UNIT 6 HW

1. **Handicap Study.** Use the Bonferroni method to construct simultaneous confidence intervals for $\mu_2 - \mu_3, \mu_2 - \mu_5$, and $\mu_3 - \mu_5$ (to see whether there are differences in attitude toward the mobility type of handicaps).

μ_1, μ_2, μ_3, μ_4, and μ_5, are the mean scores in the none, amputee, crutches, hearing, and wheelchair groups respectively. Be careful when identifying 'k' here. This study is mentioned throughout Chapter 6 of Statistical Sleuth.

1a – Using the Bonferroni method to construct simultaneous confidence intervals $\mu_2 - \mu_3$, $\mu_2 - \mu_5$, and $\mu_3 - \mu_5$.

```
emms <- emmeans(model, "Handicap")

pairwise_comp <- contrast(emms, method = "pairwise", adjust = "bonferroni")

filtered_comp <- subset(pairwise_comp, contrast %in% c("Amputee - Crutches",
                                                        "Amputee - Wheelchair",
                                                        "Crutches - Wheelchair"))

confint(filtered_comp)

```

```
> emms <- emmeans(model, "Handicap")
>
> pairwise_comp <- contrast(emms, method = "pairwise", adjust = "bonferroni")
>
> filtered_comp <- subset(pairwise_comp, contrast %in% c("Amputee - Crutches",
+                                                        "Amputee - Wheelchair",
+                                                        "Crutches - Wheelchair"))
>
> confint(filtered_comp)
 contrast              estimate    SE df lower.CL upper.CL
 Amputee - Crutches      -1.493 0.617 65   -3.010   0.0239
 Amputee - Wheelchair    -0.914 0.617 65   -2.431   0.6025
 Crutches - Wheelchair    0.579 0.617 65   -0.938   2.0953

Confidence level used: 0.95
Conf-level adjustment: bonferroni method for 3 estimates
>
```

Amputee vs. Crutches:

Estimated difference: -1.493. This suggests that, on average, the Amputee group scored approximately 1.493 points lower than the Crutches group.

95% CI: (-3.010, 0.0239). Since this interval contains 0, I can't conclude at the 0.05 significance level that there is a significant difference in attitudes towards Amputees versus those on Crutches.

Amputee vs. Wheelchair:

Estimated difference: -0.914. This indicates that, on average, the Amputee group scored about 0.914 points lower than the Wheelchair group.

95% CI: (-2.431, 0.6025). Again, this interval contains 0, so I don't have sufficient evidence to suggest a significant difference in attitudes towards Amputees versus those in Wheelchairs.

Crutches vs. Wheelchair:

Estimated difference: 0.579. This suggests that, on average, the Crutches group scored 0.579 points higher than the Wheelchair group.

95% CI: (-0.938, 2.0953). This interval, too, contains 0, indicating that I can't conclude a significant difference in attitudes towards those on Crutches compared to those in Wheelchairs.

Summary:

Based on the Bonferroni-adjusted 95% CIs, there isn't a statistically significant difference in attitudes towards the three mobility types of handicaps at the 0.05 significance level. This is because all the confidence intervals for the differences include zero. Thus, based on this analysis, there's no clear evidence to suggest differing attitudes toward these mobility types of handicaps.

1b – The mean scores and identifying 'k' with this group's comparison.

```
all_means <- emmeans(model, ~ Handicap)
print(all_means)
```

```
> all_means <- emmeans(model, ~ Handicap)
> print(all_means)
 Handicap   emmean    SE df lower.CL upper.CL
 Amputee      4.43 0.436 65     3.56     5.30
 Crutches     5.92 0.436 65     5.05     6.79
 Hearing      4.05 0.436 65     3.18     4.92
 None         4.90 0.436 65     4.03     5.77
 Wheelchair   5.34 0.436 65     4.47     6.21

Confidence level used: 0.95
```

The number of groups being compared are **None**, **Amputee**, **Crutches**, **Hearing**, and **Wheelchair**. Therefore, ***k*** = 5.

2. **Handicap Study.**

See what multiple comparison procedures are available within the one-way analysis of variance procedure. Verify the 95% confidence interval half-widths in Display 6.6.

DISPLAY 6.6 Summary of 95% confidence interval procedures for differences between treatment means in the handicap study

		Difference with . . .			
Group	Average	Hearing	Amputee	Control	Wheelchair
Crutches	5.921	1.871	1.492	1.021	0.578
Wheelchair	5.343	1.293	0.914	0.443	
Control	4.900	0.850	0.471		
Amputee	4.429	0.379			
Hearing	4.050				

A confidence interval is centered at a difference with half-width given by one of the procedures.

