Clustering players

XSEED

Michele Coaro
Daniele Piazza
Tommaso Premoli

GOAL

Cluster players who wore XSEED shin guards using both game and athletic variables.

Try to predict the role, the athletic condition and the size of the field in which they played.
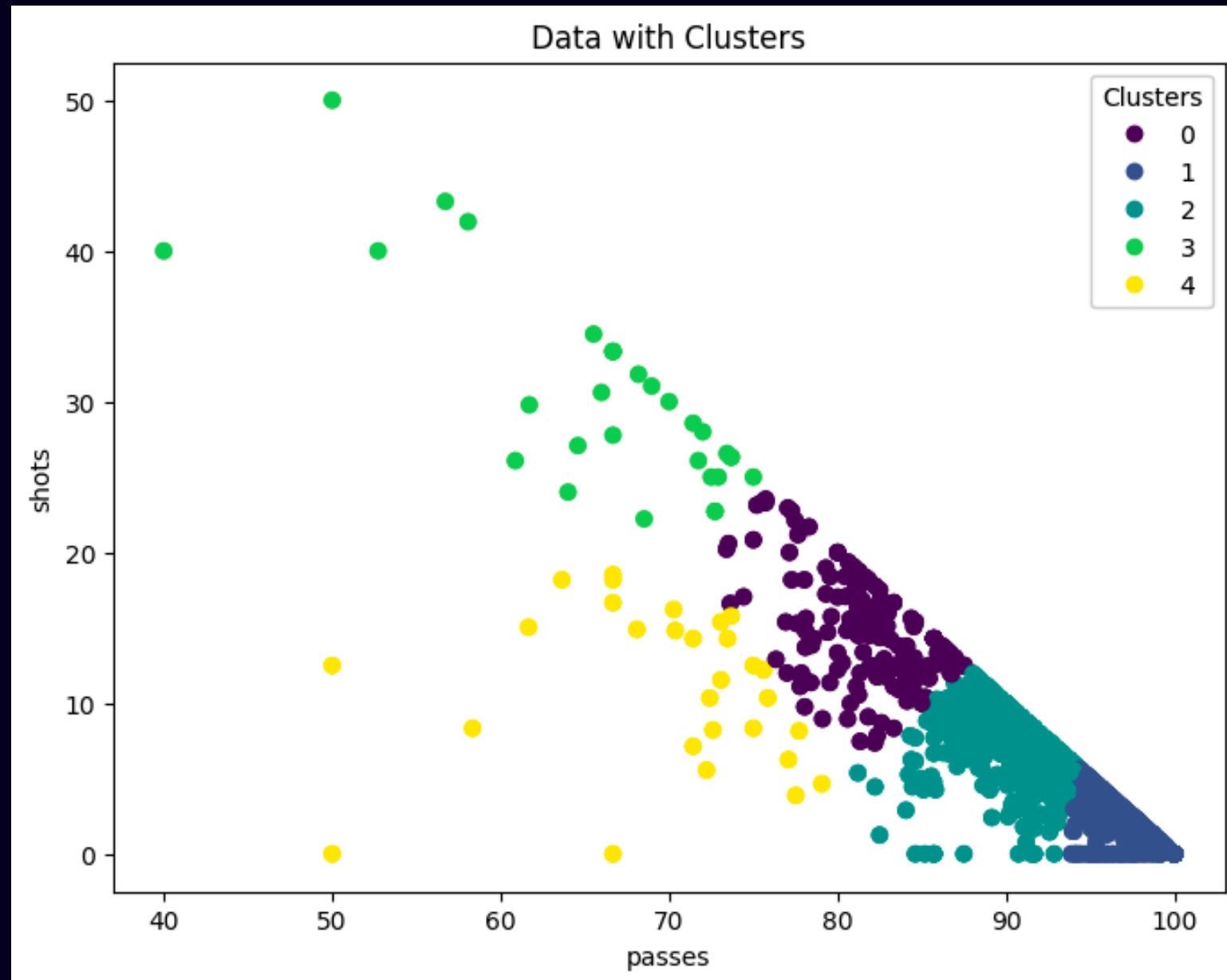
DATASET

