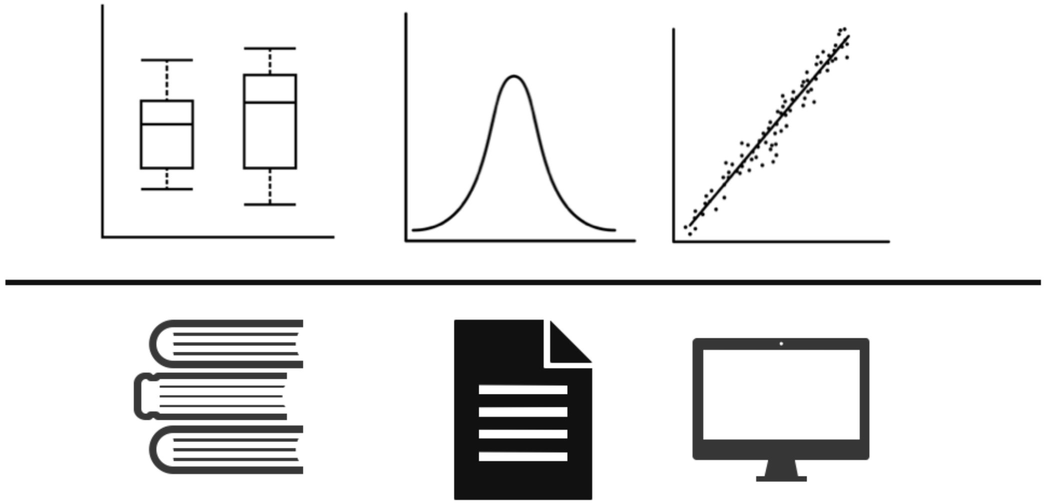
PSYC 2300

Introduction to Statistics



Lecture 09: Differences from the Population

Outline for today

- **Review**
  - Midterm Exam, Extra Credit
- **One-sample tests**
  - One-sample *z*-test
  - One-sample *t*-test
- **Calculating effect sizes**
- ***t*-tests in JASP**
  - Download `Stats Class 3 Dataset (Pets).csv`

Midterm Review

Overall (25 points)

M = 74%

SD = 15%

Conceptual (20 questions, 10 points)

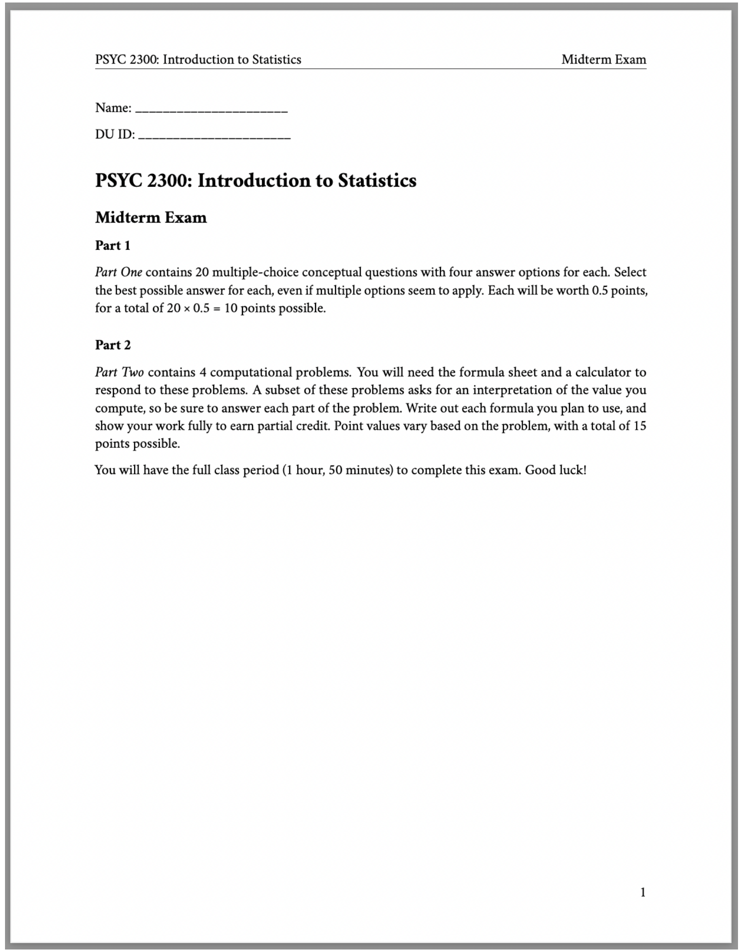
M = 63%

SD = 16%

Computational (4 problems, 15 points)

