Matt Boyas
Stat 222 Midterm Report—Deviations from Proposal

This document is intended to accompany the included work and serves to explain some of the deviations from my original research proposal.

(1)
It made sense to me to lump together my first and third research questions, for as I began investigating question #1, I was also testing many of the variables that I wanted to look at in question #3.

(2)
Some of the titles of plots and tables have been altered or completely changed because I discovered that the titles no longer really worked once I had an actual figure/table at which to look.

(3)
The originally-proposed Figures 3 & 4 were changed to scatterplots. I was not sure how to calculate the original proportions for conflicts with zero deaths, and the scatterplots seemed to convey a very similar idea.

(4)
I chose not to include the regression summary due to the fact that three of the five variables are categorical, split into 7-10 binaries. As a result, there are a lot of variables listed with names that do not mean all that much, so I plan on just summarizing the regression summary in the text of my paper and perhaps including the entire summary output as an appendix.

(5)
Many of the proposed figures/tables relating to the regression were somewhat hypothetical, and as my modeling methods changed, so did some of the plots. For example, I chose not to include normal Q-Q plots due to not running linear regression. I also did not include fitted value vs. residual plots for the training dataset in favor of focusing on clear plots/tables regarding prediction accuracy (I did not want to have too many similar looking figures/tables in my paper for fear that overall meaning and flow of the paper would be affected.)

(6)
The figure to show prediction accuracy from the number of deaths model was modified to be a fitted value vs. residual plot for the three modeling methods that I attempted.