English Premier League

Comprehensive Exploratory Data Analysis Report
2024-2025 Season

Generated: 2025-11-07 03:01

1.1 Data Preview - First 10 Rows

1.2 Dataset Information (df.info())

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 380 entries, 0 to 379
Columns: 122 entries, Div to MonthName
dtypes: datetime64[ns](1), float64(96), int32(1), int64(16), object(8)
memory usage: 360.8+ KB
```

index	1.3 Missing Data Summary (Top 30)		Missing Percentage
27	BWH	141	37.11
75	BWCH	141	37.11
29	BWA	141	37.11
28	BWD	141	37.11
77	BWCA	141	37.11
76	BWCD	141	37.11
86	WHCA	91	23.95
85	WHCD	91	23.95
36	WHH	91	23.95
37	WHD	91	23.95
38	WHA	91	23.95
84	WHCH	91	23.95
39	1XBH	9	2.37
40	1XBD	9	2.37
41	1XBA	9	2.37
101	PC>2.5	3	0.79
53	P>2.5	3	0.79
60	BFE<2.5	3	0.79
59	BFE>2.5	3	0.79
54	P<2.5	3	0.79
102	PC<2.5	3	0.79
32	BFA	1	0.26
31	BFD	1	0.26
30	BFH	1	0.26

1.4 Missing Data Pattern Heatmap
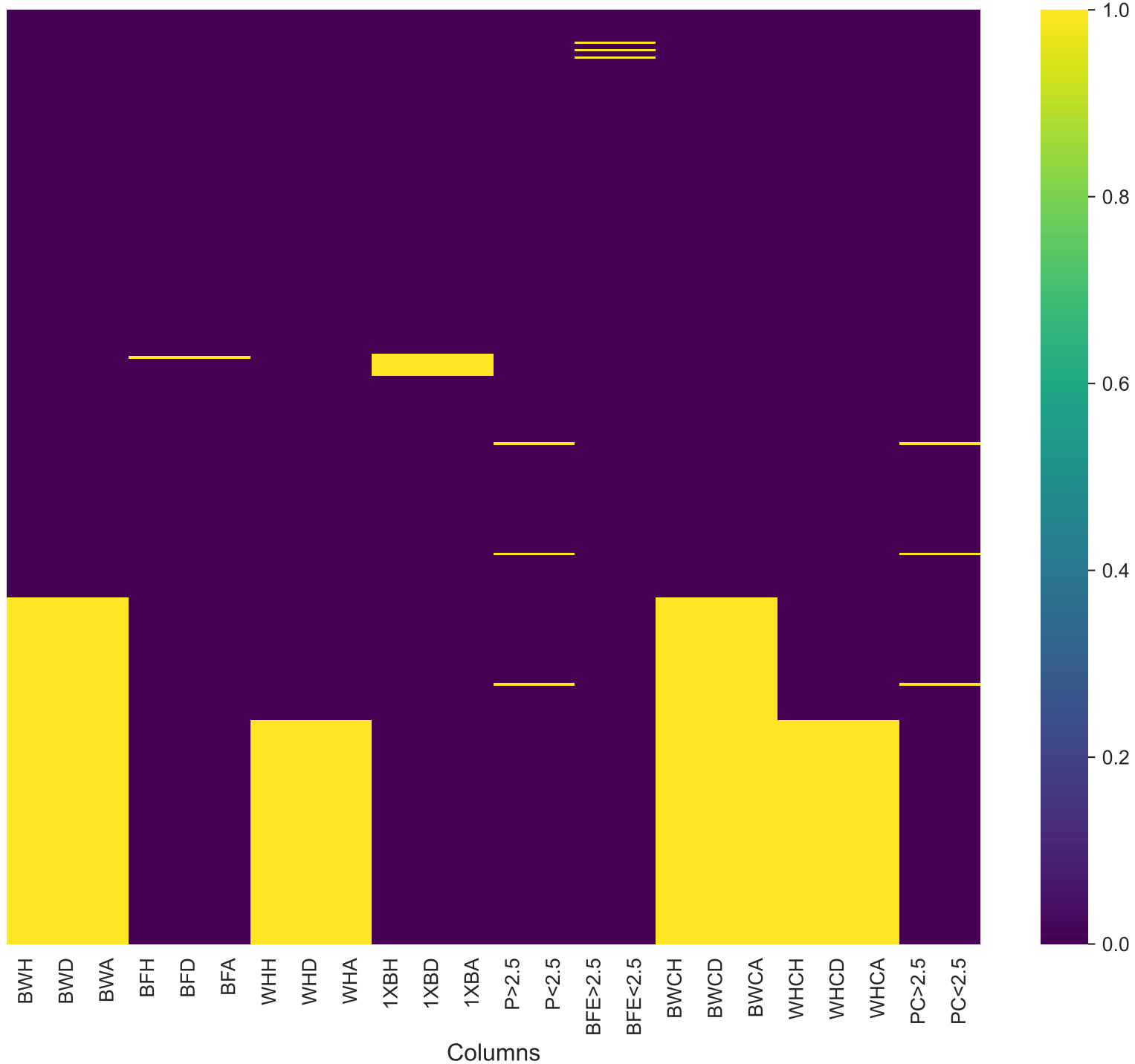
index	count	mean	std	min	25%	50%	75%	max
FTHG	380.0	1.51	1.28	0.0	1.0	1.0	2.0	7.0
FTAG	380.0	1.42	1.19	0.0	1.0	1.0	2.0	6.0
HTHG	380.0	0.75	0.88	0.0	0.0	1.0	1.0	4.0
