Two Factor Experiment

Deming

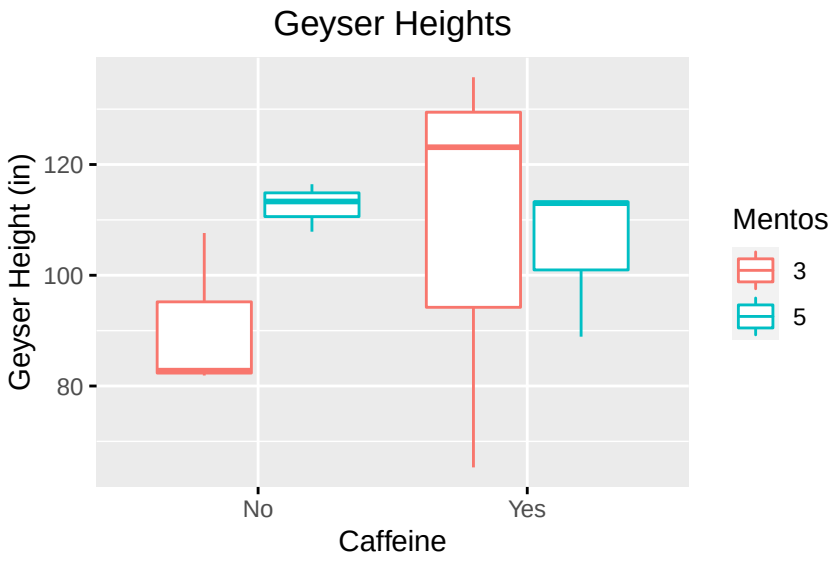
Introduction and Data

This particular experiment was conducted to determine the effect of caffeine and number of mentos on the height of a diet coke and mentos geyser. The response variable was the height of the geyser, and each of the factors (caffeine and number of mentos) had two levels: `Caffeine`, yes and no, and `Mentos`, 3 or 5. The researchers believe that increasing the number of mentos would increase the geyser height, but there may also be an interaction between the number of mentos and caffeine status. Geyser height is difficult to measure, so the explosion was recorded using a webcam to be played back frame-by-frame to measure the peak of the geyser. The experiment was randomized, and each because each of the two factors had two levels, this was a 2x2 factorial experiment. Three replicates were performed, leading to twelve total observations of geyser heights.

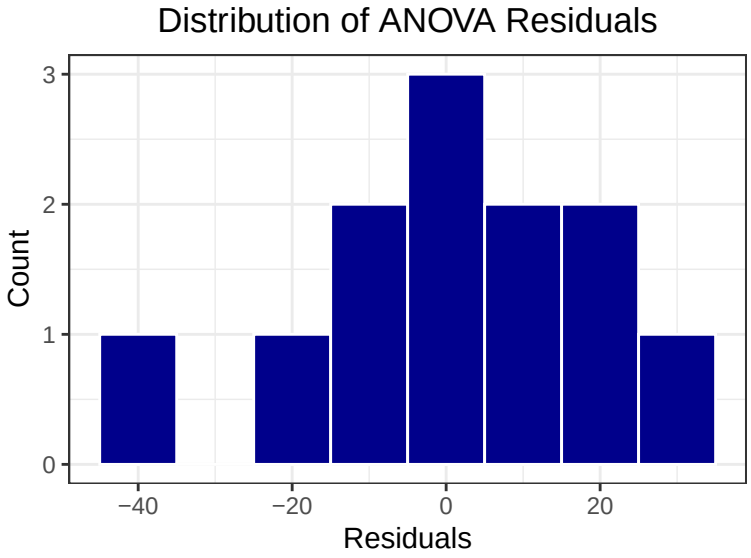
experiment	Caffeine	Mentos	Height
1	Yes	3	123.12
2	No	3	82.77
3	Yes	5	113.02
4	No	3	107.63
5	No	5	113.32
6	Yes	5	88.91
7	Yes	3	65.30
8	No	5	116.44
9	Yes	5	113.49
10	No	5	107.87
11	No	3	81.90
12	Yes	3	135.76

