

1. (10 Marks) Pick four different Canadian universities, plus Mount Allison (so you have 5 total Canadian Universities). Scrape RateMyProf for each of these schools. You should have five different csv files, each filled with the professors from each school, their rating, difficulty, the “would take again” probability, and department.
2. (10 Marks) Build a hierarchical model where you are trying to determine the modal ranking of professors for each school, **and** get an estimate of the overall population mode across the five schools. Use reasonable, mostly uninformed priors.
3. (5 Marks) Augment your hierarchical model. Add a layer of Department before School. You are now using the data to determine the modal ranking of each department, and the modal ranking of each school and the population modal ranking across all five schools. Are your school and population distributions different than they were in question 2?
4. (5 Marks) Remove the school from your hierarchy. Just departments and an overall population mode. What is the interpretation here do you think that makes this calculation qualitatively different than question 3? Are these estimates different than question 3?

Please hand in 3 R files with the three different models, the python code that generates the data, and the csv files used.