

POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale

Corso di Laurea Magistrale in Ingegneria Gestionale



TESI DI LAUREA MAGISTRALE

L'impatto del fattore campo sui risultati nel calcio nelle stagioni pre e post Covid-19.

Relatore:
Prof. Luigi Buzzacchi

Candidato:
Eugenio Sacco

Correlatore:
Dott. Francesco Luigi Milone

Anno accademico 2020/2021

INDICE

Introduzione.....	5
-------------------	---

CAPITOLO 1: LETTERATURA

1.1 L’Home Advantage nello sport.....	8
1.1.1 I fattori dell’Home Advantage.....	10
1.2 L’home Advantage nel calcio.....	17

CAPITOLO 2: IL CALCIO ITALIANO

2.1 La Serie A.....	21
2.2 La Serie B.....	22
2.3 Le infrastrutture nel mondo del calcio in Italia	24
2.3.1 Il caso Juventus Stadium	25
2.4 L’impatto del Covid-19 sul calcio italiano.....	27

CAPITOLO 3: DATASET

3.1 Raccolta dati	31
3.2 Struttura dataset	33
3.3 Statistiche descrittive	40
3.3.1 Statistiche descrittive sui tifosi.....	41
3.3.2 Statistiche descrittive sui valori economici delle squadre..	49
3.3.3 Statistiche descrittive sui goal.....	52
3.3.4 Statistiche descrittive sulle vittorie casalinghe.....	57

CAPITOLO 4: MODELLO CLARKE E NORMAN

4.1 Una prima analisi dell' Home Advantage	60
4.1.1 Modellare l'Home Advantage e l'abilità del team	60
4.2 Dati e risultati	65
4.2.1 Una prima analisi dell'impatto del Covid-19 sull' HA.....	71

CAPITOLO 5: MODELLO DI REGRESSIONE

5.1 L'analisi di regressione	73
5.1.1 Modello empirico di partenza	74
5.1.2 Risultati del modello empirico di partenza	75
5.1.1 Modello empirico che misura l'impatto del Covid-19.....	79
5.1.1 Risultati del modello che misura l'impatto del Covid-19...	79
Conclusioni.....	82
Bibliografia e sitografia.....	84
Appendice A.....	87
Appendice B.....	90

Ringraziamenti

Ormai giunto al termine di questo lungo percorso ritengo doveroso ringraziare tutte le persone che mi hanno supportato e sopportato durante questi anni.

Vorrei ringraziare il Prof. Luigi Buzzacchi e il Dott. Francesco Luigi Milone, i quali mi hanno seguito in questa ultima tappa del percorso, dandomi la possibilità di svolgere un lavoro che si addice alla mia più grande passione.

Poi, in primis grazie a Marghe, Umbe e Tato, che hanno sempre creduto in me e mi hanno sempre a loro modo spronato e stimolato ad andare avanti. Un grazie anche a tutta la mia numerosa famiglia e in particolare a mia nonna la quale mi ha sempre coccolato durante la vita universitaria e a zia Tilde, punto di riferimento assoluto per me in questi anni.

A Marzia, che con me ha intrapreso in parallelo uno stesso percorso e con cui reciprocamente ci siamo fatti forza, spronandoci e sostenendoci l'uno con l'altro.

Ai miei primi fantastici amici e compagni di Poli Lorenzo, Francesco e Daniele che mi hanno aiutato a non perdermi; e poi quelli incontrati in magistrale, con cui abbiamo costruito un grande team e che mi hanno aiutato ad alzare l'asticella e ad essere più ambizioso. Grazie Lorenzo, Simone, Catalin e Simone. Non dimenticheremo mai le giornate passate in via Busca 1. Tra i compagni di percorso del Poli, un grazie particolare va a Francesca con cui veramente per 5 anni ci siamo supportati e sopportati e con cui sono contento di condividere di nuovo insieme il raggiungimento di un altro traguardo raggiunto.

E poi, grazie ai miei due cugini Giacomo e Pietro, che mi hanno sostenuto e non mi hanno fatto mollare in un momento in cui davvero stavo perdendo le speranze.

Infine, un ultimo ringraziamento va a tutti i miei amici, che non posso citare

singolarmente, ma che in un modo o nell'altro in questi anni mi hanno dato tanto e mi hanno spinto a giungere a alla fine di questo percorso, aiutandomi a crederci. Fino alla fine.

Introduzione

Il mondo del calcio è un mondo che negli ultimi decenni sta crescendo in maniera esponenziale su molti aspetti, su tutti economici e mediatici. Infatti le principali società calcistiche professionistiche da semplice squadra di calcio si sono trasformate in vere e proprie aziende e, allo stesso modo le leghe nazionali.

Per questi motivi questo sport è diventato oggetto di studi su svariati fronti, in particolare studi economici e statistici.

Oltre alle società e ai giocatori che scendono in campo, un terzo attore che gioca un ruolo molto importante nello sport e in particolar modo nel calcio è il tifoso. A tal proposito, nell'Aprile del 2021 quando le 12 più importanti società europee fondarono la nuova competizione Superlega accessibile soltanto ad alcuni club, molti esponenti di spicco del mondo del calcio vedevano delegittimato il diritto di sognare di raggiungere traguardi importanti a tutti le altre società e soprattutto ai loro tifosi, e si alzò in coro lo slogan "Il calcio è dei tifosi". Tuttavia il calcio è diventato ormai economicamente insostenibile e per questo motivo si sta andando incontro ad una forte crisi, accelerata anche dalla pandemia, che colpisce soprattutto le società sportive di minor spessore.

È lecito domandarsi a questo punto quale ruolo ricoprano davvero i tifosi, e in particolare quelli che partecipano attivamente alle dinamiche sportive popolando gli stadi per sostenere la propria squadra; tra questi ci sono sicuramente il sostegno economico, facendo aumentare i ricavi delle società e di conseguenza la disponibilità ad investire, e il supporto che danno alla propria squadra durante gli incontri per mezzo del loro tifo.

Lo studio che viene presentato in questo elaborato nasce proprio dal desiderio

di provare a quantificare questo ultimo concetto, per misurare l'impatto che ha il tifo casalingo sui risultati degli incontri. Infatti, si pensa che questo sia il fattore principale di quello che viene definito Home Advantage (fattore campo), ovvero il vantaggio che una squadra ha quando disputa gli incontri in casa.

Il primo modello applicato in questo lavoro, descritto nel capitolo 4 prova a quantificare l'Home Advantage ed è la replica di un modello applicato in uno studio di dottorato inglese, il quale utilizza come mezzo di misura i goal segnati nell'intera stagione dalle squadre.

L'applicazione del modello originale eseguito dal Dott. George Foulds viene effettuata sulle prime quattro divisioni inglesi nelle stagioni che vanno dal 2001/2002 al 2011/2012; questo lavoro invece si concentra sulle stagioni dal 2017/2018 al 2020/2021 in Serie A e Serie B.

La scelta di applicare questo modello su entrambe le competizioni consente di effettuare un confronto molto interessante tra i risultati e osservare come in una competizione di livello inferiore il fattore campo sia molto più rilevante.

Inoltre, visto il periodo storico caratterizzato dalla pandemia dovuta dal Covid-19, il quale ha portato una situazione anomala nel mondo dello sport, ovvero la disputa degli incontri senza pubblico nelle stagioni 2019/2020 e 2020/2021, è interessante analizzare come questa situazione abbiamo impattato sull'Home Advantage. Si osserverà come in Serie B le squadre, a livello di risultati casalinghi, abbiano risentito molto di più dell'assenza del pubblico rispetto alla Serie A.

Una volta quantificato l'HA grazie all'applicazione del modello di Clarke e Normann, il lavoro proverà a definire quali fattori impattano di più sul fattore campo. A tale scopo sono stati costruiti due dataset distinti per la Serie A e Serie B, i quali raccolgono diverse informazioni, che verranno descritte nel capitolo 3, riguardo tutti

gli incontri disputati nelle stagioni suddette. A partire da questi dataset, sono state poi effettuate diverse analisi multivariate attraverso dei modelli di regressione di tipo Probit descritti nel capitolo 5, con lo scopo di definire quali fattori impattano maggiormente sui risultati casalinghi.

L'analisi di regressione mostrerà come in Serie B la probabilità di vittoria casalinga sarà maggiore rispetto alla Serie A (come già osservato dal modello precedente) e in più mostrerà come i fattori che impattano su tale probabilità siano per la Serie A solamente il valore delle rose delle squadre, mentre in Serie B risulta significativamente importante la presenza dei tifosi.

CAPITOLO 1

1.1 L'Home Advantage nello sport

Ci sono molte spiegazioni teoriche circa il vantaggio delle squadre che disputano gli incontri in casa, definito Home Advantage (HA). Queste includono teorie biologicamente fondate sulla territorialità e sui cambiamenti del ritmo circadiano, teorie psicologiche e sociali come la facilitazione sociale o il sostegno sociale percepito ed anche teorie sociologiche come l'integrazione rituale (Edwards, 1989). Al momento, non sussistono sufficienti prove in supporto di una teoria rispetto ad un'altra. Courneya e Carron (1992) hanno proposto un quadro che schematizza la ricerca sull'effetto Home Advantage sui risultati che consente l'inclusione di molti costrutti diversi (estrapolati) da teorie diverse. Questo quadro è mostrato nella Figura 1.1 di seguito.



Figura 1.1: Quadro di ricerca sulla posizione del gioco suggerito da Courneya e Carron (1992).

Il punto di partenza è la *game location*, ovvero la sede in cui si svolge l'evento. A tal proposito vi sono tre possibili alternative: casa, trasferta e campo neutro. I siti neutrali non godono di un vantaggio domestico e quindi non sono inclusi nella Figura 2.1. I fattori variabili che dipendono dalle diverse location (*game location factors*) possono essere raggruppati in quattro settori: *crowd*, *learning*, *travel* e *rules*. Le principali teorie derivano da queste quattro categorie.

Gli effetti del *crowd*, ovvero del pubblico presente allo stadio, possono dipendere da molti fattori, come per esempio dalla numerosità, dalla densità e dalla vicinanza degli spettatori. Questi effetti, che per chi gioca in casa risultano positivi e per chi gioca in trasferta negativi, riflettono il supporto sociale fornito dal pubblico di casa.

La familiarità con le caratteristiche del terreno di gioco proprie dello stadio di casa, consente al team di sviluppare fattori di *learning*, ovvero dei vantaggi dovuti, ad esempio, all'abitudine ad un tipo di superficie o alle dimensioni del campo da gioco.

La distanza tra le varie sedi delle squadre suggerisce la possibilità di avere effetti dovuti dal *travel*, ovvero al viaggio che la squadra in trasferta deve sostenere, che introducono fattori come l'affaticamento mentale e fisico, piuttosto che l'interruzione della routine.

In alcuni sport, le *rules*, ossia le regole, possono influenzare l'esito di un evento in favore della squadra di casa. Ad esempio, nel baseball, la squadra di casa batte per ultima in ogni inning.

È stato ipotizzato dagli studiosi che gli effetti del luogo di gioco, che portano a generare l'Home Advantage, si manifestino attraverso i diversi stati psicologici dei partecipanti. Infatti, i cosiddetti *Critical psychological states*, ovvero gli stati psicologici degli attori interessati possono essere influenzati dal luogo in cui si svolge un evento; questi possono manifestarsi diversamente nella stessa location per i giocatori, gli allenatori e gli arbitri. Gli stati cognitivi (es. ansia, fiducia, apprensione valutativa e aspettativa di risultato) e affettivi (es. eccitazione, rabbia, stress e orgoglio) possono portare a cambiamenti nel comportamento di uno o più dei gruppi coinvolti nella partita.

I *Critical behavioural states*, ovvero i comportamenti degli attori, sono

influenzati in presenza di un *home advantage* e quindi possono condizionare il risultato della partita.

Per i giocatori, questi stati possono includere il livello di sforzo utilizzato o la loro resistenza; gli allenatori possono mettere in atto decisioni strategiche e tattiche differenti e, infine, gli arbitri possono essere spinti a prendere decisioni soggettive.

Per chiudere questo quadro, bisogna effettuare la misura dei risultati delle prestazioni che può essere suddivisa in tre o più livelli. Una prima misura di performance può essere rappresentata dalla qualità fondamentale di una squadra, come ad esempio la media battuta nel cricket o la percentuale di tiri liberi nel basket. Le misure intermedie delle prestazioni vogliono riflettere il sistema di punteggio necessario, come ad esempio i goal segnati nel calcio o i set vinti nel volley e nel tennis. Infine, la terza misura rappresenterebbe il risultato complessivo di un incontro, come ad esempio vittoria o sconfitta o differenza di punti (Edwards, 1989; Irving e Goldstein, 1990; Schwartz e Barsky, 1977).

1.1.1 I fattori dell'Home Advantage

L'effetto della folla presente al match è il fattore più comunemente correlato al vantaggio casalingo ed è quello che, se non altro è ritenuto essere dominante da parte dei tifosi (Wolfson et al., 2005).

La dimensione della folla, la densità, l'intensità del supporto e la vicinanza ai giocatori richiedono tutte una considerazione (diversificata e) dettagliata (Nevill et al., 1996).

Tuttavia, l'estensione e le sfumature di questi effetti sono difficili da quantificare (Morris, 1981; Pollard, 1986); infatti, l'effetto di ciascun fattore non è chiaro. A tal proposito, sono stati osservati dei vantaggi casalinghi sia in presenza di un numero elevato di tifosi, sia in loro assenza, fornendo prove contrastanti e di difficile interpretazione (Courneya e Carron, 1992;

Pollard, 1986).

Dowie (1982), Pollard (1986) e Clarke e Norman (1995) dimostrato come ci fosse poca variazione del vantaggio casalingo sulle quattro divisioni della English Football League, nonostante le ovvie differenze nelle dimensioni del pubblico. Quando Neville et al. (1996) aggiunsero un'altra lega inglese inferiore e tre divisioni scozzesi al set precedentemente considerato, trovarono una relazione lineare tra la dimensione della folla e il vantaggio in casa. Sfortunatamente, la dimensione temporale del campione utilizzata in questa analisi era relativamente piccola, della durata di una sola stagione, ed hanno anche mostrato poca differenza tra le prime tre divisioni inglesi.

Non è chiaro se l'effetto dell'interazione del pubblico sia preminentemente quello di dare un vantaggio alla squadra di casa piuttosto che uno svantaggio alla squadra in trasferta. Inoltre, non è chiaro se i giocatori stessi siano direttamente interessati o se l'arbitro

veicoli l'influenza della folla (Boyko et al., 2007).

League	Level	Home advantage	Average Attendance
Premier	1	60.7%	31009
Division 1	2	61.2%	14160
Division 2	3	60.3%	6649
Division 3	4	61.9%	3757
Conference	5	56.7%	1484
Ryman Premier	6	56.7%	487
Ryman Division 1	7	54.1%	247
Ryman Division 2	8	53.3%	129
Ryman Division 3	9	55.1%	89

Figura 1.2: Vantaggio casalingo e presenze medie nelle nove divisioni calcistiche inglesi per le stagioni 1996/97 - 2001/02 (Pollard, 2006).

La figura 1.2 mostra il valore del vantaggio casalingo e della presenza media per le nove divisioni inglesi, classificati rispetto al loro livello di gioco. Il valore del vantaggio casalingo è dato dalla percentuale di tutte le partite casalinghe vinte rispetto al totale di quelle giocate in casa (Pollard, 1986).

Sebbene le dimensioni del pubblico varino notevolmente tra i vari livelli, si può notare

una minima differenza nel vantaggio casalingo tra i primi quattro, che hanno tutti un vantaggio casalingo di poco più del 60%. Nel livello subito successivo si ha un salto ad un vantaggio casalingo di circa il 55%, con squadre che sperimentano l'effetto anche con un pubblico costituito da un numero mediamente inferiore alle 100 persone. Ciò dimostra che un vantaggio casalingo si verifica in tutti i campionati di calcio inglesi e che quando la dimensione della folla aumenta al di sopra di circa 3.000 (individui) la misura in cui viene sperimentato il vantaggio in casa aumenta di circa il 5%.

Gli allenamenti e le partite ripetuti in casa generalmente si traducono in familiarità con le condizioni di gioco locali, che possono essere considerate un fattore di vantaggio in casa. Barnett e Hilditch (1993) hanno mostrato come le squadre di calcio inglesi, che giocano su campi e che utilizzano erba artificiale, ottengono un maggiore vantaggio in casa rispetto a quelle che si allenano su campi in erba tradizionale. Clarke e Norman (1995) hanno confermato questa scoperta, che ha portato al divieto di utilizzo di erba artificiale nei campionati appartenente alla lega inglese.

Tuttavia, Pollard (1986) ha prodotto prove riguardo agli effetti prodotti da un'altra caratteristica del campo di gioco, ovvero dalle sue dimensioni, dimostrando che l'uso di dimensioni anomale del campo non aumentava il vantaggio in casa normalmente atteso.

È stato inoltre dimostrato che spostare squadre e giocatori in nuovi campi di casa riduce l'Home Advantage, che può essere spiegato dalla riduzione o dalla perdita di familiarità con le condizioni del terreno di gioco (Pollard, 2002).

Ci sono numerosi fattori che possono contribuire alla familiarità delle sedi di casa. Ad esempio, gli effetti dei venti prevalenti e della luce del giorno dipendono dalla posizione di uno stadio, il design e la struttura possono fornire indicazioni visive ai giocatori di casa (specialmente nel basket dove i riferimenti per tirare sono molto importanti).

Pollard e Pollard (2005) hanno dimostrato la presenza di un calo significativo del vantaggio casalingo nella Football League of England dopo la sospensione di sette anni durante la Seconda guerra mondiale. Infatti, secondo loro, dopo aver ripreso il campionato nel 1946, le squadre partecipanti erano costituite da molti nuovi giocatori, che non avevano

familiarità con il loro ambiente di gioco locale.

Focalizzandosi ora sul fattore viaggio delle squadre in trasferta, la fatica subita da queste a causa dei viaggi è stata ampiamente analizzata rispetto alla misura in cui contribuisce all' Home Advantage (Pollard, 2006). I ridotti livelli di vantaggio casalingo mostrati nei derby locali potrebbero essere attribuiti alla mancanza di viaggi per i giocatori coinvolti, così come alla possibilità di avere maggiore supporto dei propri tifosi per la squadra in trasferta (Clarke e Norman, 1995).

Ulteriori prove sono state ottenute da Pollard (2006) dai risultati delle semifinali di Coppa dei Campioni e Champions League tra il 1960/61 e il 2003/04, dove le squadre in competizione hanno registrato un vantaggio medio casalingo del 71,7%. Infatti, questo valore elevato potrebbe essere spiegato dalle lunghe distanze percorse dalle squadre per disputare questi match.

Inoltre, Brown et al. (2002) e Clarke e Norman (1995) hanno osservato che i vantaggi in casa delle singole squadre, rispettivamente nelle competizioni internazionali e nei campionati inglesi, aumentavano in funzione della distanza geografica esistente tra le squadre.

In contrasto ai suddetti studi, quelli svolti da Pollard (1986) hanno prodotto un'argomentazione contraddittoria, non mostrando alcuna differenza nel vantaggio casalingo per le squadre distanti tra loro più o meno di 200 miglia (320 km). La Figura 2.3 riporta un grafico a barre che mostra il vantaggio casalingo nei diversi livelli del calcio inglese per il periodo 1888-2004. Si osserva che l'Home Advantage è diminuito nettamente dai primi anni rispetto all'ultimo decennio; in particolare si riscontra un drastico calo tra gli anni pre-Guerra e il dopo-Guerra, per poi mantenersi pressochè costante per i successivi tre decenni, per poi calare di nuovo nettamente tra gli anni Ottanta e Novanta. Questo potrebbe essere spiegato dalla riduzione dei tempi e dal comfort dei viaggi delle squadre ospiti, determinati dal potenziamento dei trasporti moderni (Pollard, 2006).

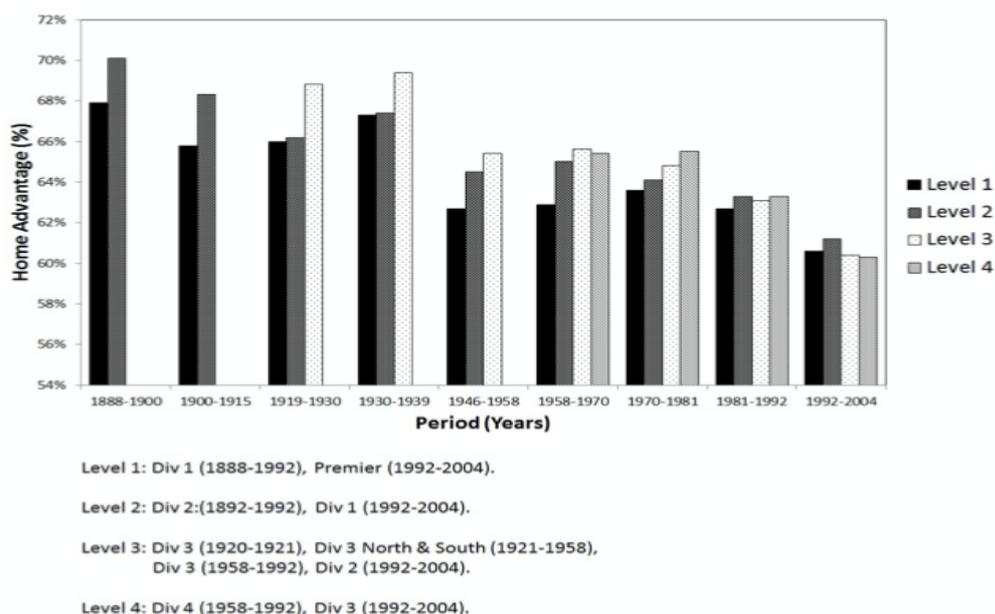


Figura 1.3: Vantaggio casalingo per quattro livelli del calcio inglese dal 1888 al 2004. (Pollard, 2006).

Per quanto riguarda, invece, il ruolo che le regole giocano nell' Home Advantage, questo non è chiaro in molti sport, poiché in molti di questi viene usato un regolamento imparziale rispetto alla game location. Pertanto, la ricerca su questo fattore è limitata a determinati sport.

Nel baseball l'opportunità di battere per ultimo può consentire di ottenere vantaggi distinti se l'incontro finisce agli extra inning, ovvero non nei minuti regolamentari di gioco. Tuttavia, Courneya e Carron (1990), usando le osservazioni sul baseball sui match disputati su campi neutri, dove sembrava non esserci alcun vantaggio fornito dall'ultima battuta, hanno dimostrato che non era come ipotizzato in precedenza.

Nell' hockey su ghiaccio, durante le interruzioni del gioco, le squadre che giocano in casa possono sostituire i giocatori soltanto dopo la squadra in trasferta, potendo così osservare e giocare strategie migliori (come nel caso del dilemma del prigioniero, dove un giocatore che può scegliere la propria mossa dopo aver osservata quella dell'altro ottiene un vantaggio). Inoltre, durante un face-off, ovvero uno

scontro tra due giocatori, il giocatore di casa ha per regola il diritto di appoggiare il suo bastone sul ghiaccio dopo l'ospite, aumentando così le probabilità di vincere lo scontro.

