

# NBA Player Types and Optimal Lineup Composition

Mark Corey  
[github.com/MarkDCorey/nba\\_player\\_clustering](https://github.com/MarkDCorey/nba_player_clustering)

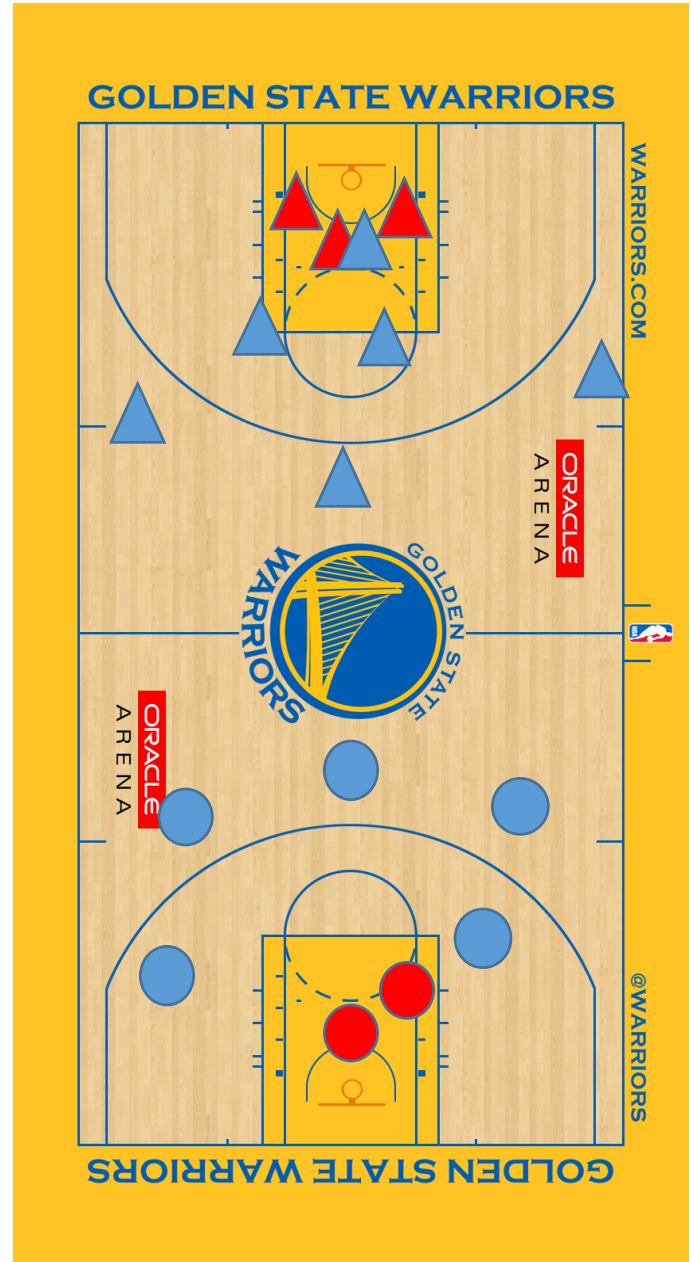


# Player Data Example

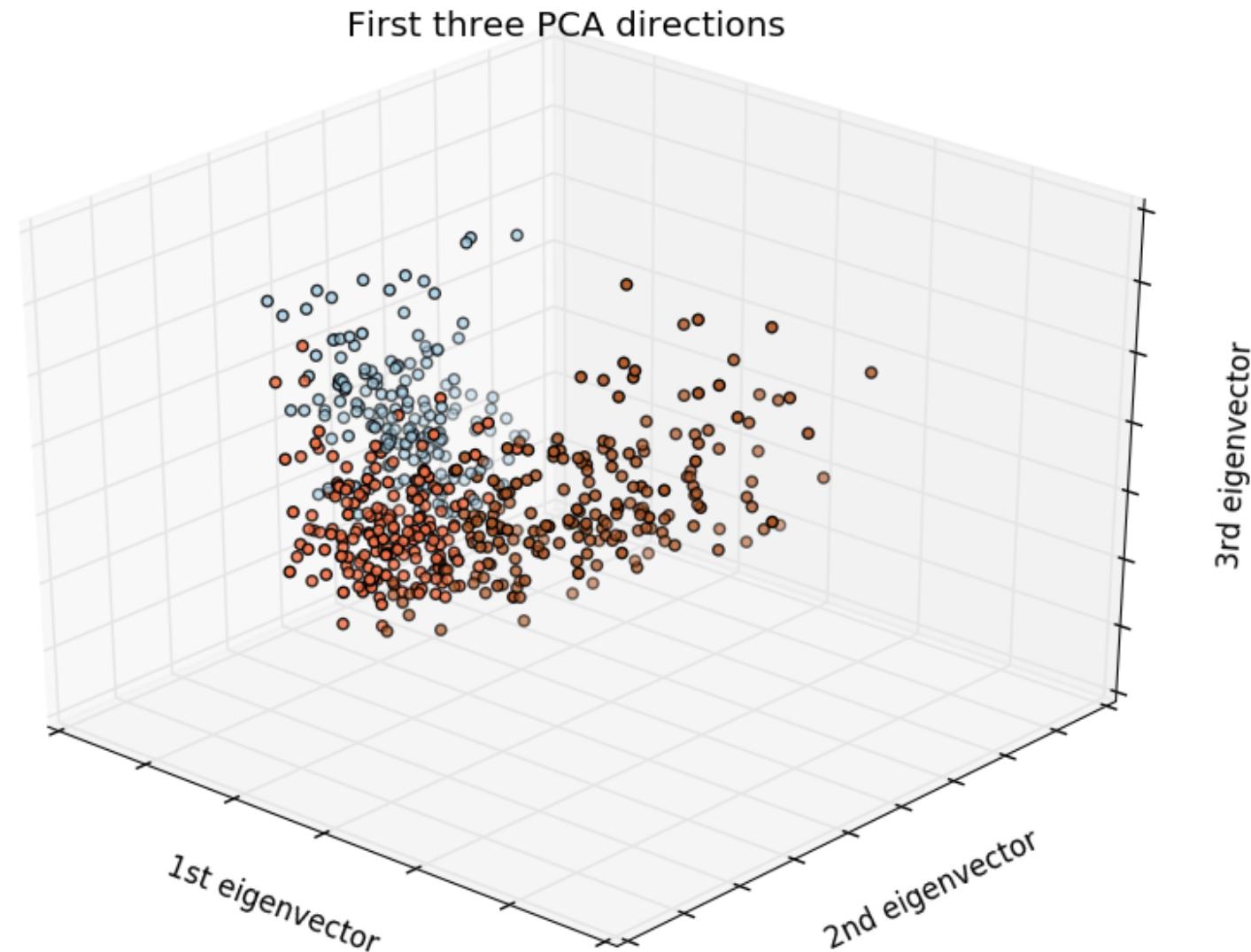
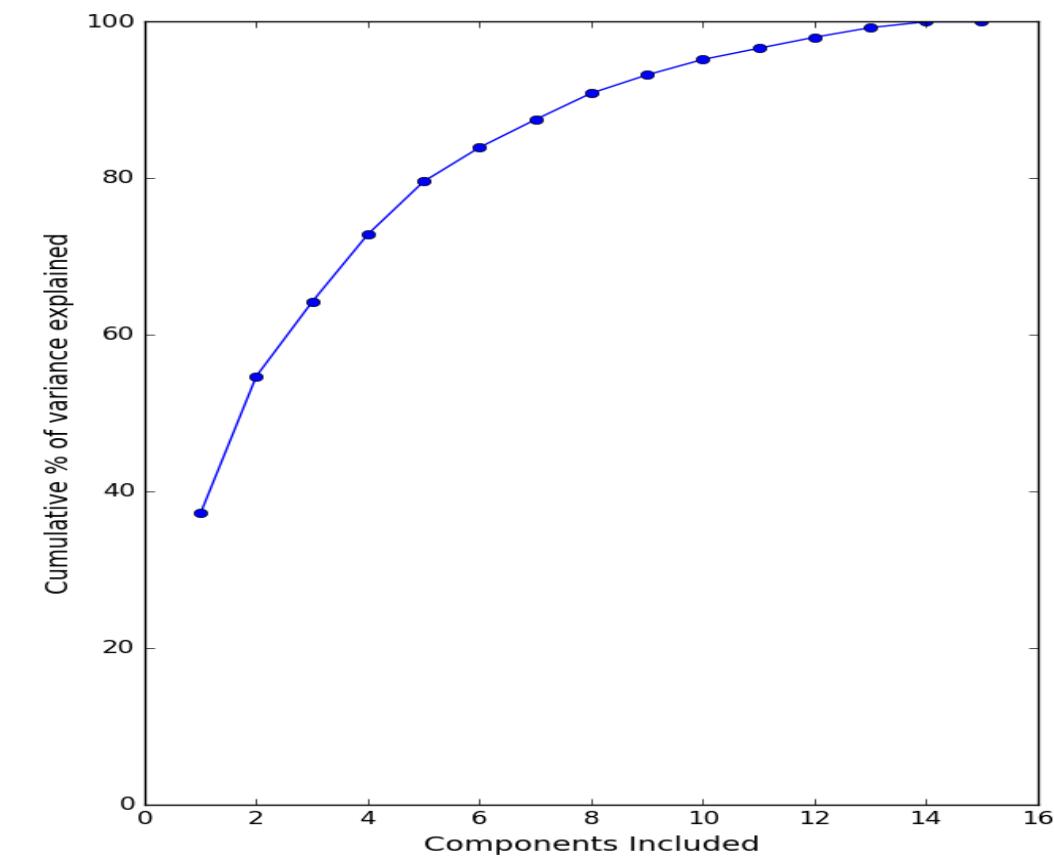
```
blk_pos : 0.012
dreb_pos : 0.204
oreb_pos : 0.04
stl_pos : 0.097
attempt_RA_pos : 0.054
attempt_paint_pos : 0.022
attempt_corner_3_pos : 0.017
attempt_non_corner_3_pos : 0.11
attempt_mid_pos : 0.051
ast_shot_pct : 0.418
