**SPSS *t*-tests/ANOVA**

**Due at the Start of Lab:** Paper 2

**Rationale for Today's Lab Session**

This tutorial is designed to ensure that you have a basic understanding of *t*-tests and Analysis of Variance (ANOVA). You will need these skills for Lab Assignment 5. These skills are also essential for academic and employment pursuits in research. Today, you will go through this tutorial with your lab instructor. You can work collaboratively on this tutorial but must work independently on the graded lab assignment.

**Instructions**

Warning
SPSS periodically changes the visual display and organization of menus. The instructions presented in this tutorial may need to be augmented marginally depending on the version of SPSS you are using. If you get stuck, use Google, or ask the lab instructor for help.

Accessing SPSS
Once you log on, go to the Start menu in the lower left corner of the screen and find SPSS. If you have difficulty finding it, ask the lab assistant for help.

Data File
For this tutorial, you will use Data Set F (see Canvas → Data Files). The data set includes 509 participants who were Tulane students and their friends and family. Download the data file (DataF.sav) and the "data dictionary" that provides more detail on the variables that were included in the survey (DataF_Dictionary.xls). The files should open in SPSS and Excel, respectively. Double-click on them to open them, or open the programs and use the file menus to locate and open these files.

**_t_-tests: The Basics**

For this section, you will learn to run a between-group *t*-test, the most commonly used type of *t*-test.

Background Information. The between-group *t*-test is used when you have two groups or categories of people. It lets you see how the groups differ in terms of their scores on some continuous variable (a variable with an ordered rating system, like 0-10 scales, age, etc.). The *t*-test is basically a means-to-an-end. It provides us with a *p*-value, the probability that a result is due to "chance" or "sampling error". The *p*-value shows up in the Output and tells us whether the result is likely due to chance. If $p < .05$, the difference between groups is reliable. If not, there is no reliable difference, and we tend to ignore the result.

Running a *t*-test.  Go to the Analyze menu, point to Compare Means, and choose "Independent-Samples T Test"
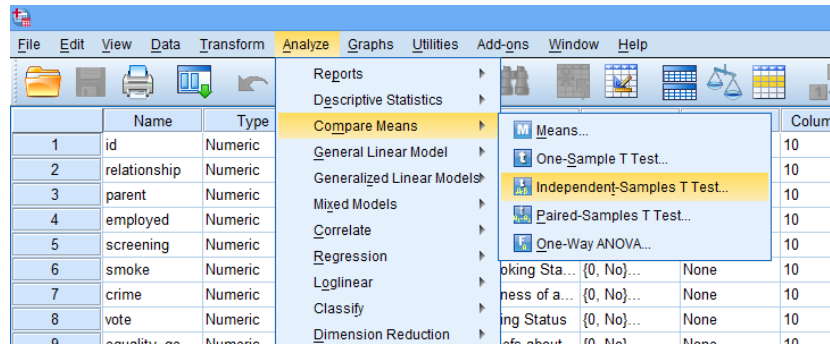


In the window that pops up, we always put the independent variable (grouping or categorical variable) in the "Grouping Variable" section of the box.  In the "Test Variable(s)" box, put any continuous dependent variables you want to examine (you can choose more than one if you like).  The analysis will tell us if the groups differ in terms of their scores on the "Test Variables".

Try putting Smoking (#6) in the "Grouping Variable" area, and put Protest Enjoyment (#55) and Openness to Criticism (#80) in the "Test Variables" section, so we can see if smokers differ on these variables.  At this point you will notice that the OK button is still gray, so we need to do one more step.

Single-click where it says "smoke(? ?)" in the Grouping Variables area.  Then, click on the Define Groups button.  SPSS needs you to tell it which numbers were used to describe the groups.  In the data file, we arbitrarily coded nonsmoker = 0 and smoker = 1, so type a 0 where is says "Group 1" and a 1 where it says "Group 2".  If you ever need to examine how variables were coded, you could use the Variable View option in SPSS or simply look in our Excel Data Guide file.
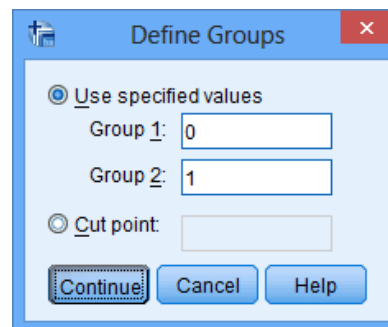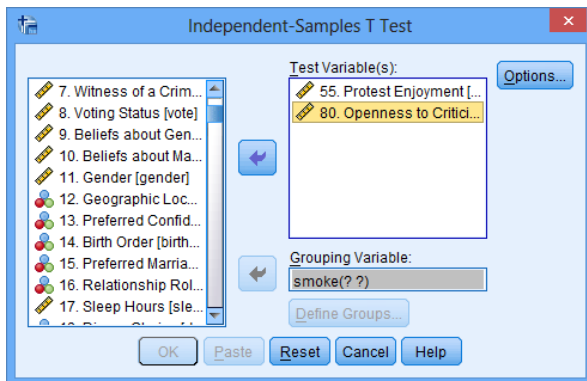


