



NBA Lineup Analysis on Clustered Player Tendencies

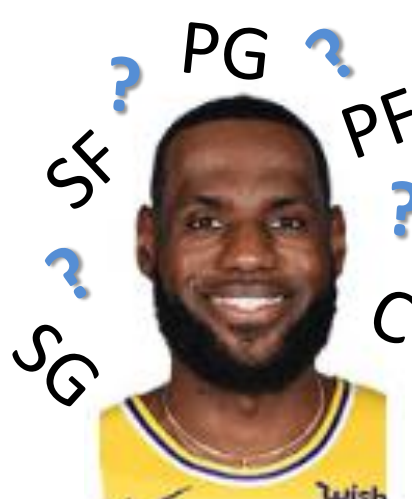
Samuel Kalman / Purdue University / kalman@purdue.edu

Jonathan Bosch / Syracuse University / jbosch@syr.edu

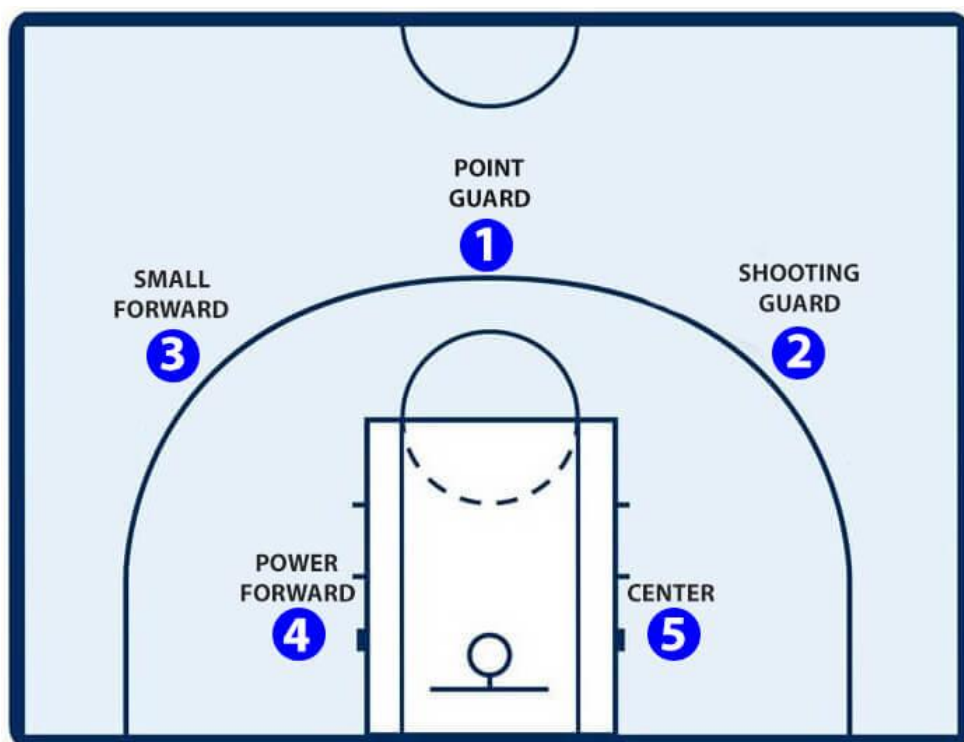
Maksim Horowitz / Atlanta Hawks / maksim.horowitz@gmail.com

Research Questions

Player role's and responsibilities are NOT captured by one single position.



LeBron James



- Can we cluster similar NBA players into new positions?
- What are the optimal lineup combinations of these new positions?

