

European League Soccer Competitiveness

A Statistical Exploration

By Dirk and Henok

Business Scenario:

Which European Soccer League is the most competitive?

- Business Investment
 - Instant success
 - ROI
- Team Development
 - Player loans
 - Schedule building
- Coach/Player Consulting
 - Develop skills
 - Increase Profile

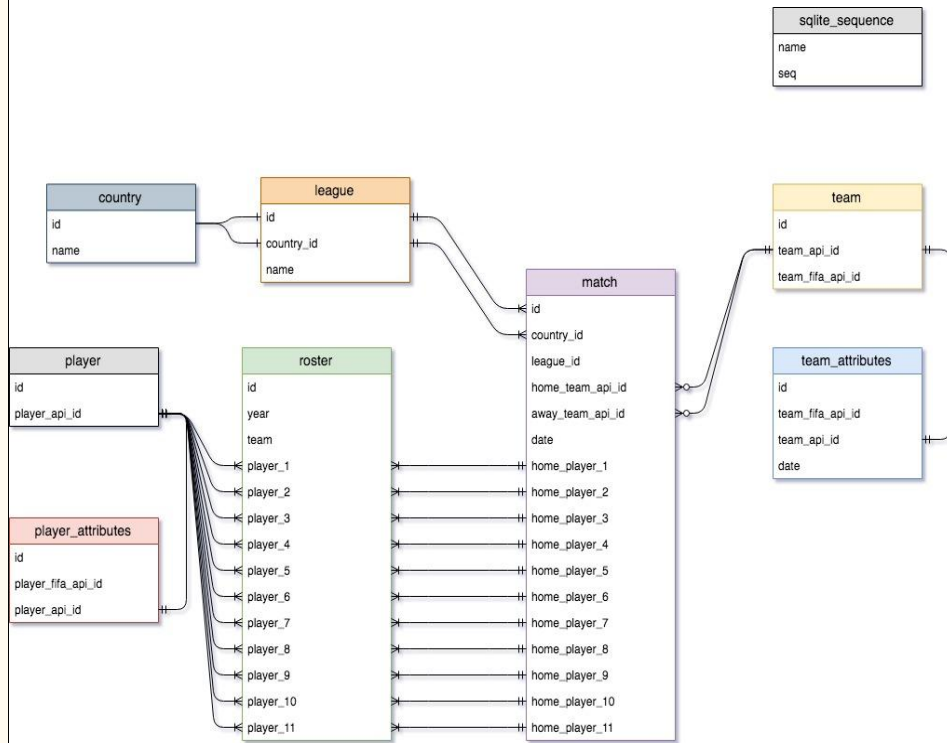




Data

European Soccer Database from Kaggle

- Data from 11 leagues over 8 years including all matches
- EA Sports FIFA video game statistics for players attributes



Approach

The English Premier League provided the baseline for competitive comparison. Our knowledge of the league as well as its reputation as the hardest league in Europe provides us with the flag bearer from which to start our investigation.

The following metrics were what we used to calculate competitiveness:

Goals Scored

Goal Differential

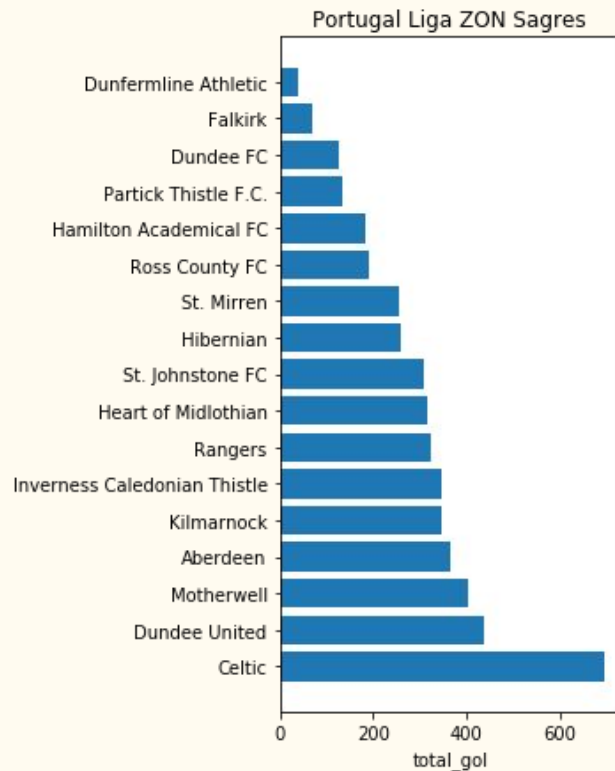
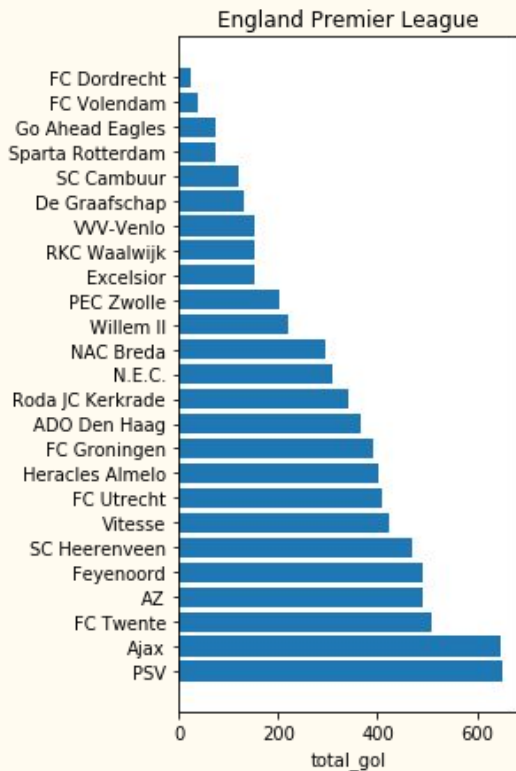
Loss %