Click the Continue button, and then the OK button to run the analyses.  Your Output should look something like this:

## Group Statistics

| | 6. Smoking Status | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| 55. Protest Enjoyment | No | 464 | .13 | 2.933 | .136 |
| | Yes | 45 | .20 | 3.027 | .451 |
| 80. Openness to Criticism | No | 464 | 2.91 | 1.798 | .083 |
| | Yes | 45 | 3.64 | 1.228 | .183 |

## Independent Samples Test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | 95% Confidence Interval of the Difference | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| 55. Protest Enjoyment | Equal variances assumed | .001 | .976 | -.145 | 507 | .885 | -.066 | .459 | -.968 | .836 |
| | Equal variances not assumed | | | -.141 | 52.333 | .889 | -.066 | .471 | -1.012 | .879 |
| 80. Openness to Criticism | Equal variances assumed | 1.478 | .225 | -2.681 | 507 | .008 | -.735 | .274 | -1.273 | -.196 |
| | Equal variances not assumed | | | -3.654 | 63.940 | .001 | -.735 | .201 | -1.137 | -.333 |

Using the top box, we see that smokers reported a marginally higher level of protest enjoyment (*M* = 0.20, *SD* = 3.03) than did non-smokers (*M* = 0.13, *SD* = 2.93), using a -5 to +5 rating scale. The second box in the Output tells us whether the difference we observed in our sample was reliable. Based on the size of the difference and the number of people in our study, would we expect this difference to hold up for the population in general, or is it just a chance finding? To make this determination, we would need a *p*-value.

We use the second box to obtain the *p*-value. Every statistical test has "assumptions," or requirements, which must be met, or the statistic will produce a biased (potentially misleading or incorrect) result. An important assumption for the *t*-test is that both groups have comparable variability on the dependent variable. Levene's test for equality of variances checks that assumption, and produces a *p*-value (see blue circle). If $p \geq .05$, there is no violation of the assumption, and we can use the results from the standard t-test (see solid red oval). If the assumption is violated (Levene's test $p < .05$), we have to use a corrected *t*-test (see dashed red oval) -- more on this later. In each of the above examples, the Levene's test shows no violation of the assumptions (p values of .976 and .225, respectively), so we can use the standard *t*-tests.

For protest enjoyment, the t-test area of the Output (solid red circle) tells us the *t*-value (-.145), the degrees of freedom (a reference number, 507). If we weren't using a computer, we would compare this observed *t*-value to a critical *t*-value or cut score to determine whether the result is significant ($p < .05$). However, SPSS gives us the exact *p*-value, or the exact probability of obtaining this result by chance (.885). Basically, we'd expect to find this type of weak group difference just by chance alone 89% of the time. Because the *p*-value is larger than .05, the result is not statistically significant. The observed difference in our sample is not a reliable or

3

trustworthy effect.  At the population level, we would not typically expect smokers and non-smokers to differ in terms of protest enjoyment.

In contrast, smokers (*M* = 3.64, *SD* = 1.23) scored higher than non-smokers (*M* = 2.91, *SD* = 1.80) on openness to criticism.  The *t*-value (-2.681), degrees of freedom (507) and *p*-value (.008) are noted.  For this analysis, *p* < .05, so the observed difference is trustworthy.  At the population level, we would expect smokers to be more open to criticism than non-smokers.

### *t*-tests: Violation of the Assumption of Equality of Variances

In the above examples, we were free to use the standard *t*-test because there were no violations of assumptions for that test. Next, we will go through an example where the assumption is violated.

Run a *t*-test examining whether Gender identity (#11, 0 = female, 1 = male) is associated with Support for Adoption Equality (#72, a scale from -5 to +5, where higher scores indicate greater support for the right of same-sex couples to adopt). The Ouput should look like the following:

**Group Statistics**

| | 11. Gender | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| 72. Support for Adoption Equality | Female | 415 | 4.34 | 1.908 | .094 |
| | Male | 94 | 3.73 | 2.239 | .231 |

**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | 95% Confidence Interval of the Difference | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| 72. Support for Adoption Equality | Equal variances assumed | 9.555 | .002 | 2.677 | 507 | .008 | .603 | .225 | .161 | 1.046 |
| | Equal variances not assumed | | | 2.421 | 125.331 | .017 | .603 | .249 | .110 | 1.097 |

Looking at the first box in the Output (red box), we see that women supported somewhat greater preferences for adoption equality than did men. We also see that the spread of responses (standard deviation) was slightly higher for men than women. The second box of the Output helps us to interpret whether there were reliable gender differences, or whether these group differences were due to chance.