ast_pos : 0.105
fta_pos : 0.065
d_fga_paint_pct : 0.18
d_fga_threes_pct : 0.365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



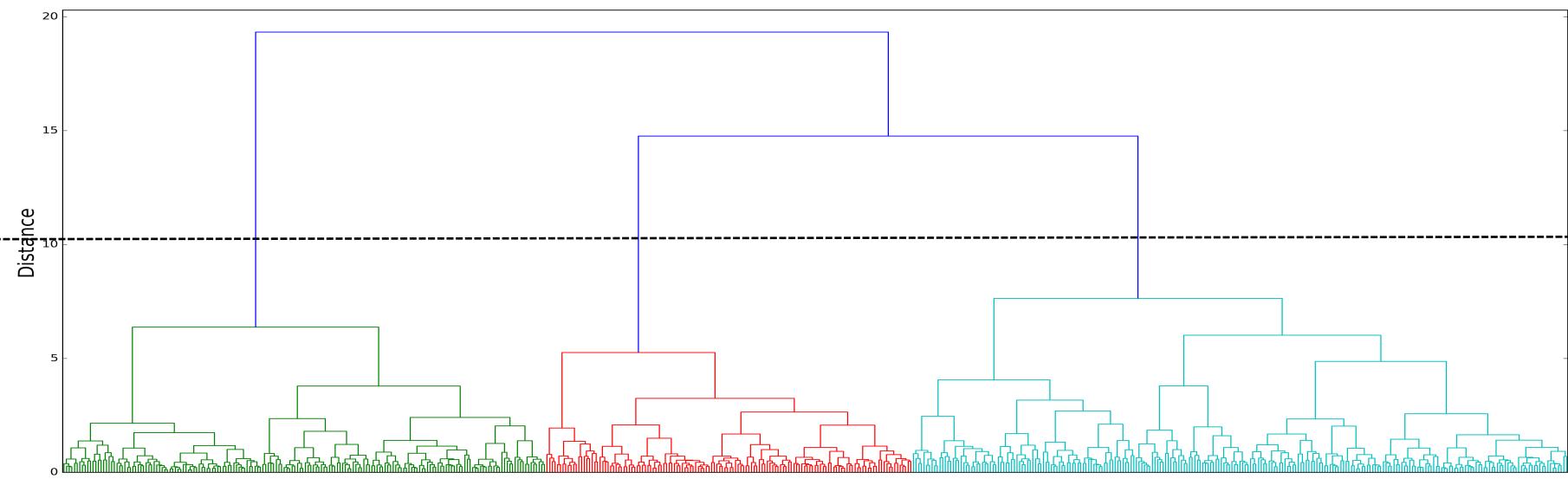
```
blk_pos : 0.166
dreb_pos : 0.296
oreb_pos : 0.201
stl_pos : 0.017
attempt_RA_pos : 0.159
attempt_paint_pos : 0.042
attempt_corner_3_pos : 0.0
attempt_non_corner_3_pos : 0.001
attempt_mid_pos : 0.031
ast_shot_pct : 0.753
ast_pos : 0.008
fta_pos : 0.07
d_fga_paint_pct : 0.509
d_fga_threes_pct : 0.102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

