Goal-Tier Predictions: Testing Hypotheses on World Football

DREW GOTBAUM

Oberlin College
Introduction to Statistics

Abstract

*I*N *an attempt to predict the outcomes of the many European football leagues, I studied an array of teams belonging to the top-level leagues of the various national football superpowers. Given a dataset of match statistics for each league's 2013-2014 season, I implemented a python software to aggregate the seasonal data-sets by individual teams to produce a modified dataset of European football teams and their respective season-wide data. Through the R programming language's Python library, I was able to run statistical tests to draw conclusions on what attributes effect a team's success/failure. I found that there are a miscellany of factors that hold strong predictive powers of a teams performance, ranging from the number of yellow cards to the average corners per game. If I were to continue my study of European football teams, I would likely examine team performance by the weighted individual performances of the players within the team, for this would likely increase the predictive power of the tests I run.*

I. INTRODUCTION