Defenders Beat by Passes Report

Project Overview

I used StatsBomb's open-source data and API to develop a statistic that quantifies how many defenders are bypassed by a player's successful forward pass, focusing on the 2022 World Cup due to its comprehensive event and event360 data. This new metric, combined with other passing statistics, offers insights into the most effective and efficient passers in the tournament. I presented tables highlighting players who bypassed the most defenders with their passes and those who did so most efficiently, based on varying criteria for completed passes.

Initial Considerations

Before defining my project goal, I explored StatsBomb's open-source data, learning that each competition's match data files contain details on the matches, along with event data files for on-ball events and event360 data files for player locations, including off-ball positions during specific events. I aimed to focus on a recent competition with substantial event data coverage.

Upon reviewing the data by year, I found limited data post-2022. The 2023 MLS had only six matches recorded, and Ligue 1 2022/23 had 32 matches, all PSG games. However, the 2022 World Cup offered complete data on all 64 matches, including rare event360 data for each game, prompting me to delve deeper into this dataset.

While brainstorming, I was drawn to the event360 data. I find that data along with StatsBomb's "under pressure" metric the most unique and interesting aspects of StatsBomb's open-source data. Given the frequency of passes within the event data, I decided to focus on them, leading to the creation of a *Defenders Beat by Pass* statistic for the 2022 World Cup, leveraging the off-ball location data.

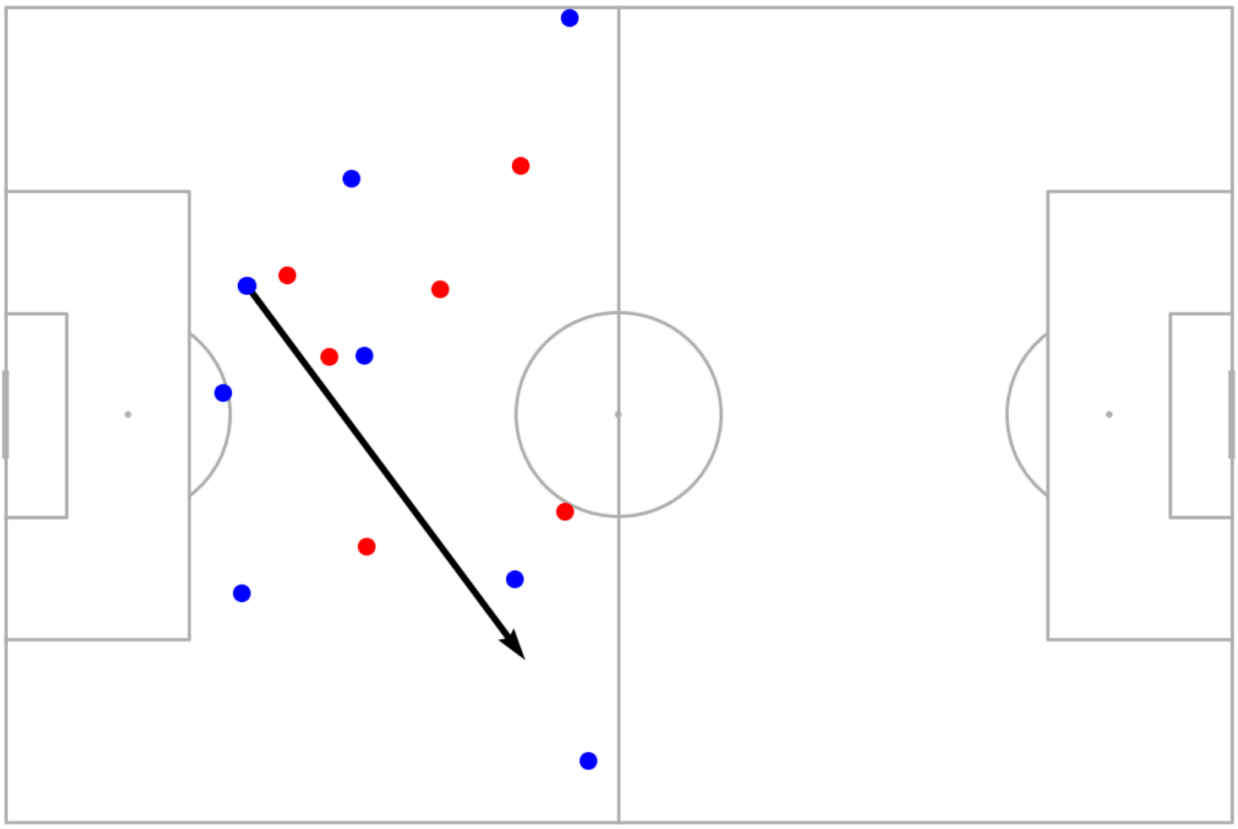
Data Exploration and Cleaning

Upon clarifying my project goal, I examined the event and event360 data for a match to understand the relevant columns and their meanings, utilizing StatsBomb's documentation for guidance. Relevant event types for my analysis included "Pass" and "Ball Receipt*", the latter indicating the pass's recipient or intended recipient. Key columns included "location" (coordinates of the passer), "pass_end_location" (coordinates of the pass's destination), "pass_outcome" (completion status), "pass_angle" (direction), and "freeze_frame" (off-ball player positions during an event).

