

Clustering NBA Players Offensively

Nehal Chigurupati

Contents

1	Objective	3
2	Methodology	4
3	Clustering Results	5
3.1	Isolation	5
3.1.1	Describing Clusters	5
3.1.2	Clusters under PCA	11
3.1.3	Variance of clusters with POSS	12
3.1.4	Variance of clusters with FREQ%	13
3.1.5	Variance of clusters with PPP	13
3.1.6	Variance of clusters with PTS	14
3.1.7	Variance of clusters with FGM	14
3.1.8	Variance of clusters with FGA	15
3.1.9	Variance of clusters with FG%	16
3.1.10	Variance of clusters with EFG%	16
3.2	Cuts	17
3.2.1	Describing clusters	17
3.2.2	Clusters under PCA	22
3.2.3	Variance of clusters with POSS	22
3.2.4	Variance of clusters with FREQ%	23
3.2.5	Variance of clusters with PPP	23
3.2.6	Variance of clusters with PTS	24
3.2.7	Variance of clusters with FGM	24
3.2.8	Variance of clusters with FGA	25
3.2.9	Variance of clusters with FG%	25
3.2.10	Variance of clusters with EFG%	26
3.3	Off screen	26
3.3.1	Describing clusters	26
3.3.2	Clusters under PCA	31
3.3.3	Variance of clusters with POSS	32
3.3.4	Variance of clusters with FREQ%	33
3.3.5	Variance of clusters with PPP	34
3.3.6	Variance of clusters with PTS	35
3.3.7	Variance of clusters with FGM	36

3.3.8	Variance of clusters with FGA	37
3.3.9	Variance of clusters with FG%	38
3.3.10	Variance of clusters with EFG%	39
3.4	Pick and roll ball handler	39
3.4.1	Describing clusters	39
3.4.2	Clusters under PCA	44
3.4.3	Variance of clusters with POSS	45
3.4.4	Variance of clusters with FREQ%	46
3.4.5	Variance of clusters with PPP	47
3.4.6	Variance of clusters with PTS	48
3.4.7	Variance of clusters with FGM	49
3.4.8	Variance of clusters with FGA	49
3.4.9	Variance of clusters with FG%	50
3.4.10	Variance of clusters with EFG%	51
3.5	Pick and roll roller	51
3.5.1	Describing clusters	51
3.5.2	Clusters under PCA	55
3.5.3	Variance of clusters with POSS	56
3.5.4	Variance of clusters with FREQ%	56
3.5.5	Variance of clusters with PPP	57
3.5.6	Variance of clusters with PTS	58
3.5.7	Variance of clusters with FGM	59
3.5.8	Variance of clusters with FGA	60
3.5.9	Variance of clusters with FG%	61
3.5.10	Variance of clusters with EFG%	61
3.6	Post up	62
3.6.1	Describing clusters	62
3.6.2	Clusters under PCA	67
3.6.3	Variance of clusters with POSS	68
3.6.4	Variance of clusters with FREQ%	69
3.6.5	Variance of clusters with PPP	70
3.6.6	Variance of clusters with PTS	70
3.6.7	Variance of clusters with FGM	71
3.6.8	Variance of clusters with FGA	71
3.6.9	Variance of clusters with FG%	72
3.6.10	Variance of clusters with EFG%	72
3.7	Spot up	73
3.7.1	Describing clusters	73
3.7.2	Clusters under PCA	76
3.7.3	Variance of clusters with POSS	77
3.7.4	Variance of clusters with FREQ%	78
3.7.5	Variance of clusters with PPP	79
3.7.6	Variance of clusters with PTS	80
3.7.7	Variance of clusters with FGM	81
3.7.8	Variance of clusters with FGA	82
3.7.9	Variance of clusters with FG%	83

3.7.10	Variance of clusters with EFG%	84
4	Analyzing Players	84
4.1	Case Study A0: Nikola Jokic	84
4.2	Case Study A1: Anthony Davis	84
4.3	Case Study A2: Tyrese Haliburton	84
4.4	Case Study A3: Anthony Edwards	84
5	Analyzing 2023-24 Lineup Compositions	84
5.1	Lineups	84
5.2	Case Study B0: Philly	85
5.3	Case Study B1: Indiana	85
5.4	Case Study B2: Phoenix	85
5.5	Case Study B3: Milwaukee	86
5.6	Case Study B4: Clippers	86
6	Conclusions	86
7	Further work	86

1 Objective

The objective of this project is to cluster together NBA players based on their offensive game, as determined by their performance on the following types of plays:

- Isolation
- Cut
- Off screen
- Pick and roll ball handler
- Pick and roll roller
- Post up
- Spot up

The clusters were then used in case studies of two types. The first type explored the clusters of a few select players, analyzing strengths and weaknesses in different types of offensive plays. The second type analyzed the cluster compositions of the five lineups in the NBA with the highest offensive rating this season.

Such clusters could be used by a team in a number of ways. If a front office is looking for a hidden gem in free agency to replace a player, they could look at everyone within the cluster associated with the departing player. Strategy-oriented coaches could use the model to craft lineups and plays, and development coaches could use clusters to find appropriate film in the off-season.

2 Methodology

For each play type, the following statistics over the 2023-24 season were collected on a per player basis (averaging over games):

- number of possessions
- frequency
- average points per possession
- total points
- field goals made
- field goals attempted
- field goal percent
- effective field goal percent
- free throw frequency
- turnover frequency
- shooting foul frequency
- and one frequency
- score frequency

For each play type, the corresponding player data was scaled and normalized. Three different types of models were tested: KMeans, spectral clustering with affinity matrix calculated via RBF, and spectral clustering with affinity matrix calculated via nearest neighbors. For each model type, seven different models were trained, one for each play type.

For KMeans, the number of clusters for each play type model was computed via the maximum of silhouette scores over a range from 2 to 20 clusters. The same technique was used for spectral clustering.

Next, all the models were trained. The average of all Davies-Bouldin scores for the KMeans models was computed, and compared to the same averages for the Spectral RBF and nearest neighbor models. The average score was lowest on the KMeans models, leading to it being chosen.

3 Clustering Results

The KMeans models were trained on the following numbers of clusters:

Iso: 3 clusters
Cut: 3 clusters
Off screen: 3 clusters
PnR Handler: 3 clusters
PnR Roller: 2 clusters
Post up: 3 clusters
Spot up: 2 clusters

The average Davies-Bouldin score over all KMeans models was 1.39, indicating a moderately strong clustering. To visualize the clusters (a difficult task, because the feature space is greater than three dimensions), two different techniques were used.

First, PCA on two components was run on the data to project into 2-space, then displayed on a scatterplot with points colored in accordance with their label. The clusters were well visible. **Second**, strip plots were produced to show how the clusters vary over a specific stat.

3.1 Isolation

3.1.1 Describing Clusters

There are three isolation clusters. The seven statistics in which the clustering was most clearly found were:

- number of isolation possessions per game (POSS)
- frequency of isolation plays every game (FREQ%)
- points per isolation possession (PPP)
- average PPG from isolation shots (PTS)
- FGM in isolation per game (FGM)
- FGA in isolation per game (FGA)
- FG% in isolation (FG%)
- EFG% in isolation (EFG%)

The range of values in these statistics for each isolation cluster are listed below.

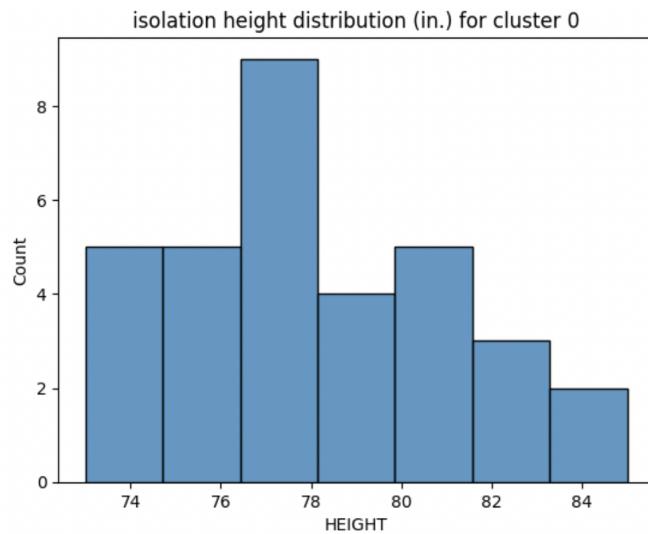
Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	0-2.5	2.5-12.5	0-.85	0-1.5	0-.6	0-2	0-42	0-42
1	2.5-7	11-22.5	.75-1.25	1.5-7.5	.55-2.4	1.9-5.2	35-60	35-58
2	0-2.5	2.5-12.5	.75-1.75	.75-3	.2-1.2	0-2	39-85	38-105

Table 1: Isolation Cluster Values

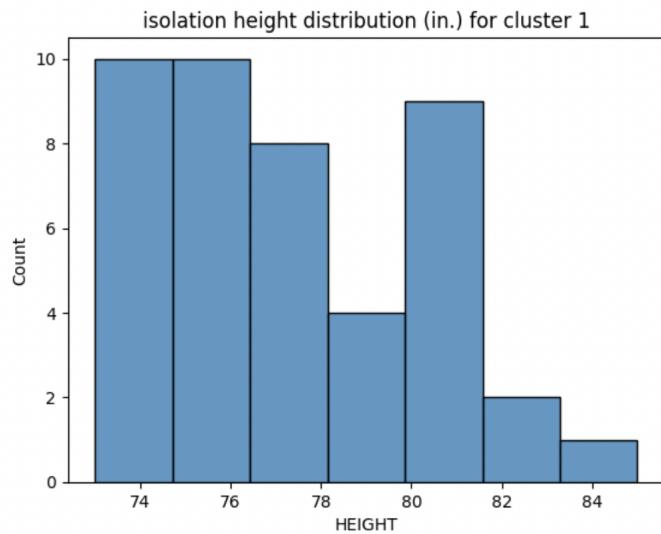
Cluster 0 players attempt isolation plays relatively infrequently, and aren't very productive in such plays. Cluster 1 players attempt a lot of isolation plays, and are moderately efficient at generating points off such possessions. Cluster 2 players don't attempt many isolation plays, but are incredibly talented scorers one-on-one.

A quick gut check with the lists of players belonging to each cluster confirms the above analysis. Cluster 2 includes skilled, tall shooting bigs like Nikola Jokic and Chet Holmgren, who can score reliably over any defender. Cluster 0 includes power and small forwards in the 6'6-6'9 range like Jayson Tatum and LeBron, who back down defenders in isolation frequently and can score moderately efficiently on such plays. Finally, cluster 1 includes shifty point guards who can generate space in isolation, but typically take tough shots that result in a basket less frequently than the post moves of players like Jokic. The fact that cluster 1 players are mostly point guards provides an explanation for why they isolate less than other players, as they are responsible for facilitating plays.

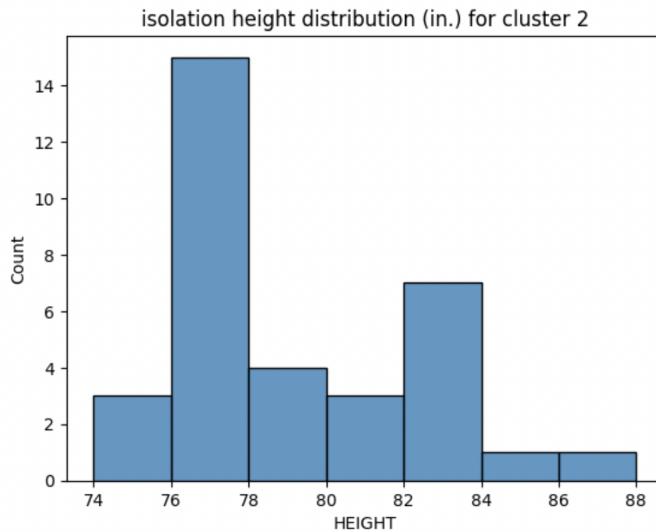
To confirm these identities, the weight and height of the players in each cluster were found. While the mean heights of all the clusters were within the 77-78 inch (6-5 to 6-6) range, there were clear left and right skews in the cluster height distributions that demonstrated the size differences between the different clusters. Cluster 0 players were mostly in the 74-82 inch range (6-1 to 6-8), while shorter players in the <78 inch interval were far more common in cluster 1, confirming the identification of cluster 0 with strong forwards in the 6-6 to 6-9 range (Luka, LeBron, Jayson Tatum) and cluster 1 with shorter, shiftier point guards. Cluster 2 clearly consists of bigs, with a right skew ranging from 78-88 inches (6-5 to 7-3):



The cluster 0 height histogram shows a relatively even distribution, centered around the 74-82 inches.

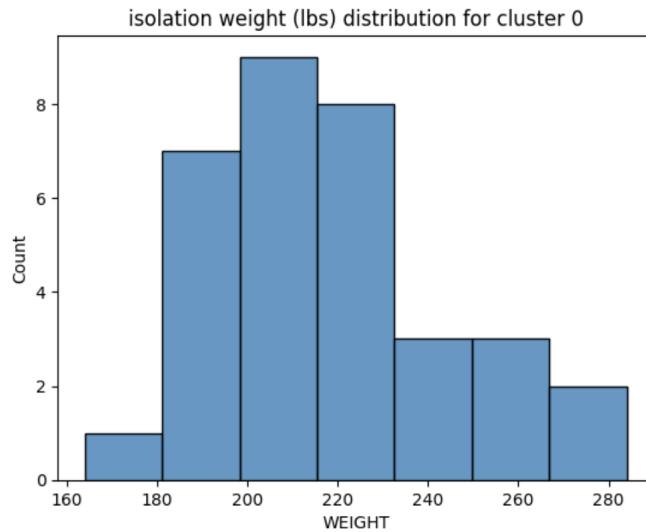


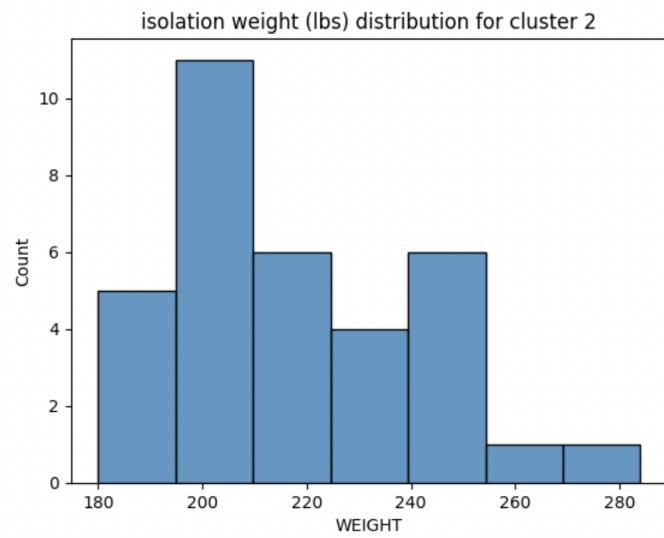
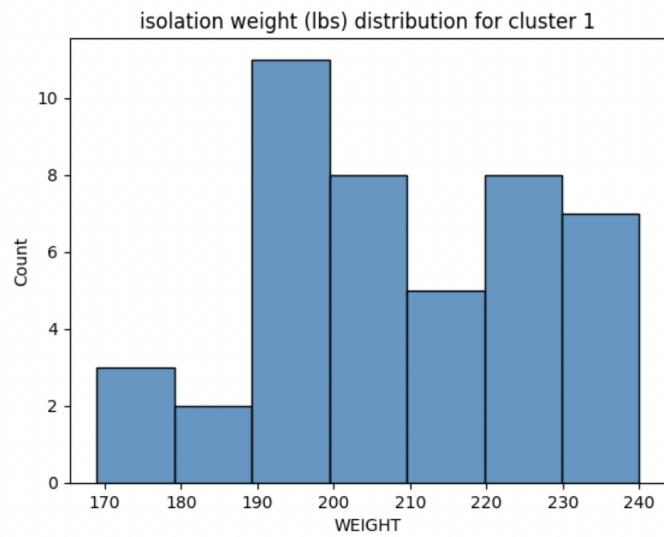
Cluster 1 is dominated by players less than 78 inches tall.



Cluster 2 has a clear right-skew, with a spike in the 82 to 84 inch range.

Weight is distributed along similar lines. Players in clusters 0 and 1 are fairly similar, ranging from 180 to 240 pounds. Cluster 2 players, on the other hand, weigh as much as 280 pounds, with virtually all players 200 pounds or heavier. Such results confirm cluster 2 as consisting of shooting big men, cluster 1 as shifty point guards, and cluster 0 as strong forwards.





Each cluster can then be summarized as follows:

Cluster	Name	Description	Height (in)	Weight (lbs)
0	Strong forward	less efficient, many attempts	74-82	180-240
1	Shifty guard	moderately efficient, few attempts	< 78	180-240
2	Shooting big	highly efficient, few attempts	78-88	200-300

Table 1: Isolation clusters descriptions

Below is a list of a few players in each cluster, in support of the identifications made above.

Cluster 0 players include:

- Luka Doncic
- LeBron James
- Jimmy Butler
- Anthony Edwards
- Kevin Durant
- Brandon Ingram
- DeMar DeRozan
- Jayson Tatum

Cluster 1 players include:

- Cade Cunningham
- Darius Garland
- Dennis Schroder
- Chris Paul
- Collin Sexton
- Jamal Murray
- Josh Giddey

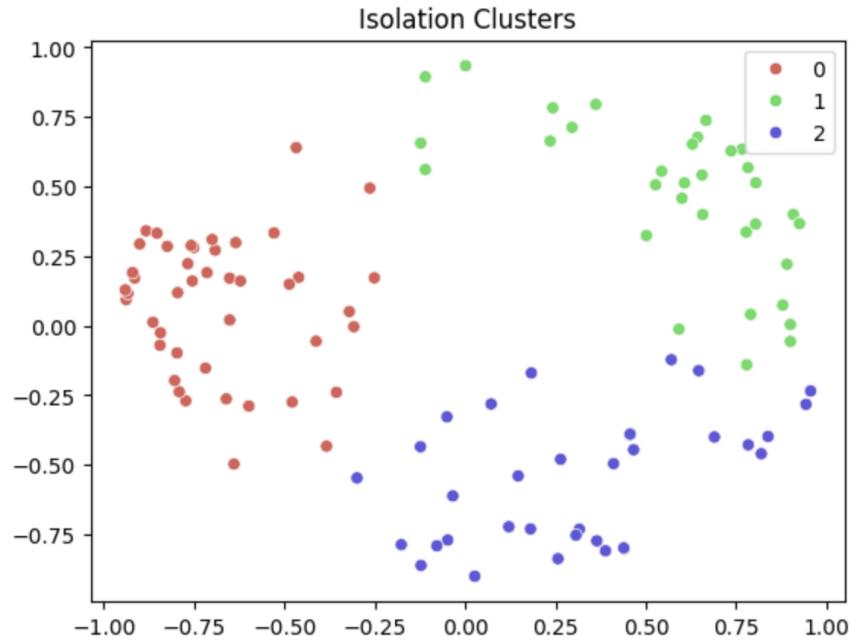
Cluster 2 players include:

- Jaren Jackson Jr.
- Nikola Jokic
- Bam Adebayo
- Anthony Davis

- Victor Wembanyama
- Chet Holmgren
- Evan Mobley

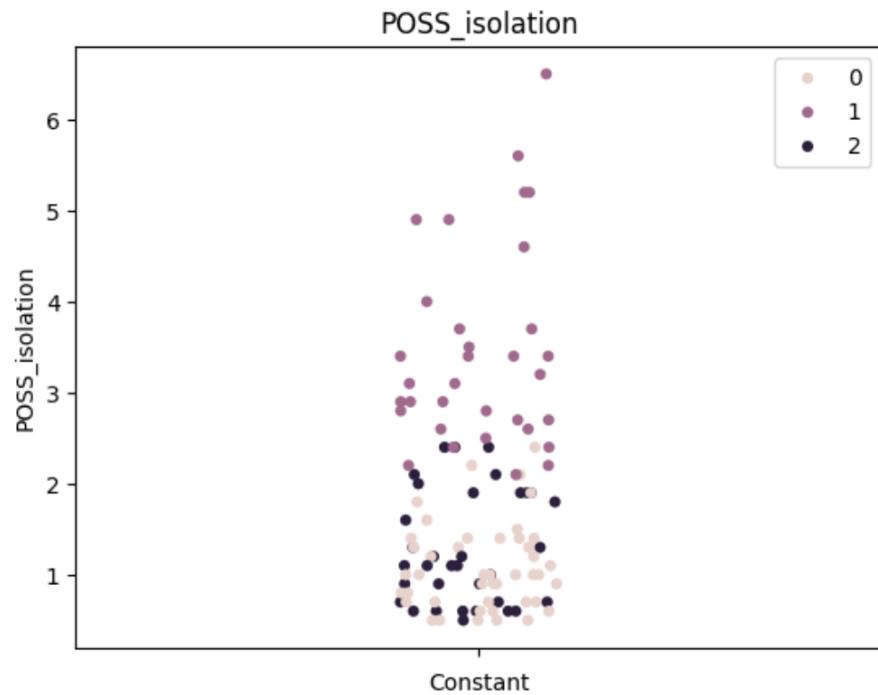
(these are not exhaustive lists)

3.1.2 Clusters under PCA

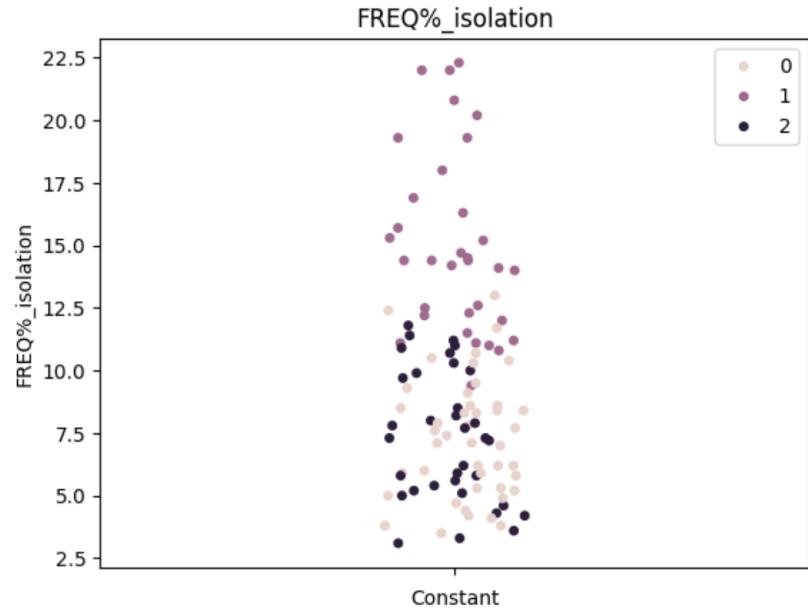


The axes are the two PCA components. It seems like the clustering algorithm is working moderately well for isolation plays.

