L'influenza dei calciatori stranieri sull'andamento delle squadre della Serie A e delle Nazionali del nostro Paese (Nazionale "maggiore" ed Under 21)

Fabio Ruini - matricola nr. 1 - 22112@unimo.it

1) Descrizione del problema e obiettivi dell'analisi statistica

A tutti coloro che sono appassionati di sport e, in particolare, di calcio è ben noto che negli ultimi anni si è assistito ad una vera e propria "invasione" di atleti stranieri nei nostri campionati. Ciò è dovuto soprattutto alla cosiddetta "Sentenza Bosman" (si veda a riguardo l'indirizzo: http://europa.eu.int/comm/sport/key_files/circ/b_bosman_it.html), che ha imposto alle federazioni calcistiche dei vari Paesi appartenenti all'Unione Europea di aprire del tutto le porte ai giocatori comunitari, rompendo con una ormai secolare tradizione fatta da rigide limitazioni.

L'entrata in vigore della nuova regolamentazione, digerita malvolentieri dalla maggior parte degli appassionati del nostro paese, abituati all'idea di vantare "il più bel campionato del mondo", ha coinciso con l'inizio di un periodo nero per il calcio italiano.

Da una parte, gli insuccessi mietuti dalle squadre italiane di club nelle varie competizioni continentali (parzialmente riscattate dall'affermazione in Champions League del Milan, alcune settimane or sono) degli ultimi anni. Dall'altra, le delusioni continue rimediate dalla nostra Nazionale (e mitigate solo in parte dal secondo posto conquistato nel campionato europeo del 2000), eliminata ai quarti di finale nei mondiali di Francia '98 ed agli ottavi nel più recente torneo di Korea/Giappone 2002.

Inutile sottolineare che, date queste premesse, il "momentaccio" del calcio italiano ha avuto violente ripercussioni sui mass-media, nonché negli innumerevoli "bar sport" di cui la Penisola è piena. Non passa giorno senza che sui quotidiani sportivi si discuta del "problema stranieri". C'è chi auspica una nuova chiusura delle frontiere, argomentando che, per i presidenti dei nostri club, l'acquisto di giocatori stranieri sia un "pallino", quando non un'azione di marketing, piuttosto che una reale necessità tecnica. E ovviamente c'è chi, al contrario, vorrebbe liberalizzare anche il tesseramento dei giocatori extra-comunitari.

Al di là di quelle che possono essere le opinioni soggettive, ritengo che possa comunque essere interessante affrontare il problema da un punto di vista statistico. In questo breve elaborato, cercherò pertanto di analizzare, da un'inquadratura quanto più oggettiva possibile, l'influenza che hanno avuto i calciatori stranieri sull'andamento delle squadre di club in cui hanno militato nella stagione da poco terminata. Al tempo stesso, prendendo in esame alcune statistiche dei campionati disputati negli ultimi 11 anni, andrò alla ricerca di un eventuale correlazione tra l'aumento di stranieri nel nostro campionato e le brucianti umiliazioni riportate dagli Azzurri.

2) Descrizione e fonte della matrice dei dati

Ai fini dell'analisi sono state realizzate due distinte matrici dei dati.

La prima matrice, relativa al campionato italiano di calcio (Serie A 2002/03), è composta da 18 righe (una per ciascuna squadra iscritta) e da 12 colonne (nome della squadra, numero di stranieri in rosa, numero di difensori, centrocampisti ed attaccanti stranieri in rosa, ampiezza della rosa, numero complessivo di difensori, centrocampisti ed attaccanti in rosa, punti totalizzati, numero reti segnate e subite).

La seconda, relativa invece alla Nazionale "maggiore" ed alla Nazionale Under 21, è formata da 11 righe (una per ciascuna stagione presa in esame) e da 6 colonne (stagione, giocatori stranieri nel campionato italiano, numero di gare disputate dalle due Nazionali, punti totalizzati dalle due Nazionali).

Le modalità associate alle singole variabili sono state prelevate dai seguenti archivi:

- rose dei giocatori del campionato di serie A: "Almanacco illustrato del calcio 2003" (ed. Panini);

- classifica finale del campionato di serie A 2002/03: http://www.lega-calcio.it/ita/atim_classifica.shtml;
- risultati della Nazionale: http://www.nazionaleitalianacalcio.it.

3) Analisi preliminari

Le matrici dei dati sono state create a partire da due fogli di lavoro di Microsoft Excel ("Serie A 2002-03.xls" e "Nazionale italiana.xls"), compilati manualmente. Tali fogli di lavoro sono poi stati convertiti in formato testo (estensione .txt e nessun delimitatore impostato) e successivamente importati in R (previa impostazione della corretta "working directory") con il duplice utilizzo della funzione "read.table()":

```
> seriea <- read.table("Serie A 2002-03.txt", header=TRUE)
> nazionale <- read.table("Nazionale italiana.txt", header=TRUE)
```

Per entrambi i data frames così ottenuti, si è provveduto all'"attach", in modo da poter disporre delle singole variabili in essi contenute:

```
> attach(seriea)
> attach(nazionale)
```

Su queste variabili è stata successivamente eseguita una serie di calcoli, che hanno permesso di ottenere nuovi parametri di confronto:

```
> propstranieri = nrstranieri / nrgiocatori
> propdifensoristranieri = nrdifensoristranieri / nrdifensori
> propcentrocampististranieri = nrcentrocampististranieri / nrcentrocampisti
> propattaccantistranieri = nrattaccantistranieri / nrattaccanti
> propdifcentstranieri = (nrdifensoristranieri + nrcentrocampististranieri) /
(nrdifensori + nrcentrocampisti)
> propcentrattstranieri = (nrcentrocampististranieri + nrattaccantistranieri) /
(nrcentrocampisti + nrattaccanti)
```

Per ottenere una visione d'insieme di questi nuovi dati, sono stati disegnati i rispettivi boxplot **[grafico nr. 1]**:

```
> boxplot(propstranieri, propdifensoristranieri, propcentrocampististranieri,
propattaccantistranieri, propdifcentstranieri, propcentrattstranieri, col=26:31,
main="Proporzione dei giocatori stranieri nelle varie rose, in base al ruolo ricoperto",
xlab="giocatori stranieri in rosa (in rosso), difensori stranieri (in verde),
centrocampisti stranieri (in blu), attaccanti stranieri (in azzurro),", sub="difensori e
centrocampisti stranieri (in viola), centrocampisti ed attaccanti stranieri (in giallo)",
ylab="proporzione di giocatori stranieri")
```