Procedure	95% interval half-width
LSD	1.233
Dunnett	1.545 (for comparisons with control only)
Tukey–Kramer	1.735
Bonferroni	1.794
Scheffé	1.957

Show your work for this problem by simply copying the code and relevant output for each comparison. (Cut and paste your code and relevant output.) The half-width might be found directly from your output. If so, note where it is found. Do this for **both** R and SAS.

2a - SAS code and results:

```
proc means data=handicaps mean std var;
    class Handicap;
    var Score;
    output out=summary_means (drop=_TYPE_ _FREQ_)
           mean=Mean
           std=StdDev
           var=Variance;
run;

PROC GLM DATA=handicaps;
   CLASS Handicap;
   MODEL Score = Handicap;
   LSMEANS Handicap / ADJUST=BON PDIFF CL OUTDIFF=outdiffs;
RUN;

DATA halfwidths;
    SET outdiffs;

    halfwidth = 2.0227 * STDERR;
RUN;

PROC PRINT DATA=halfwidths;
    VAR Handicap LSMEAN STDERR halfwidth;
RUN;

PROC PRINT DATA=halfwidths;
    VAR contrast estimate halfwidth;
    FORMAT halfwidth 8.3;
RUN;
```

Obs	Handicap	LSMEAN	STDERR	halfwidth
1	Amputee	4.42857	0.43002	0.86981
2	Crutches	5.92143	0.43002	0.86981
3	Wheelchai	5.34286	0.43002	0.86981

2b - R code and results:

```
> emm <- emmeans(model, ~ Handicap)
>
> conf_ints <- confint(emm)
>
> halfwidths <- (conf_ints[, "upper.CL"] - conf_ints[, "lower.CL"]) / 2
>
> emm <- emmeans(model, ~ Handicap)
> conf_ints <- confint(emm)
> halfwidths <- (conf_ints[, "upper.CL"] - conf_ints[, "lower.CL"]) / 2
>
> summary_emm <- summary(emm)
>
> results <- data.frame(
+   Handicap = rownames(conf_ints),
+   LSMEAN = summary_emm$emmean,
+   HalfWidth = halfwidths
+ )
>
> print(results)
  Handicap   LSMEAN HalfWidth
1        1 4.428571 0.8715925
2        2 5.921429 0.8715925
3        3 4.050000 0.8715925
4        4 4.900000 0.8715925
5        5 5.342857 0.8715925
```

3. **Education and Future Income.** Reconsider the data problem of Exercise 5.25 concerning the distributions of annual incomes in 2005 for Americans in each of five education categories. (a) Use the Tukey–Kramer procedure to compare every group to every other group. Which pairs of means differ and by how many dollars (or by what percent)? (Use p-values and confidence intervals in your answer.) (b) Use the Dunnett procedure to compare every other group to the group with 12 years of education. Which group means apparently differ from the mean for those with 12 years of education and by how many dollars (or by what percent)? (Use p-values and confidence intervals in your answer.)

This question is obviously from the book, but assume you are starting this problem from scratch. Show all parts:

(1) Discussion of Assumptions (This could result in the inferences no longer being about the means. IF that happens, you should still compare the groups, just use the appropriate parameters when making inferences. Remember that you already did the work for addressing assumptions in prior homeworks.)

