Stat 8801 Case Study 6: MPCA
March 7, 2012
Group 4: Felipe, Rachel, Yu-Feng, Xin, Bryan

Problem Recap
The MPCA uses volunteers to measure the clarity of streams. They look in tubes with stream water to see how far down they can see. The number of measurements varies both by the volunteer and by the year. Clarity varies by season and water level. They want individual reports for each volunteer to show how clarity in their stream has changed. They'd like to test if there is a trend for each stream, compare it to others and see if there is an overall trend.

Methods of Inference for Individual Volunteers
We could fit a multiple regression with season and water level as our x-variables and clarity as our response to see if there is a trend. We might have to check to see if there are any time dependence issues with season, and if that is the case we could look at something to do with time series.

Methods of Inference for Individual Streams
We could fit a multiple regression with season, water level, and volunteers (if there is more than one volunteer per stream) as our x-variables and clarity as the response to see if there is a trend. Again if there are any time dependence issues with season, and if that is the case we could look at something to do with time series.

Methods of Inference for Watersheds
We could fit a multiple regression with season, water level, volunteers, and streams as our x-variables and clarity as the response. We would be able to test if there is a difference between the streams by using a t-test to see if streams offers a significant amount of information to the model.

Presenting to Public
The easiest thing for the public to understand might just be graphs. One thing that could be used is a Trellis Plot. For each stream we could make a matrix of plots with columns being seasons, and the rows being water level. For each graph in the matrix we could plot clarity vs time. So for example the third row and second column could correspond with clarity of water for the spring season and high water level. We could do this for each watershed too if it turns out that streams are all similar, which may not be too much of a stretch if it's the same watershed. If the streams are not similar we could just plot multiple colored lines (each color corresponding for a separate stream) in our trellis plot.