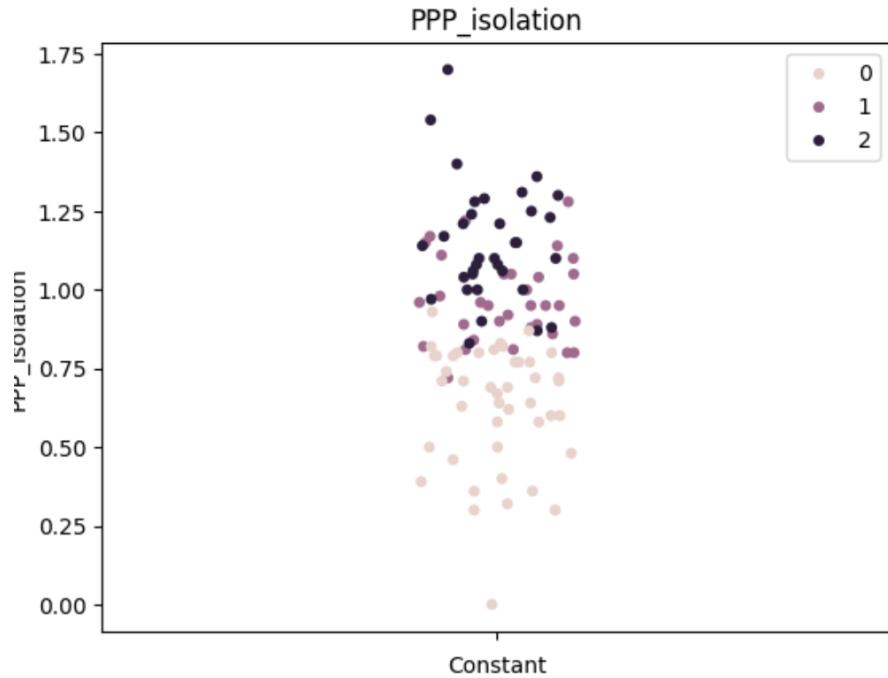
3.1.3 Variance of clusters with POSS



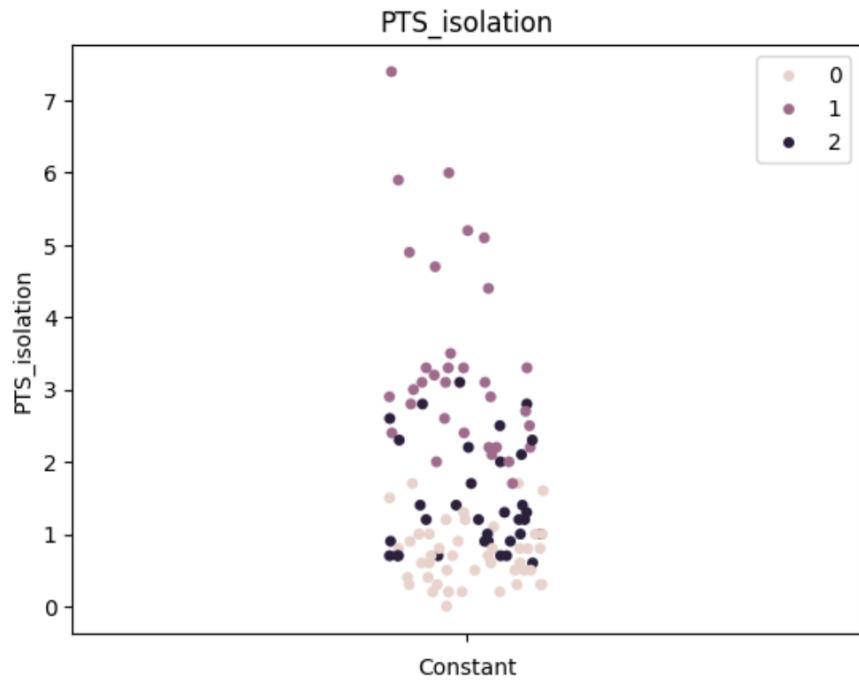
3.1.4 Variance of clusters with FREQ%



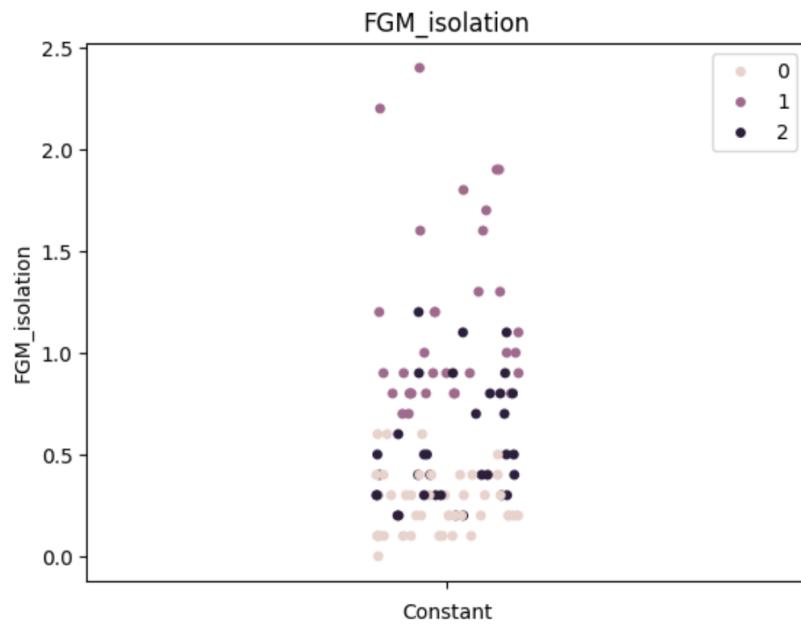
3.1.5 Variance of clusters with PPP



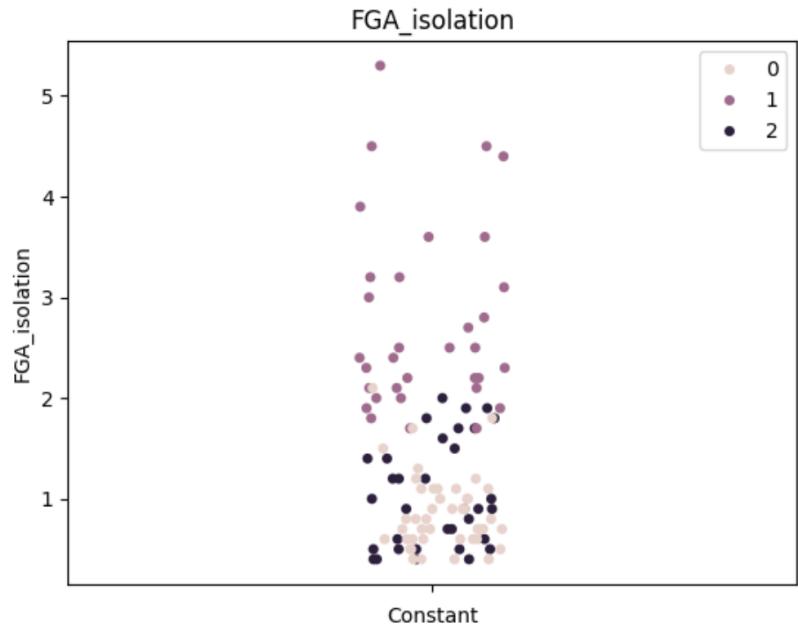
3.1.6 Variance of clusters with PTS



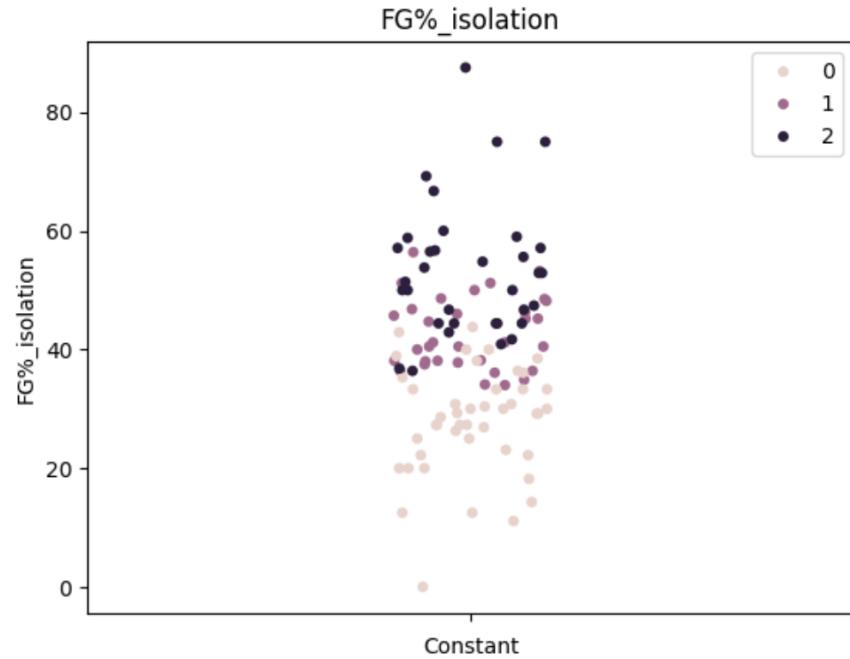
3.1.7 Variance of clusters with FGM



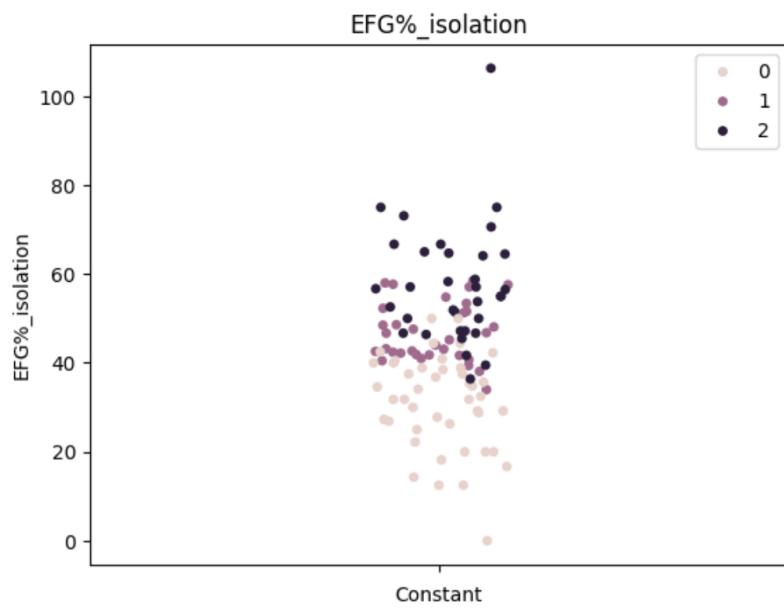
3.1.8 Variance of clusters with FGA



3.1.9 Variance of clusters with FG%



3.1.10 Variance of clusters with EFG%



3.2 Cuts

3.2.1 Describing clusters

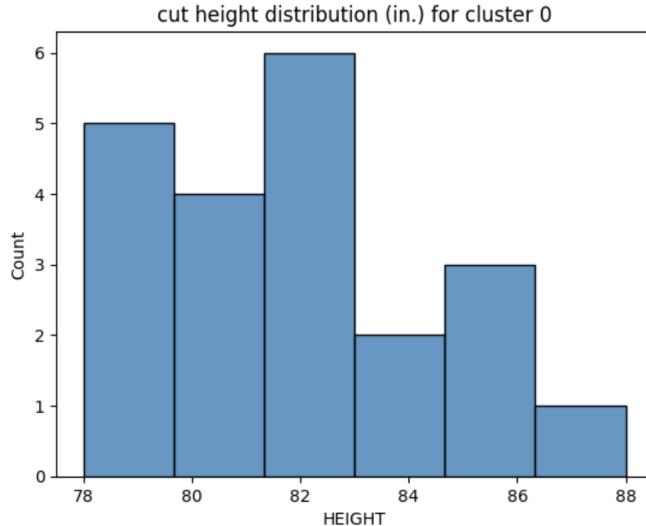
Below are the ranges of each statistic for the cut clusters.

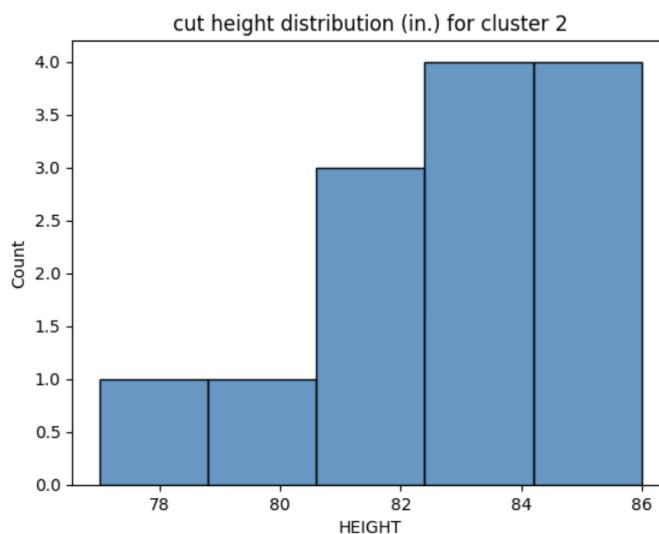
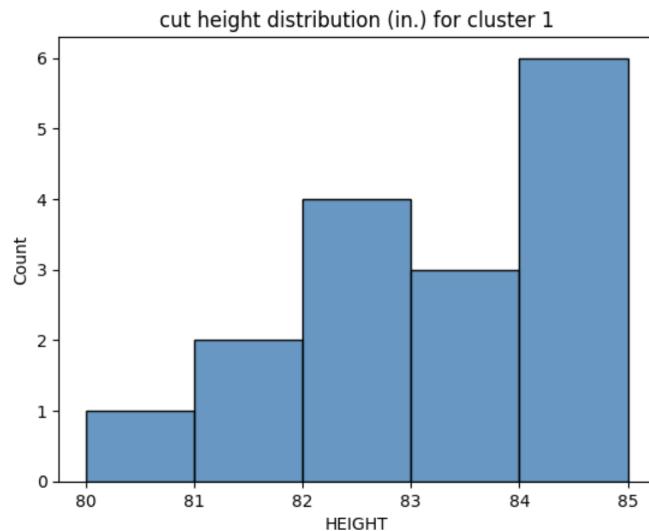
Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	.5-2.4	3.1-11.8	0.83-1.70	0.6-3.1	.2-1.2	.4-2.0	36.4-87.5	36.4-106.3
1	.5-2.4	3.5-13.0	0.00-0.93	0.0-1.7	0.0-6	4.2-2.1	0.0-43.8	0.0-50.0
2	2.1-6.5	9.4-22.3	0.72-1.28	1.7-7.4	0.7-2.4	1.7-5.3	34.0-56.4	34.0-58.5

Table 1: Cut Cluster Values

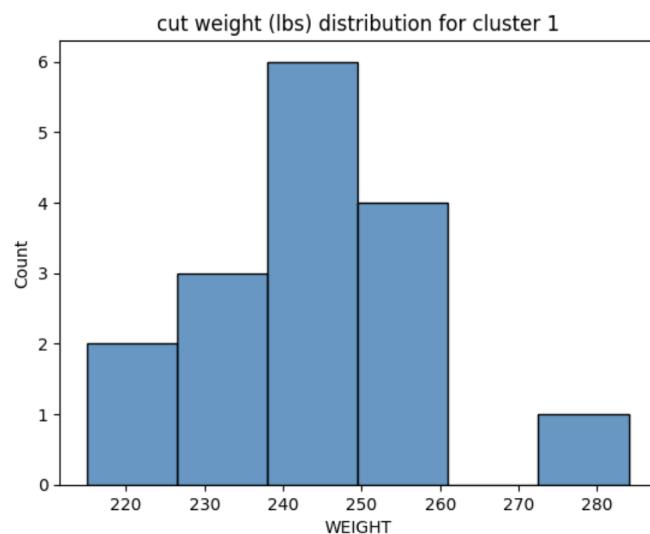
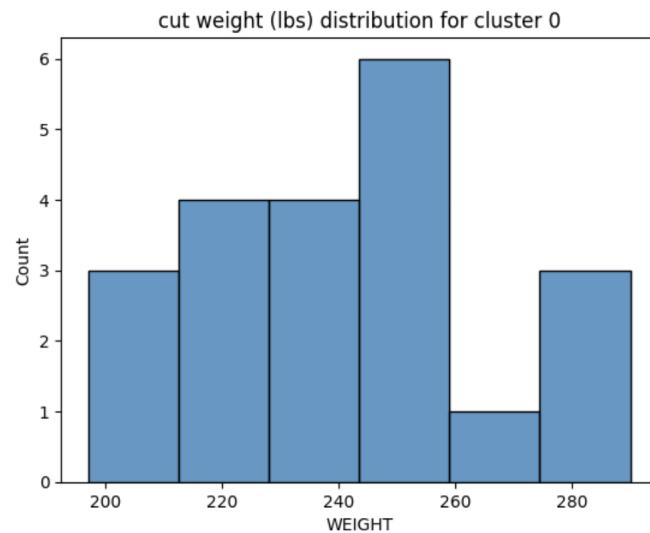
Cluster 0 players do not cut very often, yet score extremely efficiently when they do. Cluster 1 players also do not cut very often, and are also inefficient scorers on those few cuts. Cluster 2 players cut to the basket very frequently and score efficiently, although not to the same threshold as cluster 0 players. Relatively infrequent cuts for clusters 0 and 1 indicate that they either predominantly play on-ball or isolate in the low post. Cluster 2 players play off-ball, cut very frequently, and also score efficiently on cuts, suggesting that the cluster is composed of offensively talented big men who can also play on the perimeter.

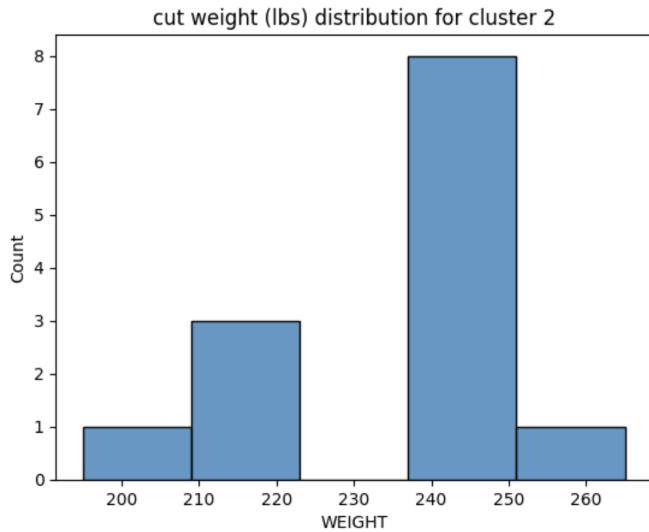
Checking the heights of players in each cluster, we can see that cluster 0 players are generally shorter than cluster 1 and 2 players. The latter two clusters are composed of players with similar heights.





With regard to weights, cluster 2 consists of significantly heavier players than cluster 1.





Cluster 0 included:

- Jusuf Nurkic
- Kevon Looney
- Brook Lopez
- Clint Capela
- Nikola Vucevic
- Marvin Bagley III
- Bam Adebayo
- Alperen Sengun

Cluster 1 included:

- Evan Mobley
- Jarrett Allen
- Domantas Sabonis
- Ivica Zubac
- Anthony Davis
- Nikola Jokic
- Rudy Gobert

- Walker Kessler

- Jalen Duren

Cluster 2 includes:

- Giannis Antetokounmpo
- Obi Toppin
- Jonas Valanciunas
- Goga Bitadze
- Kevin Durant
- Kristaps Porzingis
- De'Andre Hunter
- Keldon Johnson

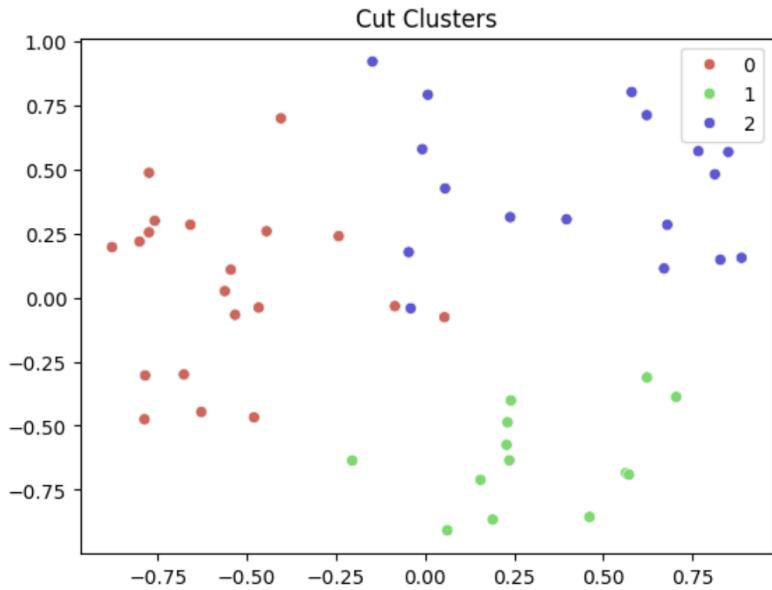
(these are not exhaustive lists)

Broadly, cluster 0 seems to be composed of traditional big men who specialize more in traditional post play than perimeter shooting. Cluster 1 includes shooting big men who perform backdoor and baseline cuts with exceptional mobility, shooting, and passing abilities. Finally, cluster 2 players are athletic and versatile, who have a unique combination of skill and size (Kevin Durant, Chet Holmgren), that allows them to finish at the rim extremely efficiently on cuts.