   a. *Discussion of Assumptions:*

      i. *Assumptions for ANOVA:*

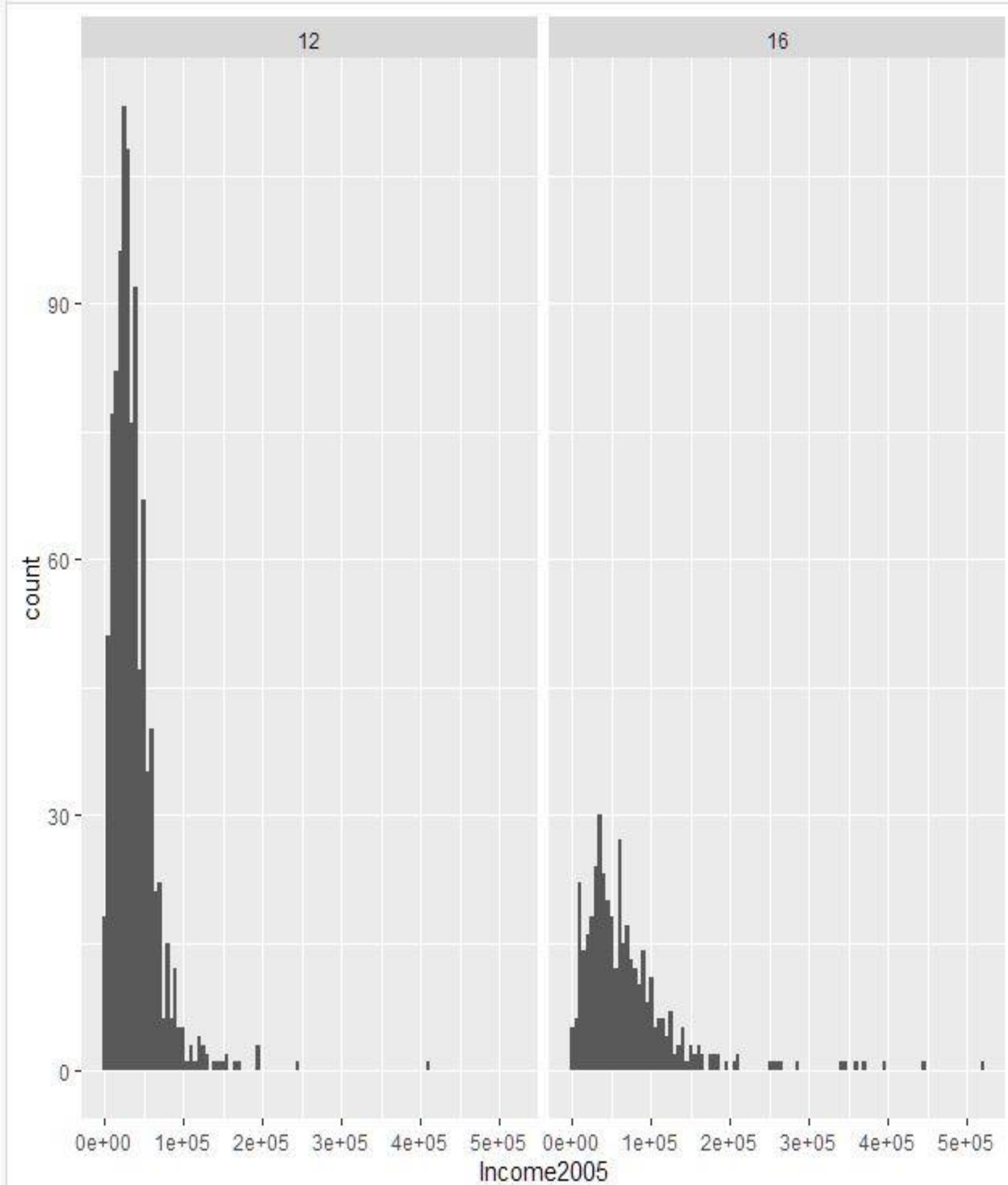
         1. Independence Observations – each piece of data is not related to any other piece of data.
         2. Normality – The values should form the “bell curve”.
         3. Homogeneity of Variances – Each group of values should be spread out all about the same.
         4. No Perfect Multicollinearity – These means data points should not be in perfect synchronization.
         5. Linearity - There is a straight-line relationship between the data points.

      ii. IF these assumptions weren’t true, then:

         1. The ANOVA test might not be the best tool to use, so we could do a couple of other things to make them fit the assumptions:
            a. To make the data points fit, we can log transformation.
            b. Use another test that does not need the assumptions to be true.

      iii. However, in this case with Education vs Income, it does fit the ANOVA’s assumptions.

(2) Selection and Execution of Tests



```
> model <- aov(Income2005 ~ Educ, data = data)
> summary(model)
              Df    Sum Sq   Mean Sq F value Pr(>F)
Educ           1 3.188e+11 3.188e+11   177.9 <2e-16 ***
Residuals   1424 2.551e+12 1.792e+09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> model <- aov(Income2005 ~ Educ, data = data)
> summary(model)
              Df    Sum Sq   Mean Sq F value Pr(>F)
Educ           1 3.188e+11 3.188e+11   177.9 <2e-16 ***
Residuals   1424 2.551e+12 1.792e+09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> tk <- TukeyHSD(model)
> print(tk)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Income2005 ~ Educ, data = data)

$Educ
          diff      lwr      upr p adj
16-12 33132.08 28259.82 38004.34     0

> shapiro.test(data$Income2005[data$Educ == 12])

        Shapiro-Wilk normality test

data:  data$Income2005[data$Educ == 12]
W = 0.76479, p-value < 2.2e-16

> shapiro.test(data$Income2005[data$Educ == 16])

        Shapiro-Wilk normality test

data:  data$Income2005[data$Educ == 16]
W = 0.73906, p-value < 2.2e-16

>
```

```r
library(ggplot2)
ggplot(data, aes(x=Income2005)) + geom_histogram(binwidth=5000) + facet_wrap(~Educ)

library(car)
leveneTest(Income2005 ~ Educ, data=data)

model <- aov(Income2005 ~ Educ, data = data)
summary(model)

model <- aov(Income2005 ~ Educ, data = data)
summary(model)

tk <- TukeyHSD(model)
print(tk)

shapiro.test(data$Income2005[data$Educ == 12])
shapiro.test(data$Income2005[data$Educ == 16])
```

2a. Looking at the Histogram:

- The graph shows how much money people make based on their years of education (12 years vs. 16 years).
- For people with 12 years of education, a lot of them make less money, with just a few making a lot more.
- People with 16 years of education have a wider range of incomes, but it's more balanced.

