CH. 12 Correlation – *using JAMOVI*

PY 221 Statistics & Research Methods I

Dr. Valenti

Outline for Ch. 12 Correlation – using JAMOVI

1. **How to run correlation using JAMOVI**
2. How to interpret JAMOVI output
3. How to write up results using APA style

Correlation using JAMOVI: Example Study

- What relationships exist between students' exam performance, exam anxiety, and the time students spend studying?
- Participants:
  - 103 students
- Measures (variable names from data file are in *italics*)
  - Time spent "revising" aka studying (in **hours**) - *Revise*
  - Exam performance (**% grade**) - *Exam*
  - Exam Anxiety (the EAQ, **score out of 100**) - *Anxiety*
  - Gender - **Gender**

Correlation: What hypotheses are we testing? (*review*)

When *r* is closer to 0, *p* will be larger.

null hypothesis?

- $H_0 : \rho = 0$
- There is **no relationship** between exam anxiety and exam performance.

alternative hypothesis?

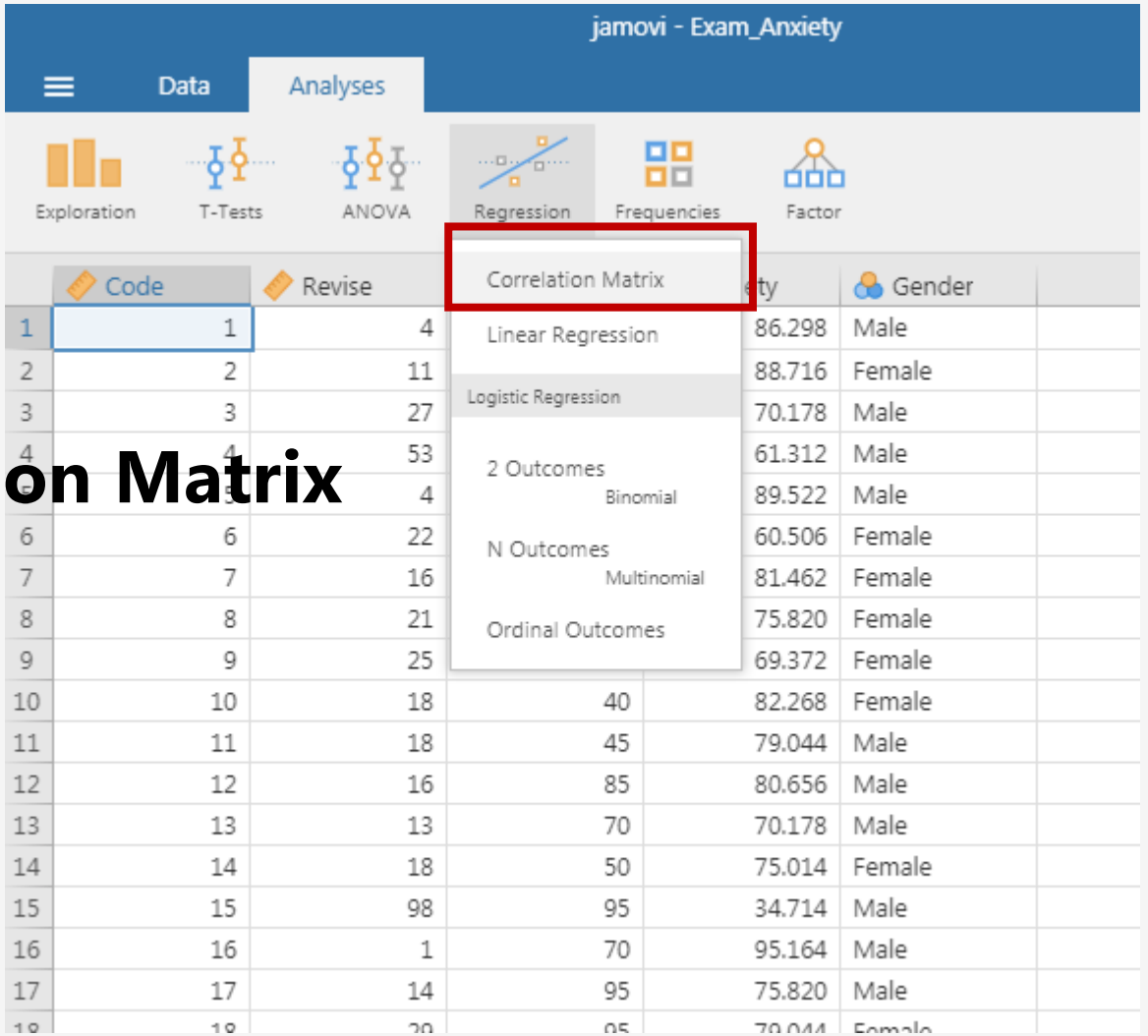
- $H_1 : \rho \neq 0$
- There **is a relationship** between exam anxiety and exam performance.

