

A.

```

1 DATA SUPPORT;
2 INFILE '/home/yeopdodo860/my_courses/tjp00/Roof_Support_stiffness_lengths.csv' delimiter=',' dsd;
3 INPUT Stiffness length ;
4 ;
5 PROC ANOVA DATA=SUPPORT;
6 CLASS length;
7 MODEL Stiffness = length;
8 MEANS length/Turkey;
9 RUN;
10
11

```

The ANOVA Procedure

Class Level Information				
Class	Levels	Values		
length	5	4	6	8 10 12

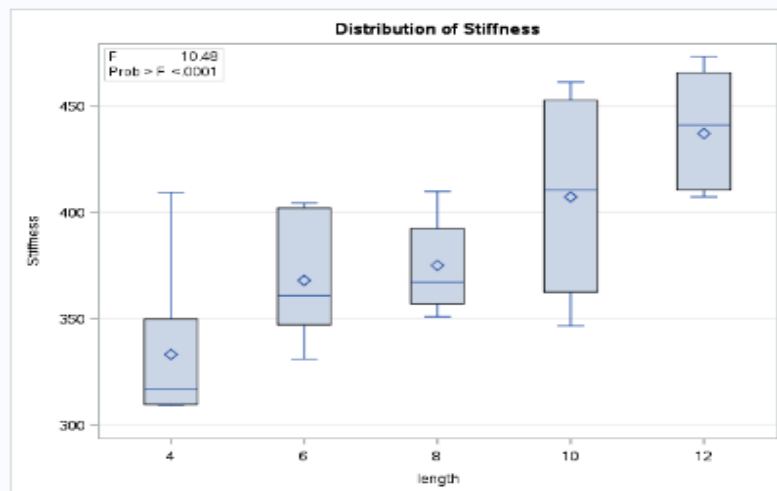
Number of Observations Read	36
Number of Observations Used	35

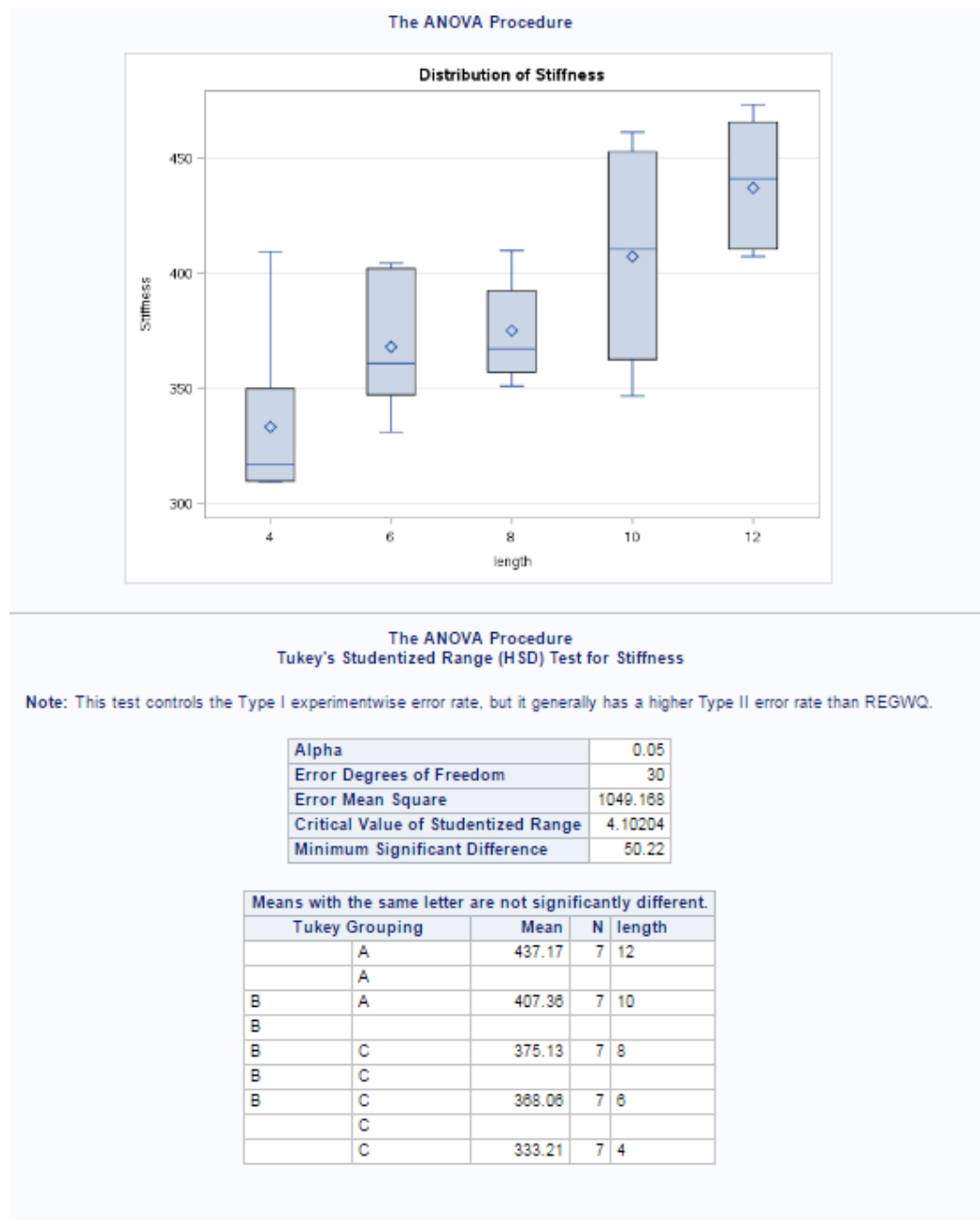
The ANOVA Procedure Dependent Variable: Stiffness

