802 Project Summary

Maksuda Aktar Toma, Jo Charbonneau, Ryan Lalicker

December 15, 2024

Introduction

This paper summarizes the consulting that was done for our assigned STAT 802 group. For more information on the experiment, the data, or any other files used in this paper see our Github page which can be found at <https://github.com/maksudatoma/Stat-802-Project>. The coding languages used in the paper are R and SAS. The corresponding code can be found in *Appendix A - R Code* and *Appendix B - SAS Code and Additional Output* respectively.

Initial Meetings

The first meeting with our clients was on September 13th. We discussed their project and what kind of data they were going to be looking at. They detailed to us their project, which is looking at the levels of salmonella in beef jerky at different inoculations and thicknesses. Prior to the meeting they sent us what their variables would be, which gave us a good idea of what might be the best experimental design. The group informed us they were avoiding a completely randomized design (CRD) at the request of their professor. With that in mind, we suggested other possible models.

Later, after receiving feedback from Dr. Howard and several PhD students within the statistics department, we suggested adding a time component to the experiment and creating multiple batches to replicate each treatment combination. This lead to us suggesting a mixed model for the analysis.

In both the initial meeting and the follow-up session the clients were more than happy to implement our suggestions. In the end the experiment involved two thickness levels (one-fourth and one-eighth of an inch), two inoculation methods (dry and wet), and five evenly spaced time points where measurements were taken (weeks 1-5) creating twenty entries per batch. The exact number of batches would not be known until after the power analysis, found in the *Power Analysis* section. We provided the client an example dataset we created to give them a better idea of what the end product may look like. This dataset had five batches.

Study Objectives and Proposed Model

The clients were most interested in the effect of the thickness levels, the inoculation method, and their interaction had on the salmonella levels. In the final model we included the week effect and subsequent interactions as well. These variables are the fixed effects in the proposed mixed model.

The other variable included in the experiment is the batch number. This is therefore treated as a random variable. As mentioned above, the exact number of batches needed was unknown prior to the power analysis, but five was used as a starting value.

Overall the study employs a 2×2 factorial design with two main factors: Inoculation Method (Dry, Wet) and Thickness (1/4-inch, 1/8-inch). Repeated measurements are taken over five equally spaced time points (Weeks 1 to 5), allowing the analysis of both main effects, their interaction, and changes over time.

The model can be written in the form

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \tau_k + (\alpha\beta)_{ij} + (\alpha\tau)_{ik} + (\beta\tau)_{jk} + (\alpha\beta\tau)_{ijk} + u_l + e_{ijkl}$$

Here, Y_{ijkl} is the salmonella level and μ is the overall mean. The fixed effects are represented by α_i for the effect of the ith inoculation method, β_j for the effect of the jth thickness level, and τ_k for the effect of the kth week. The interaction effect of the ith inoculation method and the jth thickness level is represented by $(\alpha\beta)_{ij}$, with the other two-way interactions following this form. The three-way interaction between all fixed effects is represented as $(\alpha\beta\tau)_{ijk}$. The random effect for batches is represented by u_l, which we assume are distributed as $u_l \sim N(0, \sigma_u^2)$. Lastly, the residuals are represented by e_{ijkl}, which we assume can be distributed as $e_{ijkl} \sim N(0, \sigma^2)$.

Power Analysis

Before power analysis, We reached out to the client later on in the process to determine what contrasts they were most interested in testing. They expressed they wanted to see the difference between the two levels of the inoculation method, the two levels of the thickness, and the orthogonal contrasts these variable. This resulted in six contrasts being tested.

To determine the necessary number of batches needed to increase the likelihood of detecting a true treatment effect, we performed a power analysis. To do this, probable treatment mean estimates across all five weeks and variance estimates were needed. The clients provided these metrics from Brown et al. (2024). We then used these metrics to create a dataset with five batches where the response variable was identical across the batches. This dataset was then evaluated to determine the power. The results of the power analysis perfomed in SAS are shown below.

Obs	Label	NumDF	DenDF	FValue	ProbF	Effect	ncparm	alpha	fcrit	power
1	Dry vs Wet	1	76	9.23	0.0033		9.226	0.05	3.96676	0.85058
2	1/4 vs 1/8 inches	1	76	103.28	<.0001		103.276	0.05	3.96676	1.00000
3	Dry vs Wet at 1/4 Inches	1	76	5.70	0.0195		5.695	0.05	3.96676	0.65406
4	Dry vs Wet at 1/8 Inches	1	76	44.65	<.0001		44.651	0.05	3.96676	1.00000
5	1/4 vs 1/8 inches for Dry inoculation	1	76	137.36	<.0001		137.364	0.05	3.96676	1.00000
6	1/4 vs 1/8 inches for Wet inoculation	1	76	7.03	0.0097		7.031	0.05	3.96676	0.74475
7		1	76	9.23	0.0033	Inoculation_Method	9.226	0.05	3.96676	0.85058
8		1	76	103.28	<.0001	Thickness	103.276	0.05	3.96676	1.00000
9		1	76	41.12	<.0001	Inoculatio*Thickness	41.120	0.05	3.96676	0.99999
10		4	76	7.82	<.0001	Week	31.280	0.05	2.49205	0.99659
11		4	76	13.77	<.0001	Inoculation_Met*Week	55.092	0.05	2.49205	1.00000
12		4	76	20.78	<.0001	Thickness*Week	83.132	0.05	2.49205	1.00000
13		4	76	4.04	0.0050	Inocula*Thickne*Week	16.162	0.05	2.49205	0.89525

Figure 1: Results of power analysis.

The first six rows of the table correspond to the contrasts the clients were interested in testing, while the bottom seven rows are measuring the fixed effects of the model. Many of the terms have more than 80% power. Specifically the fixed effects were all high enough for both the clients and ourselves to feel comfortable using five batches. Two of the orthogonal contrasts, `Dry vs Wet at 1/4 Inches` and `1/4 vs 1/8 inches for Wet inoculation` did have lower power scores, but after talking with both the clients and Dr. Howard about them, we felt comfortable to proceed.

Simulating Data

After finding the necessary number of batches, which was five, we proceeded with simulating the data. The estimated treatment means and variances provided by the client were used in the simulation as well. We then reviewed the simulated dataset for major issues, such as negative response values, and reran the power analysis on the new data set to ensure everything was working properly. After finding no problems with the dataset, we sent it to the clients. Note, the simulation was performed in SAS.

