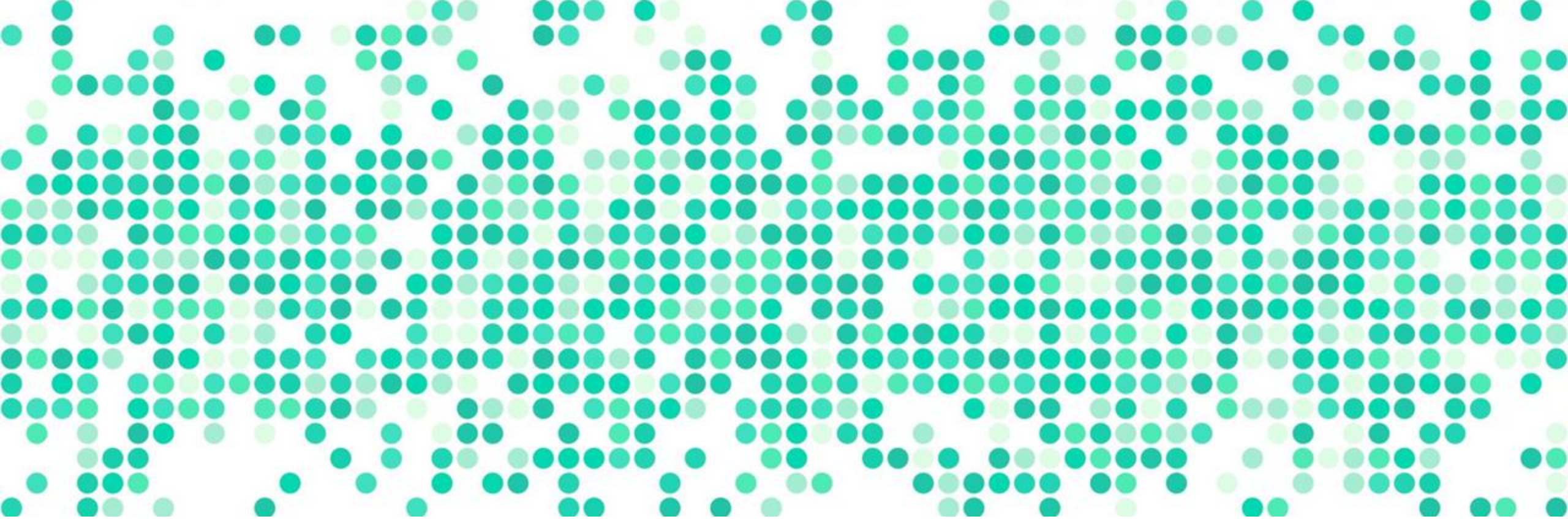


DATA ANALYSIS

Week 14: Repeated Measures

logistics: extra credit assignment (+2)

Extra Credit: Analyze a Research Paper!

Published | Assign To | Edit

This OPTIONAL assignment provides an opportunity to receive up to 2 extra credit points that will count towards your final grade in this course. In this assignment, you will identify and analyze the statistical methods from a research paper of your own choosing. Grades for this assignment will be determined as follows:

- Excellent analysis: 2 points
- Acceptable/adequate analysis: 1 point
- Incomplete or incorrect analysis / inappropriate paper: 0 points

Below are the guidelines for this assignment.

Your assignment submission should be **no more than 2 single-spaced pages and address the following aspects of the paper in clearly outlined sections:**

- Find a peer-reviewed research paper published in a legitimate scientific journal.
- Describe the basic research design and methods of the study you will be analyzing.
- Describe the key findings reported in the paper.
- Analyze and interpret the statistics reported in the paper. Your analysis must cover the following:
  - Sample characteristics (who is being analyzed?)
  - Nature of data (DV/IV, scale of variables, is the data openly available?)
  - Key statistic(s) reported and their interpretation (was it appropriate? what do the numbers mean? what are the assumptions?)
  - Key figure(s) reported and their interpretation (was it appropriate? what do the numbers mean? what are the assumptions?)
  - Statistical and practical significance of the result(s)? (Do you believe the conclusions are valid and/or important?)

Submit your analysis in PDF format.

logistics: PS policy

- 30 points total but 7 problem sets overall
- *original* policy: lowest scoring (attempted) PS dropped
- *new* policy: just add up to 30?
- **email me** if this changes your decision to opt out of problem sets

Problem Sets

Problem sets are meant to help you practice actually conducting data analysis on new datasets. Most problem sets will come from the course textbook and will involve solving problems that will resemble the kinds of questions you may see on the exams in the course. These problem sets will test your ability to work with datasets and conduct simple statistical analyses via Google Sheets. You are encouraged to look over the videos on the course website (inside each module) so that you get comfortable with using Google Sheets and using simple formulas within that platform.