Cluster	Name	Description	Height (in)	Weight (lbs)
0	Traditional post bigs	cut infrequently, strong post play	78-82	200-250
1	Versatile bigs	cut from perimeter often, mobile	82-85	220-260
2	Dynamic wing forwards	skilled and tall, high level finishers	82-86	240-260

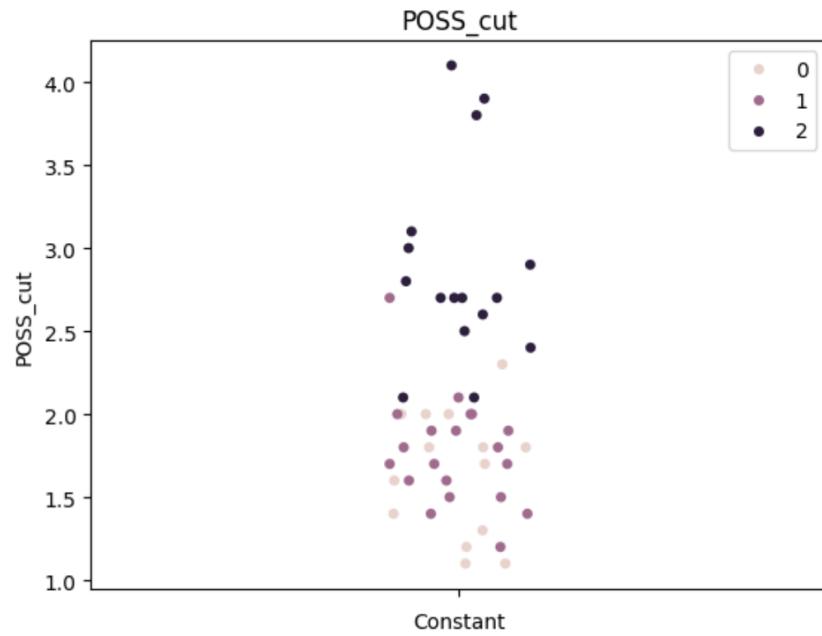
Table 1: Descriptions of cut clusters

3.2.2 Clusters under PCA

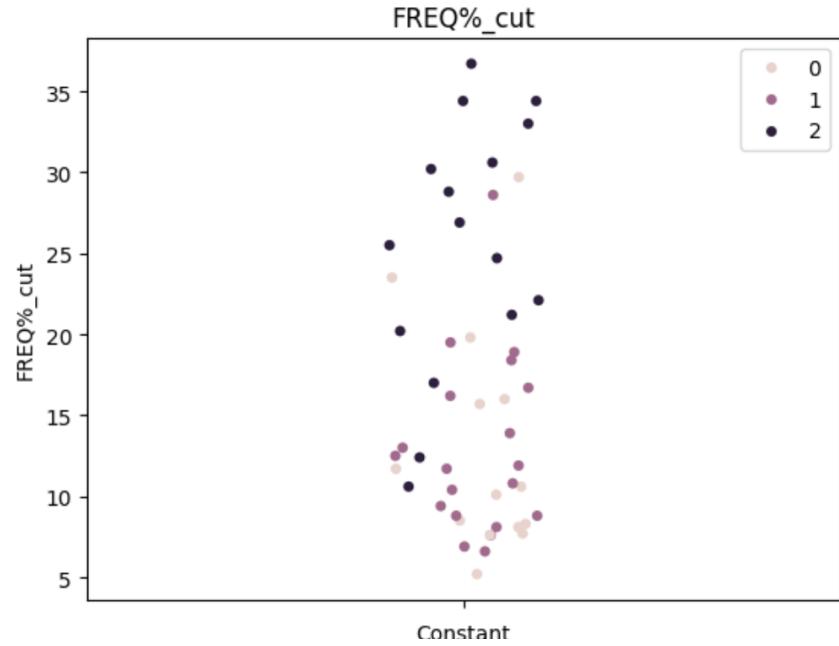


The clustering isn't working well on cuts - a symptom, it seems, of the lack of data.

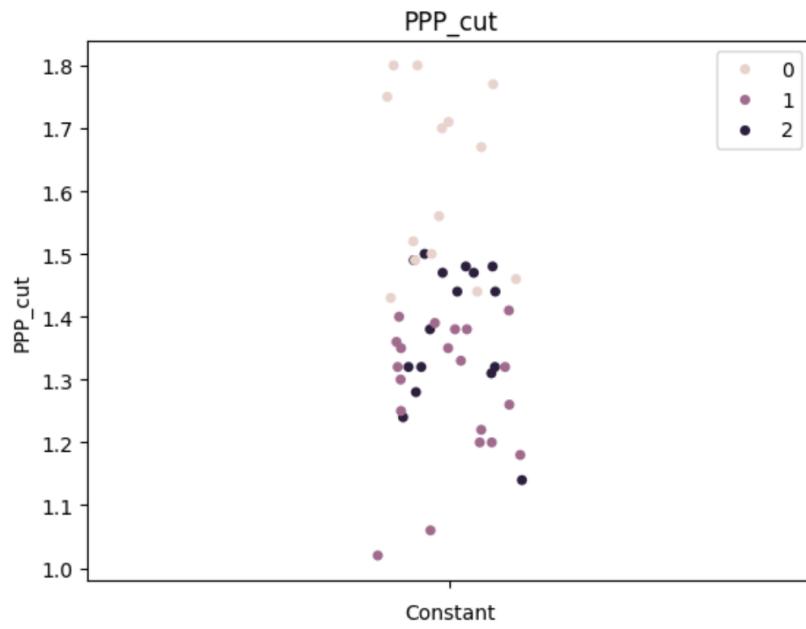
3.2.3 Variance of clusters with POSS



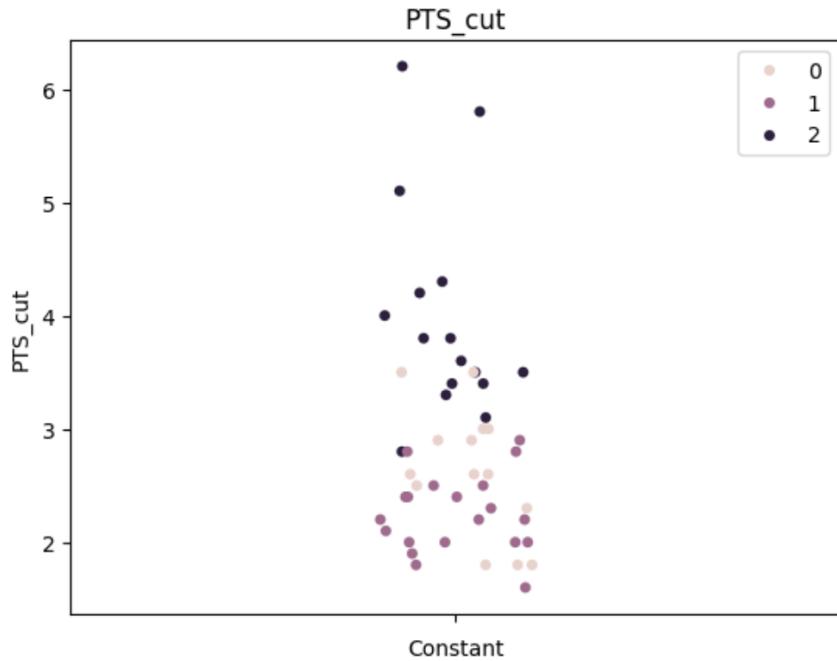
3.2.4 Variance of clusters with FREQ%



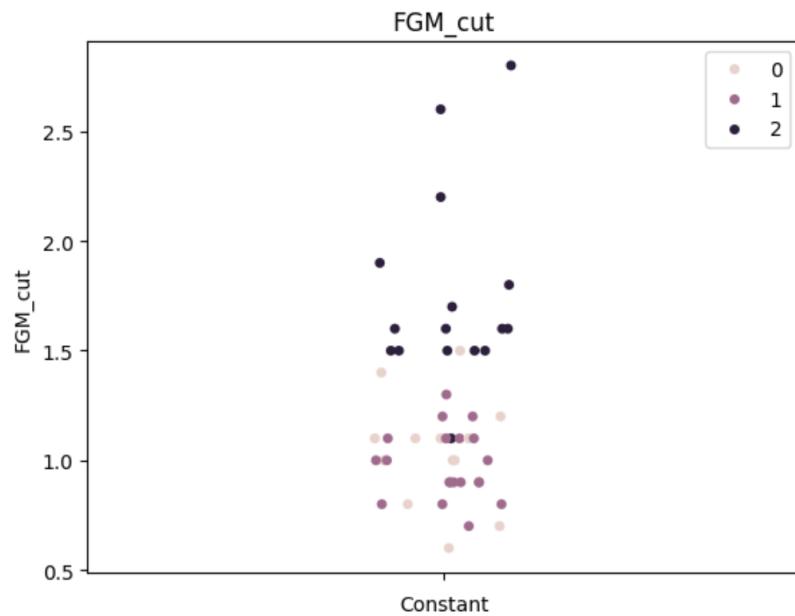
3.2.5 Variance of clusters with PPP



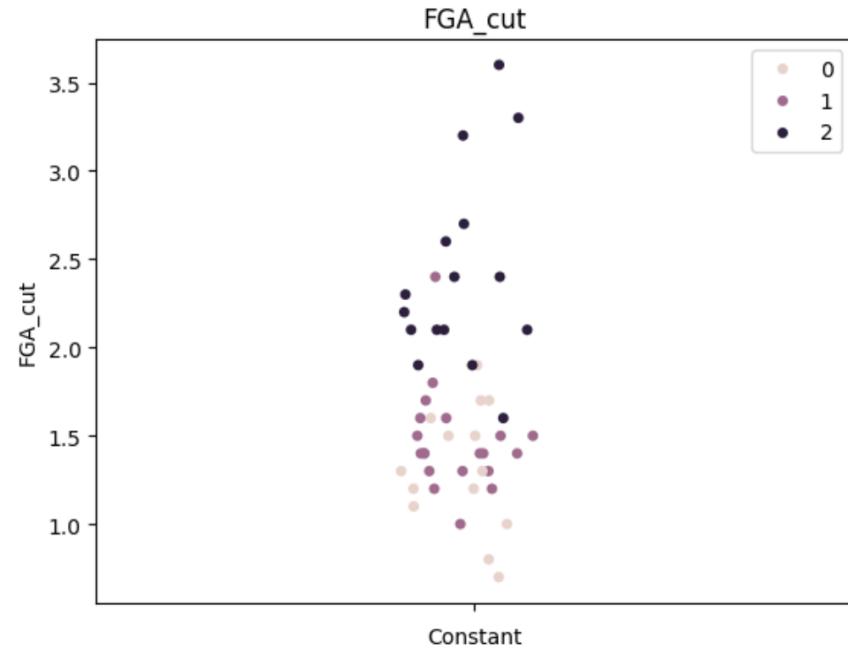
3.2.6 Variance of clusters with PTS



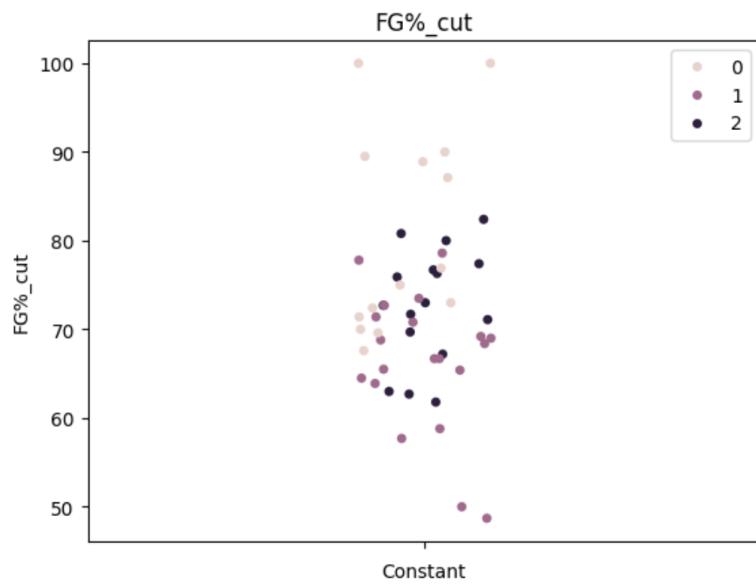
3.2.7 Variance of clusters with FGM



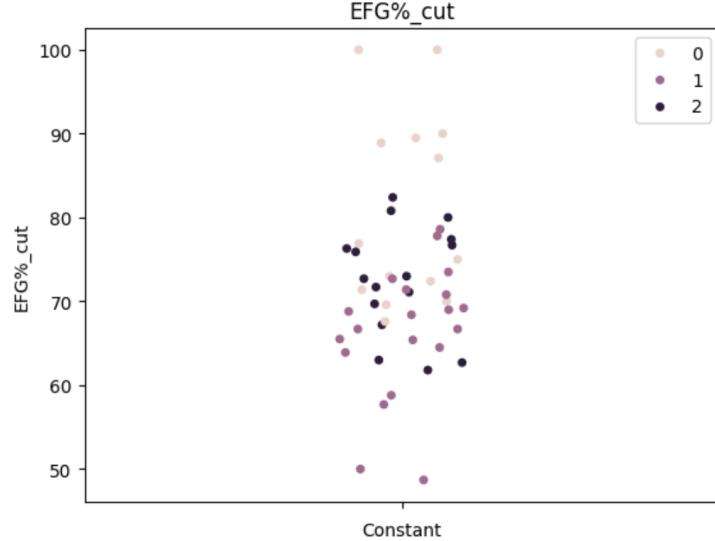
3.2.8 Variance of clusters with FGA



3.2.9 Variance of clusters with FG%



3.2.10 Variance of clusters with EFG%



3.3 Off screen

3.3.1 Describing clusters

The range of values for each cluster are detailed in the following table:

Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	.5-2.3	2.7-18.2	.36-1	.3-1.8	.1-.7	.4-1.8	19-44.4	10-60
1	1.1-5.4	5.4-33.2	.85-1.44	1.1-6.1	.4-2.1	.9-4.7	36.2-61.5	40.6-76
2	.5-1.6	1.8-9.7	1-1.64	.6-1.9	.2-8	.4-1.4	36.4-75	40.9-87.5

Table 1: Off Screen Cluster Values

Cluster 0 includes:

- Jordan Poole
- Jerami Grant
- Max Strus
- Desmond Bane
- Collin Sexton
- Tyrese Maxey
- Donovan Mitchell
- Ochai Agbaji

Cluster 1 includes:

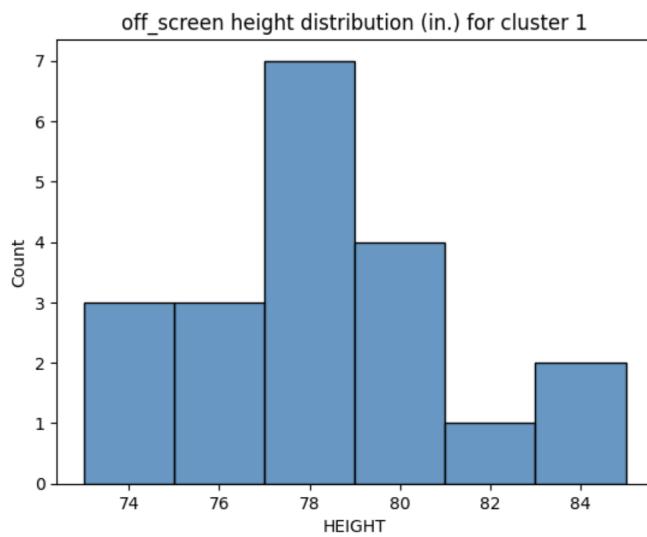
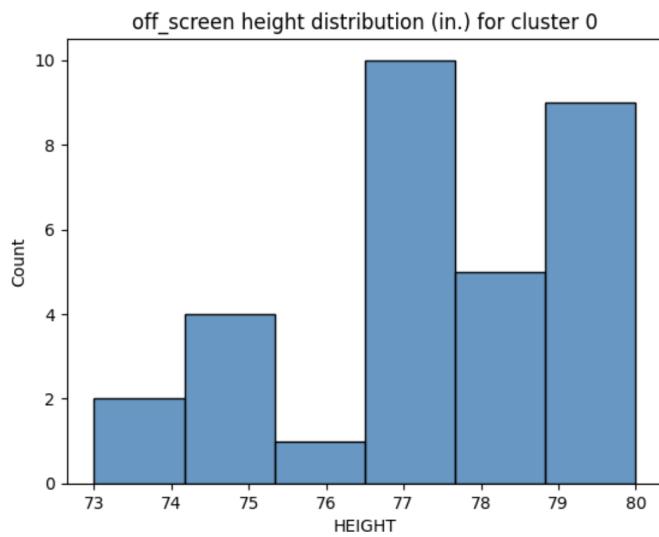
- Paul George
- Brandon Ingram
- Lauri Markkanen
- Kevin Durant
- Michael Porter Jr.
- Bogdan Bogdanovic
- Tim Hardaway Jr.

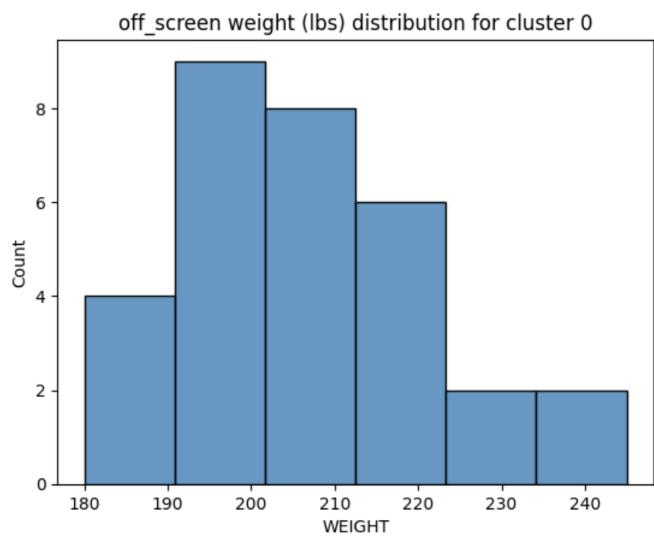
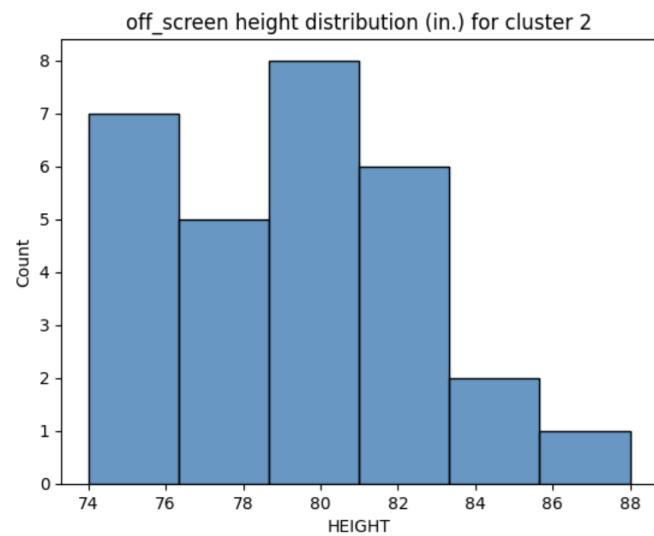
Cluster 2 includes:

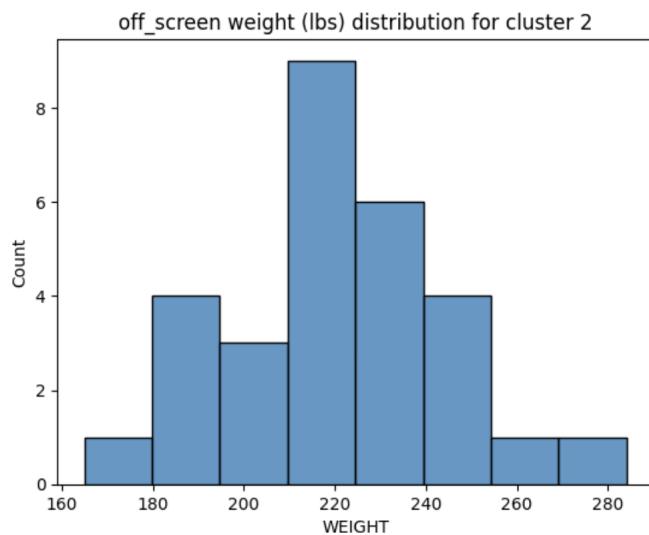
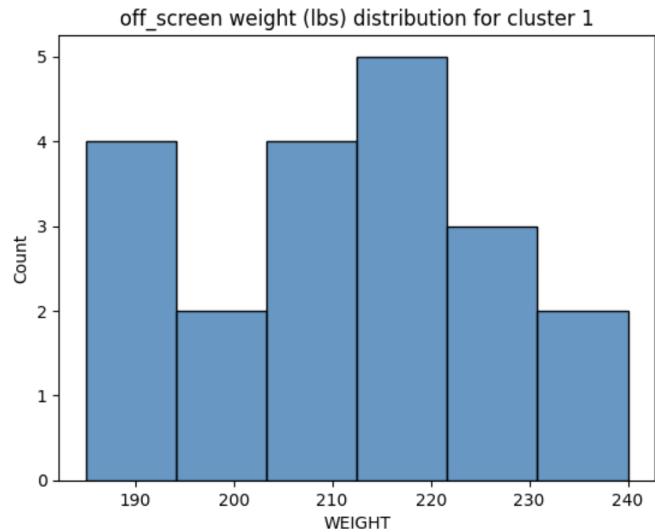
- Nikola Jokic
- Karl-Anthony Towns
- LeBron James
- Luka Doncic
- Chet Holmgren
- Victor Wembanyama
- Cade Cunningham
- Jayson Tatum
- Naz Reid

(these are not exhaustive lists)

Cluster 0 seems to be the shortest of the three clusters (73-80 in.). Cluster 1 seems to be moderately taller than cluster 0 (74-83 in.), while cluster 2 is significantly taller than cluster 1 (74-88 in.).







Cluster 0 consists of dynamic guards who are extremely agile and can score from anywhere on the floor. They generally play on-ball, reducing the frequency of shot attempts from off-ball screens. Cluster 0 players are generally shorter than cluster 1 and 2 players, explaining their lower efficiency.

Cluster 1 contains tall wing scorers who are threats from the perimeter and mid-range. These players use off-ball screens extremely frequently to free themselves up for catch-and-shoot opportunities. They are extremely efficient and

productive in such scenarios.

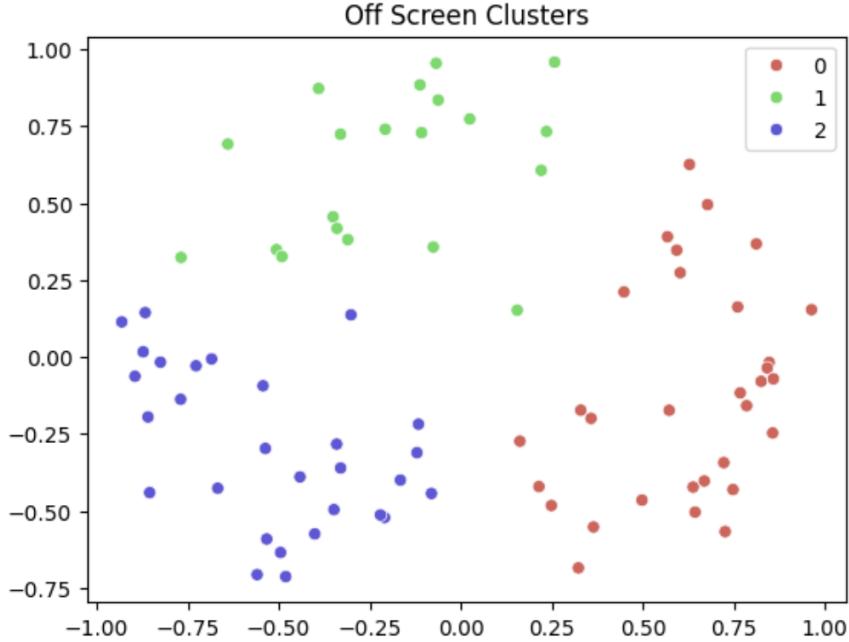
Cluster 2 features playmaking bigs and forwards. These players use off-ball screens to orchestrate plays and create mismatches, explaining the efficient production coming from their off-ball screens.