Data Analysis

Summary Statistics

As part of the project, we analyzed the simulated dataset. Before fitting out model to the dataset, we first wanted to explore some of the variables. Figure 2 shows the mean values and

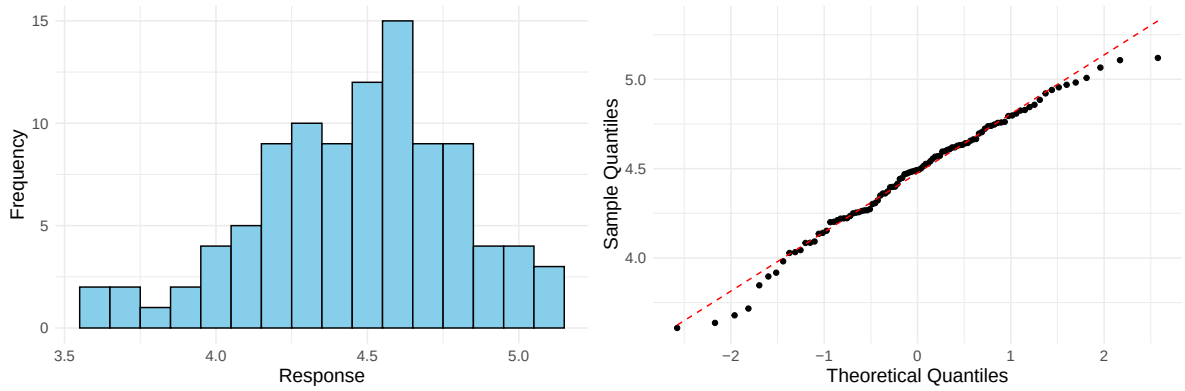
standard deviations for each treatment combination. We can see the changes in mean values are small, so further exploration and analysis are needed.

Inoculation Method	Thickness	Week	Mean Response	SD Response	Count
Dry	1/4-inch	1	4.199	0.187	5
Dry	1/4-inch	2	4.137	0.163	5
Dry	1/4-inch	3	4.187	0.294	5
Dry	1/4-inch	4	4.306	0.272	5
Dry	1/4-inch	5	4.645	0.178	5
Dry	1/8-inch	1	4.801	0.173	5
Dry	1/8-inch	2	4.801	0.238	5
Dry	1/8-inch	3	4.596	0.235	5
Dry	1/8-inch	4	4.425	0.206	5
Dry	1/8-inch	5	4.876	0.177	5
Wet	1/4-inch	1	4.240	0.307	5
Wet	1/4-inch	2	4.523	0.236	5
Wet	1/4-inch	3	4.534	0.173	5
Wet	1/4-inch	4	4.357	0.421	5
Wet	1/4-inch	5	4.299	0.200	5
Wet	1/8-inch	1	4.697	0.181	5
Wet	1/8-inch	2	4.809	0.327	5
Wet	1/8-inch	3	4.530	0.215	5
Wet	1/8-inch	4	4.223	0.361	5
Wet	1/8-inch	5	4.151	0.295	5

Figure 2: Summary of response rariable across factors and weeks

Distribution of response variable

Before continuing our investigation into the relationships among the treatment variables, we want to look into the response variable (salmonella levels). Specifically, we want to see how it is distributed. Figure 3 shows a histogram and Q-Q plot of the response variable in the left and right plots respectively. While the histogram shows a slight potential skew, this is not enough for us say the distribution is non-normal. Furthermore, the Q-Q plot indicates the response variable follows a relatively normal distribution.



(a) Histogram (b) Q-Q plot

Figure 3: Histogram and Q-Q plot of the response variable.

Exploring the Data

Now we will graphically look at how the different variables of the model impact the response variable. Figure 4 shows two interaction plots with one being for samples cut to 1/4-inches and the other for 1/8-inch samples. Each one shows the relationship between the two inoculation methods over the five weeks for the respective thickness. At the 1/4-inch level, the dry and wet methods start at a similar level but then change over time. The dry method appears to stay a lower level for the first few weeks, but then rise sharply. The opposite trend occurs for the wet method. At the 1/8-inch level, the two methods are similar until the fifth week when the measurement of the wet method decreases and the dry method increases. In both plots we seem to have interaction among the the variables, so that will need to be investigated further.

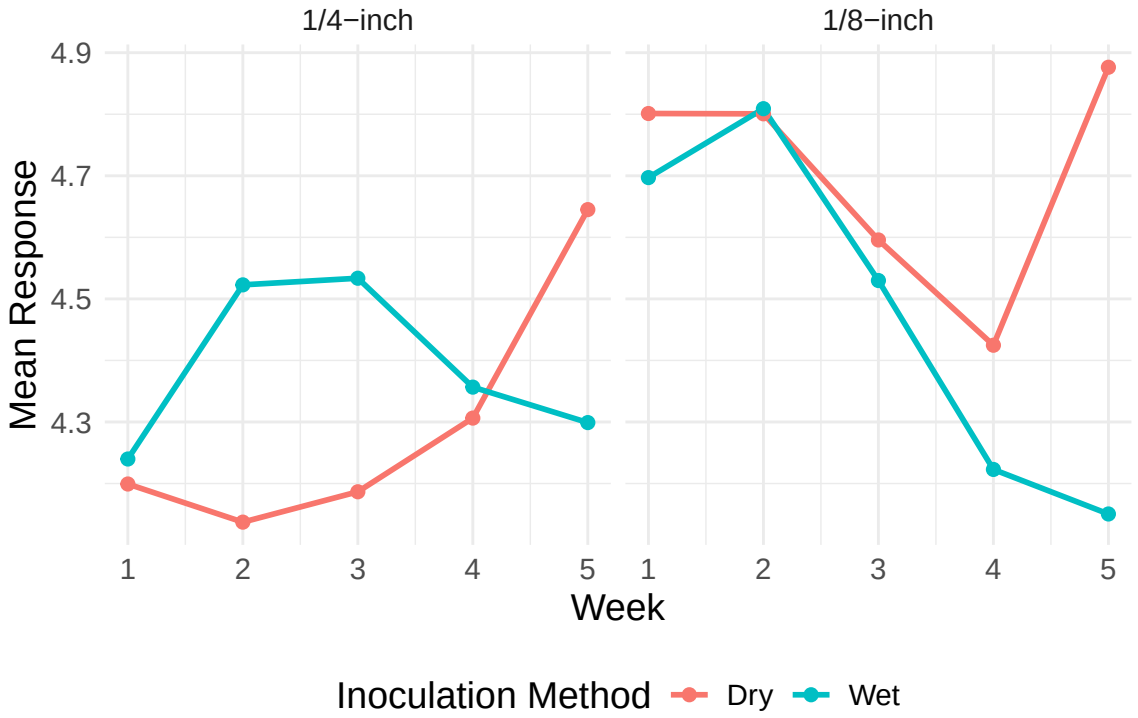


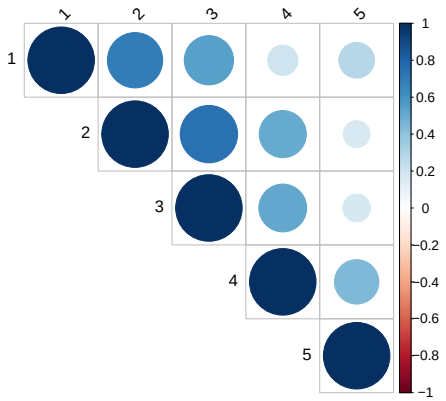
Figure 4: Interaction plots of inoculation method and time for each thickness level.

