
LS88: Sports Analytics

— Spatial Analysis in Basketball —

Outline

- Our intuitions about the game
 - ◆ Shooting occurs in space, not at distinct places for 2 or 3 points
 - ◆ Proximity to the basket should be important
 - ◆ The corner 3 technically is closer than other 3 pt shots
- What will we look at?
 - ◆ How does league/teams vary in shot choices *and* shot performance?
 - ◆ How does defense play a role?
- Our end result: a shot value model factoring in distance and defense

Basketball Review

- We've covered 4 performance metrics: FG%, eFG%, TS%, PER
- FG%, eFG%, TS% are shooting/scoring metrics
 - ◆ They ignore all other contributions beyond putting the ball in the basket
- PER is more comprehensive in evaluating a player's efficiency
 - ◆ Still flawed but isn't just solely about scoring

What do we know about basketball?

The court



What do we know about basketball?

The players in space



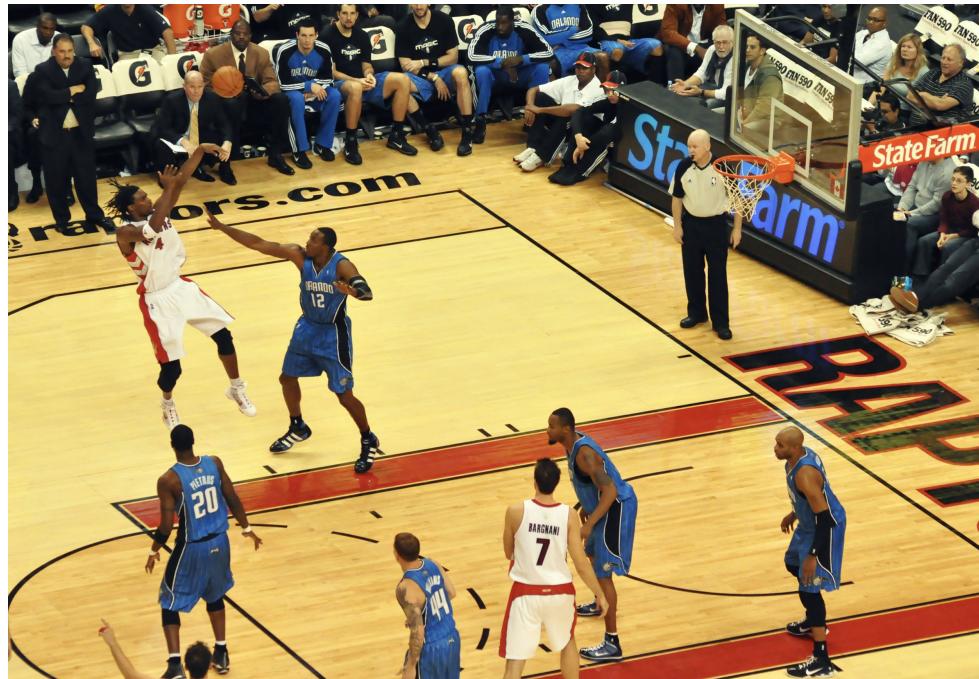
What do we know about basketball?

Short shots can be easy



What do we know about basketball?

Midrange should be harder



What do we know about basketball?

Long range 3s should be hardest



What do we know about basketball?

The corner 3 is shorter than other 3s



What do want to look at?

At the most basic level, shooting in space

Shot charts



What do want to look at?

How does the league choose to shoot?

How do teams choose to shoot?

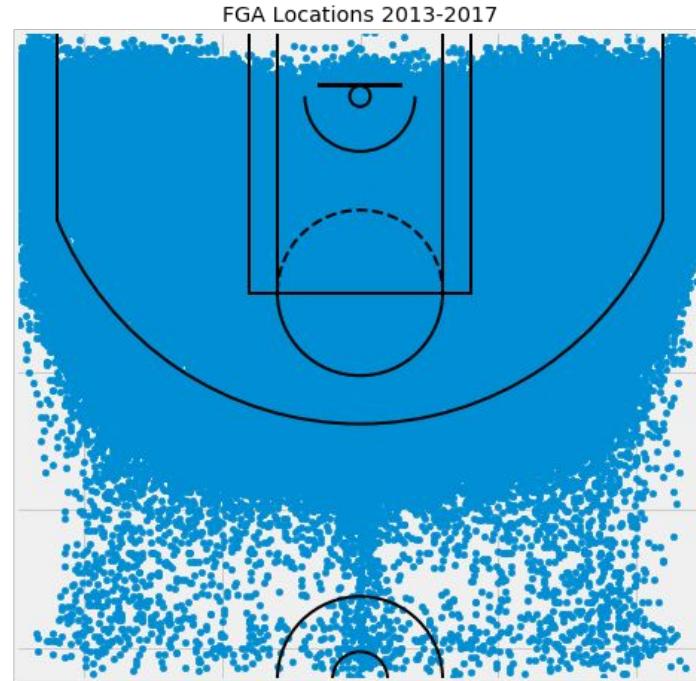
How does shooting performance vary in space?

How does defense play a role?

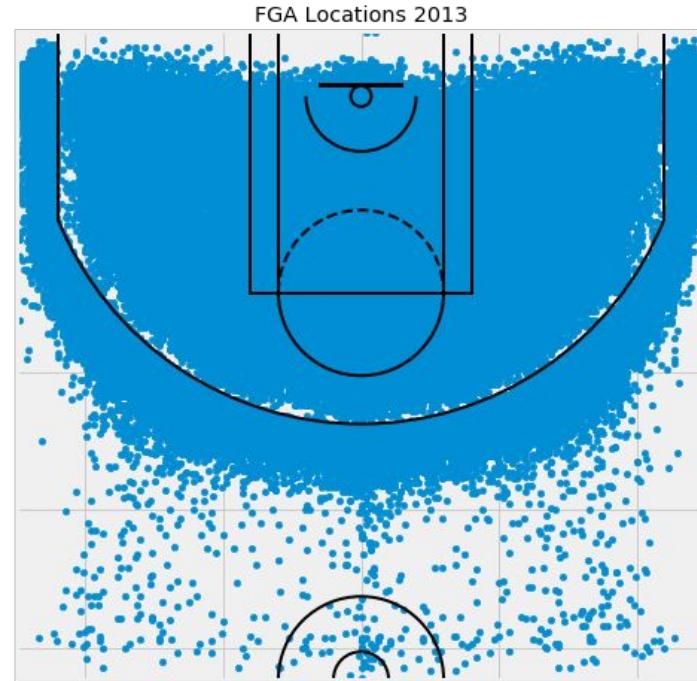
How has this changed over the last few years?

*A rough companion notebook is available that generates the following visualizations

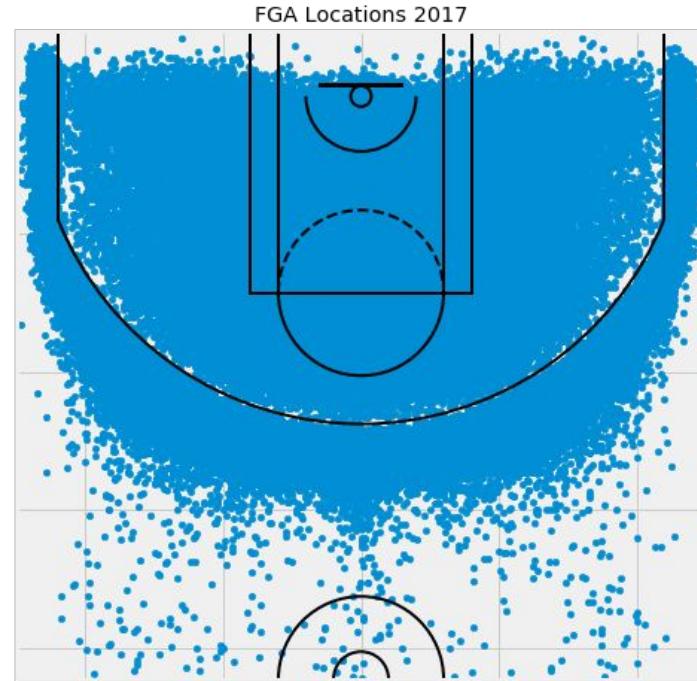
Shot Locations 2013-2017



Shot Locations 2013



Shot Locations 2017

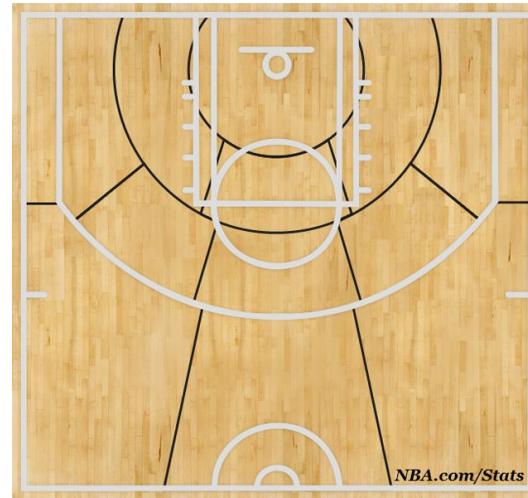


Visualizing Shooting

- Plotting so many points isn't helpful
- We need to break up the court into bins and *aggregate*
- Bins
 - ◆ We'll make a simple rectangular grid

