Statistics 36-303: Sampling, Surveys and Society
Survey Project Ideas and Proposals

1 Introduction

In this course you will be expected to make two *serious* proposals for survey projects. To help get you started I have listed some past project ideas below.

1.1 How I Will Evaluate Projects

The easiest projects typically will involve an on-campus survey of some of your fellow students, especially a subset who are easily accessible to you. The easiest mode of administration is by e-mail and/or a web service like `surveymonkey.com`, but followup and high response rates are often easier to achieve with other modes of data collection (especially interviews and other face-to-face methods), though they may require more time and effort.

In my evaluation of your project work (both proposals and completed projects), there will be a tradeoff between originality and cleverness on the one hand, and thoroughness and difficulty on the other. If you are willing to do a more complex and difficult project this may make up for lack of originality. For example, I will specifically evaluate all proposals for thoroughness, including plans for non-response followup. On the other hand, I definitely do give credit for originality. Thus a project that is really dull and boring but does hard statistics may not score higher than a project that is easier statistically but involves a clever problem or one of intrinsic scientific, social or policy interest.

1.2 Inferring Causal and Lurking Variables

Sample surveys are good at measuring things as they are, and only can be used to measure causal effects in very special and limited circumstances. Many questions of interest that are expored using sample surveys nonetheless are related to causal questions. In almost any statistical problem of comparing two or more subpopulations, or where we are studying the "effects" of some variables on others, there will be *Lurking Variables*, i.e. variables that—if not measured and accounted for in your analysis—will cause a distortion of the results, especially when you are looking at relationships. I will mention a few of these in class. I expect that your project will consider and measure lurking variables that are especially relevant for your question(s) of interest.

1.3 Thoroughness, Creatitivy and Impact

As suggested above, I will give a subjective evaluation of your proposals, based on two factors: First, is the idea really clever or really of scientific interest, and second, have you done a reasonably thorough job of things? I am hoping that this will induce you to think of a project that is both interseting and leaves scope for some statistical analysis.

For example, if you simply state that you plan to compare GPAs of male and female students, you are being neither creative nor thorough. It is easy to see that there are lurking variables (college and major), and there are also possible regression relationships with other variables of interest (e.g. SAT scores).

A good rule of thumb for whether your survey is interesting is whether anyone outside this class would pay attention to it. This might be because there is an actual client for the work, or because the work caught the interest of someone after it was completed. For example, the following 36-303 surveys were interesting:

- CMU dining services changed some of its offerings after seeing the results of a 36-303 survey on student satisfaction with food and restaurant choices on campus.
- A 36-303 survey on student attitudes toward introducing $+/-$ grading for undergraduates was written up in the *Tartan*.
- A 36-303 survey for *The Jewish Chronicle* (a client) led to changes in the Chronicle's website and publications.

Not all 36-303 surveys reach this level of impact, but if yours does, it will count positively in my evaluation of your project.

1.4 Other Notes Relating to the Project

- *I suggest that you spend approximately 80 hours as a team on this project.* You may spend more time if you are convinced that the benefits of spending more hours on this project are greater than the costs.
- *Keep a diary* (e.g. in the form of a lab book, or an on-line document), and update it regularly througout the project (no less than once a week!). Record your expenditures of time and other resources, as well as any notable events that may influence the interpretation of your data or data analyses. Keep track of which team members attend each meeting. You will not turn in this diary, but it will help you keep track of your effort and your team's effort throughout the semester.
- *You will invest a certain amount of time, energy, money and other resources, in doing this project.* As is the case with all investments, the amount of return you will receive on this project, both now and later in your career, depends on the amount you invest in it and the risk you take.

2 Some Project Ideas

Here are some topic ideas. Some of these have been used by past 36-303 teams, some have been used in other survey sampling courses, some have never been used.

You are welcome to ignore these ideas and come up with your own (more creative), or use one or more of these ideas as-is or in modified form. Either way, you will have to make a strong case that you would do a *thorough and creative* job on the topic, and that the survey could have *impact* outside this class if it is done this semester.