Queste regole si possono tradurre in un vantaggio casalingo in partite singole se la designazione casalinga viene assegnata ad una delle due squadre (IHHF, 2012).

Riguardo ai fattori psicologici e comportamentali, questi hanno impatto sia sui giocatori e allenatori che sugli arbitri degli incontri.

Lo stato psicologico di una squadra o di un individuo può essere influenzato dal luogo in cui si svolge la partita, che di conseguenza può influenzare l'esito di un evento competitivo.

Un esame della percezione del giocatore nei confronti del vantaggio di casa è stato effettuato da Jurkovic (1995). Alcuni giocatori di basket sono stati intervistati con domande che potessero consentire di conoscere il loro atteggiamento quando giocavano in casa e in trasferta. I risultati del sondaggio hanno mostrato che il 97% dei giocatori intervistati sentiva di aver giocato meglio quando era supportato da una folla rumorosa e attiva in casa e il 74% si sentiva allo stesso modo quando giocava in trasferta. Inoltre, il 47% riteneva che le proprie statistiche personali fossero migliorate giocando in casa e il 76% si sentiva più sicuro.

Gli effetti psicologici sono spesso il risultato di reazioni fisiologiche agli stimoli ambientali, come il meccanismo di attacco o fuga (Gleitman et al., 2010); infatti ormoni come il testosterone e l'adrenalina possono agire per alterare lo stato mentale di un giocatore. A tal proposito, sono state condotte molte ricerche sull'effetto dell'ormone steroideo testosterone, dimostrando così come le variazioni dell'aggressività degli animali maschi coincidono con le variazioni stagionali del testosterone (Lincoln et al., 1972). Monaghan e Glickman (1992) hanno mostrato

come l'aumento artificiale del testosterone incrementi l'aggressività e, inoltre, è stato dimostrato da Mazur e Booth (1998) che i livelli di testosterone, aumentando quando si è tenuti ad affrontare una sfida, incrementano la competitività e gli atteggiamenti dominanti negli esseri umani.

Varca (1980), invece, ha misurato l'aggressività degli giocatori per poter analizzare il comportamento della squadra in casa e in trasferta nelle partite di basket collegiali utilizzando statistiche riguardanti i falli, i recuperi, i rimbalzi e le stoppate. Queste misure sono state poi suddivise in due sottoinsiemi: misure funzionali (vantaggiose) e disfunzionali (svantaggiose) per ogni team. Dalla sua analisi, Varca osservò che le squadre che giocano in casa hanno mostrato un comportamento aggressivo più funzionale, stoppando i tiri e ottenendo più rimbalzi, mentre le squadre in trasferta hanno mostrato un comportamento più disfunzionale, commettendo un numero maggiore di falli.

Per quanto riguarda gli arbitri, invece, essi hanno il compito di fare rispettare le regole con autorità e a causa della natura di alcune di queste, devono essere prese decisioni soggettive. Per questo motivo sono state effettuate più indagini rispetto all'effetto di queste scelte sull' Home Advantage. Sono stati documentati vari studi che mostrano una propensione verso maggiori decisioni a favore delle squadre di casa o contro le squadre in trasferta (Lefebvre e Passer, 1974; Varca, 1980; Sumner e Mobley, 1981; Greer, 1983; Glamsner, 1990).

Va considerato, tuttavia, che una maggiore percentuale di decisioni soggettive prese a discapito della squadra ospite potrebbe non essere causa dell'Home Advantage, ma piuttosto una conseguenza (Sumner e Mobley, 1981). Infatti, le squadre che giocano in trasferta possono trascorrere più tempo in difesa ed essere più disfunzionalmente aggressivi.

Lehman e Reifman (1987) hanno ipotizzato che gli arbitri delle partite di basket professionistico fossero maggiormente propensi a favorire i fuoriclasse della squadra di casa, consentendo loro più libertà. I due studiosi pensavano che questo comportamento avesse origine dall'influenza esercitata dalla folla e credevano che la loro indagine fosse in grado di dimostrare quanto detto. Il set di dati utilizzato per testare questa ipotesi riguardava le partite che coinvolgevano i Los Angeles Lakers durante la stagione 1984/85. Lehman e Reifman hanno concluso che il loro ragionamento fosse corretto, poiché ai fuoriclasse della squadra di casa venivano assegnati meno falli rispetto alla squadra in trasferta; mentre nessuna differenza è stata trovata per i giocatori classificati come non fuoriclasse.

1.2 Home Advantage nel calcio

Il calcio è uno degli sport in cui è stato maggiormente documentato e riscontrato in letteratura l'Home Advantage (Jamieson 2010, Pollard & Pollard 2005); esistono però alcune differenze tra diversi paesi (Pollard 2006) e tra le leghe maschili e quelle femminili (Pollard & Gomez 2012). Sebbene il vantaggio di casa sia diffuso e ben noto, i suoi driver e le loro dimensioni sono ancora oggetto di un intenso dibattito.

Tuttavia, la letteratura concorda ampiamente sul fatto che il vantaggio casalingo non sia guidato da un'unica fonte, ma come detto in precedenza sia dovuto a molti fattori che si influenzano anche a vicenda.

Probabilmente non sono molti i fenomeni che attirano l'attenzione di tante discipline scientifiche come lo studio del vantaggio casalingo nel calcio. Considerando, come accennato in precedenza, che le condizioni psicologiche e

mediche includono differenze ormonali nei giocatori tra le partite in casa e quelle in trasferta (Bray et al. 2002, Neave & Wolfson 2003, Terry et al. 1998), la ricerca economica si è concentrata sulle aspettative dei risultati a seconda del luogo in cui si svolge una partita (Carmichael & Thomas 2005), e della pressione sociale esercitata dai tifosi (Dohmen 2008B, Dohmen & Sauermann 2016, Garicano et al. 2005).

Nella presente analisi non si terrà conto di alcuni fattori sopracitati, come il ruolo dell'arbitro e gli stati psicologici dei giocatori, in quanto non sono facilmente quantificabili e analizzabili.

Infatti, il lavoro si concentra sul ruolo svolto dal supporto della folla sui risultati sportivi, con un focus particolare all'impatto che l'assenza totale dei supporter allo stadio ha avuto sui risultati, dovuto al periodo storico in cui viviamo.

Nel calcio, principalmente, dove quasi tutte le competizioni contano di sfide andata-ritorno, ovvero incontri tra due squadre che si svolgono prima in casa di una e poi in casa dell'altra, entrambe hanno molto interesse ad ottenere un risultato positivo quando giocano tra le "mura amiche", consci del fatto che un match in trasferta può comportare sempre delle difficoltà in più. Da un punto di vista metaforico o da una prospettiva psicologica, le squadre in trasferta possono essere considerate "invasori", facendo sì che i giocatori della squadra di casa stimolino ulteriori forze di resistenza, misurate, ad esempio, attraverso livelli più elevati di testosterone nei giocatori di casa (Neave & Wolfson 2003, Carre et al. 2006). Pollard & Pollard (2005) cercano di affrontare questo problema mostrando come le differenze di vantaggio domestico tra squadre di diversi paesi possano, in una certa misura, essere spiegate dalla storia. Ad esempio, se c'è stata una guerra civile in uno stato, le rivalità sembrano essere più severe e l'effetto di territorialità più forte.

Staufenbiel et al. (2018) suggeriscono che la consapevolezza dei meccanismi

alla base del vantaggio casalingo sia meno presente nei giocatori più giovani, anche se questa possa essere parzialmente acquisita e insegnata nel tempo.

La familiarità con il luogo in cui si svolge un incontro, nel calcio, significa che le squadre di casa sono abituate alle condizioni e alle circostanze tipiche della loro location. Ciò include una migliore conoscenza delle condizioni meteorologiche e del (micro) clima locale durante la partita o fattori più specifici come la lunghezza del campo o la tipologia del manto erboso, che può variare a seconda della tipologia (sintetico o erba) e dell'altezza dell'erba.

Tutte queste condizioni implicano che lo stadio sia di grande importanza nel calcio e, infatti, Pollard (2002) osserva come le squadre di casa abbiano prestazioni peggiori nel periodo immediatamente successivo alla costruzione di un nuovo stadio, indicando come causa la riduzione della familiarità.

Tuttavia, nel caso di squadre appena promosse e che sono nuove in un campionato, non è stato osservato alcuno svantaggio per le loro squadre ospiti, anche se dovrebbero avere meno familiarità con lo stadio della nuova squadra (Clarke e Norman 1995).

Da sempre nel mondo del calcio uno dei fattori che si pensa influisca sui risultati è il suddetto *Home Advantage* (HA), cioè il vantaggio nel giocare una partita in casa, nel proprio stadio e di fronte ai propri tifosi.

Il modello formulato ha l'obiettivo di comprendere se le performance sportive di un club di calcio siano solo determinate da variabili legate al valore delle squadre oppure se anche i tifosi e il fattore campo abbiano un'influenza su di esse.

A tal fine è condotta un'analisi di regressione sui risultati ottenuti dalle società partecipanti al campionato di Serie A e serie B dalla stagione 2017-18 alla stagione 2020-21, scelte per poter ottenere un numero più consistente di dati e poter effettuare

un confronto tra queste.

I dati e le relative descrittive saranno presentate nei paragrafi successivi.

CAPITOLO 2

2.1 La Serie A

Il 1898 è considerato il punto di partenza della storia del calcio italiano in quanto è il l'anno in cui sono state disputate le prime partite ufficiali, anche se con una organizzazione ancora scarsa. La forma odierna, infatti, per quanto concerne caratteristiche e modalità, viene assunta nell'ottobre del 1929.

Prima di quell'anno infatti si parlava piuttosto di un semplice torneo organizzato su scala amatoriale fatto di gironi plurimi su base regionale, ad eliminazione diretta, al cui termine c'era in palio una coppa. Invece, nella stagione 1929-1930 si decise di creare un girone unico, detto all'italiana. Furono istituiti l'andata e il ritorno e fu ufficializzato il nome serie A per distinguerlo da altre competizioni italiane di minore importanza. Fin dall'inizio, la Juventus, l'Inter ed il Milan riuscirono ad affermarsi sulle rivali ed il predominio durò fino alla Seconda guerra mondiale, quando squadre come il Bologna, il Torino, la Fiorentina, la Lazio e la Roma riuscirono a contrastare l'egemonia dell'asse Milano-Torino. Attualmente il club più titolato d'Italia è la Juventus con 38 scudetti dal 1905 ad oggi.

Il 16 maggio 1946, nella sede della Fiorentina, alla presenza di 16 delegati di Serie A e 11 di Serie B viene costituita la Lega Nazionale Professionisti, fissandone la sede a Milano. Nella stagione '47-'48 venne disputato il più lungo campionato della storia. Gli anni successivi videro la riforma dei campionati di serie "A" e "B", riportandoli a 18 squadre.

Un altro step fondamentale avvenne nel 1966 quando la Lega Nazionale Professionisti in seguito alla disfatta ai Mondiali vietò l'ingaggio di giocatori e tecnici

di campionati esteri. Questo evento portò ad un declino della Serie A , aggravato agli inizi degli anni Ottanta dal primo grande scandalo di calcioscommesse.

Gli anni a cavallo tra Ottanta e Novanta invece segnarono la rinascita della Serie A, che nel 1989-90 raggiunse la prima posizione nel ranking continentale grazie alla vittoria delle tre competizioni confederali più importanti (Champions League, Coppa Uefa e Coppa delle Coppe).

Nel 1994-95 la Lega seguì l'esempio di alcuni altri campionati pionieri, introducendo un metodo diverso di calcolo dei punti per la classifica finale; infatti in precedenza ogni vittoria portava ad una squadra 2 punti, il pareggio 1 punto e la sconfitta 0 punti, mentre da quell'anno i 2 punti in casa di esito positivo furono sostituiti dagli ancora attuali 3 punti.

Il più recente cambiamento riguardo la struttura del campionato di serie A avvenne nella stagione 2005-2006 quando si portò il numero di squadre partecipanti a 20, mentre in precedenza prendevano parte alla competizione 18 squadre.

A partire dal 2005 il calcio italiano iniziò un lento declino dovuto al verificarsi di una serie di fattori, come lo scandalo relativo all'ambito delle scommesse e a Calciopoli (2006) e la precaria situazione finanziaria dei club che ha portato a diversi fallimenti societari (Parma nel 2010, Torino nel 2005).

Negli ultimi anni, per stare al passo con i principali campionati europei, molte società Serie A ha iniziato ad aumentare nettamente gli investimenti e ciò ha portato ad una ripresa, riportando molti campioni ad ambire a giocare nel campionato italiano.

2.2 La Serie B

I primi campionati italiani di calcio erano composti da poche squadre ed erano

prettamente elitari. Nel 1904 il torneo si allargò con la prima edizione della Seconda Categoria a cui partecipavano le formazioni della riserve delle società che disputavano la Prima Categoria, e quelle associazioni provinciali che erano da poco entrati nella Federazione.

Fino al 1912 alle squadre provinciali non bastava vincere questa competizione per poter essere ammessa alla Prima Categoria, ma da quell'anno venne introdotto il meccanismo promozione-retrocessione. Venne così trasformato il torneo di Seconda Categoria in Promozione, dove partecipavano le compagini provinciali, e venne creato il torneo un campionato dedicato alle Riserve.

Fu solo nel 1921 che si creò un vero e proprio campionato cadetto nazionale istituendo la Seconda Divisione, torneo in cui confluirono tutte quelle società rimaste tagliate dalla scrematura cui andò incontro la Prima Categoria ora ribattezzata come Prima Divisione. La nuova normativa introdotta prevedeva la suddivisione dell'Italia in 2 grosse aree geografiche gestite da due Leghe Nord e Sud. Solo le vincenti le qualificazioni regionali si incontravano in uno o più gironi di finale per la promozione alla categoria superiore.

La riforma epocale fu portata a compimento nel 1928: la creazione di un campionato cadetto radicalmente diverso, e cioè non più un torneo interregionale, bensì un girone unico nazionale esattamente identico a quello parimenti progettato per il campionato maggiore: nacque così, nel 1929, la Serie B della Divisione Nazionale, che vedeva iscritte 18 squadre.

Nel 1948-1949, dopo l'annullamento della competizione dovuta alla Guerra e ai problemi economici che ne scaturirono, finalmente la F.I.G.C. riuscì a ricreare un girone unico a 22 squadre, ridotte poi nuovamente a 18 nel 1952-1953. Nel 1958-1959 si decise un allargamento dell'organico a 20 società, stabilendo un format che, tranne

in un'occasione rimase inalterato per 35 anni. Furono le conseguenze del cosiddetto Caso Catania a sconvolgere nel 2003-2004 la consolidata tradizione e a portare il lotto delle iscritte al record assoluto di 24 squadre, poi scese a 22 l'anno successivo. Nel 2018-2019 a causa della mancata iscrizione di Avellino, Bari e Cesena, per l'impossibilità di ripescare squadre in tempi brevi (entro la presentazione dei calendari per la stagione 2018-2019), il campionato cadetto torna ad avere un numero dispari di squadre partecipanti, da 22 scende a 19. Su disposizione della FIGC nella stagione 2019-2020 il campionato è tornato ad avere 20 squadre.

2.3 Le infrastrutture nel mondo del calcio in Italia

Ogni anno, a partire dal 2011, la FIGC rende pubblica un report per analizzare tutti i dati relativi ai campionati professionistici nazionali e confrontarli con gli anni precedenti e i principali campionati europei.

In Italia gli stadi sono quasi tutti di proprietà del comune della città in cui è insediato e questa è una delle più grande pecche delle società calcistiche italiane, in quanto non permette ad esse di poter godere di un asset e di tutti i vantaggi economici che ne deriverebbe. In particolare, ad oggi, gli stadi di proprietà sono soltanto 5 sui 37 utilizzati tra serie A e serie B e sono: l'Allianz Stadium della Juventus, il Gewiss Stadium dell' Atalanta, la Dacia Arena dell' Udinese, il Benito Stirpe del Frosinone e il Mapei Stadium del Sassuolo. È interessante osservare come tali stadi abbiano il nome che fa riferimento ad una società che paga uno sponsor per averne diritto che entra direttamente nelle casse delle squadre. Tutti gli altri stadi in Italia, eccetto lo Stadio Olimpico che è di proprietà del CONI, sono come detto comunali e quindi le società sono costrette a pagare un affitto per usufruirne e i ricavi da stadio ovviamente

serviranno in parte per saldare questo costo.

Il primo stadio di proprietà è stato acquisito e costruito dalla Juventus nel 2011 (molto dopo rispetto il resto dell' Europa) e negli anni successivi le altre quattro società hanno seguito il modello; nei prossimi anni altre società si stanno muovendo verso tale direzione, come ad esempio Inter e Milan, i quali hanno diversi progetti per rivitalizzare lo storico San Siro di Milano.

Nei Top 5 campionati europei, esclusa la già citata Serie A, le realtà sono diverse, infatti la Premier League e la Bundesliga sono sempre state considerate dei modelli di riferimento e di avanguardismo per quanto riguarda gli impianti sportivi e a confermarlo sono il numero degli stadi di proprietà:

- Bundesliga: 16 su 18
- Premier League: 17 su 20
- La Liga: 11 su 20
- Ligue 1: 1 su 20 (Groupama Stadium – Olympique Lione)

Questi dati si riflettono sulla potenza economica dei vari campionati; infatti è riscontrato che la Premier League e la Bundesliga e la Liga siano ormai negli ultimi anni campionati seguiti con più interesse rispetto agli altri e sono competizioni dove le società investono cifre nettamente superiori.

2.3.1 Il caso Juventus Stadium

A partire dalla stagione 2011, come detto in precedenza la Juventus ha iniziato a disputare i suoi incontri casalinghi a quello che era chiamato Juventus Stadium, costruito sul suolo dove era presente il vecchio stadio Delle Alpi di proprietà del comune di Torino.

Il progetto iniziato nel 2003 e concluso 8 anni dopo ha richiesto alla società torinese un esborso di circa 155 milioni di euro. In quel periodo, la Juventus si è spostata a disputare i match casalinghi nell'altro stadio comunale Olimpico di Torino, ora chiamato Grande Torino.

In molti in queste stagioni hanno indicato lo Stadium, diventato nel frattempo Allianz Stadium, come una delle armi principali che hanno portato i bianconeri al successo.

Nell'ultima stagione disputata all'Olimpico, i ricavi da stadio hanno fatto incassare alla Juventus 11.9 milioni di euro, compresi i ricavi dalle attività commerciali legate allo stadio.

I ricavi da stadio dal 2011 ad oggi hanno superato il mezzo miliardo di euro.

Considerato quindi il punto di partenza prima della fondazione dello Juventus Stadium, l'aumento fino alla stagione 2018/2019 è stato pari a quasi il 600%, con un picco raggiunto in questa ultima stagione dove si è arrivati a quota 83 milioni di euro, suddivisi in 62 derivanti dagli incontri e 21 legati all'extra campo.

Ovviamente l'impatto del Covid-19 ha fatto crollare questi numeri, nonostante fino a metà stagione 2019/2020 i numeri fossero ancora in crescita; infatti con lo Stadium aperto fino a metà febbraio 2019 si era arrivati a 54 milioni di euro di ricavi da stadio.

L' Allianz Stadium ha garantito alla Juventus circa il 13% dei ricavi complessivi (plusvalenze comprese): si va dal massimo del 14.9% del 2012/13 ad un minimo dell'11.5% nel 2016/17, con una percentuale stabile nelle ultime due stagioni intorno al 13%. Questi numeri sono il doppio rispetto all'ultima stagione di stadio Olimpico, quando i ricavi da stadio erano il 6.9% sul fatturato.

Tuttavia, nonostante quelli della Juventus sembrano essere numeri

impressionanti, questi sono lontani da quelli delle big europee. Infatti nel 2018/2019 cinque squadre (Real Madrid, Barcellona, Manchester United, PSG e Arsenal) hanno sfondato il tetto dei 100 milioni di euro di ricavi da stadio. Una causa di questi numeri potrebbe essere anche la capienza degli stadi di queste squadre, che il doppio se non di più rispetto ai 41 milioni di spettatori che può ospitare l'Allianz Stadium di Torino.

I numeri descritti in precedenza possono essere osservati nella figura 2.3.

STAGIONE	ABBONATI	MEDIA SPETTATORI	RICAVI DA STADIO	RICAVI NO MATCHDAY	TOTALE RICAVI STADIO	FATTURATO	IMPATTO STADIO SU RICAVI
2010/11 (*)	14.290	21.966	10,1	1,8	11,9	172,1	6,9%
2011/12	24.531	37.044	27,4	4,2	31,6	213,8	14,8%
2012/13	27.400	36.931	34,9	7,4	42,3	283,8	14,9%
2013/14	28.000	37.635	38,2	8,2	46,4	315,8	14,7%
2014/15	28.000	38.687	41,2	9,5	50,7	348,2	14,6%
2015/16	28.000	38.554	39,3	10,0	49,3	387,9	12,7%
2016/17	29.300	39.673	49,2	15,7	64,9	562,7	11,5%
2017/18	29.300	39.106	49,6	15,4	65,0	504,7	12,9%
2018/19	29.300	39.561	62,5	20,8	83,4	621,5	13,4%
2019/20	27.700	39.777	39,4	14,7	54,0	573,4	9,4%
2020/21**	0	0	5,2	2,0	7,1	258,2	2,8%
TOTALE STADIUM			397,0	109,7	506,7	4.242,0	11,9%

Dati in milioni di euro
 Elaborazione CF su dati Fc Juventus Spa
 Ricavi stadio: abbonamenti+biglietteria+servizi aggiuntivi gare
 (*) ultima stagione all'Olimpico
 (**) dati al 31/12/2020
 In rosso: stagione con impatto Covid

Figura 2.1: Riepilogo dati riguardanti l'Allianz Stadium (fonte Calcioefinanza)

Questo excursus sulla Juventus vuole fare vedere come un investimento così importante di una società può aiutare ad ottenere grandi risultati. Infatti con l'aumento di ricavi si sono potuti effettuare molti investimenti in ambito sportivo e i risultati sportivi sono stati molto positivi con la vittoria di 9 campionati consecutivi.

2.4 L'impatto del Covid-19 sul calcio italiano

8 marzo 2020, Parma-Spal: recupero della 25esima giornata di serie A è l'inizio dell'era calcio al tempo della pandemia, con partite disputate a porte chiuse e continui rinvii di incontri a causa del contagio di membri delle squadre.

Infatti il Dpcm emanato il 4 Marzo 2020 dichiara in tutti i comuni delle regioni Emilia Romagna, Friuli Venezia Giulia, Lombardia, Veneto, Piemonte e Liguria la sospensione di eventi e competizioni sportivi in luoghi pubblici o privati e consente

solo le partite “all’interno di impianti sportivi utilizzati a porte chiuse”. E fu così che la ventiseiesima giornata di Serie A si disputò con 6 partite su 10 a porte chiuse.

