wrapping up

Adam Okulicz-Kozaryn

adam.okulicz.kozaryn@gmail.com

this version: Tuesday 3rd December, 2024 17:25

outline

reg ps comments

- see in ps5.pdf

general thought about research

- research is hard! specifically, overwhelming and time consuming
- it's normal, and the only way to persist is to study something you're passionate about
  - and then (after long time, months, or years), it will pay off, don't give up
- take a break, don't get burnt out
- don't keep it in a drawer, share it! go to conferences, post it online, try to publish it

specific ideas

- no relationship in social science is bivariate
- begin with main bivariate relationship of interest between the main Independent Variable of interest (IV) affecting the Dependent variable (DV): IV-¿DV, that's your hypothesis, say unemployment-¿crime
  - but there are other predictors of crime and retention–thats where the literature comes in!
  - you have to take the additional other variables into account! do descriptive stats such as line graphs and scatterplots of these other variables as well

specific ideas

- and include additional predictors as control variables in your regression model-many folks confuse IV and DV, be clear what affects what!
- be clear about the unit of analysis! what do you study? what is in each row of your spreadsheet? persons, municipalities, counties, schools, states?

data, data everywhere

- again see: www.economist.com/node/15557443
- Wheelan (2013) discusses uses of data, eg:
- Target predicts better pregnancy of your daughter
  - she buys unscented lotions, vitamins, etc [p252-3]
-
- pub adm application: geolocated tweets in same loction about potholes, or food making you sick
- conclusion (p.240-254 Wheelan, 2013)

so what?

- use data! (do stats)
- or read about using it (lit rev)
  - AND ALWAYS think about it! (critique research)
  - this is *important* for final project in this class
  - and use stat software (Python, etc): a job skill!

remember stats is positive, not normative

- it says what it is
- not what it should be
- for the latter we need something like philosophy or religion
- https://en.wikipedia.org/wiki/Positive_statement
- research can help evaluate damage from say

  cigars v cigarettes
- but cannot tell us what to do about it

be skeptical

- eg correlation $\neq$ causation
  - MMR vaccine, autism (p245,246)
- also: measurement
  - many ways to measure the same thing
  - no measure is perfect
  - all measures oversimplify
- eg: teacher ratings, school ratings (p246-249)

do experiments!

- again, experiment is the gold standard
  - (superb internal validity, but usually poor external)
- eg: force Indian teachers to show up by recording them
  - randomly assign cameras (p250)

Mohr, L. B. (1995): Impact Analysis for Program Evaluation, Sage, Beverly Hills CA, second edition ed.

Shadish, W. R., T. D. Cook, and D. T. Campbell (2002): Experimental and quasi-experimental designs for generalized causal inference, Wadsworth Cengage learning.

Wheelan, C. (2013): Naked statistics: stripping the dread from the data, WW Norton & Company.