

FE 582

Lecture 3: Case Study: Data and Web Technologies

Dragos Bozdog

For academic use only.

Exercise: Modeling Runners' Times



<http://cherryblossom.org/>

CREDIT UNION
CHERRY BLOSSOM
TEN MILE RUN

Credit Union
Cherry Blossom
Ten Mile Run

April 6, 2014 in Washington, D.C at 7:30 A.M. "The Runner's Rite of Spring"®

Become an Online Fundraiser
Support Children's Miracle Network
Hospitals

Race Results

Searchable Results (1973-2013)
RunPix Enhanced Results
Results for 1973 - 2013

2013 Results
2012 Results
2011 Results
2010 Results
2009 Results
2008 Results
2007 Results
2006 Results
2005 Results
2004 Results
2003 Results
2002 Results
2001 Results
2000 Results
1999 Results

THE RACES
10 Mile Run
5K Run-Walk
1/2 Mile Kids' Run

RACE RESULTS
RACE RESULTS

GENERAL INFORMATION
The Basics
Bag Check
Elite/Seeded Runner Qualifying
Entry Information
Expo/Packet Pickup Info
Getting to the Race

COUNTDOWN TO RACE DAY
49 days 12h 50m 50s

THE LATEST NEWS
USA Men's and Women's 10 Mile Championships
Transfer Process Now Open
Kids Run Registration Now Open
Team Registration Now Open
Training Program Registration Now Open
2014 Entry Closed - Lottery Results Announced
2015 Lottery Notification Signup
Important Hotel Notice - Renaissance Hotel Full
Number Transfer Procedures Open Feb. 1
Close-Out Sale on 2013 Merchandise!
Credit Union Cherry Blossom Blog
October 2013 E-Newsletter
Full 2013 Race Results
1973-2013 Searchable Results Database
2013 Participant Photos
Make Sure You Receive Our Emails

STAY CONNECTED
twitter facebook



Exercise: Modeling Runners' Times

Topics:

- Scrape simple Web pages for text content
- Use regular expressions to extract and clean messy data from pre-formatted text tables and to create unique identifiers for matching records that belong to the same individual.
- Employ statistical techniques to identify bad data and to confirm these problems have been corrected.
- Visualize data that have a large number of observations (~150,000 records).
- Gain experience with the R formula language for plots and modeling.
- Fit piecewise linear models using least squares and non-parametric curves using local averaging.
- Compare data structures, e.g., a data frame and a list of data frames, for holding and working with longitudinal data. This includes the application of 'apply' functions such as `tapply()`, `mapply()`, `sapply()`, and `lapply()`.



Thank You

Dragos Bozdog

For academic use only.