Overview of Sports Analytics

INEFC Institut Nacional d'Educació Física de Catalunya Barcelona
Generalitat de Catalunya
2024
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
Martí Casals STATSTHINKING

1

Outline

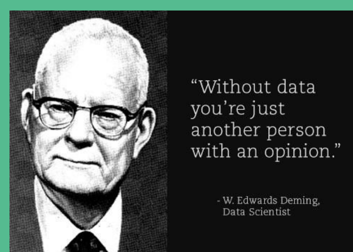
1. History of Sports Analytics
2. Sports Statistics or Sports Analytics?
3. Resources of Sports Analytics
4. Specializations in sports and data
5. Technology and Visualization
6. Types of data, variables, and c-speak
7. Practical Tips for Real-World Applications

2 @CasalsTMarti

2

What is the objective of today's session?

To Develop Critical Thinking in Sports !!!



3 @CasalsTMarti

3

0. My academic and profesional point of view

"We believe that the **leading thinkers of the next decade will be those who seamlessly knit together tools from both statistics and computing** and that how we think about statistics will be informed by complementary computational thinking**"**

Nicholas Horton & Johanna S. Hardin

4 @CasalsTMarti

4

My current work and experience

- Lecturer of Statistics at INEFC (University of Barcelona). (2021 – Present)
- Lecturer of Biostatistics and Epidemiology at Faculty of Medicine, University of Vic – Central University of Catalonia (UVic-UCC). (2017 – Present)
- Associate Editor of British Journal of Sports Medicine (BJSM) and Injury Prevention and Rehabilitation (specialty section of Frontiers in Sports and Active Living) (2020 – Present)
- Researcher in the Sport Performance Analysis Research Group (SPARG) of the Sport and Physical Activity Studies Centre (CEEAF), University of Vic (UVic-UCC). (2015- Present)
- External biostatistician and basketball analyst for the Memphis Grizzlies (NBA). (2016 - 2018)
- Team member of Barça Innovation Hub – Universitas. Statistician Researcher, Sport Science Department. F.C. Barcelona (2018 – 2021)
- More info: LinkedIn

5

Why I decided to study statistics...

I had also planned to study a sports-related degree

6

Some of my contributions: Sports Analytics

MUNDO DEPORTIVO

BALONCESTO · NBA

Un estudio español revela las claves del rendimiento

Damian Lillard trata de superar a Jose Calderon

Journal

International Journal of Performance Analysis in Sport ›

Volume 13, 2013 - Issue 1

61 Views

1 CrossRef citations to date

0 Altmetric

Articles

Modelling player performance in basketball through mixed models

Martí Casals & A. Jose Martinez

Pages 64-82 | Published online: 03 Apr 2017

Download citation https://doi.org/10.1080/24748668.2013.11868632

Annals of Operations Research
https://doi.org/10.1007/s10479-022-04733-0

ORIGINAL RESEARCH

Influence of Red and Yellow cards on team performance in elite soccer

Llorenç Badiella[1,2] · Pedro Puig[2,3] · Carlos Lago-Peñas[4] · Martí Casals[5,6,7]

Accepted: 12 April 2022

7

Some of my contributions: Sports Medicine

Appl. Sci. **2019**, *9*(3), 500; https://doi.org/10.3390/app9030500 Open Access Article

Mortality of NBA Players: Risk Factors and Comparison with the General US Population

Jose A. Martinez [1], Klaus Langohr [2], Julián Felipo [3] and Martí Casals [4,5,*]

1. Department of Business Economics, Universidad Politécnica de Cartagena, 30201 Cartagena, Spain
2. Department of Statistics and Operations Research, Universitat Politècnica de Catalunya/Barcelonatech, 08034 Barcelona, Spain
3. Newsroom, Basketball Departament, Mundo Deportivo, 08036 Barcelona, Spain
4. Sport and Physical Activity Studies Centre (CEEAF), University of Vic—Central University of Catalonia (UVic-UCC), 08500 Catalonia, Spain
5. Medical Department, Futbol Club Barcelona, Barça Innovation Hub, 08028 Barcelona, Spain

* Author to whom correspondence should be addressed.

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

ELSEVIER

Data Article

Data set on mortality of national basketball association (NBA) players

Jose A. Martínez [a], Klaus Langohr [b], Julián Felipo [c], Luciano Consuegra [d], Martí Casals [e,f,*]

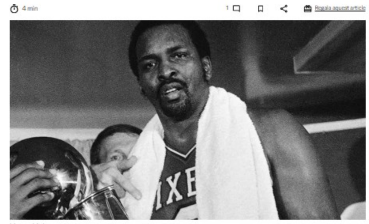
ara

Bàsquet - NBA

Els exjugadors de la NBA afroamericans i alts tenen més probabilitat de morir abans que la resta

Un estadi associa l'alçada i l'ètnia amb la mortalitat

Àlex Gozalbo



Moses Malone

8

Some of my contributions: Stats & Injury Epidemiology

SORT 39 (2) July-December 2015, 281-308

Parameter estimation of Poisson generalized linear mixed models based on three different statistical principles: a simulation study

Martí Casals[1,2,3,4], Klaus Langohr[5], Josep Lluís Carrasco[1] and Lars Rönnegård[6]

Abstract

Generalized linear mixed models are flexible tools for modeling non-normal data and are useful for accommodating overdispersion in Poisson regression models with random effects. Their main difficulty resides in the parameter estimation because there is no analytic solution for the maximization of the marginal likelihood. Many methods have been proposed for this purpose and many of them are implemented in software packages. The purpose of this study is to compare the performance of three different statistical principles – marginal likelihood, extended likelihood, Bayesian analysis – via simulation studies. Real data on contact wrestling are used for illustration.

MSC: 62J12, 62P99, 62F99.

Keywords: Estimation methods, overdispersion, Poisson generalized linear mixed models, simulation study, statistical principles, sport injuries.

9 @CasalsTMarti

9

Some of my contributions: Sports Science and Medicine

©Journal of Sports Science and Medicine (2018) **17**, 289-297
http://www.jssm.org

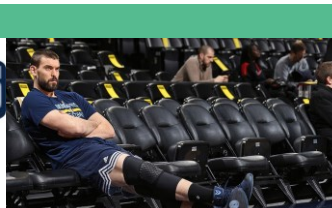
Research article

Low External Workloads Are Related to Higher Injury Risk in Professional Male Basketball Games

Toni Caparrós [1,2]✉, Martí Casals [2], Álvaro Solana and Javier Peña [2,3]
[1] Institut Nacional d'Educació Física de Catalunya (INEFC), Barcelona, Spain; [2] Sport Performance Analysis Research Group (SPARG), Universitat de Vic, Vic, Spain; [3] UVic-UCC Sport and Physical Activity Studies Centre (CEEAF), Universitat de Vic, Vic, Spain

Toni Caparrós. Sports Scientist Advisor. S&C coach.

MEMPHIS GRIZZLIES



10 @CasalsTMarti

10

Some of my contributions: Sports Analytics and Education

Original research article

International Journal of Sports Science & Coaching

Violinboxplot and enhanced radar plot as components of effective graphical dashboards: An educational example of sports analytics

Martí Casals[1,2,3] and Pepus Daunis-i-Estadella[4]

International Journal of Sports Science & Coaching
1–12
© The Author(s) 2022
Article reuse guidelines: sagepub.com/journals-permissions
DOI: 10.1177/17479541221099638
journals.sagepub.com/home/spo
SAGE

Check for updates

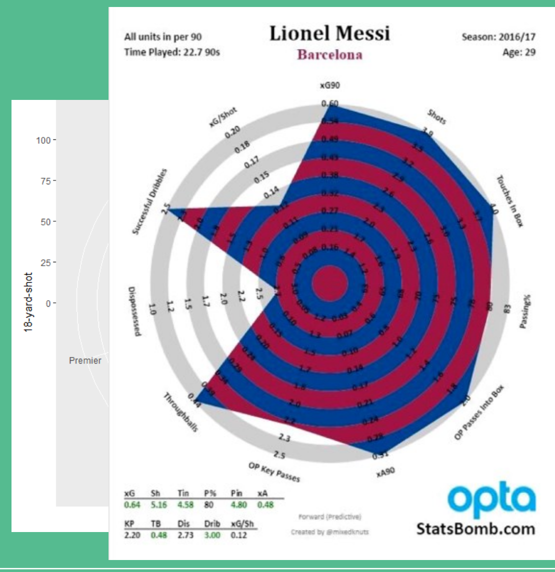
Review

International Journal of Sports Science & Coaching

A systematic review of sport-related packages within the R CRAN repository

Martí Casals[1,2,3], José Fernández[4], Victor Martínez[5], Michael Lopez[6], Klaus Langohr[7], and Jordi Cortés[7]

International Journal of Sports Science & Coaching
1–9
© The Author(s) 2022
Article reuse guidelines: sagepub.com/journals-permissions
DOI: 10.1177/17479541221136238
journals.sagepub.com/home/spo
SAGE



11

Some of my contributions of Tennis Research

International Journal of Sports Physiology and Performance, (Ahead of Print)
https://doi.org/10.1123/ijspp.2023-0385
© 2024 Human Kinetics, Inc.
First Published Online: Apr. 29, 2024

Human Kinetics
BRIEF REVIEW

Quantifying Hitting Load in Racket Sports: A Scoping Review of Key Technologies

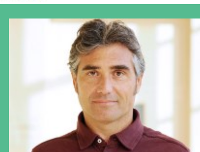
Quim Brich,[1] Martí Casals,[1,2] Miguel Crespo,[3] Machar Reid,[4] and Ernest Baiget[1]
[1]National Institute of Physical Education of Catalonia (INEFC), University of Barcelona (UB), Barcelona, Spain; [2]Faculty of Medicine, Sport and Physical Activity Studies Center (CEEAF), University of Vic-Central University of Catalonia (UVic-UCC), Barcelona, Spain; [3]Development Department, International Tennis Federation, London, United Kingdom; [4]Tennis Australia, Melbourne, VIC, Australia

PLOS ONE

RESEARCH ARTICLE

Retirements of professional tennis players in second- and third-tier tournaments on the ATP and WTA tours

Maria Palau[1], Ernest Baiget[2], Jordi Cortés[3], Joan Martinez[4], Miguel Crespo[5], Martí Casals[2,6,7]



Ernest Baiget

Jordi Cortés

Retirements, defaults and walkovers of professional tennis players in ATP and WTA events. (Under review)

12

Some of my contributions of Tennis Research

Rublev calls on ATP to review rule that led to default in Dubai

By Reuters

March 5, 2024 5:32 AM GMT+1 · Updated 2 months ago

Auger-Aliassime reaches first Masters final in Madrid with another walkover

MADRID — Montreal's Felix Auger-Aliassime has advanced to his first ATP Masters final, and he hasn't had to play all that much tennis to do it.

Canadian Press

May 3, 2024 12:27 PM

Updated May 3, 2024 12:35 PM

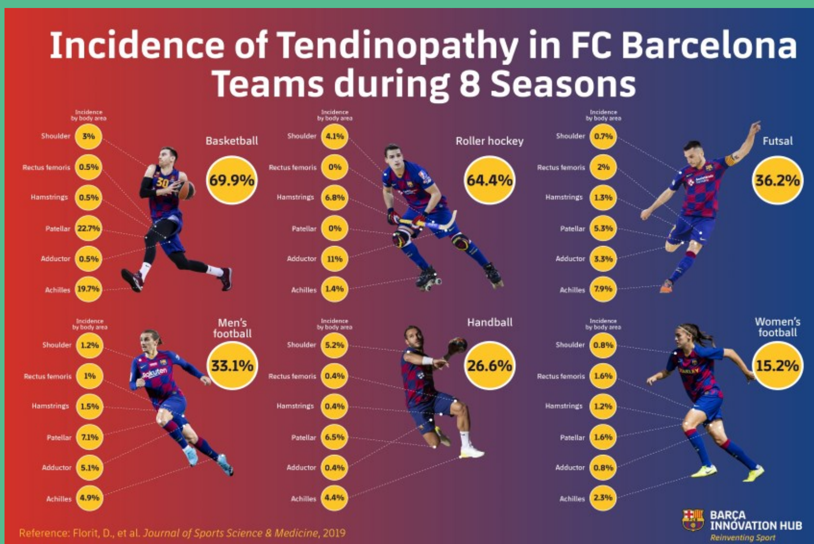
Sports Analytics and Biostatistics applied in FCB

- **Statistics & Education**
- Statistical training and awareness among differe professionals to help them ask better questions lead us to improved designs and decisions.

- **Performance and Sports Analytics**
- Group meetings (internal projects)

- **Research on Injury prevention & epidemiology**

Some of my contributions: Sports Science and Medicine



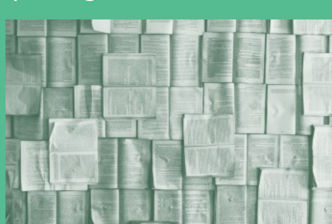
15 @CasalsTMarti

15

My point of view of Sports Analytics

Evolution of Sports Analytics
Basic statistics
Advanced statistics
Data literacy
Maths
Sports geeks

Ecological Physics and Motor Control
Sports Science
Computer science & Technology
Programmers
Visualization
Communication



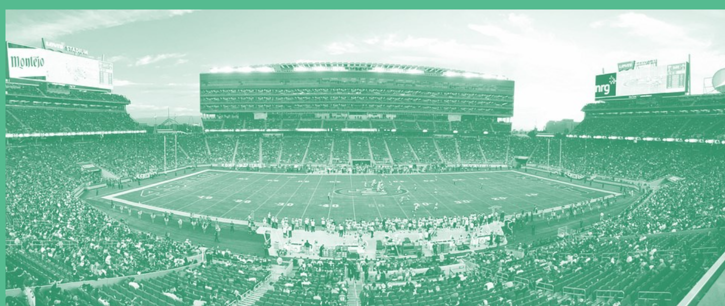
AND EDUCATION ??

16 @CasalsTMarti

16

1. History of Sports Analytics

More people are exercising and watching athletes on TV and in stadiums



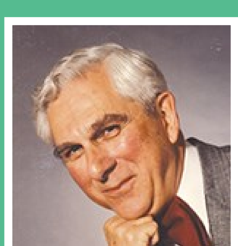
17

@CasalsTMarti

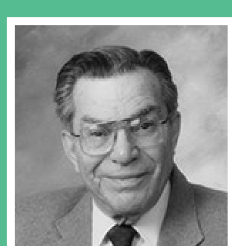
17

Sports Analytics /

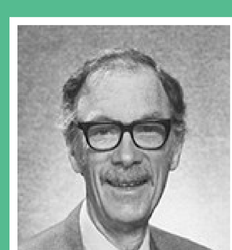
Statisticians in History



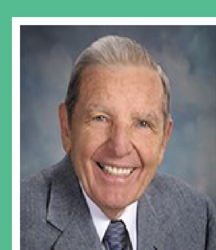
Richard L. Anderson (1915–2003)



Samuel W. Greenhouse (1918–2000)



Monroe Sirken



Joe Ward

Arnold Zellner

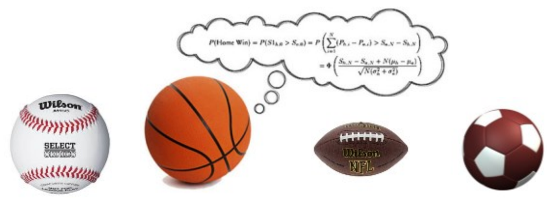
18

@CasalsTMarti

18

Sports Analytics / 2

History of Sports Analytics. A new trending topic?

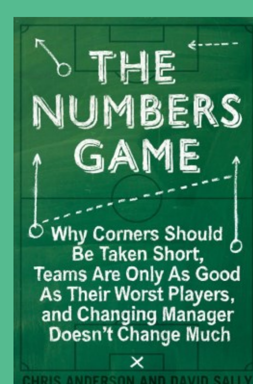


19 @CasalsTMarti

19

Do you know Charles Reep? (1930)

Charles Reep, a pioneer of football and performance analyst



Analytics in Soccer (2016).
https://sites.duke.edu/wcwp/2016/04/27/analytics-in-soccer/

The first xG calculations are surprisingly older

20 @CasalsTMarti

20

Richard Pollard. History Sports Analytics

Richard Pollard

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Measuring the effectiveness of playing strategies at soccer

Richard Pollard, Charles Reep,

First published: 05 January 2002 | https://doi.org/10.1111/1467-9884.00108 | Citations: 34

PDF TOOLS SHARE

Abstract

Using a notational system which records on-the-ball events taking place throughout a soccer match, the game can be broken down into a series of team possessions. To assess the effectiveness of a team possession, a quantitative variable is developed representing the probability of a goal being scored, minus the probability of one being conceded. This variable, called the yield, can be used to evaluate both the expected outcome of a team possession originating in a given situation, as well as the actual outcome of the possession. In this way, the effectiveness of different strategies occurring during the possession can be quantified and compared.