M = 81%

SD = 20%



Name: ______________________

DU ID: ______________________

PSYC 2300: Introduction to Statistics

Midterm Exam

Part 1

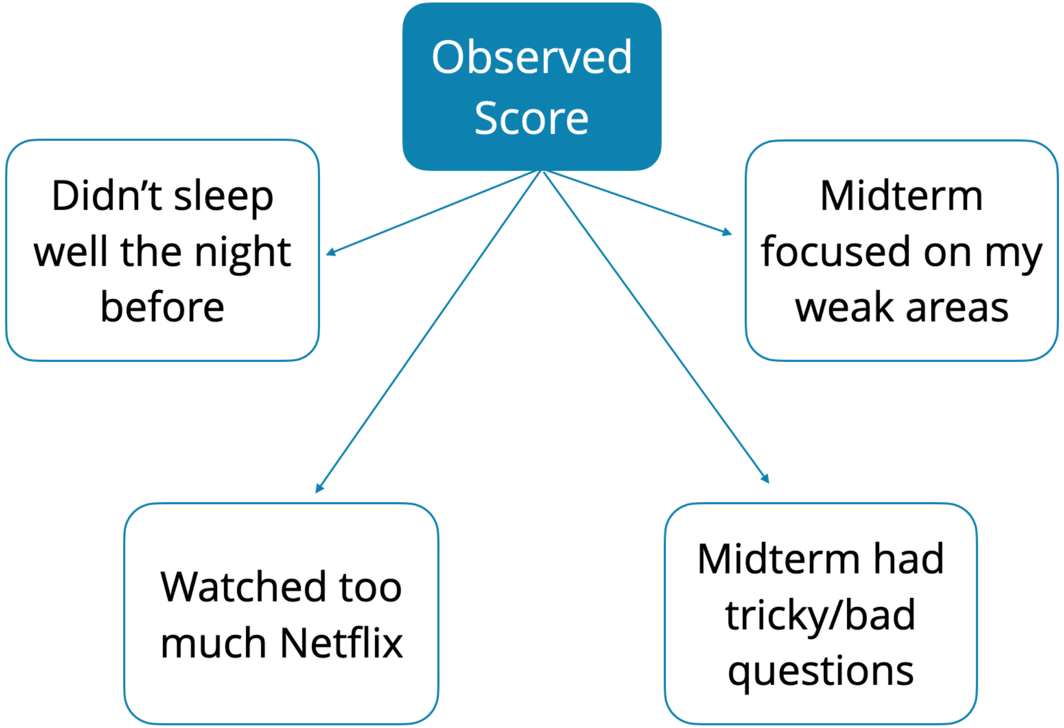
Part One contains 20 multiple-choice conceptual questions with four answer options for each. Select the best possible answer for each, even if multiple options seem to apply. Each will be worth 0.5 points, for a total of $20 \times 0.5 = 10$ points possible.

Part 2

Part Two contains 4 computational problems. You will need the formula sheet and a calculator to respond to these problems. A subset of these problems asks for an interpretation of the value you compute, so be sure to answer each part of the problem. Write out each formula you plan to use, and show your work fully to earn partial credit. Point values vary based on the problem, with a total of 15 points possible.

You will have the full class period (1 hour, 50 minutes) to complete this exam. Good luck!

Classical Test Theory



Question 2

A dataset containing responses to a question in which participants can select between "*Not Angry*", "*Sort of Angry*", and "*Very Angry*" would be what scale of measurement?

A. **Ordinal**

B. Nominal

C. Ratio

D. Interval

$$\text{Not angry} < \text{sort of angry} < \text{very angry}$$

Question 10

A researcher develops a new test to measure people's reaction times. To investigate whether his test is performing well, he asks the same group of 10 people to take the test once this week and another time next month. He correlates the participants' scores from these two testing sessions and **finds a correlation coefficient of 0.09**. Which of the following statements would be reasonable to conclude about this researcher's test?

A. **The test is probably not a valid measure of people's reaction times.**

B. The test has acceptable test-retest reliability.

C. The test has acceptable parallel forms reliability.

D. The researcher should have calculated internal consistency instead.

Question 12

A professor wants to know whether her final exam actually does a good job at measuring her students' current level of understanding. The professor compares her final exam to a **well-established, standardized test** on the same topic. What form of validity is the professor interested in assessing?

A. **Concurrent criterion validity.**

B. Predictive validity.

C. Construct validity.

D. Content validity.

Question 13

Which of the following is the best interpretation of an alpha (α) level of 0.10?

A. **We will only accept the 10% most extreme values under the null hypothesis as evidence to reject the null hypothesis.**

B. We will only accept the 10% most extreme values under the alternative hypothesis as evidence to reject the alternative hypothesis.

C. There is a 10% chance that we will find a value extreme enough to reject the alternative hypothesis

D. There is a 10% chance that we will find a value extreme enough to reject the null hypothesis.

Question 19

What is the probability of committing a Type II Error *assuming the null hypothesis is true?*

A. **0**

B. Alpha

C. Beta

D. There is not enough information to know.

Instructions

> "*Part One* contains 20 multiple-choice conceptual questions with four answer options for each. Select the best possible answer for each, even if multiple options seem to apply. Each will be worth 0.5 points, for a total of 20 × 0.5 = 10 points possible."

Extra Credit Opportunities: SONA

PSYC-2300-2 › Pages › Instructions for the SONA Subject Pool

Student View

Winter Quarter 2022

- Home
- Announcements
- Assignments
- Discussions
- Grades
- People
- Pages
- Files
- Syllabus
- Outcomes
- Rubrics
- Quizzes
- Modules

View All Pages | Published | Edit

Instructions for the SONA Subject Pool

The SONA subject pool allows you to participate in research studies, which provides University of Denver researchers with valuable data and provides you with extra credit to buffer your grade in my course!

The following PDF contains **step-by-step instructions** (with screenshots) for how to register for and use the subject pool: sona-guidelines-for-students.pdf

If you are **already familiar** with how to use the subject pool, you can click the following link to begin: https://du.sona-systems.com/

When registering for the subject pool, be sure to select the appropriate course and section! You should select: **PSYC 2300-1**

Keep in mind that, in order to receive full subject pool (extra) credit for my course, you will need to complete a brief **Subject Pool Reflection** assignment. Failing to do so will result in a penalty to the amount of extra credit points you earn for participating in the subject pool. This assignment is posted on Canvas and is due on the last day of classes by 11:59 PM.

