11/20/17

* Questions:
  + How should I process the charge descriptions? There are 2349 unique charge descriptions (but probably just different in how they’re written down)
    - Idea: do it by charge degree b/c there are only 19 charge degrees?
  + What does it mean to clean the data? Keep seeing this term
  + How to save a sqlite db to somewhere so I can access it in another r file?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

11/3/17

to do for next time

* features table
  + recompute some of compas’ data
* data api broward county

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

10/18/17

* Use charges\_df to generate recidivated/not recidivated
  + I think COMPAS only took charges within 30 days of screenings (for their c\_charge) and only counted non-traffic, 1st charges from a different case for their history of offenses and recidivated (will not count secondary offenses as a recidivation, even though that’s questionable)
    - Thought that will not count secondary offenses in a case, but did so for aaron Epstein—not consistent
  + COMPAS throws out code (0) charges, as traffic charges

To do:

* Check where they got the data from
* Visualizing each person as a case study
* Combine all information into same table—done
* Check COMPAS scores were calculated only once per person
* Begin generating a features table for machine learning

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

10/10/17

* Brainstormed features
  + Suggested looking at features computed in Jiaming’s paper
* Code to take underlying data into something we could use
  + Get all people in compas score database, keep all 3 scores
  + Get most recent offence (the offence the compas score was calculated for)
  + Lines 87-96 in reengineering
* Trying to see what compas did

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

10/11/17