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	43992.55143	10998.13786	10.48	<.0001
Error	30	31475.03143	1049.16771		
Corrected Total	34	75467.58286			

R-Square	Coeff Var	Root MSE	Stiffness Mean
0.582933	8.431042	32.39086	384.1857

Source	DF	Anova SS	Mean Square	F Value	Pr > F
length	4	43992.55143	10998.13786	10.48	<.0001





B. Answer the following.

1. State the null and alternate hypotheses.

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D = \mu_E$$

H_1 : At least one population mean is different from the others

2. State the sample means for each of the independent variables.

A: 333.214286

B: 368.057143
C: 375.128571
D: 407.357143
E: 437.171429

3. State the value of the hypothesis test statistic found by SAS.

F value : 10.48

4. State the p -value found by SAS.

<.0001

5. State whether this p -value would indicate “reject the null hypothesis” or “fail to reject the null hypothesis”
at the given level of significance.

p value would indicate reject the null hypothesis since alpha is 0.01 which is bigger than p value

6. State a conclusion in words, in the context of the given situation, i.e. using the variable names.
II.

There is at least one different group in A, B, C, D, E groups from the length variables.

II. Multiple-Comparison ANOVA

```
1 DATA HORMONE;  
2 DO ROW = 1 TO 5;  
3 DO HORM= 1 TO 4;  
4 INPUT HPERIOD @@;  
5 OUTPUT;  
6 END;  
7 END;  
8 DATALINES;  
9 13 17 7 14  
10 21 13 20 17  
11 18 15 20 17  
12 7 11 18 10  
13 6 11 15 8  
14 ;  
15 RUN;  
16 PROC ANOVA DATA = HORMONE;  
17 CLASS ROW HORM;  
18 MODEL HPERIOD = ROW HORM;  
19 MEANS ROW HORM/ TUKEY;  
20 RUN;  
21  
22
```

The ANOVA Procedure

Class Level Information				
Class	Levels	Values		
ROW	5	1	2	3 4 5
HORM	4	1	2	3 4

Number of Observations Read	20
Number of Observations Used	20

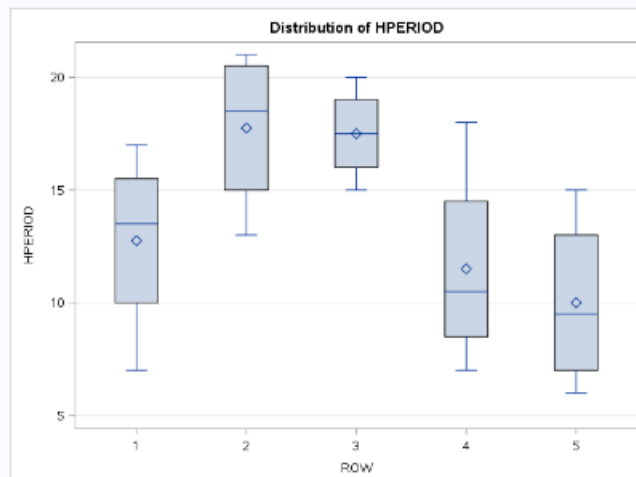
The ANOVA Procedure Dependent Variable: HPERIOD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	230.1000000	32.8714286	2.12	0.1202
Error	12	185.7000000	15.4750000		
Corrected Total	19	415.8000000			

R-Square	Coeff Var	Root MSE	HPERIOD Mean
0.553391	28.30092	3.933828	13.90000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
ROW	4	200.3000000	50.0750000	3.24	0.0510
HORM	3	29.8000000	9.9333333	0.64	0.6028

The ANOVA Procedure



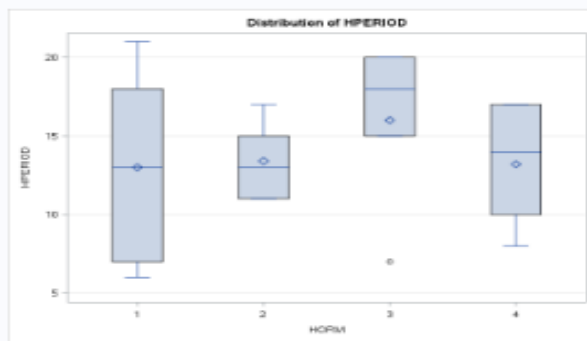
The ANOVA Procedure
Tukey's Studentized Range (HSD) Test for HPERIOD

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	15.475
Critical Value of Studentized Range	4.50788
Minimum Significant Difference	5.9952

Means with the same letter are not significantly different.			
Tukey Grouping	Mean	N	HCRM
A	17.750	4	2
A			
A	17.500	4	3
A			
A	12.750	4	1
A			
A	15.500	4	4
A			
A	10.000	4	5