Ball

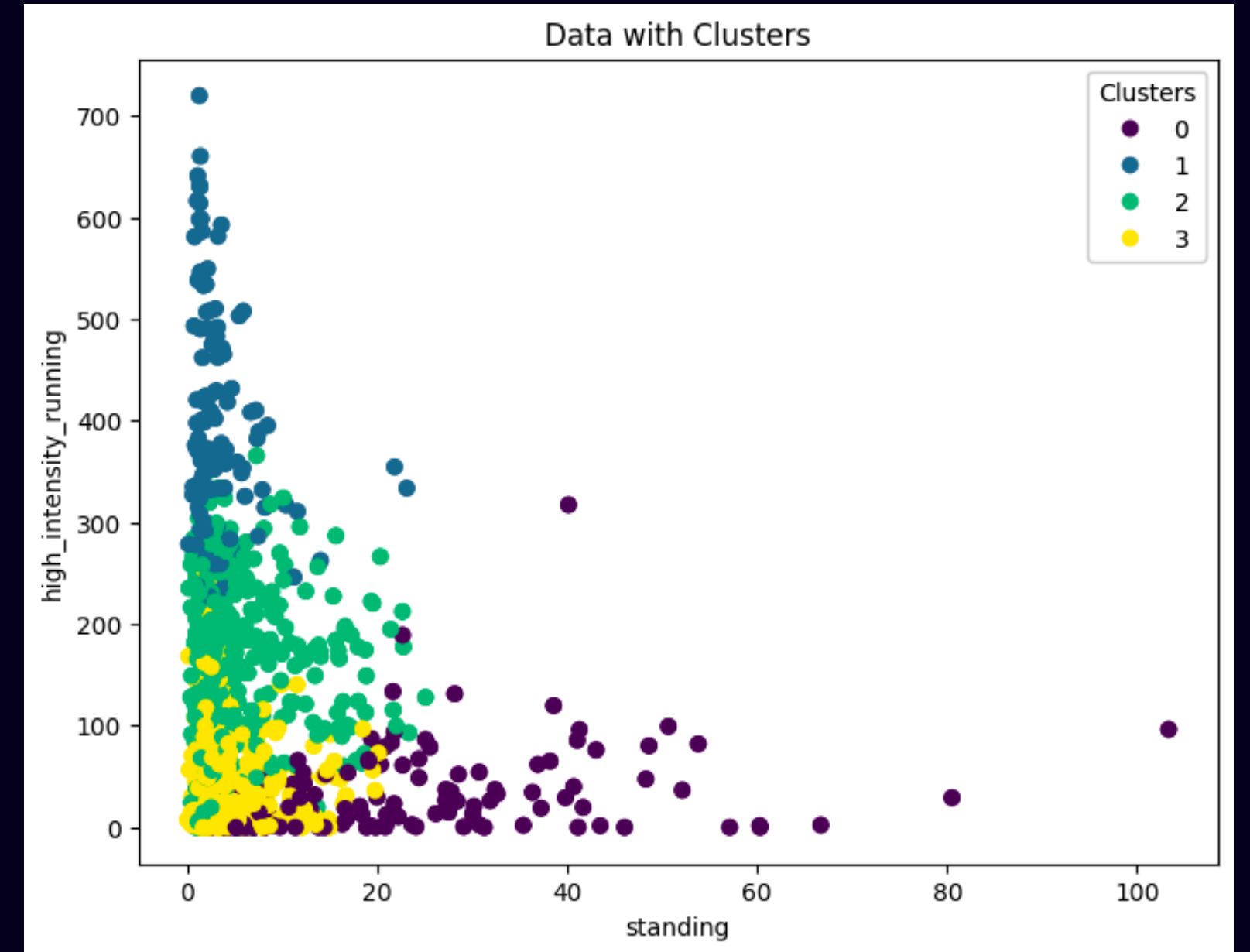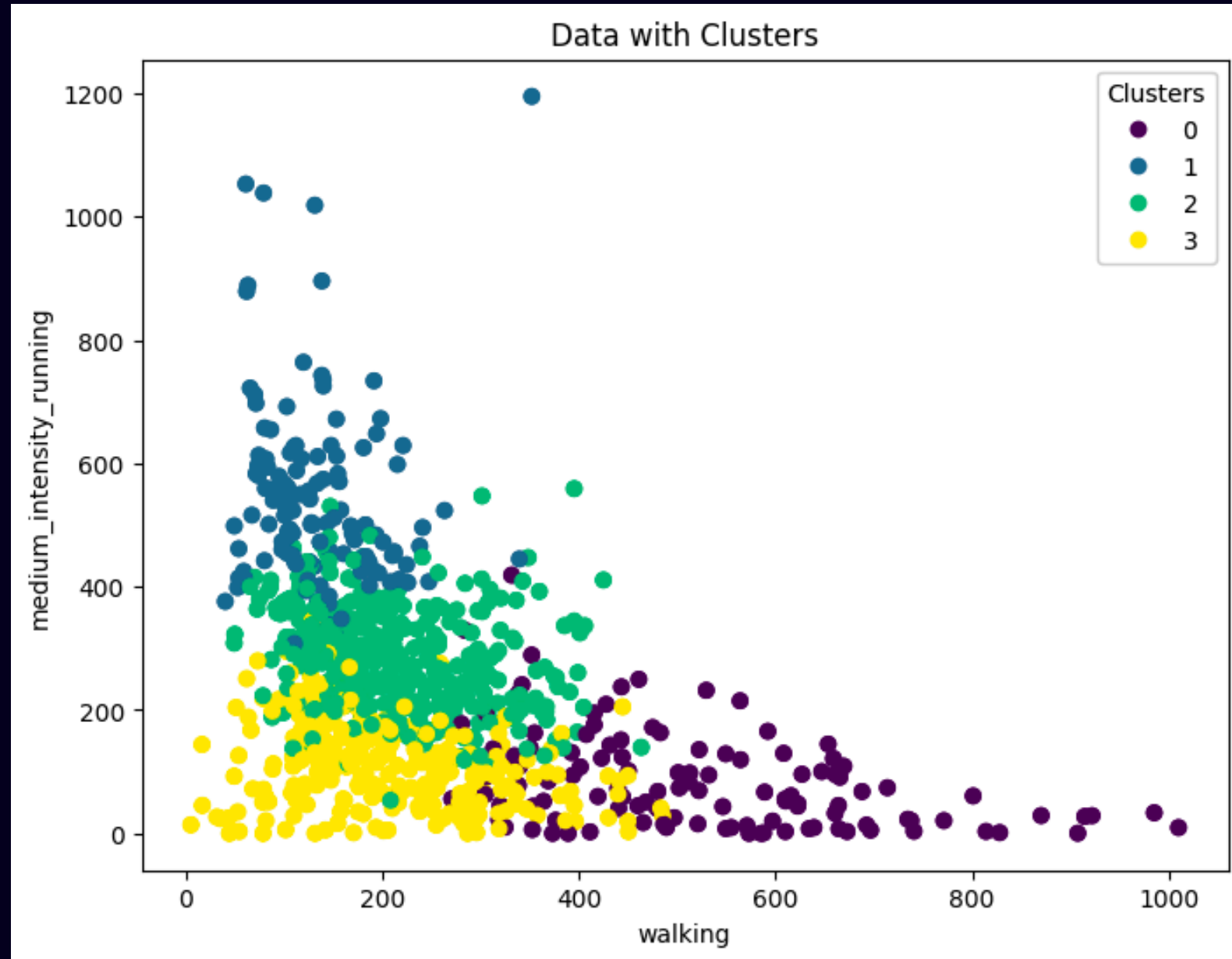|	time_played	touches	passes	shots	crosses	hash
0	540	0.200000	66.666667	27.777778	5.555556	5654103485b27c9f4e923c50ca18e5ab
3	508	0.326772	82.530120	16.867470	0.602410	7c812694b209c87c9f29447feb6c8f56
4	238	0.432773	85.436893	11.650485	2.912621	4d0ec9b39b8a1626ed99c1e79fcb3e77
8	367	0.256131	97.872340	1.063830	1.063830	893a6ce00d5690a7926495272ad4b4f6
9	2328	0.368986	99.534342	0.349243	0.116414	ade4e9b4c35949bf8b0b762254c8600e

