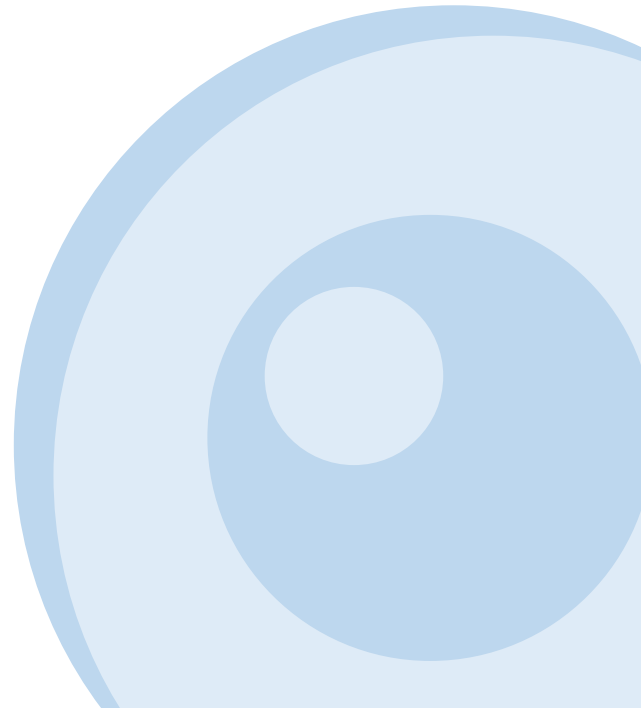


PSYC 193R:

Data Analysis & Visualization with R

Lecture 1: Statistics Concept Review

Jonas S.-H. Lau [silau -et- ucsd.edu]
psyc193r.ucsd.edu



Agenda

- Some class logistics
- Statistics Concept Review
- Preparation for the Next Class



The 10 skills in demand around the world that can get you hired in 2017



Jacquelyn Smith



Oct. 20, 2016, 10:40 AM 67,134



BUSINESS
INSIDER

TECH

FINANCE

POLITICS

STRATEGY

Here are the hottest, most in-demand skills around the globe:

1. Cloud and Distributed Computing
2. Statistical Analysis and Data Mining
3. Web Architecture and Development Framework
4. Middleware and Integration Software
5. User Interface Design
6. Network and Information Security
7. Mobile Development
8. Data Presentation
9. SEO/SEM Marketing
10. Storage Systems and Management

Essential skills for graduate school

Knowledge of the field

(Cognitive, Sensation/Perception,
Developmental, Social, Neural)

Programming Skills

(generating / presenting
stimulus, automate data
collection, collecting reaction
time data, etc)

Data Analysis

(data collection, data structure,
analysis and presentation)

+

Some Creativity

Who is this class for?

If you...

- are working / volunteering in a lab
- are planning to do grad school in psychology or related fields
- want to become a researcher in the industry, e.g., data scientist, user experience designer, marketing researcher, etc...
- or are curious to know what you can do with some statistical programming



You as a Statistics Consultant

PSYC 60 / PSYC 70

- You think you know everything
- You're given a Research Question & data set
- Find *the single* correct answer

PSYC 193R

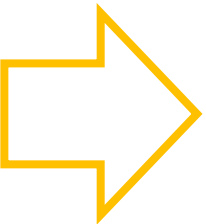
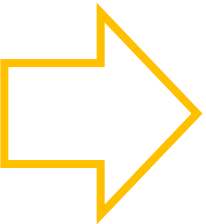
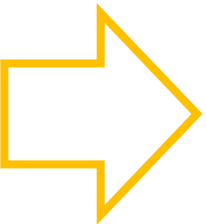
- **learning how to learn**
 - Given Research Questions
 - You will pick a statistical design & try to communicate your decision
 - Assumptions, Power & Effect Size
 - Communicating the results
- 

You'll be able to do ~70% of the common analyses

- And you will be able to learn on your own for the reminding 30%
- Online classes, books, stackoverflow.com etc



Psychology as a Science

- Based on Observation:
Falsifiable  • You can show which of the hypothesis is more plausible
[statistical test]
- Generalization:
the world is messy  • You should know how
representative your sample
it [Sampling]
- Replicable  • You can show how
reliable your
observation is
[Confident Intervals]



Class Style

- Active learning
 - Discussion-based
 - In-class experiments / quick poll
 - In-class challenges (programming problems)
- Practical
 - This is what people actually do
- Interesting
 - I'll try