Another element to consider is correlation over time. Since this is a repeated measures experiment we need to account for this correlation by selecting a type of covariance structure. There are several types of structures such as variance components (VC), unstructured (UN), compound symmetry (CS) and heterogeneous compound symmetry (CSH), p-order auto-regressive (AR(p)) and heterogeneous auto-regressive (ARH(p)), p-order ante-dependence (ANTE(p)), and Toeplitz (TOEP) among others. For more information on these see Lipka and Tyner (2004) and USDA (n.d.). The AR(p) structure fits data that is ordered through time an equally spaced. For that reason, our initial plan was to use this as the covariance structure with $p = 1$.

To see if this first-order auto-regressive structure might fit the data, let's consider the table and plot in Figure 5. The correlation matrix (left) and plot (right) show the relationships between repeated measurements over weeks one to five. Strong correlations are observed between adjacent weeks (e.g., Week 1 vs. Week 2, r=0.69, Week 2 vs. Week 3, r=0.74), indicating temporal dependency. Correlations weaken as the time gap increases (e.g., Week 1 vs. Week 5, r=0.29), leading us to believe orders of $p > 1$ are not necessary. This can be seen visually in the plot which uses circle size and color to model the correlation metrics seen in the matrix. We can see as the gap between weeks increases, the circles become smaller and lighter. This pattern supports the use of models like AR(1).

Time	1	2	3	4	5
1	1.00	0.69	0.55	0.21	0.29
2	0.69	1.00	0.74	0.50	0.16
3	0.55	0.74	1.00	0.52	0.17
4	0.21	0.50	0.52	1.00	0.45
5	0.29	0.16	0.17	0.45	1.00

(a) Correlation matrix



(b) Correlation matrix

Figure 5: Table and plots to see the correlation across time.

Model Comparison

After exploring the data we can move on to fitting the model. While we were confident in using the AR(1), we chose to fit the model using other covariance structures as well so we could see how the fit compares. The results are shown in Figure 6. For each of the fit statistics in this table a lower score is better, even when looking at negative values. (Bobbit (2021)). This means the model fit using an AR(1) structure had the best AIC and AICC scores and a respectable BIC score. This verifies our choice in the AR(1) covariance structure.

Model	AIC	AICC	BIC
VC	-12.25	-12.09	-13.03
UN	-12.22	-3.58	-18.47
CS	-11.60	-11.29	-12.77
AR(1)	-12.86	-12.55	-14.03
ARH(1)	-12.56	-11.01	-15.30
ANTE(1)	-8.57	-5.39	-12.48
TOEP	-7.34	-6.19	-9.68

Figure 6: Model comparison table.

After verifying the fit of the AR(1) covariance structure for the repeated measures, we needed to see if the assumptions for a linear mixed model are violated or not. These include the residuals being normally distributed and homogeneous. The plots in Figure 7 allow us to evaluate these assumptions. To graphically test normality, we can look at both the histogram (top right) and the Q-Q plot (bottom left). These both appear approximately normal, indicating the assumption holds. The boxplot (bottom right) can also show normality as well as potential

outliers. It appears there is one outlier, but the normality assumption still holds. The residual plot (top left) allows us to check if the homogeneous assumption holds, and it appears to since the points seem somewhat randomly distributed with no clear pattern. Since the assumptions are holding, we can proceed with the linear mixed model.

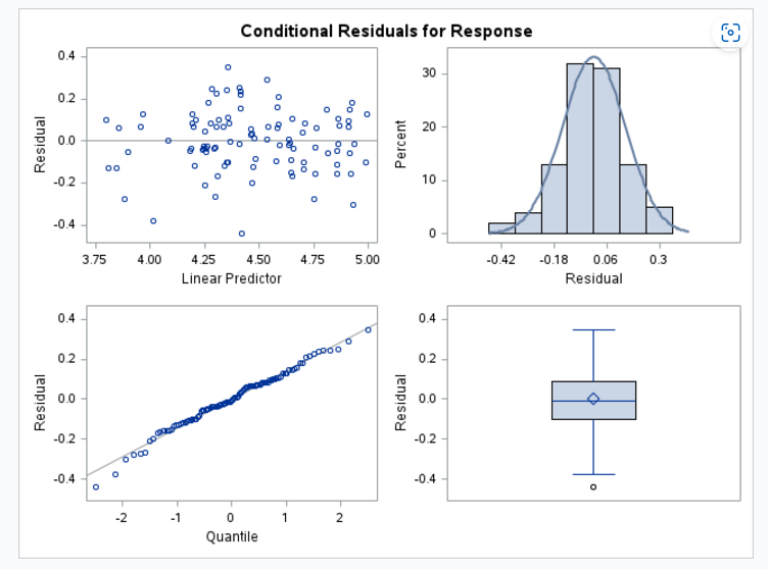


Figure 7: Residual plots for checking assumptions.

Model Output

We can now look at some of the output from fitting the linear mixed model in SAS. Figure 8 shows three tables from this output. The first is the *Fit Statistics* table, which is where the metrics used in Figure 6 come from. These metrics indicate the fit of the model, which seems fine to us. The second table is the *Covariance Parameter Estimates* table and it shows how much of the variance in the model is explained by the random terms. We can see all three values are fairly small, with the estimated variance for the AR(1) autocorrelation structure actually being negative (-0.2140). This negative value shows a weak negative correlation between adjacent time points. The small values for the estimated residual variance (0.02664) is good news as it means there is minimal unexplained variability in this model.

Fit Statistics	
-2 Res Log Likelihood	-18.86
AIC (smaller is better)	-12.86
AICC (smaller is better)	-12.55
BIC (smaller is better)	-14.03
CAIC (smaller is better)	-11.03
HQIC (smaller is better)	-16.01
Generalized Chi-Square	2.13
Gener. Chi-Square / DF	0.03

Covariance Parameter Estimates			
Cov Parm	Subject	Estimate	Standard Error
Intercept	Batches	0.03802	0.02755
AR(1)	Batche*Inocul*Thickn	-0.2140	0.1286
Residual		0.02664	0.004423

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Inoculation_Method	1	76	4.98	0.0286
Thickness	1	76	81.92	<.0001
Inoculatio*Thickness	1	76	32.65	<.0001
Week	4	76	6.12	0.0003
Inoculation_Met*Week	4	76	17.90	<.0001
Thickness*Week	4	76	14.59	<.0001
Inocula*Thickne*Week	4	76	0.55	0.6970

Figure 8: *Fit Statistics*, *Covariance Parameter Estimates*, and *Type III Tests of Fixed Effects* tables.