HTAG	380.0	0.61	0.83	0.0	0.0	0.0	1.0	5.0
HS	380.0	13.75	5.59	2.0	10.0	13.0	17.0	36.0
AS	380.0	12.17	5.41	1.0	9.0	11.5	15.0	37.0
HST	380.0	4.83	2.49	0.0	3.0	5.0	6.0	16.0
AST	380.0	4.27	2.34	0.0	3.0	4.0	6.0	13.0
HF	380.0	10.79	3.64	2.0	8.0	11.0	13.0	21.0
AF	380.0	11.28	3.51	1.0	9.0	11.0	14.0	21.0
HC	380.0	5.43	3.18	0.0	3.0	5.0	7.0	17.0
AC	380.0	4.87	3.03	0.0	3.0	4.0	7.0	18.0
HY	380.0	1.91	1.4	0.0	1.0	2.0	3.0	7.0
AY	380.0	2.14	1.43	0.0	1.0	2.0	3.0	8.0
HR	380.0	0.07	0.26	0.0	0.0	0.0	0.0	2.0
AR	380.0	0.07	0.25	0.0	0.0	0.0	0.0	1.0
B365H	380.0	2.75	1.73	1.08	1.6	2.2	3.0	11.0
B365D	380.0	4.33	1.28	3.1	3.6	3.9	4.5	12.0
B365A	380.0	4.24	3.17	1.2	2.2	3.2	5.25	23.0
BWH	239.0	2.74	1.65	1.09	1.61	2.25	3.1	10.0
BWD	239.0	4.42	1.39	3.2	3.6	4.0	4.75	13.0
BWA	239.0	4.33	3.55	1.28	2.2	3.0	5.25	23.0
BFH	379.0	2.78	1.84	1.08	1.61	2.2	3.1	13.0
BFD	379.0	4.36	1.39	3.1	3.6	3.9	4.6	13.0
BFA	379.0	4.51	3.8	1.17	2.2	3.2	5.0	31.0

1.5 Summary Statistics - Numerical (Part 1)

index	count	mean	std	min	25%	50%	75%	max
PSH	380.0	2.81	1.84	1.08	1.63	2.23	3.15	14.01
