

Are performance and salary in the NBA related?

Helen Belogolova
Peter Elliott
Yuting Wang

Motivation

- Kobe Bryant made \$25M in 2011
- Main question: Is this salary justified?
- Based on historical data from 2009, 2010, 2011 we can look at:
 - Salaries for each player per year
 - Specific statistics per player such as
 - Positions (Forward, Center, Guard)
 - # of Wins in the season
 - PPG (points per game)
 - RPG (rebounds per game)
 - MPG (minutes per game)
 - APG (assists per game)

Data Collection 1

Data available online on: <http://www.basketball-reference.com>

Roster

[Glossary](#) · [SHARE](#) · [Embed](#) · [CSV](#) · [PRE](#) · [LINK](#) · ?

No.	Player	Pos	Ht	Wt	Birth Date	Exp	College
9	Matt Barnes	F	6-7	235	March 9, 1980	7	University of California, Los Angeles
5	Steve Blake	G	6-3	172	February 26, 1980	7	University of Maryland
12	Shannon Brown	G	6-4	205	November 29, 1985	4	Michigan State University
24	Kobe Bryant	G	6-6	200	August 23, 1978	14	
17	Andrew Bynum	C	7-0	285	October 27, 1987	5	
45	Derrick Caracter	F	6-9	275	May 4, 1988	R	University of Texas at El Paso
3	Devin Ebanks	F	6-9	215	October 28, 1989	R	West Virginia University
2	Derek Fisher	G	6-1	200	August 9, 1974	14	University of Arkansas at Little Rock
16	Pau Gasol	F-C	7-0	227	July 6, 1980	9	
10	Trey Johnson	G	6-5	218	August 30, 1984	1	Jackson State University
7	Lamar Odom	F	6-10	220	November 6, 1979	11	University of Rhode Island
50	Theo Ratliff	C-F	6-10	225	April 17, 1973	15	University of Wyoming
1	Joe Smith	F	6-10	225	July 26, 1975	15	University of Maryland
18	Sasha Vujacic	G	6-7	193	March 8, 1984	6	
4	Luke Walton	F	6-8	235	March 28, 1980	7	University of Arizona
15	Metta World Peace	F	6-6	244	November 13, 1979	11	St. John's University

