

# DS 501 CASE STUDY 2

## - Data Science in Sports Games

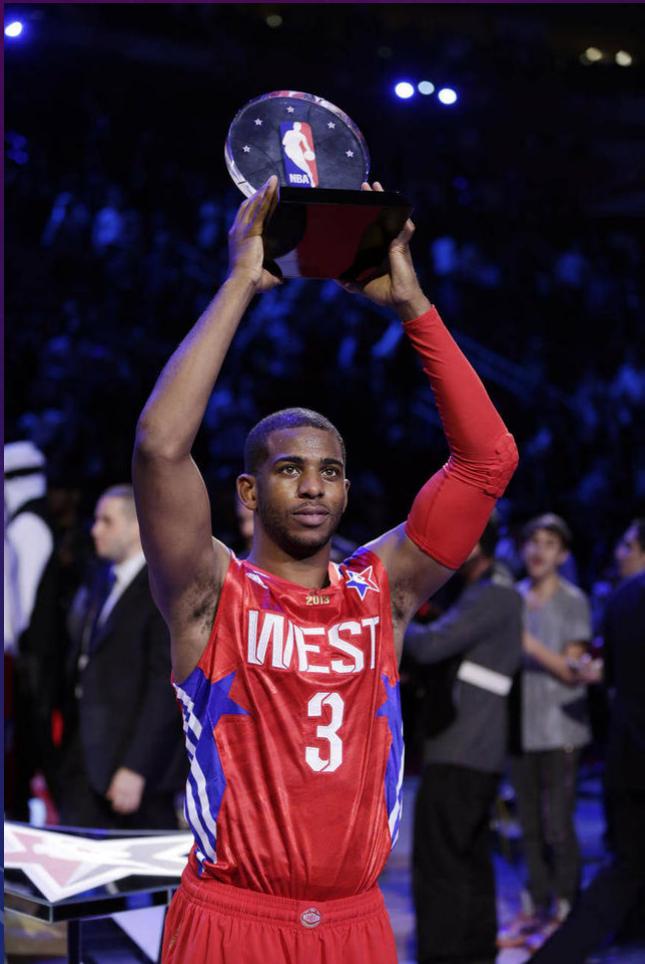
GROUP 6: QIAN WANG, ELIZABETH K KARPINSKI, JINGNAN XU, XIAOSHUAI LI

# DATA COLLECTION

- Obtained a list of players information through using py-Goldsberry.
- Accessed the API that the NBA uses for *stats.nba.com*.
  1. Collected data about individual NBA players via **Shots API**.
  2. **PlayerProfile API & PlayerDashPtShotLog API**: Help explore the data with detailed players information (Running additional experiments).



