

# **Data Driven Analysis of NCAA Basketball**

**Karl Shiffler**

# Hypotheses

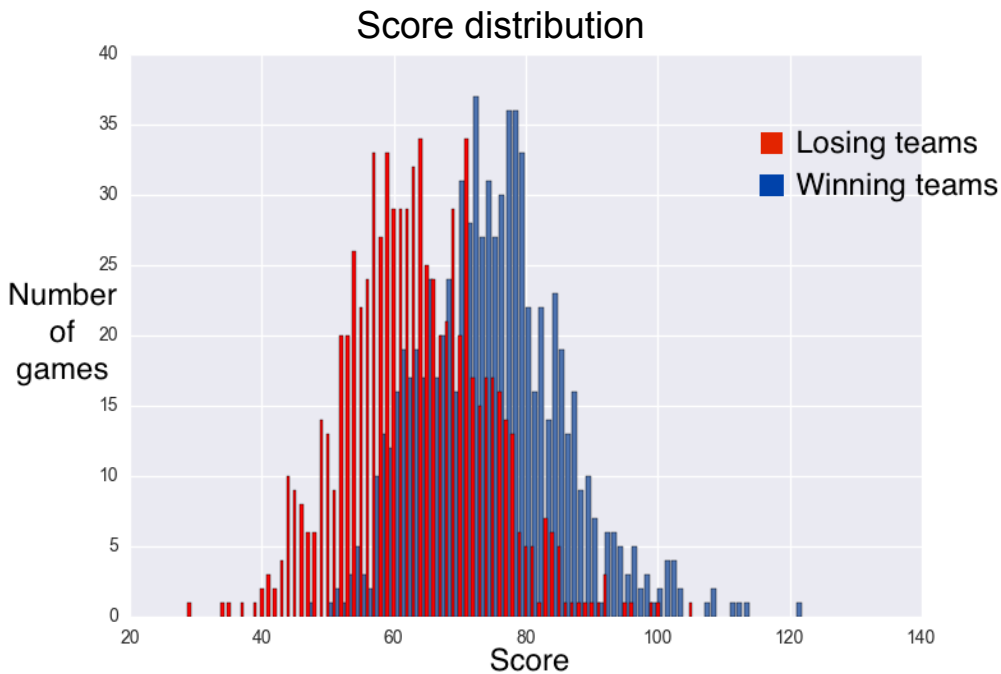
- Rule changes over the past 10 years have had significant effects on stats.
- Free throw percentages are integral to winning.
- Getting rebounds is integral to winning.
- Its better to play as a team than depend on a few stars.
- Efficiency is more important than volume.
- The tournament adds stress that leads to worse performance.

