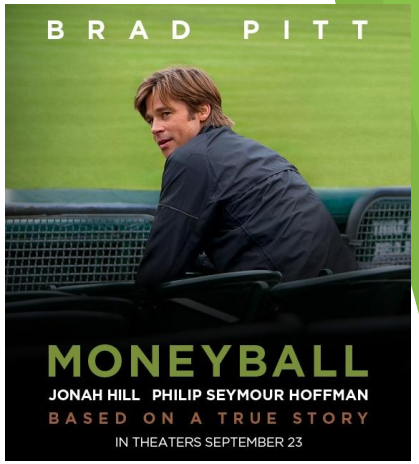
Predicting Postseason Success in the NBA, MLB, and NFL

Bobby He, Brian Li, Connor Regan, Richard Shen

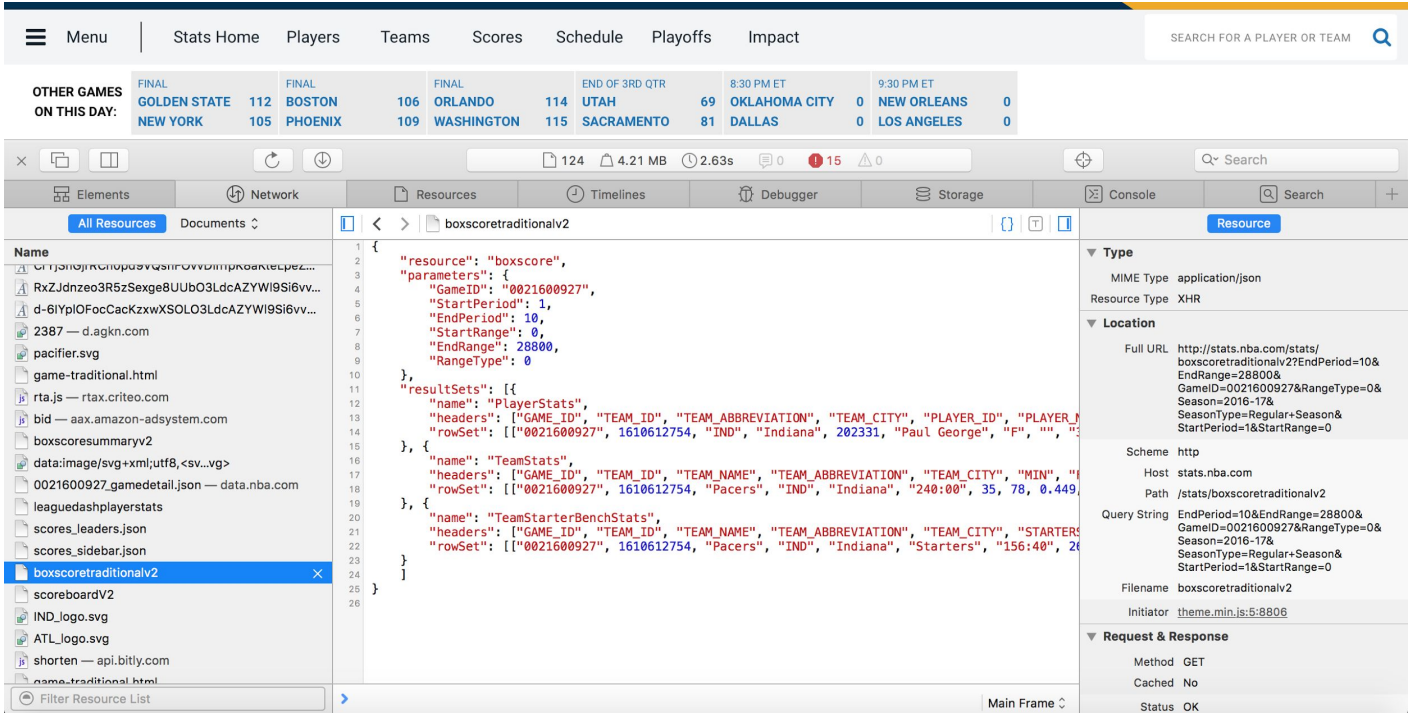
Motivation


Americans Lose 119 Billion Through Gambling
Biggest gambling losses by country in 2013 (in billion U.S. dollars)
U.S. $119.0b
China $76.0b
Japan $31.4b
Italy $23.9b
Britain $19.9b
Australia $18.4b
Germany $14.8b
France $13.3b
Canada $13.0b
Spain $12.4b
Source: H2 Gambling Capital via The Economist
Mashable
statista


