26 February 2022

Tea Time! The Effects of Water and Brewing Time on the PH of Tea

Brandon Kim, Statistics Major, California Polytechnic State University
Brendan Scott Callender, Statistics Major, California Polytechnic State University
Ken Xie, Statistics Major, California Polytechnic State University

Abstract

We conducted a completely randomized design to investigate whether differences in type of water and brewing time cause differences in the pH of black tea. Our experiment was conducted using Lipton black tea, Kirkland bottled water, and tap water from our apartment. We found using bottled water produced tea that was significantly more acidic than tea that was brewed with tap water. We also found brewing the tea for 5 and 8 minutes produced tea that was significantly more acidic than tea that was only brewed for 2 minutes.

1. Introduction

Water is one of the most basic necessities in everyday life with a vast amount of usages. Whether it's through hydrating plants, basic forms of cooking or even quenching our thirst, water is basically everywhere. Such an imperative essential leads to an equally imperative question: where should we be getting our water from?

The most typical answer to that question is either bottled water or tap. However, as of recent, it has been shown that the growing use of bottled water here in America could be as a result of taste. This is backed up by a particular study by Miguel F. Doria [1], but this is more for simple drinking water. What about the different usages of water, such as cooking, does it taste different? Another study done by Medical News Today has shown that bottled water costs as much as 2000 times the price of tap water. [2] So if someone is paying that much more for bottled water in cooking, it surely has to taste different right?

Cooking with water usually comes in the form of boiling or brewing, and the most simple form of this is in the form of tea. Tea is also an extremely popular form of beverage that is especially popular amongst college students, so we had a personal interest in finding which treatment combination is the best. This also led to us creating another factor in this experiment, which was brewing time, as a very popular method of preparing tea is to over steep it as it leads to a stronger and richer taste. We measured our "taste" in pH, since sensitivity of taste buds can vary from person to person, so measuring "taste" through a numerical scale such as pH was more quantifiable.

2. Materials and Methods

This experiment was conducted at our apartment during the middle of the day (approx. 11am - 2pm). We began by filling two equally sized pots with the two types of water: one pot with tap water and the other pot with bottled water. We then boiled both waters at the same time at the same temperature until they both reached 180 degrees Fahrenheit. We maintained the two pots at 180 degrees with a thermometer. Before pouring the water inside the tea cups with tea bags, we randomly assigned tea bags to each of the cups. We collected the data for the experiment by brewing 6 cups of tea at a time. For each set of 6 cups, we followed JMP's randomized run order of type of water and brewing time and wrote them on post-it notes in front of each cup to know what water to pour in as well as the time to let it brew. We then put each assigned tea bag into the water at the same time and started the timer. Each cup of tea received 5 up and down motions of the tea bag at the beginning of the brew. At the end of each brewing time, the tea bags were taken out. Once all the tea bags were taken out, we used a digital pH pen (that read up to 2 decimal places) to measure the pH of each cup. We then recorded the pH value on our table. Between measurements we rinsed the pH pen in clean tap water and then wiped the reader dry. We repeated this process 3 additional times for 24 total cups of tea brewed.

2.1 Treatment Structure

The Treatment Structure used in our experiment was a 2x3 factorial treatment structure with 6 treatments. Our first factor with two levels was the type of water, either tap water or bottled water. Our second factor with three levels was the brewing time, either two, five, or eight minutes. We specifically chose these three brewing times as they were centered around the recommended brewing time of five minutes for black tea. We defined brewing time as the amount of time the tea bag sits in the hot water to make the tea.

2.2 Response Variable(s)

The response variable for our experiment is the pH of the brewed tea which was measured with a digital pH pen reader. Our unit of measurement is pH. We expected pH values that are similar to that of the average pH level of black tea, which is around 4.50-5.50.

2.3 Experimental Unit

The experimental units in our experiment are each cup with a teabag in it.

2.4 Design Structure

Our experiment follows a completely randomized design (CRD). Random assignment of treatments to each cup was done through using a JMP randomization table. Each treatment has 4 replicates for a total of 24 cups used in this study. There is no blocking used. We decided on 4 replicates because the assignment required at least 3 replicates. Because it was within our budget and capacity we decided to add 1 extra replicate to increase the power of the study.