The last table in Figure 8 is the *Type III Tests of Fixed Effects* table. This allows us to see if the fixed effects are significant by looking at the p-values reported in the `Pr > F` column. We must first consider the significance of the interaction effects and only look at main effects if the interaction effects are insignificant. While the three-way interaction between the inoculation method, thickness, and week is not significant, all three two-way interactions between these variables are highly significant ($p<0.0001$). Therefore, we need to consider simple effects. Overall we can say each respective two-way interaction does have an impact on the data. This validates what we saw in Figure 4.

Figure 9 shows the output for the contrasts we are evaluating at the client's request. While both the inoculation method and thickness are significant, this is just repeating the main effects for each variable that we saw in Figure 8. Therefore, we need to focus on the interaction terms. All interaction contrasts are significant, with `Dry vs. Wet at 1/8 inches` and `1/4 vs. 1/8 inches for Dry inoculation` being highly significant. This requires further investigation.

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
Dry vs Wet	1	76	4.98	0.0286
1/4 vs 1/8 inches	1	76	81.92	<.0001
Dry vs Wet at 1/4 Inches	1	76	6.06	0.0161
Dry vs Wet at 1/8 Inches	1	76	31.56	<.0001
1/4 vs 1/8 inches for Dry inoculation	1	76	108.99	<.0001
1/4 vs 1/8 inches for Wet inoculation	1	76	5.57	0.0209

Figure 9: Contrasts table.

Since we found three significant two-way interactions we will need to look at the *Least Squares Means* output for each interaction. The first one we are looking at is for the interaction between inoculation method and thickness. This can be seen in Figure 10. These tables allow us to see how the variables interact at each level.

Inoculatio*Thickness Least Squares Means									
Inoculation_Method	Thickness	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Dry	1/4-inch	4.2949	0.09141	76	46.98	<.0001	0.05	4.1128	4.4770
Dry	1/8-inch	4.6997	0.09141	76	51.41	<.0001	0.05	4.5177	4.8818
Wet	1/4-inch	4.3904	0.09141	76	48.03	<.0001	0.05	4.2083	4.5724
Wet	1/8-inch	4.4819	0.09141	76	49.03	<.0001	0.05	4.2998	4.6639

Differences of Inoculatio*Thickness Least Squares Means											
Inoculation_Method	Thickness	_Inoculation_Method	_Thickness	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Dry	1/4-inch	Dry	1/8-inch	-0.4048	0.03878	76	-10.44	<.0001	0.05	-0.4821	-0.3276
Dry	1/4-inch	Wet	1/4-inch	-0.09547	0.03878	76	-2.46	0.0161	0.05	-0.1727	-0.01824
Dry	1/4-inch	Wet	1/8-inch	-0.1870	0.03878	76	-4.82	<.0001	0.05	-0.2642	-0.1097
Dry	1/8-inch	Wet	1/4-inch	0.3094	0.03878	76	7.98	<.0001	0.05	0.2321	0.3866
Dry	1/8-inch	Wet	1/8-inch	0.2179	0.03878	76	5.62	<.0001	0.05	0.1406	0.2951
Wet	1/4-inch	Wet	1/8-inch	-0.09150	0.03878	76	-2.36	0.0209	0.05	-0.1687	-0.01427

Tests of Effect Slices for Inoculatio*Thickness Sliced By Inoculation_Method				
Inoculation_Method	Num DF	Den DF	F Value	Pr > F
Dry	1	76	108.99	<.0001
Wet	1	76	5.57	0.0209

Tests of Effect Slices for Inoculatio*Thickness Sliced By Thickness				
Thickness	Num DF	Den DF	F Value	Pr > F
1/4-inch	1	76	6.06	0.0161
1/8-inch	1	76	31.56	<.0001

Figure 10: *Least Squares Means* output for inoculation method and thickness interaction.

We can see from the *Least Squares Means* table that all the combinations are significant, but we want to look closer at the differences in the second table. Once again all the differences between each possible combination of inoculation method and thickness are significant since the largest p-value is 0.0209. We can see the largest difference is between the combinations

of 1/4-inch with dry inoculation and 1/8-inch with dry inoculation. The negative difference implies the 1/8-inch pieces have significantly lower values, which is what we want. The smallest difference was involved the two thicknesses with wet inoculation. This indicates both perform in a similar way with the 1/8-inch pieces performing slightly but significantly better.

Inoculation_Met*Week Least Squares Means

Inoculation_Method	Week	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Dry	1	4.5002	0.1013	76	44.41	<.0001	0.05	4.2984	4.7020
Dry	2	4.4690	0.1013	76	44.10	<.0001	0.05	4.2671	4.6708
Dry	3	4.3912	0.1013	76	43.33	<.0001	0.05	4.1894	4.5930
Dry	4	4.3655	0.1013	76	43.08	<.0001	0.05	4.1637	4.5673
Dry	5	4.7607	0.1013	76	46.98	<.0001	0.05	4.5589	4.9625
Wet	1	4.4685	0.1013	76	44.10	<.0001	0.05	4.2666	4.6703
Wet	2	4.6659	0.1013	76	46.05	<.0001	0.05	4.4641	4.8677
Wet	3	4.5317	0.1013	76	44.72	<.0001	0.05	4.3299	4.7335
Wet	4	4.2898	0.1013	76	42.33	<.0001	0.05	4.0879	4.4916
Wet	5	4.2248	0.1013	76	41.69	<.0001	0.05	4.0230	4.4266

Figure 11: *Least Squares Means* table for inoculation method and week interaction.

Moving on to the inoculation method and week interaction, Figure 11 shows us that the estimated response variables differ for the treatment combinations. The largest estimate is the dry samples in week 5 (4.7607), while that same week is the lowest estimate for the wet samples (4.2249). Additionally week 2 has the highest estimate for the wet samples. Similarly, Figure 12 shows that the interaction between the two variables is most significant at week 5 (p-value$<$0.0001), followed by week 2 (p-value=0.0086). This suggests the impact of the inoculation method depends on the week of measurement.

Figure 15, which shows the differences among the different combinations of inoculation methods and weeks, can be found in *Appendix B - SAS Code and Additional Output*. Note that the second table in Figure 12 comes from the comparison of the dry versus wet method for a given week found in Figure 15. For the comparisons involving different weeks, most of the significant differences involve cases where there are two or more weeks between the treatment combinations, but some of the largest significant estimated differences occurred in weeks 4 and 5.

