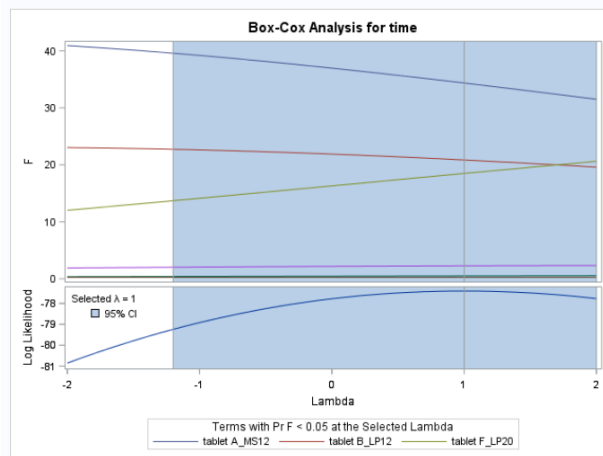


1. (a) Neither method suggests a transformation.

The Box Cox procedure on the right suggests a transformation of $\hat{\lambda} = 1$, which does not change the response.

The log standard deviation was regressed on the log mean. There was no evidence the slope differed from 0 indicating (although we do not accept the null hypothesis that the slope is 0) the transformation suggested is $1-0=1$, or again, no transformation.

```
PROC GLM DATA=yset;
  MODEL logstd=logmean / SS3 solution;
run;
```



Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	2.844531103	4.15082288	0.69	0.5236
logmean	-0.422577086	0.99524471	-0.42	0.6888

- (b) Let $|\bar{\epsilon}_j|$ = the absolute value of true mean error of tablets of type j for $j = 1, \dots, 6$

$$H_o: |\bar{\epsilon}_1| = \dots = |\bar{\epsilon}_6|$$

$$H_a: \text{At least one } |\bar{\epsilon}_j| \text{ differs from the rest for } j = 1, \dots, 6$$

Based on a p-value of 0.3874 there is no evidence against the equal variance assumption.

```
proc glm data=in1;
  class tablet;
  model time = tablet /ss3;
  means tablet / hovtest=levене(type=abs);
run;
```

Levene's Test for Homogeneity of time Variance ANOVA of Absolute Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
tablet	6	12.8540	2.1423	1.07	0.3874
Error	63	125.6	1.9941		

$$(c) y_{ij} = \mu + \tau_j + \epsilon_{ij}$$

y_{ij} = the i th reponse in the j th tablet type

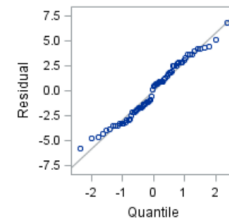
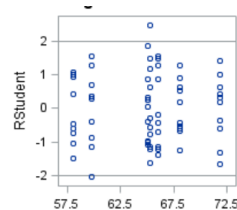
μ = the true mean disintegration time averaged over all tablet types in the study

τ_j = the effect of the j th tablet type on the true mean disintegration time

ϵ_{ij} = the true random error associated with disintegration time of the i th tablet of type j

$\epsilon_{ij} \sim N(0, \sigma^2)$ and are independent

Two diagnostic plots are below. The standardized residuals fall very close to the associated quantiles from the a standard normal distribution, indicating it is reasonable to assume the residuals follow a normal distribution. Trends are not shown in the residuals vs. fitted plot aside from the grouping by treatment and also because of the randomization, independence in the residuals is reasonable. Based on the test from 1b assessing HOV and on the residuals vs. fitted plot, HOV is reasonable. Although, looking at the scale, we do see a few standardized residuals quite far above the ± 2 general scope from the empirical rule.



