CAPSTONE

Defining modern day NBA player positions by using clustering methods.

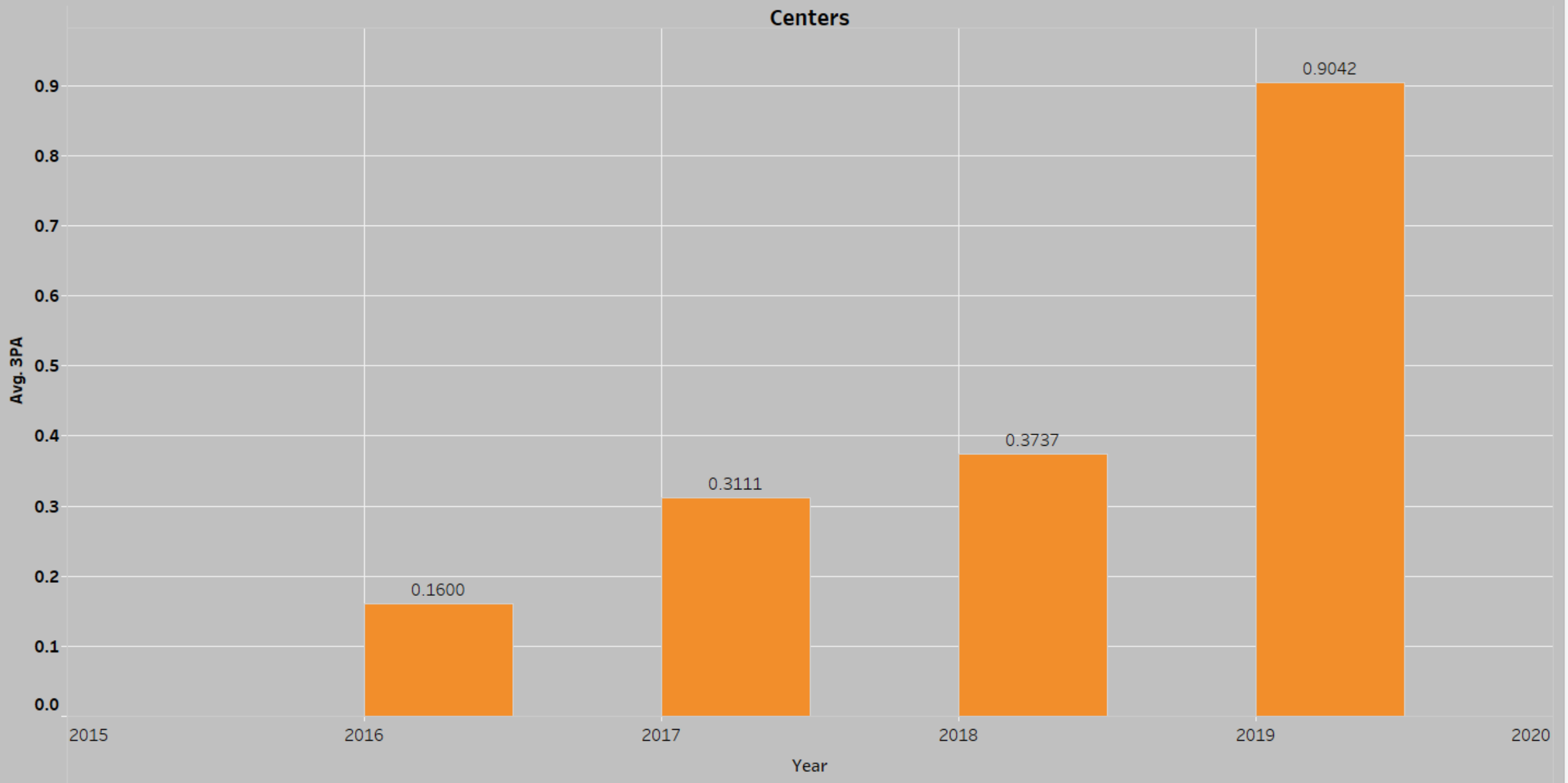
By: Christopher Homsi

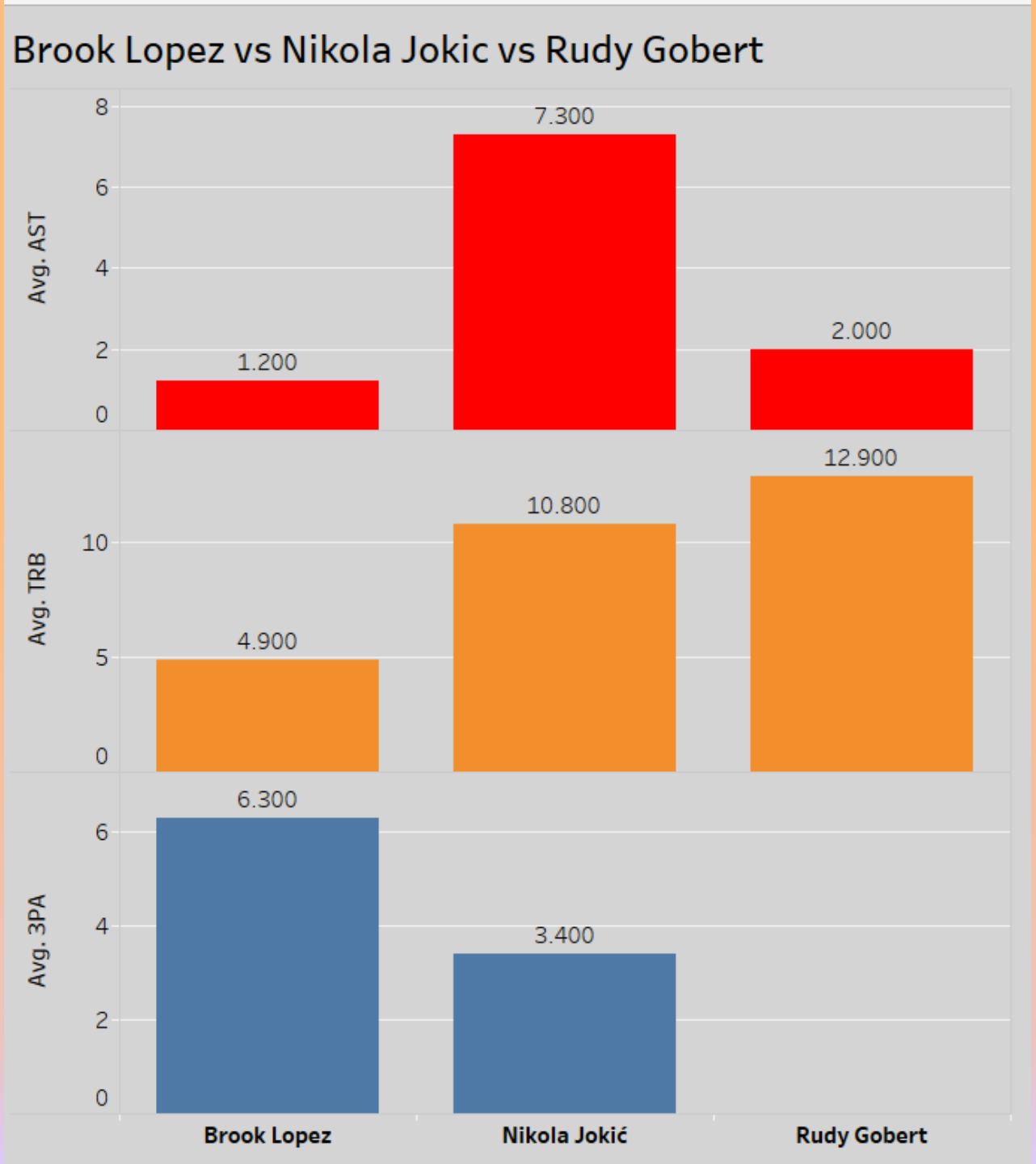
The Evolution of Basketball

“Player positions were created to help fans understand the game” – Jalen Rose

- PG (Point Guard)
- SG (Shooting Guard)
- SF (Small Forward)
- PF (Power Forward)
- C (Center)