The ANOVA Procedure



The ANOVA Procedure
Tukey's Studentized Range (HSD) Test for HPERIOD

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	15.475
Critical Value of Studentized Range	4.19851
Minimum Significant Difference	7.3953

Means with the same letter are not significantly different.			
Tukey Grouping	Mean	N	HCRM
A	16.000	5	3
A			
A	13.400	5	2
A			
A	13.200	5	4
A			
A	13.000	5	1

B. Answer the following.

1. State the null and alternate hypotheses.

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D = \mu_E$$

H_1 : At least one population mean is different from the others

2. State the sample means for each of the independent variables.

ROW 1: 17.750

ROW 2: 17.500

ROW 3: 12.750

ROW 4: 11.500

ROW 5: 10.000

3. State the value of the ANOVA hypothesis test statistic found by SAS.

F value: 2.12

4. State the ANOVA p -value found by SAS.

0.1202

5. State whether this ANOVA p -value would indicate “reject the null hypothesis” or “fail to reject the null hypothesis” at the given level of significance.

The p value would indicate fail to reject the null hypothesis since the alpha is $0.05 < 0.1202$

6. State a conclusion for the ANOVA in words, in the context of the given situation.

There is no group that has a significant difference with other groups

7. Interpret the Tukey method analysis to determine any significant differences.

a. Write the 5 sample means in increasing order.

10.0 11.5 12.75 17.5 17.75

b. Use the SAS grouping information to determine whether each pairwise comparison indicates a significant difference: Underline each pair of means if there is *not* a significant difference indicated. Do

not underline a pair if there is a significant difference indicated.

10.0 11.5 12.75 17.5 17.75

9. Answer the questions, “What happens when Tukey’s procedure is applied? Does the Tukey analysis indicate which of the means is different from the rest?”

The Turkey's procedure makes easy to visualize the differences between the mean values and separate them in multiple groups by the differences.

III. Single-Factor ANOVA with differing sample sizes

A.

```
1 DATA HORMONE;
2 DO ROW = 1 TO 5;
3 DO tomato = 1 TO 4;
4 INPUT salinity @@;
5 OUTPUT;
6 END;
7 END;
8 DATALINES;
9
10 59.5 55.2 51.7 44.6
11 53.3 59.1 48.8 48.5
12 56.8 52.8 53.9 41.0
13 63.1 54.5 49.0 47.3
14 58.7 46.1
15 ;
16 RUN;
17 PROC GLM DATA = HORMONE;
18 CLASS ROW tomato ;
19 MODEL salinity = ROW tomato ;
20 LSMEANS ROW tomato / ADJUST = TUKEY;
21 RUN;
22
23
```

The GLM Procedure

Class Level Information		
Class	Levels	Values
ROW	5	1 2 3 4 5
formto	4	1 2 3 4

Number of Observations Read	15
Number of Observations Used	15

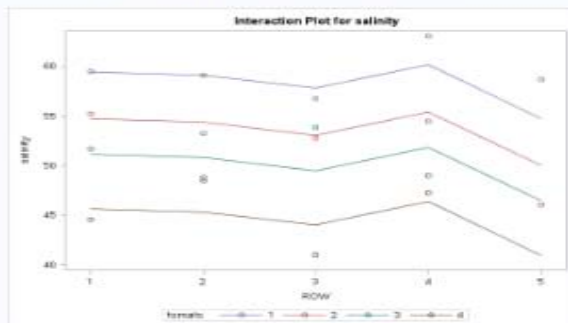
The GLM Procedure
Dependent Variable: salinity

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	433.1469125	433.1469125	4.07	0.0227
Error	13	135.8525750	10.4501977		
Corrected Total	14	569.0000000			

R-Square	Coef Var	Root MSE	salinity Mean
0.762355	7.409575	3.232775	52.43559

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ROW	4	11.5932775	2.8983194	0.19	0.9373
formto	3	418.5562500	139.5187500	9.25	0.0011