2.5 Dealing with other sources of variation

We directly controlled the temperature of the tea at the start of brewing to be 180 degrees Fahrenheit. We also measured 1 cup of water into each tea cup so differences in pH wouldn't be from the tea being diluted by the amount of excess water in each cup initially. The final variable we controlled was the temperature at the time of measurement. We measured the pH of each cup 11 minutes after the start of brewing so the temperatures at the time of measurement remained fairly constant across levels of brewing time. Randomization of run order was used to ensure outside sources of variation such as outside environment temperature and water quality were spread equally across all groups.

2.6 Statistical model and data analysis

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

Where y_{ijk} is the pH for the kth cup of tea that received the ith treatment for brewing time and the jth treatment for the type of water. Since there were four replicates in each of the treatment combinations, k can take up values from k = 1,2,3,4.

α_i is defined to be the effect of ith treatment for brewing time. Where i = 1 represents 3 minutes, i = 2 represents 5 minutes, and i = 3 represents 8 minutes. So i can take values i = 1, 2, 3.

β_j is defined to be the effect of the jth treatment for the type of water. Where j = 1 represents tap water and j = 2 represents bottled water. So j can take values j = 1, 2

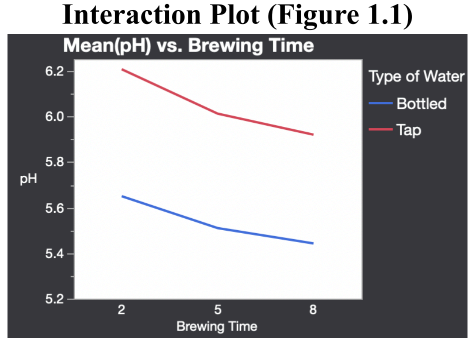
$\alpha\beta_{ij}$ is defined to be the interaction effect of the ith treatment for brewing time and the jth treatment for type of water. i and j can take the same values defined above. The interaction effect looks into how the effect of brewing time changes depending on the type of water used.

ε_{ijk} represents the random error (natural variation) for the kth cup of tea that received the ith treatment for brewing time and the jth treatment for the type of water.

3. Analysis and Results

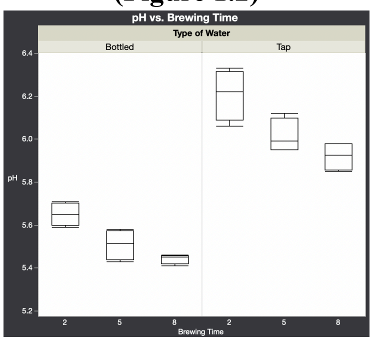
3.1 Descriptive statistics

Interaction Plot (Figure 1.1)



From the interaction plot shown in figure 1.1, we can see that the mean pH for tap water is higher than the mean pH for bottled water across all levels of brewing time. We also see that for both types of water, 2 minute brewing time has the largest mean pH followed by 5 minute brewing time, then 8 minutes. Since the two lines appear to follow the same trend it does not appear that the effect of brewing time changes based on which type of water was used and vice versa.

Side by side boxplots of pH vs. Brewing Time (Figure 1.2)



From the side by side boxplots shown in Figure 1.2, we see that there appear to be visible differences in mean pH across the type of water. There also appears to be differences in mean pH across the different brewing times within each type of water however these differences are much harder to spot.

3.2 Inferential findings

The test of choice for analyzing our data was the analysis of variance (a.k.a ANOVA). The ANOVA test tells us if the differences in the data are being caused by our treatment factors or by random variation. If the differences are being caused by our factors, ANOVA will also tell us the size of the differences in the post-hoc comparison phase.

The F-ratio is the ratio of the mean differences caused by the different treatments compared to the mean of the differences caused by random variation.

F-Ratio	66.0148

The probability of the F-Ratio being this high purely by chance was calculated to be <0.0001. This means we have significant evidence that there is at least 1 difference in mean pH for one of the treatments. Now we can look into which treatment or treatments are causing major differences in pH.

Effect Tests (Table 1.3)

Source	F-Ratio	P-value
Brewing Time	22.9194	<0.0001
Type of Water	283.0319	<0.0001
Brew*Water	0.6016	0.5586

Because the F-ratio for both brewing time and type of water, in Table 1.3 above, are large, there is significant evidence to show that both brewing time and type of water causes a change in pH of black tea. Now since we know both brewing time and type of water cause a difference in pH, our next logical step is to check how much of a difference both factors make. Since the F-ratio for the interaction effect is small, we will look into the main effects of brewing time and type of water separately since they have no effect on each other.