When *r* is farther from 0, *p* will be smaller.

- **Our *r* and associated *p*-value tell us about the credibility of the null hypothesis.**
- Let's set our $\alpha = .01$.

Correlation: How to run in JAMOVI

Analyses → Regression → Correlation Matrix



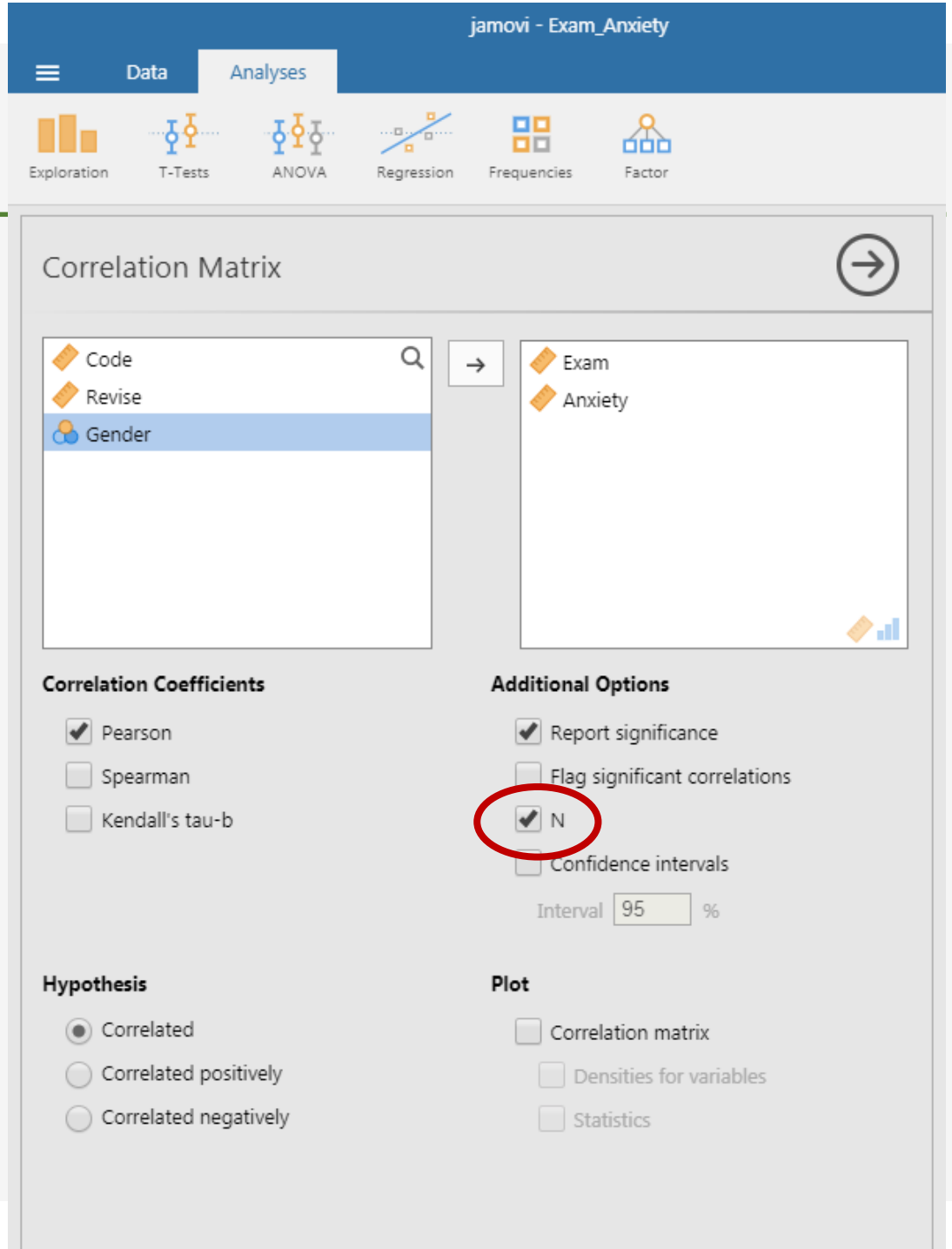
Correlation:
How to run in JAMOVI

Click over two relevant variables (or more than two, if you're asked to examine the relationships between more than one pair of variables).

In this first example, let's examine whether exam anxiety is related to exam performance.

Click the box for N in order to get the sample size.

Leave everything else the same. (Your screen should look like the one on the right.)



Outline for Ch. 12 Correlation – using JAMOVI

1. How to run correlation using JAMOVI
2. **How to interpret JAMOVI output**
3. How to write up results using APA style

*To help us **interpret** r, we can start by looking at our p-value & comparing it to the alpha we set.*
Remember the guidelines:

$p < \alpha$ → reject null; there's a significant relationship
$p > \alpha$ → retain null; there is no significant relationships

Correlation Matrix

The correlation coefficient, r = **-.44**
(**Remember to pay attention to the positive or negative sign.)

The two variables we are examining the relationship between.

The p-value associated with our r of -.44. Indicates whether to reject or retain the null hypothesis.

Sample size, N = 103

Correlation Matrix

		Exam	Anxiety
Exam	Pearson's r	—	
	p-value	—	
	N	—	
Anxiety	Pearson's r	−0.44	—
	p-value	< .001	—
	N	103	—

Interpreting output from a correlational analysis

Interpret *r*. *(from handout)*

We *reject* the null hypothesis b/c $p < .001$, which is also less than our α of .01.

There is a significant negative relationship between exam performance and exam anxiety.

- As exam anxiety goes up, exam performance worsens.
- The effect size is medium-to-large (*or* There is a moderate-to-strong relationship.
- 19% of the variability in exam performance can be explained by a person's exam anxiety.