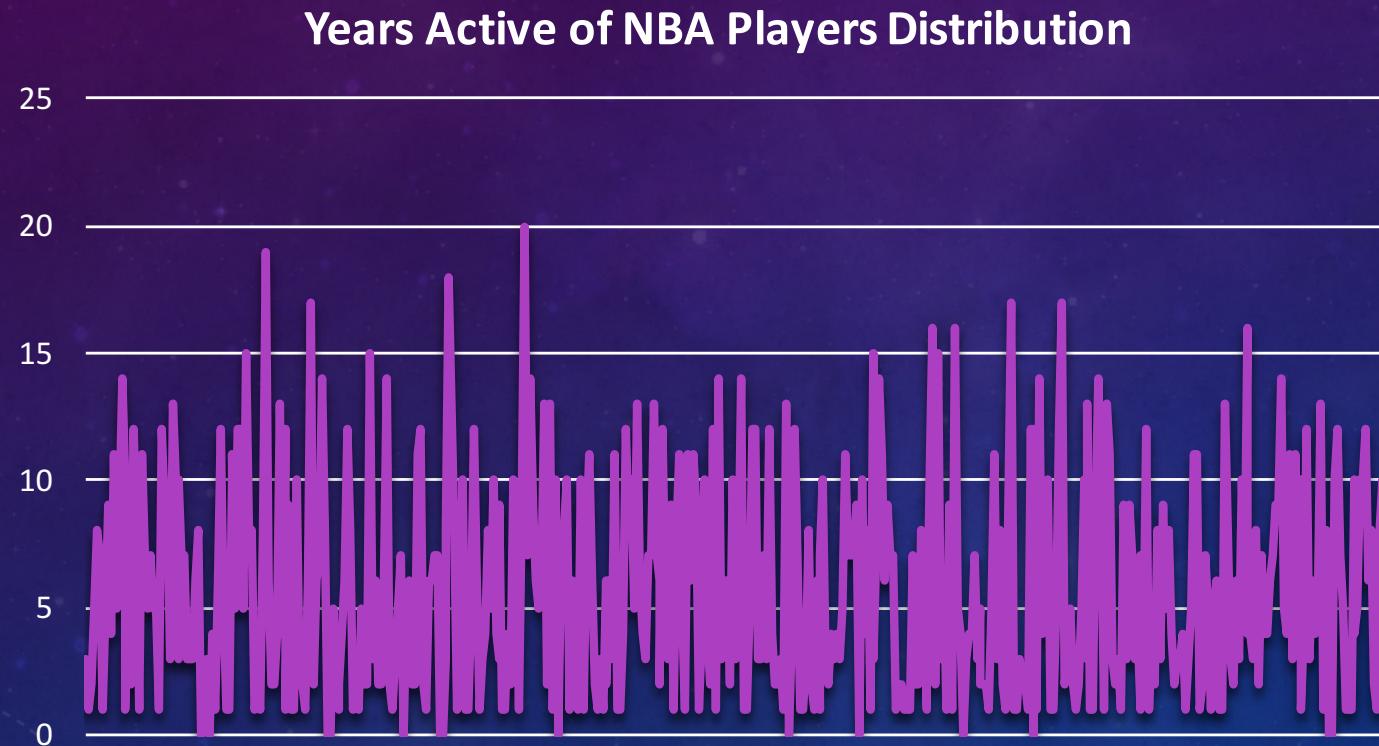
# OUR INTEREST IN THE TOPIC/ MOTIVATIONS



- **CHRIS PAUL, *We Choose You!***
- Chris Paul is interesting!
  - Tall: 6' & Weight: 175 lbs.
  - Older than average: more experienced but at a handicap of being less youthful
  - Olympic gold medal & NBA All-Star Game 2013 MVP
  - One of the highest paid athletes in the world
- We analyzed data on where he took his shots for the goal.
- Relative movements of the players
  - Focused on a **particular controversial moment** in a particular game: Chris Paul taking a serious risk by taking a shot himself, while being guarded by two defenders, despite three open teammates.

# DATA ANALYSIS & RESULTS

## - 2014 Player List



NBA player data:

- Full list of 493 players.
- Analyzed their active years  
(approximate distribution shown in figure)

# DATA ANALYSIS & RESULTS

## - 2014 Player List

NBA Players Start Year Distribution



- Helps to understand the general trend of NBA players' careers.
- Much easier to see how comparatively few players started before the year 2000.
- Show the relative proportions of NBA player start years (better sense of the relative proportion of player experiences.)

