghirardelli	Extra	Male
5	4	1	2	Ghirardelli	Normal	Male
6	5	2	2	Ghirardelli	Normal	Male
7	6	2	4	Ghirardelli	Extra	Male
8	7	2	1	BettyCrocker	Normal	Male
9	8	2	3	BettyCrocker	Extra	Male
10	9	3	4	Ghirardelli	Extra	Male
11	10	3	1	BettyCrocker	Normal	Male
12	11	3	2	Ghirardelli	Normal	Male
13	12	3	3	BettyCrocker	Extra	Male
14	13	4	1	BettyCrocker	Normal	Male
15	14	4	4	Ghirardelli	Extra	Male
16	15	4	3	BettyCrocker	Extra	Male
17	16	4	2	Ghirardelli	Normal	Male

94	93	24	5	BettyCrocker	Normal	Female
95	94	24	6	Ghirardelli	Normal	Female
96	95	24	8	Ghirardelli	Extra	Female
97	96	24	7	BettyCrocker	Extra	Female
98	97	25	5	BettyCrocker	Normal	Female
99	98	25	7	BettyCrocker	Extra	Female
100	99	25	8	Ghirardelli	Extra	Female
101	100	25	6	Ghirardelli	Normal	Female
102	101	26	6	Ghirardelli	Normal	Female
103	102	26	8	Ghirardelli	Extra	Female
104	103	26	5	BettyCrocker	Normal	Female
105	104	26	7	BettyCrocker	Extra	Female
106	105	27	8	Ghirardelli	Extra	Female
107	106	27	5	BettyCrocker	Normal	Female
108	107	27	6	Ghirardelli	Normal	Female
109	108	27	7	BettyCrocker	Extra	Female
110	109	28	5	BettyCrocker	Normal	Female
111	110	28	8	Ghirardelli	Extra	Female
112	111	28	7	BettyCrocker	Extra	Female
113	112	28	6	Ghirardelli	Normal	Female

Power Analysis - Determining Sample Size

```
15  scale = 10; difference = 1.5
16
17  treatments = nrow(combinations)
18  brownie_scores = c(rep(0, treatments-1), difference)
19  #We want to detect a different means significantly by 1.5 points
20  between_var = var(brownie_scores)
21
22  within_var = ((scale-(scale/5))/6)^2
23  #We assume the 99% range is roughly around 8 because
24  #people will give ratings at least 2 or higher
25  n = ceiling(power.anova.test(groups=treatments,
26                               between.var = between_var,
27                               within.var = within_var,
28                               sig.level = 0.05,
29                               power=0.80)$n)
```

```
     Balanced one-way analysis of variance power calculation

         groups = 8
              n = 13.86214
    between.var = 0.28125
     within.var = 1.777778
      sig.level = 0.05
          power = 0.8

NOTE: n is number in each group

> 14*8
[1] 112
>
```

```
> combinations
# A tibble: 8 x 4
     ID Brand        OvenTime Gender
  <int> <fct>        <fct>    <fct>
1     1 BettyCrocker Normal   Male
2     2 Ghirardelli  Normal   Male
3     3 BettyCrocker Extra    Male
4     4 Ghirardelli  Extra    Male
5     5 BettyCrocker Normal   Female
6     6 Ghirardelli  Normal   Female
7     7 BettyCrocker Extra    Female
8     8 Ghirardelli  Extra    Female
```

EXPERIMENTAL PROCESS

- Pan size
- Temperature (post cooling)
- Temperature (oven)
- Container and utensils used for mixing
- Size of the brownie samples
- Edge vs. Center
- Cooking spray

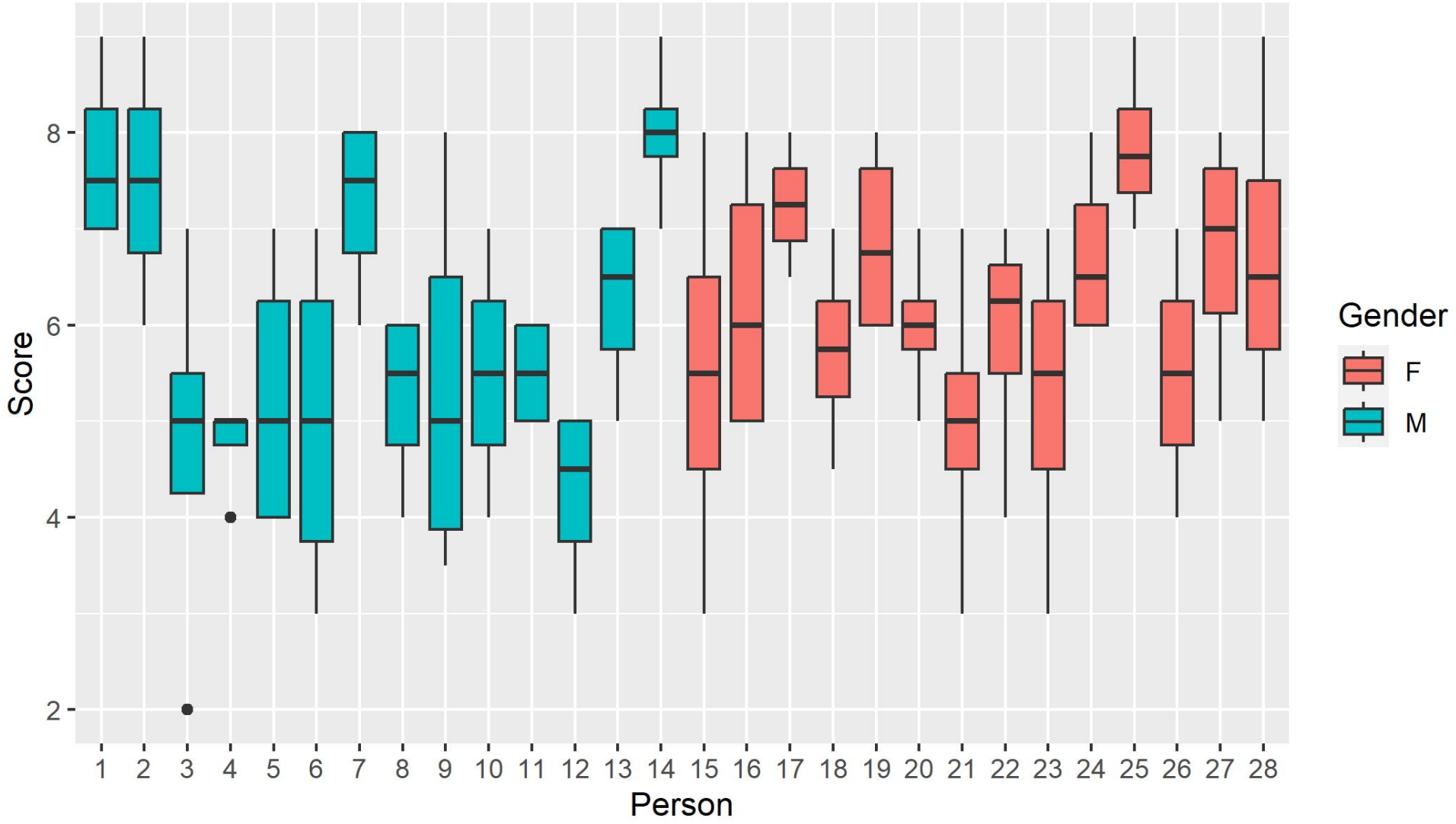
EXPERIMENTAL PROCEDURES

- preheating
- mix preparation (brand, container)
- baking (by brand)
- cooling
- trimming
- portioning

