

class wrap up

adam okulicz-kozaryn

`adam.okulicz.kozaryn@gmail.com`

this version: Tuesday 22nd April, 2025 16:07

outline

ps5

- do not overcomplicate!
 - better to have simple clean code that does the job
 - than messy complex fancy code that is wrong
- if parts of code take long time to run, say $> 10min$
 - optimize it, take random sample, etc
- again, it always must start from the very raw data!
- easy to make mistake:
 - think about it AND cross check; correctness is important!

ps5

- explain what you are doing!
- interpret things!
- eg when you run descriptive stats, *and* find sth interesting, put a comment and say what you have found in few words
- (don't comment output of every command)

ps5

- google things!
- before writing the code check if someone already wrote it
- and build on others work! ie copy and adapt and improve
- eg googling "python notebook basketball analysis" yields:
https:
`//www.google.com/search?client=firefox-b-1-d&q=python+notebook+basketball+analysis`
- first couple hits look good

get into flow with programming!

https:

[`//en.wikipedia.org/wiki/Flow_\(psychology\)`](https://en.wikipedia.org/wiki/Flow_(psychology))

publish or perish

- coding is fun
- but don't lose the end goal from your sight
- yes, you are an amateur software engineer—it helps a lot
- it's actually necessary these days!
- but you are a professional soc sci
- your end goal is a publication
- 'keep article pipeline full'
- at any time i have at least 3 articles under review
- how about you? submit paper right now!
- start simple, even just some des stats...but keep on submitting papers

this is super important! remember this!!

- publishing (and maybe conferences) is
 - *the only way* to get in touch with people exactly in your area
- there's just a handful of them,
 - almost never at you university, sometimes at a conference
 - usually at a journal where you submit;
(if you pick the right one, almost always at a journal)
- this is *the only way* to take your work to next level!!
 - it does take time; start now; otherwise you may never make it
- or even ssrn or arxiv etc

future research

- you've probably realized that i am into
 - Stata, Python, data
- and always happy to discuss them
- let's stay in touch!

what next?

- if didn't yet, print out box1 from <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001745> and hang on your fridge, over your desk, etc
- make use of your data management skills...
- use it or lose it
- apply for postdocs
- collaborate with faculty
 - faculty need your data management skills
 - they know much less about data mgmt than you!
 - and in general, just make \$!

make \$

- industry data jobs usually require SAS, SQL, Python, Java
- a ton of data science jobs:
 - <http://www.icrunchdata.com/>
 - <http://www.cybercoders.com/>

see again

- theory.pdf
- intro slides