# Pandas exercise

Git clone phd_nikhef repository to your computer
https://github.com/twalstra/phd_nikhef

1.  Start your python script with importing pandas and load the csv file seasonresults.csv. Use the variable df as the pandas datafile.
2.  print the top 5 lines of the datafile using pandas
3.  calculate the size of the datafile in rows and columns using pandas
4.  use describe() to have some statistics
5.  Determine which line contains the maximum for the column "Wscore" using argmax()
6.  Interesting is the iloc() function for indexing: print the following:
    ```
    df.iloc[[df['Wscore'].argmax()]]
    ```

    The result is a pandas DataFrame object
7.  df.iloc[:3] gives you the first 3 lines of the pandas data set print the result and print also the result of df.loc[:3] (without the 'i'). What is the difference? (no, it's not a major difference)
8.  `df.sort_values('Lscore')` will sort the column. Print the first 5 lines of this one
9.  print the first 5 rows of the columns 'Wscore' and 'Lscore'