Sports Statistics or Sports Analytics? Sports Analytics

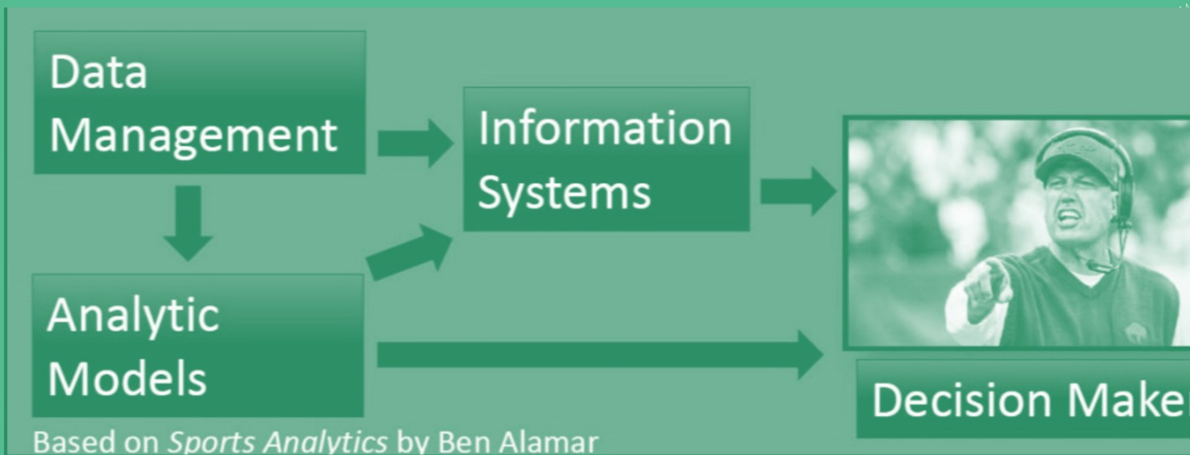
... brings together professionals of the same staff

(physicians, strength and Conditioning (S&C), physiotherapists, analysts, coaches)

... is not individual

... is a multidisciplinary team in sports clubs

where Sport Science, Behavioral Science, Sports Medicine and Data Science & Visualization can all benefit to better using and understanding team sports data.



What is Sports Analytics?

Data Mangement, Analytical Model and Information System for Decision maker is Sports Analytics

23

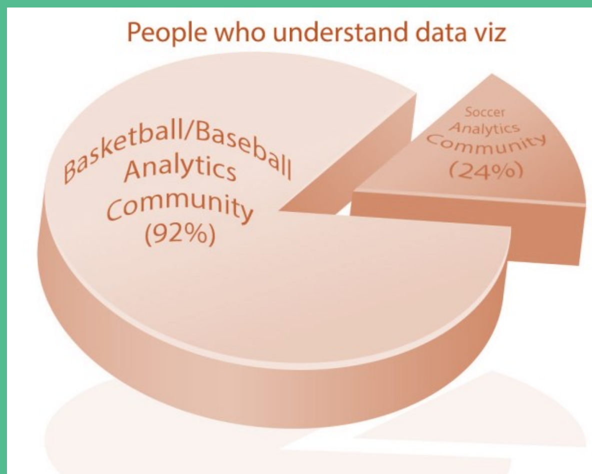
10 Steps to Get Started in Sports Analytics

1. Love Sports and Be Curious
2. Read, listen, and follow the experts
3. Learn to Code
4. Find Data Sources
5. Understand the Analytics Techniques that Work for Sports
6. Become a Storyteller
7. Analyze Everything in Your Path
8. Build your Personal Brand – Showcase Your Analyses
9. Network
10. Hustle

24 **Source:** Bill Kapatsoulias @CasalsTMarti

24

Different culture of Data Viz according to sport?



Source: Luke Bornn (Sport Analyst of Sacramento Knings and exanalyst (Roma; soccer)

25 @CasalsTMarti

25

Sports Analytics / 2

Community of sports statistics

The American Statistical Association (ASA) began a Section on Statistics in Sports in 1992 (https://community.amstat.org/sis/home)

A number of journals, such as Chance, The American Statistician and The Statistician, have regularly featured articles on the statistical analysis of sports data.

The Journal of Quantitative Analysis in Sports was launched in 2005 to be the first academic journal devoted to the statistical analysis of sports.

A number of research groups have been formed that are dedicated to sports statistics (Harvard, Simon Fraser, Carnegie Mellon, California, and Berkeley universities)

26 @CasalsTMarti

26

Influencers of Sports Statistics

Large list of sports statisticians:

Gilbert Fellingham, Mark Glickman , Tim Swartz, Luke Bornn , Nate Silver, Stephanie Kovalchik, Sam Ventura, Ben Baumer, Brian Macdonald, Katherine L Evans, Rebecca Nugent, Dimitrios Karlis John Newell, Marica Manisera, Paola Zuccolotto,....

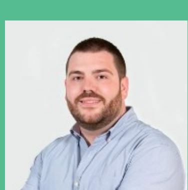
Jim Albert

Michael Lopez

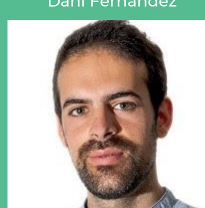
Sports statisticians in Spain?

Jordi Cortés

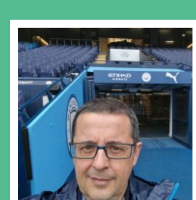
Dani Fernández



Marcos Matabuena



David Blanco



Klaus Langohr, José María Fernández Ponce, Román Salmerón, Dae-Jin Lee, Lore Zumeta

Networks of sports statistics

- ISI Sports Statistics
- BDsports Group (Italy)
- SCORE Network for Statistical Education in Sports
- Statistics In Sports Research Group (Sri Lanka)
- Carnegie Mellon Sports Analytics
- Sports Analytics Group at Berkeley

Journals to explore

- Journal of Quantitative Analysis in Sports (JQAS)
- Journal of Sports Analytics
- Statistical Analysis and Data Mining: The ASA Data Science Journal –Volume 9, Issue 5, 2016
- Data Mining and Knowledge Discovery–Volume 31, Issue 6, 2017
- Electronic Journal of Applied Statistical Analysis – Volume 10, Issue 3, 2017
- Machine Learning–2018
- International Journal of Forecasting -2018

Swartz, T. B. (2020). Where should i publish my sports paper?. The American Statistician, 74(2), 103-108.

Sports Analytics / 2

A Great Starting Point for Reading Some Basic Works

Home › AStA Advances in Statistical Analysis › Article

Estimating the change in soccer's home advantage during the Covid-19 pandemic using bivariate Poisson regression

Original Paper | Published: 27 July 2021

Volume 107, pages 205–232, (2023) Cite this article

WIREs COMPUTATIONAL STATISTICS

Overview

Big ideas in sports analytics and statistical tools for their investigation

Benjamin S. Baumer, Gregory J. Matthews, Quang Nguyen

First published: 10 May 2023 | https://doi.org/10.1002/wics.1612 | Citations: 2

Edited by: Frieben, Commissioning Editor and David Scott, Review Editor and Co-Editor-in-Chief

What was lost? A causal estimate of fourth down behavior in the National Football League

Cite

Article type: Research Article

Authors: Yam, Derrick R.[a] | Lopez, Michael J.[b,*]

Affiliations: [a] Brown University, RI, USA | [b] Skidmore College, NY, USA

Correspondence: [*] Corresponding author: Michael J. Lopez, Skidmore College, NY, USA. E-mail: mlopez1@skidmore.edu.

Keywords: "There's so much more involved with the game than just sitting there, looking at the numbers and saying, 'OK, these are my percentages, then I'm going to do it this way,' because that one time it doesn't work could cost your team a football game, and that's the thing a head coach has to live with, not the professor." - Bill Cowher, former head coach of the Pittsburgh Steelers (Garber, 2002)

DOI: 10.3233/JSA-190294

Journal: Journal of Sports Analytics, vol. 5, no. 3, pp. 153-167, 2019

Published: 26 August 2019

JOURNAL ARTICLE

HOW OFTEN DOES THE BEST TEAM WIN? A UNIFIED APPROACH TO UNDERSTANDING RANDOMNESS IN NORTH AMERICAN SPORT

Michael J. Lopez, Gregory J. Matthews and Benjamin S. Baumer

The Annals of Applied Statistics

Vol. 12, No. 4 (December 2018), pp. 2483-2516 (34 pages)

Published By: Institute of Mathematical Statistics

Unlicensed Published by De Gruyter May 29, 2020

Bigger data, better questions, and a return to fourth down behavior: an introduction to a special issue on tracking datain the National football League

Michael J. Lopez

From the journal Journal of Quantitative Analysis in Sports

https://doi.org/10.1515/jqas-2020-0057

WJARR

World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

Data science in sports analytics: A review of performance optimization and fan engagement

Ogugua Chimezie Obi [1], Samuel Onimisi Dawodu [2], Shedrack Onwusinkwue [3], Femi Osasona [4], Akoh Atadoga [5] and Andrew Ifesinachi Daraojimba [6,*]

[1] *Independent Researcher, Lagos, Nigeria.*
[2] *NDIC, Nigeria.*
[3] *Department of Physics, University of Benin, Nigeria.*
[4] *Scottish Water, UK.*
[5] *Independent Researcher, San Francisco, USA.*
[6] *Department of Information Management, Ahmadu Bello University, Zaria, Nigeria.*

World Journal of Advanced Research and Reviews, 2024, 21(01), 2663–2670

Publication history: Received on 20 December 2023; revised on 27 January 2024; accepted on 29 January 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.1.0370

31 @CasalsTMarti

31

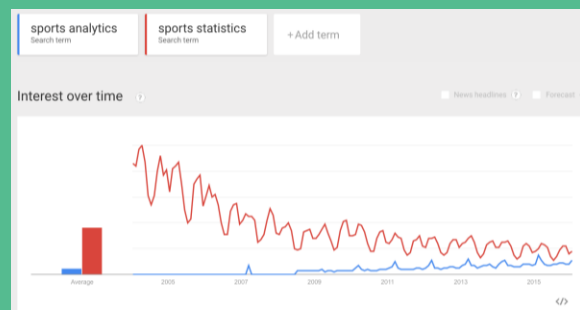
The future of Sports Science in the Sky News

32 @CasalsTMarti

32

2. Sports Statistics or Sports Analytics?

Sports and Statistics have a long history. Sports and statistics becomes sports analytics



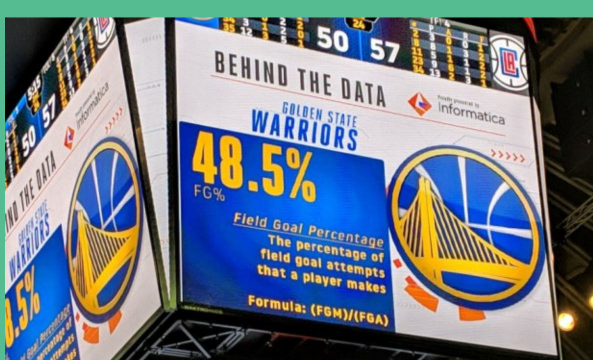
Source: Michael Lopez

33

@CasalsTMarti

33

What is really Statistics?



Science of **collecting, organizing, summarizing, presenting, analyzing data** (not just using) and situations of **uncertainty**, as well as with **drawing** valid conclusions and making reasonable decisions on the basis of such analysis. (M. R. Spiegel)

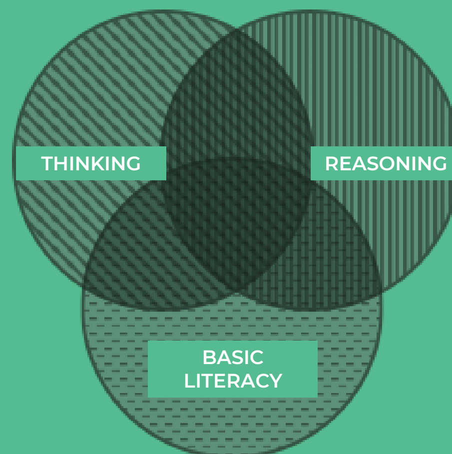
34

@CasalsTMarti

34

Data literacy

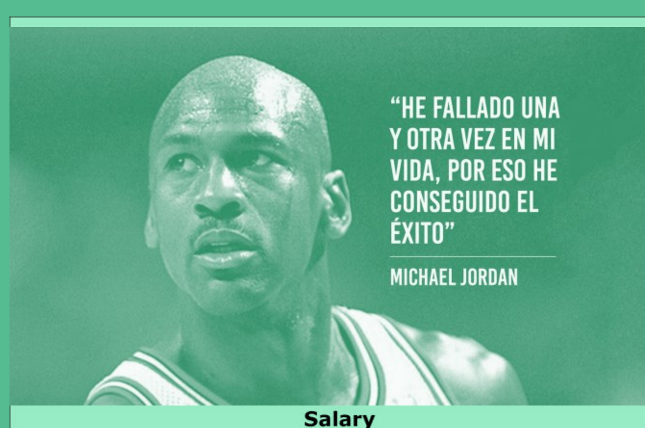
- Statistical literacy
- Statistical reasoning
- Statistical thinking



35

Statistical literacy

Statistical literacy includes basic and important skills that may be used in understanding statistical information or research results.



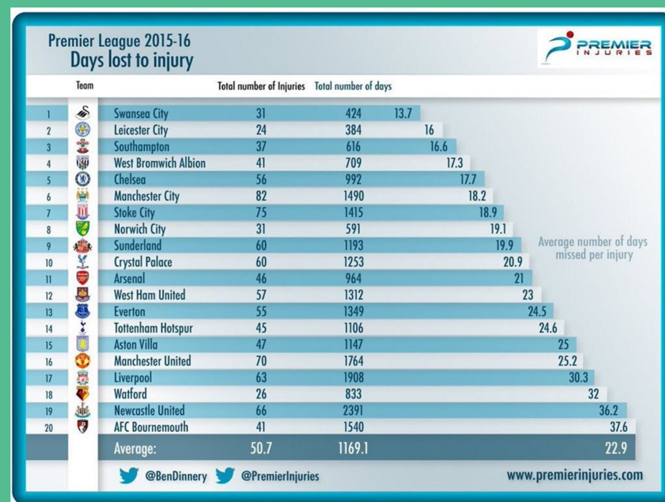
Michael Jordan's Anecdote : Mean and Median measures.
http://rpdp.net/admin/images/uploads/251H-EffectofOutliersonCenter.pdf

 Source: Ben-Zvi and Garfield (2004), and StaceyHancock (2018)

36

Statistical reasoning

Statistical reasoning may be defined as the way people reason with statistical ideas and make sense of statistical information.



Premier League 2015-16
Days lost to injury

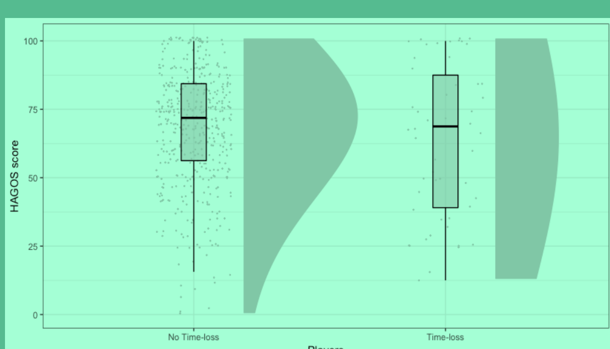
	Team	Total number of Injuries	Total number of days	Average number of days missed per injury
1	Swansea City	31	424	13.7
2	Leicester City	24	384	16
3	Southampton	37	616	16.6
4	West Bromwich Albion	41	709	17.3
5	Chelsea	56	992	17.7
6	Manchester City	82	1490	18.2
7	Stoke City	75	1415	18.9
8	Norwich City	31	591	19.1
9	Sunderland	60	1193	19.9
10	Crystal Palace	60	1253	20.9
11	Arsenal	46	964	21
12	West Ham United	57	1312	23
13	Everton	55	1349	24.5
14	Tottenham Hotspur	45	1106	24.6
15	Aston Villa	47	1147	25
16	Manchester United	70	1764	25.2
17	Liverpool	63	1908	30.3
18	Watford	26	833	32
19	Newcastle United	66	2391	36.2
20	AFC Bournemouth	41	1540	37.6
	Average:	50.7	1169.1	22.9

@BenDinnery @PremierInjuries www.premierinjuries.com

 Source: Ben-Zvi and Garfield (2004), and StaceyHancock (2018)

@CasalsTMarti

Statistical thinking



Source: Data of Esteve, E, et al (Journal of Medicine & Science in Sports; 2020). **Figure:** Raincloud plot with box plots of HAGOS, Sport/Rec subscale scores from players reporting time-loss and no time-loss groin problems.

