



FEDERAZIONE ITALIANA GIUOCO CALCIO
SETTORE TECNICO

Corso Match Analyst - Stagione Sportiva 2021/2022

**EURO 2020: SVILUPPO DI
METRICHE TECNICO-TATTICHE ED
ANALISI DESCRITTIVE CON I DATI
POSIZIONALI STATSOMB360**

Corsista: FEDERICO NINIVAGGI

INDICE

PREMESSA AL LAVORO	3
DATI UTILIZZATI	4
ELABORAZIONI STATISTICHE	8
VISUALIZZAZIONE	14
ANALISI DESCRITTIVE	18
CONCLUSIONI	27

PREMESSA AL LAVORO

Quello che verrà presentato è un lavoro che vuole essere espressione anche della mia storia personale con il calcio e i dati. Da sempre appassionato dello sport, pur non avendolo mai praticato se non per brevissimi periodi, l'interesse per il calcio ha subito un'accelerazione nei miei anni di studi universitari e poi nel primo periodo lavorativo. Il contributo del percorso di studi, che tutt'ora porto avanti come dottorando di ricerca, e dell'esperienza lavorativa è stato quello di avermi fornito gli strumenti utili per esplorare in prima persona il mondo dei dati legati al calcio. Le tecniche di analisi statistica e di programmazione apprese in ambito economico sono state riversate anche nello studio dello sport che da sempre mi appassiona. Ne è nato un esperimento sui social network (@mxd_zone) legato all'analisi di squadre e giocatori partendo dai dati resi pubblici dal sito FBREF.

Il lavoro di tesina in occasione della prova di ambito statistico per l'abilitazione a Match Analyst è l'occasione per cimentarmi con un nuovo tipo di dati e portare l'integrazione delle mie conoscenze di programmazione e statistiche ad un nuovo livello. Con questa motivazione nasce il lavoro che viene presentato. Una sfida verso me stesso a cercare di affrontare nuove basi dati e cominciare ad esplorarne il potenziale applicativo verso una migliore comprensione del gioco.

DATI UTILIZZATI

Il lavoro sviluppato vuole essere quindi un personale primo approccio ai dati di tracking posizionale. Sono stati analizzati i dati di questo tipo offerti gratuitamente da StatsBomb denominati StatsBomb360 riferiti all'ultimo Europeo, Euro2020, di calcio vinto dalla Nazionale Italiana nell'estate del 2021.

StatsBomb 360 offre per ogni evento raccolto il dato posizionale sia per l'attore dell'evento sia per compagni e avversari in zona "palla". Si riescono così ad avere le coordinate geografiche x-y dei giocatori in un certo intorno dell'attore protagonista dell'evento registrato, l'area visibile ha carattere variabile ma mai raccoglie tutti e 22 i giocatori.

Tutte le elaborazioni, sia quelle preliminari sui dati StatsBomb che quelle successive per la produzione delle grafiche sono state realizzate con il linguaggio di programmazione R nell'ambiente di sviluppo specifico per l'analisi statistica dei dati.

Nel dataset StatsBomb ogni riga corrisponde ad un evento. Per ogni evento raccolto si hanno a disposizione numerose informazioni, sia di carattere generale, sia proprie degli specifici eventi queste ultime in grado di fornire maggior dettaglio dell'evento oltre a descriverne anche l'outcome. Tra le informazioni di carattere generali si trovano il minuto e il secondo della partita in cui l'evento si è avvenuto, l'identificativo della partita a cui l'evento si riferisce, l'identificativo univoco e la sua posizione nella sequenza ordinata di sviluppo degli eventi per ciascuna partita. Altri attributi generali utili all'identificazione e classificazione degli eventi sono quelli riferiti alla squadra cui appartiene il giocatore protagonista, quella avversaria, il nome del giocatore, la sua posizione sia secondo coordinate x-y che secondo il ruolo ricoperto e la durata dell'evento.

Ogni osservazione nei dati StatsBomb corrisponde ad un evento, ognuno viene classificato a seconda del "Tipo" cui appartiene. Di seguito l'elenco dei tipi di evento raccolti dall'azienda.

Event Type	Event Type ID	Event Description	Average Percent of Events
50/50	33	2 players challenging to recover a loose ball.	0.05%
Bad Behaviour	24	When a player receives a card due to an infringement outside of play.	0.02%
Ball Receipt*	42	The receipt or intended receipt of a pass.	25.84%
Ball Recovery	2	An attempt to recover a loose ball	2.95%
Block	6	Blocking the ball by standing in its path.	1.08%
Camera off: Deprecated in version 1.03	29	Video broadcast leaves gameplay and show a replay or video is cut.	
Camera On: Deprecated in version 1.03	5	Signals the stop of the camera to capture gameplay for a replay/video cut.	
Carry	43	A player controls the ball at their feet while moving or standing still.	23.11%
Clearance	9	Action by a defending player to clear the danger without an intention to deliver it to a teammate.	0.92%
Dispossessed	3	Player loses ball to an opponent as a result of being tackled by a defender without attempting a dribble	0.68%
Dribble	14	An attempt by a player to beat an opponent	0.94%
Dribbled Past	39	Player is dribbled past by an opponent.	0.62%
Duel	4	A duel is an 50-50 contest between two players of opposing sides in the match.	1.37%
Error	37	When a player is judged to make an on-the-ball mistake that leads to a shot on goal.	0.02%
Foul Committed	22	Any infringement that is penalised as foul play by a referee. Offside are not tagged as a foul committed.	0.84%
Foul Won	21	A foul won is defined as where a player wins a free-kick or penalty for their team after being fouled by an opposing player.	0.79%
Goal Keeper	23	Actions that can be done by the goalkeeper.	0.95%
Half End	34	Signals the referee whistle to finish a match part.	0.12%
Half Start	18	Signals referee whistle to start a match period.	0.12%
Injury Stoppage	40	A stop in play due to an injury.	0.1%
Interception	10	Preventing an opponent's pass from reaching their teammates by moving to the passing lane/reacting to intercept it.	0.7%
Miscontrol	38	Player loses ball due to bad touch	0.79%

Offside	8	Offside infringement. Cases resulting from a shot or clearance (non-pass). For passes resulting in an offside check pass outcomes section	0.02%
Own Goal Against	20	An own goal scored against the team	0%
Own Goal For	25	An own goal scored for the team.	0%
Pass	30	An intended kick from one player to his teammate	28.71%
Player Off	27	A player goes/ is carried out of the pitch without a substitution.	0.03%
Player On	26	A player returns to the pitch after a Player Off event.	0.03%
Pressure	17	Field Position and player 1 must always have a value (Field position shows the spot where the player started to apply pressure & the duration of the event shows the duration that he kept applying pressure)	8.1%
Referee Ball-Drop	41	Referee drops the ball to continue the game after an injury stoppage.	0.02%
Shield	28	Player shields ball going out of bounds to prevent opponent from keeping it in play.	0.03%
Shot	16	An attempt to score a goal, made with any (legal) part of the body.	0.79%
Starting XI	35	Indicates the players in the starting 11, their position and the team's formation.	0.06%
Substitution	19	A player is substituted off the field for various reasons.	0.17%
Tactical Shift	36	Indicates a tactical shift made by the team, shows the players' new positions and the team's new formation.	0.05%

Gli attributi specifici vengono quindi registrati a seconda del tipo di evento. Il lavoro che sarà presentato sarà focalizzato sui passaggi. È quindi di seguito riportata la tabella riferita agli attributi specifici raccolti da StatsBomb rispetto all'evento passaggio.

Attribute	Attribute Type	Attribute Description	Attribute Values	Value Descriptions	Constraints
recipient	object (id (integer) / name (text))	Id / Name specifying the player id of the recipient of the pass, or for whom an incomplete pass was intended.	e.g., 2934 / "Erik Lamela"		Recorded whenever the recipient of the pass is clear, regardless of whether or not the pass was complete.
length	decimal	The length in yards of the pass, from its origin to its destination.	e.g., 16.03		Recorded for all passes.
angle	decimal	The angle of the pass in radians, with 0 pointing towards the team's attacking goal, positive values between 0 and π indicating an angle clockwise, and negative values between 0 and $-\pi$ representing an angle anti-clockwise.	e.g., -2.49		Recorded for all passes.
height	object (id (integer) / name (text))	Id / Name specifying the height of the pass.	1 / "Ground Pass"	Ball doesn't come off the ground.	Recorded for all passes.
			2 / "Low Pass"	Ball comes off the ground but is under shoulder level at peak height.	
			3 / "High Pass"	Ball goes above shoulder level at peak height.	
end_location	array [x,y]	Array containing two integer values. These are the x and y coordinates at which the pass ended.	e.g., (100,45)		Recorded for all passes.
assisted_shot_id	uuid	Reference to the shot this pass assisted.	e.g. "0052d1b5-e2b0-4629-bbea-c18c884ab103"	See Appendix 2 for more information on location coordinates.	Recorded for all passes that assist a shot.
backheel	boolean	Added if the pass was made by using a backheel.	TRUE		Recorded only when true.

