



LA NOTTE EUROPEA

DEI RICERCATORI



venerdì **29**

settembre **2017**



Sports Analytics: la Statistica divertente



**Marica Manisera
e Paola Zuccolotto**

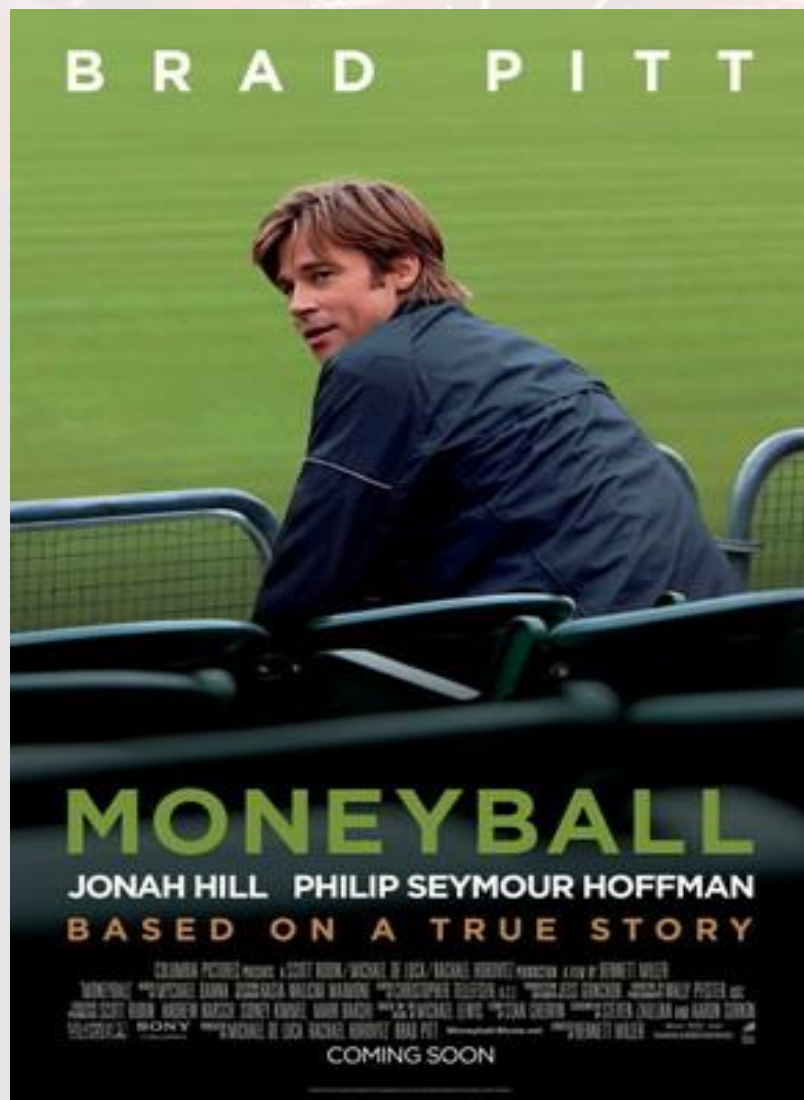
Università degli Studi di Brescia
Big&Open Data Innovation
Laboratory



Introduzione

Importanza della **Statistica nello Sport**

- nascita di aziende e startup
- coinvolgimento di analisti nelle squadre
- conferenze a tema
- pubblicazioni scientifiche, divulgative, su web e social media





BDSports

In questo quadro è nato il progetto internazionale
Big Data Analytics in Sports (BDSports)

nell'ambito delle attività del
Big&Open Data Innovation Laboratory
dell'Università di Brescia
(BODaI-Lab, bodai.unibs.it)





BDSports

Il **Big&Open Data Innovation Laboratory** è un laboratorio di ricerca scientifica con lo «scopo di creare gruppi di lavoro che predispongano - all'interno di specifici progetti - metodi, tecniche e strumenti innovativi per il reperimento, la gestione e l'analisi dei dati con approccio multidisciplinare, favorendo il trasferimento tecnologico verso la PA e il mondo produttivo»



BDSports

Il **Big Data Analytics in Sports** è un progetto internazionale che mira a creare una collaborazione tra esperti interessati alle analisi quantitative nello Sport

<http://bodai.unibs.it/BDSports/>

Le competenze dei partecipanti: modelli statistici, analisi multivariata, data mining, algoritmi di intelligenza artificiale e apprendimento automatico, big data



BDSports

1) Ricerca scientifica

pubblicazione di articoli scientifici e numeri speciali di riviste scientifiche, organizzazione di workshop e conferenze, condivisione di dati e software

2) Applicazione sul campo

stage di studenti universitari, tavole rotonde, contatti con aziende che si occupano di rilevazione dati nello sport

3) Didattica

4) Divulgazione

incontri di orientamento per studenti, pubblicazione articoli divulgativi su riviste di settore o su social media, Notte dei Ricercatori, attività ludiche (StatisticAll)



BDSports

Didattica

Obiettivo:

Insegnare agli studenti che le materie quantitative possono anche essere divertenti

Metodi didattici alternativi

Progetti applicativi anche multidisciplinari



Incontro di oggi



BDSports

Proposta: **3 progetti** da realizzare nell'ambito di un percorso formativo scolastico pensato per la scuola secondaria superiore

Si tratta di proposte «personalizzabili» (+ incontri di formazione e creazione progetti su misura)

Il materiale didattico a supporto degli insegnanti è **disponibile on line** (area protetta da password)

<http://bodai.unibs.it/BDSports/Students.htm>



Progetto 1

ESAME DEI DATI E DELLE STATISTICHE PRESENTI ON LINE

Esplorare siti web e analizzare il contenuto delle pagine dedicate alle statistiche on line

Obiettivo:

Fare scoprire agli studenti il mondo della Statistica nello Sport

Competenze sviluppate:

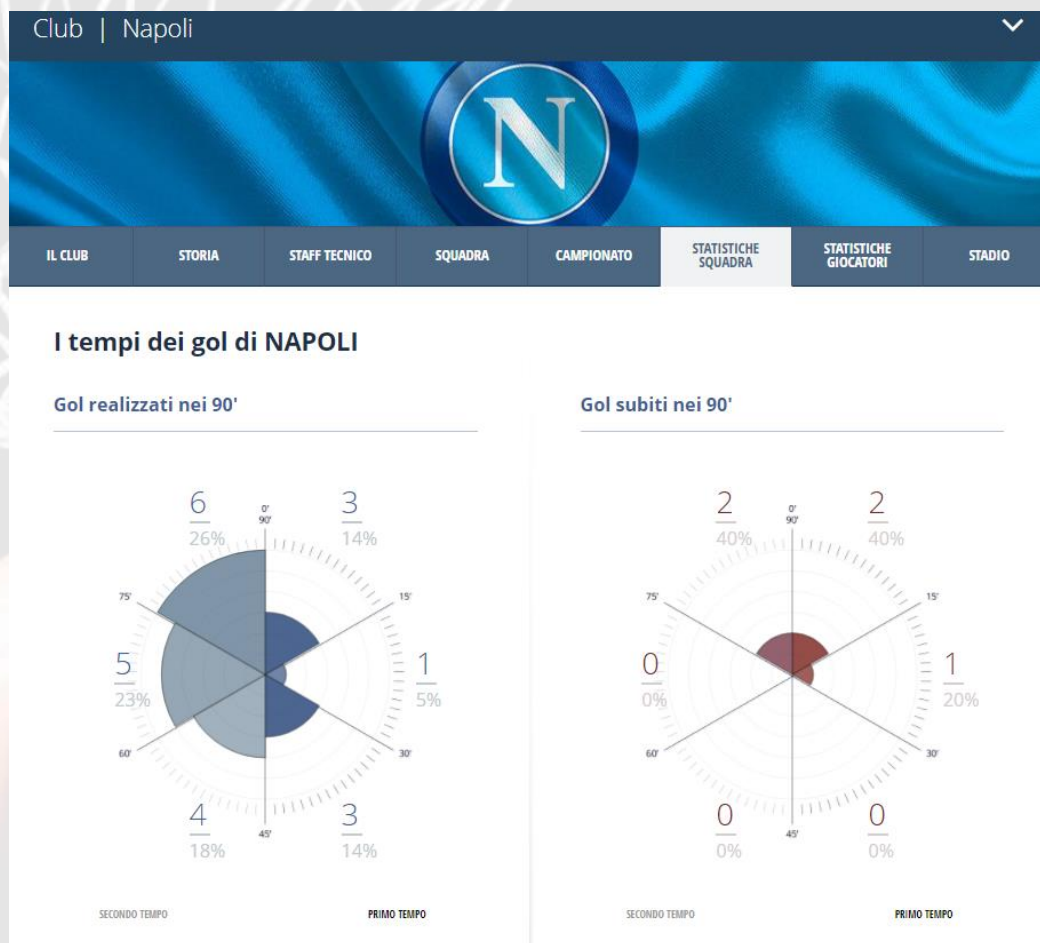
- Ricerca e consultazione web
- Interpretazione dati e grafici statistici