The last of the three two-way interactions is for the thickness and the weeks. Figure 13 shows that the largest estimated response variable is 1/8-inch samples in week 2 (4.8049) followed by the same thickness the week prior (4.7491). Meanwhile, the 1/4-inch samples over the five weeks account for five of the six lowest estimates. This indicates the 1/4-inch samples are generally better at slowing bacterial growth over time. They are also more consistent than the 1/8-inch pieces.

Adding to this, Figure 16 in *Appendix B - SAS Code and Additional Output* shows the differences among the various combinations of the thicknesses and the weeks. The differences

Tests of Effect Slices for Inoculation_Met*Week Sliced By Inoculation_Method				
Inoculation_Method	Num DF	Den DF	F Value	Pr > F
Dry	4	76	9.04	<.0001
Wet	4	76	14.99	<.0001

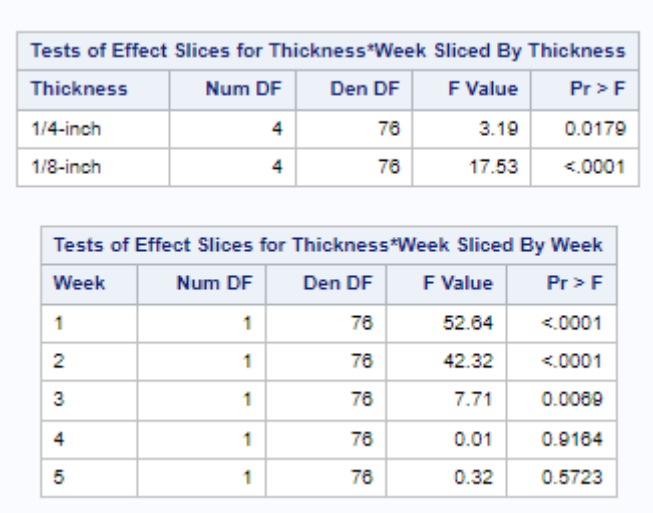
Tests of Effect Slices for Inoculation_Met*Week Sliced By Week				
Week	Num DF	Den DF	F Value	Pr > F
1	1	76	0.19	0.6649
2	1	76	7.28	0.0086
3	1	76	3.71	0.0579
4	1	76	1.08	0.3026
5	1	76	53.91	<.0001

Figure 12: *Test of Effect Slices* table for the two-way interaction with respect to the inoculation method and the week respectively.

Thickness*Week Least Squares Means									
Thickness	Week	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
1/4-inch	1	4.2195	0.1013	76	41.64	<.0001	0.05	4.0177	4.4214
1/4-inch	2	4.3300	0.1013	76	42.73	<.0001	0.05	4.1282	4.5318
1/4-inch	3	4.3601	0.1013	76	43.03	<.0001	0.05	4.1583	4.5619
1/4-inch	4	4.3315	0.1013	76	42.75	<.0001	0.05	4.1297	4.5333
1/4-inch	5	4.4721	0.1013	76	44.13	<.0001	0.05	4.2702	4.6739
1/8-inch	1	4.7491	0.1013	76	46.87	<.0001	0.05	4.5473	4.9510
1/8-inch	2	4.8049	0.1013	76	47.42	<.0001	0.05	4.6030	5.0067
1/8-inch	3	4.5628	0.1013	76	45.03	<.0001	0.05	4.3610	4.7646
1/8-inch	4	4.3238	0.1013	76	42.67	<.0001	0.05	4.1220	4.5256
1/8-inch	5	4.5135	0.1013	76	44.54	<.0001	0.05	4.3116	4.7153

Figure 13: *Least Squares Means* table for thickness and week interaction.

involving combinations that hold one of the two variables constant (eg. 1/4-inch at week 1 versus 1/8-inch at week 1 where the week is constant) are shown in Figure 13. The second table in Figure 13 shows us the interaction of these variables is most significant in the first two weeks (p-value<0.0001) followed by the third week (p-value=0.0069). When looking at these same cases in Figure 16 we can see negative estimates, implying the 1/4-inch thickness returns a lower response value for the first three weeks. Weeks 4 and 5 are not significant though. Additionally, in the first table we can see the interaction at the 1/8-inch level is highly significant (p-value<0.0001), while the 1/4-inch level has a higher p-value (0.0179) but is still significant.



Tests of Effect Slices for Thickness*Week Sliced By Thickness

Thickness	Num DF	Den DF	F Value	Pr > F
1/4-inch	4	76	3.19	0.0179
1/8-inch	4	76	17.53	<.0001

Tests of Effect Slices for Thickness*Week Sliced By Week

Week	Num DF	Den DF	F Value	Pr > F
1	1	76	52.64	<.0001
2	1	76	42.32	<.0001
3	1	76	7.71	0.0069
4	1	76	0.01	0.9164
5	1	76	0.32	0.5723

Figure 14: *Test of Effect Slices* table for the two-way interaction with respect to the thickness and the week respectively.

Recommendation

In terms of how the results of this experiment would affect a producer looking to sell beef jerky we need to make a key assumption. Since they would obviously not be inoculating it with salmonella, let's assume the dry and wet inoculation methods reflect storage conditions. With that in mind, our recommendation depends on a few factors.

If a producer will have both dry and wet storage methods, we suggest prioritizing cutting the jerky to 1/4-inch. This length was more consistent and had generally lower estimates than the 1/8-inch slices did. On the other hand, if a producer wants to use both thicknesses we recommend more wet storage conditions as the wet inoculation method saw lower estimates of bacterial growth over time. Of course, this is simulated data so these recommendations are not meant for real-world application.

Conclusion

This report highlights the work done in the consultation of our STAT 802 clients. We first met with our group to learn more about the experiment they wished to perform. After consulting with others we were able to provide recommendations for the framework of the experiment, including a model. By completing a power analysis using data provided by the client we were able to verify the necessary requirements for the size of the experiment in terms. Then we carried out a simulation to create a dataset that mimics the results of a real experiment. We were then able to explore and analyze the data using R and SAS.

Acknowledge

We want to thank Dr. Reka Howard, Malith Premarathna, and other members of the SC3L staff for helping us with the project. We're also grateful to the people of STAT 802 group for sharing their study ideas and letting us work with them on a project in their area of expertise.

References

Bobbit, Zach. 2021. "How to Interpret Negative AIC Values." Statology. 2021. https://www.statology.org/negative-aic/.

Brown, Jessica, Taylor Bannon, Arie Havelaar, Chad Carr, Barbara B. Kowalcyk, and Jason M. Scheffler. 2024. "Inactivation of Salmonella Enterica, Escherichia Coli O157:H7, and Campylobacter Jejuni in a Restructured Beef Jerky Developed for Production in Ethiopia." *Meat and Muscle Biology* 8 (April): 1–12. https://doi.org/10.22175/mmb.16091.

