



7-8-9
OTTOBRE 2016
TREVISO



LA STATISTICA
NEL CAMPO
di BASKET

EVENTO ORGANIZZATO DAL

 **BODAI – LAB**

Big&Open DATA INNOVATION LABORATORY

UNIVERSITÀ degli STUDI di BRESCIA



Big Data Analytics in Sports

DATA SCIENCE

University of Brescia, Pavia,
Naples Federico II
KPA Ltd., Raanana, Israel
Vrije Universiteit Amsterdam
Plekhanov University, Russia
Financial University of Moscow

DATA MANAGEMENT

SPORT MANAGEMENT, MARKETING AND FINANCE

StatBasket, Italy
University of Brescia
Plekhanov University, Russia



Big DATA Analytics in Sports

DATA SCIENCE

University of Brescia
University of Naples Federico II
KPA Ltd., Raanana, Israel
Vrije Universiteit Amsterdam
Plekhanov University, Russia
Financial University of Moscow

DATA MANAGEMENT

SPORT MANAGEMENT, MARKETING AND FINANCE

StatBasket, Italy
University of Brescia
Plekhanov University, Russia

Gli ambiti di competenza dei data scientist coprono un'ampia gamma di metodi quantitativi nel campo dello **sviluppo di modelli statistici, analisi multivariata, data mining, modellizzazione algoritmica, machine learning e intelligenza artificiale.**

TEMATICHE PRINCIPALI:

- Statistiche di base e analisi più complesse di una partita o un campionato
- Analisi della performance (di squadre, giocatori, atleti)
- Identificazione dei fattori di successo e delle strategie di gioco ottimali
- Previsione
- Misurazione di tratti psicologici (mental toughness, coping strategies, ...)
- Analisi di mercato per il marketing dello sport
- Valutazione finanziaria delle società sportive e dei progetti collegati allo sport

PARTECIPANTI

Gianluca Berti (DS Pallacanestro Cantù)

Guido Corti (**BODAI-Lab**, CEO StatBasket, Statistico certificato FIBA)

Andrea Gracis (DS Treviso Basket, ex cestista)

Marica Manisera (**BODAI-Lab**, Professore di Statistica, Università di Brescia)

Marco Mian (DT Benetton Basket, coach, ex cestista)

Paolo Pressacco (Responsabile del settore giovanile Benetton Basket)

Paola Zuccolotto (**BODAI-Lab**, Professore di Statistica, Università di Brescia)

AGENDA

L'ORIGINE DELLA STATISTICA: i dati

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

ANALYTICS IN SPORTS

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

L'ORIGINE DELLA STATISTICA: i DATI

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

13/06/2016 - 20:45, Impianto: PalaBigi



Grissin Bon Reggio
Emilia

70 - 74

EA7 Emporio
Armani Milano

OLIMPIA MILANO
1936

EA7 Emporio Armani Milano

EA7 Emporio Armani Milano		Falli			Tiri da 2			Sc	Tiri da 3			Tiri Liberi			Rimbalzi			Stoppate		Palle		Ass	Valutaz.		+/-
All.	Pt	Min	C	S	R	T	%		R	T	%	R	T	%	Off	Dif	Tot	Dat	Sub	Per	Rec		Lega	OER	
1 McLean Jamel	0	13	0	1	0	1	0.0	0	0	0.0	0	2	0.0	0	0	0	1	0	1	0	0	-2	0.00	4	
3 Lafayette Oliver	3	25	1	3	0	2	0.0	0	1	100.0	0	0	0.0	0	4	4	1	0	1	1	1	9	0.75	5	
• 5 Gentile Alessandro	16	31	2	2	7	15	46.7	1	0	2	2	2	100.0	0	4	4	0	1	0	1	2	12	0.89	-8	
7 Cerella Bruno	0	4	2	0	0	2	0.0	0	0	0.0	0	0	0.0	2	0	2	0	0	0	0	0	-2	0.00	4	
• 9 Kalnietis Mantas	7	25	4	2	1	4	25.0	0	1	2	2	2	100.0	0	2	2	0	1	3	1	1	1	0.70	-1	
13 Macvan Milan	7	15	3	2	2	2	100.0	0	1	1	100.0	0	0	0.0	0	2	2	1	0	3	1	2	9	117	10
14 Pecchia Andrea	0	0	0	0	0	0	0.0	0	0	0.0	0	0	0.0	0	0	0	0	0	0	0	0	0	0.00	0	
15 Magro Daniele	0	0	0	0	0	0	0.0	0	0	0.0	0	0	0.0	0	0	0	0	0	0	0	0	0	0.00	0	
• 20 Cinciarini Andrea	3	15	4	5	1	3	33.3	0	0	2	1	4	25.0	2	2	4	1	0	0	1	2	5	0.43	4	
• 21 Sanders Rakim	13	28	1	3	2	4	50.0	1	2	4	3	4	75.0	1	1	2	0	0	0	1	5	18	1.30	2	
43 Simon Krunoslav	15	22	1	5	2	3	66.7	0	2	6	5	8	62.5	1	6	7	1	1	2	0	4	20	1.00	2	
• 51 Batista Esteban	10	22	2	4	5	7	71.4	0	0	0	0	0	0.0	4	5	9	0	1	0	1	4	23	1.43	-2	
Squadra	0	0	0	0	0	0	0.0	0	0	0	0	0	0.0	2	2	4	0	0	1	0	0	3	0.00		
Totale	74	200	20	27	20	43	46.5	2	7	18	13	22	59.1	12	28	40	5	4	11	7	21	96	0.89		

L'ORIGINE DELLA STATISTICA: i DATI

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

Grissin Bon Reggio Emilia	Min.	EA7 Emporio Armani Milano
	1. min	
Polonara Achille - Canestro da fuori De Nicolao Andrea - Assist (1)	2-0	Batista Esteban - Tiro sbagliato da sotto Cinciarini Andrea - Rimbalzo offensivo (1)
	2-2	Batista Esteban - Canestro da sotto Sanders Rakim - Assist (1) Cinciarini Andrea - Fallo commesso (1)
Della Valle Amedeo - Fallo subito (1) Veremeenko Vladimir - Canestro da sotto De Nicolao Andrea - Assist (2)	4-2	
	2. min	
	4-4	Kalnietis Mantas - Canestro da sotto Kalnietis Mantas - Fallo commesso (1)
Kaukenas Rimantas - Fallo subito (1) De Nicolao Andrea - Canestro da 3 punti Veremeenko Vladimir - Assist (1)	7-4	
De Nicolao Andrea - Palla recuperata (1) De Nicolao Andrea - Fallo subito (1)		Kalnietis Mantas - Palla persa (1)
De Nicolao Andrea - Fallo commesso (1) De Nicolao Andrea - Palla persa (1)		Batista Esteban - Fallo commesso (1)

