ST_518 Project

MA, HK, BA, JF

2022-12-03

Executive Summary

We were tasked with analyzing experimental data for dissolving times of brand name versus store brand cold medicine at different water temperatures and stirring factors. We will explore the results to determine if differences do exist between brand name and store brand medicine; if so, are they the result of temperature differences, stirring factors, or both?

The data was collected by dropping the medicine tablets into 60 mL of water of varying temperature (6°, 23°, 40°) from a fixed height and recording the time it took the tablet to completely dissolve from the time it was dropped. The stirring factor was used as a blocking method; Block I was stirred while Block II was not. Four different individuals performed the experiment and recorded their results; the average of these results were recorded as the observations.

A summary statistics table shows there is, indeed, a difference in mean dissolve time between name brand and store brand. A means table shows name brand has a mean of 73.443 with variance of 21.131, while store brand has a mean of 68.104 and variance of 67.032. Our findings also showed that stirring appears to increase variability in name brand while decreasing variability in store brand. Additionally, plots of the data show a meaningful interaction between brand and temperature, which becomes more prominent at higher temperatures.

8 models were fit in total: 3 for fixed effects, 3 for mixed effects, and 2 for adding the `order` variable as a covariate. Model 1, which we have referred to as "the full model", includes a three-way interaction between brand, temp, and stirred and produced the lowest Root Mean Square Error (RMSE) of 1.075. Model 8, which used the full model but added `order` as covariate measure, produced a RMSE of 1.083 (the 2nd lowest RMSE). We chose the best model by using the common fit criteria RMSE, Adjusted R^2, AIC, and BIC. For RMSE, AIC, and BIC, lower values are better, while a higher value of Adjusted R^2 is preferred as it explains the percentage of variation explained by the model. Model 1 produced the best results across all four model selection criteria with RMSE of 1.075, Adjusted R^2 of 0.977, AIC of 155.337, and BIC of 179.663. Model 1's assumptions of normality and constant variance are generally met, despite having slightly heavy tails and mild heteroscedasticity. The Cook's Distance plot shows observations 2, 3, and 8 potentially being problematic outliers, but because the values were all under 0.5, we chose to leave them in the analysis to avoid creating imbalances.

An Analysis of Variance (ANOVA) was conducted for the full model to check differences and run several hypothesis tests. A significant result is declared when the F critical value (F-crit) is less than the calculated F-score. The results showed that there is at least one difference in means (F-crit = 2.06, F-score = 183.21) and there are significant interactions between brand and temperature (F-crit = 3.2594, F-score = 100.3688), brand and stirring (F-crit = 4.1132, F-score = 17.758), and brand, temperature, and stirring (F-crit = 3.2594, F-score = 3.919). The same analysis also showed that there is not sufficient evidence to claim meaningful interactions between temperature and stirring (F-crit = 3.2594, F-score = 0.0537).

A number of contrast analyses were conducted to see how much the significant means differed from one another. Our findings showed that, on average, stirring medicine reduces dissolving time by 1.78 to 3.04 seconds, name brand medicine dissolves between 4.71 and 5.97 seconds slower than store brand, and stirring reduces name brand dissolving time by 2.83 to 4.61 seconds compared to 0.22 to 2 seconds for store brand. Pairwise analyses were conducted for the three levels of temperature using Bonferroni correction which showed significant differences for each of the pairs. The 95% confidence limits were (6.25, 8.38), (13.32, 15.44), and (6.00, 8.13) for the pair wise comparisons of $6°C\ vs\ 23°C$, $6°C\ vs\ 40°C$, and $23°C\ vs\ 40°C$, respectively. When comparing each temperature to the remainder of the group, $23°C$ was not significantly different.

To recap, we have determined that, on average, name brand dissolving times are between 4.71 and 5.97 seconds slower than store brands, water temperature has an inverse relationship with dissolving times, there is a highly significant interaction between brand and temperature, and stirring reduces dissolving time by 1.78 and 3.04 seconds. Stirring was found to have a highly significant interaction with brand (p=0.0002) and a significant interaction with both brand and temperature (p=0.029); because these interactions are so significant, we suggest using stirring as a main effect rather than a blocking variable.

Introduction

For this paper we have been presented with data gathered on the dissolving times of cold medicine in water. The dataset contains dissolving characteristics of different cold medicine brands done under various environmental conditions. The goal of this paper is to answer the following questions:

- Are the dissolving characteristics different between brands?
- Does temperature of the water influence dissolving characteristics? If so, is there an interaction effect between brand and temperature?
- Does stirring influence dissolving times and is there an interaction with the other two effects? What is the proper role for stirring?

Experimental Design

The data collected consisted of name brand versus store brand cold medicine tablets that were dissolved at varying water temperatures (6° C, 23° C, 40° C). A complete block design was formulated using 2 blocks with 4 observations on each of the treatment combinations in each block. Block I was stirred with a magnetic stirring plate at 350 rpm while Block II was not stirred.

Each tablet was dropped into 60 mL of water from a fixed height and dissolving time was measured from the time the tablet was dropped until it was completely dissolved. This process was completed by 4 different experimenters and the average time (t/4) was recorded as the observation.

Exploratory Data Analysis

Summary

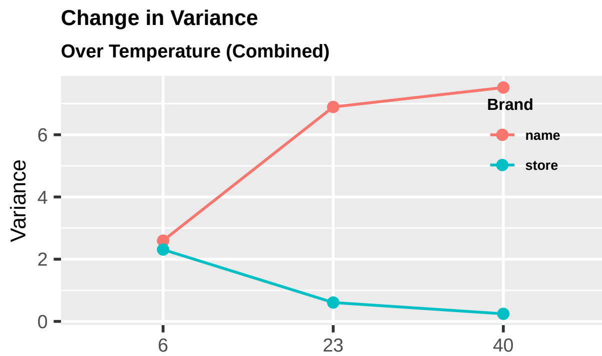
In total, the provided dataset contains 48 rows and 5 columns. The 5 columns include 3 explanatory categorical variables (Brand, Temp in °C, and Stirring), a single continuous response variable (Time, in seconds) and one descriptor (order). Prior to analysis, the data will be explored to gain a better understanding of what to expect and, more importantly, check for any potential violations of analytic assumptions.

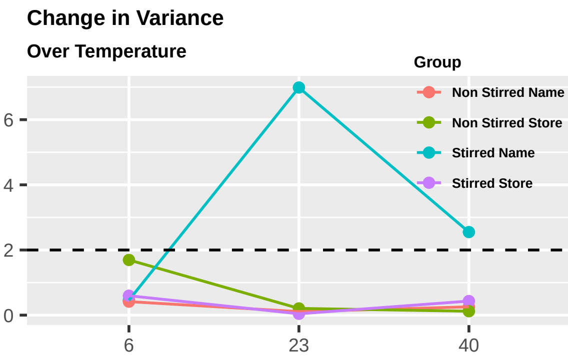
From the summary statistics table (see Appendix I, Table 4), we can see that each group has exactly 4 entries, eliminating concerns with respect to design imbalance. Constructing a means table for the data without taking into account stirring, we can see that there does appear to be a disparity between the mean dissolving times of store brand cold medications when compared to name brand.

Table 1: Means Table

Temp	Name	Store	Temp. Mean
6	77.60	78.41	78.01
23	74.53	66.85	70.69
40	68.20	59.04	63.62
Brand Mean	73.44	68.10	70.77

When inspecting the marginal means of store versus name brand, we find that store brand dissolves in less time, on average, than the name brand. This disparity becomes more pronounced as the effect of temperature is introduced. It was observed that increasing temperature has a more dramatic effect on store brand medicine than name brand. Dissolving time for store brand medicine drops from 78.42 to 59.04 (Δ of -19.38 seconds) across a temperature change from 6°C to 40°C. Whereas, name brand medicine only drops from 77.60 to 68.20 (Δ of -9.40 seconds) across the same temperature change.



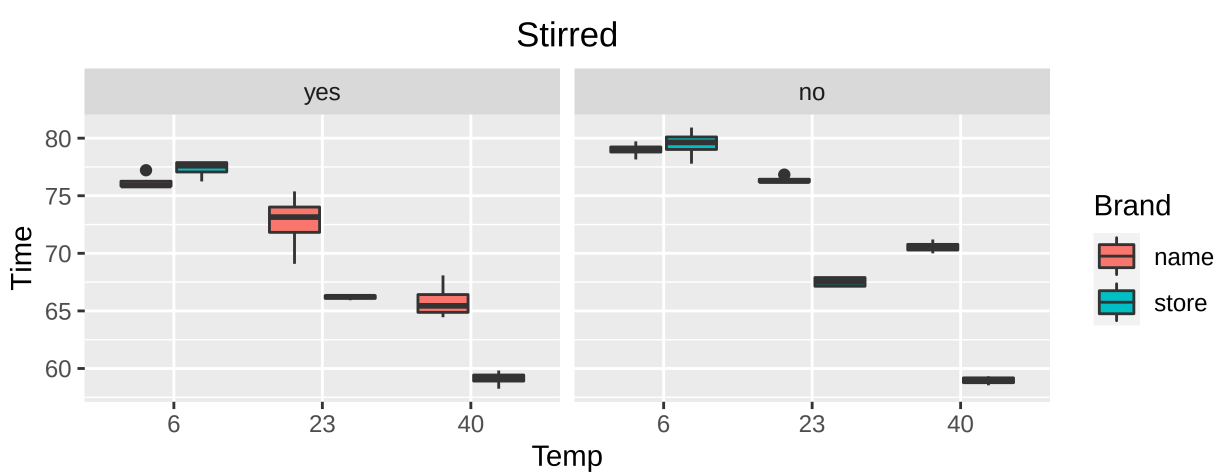


While combing over the summary table, the variability between factors caught our attention. We noticed that some of the variances behaved differently between stirring condition while holding brand type and temperatures constant. Two plots were created to illustrate these differences. In the left plot above, we can see a clear departure in differences of variances between brand types. On the right, stirred name brand medicines at higher temperatures had much greater differences.