data: <https://www.kaggle.com/c/march-machine-learning-mania-2015/data>

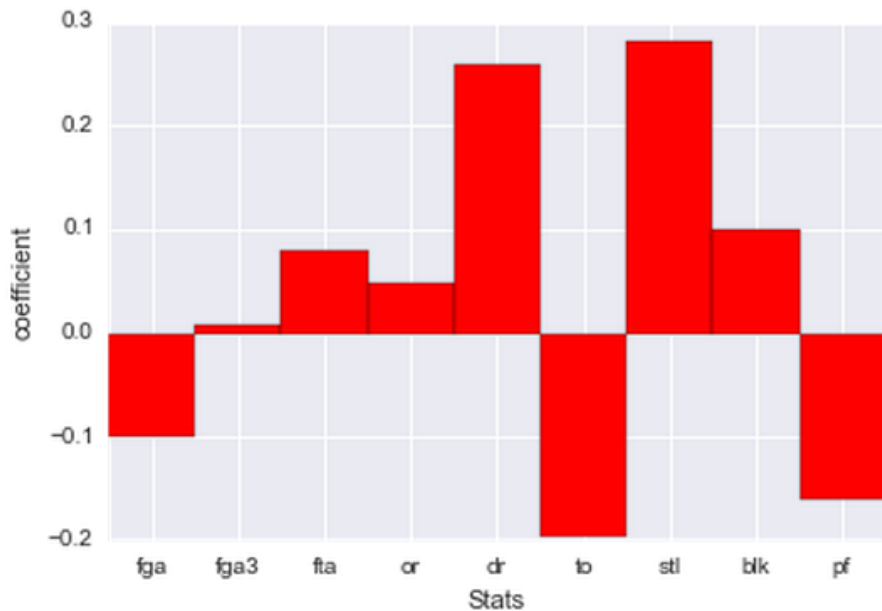
# Understanding the data

We have access to:

field goals made/attempted
three pointers made/attempted
free throws made/attempted
offensive/defensive rebounds
assists
turnovers
steals
blocks
personal fouls

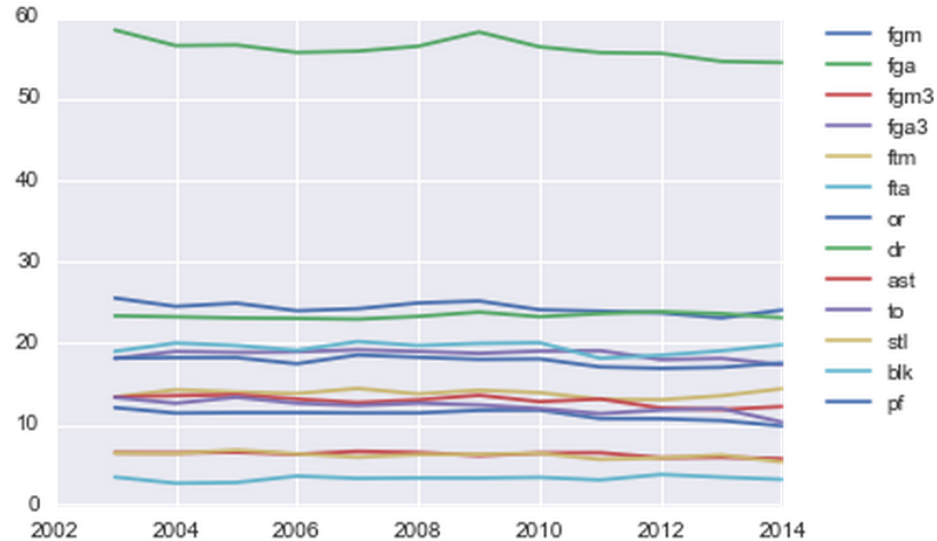
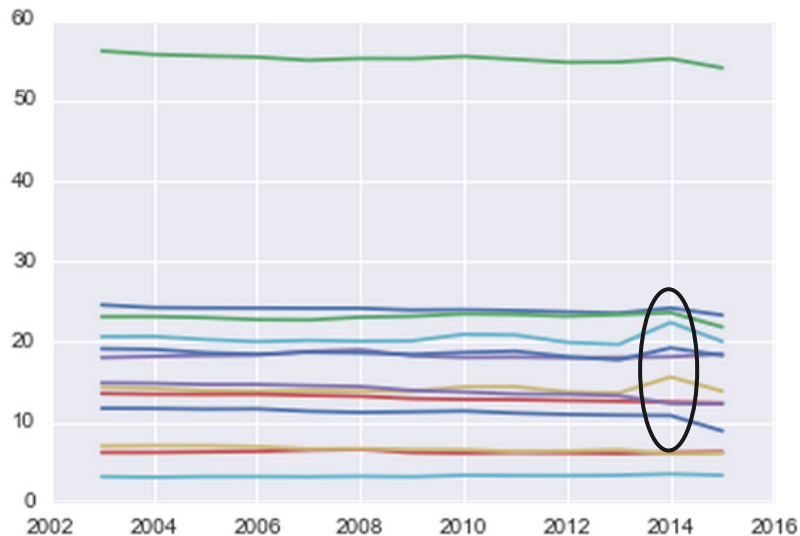


