

Lab 7

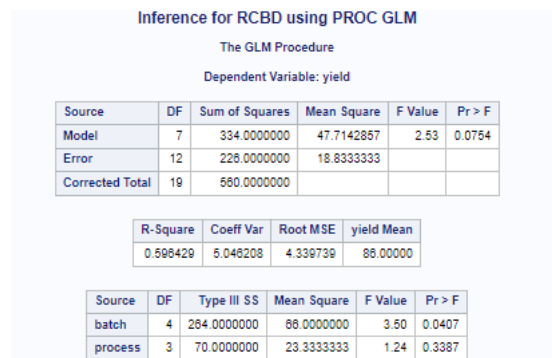
2024-10-22

Q1

The full SAS program to analyze the penicillin data is provided in the penicillin Lab7.sas file located in our course's shared folder in SAS Studio. Refer to the output (and modify the code where necessary) to complete the following exercises:

(a)

From the SAS output, find the full ANOVA table and provide an appropriate summary of results for analyzing the different processes on the yield of penicillin.



Inference for RCBD using PROC GLM

The GLM Procedure

Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	334.0000000	47.7142857	2.53	0.0754
Error	12	226.0000000	18.8333333		
Corrected Total	19	560.0000000			

R-Square	Coeff Var	Root MSE	yield Mean
0.598429	5.046208	4.339739	86.00000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
batch	4	264.0000000	66.0000000	3.50	0.0407
process	3	70.0000000	23.3333333	1.24	0.3387

Figure 1: Img

Overall F-test, with a p-value of 0.0754 we would consider having evidence at the $\alpha = 0.10$ level of rejecting the null hypothesis that the mean of the processes are equal averaged across all the batches and that the mean of the batches are equal averaged across all the processes for the response variable of penicillin yields (all combinations of the levels of the variables have the same mean penicillin yields).

With a p-value of 0.0407, we would reject the null hypothesis at the $\alpha = 0.05$ level that the mean of the batches are equal when averaged across all the processes; so we evidence of blocking effects on average penicillin yields across processes for the response variable of penicillin yields.

However, with a p-value of 0.3387 we do not reject the null hypothesis that the mean of the batches are equal averaged across all the processes for the response variable of penicillin yields.

(b)

Perform all pairwise comparisons of treatment means using Tukey's HSD method. Write a summary of your findings

Differences of Least Squares Means									
Effect	process	_process	Estimate	Standard Error	DF	t Value	Pr > t	Adjustment	Adj P
process	A	B	-1.0000	2.7447	12	-0.36	0.7219	Tukey-Kramer	0.9827
process	A	C	-5.0000	2.7447	12	-1.82	0.0935	Tukey-Kramer	0.3105
process	A	D	-2.0000	2.7447	12	-0.73	0.4802	Tukey-Kramer	0.8838
process	B	C	-4.0000	2.7447	12	-1.46	0.1707	Tukey-Kramer	0.4905
process	B	D	-1.0000	2.7447	12	-0.36	0.7219	Tukey-Kramer	0.9827
process	C	D	3.0000	2.7447	12	1.09	0.2958	Tukey-Kramer	0.7002

Figure 2: Img

All pairwise comparisons of treatment (process) have relatively large p-values, such that we do not reject the null hypothesis that any of the processes are on average different from any other average process used, with reference to their average penicillin yields. This is consistent with the previous results where we found overall the impact of process is not significantly different from zero, i.e. we showed that process overall does not appear to have an average effect on penicillin yield.

(c)

There are three orthogonal contrasts specified in the SAS code. Describe the analysis provided by these contrasts and determine which are statistically significant. Write a summary of your findings.

Inference for RCBD using PROC GLM

The GLM Procedure

Dependent Variable: yield

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
A-B	1	2.50000000	2.50000000	0.13	0.7219
C-(A+B)/2	1	87.50000000	87.50000000	3.58	0.0827
D-(A+B+C)/3	1	0.00000000	0.00000000	0.00	1.0000

Parameter	Estimate	Standard Error	t Value	Pr > t
A-B	-1.00000000	2.74489185	-0.38	0.7219
C-(A+B)/2	4.50000000	2.37897288	1.89	0.0827
D-(A+B+C)/3	0.00000000	2.24103151	0.00	1.0000

Figure 3: Img

The contrasts generally are: (1) whether the average yield of penicillin for process A is different from the mean amount of yield of penicillin for process B, (2) that the average yield of penicillin for process C is different from the average of the mean amount of penicillin yield of processes A and B, and (3) that the average yield of penicillin for process D is different from the average of the mean amount of penicillin yield of processes A, B, and C.