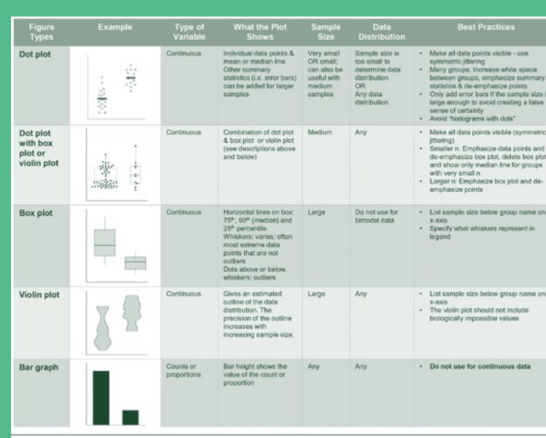


Figure Types	Example	Type of Variable	What the Plot Shows	Sample Size	Data Distribution	Best Practices
Dot plot		Continuous	Individual data points & mean or median line Other summary statistics (i.e. error bars) can be added for larger samples	Very small OR small; can also be useful with medium samples	Sample size is too small to determine data distribution OR Any data distribution	• Make all data points visible - use symmetric jittering • Many groups: Increase white space between groups, emphasize summary statistics & de-emphasize points • Only add error bars if the sample size is large enough to avoid creating a false sense of certainty • Avoid "histograms with dots"
Dot plot with box plot or violin plot		Continuous	Combination of dot plot & box plot or violin plot (see descriptions above and below)	Medium	Any	• Make all data points visible (symmetric jittering) • Smaller n: Emphasize data points and de-emphasize box plot, delete box plot and show only median line for groups with very small n • Larger n: Emphasize box plot and de-emphasize points
Box plot		Continuous	Horizontal lines on box: 75th, 50th (median) and 25th percentile Whiskers: varies; often most extreme data points that are not outliers Dots above or below whiskers: outliers	Large	Do not use for bimodal data	• List sample size below group name on x-axis • Specify what whiskers represent in legend
Violin plot		Continuous	Gives an estimated outline of the data distribution. The precision of the outline increases with increasing sample size.	Large	Any	• List sample size below group name on x-axis • The violin plot should not include biologically impossible values
Bar graph		Counts or proportions	Bar height shows the value of the count or proportion	Any	Any	• **Do not use for continuous data**

e 2. Figures for comparing groups in cross-sectional or experimental studies.

Weissgerber, et al. (2019). Reveal, don't conceal: transforming data visualization to improve transparency. Circulation, 140(18), 1506-1518.

 Source: Ben-Zvi and Garfield (2004), and StaceyHancock (2018)

@CasalsTMarti

Statistical thinking in Basketball

Lonzo Ball's Career Shooting Splits By Month

Month	G	MP	FGA	FG%	3PA	3P%	FTA	FT%	TS%
November	34	975	321	33.6%	157	26.8%	30	43.3%	40.5%
December	45	1443	496	40.3%	261	34.1%	68	47.1%	49.5%
January	40	1355	480	40.2%	276	35.1%	36	47.2%	50.4%
February	27	883	289	45.0%	186	44.6%	34	85.3%	61.2%
March	27	951	318	41.2%	196	35.2%	35	45.7%	52.0%
April	11	332	138	39.9%	94	34.0%	7	85.7%	52.5%
May	6	210	101	38.6%	65	40.0%	14	78.6%	53.7%

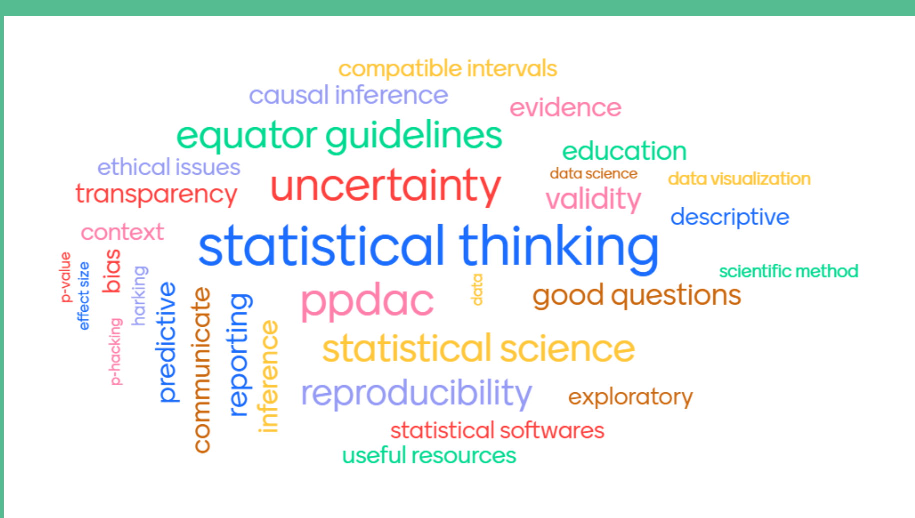
Source: Basketball-Reference.com
Graphic Credit: Owen Phillips

Can we think of a reason why Lonzo would be so much better in February than all the other months?

So how consistent has this February pattern been in Lonzo's career?

 Source: Ben Falk (2021)

Statistical thinking

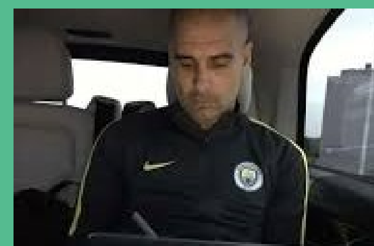


 Poldrack, R. A. (2019). *Statistical Thinking for the 21st Century*.

Observation

Report ideas

Report information



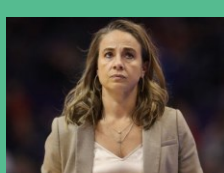
"Intelligence is the ability to adapt to change."

Stephen Hawking

41

The thinking of coaches

Analysis on overall performance metrics and on sufficiently large sets of players are difficult to apply. Within a team, what is important is to study very specific questions about players or very specific sets of players <Sergi Oliva>



Less impactful	Sometimes impactful	Most impactful
Hypothesis testing		Finding best estimand
p-values	Machine Learning	Relatable metrics
Traditional plots		Beeswarm, Joy/Ridge plots

Source: Michael Lopez

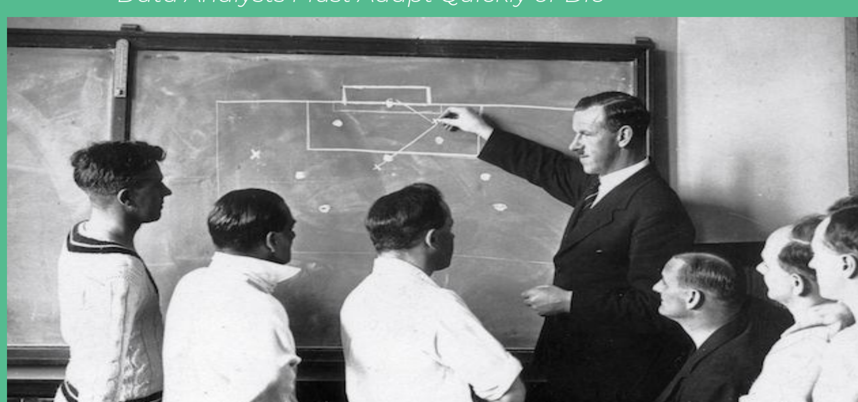
42

“Does science really help? Not much! When scientists begin to think they are coaches, the results can be a disaster”
Dr. Peter Snell

43

The Analyst

Data Analysts Must Adapt Quickly or Die

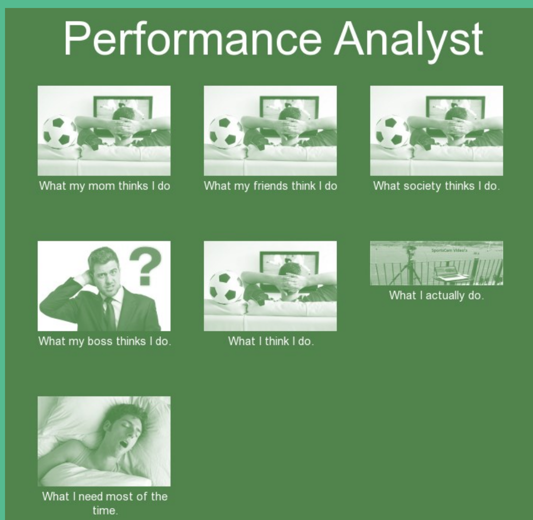


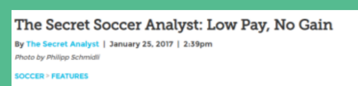
“It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is most adaptable to change.” <C.Darwin>

44 @CasalsTMarti

44

What idea is there or was there?




The Secret Soccer Analyst: Low Pay, No Gain

By The Secret Analyst | January 25, 2017 | 2:39pm

Photo by Philipp Schmidli

SOCCER > FEATURES

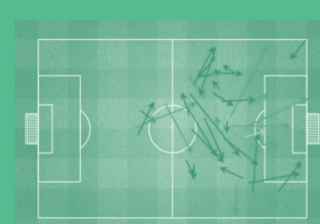
What idea is there?

What skills should an analyst have?

- Know the sport in depth (not only tactical or technical aspects, but also physical, psychological, social ...)
- Knowledge of the coaching staff to which it belongs, and the competition and history in which the club is involved.
- Entrepreneur and technological. It is essential that you know how to adapt to all the software options, platforms and tools that are available today.
- Quantitative and qualitative expert, able to make sense of the data, statistics and numbers related to the sport to correctly transmit to the staff.
- Be aligned with the coach and club in terms of the game model.

What can we analyze?

- **Physical performance** (total distance, high-intensity distance, sprinting distance, High Speed Run, accelerations, ...)
- **Technical performance** (number of passes, pass accuracy, tackles won/lost, Shots on/off target, any technical event)
- **Tactical performance** (Ball possession in a certain are of the pitch, Transition, Additional thoughts [Ecological Dynamics],...)
- **Psychological performance** (Emotional intelligence,..)
- **Injury prevention** (previous injury, time-loss, injury burden, incidence, load, ...)
- **And much more....** (spatial- temporal facors, latent factors,...)



The analyst: “Do I have to know about everything?”

- Initially, photos or video screenshots were taken.
- Many hours of work. Video analysis software facilitated this work (ex: Nacsport, Longomatch, ERIC Sport, Sportcode, ...)
- New technologies, the information boom and the professionalization of sport have created the need to ask ourselves about how to train better analysts

Video Analyst or Photo Analyst ?

Tactical Analyst?

Performance Analyst or Data Analyst?

49 @CasalsTMarti

Source: Dani Pérez (Objetivo analista)

49

Reflections of the Soccer or Football Analyst

- There has been a **change in how performance is observed**, recorded and analyzed in sport.
- The role of the Video Analyst by Rob Carroll. (http://thevideoanalyst.com/contact/)

- We tend to hear and read **less about** notational analysis now and talk more about analytics. This indicates an important change in the community of practice that analyses performance in sport.
- Ben Alamar (2011) “Sports Analyst” has three components: **Data management, Predictive models and Information systems.**
- Jeremy Abramson (2014) : Sports Analytics is “the **discovery and communication of meaningful patterns in data**”
- Bill Gerard (2016) : Sports Analytics means “**analysis of tactical data to support tactics-related sporting decisions**”.

50 @CasalsTMarti

50

Reflections and tips of the Bill Gerrard (Soccer Analyst)

- Start simple when first introducing data analytics as a coaching tool.
- Analytical results are usually presented most effectively to coaches by using data visualisation and story-telling.
- Data analytics is only one input into decision making by coaches, albeit a potentially very important one if used effectively.
- Data analytics is suffering from a fixation with big-data analytics. Big-data, context-generic statistical analysis must be translated into practical solutions to small-data.
- Sports analytics is most effective when the analyst understands the specific operational context of the coach, produces relevant data analysis and translates that analysis into practical recommendations.

51 @CasalsTMarti

51

Soccer Analytics is more difficult….

Soccer Analytics is hard …..

- ✓ 22 players on the field
- ✓ Limited Player substitutions
- ✓ ¬ 38 games per season
- ✓ Free flowing action
- ✓ Large spaces and deep tactical development
- ✓ Teammate interactions and responsabilities
- ✓ Goals are rare

But it's not impossible…

52 Fuente: **Dan Cervone** @CasalsTMarti

52

Reflections and tips of Basketball Analytics

- Basketball analytics is not just fancy statistics.
- Basketball analytics is about asking the right questions.
- Asking the right questions is a more important basketball analytics skill than manipulating numbers.
- Context is important

53 @CasalsTMarti

53

3. Resources of Sports Analytics

More people are exercising and watching athletes on TV and in stadiums

61 @CasalsTMarti

61

Guide to Sports Analytics

In addition to 200+ R resources, new sports-focused Python links are now included! Over 100+ python tutorials, 30+ packages, 25+ accounts to follow, 10 cheatsheets, and several free books & blogs now all in one place !

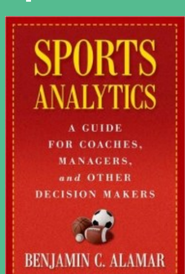
Source: @DSamangy

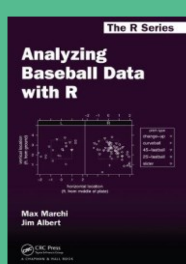
Blogger // Resource	Description	Link	Type
Ben Alamar	Sports Analytics: A Guide for Coaches, Managers, and Other Decision Makers	https://www.amazon.com/gp/produ	Book
Wayne L. Winston	Mathletics: How Gamblers, Managers, and Sports Enthusiasts Use Mathematics in Base	https://www.amazon.com/dp/06911	Book
Tobias Moskowitz & Lon Wertheim	Scorecasting: The Hidden Influences Behind How Sports Are Played and Games Are W	https://www.amazon.com/Scorecas	Book
Dean Oliver	Basketball on Paper: Rules and Tools for Performance Analysis	https://www.amazon.com/Basketba	Book
Seth Partnow	The Midrange Theory	https://www.amazon.com/Midrange	Book
Kirk Goldsberry	Sprawlball: A Visual Tour of the New Era of the NBA	https://www.amazon.com/SprawlBa	Book
Ben Taylor	Thinking Basketball	https://www.amazon.com/Thinking-	Book
Stephen Shea & Christopher Baker	Basketball Analytics: Objective and Efficient Strategies for Understanding How Teams W	https://www.amazon.com/Basketba	Book
Stephen Shea	Basketball Analytics: Spatial Tracking	https://www.amazon.com/Basketba	Book
David Sumpter	Soccermatics: Mathematical Adventures in the Beautiful Game Pro-Edition	https://www.amazon.co.uk/Soccern	Book
Christoph Biermann	Football Hackers: The Science and Art of a Data Revolution	https://www.amazon.com/Football-l	Book
Chris Anderson	The Numbers Game: Why Everything You Know About Soccer Is Wrong	https://www.amazon.com/Numbers	Book
James Tippett	The Expected Goals Philosophy: A Game-Changing Way of Analysing Football	https://www.amazon.com/Expected	Book
on Kuper	Soccernomics	https://www.amazon.com/Soccerno	Book

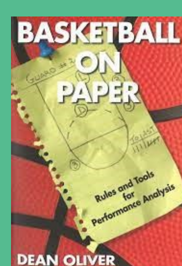
Py Accounts · Data Resources · Cheatsheets · Blogs // Books · Shiny Apps · Researc

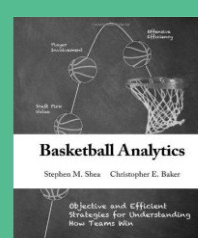
https://docs.google.com/spreadsheets/d/16Xvhl7fCKEs1JTr-VXPZDmctO2gq4TcmuNmAhoHQQs0/edit#gid=507593318

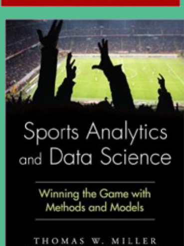
Sports Statistics literature

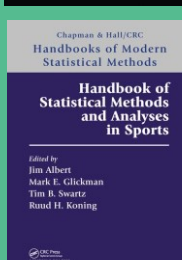


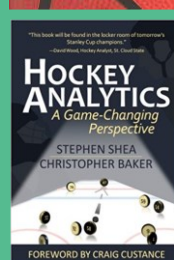


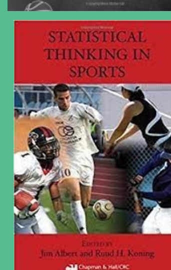












Who to follow in Soccer Analytics?