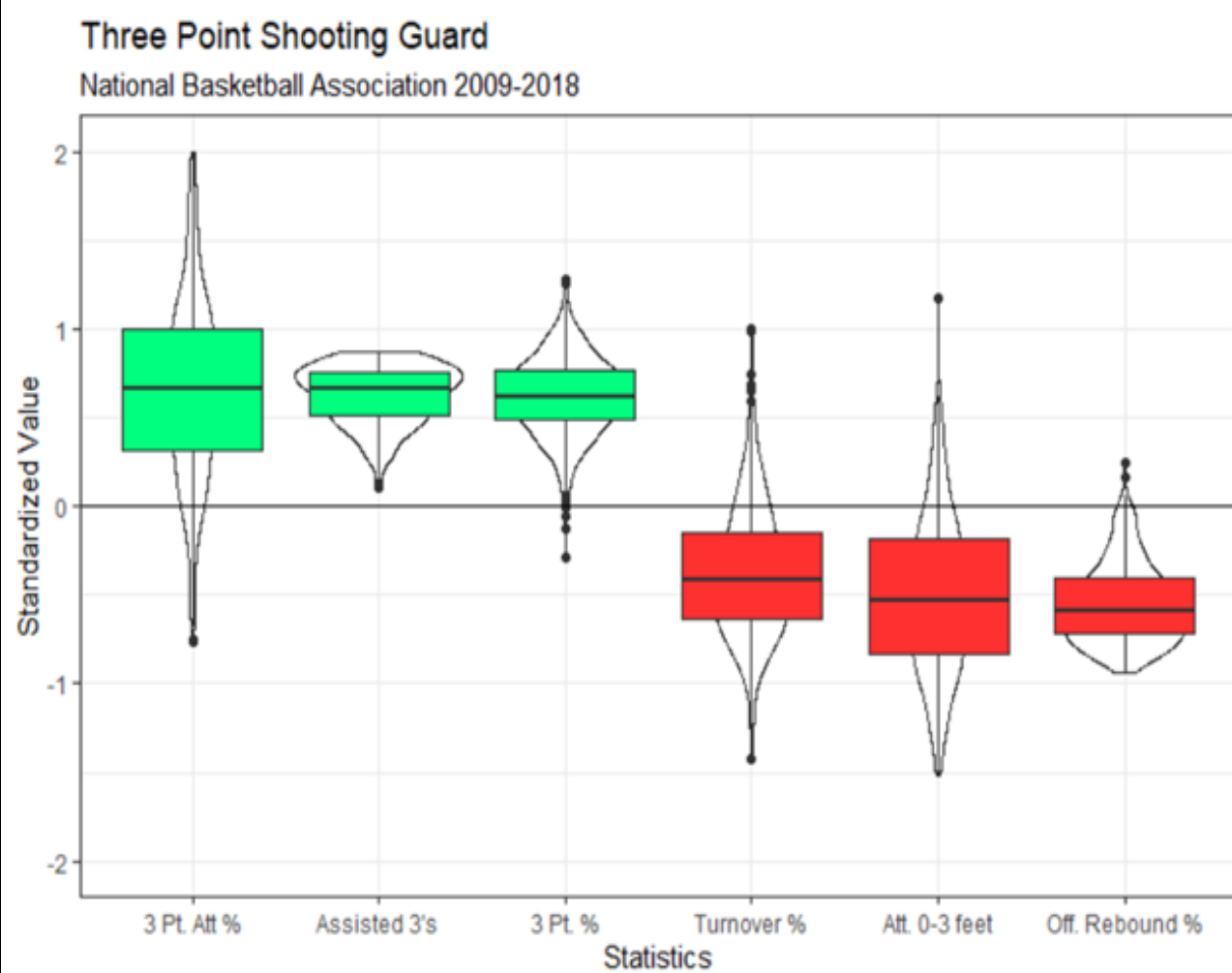
Player Clustering

- Selected & scaled **23** variables that account for tendencies, opportunity, and skill
- Model-Based clustering (**mclust** package)

Height	PER	2FG Ast Rate
Block Rate	FTTr	3FG%
Oreb Rate	FT%	3FGA%
Dreb Rate	FGA*	Corner 3FGA%
Ast Rate	2FG%	3FG Ast Rate
Steal Rate		
Usage Rate		
Turnover Rate		
Points*		
Dunk Att. Rate		
0-3ft FGA%		
3-10ft FGA%		
10ft-3p FGA%		

- Expectation-Maximization Algorithm chose nine clusters (VEV)
- Each player receives a **probability** of membership in each cluster

- Ex: Player X is 0.4 c_1 , 0.5 c_2 , 0.1 c_3 , 0 for all other c_n



Ex. Cluster #3:

“Three Point Shooting Guard”

- ❖ Shoot mostly assisted threes
- ❖ Converts threes at a high %
- ❖ Doesn't shoot near the rim
- ❖ Doesn't grab offensive rebounds

These players loiter at the three point line, with the simple offensive role of shooting threes.