Action

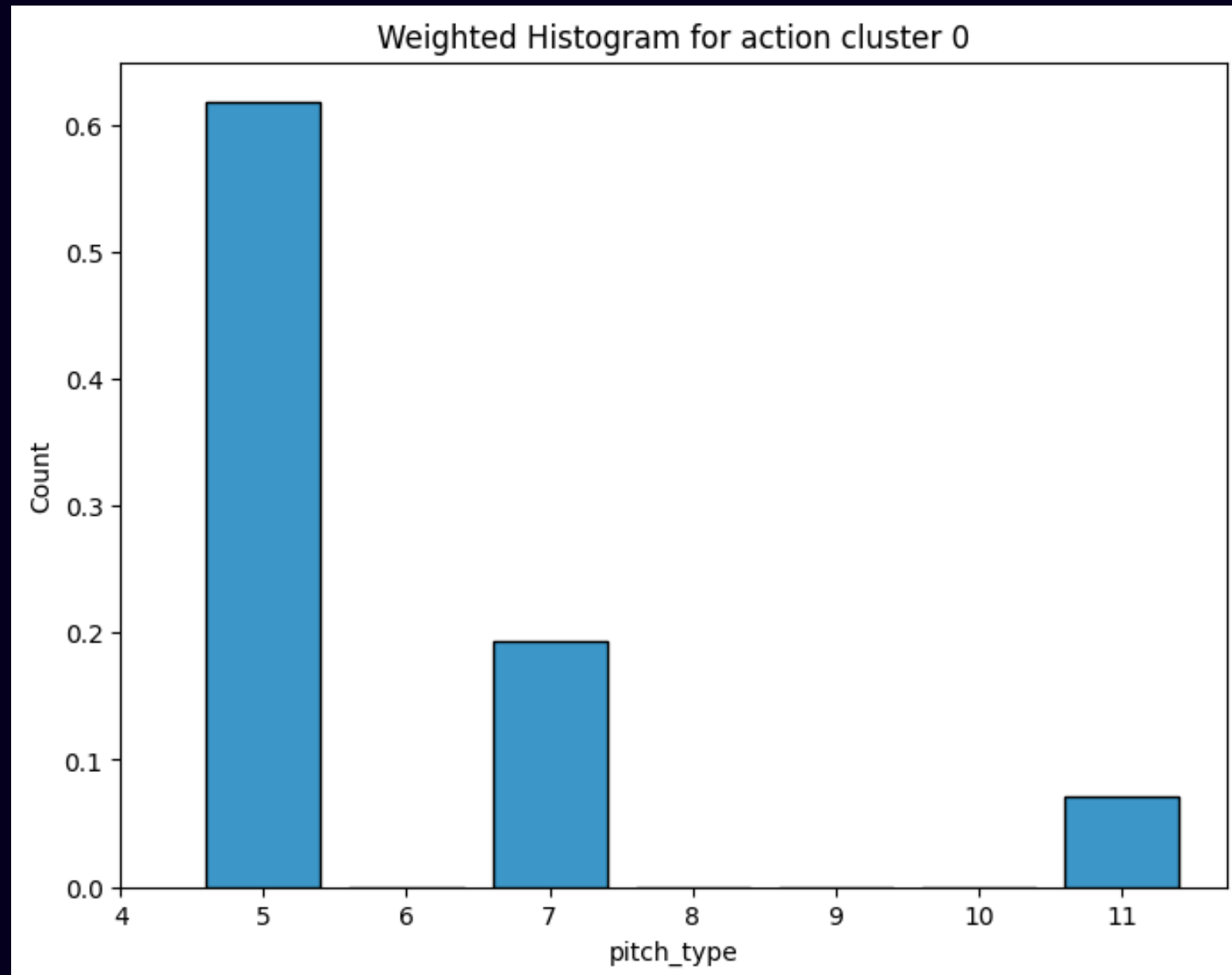|	total_distance	standing	walking	low_intensity_running	medium_intensity_running	high_intensity_running
0	5930.678600	8.551333	135.297333	5331.074600	324.364000	131.391333
3	2513.751720	29.084764	318.748583	2143.423807	22.004528	0.490039
4	3490.971000	1.609160	160.632605	3267.343185	55.801008	5.585042
8	6537.639057	4.163869	119.082180	5812.112191	405.055259	197.225559
9	5427.326363	6.376572	309.596495	4704.429276	255.008222	151.915799