DATA COLLECTION

- 10-point scale for enjoyment
- Randomized trials and treatment orders
- Multiple locations
- Same time period
- Use of a palate cleanser

Distribution of Scores Across People



TRANSFORMATION OF DATA

Can we assume the population variances are equal?

```
85  stds = data %>% group_by(Person) %>%
86    summarize(
87      std = sd(Score)
88    )
89
90  stds
91
92  #This is greater than 2
93  max(stds$std)/min(stds$std)
```

```
> max(stds$std)/min(stds$std)
[1] 4.163332
```

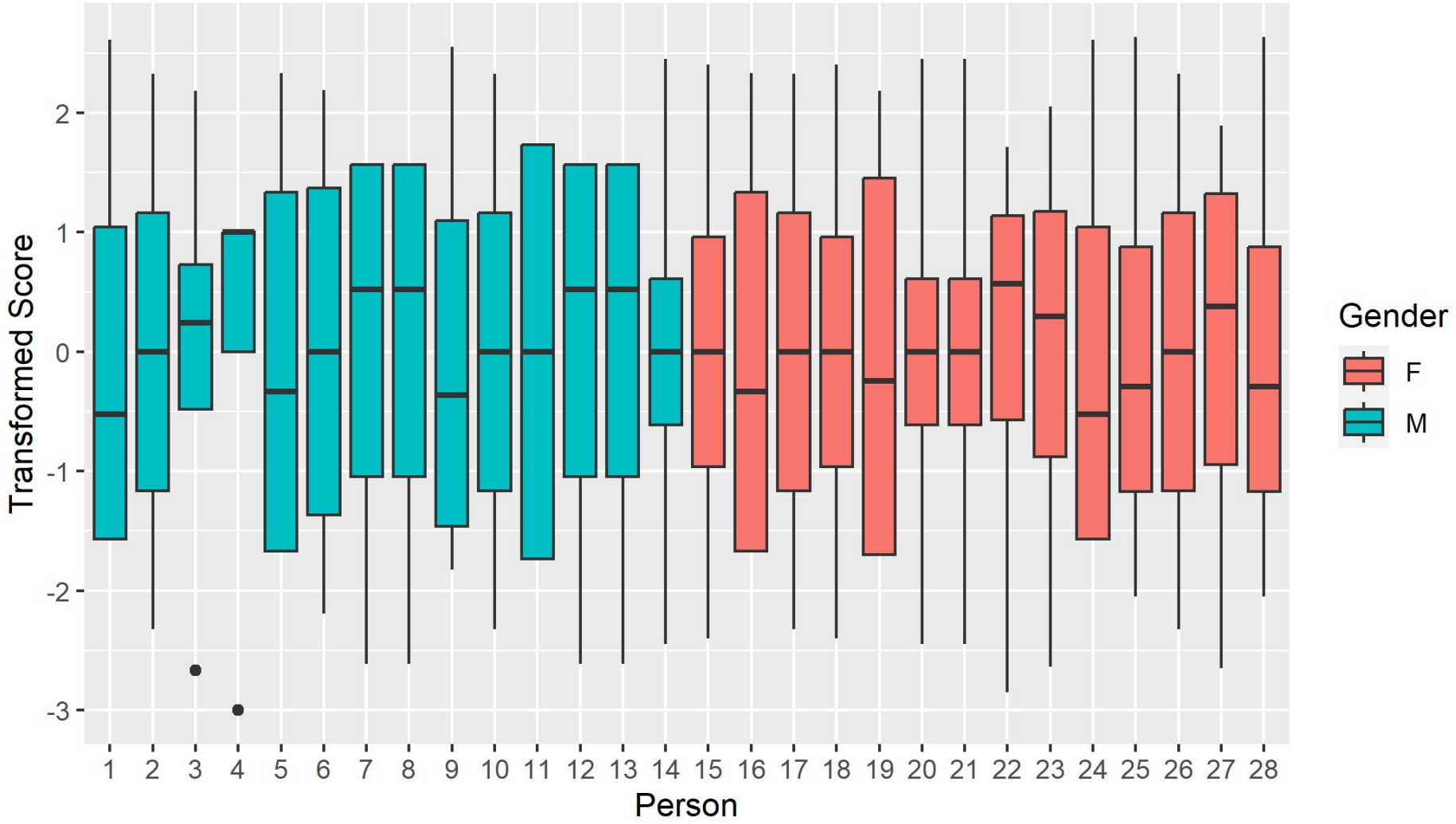
A log transformation does not work with our data

```
#This does not work with our data
log_stds = data %>% group_by(Person) %>%
  summarize(
    std = sd(log(Score))
  )

#This is also greater than 2
max(log_stds$std)/min(log_stds$std)
```

```
> #This is also greater than 2
> max(log_stds$std)/min(log_stds$std)
[1] 6.021513
>
```