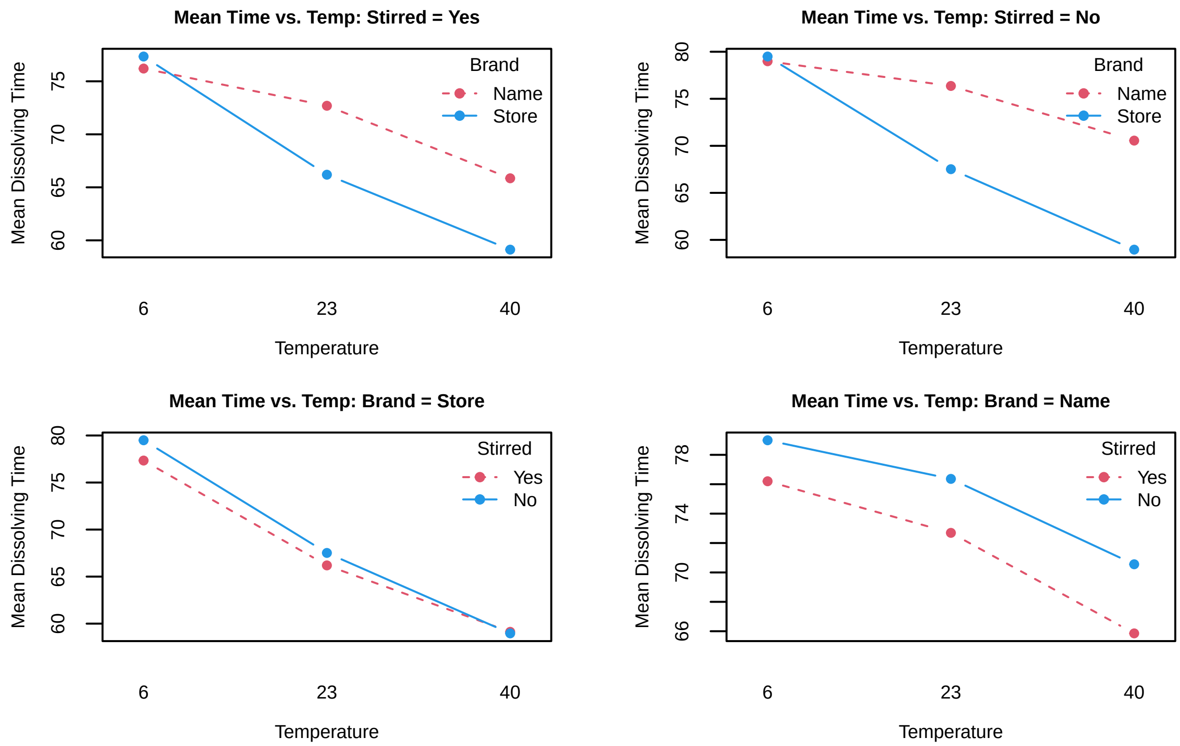
Interactions

From the boxplots below, we can immediately see that stirring seems to increase the variance of the name-brand medicine–while also decreasing the mean differences of brand observations within each temperature grouping. An interaction effect between temperature and brand can be deduced if lines are drawn through the centers of the boxes. Earlier, we had introduced an insight from the summary statistics output indicating an inverse relationship between temperature and dissolve time–as temperature increases dissolve time decreases. The boxplot reinforces this idea. We can also claim that temperature has an inverse effect on dissolving times whether stirring is present or not–indications of temperature having a strong effect on dissolving time by itself. Stirring might have an additive effect regardless of temperature.

It is simply conjecture at this point, however, we noticed that observations of dissolve time while stirring the water seems to have increased the name brand variability, while not stirring the water seems to have increased the store brand variability. Perhaps it is worth looking into the blocking effects of stirred on variability at various temperatures.

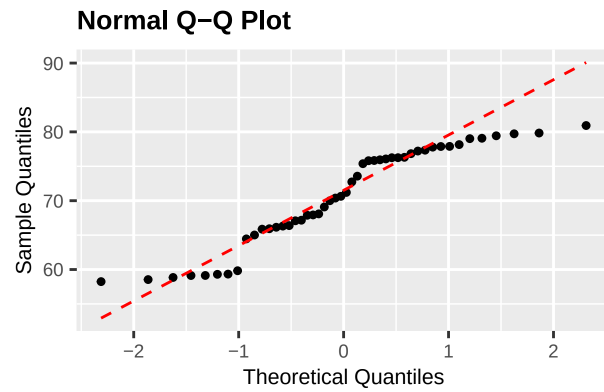


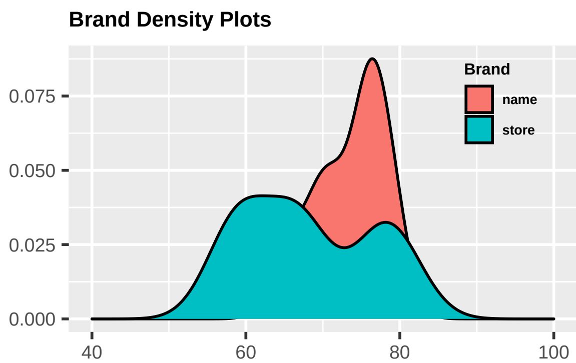
The possible interaction between brand and temperature becomes even more noticeable in the three-factor interaction plots below. Specifically, the brand and temperature interaction can be seen when the temperature increases. The slope for the store brand has a more pronounced negative slope than the slope of the name brand. In addition, there might be a slight three-factor interaction between brand, temperature, and stirring as the name and store brand lines appear to be closer together in the stirred=yes plot than the stirred=no plot.

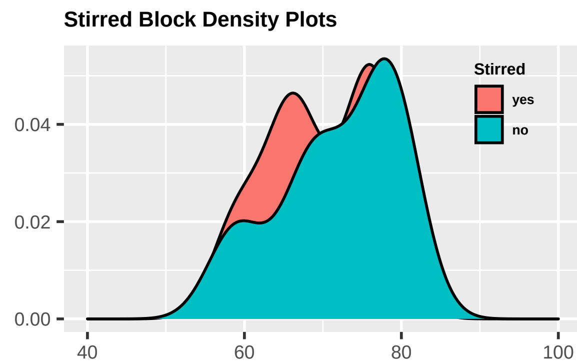


Assumptions and Violations

The Time versus Order scatter shows no obvious grouping, whether by brand or blocking factor. Density plots do reveal multiple peaks in the data indicating the presence of multiple populations. These could be based on temperature, or some unaccounted for lurking variable.







Finally, we check the normality of the data. Here a Q-Q plot is generated for the full model residuals. The data appears to suffer from heavy tails, multimodality and/or gaps in data between the left tail and the center. Since downstream analysis hinges on the assumption that our data is normally distributed, these issues may pose a problem.

Analysis and Results

Model Development

The following models were developed and analyzed for this paper:

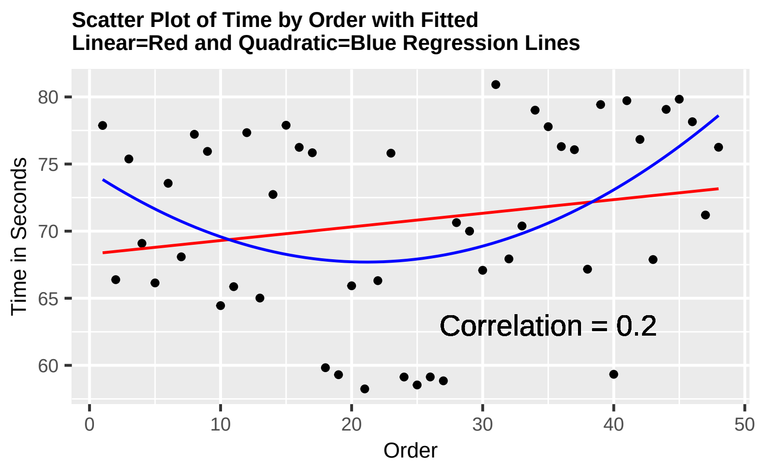
Fixed Effects Models:

$Model\ 1 : Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\gamma\beta)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
$Model\ 2 : Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + \epsilon_{ijkl}$
$Model\ 3 : Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijkl}$

Mixed Effects models:
$Model\ 4 : Y_{ijk} = \mu + \alpha_i + B_j + (\alpha B)_{ij} + \epsilon_{ijk}$
Where brand is fixed and temperature is random.
$Model\ 5 : Y_{ijk} = \mu + A_i + \beta_j + (A\beta)_{ij} + \epsilon_{ijk}$
Where brand is random and temperature is fixed.
$Model\ 6 : Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \epsilon_{ijk}$
Where both brand and temperature are random.

Where α is brand effect, β is temperature effect, γ is stir effect. i, j, k are (1, 2), (1, 2, 3), and (1, 2), respectively. ϵ_{ijkl} is assumed to be normally distributed with a μ_ϵ of 0 and a variance of σ^2_ϵ. μ is the overall mean and is an unknown value.

With Order as a Model covariate:
We developed Models 7 and 8 to account for run order variable Order as a covariate. We first considered the linear relationship between only Order and the response Time.



The preceding scatter plot shows only a weak positive linear relationship between Order and Time with a small correlation=0.2. However, the plot does show a plausible quadratic relationship between Order and Time.

A fitted quadratic regression model of Time versus additive quadratic Order was significant with overall overall F test p-value=0.013 and all coefficients having p-values less than .05 (See Appendix I on pp. 18-19). Thus, we decided to include both Order and $Order^2$ as additive covariates in Models 7 and 8.

$Model\ 7: Y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + \nu_l + \nu_l^2 + (\alpha\beta)_{ij} + \epsilon_{ijklm}$
Similar to model 2, but with an introduced covariate effect, ν, to represent order.

$Model\ 8: Y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + \nu_l + \nu_l^2 + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\gamma\beta)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijklm}$
Similar to model 1, but with an introduced effect, ν, to represent order.