BRAD PITT
MONEYBALL
JONAH HILL PHILIP SEYMOUR HOFFMAN
BASED ON A TRUE STORY
IN THEATERS SEPTEMBER 23


TOYOTA CENTER
TOYOTA CENTER
ROCKETS

Data Collection Methods



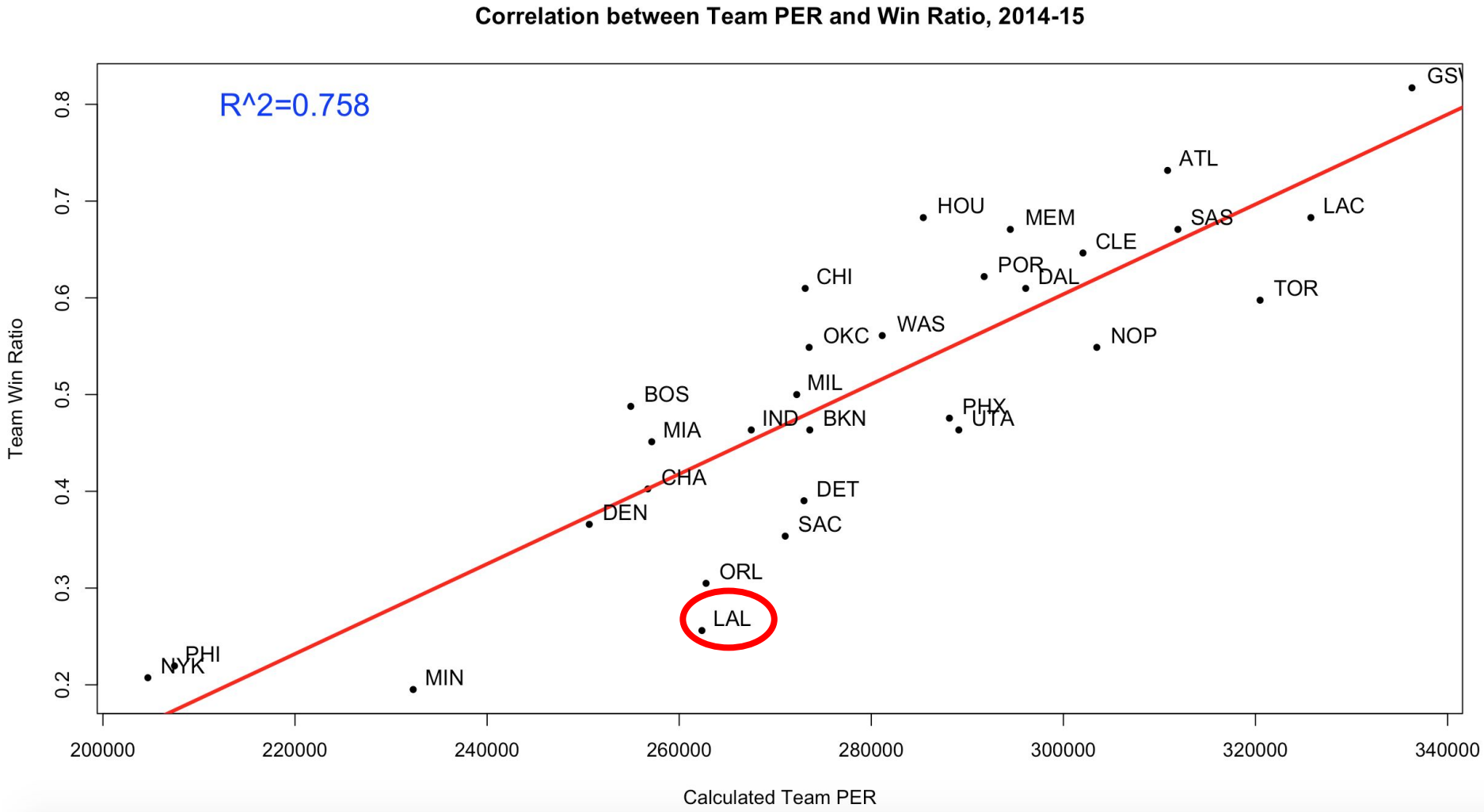
Linear Regression - Using Team PER to Predict Playoff Teams