Lipka, Alex, and Benjamin Tyner. 2004. "Repeated Measures Covariance Structure." Statistical Consulting Service - Purdue University. 2004. https://www.stat.purdue.edu/scs/help/method/Repeated%20measures%20analysis.pdf.

USDA. n.d. "Appendix e: Selecting an Appropriate Covariance Structure," n.d., E.1–8. https://www.ars.usda.gov/ARSUserFiles/80000000/StatisticsGroupWebinars/Appendix%20E%20-%20Selecting%20a%20Covariance%20Structure.pdf.

Appendix A - R Code

```
library(knitr)
library(dplyr)
library(ggplot2)
library(naniar)
library(reshape2)
library(GGally)
library(janitor)
library(emmeans)
library(MASS)
library(multcomp)
library(lme4)
library(nnet)
library(tidyr)
library(knitr)
library(kableExtra)
library(corrplot)

data <- read.csv("rptm_simulation.csv")

response_summary <- data %>%
  group_by(Inoculation_Method, Thickness, Week) %>%
  summarise(
    Mean_Response = mean(Response, na.rm = TRUE),
    SD_Response = sd(Response, na.rm = TRUE),
    Count = n()
  )

knitr::kable(
  response_summary,
  caption = "Summary of response variable across factors and weeks",
  digits = 3,
  col.names = c("Inoculation Method", "Thickness", "Week", "Mean Response",
                "SD Response", "Count"),
  format = "markdown"
)

ggplot(data, aes(x = Response)) +
  geom_histogram(binwidth = 0.1, color = "black", fill = "skyblue") +
  labs(
```

```
    x = "Response",
    y = "Frequency"
  ) +
  theme_minimal() +
  theme(text = element_text(size = 14))

ggplot(data, aes(sample = Response)) +
  stat_qq(color = "black") +
  stat_qq_line(color = "red", linetype = "dashed") +
  labs(
    x = "Theoretical Quantiles",
    y = "Sample Quantiles"
  ) +
  theme_minimal() +
  theme(text = element_text(size = 14))

ggplot(data, aes(x = Week, y = Response,
                 color = Inoculation_Method,
                 group = interaction(Inoculation_Method, Thickness))) +
  geom_line(stat = "summary", fun = mean, size = 1) +
  geom_point(stat = "summary", fun = mean, size = 2) +
  facet_wrap(~ Thickness) +
  labs(
    x = "Week",
    y = "Mean Response",
    color = "Inoculation Method"
  ) +
  theme_minimal() +
  theme(
    text = element_text(size = 14),
    legend.position = "bottom"
  )

wide_data <- pivot_wider(data, names_from = Week, values_from = Response)

time_data <- wide_data[ , -1]
time_data <- time_data[sapply(time_data, is.numeric)]

time_cor_matrix <- cor(time_data, use = "pairwise.complete.obs")

cor_df <- as.data.frame(time_cor_matrix)
cor_df <- cbind(Time = rownames(cor_df), cor_df)
```

```r
kable(
  cor_df,
  digits = 2,
  col.names = c("Time", colnames(time_cor_matrix)))


corrplot(time_cor_matrix, method = "circle", type = "upper",
         tl.col = "black", tl.srt = 45)

# Model comparison data
covstruct <- data.frame(
  Model = c("VC", "UN", "CS", "AR(1)", "ARH(1)", "ANTE(1)", "TOEP"),
  AIC = c(-12.25, -12.22, -11.6, -12.86, -12.56, -8.57, -7.34),
  AICC = c(-12.09, -3.58, -11.29, -12.55, -11.01, -5.39, -6.19),
  BIC = c(-13.03, -18.47, -12.77, -14.03, -15.30, -12.48, -9.68)
)

knitr::kable(
  covstruct,
  format = "markdown",
  align = "c"
)
```

Appendix B - SAS Code and Additional Output

Power Analysis

```
data rptm_means;
input Inoculation_Method $ Thickness $ @@;
do Week=1 to 5 by 1;
    input mu @@;
    output;
end;
datalines;
Dry 1/4 4.26 4.25 4.47 4.33 4.54
Dry 1/8 4.91 4.95 4.67 4.56 4.97
Wet 1/4 4.21 4.57 4.65 4.49 4.38
Wet 1/8 4.86 4.78 4.62 4.32 4.22
;

data rptm_design;
 set rptm_means;
 do Batches = 1 to 5; /* Creating 3 blocks (batches) */
  output;
 end;
run;

proc print data=rptm_design;
run;

/* Creating Model */

proc glimmix data=rptm_design;
    class Batches Inoculation_Method Thickness Week;
    model mu = Inoculation_Method|Thickness|Week;
    random intercept / subject=Batches;
    random Week / subject=Batches*Inoculation_Method*Thickness type=ar(1) residual;
    parms (.029)(0.017)(.028)/hold=1,2,3;
    /* Provide 3 parameters for variance components */
    lsmeans Inoculation_Method*Thickness*Week / slicediff=Week cl;
    /* Define main effect contrasts */
    contrast 'Dry vs Wet'
        Inoculation_Method 1 -1;
    contrast '1/4 vs 1/8 inches'
```

```
        Thickness 1 -1;

    /* Define interaction contrasts */
   contrast 'Dry vs Wet at 1/4 Inches'
        Inoculation_Method 1 -1 Inoculation_Method*Thickness 1 0 -1 0;
    contrast 'Dry vs Wet at 1/8 Inches'
        Inoculation_Method 1 -1 Inoculation_Method*Thickness 0 1 0 -1;
    contrast '1/4 vs 1/8 inches for Dry inoculation'
        Thickness 1 -1 Inoculation_Method*Thickness 1 -1 0 0;
    contrast '1/4 vs 1/8 inches for Wet inoculation'
        Thickness 1 -1 Inoculation_Method*Thickness 0 0 1 -1;

    ods output contrasts=f_contrast tests3=f_anova;
run;

/*Power*/
data power;
    set f_contrast f_anova;
    ncparm = numdf * fvalue;
    alpha = 0.05;
    fcrit = finv(1-alpha, numdf, dendf, 0);
    power = 1 - probf(fcrit, numdf, dendf, ncparm);
run;

proc print data=power;
run;
```