To consolidate the data, I merged event and event360 data for all 2022 World Cup matches by their IDs and compiled them into a single dataset. I then created separate data frames for passing and ball receipt events, focusing on forward completed passes that could bypass defenders. By merging these data frames, I aligned pass details with corresponding ball receipts. Lastly, I filtered for events with available off-ball player data. The merge and cleaning resulted in 79.25% of all forward completed passes, which is a significant portion of the data.

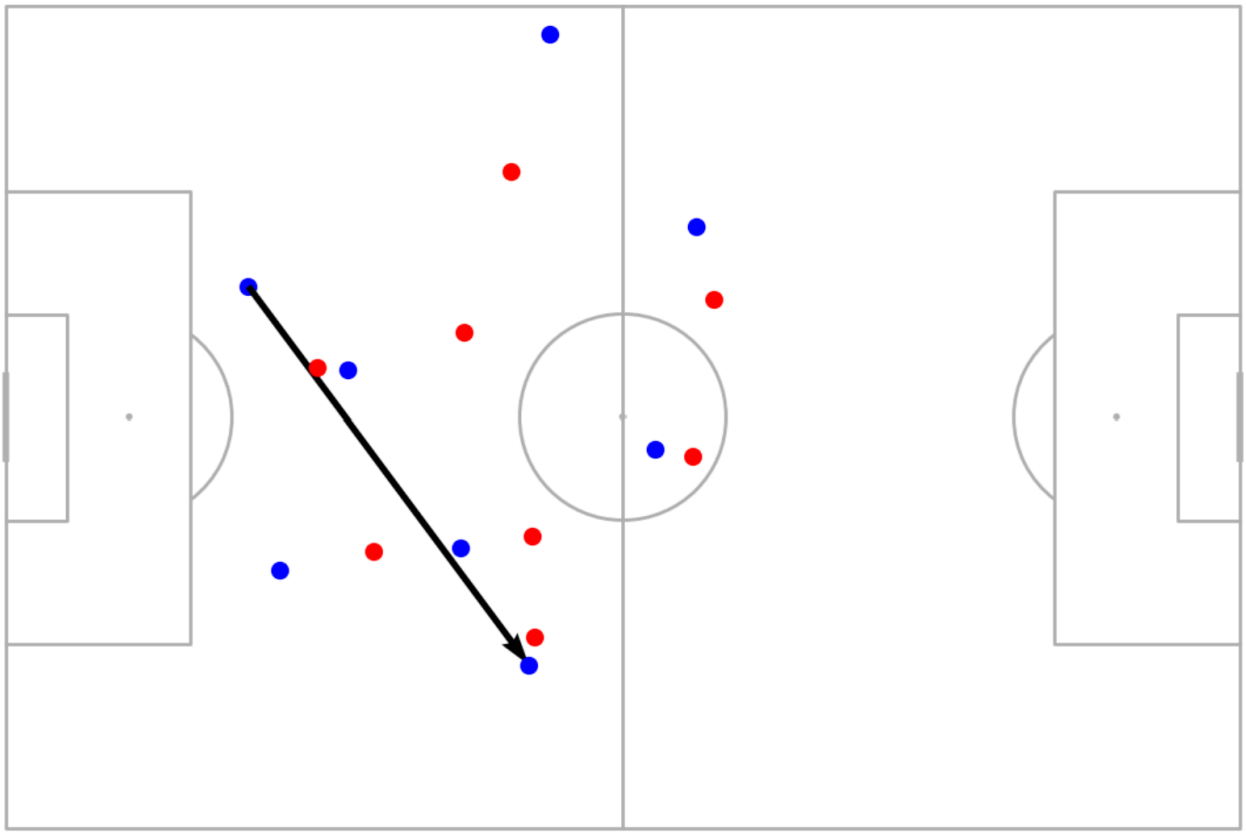
Project Implementation

To develop the *Defenders Beat by Pass* statistic, I started by calculating it for a specific instance and created visualizations to confirm my calculation's accuracy. I selected a single pass, and using the mplsoccer API, which can interpret StatsBomb’s coordinate system, I plotted the pass vector and player positions at the moment the pass was executed.



Plot of Pass and Player Locations at the Time the Ball is Played

The pass vector, indicated by the start and end locations of the pass, is displayed in black. The teammates of the player in possession are shown in blue, and the opponents are depicted in red. I then plotted the pass again, this time using player location data from when the ball was received.



Plot of Pass and Player Locations at the Time the Ball is Received

I chose to use player location data when the ball is received to account for defenders' movements during the pass. This approach helps avoid overestimating the number of defenders bypassed, which can happen when using player locations at the time the pass is made. To calculate, I compared the number of defenders behind the ball using the x-coordinates before and after the pass is received.

However, this calculation doesn't perfectly capture all defenders bypassed by a pass. Future improvements could include considering horizontal passes that lead to the recipient dribbling into open space, effectively bypassing defenders. Additionally, I might exclude players far from the play, as their distance gives them more time to reposition, reducing their relevance to the immediate play.