Below are two letters plots (Table 1.4). Letter plots show the means of the levels of each factor and assign each level letters. Levels that share a letter are not significantly different. On the other hand, levels that don’t share a letter are significantly different.

Letters Plot: Brewing Time (Table 1.4)

Brewing Time	Letter	Mean
2	A	5.929
5	B	5.761
8	B	5.681

Table 1.4 is the letter plot for the factor brewing time. We see that brewing tea for 5 and 8 minutes makes tea that is significantly more acidic than tea that was brewed for only 2 minutes.

Letters Plot: Type of Water (Table 1.5)

Type of Water	Letter	Mean
Tap	A	6.047
Bottle	B	5.534

Table 1.5 is the letter plot for the factor type of water. We see that tap water makes tea that is significantly more basic than tea that is made from bottled water.

In order to use ANOVA to analyze our data, our data must pass certain assumptions:
The three assumptions are regarding ε_{ijk}.

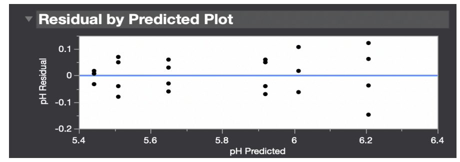
1. Each ε_{ijk} must be independent
2. ε_{ijk} are normally distributed
3. ε_{ijk} have equal variances

Because the data comes from a study in which we used random assignment and randomization, we are allowed to assume this assumption is satisfied.

In order to conclude normality, we run the estimated errors (actual values - predicted values) through the Shapiro-Wilk test which finds the probability of the given data occurring given the data came from a normal distribution. Low probabilities indicate it is unlikely the data follows a normal trend. Since our data produced a p-value of 0.6713, we are safe to assume the ε_{ijk} came from a normal distribution.

In order to conclude equal variance we need to look at Figure 1.3. Figure 1.7 (below) shows the predicted values against the difference of the predicted value and the actual value. We want the heights of the columns in the graph to be roughly the same. For the case below, this is not necessarily satisfied. However, because our experiment was a balanced design, meaning all treatment combinations had the same number of observations, we are safe to carry on with ANOVA and this assumption is not violated.

Residuals by Predicted Plot (Figure 1.7):



4. Conclusion

We have strong evidence to show that using tap water will on average lead to more basic water (higher pH) than using bottled water. We also found evidence that steeping the tea for 2 minutes will lead to on average more basic water than steeping the tea for 5 minutes or 8 minutes, with both 5 minutes and 8 minutes leading to results that didn't significantly differ. We don't have any evidence of an existence of an interaction between these two factors.

In terms of what's "best," we personally believe that combination of bottled water and 8 minute steeping time was the best, as it led to the result that was within the average black tea pH range (4.50-5.50 pH). The treatment that had the largest effect (most different from overall mean) was using tap water and was steeped for 2 minutes and the treatment that had the smallest effect (least different from overall mean) was using bottled water and was steeped for 2 minutes.

5. Next steps

If we were to conduct this study again, there would be a few things we would change. Next time, we would try to brew more tea at each batch, so it would eliminate as much error caused by having multiple batches of tea brewed at different times. Since our data was collected for only black tea, the main next step would be to look at how these results change for different types of tea. Perhaps a study similar to this one that blocked on type of tea would be an interesting follow up.

References

1. Doria MF. Bottled water versus tap water: understanding consumers' preferences. J Water Health. 2006 Jun;4(2):271-6. PMID: 16813019.
2. Leonard , Jayne. "Bottled Water vs. Tap Water: Pros and Cons." Edited by Natalie Butler, *Medical News Today*,

MediLexicon International, 2 Jan. 2020, https://www.medicalnewstoday.com/articles/327395#pros-and-cons-of-tap-wate

Appendix

ANOVA and Effect tests:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	1.8378708	0.367574	66.0148
Error	18	0.1002250	0.005568	Prob > F
C. Total	23	1.9380958		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	5.7904167	0.015232	380.16	<.0001*
Brewing Time[2]	0.1383333	0.021541	6.42	<.0001*
Brewing Time[5]	-0.029167	0.021541	-1.35	0.1925
Type of Water[Bottled]	-0.25625	0.015232	-16.82	<.0001*
Type of Water[Bottled]*Brewing Time[2]	-0.0225	0.021541	-1.04	0.3101
Type of Water[Bottled]*Brewing Time[5]	0.005	0.021541	0.23	0.8191

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Brewing Time	2	2	0.2552333	22.9194	<.0001*
Type of Water	1	1	1.5759375	283.0319	<.0001*
Type of Water*Brewing Time	2	2	0.0067000	0.6016	0.5586