Draw %

Total Goals

Total goals proved to have no significant statistical difference to the English Premier League.

Number of goals scored in a season per team was mostly random regardless of league and therefore could not be trusted in predictive analysis for competition.

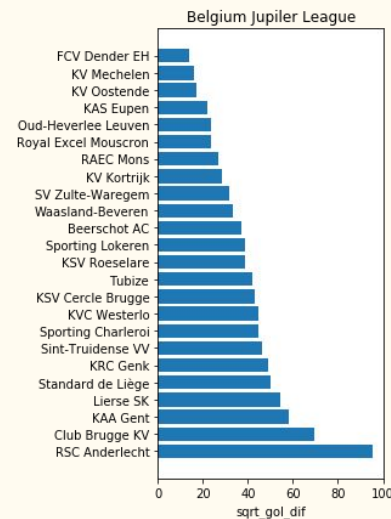
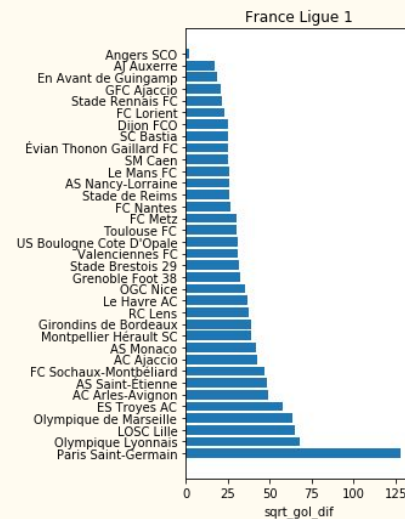
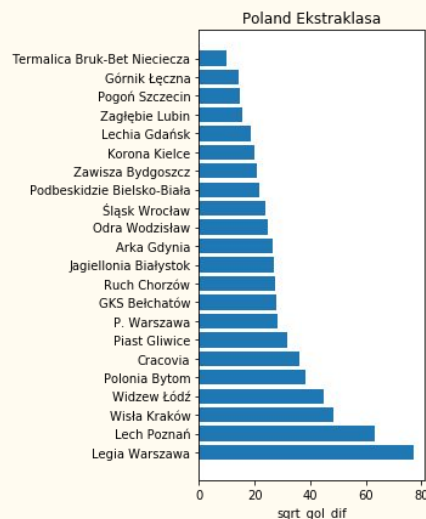
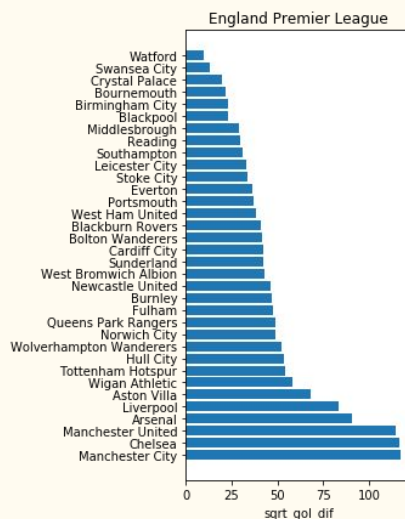


Goal Differential

Goal differential showed significant difference with the French, Belgian, and Polish leagues.

In an effort to help standardize our results we took the square root of the goal differentials of each team within each league.

The leagues with high variance of goal differential illustrates the vast gap in team results making the competition low on average between teams.