L'ORIGINE DELLA STATISTICA: i DATI

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

Tempo	Moncada Energy Group AG	Punti	Margine	Manital Torino
00:00	Fine partita	76-65	11	
00:45	Alessandro Piazza, esce	74-63	11	
00:45	Andrea Portannese, entra	74-63	11	
00:45	Andrea Saccaggi, esce	74-63	11	
00:45	Federico Vai, entra	74-63	11	
00:45	Andrea Portannese, tiro da due punti realizzato	76-63	13	
00:45		76-63	13	Tommaso Pichi, tiro da tre punti sbagliato
00:45		76-63	13	Tommaso Fantoni, rimbalzo offensivo
00:45	Quirino De Laurentiis, fallo	76-63	13	
00:45		76-63	13	Tommaso Fantoni, fallo subito
00:45	Albano Maximo Chiarastella, esce	76-63	13	
00:45	Quirino De Laurentiis, esce	76-63	13	
00:45	Marco Evangelisti, entra	76-63	13	

L'ORIGINE DELLA STATISTICA: i DATI

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

Finale [9]
sabato 9 luglio 2016
21:00 (GMT+2)
Città, Arena:
Turin, PalaIsozaki

CROAZIA ITALIA

84 78

Q1 Q2 Q3 Q4 OT1

PLAY BY PLAY

10:00
#15 M. BILAN
jump ball won

+3
9:48
3-0 #7 K. SIMON
3pt jump shot made

9:29
#3 M. BELINELLI
3pt jump shot from center missed

9:29
#15 M. BILAN
defensive rebound

9:14
#7 K. SIMON
3pt jump shot missed

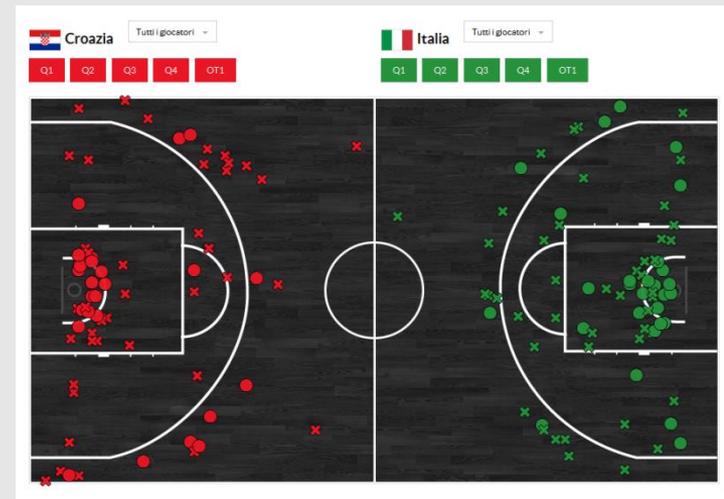
9:14
#8 D. GALLINARI
defensive rebound

+2
8:57
3-2 #5 A. GENTILE
layup made

8:41
#8 D. GALLINARI
personal foul

Italia

#	Players	MIN	P	FG	2PTS	3PTS	TL	OREB	DREB	REB	AST	F	PP	STL	BLK	+/-	EF
2	G. POETA																
Did not play																	
3	*M. BELINELLI	38:25	18	4/15 26.7%	2/10 20%	2/5 40%	8/9 88.9%	0	3	3	2	5	4	2	0	-5	9
4	P. ARADORI	20:19	3	1/6 16.7%	0/1 0%	1/5 20%	0/0 -%	0	5	5	1	0	0	0	0	-5	4
5	*A. GENTILE	28:20	6	3/8 37.5%	3/7 42.9%	0/1 0%	0/0 -%	0	1	1	4	1	5	0	0	-20	1
7	*A. BARGNANI	15:28	4	2/7 28.6%	2/7 28.6%	0/0 0%	0/0 -%	1	2	3	0	3	1	0	0	0	1
8	*D. GALLINARI	25:11	12	5/13 38.5%	5/10 50%	0/3 0%	2/2 100%	1	7	8	4	5	0	1	0	11	17
9	N. MELLI	25:01	11	5/8 62.5%	5/7 71.4%	0/1 0%	1/1 100%	5	3	8	0	5	1	0	2	-5	17
12	M. CUSIN	12:20	2	0/1 0%	0/1 0%	0/0 -%	2/2 100%	0	1	1	1	2	1	0	3	-8	5
13	L. DATOME(C)	27:59	12	5/9 55.6%	4/6 66.7%	1/3 33.3%	1/1 100%	1	2	3	1	5	0	0	2	-3	14
14	R. CERVI																
Did not play																	
23	*D. HACKETT	30:51	10	3/6 50%	2/3 66.7%	1/3 33.3%	3/4 75%	1	5	6	4	4	0	2	0	6	18
27	S. TONUT	01:06	0	0/0 -%	0/0 -%	0/0 -%	0/0 -%	0	0	0	0	0	0	0	0	-1	0
Squadra/Allenatori																	
TOTALI		225	78	28/73 38.4%	23/52 44.2%	5/21 23.8%	17/19 89.5%	10	29	39	17	30	12	5	7	-6	86



L'ORIGINE DELLA STATISTICA: i DATI

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

PLAY BY PLAY

Q1	Q2	Q3	Q4	OT1
10:00	#15 M. BILAN jump ball won			
9:46	+3 3-0 #7 K. SIMON 3pt jump shot made			
9:29	#3 M. BELINELLI 3pt jump shot from center missed			
9:25	#15 M. BILAN defensive rebound			
9:14	#7 K. SIMON			



EA7 EMPORIO ARMANI MILAN	88
ANADOLU EFES ISTANBUL	84

Report Video Boxscore **Play by play** Graphicstats Shooting chart

1st Quarter	2nd Quarter	3rd Quarter
EA7 Emporio Armani Milan		
Min		Score
	Begin Period	0-0
09:41	McLEAN, JAMEL Foul Drawn (1)	0-0
09:41	McLEAN, JAMEL Free Throw In (1/1 - 1 pt)	1-0
09:41	McLEAN, JAMEL Free Throw In (2/2 - 2 pt)	2-0
09:41		2-2
09:41		2-2
08:59	BARAC, STANKO Missed Two Pointer (0/1 - 0 pt)	2-2
08:56		2-2
08:50		2-2



