Project Proposal

Topic:

We are working with BENCHMARKS: A Citizen’s Scorecard on Judicial Accountability in Massachusetts to determine conditions in which judicial rulings are more or less likely to occur when a case is reversed. Ideally, the project will uncover never before-seen patterns of judicial behavior and decision-making in criminal and civil cases, which will have the effect of recasting and spurring more aggressive local news coverage of the state court system. As a result, the project should become a valuable tool for judicial accountability. This project is mentored by Maggie Mulvihill from the College of Communication Journalism Department and is partnered with the Boston Globe. It seeks to combine data analysis and investigative journalism to create results that are more impactful than just using one of the two. Specifically, this project should serve as a building block for a new Boston University interdisciplinary news team which will examine case reversals in all 50 states over the past decade.

Questions:

* What are similarities between cases that get reversed?
  + Are there certain judges, police departments, or districts that have a higher rate of reversal than average?
  + Are there certain types of crimes that are reversed more often than others?
  + Do the judges who have more reversals have things in common such as race, age, gender, or courtroom management style?
  + What is the base rate of reversal? What is the distribution of this?
  + For a given locality with a higher than average number of reversals, what is the likely cause (i.e. is it random chance or is it correlated to certain features)?
  + Are the reasons for reversal dependent on any of these other factors?
* What factors are most strongly correlated with the reversal of a criminal conviction?
  + Is this an instance where Gaussian Mixture would be useful?
  + If the distribution of reversed cases has some identifiable features, we could possibly find it using this method.
* (added by James on Sunday) Are there any other conviction or sentencing disparities observable in the data? If so, how well do they correlate with reversals?
  + This wasn’t mentioned in the project pitch, but I think it wouldn’t be too much of a stretch beyond our goal.

Questions to ask Maggie:

* What is the lifecycle of a criminal case (possibly a flowchart to help in analysis)?
* How can we interpret the data on the Appellate court website, including what fields indicate a reversal and what fields are important?
* What methods of compiling data from the website were used last year? Were they useful, or could they be improved upon?
* What key search terms do you have in mind for us to focus on?
* (added by James on Saturday) Would you agree that the previous effort toward this focused more on constructing the database than analyzing the data? If so, ought we to focus our efforts on the latter?

Datasets:

* Data collected from the previous team that worked on this project. These data are in CSV format. We have 4 different CSV files, ranging in size from 5-50 MB, which correspond to decisions rendered at different levels in the judicial system.
* Appellate court website: <https://www.mass.gov/appellate-opinion-portal>. This website contains information about court cases in the state of MA. We would have to do some web scraping to get this data into a form that we can analyze. I’m not sure about the size of this dataset and what exact type of information it contains, we will need to talk to Maggie about this.