I was involved in two of the consulting projects: (1) Investigation on math anxiety among high school students. (2) Investigation on patients’ responses to doctors’ recommendations of taking eye surgeries. SAS software was used to perform statistical analysis in both of the two projects.

(1) **Math anxiety project**

**Introduction**
This consulting project is to assist a M.S. student majoring in math education to write her Master’s thesis. The objective of our project is to (a) Estimate the prevalence and intensity of math anxiety among high school students. (b) Identify pedagogical strategies that may help alleviate math anxiety. Three surveys were issued to two groups of high school students. Group 1 students had one missing value, group 2 students had two missing values, the observations that have missing values were deleted and were not included in our analysis. After we manipulated the raw data, we have group 1 consisted of 20 students (80% female, 20% male) and group 2 consisted of 23 students (65% female, 35% male).

**Step 1: Estimate the prevalence and intensity of math anxiety among high school students**

The first survey had 10 questions, the first 5 questions were positively worded and the last 5 were negatively worded, each question had choices of 5 scores ranging from 1 to 5. We reversed the scores of the first 5 questions and then checked the frequency table of the first 5 question and the last 5 questions. According to our results, the answers made by the students were quite consistent when we compare the questions that were differently worded but actually had similar meanings. Therefore we can safely use the scores picked by the students. After this step we summed the scores of the 10 questions to get the total scores of the students respectively. We defined initial math anxiety scale of the students by the following criteria and obtained the following results: low (total score<=10, 0.00%), medium low (11<=total score<=20, 2.33%), medium (21<=total score<=30, 83.72%), medium high (31<=total score<=40, 11.63%), high (41<=total score<=50, 2.33%). Therefore we concluded that math anxiety dose exist but most of the students were in medium and medium high range. We also checked the quantitative ability scores given by the teacher and found that no correlations between the students’ quantitative ability and their initial math anxiety scale.

**Step 2: Pre-teaching and post-teaching comparison (Pre & post results from the same sample were taken from the second and the third surveys.)**

(a) **Consider using paired t test**
Assumption for paired t test: The distribution of the (post-pre) difference is normal. If the normality check of SAS UNIVARIATE procedure shows p-value>=0.05 then the normality assumption is accepted and we can use paired t test. On the contrary, if the normality assumption was rejected then we need to use the other statistical methods instead.

(b) **Consider using Non-parametric method**
If the distribution of the (post-pre) difference is not normal, we can’t use paired t test and we can try non-parametric method. SAS UNIVARIATE procedure will perform some testes such as “sign test”, “student’s t test”, “signed rank test” to test the null
hypothesis that location for the mean=0. If p-value of the tests >= 0.05 then there is no significant difference between the mean of the post and pre teaching scores.

(c) Discussion
The tests we performed above showed there were no significant differences between the mean of the pre and post scores for the two groups respectively, but when we looked at the two samples we found that there were very large variation within the groups (Group 1 Period A: mean=-1.75, std=8.87; Group 2 Period B: mean=1.30, std=4.54) and box plots of the two samples show that there were low outliers in group 1 and high outlier in group 2. There must exist some other factors we should include in our tests to investigate the effects of different teaching methods.

Step 3: Compare the post-pre differences of two different teaching methods.

Strong evidence was shown to accept the null hypothesis that there is no significant difference among the mean of the initial math anxiety scale of the two groups. (From SAS PROC TTEST output we can see that equality of variance assumption of pooled t test was satisfied and the p-value of the test is 0.5029.) After this test we will carry on tests to compare effects of different teaching methods (Period A and Period B). The second and the third survey were the same survey consisting of 15 questions and were answered by the students before and after the lectures respectively. To perform the research to identify effects of different pedagogical strategies, each of the two groups of high school students was given a different lecture respectively. One is manipulated (Period A) and one is more abstract (Period B).

(a) Analysis of variance
SAS GLM procedure was performed to analyze source of variances. The coefficient of determination (R^2) of our model is 0.569, which means that 56.9% of the variance of the response variable (post-pre difference) can be explained by our model. There were no significant differences between the mean of the post-pre difference of the 2 groups of students but there were significant difference within the groups and the interaction effects were significant.

(b) Consider using logistic regression model
The response variable here is the difference of post-pre scores (survey 3 scores subtract survey 2 scores) and it takes m>2 discrete values. Results of SAS LOGISTIC procedure showed that teaching method of period A can help to alleviate the students’ math anxiety more than period B method (lower math anxiety after the lecture), students who initially felt higher math anxiety or have lower quantitative ability will alleviate their math anxiety more after the lectures than the other students.

(2) Patients’ responses project
This consulting project is to assist a M.D. student majoring in health and public administration to write her Doctor’s thesis. Survey data from 3 different hospitals were analyzed to investigate factors affecting patients’ responses to doctors’ recommendations to take eye surgeries. Chi-square test was performed to check the homogeneity of the affects of race and hospitals respectively. Dummy variables of race, hospital, doctor type and their interaction effects were created and logistic regression model containing both the discrete variable and continuous variable (age, eye test score) was built by SAS PROC Logistic. Parameters were estimated to calculate dichotomous outcomes based on a patient’s decision to undergo surgery or not.