Based on a p-value of <0.001 , there is strong evidence at least one of the tablets had a different true mean time to disintegration.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1311.970857	218.661810	23.95	<.0001
Error	63	575.180000	9.129841		
Corrected Total	69	1887.150857			

For the multiple comparisons procedure with G_TL16 as the control, each line of the MC procedure using Dunnett's method is testing:

H_o : there is no difference in the means between each of the other groups individually (each line is a different group) and the control group (G_TL16)

H_a : there is a difference in the means between the particular group we are on for that line and the control group

DATA IN11;

DO tablet='AZ_MS12', 'B_LP12', 'C_MS16', 'D_LP16', 'E_MS20', 'F_LP20', 'AG_TL_16';

```

DO rep = 1 TO 10;
  INPUT time @@; OUTPUT;
END; END;
LINES;
56.3 61.1 60.9 53.8 59.3 56.7 60.8 55.9 60.9 55.1
57.3 61.8 60.8 63.5 60.6 58.7 56.5 54.1 64.2 60.8
62.1 63.9 67.5 65.7 65.9 61.9 62.2 65.2 70.2 65.7
63.6 62.0 64.4 63.1 69.5 68.6 61.8 72.1 60.7 67.1
69.5 69.8 70.6 68.6 66.3 64.5 66.8 66.1 71.7 66.5
70.8 74.7 72.3 73.6 73.0 67.1 75.8 72.7 70.1 68.0
64.7 62.0 65.4 64.0 70.4 69.6 62.7 70.2 62.5 68.5
;
proc glm data=in11;
class tablet;
model time = tablet /ss3;
means tablet / DUNNETT alpha=0.05;
run;

```

Comparisons significant at the 0.05 level are indicated by ***.				
tablet Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
F_LP20 - AG_TL_1	5.810	2.245	9.375	***
E_MS20 - AG_TL_1	2.040	-1.525	5.605	
D_LP16 - AG_TL_1	-0.710	-4.275	2.855	
C_MS16 - AG_TL_1	-0.970	-4.535	2.595	
B_LP12 - AG_TL_1	-6.170	-9.735	-2.605	***
AZ_MS12 - AG_TL_1	-7.920	-11.485	-4.355	***

- (d) I did not transform the responses. Normality was reasonable based on the Normal Q-Q plot. The residuals vs. fitted plot showed slight non-constant variance, particularly variation in the D_LP16 group was slightly larger than the other groups. However, it was not larger by much and may have occurred by chance. The HOV assumption was reasonable based on Levene's test. Furthermore, the variation was not increasing as a function of the mean, so standard transformations such as those suggested by the Box Cox method would not fix a non-constant variance assumption. In fact, neither the Box Cox method nor the empirical method suggested a transformation of the response.

(e) **CONTRAST CODE AND OUTPUT (e-h)**

```

proc glm data=in1;
class tablet;
model time = tablet;
means / diff;
/* note n = 10 balanced */
contrast 'mean(ms) - mean(lp)' tablet 1/3 -1/3 1/3 -1/3 1/3 -1/3 0;

```

```
contrast 'mean(12) - mean(20)' tablet 1/2 1/2 0 0 -1/2 -1/2 0;
contrast 'linear ms' tablet -1 0 0 0 1 0 0;
contrast 'quad ms' tablet 1 0 -2 0 1 0 0;
contrast 'linear lp' tablet 0 -1 0 0 0 1 0;
contrast 'quad lp' tablet 0 1 0 -2 0 1 0;
run;
```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
mean(ms) - mean(lp)	1	55.680667	55.680667	6.10	0.0162
mean(12) - mean(20)	1	1203.409000	1203.409000	131.81	<.0001
linear ms	1	496.008000	496.008000	54.33	<.0001
quad ms	1	25.872667	25.872667	2.83	0.0972
linear lp	1	717.602000	717.602000	78.60	<.0001
quad lp	1	1.872667	1.872667	0.21	0.6522

(1.5pt) Perform a test for the contrast comparing the mean of the three disintegration time means for the magnesium stearate (MS) treatments to the mean of the three disintegration time means for the three liquid petrolatum (LP) treatments. Be sure to state the hypotheses in terms of the means (the μ 's) and include the associated p-value for the test.

$$H_0: \frac{1}{3}[\mu_{ams12} + \mu_{cms16} + \mu_{ems20}] = \frac{1}{3}[\mu_{blp12} + \mu_{dlp16} + \mu_{flp20}]$$

$$H_a: \frac{1}{3}[\mu_{ams12} + \mu_{cms16} + \mu_{ems20}] \neq \frac{1}{3}[\mu_{blp12} + \mu_{dlp16} + \mu_{flp20}]$$

Based on a p-value of <0.0162, there is strong evidence the mean disintegration time of the magnesium stearate treatments differed from the mean disintegration time of the liquid petrolatum treatment.