In base alla classifica finale, si è poi provveduto a creare due vettori unidimensionali, contenenti le proporzioni di stranieri in rosa per le prime quattro squadre classificate e, rispettivamente, per le ultime quattro (ossia quelle retrocesse in Serie B):

```
> prime4 = c(propstranieri[8], propstranieri[7], propstranieri[10], propstranieri[9])
> ultime4 = c(propstranieri[17], propstranieri[5], propstranieri[14], propstranieri[1])
```

Lo studio si è evoluto mettendo ovviamente a confronto i boxplot di queste due nuove variabili **[grafico nr. 2]**:

```
> boxplot(prime4,ultime4,col=3:2,main="Proporzione di giocatori stranieri per le prime
quattro squadre classificate e per quelle retrocesse",ylab="proporzione di stranieri
nella rosa",xlab="squadre posizionate ai primi 4 posti della classifica (in verde) -
squadre retrocesse in serie B (in rosso)")
```

Successivamente si è preso in considerazione l'insieme di squadre con una proporzione di stranieri superiore al 40% (valore limite, che divide esattamente a metà le diciotto osservazioni) e lo si è messo a confronto con quello delle squadre con una proporzione minore di 0.4, creando un doppio boxplot focalizzato sui punti totalizzati **[grafico nr. 3]**:

```
> piudi40 = c(punti[3], punti[7], punti[8], punti[9], punti[10], punti[12], punti[15],
punti[16], punti[18])
> menodi40 = c(punti[1], punti[2], punti[4], punti[5], punti[6], punti[11], punti[13],
punti[14], punti[17])
```

```
> boxplot(piudi40, menodi40, col=5:9, main="Punti totalizzati dalle squadre con più o meno del 40% di stranieri in rosa", ylab="numero di punti totalizzati", xlab="squadre con una proporzione di stranieri in rosa >= 0.4 (in azzurro) - squadre con una proporzione di stranieri in rosa < 0.4 (in viola)")
```

Sul contenuto del data frame "seriea" è stata effettuata un'ultima operazione in modo da ottenere, per ciascuna squadra, la differenza reti fatta registrare al termine del campionato:

```
> diffreti = nrretisegnate – nrretisubite
```

Per quanto riguarda invece la parte dell'analisi relativa alle nazionali, si è innanzitutto provveduto a tracciare il grafico **[grafico nr. 4]** che mostra l'andamento del numero di stranieri presenti nella Serie A del campionato italiano, nelle ultime undici stagioni calcistiche

```
> plot.default(stagione, giocatoristranieri, col="blue", main="Numero di stranieri nel campionato italiano, dalla stagione 1992/93 alla stagione 2002/03", sub="(le stagioni sono numerate progressivamente: 2 = 1993/94, 4 = 1995/96, ecc)", xlab="Stagione", ylab="Numero di stranieri presenti nel campionato",type="b")
```

Per sottolineare l'effetto "devastante" del "caso Bosman" si è aggiunta al grafico una retta, intersecata, sull'asse delle X, con l'anno di emissione della relativa sentenza:

```
> abline(v=4.5,col="red")
```

Sono state inoltre calcolate le medie dei punti a partita totalizzati dalle due squadre nelle stagioni in esame:

```
> puntiapartitanazionalea = puntinazionalea / garenazionalea
> puntiapartitaunder21 = puntiunder21 / garenazionaleu21
```

E' bene sottolineare che i dati contenuti nel data frame "nazionale" sono stati calcolati in un modo un po' "particolare" rispetto a quanto non si faccia solitamente nelle analisi di questo genere. E' infatti consuetudine prendere in considerazione i risultati di una nazionale utilizzando l'anno solare come arco temporale di riferimento (quindi dal 1° gennaio al 31 dicembre). Nel nostro caso, invece, per poter osservare eventuali correlazioni con quanto avviene nel campionato italiano, si è preferito "adattare" i risultati delle nazionali all'arco temporale del torneo di Serie A (indicativamente compreso tra il 1° settembre ed il 31 agosto dell'anno solare successivo). Nella stagione 2001/02, ad esempio, sono state considerate tutte le partite disputate dalle due nazionali nel periodo compreso tra il 1° settembre 2001 ed il 31 agosto 2002.

4) Risultati

Le analisi effettuate sul data frame "seriea" al punto precedente sono servite soprattutto come spunto per studiare le correlazioni tra le varie variabili raccolte/dedotte. Tra le associazioni più significative, per le quali si è provveduto a disegnare il diagramma di dispersione (arricchito anche dalla retta di regressione, qui utilizzata non a scopi previsionali quanto per una miglior chiarezza espositiva del grafico), meritano almeno una citazione:

- proporzione di giocatori stranieri in rosa e punti totalizzati dalla squadra (coefficiente di correlazione 0.67) **[grafico numero 5]**:

  ```
  > plot(propstranieri~punti, main="Proporzione di stranieri in rosa e punti totalizzati nel campionato", xlab="punti totalizzati al termine del campionato", ylab="proporzione di stranieri in rosa",col="blue")
  > modstranieripunti = lm(propstranieri~punti)
  > abline(modstranieripunti)
  ```

  Per completezza, è stato tracciato anche il grafico dei residui per tale correlazione **[grafico numero 6]**:

  ```
  > plot(resid(modstranieripunti), main="Grafico dei residui per la correlazione tra giocatori stranieri nel campionato e media punti nazionale A", xlab="", ylab="Residui")
  > abline(h=0)
  ```

- proporzione di difensori stranieri in rosa e differenza reti della squadra (coefficiente di correlazione 0.66) **[grafico numero 7]**:

```
> plot(propdifensoristranieri~diffreti, main="Proporzione di difensori stranieri
in rosa e differenza reti", xlab="differenza reti fatta registrare al termine del
campionato", ylab="proporzione di difensori stranieri in rosa",col="darkgreen")
> moddifstranieridiffreti = lm(propdifensoristranieri~diffreti)
> abline(moddifstranieridiffreti)
```

E' invece pressoché inesistente una qualsiasi relazione tra la proporzione di attaccanti stranieri e le reti segnate dalla squadra, la differenza reti ed i punti totalizzati (coefficienti di correlazione: 0.18, 0.25 e 0.27).