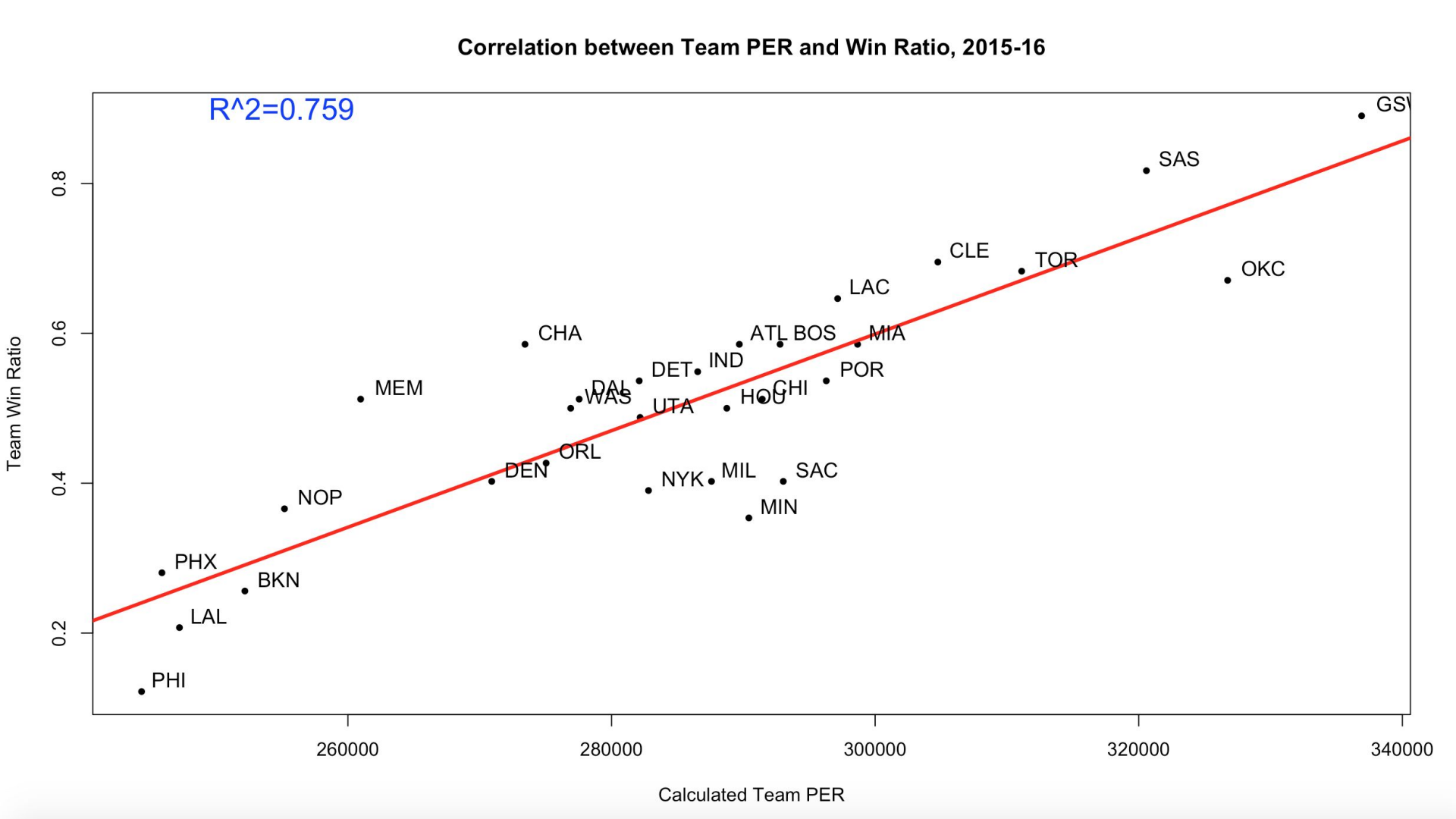
- Player Efficiency Rating (PER) is a widely available metric originally calculated by John Hollinger. It is a metric used to rate players with a single number such that the league average is 15.
- We create a new metric called Team Player Efficiency Rating (Team PER) which is a weighted average of PER for the 12 most active players for each NBA team.