- (f) (1.5pt) The researcher wanted to check for a mesh size effect. To do this, she wants to perform a test for the contrast comparing the mean of the two disintegration time means for the 12 mesh granule size treatments to the mean of the two disintegration time means for the 20 mesh granule size treatments. Be sure to state the hypotheses in terms of the means (the μ 's) and include the associated p-value for the test.

$$H_0: \frac{1}{2}[\mu_{ams12} + \mu_{blp12}] = \frac{1}{2}[\mu_{ems20} + \mu_{fms20}]$$

$$H_a: \frac{1}{2}[\mu_{ams12} + \mu_{blp12}] \neq \frac{1}{2}[\mu_{ems20} + \mu_{fms20}]$$

Based on a p-value of <0.0001, there is strong evidence the mean disintegration time of the mesh size of 12 treatment differed from the mean disintegration time of the mesh size of 20 treatments.

- (g) (1.5pt) The researcher wanted to check for linear and quadratic trends for the means across the four granule sizes for the MS tablets. To do this, she wants to use two orthogonal contrasts. Be

sure to state the hypotheses in terms of the means (the μ 's) and include the associated p -values for the two contrast tests.

$$H_0: \mu_{ams12} + \mu_{ems20} = \mu_{cms16}$$

$$H_a: \mu_{ams12} + \mu_{ems20} \neq \mu_{cms16}$$

Based on a p -value of <0.0001 , there is strong evidence of a linear relationship between mean disintegration time and the mesh size of the magnesium searate treatment.

$$H_0: \frac{1}{2}[\mu_{ams12} + \mu_{ems20}] = \mu_{cms16}$$

$$H_a: \frac{1}{2}[\mu_{ams12} + \mu_{ems20}] \neq \mu_{cms16}$$

Based on a p -value of 0.0972, there is some evidence of a quadratic relationship between mean disintegration time and mesh size in the magnesium searate treatment.

- (h) (1.5pt) The researcher wanted to check for linear and quadratic trends for the means across the four granule sizes for the LP tablets. To do this, she wants to use two orthogonal contrasts. Be sure to state the hypotheses in terms of the means (the μ 's) and include the associated p -values for the two contrast tests.

$$H_0: \mu_{blp12} + \mu_{dlp20} = \mu_{flp16}$$

$$H_a: \mu_{blp12} + \mu_{dlp20} \neq \mu_{flp16}$$

Based on a p -value of <0.0001 , there is strong evidence of a linear relationship between the mean disintegration time and the mesh size in the liquid petrolatum treatment.

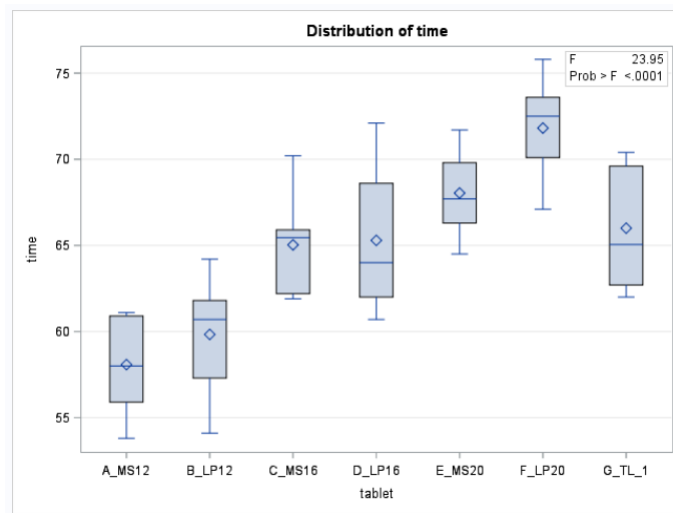
$$H_0: \frac{1}{2}[\mu_{blp12} + \mu_{dlp20}] = \mu_{flp16}$$

$$H_a: \frac{1}{2}[\mu_{blp12} + \mu_{dlp20}] \neq \mu_{flp16}$$

Based on a p -value of 0.6522, there is no evidence of a quadratic relationship between mean disintegration time and the mesh size in the liquid petrolatum treatment.

