Clustering NBA Players Offensively

Nehal Chigurupati

1 Summary

This document contains the technical details of the project. For a non-technical writeup, see here:
https://medium.com/@jxuwrsb/using-clustering-to-determine-crucial-skills-in-crafting-lineups-3b8f35681f5d

The objective of this project is to determine groups of NBA players who perform similarly on certain play-types (isolation, cut, off-ball screen, pick and roll handler, pick and roll roller, spot-up shooter, and post-up shooter). These groups/clusters were then used to break down the composition of the three lineups in the NBA with the highest offensive ratings, yielding three lessons for team-building:

- In all three lineups, at least three out of the five players on the floor were extremely efficient ($>$ .9 PPP) in high volume at spot-up shooting, suggesting that **teams should establish a high baseline level of shooting for players on the floor**.
- No more than two players in each lineup attempted post-up shots in volume, and the second and third most efficient offenses only had two players who attempted isolation shots in volume. The lesson seems to be that **teams should de-emphasize isolation and post-up talent when crafting rosters, prioritizing spot-up shooting talent.**
- All three lineups had an extremely efficient pick and roll combo using a big man and a guard. The big men typically were excellent shooters, passers, and finishers, permitting inverted pick and rolls. **Teams should target these versatile big men during free agency and trading season.**

Contents

2 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-ball screen
- Pick and roll ball handler
- Pick and roll roller
- Post up
- Spot up

The clusters were then used to study the compositions of the three lineups in the NBA with the highest offensive ratings this season (minimum 100+ minutes).

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.

3 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
- 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 clusterings were produced, 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 its selection.

4 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. For the strip plots, refer to the appendix.

4.1 Isolation

4.1.1 Describing Clusters

There are three isolation clusters. The range of values attained by 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 a lot of isolation plays, and are moderately efficient at generating points off such possessions. Cluster 1 players attempt isolation plays relatively infrequently, and aren't very productive in such plays. 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 weights and heights 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 range 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 consists of bigs, with a right skew ranging from 78-88 inches (6-5 to 7-3).

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:

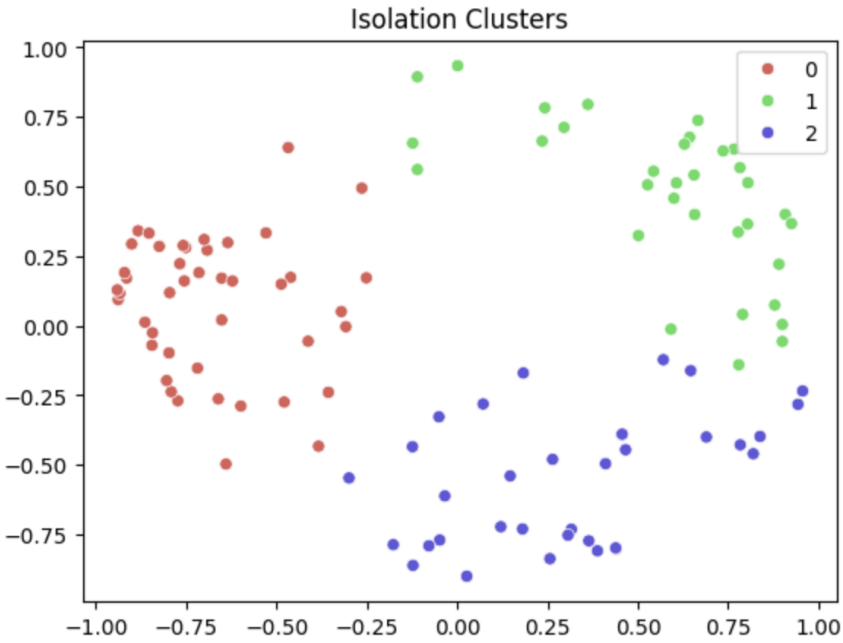
- Cade Cunningham
- Darius Garland
- Dennis Schroder
- Chris Paul
- Collin Sexton
- Jamal Murray
- Josh Giddey

Cluster 2 players include:

- Jaren Jackson Jr.
- Nikola Jokic
- Bam Adebayo
- Anthony Davis
- Victor Wembanyama
- Chet Holmgren
- Evan Mobley

(these are not exhaustive lists)

4.1.2 Clusters under PCA



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

4.2 Cuts

4.2.1 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	.4-2.1	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 at the same clip 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 and weights of players in each cluster, it appears 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

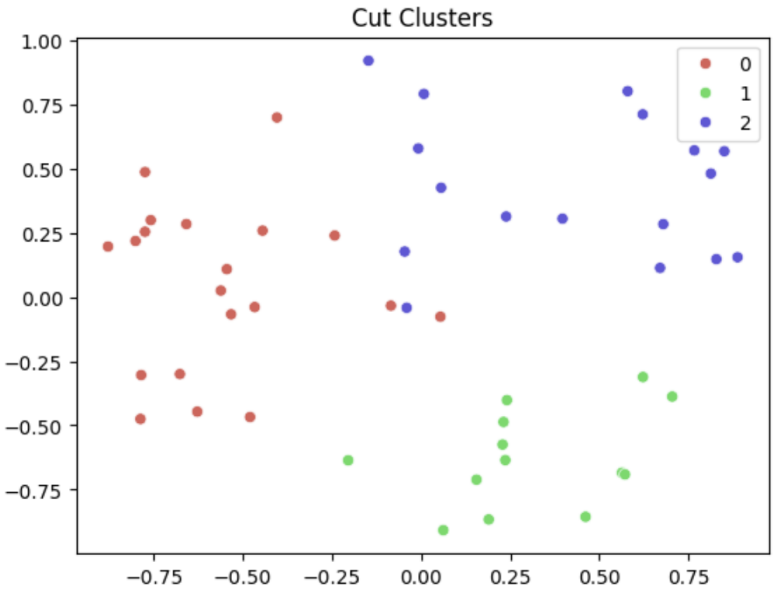
(these are not exhaustive lists)

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

4.2.2 Clusters under PCA



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

4.3 Off screen

4.3.1 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

4.3.2 Clusters under PCA

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

4.4 Pick and roll ball handler

4.4.1 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 2 players are similar in height to cluster 1 players, but weigh significantly more.

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).

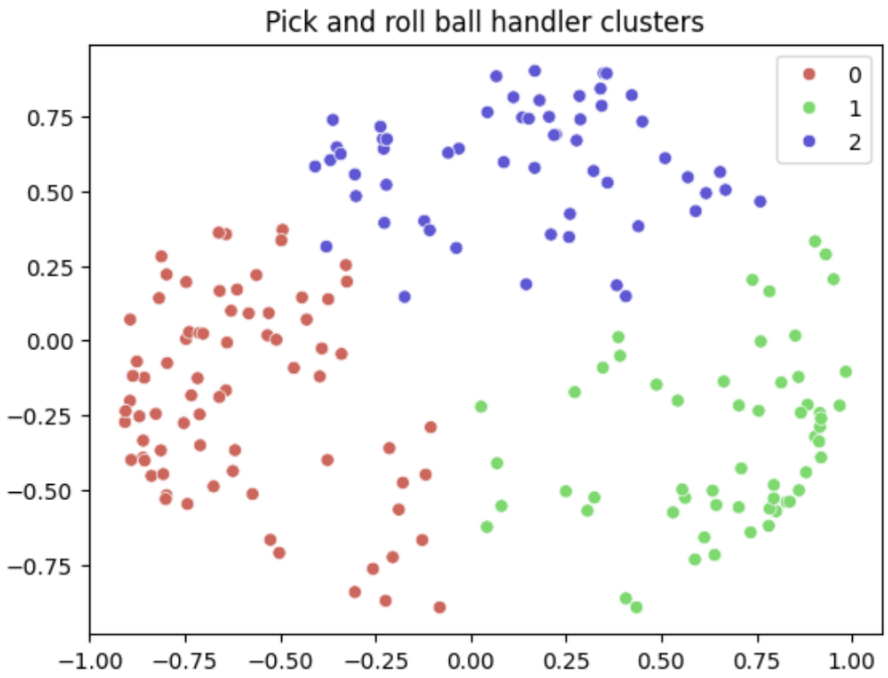
Finally, cluster 2 includes large forwards who use the pick and roll to create size mismatches, overpowering their defender to get to the rim. Because of their size advantage, they are extremely efficient as the ball handler in pick and rolls, generating between .78 and 1.60 PPP. However, they don't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

4.4.2 Clusters under PCA



4.5 Pick and roll roller

4.5.1 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.

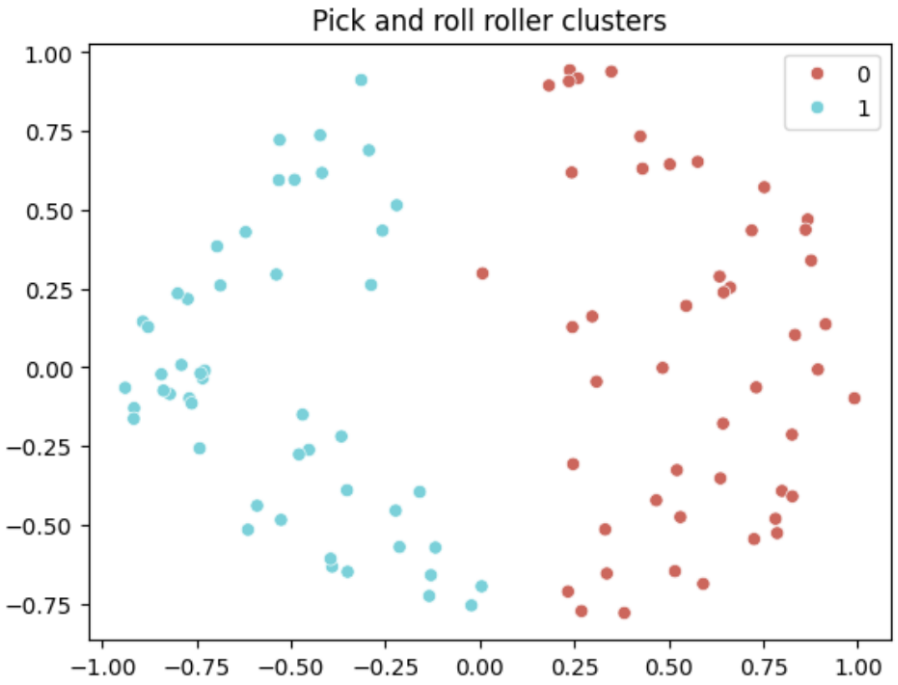
Cluster 1 consists of stretch bigs who are efficient shooters. As rollers, they stretch the floor, forcing defenders into tough decisions when hedging or defending the roll.

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

4.5.2 Clusters under PCA



4.6 Post up

4.6.1 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.

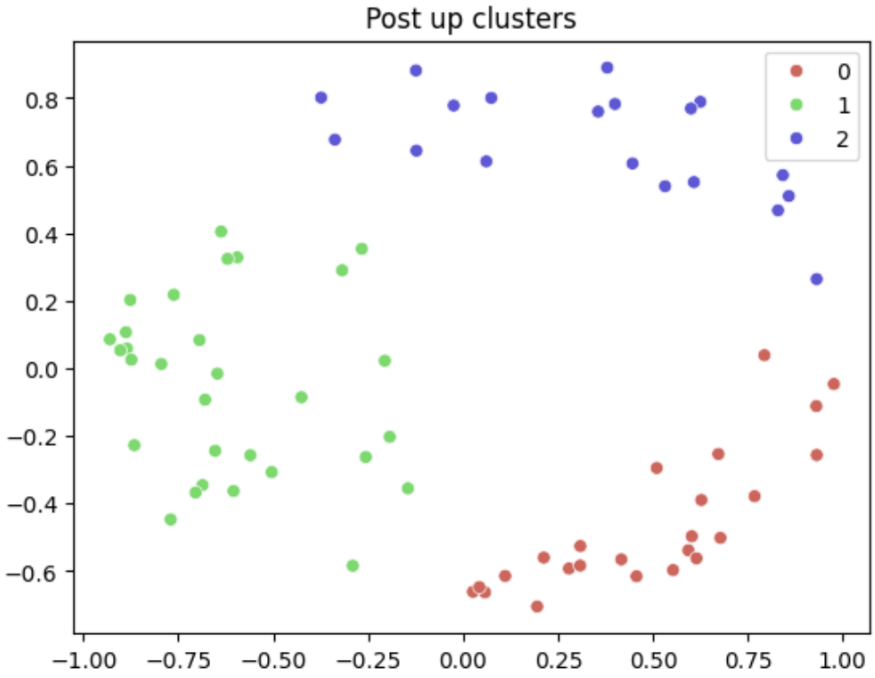
Cluster 1 players are the least efficient of the group. They attempt fadeaways and turnaround jumpers, but aren't as efficient on such attempts as cluster 0 players. The strength of these players is in their power: Jimmy Butler and Kawhi Leonard, for instance, can absorb contact in the post and finish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

4.6.2 Clusters under PCA



4.7 Spot up

4.7.1 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 composed of 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.

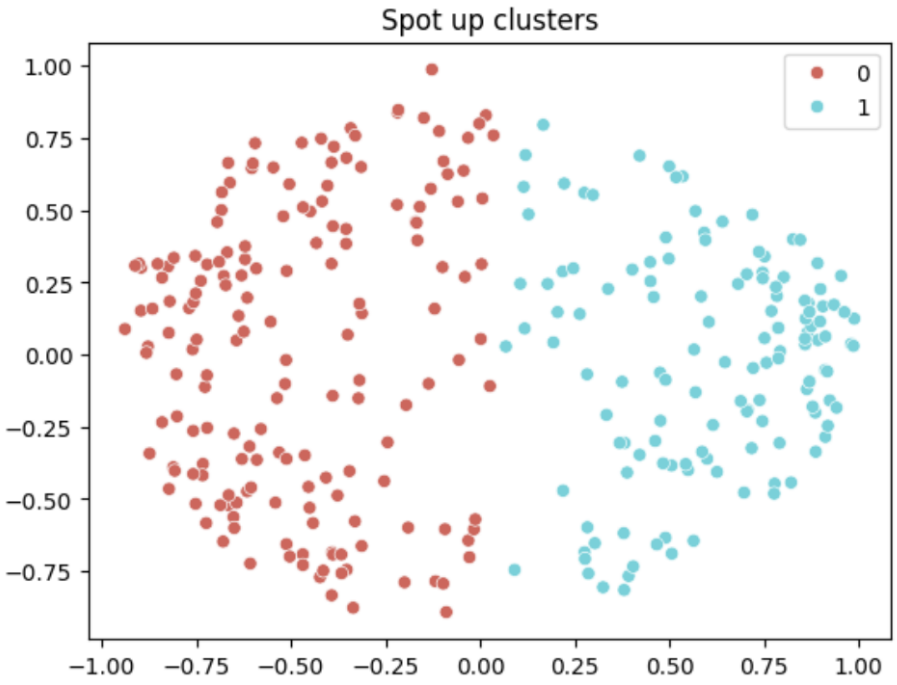
Cluster 1 consists of large forwards and centers who primarily take shots from the high to mid post. They generally are less effective spot up shooters than cluster 0 players, preferring to leverage their size by driving after catching the ball.