Study Information	Approved?	View
Attitudes & Perceptions (2 Chits) This study is open to BLACK PARTICIPANTS ONLY. Those who are not AT LEAST PART BLACK will be turned away.	Approved	• Study Info • Timeslots
College Student Social Support (2 Chits) (Timeslots Available) (Online Study) This 60-minute online survey involves a nationwide study of social media interactions amongst college students.	Approved	• Study Info • Timeslots
Emotion Study (2 Chits) (Timeslots Available) (Online Study) Online survey asking questions about emotions and thoughts	Approved	• Study Info • Timeslots
Judging Emotional Facial Expressions (1 Chits) (Online Study) Judge the perceived genuineness of emotional facial expressions	Approved	• Study Info • Timeslots
Mechanisms of Conscious Perception - Tall & Flat Judgments (2 Chits) (Online Study) In the current study, you'll be viewing shapes on the screen. Your task will be to make judgments about how tall or flat those shapes looked.	Approved	• Study Info • Timeslots
Mental Impressions of Photographs (2 Chits) (Online Study) In this online study, you will make judgments about the content of photographs and the meaning they convey.	Approved	• Study Info • Timeslots
Pain Experiences and Care Seeking Study (1 Chits) (Timeslots Available) In this study, we are interested in how people experience pain and make care seeking decisions. You will participate in a brief pain induction task and complete brief questionnaires about yourself.	Approved	• Study Info • Timeslots
Perceptions of Groups and Group Members (1 Chits) (Online Study) In this study you will view members of a hypothetical group and form judgments of other members of that group that you have not yet seen.	Approved	• Study Info • Timeslots
Perceptions of Physical Space - W22 (2 Chits) (Timeslots Available) At Frontier Hall. Go DOWNSTAIRS to the basement and sit in the chairs across from the elevator	Approved	• Study Info • Timeslots
Person Judgement (1 Chits) (Online Study) Learn about hypothetical diseases and rate patients on vulnerability to these diseases.	Approved	• Study Info • Timeslots

See Syllabus for more details

One-sample z-test

Goal

- To develop a way to investigate whether a *sample* mean differs significantly from a *population* mean

- Both the one-sample *z*-test and the one-sample *t*-test allow you to do this

$$z_{\overline{x}} \qquad t_{\overline{x}}$$

Test Statistic

Test statistic: A numerical summary of the degree to which a sample is unlike the samples predicted by the null hypothesis, H_0

z-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

Standard Error

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Back to our experiment

In the population

$\mu = 100$

$\sigma = 15$

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$\begin{array}{c} 101, 122, 132 \\ 94, 129, 89 \\ 109, 92, 100 \\ 125, 103, 91 \\ 94, 116, 92 \end{array}$$

Back to our experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

$$z_{\overline{x}} = \frac{105.9 - 100}{\frac{15}{\sqrt{15}}} = 1.53$$

Hypothesis Testing

Null Hypothesis H_0

$$\mu_{NeuroIQ} = 100$$

Alternative Hypothesis H_a

$$\mu_{NeuroIQ} \neq 100$$

Alpha

$$\alpha = .05$$

Test Statistic

$$z_{\overline{x}} = 1.53$$

Determining significance: Software

Your test statistic is a description of how extreme your sample mean is, which is why we can use it to get a *p*-value (from software, tables, etc.)

If $p > \alpha$, fail to reject the null hypothesis

If $p < \alpha$, reject the null hypothesis

One Sample T–Test

One Sample T–Test

	Z	p
neuro_IQ	1.532	0.126

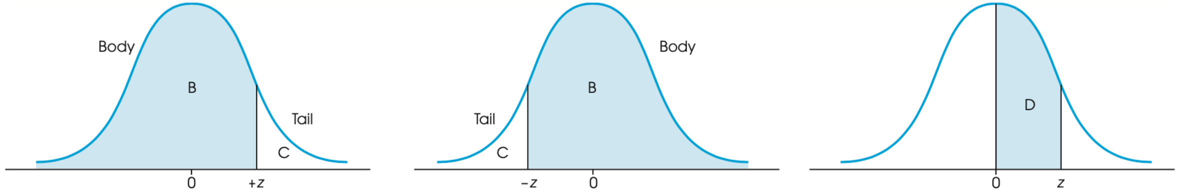
Note. For the Z–test, the alternative hypothesis specifies that the mean is different from 100.
Note. Z test.

Descriptives

Descriptives

	N	Mean	SD	SE
neuro_IQ	15	105.933	15.107	3.901

Determining significance: z table



(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z	(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z
0.00	.5000	.5000	.0000	0.25	.5987	.4013	.0987
0.01	.5040	.4960	.0040	0.26	.6026	.3974	.1026
0.02	.5080	.4920	.0080	0.27	.6064	.3936	.1064
0.03	.5120	.4880	.0120	0.28	.6103	.3897	.1103
0.04	.5160	.4840	.0160	0.29	.6141	.3859	.1141
0.05	.5199	.4801	.0199	0.30	.6179	.3821	.1179
0.06	.5239	.4761	.0239	0.31	.6217	.3783	.1217
0.07	.5279	.4721	.0279	0.32	.6255	.3745	.1255
0.08	.5319	.4681	.0319	0.33	.6293	.3707	.1293
0.09	.5359	.4641	.0359	0.34	.6331	.3669	.1331
0.10	.5398	.4602	.0398	0.35	.6368	.3632	.1368
0.11	.5438	.4562	.0438	0.36	.6406	.3594	.1406
0.12	.5478	.4522	.0478	0.37	.6443	.3557	.1443
0.13	.5517	.4483	.0517	0.38	.6480	.3520	.1480
0.14	.5557	.4443	.0557	0.39	.6517	.3483	.1517
0.15	.5596	.4404	.0596	0.40	.6554	.3446	.1554
0.16	.5636	.4364	.0636	0.41	.6591	.3409	.1591
0.17	.5675	.4325	.0675	0.42	.6628	.3372	.1628
0.18	.5714	.4286	.0714	0.43	.6664	.3336	.1664
0.19	.5753	.4247	.0753	0.44	.6700	.3300	.1700
0.20	.5793	.4207	.0793	0.45	.6736	.3264	.1736
0.21	.5832	.4168	.0832	0.46	.6772	.3228	.1772
0.22	.5871	.4129	.0871	0.47	.6808	.3192	.1808
0.23	.5910	.4090	.0910	0.48	.6844	.3156	.1844
0.24	.5948	.4052	.0948	0.49	.6879	.3121	.1879

One-sample *t*-test

z–test vs. *t*–test

In practice, the *z*-test is rarely used. Why?

