In order for us to find the best variables to review, we needed to implement a few regression models. We decided on using these methods due to each variable being graphed against DOC not resulting in any graph visually standing out as most significant. We chose to use three regression models: **STEP BACK, STEP FORWARD, and LASSO(least absolute shrinkage and selection operator).** After running each of the methods on the data set, we compared the variables that were kept in each one. Each of these variables were grouped by number of selections and the higher chosen variables were given higher consideration. After the groups were made, each group's variables were graphed against the DOC. From there, a few things were considered: the data represented by the variables themselves, the graph and slope line, and the general relationship between the variable and the DOC. From there, two variables were chosen from the ten that were selected by the models as the most **INTRIGUING/INTERESTING/IMPORTANT**and the ones that would be considered for further analysis. The two variables that were chosen were Mean Elevation and Waste Water Source Point.