Gli altri coefficienti di correlazione misurati e che hanno mostrato un buon livello di associazione tra le due variabili prese, di volta in volta, in esame, sono i seguenti:

- tra proporzione di difensori stranieri e numero di reti segnate, subite e punti totalizzati: 0.59, -0.55, 0.61;
- tra proporzione di centrocampisti stranieri e numero di reti segnate, subite, differenza reti e punti totalizzati: 0.53, -0.51, 0.57, 0.59;
- tra proporzione di stranieri in rosa e numero di reti segnate, subite e differenza reti: 0.60, -0.61, 0.66;
- tra proporzione di difensori e centrocampisti stranieri e numero di reti segnate, subite, differenza reti e punti totalizzati: 0.63, -0.58, 0.66, 0.66;
- tra proporzione di centrocampisti ed attaccanti stranieri e numero di reti segnate, subite, differenza reti e punti totalizzati: 0.47, -0.50, 0.53, 0.55.

In merito ai dati contenuti all'interno del data frame "nazionale", è stato tracciato un altro diagramma di dispersione **[grafico nr. 8]**, riguardante la presunta correlazione (che si è in realtà rivelata molto scarsa, dato il coefficiente di correlazione pari a -0.22) tra la media punti stagionale della nazionale maggiore ed il numero di stranieri presenti nel campionato italiano nella stagione.

```
> plot(puntiapartitanazionalea~giocatoristranieri, main="Correlazione tra giocatori
stranieri nel campionato e media punti fatta registrare dalla nazionale maggiore",
xlab="numero di giocatori stranieri presenti nel campionato", ylab="media punti della
nazionale maggiore", col="darkred")
> modpuntinmstranieri = lm(puntiapartitanazionalea~giocatoristranieri)
> abline(modpuntinmstranieri)
```

Risultati ancora meno significativi (coefficiente di correlazione: -0.12) sono stati ottenuti analizzando la correlazione tra la media punti fatta registrare dalla nazionale Under 21 ed il numero di atleti stranieri militanti nel nostro campionato.

5) Conclusioni

I giocatori stranieri, statisticamente, hanno un influenza positiva sull'andamento delle squadre nelle quali giocano. Si è visto, infatti, come le formazioni che hanno una maggiore proporzione di stranieri in rosa, totalizzino mediamente più punti di quanti non ne accumulino le squadre con una maggior proporzione di atleti indigeni. Al di là dei punti in classifica, comunque, le squadre che possono vantare le maggiori proporzioni di difensori e centrocampisti stranieri (in rapporto a difensori e centrocampisti totali) incassano mediamente meno segnature rispetto alle controparti, ne segnano di più e possono pertanto vantare una miglior differenza reti.

Il discorso sin qui abbozzato è valido per la proporzione complessiva di stranieri, così come per quella di difensori e/o centrocampisti. Del tutto diversa, invece, è la situazione per quanto concerne gli attaccanti stranieri. Le misurazioni effettuate, infatti, ci mostrano che gli attaccanti stranieri non apportano benefici significativi sull'andamento delle formazioni in cui militano. E' pertanto plausibile ritenere che i milioni di euro spesi dai presidenti dei grandi club per aggiudicarsi le prestazioni dei più grandi centravanti del mondo, siano in realtà soldi investiti malamente.

Per quanto riguarda il problema della Nazionale italiana, l'analisi non ci fornisce elementi che possano far pensare all'esistenza di un qualsiasi tipo di correlazione tra il numero di stranieri presenti nel nostro campionato e gli insuccessi fatti registrare dalla rappresentativa azzurra. Da un punto di vista statistico, in sostanza, non c'è motivo di ritenere che l'andamento "altalenante" della nostra Nazionale di calcio nel corso degli ultimi undici anni possa essere imputato ad un eccessiva presenza di calciatori stranieri nel campionato italiano. I motivi del periodo buio attraversato dal calcio di casa nostra sono pertanto da ricercare altrove. Non dimentichiamoci, ad esempio, della violenta crisi finanziaria che ha recentemente sconvolto il mondo del pallone, facendo dichiarare il fallimento (limitando, volutamente, l'analisi alla sola serie A) di una società importante come la Fiorentina e la messa in mora della Lazio. Una crisi che ha avuto ripercussioni su tutto il movimento calcistico italiano, andando ad intaccare anche gli organici delle formazioni giovanili, tradizionali serbatoi di talenti per le Nazionali del futuro.

6) Tabella Dati (data frame "seriea")

squadre	nrstranieri	nrdifensoristranieri	nrcentrocampististranieri	nrattaccantistranieri	nrgiocatori	nrdifensori	nrcentrocampisti	nrattaccanti	punti	nrretisubite	nrretisegnate
Atalanta	7	1	4	2	32	11	11	5	38	47	35
Bologna	3	1	1	1	25	8	9	5	41	47	39
Brescia	11	2	4	4	25	8	7	7	42	38	36
Chievo	8	1	5	2	30	10	11	6	55	39	51
Como	6	2	2	2	24	8	8	5	24	57	29
Empoli	2	1	1	0	23	6	8	6	38	46	36
Inter	16	4	8	3	30	8	14	5	65	38	64
Juventus	11	3	5	3	26	9	9	5	72	29	64
Lazio	11	5	4	1	26	10	9	4	60	32	57
Milan	16	6	5	4	28	9	10	6	61	30	55
Modena	2	0	0	1	24	7	9	5	38	48	30
Parma	11	4	3	2	26	10	8	5	56	36	55
Perugia	6	2	2	1	24	6	9	6	42	48	40
Piacenza	5	1	2	2	25	9	8	5	30	62	44
Reggina	10	2	5	2	22	7	7	5	38	53	38
Roma	14	8	5	1	27	10	9	5	49	46	55
Torino	4	0	1	3	27	10	9	5	21	58	23
Udinese	18	6	8	3	29	9	12	5	56	35	38

Tabella Dati (data frame "nazionale")

stagione	giocatoristranieri	garenazionalea	garenazionaleu21	puntinazionalea	puntiunder21
1992-93	71	9	8	20	18
1993-94	68	15	10	31	17
1994-95	63	9	8	19	19
1995-96	67	10	10	21	21
1996-97	84	11	8	21	14
1997-98	113	12	5	22	9
1998-99	148	10	9	20	16
1999-00	152	13	12	22	29
2000-01	185	9	12	22	30
2001-02	189	13	11	19	13
2002-03	161	9	7	20	15

