ST 518 Project

Rachel Hencher, Dane Korver, Mohammed Mira

2022-12-03

Contents

Executive Summary

A one or two paragraph summary that includes a description of the experiment, significant results (including any interesting numerical results), and any conclusions you draw. The reader should be able to glean all the important aspects of your work from the executive summary. Effectively and succinctly convey objectives, summary of experimental design, and results and conclusions drawn from experiment.

Introduction

The purpose of this experiment is to investigate specific factors and their effect on the amount of time it takes to dissolve a cold medicine tablet in water. The data is from an "Effervescent Experiment" designed to compare dissolving times of two different brands of tablets (name brand and store brand) at three different equally spaced water temperatures (6°C, 23°C, and 40°C). The run order number, as well as whether or not the sample was stirred, were also recorded and are investigated in the analysis portions of this report.

Below, we have a brief look at the first 10 of 48 rows of data in order to begin to gain an understanding of the data set we are working with.

Table 1: First 10 Rows of Effervescence Data

Brand	Temp	Stirred	Order	Time
name	6	yes	8	77.21547
name	23	yes	3	75.37855
name	40	yes	7	68.08492
store	6	yes	1	77.87371
store	23	yes	2	66.38436
store	40	yes	18	59.82388
name	6	yes	9	75.94293
name	23	yes	4	69.08937
name	40	yes	10	64.45156
store	6	yes	12	77.33947

Experimental Design

The experiment carried out was a complete block design where $b = 2$ blocks (by stirred status) were selected with $n = 4$ observations on each of the treatment combinations in each block. In Block I, the water was stirred using a magnetic stirring plate at 350 revolutions per minute, whereas in Block II, the water was not stirred. The time for the tablet to dissolve was measured from the moment the tablet was dropped into the water to the time the tablet was completely dissolved, and each tablet was dropped from a fixed height into 60mL of water. The observation was taken as an average of the times as measured by four experimenters and was recorded, along with the run order for each observation. The primary effects being investigated are `Brand` and `Temp`, though `Stirred` and `Order` are also considered in the analysis portion below.

Exploratory Analysis

Summary statistics for each variable can be seen below. For the `Brand`, `Temp`, `Stirred`, and `Order` variables, we can see counts for each level. For the `Time` variable, we can see a five-number summary for the variable.

Table 2: Summary Stats for Variables

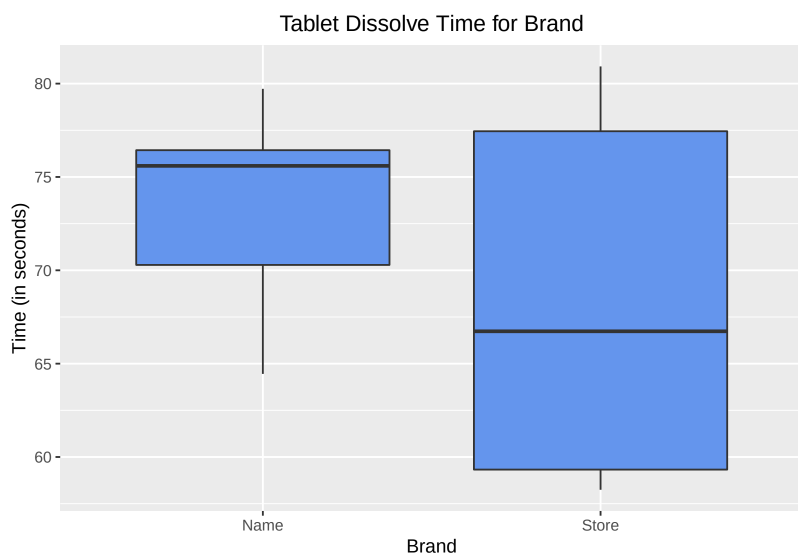
Brand	Temp	Stirred	Order	Time
name :24	6 :16	no :24	1 : 1	Min. :58.24
store:24	23:16	yes:24	2 : 1	1st Qu.:66.09
NA	40:16	NA	3 : 1	Median :70.92
NA	NA	NA	4 : 1	Mean :70.77
NA	NA	NA	5 : 1	3rd Qu.:76.93
NA	NA	NA	6 : 1	Max. :80.92
NA	NA	NA	(Other):42	NA

The summary statistics for the `Time by Group Table` below shows a decrease in mean dissolving time as temperature increases regardless of `Brand` or `Stirred` status. "Name Brand" generally had a higher mean dissolving time, but occasionally, "Store Brand" had a higher mean dissolve time. The same was true for when the water was not stirred, both instances at the 6 degrees Celsius. There were also some unusually high standard deviations for "Name Brand" at 23 and 40 degrees with their standard deviations at 2.64 and 1.6, respectively. In both of these instances, the water was stirred. "Store Brand" no stirred at 6 degrees also had a higher standard deviation of 1.3 whereas, overall, the standard deviations tended to range between 0.2 and 0.77.

