

Bassem Mikhael  
Christopher Loncke  
Abigail Daughtrey  
Perna Agarwal  
Yong-Gyun Choi  
February 22, 2011

**Background:** We are measuring student consumption of caffeine on campus, and their perception of average student consumption. We are also collecting other data, such as GPA, frequency of exercise, etc., to determine if there is any correlation between caffeine consumption and performance. Further, we will determine what the preferred method of caffeine consumption is on campus.

**K.** The sampling scheme we will use is a simple random survey without replacement. We will be using a random number generator to generate two numbers, which we will use to randomly draw names from our sampling frame, the C-book. This will be done for 500 iterations, and 1 email will be sent to each person selected. The CMU undergraduate population is comprised of approximately 6,000 students, so a sample of 500 individuals represents about 8% of the entire target population. Since this is a relatively large proportion, we do not feel the need at this point to stratify our survey or use clustered sampling. Implementing a stratified survey would be relatively straightforward given our target population; we could easily divide the sample by class year (freshman, sophomore, junior, senior) or by college (H&SS, MCS, CIT, etc.). However, at this point, this could introduce selection bias. Therefore, we will maintain a simple random survey without replacement. If it is necessary to stratify our survey, then we will do so according to the class year guideline. We would prefer this as opposed to doing some post-survey adjustment, such as weighting underrepresented components of the sample, so as to avoid introducing any adjustment error into our calculations.

**L.** Questions for the survey:

1. What is your year at CMU?
  - a. Freshman
  - b. Sophomore
  - c. Junior
  - d. Senior
2. What college is your primary major in? (For interdisciplinary, select both)
  - a. CIT
  - b. H&SS
  - c. CFA
  - d. SCS
  - e. MCS
  - f. Heinz
  - g. Tepper
3. What is your GPA?
  - a. 0.0-2.5
  - b. 2.6-3.0
  - c. 3.0-3.5
  - d. 3.6-4.0

post-survey weighting is fine too, though you won't be able to make separate inferences about different classes if you post-stratify

Stratification reduces bias, it doesn't introduce bias. You may wish to stratify, since caffeine consumption may be different at different stages in college.

4. What is your gender?
  - a. Male
  - b. Female
5. How many units are you taking this semester?
  - a. 36 or less
  - b. 37-45
  - c. 45-55
  - d. More than 55
6. Do you have any additional majors or minors?
  - a. Yes
  - b. No
7. If so, which statement best describes your situation?
  - a. I have a minor
  - b. I have more than one minor
  - c. I have an additional major
  - d. I have an additional major and an additional minor
  - e. I have more than one additional major
  - f. I have more than one additional major and more than one minor
8. About how many hours a week, on average, do you spend on extracurricular activities such as clubs, student organizations, or inter-mural sports (this includes attending and preparing for meetings or events)?
  - a. 0-3 hours
  - b. 3-6 hours
  - c. 6-9 hours
  - d. 9-12 hours
  - e. More than 12 hours
9. How much personal leisure time do you have in a given week? This is defined as any time available for all non-work related activities (classes, extracurricular activities, employment, etc.)
  - a. None
  - b. Less than 2 hours per day
  - c. 2-4 hours per day
  - d. More than 4 hours per day
10. How many hours outside of class, on average, do you spend doing classwork?
  - a. None
  - b. Less than 2 hours per day
  - c. 2-4 hours per day
  - d. More than 4 hours per day
11. Do you have a part-time job (employed on or off campus)?
  - a. Yes; I work for less than 5 hours per week.
  - b. Yes; I work for less than 10 hours per week.
  - c. Yes; I work for less than 15 hours per week.
  - d. No; I do not have a job.
12. How often do you exercise in an average week?
  - a. None
  - b. Once a week
  - c. 2-4 times per week
  - d. More than 4 times per week

Though it would be a bit harder to code, "list your majors (including primary, second, dual, etc.)" and "list your minors" might get you more both more accurate information (that could be coded into these categories) and deeper info (you might find out what majors or groups of majors go with higher caffeine consumption).

All of the prompts here are "per week" but the options flip back and forth between "per day" and "per week". This could easily be confusing...

13. How many hours do you sleep on an average night?
- Less than 3 hours
  - 3-5 hours
  - 5-8 hours
  - More than 8 hours
14. How much caffeine do you consume on a daily basis, where 1 serving is one item (eg. One cup of coffee, one espresso shot, etc.)
- None
  - Less than 2
  - 2-4
  - 4-6
  - More than 6
15. What is your primary motivation for consuming caffeine?
- Pleasure/Leisure
  - To stay awake/maintain energy
  - Partying
  - Other \_\_\_\_\_ **that**
16. What types of caffeine beverages do you consume? Select all **which** are relevant.
- Coffee
  - Espresso
  - Caffeine tablets
  - Redbull/energy drinks
  - Soda
  - Chocolate
  - Other \_\_\_\_\_
17. Are you consuming more or less caffeine than you are comfortable with?
- I am consuming more caffeine than I am comfortable with.
  - I am consuming a comfortable amount of caffeine.
  - I am drinking less caffeine than I am comfortable with
18. Do you feel that your personal caffeine consumption is more or less than the average consumption of your undergraduate peers at Carnegie Mellon?
- I believe I consume an above average amount of caffeine.
  - I believe I consume an average amount of caffeine.
  - I believe I consume a below average amount of caffeine.
19. Do you believe that consuming **more** caffeine positively impacts your performance in school work, employment, athletics, or extracurricular activities?
- Yes
  - No
20. Do you believe that consuming **more** caffeine negatively impacts your health?
- Yes
  - No
21. Do you experience any of the following on a regular basis (on most days)? Choose all that apply.
- Insomnia
  - Nervousness/anxiety throughout the day
  - Restlessness
  - Irritability
  - Nausea/gastrointestinal problems
  - Fast or irregular heartbeat

**more than what? And how much more? maybe better would be "consuming caffeine at your current level" or something like that.**

- g. Muscle Tremors
  - h. Headaches
22. Do you smoke?
- a. Yes
  - b. No
  - c. I do not wish to disclose ~~~~~ why?
23. How frequently do you consume alcohol?
- a. None
  - b. Less than twice a week
  - c. 2-4 times per week
  - d. More than 4 times per week

**M.** Since this is a simple random survey without replacement, the formula below is used to calculate the minimum number of respondents. We assume a margin of error of 0.1, and a standard deviation of 0.5. We arrived at this value for standard deviation considering questions such as the number of hours per day the student spends on various activities, such as exercise, studying, etc. We think, based on just informally polling our group members, that the standard deviation will be approximately 0.5.

$$n_{\text{minimum}} > \frac{N * n_0}{N + n_0}$$

$$n_0 = \frac{(z_{\alpha/2})^2 * (SD)^2}{(ME)^2}$$

$$n_0 = \frac{(1.96)^2 * (0.5)^2}{(0.1)^2} = 96.04$$

$$n_{\text{minimum}} > \frac{5951 * 97}{5951 + 97} = 95 \text{ people}$$

Given that we have a \$50 Amazon gift card incentive for response, we estimate that our response rate will be approximately 25%. Therefore, given an ideal sample response of  $n_0 = 95$  people and a response rate of 25%, we will need 4 times this number, or 380 people. Therefore, polling 500 should be ample, even if our response rate is slightly below 25%. With 500 people, we can afford to have a response rate of about 19%.

looks fine

-BJ