Visualizing Shooting

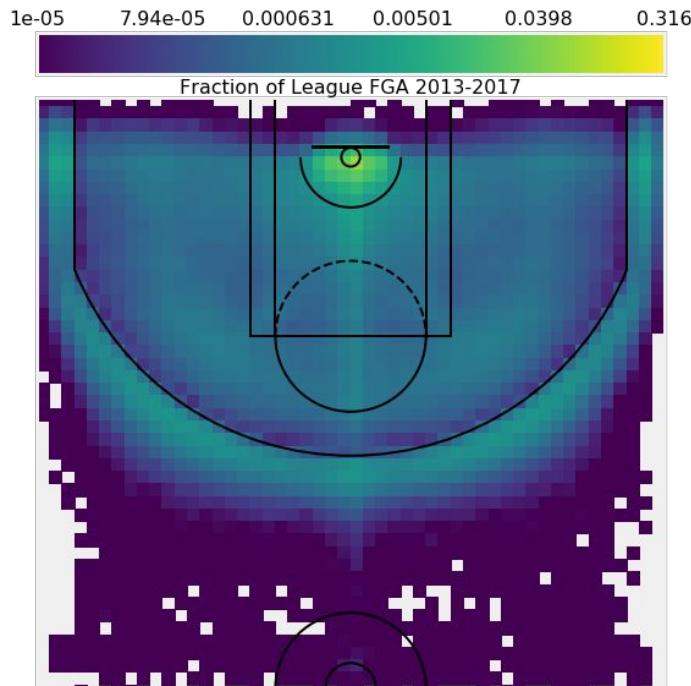
- We'll make a simple rectangular grid of shot locations
- Our bins don't have to be rectangular



Visualizing Shooting

- We'll aggregate three quantities
 - ◆ Fraction of FGA in a location
 - ◆ FG% in a location
 - ◆ Pts per FGA (or $2 \times \text{eFG\%}$) in a location

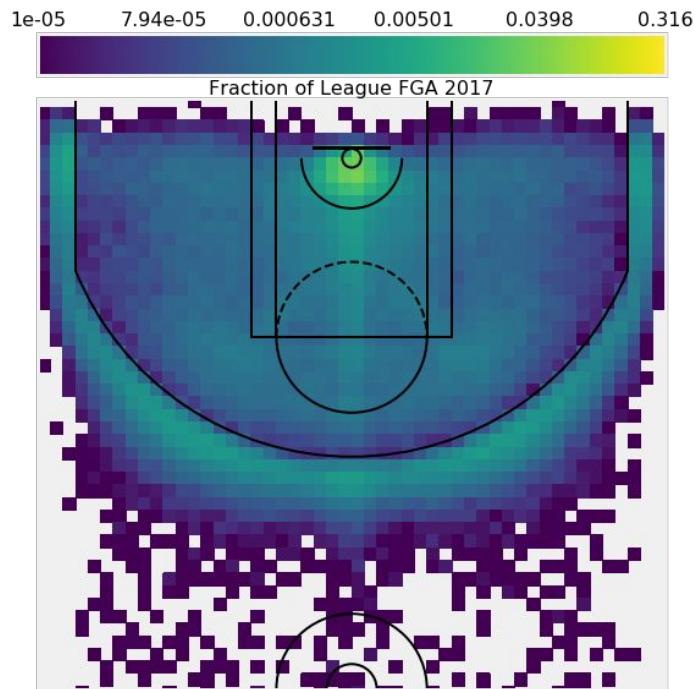
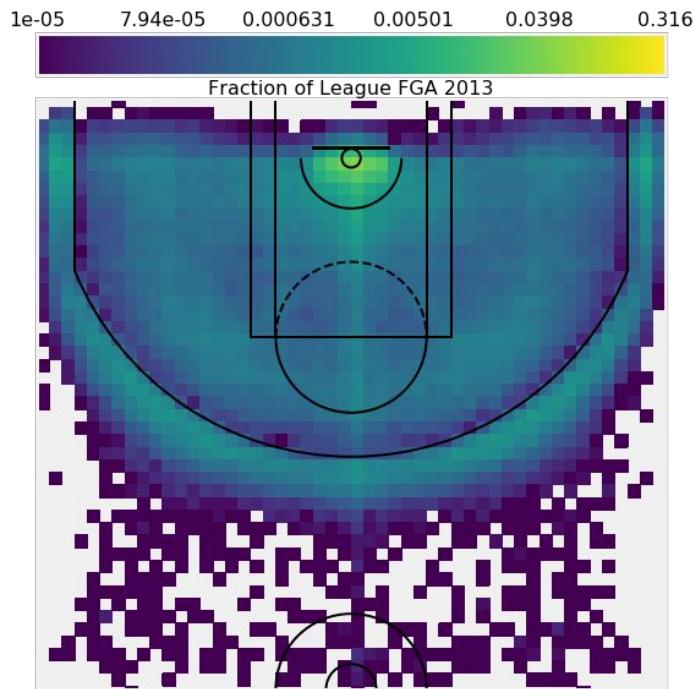
Shot Locations 2013-2017



Shot Locations 2013-2017

Has the league evolved from 2013 to 2017?

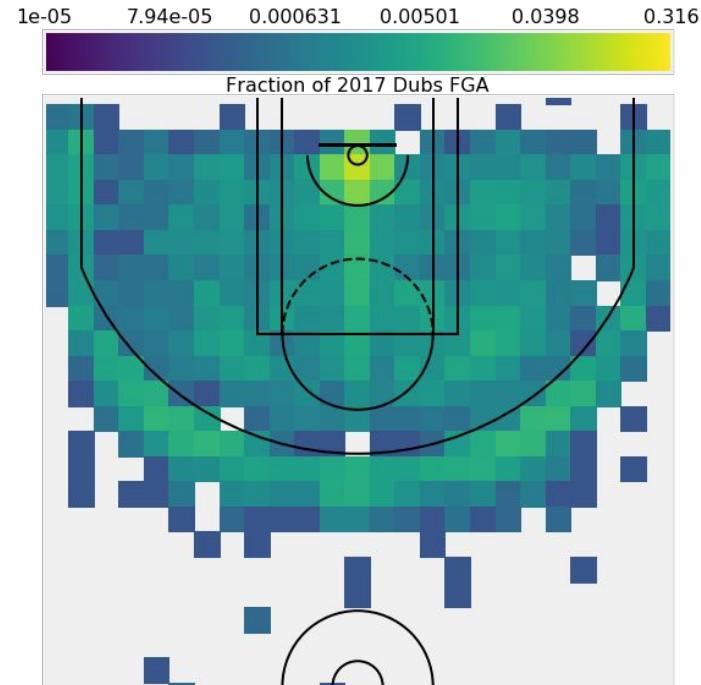
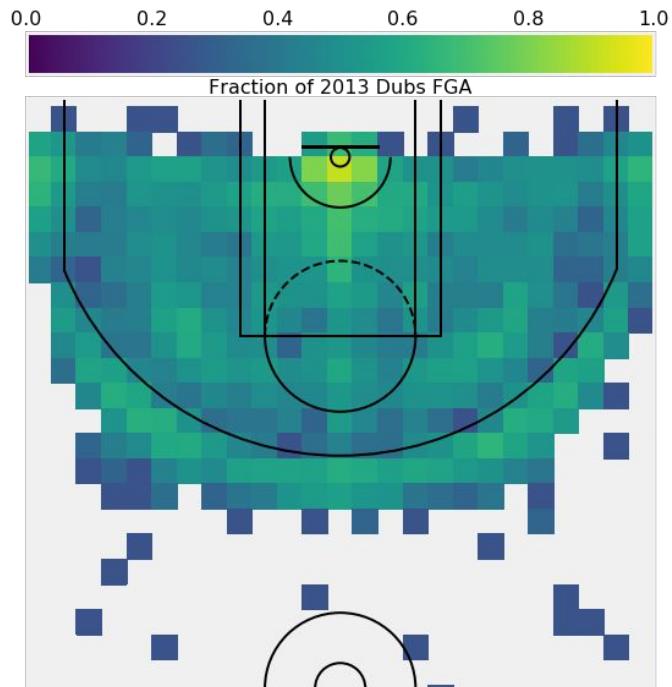
Shot Locations 2013 vs 2017