Each problem set will be graded on 5 points (2.5 points per attempt) and there will be 7 problem sets throughout the semester. Your lowest scoring (*attempted*) problem set will be dropped. Problem sets will be due roughly every week and will contain questions that will require you to do basic calculations and/or analyses. They may also contain some short answer questions. Please refer to Canvas for exact deadlines - they will usually be due the Monday midnight before class.

Points

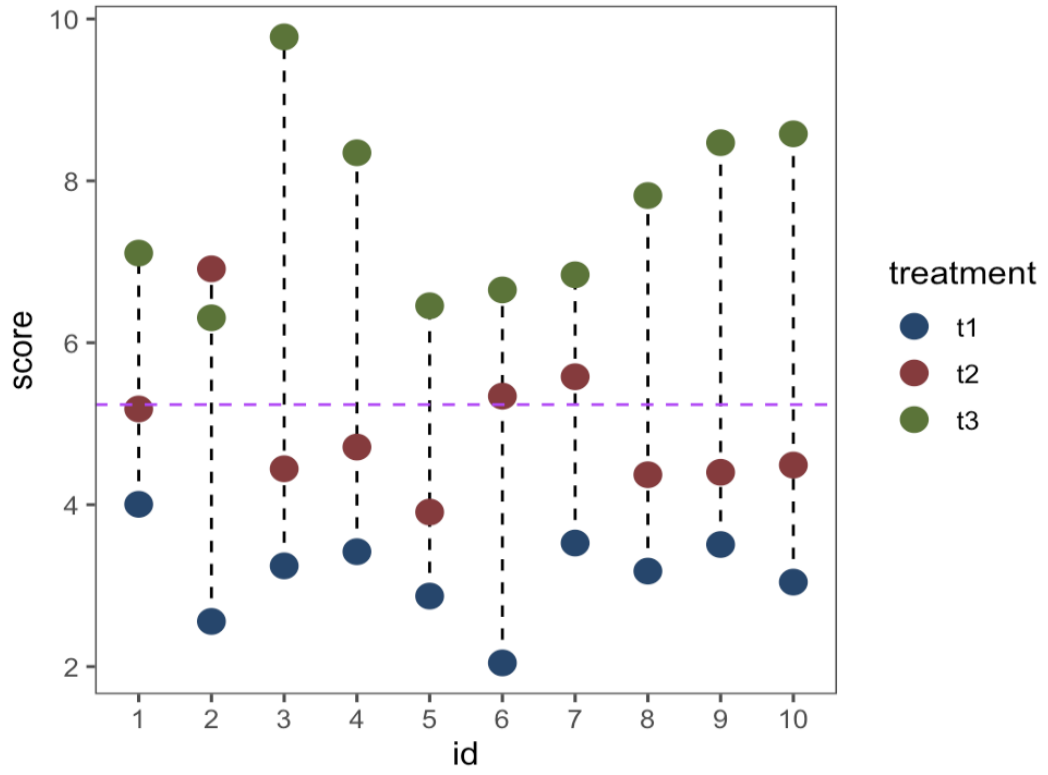
Component	Points
Weekly quizzes	10
Problem sets	30
Exam: Midterm 1	15
Exam: Midterm 2	15
Exam: Final	20
Class participation	10
Extra credit	5
Total	105

lingering question(s)

