

APPM2720 HW5

The file **BB200.txt** in the Homework Directory contains results on the top 200 finishers from the 2014 Bolder Boulder 10K (6.2 miles) footrace. The columns are: place, names, sex/age division (M or F and age category), place in the division, times for each mile, average mile time and the total time.

In the file HW5.R is a handy function **convertTime** to convert the times in the data set into the number of minutes e.g. `Mile1<- convertTime(results[,6])`

Don't forget to source this function in order to use it!

- (1) Read in the data **BB200.txt** as a data frame.
- (2) Convert the times to numerical values. Explain whether the first mile split time (MILE1) is predictive or not predictive of the place (PLACE). Include a figure to support your answer.
- (3) Use the **substr** and **as.numeric** functions to make a new variable AGE from the division (DIV) variable. Also make another variable SEX based on the M/F division code. Make a plot of TIME as a function of AGE and color the points based on the SEX.
- (4) The file **BB200Raw.txt** is closer to the raw data that was extracted from the web. `read.table` does not work on this file. What are the three separate aspects that need to be corrected for this file to be read in by **read.table** ?