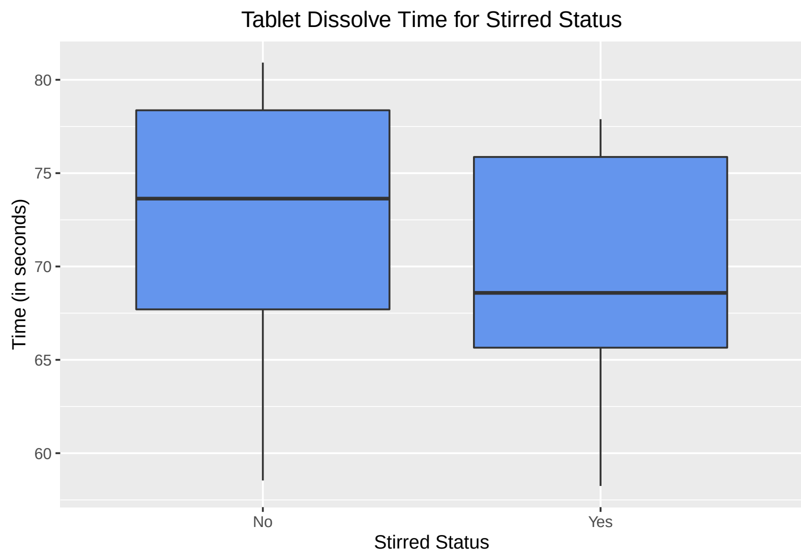
Table 3: Summary Stats for Time by Group

Brand	Stirred	Temp	Mean	SD	Min	Q1	Median	Q3	Max
name	no	6	78.99	0.64	78.15	78.80	79.04	79.24	79.72
name	no	23	76.36	0.33	76.07	76.20	76.28	76.43	76.83
name	no	40	70.56	0.50	70.00	70.29	70.51	70.78	71.20
name	yes	6	76.20	0.68	75.81	75.83	75.89	76.26	77.22
name	yes	23	72.69	2.64	69.09	71.82	73.15	74.02	75.38
name	yes	40	65.85	1.60	64.45	64.87	65.44	66.42	68.08
store	no	6	79.49	1.30	77.78	79.02	79.63	80.10	80.92
store	no	23	67.52	0.45	67.08	67.14	67.52	67.90	67.93
store	no	40	58.96	0.35	58.54	58.77	58.99	59.19	59.33
store	yes	6	77.34	0.77	76.24	77.07	77.61	77.88	77.89
store	yes	23	66.19	0.20	65.93	66.09	66.23	66.33	66.38
store	yes	40	59.13	0.66	58.24	58.91	59.22	59.43	59.82

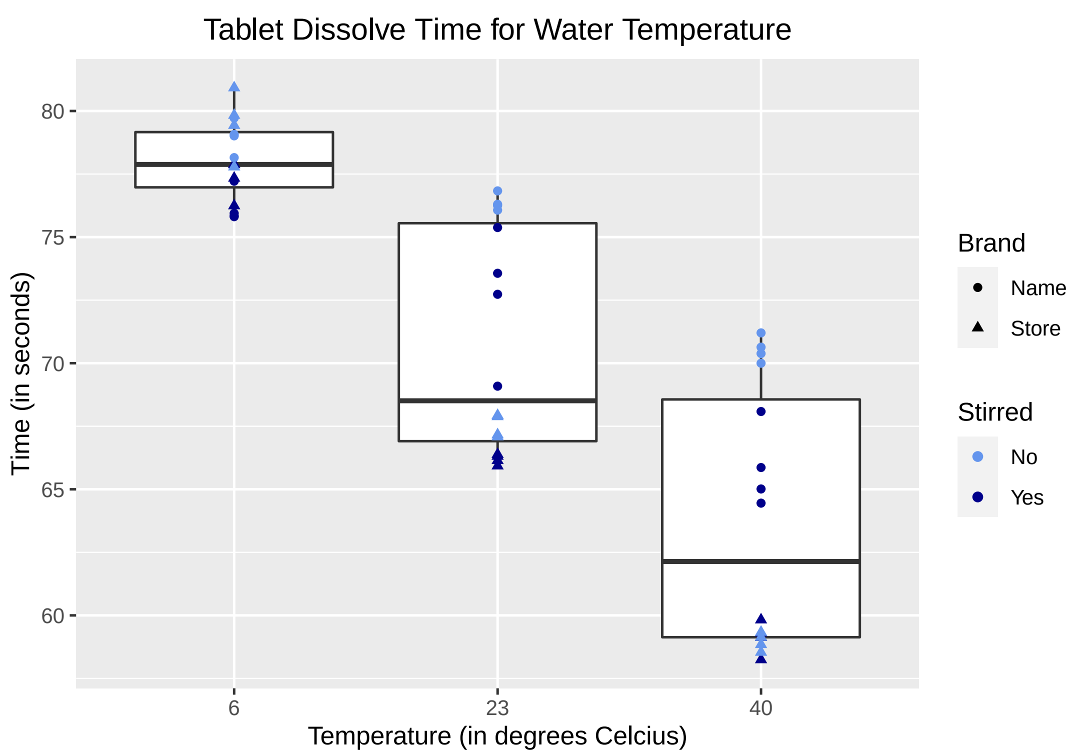
The box plot below displays a five-number summary of dissolving time for each brand of tablet. The plot displays `Time` as a function of `Brand` and indicates that there is an effect of the brand on time.