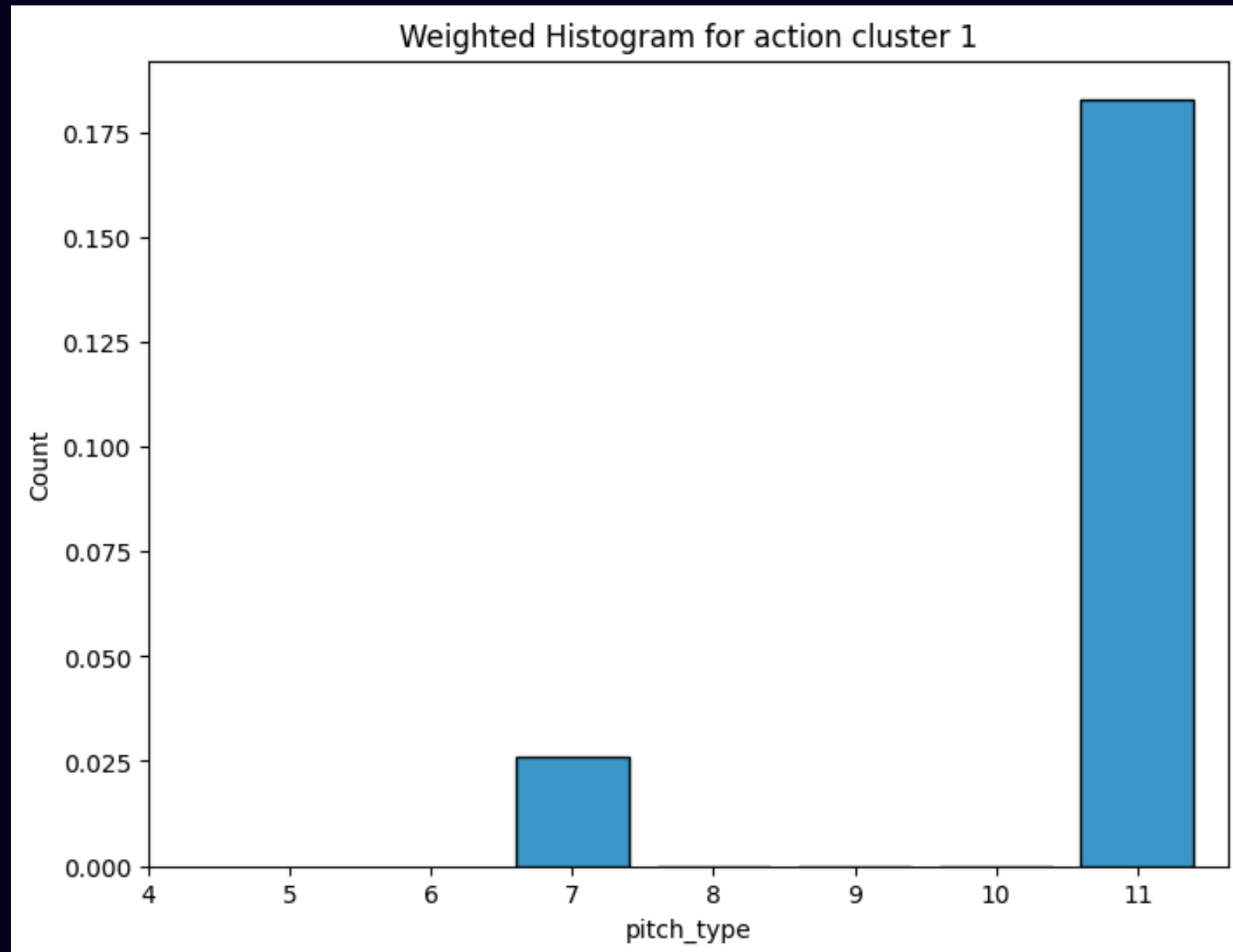# BALL CLUSTERING

ACTION CLUSTERING

MAPPING DATASET

	pitch_type	hash	gender	soccRole	age_group	age
0	11-a-side	5654103485b27c9f4e923c50ca18e5ab	male	FW	38+	52.840520
1	7-9-a-side	60faf5552fd0676c1fb426e450013006	male	NaN	18-37	24.082136
2	7-9-a-side	7e98a17764eef1584f1fec87e7a8182c	male	CM	18-37	33.084189
3	5-a-side	7c812694b209c87c9f29447feb6c8f56	male	FW	<14	8.928131
4	7-9-a-side	4d0ec9b39b8a1626ed99c1e79fcb3e77	male	FW	<14	8.928131