Jan Van Haaren
2.409 Tuits

Jan Van Haaren
@JanVanHaaren Et segueix

Seguint

Football Data Scientist @ClubBrugge. Research Fellow @KU_Leuven. Data Editor for FIFA series @EASports. PhD in Machine Learning. MSc in Computer Science.

Knokke-Heist, Belgium janvanhaaren.be
Data en què s'hi va unir: juny de 2009

1.490 Seguint **5.485** Seguidors

You should read it: **"Soccer Analytics 2020 Review"**

David Sumpter
@Soccermatics

Here is my list, based on some research I made during my course, who should get credit for first application of things in football analytics. Expected goals, Expected Threat, Pitch Control etc.....

12:52 p. m. · 20 de des. de 2020 · Twitter Web App

Luke Bornn
@LukeBornn

My lab has had 11 Sloan papers over the last 5 years:

'14: EPV
'15: Counterpoints, Move or Die
'16: Pressing Game, Court Realty
'17: Possession Sketches, Scorekeeper Bias
'18: Open Spaces, NFL Injury, NBA Replay, Deep Learning Trajectories

here's a summary thread of them all:

5:27 p. m. · 20 de febr. de 2018 · Twitter Web Client

More Resources….

Podcast Luke Born (2021)

https://podcasts.apple.com/gb/podcast/training-ground-guru-podcast/id1458881321?i=1000532842141&s=03

Sports Analytics Conferences

http://sportsanalyticsconferences.com/

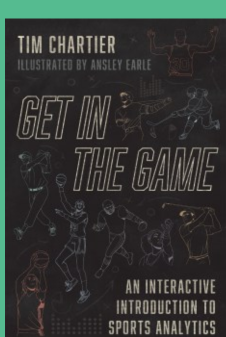
Sports Analytics Resources

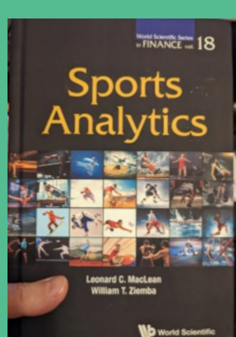
https://www.samford.edu/sports-analytics/resources

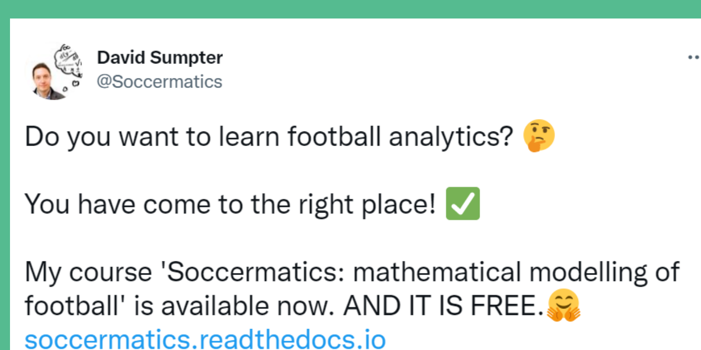
66 @CasalsTMarti

66

NEW BOOKS and COURSE☺





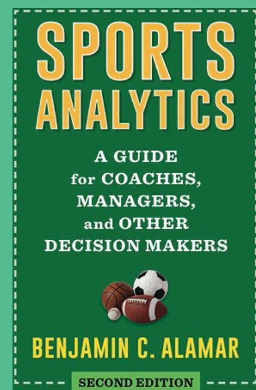


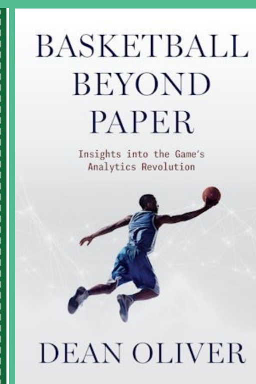
67 @CasalsTMarti

67

Future of Basketball Analytics?

- 53% use statistical analysis regularly in decision making.
- Only 32% report data are presented clearly and consistently.
- 83% expect their analytics teams to grow in the next five years.
- Improving data management and integration is crucial.
- Decision-making processes need to better utilize available data.

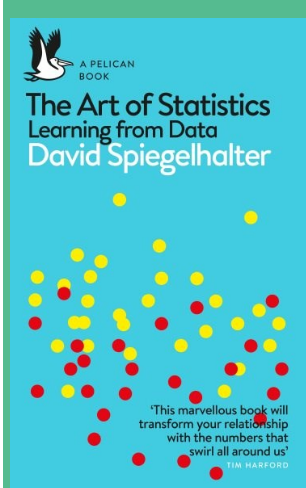


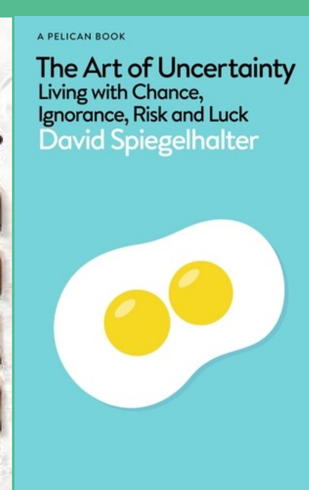


68 @CasalsTMarti

68

Recommended books of Statistics

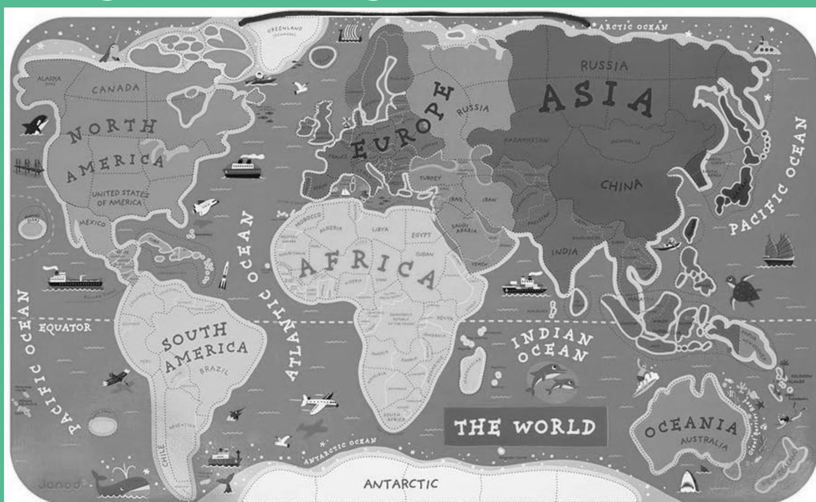




69 @CasalsTMarti

69

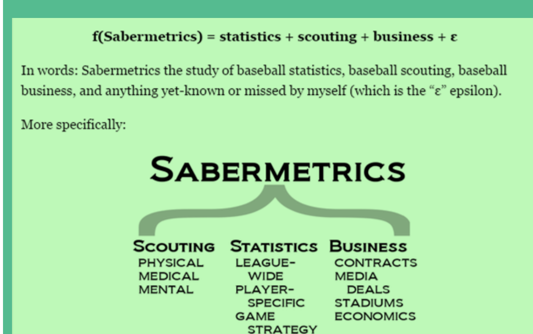
Where to Study Sports Analytics: Global Degree and Postgraduate Programs



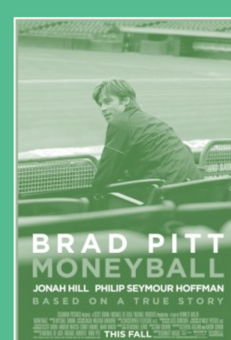
70 @CasalsTMarti

70

4. Specialitzations in Sports and data



The revolution of the roles in the world of data and sports!!



71 @CasalsTMarti

71

Who could be interested in statistics applied to sport?

- Sports managers and Decision-makers
- Players
- Journalists
- Bookmakers
- Scouts and video analysts
- Researchers (in other fields)
- Fans
- Mathematicians and Statisticians
- Physiotherapists
- Psychologists
- Sports Scientists and Sports medicine experts
- **Trainers and** coaches
- S&C coaches
- Fitness coaches
- Academics and students

Casals M, Finch CF. Sports Biostatisticians – a critical member of all sports science and medicine teams for injury prevention. *Injury Prevention. DOI: 0.1136/injuryprev-2016-042211*

72 @CasalsTMarti

72

Sports Analytics / 2

Specialisations in the field of statistics and sports statistics

BOX 1

Most common specialisations in the field of statistics

BIOSTATISTICS

Statistics and Epidemiology, and Public Health | Medicine

BIOINFORMATICS

Statistics and Computer Science, and Biology and Genetics

GEOSTATISTICS

Statistics and Geography

PSYCHOMETRICS

Statistics and Psychology

ECONOMETRICS

Statistics and Economics

73 @CasalsTMarti

73

Sports Analytics / 2

Specialisations in the field of statistics and sports statistics

BOX 2

Statistics and sports science specialisations

SABERMETRICS

Statistics and Baseball

MONEYBALL

Statistics and Sport Science, and Economics and Computer Science

SPORTS ANALYST

Statistics and Sport Science, and Video Analyst and Computer Science

SPORTS BIOSTATISTICIAN

Statistics and Epidemiology, and Public Health |Medicine and Sports Science

Casals M, Finch CF. Sports Biostatisticians – a critical member of all sports science and medicine teams for injury prevention. *Injury Prevention. DOI: 0.1136/injuryprev-2016-042211*

74 @CasalsTMarti

74

One of my contributions: Sports Biostatistics

IP Online First, published on December 30, 2016 as 10.1136/injuryprev-2016-042211

Special feature

Sports Biostatistician: a critical member of all sports science and medicine teams for injury prevention

Martí Casals,[1,2,3] Caroline F Finch[4]

[1]Sport Performance Analysis Research Group, University of Vic, Barcelona, Spain
[2]Research Centre Network for Epidemiology and Public Health (CIBERESP), Spain
[3]Epidemiology Service, Public Health Agency of Barcelona, Barcelona, Spain
[4]Australian Collaboration for Research into Sports and its Prevention, Federation University Australia, Ballarat, Australia

Correspondence to
Dr Martí Casals, Epidemiology Service, Public Health Agency of Barcelona, Pza Lesseps, 1, Barcelona 08023, Spain;

ABSTRACT
Sports science and medicine need specialists to solve the challenges that arise with injury data. In the sports injury field, it is important to be able to optimise injury data to quantify injury occurrences, understand their aetiology and most importantly, prevent them. One of these specialty professions is that of Sports Biostatistician. The aim of this paper is to describe the emergent field of Sports Biostatistics and its relevance to injury prevention. A number of important issues regarding this profession and the science of sports injury prevention are highlighted. There is a clear need for more multidisciplinary teams that incorporate biostatistics, epidemiology and public health in the sports injury area.

researchers (in other fields) and statisticians (to design and develop new statistical analysis models).

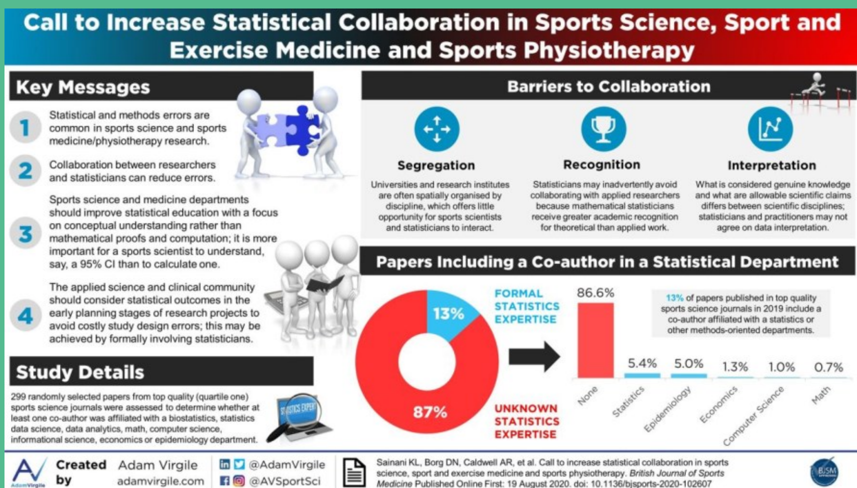
Although scouting, sabermetrics and the book *Moneyball* had already previously given importance to statistics in this field, the Hollywood film of *Moneyball* was the trigger to help awaken the interest in analytics in sports science.[5] In sports science, for example, the combination of statistics with the passion and knowledge of baseball has led to a new profession called sabermetricians. Similarly, the combination of strong statistical skills with the passion and knowledge of sports science and with economics and computer science would enable more sports data scientists to become like the stars of *Moneyball* (box 2).

Currently, most of the discussion about data in

75 @CasalsTMarti

75

Sports Biostatisticians can help…



76

Sports Biostatistics applied in the scientific field

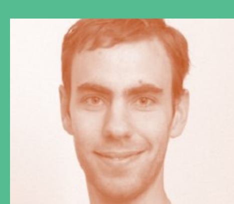
When is a study result important for athletes, clinicians and team coaches/staff? (2017)

Seven sins when interpreting statistics in sports injury science (2018)

Training load and structure-specific load: applications for sport injury causality and data analyses (2018)

Are prevalence measures better than incidence measures in sports injury research? (2019)

Time-to-event analysis for sports injury research part 1: time-varying exposures (2019)



Rasmus Nielsen

BJSM The Journal of Sport & Exercise Medicine

Time-to-event analysis for sports injury research part 2: time-varying outcomes (2019)

Randomised controlled trials (RCTs) in sports injury research: authors—please report the compliance with the intervention (2020)

Methods matter: exploring the 'too much, too soon' theory, part 1: causal questions in sports injury research (2020)

Statement on methods in sport injury research from the 1st METHODS MATTER Meeting, Copenhagen, 2019 (2020)

77

Recent papers of Sports Biostatistics

Methods matter: clinical prediction models will benefit sports medicine practice, but only if they are properly developed and validated

Garrett S Bullock,[1,2] Tom Hughes,[3,4] Jamie C Sergeant,[5,6] Michael J Callaghan,[3,4,5] Richard Riley,[7] Gary Collins[8,9]

Estimating unbiased sports injury rates: a compendium of injury rates calculated by athlete exposure and athlete at risk methods

Joseph El-Khoury,[1] Steven D. Stovitz,[2] Ian Shrier[3]

AStA Advances in Statistical Analysis
https://doi.org/10.1007/s10182-021-00428-2

ORIGINAL PAPER

Prediction of sports injuries in football: a recurrent time-to-event approach using regularized Cox models

Lore Zumeta-Olaskoaga[1,2] · Maximilian Weigert[3] · Jon Larruskain[4] · Eder Bikandi[4] · Igor Setuain[5] · Josean Lekue[4] · Helmut Küchenhoff[3] · Dae-Jin Lee[1]

78

A sports biostatistician as head the research or analytics department of an NBA franchise

- Evans has a **doctorate in biostatistics** from Harvard.
- Katherine Evans is joining the Wizards as the first woman to head the research or analytics department of an NBA franchise.