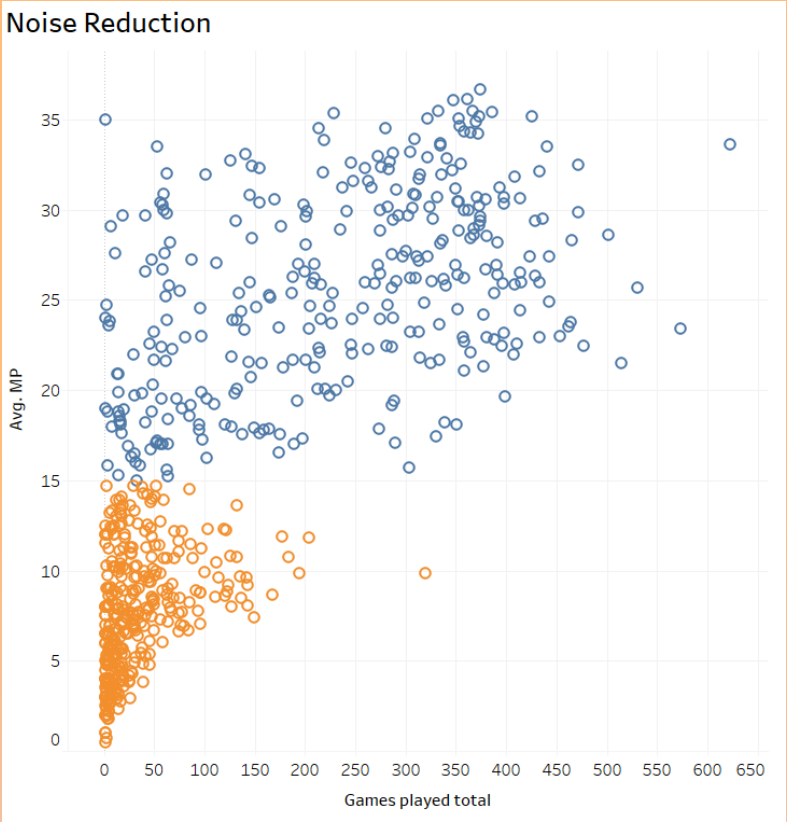
Centers 3pt Attempts




Brook Lopez vs Nikola Jokic vs Rudy Gobert
Avg. AST
8
6
4
2
0
1.200
7.300
2.000
Avg. TRB
10
5
0
4.900
10.800
12.900
Avg. 3PA
6
4
2
0
6.300
3.400
Brook Lopez
Nikola Jokić
Rudy Gobert