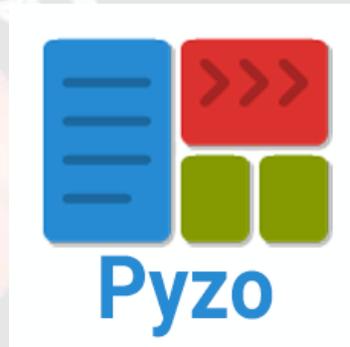
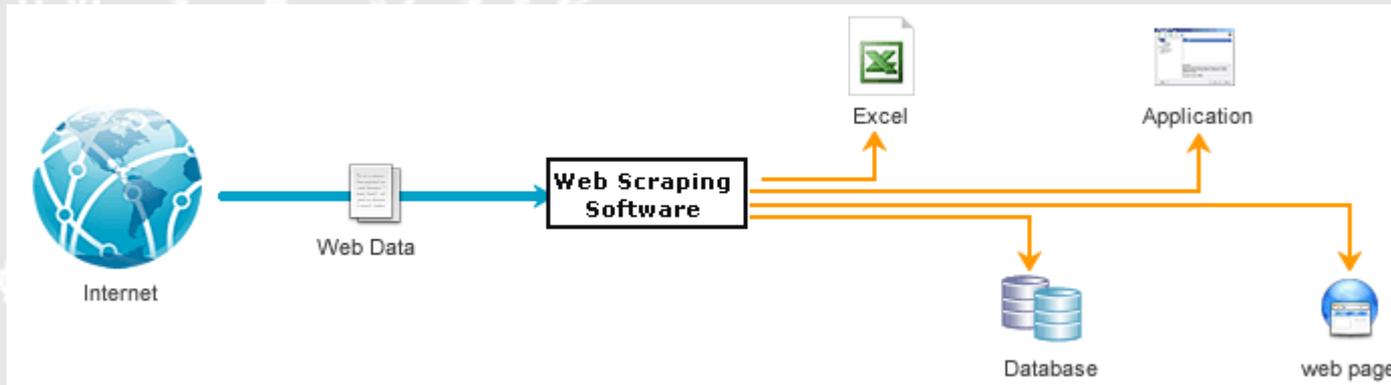
09:34	Curry 3pt Shot Missed	2-2
09:24	K Thompson Rebound (Off 1 Def 0)	2-2
09:15	Green Driving Dunk Shot: Made (2 PTS) Assist: Ezeli (1 AST)	3-3
09:07	Barnes 3pt Shot: Made (5 PTS) Assist: Barnes (1 AST)	3-3
09:07	Barnes Rebound (Off 0 Def 1)	3-3
09:42		3-3
09:36	Barnes 3pt Shot: Made (3 PTS) Assist: Green (1 AST)	3-3
09:20		3-3
09:19	Curry Rebound (Off 0 Def 2)	3-3
09:06	Green 3pt Shot: Made (5 PTS) Assist: Barnes (1 AST)	3-3
09:06		3-3
08:51		3-3
08:50	Ezeli Rebound (Off 0 Def 1)	3-3
08:43	K Thompson 3pt Shot: Missed	3-3
08:42		3-3
08:32		3-3

EA7 Emporio Armani Milano	
Batista Esteban - Tiro sbagliato da sotto	
Cinciarini Andrea - Rimbalzo offensivo (1)	
Batista Esteban - Canestro da sotto	
Sanders Rakim - Assist (1)	
Cinciarini Andrea - Fallo commesso (1)	
Kalnetis Mantas - Canestro da sotto	

L'ORIGINE DELLA STATISTICA: i dati

- Rilevazione di dati e statistiche ufficiali in Italia
- Le certificazioni FIBA
- Dal punto di vista del Data Scientist: Big&Open Data

WEBSCRAPING



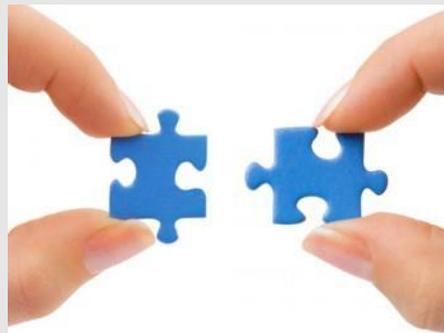
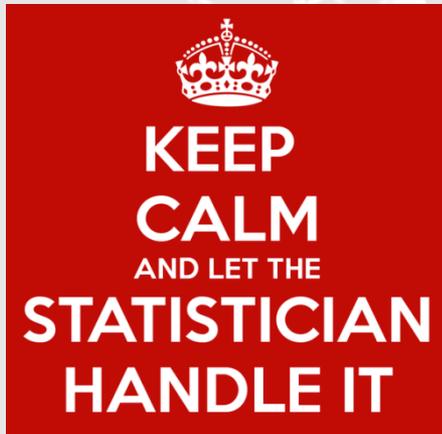
Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale



Statistiche
Dati

Quanti tiri da 3 punti ha tentato il giocatore XY quando il margine rispetto all'avversario era tra -5 e +5 punti?

Quanti canestri?



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale



Discussione



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

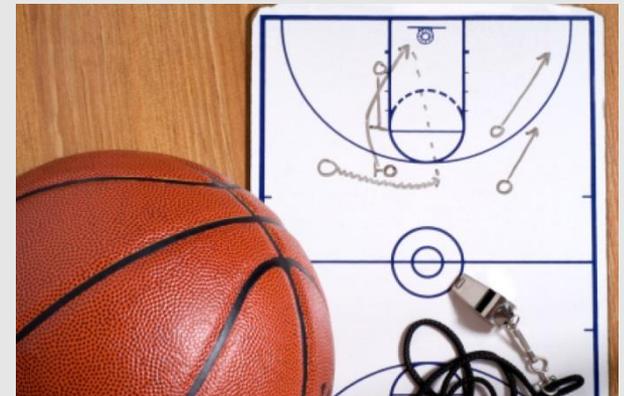


Il giocatore XY come vive le situazioni stressanti durante la partita?

Rinuncia a tirare?

La sua performance migliora o peggiora?

Domande

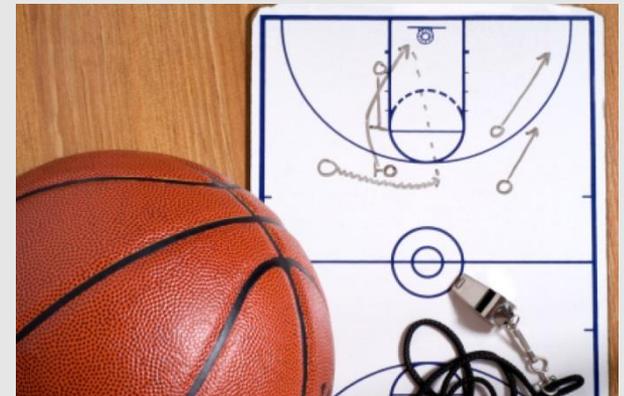


Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale



Data Mining
Informazione



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

STATISTICHE

A2 Italy 2015/16

Margine - IP	Min	2pM	2pA	2p%	2pM40	3pM	3pA	3p%	3pM40	FTM	FTA	FT%	FT40	Pts40	PtsOpp40
A: sotto la forbice	158	81	179	45.25	20.50	26	83	31.33	6.58	79	114	69.30	20.00	80.74	64.29
B: interno alla forbice	408	243	458	53.06	23.81	71	214	33.18	6.96	95	135	70.37	9.31	77.80	68.40
C: sopra la forbice	503	259	507	51.08	20.60	87	244	35.66	6.92	125	186	67.20	9.94	71.91	71.04

Margine - IP	Min	FS	FS40	RO	RO40	TO	TO40
A: sotto la forbice	158	102	25.82	67	16.96	57	14.43
B: interno alla forbice	408	190	18.62	127	12.44	134	13.13
C: sopra la forbice	503	217	17.26	143	11.38	157	12.49