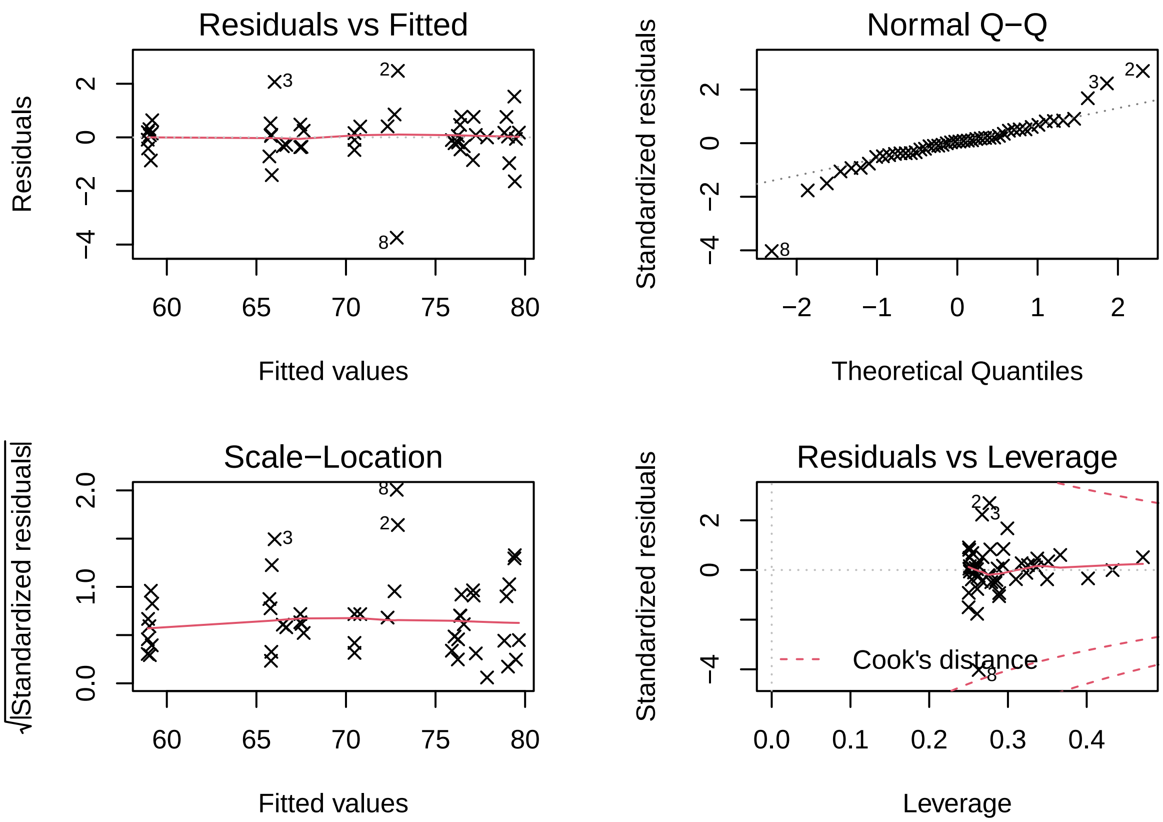
Model Selection

	Root MSE	R^2	adj R^2	AIC	BIC
Model1	1.075	0.982	0.977	155.337	179.663
Model2	1.319	0.970	0.966	171.198	186.168
Model3	2.655	0.872	0.860	236.681	247.908
Model7	1.312	0.972	0.966	172.342	191.054
Model8	1.083	0.983	0.977	157.322	185.390

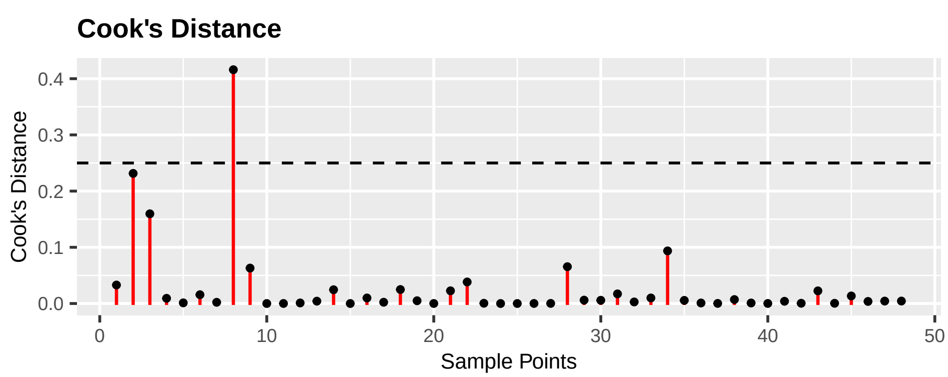
The mixed/random effects Models 4, 5, 6 all had Root MSE=1.833 (see Appendix I, p. 14).

Several criteria were used to select the final model including Root MSE(error standard deviation estimate), Adjusted R^2, AIC, and BIC. Model 1 had the best outcome in all these criteria with Root MSE=1.075, Adjusted R^2=0.9770, AIC= 155.34, and BIC= 179.66.

Next, model assumptions are checked for Model 1 using diagnostic plots.



Multiple diagnostic plots indicate observations 2, 3, and 8 could be outliers. Evidence for outliers was seen in the EDA boxplots. Plotting the Cook's Distance for each point based on our model. We have chosen not to remove any outliers from the data because the calculated Cook's distance is below 0.5 and because we do not want to introduce imbalances in the analysis.



The residual-to-fitted plot also shows that there may be an issue with mild heteroscedasticity. The implication is that the criteria that residuals are drawn from a population of constant variance may not be met. The QQ plot indicates the normality assumption is likely met although there are somewhat heavy tails due to outlier observations.

Overall the plots indicate the model assumptions have been met with the caveat that there are a few outliers in the data.

ANOVA Analysis

Table 3: Model 1: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	1	342.007	342.007	296.041	0.000
Temp	2	1654.737	827.368	716.169	0.000
Stirred	1	69.888	69.888	60.495	0.000
Brand:Temp	2	231.852	115.926	100.345	0.000
Brand:Stirred	1	20.510	20.510	17.753	0.000
Temp:Stirred	2	0.125	0.062	0.054	0.948
Brand:Temp:Stirred	2	9.056	4.528	3.919	0.029
Residuals	36	41.590	1.155	NA	NA

Hypothesis Testing

The following hypothesis testing is based on the results of the Model 1 ANOVA table. All rejection criteria is based on rejecting H_0 when $F_{crit} < F_{score}$:

Full model ANOVA (see Appendix I, p. 14 for details):
$H_0 : \mu_1 = \mu_2 = ... = \mu_i = 0$
H_a : Not all means are equal to zero.
$F_{score} = \frac{MS[Treatment]}{MS(E)} = \frac{211.65}{1.155} = 183.21$
Degrees of freedom = 11 and 36
$F_{crit} = 2.06$
Decision: Reject H_0. There is sufficient evidence to support the claim that there is at least one mean that differs with a significance level of $\alpha = 0.05$

Interaction Effects

$H_0 : (\alpha\beta)_{ij} = 0$
$H_a : (\alpha\beta)_{ij} \neq 0$
$F_{score} = 100.3688$
Degrees of freedom = 2 and 36
$F_{crit} = 3.2594$
Decision: $Reject H_0$. There is sufficient evidence to support the claim that there is an interaction effect between brand type and temperature with a significance level of $\alpha = 0.05$.

$H_0 : (\alpha\gamma)_{ik} = 0$
$H_a : (\alpha\gamma)_{ik} \neq 0$
$F_{score} = 17.758$
Degrees of freedom = 1 and 36
$F_{crit} = 4.1132$
Decision: $Reject H_0$. There is sufficient evidence to support the claim that there is an interaction effect between brand type and stirring with a significance level of $\alpha = 0.05$.

$H_0 : (\beta\gamma)_{ij} = 0$
$H_a : (\beta\gamma)_{ij} \neq 0$
$F_{score} = 0.0537$
Degrees of freedom = 2 and 36
$F_{crit} = 3.259446$
Decision: Fail to Reject H_0. There is insufficient evidence to support the claim that there is an interaction

effect between temperature and stirring with a significance level of $\alpha = 0.05$.

$H_0 : (\alpha\beta\gamma)_{ijk} = 0$
$H_a : (\alpha\beta\gamma)_{ijk} \neq 0$
$F_{score} = 3.919$
Degrees of freedom = 2 and 36
$F_{crit} = 3.259446$
Decision: Reject H_0. There is sufficient evidence to support the claim that there is an interaction effect between brand, temperature, and stirring with a significance level of $\alpha = 0.05$.

We saw that three factor interaction exists from earlier interaction plots. Hypothesis tests were used to determine which interactions were plausible. As a result, we determined that there is likely an interaction between brand and temperature and brand and stirring. We also found that there was a difference in the means within treatments.

With regards to Table 15, type III sums of squares show that temperature explains the most amount of variance. Temperature adds 221.58 sum of squares given all the other variables in the model. The next highest contributor to variance is the interaction effect between brand and temperature, which adds 80.11 given all the other variables in the model. Brand adds little variance add a value 2.6 to the sum of squares given the other variables. The next lowest are both interaction effects involving stirring. The interaction effect for stirring & temperature and stirring & brand only add 3.67 and 0.4 to variance respectively. They are also not significant since their respective p-values are above 0.05.

Contrasts

Conducting a linear contrast analysis on each of the explanatory variables reveals that there are significant differences between groups based on factors (see Appendix I, p. 12)

In the first case, we contrasted the means of stirred versus not stirred. Here the difference in means is -2.41 with an upper 95% confidence limit of -3.04 and a lower 95% CI limit of -1.78. In other words, on average stirring medicine reduces dissolving time by between 1.78 and 3.04 seconds regardless of brand or temperature. When looking only at brand, name brand dissolving times were on average between 4.71 and 5.97 (95% CI) seconds slower than store brand. Since neither of the intervals contained zero we can conclude that there is a difference between brands and between the presence of stirring.

While significant for both store and name brands, stirring had more of an impact on dissolving times for name brand than it did for the store brand. Stirring reduced name brand dissolving times by 2.83 and 4.61 seconds whereas for the store brand that interval was 0.22 and 2 seconds.

A similar analysis was completed for the three levels of temperature. Completing a contrast analysis using a Bonferroni correction we found that in pairwise cases each level was significantly different from the other. The 95% confidence limits were (6.25, 8.38), (13.32, 15.44), and (6.00, 8.13) for the pair wise comparisons of $6°C\ vs\ 23°C$, $6°C\ vs\ 40°C$, and $23°C\ vs\ 40°C$, respectively. Zero did not fall in any of those ranges. When comparing individual levels versus the remainder of the group, $23°C$ was found not to be significantly different from the rest of the levels. That confidence interval ranged from -1.04 to 0.80 seconds of dissolving time. Due to that, we do not have enough evidence to say $23°C$ is different from either $6°C$ or $40°C$.

Conclusion

Based on lowest RMSE criteria and adequate diagnostic assumption checks we chose the following three factor interaction model as our final model. In this model, the three factors are assumed to be fixed effects.

$Model\ 1 : Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\gamma\beta)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$

Where α is brand effect, β is temperature effect, γ is stir effect. i, j, k are (1, 2), (1,2,3), and (1,2), respectively. ϵ_{ijkl} is assumed to be normally distributed with a μ_ϵ of 0 and a variance of σ^2_ϵ. μ is the overall mean and is an unknown value.

Next we address the primary research questions.

- Are the dissolving characteristics different between brands?
  When looking only at brand, name brand dissolving times were on average between 4.71 and 5.97 (95% CI) seconds slower than store brand.