PSD	380.0	4.47	1.41	3.15	3.65	4.0	4.74	13.84
PSA	380.0	4.44	3.59	1.19	2.26	3.23	5.4	25.71
WHH	289.0	2.76	1.65	1.08	1.65	2.3	3.2	10.0
WHD	289.0	4.28	1.2	3.1	3.6	3.9	4.5	12.0
WHA	289.0	4.37	3.78	1.29	2.15	3.0	5.0	26.0
1XBH	371.0	2.8	1.85	1.09	1.65	2.22	3.12	15.0
1XBD	371.0	4.52	1.31	3.24	3.75	4.12	4.81	13.6
1XBA	371.0	4.55	3.83	1.2	2.28	3.24	5.34	30.0
MaxH	380.0	2.9	1.99	1.1	1.66	2.28	3.2	15.0
MaxD	380.0	4.63	1.49	3.24	3.78	4.13	4.9	14.0
MaxA	380.0	4.77	4.38	1.2	2.3	3.26	5.52	34.0
AvgH	380.0	2.77	1.79	1.08	1.63	2.22	3.09	13.37
AvgD	380.0	4.39	1.29	3.14	3.62	3.97	4.66	12.4
AvgA	380.0	4.41	3.61	1.19	2.24	3.16	5.26	26.97
BFEH	380.0	2.94	2.05	1.1	1.67	2.29	3.21	15.0
BFED	380.0	4.64	1.6	3.2	3.75	4.1	4.9	15.0
BFEA	380.0	4.87	4.42	1.22	2.32	3.3	5.65	34.0
B365>2.5	380.0	1.64	0.21	1.2	1.5	1.62	1.73	2.38
B365<2.5	380.0	2.4	0.45	1.57	2.1	2.3	2.63	4.5
P>2.5	377.0	1.66	0.21	1.29	1.5	1.63	1.79	2.44
P<2.5	377.0	2.42	0.41	1.6	2.1	2.36	2.67	3.49
Max>2.5	380.0	1.68	0.22	1.21	1.52	1.66	1.81	2.44
Max<2.5	380.0	2.5	0.49	1.66	2.16	2.41	2.79	5.0
Avg>2.5	380.0	1.64	0.21	1.19	1.48	1.61	1.76	2.36