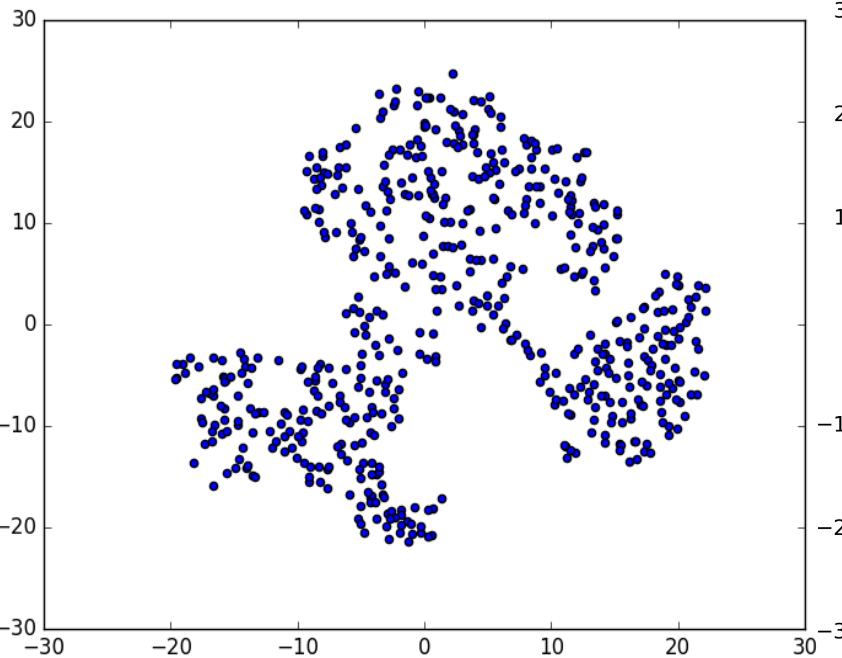
feature	std_from_avg
attempt_corner_3_pos	0.826298
ast_shot_pct	0.704936
attempt_non_corner_3_pos	0.419247
d_fga_threes_pct	0.396999
d_fga_paint_pct	-0.263190
stl_pos	-0.269688
attempt_mid_pos	-0.291332
d_fga_mid_pct	-0.336092
blk_pos	-0.382645
ast_pos	-0.438954
dreb_pos	-0.462606
oreb_pos	-0.486201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