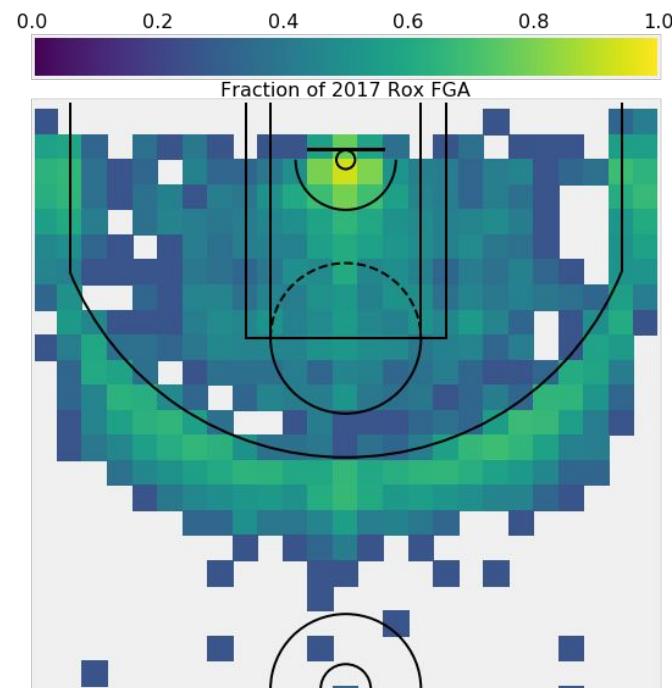
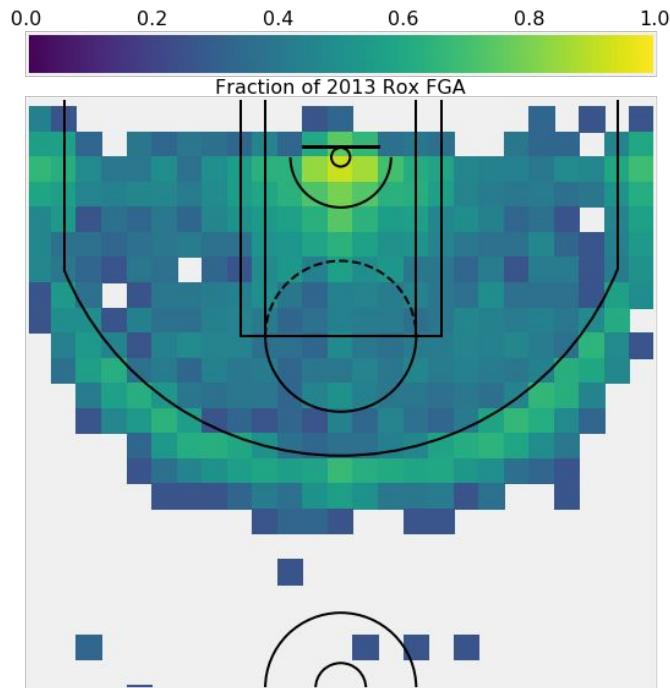
How have teams evolved?

- The league has evolved
 - ◆ Our knowledge of the league from reporting says its true
 - ◆ Data supports it
- What about teams?
 - ◆ The 2013 Warriors were coached by Mark Jackson and known to be a bad offense
 - ◆ The Rockets are renowned for the 3 pt shooting volume
 - ◆ The 2013 Lakers featured Kobe
- The Warriors and Lakers have changed a lot since 2013, does the data show it?

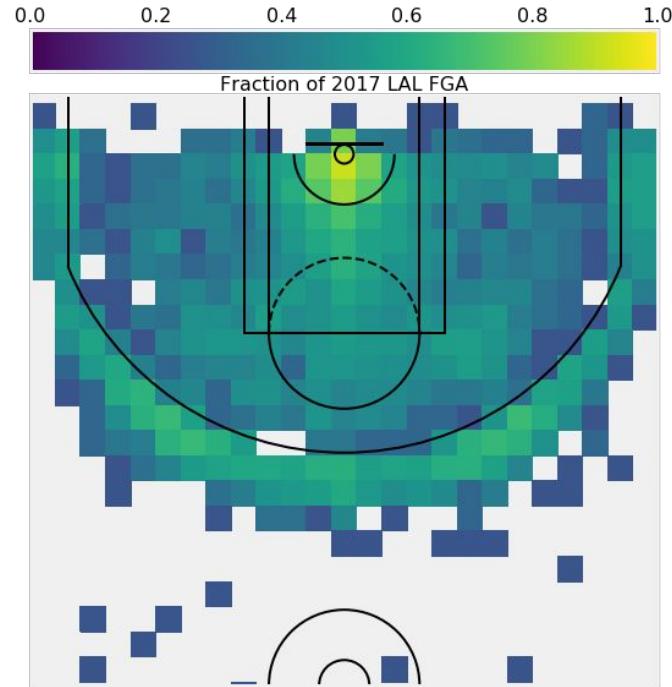
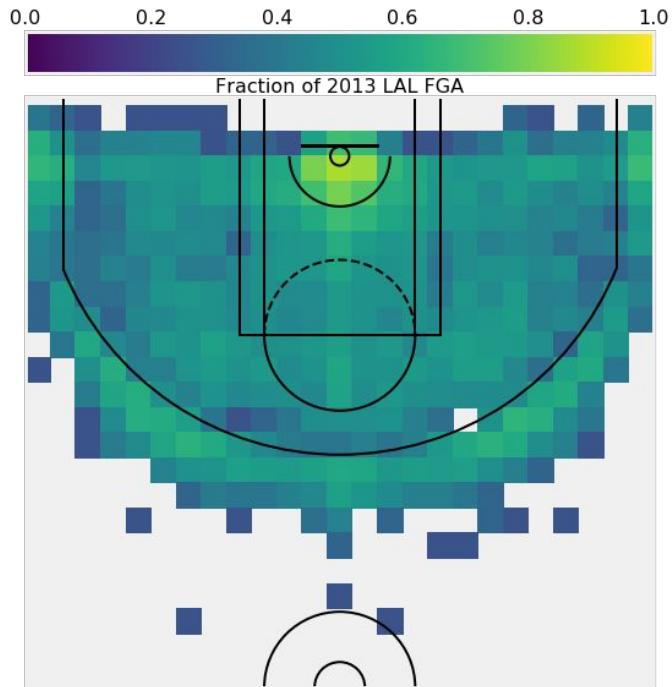
Warriors 2013 vs 2017



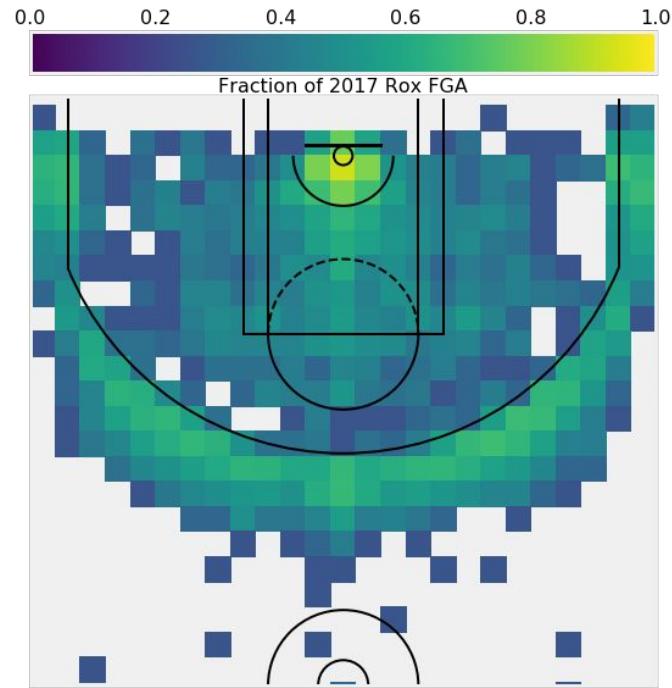
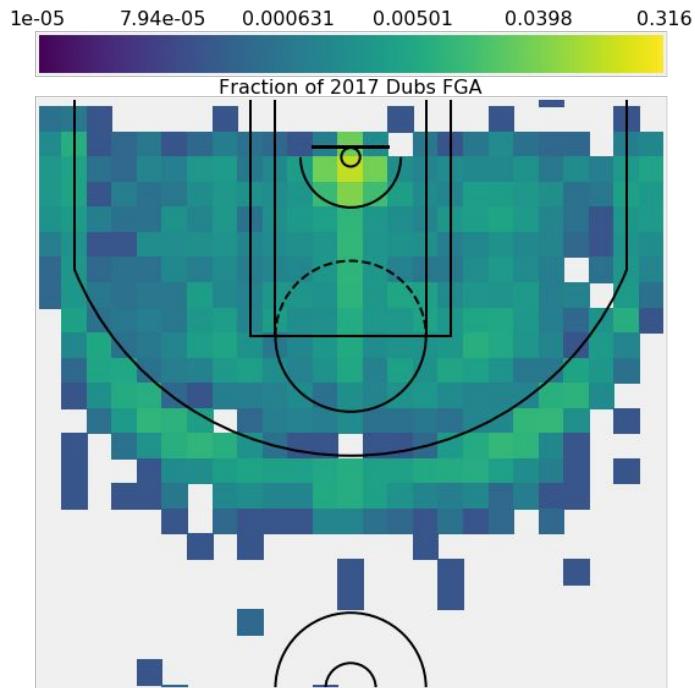
Rockets 2013 vs 2017