Letter Plots:

Brewing Time

Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
2	5.9287500	0.02638194	5.92875
5	5.7612500	0.02638194	5.76125
8	5.6812500	0.02638194	5.68125

LSMeans Differences Student's t

α= 0.050 t= 2.10092

LSMean[j]

LSMean[i]	Mean[i]-Mean[j] / Std Err Dif / Lower CL Dif / Upper CL Dif	2	5	8
2	Mean[i]-Mean[j]	0	0.1675	0.2475
	Std Err Dif	0	0.03731	0.03731
	Lower CL Dif	0	0.08912	0.16912
	Upper CL Dif	0	0.24588	0.32588
5	Mean[i]-Mean[j]	-0.1675	0	0.08
	Std Err Dif	0.03731	0	0.03731
	Lower CL Dif	-0.2459	0	0.00162
	Upper CL Dif	-0.0891	0	0.15838
8	Mean[i]-Mean[j]	-0.2475	-0.08	0
	Std Err Dif	0.03731	0.03731	0
	Lower CL Dif	-0.3259	-0.1584	0
	Upper CL Dif	-0.1691	-0.0016	0

Level				Least Sq Mean
2	A			5.9287500
5		B		5.7612500
8			C	5.6812500

Levels not connected by same letter are significantly different.

Type of Water

Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
Bottled	5.5341667	0.02154077	5.53417
Tap	6.0466667	0.02154077	6.04667

LSMeans Differences Student's t

α= 0.050 t= 2.10092

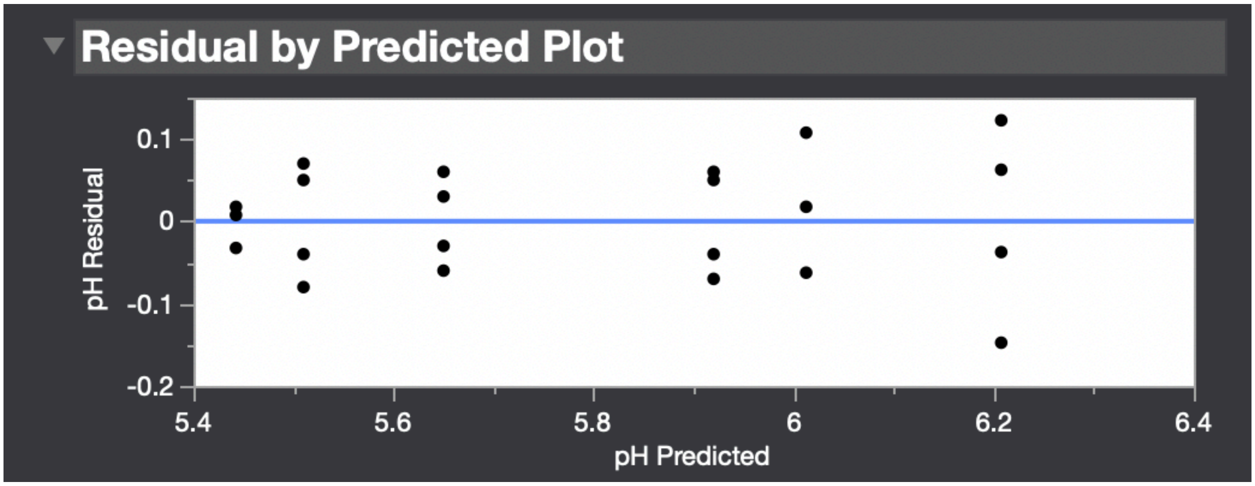
LSMean[j]

LSMean[i]	Mean[i]-Mean[j] / Std Err Dif / Lower CL Dif / Upper CL Dif	Bottled	Tap
Bottled	Mean[i]-Mean[j]	0	-0.5125
	Std Err Dif	0	0.03046
	Lower CL Dif	0	-0.5765
	Upper CL Dif	0	-0.4485
Tap	Mean[i]-Mean[j]	0.5125	0
	Std Err Dif	0.03046	0
	Lower CL Dif	0.4485	0
	Upper CL Dif	0.5765	0

Level			Least Sq Mean
Tap	A		6.0466667
Bottled		B	5.5341667

Levels not connected by same letter are significantly different.

Assumption Output:



Goodness-of-Fit Test

	W	Prob<W
Shapiro-Wilk	0.9701761	0.6713