**Distributors and shot creators (203)**

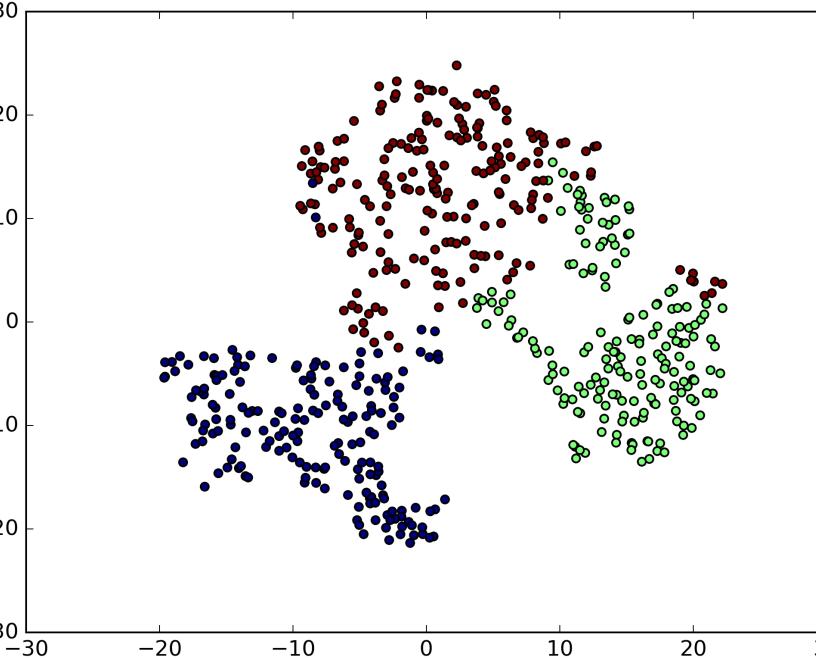
feature	std_from_avg
ast_pos	0.928268
d_fga_threes_pct	0.738274
stl_pos	0.422507
attempt_non_corner_3_pos	0.415379
attempt_mid_pos	0.290623
fta_pos	0.128286
attempt_paint_pos	0.092343
d_fga_mid_pct	0.090562
attempt_corner_3_pos	0.045846
attempt_RA_pos	-0.147727
dreb_pos	-0.322438
blk_pos	-0.476669
oreb_pos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