# The big question: how do we win?



fga	Field Goal Attempt
fga3	3 pt Field Goal attempt
fta	Free throw attempt
or	Offensive rebound
dr	defensive rebound
to	turnover
stl	steal
blk	block
pf	personal foul

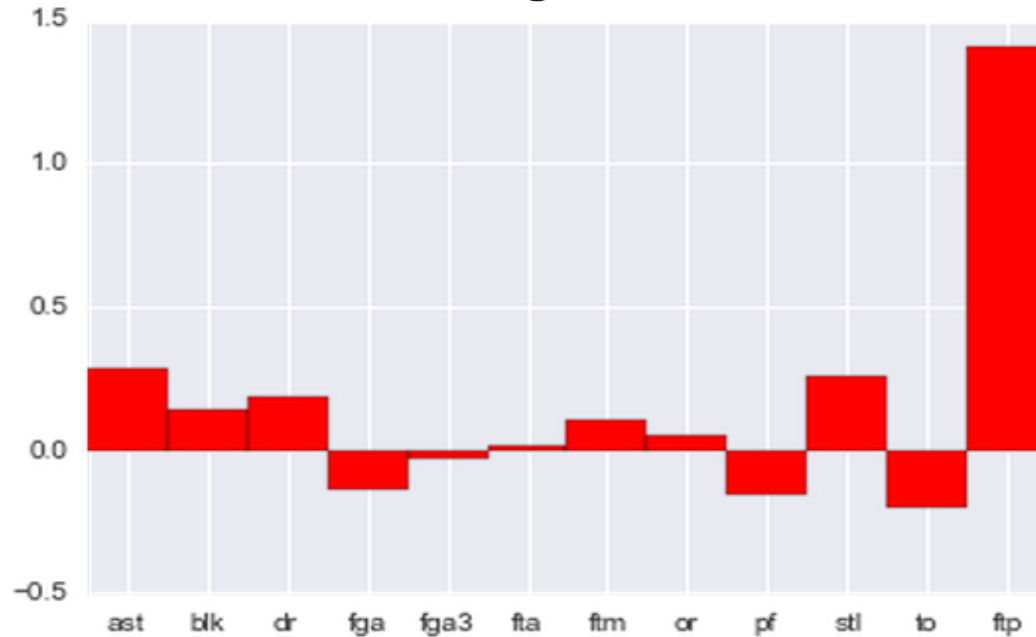
# How has the game changed over time?



Season Averages	FG Attempts	Offensive Rebounds	Turnovers
<b>2003</b>	58.406250	12.101562	13.351562
<b>2014</b>	54.417910	9.858209	10.313433