Lakers 2013 vs 2017



Warriors vs Rockets 2017

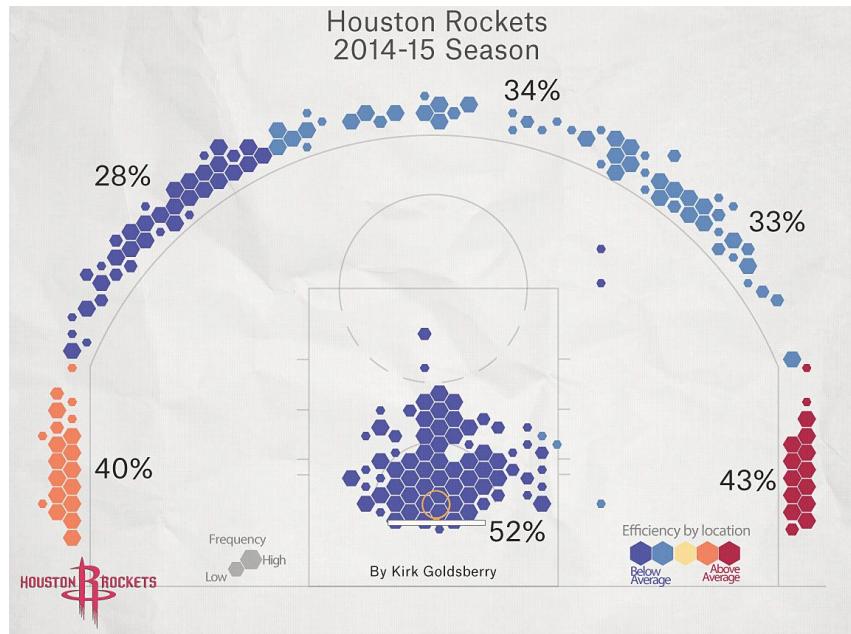


Rockets No-Midrange Game

Well-known strategic approach in the last several years

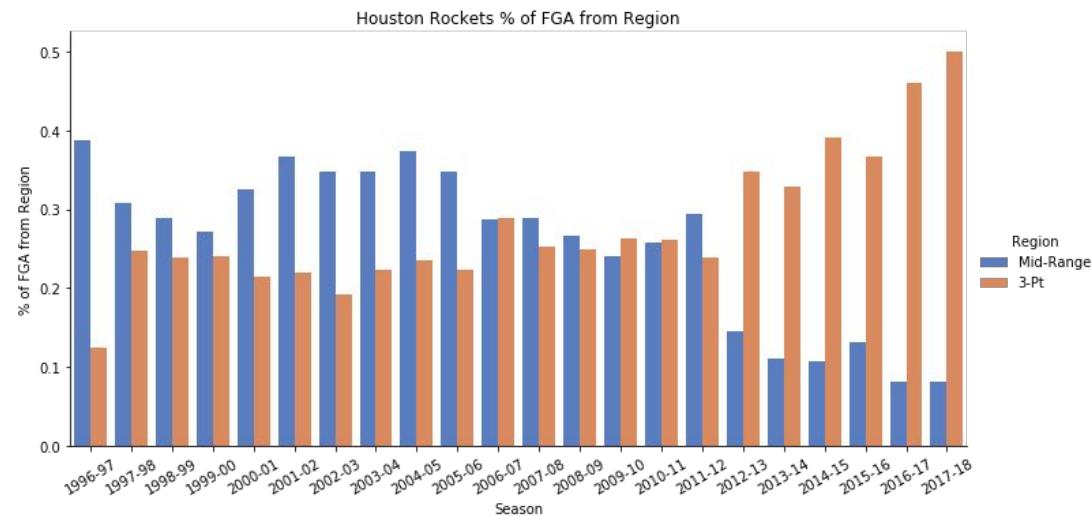
Comment 1: We'll return to Goldsberry's charts at the end

Comment 2: Kirk Goldsberry is a big innovator in NBA visualizations, but it's frustrating that he doesn't show anything in the midrange. The Rockets took shots there. He should explain the criterion for dropping out a zone.



Rockets No-Midrange Game

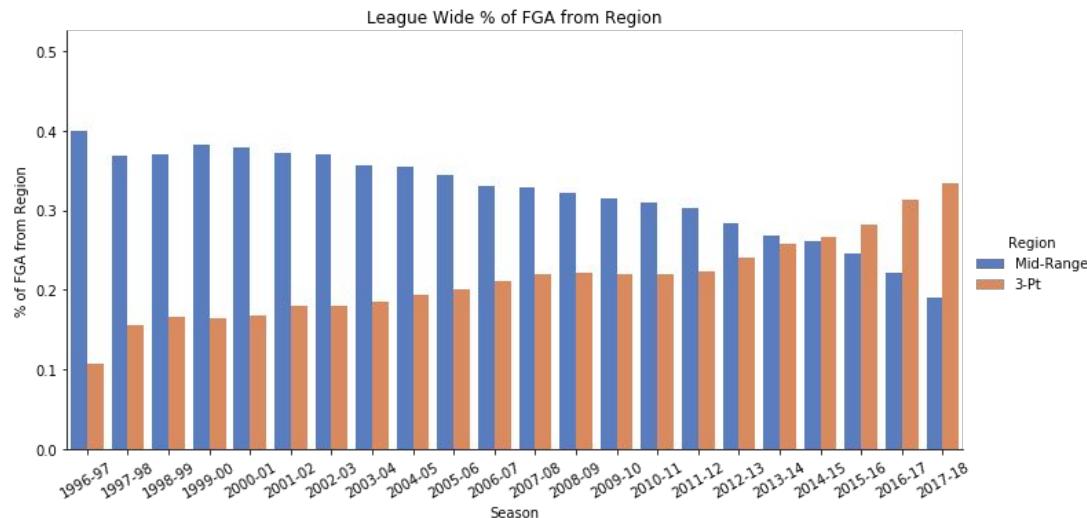
Well-known strategic approach in
the last several years



Recreation and extension of a chart from *Basketball Analytics: Spatial Tracking* by Stephen Shea

League Trend

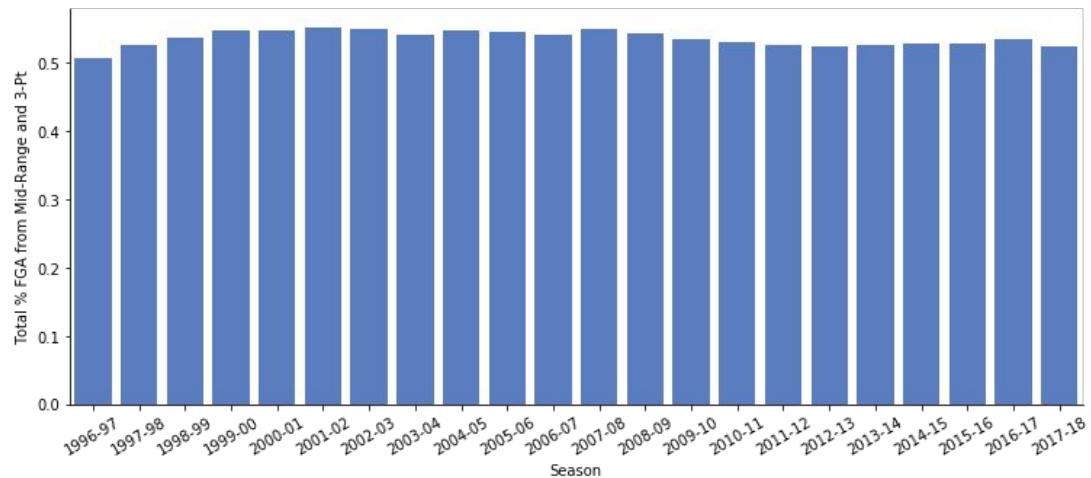
The league had already been trending this way



Recreation and extension of a chart from *Basketball Analytics: Spatial Tracking* by Stephen Shea

League Trend

No trend in Mid-range + 3pt means direct trade-off

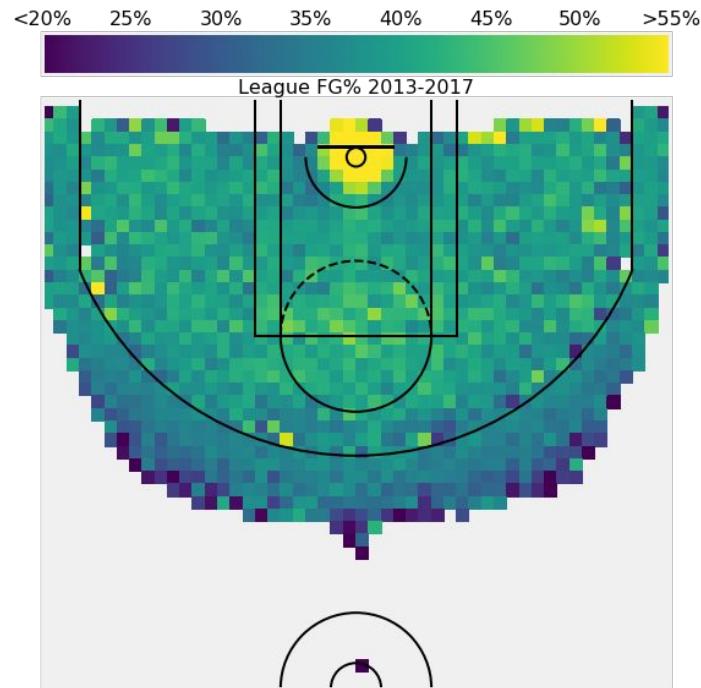


Recreation and extension of a chart from *Basketball Analytics: Spatial Tracking* by Stephen Shea

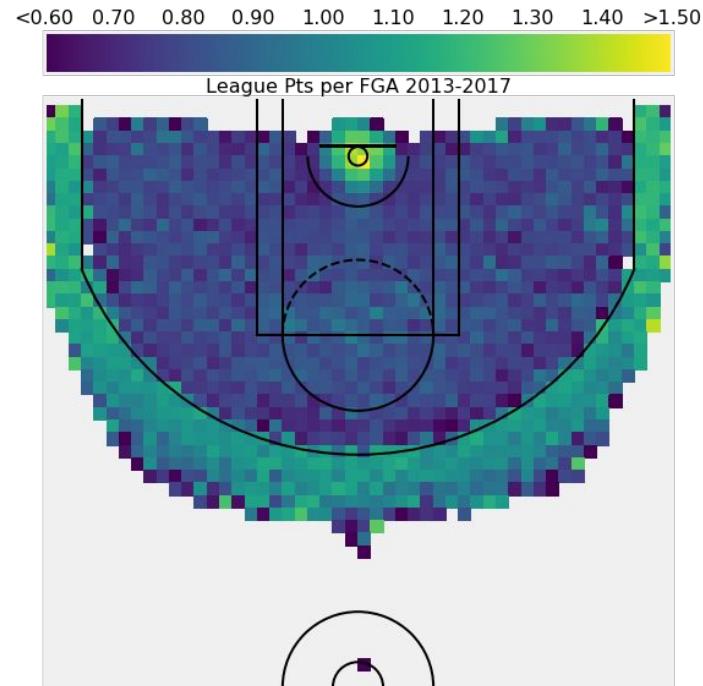
What about FG% and Scoring?