deflected	boolean	Added if the pass was deflected.	TRUE	Pass was deflected, can occur on complete and incomplete passes.	Recorded only when true.
miscommunication	boolean	Added if the pass was a miscommunication.	TRUE	Pass was accurate but receiving player misread the pass OR the run from the receiving player was correct and passer played in his previous position	Recorded only when true.
cross	boolean	Added if the pass was a cross.	TRUE	See Appendix 6 for more on how crosses are defined	Recorded only when true.
cut_back	boolean	Added if the pass was a "cut back" (ball passed low backwards, within the opposition's penalty box).	TRUE	See Appendix 5 for more information on how cut backs are depicted.	Recorded only when true.
switch	boolean	Added if the pass was a switch (ball transitioned at least 50% of the pitch vertically).	TRUE	Whether or not the passer switched the ball to the other side of the pitch. A switch is any pass that travels more than 40 yards of the width of the pitch.	Recorded only when true.
shot_assist	boolean	Added if the pass was an assist to a shot (that did not score a goal).	TRUE		Recorded only when true.
goal_assist	boolean	Added if the pass was an assist to a goal.	TRUE		Recorded only when true.
body_part	object (id (integer) / name (text))	Id / Name of the body part used to make this pass.	68 / "Drop Kick"	Pass is a keeper drop kick	Recorded for most passes.
			37 / "Head"	Pass with head	
			69 / "Keeper Arm"	Pass thrown from keepers hands	
			38 / "Left Foot"	Pass with left foot	
			70 / "Other"	Other body part (chest,back, etc.)	
			40 / "Right Foot"	Pass with right foot	
			106 / "No Touch"	A player deliberately let the pass go past him instead of receiving it to deliver it to a teammate behind him. (Also known as a "dummy").	
type	object (id (integer) / name (text))	Id / Name for the attribute option specifying the type of pass.	61 / "Corner"	A pass from a corner kick	Recorded for all passes with one of the special types listed to the left. Passes missing a "type" are
			62 / "Free Kick"	A pass from a free kick	
			63 / "Goal Kick"	A pass from a goal kick	
			64 / "Interception"	One touch pass off an interception	

			65 / "Kick Off"	A pass from a kick off at the beginning of a match or after scoring	standard passes in the run of play.
			66 / "Recovery"	One touch pass off a loose ball recovery	
			67 / "Throw-in"	A pass from a throw-in	
outcome	object (id (integer) / name (text))	ID / Name of the outcome of the pass.	9 / "Incomplete"	Ball does not reach a teammate and is still in play	Recorded for all passes that are incomplete for a specific reason or their outcome is unknown. Passes missing an "outcome" are completed passes.
			74 / "Injury Clearance"	Ball is played out of bounds to stop play for an injury	
			75 / "Out"	Ball goes out of bounds	
			76 / "Pass Offside"	Ball reaches teammate but pass is judged offside	
			77 / "Unknown"	Outcome is unknown (i.e. foul was called while in mid-flight)	
Technique	object (id (integer) / name (text))	ID / Name of the technique of the pass.	104 / "Inswinging"	For inswinging corner kicks.	Recorded only for corner kicks of low or high height.
			105 / "Outswinging"	For outswinging corner kicks.	
			107 / "Straight"	For neither inswinging nor outswinging corner kicks.	
			108 / "Through Ball"	Pass cuts last line of defence	
					Recorded only when true.

A queste informazioni StatsBomb360 aggiunge la posizione di tutti i giocatori localizzati nell'area visibile per ogni evento raccolto.

Dall'integrazione dei dati evento riferiti ai soli passaggi con quelli posizionali forniti dalla tecnologia StatsBomb360 è stato sviluppato il lavoro. L'idea è di costruire misure in grado di cogliere aspetti spaziali riferiti alle condizioni di gioco che potrebbero essere determinanti sia nell'analisi della scelta fatta dal giocatore che effettua il passaggio, sia rispetto ad aspetti tattici della partita. A questo proposito sono stati quindi analizzati tutti i passaggi effettuati nel corso dell'Europeo, sviluppando su di essi alcune elaborazioni al fine di ottenere informazioni e metriche non riportate nel dataset di partenza. Si è anche sviluppata una visualizzazione per ogni evento in grado di esprimere le elaborazioni svolte e in ultimo le nuove misure sono state aggiunte alle già raccolte da StatsBomb e su di esse sono state sviluppate analisi descrittive.

Il lavoro seguirà nella struttura quanto scritto nelle righe precedenti. Ad una prima sezione descrittiva delle elaborazioni svolte ne seguirà una di presentazione della visualizzazione realizzata e poi una inerente alle analisi descrittive condotte. Il lavoro si concluderà con considerazioni finali circa possibili sviluppi futuri del lavoro.

ELABORAZIONI STATISTICHE

Tre sono le direttrici seguite dalle elaborazioni statistiche implementate. Per ogni passaggio l'interesse era quello di poter calcolare:

- 1- Lo spazio imputabile sotto il controllo della squadra in possesso e in maniera complementare quello sotto il controllo della difesa all'interno della zona di campo resa visibile
- 2- Visualizzare e conteggiare le linee di passaggio libere disponibili al giocatore in possesso tra quelle calcolabili nella finestra di campo resa visibile
- 3- Disegnare e contare i difendenti bypassati dal passaggio effettuato

IL CONTROLLO DELLO SPAZIO

Uno degli elementi centrali del calcio dal punto di vista tattico è il controllo dello spazio e come questo viene modificato dai movimenti dei giocatori delle due squadre. Con l'obiettivo di quantificare l'area sotto il controllo delle due squadre è stata implementata una tecnica nota come tassellazione di Voronoi. Si tratta di una tecnica di decomposizione dello spazio determinata dalle distanze rispetto ad un determinato insieme discreto di elementi dello spazio. Nel caso di un piano, come in questo caso, dato un insieme finito di punti, i giocatori nell'area visibile, la partizione del piano associa una regione di campo ad ogni giocatore, in modo tale che tutti i punti all'interno del perimetro della regione determinata siano più vicini al giocatore in questione che ad ogni altro. I punti dello spazio che compongono i perimetri delle regioni sono quindi quelli equidistanti tra due giocatori.

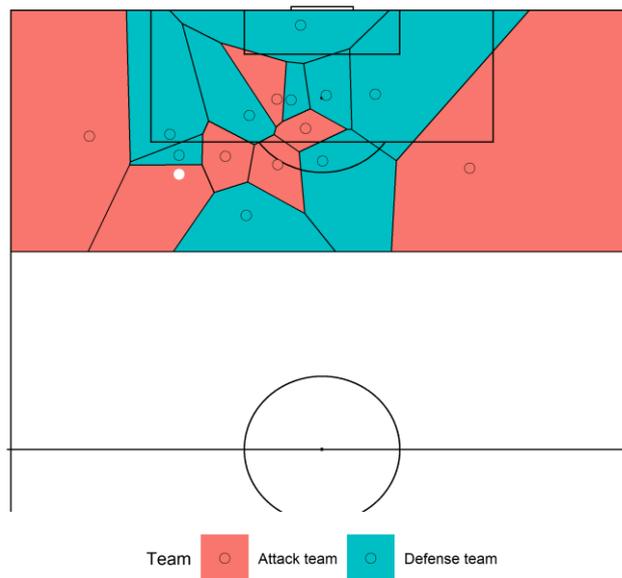
La tecnica di tassellazione di Voronoi considera solamente l'aspetto spaziale, in termini di distanze, tralasciando del tutto questioni relative ai singoli giocatori di natura tecnica e atletica.

Una volta calcolata per ogni giocatore la regione in controllo, sono state sommate le aree dei giocatori ed ottenute quelle di squadra, in possesso e difendente. Successivamente sono state ricavate le percentuali controllate per squadra sul totale dell'area coperta nel frame. Al dataset finale, sul quale saranno sviluppate le analisi descrittive, è stata anche aggiunta l'area, espressa in valore assoluto, controllata dal giocatore in possesso di palla.

Di seguito un esempio di tassellazione di Voronoi su un passaggio effettuato da Chiellini. Il giocatore che effettua il passaggio è colorato in bianco, gli altri giocatori con il rispettivo colore assegnato alla squadra. In tonalità di rosso la squadra in possesso palla e in blu quella difendente. Per ogni giocatore l'algoritmo implementato con funzioni R calcola l'area da assegnare a ciascun giocatore definendo l'immagine finale.