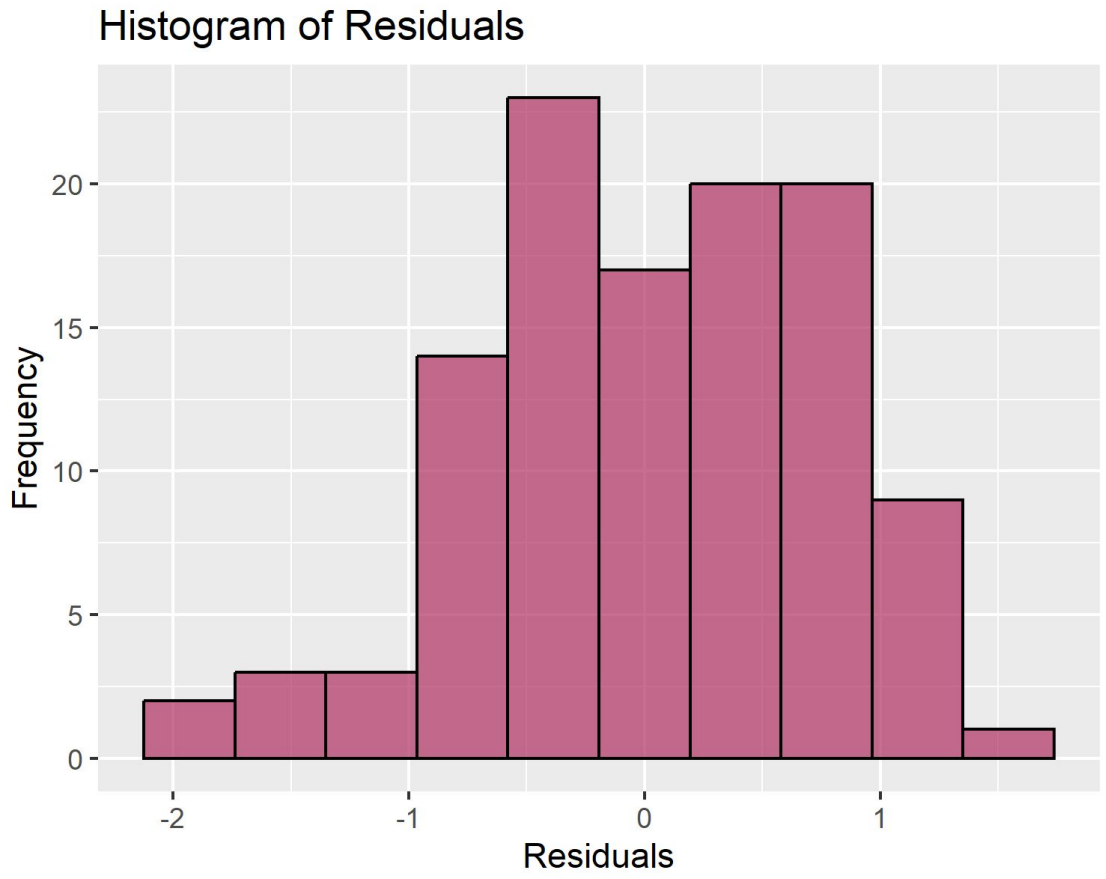
Z TRANSFORMATION

```
105  #Z Transformation
106  data = data %>% group_by(Person) %>% mutate(
107    Z = (Score-mean(Score))/(sd(Score))
108  )
```



STATISTICAL MODEL

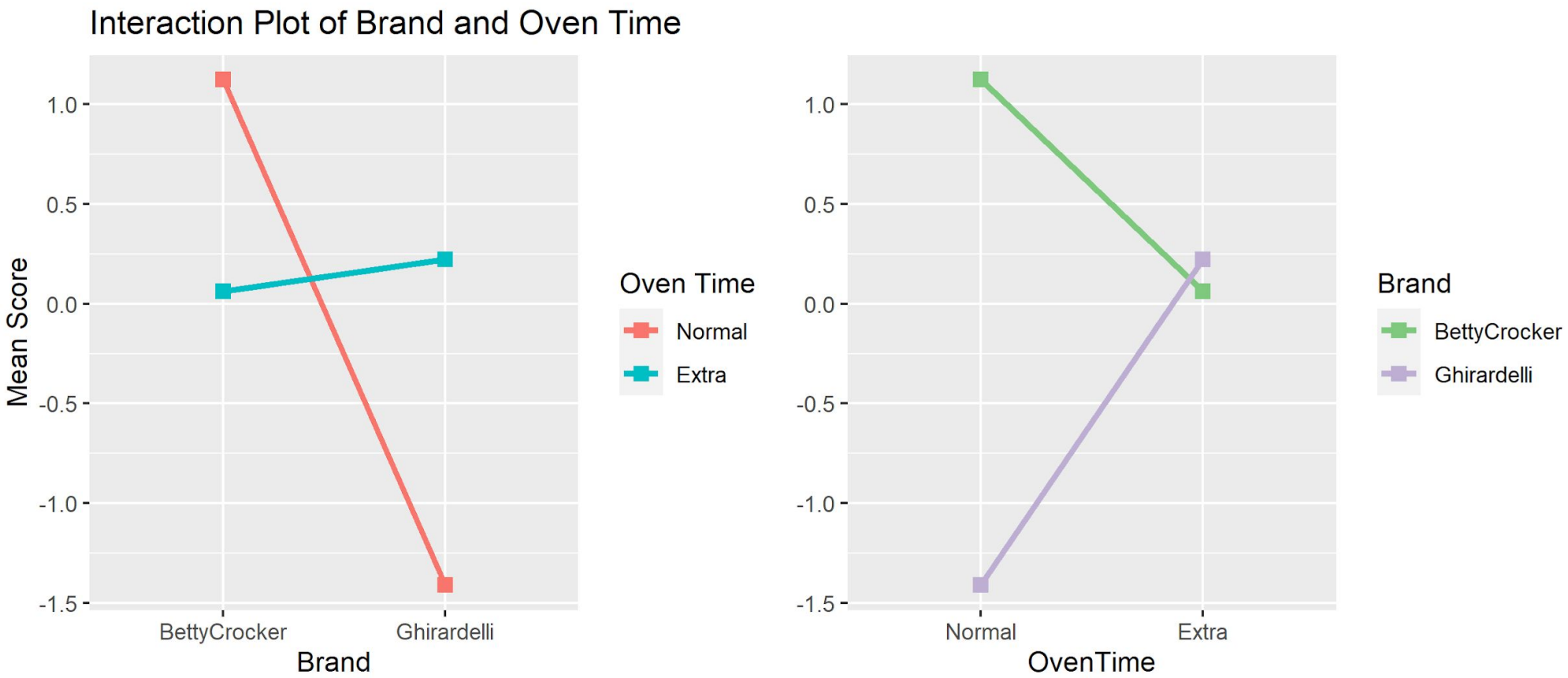
```
fit <- lm(Z ~ Brand + OvenTime + Brand:OvenTime, data = data)

results = anova(fit)
results
```



ANOVA TABLE

	Df	Sum of Squares	Mean Sq	F value	P(>F)
Brand	1.0000	9.8658	9.8658	17.5067	0.0001
OvenTime	1.0000	0.5674	0.5674	1.0068	0.3179
Brand:OvenTime	1.0000	12.7044	12.7044	22.5440	0.0000
Residuals	108.0000	60.8624	0.5635		