- *Surveys that have been or could be conducted using campus email and websites like SurveyMonkey.* **Because of the large number of groups in this class, I will limit the number of groups that will be able to conduct this type of survey.**
  - CMU community smoking habits and attitudes toward smoking
  - Workload at Carnegie Mellon vs other universities
  - Student attitudes towards plus/minus grading at CMU
  - Student attitudes toward PAT bus passes
  - Perceived versus actual use of alcohol, tobacco, and illicit drugs
  - What makes Carnegie Mellon students read email?
  - CMU student perceptions of academic integrity and old course-materials archives
  - Student attitudes toward CMU dining facilities
  - Post-Graduation plans of Carnegie Mellon University's undergraduate class of 2010
  - Knowledge and utilization of CMU SafeWalk and Escort Services
  - Student perception and utilization of mandatory student fees
- *Surveys that have been or could be conducted with other methodologies (not on-campus email/SurveyMonkey surveys).*
  - Academic departments' uses of information technology
  - Randomly sample classes/classrooms and give surveys to everyone in each sampled class. For example: Students' use of and satisfaction with Faculty Course Evaluations and other means of evaluating faculty teaching
  - Measure use or visibility of bulletin boards by posting flyers with tear-off strips people could use to enter a contest.
  - Measure the danger of an intersection by sampling or designating 30 minute blocks of time over a couple of days, and then counting the number of traffic violations (e.g. bikes running red lights) in each 30 minute block.

- Impact of the Consol Energy Center (new hockey stadium) or the Rivers Casino xon businesses in downtown Pittsburgh
- Randomly sample area bars (and possibly blocks of time) and count the number of people of a certain type (men, women, age range, etc.) in each bar during that time.
- Man-on-the-street surveys. For example, ask people to answer questions to assess their knowledge of world geography
- Measure facility usage by interviewing or observing people who use the facility. Example include
  * Sample building entrances and count the number of smokers who do not stand far enough away from the doors when smoking.
  * How often do students break rules like no food/drinks in the computer clusters?
  * Who eats lunch at the lunch trucks on Margaret Morrison?
  * Who uses the exercise rooms at the UC?

• *Special project(s): Surveys for the Jewish Chronicle and "J: Jewish Pittsburgh Living"*

  The Jewish Chronicle is an independent multimedia Jewish news organization serving the Jewish communities of Pittsburgh, western Pennsylvania and West Virginia. A sister publication, *J: Jewish Pittsburgh Living*, focuses on Pittsburgh specifically. You can find out more about both at `http://thejewishchronicle.net/`.

  Several years ago a 36-303 team conducted a survey of subscribers for the Chronicle. Since several years have passed and many changes have been made, the time is right for a followup survey. Some possible topics for the survey would include the updated website, introduction of *J* magazine, etc. The survey would most likely involve a mail-out, mail-back questionnaire, but might also involve telephone and/or web-based components.

  - The advantage of working on this project is that the focus of the survey would be defined by the client, the *Jewish Chronicle* organization, and there would be high chances of positive impact.
  - The disadvantage is that it would be more work than other possible surveys. However, the extra impact and positive impact would be considered in your grade.

  I can imagine one or more *groups working on this project.*

3 Your Project Proposals

1. Please indicate on the first page whether your group is interested in the *Jewish Chronicle* survey project(s).

2. Whether or not you indicate interest in the *Jewish Chronicle* survey project(s), your group must submit project proposals on **two** additional topics.

   For each proposal, each <u>team</u> should prepare a 1–2 page proposal. Each proposal should begin with a title that indicates the topic, and the names of every team member. Your description should include the following three section heads:

   A. Why is this topic interesting? Why does *this* survey need to be done *now*? Is there a client for whom you might do the survey?

   B. What question(s) do you propose to study? Give a brief answer that would have been understandable by a non-statistician.

   C. What research has already been done on the topic or on the theoretical construct of central importance to your topic? What could be learned from survey results? Each group member should locate and review 1 relevant item of research (e.g., article, report, book, etc.). For each item, provide:

      - Full author name(s), date, title, where it appeared (book title and publisher, if it appeared in a book; magazine title, date, page numbers, if it appeared in a magazine; website title and web address if it appeared online, etc.)
      - A sentence or two summarizing the item and what its connection is to the survey you are proposing.
      - The name of the team member who found this item.

   There is no restriction on what kind of surveys you propose, except that they should be doable by your team, in this semester.