

NBA Player Types and Optimal Lineup Composition

Mark Corey
github.com/MarkDCorey/nba_player_clustering

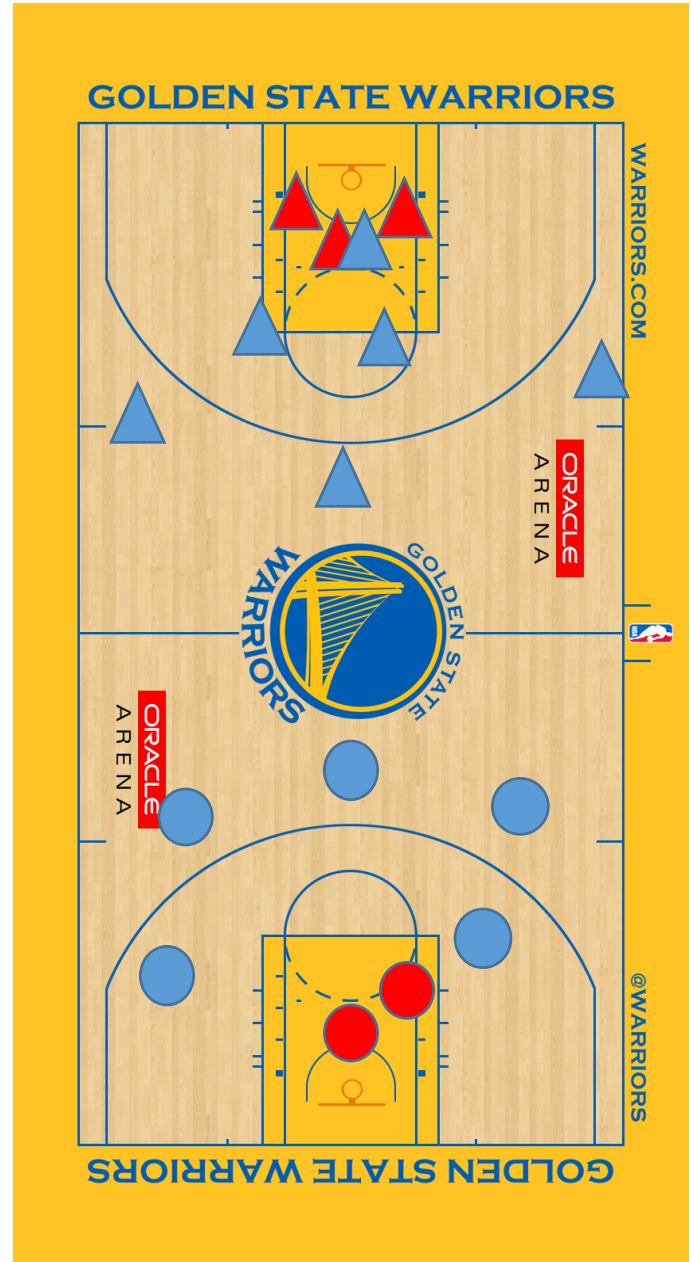


Player Data Example

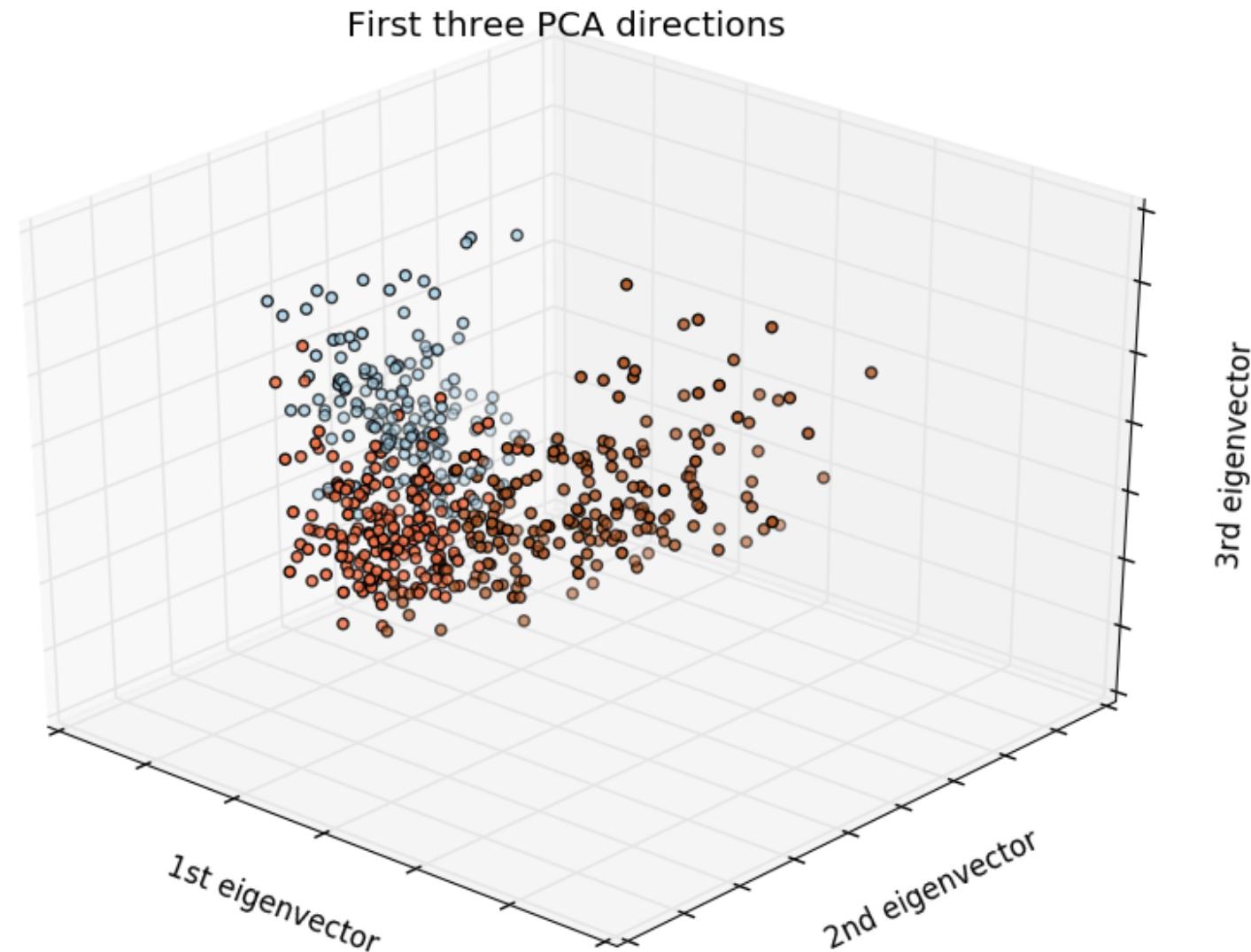
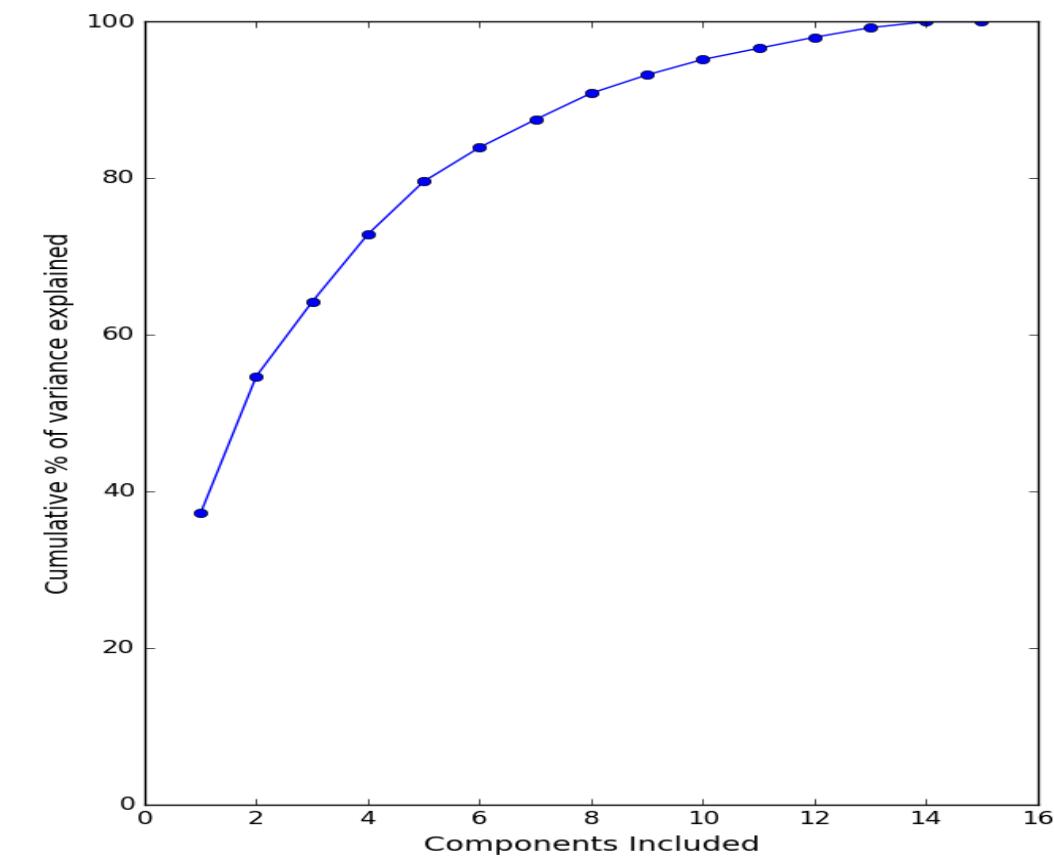
```
blk_pos : 0.012
dreb_pos : 0.204
oreb_pos : 0.04
stl_pos : 0.097
attempt_RA_pos : 0.054
attempt_paint_pos : 0.022
attempt_corner_3_pos : 0.017
attempt_non_corner_3_pos : 0.11
attempt_mid_pos : 0.051
ast_shot_pct : 0.418
ast_pos : 0.105
fta_pos : 0.065
d_fga_paint_pct : 0.18
d_fga_threes_pct : 0.365
d_fga_mid_pct : 0.455
```