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

4.7.2 Clusters under PCA



5 Analyzing 2023-24 Lineup Compositions

5.1 Lineups

The top three lineups in the NBA this season (as of December 18th, 2023) 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)

Note: in the lineup composition tables below, "INS" means that the given player does not partake in possessions of a certain kind frequently enough for the data to be recorded by the NBA.

5.2 Case Study 0: 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

Almost all players in the Philadelphia 76ers lineup can play isolation. Joel Embiid and Tyrese Maxey, the focal points of the Sixers' offense, are associated with the "less efficient, many attempts" cluster of isolation players. Tobias Harris and De'Anthony Melton don't attempt as many isolation plays, but they score at a higher rate on such attempts.

Harris and Embiid are the only two players on the Sixers who consistently cut. They both belong to cluster 0, meaning they cut infrequently but are able to score efficiently on cuts due to their strong post play. This cluster also includes big men like Brook Lopez and Bam Adebayo.

Only Melton and Maxey frequently employ off-ball screens. Melton, in cluster 2, has the ability to create mismatches and convert, similar to players like Nikola Jokic, Karl-Anthony Towns, and LeBron James. Maxey, in cluster 0, is agile in navigating off-ball screens to create space but is relatively inefficient in shot attempts out of such screens. This inefficiency isn't necessarily a problem for Maxey, who primarily handles the ball.

Philadelphia primarily employs Harris, Melton, and Maxey as ball handlers in pick-and-rolls. Harris is in cluster 0, exhibiting a unique athletic ability to finish through extreme coverages (doubles, etc.). This cluster, the least efficient of the three for pick-and-roll ball handlers, also includes Paolo Banchero, Anthony Davis, and Shaedon Sharpe. Maxey is in cluster 1, alongside guards like Trae Young and Luka Doncic, indicating effectiveness at weaving to the rim, shooting, and passing out of the pick and roll to open shooters. Cluster 1 is significantly more efficient than cluster 0. Lastly, Melton is in cluster 2, the most efficient group, similar in performance to players like Grayson Allen, Lonnie Walker, and Benedict Mathurin.

The focus of pick and rolls in the Philly offense is to get the ball to Joel Embiid down deep. Embiid, as a roller, is in cluster 0, producing at an extremely efficient rate when getting deep in the paint. This cluster also includes Chet Holmgren, Nikola Jokic, and Alperen Sengun, known for their near-unstoppable performance when rolling to the rim.

Harris and Embiid are the only two Sixers who frequently post up. Both are in

efficient post-up clusters. Harris shares a cluster with players known for their arsenal of fadeaways and turnaround jumpers, like DeMar DeRozan, LeBron James, and Jayson Tatum. Embiid is in cluster 2, composed of "post bigs" who can reliably get up a quality shot over any defender. This cluster also includes Nikola Jokic, Karl-Anthony Towns, Bam Adebayo, and Deandre Ayton.

A standout feature of this lineup is the quality of spot-up shooters. All players are in cluster 0, indicating high efficiency (.88-1.73 points per possession from 45.1-92.3% effective field goal percentage), ranking them among excellent catch-and-shoot players like Michael Porter Jr.

Philly's spot-up shooting is particularly effective considering the deadly pick-and-roll combos and post-up ability in the lineup. Any extra attention defenders pay to pick-and-rolls (with Maxey/Melton handling and Embiid rolling) or post-ups (by Embiid and Harris) can be easily exploited by passing out to efficient perimeter spot-up shooters.

5.3 Case Study 1: 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

Bruce Brown and Tyrese Haliburton are the only Pacers players in the lineup that frequently employ isolation play. Interestingly, both have the efficiency in isolation typical of big men, despite being guards. Players in the same cluster include Nikola Jokic and Bam Adebayo.

The Pacers have two excellent cutters in Myles Turner and Obi Toppin. Turner belongs to cluster 0, indicating efficiency and strong post play, while Toppin is in cluster 2, also denoting skilled, athletic finishing and high efficiency off cuts. Other players in cluster 0 include Brook Lopez and Alperen Sengun, with Giannis Antetokounmpo and De'Andre Hunter in cluster 2. Tyrese Haliburton, an excellent passing point guard for the Pacers, serves Turner and Toppin with passes on cuts that they convert with their athleticism and post skills.

Buddy Hield heavily utilizes off-ball screens and belongs to cluster 1, primarily composed of catch-and-shoot specialists. He uses even small spatial concessions by the defense to efficiently shoot threes. This cluster also includes Michael Porter Jr. and Kevin Durant.

Indiana relies primarily on Buddy Hield, Bruce Brown, and Tyrese Halibur-

ton as ball handlers in pick-and-roll situations. Hield, part of cluster 0, finishes well as a ball handler, even when doubled, though he isn't as efficient as players in clusters 1 and 2. Brown, in cluster 1, demonstrates the ability to weave to the rim, shoot, and pass out to the perimeter effectively, sharing this cluster with guards like Trae Young and Luka Doncic. Haliburton, in the most efficient cluster 2, shows an efficiency resembling that of Paul George and LeBron James, producing reliable, high-quality shots by creating and overpowering mismatches.

Indiana uses Myles Turner and Buddy Hield for rolls. Turner, in cluster 0, the more efficient of the two, is among strong post finishers like LeBron James, Paul George, Lauri Markkanen, and Jayson Tatum.

The post-up is not a significant part of Indy's game, with Myles Turner being the only frequent user. He is in cluster 0, associated with the highest-tier efficiency, alongside Tobias Harris, Kristaps Porzingis, and Paul George.

Lastly, the Indy lineup, similar to Philly, features extremely efficient spot-up shooters like Hield, Brown, Toppin, and Haliburton, all in cluster 0, indicating great efficiency from 3 on spot-up shots. Haliburton's passing ability and the availability of efficient spot-up shooters open up the middle of the floor for Turner's post-ups.

5.4 Case Study 2: 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

The two primary isolation players on the Phoenix Suns are Kevin Durant and Devin Booker. Both belong to cluster 0, suggesting they are somewhat inefficient despite many attempts. This cluster also includes players like Jimmy Butler and Jayson Tatum.

Phoenix relies on Durant and Jusuf Nurkic for cuts. Durant is in cluster 2, indicating mid-level efficiency, while Nurkic is a part of cluster 0, known for extreme productivity on shot attempts from cuts. Nurkic shares this cluster with Bam Adebayo, Alperen Sengun, and Nikola Vucevic, while Durant's group includes De'Andre Hunter, Kristaps Porzingis, Keldon Johnson, and Obi Toppin.

For off-ball screens, Durant and Booker, two prolific catch-and-shoot players, are heavily utilized. Both are in cluster 1, representing incredible efficiency,

alongside players like Paul George, Brandon Ingram, Michael Porter Jr., and Lauri Markkanen.

In pick-and-roll scenarios, the Suns frequently use Durant, Eric Gordon, and Devin Booker. Durant, in cluster 2, is the most efficient, capable of overpowering defenders. This cluster also includes LeBron James and Paul George. Eric Gordon is in cluster 0, indicating lower efficiency, and Booker is in cluster 1, proficient at penetrating to the rim and shooting, punishing defenders who go under the screen. He is more efficient than cluster 0 but less so than cluster 2 players.

The Suns rely on Durant and Nurkic for rolling to the rim. Durant, in cluster 0, signifies top-tier efficiency, while Nurkic in cluster 1 is slightly less efficient. Other players in cluster 0 include Chet Holmgren, Myles Turner, and Nikola Jokic, whereas Nurkic shares cluster 1 with players like Lauri Markkanen, Mitchell Robinson, Kyle Kuzma, and Zach Collins. Ideal Suns pick-and-roll situations involve Durant setting a screen for Booker or Allen, then rolling to the rim.

The Suns utilize post-up plays more than the previous two lineups. Durant, Nurkic, and Booker frequently engage in this kind of play. Nurkic and Durant are in cluster 2, indicating mid-level efficiency, in the company of players like Deandre Ayton, Pascal Siakam, and Bam Adebayo. Booker, with a variety of jumpers and fadeaways, is firmly in cluster 0, the most efficient group, alongside DeMar DeRozan, Lauri Markkanen, Paul George, and Kristaps Porzingis.

Finally, like Indiana and Philadelphia, Phoenix is replete with proficient spot-up shooters. Most notably, Durant, Gordon, and Allen belong to cluster 0, indicating extreme efficiency in spot-up situations from the perimeter. Jusuf Nurkic and Devin Booker, the only two in cluster 1, excel in pick-and-roll situations, suggesting an optimal setup where Nurkic sets a screen for Booker, initiating a pick-and-roll, with other players spread around the perimeter, ready to shoot.

In summary, the key to success for all three lineups appears to be having excellent pick-and-roll combinations with big men who can also score efficiently on post-ups and at the rim, while maintaining effective spot-up shooting on the perimeter. This strategy optimizes for the two most efficient shots in basketball: open three pointers and layups.

6 Further work

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

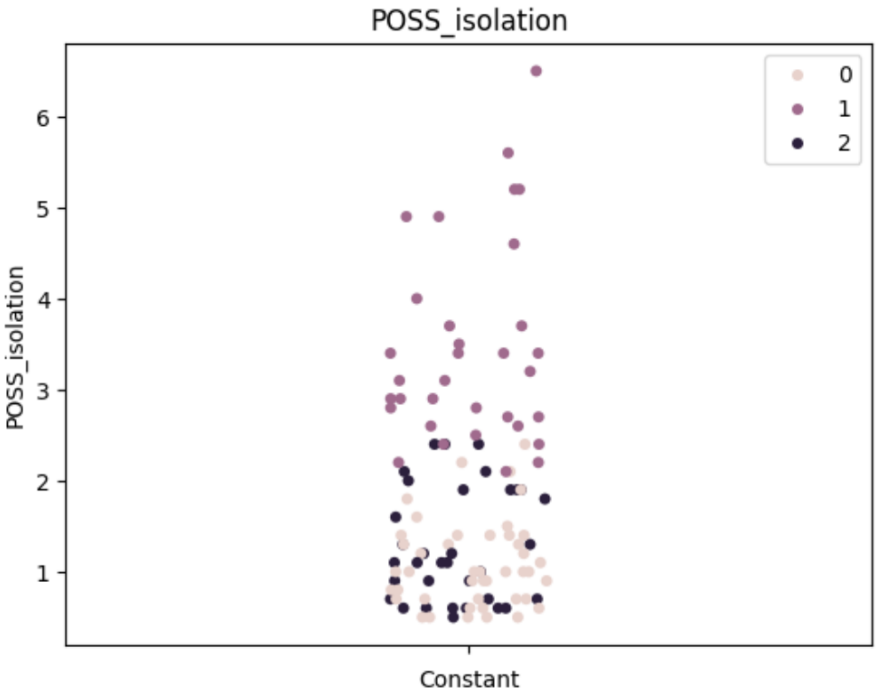
- Graphics to holistically visualize a player's strengths/weaknesses as determined by their cluster
- A similar analysis on 2022-23 data to determine areas of improvement