# DATA ANALYSIS & RESULTS

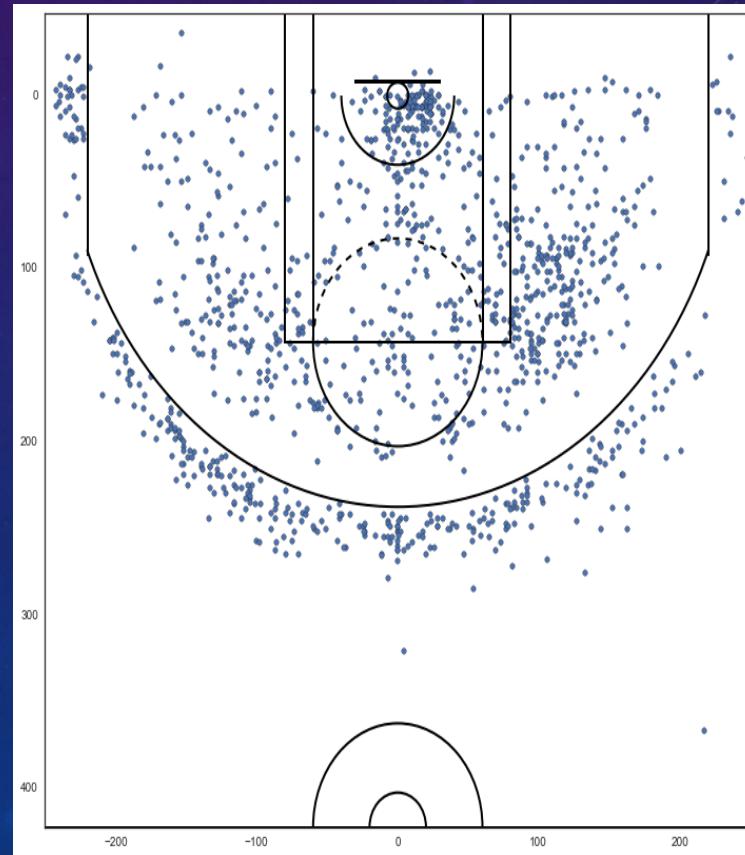
## - Chris Paul Shot Chart

Sample of Chris Paul Shot Chart

GRID_TYPE	GAME_ID	GAME_EVENT_ID	PLAYER_ID	PLAYER_NAME	TEAM_ID	TEAM_NAME	PERIOD
Shot Chart Detail	21400020	42	101108	Chris Paul	1610612746	Los Angeles Clippers	1
Shot Chart Detail	21400020	48	101108	Chris Paul	1610612746	Los Angeles Clippers	1
Shot Chart Detail	21400020	97	101108	Chris Paul	1610612746	Los Angeles Clippers	1
Shot Chart Detail	21400020	229	101108	Chris Paul	1610612746	Los Angeles Clippers	2
Shot Chart Detail	21400020	268	101108	Chris Paul	1610612746	Los Angeles Clippers	2
Shot Chart Detail	21400020	299	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	302	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	303	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	320	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	332	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	350	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	371	101108	Chris Paul	1610612746	Los Angeles Clippers	3
Shot Chart Detail	21400020	456	101108	Chris Paul	1610612746	Los Angeles Clippers	4
Shot Chart Detail	21400020	467	101108	Chris Paul	1610612746	Los Angeles Clippers	4
Shot Chart Detail	21400020	469	101108	Chris Paul	1610612746	Los Angeles Clippers	4
Shot Chart Detail	21400020	490	101108	Chris Paul	1610612746	Los Angeles Clippers	4
Shot Chart Detail	21400020	493	101108	Chris Paul	1610612746	Los Angeles Clippers	4
Shot Chart Detail	21400020	515	101108	Chris Paul	1610612746	Los Angeles Clippers	4
Shot Chart Detail	21400026	10	101108	Chris Paul	1610612746	Los Angeles Clippers	1
Shot Chart Detail	21400026	39	101108	Chris Paul	1610612746	Los Angeles Clippers	1