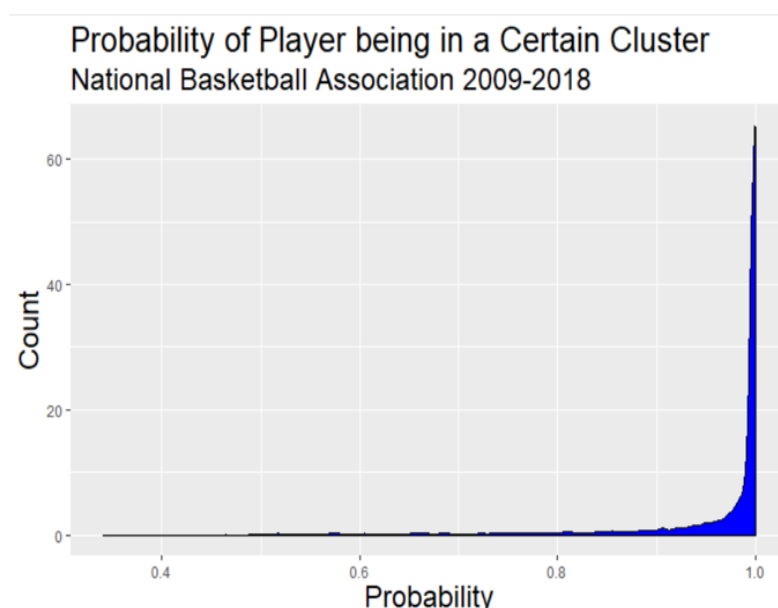
Cluster	Traits Overrepresented	Traits Underrepresented	Core Example Players
High Usage Guard	Ast Rate Usage Rate	2FG Ast Rate Height	'14 Trey Burke '16 Chris Paul
Stretch Forward	3FGA% Height	Usage Rate FTTr	'12 Shane Battier '13 Steve Novak
Three Point Shooting Guard	3FG% 3FG Ast Rate	Oreb Rate Turnover Rate	'17 Klay Thompson '18 JJ Redick
Traditional Center	Dunk Att. Rate Oreb Rate	3FGA% 3FG%	'15 DeAndre Jordan '18 Tyson Chandler
Versatile Role Player	2FG Ast Rate Oreb Rate	Points* FGA*	'13 Myers Leonard '14 Shaun Livingston
Floor General	Ast Rate Turnover Rate	Height 2FG Ast Rate	'11 Jason Kidd '12 Rajon Rondo
Mid-Range Big	10ft-3p FGA% Dreb Rate	3FGA% 3FG%	'09 Pau Gasol '15 Tiago Splitter
Skilled Forward	Dreb Rate 3FG Ast Rate	Ast Rate Steal Rate	'09 Amar'e Stoudamire '14 Anthony Davis
Ball Dominant Scorer	Points* Usage Rate	Corner 3FGA% 2FG Ast Rate	'18 James Harden '18 LeBron James

“In-between” Players:

- '13 **Matt Barnes** – 0.51 Stretch Forward, 0.49 Three Point Shooting Guard
- '17 **Dirk Nowitzki** – 0.50 Skilled Forward, 0.50 Stretch Forward

Notable Cluster Changes:

- **Steph Curry** – High Usage Guard → Ball Dominant Scorer
- **Brook Lopez** – Mid-Range Big → Skilled Forward
- **Kawhi Leonard** – Stretch Forward → Ball Dominant Scorer



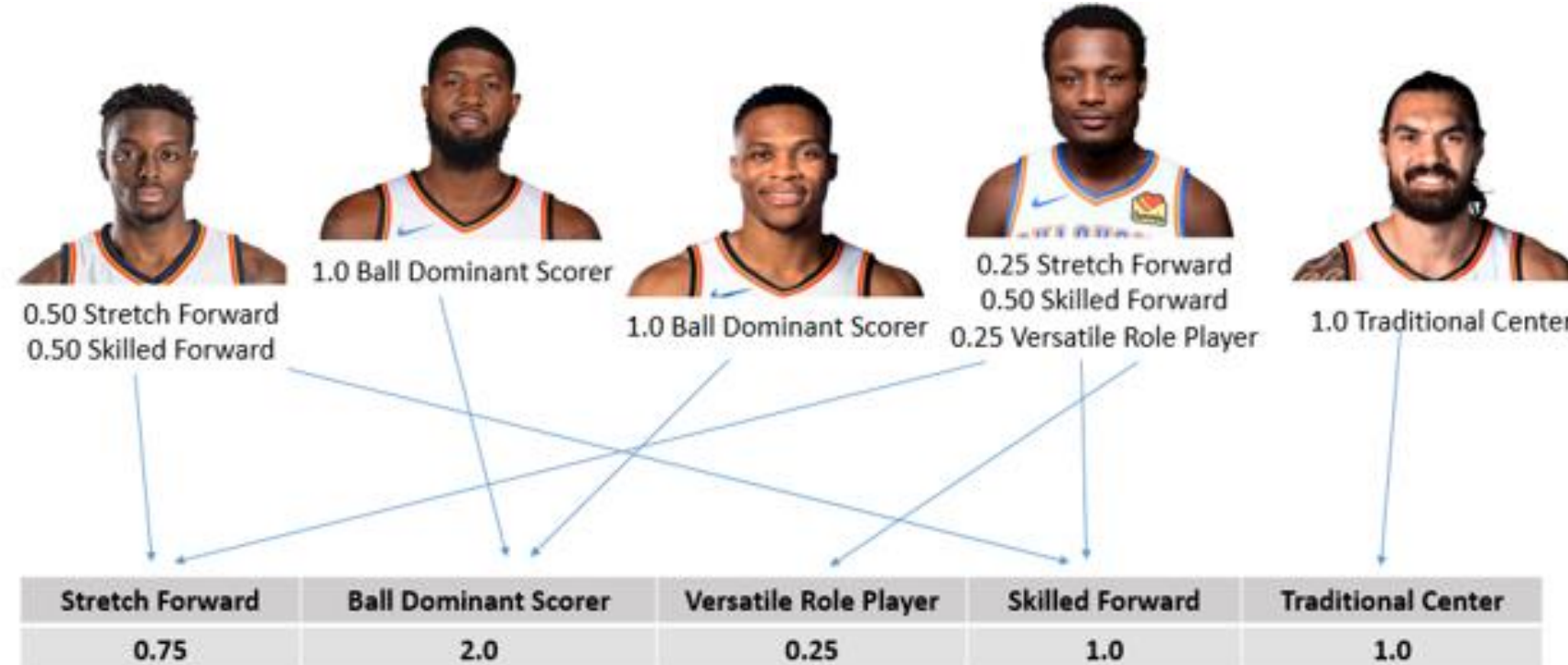
Notable Cluster Consistencies:

- **LeBron James** – Ball Dominant Scorer
- **Klay Thompson** – Three Point Shooting Guard
- **Chris Paul** – High Usage Guard

← Most players had high probabilities within 1 cluster

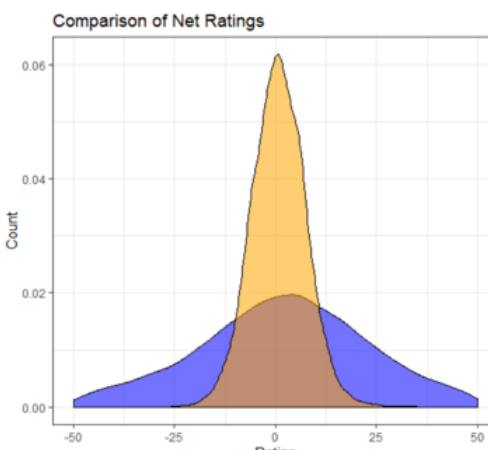
Lineup Analysis

- Nine predictor variables; “**soft-lineups**” built from cluster probabilities



Response Variable (R) = Bayesian Net Rating

Need to weight lineups to a baseline based on total possessions



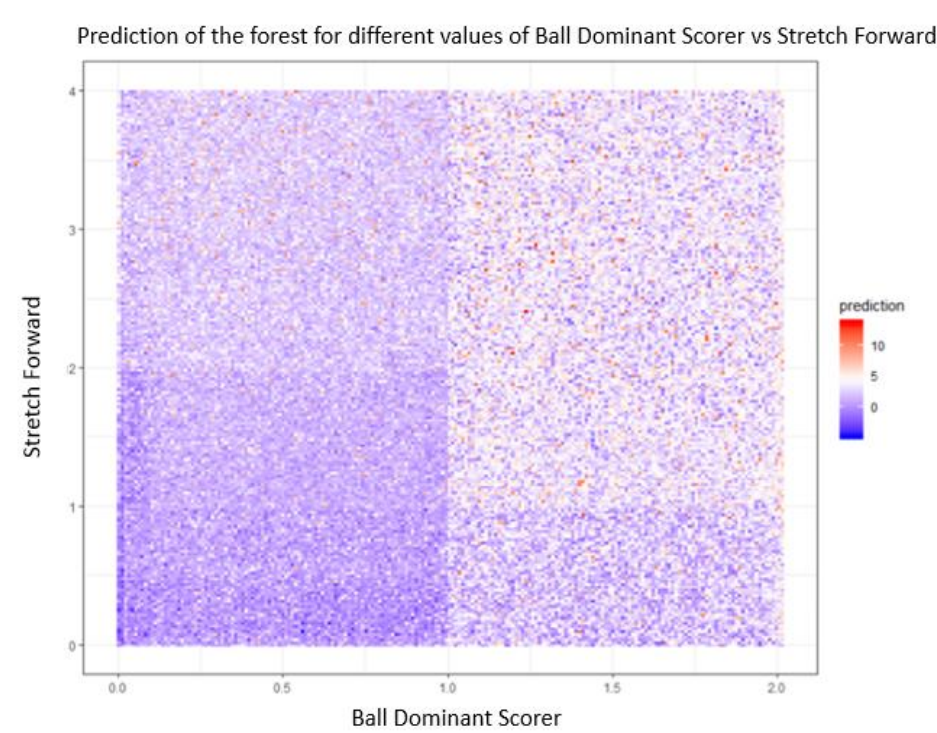
$$\text{If } \frac{\text{Possessions}}{600} \geq 1, \quad R = \text{Net Rating}$$
$$\text{If } \frac{\text{Possessions}}{600} < 1, \quad R = \left(\frac{\text{Possessions}}{600}\right) * \text{Net Rating} + \left(1 - \frac{\text{Possessions}}{600}\right) * \text{Team Net Rating}$$

