Amy Zhang

Stat 511

Project Proposal

My plan is to recreate, as best I can, Nate Silver’s forecasts for the 2012 presidential election. My methodology will follow what is described on the FiveThirtyEight blog, where polls are collected together and weighted according to three factors: how recent they are, how large the sample size of the poll was, and how well that particular poll has done in the past. How well the poll has done is determined by a rating, which Nate Silver described at length on FiveThirtyEight. In this way, we can model the predictive power of each poll. An average of the polls is then taken, and the polling average is used along with other non-poll factors, such as the political makeup of each state and the sum of individual contributions received by each candidate.

I will use a dataset made available on GitHub, which may not be the same as Nate Silver’s dataset but was created to replicate Nate Silver’s work. The dataset includes presidential poll data for 2004, 2008 and 2012, as well as 2000 census data and partisan voting data.