Problem Set 2 - 'Learning to crawl' (15%): due midnight, 7 October 2020

- Working in teams of one to four people, please consider this scenario: 'You are employed as a junior statistician at Petit Poll - a Canadian polling company. Petit Poll has a contract with a Canadian political party to provide them with monthly polling updates.'
- Please pick a political party that you are 'working for', and pick a geographic focus: 1) the overall election, 2) a particular province, or 3) a specific riding.
- Then decide on a survey methodology (hint: p. 13 of Wu & Thompson provides a handy checklist). Some questions you should address include:
  - What is the population, frame, and sample?
  - What sampling methods will you use and why (e.g. you could choose SRSWOR, stratified, etc). What are some of the statistical properties that the method brings to the table (e.g. for SRSWOR you could discuss Wu & Thompson, Theorem 2.2, etc, as appropriate)?
  - How are you going to reach your desired respondents?
  - How much do you estimate this will cost?
  - What steps will you take to deal with non-response and how will non-response affect your survey?
  - How are you going to protect respondent privacy?

  Remember to consider all of this in the context of your 'client' - for instance, who would be more interested in Alberta ridings: Bloc Québécois or the Conservatives? Who likely has more money to spend - the Liberals or the Greens?
- Develop a survey on a platform that was introduced in class. Be sure to test it yourselves. You will want to test this as much as possible, maybe even swap informally with another group?
- Now release the surveys into the (simulated) 'field'. Please do this by simulating an appropriate number of responses to your survey in R. Don't forget to simulate in relation to the survey methodology that you proposed. Show the results and discuss your 'findings'. Everything must be entirely reproducible. You may like to consider linking your survey 'responses' with other data such as the census or GSS.
- Use R Markdown to write a PDF report about all of this. Discuss your results and findings, your survey design and motivations, etc - all of it. You are writing a report that will eventually go to the 'client', so you must set the scene, and use language that demonstrates your command of statistical concepts but brings the reader along with you. Be sure to include graphs and tables and reference them in your discussion. Be sure

to be clear about weaknesses and biases, and opportunities for future work.

- Your report must be well written. You are allowed to, and should, use mathematical notation, but you must explain all of it in plain english. Similarly, you can, and should, use surveys/sampling/observational data terminology, but again, you need to explain it.
- Your report must include at least the following aspects: title, date, authorship, non-technical executive summary, introduction, survey methodology, results, discussion, appendices that detail the survey, and references. Your 'client' has stats graduates working for it who need to be impressed by the main content of the report, but also has people who barely know what an average is and these people need to be impressed also. This is why your report should include a non-technical executive summary. In terms of length, this would typically be roughly 10 per cent of the report. It would be more detailed than an introduction, but still at a high level.
- Your graphs must be of an extremely high standard.
- Check that you have referenced everything. Strong submissions will draw on related literature in the discussion and would be sure to also reference those. The style of references does not matter, provided it is consistent.
- Via Quercus, submit a link to your PDF report which is hosted on GitHub. At some point in the introduction to your report, you must provide a link to the GitHub repo where the code that you used for this assignment lives (Hint: Comment. Your. Code.). Your entire workflow must be entirely reproducible.
- Please work in public where it is at all possible. If you have a personal situation for not wanting your name to be public, then please email me a request for this - please do NOT reveal your personal situation to me - just make the request and suggest a work-around.
- Please be sure to include a link to your survey in your report and screenshots of the survey in the appendix of your report.
- There should be no evidence that this is a class assignment.