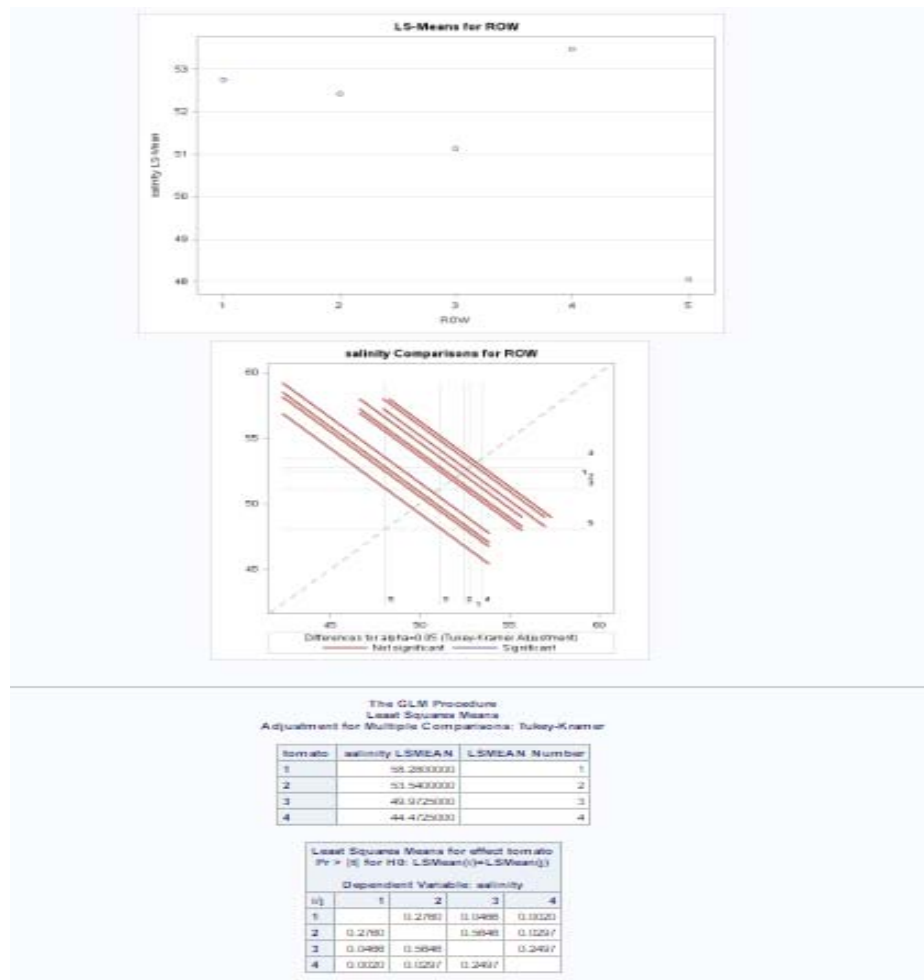
Source	DF	Type III SS	Mean Square	F Value	Pr > F
ROW	4	42.3575250	10.5893125	0.70	0.6079
formto	3	418.5562500	139.5187500	9.25	0.0011



The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

ROW	salinity LSMEAN	LSMEAN Number
1	52.7500000	1
2	52.4250000	2
3	51.1250000	3
4	51.4750000	4
5	48.0562500	5

Least Squares Means for effect ROW Pr > t for H0: LSMean(i)=LSMean(j)					
Dependent Variable: salinity					
i\j	1	2	3	4	5
1		0.9999	0.9733	0.9957	0.6748
2	0.9999		0.9852	0.9947	0.7281
3	0.9733	0.9852		0.9998	0.8552
4	0.9957	0.9947	0.9998		0.5575
5	0.6748	0.7281	0.8552	0.5575	



B. Answer the following.

1. State the null and alternate hypotheses.

$$H_0: \mu A = \mu B = \mu C = \mu D$$

H_1 : At least one population mean is different from the others

2. State the sample means for each of the independent variables.

58.2800000

53.5400000

49.9725000

44.4725000

3. State the value of the ANOVA hypothesis test statistic found by SAS.

9.25

4. State the ANOVA p -value found by SAS.

0.0031

5. State whether this ANOVA p -value would indicate “reject the null hypothesis” or “fail to reject the null hypothesis” at the given level of significance.

reject the null hypothesis since p value is $0.0031 < 0.05$

6. State a conclusion for the ANOVA in words, in the context of the given situation.

There is at least one group is significantly different from others.

7. Interpret the Tukey method analysis to determine any significant differences by writing the sample means in increasing order, and underlining (or not) as appropriate (as was done for part II).