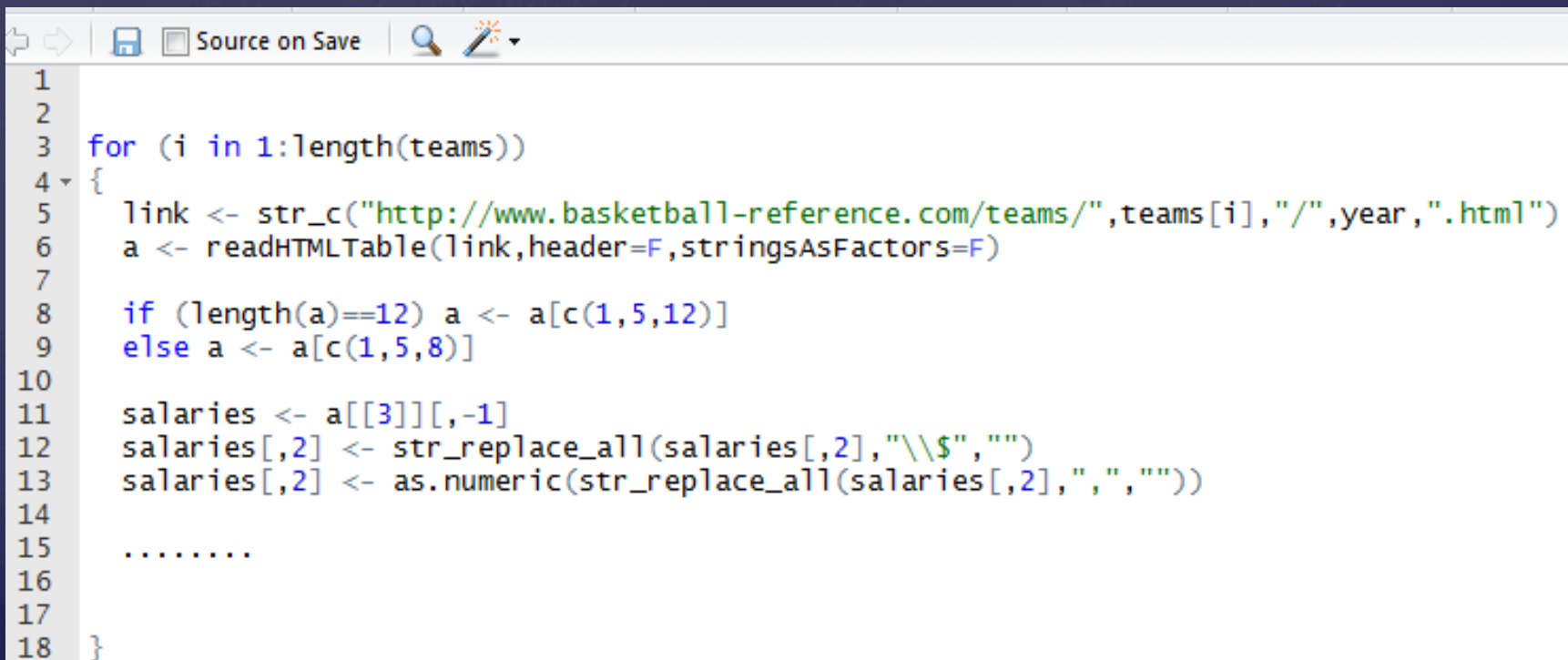
Salaries

Rk	Player	Salary
1	Kobe Bryant	\$24,806,250
2	Pau Gasol	\$17,823,000
3	Andrew Bynum	\$13,700,000
4	Lamar Odom	\$8,200,000
5	Metta World Peace	\$6,322,320
6	Luke Walton	\$5,260,000
7	Steve Blake	\$4,000,000
8	Derek Fisher	\$3,700,000
9	Shannon Brown	\$2,200,000
10	Matt Barnes	\$1,765,000
11	Theo Ratliff	\$1,352,181
12	Joe Smith	\$1,352,181
13	Devin Ebanks	\$473,604
14	Derrick Caracter	\$473,604

```
link <- str_c("http://www.basketball-reference.com/teams/" teams[i], "/" year ".html")
a <- readHTMLTable(link,header=F,stringsAsFactors=F)
```

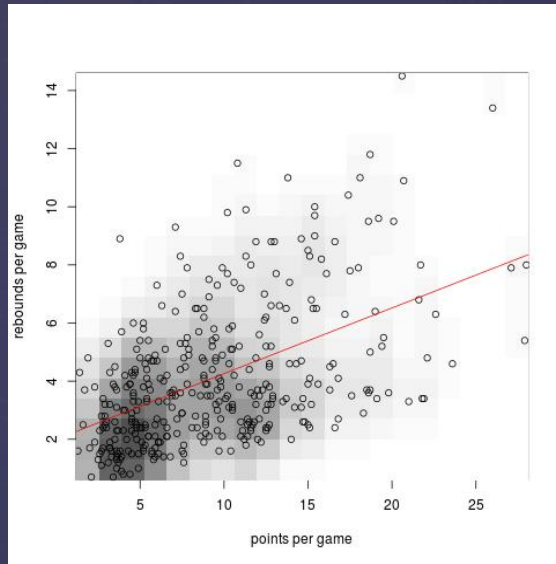
Data Collection 2

How we got the data into R – extracting data from online HTML tables:

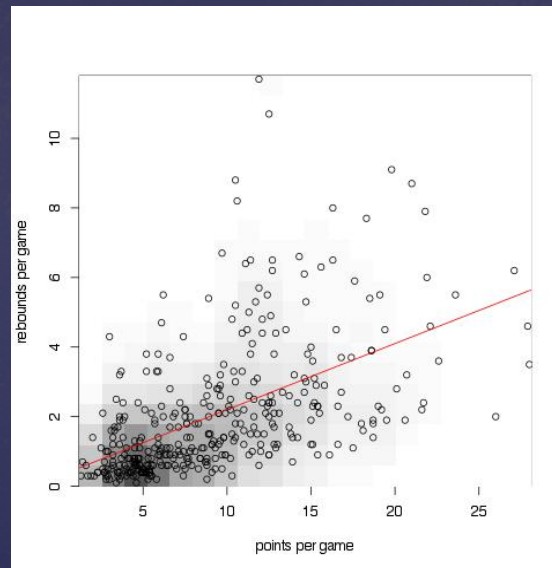
A screenshot of an R script editor window. The window has a title bar with a file icon, a 'Source on Save' checkbox, and a search icon. The script is written in R and is enclosed in a for loop that iterates over a vector named 'teams'. Inside the loop, a URL is constructed using 'str_c', and an HTML table is read using 'readHTMLTable'. The table is then processed to extract salary data, with some cleaning steps using 'str_replace_all' and 'as.numeric'. The script ends with an ellipsis and a closing brace.

```
1
2
3 for (i in 1:length(teams))
4 {
5   link <- str_c("http://www.basketball-reference.com/teams/" ,teams[i], "/" ,year, ".html")
6   a <- readHTMLTable(link,header=F,stringsAsFactors=F)
7
8   if (length(a)==12) a <- a[c(1,5,12)]
9   else a <- a[c(1,5,8)]
10
11   salaries <- a[[3]][,-1]
12   salaries[,2] <- str_replace_all(salaries[,2],"\\$", "")
13   salaries[,2] <- as.numeric(str_replace_all(salaries[,2],",", ""))
14
15   .....
16
17
18 }
```