- Does temperature of the water influence dissolving characteristics? If so, is there an interaction effect between brand and temperature?
  When looking only at temperature, the general trend was a decrease in dissolving times as temperature increased. Comparing specific temperatures, we found a significant difference between dissolving mean times for the lowest and highest temperatures $6°C$ and $40°C$.
  We found a highly significant two factor interaction (p-value$<.0001$) between brand and temperature. Stirring reduced name brand dissolving times by 2.83 and 4.61 seconds whereas for the store brand that interval was 0.22 and 2 seconds.

- Does stirring influence dissolving times and is there an interaction with the other two effects? What is the proper role for stirring?
  Looking only at stirring, dissolving time was reduced for stirred medicine by between 3.04 and 1.78 seconds (95% CI).
  We found a significant three factor interaction (p-value= 0.029) between stirring, brand, and temperature confirming earlier insights from three factor interaction plots. A highly significant two factor interaction (p-value=0.0002) exists between stirring and brand.
  Rather than treating stirring as a blocking variable, we suggest treating stirring as a third interacting factor as we did in our final model.

Finally, we pose some questions and possible routes for future research.

Would different name and store brands have similar dissolving times?
Our final model assumed all fixed effects meaning we assumed particular name and store brands were used in this study. A future study could incorporate additional specific name and store brands to examine whether additional types of name and store brands would have similar dissolving times.

What question are the researchers wanting to answer with the chosen temperature range?
This experiment included a wide range of temperatures(6°C, 23°C, 40°C). 6°C is below and 40°C is above typical room temperature. Are the researchers trying to include possible outdoor usage of the medicines? Answers to these questions could lead to different ranges of chosen temperatures for a future study.

Appendix I: Analysis Tables and Figures

Table 4: Data Summary Table

Brand	Temp	Stirred	Min	25%	Mean	Median	75%	Max	Range	Var	n
name	6	yes	75.80973	75.83358	76.20241	75.89223	76.26107	77.21547	1.4057377	0.4593492	4
name	6	no	78.15246	78.79910	78.99061	79.04435	79.23586	79.72130	1.5688327	0.4146440	4
name	23	yes	69.08937	71.82180	72.69145	73.14894	74.01859	75.37855	6.2891789	6.9869087	4
name	23	no	76.06895	76.20492	76.36351	76.27622	76.43481	76.83265	0.7636940	0.1078134	4
name	40	yes	64.45156	64.87321	65.85343	65.43863	66.41886	68.08492	3.6333543	2.5499751	4
name	40	no	69.99943	70.28754	70.55511	70.50947	70.77705	71.20207	1.2026434	0.2544033	4
store	6	yes	76.24402	77.06561	77.33703	77.60659	77.87801	77.89089	1.6468708	0.5964884	4
store	6	no	77.78345	79.01994	79.49240	79.63219	80.10465	80.92176	3.1383169	1.6942517	4
store	23	yes	65.92809	66.08831	66.19126	66.22629	66.32923	66.38436	0.4562787	0.0411024	4
store	23	no	67.08353	67.14393	67.51552	67.52360	67.89520	67.93138	0.8478521	0.2060739	4
store	40	yes	58.24407	58.90895	59.12529	59.21659	59.43293	59.82388	1.5798100	0.4320148	4
store	40	no	58.53920	58.76884	58.96347	58.99050	59.18513	59.33370	0.7945066	0.1202191	4

Brand	Mean	Var	Max	Min	Spread
name	73.44276	21.13144	79.72130	64.45156	15.26973
store	68.10416	67.03190	80.92176	58.24407	22.67769

Temp	Mean	Var	Max	Min	Spread
6	78.00561	2.467426	80.92176	75.80973	5.112037
23	70.69044	19.204515	76.83265	65.92809	10.904559
40	63.62433	25.996561	71.20207	58.24407	12.958001

Brand	Temp	Mean	Var	Max	Min	Spread
name	6	77.59651	2.5957290	79.72130	75.80973	3.911567
name	23	74.52748	6.8931729	76.83265	69.08937	7.743272
name	40	68.20427	7.5178162	71.20207	64.45156	6.750513
store	6	78.41471	2.3090692	80.92176	76.24402	4.677741
store	23	66.85339	0.6069845	67.93138	65.92809	2.003292
store	40	59.04438	0.2441527	59.82388	58.24407	1.579810

Brand	Stirred	Mean	Var	Max	Min	Spread
name	yes	71.58243	22.87009	77.21547	64.45156	12.763904
name	no	75.30308	13.76300	79.72130	69.99943	9.721863
store	yes	67.55119	61.60366	77.89089	58.24407	19.646821
store	no	68.65713	77.88680	80.92176	58.53920	22.382569

Table 9: Least Squares Means

Temp	Stirred	Mean	Var	Max	Min	Spread
6	yes	76.76972	0.8203152	77.89089	75.80973	2.081166
6	no	79.24151	0.9757520	80.92176	77.78345	3.138317
23	yes	69.44135	15.0841513	75.37855	65.92809	9.450466
23	no	71.93952	22.5021998	76.83265	67.08353	9.749120
40	yes	62.48936	14.2117017	68.08492	58.24407	9.840843
40	no	64.75929	38.5508753	71.20207	58.53920	12.662879

Table 10: Contrast Stirred and Brand

Brand	Temp	Stirred	emmean	SE	df	lower.CL	upper.CL
name	6	yes	76.20241	0.5374175	36	75.11248	77.29235
store	6	yes	77.33703	0.5374175	36	76.24709	78.42696
name	23	yes	72.69145	0.5374175	36	71.60152	73.78138
store	23	yes	66.19126	0.5374175	36	65.10132	67.28119
name	40	yes	65.85343	0.5374175	36	64.76350	66.94337
store	40	yes	59.12529	0.5374175	36	58.03535	60.21522
name	6	no	78.99061	0.5374175	36	77.90068	80.08055
store	6	no	79.49240	0.5374175	36	78.40247	80.58233
name	23	no	76.36351	0.5374175	36	75.27358	77.45344
store	23	no	67.51552	0.5374175	36	66.42559	68.60546
name	40	no	70.55511	0.5374175	36	69.46518	71.64505
store	40	no	58.96347	0.5374175	36	57.87354	60.05341

Table 11: Contrast Stirred versus Brand

contrast	estimate	SE	df	lower.CL	upper.CL
stirred	-2.413294	0.3102781	36	-3.042567	-1.784021
branding	5.338595	0.3102781	36	4.709322	5.967868

Table 12: Contrast Temperatures

contrast	estimate	SE	df	lower.CL	upper.CL
stirredbrand	-3.720646	0.4387996	36	-4.610573	-2.8307197
stirredstore	-1.105942	0.4387996	36	-1.995869	-0.2160151

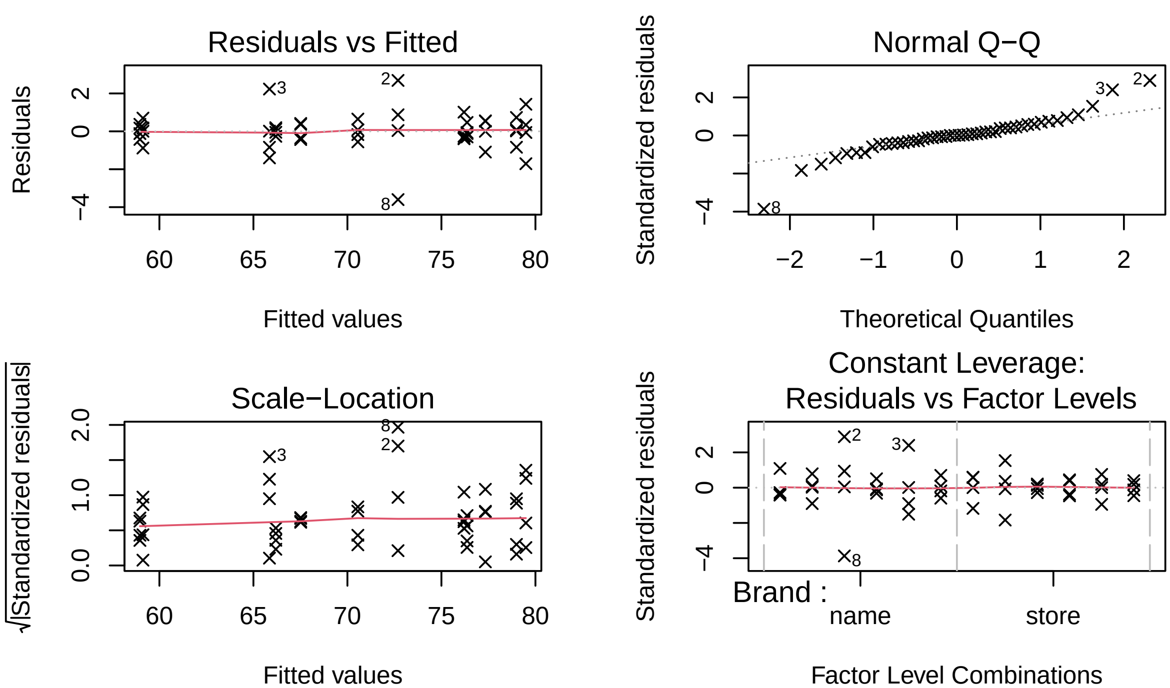
contrast	estimate	SE	df	lower.CL	upper.CL
temp6_23	7.3151767	0.3800116	36	6.254195	8.3761584
temp6_40	14.3812861	0.3800116	36	13.320305	15.4422678
temp23_40	7.0661094	0.3800116	36	6.005128	8.1270910
temp6_rest	10.8482314	0.3290997	36	9.929394	11.7670685
temp23_rest	-0.1245337	0.3290997	36	-1.043371	0.7943034
temp40_rest	-10.7236978	0.3290997	36	-11.642535	-9.8048607

Table 14: Model 1: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	1	342.007	342.007	296.041	0.000
Temp	2	1654.737	827.368	716.169	0.000
Stirred	1	69.888	69.888	60.495	0.000
Brand:Temp	2	231.852	115.926	100.345	0.000
Brand:Stirred	1	20.510	20.510	17.753	0.000
Temp:Stirred	2	0.125	0.062	0.054	0.948
Brand:Temp:Stirred	2	9.056	4.528	3.919	0.029
Residuals	36	41.590	1.155	NA	NA

