Assignment 9: ANOVA

Amartya Dutta

Problem 1:

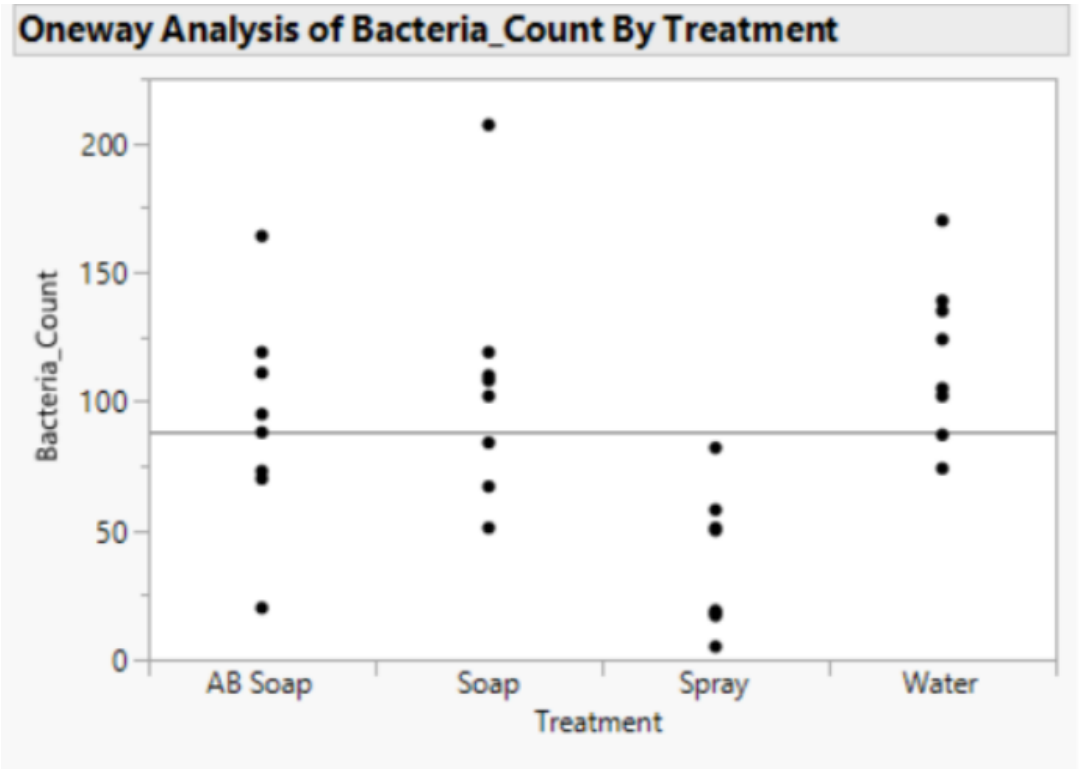
Part A:

Hypothesis:

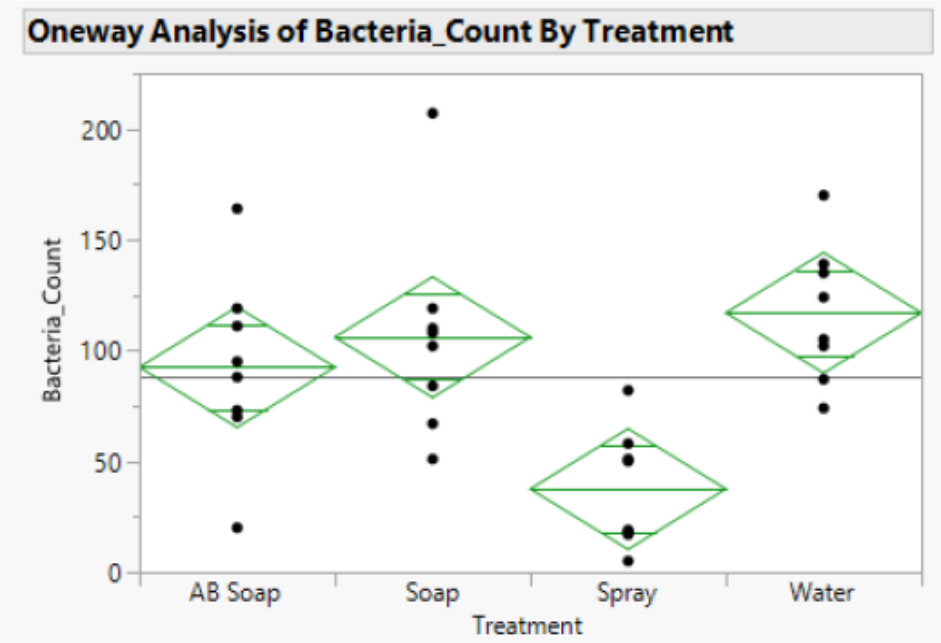
Null Hypothesis (H_0): The mean bacteria count is the same for all four washing methods.

Alternative Hypothesis (H_a): The mean bacteria count is different for at least one pair of washing methods.