The *t*-test assumes that both groups should be comparable in terms of their variability (standard deviation, variance, etc.) on the dependent variable (Support for Adoption Equality). Levene's test suggests that this assumption is not met here because the *p*-value for the Levene's test is < .05 (Sig. value in the blue circle, *p* = .002). This is consistent with our anecdotal observation fro the first box in the Output that the spread of responses (standard deviation) was slightly higher for men than women (2.239 vs. 1.908); the Levene's test merely tells us this is a statistically significant difference, so we need to use a modified version of the *t*-test (dashed red oval).

Looking at that section (dashed red circle), the *t*-value (2.421), degrees of freedom (125.331) and *p*-value (.017) are noted.  For this analysis, *p* < .05, so the observed difference is trustworthy.  At the population level, we would expect women to be more supportive of adoption equality than men.

Practice Questions

1) Conduct a *t*-test examining whether Parental Status (#3) is related to Financial Distress (#59). How would you report the result in APA style, using the APA Style Guide at the end of this assignment? Cohen's *d* can be calculated by using this calculator: http://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html
2) Conduct a *t*-test examining whether Relationship Status (#2) is related to Insomnia (#40). How would you report the result in APA style, using the APA Style Guide at the end of this assignment?
3) Conduct a *t*-test examining whether having been a Witness to a Crime (#7) is associated with Anxiety (#24). How would you report the result in APA style, using the APA Style Guide at the end of this assignment?
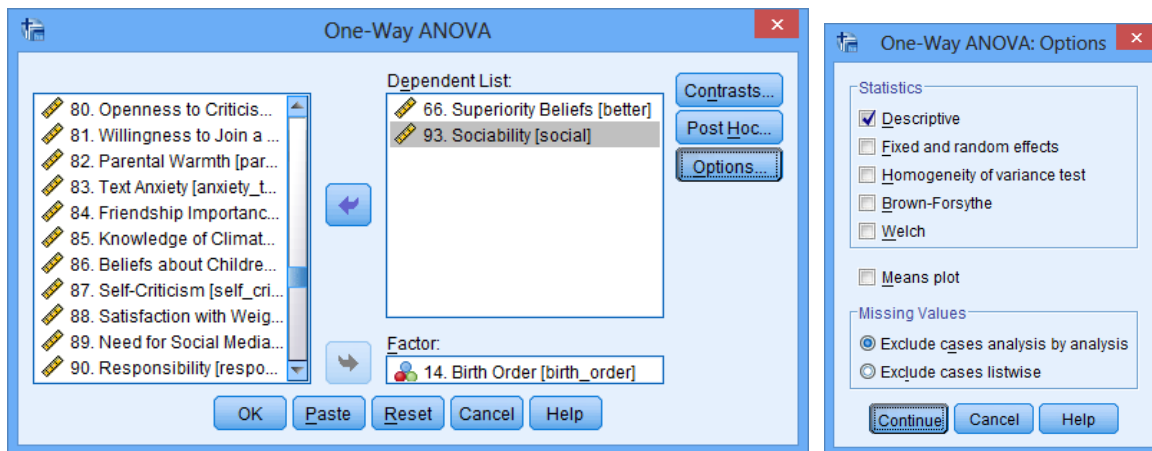
## ANOVA

ANOVA is similar to the between-group *t*-test. It is used when there are more than two categories or groups of people. Again, these groups are compared in terms of their scores on some continuous variable.

Go to the Analyze menu, point to Compare Means, and choose One-Way ANOVA.
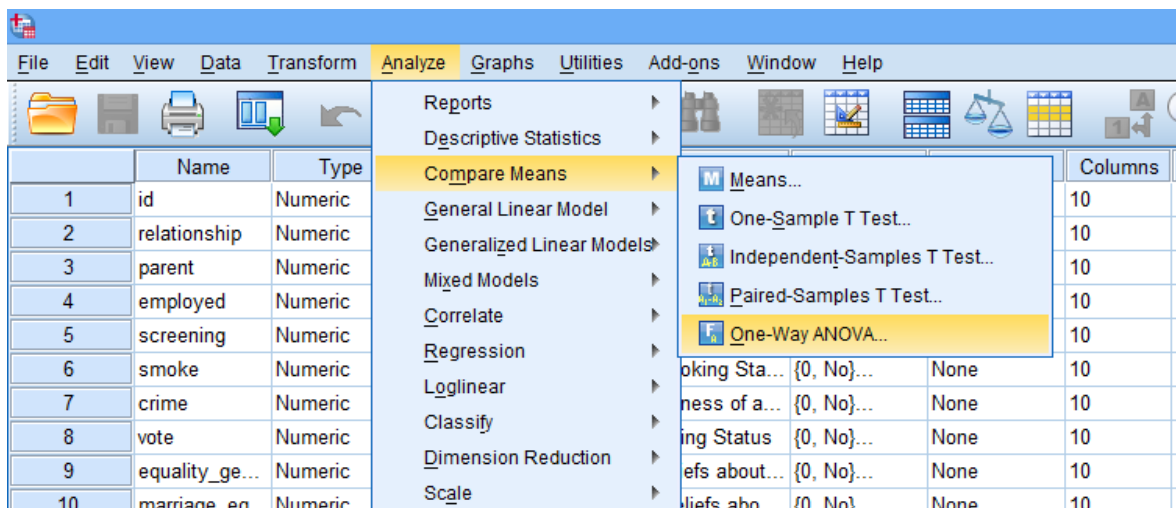