- A similar analysis, but for player defense against the same types of plays

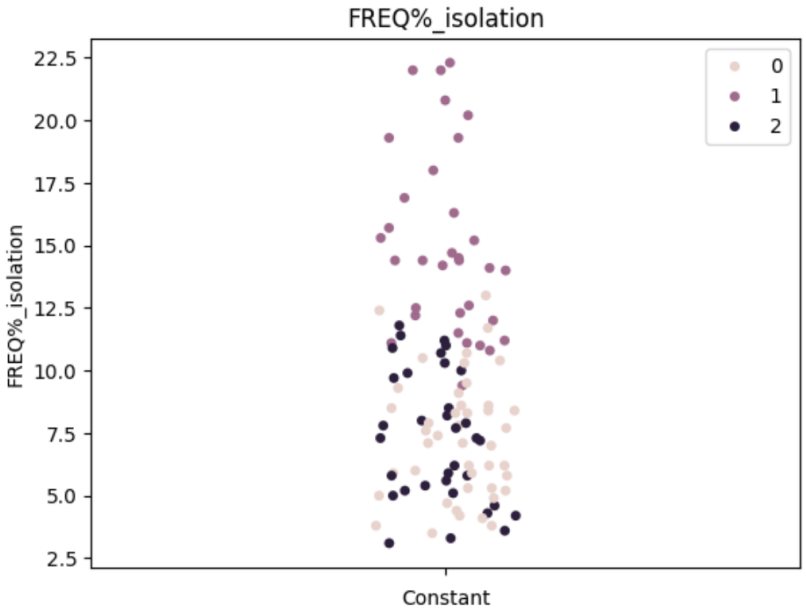
7 Appendix

7.1 Isolation strip plots

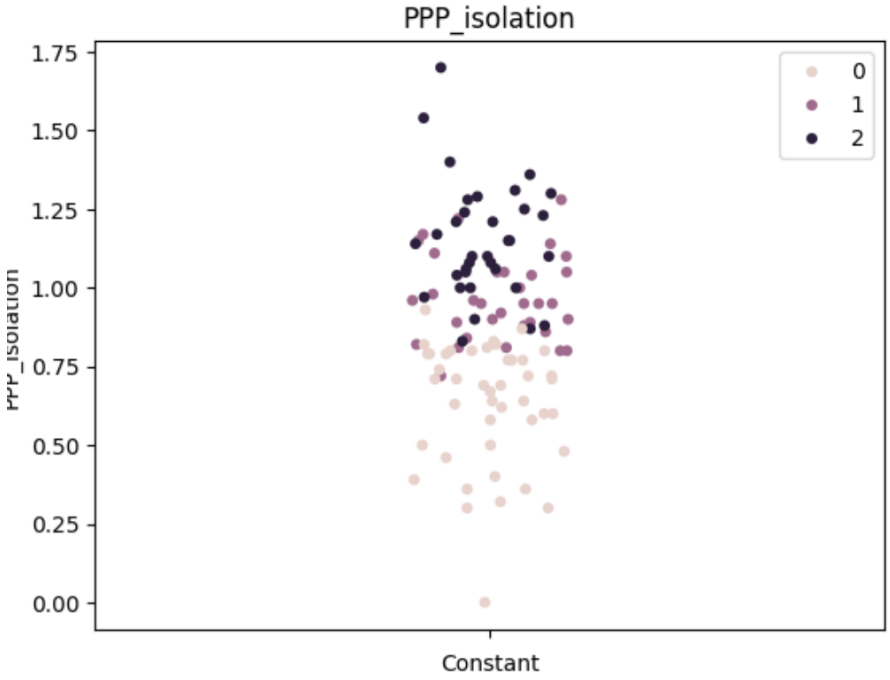
7.1.1 Variance of clusters with POSS



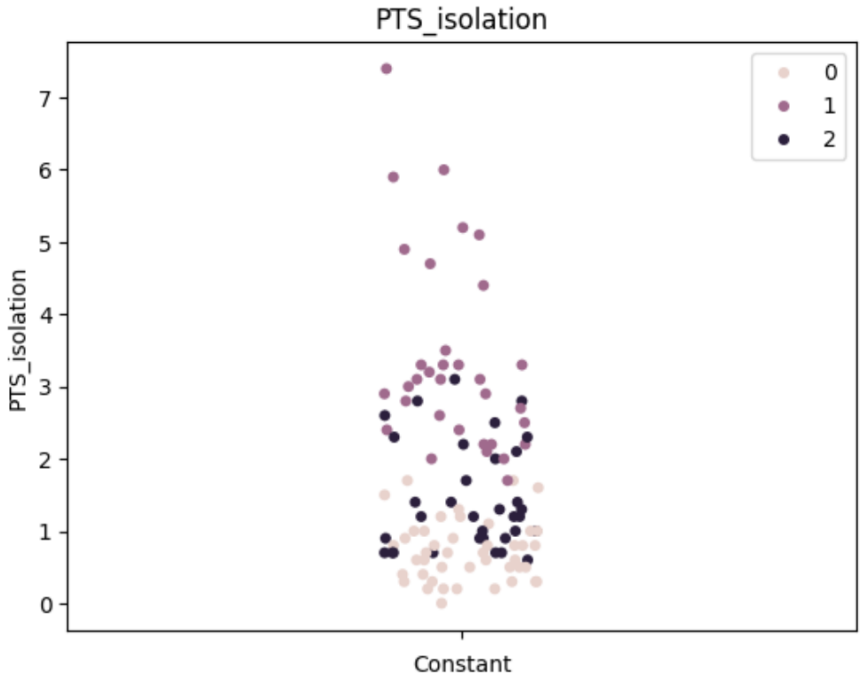
7.1.2 Variance of clusters with FREQ%



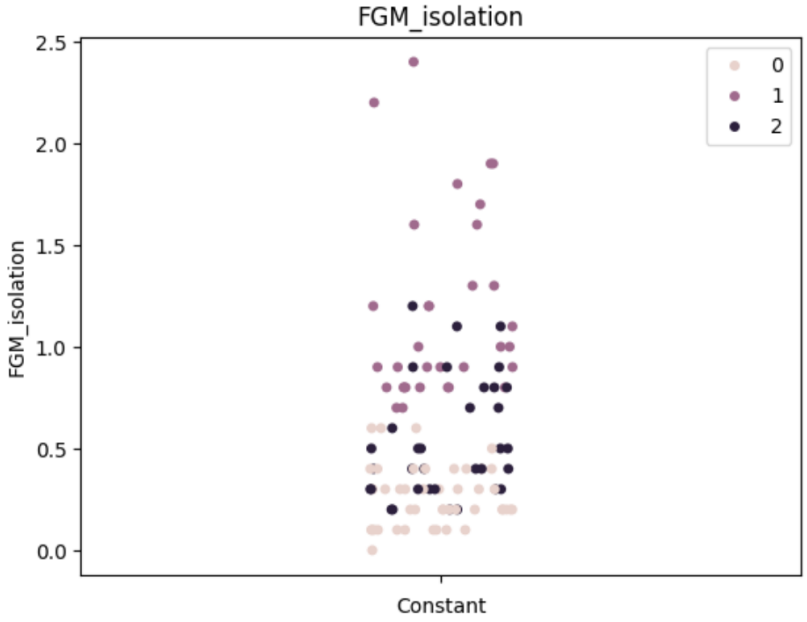
7.1.3 Variance of clusters with PPP



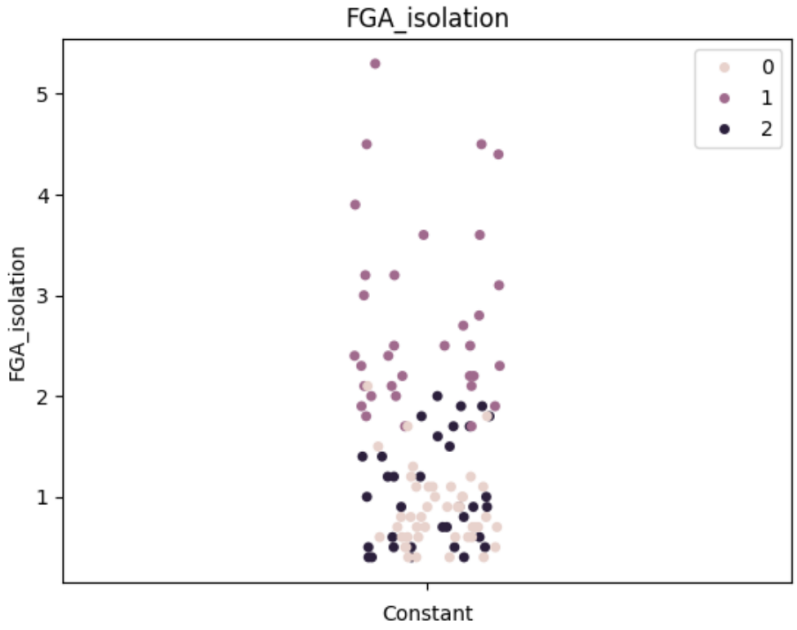
7.1.4 Variance of clusters with PTS



7.1.5 Variance of clusters with FGM

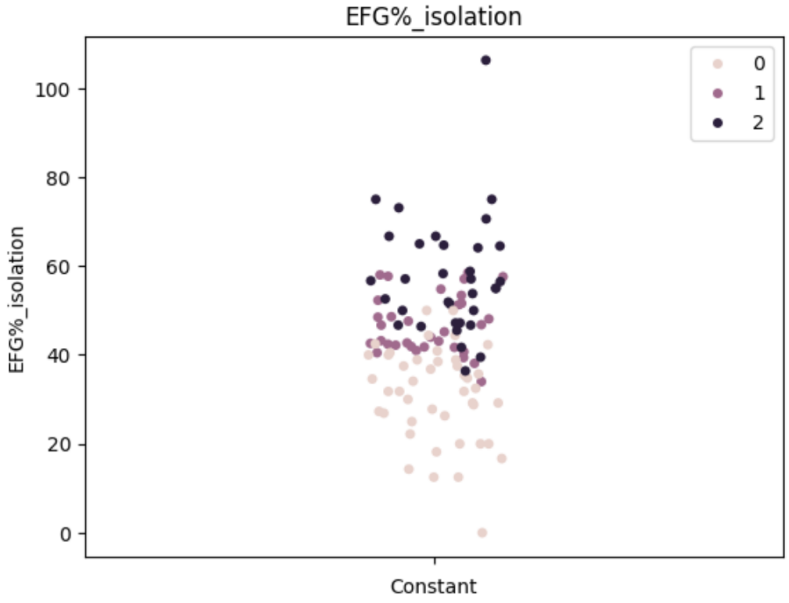


7.1.6 Variance of clusters with FGA



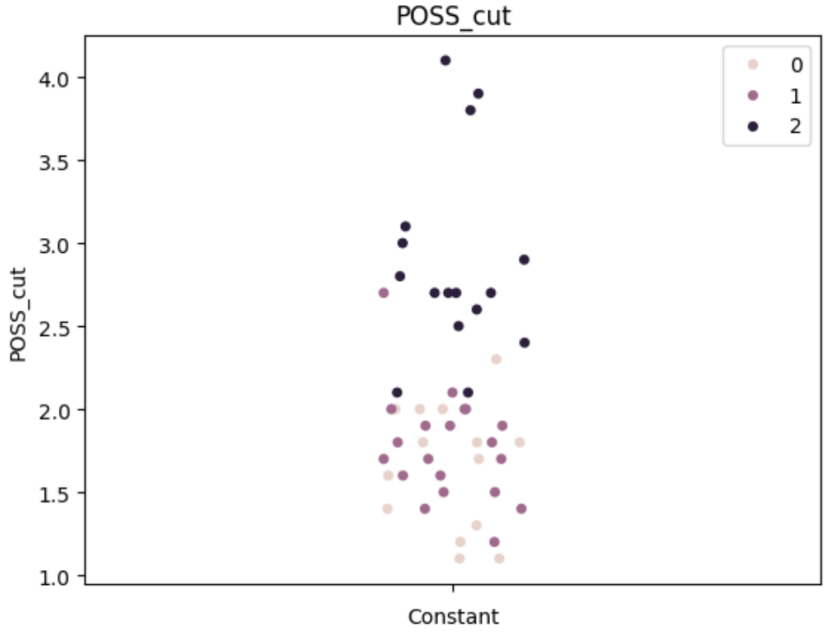
7.1.7 Variance of clusters with FG%

7.1.8 Variance of clusters with EFG%

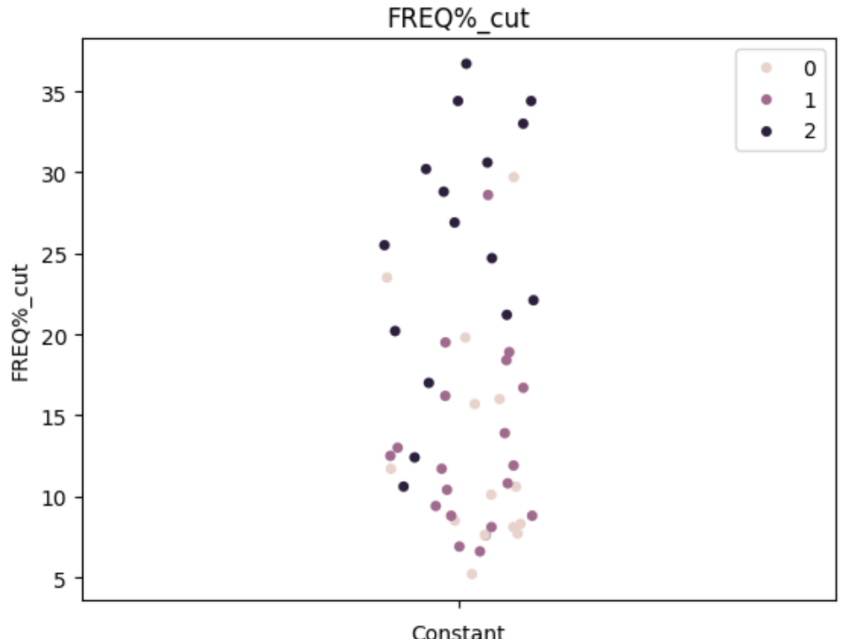


