**CHAPTER 13 SOLUTIONS**

**13.1.** The R commands used to obtain the output are:

**boxplot(NELS$achmat08~NELS$hsprog)**

**table(NELS$hsprog)**

**skew(NELS$achmat08[NELS$hsprog == "Rigorous Academic"])**

**se.skew(NELS$achmat08[NELS$hsprog == "Rigorous Academic"])**

**skew.ratio(NELS$achmat08[NELS$hsprog == "Rigorous Academic"])**

**skew(NELS$achmat08[NELS$hsprog == "Academic"])**

**se.skew(NELS$achmat08[NELS$hsprog == "Academic"])**

**skew.ratio(NELS$achmat08[NELS$hsprog == "Academic"])**

**skew(NELS$achmat08[NELS$hsprog == "Some Vocational"])**

**se.skew(NELS$achmat08[NELS$hsprog == "Some Vocational"])**

**skew.ratio(NELS$achmat08[NELS$hsprog == "Some Vocational"])**

**skew(NELS$achmat08[NELS$hsprog == "Other"])**

**se.skew(NELS$achmat08[NELS$hsprog == "Other"])**

**skew.ratio(NELS$achmat08[NELS$hsprog == "Other"])**

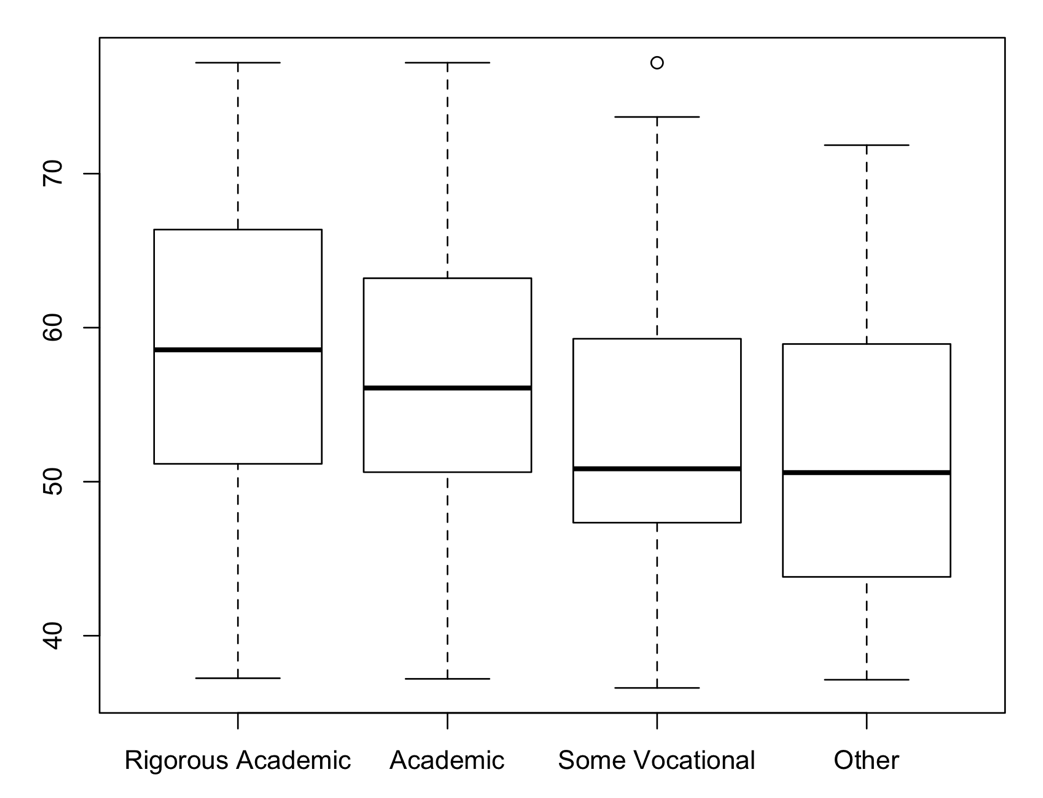
**levenes.test(NELS$achmat08, NELS$hsprog)**