Legenda prima tabella	
Margine nelle tre categorie - Intera Partita	Margine - IP
Tempo di permanenza in minuti nelle tre categorie di margine	Min
Tiri da 2 riusciti	2pM
Tiri da 2 totali	2pA
Tiri da 2 % successo	2p%
Tiri da 2 rapportati al tempo (40 minuti)	2pM40
Tiri da 3 riusciti	3pM
Tiri da 3 totali	3pA
Tiri da 3 % successo	3p%
Tiri da 3 rapportati al tempo (40 minuti)	3pM40
Tiri liberi riusciti	FTM
Tiri liberi totali	FTA
Tiri liberi % successo	FT%
Tiri liberi rapportati al tempo (40 minuti)	FT40
Punti fatti rapportati al tempo (40 minuti)	Pts40
Punti subiti rapportati al tempo (40 minuti)	PtsOpp40

Legenda seconda tabella	
Margine nelle tre categorie - Intera Partita	Margine - IP
Tempo di permanenza in minuti nelle tre categorie di margine	Min
Falli subiti	FS
Falli subiti rapportati al tempo (40 minuti)	FS40
Rimbalzi offensivi	RO
Rimbalzi offensivi rapportati al tempo (40 minuti)	RO40
Palle perse	TO
Palle perse rapportate al tempo (40 minuti)	TO40

