Sam Edgemon, SAS Senior Consultant

Hoop Analytics 101

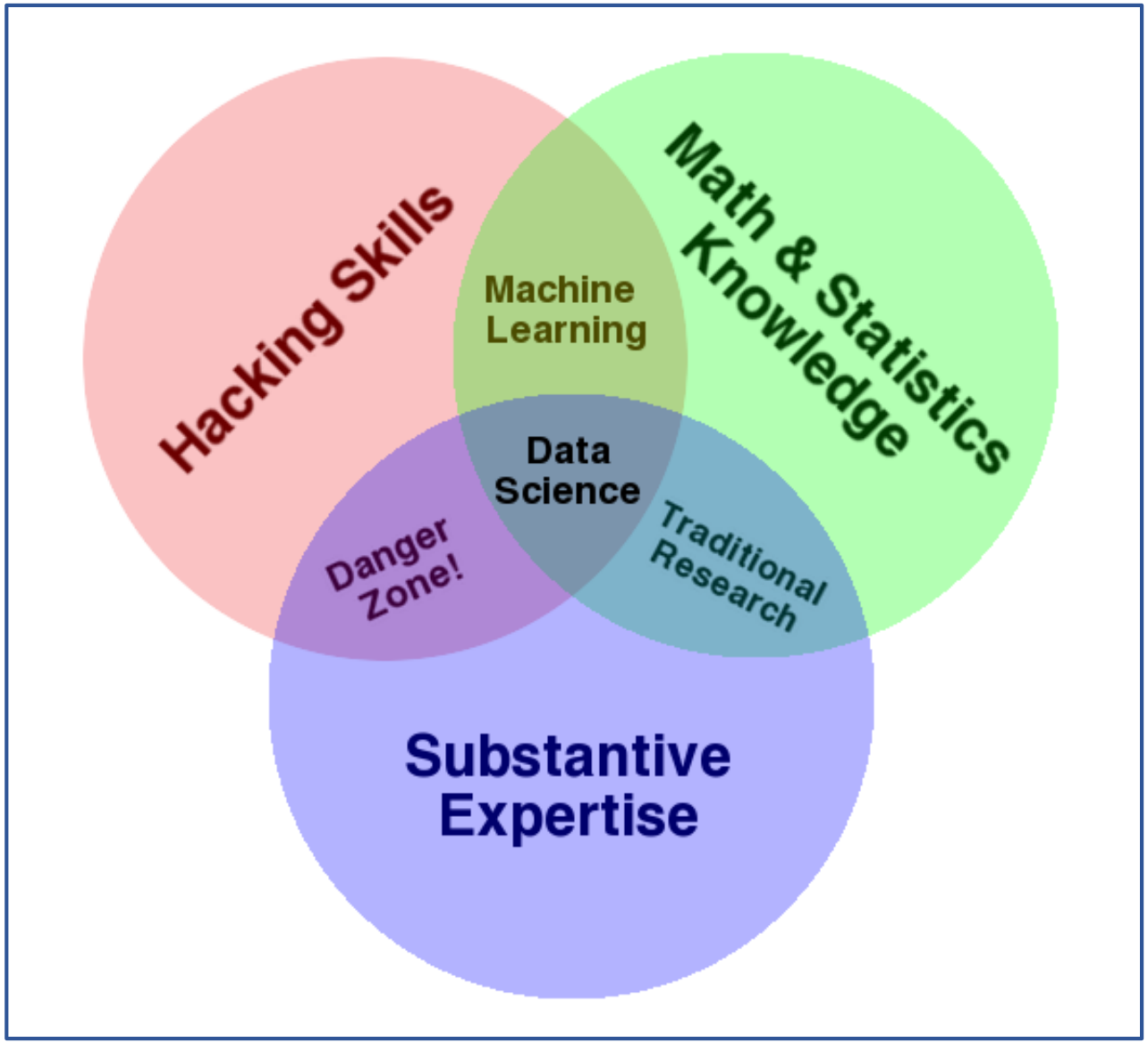
An Introduction to Basketball Analytics

(the opportunity is there!)