1.5 Summary Statistics - Numerical (Part 2)

index	count	mean	std	min	25%	50%	75%	max
Avg<2.5	380.0	2.4	0.44	1.59	2.08	2.34	2.64	4.41
BFE>2.5	377.0	1.68	0.22	1.04	1.53	1.65	1.82	2.44
BFE<2.5	377.0	2.52	0.5	1.16	2.16	2.44	2.78	4.8
AHh	380.0	-0.31	0.9	-2.75	-1.0	-0.25	0.25	2.0
B365AHH	380.0	1.94	0.09	1.73	1.87	1.95	2.02	2.11
B365AHA	380.0	1.95	0.09	1.7	1.88	1.95	2.03	2.11
PAHH	380.0	1.95	0.09	1.77	1.88	1.94	2.03	2.16
PAHA	380.0	1.96	0.09	1.76	1.88	1.96	2.03	2.16
MaxAHH	380.0	1.97	0.09	1.81	1.89	1.97	2.04	2.16
MaxAHA	380.0	1.98	0.09	1.8	1.9	1.98	2.05	2.2
AvgAHH	380.0	1.92	0.08	1.76	1.85	1.92	1.99	2.1
AvgAHA	380.0	1.93	0.08	1.77	1.85	1.93	1.99	2.09
BFEAHH	380.0	1.97	0.1	1.37	1.9	1.97	2.06	2.24
BFEAHA	380.0	1.99	0.11	1.14	1.91	2.0	2.06	2.22
B365CH	380.0	2.75	1.78	1.07	1.62	2.17	3.2	13.0
B365CD	380.0	4.3	1.24	3.0	3.6	3.9	4.5	12.0
B365CA	380.0	4.32	3.32	1.2	2.2	3.2	5.06	23.0
BWCH	239.0	2.73	1.61	1.08	1.62	2.2	3.25	9.5
BWCD	239.0	4.36	1.39	3.1	3.6	3.9	4.68	13.5
BWCA	239.0	4.38	3.66	1.28	2.17	3.2	5.12	29.0
BFCH	380.0	2.77	1.86	1.06	1.6	2.2	3.25	14.0
