March Mathness Worksheet 4

Massey Method

Due: Mon. Feb. 29th by 11:59PM

In this worksheet, you will conduct you own data mining and Massey Ranking Analysis for your assigned conference.

* First, Watch Lectures 7-8 and 13-16 from Tim Chartier on Udemy
* Also read this introduction from Ken Massey himself on his method of bracketology selection: <http://blog.press.princeton.edu/2012/02/29/march-mathness-the-massey-method/>
* Be prepared to review your notes on the Massey Method from class.
* Now, we’re ready to generate our own conference rankings using Massey’s method. Go to <http://www.masseyratings.com> and select **“Basketball College Men | NCAA I”** on the left hand side of the page. You will come to the official Massey ranking page for all NCAA D-1 basketball teams. This site is great for looking at all of the D-1 teams, but we’re focused on our assigned conferences today.
* Select the dropdown menu that says **“NCAA 1”.** Change that to your assigned conference.
* Then click on the **“More”** dropdown menu from the rankings page. Select **“Scores/Schedule data”**
* The screen may go blank with no data (or a small selection of data) displayed. Click the option **“All”** at the top of the screen.
* Now you should have a LOT of game data displayed. These are all the games played by members of your conference at any point in the season. This data includes conference season, non-conference season, and any pre-season tournaments. You will be able to see the date, teams playing, the scores, and even the location if it is a neutral site game or if the game went to overtime.
* Note: The winning team in a given matchup is displayed on the left hand side. Also, teams that are the “home team” for a given game have a “@” at the beginning of their name.
* In today’s exercise, we’re only going to be concerned with games between conference members. This page shows game data for any game involving a team from your conference. We can change that by selecting **“Intra”** instead of “All” at the top of the screen. Also, do not select “Scheduled” or “Exhibition”, they are not necessary. By selecting “Intra”, the website will only include conference season games. Now our game data looks confined to just the conference season. For this module, we’ll use excel to manipulate the data.
* Open the Massey\_Module.xlsm spreadsheet attached in the email. Many elements of this spreadsheet will be locked. You only have to worry about certain cells. We’ll get to that later.
* Highlight all the game data (AND JUST THE GAME DATA), copy, and paste into cell A1 in the excel spreadsheet (an example game has been provided in cells A1:H1. We’ll remove it in a minute).
* You’ll notice that all the data got bunched into one column. We need to spread all that game data into multiple cells. There are two ways to do this:

**METHOD ONE:**

* If your computer gives you the option to “Enable Macros” then say yes or “Enable”.
* Then, once all of the data has been entered into the first cell A1 and extends downwards, click on the “Text to Columns” Figure. It’s big and blue, you can’t miss it.
* The Macro should populate cells A-H with your new games data. From there, go on to **NEXT STEPS** below, after Method 2.

**METHOD TWO:**

* Go to the Excel tab titled “Review” at the top. Click “Unprotect Sheet”.
* Then, click the Column A heading to highlight all the values in column A. Then, go to the Data Tab and select the “Text-to-Columns” feature.
* Select the Fixed width option as shown:



* Press next. You now have the opportunity to select your barriers between cells. It is essential that you break the data into exactly 8 columns. You need to create a barrier between the @ symbol and the winning teams, as well as between the @ symbol and the losing teams. See below:



* Press Finish. You’ve now got your data in workable order. Next, go to Cells I4:I19.
* Now, go back to the “Review” tab and select “Protect Sheet” and press OK when prompted. This will protect some of the formula sensitive cells from an accidental user error.

**NEXT STEPS:**

* Enter in each team name in your conference, exactly as it appears in the game data. You’ll know if you spelled it correctly when columns L,M,N,O fill up with numbers.
* Notice that column O denotes a team net points, which is represents matrix **p** in the **M\*r = p** Massey equation.
* The green table represents the **M** matrix of the Massey equation. Some of it has been filled in. It will be your job to fill in the rest.

**Assignment:**

Now that you have all the data you need to successful complete the Adjusted Massey Matrix equations, use *Mathematica* to enter in the matrices (remember to replace the last line in **M**  with a row of 1’s in the Adjusted Massey) and solve for the system of linear equations to generate a rating **ri** for each team. Enter the rating for each team in the Excel document, then under the ranking heading, list the team in order by rating. This is your Adjusted Massey Ranking!

Then write up a report on the differences between the conference standings, the Massey rankings and the Colley Rankings. What do you feel is the best indicator for future success?

Submit the excel spreadsheet, the write up in Word, and the *Mathematica* code to your TA (agormley@nd.edu) no later than 11:59PM on Monday, Feb. 29th.