Simulation

```
/* Step 1: Define AR(1) Covariance Structure in PROC IML */
proc iml;
    n = 20;                    /* Number of subjects per treatment*/
    mean = {0 0 0 0 0};        /* Mean for each week */
    T = 5;                     /* Number of repeated measures (weeks) */
    rho = 0.2;                 /* AR(1) correlation parameter */
    sigma2 = {0.29 0.29 0.29 0.29 0.29};      /* Variance for each week */

    /* Construct AR(1) covariance matrix */
    cov = j(T, T, 0);
    do i = 1 to T;
```

```
        do j = 1 to T;
            cov[i, j] = sqrt(sigma2[i] * sigma2[j]) * rho**abs(i - j);
        end;
    end;

    /* Print covariance matrix */
    print "Covariance Matrix:", cov;

    /* Generate simulated data using the covariance matrix */
    call randseed(12349);      /* Set random seed */
    x = randnormal(n, mean, cov); /* Simulate AR(1) correlated data */
    cname = {"t1", "t2", "t3", "t4", "t5"};

    /* Print the simulated data matrix directly */
    print "Simulated Data Matrix (x):", x;
    /* Print Sample mean */
    samplemean = x[:,];
    print samplemean n;

    /* Create dataset from simulated data */
    create inputdatacb from x[colname=cname];
    append from x;
close inputdatacb;
quit;

/* Step 2: Display the Simulated Data as a SAS Table */
proc print data=inputdatacb label;
    title "Simulated Data with AR(1) Covariance Structure";
run;

/* Step 3: Define Treatment Structure and Random Effects */
data rptm_simulation;
    retain Subject 0;
    keep Inoculation_Method Thickness Week Batches Response;

    array weeks[5] t1-t5;

    /* Define mean values for each combination of factors and week */
    if _n_ = 1 then do;
        array mean_values[4,2,5] _temporary_ (
            /* Dry, 1/4 inch */
            4.26, 4.25, 4.47, 4.33, 4.54,
```

```
            /* Dry, 1/8 inch */
            4.91, 4.95, 4.67, 4.56, 4.97,
            /* Wet, 1/4 inch */
            4.21, 4.57, 4.65, 4.49, 4.38,
            /* Wet, 1/8 inch */
            4.86, 4.78, 4.62, 4.32, 4.22
        );
    end;

    /* Simulation parameters */
    sigma_batch = sqrt(0.029); /* Batch variance */
    sigma_resid = sqrt(0.017); /* Residual variance */

    /* Loop through each combination of factors */
    do Batches = 1 to 5; /* Number of batches */
        batch_effect = rand("Normal", 0, sigma_batch); /*Random batch effect*/

        do Inoculation_Method = "Dry", "Wet";
            do Thickness = "1/4-inch", "1/8-inch";
                Subject + 1;
                set inputdatacb;

                /* Generate response for each week with AR(1) structure */
                do Week = 1 to 5;
                    Mean_Value = mean_values[
                      (Inoculation_Method="Dry")*1+(Inoculation_Method="Wet")*2,
                      (Thickness="1/4-inch")*1 + (Thickness="1/8-inch")*2,
                        Week
                    ];
                    Response = Mean_Value + batch_effect + weeks[Week];
                    output;
                end;
            end;
        end;
    end;
run;

/* Step 4: Display the Simulated Data in a Structured Format */
proc print data=rptm_simulation label;
    title "Simulated Data for 2x2 Factorial Design with Repeated Measures";
run;
```

Analysis

```
proc glimmix data=work.import plots=residualpanel;
    class Batches Inoculation_Method Thickness Week;
    model Response = Inoculation_Method|Thickness|Week;
    random intercept / subject=Batches;
    random Week / subject=Batches*Inoculation_Method*Thickness type=ar(1) residual;
    lsmeans Inoculation_Method*Thickness / diff slicediff = week;
    lsmeans Inoculation_Method*Week / diff slicediff = Thickness;
    lsmeans Thickness*Week / diff slicediff = Inoculation_Method;
    lsmeans Inoculation_Method*Thickness / slice=week;
    lsmeans Inoculation_Method*Week / slice=Thickness;
    lsmeans Thickness*Week / slice=Inoculation_Method;

    /* Define main effect contrasts */
    contrast 'Dry vs Wet'
        Inoculation_Method 1 -1;
    contrast '1/4 vs 1/8 inches'
        Thickness 1 -1;

    /* Define interaction contrasts */
   contrast 'Dry vs Wet at 1/4 Inches'
        Inoculation_Method 1 -1 Inoculation_Method*Thickness 1 0 -1 0;
    contrast 'Dry vs Wet at 1/8 Inches'
        Inoculation_Method 1 -1 Inoculation_Method*Thickness 0 1 0 -1;
    contrast '1/4 vs 1/8 inches for Dry inoculation'
        Thickness 1 -1 Inoculation_Method*Thickness 1 -1 0 0;
    contrast '1/4 vs 1/8 inches for Wet inoculation'
        Thickness 1 -1 Inoculation_Method*Thickness 0 0 1 -1;

    ods output contrasts=f_contrast tests3=f_anova;
run;
```

Additional Output

Inoc	Week	Inoc_	Week___	Estimate	Std_Error	DF	t_Value	Pr…t
Dry	1	Dry	2	0.031250	0.08042	76	0.39	0.6987
Dry	1	Dry	3	0.109000	0.07130	76	1.53	0.1304
Dry	1	Dry	4	0.134700	0.07335	76	1.84	0.0702
Dry	1	Dry	5	-0.260500	0.07291	76	-3.57	6e-04
Dry	1	Wet	1	0.031740	0.07299	76	0.43	0.6649
Dry	1	Wet	2	-0.165700	0.07299	76	-2.27	0.026
Dry	1	Wet	3	-0.031500	0.07299	76	-0.43	0.6673
Dry	1	Wet	4	0.210400	0.07299	76	2.88	0.0051
Dry	1	Wet	5	0.275400	0.07299	76	3.77	3e-04
Dry	2	Dry	3	0.077780	0.08042	76	0.97	0.3365
Dry	2	Dry	4	0.103400	0.07130	76	1.45	0.151
Dry	2	Dry	5	-0.291800	0.07335	76	-3.98	2e-04
Dry	2	Wet	1	0.000498	0.07299	76	0.01	0.9946
Dry	2	Wet	2	-0.197000	0.07299	76	-2.70	0.0086
Dry	2	Wet	3	-0.062740	0.07299	76	-0.86	0.3927
Dry	2	Wet	4	0.179200	0.07299	76	2.46	0.0164
Dry	2	Wet	5	0.244200	0.07299	76	3.35	0.0013
Dry	3	Dry	4	0.025660	0.08042	76	0.32	0.7506
Dry	3	Dry	5	-0.369500	0.07130	76	-5.18	<.0001
Dry	3	Wet	1	-0.077280	0.07299	76	-1.06	0.293
Dry	3	Wet	2	-0.274800	0.07299	76	-3.76	3e-04
Dry	3	Wet	3	-0.140500	0.07299	76	-1.93	0.0579
Dry	3	Wet	4	0.101400	0.07299	76	1.39	0.1687
Dry	3	Wet	5	0.166400	0.07299	76	2.28	0.0254
Dry	4	Dry	5	-0.395200	0.08042	76	-4.91	<.0001
Dry	4	Wet	1	-0.102900	0.07299	76	-1.41	0.1625
Dry	4	Wet	2	-0.300400	0.07299	76	-4.12	<.0001
Dry	4	Wet	3	-0.166200	0.07299	76	-2.28	0.0256
Dry	4	Wet	4	0.075760	0.07299	76	1.04	0.3026
Dry	4	Wet	5	0.140700	0.07299	76	1.93	0.0576
Dry	5	Wet	1	0.292300	0.07299	76	4.00	1e-04
Dry	5	Wet	2	0.094800	0.07299	76	1.30	0.198
Dry	5	Wet	3	0.229000	0.07299	76	3.14	0.0024
Dry	5	Wet	4	0.471000	0.07299	76	6.45	<.0001
Dry	5	Wet	5	0.535900	0.07299	76	7.34	<.0001
Wet	1	Wet	2	-0.197500	0.08042	76	-2.46	0.0164
Wet	1	Wet	3	-0.063240	0.07130	76	-0.89	0.3779
Wet	1	Wet	4	0.178700	0.07335	76	2.44	0.0172
Wet	1	Wet	5	0.243700	0.07291	76	3.34	0.0013
Wet	2	Wet	3	0.134200	0.08042	76	1.67	0.0992
Wet	2	Wet	4	0.376200	0.07130	76	5.28	<.0001
Wet	2	Wet	5	0.441100	0.07335	76	6.01	<.0001
Wet	3	Wet	4	0.241900	0.08042	76	3.01	0.0036
Wet	3	Wet	5	0.306940	0.07130	76	4.30	<.0001
Wet	4	Wet	5	0.064970	0.08042	76	0.81	0.4217