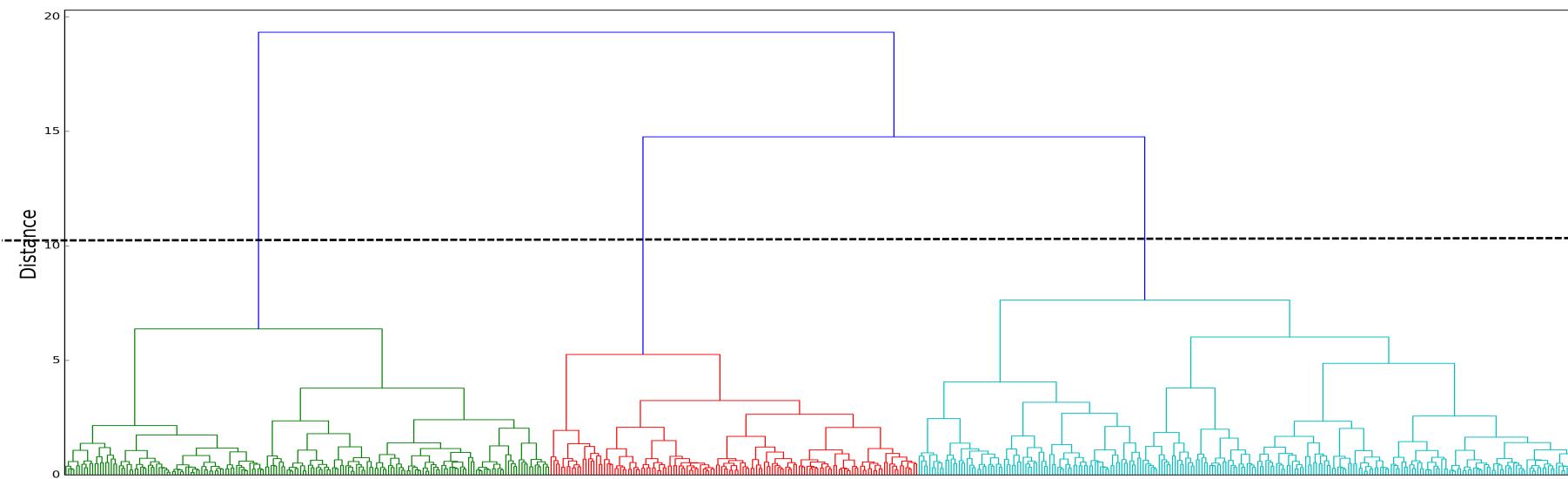
```
blk_pos : 0.166
dreb_pos : 0.296
oreb_pos : 0.201
stl_pos : 0.017
attempt_RA_pos : 0.159
attempt_paint_pos : 0.042
attempt_corner_3_pos : 0.0
attempt_non_corner_3_pos : 0.001
attempt_mid_pos : 0.031
ast_shot_pct : 0.753
ast_pos : 0.008
fta_pos : 0.07
d_fga_paint_pct : 0.509
d_fga_threes_pct : 0.102
d_fga_mid_pct : 0.389
```



Player Data: Feature Reduction and First Clustering Attempt



Hierarchical Structure?