Web scraping for Data

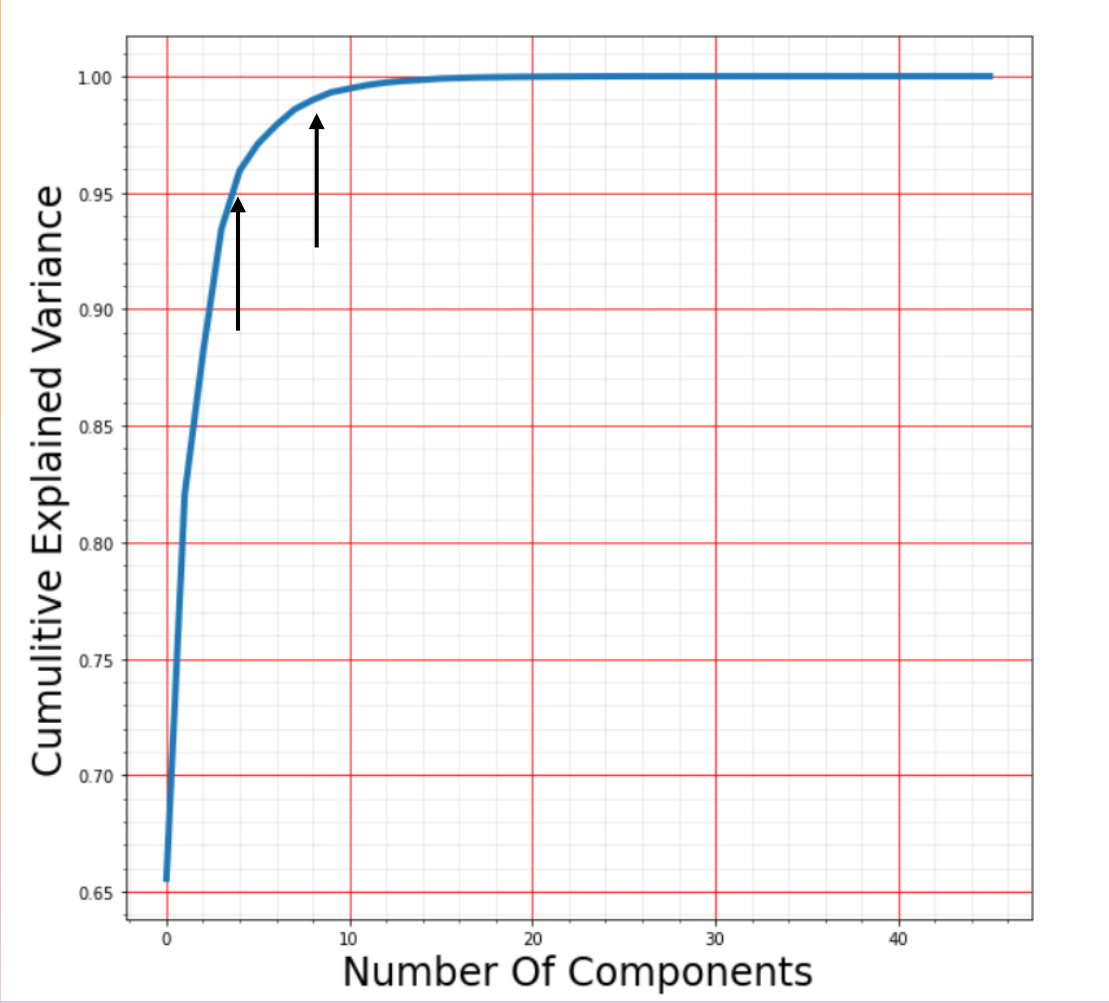
- Beautiful Soup
- Scraped 2016-2019 NBA Seasons
(basic stats merged with advanced stats)
- Got rid of players with less than 15 min played to reduce noise



Modeling using K-Means Clustering

- Step: 1 Standardize our Data
- Step: 2 Apply PCA on our standardized Data
- Step 3: Find number of components of PCA'd data to use