Each cluster can be summarized as follows:

Cluster	Name	Description	Height (in)	Weight (lbs)
0	dynamic guards	on-ball focus, agile but inefficient	73-80	180-240
1	tall wings	catch-and-shoot specialists	74-83	190-240
2	playmaking bigs/forwards	efficient, create mismatches	74-88	170-280

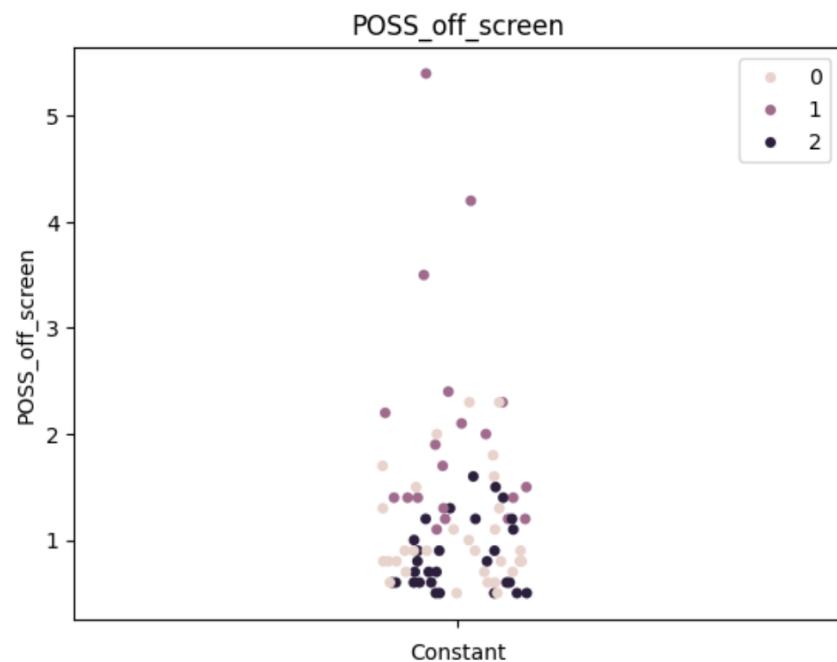
Table 1: Descriptions of off-screen clusters

3.3.2 Clusters under PCA

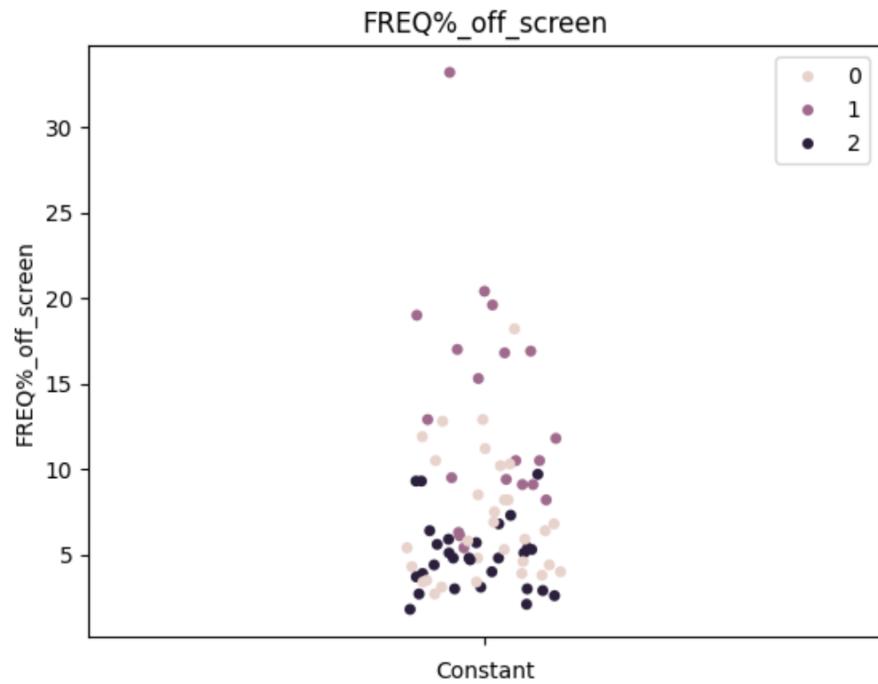


The clusters are visible and clearly delineated under PCA, suggesting KMeans is operating well.

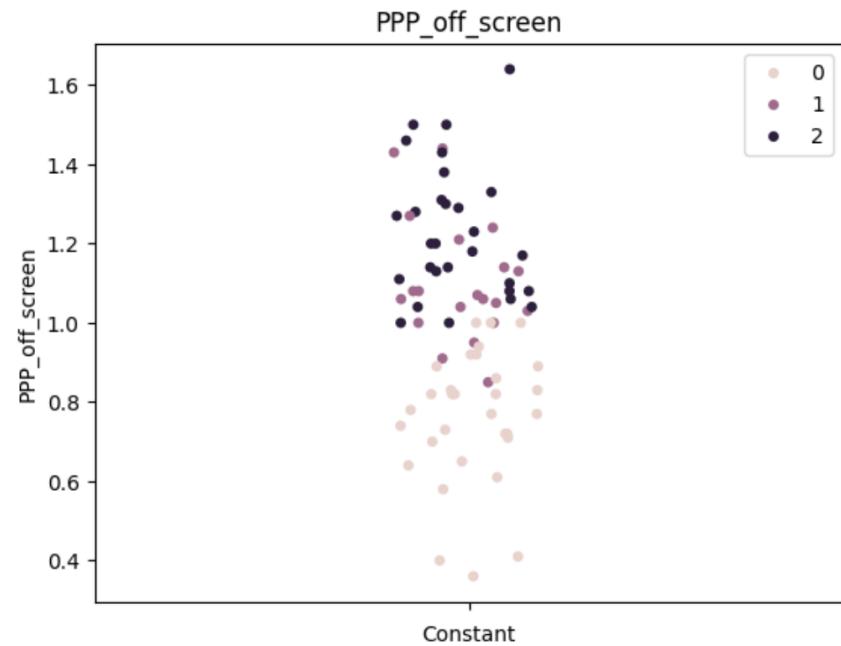
3.3.3 Variance of clusters with POSS



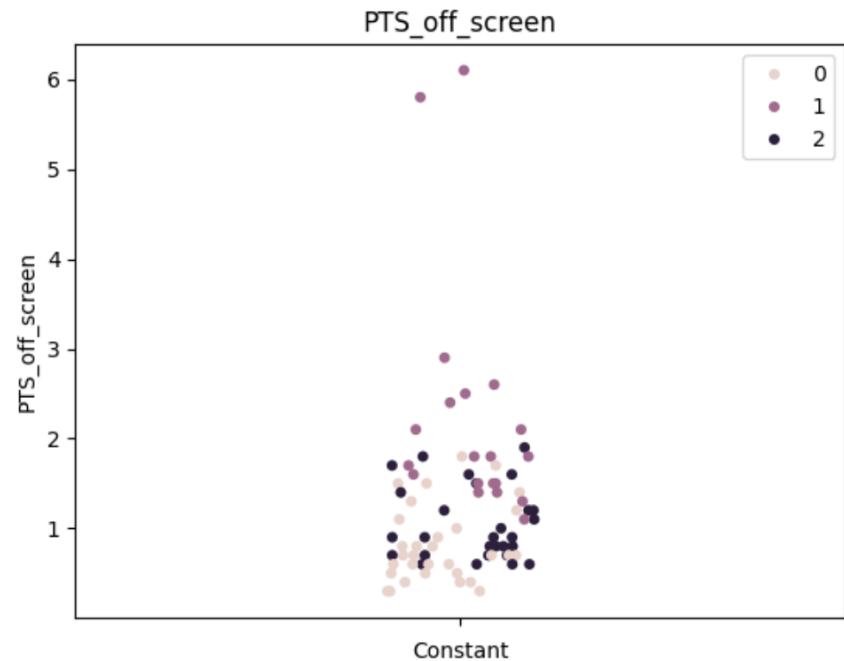
3.3.4 Variance of clusters with FREQ%



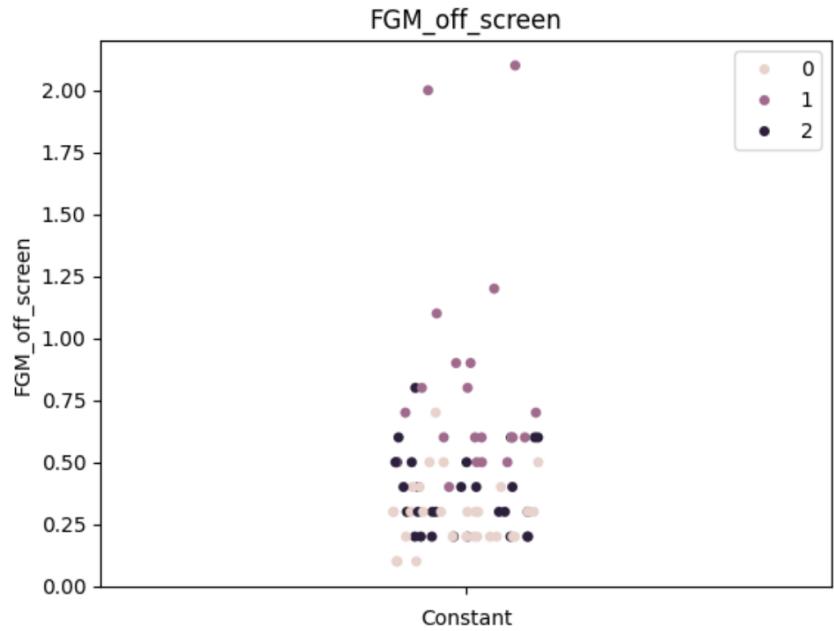
3.3.5 Variance of clusters with PPP



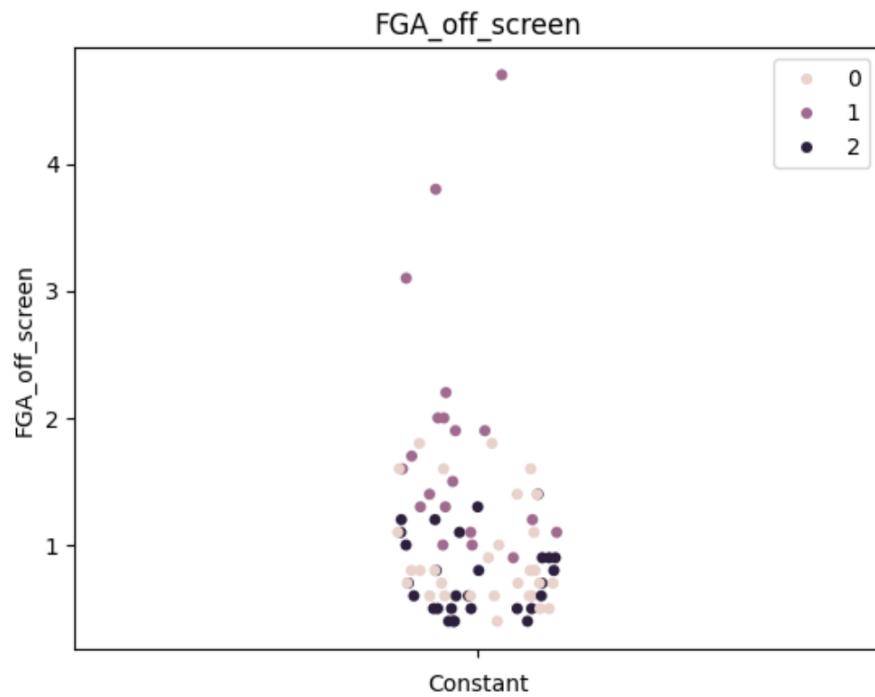
3.3.6 Variance of clusters with PTS



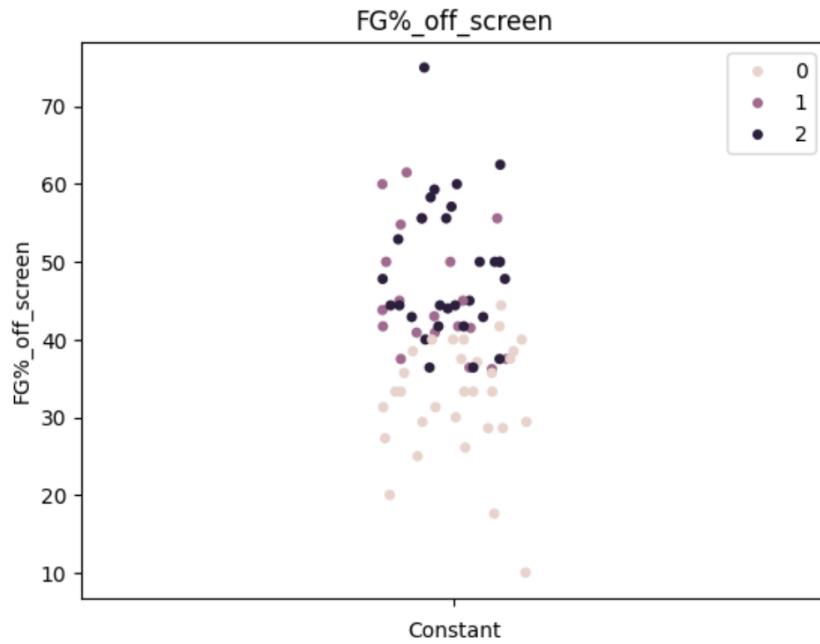
3.3.7 Variance of clusters with FGM



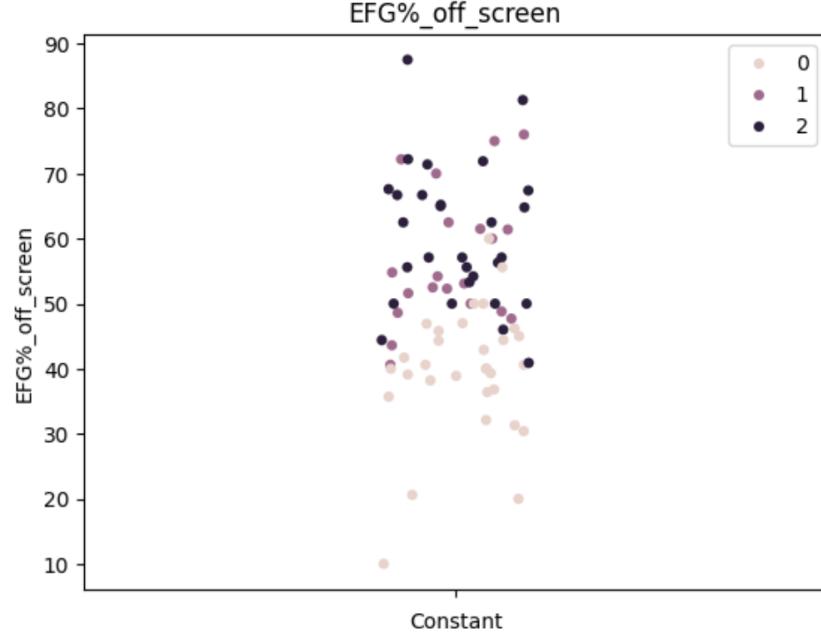
3.3.8 Variance of clusters with FGA



3.3.9 Variance of clusters with FG%



3.3.10 Variance of clusters with EFG%



3.4 Pick and roll ball handler

3.4.1 Describing clusters

The clusters range over the following values:

Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	.5-6.4	2.4-42.2	.27-.91	.2-4.5	0.0-1.7	.3-4.7	0.0-50.0	0.0-58.3
1	3-13.5	21.7-57.5	.75-1.24	2.7-13.5	1.1-4.5	2.5-9.8	31.3-65.9	34.4-66.5
2	.5-4.4	3.2-43.9	.78-1.60	.5-4.9	.2-1.8	.4-3.4	36.4-66.7	38.5-77.8

Table 1: Pick and roll ball handler clusters

Cluster 0 includes:

- Shaedon Sharpe
- Jayson Tatum
- Paolo Banchero
- Anthony Davis
- Aaron Nesmith

- Donte DiVincenzo
- Kira Lewis Jr.

Cluster 1 includes:

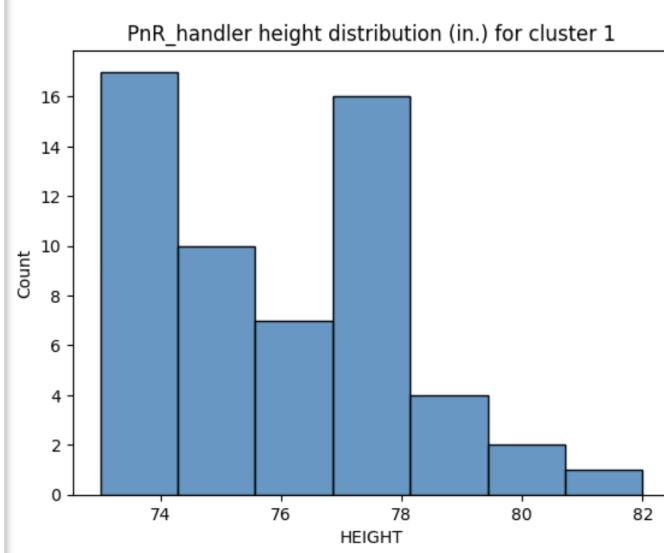
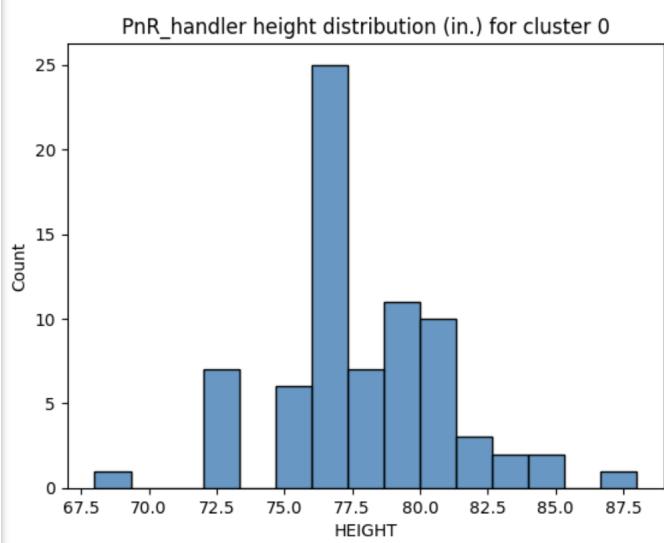
- Trae Young
- Luka Doncic
- Damian Lillard
- Cade Cunningham
- Tyrese Haliburton
- Jalen Brunson
- LaMelo Ball
- De'Aaron Fox

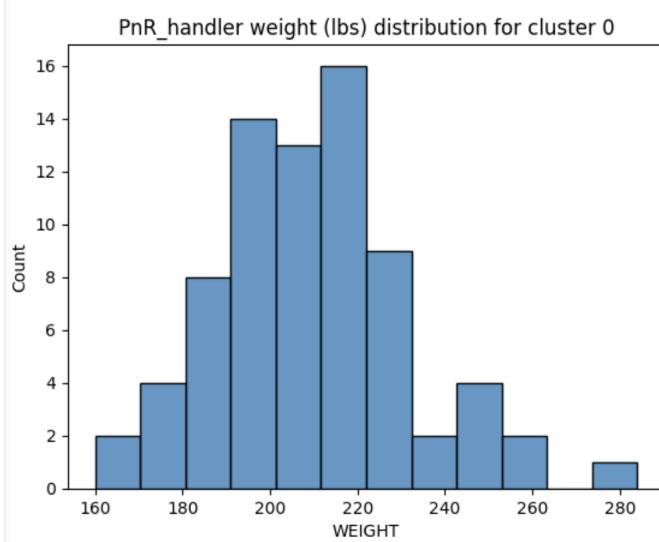
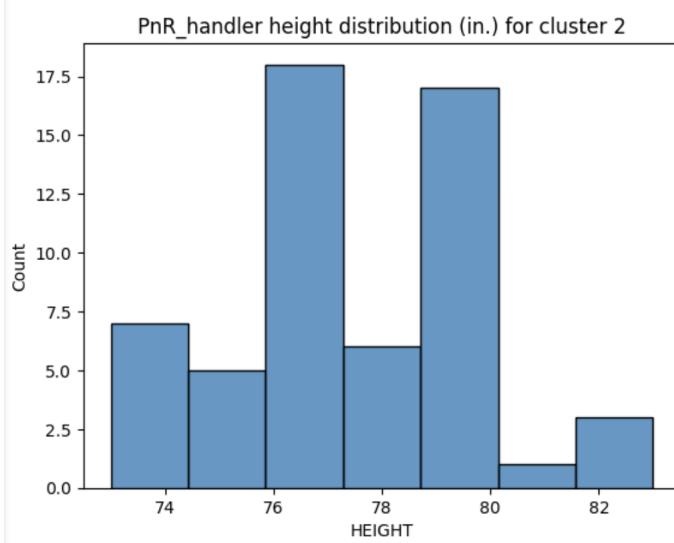
Cluster 2 includes:

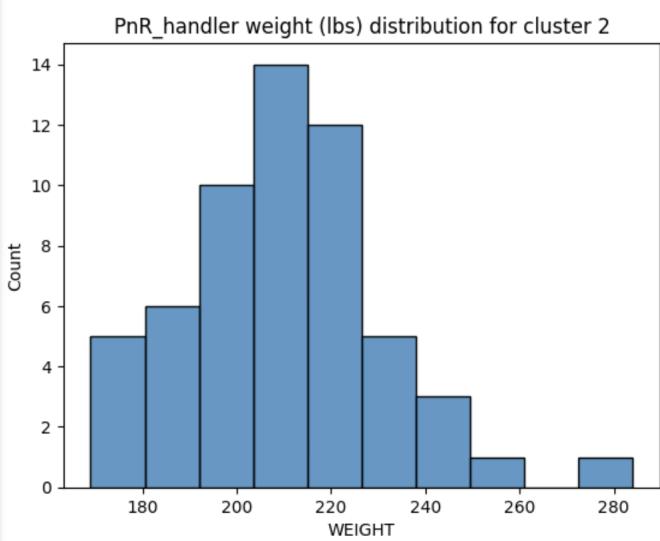
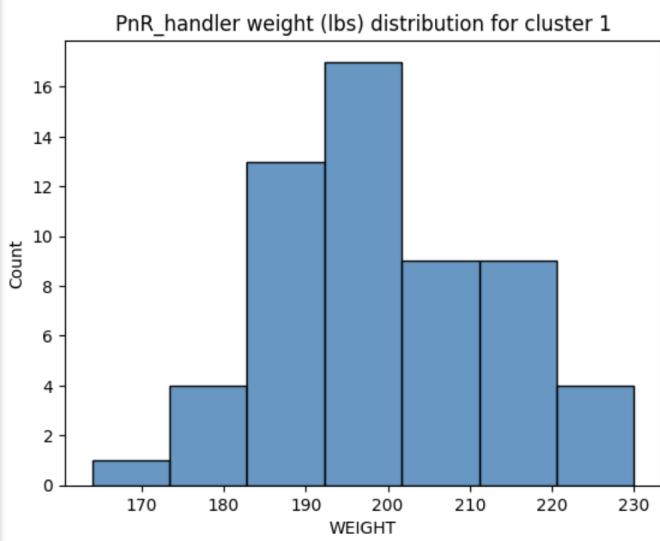
- Paul George
- Kevin Durant
- LeBron James
- Grayson Allen
- Lonnie Walker IV
- Jaden Ivey
- Benedict Mathurin
- Scottie Barnes

(these are not exhaustive lists)

Cluster 1 players are the shortest of the bunch, while cluster 0 players are the tallest.