7) Grafico nr. 1

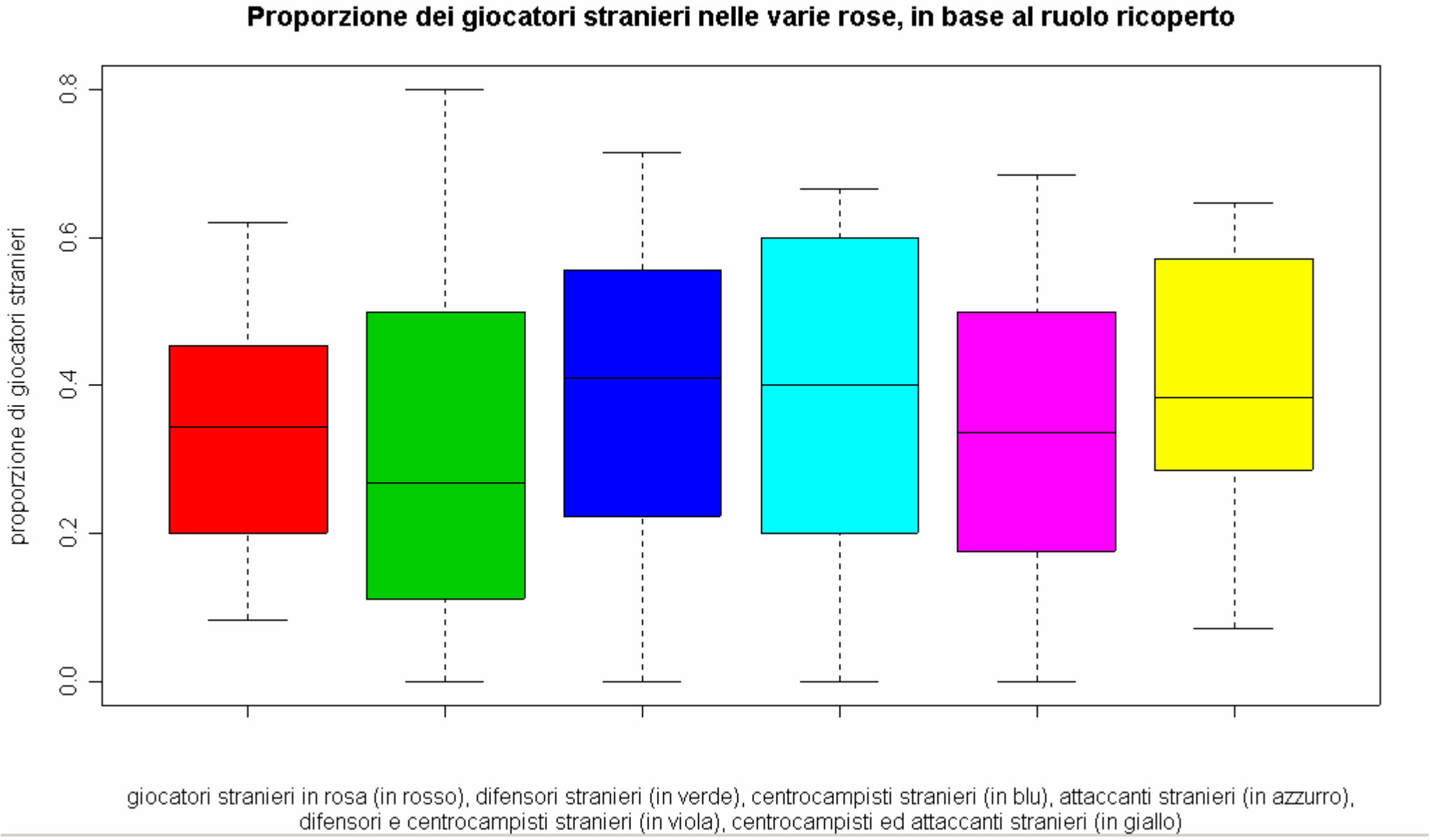


Grafico nr. 2

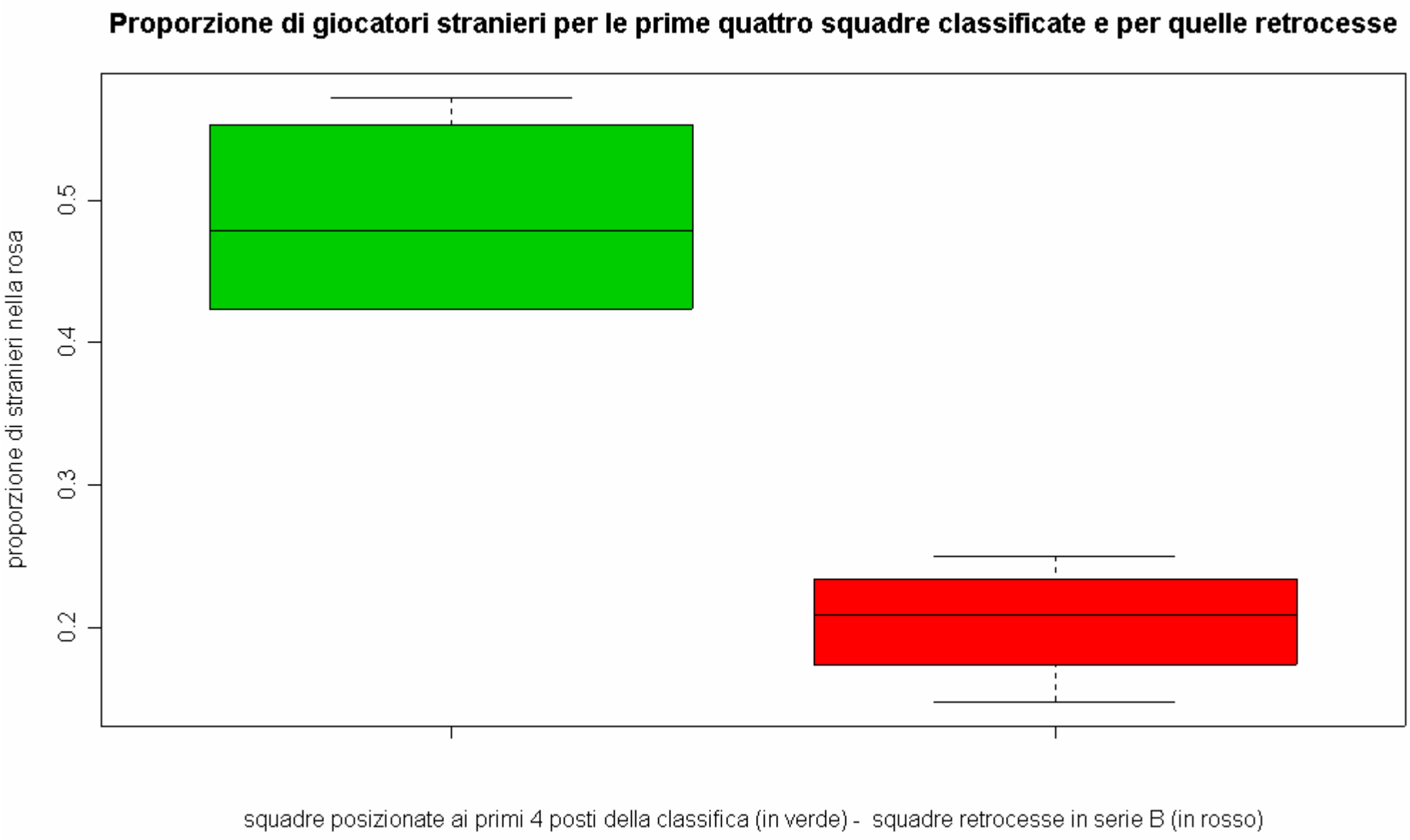

Proporzione di giocatori stranieri per le prime quattro squadre classificate e per quelle retrocesse
proporzione di stranieri nella rosa
0.2
0.3
0.4
0.5
squadre posizionate ai primi 4 posti della classifica (in verde) - squadre retrocesse in serie B (in rosso)