“We can’t get data into the thing.”

The Data Scientist is going to be the person that **understands the client needs**, and then how to position those needs as **a question that can be answered with data** found in the database.

The Data Scientist **must be good** at **statistics** and **data analytics**, **computer science** and **database** work



Drew Conway's Data Science Venn Diagram

http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram

Communication.

The Data Step

Procedures (PROCs)

SAS

- **Website:** www.sas.com
- **SAS University for Edition**
- https://www.sas.com/en_us/software/university-edition.html

JMP

- **Website:** www.jmp.com
- **JMP Trial and Student Edition**
- https://www.jmp.com/en_us/download-jmp-free-trial.html
- https://www.jmp.com/en_us/academic/jmp-student-edition.html

- https://www.basketball-reference.com/
- https://www.kaggle.com/andrewsundberg/college-basketball-dataset

The programs, slides and data we use in these sessions will be available at:

github.com/AtlantaSam/BasketballAnalytics

SAS Institute

Cary, NC

www.sas.com

How did I get in the sports analytics)?

Ans. Kobe Bryar

Game Plan

- Foundations of Analytics – Preparing Data for Analysis
- The History of Basketball Analytics
  - The Bill James (Analytics) Revolution
  - Considering interpretations of metrics
- Modeling Basketball Data
- Evaluate and find players using what was learned modeling the data
- Bracketology and the Midway Mules
- Discussion

Preparing Data for Analysis

- https://www.basketball-reference.com/
- Create the dataset(s)
- TeamGo, TeamStop, MiscStats, TeamPossStats
- Exploratory Data Analysis (EDA)
  - Data cleansing (Timberwolve vs. Timberwolves)
  - Look for outliers, missing values, etc.

Sam Edgemon
sam.edgemon@sas.com
www.linkedin.com/in/samedgemon

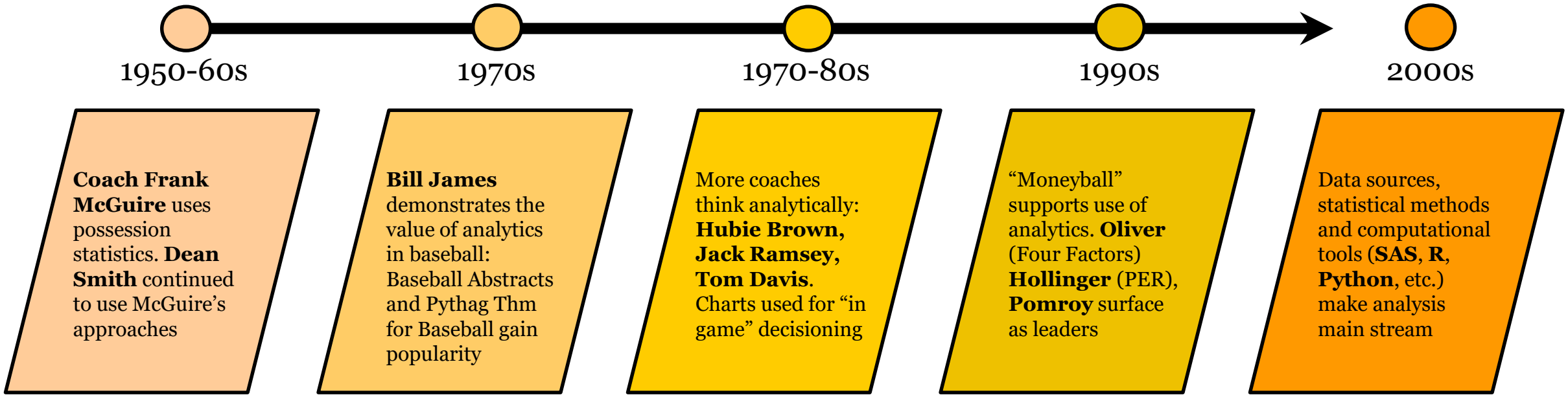
www.github.com/AtlantaSam/BasketballAnalytics

(Friday, 12:00)

