

DOCUMENTATION

The FIFA 19 dataset is a comprehensive collection of data on soccer players from around the world. It contains information on various attributes such as age, nationality, club, and performance metrics.

The data can be loaded into Python using the `read_csv()` function from the pandas module.

The dataset contains a total of 18,207 observations and 89 columns. Each row represents a unique player, and each column represents a different attribute of the player. The attributes include both categorical and numerical variables.

Some of the categorical variables in the dataset are player name, nationality, club, preferred foot, and player position. The numerical variables include the player's age, potential, overall rating, and various performance metrics such as pace, shooting, and passing.

The age of the players in the dataset ranges from 16 to 45 years old, with the majority of players falling between the ages of 21 and 29. The dataset also contains players from over 164 different nationalities, with the majority of players coming from England, Spain, and Brazil.

The club variable in the dataset contains information on the team that the player is currently playing for. The dataset contains players from over 650 different clubs, with the majority of players coming from the top leagues in Europe such as the English Premier League, Spanish La Liga, and German Bundesliga.

The performance metrics in the dataset are measured on a scale of 0 to 100, with 100 being the highest possible score. These metrics include pace, shooting, passing, dribbling, defending, and physical attributes.

Overall, the FIFA 19 dataset is a comprehensive collection of data on soccer players from around the world. It contains a wide range of attributes,

making it useful for analyzing and understanding the performance of individual players and teams.

This code uses the pandas library to analyze and manipulate a FIFA19 dataset. The dataset is loaded into a pandas DataFrame using the `pd.read_csv()` method.

The first few lines of the code print out some basic information about the dataset, including the number of rows and columns, and displays the first 10 rows of the dataset.

The code then performs several grouping operations to count the number of players by nationality and club, and to calculate the average age, rating, and cost of players by nationality, club, and stadium.

Finally, the code displays the coaches, captains, and titles/victories of each club.