Grafico nr. 3

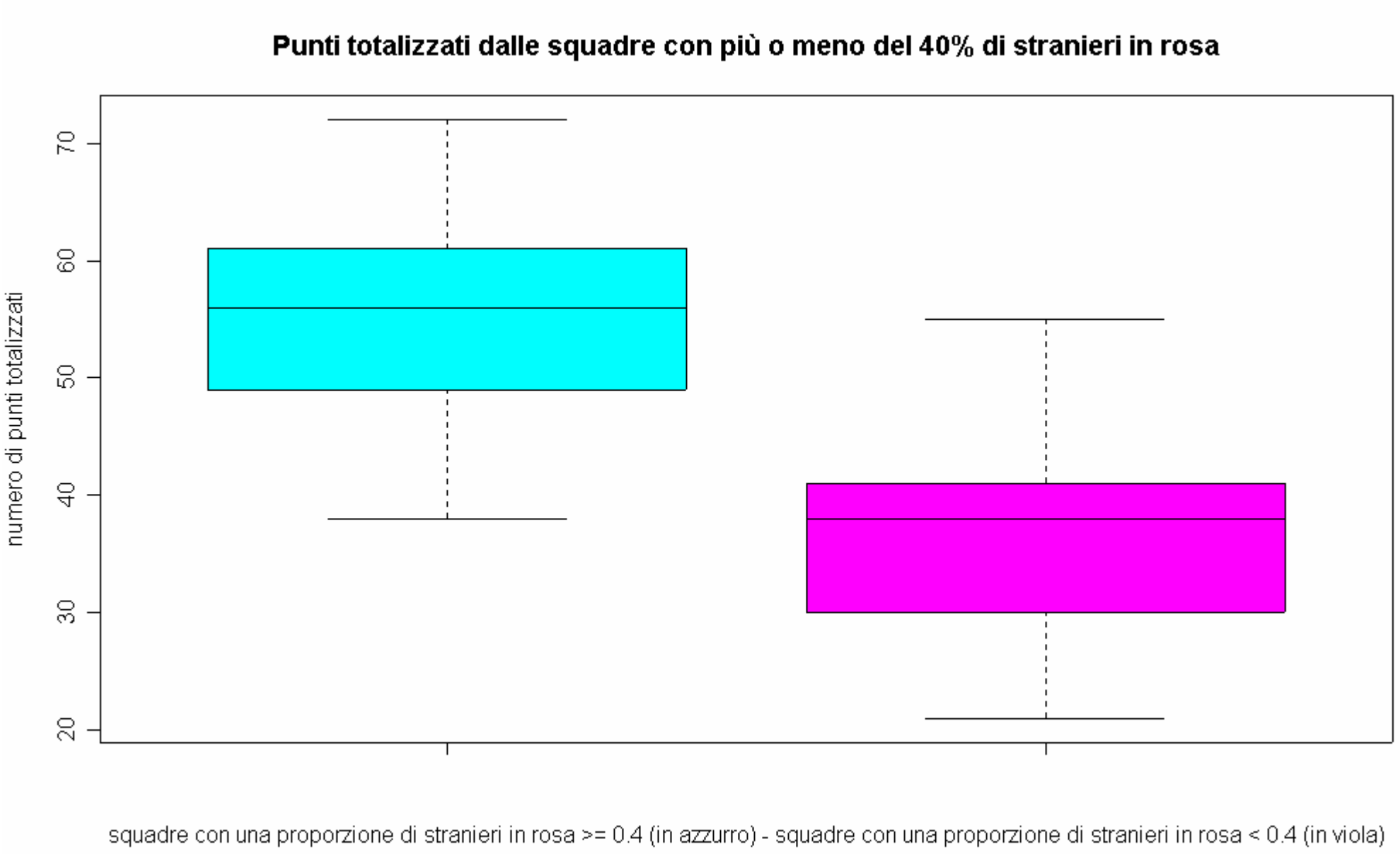

Punti totalizzati dalle squadre con più o meno del 40% di stranieri in rosa
numero di punti totalizzati
20
30
40
50
60
70
squadre con una proporzione di stranieri in rosa >= 0.4 (in azzurro) - squadre con una proporzione di stranieri in rosa < 0.4 (in viola)