_z_-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

Standard Error

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Motivation for the *t*-test

In practice, the z-test is rarely used. Why?

- Because you usually **don't know the true standard deviation** in the population! We need a more general type of hypothesis test that will work for regular science
- Usually, we use *t*-tests instead, since they do not require knowledge of variability statistics in the population
- Versions of *t*-tests can also be used for other situations (e.g., comparing 2 groups), making it useful and versatile

Estimating population variability

We can use our sample's variability to make a best guess about the population's variability

Population Variance

$$\sigma^2 = \frac{\sum (X_i - \mu)^2}{N}$$

Sample Variance

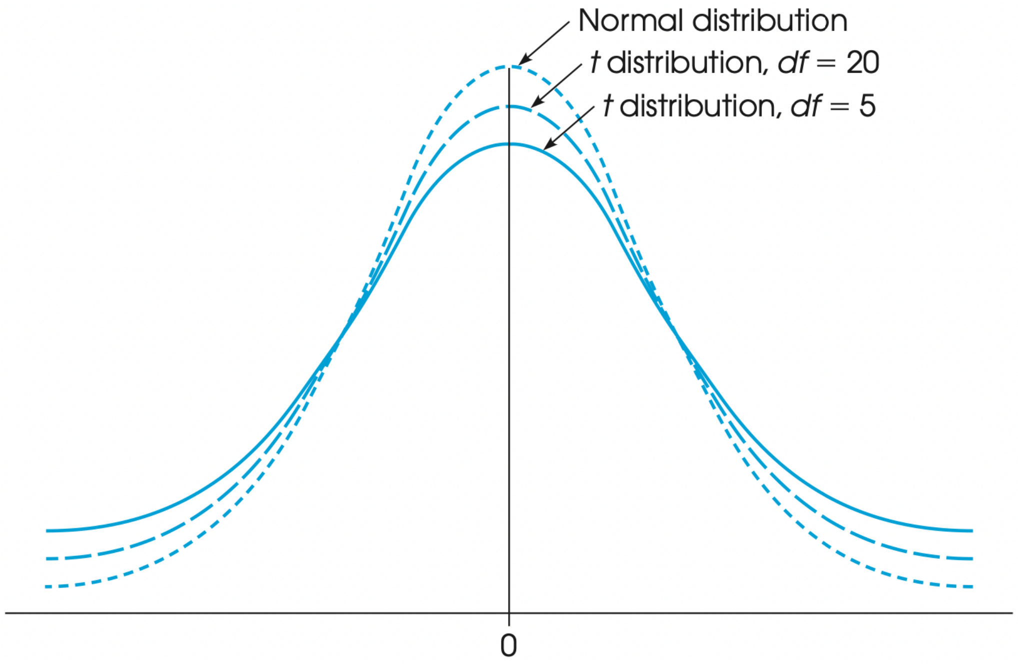
$$s^2 = \frac{\sum (X_i - \bar{x})^2}{n-1}$$

$$s^2 \mathrel{\hat{=}} \sigma^2$$

$\hat{=}$ means "is an unbiased estimator of"

t-distribution

t distribution: the complete set of *t* values computed for every possible random sample for a specific sample size, n. The *t* distribution approximates the shape of a normal distribution, especially for large samples or samples from a normal population.



Main point: the *t*-distribution resembles a normal distribution with sufficiently large samples

Estimating Standard Error

Estimated standard error: $s_{\bar{x}}$ is an estimate of the real standard error, $\sigma_{\bar{x}}$, when the value of σ is unknown. It provides an estimate of the standard distance between a sample mean, $\bar{x}$, and the population mean, μ.

Population

True standard error

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Sample

Estimated standard error

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

One-sample *z*-test

z-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

Standard Error

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

One-sample *t*-test statistic

***t* statistic**: is used to test hypotheses about an unknown population mean, μ, when the value of σ is unknown. The formula for the *t* statistic has the same structure as the *z*-test statistic formula, except that the *t* statistic uses the estimated standard error in the denominator.

One-sample *t*-test statistic

$$t_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{s_{\overline{x}}}$$

Estimated standard error

$$s_{\overline{x}} = \frac{s}{\sqrt{n}}$$

t-test statistic

General form of t-test statistics

$$t = \frac{sample\ statistic\ - population\ parameter}{estimated\ standard\ error\ of\ statistic}$$

Three types we'll learn:

- One sample t-test
- Independent samples t-test
- Dependent samples t-test

One sample *t*-test

General form of t-test statistics

$$t = \frac{sample\ statistic\ - population\ parameter}{estimated\ standard\ error\ of\ statistic}$$

General form of the ***one sample t-test statistic***

$$t = \frac{sample\ mean - population\ mean}{estimated\ standard\ error\ of\ the\ mean}$$

Lie detection demo

Mock crime video 1

Mock crime video 2

Mock crime video 3

Mock crime video 4

Results

Video	Veracity
Mock crime 1	Liar
Mock crime 2	Truth-teller
Mock crime 3	Liar
Mock crime 4	Truth-teller

Lie Detection Demo

Null hypothesis

$$H_0 : \mu = 0.50$$

Lie detection accuracy does not differ from chance.

Alternative hypothesis

$$H_a : \mu \neq 0.50$$

Lie detection accuracy is different from chance.