Chris Paul Shot Chart in court

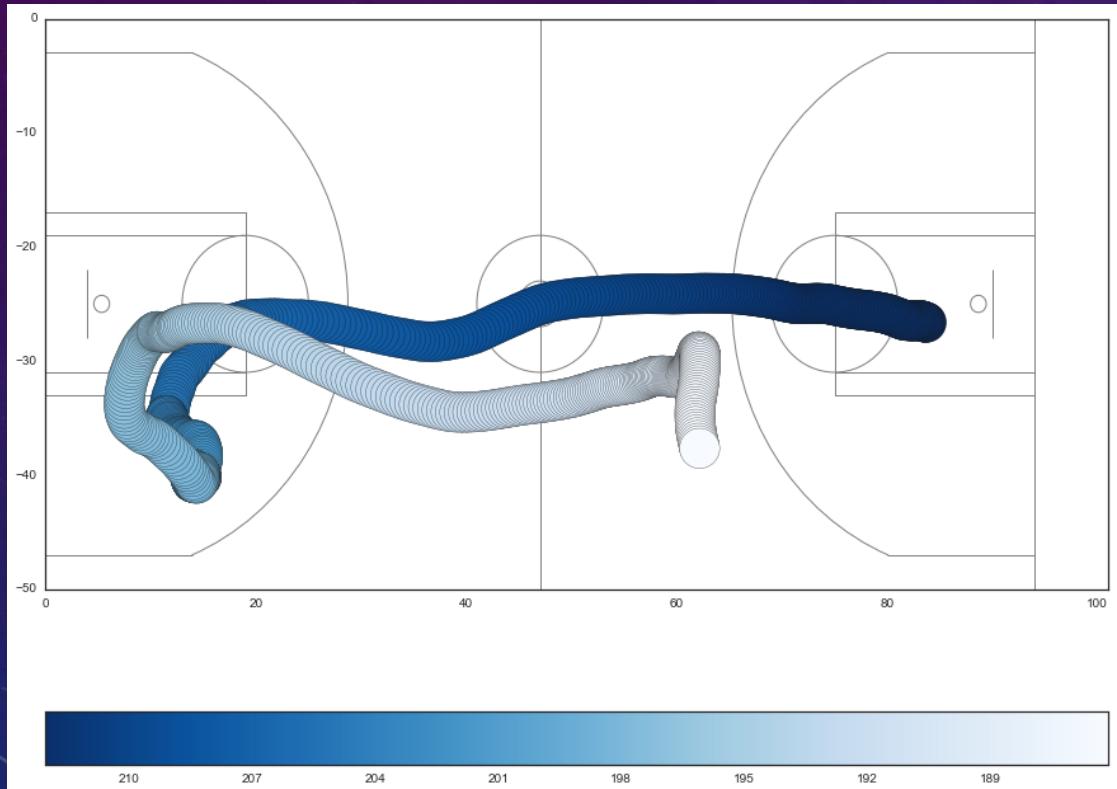
A better representation of the locations with respect to real world items.



# DATA ANALYSIS & RESULTS

## - Chris Paul Movements

Chris Paul Movements during the moment



Looked at in depth was the specific moments and movements of game.