Grafico nr. 4

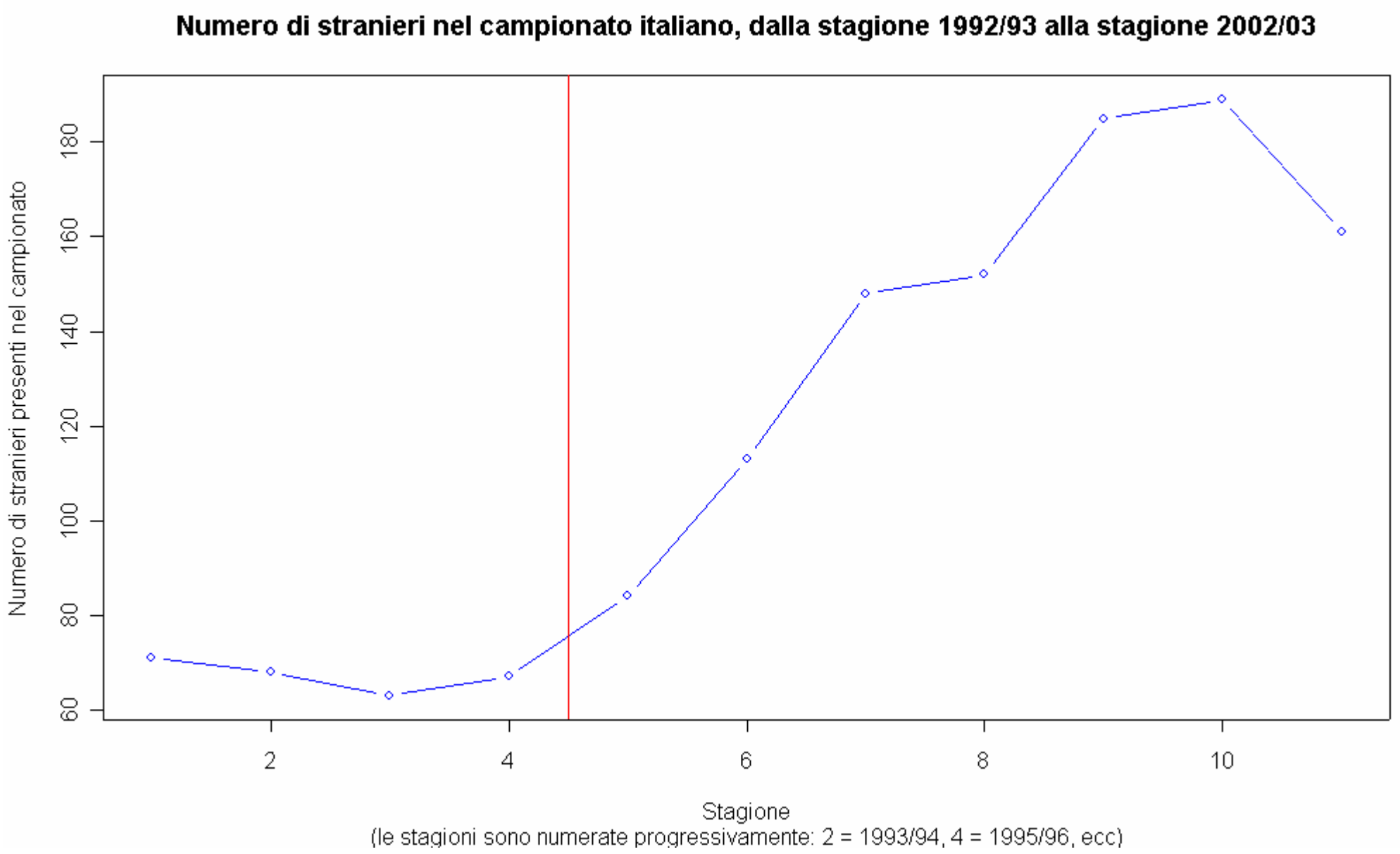

Numero di stranieri nel campionato italiano, dalla stagione 1992/93 alla stagione 2002/03
Numero di stranieri presenti nel campionato
60
80
100
120
140
160
180
2
4
6
8
10
Stagione
(le stagioni sono numerate progressivamente: 2 = 1993/94, 4 = 1995/96, ecc)

Grafico nr. 5

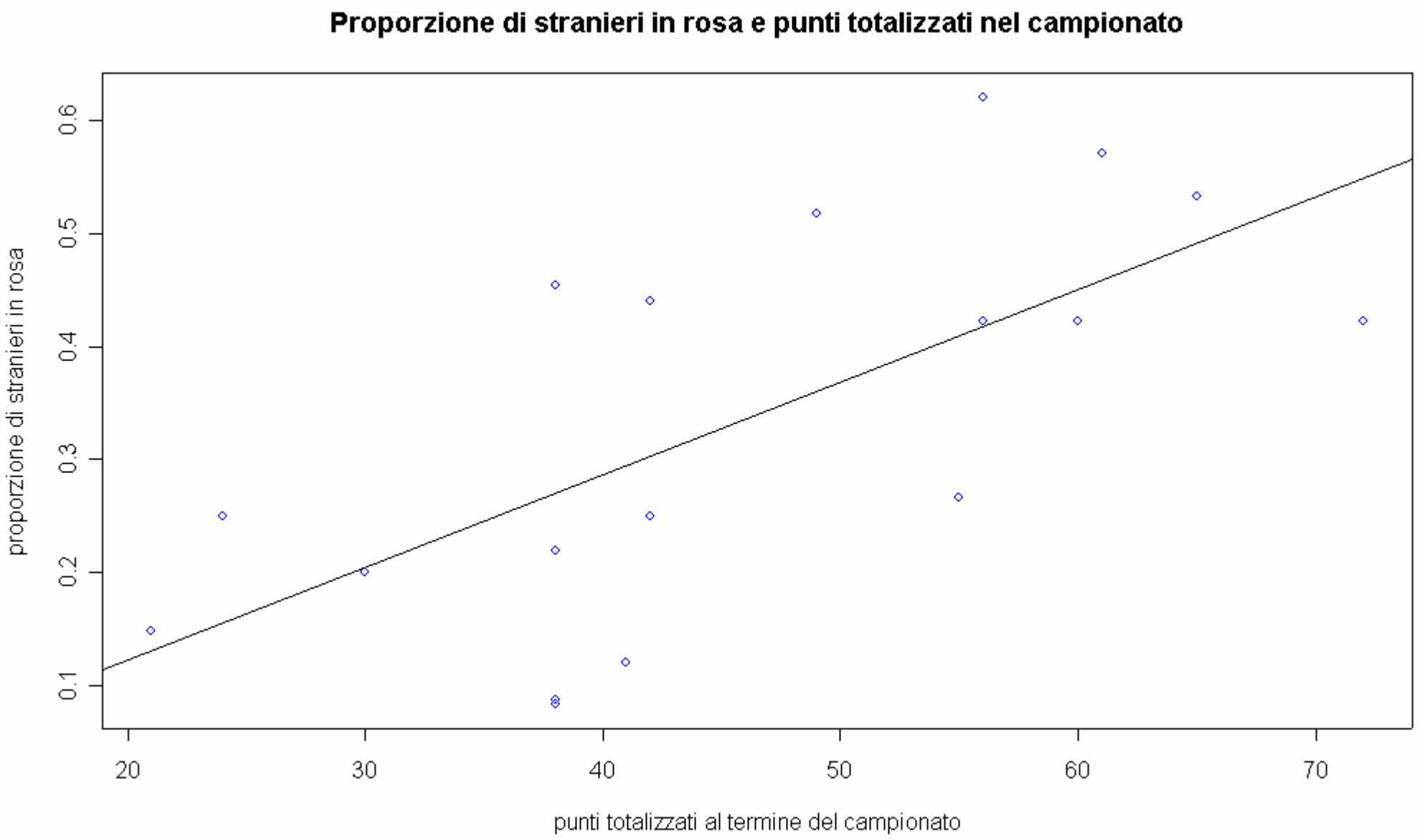

Proporzione di stranieri in rosa e punti totalizzati nel campionato
proporzione di stranieri in rosa
0.1
0.2
0.3
0.4
0.5
0.6
20
30
40
50
60
70
punti totalizzati al termine del campionato