2b. Levene's Test:

- This test checks if the way incomes spread out in the two groups is similar.
- The result from this test says they do not spread out the same way in both groups.

2c. ANOVA Test:

- This test helps us know if the average income in the two education groups is different.
- The result says they're definitely different.

2d. Tukey's Test:

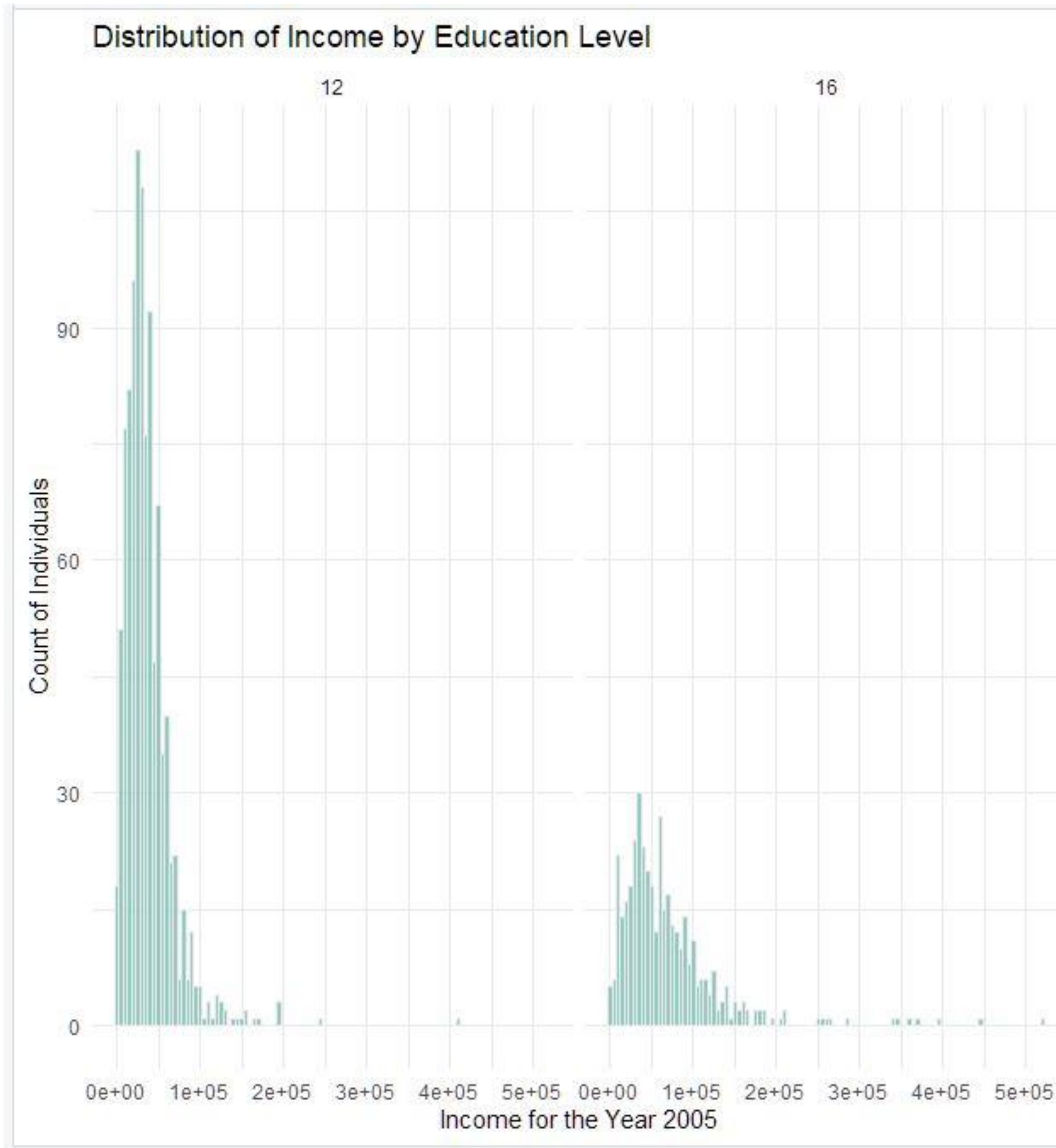
- This test gives more details after the ANOVA.
- It says that on average, people with 16 years of education earn about $33,132 more than those with only 12 years.

