Math 222 - Project 4   **Due Friday, March 29**

1. The General Social Survey (GSS) asked this question: "Have you attended religious services in the last week?" The responses for individuals with at least a high school degree are contained in the file services.csv.

   (a) Make a 2-way table showing the relationship between education level and religious attendance. Display the association with a segmented bar graph and/or a mosaic plot. Explain what you see.

   (b) Is there a statistically significant association between religious attendance and education level? Describe the nature of the association and its strength. Clearly state your hypotheses and conclusions.

   (c) Remove the column for people whose highest degree was high school. Repeat the test for association between religious attendance and education. Do you still get the same results? Explain.

   (d) Combine all three columns of data from the last part, and compare individuals whose highest degree was high school with individuals who have completed a degree beyond high school. Estimate the size of the difference in religious attendance between these two groups.

   (e) Write a short summary explaining the overall relationship between religious attendance and education level.

2. A 2013 study tested the effectiveness of an influenza (flu) vaccine in children aged 3-8 years old. 5168 children from 8 countries were randomly assigned to receive either an influenza vaccine or a control vaccine (for Hepatitis A). The results of the study are shown in the two way table below.

	Control Vaccine	Flu Vaccine
Developed influenza	148	62
Did not develop influenza	2436	2522

   One way to measure the effectiveness of the flu vaccine is to calculate the **odds ratio**, which is the ratio of the odds of developing flu for the control group divided by the odds for the treatment group. Recall that the the odds are the number of successes divided by the number of failures. In this study, the odds of developing the flu were 2.47 times higher for children in the control group than in the treatment group.

   (a) Use the bootstrap method to create a bootstrap distribution for the odds ratio. Describe the bootstrap distribution (shape, center, spread, etc.). Hint: To get the bootstrap distribution, you need to generate bootstrap samples of 5168 individuals randomly selected from the original data with replacement. This file: fluVaccine.csv contains the original sample. You can randomly select rows by adapting the following R-code snippet (which selects 3 random rows from a data frame `df`:

   $$df[sample(nrow(df),3),]$$

   (b) Make a bootstrap confidence interval for the odds ratio in the population.

   (c) Based on your confidence interval, what can you say about whether the flu vaccine is effective?