BFCD	380.0	4.27	1.33	2.9	3.5	3.9	4.53	13.0
BFCA	380.0	4.54	3.87	1.15	2.2	3.22	5.0	26.0
PSCH	380.0	2.85	1.95	1.09	1.65	2.24	3.3	15.0
PSCD	380.0	4.44	1.37	3.0	3.63	3.97	4.7	12.25

index	count	mean	std	min	25%	50%	75%	max
PSCA	380.0	4.69	4.12	1.19	2.25	3.27	5.4	31.0
WHCH	289.0	2.77	1.67	1.08	1.65	2.25	3.3	10.0
WHCD	289.0	4.24	1.24	2.9	3.5	3.8	4.5	13.0
WHCA	289.0	4.46	3.9	1.25	2.15	3.1	5.0	29.0
1XBCH	380.0	2.88	1.92	1.08	1.65	2.26	3.38	14.1
1XBCD	380.0	4.43	1.38	3.05	3.64	3.99	4.62	14.7
1XBCA	380.0	4.5	3.66	1.22	2.18	3.26	5.37	28.0
MaxCH	380.0	2.97	2.05	1.1	1.68	2.31	3.39	15.0
MaxCD	380.0	4.59	1.53	3.05	3.71	4.1	4.86	15.0
MaxCA	380.0	4.94	4.61	1.22	2.29	3.34	5.61	38.0
AvgCH	380.0	2.8	1.85	1.08	1.63	2.22	3.21	13.84
AvgCD	380.0	4.37	1.32	2.98	3.58	3.92	4.64	12.79
AvgCA	380.0	4.52	3.8	1.19	2.22	3.2	5.26	27.5
BFECH	380.0	2.96	2.14	1.1	1.66	2.29	3.36	17.0
BFECD	380.0	4.59	1.52	3.1	3.7	4.0	4.82	15.0
BFECA	380.0	5.0	4.65	1.21	2.28	3.35	5.7	32.0
B365C>2.5	380.0	1.66	0.24	1.17	1.5	1.62	1.8	2.63
B365C<2.5	380.0	2.39	0.5	1.5	2.03	2.3	2.63	5.0
PC>2.5	377.0	1.69	0.24	1.29	1.51	1.66	1.83	2.62