Under-the-Basket Players (177)

feature	std_from_avg
d_fga_paint_pct	1.215297
oreb_pos	1.106488
blk_pos	0.950951
dreb_pos	0.858543
attempt_RA_pos	0.754931
attempt_paint_pos	0.531102
ast_shot_pct	0.409722
fta_pos	0.386286
d_fga_mid_pct	0.251215
attempt_mid_pos	-0.025522
stl_pos	-0.199645
ast_pos	-0.600870
attempt_non_corner_3_pos	-0.919328
attempt_corner_3_pos	-0.925562
d_fga_threes_pct	-1.266150

Perimeter shooters/defenders (187)

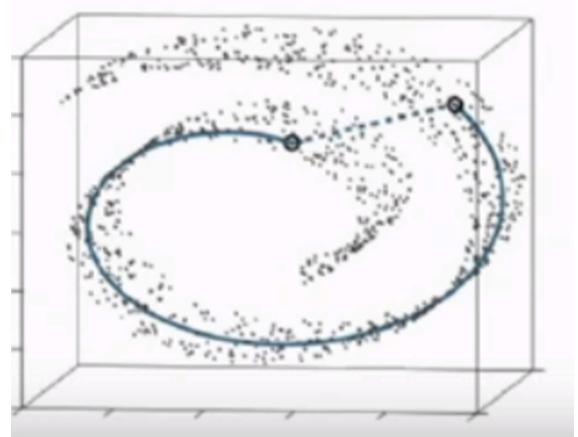
feature	std_from_avg
attempt_corner_3_pos	0.826298
ast_shot_pct	0.704936
attempt_non_corner_3_pos	0.419247
d_fga_threes_pct	0.396999
d_fga_paint_pct	-0.263190
stl_pos	-0.269688
attempt_mid_pos	-0.291332
d_fga_mid_pct	-0.336092
blk_pos	-0.382645
ast_pos	-0.438954
dreb_pos	-0.462606
oreb_pos	-0.486201
fta_pos	-0.504891
attempt_RA_pos	-0.554194
attempt_paint_pos	-0.602945

Distributors and shot creators (203)

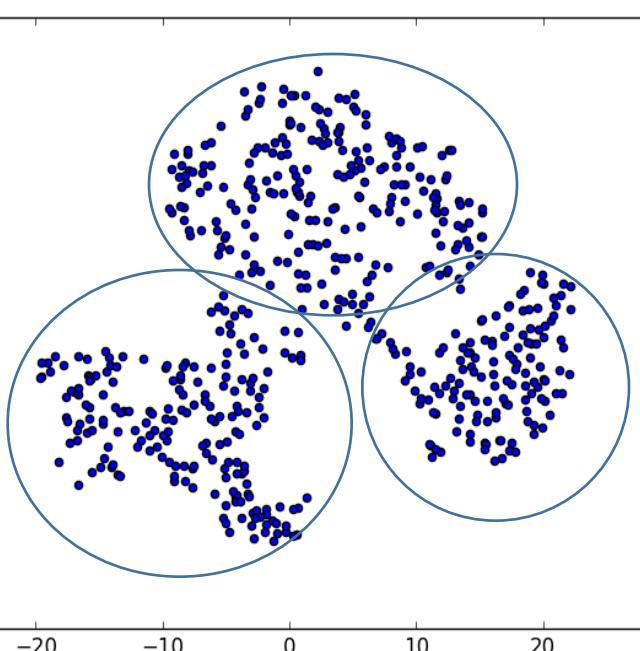
feature	std_from_avg
ast_pos	0.928268
d_fga_threes_pct	0.738274
stl_pos	0.422507
attempt_non_corner_3_pos	0.415379
attempt_mid_pos	0.290623
fta_pos	0.128286
attempt_paint_pos	0.092343
d_fga_mid_pct	0.090562
attempt_corner_3_pos	0.045846
attempt_RA_pos	-0.147727
dreb_pos	-0.322438
blk_pos	-0.476669
oreb_pos	-0.516891
d_fga_paint_pct	-0.817197
ast_shot_pct	-1.006620