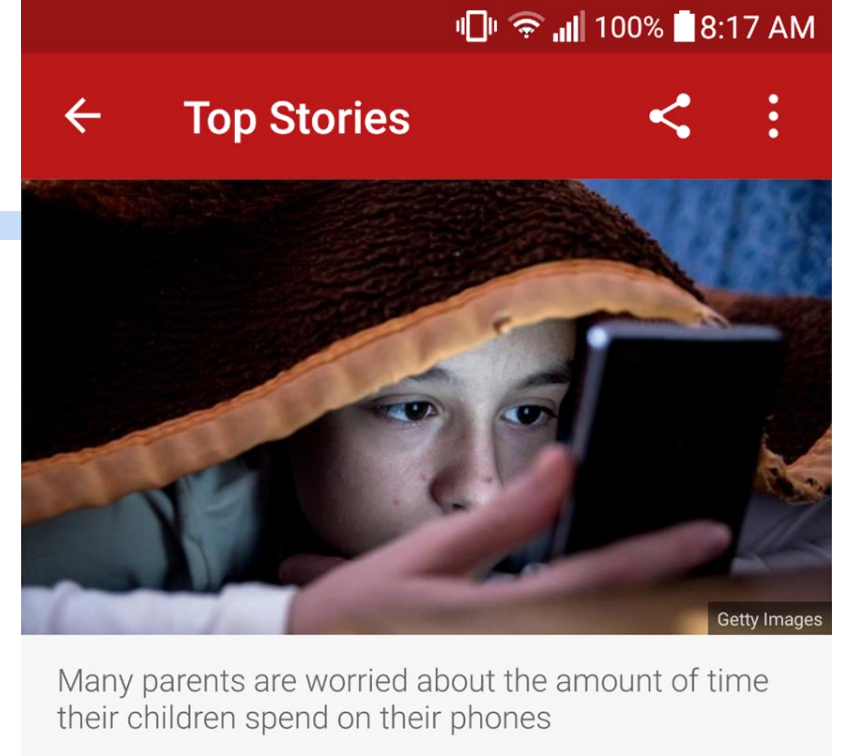
Resources

- Office Hours: Friday 2-4 @ Mandler 1586, or by appointment
- Optional Statistics Book (free for all)
Dean, S. & Illowsky, B. (2018). Collaborative Statistics.
<https://cnx.org/contents/XgdE-Z55@40.9:LnCgyaMt@17/>
- Optional Introduction to R Book (free for UC San Diegans)
Dalgaard, P. (2008). Introductory statistics with R. Springer: New York.
<http://roger.ucsd.edu/record=b6618095~S7>
- Graphing with R (free for UC San Diegans)
Chang, W. (2012). *R graphics cookbook*. O'Reilly Media, Inc.
<http://roger.ucsd.edu/record=b8037800~S9>



Cell phone / Facebook Policy

- You can survive without social media for 80 minutes
- Try not to distract others in class



Apple investors urge action on 'smartphone addiction'

1 hour ago | Technology

Big investors have called on Apple to develop software that limits how long children can use its smartphones.

The call came from two investment groups that held \$2bn (£1.48bn) of Apple stock

Best Learning Style

- Understand
- Practice
- Practice on Mixed Problem Sets (Metacognition Problem)



Statistics Concept Review

Population vs Sample

Population

- Depends on the research question
- Includes everyone of interest in your research question
 1. The rating of the US president
 2. Well-being of UC San Diego students
- Sometimes it's less obvious
 1. A drug for prostate cancer
 2. A memory drug
 3. A search for happiness

Sample

- A subset of the population
- Ideally recruited through random sampling
[or other sampling methods that best represent the population]



Variables, Values, Scores

- Score: a property of a particular person
e.g. 1, Peter is a 45: “45” is the score of Peter.
e.g. 2, Mary has 2 little lambs: “2” the score of Mary.
- Values are all possible outcomes of a Variable
e.g. 1, Values of the variable age: 0, 1, 2, ..., 110
e.g. 2, Values of gender: female, male (in the traditional sense)
- Variable: is a collection of values.
e.g. 1, Gender has two values: Female & Male
e.g. 2, Color has multiple values: depending on how you define it