PC<2.5	377.0	2.4	0.44	1.53	2.07	2.32	2.66	3.54
MaxC>2.5	380.0	1.71	0.25	1.24	1.53	1.68	1.87	2.63
MaxC<2.5	380.0	2.53	0.54	1.59	2.13	2.42	2.79	5.4
AvgC>2.5	380.0	1.66	0.23	1.18	1.49	1.63	1.8	2.52
AvgC<2.5	380.0	2.4	0.49	1.54	2.05	2.31	2.66	4.82
BFEC>2.5	380.0	1.72	0.25	1.21	1.54	1.69	1.86	2.66

1.5 Summary Statistics - Numerical (Part 4)

1.5 Summary Statistics - Numerical (Part 5)

index	count	mean	std	min	25%	50%	75%	max
BFEC<2.5	380.0	2.54	0.56	1.59	2.14	2.43	2.82	5.5
AHCh	380.0	-0.31	0.91	-3.0	-1.0	-0.25	0.25	2.0
B365CAHH	380.0	1.95	0.09	1.68	1.88	1.95	2.02	2.1
B365CAHA	380.0	1.96	0.09	1.73	1.91	1.97	2.04	2.15
PCAHH	380.0	1.95	0.08	1.71	1.88	1.95	2.02	2.17
PCAHA	380.0	1.97	0.09	1.77	1.9	1.97	2.04	2.26
MaxCAHH	380.0	1.99	0.09	1.83	1.91	1.99	2.05	2.34
MaxCAHA	380.0	2.0	0.09	1.8	1.94	2.0	2.07	2.31
AvgCAHH	380.0	1.93	0.08	1.73	1.86	1.92	1.99	2.13
AvgCAHA	380.0	1.94	0.08	1.75	1.87	1.94	2.01	2.14
BFECAHH	380.0	1.98	0.09	1.75	1.91	1.98	2.05	2.21
BFECAHA	380.0	2.01	0.09	1.81	1.93	2.01	2.07	2.3
Month	380.0	6.68	3.96	1.0	3.0	5.0	11.0	12.0

1.6 Summary Statistics - Categorical