In the window that pops up, you put the categorical variable in the Factor area and the continuous dependent variable in the Dependent List area (feel free to include more than one). For practice, put Birth Order (#14) in the Factor area, and put Superiority Beliefs (#66) and Sociability (93) in the Dependent List area. Before clicking OK, click on the Options button. Choose Descriptive in the pop-up window, telling SPSS to include basic descriptive statistics with our Output. Click Continue and then OK.

The Output should look something like this:

**Descriptives**

| | | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean Lower Bound | 95% Confidence Interval for Mean Upper Bound | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|---|
| 66. Superiority Beliefs | Only child | 51 | -2.49 | 2.745 | .384 | -3.26 | -1.72 | -5 | 5 |
| | Youngest | 168 | -2.42 | 2.865 | .221 | -2.86 | -1.99 | -5 | 5 |
| | In the middle | 110 | -2.38 | 2.952 | .281 | -2.94 | -1.82 | -5 | 5 |
| | Oldest | 180 | -2.47 | 2.888 | .215 | -2.89 | -2.04 | -5 | 5 |
| | Total | 509 | -2.44 | 2.872 | .127 | -2.69 | -2.19 | -5 | 5 |
| 93. Sociability | Only child | 51 | 5.84 | 2.485 | .348 | 5.14 | 6.54 | 0 | 10 |
| | Youngest | 168 | 7.20 | 2.089 | .161 | 6.88 | 7.52 | 2 | 10 |
| | In the middle | 110 | 6.39 | 2.563 | .244 | 5.91 | 6.88 | 0 | 10 |
| | Oldest | 180 | 6.82 | 2.394 | .178 | 6.47 | 7.17 | 0 | 10 |
| | Total | 509 | 6.76 | 2.378 | .105 | 6.55 | 6.96 | 0 | 10 |

**ANOVA**

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 66. Superiority Beliefs | Between Groups | .672 | 3 | .224 | .027 | .994 |
| | Within Groups | 4190.503 | 505 | 8.298 | | |
| | Total | 4191.175 | 508 | | | |
| 93. Sociability | Between Groups | 91.426 | 3 | 30.475 | 5.535 | .001 |
| | Within Groups | 2780.366 | 505 | 5.506 | | |
| | Total | 2871.792 | 508 | | | |

The top box is very detailed, so examine it slowly. The box tells how many people were in each group, and provides their basic descriptive statistics. For example, 51 people reported being an only child, and as a group, they had the lowest score for superiority beliefs ($M = -2.49$, $SD = 2.75$) on a -5 to +5 scale.

The second box in the SPSS Output provides the results of the ANOVA analysis, which are used to indicate whether there were any reliable differences across the groups of people. The $F$-value and degrees of freedom values are provided. If you are unfamiliar with these statistics from prior

statistics courses, just be aware that they are used in the process of calculating a *p*-value. The *p*-value is .99. Because *p* is larger than .05 (and by a lot), there are no reliable differences for superiority beliefs.

However, there was a reliable group difference for sociability ($p$ = .001), meaning that at least one of the groups differs from the others in terms of their sociability.

## ANOVA Post-hoc Tests

The preceding ANOVA was simple to conduct, but it only provides limited information. The significantly *p*-value in the Sociability example merely indicates that at least one of the groups reliably differed from another group. It could be that all four of the groups reliably differed from each other. It could be that only two differed from each other. Alternatively, one group might reliably differ from two other groups but not reliably differ from another group.

Look back at the first block of Output. Specifically, look at the mean Sociability scores for the different groups (Only Child, Youngest, Middle, Oldest). Do you think that Only Children are reliably less sociable than Youngest Children? Probably, the difference is pretty big. Do you think Middle Children reliably differ from Oldest Children on sociability? That's more difficult to say, as the difference in means looks a lot smaller. The statistically significant ANOVA simply says that at least two groups differ – it is a signal that we need to conduct a follow-up test, called a "post-hoc test" to determine more specifically which groups reliably differ.

