Stat 446 **Name:** _____
Take home midterm
Due 10/17/16 at 1:10 PM

For the take home exam, you may use the textbook, any course materials provided on D2L, homeworks, and labs. You **may not** discuss questions or work together with classmates. You are welcome to contact the instructor with any questions related to better explantion or understanding of the questions themselves. Any relevant material from questions will be posted to D2L for the benefit of the entire class. For complete (and partial credit) please show all work, whether that be by hand or printed R code.

1. *Continuing from the in class midterm.* The Procrastinator Theater is interested in estimating the number of Bozeman residents (both MSU students and town citizens) that would attend movies over the summer.

   Assume your final recommendation was to conduct a stratified random sample. There were four strata: MSU students, age 0 - 22, age 23-50, and age 50+. The strata are mutually exclusive, that is MSU students are not included in the other strata. Use the following information for the rest of the question.

Stratum	Total Population (N_h)	Sample Size (n_h)	Number of positive responses in sample
0 - 22	6,000	1,000	151
23 - 50	12,000	2,000	372
50 +	6,000	1,000	112
MSU Students	16,000	4,000	803
Total	40,000	8,000	1,438

   (a) (10 points) Compute 95% confidence intervals for the proportion that would attend a summer movie at the Procrastinator for each stratum. That is you should calculate 4 confidence intervals, one for each stratum.

(b) (6 points) Compute a 95% confidence interval for the proportion of all Bozeman residents that would see a movie at the Procrastinator. Given that N is very large, use a critical value from a normal distribution rather than Sattertwaithe's approximation to construct your confidence interval.

(c) (6 points) Summarize your results for the director of the Procrastinator. Include an interpretation of the confidence intervals in parts (a) and (b).

2. Your second task as the lead statistician at Capital Bikeshare is to estimate the number of bikes checked out on a given day. This is important as Capital Bikeshare wants to make sure there is always a bike available to be rented or they will lose business. You have been asked to conduct an SRS to estimate the average number of bikes checked out at a station. You ask for an estimate of the population variance, S^2, and are told that another study came up with a value of $s^2 = 775$.

   (a) (13 points) Your study should estimate the average number of bikes rented at a station to within 3 bikes. Given that there are 392 stations, compute the sample size you will need to get this level of precision with 95% probability.

   (b) (13 points) Using the 'MidtermBikes.csv' dataset on D2L take a sample with the sample size you calculated in part (a). Using this single sample compute a 95% confidence interval for the total number of bikes checked out daily. Your supervisors are not interested in the minimum number of bikes, so compute a confidence interval such that the lower limit is 0 (a 1-sided confidence interval).

(c) (6 points) Write a written summary of your results. Include a discussion of the sample size calculation and an interpretation of the confidence interval.