

[← Go Back to Data Analysis & Visualization](#)[☰ Course Content](#)

Problem Statement - Fantasy Sports Clustering Analysis

Context:

Fantasy sports are online gaming platforms where participants draft and manage virtual teams of real professional sports players. Based on the performance of the players in the real world, players are allotted points in the fantasy sports platform every match. The objective is to create the best possible team with a fixed budget to score maximum fantasy points, and users compete against each other over an entire sports league or season. Some of these fantasy sports require actual financial investments for participation, with the chances of winning monetary rewards as well as free matchday tickets on a periodic basis.

The fantasy sports market has seen tremendous growth over the past few years, with a valuation of $\$18.6$ billion in 2019. The football (soccer) segment led in terms of market share in 2019, with over 8 million participants worldwide, and is expected to retain its dominance over the next couple of years. Digitalization is one of the primary factors driving the growth of the fantasy sports market as it allows participants the opportunity to compete on a global level and test their skills. With an increase in smartphone usage and availability of fantasy sports apps, this market is expected to witness a globe surge and reach a $\$48.6$ billion valuation by 2027.

Objective:

OnSports is a fantasy sports platform that has fantasy leagues for many different sports and has witnessed an increasing number of participants globally over the past 5 years. For each player, a price is set at the start, and the price keeps changing over time based on the performance of the players in the real world. With the new English Premier League season about to start, they have collected data from the past season and want to analyze it to determine the price of each player for the start of the new season. OnSports have hired you as a data scientist and asked you to conduct a cluster analysis to identify players of different potentials of each player based on previous season performance. This will help them understand the patterns in player performances and fantasy returns and decide the exact price to be set for each player for the upcoming football season.

Data Description:

- Player_Name: Name of the player.
- Club: Club in which the player plays.
- Position: Position in which the player plays.
- Goals_Scored: Number of goals scored by the player in the previous season.
- Assists: Number of passes made by the player leading to goals in the previous season.
- Total_Points: Total number of fantasy points scored by the player in the previous season.
- Minutes: Number of minutes played by the player in the previous season.
- Goals_Conceded: Number of goals conceded by the player in the previous season.
- Creativity: A score, computed using a range of stats, that assesses player performance in terms of producing goalscoring opportunities for other players.

- Influence: A score, computed using a range of stats, that evaluates a player's impact on a match, taking into account actions that could directly or indirectly affect the match outcome.
- Threat: A score, computed using a range of stats, that gauges players who are most likely to score goals.
- Bonus: Total bonus points received. The three best performing players in each match receive additional bonus points based on a score computed using a range of stats. 3 points are awarded to the highest scoring player, 2 to the second best, and 1 to the third.
- Clean_Sheets: Number of matches without conceding a goal in the previous season.

[< Previous](#)[Next >](#)

Proprietary content.©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

© 2023 All rights reserved.

[Help](#)