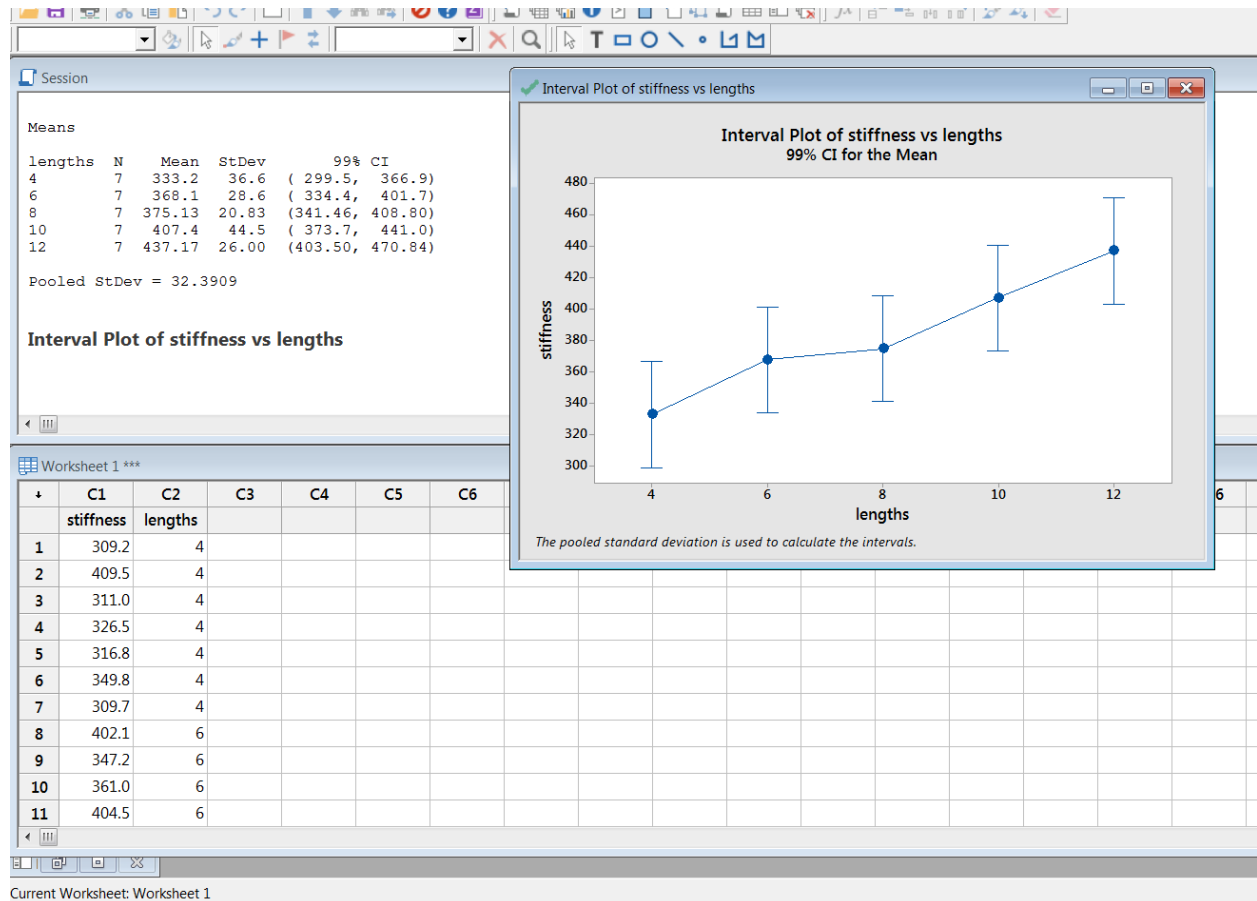
1	2	3	4
<hr/>			—

8. State an appropriate conclusion to Tukey’s procedure. Specifically, what do the overlapping significance connections imply?

The group 1,2,3 are significantly different from a group 4.

MINITAB

A



STAT>ANOVA>ONEWAY

One-way ANOVA: stiffness versus lengths

Method

Null hypothesis All means are equal
 Alternative hypothesis At least one mean is different
 Significance level $\alpha = 0.01$

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
lengths	5	4, 6, 8, 10, 12

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
lengths	4	43993	10998	10.48	0.000
Error	30	31475	1049		
Total	34	75468			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
32.3909	58.29%	52.73%	43.23%

Means

lengths	N	Mean	StDev	99% CI
4	7	333.2	36.6	(299.5, 366.9)
6	7	368.1	28.6	(334.4, 401.7)
8	7	375.13	20.83	(341.46, 408.80)
10	7	407.4	44.5	(373.7, 441.0)
12	7	437.17	26.00	(403.50, 470.84)

Pooled StDev = 32.3909

B. Answer the following.

1. State the null and alternate hypotheses.

Null hypothesis : All means are equal

Alternative hypothesis : At least one mean is different

2. State the sample means for each of the independent variables.

A: 333.2

B: 368.1

C: 375.13

D: 407.4

E: 437.17

3. State the value of the hypothesis test statistic found by SAS.

F value : 10.48

4. State the p -value found by SAS.

<.0000

5. State whether this p -value would indicate “reject the null hypothesis” or “fail to reject the null hypothesis” at the given level of significance.

p value would indicate reject the null hypothesis since alpha is 0.01 which is bigger than p value