A seguito di questa giornata fu emanato un nuovo Dpcm del 9 Marzo 2020 che dichiarò:

- la sospensione degli eventi e delle competizioni sportive di ogni ordine e disciplina, in luoghi pubblici o privati. Restava consentito lo svolgimento dei predetti eventi e competizioni, nonché delle sedute di allenamento degli atleti professionisti e atleti di categoria assoluta che partecipano ai giochi olimpici o a manifestazioni nazionali o internazionali, all'interno di impianti sportivi utilizzati a porte chiuse, ovvero all'aperto senza la presenza di pubblico.

Il 14 Marzo Dopo un breve tira e molla, tra zone rosse e arancioni, partite rinviate e a porte chiuse, la Lega di Serie A si è dovuta arrendere alla realtà accettando lo stop del campionato imposto dal Governo e con il Dpcm del 1 Aprile 2020 furono sospesi anche gli allenamenti individuali degli atleti professionisti.

Non era mai successo, se non nel corso della Prima Guerra mondiale, un'interruzione a stagione in corso.

Da quel momento iniziarono a presentarsi la possibilità di non riuscire a portare a termine la stagione ormai avviata e si iniziò a pensare a delle soluzioni alternative come i playoff e playout per stabilire il vincitore e le retrocessioni o il congelamento della classifica.

Un nodo critico era quello relativo ai contratti dei calciatori, che prevedono la scadenza al 30 giugno. Infatti, andando verso uno scenario di proseguimento oltre quella data, bisognava costituire norme ad hoc per evitare dei problemi.

Ma dopo un tira e molla con il Comitato Tecnico Scientifico , il Ministero della Salute, quello dello Sport, alla fine il calcio riesce a ripartire. Infatti, il Dpcm dell'11

giugno 2020 aveva disposto che:

- a decorrere dal 12 giugno 2020, gli eventi e le competizioni sportive riconosciuti di interesse nazionale dal CONI, dal CIP e dalle rispettive Federazioni, ovvero organizzati da organismi sportivi internazionali, erano consentiti a porte chiuse ovvero all'aperto senza la presenza di pubblico, nel rispetto dei protocolli emanati dalle rispettive Federazioni sportive nazionali, discipline sportive associate ed enti di promozione sportiva; anche le sessioni di allenamento degli atleti, professionisti e non professionisti, degli sport individuali e di squadra, erano consentite a porte chiuse, nel rispetto dei medesimi protocolli;

Il 20 giugno, così, i campionati di Serie A e Serie B riprendono, sempre a porte chiuse e con un rigido protocollo anti covid-19 da rispettare.

Tra le novità, fu introdotta la possibilità di fare cinque sostituzioni. Questo perché il calendario era molto fitto con 124 partite concentrate in 43 giorni (dal 20 giugno al 2 agosto).

Il 19 settembre comincia la stagione 2020/21; grazie al Dpcm del 7 agosto 2020, era consentita la partecipazione del pubblico per un numero massimo di 1.000 spettatori per gli stadi all'aperto e di 200 spettatori per impianti sportivi al chiuso. Tuttavia, dal 3 novembre quando è stato instaurato il sistema a colori delle regioni in base alla fascia di rischio, nelle zone rosse e arancio fu reso nuovamente obbligatorio la disputa a porte chiuse delle partite.

Fu così che per tutta la stagione 2020/2021 si alternarono partite con 0 o 1000 spettatori, e anche per questo motivo nell'analisi che verrà effettuata successivamente si considerano a porte chiuse gli incontri con 1000 tifosi.

Solamente con il Decreto Legge del 1 giugno 2021 furono approvate le nuove "Linee Guida per l'attività sportiva di base e l'attività motoria in genere" che

consentivano in zona gialla la presenza del pubblico con posti a sedere preassegnati e a condizione che fosse assicurato il rispetto della distanza interpersonale di almeno un metro sia per gli spettatori che non fossero abitualmente conviventi, sia per il personale. La capienza consentita non poteva essere superiore al 25% di quella massima autorizzata e, comunque, il numero massimo di spettatori non poteva essere superiore a 1.000 per impianti all'aperto e a 500 per impianti al chiuso.

Dal 23 luglio 2021, è stato anzitutto previsto un limite di capienza per la partecipazione del pubblico agli eventi sportivi di qualsiasi genere anche nelle zone bianche.

In particolare, è stato previsto che:

- nelle zone bianche la capienza consentita non poteva essere superiore al 50% di quella massima autorizzata se l'evento si svolgeva all'aperto. Per gli eventi svolti al chiuso, la capienza consentita non poteva essere superiore al 25% dal 23 luglio al 6 agosto 2021, aumentata al 35% dal 7 agosto 2021;
- nelle zone gialle, la capienza consentita non poteva essere superiore al 25% di quella massima autorizzata, con un limite di 2.500 spettatori per impianti all'aperto e di 1.000 per impianti al chiuso.

Durante questo periodo drammatico, le società sportive sono state costrette a subire un crollo del fatturato dovuto ai mancati incassi da stadio e ad adattarsi ai vari protocolli sanitari che hanno richiesto molti investimenti.

I risultati dei campionati disputati in questo anno e mezzo possono essere stati influenzati da tutti questi fattori, tra partite disputate ogni 3 giorni, squadre dimezzate per diffusione del Covid-19, assenza del pubblico allo stadio e crollo degli investimenti in ambito sportivo.

CAPITOLO 3

3.1 Raccolta dati

Realizzata una panoramica generale del in questo capitolo verranno descritti i dati raccolti per costruire i vari modelli. Per l'analisi sono stati individuati, su diversi siti web, i dati delle società calcistiche che almeno in una stagione dal 2017-18 al 2020-21 hanno preso parte al campionato di Serie A o Serie B e le informazioni sui singoli match disputati nelle suddette stagioni (risultato, quote sportive, giornata, numero spettatori, ecc.).

Il database di riferimento è stato realizzato attraverso informazioni ottenute principalmente dalle seguenti fonti:

- Transfermarkt;
- Stadiopostcards;
- Football-data;
- Wikipedia;
- Distanza

In particolare, Transfermarkt è diventato negli ultimi anni il principale sito online di informazioni calcistiche ed in questa ricerca rappresenta una delle centrali fonti di raccolta dati. Sul sito, infatti, sono rese disponibili informazioni che riguardano campionati e coppe di tutto il mondo, statistiche di giocatori, società calcistiche ed allenatori.

Transfermarkt consente di reperire dati in continuo aggiornamento grazie al contributo di più di trenta addetti ai lavori. Tali rapporti riguardano non solo gli aspetti più classici del calcio come le classifiche delle varie competizioni o le composizioni

delle rose dei club, ma anche dati più di dettaglio come il valore di mercato dei vari calciatori o delle rose, il numero di cartellini per ciascun calciatore, il numero di partite con porte inviolate di ogni squadra, ecc.

Il dato relativo al valore dei calciatori e quindi delle rose di ogni squadra non è un valore ufficiale della UEFA, ma si tratta di un valore ottenuto dalla piattaforma attraverso l'utilizzo di statistiche e calcoli parametrici dettagliati. In particolare, sfruttando le performance sportive, le quotazioni attuali, i valori di trasferimento ed il giudizio dei vari membri iscritti al sito, alcuni esperti determinano questi valori. Tale meccanismo prende il nome di community market value che da una parte consente ai membri di esprimere la propria opinione e dall'altra garantisce che la stima finale venga comunque realizzata da esperti. Come sostenuto da Herm et al. (2014), il community market value è uno degli stimatori più accurati del vero valore del calciatore e quindi delle rose.

Stadiopostcards è un sito online che dal 2001 ogni settimana offre un completo riepilogo, club per club, partita per partita, sulle presenze degli spettatori negli stadi italiani di Serie A (dal 1963/1964), B & C. Dunque sono stati reperiti da questa pagina i dati relativi al numero di spettatori, lo stadio in cui è stato disputato ogni match e eventuali informazioni circa il motivo della decisione di disputare l'incontro a porte chiuse, ovvero senza la presenza del pubblico. Inoltre, per ogni stagione di ogni competizione, sono riportati i dati cumulativi e riassuntivi per ogni squadra, come ad esempio il numero medio di tifosi presenti ed il numero complessivo.

La terza fonte di informazioni utilizzata è Football-data, sito online inglese ricco di dati su ogni match di molte competizioni calcistiche a livello mondiale; infatti si possono trovare i file di dati CSV per ogni stagione riportanti le quote degli incontri e i risultati parziali e finali di essi. Le quote sportive sono prese da diversi bookmakers

di riferimento e possono essere utili in quanto la quota nelle scommesse sportive corrisponde alla probabilità di un determinato esito di una scommessa ed è determinata dai bookmakers prima di ogni evento.

Un'altra fonte di dati è Wikipedia, famosa piattaforma online che contiene informazioni di ogni genere. Tale risorsa è stata utilizzata per reperire un cospicuo numero di informazioni utili, come la classifica generale ed i principali eventi che si sono verificati, ma anche la lista di tutti gli stadi presenti in Italia, le città in cui essi si trovano e la capienza massima disponibile.

Infine, l'ultima fonte di dati è Distanza, sito online che permette di calcolare la distanza in linea d'aria o in auto tra due città; perciò è stato utilizzato per definire i chilometri percorsi da ogni squadra per andare a giocare la partita in trasferta.

3.2 Struttura dataset

Prima di presentare tutte le informazioni a disposizione è opportuno sottolineare alcuni aspetti essenziali ai fini dello studio.

La Serie A è una competizione a cui partecipano ogni stagione 20 squadre e dunque in totale sono disputate 38 giornate e ogni squadra disputa metà degli incontri (19) in casa e l'altra metà in trasferta.

La Serie B invece nel corso degli anni ha subito diverse modifiche: la prima stagione presa in considerazione (2017-18) ha 22 squadre partecipanti e sono state disputate 42 giornate, con metà dei match giocati in casa e metà in trasferta; la stagione 2018-19 ha 19 squadre partecipanti a causa di alcuni tagli di squadre inadempienti e quindi le giornate in totale sono 38, ma ogni squadra disputa 36 incontri, metà in casa e metà in trasferta, e 2 turni di riposo; le ultime due stagioni di interesse (2019-20,

2020-21), a seguito della riforma della Serie B voluta dalla Lega Calcio, ha 20 squadre partecipanti e dunque in totale sono disputate 38 giornate e ogni squadra disputa metà degli incontri (19) in casa e l'altra metà in trasferta.

Il *database* utilizzato è un file Excel costituito da due fogli: il primo per la Serie A e il secondo per la serie B.

Il foglio della Serie A contiene 1520 righe che rappresentano le osservazioni e N colonne che invece indicano gli attributi dei vari record; quello della serie B contiene 1564 righe e N colonne . Come sarà approfondito successivamente tale database è stato “pulito” di alcune osservazioni al fine di avere dei dati affidabili per l'analisi effettuata nei capitoli successivi.

Nelle figure sottostanti è mostrata la struttura di alcune righe dei database costruito relativo alla serie A; allo stesso modo è strutturato quello relativo alla serie B.

Le prime quattro colonne, osservabili in Figura 3.1 riportano informazioni riguardo alla partita corrispondente, ovvero la competizione (serie A o serie B), la stagione, la data e l'orario di disputa dell'incontro l'orario e la giornata di campionato di appartenenza.

INFORMAZIONI PARTITA				
CAMPIONATO	STAGIONE	DATA	ORA	GIORNATA
SERIE A	2017/2018	27/11/17	20:45	14
SERIE A	2017/2018	22/10/17	15:00	9
SERIE A	2017/2018	30/12/17	15:00	19
SERIE A	2017/2018	04/02/18	15:00	23
SERIE A	2017/2018	20/09/17	20:45	5
SERIE A	2017/2018	18/02/18	18:00	25
SERIE A	2017/2018	29/04/18	15:00	35
SERIE A	2017/2018	14/04/18	20:45	32
SERIE A	2017/2018	01/10/17	20:45	7
SERIE A	2017/2018	17/12/17	20:45	17
SERIE A	2017/2018	13/05/18	18:00	37

Figura 3.1: parte del dataset su informazioni partita

Successivamente sono riportate le informazioni riguardanti le squadre che di ogni incontro, la squadra di casa e la squadra ospite. Per ognuna di esse è riportata l'informazione del nome e dell'ID che gli è stato associato (in ordine crescente rispetto all'ordine alfabetico per ogni competizione).

Poi per ogni squadra sono stati riportati i dati relativi agli anni di esperienza nella competizione, ovvero il numero di anni consecutivi di partecipazione della squadra in Serie A e , quindi, con la variabile booleana *NEOPROMOSSA* si indica se una squadra proviene dalla Serie B o ha alle spalle anni di esperienza in serie A. Pertanto se la variabile anni di esperienza assume il valore 0, la variabile boolean sarà vero (uguale a 1).

L' ultima variabile riporta il valore di mercato della rosa della squadra in milioni di euro ed è il dato estrapolato da Trasfermarket riferito all'inizio di ogni stagione. Avendo questi valori per la squadra di casa e ospite si è costruita la variabile che le correla per ogni partita, ovvero la differenza di valore tra la prima e la seconda, che quindi è maggiore di 0 se la squadra casalinga ha una rosa più strutturata e minore di 0 nel caso contrario.

Questa parte del dataset è riportata in Figura 3.2 e Figura 3.3.

SQUADRA CASA				
ID SQUADRA CASA	SQUADRA CASA	NEOPROMOSSA (1-SI, 0-NO)	ANNI ESPERIENZA	VALORE MERCATO SQUADRA CASA (in milioni di euro)
1	Atalanta	0	6	126,20
3	Bologna	0	2	70,75
7	Crotone	0	1	27,40
12	Inter	0	85	319,70
13	Juventus	0	10	496,00
14	Lazio	0	29	206,90
20	Sampdoria	0	5	100,60
21	Sassuolo	0	4	102,95
25	Udinese	0	22	81,30
26	Verona	1	0	36,43

Figura 3.2: parte dataset su squadra casa

SQUADRA OSPITE						
ID SQUADRA OSPITE	SQUADRA OSPITE	NEOPROMOSSA OSPITE (1-SI, 0-NO)	ANNI ESPERIENZA OSPITE	VALORE MERCATO SQUADRA OSPITE (in milioni di euro)	DIFFERENZA VALORE MERCATO	
19	Roma	0	64	323,43		-197,23
24	Torino	0	5	122,00		-51,25
16	Milan	0	34	350,05		-322,85
9	Florentina	0	13	136,65		183,05
5	Cagliari	0	1	52,55		443,45
22	Spal	1	0	32,28		174,62
2	Benevento	1	0	32,95		67,65
11	Genoa	0	10	93,20		9,75
6	Chievo	0	9	43,28		38,02
17	Napoli	0	10	371,58		-335,15

Figura 3.3: parte dataset su squadra ospite

Nella parte successiva del dataset, riportata nella Figura 3.4, si riassume l'esito del match sotto osservazione. A tale scopo, si inseriscono le variabili relative ai goal segnati dalla squadra di casa e quelli segnati dalla squadra in trasferta; da queste due si genera la differenza reti (goal squadra casa – goal squadra ospite) che determina così il risultato dell'incontro. Questo è descritto nella variabile *Risultato* che può essere H (home) in caso di vittoria casalinga, D (draw) in caso di pareggio e A (away) in caso di vittoria degli ospiti. Quindi le due colonne successive contano i punti ottenuti dalle due squadre nell'incontro in esame, ovviamente 3 punti per la squadra vincente e 0 per quella sconfitta e 1 punto ad entrambe in caso di parità.

RISULTATI PARTITA						
GOAL SQUADRA CASA	GOL SQUADRA OSPITE	DIFFERENZA GOL	RISULTATO	PUNTI SQUADRA CASA	PUNTI SQUADRA OSPITE	
0	1	-1	A	0	3	
1	1	0	D	1	1	
0	3	-3	A	0	3	
3	0	3	H	3	0	
3	0	3	H	3	0	
0	0	0	D	1	1	
2	1	1	H	3	0	
0	0	0	D	1	1	
1	2	-1	A	0	3	
1	3	-2	A	0	3	

Figura 3.4: parte dataset su risultati partite

Le tre colonne che seguono (Figura 3.5) riassumono il risultato alla fine del primo tempo per ogni incontro e, come nella parte appena descritta, lo si fa contando i gol fatti dalla squadra di casa e la squadra ospite e quindi il risultato che ne consegue (H, D o A).

RISULTATI PRIMO TEMPO		
GOL SQUADRA CASA HT	GOL SQUADRA OSPITE HT	RISULTATO HT
0		1 A
1		1 D
0		3 A
2		0 H
2		0 H
0		0 D
1		1 D
0		0 D
1		1 D
0		2 A

Figura 3.5: parte dataset su risultati primi tempi

La parte del dataset riportata in Figura 3.6 riporta le informazioni dei match secondo i bookmakers; in particolare per ogni incontro sono state prese le quote delle scommesse riguardanti l'esito 1, X, 2. Per avere un valore più dettagliato possibile, sono state prese in considerazione diversi bookmakers quali Bet365, Bet&Win, Interwetten, Pinnacle, William Hill e VC Bet. Da queste si sono ricavate i valori medi delle quote relative all'esito 1, X e 2 e inseriti nelle variabili AvgH, AvgD e AvgA; queste sono state utilizzate per definire quale delle due squadre era favorita per i bookmakers, utilizzando la variabile binaria Squadra Favorita, che assume valore 1 nel caso di squadra casa favorita ($AvgH > AvgA$) e 0 nel caso opposto ($AvgH < AvgA$).

QUOTE DELLA PARTITA																				SQUADRA FAVORITA (1-CASA, 0-TRASFERITA)	
B365H	B365D	B365A	BW1	BWD	BWA	BW1	BWD	BWA	P1H	P1D	P1A	WH1	WHD	WHA	VCH	VCD	VCA	AvgH	AvgD	AvgA	
3.20	3.30	2.30	3.20	3.40	2.25	3.10	3.30	2.20	3.17	3.58	2.36	3.10	3.40	2.25	3.20	3.40	2.30	3.16	3.39	2.28	0
3.10	3.30	2.38	3.20	3.30	2.30	2.75	3.30	2.40	3.29	3.42	2.36	3.10	3.30	2.30	3.25	3.20	2.38	3.12	3.30	2.35	0
6.00	4.00	1.57	5.50	4.10	1.60	5.60	3.90	1.55	6.04	4.20	1.62	6.00	3.75	1.60	5.50	4.00	1.62	5.77	3.99	1.59	0
1.57	4.00	6.00	1.57	4.20	5.75	1.60	3.70	5.40	1.57	4.42	6.33	1.53	4.00	6.50	1.55	4.20	6.50	1.57	4.09	6.08	1
1.14	8.00	19.00	1.16	7.50	17.00	1.15	7.00	15.00	1.16	8.29	24.44	1.14	7.50	21.00	1.14	8.00	21.00	1.15	7.72	19.57	1

Figura 3.6: parte dataset sulle quote delle scommesse

In Figura 3.7 è riportata la sezione del dataset riferito alle informazioni sull'affluenza degli spettatori allo stadio. In particolare le prime due variabili numeriche sono riferite al numero di persone presenti allo stadio per l'incontro e a quella che viene definita densità del pubblico, ovvero il numero di spettatori presenti rispetto alla capienza massima dello stadio in termini percentuali. Quest'ultima variabile è un modo per riassumere e descrivere il calore e la partecipazione dei tifosi, in quanto più la densità è maggiore più lo l'ambiente diventa ostile per la squadra

ospite.

La terza variabile in questa sezione è di tipo binario e assume il valore 1 nel caso in cui lo stadio in cui si disputa l'incontro è alla prima stagione dalla fondazione e 0 se era già presente nelle stagioni precedenti.

Le variabili successive sono variabili descrittive che spiegano se l'incontro si è disputato a porte aperte (regolare) o a porte chiuse e in questo ultimo caso il motivo di tale condizione che può essere per motivi disciplinari nelle prime due stagioni e causa Covid-19 nelle stagioni seguenti. Si ricorda, come già detto in precedenza, che nelle stagioni 2019/2020 e 2020/2021 si è considerato a porte chiuse anche i match disputatosi con capienza massima 1000 persone.

INFORMAZIONI SPETTATORI PARTITA						
NUMERO SPETTATORI	DENSITA' STADIO	STADIO NUOVO (1:SI, 0:NO)	AFFLUSSO STADIO	MOTIVO PORTE CHIUSE	BOOLEAN PORTE CHIUSE	
19108	84.88%		0 regolare			1
19033	82.20%		0 regolare			1
0	0.00%		0 porte chiuse	MOTIVI DISCIPLINARI		0
0	0.00%		0 porte chiuse	COVID-19		0

Figura 3.7: parte del dataset sul pubblico allo stadio

Le ultime colonne del dataset costruito (Figura 3.8) riportano due variabili la prima riporta la distanza in chilometri della trasferta che la squadra ospite deve ricoprire per disputare la partita, la seconda è la variabile binaria che assume valore 1 se il match si disputa in orario serale, ovvero dopo le 18.00, e 0 nel caso di orario diurno.

DISTANZA TRASFERTA	BOOLEAN ORARIO (1: serale; 0: giorno)
479	0
296	1
965	1
250	1
661	0
334	1

Figura 3.8: parte finale dataset

Prima di aver effettuato le diverse analisi, sono state fatte alcune ipotesi circa i

dati raccolti e sono state rimosse alcune osservazioni.

La prima, come già accennato in precedenza, è stata quella di considerare a porte chiuse e quindi con $NumeroSpettatori = 0$ anche gli incontri disputati con capienza ridotta a 1000 persone nella stagione 2020/2021. Questo perché si è ipotizzato che la presenza di questo numero esiguo di tifosi (nella maggior parte dei casi neanche spettatori paganti, ma invitati) in stadi di capienza media di 30 mila persone non sia significativa.

Inoltre, sempre riguardo la variabile *NumeroSpettatori*, il dato contiene il numero di tifosi sia della squadra di casa che degli ospiti, in quanto non si ha il dettaglio della suddivisione precisa. A tal proposito, è importante osservare che dal “REGOLAMENTO DEGLI STADI DELLA LEGA NAZIONALE PROFESSIONISTI” (Comunicato Ufficiale n. 109 del 9 novembre 2006 e successive modifiche pubblicate con Comunicato Ufficiale n. 1 del 4 luglio 2007) negli Stadi di Serie A e Serie B deve essere riservato un settore avente una capienza minima pari al 5% della capienza totale per i sostenitori della squadra ospite.

Nel dettaglio, per ogni stadio delle squadre che hanno disputato almeno una stagione nella Serie A, si può osservare nella Figura 3.9 la capienza del settore ospiti rispetto alla capienza totale. Visti questi numeri, essendo la capienza massima del settore ospiti quella dello stadio San Siro di Milano con 4500 posti (in percentuale sul totale è solo il 6%) e la capienza massima percentuale del 19% negli stadi del Sassuolo e dell’Empoli, si è deciso di non effettuare nessuna approssimazione della variabile *Numero Spettatori*.