FORBICE: tra-4 e 4

ANALYTICS IN SPORTS

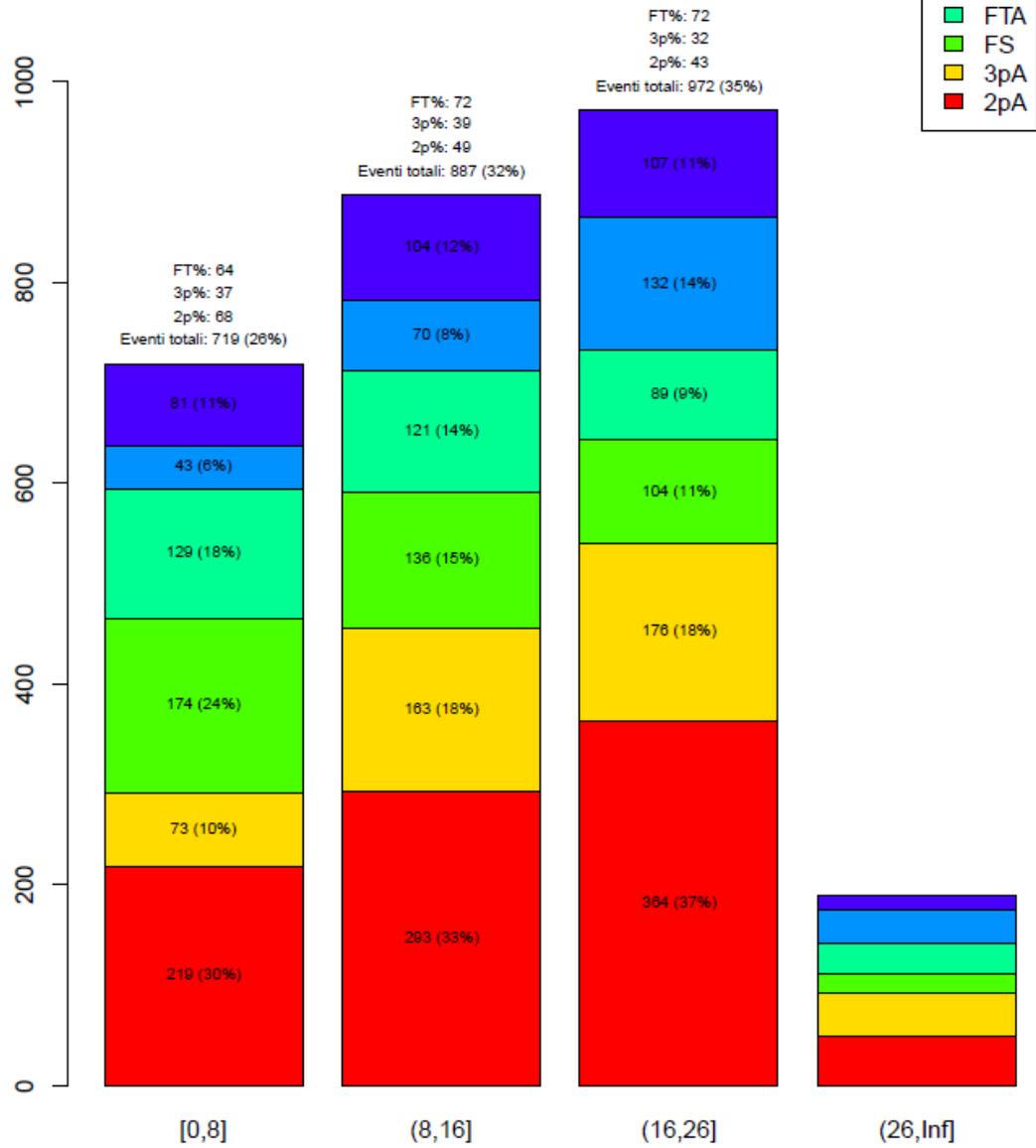
- Analisi s
- Esempi d
- Nuove fr



STATISTICHE

A2 Italy
2015/16

- TO
- RO
- FTA
- FS
- 3pA
- 2pA



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

STATISTICHE

A2 Italy 2015/16



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

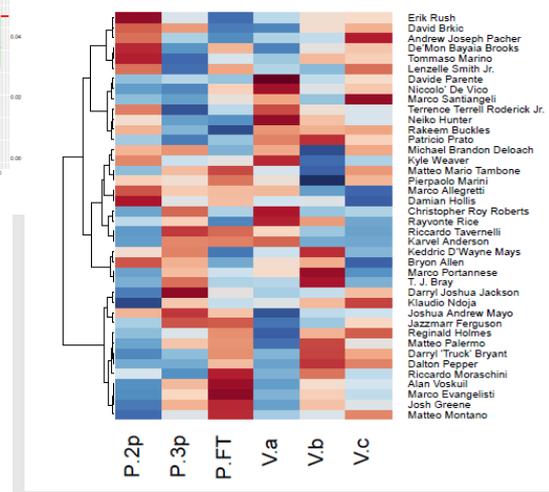
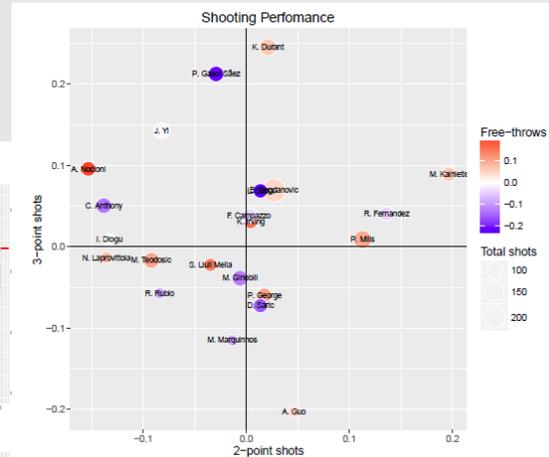
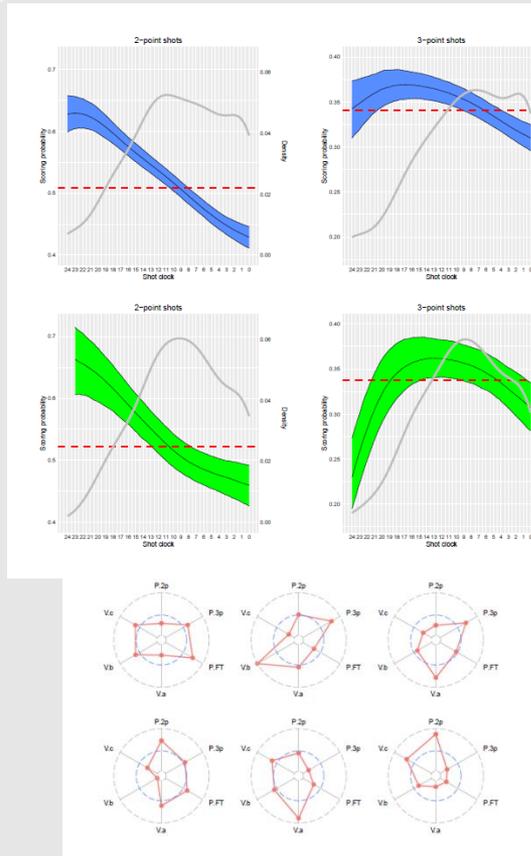
DATA MINING

Big data analytics to model scoring probability in basketball: the effect of shooting under high-pressure conditions

1 Introduction

In the last decades the idea of a statistical thinking in sports has gained a rapidly growing interest, as documented by the wide scientific production on this theme and also by the publication of some insightful collections of statistical analyses applied to data from a wide range of sports, including American football, baseball, basketball, and ice hockey (Albert, Bennett, and Cochran, 2005, Albert and Koning, 2007).

As concerns basketball, several statistical techniques have been applied to analyze data with a great variety of different aims, ranging from simply depicting the main features of a game by means of descriptive statistics (Kubatko, Oliver, Pelton, Rosenbaum et al., 2007) to the investigation of more complex problems, such as forecasting the outcomes of a game or a tournament (West et al., 2008, Loeffelholz, Bednar, Bauer et al., 2009, Brown, Sokol et al., 2010, Gupta, 2015, Lopez and Matthews, 2015, Ruiz and Perez-Cruz, 2015, Yuan, Liu, Yeh, Kaufman, Reece, Bull, Franks, Wang, Illushin, and Bornn, 2015, Manner, 2016), analysing players' performance (Page, Fellingham, and Reese, 2007, Cooper, Ruiz, and Sirvent, 2009, Piette, Anand, and Zhang, 2010, Fearnhead and Taylor, 2011, Ozmen, 2012, Page, Barney, and McGuire, 2013, Deshpande and Jensen, 2016), studying the network of players' pathways from the in-bounds pass to the basket (Skinner, 2010) and their spatial positioning (Shorridge, Goldsberry, and Adams, 2014), or identifying optimal game strategies (Annis et al., 2006).