INTERACTION PLOTS

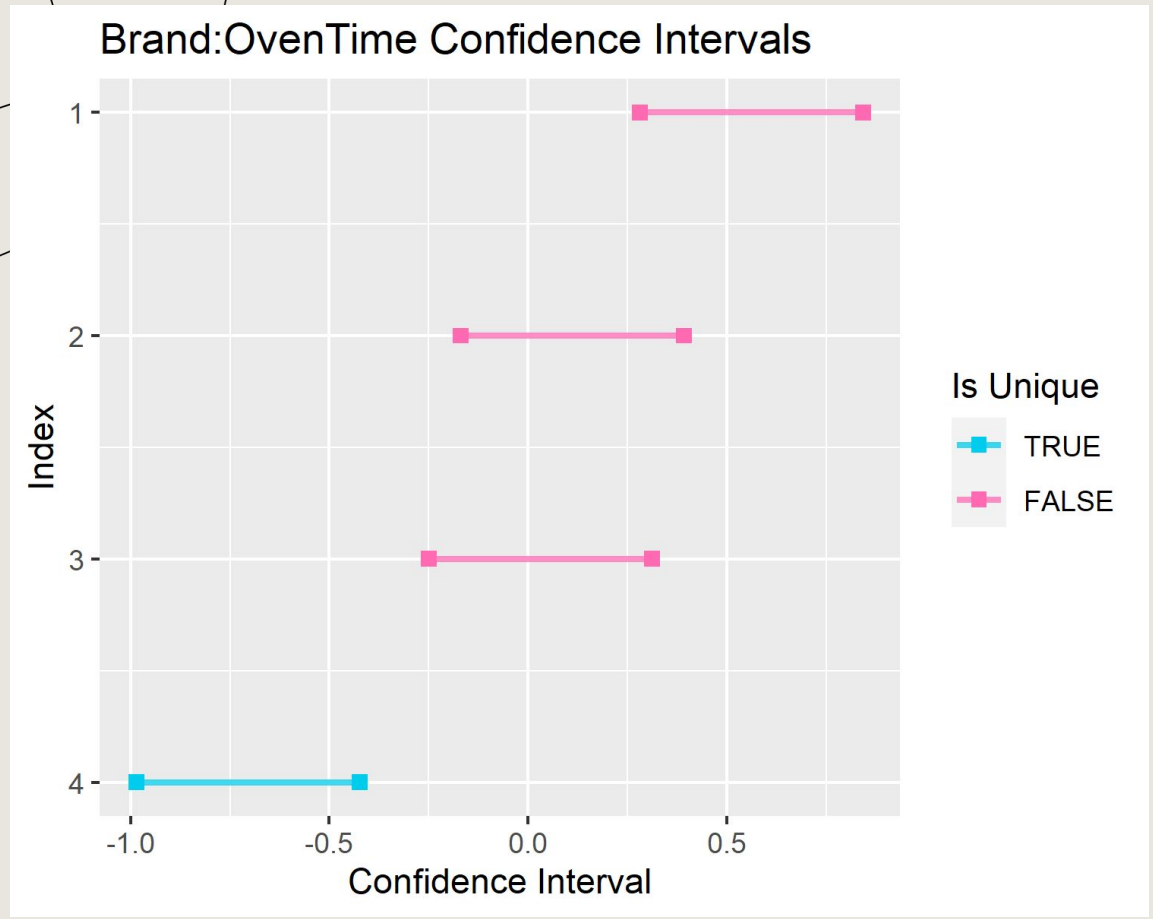


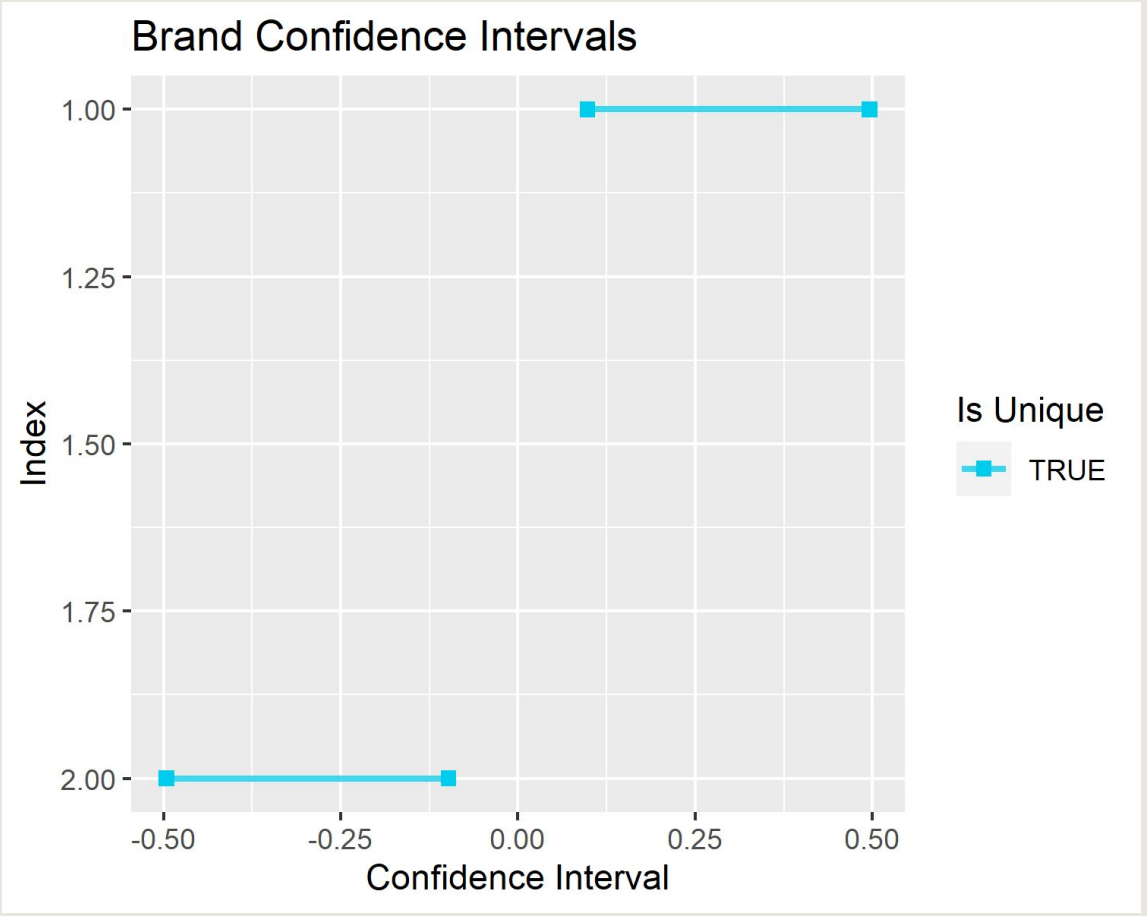
CONFIDENCE INTERVALS

	Brand	Mean	SE	df	lower.CL	upper.CL
1	BettyCrocker	0.2968	0.1003	108	0.0980	0.4956
2	Ghirardelli	-0.2968	0.1003	108	-0.4956	-0.0980

	Brand	OvenTime	Mean	SE	df	lower.CL	upper.CL
1	BettyCrocker	Normal	0.5624	0.1419	108	0.2812	0.8436
2	Ghirardelli	Extra	0.1112	0.1419	108	-0.1700	0.3924
3	BettyCrocker	Extra	0.0312	0.1419	108	-0.2500	0.3124
4	Ghirardelli	Normal	-0.7048	0.1419	108	-0.9860	-0.4236

CONFIDENCE INTERVALS





INTERPRETATION

- Failed to reject the null hypothesis regarding oven time. We did not observe a significant difference in the preference for baking time in the oven.