Cluster 0 is composed of tall, skilled, athletic finishers, who can finish as the ball handler on their own, even when trapped in the pick and roll.

Cluster 1 generally contains agile, playmaking guards. These players can navigate through the defense in pick and roll scenarios, but can also punish defenses for going under screens through deadly shooting. They are also excellent at making tough passes through any type of coverage, allowing them to get the

ball to the roller or to a newly open shooter on the perimeter. As a result of all these faculties, they are extremely efficient as ball handlers in a pick and roll (.75-1.24 PPP).

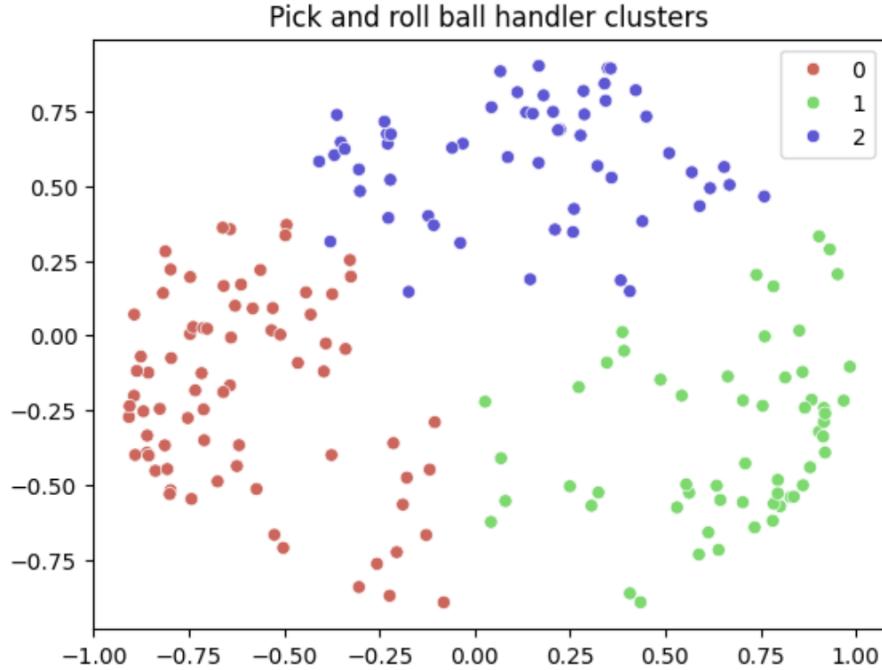
Finally, cluster 2 includes large forwards who use the pick and roll to create size mismatches, overpowering their defender to get to the rim. Because of their size advantage, they are extremely efficient as the ball handler in pick and rolls, generating between .78 and 1.60 PPP. However, they don't initiate pick and rolls as ball handlers as frequently as the guards in cluster 1.

The clusters are summarized below:

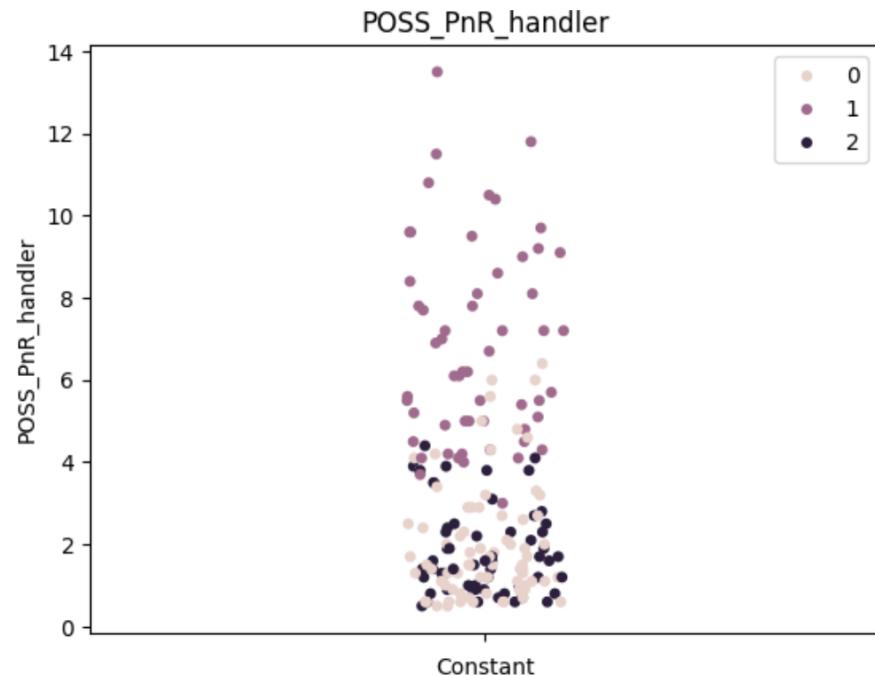
Cluster	Name	Description	Height (in.)	Weight (lbs)
0	tall, athletic finishers	finish through any coverage	72-87	160-280
1	playmaking guards	weave to rim, shooters, pass well	74-82	170-230
2	strong forwards	get mismatches, overpower them	74-82	180-280

Table 1: Descriptions of pick and roll ball handler clusters

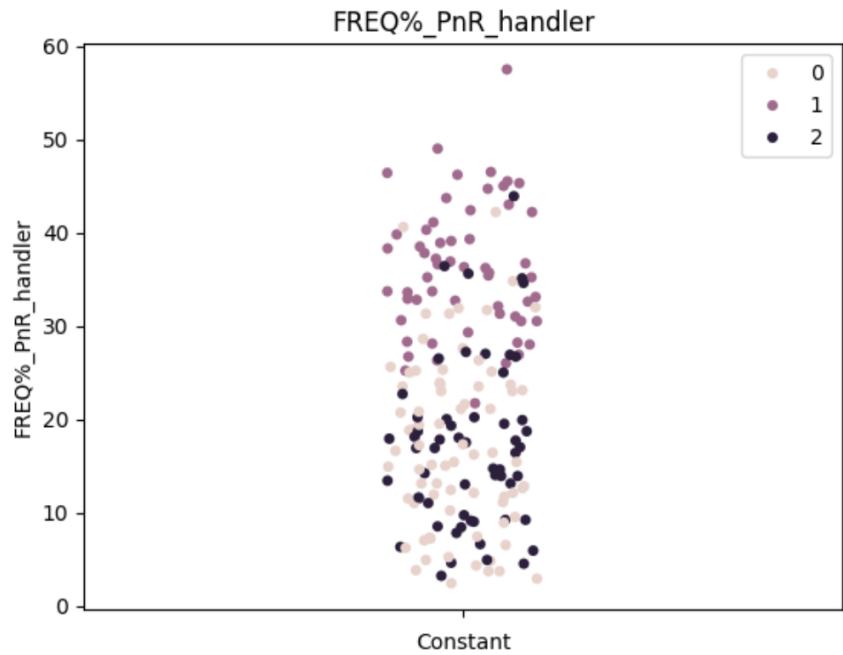
3.4.2 Clusters under PCA



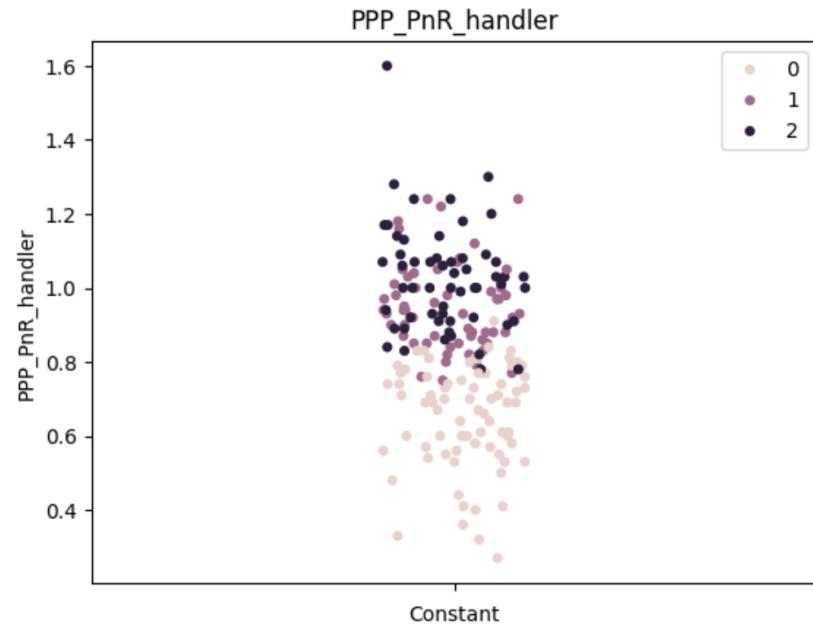
3.4.3 Variance of clusters with POSS



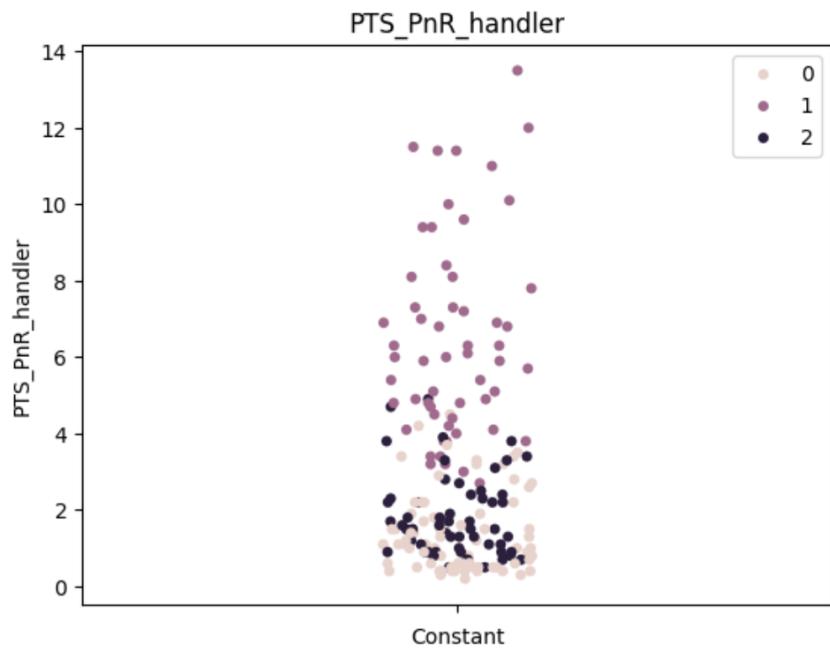
3.4.4 Variance of clusters with FREQ%



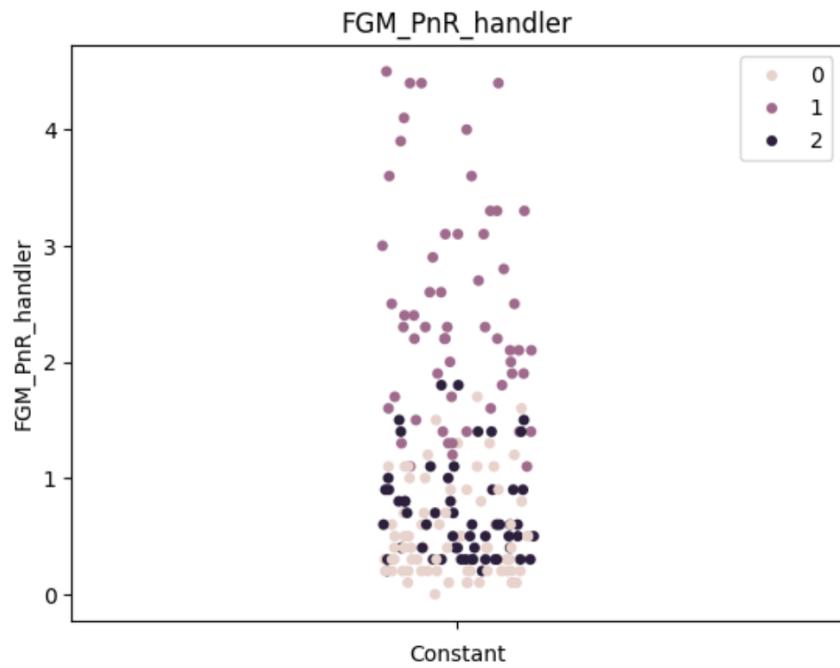
3.4.5 Variance of clusters with PPP



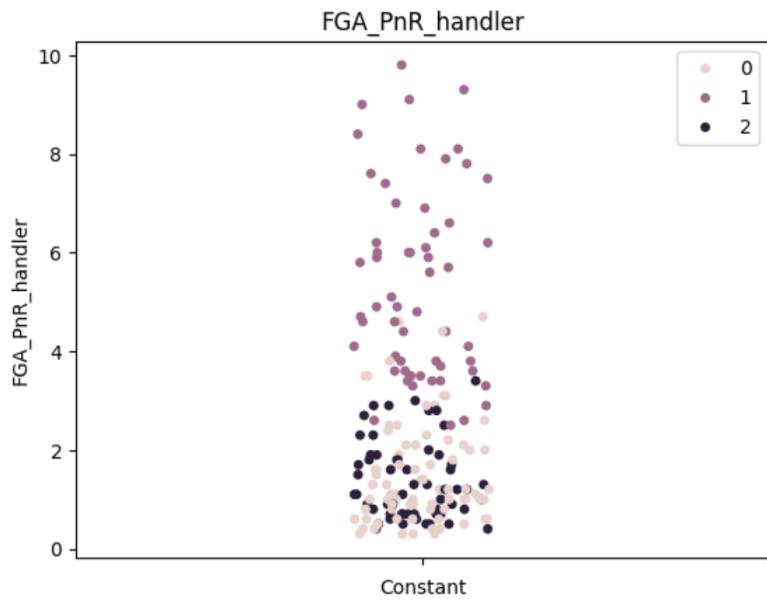
3.4.6 Variance of clusters with PTS



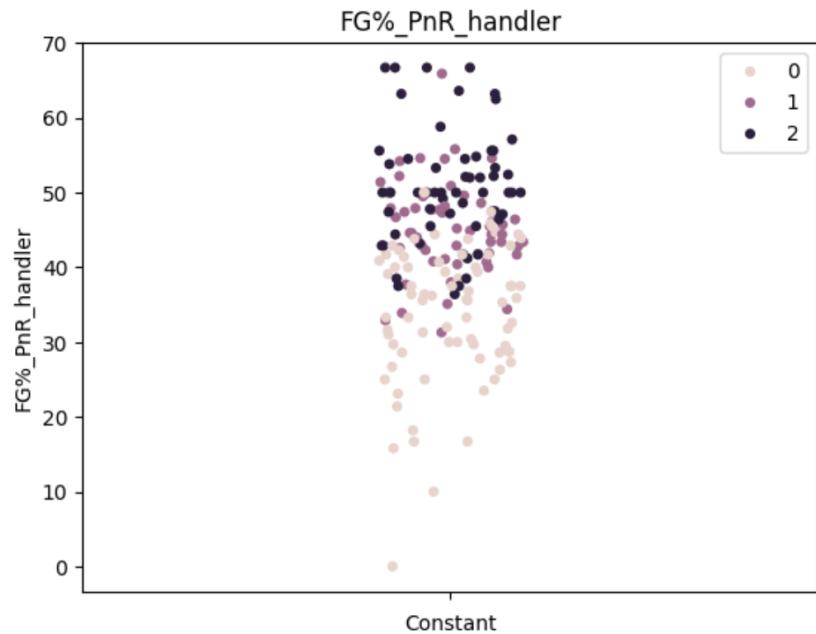
3.4.7 Variance of clusters with FGM



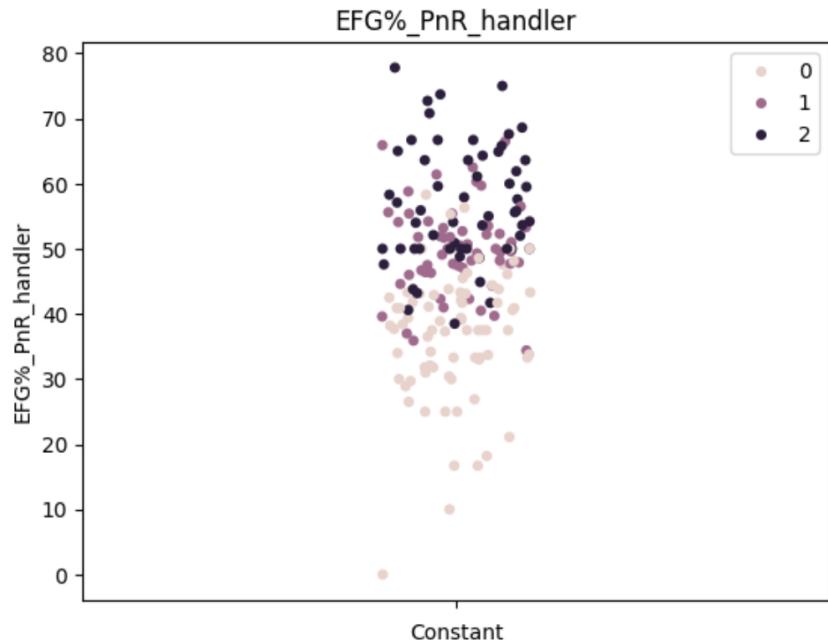
3.4.8 Variance of clusters with FGA



3.4.9 Variance of clusters with FG%



3.4.10 Variance of clusters with EFG%



3.5 Pick and roll roller

3.5.1 Describing clusters

The clusters range over the following values:

Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	.6-6.9	2-36.8	1.94-2	.8-7.4	.3-3.1	.4-5.7	45.7-100	50-100
1	.5-3.9	3.1-33	.43-1.35	.3-3.4	.1-1.4	.3-3.4	25-70	25-70

Table 1: Pick and roll roller cluster values

Cluster 0 includes:

- Nikola Vucevic
- Alperen Sengun
- Myles Turner
- Nikola Jokic
- Chet Holmgren

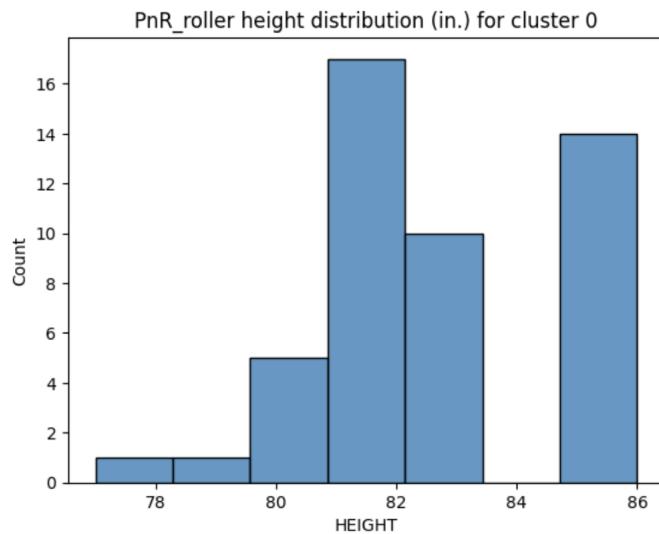
- Deandre Ayton
- Jarrett Allen

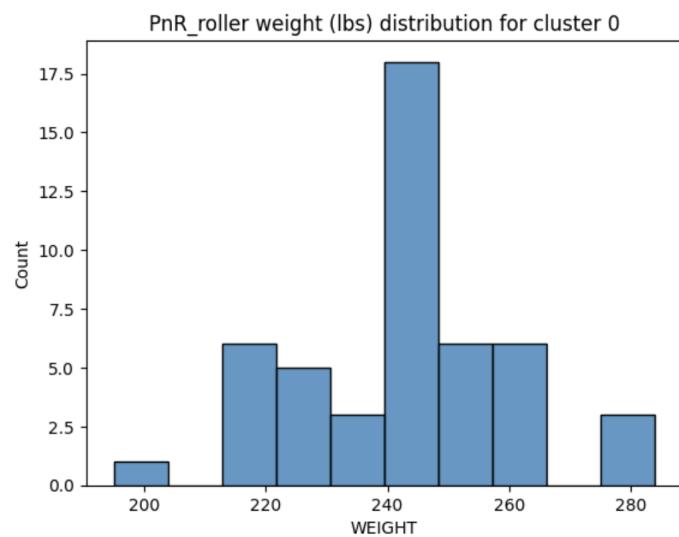
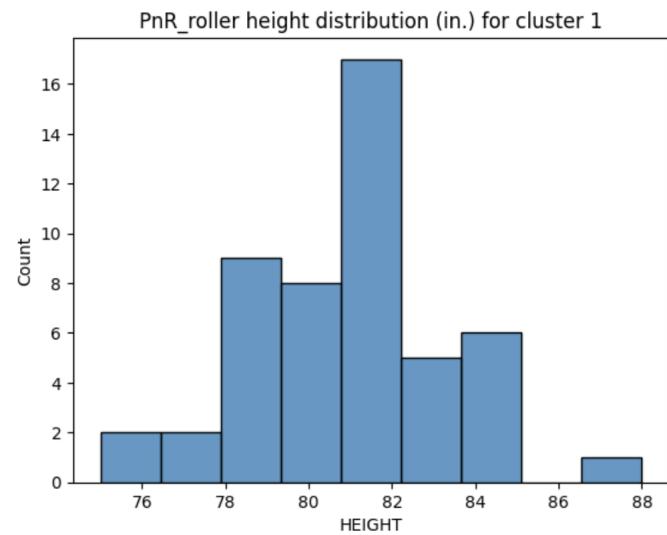
Cluster 1 includes:

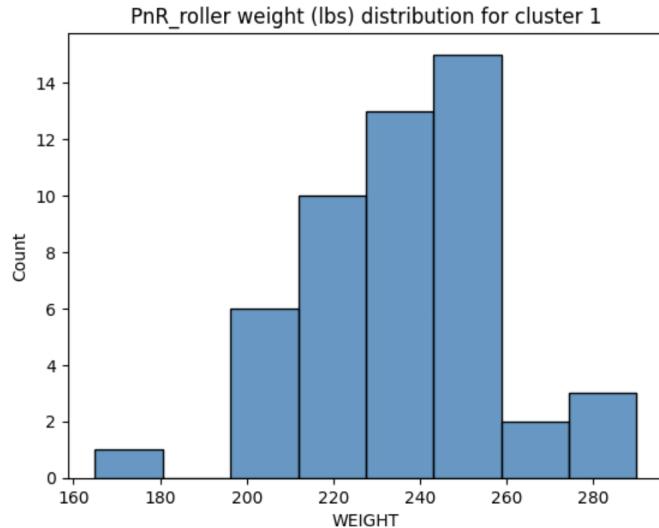
- Zach Collins
- Dario Saric
- Evan Mobley
- Lauri Markkanen
- Victor Wembanyama
- Kyle Kuzma
- Mitchell Robinson
- Kevin Love

(these are not exhaustive lists)

The heights and weights of the two clusters are similar. Cluster 0 consists of players that are 78-86 inches tall, while cluster 1 players are typically between 76 and 88 inches tall. Both cluster 0 and cluster 1 players weigh between 200 and 280 pounds.