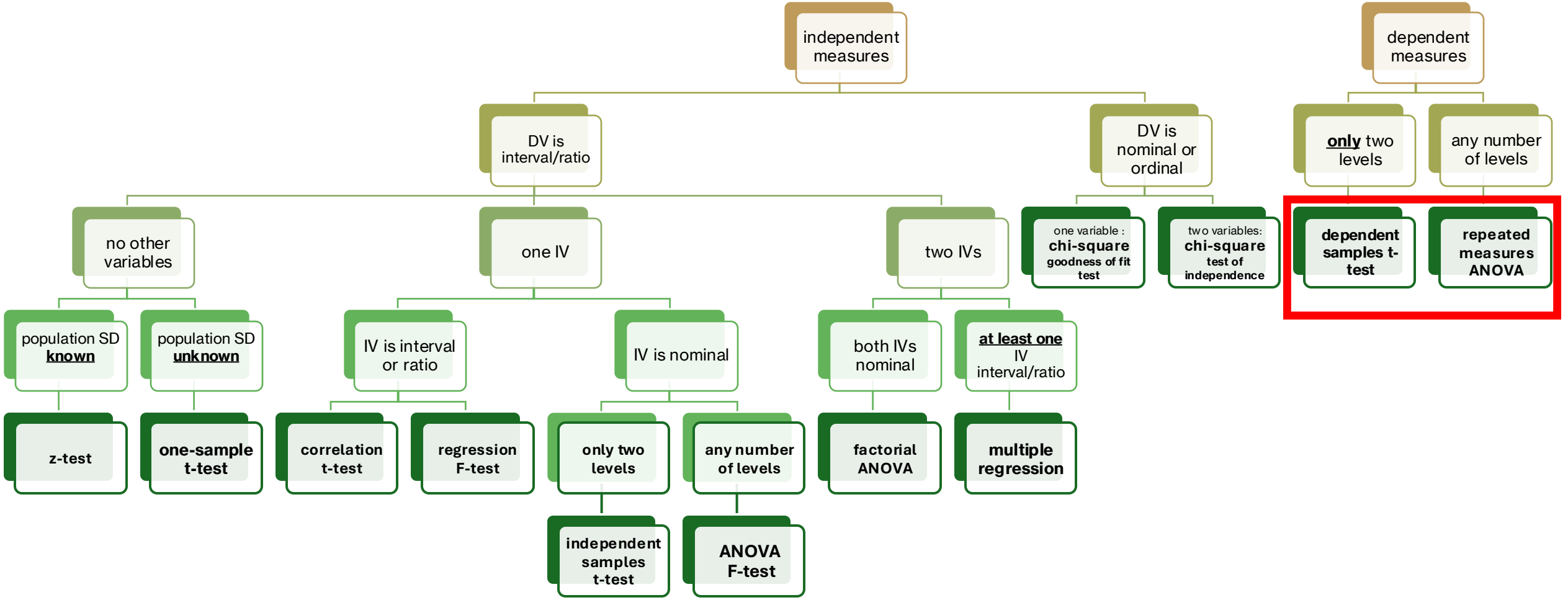
- Do we have to write a sentence explaining what a hypothesis means in the "stating the hypotheses" step? (example: the answer key for the mtcars problem says "H_0: The correlation between miles per gallon and horsepower is zero, $\rho = 0$", while a different problem just says "H_0: $\rho = 0$" for the same step
- Is there a difference between H_1 and H_a for hypothesis testing? (asking because answer keys include both)

self esteem data

- the self-esteem dataset in R contains results from an experiment comparing self-esteem scores from a group of participants who were all exposed to three different treatment conditions
- this dataset contains repeated observations from the same patient and therefore, the data are not independent
  - also called a within-subject or within-participant design
- we have not covered any statistical tests that we can use to analyze such data! ☹