- (i) (1.5pt) Compare the results of the orthogonal contrasts in (g) and (h) to the side-by-side boxplots. Briefly describe why the contrast results are or are not consistent with the pattern of the boxplots.

DRAW LINES IN BOXPLOTS AND DISCUSS



2. PROBLEM 2 SAS CODE

```
proc glm data=in;
class glass subject;
model guess= subject glass;
random subject;
contrast 'mean(s) - mean(b)' glass 1 -1 -1 0 1 0;
contrast 'mean(others) - mean(tea)' glass 1 1 1 1 1 -5;
run;
```

REDO WITH BLOCKING VARIABLE FIRST

- (a) (6.5pt) Perform a complete analysis comparing the means of the student guesses. Be sure you include the model and state all hypotheses tested. Also include Bonferroni's multiple comparison test. You do not need to transform the response. Use $\alpha = .10$ for testing.

Let $i = 1, \dots, 6$ and $j = 1, \dots, 15$

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

Y_{ij} = the guess in ounces of student j for glass i

μ = the true mean guess in ounces averaged over all students and all glasses

τ_i = the true change in mean guess for glass i from the overall mean guess averaged over all students

$\beta_j \sim N(0, \sigma_b^2)$ = the random effect associated with student j , where students are assumed independent of one another and of the random error (ϵ_{ij})

$\epsilon_{ij} \sim N(0, \sigma^2)$ = the random error associated with guess i from student j , where each ϵ_{ij} is

assumed to be independent of all others and independent of the random student effect (β_j)

$$H_o: \tau_1 = \dots = \tau_6$$

H_a : at least one $\tau_i \neq \tau_{i'}$ among all $i \neq i'$ for i in 1,...,6

Based on a p-value of <0.0001 , after accounting for variation among students, there is strong evidence the mean guesses were different for at least two glasses.

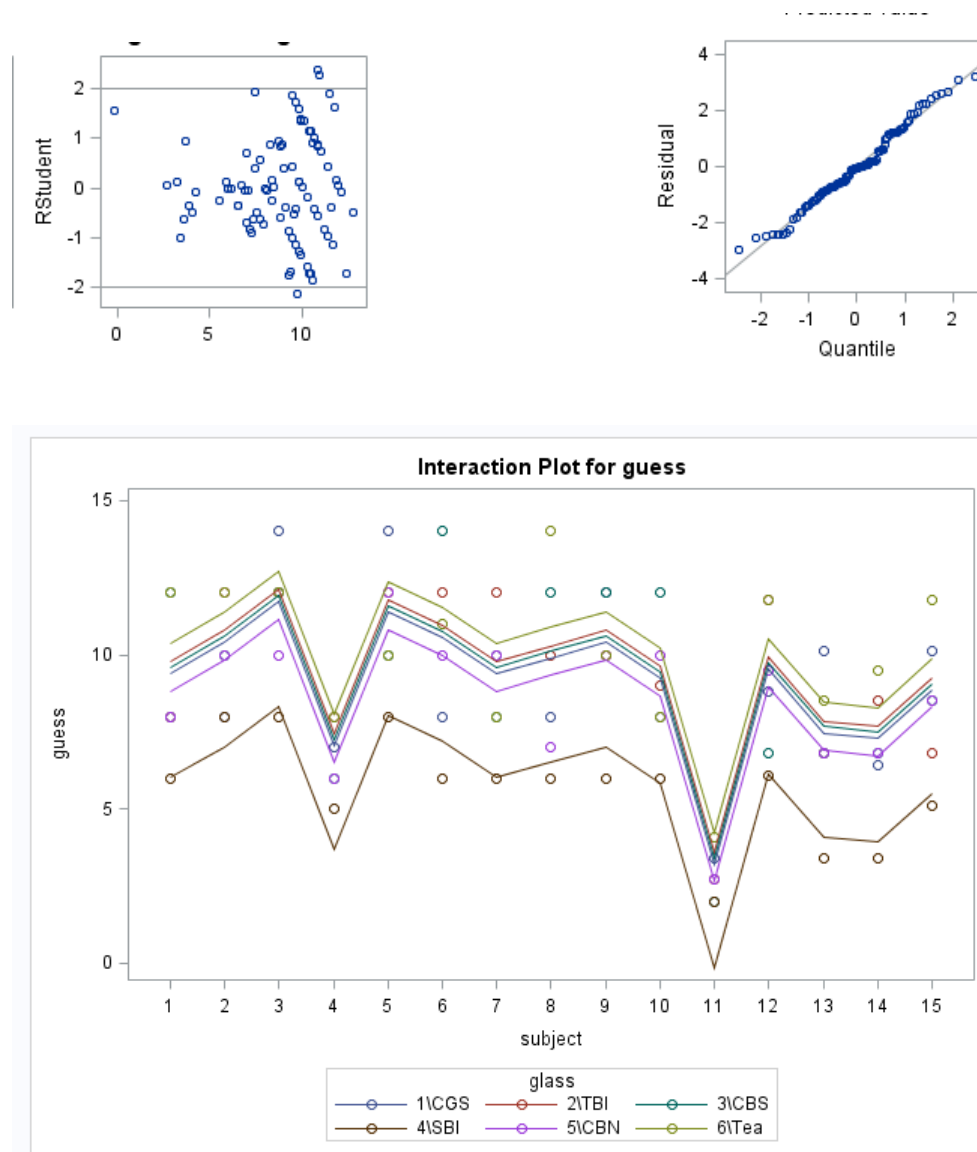
Bonferroni Comparison Lines for Least Squares Means of glass			
LS-means with the same letter are not significantly different.			
	guess LSMEAN	glass	LSMEAN Number
A	10.046667	6\Tea	6
A			
A	9.440000	2\TBI	2
A			
A	9.260000	3\CBS	3
A			
A	9.053333	1\CGS	1
A			
A	8.486667	5\CBN	5
B	5.666667	4\SBI	4