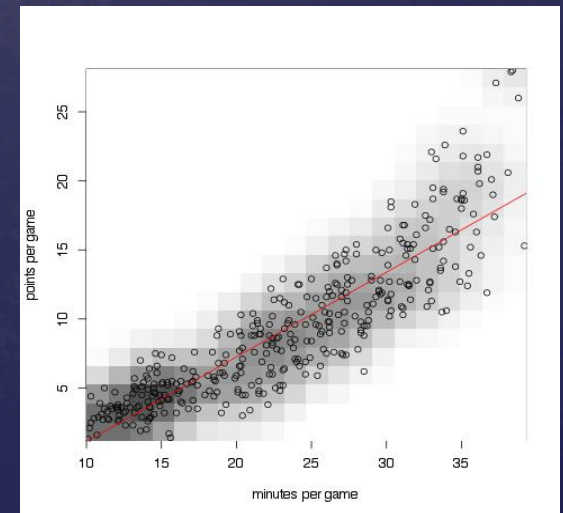
Player's performance \neq Athletic Ability



PPG vs RPG ($p < 2e-16$)

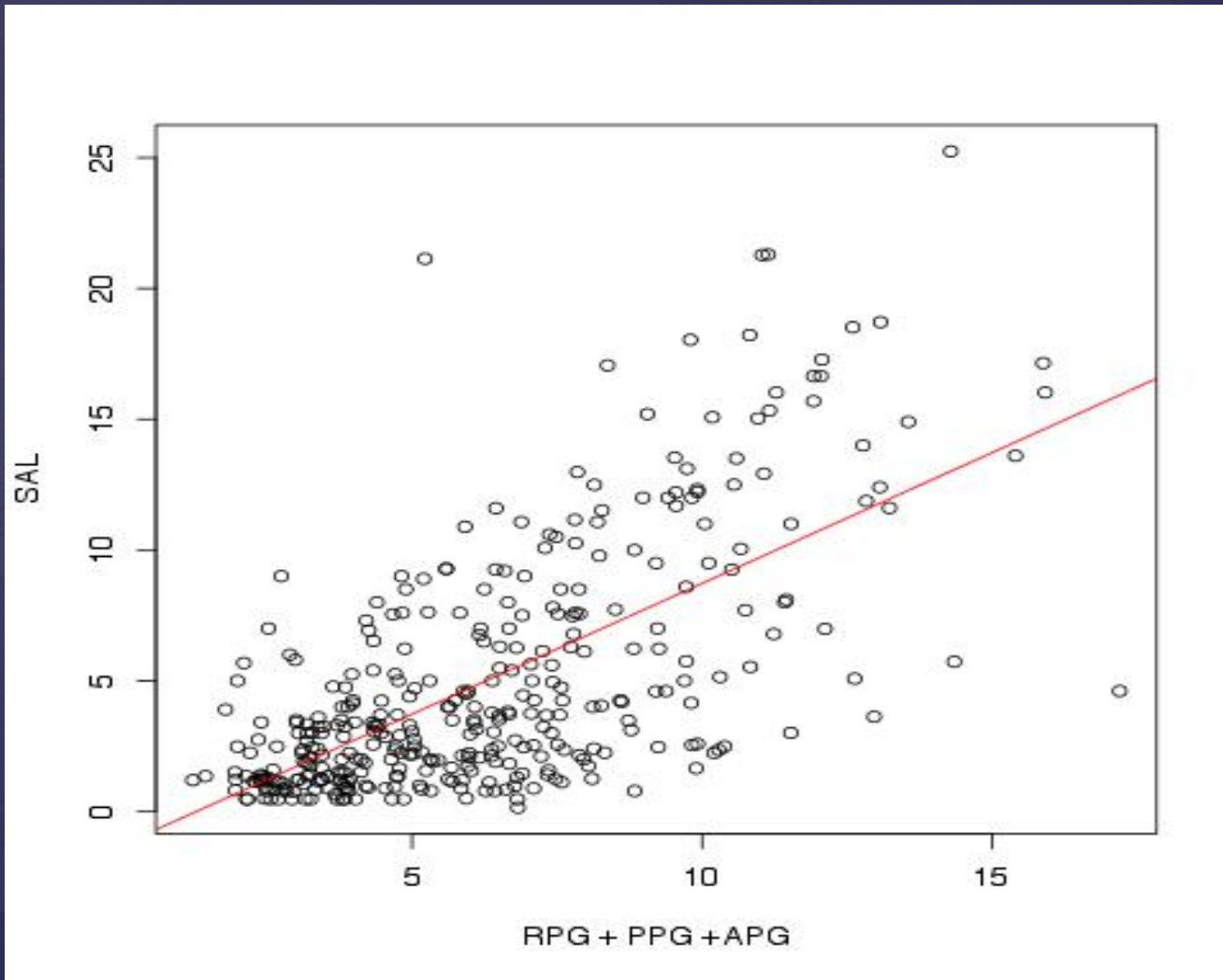


PPG vs APG ($p < 2e-16$)



MPG vs PPG ($p < 2e-16$)