- Rejected the null hypothesis regarding brand. We found significant evidence that there is a difference in the preference for Betty Crocker brownies and Ghirardelli.

- Rejected the null hypothesis regarding the interaction between oven time and brand. We found significant evidence to conclude that at least one group (Ghirardelli brownies baked at normal time in this case) differs from the rest.

CONCLUSION

The findings may be useful for consumers and bakers who want to make a decision about which brownie mix to purchase and how long to bake their brownies.

FUTURE STUDIES

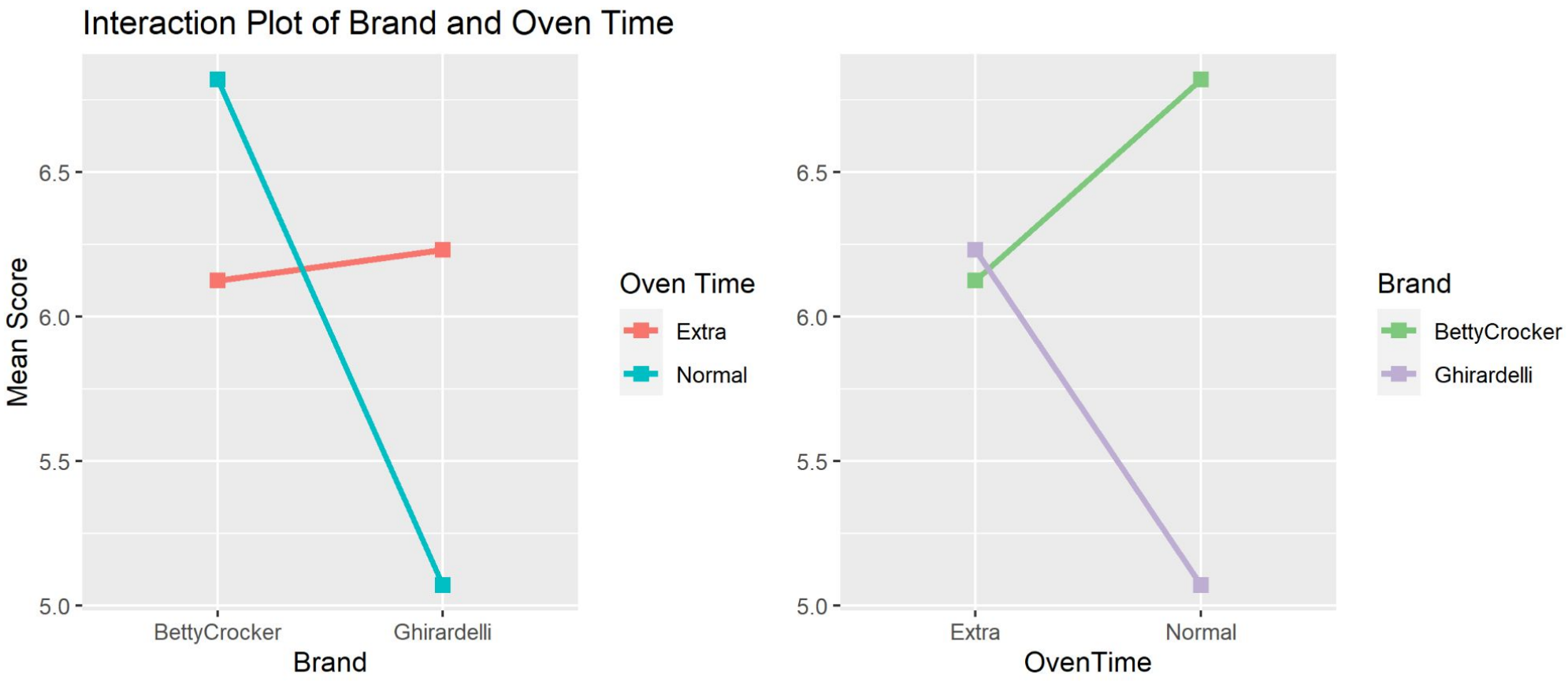
- More types of brownie mix
- More levels for oven time
- Different types of palate cleansers
- Different pan sizes / material (glass)
- Different time period (morning, afternoon, evening)

APPENDIX

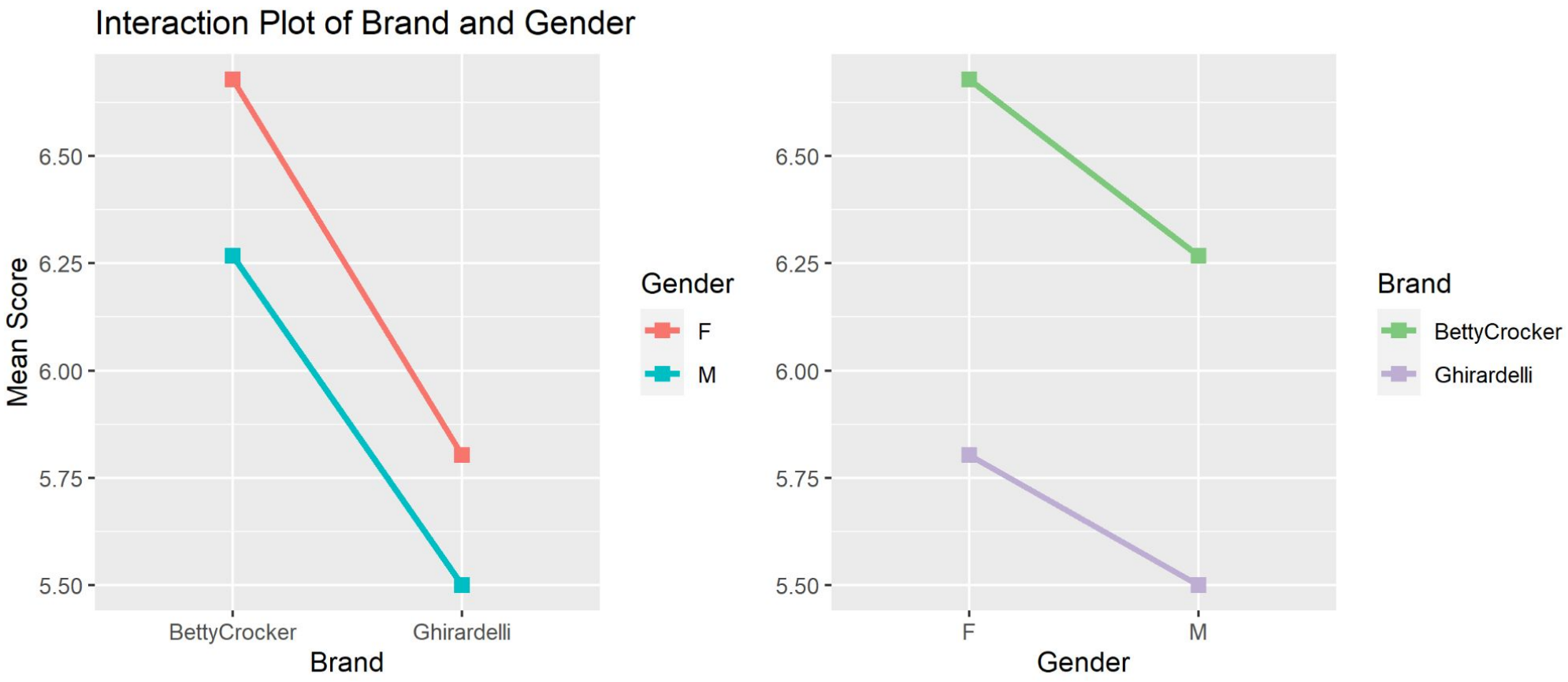
Had we done the analysis without transforming the data…

This relies on the critical assumption the population variances are equal notwithstanding the deviation that we observed between the scores.