7.2 Cut strip plots

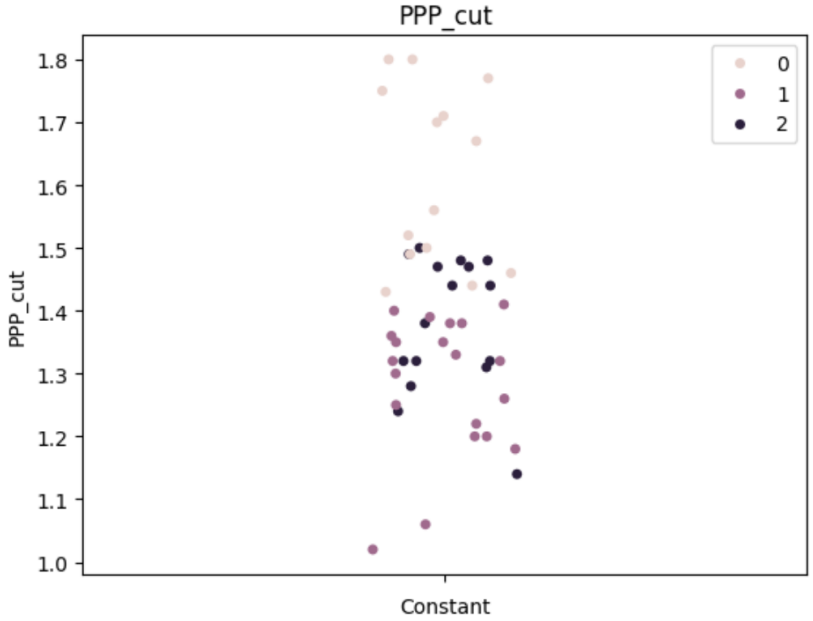
7.2.1 Variance of clusters with POSS



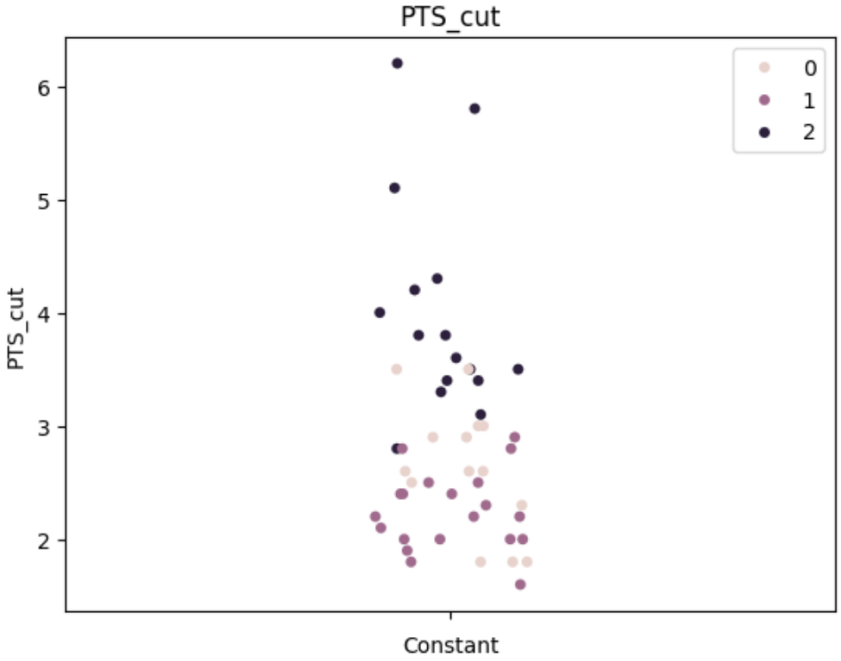
7.2.2 Variance of clusters with FREQ%



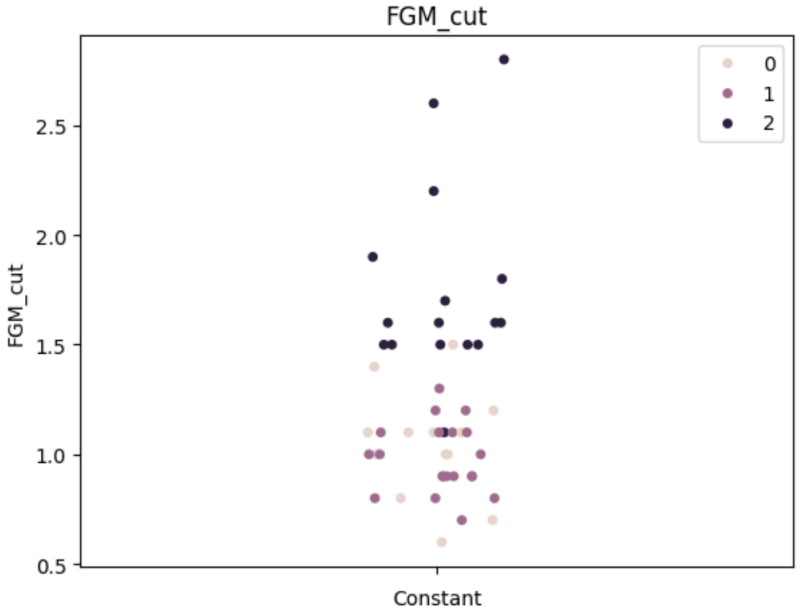
7.2.3 Variance of clusters with PPP



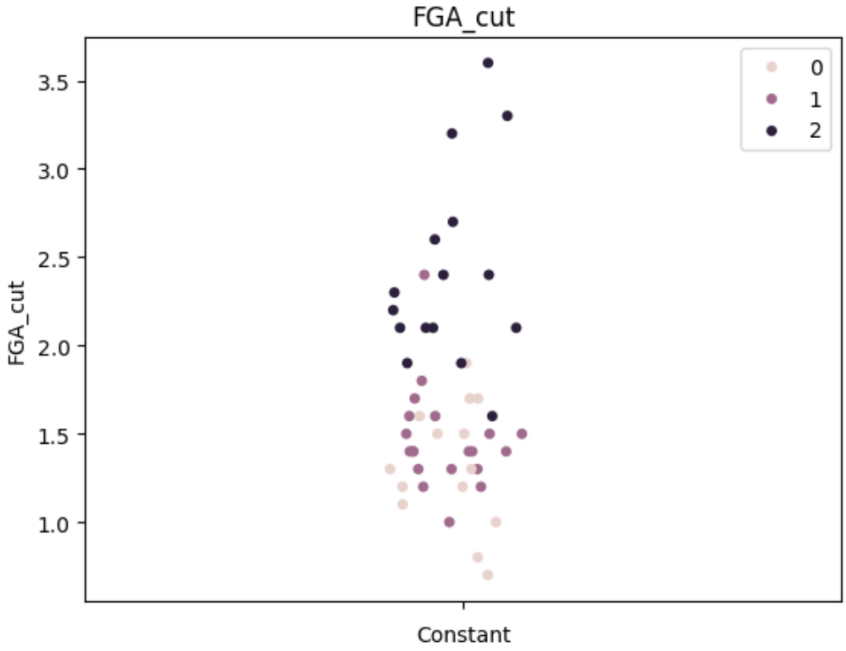
7.2.4 Variance of clusters with PTS



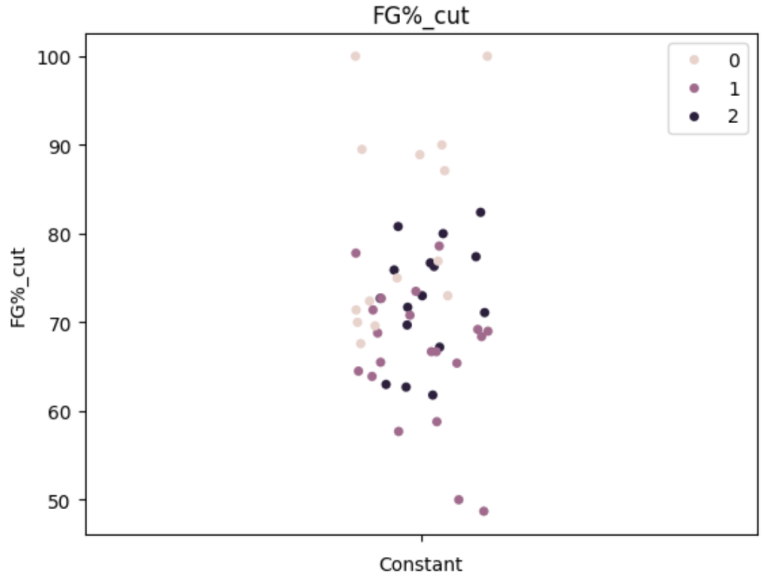
7.2.5 Variance of clusters with FGM



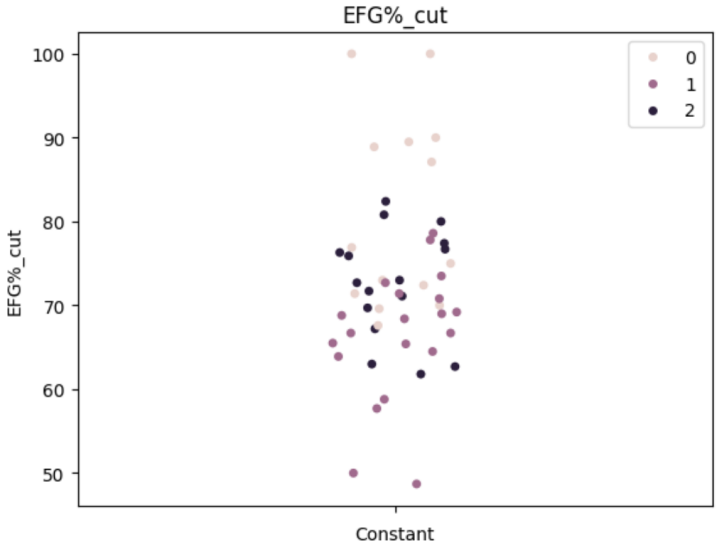
7.2.6 Variance of clusters with FGA



7.2.7 Variance of clusters with FG%

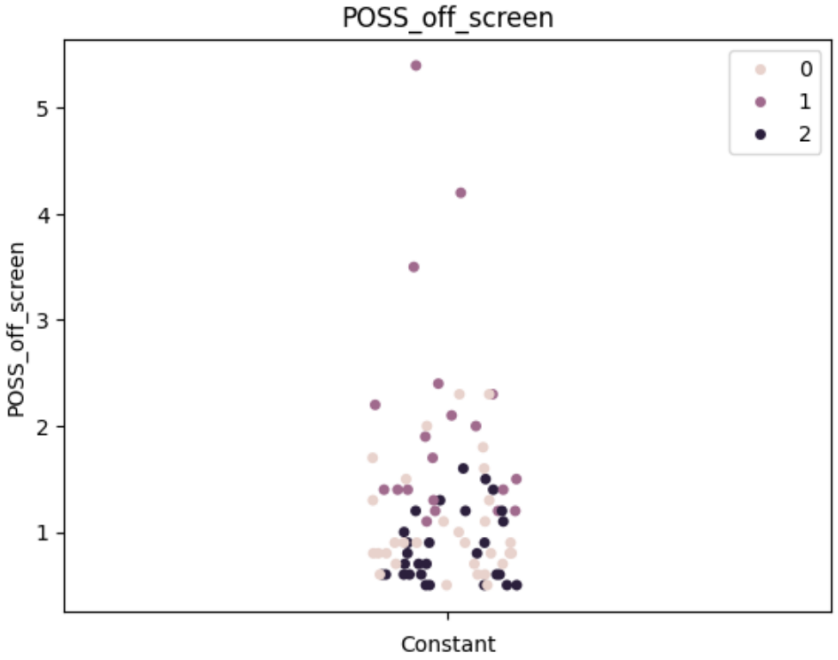


7.2.8 Variance of clusters with EFG%



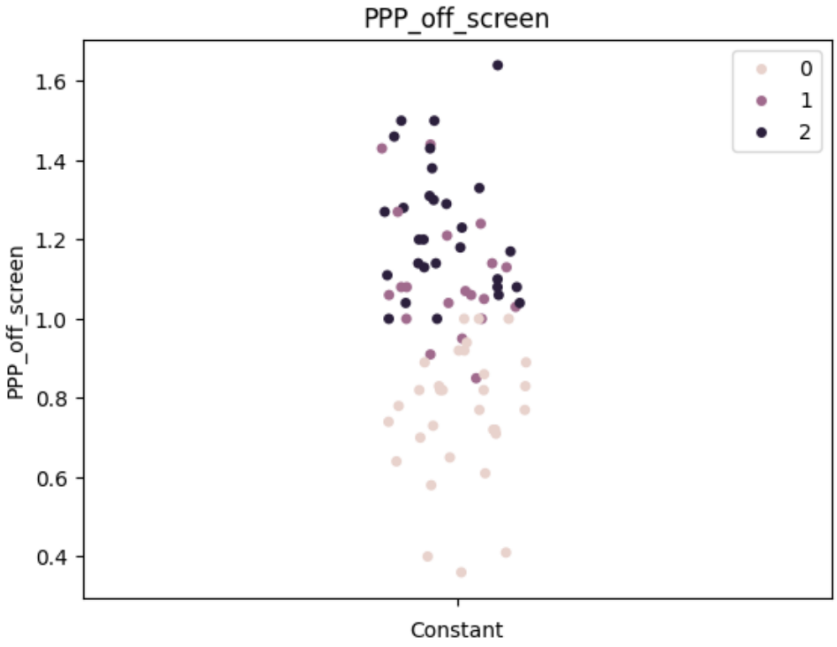
7.3 Off screen strip plots

7.3.1 Variance of clusters with POSS



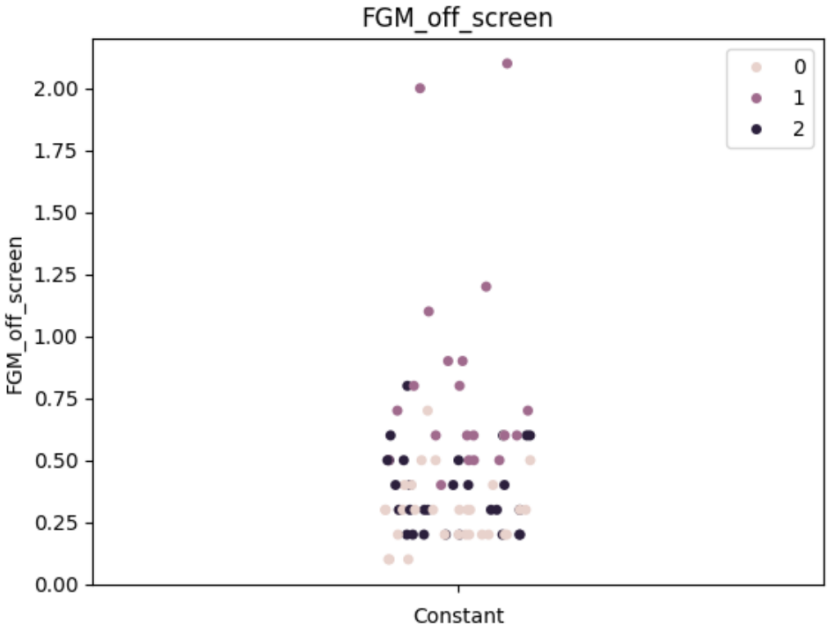