Statisticians learn to collaborate early on in their training, and that ability to listen and take the other’s perspective is essential for good leadership. <Judith D. Singer>

MUNDO DEPORTIVO

Última hora Fútbol Fútbol Internacional Baloncesto Motor Opinión Resultados Otro Mundo UnComo Vidae

Mercado de fichajes | Última hora de rumores y traspasos, en directo

Los Wizards fichan a una científica para liderar su departamento de estadística

- La doctora Katherine Evans se convierte en la primera mujer en liderar el departamento de análisis e investigación de una franquicia
- Con estudios en Harvard y Berkeley, esta científica ya colaboró en el pasado con los Toronto Raptors como directora de estrategia

La doctora Katherine Evans será pionera al liderar el departamento de estadística avanzada de los Washington Wizards. (Monumental Basketball)

Katherine Evans

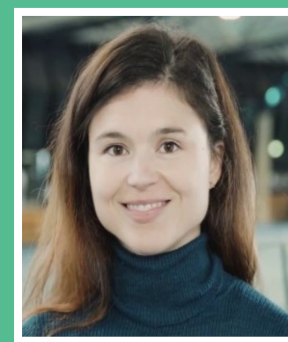
79

Influencers of Sports Biostatistics

F. Nigthtingale

Caroline Finch

Kristin L. Sainani



Stephanie Kovalchik

"There are so many females [in statistics], you never feel like you're alone in a man's world.
You see other women, and think, 'Oh, they can do it, so I can, too." Schulte

80 @CasalsTMarti

80

Conclusions

Sports Biostatisticians can help...

- To ask better questions that will improve the decision-making process in the future.
- To think in a different way taking into account study design.
- To use appropriate methodology according to the needs and objectives.
- To better interpret and understand results.
- To create a stronger relationship between the community of Sports Science & Medicine and Statistics

81 @CasalsTMarti

81

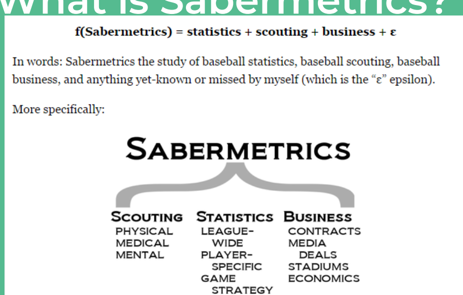
The three fields of monitoring load

LOAD	DATA SCIENCE	SPORTS & EXERCISE MEDICINE
Monitoring Training load	Effective analysis	Injury Prevention
Artefacts	Visualization	Performance
Workload	ML	Programs
External	Statistics	Recurrence
Internal	Epidemiology	Subsequent
EDUCATION	Scientific Method	Severity
	EDUCATION	EDUCATION

82 @CasalsTMarti

82

What is Sabermetrics?


f(Sabermetrics) = statistics + scouting + business + ε

In words: Sabermetrics the study of baseball statistics, baseball scouting, baseball business, and anything yet-known or missed by myself (which is the "ε" epsilon).

More specifically:

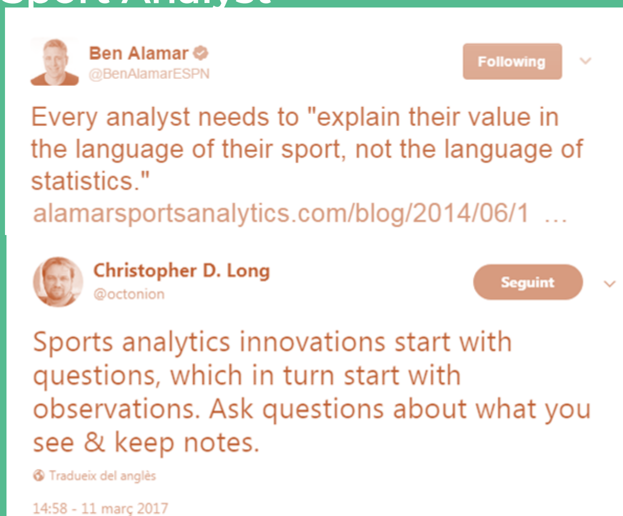
Sabermetrics is "the search for objective knowledge about baseball", often using statistical analysis to question traditional measures of baseball evaluation such as batting average and pitcher wins. < Bill James in 1980 >

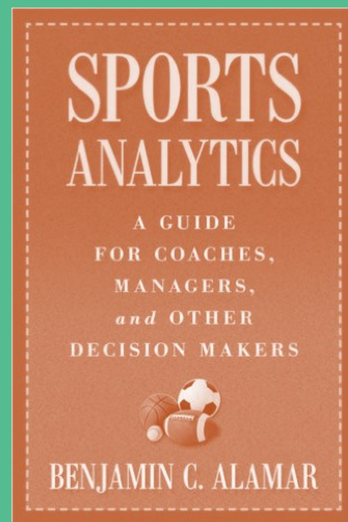
Analytics means "sabermetrics for business."

83 @CasalsTMarti

83

Sport Analyst





84 @CasalsTMarti

84

"It's not necessarily, 'does analytics create more wins?' It's 'does *not* having analytics create more losses?' And there I think that the case is really clear that it does."

— BEN ALAMAR, FORMER DIRECTOR OF SPORTS ANALYTICS AT ESPN

A Sports Analytics Podcast from *MIT SMR*

Listen on Apple Podcasts

85 @CasalsTMarti

85

Basketball knowledge from the Houston to the Sixers

Daryl Morey (Houston) and Sam Hinkie (Sixers). Both were important influencers to build basketball analytics departments with academic Phd Data scientists.

Sports Industry: Analytics staffs and current oportunities

Charlie Adkins	Coordinator, Football Analytics & Research
Meredith Manley	Football Analytics Assistant
Kevin Jordan	Football Analytics Assistant
John Taormina	Manager of Football Analytics
Emily Badis	Salary Cap and Contract Analyst
Danny Leskin	Football Analyst
David McDonald	Director of Research & Development
Scott Cohen	Director of Football Research
Corey Krawiec	Manager, Player Evaluations & Analytics
Sarah Mallepalle	Player Personnel Analyst
Derrick Yam	Quantitative Analyst
Daniel Stern	Football Analyst
Dennis Lock	Director of Football Research and Strategy
Luis Guilamo	Director of Analytics & Application Development
Shuler Cotton	Data Analyst
Evan Weiss	Football Analyst
Taylor Rajack	Director of Football Analytics
Bennie Contrino	Assistant - Football Analytics
Brad Goldsberry	Director of Football Analytics/Research
Sam Francis	Football Data Analyst
Paul DePodesta	Chief Strategy Officer
Kwesi Adofo-Mensah	Vice President, Football Operations
Ken Kovash	Vice President, Player Personnel Process & Development
Andrew Healy	Vice President, Research & Strategy
Dave Giuliani	Director, Research & Strategy
Nate Sterken	Lead Data Scientist
Sam Schmall	Football Research Analyst
Tom Robinson	Director of Football Research
Adam Vonder Haar	Football Research Analyst
Alok Pattani	Football Research Consultant
Tony Lazzaro	Senior Director, Football Technology and Research
Scott Flaska	Senior Manager, Football Analytics
Emily Kuehler	Data Scientist
Caio Brighenti	Analyst, Football Information
Michael Pelfrey	Football Analytics Assistant
Mike Halbach	Director of Football Technology
Jack Prominski	Football Analytics Manager
Dawson Friedland	Football Data Analyst
Weller Ross	Assistant Director of Football Information Systems for Analytics
Curtis Goodwin	Sports Performance Data Scientist
Kevin Clark	Football Data and Applications Engineer
George Li	Senior Football Strategy Analyst/Game Management
John Park	Manager of Football Research & Strategy
Tony Khan	Chief Football Strategy Officer
Karim Kassam	Senior Vice President of Football Operations and Strategy
Eugene Shen	Vice President of Football Analytics
Arri Landsman-Roos	Vice President of Decision Science
Momin Ghaffar	Director of Strategic Research & Development
Victor Li	Quantitative Research Manager
Sam Burgess	Data Analyst
Zach Beistline	Football Database Analyst
Brandt Tilis	Director of Football Administration
Michael Frazier	Statistical Analysis Coordinator
David Christoff	Director of Football Analytics
Walter King	Football R&D
Aditya Krishnan	Director of Football Research & Analytics
Ryan Garlisch	Manager, Software Development
Jake Temme	Manager, Data and Analytics
Sarah Bailey	Manager, Data and Analytics
Max Mulitz	Manager, Coaching Analytics
Harrison Freid	Football Research Assistant
Scott Kuhn	Director, Analytics/Pro Scout
Rex Johnson	Manager, Analytics
David Blando	Football Data Analyst
Chris French	Football Analyst
Richard Miller	Director of Research
Ryan Herman	Football Research & Strategy
Ty Siam	Director of Football Data & Innovation
Courtney Kennedy	Football Data Analyst
Brian Shields	Senior Manager, Football Scouting Research Analytics
Jason Feldman	Coordinator, Football Analytics
Zach Stuart	Coordinator, Football Analytics
Alec Halaby	Vice President of Football Operations and Strategy
James Gilman	Assistant Director of Analytics
Chase Perlen	Assistant Director of Analytics
Jon Liu	Football Operations Analyst
Tosin Kazeem	Football Analyst
Jay Whitmire	Football Analyst
Brian Hampton	Vice President of Football Administration
Demitrius Washington	Manager, Football R&D
Matt Ploenzke	Football R&D Analyst
Patrick Ward	Director of Research and Development
Brian Eayrs	Research Analyst
Josh Smith	Database Architect/Application Developer
Jacqueline Davidson	Director of Football Research
N/A	
Douglas Drewry	Football Research Analyst

@SethWalder, ESPN

Ed Kupfer Basketball Analyst
Soumya Kambhampati Data Scientist
Dylan Murphy Basketball Operations Analyst
Anna Zhao Data Scientist/Developer
@SethPartnow

Sports Analytics / 2

Europe and “Futbol” Analytics?

Soccer Analytics is hard .....

22 players on the field

Limited Player substitutions

¬ 38 games per season

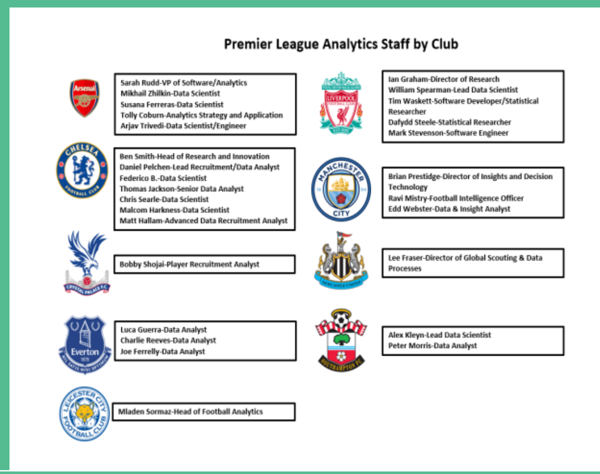
Free flowing action

Large spaces and deep tactical development

Teammate interactions and responsabilities

Goals are rare

But it’s not impossible...



Source: Dan Cervone

88 @CasalsTMarti

88

Sports Industry: Analytics staff

Houston Astros	Philadelphia Phillies
Research and Development Director (1) Research and Development Program Manager (1) Research and Development Manager (1) Advance information Director (1) Player Evaluation & Economics (1) Major League Video and Technologies (1) Baseball Technology Manager (1) Senior Research Scientist (1) Senior Analyst (1) Senior Architect (1) Senior Developer (1) Front End Engineer (1) Full Stack Developer (2) Data Modeler (1) Machine Learning Engineer (1) Player Evaluation Senior Director (1) Player Evaluation Director (1)	Strategic Initiatives Director (1) Senior Quantitative Analyst (1) Principal Engineer (1) Lead Quantitative Analyst, Player Evaluation (1) Lead Quantitative Analyst, Amateur Scouting (1) Lead Software Engineer, Application Development (1) Lead Software Engineer, Infrastructure (1) Quantitative Analyst (3) Software Engineer (4) Director, Integrative Baseball Performance (1) Coordinator, Integrative Baseball Performance (1) Assistant, Integrative Baseball Performance (1) Lead Quantitative Analyst, Integrative Baseball Performance (1)

Frequency Rank	Text	Mentions	Frequency Rank	Text	Mentions
1	Analyst	121	20	Scouting	6
2	Research and Development	93	21	Informatics	6
3	Systems	55	22	Pitching	6
4	Developer	52	23	Sports	6
5	Engineer	38	24	System	5
6	Coordinator	34	25	Database	5

Foster, G., Naidu, Z., & O'Reilly, N. (2021). Playing-Side Analytics in Team Sports: Multiple Directions, Opportunities, and Challenges. *Frontiers in Sports and Active Living*, *3*, 173.

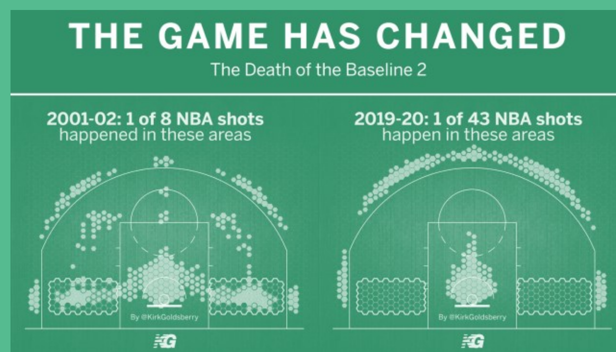
89 @CasalsTMarti

89

5. Technology and Visualization

"By prioritizing data visualization and making our reports look better, we increase the likelihood that coaches and athletes linger, learn and remember"

<Johann Windt 2022>



90 @CasalsTMarti

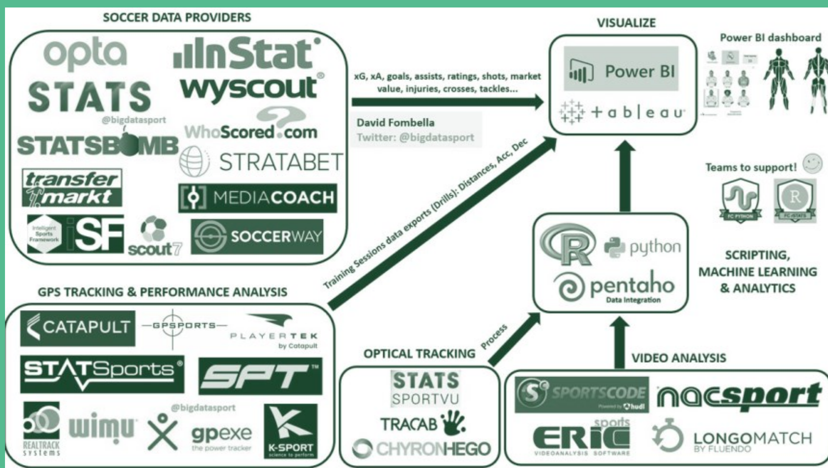
90

Video Clip Wimu

91 @CasalsTMarti

91

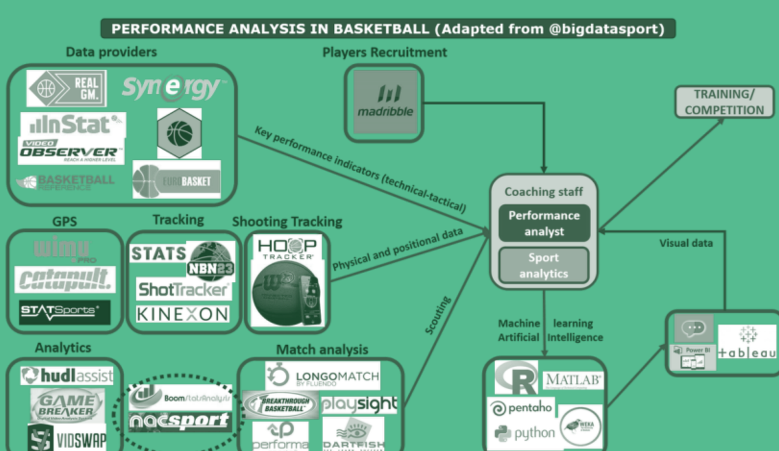
Soccer Data Providers



Source: David Fombella

@CasalsTMarti

Basketball Data Providers



Source: David Fombella

@CasalsTMarti

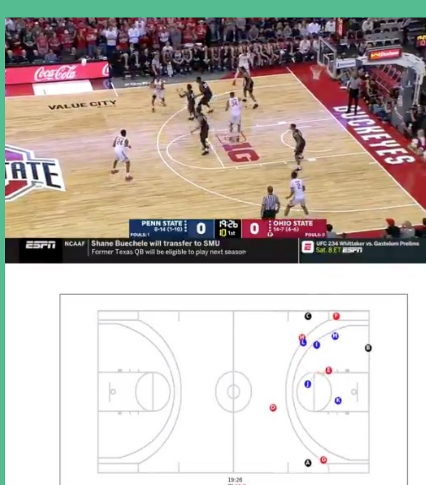
The technology: ”Measurement error”, “Validations”

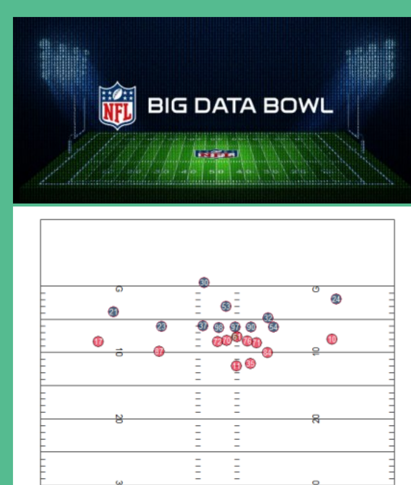
- In sport many validations are made using tests or other instruments.
- Validity and reliability cannot be assured many times even though many use it. <u>**Less than 15% of the technologies are validated. We have to know what we are measuring.**</u>
- Can we ensure that the value we register after an appraisal is reliable? If not, it depends on the person responsible for measuring it.
- Are we treating data based on records that we don’t know are true? Do we have responsibility?