Post-hoc tests. We only conduct post-hoc tests if the ANOVA is statistically significant. Bring up the ANOVA window again (Analyze → Compare Means → One-Way ANOVA). The ANOVA involving Superiority Beliefs (#66) was non-significant, so select it and click on the arrow in the middle of the box to remove it from the analyses.
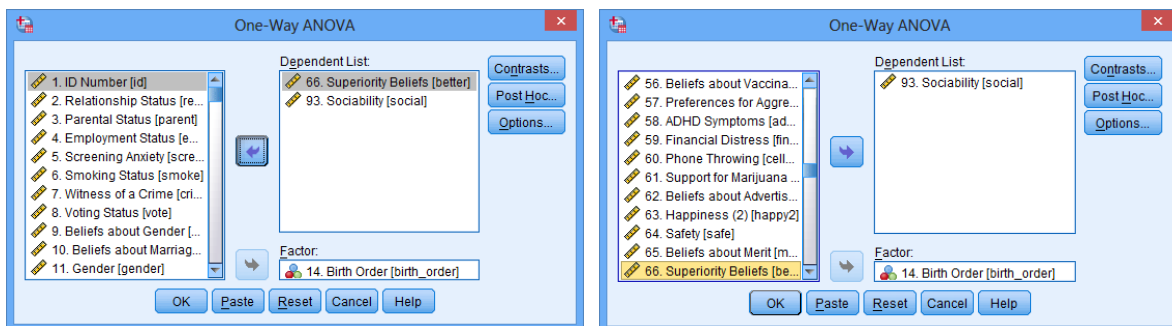


Now, click on the "Post Hoc" button. Another window pops up. There are many different types of post-hoc tests. Some post-hoc tests are very liberal – they more easily declare a group difference to be statistically significant (e.g., Least Significant Difference, or "LSD"). Other post-hoc tests are more conservative – they are more cautious and tend to declare fewer group differences to be statistically significant (e.g., "Scheffe"). Many are somewhere in between. The LSD and Sheffe are very commonly used, and the most common "inbetween" option is the "Tukey," which is neither too liberal nor conservative (a moderate). Statisticians have and will continue to debate the merits of each post-hoc test. Acknowledging that's probably a very

exciting debate, we'll just stick with the Tukey for these analyses. Click Tukey, then Continue, then OK.
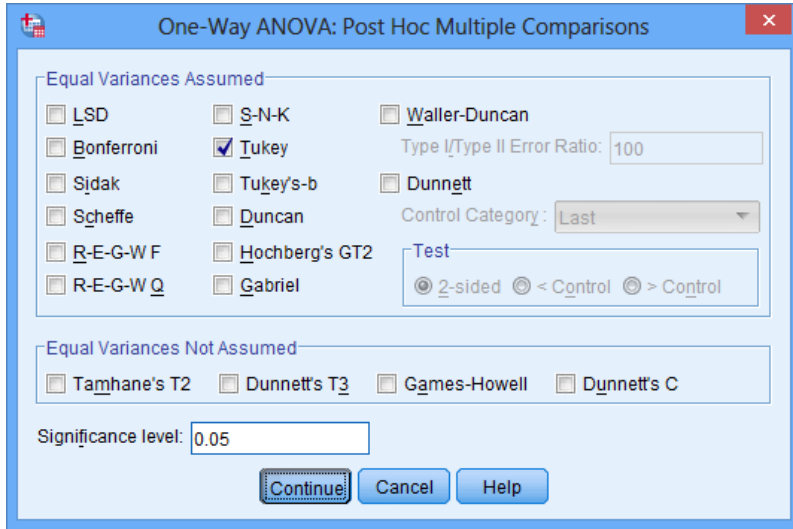


The Output should look familiar. It added two additional boxes.

**Descriptives**

93. Sociability

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| Only child | 51 | 5.84 | 2.485 | .348 | 5.14 | 6.54 | 0 | 10 |
| Youngest | 168 | 7.20 | 2.089 | .161 | 6.88 | 7.52 | 2 | 10 |
| In the middle | 110 | 6.39 | 2.563 | .244 | 5.91 | 6.88 | 0 | 10 |
| Oldest | 180 | 6.82 | 2.394 | .178 | 6.47 | 7.17 | 0 | 10 |
| Total | 509 | 6.76 | 2.378 | .105 | 6.55 | 6.96 | 0 | 10 |

**ANOVA**

93. Sociability

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 91.426 | 3 | 30.475 | 5.535 | .001 |
| Within Groups | 2780.366 | 505 | 5.506 | | |
| Total | 2871.792 | 508 | | | |