Progetto 1

ESEMPIO 1

<http://www.legaseriea.it/it/serie-a-tim/squadre/napoli/statistiche-squadra>





Progetto 1

ESEMPIO 2

<http://www.legab.it/campionato-serie-b/club/salernitana>

VIDEO PANINI DIGITAL:

<https://www.youtube.com/watch?v=gxEa6B6eW5o>

CAMPIONATO ORGANIGRAMMA STORIA STAFF TECNICO TEAM **STATISTICHE** STADIO BIGLIETTERIA

STAGIONE 2016/2017

Selezione: Squadra ▼

I TEMPI DEL GOAL

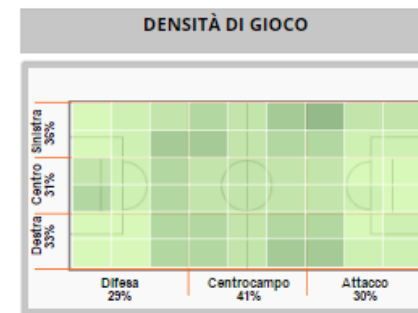


VALORI MEDI DI STAGIONE

	Media
Possesso palla	23:56"
Tiri totali	11.9
Tiri dentro	4.5
Palle giocate	539.1
Passaggi riusciti	60.5
Fall commessi	17.0
Angoli	5.1
Protezione area	54.7
Supremazia territoriale	10:06"
Attacco alla porta	47.2
Pericolosità	42.9
Giocate utili	81.3
Gol Fatti	1.1

POSIZIONE IN CLASSIFICA

POSIZIONE	NOME SQUADRA	PUNTI
10	Ascoli	33
11	Carpi	33
12	Novara	31
13	Salernitana	31
14	Avellino	29
15	Cesena	28
16	Brescia	28

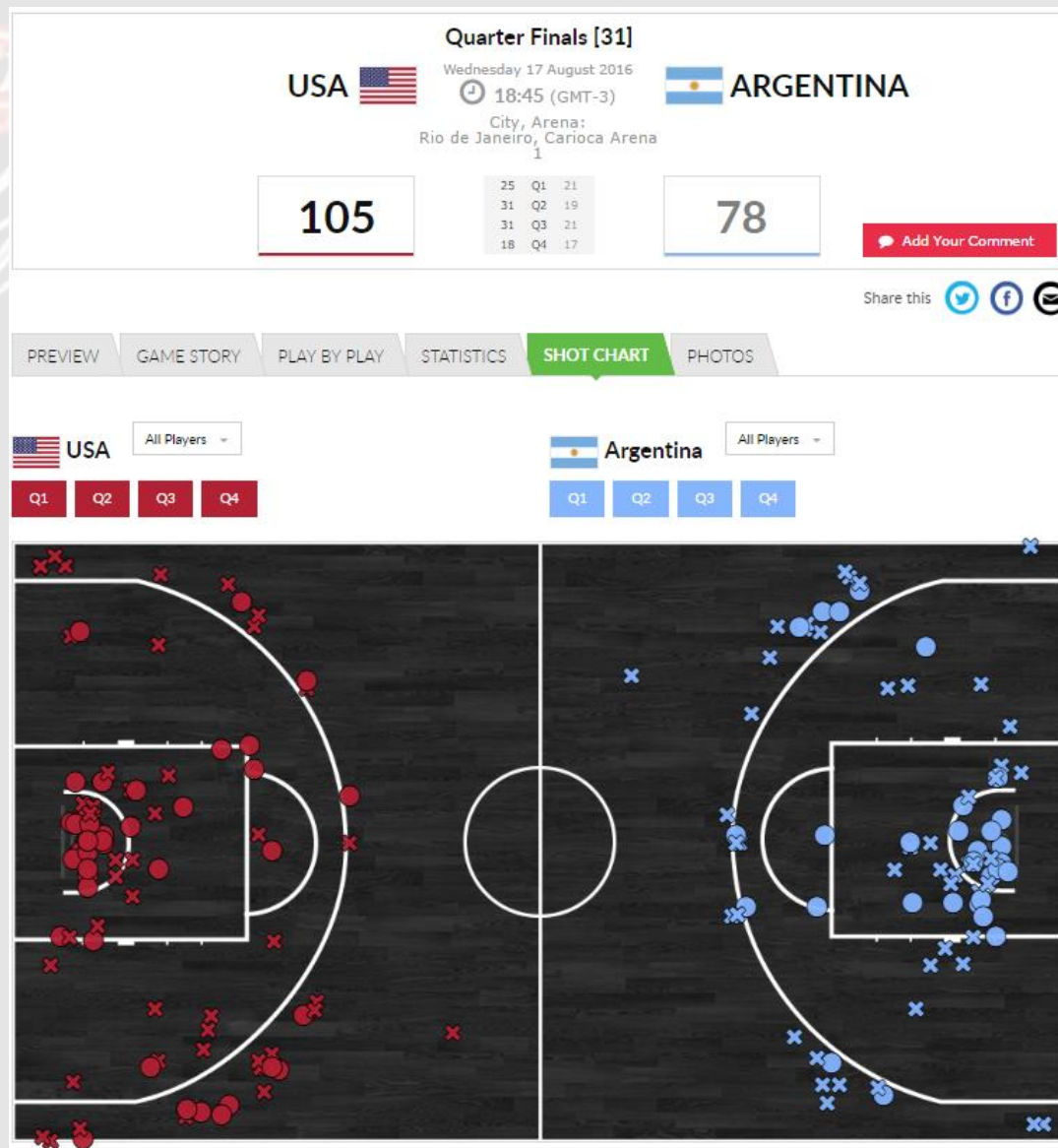




Progetto 1

ESEMPIO 3

http://www.fiba.com/olympics/2016/1708/USA-Argentina#tab=shot_chart





Progetto 1

ESEMPIO 4

http://www.wimbledon.com/en_GB/players/overview/wta310761.html

WIMBLEDON
 3 JULY - 16 JULY 2017

Players
 SEARCH
RADIO
IBM

FIRST ROUND
COURT 14 • LADIES' SINGLES • COMPLETE

S. Errani (ITA) [20]	PTS	1	2	3	4	5
P. Tig (ROU)		6	6			
		4	4			

MATCH DURATION: 1:31

Sara Errani

MATCH STATISTICS

Patricia Maria Tig

0	ACES	3
1	DOUBLE FAULTS	7
56/67 (84 %)	1ST SERVES IN	38/66 (58 %)
34/56 (61 %)	1ST SERVE POINTS WON	27/38 (71 %)
7/11 (64 %)	2ND SERVE POINTS WON	9/28 (32 %)
8/14 (57 %)	NET POINTS WON	11/15 (73 %)
4/9 (44 %)	BREAK POINTS WON	2/10 (20 %)
30/66 (45 %)	RECEIVING POINTS WON	26/67 (39 %)
12	WINNERS	31
14	UNFORCED ERRORS	37
71	TOTAL POINTS WON	62

