Programming for Data Science

Project Proposal

Group Members:

Varun Muralidharan (VXM180016)
Karthick Ganesan (KXG170003)
Parthasarathy Sundar Karthikeyan (PXS180018)
Abhinandan Mohan Raj (AXM180090)

Data set:

FIFA 19 Complete Player Dataset obtained from Kaggle.

Sample of the attributes in the data set:

Age, Nationality, Overall, Potential, Club, Value, Wage, Preferred Foot, International Reputation, Weak Foot, Skill Moves, Work Rate, Position, Jersey Number, Joined, Loaned From, Contract Valid Until, Height, Weight.

Overview:

One of the most anticipated and sought out games, based on the sport Soccer, is FIFA 19, developed by Electronic Arts. A simulation-based game, where users can play with their favourite team or build a team to their own desires. The game consists of a plethora of different playing modes including:

- Training games
- Arena
- Kick-off
- Manager career
- Player career

Objective:

Our objective is to analyze the data set to arrive at the optimal success rate at excelling in the Manager Career play mode. We plan to use the **Player Valuation, Player Wage, Player Positions, Player age,** and **various player performance parameters** (Passing, Dribbling, Shot Accuracy, Ball Control, Agility, Reactions, Long Shots, Stamina among the few) data for this analysis.

Our analysis will include:

- Finding highly correlated features based on the overall rating by position
- Analysis of the differences in the players current and their potential rating
- Teams with the highest potential
- The youngest teams and oldest teams and their rating
- Use clustering to try to find “bargains”; ie. if there is someone with the same skills/potential, can they be found for a bargain?