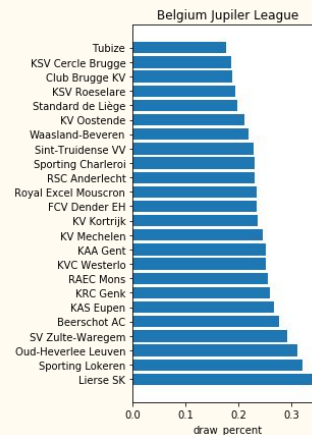
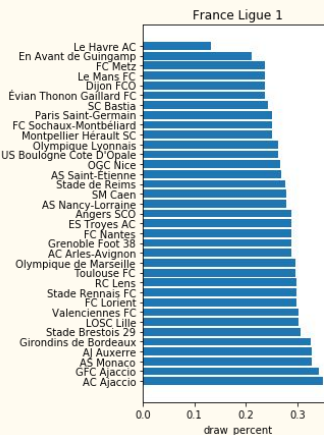
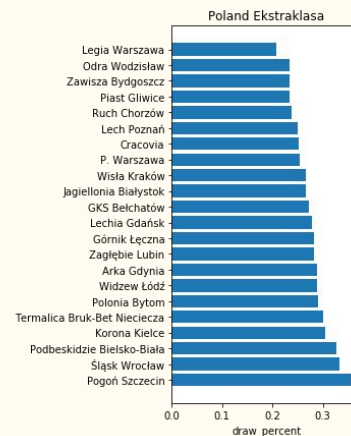
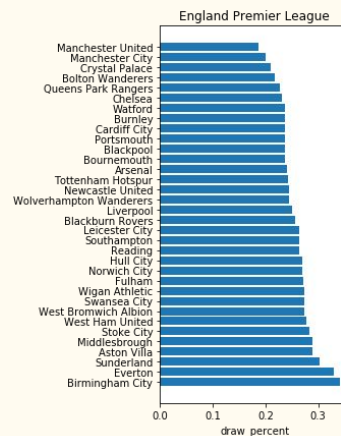


Game Draw Percentage

Draw percentage for several leagues had significant statistical difference.

The Eredivisie, Ekstraklasa, La Liga, and Ligue 1 showed significant statistical difference to the Premier League. These leagues generally have a couple dominant teams which led to less competition on the whole.

Higher draw percentages in a league showed that teams on the whole were evenly matched throughout, much like the Premier League.

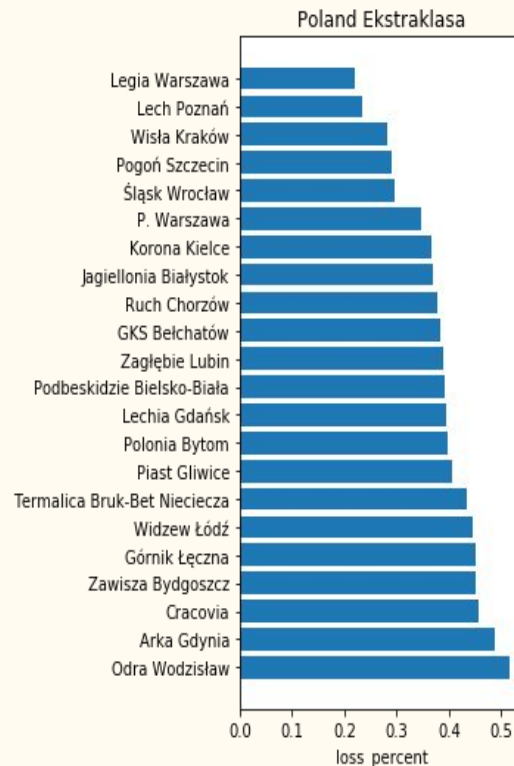
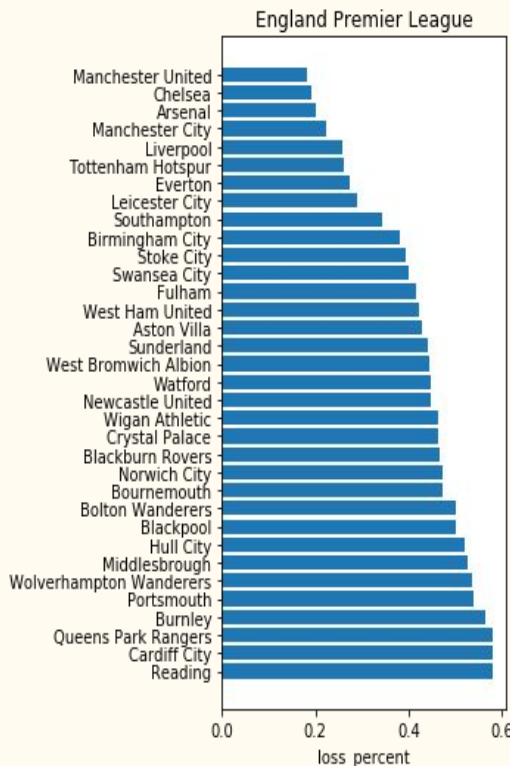


*Eredivisie not shown

Game Loss Percentage

Loss percentage had statistical significance in showing the competitiveness of only the Polish League to the English League.