Sara Errani: First round

[Player Profiles](#)
[Competitor List](#)
[Player Statistics](#)
[Country Statistics](#)
[Player Challenge](#)
[Rolls of Honour](#)
[Archive](#)
[NEWS](#)
[SCORES & SCHEDULE](#)
[DRAWS](#)
[PLAYERS](#)
[VIDEO & RADIO](#)
[VISIT & TICKETS](#)
[ABOUT](#)
[FOUNDATION](#)
[SHOP](#)



Progetto 2

ANALISI DELLE PRESTAZIONI DI SQUADRE E GIOCATORI

Utilizzare strumenti di statistica descrittiva per analizzare le prestazioni di una squadra di pallacanestro e dei suoi giocatori

Obiettivo:

Insegnare agli studenti nozioni non banali di Statistica e la loro implementazione pratica in Excel

Competenze sviluppate:

- Reperimento dati sul web
- Analisi statistiche descrittive di base: teoria ed applicazione con interpretazione dei risultati
- Costruzione matrici di dati, utilizzo di formule e grafici in Excel



Progetto 2

File: 'Progetto2.xlsx' disponibile on line sul sito BDSports

Esempio: Analisi della squadra **Virtus Bologna**

serie A2 2016-2017, dati del girone di andata (dal 2/10/2016 al 6/1/2017)

FASE 1 -

Reperimento dati, descrizione variabili, organizzazione in tabella su Excel

www.legapallacanestro.com/serie-a2/segafredo-virtus-bologna



Progetto 2

Progetto2.xlsx - Microsoft Excel

File Home Inserisci Layout di pagina Formule Dati Revisione Visualizza

Calibri 11 A A

Testo a capo

Generale

Unisci e allinea al centro

Formattazione condizionale

Formatta come tabella

Stili

Normale Neutrale

Valore non v... Valore valido

Somma automatica

Riempimento

Cancella

Ordina e filtra

Modifica

Appunti

Carattere

Allineamento

Numeri

Celle

Modifica

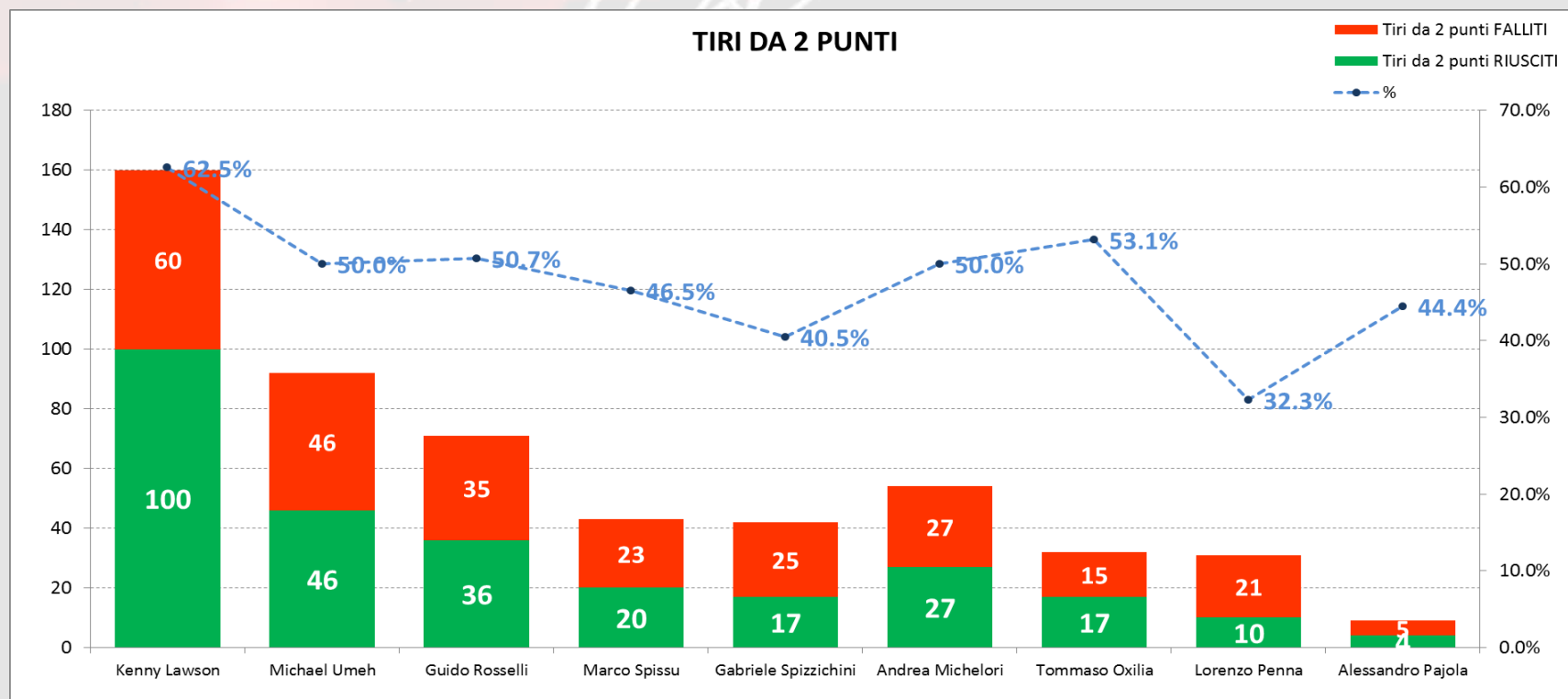
A1	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
	Pun	G	Min	Falli		Tiri da 2		Tiri da 3		Tiri Liberi		Rimbalzi		Stop		Palle		Ass									
				C	S	R	T	%	R	T	%	R	T	%	O	D	T	D	S	P	R						
1																											
2																											
3	Kenny Lawson	287	14	454	30	47	100	160	63	16	40	40	39	53	74	45	82	127	6	5	20	10	17				
4	Michael Umeh	248	14	435	19	27	46	92	50	44	103	43	24	32	75	8	41	49	0	9	23	15	30				
5	Guido Rosselli	155	14	419	36	47	36	71	51	15	39	38	38	44	86	14	61	75	7	4	22	21	59				
6	Marco Spissu	138	14	372	32	25	20	43	47	25	71	35	23	28	82	9	21	30	0	6	17	29	46				
7	Gabriele Spizzichini	62	14	286	32	37	17	42	40	5	26	19	13	23	57	10	30	40	3	4	14	14	34				
8	Andrea Michelori	72	13	232	34	39	27	54	50	0	0	0	18	26	69	25	35	60	2	1	18	8	9				
9	Tommaso Oxilia	53	14	171	20	16	17	32	53	5	17	29	4	11	36	12	21	33	1	2	13	3	9				
10	Lorenzo Penna	50	14	199	21	20	10	31	32	7	30	23	9	13	69	7	14	21	0	3	11	11	16				
11	Alessandro Pajola	35	13	113	21	6	4	9	44	7	11	64	6	10	60	6	8	14	0	1	1	6	6				
12																											
13																											
14	Fonte dati	http://www.legapallacanestro.com/serie-a2/segafredo-virtus-bologna																									
15	Aggiornati al	07-gen-17																									
16																											
17	RACCOMANDAZIONE: IN TUTTI I FOGLI DI LAVORO SI POSSONO MODIFICARE SOLAMENTE LE CELLE CON SFONDO ROSSO E CARATTERI BIANCHI , SEGUENDO LE INDICAZIONI RIPORTATE NEI COMMENTI																										
18	OGNI ALTRA MODIFICA POTREBBE DANNEGGIARE LE ELABORAZIONI																										
19																											
20	Significato delle variabili																										
21		Pun	Punti segnati																								
22		G	Partite giocate																								
23		Min	Minuti giocati																								
24		Falli	C	Commessi																							
25			S	Subiti																							
26		Tiri da 2	R	Riusciti																							
27			T	Tentati																							
28			%	R/T*100																							
29		Tiri da 3	R	Riusciti																							
30			T	Tentati																							
31			%	R/T*100																							
32		Tiri Liberi	R	Riusciti																							
33			T	Tentati																							
34			%	R/T*100																							
35		Rimbalzi	O	Offensivi																							
36			D	Difensivi																							
37			T	Totali																							
38		Stop	D	Difensivi																							
39			S	Subiti																							
40		Palle	P	Perse																							
41			R	Recuperate																							
42		Ass		Assist																							