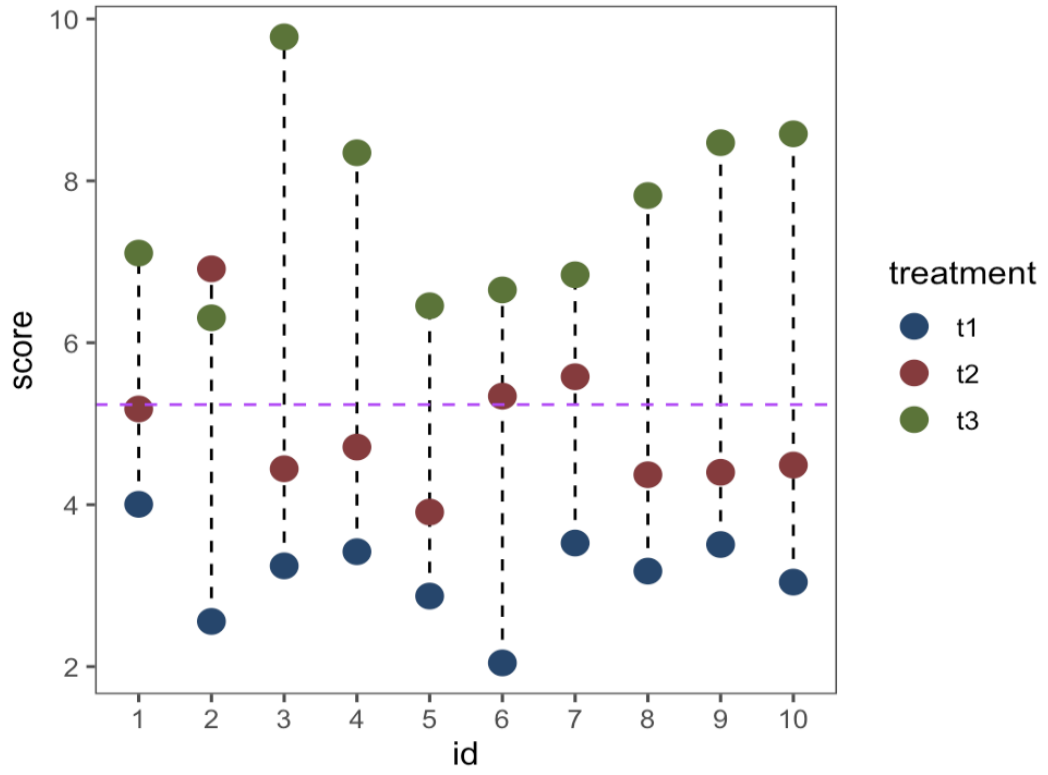


hypothesis testing flowchart



self esteem data

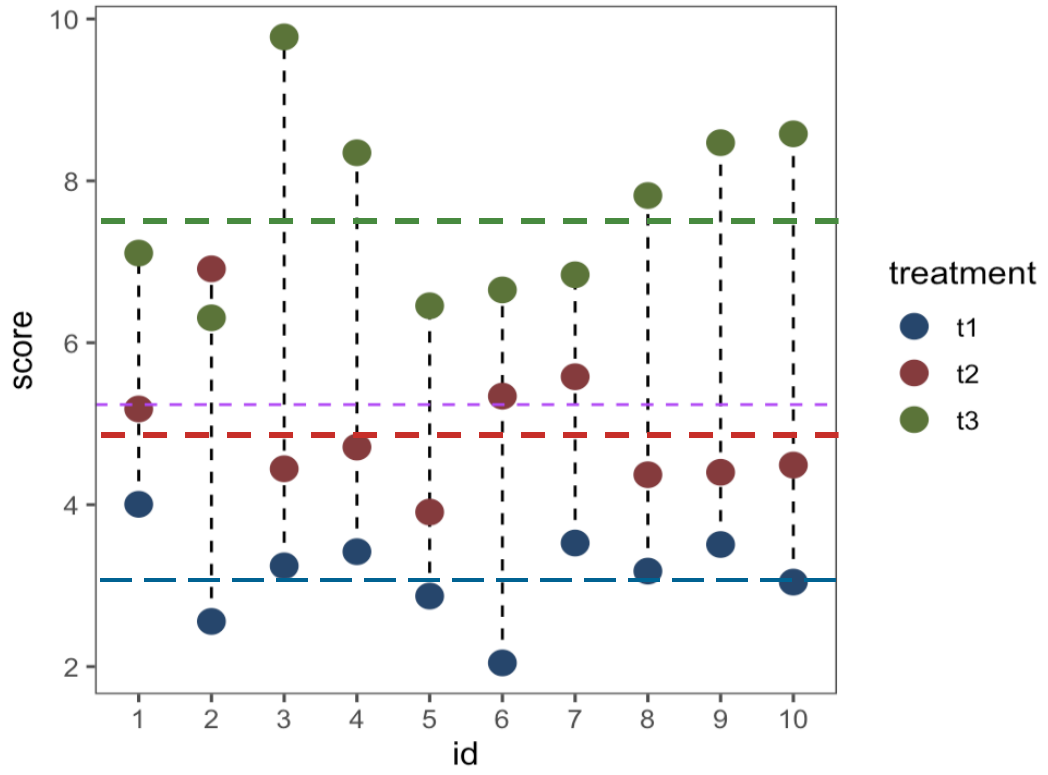
- **research question**: are there differences in self-esteem scores across treatments?
- how do we start building a model?
- how would we proceed if the data were coming from independent samples? which test/model would be appropriate?



repeated-measures F-test

- **step 1: grand mean model**
  - compute grand mean = M_y = 5.2368
  - obtain $SS_{total} = \sum(Y - M_y)^2$ = 123.65
- **step 2: treatment mean model**
  - get $\hat{Y}$ by substituting treatment means
    - $M_{t1} = 3.14, M_{t2} = 4.93, M_{t3} = 7.64$
  - obtain $SS_{treatment\_error} = \sum(Y - \hat{Y})^2$= 21.19
  - obtain $SS_{treatment\_model}$