The players in cluster 0 are strong finishers who can establish deep position in the paint. Once established, they can call upon wide arsenals of post moves to get high quality shots up. Because they are strong enough to get deep into the paint when rolling, they generally are more efficient than cluster 1 players, generating 1.94-2 points per possession to .43-1.35 for cluster 1.

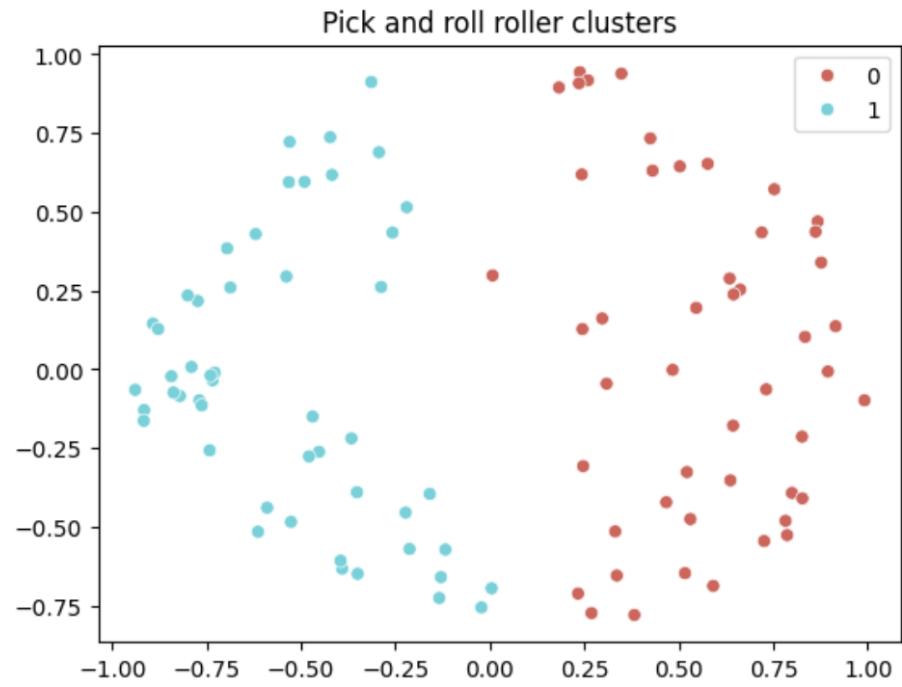
Cluster 1 consists of stretch bigs who are efficient shooters. As rollers, they stretch the floor, forcing defenders into tough decisions when hedging or defending the roll.

The clusters are summarized in the following table:

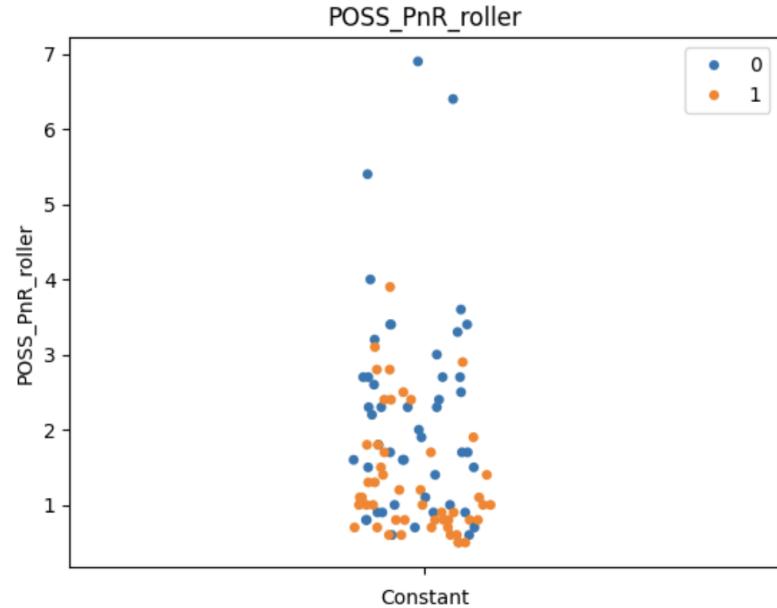
Cluster	Name	Description	Height	Weight
0	strong post finishers	get deep in paint, efficient production	78-86	200-280
1	stretch bigs	shoot, prevent hedging, moderate efficiency	76-88	200-280

Table 1: Descriptions of pick and roll roller clusters

3.5.2 Clusters under PCA



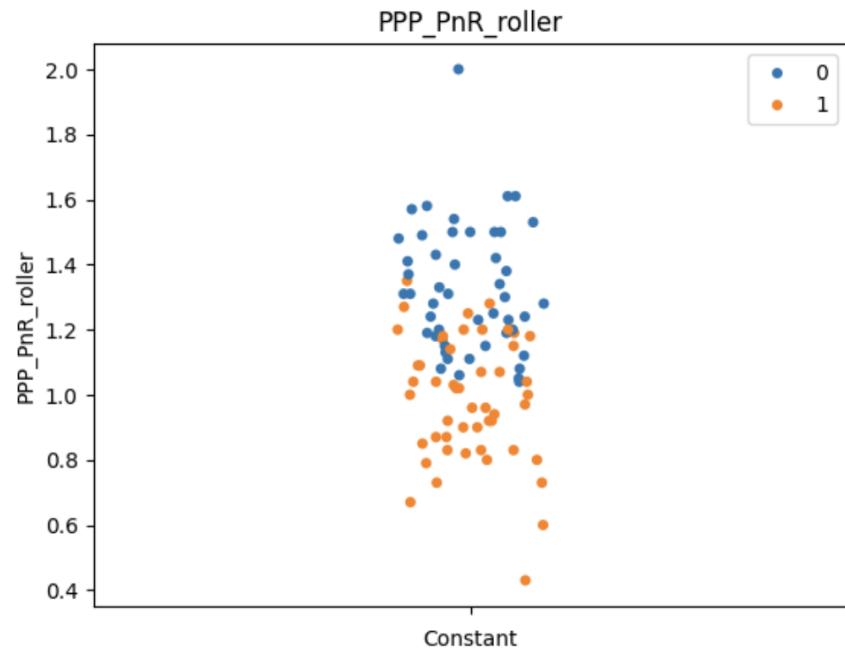
3.5.3 Variance of clusters with POSS



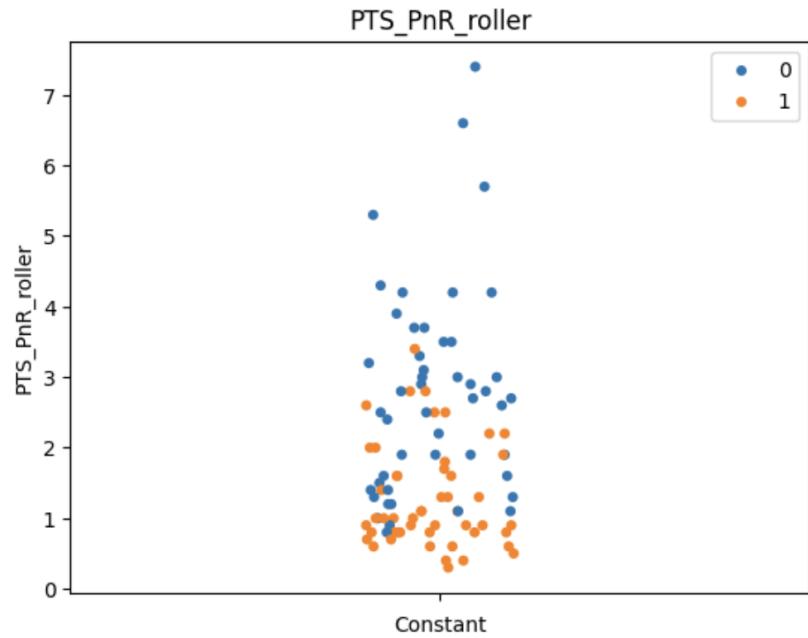
3.5.4 Variance of clusters with FREQ%



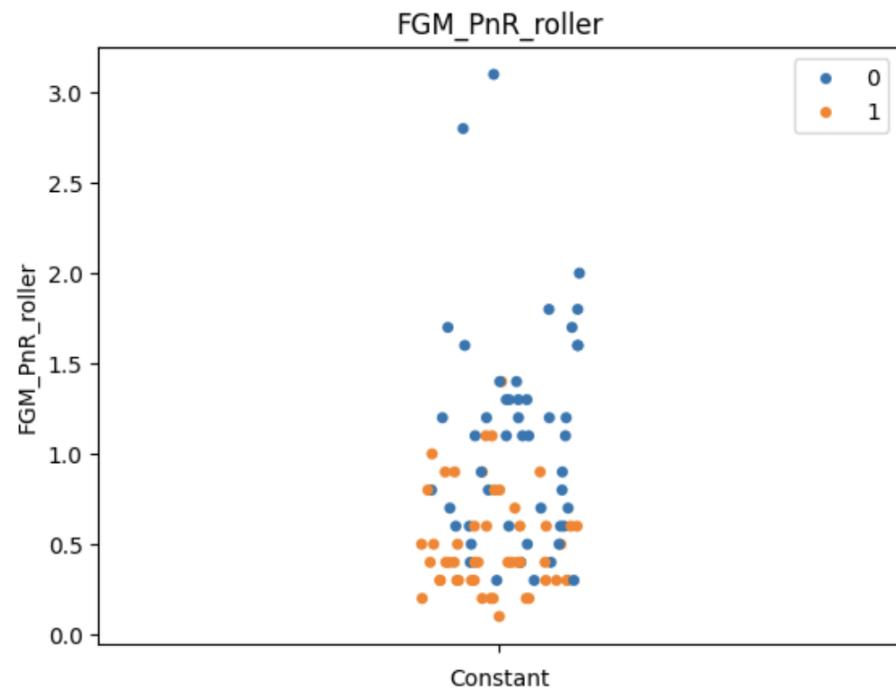
3.5.5 Variance of clusters with PPP



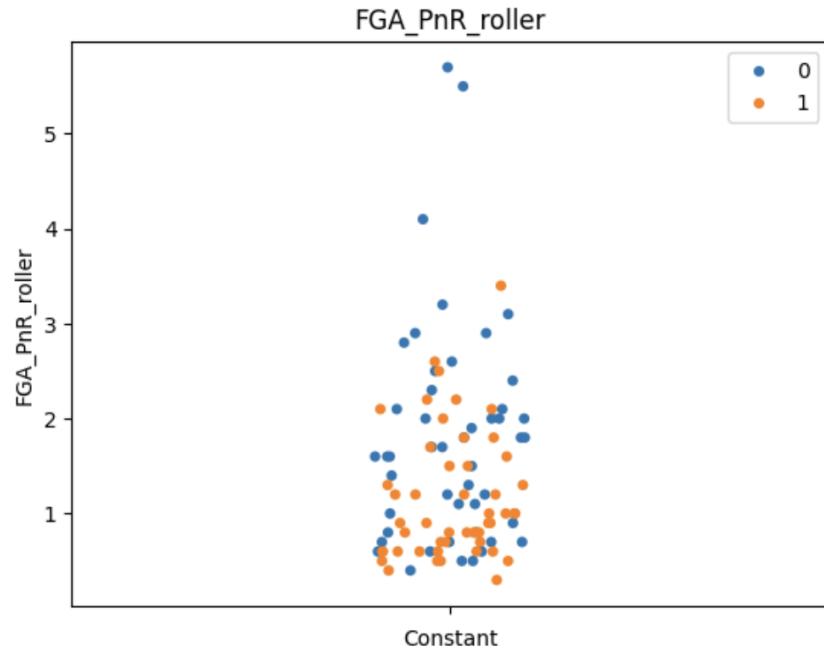
3.5.6 Variance of clusters with PTS



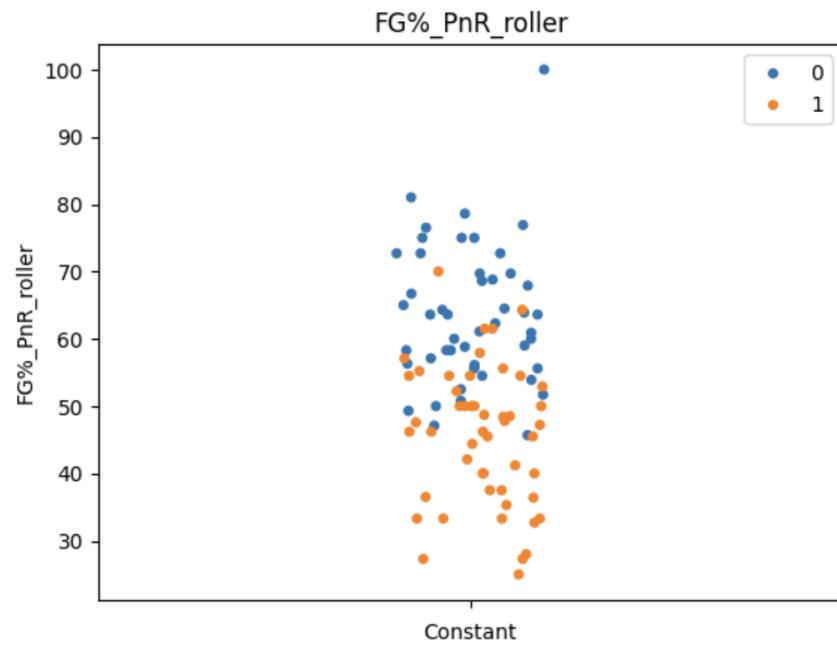
3.5.7 Variance of clusters with FGM



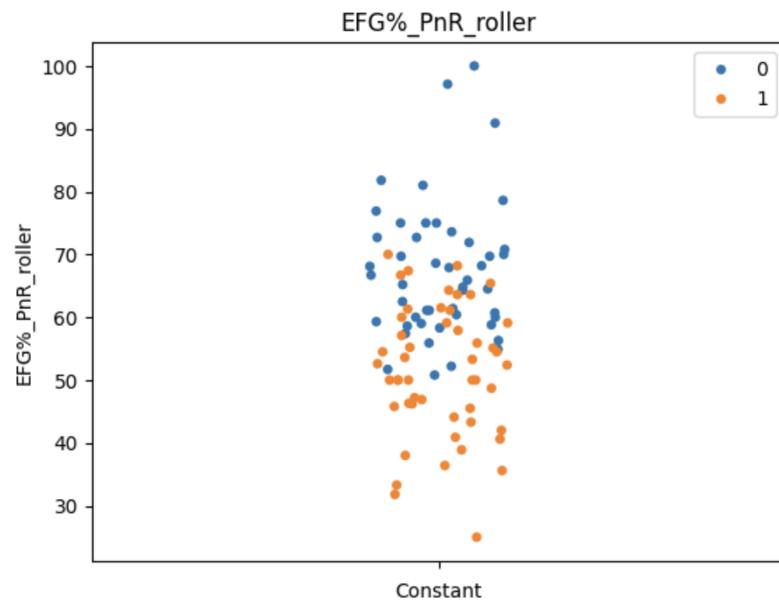
3.5.8 Variance of clusters with FGA



3.5.9 Variance of clusters with FG%



3.5.10 Variance of clusters with EFG%



3.6 Post up

3.6.1 Describing clusters

The clusters occupy the following ranges:

Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	.6-2.9	3.9-13.9	1.03-1.56	.6-3.7	.1-1.3	.2-2.1	42.9-79.2	42.9-79.2
1	.5-2.6	2.8-15.8	.4-1	.2-2.3	0-7	.3-1.9	14.3-57.1	14.3-57.1
2	2.3-6.6	10.8-25.5	.73-1.21	2.1-7.2	.8-2.8	1.7-5.1	39.5-64.6	39.5-64.6

Table 1: Post up cluster ranges

Cluster 0 includes:

- Jayson Tatum
- DeMar DeRozan
- Tobias Harris
- LeBron James
- Jonathan Kuminga
- Kyle Kuzma
- Lauri Markkanen
- Myles Turner
- Kristaps Porzingis
- Paul George

Cluster 1 includes:

- Luka Doncic
- Jimmy Butler
- Kawhi Leonard
- Scottie Barnes
- Jaylen Brown
- Anthony Edwards
- Jamal Murray
- Aaron Gordon

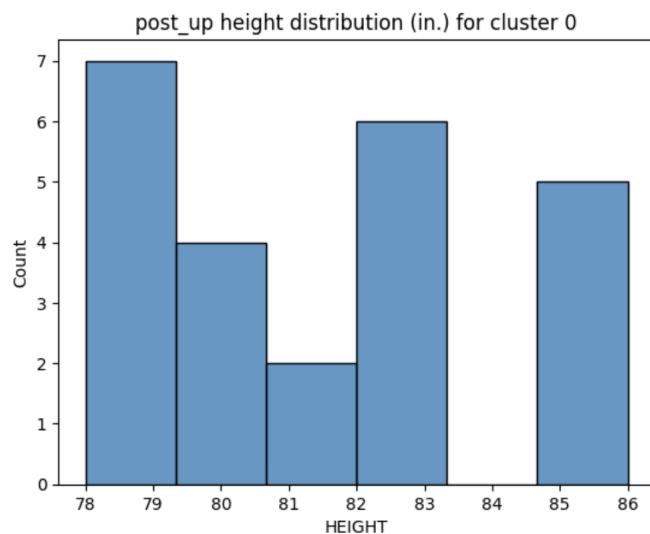
- Christian Wood

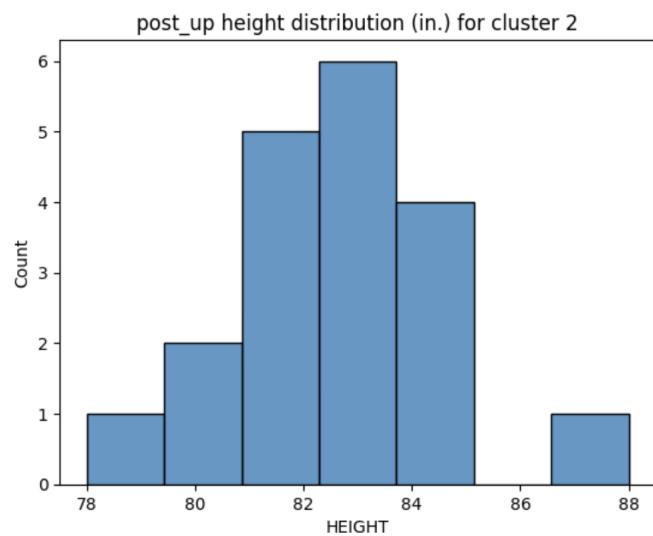
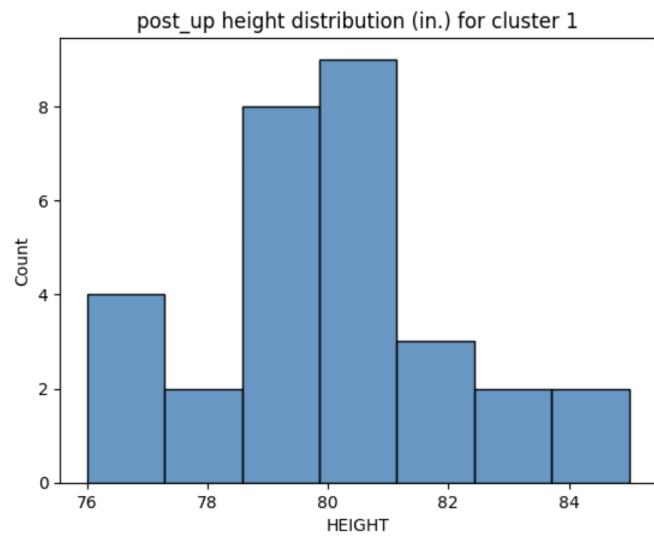
Cluster 2 includes:

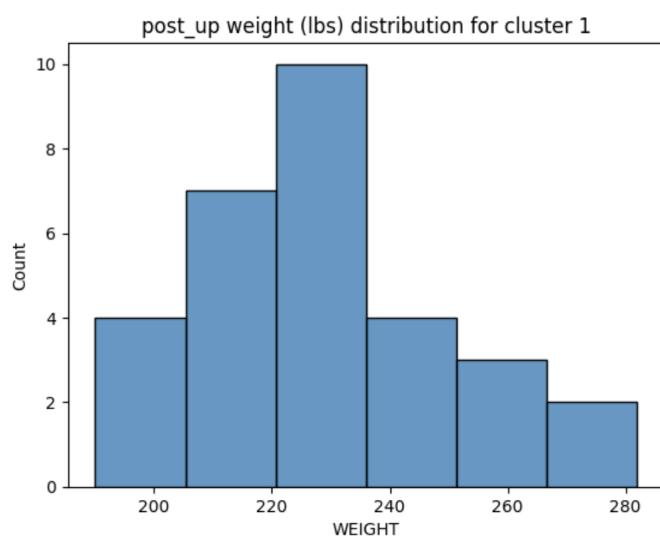
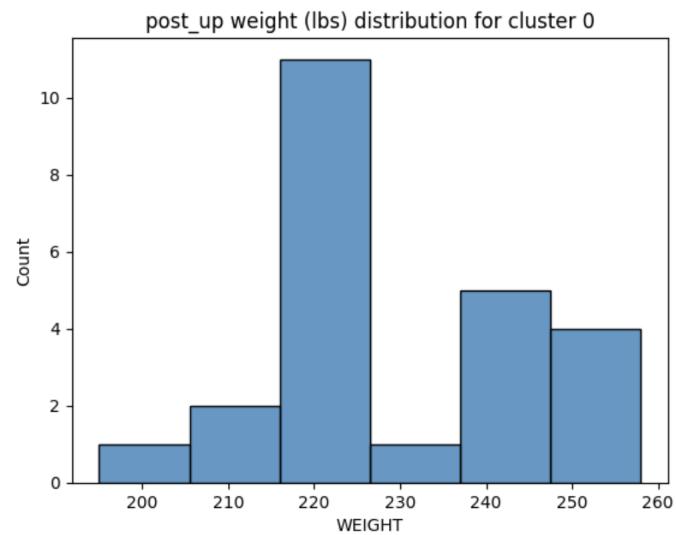
- Nikola Jokic
- Joel Embiid
- Anthony Davis
- Jaren Jackson Jr.
- Karl-Anthony Towns
- Bam Adebayo
- Pascal Siakam
- Deandre Ayton
- Giannis Antetokounmpo

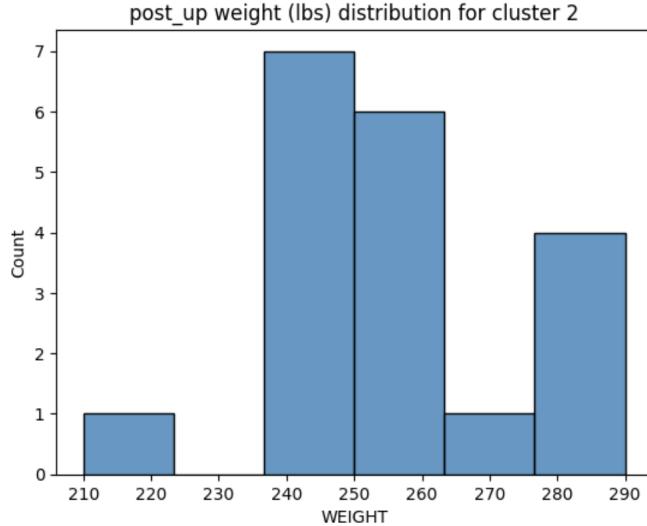
(these are not exhaustive lists)

Cluster 1 seems to be composed mostly of shorter guards (76-84 inches). Cluster 0 includes taller forwards (78-86 inches). Finally, cluster 2 includes bigs (78-88 inches).