- Within each bin, we can compute FG% and Pts per FGA
 - ◆ Require a minimum of 25 shots in a bin
- We should expect to see FG% drop as the distance to the hoop decreases
- We should see a jump in Pts per FGA at the 3-pt line

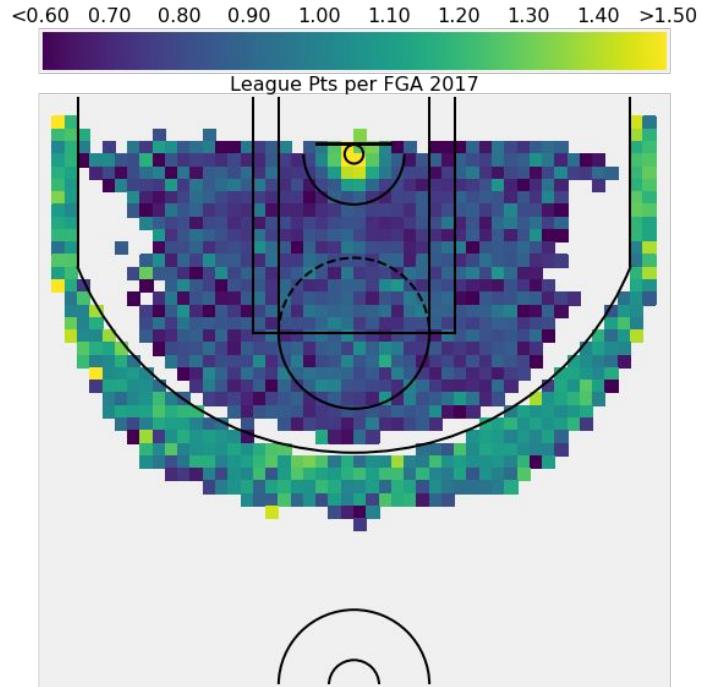
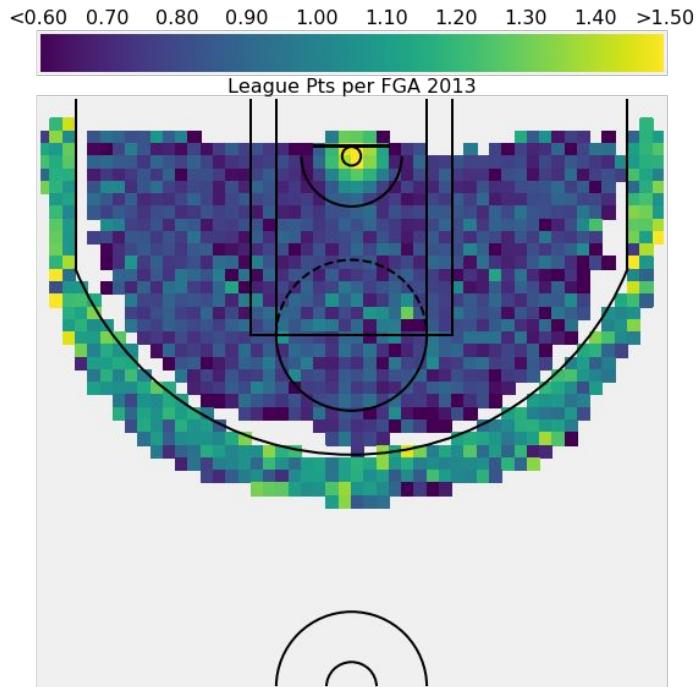
FG% 2013-2017



Pts per FGA (eFG%) 2013-2017



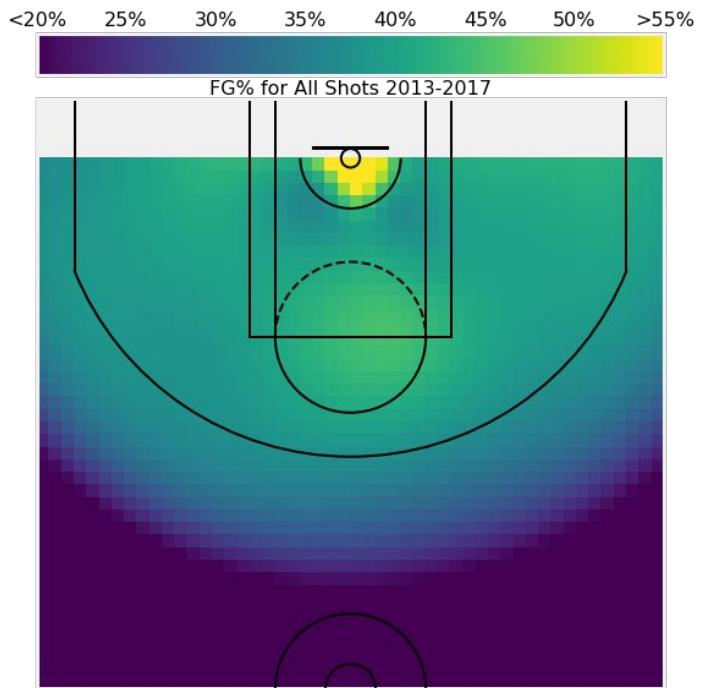
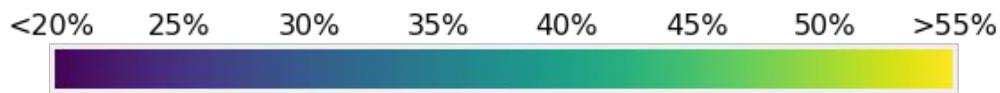
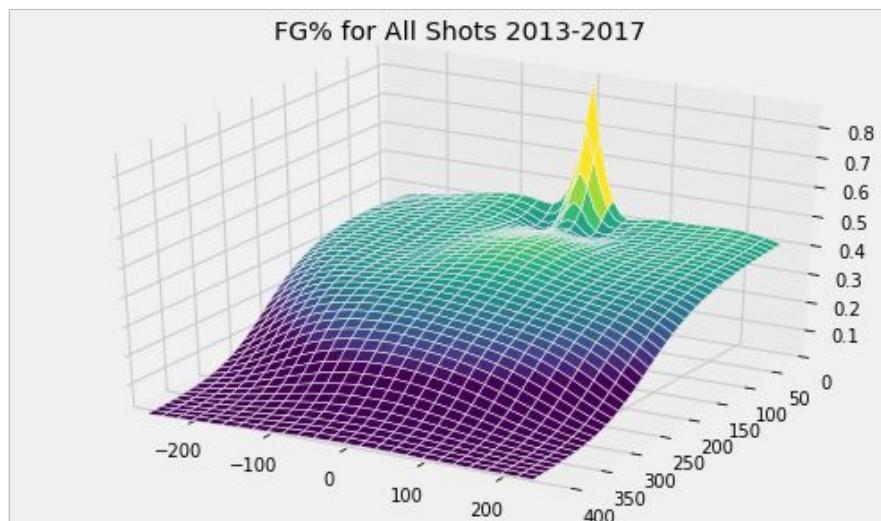
Pts per FGA (eFG%) 2013 vs 2017



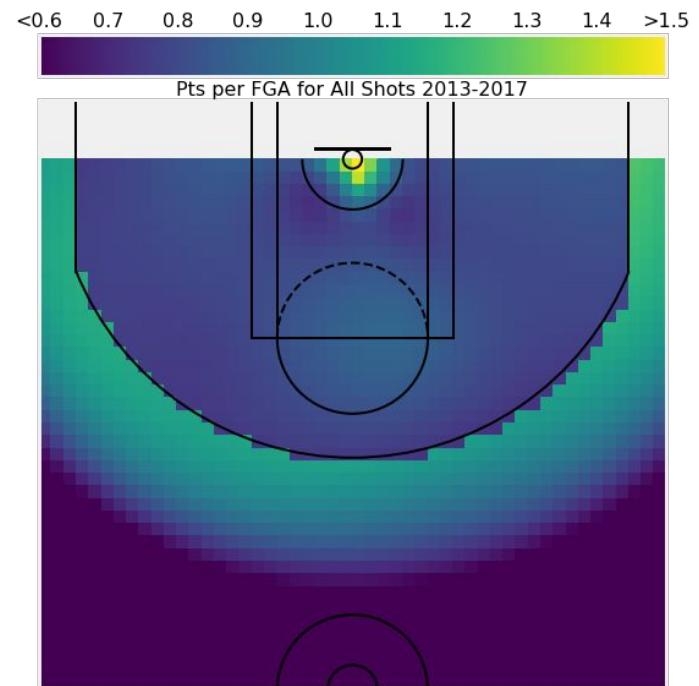
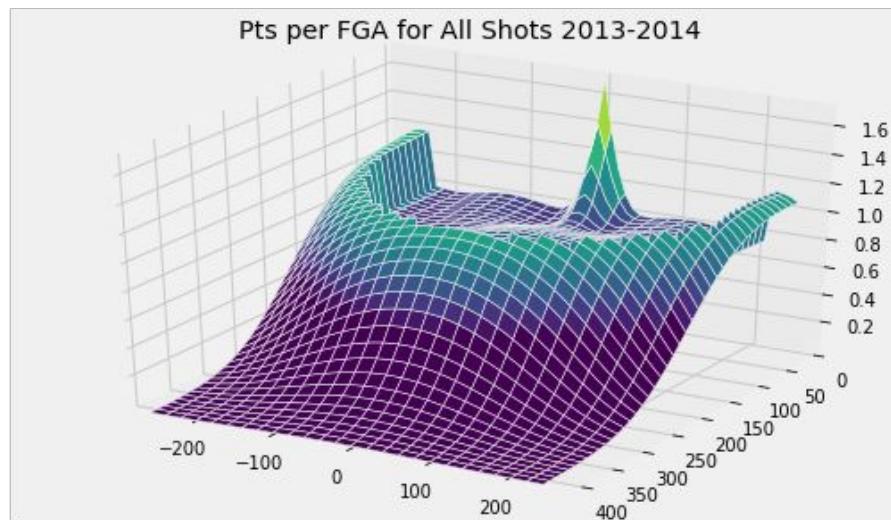
Smoothed 2-D Model