SQUADRA	CITTA'	CAPIENZA STADIO	CAPIENZA SETTORE OSPITI	PERCENTUALE DIMENSIONE SETT. OSPITI
Atalanta	Bergamo	22512	2000	9%
Benevento	Benevento	16867	1663	10%
Bologna	Bologna	36462	2500	7%
Brescia	Brescia	19533	1000	5%
Cagliari	Cagliari	16416	415	3%
Chievo	Verona	31045	1595	5%
Crotone	Crotone	16108	834	5%
Empoli	Empoli	16800	3144	19%
Fiorentina	Firenze	43147	2450	6%
Frosinone	Frosinone	16227	2997	18%
Genoa	Genova	36599	2032	6%
Inter	Milano	78275	4500	6%
Juventus	Torino	41507	2100	5%
Lazio	Roma	70634	5800	8%
Lecce	Lecce	31533	1630	5%
Milan	Milano	78275	4500	6%
Napoli	Napoli	54726	4000	7%
Parma	Parma	22352	1152	5%
Roma	Roma	70634	5800	8%
Sampdoria	Sampdoria	36599	2032	6%
Sassuolo	Reggio Emilia	21525	4000	19%
Spal	Ferrara	16134	1490	9%
Spezia	Spezia	10336	1600	15%
Torino	Torino	28177	1495	5%
Udinese	Udine	25144	1311	5%
Verona	Verona	31045	1595	5%

Tabella 3.9: tabella che ricapitola la capienza del settore ospiti degli stadi di Serie A

Inoltre, alcune partite sono state disputate eccezionalmente non nello stadio della squadra di casa, ma in campi neutri; in particolare si tratta di 4 eventi relativi all'Atalanta tra la fine della stagione 2018/2019 e l'inizio della stagione 2019/2020, periodo in cui era in corso la ristrutturazione dello stadio di Bergamo. Queste osservazioni, quindi, sono state rimosse dal dataset per la successiva analisi di regressione.

3.3 Statistiche descrittive

In questo paragrafo analizzeremo, con l'aiuto di qualche statistica descrittiva, il dataset costruito. In particolare il focus è sulle variabili che sono ritenute più importanti nello studio, ovvero:

- *NumeroSpettatori*, *DenistàStadio*: la prima variabile conta il numero di spettatori presenti allo stadio, la seconda misura il "calore" del pubblico, misurandolo come capienza raggiunta per ogni incontro dello

stadio in percentuale.

- *Valoresqcasa, Valuesqospite, Diffvalori*: variabili che per ogni osservazione (partita) misura il valore di mercato della rosa in milioni di euro rispettivamente della squadra di casa, di quella ospite e la differenza tra queste.
- *GoalSquadracasa, GoalSquadraospite*: variabili che contano i gol rispettivamente segnati dalla squadra di casa e dalla squadra ospite per ogni partita.
- *Risultato*: variabile che descrive il risultato dell'incontro in considerazione.

3.3.1 Statistiche descrittive sui tifosi

In questa prima parte, si analizza la parte del dataset che contiene i dati relativi ai tifosi.

In primis, è importante osservare come siano presenti nel dataset della Serie A 1007 incontri con un afflusso regolare di tifosi allo stadio e, invece, 513 match disputati a porte chiuse, di cui 510 a causa del Covid-19 e delle restrizioni descritte in precedenza e 3 a causa di divieti decisi dal Giudice Sportivo nei confronti della tifoseria di casa. Invece nelle osservazioni relative alla Serie B sono presenti 499 match disputati a porte chiuse (2 per motivi disciplinari e 497 causa Covid-19) e 1065 disputati con il pubblico presente.

Nelle Figure 3.10 e 3.11 si possono osservare i pie chart che rappresentano la suddivisione delle osservazioni appena descritta per la Serie A e Serie B.

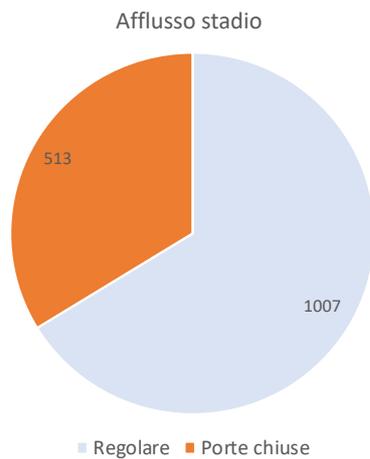


Figura 3.10: Afflusso tifosi allo stadio Serie A

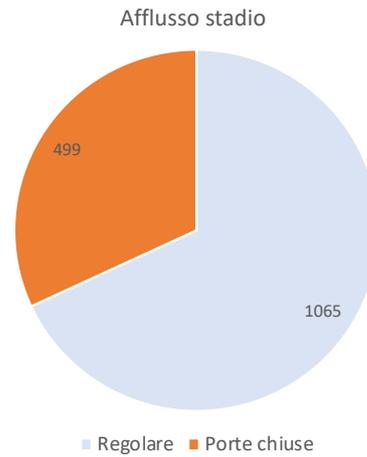


Figura 3.11: Afflusso tifosi allo stadio Serie B

Non considerando dunque le partite disputate a porte chiuse, la Tabella 3.2 riporta per le squadre della Serie A le statistiche descrittive relative al numero di spettatori allo stadio. In particolare sono stati raggruppati i dati per ogni singola squadra, ognuna delle quali ha un numero di osservazioni diverso dovuto al fatto che non tutte hanno disputato tutte le stagioni in Serie A. Quindi si possono osservare per ogni squadra il numero medio, la deviazione standard, la mediana e i valori massimo e minimo di spettatori. La tabella ordina le squadre da quella con più seguito allo stadio a quella meno e si può notare come in cima alla classifica ci sia l'Inter con 60983 tifosi di media allo stadio e in fondo l'Empoli con 9506. La capienza massima assoluta per un incontro di serie A in questi anni è stata di 78328 spettatori a San Siro (tutto esaurito) per Inter-Milan, Milan-Juventus e Inter-Juventus nel 2017/2018; il match che il numero più basso di tifosi è stato invece Empoli-Chievo nella stagione 2018/2019 con 7030 spettatori presenti.

Tuttavia, è importante osservare come questi dati abbiano una grossa dispersione come si può osservare dai valori mediamente alti delle deviazioni standard; ad esempio, Lazio e Napoli presentano deviazioni standard pari rispettivamente a 11829 e 10286 rispetto alla loro media di 36325 e 43849 spettatori. Questo è dovuto

al fatto che la “voglia di stadio” dei tifosi riflette molto l’andamento della squadra; infatti se una squadra non sta ottenendo i risultati desiderati, può venir meno l’afflusso dei tifosi allo stadio.

È utile osservare che le squadre con deviazione standard più basse siano Cagliari, Frosinone e Spal. Tuttavia per provare a capire meglio il significato di questi numeri conviene effettuare le stesse statistiche descrittive sulla densità dei tifosi allo stadio, in quanto in termini di numero assoluto di spettatori, senza considerare la capienza massima dello stadio, si rischia di fare confronti non precisi.

La Tabella 3.3 riporta queste informazioni e si può appunto osservare come le tre squadre sopracitate siano nelle top 5 di quelle con densità di tifosi maggiore. La squadra che riempie mediamente di più lo stadio è la Juventus (densità media pari a 0,95 con deviazione standard pari a 0,055) e questa è una conferma circa la correlazioni tra risultati desiderati e “voglia di stadio” da parte dei tifosi. Invece, al fondo di questa classifica c’è il Chievo con una densità media pari a 0,41 e valore minimo raggiunto pari al 26% di tifosi sul totale della capienza massima raggiungibile.

È interessante osservare come le squadre che hanno raggiunto il “tutto esaurito” (valore massimo = 1) in almeno un match in queste stagioni siano: Juventus, Cagliari, Frosinone, Spal, Udinese, Inter, Brescia, Benevento, Milan, Napoli, Sassuolo.

A tal proposito, è interessante provare a capire per quali incontri vi è un riempimento dello stadio maggiore in termini di densità; per questo motivo si è osservato se è presente una correlazione tra la densità di tifosi in un incontro e il valore della rosa della squadra ospite e la differenza di valore rispetto alla squadra casa; osservando la matrice di correlazione riportata in Figura 3.12 dimostra come non vi sia una stretta correlazione questi.

SquadraCasa	Numero osservazioni	Media	Mediana	Deviazione Standard	Valore minimo	Valore massimo
Inter	49	60983	59648	8653	46807	78328
Milan	50	53745	50393	9568	39504	78328
Juventus	50	39399	40091	2296	29412	41495
Roma	51	38382	35609	8276	29045	62304
Lazio	52	36325	34773	11829	17956	68236
Napoli	51	34849	34474	10246	16171	55567
Fiorentina	51	30286	30535	5224	21043	40872
Lecce	13	23295	22087	1997	21321	26591
Genoa	50	21543	20509	3447	18281	32288
Bologna	51	21536	21241	3022	16076	29270
Torino	50	20324	19602	3641	14605	26664
Sampdoria	51	20205	19260	2875	17563	30548
Udinese	50	19783	19677	3309	13948	25021
Atalanta	46	18634	18640	1264	16120	20940
Verona	31	17629	16282	4377	13724	29215
Parma	31	16381	15299	2161	13141	20814
Cagliari	51	15204	15378	963	13213	16412
Brescia	13	14919	14574	2012	12133	19451
Frosinone	18	13453	12848	1365	12167	16310
Spal	49	12989	12906	1349	10338	16146
Chievo	38	12839	10650	5212	8000	28375
Benevento	19	12131	11054	2448	9443	16867
Sassuolo	50	12053	11376	3318	7563	21584
Crotone	19	10580	9158	2334	8564	15021
Empoli	19	9506	8736	2340	7030	15889

Tabella 3.2: statistiche descrittive tifosi Serie A

SquadraCasa	Numero osservazioni	Media	Mediana	Deviazione Standard	Valore minimo	Valore massimo
Juventus	50,00	0,95	0,97	0,055	0,71	1,00
Cagliari	51,00	0,93	0,94	0,058	0,80	1,00
Frosinone	18,00	0,83	0,79	0,085	0,75	1,00
Atalanta	46,00	0,83	0,83	0,056	0,72	0,93
Spal	49,00	0,80	0,80	0,084	0,64	1,00
Udinese	50,00	0,79	0,79	0,132	0,55	1,00
Inter	49,00	0,78	0,76	0,110	0,60	1,00
Brescia	13,00	0,76	0,75	0,103	0,62	1,00
Lecce	13,00	0,74	0,70	0,062	0,68	0,84
Parma	31,00	0,73	0,68	0,097	0,59	0,93
Torino	50,00	0,72	0,70	0,129	0,52	0,95
Benevento	19,00	0,72	0,66	0,145	0,56	1,00
Fiorentina	51,00	0,70	0,71	0,121	0,49	0,95
Milan	50,00	0,69	0,65	0,123	0,50	1,00
Crotone	19,00	0,66	0,57	0,144	0,53	0,93
Napoli	51,00	0,64	0,64	0,187	0,30	1,00
Bologna	51,00	0,59	0,58	0,083	0,44	0,80
Genoa	50,00	0,59	0,56	0,094	0,50	0,88
Verona	31,00	0,57	0,52	0,141	0,44	0,94
Empoli	19,00	0,57	0,52	0,140	0,42	0,95
Sassuolo	50,00	0,56	0,53	0,154	0,35	1,00
Sampdoria	51,00	0,55	0,53	0,078	0,48	0,83
Roma	51,00	0,54	0,50	0,117	0,41	0,88
Lazio	52,00	0,51	0,50	0,168	0,25	0,97
Chievo	38,00	0,41	0,34	0,168	0,26	0,91

Tabella 3.3: statistiche descrittive densità tifosi Serie A

	ValoreSquadraOspite	DifferenzaValori	DensitàStadio
ValoreSquadraOspite	1		0,0142
DifferenzaValori		1	-0,0416
DensitàStadio	0,0142	-0,0416	1

Figura 3.12: matrice correlazione

Le stesse statistiche descrittive circa il numero di spettatori e la densità sono state effettuate per la Serie B e i dati sono riportati rispettivamente nella Tabella 3.4 e 3.5.

Si può osservare come mediamente il numero di spettatori nella categoria cadetta sia nettamente più basso; infatti la squadra con maggior seguito è il Bari con 15800 tifosi. Inoltre i team aventi numero medio di spettatori maggiore di 10mila sono soltanto 8 su 34.

In questa competizione il picco di numero di tifosi è stato raggiunto dal Bari con 33567 e subito a seguire il Palermo con 28351, ma alcuni incontri in casa del Carpi e della Virtus Entella sono stati seguiti solamente da poco più di 1000 tifosi.

È interessante osservare come il discostamento dai valori medi sia molto più basso rispetto a quelli trovati in Serie A; tuttavia alcune squadre come Palermo, Lecce e Bari mostrano una deviazione standard molto elevata.

Inoltre, soffermandosi sulle squadre che hanno partecipato ad entrambe le competizioni, si può osservare come i tifosi in Serie B abbiano molto meno interesse nel seguire il match allo stadio; ad esempio il Chievo è passato da 12839 tifosi di media in Serie A a solamente 3808 nella categoria cadetta, il Lecce da 23295 a 12115, l'Empoli da 9506 a 5154.

SquadraCasa	Numero osservazioni	Media	Mediana	Deviazione Standard	Valore minimo	Valore massimo
Bari	21	15800	14302	4690	12530	33567
Lecce	18	12115	10990	4444	8198	25135
Cesena	21	12105	11581	1511	10187	16234
Parma	21	11385	11261	1278	9850	14820
Foggia	38	10660	10554	1299	8836	14343
Frosinone	34	10596	10869	2672	912	16286
Verona	18	10574	9881	1741	8258	14517
Benevento	32	10338	9989	1121	9019	13622
Palermo	39	9197	6844	5449	4513	28351
Salernitana	52	8493	7842	2750	5019	18003
Perugia	53	8182	7885	1222	6483	13696
Pisa	14	7858	7792	581	7038	9133
Padova	18	7543	7371	933	6405	9338
Brescia	39	7494	7053	1860	5081	14000
Cremonese	51	7036	6528	1447	5416	11840
Pescara	53	6966	6579	1401	5162	12672
Cosenza	30	6938	6547	2206	3032	12375
Crotone	32	6324	5907	1336	4879	10565
Ascoli	50	5900	5707	1026	3903	8764
Livorno	31	5496	5540	1259	3946	10664
Spezia	53	5482	5330	648	4660	8270
Empoli	35	5154	5156	1150	3963	9132
Termana	21	4768	4129	1820	2455	9509
Avellino	21	4420	3783	1859	2964	10000
Trapani	13	4331	4337	3208	3908	4851
Novara	21	4044	3854	758	2910	5735
Cittadella	52	3933	3889	719	2892	6605
Venezia	51	3809	3533	9436	2181	5944
Chievo	12	3808	3500	777	3000	6002
Pordenone	14	3491	3232	1078	2345	6856
Juve_Stabia	14	3473	3399	501	2846	4454
Pro_Vercelli	21	2753	2484	646	2092	4446
Carpi	39	2221	2126	575	1417	3812
Virtus_Ent^a	33	2082	1849	592	1262	3813

Tabella 3.4: statistiche descrittive tifosi Serie B

Per quanto riguarda il riempimento degli stadi in Serie B, si può osservare dalla Tabella 4.14 come la squadra che mediamente lascia meno posti vuoti sia stato il Pisa che presenta una densità media pari a 0.91.

Tuttavia, come dai dati in termini assoluti esaminati in precedenza, si può dedurre come in questa competizione la densità media è nettamente minore rispetto alla Serie A. Infatti solamente 5 squadre mediamente riempie per almeno il 60% lo stadio di casa e ben 5 squadre presentano valori di densità medi minori dello 0.25; alcune partite addirittura hanno visto lo stadio occupato per meno del 10% (Pordenone 9%) e solamente Pisa e Frosinone hanno raggiunto il tutto esaurito almeno una volta.

SquadraCasa	Numero osservazioni	Media	Mediana	Deviazione Standard	Valore minimo	Valore massimo
Pisa	14	0,91	0,91	0,067	0,82	1,00
Foggia	38	0,73	0,73	0,089	0,61	0,99
Frosinone	34	0,65	0,67	0,653	0,06	1,00
Benevento	32	0,61	0,59	0,068	0,53	0,81
Cesena	21	0,60	0,57	0,075	0,50	0,80
Trapani	13	0,56	0,56	0,556	0,56	0,56
Spezia	53	0,53	0,52	0,063	0,53	0,80
Ascoli	50	0,52	0,50	0,091	0,34	0,77
Venezia	51	0,52	0,48	0,127	0,30	0,81
Cittadella	52	0,52	0,51	0,095	0,38	0,87
Parma	21	0,51	0,50	0,057	0,44	0,66
Pro_Vercelli	21	0,50	0,45	0,118	0,38	0,81
Juve_Stabia	14	0,45	0,45	0,066	0,37	0,58
Cremonese	51	0,44	0,41	0,091	0,34	0,74
Avellino	21	0,43	0,37	0,182	0,29	0,98
Padova	18	0,41	0,41	0,051	0,35	0,51
Carpi	39	0,40	0,39	0,104	0,26	0,69
Crotone	32	0,39	0,37	0,083	0,30	0,66
Lecce	18	0,38	0,35	0,141	0,26	0,80
Brescia	39	0,38	0,36	0,096	0,26	0,72
Virtus_Entella	33	0,37	0,33	0,106	0,23	0,68
Perugia	53	0,35	0,33	0,052	0,27	0,58
Verona	18	0,34	0,32	0,056	0,27	0,47
Pescara	53	0,34	0,32	0,068	0,25	0,62
Empoli	35	0,31	0,31	0,068	0,24	0,54
Cosenza	30	0,29	0,27	0,090	0,13	0,51
Livorno	31	0,29	0,29	0,065	0,21	0,55
Ternana	21	0,27	0,24	0,103	0,14	0,54
Bari	21	0,27	0,25	0,081	0,22	0,58
Palermo	39	0,25	0,19	0,150	0,12	0,78
Salernitana	52	0,23	0,21	0,073	0,13	0,48
Novara	21	0,23	0,23	0,226	0,16	0,32
Pordenone	14	0,14	0,13	0,043	0,09	0,27
Chievo	12	0,12	0,11	0,024	0,10	0,19

Tabella 3.5: statistiche descrittive densità tifosi Serie B

La figura 4.15 riporta il numero di spettatori per le squadre che hanno partecipato con continuità alla Serie A nelle stagioni pre-Covid (17/18, 18/19, 19/20) e se ne può osservare il trend. In particolare, per tutti i team eccetto il Napoli si può notare un aumento del numero di spettatori; è particolarmente netto l'incremento del seguito dei tifosi per l'Inter, la Fiorentina e la Lazio.

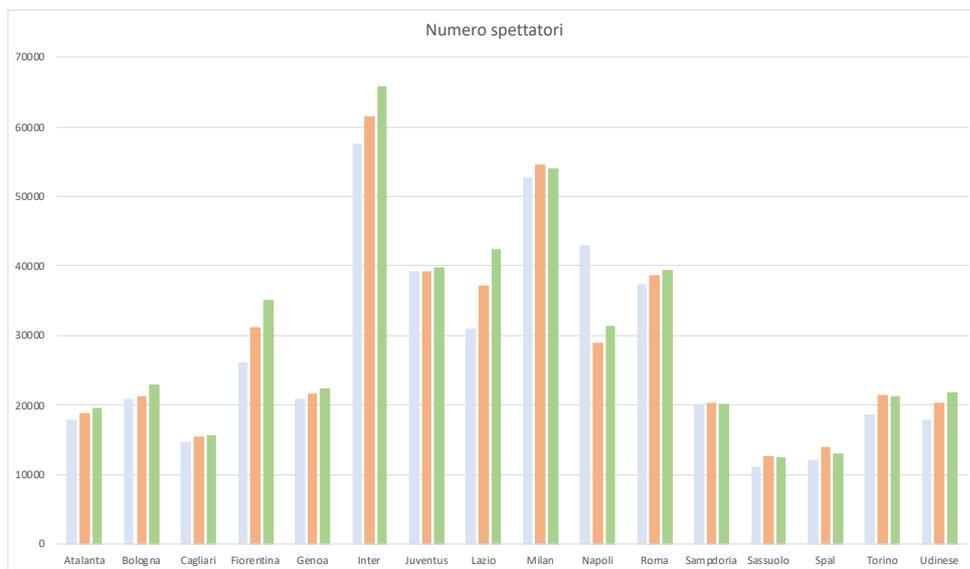


Figura 3.12: trend tifosi Serie A per squadra

La Figura 3.13 e la Figura 3.14 rappresentano due istogrammi estrapolati da STATA circa la distribuzione dei tifosi nelle osservazioni del dataset rispettivamente in Serie A e Serie B; da tali grafici si può osservare la dispersione troppo grande che rischierebbe di comportare criticità nell'analisi di regressione. Per questo motivo si è deciso di effettuare la trasformazione logaritmica della variabile *NumeroSpettatori*, generando la nuova denominata *ln_NumeroSpettatori*, che presenta una distribuzione molto più concentrata come si può osservare in Figura 3.15 per la Serie a e in Figura 3.16 per la Serie B.

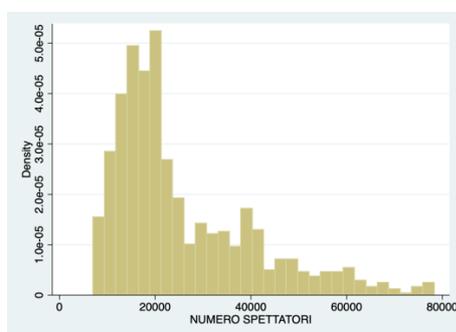


Figura 3.13: distribuzione NumeroSpettatori Serie A

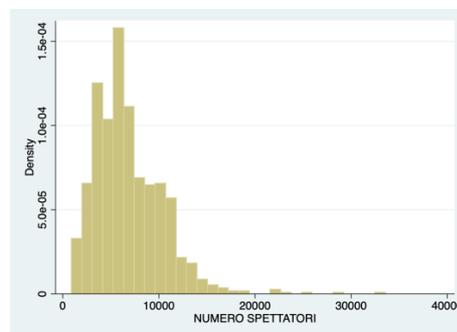


Figura 3.14: distribuzione NumeroSpettatori Serie B

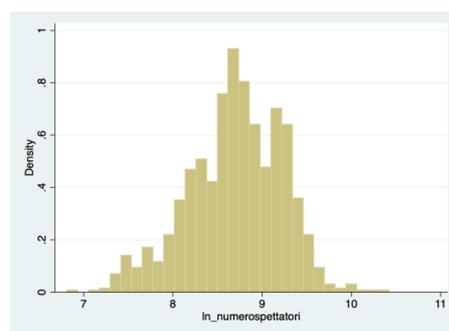
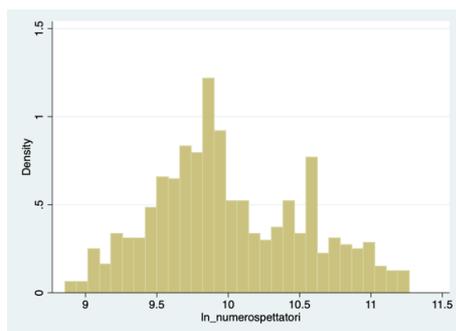


Figura 3.15: distribuzione $\ln_NumeroSpettatori$ Serie A Figura 3.16: distribuzione $\ln_NumeroSpettatori$ Serie B

Seguendo lo stesso ragionamento, sono state effettuate le trasformazioni logaritmiche anche per le variabili continue *DistanzaTrasferta*, *ValoremercatoSquadraCasa*, *ValoremercatoSquadraOspite*, *DifferenzaValori* generando rispettivamente le variabili $\ln_Distanza$, $\ln_Valoresqcasa$, $\ln_Valoresqospite$ e $\ln_Diffvalori$.

