

2013 Chicago Marathon Numbers

Peter Chng

I like running and numbers

Introduction

- Scraper: Simple BS4-based script run against <http://results.chicagomarathon.com/2013/>
- Output results into “CSV” format; read into Pandas
- Everything available at:
<https://github.com/pchng/cm2013>
(I really should have used csv module rather than ‘,’.join()...)
- Learning exercise, not rigorous
- Structure of data: Finish place, division, half split time, finishing time

Sample size and Demographics

- Mass participation event; likely a good sample of runners that choose to run a marathon (n=38,883)
- No qualification requirements for 2013
- Bias is toward people from {Chicago, Illinois, USA, North America}
- Focused on finishing times
- Age/gender breakdown: Five-year M/F divisions; differences across divisions/gender

Finishing time distribution

- Not smooth; likely indicates running to a specific goal time rather than “all out”
- Cherry-picked results/graphs
- Affected by specific goal times: Vary by age, BQs, etc.

Boston Qualifiers

- No male/female advantage
- Percentage qualifying increases with age (interpretation?)
- Contrast with new Chicago Qualifying standards: Not fair to older folks!
(Put into place following 2013 registration debacle)

Positive/negative split

- Positive: Second half slower; Negative: Second half faster
- Accepted dogma for most runners (i.e. not pros) is to aim for a negative split: Will run faster overall, feel better, etc. (Hold yourself back)
- Overall, mean finish for negative split was lower than mean for positive split, but this likely confuses cause and effect – runner could have “blew up” (Always be aware of interpretation!)
- This positive/negative disappears if limit to faster runners (this eliminates the runners who “blew up”)