Table 15: Model 1: Type III ANOVA Table

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	23227.231	1	20105.451	0.000
Brand	2.575	1	2.229	0.144
Temp	221.582	2	95.901	0.000
Stirred	15.548	1	13.458	0.001
Brand:Temp	80.110	2	34.672	0.000
Brand:Stirred	0.400	1	0.347	0.560
Temp:Stirred	3.668	2	1.588	0.218
Brand:Temp:Stirred	9.056	2	3.919	0.029
Residuals	41.590	36	NA	NA



SAS Overall ANOVA Model Table

Source	DF	Sum of Squares	Mean Square	F Value	Pr>F
Model	11	2328.174357	211.652214	183.21	<.0001
Error	36	41.589732	1.155270		
CorrectedTotal	47	2369.764090			

Produced from the following code:

```
proc import datafile = "/home/u39732161/data/effervescence.csv"
     out = proj_data
     dbms = csv replace;
run;

proc anova data = proj_data;
    class Brand Temp Stirred;
    model Time = Brand|Temp|Stirred;
run;
```

Table 17: Model 2: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	1	342.00715	342.007154	196.71962	0e+00
Temp	2	1654.73655	827.368276	475.89522	0e+00
Stirred	1	69.88787	69.887866	40.19891	1e-07
Brand:Temp	2	231.85191	115.925956	66.67963	0e+00
Residuals	41	71.28061	1.738551	NA	NA

Table 18: Model 2: Type III ANOVA Table

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	40014.251	1	23015.858	0.000
Brand	2.678	1	1.540	0.222
Temp	366.976	2	105.541	0.000
Stirred	69.888	1	40.199	0.000
Brand:Temp	231.852	2	66.680	0.000
Residuals	71.281	41	NA	NA

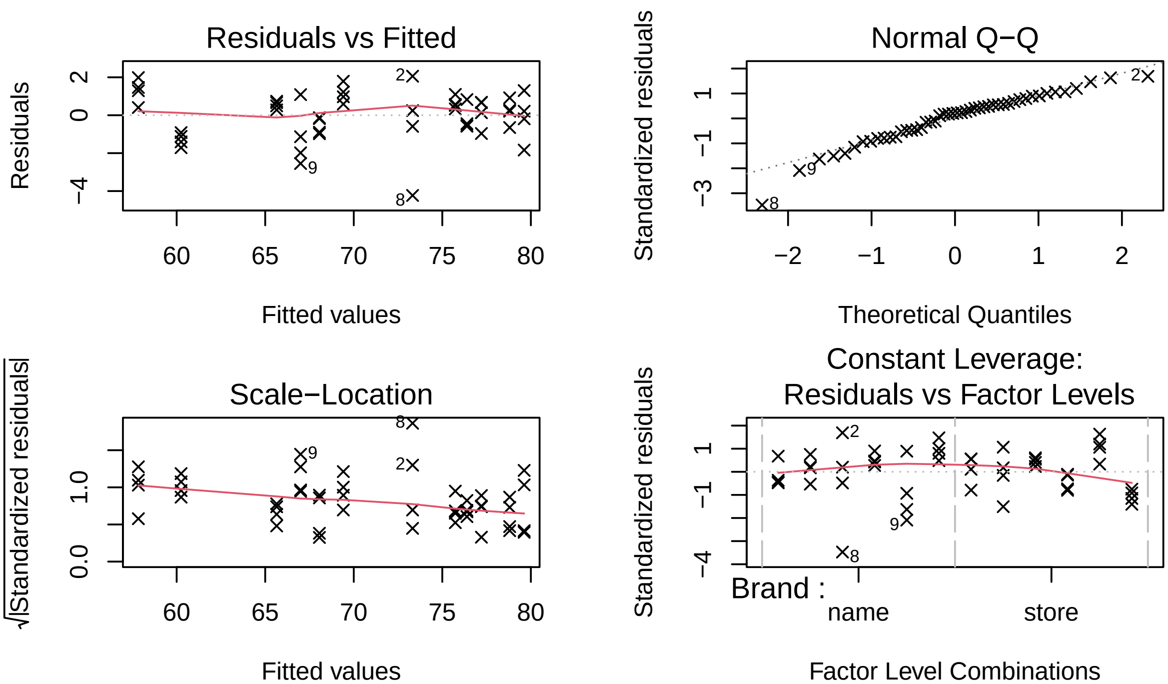


Table 19: Model 3: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	1	342.00715	342.007154	48.514451	0.0000000
Temp	2	1654.73655	827.368276	117.363970	0.0000000
Stirred	1	69.88787	69.887866	9.913744	0.0029802
Residuals	43	303.13252	7.049593	NA	NA

Table 20: Model 3: Type III ANOVA Table

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	60625.967	1	8599.924	0.000
Brand	342.007	1	48.514	0.000
Temp	1654.737	2	117.364	0.000
Stirred	69.888	1	9.914	0.003
Residuals	303.133	43	NA	NA

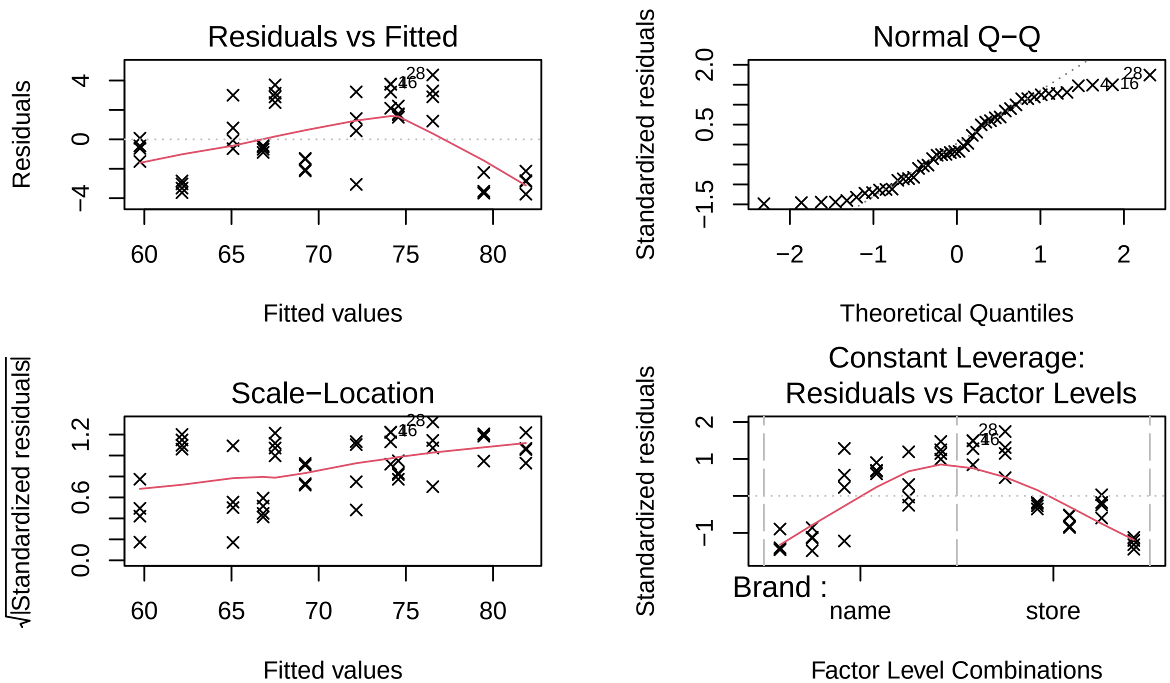


Table 21: Mixed Effects Models

Source	Df	Sum Sq	Mean Sq	Error Term	Error Df	F Score	Pr>F
Brand	1	342.0072	342.007154	MSAB	2	2.950221	0.2280
Temp	2	1654.7366	827.368276	MSAB	2	7.137041	0.1229
Brand*Temp	2	231.8519	115.925956	MSE	42	34.491507	0.0000
Residual	42	141.1685	3.361154	NA	NA	NA	NA

Table 22: Model4: Brand is Fixed and Temperature is Random

Brand	Mean	Effect
name	73.443	0.000
store	68.104	-5.339

Cov Parm	Estimate
Temp	44.465
Brand*Temp	14.071
Residual	3.361

Table 24: Model5: Brand is Random and Temperature is Fixed

Temp	Mean	Effect
6	78.006	0.000
23	70.690	-7.315
40	63.624	-14.381

Cov Parm	Estimate
Brand	9.42
Brand*Temp	14.071
Residual	3.361

Table 26: Model6: Brand and Temperature are Random

Cov Parm	Estimate
Brand	9.42
Temp	44.465
Brand*Temp	14.071
Residual	3.361

Table 27: Model 7: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	1	342.0071543	342.0071543	198.5967135	0.0000000
Temp	2	1654.7365514	827.3682757	480.4362083	0.0000000
Stirred	1	69.8878657	69.8878657	40.5824857	0.0000002
Order	1	0.9059095	0.9059095	0.5260435	0.4726048
Order2	1	12.3382937	12.3382937	7.1646004	0.0108189
Brand:Temp	2	222.7256778	111.3628389	64.6661730	0.0000000
Residuals	39	67.1626372	1.7221189	NA	NA

Table 28: Model 7: Type III ANOVA Table

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	14804.990	1	8596.961	0.000
Brand	3.665	1	2.128	0.153
Temp	326.992	2	94.939	0.000
Stirred	6.182	1	3.590	0.066
Order	0.000	1	0.000	0.990
Order2	0.746	1	0.433	0.514
Brand:Temp	222.726	2	64.666	0.000
Residuals	67.163	39	NA	NA

