**Clustering**

In the FIFA data, finding the best set of players is imperative. This information is crucial to properly gauge the ability and value.

The dataset consists of each players ability, based on the performance rating or player attributes. Most of the attributes are numerical and are having a range between 0 to 100. There are some other attributes such as weight and height which were not to scale.

It was prudent to use unsupervised learning, since there were no past data and to use supervised models. Clustering was the best method to separate the best players from the average.

The data had very low outliers and numeric values.

PCA was used to reduce the dimension of the data capturing 90% of the information in 17 components.

K- means seemed to be the idle method to apply and separate the players.

The data was scaled or normalized using the standard scaler and the categorical variables were converted into numeric variables using label encoder.

Kmeans clustering algorithm was run with 3 clusters to separate the players.

The best cluster were ranked using cluster score.

Each cluster’s background was analyzed and interesting outcomes were obtained.

**Best cluster**

England, Germany and Argentina were found to have the most number of best players in fifa 19.

**Good Cluster**

England , Germany, France were found to have the most number of good players in fifa 19.

**Average Cluster**

Spain, Brazil and England were found to have the most number of good players in fifa 19.

Striker position seems to dominate the best cluster followed by right and left mid field positions.

Center back position player seems to dominate the good cluster followed by right back and left back.

Center mid field has the most number of worst players.