- (1) A-B: With a p-value of 0.7219, we do not have evidence to reject the null hypothesis that the average yield of penicillin for process A is not different from the mean amount of yield of penicillin for process B.
- (2) C - (A+B)/2: This contrast is whether process C is different from the average of processes A and B. With a p-value of 0.0827, we have evidence at the $\alpha = 0.10$ level to reject the null hypothesis the average yield of penicillin for process C is not different from the average of the mean amount of penicillin yield of processes A and B, which is evidence in favor of the null hypothesis that the average yield of penicillin for process C is different from the average of the mean amount of penicillin yield of processes A and B.
- (3) D - (A+B+C)/3: This contrast is whether the average yield of penicillin for process D is different from the average of the mean amount of penicillin yield of processes A, B, and C. With a p-value of 1.00, we do not have evidence to reject the null hypothesis that the average yield of penicillin for process D is not different from the average of the mean amount of penicillin yield of processes A, B, and C.

(d)

Check the assumptions for the RCBD analysis using SAS output. Write a summary of your findings.

There are 4 assumptions in the RCBD: 1. Independence of residuals: assessed through the study design principles

2. Homogeneous (equal) residual variance: assessed through boxplots of residuals within each treatment, ratio of standard deviations, and equal variance statistical tests such as the Brown Forsythe
3. Normality of residuals: assessed through histograms, normal Q-Q plots, summary statistics (mean, median, skewness, excess kurtosis), and tests for normality of the residuals
4. Additive block and treatment effects: Using the penicillin example as illustration, the additivity assumption for the RCBD can be checked in SAS using both graphs and statistical tests • Residual vs. Fitted Graph (no replication required): The easiest method for diagnosing this new assumption is to plot the fitted values on the x-axis and the residuals on the y-axis and then examine this plot for any trends (trends indicate an interaction between blocks and treatments, whereas random scatter indicates additive effects).

Overall: We do not have evidence to suspect that our key assumptions for the analysis are being violated. However, the assumption of additivity is somewhat suspect, as there may be some evidence of there being interaction effects, though this evidence is limited.

Independence: We do not have reason to believe this assumption is violated, as 5 random batches were selected for each process. Though not explicitly stated, we may reasonably infer that batches were not reused/recycled or otherwise “repeated” in this process; if this were violated however, our independence assumption would likely be violated.

```
sds <- c(5.1231, 5.9582, 6.2849, 5.5227)
max(sds)/min(sds)
```

```
## [1] 1.226777
```

Homogenous variances: The ratio of standard deviations by process is 1.2267 (see above code and below outputs), which is not especially problematic for our equal variances assumption. Turning to the residuals by group (process), we observe the spread of residuals is similar across groups, with a minor exception being the residuals for process “A” being relatively more narrow compared to groups “B”, “C”, and “D”. Finally, turning to statistical tests we utilize the Brown-Forsythe test and do not reject the null hypothesis of homogenous variances between treatments (p-value of 0.9388). Use of Leven’s Test provides consistent conclusions as well (not rejecting the null hypothesis of equal variances, albeit with p-value of 0.7620).

Normality: The QQ plot appears generally normal (closely aligns with the reference line), with a minor exception being the tails of the distribution. Via Shapiro-Wilk, we do not reject the null hypothesis of normally distributed (via a p-value of 0.95). Looking at the summary statistics: While Mean is not equal to median, it is close (0 vs. -0.5 resp.); additionally skewness and excess kurtosis are close to zero (their magnitude is no greater than 1). Taken together, there is nothing glaringly obvious that something is being violated for this assumption.

Block and Treatment effects: To test this we primarily use visual inspection of the fitted vs. residual plot. I want to be completely honest: I believe this could go either way and I’d recommend we actually estimate whether there are interactions in order to be more certain.

