ST 518 Project

Rachel Hencher, Dane Korver, Mohammed Mira

2022-11-16

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

Explain what you are trying to learn from the experiment – you may borrow heavily from my description. Effectively describe the purpose of the experiment.

Experimental Design

Include a description of the experiment and the data that was collected – you may borrow heavily from my description. Effectively describe the experimental design and factors.

Below, we have a brief look at the first ten rows of the 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

Exploratory Analysis

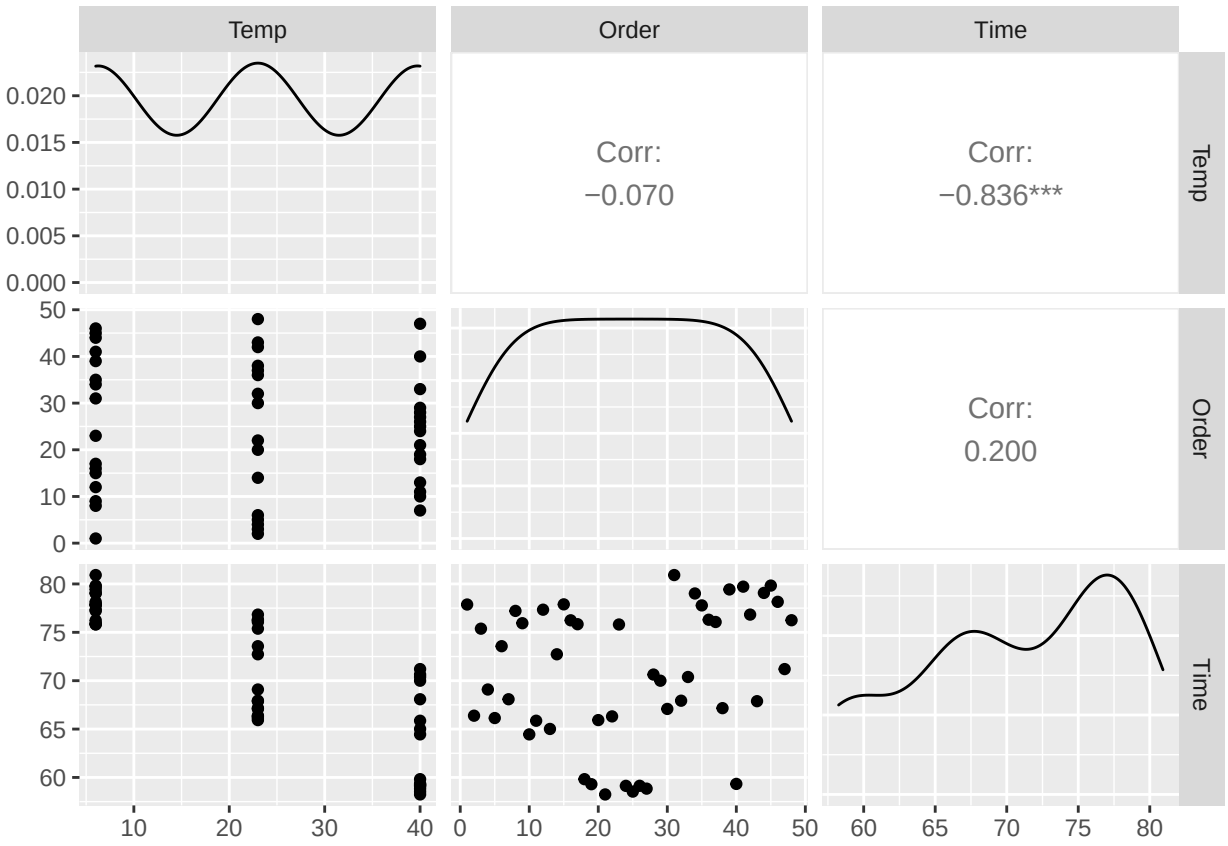
Make appropriate plots of the data and comment on any interesting features/associations you observe. You may also want to augment your visualizations with numerical computations (summary stats, correlation coefficients, etc.). Effectively describe results and insights gained from exploratory analysis; include any appropriate visualizations or computations to support your assessment.