t-SNE

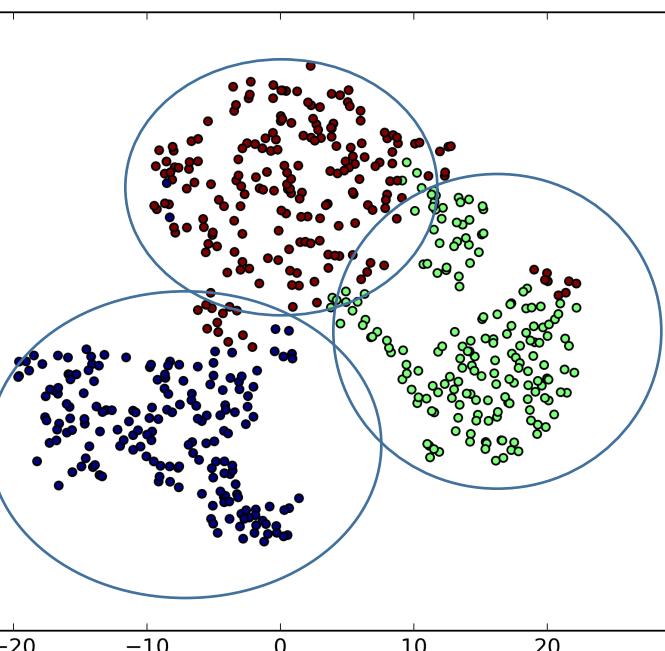
t-SNE example



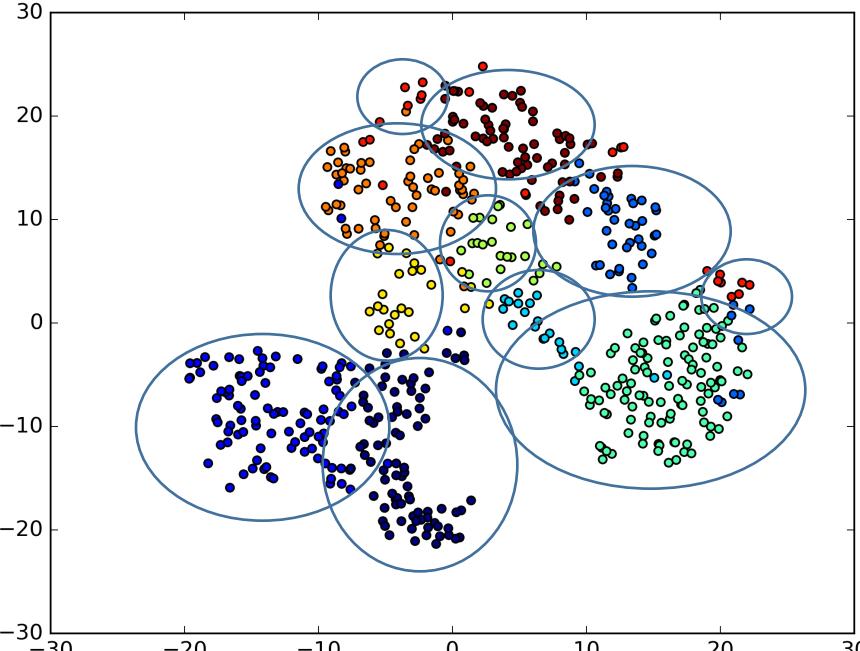
Full data set reduced to two dimensions



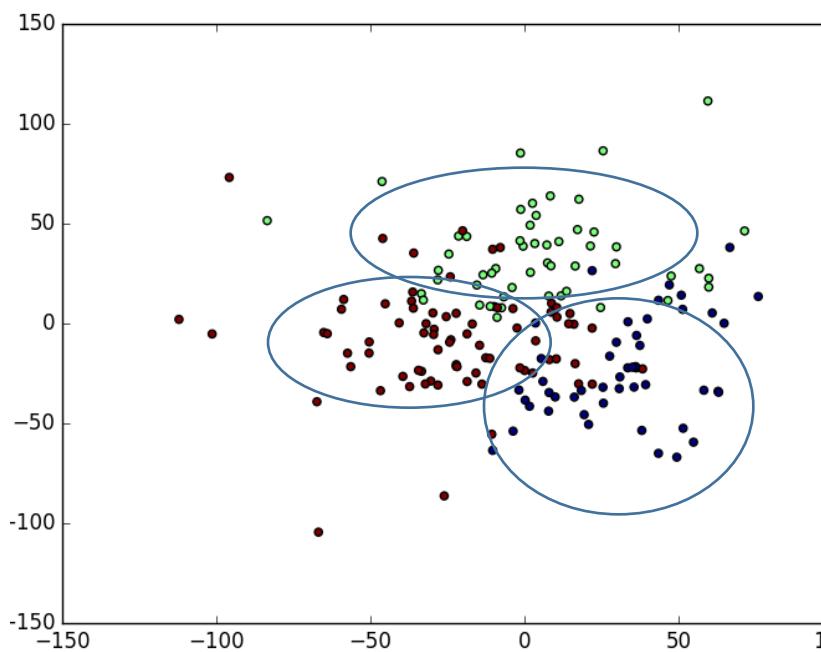
