Datakick Inc.

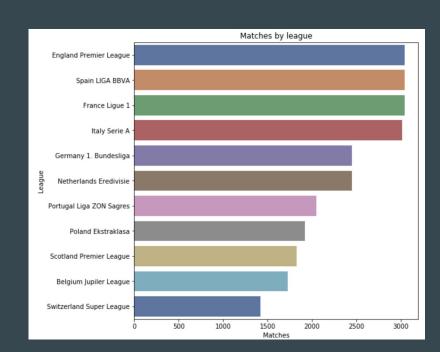
•••

By: Andy Luc & Taeho Jeon

European Soccer Dataset

What is Included in our calculations:

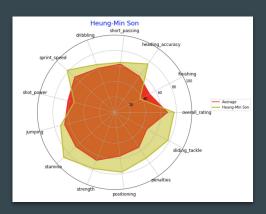
- 11 European Countries with their lead championship
- Seasons 2008 to 2016
- +25,000 matches
- Betting odds from up to 10 providers
- Detailed match events (for +10,000 matches)
 - a. goal types
 - b. possession
 - c. corner
 - d. cross
 - e. fouls
 - f. cards



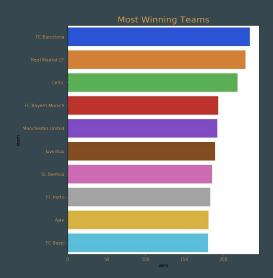
European Soccer Dataset

What is Included in our calculations:

- +10,000 Players
 - a. Height / Weight
 - b. Overall
 - c. Attacking & Defensive Work ...



- **+290** Teams
 - a. Line up with Squad Formation
 - b. Offense (shooting, passing)
 - c. Defense (pressure, aggression)



Focused Sample Questions

Is there a statistical difference in the odds of **winning** or **losing** a match when....

- 1. A team is playing in front of their home crowd?
- 2. The players' height is taller / shorter than other players in the league?
- 3. A team brings stronger defensive pressure during the match?
- 4. A team brings stronger offensive attributes during the match?
- 5. A team is playing in front of their home crowd vs. away crowd?



Methodology - Hypothesis Testing

Brainstorm several Created tables in SQL factors that Set Null and Alternative → Python Dataframe formalizes odds of Hypotheses in response representing each winning / losing to the sample questions hypothesis soccer matches Run through 1-sample Document and 2-sample T-tests, Results with a threshold of 5%

Results - How Can You Benefit?

Result 1: There is no difference when a team wins in front of a home crowd

Result 2: There *is no* difference when a team wins based on the height of team's players.

Result 3: The defense pressure of a team *does* contribute to a win/loss of a match.

Result 4: There *is no* statistical difference when a team wins based on the team's offensive attributes(passing & shooting)

Result 5: There *is* a difference when a team wins in front of a home crowd vs. away crowd

Based on the 4 hypothesis tests that were performed, we have concluded that there are statistical differences when looking at various attributes pertaining to players or the entire team. This data will provide either confidence or additional information to accurately set your betting odds!