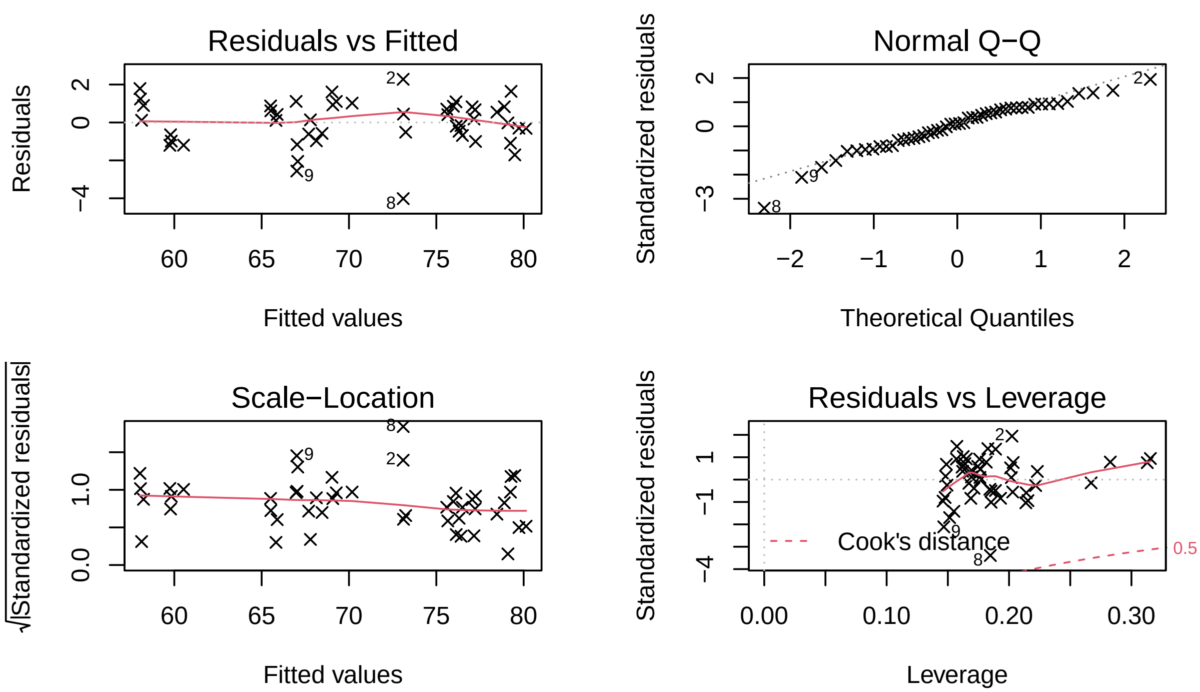
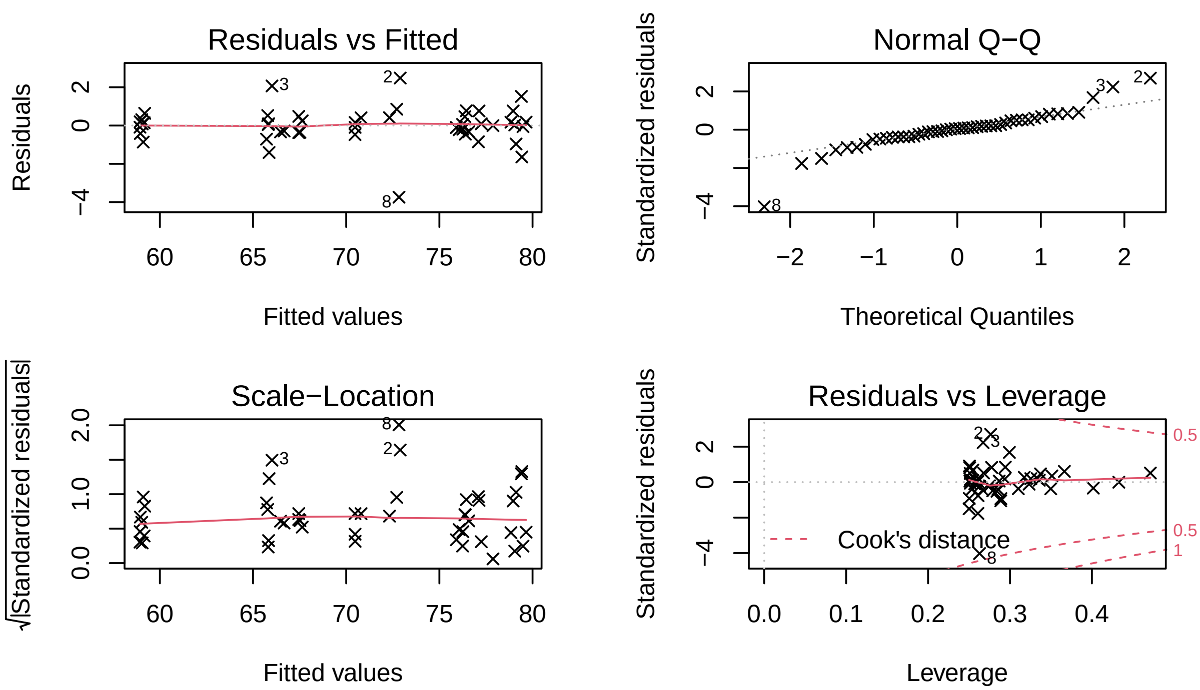


Table 29: Model 8: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	1	342.0071543	342.0071543	291.5845291	0.0000000
Temp	2	1654.7365514	827.3682757	705.3881360	0.0000000
Stirred	1	69.8878657	69.8878657	59.5841933	0.0000000
Order	1	0.9059095	0.9059095	0.7723499	0.3856600
Order2	1	12.3382937	12.3382937	10.5192407	0.0026496
Brand:Temp	2	222.7256778	111.3628389	94.9444493	0.0000000
Brand:Stirred	1	16.9126349	16.9126349	14.4191799	0.0005766
Temp:Stirred	2	0.0511659	0.0255830	0.0218112	0.9784386
Brand:Temp:Stirred	2	10.3193431	5.1596716	4.3989735	0.0199964
Residuals	34	39.8794932	1.1729263	NA	NA

Table 30: Model 8: Type III ANOVA Table

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	9812.832	1	8366.111	0.000
Brand	1.947	1	1.660	0.206
Temp	221.999	2	94.635	0.000
Stirred	8.814	1	7.515	0.010
Order	1.694	1	1.444	0.238
Order2	1.467	1	1.251	0.271
Brand:Temp	60.773	2	25.907	0.000
Brand:Stirred	0.175	1	0.149	0.702
Temp:Stirred	4.338	2	1.849	0.173
Brand:Temp:Stirred	10.319	2	4.399	0.020
Residuals	39.879	34	NA	NA