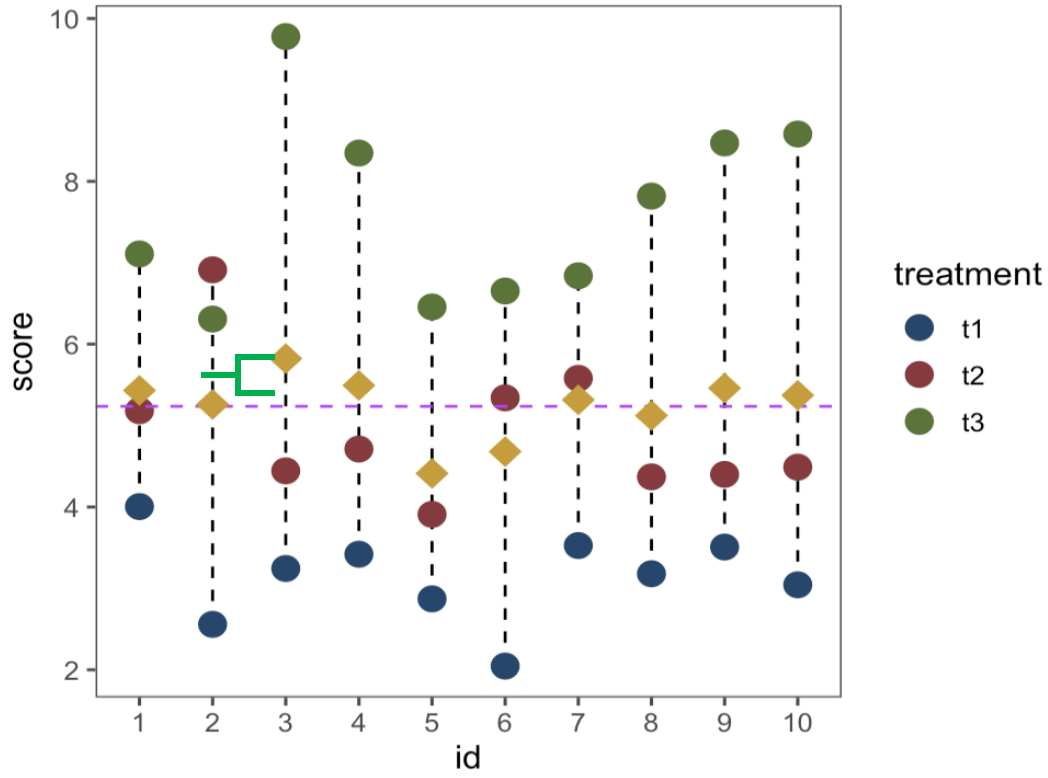
  $= SS_{total} - SS_{treatment\_error}$ = 102.46



data

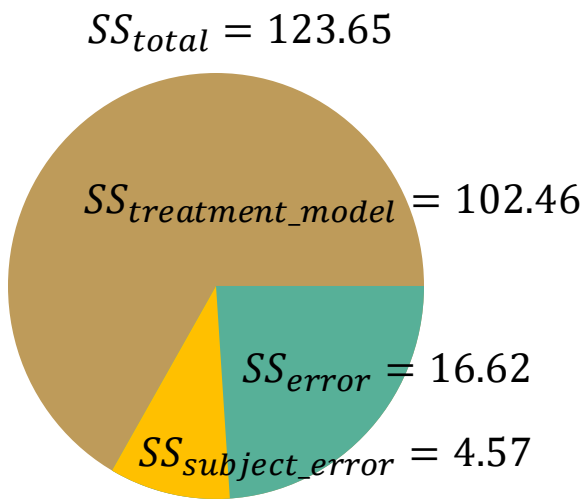
building a subject-level model

- our goal is to further reduce $SS_{treatment\_error}$ by utilizing information about the subject that is implicitly part of the treatment model
- we start by calculating a mean for each subject $M_{subject_i}$
- next, we look at how much we gain by using a subject-level mean relative to the grand mean
- we do this for ALL subjects across ALL groups in our data and then look at how much "error" is explained by this subject-level model relative to the grand mean
  - $SS_{subject} = \sum k\ (M_{subject_i} - M_y)^2$
  - k: number of levels of IV