index	count	unique	top	freq
Div	380	1	E0	380
Time	380	16	15:00	137
HomeTeam	380	20	Man United	19
AwayTeam	380	20	Fulham	19
FTR	380	3	H	155
HTR	380	3	H	142
Referee	380	26	A Taylor	30
MonthName	380	10	December	62

2.1 Full-Time Results Distribution
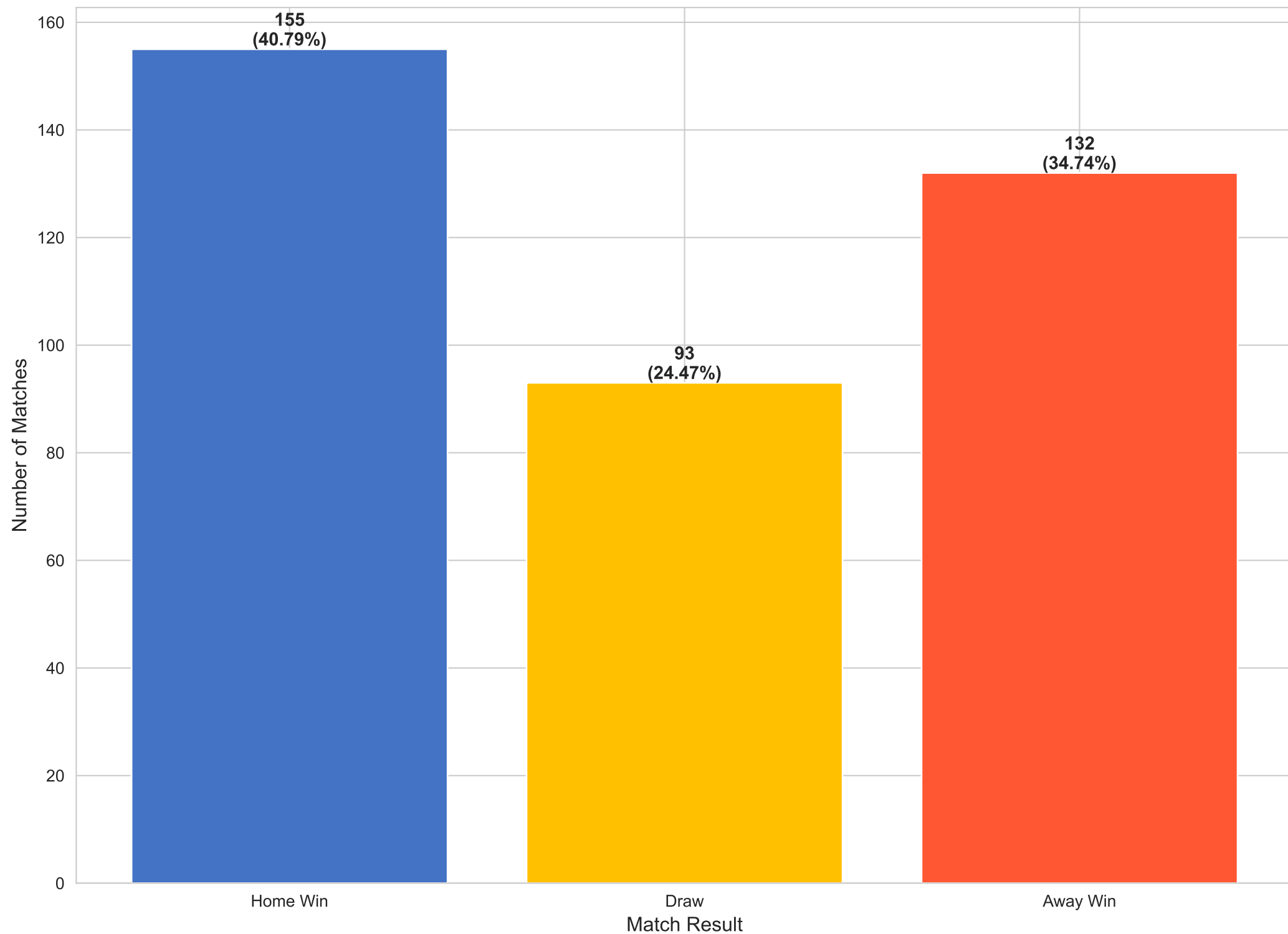
2.2 Half-Time Results Distribution
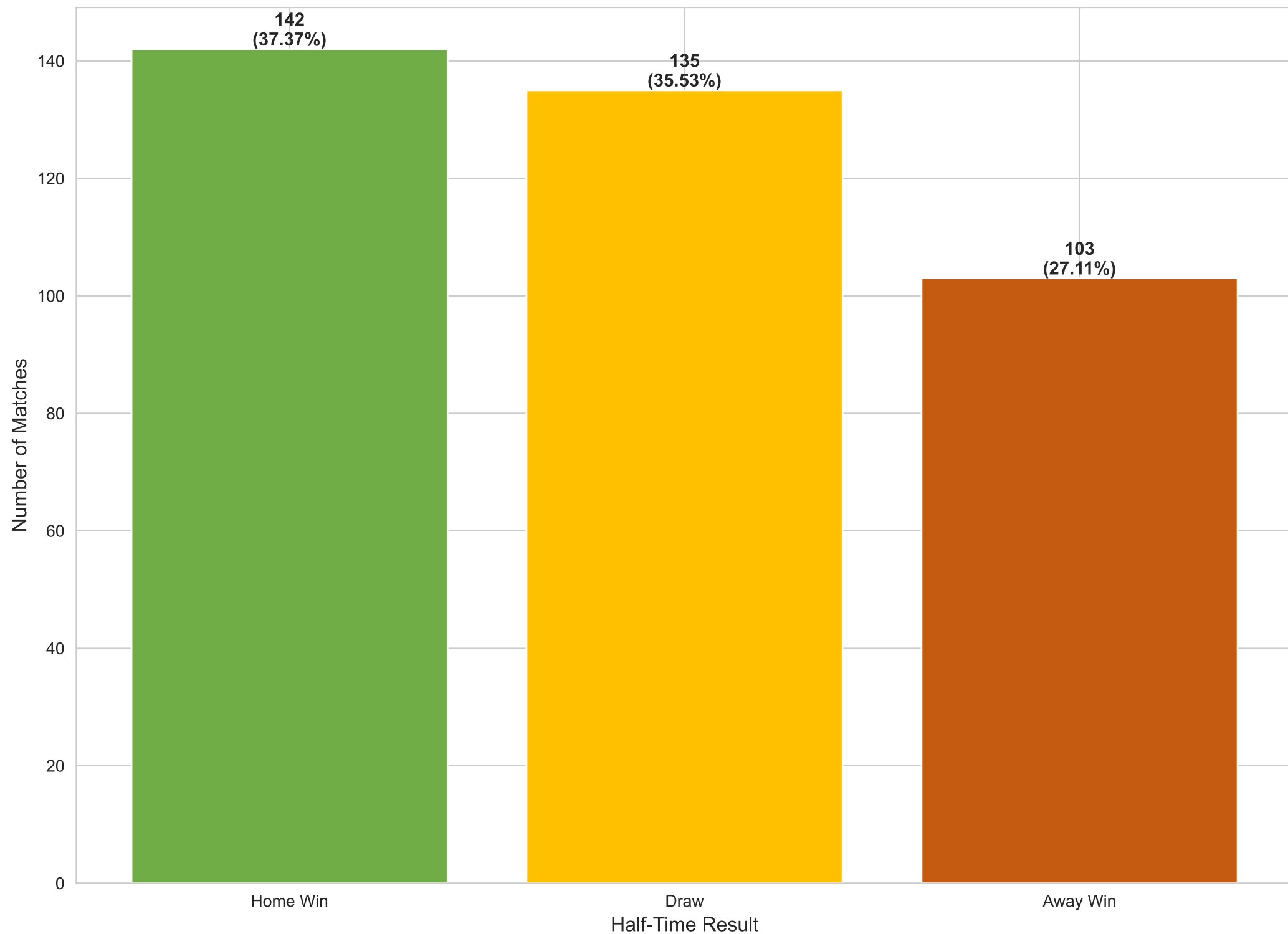
2.3 Half-Time vs Full-Time Results Cross-Tabulation
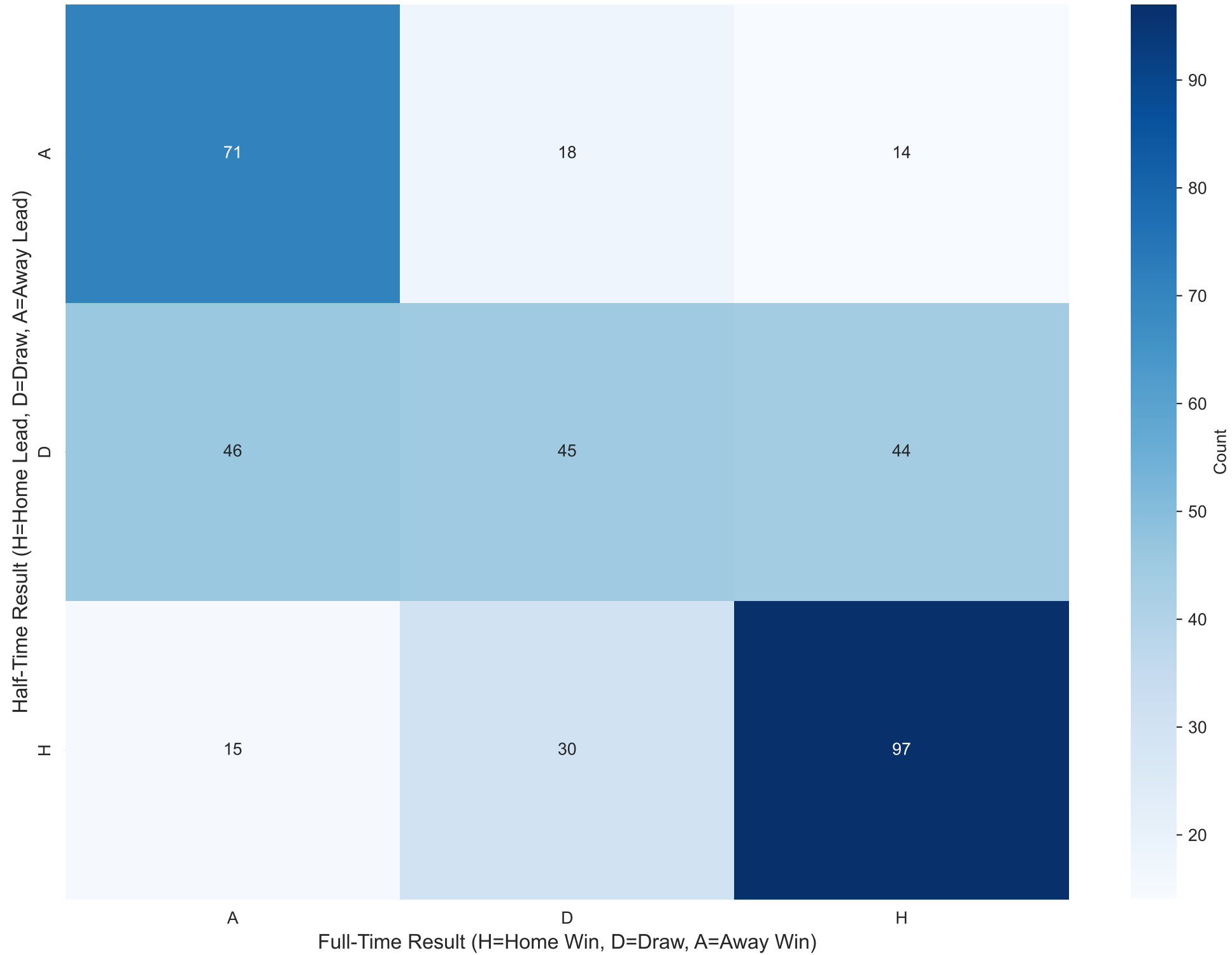
Full-Time Result (H=Home Win, D=Draw, A=Away Win)
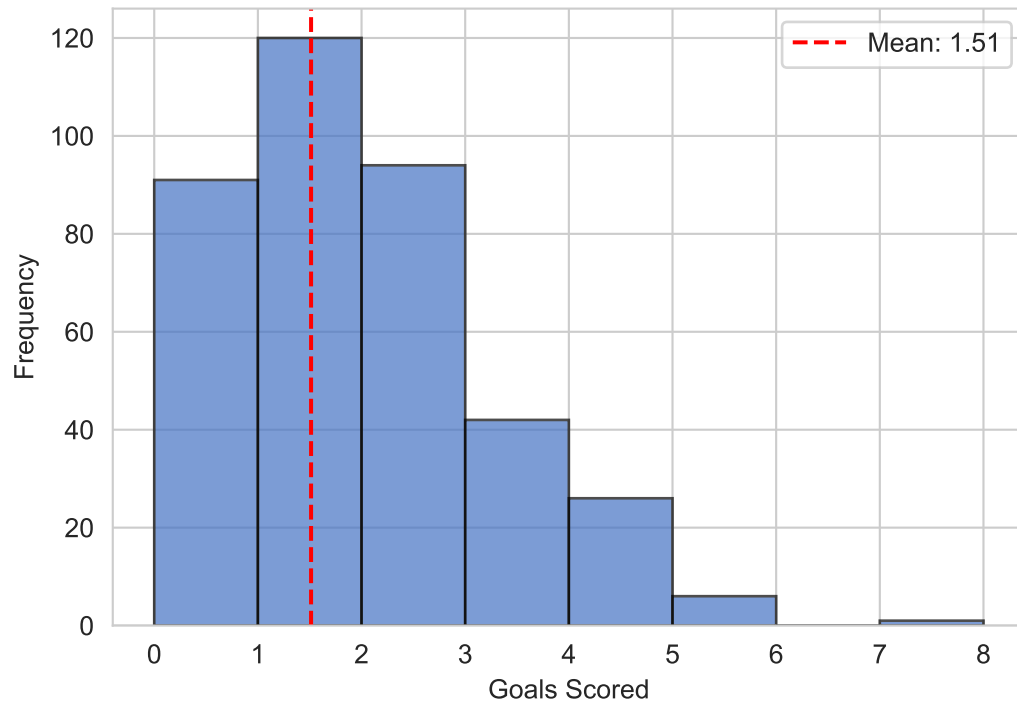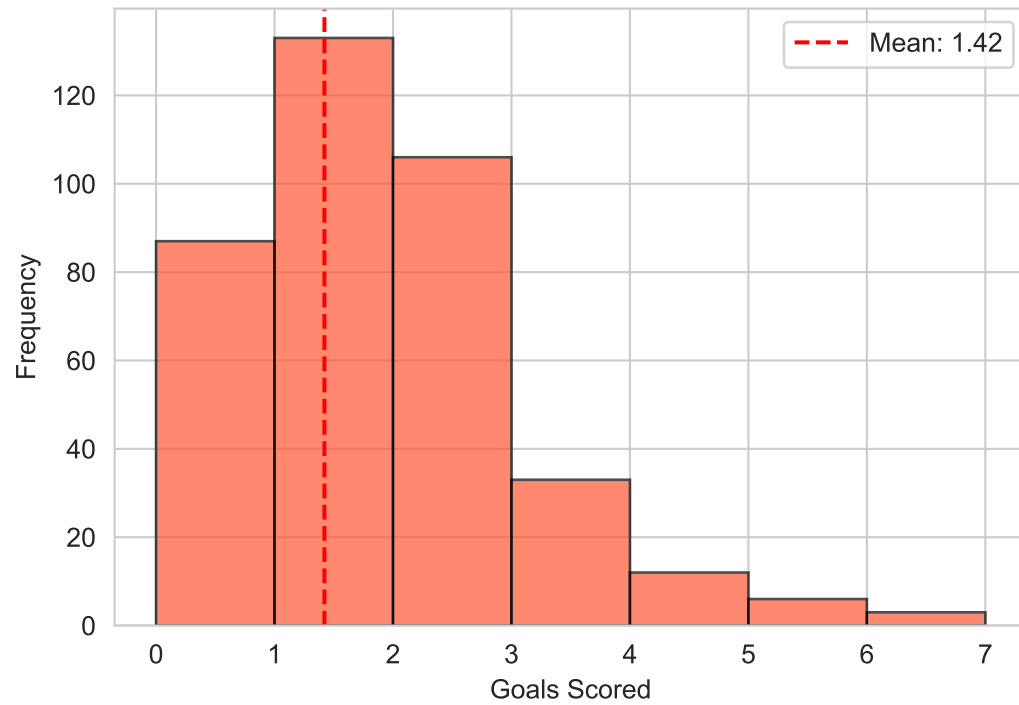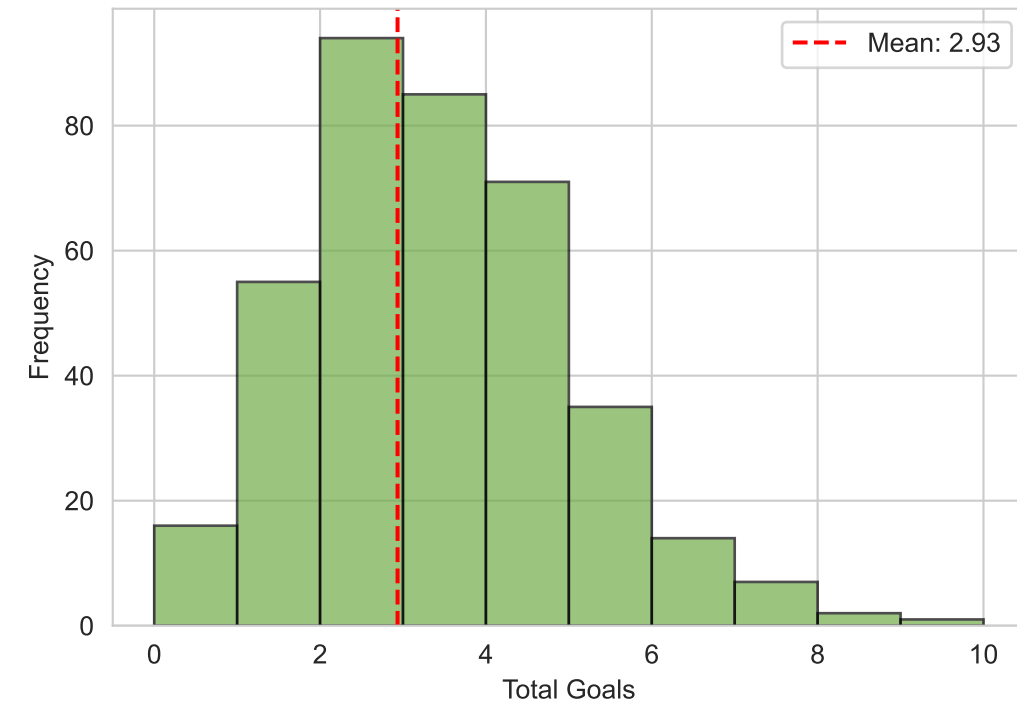
2.4 Goal Distributions
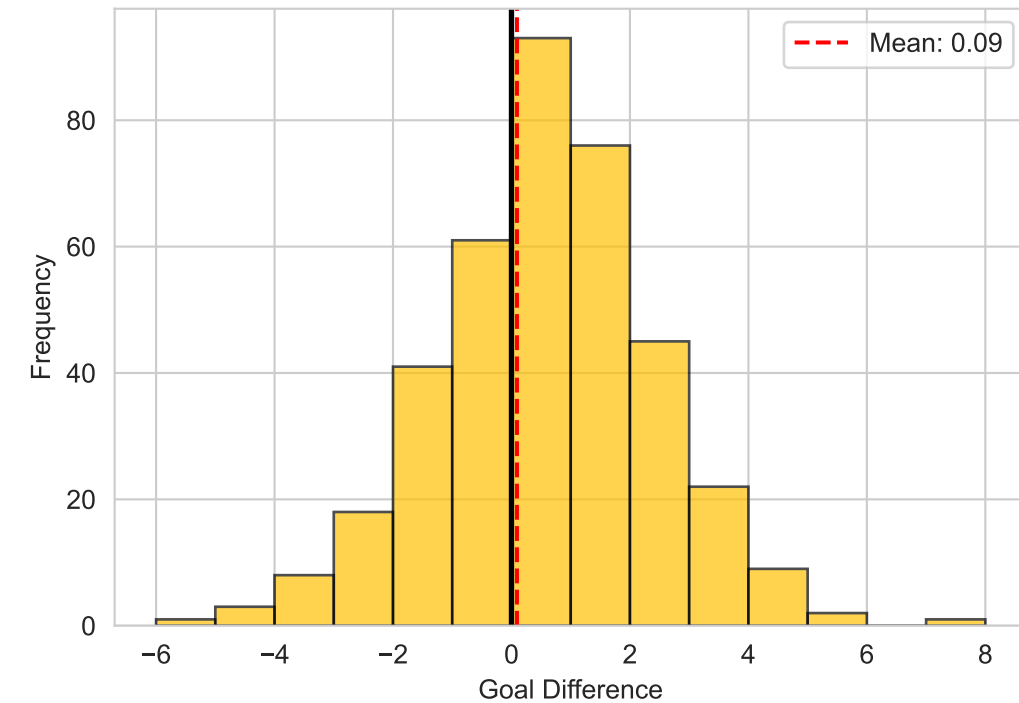
2.5 Home vs Away Goals Distribution
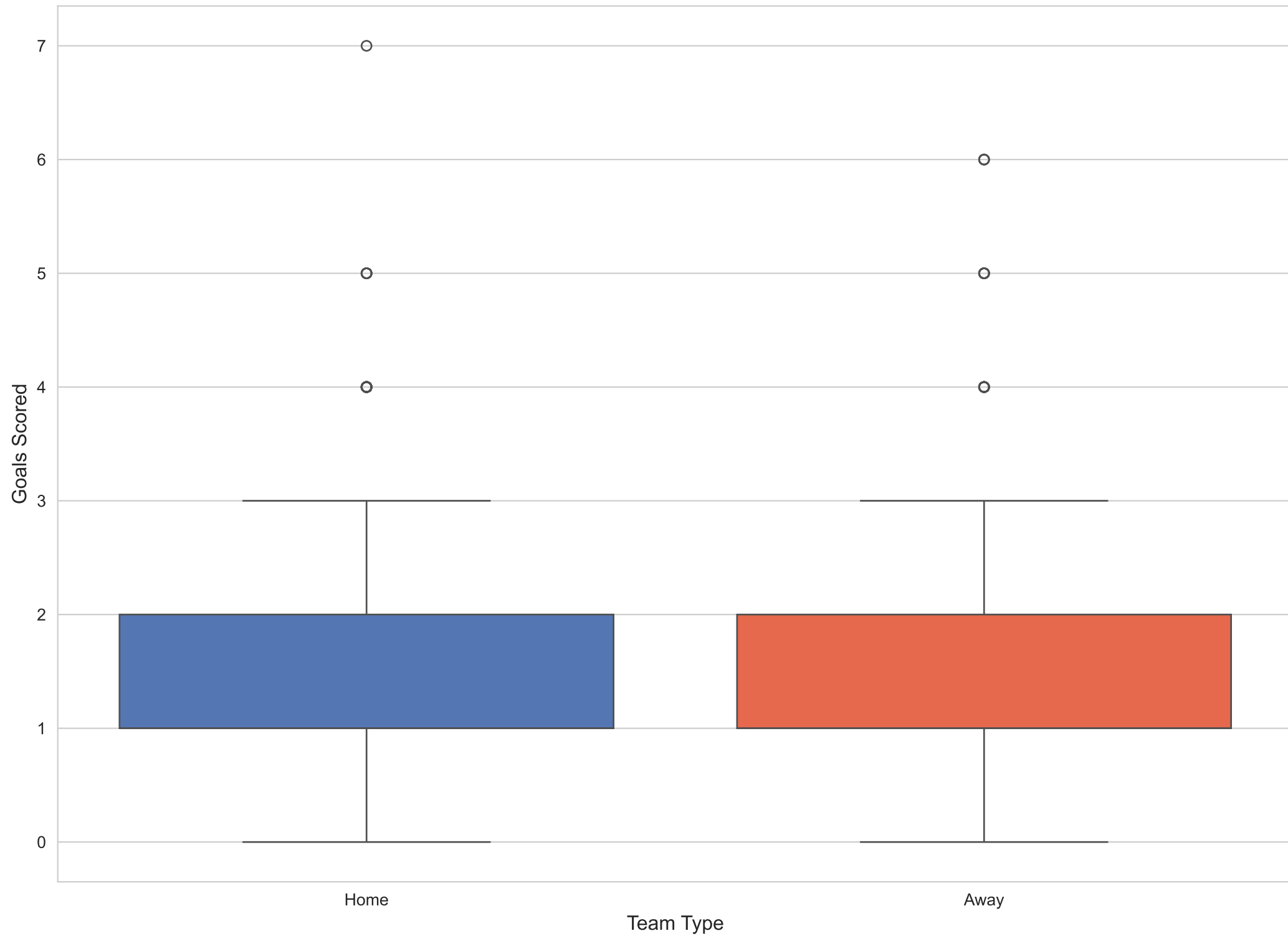
3.1 Top 10 Teams by Total Wins

Team	TotalWins	HomeWins	AwayWins
Liverpool	25	14	11
Man City	21	13	8
Arsenal	20	11	9
Chelsea	20	12	8
Newcastle	20	12	8
Nott'm Forest	19	9	10
Aston Villa	19	11	8
Brentford	16	9	7
Brighton	16	8	8
Bournemouth	15	8	7

3.2 Top 10 Teams by Total Goals Scored

