Predictive Analytics II
Individual Project Proposal
Daniel Burkhalter
10/06/2023

Proposal

Competitive running is a sport which garners little attention from the public in the United States other than weird looks, yet often compliments. However, recreational running is a common practice in summer 5k's and marathons. My goal is to promote competitive running in South Dakota for high school students by creating a power rankings system for cross country (otherwise known as XC).

Definition

What is a power ranking system? A power ranking system takes cross country results from all XC meets in SD and ranks individuals (and teams) based on head-to-head competition. For example, runners in Northwest SD (middle of nowhere, Bison SD) will occasionally run against Rapid City schools but will never see Sioux Falls schools until the end of the year. Rapid City schools will race Sioux Falls schools several times throughout the year. My goal is to compare the matchup between Bison-Rapid City and Rapid City-Sioux Falls, to determine Bison-Sioux Falls matchup.

Current System

The current system ranks individuals based on 5k times throughout the state. There are two major drawbacks:

1) The difficulty (and distance) of courses varies widely throughout the state. A flat course in Sioux Falls is not comparable to a course at altitude in Custer, SD. Hence times are not standardized.
2) Head-to-head competitions are not considered. If individual A beats individual B three times throughout the year in Custer, SD. Yet, individual B races in Sioux Falls (individual A doesn't race in Sioux Falls) and runs a "faster time" than individual A, then individual B will be ranked higher than individual A.

Data

There is a centralized database for all SD high school cross country meets housed by Athletic.net. The displayed results are proprietary information and permission must be granted before obtaining/scraping results.

Given denied permission from Athletic.net to access their database, it is assumed Athltic.net does not own the rights to SD high school cross country meets. The South Dakota High School Activities Association (SDHSAA) owns the rights to all (sponsored) high school sports in SD and permission to use results for academic purposes will likely be granted.

Given permission from SDHSAA, I will contact all cross-country coaches/administrators in SD and request their results from 2023.

Ranking Methodology

The proposed ranking methodology is the ELO rating system (used for Chess rankings). Custom additions to ELO are possible due to the complexity of cross country. For example: Runner A beats Runner B and Runner B beats Runner C, how do we rank when Runner C beats runner A?

Deployment

End users will obtain a weekly refreshed power ranking system (individual and team), housed on a Shiny App powered by R.