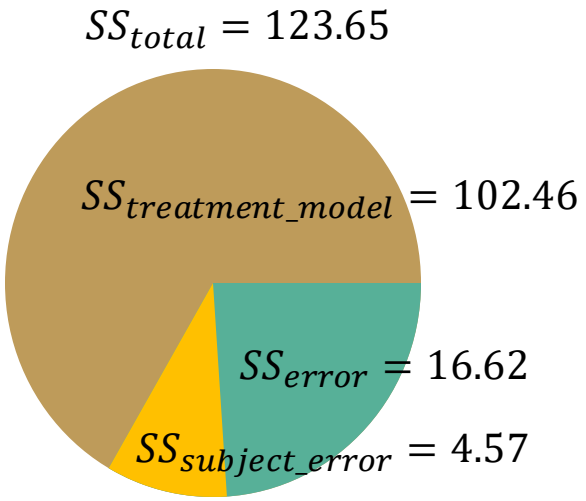


factoring out $SS_{subject}$

- step 1: from grand mean model
  - $SS_{total} = \sum(Y - M_y)^2$ = 123.65
- step 2: from drug mean model
  - $SS_{treatment\_error} = \sum(Y - \hat{Y})^2$= 21.19
  - $SS_{treatment\_model} = SS_{total} - SS_{treatment\_error}$ = 102.46
- step 3: subject-level model
  - $SS_{subject} = \sum k\ (M_{subject_i} - M_y)^2$= 4.57
- step 4: remove this estimate from remaining error
  - final SS_{error} = $SS_{treatment\_error} - SS_{subject}$ = 16.62373649



F table



- n: number of observations (data points)
- $n_{subjects}$: number of subjects or participants

	SS	df	MS	F	p-value
between-subjects (treatment)	102.46	$k - 1 = 3 - 1 = 2$	51.23	55.47	<.001
within-subjects					
• subject error	4.57	$n_{subjects} - 1 = 10 - 1 = 9$			
• residual error	16.62	$(k - 1)(n_{subjects} - 1) = (2)(9) = 18$	0.92		
total	123.65				

Sheets solution
Video tutorial

NHST for repeated measures ANOVA

step 1: build grand mean model

(1) "summarizing" data using a single grand mean (ignoring all group labels)

(2) compute

$SS_{total} = \sum (Y - M_y)^2$

step 2: build group means model

(1) find group means M_{group}

(2) compute $SS_{model\_error} = \sum (Y - M_{group})^2$

(3) compute $SS_{model} = SS_{total} - SS_{model\_error}$

step 3: build subject means model

(1) find subject-level means ($M_{subject_i}$)