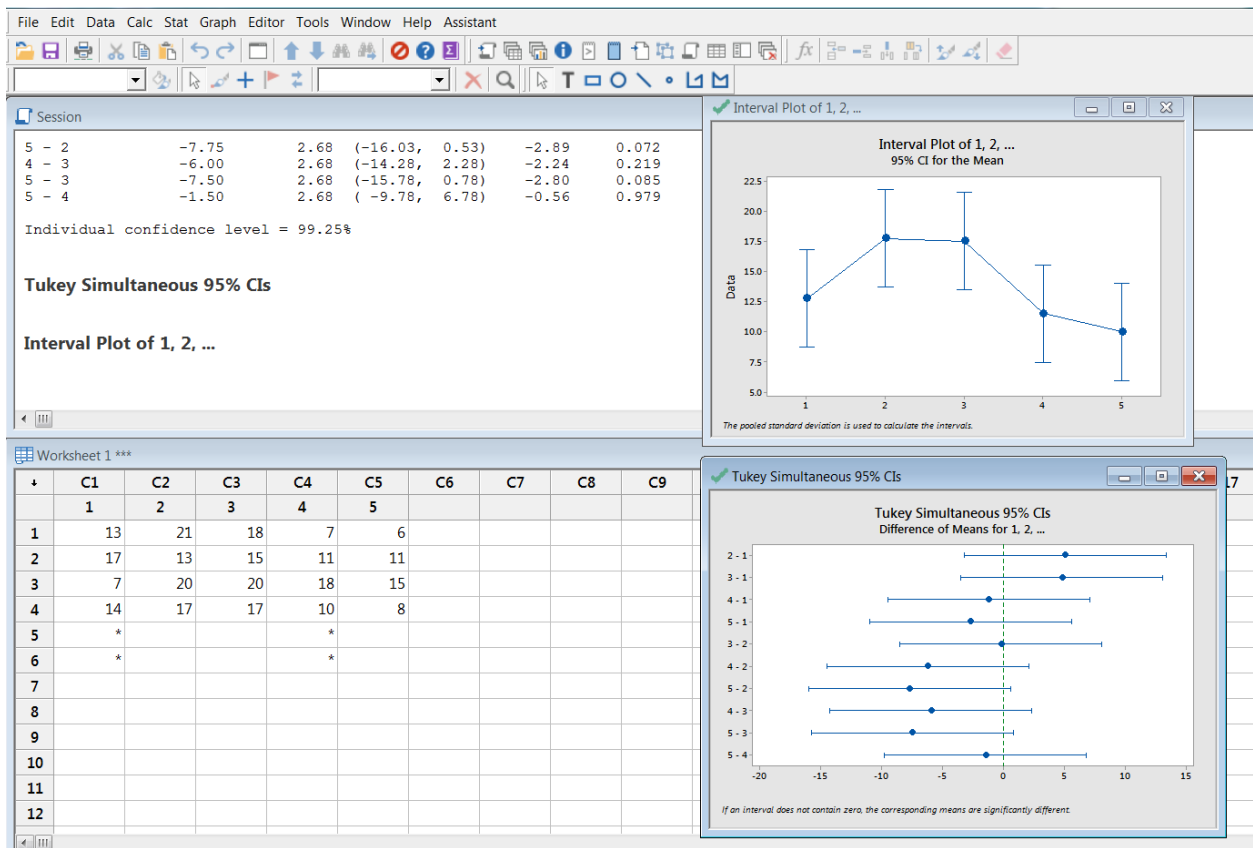
6. State a conclusion in words, in the context of the given situation, i.e. using the variable names.

II.

There is at least one different group in A, B, C, D, E groups from the length variables.

II. Multiple-Comparison ANOVA

A.



One-way ANOVA: 1.6, 3.8, 6, 10.2

Method

Null hypothesis All means are equal
Alternative hypothesis At least one mean is different
Significance level $\alpha = 0.05$
Rows unused 2

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	4	1.6, 3.8, 6, 10.2

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	456.5	152.168	17.11	0.000
Error	14	124.5	8.893		
Total	17	581.0			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.98207	78.57%	73.98%	65.08%

Means

Factor	N	Mean	StDev	95% CI
1.6	5	58.28	3.60	(55.42, 61.14)
3.8	4	55.40	2.66	(52.20, 58.60)
6	4	50.85	2.43	(47.65, 54.05)
10.2	5	45.50	2.90	(42.64, 48.36)

Pooled StDev = 2.98207

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping
1.6	5	58.28	A
3.8	4	55.40	A B
6	4	50.85	B C
10.2	5	45.50	C

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
3.8 - 1.6	-2.88	2.00	(-8.69, 2.93)	-1.44	0.497
6 - 1.6	-7.43	2.00	(-13.24, -1.62)	-3.71	0.011
10.2 - 1.6	-12.78	1.89	(-18.26, -7.30)	-6.78	0.000
6 - 3.8	-4.55	2.11	(-10.68, 1.58)	-2.16	0.183
10.2 - 3.8	-9.90	2.00	(-15.71, -4.09)	-4.95	0.001
10.2 - 6	-5.35	2.00	(-11.16, 0.46)	-2.67	0.076

Individual confidence level = 98.85%

Tukey Simultaneous 95% CIs

Interval Plot of 1.6, 3.8, ...

One-way ANOVA: 1, 2, 3, 4, 5

Method

Null hypothesis	All means are equal
Alternative hypothesis	At least one mean is different
Significance level	$\alpha = 0.05$

Factor Information

Factor	Levels	Values
Factor	5	1, 2, 3, 4, 5

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	4	200.3	50.08	3.49	0.033
Error	15	215.5	14.37		
Total	19	415.8			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
3.79034	48.17%	34.35%	7.86%

Means

Factor	N	Mean	StDev	95% CI
1	4	12.75	4.19	(8.71, 16.79)
2	4	17.75	3.59	(13.71, 21.79)
3	4	17.50	2.08	(13.46, 21.54)
4	4	11.50	4.65	(7.46, 15.54)
5	4	10.00	3.92	(5.96, 14.04)

Pooled StDev = 3.79034

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping
2	4	17.75	A
3	4	17.50	A
1	4	12.75	A
4	4	11.50	A
5	4	10.00	A

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping
2	4	17.75	A
3	4	17.50	A
1	4	12.75	A
4	4	11.50	A
5	4	10.00	A

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
2 - 1	5.00	2.68	(-3.28, 13.28)	1.87	0.375
3 - 1	4.75	2.68	(-3.53, 13.03)	1.77	0.424
4 - 1	-1.25	2.68	(-9.53, 7.03)	-0.47	0.989
5 - 1	-2.75	2.68	(-11.03, 5.53)	-1.03	0.840
3 - 2	-0.25	2.68	(-8.53, 8.03)	-0.09	1.000
4 - 2	-6.25	2.68	(-14.53, 2.03)	-2.33	0.188
5 - 2	-7.75	2.68	(-16.03, 0.53)	-2.89	0.072
4 - 3	-6.00	2.68	(-14.28, 2.28)	-2.24	0.219
5 - 3	-7.50	2.68	(-15.78, 0.78)	-2.80	0.085
5 - 4	-1.50	2.68	(-9.78, 6.78)	-0.56	0.979

Individual confidence level = 99.25%

Tukey Simultaneous 95% CIs

|

Interval Plot of 1, 2, ...