PITCH TYPE

Weighted Histogram for action cluster 0

Cluster 0 walkers all played on 5-a-side pitches
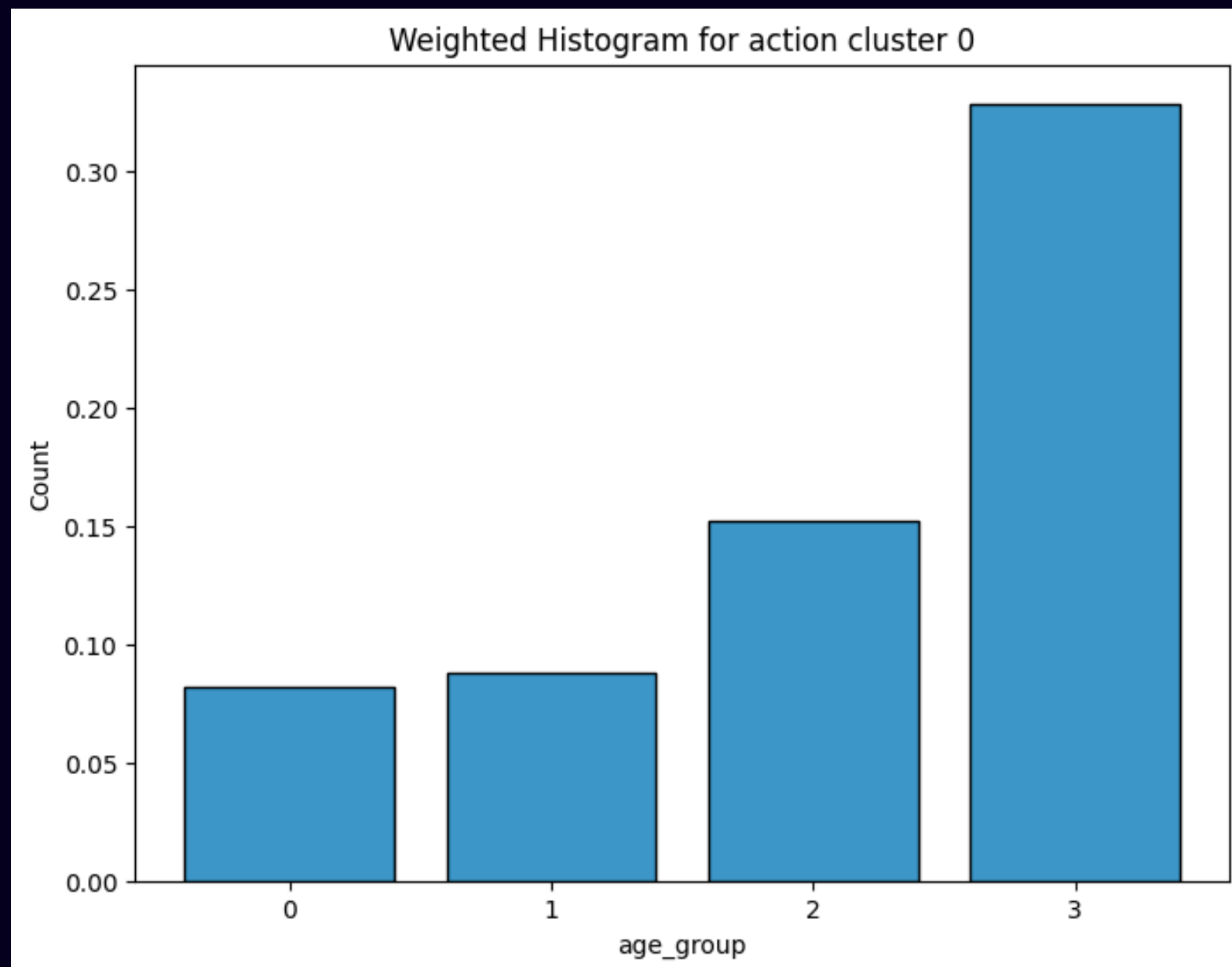
PITCH TYPE