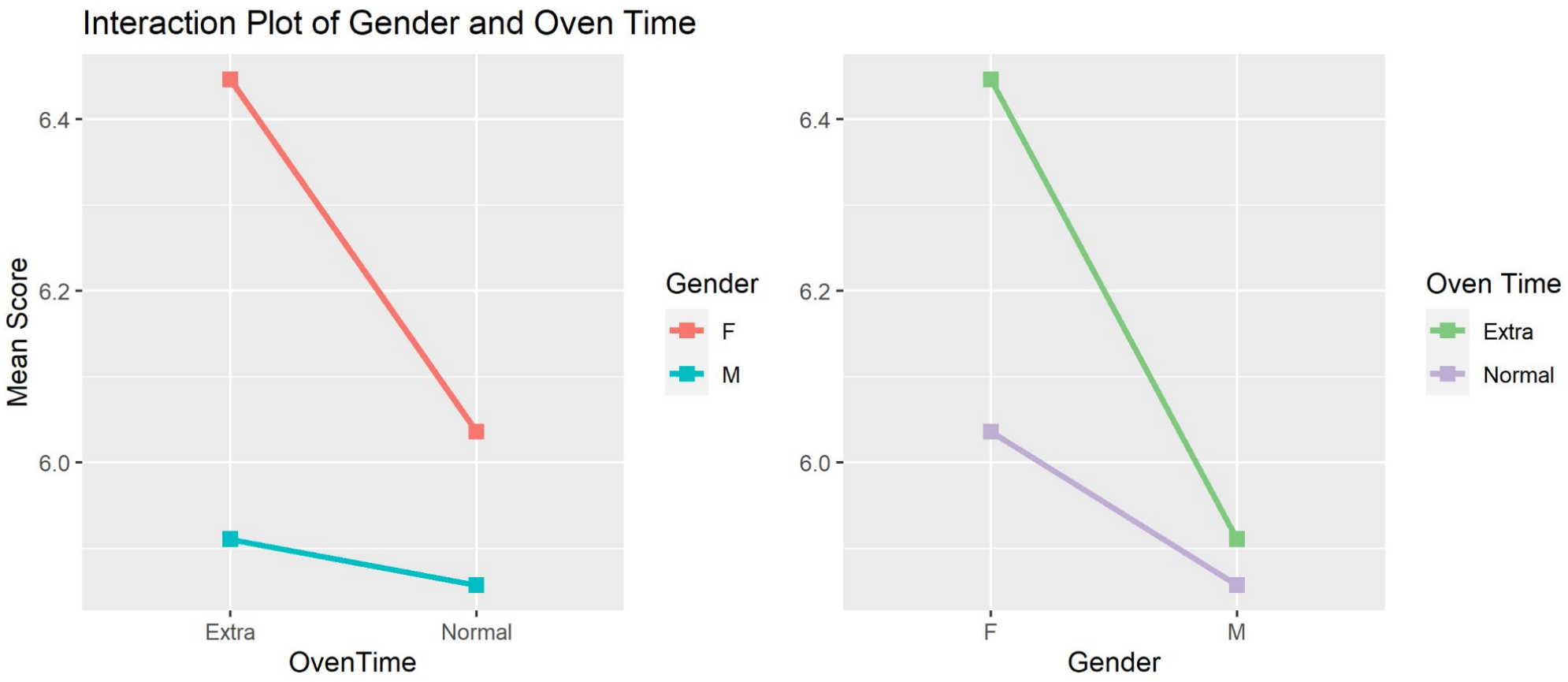
INTERACTION PLOTS



INTERACTION PLOTS

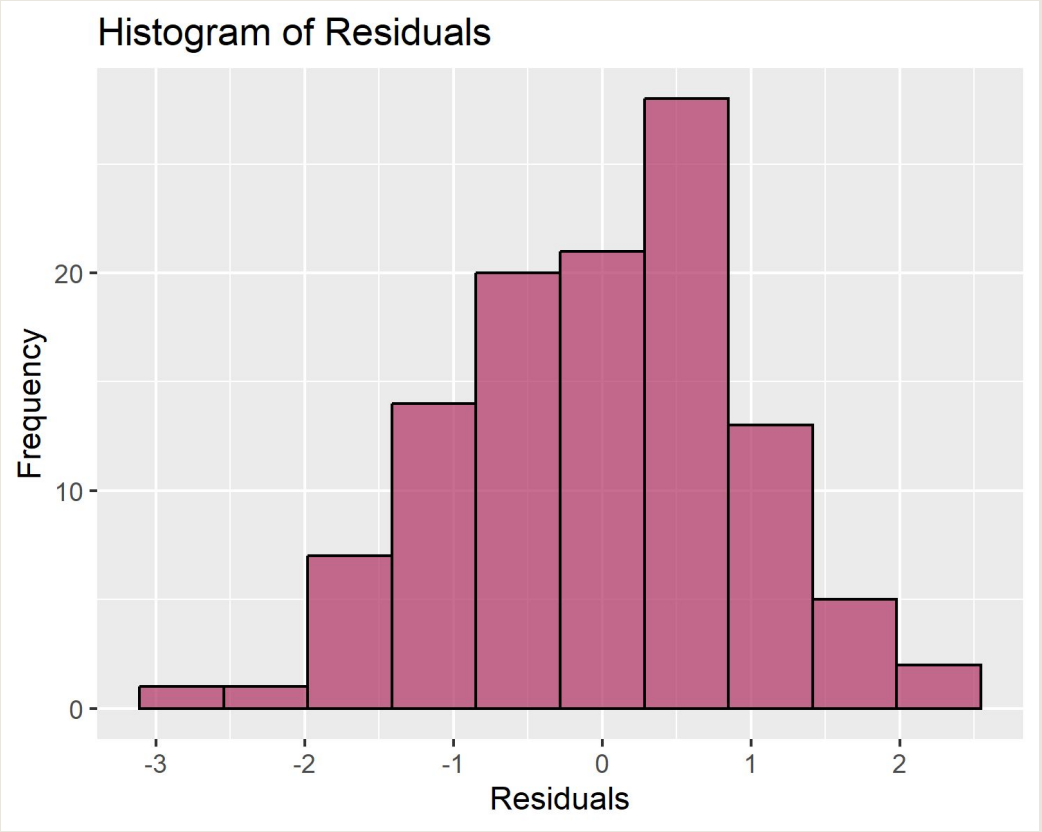


INTERACTION PLOTS

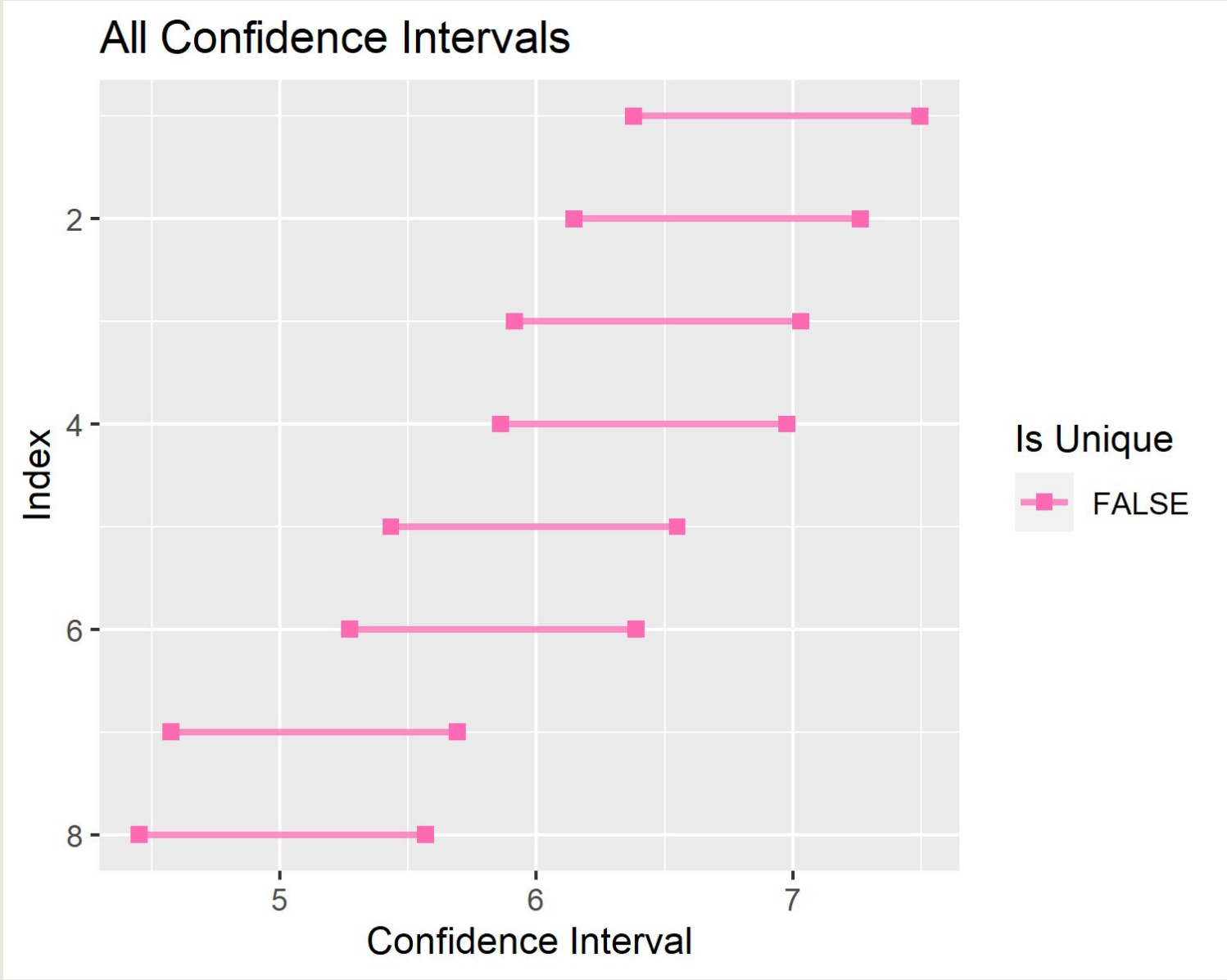


STATISTICAL MODEL

```
fit = lm(Score ~ Brand + OvenTime + Gender + as.factor(Person) +
         Brand:OvenTime + Gender:Brand + Gender:OvenTime, data = data)
results = anova(fit)
results
```