**results = aov(NELS$achmat08 ~ NELS$hsprog)**

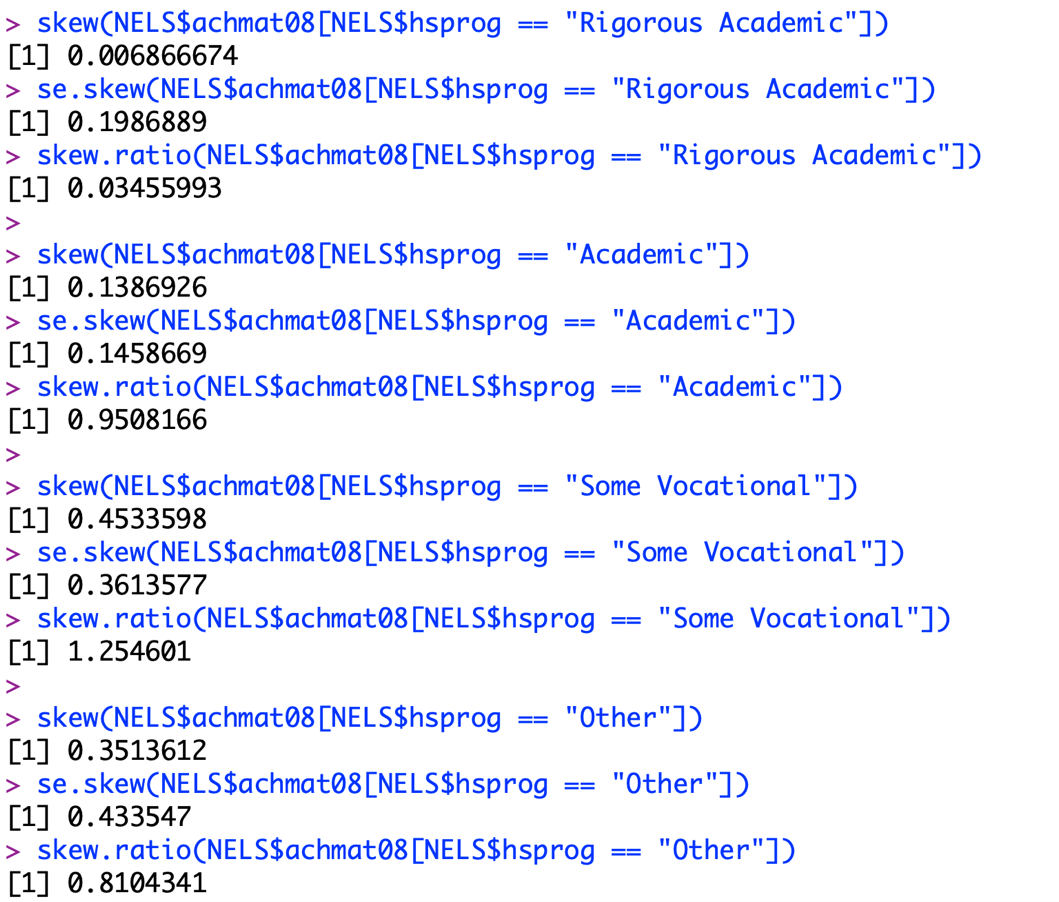
**summary(results)**

**PostHocTest(results, method = "hsd")**

1. One-way ANOVA is used to compare unrelated means of an at least interval leveled variable for two or more independent groups. In this case, we are comparing the population mean math achievement for *four* different groups based on type of high school program. Assuming that the normality and homogeneity of variances assumptions are tenable or that the test is robust to their violations (evaluated in later parts of this exercise), one-way ANOVA is an appropriate inferential test.
2. Based on the boxplot, the normality assumption appears to be tenable for all but possibly the some vocational program, because the distributions are fairly symmetric. The homogeneity of variance assumption appears tenable as well because the interquartile ranges are similar. Finally, the results of the ANOVA are likely to mirror the pattern of the medians: the median eighth grade math achievement for rigorous academic is the highest, followed by academic. The respective medians for some vocational and other program types appear to be approximately the same and lower than the other two program types.

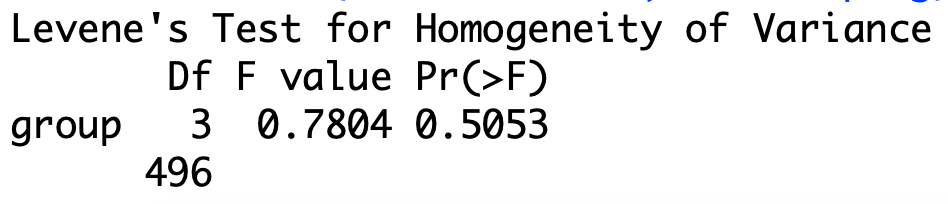


1. A test of the skewness must be performed to evaluate the normality assumption because not all of the groups have at least 30 students. The results appear below.

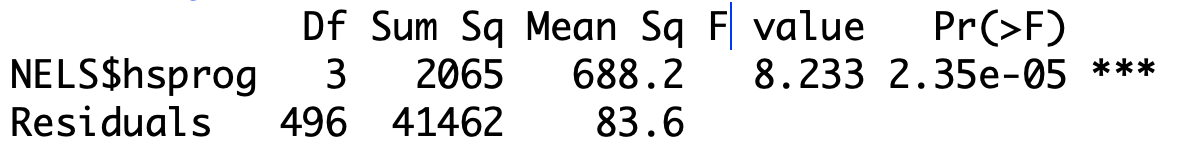
****

For all types of high school programs, the skewness is less than 1 in magnitude and the skewness ratio is less than 2 in magnitude, so we conclude that the normality assumption is tenable.

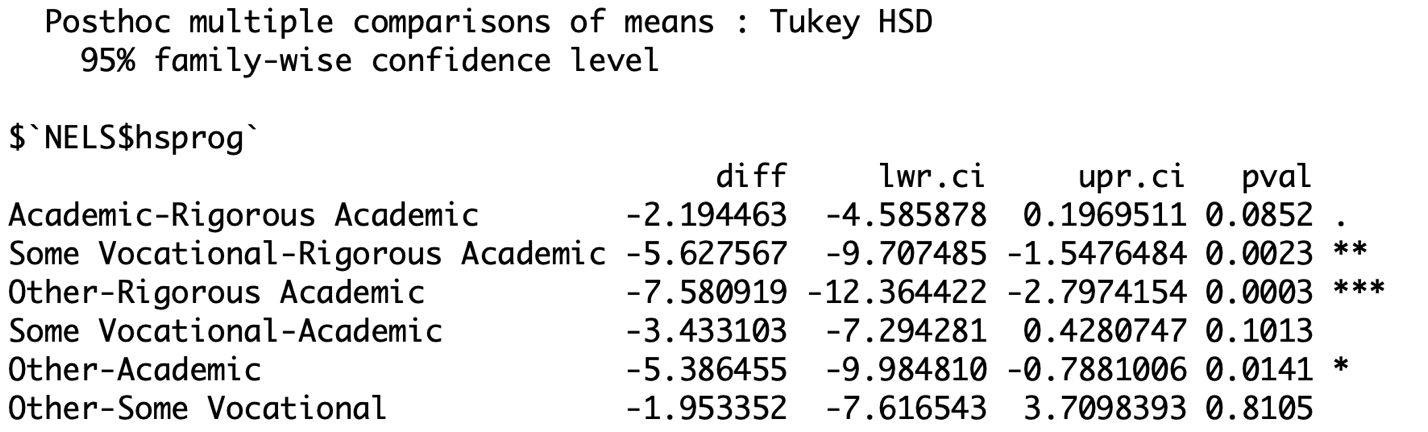
1. Levene’s test must be performed to test the homogeneity of variance assumption because the sample sizes are not equal for the four groups. The results of Levene’s test indicate that the homogeneity of variance assumption is tenable, *F*(3, 496) = .78, *p* = .51.