Dati Analisi variabilità Analisi concentrazione Tiri Radial Plot



Progetto 2

FASE 2 - Analisi delle prestazioni di tiro

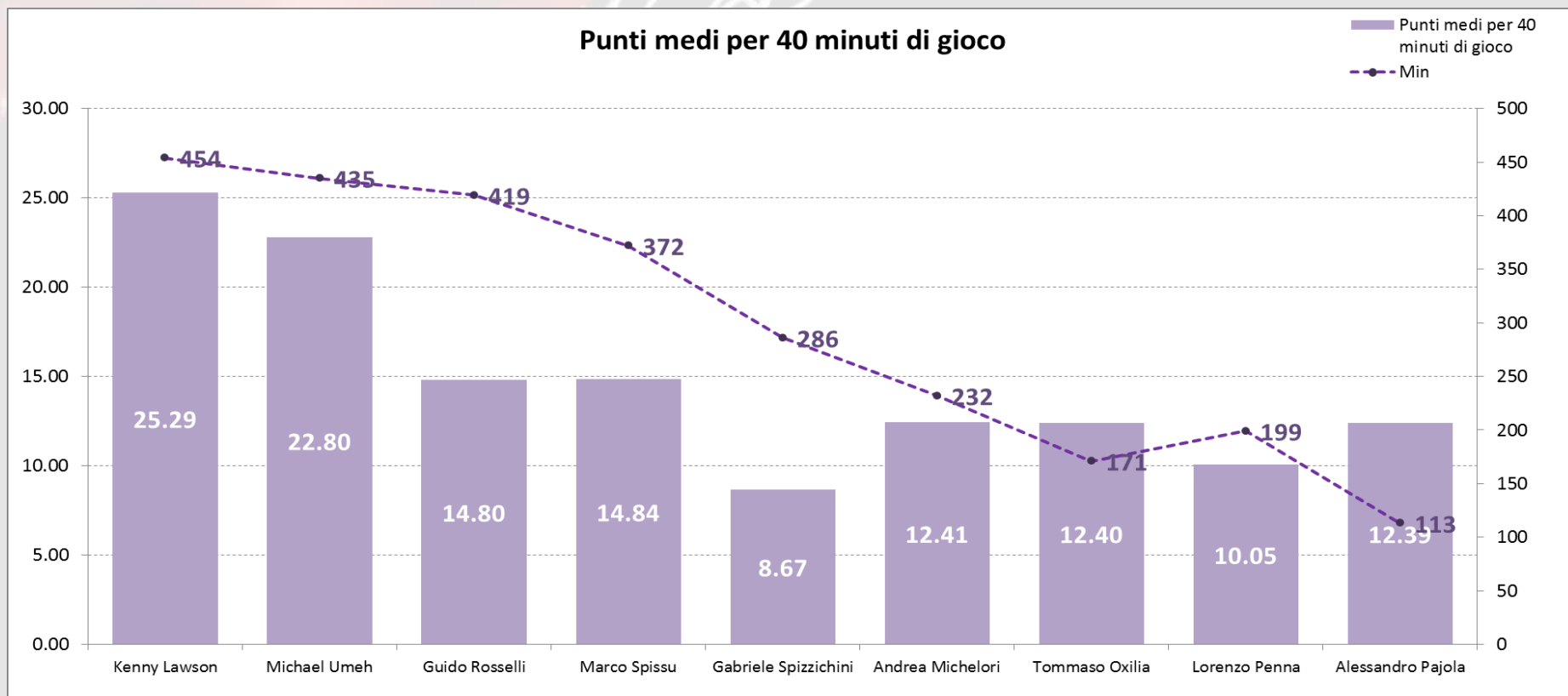


Tiri da 2 punti: tiri riusciti (barre verdi), tiri falliti (barre rosse), tiri totali (altezza totale delle barre verdi + rosse), percentuale di riuscita dei tiri ($\text{riusciti} / \text{totali} \times 100$, riga blu tratteggiata)



Progetto 2

FASE 3 - Analisi della produttività in termini di punti

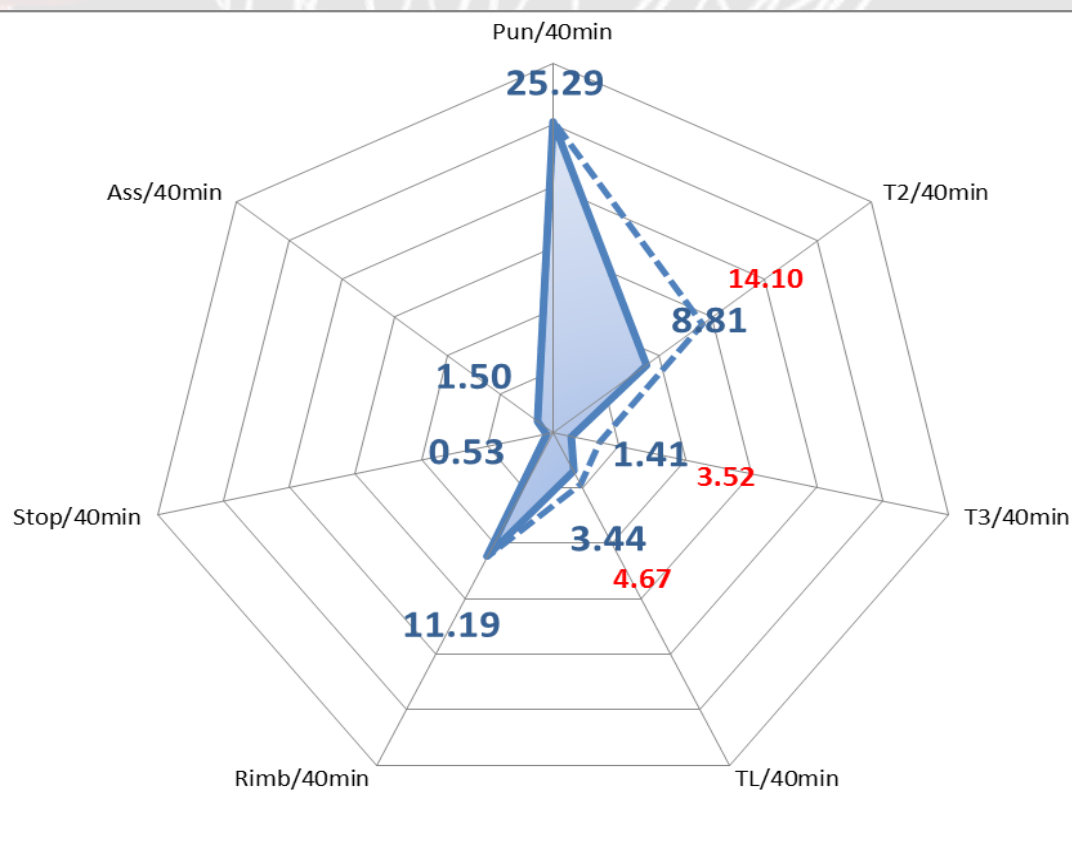


Punti medi per 40 minuti giocati (barre viola)
Minuti complessivamente giocati (linea tratteggiata)



Progetto 2

FASE 4 - Profilo complessivo del giocatore



Kenny Lawson - variabili (rapportate a 40 minuti di gioco)
Punti, Tiri da 2 punti riusciti, Tiri da 3 punti riusciti, Tiri liberi riusciti, Rimbalzi totali, Stop difensivi, Assist;

Tiri da 2 e 3 punti e tiri liberi tentati rappresentati dalla linea tratteggiata e dalle etichette rosse.



