How to Think of a Project for ISTA 370

Paul Cohen
January 12, 2012

The first project requirement for ISTA is to think of a project. This can be difficult, so here are some examples and hints.

**Hint: Start with a question**  Questions focus projects. Perhaps you want to do a project on software piracy. What do you want to find out about it? Who does it? Why do they do it? How often? Do they think they’ll get caught? Do they think it’s ethical? How do they justify it? Suppose you want to do a project on cooking. What do you want to find out? Do browned onions taste better if they are cooked quickly or slowly? Does pineapple juice make meat mushy? What’s the average time to cook an evening meal in your house? Start with questions.

Some questions are too slight and insubstantial to sustain a semester-long project; for example:

- What is the distribution of cooking times of spaghetti? The experimental work is pretty simple: just cook a bunch of boxes of spaghetti and measure how long each requires to cook. There are no interesting experiment design issues or data analysis issues, or, for that matter, culinary issues.

- Do more men than women oppose the death penalty? This question is a bit more challenging because you have to think about your samples of men and women, and ensure that they are representative samples. Also, unlike the spaghetti example, which involves just one variable (cooking time) this one involves both gender and attitudes about the death penalty. So it is a better question, but probably still isn’t substantial enough for a semester-long project.

**Hint: Many interesting and substantial studies are about how things change over time, particularly via learning.** Classic examples are studies of the time it takes people to perform a task repeatedly. In general, the time decreases, as people learn to perform the task. More recent examples concern how opinions change in response to messages from friends in social networks, and why some Twitter tweets are long-lived (i.e., they are re-tweeted often) while others are not.

**Hint: Many interesting and substantial studies are about causes.** Instead of asking, “Do onions taste better if they are browned quickly/slowly?” ask “Why do onions taste better if they are browned quickly/slowly?” Instead of asking whether men or women are more averse to the death penalty, dig deeper and ask why. Here are some examples of questions that get at causes:

- Which foods at the cafeteria cause most extreme flatulence, and what do they have in common?

- Which attributes of students and their behaviors best predict grades in ISTA 370?

- Which attributes of YouTube videos best predict whether they will go viral?
**Hint: Some studies hint at or suggest causes, others test causes.** The previous three examples suggest causes; specifically, attributes of food, students, and videos that affect digestion, scholarship and popularity, respectively. They do not *test* these suggested causes. You could get the *suggestion* that coffee improves performance on tests by conducting a survey of your friends, but to *test* whether coffee helps, you’d have to go further, and do a *controlled* study in which some test-takers had coffee and others didn’t. And to prove that a particular *component* of coffee affects performance, you’d have to compare conditions that don’t include this component with conditions that do.

Not all empirical science tests causes (so your project needn’t), and sometimes it is impossible to test causes, for ethical reasons among others. But testing causes is an important goal of science.

**Hint: Avoid testing too many causes at once.** If you think that gender, training time, caffeinated and decaffeinated coffee, whether the student is a freshman, sophomore ..., and whether the subject is Geography, Math, History or Biology *all* affect performance on a test, then you could end up with an experiment that has dozens of conditions, and you will not have time to run the experiment. Later in the term you will learn when you need to test multiple causes in one experiment, but for now, limit yourself to projects that test at most two or three possible causes.

**Hint: Your project should involve data, and you should gather some of it yourself.** One way to answer a question is to look up the answer on Wikipedia or someplace like it. Another way is to formulate and prove a mathematical theorem. Both are important to research, but neither qualifies as a project in this class because this class is about *empirical* research – or research that involves gathering evidence. By the way, it’s fine to *replicate* a study that someone else has done, or include data from other people as part of the justification for your study. But you should gather some data (and analyze it) yourself.