CMPE-138: Group-14 Project Document

Title

Most Valuable Player(MVP) for a season

Group Members

Amarender Reddy Jakka - 017536610
Pranav Tadepu - 017437368
Harshini Pothireddy - 017513548

Abstract

The objective of this project is to analyze tournament **“ncaa_basketball”** dataset from the bigquery-public-data repository to identify the Most Valuable Player (MVP) for a season. The project will focus on leveraging the knowledge gained on complex queries and a few analytics techniques to extract valuable insights from the dataset.

The project will involve selecting **mbb_games_sr, mbb_teams_sr and mbb_players_games_sr** from the **ncaa_basketball** dataset. These tables contain comprehensive information about tournaments, matches, players, and game statistics.

The project will involve preprocessing and cleaning the data using simple queries to ensure its quality and consistency. This may include handling missing values, standardizing data formats, and removing outliers or irrelevant data points.

The next step will be to design an algorithm to give a score to each player using multiple player attributes like different types of points viz. 3 pointers and 2 pointers scored, fouls, blocks and other player and game level statistics.

Finally, the project will conclude with the identification and validation of the Most Valuable Player for the season based on the analysis of the above mentioned data. The findings will be presented in a comprehensive report, highlighting the methodology used, key insights obtained, and recommendations for further analysis or improvements.

Overall, this project aims to demonstrate the application of data analytics to sports data, specifically focusing on identifying the Most Valuable Player for a tournament season using a dataset from the "bigquery-public-data" repository.

Weekly Schedule:

Week	Activity
03/4 - 10/4	● Explore the datasets from bigquery-public-data repository and select a dataset ● Brainstorm on ideas for the project by going through dataset
11/4 - 17/4	● ERD creation ● Algorithm Design to score a player using player and game level statistics ● Create queries for Pre-processing and data cleaning
18/4 - 24/4	● Create queries for selecting the MVP for each team for a season based on the above algorithm or formula ● Check for any query optimizations and document the performance improvement
25/4 - 02/05	● Report completion ● Prepare for presentation