Grafico nr. 6

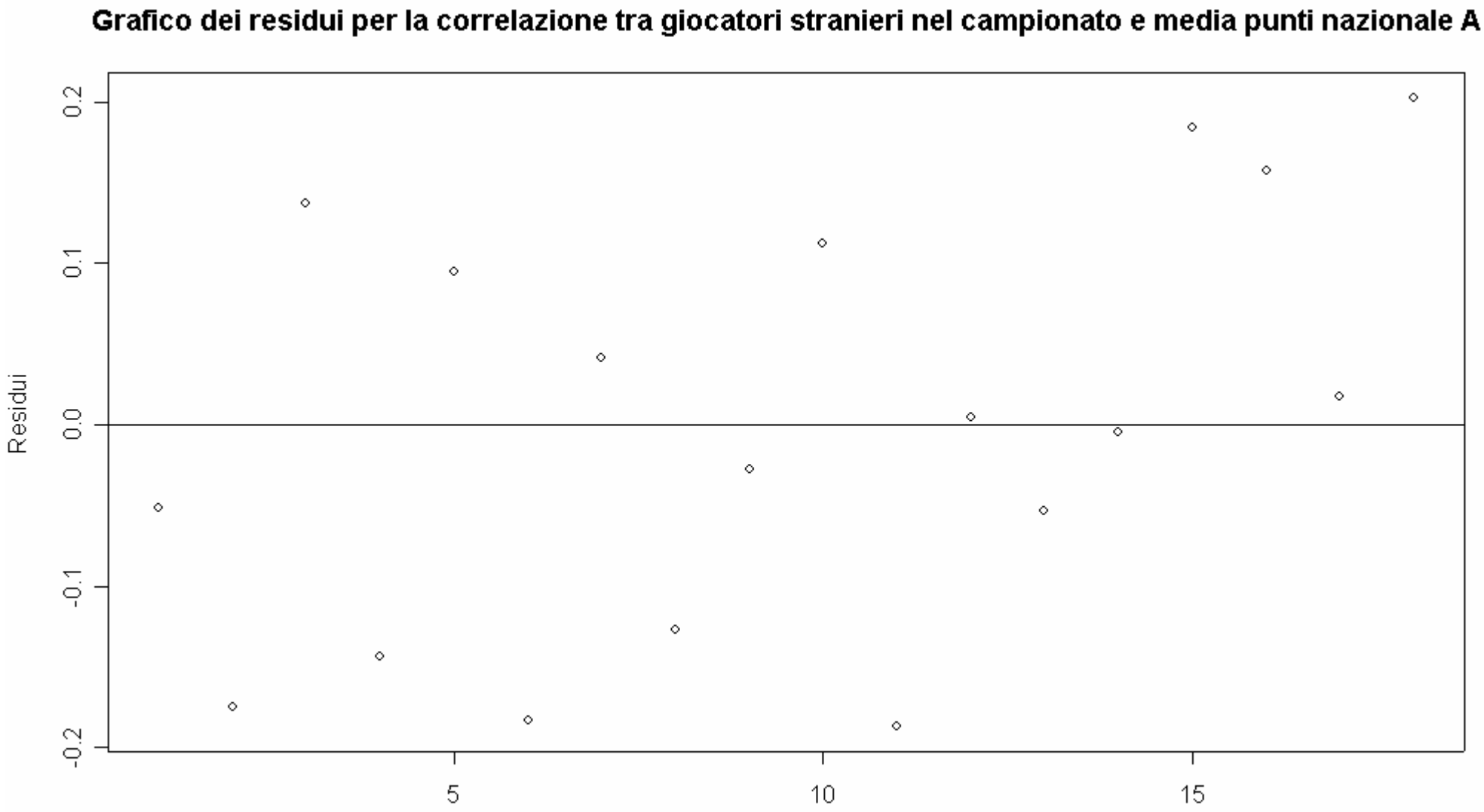


Grafico dei residui per la correlazione tra giocatori stranieri nel campionato e media punti nazionale A
Residui
0.2
0.1
0.0
-0.1
-0.2
5
10
15