My gut instinct on the visual inspection is that there is a bit of a trend, as we have a more narrow spread in larger fitted values (around 90-95), though overall there appears to be a random scatter of points. If we see no trend then we would say we likely have additive effects, but as I believe there is a slight trend I would

argue there is some (though sparing) evidence that there is an interaction between blocks and treatments. The visual evidence is not especially striking.

We also can see from the “Interaction Plot” that the lines do not cross, i.e. there is a consistent pattern across batches. The interpretation of this is: As the lines do not cross, we do not have evidence that additivity is being violated, further reinforcing the above interpretation.

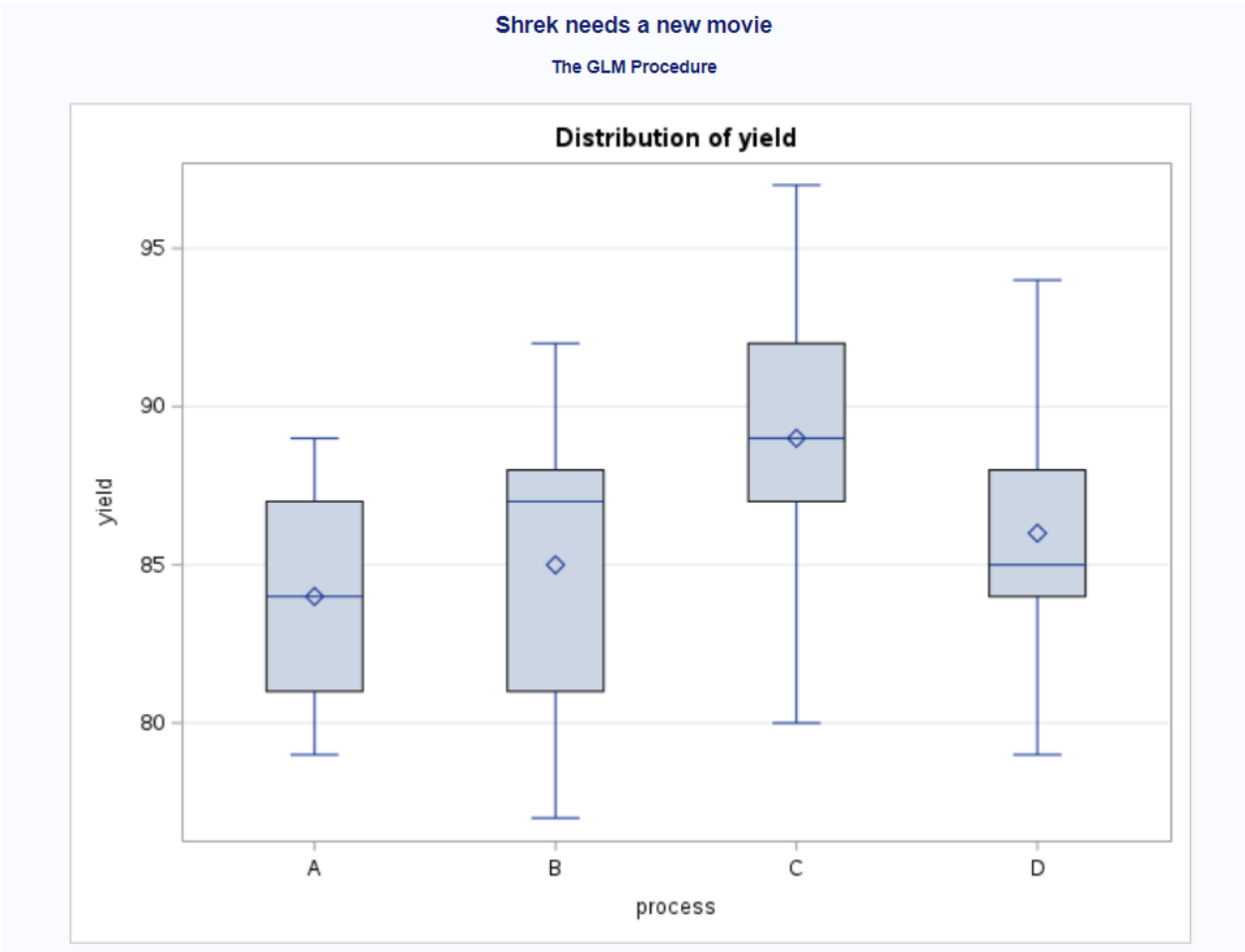


Figure 4: Img

Shrek needs a new movie

The GLM Procedure

Levene's Test for Homogeneity of yield Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
process	3	922.2	307.4	0.39	0.7620
Error	16	12618.8	788.7		