$SAL \sim PPG + RPG + APG$



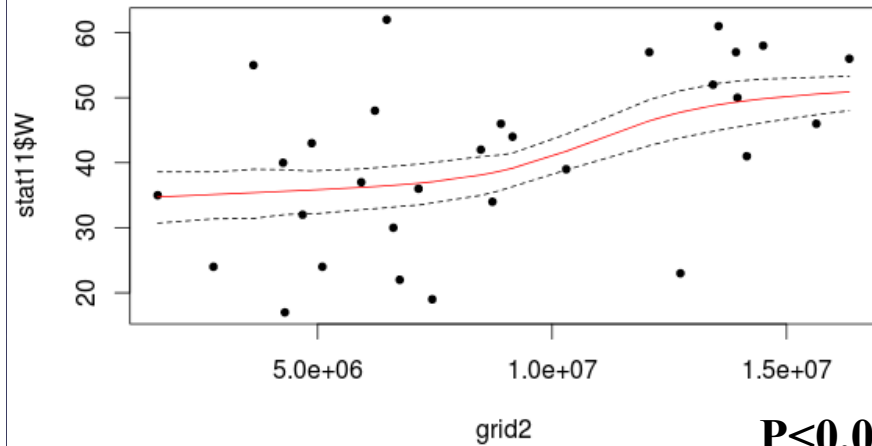
PPG ($p < 3.74e-11$)

RPG ($p < 2.13e08$)

APG ($p < 0.009$)

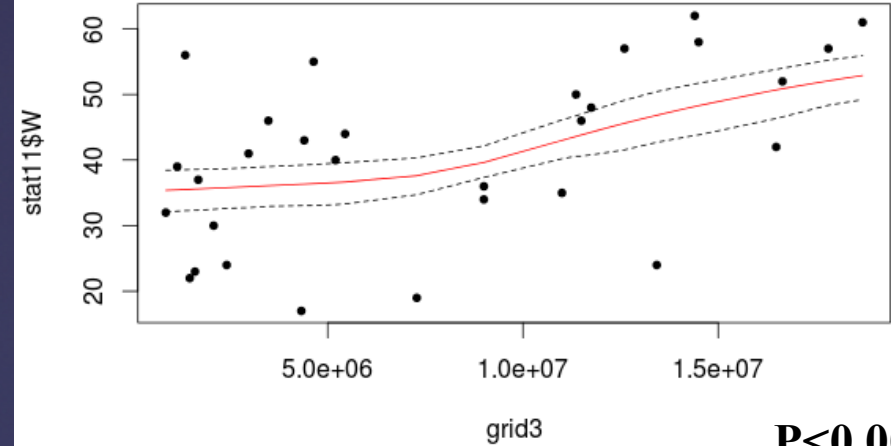
Salary and Wins relationship (2011)