Based on the Bonferroni correction for multiple comparisons, at the family-wise $\alpha = 0.1$ significance threshold, there is strong evidence that the true mean guess for glass SBI differed from the true mean guess for all other glasses. There was no evidence of a difference in true mean guess between all other glasses at the 90% family-wise Bonferroni adjusted confidence level.

ASSESSING MODEL ASSUMPTIONS

The HOV assumption is likely violated as the studentized error residuals vs. fitted plot shows increasing variance with increasing mean. Normality in error residuals appears reasonable to assume as the residuals fall close to the line based on theoretical normal quantiles.

It is reasonable to assume the variation is the same within each subject.



Glass placement order was not randomized, so there may be a carry-over effect within each subject, making observations within each subject not independent. Unless students looked at each others' answers, it is reasonable to believe that students' guesses were independent. We also assume random errors in guesses do not depend on random student effects around the overall mean guess.

- (b) (1.5pt) Perform a test for the contrast comparing the mean of the two means for the stemmed glasses to the mean of the two means for the blue glasses. Be sure to include the p-value.

Based on a p-value of 0.0002, there is strong evidence of a difference in mean guess between the stemmed and blue glasses.

```
proc glm data=in;
class glass subject;
model guess= subject glass;
```



```

random subject;
contrast 'mean(s) - mean(b)' glass 1 -1 1 -1 0 0;
contrast 'mean(others) - mean(tea)' glass 1 1 1 1 1 -5;
run;

```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
mean(s) - mean(b)	1	38.56016667	38.56016667	15.13	0.0002
mean(others) - mean(tea)	1	34.66668889	34.66668889	13.60	0.0004

- (c) (1.5pt) Perform a test for the contrast comparing the tea mug mean to the mean of the five means for the other glasses. Be sure to include the p-value.

Based on a p-value of 0.0004, there is strong evidence of a difference in the mean tea guess and between the mean of guesses of all other glasses.

- (d) The contrasts are orthogonal because the sum of the coefficients is 0 in both cases.
- (e) (2pt) Suppose you are asked to compare how close the guesses were to the true glass volumes across the six glass treatments. Briefly describe how (if possible) you could address this request from the data provided. No analysis is required. If you do not believe this can be addressed, briefly give a reason why.