Figure 5: Img

Cars is the Perfect Trilogy					
The GLM Procedure					
Brown and Forsythe's Test for Homogeneity of yield Variance ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
process	3	4.9500	1.6500	0.13	0.9388
Error	18	198.0	12.3750		

Figure 6: Img

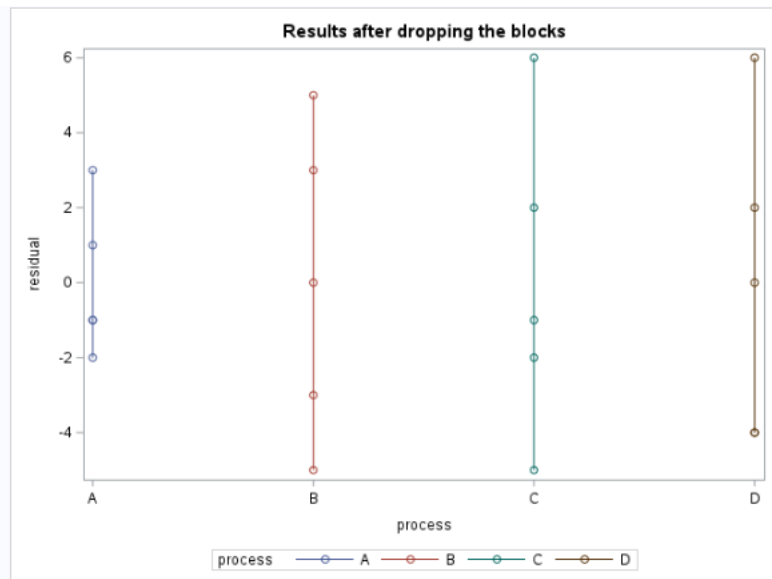


Figure 7: Img

Level of process	N	yield	
		Mean	Std Dev
A	5	84.0000000	4.12310583
B	5	85.0000000	5.95818784
C	5	89.0000000	6.28490284
D	5	88.0000000	5.52288051