Exploratory Data Analysis

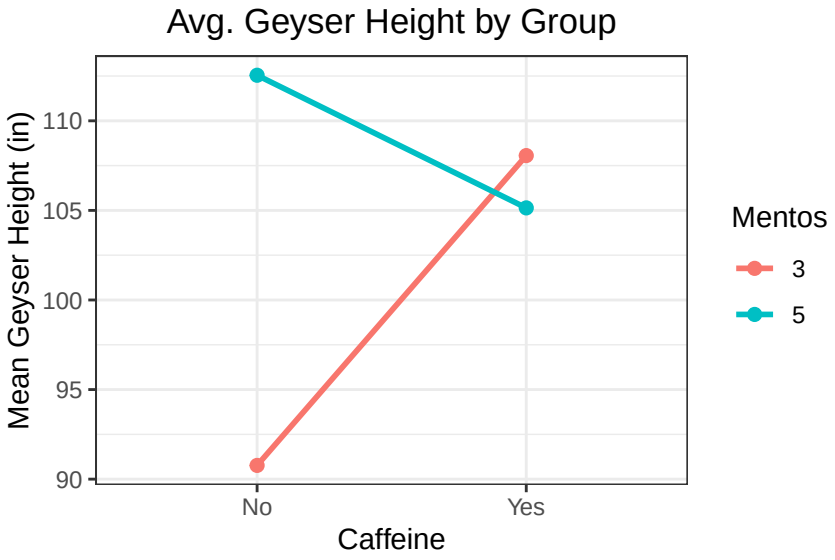
Because this experiment is particularly difficult to perform under controlled conditions, we first wanted to check whether there is any evidence of violating model assumptions. Because this is a two-factor experiment with quantitative response, we will perform a two-way ANOVA test. The assumptions of this test are that there is a constant variance between each of the four Caffeine/Mentos groups, that there is independence between observations, and that the distribution of heights (or residuals) is normal for each group. The following side-by-side boxplots show that there is certainly a difference in the spread of the geyser heights within each factor level combination. However, in order to determine whether this difference is significant, we conducted a Levene test for homogeneity of variance across groups. The null hypothesis of this test is that the variances are equal across groups, and the alternative is that the variances are significantly different across groups. We calculated a p-value of 0.571, which is not significant at the .05 level and leads us to fail to reject the hypothesis that the variances are equal across groups. The equal variance assumption appears to be met.



The final two assumptions are independence and normality. We can safely assume that each trial of geyser heights is independent of the previous trial (one geyser height measurement likely does not affect the height of other geysers), so we feel confident in saying that the independence assumption is met. Lastly, we checked the normality assumption using a Shapiro-Wilk test for normality of residuals. The null hypothesis of this test is that the data (residuals in our case) are normally distributed, and the alternative is that they are not normally distributed. The p-value of the test was 0.589. We also plotted the residuals of the ANOVA model, which we will desribe in the next section. This histogram does not appear to have very significant skewness. Based on the Shapiro-Wilk test and histogram, I believe we can safely say that the residuals are approximately normally distributed and continue with our analysis.



We also were interested to se whether there is an interaction between the soda's caffeine status and the number of mentos used in the trial. The interaction plot below seems to imply that the effect of caffeine status on geyser height depends on the number of mentos used in the trial. For example, with 5 mentos, the average geyser height seems to decrease between the caffeine free and caffiene diet cokes, while the average height increases for the 3 mentos group between the caffeine free and non-caffeine free sodas.



A table of the within-group means and standard deviations is provided below. Of note is that there are three observations of each caffeine/mentos combination (which we discussed in the first section), and this means that we have a balanced design and do not need to worry about differing sample sizes in our analysis.