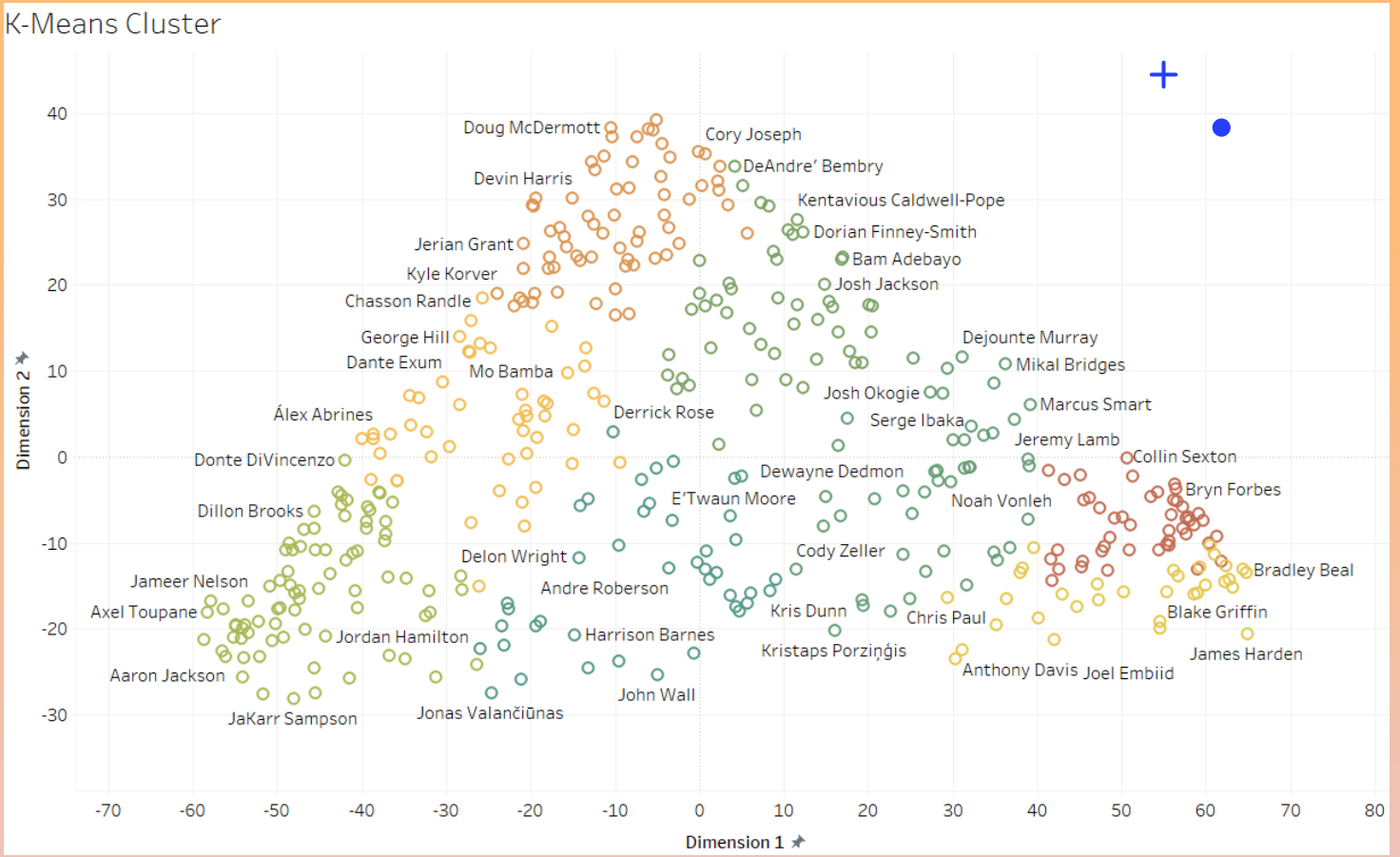
- The First 5 PCA components account for more than 95% of variance in the data.



Modeling using K-Means Clustering cont:

- Step 4: Fit our K-Means Model on the first 5 principal components

- Step 5: Decide how many clusters we want to use. After many attempts between 7-12 clusters, the number of clusters that made the most sense was 8.



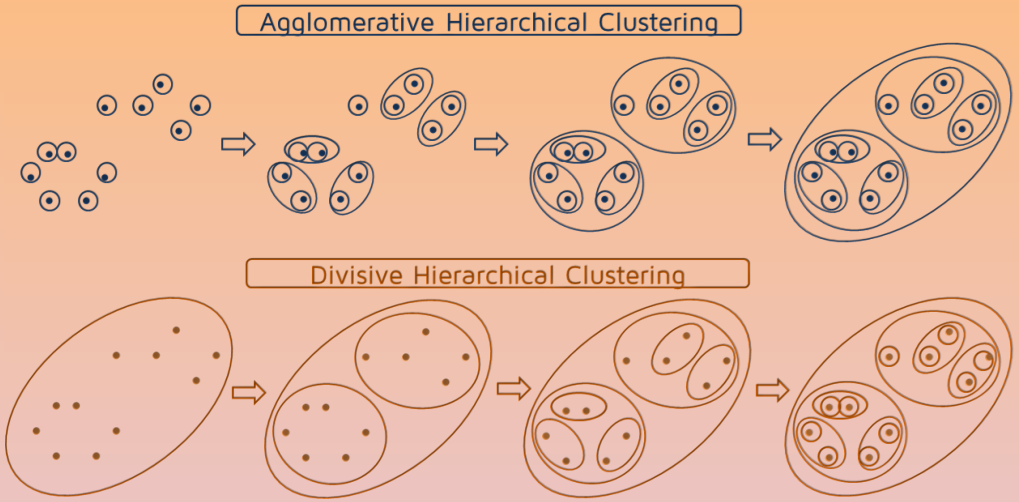
Conclusion: K-Means was not doing a phenomenal job, it was time to try other clustering methods