Four levels of measurement

	Comparison	Arithmetic (Computable)	With Absolute Zero
Nominal Scale			

e.g., Gender, Country



Four levels of measurement

	Comparison	Arithmetic (Computable)	With Absolute Zero
Nominal Scale			
Ordinal Scale	○		

e.g., Class Standing, Birth Order



Four levels of measurement

	Comparison	Arithmetic (Computable)	With Absolute Zero
Nominal Scale			
Ordinal Scale	○		
Interval Scale	○	○	

e.g., Anxiety level, Temperature (F)

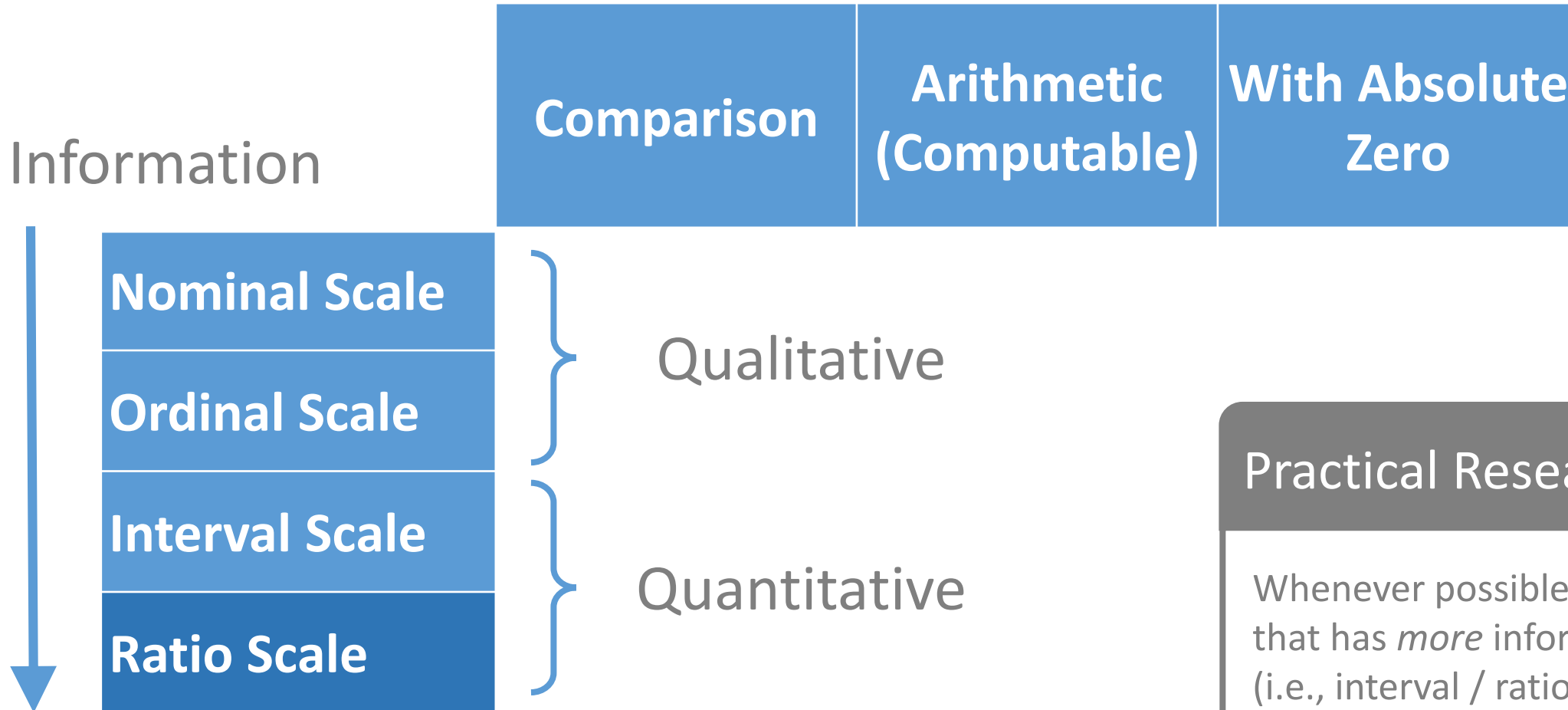


Four levels of measurement

	Comparison	Arithmetic (Computable)	With Absolute Zero
Nominal Scale			
Ordinal Scale	○		
Interval Scale	○	○	
Ratio Scale	○	○	○

e.g., Height, Amount of \$

Four levels of measurement



Practical Research Tips:

Whenever possible, collect data that has *more* information (i.e., interval / ratio scale)

