Final Project

Omar Ahmed Anwar Daif : 20191615312

Ahmed Mustafa Helmy Ahmed Otifi: 20191616312

---

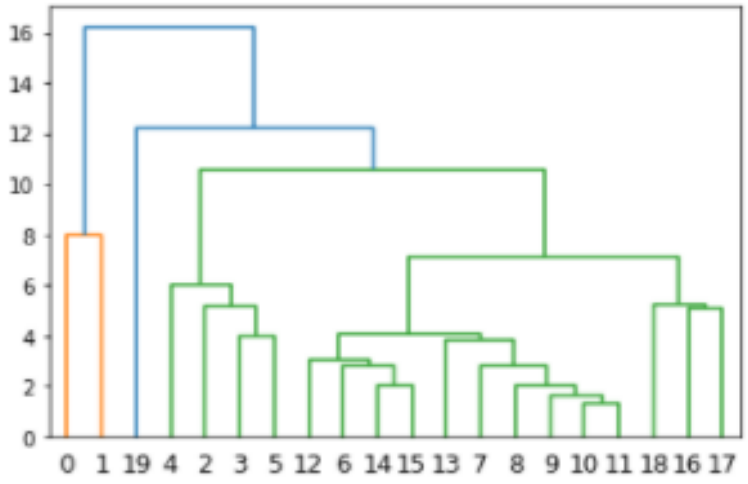
dataset link: https://www.kaggle.com/thesiff/premierleague1819

Dataset description:

Premier League data pertaining to finances, general league table, shots, possession, passes etc… .

Our target is to show what is the effect of the attack possession on the rate of scoring.

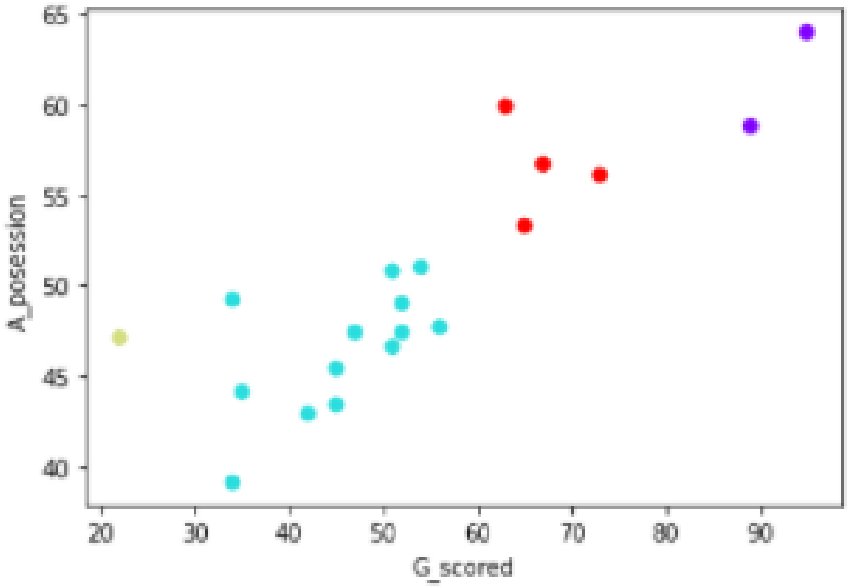
Dendrogram :




```
Wall time: 588 ms
```

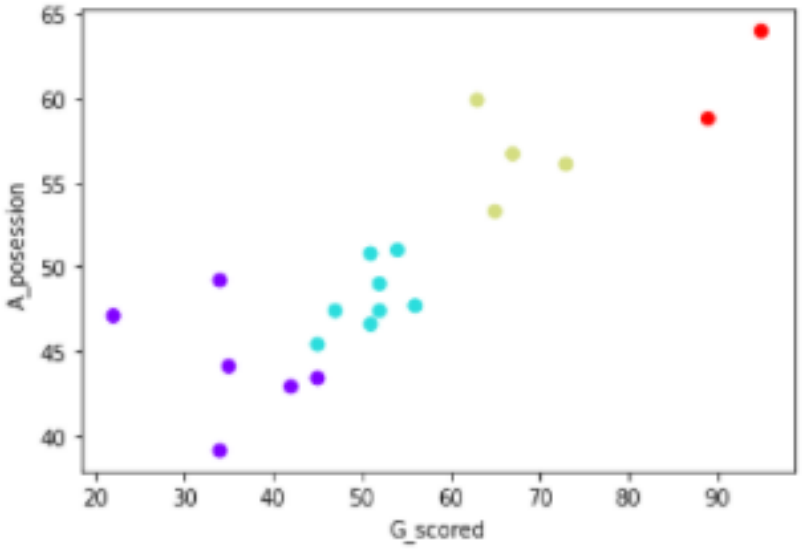
Hierarchical Clustering Graph:

The three cluster groups that represented by light blue , red and purple have positive correlation that means the more the attack have the ball the more they score ,
But for the last cluster group that represented by yellow have negative possession which means that the attack have high possession but have poor scoring record .



K-Medoids Clustering Graph:

The two cluster groups that represented by yellow and red have positive correlation that means the more the attack have the ball the more they score ,
But for the last two cluster groups that represented by purple and light blue have no correlation Which means there is some teams have high attacking possession but poor scoring record and for the other teams have better scoring record but not good enough and poor attacking possession .



The Hierarchical Clustering with measuring centers by Euclidean Distance has better groups than the K-Medoids Clustering because it has no outliers .

Roles:

Ahmed made the Hierarchical Clustering code and had plotted it and did the its report .
Omar made K-Medoids Clustering code and had plotted it and did its report.