

Lab 1: Let's collect some data!

This lab will guide you and your group through the process of designing and collecting some data on your fellow students at UMass. The main goals of this lab are to get you to think carefully about what the process of moving from a specific question, to data collection, to preparing the data for analysis.

Lab writeup for the group is due on Wednesday, September 14 at 5pm. Each group will give a 3-5 minute presentation (strict limit!) on Thursday September 15th summarizing their hypothesis, the data collection process, and qualitative results. Since we have not covered yet how to efficiently summarize or display your data, it is not required for this presentation. Although you should give us some basic quantitative information about your process: how long did you collect for? how much data did you collect? ... Tell us a complete (but brief) story about what data you collected and why. I encourage each team to use 1-2 slides to describe the study.

Questions to answer on your own

Exercise 1 Look around the room at your classmates. What characteristics and features can you reliably and quickly identify about them?

Exercise 2 Pick three of these features, and sketch out below a table and collect some data on a few of your classmates.

Exercise 3 Take a minute and reflect on the choices that you made when creating this table. What defines a row and column in your small dataset? What types of variables were you collecting and how did you “encode” them in your table?

Questions to answer with your group

Exercise 4 Give each member of the group 2-3 minutes to present the data they collected and how they organized their table. Discuss similarities and differences between how you each collected data.

Exercise 5 As a group, brainstorm different settings on campus where you might expect the composition of students or groups of students (according to features that you think you can measure accurately) will be different. Write down a few below.

Exercise 6 Take the list you just wrote down and turn it into a series of hypotheses, with a proposed mechanism. For example, “we hypothesize that group A is different from group B because XX.”.

Exercise 7 Plan out, as a group, how you will collect data to test one or two of the hypotheses you define above. This should include logistics (when and where will each member collect data? what data will they collect?) and technical details (how will you create a dataset that you all can work from and enter data into?).

Things to hand in and present

Each group should prepare a brief written report (no more than 2 pages) describing their study. The report should include a description of the hypothesis and the data collection protocol. Figures or diagrams may be included for illustrative purposes. Additionally, include a description of the data collected as well as a proposal for how this data could be used to evaluate the hypothesis.