Power Percentage Rankings

for Collegiate Bowling Teams

This ranking system was developed by Frank Barnhart for use as a tool when figuring out his IBMA Poll ballot. (See page 2 for the calculation process.) After trying several several systems and formulas, he came to the realization that the shortcoming in those systems was they did not account for the strength of the field in each tournament.

- This system is purely a "Power Ranking" in the same sense of those you find, for example, on SI.com (Sports Illustrated, © Time Inc.). Wins & losses (and/or bonus points) in the case of match-play events are ignored. Also, while 1st place always counts for more than 2nd place, how much more depends on the winning margin. So, it is pinfall that counts in this system, not position
- So, it should not be confused with a "Poll Ranking," although it was developed it to help figure out poll ballots. If you asked a college coach in a match-play tournament, "If you could only achieve one, which would you choose: Win the tournament, or be the team with the highest average in the tournament?" I believe most all coaches would answer "Win." So, it would be quite reasonable for there to be differences between these rankings and the poll rankings of someone using these rankings as a guide.
- The formulas used are circular:
    - The results of the last step feed back into the first step.
    - With each iteration of the calculations, you get closer to a "solution," with the changes getting smaller and smaller with each iteration.
    - One could repeat iterations indefinitely, but one only needs to repeat the calculations until the changes from the prior iteration are sufficiently small.
    - Power % rankings are shown to one decimal place (e.g., 102.3%). Calculation iterations are repeated until the largest change for any one team does not exceed plus or minus 0.010% from the prior iteration.
- 100.0% represents a "national average" team, i.e., it is the average of all the teams in the country.
- Results from tournaments with less than four teams are omitted.
- Also omitted is data for teams that do not bowl all of a tournament.
- Teams with data from less than 4 tournaments are not ranked.
- There is no attempt to track roster changes from tournament to tournament. A team from a school is considered to be the "same" team each time they bowl.
- A key feature of this system is the calculation of an "Adjusted Field Average" for each tournament. This is what a hypothetical "national average" team would be expected to bowl in this event. While the purpose here was to take Field Strength into account when calculating each team's Power % Score (see Step 4 below); this also provides a comparison of the scoring level of each event. This is included both in the Tournament Rankings.
- All teams are included regardless of which, if any, eligibility requirements they have met. With this system, the more data you have, the more reliable the results become. If you want only NAIA or USBC eligible teams, for example, the results are more accurate if you eliminate teams not meeting that eligibility after calculating the power % rankings, not before.

Methodology: (Each iteration involves these steps.)

1. Each team has a Power % Rating, which is the result of Step 6 below from the previous iteration. When a team is first added to the spreadsheet, it is assigned a Power % Rating of 100.0.
2. For each tournament, a Field Strength Rating is calculated. This is the average of the Power % Ratings of all the teams in the tournament.
3. For each tournament, the Field Average is calculated, which is the average of all the participating teams' averages for the event. The Field Average is then divided by the Field Strength Rating to determine the event's Adjusted Field Average.
4. Each team in each tournament then receives a Power % Score for the event, which is their average in the event, divided by the Adjusted Field Average.
5. Each team's Power % Rating is then calculated by averaging their Power % Score for all the events in which they participated. This is a weighted average, with the weight given to a tournament being (n-1)/n, where n is the # of teams in the tournament.
6. All the teams' Power % Ratings are then re-normalized to the 100.0% "national average" reference point, if needed, by adjusting all of them up or down an equal percentage so that they all average out to 100.0.
7. The calculations are then repeated (i.e., another iteration is run) by feeding the #'s from Step 6 into Step 1. (On my Excel spreadsheet, I have two columns, one for each of these. To run another iteration, I copy the cells in the column from Step 6, and paste them into the column for Step 1 by using "Paste Values".)

Note: Sectionals are the only post-season tournaments included in these rankings.

Some Notes About "Cross-Over" Teams

Note: The data presented below comes from my women's Power Percentage Rating spreadsheet containing data available as of the second IBMA Poll of the 2013-2014 season, with addition of the results of the 2014 ITC Sectional tournaments.

A number of NJCAA teams (men & women) bowl only in NJCAA events and a number of NCAA teams (women) bowl only in NCAA events.

To have a meaningful comparison between these teams and other teams not bowling in these events requires having some "cross-over" teams. These teams provide what I call "points of intersection" when trying to make these comparisons. Here is a very simple data set to illustrate:

- Two tournaments, X and Y
- Two teams in each tournament
- Team A bowls only in Tournament X. Team C bowls only in Tournament Y. Team B bowls in both tournaments.
- Team B is the "point of intersection" needed to compare Team A to Team C.

Now to the risk: If Team B has an unusually good performance in Tournament X (and a typical performance in Tournament Y), that will not only give them a good ranking in comparison to Team A (which I am fine with, if that is all the data you have, they out-bowled Team A), but it will also distort the comparison between Teams A and C, causing Team C to look better in comparison to Team A.

The answer to this problem is to have enough "points of intersection" so that any such variances are either offset by variances in the opposite direction and/or are diluted by the quantity of data available.

Most of the discussion on this point has centered on the NCAA women's teams as there are more of those teams that would potentially be ranked teams (when combining everyone together) than there are NJCAA teams.

So, the question I asked myself (see note above as to the data on hand at this point in time) was: Is there enough cross-over team data at this point to reliably compare NCAA women's teams to non-NCAA women's teams?

To check this out, I picked one of the NCAA "cross-over" teams (McKendree), two non-NCAA teams that had participated in a non-NCAA event in which McKendree participated (Robert Morris-Illinois and Rock Valley in the McKendree Challenge), and two NCAA teams which had participated only in NCAA events including one in which McKendree participated (Bethune-Cookman and Vanderbilt in the NCAA Mid-Winter).

McKendree averaged 895.6 in the McKendree Challenge. My test was to see how the comparison between the NCAA teams and the non-NCAA teams was affected a) if McKendree had averaged 50 pins higher (945.6) in the McKendree Challenge, and b) if McKendree had averaged 50 pins lower (845.6) in the McKendree Challenge. When doing these two tests, I again repeated my calculation iterations until no team had a change from one iteration to the next of more than plus or minus 0.010.

- The changes in McKendree's data changed their Power % Score for the McKendree Challenge from 112.2 to 118.1 and to 106.1, respectively. (This is a fairly large variance in a team's performance. Of all the teams on my spreadsheet, about 2/3rds of them had a maximum Power % Score that was less than 5 percentage points above their Power % Rating.)
- McKendree's Power % Rating changed from 108.7 to 109.6 and to 107.7, respectively.
- The Field Strength ratings for the two tournaments changed by +0.1 and -0.1, respectively.
- The Power % Ratings for Robert Morris-Illinois and Rock Valley did not change.
- The Power % Ratings for Bethune-Cookman and Vanderbilt changed by only 0.1 (interestingly, by +0.1 in both cases).

So, given how small the change in the Power % Ratings is due to one large variance in one team's performance in one event, my conclusion is that the data set used to conduct this test contains sufficient cross-over team data to reliably compare NCAA and non-NCAA women's teams. I plan to repeat this test next season using only the data that is available for the first poll. The changes from the test should be larger with less cross-over team data. The question will be how much larger?