STAT 217: Project 4
Due Friday, March 20 in class

You must work in a group of 2-4. There will be a point deduction if you do not work in a group.

One method for assessing avalanche hazard is to conduct a snowpack stability test known as a compression test. Avalanche forecasters, recreational skiiers, snowmobilers, and others use compression tests to assess the likelihood of triggering an avalanche. A compression test is conducted by isolating a thirty centimeter by thirty centimeter column of snow, tapping on the column, and watching for a failure in the snowpack.

Watch a short video about compression testing at the following url: https://vimeo.com/30996756. We will focus on the first 7 minutes and 30 seconds of the video.

There are two variables that have been considered useful or important to consider when assessing the avalanche hazard with a compression test. The first is the number of taps. If the failure occurs when tapping from the wrist, it is recorded as 'easy'. If the failure occurs when tapping from the elbow, it is recorded as 'medium', and if the failure occurs when tapping from the whole arm, it is recorded as 'hard'. The second is fracture character. Fracture character describes the shear quality of the failure and is categorized into one of five types - 'progressive compression', 'sudden planar', 'sudden collapse', 'resistant planar', or 'break'.

In the video, Mr. Jamieson points out,

> Fracture character is distinguishing between whether skiier triggering is likely or unlikely, and it does so much better than by paying attention to the number of taps only.

Avalanche experts are intrigued and would like you to further investigate this topic. Conduct a chi squared test for fracture character. Follow the project writing guidelines. In addition to the project writing guidelines, include the following:

- *Introduction*
  There is very little information about how the data were collected, but give a brief description of the sample data that is available to you. Assume that each time the researchers went into the field, they conducted a compression test and then recorded number of taps, fracture character, and whether or not the skiier triggered an avalanche on the slope. Clearly state the question of interest that you would use for the appropriate chi-squared test (homogeneity or independence).
- *Statistical Procedures Used*
  In addition to what you normally include in this section (outlined in the *Project Writing Guidelines*), include the following:
  - State the appropriate statistical test (homogeneity or independence).

  - Discuss what you see in the contingency table and the mosaic plot (or stacked bar chart depending on which type of test you perform).

- *Summary of Statistical Findings*
  Write a decision sentence and an evidence sentence as usual. Instead of including a traditional 'estimate' sentence, discuss what you see in the standardized residuals plot.

- *Scope of Inference*
  Follow the examples in the *Project Writing Guidelines*.

- *Overall Summary*
  After you have finished the report, consider the second variable, number of taps. The code below produces a contingency table and a standardized residual plot for number of taps. Based on these plots, discuss which variable (fracture or taps) you would use to predict whether or not a slope will avalanche and why.

  - **Extra Credit**: The last line of code produces two tableplots that show the relationship between fracture, taps, and the likelihood of a triggering a slide. You will receive extra credit if you include an extra paragraph to describe the relationships you see in these tableplots. How do these plots help display which variable is most useful in predicting avalanches?

R Code
Import the "slides" dataset from D2L and run the following code. Before you begin, install the `mosaic` and `tabplot` packages if you have not done so already.

```
require(mosaic)
#First look at the raw dataset
slides

#For fracture
frac.tally <- tally(~fracture+trigger, data=slides)
frac.tally
mosaicplot(frac.tally)
chisq.test(frac.tally)
chisq.test(frac.tally)$expected
chisq.test(frac.tally)$residuals
mosaicplot(frac.tally, shade=T)

#For number of taps
taps.tally <- tally(~taps+trigger, data=slides)
taps.tally
mosaicplot(taps.tally, shade=T)

#tableplot
require(tabplot)
```

```
tableplot(slides[,c(2,1,3)])
tableplot(slides)
```