History of Basketball Analytics

- Who and what?
- Bill James
- Pythagorean Theorem for Basketball
- Explore Basketball Metrics like Field Goal Percentage, True Shooting Percentage, and the Effective Field Goal Percentage

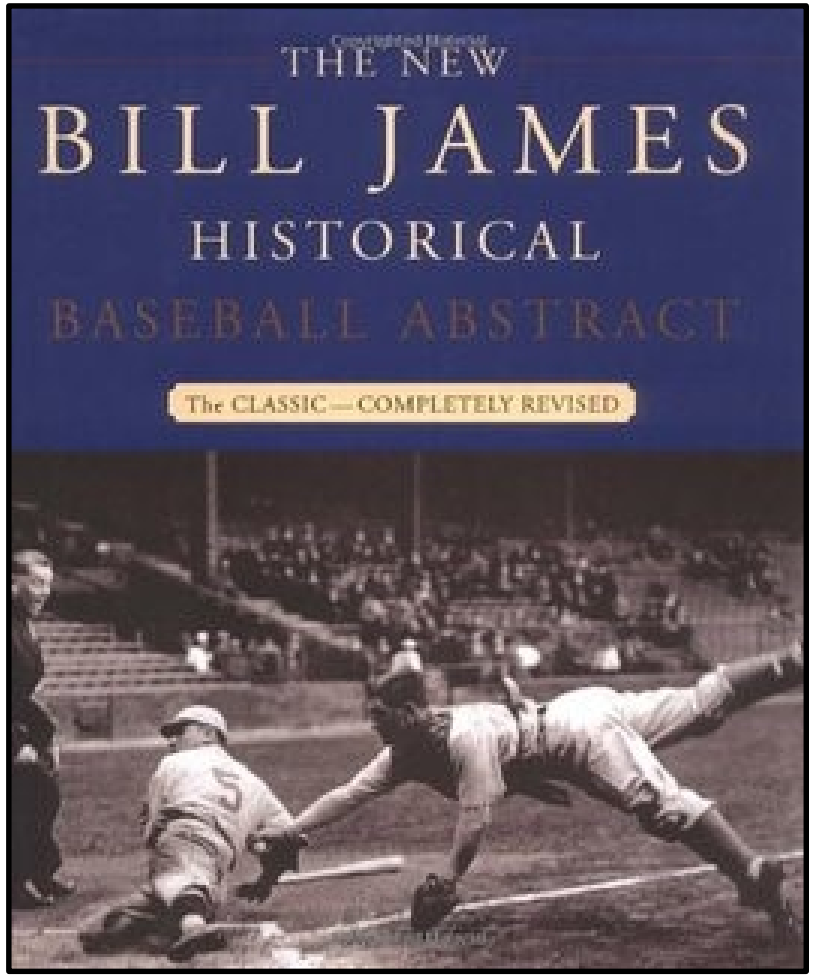
The History of Basketball Analytics



THE BILL JAMES (ANALYTICS) REVOLUTION

Bill James

- Runs Created
- Range Factor
- Baseball's Pythagorean Theorem



First published in 1977

“What James demands is that we take the time to listen to what the game is telling us over and above what we are predisposed to believe”

The 2006 TIME 100

Here's our list of the 100 men and women whose power, talent or moral example is transforming our world

Like 61 Tweet G+1 0 Share

SCIENTISTS & THINKERS

Bill James

92 of 102
VIEW ALL

By John Henry | Monday, May 08, 2006

The Red Sox raised a few eyebrows by hiring maverick statistician Bill James in 2002. That was before the team won its first World Series in 86 years. When I asked our general manager, Theo Epstein, what he thought about James' impact on the game, he said, "The thing that stands out for me is Bill's humility. He was an outsider, self-publishing invisible truths about baseball while the Establishment ignored him. Now 25 years later, his ideas have become part of the foundation of baseball strategy. But where's Bill? Where's the gloating? Where's the publicist? He's like somebody outlining the Internet in the '80s and watching silently as it comes to pass."

