

I.4, Group H: Atishé Chordia, Stafford Brunk, Tony Poor, Henry Wu, SeungJin Park

Target Population

target pop	25/25
samp frame	25/25
using frame	25/25
nonresponse	25/25
total	100/100

✓ The target population includes all undergraduate students who are currently freshman, sophomores, juniors and seniors studying in Carnegie Mellon University's main campus in Pittsburgh.

Sampling Plan

✓ Our sampling frame consists of all undergraduate students in the online CMU student directory. We will take a stratified random sampling approach (based on year-freshmen, sophomores, juniors, seniors and departments-CIT, MCS, SCS, Tepper, HSS, CFA) in order to obtain a representative sample, where we randomly select x_1 SCS freshmen, x_2 SCS sophomores, etc. We will define the sample size of each strata based on the proportion to the student body, and use the email address listed on the CMU student directory to contact our randomly selected students for an email-and-web survey.

your ruby code looks great as well. nice approach.

Using the CMU student directory instead of the C-Book offers us the advantage of being able to easily create a program that randomly selects students based on the strata that we defined. However, it does come with a few caveats: for example, the student directory includes ineligible units like Qatar and Silicon Valley students. Thus, our program will select students from these campuses as well, even if they aren't included in our target population. This is the coverage error that can be fixed by using a survey question that asks for their home campus and filters out non-Pittsburgh students.

All good thinking...

After creating our survey and pretesting it out (field tests, experts in the topic) we will start the survey process by sending out an email to the selected sample and asking them to complete an email-and-web survey. After getting our preliminary results we can identify the nonresponders of the initial survey. We will start by sending them a friendly reminder e-mail regarding the survey and reiterate our need of a quick response from them, as the first wave of nonresponders may have intended to complete the survey, but forgot about it as time passed. This will be sent out 3 days after the first survey and it will highlight that the survey is designed to take very little time to complete. We will send more proactive follow up e-mails (every three days for the next 2 weeks or 4-5 emails) with repeat nonresponders to try to gauge their leverage and determine the effort and incentives we need to give to persuade them to complete the survey. We will try to increase their interest in the topic and reassure them of the legitimacy of our survey and the possibility of incorporating the survey results to allow for better pizza quality and service.

this sounds great. you may wish to change the # of followup msgs depending on how well you think they're working

After respondents provide data to ensure that the quality of our estimates are high we may need to make post survey adjustments, which include weighting underrepresented populations or imputing missing values in the surveys. ✓

If the stratified sample is not representative of the target population, we will use our program to randomly select an additional sample from underrepresented strata that will produce enough expected respondents to make our overall sample reflective of the target population. We will then send out surveys and follow up on the new sample using the same methods we applied to the original sample. If there is inadequate time to complete an additional sample, then we will use a simple weighted approach for the existing answers, putting more weight on the underrepresented strata.

both of these are great ideas.