Could an Independent Yorkshire Win the World Cup - Simulate World Cups

Robert Hickman

2018-06-07

Recently, a Yorkshire national football team appeared in a league of national teams for stateless people. This got me wondering how the historic counties of the UK would do at the world cup. Could any of them compete with full international teams?

This is the complete script for an short article I wrote for CityMetric on the topic. It's split over 5 separate parts and is pretty hefty but contains pretty much everything you need to clone the article. In the last post, we located the place and county of birth for British players, which we'll use to pick teams for counties now.

```
library(dplyr)
library(magrittr)
library(data.table)
library(ggplot2)
```

Get County Rankings

Now that we have the teams for each county, we want to work out how well they would do at a world cup. For this, we need to know roughly what their ranking would be compared to actual nations.

Two sources of rankings of nations are the official FIFA world rankings, and also the world ELO ratings of each nation at www.eloratings.net.

I scraped both of these (accurate to mid-May) and cleaned the data to match the nation names to those in the player dataset we're using.

```
#scraped world rankings from FIFA and world ELO
#http://www.fifa.com/fifa-world-ranking/ranking-table/men/index.html
#https://www.eloratings.net/
#accurate for mid-May
#have matched country names between world rankings and FIFA player data
world_rankings <- readRDS("national_rankings.rds")
```

```
#glimpse the data
head(world_rankings)
```

```
##      nation  ELO FIFA
## 1    Brazil 2131    2
## 2   Germany 2092    1
## 3     Spain 2049    8
## 4 Argentina 1985    5
## 5    France 1984    7
## 6  Portugal 1975    4
```

ELO is a chess rating mechanism which can be used to make predictions about which team would win in a matchup. If we compare it to the FIFA rankings, we can see there's a clear negative correlation (the lower the ranking (e.g. top 10 teams in the world), the higher the ELO)

```r
#plot FIFA rankings vs. ELO
p <- ggplot(data = world_rankings, aes(x = FIFA, y = ELO)) +
  geom_text(aes(label = nation))

plot(p)
```