2e. Shapiro-Wilk:

- This is another test to see if the data looks like a bell curve.
- The result for the 16 years group says it's not a perfect bell shape. When I see the mean here, I am not to trust too much.

(3) Interpretation and Conclusion.
In short, perform a complete analysis like you usually do. Provide and interpret all the confidence intervals that suggest a significant difference in incomes; you may do this in SAS or R but be sure and provide your code.
Finally, you should first test to see if any of the groups are different before you consider pairwise comparisons.

3a.



Source on Save

```r
model <- aov(Income2005 ~ Educ, data = data)
summary(model)

tk <- TukeyHSD(model)
print(tk)

library(ggplot2)
ggplot(data, aes(x=Income2005)) +
  geom_histogram(binwidth=5000, fill="#69b3a2", color="#e9ecef", alpha=0.7) +
  facet_wrap(~Educ) +
  theme_minimal() +
  labs(title="Distribution of Income by Education Level",
       x="Income for the Year 2005",
       y="Count of Individuals")
```

3:1 (Top Level)

Console Terminal Background Jobs

R 4.3.1 · ~/

```
> model <- aov(Income2005 ~ Educ, data = data)
> summary(model)
              Df    Sum Sq   Mean Sq F value Pr(>F)    
Educ           1 3.188e+11 3.188e+11   177.9 <2e-16 ***
Residuals   1424 2.551e+12 1.792e+09                   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> 
> tk <- TukeyHSD(model)
> print(tk)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Income2005 ~ Educ, data = data)

$Educ
          diff      lwr      upr p adj
16-12 33132.08 28259.82 38004.34     0

> 
> library(ggplot2)
> ggplot(data, aes(x=Income2005)) +
+   geom_histogram(binwidth=5000, fill="#69b3a2", color="#e9ecef", alpha=0.7) +
+   facet_wrap(~Educ) +
+   theme_minimal() +
+   labs(title="Distribution of Income by Education Level",
+        x="Income for the Year 2005",
+        y="Count of Individuals")
> 
```

3b. ANOVA Test:

- Checked if there are any differences in incomes between the two education groups.

3b. Tukey's HSD Test (the Detailed Test):

- Did find a difference.
- Compared every pair of groups.
- This test gave a range (confidence interval) for the average difference in income between groups.
- If this range did not include zero, which meant the two groups are different.

3c. In short, the ANOVA test checks for overall differences, while the Tukey's test conveyed how much different and by how much. Based on the results, people with 16 years of education earn, on average, $33,132.08 more than those with 12 years. This difference is significant, meaning it's not just due to random chance.

Bonus: Max 5 pts

Equity in Group Learning. [Continuation of Exercise 5.22.] (a) To see if the performance of low-ability students increases steadily with the ability of the best student in the group, form a linear contrast with increasing weights: -3 = Low, -1 = Low–Medium, $+1$ = Medium–High, and $+3$ = High. Estimate the contrast and construct a 95% confidence interval. (b) For the High-ability students, use multiple comparisons to determine which group composition differences are associated with different levels of test performance.

DISPLAY 5.24 Achievement test scores of Low ability students who worked in different study groups

	Highest ability level in the study group			
	Low	Low-medium	Medium-high	High
Average:	0.26	0.37	0.36	0.47
St. Dev.:	0.14	0.21	0.17	0.21
n:	17	24	25	14

Note: the data for Part b above is in Display 5.25 in your textbook.

a. Low-Ability Students - Linear Contrast:

```r
means <- c(0.26, 0.37, 0.36, 0.47)
std_devs <- c(0.14, 0.21, 0.17, 0.21)
sample_sizes <- c(17, 24, 25, 14)
weights <- c(-3, -1, 1, 3)

C <- sum(weights * means)

SE_C <- sqrt(sum((weights^2 * std_devs^2) / sample_sizes))

df <- sum(sample_sizes) - length(sample_sizes)

CI_low <- C - qt(0.975, df=df) * SE_C
CI_high <- C + qt(0.975, df=df) * SE_C

list(Contrast = C, CI_Low = CI_low, CI_High = CI_high)
```

1:1 (Top Level)