Giorgio Chiellini Pass 36:2

Action: Pass
Italy-Turkey, UEFA EURO 2020
xG:NA, Outcome: NA



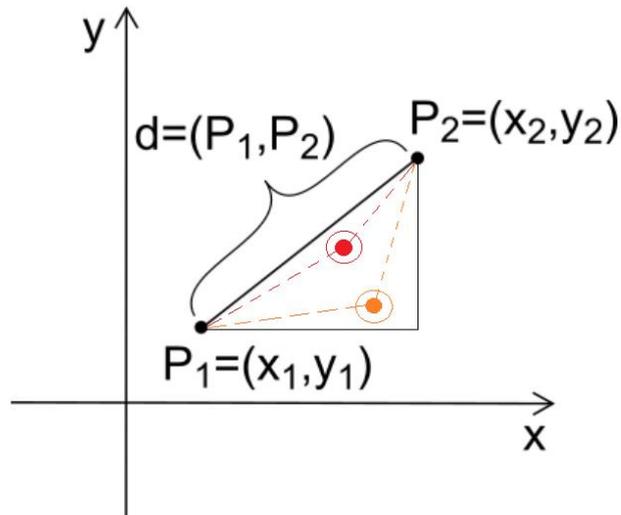
Data from StatsBomb
<https://github.com/statsbomb/open-data>
By: @mxd zone, Match & Data

LE LINEE DI PASSAGGIO DISPONIBILI

La seconda elaborazione svolta riguarda la possibilità di visualizzare e contare le linee di passaggio che il possessore di palla ha a disposizione. In un'ottica di analisi le informazioni ricavabili in questo contesto potrebbero risultare utili nella valutazione delle decisioni prese dai singoli.

Il primo passo è stato quello di calcolare, e disegnare, la distanza euclidea tra il giocatore in possesso di palla e tutti i compagni nell'area. Nel calcolare la distanza tra giocatori è possibile visualizzare anche le rispettive linee di passaggio. L'idea di linea di passaggio libera o chiusa discende dall'esistenza o meno di un avversario posizionato in maniera tale da intercettare il passaggio. Per spiegare come si è implementata la regola a livello di codice nell'analisi è d'aiuto l'immagine che segue.

$$P_1 = (x_1, y_1) \quad ; \quad P_2 = (x_2, y_2)$$



Considerando P1 e P2 due compagni di squadra, il segmento che li congiunge definito d corrisponde alla distanza euclidea tra i due punti nonché alla linea di passaggio tra i compagni. Per poterla calcolare è necessario sviluppare il teorema di Pitagora, identificando la distanza tra i due giocatori come l'ipotenusa di un triangolo rettangolo.

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Si considerino ora i due punti di colore *Arancione* e *Rosso* come due avversari. La condizione inizialmente impostata per definire una linea di passaggio come libera o chiusa è stata così scritta:

$$\text{distanza [P1, AVV 1]} + \text{distanza [AVV 1, P2]} == \text{distanza [P1, P2]}$$

Se la somma tra le distanze tra l'avversario e i due giocatori della squadra in possesso è uguale alla distanza d tra i due giocatori P1 e P2, allora l'avversario considerato può dirsi nella linea di passaggio tra P1 e P2. All'implementazione di questa regola però sia l'avversario *Rosso* che quello *Arancione* risulterebbero entrambi in grado di intercettare e chiudere la linea di passaggio tra P1 e P2. Questo perché pur essendo posizionati in maniera diversa rispetto all'ipotenusa, linea di passaggio d , entrambi ricadono nello spazio interno al triangolo rettangolo costruito per calcolare la distanza tra P1 e P2. Dato il loro posizionamento interno al triangolo il calcolo della somma delle distanze dai due giocatori in possesso altro non è che una scomposizione della

distanza originaria. Entrambe le distanze calcolate ricostruirebbero la distanza originaria d , pur essendo il giocatore *Rosso* molto più vicino alla linea di passaggio rispetto al giocatore *Aranzone*. Per ovviare a questo problema la condizione è stata modificata impostando un fattore correttivo.

La nuova condizione anziché l'uguaglianza delle distanze imposta una banda di vicinanza al valore della distanza tra P1 e P2 maggiorato o diminuito di una percentuale della distanza d . La regola di determinazione tra linea libera o chiusa è stata così riscritta:

$$\text{distanza}[P1,AVV 1] + \text{distanza}[AVV 1, P2] \leq \text{distanza}[P1, P2] + (0.025 * \text{distanza}[P1, P2])$$

$$\& \text{distanza}[P1, AVV 1] + \text{distanza}[AVV 1, P2] \geq \text{distanza}[P1, P2] - (0.025 * \text{distanza}[P1, P2])$$

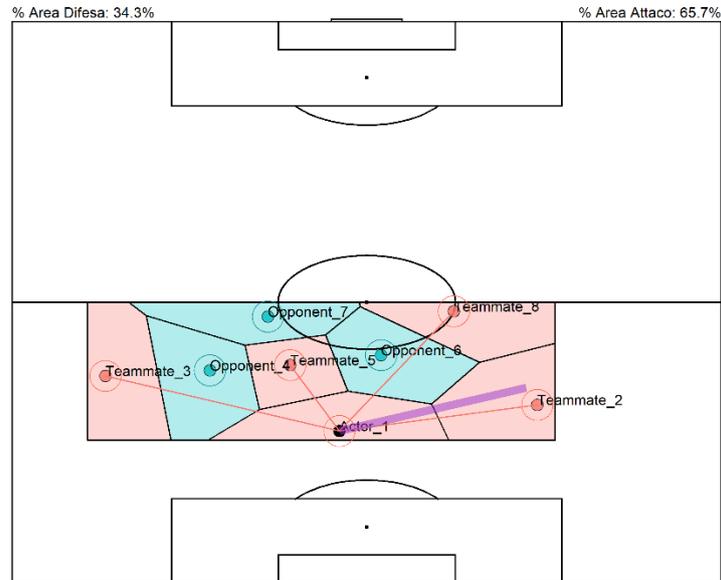
Se la somma delle distanze tra l'Avversario e i due giocatori in possesso P1 e P2 è compresa nell'intervallo tra la distanza d tra i due giocatori P1 e P2 +/- il 2.5% della stessa distanza d , allora la linea di passaggio sarà considerata chiusa, altrimenti libera.

Viene così stabilita una banda in cui se l'avversario è posizionato all'interno sarà considerato in grado di intercettare il passaggio. Il fatto del 2.5% è stato stabilito in maniera arbitraria dopo numerosi tentativi nel confronto tra l'immagine posizionale e i risultati forniti dalla condizione stabilita.

Di seguito è riportata la grafica finale prodotta dall'elaborazione. L'esempio fa riferimento ad un passaggio effettuato da Tomas Kalas al minuto 13:35 della partita Repubblica Ceca – Croazia. Oltre alla tassellazione di Voronoi con le espressioni percentuali dell'area controllata da ogni squadra, sono disegnate, in tono di rosso, le linee di passaggio a disposizione di Kalas, denominato Actor1. Ad ogni giocatore che sia della squadra in possesso o di quella difendete viene assegnato un numero, in questo caso da 1 a 8. In viola il passaggio effettuato da Kalas. Nel sottotitolo sono riportati i conteggi dei giocatori coinvolti, compagni e avversari, il numero dei difendenti superati e le linee di passaggio considerate libere e chiuse. Nel caso in esame 3 linee di passaggio valutate libere e una considerata bloccata stante la condizione precedentemente illustrata.

Tomáš Kalas - Pass - 13:35

Czech Republic-Croatia, UEFA EURO 2020
 BYPASSED OPPONENTS: 0 | OPPONENTS: 3 | TEAMMATES: 4
 LINEE PASS FREE: 3 | LINEE PASS BLOCKED: 1
 PASS OUTCOME: Completed



as.character(Colteam) — Attack team Team — Attack team — Defense team

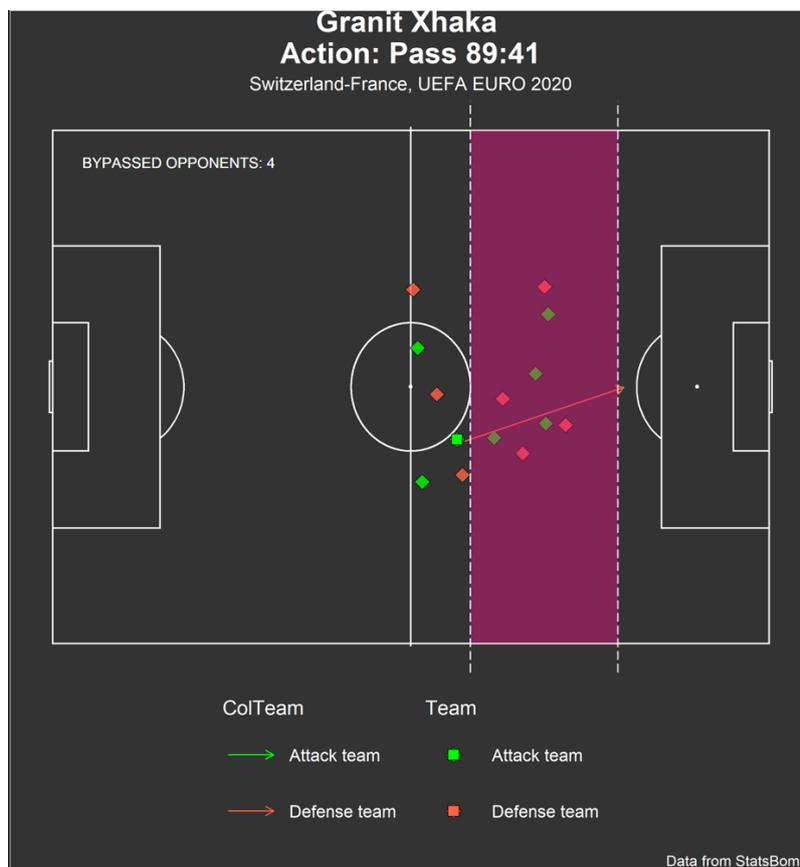
Data: StatsBomb
 By: @mxzone, Match & Data