$$\text{Team PER} = \sum_{i=1}^{12}[(\text{MP}_i)(\text{PER}_i)]$$

Where MP = minutes played in a season

- This player-centric metric accounts for many variables such as…
  - A single player's contribution to his team
  - The effect of "all-star" players
  - Roster changes due to injury, mid-season trades, or coaching decisions
- We compare this metric to win ratio (simply the ratio of each teams wins and losses in a regular season) in order to measure its effectiveness.
- In the following slides Team PER is used to predict which teams will make the NBA playoffs for the 2011-12 through 2016-17 seasons, boasting 100% accuracy in predicting the top 4 seeded teams in each conference and 89% accuracy overall.


Correlation between Team PER and Win Ratio, 2014-15
R^2=0.758
Team Win Ratio
Calculated Team PER
0.2
0.3
0.4
0.5
0.6
0.7
0.8
200000
220000
240000
260000
280000
300000
320000
340000
GSW
ATL
HOU
MEM
SAS
LAC
CLE
POR
DAL
CHI
TOR
WAS
OKC
NOP
MIL
BOS
PHX
UTA
IND
BKN
MIA
CHA
DET
DEN
SAC
ORL
LAL
PHI
NYK
MIN


Correlation between Team PER and Win Ratio, 2015-16
R^2=0.759
Team Win Ratio
Calculated Team PER
0.2
0.4
0.6
0.8
260000
280000
300000
320000
340000
GSW
SAS
CLE
TOR
OKC
LAC
CHA
ATL BOS
MIA
IND
DET
POR
MEM
DAL
CHI
WAS
HOU
UTA
ORL
DEN
NYK
MIL
SAC
NOP
MIN
PHX
BKN
LAL
PHI