Progetto 2

FASE 5 - Analisi della variabilità di tiro

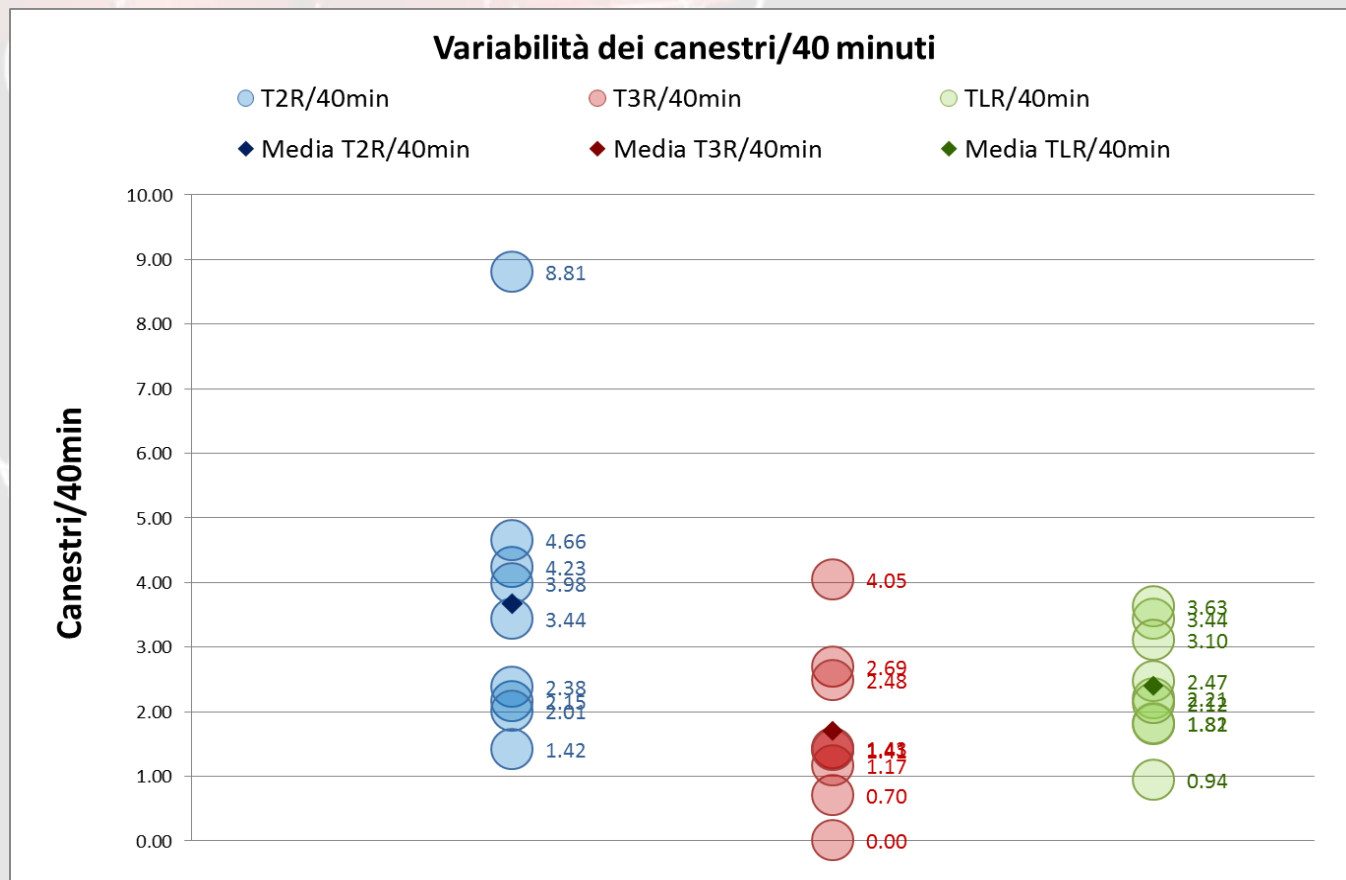
	T2R/40min	T3R/40min	TLR/40min
Kenny Lawson	8.81	1.41	3.44
Michael Umeh	4.23	4.05	2.21
Guido Rosselli	3.44	1.43	3.63
Marco Spissu	2.15	2.69	2.47
Gabriele Spizzichini	2.38	0.70	1.82
Andrea Michelori	4.66	0.00	3.10
Tommaso Oxilia	3.98	1.17	0.94
Lorenzo Penna	2.01	1.41	1.81
Alessandro Pajola	1.42	2.48	2.12
MEDIA	3.6737	1.7033	2.3927
MEDIANA	3.4368	1.4097	2.2069
SQM	2.0996	1.1304	0.8183
COEFF. VARIAZIONE	0.5715	0.6636	0.3420

Numero medio di tiri per 40 minuti giocati



Progetto 2

FASE 5 - Analisi della variabilità di tiro



Tiri riusciti da 2 punti (blu), da 3 punti (rosso), liberi (verde), in media sui 40 minuti di gioco; valore medio della squadra (rombo colorato) e valori dei singoli giocatori (bolle)



Progetto 2

FASE 6 - Analisi della concentrazione dei punti

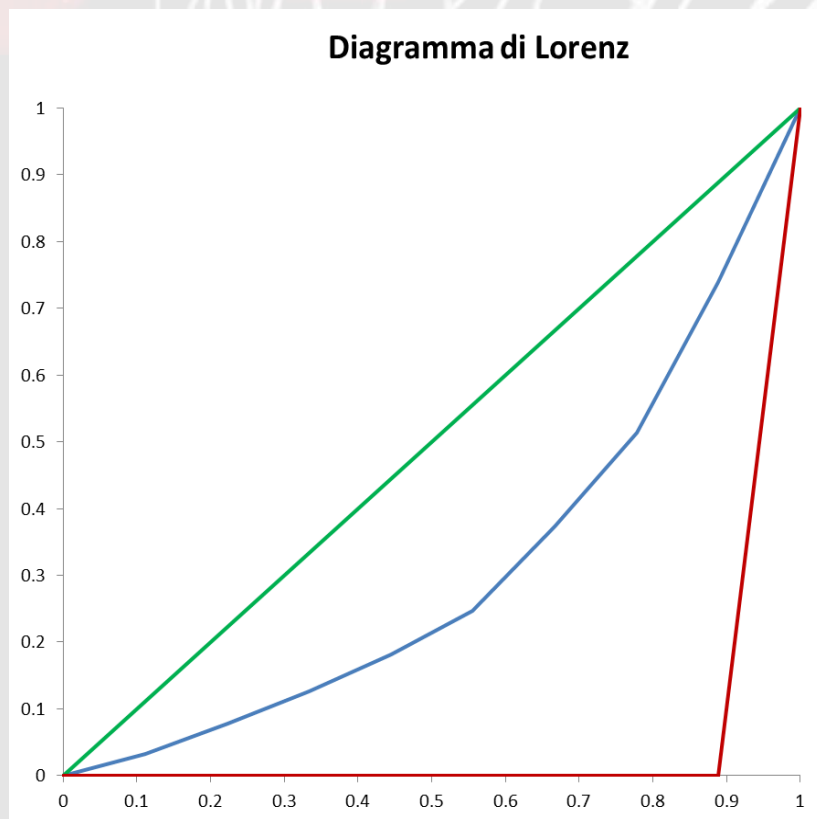


Diagramma di Lorenz dei punti segnati
retta di equidistribuzione (verde), spezzata di max concentrazione (rosso), spezzata di Lorenz (blu)