B. Answer the following.

1. State the null and alternate hypotheses.

$H_0: \mu A = \mu B = \mu C = \mu D = \mu E$

H_1 : At least one population mean is different from the others

2. State the sample means for each of the independent variables.

ROW 1: 17.750

ROW 2: 17.500

ROW 3: 12.750

ROW 4: 11.500

ROW 5: 10.000

3. State the value of the ANOVA hypothesis test statistic found by SAS.

F value: 2.12

4. State the ANOVA p -value found by SAS.

0.1202

5. State whether this ANOVA p -value would indicate “reject the null hypothesis” or “fail to reject the null hypothesis” at the given level of significance.

The p value would indicate fail to reject the null hypothesis since the α is $0.05 < 0.1202$

6. State a conclusion for the ANOVA in words, in the context of the given situation.

There is no group that has a significant difference with other groups

7. Interpret the Tukey method analysis to determine any significant differences.

a. Write the 5 sample means in increasing order.

10.0 11.5 12.75 17.5 17.75

b. Use the SAS grouping information to determine whether each pairwise comparison indicates a significant difference: Underline each pair of means if there is *not* a significant difference indicated. Do

not underline a pair if there is a significant difference indicated.

10.0 11.5 12.75 17.5 17.75

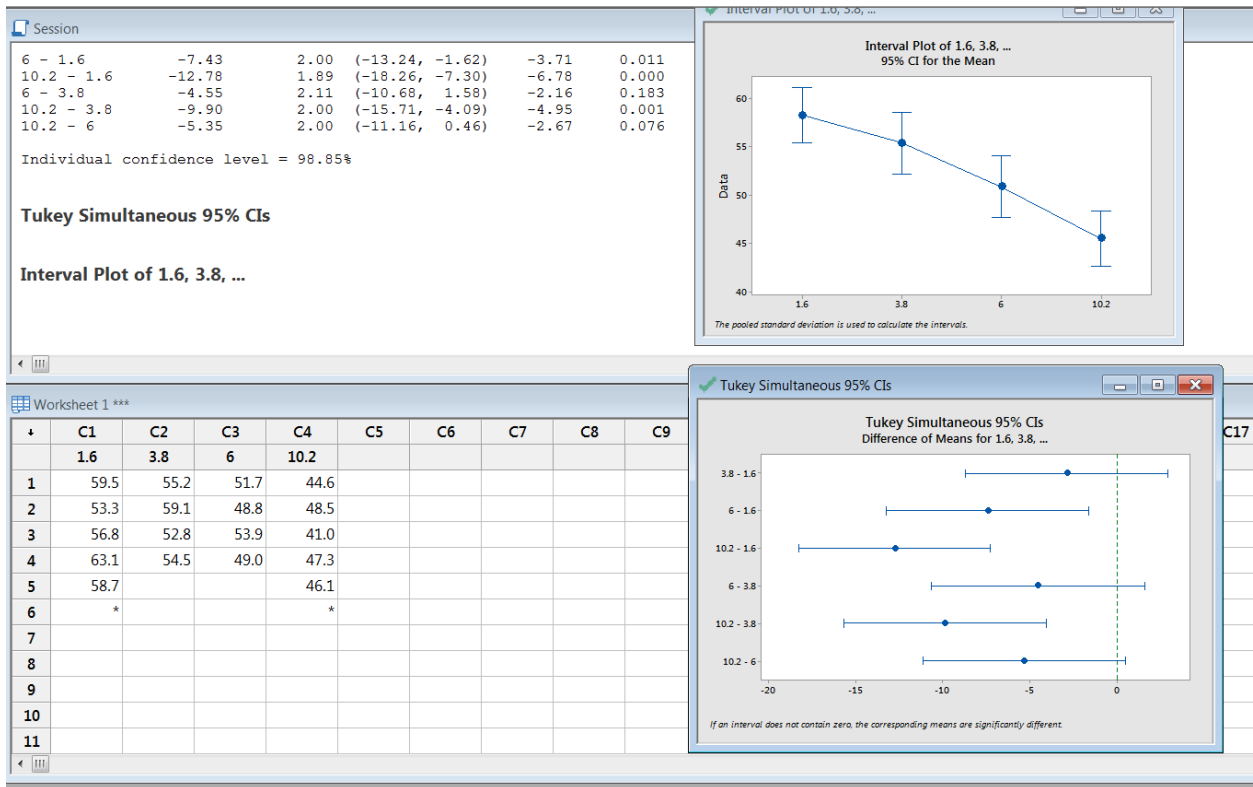
9. Answer the questions, “What happens when Tukey’s procedure is applied? Does the Tukey analysis

indicate which of the means is different from the rest?”