Forward Salaries vs Wins



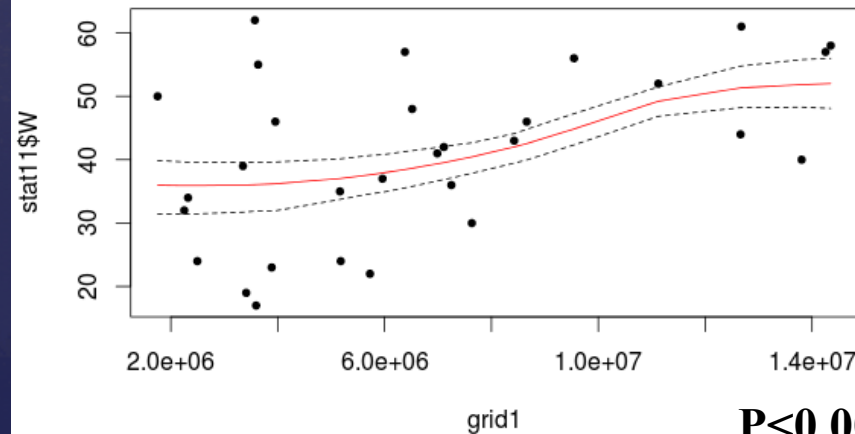
P<0.005

Center Salaries vs Wins



P<0.003

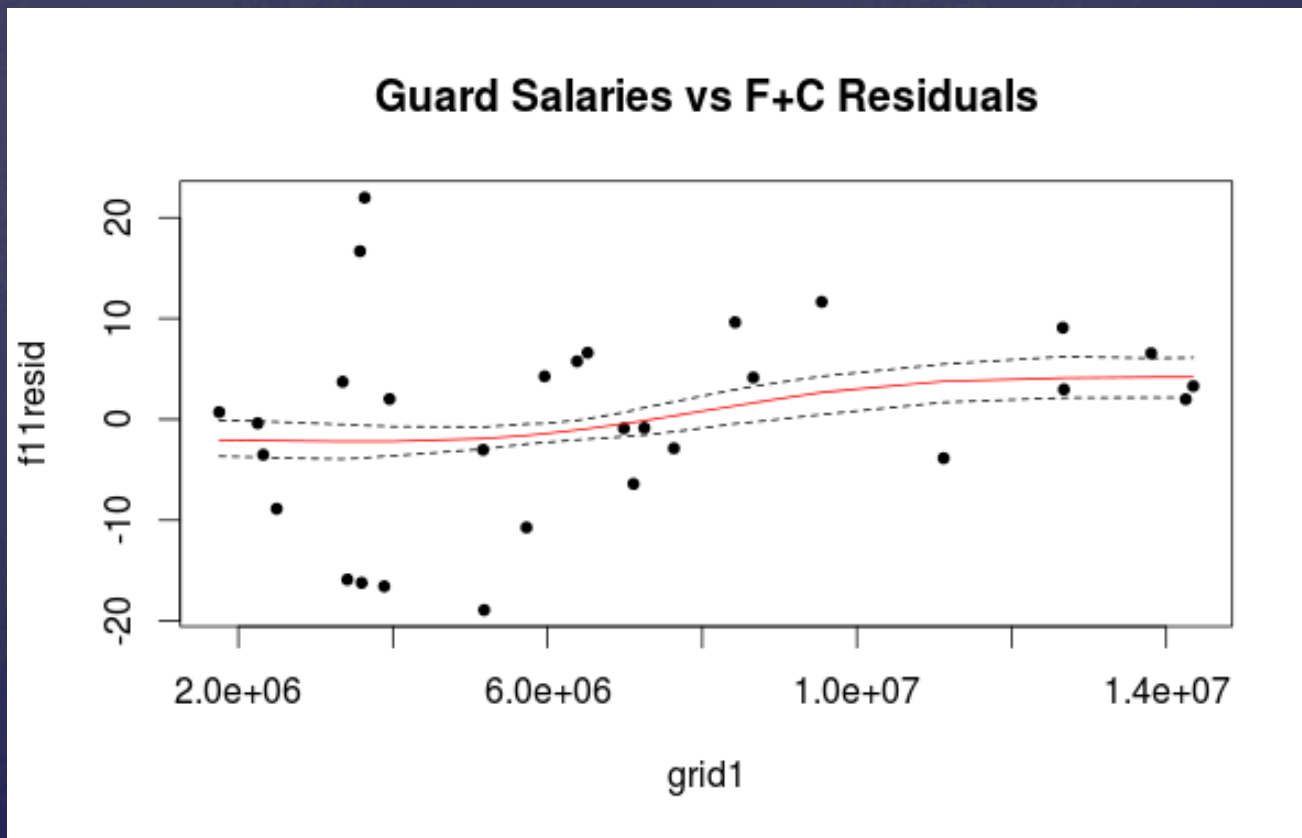
Guard Salaries vs Wins



P<0.008

Salary and Wins relationship

Does position matter? Likely not



Summary of Results

Statistically significant linear relationships between:

- ✓ Performance variables RPG, PPG, MPG, APG
- ✓ Salary and individual player performance/ability as measured by PPG, RPG, and APG
- ✓ Salary and team performance as measured by wins

Limitations of Current Analysis/Future Directions

- Tracking team roster changes
- Outlier players (e.g. Kobe Bryant)
- Outlier teams (e.g. Lakers)
- Other individual differences between players such as age
- Are contracts based on past performance?

THANKS!