Nell'elaborazione statistica condotta dal codice R scritto, la condizione di determinazione sulle linee di passaggio produce per ogni passaggio considerato una matrice specifica. In colonna sono identificati gli avversari, per cui vengono valorizzate le colonne con il numero dei soli avversari presenti (4, 6 e 7). In riga i giocatori della squadra in possesso, con valorizzate solamente le righe con riferimento ai compagni presenti (1,2,3, 5 e 8). Le celle sono valorizzate con 0 se l'incrocio attaccante in riga e difendente in colonna determina una linea di passaggio libera e con 1 qualora l'incrocio determini una condizione di linea di passaggio bloccata. Nell'esempio specifico l'avversario 6 (Opponent_6) si posiziona in maniera tale da rendere chiuso a Kalas il passaggio verso il compagno 8 (Teammate_8).

	V1	V2	V3	V4	V5	V6	V7	V8
1	NA	NA	NA	0	NA	0	0	NA
2	NA	NA	NA	0	NA	0	0	NA
3	NA	NA	NA	0	NA	0	0	NA
4	NA							
5	NA	NA	NA	0	NA	0	0	NA
6	NA							
7	NA							
8	NA	NA	NA	0	NA	1	0	NA

I DIFENDENTI BYPASSATI DAL PASSAGGIO EFFETTUATO

Il terzo elemento calcolato nel lavoro è quello dei difendenti superati dal passaggio effettuato in direzione della porta avversario. I dati StatsBomb per il passaggio forniscono le coordinate x-y finali di arrivo, i dati StatsBomb360 permettono di localizzare nell'area visibile fornita i giocatori della squadra difendente. Un avversario viene quindi considerato come bypassato qualora venga superato dal passaggio effettuato.



Il passaggio di Xhaka, quadrato verde, al minuto 89:41 di Svizzera-Francia, supera 4 giocatori francesi considerati nell'area visibile fornita da SatsBomb360.

VISUALIZZAZIONE

Il processo di elaborazione sopra descritto è stato calcolato per ogni passaggio svolto nel corso di Euro 2020. Per ognuno dei 54820 passaggi analizzati è stata prodotta una visualizzazione come quella che segue. In ogni visualizzazione sono riportate delle informazioni:

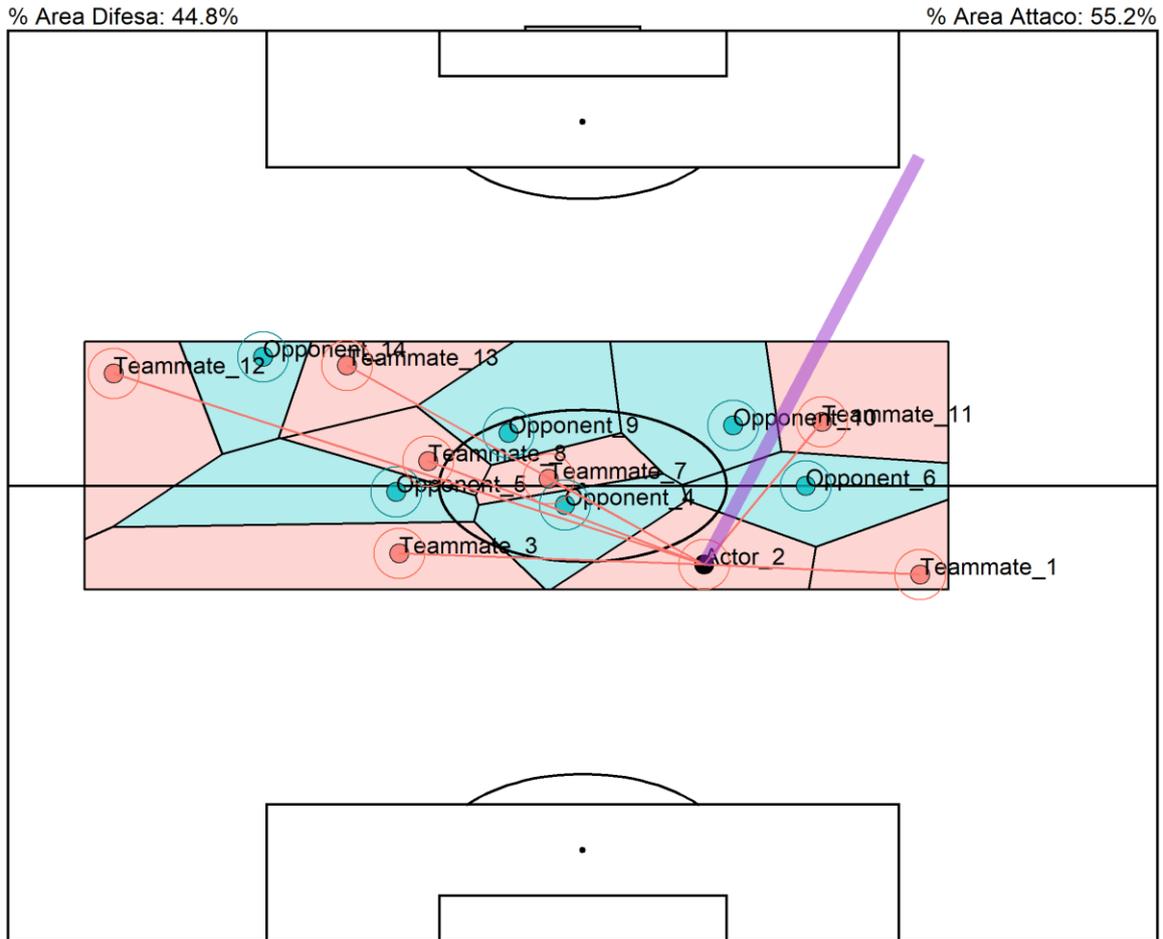
- TITOLO: Autore del passaggio e riferimento temporale dell'evento
- SOTTOTITOLO:
 - Partita cui si riferisce l'evento
 - Numero di Difendenti Bypassati | N totale Difendenti | N totale Compagni di squadra
 - N di linee di passaggi libere | N linee di passaggio chiuse
 - Outcome del passaggio

In alto a sinistra della grafica è riportata la percentuale dell'area visibile fornita da StatsBomb per l'evento, colorata, controllata dalla squadra in difesa, al lato opposto la stessa misura per la squadra in attacco.

Nella visualizzazione grafica la tassellazione di Voronoi è sviluppata solamente all'interno dell'area resa visibile da StatsBomb. È quindi colorata l'area all'interno del rettangolo di campo visibile. In viola la traiettoria del passaggio effettuato.

Leonardo Bonucci - Pass - 0:38

Italy-Wales, UEFA EURO 2020
 BYPASSED OPPONENTS: 6 | OPPONENTS: 6 | TEAMMATES: 7
 LINEE PASS FREE: 3 | LINEE PASS BLOCKED: 4
 PASS OUTCOME: Pass Offside



as.character(Colteam) — Attack team Team ● Attack team ● Defense team

Data: StatsBomb
 By: @mxd zone, Match & Data

Oltre alla grafica per ogni passaggio su cui viene sviluppata l'analisi vengono raccolte in un Excel le principali informazioni. Di seguito attraverso un esempio l'elenco completo delle informazioni raccolte.

player.name	Jakub Moder
type.name	Pass
second	34
minute	30
team.name	Poland
OpposingTeam	Spain
match_id	3788762
ff_location.x	73,207
ff_location.y	5,38157
ColTeam	Attack team
Player_Type_Key	Actor
pass.end_location.x	71,9
pass.end_location.y	20,3
count	4
ID	Actor_4
Distance_to_Actor	0
dir.area	137,920526
opp_in_trajectory	0
area_difesa	40,8
area_attacco	59,2
avversari	9
compagni	8
bypassed_opp	0
mate_free	3
mate_closed	5
under_pressure	
play_pattern.name	From Throw In
position.name	Left Center Midfield
pass.length	13,191285
pass.angle	1,9189197
pass.recipient.name	Mateusz Andrzej Klich
pass.height.name	Ground Pass
pass.body_part.name	Right Foot
pass.type.name	
pass.outcome.name	Completed
pass.technique.name	
pass.inswinging	
pass.deflected	
id	2eb3076c-e70a-4141-8b35- da3fa1067dda
match_index	37887621339
start_Zone	Zona 2
end_Zone	Zona 2

Oltre alle misure già presentate, nel foglio Excel vengono riportate anche le informazioni circa l'area che il giocatore in possesso di palla ha sotto il proprio

controllo (**dir.area**), il nome del giocatore destinatario del passaggio (**pass.recipient.name**), se il passaggio è avvenuto sotto pressione avversaria o meno (**under_pressure**), il piede con cui ha effettuato il passaggio (**pass.body_part.name**), oltre alla zona di campo da cui parte il passaggio (**start_Zone**) e la zona in cui è destinato il passaggio (**end_Zone**).