Figure 8: Img

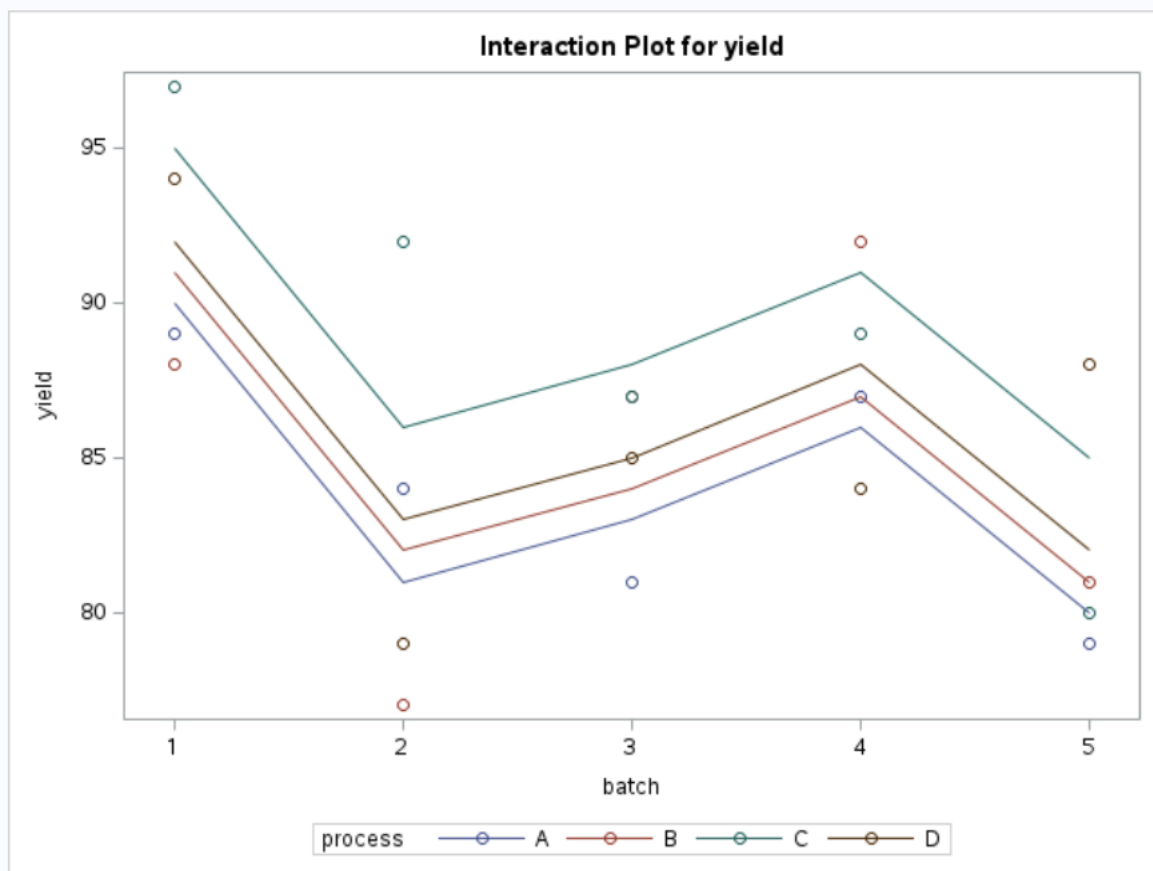


Figure 9: Img

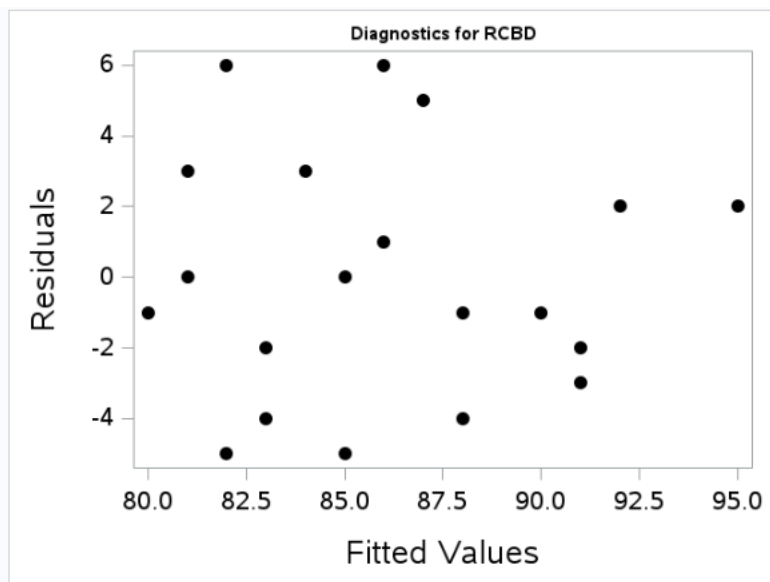


Figure 10: Img

Moments			
N	20	Sum Weights	20
Mean	0	Sum Observations	0
Std Deviation	3.44887472	Variance	11.8947388
Skewness	0.29080495	Kurtosis	-0.8382426
Uncorrected SS	226	Corrected SS	226
Coeff Variation	.	Std Error Mean	0.77119183

Basic Statistical Measures			
Location		Variability	
Mean	0.00000	Std Deviation	3.44887
Median	-0.50000	Variance	11.89474
Mode	-1.00000	Range	11.00000
		Interquartile Range	5.00000

Figure 11: Img

Tests for Normality				
Test	Statistic	p Value		
Shapiro-Wilk	W	0.950472	Pr < W	0.3743
Kolmogorov-Smirnov	D	0.114073	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.03563	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.275682	Pr > A-Sq	>0.2500

