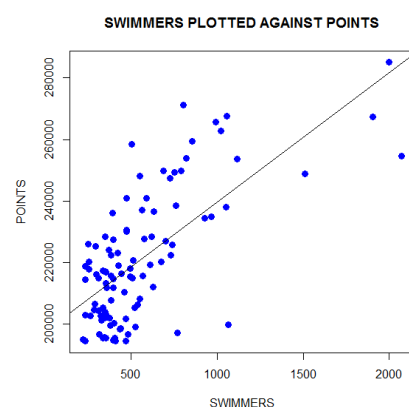


Executive Summary - USA Swimming Virtual Club Championship Correlation Project
Mike Hruska - Geog 560 Final Project - May 3, 2018

This project will look at the USA Swimming Virtual Club Championship (VCC) for the short course 2017-18 season. We will discover if there is a correlation between our dependent variable, 2018 VCC point totals for the top 100 teams in the USA, and our independent variables: 1) the population of the team's metropolitan area, 2) the number of registered swimmers on the team, and 3) the number of registered coaches on the team. After the initial analyses we will also factor in two new variables: 4) last years, 2017, VCC points and 5) the region the team is located. The goal of this project is to find if there is any kind of correlation between the point total of a team and: the population numbers they pull from, the number of swimmers they have, and the number of coaches they have.

Based on our methods we can state that, when comparing the population size they draw from, their number of coaches, and their number of registered swimmers the largest individual predictor for a team's VCC points is number of registered swimmers (as shown to the right). When all three variables are put together we show an increased predictor overall. We then analyzed all five variables, including the previous year's points and zones. The analysis shows that the previous year's point total shows the highest correlation, an even greater correlation than all of our three original variables combined. All five variables together show the greatest of all, although zone, also known as the region a team is located in, seems to have very little influence on points.



It is expected that a team which draws from a larger population is likely to have more swimmers, and will thus need more coaches, and will likely score more points. My estimate is that the further down on the VCC point list we may have went, say top 200, 500, or even 1000 teams, is that all variables, swimmers, coaches, and population, would have much higher correlations to point total. I believe this would happen for a number of reasons, not just from a performance standpoint, but also from a purely numbers game. As a team size gets smaller their chance of filling all point scoring slots with higher level swims become increasingly more difficult.

One factor which really stood out when analyzing all the data is that each of the top 100 teams are within the top 185 metropolitan areas in the US, with most being in the top 50. A number of the teams cluster around specific areas, specifically 12 metro areas account for 55 of the top 100 teams in the US, which is over half of our list. I think this would especially prove interesting for future follow-up and analysis pertaining to population and location analysis. Looking at how top ranking teams tend to cluster around specific locations. Does having top ranking teams, and swimmers, in an area help to create other top ranking teams?