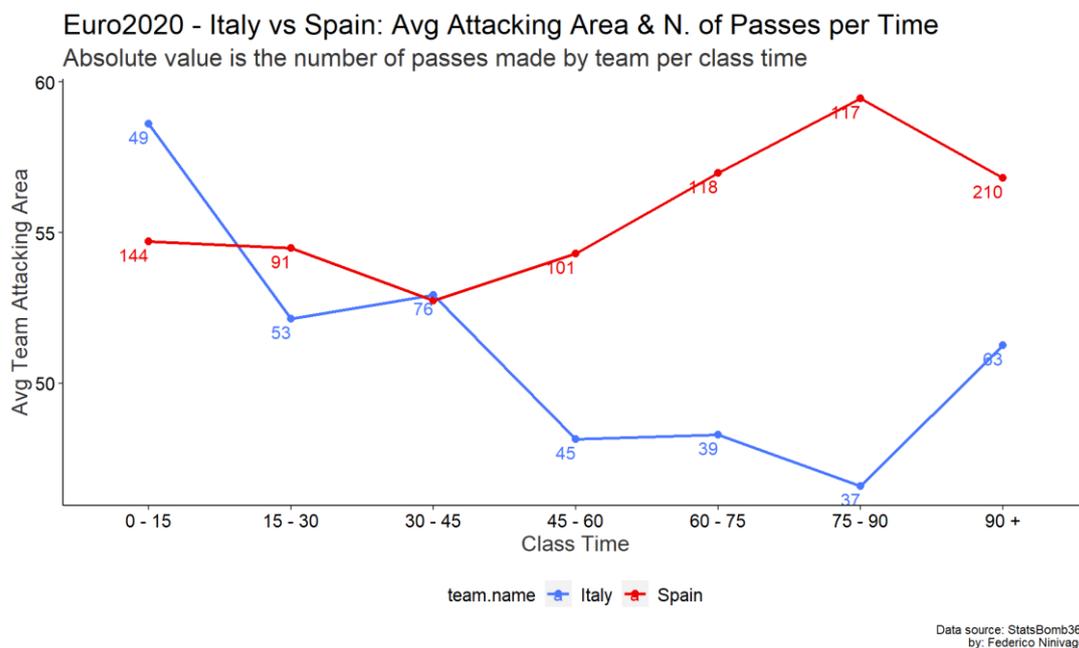
I file Excel prodotti per ogni passaggio, vengono raggruppati in un unico dataset dove per ogni riga c'è un passaggio e in colonna le informazioni raccolte. Da questo file aggregato sono sviluppate le analisi descrittive che saranno presentate nella prossima sezione.

ANALISI DESCRITTIVE

Il dataset raggruppa le informazioni descritte in precedenza avendo in riga il singolo passaggio. Aggregando i dati rispetto ad alcune dimensioni è stato possibile visualizzare alcune evidenze sia rispetto ai singoli giocatori, che rispetto alle squadre, per partita o per l'intero torneo. Di seguito vengono proposte alcune delle possibili visualizzazioni focalizzandosi soprattutto su quelle che sono state le performance della Nazionale Italiana.

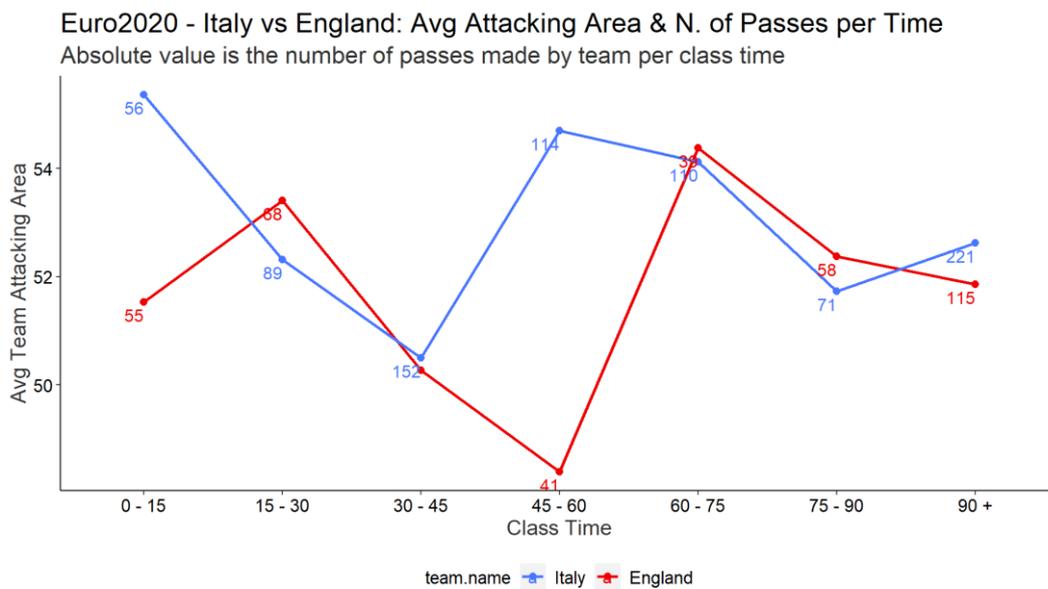
La prima analisi si propone di visualizzare l'andamento per intervallo di tempo dell'area media occupata in attacco dalle squadre durante la partita.

Sull'asse delle y è riportato la scala di misura del valore medio della percentuale di area occupata da ogni squadra quando in possesso della palla. Sull'asse x gli intervalli di tempo, di 15 minuti, in cui è stato suddiviso l'andamento della partita. Viene così disegnato l'andamento dell'occupazione offensiva media degli spazi in attacco da parte delle squadre. Ulteriore elemento sono il numero totale di passaggi tentati dalle due squadre nel rispettivo periodo di tempo. Le grafiche proposte si riferiscono alle partite della Nazionale contro Spagna e Inghilterra.



Contro la Spagna la Nazionale ha visto con il passare del tempo diminuire l'area controllata quando in possesso, così come il numero di passaggi tentati. Al contrario la nazionale spagnola nel corso del secondo tempo ha incrementato sia il numero di passaggi, sia l'area media sotto controllo in fase offensiva, portandola abbondantemente oltre il 50%.

In occasione della finale di Euro2020 l'andamento del controllo territoriale in fase offensiva per le due squadre ha avuto un andamento molto più altalenante rispetto a quello avuto nella semifinale tra Italia e Spagna.

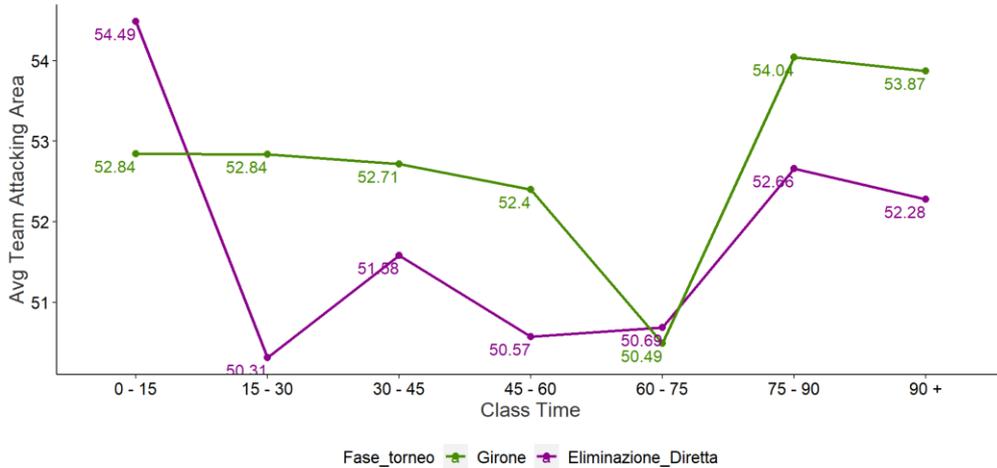


Data source: StatsBomb360
 by: Federico Ninivaggi

A inizio secondo tempo tra il 45' e il 60' si è registrato il picco per l'Italia sia in termini di passaggi effettuati sia di area media controllata in possesso, oltre il 54%. Al contrario nello stesso periodo di tempo l'Inghilterra ha tentato 41 passaggi riuscendo a controllare in quelle occasioni meno del 50% dell'area su cui è stata sviluppata l'analisi.

L'occupazione media dell'area in fase di possesso della Nazionale Italiana nel corso del torneo è riassunta nel grafico seguente. Nella figura viene visualizzato l'andamento temporale nel corso della partita del valore medio dell'area d'attacco fatta registrare dalla squadra di Mancini nelle partite del girone eliminatorio e nelle partite ad eliminazione diretta.