Il rapporto di concentrazione di Gini risulta pari a 42,77% del suo massimo teorico



Progetto 3

MONITORAGGIO PRESTAZIONI DEGLI STUDENTI

Gli studenti analizzano le proprie prestazioni sportive (+ confronti tra studenti e nel tempo)

Obiettivo:

Insegnare agli studenti come monitorare e analizzare le proprie prestazioni sportive

Competenze sviluppate:

- Costruzione e gestione di form online
- Utilizzo funzioni Tabella e Grafico Pivot di Excel



Progetto 3

File: 'Progetto3.xlsx' disponibile on line sul sito BDSports
Su richiesta, condivisione del form online già predisposto

Esempio:

Atletica leggera - 80 m piani, 200 m piani, salto in lungo

FASE 1 - Costruzione form online e raccolta dati

Google forms: goo.gl/forms/60Te0a82ZAVv9qDF2



Progetto 3

1. Cognome dello studente * _____ 2. Classe * _____ 3. Sezione * _____ 4. Genere <i>Contrassegna solo un ovale.</i> ☐ Maschio ☐ Femmina 5. Svolgi attività sportiva agonistica * <i>Contrassegna solo un ovale.</i> ☐ Sì, atletica leggera ☐ Sì, altro sport ☐ No 6. Quanto ti sei allenato rispetto al test precedente? * <i>Contrassegna solo un ovale.</i> ☐ Per nulla ☐ Poco ☐ Abbastanza ☐ Molto	7. Numero sessione di misurazione * _____ 8. 80 m piani (centesimi di secondo) * _____ 9. 200 m piani (centesimi di secondo) * _____ 10. Salto in lungo (centimetri) * _____
---	--

Scheda di sintesi del form online
(SINISTRA: Informazioni sullo studente; DESTRA: Risultati dei test)



Progetto 3

FASE 2 - Elaborazione dati

Tabella e **Grafico Pivot** di Microsoft Excel per:

1. valutare le prestazioni medie degli studenti nelle varie discipline, complessivamente e separatamente per le varie sessioni di misurazione;
2. monitorare l'andamento nel tempo delle prestazioni per singolo studente;
3. mettere in luce le differenze nelle prestazioni medie per attività sportiva svolta all'esterno della scuola, classe, genere.



Progetto 3

Elenco campi tabella pivot

Selezionare i campi da aggiungere al rapporto:

- ☐ Informazioni cronologiche
- ☐ Cognome dello studente
- ☐ Classe
- ☐ Sezione
- ☐ Genere
- ☐ Svolgi attività sportiva agonistica?
- ☐ Quanto ti sei allenato rispetto al test precedente?
- ☒ **Numero sessione di misurazione**
- ☒ **80 m piani (centesimi di secondo)**
- ☐ 200 m piani (centesimi di secondo)
- ☐ Salto in lungo (centimetri)

Trascinare i campi nelle aree sottostanti:

Filtro rapporto

Campi legenda (serie)

Campi asse (categorie)

Numero sessione di misurazione

Valori

Media di 80 m piani (centesimi di s...)

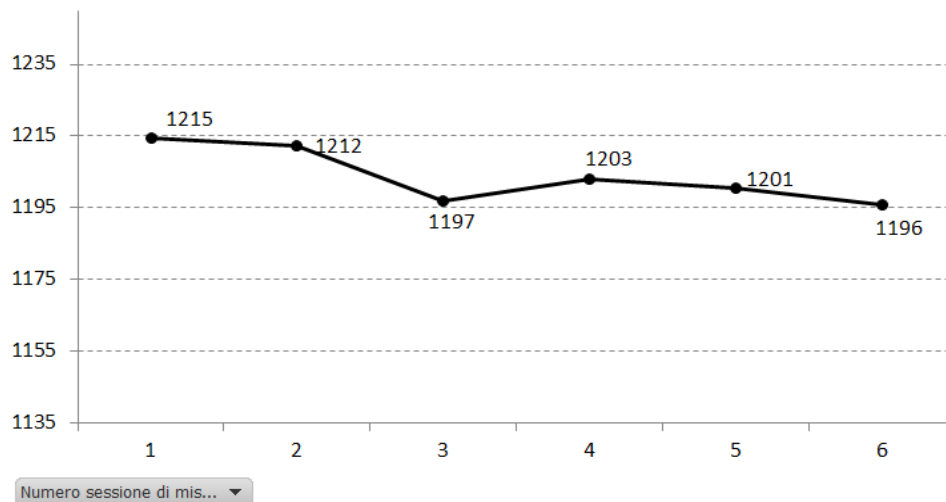
☐ Rinvia aggiornamento layout

Aggiornamento

Etichette di riga	Media di 80 m piani (centesimi di secondo)
1	1215
2	1212
3	1197
4	1203
5	1201
6	1196
Totale complessivo	1204

Media di 80 m piani (centesimi di...

80 m piani - centesimi di secondo



Analisi delle prestazioni medie negli 80 m piani lungo le varie sessioni di misurazione



Progetto 3

Elenco campi tabella pivot

Selezionare i campi da aggiungere al rapporto:

- ☐ Informazioni cronologiche
- ☒ **Cognome dello studente**
- ☐ Classe
- ☐ Sezione
- ☐ Genere
- ☐ Svolgi attività sportiva agonistica?
- ☐ Quanto ti sei allenato rispetto al test precedente?
- ☒ **Numero sessione di misurazione**
- ☒ **80 m piani (centesimi di secondo)**
- ☐ 200 m piani (centesimi di secondo)
- ☐ Salto in lungo (centimetri)

Trascinare i campi nelle aree sottostanti:

Filtro rapporto

Cognome dello studente

Campi legenda (serie)

Campi asse (categorie)

Numero sessione di misurazione

Valori

Media di 80 m piani (centesimi di s...)

☐ Rinvia aggiornamento layout

Aggiornamento

Cognome dello studente Verdi

Etichette di riga

1
2
3
4
5
6

Totale complessivo

Cerca

(Tutto)
Bianchi
Neri
Rossi
Verdi

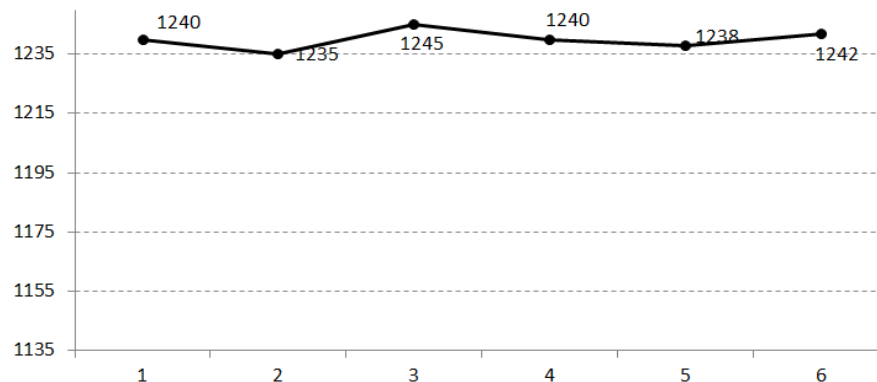
☐ Seleziona più elementi

OK Annulla

Cognome dello stu...

Media di 80 m piani (centesimi di s...)

80 m piani - centesimi di secondo



Numero sessione di mis...