Alpha

$$\alpha = .05$$

Lie Detection Demo

Class Performance

$$\bar{x} = 0.56$$

$$s = 0.12$$

$$n = 40$$

One-sample *t*-test statistic

$$t_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{s_{\bar{x}}}$$

Estimated standard error

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Lie Detection Demo

$$t_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{s_{\overline{x}}}$$

$$t_{\overline{x}} = \frac{0.56 - 0.50}{\frac{0.12}{\sqrt{40}}}$$

$$t_{\overline{x}} = \frac{0.06}{.01897367} = 3.16$$

Lie Detection Demo

Hypotheses

$H_0 : \mu = 0.50$

$H_a : \mu \neq 0.50$

$\alpha = .05$

Class lie detection performance

$\overline{x} = 0.56 \qquad s = 0.12 \qquad n = 40$

$t_{\overline{x}} = 3.16 \qquad p = 0.0030$

Reject the null hypothesis
Lie detection accuracy of the class (56%) was significantly different from chance (50%), $p < .05$

Effect Size

Effect size

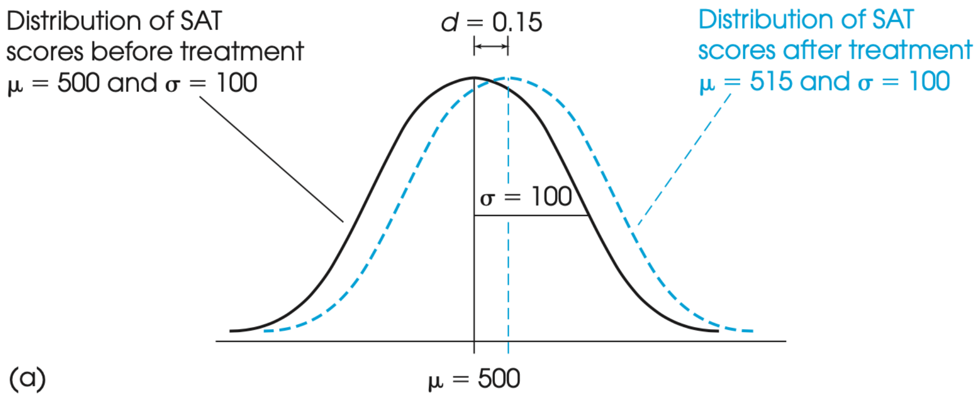
A measure of how different two groups are from one another (or a measure of the magnitude of treatment) Represented by **Cohen's *d***

- "Small" effect sizes range from 0 to .2
- "Medium" effect sizes range from .2 to .5
- "Large" effect sizes are values about .5

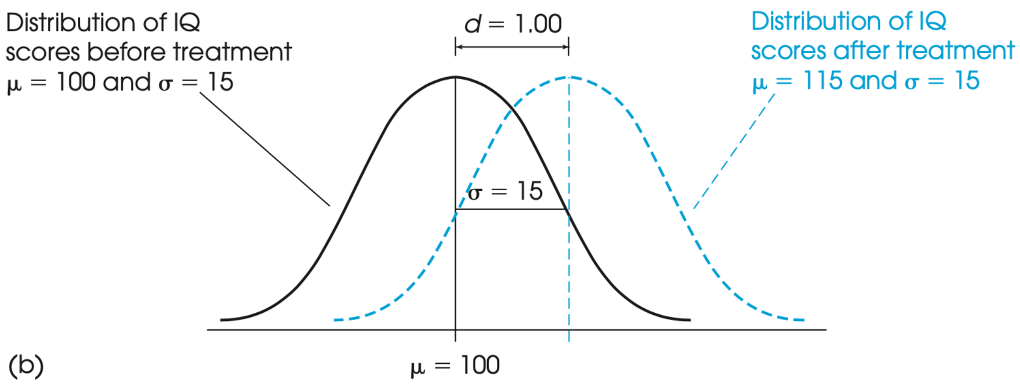
*Same ranges in the negative direction, too

Effect size

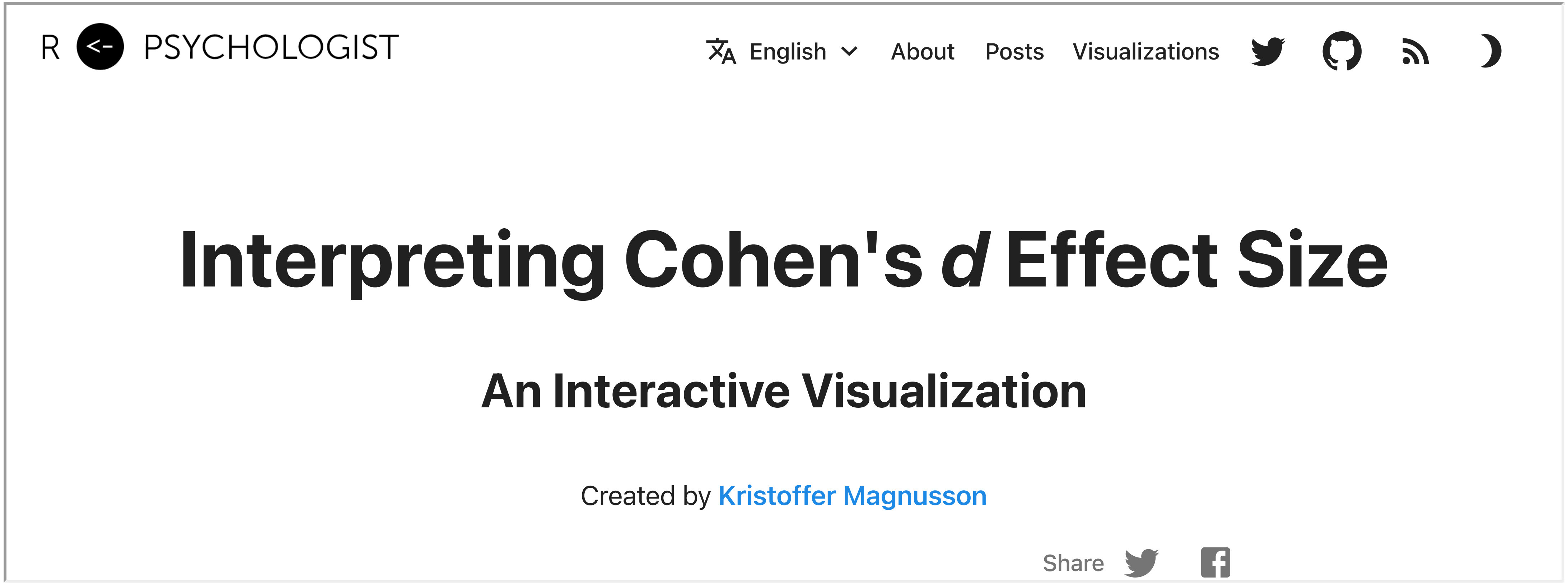
Small effect size



Large effect size



Effect Size



https://rpsychologist.com/cohend/

Calculating Effect Sizes

***z*-test effect size**

$$d_z = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma}$$

***t*-test effect size**

$$d_t = \frac{\overline{x} - \mu_{\overline{x}}}{s}$$

*Population standard deviation
(not standard error)

