

CAROLINAS SPORTS ANALYTICS MEETING 2014

POSTER ABSTRACTS

Analyzing NCAA Tournament Matchups Using Cluster Analysis

Miles Abbett, Davidson College

Despite the existence of many predictive models in the world of sports, few are able to quantify how certain teams matchup against one another. More often than not, the better team wins, though upsets occur all the time due to variations in style of play. Using k-means clustering and analysis of 756 NCAA basketball teams, I determined the historical win probabilities between different types of teams and attempted to predict this year's NCAA tournament with the results

A Graphic View of Davidson Basketball

Miles Abbet and Seth Kindig, Davidson College

Coauthor: Tim Chartier

Students and faculty came together this year to form CatsStats, a group of mathematically-focused thinkers with the goal of providing the Davidson men's basketball team with an analytical advantage. The group worked hand in hand with the coaching staff during the course of the season, providing the team with data-driven scouting reports, lineup analysis, and shot chart visualizations. Senior math majors Miles Abbett and Seth Kindig designed an infographic highlighting the Wildcats' notable statistical feats from the 2013-2014 season.

The "Power" of a Power Play in the NHL

William Baker, Radford University

Coauthor: Gary Fetter

The power play in hockey can be a golden opportunity for a team to score on their opponent, if they can take advantage of it. The power play gives a team the advantage of having more players on the ice for a certain amount of time than their opponent. It is generally agreed that it is easier to score when a team has a player advantage as compared to when each team has the usual 5 players and a goalie. However, having a power play opportunity does not guarantee that a team will score, and whether or not a team scores during a power play opportunity does not guarantee that they will win the game. Although the power play is an important opportunity in the game of hockey, there are few research studies that have examined its quantitatively. This study aims to evaluate the impact that power plays in hockey have upon winning or losing a game. We present results from our statistical analysis using data from the 2010-11, 11-12 and 12-13 season's National Hockey League (NHL) games.

Heads or Tails when Flipping and Spinning

Jeremiah Bartz, Francis Marion University

It is common folklore that flipping a penny results in heads roughly half the time and tails the other half. On the other hand, some claim that spinning a penny will result in a different proportion of heads and tails than in the flipping situation. This poster investigates these situations by analyzing experimental data collected by students enrolled in an introductory probability and statistics course at Francis Marion University.

Course Management in Golf

James Bramlett, Davidson College

I will take on a multivariable approach of certain golf statistics to see if there is correlation with strong course management and consistency at varying levels.

What's in a Ranking?

Andrea Fant and Laura Peckham, Furman University

There are over 250 college football teams in Division I alone. This poses a difficulty in ranking the teams, as any good ranking must be able to account for different the different conferences, and compare teams that have never played each other. The existing methods for rating and ranking are very different. They each take into account a variety of factors, and may emphasize wins, losses, point differentials, or more specific elements of a season. Focusing on the 120 teams in the Football Bowl Subdivision (FBS), we explored several questions. When ranking, which factors seem to be most important? Is there a simple strategy that shows promise as an effective ranking method? Finally, how do we measure the quality and success of a ranking method?

Keys to Success: Measuring Significant Statistics in Basketball

Ford Higgins, Davidson College

Coauthors: Miles Abbett and Quinn Morrison

Over the past 15 years, Coach Bob McKillop of the Davidson College men's basketball team has used a set of 15 goals to motivate and evaluate his team's performance. Our objective is to measure how strongly each goal correlates with a win, if there is a minimum number of goals to expect a win, and if there are any improvements that we can make to the goals.

Exploring Exercise Science with Statistics

Andrew Hirst, St. Olaf College

Coauthors: Steve Papciak and Sam Walczak

Each year at Saint Olaf College, senior exercise science majors must design a scientific experiment as a graduation requirement. The experiments of the graduating class of 2014 include the effects of carbohydrate supplementation in soccer players, protein synthesis in the elderly, and cross training for long distance runners, among other topics. Lacking the necessary experience in research design and statistical analysis, these students need assistance in the technical aspects of their study. We assist these seniors with their research methods, experimental design, data collection, modeling, and interpretation of results. This entails meeting with the students on a regular basis to explain statistical ideas as well as demonstrate how certain concepts can be implemented within the context of their study. The goal of our project is to ensure that the studies will be statistically sound in hopes that the results produced will be useful within the relevant fields of exercise science.

Using Similarity Scores for Tournament Prediction

Seth Kindig, Davidson College

Similarity Scores are a concept developed by Bill James in SABERmetrics to compare teams across time. What if we applied this concept to teams in the same season? Using advanced analytics, we attempted to compare teams based on style, and then use how those styles performed against each other in the regular season of the 2014 NCAA basketball season to predict matchups in the 2014 NCAA tournament. The results appear to reveal that this method is good at picking out potential upsets as well as predicting how games will be played.

The NFL Schedule: Investigations of the Colley Inverse

Daniel Kuzbary, Furman University

Coauthor: Cassandra Chan

Traditional ranking systems are based primarily upon win-loss ratios. However, such an emphasis ignores factors like the ease or difficulty of one team's schedule in comparison to that of another. The Colley ranking system resolves this issue by taking into account both win-loss ratios and strength of schedule. Our research project investigated how the structure of a league's overall schedule affects their teams' Colley rankings, with particular application to the 32-team National Football League schedule.

March Mathness

Andrew Liu, Davidson College

We created a website, <http://marchmathness.davidson.edu>, that allowed users to explore the mathematics of ranking in order to improve their March Madness brackets. The website offered the Colley or Massey methods, both utilizing linear algebra. Users could input parameters to weight the recency of games. Then, the resulting weighted Colley or Massey method would compute a rating for all Division I NCAA men's basketball teams. Teams with higher ratings are predicted to win over teams with lower ratings. In all, the code was run over 16,000 in over 25 countries. We aggregated all user parameters and the resulting bracket rankings produced by running the code available on the site. This poster will discuss the process of developing the website and explore the types of parameters entered by people in order to improve their brackets.

Player Synchronicity: Using Heart Rates to Determine Team Cooperation

Hannah Vincent, Furman University

Teams across all sports look to quantify social and psychological factors that impact a team's and individual's performance. This poster details and investigates the use of heart rate data to determine a method to quantify these factors, and to determine if "synchronicity" is a good measure of team cooperation and performance.

A Decision Tree Approach to Pitch Prediction in Major League Baseball

Noah Woodward, Davidson College

The project makes use of PITCHf/x pitch data to identify pitch selection patterns that various MLB starting pitchers exhibit. These patterns are identified through machine learning algorithms in the statistical package C5.0, and displayed through a decision tree framework. (Research results will be published in an article on hardballtimes.com on Monday, March 17th.)

Elo Analysis of Quarterbacks in the NFL

Cameron Zima, Wake Forest University

Coauthors: Steven Thompson and Jason Parsley

There is a debate about how quarterbacks should be ranked and valued in the NFL. Most rankings are defined solely by personal performance and statistics; however, we will be approaching it from a different perspective. Our goal is to rank NFL quarterbacks by a variation of the Elo Ranking System. Popularly used in chess, the ranking system will be evaluated by the difference in a player's actual to expected performance against a given opponent. We use econometric regressions to isolate quarterback variables that are most important to wins. This ranking offers a global comparison of quarterbacks and the result will be distinct comparable statistics.