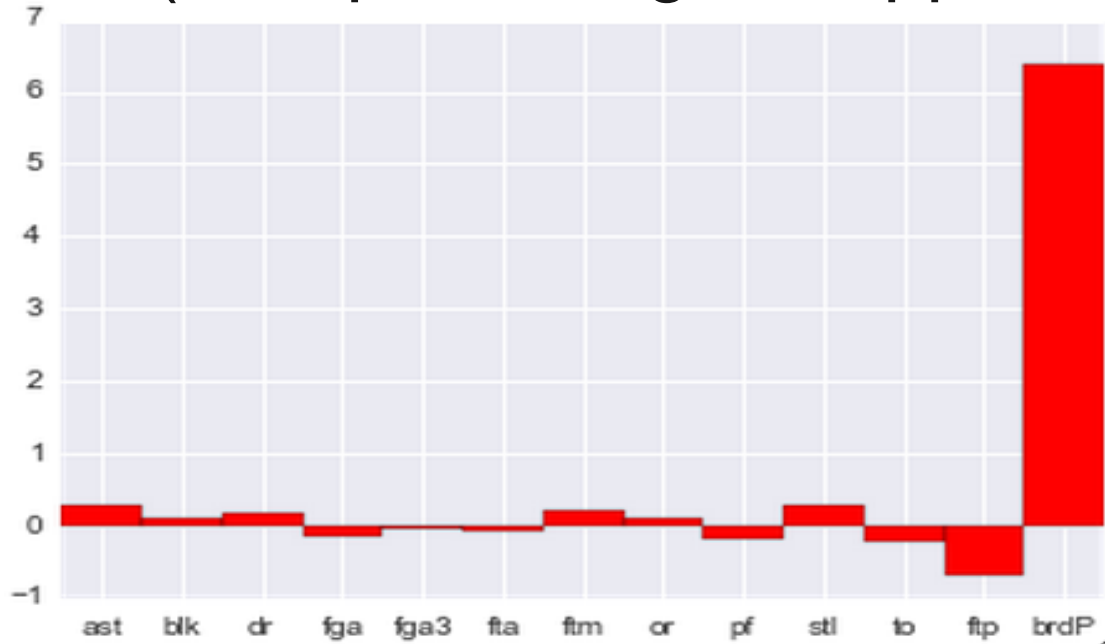
# Conventional Basketball Wisdom

## Free Throw Percentage



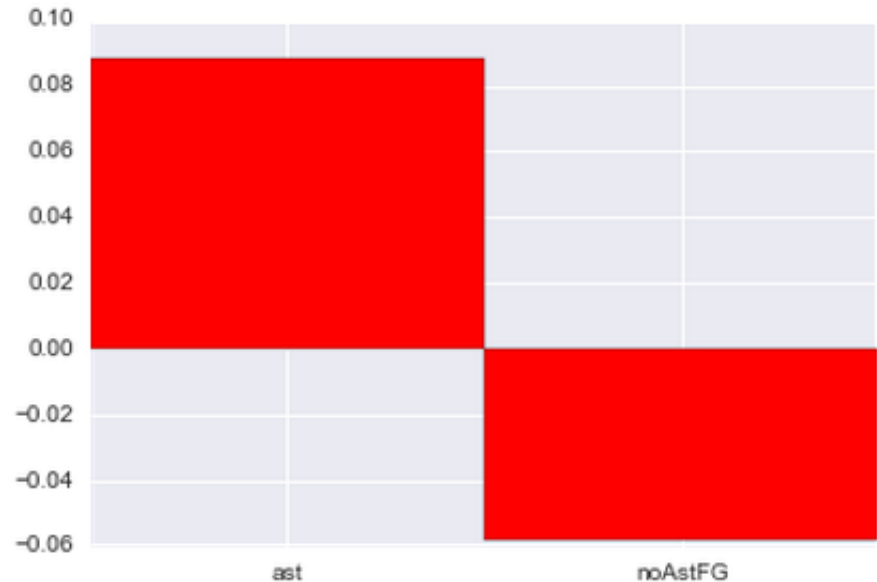
# Conventional Basketball Wisdom

Rebounds (as a percentage of opponents' shots)