Season					Season				
2011-12	Actual EC	Predicted EC	Actual WC	Predicted WC	2014-2015	Actual EC	Predicted EC	Actual WC	Predicted WC
1	CHI	CHI 253877	SAS	OKC 260523	1	ATL	TOR 320476	GSW	GSW 336289
2	MIA	MIA 251806	OKC	SAS 253398	2	CLE	ATL 310846	HOU	LAC 325763
3	IND	PHI 249420	LAL	LAC 249391	3	CHI	CHI 307874	LAC	SAS 311919
4	BOS	IND 244012	MEM	UTA 248077	4	TOR	CLE 302036	POR	NOP 303474
5	ATL	ATL 242936	LAC	LAL 243624	5	WAS	WAS 281136	MEM	DAL 296073
6	ORL	ORL 232708	DEN	DEN 243412	6	MIL	BKN 273592	SAS	MEM 294471
7	NYK	NYK 219837	DAL	PHX 238490	7	BOS	DET 272990	DAL	POR 291749
8	PHI	BOS 213525	UTA	DAL 235771	8	BKN	MIL 272233	NOP	UTA 289107
2012-13	Actual EC	Predicted EC	Actual WC	Predicted WC	2015-2016	Actual EC	Predicted EC	Actual WC	Predicted WC
1	MIA	MIA 338151	OKC	OKC 336245	1	CLE	TOR 311118	GSW	GSW 336908
2	NYK	BKN 303754	SAS	DEN 326531	2	TOR	CLE 304757	SAS	OKC 326744
3	IND	NYK 299704	DEN	LAC 322063	3	MIA	MIA 298667	OKC	SAS 320583
4	BKN	IND 285194	LAC	SAS 304600	4	ATL	BOS 292786	LAC	LAC 297154
5	CHI	NOP 282768	MEM	LAL 302604	5	BOS	CHA 291437	POR	POR 296294
6	ATL	MIL 282358	GSW	UTA 301676	6	CHA	ATL 289697	DAL	SAC 293030
7	BOS	CHI 282122	LAL	HOU 292760	7	IND	MIL 287583	MEM	MIN 290420
8	MIL	ATL 278504	HOU	GSW 289922	8	DET	IND 286537	HOU	HOU 288746
2013-14	Actual EC	Predicted EC	Actual WC	Predicted WC	2016-2017	Actual EC	Predicted EC	Actual WC	Predicted WC
1	IND	MIA 319437	SAS	DAL 323840	1		TOR 243334		GSW 255098
2	MIA	TOR 291083	OKC	OKC 318545	2		BOS 226423		HOU 245325
3	TOR	WAS 286615	LAC	LAC 318206	3		CLE 223718		SAS 238105
4	CHI	ATL 284893	HOU	SAS 314618	4	???	WAS 222316	???	LAC 234133
5	WAS	CHI 279594	POR	HOU 314283	5		IND 220815		MIN 230704
6	BKN	IND 274294	GSW	POR 312229	6		MIL 220563		DEN 226816
7	CHA	CHA 274041	MEM	PHX 305993	7		CHI 220144		UTA 219905
8	ATL	BKN 271046	DAL	GSW 299413	8		DET 220089		POR 215618

Markov Chain Monte Carlo

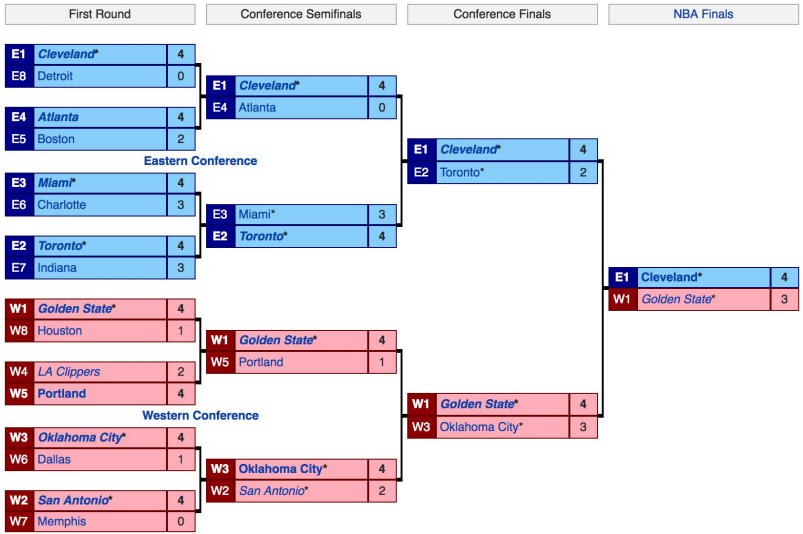
- Distribution sampling method involving simulating the construction of a Markov chain
  - Equilibrium distribution of the simulated Markov chain is the desired distribution

- Applications towards sports: estimating the probability of a particular team being the “top team” in a given set (playoff teams)
  - NBA, MLB, and NFL
  - Probabilistic rankings

- Markov chain: each state represents the current belief as to which team in a particular playoffs is the actual “top team”
  - State transition probabilities: comparing the current “top team” against a randomly chosen “check team”
  - Head-to-head scoring differential

