**Activity for Trainer selection**

**English Premier League Soccer Standings**

The English Premier League (EPL) is a major soccer league in Great Britain consisting of 20 teams. The season begins in August and concludes in May with each team playing each other team exactly twice (home and away). Each team plays 38 games in a season while the total number of games is 380. A team receives 3 points for a win and if the game is tied, both teams receive 1 point; no points are awarded for a loss.

The EPL game results are found at <http://www.football-data.co.uk/englandm.php> under the heading *Premier League*.

Develop a function with the inputs of date and season that returns the league standings for the date and season specified. The season specified should, at a minimum, accept the current seasons and both of the two most recently completed seasons

**Data**:

* The csv data files available via the weblink provided above contain a lot of information that is not needed. The critical information is:
  + Date (in day/month/year format)
  + Home Team
  + Away Team
  + FTHG – the number of goals scored by the home team
  + FTAG – the number of goals scored by the away team
  + FTR – the result of the match (*H* indicates home team won, *A* indicates the away team won, *D* indicates a draw or tie)
* The data for previous seasons is static (unchanging) but the data for the current season changes whenever a game is played. Your function should read the data directly from the web pages and not from a csv file stored on your hard drive. The *read.csv*()command will do this easily where the url is used instead of the filename.
* Notice that the year in the date column sometimes appears as a 2-digit value but sometimes it appears as 4-digit value. Your function will need to address this issue.
* Check that the data column names are consistent from season to season; if they are not, you need to address this issue in your function.

**Requirements**:

* Your function should be named *EPL\_Standings*. Any other function name will lose points in grading. Consistent naming is very important, as I will be testing each of your submissions with a grading script that will only look for the exact function name listed above.
* The parameters of the function are, in order,
  + Date, in the *mm/dd/yy* format, and
  + Season specified by *yyyy/yy*, such as 2014/15

***Example Function Usage*:**

If I were to load your script and enter the following command in my R console:

*EPL\_Standings(“04/25/2017”, “2016/17”)*

Your function should return information specified below on the 2016/17 season at the conclusion of all matches played on April 25, 2017.

**Function Specifications:**

* The function should accomplish two goals. It should
  + return a data frame described in the next bullet, and
  + display the standings in descending order according to the number of points per match earned up to and including the date parameter value. When two teams have the same number of points per match, the teams should appear in descending order according to the number of points per match, then wins, then goals scored per match and finally ascending according to goals allowed per match.
* Along with the team name (named *TeamName*), the data frame returned by the function will contain the following columns where the column name is specified in parentheses:
  + record as *wins-loses-ties* *(Record*)
  + home record (*HomeRec*)
  + away record (*AwayRec*)
  + matches played (*MathchesPlayed*)
  + points (*Points*),
  + points per match (*PPM*),
  + point percentage = points / 3 \* the number of games played, (*PtPct*)
  + goals scored (*GS*),
  + goals scored per match (*GSM*)
  + goals allowed (*GA*)
  + goals allowed per match (*GAM*)
  + the team’s record over the last 10 games played (*Last10*)
  + the team’s current streak where *W6* would mean the team has won their last six games but they have not won their last 7 games, *L2* indicates a team has lost their last two games but they have not lost their last 3 games, while *T1* denotes the team has tied their most recent game but did not tie their last two games. (*Streak*)
* **Pls note it must be programmed entirely in R** and may not rely on packages that permit coding in another language such as SQL, VB, Python, etc.

**Activity:**

* Send the correct standings for any specified dates and seasons
* Clear documentation of the function through code comments