****

1. H0: *RA**A***SV**Oand H1: Not H0
2. *p = 2.35e-05*.



1. We reject the null hypothesis in favor of the alternative.
2. According to the results of the one-way analysis of variance, the average math achievement in eighth grade of college-bound students who have always been at grade level does vary as a function of their type of high school program, *F*(3, 496) = 8.23, *p* < .0005.
3. Approximately 4.74 percent of the variance in eighth grade math achievement can be explained by type of high school program, a relatively small effect, according to Cohen’s rule of thumb guidelines.
4. Post-hoc tests are necessary, in this case, because the results of the ANOVA are statistically significant and more than two independent means are being compared.

****

According to the results of the Tukey Test, the students in the rigorous academic program outperform those in the vocational and those in the other programs, but not those in the academic programs. The students in the academic program perform significantly better than those in the other programs, but not those in the vocational programs.

* 1. The R commands for generating the output for these analyses are:

**boxplot(NELS$slfcnc08 ~ NELS$urban)**

**table(NELS$urban)**

**levenes.test(NELS$slfcnc08, NELS$urban)**

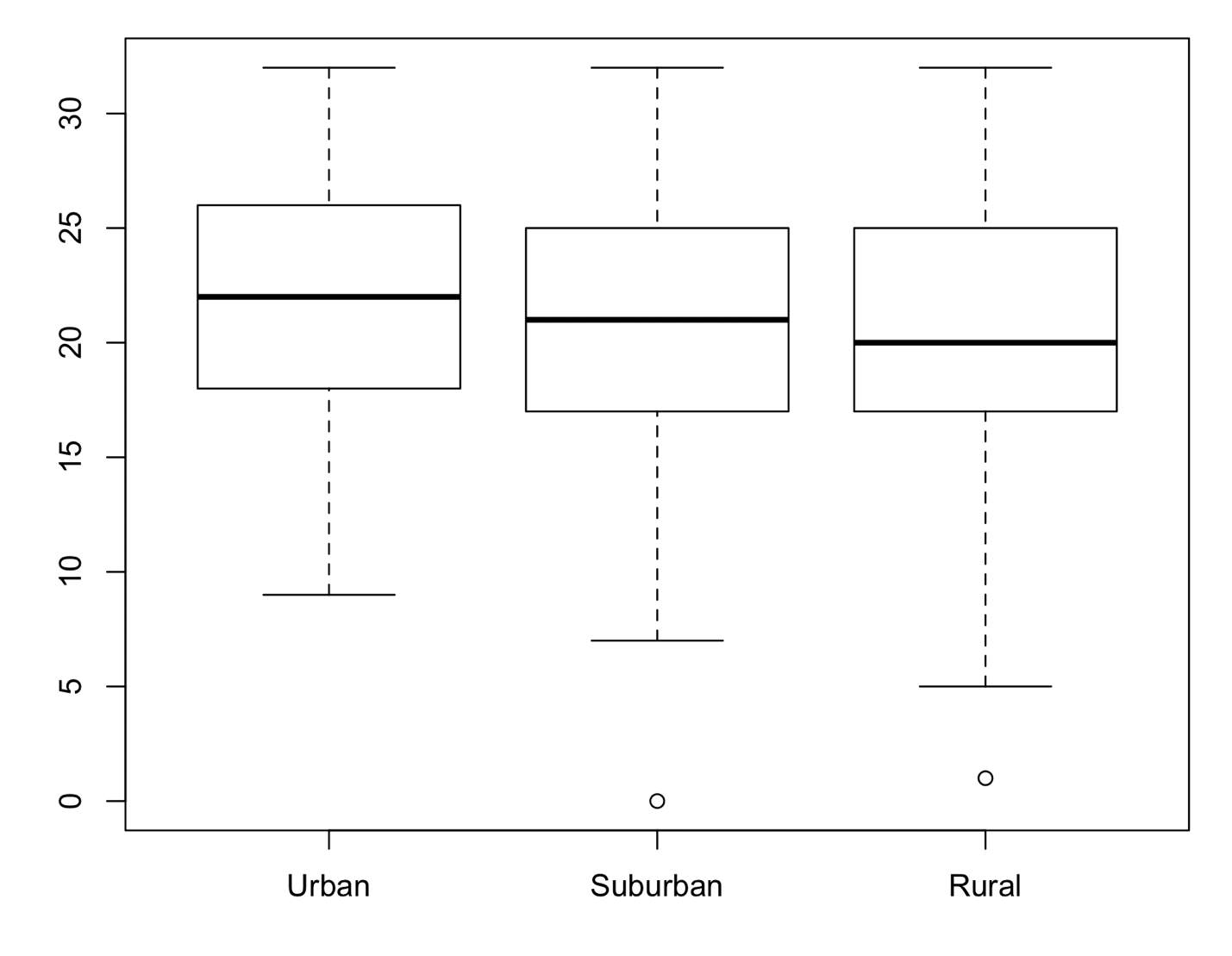
**results = aov(NELS$slfcnc08 ~ NELS$urban)**

**summary(results)**

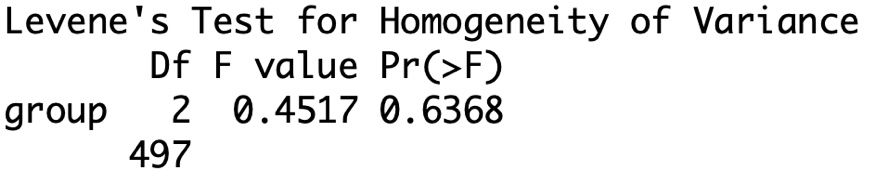
**PostHocTest(results, method = "hsd")**