# Team Players or Stars?

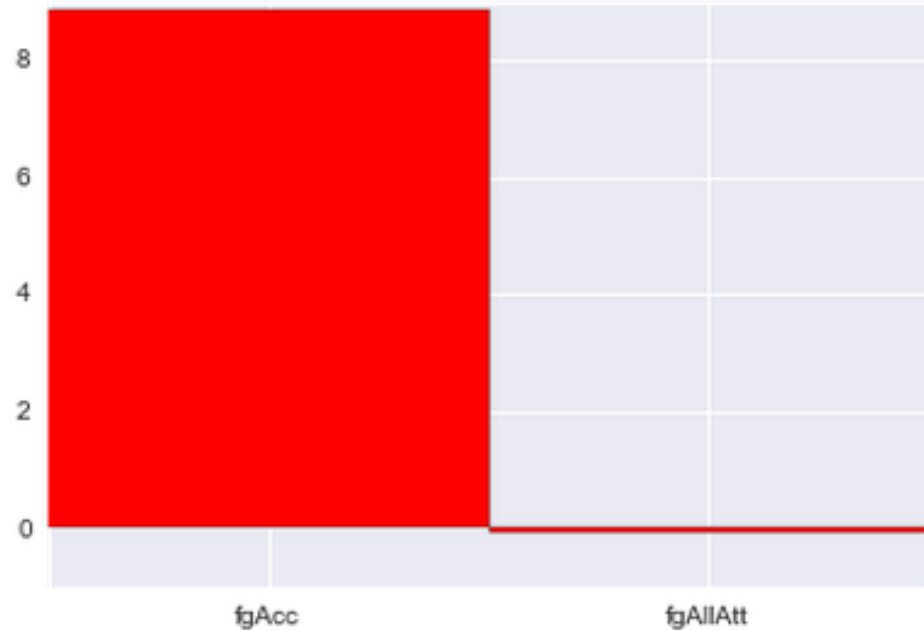
Conventional basketball wisdom dictates “one player can’t carry a team,” and it appears that in college basketball, this holds true.





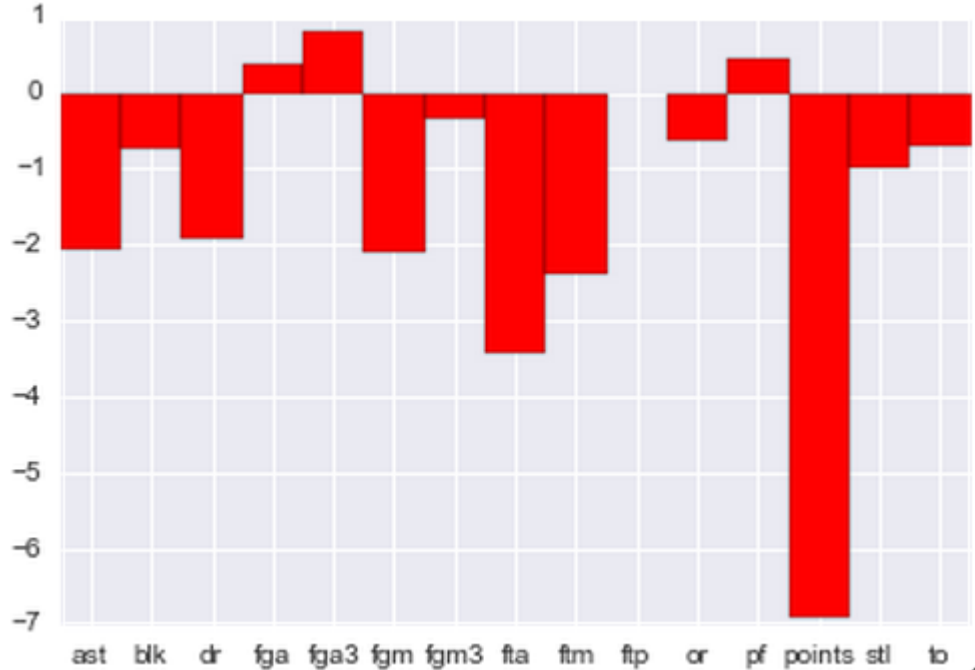
# Efficiency or high volume?

Yet another basketball trope: “Take care of the ball.”



# Regular Season vs Tournament

Controlled by team and season, subtract regular season averages from tournament averages.



# Conclusions

- Retaining possession and limiting the possessions of your opponents is the key to winning.
- Stats have remained largely stable over the last ten years.
- The conventional basketball wisdom regarding free throws and rebounds is supported by the data
- Team play and a focus on efficiency is also important, as supported by the data.