Cluster 1 athletes almost all played on 11-a-side pitches

AGE GROUP

```
age_group = {'<14': 0, '14-17': 1, '18-37': 2, '38+': 3}
```
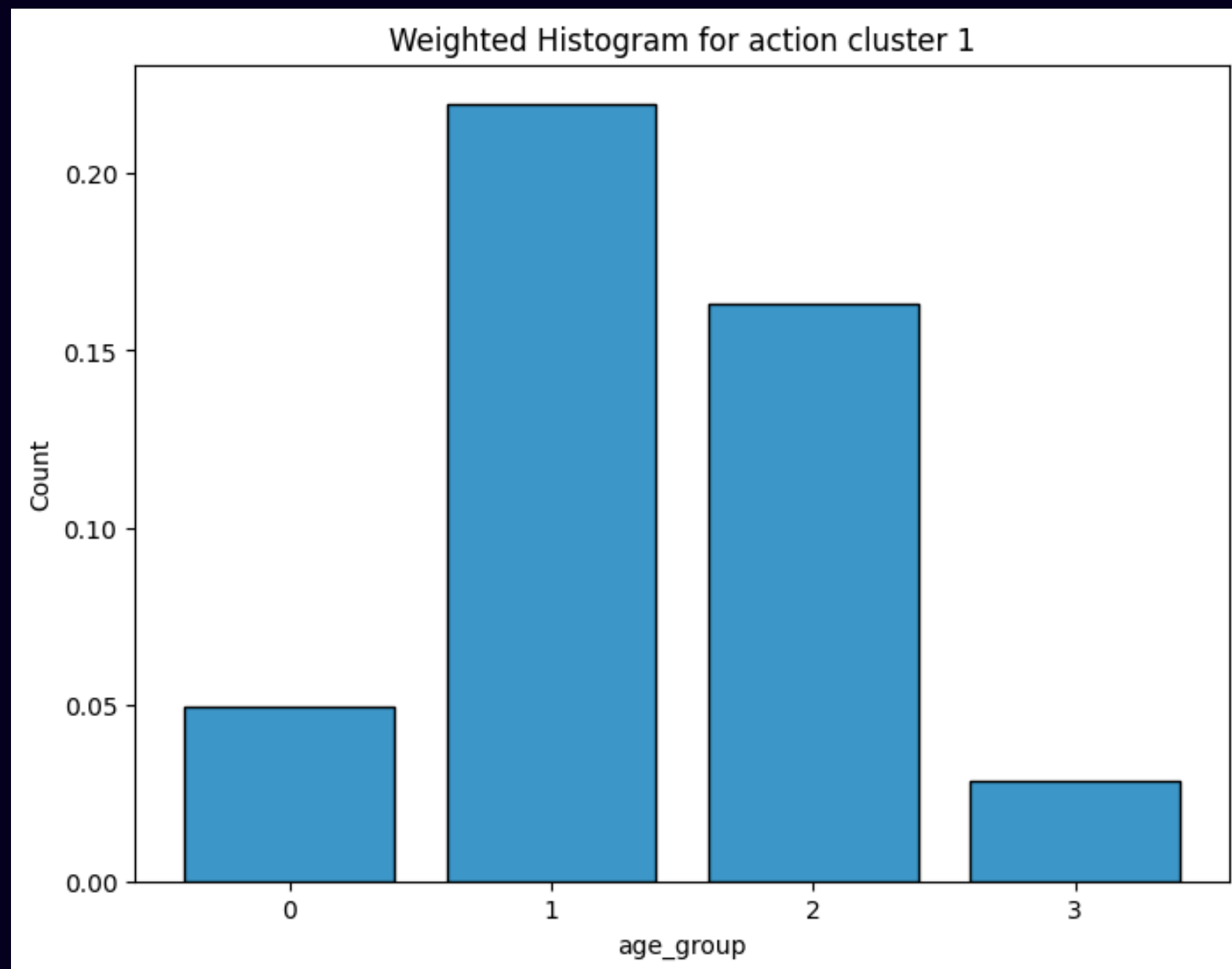

Weighted Histogram for action cluster 0

Cluster 0 of walkers are mostly over 38 years old and, as a consequence, they play more on 5-a-side pitches

AGE GROUP

```python
age_group = {'<14': 0, '14-17': 1, '18-37': 2, '38+': 3}
```


Weighted Histogram for action cluster 1

Athletes in cluster 1 are between 14 and 37 years old

PITCH DIVISION

According to the article *"Football and Futsal: In which do you run more?"* by *Oliver Sports,* in an 11-a-side field you reach an average of 31 and 35 km/h, while in a 5-a-side field you reach between 26 and 27 km/h.

```
pitch_5 = norm[norm['max_speed'] < 22]
pitch_7 = norm[(norm['max_speed'] >= 22) & (norm['max_speed'] < 27)]
pitch_11 = norm[norm['max_speed'] >= 27]
```

```
pitch_type
11-a-side      587
7-9-a-side      82
5-a-side         5
```
11-a-side field

```
pitch_type
5-a-side        26
7-9-a-side      20
11-a-side       11
```
5-a-side field

```
pitch_type
11-a-side       91
7-9-a-side      90
5-a-side        24
```
7-a-side field