# t-SNE

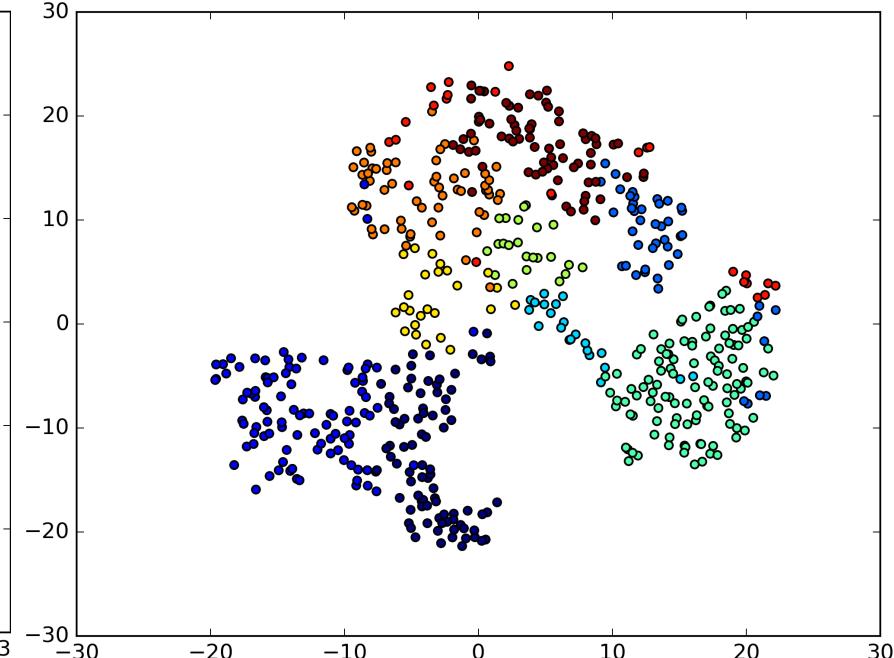
Full data set reduced to two dimensions



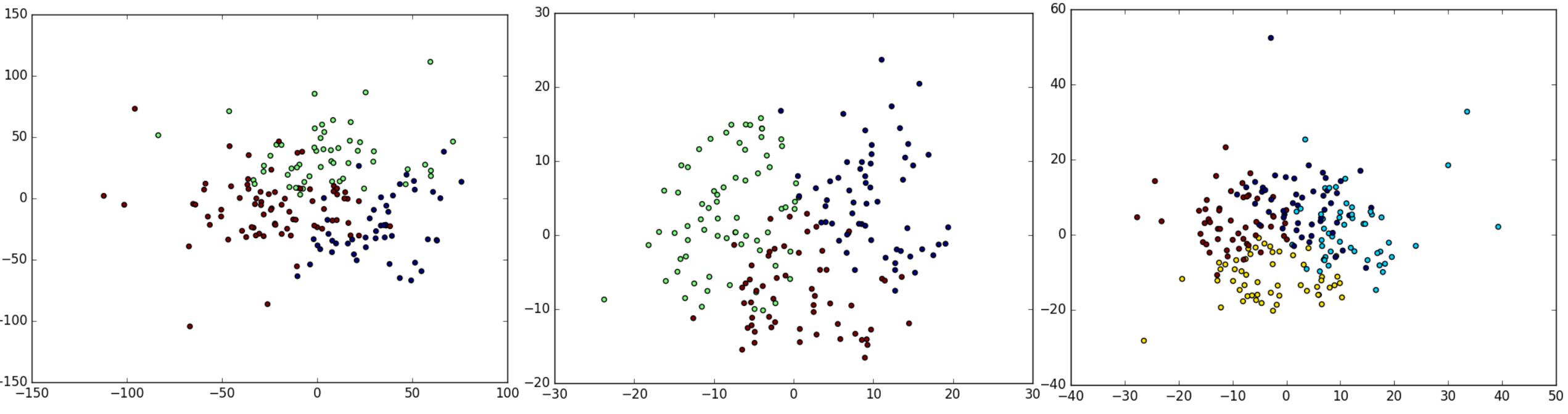
3 Hierarchical Clusters Overlaid



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
- Takeaway:

## 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

## Cluster: 0

	feature	std_from_avg
d_fga_paint_pct	blk_pos	1.564289
	attempt_RA_pos	1.181094
	oreb_pos	0.776865
	ast_shot_pct	0.768222
	dreb_pos	0.663659
	fta_pos	0.301395
	attempt_paint_pos	0.299988
	d_fga_mid_pct	0.112458
	stl_pos	-0.068249
	attempt_mid_pos	-0.451555
	ast_pos	-0.465546
attempt_non_corner_3_pos		-0.698354
attempt_corner_3_pos		-1.005966
d_fga_threes_pct		-1.030197

## Cluster: 1

	feature	std_from_avg
d_fga_paint_pct	oreb_pos	1.223622
	attempt_paint_pos	0.866471
	blk_pos	0.821876
	dreb_pos	0.742007
	attempt_RA_pos	0.646996
	fta_pos	0.621512
	d_fga_mid_pct	0.443524
	ast_shot_pct	0.428386
attempt_mid_pos		0.301087
stl_pos		0.289865
ast_pos		-0.317100
attempt_non_corner_3_pos		-0.589720
attempt_corner_3_pos		-0.936587
d_fga_threes_pct		-0.944691

## Cluster: 2

	feature	std_from_avg
d_fga_paint_pct	oreb_pos	1.478169
	dreb_pos	1.344231
	blk_pos	0.996545
	attempt_RA_pos	0.954293
	attempt_paint_pos	0.833377
	fta_pos	0.586495
	ast_shot_pct	0.399563
	d_fga_mid_pct	0.324264
	stl_pos	0.324264
attempt_mid_pos		0.034964
ast_pos		0.025979
attempt_corner_3_pos		-0.549041
attempt_non_corner_3_pos		-0.848527
d_fga_threes_pct		-0.854565

## Cluster: 3

	feature	std_from_avg
attempt_corner_3_pos	attempt_mid_pos	0.610763
	ast_shot_pct	0.543357
attempt_non_corner_3_pos	d_fga_threes_pct	0.516315
	stl_pos	0.479074
	d_fga_paint_pct	-0.173551
	attempt_mid_pos	-0.195446
	blk_pos	-0.233918
attempt_RA_pos		-0.280932
	dreb_pos	-0.281711
	fta_pos	-0.284547
	oreb_pos	-0.299535
	ast_pos	-0.302523
attempt_paint_pos		-0.406817
	d_fga_mid_pct	-0.444669

## Cluster: 4

	feature	std_from_avg
attempt_corner_3_pos	attempt_mid_pos	1.082302
	ast_shot_pct	0.963969
	d_fga_threes_pct	0.574137
attempt_non_corner_3_pos	stl_pos	0.383286
	d_fga_mid_pct	-0.135884
	attempt_mid_pos	-0.159474
	blk_pos	-0.335940
	dreb_pos	-0.341594
	ast_pos	-0.417212
d_fga_paint_pct		-0.469248
	oreb_pos	-0.530209
	fta_pos	-0.534552
	attempt_RA_pos	-0.605759

## Cluster: 5

	feature	std_from_avg
attempt_corner_3_pos	attempt_mid_pos	0.730513
	ast_shot_pct	0.549530
attempt_non_corner_3_pos	d_fga_threes_pct	0.363566
	d_fga_paint_pct	0.091399
	attempt_mid_pos	-0.000708
	ast_pos	-0.208119
attempt_paint_pos		-0.295196
	stl_pos	-0.434486
	blk_pos	-0.487610
attempt_RA_pos		-0.535867
	fta_pos	-0.539060
attempt_RA_pos		-0.580294
	oreb_pos	-0.591811
d_fga_threes_pct		-0.616097