7.3.2 Variance of clusters with FREQ%

7.3.3 Variance of clusters with PPP

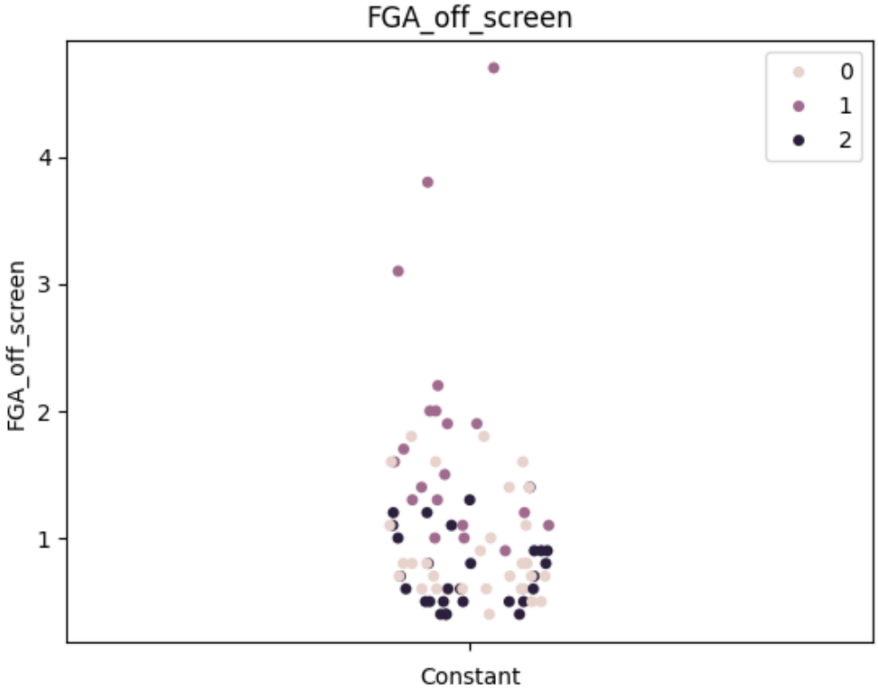


7.3.4 Variance of clusters with PTS

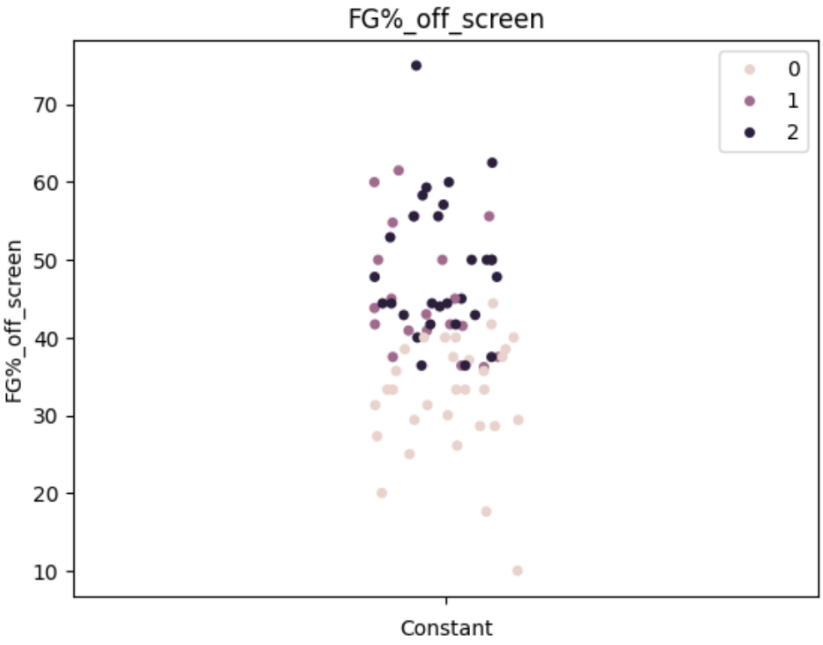
7.3.5 Variance of clusters with FGM



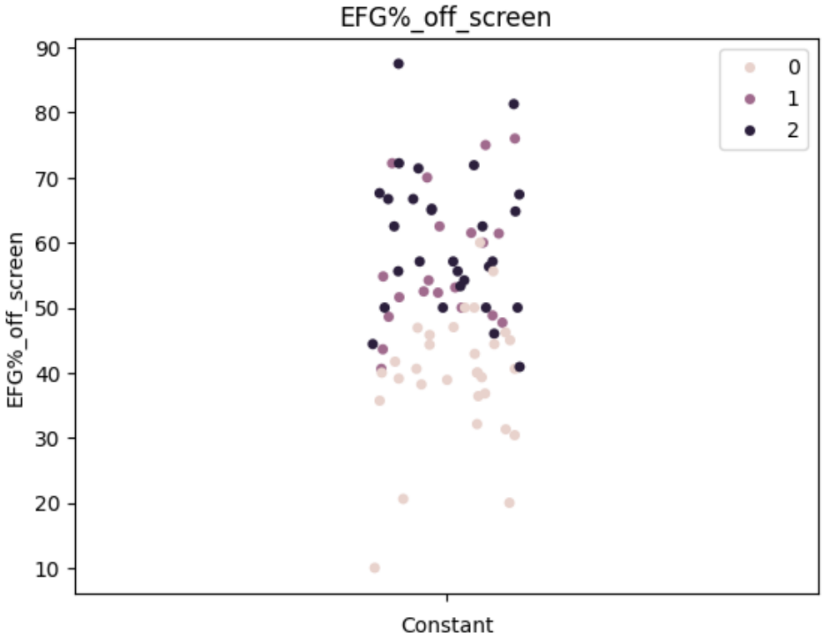
7.3.6 Variance of clusters with FGA



7.3.7 Variance of clusters with FG%

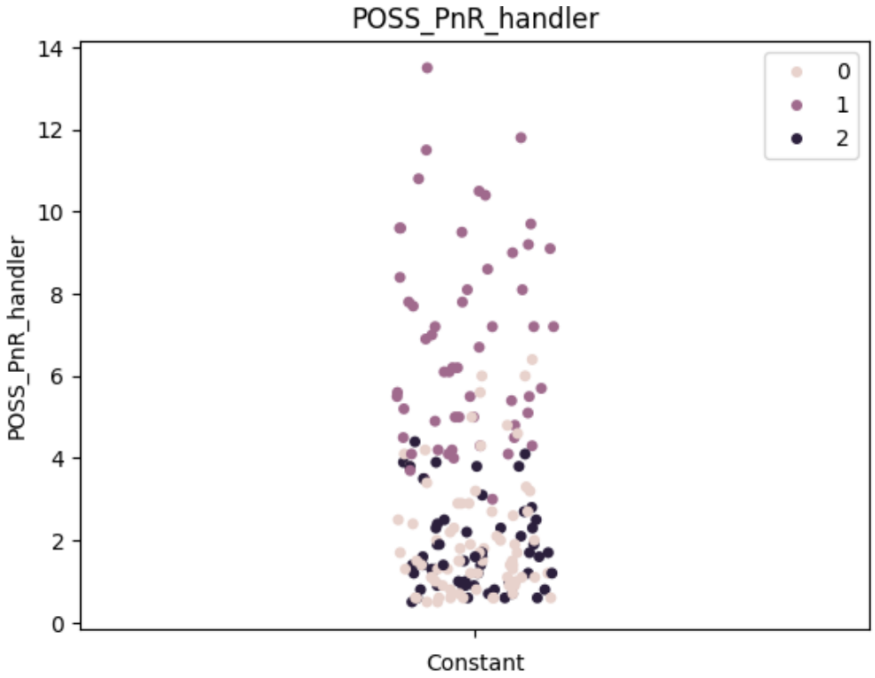


7.3.8 Variance of clusters with EFG%

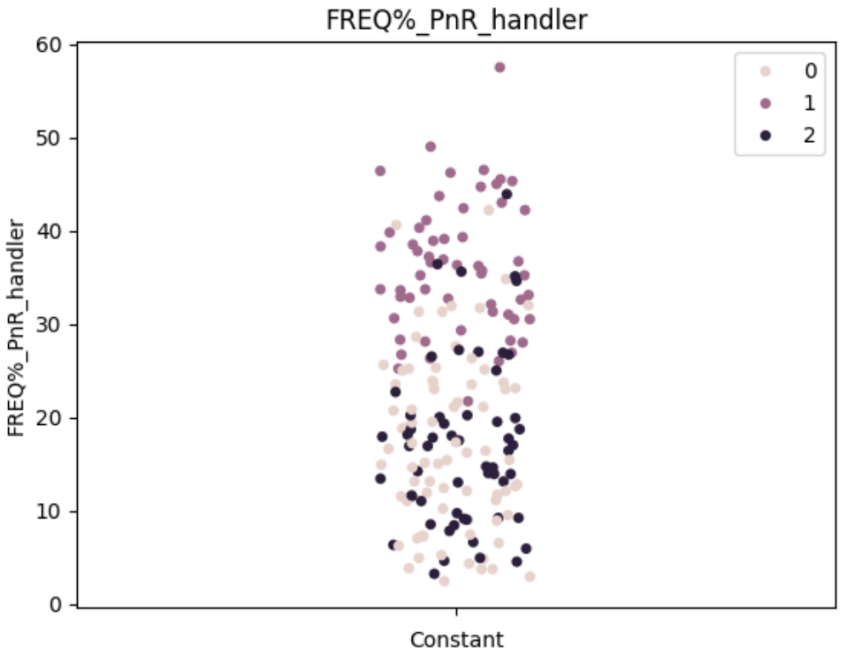


7.4 Pick and roll handler strip plots

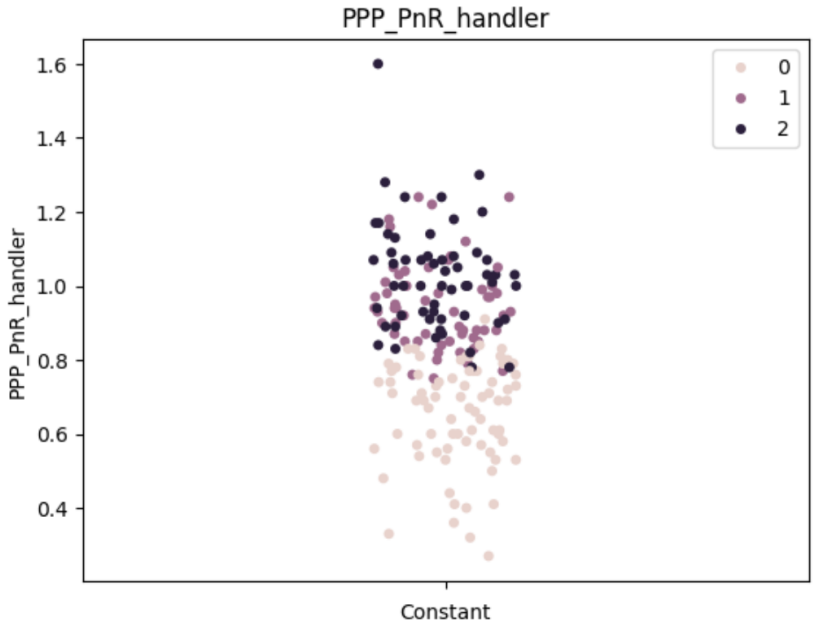
7.4.1 Variance of clusters with POSS



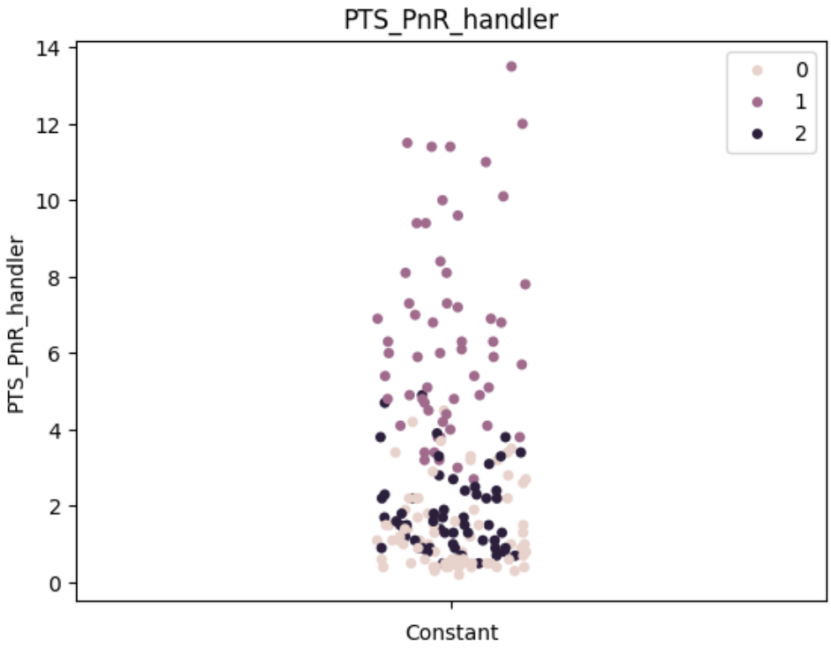
7.4.2 Variance of clusters with FREQ%



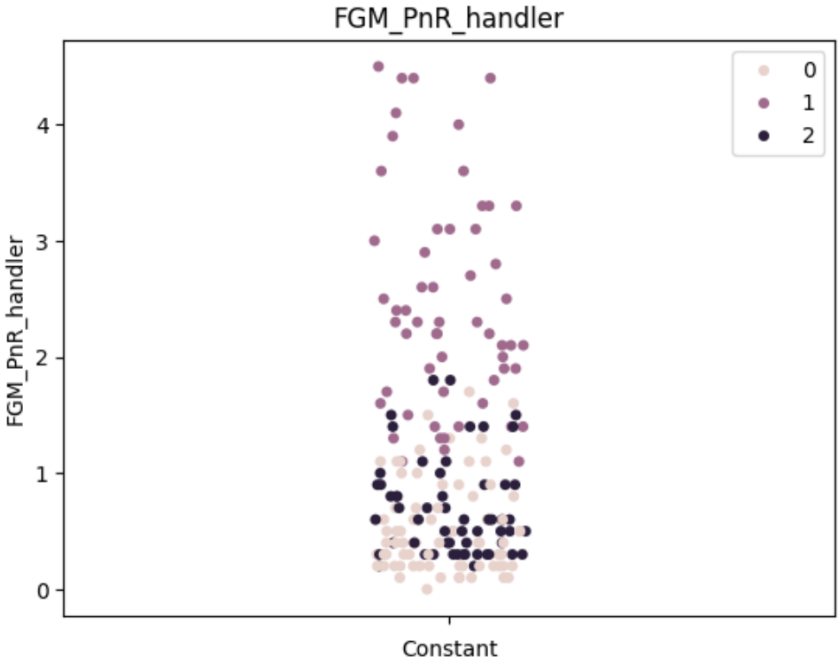
7.4.3 Variance of clusters with PPP