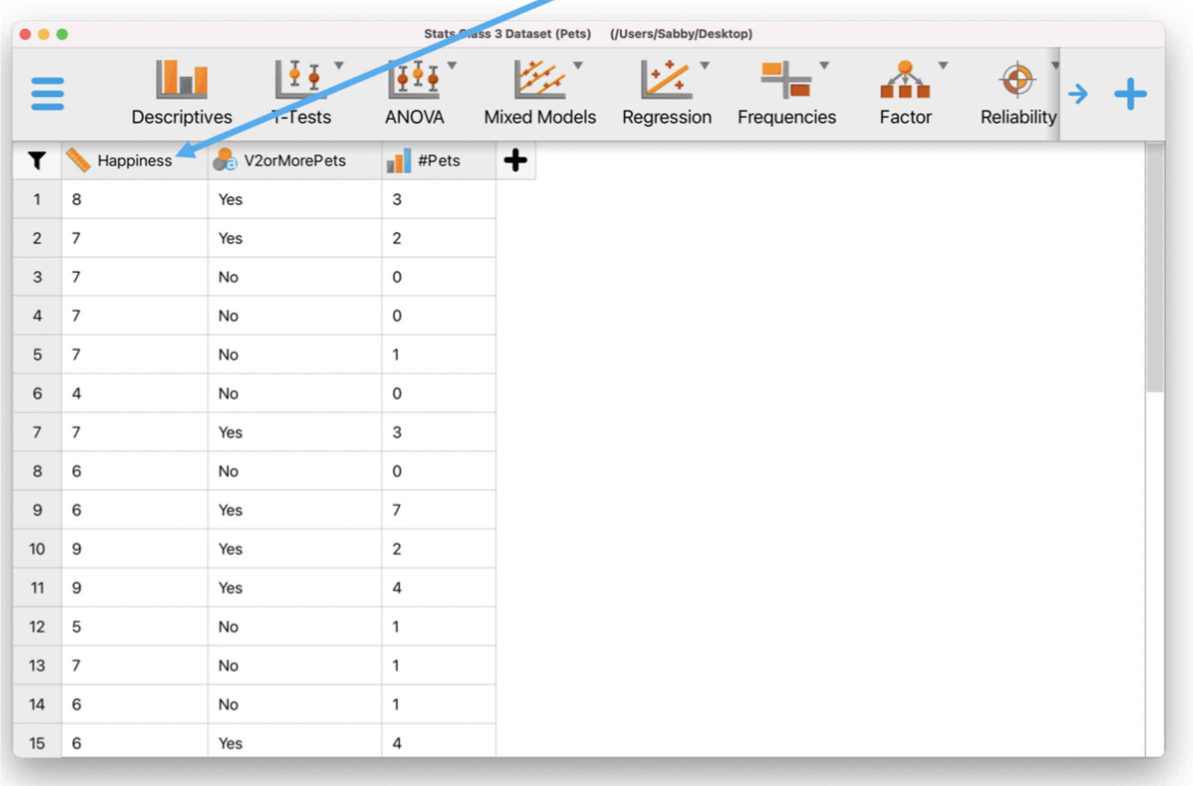
*Sample standard deviation
(not estimated standard error)

The difference is an effect size (vs. test statistic) does not take into account *sample size*

One-sample *t*-test in JASP

JASP: One-sample *t*-test

Research question from our “Pets” in-class demonstration:
Is our class happier than “average,” on a scale of 1 to 10?