Yes, this can be addressed. The **null** hypotheses change to:

- 1) $\mu_{2TBI} = 9$
- 2) $\mu_{3CBS} = 9$
- 3) $\mu_{4TEA} = 9$
- 4) $\mu_{5SBI} = 9$
- 5) $\mu_{1CGS} = 10$
- 6) $\mu_{6CBN} = 11$

I would create confidence intervals and see whether the true value was in the confidence interval for each mean. We can adjust for multiple comparisons if that is necessary.

3. (2.5pt) Look at the attached SAS output. It contains two ANOVA tables. On the top is the ANOVA for a RCBD and on the bottom is the ANOVA for the RCBD data but with blocks removed from the model. At an $\alpha = .05$ level, you would draw different conclusions regarding the equality of treatment effects. Suppose you hear the following: "By systematically controlling outside effects, the residual mean square error will be reduced. This increases the precision of the experiment making detection of significant effects more likely." This example may appear to be inconsistent with this statement. Explain why you reject

the null hypothesis without blocks and fail to reject the null hypothesis with blocks (assuming $\alpha = .05$).

The sums of squares associated with the blocking variable do not explain enough variation in the model to “pay off” for the additional number of parameters that need to be estimated to include block in the model.

4. Use the sample means and an estimate of σ based on the oneway ANOVA to estimate the sample size needed for a CRD to achieve

Note the estimate of $\sigma = \text{sqrt}(46.9054167) = 6.8487529$ which is the RMSE from the one way ANOVA.

```
PROC GLMPower DATA = p4;
CLASS nitrogen;
MODEL yield=nitrogen;
POWER
STDDEV = 6.848753
ALPHA = 0.05 0.10
NTOTAL = .
POWER = 0.7 0.71 0.72 0.73 0.74 0.75 0.8 0.85 0.9 0.91 0.92 0.93 0.94 0.95;
TITLE 'Power Calc';
RUN;
```

Computed N Total					
Index	Alpha	Nominal Power	Error DF	Actual Power	N Total
1	0.05	0.70	66	0.833	72
2	0.05	0.71	66	0.833	72
3	0.05	0.72	66	0.833	72
4	0.05	0.73	66	0.833	72
5	0.05	0.74	66	0.833	72
6	0.05	0.75	66	0.833	72
7	0.05	0.80	66	0.833	72
8	0.05	0.85	90	0.938	96
9	0.05	0.90	90	0.938	96
10	0.05	0.91	90	0.938	96
11	0.05	0.92	90	0.938	96
12	0.05	0.93	90	0.938	96
13	0.05	0.94	114	0.980	120
14	0.05	0.95	114	0.980	120
15	0.10	0.70	42	0.736	48
16	0.10	0.71	42	0.736	48
17	0.10	0.72	42	0.736	48
18	0.10	0.73	42	0.736	48
19	0.10	0.74	66	0.904	72
20	0.10	0.75	66	0.904	72
21	0.10	0.80	66	0.904	72
22	0.10	0.85	66	0.904	72
23	0.10	0.90	66	0.904	72
24	0.10	0.91	90	0.970	96
25	0.10	0.92	90	0.970	96
26	0.10	0.93	90	0.970	96
27	0.10	0.94	90	0.970	96
28	0.10	0.95	90	0.970	96

- (a) A power of at least .95 for the ANOVA F-test assuming $\alpha = .05$.

N = 120 or 20 observations per nitrogen treatment with a balanced design will give a power of 0.98

- (b) A power of at least .90 for the ANOVA F-test assuming $\alpha = .01$.

N = 72 or 12 observations per nitrogen treatment with a balanced design will give a power of 0.94

- (c) What would be the estimated power for the ANOVA F-test assuming $n = 2$ replicates are to be taken for each treatment and assuming $\alpha = .05$?