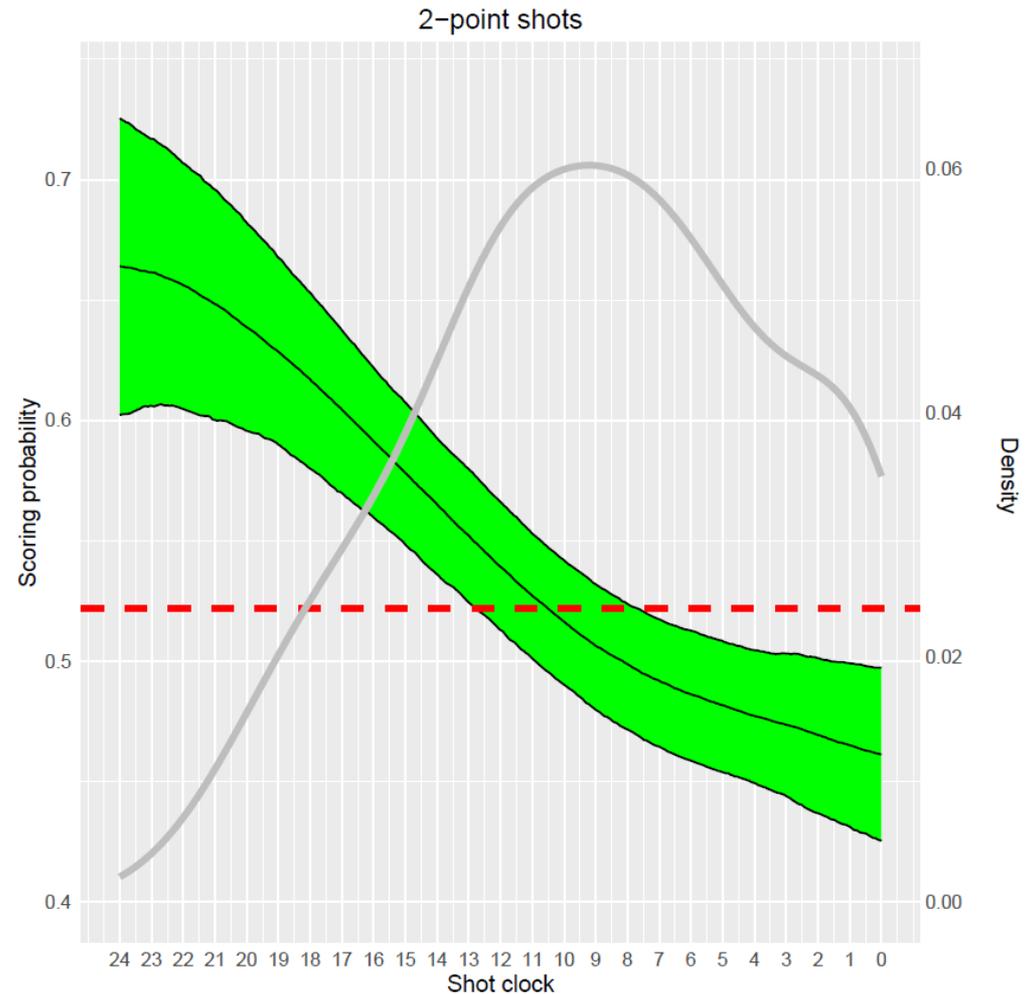


Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

DATA MINING

Rio 2016

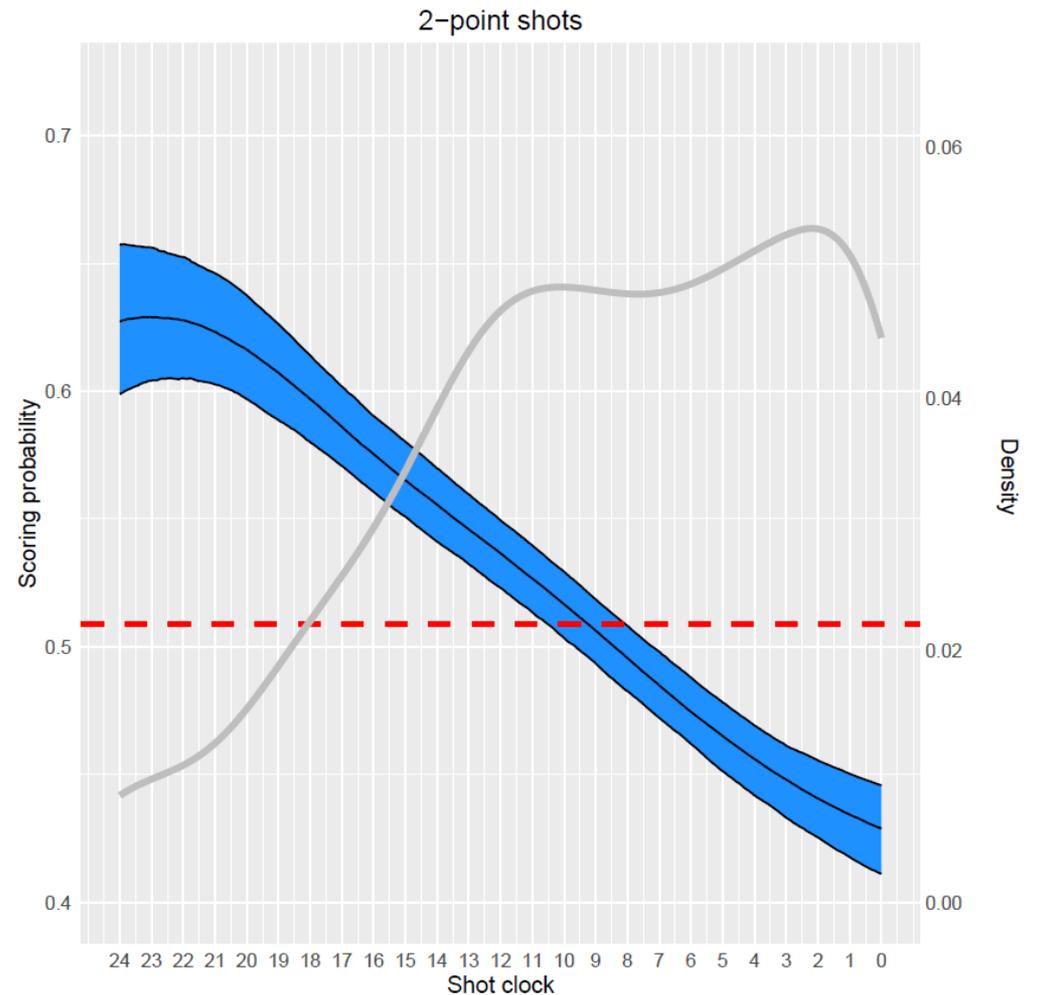


Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

DATA MINING

A2 Italy
2015/16

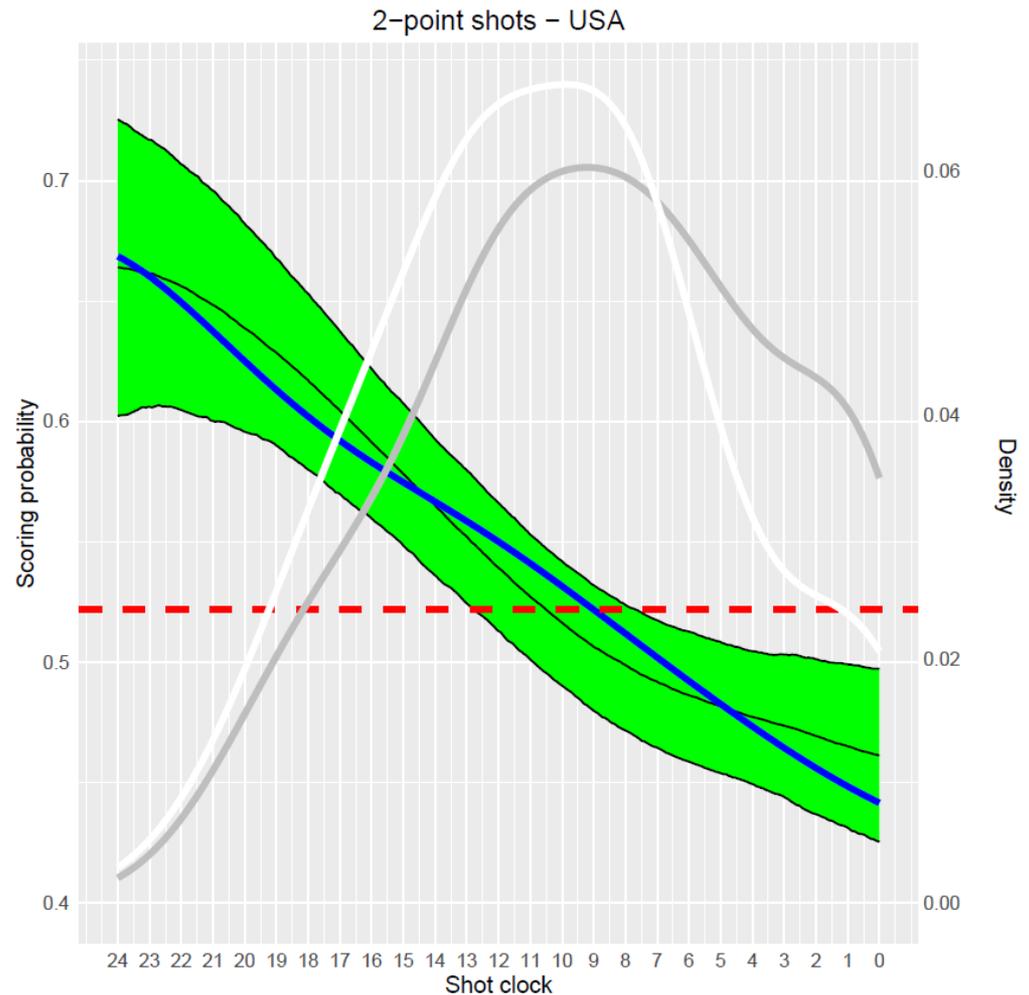


Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

DATA MINING

Rio 2016



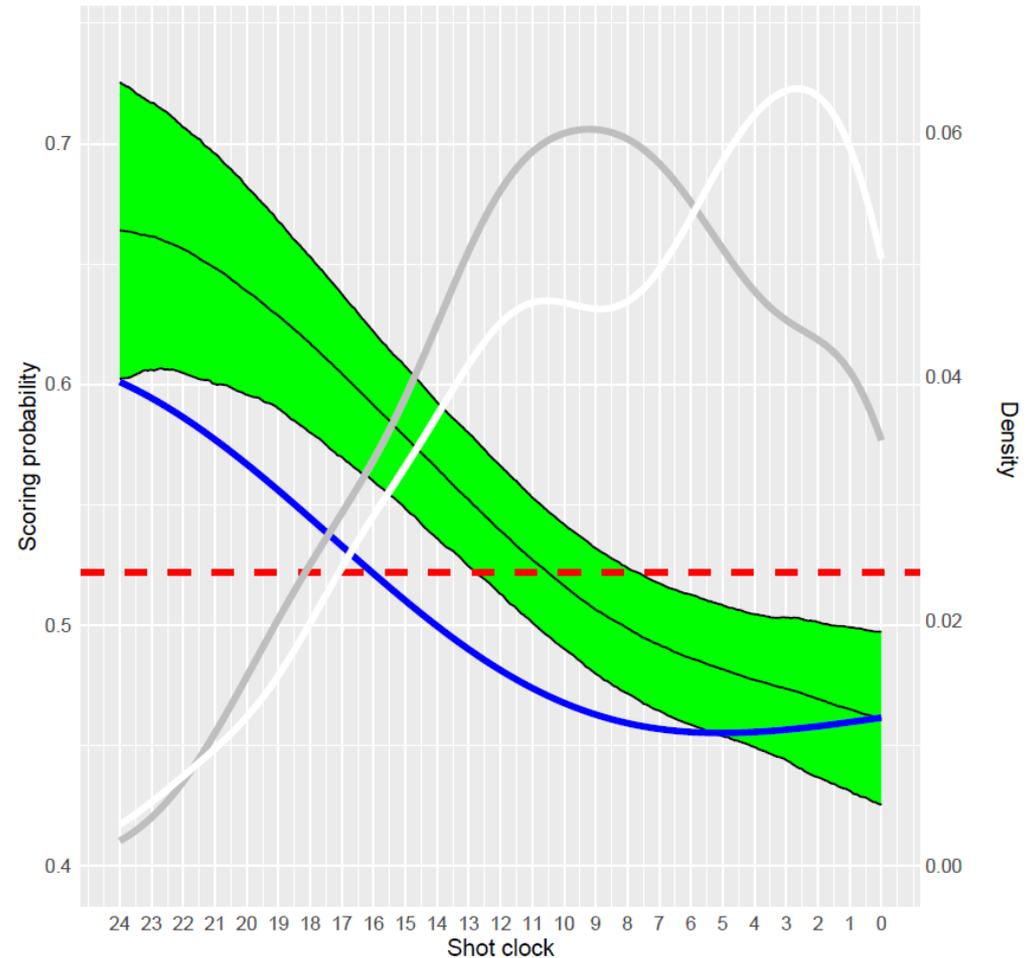
Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

DATA MINING

Rio 2016

2-point shots - CHINA



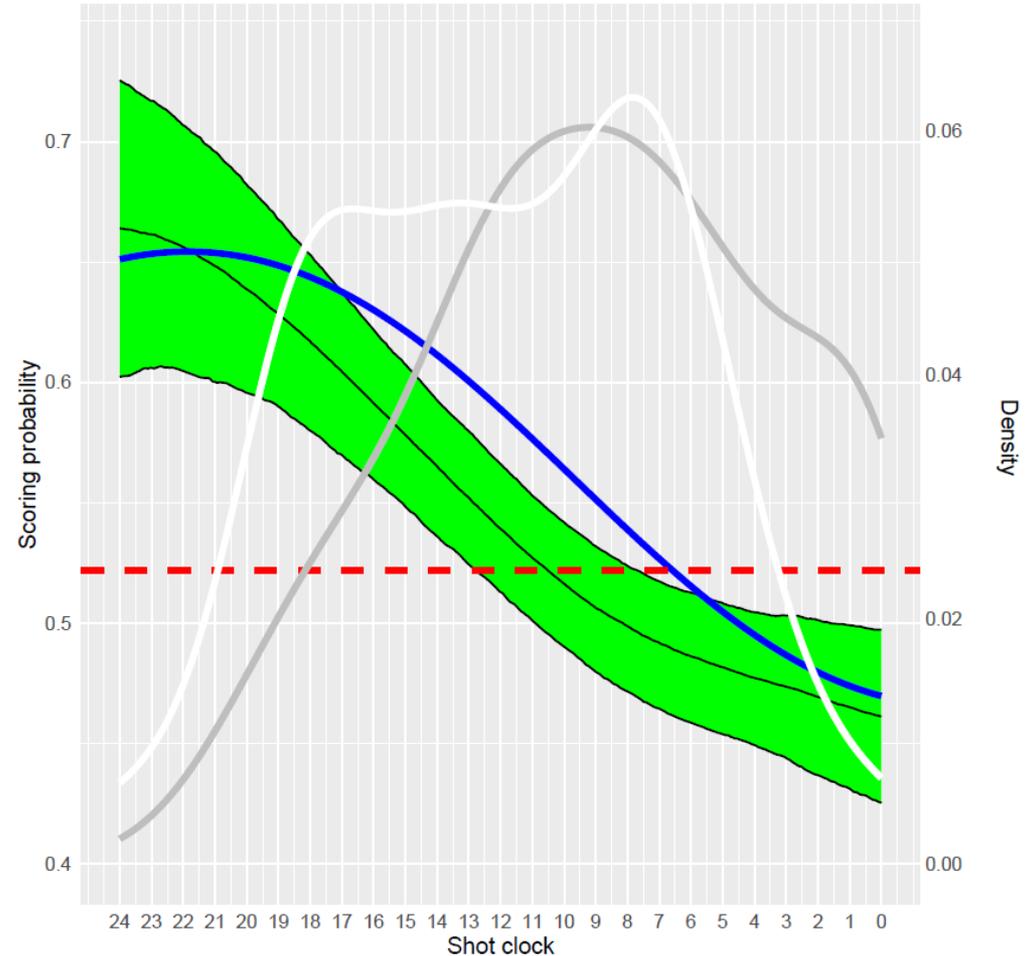
Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