94 Peake, J., Kerr, G. K., & Sullivan, J. P. (2018). A critical review of consumer wearables, mobile applications and equipment for providing biofeedback, monitoring stress and sleep in physically active populations. Frontiers in Physiology. @CasalsTMarti

94

The technology: Computer vision and science



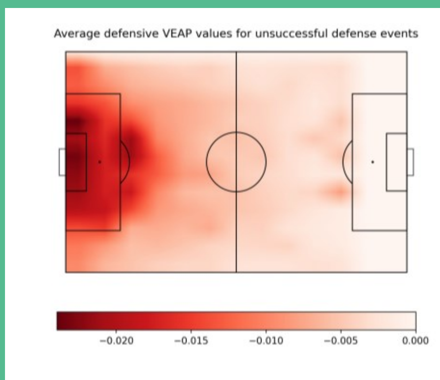


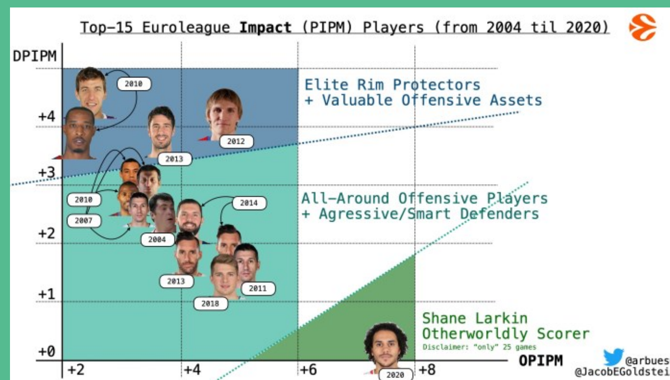
95 https://github.com/nfl-football-ops/Big-Data-Bowl @CasalsTMarti

95

Sports Analytics / 3

Visualizations and Metrics with Contextual Constraints





Source: Van Hareen

Source: Adrià Arbués

96 @CasalsTMarti

96

Now, a new revolution in Sports Analytics? Zelus Analytics

A sports intelligence platform that intend to help the professional teams in our exclusive partner network compete and win championships.

ZELUS

Which professional members work in the team of Zelus Analytics?

You can answer it☺

97 @CasalsTMarti

97

Now, a new revolution in Soccer Analytics? Kognia

Our Value Proposition

Kognia Sports Intelligence embrace a combination of proprietary, cutting-edge technologies and world-class football knowledge to bring to the market a disruptive proposition based on three main differentiators:

- **Tactical analysis of the game**
- **Automated process**
- **In real time**

Kognia
SPORTS INTELLIGENCE

Do we have enough with Technology in sports?

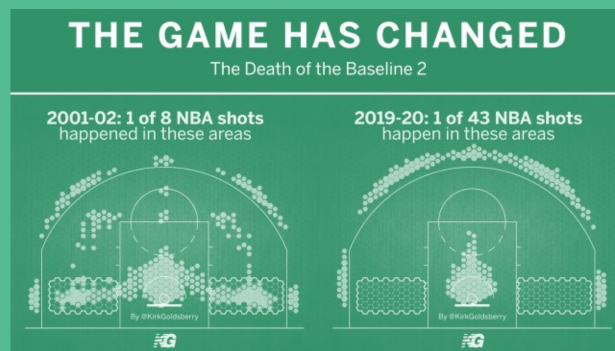
Success Comes From Better Data, Not Better Analysis

<Daryl Morey 2011>

6. Types of data, variables, and c-speak

"By prioritizing data visualization and making our reports look better, we increase the likelihood that coaches and athletes linger, learn and remember"

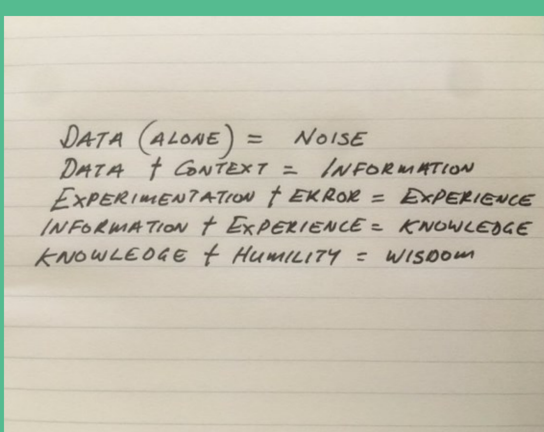
<Johann Windt 2022>



101 @CasalsTMarti

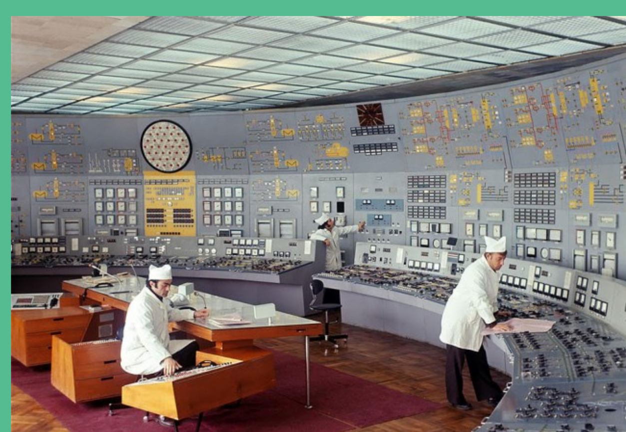
101

Type of Data, Variables and C Speak



Without data, you're just another person with an opinion."

<W. Edwards Deming>



Big information needs a big control room

102 @CasalsTMarti

102

Type of Data: *Boxscore* in Basketball

- ✓ 30 records approx
- ✓ Summary of data in frequency table
- ✓ Little context in basketball
- ✓ Pace and possessions matter!
- ✓ Standardize factors by minutes played
- ✓ Measurement, Metric, Effectiveness, Efficiency are not the same!

J 46 | 26/06/2014 | 19:15 | Palau Blaugrana | Público:7537 — 1 2 3 4

Árb: Daniel Hierrezuelo, M.A. Pérez Pérez, J.R. García Ortiz — 25|19 22|23 18|15 18|24

FC BARCELONA 83										REB						TAP			FP			
D	Nombre	Min	P	T2	T2 %	T3	T3 %	T1	T1 %	T	D+O	A	BR	BP	C	F	C	M	F	C	+/-	V
0	Pullen, Jacob																					
6	Dorsey, Joey	8:31	2	1/2	50%	0/0	0%	0/0	0%	4	3+1	0	0	2	0	0	1	0	5	0	-2	-3
8	Sada, Víctor	17:7	2	1/2	50%	0/0	0%	0/0	0%	4	4+0	3	0	0	0	0	0	0	0	0	-7	8
9	Huertas, M.	22:53	9	1/3	33%	1/2	50%	4/4	100%	2	1+1	7	0	3	0	0	0	0	3	2	9	11
10	Abrines, Álex	11:25	8	0/0	0%	2/3	67%	2/2	100%	1	1+0	0	0	1	0	0	0	0	4	1	2	4
11	Navarro, J.C.	31:24	14	1/6	17%	1/3	33%	9/10	90%	3	2+1	2	2	7	0	0	0	0	3	8	6	11
14	Todorovic, M.																					
16	Papanikolaou, K.	18:8	3	1/3	33%	0/1	0%	1/2	50%	3	2+1	0	0	1	0	1	0	1	2	1	6	1
24	Oleson, Brad	19:17	10	0/2	0%	2/3	67%	4/4	100%	0	0+0	1	0	0	0	0	1	0	0	2	-11	9
30	Lampe, Maciej	19:45	3	0/1	0%	1/2	50%	0/0	0%	6	6+0	1	0	1	0	2	1	0	4	1	-8	5
34	Nachbar, Bostjan	21:8	10	2/2	100%	2/5	40%	0/0	0%	5	4+1	0	0	1	0	0	0	0	5	1	7	7
44	Tomic, Ante	30:22	22	8/12	67%	0/0	0%	6/8	75%	9	7+2	0	2	5	0	0	1	1	3	10	8	28
	Equipo		0	0/0	0%	0/0	0%	0/0	0%	1	1+0	0	0	2	0	0	0	0	1	2	0	0
	Total	200:0	83	15/33	45%	9/19	47%	26/30	87%	38	31+7	14	4	23	0	3	4	2	30	28	2	81
E	Pascual, Xavi																					
5f	Dorsey, Joey, Nachbar, Bostjan																					

REAL MADRID 81										REB						TAP			FP			
D	Nombre	Min	P	T2	T2 %	T3	T3 %	T1	T1 %	T	D+O	A	BR	BP	C	F	C	M	F	C	+/-	V
4	Draper, Dontaye																					
5	Fernández, Rudy	28:8	13	3/5	60%	1/5	20%	4/6	67%	3	1+2	2	2	0	1	0	0	1	5	8	-5	15
9	Reyes, Felipe	14:29	11	4/10	40%	0/0	0%	3/4	75%	4	0+4	0	2	1	1	0	1	0	3	4	-6	9
11	Díez, Daniel	0:24	0	0/0	0%	0/0	0%	0/0	0%	0	0+0	0	0	0	0	0	0	0	0	0	0	0
12	Mirotic, Nikola	20:28	11	3/6	50%	0/1	0%	5/5	100%	0	0+0	0	1	3	0	0	0	0	5	5	2	5
13	Rodríguez, S.	19:20	5	0/5	0%	1/5	20%	2/2	100%	1	0+1	3	1	3	0	0	2	0	3	1	8	-6
20	Carroll, Jaycee	14:56	5	1/2	50%	0/2	0%	3/3	100%	1	1+0	2	3	1	0	0	0	0	1	3	0	9
21	Darden, Tremmell	22:37	6	1/3	33%	1/3	33%	1/1	100%	7	5+2	0	2	0	0	1	0	0	1	1	0	12
23	Llull, Sergio	34:48	8	2/6	33%	1/5	20%	1/1	100%	5	4+1	3	1	1	0	0	0	0	2	3	-6	9
30	Bourousis, I.	24:36	12	5/6	83%	0/0	0%	2/2	100%	7	5+2	0	0	0	0	1	0	0	3	1	-14	17
44	Slaughter, M.	5:26	0	0/0	0%	0/0	0%	0/0	0%	1	1+0	0	1	0	0	0	0	0	1	0	2	1
50	Mejri, Salah	14:48	10	4/4	100%	0/0	0%	2/5	40%	4	4+0	1	0	0	0	2	0	2	2	3	9	15
	Equipo		0	0/0	0%	0/0	0%	0/0	0%	1	1+0	0	1	3	0	0	0	0	2	1	0	-2
	Total	200:0	81	23/47	49%	4/21	19%	23/29	79%	34	22+12	11	14	12	2	4	3	3	28	30	-2	84
E	Laso, Pablo																					
5f	Fernández, Rudy, Mirotic, Nikola																					

103 Source: Kubatko, J., Oliver, D., Pelton, K., & Rosenbaum, D. T. (2007). A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports*, *3*(3), 1-22.

@CasalsTMarti

103

Boxscore: *Descriptive, ratios and metrics*

- ✓ All quotients are not proportions
- ✓ Example - D. Adams:

- **Ratio explanation: 3 (3P)/ 6 (3PA)**

Ratio: 50 three-pointers made per 100 attempts

- **Rate explanation: 3 (3P) in 21 min**

For every 10 minutes played by Darius Adams, he makes 1.4 three-pointers.

- ✓ OER and DER rates?

Just tools for staff to understand the value of each possession

Bilbao Basket

Starter	MP	FG	FGA	3P	3PA	FT	FTA	ORB	TRB	AST	STL	BLK	TOV	PF	PTS
Dairis Bertans	33	4	13	1	9	2	2	0	0	1	0	0	3	3	11
Alex Mumbru	31	5	16	0	8	11	12	0	3	2	2	0	3	3	21
Clevin Hannah	21	1	3	1	3	4	4	1	3	1	1	0	2	5	7
Axel Hervelle	20	1	2	0	1	2	3	1	3	0	0	0	2	5	4
Giorgos Bogris	16	1	6	0	0	0	0	0	0	1	0	0	1	4	2
Reserves	**MP**	**FG**	**FGA**	**3P**	**3PA**	**FT**	**FTA**	**ORB**	**TRB**	**AST**	**STL**	**BLK**	**TOV**	**PF**	**PTS**
Alex Ruoff	17	1	3	0	1	2	2	1	2	2	1	0	2	5	4
Raul Lopez	15	1	3	1	1	2	2	0	2	3	0	0	4	2	5
Alejandro Suarez	14	2	3	2	3	0	0	0	1	1	1	0	0	0	6
Tautvydas Slezas	13	1	3	0	0	0	0	2	3	0	0	0	1	1	2
Mirza Begic	11	0	1	0	0	0	0	0	2	0	0	1	2	2	0
Dejan Todorovic	5	0	0	0	0	0	0	0	0	0	0	0	2	1	0
Borja Mendia	4	0	2	0	1	0	0	0	0	0	0	0	0	1	0
Totals		**17**	**55**	**5**	**27**	**23**	**25**	**5**	**19**	**11**	**5**	**1**	**22**	**32**	**62**

Laboral Kutxa Baskonia

Starter	MP	FG	FGA	3P	3PA	FT	FTA	ORB	TRB	AST	STL	BLK	TOV	PF	PTS
Adam Hanga	23	2	4	0	2	2	2	1	5	2	2	0	1	2	6
Darius Adams	21	9	15	3	6	9	10	0	3	1	3	0	1	2	30
Ilimane Diop	21	2	2	0	0	4	4	0	4	0	0	0	2	3	8
Fabien Causeur	20	2	6	0	2	1	3	0	2	2	1	0	2	2	5
Kim Tillie	17	3	3	0	0	0	0	3	3	0	1	0	1	2	6
Reserves	**MP**	**FG**	**FGA**	**3P**	**3PA**	**FT**	**FTA**	**ORB**	**TRB**	**AST**	**STL**	**BLK**	**TOV**	**PF**	**PTS**
Jaka Blazic	20	3	7	0	1	3	4	1	3	1	4	0	2	4	9
Ioannis Bourousis	19	4	8	2	4	4	4	0	3	1	1	0	0	1	14
Davis Bertans	19	3	5	1	3	5	6	1	5	0	0	0	1	4	12
Mike James	19	3	7	1	2	6	6	0	2	6	0	0	1	3	13
Alberto Corbacho	16	1	3	1	3	0	0	0	2	0	0	1	0	1	3
Mamadou Diop	6	0	0	0	0	2	2	0	3	0	0	0	0	0	2
Totals		**32**	**60**	**8**	**23**	**36**	**41**	**6**	**35**	**13**	**12**	**1**	**11**	**24**	**108**