# Future work

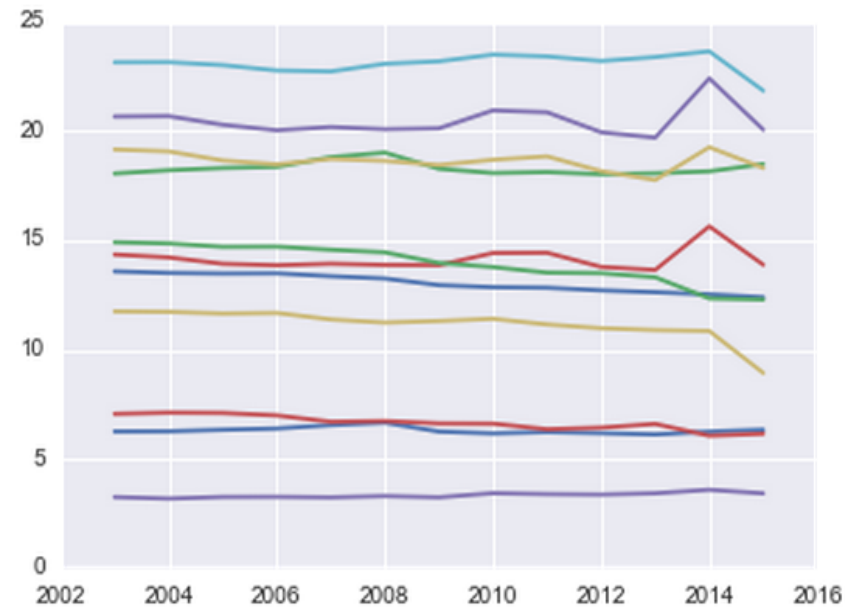
- Delve further into individual teams
- Add more data: individual player stats, Vegas spreads
- Create a predictor for the March Madness tournament

**Thank you!**

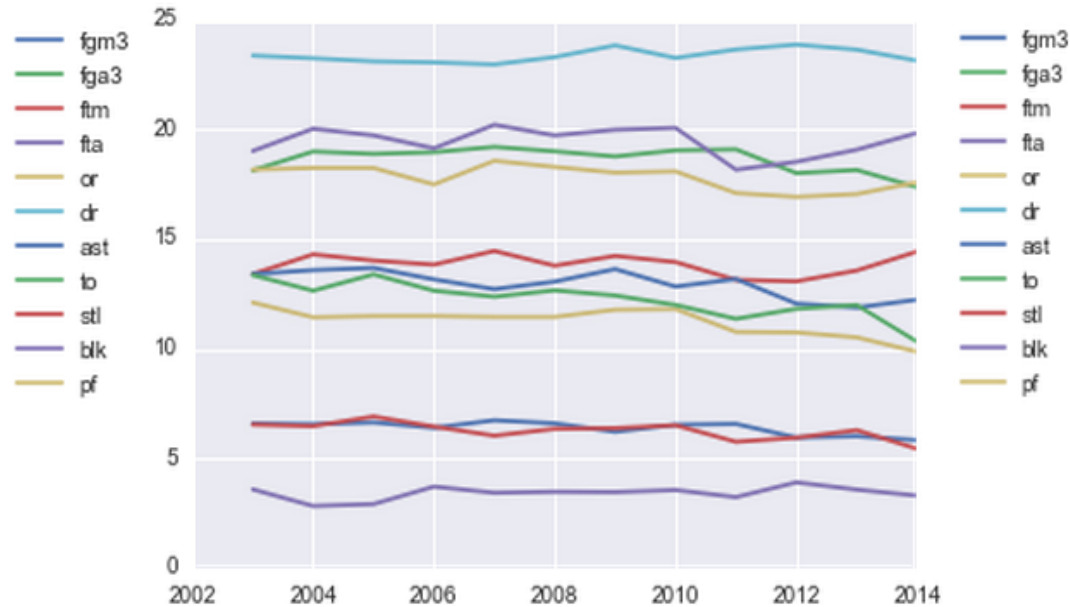


# Timeseries Zoomed

Regular Season



Tournament



# FT % and Board % Logscale