Euro2020 - Italy Avg Attacking Area: Group Stage vs Knockout stage per Time
 Group Stage: Turkey, Switzerland, Wales
 Knockout stage: Austria, Belgium, Spain, England

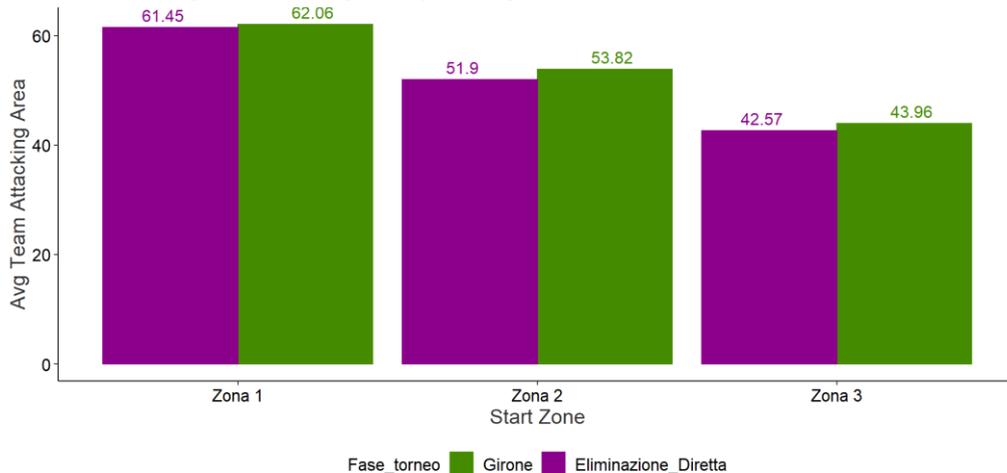


Data source: StatsBomb360
 by: Federico Ninivaggi

L'Italia è sempre riuscita a fare registrare valori medi di occupazione maggiori del 50%. Per le partite del girone eliminatorio con eccezione del minimo avuto nel periodo tra il 45-60 minuto l'Italia ha avuto valori migliori rispetto a quanto fatto per gli stessi periodi nella fase ad eliminazione diretta. Questo è probabilmente imputabile alla maggior livello tecnico e tattico degli avversari incontrati nella seconda fase del torneo.

La stessa misura è di seguito rappresentata rispetto alla zona di campo in cui è avvenuto il possesso.

Euro2020 - Italy Avg Attacking Area: Group Stage vs Knockout stage per Zone
 Group Stage: Turkey, Switzerland, Wales
 Knockout stage: Austria, Belgium, Spain, England

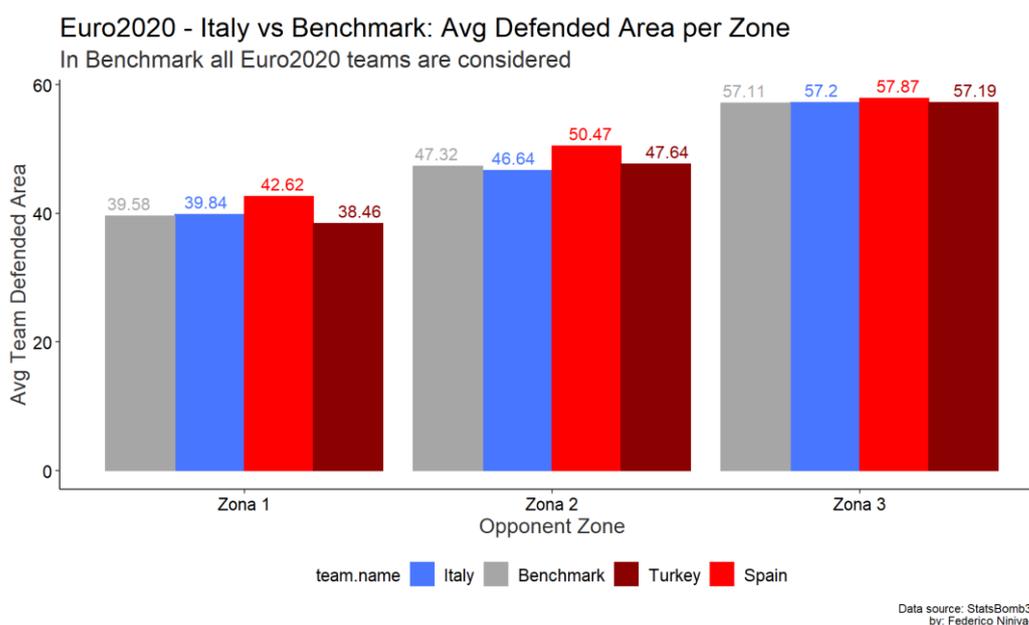


Data source: StatsBomb360
 by: Federico Ninivaggi

Come intuibile man mano che il possesso della Nazionale si sposta verso la porta avversaria la percentuale media dell'area controllata tende a diminuire. Anche in

questo caso i valori registrati per la fase a gironi sono superiori di quelli della fase ad eliminazione diretta.

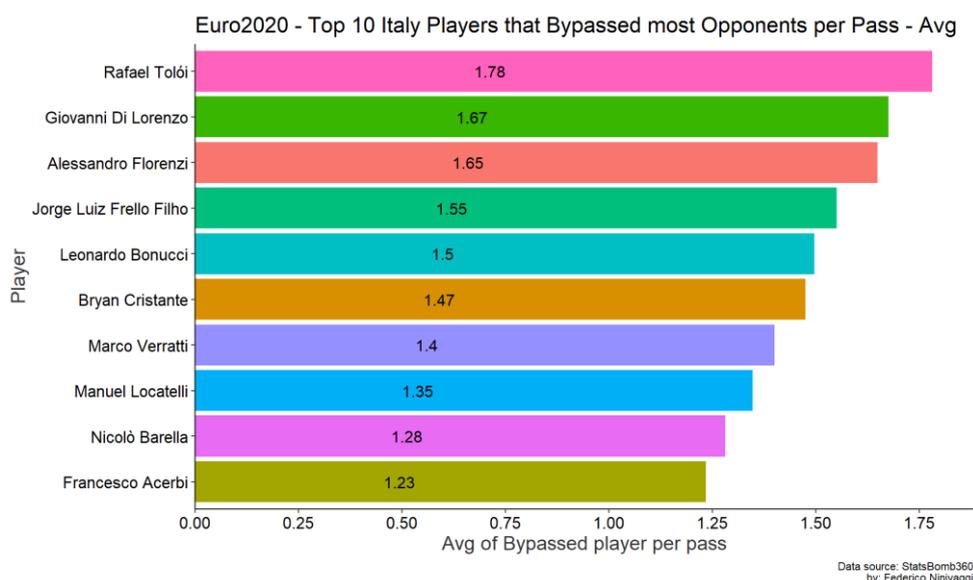
Analisi dello stesso genere possono essere condotte rispetto alla percentuale di area controllata dalla squadra durante la fase difensiva. Nel riferirsi alla distribuzione geografica del campo come nella visualizzazione successiva, va modificata la modalità di lettura delle zone espresse sull'asse x. Il riferimento è rispetto all'azione condotta dalla squadra avversaria che è in possesso del pallone. La Zona 1 sarà quindi la zona in cui generalmente avviene la costruzione della squadra avversaria quando in possesso. La Zona 3 si riferisce alla zona di rifinitura e finalizzazione per la squadra in possesso.



Nel grafico ad ogni colore corrisponde una squadra, in grigio il valore di Benchmark calcolato aggregando i valori fatti registrare da tutte le squadre che hanno preso parte ad Euro2020. La Spagna fa registrare in tutte e tre le zone del campo il valore medio più alto tra le squadre presentate nel grafico. L'Italia si attesta intorno ai valori del benchmark in tutte le zone del campo. I dati sempre superiori alle altre squadre della Spagna indicano un'ottima difesa degli spazi soprattutto in zona palla, essendo l'area resa visibile dai dati StatsBomb360, anche rispetto alla Zona 1 avversaria, quest'ultimo valore riconducibile allo stile di riaggresione e pressing tipico della squadra di Luis Enrique.

Le elaborazioni svolte e le metriche calcolate per ciascun passaggio possono essere applicate anche alle performance dei singoli giocatori. Di seguito vengono proposte alcune delle possibili visualizzazioni.