## Post Hoc Tests

8

**Multiple Comparisons**

93. Sociability
Tukey HSD

| (I) 14. Birth Order | (J) 14. Birth Order | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound |
|---|---|---|---|---|---|---|
| Only child | Youngest | -1.359* | .375 | .002 | -2.33 | -.39 |
| | In the middle | -.548 | .397 | .514 | -1.57 | .48 |
| | Oldest | -.979* | .372 | .043 | -1.94 | -.02 |
| Youngest | Only child | 1.359* | .375 | .002 | .39 | 2.33 |
| | In the middle | .811* | .288 | .026 | .07 | 1.55 |
| | Oldest | .380 | .252 | .432 | -.27 | 1.03 |
| In the middle | Only child | .548 | .397 | .514 | -.48 | 1.57 |
| | Youngest | -.811* | .288 | .026 | -1.55 | -.07 |
| | Oldest | -.431 | .284 | .427 | -1.16 | .30 |
| Oldest | Only child | .979* | .372 | .043 | .02 | 1.94 |
| | Youngest | -.380 | .252 | .432 | -1.03 | .27 |
| | In the middle | .431 | .284 | .427 | -.30 | 1.16 |

*. The mean difference is significant at the 0.05 level.

## Homogeneous Subsets

93. Sociability

Tukey HSD[a,b]

| 14. Birth Order | N | Subset for alpha = 0.05 1 | Subset for alpha = 0.05 2 |
|---|---|---|---|
| Only child | 51 | 5.84 | |
| In the middle | 110 | 6.39 | 6.39 |
| Oldest | 180 | | 6.82 |
| Youngest | 168 | | 7.20 |
| Sig. | | .354 | .071 |

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 99.486.
b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

The third box of the Output provides the most important information. The Blue box compares the means and the Red oval provides *p*-values indicating which group differences are statistically significant. These tests are similar to running *t*-tests on every possible two-group comparison.

The results are complicated, so it's good to have a general strategy for how to proceed. Usually, it's useful to begin with the highest-scoring group (in this case Youngest Children) and examine how they compare to each group. Then, look at how the second-highest group (Oldest Children) compares to each group with a lower mean, and so on.

- Youngest Children were reliably more sociable than Only Children ($p = .002$) and Middle Children ($p = .03$) but do not significantly differ from Oldest Children ($p = .43$).
- Oldest Children were more sociable than Only Children ($p = .04$) but did not significantly differ from the other groups.
- Middle Children were less sociable than Youngest Children ($p = .03$) but did not differ from the other groups.

9

- Only Children were less sociable Youngest Children ($p = .002$) and Oldest Children ($p = .04$) but could not be differentiated from Middle Children.

When there are many groups present (4, 5, 6…), it's easy for beginners to get overwhelmed by these post-hoc tests, which is why it's useful to use the above procedure of working from highest-to-lowest. It's also useful to end by summarizing the results as simply as possible. "The ANOVA indicated that birth order was significantly related to sociability. The biggest statistically significant difference found in post-hoc analyses was that Youngest Children were more sociable than Only Children, with Middle and Oldest Children falling in between."

---

Practice Questions

4) Conduct an ANOVA analysis examining whether Geographic Region (#12) is associated with Satisfaction with Appearance (#38). How would you report the result in APA style, using the APA Style Guide at the end of this assignment? Cohen's *d* calculations are optional, though sometimes helpful for writing up ANOVA results.
5) Conduct an ANOVA analysis examining whether one's Preferred Depression Treatment (#22) is related to Self-esteem (#46). How would you report the result in APA style, using the APA Style Guide at the end of this assignment?
6) Conduct an ANOVA analysis examining whether one's News Media Choice (#19) is related to Beliefs about Sexual Orientation (#36). How would you report the result in APA style, using the APA Style Guide at the end of this assignment? Is it appropriate to examine the Post-Hoc results at all?

---

## Analytic Decision Making: Choosing the Best Statistical Test

A number of statistical tests have been covered in this course: correlation, regression, between-group *t*-test, and ANOVA. The guide below provides input on when to use each test.

| Independent Variable(s) | Dependent Variable | Test |
|---|---|---|
| 1 dichotomous variable (1 variable with 2 categories) | 1 continuous variable | *t*-test [a][b] |
| 1 polytomous variable (1 variable with more than 2 categories) | 1 continuous variable | ANOVA [b] |
| 1 continuous variable | 1 continuous variable | Correlation [b] |
| Several dichotomous or continuous variables | 1 continuous variable | Multiple Regression [b] |

[a] Could use a correlation under those circumstances, though the *t*-test is generally preferred

[b] Could conduct these analyses when the dependent variable is dichotomous. Technically, other tests, such as chi-square and logistic regression are more appropriate, but no need to worry about those for this class.