Figure 12: Img

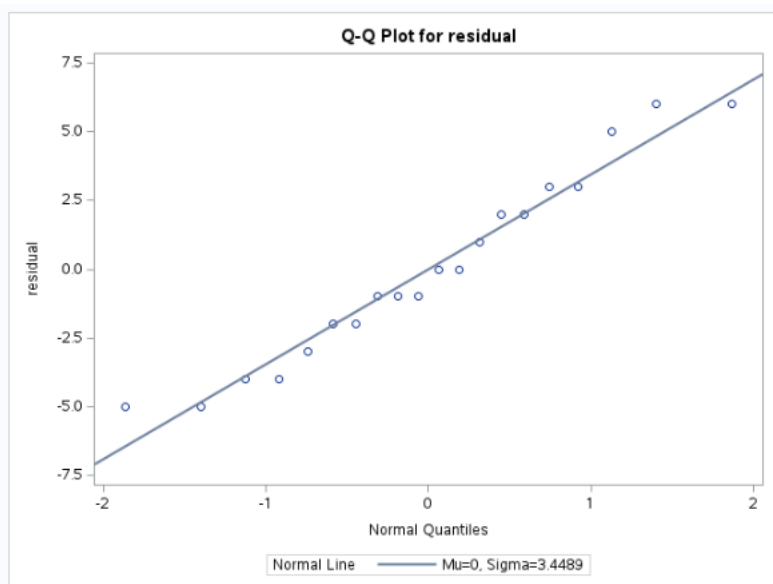


Figure 13: Img

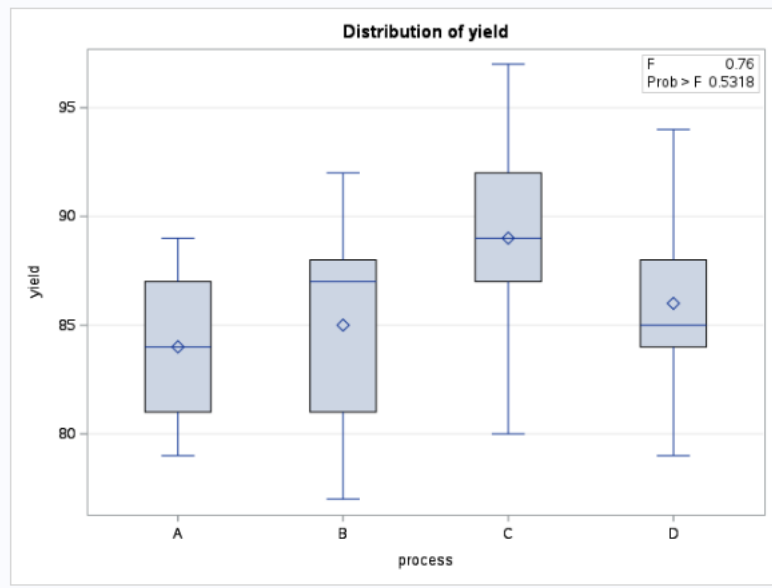


Figure 14: Img

(e)

Did blocking help? Compare the RCBD design with a design dropping the block effect and summarize your findings. Include a discussion of efficiency

Inference for RCBD using PROC GLM
The GLM Procedure
Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	334.0000000	47.7142857	2.53	0.0754
Error	12	228.0000000	18.8333333		
Corrected Total	19	560.0000000			

R-Square	Coeff Var	Root MSE	yield Mean
0.598429	5.048208	4.336739	86.00000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
batch	4	284.0000000	66.0000000	3.50	0.0407
process	3	70.0000000	23.3333333	1.24	0.3387

Figure 15: Img

Blocking (inclusion of the “batch” variable) significantly improves the efficiency of the model, as evidenced by the relatively higher R-squared (0.5984 vs. 0.125) and a lower Root MSE (4.33 vs. 5.53). Furthermore, we saw that the variable for block effect, “batch”, is statistically significant at the $\alpha = 0.05$ level, such that we had evidence to reject the null hypothesis that the blocking had no impact on the average penicillin yields. We additionally see that our Model Sum of Squares is higher when we include “batch” and Model MSE (not Error or Root) is also greater when we incorporate the blocking factor in our model, which indicates that this model is able to better explain the variability in our data, i.e. is more precise and efficient.

This, despite the “process” factor not not being significant (we do not reject the null hypothesis of mean levels of penicillin yields for a level of “process” averaged across batches having no differences) in either model. With such reasoning, primarily the notes of the prior paragraph, we find the initial RCBD design with blocks is more effective than the model without blocks.