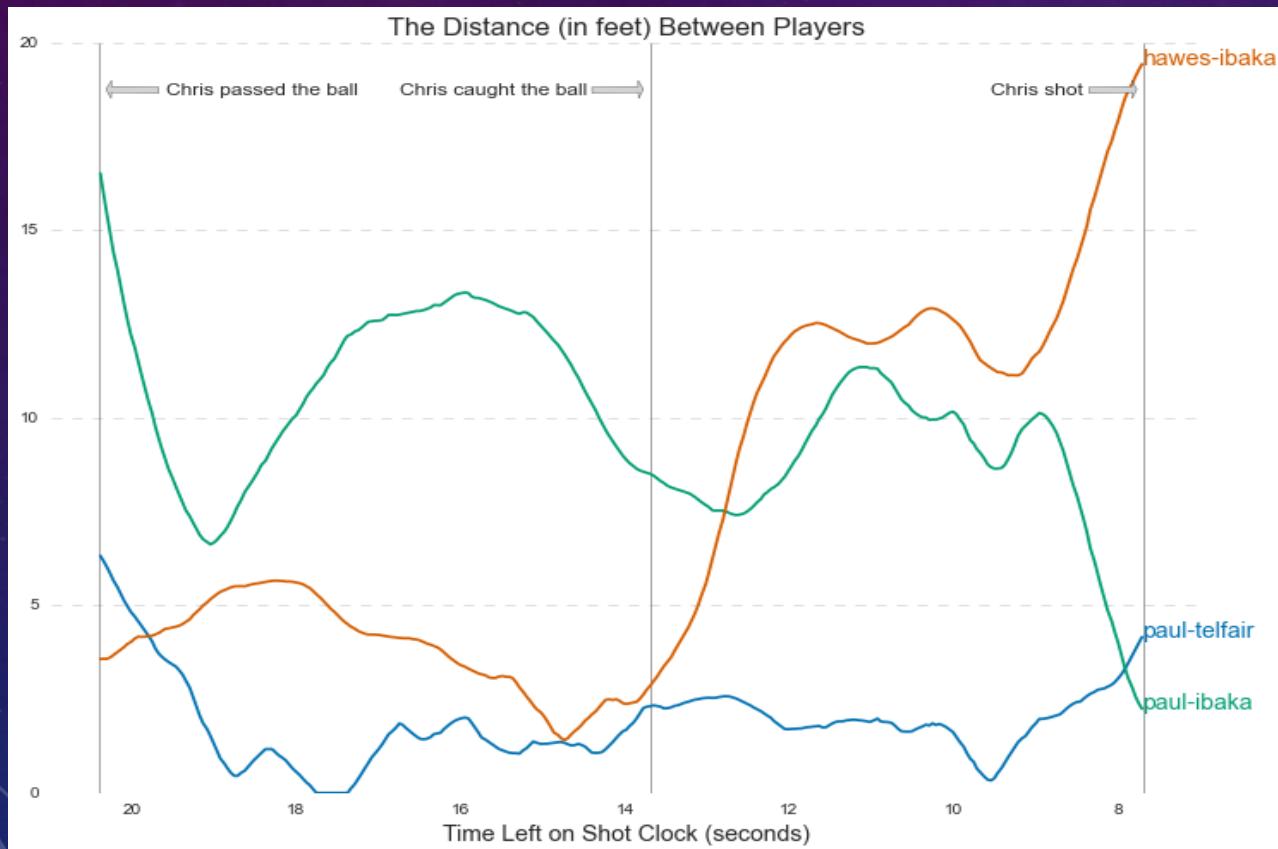
The data could tell us precisely what was happening: Chris Paul makes a risky shot instead of passing to open teammates.

A plot of Chris Paul's movements during this moment is shown in figure.

# DATA ANALYSIS & RESULTS

## - Chris Paul Movements

Distance between players during the moment



**A** plot of the distance between players: Chris Paul was being heavily guarded by two opposing players, while his teammate (seen in the orange) was wide open, at a considerable distance from the nearest defender.

**W**hy did Chris Paul make the shot instead of attempting pass to his teammates?

- Chris Paul was in a position where he habitually made shot.

# DATA ANALYSIS & RESULTS

## - 2014 Players average statistics

	name	avg_defender_distance	avg_dribbles	avg_shot_distance	avg_touch_time
0	Adams, Jordan	4.545349	1.174419	13.524419	1.784884
1	Acy, Quincy	4.924159	0.318043	14.162080	1.544343
2	Adams, Jordan	4.545349	1.174419	13.524419	1.784884
3	Adams, Steven	2.731122	0.375000	4.996684	1.508163
4	Adrien, Jeff	2.704651	0.511628	6.527907	1.793023

Touch time is an useful measurement:

- It measures how many seconds the player has control of the ball before they take the shot.
- According to the blog *Basketball Analytics*, a shorter touch time is better than a longer one: It is correlated with more successful scoring: shots taken after more than two seconds are significantly less likely to score.
- This correlation helps to make touch time a valuable way of assessing the offensive strength of a player.