(2) compute $SS_{subject} = \sum k \ (M_{subject_i} - M_y)^2$

k: number of levels of IV

(3) compute final $SS_{error} = SS_{model\_error} - SS_{subject}$

step 4: conduct F test

(1) create F table

(2) find $F_{critical}$

(3) compute $F_{observed} = \frac{MS_{model}}{MS_{error}}$

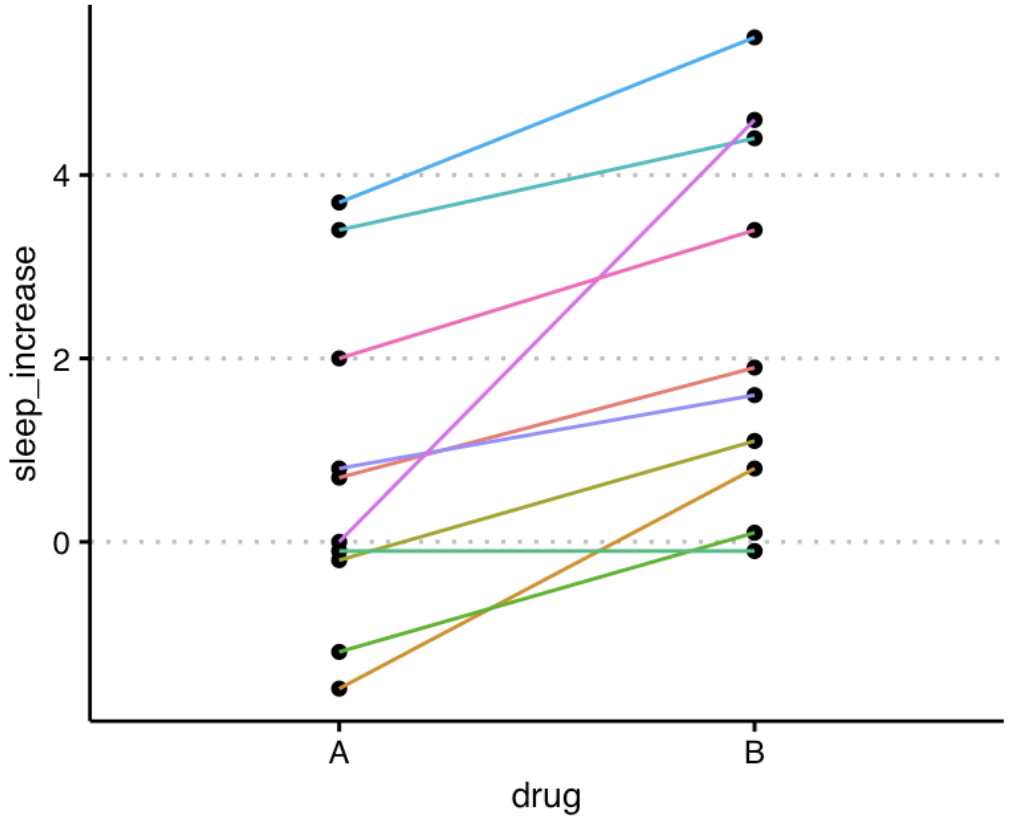
(4) find p-value for F-score

(4) decide!

questions

W14 Activity 1

- data about “the effect of two soporific drugs (increase in hours of sleep compared to control) on 10 patients”
- is there a difference in the effectiveness of the two drugs?
- conduct a repeated measures ANOVA for these data



RM-ANOVA assumptions

- interval/ratio dependent variable
- normality within each level
- sphericity: the variances of the differences between all possible pairs of within-subject conditions are equal
  - Mauchly's test is typically performed to test for sphericity

id	t1	t2	t3	t1-t2	t1-t3	t2-t3
1	4.005027	5.182286	7.107831	-1.177259	-3.102804	-1.925545
2	2.558124	6.912915	6.308434	-4.354791	-3.75031	0.604481
3	3.244241	4.443434	9.77841	-1.199193	-6.534169	-5.334976
4	3.419538	4.711696	8.347124	-1.292158	-4.927586	-3.635428
5	2.871243	3.908429	6.457287	-1.037186	-3.586044	-2.548858
6	2.045868	5.340549	6.653224	-3.294681	-4.607356	-1.312675
7	3.525992	5.580695	6.840157	-2.054703	-3.314165	-1.259462
8	3.179425	4.370234	7.818623	-1.190809	-4.639198	-3.448389
9	3.507964	4.399808	8.471229	-0.891844	-4.963265	-4.071421
10	3.043798	4.489376	8.5811	-1.445578	-5.537302	-4.091724
				var	var	var
				1.303951124	**1.155305965**	**3.081987038**

within vs. between participant designs

- both designs have advantages and disadvantages
- what are the key variables of interest in this study?
- what are some factors that would affect performance in this study?

Younger and older adults attempted to retrieve words (e.g., abdicate) from low-frequency word definitions (e.g., *to formally renounce a throne*). Retrieval was preceded by primes that were "both" semantically and phonologically related(e.g., *abandon*), phonologically related (e.g., *abdomen*), semantically related (e.g., *resign*), or unrelated (e.g., *pink*). Each participant received definitions for all 100 target words, presented in a random order, with each type of prime word shown 25 times (i.e., they received the phonological prime on 25 of the 100 trials, the semantic prime on 25 of the 100 trials, etc.). Younger and older adults both benefited from phonological primes in retrieval, and also showed reduced, but reliable, facilitation from "both" primes compared to semantic and unrelated primes. There was no difference in retrieval for semantic and unrelated primes.

within vs. between participant designs

- both designs have advantages and disadvantages
- sample size
- capturing temporal changes
- individual differences
- order/practice effects
- counterbalancing
- random assignment

Younger and older adults attempted to retrieve words (e.g., abdicate) from low-frequency word definitions (e.g., *to formally renounce a throne*). Retrieval was preceded by primes that were "both" semantically and phonologically related(e.g., *abandon*), phonologically related (e.g., *abdomen*), semantically related (e.g., *resign*), or unrelated (e.g., *pink*). Each participant received definitions for all 100 target words, presented in a random order, with each type of prime word shown 25 times (i.e., they received the phonological prime on 25 of the 100 trials, the semantic prime on 25 of the 100 trials, etc.). Younger and older adults both benefited from phonological primes in retrieval, and also showed reduced, but reliable, facilitation from "both" primes compared to semantic and unrelated primes. There was no difference in retrieval for semantic and unrelated primes.

an analysis of variance(s)