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

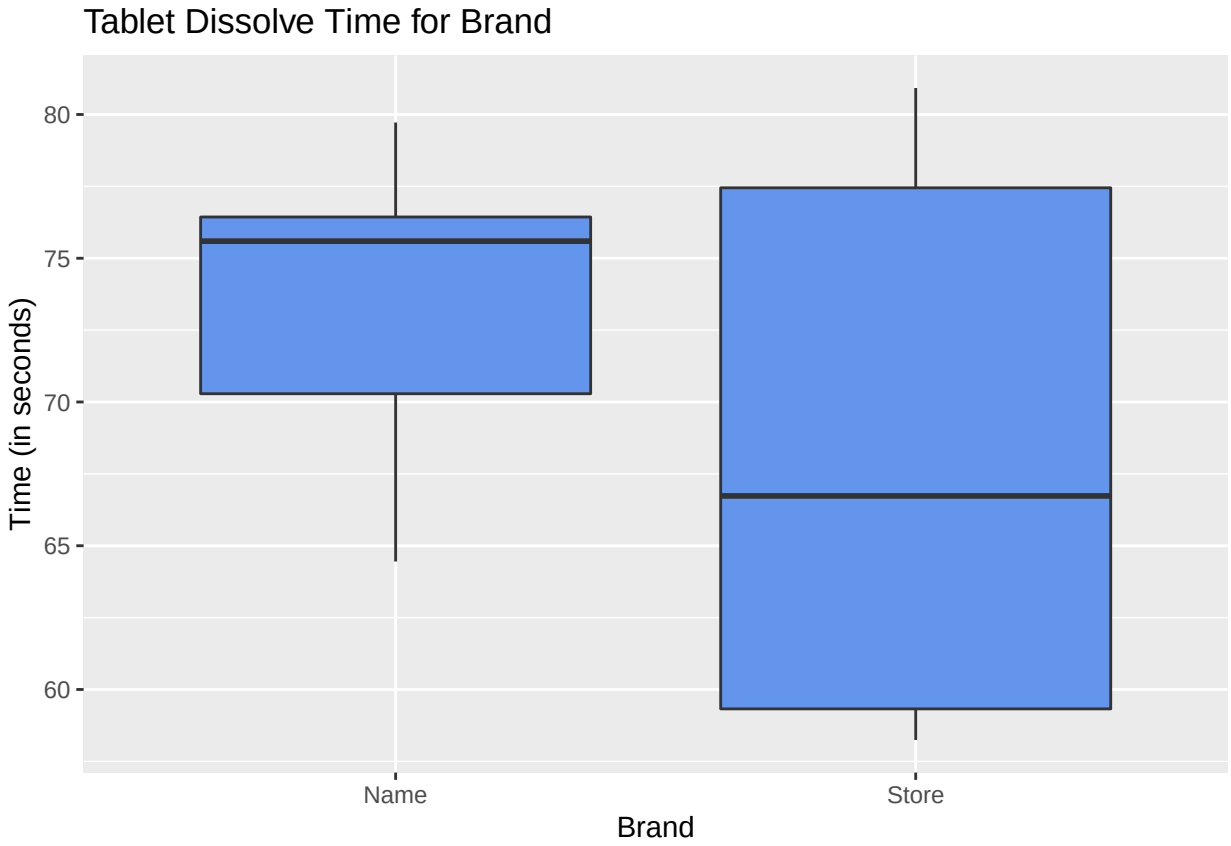
Table 2: Summary Stats for Variables

Brand	Temp	Stirred	Order	Time
name :24	Min. : 6	no :24	Min. : 1.00	Min. :58.24
store:24	1st Qu.: 6	yes:24	1st Qu.:12.75	1st Qu.:66.09
NA	Median :23	NA	Median :24.50	Median :70.92
NA	Mean :23	NA	Mean :24.50	Mean :70.77
NA	3rd Qu.:40	NA	3rd Qu.:36.25	3rd Qu.:76.93
NA	Max. :40	NA	Max. :48.00	Max. :80.92

