

Project Title: Boston Marathon Run Times

Date: 2017/09/28

Group Members:

- Ash Mohan
- Nathaniel Medina
- Nicolas Pappas
- Rick Yang

Description:

- The Github repository below contains Boston Marathon data from the years 2001-2014.
- Each year contains individual data from thousands of marathon participants, including each runner's name, age, gender, country of origin, official run time, official halfway time, etc.
- With so many columns of data for every runner, there are LOTS of things we can do with the data in a given year.
 - For instance, we can look at the distribution of all official run times in 2014 and analyze how different variables (age, gender, country, etc) correlate to run time. R and Tableau can be used to visualize these ideas.
 - Another idea is to compare different years against each other, looking for trends that occur over time (number of participants, average run time, men/women ratio, etc).
 - Another idea is to look at runners who ran the marathon in consecutive years, and visualize the change in their run time. Each year typically has a few thousand runners from the previous year.
- Ultimately, our goal is to find insights and key takeaways from the provided Boston Marathon data, and then visualize our findings with the tools learned in this course.

Potential Dataset:

<https://github.com/llimllib/bostonmarathon/tree/master/results>

- Excluding 2013 (Boston Marathon Bombing)