Keith Philpott for TIME

What we now know as Moneyball and sabermetrics came from James, 56. He taught us, among other things, that individual ballparks have a profound effect on a ballplayer's production, that the largest variable determining how many runs a team will score is how many times the leadoff hitter gets on base, that much of what we perceive as pitching is actually defense. What I call Jamesian principles infuse our thinking with a perspective that is objective rather than subjective. What James demands is that we take the time to listen to what the game is telling us over and above what we are predisposed to believe.

Sabermetrics

Society for American
Baseball Research

"The mathematical and
statistical analysis of baseball"

APBRmetrics

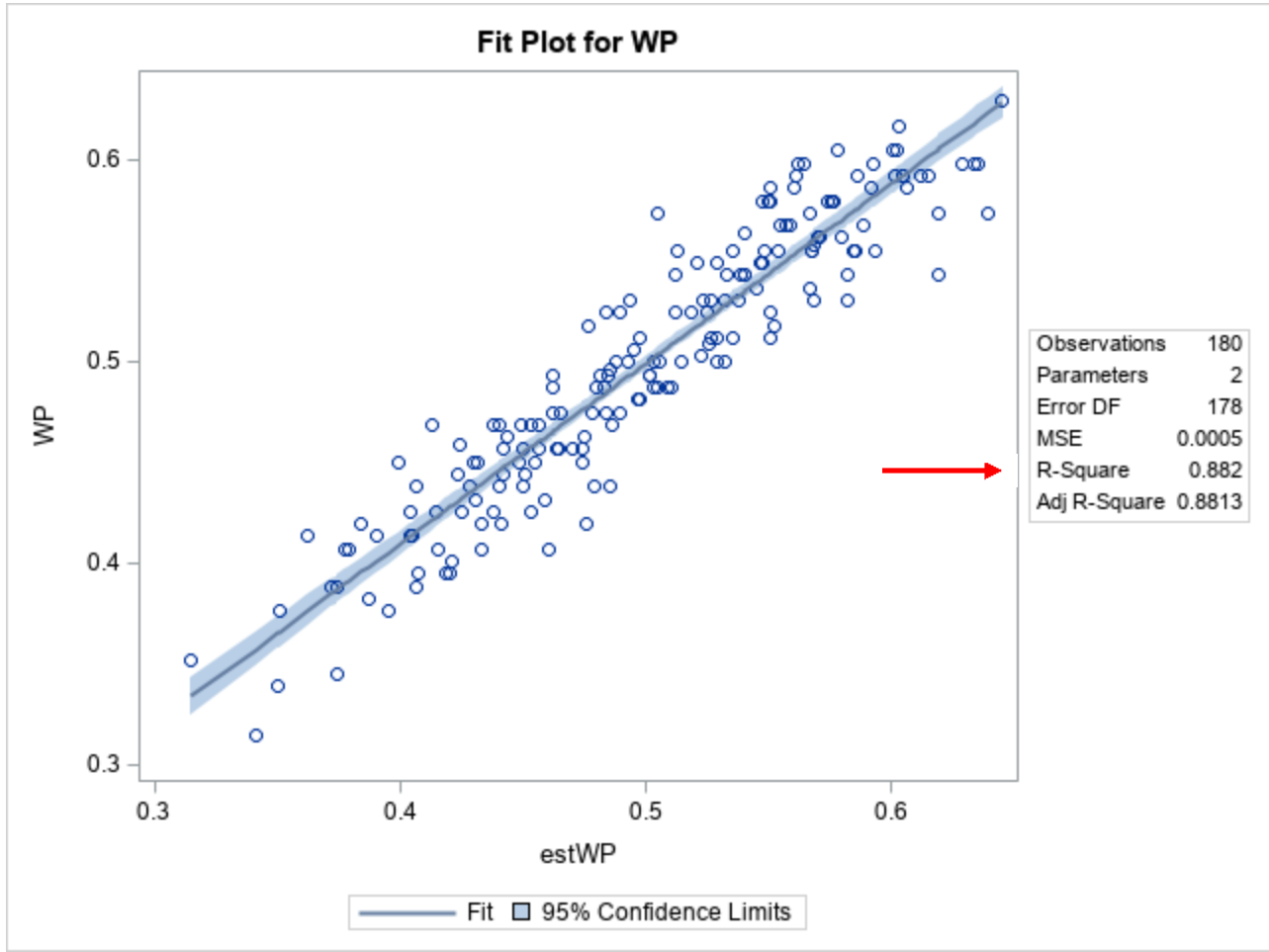
Association for Professional
Basketball Research

