

# Intro

Adam Okulicz-Kozaryn

`adam.okulicz.kozaryn@gmail.com`

this version: Tuesday 23<sup>rd</sup> January, 2024 11:44

# outline

overview, why this class?

descriptive statistics

bonus—data sources

## intros

- `theaok.github.io`
- data: mostly surveys: gss, wvs
- your research interests?
- data you're using?
- expectations for this class?
- [do note if someone doing sth you're into and collaborate; not necessarily sth similar, can be cross-disciplinary!]

## find topic and data and make it easy

- what to study?: study sth:
  - you're passionate about (and usually knowledgeable about)
  - is doable (accessible and easy data)
  - will further your career (think beyond graduation)
  - [sth familiar/local/work related, applied, relevant]
- at the beginning and or if learning/difficult/stuck:
  - subset dataset to sth like 3-10vars and 30-100obs

## finding your own data

- be opportunistic:
  - cannot find data for neighborhoods?
  - study cities, counties, or states
- do not have exact variable you need?
  - study sth similar!
- read literature! and follow it! (in terms of data too)
  - your first studies should closely follow published examples
  - just tweak them a little bit

# outline

overview, why this class?

descriptive statistics

bonus—data sources

# my teaching

- informal, applied
- ask questions! interrupt me
- class is not useful if you do not use the material
- apply tools to your data
- min math, max application

## warning!

- difficult! if math/stats not your thing
- back-and-forth/marinate/digest (time consuming)
- set lots of time aside
- do not get behind—very difficult to catch up later
- i'll be harsh in grading
- but i will spend as much time with you as needed!
- ask questions; email me; stop by; do it early!
- 2 KEYS TO SUCCESS: 1 start early; 2 ask q



## how qm2 differs from qm1?

- qm1 just prepares for qm2
- qm1 is not very useful
- qm2 is more useful than qm1
- qm2 requires much more work than qm1
- qm2 is more fun than qm1

# mechanics

- read slides carefully
- make sure you understand **everything crystal clear**
- if any slightest doubts, mark it up, stop by my office
- unlike most other classes, some stuff is non-intuitive
- must let it digest, set it aside, come back to it several times
- practice, practice, practice

# why regression?

- many jobs require stats
- we have more and more data
- analysis is usually done (in social science) with regression

## this class is about research

- you do PhD to do research and research only!
- otherwise it doesn't make sense!
- if you want to \*do\* public administration, community development etc, just get BA/MA
- if you want to \*research\* them and other things do PhD
- PhDs become academics or researchers in orgs
- if you want to be manager, leader etc, just get MBA!
- PhDs must have some research ideas
- regression is an excellent tool for research
- use it (or lose it)

# outline

overview, why this class?

descriptive statistics

bonus—data sources

## rule

- never run regressions without doing descriptive statistics first
- you need to understand data before making inferences from it

## descriptive statistics first!

- what is the level of measurement ?
- interval/ratio
- ordinal
- nominal

examples ?

- in ols regression:  $DV (Y) < \text{---} IVs (Xs)$ 
  - dv (outcome var) has to be interval/ratio
  - all rhs (predictors) var have to be interval/ratio or dummies (1/0); shouldn't use ordinal variables

# outline

overview, why this class?

descriptive statistics

bonus—data sources



## census data: 5-yr ACS

- census is a good source of data, even at neighborhood level!
- for city/neighborhood level probably want 5-yr ACS
- <https://www.socialexplorer.com> (on-campus or vpn)

## data sources

- <http://www.worldvaluessurvey.org/>
- <https://gss.norc.org/>
- <http://www.icpsr.umich.edu/icpsrweb/ICPSR/> and see Find Data-Search/Compare Vars
- <http://www.thearda.com/>
- <https://www.pippanorris.com/data/>
- <https://www.statepolicyindex.com/>
- <https://fred.stlouisfed.org/>
- [https://www.countyhealthrankings.org/  
explore-health-rankings/rankings-data-documentation](https://www.countyhealthrankings.org/explore-health-rankings/rankings-data-documentation)

## some cool res ideas

- (long congested) commute is the worst thing for your happiness <https://www.science.org/doi/epdf/10.1126/science.1103572>
- us whites poisoning themselves <https://www.pnas.org/doi/epdf/10.1073/pnas.1518393112>
- again, first find some cool published research, and then follow it closely and tweak a bit to produce your first research