On the topic of efficiency gains: The model with blocking is more efficient. By including blocking we

Results after dropping the blocks					
The GLM Procedure					
Dependent Variable: yield					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	70.0000000	23.3333333	0.78	0.5318
Error	16	490.0000000	30.6250000		
Corrected Total	19	560.0000000			

R-Square	Coeff Var	Root MSE	yield Mean
0.125000	6.434867	5.533986	86.00000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
process	3	70.0000000	23.3333333	0.78	0.5318

Figure 16: Img

effectively reduce the Sum of Squares error, which in turn indicates that the estimates being produced are more efficient (precise). This is to be expected though, to a certain extent. During lectures we made note that the RCBD design is more efficient than CRD, for a fixed number of units/observations (in one example we required 1.48 times the number of units/observations of the RCBD study in order to have similar/equal efficiency).

Efficiency calculation:

$$Efficiency = \sigma_{CRD}^2 / \sigma_{RCBD}^2 = 30.625 / 18.833 = 1.626135$$

This implies that the CRD design is less efficient than the RCBD design as this ratio is greater than 1.

Q2

The full SAS program to analyze the brome data is provided in the brome Lab7.sas file located in our course's shared folder in SAS Studio. Refer to the output (and modify the code where necessary) to complete the following exercises:

(a)

Find the full ANOVA table and provide an appropriate summary of results for analyzing the different management plans on the yield of brome.

Analysis of Latin Square
The GLM Procedure
Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	2387.749175	265.305484	109.11	<.0001
Error	6	14.589400	2.431587		
Corrected Total	15	2402.338575			

R-Square	Coeff Var	Root MSE	yield Mean
0.993927	3.142342	1.559348	49.62375

Source	DF	Type I SS	Mean Square	F Value	Pr > F
freeway	3	3.015725	1.005242	0.41	0.7497
stream	3	108.960825	36.320275	14.94	0.0034
trt	3	2275.772825	758.590875	311.98	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
freeway	3	3.015725	1.005242	0.41	0.7497
stream	3	108.960825	36.320275	14.94	0.0034
trt	3	2275.772825	758.590875	311.98	<.0001

Figure 17: Img

Overall F-test, with a p-value of <0.0001 we would consider having evidence at the $\alpha = 0.001$ level of rejecting the null hypothesis that following overall holds: The mean of the brome levels for freeway are equal averaged across all the streams and treatments, the mean of the brome levels for streams are equal averaged across all the freeways and treatments, and the the mean of the brome levels for treatments are equal averaged across all the streams and freeways. Said otherwise, there is evidence that there exists some difference in the average effect of freeway/stream/treatment level when averaged across the other respective variables, e.g. there is a difference in average brome levels across “freeway” when averaged across “stream” and “treatment” levels.

For specific variables of interest:

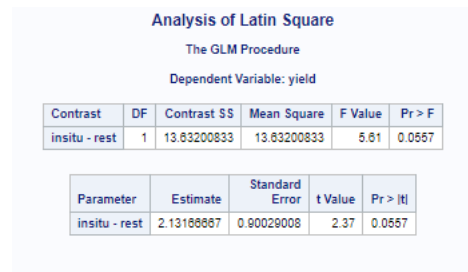
With a p-value of 0.0034, we would reject the null hypothesis at the $\alpha = 0.01$ level that the mean of brome levels across levels of stream are equal when averaged across all the freeways and treatments; so we evidence of stream blocking effects on average brome levels.

Similarly, with a p-value of <0.0001, we would reject the null hypothesis at the $\alpha = 0.01$ level that the mean of brome levels across levels of treatment are equal when averaged across all the freeways and streams; so we evidence of treatment blocking effects on average brome levels.

However, with a p-value of 0.7497 we do not reject the null hypothesis that mean brome levels across freeway levels are equal when averaged across all the streams and treatments.

(b)

Is there any difference between the management plan of “in situ” versus the other plans? Provide appropriate output from SAS and summarize your findings.



The figure displays two SAS output tables. The first table, titled 'Analysis of Latin Square', shows the results of a GLM procedure for the dependent variable 'yield'. It includes a contrast for 'insitu - rest' with 1 degree of freedom, a contrast sum of squares of 13.63200833, a mean square of 13.63200833, an F value of 5.61, and a p-value of 0.0557. The second table shows the parameter estimate for the 'insitu - rest' contrast, with an estimate of 2.13166667, a standard error of 0.90029008, a t value of 2.37, and a p-value of 0.0557.