Correlation Matrix

$R^2 = (-.44)(-.44) = .19$ or 19%

Correlation Matrix

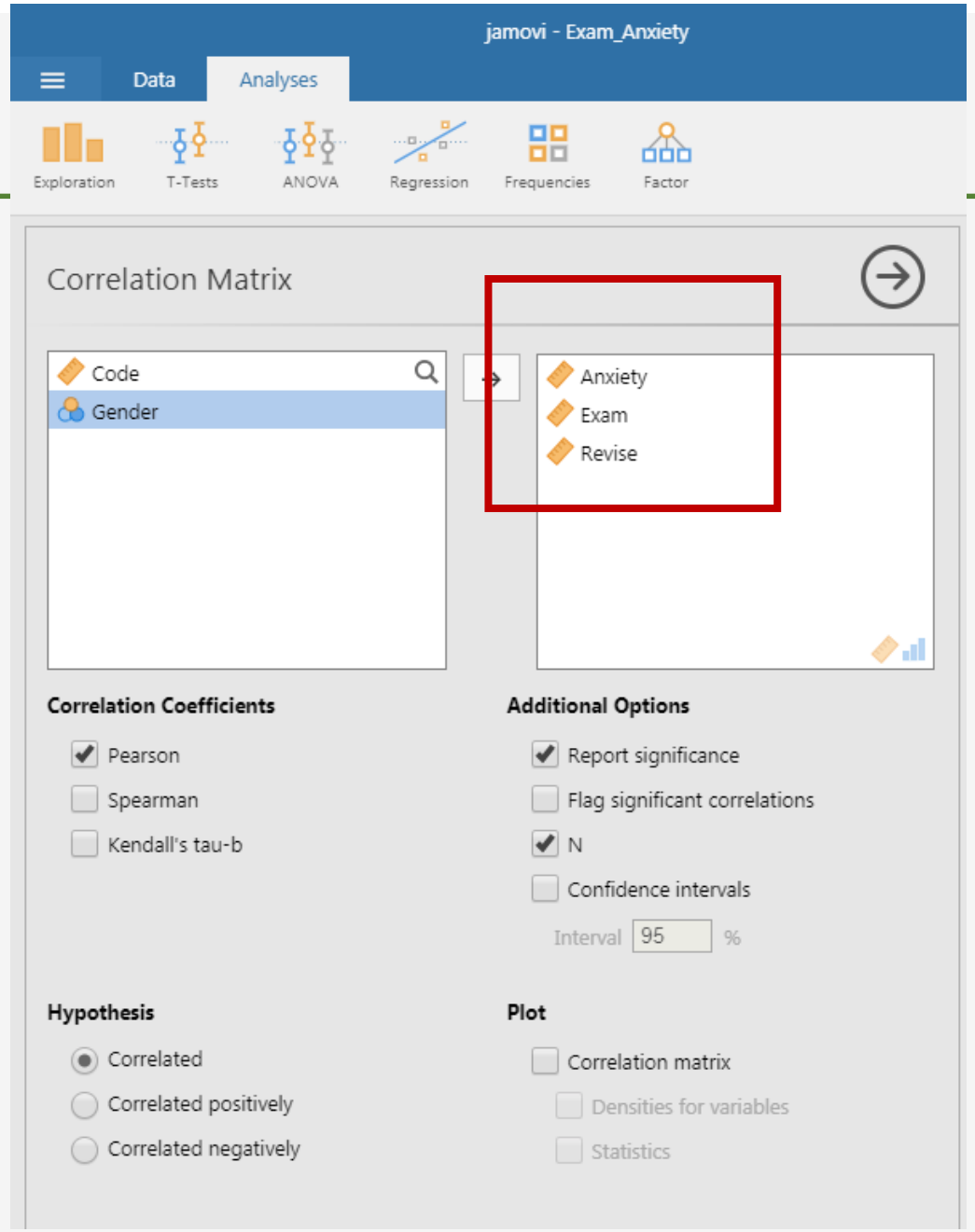
		Exam	Anxiety
Exam	Pearson's r	—	
	p-value	—	
	N	—	
Anxiety	Pearson's r	−0.44	—
	p-value	< .001	—
	N	103	—

Assume α = .01.

You can run correlations for multiple pairs of variables at a time.

Use the same procedure as before, just click over whichever variables you are interested in.

JAMOVI will run a correlational analysis for each pair, separately.



Correlation in JAMOVI:
Output for correlation between exam performance, exam anxiety, and time spent revising

1. What is the correlation between exam anxiety and time spent "revising" aka studying?
2. What's the p-value for that correlation? Assume α=.01. Is the correlation statistically significant? How do you know?
3. Is the relationship weak, moderate, or strong? Should you interpret the effect size at all?
4. What % of the variance in time spent studying is shared by exam anxiety?
5. Which pair of variables has the weakest relationship?
6. For the pair of variables you identified in question 5, finish this statement indicating the direction of their relationship: *As time spent studying increases, …*

Correlation Matrix

Correlation Matrix

		Anxiety	Exam	Revise
Anxiety	Pearson's r	—		
	p-value	—		
	N	—		
Exam	Pearson's r	−0.44	—	
	p-value	< .001	—	
	N	103	—	
Revise	Pearson's r	−0.71	0.40	—
	p-value	< .001	< .001	—
	N	103	103	—

Correlation in JAMOVI:
Output for correlation between exam performance, exam anxiety, and time spent reviewing

1. What is the correlation between exam anxiety and time spent revising?

***r* = -.71** (don't forget the negative sign!)

2. What is the p-value? Assuming α = .01, is that correlation statistically significant? How do you know?

***p* < .001. Yes, it is significant because it is less than .01**

3. Is that relationship weak, moderate, or strong? Should you interpret the effect size at all?

Large, because the absolute value of *r* is at least .50.