- The bucketing and visualizations are great
- But I don't like how noisy they are, so I created a smoothed model
 - ◆ Only for shots in the half court and in front of the backboard

FG% Surface & Heatmap 2013-2017



Pts per FGA Surface & Heatmap 2013-2017



Defense

- The defender *definitely* matters when factoring in a shot's performance
- Quality defense is being near the shooter and getting a hand up to contest the shot



Defense

- We only have defender proximity data
- Being close isn't a guarantee that the shot will be contested (and vice versa) but it should be a good proxy



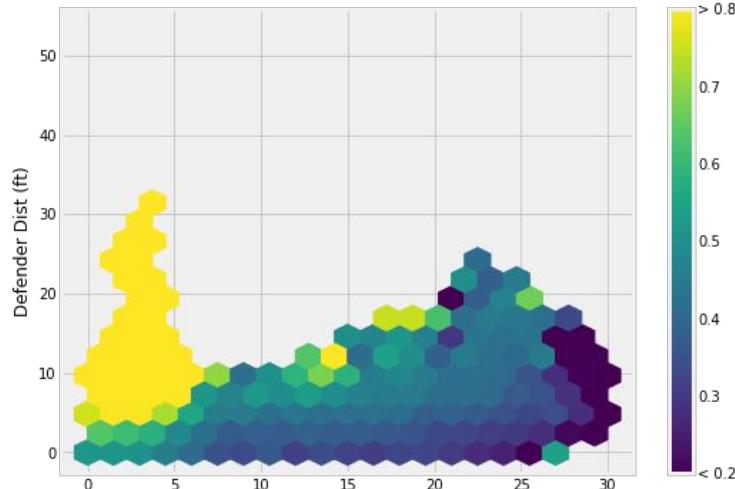
Shooting, Defense, and Expected Value

Demo

Expected Shot Value

For a shot at distance R and defender distance X

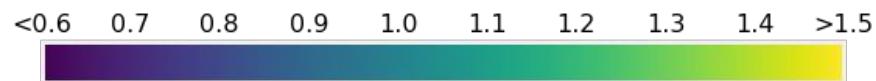
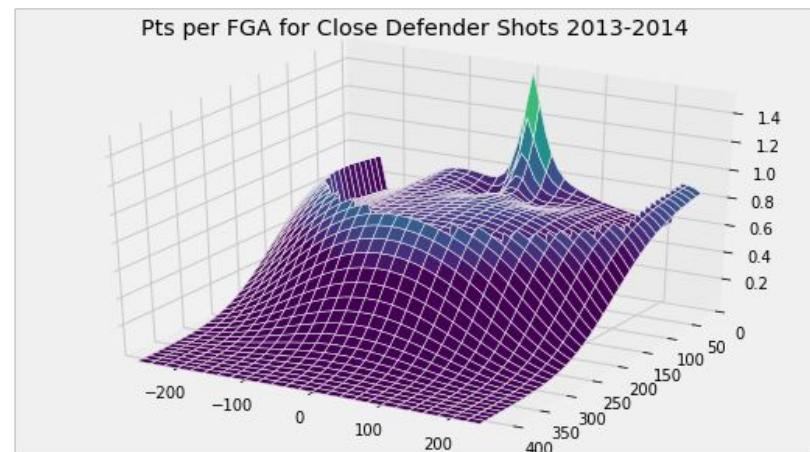
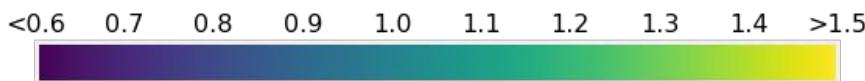
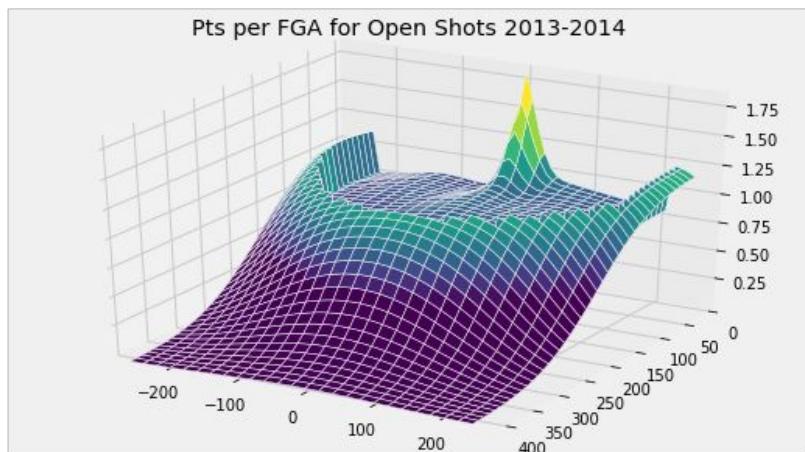
$$\text{ESV} = \text{Point Value} \times \text{Expected FG\% at } (R, X)$$



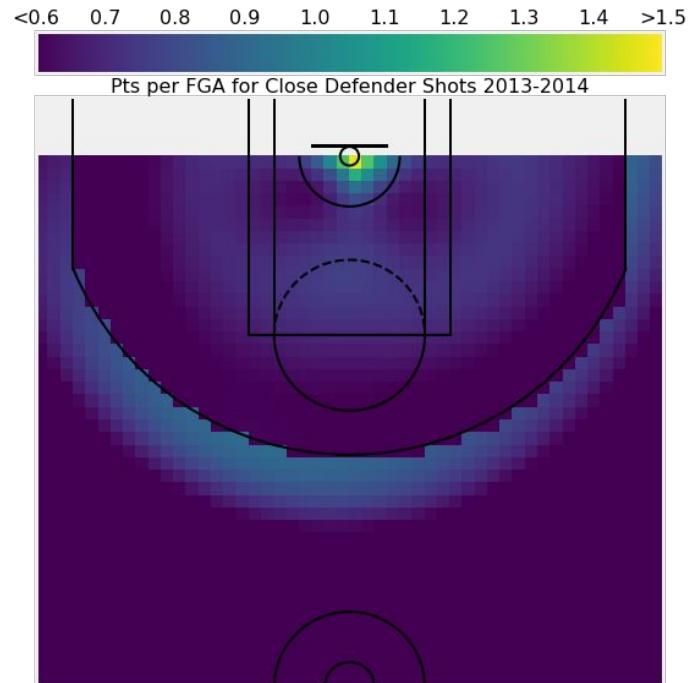
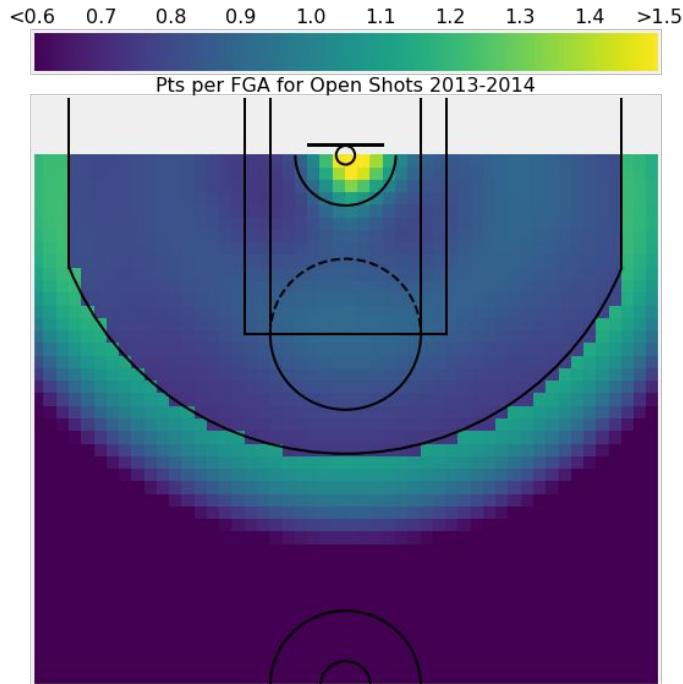
Shooting, Defense, and Expected Value

