

# Spring 2021 SKKU Biostats and Big data







# Lecture 09 Designing experiments!

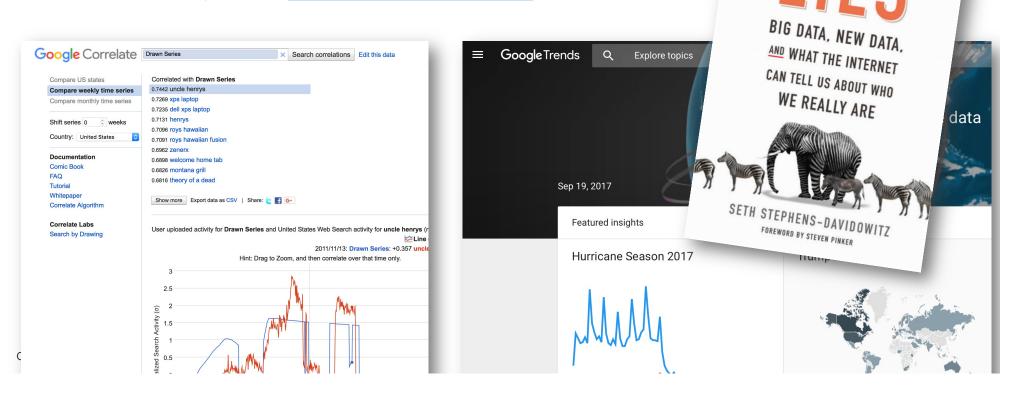






#### **Observational Studies**

- Researchers do not assign choices, they simply observe them.
- An extreme case: https://www.google.com/trends/correlate/draw?p=us
- Also Google trends: <a href="https://trends.google.com/trends/">https://trends.google.com/trends/</a>



#### **Observational Studies**

- Valuable for discovering trends and possible relationships
- Widely used in public health and marketing
- Often retrospective (retrospective study <-> prospective study)
- Lots of lurking variables
- Noise, errors
- correlation ≠ causation



https://twitter.com/tomfishburne/status/455486271339511808







#### Randomized, Comparative Experiments

- Experiments require:
  - Random assignment of subjects to treatments
  - Manipulate explanatory variables (or factors)
  - Measure response variables
- Four principles of experimental design
  - Control
  - Randomize
  - Replicate
  - Block







## Experimental design: Control

- Control sources of variation other than the factors we're manipulating
  - Making conditions as similar as possible for all treatment groups
  - To reduce the variability of the responses
- Bad for generalizability?
  - Consideration: generalizability and ecological validity







## Experimental design: Randomize

- To equalize the effects of unknown or uncontrollable sources of variation
- Not eliminating the effects of these sources, but distributing them equally across treatment levels
- Protect against factors we might be missing
- "Control what you can, and randomize the rest"







#### Experimental design: Replicate

- Subject-level replication
  - Repeat treatments to many subjects
  - The outcome of an experiment on a single subject is an anecdote (not data)
- Study-level replication
  - The entire experiment is repeated on a different population of experimental units.
- Within-subject level replication
  - Repeated measures
  - Can measure the reliability of a measure
  - Increase signal-to-noise ratio







#### Experimental design: Block

- There are things that we are not studying, but we can't control (e.g., male, female).
- We can group similar individuals together (block), and randomize within each of these blocks.
- This is a compromise between randomization and control, and this is not required, different than other three, Control, Randomize, Replicate.
- This is an idea similar to stratified random sampling.
- This design is called "Randomized block design"







### Blinding

- Humans are susceptible to errors in judgement (e.g., preference in cola tasting)
- Blinding the participants to the treatment could prevent potential biases.
- Experimenters can subconsciously provide contextual cues related to treatment.
- Two classes of individuals who can affect the outcome of the experiment:
  - Those who could **influence** the results (e.g., subjects, treatment administrators, technicians, etc.)
  - Those who evaluate the results (e.g., judges, treating physicians, etc.)
- Single-blind: either one of these classes are blinded
- Double-blind: both classes are blinded



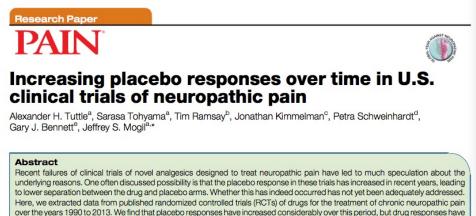




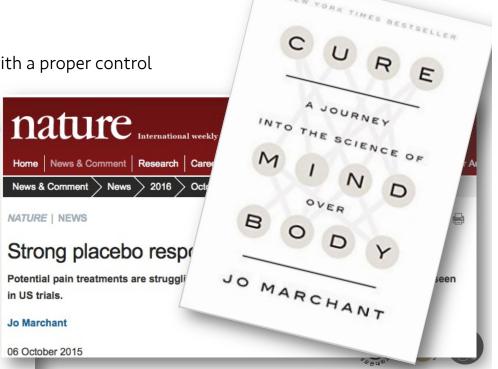
#### **Placebos**

- Simply applying any treatment can cause an improvement.
- Placebo: "fake" treatment that looks just like the real treatment
- Placebo effects tells us
  - The importance of effective blinding
  - The importance of comparing treatments with a proper control

(e.g., active control)

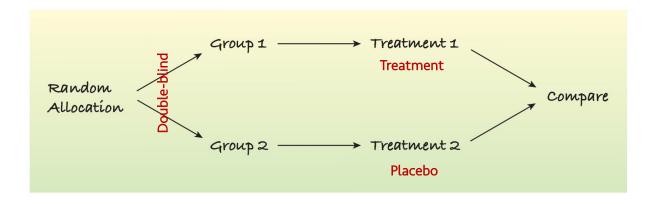


remained stable, leading to diminished treatment advantage. This trend has been driven by studies conducted in the United States.



### Best experiments:

- Randomized
- Comparative
- Double-blind
- Placebo-controlled









## Confounding

- Two factors are **confounded**: when the levels of one factor are associated with the level of another factor.
- No way to tease apart the two effects:
  - E.g., When examining the effects of different teaching styles, one class was during the fall semester, and the other class was spring. In this case, different teaching style was confounded with the season.
  - E.g., what if two different professors taught two different classes? Different teaching styles was confounded with professors and the seasons.
  - E.g., When examining the differences in brain signal for no pain and pain runs, and if no pain runs were always followed by pain runs, the pain manipulation was confounded with the order effects.
- It's super important to use the "control" and "randomize" strategies wisely.







#### **Key Points**

#### Chapter 13: Experiments and Observational Studies

- Observational studies: Valuable for discovering trends and relationships, but correlation ≠ causation
- Four principles of experimental design: Control. Randomize, Replicate, Block
- Good experiments: Randomized, Comparative (control), Double-blind, Placebo-controlled
- Confounding