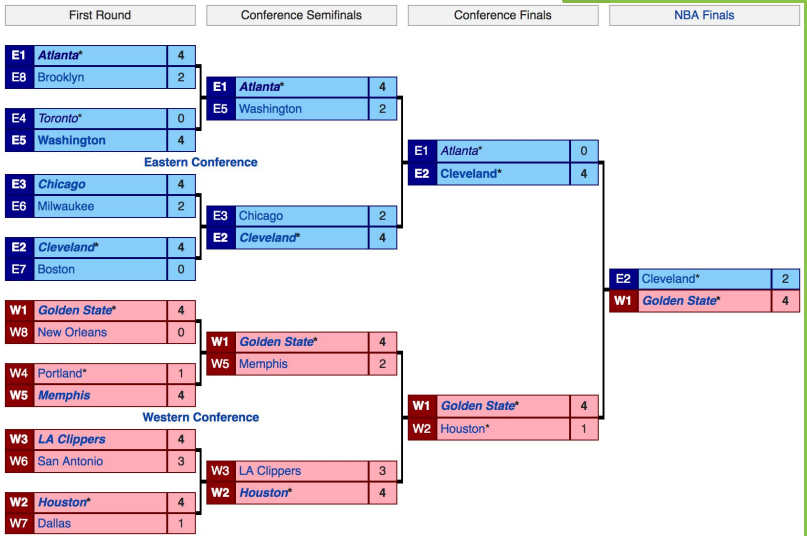
Results: NBA

```
[1] "Year: 2015-2016"

       MEM        DAL        HOU        MIA        CHA        ATL        LAC        BOS
0.01419716 0.02199560 0.02519496 0.02659468 0.02959408 0.03199360 0.03199360 0.03259348
       POR        IND        DET        TOR        CLE        OKC        SAS        GSW
0.03499300 0.03559288 0.04239152 0.05278944 0.08198360 0.08558288 0.12157568 0.33093381
```



Source: wikipedia.org/wiki/2016_NBA_Playoffs

```
[1] "Year: 2014-2015"

       BOS        MIL        BKN        TOR        WAS        NOP        POR        DAL
0.01299740 0.01619676 0.01659668 0.02959408 0.03079384 0.03659268 0.04119176 0.04419116
       LAC        HOU        CHI        MEM        SAS        CLE        ATL        GSW
0.04519096 0.05478904 0.05638872 0.05958808 0.08858228 0.09058188 0.10397920 0.27274545
```



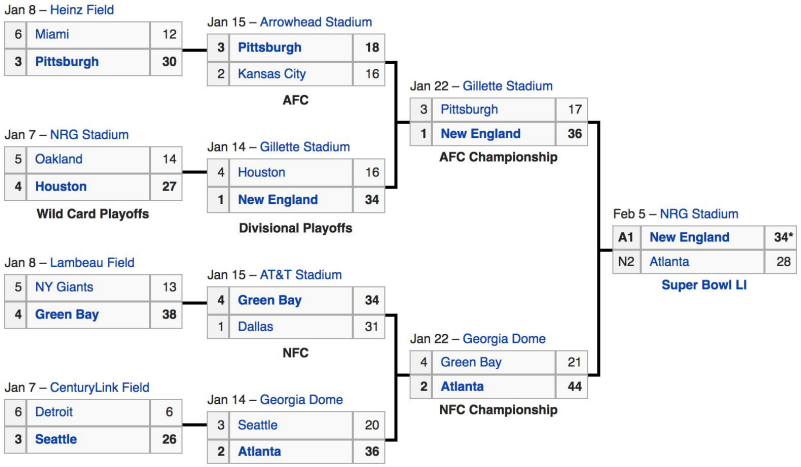
Source: wikipedia.org/wiki/2015_NBA_Playoffs

Results: MLB and NFL

```
[1] "Year: 2016"

       CLE        NYM        WSH        TEX         SF        CHC        BAL        LAD
0.04039192 0.05658868 0.06918616 0.06938612 0.07178564 0.08778244 0.12377524 0.12597481
       TOR        BOS
0.16836633 0.18676265
```