Finish Demo

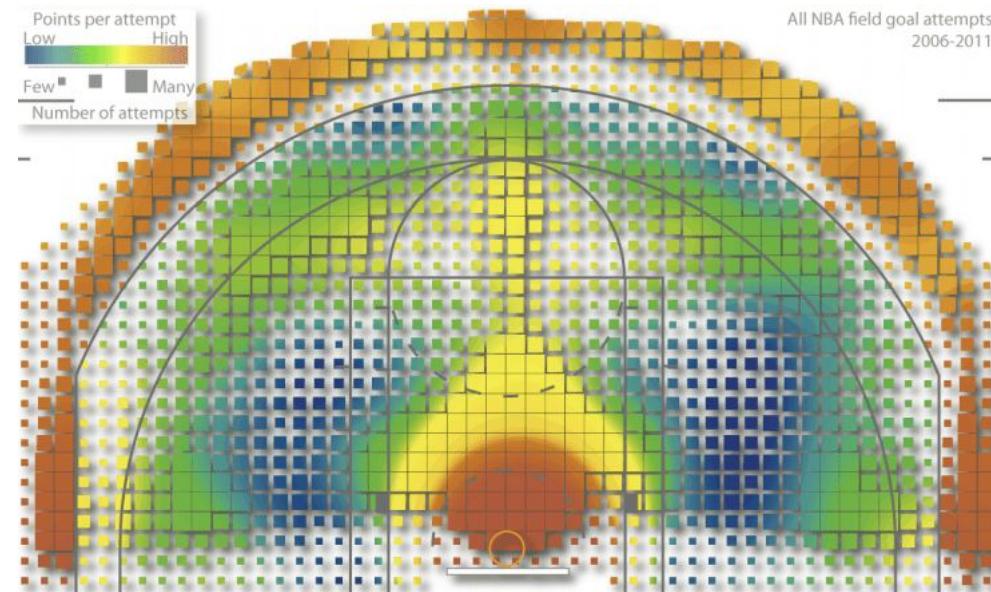
Smoothed Models Again



Smoothed Models Again



Goldsberry's Charts

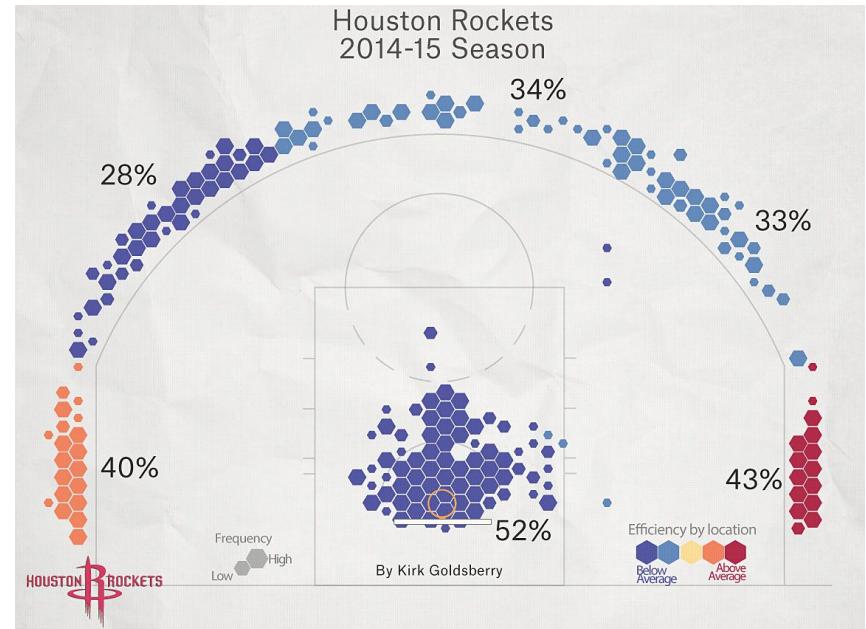


CourtVision: New Visual and Spatial Analytics for the NBA

Goldsberry's Charts: Location, Size, Color

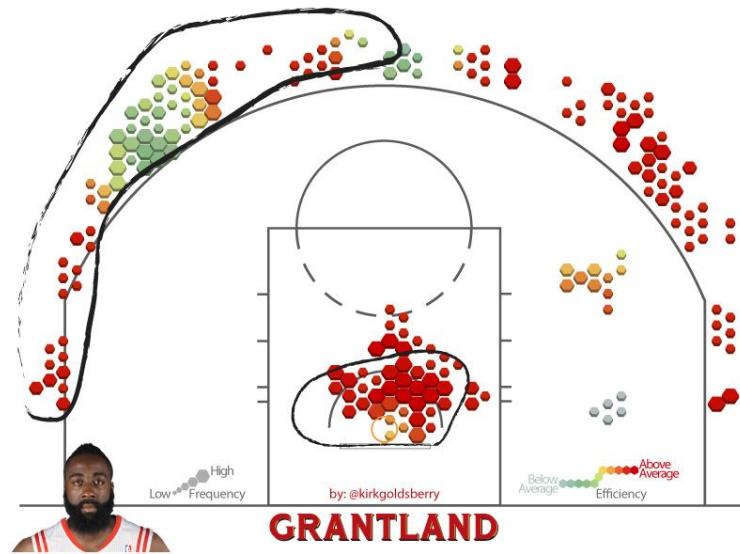
- *Location*: where the shots were taken
- *Size*: relatively how often the shots were taken
- *Color*: relatively how efficient the shots were (factoring in point value, naturally)

Blog post for producing the charts in Python:
www.eyalshafran.com/grantland_shotchart.html