## Cluster: 6

	feature	std_from_avg
ast_pos	d_fga_mid_pct	1.271971
	d_fga_threes_pct	0.733428
attempt_non_corner_3_pos	attempt_mid_pos	0.597855
	attempt_paint_pos	0.384682
	stl_pos	0.309780
attempt_corner_3_pos	fta_pos	0.103640
	attempt_RA_pos	0.095019
	blk_pos	-0.068311
attempt_RA_pos	dreb_pos	-0.199112
	oreb_pos	-0.362397
	blk_pos	-0.655046
	oreb_pos	-0.685833
d_fga_paint_pct	fta_pos	-0.709880
	ast_shot_pct	-0.964114

## Cluster: 7

	feature	std_from_avg
d_fga_threes_pct	stl_pos	0.859098
	ast_pos	0.731250
	fta_pos	0.488630
	attempt_mid_pos	0.375438
attempt_non_corner_3_pos	attempt_paint_pos	0.324607
	stl_pos	0.277377
attempt_corner_3_pos	dreb_pos	0.141296
	blk_pos	0.079630
attempt_paint_pos	fta_pos	0.050408
	attempt_RA_pos	-0.009734
	blk_pos	-0.245310
d_fga_paint_pct	oreb_pos	-0.349141
	ast_shot_pct	-0.596791

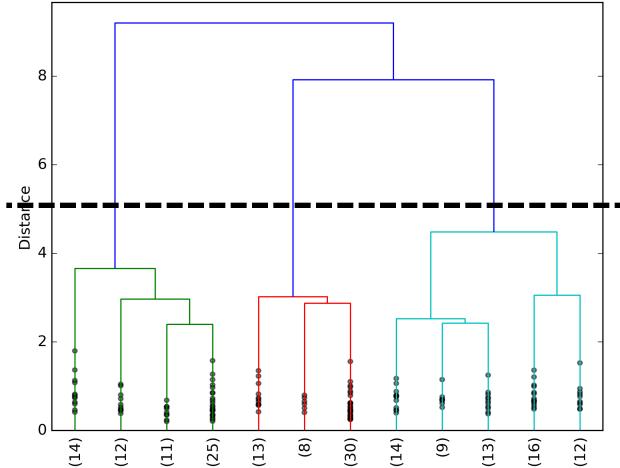
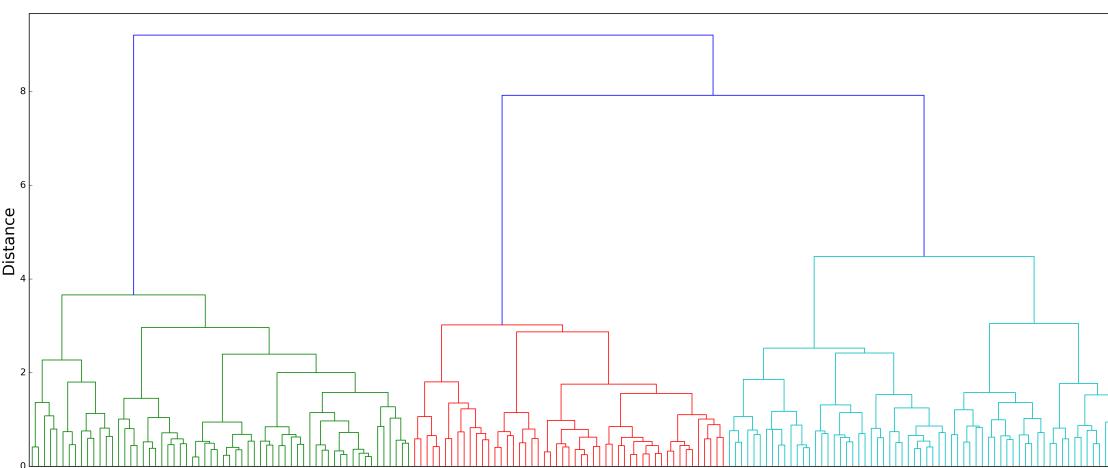
## Cluster: 8

	feature	std_from_avg
d_fga_threes_pct	ast_pos	0.0
	attempt_paint_pos	0.0
attempt_non_corner_3_pos	stl_pos	0.0
	attempt_paint_pos	0.0
attempt_corner_3_pos	d_fga_mid_pct	0.0
	attempt_mid_pos	0.0
attempt_paint_pos	fta_pos	0.0
	dreb_pos	0.0
attempt_RA_pos	blk_pos	0.0
	ast_pos	0.0
d_fga_paint_pct	oreb_pos	0.0
	attempt_paint_pos	0.0