# DATA ANALYSIS & RESULTS

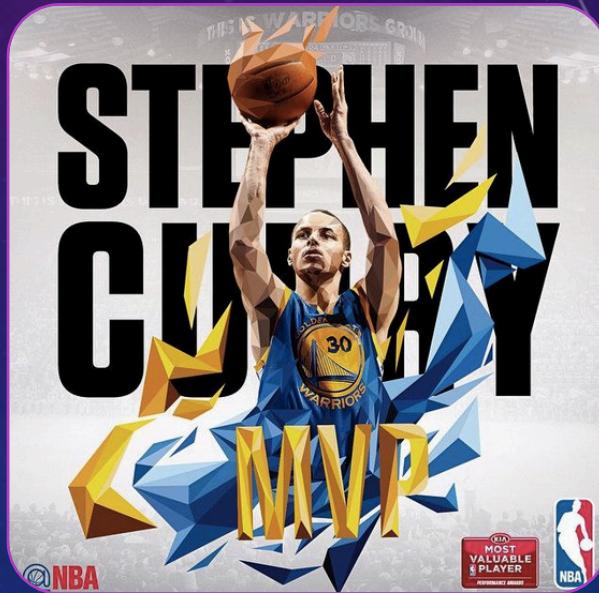
## - MVP

### Predict MVP in 2014-2015 Season

- The modified formula is:
  - $(Points + Rebounds + Assists + Steals + Blocks) - (Missed Field Goals + Missed Free Throws + Turnovers)$ .

	name	PTS	REB	AST	STL	BLK	PFD	FTA	FTM	FGA	FGM	TOV	Score
103	Cousins, DeMarcus	24.1	12.7	3.6	1.5	1.7	8.8	9.2	7.2	18.1	8.4	4.3	36.4
116	Davis, Anthony	24.4	10.2	2.2	1.5	2.9	5.5	6.8	5.5	17.6	9.4	1.4	35.8
469	Westbrook, Russell	28.1	7.3	8.6	2.1	0.2	6.6	9.8	8.1	22.0	9.4	4.4	34.2
196	Harden, James	27.4	5.7	7.0	1.9	0.7	6.7	10.2	8.8	18.1	8.0	4.0	33.9
234	James, LeBron	25.3	6.0	7.4	1.6	0.7	6.0	7.7	5.4	18.5	9.0	3.9	31.3
138	Durant, Kevin	25.4	6.6	4.1	0.9	0.9	4.9	6.3	5.4	17.3	8.8	2.7	30.7
111	Curry, Stephen	23.8	4.3	7.7	2.0	0.2	3.7	4.2	3.9	16.8	8.2	3.1	29.7
360	Paul, Chris	19.1	4.6	10.2	1.9	0.2	3.6	3.9	3.5	14.3	6.9	2.3	29.5
189	Griffin, Blake	21.9	7.6	5.3	0.9	0.5	5.9	6.4	4.6	17.1	8.6	2.3	29.5
166	Gasol, Pau	18.5	11.8	2.7	0.3	1.9	4.2	4.7	3.8	14.8	7.3	2.0	29.0

Candidates for MVP in 2014-2015 season



The winner of MVP for the 2014-2015 season, ranks 7th in our result.

# DATA SCIENCE IN HEALTHCARE INDUSTRY

- U.S. health care industry is an overpriced, inefficient mess.
- By 2020, health care expenses is estimated to rise to nearly 20% of GDP. The country ranks 37th out of developed economies in life expectancy and other measures of health.
- Big data is the answer!
- That situation puts data scientists in a prime position.
- Used wisely, big data has the potential to help physicians make better decisions across the board – from personalized treatments to preventive care.



# OPPORTUNITIES IN HEALTH CARE DATA SCIENCE



Personalized Medicine



Genomics



Predictive Analytics & Preventive Measures



Patient Monitoring & Home Devices



Self-Motivated Care



Disease Modeling & Mapping



The Ultimate EHR

# DATA RISKS AND REGULATIONS



- The Challenges
- Patient Privacy



THANK YOU!