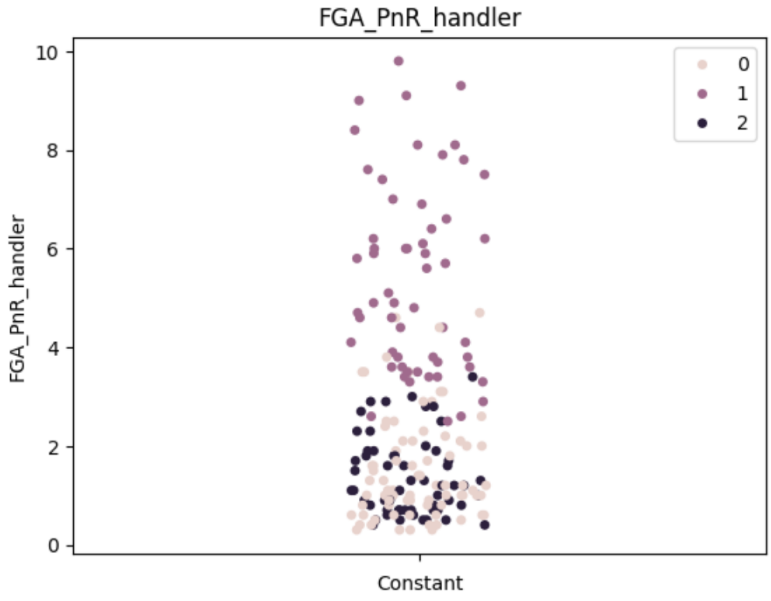
7.4.4 Variance of clusters with PTS



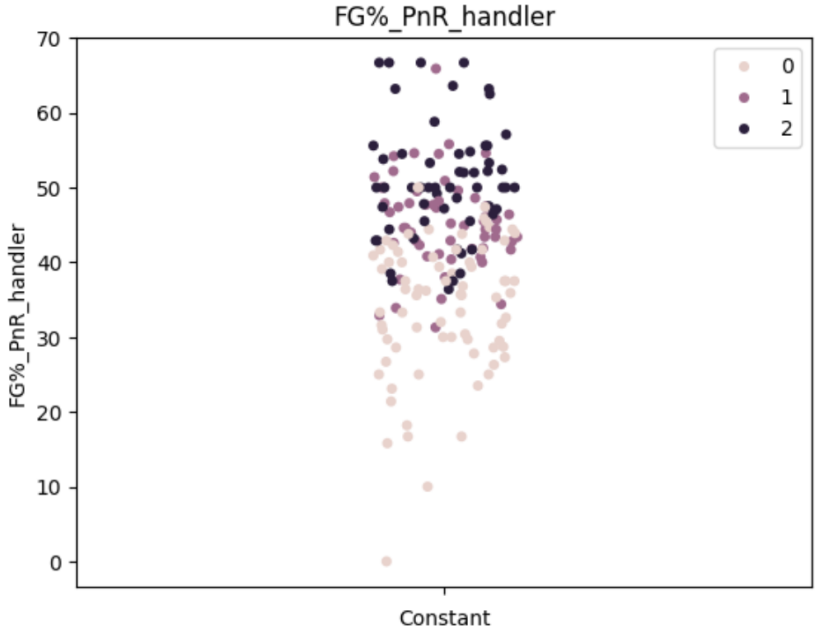
7.4.5 Variance of clusters with FGM



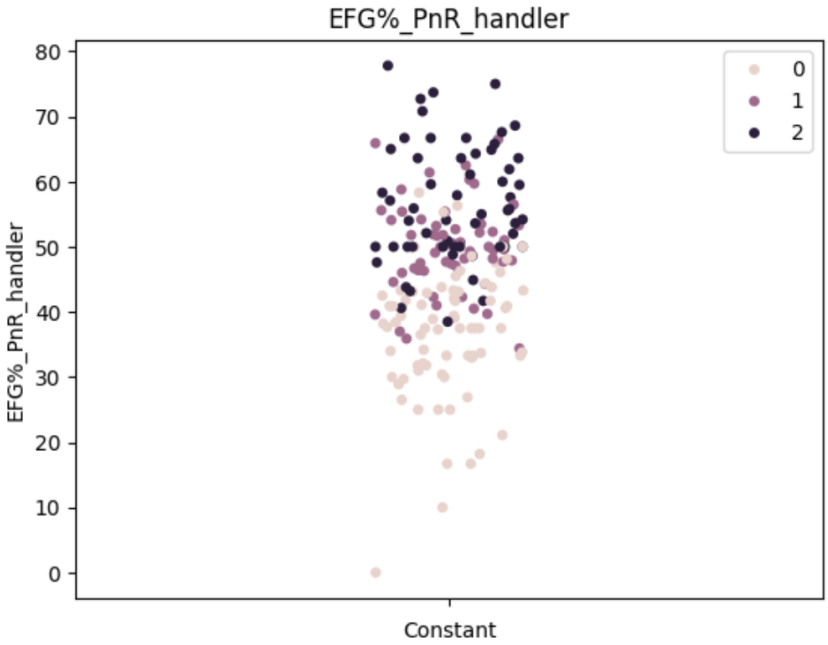
7.4.6 Variance of clusters with FGA



7.4.7 Variance of clusters with FG%

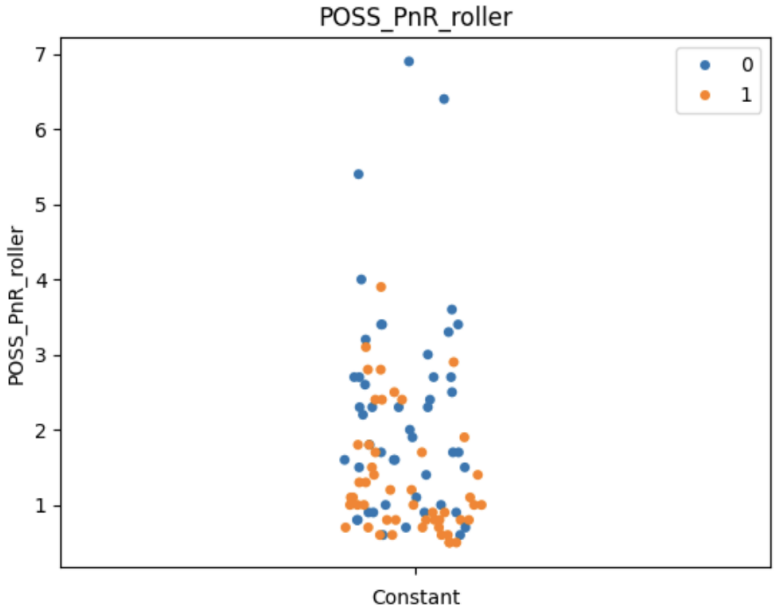


7.4.8 Variance of clusters with EFG%

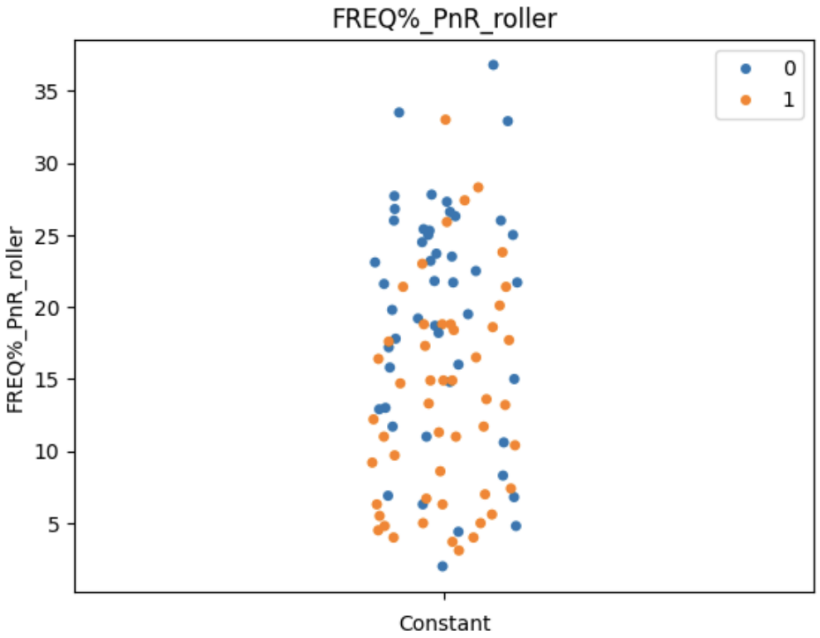


7.5 Pick and roll roller strip plots

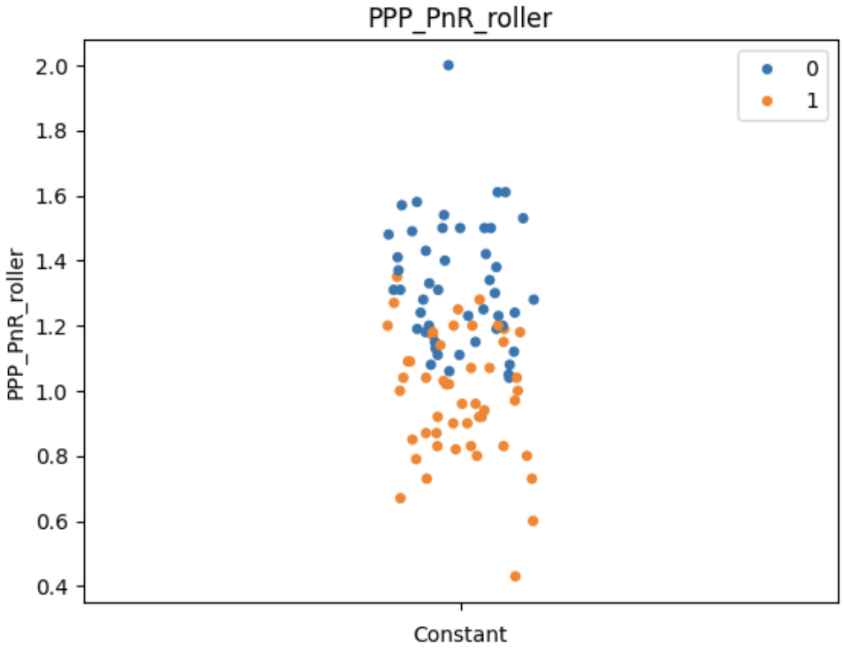
7.5.1 Variance of clusters with POSS



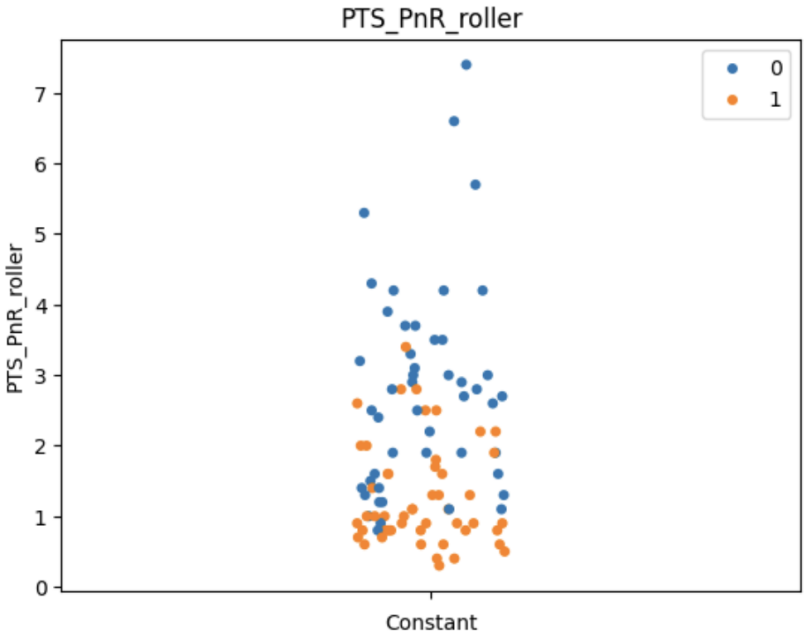
7.5.2 Variance of clusters with FREQ%



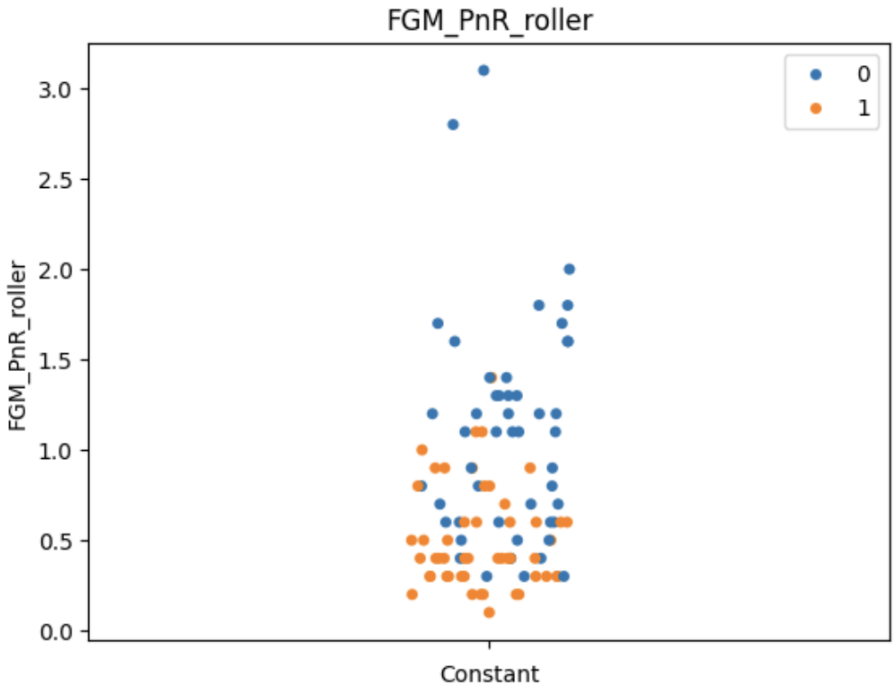
7.5.3 Variance of clusters with PPP



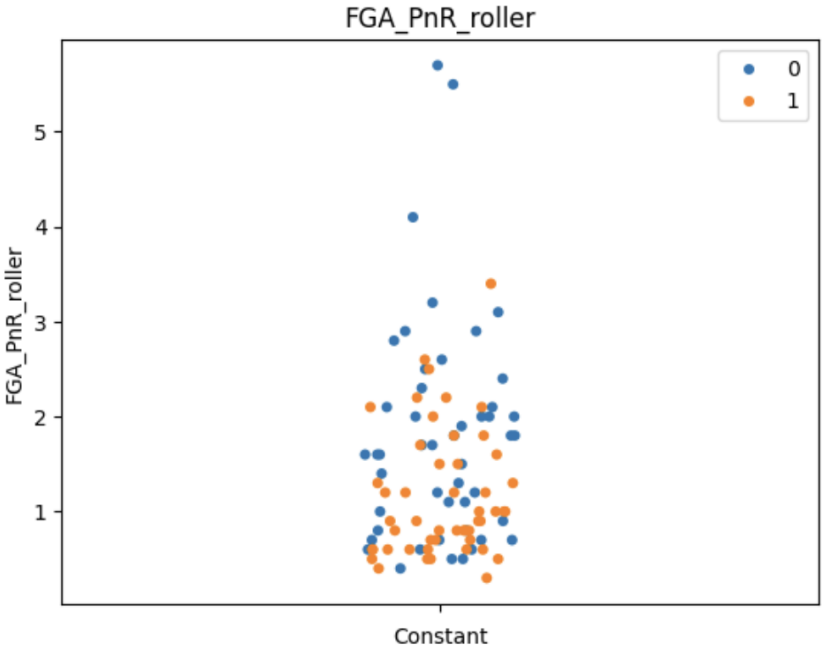
7.5.4 Variance of clusters with PTS