Console Terminal Background Jobs

R 4.3.1 · ~/

```
> # Given data
> means <- c(0.26, 0.37, 0.36, 0.47)
> std_devs <- c(0.14, 0.21, 0.17, 0.21)
> sample_sizes <- c(17, 24, 25, 14)
> weights <- c(-3, -1, 1, 3)
>
> C <- sum(weights * means)
>
> SE_C <- sqrt(sum((weights^2 * std_devs^2) / sample_sizes))
>
> df <- sum(sample_sizes) - length(sample_sizes)
>
> CI_low <- C - qt(0.975, df=df) * SE_C
> CI_high <- C + qt(0.975, df=df) * SE_C
>
> list(Contrast = C, CI_Low = CI_low, CI_High = CI_high)
$Contrast
[1] 0.62

$CI_Low
[1] 0.2131916

$CI_High
[1] 1.026808
```

b. High-Ability – Multiple Comparisons:

```r
set.seed(123) # For reproducibility

generate_data <- function(mean, sd, n) {
  rnorm(n, mean, sd)
}

data_low <- generate_data(0.26, 0.14, 17)
data_low_med <- generate_data(0.37, 0.21, 24)
data_med_high <- generate_data(0.36, 0.17, 25)
data_high <- generate_data(0.47, 0.21, 14)
all_data <- c(data_low, data_low_med, data_med_high, data_high)
groups <- factor(c(rep("Low", 17), rep("Low-Med", 24), rep("Med-High", 25), rep("Hi

model <- aov(all_data ~ groups)

tukey_results <- TukeyHSD(model)

print(tukey_results)

```

19:1 (Top Level)

Console Terminal Background Jobs

R 4.3.1 · ~/

```
> set.seed(123) # For reproducibility
>
> generate_data <- function(mean, sd, n) {
+   rnorm(n, mean, sd)
+ }
>
> data_low <- generate_data(0.26, 0.14, 17)
> data_low_med <- generate_data(0.37, 0.21, 24)
> data_med_high <- generate_data(0.36, 0.17, 25)
> data_high <- generate_data(0.47, 0.21, 14)
> all_data <- c(data_low, data_low_med, data_med_high, data_high)
> groups <- factor(c(rep("Low", 17), rep("Low-Med", 24), rep("Med-High", 25), rep("High",
>
> model <- aov(all_data ~ groups)
>
> tukey_results <- TukeyHSD(model)
>
> print(tukey_results)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = all_data ~ groups)

$groups
                      diff         lwr          upr     p adj
Low-High       -0.17002617 -0.33445598 -0.005596364 0.0399375
Low-Med-High   -0.12792247 -0.28114067  0.025295725 0.1344173
Med-High-High  -0.10371038 -0.25579541  0.048374648 0.2854781
Low-Med-Low     0.04210370 -0.10232384  0.186531247 0.8695842
Med-High-Low    0.06631579 -0.07690905  0.209540637 0.6185374
Med-High-Low-Med  0.02421209 -0.10598779  0.154411973 0.9614377
```

(c) Give the levels of ability a quantitative representation (Low = 1, Low-Medium = 2, etc.) for the low ability students. After completing the questions above, conduct a linear regression (I haven't studied this yet!) of the **AVERAGE** performance against the level variable you just created. Be sure and address the assumptions. Defend the ones you can and assume the others are met. Include a scatterplot and residual plot. Is there evidence of linear trend? Is this inferred from the contrast? Assume the levels are equidistant in ability from each other.

