What Determines Involvement at Carnegie Mellon?

Ellie Gurary, Bruce Jackson, Maggie Soderholm, Chrissy Swierkocki, Jen Sung

A. Why is this topic interesting? Is there a client for whom you might do the survey?

Carnegie Mellon is not known for its school spirit. In particular, we have noticed low attendance at sporting events and the creation of the Tartan Rewards Program, which aims to increase attendance at Carnegie Mellon events. Understanding how groups form on campus can be used to unify the student body and improve attendance at school events. CMU sports teams and clubs and the Tartan Rewards Program could be possible clients of this survey because the information could improve their attendance rates. This study needs to be done now to benefit incoming classes of students and attract more students to come to Carnegie Mellon. We are going to

B. What question(s) do you propose to study?

Our survey would ask questions about the student's current involvement on campus. Questions would include which factors student's find most influential in choosing groups to participate in:
- Friends involved
- Advertising
- Major
- Past participation
- Career goals
- Dorm activities
- What main mode of contact do you use for involvement (email, facebook, text, call, in person)
- Do you live on or off campus? Have a car?
- Nationality/International Status?

We would also ask about a student's grade level, hometown and gender to see if these are confounding variables. We can do additional research into the size of students' virtual social networks versus "real" social networks. How do you mainly interact with others: classes, clubs, social life, office hours/study groups, other?

K. Sampling Scheme

Undergraduate Students in randomly sampled classes. We will use a SRS without replacement sampling scheme in order to conduct our survey. We will randomly select classes from this website: < https://enr-apps.as.cmu.edu/open/SOC/web/images/documents.htm> through a random number generator in R. We will email the respective professors for permission to come in and administer the survey (hopefully at the end of class, so as not to disrupt the lecture). Once we gain permission, we will be conducting self-administered pen-and-paper questionnaires that the students can take.

We decided on SRS without replacement because we would like to give equal chances to all the undergraduate classes (minus the graduate classes). Since our target population is

Carnegie Mellon Undergraduates, it would be best to not stratify and give all the undergraduates an equal chance. We chose to do SRS without replacement since once we choose a class for our sample we are not replacing the chosen class back into our random number generator.

L. Questions

What year were you born?

What is/are your major(s)? Any minor(s)?

What is your gender?

What is your ethnicity/race?

Where is your permanent residence? Is it a suburb/urban/rural area?

Are you currently living off campus/on campus? (address)
-Do you have a roommate?

What is your childhood household size? (childhood=age 0-12)

Are you currently employed?
-On campus/off campus?
-How many hours a week do you work?

Are you currently involved in a campus organization? (athletics, greek life, clubs)

How often do you communicate with others through these modes:
-Facebook
-Twitter
-Email
-Cellphone
-Gchat
-AIM
-Skype

Do you currently have a social networking account? If so, in which website(s)?

How many athletic events have you attended since the beginning of the school year? (range)

How many Greek events have you attended since the beginning of the school year? (range)

How likely are you to go to a campus event if publicized through email? (likert scale)

How likely are you to go to a campus event if publicized through flyers? (likert scale)

[Handwritten annotations in red:]

these questions are a good start but they are still somewhat vague (both in wording and in the question format you will use).

Moreover they do not seem to really get at your main research question aobut "understanding how groups form" etc.

If you do not work hard on your survey questions, you will not be able to get data to answer the research qiuestion you have posed, and this will make for a less good final project.

why are these relevant questions?

there is some redundancy in these two questions. separate out the issues and put separate issues in separate questions.

These questions are not silly but I don't see how you string them together to get a picture of "how" students form groups. What modes of communication do students use to find out abotu or form groups? Is that the major issue or is it how much study time they have to put in? Or they feel they ahve enough personal friends tehy don't need formal groups? Or... etc. etc. You need to think much harder about extracting information that will help asnwer your reseearch questions!!

?

How likely are you to go to a campus event if publicized through <u>WOM</u>? (likert scale)

How likely are you willing to go to a campus event if publicized through social networking sites? (likert scale)

Where do you prefer to study? Alone or with a group?

What percentage of your friends is in the same major as you? (range)
-With this set of friends, do you primarily do schoolwork or activities not related to school? Or both? (this could be changed, the wording is off)

When you do attend campus events, do you go alone or with a group of friends?

How did you meet your closest friends?
-Dorm
-Campus events
-Classes
-Parties

How many days per week do you go out to socialize (excluding schoolwork)? (range)

same comments as above. These are "OK" questions about socializing but not about *your* research question!

M. Sample Size

We assumed that there 5,705 CMU undergraduates on campus this year (which was given to us in lecture). We decided on a 40% response rate (to oversample) and 5% margin of error (p).

Standard Deviation $= \sqrt{(0.4)(0.6)} = 0.49$ **why not just use the worst-case SD=0.50?**

$N_0 = \dfrac{(1.96)^2(0.24)}{0.0025} = 368.8 = 369$
N_0 = Sample size for SRS with replacement

BUT since we're conducting SRS without replacement:

$N \geq \dfrac{5707 * 369}{5705 + 369} = 346.6 = 347$
N = Sample size for SRS without replacement

with 40% response rate you would need to contact 347/0.40 = 868 students.

Since we assumed 40% response rate, we believe that this is a good sample size. However, if our actual response rate is different from our hypothetical one, then we will adjust our sample size to reflect that.

In addition, because you are doing clusetered sampling, you should inflate your sample size by another 20% or so, so 1.2*347 = 441 students needed; with the 40% resp rate this means 441/0.40 = 1101 students to contact

We are assuming that there are at least 20 students per class at CMU. There may be lectures that we may be sampling, but there also may be classes where there are less than 10 students. So, with that low average, we are randomly choosing 18 classes. If we see that that is not the case,

we will adjust the number of classes accordingly. The 18 classes that we chose randomly through a random number generator are given in a different document.

Make use of data from the schedule of classes, factbook, etc., to get a better estimate of average class size, so you know about how many classes you will need to sample.

If the average class size were 20 (I think this is small btw) then you would need to contact a minumum of about 22 classes (for 440 students) and a maximum of about 55 classes (for the full 1100 students). Note that as soon as you have about 440 students you can stop (so if the resp rate is better than 40% you won't have to contact all 55 classes in your random sample).