

Phase 1

Jonathan Zak 1003146464

Shadman Aziz 1003087991

Domain: Premier (Soccer) league stats for the season 2019/2020

Dataset:

2019-2020 premier league season stats <https://www.football-data.co.uk/englandm.php>

Questions:

Does Chelsea do better in away games?

Does Manchester United do better when the betting odds are against them?

Does M. Oliver give out more cards to away teams compared to home teams?

Schema Definition

Name: 19-20 premier league stats

Games(HomeTeam, AwayTeam, Date, Time, Referee, odds)

HomeStats(HomeTeam, Full Time Home Team Goals, Full Time Away Team Goals, Full Time Result, Half Time Result)

AwayStats(AwayTeam, Full Time Home Team Goals, Full Time Away Team Goals, Full Time Result, Half Time Result)

Referee(Referee, HomeTeam, AwayTeam, Home team foul committed, Away team foul committed, Home team red card, Away team red card, Home team yellow card, Away team yellow card)

TeamStats(Team, wins, losses, draw)

Betting(company, home team, away team, home win odds, draw odds, away win odds)

HomeStats[HomeTeam] \subseteq TeamStats[Team]

AwayStats[AwayTeam] \subseteq TeamStats[Team]

Games[Referee] \subseteq Referee[referee]

Games[HomeTeam] \subseteq TeamStats[Team]

Games[AwayTeam] \subseteq TeamStats[Team]

Betting[HomeTeam] \subseteq TeamStats[Team]

Betting[AwayTeam] \subseteq TeamStats[Team]

Referee[HomeTeam] \subseteq TeamStats[Team]

Referee[AwayTeam] \subseteq TeamStats[Team]