Appendix: Code

```
library(tidyverse)
library(emmeans)
library(lme4)
library(lmerTest)
library(olsrr)
library(car)
library(cowplot)
library(lmtest)
df_eff <- read_csv('effervescence.csv', col_types = 'fffnnn')
df_stats <-
df_eff %>% group_by(Brand, Temp, Stirred) %>%
summarise('Min' = min(Time),
          '25%' = quantile(Time, probs = 0.25),
          'Mean' = mean(Time),
          'Median' = median(Time),
          '75%' = quantile(Time, probs = 0.75),
          'Max' = max(Time),
          'Range' = Max - Min,
          'Var' = var(Time),
          'n' = n())

means_table <- df_eff %>% group_by(Brand, Temp) %>% summarise('Mean' = mean(Time))
means_table <- means_table %>% pivot_wider(names_from = Brand, values_from = Mean)
means_table$TempMean <- rowMeans(means_table[,2:3])
means_table <- cbind('Temp' = c('6', '23', '40', 'Brand Mean'),rbind(means_table[,2:4], colMeans(means_t
knitr::kable(means_table, digits = 2, col.names = c("Temp", "Name", "Store", "Temp. Mean"),
             caption = "Means Table")
change_var <-
df_eff %>% group_by(Brand, Temp) %>% summarise('Mean' = mean(Time),
                                               'Var' = var(Time),
                                               'Max' = max(Time),
                                               'Min' = min(Time),
                                               'Spread' = Max - Min)

var_table <- df_eff %>% group_by(Brand, Temp, Stirred) %>% summarise('Var' = var(Time))

var_table$Group <- ifelse(var_table$Brand == 'name' & var_table$Stirred == 'yes', 'Stirred Name',
                          ifelse(var_table$Brand == 'name' & var_table$Stirred == 'no', 'Non Stirred Nam
                                 ifelse(var_table$Brand == 'store' & var_table$Stirred == 'yes', 'Stirre
                                        ifelse(var_table$Brand == 'store' & var_table$Stirred == 'no',

change_varplot <-
ggplot(change_var) + geom_point(aes(x = Temp, y = Var, col = Brand)) +
                     geom_line(aes(x = Temp, y = Var, col = Brand, group = Brand)) +
                     labs(title = "Change in Variance",
                          subtitle = "Over Temperature (Combined)",
                          x = '',
                          y = 'Variance') +
                     theme(legend.position = c(0.85, 0.75),
                           plot.title = element_text(size = 8, face = "bold"),
                           plot.subtitle = element_text(size = 7, face = "bold"),
```

```
                           axis.text = element_text(size = 7),
                           axis.title = element_text(size = 8),
                           legend.text = element_text(size=5, face="bold"),
                           legend.title = element_text(size=6, face="bold"),
                           legend.key.size = unit(0.4, 'cm'),
                           legend.key = element_rect(colour = "transparent",
                                                     fill = alpha("white", 0)),
                                                     fill = alpha("white", 0)))

change_var2 <-
df_eff %>% filter(Stirred == "yes") %>% group_by(Brand, Temp) %>% summarise('Mean' = mean(Time),
                                               'Var' = var(Time),
                                               'Max' = max(Time),
                                               'Min' = min(Time),
                                               'Spread' = Max - Min)
change_var2plot <-
ggplot(change_var2) + geom_point(aes(x = Temp, y = Var, col = Brand)) +
                     geom_line(aes(x = Temp, y = Var, col = Brand, group = Brand)) +
                     labs(title = "Change in Variance",
                          subtitle = "Over Temperature (Stirred)",
                          x = '',
                          y = '') +
                     theme(legend.position = c(0.85, 0.8),
                           plot.title = element_text(size = 8, face = "bold"),
                           plot.subtitle = element_text(size = 7, face = "bold"),
                           axis.text = element_text(size = 7),
                           axis.title = element_text(size = 8),
                           legend.text = element_text(size=5, face="bold"),
                           legend.title = element_text(size=6, face="bold"),
                           legend.key.size = unit(0.4, 'cm'),
                           legend.key = element_rect(colour = "transparent",
                                                     fill = alpha("white", 0)),
                                                     fill = alpha("white", 0)))

change_var3 <-
df_eff %>% filter(Stirred == "no") %>% group_by(Brand, Temp) %>% summarise('Mean' = mean(Time),
                                               'Var' = var(Time),
                                               'Max' = max(Time),
                                               'Min' = min(Time),
                                               'Spread' = Max - Min)
change_var3plot <-
ggplot(change_var3) + geom_point(aes(x = Temp, y = Var, col = Brand)) +
                     geom_line(aes(x = Temp, y = Var, col = Brand, group = Brand)) +
                     labs(title = "Change in Variance",
                          subtitle = "Over Temperature (Not Stirred)",
                          x = '',
                          y = '') +
                     theme(legend.position = c(0.85, 0.8),
                           plot.title = element_text(size = 8, face = "bold"),
                           plot.subtitle = element_text(size = 7, face = "bold"),
                           axis.text = element_text(size = 7),
                           axis.title = element_text(size = 8),
                           legend.text = element_text(size=5, face="bold"),
```

```
                          legend.title = element_text(size=6, face="bold"),
                          legend.key.size = unit(0.4, 'cm'),
                          legend.key = element_rect(colour = "transparent", fill = alpha("white", 0)),
                          )

change_var4plot <-
ggplot(var_table) + geom_point(aes(x = Temp, y = Var, col = Group)) +
    geom_line(aes(x = Temp, y = Var, col = Group, group = Group)) +
    geom_hline(yintercept = 2, lty = 'dashed') +
         labs(title = "Change in Variance",
              subtitle = "Over Temperature",
              x = '',
              y = '') +
         theme(legend.position = c(0.85, 0.8),
               plot.title = element_text(size = 8, face = "bold"),
               plot.subtitle = element_text(size = 7, face = "bold"),
               axis.text = element_text(size = 7),
               axis.title = element_text(size = 8),
               legend.text = element_text(size=5, face="bold"),
               legend.title = element_text(size=6, face="bold"),
               legend.key.size = unit(0.4, 'cm'),
               legend.key = element_rect(colour = "transparent", fill = alpha("white", 0)),
               legend.background = element_rect(fill = alpha("white", 0))
               )

plot_grid(change_varplot,change_var4plot, nrow = 1 , ncol = 2)
df_eff %>% ggplot() + geom_boxplot(aes(fill = Brand, y = Time, x = Temp)) +
  facet_grid(cols = vars(Stirred)) + labs(title = "Stirred") + theme(
  plot.title = element_text(hjust = 0.5)
)
##3 factor interaction plot based on HW7 code
par(mfrow=c(2,2), mar = c(3.5,3.5,2,2))
with(df_eff%>%filter(Stirred=="yes"),interaction.plot(Temp,Brand,Time,
            type="b", pch=20, col=c(2,4), ylab="", xlab = "",
            main="Mean Time vs. Temp: Stirred = Yes",
            cex.main = 0.75, cex.axis = 0.7, legend = FALSE))
legend("topright",
       title = "Brand",
       c("Name", "Store"),
       bty = "n",
       cex = 0.7,
       col = c("#DF536B","#2297E6"),
       pch = c(19,19), lty = c(2,1))
title(xlab = "Temperature", ylab = "Mean Dissolving Time", line = 2.25, cex.lab = 0.7)

with(df_eff%>%filter(Stirred=="no"),interaction.plot(Temp,Brand,Time,
          type="b", pch=20, col=c(2,4), ylab="", xlab = "",
          main="Mean Time vs. Temp: Stirred = No",
          cex.main = 0.75, cex.axis = 0.7, legend = FALSE))
legend("topright",
       title = "Brand",
       c("Name", "Store"),
       bty = "n",
```

```r
       cex = 0.7,
       col = c("#DF536B","#2297E6"),
       pch = c(19,19), lty = c(2,1))
title(xlab = 'Temperature', ylab = "Mean Dissolving Time", line = 2.25, cex.lab = 0.7)
with(df_eff%>%filter(Brand=="store"),interaction.plot(Temp,Stirred,Time,
            type="b", pch=20, col=c(2,4), ylab="", xlab = "",
            main="Mean Time vs. Temp: Brand = Store",
            cex.main = 0.75, cex.axis = 0.7, legend = FALSE))
legend("topright",
       title = "Stirred",
       c("Yes", "No"),
       bty = "n",
       cex = 0.7,
       col = c("#DF536B","#2297E6"),
       pch = c(19,19), lty = c(2,1))
title(xlab = "Temperature", ylab = "Mean Dissolving Time", line = 2.25, cex.lab = 0.7)

with(df_eff%>%filter(Brand=="name"),interaction.plot(Temp,Stirred,Time,
          type="b", pch=20, col=c(2,4), ylab="", xlab = "",
          main="Mean Time vs. Temp: Brand = Name",
          cex.main = 0.75, cex.axis = 0.7, legend = FALSE))
legend("topright",
       title = "Stirred",
       c("Yes", "No"),
       bty = "n",
       cex = 0.7,
       col = c("#DF536B","#2297E6"),
       pch = c(19,19), lty = c(2,1))
title(xlab = 'Temperature', ylab = "Mean Dissolving Time", line = 2.25, cex.lab = 0.7)
#model1
aov_eff <- aov(lm_eff <- lm(Time ~ Brand * Temp * Stirred, data = df_eff))

cooksD_values <- cooks.distance(lm_eff)

CD_plot <- ggplot() +
  geom_col(aes(y = cooksD_values, x = 1:length(cooksD_values)),
  width = 0.025, col = 'red')  +
  geom_point(aes(y = cooksD_values, x = 1:length(cooksD_values)), shape = 20) +
  xlab('Sample Points') + ylab("Cook's Distance") +
  geom_hline(yintercept = 0.25, lty = 2) +
  labs(title = "Cook's Distance") +
  theme(
      plot.title = element_text(size = 10, face = "bold"),
      axis.text = element_text(size = 7),
      axis.title = element_text(size = 8)
)

scttr_plot <-
ggplot(df_eff) + geom_point(aes(x = Order, y = Time, col = Brand, pch = Stirred)) +
 labs(title = "Time versus Order",
              x = 'Order',
              y = 'Time (seconds)') +
         theme(#legend.position = c(0.9, 0.15),
```

```r
               plot.title = element_text(size = 8, face = "bold"),
               plot.subtitle = element_text(size = 7, face = "bold"),
               axis.text = element_text(size = 7),
               axis.title = element_text(size = 8),
               legend.text = element_text(size=5, face="bold"),
               legend.title = element_text(size=6, face="bold"),
               legend.key.size = unit(0.4, 'cm'),
               legend.key = element_rect( fill = alpha("white", 0.5)),
               legend.background = element_rect(fill = alpha("white", 0.5)))

#CD_plot <- ols_plot_cooksd_chart(lm_eff)
qqplot1 <- ggplot(df_eff, aes(sample = Time)) +
  stat_qq(shape = 20) +
  stat_qq_line(linetype = "dashed", col = 'red') +
  labs(x = "Theoretical Quantiles",
       y = "Sample Quantiles",
       title = "Normal Q-Q Plot") +
  theme(
      plot.title = element_text(size = 10, face = "bold"),
      axis.text = element_text(size = 7),
      axis.title = element_text(size = 8)
)

d_plot1 <-
ggplot(df_eff) + geom_density(aes(Time, fill = Brand), adjust = 1) + xlim(c(40, 100)) +
 labs(title = "Brand Density Plots",
             x = '',
             y = '') +