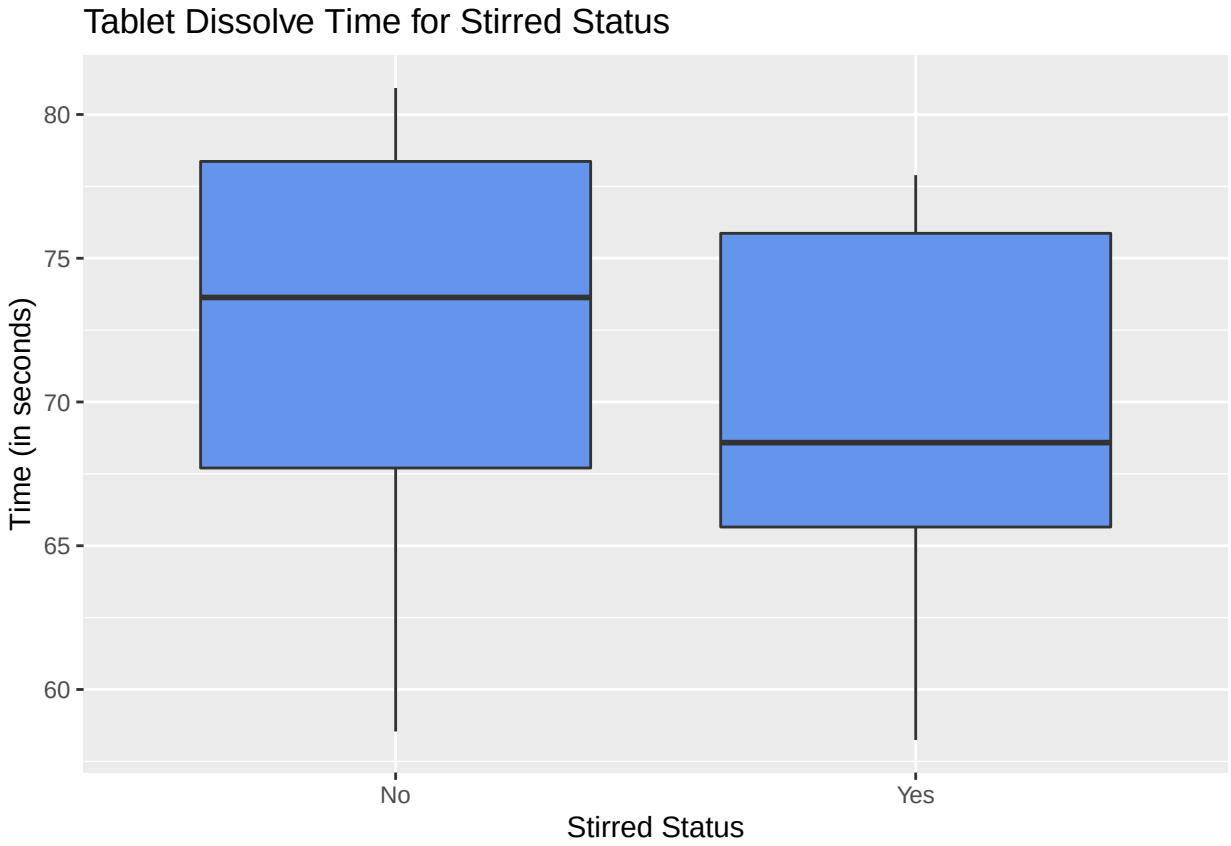
Correlations between each numeric variable can be seen below. A correlation coefficient close to -1 or 1 indicates a strong correlation between two variables and a correlation coefficient close to 0 indicates little to no correlation between two variables. For our data, we can see that there is a relatively strong correlation between `Time` and `Temperature`.