3.3.2 Statistiche descrittive sui valori economici delle squadre

Analizzando ora il dataset da un punto di vista più economico, si focalizza l'attenzione sul valore di mercato delle rose nelle diverse stagioni. La Tabella 3.6 riporta questi dati per tutte le squadre che hanno preso parte almeno un anno alla Serie A.

Si può osservare come la Juventus in media abbia un valore nettamente più alto rispetto a tutte le altre società, bensì nell'ultima stagione il valore della propria rosa è passato da 869 milioni di euro a 690 milioni di euro, quindi raggiunto da quello dell'Inter che è aumentato fino a 685 milioni di euro. In fondo a questa classifica vi è lo Spezia, che è riuscita ad arrivare in Serie A soltanto nell'ultima stagione e la società per provare a competere in questa competizione ha aumentato del 144% il valore della rosa, investendo sul mercato, arrivando a circa 47 milioni di euro.

L'esempio dello Spezia si ripete anche per tutte le altre squadre che nell'anno in cui sono state promosse hanno aumentato nettamente il loro valore della rosa. A tal proposito in Tabella 3.7 sono riportati i valori delle rose di tali squadre nelle due categorie e l'incremento percentuale.

	2017/2018	2018/2019	2019/2020	2020/2021	Media
Juventus	496,00	776,50	869,00	690,20	707,93
Inter	319,70	560,70	601,90	685,10	541,85
Napoli	371,58	490,48	620,00	593,98	519,01
Milan	350,05	454,50	503,50	409,90	429,49
Roma	323,43	404,70	409,90	336,65	368,67
Lazio	206,90	345,15	298,55	361,70	303,08
Atalanta	126,20	161,70	252,00	372,95	228,21
Fiorentina	136,65	235,15	239,10	242,70	213,40
Torino	122,00	157,60	196,83	161,90	159,58
Sassuolo	102,95	120,43	154,93	195,00	143,33
Sampdoria	100,60	137,55	150,38	110,45	124,75
Udinese	81,30	97,15	153,30	148,05	119,95
Cagliari	52,55	108,28	157,38	154,38	118,15
Genoa	93,20	84,85	138,15	100,25	104,11
Bologna	70,75	71,80	106,93	138,00	96,87
Parma	21,00	47,20	78,98	107,88	63,77
Verona	36,43	39,90	38,73	109,05	56,03
Spal	32,28	51,48	72,30	49,90	51,49
Brescia	19,30	29,83	69,18	47,70	41,50
Empoli	28,15	34,60	46,98	27,53	34,32
Benevento	32,95	32,25	23,28	44,56	33,26
Chievo	43,28	44,65	25,05	15,70	32,17
Crotone	27,40	32,98	21,70	40,95	30,76
Frosinone	29,08	39,30	25,65	27,30	30,33
Spezia	22,25	18,55	19,25	46,95	26,75

Tabella 3.6: valore di mercato delle rose Serie A

	Serie B	Serie A	Aumento %
Verona	38,73	109,05	182%
Cagliari	52,55	108,28	106%
Brescia	29,83	69,18	132%
Parma	21,00	47,20	125%
Spezia	19,25	46,95	144%
Chievo	25,05	44,65	78%
Benevento	23,28	44,56	91%
Crotone	21,70	40,95	89%
Frosinone	29,08	39,30	35%
Empoli	28,15	34,60	23%

Tabella 3.7: scostamento valore di mercato rose tra Serie B e Serie A

I valori delle rose delle squadre in Serie B per ogni stagione sono riportati invece il Tabella 3.8, ripotando soltanto quelle che hanno disputato almeno tre stagioni in tale competizione.

Squadra	2017/2018	2018/2019	2019/2020	2020/2021	Media
Ascoli	14,630	22,930	28,580	22,530	22,168
Brescia	19,300	29,830		47,700	32,277
Cittadella	13,680	14,850	12,530	10,500	12,890
Cosenza		11,950	12,800	12,680	12,477
Cremonese	19,080	20,230	21,280	15,700	19,073
Empoli	28,150		46,980	27,530	34,220
Frosinone	29,080		25,650	27,300	27,343
Perugia	23,130	31,550	28,530		27,737
Pescara	29,150	23,630	22,600	23,180	24,640
Salernitana	17,950	20,200	17,680	23,980	19,953
Spezia	22,250	18,550	19,250		20,017
Venezia	9,900	16,800	13,530	21,250	15,370
Virtus Entella	19,580		19,630	14,050	17,753

Tabella 3.8: valore di mercato delle rose Serie B

In Figura 3.17 si può osservare quale è stato il trend medio dei valori delle rose nelle due competizioni in questi anni; si osserva come tra le prime due stagioni in entrambe le categorie ci sia un aumento netto degli investimenti sul mercato, per poi stabilizzarsi tra la stagione 18/19 e 19/20. Come era ovvio aspettarsi il trend si inverte nell'ultima stagione a causa principalmente del sopraggiungere della pandemia che ha portato ad una crisi globale e quindi anche ad una riduzione degli investimenti.

I dati della stagione corrente ci si aspetta seguano questo trend, sperando ad una ripresa dell'economia mondiale e di conseguenza agli investimenti in tutti i settori.

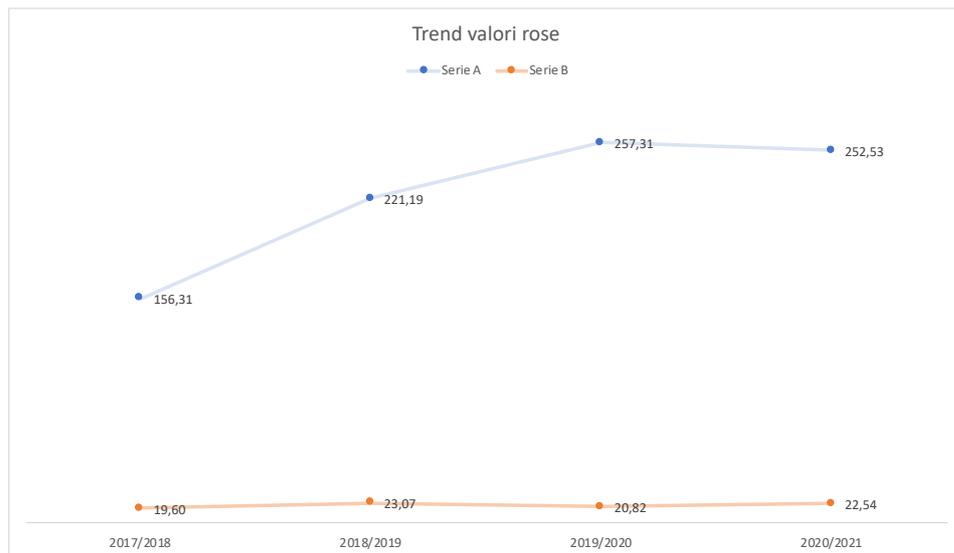


Figura 3.17: trend valori rose squadre per le due competizioni

3.3.3 Statistiche descrittive sui goal

Le ultime statistiche descrittive riportate riguardano i goal segnati e subiti dalle squadre in casa, che sono la misura sulla quale si basa il modello descritto nel capitolo successivo.

La Tabella 3.9 riporta per ogni squadra media, mediana, deviazione standard, massimo e minimo dei goal fatti e subiti in casa per ogni singola squadra in Serie A nelle stagioni in esame. Si può osservare come per tutti i team i valori della mediana sono 1 o 2 sia per i goal fatti che per quelli subiti. Per quanto riguarda la media, invece, questa assume valori molto diversi; la squadra con la media di goal fatti più alta è la Lazio con un valore pari a 2.31 reti, di poco superiore rispetto alla Juventus (2.26). Queste due squadre sono le uniche a raggiungere un valore medio maggiore di 2; tutte le altre hanno segnato in media almeno un goal, eccetto il Chievo e il Frosinone che sono la squadra con i valori medi più bassi, rispettivamente 0.95 e 0.78 goal fatti.

Un dato abbastanza sorprendente che si può dedurre da questa tabella è il

seguito: solamente 12 squadre su 25, quindi meno del 50%, ha segnato più goal rispetto a quanti ne ha subiti in casa; tuttavia la media totale riporta un valore maggiore di goal segnati (1.489 contro 1.256).

Inoltre i valori della deviazione standard dei goal fatti di ogni squadra si distribuiscono tra 0.8 e 1.6 senza nessun caso particolare e tutti i team presentano come valore minimo 0 goal fatti, ovvero tutte le squadre hanno almeno una volta finito la partita in casa senza segnare. Il maggior numero di reti segnate in casa in una partita è stato raggiunto dalla Juventus e dall'Atalanta (7 goal).

Circa i goal subiti dalle squadre in casa, il massimo è stato 7 reti subite dal Torino e dal Lecce e, come per quelli segnati, tutte le squadre hanno almeno in un'occasione subito 0 goal in casa.

La Juventus è la squadra avente il miglior dato medio e con 0.620 goal per partita di media è la miglior difesa casalinga, davanti all'Inter (0.735) e al Milan (0.920). Queste sono le uniche tre squadre ad avere un valore minore di 1; tutte le altre mediamente subiscono tra 1 e 2 reti a partite, eccetto Benevento e Frosinone che prendono entrambe 2.11 goal a incontro in Serie A.

	Media	Mediana	Deviazione sta	Min	Max		Media	Mediana	Dev.Standard	Minimo	Massimo
Atalanta						Lazio					
Goal Fatti	1,957	2	1,619	0	7	Goal Fatti	2,308	2	1,627	0	6
Goal Subiti	1,065	1	0,929	0	3	Goal Subiti	1,077	1	1,082	0	4
Benevento						Lecce					
Goal Fatti	1,211	1	1,134	0	3	Goal Fatti	1,462	2	1,127	0	4
Goal Subiti	2,105	2	1,370	0	5	Goal Subiti	2,000	1	1,826	0	7
Bologna						Milan					
Goal Fatti	1,392	1	1,021	0	4	Goal Fatti	1,360	1	1,208	0	5
Goal Subiti	1,333	1	1,030	0	4	Goal Subiti	0,920	1	0,877	0	3
Brescia						Napoli					
Goal Fatti	1,000	1	1,080	0	3	Goal Fatti	1,980	2	1,393	0	6
Goal Subiti	1,923	2	1,256	0	4	Goal Subiti	1,020	1	0,969	0	4
Cagliari						Parma					
Goal Fatti	1,294	1	1,154	0	5	Goal Fatti	1,129	1	1,176	0	5
Goal Subiti	1,431	1	1,118	0	5	Goal Subiti	1,161	1	1,158	0	4
Chievo						Roma					
Goal Fatti	0,947	1	0,899	0	3	Goal Fatti	1,961	2	1,310	0	5
Goal Subiti	1,605	1	1,326	0	5	Goal Subiti	1,255	1	0,977	0	4
Crotone						Sampdoria					
Goal Fatti	1,211	1	1,273	0	4	Goal Fatti	1,529	1	1,376	0	5
Goal Subiti	1,316	1	1,003	0	3	Goal Subiti	1,118	1	1,291	0	5
Empoli						Sassuolo					
Goal Fatti	1,737	2	1,098	0	4	Goal Fatti	1,380	1	1,323	0	5
Goal Subiti	1,579	1	1,216	0	4	Goal Subiti	1,540	1	1,432	0	6
Fiorentina						Spal					
Goal Fatti	1,275	1	1,250	0	6	Goal Fatti	1,061	1	0,852	0	3
Goal Subiti	1,137	1	1,249	0	4	Goal Subiti	1,531	1	1,226	0	5
Frosinone						Torino					
Goal Fatti	0,778	1	1,003	0	3	Goal Fatti	1,440	1	1,053	0	4
Goal Subiti	2,111	2	1,410	0	5	Goal Subiti	1,180	1	1,257	0	7
Genoa						Udinese					
Goal Fatti	1,160	1	0,866	0	3	Goal Fatti	1,120	1	1,023	0	4
Goal Subiti	1,340	1	1,099	0	4	Goal Subiti	1,280	1	1,310	0	6
Inter						Verona					
Goal Fatti	1,776	2	1,295	0	5	Goal Fatti	1,000	1	1,000	0	3
Goal Subiti	0,735	1	0,861	0	3	Goal Subiti	1,516	1	1,480	0	5
Juventus						TOTALE					
Goal Fatti	2,260	2	1,226	0	7	Goal Fatti	1,489	1	1,266	0	7
Goal Subiti	0,620	0,500	0,753	0,000	3,000	Goal Subiti	1,256	1	1,187	0	7

Tabella 3.9: statistiche descrittive goal subiti e segnati squadre in Serie A

Le stesse statistiche descrittive sono state calcolate per la Serie B e la Tabella 3.10 riepiloga tali dati per le squadre in suddetta competizione.

La prima osservazione che si può fare rispetto alla Serie A è quella relativa alla media complessiva di goal fatti e subiti dalle squadre di casa; nella categoria cadetta entrambi i valori sono leggermente minori (1.412 contro 1.489 goal fatti e 1.105 contro 1.256 goal subiti). Quindi in generale in Serie B le squadre segnano meno, sia in casa che in trasferta.

È interessante inoltre osservare come le squadre promosse in Serie A già citate in precedenza abbiano in tale competizione una media goal fatti più bassa e una media goal subiti più alta rispetto alle stagioni in Serie B.

Dalla Tabella 3.10 si può notare come solamente il Benevento abbia segnato più di 2 goal a partita in media (2.054) e la maggior parte delle altre, come in Serie A,

abbiano questo valore compreso tra 1 e 2, eccetto il Novara e il Padova che rispettivamente hanno segnato 0.952 e 0.778 goal a partita. Circa le reti subite dalle squadre in casa, invece, ben 11 squadre su 39 hanno subito in media meno di 1 goal a partita (molte di più rispetto alla Serie A) e le restanti 28 subiscono meno di 2 reti; sembra quindi che in Serie B le squadre di casa presentino un'attitudine a difendere meglio.

Un'altra differenza rispetto ai risultati ottenuti dalla Serie A riguarda un numero di gran lunga superiore di squadre con più goal segnati in casa in media rispetto a quelli subiti (28 squadre su 39).

Una caratteristica comune, invece, tra le due competizioni riguarda il massimo numero di goal segnati da una squadra in casa in uno stesso incontro (7 reti); mentre in Serie B il maggior numero di reti subite in una evento è stato 5 contro le 7 della Serie A.

	Media	Mediana	Deviazione standard	Min	Max		Media	Mediana	Deviazione standard	Min	Max
Ascoli						Padova					
Goal Fatti	1,234	1	0,944	0	3	Goal Fatti	0,778	1	0,878	0	3
Goal Subiti	1,182	1	1,022	0	4	Goal Subiti	0,883	1	0,857	0	3
Avellino						Palermo					
Goal Fatti	1,381	1	1,161	0	5	Goal Fatti	1,641	1	1,203	0	4
Goal Subiti	1,238	1	0,944	0	3	Goal Subiti	0,821	1	0,885	0	3
Bari						Parma					
Goal Fatti	1,571	1	1,207	0	4	Goal Fatti	1,619	2	1,244	0	4
Goal Subiti	0,857	0	1,153	0	4	Goal Subiti	0,619	0	0,740	0	2
Benevento						Perugia					
Goal Fatti	2,054	2	1,268	0	5	Goal Fatti	1,431	1	1,141	0	5
Goal Subiti	0,892	1	0,996	0	3	Goal Subiti	1,172	1	1,216	0	5
Brescia						Pescara					
Goal Fatti	1,690	2	1,188	0	4	Goal Fatti	1,390	1	1,194	0	5
Goal Subiti	1,172	1	0,994	0	4	Goal Subiti	1,299	1	1,065	0	5
Carpi						Pisa					
Goal Fatti	1,077	1	0,984	0	4	Goal Fatti	1,421	1	1,030	0	4
Goal Subiti	1,256	1	1,019	0	4	Goal Subiti	1,026	1	1,052	0	4
Cesena						Pordenone					
Goal Fatti	1,667	1	1,278	0	4	Goal Fatti	1,368	1	0,942	0	3
Goal Subiti	1,238	1	1,136	0	3	Goal Subiti	0,816	1	0,865	0	3
Chievo						Pro Vercelli					
Goal Fatti	1,526	1	0,922	0	4	Goal Fatti	1,048	1	1,284	0	5
Goal Subiti	0,842	1	0,855	0	3	Goal Subiti	1,286	1	1,231	0	4
Cittadella						Reggiana					
Goal Fatti	1,504	1	1,131	0	4	Goal Fatti	1,000	1	0,943	0	3
Goal Subiti	1,104	1	0,995	0	3	Goal Subiti	1,211	1	0,976	0	4
Cosenza						Reggina					
Goal Fatti	1,000	1	0,831	0	3	Goal Fatti	1,053	1	0,970	0	3
Goal Subiti	1,018	1	0,863	0	3	Goal Subiti	1,158	1	1,214	0	4
Cremonese						Salernitana					
Goal Fatti	1,273	1	1,210	0	5	Goal Fatti	1,494	1	1,084	0	4
Goal Subiti	0,974	1	1,038	0	4	Goal Subiti	1,143	1	1,121	0	5
Crotone						Spal					
Goal Fatti	1,541	1	1,169	0	4	Goal Fatti	1,316	1	1,003	0	4
Goal Subiti	0,865	1	0,887	0	3	Goal Subiti	1,000	1	1,247	0	4
Empoli						Spezia					
Goal Fatti	1,898	2	1,227	0	5	Goal Fatti	1,552	1	1,273	0	5
Goal Subiti	1,034	1	0,999	0	5	Goal Subiti	0,879	1	0,900	0	4
Foggia						Ternana					
Goal Fatti	1,564	1	0,940	0	4	Goal Fatti	1,524	1	1,365	0	5
Goal Subiti	1,359	1	0,873	0	3	Goal Subiti	1,476	1	0,981	0	3
Frosinone						Trapani					
Goal Fatti	1,441	1	1,149	0	4	Goal Fatti	1,421	1	1,170	0	4
Goal Subiti	1,068	1	1,065	0	4	Goal Subiti	1,053	1	1,026	0	3
Juve Stabia						Venezia					
Goal Fatti	1,579	1	1,121	0	4	Goal Fatti	1,247	1	0,948	0	4
Goal Subiti	1,526	2	1,219	0	5	Goal Subiti	1,013	1	0,953	0	4
Lecce						Verona					
Goal Fatti	2,000	2	1,599	0	7	Goal Fatti	1,667	2	1,138	0	4
Goal Subiti	1,162	1	0,928	0	3	Goal Subiti	1,222	1	0,878	0	3
Livorno						Vicenza					
Goal Fatti	1,000	1	1,054	0	4	Goal Fatti	1,158	1	1,068	0	4
Goal Subiti	1,622	2	1,361	0	5	Goal Subiti	1,263	1	1,046	0	4
Monza						Virtus Entella					
Goal Fatti	1,421	1	1,216	0	5	Goal Fatti	1,085	1	0,952	0	3
Goal Subiti	0,842	0	1,119	0	4	Goal Subiti	1,356	1	1,310	0	5
Novara						TOTALE					
Goal Fatti	0,952	1	0,660	0	2	Goal Fatti	1,412	1	1,132	0	7
Goal Subiti	1,333	1	0,913	0	3	Goal Subiti	1,105	1	1,038	0	5

Tabella 3.10: statistiche descrittive goal subiti e segnati squadre in Serie B

Un'ultima analisi è riportata nella Figura 3.18 che mostra i grafici relativi ai trend dei goal fatti e subiti rispettivamente in la Serie A e Serie B, per osservare se con l'avvento del Covid-19 si può osservare un inversione di tendenza. A tal proposito, la stagione 2018/2019 è stata suddivisa in pre e post Covid.



Figura 3.18: trend goal fatti e subiti Serie A e Serie B

Si può osservare dai grafici come in Serie A nelle stagioni prima della pandemia il numero dei goal segnati in casa fosse in aumento stagione dopo stagione, ma il presentarsi di tale evento esogeno sembra aver interrotto e invertito questo trend. Tale osservazione è confermata anche dall'analisi del grafico relativo alla Serie B, dove il numero dei goal fatti dalle squadre in casa era già in diminuzione nelle stagioni pre-Covid e il disputare i match a porte chiuse durante la pandemia ha rafforzato questo trend.

Per quanto riguarda invece i goal subiti dalle squadre di casa, in entrambe le competizioni si osserva come la tendenza a subire sempre più reti sia rafforzata nel periodo pandemico, dimostrando come l'assenza dei tifosi casalinghi sembra rafforzare gli ospiti.

Nei capitoli successivi si analizzerà più nel dettaglio se queste prime osservazioni troveranno conferme o meno.

3.3.4 Statistiche descrittive sulle vittorie casalinghe

Le ultime statistiche descrittive riguardano i risultati degli incontri, con un focus particolare sull'andamento delle vittorie casalinghe; infatti la probabilità di vittoria in casa sarà al centro dell'analisi di regressione effettuata in questo lavoro di

tesi.

In Tabella 3.11 e in Tabella 3.12 rispettivamente per la Serie A e la Serie B sono riportati il numero di vittorie casalinghe per ogni squadra nelle diverse stagioni. Si osserva come in media il numero di vittorie in casa sia in Serie A che in Serie B sia diminuito nell'ultima stagione quando gli incontri si sono disputati a porte chiuse. In particolare nella categoria cadetta le squadre in casa vincono più partite rispetto alla Serie A, ma nell'ultima stagione l'impatto delle porte chiuse ha stravolto questo dato comportando soltanto 143 vittorie in casa contro le 155 avvenute in prima categoria.