**PostHocTest(results, method = "bonferroni")**

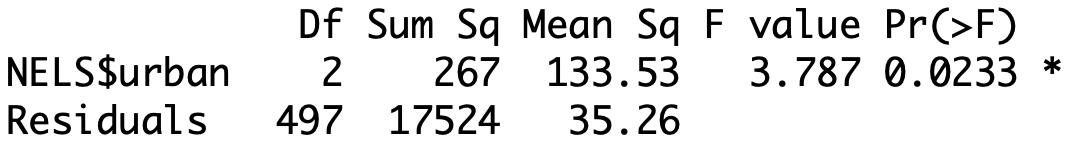
1. In this case, we perform a one-way ANOVA over an independent samples *t*-test because there are three independent levels of urbanicity.
2. Based on the boxplot, the normality assumption appears to be tenable for those from an urban environment. Although those from a suburban and rural environment have outliers, the sample size in each of these groups is large enough to prevent these outliers from having an impact on the validity of the test results. The homogeneity of variance assumption appears tenable as well because the interquartile ranges appear similar. Finally, the results of the ANOVA are likely to indicate that there are no statistically significant differences between the group means because the medians are all similar.



1. There are 123 students in the urban group, 215 in the suburban group, and 162 students in the rural group. Because all of these sample sizes are larger than 30, the ANOVA is robust to violations of the normality assumption, if it is violated. Levene’s test must be performed to test the homogeneity of variances assumption because the sample sizes are not equal for the three types of environments. The results of Levene’s test indicate that the homogeneity of variances assumption is tenable, *F*(2, 497) = .45, *p* = .64.

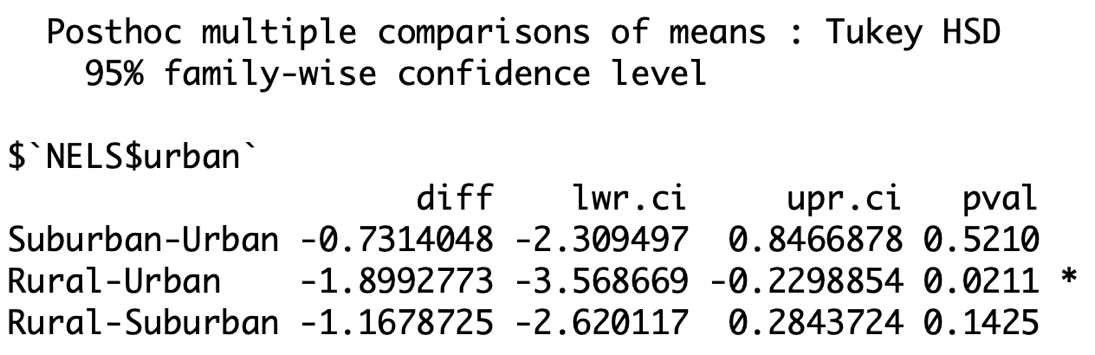


1. The results of the one-way ANOVA are summarized in the following table.

****

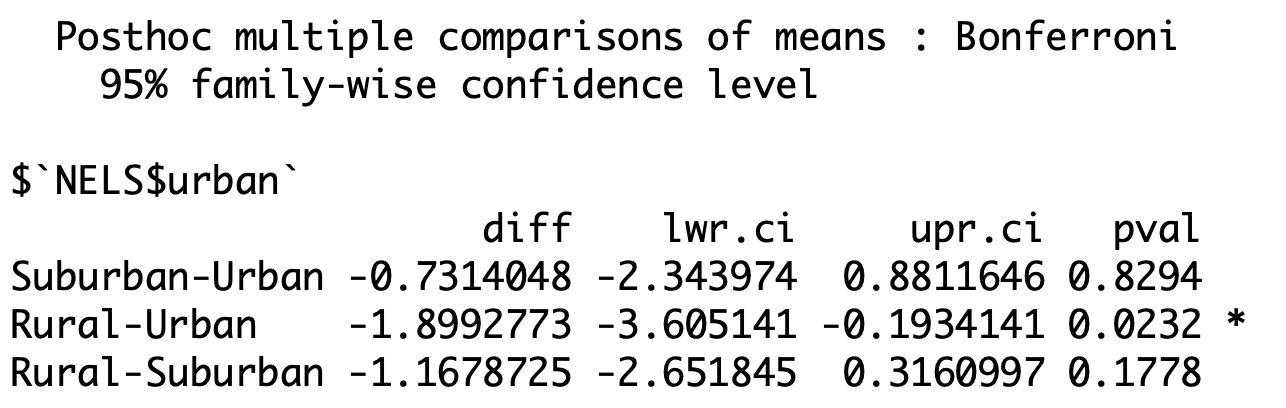
According to the results of the one-way analysis of variance, the average self-concept score in eighth grade of college-bound students who have always been at grade level does vary as a function of the type of setting in which these are from, *F*(2, 497) = 3.79, *p* = .02.

1. Approximately 1.5 percent of the variance in eighth grade self-concept is explained by urbanicity, a small effect.
2. Post-hoc tests are necessary, in this case, because the results of the ANOVA are statistically significant and more than two independent means are being compared. The following table provides the summary statistics and the results of the Tukey post-hoc test.



According to the Tukey post-hoc test, the students from urban environments have statistically significantly higher eighth grade self-concept scores than the students from rural environments. No other statistically significant differences are detected between the different environments.

1. According to the Bonferroni adjusted post-hoc tests, the students from the Urban environments have significantly higher eight grade self-concept scores than the students from the rural environments. No other statistically significant differences are detected between the different environments.