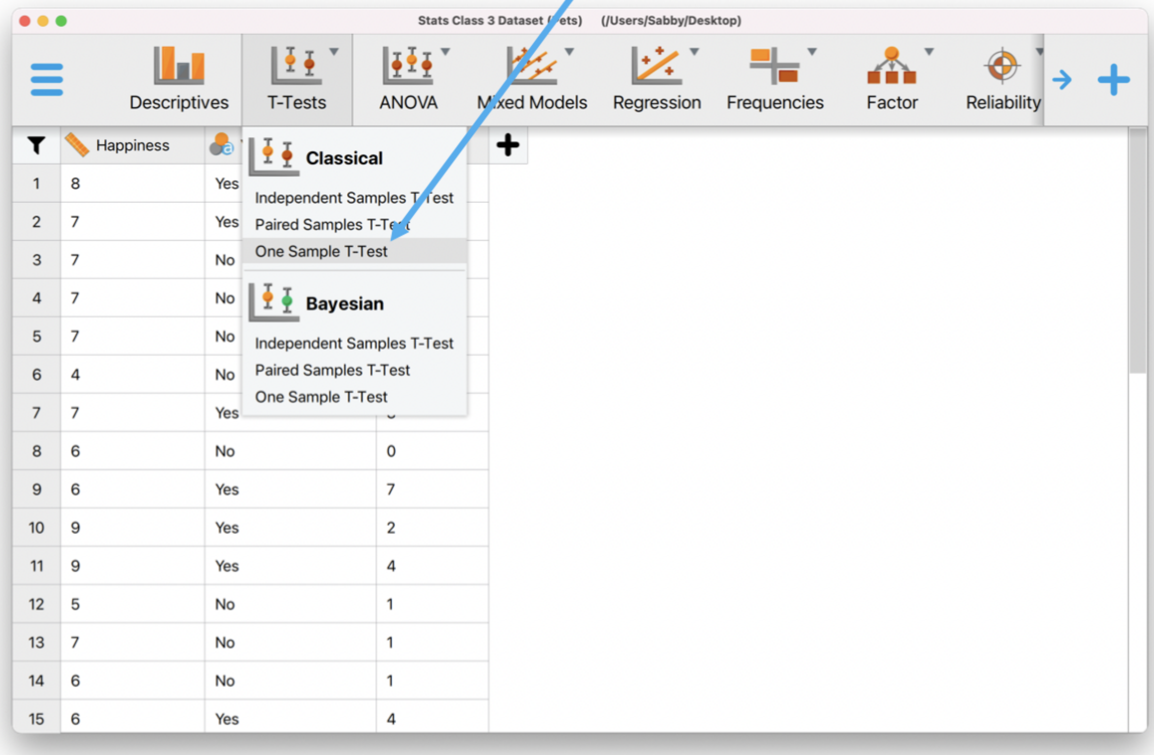


	Happiness	V2orMorePets	#Pets
1	8	Yes	3
2	7	Yes	2
3	7	No	0
4	7	No	0
5	7	No	1
6	4	No	0
7	7	Yes	3
8	6	No	0
9	6	Yes	7
10	9	Yes	2
11	9	Yes	4
12	5	No	1
13	7	No	1
14	6	No	1
15	6	Yes	4

1

JASP: One-sample *t*-test

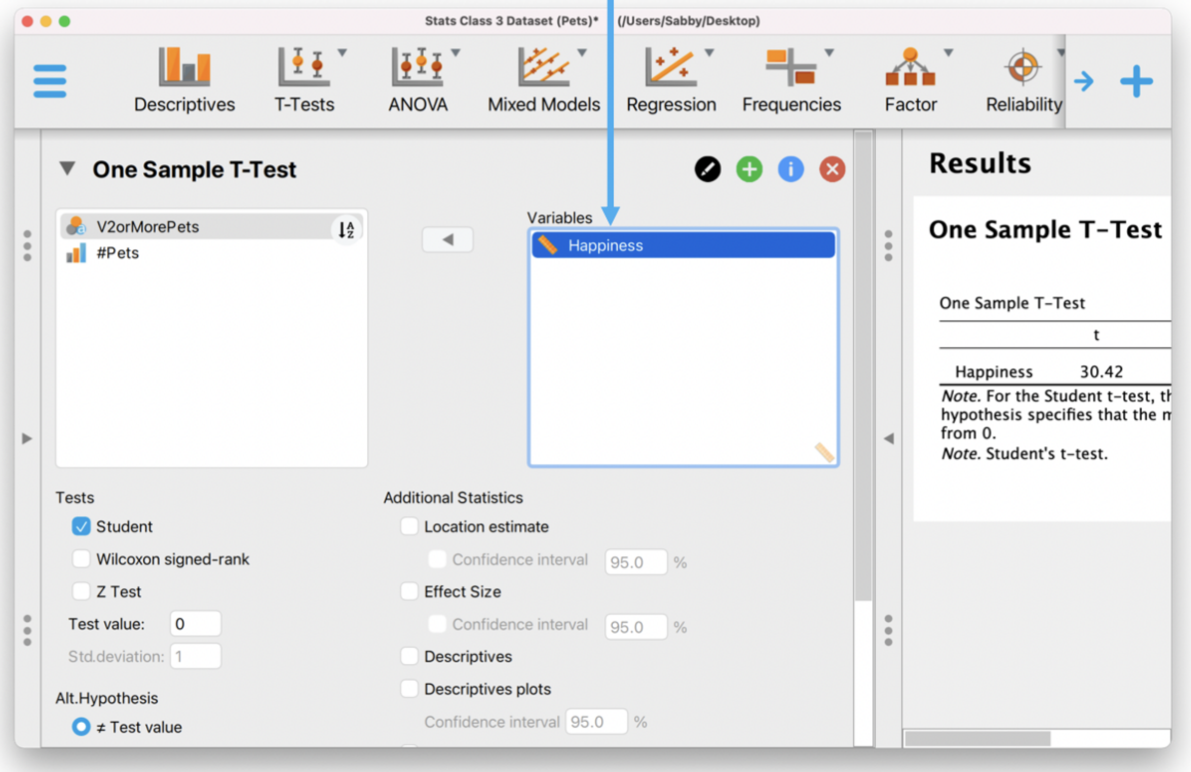
Select “T-Tests” and then click on the “One-Sample T-Test,” since we’re comparing one sample (y’all) against an “average” of 5