The box plot below displays a five-number summary of dissolving time for each stirred status. The box plot displays `Time` as a function of `Stirred` and indicates that there is an effect of the stirred status on time.

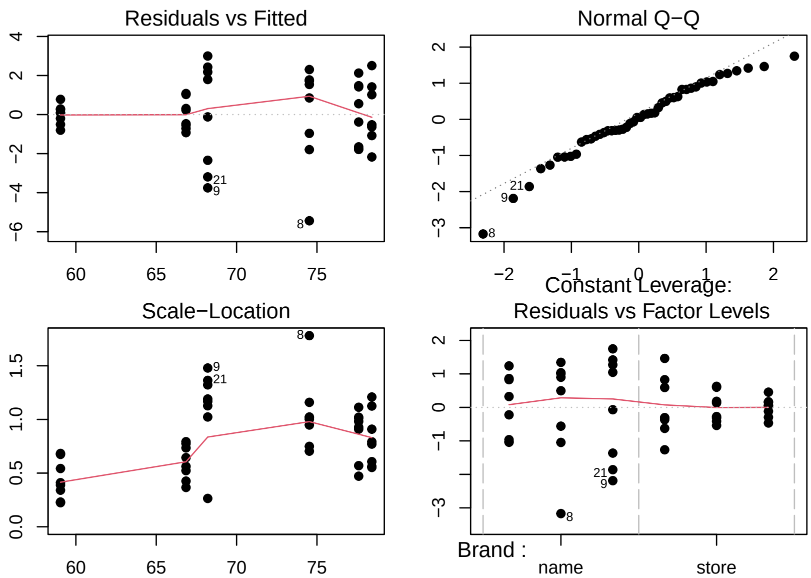


The box plot below displays the dissolving time for different water temperatures. `Time` is displayed as a function of `Temperature`, however, we can also see how the `Brand` and `Stirred` variables affect the dissolving time by observing the color and shape of the points. It is clear that a warmer temperature reduces the dissolving time. It also appears that, as mentioned above, stirring the water reduces the dissolving time and that at the higher temperatures, the store brand dissolves more quickly than the name brand tablets.