* 1. The R commands for generating the output for these analyses are:

**boxplot(NELS$achmat12 ~ NELS$cigarett)**

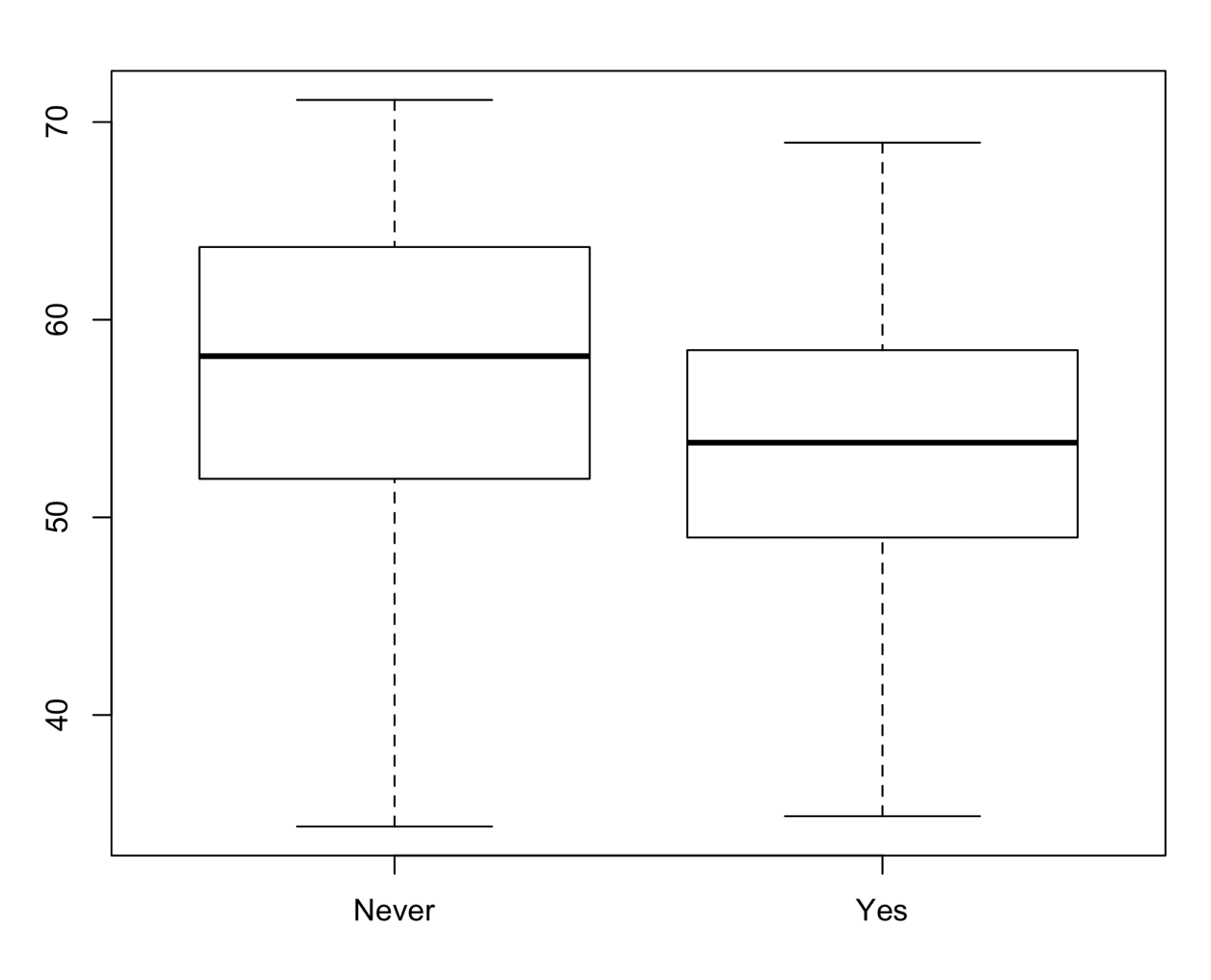
**levenes.test(NELS$achmat12, NELS$cigarett)**

**results = aov(NELS$achmat12 ~ NELS$cigarett)**

**summary(results)**

**t.test(NELS$achmat12[NELS$cigarett == "Never"], NELS$achmat12[NELS$cigarett == "Yes"], var.equal = T)**

1. Based on the boxplot, the normality assumption appears to be tenable because both distributions are reasonably symmetric. The homogeneity of variance assumption appears tenable as well because the interquartile ranges are similar. Finally, based on the medians of the two groups, the results of the ANOVA are likely to indicate that the twelfth grade math achievement of those who never smoked cigarettes is higher.



1. The sample sizes are large enough (*n* = 429 for those who have never smoked cigarettes and *n* = 71 for those who tried cigarettes at least once) to suggest that any deviations from normality that occur in the population for these two groups will not compromise the validity of the results of the analysis of variance. The results of Levene’s test indicate that the homogeneity of variance assumption is tenable, *F*(1, 498) = .001, *p* = .9796.
2. According to the results of the one-way analysis of variance, the average twelfth grade math achievement score of college-bound students who have always been at grade level does differ depending on whether or not they smoked cigarettes in high school, *F*(1, 498) = 18.53, *p* = 2.01e-05.
3. According to the ratio of the sums of squares, approximately 3.59 percent of the variance in twelfth grade math achievement scores can be explained by cigarette use, a relatively small effect.
4. It is not necessary to perform post-hoc testing in this case. Because there are only two groups, we may use the sample means to determine the nature of the mean difference. Students who never smoked cigarettes (*M* = 57.51, *SD* = 7.71) performed statistically significantly better, on average than those who did (*M* = 53.24, *SD* = 8.00).
5. According to the results of the independent samples *t*-test, students who never smoked cigarettes performed statistically significantly better, on average, than those who did, *t*(498) = 4.31, *p* = 2.01e-05. As expected, when there are only two groups, the results of the one-way ANOVA are consistent with those of the independent samples *t*-test.
6. Yes, but the *p*-value from the ANOVA must be divided in half because the one associated with the ANOVA is two-tailed and the question is one-tailed. The *p*-value associated with the one-tailed question is *p* = 1.005e-05.
   1. The R commands for generating the output for these analyses are:

**boxplot(NELS$unitmath ~ NELS$edexpect)**

**table(NELS$edexpect)**

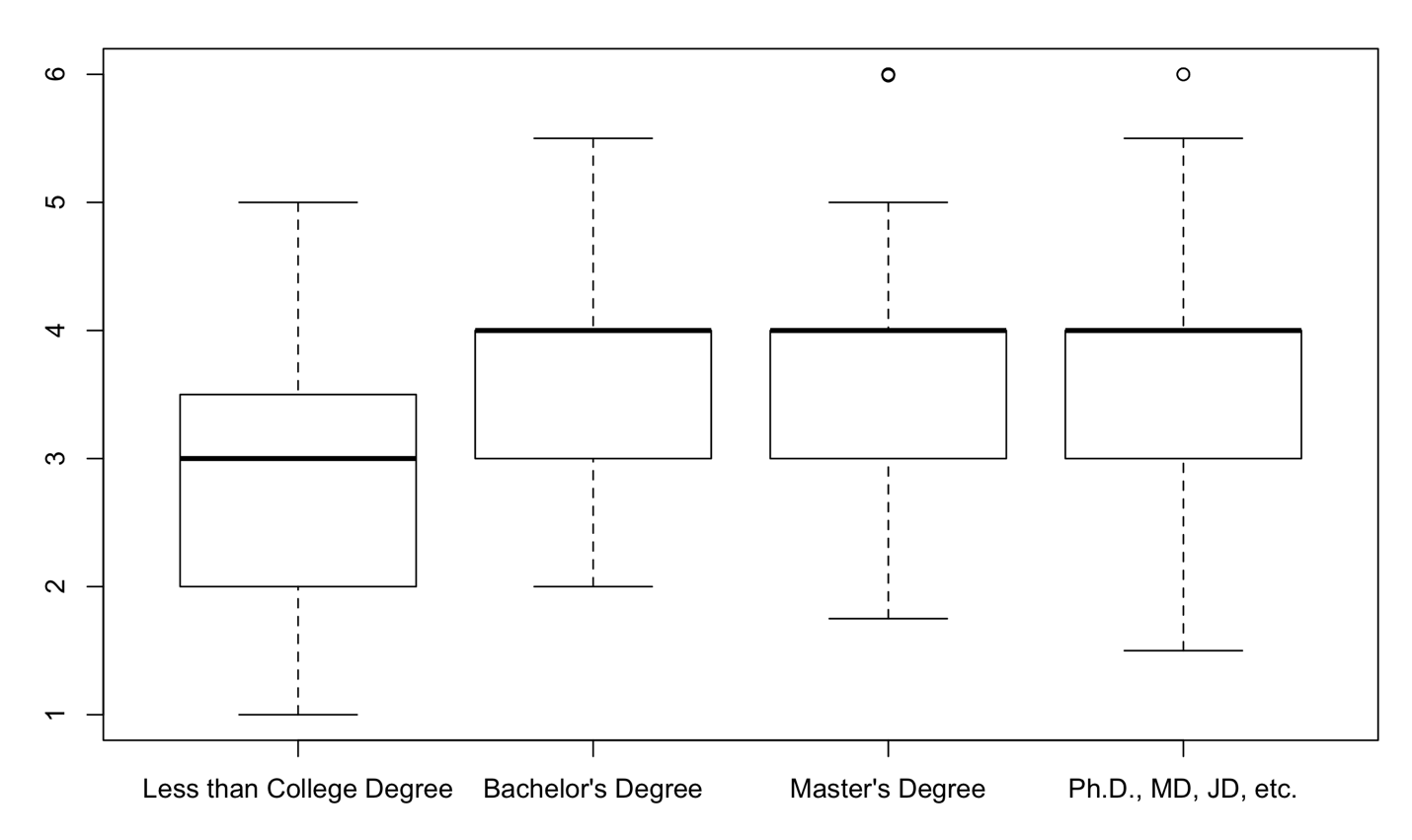
**levenes.test(NELS$unitmath, NELS$edexpect)**

**results = aov(NELS$unitmath ~ NELS$edexpect)**

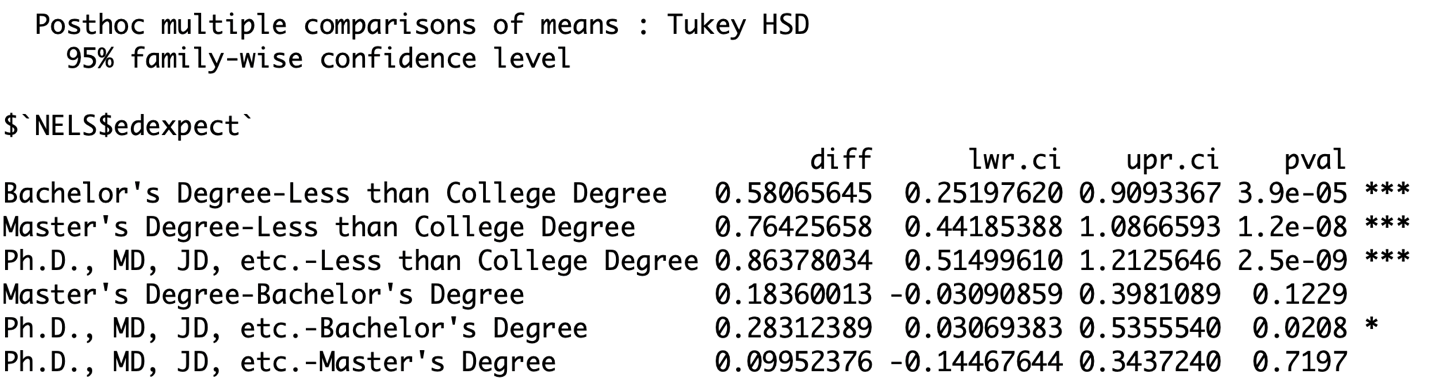
**summary(results)**

**PostHocTest(results, method = "hsd")**

1. Based on the boxplot, the normality assumption appears to be tenable for those who expect to attain a Bachelor’s degree or less. Although the distributions for those who expect to attain more than a Bachelor’s degree have several outliers each, the samples are large enough in each of these distributions that these outliers are unlikely to impact the validity of the test results. The homogeneity of variance assumption appears tenable as well because the interquartile ranges are similar. Finally, based on the medians, the results of the ANOVA are likely to indicate that those who expect to attain less than a college degree take statistically significantly fewer units of math than the other groups.