Random Forest Model

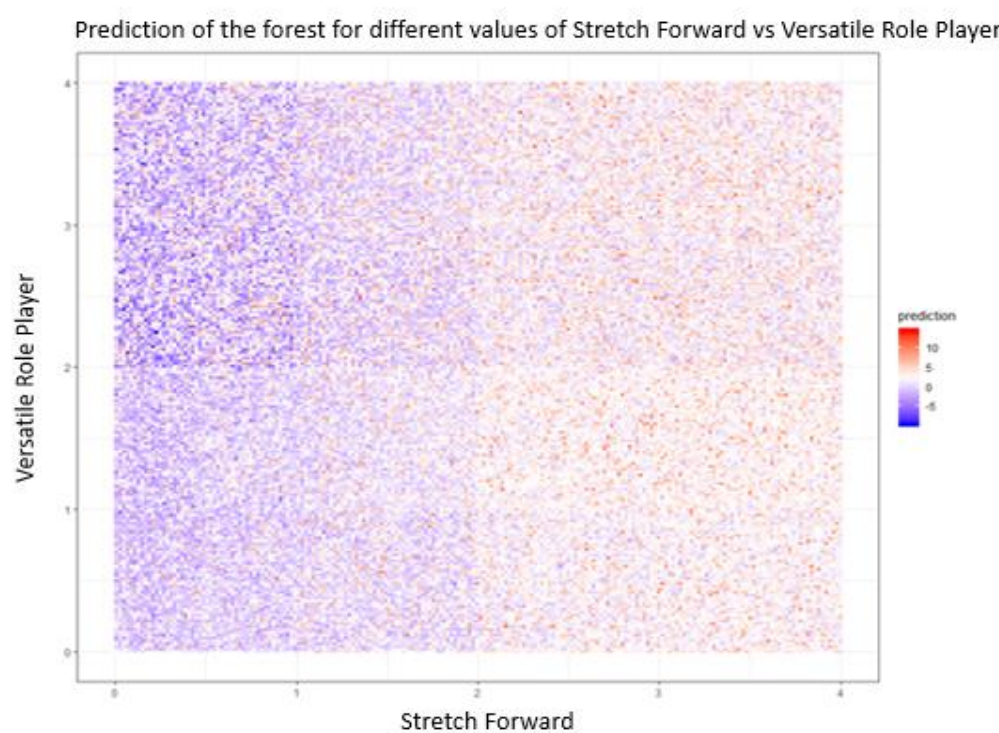
- Used all data to bootstrap 100 different random forest models
- Created prediction frame; all ~3.1 million possible soft lineups combinations

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Prediction
5	0	0	0	0	0	0	0	0	0
4.75	0.25	0	0	0	0	0	0	0	0
4.5	0.25	0.25	0	0	0	0	0	0	0
4.25	0.25	0.25	0.25	0	0	0	0	0	0
...
0	0	0	0	0	0	0	0	5	5

100 Bayesian Net Rating predictions for ALL ~3.1 million lineups



Most underperforming lineups have less than 2 stretch forwards, and less than 1 ball dominant scorer. Combining floor spacing with a high-usage, ball dominant player is very effective.



The amount of versatile role players is not nearly as important as the presence of stretch forwards, with the ability of spacing the floor.

Best Performing (predicted) Lineup

1.25 Ball Dominant Scorer	2.25 Versatile Role Player	1.0 Traditional Center	0.5 Stretch Forward	Net Rating
				14.5-15.5

'18-'19 Warriors “Death Lineup” Prediction

Steph Curry	Klay Thompson	Kevin Durant	Andre Iguodala	Draymond Green	12.4
-------------	---------------	--------------	----------------	----------------	------

- NBA players can be clustered into better role-defining positions than the traditional five
- Combinations of the new positions do matter in terms of lineup efficiency:
 - The presence of a high usage player combined with players who have the shooting ability to space the floor yields highly efficient lineups