how to read quantitative papers when you don't know stats

some tips

Charlotte Lloyd, Harvard Sociology Department Writing Fellow, 2017-2018

start with the "big picture"

- even if you don't have a stats background, you can always understand the "big picture"
- identify the inputs and outcomes, don't worry about the math

input ➡ terrifying statistical model ➡ outcome

- WARNING: be aware that papers may present MULTIPLE models, and that these models may have different inputs (and outcomes!)

input: what is it?

○ models basically take input in the form of a data table

  ○ *n* number of observations (rows)

  ○ *x* number of variables (columns)

observation #	fruit	month	day	year	city	state
1	apple	June	8	1996	Denver	CO
2	banana	April	27	1988	Boulder	CO
...	...	...	...	...	...	...
100	kiwi	May	4	1980	Aspen	CO

input: be able to answer these questions

- observations
  - what is the unit of n? (people, cities, countries, etc.)
  - how many are there? (i.e. what's the "n"?)

- variables
  - what is the independent variable? what are the dependent variables?
  - what is the source of data for each variable? (each variable may have a different source)
  - does the author mention serious missing data problems for any variables?

outcomes: some major types

- **regression models**
  - outcome will be numerical "coefficients" that describe the relationship between variables
  - positive coefficients mean there is a positive relationship and vice-versa

- **logit / probit models**
  - primary outcome will be a "probability" or "risk" of an event happening (scale 0-1)
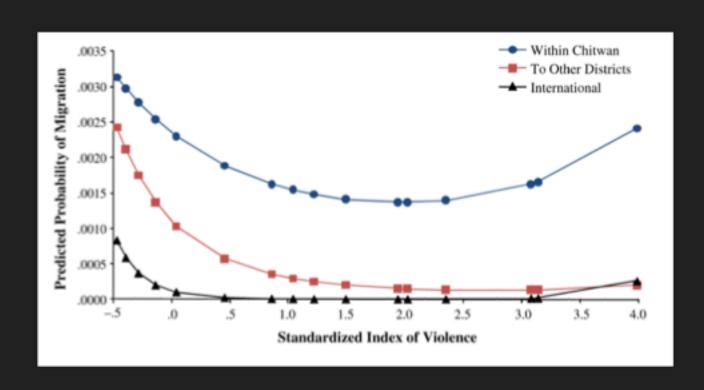 - will also generate numerical coefficients to show how variables contribute to outcomes

- **clustering, PCA, topic models**
  - outcome will be assigning each observation to a group
  - the point is to understand differences between the groups of observations
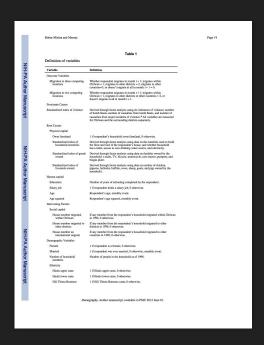
trust that they did the math correctly

○ if it's a peer-reviewed paper, the math is probably fine… or close enough ☺

○ if you don't know stats, it's fine to ignore mathy things like p-values, R-squared, model fit, heteroskedasticity, and variance. just focus on the inputs and outcomes!
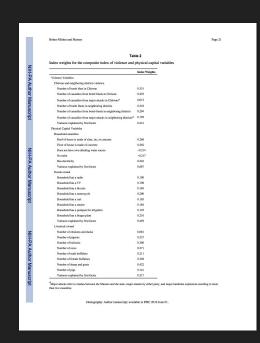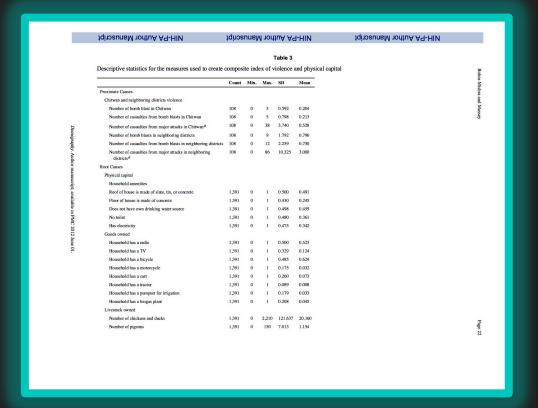
tip: look for an actual "big picture"

- generally papers provide figures that you can use to jumpstart your understanding of the "big picture"

- figures most often show the relationship between the independent variable (one of the inputs) and the outcome

Bohra-Mishra, Pratikshya, and Douglas S. Massey. "Individual decisions to migrate during civil conflict." *Demography* 48.2 (2011): 401-424.

tip: look for tables summarizing variables

Bohra-Mishra, Pratikshya, and Douglas S. Massey. "Individual decisions to migrate during civil conflict." *Demography* 48.2 (2011): 401-424.

tip: you *can* look at the outcomes table...

○ ...but always read the findings section where authors *explain and interpret* the outcomes.