Squadra	2017-18	2018-19	2019-20	2020-21
Atalanta	9	10	12	12
Benevento	5			2
Bologna	6	9	5	7
Brescia			3	
Cagliari	6	8	7	5
Chievo	7	1		
Crotone	6			5
Empoli		9		
Fiorentina	8	5	5	6
Frosinone		1		
Genoa	6	6	7	5
Inter	11	11	11	17
Juventus	16	15	16	14
Lazio	9	9	14	13
Lecce			4	
Milan	10	12	9	8
Napoli	14	13	10	12
Parma		5	7	2
Roma	11	12	10	13
Sampdoria	12	10	6	9
Sassuolo	4	5	8	7
Spal	5	5	2	
Spezia				5
Torino	8	12	7	3
Udinese	6	8	6	5
Verona	5		9	5
Totale complessivo	164	166	158	155
Media	8,2	8,3	7,9	7,8

Tabella 3.11: Vittorie in casa Serie A

Squadra	2017-18	2018-19	2019-20	2020-21
Ascoli	6	7	9	8
Avellino	8			
Bari	13			
Benevento		10	14	
Brescia	6	13		9
Carpi	8	3		
Cesena	9			
Chievo			10	10
Cittadella	7	9	7	9
Cosenza		7	6	3
Cremonese	6	9	7	7
Crotone		6	12	
Empoli	14		7	11
Foggia	8	8		
Frosinone	12		10	4
Juve Stabia			9	
Lecce		13		6
Livorno		7	3	
Monza				9
Novara	5			
Padova		3		
Palermo	11	8		
Parma	12			
Perugia	11	9	7	
Pescara	7	11	7	4
Pisa			10	7
Pordenone			11	5
Pro Vercelli	6			
Reggiana				6
Reggina				8
Salernitana	7	8	10	11
Spal				8
Spezia	10	10	10	
Ternana	5			
Trapani			8	
Venezia	12	6	5	10
Verona		8		
Vicenza				5
Virtus Entella	6		8	3
Totale complessivo	189	155	170	143
Media	8,6	8,2	8,5	7,2

Tabella 3.12: Vittorie in casa Serie B

Per completezza, è interessante osservare come in Serie A la squadra con più vittorie casalinghe in tutte le stagioni sia stata la Juventus, tranne nell'ultima stagione

dove l'Inter ha ottenuto il valore massimo (ben 17 vittorie su 19 incontri disputati); inoltre mai nessuno ha ottenuto zero vittorie in casa in una stagione; infatti al minimo il Chievo e il Frosinone hanno ottenuto una sola vittoria nella stagione 2018/2019.

In Serie B, invece, il numero massimo di vittorie ottenute in casa da una squadra in una singola stagione è stato raggiunto dal Benevento nella stagione 2019/2020 (14 vittorie su 19 incontri disputati), ma mediamente tutte le squadre vincono in casa un numero maggiore di partite.

CAPITOLO 4

4.1 Una prima analisi dell' Home Advantage

In questa prima analisi dell'HA si segue un modello applicato in una tesi di dottorato che analizza le competizioni inglesi dalla stagione 2001/2002 alla stagione 2011/2012 (Foulds, 2016).

Pollard (1986) ha creato una definizione quantificabile di vantaggio casalingo, per un numero equilibrato di partite, ovvero per ogni squadra lo stesso numero di partite giocate in casa e in trasferta; infatti ha definito l'home advantage come la percentuale di partite vinte in casa rispetto al totale delle partite disputate in casa omettendo i pareggi, indicando con il valore del 50% lo zero del vantaggio casalingo. Tuttavia, questo è un metodo inadatto per modellare i singoli club, infatti deve essere considerata la propria abilità nella loro lega o divisione.

Infatti, Snyder e Purdy (1985) hanno analizzato il vantaggio casalingo nel basket collegiale e hanno scoperto che la qualità dell' avversario ha influito di più rispetto all' home advantage sul risultato. Per questo motivo, per partite equilibrate come la maggior parte di quelle del calcio inglese o italiano, ha senso prendere in considerazione la forza della squadra e misurare l'home advantage confrontando i risultati in casa e in trasferta.

4.1.1 Modellare l'Home Advantage e l'abilità del team

Per sviluppare questo primo modello, serve avere a disposizione i dati a livello stagionale per le competizioni di interesse, ovvero la Serie A e la Serie B.

In particolare, per ogni annata, sono stati recuperati i dati relativi al numero di vittorie pareggi e sconfitte e goal segnati e subiti da ciascuna squadra dividendoli per i match disputati in casa e quelli in trasferta.

La Tabella 4.1 è lo scenario finale della Serie A 2018/2019. Si può osservare appunto che i risultati in casa e in trasferta sono separati. Ad esempio, senza considerare le ultime 5 colonne che verranno spiegate successivamente e tengono conto di una stima dell'abilità delle squadre, si osserva come la Fiorentina sembrerebbe avere Home Advantage nullo in quanto in casa ha lo stesso numero di vittorie e sconfitte, ottenendo così il 50%, ovvero lo zero del vantaggio casalingo.

Tuttavia, tenendo conto anche delle prestazioni in trasferta, si può ottenere la *team ability u* e considerarla per il conteggio dell'*home advantage h*. Di nuovo, osservando la tabella si nota come la Fiorentina abbia ottenuto soltanto 3 vittorie nelle 19 partite giocate in trasferta, con una differenza reti di -4.

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Atalanta	10	5	4	36	19	10	4	5	41	27	69	17	14	-0,152	0,994
Bologna	9	4	6	28	22	2	7	10	20	34	44	6	-14	0,792	-0,453
Cagliari	8	6	5	23	20	2	5	12	13	34	41	3	-21	1,015	-0,814
Chievo	1	7	11	14	37	1	7	11	11	38	20	-23	-27	-0,096	-1,058
Empoli	9	3	7	33	30	1	5	13	18	40	38	3	-22	1,070	-0,867
Fiorentina	5	9	5	25	19	3	8	8	22	26	41	6	-4	0,237	0,075
Frosinone	1	6	12	14	38	4	4	11	15	31	25	-24	-16	-0,763	-0,475
Genoa	6	8	5	21	22	2	6	11	18	35	38	-1	-17	0,570	-0,592
Inter	11	5	3	27	10	9	4	6	30	23	69	17	7	0,237	0,625
Juventus	15	4	0	39	13	13	2	4	31	17	90	26	14	0,348	0,969
Lazio	9	5	5	36	25	8	3	8	20	21	59	11	-1	0,348	0,219
Milan	12	4	3	31	17	7	7	5	24	19	68	14	5	0,181	0,528
Napoli	13	4	2	41	16	11	3	5	33	20	79	25	-13	0,348	0,919
Parma	5	7	7	18	25	5	4	10	23	36	41	-7	-13	0,015	-0,364
Roma	12	5	2	43	26	6	7	6	23	22	66	17	1	0,570	0,308
Sampdoria	10	2	7	28	16	5	6	8	32	35	53	12	-3	0,515	0,111
Sassuolo	5	10	4	33	33	4	6	9	20	27	43	0	-7	0,070	-0,067
Spal	5	7	7	20	26	6	2	11	24	30	42	-6	-6	-0,319	0,003
Torino	12	2	5	32	19	4	13	2	20	18	63	13	2	0,292	0,372
Udinese	8	5	6	22	22	3	5	11	17	31	43	0	-14	0,459	-0,436
Totale	166	108	106	564	455	106	108	166	455	564	1032	109	-109	5,737	0,000

Tabella 4.1: Tabella finale della Serie A 2018/19, compreso il vantaggio casalingo dei singoli club (h) e qualità (u).

Clarke e Norman (1995) hanno considerato tre squadre, A, B e C classificate in base alla loro abilità in ordine alfabetico (cioè A è migliore di B che è migliore di C), ognuna senza alcun vantaggio in casa. Se si dovesse supporre che i risultati fossero come nella Tabella 3.2 e la tabella finale assumesse la forma della Tabella 3.3, dove i

pareggi sono contati come 0,5 di una vittoria, è ovvio che ogni squadra ottiene lo stesso risultato in casa e in trasferta per quanto riguarda i gol segnati e le vittorie.

Tuttavia, se la squadra C dovesse partire con due gol in più giocando in casa, i risultati finali sarebbero come nella Tabella 3.4 e la Tabella 3.5 che rappresenta la tabella finale. Da quest'ultima si può osservare come tutte le squadre ora hanno risultati migliori sia in termini di differenza reti che di partite vinte in casa e in trasferta.

Home team	Away team		
	A	B	C
A	-	2-1	3-1
B	1-2	-	2-1
C	1-3	1-2	-

Tabella 3.2: Risultati finali modello Clarke e Norman

Team	Home				Away				Home-away	
	Wins	Draws	Losses	Goals	Wins	Draws	Losses	Goals	Wins	Goals
A	2	0	0	5-2	2	0	0	5-4	0	0
B	1	0	1	3-3	1	0	1	3-5	0	0
C	0	0	2	6-5	0	0	2	2-5	0	0

Tabella 3.3: Classifica finale modello Clarke e Norman

Ora è facile vedere come si possano trarre false conclusioni sul vantaggio casalingo delle singole squadre. Anche se solo C ha un vero vantaggio in casa, le squadre A e B sperimentano un falso vantaggio in casa.

Home team	Away team		
	A	B	C
A	-	2-1	3-1
B	1-2	-	2-1
C	3-3	3-2	-

Tabella 3.4: Risultati finali modello Clarke e Norman con 2 goal di vantaggio alla squadra C

Team	Home				Away				Home-away	
	Wins	Draws	Losses	Goals	Wins	Draws	Losses	Goals	Wins	Goals
A	2	0	0	5-2	1	1	0	5-2	0.5	2
B	1	0	1	3-3	0	0	2	3-3	1	2
C	1	1	0	2-5	0	0	2	2-5	1.5	4

Tabella 3.5: Classifica finale modello Clarke e Norman con 2 goal di vantaggio alla squadra C

La stima del vantaggio in casa di una squadra individuale richiede un modello per quantificare l'abilità delle squadre (Clarke, 1993; Stefani, 1983; Stefani e Clarke, 1987). Clarke e Norman (1995) hanno suggerito una derivazione di una formula per il calcolo del vantaggio casalingo e quello della prestazione di squadra mediante l'uso dei minimi quadrati. Supponiamo che il margine di vincita per la squadra di casa I contro la squadra ospite J sia rappresentato da w_{ij} , che è negativo in caso di perdita.

Ci sono due possibili definizioni di w_{ij} che potrebbero essere utilizzate. In primo luogo, può essere definito come 1 se la squadra di casa ha vinto, 0 se la partita termina in pareggio o -1 se la squadra di casa ha perso.

In alternativa, il metodo preferito per questo modello di previsione, è definire come margine (o differenza) dei gol, in quanto ciò consente di definire con una maggiore sensibilità il vantaggio casalingo. Per illustrare questo, si immagini ad esempio che una squadra in particolare vinca 4-0 in casa e vinca anche 2-1 in trasferta; con l'utilizzo del metodo vittoria, sconfitta o pareggio si otterrebbe zero vantaggio in casa; tuttavia con l'altro metodo, usando la differenza reti, ci sarebbe un vantaggio casalingo di 3 gol.

Questo secondo metodo produrrebbe una matrice $N \times N$ per N squadre, senza diagonali. Sommando lungo la i -esima riga si ottiene il valore della differenza reti in casa (HGD_I) e sommando lungo la i -esima colonna si ricava il valore della differenza reti in trasferta (AGD_I) per il team I , come illustrato di seguito:

$$HGD_I = \sum_{j=1(j \neq i)}^N w_{Ij} \qquad AGD_I = - \sum_{i=1(i \neq i)}^N w_{Ij}$$

Questo può essere modificato come segue:

$$\sum_{i=1}^N HGD_i = - \sum_{i=1}^N AGD_i$$

Definendo h_i come misura del vantaggio in casa di una singola squadra i , u_i come misura della loro qualità e ε_{ij} come un errore casuale a media zero, Clarke e Norman (1995) hanno modellato w_{ij} come:

$$w_{ij} = u_i - u_j + h_i + \varepsilon_{ij}$$

Si assume in questo modello che u_i e h_i siano costanti per tutta la durata della stagione.

Per rendere identificabile il modello, viene aggiunto un vincolo aggiuntivo, che assicura $\sum_{i=1}^N u_i = 0$ per N squadre. Il modello può quindi essere adattato ai dati utilizzando un semplice modello di regressione. Vengono utilizzate delle variabili dummy per u_i (1 per la squadra di casa, -1 per la squadra in trasferta e 0 per le altre squadre del campionato) e h_i (1 per la squadra di casa e 0 per le altre squadre).

Se si analizzano i dati di stagioni complete e bilanciate, ovvero con lo stesso numero di partite giocate in casa ed in trasferta, non è necessario utilizzare complicati metodi di regressione a favore del calcolo semplice utilizzando la classifica finale. Quest'ultimo è il metodo utilizzato nella *Tabella 3.1*.

In un campionato equilibrato di N squadre, ogni squadra gioca contro le altre $N-1$ squadre una volta in casa e una volta in trasferta. Il vantaggio casalingo totale dell'intero campionato è rappresentato da H . Questo è calcolato nel seguente modo:

$$H = \sum_{i=1}^N h_i = \frac{\sum_{i=1}^N HGD_i}{N - 1}$$

L'home advantage del team i , h_i , è data dalla differenza tra la differenza goal in casa e fuori, meno il vantaggio totale in casa H , diviso per $N-2$, come mostrato da

$$h_i = \frac{HGD_i - AGD_i - H}{N - 2}$$

Infine, la misura della capacità individuale del team, u_i , è dato da

$$u_i = \frac{HGD_i - (N - 1)h_i}{N}$$

4.2 Dati e risultati

I dati sono stati organizzati nelle tabelle di fine stagione come la Tabella 3.1 per e sono osservabili nell' Appendice A in ordine cronologico rispettivamente per la serie A e serie B.

Come spiegato precedentemente per poter utilizzare questo modello, le partite delle N squadre devono essere disputate rispettivamente metà in casa e metà i trasferta contro i rispettivi avversari e, per questo, i pochi casi di partite giocate in campo neutro (esempio Atalanta-Genoa della stagione 2018-2019 che è stata giocata a Reggio Emilia per indisponibilità dello stadio della squadra di casa), è stata considerato come match disputato in casa dalla squadra che da calendario ne aveva diritto.

Per le 3084 partite giocate in tutti e due le competizioni esaminate (1520 per la

serie A e 1564 per la serie B), sono state registrate 1300 (42.15%) vittorie casalinghe, 877 (28.44%) pareggi e 907 (29.41%) sconfitte. Più nel dettaglio per la serie A si sono verificate 643 vittorie della squadra di casa (42.30%), 373 pareggi (24.54%) e 504 vittorie della squadra ospite (33.16%); mentre per la serie B si possono osservare 657 vittorie in casa (42%), 504 pareggi (32.22%) e 403 vittorie per la squadra ospite (25.78%).

In termini di gol, le squadre di casa hanno segnato 4560 (55.02%) dei 8287 gol totali. In particolare per la serie A, i goal segnati dalla squadra di casa sono 2351 (54.05%) contro i 1999 della squadra ospite; mentre in serie B i goal realizzati dalla squadra di casa sono 2209 (56.10%) contro i 1728 subiti.

Seguendo il modello di Clarke (1996) illustrato sopra, la percentuale di gol fatti dalla squadra di casa precedentemente calcolati (55.02%; 54.95%; 56.10%) sono rispettivamente minori della percentuale di vittoria calcolata utilizzando i parametri di Clarke (1 per la vittoria, 0.5 per il pareggio). Infatti con questo metodo si ottengono rispettivamente $42.15 + 0.5 \times 29.41 = 56.37\%$, $42.3 + 0.5 \times 24.54 = 55.3\%$, $42 + 0.5 \times 32.22 = 58.11\%$.

È stato inoltre riscontrato che il vantaggio medio casalingo per squadra dalle stagioni 2017/18 al 2020/21 è maggiore nella lega di livello inferiore, ovvero la Serie B, rispetto alla Serie A, come si può evincere dalla tabella in Tabella 3.6.

Competizione	Average Home Advantage H				Media totale
	2017/18	2018/19	2019/20	2020/21	
Serie A	0,234	0,287	0,205	0,200	0,232
Serie B	0,307	0,436	0,297	0,203	0,311

Tabella 3.6: HA medio per Serie A e Serie B

Squadra	2017-18	2018-19	2019-20	2020-21	Media	Standard error
Empoli	-	1,070	-	-	-	1,070
Crotone	0,962	-	-	1,056	1,009	0,033
Brescia	-	-	0,772	-	0,772	-
Sampdoria	1,740	0,515	0,161	0,556	0,743	0,172
Lazio	0,184	0,348	0,939	0,889	0,590	0,095
Juventus	0,406	0,348	1,161	0,056	0,493	0,118
Bologna	0,295	0,792	0,050	0,556	0,423	0,080
Lecce	-	-	0,383	-	0,383	-
Chievo	0,795	-0,096	-	-	0,349	0,315
Spezia	-	-	-	0,333	0,333	-
Genoa	-0,149	0,570	0,439	0,278	0,284	0,078
Cagliari	-0,038	1,015	0,439	-0,444	0,243	0,157
Roma	-0,760	0,570	-0,673	1,722	0,215	0,294
Inter	0,073	0,237	-0,173	0,667	0,201	0,088
Napoli	-0,149	0,348	-0,061	0,611	0,187	0,089
Sassuolo	0,184	0,070	0,661	-0,222	0,173	0,092
Torino	0,518	0,292	0,216	-0,389	0,159	0,097
Udinese	-0,094	0,459	-0,006	0,111	0,118	0,061
Benevento	0,684	-	-	-0,500	0,092	0,419
Verona	0,073	-	0,439	-0,278	0,078	0,119
Fiorentina	-0,149	0,237	-0,395	0,444	0,034	0,094
Atalanta	0,073	-0,152	-0,228	0,167	-0,035	0,046
Parma	-	0,015	0,050	-0,333	-0,090	0,071
Spal	0,073	-0,319	-0,117	-	-0,121	0,065
Milan	-0,038	0,181	0,050	-1,278	-0,271	0,169
Frosinone	-	-0,763	-	-	-0,763	-
Media	0,234	0,287	0,205	0,200	0,257	0,131

Tabella 3.7: HA squadre in serie A

Squadra	2017-18	2018-19	2019-20	2020-21	Media	Standard error
Trapani	-	-	1,114	-	-	1,114
Cesena	0,862	-	-	-	0,862	-
Pordenone	-	-	1,003	0,608	0,805	0,066
Reggiana	-	-	-	0,775	0,775	-
Parma	0,762	-	-	-	0,762	-
Juve Stabia	-	-	0,670	-	0,670	-
Spezia	1,012	0,631	0,225	-	0,623	0,080
Bari	0,612	-	-	-	0,612	-
Pisa	-	-	0,336	0,830	0,583	0,082
Chievo	-	-	0,336	0,719	0,528	0,064
Ascoli	0,562	0,572	0,670	0,275	0,520	0,029
Salernitana	0,112	0,572	0,836	0,553	0,518	0,052
Pescara	0,262	0,572	0,947	0,275	0,514	0,056
Ternana	0,512	-	-	-	0,512	-
Avellino	0,512	-	-	-	0,512	-
Benevento	-	0,101	0,781	-	0,441	0,113
Crotone	-	0,337	0,503	-	0,420	0,028
Spal	-	-	-	0,330	0,330	-
Carpi	0,762	-0,134	-	-	0,314	0,149
Pro Vercelli	0,312	-	-	-	0,312	-
Brescia	0,312	0,278	-	0,330	0,307	0,005
Verona	-	0,278	-	-	0,278	-
Palermo	0,362	0,160	-	-	0,261	0,034
Empoli	0,212	-	0,058	0,497	0,256	0,045
Foggia	-0,238	0,748	-	-	0,255	0,164
Lecce	-	0,984	-	-0,503	0,241	0,248
Cremonese	-0,188	0,807	0,058	0,219	0,224	0,073
Perugia	0,112	0,219	0,281	-	0,204	0,017
Frosinone	0,362	-	0,725	-0,558	0,176	0,135
Padova	-	0,160	-	-	0,160	-
Cosenza	-	0,572	-0,497	0,219	0,098	0,111
Virtus Entell	0,212	-	0,225	-0,336	0,034	0,065
Cittadella	-0,488	0,984	-0,664	0,275	0,027	0,131
Monza	-	-	-	-0,003	-0,003	-
Venezia	0,462	0,160	-0,830	-0,003	-0,053	0,096
Vicenza	-	-	-	-0,170	-0,170	-
Livorno	-	0,278	-0,830	-	-0,276	0,185
Reggina	-	-	-	-0,281	-0,281	-
Novara	-0,638	-	-	-	-0,638	-
Media	0,307	0,436	0,297	0,203	0,329	0,088

Tabella 3.8: HA squadre serie B

Utilizzando le modalità esposte nel paragrafo precedente, i vantaggi casalinghi h ottenuti da ciascuna squadra in serie A sono illustrate in Tabella 3.7 e quelli ottenuti da ciascuna squadra in serie B si osservano in Tabella 3.8, dove i risultati sono disposti

in ordine decrescente rispetto all' Home Advantage medio per squadra.

Quindi le squadre che godono di un vantaggio casalingo medio più alto nel periodo analizzato sono l'Empoli in serie A e il Trapani in serie B; il Frosinone in A e il Novara in B sono invece le squadre col più basso Home Advantage medio.

A causa della natura variabile intrinseca delle partite di calcio, è necessario un valore medio del vantaggio casalingo su un certo numero di anni per consentire conclusioni ragionevoli. È evidente dalle tabelle che alcune squadre sono state coinvolte nelle due competizioni, quindi avranno meno osservazioni in ciascuna delle due categorie, ma si può osservare il loro diverso andamento in esse.

Inoltre si può notare che alcune squadre mostrano un vantaggio casalingo negativo. Ciò è degno di nota, poiché si pensa che un vantaggio casalingo positivo sia presente su tutte le squadre (Pollard, 1986).

Una possibile causa del vantaggio casalingo che è già stata discussa in precedenza è la familiarità della squadra con il proprio terreno di casa. Tuttavia, ciò potrebbe anche essere visto al contrario, ovvero come non familiarità delle squadre ospiti con il terreno dove si disputa il match. Pertanto, ci si può aspettare che le squadre che sono nuove nella divisione, o che sono rientrate dopo un periodo di assenza, abbiano un elevato vantaggio in casa.

Clarke e Norman (1995) hanno scoperto che questa teoria è l'opposto di ciò che mostra l'evidenza empirica, con una piccola correlazione positiva non significativa tra il vantaggio in casa e gli anni nella divisione. I dati analizzati nella tesi di dottorato (Foulds, 2016) dai primi quattro campionati inglesi sembrano seguire lo stesso risultato.

Tuttavia, ci sono alcune squadre che mostrano il comportamento ipotizzato. Ad esempio, Sheld United e Stoke City sono entrati in Premiership rispettivamente nel

2006 e nel 2008, mostrando entrambi un elevato vantaggio in casa. Anche il Birmingham City dopo un periodo di assenza di una stagione ha mostrato il suo più alto vantaggio in casa nel periodo analizzato. Ciò può essere contrastato con gli esempi di Reading e Derby County, che hanno mostrato un vantaggio casalingo più basso o negativo al momento del ritorno in Premier League.