Project Results

I produced multiple tables showcasing the players in the 2022 World Cup, who were most influential in regards to interpretations of the *Defenders Beat by Pass* statistic.

Top 15 Players with Highest Total Defenders Beat by Passes

	Player	Team	Total_Defenders_Beat	Foward_Completed_Passes	Avg_Defenders_Beat_per_Pass
0	Luka Modrić	Croatia	706	267	2.64
1	Rodrigo Hernández Cascante	Spain	583	380	1.53
2	Pedro González López	Spain	549	198	2.77
3	Lionel Andrés Messi Cuccittini	Argentina	526	152	3.46
4	Marcelo Brozović	Croatia	471	244	1.93
5	Joško Gvardiol	Croatia	463	311	1.49
6	Nicolás Hernán Otamendi	Argentina	454	326	1.39
7	Rodrigo Javier De Paul	Argentina	443	173	2.56
8	Enzo Fernandez	Argentina	435	227	1.92
9	John Stones	England	432	281	1.54
10	Achraf Hakimi Mouh	Morocco	431	168	2.57
11	Aurélien Djani Tchouaméni	France	424	218	1.94
12	Mateo Kovačić	Croatia	418	175	2.39
13	Aymeric Laporte	Spain	391	287	1.36
14	Luke Shaw	England	390	147	2.65

The players in this table were the most influential in terms of volume in bypassing defenders with passes. Those with a low "Avg_Defenders_Beat_per_Pass" might not have excelled at finding spaces and bypassing defenders during the 2022 World Cup but rather had numerous opportunities on the ball in possession-based teams.

I next wanted to sort by “Avg_Defenders_Beat_per_Pass” (with a minimum of 100 forward completed passes) to provide insight on which players were the most efficient with their passes.

Top 15 Players with Highest Avg. Defenders Beat per Completed Foward Pass

*minimum 100 completed foward passes

	Player	Team	Total_Defenders_Beat	Foward_Completed_Passes	Avg_Defenders_Beat_per_Pass
0	Lionel Andrés Messi Cuccittini	Argentina	526	152	3.46
1	Antoine Griezmann	France	332	102	3.25
2	Theo Bernard François Hernández	France	328	108	3.04
3	Daley Blind	Netherlands	326	114	2.86
4	Borna Sosa	Croatia	302	107	2.82
5	Jordi Alba Ramos	Spain	353	127	2.78
6	Pedro González López	Spain	549	198	2.77
7	Jules Koundé	France	329	121	2.72
8	Carlos Henrique Casimiro	Brazil	288	108	2.67
9	Luke Shaw	England	390	147	2.65
10	Luka Modrić	Croatia	706	267	2.64
11	Josip Juranović	Croatia	386	146	2.64
12	Éder Gabriel Militão	Brazil	346	132	2.62
13	Achraf Hakimi Mouh	Morocco	431	168	2.57
14	Rodrigo Javier De Paul	Argentina	443	173	2.56

Similar to the previous table, this above table consists of all technically elite players on top international teams. Because the previous two tables consist entirely of players on top-performing teams who made it to the knockout-phase of the tournament, I produced another table with a lower minimum forward completed passes threshold, to provide insight on potentially underappreciated players at the 2022 World Cup.

Top 15 Players with Highest Avg. Defenders Beat per Completed Foward Pass

*minimum 50 completed foward passes

	Player	Team	Total_Defenders_Beat	Foward_Completed_Passes	Avg_Defenders_Beat_per_Pass
0	Lionel Andrés Messi Cuccittini	Argentina	526	152	3.46
1	Ngoran Suiru Fai Collins	Cameroon	194	57	3.40
2	Weston McKennie	United States	212	65	3.26
3	Antoine Griezmann	France	332	102	3.25
4	Rasmus Nissen Kristensen	Denmark	196	61	3.21
5	Yahia Attiyat allah	Morocco	163	51	3.20
6	Abdulelah Al Amri	Saudi Arabia	183	58	3.16
7	Steven Berghuis	Netherlands	163	52	3.13
8	Kevin De Bruyne	Belgium	162	52	3.12
9	Neymar da Silva Santos Junior	Brazil	213	69	3.09
10	Silvan Widmer	Switzerland	194	63	3.08
11	Antonee Robinson	United States	250	82	3.05
12	Theo Bernard François Hernández	France	328	108	3.04
13	Mohammed Kanoo	Saudi Arabia	176	58	3.03
14	Bruno Miguel Borges Fernandes	Portugal	245	81	3.02

With the new threshold, there are clearly a lot of new players in the above table. Although a lot of these players were not as involved in the tournament as the players in the previous tables, it is important to recognize their effective passing abilities displayed in the tournament.

Conclusion

The *Defenders Beat by Pass* statistic, which quantifies the effectiveness of players in bypassing defenders with forward passes, enriches the analysis of passing efficiency and directness in play. The research highlighted not only the top performers in terms of volume but also those who demonstrated exceptional efficiency in their passing game. The inclusion of

players from teams that did not advance beyond the group stages underscores the statistic's ability to uncover underappreciated talents.