Stat 322 (F17) - Assignment #3

(Due ~~Wed. Nov. 15th at 4:00~~ Fri. Nov. 17th at 12:00 pm (noon) in appropriate STAT 322 slot in assignment box #15 outside the Math Tutorial Centre (MC 4066/4067). Electronic submissions or in-class submissions will not be accepted.

1) Read the Waterloo News (March 5/2014) article *Energy drinks linked to teen health risks* posted in the articles on LEARN.
   a) Based on the article, define the following:
      i) The target population.
      ii) The study population (i.e. the sampling frame)
      iii) The sample.
   b) Give two population attributes, and the values of their estimates obtained from the sample, referred to in the article.
   c) Describe each of the following errors in the context of the study, and for each error, discuss whether you feel the error would be negligible or considerable.
      i) Frame error (or Study error)
      ii) Sample error
      iii) Measurement error
2) Open the associated journal article through the 'published in Preventative Medicine' link in the news article (also posted on LEARN), and read about the methodology in the Methods section (primarily, the *Participants* section) to answer the following questions:
   (Note: for consistency, use the sample size provided in the Waterloo News article for calculation purposes)
   a) Update your definition of the study population from question 1).
   b) Based on the description of the sample design, clearly describe how stratified sampling, cluster sampling, and simple random sampling were used in this survey. Be sure to clearly define the strata and the clusters.
   c) According to the webpage, https://www.populationpyramid.net/canada/2012/, approx. 5% of Canadians are of the age from which the sample was taken. If we assume that this proportion is consistent with the provinces involved in this survey, calculate the finite population correction factor associated with the standard deviation of estimates associated with this survey (note that the total population of the relevant provinces is given).
   d) Based on your results in 2c) and on the information in the Results section, calculate an approx. 95% confidence interval for:
      i) The proportion of Canadian teens who used energy drinks at least once in the past year.
      ii) The proportion of Canadian teens who used energy drinks at least once a month in the past year.
      iii) The average age of Canadians from the age demographic in the frame (Grades 7, 9, 10, 12).
   e) Which of the intervals calculated for the proportions in d) is wider? Why?
   f) For the proportion in d i), calculate the sample size required for the results to be 'accurate to within 1 percentage point, 99 times out of 100'.