Osservando i risultati riportati nella Tabella 3.6 riguardanti la Serie A si trova la stessa situazione della Premier League sopracitata; infatti in giallo sono evidenziati i valori dell'Home Advantage per le squadre neopromosse dalla serie inferiore. Ad esempio, nella stagione 2018/2019 si può osservare come le tre squadre neopromosse presentano risultati opposte, con l'Empoli avente l'Home Advantage più alto di tutti (1.070), il Frosinone il più basso di tutti (-0.763) e il Parma un valore sotto la media (0.015). Risulta interessante confrontare l'*h* ottenuto dalle stesse squadre nella stagione precedente (2017/2018) in Serie B, e si osserva come si capovolga la situazione; infatti l'Empoli ottiene l'*h* più basso e sotto media (0.212), il Parma un valore molto sopra la media (0.762) e il Frosinone un valore di poco sopra media (0.362), ma comunque nettamente superiore al risultato che ha. In serie A la stagione successiva.

Per ricapitolare i risultati ottenuti distinguendoli per squadre nuove, cioè promosse dalla serie B nel caso della competizione serie A e neopromosse dalla Lega Pro e retrocesse dalla serie A nel caso della competizione serie B, si può osservare rispettivamente la tabella in Tabella 3.7 e Tabella 3.9 e, invece, per le squadre aventi continuità nella stessa categoria, la tabella in Tabella 3.8 e Tabella 3.10.

NEOPROMOSSE	2017-18	2018-19	2019-20	2020-21	Media totale
h medio	0,277	0,107	0,531	0,296	0,303

Figura 3.8: HA medio per squadre neopromosse in serie A

NON NEOPROMOSSE	2017-18	2018-19	2019-20	2020-21	Media totale
h medio	0,227	0,319	0,148	0,183	0,219

Figura 3.9: HA medio per squadre già presenti in serie A nella stagione precedente

NEOPROMOSSE	2017-18	2018-19	2019-20	2020-21	Media totale
h medio	0,233	0,387	0,558	0,108	0,303

Figura 3.10: HA medio per squadre neopromosse o neoretrocesse in serie B

NON NEOPROMOSSE	2017-18	2018-19	2019-20	2020-21	Media totale
h medio	0,342	0,464	0,123	0,245	0,294

Figura 3.11: HA medio per squadre già presenti in serie B nella stagione precedente

Si può concludere da questa breve analisi, quanto detto all'inizio, ovvero che le squadre neopromosse non hanno sempre un HA maggiore rispetto ai team rivali. Infatti, in serie B, eccetto per la stagione 2019/2020, le squadre già presenti nella competizione dall'anno precedente presentano un valore HA maggiore. In serie A, invece, eccetto per la stagione 2019/2020, le squadre neopromosse presentano un HA più alto rispetto alle altre, andando a confermare la teoria sopracitata.

Come in questa analisi, molti studi provano ad analizzare e cercare delle correlazioni tra vari fattori e i risultati analizzandoli separatamente. Tuttavia, attraverso un approccio di modellazione più sofisticato, più fattori possono essere valutati simultaneamente e questo potrebbe potenzialmente rivelare molte più informazioni.

4.2.1 Una prima analisi dell'impatto del Covid-19 sull' Home Advantage

Da questa prima analisi si può iniziare a valutare se il Covid-19 ha comportato una riduzione dell'HA, come sarebbe lecito aspettarsi in quanto come detto in precedenza ha comportato la disputa di match a porte chiuse, facendo venir meno la presenza del tifo delle squadre negli stadi di casa.

In generale, osservando la Tabella 3.6, il trend dell'HA era in crescita nelle stagioni precedenti alla 2019/2020. Tuttavia. Sarebbe interessante anche replicare l'analisi per le stagioni precedenti rispetto a quella di partenza di questo studio.

Questo trend pare essere opposto rispetto a quello delle divisioni inglesi analizzate nell'elaborato di Foulds, dove mediamente l'HA sembrava decrescere con il passare degli anni.

Focalizzandosi, invece, sull'impatto del Covid-19, questo dai risultati ottenuti sembrerebbe avere portato una drastica riduzione del vantaggio casalingo, come ipotizzato. Infatti nelle stagione 2019/2020 il trend dell'Home Advantage si è invertito sia in serie A che in serie B; in particolare, in quest'ultima si ha avuto un crollo passando da un valore pari a 0.436 nella stagione 2018/2019 a 0.297 nella stagione seguente dove il Covid ha avuto impatto da circa metà stagione, fino 0.203 nella scorsa stagione dove si sono disputati tutti gli incontri a porte chiuse. In Serie A il trend è stato lo stesso, ma con una variazione molto minore portando l'HA da 0.287 della stagione 2018/2019 a 0.200 della stagione 2020/2021.

A tal proposito, è interessante osservare come nelle ultime due stagioni, in entrambe le competizioni ci sia un aumento, seppur leggero, della deviazione standard dell'h di ogni squadra, che passa come si può osservare in Tabella 3.7 e Tabella 3.8 rispettivamente da 0.021 e 0.16 del 2018/2019 fino a 0.033 e 0.021 nella stagione

finale.

Da questa analisi, non è semplice definire come il Covid-19 abbia impattato sui risultati, in quanto appunto sembrerebbe che alcune squadre abbiano aumentato il loro Home Advantage e altri lo abbiano ridotto, ma mediamente si può concludere che le porte chiuse abbiano ridotto il vantaggio casalingo.

CAPITOLO 5

5.1 L'analisi di regressione

Dopo aver osservato i risultati ottenuti dal modello di Clarke e Norman, in questo capitolo si prova a dare una spiegazione più dettagliata di essi. Infatti, per mezzo di modelli di regressione si tenta di capire e quantificare i contributi che hanno le diverse variabili relative al fattore campo sui risultati sportivi.

La scelta del modello di regressione più adatto da utilizzare dipende dal tipo di variabili, in particolar modo quella dipendente (Y).

Nel caso in esame, la variabile dipendente è di tipo dicotomica $DummyVittoriaCasa$, generata da $PuntiSquadracasa$ che tiene conto se la squadra in casa vince o no il match. Per questo motivo il modello più adatto in econometria è la regressione multipla Probit.

Il modello di regressione probit è:

$$\Pr(Y=1|X_1, X_2, \dots, X_n) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)$$

dove:

- Φ è la funzione cumulata di una normale standardizzata.
- Y è la variabile dipendente dicotomica con una distribuzione bernoulliana ;
- $z = \beta_0 + \beta_1 X + \beta_2 X_2$ è il “z-value” o “z-index” del modello probit;
- β_1 è l'effetto su “z-valore” di un cambiamento di una unità in X_1 , tenendo costante X_2 .

Nel caso di modello Probit, a differenza del modello di regressione lineare che utilizza

il classico metodo OLS, per stimare i coefficienti β_i si usa la funzione di massimo verosimiglianza MLE che è la funzione di densità di Y_1, \dots, Y_n date X_1, \dots, X_n , trattata come una funzione di parametri sconosciuti β_0 e β_1 .

Come già descritto in precedenza i dati a disposizione costituiscono un panel, riguardano cioè diversi eventi osservati (contemporaneamente) in diversi istanti temporali.

Tutte i modelli di regressioni sono stati applicati sia per la Serie A, sia per la Serie B.

5.1.1 Modello empirico di partenza

Per il modello Probit scelto dunque sono state effettuate numerose prove, cercando di utilizzare tutte le variabili disponibili nel dataset con le trasformazioni in variabili di tipo dummy o logaritmiche come descritto in precedenza. Il primo modello semplificato non tiene conto di raggruppamento e interazioni tra variabili ed è composto da:

- Variabile dipendente: *DummyVittoriaCasa*
- Variabili indipendenti: *ln_diffvalori*, *ln_numerospettatori*, *densitastadio*, *ln_distanza*, *dummy_serale*.

In particolare si è partiti dal più semplice modello:

$$\Pr(DummyVittoriaCasa = 1 | ln_diffvalori, ln_numerospettatori) = \Phi(\beta_0 + \beta_1 ln_diffvalori + \beta_2 ln_numerospettatori)$$

In seguito si sono aggiunte tutte le altre variabili indipendenti.

5.1.2 Risultati del modello empirico di partenza

I risultati di questi primi modelli si possono osservare dettagliatamente nell'Appendice B dove sono riportati gli output completi di Stata. Le tabelle 5.1 e 5.2 riportano i risultati riassuntivi di tali regressioni rispettivamente per la Serie A e la Serie B.

Dall'analisi dei primi esiti, si è osservata un'alta collinearità tra le variabili *ln_numerospettatori* e *densitastadio*; per questo motivo non sono rappresentati modelli con entrambi le variabili presenti.

	dummyVittoriaCasa [1]	dummyVittoriaCasa [2]	dummyVittoriaCasa [3]	dummyVittoriaCasa [4]	dummyVittoriaCasa [5]	dummyVittoriaCasa [6]
ln_numerospettatori	0.002 (0.007)	0.003 (0.007)	0.001 (0.007)			
ln_diffvalori	0.465*** (0.029)	0.465*** (0.029)	0.467*** (0.029)	0.468*** (0.029)	0.466*** (0.029)	0.467*** (0.029)
ln_distanza		0.025 (0.031)	0.023 (0.031)	0.024 (0.031)	0.025 (0.031)	
dummy_serale			-0.100 (0.071)	-0.097 (0.071)		
densitastadio				0.052 (0.099)	0.069 (0.099)	0.065 (0.098)
Constant	-0.241*** (0.059)	-0.378** (0.181)	-0.320* (0.186)	-0.340* (0.185)	-0.395** (0.180)	-0.254*** (0.056)
Observations	1513	1513	1513	1513	1513	1513
R-squared						
Adjusted R-squared	0.137	0.138	0.139	0.139	0.138	0.138

Note: *** significatività all' 1%, ** significatività al 5%, * significatività al 10%

Tabella 5.1: i determinanti della probabilità di vittoria casalinga in Serie A (modello base)

Come si evince dalla Tabella 5.1, in Serie A i valori delle rose impattano in maniera dominante sulla probabilità di vittoria della squadra di casa. Dunque la variabile *ln_diffvalori* è significativa addirittura all'1% e non può mai essere estromessa dal modello, essendo pertanto una variabile di controllo. Tutte le altre variabili, invece, non sono significative per nessun livello e dal valore dei coefficienti si può osservare come la densità dello stadio sia più impattante sulla probabilità di

vittoria della squadra casa rispetto al numero di spettatori totali. È interessante osservare come vi sia, pur se non significativa, una correlazione positiva tra la distanza che la squadra ospite deve coprire per la trasferta e la probabilità di vittoria della squadra di casa; invece dai risultati sembrerebbe che negli incontri serali (dopo le 18.00) la probabilità di vittoria delle squadre di casa si riduca.

I risultati relativi alla Serie A non trovano conferma in Serie B dove, come già osservato dall'applicazione del modello di Clarke e Normann, il fattore campo sembra avere un ruolo più importante nei risultati.

	dummyVittoriaCasa [1]	dummyVittoriaCasa [2]	dummyVittoriaCasa [3]	dummyVittoriaCasa [4]	dummyVittoriaCasa [5]	dummyVittoriaCasa [6]
ln_numerospettatori	0.018** (0.008)	0.017** (0.008)	0.018** (0.008)			
ln_diffvalori	0.277*** (0.061)	0.277*** (0.061)	0.277*** (0.061)	0.284*** (0.061)	0.283*** (0.061)	0.283*** (0.061)
ln_distanza		0.018 (0.042)	0.019 (0.042)	0.025 (0.042)	0.025 (0.042)	
dummy_serale			0.058 (0.066)	0.056 (0.066)		
densitastadio				0.237* (0.133)	0.228* (0.132)	0.223* (0.132)
Constant	-0.307*** (0.057)	-0.412* (0.245)	-0.439* (0.247)	-0.438* (0.249)	-0.411* (0.247)	-0.267*** (0.050)
Observations	1564	1564	1564	1564	1564	1564
R-squared						
Adjusted R-squared	0.012	0.012	0.013	0.012	0.011	0.011

Tabella 5.2: i determinanti della probabilità di vittoria casalinga in Serie B (modello base)

Dalla Tabella 5.2 riepilogativa si può evincere come la correlazione tra la differenza dei valori della rosa e la probabilità di vittoria casalinga anche in Serie B sia significativa fino all'1%, con valori di beta comunque nettamente inferiore rispetto alla Serie A.

Nella competizione in esame, quindi, assumono significatività le variabili relative al pubblico presente allo stadio; in particolare *ln_numerospettatori* in tutti e tre i modelli è significativa al 5%, mentre *densitastadio* è significativa al 10%. Ciò significa che a parità di differenza dei valori tra le rose, l'aumento del numero di spettatori presenti e del loro "calore", impattano significativamente sull'aumento della

probabilità di vittoria casalinga.

Rispetto alla Serie A, anche se non significativa, la probabilità di vittoria della squadra di casa negli incontri serali è più alta (coefficiente di *dummy_serale* maggiore di 0) a parità di differenza dei valori delle rose.

È molto interessante osservare come si ottengano dei valori di R-quadro diversi tra le analisi della Serie A rispetto a quelle della serie B. In particolare tale valore è molto basso nei modelli della Serie B (0.011-0.012) rispetto a quello nei modelli della Serie A (0.137-0.138); questo parametro indica come la variazione della variabile dipendente del modello sia spiegata solamente per l'1% dal modello costruito per la Serie B e, invece, per il 13-14% dal modello della Serie A.

In seguito, per la Serie A nel modello Probit sono state inserite altre variabili indipendenti raccolte nel dataset per vedere se si riesce ad ottenere nuove relazioni con la probabilità di vittoria casalinga. In particolare le variabili aggiunte sono:

- *Anniesperienza*: variabile che conta gli anni consecutivi di presenza in Serie A della squadra di casa;
- *Neopromossa*: variabile dicotomica che indica se una squadra di casa è nuova nella categoria o meno;
- *Stadionuovo*: variabile dicotomica che assume valore 1 se la squadra di casa è il primo anno che disputa gli incontri in tale struttura.

	dummyVittoriaCasa [7]	dummyVittoriaCasa [8]
ln_numerospettatori	0.001 (0.007)	
ln_diffvalori	0.482*** (0.035)	0.483*** (0.035)
ln_distanza	0.022 (0.031)	0.023 (0.031)
dummy_serale	-0.095 (0.072)	-0.092 (0.071)
NEOPROMOSSA	0.191* (0.113)	0.192* (0.113)
ANNI ESPERIENZA	0.001 (0.002)	0.001 (0.002)
STADIO NUOVO	0.044 (0.328)	0.023 (0.330)
densitastadio		0.055 (0.100)
Constant	-0.357* (0.191)	-0.380** (0.191)
Observations	1513	1513
R-squared		
Adjusted R-squared	0.140	0.140

Tabella 5.3: i determinanti della probabilità di vittoria casalinga in Serie A (modello più strutturato)

In Tabella 5.3 sono riportati i risultati del modello Probit contenente anche le variabili appena descritte. È interessante osservare come gli anni di esperienza della squadra di casa non abbiano un'influenza significativa sulla probabilità di vittoria casalinga, così come la presenza del nuovo stadio (a differenza dello studio di Pollard (2002) che osservava un impatto negativo sui risultati della squadra di casa in presenza di un nuovo stadio). Invece, la variabile che classifica una squadra neopromossa in Serie A risulta significativa al 10% in entrambi i modelli e impatta positivamente sulla probabilità di vittoria di casa; ciò significa, e di nuovo è una conferma rispetto quanto osservato dal modello di Clark e Normann, che la presenza di una squadra neopromossa con i suoi match casalinghi aumenta la probabilità di vittoria a parità di differenza dei valori delle rose.

Inoltre, con l'aggiunta di queste variabili si può osservare un aumento del valore del R-quadro, a dimostrazione di un raggiungimento di maggiore robustezza del modello.

5.1.3 *Modello empirico che misura l'impatto del Covid-19*

Partendo dal modello base utilizzato per misurare l'impatto dei vari fattori sulla probabilità di vittoria casalinga, tale Probit è stato replicato per osservare se il Covid-19, che ha comportato come già detto la disputa degli incontri a porte chiuse, ha avuto un impatto significativo sui risultati.

Il modello utilizzato è una replica di quello descritto precedentemente; l'unica differenza è la variabile indipendente principale utilizzata, non più *ln_numerospettatori* o *densitastadio*, bensì la dummy che descrive il disputarsi di incontri a porte chiuse a causa della pandemia.

Dunque il modello presenterà la seguente struttura:

- Variabile dipendente: *DummyVittoriaCasa*
- Variabili indipendenti: *ln_diffvalori*, *booleanportechiuse*, *ln_distanza*, *dummy_serale*, *anniesperienza*, *neopromossa*, *stadionuovo*.

5.1.4 *Risultati del modello che misura l'impatto del Covid-19*

I risultati di questi modelli si possono osservare dettagliatamente nell'Appendice B; in Tabella 5.4 e Tabella 5.5 rappresentano le tabelle riepilogative dei modelli Probit rispettivamente per la Serie A e per la Serie B.

Così come nei modelli precedenti, in Serie A si è approfondita l'analisi con un modello Probit contenente le variabili indipendenti *Neopromossa*, *AnniEsperienza*, *StadioNuovo*.

Come era lecito aspettarsi, in tutti gli output si può osservare una collinearità

negativa tra la variabile *dummyPortechiuse* e la probabilità di vittoria casalinga; ciò significa che nelle stagioni caratterizzate dalla disputa degli incontri a porte chiuse a causa della pandemia, la probabilità di incontri con esiti favorevoli alla squadra di casa è diminuita.

	dummyVittoriaCasa [9]	dummyVittoriaCasa [10]	dummyVittoriaCasa [11]	dummyVittoriaCasa [12]
<i>dummyPortechiuse</i>	-0.031 (0.073)	-0.032 (0.073)	-0.016 (0.074)	-0.016 (0.074)
<i>ln_diffvalori</i>	0.466*** (0.029)	0.466*** (0.029)	0.467*** (0.029)	0.506*** (0.040)
<i>ln_distanza</i>		0.025 (0.031)	0.023 (0.031)	0.023 (0.031)
<i>dummy_serale</i>			-0.099 (0.071)	-0.097 (0.072)
NEOPROMOSSA (1:SI, 0:NO)				0.213* (0.115)
ANNI ESPERIENZA				0.001 (0.002)
STADIO NUOVO (1:SI, 0:NO)				0.057 (0.327)
Constant	-0.214*** (0.042)	-0.349** (0.174)	-0.307* (0.176)	-0.327* (0.182)
Observations	1513	1513	1513	1513
R-squared				
Adjusted R-squared	0.138	0.138	0.139	0.141

Tabella 5.4: Risultati modello che misura effetto del Covid-19 sulla probabilità di vittoria casalinga in Serie A

Più nel dettaglio, analizzando la Tabella 5.4 relativa alla competizione Serie A, si può osservare, come nei modelli precedenti, come l'impatto della presenza dei tifosi sulla probabilità di vittoria non sia significativa sulla probabilità di vittoria della squadra casalinga. In particolare, questi modelli sono interessanti in quanto consentono di dedurre che in Serie A, giocare a porte chiuse o aperte non ha impatto significativo sui risultati. Questo si può affermare osservando il valore del coefficiente della variabile *dummyPortechiuse* e il suo relativo errore. Tuttavia, confrontando tali coefficienti con i coefficienti delle variabili *ln_numerospettatori* e *densitàstadio* presenti negli altri modelli Probit della Serie A ([1], [2], [3], [4], [5], [6], [7], [8]), si può notare come l'assenza o la presenza dei tifosi abbia una correlazione maggiore con l'aumento della probabilità di vittoria casalinga rispetto a quella che si presenta all'aumentare del numero di tifosi o della densità dei tifosi allo stadio.

	dummyVittoriaCasa [7]	dummyVittoriaCasa [8]	dummyVittoriaCasa [9]
dummyPortechiuse	-0.137** (0.069)	-0.137** (0.069)	-0.142** (0.069)
ln_diffvalori	0.279*** (0.061)	0.279*** (0.061)	0.279*** (0.061)
ln_distanza		0.019 (0.042)	0.019 (0.042)
dummy_serale			0.058 (0.066)
Constant	-0.160*** (0.039)	-0.266 (0.243)	-0.288 (0.244)
Observations	1564	1564	1564
R-squared			
Adjusted R-squared	0.012	0.012	0.012

Tabella 5.5: risultati modello che misura effetto del Covid-19 sulla probabilità di vittoria casalinga in Serie B

Così come in Serie A, anche dai risultati riportati in Tabella 5.5 riguardanti i modelli della Serie B si può notare una continuità di risultati rispetto alle analisi precedenti. Infatti, osservando il coefficiente della variabile *dummyPortechiuse*, che presenta valori -0.137 e -0.142, si può dedurre una correlazione negativa significativa al 5% tra l'assenza dei tifosi allo stadio e la probabilità di vittoria casalinga.

Questa è l'ennesima dimostrazione di come in questa competizione il tifo possa influenzare i risultati delle partite; addirittura estremizzando, la presenza di anche soltanto poche centinaia di tifosi, fa aumentare la probabilità di vittoria della squadra locale.

Conclusioni

La presente dissertazione ha voluto osservare quanto nei campionati di Serie A e Serie B il fattore campo ha avuto impatto sui risultati nelle stagioni che vanno dal 2017/2018 al 2020/2021 e nel dettaglio quali sono le componenti che influiscono sull'Home Advantage.

Dalla prima analisi più macroscopica, effettuata a livello di stagione, è risultato molto più presente il fattore campo in Serie B rispetto alla Serie A. Questo è dovuto sicuramente al fatto che in tale competizione ci sia molta più competitività, dimostrata dal fatto che le squadre presentano un valore delle rose molto più equamente distribuito. Per questo motivo, visto il maggiore equilibrio tra le squadre sfidanti, il fattore campo risulta più decisivo in Serie B, dove si osservano valori dell'Home Advantage maggiori.

In Serie B, dunque, ci sono mediamente più vittorie casalinghe rispetto alla Serie A.

Dall'analisi di regressione si osserva in particolare che ciò che influenza significativamente il fattore campo, dove è presente, sono il numero di spettatori e il calore di essi nei confronti della squadra di casa. Invece, la probabilità di vittoria casalinga e quindi il fattore campo non sono influenzati da altri fattori come si poteva immaginare. Infatti, la distanza della trasferta che la squadra ospite deve coprire per disputare l'incontro non impatta significativamente; questo può essere dovuto al fatto che, come accennato all'inizio, ormai le squadre possono godere di viaggi comunque molto confortevoli.

Inoltre, una conclusione a cui si può giungere dall'analisi della Serie A è quella che le squadre neopromosse, che quindi risultano essere una nuova sfidante per le

squadre già presenti, presentano un vantaggio casalingo mediamente più consistente.

Alla fine di questo lavoro, è stato interessante osservare come l'applicazione di due diversi modelli abbiano portato a risultati concordanti e in particolare l'analisi multivariata di regressione ha spiegato i risultati ottenuti dal modello di Clarke e Normann.