Source: wikipedia.org/wiki/2016_Major_League_Baseball_season

```
[1] "Year: 2016"

       HOU        SEA         GB        NYG        DAL        ATL         KC        MIA
0.07958408 0.08078384 0.08178364 0.08178364 0.08198360 0.08298340 0.08318336 0.08478304
        NE        OAK        PIT        DET
0.08498300 0.08498300 0.08618276 0.08698260
```



Source: wikipedia.org/wiki/2016-17_NFL_playoffs

Discussion

Regular Season as Indicator of Playoff Success

- NBA
  - 82 game regular season: 30 teams and 15 make playoffs
    - Over past five years, top seed average record was .807 and bottom seed was .482
  - Seven game series and least stochastic playoff results
    - Average of 100 possessions per team
    - Steady accumulation of points over game
- MLB
  - 162 game regular season: 30 teams and 10 make playoffs
    - Over past five years, top seed average record was .611 and bottom seed was .542
  - Seven game series but each individual game is a very stochastic result
    - Average of 34 at bats per team
    - Number of runs close to a random walk
    - Greater variance game to game because scoring comes in bunches
- NFL
  - 16 game regular season: 32 teams and 12 make playoffs
    - Over past five years, top seed average record was .8375 and bottom seed was .5375
  - One game matchups each round of playoffs
    - Average of 12 possessions per team
    - Big play potential
- Conclusion - NBA produces postseason champions most aligned with regular season performance

Quality of Champions Across Sports

Pythagorean Expectation

$$\text{Win Ratio} = \frac{\text{runs scored}^2}{\text{runs scored}^2 + \text{runs allowed}^2} = \frac{1}{1 + (\text{runs allowed}/\text{runs scored})^2}$$

Simple Rating System

Compares runs/goals/points to league average for the season, while taking into account a team's strength of schedule

	Pythagorean Expectation			Simple Rating System		
League Champions from 1995-2013	Mean Rank	Median Rank	Mean Z-Score	Mean Rank	Median Rank	Mean Z-Score
NFL	5.3	4.5	1.34	4.7	2.5	1.37
MLB	5.8	4	1.10	6.5	7	1.05
NBA	3.3	2	1.47	3.3	2	1.48

Source: Harvard Sports Analysis: Undeserving Champions: Examining Variance in the Postseason

Bibliography

- https://en.wikipedia.org/wiki/2016_NBA_Playoffs
- https://en.wikipedia.org/wiki/2015_NBA_Playoffs
- https://en.wikipedia.org/wiki/2016-17_NFL_playoffs
- https://en.wikipedia.org/wiki/2016_Major_League_Baseball_season
- http://harvardsportsanalysis.org/2013/09/undeserving-champions-examining-variance-in-the-postseason/

Appendix: MCMC Results I

```
[1] "Year: 2011-2012"

       ORL        NYK        PHI        UTA        ATL        BOS        DAL        LAC
0.01459708 0.02399520 0.02619476 0.03179364 0.03199360 0.03399320 0.03639272 0.05138972
       IND        MIA        LAL        MEM        CHI        DEN        OKC        SAS
0.05178964 0.05758848 0.05838832 0.06398720 0.11197760 0.11297740 0.12517497 0.16776645
[1] "Year: 2012-2013"

       BOS        MIL        BKN        GSW        ATL        CHI        LAL        HOU
0.02559488 0.02739452 0.03039392 0.03079384 0.03279344 0.03459308 0.03759248 0.04659068
       IND        NYK        LAC        MEM        OKC        SAS        DEN        MIA
0.04899020 0.07058588 0.07938412 0.08198360 0.08258348 0.08538292 0.10717856 0.17816437
[1] "Year: 2013-2014"

       ATL        CHA        WAS        CHI        DAL        MEM        TOR        BKN
0.02139572 0.03039392 0.03559288 0.03979204 0.04139172 0.04419116 0.04819036 0.04939012
       GSW        LAC        POR        IND        HOU        SAS        MIA        OKC
0.05178964 0.06658668 0.07238552 0.07478504 0.08398320 0.10457908 0.10637872 0.12917417
[1] "Year: 2014-2015"

       BOS        MIL        BKN        TOR        WAS        NOP        POR        DAL
0.01299740 0.01619676 0.01659668 0.02959408 0.03079384 0.03659268 0.04119176 0.04419116
       LAC        HOU        CHI        MEM        SAS        CLE        ATL        GSW
0.04519096 0.05478904 0.05638872 0.05958808 0.08858228 0.09058188 0.10397920 0.27274545
[1] "Year: 2015-2016"

       MEM        DAL        HOU        MIA        CHA        ATL        LAC        BOS
0.01419716 0.02199560 0.02519496 0.02659468 0.02959408 0.03199360 0.03199360 0.03259348
       POR        IND        DET        TOR        CLE        OKC        SAS        GSW
0.03499300 0.03559288 0.04239152 0.05278944 0.08198360 0.08558288 0.12157568 0.33093381
```