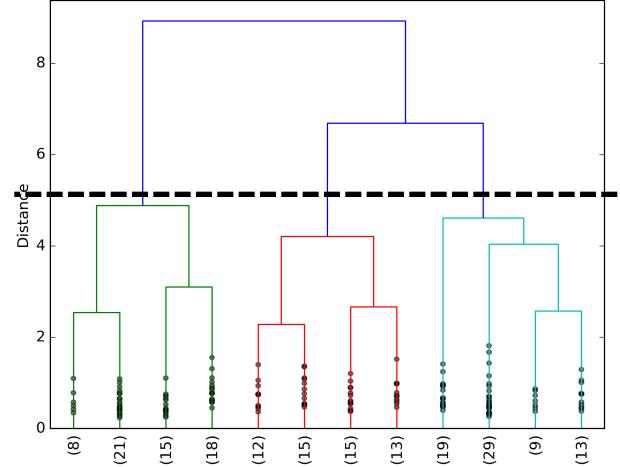
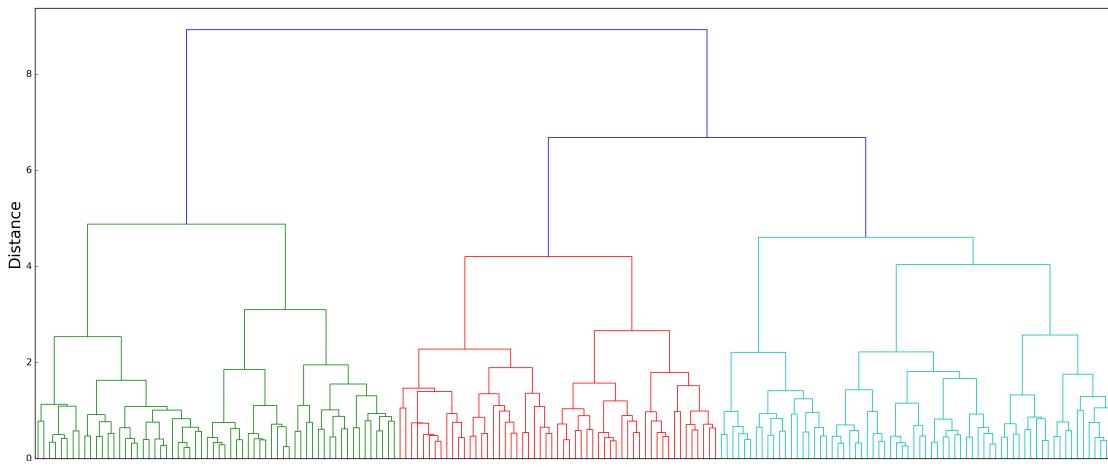
## Cluster: 9

	feature	std_from_avg
ast_pos	d_fga_threes_pct	1.0
	fta_pos	0.0
attempt_mid_pos	attempt_paint_pos	0.0
	stl_pos	0.0
attempt_non_corner_3_pos	attempt_RA_pos	0.0
	d_fga_mid_pct	0.0
attempt_corner_3_pos	dreb_pos	0.0
	blk_pos	0.0
attempt_RA_pos	oreb_pos	0.0
	blk_pos	0.0
d_fga_paint_pct	oreb_pos	0.0
	ast_shot_pct	0.0

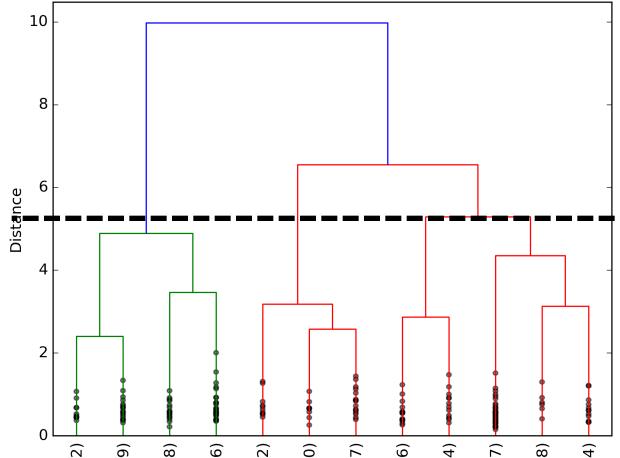
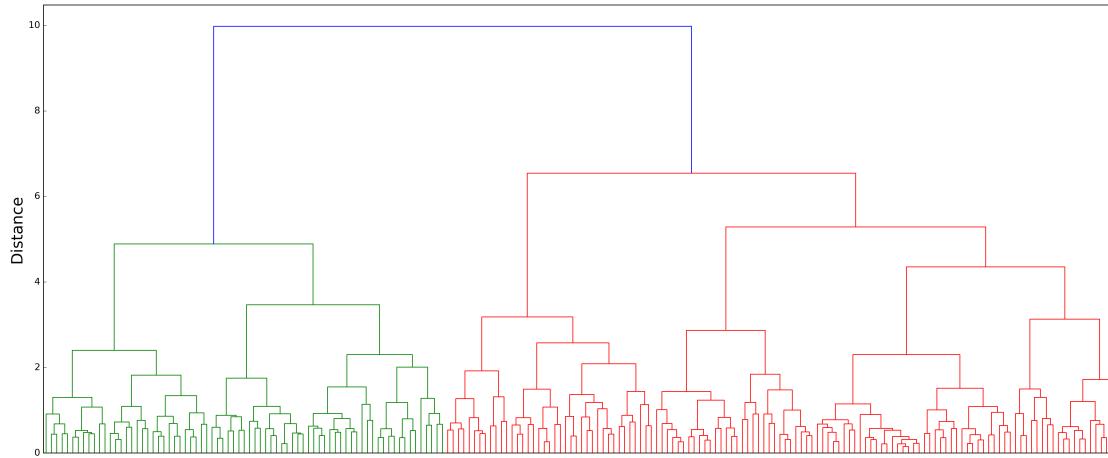
# Under-the-Basket Player Cluster (177) PC = 8