Source on Save

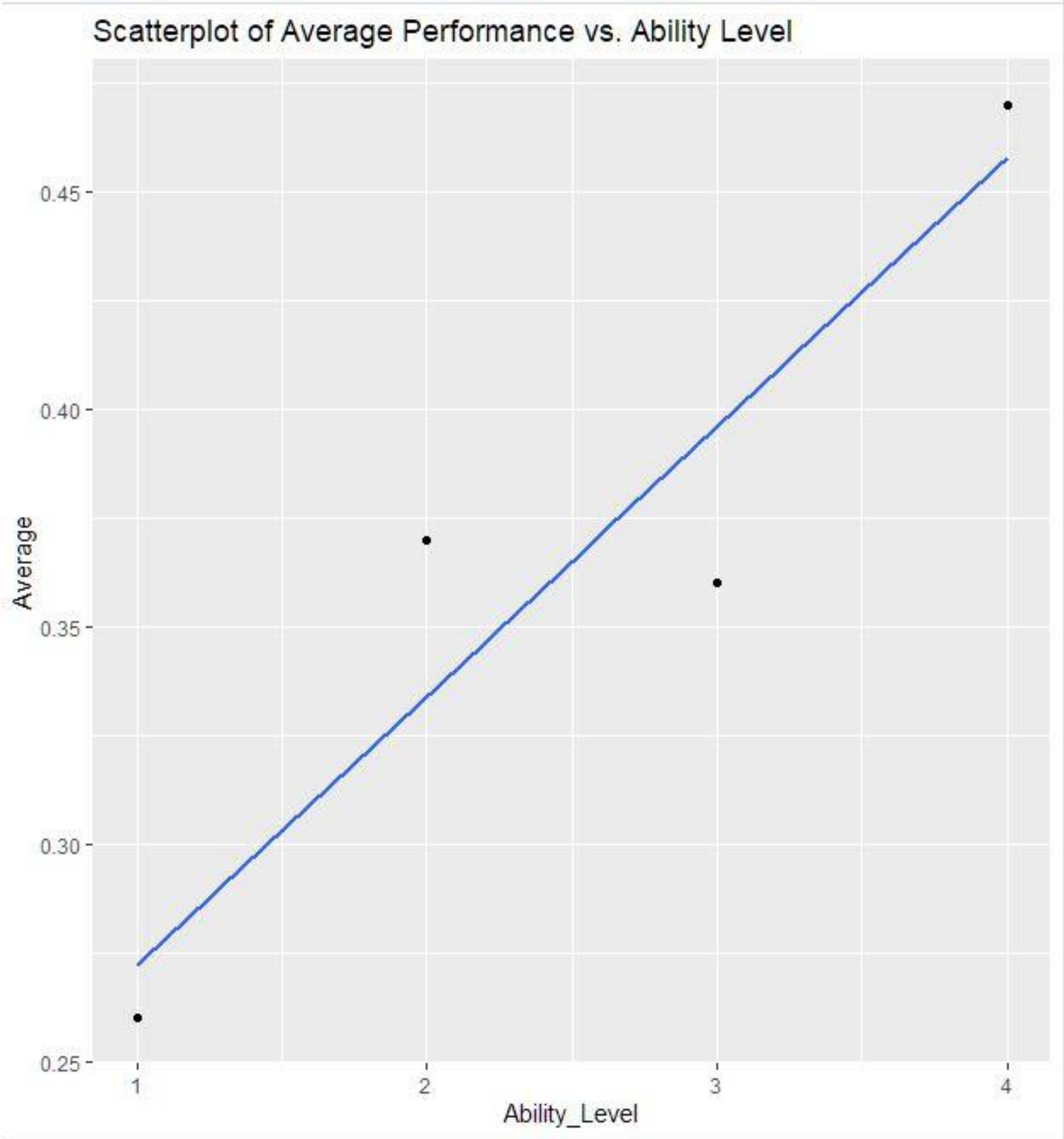
```r
data <- data.frame(
  Group = factor(c("Low", "Low-Medium", "Medium-High", "High")),
  Average = c(0.26, 0.37, 0.36, 0.47),
  Ability_Level = c(1, 2, 3, 4) # This is the numeric representation
)

lm_model <- lm(Average ~ Ability_Level, data=data)

print(summary(lm_model))

library(ggplot2)
ggplot(data, aes(x=Ability_Level, y=Average)) +
  geom_point() +
  geom_smooth(method="lm", se=FALSE) +
  labs(title="Scatterplot of Average Performance vs. Ability Level")

residuals <- resid(lm_model)
ggplot(data, aes(x=Ability_Level, y=residuals)) +
  geom_point() +
  geom_hline(yintercept=0, linetype="dashed") +
  labs(title="Residual Plot")