Analysis of Latin Square					
The GLM Procedure					
Dependent Variable: yield					
Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
insitu - rest	1	13.63200833	13.63200833	5.61	0.0557

Parameter	Estimate	Standard Error	t Value	Pr > t
insitu - rest	2.13166667	0.90029008	2.37	0.0557

Figure 18: Img

Directly comparing “in situ” to the rest of the plans, we do have evidence to reject the null hypothesis at the $\alpha = 0.10$ level (p-value of 0.0557) in favor of the alternative hypothesis that the rest of the plans on average have different brome yields compared to “in situ”. Specifically, we believe that average brome levels “in situ” are different than the average brome levels when averaged across the other levels of freeway, stream, and treatments.

(c)

Did the consideration of the stream block help? Compare the LS design with a design dropping the stream effect and summarize your findings. Include a discussion of efficiency.

RCBD Analysis after dropping stream
The GLM Procedure
Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2278.788350	379.798058	27.67	<.0001
Error	9	123.560225	13.727803		
Corrected Total	15	2402.338575			

R-Square	Coeff Var	Root MSE	yield Mean
0.948571	7.466395	3.705105	49.62375

Source	DF	Type I SS	Mean Square	F Value	Pr > F
freeway	3	3.015725	1.005242	0.07	0.9729
trt	3	2275.772825	758.590875	55.26	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
freeway	3	3.015725	1.005242	0.07	0.9729
trt	3	2275.772825	758.590875	55.26	<.0001

Figure 19: Img

Analysis of Latin Square
The GLM Procedure
Least Squares Means

trt	yield LSMEAN	Standard Error	Pr > t
harvest	62.5250000	0.7798741	<.0001
insitu	51.2225000	0.7798741	<.0001
uncut	30.2325000	0.7798741	<.0001
windrow	54.5150000	0.7798741	<.0001

Figure 20: Img

RCBD Analysis after dropping stream
The GLM Procedure
Least Squares Means

trt	yield LSMEAN	Standard Error	Pr > t
harvest	62.5250000	1.8525525	<.0001
insitu	51.2225000	1.8525525	<.0001
uncut	30.2325000	1.8525525	<.0001
windrow	54.5150000	1.8525525	<.0001

Figure 21: Img

As the “stream” variable is significant in the Latin Squares design (significant as in has a p-value of 0.0041 leading us to reject the null hypothesis of no difference between average brome levels across “stream” levels at the $\alpha = 0.01$ level, when averaged across the other blocking variables), we believe that dropping this from our model will decrease efficiency, increase the Sum of Squares Error (decrease Model Sum of Squares), and overall decrease the precision of our estimates.

On the other hand, in both models we do not reject the null hypothesis that average brome levels do not differ significantly across levels of “highway”, meaning that the continued inclusion of this variable will have very little (if any) impact on our overall model in terms of efficiency (though this is beside the point of the question at hand).

Taken together, we do believe, or at least have some evidence, to suggest that the inclusion of the “stream”

variable helps, insomuch as it provides greater efficiency, precision, and predictive power in our overall model. Specifically, when comparing overall model diagnostics such as (Model) MSE/R-Squared, we observe relatively higher R-Squared in our Latin Squares Model compared to the one when stream is dropped (0.9939 vs. 0.94857) and the opposite, or lower comparative, Root MSE (1.5590 vs. 3.7051) though there is some oddity in the above outputs when comparing Model MSE.

By considering the stream as a block, the model becomes more effective at estimating treatment effects, as evidenced by the much higher F value for treatment (“trt”) in the Latin Squares design (311.98 vs. 55.68 specifically for estimating “trt”). Furthermore, when directly comparing the Standard Errors of the different treatments, we see significantly lower SEs when utilizing “stream” (0.77967 vs. 1.85255).

Efficiency:

$$Efficiency = \sigma_{CRD}^2 / \sigma_{LS}^2 = 13.7278 / 2.43157 = 5.645653$$

This implies that the CRD design is less efficient than the LS design as this ratio is greater than 1.