Figure 15: *Differences of Inoculation Method * Week Least Squares Means* table.

Thickness	Week	Thickness_	Week_	Estimate	Std_Error	DF	t_Value	Pr…t
1/4-inch	1	1/4-inch	2	-0.110500	0.08042	76	-1.37	0.1735
1/4-inch	1	1/4-inch	3	-0.140600	0.07130	76	-1.97	0.0523
1/4-inch	1	1/4-inch	4	-0.111900	0.07335	76	-1.53	0.1311
1/4-inch	1	1/4-inch	5	-0.252500	0.07291	76	-3.46	9e-04
1/4-inch	1	1/8-inch	1	-0.529600	0.07299	76	-7.26	<.0001
1/4-inch	1	1/8-inch	2	-0.585300	0.07299	76	-8.02	<.0001
1/4-inch	1	1/8-inch	3	-0.343300	0.07299	76	-4.70	<.0001
1/4-inch	1	1/8-inch	4	-0.104300	0.07299	76	-1.43	0.1573
1/4-inch	1	1/8-inch	5	-0.293900	0.07299	76	-4.03	1e-04
1/4-inch	2	1/4-inch	3	-0.030060	0.08042	76	-0.37	0.7096
1/4-inch	2	1/4-inch	4	-0.001450	0.07130	76	-0.02	0.9838
1/4-inch	2	1/4-inch	5	-0.142000	0.07335	76	-1.94	0.0565
1/4-inch	2	1/8-inch	1	-0.419100	0.07299	76	-5.74	<.0001
1/4-inch	2	1/8-inch	2	-0.474800	0.07299	76	-6.51	<.0001
1/4-inch	2	1/8-inch	3	-0.232800	0.07299	76	-3.19	0.0021
1/4-inch	2	1/8-inch	4	0.006232	0.07299	76	0.09	0.9322
1/4-inch	2	1/8-inch	5	-0.183400	0.07299	76	-2.51	0.0141
1/4-inch	3	1/4-inch	4	0.028610	0.08042	76	0.36	0.723
1/4-inch	3	1/4-inch	5	-0.112000	0.07130	76	-1.57	0.1205
1/4-inch	3	1/8-inch	1	-0.389000	0.07299	76	-5.33	<.0001
1/4-inch	3	1/8-inch	2	-0.444800	0.07299	76	-6.09	<.0001
1/4-inch	3	1/8-inch	3	-0.202700	0.07299	76	-2.78	0.0069
1/4-inch	3	1/8-inch	4	0.036290	0.07299	76	0.50	0.6205
1/4-inch	3	1/8-inch	5	-0.153400	0.07299	76	-2.10	0.0389
1/4-inch	3	1/4-inch	5	-0.140600	0.08042	76	-1.75	0.0845
1/4-inch	4	1/8-inch	1	-0.417700	0.07299	76	-5.72	<.0001
1/4-inch	4	1/8-inch	2	-0.473400	0.07299	76	-6.49	<.0001
1/4-inch	4	1/8-inch	3	-0.231300	0.07299	76	-3.17	0.0022
1/4-inch	4	1/8-inch	4	0.007684	0.07299	76	0.11	0.9164
1/4-inch	4	1/8-inch	5	-0.182000	0.07299	76	-2.49	0.0148
1/4-inch	5	1/8-inch	1	-0.277100	0.07299	76	-3.80	3e-04
1/4-inch	5	1/8-inch	2	-0.332800	0.07299	76	-4.56	<.0001
1/4-inch	5	1/8-inch	3	-0.090740	0.07299	76	-1.24	0.2176
1/4-inch	5	1/8-inch	4	0.148300	0.07299	76	2.03	0.0457
1/4-inch	5	1/8-inch	5	-0.041400	0.07299	76	-0.57	0.5723
1/8-inch	1	1/8-inch	2	-0.055730	0.08042	76	-0.69	0.4904
1/8-inch	1	1/8-inch	3	0.186300	0.07130	76	2.61	0.0108
1/8-inch	1	1/8-inch	4	0.425300	0.07335	76	5.80	<.0001
1/8-inch	1	1/8-inch	5	0.235700	0.07291	76	3.23	0.0018
1/8-inch	2	1/8-inch	3	0.242100	0.08042	76	3.01	0.0035
1/8-inch	2	1/8-inch	4	0.481100	0.07130	76	6.75	<.0001
1/8-inch	2	1/8-inch	5	0.291400	0.07335	76	3.97	2e-04
1/8-inch	3	1/8-inch	4	0.239000	0.08042	76	2.97	0.004
1/8-inch	3	1/8-inch	5	0.049340	0.07130	76	0.69	0.4911
1/8-inch	4	1/8-inch	5	-0.189700	0.08042	76	-2.36	0.0209

Figure 16: *Differences of Thickness * Week Least Squares Means* table.