Cluster 0 players are the most efficient in post up scenarios ($1.03=1.56$ PPP). Many of these players are excellent at turnaround jumpers and fadeaways, allowing them to gain space and get up quality shots against any defender. These players also have excellent footwork (spin moves, drop steps, etc.), allowing them to evade their defender in post-up situations.

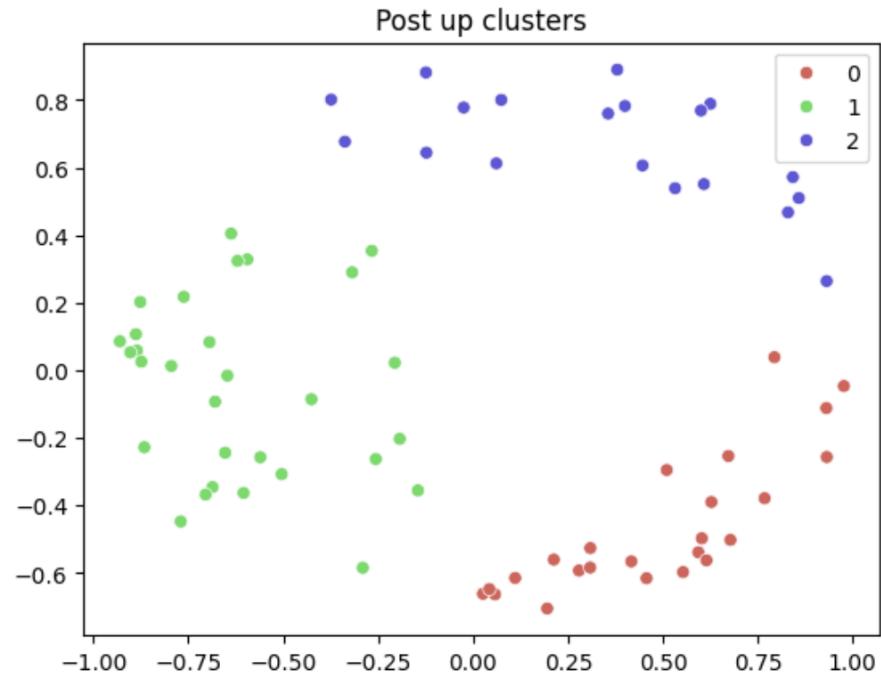
Cluster 1 players are the least efficient of the group. They attempt fadeaways and turnaround jumpers, but aren't as efficient on such attempts as cluster 0 players. The strength of these players is in their power: Jimmy Butler and Kawhi Leonard, for instance, can absorb contact in the post and finish on any defender.

Cluster 2 consists of shooting bigs. These players use a variety of post moves with their back to the basket, including hook shots and drop steps. They are tall enough to rise over any defender and get off high quality shots. Cluster 2 players rely on post-ups far more than players in the other two clusters, at efficiencies better than cluster 1 but slightly worse than cluster 0.

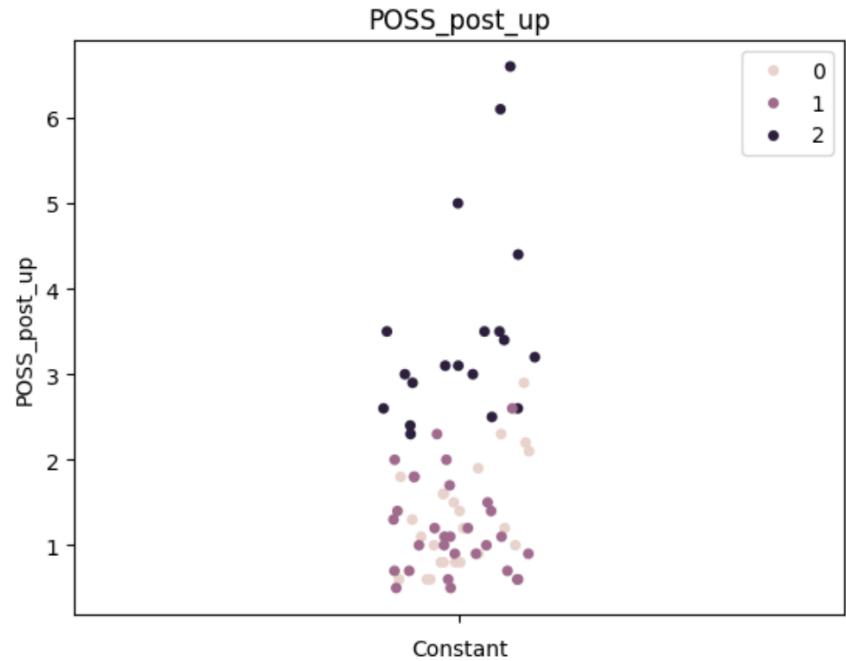
Cluster	Name	Description	Height (in.)	Weight (lbs)
0	versatile forwards	jumper arsenal, great footwork, efficient	76-84	200-260
1	strong guards	strong finishers, inefficient	78-86	200-280
2	post bigs	post moves, rise over defender	78-88	210-290

Table 1: Descriptions of post up clusters

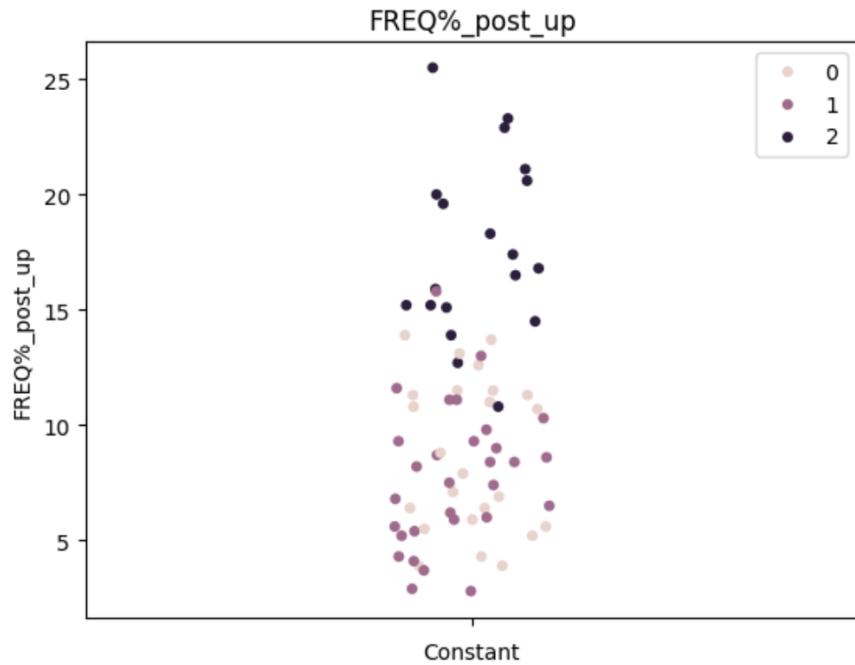
3.6.2 Clusters under PCA



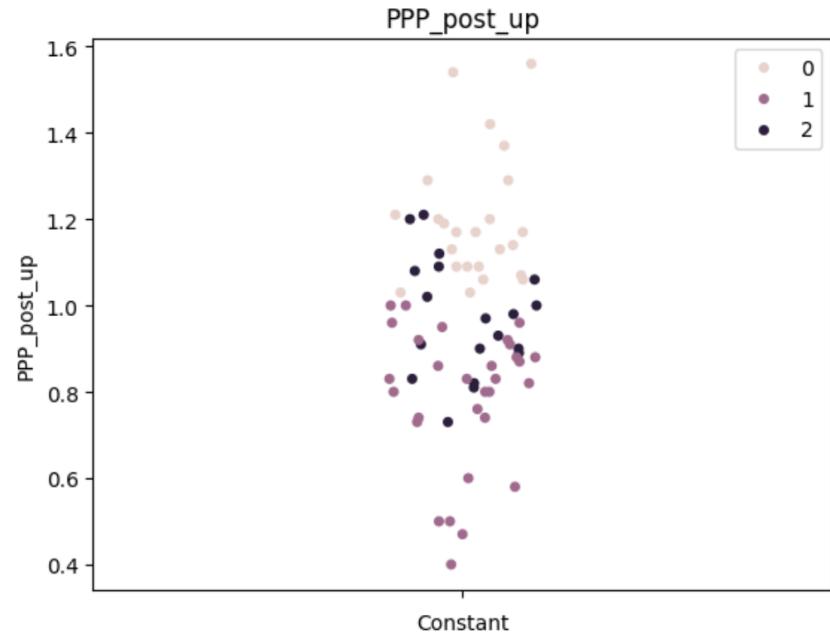
3.6.3 Variance of clusters with POSS



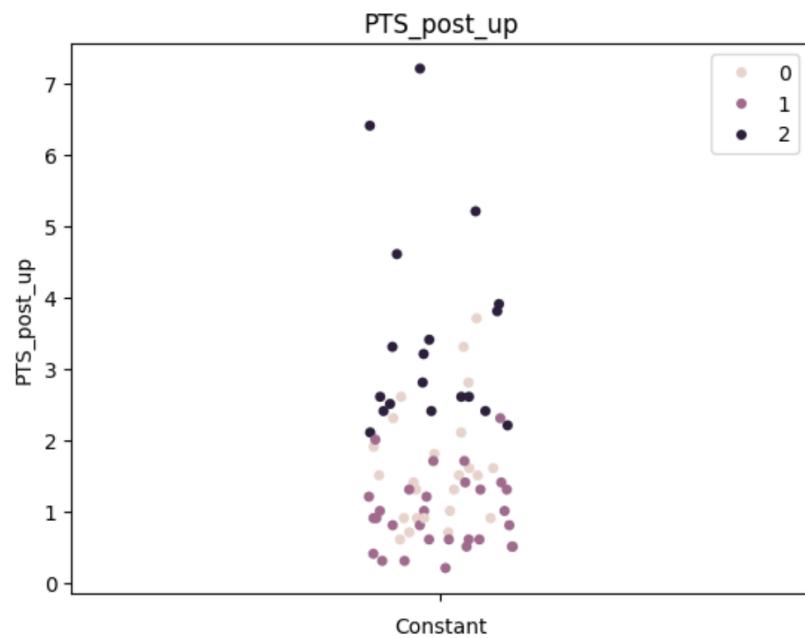
3.6.4 Variance of clusters with FREQ%



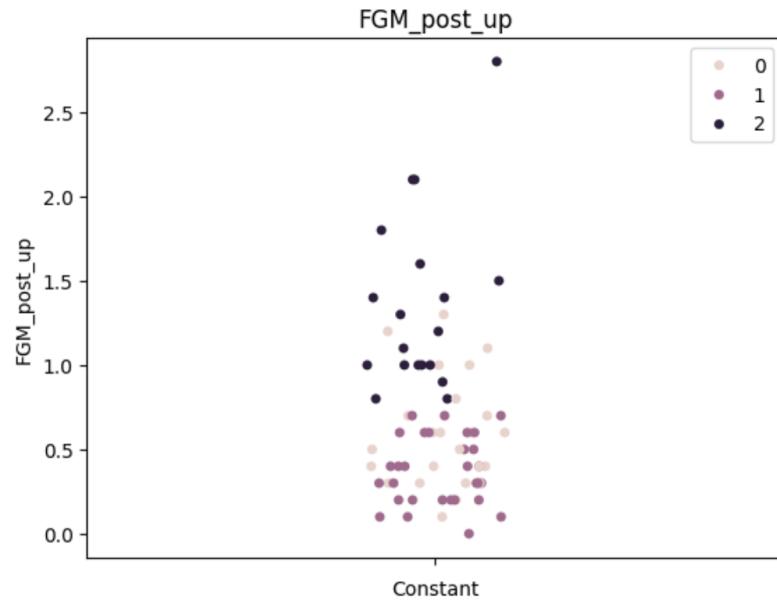
3.6.5 Variance of clusters with PPP



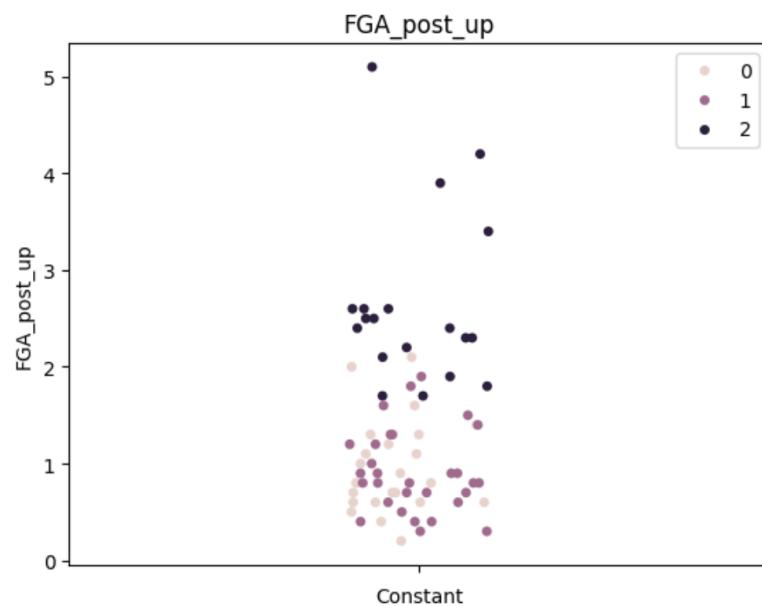
3.6.6 Variance of clusters with PTS



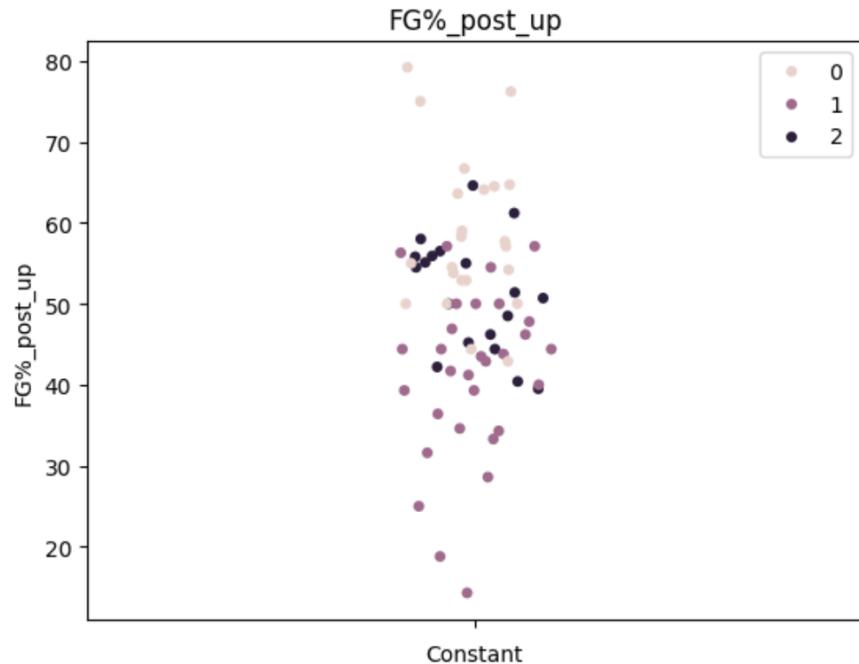
3.6.7 Variance of clusters with FGM



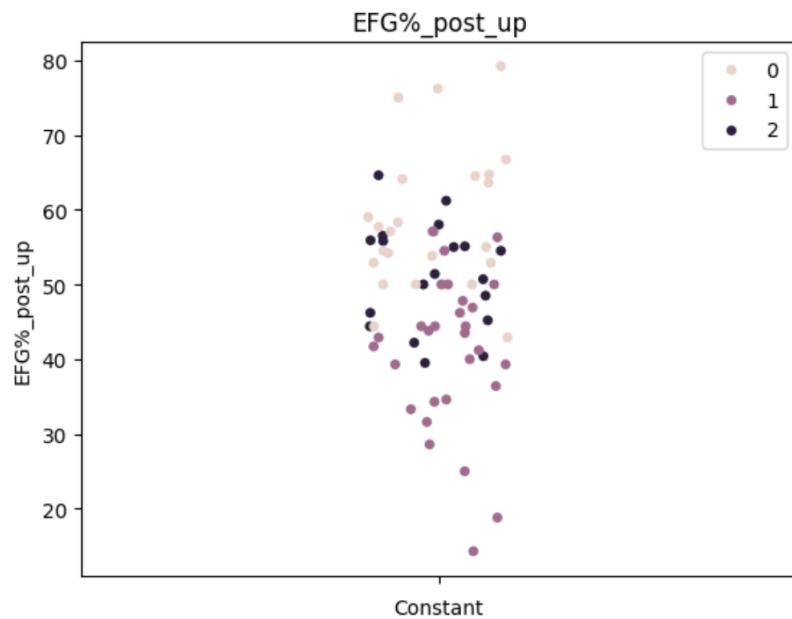
3.6.8 Variance of clusters with FGA



3.6.9 Variance of clusters with FG%



3.6.10 Variance of clusters with EFG%



3.7 Spot up

3.7.1 Describing clusters

The clusters vary over the following ranges:

Cluster	POSS	FREQ%	PPP	PTS	FGM	FGA	FG%	EFG%
0	1.1-6.5	7-56.2	.88-1.73	1.5-7.7	.5-2.7	1-6	33.8-69.2	45.1-92.3
1	.7-5.5	3.9-70.6	.37-1.13	.4-3.6	.1-1.3	.6-6	14.8-53.3	18.9-64.3

Table 1: Spot up cluster ranges

Cluster 0 includes:

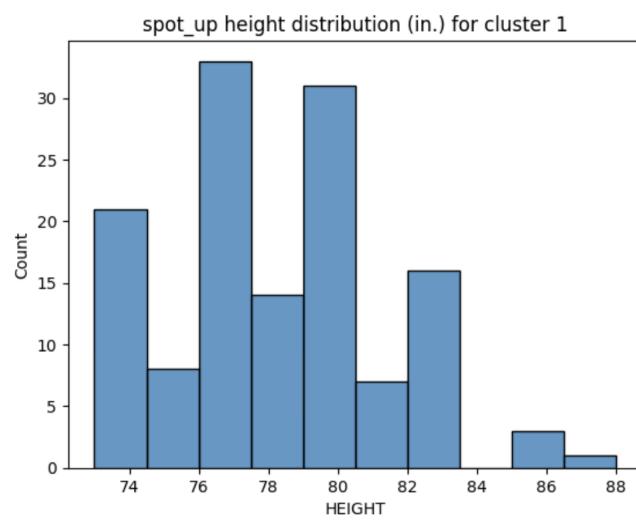
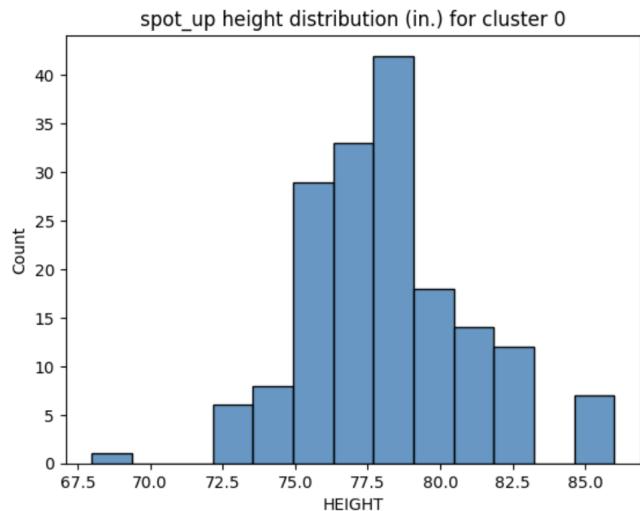
- Coby White
- Michael Porter Jr.
- Jerami Grant
- Eric Gordon
- Jalen Brunson
- Jalen Smith
- Bradley Beal

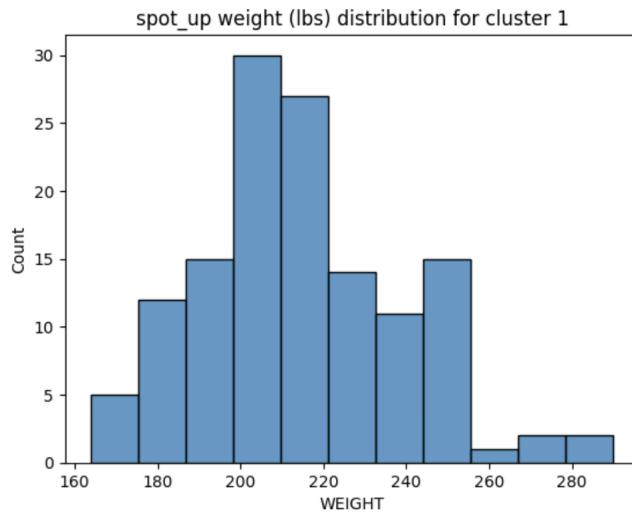
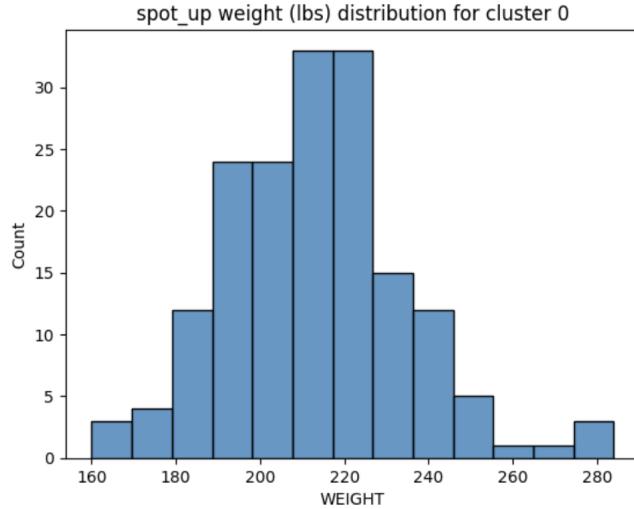
Cluster 1 includes:

- Pascal Siakam
- Jaren Jackson Jr.
- Patrick Williams
- Luguentz Dort
- Jalen Duren
- Walker Kessler
- Kira Lewis Jr.

(these are not exhaustive lists)

Cluster 1 (74-88 inches, 170-280 lbs) is generally taller than cluster 0 (72-83 inches, 160-280 pounds).





Cluster 0 is generally composed of more efficient shooters, who are also excellent at moving off-ball. Players like Jalen Brunson can navigate through screens to find open spaces with ease. These players also tend to be effective both in the mid-range and from the three point line. Cluster 0 players are generally more efficient from the field than cluster 1 players, although spot-up shots make up a smaller proportion of their aggregate attempts.

