DATA-437 IN-CLASS ASSIGNMENT OCTOBER 11, 2021

1. Consider the `standings.csv` data file. This file contains information on Baseball teams from 1980 to 2006. **In Python**, answer the following:

   (a) (4 points) Using the pandas library, read the csv file and create a data-frame called `baseball`.

   (b) (4 points) Create a scatter-plot between `Runs Scored` and `Wins`. Comment on the plot.

   (c) (6 points) Compute the correlation between `Runs Scored` and `Wins`. Is this correlation significant?

2. Consider the `Dataset_2_4.csv` data file. This file contains information on 2011 MLB players with a qualifying number of plate appearances (502 or more). **In R**, answer the following:

   (a) (4 points) Using the `read.csv` function, read the csv file and create a data-frame called `mlb`.

   (b) (4 points) Create a scatter-plot between `Hits` and `Runs`. Comment on the plot.

   (c) (6 points) Compute the correlation between `Hits` and `Runs`. Is this correlation significant?