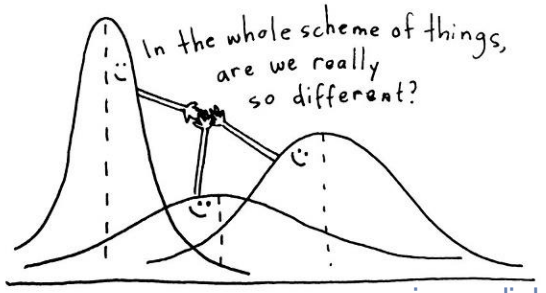
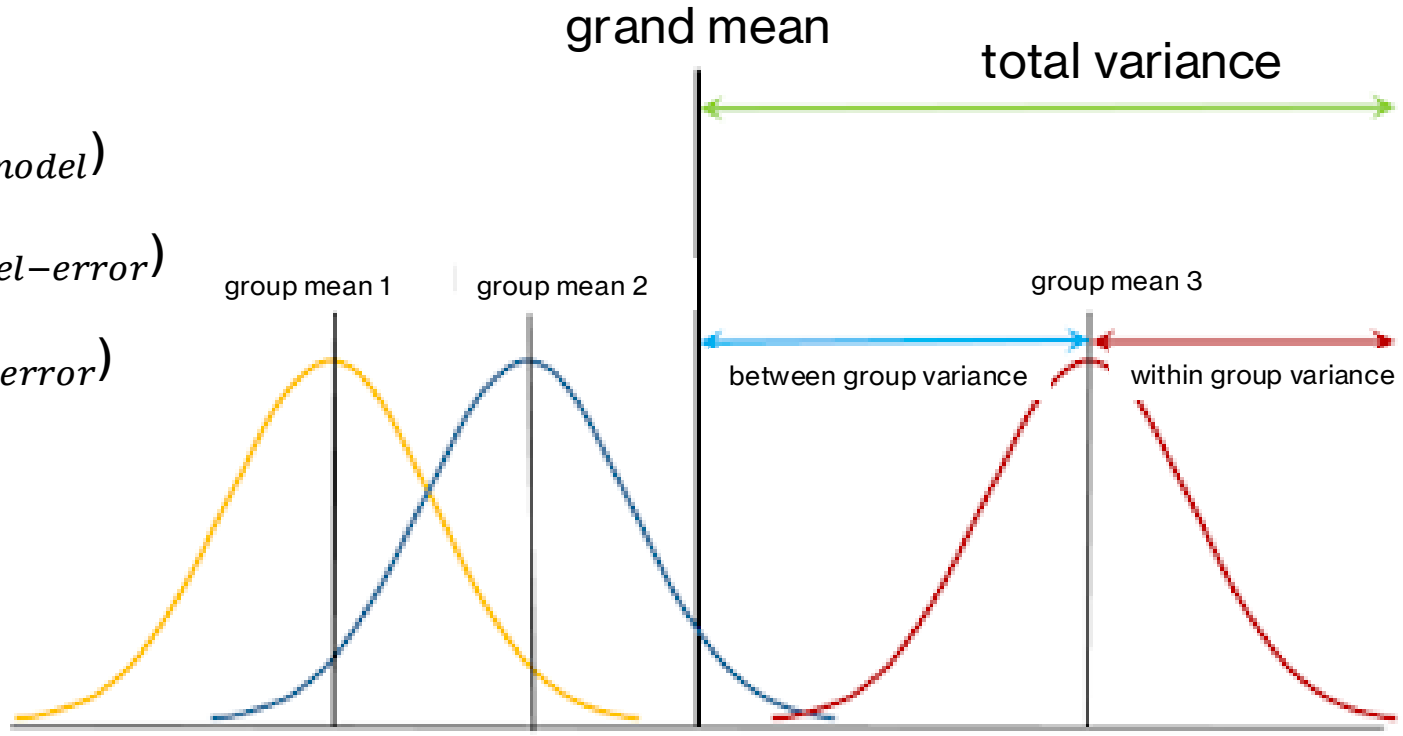


image link

- total variance (SS_{total})
- between-group variance (SS_{model})
- within-group variance ($SS_{model-error}$)
- individual variance ($SS_{subject-error}$)
- residual variance (SS_{error})

$$F_{observed} = \frac{MS_{model}}{MS_{error}}$$



next time

- dependent samples t-test (special case!)

Before Thursday

- Watch: Dependent samples t-test.
  - Practice Data
  - Solution Sheet

After Thursday

- See Apply section.

Here are the to-do's for this week:

- Submit Week 14 Quiz
- Submit revision for Problem Set 5
- Submit revision for Problem Set 6
- Submit any lingering questions here!
- Extra credit opportunities:
  - Submit Exra Credit Questions
  - Submit Optional Meme Submission