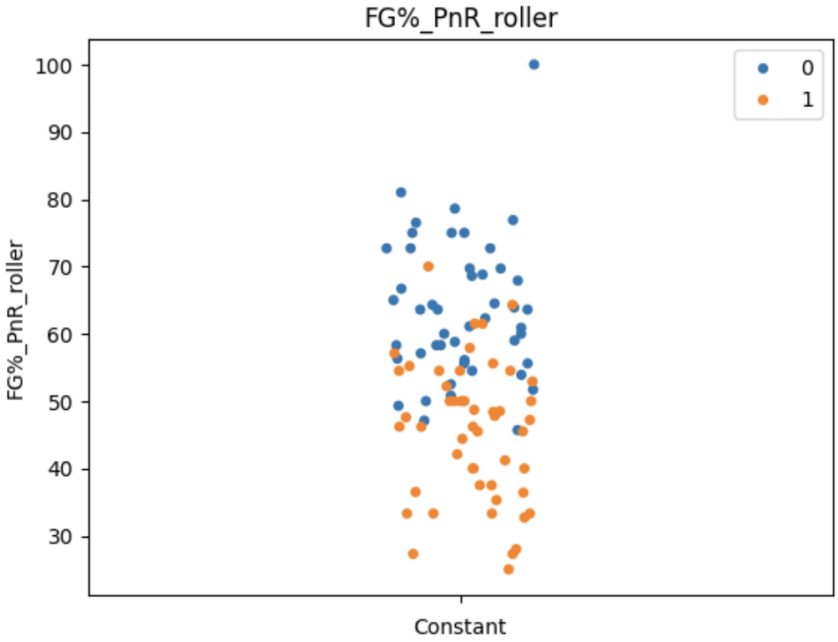
7.5.5 Variance of clusters with FGM



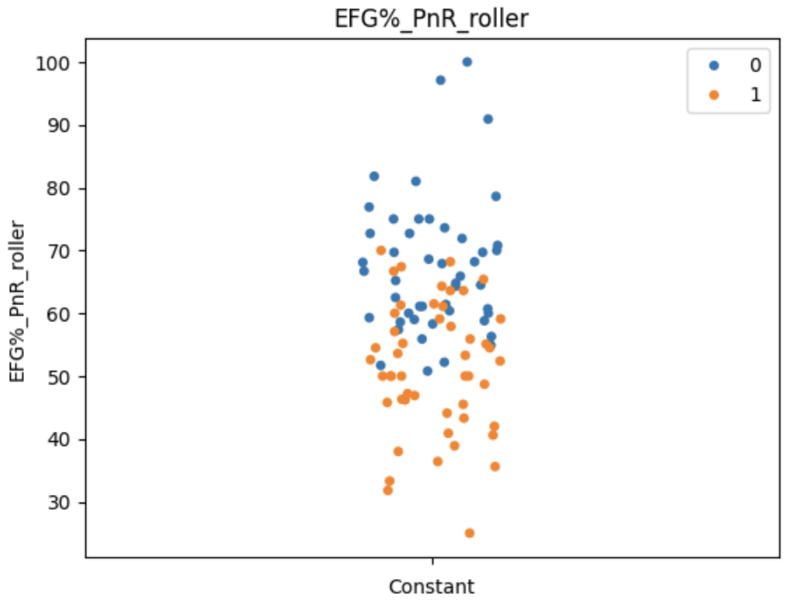
7.5.6 Variance of clusters with FGA



7.5.7 Variance of clusters with FG%

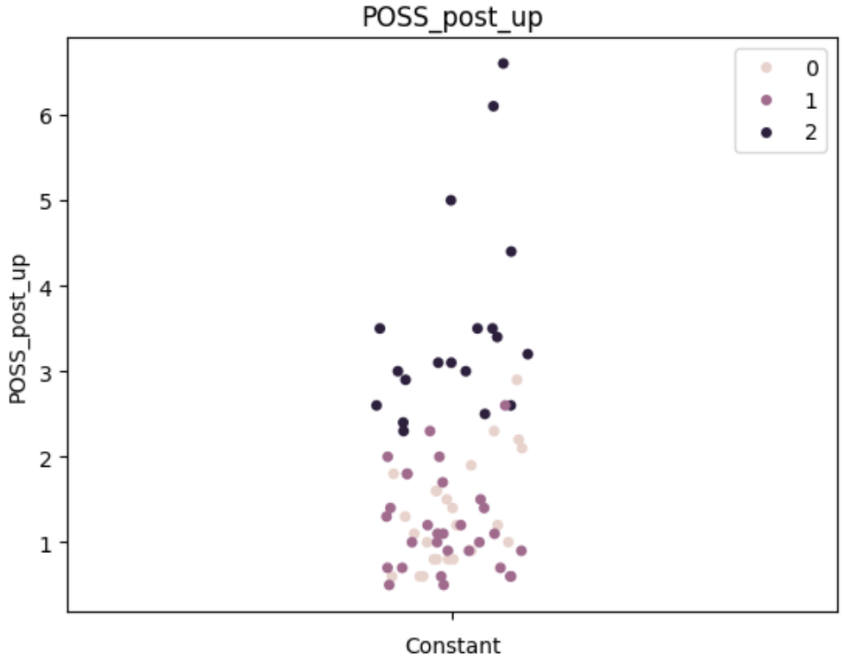


7.5.8 Variance of clusters with EFG%

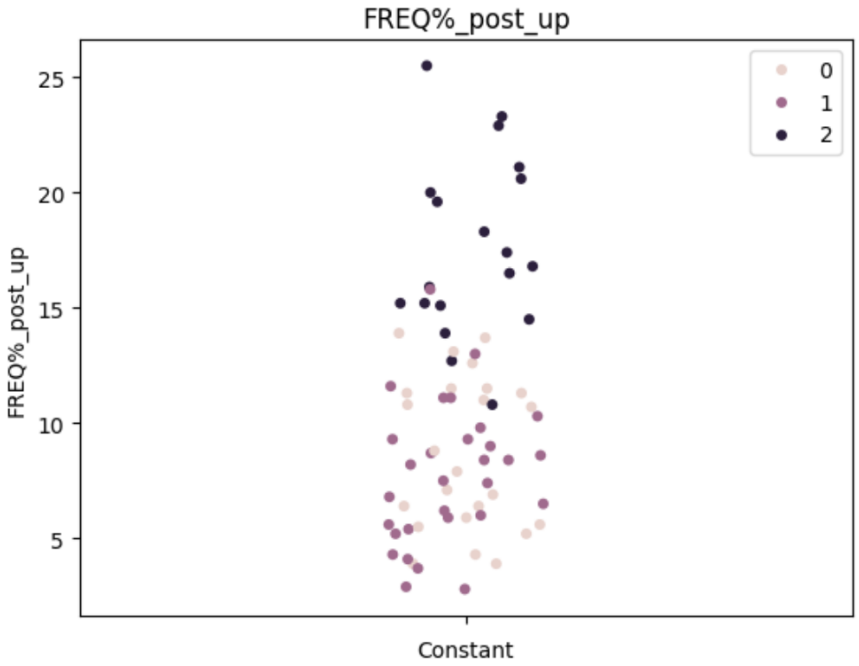


7.6 Post up strip plots

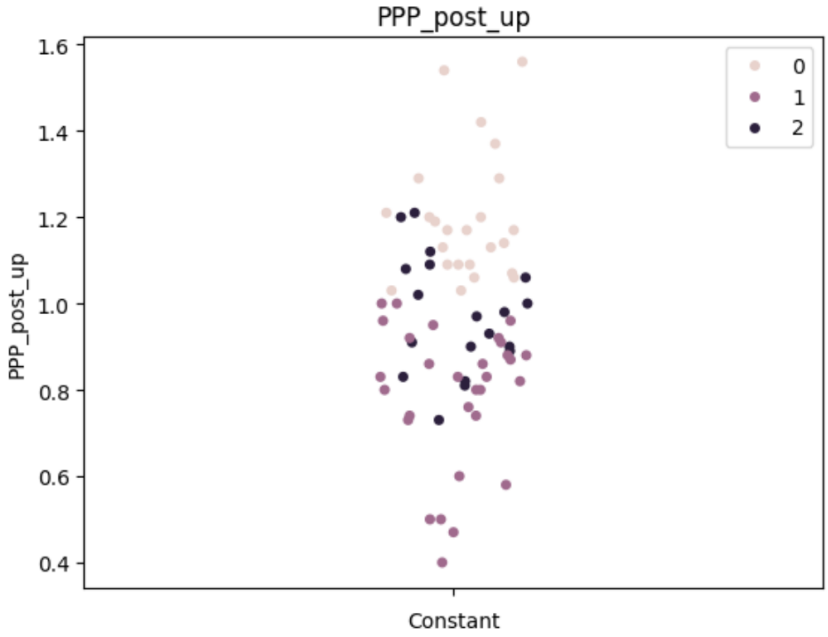
7.6.1 Variance of clusters with POSS



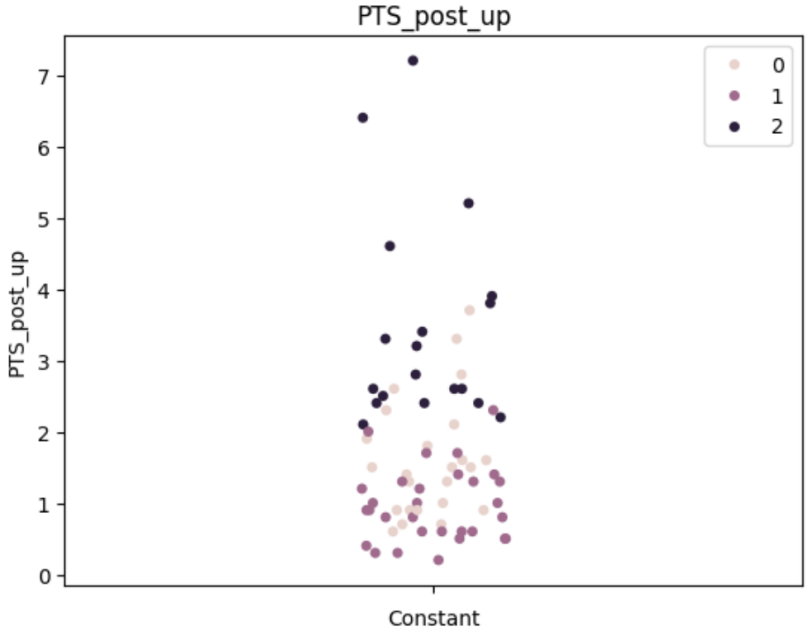
7.6.2 Variance of clusters with FREQ%



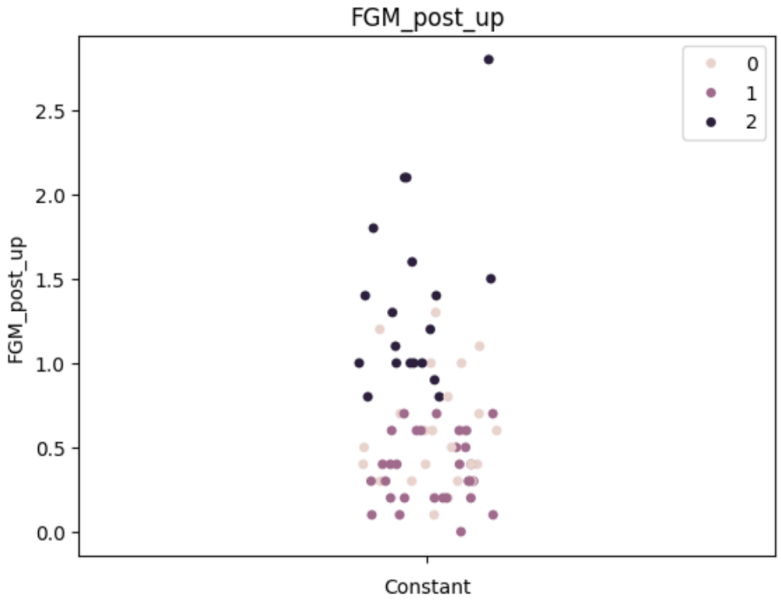
7.6.3 Variance of clusters with PPP



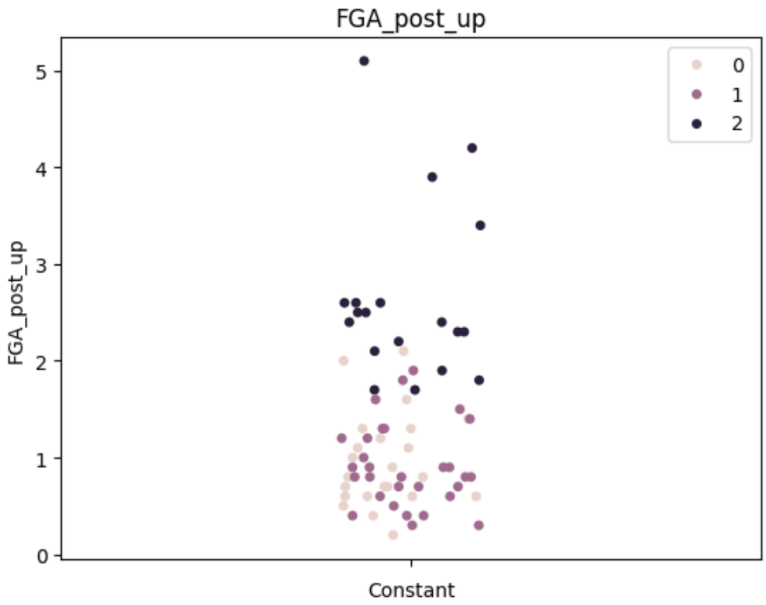
7.6.4 Variance of clusters with PTS



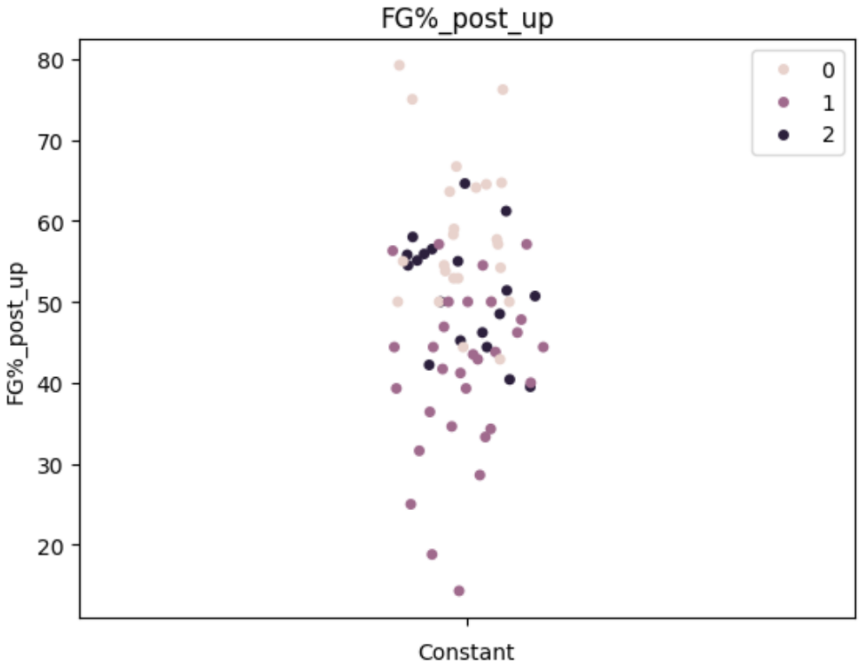
7.6.5 Variance of clusters with FGM