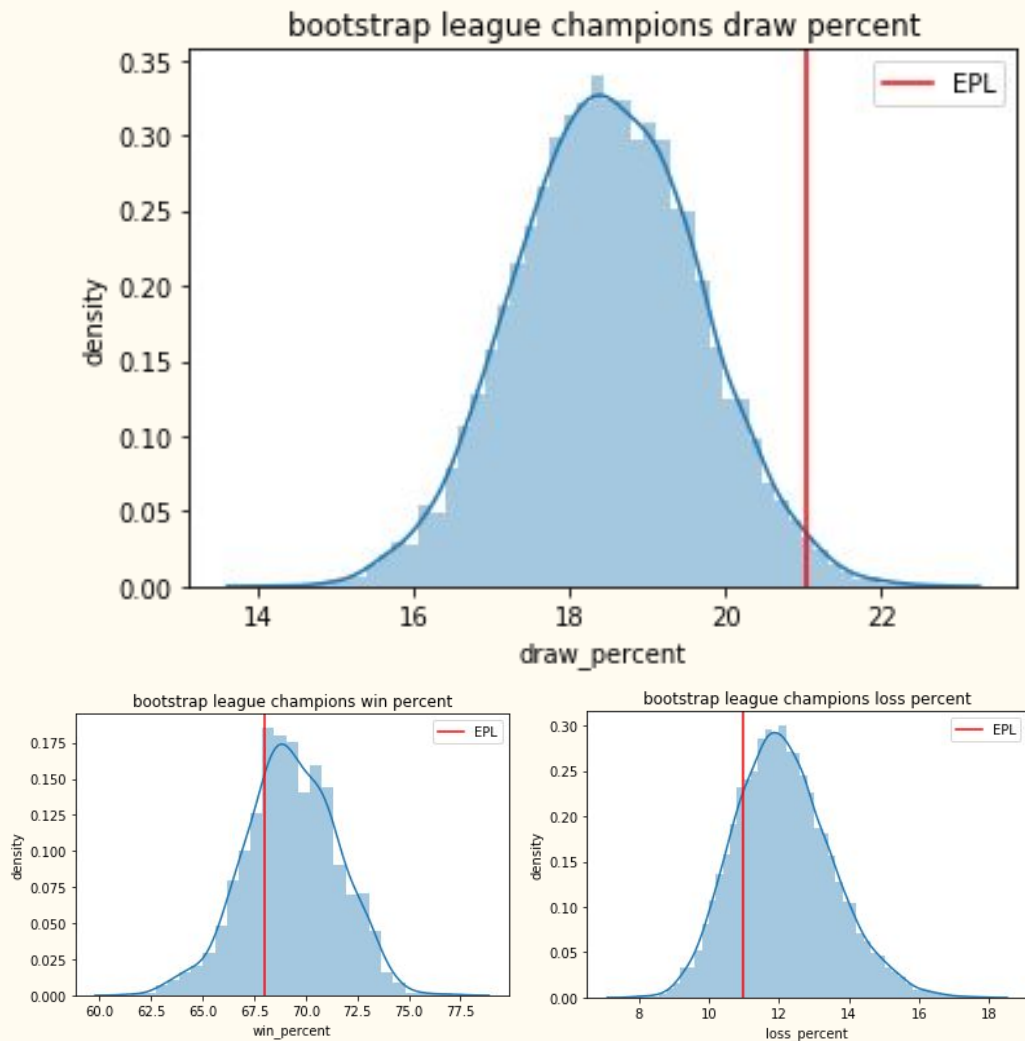
Leagues in general have a fairly equal distribution of losses resulting in no significant difference than the English League except for the Polish League.



Champion Comparison

As further illustration of competitive nature of the Premier League, here we see the extrapolated win, loss, and draw percentage of each league if simulated 10k times.

The champion draw percentage in this scenario falls as an outlier in comparison to other leagues showing even the most successful team can struggle to get wins against the rest of the competition.



Conclusions

- Differences in league Loss %, Total Goals, and Goal Differential varied but not enough for any statistical significance to suggest it as a good overarching indicator of the competitive nature of a league.
- Draw % was our best indicator in showing which leagues had competition like the Premier League.
- Comparing champions over multiple seasons and then extrapolating that data give more significant indication of the competition level for the best team in each league. Ultimately giving us clues to overall ability for any team to compete in said league.

Next Steps

- Mapping the player database to teams per year and continue testing.
- Obtain more in depth team analytical data to help determine other factors that compare to the English Premier League.
- Explore eliminating outliers as the winners and losers in a given season may not represent overall competition.