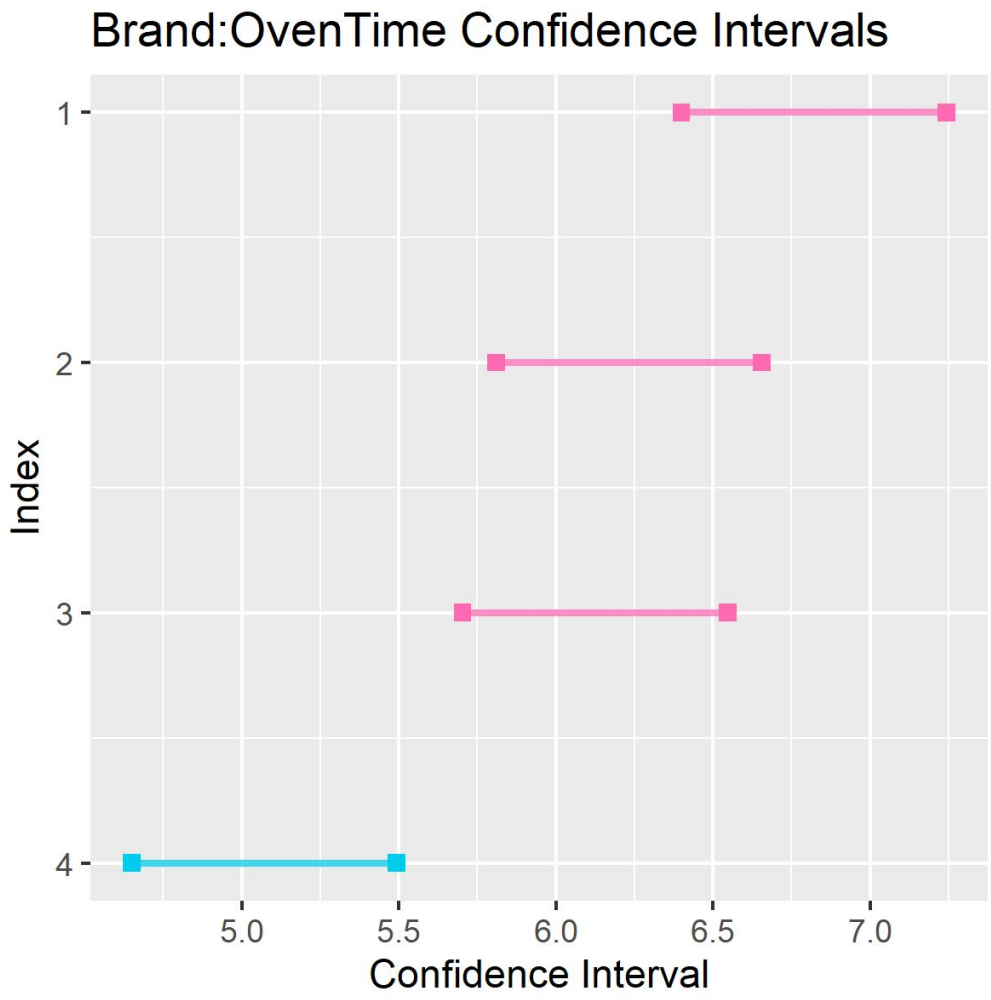
CONFIDENCE INTERVALS

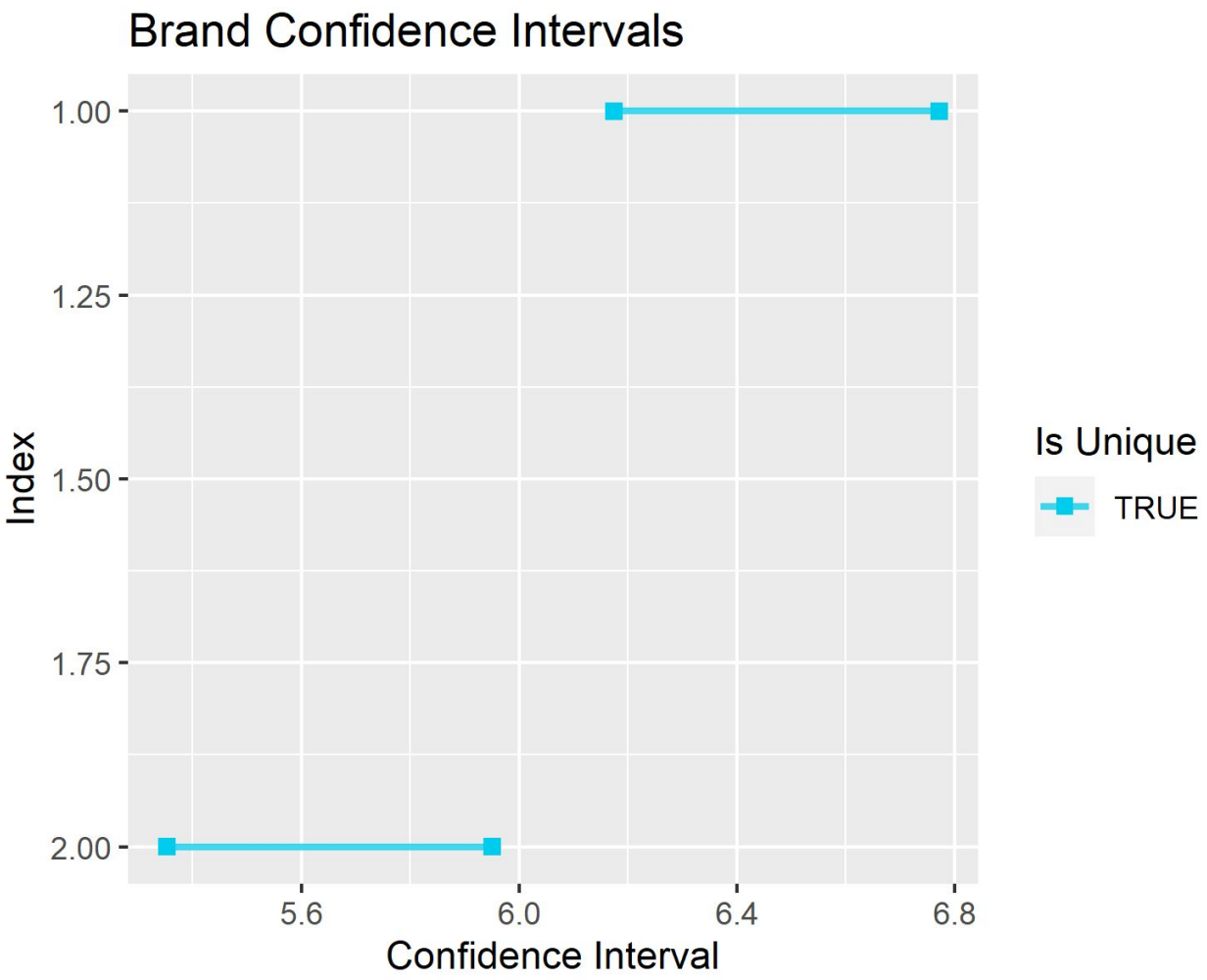


ALL CONFIDENCE INTERVALS

	Brand	OvenTime	Gender	Mean	SE	df	lower.CL	upper.CL
1	BettyCrocker	Normal	F	6.9375	0.2805	79	6.3791	7.4959
2	BettyCrocker	Normal	M	6.7054	0.2805	79	6.1470	7.2638
3	Ghirardelli	Extra	F	6.4732	0.2805	79	5.9148	7.0316
4	BettyCrocker	Extra	F	6.4196	0.2805	79	5.8612	6.9780
5	Ghirardelli	Extra	M	5.9911	0.2805	79	5.4327	6.5495
6	BettyCrocker	Extra	M	5.8304	0.2805	79	5.2720	6.3888
7	Ghirardelli	Normal	F	5.1339	0.2805	79	4.5755	5.6923
8	Ghirardelli	Normal	M	5.0089	0.2805	79	4.4505	5.5673

SIGNIFICANT CONFIDENCE INTERVALS





SIGNIFICANT CONFIDENCE INTERVALS

	Brand	Mean	SE	df	lower.CL	upper.CL
1	BettyCrocker	6.4732	0.1500	79	6.1747	6.7717
2	Ghirardelli	5.6518	0.1500	79	5.3533	5.9503

	Brand	OvenTime	Mean	SE	df	lower.CL	upper.CL
1	BettyCrocker	Normal	6.8214	0.2121	79	6.3993	7.2435
2	Ghirardelli	Extra	6.2321	0.2121	79	5.8100	6.6543
3	BettyCrocker	Extra	6.1250	0.2121	79	5.7029	6.5471
4	Ghirardelli	Normal	5.0714	0.2121	79	4.6493	5.4935

ANOVA TABLE

	Df	Sum of Squares	Mean Sq	F value	P(>F)
Brand	1	18.8929	18.8929	15.0031	0.0002
Oven Time	1	1.5089	1.5089	1.1983	0.2770
Gender	1	3.5714	3.5714	2.8361	0.0961
Person	26	115.4911	4.4420	3.5274	0.0000
Brand:OvenTime	1	24.1429	24.1429	19.1721	0.0000
Brand:Gender	1	0.0804	0.0804	0.0638	0.8012
OvenTime:Gender	1	0.8929	0.8929	0.7090	0.4023
Residuals	79	99.4821	1.2593		