Inoltre, si è potuto osservare come la pandemia che sta stravolgendo le nostre vite, abbia avuto un impatto anche sui risultati calcistici e in particolare in Serie B l'assenza dei tifosi allo stadio ha portato una drastica riduzione dell'impatto del fattore campo sulle vittorie casalinghe. Sperando a breve in un ritorno completo alla normalità, sarebbe interessante osservare se il mondo del calcio ritornerà ad avere risultati simili a quelli pre-Covid o se questo stravolgimento comporterà un cambiamento del trend.

Partendo dalla base di questa analisi, si potrebbe sviluppare uno studio più approfondito, andando ad inserire nel modello Probit ulteriori variabili riguardanti alcune caratteristiche dei vari stadi (ad esempio la presenza o meno della pista d'atletica) oppure alcune caratteristiche intrinseche delle squadre (ad esempio l'età media dei giocatori). Tutto questo con lo scopo di provare a spiegare con i numeri l'andamento dei risultati sportivi che da sempre risulta sembrare totalmente casuale.

Bigliografia

- Clarke, A. R. (1996). *Home advantage in balanced competitions - English soccer 1990-1996*. *Mathematics and Computers in Sport*, 111–116.
- Clarke, S. R. and Norman, J. M. (1995). *Home ground advantage of individual clubs in English soccer*. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 509–521.
- Courneya, K. S. and Carron, A. V. (1992). *The home advantage in sport competitions: a literature review*. *Journal of Sport & Exercise Psychology*, 14, 13–27.
- Gleitman, H., Gross, J., and Reisberg, D. (2010). *Psychology*. W. W. Norton & Company.
- Greer, D. L. (1983). *Spectator booing and the home advantage: A study of social influence in the basketball arena*. *Social Psychology Quarterly*, 46, 252–261.
- Jurkovic, T. (1995). *Collegiate basketball players' perceptions of the home advantage*. Unpublished Master's thesis, Bowling Green State University, Bowling Green, OH.
- Lefebvre, L. and Passer, M. (1974). *The effects of game location and importance on aggression in team sport*. *International Journal of Sport Psychology*, 5, 102–110.
- Maher, M. J. (1982). *Modelling association football scores*. *Statistica Neerlandica*, 36, 109–118.
- Monaghan, E. P. and Glickman, S. E. (1992). *Hormones and aggressive behaviour*. In Becker, J., Breedlove, A. and Crews, D. (Eds.) *Behavioural endocrinology*. MIT Press, 9, 261-85.
- Pollard, R. (1986). *Home advantage in soccer: a retrospective analysis*. *Journal of Sports Sciences*, 4, 237–48.
- Pollard, R. (2002). *Evidence of a reduced home advantage when a team moves to a*

- new stadium*. J Sports Sci, 20, 969–973.
- Pollard, R. (2006). *Home advantage in soccer: variations in its magnitude and a literature review of the inter-related factors associated with its existence*. Journal of Sports Behavior, 29, 169–189.
- Pollard, R. (2008). *Home advantage in football: A current review of an unsolved puzzle*. The Open Sports Sciences Journal, 1, 12–14.
- Pollard, R. and Pollard, G. (2005). *Long-term trends in home advantage in professional team sports in North America and England (1876-2003)*. J Sports Sci, 23, 337–350.
- Schwartz, B. and Barsky, S. F. (1977). *The home advantage*. Social Forces, 55, 641–661.
- Varca, P. E. (1980). *An analysis of home and away game performance of male college basketball teams*. Journal of Sport Psychology, 2, 245–257.
- Wolfson, S., Wakelin, D., and Lewis, M. (2005). *Football supporters' perceptions of their role in the home advantage*. J Sports Sci, 23, 365–374.
- Sors F., Grassi M., Agostini T., Murgia M. (2020). *The sound of silence in association football: Home advantage and referee bias decrease in matches played without spectators*. European Journal of Sport Science.
- Fischer K., Haucap J. (2020). *Does crowd support drive the home advantage in professional soccer? Evidence from German ghost games during the COVID-19 pandemic*. DICE Discussion Paper, No. 344.
- Wunderlich F, Weigelt M, Rein R, Memmert D (2021). *How does spectator presence affect football? Home advantage remains in European top-class football matches played without spectators during the COVID-19 pandemic*. PLoS ONE 16(3).
- Foulds G. (2016). *A Study of Home Advantage in Football and Other Contributions to*

Sports Data Analysis. Master's thesis, Lancaster University.

Titman A. C., Costain D. A., Ridall P. G. and Gregory K. (2015). *Joint modelling of goals and bookings in association football*. J. R. Statist. Soc. A 178, Part 3, 659–683.

Forrest D., Beaumont J., Goddard J. and Simmons R. (2005). *Home advantage and the debate about competitive balance in professional sports leagues*. Journal of Sports Sciences, 23:4, 439-445

Sitografia

<http://football-data.co.uk/data>.

<https://www.stadiapostcards.com/archivio.htm>

<https://www.transfermarkt.it/serie-a/marktwerteverein/wettbewerb/IT1>

<https://www.transfermarkt.it/serie-b/marktwerteverein/wettbewerb/IT1>

<https://www.distanza.org>

<https://www.calciofinanza.it/2021/09/04/juventus-impatto-ricavi-stadium-10-anni/>

https://www.figc.it/media/uploads/federazione/trasparenza/REPORT_CALCIO_2021_Alta.pdf

https://www.figc.it/media/uploads/federazione/trasparenza/report_calcio_2020_alta.pdf

https://www.figc.it/media/109475/2018_0673_rc_ita_bassa.pdf

https://www.figc.it/media/1128/report_calcio_2018_it.pdf

https://www.figc.it/media/1123/report_calcio_2017_it.pdf

https://www.figc.it/media/1124/report_calcio_2016_it.pdf

APPENDICE A

Tabella finale Serie A stagione 2017/2018

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Atalanta	9	6	4	30	18	7	6	6	27	21	60	12	6	0,073	0,531
Benevento	5	2	12	23	40	1	1	17	10	44	21	-17	-34	0,684	-1,500
Bologna	6	4	9	25	26	5	2	12	15	26	39	-1	-11	0,295	-0,331
Cagliari	6	3	10	18	30	5	3	11	15	31	39	-12	-16	-0,038	-0,564
Chievo	7	6	6	22	24	3	4	12	14	35	40	-2	-21	0,795	-0,856
Crotone	6	6	7	23	25	3	2	14	17	41	35	-2	-24	0,962	-1,014
Fiorentina	8	5	6	27	22	8	4	7	27	24	57	5	3	-0,149	0,392
Genoa	6	3	10	23	27	5	5	9	10	16	41	-4	-6	-0,149	-0,058
Inter	11	5	3	37	16	9	7	3	29	14	72	21	15	0,073	0,981
Juventus	16	1	2	45	8	14	4	1	41	16	95	37	25	0,406	1,464
Lazio	9	5	5	45	21	12	4	3	44	28	72	24	16	0,184	1,025
Milan	10	5	4	25	16	8	5	6	31	26	64	9	5	-0,038	0,486
Napoli	14	3	2	43	18	14	4	1	34	11	91	25	23	-0,149	1,392
Roma	11	2	6	31	19	12	6	1	30	9	77	12	21	-0,760	1,322
Sampdoria	12	3	4	36	20	4	3	12	20	40	54	16	-20	1,740	-0,853
Sassuolo	4	7	8	11	22	7	3	9	18	37	43	-11	-19	0,184	-0,725
Spal	5	8	6	22	29	3	6	10	17	30	38	-7	-13	0,073	-0,419
Torino	8	6	5	29	18	5	9	5	25	28	54	11	-3	0,518	0,058
Udinese	6	2	11	24	30	6	2	11	24	33	40	-6	-9	-0,094	-0,211
Verona	5	1	13	14	35	2	3	14	16	43	25	-21	-27	0,073	-1,119
Totale compi	164	83	133	553	464	133	83	164	464	553	1057	89	-89	4,684	0,000

Tabella finale Serie A stagione 2018/2019

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Atalanta	10	5	4	36	19	10	4	5	41	27	69	17	14	-0,152	0,994
Bologna	9	4	6	28	22	2	7	10	20	34	44	6	-14	0,792	-0,453
Cagliari	8	6	5	23	20	2	5	12	13	34	41	3	-21	1,015	-0,814
Chievo	1	7	11	14	37	1	7	11	11	38	20	-23	-27	-0,096	-1,058
Empoli	9	3	7	33	30	1	5	13	18	40	38	3	-22	1,070	-0,867
Fiorentina	5	9	5	25	19	3	8	8	22	26	41	6	-4	0,237	0,075
Frosinone	1	6	12	14	38	4	4	11	15	31	25	-24	-16	-0,763	-0,475
Genoa	6	8	5	21	22	2	6	11	18	35	38	-1	-17	0,570	-0,592
Inter	11	5	3	27	10	9	4	6	30	23	69	17	7	0,237	0,625
Juventus	15	4	0	39	13	13	2	4	31	17	90	26	14	0,348	0,969
Lazio	9	5	5	36	25	8	3	8	20	21	59	11	-1	0,348	0,219
Milan	12	4	3	31	17	7	5	24	19	68	14	5	0,181	0,528	
Napoli	13	4	2	41	16	11	3	5	33	20	79	25	13	0,348	0,919
Parma	5	7	7	18	25	5	4	10	23	36	41	-7	-13	0,015	-0,364
Roma	12	5	2	43	26	6	7	6	23	22	66	17	1	0,570	0,308
Sampdoria	10	2	7	28	16	5	6	8	32	35	53	12	-3	0,515	0,111
Sassuolo	5	10	4	33	33	4	6	9	20	27	43	0	-7	0,070	-0,067
Spal	5	7	7	20	26	6	2	11	24	30	42	-6	-6	-0,319	0,003
Torino	12	2	5	32	19	4	13	2	20	18	63	13	2	0,292	0,372
Udinese	8	5	6	22	22	3	5	11	17	31	43	0	-14	0,459	-0,436
Totale	166	108	106	564	455	106	108	166	455	564	1032	109	-109	5,737	0,000

Tabella finale Serie A stagione 2019/2020

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Atalanta	12	2	5	51	26	11	7	1	47	22	78	25	25	-0,228	1,467
Bologna	5	8	6	25	29	7	3	9	27	36	47	-4	-9	0,050	-0,247
Brescia	3	5	11	21	34	3	2	14	14	45	25	-13	-31	0,772	-1,383
Cagliari	7	4	8	32	28	4	8	7	20	28	45	4	-8	0,439	-0,217
Fiorentina	5	8	6	22	22	7	5	7	29	26	49	0	3	-0,395	0,375
Genoa	7	1	11	24	31	3	8	8	23	42	39	-7	-19	0,439	-0,767
Inter	11	6	2	40	17	13	4	2	41	19	82	23	22	-0,173	1,314
Juventus	16	2	1	46	17	10	3	6	30	26	83	29	4	1,161	0,347
Lazio	14	3	2	46	17	10	3	6	33	25	78	29	8	0,939	0,558
Lecce	4	5	10	30	41	5	3	11	22	44	35	-11	-22	0,383	-0,914
Milan	9	6	4	31	20	10	3	6	32	26	66	11	6	0,050	0,503
Napoli	10	3	6	31	24	8	5	6	30	26	62	7	4	-0,061	0,408
Parma	7	2	10	26	24	7	5	7	30	33	49	2	-3	0,050	0,053
Roma	10	4	5	36	27	11	3	5	41	24	70	9	17	-0,673	1,089
Sampdoria	6	4	9	25	30	6	2	11	23	35	42	-5	-12	0,161	-0,403
Sassuolo	8	3	8	44	33	6	6	7	25	30	51	11	-5	0,661	-0,078
Spal	2	4	13	16	40	3	1	15	11	37	20	-24	-26	-0,117	-1,089
Torino	7	4	8	23	30	4	3	12	23	38	40	-7	-15	0,216	-0,556
Udinese	6	6	7	18	23	6	3	10	19	28	45	-5	-9	-0,006	-0,244
Verona	9	5	5	29	25	3	8	8	18	26	49	4	-8	0,439	-0,217
Totale complessivo	158	85	137	616	538	137	85	158	538	616	1055	78	-78	4,105	0,000

Tabella finale Serie B stagione 2020/2021

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Atalanta	12	3	4	49	24	11	6	2	41	23	78	25	18	0,167	1,092
Benevento	2	7	10	19	39	5	5	9	21	36	33	-20	-15	-0,500	-0,525
Bologna	7	5	7	33	33	3	6	10	18	32	41	0	-14	0,556	-0,528
Cagliari	5	4	10	22	32	4	6	9	21	27	37	-10	-6	-0,444	-0,078
Crotone	5	2	12	23	35	1	3	15	22	57	23	-12	-35	1,056	-1,603
Fiorentina	6	6	7	25	25	3	7	9	22	34	40	0	-12	0,444	-0,422
Genoa	5	7	7	29	30	5	5	9	18	28	42	-1	-10	0,278	-0,314
Inter	17	1	1	53	18	11	6	2	36	17	91	35	19	0,667	1,117
Juventus	14	2	3	40	18	9	7	3	37	20	78	22	17	0,056	1,047
Lazio	13	3	3	36	23	8	2	9	25	32	68	13	-7	0,889	-0,194
Milan	8	6	5	31	24	16	1	2	43	17	79	7	26	-1,278	1,564
Napoli	12	4	3	50	20	12	1	6	36	21	77	30	15	0,611	0,919
Parma	2	5	12	16	39	1	6	12	23	44	20	-23	-21	-0,333	-0,833
Roma	13	4	2	42	18	5	5	9	26	37	63	24	-11	1,722	-0,436
Sampdoria	9	3	7	32	26	6	4	9	20	28	52	6	-8	0,556	-0,228
Sassuolo	7	8	4	31	27	10	3	6	33	29	62	4	4	-0,222	0,411
Spezia	5	7	7	28	33	4	5	10	24	39	39	-5	-15	0,333	-0,567
Torino	3	9	7	25	36	4	7	8	25	33	37	-11	-8	-0,389	-0,181
Udinese	5	4	10	14	19	5	6	8	28	39	40	-5	-11	0,111	-0,356
Verona	5	7	7	20	23	5	6	8	23	25	43	-3	-2	-0,278	0,114
Totale complessivo	155	97	128	618	542	128	97	155	542	618	1043	76	-76	4,000	0,000

Tabella finale Serie B stagione 2017/2018

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Accoli	6	8	7	23	24	5	5	11	17	36	46	-1	-19	0,562	-0,582
Avellino	8	7	6	29	26	3	8	10	20	34	48	3	-14	0,512	-0,352
Bari	13	5	3	33	18	5	8	8	26	30	67	15	-4	0,612	0,098
Brescia	6	9	6	23	22	5	6	10	18	30	48	1	-12	0,312	-0,252
Carpi	8	9	4	23	19	4	7	10	9	27	52	4	-18	0,762	-0,545
Cesena	9	9	3	35	26	2	8	11	20	35	50	9	-15	0,862	-0,414
Cittadella	7	7	7	30	25	11	5	5	31	23	66	5	8	-0,488	0,693
Cremonese	6	9	6	30	28	3	12	6	18	19	48	2	-1	-0,188	0,270
Empoli	14	6	1	48	23	10	7	4	40	26	85	25	14	0,212	0,934
Foggia	8	6	7	30	30	8	4	9	36	38	58	0	-2	-0,238	0,227
Frosinone	12	7	2	40	24	7	8	6	25	23	72	16	2	0,362	0,382
Novara	5	6	10	20	28	5	8	8	22	24	44	-8	-2	-0,638	0,245
Palermo	11	7	3	33	16	7	10	4	26	23	71	17	3	0,362	0,427
Parma	12	6	3	34	13	9	3	9	23	24	72	21	-1	0,762	0,227
Perugia	11	4	6	36	27	5	8	8	31	31	60	9	0	0,112	0,302
Pescara	7	7	7	30	31	4	8	9	20	33	48	-1	-13	0,262	-0,295
Pro Vercelli	6	8	7	22	27	3	5	13	25	43	40	-5	-18	0,312	-0,525
Salernitana	7	9	5	28	27	4	9	8	23	31	51	1	-8	0,112	-0,061
Spezia	10	7	4	29	15	3	7	11	17	30	53	14	-13	0,012	-0,330
Terana	5	10	6	32	31	2	6	13	30	46	37	1	-16	0,512	-0,443
Venezia	12	6	3	28	13	5	10	6	28	29	67	15	-1	0,462	0,241
Virtus Entell	6	7	8	26	27	4	7	10	15	27	44	-1	-12	0,212	-0,248
Totale compi	189	159	114	662	520	114	159	189	520	662	1227	142	-142	6,76190476	0,000

Tabella finale Serie B stagione 2018/2019

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Ascoli	7	7	4	25	24	3	6	9	15	32	43	1	-17	0,572	-0,489
Benevento	10	3	5	37	24	7	6	5	24	21	60	13	3	0,101	0,588
Brescia	13	3	2	39	19	5	10	3	30	23	67	20	7	0,278	0,789
Carpi	3	6	9	19	30	4	2	12	20	37	29	-11	-17	-0,134	-0,452
Cittadella	9	5	4	38	20	3	10	5	11	18	51	18	-7	0,984	0,015
Cosenza	7	8	3	20	15	4	5	9	14	27	46	5	-13	0,572	-0,279
Cremonese	9	6	3	19	6	3	7	8	18	27	49	13	-9	0,807	-0,080
Crotone	6	8	4	24	18	5	2	11	16	24	43	6	-8	0,337	-0,003
Foggia	8	8	2	31	23	2	5	11	13	26	43	8	-13	0,748	-0,288
Lecce	13	4	1	38	15	6	5	7	28	30	66	23	-2	0,984	0,279
Livorno	7	5	6	18	18	2	7	9	20	33	39	0	-13	0,278	-0,263
Padova	3	10	5	14	15	2	6	10	22	34	31	-1	-12	0,160	-0,204
Palermo	8	9	1	31	16	8	6	4	26	22	63	15	4	0,160	0,638
Perugia	9	2	7	28	22	5	6	7	21	27	50	6	-6	0,219	0,108
Pescara	11	5	2	28	17	3	8	7	22	29	55	11	-7	0,572	0,037
Salernitana	8	3	7	26	25	2	5	11	15	32	38	1	-17	0,572	-0,489
Spezia	10	5	3	31	18	4	4	10	22	28	51	13	-6	0,631	0,087
Venezia	6	8	4	22	22	2	6	10	13	24	38	0	-11	0,160	-0,152
Verona	8	7	3	30	22	5	6	7	19	24	52	8	-5	0,278	0,158
Totale compi	155	112	75	518	369	75	112	155	369	518	914	149	-149	8,2777778	0,00

Tabella finale Serie B stagione 2019/2020

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Ascoli	9	5	5	29	24	4	2	13	21	34	46	5	-13	0,670	-0,386
Benevento	14	4	1	39	9	12	4	3	28	18	86	30	10	0,781	0,758
Chievo	10	4	5	29	18	4	10	5	19	20	56	11	-1	0,336	0,231
Cittadella	7	4	8	23	26	10	3	6	26	23	58	-3	3	-0,664	0,481
Cosenza	6	5	8	21	22	6	5	8	29	27	46	-1	2	-0,497	0,422
Cremonese	7	6	6	22	19	5	7	7	20	24	49	3	-4	0,058	0,094
Crotone	12	5	2	33	14	8	3	8	30	26	68	19	4	0,503	0,472
Empoli	7	8	4	26	23	7	4	8	21	25	54	3	-4	0,058	0,094
Frosinone	10	5	4	29	18	4	7	8	12	20	54	11	-8	0,725	-0,139
Juve Stabia	9	3	7	30	29	2	5	12	17	34	41	1	-17	0,670	-0,586
Livorno	3	2	14	19	42	2	4	13	11	25	21	-23	-14	-0,830	-0,361
Perugia	7	5	7	19	19	5	4	10	19	30	45	0	-11	0,281	-0,267
Pescara	7	8	4	31	23	5	1	13	17	32	45	8	-15	0,947	-0,500
Pisa	10	5	4	28	20	4	7	8	21	25	54	8	-4	0,336	0,081
Pordenone	11	5	3	25	12	5	5	9	23	34	58	13	-11	1,003	-0,303
Salernitana	10	6	3	35	23	4	4	11	18	27	52	12	-9	0,836	-0,194
Spezia	10	4	5	30	18	7	6	6	24	22	61	12	2	0,225	0,386
Trapani	8	6	5	27	20	3	7	9	21	40	46	7	-19	1,114	-0,708
Venezia	5	6	8	18	24	7	8	4	19	16	50	-6	3	-0,830	0,489
Virtus Entell	8	6	5	21	18	4	6	9	25	32	48	3	-7	0,225	-0,064
Totale compi	170	102	108	534	421	108	102	170	421	534	1038	113	-113	5,94736842	0,00

Tabella finale Serie B stagione 2020/2021

Squadra	Casa					Trasferta					Punti	HGD	AGD	h	u
	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti	Vittorie	Pareggi	Sconfitte	Goal fatti	Goal subiti					
Ascoli	8	5	6	18	19	3	6	10	19	29	44	-1	-10	0,275	-0,311
Brescia	9	5	5	36	27	6	6	7	25	26	56	9	-1	0,330	0,136
Chievo	10	5	4	29	14	4	9	6	21	23	56	15	-2	0,719	0,067
Cittadella	9	7	3	25	14	6	5	8	23	21	57	11	2	0,275	0,289
Cosenza	3	9	7	15	20	3	8	8	14	27	35	-5	-13	0,219	-0,458
Cremonese	7	6	6	27	22	5	6	8	19	22	48	5	-3	0,219	0,042
Empoli	11	8	3	38	15	8	8	3	30	20	73	23	10	0,497	0,678
Frosinone	4	9	6	16	21	8	5	6	22	21	50	-5	1	-0,558	0,281
Lecce	6	9	4	36	28	10	5	4	32	19	62	8	13	-0,503	0,878
Monza	9	6	4	27	16	8	7	4	24	17	64	11	7	-0,003	0,553
Pescara	4	5	10	18	29	3	6	10	11	31	32	-11	-20	0,275	-0,811
Pisa	7	10	2	26	19	4	5	10	28	40	48	7	-12	0,830	-0,439
Pordenone	5	10	4	27	19	5	5	9	13	20	45	8	-7	0,608	-0,178
Reggiana	6	5	8	19	23	3	2	14	12	34	34	-4	-22	0,775	-0,936
Reggina	8	4	7	20	22	4	10	5	22	23	50	-2	-1	-0,281	0,167
Salernitana	11	6	2	26	13	8	6	5	20	21	69	13	-1	0,553	0,125
Spal	8	6	5	25	19	6	8	5	19	23	56	6	-4	0,330	-0,014
Venezia	10	3	6	28	19	5	11	3	25	20	59	9	5	-0,003	0,453
Vicenza	5	9	5	22	24	6	6	7	26	29	48	-2	-3	-0,170	0,061
Virtus Entell	3	4	12	17	35	1	7	11	13	29	23	-18	-16	-0,336	-0,581
Totale compi	143	131	106	495	418	106	131	143	418	495	1009	77	-77	4,05263158	0,00