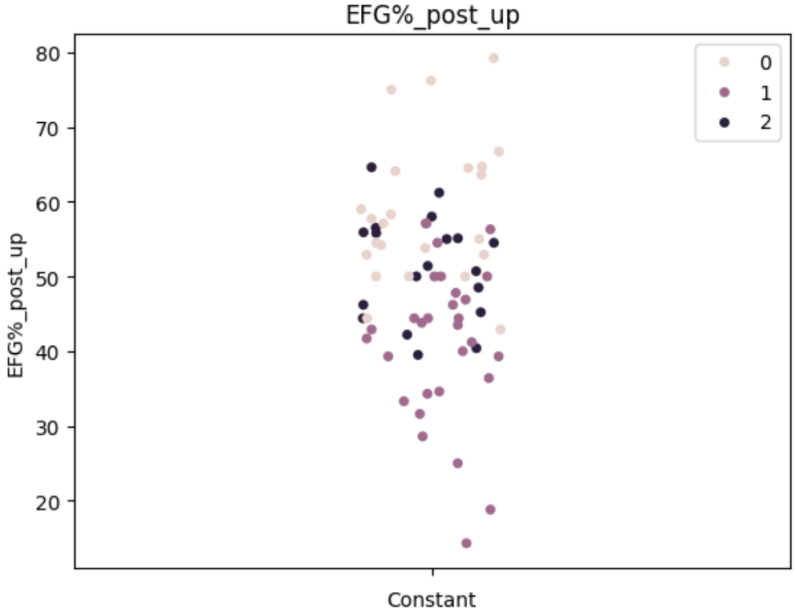
7.6.6 Variance of clusters with FGA



7.6.7 Variance of clusters with FG%

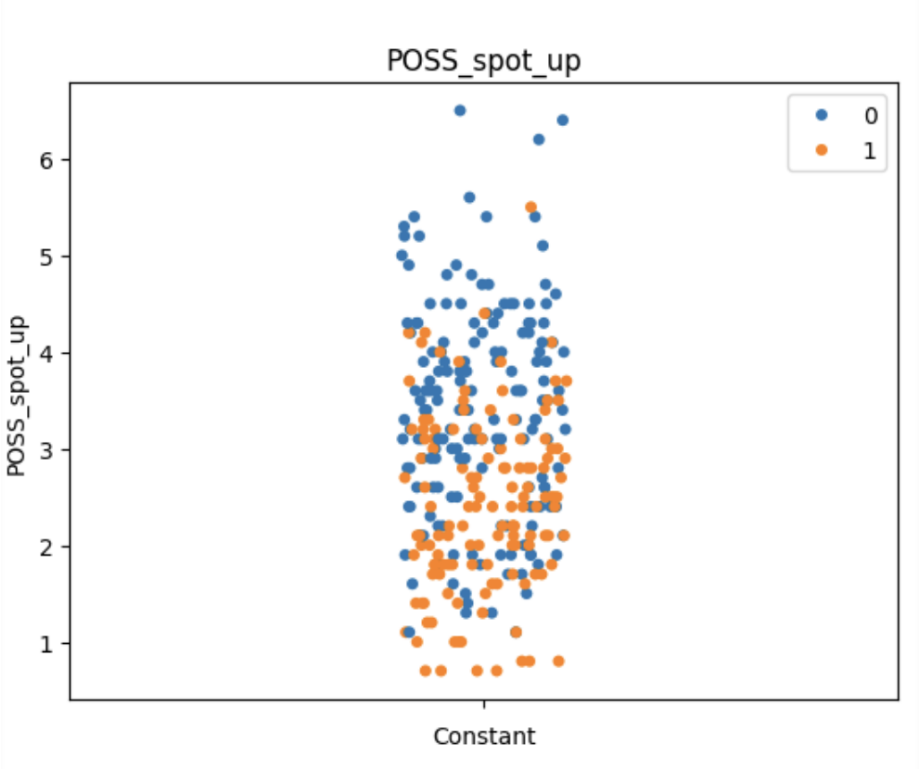


7.6.8 Variance of clusters with EFG%

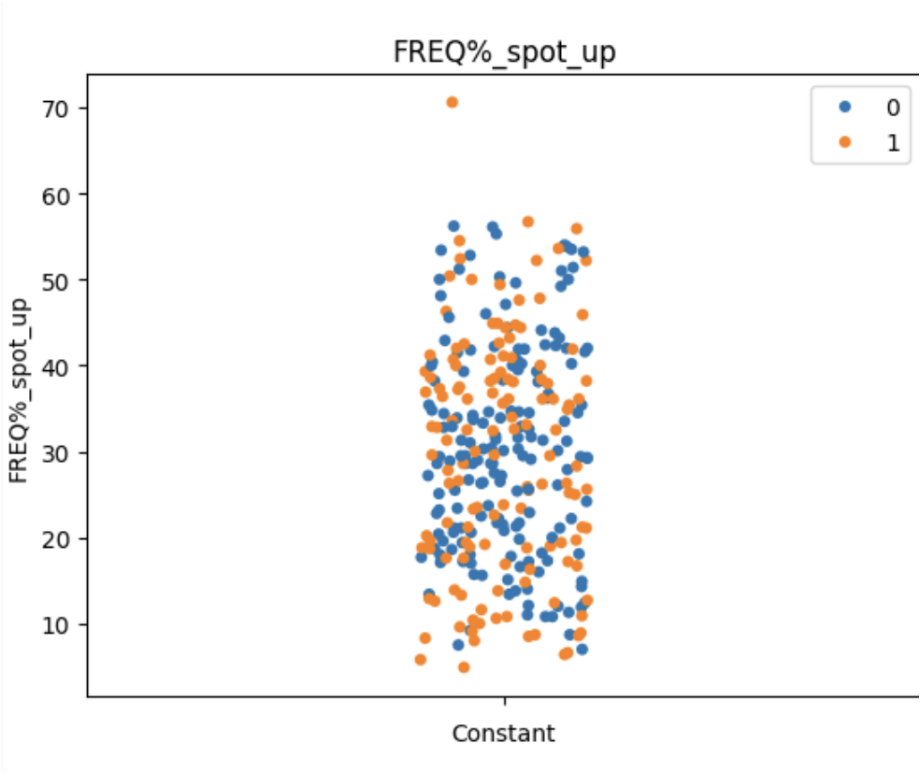


7.7 Spot up strip plots

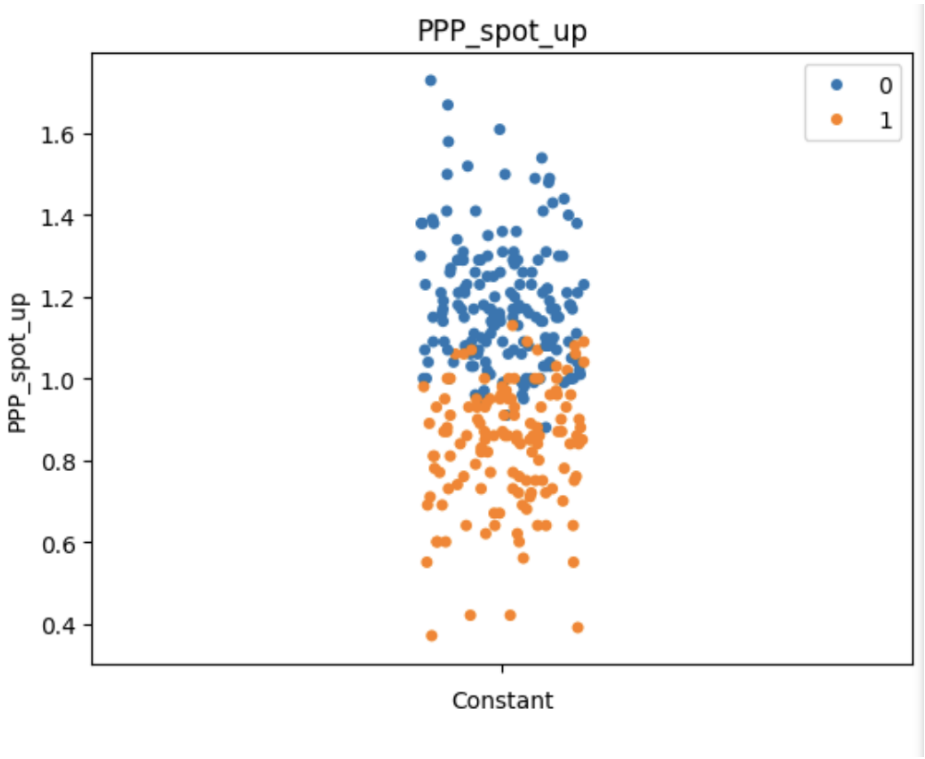
7.7.1 Variance of clusters with POSS



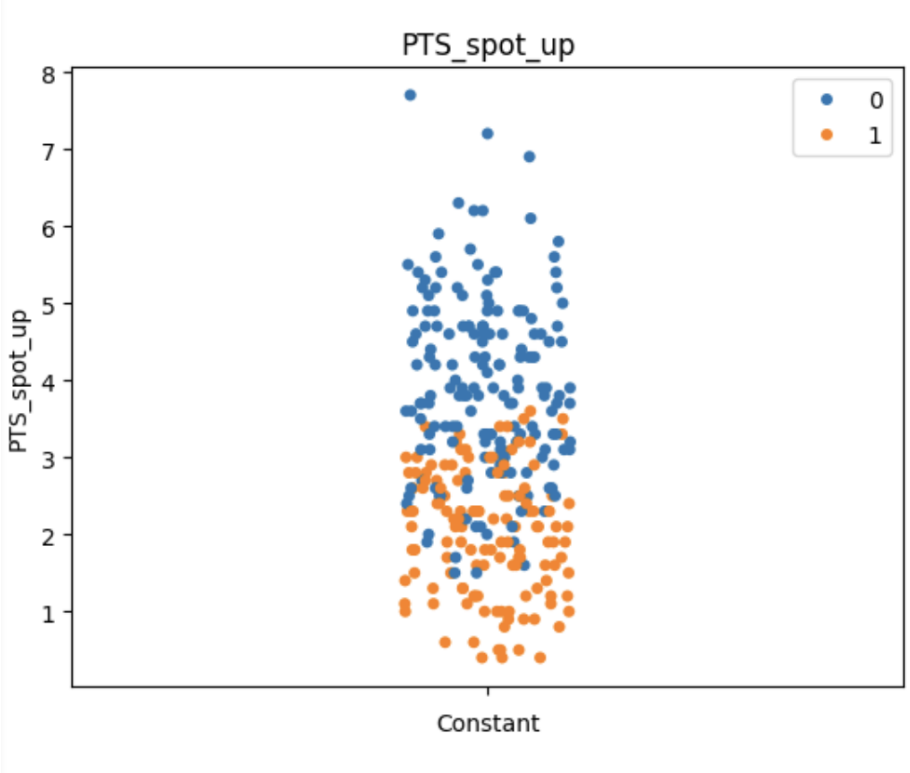
7.7.2 Variance of clusters with FREQ%



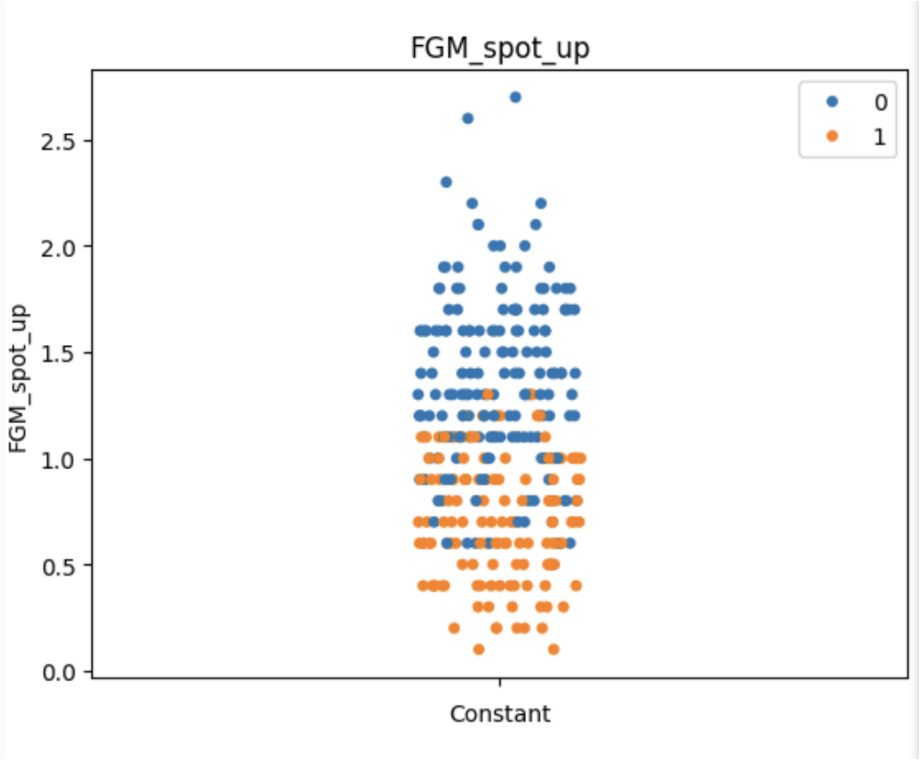
7.7.3 Variance of clusters with PPP



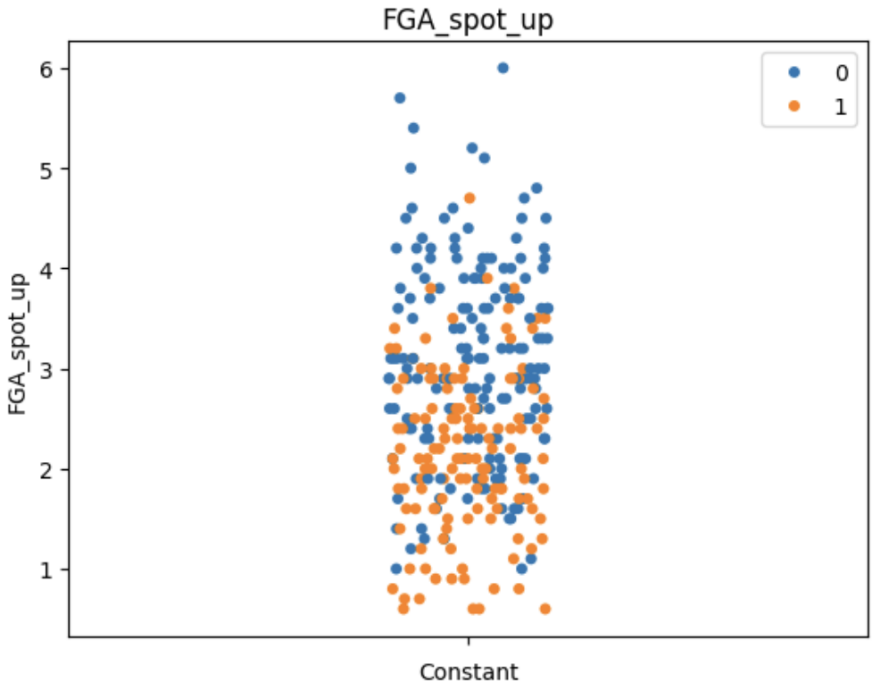
7.7.4 Variance of clusters with PTS



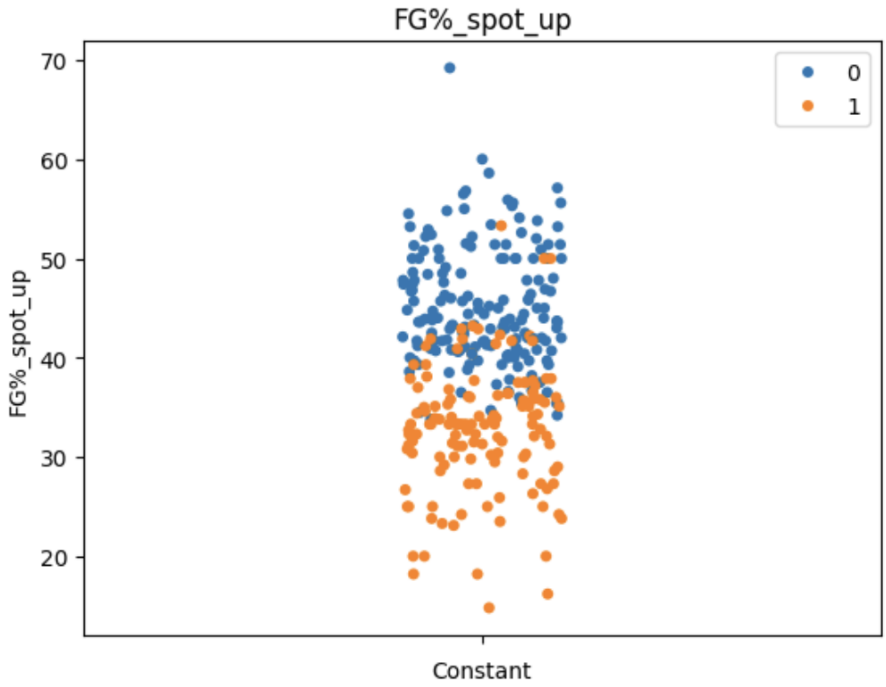
7.7.5 Variance of clusters with FGM



7.7.6 Variance of clusters with FGA



7.7.7 Variance of clusters with FG%



7.7.8 Variance of clusters with EFG%