2

JASP: One-sample *t*-test

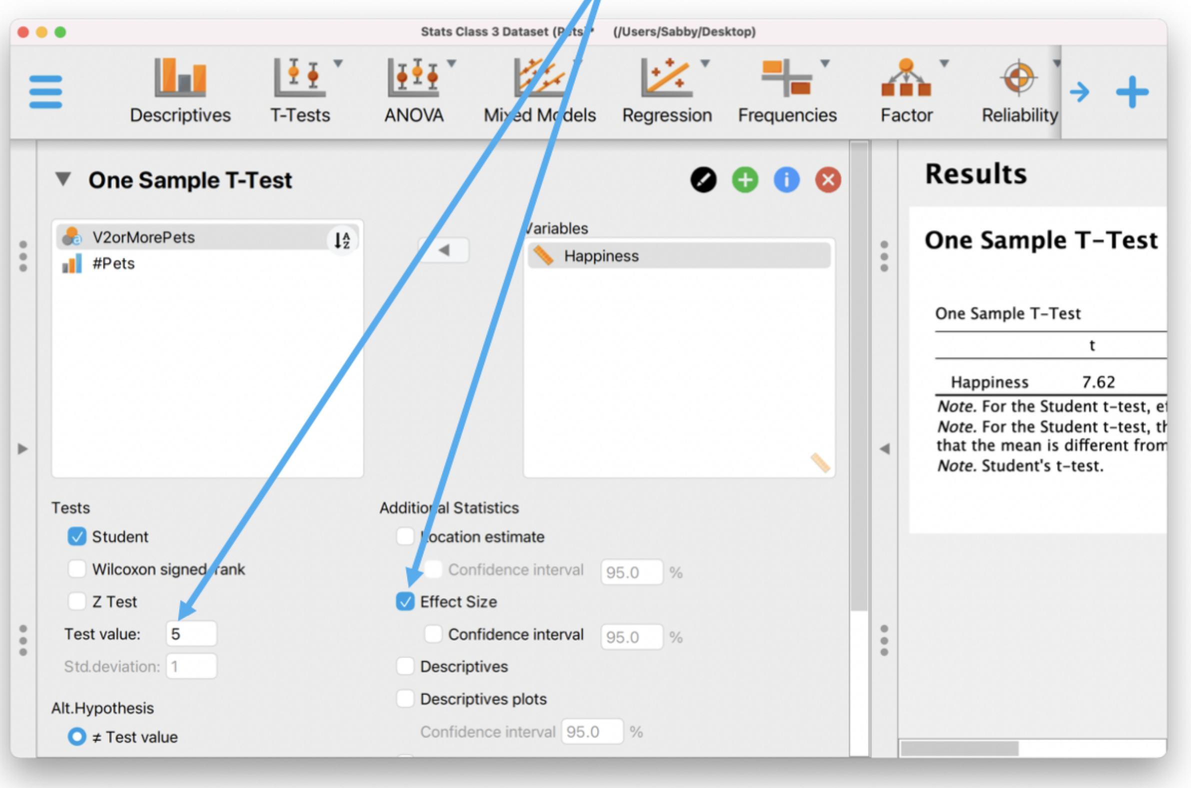
Add the variable “Happiness” to the “Variables” box to begin conducting the One-Sample T-Test



3

JASP: One-sample *t*-test

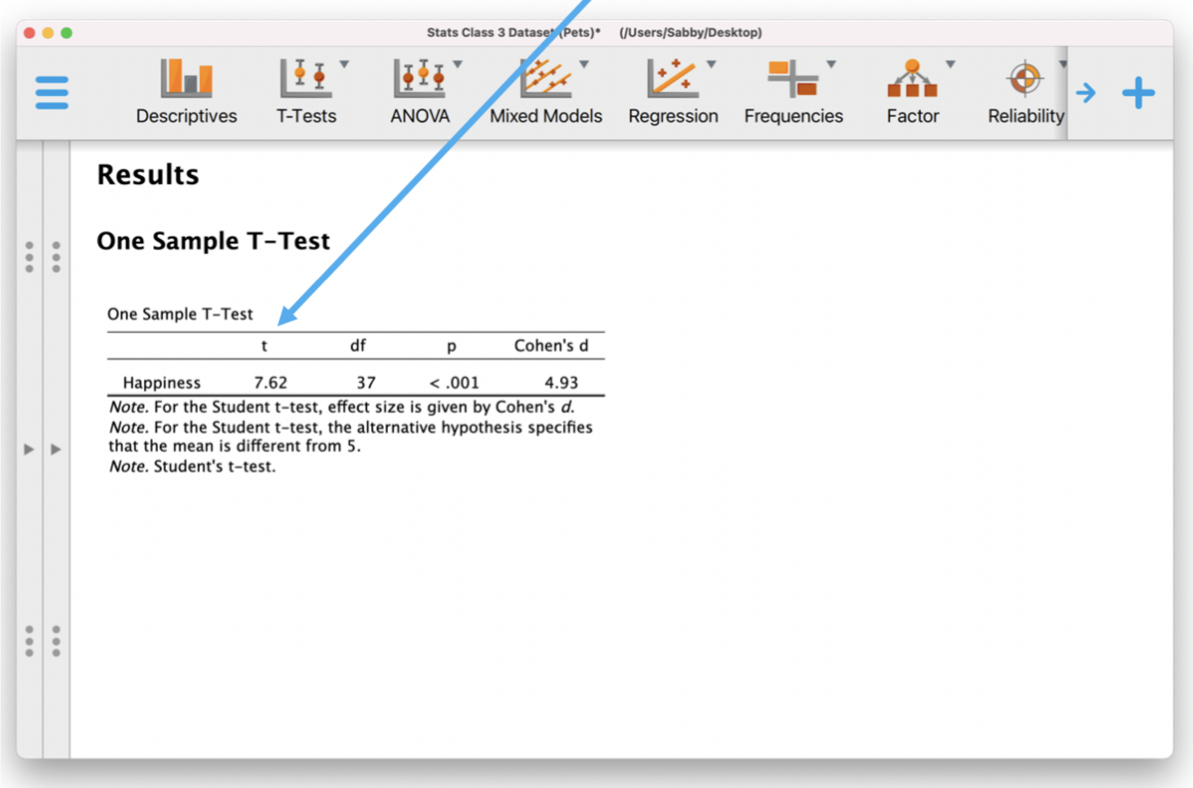
Test against 5, since that's the midpoint of our happiness scale (i.e., a kind of "average") and tick the "Effect Size" box as well



4

JASP: One-sample *t*-test

We now have our t-test statistic, a *p*-value, as well as Cohen's *d* effect size. That's all, folks! ☺



One Sample T-Test

	t	df	p	Cohen's d
Happiness	7.62	37	< .001	4.93

Note. For the Student t-test, effect size is given by Cohen's *d*.
Note. For the Student t-test, the alternative hypothesis specifies that the mean is different from 5.
Note. Student's t-test.

5

Next time

Lecture

- Differences between two groups I

Reading

- Ch.10

Quiz 3

- Due Wednesday 2/9/22 11:59pm MT (Ch.9, 10)