Analisi delle prestazioni medie negli 80 m piani lungo le varie sessioni di misurazione – selezione dello studente Verdi



Progetto 3

Elenco campi tabella pivot

Selezionare i campi da aggiungere al rapporto:

- ☐ Informazioni cronologiche
- ☒ **Cognome dello studente**
- ☐ Classe
- ☐ Sezione
- ☐ Genere
- ☐ Svolgi attività sportiva agonistica?
- ☐ Quanto ti sei allenato rispetto al test precedente?
- ☒ **Numero sessione di misurazione**
- ☒ **80 m piani (centesimi di secondo)**
- ☐ 200 m piani (centesimi di secondo)
- ☐ Salto in lungo (centimetri)

Trascinare i campi nelle aree sottostanti:

Filtro rapporto: **Cognome dello studente**

Campi legenda (serie):

Campi asse (categorie): **Numero sessione di misurazione**

Σ Valori: **Media di 80 m piani (centesimi di s...)**

☐ Rinvia aggiornamento layout Aggiornamento

Cognome dello studente **Bianchi**

Etichette di riga: **Media**

Cerca: (Tutto) **Bianchi**, Neri, Rossi, Verdi

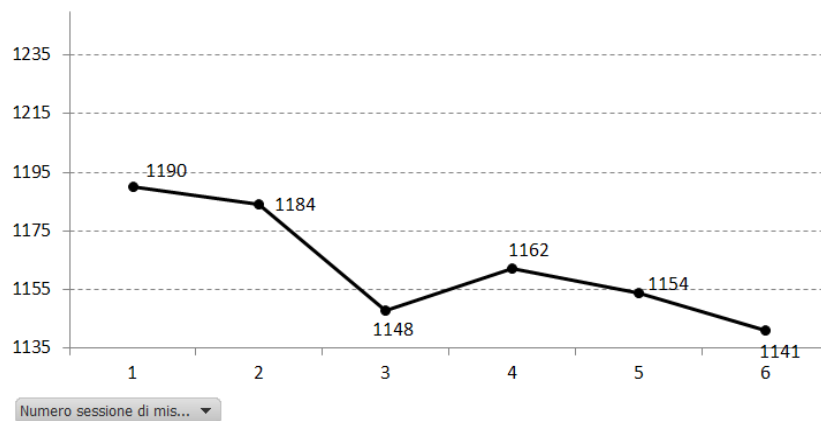
☐ Seleziona più elementi

OK Annulla

Cognome dello stu... ↗

Media di 80 m piani (centesimi di s...)

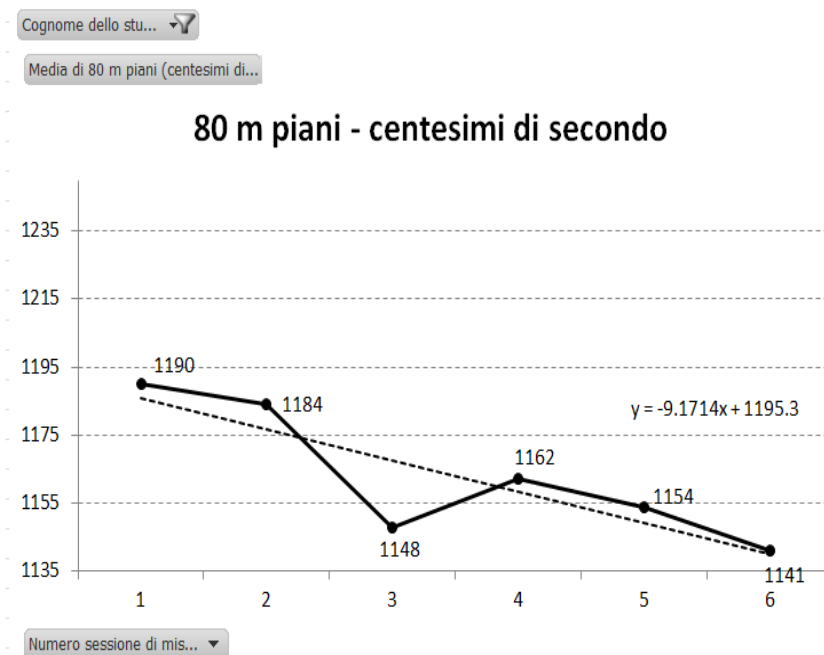
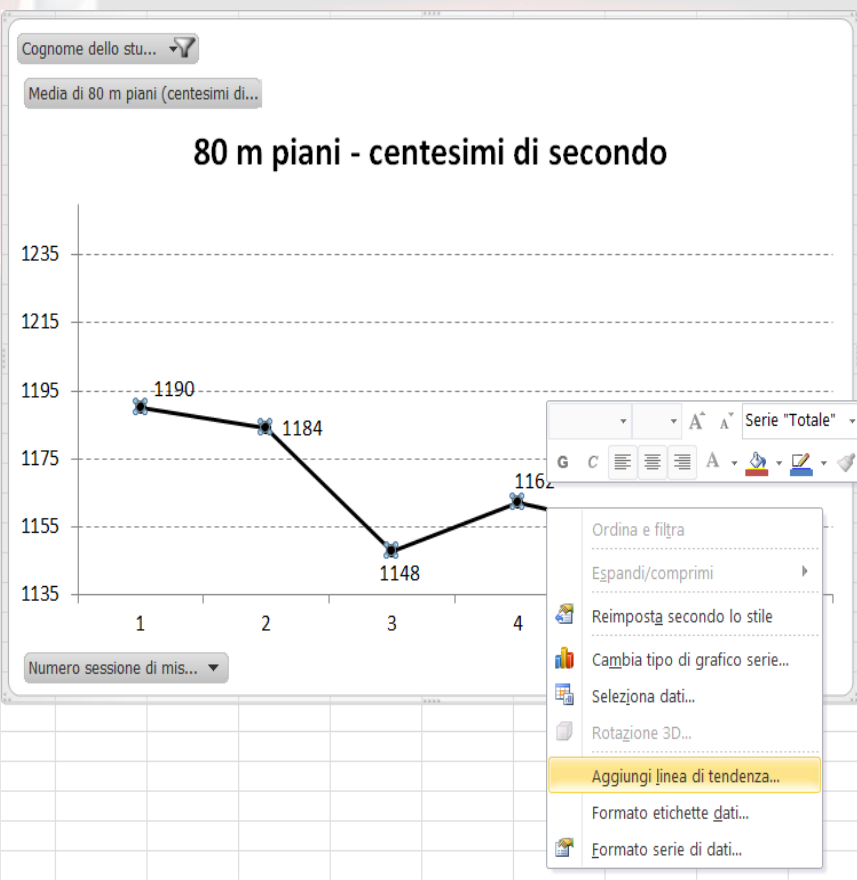
80 m piani - centesimi di secondo



Analisi delle prestazioni medie negli 80 m piani lungo le varie sessioni di misurazione – selezione dello studente Bianchi



Progetto 3



Interpolazione lineare per sintetizzare l'andamento delle prestazioni dei singoli studenti



Progetto 3

Elenco campi tabella pivot

Selezionare i campi da aggiungere al rapporto:

- ☐ Informazioni cronologiche
- ☐ Cognome dello studente
- ☐ Classe
- ☐ Sezione
- ☐ Genere
- ☒ Svolgi attività sportiva agonistica?
- ☐ Quanto ti sei allenato rispetto al test precedente?
- ☒ Numero sessione di misurazione
- ☒ 80 m piani (centesimi di secondo)
- ☐ 200 m piani (centesimi di secondo)
- ☐ Salto in lungo (centimetri)

Trascinare i campi nelle aree sottostanti:

Filtro rapporto

Etichette di colonna

Svolgi attività sportiva agonistica?

Etichette di riga

Numero sessione di misurazione

Valori

Media di 80 m piani (centesimi di s...)