The Turkey’s procedure makes easy to visualize the differences between the mean values and separate them in multiple groups by the differences.

III. Single-Factor ANOVA with differing sample sizes

A.



Session

Method

Null hypothesis All means are equal
Alternative hypothesis At least one mean is different
Significance level $\alpha = 0.05$
Rows unused 2

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	4	1.6, 3.8, 6, 10.2

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	456.5	152.168	17.11	0.000
Error	14	124.5	8.893		
Total	17	581.0			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.98207	78.57%	73.98%	65.08%

Means

Factor	N	Mean	StDev	95% CI
1.6	5	58.28	3.60	(55.42, 61.14)
3.8	4	55.40	2.66	(52.20, 58.60)
6	4	50.85	2.43	(47.65, 54.05)
10.2	5	45.50	2.90	(42.64, 48.36)

Pooled StDev = 2.98207

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping
2	4	17.75	A
3	4	17.50	A
1	4	12.75	A
4	4	11.50	A
5	4	10.00	A

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
2 - 1	5.00	2.68	(-3.28, 13.28)	1.87	0.375
3 - 1	4.75	2.68	(-3.53, 13.03)	1.77	0.424
4 - 1	-1.25	2.68	(-9.53, 7.03)	-0.47	0.989
5 - 1	-2.75	2.68	(-11.03, 5.53)	-1.03	0.840
3 - 2	-0.25	2.68	(-8.53, 8.03)	-0.09	1.000
4 - 2	-6.25	2.68	(-14.53, 2.03)	-2.33	0.188
5 - 2	-7.75	2.68	(-16.03, 0.53)	-2.89	0.072
4 - 3	-6.00	2.68	(-14.28, 2.28)	-2.24	0.219
5 - 3	-7.50	2.68	(-15.78, 0.78)	-2.80	0.085
5 - 4	-1.50	2.68	(-9.78, 6.78)	-0.56	0.979

Individual confidence level = 99.25%

Tukey Simultaneous 95% CIs

Interval Plot of 1, 2, ...

One-way ANOVA: 1.6, 3.8, 6, 10.2

Method

One-way ANOVA: 1, 2, 3, 4, 5

Method

Null hypothesis	All means are equal
Alternative hypothesis	At least one mean is different
Significance level	$\alpha = 0.05$
Rows unused	4

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	5	1, 2, 3, 4, 5

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	4	200.3	50.08	3.49	0.033
Error	15	215.5	14.37		
Total	19	415.8			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
3.79034	48.17%	34.35%	7.86%

Means

Factor	N	Mean	StDev	95% CI
1	4	12.75	4.19	(8.71, 16.79)
2	4	17.75	3.59	(13.71, 21.79)
3	4	17.50	2.08	(13.46, 21.54)
4	4	11.50	4.65	(7.46, 15.54)
5	4	10.00	3.92	(5.96, 14.04)

Pooled StDev = 3.79034

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping
2	4	17.75	A
3	4	17.50	A
1	4	12.75	A
4	4	11.50	A
5	4	10.00	A

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
2 - 1	5.00	2.68	(-3.28, 13.28)	1.87	0.375
3 - 1	4.75	2.68	(-3.53, 13.03)	1.77	0.424
4 - 1	-1.25	2.68	(-9.53, 7.03)	-0.47	0.989
5 - 1	-2.75	2.68	(-11.03, 5.53)	-1.03	0.840
3 - 2	-0.25	2.68	(-8.53, 8.03)	-0.09	1.000
4 - 2	-6.25	2.68	(-14.53, 2.03)	-2.33	0.188
5 - 2	-7.75	2.68	(-16.03, 0.53)	-2.89	0.072
4 - 3	-6.00	2.68	(-14.28, 2.28)	-2.24	0.219
5 - 3	-7.50	2.68	(-15.78, 0.78)	-2.80	0.085
5 - 4	-1.50	2.68	(-9.78, 6.78)	-0.56	0.979

Individual confidence level = 99.25%

|

Tukey Simultaneous 95% CIs

Interval Plot of 1, 2, ...

One-way ANOVA: 1.6, 3.8, 6, 10.2

B. Answer the following.

1. State the null and alternate hypotheses.

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D$$

H_1 : At least one population mean is different from the others

2. State the sample means for each of the independent variables.

58.2800000

53.5400000

49.9725000

44.4725000

3. State the value of the ANOVA hypothesis test statistic found by SAS.

9.25

4. State the ANOVA p -value found by SAS.

0.0031

5. State whether this ANOVA p -value would indicate “reject the null hypothesis” or “fail to reject the null hypothesis” at the given level of significance.

reject the null hypothesis since p value is $0.0031 < 0.05$

6. State a conclusion for the ANOVA in words, in the context of the given situation.

There is at least one group is significantly different from others.

7. Interpret the Tukey method analysis to determine any significant differences by writing the sample means in increasing order, and underlining (or not) as appropriate (as was done for part II).

1	2	3	4
<hr/>			—

8. State an appropriate conclusion to Tukey’s procedure. Specifically, what do the overlapping significance connections imply?

The group 1,2,3 are significantly different from a group 4.