3 Hierarchical Clusters Overlaid



10 Hierarchical Clusters Overlaid

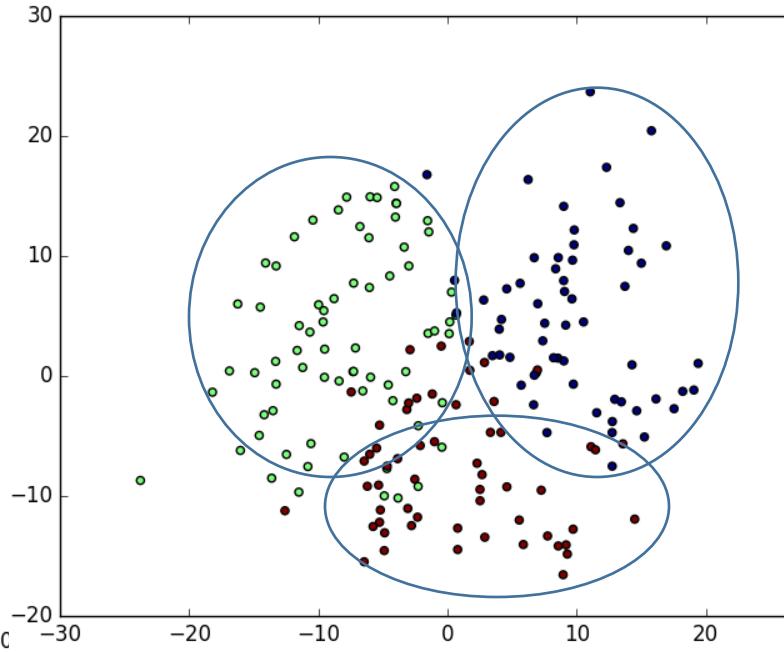


Final Clustering Step



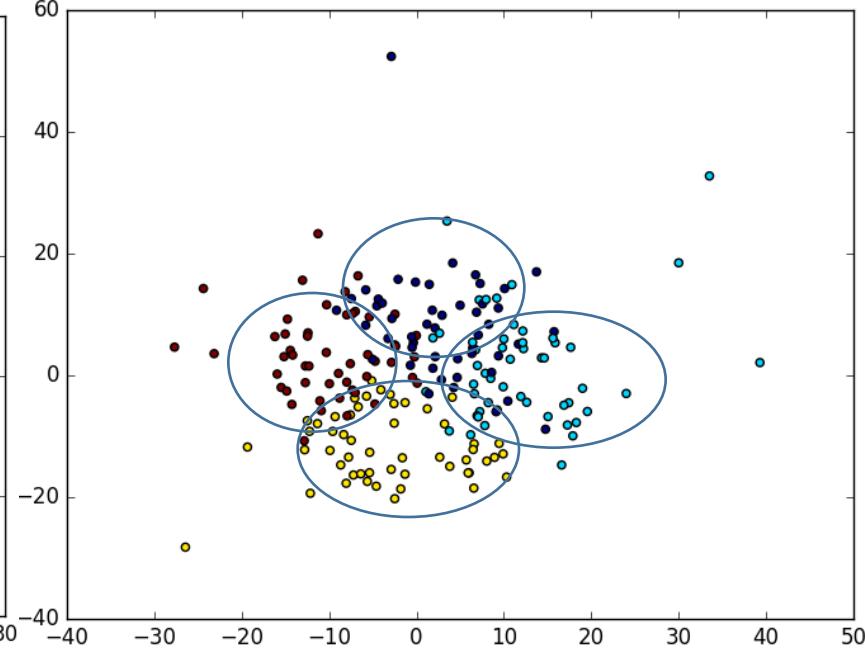
Under-the-Basket Players (177)