Appendix: MCMC Results II

```
[1] "Year: 2012"

       STL        ATL         SF        BAL        CIN        WSH        OAK        DET
0.04539092 0.05218956 0.06118776 0.08078384 0.08558288 0.08958208 0.09478104 0.10757848
       TEX        NYY
0.14997001 0.23295341
[1] "Year: 2013"

        TB        CLE        CIN        STL        LAD        OAK        BOS        ATL
0.04539092 0.05158968 0.06638672 0.09258148 0.10717856 0.11257748 0.11397720 0.12117576
       PIT        DET
0.13957209 0.14957009
[1] "Year: 2014"

       STL        WSH         SF        PIT        LAD        BAL         KC        LAA
0.03659268 0.04299140 0.05298940 0.06878624 0.08738252 0.10277944 0.12357528 0.12897421
       OAK        DET
0.17396521 0.18196361
[1] "Year: 2015"

       NYM        LAD        CHC        STL        PIT        NYY         KC        TEX
0.02099580 0.02239552 0.04919016 0.05558888 0.06078784 0.08698260 0.09738052 0.15596881
       HOU        TOR
0.17396521 0.27674465
[1] "Year: 2016"

       CLE        NYM        WSH        TEX         SF        CHC        BAL        LAD
0.04039192 0.05658868 0.06918616 0.06938612 0.07178564 0.08778244 0.12377524 0.12597481
       TOR        BOS
0.16836633 0.18676265
```

Appendix: MCMC Results III

```
[1] "Year: 2012"

        NE        HOU        BAL        CIN        ATL        WAS        MIN         GB
0.07638472 0.07898420 0.08318336 0.08318336 0.08378324 0.08398320 0.08418316 0.08458308
        SF        SEA        IND        DEN
0.08458308 0.08478304 0.08518296 0.08718256
[1] "Year: 2013"

       CIN         SF        IND        DEN         SD         NE         GB         KC
0.07458508 0.07878424 0.08078384 0.08178364 0.08198360 0.08238352 0.08358328 0.08418316
       SEA        CAR         NO        PHI
0.08458308 0.08638272 0.08778244 0.09318136
[1] "Year: 2014"

       DAL        SEA         GB        IND         NE        CAR        CIN        ARI
0.07958408 0.07958408 0.07978404 0.08138372 0.08138372 0.08298340 0.08318336 0.08398320
       DET        PIT        DEN        BAL
0.08498300 0.08658268 0.08798240 0.08858228
[1] "Year: 2015"

       DEN         GB        PIT        HOU        SEA        CIN        MIN        WAS
0.07718456 0.07878424 0.07918416 0.07978404 0.07998400 0.08198360 0.08298340 0.08438312
       CAR        ARI         NE         KC
0.08738252 0.08918216 0.08918216 0.08998200
[1] "Year: 2016"

       HOU        SEA         GB        NYG        DAL        ATL         KC        MIA
0.07958408 0.08078384 0.08178364 0.08178364 0.08198360 0.08298340 0.08318336 0.08478304
        NE        OAK        PIT        DET
0.08498300 0.08498300 0.08618276 0.08698260
```