Practice Questions

7) What statistical analysis would be most appropriate for examining the relationship between the following variables?  Employment Status (#4) to predict Depression (#29)
8) What statistical analysis would be most appropriate for examining the relationship between the following variables?  Anger (#44) to predict Healthy Food Choice (35)
9) What statistical analysis would be most appropriate for examining the relationship between the following variables?  Extraversion (#108) and Neuroticism (#109) to predict Health (#106)
10) What statistical analysis would be most appropriate for examining the relationship between the following variables?  Preferred Marriage Age (#15) to predict Social Anxiety (#41)

Lab Assignment

You are now ready to begin working on your next homework assignment, "Lab Assignment 5" (see "Due" column on the Course Calendar).

**Dismissal**

The lab instructor can dismiss students if they have completed LA5 in its entirety

## APA Style Guide

*Note*: You have my permission to copy any or all of this writing for this or future assignments.

### Style and Rounding

Rules governing rounding vary considerably from discipline to discipline. These guidelines reflect the current norms in psychology.

*p*-values. Historically, published articles either reported statistically significant findings as "$p <$ .05" and non-significant findings as "*ns*" – this was a very imprecise way of reporting the results. All major psychology journals now advocate reporting actual *p*-values when they are provided in text (in tables and figures asterisks are still common). In general, *p*-values should be reported rounded to two decimals (e.g., $p = .08$ or $p = .02$). However, if the *p*-value is less than .01, report three decimals (e.g., $p = .009$ or $p = .002$). If the p-value is less than .001, simply report as "$p <$ .001" (note, SPSS strangely reports these as .000, but it is impossible to have a probability of zero, so do not report it that way).

Percentages. Usually percentages are rounded to one decimal place (e.g. 88.6% or 1.1%).

Other statistics. In general, all other statistics are rounded to two decimal places (e.g., $M = 1.46$ or $r = -.33$)

Leading zero. If a statistic is a decimal, people usually include a leading zero only if the statistic can commonly exceed 1.0. For example, most descriptive statistics, as well as *t*-scores, *Z*-scores, and *F*-scores commonly exceed 1.0, so even if an observed value is a decimal, a leading zero is included (e.g., $Z = 0.23$ or $t = 0.96$). In contrast, correlations cannot exceed 1.0, so no leading zero is included (e.g., $r = .23$ or $r = .96$).

Italics. Statistical symbols should be italicized (e.g., *M*, *SD*, *r*, *t*, *d*, *F*, *p*, etc.) but not the numbers following them.

### Descriptive Statistics

On the 1-9 depression severity scale, the sample reported a mean score of 4.66 (*SD* = 1.59), with 12.3% reporting a "1" (not at all depressed) and 2.1% reporting a "9" (completely depressed).

The sample was predominantly white (94.5%) and college-educated (86.6%), more often female (61.3%), and distributed relatively evenly across the U.S. (North: 24%, South: 30%, Midwest: 20%, West: 26%).

Participants varied considerably in age (*M* = 35.5 years, *SD* = 10.2, ranging from 18 to 77).

Participants identified as Democrats (30.2%), Republicans (19.8%), or Independents (50.0%).

## Correlation (Significant, $p < .05$)

*Note:* Include the correlation, *p*-value, a description of the direction (more, less, positively, negatively, directly, inversely, etc.), and a description of the effect size (e.g., near-zero/marginal, small/slight, medium/moderate/modest, strong/large/sizeable). If the finding might be confusing to a non-statistician, include a second sentence explaining the finding in simpler terms.

Participants who were more neurotic reported exercising moderately less often, which was statistically significant, $r = -.35$, $p = .02$.

Quarterbacks who were taller had marginally better completion rates, $r = .09$, $p = .04$. Thus, tall quarterback throw completed (caught) passes more often than short quarterbacks.

Anxiety and depression were strongly positively correlated ($r = .71$, $p = .007$). Therefore, it could be difficult to distinguish between whether someone's primary diagnosis should be an anxiety disorder or a mood disorder.

## Correlation (Non-Significant, $p > .05$)

Age was not significantly associated with income ($r = .13$, $p = .23$), political views ($r = .01$, $p = .99$), or vocabulary ($r = .06$, $p = .62$).

The present study failed to find an association between wealth and happiness, $r = .08$, $p = .64$.

## Several Correlations, followed by Multiple Regression

*Note:* First, describe the correlational results, where you compare each of the predictors to the dependent variable. Then, provide a rationale for the regression analyses. In reporting the results, people usually include $R$, $R^2$, or both, followed by the *p*-value. Then, describe the results in plain English, if needed.

Family stress ($r = .48$, $p = .008$), work stress ($r = .56$, $p < .001$), and school stress ($r = .21$, $p = .04$) all significantly predicted overall life stress. However, social support did not predict level of life stress, $r = .03$, $p = .64$. Thus, although social support was not related to life stress, one's level of school stress was slightly related, family stress was modestly related, and work stress was strongly related to level of life stress. To examine the overall contribution of the three significant predictors (school stress, family stress, and life stress) in accounting for life stress, multiple regression was used. The results of the multiple regression analysis indicated that these three predictors accounted for a large proportion of the variance in life stress, $R^2 = .40$, $p < .001$. Thus, school stress, family stress, and work stress together account for 40% of the differences in overall life stress.

