

# PSYCH308A - Data Analysis 4 (DA4)

Brady C. Jackson

2024/10/04

## Contents

<b>Question 01 (Q01):</b>	<b>1</b>
Answer to Q01: . . . . .	2
<b>Question 02 (Q02):</b>	<b>2</b>
Answer to Q02: . . . . .	2
<b>Question 03 (Q03):</b>	<b>2</b>
Answer to Q03: . . . . .	2
<b>Research Prompt for Q04 through Q09:</b>	<b>3</b>
Research Question 1 (RQ1): Q04 - Q06: . . . . .	3
Research Question 2 (RQ2): Q07 - Q09: . . . . .	4
<b>Question 10 (Q10):</b>	<b>5</b>
Answer to Q10: . . . . .	5
<b>Question 11 (Q11):</b>	<b>5</b>
Answer to Q11: . . . . .	5
<b>Question 12 (Q12):</b>	<b>5</b>
Answer to Q12: . . . . .	5

---

```
# Load packages. Set messages and warnings to FALSE so I don't have to see the
# masking messages in the output.
library(psych)
library(jmv)
library(ggplot2)
library(dplyr)
library(magrittr)
```

---

## Question 01 (Q01):

1. Interpret a p-value of .042. (This question is **not** asking for the decision this p-value results in, rather what does this value mean?)

### Answer to Q01:

A p-value of .042 implies that the finding of some statistical test (z-test, t-test, etc.) that the probability of obtaining the test statistic found or something more extreme was .042 (a.k.a. 4.2%) assuming the null hypothesis was true

---

### Question 02 (Q02):

2. *In no more than two sentences*, what is the relationship between sample size, effect size, and power?

### Answer to Q02:

Power is positively correlated with both sample size and effect size in a manner that convolutes the two (sample and effect). That is to say, the power of a finding will increase if either sample size increases or effect size increases, and if one (e.g. sample size) is necessarily small, a higher power can be achieved by increasing the other (e.g. effect size).

---

### Question 03 (Q03):

3. A researcher records the number of words recalled by students presented with a list of words for 1 minute. In one group, students were presented with the list of words in color; in a second group, the same words were presented in black and white. An equal number of students were in each group. The researcher reports the following: *Participants recalled significantly more words when the words were presented in color ( $M = 12.4$  words) versus black and white ( $M = 10.9$  words),  $t(48) = 2.01$ ,  $p = .035$ ,  $d = 0.18$ . Based on the previous statement, what is the sample size in each group?*

### Answer to Q03:

Given the wording of the prompt, I'm assuming this test was conducted as an independent sample t-test ("In one group ... in a second group" instead of phrasing like, "... the group of students looked at colored words and then later looked at black and white words...").

The degrees of freedom for the t-test ran is given as 48 ( $t(48)$ ).

As this is an independent t-test, two parameters had to be estimated, the mean of the population from which group 1 was drawn, and the mean of the population from which group 2 was drawn. Therefore, we have the relationship

$$\begin{aligned} df &= [Things\ we\ know\ (n_{total})] - [Estimated\ Parameters\ (n_{P,est})] \\ df &= n_{total} - n_{P,est} \\ 48 &= n_{total} - 2 \\ 48 + 2 &= n_{total} \\ 50 &= n_{total} \end{aligned}$$

Therefore, there were a total of 50 students in the study, with 25 in each group given that the prompt states the same number of students were in each group.

---

## Research Prompt for Q04 through Q09:

You are teaching your first Intro to Psychology course! After the midterm, you are disappointed with your students' overall test scores. You decide to implement two different required study techniques. There are 100 students in the class; 50 of them will be required to meet in groups to study right before the final (Group A) and the other 50 will be required to create flashcards to aid in memorization (Group B). You are interested in two primary research questions:

---

### Research Question 1 (RQ1): Q04 - Q06:

**Did student test scores improve significantly from the midterm to the final?**

**Data:** *308A.RQ1 Data.DA4.csv*

---

#### Question 04 (Q04):

4. Visualize your data for this research question. Include your visualization here.

#### Answer to Q04:

**Loading Data for Q04** The first step to visualizing the data is to load it. See code below:

```
# In order to visualize the data we must first load it. To do so we look  
# in the present working directory for CSV files, and take the one that  
# has RQ1 in its filename. Append it to the current working directory  
# to create the fullpath for loading  
here <- getwd()  
rq1_name <- list.files(here, pattern = ".*RQ1.*csv")  
  
# Use the file.path function to create a platform appropriate fullpath / filename  
# to the Research Question 1 data.  
rq1_file = file.path(here, rq1_name)  
  
# Read the CSV data in as rq1_dat  
rq1_dat = read.csv(rq1_file, header = TRUE)  
  
# Lower-case all column names for convenience  
colnames(rq1_dat) <- tolower(colnames(rq1_dat))
```

**Data Prep Q04** Ultimately we're going to answer Research Question 1 as a dependent T-test since we're evaluating change in student test scores over time (regardless) of study method. So we're going to prepare the data further with some math. See code below:

```
# Need to delta the final and midterm exam scores
```

**Visualize Prepared Data Q04** See the histograms below for visualization of the data pertinent to this research question.

*# Use descriptives here*

---

**Question 05 (Q05):**

5. Did student test scores improve significantly from the midterm to the final? Using RStudio to analyze, conduct a hypothesis test to evaluate this question. Organize your answer according to the 4 steps of hypothesis testing.

**Answer to Q05:** This is my answer

---

**Question 06 (Q06):**

6. Report your findings in APA format. (Hint: make sure to answer the research question!)

**Answer to Q06:** This is my answer

---

**Research Question 2 (RQ2): Q07 - Q09:**

**Does the study technique used predict scores on the final exam?**

**Data:** *308A.RQ2 Data.DA4.csv*

---

**Question 07 (Q07):**

7. Visualize your data for this research question. Include your visualization here.

**Answer to Q07:** This is my answer

---

**Question 08 (Q08):**

8. Does the study technique used predict scores on the final exam? Using RStudio to analyze, conduct a hypothesis test to evaluate this question. Organize your answer according to the 4 steps of hypothesis testing.

**Answer to Q08:** This is my answer

---

**Question 09 (Q09):**

9. ...
- a. Report your findings in APA format. (Hint: make sure to answer the research question!)

- b. The Dean of the university was also interested in your results, as this may help to raise scores in other departments. Unfortunately, she does not understand statistical language. Please interpret your findings for the Dean. Did scores improve? Which technique is better?

**Answer to Q09:** This is my answer

---

### Question 10 (Q10):

10. A developmental psychologist is interested in the effect of a positive psychology intervention on the well-being of aging adults. She administers the intervention, collects well-being scores from a sample of 100 participants, and tests whether their well-being differs significantly from the national average. Using G\*Power, she determines the power for her test is .80.
- a. Interpret this value
- b. What suggestion would you give her if she wants a higher probability of detecting a true effect?

**Answer to Q10:**

This is my answer

---

### Question 11 (Q11):

11. What would it mean if your analysis returned the following values? Consider the meaning of  $t$  - not the decision associated with it.
- a.  $t(24) = 0.35$
- b.  $t(24) = 1.00$
- c.  $t(24) = 3.2$

**Answer to Q11:**

This is my answer

---

### Question 12 (Q12):

12. Draw and annotate all the properties of the null and alternative curves: power, beta, alpha, type 1 error, type 2 error.

**Answer to Q12:**

This is my answer