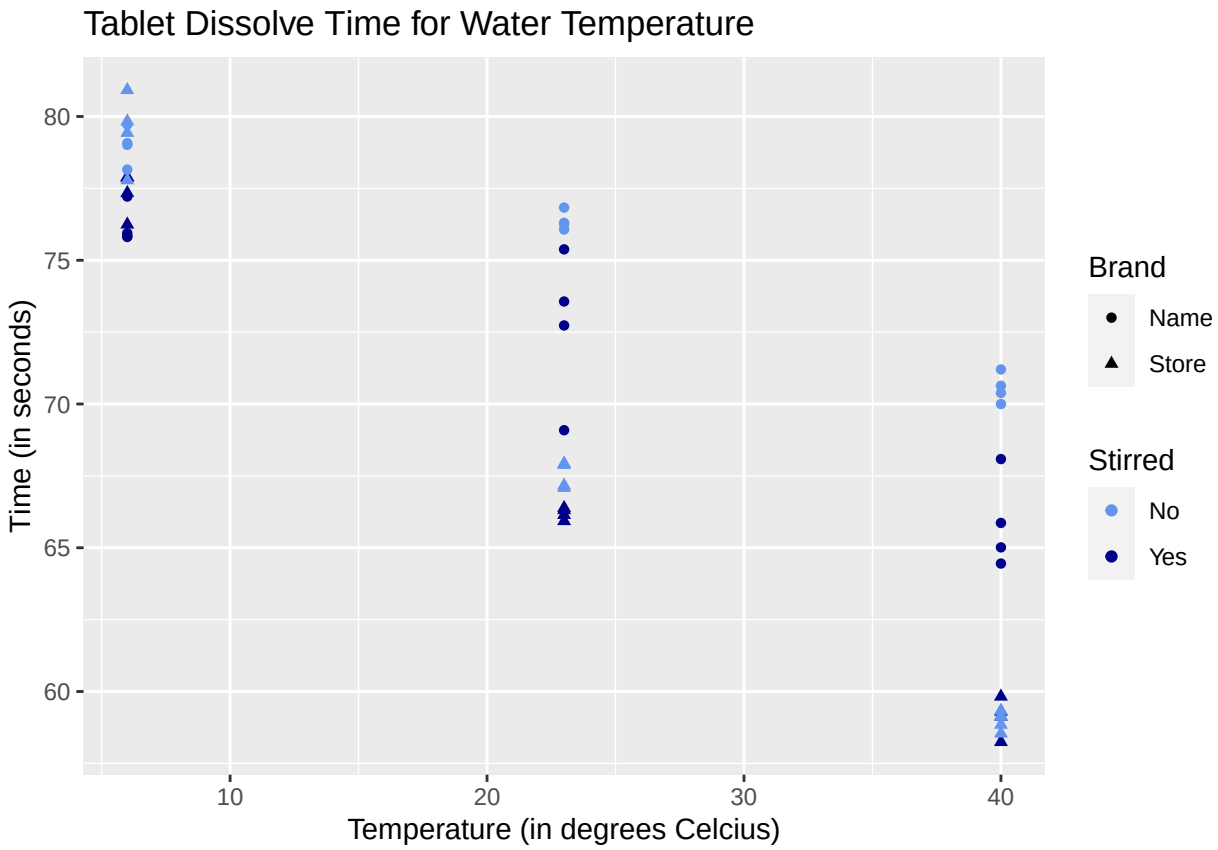
The box plot below displays a five-number summary of dissolving time for each brand of tablet. The plot displays `Time` as a function of `Brand` and indicates that there is an effect of the brand on time.



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 scatterplot 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

Fit appropriate models and follow good statistical analysis process to determine the best model to use. Make use of proper diagnostics. Choose the appropriate effects to compare, correctly estimate and test significance of the effects and trends.

Conclusion

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