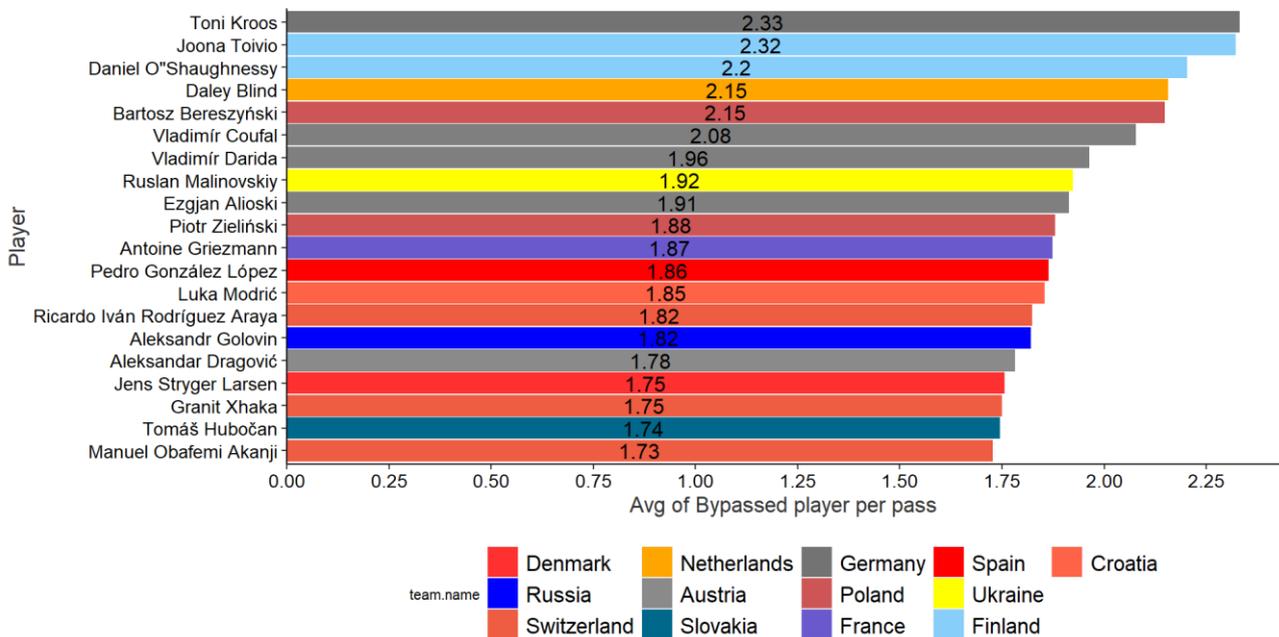
I primi dieci giocatori per numero medio di avversari superati con un passaggio sono di seguito riportati. Al primo posto Tolòì seguito da Di Lorenzo, Florenzi e poi Jorginho. In top 10 ci sono i centrocampisti centrali e le mezzali, Verratti, Barella Cristante e Locatelli, mancano attaccanti e esterni offensivi.



Estendendo il campo d'indagine a tutti i giocatori che hanno preso parte alla competizione e hanno completato almeno 75 passaggi nel torneo nelle prime 20 posizioni non troviamo nessun giocatore della Nazionale. Tolòì, infatti, nel corso del torneo non ha completato il numero minimo di passaggi (75) posto come filtro a monte. A guidare il ranking è il centrocampista tedesco Toni Kroos che in media con ogni passaggio ha superato 2.33 avversari.

Euro2020 - Top 20 Players that Bypassed most Opponents per Pass (Avg)

Players with at least 75 passes Completed

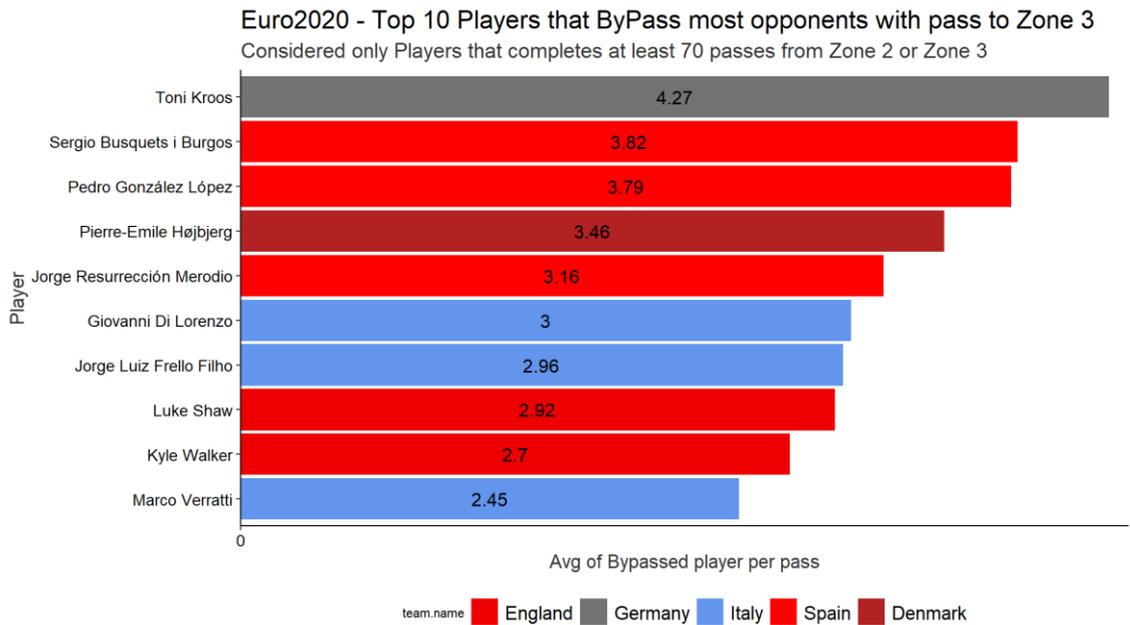


Data source: StatsBomb360
by: Federico Ninivaggi

A Kroos seguono due difensori della nazionale finlandese. In classifica sono molti i difensori presenti oltre ai centrocampisti. Tra i giocatori offensivi ed attaccanti compaiono il regista offensivo ucraino Malinovskiy, l'attaccante francese Griezmann e l'esterno offensivo russo Golovin. Oltre all'Italia mancano in classifica giocatori per il Belgio e l'Inghilterra. Per la Spagna il solo Pedri risulta nella top 20.

La domanda d'indagine potrebbe essere specificata facendo riferimento ad una specifica zona di campo. Quali sono i giocatori che completano il maggior numero di passaggi diretti in zona 3 partendo da zona 2 o 3? Chi per la stessa tipologia di passaggi supera più avversari in media?

I giocatori che fanno registrare le migliori performance rispetto alle due domande sono riportati nelle grafiche che seguono.

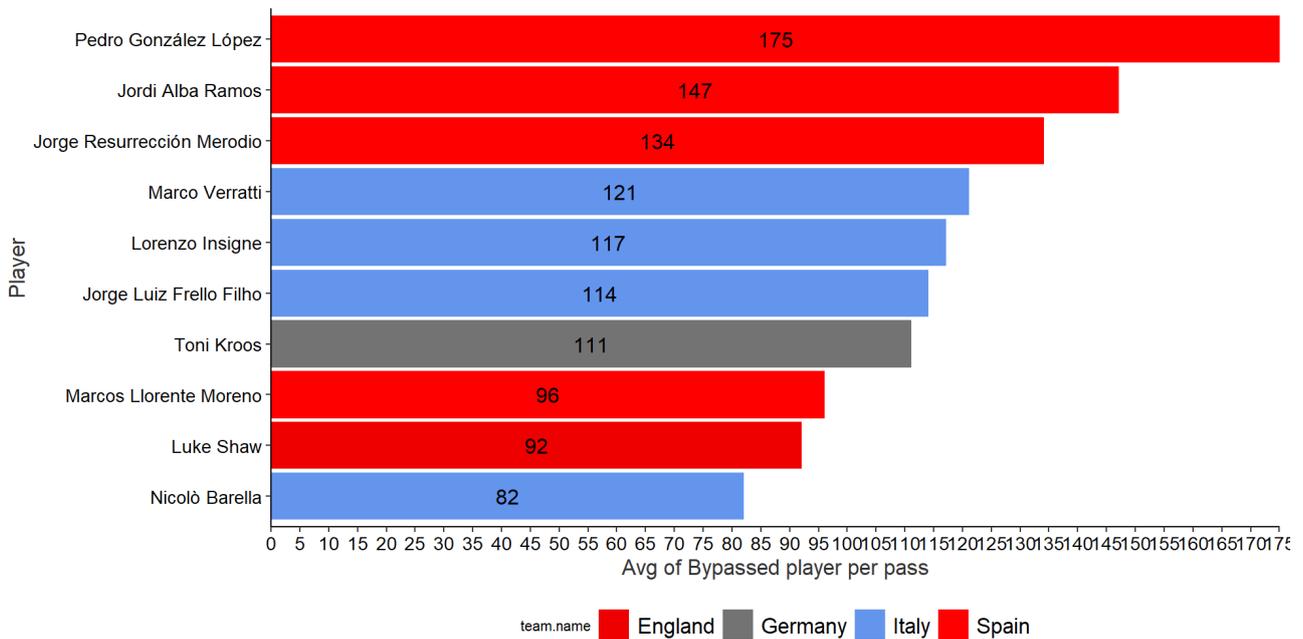


A dominare le due top10 sono i giocatori di Inghilterra Italia e Spagna. A guidare la classifica è ancora una volta Toni Kroos. I passaggi diretti in zona 3 del tedesco in media superano 4.27 avversari. In classifica i 3 centrocampisti della Spagna Busquets, Pedri e Koke. Per l'Italia il primo è Di Lorenzo, in classifica anche i due play azzurri Jorginho e Verratti. Per l'altra finalista del torneo ad entrare in top 10 sono i due terzini Walker e Shaw.