Cluster 1 consists of large forwards and centers who primarily take shots from the high to mid post. They generally are less effective spot up shooters than cluster 0 players, preferring to leverage their size by driving after catching the

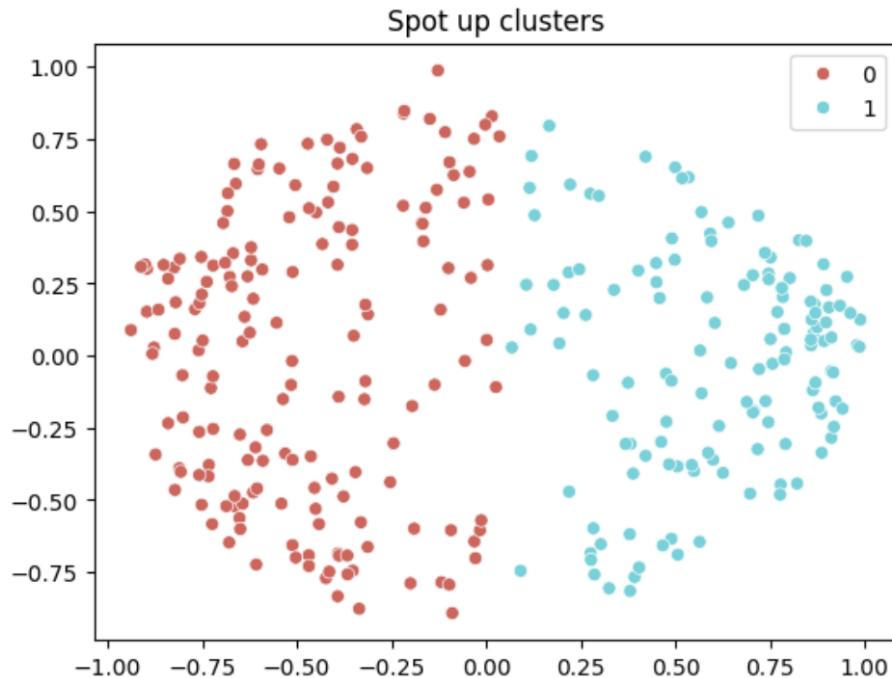
ball.

The qualities of each cluster are summarized below:

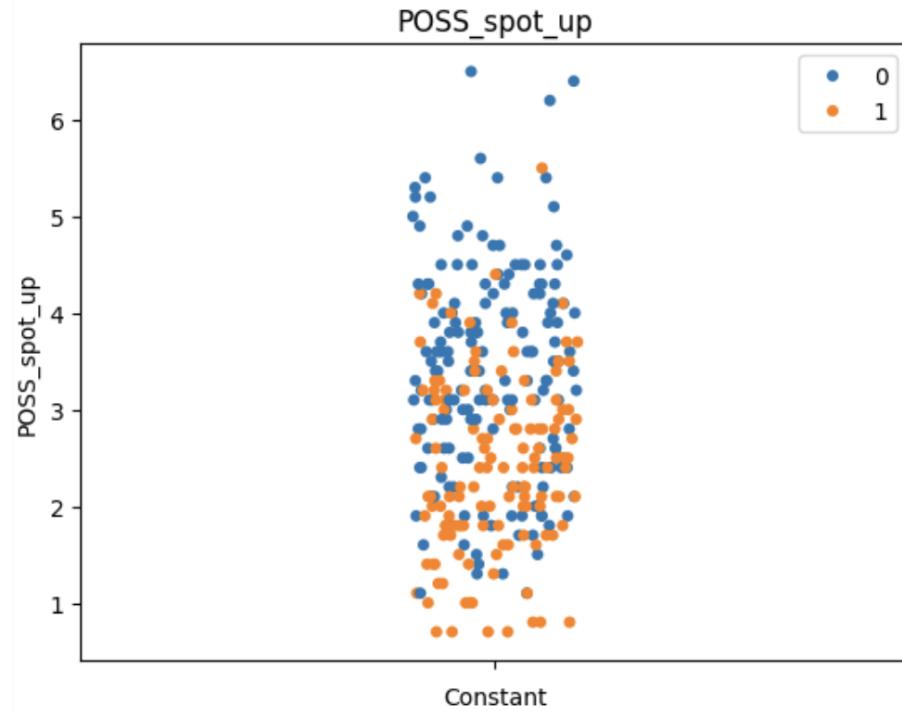
Cluster	Name	Description	Height (in.)	Weight (lbs)
0	agile perimeter shooters	move off-ball, great from 3, midrange	74-88	170-280
1	midrange forwards	post shooters, catch and drive	72-83	160-280

Table 1: Descriptions of spot up clusters

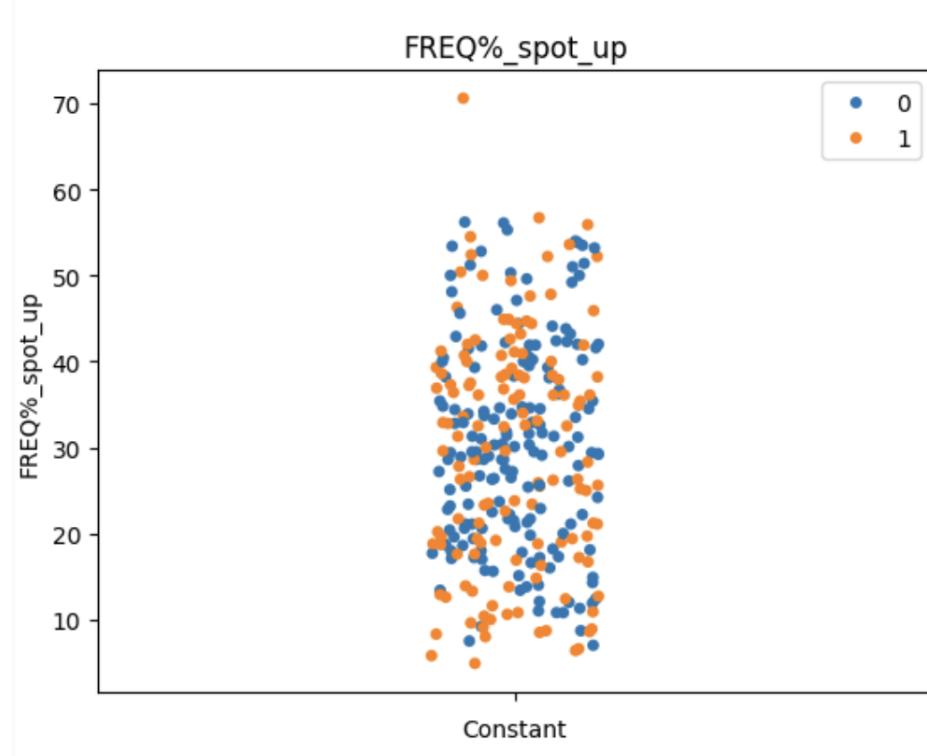
3.7.2 Clusters under PCA



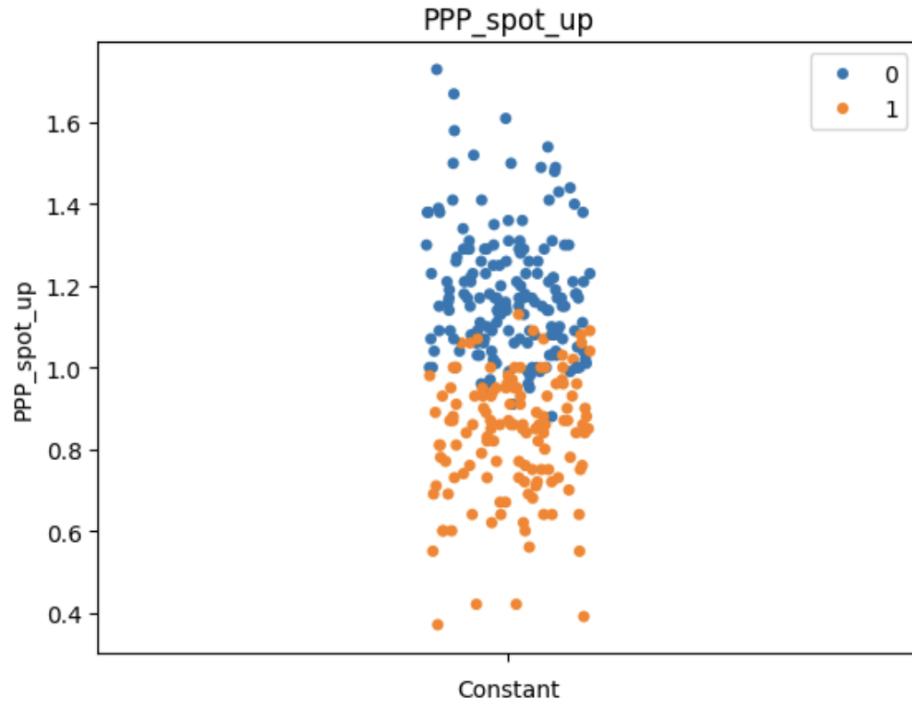
3.7.3 Variance of clusters with POSS



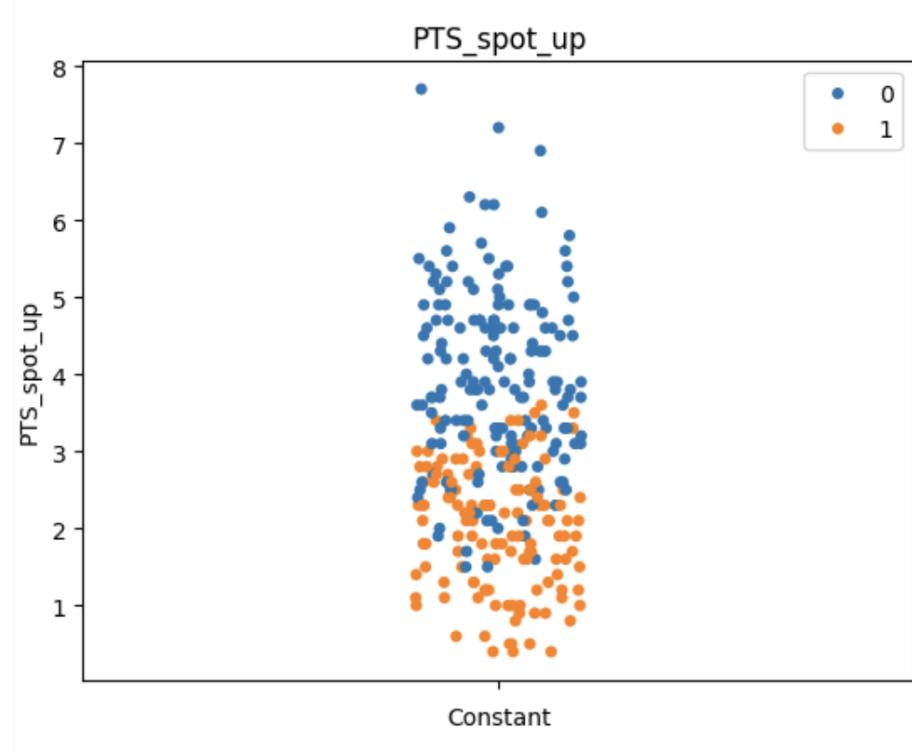
3.7.4 Variance of clusters with FREQ%



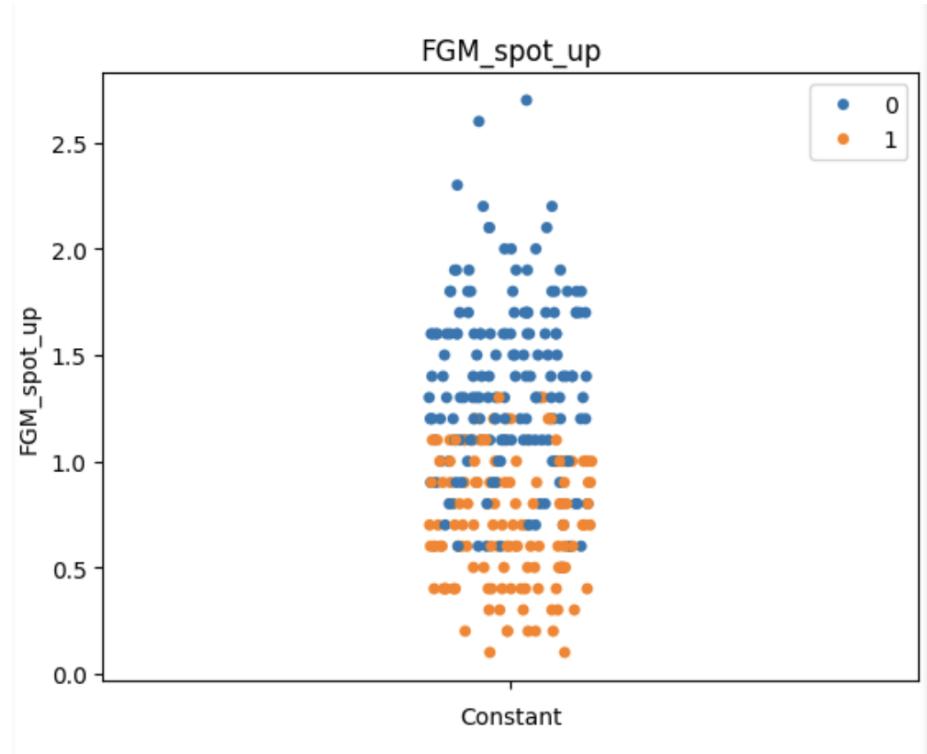
3.7.5 Variance of clusters with PPP



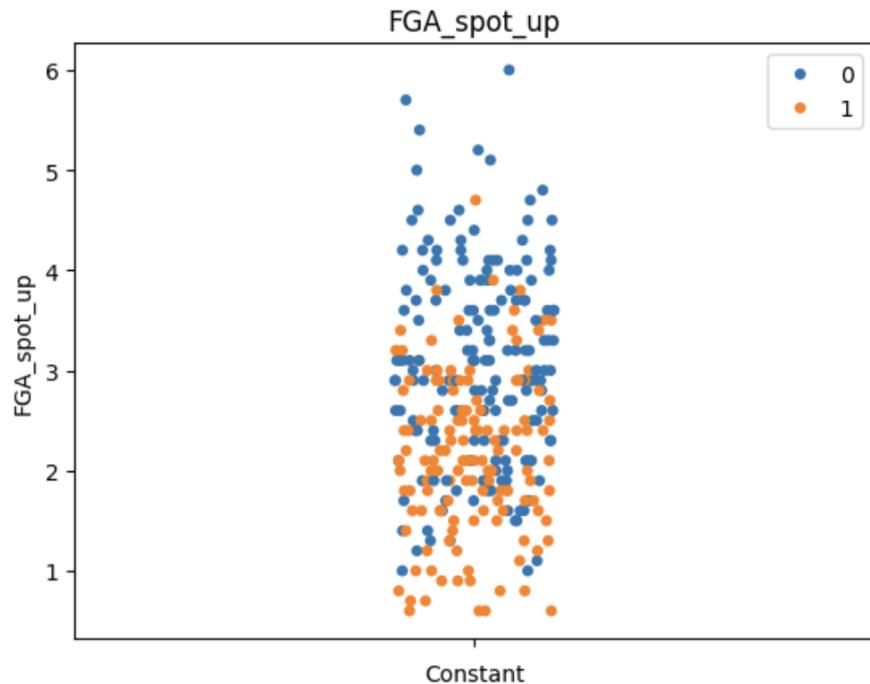
3.7.6 Variance of clusters with PTS



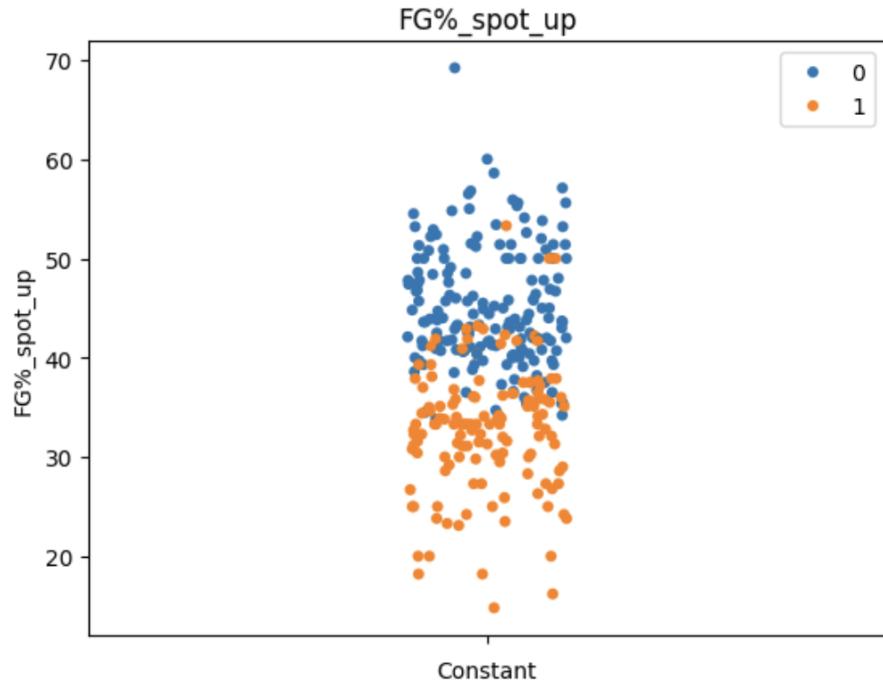
3.7.7 Variance of clusters with FGM



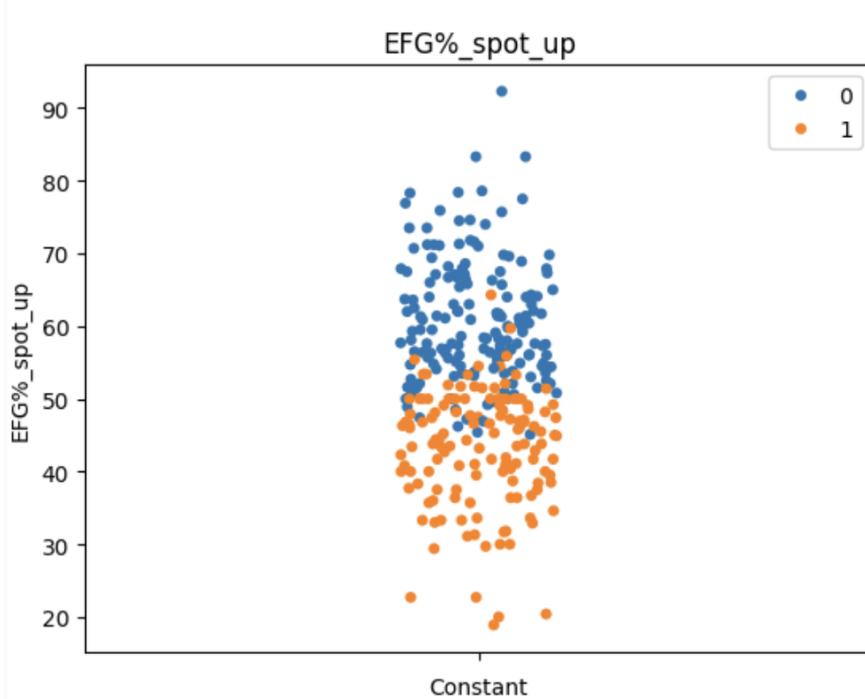
3.7.8 Variance of clusters with FGA



3.7.9 Variance of clusters with FG%



3.7.10 Variance of clusters with EFG%



4 Analyzing Players

4.1 Case Study A0: Nikola Jokic

4.2 Case Study A1: Anthony Davis

4.3 Case Study A2: Tyrese Haliburton

4.4 Case Study A3: Anthony Edwards

5 Analyzing 2023-24 Lineup Compositions

5.1 Lineups

The top five lineups in the NBA this season in terms of offensive rating (over 100 minutes played together) are:

- Nicolas Batum, Tobias Harris, Joel Embiid, De'Anthony Melton, Tyrese Maxey @ PHI (136.4)

- Myles Turner, Buddy Hield, Bruce Brown, Obi Toppin, Tyrese Haliburton @ IND (134.4)
- Kevin Durant, Eric Gordon, Jusuf Nurkic, Devin Booker, Grayson Allen @ PHX (126.6)
- Brook Lopez, Damian Lillard, Khris Middleton, Giannis Antetokounmpo, Malik Beasley @ MIL (125.2)
- James Harden, Paul George, Kawhi Leonard, Ivica Zubac, Terance Mann @ LAC (122.3)

5.2 Case Study B0: Philly

Player	Isolation	Cut	Off screen	PnR handler	PnR roller	Post up	Spot up
Batum	INS.	INS.	INS.	INS.	INS.	INS.	0
Harris	1	0	INS.	0	INS.	0	0
Embiid	0	0	INS.	INS.	0	2	0
Melton	1	INS.	2	2	INS.	INS.	0
Maxey	0	INS.	0	1	INS.	INS.	0

Table 1: Philly Lineup Composition

5.3 Case Study B1: Indiana

Player	Isolation	Cut	Off screen	PnR handler	PnR roller	Post up	Spot up
Turner	INS.	0	INS.	INS.	0	0	1
Hield	INS.	INS.	1	0	1	INS.	0
Brown	2	INS.	INS.	2	INS.	INS.	0
Toppin	INS.	2	INS.	INS.	INS.	INS.	0
Haliburton	2	INS.	INS.	1	INS.	INS.	0

Table 1: Indiana Lineup Composition

5.4 Case Study B2: Phoenix

Player	Isolation	Cut	Off screen	PnR handler	PnR roller	Post up	Spot up
Durant	0	2	1	2	0	2	0
Gordon	INS.	INS.	INS.	0	INS.	INS.	0
Nurkic	INS.	0	INS.	INS.	1	2	1
Booker	0	INS.	1	1	INS.	0	1
Allen	INS.	INS.	INS.	2	INS.	INS.	0

Table 1: Phoenix Lineup Composition

5.5 Case Study B3: Milwaukee

Player	Isolation	Cut	Off screen	PnR handler	PnR roller	Post up	Spot up
Lopez	INS.	0	INS.	INS.	0	1	0
Lillard	0	INS.	INS.	1	INS.	INS.	0
Middleton	2	INS.	INS.	0	INS.	0	1
Antetokounmpo	0	2	INS.	2	0	2	1
Beasley	INS.	INS.	INS.	0	INS.	INS.	0

Table 1: Milwaukee Lineup Composition

5.6 Case Study B4: Clippers

Player	Isolation	Cut	Off screen	PnR handler	PnR roller	Post up	Spot up
Harden	0	INS.	INS.	1	INS.	INS.	0
George	0	INS.	1	2	INS.	0	0
Leonard	0	INS.	0	2	INS.	1	0
Zubac	INS.	1	INS.	INS.	0	0	INS.
Mann	INS.	1	INS.	INS.	0	0	INS.

Table 1: Clippers Lineup Composition

6 Conclusions

7 Further work

Right now, I'm working on a few different extensions of this project:

- Breaking down most effective combos of clusters in starting lineups
- Graphics to holistically visualize a player's strengths/weaknesses as determined by their cluster
- A similar analysis on 2022-23 data to determine areas of improvement
- A similar analysis, but for player defense against the same types of plays