104

@CasalsTMarti

104

Basketball metrics: Teams

Four factors for the 2017 NBA season. **8 FACTORS?**

Team	W	Possessions	eFG%	TO%	OREB%	FTRate	OppeFG%	OppTO%	DREB%	OppFTRate
Golden State Warriors	67	8169	0.563445378	0.148243359	0.228215768	0.204061625	0.485501	0.15571	0.749	0.1983313
San Antonio Spurs	61	7700	0.524256993	0.142987013	0.239988308	0.20979021	0.492177	0.15377	0.7763	0.1922796
Houston Rockets	55	8160	0.544672819	0.151960784	0.245933278	0.233081655	0.519489	0.15135	0.7577	0.1942081
Boston Celtics	53	7899	0.524577243	0.137738954	0.211784799	0.220120378	0.502559	0.14445	0.7534	0.2231699
Cleveland Cavaliers	51	7885	0.546962516	0.142168675	0.219484241	0.205514864	0.515925	0.13101	0.7577	0.174492
LA Clippers	51	7914	0.53702346	0.13419257	0.214659686	0.23255132	0.505715	0.13697	0.7692	0.2114588
Toronto Raptors	51	7757	0.516551026	0.134201367	0.249570201	0.233448974	0.507131	0.1547	0.763	0.2233495
Utah Jazz	51	7509	0.526247122	0.148888001	0.231879699	0.21504221	0.49286	0.13251	0.789	0.194556
Washington Wizards	49	8000	0.527672692	0.14525	0.240582192	0.199103265	0.524219	0.15713	0.7547	0.2133849
Oklahoma City Thunder	47	7915	0.499790766	0.155401137	0.279497908	0.219695913	0.511146	0.14264	0.79	0.2179194
Atlanta Hawks	43	7908	0.504119688	0.163631765	0.236251403	0.214512865	0.506569	0.15959	0.7612	0.1788959
Memphis Grizzlies	43	7531	0.491318938	0.140618776	0.247624371	0.219433907	0.505878	0.15549	0.775	0.2564578
Indiana Pacers	42	7878	0.515870726	0.143437421	0.21242485	0.211657769	0.512303	0.1555	0.7545	0.2253582
Milwaukee Bucks	42	7774	0.5274758	0.147285825	0.215242632	0.210424423	0.518083	0.155	0.7539	0.2152308
Chicago Bulls	41	7800	0.487396723	0.142564103	0.270219097	0.206413668	0.507325	0.14103	0.7684	0.1686153
Miami Heat	41	7750	0.512434276	0.142193548	0.241618177	0.177348302	0.496451	0.14465	0.7637	0.2198725
Portland Trail Blazers	41	7922	0.519762006	0.142135824	0.230044346	0.214336308	0.508417	0.13279	0.7682	0.2413354
Denver Nuggets	40	8047	0.529886016	0.152354915	0.272727273	0.213650264	0.532285	0.12203	0.7865	0.1935706
Detroit Pistons	37	7735	0.492241143	0.125791855	0.240912709	0.156550398	0.516471	0.13523	0.8122	0.1970717
Charlotte Hornets	36	7804	0.500714286	0.12070733	0.19862259	0.227285714	0.523407	0.13724	0.7958	0.1637056
New Orleans Pelicans	34	8089	0.504067321	0.130300408	0.185292567	0.191865358	0.509316	0.14143	0.7675	0.1771476
Dallas Mavericks	33	7539	0.504740741	0.128929566	0.181232493	0.18	0.529412	0.16249	0.7758	0.2178478
Sacramento Kings	32	7801	0.515815266	0.153954621	0.210047004	0.22022572	0.527696	0.14678	0.7625	0.2265533
Minnesota Timberwolves	31	7768	0.510762785	0.148043254	0.272199652	0.22912453	0.534705	0.15049	0.7591	0.2115976
New York Knicks	31	7837	0.495589857	0.145336226	0.265819362	0.187982359	0.511238	0.13526	0.7408	0.2140165
Orlando Magic	29	7884	0.489205103	0.138381532	0.215717723	0.183513248	0.522666	0.13559	0.7738	0.2012428
Philadelphia 76ers	28	8012	0.501072654	0.170494259	0.223607648	0.199227689	0.512353	0.15489	0.7533	0.2342752
Los Angeles Lakers	26	8048	0.500977108	0.15444831	0.249866095	0.195002792	0.541613	0.14761	0.7577	0.2125337
Phoenix Suns	24	8195	0.492768595	0.154362416	0.26040555	0.230853994	0.525106	0.14912	0.7635	0.264215
Brooklyn Nets	20	8317	0.506726778	0.163039558	0.195604992	0.227851725	0.513089	0.1313	0.7614	0.2122433

105 Oliver, D. (2004). *Basketball on paper: rules and tools for performance analysis*. Potomac Books, Inc.. @CasalsTMarti

105

Basketball metrics: Players

Basketball Player Evaluation Metrics

Adjusted Player Efficiency Rating (APER)	Adjusted Plus-Minus (APM)	Approximate Value (AV)	Assist Percentage
Box Plus-Minus (BPM)	CARMELO	Daily Updated Ranking of Individual Performance (DRIP)	DARKO (DPM)
			Defensive Plus-Minus (DPM)
Defensive Rating	Defensive Stop	Diamond Rating	Effective Field Goal Percentage (eFG%)
Estimated Plus-Minus (EPM)	Fantasy Points Scored for DraftKings (DK FPTS)	Fantasy Points Scored for FanDuel (FD FPTS)	Floor Impact Counter (FIC)
Game Score	Individual Floor Percentage	Individual Player Value (IPV)	LEBRON
NBA Efficiency	NBA Plus-Minus (+/-) & Player Impact Metrics	Net Plus-Minus (Roland Rating)	Net Points
Non-Scoring Player Possessions	Offensive Conversion Rate	Offensive Rating	Per-Minute Ratings
Personal Foul Efficiency	Player Efficiency Rating (PER)	Player Impact Estimate (PIE)	Player Tracking Plus Minus (PT-PM)
Points Created	Points Per Possession (PPP)	Points Per Shot Attempt (PTS/FGA)	Points Produced
Position Adjusted Win Score (PAWS)	Quantified Shooter Impact (qSI)	Quantified Shot Quality (qSQ)	RAPTOR
Real Plus-Minus (RPM)	Rebound Percentage	Regularized Adjusted Plus Minus (RAPM)	Scoring Player Possessions
Seasons Left	Simple Projection System (SPS)	Simple Rating System (SRS)	Statistical Player Value, SPV
Statistical Plus-Minus (SPM)	Steal Percentage	Tendex	Total Player Possessions
Touches	Trade Value	True Shooting Percentage (TS%)	Turnover Ratio
Usage Rate	Versatility Index	Win Probability Added (WPA)	Win Score
Win Share (WS/48)	Wins Above Replacement Player (WARP)		

https://www.nbastuffer.com/

106 Martínez, José Antonio. "Una revisión de los sistemas de valoración de los jugadores de baloncesto (I): Descripción de los métodos existentes." Revista Internacional de Derecho y Gestión del Deporte 10 (2010): 37-77. @CasalsTMarti

106

Basketball metrics: Players

- ✓ Subjective context
- ✓ Bias and Limitations
- ✓ On/Off
- ✓ And defensive metrics?

Defensive Ratings: Estimation vs. Counting

squared2020 / 2 weeks ago

Defensive rating, a box score calculation, is an estimation procedure that attempts to identify the **points per 100 possessions** that an NBA player yields in a game. In this calculation, a player's defensive rating is effectively **eighty percent of their team's defensive rating** plus **twenty percent of defensive points per scoring possessions when on the court.** In terms of equations, this is written as

$$\text{Defensive Rating} = 0.8 * \text{Team Defensive Rating} + 0.2 * (100 * \text{Defensive Points Per Scoring Possession} * (1 - \text{Stop}\,\%))$$

This requires construction of a **Team Defensive Rating**, a **Defensive Points Per Scoring Possession**, and a **Stop Percentage**. In this article, we take a look at the construction of defensive rating. But more importantly, as it is a box score calculation, we look to see how it compares to truth by using play-by-play data.

Kidd Score $\text{Kidd Score} = \sqrt{a \cdot r}$

 @CasalsTMarti

Proposed alternative metrics: Players

¿La Valoración ACB nos sirve?

Alternativas de métricas de rendimiento y regularidad del jugador

Ampliando horizontes sobre medición del rendimiento y regularidad en el baloncesto profesional

Expanding horizons on performance measurement and regularity in professional basketball

Román Salmerón-Gómez, Samuel Gómez-Haro
Facultad Ciencias Económicas y Empresariales. Universidad de Granada

$$RR = \overline{x} + \frac{\overline{x}}{S_x} = \overline{x} + CV^{-1},$$

Tabla 1. Valores de los índices de regularidad y rendimiento para datos inventados

Índice	J1	J2	J3
Media	10	11.2	2.8
Desviación Típica	2	8.074	0.836
Coeficiente de Variación	0.2	0.721	0.298
Media + Desviación Típica	12	19.274	3.636
Media + Coeficiente de Variación	10.2	11.921	3.098
Rendimiento-Regularidad (*RR*)	15	12.587	6.146

- ✓ Abuse of calculation of means
- ✓ Dispersion measures matter!

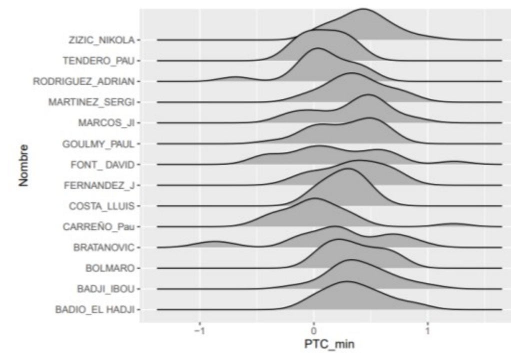
Salmerón-Gómez, R., & Gómez-Haro, S. (2016). Ampliando horizontes sobre medición del rendimiento y regularidad en el baloncesto profesional.[Expanding horizons on performance measurement and regularity in professional basketball]. RICYDE. Revista Internacional de Ciencias del Deporte. doi: 10.5232/ricyde, 12(45), 234-249.

@CasalsTMarti

Proposed alternative metrics: Players

- ✓ **Player Total Contribution (PTC/min)**

- ✓ **PIR Limitations:** Subjective and criticized
- ✓ **Effective Metrics:** Linked to game outcome
- ✓ **Regression Models:** Winscore and others
- ✓ **PTC per Minute:** Evaluates production
- ✓ **Action Importance:** Varies in impact
- ✓ **Normalization:** Points per (minutes, pace, usage)
- ✓ **Comparison Caution:** Role-specific evaluation
- ✓ **Sample Size:** Finite populations
- ✓ **Uncertainty Measures:** Errors, confidence intervals



La productividad y PTC por minuto nos da la idea de la distribución en cada uno de los jugadores respecto PTC por minuto. Parece en algunos jugadores como Jaime Fernandez, Badio o Badji que aunque tienen un PTC por minuto elevado en comparación con otros jugadores, la función de densidad nos marca su punto más alto pero también al tener una base con una amplitud grande, parecen ser jugadores más imprevisibles con esta métrica que otros a priori más **regulares como Lluis Costa o Zizic**.

109 Martínez, J. A. (2019). A more robust estimation and extension of factors determining production (FDP) of basketball players. International Journal of Physical Education, Sports and Health, 6(3),81-85

@CasalsTMarti

109

Proposed alternative metrics: Players

PTS_{mpu}= [48* (PTS/MIN)] / [PACE * USG]

PTC_{mp}= PTC / [MIN*POSS]

- ✓ Identifying Undervalued Players

- ✓ Incorporating Uncertainty in Estimates

- ✓ Normalized Metrics and Budget Management

110 Martínez, J. A. (2024). Comparaciones estadísticas y elaboración de rankings de rendimiento de jugadores de baloncesto. Retos: nuevas tendencias en educación física, deporte y recreación, (55), 170-176.

@CasalsTMarti

110

Type of Data: *Play-by-play (eventing)* in Basketball

- ✓ 400 records approx
- ✓ Variables: Time and events (Longitudinal)
- ✓ More basketball context than *boxscore*

Season	Match	Team	H / A	Player	Player Role	Quarter	Clock	Event	Home	Away	Diff.
									89	67	
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	HEURTEL, THOMAS	1	Q1	00:15	Foul	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	HEURTEL, THOMAS	1	Q1	00:15	Out	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	PRESSEY, PHIL	1	Q1	00:15	In	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	PRESSEY, PHIL	1	Q1	00:15	Turnover	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	KOPONEN, PETTERI	2	Q1	00:15	In	0	0	0
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	COLE, NORRIS	2	Q1	00:15	Foul Drawn	0	0	0
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	COLE, NORRIS	2	Q1	00:15	Missed Free Throw	0	0	0
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	COLE, NORRIS	2	Q1	00:15	Missed Two Pointer	0	0	0
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	COLE, NORRIS	2	Q1	00:15	Steal	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	HANGA, ADAM	3	Q1	00:15	Out	0	0	0
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	ROLL, MICHAEL	3	Q1	00:15	In	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	CLAVER, VICTOR	4	Q1	00:15	Def Rebound	0	0	0
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	BOLDEN, JONAH	4	Q1	00:15	Out	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	HEURTEL, THOMAS	1	Q1	00:38	Assist	0	0	0
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	CLAVER, VICTOR	4	Q1	00:41	Three Pointer	3	0	3
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	TOMIC, ANTE	5	Q1	00:53	Def Rebound	3	0	3
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	THOMAS, DESHAUN	3	Q1	00:56	Missed Three Pointer	3	0	3
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	TYUS, ALEX	5	Q1	01:02	Def Rebound	3	0	3
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	MOERMAN, ADRIEN	4	Q1	01:05	Missed Three Pointer	3	0	3
E2017	67. BAR-TEL	FC Barcelona Lassa	Home	HEURTEL, THOMAS	1	Q1	01:09	Steal	3	0	3
E2017	67. BAR-TEL	Maccabi FOX Tel Aviv	Away	COLE, NORRIS	2	Q1	01:10	Turnover	3	0	3

111 @CasalsTMarti

111

Type of Data: *Play-by-play (eventing)* in Basketball

- ✓ Different questions
- ✓ Time-to-event models

Who assists who?

	passer	shooter	n
1	CALATHES, NICK	THOMAS, DESHAUN	22
2	CALATHES, NICK	PAPAPETROU, IOANNIS	21
3	CALATHES, NICK	FREDETTE, JIMMER	16
4	CALATHES, NICK	PAPAGIANNIS, GEORGIOS	12
5	CALATHES, NICK	WILEY, JACOB	11
6	CALATHES, NICK	RICE, TYRESE	10
7	RICE, TYRESE	PAPAPETROU, IOANNIS	6
8	RICE, TYRESE	THOMAS, DESHAUN	6
9	CALATHES, NICK	MITOGLOU, KONSTANTINOS	5
10	FREDETTE, JIMMER	PAPAPETROU, IOANNIS	5
11	RICE, TYRESE	WILEY, JACOB	5
12	CALATHES, NICK	VOUGIOUKAS, IAN	4

...thes genera 3,5 (quasi 4) triples.

més punts generen per assitencies fetes
i això cal tenir en compte.

112 @CasalsTMarti

112

Type of Data: *Tracking*

✓ ✓ ✓

Expected Possession Value (EPV)

Purpose: Using film or player tracking data (Sport Vu), a quantitative analysis representative of the entire possession will occur where each moment in the possession is summarized on the basis of the expected value of the possession at each moment. For each moment of a possession a value is assigned to each of the individual tactical moves a player can make, allowing analysts to evaluate each decision that a player makes. Under this model, dishing the ball to an open shooter at the key or near the basket is worth more expected points than to a covered player in the corner. EPV is an extremely new avenue of basketball analysis focusing on decision-making, opportunity creation and prevention.

Formula: EPV is a conditional expectation – the expected number of points the offense will score, given the spatial configuration of the players and ball at time during the possession. The expected value of different plays in a possession in some sense is subjective. The current EPV of a possession is the weighted average of the outcomes of all future paths that the possession could take.

To evaluate metrics based on three criteria:

(1) Stability: Does the metric measure the same thing over time?

(2) Discrimination: Does the metric reliably differentiate between players?

(3) Independence: Does the metric provide new information?

Cervone, D, et al (2016). A multiresolution stochastic process model for predicting basketball possession outcomes. Journal of the American Statistical Association, 111(514), 585-599.