Perimeter  
Shooter/Defender  
Player Cluster  
(187)  
PC = 9



# Distributors and Shot-Creators Player Cluster (203) PC = 9



## Top 25 Lineups (90% confidence)

cluster_combo	net_per_min	p_val	t_score	c3_combo	net_per_min	observations	p_val	total_min_combo	z_score
2,3,4,5,8	0.689033	0.00317	-4.736982	[4, 5, 6]	0.151517	153	0.00000	8131.1	4.635159
1,2,7,7,7	0.635183	0.03032	-3.874287	[0, 1, 5]	0.143944	64	0.00001	4851.9	4.403476
1,2,4,7,7	0.547111	0.00419	-7.801654	[2, 4, 6]	0.104688	545	0.00136	29402.6	3.202582
0,1,3,8,9	0.542409	0.03248	-2.767297	[2, 4, 5]	0.100090	269	0.00220	17041.6	3.061932
1,2,5,8,8	0.531968	0.09761	-2.378017	[4, 5, 9]	0.099370	157	0.00237	7401.6	3.039907
0,3,5,7,7	0.523364	0.07213	-2.272588	[0, 5, 7]	0.093259	139	0.00433	7771.9	2.852950
0,4,6,6,7	0.454370	0.00199	-7.069967	[6, 8, 9]	0.087829	129	0.00721	5020.0	2.686827
2,5,6,7,7	0.391993	0.05897	-2.614345	[5, 7, 8]	0.084172	187	0.01003	10642.5	2.574961
2,3,4,4,8	0.329866	0.05741	-2.638444	[2, 5, 6]	0.083044	233	0.01107	14251.5	2.540452
1,2,3,5,8	0.307356	0.03443	-2.879493	[1, 4, 5]	0.081674	166	0.01247	5594.2	2.498541
0,0,6,6,7	0.306691	0.04739	-2.482844	[1, 8, 9]	0.081121	347	0.01308	19760.5	2.481640
2,4,6,7,9	0.296714	0.08513	-2.529298	[0, 5, 8]	0.080149	145	0.01421	9380.0	2.451900
2,4,7,8,9	0.292763	0.04785	-2.287872	[1, 5, 8]	0.078269	229	0.01665	15233.4	2.394374
1,2,3,6,7	0.275272	0.00629	-3.203810	[1, 4, 9]	0.077344	419	0.01798	20971.3	2.366076
2,6,7,7,8	0.230789	0.01997	-2.816616	[3, 8, 9]	0.075971	293	0.02012	15176.9	2.324071
1,2,7,7,9	0.229255	0.08700	-1.894264	[4, 7, 9]	0.075019	445	0.02174	26294.6	2.294963
2,4,5,7,9	0.219718	0.09571	-1.973378	[3, 4, 7]	0.070650	280	0.03067	11341.9	2.161288
0,5,6,7,7	0.218533	0.05425	-2.132173	[2, 4, 9]	0.068273	761	0.03674	43743.2	2.088600
0,2,3,4,8	0.217417	0.08655	-1.762615	[5, 6, 9]	0.067529	79	0.03884	3687.3	2.065828
0,5,7,7,8	0.212511	0.06054	-1.965028	[1, 4, 8]	0.065667	348	0.04455	14151.8	2.008850
0,2,5,5,8	0.195707	0.05601	-2.113906	[3, 7, 9]	0.063921	361	0.05053	17882.9	1.955463
0,0,6,7,8	0.161418	0.04664	-2.144474	[5, 8, 9]	0.062861	102	0.05448	4175.9	1.923014
2,4,6,7,7	0.150549	0.01963	-2.765107	[1, 2, 9]	0.061625	421	0.05940	28212.8	1.885197
1,3,4,7,8	0.132776	0.06125	-2.011992	[1, 2, 4]	0.060803	442	0.06288	22545.0	1.860061
0,4,7,8,8	0.107432	0.03651	-2.186198	[1, 5, 7]	0.060768	233	0.06303	11014.0	1.858998
				[4, 6, 9]	0.060561	166	0.06393	5515.1	1.852663
				[1, 3, 9]	0.058057	277	0.07572	15503.6	1.776076
				[2, 5, 9]	0.055172	266	0.09145	15593.1	1.687799
				[1, 6, 8]	0.055000	303	0.09246	14392.6	1.682557
				[3, 4, 8]	0.054911	253	0.09299	10347.7	1.679812
				[1, 2, 8]	0.054247	481	0.09701	29649.3	1.659521
				[2, 3, 7]	0.054202	676	0.09729	33067.0	1.658132