Several factors were hypothesized to predict college GPA. Being encouraged to read ($r = .19$, $p = .002$) and conscientiousness ($r = .26$, $p < .001$) had small positive relationships with college GPA. ADHD symptoms had a small negative relationship ($r = -.17$, $p = .007$). Hours of work per week was not correlated with GPA ($r = .08$, $p = .22$). Thus, being encouraged to read and being conscientious are related to better grades, but having ADHD symptoms is related to lower grades. The number of hours people spend on employment was not related to grades. Multiple regression was used to examine the combined effect of being encourages to read, conscientiousness, and ADHD symptoms on college GPA. These three predictors combined to modestly predict GPA, $R = .33$, $R^2 = .11$, $p < .001$. Therefore, being encouraged to read, conscientiousness, and ADHD symptoms explain 11% of the differences in college grades.

**_t_-test (Levene's test Non-Significant, _p_ > .05, and _t_-test Significant, _p_ < .05)**

_Note_: Be sure to include the basic descriptives (_M_ and _SD_), _d_ (calculated), _t_, _df_ (in parentheses), and the _p_-value.

Females (_M_ = 2.34, _SD_ = 2.06) tan slightly more often than males (_M_ = 1.60, _SD_ = 1.46), which was a significant effect, _d_ = 0.42, _t_(298) = 3.11, _p_ = .002.  Thus, women are more likely to go tanning than men.

**_t_-test (Levene's test Non-Significant, _p_ > .05, and _t_-test Non-Significant, _p_ > .05)**

Smokers (_M_ = 5.01, _SD_ = 2.38) were slightly moodier than non-smokers (_M_ = 4.50, _SD_ = 2.18); however, this differences was non-significant, _d_ = 0.26, _t_(298) = 1.64, _p_ = .10.  That is, smoking is unrelated to moodiness.

**_t_-test (Levene's test Significant, _p_ < .05, and _t_-test Significant, _p_ < .05)**

_Note_: Find and report the _p_-value from the Levene's test, note that a modified _t_-test was used, and report the _t_-test results from the lower row of the Output (Equal variances not assumed). The degrees-of-freedom (df) value for the t-test will often include decimals, and for simplicity, just round that to the nearest whole number.

A Levene's test showed that equality of variances could not be assumed, _p_ = .006. Due to this violated assumption, a _t_-test not assuming homogeneity of variance was used. This showed that females (_M_ = 2.34, _SD_ = 2.06) tan slightly more often than males (_M_ = 1.60, _SD_ = 1.46), which was a significant effect, _d_ = 0.42, _t_(59) = 2.46, _p_ = .02.  Thus, women are more likely to go tanning than men.

**_t_-test (Levene's test Significant, _p_ < .05, and _t_-test Significant, _p_ > .05)**

_Note:_ If Levene's test indicates the violated assumption (p < .05), use the t-test Output in the lower row of the Output (Equal variances not assumed). If the analysis is an important one, note the result of the Levene's test and the _t_-test. If the analysis is not central, just report the _p_-value associated with the _t_-test.

A Levene's test showed that equality of variances could not be assumed, _p_ = .006. Due to this violated assumption, a _t_-test not assuming homogeneity of variance was used. However, this showed that females (_M_ = 2.34, _SD_ = 2.06) and males (_M_ = 1.60, _SD_ = 1.46) did not differ in terms of extraversion, _d_ = 0.08, _t_(36) = 0.99, _p_ = .78.

_Or more simply:_  There were no gender differences in extraversion (_p_ = .78).

**ANOVA (Significant, _p_ < .05):**

_Note:_ Include the degrees of freedom (the top two _df_ values in the Output), the _F_-value, and the _p_-value.  Include basic descriptive statistics as well.  You may calculate Cohen's _d_ to compare any two of the specific groups, if that would be informative, but if you have many groups, listing out all of the _d_-values could be tedious.

Music device preference was significantly related to openness to experience, $F(2,297) = 4.30$, $p = .02$. People who listen to vinyl or cassettes were highest ($M = 8.33$, $SD = 0.78$) on openness to experience, followed by .mp3 listeners ($M = 6.91$, $SD = 1.75$), followed by CD listeners ($M = 6.86$, $SD = 1.66$). A post-hoc test revealed that each of the group differences was statistically significant. People who use older music devices are more open.

## ANOVA (Non-Significant, $p > .05$):

Drivers, walkers, and bikers did not differ significantly in terms of religious involvement, $F(2,297) = 1.65$, $p = .19$. Thus, transportation mode is not related to involvement with religion.