1. There are 48 students who expect to attain less than a college degree, 159 expect a Bachelor’s degree, 190 a Master’s degree, and 103 a terminal degree. Because all of these sample sizes are larger than 30, the ANOVA is robust to violations of the normality assumption, if it is violated. Levene’s test must be performed to test the homogeneity of variance assumption because the sample sizes are not equal for the four expected levels of education. The results of Levene’s test indicate that the homogeneity of variance assumption is tenable, *F*(3, 496) = .37, *p* = .77.
2. According to the results of the one-way analysis of variance, the average number of years of mathematics taken by college-bound students who have always been at grade level does vary as a function of their expected educational attainment, *F*(3, 496) = 15.70, *p* = 9.1e-10.
3. Approximately 8.67 percent of the variance in units of math taken can be explained by the highest degree the student anticipates earning, a moderate effect.
4. Post-hoc tests are necessary, in this case, because the results of the ANOVA are statistically significant and more than two independent means are being compared. The following tables provide the summary statistics and the results of the two post-hoc tests.

****

According to the Tukey post-hoc test, students who anticipate earning less than a college degree take statistically significantly fewer units of math than any other degree type. Also, students who anticipate earning a BA take statistically significantly fewer units of math than those in the PhD group. No other statistically significant differences are detected.

* 1. The R commands for generating the output for these analyses are

**boxplot(NELS$schattrt ~ NELS$cigarett)**

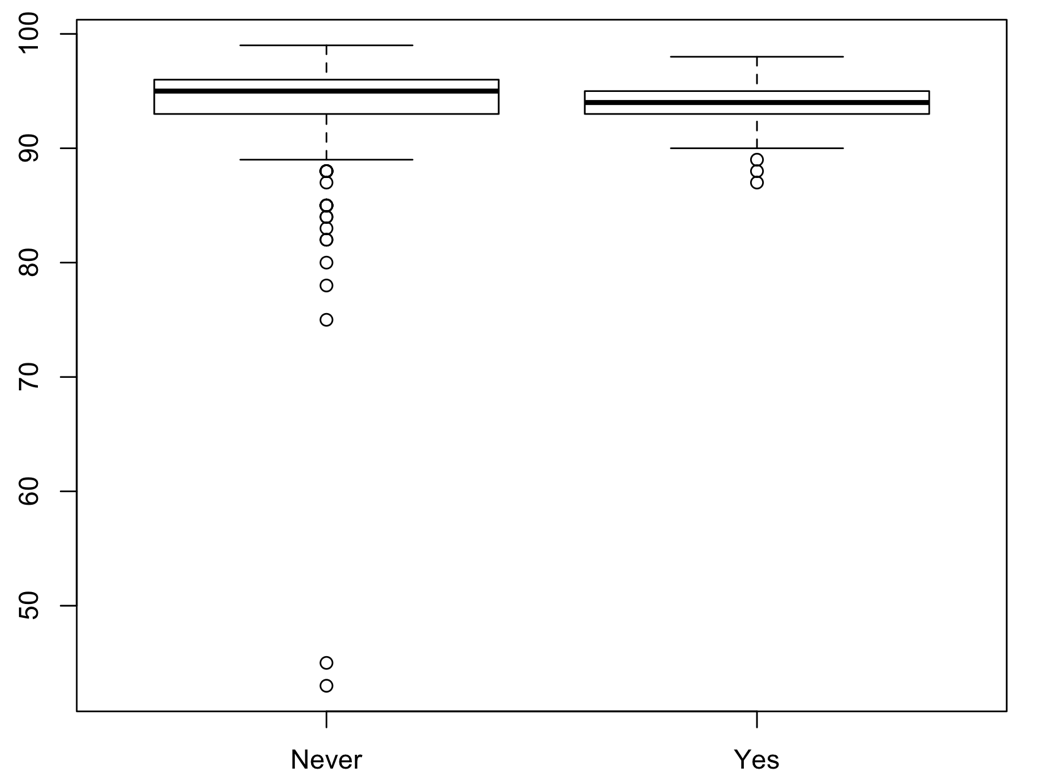
**table(NELS$cigarett)**

**levenes.test(NELS$schattrt, NELS$cigarett)**

**results = aov(NELS$schattrt ~ NELS$cigarett)**

**summary(results)**

1. Based on the boxplot, the normality assumption does not appear to be tenable. However, the samples are large enough that these outliers are unlikely to impact the validity of the test results. The homogeneity of variance assumption appears tenable as well because the interquartile ranges are similar. Finally, based on the medians, the results of the ANOVA are likely to indicate that those who never smoked cigarettes attended schools with a similar school attendance rate than those who smoked cigarettes at least once.



1. There are 359 students that reported that they have never smoked cigarettes while 58 reported that they had at least once. Because both of these sample sizes are larger than 30, the ANOVA is robust to any violation of the normality assumption. Levene’s test must be performed to test the homogeneity of variance assumption because the sample sizes are not equal for the two groups. The results of Levene’s test indicate that the homogeneity of variance assumption is tenable, *F*(1, 415) = 2.30, *p* = .13.
2. According to the results of the one-way analysis of variance, among college-bound students who have always been at grade level, the average school attendance rate of those who ever smoked (*M* = 94.12, *SD* = 2.43) does not differ statistically significantly from that of those who never smoked (*M* = 93.57, *SD* = 4.96), *F*(1, 415) = .68, *p* = .41.
3. Post-hoc tests are not necessary, in this case, because the results of the ANOVA are not statistically significant, and cigarett only has two levels.
4. No. Because the *p-*value at .41/2 = .21 still exceeds .05.