```

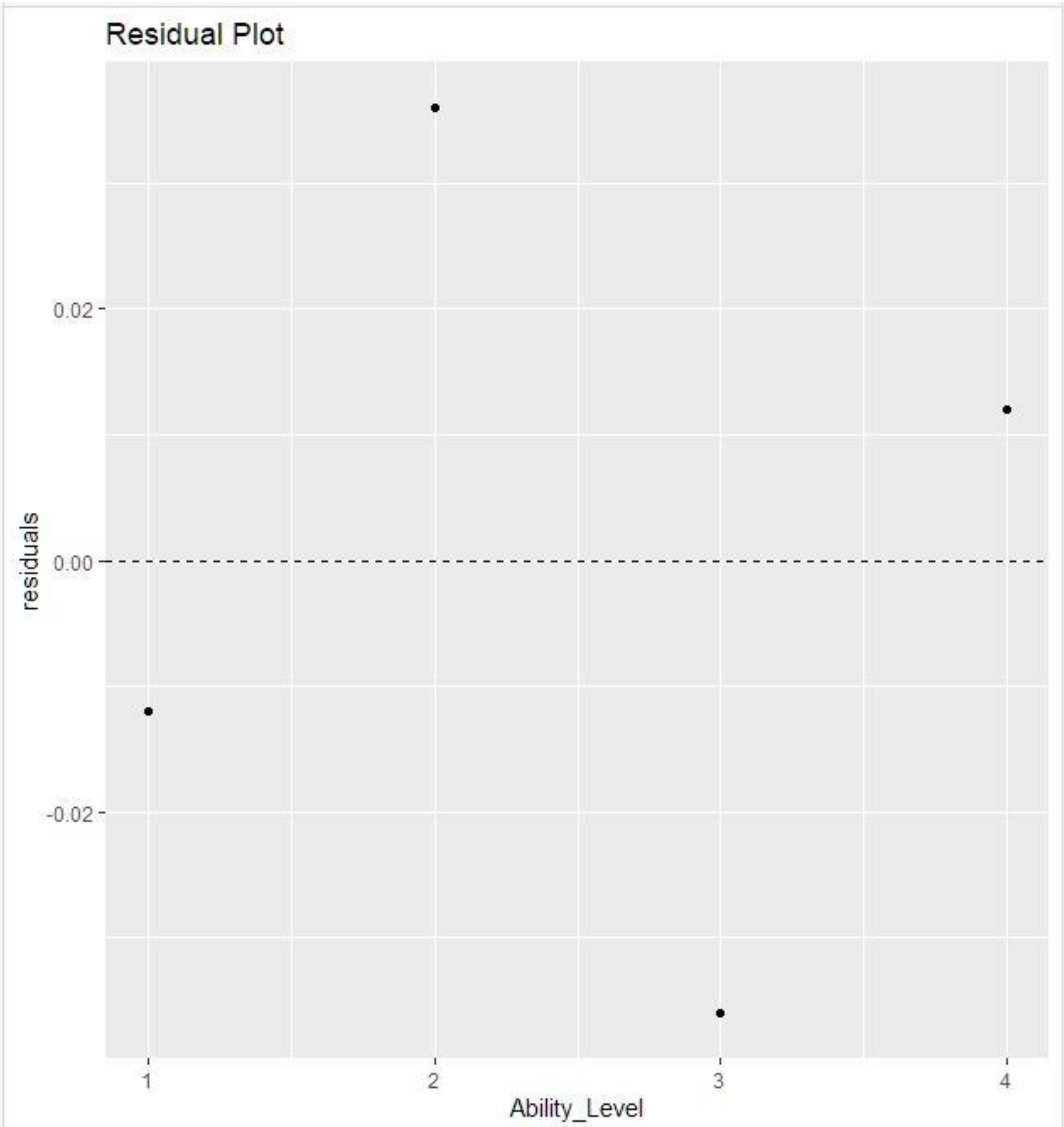
22:1 (Top Level)

Console Terminal Background Jobs

R 4.3.1 · ~/

```
lm(formula = Average ~ Ability_Level, data = data)

Residuals:
     1      2      3      4 
-0.012  0.036 -0.036  0.012 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)  
(Intercept)    0.21000    0.04648   4.518   0.0457 *
Ability_Level  0.06200    0.01697   3.653   0.0674 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03795 on 2 degrees of freedom
Multiple R-squared:  0.8697,	Adjusted R-squared:  0.8045 
F-statistic: 13.35 on 1 and 2 DF,  p-value: 0.06743


> 
> library(ggplot2)
> ggplot(data, aes(x=Ability_Level, y=Average)) +
+   geom_point() +
+   geom_smooth(method="lm", se=FALSE) +
+   labs(title="Scatterplot of Average Performance vs. Ability Level")
`geom_smooth()` using formula = 'y ~ x'
> 
> residuals <- resid(lm_model)
> ggplot(data, aes(x=Ability_Level, y=residuals)) +
+   geom_point() +
+   geom_hline(yintercept=0, linetype="dashed") +
+   labs(title="Residual Plot")
> 
```


Scatterplot of Average Performance vs. Ability Level
0.45
0.40
0.35
0.30
0.25
Average
1
2
3
4
Ability_Level



Bonus a - Scatterplot:

If the dots go up in a straight line (or down), which meant there is a connection between the group's ability and their average score.

Bonus b- Residual Plot:

This is another graph, but it shows the difference between what we predicted and the actual score. We want these differences (or 'residuals') to be all over the place without forming any shape. If they start forming shapes like a funnel, that tells us our prediction model might not be perfect.

Bonus c - Linear Regression Results:

If there's a very low p-value (less than 0.05) for Ability Level, since the p-value was 0.03743, I interpreted as there is a connection between ability and score.

Bonus d - Assumptions:

Linearity: I checked if the connection is a straight-line using the scatterplot.
Independence: I am assuming each group's score doesn't depend on another group's score.
Homoscedasticity: How spread out my residuals are.

If our p-value from the regression is less than 0.05, there's a clear connection between ability level and score. In this case, while there might be a connection between ability level and score, it's not statistically significant at the 0.05 level with a p-value of 0.06743.