

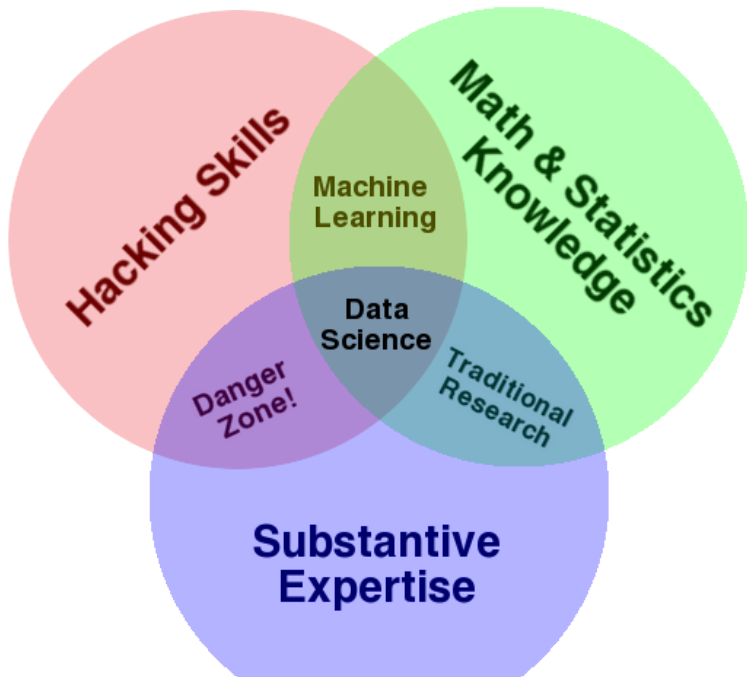
class wrap up

adam okulicz-kozaryn

`adam.okulicz.kozaryn@gmail.com`

this version: Friday 22nd November, 2024 12:54

substantive exp! the only way to beat CS



get into flow with Python!

https:

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

A simple stick figure is shown in mid-air, with its arms and legs outstretched as if it is flying or falling. The figure is positioned in the upper right quadrant of the frame.

PYTHON!

YOU'RE FLYING!
HOW?





No thanks!

We are
too busy

go public!

- make these maps public!
- show to relevant stakeholders, email to local organizations, institutions, local newspapers, etc
- it's a waste to keep such revealing maps in a drawer

publish or perish

- gis is fun
- but don't lose the end goal from your sight
- your end goal is a publication
- 'keep article pipeline full'
- at any time i have at least 3 articles under review
- start simple, even just some quick maps...then add complexity and nuance; keep on submitting papers

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!!
 - takes time; start now; otherwise may never make it
- and can just put online, eg ssrn or arxiv etc

likewise for non-academia: for-profit and non-profits

- may try to get in touch with people who do similar work/analysis
- again, first step is just to google what youare doing with keywords 'visualization' 'python' etc, and look at code and images; like lit rev in academia

in general: make it public, show to stakeholders

- the worst thing you can do is to keep it in a drawer
- when you share it (locally/globally)
- you get ideas and directions
- become part of decision making
- find mistakes and misconceptions

protect your organization

- just remember (rightfully so) each organization is scared to get hit on the head with their own data
 - so they're scared to share data and make it public
 - so make sure you'll deidentify it! and maybe fake it too! say on github your org is in chickasaw county mississippi!
 - and do not share any org specific info
 - in addition to deidentifying stuff like dropping geo locations, may take random sample of the data

use vis in other classes and thesis; and merge!

- again, merge with other data
- it could be thesis/dissertation
- often time great insight come from relating data from variety of sources eg <https://freakonomics.com>

GIGO: dont trust anybody! esp ur org

- say if you have data from census, many people use and probably found most mistakes and fixed it
- but your organization's data—probably nobody is looking at these data or very few people
- so almost for sure there are many mistakes and problems
- eg just mistake-mistake age of 20 miscoded as 200 or zip 08102 coded as 8102
- or problems: data not representative, missing data, etc etc
- in addition to vis do:
 - `info()`
 - `value_counts(dropna=False)`

future research

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