"Analyzing basketball statistics
through objective evidence"

The Pythagorean Theorem of Baseball

$$Winning\ Percentage = \frac{W}{(W+L)} \cong \frac{R^2}{(R^2+RA^2)}$$

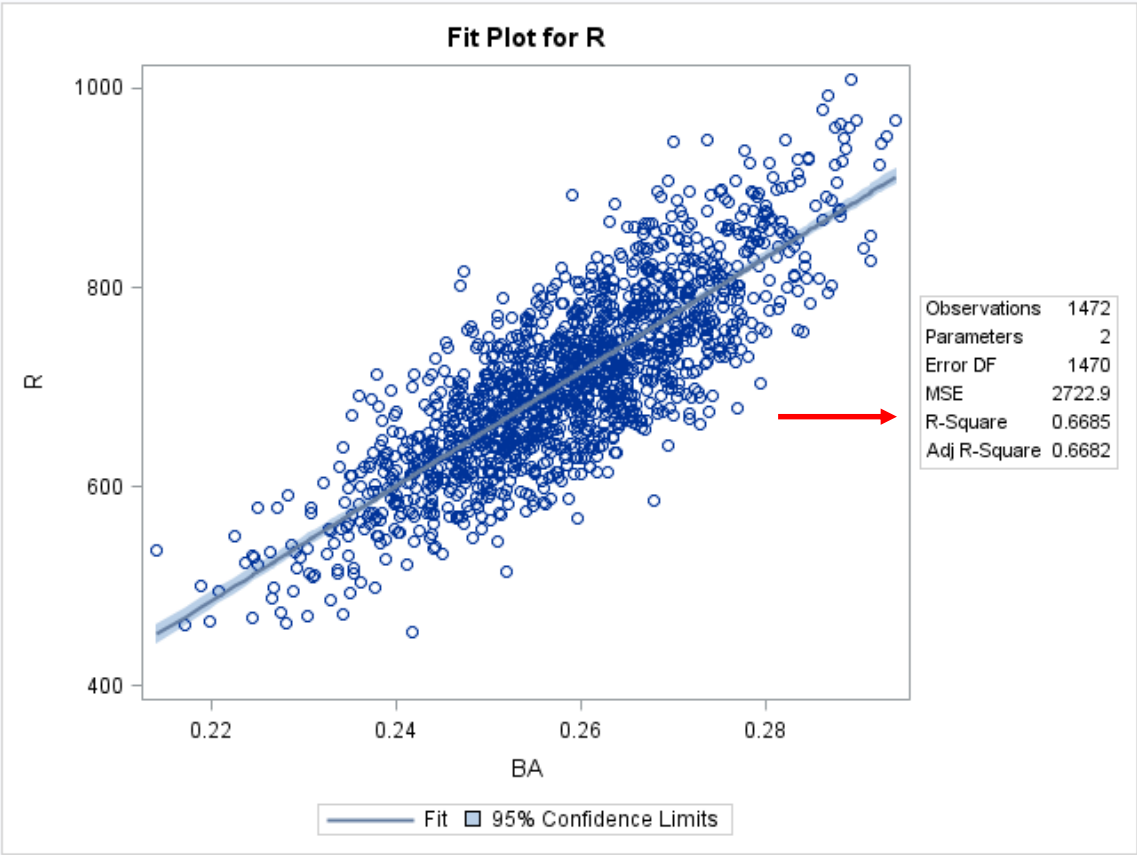
Actual Winning Percentage (WP) vs the Pythagorean Estimate (estWP) for Winning Percentage



“People who run ball clubs, they think in terms of buying players. Your goal shouldn’t be to buy players; your goal should be to buy wins. And in order to buy wins, you need to buy runs …”