Analysis and Results

Two Factor Crossed Mixed Effects Model




```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## Brand         1  342.0   342.0  101.75 8.65e-13 ***
## Temp          2 1654.7   827.4  246.16  < 2e-16 ***
## Brand:Temp    2  231.9   115.9   34.49 1.37e-09 ***
## Residuals    42  141.2     3.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

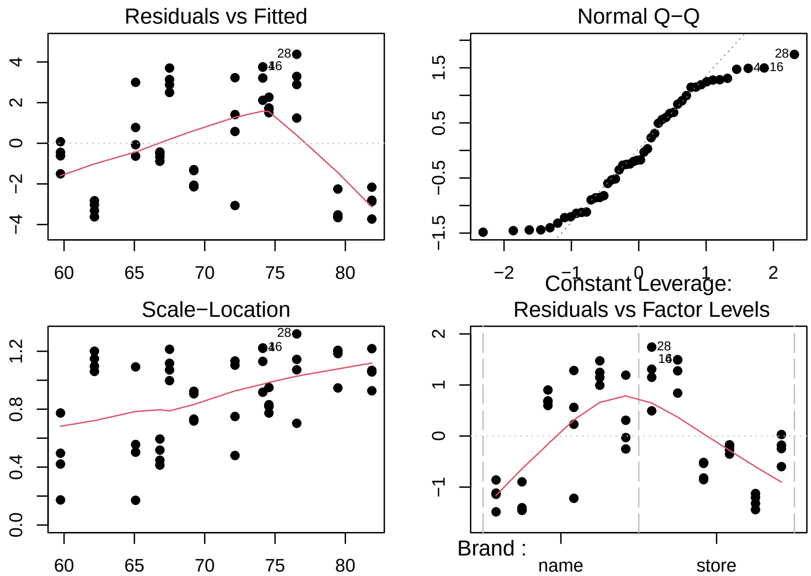
Our model is $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + E_{ijk}$

for $i = 1, 2$; $j = 1, 2, 3$; $k = 1, 2, 3, 4$

where $\alpha_i \sim iid\ N(0, \sigma^2_\alpha)$; $\beta_j \sim iid\ N(0, \sigma^2_\beta)$; $(\alpha\beta)_{ij} \sim iid\ N(0, \sigma^2_{\alpha\beta})$; $E_{ijk} \sim iid\ N(0, \sigma^2)$
We also require that each of the above are independent of each other.

Y_{ijk} is the dissolving time
μ is the overall mean
α_i is the fixed effect due to the ith brand
β_j is the random effect due to the jth temperature
$\alpha\beta_{ij}$ is the interaction effect between the ith brand and jth temp
E_{ijk} is the error term

Three Factor Analysis of Variance Model

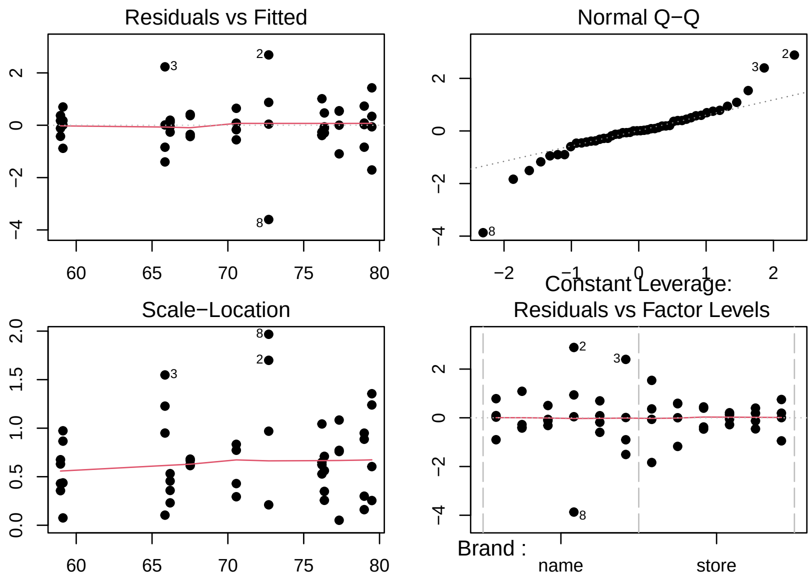



```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## Brand         1  342.0   342.0  48.514 1.44e-08 ***
## Temp          2 1654.7   827.4 117.364  < 2e-16 ***
## Stirred       1   69.9    69.9   9.914  0.00298 **
## Residuals    43  303.1     7.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Our model is $Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijkl}$