Two replicates per group is not enough to detect a difference and so the power cannot be estimated.

```
PROC GLMPower DATA = p4;
CLASS nitrogen;
MODEL yield=nitrogen;
POWER
STDDEV = 6.848753
ALPHA = 0.05
NTOTAL = 18 24 36 48
POWER = .;
TITLE 'Power Calc';
RUN;
```

Computed Power					
Index	Nominal N Total	Actual N Total	Error DF	Power	Error
1	18	0	0	.	Invalid input
2	24	24	18	0.275	
3	36	24	18	0.275	
4	48	48	42	0.612	

5. (5pt) Stat 541 Students: Read Section 4.3 (pages 165-168) on Graeco-Latin squares. Consider the following experiment. A food processor wants to determine the effect of package design on the sale of one of her company's products. She had five designs to be tested: A, B, C, D, E. There were several sources of variation, however, whose possible effects were to be accounted for in the study. These included (1) day of the week, (2) differences among stores, and (3) effect of shelf height. She decided to use a Graeco-Latin square with five weekdays (M, Tu, W, Th, F,) for the row blocks, five stores for the column blocks, and five shelf heights for the Greek letter blocks. The experimental design and sales

in dollars are shown in the following table.

```
# create latin squares data set

rowsday <- c(rep(c("Mon", "Tues", "Wed", "Thurs", "Fri"),5))
colstore <- c(rep("I", 5), rep("II", 5), rep("III", 5),
              rep("IV", 5), rep("V", 5))
latin <- c("E", "D", "B", "C", "A", "C", "B", "E", "A", "D",
           "B", "A", "D", "E", "C", "A", "E", "C", "D", "B")

greek <- c("alpha", "delta", "epsilon", "beta", "gamma", "delta",
           "beta", "gamma", "epsilon", "alpha", "gamma", "alpha",
           "beta", "delta", "epsilon", "epsilon", "gamma", "delta",
           "alpha", "beta", "beta", "epsilon", "alpha", "gamma", "delta")

dollars <- c(238,149,222,187,65,228,220,295,66,118,158,92,104,242,
            279,188,169,54,122,278,74,282,213,90,176)

p5data <- data.frame(cbind(rowsday, colstore, latin, greek, dollars))

Warning in cbind(rowsday, colstore, latin, greek, dollars): number of rows of result is not a multiple of
vector length (arg 3)

dim(p5data)

[1] 25  5

aggregate(as.numeric(as.character(p5data$dollars)), list(as.factor(p5data$rowsday)), mean)

  Group.1      x
1     Fri 183.2
2     Mon 177.2
3    Thurs 141.4
4     Tues 182.4
5     Wed 177.6

#write.csv(p5data, file = "p5data.csv")

mon <- c(238,228,158,188,74)
tues <- c(149,220,92,169,282)
wed <- c(222,295,104,54,213)
thurs <- c(187,66,242,122,90)
fri <- c(65,118,279,278,176)
c(mean(mon), mean(tues), mean(wed), mean(thurs), mean(fri))

[1] 177.2 182.4 177.6 141.4 183.2
```

- (a) Perform an ANOVA. State the null and alternative hypotheses of interest, and your conclusions. You can use $\alpha = .05$ for testing.

$$Y_{ijkl} = \mu + \beta_i + \gamma_j + \psi_k + \tau_l + \epsilon_{ijkl}$$

Let τ_l represents the effect of design l which is an element of the set A,B,C,D,E.

$$H_o: \tau_A = \tau_B = \tau_C = \tau_D = \tau_E$$

H_a : at least one $\tau_l \neq \tau_{l'}$ for $l \neq l'$

Based on a p-value of 0.2828, there is no evidence of a design effect on the true mean price. At the $\alpha = 0.05$ cutoff, the design did not have an effect on the true mean price after accounting for the store, shelf, and weekday effects.

```
proc glm data=p5data;
title 'Graeco-Latin Square';
class rowsday colstore latin greek;
model newdollar = rowsday colstore latin greek;
means latin / tukey alpha=0.05;
run;
```

Source	DF	Type III SS	Mean Square	F Value	Pr > F
rowsday	4	12707.47106	3176.86777	0.36	0.8274
colstore	4	1544.96000	386.24000	0.04	0.9955
latin	4	53168.06987	13292.01747	1.53	0.2828
greek	4	26908.47700	6727.11925	0.77	0.5727

(b) *Include a Tukey multiple comparison procedure, and your conclusions from it.*

Using the 95% Tukey adjusted family-wise confidence level, there is no evidence the mean cost differs for any of the designs.

Means with the same letter are not significantly different.			
Tukey Grouping	Mean	N	latin
A	218.20	5	B
A			
A	203.60	5	E
A			
A	167.60	5	C
A			
A	155.00	5	D
A			
A	117.40	5	A

6. See attached.