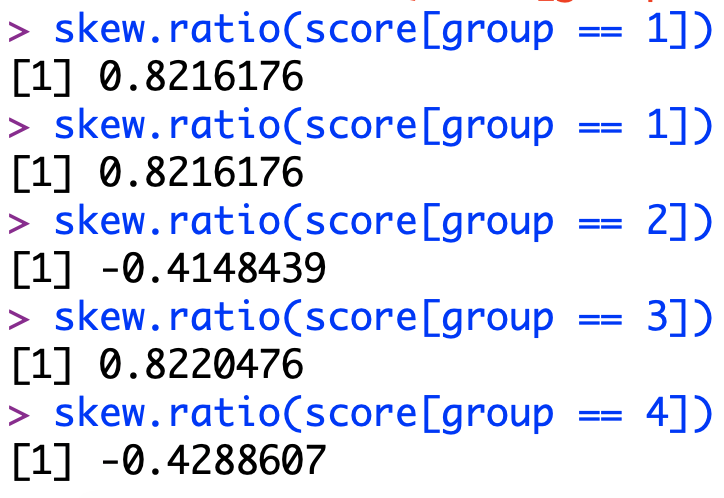
* 1. a) 3 b) 2 c) 6 d) 7 e) 5

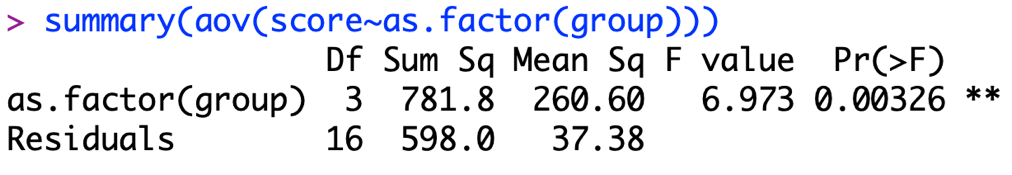
ANOVA

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 2,240 | 2 | 1120 | 5 | .001 < *p* < .01 |
| Within Groups | 26,208 | 117 | 224 |  |  |
| Total | 28448 | 119 |  |  |  |

The results of the one-way analysis of variance indicate that there is at least one statistically significant difference in serotonin levels among bulimics, recovering bulimics, and the control group, *F*(2, 117) = 5, *p* = .01.

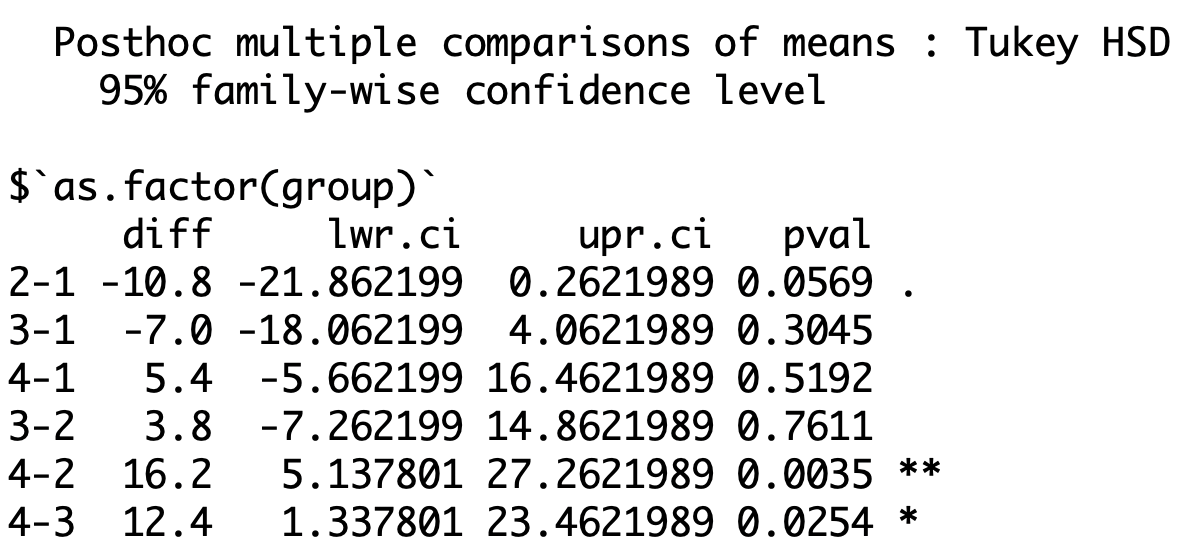
1. The one-way ANOVA can be performed by hand. Alternatively, it can be performed using R. Using R, we obtain the following output:





None of the distributions are significantly skewed and the results of Levene’s test indicate that the homogeneity of variance assumption is tenable, in this case, *F*(3, 16) = .31, *p* = .82. The results of the one-way ANOVA indicate that the mean score in the population is not the same across all four groups, *F*(3, 16) = 6.97, *p* = .003.

1. A post-hoc analysis should be conducted because the results of the ANOVA are statistically significant and there are more than two groups. These can be performed by hand. Alternatively, they can be performed using R. Using R, we obtain the following output for the Tukey test:

****

The results of the Tukey test indicate that, on average, students in Group 4 score statistically significantly higher than students in both Group 2 and Group 3. There were no other statistically significant differences detected by the Tukey test.

* 1. Using the R command **pwr.anova.test(k = 3, f = 0.4, power = 0.8, sig.level = 0.05)**, we see that the appropriate sample size is *n* = 22 per group, for a total *N* = 66.