We *should* interpret the effect size since the relationship is statistically significant (i.e, since p < α).

Correlation Matrix

Correlation Matrix

		Anxiety	Exam	Revise
Anxiety	Pearson's r	—		
	p-value	—		
	N	—		
Exam	Pearson's r	−0.44	—	
	p-value	< .001	—	
	N	103	—	
Revise	Pearson's r	−0.71	0.40	—
	p-value	< .001	< .001	—
	N	103	103	—

Correlation in JAMOVI:
Output for correlation between exam performance, exam anxiety, and time spent reviewing

4. What % of the variance in time spent studying is shared by exam anxiety?

   Calculate the *coefficient of determination*,
   R^2 = -.71 * -.71 = .5041 ~ **50%**

5. Which pair of variables has the *weakest* relationship?

Exam *performance* and time spent revising (studying).

The *r* = .40, which has the smallest absolute value compared with -.44, and -.71. It is the closest to zero; however it is still statistically significant.

6. *As time spent studying (revising) increases,*
   exam performance increases.

You could also describe this relationship as: *As time spent studying decreases, exam performance decreases*. **OR**

As exam performance increases, time spent studying increases.

Correlation Matrix

Correlation Matrix

		Anxiety	Exam	Revise
Anxiety	Pearson's r	—		
	p-value	—		
	N	—		
Exam	Pearson's r	−0.44	—	
	p-value	< .001	—	
	N	103	—	
Revise	Pearson's r	−0.71	0.40	—
	p-value	< .001	< .001	—
	N	103	103	—

Flip the page over on your handout for a new example . . .

- A record label collected data for each band on four variables:
  - Sales – album sales for the band
  - Airplay – # of plays on the radio for the band's songs
  - Adverts – advertising budget for the band's albums
  - Attract – how attractive the band members are

1. What is the sample size?

N = 200 bands

2. What null & alternative hypotheses (in symbols) are being tested for each pair of variables?

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

Ad budget

Correlation Matrix

Ad budget

Correlation Matrix

		Sales	Airplay	Adverts	Attract
Sales	Pearson's r	—			
	p-value	—			
	N	—			
Airplay	Pearson's r	0.60	—		
	p-value	< .001	—		
	N	200	—		
Adverts	Pearson's r	0.58	0.10	—	
	p-value	< .001	0.151	—	
	N	200	200	—	
Attract	Pearson's r	0.33	0.18	0.08	—
	p-value	< .001	0.010	0.256	—
	N	200	200	200	—

PRACTICE interpreting output

Q3 from handout

	r	direction	*p*	Stat. sig?	Reject or retain null?	Effect size	Coefficient of determin-ation
Attractiveness of band & # of plays on radio	.18	positive	$p = .01$	Yes	REJECT	weak	.03 or 3%
Album sales and # of plays	.60	positive	$p < .001$	Yes	REJECT	strong	.36 or 36%
Ad budget and attractiveness	.08	positive	$p = .256$	NO	retain	very weak / no relationship	.006 or 0.60%
Album sales & attractiveness	.33	positive	$p < .001$	Yes	REJECT	moderate	.11 or 11%
Album sales & ad budget	.58	Positive	$p < .001$	Yes	REJECT	strong	.34 or 34%
# plays on radio and ad budget	.10	positive	$p = .151$	NO	retain	Weak	.01 or 1%

4. Write a sentence interpreting the R^2 for the relationship between advertising budget and album sales.

We already calculated that $R^2 = 34\%$.

Thirty-four percent of the variance in album sales is explained by advertising budget.

