

Problem 1

b) Three ways of R code to get F test statistic value 6.062, and p-value is 0.0215.

Since $0.0215 < 0.05$, thus Reject H_0 at $\alpha = 5\%$ significance level.

First Way:

```
> Drama = c(1.8, 0.9, 1.5, 2.4)
> Writing = c(2.4, 3.3, 3.9, 3.6)
> Statistics = c(2.1, 2.4, 3.0, 3.9)
> Clubs = c(rep("Drama", 4), rep("Writing", 4), rep("Statistics", 4))
> GPA = c(Drama, Writing, Statistics)
> results = glm(GPA ~ factor(Clubs))
> summary(aov(results))
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(Clubs)	2	5.82	2.91	6.062	0.0215 *
Residuals	9	4.32	0.48		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Second Way:

```
> Drama <- c(rep(1, 4), rep(0, 8)); Drama
[1] 1 1 1 1 0 0 0 0 0 0 0 0
> Writing <- c(rep(0, 4), rep(1, 4), rep(0, 4)); Writing
[1] 0 0 0 0 1 1 1 1 1 0 0 0
> Statistics <- c(rep(0, 8), rep(1, 4)); Statistics
[1] 0 0 0 0 0 0 0 0 1 1 1 1
> results2 = lm(GPA ~ Drama + Writing + Statistics)
> summary(results2)
```

Call:

```
lm(formula = GPA ~ Drama + Writing + Statistics)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.900	-0.525	0.075	0.375	1.050

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.8500	0.3464	8.227	1.77e-05 ***
Drama	-1.2000	0.4899	-2.449	0.0368 *
Writing	0.4500	0.4899	0.919	0.3823
Statistics	NA	NA	NA	NA

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6928 on 9 degrees of freedom

Multiple R-squared: 0.574, Adjusted R-squared: 0.4793

F-statistic: 6.062 on 2 and 9 DF, p-value: 0.0215

Third Way:

```
> anova(lm(GPA ~ 1), results2)
```

Analysis of Variance Table

Model 1: GPA ~ 1

Model 2: GPA ~ Drama + Writing + Statistics

	Res. Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	11	10.14				
2	9	4.32	2	5.82	6.0625	0.0215 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

e)

```
> Drama = c(1.8, 0.9, 1.5, 2.4)
> Writing = c(2.4, 3.3, 3.9, 3.6)
> Statistics = c(2.1, 2.4, 3.0, 3.9)
> Clubs = c(rep("Drama", 4), rep("Writing", 4), rep("Statistics", 4))
> GPA = c(Drama, Writing, Statistics)
> results = glm(GPA ~ factor(Clubs))
> TukeyHSD(aov(results))
Tukey multiple comparisons of means
95% family-wise confidence level
```

Fit: aov(formula = results)

	diff	lwr	upr	p adj
Statistics-Drama	1.20	-0.1677978	2.567798	0.0851150
Writing-Drama	1.65	0.2822022	3.017798	0.0204285
Writing-Statistics	0.45	-0.9177978	1.817798	0.6428419

g)

```
> qt(1-0.05/6, 9)
[1] 2.933324
```

j)

```
> Drama = c(1.8, 0.9, 1.5, 2.4)
> Writing = c(2.4, 3.3, 3.9, 3.6)
> Statistics = c(2.1, 2.4, 3.0, 3.9)
> Clubs = c(rep("Drama", 4), rep("Writing", 4), rep("Statistics", 4))
> GPA = c(Drama, Writing, Statistics)
> kruskal.test(GPA ~ factor(Clubs))
```

Kruskal-Wallis rank sum test

data: GPA by factor(Clubs)

Kruskal-Wallis chi-squared = 6.234, df = 2, p-value = 0.04429

Since p-value = 0.04429 < 0.05, thus Reject H_0 at $\alpha = 5\%$ significance level.