Data Distribution Plot:



ANOVA Analysis:



Oneway Anova

Summary of Fit

Rsquare	0.430787
Adj Rsquare	0.3698
Root Mean Square Error	37.55187
Mean of Response	88.25
Observations (or Sum Wgts)	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment	3	29882.000	9960.67	7.0636	0.0011*
Error	28	39484.000	1410.14		
C. Total	31	69366.000			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
AB Soap	8	92.500	13.277	65.304	119.70
Soap	8	106.000	13.277	78.804	133.20
Spray	8	37.500	13.277	10.304	64.70
Water	8	117.000	13.277	89.804	144.20

Std Error uses a pooled estimate of error variance

F statistic: **7.0636**

p Value: **0.0011**

Conclusion: Since the p-value is less than 0.05, we can reject the null hypothesis, and say that there is at least one mean that is different from the rest.

Part B:

Analysis using Turkey's HSD Method:

Means Comparisons

Comparisons for all pairs using Tukey-Kramer HSD

Confidence Quantile

q*	Alpha
2.73031	0.05

HSD Threshold Matrix

Abs(Dif)-HSD

	Water	Soap	AB Soap	Spray
Water	-51.264	-40.264	-26.764	28.236
Soap	-40.264	-51.264	-37.764	17.236
AB Soap	-26.764	-37.764	-51.264	3.736
Spray	28.236	17.236	3.736	-51.264

Positive values show pairs of means that are significantly different.

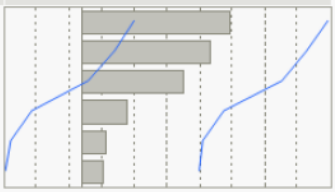
Connecting Letters Report

Level		Mean
Water	A	117.00000
Soap	A	106.00000
AB Soap	A	92.50000
Spray	B	37.50000

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
Water	Spray	79.50000	18.77593	28.2358	130.7642	0.0012*
Soap	Spray	68.50000	18.77593	17.2358	119.7642	0.0056*
AB Soap	Spray	55.00000	18.77593	3.7358	106.2642	0.0320*
Water	AB Soap	24.50000	18.77593	-26.7642	75.7642	0.5676
Soap	AB Soap	13.50000	18.77593	-37.7642	64.7642	0.8887
Water	Soap	11.00000	18.77593	-40.2642	62.2642	0.9355



Conclusion: From the results obtained we can see that ab_soap -spray, soap-spray and water- spray have p-value < 0.05. Hence, we reject the null hypothesis and they have significant differences. While, soap-ab_soap, water-ab_soap, and water-soap have p-value > 0.05. Thus, we do not reject the null hypothesis and they have very similar values.

Problem 2:

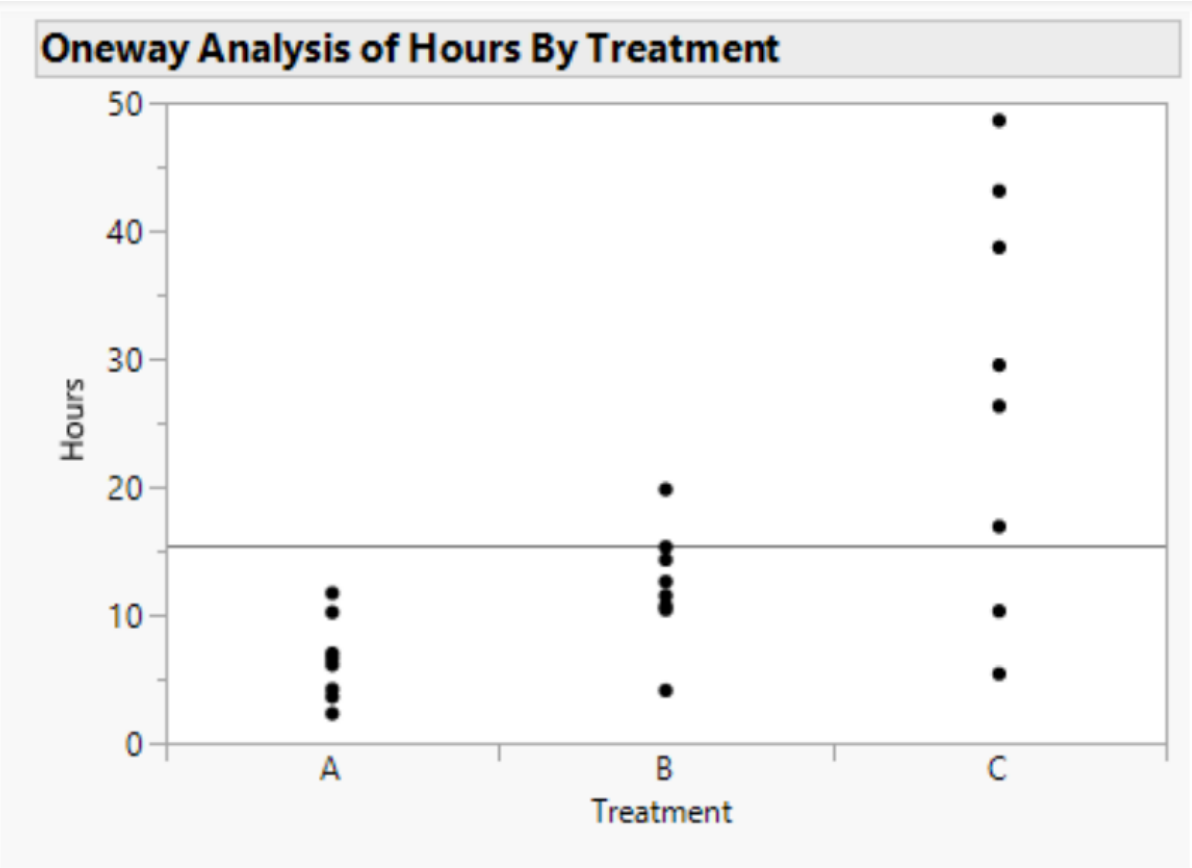
Part A:

Hypothesis:

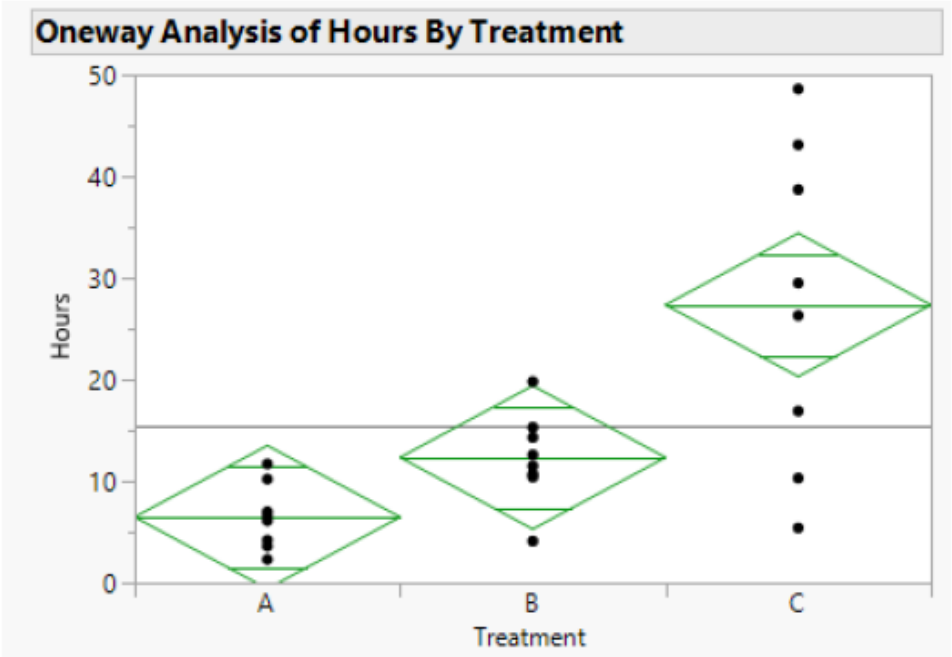
Null Hypothesis (H_0): The mean hours count is the same for all three treatment methods.

Alternative Hypothesis (H_a): The mean hours count is different for different treatment methods.

Data Distribution Plot:



ANOVA Analysis:



Oneway Anova

Summary of Fit

Rsquare	0.490066
Adj Rsquare	0.441501
Root Mean Square Error	9.591011
Mean of Response	15.38333
Observations (or Sum Wgts)	24

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment	2	1856.4758	928.238	10.0909	0.0008*
Error	21	1931.7375	91.988		
C. Total	23	3788.2133			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
A	8	6.4625	3.3909	-0.59	13.514
B	8	12.3375	3.3909	5.29	19.389
C	8	27.3500	3.3909	20.30	34.402

Std Error uses a pooled estimate of error variance

F statistic: 10.0909

p Value: **0.0008**

Conclusion: Since the p-value is less than 0.05, we can reject the null hypothesis, and say that there is at least one mean that is different from the rest.

Part B:

Analysis using Turkey's HSD Method:

Means Comparisons

Comparisons for all pairs using Tukey-Kramer HSD

Confidence Quantile

q*	Alpha
2.52057	0.05

HSD Threshold Matrix

Abs(Dif)-HSD

	C	B	A
C	-12.087	2.925	8.800
B	2.925	-12.087	-6.212
A	8.800	-6.212	-12.087

Positive values show pairs of means that are significantly different.

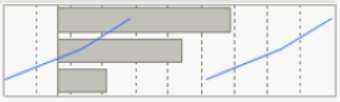
Connecting Letters Report

Level			Mean
C	A		27.350000
B		B	12.337500
A		B	6.462500

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
C	A	20.88750	4.795506	8.80009	32.97491	0.0008*
C	B	15.01250	4.795506	2.92509	27.09991	0.0134*
B	A	5.87500	4.795506	-6.21241	17.96241	0.4520



Conclusion: From the above table we can see that the B-A pair has a p-value >0.05. So, we do not reject the null hypothesis for this pair and thus they have similar mean treatment hours.

Whereas, pairs C-A and C-B have p-values < 0.05 and as a result we reject the null-hypothesis. Thus, they also have different mean treatment hours