for $i = 1, 2$; $j = 1, 2, 3$; $k = 1, 2$; $l = 1, ...48$

Three Factor Crossed Mixed Effects Model



```
##                     Df Sum Sq Mean Sq F value   Pr(>F)    
## Brand                1  342.0   342.0 296.041  < 2e-16 ***
## Temp                 2 1654.7   827.4 716.169  < 2e-16 ***
## Stirred              1   69.9    69.9  60.495 3.22e-09 ***
## Brand:Temp           2  231.9   115.9 100.345 1.90e-15 ***
## Brand:Stirred        1   20.5    20.5  17.753 0.000161 ***
## Temp:Stirred         2    0.1     0.1   0.054 0.947535    
## Brand:Temp:Stirred   2    9.1     4.5   3.919 0.028838 *  
## Residuals           36   41.6     1.2                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

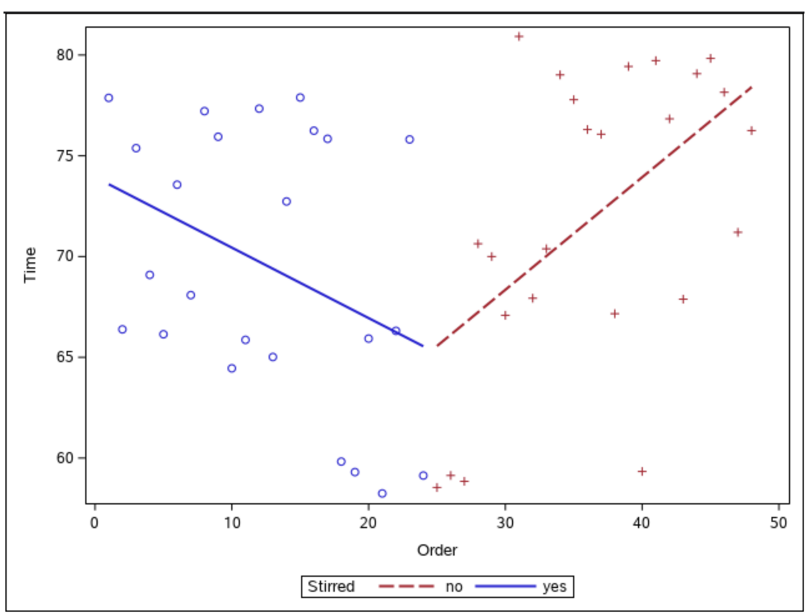
Our model is $Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + E_{ijkl}$

for $i = 1, 2$; $j = 1, 2, 3$; $k = 1, 2, 3, 4$; $l = 1, ..., 48$

Y_{ijkl} is the dissolving time
μ is the overall mean
α_i is the fixed effect due to the ith brand
β_j is the random effect due to the jth temperature
γ_k is the fixed effect due to the kth stirred status
$\alpha\beta_{ij}$ is the interaction effect between the ith brand and jth temp
$\alpha\gamma_{ik}$ is the interaction effect between the ith brand and kth stirred status
$\beta\gamma_{jk}$ is the interaction effect between the jth temp and kth stirred status
E_{ijkl} is the error term

Using Order as a Covariate with Time

When we graph the Order variable with respect to Time and group it by Stirred, there appears to be a linear association between the two blocks, albeit weak, but when we look at the two blocks separately, we can see the potential of using Order as a covariate to assist us in determining if there is a difference in dissolving times between the two blocks.



For Block 1, where the liquid was stirred, there appears to be a downward linear trend in the dissolving time of the cold medicine tablets. For Block 2, where the liquid was not stirred, there appears to be an upward linear trend in the dissolving time of the cold medicine tablets. Since the `Order` variable directly matches the `Stirred` variable, meaning the first 24 values of `Order` coincide with Block 1 and the second 24 values of `Order` coincide with Block 2 and the slopes for the two Blocks are not moving in the same direction as we can observe looking at our initial graph, we can write our model as follows:

$$Time = Mean + Brand + Temp + Brand * Temp + Error$$

When we look at our model for Block 1 and Block 2, we see that our regression slopes are not all equal to 0 given the p-values (<0.001) for both models are less than our significance level of 0.05 and that there is a significant interaction effect. Since we have different slopes we can look at the differences between `Brand` and `Temp` at the different levels. For both Blocks, we can see that there was not a significant difference between the Brands at 6 degrees Celsius.

Conclusion

Effectively describe conclusions and reasons for recommendation, analysis limitations, and future work. Address the proper role of the Stirred variable in this analysis.

The model which describes this experiment best is the Three Factor Crossed Mixed Effects Model. When comparing MSE between the models, this model has the lowest value at 1.2. MSE is the average of the squared errors, so a lower value indicates a lower error and a better model fit. This model includes three explanatory variables, `Brand`, `Temp`, and `Stirred`. While `Brand` and `Stirred` are fixed effects because we are only interested in comparing "Name Brand" vs "Store Brand" and "Stirred" vs "Not Stirred", `Temp` is a random effect since the three temperatures have presumably been selected as a sample from the total population of temperatures at which one might dissolve a tablet. In reviewing the analysis of variance table for this model, we can see that all model terms are significant at the 5% level aside from one, the interaction effect between `Temperature` and `Stirred`. This is confirmed in the interaction plot below, where we observe the two lines to be parallel. We can see that the temperature reduces the dissolve time and whether or not it is stirred also affects dissolve time, but that temperature and stirred status do not depend on the value of one another.