Caffeine	Mentos	Count	Mean	SD
No	3	3	90.77	14.61
No	5	3	112.54	4.34
Yes	3	3	108.06	37.57
Yes	5	3	105.14	14.06

Analysis

We want to find out if there is a statistically significant interaction between `Caffeine` and `Mentos`. To do this, we will test the following hypotheses:

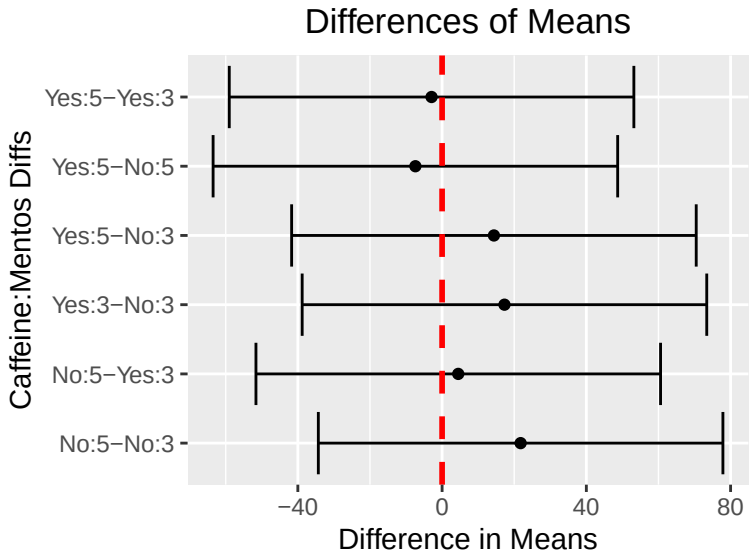
$$H_o: \textit{There is no interaction between Caffeiene and Number of Mentos}$$
$$H_a: \textit{There is an interaction between Caffeiene and Number of Mentos}$$

The following table shows the results of the ANOVA test stated above. We can see the effect of the interaction between the two factors on the third line of this table. The p-value is 0.348, which is obviously larger than the .05 significance level and causes us to fail to reject the null hypothesis. There is not significant evidence of an interaction between caffeine and the number of mentos on the effect of geyser height. In other words, the observed effect of the interaction was not far away enough from what we would expect to see if there truly was not an interaction, so we cannot conclusively say that there is a significant interaction between the two factors.

```
##                 Df Sum Sq Mean Sq F value Pr(>F)
## Caffeine         1     73    73.4   0.159  0.700
## Mentos           1    267   266.7   0.579  0.468
## Caffeine:Mentos  1    457   457.4   0.994  0.348
## Residuals        8   3682   460.3
```

A table of pairwise confidence intervals for each the difference in mean geyser height for each combination of `Caffeine` and `Mentos` is shown below. We can see that each interval contains zero, further supporting our previous conclusion of a non-significant interaction. If there was a signficant interaction, we would expect one of the intervals below to not contain zero. A plot of these confidence intervals is also displayed below.

	diff	lwr	upr
Yes:3-No:3	17.293	-38.803	73.390
No:5-No:3	21.777	-34.320	77.873
Yes:5-No:3	14.373	-41.723	70.470
No:5-Yes:3	4.483	-51.613	60.580
Yes:5-Yes:3	-2.920	-59.017	53.177
Yes:5-No:5	-7.403	-63.500	48.693



The main effects of this model, which are the separate effects of mentos and caffeine on geyser height, do not tell the entire story. Interactions between factors imply that the effect of one factor actually depends on one or more factor present in the experiment. When one only considers main effects of factors, we are unable to see these interactions, leading to incorrect conclusions about the experiment. It should also be noted that because of the small sample size of each of the replicate groups (3), this data is prone to variability in the results–as we saw with the large difference in standard deviation between each of the four groups. Based on the interaction plot generated in the Exploratory Data Analysis section, we would have expected to see a significant interaction between `Caffeine` and `Mentos`, but the p-value from the ANOVA test did not support the intuition. It could be that the sample size was just too small to detect the significant interaction, so I would recommend more replicates in future analyses.