- 0)** Rim protectors, limited offense versatility (47)
- 1)** Strong rebounders, versatile offensive skill set (53)
- 2)** Dominant rebounders, shot blockers (77)



Perimeter shooters/defenders (187)

- 3)** Perimeter only defenders, 3pt shooters (59)
- 4)** 3pt shooters, volume catch-and-shoot scoring, versatile defenders (71)
- 5)** Versatile interior and perimeter defenders, 3pt shooters (57)



Distributors and shot creators (203)

- 6)** Ball dominant distributors and shooters (49)
- 7)** Secondary distributors, perimeter defenders, versatile scorers (49)
- 8)** Perimeter defenders, mid-range and three shooters, distributors (54)
- 9)** Pure creators/distributors (51)

Lineup Example

lineup: Curry,Stephen - Durant,Kevin - Green,Draymond - Iguodala,Andre - Thompson,Klay

clusters: 6,7,2,4,4

points scored: 206.7

points allowed: 157.3

minutes played: 67.6

net per min: 0.7308



3 cluster combo performance results

Top cluster combo: 4,5,6

- 4 = 3pt shooters, volume catch-and-shoot scoring, versatile defenders
- 5 = Versatile interior and perimeter defenders, 3pt shooters
- 6 = Ball dominant distributors and shooters

Critical player types

- 4s appear in 10 of the top 20 combos
- 5s appear in 9 of the top combos
- 9s appear in 8 of the top combos
- Best guess: shot creation and three point shooting are the highest value skill sets

Underrepresented player types

- 3s only appear in 2 of the top combos
- 0s only appear in 3 of the top combos
- Best guess: 3s have limited defensive versatility (rarely leave the perimeter) and 0s have limited offensive versatility (rarely leave paint)

c3_combo	net_per_min	observations	p_val	total	min_combo	z_score
[4, 5, 6]	0.151517	153	0.00000	8131.1	4.635159	
[0, 1, 5]	0.143944	64	0.00001	4851.9	4.403476	
[2, 4, 6]	0.104688	545	0.00136	29402.6	3.202582	
[2, 4, 5]	0.100090	269	0.00220	17041.6	3.061932	
[4, 5, 9]	0.099370	157	0.00237	7401.6	3.039907	
[0, 5, 7]	0.093259	139	0.00433	7771.9	2.852950	
[6, 8, 9]	0.087829	129	0.00721	5020.0	2.686827	
[5, 7, 8]	0.084172	187	0.01003	10642.5	2.574961	
[2, 5, 6]	0.083044	233	0.01107	14251.5	2.540452	
[1, 4, 5]	0.081674	166	0.01247	5594.2	2.498541	
[1, 8, 9]	0.081121	347	0.01308	19760.5	2.481640	
[0, 5, 8]	0.080149	145	0.01421	9380.0	2.451900	
[1, 5, 8]	0.078269	229	0.01665	15233.4	2.394374	
[1, 4, 9]	0.077344	419	0.01798	20971.3	2.366076	
[3, 8, 9]	0.075971	293	0.02012	15176.9	2.324071	
[4, 7, 9]	0.075019	445	0.02174	26294.6	2.294963	
[3, 4, 7]	0.070650	280	0.03067	11341.9	2.161288	
[2, 4, 9]	0.068273	761	0.03674	43743.2	2.088600	
[5, 6, 9]	0.067529	79	0.03884	3687.3	2.065828	
[1, 4, 8]	0.065667	348	0.04455	14151.8	2.008850	

Next Steps

- More data!
- Test Gaussian Mixture model
- Applications: team-by-team recommendations, projection of player type and NBA comps for rookies or college players