         theme(legend.position = c(0.85, 0.8),
               plot.title = element_text(size = 8, face = "bold"),
               plot.subtitle = element_text(size = 7, face = "bold"),
               axis.text = element_text(size = 7),
               axis.title = element_text(size = 8),
               legend.text = element_text(size=5, face="bold"),
               legend.title = element_text(size=6, face="bold"),
               legend.key.size = unit(0.4, 'cm'),
               legend.key = element_rect(colour = "transparent", fill = alpha("white", 0)),
               legend.background = element_rect(fill = alpha("white", 0)))
d_plot2 <-
ggplot(df_eff) + geom_density(aes(Time, fill = Stirred), adjust = 1) + xlim(c(40, 100)) +
 labs(title = "Stirred Block Density Plots",
             x = '',
             y = '') +
         theme(legend.position = c(0.85, 0.8),
               plot.title = element_text(size = 8, face = "bold"),
               plot.subtitle = element_text(size = 7, face = "bold"),
               axis.text = element_text(size = 7),
               axis.title = element_text(size = 8),
               legend.text = element_text(size=5, face="bold"),
               legend.title = element_text(size=6, face="bold"),
               legend.key.size = unit(0.4, 'cm'),
               legend.key = element_rect(colour = "transparent", fill = alpha("white", 0)),
               legend.background = element_rect(fill = alpha("white", 0)))
```

```r
plot_grid(scttr_plot, qqplot1, d_plot1, d_plot2, nrow = 2, ncol = 2)
#model with stirred as block effect without interaction
#model2
aov_block_eff <- aov(lm_block_eff <- lm(Time ~ Brand * Temp + Stirred, data = df_eff))
#model3
aov_block_eff_noint <- aov(lm_block_eff <- lm(Time ~ Brand + Temp + Stirred, data = df_eff))
#added covariate Order model with stirred as block effect without interaction
#model7
##Create new Order^2 variable
df_eff$Order2<-df_eff$Order^2

aov_block_order_eff <- aov(lm_block_order_eff <- lm(Time ~ Brand * Temp + Stirred + Order +Order2, data
#model8
aov_three_order_eff <- aov(lm_three_order_eff <- lm(Time ~ Brand * Temp * Stirred + Order +Order2, data
##get correlation to add to plot
correlation <- cor(df_eff$Time,df_eff$Order)

##construct scatter plot of Time by Order with fitted SLR line
## add correlation result to this plot
gOrderTime <- ggplot(df_eff,aes(x=Order,y=Time))
gOrderTime +
    geom_point(pch = 20) + geom_smooth(method = lm, col = "Red", se = F, size = 0.5) +
    geom_text(x = 35, y = 63, size = 4, label = paste0("Correlation = ",round(correlation, 2))) +
    geom_smooth(method = lm, formula = y~poly(x,2), col="Blue", se = F, size = 0.5) +

    labs(title = "Scatter Plot of Time by Order with Fitted \nLinear=Red and Quadratic=Blue Regression L
             x = 'Order',
             y = 'Time in Seconds') +
    theme(legend.position = c(0.85, 0.8),
          plot.title = element_text(size = 8, face = "bold"),
          plot.subtitle = element_text(size = 7, face = "bold"),
          axis.text = element_text(size = 7),
          axis.title = element_text(size = 8),
          legend.text = element_text(size=5, face="bold"),
          legend.title = element_text(size=6, face="bold"),
          legend.key.size = unit(0.4, 'cm'),
          legend.key = element_rect(colour = "transparent", fill = alpha("white", 0)),
          legend.background = element_rect(fill = alpha("white", 0)))
RMSE_function <- function(df_aov){

    r_mse <- sqrt(sum(df_aov$residuals^2)/df_aov$df)

    r_s <- 1 - tail(summary(df_aov)[1][[1]][[2]], n = 1) / sum(summary(df_aov)[1][[1]][2])

    a_r_s <- 1 - (1 - r_s)*(nrow(df_aov$model) - 1)/(df_aov$df)

    aic_ <- AIC(df_aov)

    bic_ <- BIC(df_aov)

    output_stats <- c(
        r_mse,
        r_s,
```

```
        a_r_s,
        aic_,
        bic_
    )
    return(output_stats)
}


Model1 <- RMSE_function(aov_eff)
Model2 <- RMSE_function(aov_block_eff)
Model3 <- RMSE_function(aov_block_eff_noint)
Model7 <- RMSE_function(aov_block_order_eff)
Model8 <- RMSE_function(aov_three_order_eff)

d <- rbind(Model1, Model2, Model3, Model7, Model8)
colnames(d) <- c('Root MSE', '$R^2$', 'adj $R^2$', 'AIC', 'BIC')
knitr::kable(d, escape = FALSE, digits = 3, align = "ccccc")
par(mfrow=c(2,2), mar = c(5,5,2,2))
plot(lm_three_order_eff, pch = 4)
CD_plot
knitr::kable(summary(aov_eff)[[1]], 'simple', digits = 3, caption = 'Model 1: ANOVA Table')
means_eff <- emmeans(aov_eff, specs = c('Brand', 'Temp', 'Stirred'))
#summary(means_eff)
cont_str_brd <-
contrast(means_eff, list(stirred = c(1/6, 1/6, 1/6, 1/6, 1/6, 1/6, -1/6, -1/6, -1/6, -1/6, -1/6, -1/6),
                         branding = rep(c(1/6,-1/6), 6)
                         )
         )

cont_strbrd <-

contrast(means_eff, list(stirredbrand = c(1/3, 0,  1/3, 0, 1/3, 0, -1/3, 0, -1/3, 0, -1/3, 0),
                         stirredstore = c(0, 1/3, 0,  1/3, 0, 1/3, 0, -1/3, 0, -1/3, 0, -1/3)
                         )
         )
cont_temp <-
contrast(means_eff, list(temp6_23 = c(1/4, 1/4, -1/4, -1/4, 0, 0, 1/4, 1/4, -1/4, -1/4, 0, 0),
                         temp6_40 = c(1/4, 1/4, 0, 0, -1/4, -1/4, 1/4, 1/4, 0, 0, -1/4, -1/4),
                         temp23_40 = c(0, 0, 1/4, 1/4, -1/4, -1/4, 0, 0, 1/4, 1/4, -1/4, -1/4),
                         temp6_rest = c(1/4, 1/4, -1/8, -1/8, -1/8, -1/8, 1/4, 1/4, -1/8, -1/8, -1/8, -
                         temp23_rest = c(-1/8, -1/8, 1/4, 1/4, -1/8, -1/8, -1/8, -1/8, 1/4, 1/4, -1/8, -
                         temp40_rest = c(-1/8, -1/8, -1/8, -1/8, 1/4, 1/4, -1/8, -1/8, -1/8, -1/8, 1/4,
                         ), options=list(adjust="bonferroni")
         )
lm_eff_me <- lm(Time ~ Brand * Temp, data = df_eff)
aov_eff_me <- aov(lm_eff_me)
anova_eff_me <- anova(lm_eff_me)

me_table <- as_tibble(summary(aov_eff_me)[[1]][,1:3])
me_table <- cbind('Source' = c('Brand', 'Temp', 'Brand*Temp', 'Residual'),
                  me_table,
                  'Error Term' = c("MSAB", "MSAB", "MSE", "NA"))

MSA <-  me_table[1, 4]
```

```
MSB <-  me_table[2, 4]
MSAB <- me_table[3, 4]
MSE <-  round(me_table[4, 4],3)

a <- length(levels(df_eff$Brand))
b <- length(levels(df_eff$Temp))
n <- nrow(df_eff)/(a*b)

sigma_ab <- round((MSAB - MSE) / n,3)
sigma_a <- round((MSA - MSAB) / (b * n), 3)
sigma_b <- round((MSB - MSAB) / (a*n), 3)

f_scores <- c(MSA/MSAB,MSB/MSAB,MSAB/MSE, NA)

error_dof <- c(rep(tail(me_table$Df, n = 2)[1], 2), tail(me_table$Df, n = 1)[1], NA)

me_table['Error Df'] <- error_dof

me_table["F Score"] <- f_scores
f_test <- round(1 - pf(me_table['F Score'][[1]],
                       me_table['Df'][[1]],
                       me_table['Error Df'][[1]]),
                4)

me_table['Pr>F'] <- f_test

se_mu <- sqrt((MSA+MSB-MSAB)/(a*b*n))

dof_app <- (MSA + MSB - MSAB)^2 / (MSA^2/(a-1) + MSB^2/(b-1) + MSAB^2 / ((a-1)*(b-1)))

CV_hat <- sqrt(sum(c(sigma_a, sigma_b,sigma_ab, MSE)))/mean(df_eff$Time)

re_table <- as_tibble(cbind(c('Brand', 'Temp' , 'Brand*Temp', 'Residual'), round(c(sigma_a, sigma_b,sig
colnames(re_table) <- c('Cov Parm', 'Estimate')
re_table$Estimate <- as.numeric(re_table$Estimate)
re_table$Portion <- round(as.numeric(re_table$Estimate) / sum(as.numeric(re_table$Estimate)),3)
knitr::kable(df_stats, caption = "Data Summary Table")
df_eff %>% group_by(Brand) %>% summarise('Mean' = mean(Time), 'Var' = var(Time),'Max' = max(Time), 'Min
df_eff %>% group_by(Temp) %>% summarise('Mean' = mean(Time), 'Var' = var(Time), 'Max' = max(Time), 'Min
df_eff %>% group_by(Brand, Temp) %>% summarise('Mean' = mean(Time), 'Var' = var(Time), 'Max' = max(Time
df_eff %>% group_by(Brand, Stirred) %>% summarise('Mean' = mean(Time), 'Var' = var(Time), 'Max' = max(T
df_eff %>% group_by(Temp, Stirred) %>% summarise('Mean' = mean(Time), 'Var' = var(Time), 'Max' = max(Ti
knitr::kable(summary(means_eff), caption = "Least Squares Means")
knitr::kable(confint(cont_str_brd), caption = "Contrast Stirred and Brand")
knitr::kable(confint(cont_strbrd), caption = "Contrast Stirred versus Brand")
knitr::kable(confint(cont_temp), caption = "Contrast Temperatures")
knitr::kable(summary(aov_eff)[[1]], 'simple', digits = 3, caption = 'Model 1: ANOVA Table')
knitr::kable(Anova(aov_eff, type=3), 'simple', digits = 3, caption = 'Model 1: Type III ANOVA Table') #
par(mfrow=c(2,2), mar = c(5,5,2,2))
plot(aov_eff, pch = 4)
knitr::kable(summary(aov_block_eff)[[1]],
             'simple' , caption = "Model 2: ANOVA Table")
knitr::kable(Anova(aov_block_eff, type=3),
```

```
             'simple', digits = 3, caption = 'Model 2: Type III ANOVA Table')
par(mfrow=c(2,2), mar = c(5,5,2,2))
plot(aov_block_eff, pch = 4)
knitr::kable(summary(aov_block_eff_noint)[[1]],
             'simple', caption = "Model 3: ANOVA Table")
knitr::kable(Anova(aov_block_eff_noint, type=3),
             'simple', digits = 3, caption = 'Model 3: Type III ANOVA Table')
par(mfrow=c(2,2), mar = c(5,5,2,2))
plot(aov_block_eff_noint, pch = 4)

name_mean <- df_eff %>% filter(Brand == 'name') %>% select(Time) %>% unlist() %>% mean()
brand_fe <- df_eff %>% group_by(Brand) %>% summarise('Mean' = mean(Time), 'Effect' = Mean - name_mean)

name_mean <- df_eff %>% filter(Temp == '6') %>% select(Time) %>% unlist() %>% mean()
temp_fe <- df_eff %>% group_by(Temp) %>% summarise('Mean' = mean(Time), 'Effect' = Mean - name_mean)
knitr::kable(me_table, caption = 'Mixed Effects Models')

knitr::kable(brand_fe, digits = 3,
             caption = 'Model4: Brand is Fixed and Temperature is Random')

knitr::kable(cbind(c('Temp', 'Brand*Temp', 'Residual'),
                   c(sigma_b, sigma_ab, MSE)), col.names = c('Cov Parm', 'Estimate'), caption = "")
knitr::kable(temp_fe, caption = 'Model5: Brand is Random and Temperature is Fixed',
             digits = 3)
knitr::kable(cbind(c('Brand', 'Brand*Temp', 'Residual'),
                   c(sigma_a, sigma_ab, MSE)), col.names = c('Cov Parm', 'Estimate'),
             caption = "")
knitr::kable(cbind(c('Brand', 'Temp' , 'Brand*Temp', 'Residual'),
                   c(sigma_a, sigma_b,sigma_ab, MSE)), col.names = c('Cov Parm', 'Estimate'),
             caption = 'Model6: Brand and Temperature are Random')
##get correlation to add to plot
#correlation <- cor(df_eff$Time,df_eff$Order)

##construct scatter plot of Time by Order with fitted SLR line
## add correlation result to this plot
#gOrderTime <- ggplot(df_eff,aes(x=Order,y=Time))
#gOrderTime + geom_point() +geom_smooth(method=lm,col="Red",se = F) +
#  geom_text(x=35,y=63,size=6, label = paste0("Correlation = ",round(correlation, 2))) +
#  labs(title = "Scatter Plot of Time by Order with Fitted \nLinear=Red and Quadratic=Blue Regression L
#  geom_smooth(method=lm,formula = y~poly(x,2),col="Blue",se=F)


##Create new Order^2 variable
df_eff$Order2<-df_eff$Order^2

#Fit quadratic model with order and order^2
slrOrderQ<-lm(Time~Order+ Order2,data = df_eff)
#summary(slrOrderQ)

knitr::kable(summary(aov_block_order_eff)[[1]],
             'simple', caption = "Model 7: ANOVA Table")
knitr::kable(Anova(aov_block_order_eff, type=3), 'simple',
             digits = 3, caption = 'Model 7: Type III ANOVA Table')
```

```r
par(mfrow=c(2,2), mar = c(5,5,2,2))
plot(aov_block_order_eff, pch = 4)
knitr::kable(summary(aov_three_order_eff)[[1]],
             'simple', caption = "Model 8: ANOVA Table")
knitr::kable(Anova(aov_three_order_eff, type=3),
             'simple', digits = 3, caption = 'Model 8: Type III ANOVA Table')
par(mfrow=c(2,2), mar = c(5,5,2,2))
plot(lm_three_order_eff, pch = 4)
```