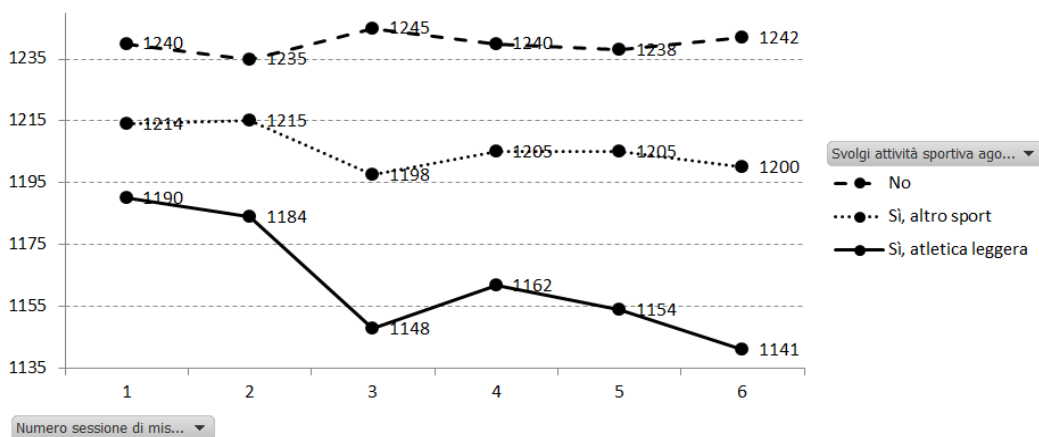
Rinvia aggiornamento layout

Aggiornamento

Media di 80 m piani (centesimi di secondo)		Etichette di colonna			
Etichette di riga		No	Sì, altro sport	Sì, atletica leggera	Totale complessivo
1		1240	1214	1190	1215
2		1235	1215	1184	1212
3		1245	1198	1148	1197
4		1240	1205	1162	1203
5		1238	1205	1154	1201
6		1242	1200	1141	1196
Totale complessivo		1240	1206	1163	1204

Media di 80 m piani (centesimi di...)

80 m piani - centesimi di secondo



Analisi delle prestazioni medie negli 80 m piani lungo le varie sessioni di misurazione – influenza dell'attività sportiva agonistica extrascolastica



Progetto 3

Elenco campi tabella pivot

Selezionare i campi da aggiungere al rapporto:

- ☐ Informazioni cronologiche
- ☐ Cognome dello studente
- ☒ **Classe**
- ☐ Sezione
- ☐ Genere
- ☐ Svolgi attività sportiva agonistica?
- ☐ Quanto ti sei allenato rispetto al test precedente?
- ☐ Numero sessione di misurazione
- ☒ **80 m piani (centesimi di secondo)**
- ☐ 200 m piani (centesimi di secondo)
- ☐ Salto in lungo (centimetri)

Trascinare i campi nelle aree sottostanti:

Filtro rapporto

Etichette di colonna

Etichette di riga

Classe

Valori

Media di 80 m piani (centesimi di ...)

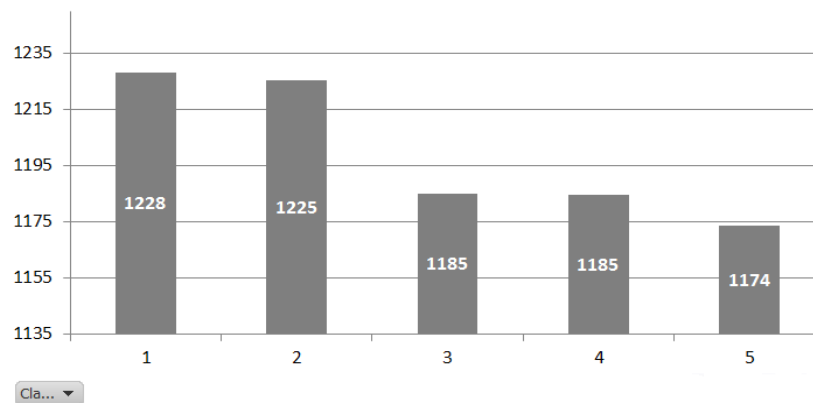
Rinvia aggiornamento layout

Aggiornamento

Etichette di riga	Media di 80 m piani (centesimi di secondo)
1	1228
2	1225
3	1185
4	1185
5	1174
Totale complessivo	1204

Media di 80 m piani (centesimi di ...)

80 m piani - centesimi di secondo



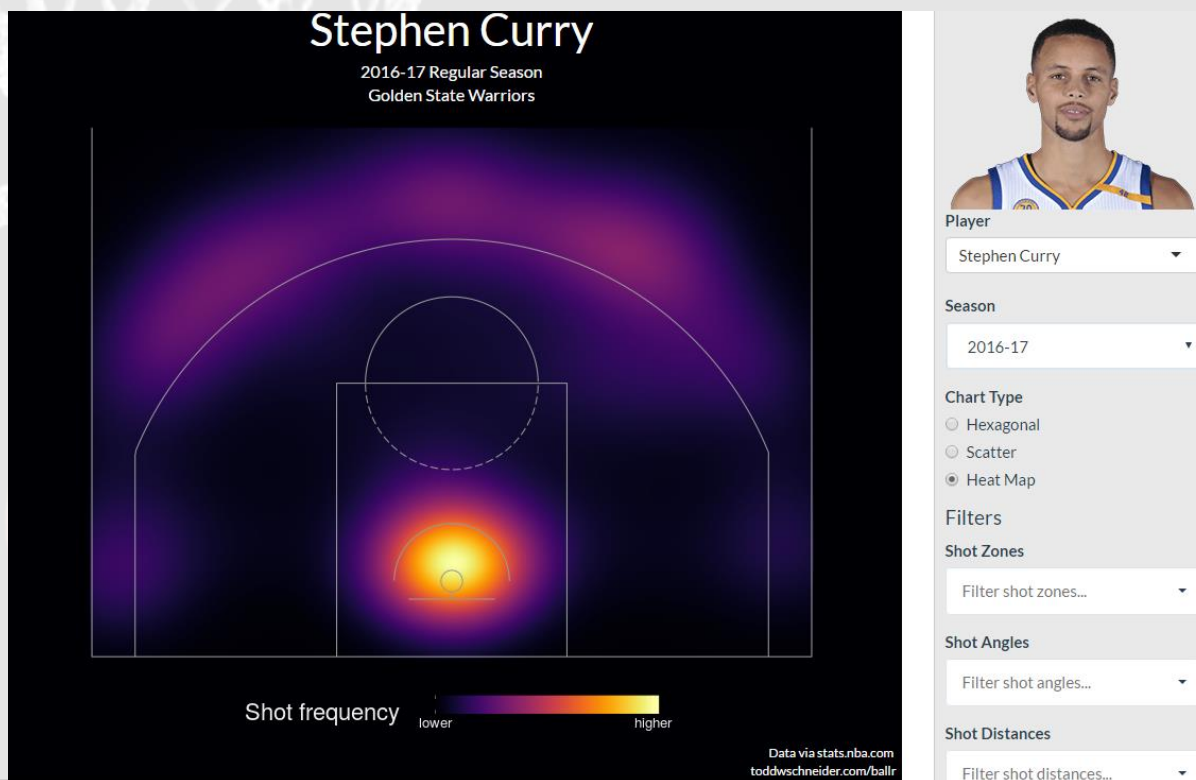
Prestazioni medie negli 80 m piani suddivise per classe frequentata



Approfondimenti

Esistono, anche on line, vari strumenti interattivi di visualizzazione di dati, risultati di analisi, traiettorie dei giocatori,

Esempi al link <http://bodai.unibs.it/BDSports/basketball.htm>





Conclusioni

I progetti presentati sono rivolti agli insegnanti che volessero sviluppare, con i propri studenti, un percorso formativo appassionante e, in taluni casi, multidisciplinare, coinvolgendo insegnanti di matematica, informatica, educazione fisica

Il materiale didattico di ciascun progetto è on line (tutte le informazioni utili si trovano al link bodai.unibs.it/BDSports/Students.htm)

Dettagli sul sito BDSports e nell'articolo:

P. Zuccolotto e M. Manisera (2017). Sport Analytics: la statistica divertente, *Induzioni*