Four levels of measurement

- While constructing a survey, you want to know what “life stage” people are in:

Option 1:

- Infant
- Toddler
- Adolescence
- Young adult
- Mid-life
- Senior

Option 2:

- Age

Practical Research Tips:

Whenever possible, collect data that has *more* information (i.e., interval / ratio scale)

Operational Definition

- Make the variables measureable (usually that means quantifiable)
- For example, how much you've mastered the material in a class
 - An exam, scores from 0 – 100
- How fast you can run:
 - The time needed to complete a 400m track
- Sometimes it's harder
 - Happiness
 - Health
 - Attractiveness



Random Sampling vs Random Assignment

Random Sampling

- Goal: To produce a *representative sample* of the population
- Everyone in the population has an equal chance of being selected
- So that the properties of the sample would resemble that of the population
(there are other ways of sampling)

Random Assignment

- Goal: To produce *comparable samples* for (experimental) testing



Random Sampling vs Random Assignment

Random Sampling

- Goal: To produce a *representative sample* of the population
 - Jonas wants to devise a program to cure cellphone addiction. He wants his program to work on different ages. How should he recruit his participants into the study?
 - Step 1: Define the Research Question
 - Step 2: Define the population: and look for some key properties
 - Step 3: Devise a recruitment plan that everyone in the population has an equal chance of being selected
 - Step 4: Check the properties of the sample against those of the population (e.g., age, gender, ...)
- 

Random Sampling vs Random Assignment

- Divide the 50 students into two groups using some arbitrary methods
 - Coin tossing (head → experimental, tail → control)
 - Randomly hand out stickers to the students (Sticker A → experimental, Sticker B → Control)
- The two groups then only differ by chance

Random Assignment

- Goal: To produce *comparable samples* for (experimental) testing
- Jonas wants to test whether drinking soda leads to weight gain among elementary students. He recruits 50 students, and put them into either a control or experimental group.

Experimental, Quasi-Experimental, & Observational Studies

Experimental

- Manipulation of the variable(s) of interest
- Random Assignment

Q1.
Effect of IQ on academic performance

Quasi-Experimental

- Manipulation of the variable(s) of interest
- No Random Assignment

Q2.
Soda on Weight-gain Experiment

Observational

- No Manipulation
- No Random Assignment

Q3.
Teaching Methods:
Classroom study

Experimental, Quasi-Experimental...

- Obesity → Bad memory?
- Bad memory → Obesity?
- (Something else)



Bad memory

Obesity

Practical Research Tips:

Whenever you can, you want to do an experiment to answer a research question



Obesity linked to 'worse memory'



By James Gallagher

Health editor, BBC News website

7 hours ago | **Health**

People who are obese have a worse memory than their thinner friends, a small study shows.

Tests on 50 people showed being overweight was linked to worse "episodic memory" or the ability to remember past

Independent vs Dependent Variables

Independent

- The variable being manipulated in an experiment

Dependent

- The outcome of an experiment

E.g., In an experiment that test the effect of alcohol on driving ability, 30 participants were put into 2 groups: 3 glasses of beers vs no beer. Participants performed a driving test in the simulator, which had multiple driving scenarios. The number of errors that the participants made were recorded, as well as the average reaction time each participant took when unexpected events happened.



Between- / Within-Subjects Designs

Between

- Each participant encounters one condition of the Independent variable

Within

- Each participant encounters multiple condition of the Independent variable

E.g., In an experiment that test the effect of alcohol on driving ability, 30 participants were put into 2 groups: 3 glasses of beers vs no beer.

Practical Research Tips:

If you can design an experiment using either design, a within-subjects design generally preferable

and a driving test in the simulator, which had 2 conditions. The number of errors that the participants made as well as the average reaction time each participant was recorded. Unexpected events happened.

Class Activity: Design a study

Research Question:

Does coffee drinking improve study efficiency?

Decide:

- What's the population?
 - An experiment, quasi-experiment, or observational study
 - Independent / Dependent Variables
 - Operational definitions of the Variables
 - Scale of Measurement of the Variables
 - Between- / Within-subjects design
 - What is the study like?
- 

Preparation for Next Class

- Debrief of the first Class Experiment
- Read the Syllabus
- Bring a laptop
- Install R [r-project.org] & RStudio [www.rstudio.com/products/RStudio/]
 - More instructions (optional): https://www.youtube.com/watch?v=cX532N_XLIs
- Consent for participating in in-class experiments