Modeling using DBSCAN method

- Just using standardized Data here not PCA!
- Using epsilon and min_samples
- Maximum silhouette score was a .05
- DBSCAN was not good!

*Image source https://miro.medium.com/max/1400/1*r1YriAFjwJokgcdtaodQAQ.png*

Hierarchical Clustering



- 2 methods (Divisive and Agglomerative)
- I used Agglomerative with 8 clusters using Euclidian distance and a parameter called affinity which computes the linkage.
- And another parameter called “linkage”. I used was the default called “ward” ward minimizes the variance of the clusters being merged.

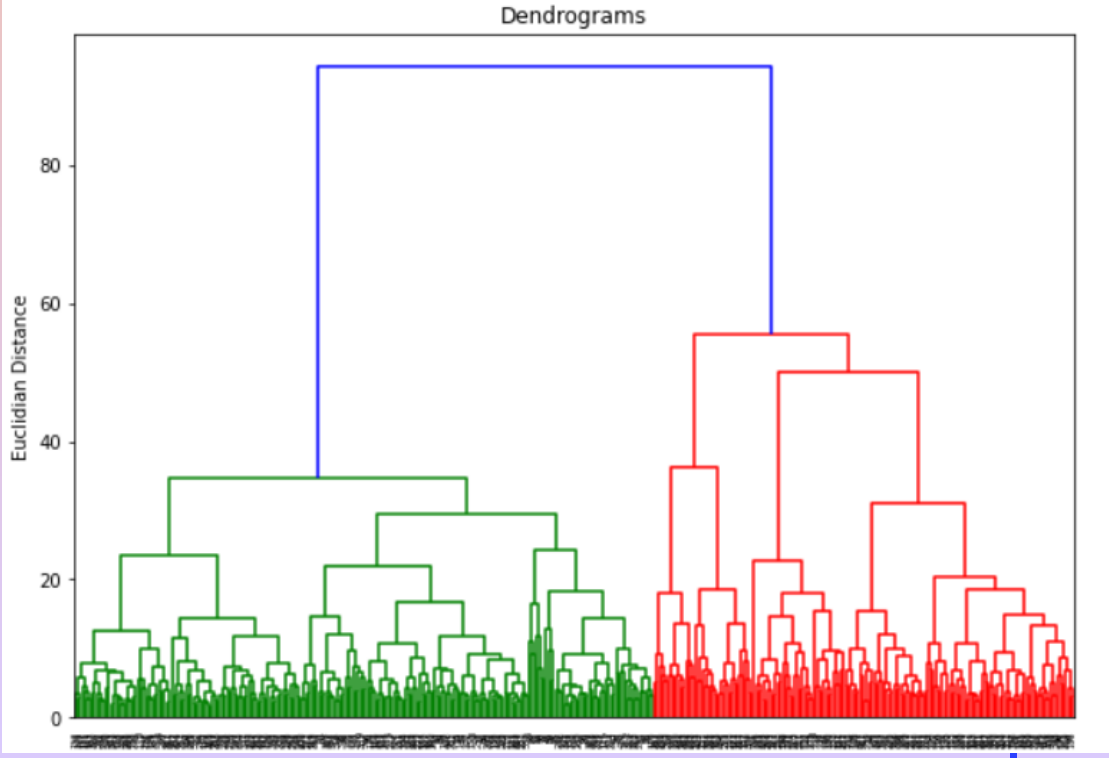


Tableau Interactive Visual of Clusters

- https://public.tableau.com/profile/christopher.homsi#!/vizhome/NBAStats2016-2019playerslatestseasons/Dashboard1

Thank you!

- Any questions?!