Correlation Matrix

Correlation Matrix

		Sales	Airplay	Adverts	Attract
Sales	Pearson's r	—			
	p-value	—			
	N	—			
Airplay	Pearson's r	0.60	—		
	p-value	< .001	—		
	N	200	—		
Adverts	Pearson's r	0.58	0.10	—	
	p-value	< .001	0.151	—	
	N	200	200	—	
Attract	Pearson's r	0.33	0.18	0.08	—
	p-value	< .001	0.010	0.256	—
	N	200	200	200	—

PRACTICE interpreting output

Research questions based on student survey (see handout p.3)

Research Question #1

Is the amount of Netflix you watch related to how much you exercise?

Research Question #2

Is the amount of alcohol you consume in a night related to how happy you are?

Is the amount of Netflix you watch related to how much you exercise?

Correlation Matrix

Correlation Matrix

		netflixhr_1	exercisemin_1
netflixhr_1	Pearson's r	—	
	p-value	—	
	N	—	
exercisemin_1	Pearson's r	−0.21	—
	p-value	0.456	—
	N	15	—

Is the amount of Netflix you watch related to how much you exercise?

1. What is the correlation between the number of minutes a person exercises during a typical day and the number of hours of Netflix a person watches in a typical week?

***r* = -.21**

2. What is the p-value associated with this correlation? Is this correlation statistically significant, assuming α = .05? Do you reject or retain the null hypothesis?

***p* = .456. Since the p-value is greater than .05, the correlation is not statistically significant and we retain the null.**

Is the amount of alcohol you consume in a night related to how happy you are?

Correlation Matrix

Correlation Matrix

		numdrinkyou_1	happynow
numdrinkyou_1	Pearson's r	—	
	p-value	—	
	N	—	
happynow	Pearson's r	−0.58	—
	p-value	0.024	—
	N	15	—

Is the amount of alcohol you consume in a night related to how happy you are?

3. What is the correlation between the number of drinks a person normally has when he/she drinks and how happy that person was on the day he/she completed the survey? ***r* = -.58**
4. What is the p-value associated with this correlation? Is this correlation statistically significant, assuming α = .05? Do you reject or retain the null hypothesis? ***p* = .024, so yes, the correlation is statistically significant and we will <u>reject</u> the null hypothesis that there's no relationship.**
5. How would you describe this relationship in words? That is, interpret r (referring to significance (or not), and, if significant, referring to direction, strength (effect size), and coefficient of determination). Write in full sentences

The amount of alcohol someone drinks is significantly negatively related to how happy they are; the more someone drinks, the less happy they report being. This effect represents a large effect size.

R^2 = -.58 * -.58 = .3364 ~ 34%

Thirty-four percent of the variability in people's happiness is explained by how much they drink.

***OR,* Thirty-four percent of the variability in people's drinking habits is explained by how happy they are.**

Outline for Ch. 12 - Correlation

1. How to run correlation using JAMOVI
2. How to interpret JAMOVI output
3. **How to write up results using APA style**
   - **Please see p. 4 in your handout**

Watch Commercials, Buy Candy?

Imagine that 5 people were asked how many advert-isements (commercials) for a certain candy they saw one week, and then researchers measured how many packets of that candy they purchased the next week . . .

Relevant research question: Is the # of candy commercials you saw one week related to the # of candy packets you purchased the next week?

TABLE 6.1

Subject	1	2	3	4	5	Mean	*S*
Adverts Watched	5	4	4	6	8	**5.4**	**1.67**
Packets Bought	8	9	10	13	15	**11.0**	**2.92**

Reporting correlations in APA-style p

Results

To examine whether there was a relationship between number of commercials watched and the amount of candy purchased, we ran a correlational analysis using these two variables. There is a significant positive correlation between number of commercials watched and amount of candy purchased, $r(5) = .87$, $p = .034$. As the number of commercials watched increases, the amount of candy purchased increases. This relationship has a large effect size. Also, 76% of the variance in how much candy people purchased is explained by the number of commercials people watched.

Say that you ran a correlational analysis and list the two variables that went into it.