DATA MINING

Rio 2016

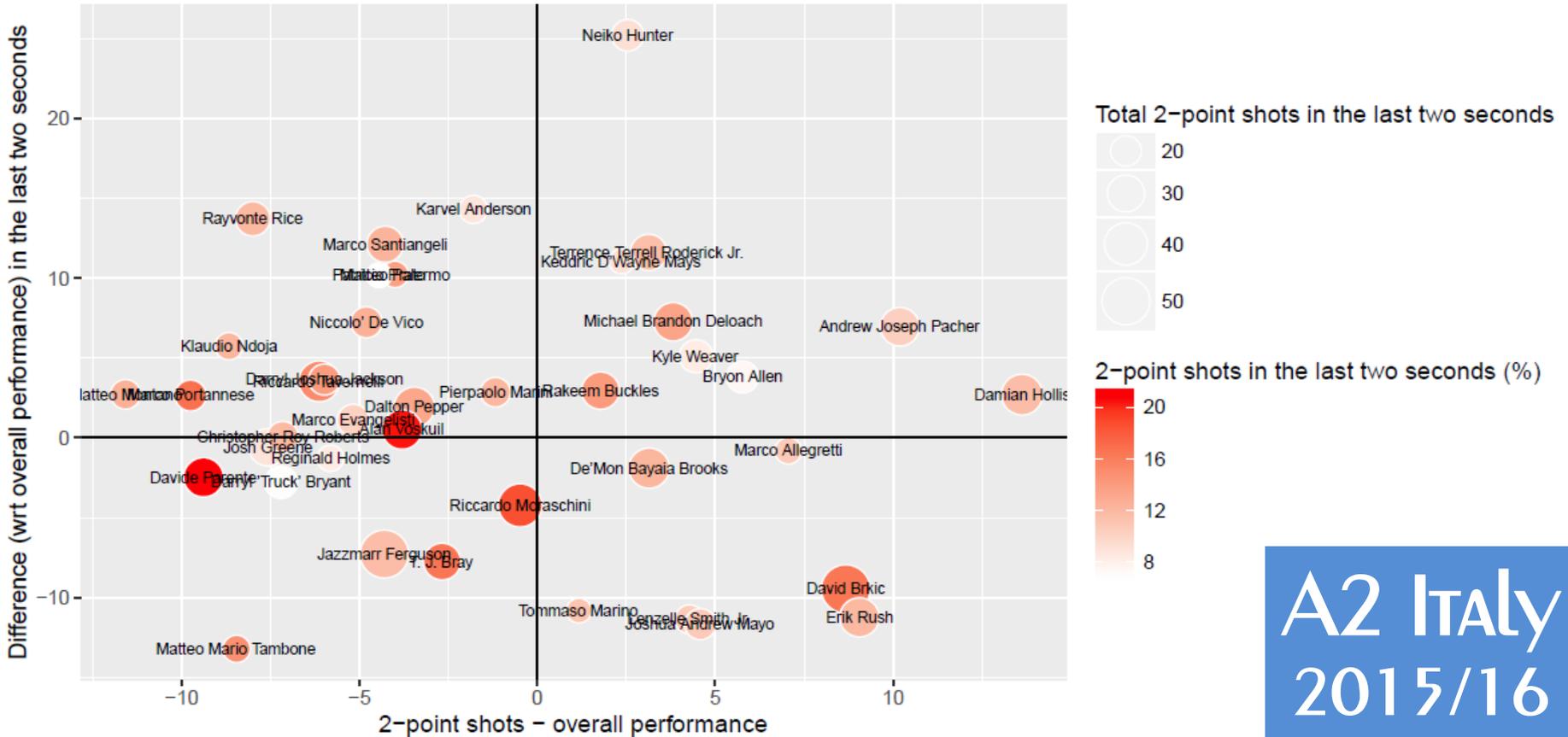
2-point shots - KEVIN DURANT



Analytics in Sports

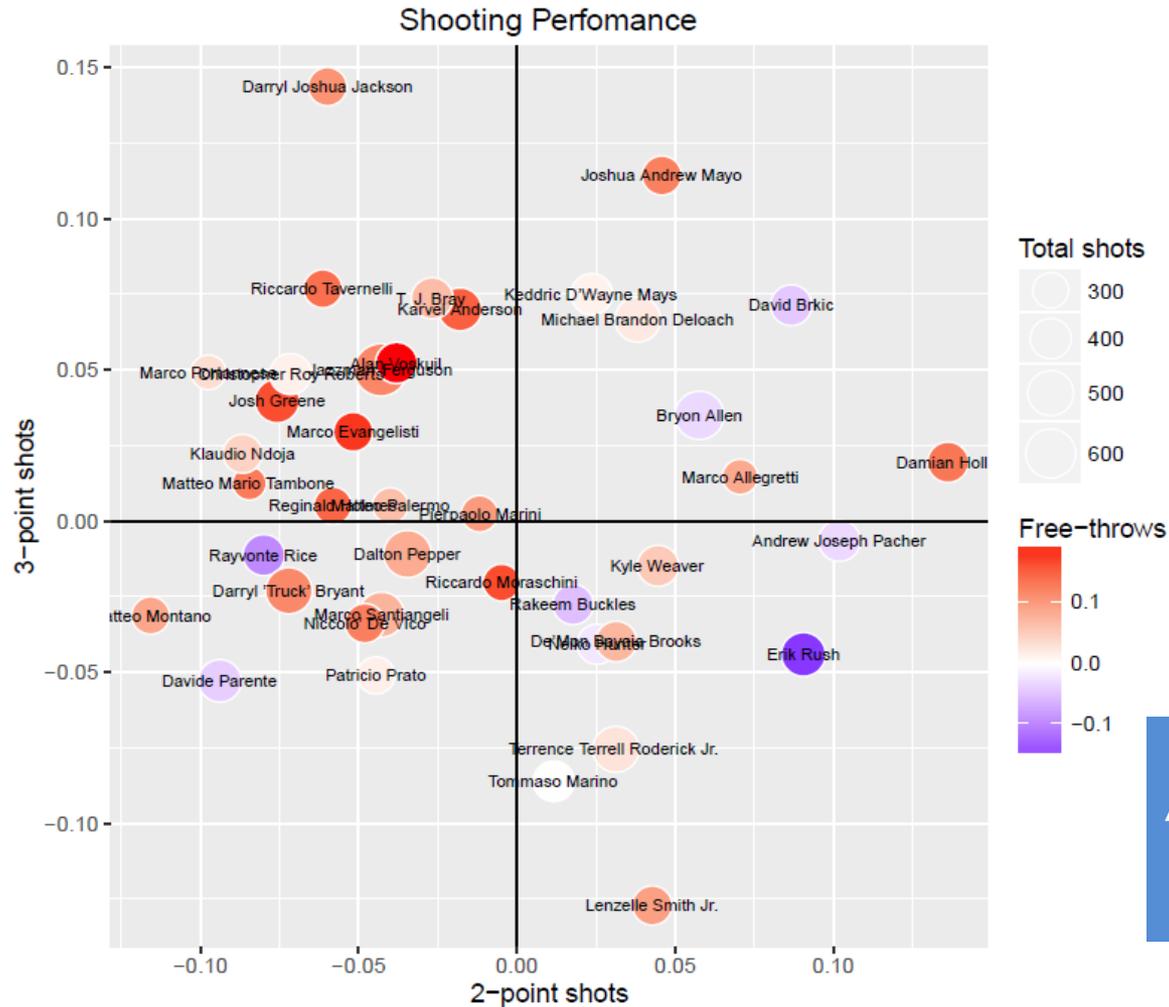
- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale

2-point shots – performance in the last two seconds



Analytics in Sports

- Analisi sul campo e ricerca scientifica: stabilire un contatto tra i due mondi
- Esempi di analytics su dati play-by-play
- Nuove frontiere della tecnologia: sensori e monitoraggio in tempo reale



A2 Italy
2015/16

