

where do i get data?

adam okulicz-kozaryn

`adam.okulicz.kozaryn@gmail.com`

this version: Tuesday 28<sup>th</sup> January, 2025 12:26

## outline

where?

census data

good practices

# outline

where?

census data

good practices

## no need to collect—data everywhere

- used to be that you had to collect data yourself
- some people still do it
- but almost always the data you need already collected
- it is not easy to collect/produce data
  - a separate degree or at least a class on how to do it
- data are everywhere
  - whatever you study, can get data for it

## find data—the fast way

- email the listserv
- again, data questions are perfectly fine
- google/AI
- requires skill in knowing what to google/prompt

## find data—the slow (classic) way: a website

- eg these listed last week <https://theaok.github.io/datManPy/intro.pdf#page=11.98> or census

# outline

where?

census data

good practices

# socialexplorer

- i like `http://socialexplorer.com`
- (on campus or vpn)



## census website

- a similar interface on census website
- `https://data.census.gov/advanced`
- and click geographies on the left
- then at the top can search in box 'topic or table name'
  - say for 'income'

# outline

where?

census data

good practices

## cite it!

- always cite your data
  - at least full name and **specific** url [not census.gov etc]
- i do not care about the format; eg
  - cite at the end among regular references
  - cite in text
  - have an appendix
  - cite in footnote

## be careful about confidential data!

- anything that can be used to identify a person
  - ssn, dob, name, any ID #, etc
- just de-identify:
  - either drop sensitive vars
  - or create 2 files with your own new ID
    - one with sensitive info and ID, and keep it secure
    - the other with other info that you can easily carry around