Your Pearson's *r*

Specify not only that you have a significant correlation, but that the direction is *positive*.

Italicize *r*

Sample size in parentheses

Italicize *p*, but not the value of p.

Since our p-value is less than α (which is .05), this correlation is *statistically significant*.

Results

To examine whether there was a relationship between number of commercials watched and the amount of candy purchased, we ran a correlational analysis using these two variables. There is a significant positive correlation between number of commercials watched and amount of candy purchased, $r(5) = .87$, $p = .034$. As the number of commercials watched increases, the amount of candy purchased increases. This relationship has a large effect size. Also, 76% of the variance in how much candy people purchased is explained by the number of commercials people watched.

Effect size is large b/c any *r* that is .50 or greater is considered large, and here, *r* is .87.
Alternatively, you can state:
These variables have a strong relationship.

This is the R^2. Turn the R^2 decimal into a % by multiplying by 100.

The format of this final statement will always be the same. You state that the "% of the variance in [insert one of your two variables] is shared by (or 'is explained by') [insert other variable]."
Use the variable names here, rather than saying more/less of that variable. e.g., say "how much candy" vs. "more candy" or "less candy"

Correlation Matrix

- Example from earlier in this lecture, on exam anxiety, performance and studying.
- See page 5 of the handout from class.
- Results section on next slide

Correlation Matrix

		Anxiety	Exam	Revise
Anxiety	Pearson's r	—		
	p-value	—		
	N	—		
Exam	Pearson's r	−0.44	—	
	p-value	< .001	—	
	N	103	—	
Revise	Pearson's r	−0.71	0.40	—
	p-value	< .001	< .001	—
	N	103	103	—

Results

To examine whether exam anxiety, exam performance, and time studying were related, we ran correlational analyses using these three variables. We observed a significant negative correlation between exam anxiety and exam performance ($r(103) = -.44$, $p < .001$), such that higher anxiety was associated with worse exam performance. This is a moderate-to-strong relationship, and approximately 19% of the variation in exam performance can be explained by exam anxiety. We also observed a significant positive relationship between time spent studying and exam performance ($r(103) = .40$, $p < .001$), such that more time spent studying was associated with better exam performance, a relationship that was moderate to strong (or just *moderate)* in size. About 16% of the variation in exam performance can be explained by how much time students spent studying. Finally, we observed a significant negative correlation between exam anxiety and time spent studying ($r(103) = -.71$, $p < .001$), a strong relationship. Greater exam anxiety was associated with less time spent studying. About half the variation in time spent studying is shared by exam anxiety.

Data from page 6 of handout

Correlation Matrix

Correlation Matrix

		numdrinkyou_1	numdrinkpeer_1
numdrinkyou_1	Pearson's r	—	
	p-value	—	
	N	—	
numdrinkpeer_1	Pearson's r	0.09	—
	p-value	0.745	—
	N	15	—

Correlation Matrix

Correlation Matrix

		numdrinkyou_1	happynow
numdrinkyou_1	Pearson's r	—	
	p-value	—	
	N	—	
happynow	Pearson's r	−0.58	—
	p-value	0.024	—
	N	15	—

APA-STYLE RESULTS SECTION FROM PAGE 6 OF HANDOUT

Results

We ran two correlational analyses to examine, first, whether the number of alcoholic drinks a person themselves drinks is related to the number of drinks they believe their typical peer drinks per night of drinking, and second, whether the number of alcoholic drinks a person themselves drinks is related to that person's self-reported happiness at the time of the survey. There was no significant relationship between the number of alcoholic drinks people normally drink and the number of drinks they believe others normally drink, $r(15) = .09$, $p = .75$. In other words, whether a person normally drinks a lot or a little is not related to their perceptions of how much others normally drink. There was, however, a significant negative relationship between how much a person normally drinks and how happy they were on the day of the survey, $r(15) = -.58$, $p = .024$. In particular, the more a person normally drinks, the less happy they were on the day they took the survey. The number of drinks a person normally drinks shared 34% of its variance with how happy a person was on the day of the survey, and the relationship between the two variables was strong.