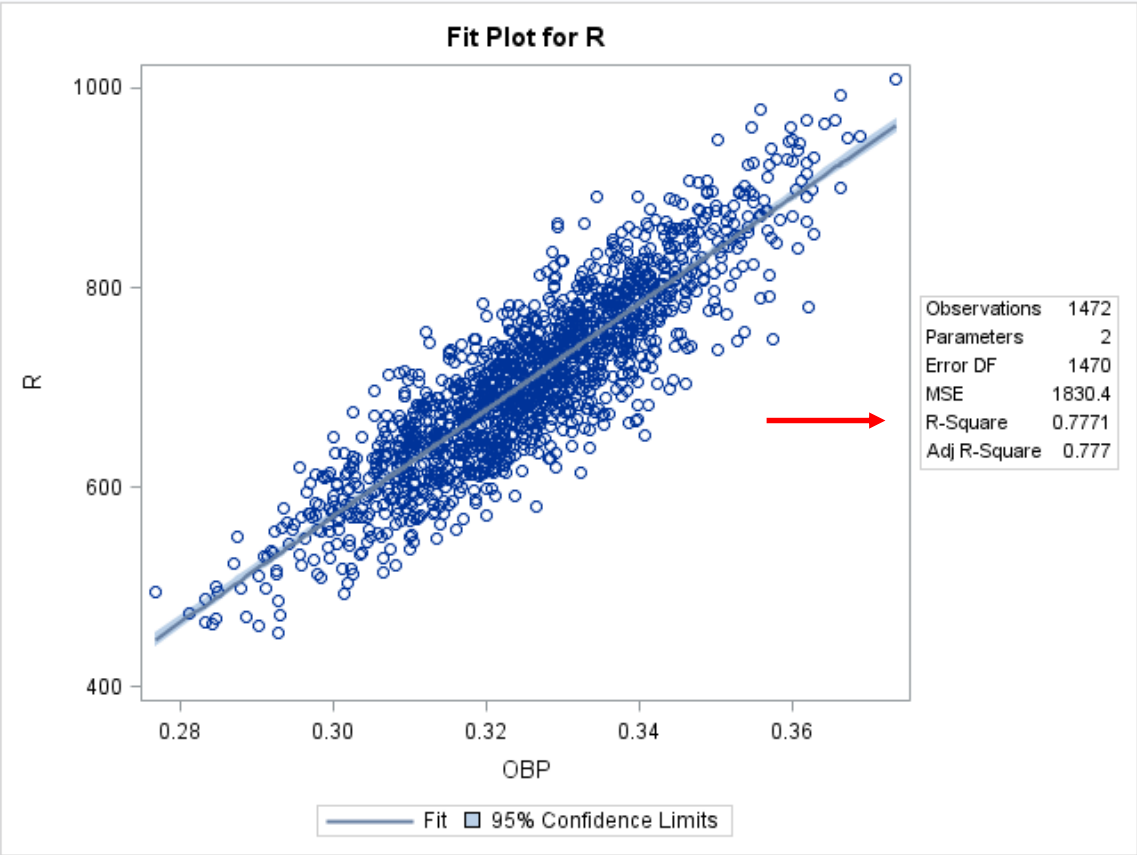
--- Peter Brand

Batting Average (BA)



$$BA = \frac{H}{AB}$$

On Base Percentage (OBP)



$$OBP = \frac{(H + BB + HBP)}{(AB + BB + HBP + SF)}$$

Will these ideas apply to the game of basketball?

Modeling Basketball Data

- Testing ideas – consider that if want to build a team … we want to know how to evaluate players
- Dean Oliver's Four Factors
- Predicting Winning Percentages using Statistical Models

Evaluating Basketball Talent

- Considering what we've learned from the modeling effort, can we find players that meet our goals?

Bracketology

- The Midway Mules were great in the “long haul” of tournaments
- It’s not about picking which team will win “a single game,” but the challenge is to find which teams are best equipped to win 6 tournament games

How do basketball teams win games?

https://www.basketball-reference.com/about/factors.html

Thank You!

Sam Edgemon
sam.edgemon@sas.com
www.linkedin.com/in/samedgemon