Guardando invece al numero complessivo di passaggi completati destinati in Zona 3 da Zona 2 o 3 nelle prime tre posizioni ci sono tre giocatori della Spagna: Pedri con 175 passaggi completati verso Zona 3, seguito da Jordi Alba e Koke. Al tritico spagnolo fa seguito un trio di giocatori italiani, Verratti, Insigne e Jorginho. Le due nazionali portano in classifica un altro giocatore per parte. Unici in Top 10 a non appartenere alle due squadre sono ancora una volta il tedesco Kroos e l'esterno sinistro inglese Shaw.

Euro2020 - Top 10 Players that complete most passes to Zone 3

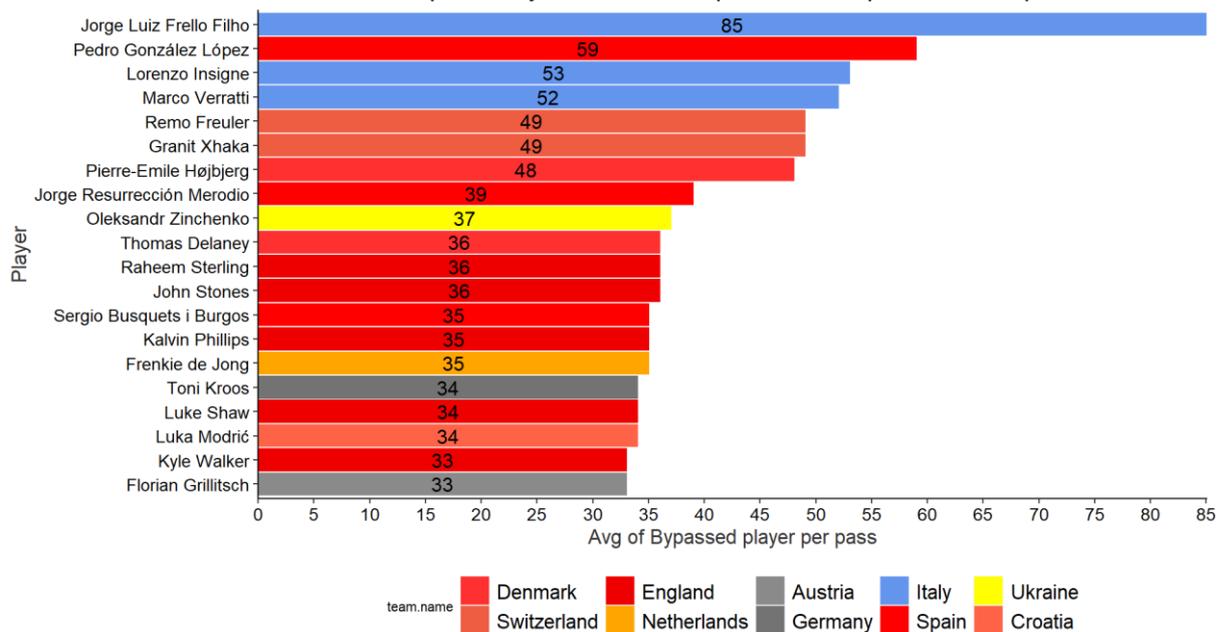
Considered only completed passes from Zone 2 or Zone 3



Data source: StatsBomb360
by: Federico Ninivaggi

Un ulteriore approfondimento può essere condotto integrando le misure ricavate dai dati posizionali con gli attributi forniti dai dati StatsBomb relativi allo specifico evento. Tra questi uno dei più interessanti è la qualifica secondo il quale il passaggio sarebbe avvenuto o meno sotto la pressione di un avversario.

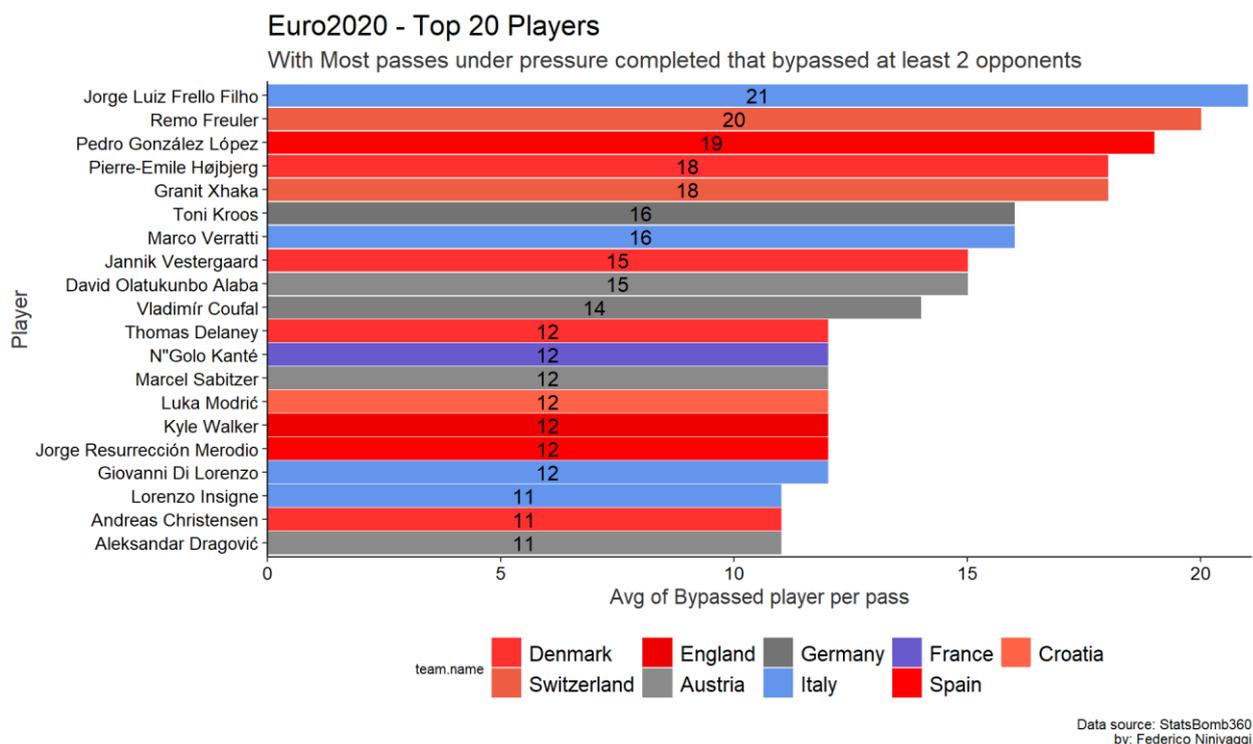
Euro2020 - Top 20 Players With Most passes Completed under pressure



Data source: StatsBomb360
by: Federico Ninivaggi

Ad aver completato più passaggi sotto pressione è stato il regista azzurro Jorginho che con 85 passaggi stacca di molto il secondo nel ranking, lo spagnolo Pedri. Ottime le performance fatte registrare anche da Insigne e Verratti. Nelle prime 4 posizioni bene 3 giocatori della Nazionale.

Se si guarda ai giocatori che nel torneo sono riusciti a completare più passaggi sotto pressione avversaria superando almeno due avversari la classifica si modifica lasciando però in testa Jorginho. Scala verso il basso di una posizione Pedri lasciando il secondo posto allo svizzero Freuler.



CONCLUSIONI

Il lavoro fin qui presentato rappresenta il primo personale tentativo di analisi dei dati posizionali. Si tratta con certezza di un'esplorazione dello strato superficiale della potenza informativa della tipologia di dato. Le estensioni che il lavoro potrebbe avere sarebbero molteplici. Dal lato dell'elaborazione di metriche da sviluppare, oltre a quelle presentate si potrebbero aggiungere informazioni circa le distanze tra giocatori e reparti, il calcolo degli involucri convessi delle squadre nelle due fasi, così come la lunghezza e ampiezza assunta dalle formazioni, tutte informazioni utili nell'identificare stili di gioco e tipicità tattiche delle squadre. Altro ambito di estensione del lavoro sarebbe quello degli eventi coinvolti, dalle conduzioni alle azioni difensive. Il lavoro come si è visto si è limitato ad analizzare il passaggio essendo l'evento più numeroso tra quelli StatsBomb. Integrare tutti gli eventi ed estendere la copertura all'intero campo permetterebbe poi uno sviluppo totale sia geografico che temporale della partita. Aggregare dati e metriche ricavate dai dati posizionali, come quelle presentate, per interi campionati permetterebbe di avere una visione più specifica e dettagliata degli stili di gioco di giocatori e squadre evidenziandone pregi e difetti tattici e tecnici nell'arco di un'intera stagione.