

# tips, tricks, ethics

adam okulicz-kozaryn

adam.okulicz.kozaryn@gmail.com

this version: Friday 7<sup>th</sup> November, 2025 18:05

## outline

tips and tricks

ethics

research design again: important from now on



## outline

tips and tricks

ethics

research design again: important from now on

## have a big screen

- a big screen is key for gis work
- around \$250
  - will get 30-100% more productive
  - return on investment in couple weeks, at min wag \$15/hr
  - not just size but resolution matters!
  - $\geq$  4K, say UHD 3840x2160
- and get a nice mouse too

## google for maps; like lit rev!

- whatever you're mapping, someone has already done it!
- google and see images, say: 'nj counties contamination sites' <https://www.google.com/search?q=nj+counties+contamination+sites&tbo=isch>
- or "Philadelphia healthy stores map" (sometimes need word 'map' otherwise get pics of healthy food)
  - <https://www.google.com/search?q=philadelphia+healthy+stores+map&tbo=isch>
- get ideas, inspiration, make your map better, like lit rev

## google for shp, kml, geoPackage

- eg “what you are looking for, shapefile”
- eg “new jersey public schools, shapefile”
  - or kml, geoPackage (gpkg)
- tips:
  - may need to look for a higher level; eg NJ schools instead of Depford Twshp schools
  - if you can't find it, contact govt; eg city of Camden, state of NJ, etc—they'll be happy you're using their data
  - again, may find only traditional data and need to join

## join data

- the real value comes from joining data!
- again, a map about any single var has already been made
  - but 2 or more vars from varied sources: rare
- ton of datasets and vars out there!
  - use your creativity and imagination
  - and you'll easily come up with something that no one did

## its geo in maps, but map time too

- $POP10 - POP00$
- $(POP10 - POP00)/POP00$
- another time issue is that things change, say business cycle
- use say avg. 5-yr unemployment rate

**make maps, find things; go there see with your eyes**

- drove through MS from TX to NJ to see lowest lexp
- went to Pine Bluff AR to see fastest shrinking town
- map a place you grew up, or where you want to live etc

## outline

tips and tricks

ethics

research design again: important from now on

## actually important

- not just some second thought sidenote
- fine, do consulting, make \$, don't be broke
  - but add: potential/perceived conflict of interest/funding
  - and do at least some work that's independent

## integrity/honesty

- be explicit about problems in your data
  - eg non-joins, missing data, miscodings
- be explicit about holes in your story:
  - eg don't hide maps bc they contradict your story
  - discuss: how, why; ask audience to comment/criticize
- instead of forcing data to tell your story,
  - even be devil's advocate: try to find holes yourself
  - listen carefully; let data tell you her story!

## ethics

- everybody wants to sell something
- including academics, thinkers, and students!
  - we try to sell some idea or point of view
- nobody's 100% objective
- try to present alternative/opposite points of view
- present the whole picture
- do force yourself to be objective, because humans aren't
- see eye-opening [righteousmind.com](http://righteousmind.com)

## bad examples

- cherry picking of vars or samples or timeframes, etc
  - eg using only vars/operationalizations that fit your story
  - eg using year in which you find what you wanted to find
  - classification: playing with bins to fit your story

## outline

tips and tricks

ethics

research design again: important from now on

## quality

- GIGO: Garbage In, Garbage Out
- 'Cos it's in the computer, don't mean it's right
  - double, triple check
  - ask yourself if it makes sense...
  - (Camden richer than Cherry Hill?)
- and who produced the data? Trust US or Indian govt?  
Bias in what direction?

## **construct validity; seven sins**

- are you measuring what you say you are measuring?
- say want IQ, but you only have education data
- <http://www.socialresearchmethods.net/kb/constval.php>
- [http://2.bp.blogspot.com/\\_R3SXJVojagU/SwLzZJL1E2I/AAAAAAAIAE/7GbMzcZPDDk/s1600/sevendeadlysins.bmp](http://2.bp.blogspot.com/_R3SXJVojagU/SwLzZJL1E2I/AAAAAAAIAE/7GbMzcZPDDk/s1600/sevendeadlysins.bmp)
- <https://creativecommons.org/licenses/by-sa/4.0/>
- <https://creativecommons.org/licenses/by-sa/4.0/>
- <https://creativecommons.org/licenses/by-sa/4.0/>
- <https://creativecommons.org/licenses/by-sa/4.0/>
- how measure singles? polygamy, cheating, etc
- much downwards bias too eg: crime, vote (shy Trump)

## triangulate

- use several datasources and/or several variables to measure the same thing (triangulation)
- triangulation=use different measures for the same concept
- eg education:
  - years of schooling
  - highest degree obtained
  - avg SAT score
  - avg ranking of schools in the area
  - etc etc

## external validity

- are your data representative ?
- how big is the sample ?
- eg I was geocoding WVS at province level only to find out it was unrepresentative