Franks, A. M., D'Amour, A., Cervone, D., & Bornn, L. (2016). Meta-analytics: tools for understanding the statistical properties of sports metrics. Journal of Quantitative Analysis in Sports, 12(4), 151-165.

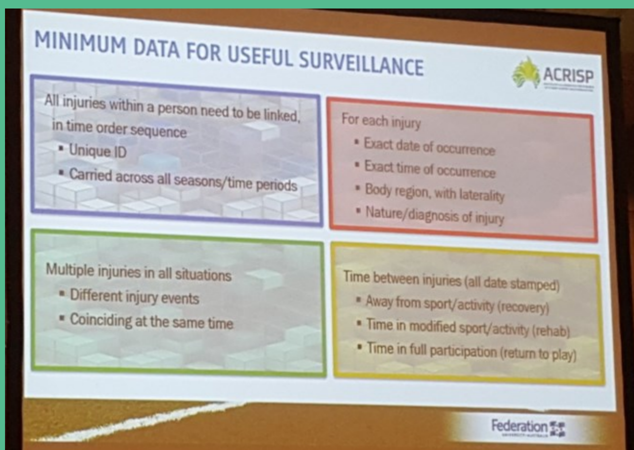
What is the unit of measurement of the data base?

- ✓ Futsal example: 1267 records of player-season-injury (or not injury)
- ✓ Tennis example: 6060 Davis Cup singles matches

tourney_id	tourney_name	group	round	surface	tourney_level	match_num	winner_name	loser_name	score	best_of	Retirement	sum_games
2000-D078	Davis Cup G1 QF: JPN vs KOR	No Mundial 1	QF	Carpet	Davis Cup	1	Hyung Taik Lee	Yaoki Ishii	4-6 4-6 6-2 6-1 10-8	5 sets	No	53
2000-D078	Davis Cup G1 QF: JPN vs KOR	No Mundial 1	QF	Carpet	Davis Cup	2	Gouichi Motomura	Yong Il Yoon	6-3 4-6 6-4 2-6 6-3	5 sets	No	46
2000-D078	Davis Cup G1 QF: JPN vs KOR	No Mundial 1	QF	Carpet	Davis Cup	3	Hyung Taik Lee	Gouichi Motomura	7-5 6-4 7-5	5 sets	No	34
2000-D078	Davis Cup G1 QF: JPN vs KOR	No Mundial 1	QF	Carpet	Davis Cup	4	Yaoki Ishii	Hee Sung Chung	6-4 6-4	3 sets	No	20
2000-D081	Davis Cup G1 R1: UKR vs POR	No Mundial 1	R1	Carpet	Davis Cup	1	Andrey Dernovskiy	Joao Cunha Silva	7-6(6) 6-4 3-6 3-6 6-4	5 sets	No	51
2000-D081	Davis Cup G1 R1: UKR vs POR	No Mundial 1	R1	Carpet	Davis Cup	2	Andrei Medvedev	Bernardo Mota	6-1 6-3 4-6 6-2	5 sets	No	34
2000-D081	Davis Cup G1 R1: UKR vs POR	No Mundial 1	R1	Carpet	Davis Cup	3	Orest Tereshchuk	Joao Cunha Silva	6-3 6-4	3 sets	No	19
2000-D081	Davis Cup G1 R1: UKR vs POR	No Mundial 1	R1	Carpet	Davis Cup	4	Bernardo Mota	Andrey Dernovskiy	7-5 3-6 7-6(5)	3 sets	No	34
2000-D073	Davis Cup WG R1: GER vs NED	Mundial 1	R1	Carpet	Davis Cup	1	Tommy Haas	John Van Lottum	4-6 7-6(4) 6-3 6-2	5 sets	No	40
2000-D073	Davis Cup WG R1: GER vs NED	Mundial 1	R1	Carpet	Davis Cup	2	Sjeng Schalken	Rainer Schuettler	3-6 7-6(2) 6-1 6-0	5 sets	No	35
2000-D073	Davis Cup WG R1: GER vs NED	Mundial 1	R1	Carpet	Davis Cup	3	Tommy Haas	Sjeng Schalken	6-2 6-2 6-3	5 sets	No	25
2000-D073	Davis Cup WG R1: GER vs NED	Mundial 1	R1	Carpet	Davis Cup	4	David Prinosil	John Van Lottum	6-3 6-3	3 sets	No	18
2000-D063	Davis Cup WG R1: RUS vs BEL	Mundial 1	R1	Carpet	Davis Cup	1	Yevgeny Kafelnikov	Filip Dewulf	6-7(3) 6-4 7-5 6-2	5 sets	No	43
2009-D020	Davis Cup G1 PO: PER vs CAN	No Mundial 1	PO	Clay	Davis Cup	3	Luis Horna	Frederic Niemeyer	7-6(4) 4-6 6-4 7-5	5 sets	No	45
2009-D020	Davis Cup G1 PO: PER vs CAN	No Mundial 1	PO	Clay	Davis Cup	4	Bruno Agostinelli	Ivan Miranda	7-6(5) 1-6 6-3 6-4	5 sets	No	39
2009-D069	Davis Cup G2 PO: ALG vs POR	No Mundial 1	PO	Clay	Davis Cup	1	Rui Machado	Valentin Rahmine	6-0 6-0 6-0	5 sets	No	18
2009-D069	Davis Cup G2 PO: ALG vs POR	No Mundial 1	PO	Clay	Davis Cup	2	Frederico Gil	Abdelhak Hameurlaine	6-1 6-2 6-4	5 sets	No	25
2009-D069	Davis Cup G2 PO: ALG vs POR	No Mundial 1	PO	Clay	Davis Cup	3	Leonardo Tavares	Valentin Rahmine	6-1 6-0	3 sets	No	13
2009-D069	Davis Cup G2 PO: ALG vs POR	No Mundial 1	PO	Clay	Davis Cup	4	Joao Sousa	Sid Ali Akkal	6-3 6-0	3 sets	No	15
2009-D068	Davis Cup G2 PO: DEN vs MNE	No Mundial 1	PO	Clay	Davis Cup	1	Goran Tosic	Frederik Nielsen	5-7 6-4 7-6(5) 6-4	5 sets	No	45
2009-D068	Davis Cup G2 PO: DEN vs MNE	No Mundial 1	PO	Clay	Davis Cup	2	Martin Pedersen	Daniel Danilovic	7-6(5) 6-4 6-1	5 sets	No	30
2009-D068	Davis Cup G2 PO: DEN vs MNE	No Mundial 1	PO	Clay	Davis Cup	3	Frederik Nielsen	Daniel Danilovic	3-6 6-1 6-2 6-4	5 sets	No	34
2009-D068	Davis Cup G2 PO: DEN vs MNE	No Mundial 1	PO	Clay	Davis Cup	4	Martin Pedersen	Goran Tosic	6-4 6-3 6-2	5 sets	No	27
2001-D010	Davis Cup G2 PO: EGY vs GEO	No Mundial 1	PO	Clay	Davis Cup	1	Karim Maamoun	George Khrikadze	6-0 5-0 RET	5 sets	Yes	11
2009-D066	Davis Cup G2 PO: EGY vs GEO	No Mundial 1	PO	Clay	Davis Cup	2	Sherif Sabry	Nodar Itonishvili	6-2 6-4 6-0	5 sets	No	24
2009-D066	Davis Cup G2 PO: EGY vs GEO	No Mundial 1	PO	Clay	Davis Cup	3	Mohamed Mamoun	Nodar Itonishvili	6-1 6-3	3 sets	No	16
2009-D066	Davis Cup G2 PO: EGY vs GEO	No Mundial 1	PO	Clay	Davis Cup	4	Omar Hedayet	Alexander Tavkhelidze	1-6 6-2 6-2	3 sets	No	23

Quadre dades Lesio + No Lesio | Hoja1 | +

Injury Data Surveillance



Lauren Fortington - Australian Centre for Research into Injury in Sports and its Prevention

115 @CasalsTMarti

115

Glossary and C Speak in Sports Analytics

- **Ratios and some metrics** have blind spots (e.g: odds, Home runs, Batting average, corsi, expected goals (xG), expected threat (xT), Valuing On-the-Ball Actions (VAEP) ...)
- **"Larking" factors** (e.g: Team strength, playing at home, rest, body clock, travel, Injuries, Pace, Weather, Arena/Playing Surface, Referees,..)
- **Myths or reality?** (e.g: Momentum, Hot hand, Team chemistry , Serve First, New Balls, Winning mood, Big points,...)
- **Statistical models** (e.g: Markov-switching models and Bayesian Learning, ordered Bradley–Terry model, ELO-type rating models , Null's nested Dirichlet, Bivariate Poisson models,...)

Michael J. Lopez, Gregory J. Matthews, and Benjamin S. Baumer. "How often does the best team win? A unified approach to understanding randomness in North American sport," Annals of Applied Statistics, 2018

116 @CasalsTMarti

116

Different metrics
Measuring *Clutch* Play in the NBA, *hot hand* or *momentum*

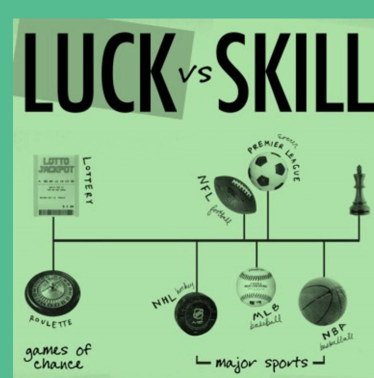
"When a measure becomes a target, it ceases to be a good measure"
Goodhart's Law

117 @CasalsTMarti

117

7. Practical Tips for Real-World Applications

"Researchers are humans and errors and biases are normal. Researchers in Sport Science should provide objective appreciation of questions/uncertainties and an unbiased attempt to address these." Alan McCall 2021

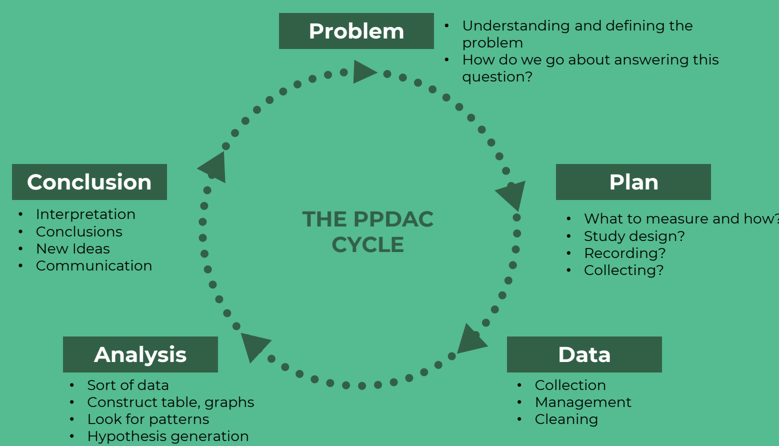


118 @CasalsTMarti

118

The scientific method of the industry

"Application of PPDAC cycle"



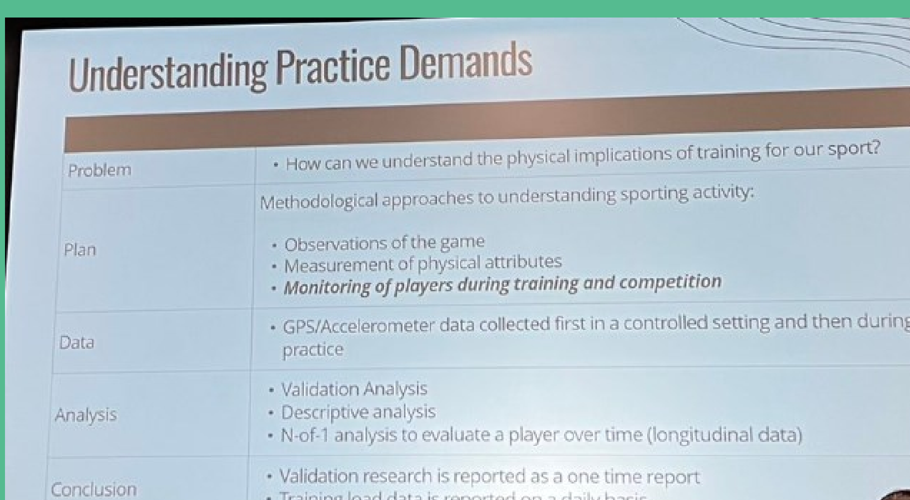
Source: Wild, C. J., & Pfannkuch, M. (1999). Adapted by David Spiegelhalter

119 @CasalsTMarti

119

Vols fer bones preguntes? **PICO**

Example of PPDAC in Sports Sciences



Understanding Practice Demands

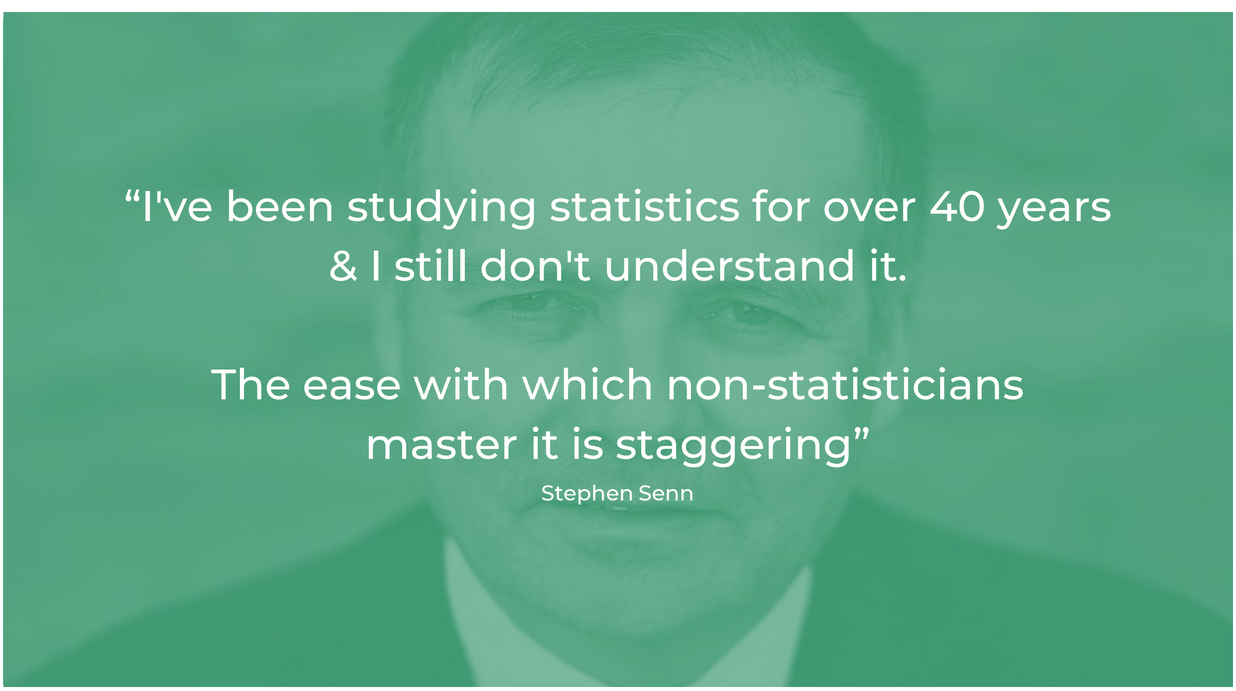
Problem	• How can we understand the physical implications of training for our sport?
Plan	Methodological approaches to understanding sporting activity: • Observations of the game • Measurement of physical attributes • ***Monitoring of players during training and competition***
Data	• GPS/Accelerometer data collected first in a controlled setting and then during practice
Analysis	• Validation Analysis • Descriptive analysis • N-of-1 analysis to evaluate a player over time (longitudinal data)
Conclusion	• Validation research is reported as a one time report • Training load data is reported on a daily basis

120 @CasalsTMarti

120


We need less research,
better research,
and research done
for the right reasons.
Doug Altman

128


“I've been studying statistics for over 40 years
& I still don't understand it.
The ease with which non-statisticians
master it is staggering”
Stephen Senn

129

Move physically
and mentally.
Sports statistics
does the same.

130 @CasalsTMarti

130

Martí Casals
STATSTHINKING

Follow me on: Linkedin &Twitter

@CasalsTMarti

131