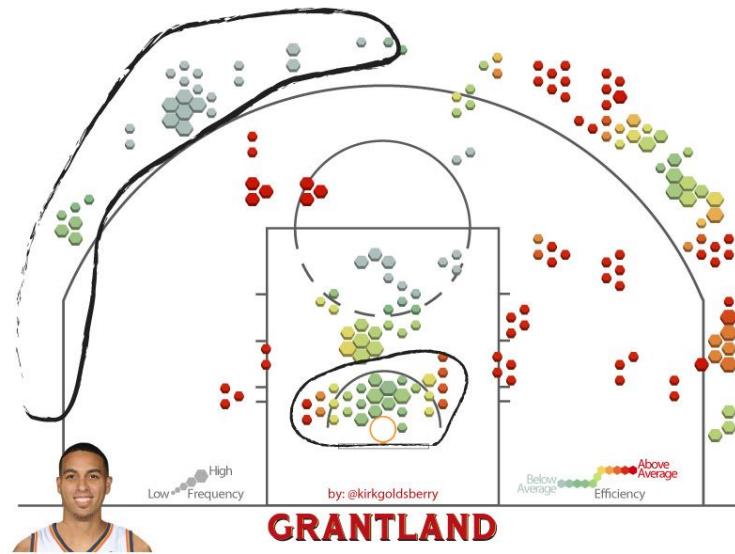


Goldsberry's Charts

James Harden



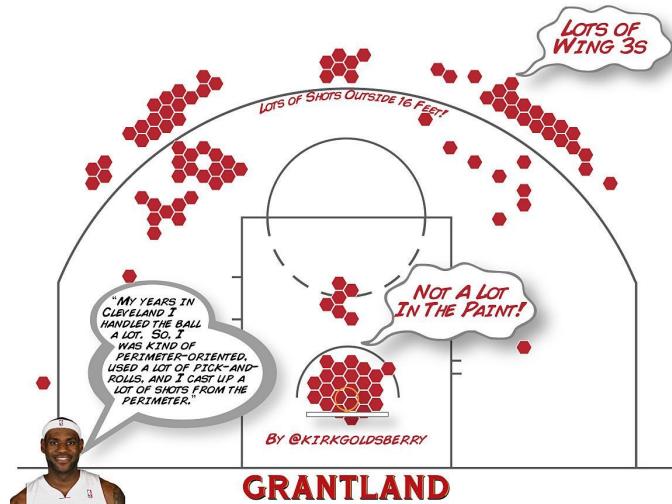
Kevin Martin



GRANTLAND

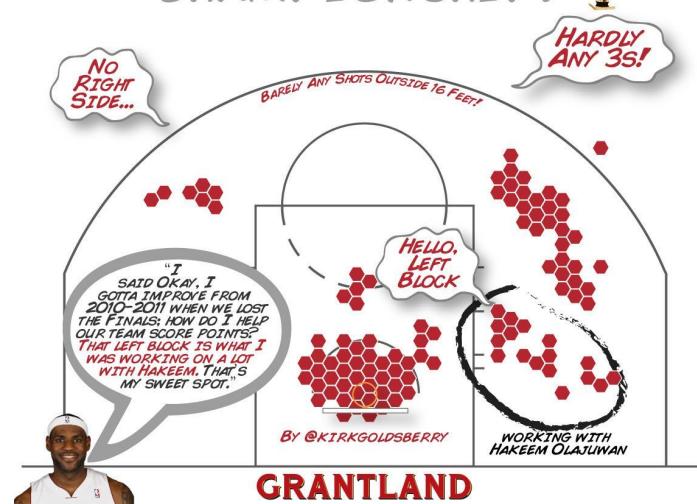
Goldsberry's Charts

LAST YEAR IN CLEVELAND



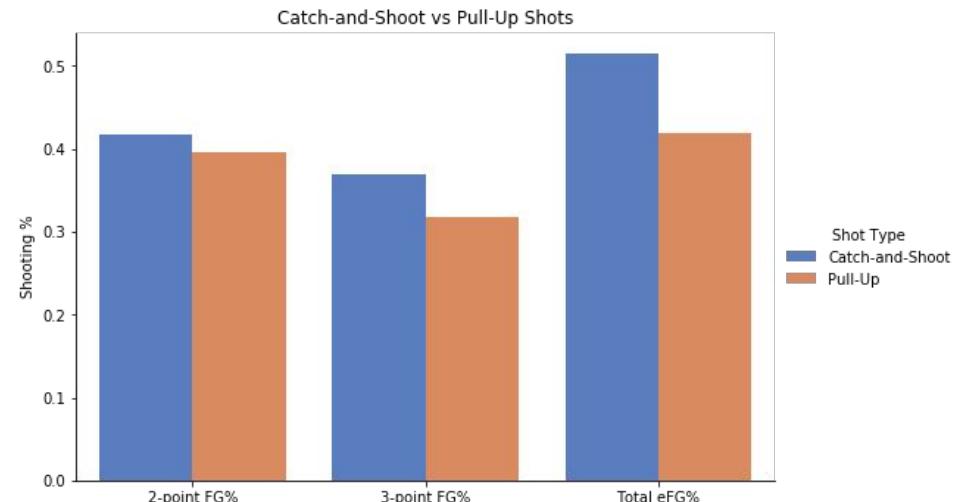
The Evolution of King James

SECOND YEAR IN MIAMI
CHAMPIONSHIP!



Taking it Further: Dribbling/Catch-and-Shoot

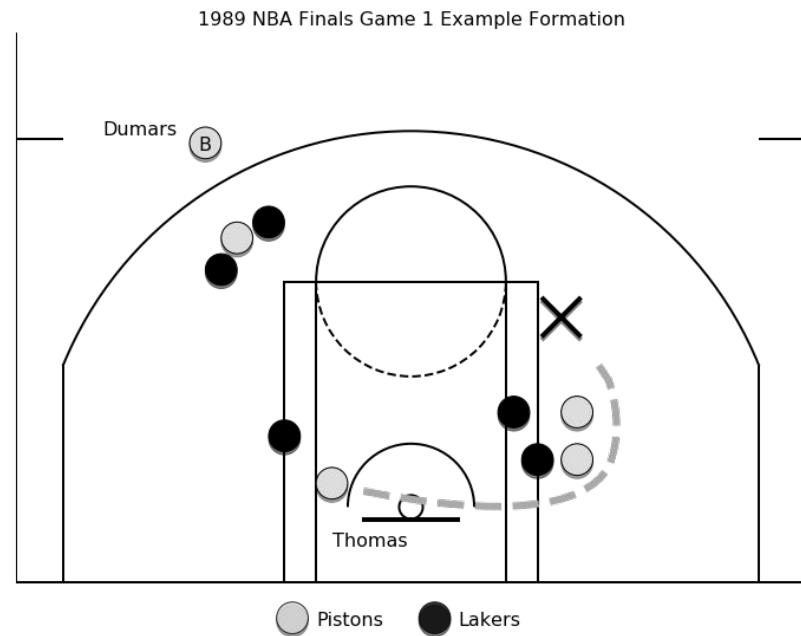
- Utilize more information about the style of the shot
- Catch-and-Shoot is a friend to any shooter



Recreation and extension of a chart from *Basketball Analytics: Spatial Tracking* by Stephen Shea

Taking it Further: Floor Spacing

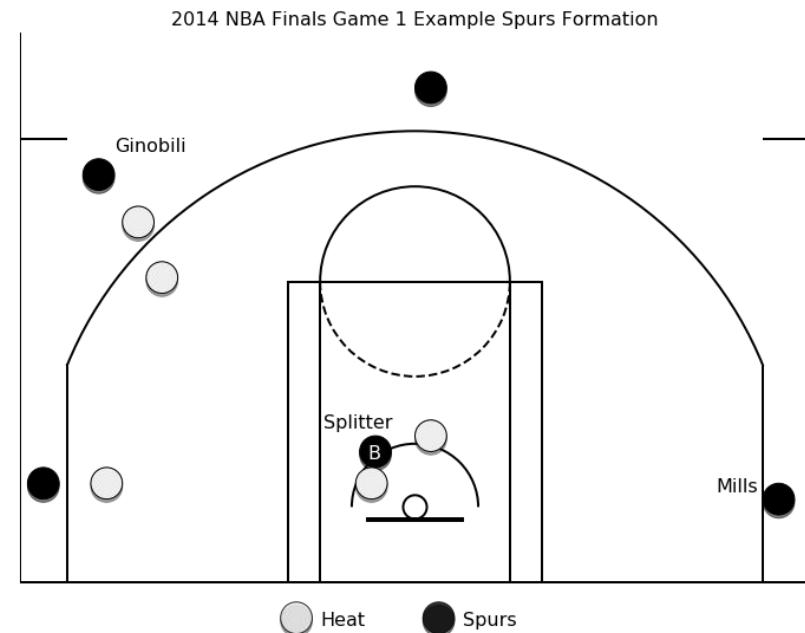
- Compact floor spacing from the 1989 Finals
- Noone set up in the corner for a 3 attempt
- The Pistons eventually took a mid-range shot (the X)



Recreation of a diagram from *Basketball Analytics: Spatial Tracking* by Stephen Shea

Taking it Further: Floor Spacing

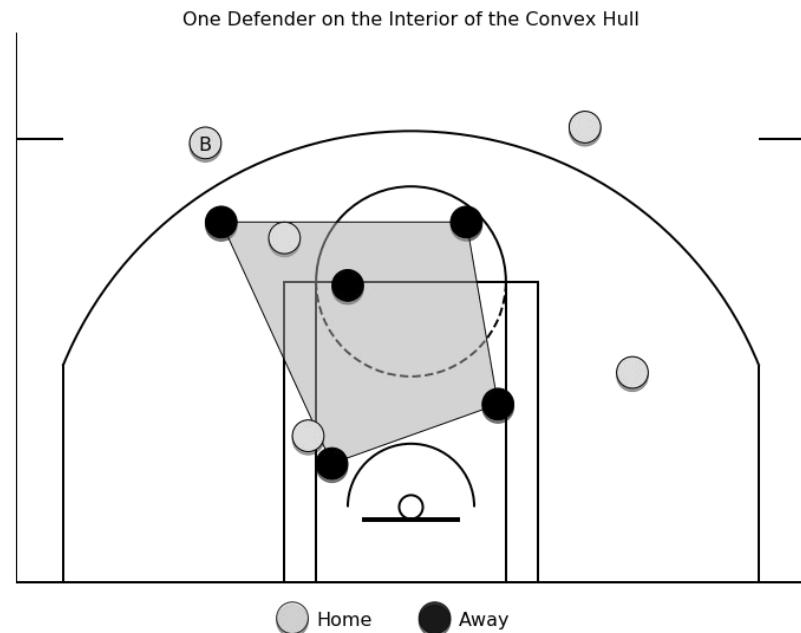
- Spread floor spacing from the 2014 Finals
- Screen and Roll to Splitter
- Spurs space the floor with many shooters



Recreation of a diagram from *Basketball Analytics: Spatial Tracking* by Stephen Shea

Taking it Further: Floor Spacing

- Quantifying spacing
- The extent to which a defense covers an area

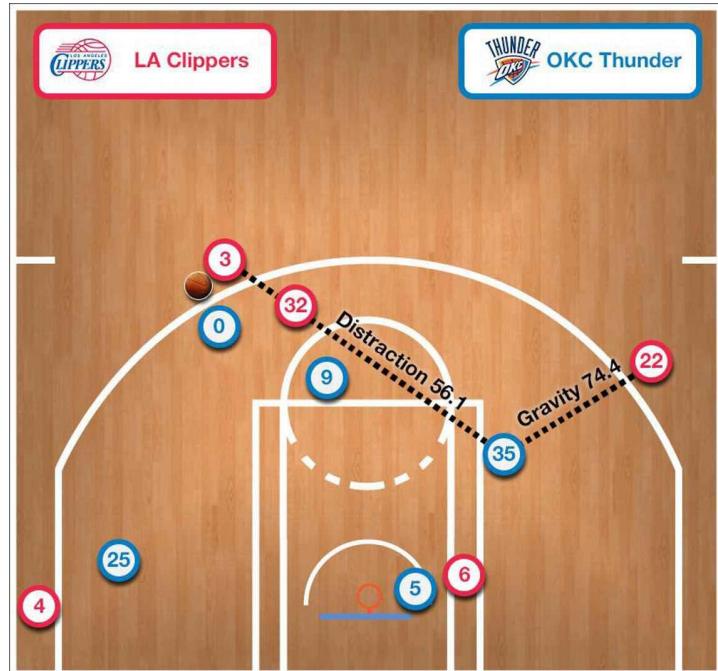


Recreation of a diagram from *Basketball Analytics: Spatial Tracking* by Stephen Shea

Taking it Further: Gravity

- How much respect a player receives from the defense

Explaining 'gravity' in basketball by Kevin Pelton



Summary

- Visualizing shot location and performance can be powerful
- The league is adapting to more 3s
- The midrange is death for an offense
 - ◆ Better to be underperforming in a high efficiency area (at the basket, corner 3) than overperforming in a low efficiency area (midrange)
- Defense matters and affects shot making
- We can also model shooting efficiency through expected values
- This is only the tip of the iceberg for spatial analysis of basketball