Grafico nr. 7

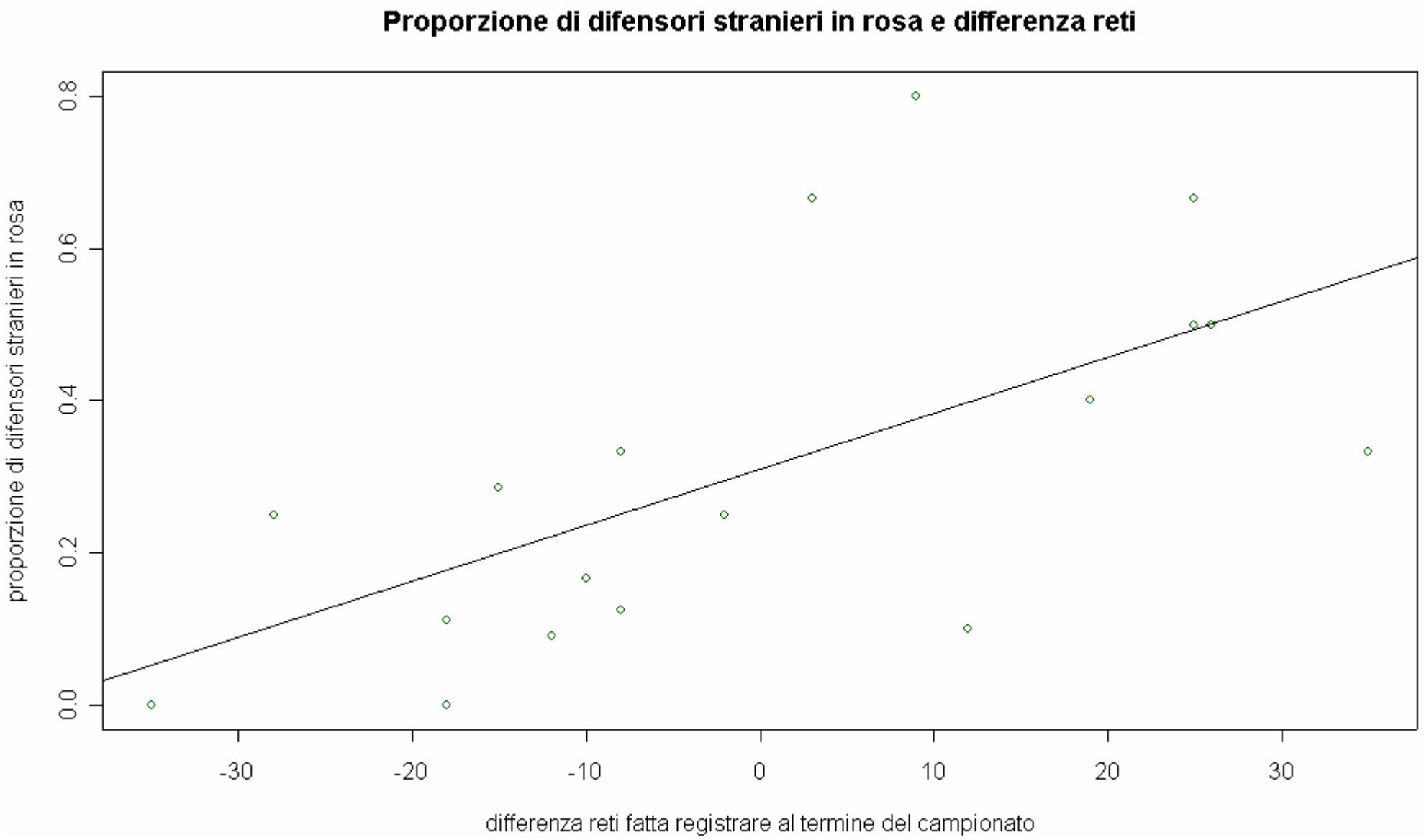
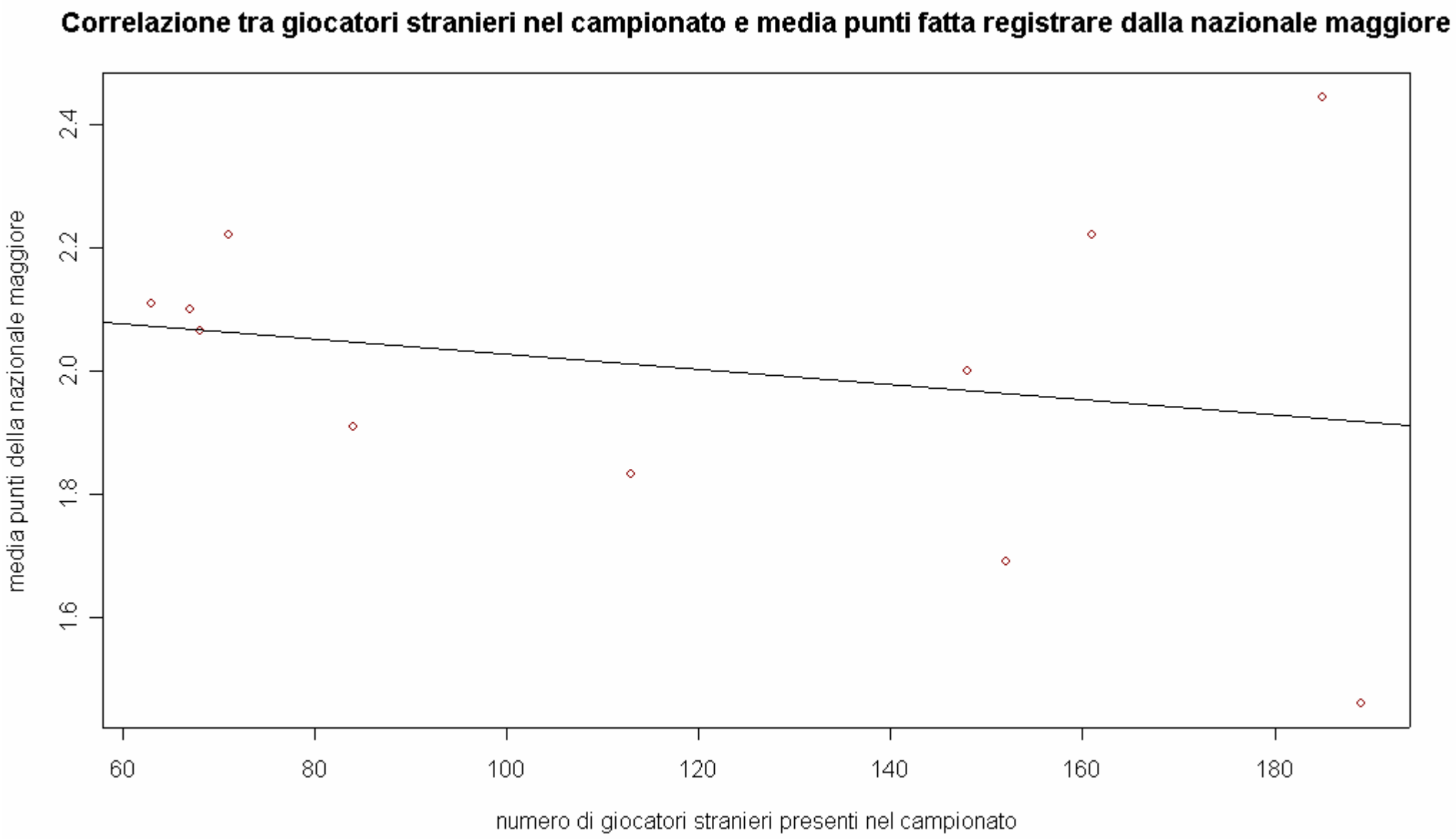

Proporzione di difensori stranieri in rosa e differenza reti
proporzione di difensori stranieri in rosa
0.0
0.2
0.4
0.6
0.8
-30
-20
-10
0
10
20
30
differenza reti fatta registrare al termine del campionato

Grafico nr. 8


Correlazione tra giocatori stranieri nel campionato e media punti fatta registrare dalla nazionale maggiore
media punti della nazionale maggiore
1.6
1.8
2.0
2.2
2.4
60
80
100
120
140
160
180
numero di giocatori stranieri presenti nel campionato