Team	TotalGoalsScored	HomeGoalsScored	AwayGoalsScored
Liverpool	86	42	44
Man City	72	43	29
Arsenal	69	35	34
Newcastle	68	40	28
Brentford	66	40	26
Brighton	66	30	36
Chelsea	64	35	29
Tottenham	64	35	29
Bournemouth	58	23	35
Nott'm Forest	58	26	32

3.3 Top 10 Teams by Fewest Goals Conceded

Team	TotalGoalsConceded	HomeGoalsConceded	AwayGoalsConceded
Arsenal	34	17	17
Liverpool	41	16	25
Chelsea	43	18	25
Everton	44	23	21
Man City	44	23	21
Bournemouth	46	16	30
Nott'm Forest	46	16	30
Newcastle	47	20	27
Aston Villa	51	20	31
Crystal Palace	51	26	25

3.4 Average Goals Scored: Home vs Away (by Team)
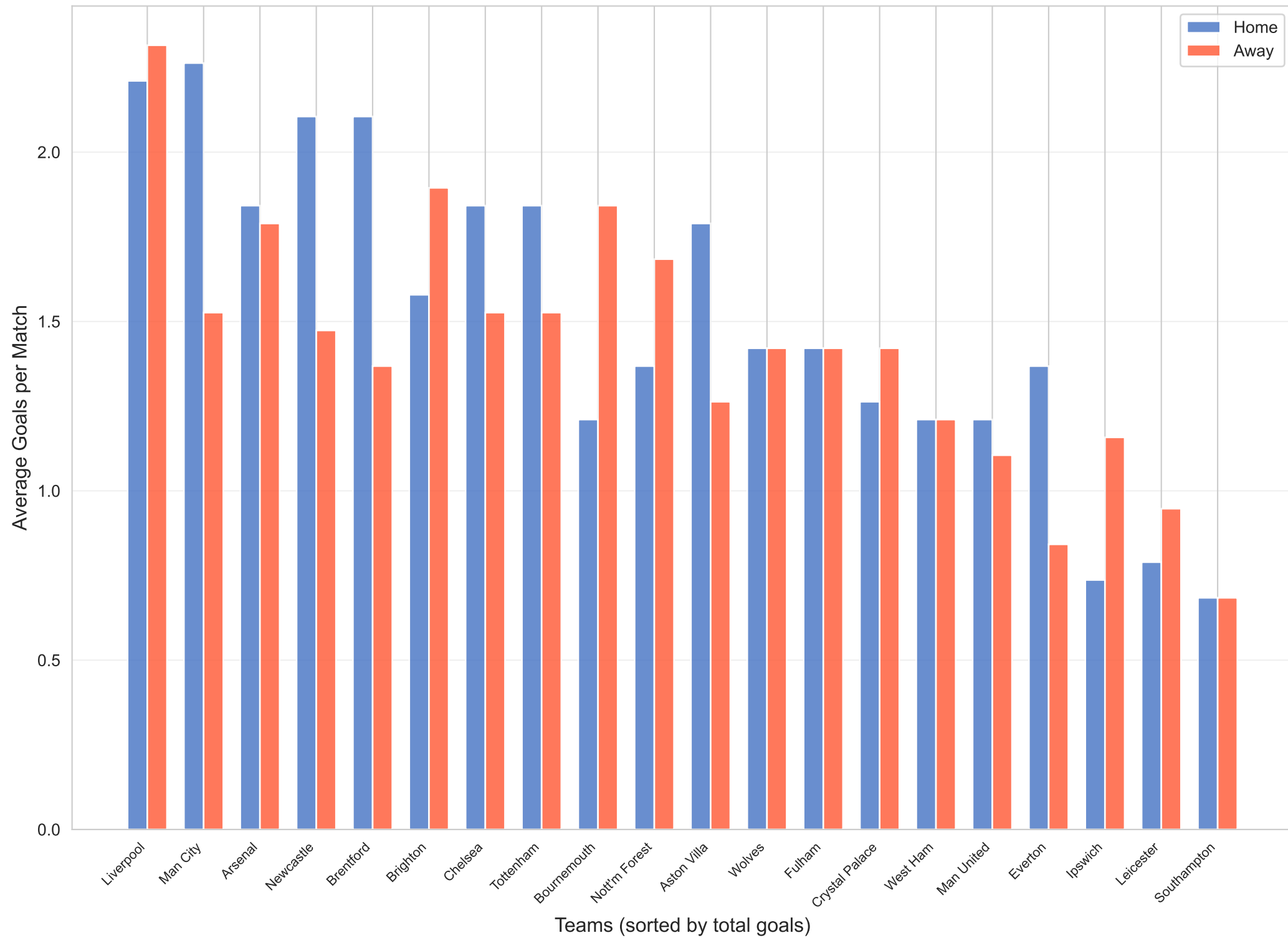
3.5 Average Goals Conceded: Home vs Away (by Team)
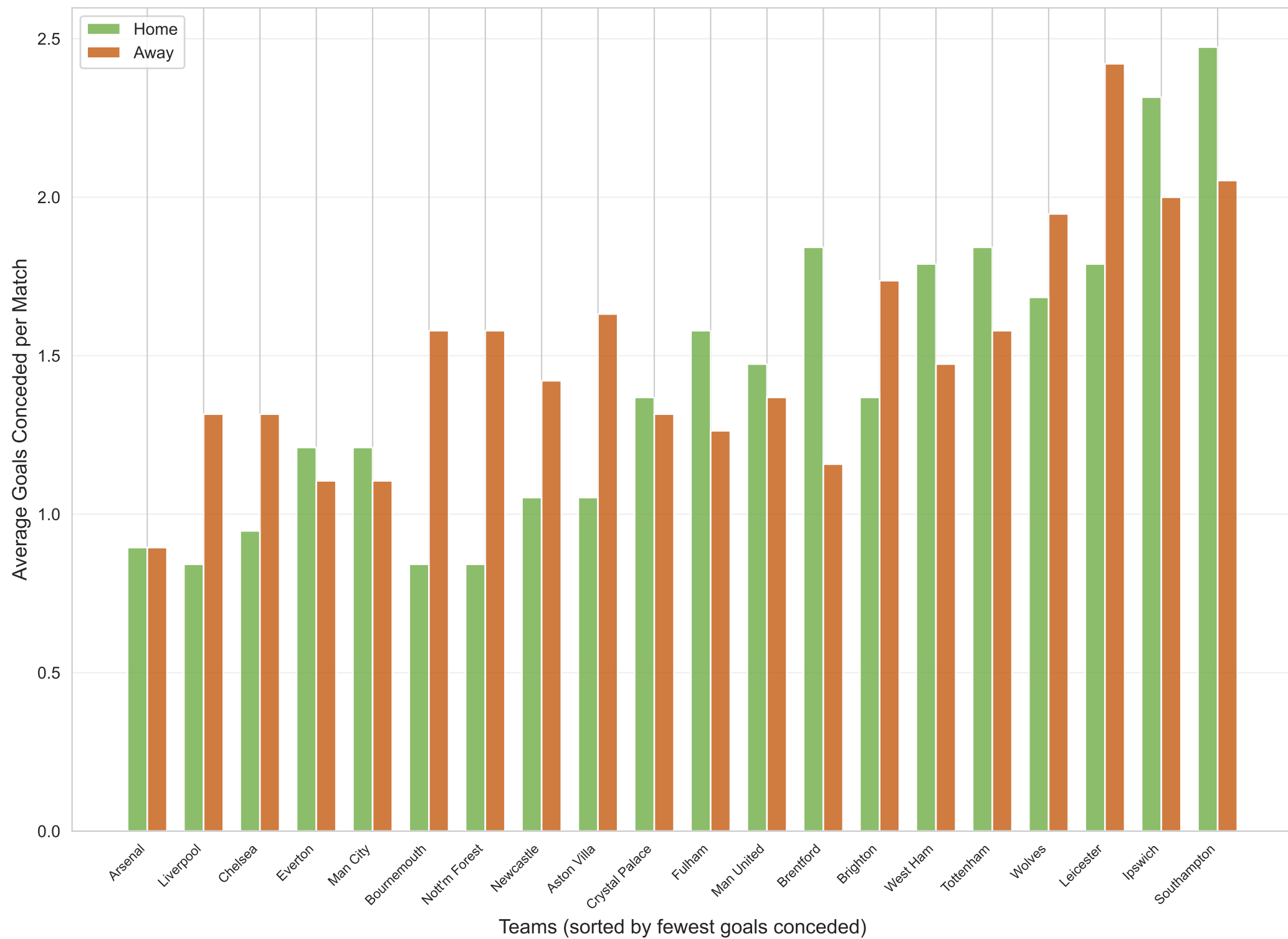
3.6 Goal Difference by Team (Goals Scored - Goals Conceded)
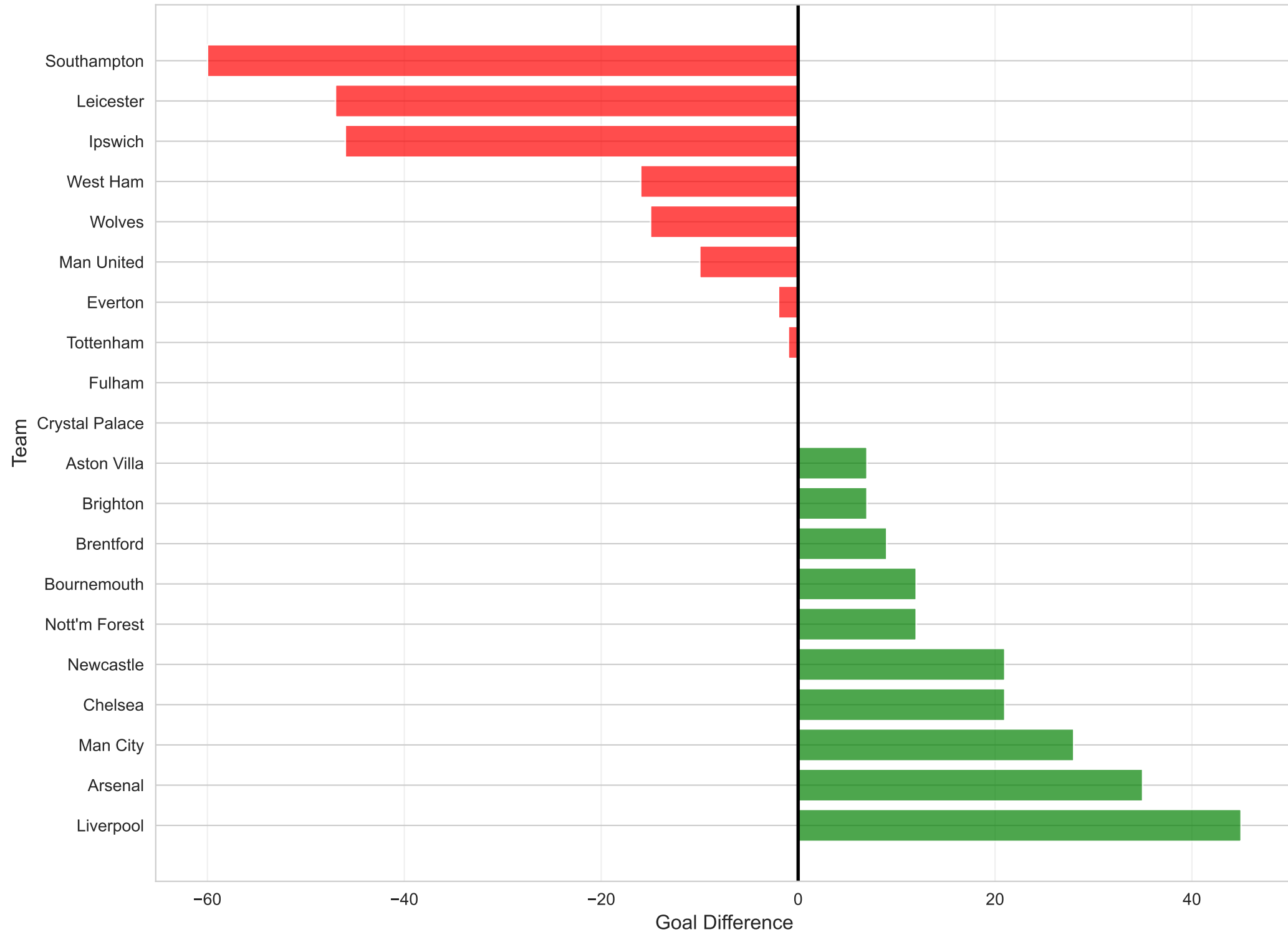
4.1 Match Statistics Distribution
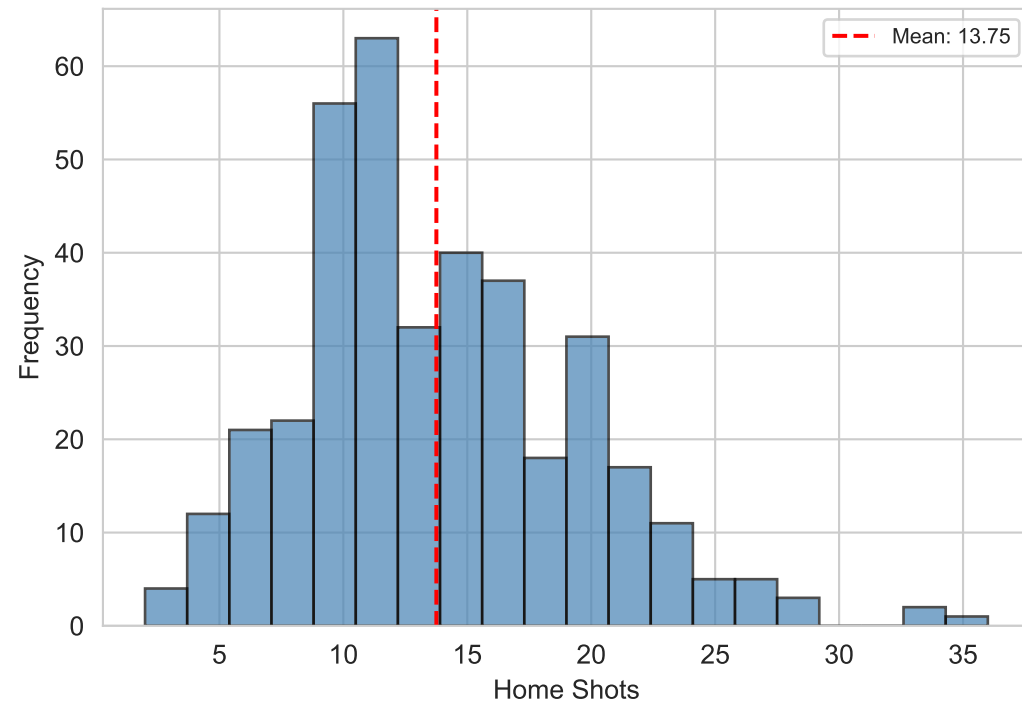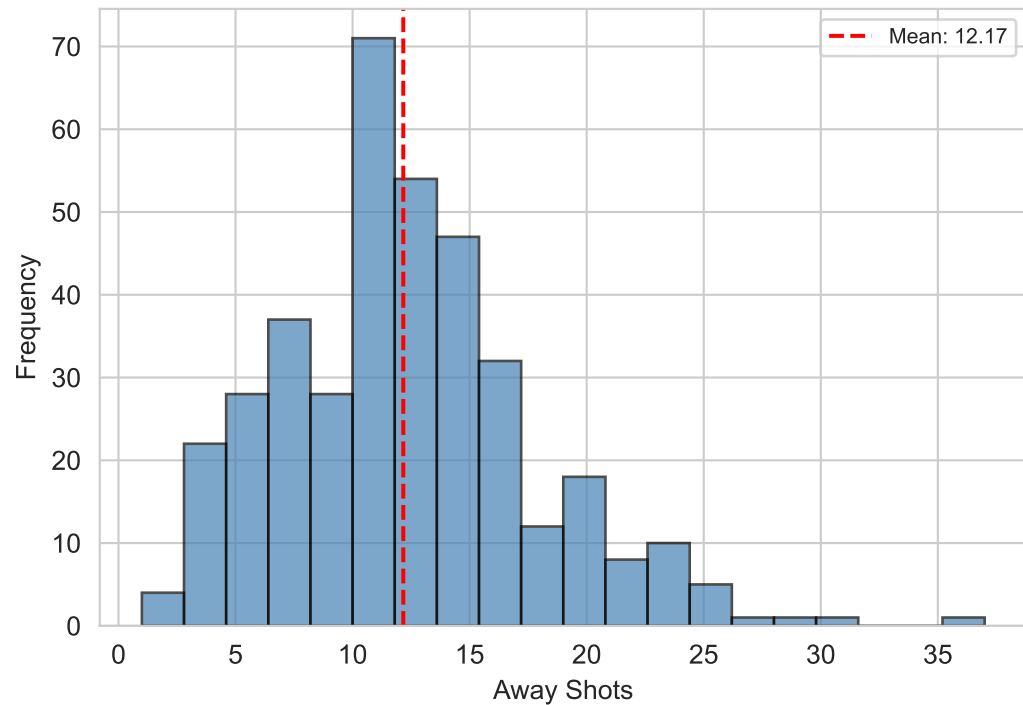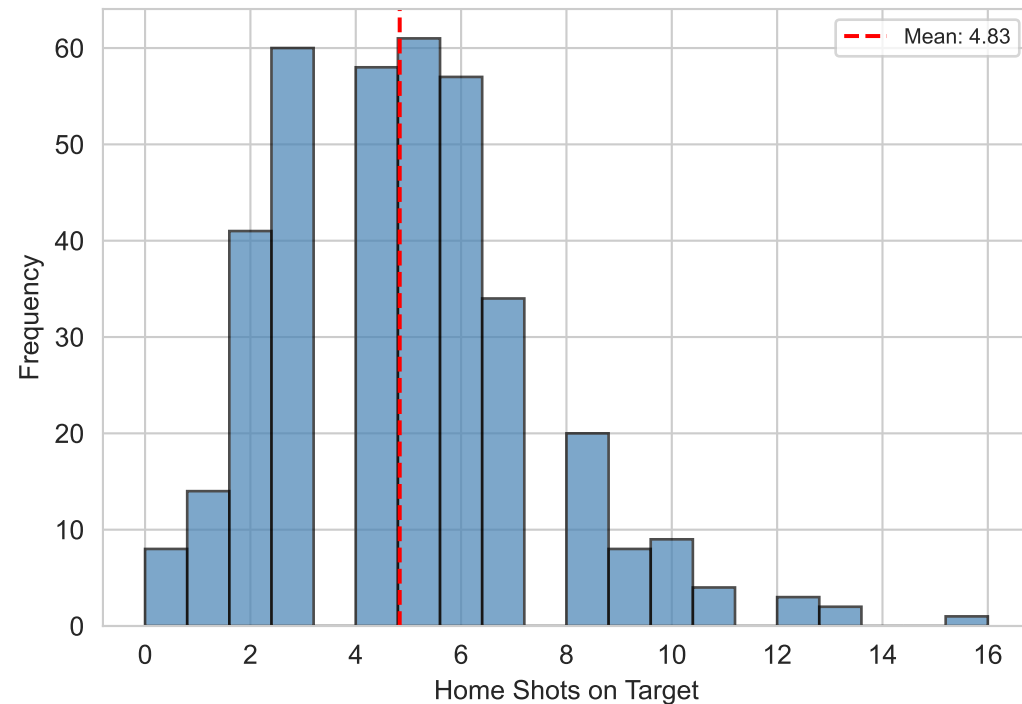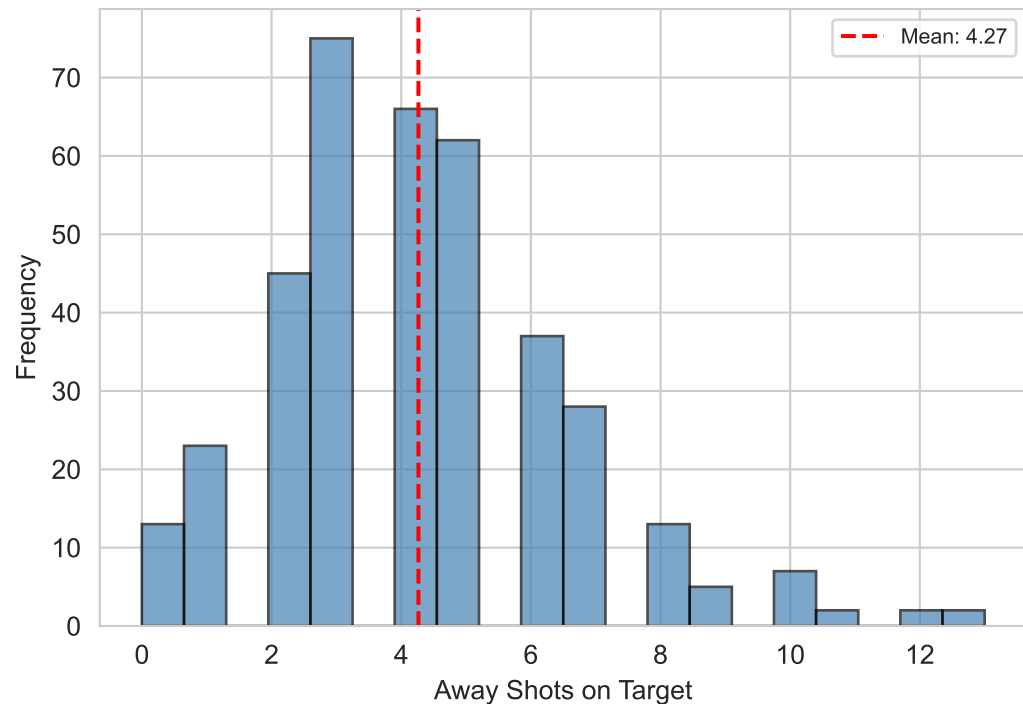
4.2 Match Statistics Distribution
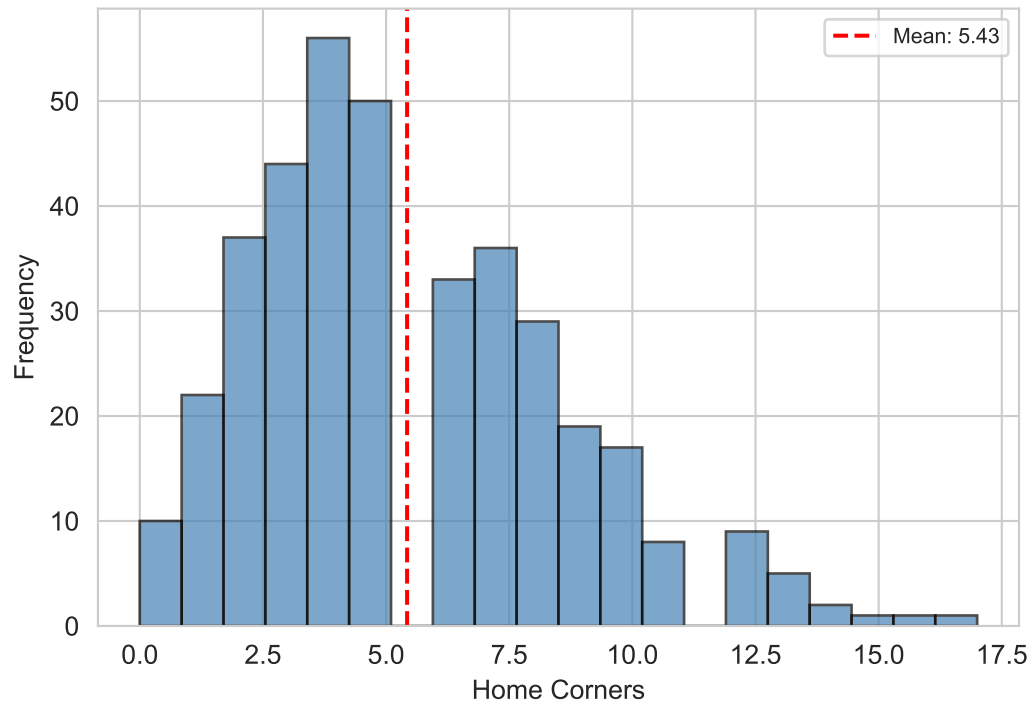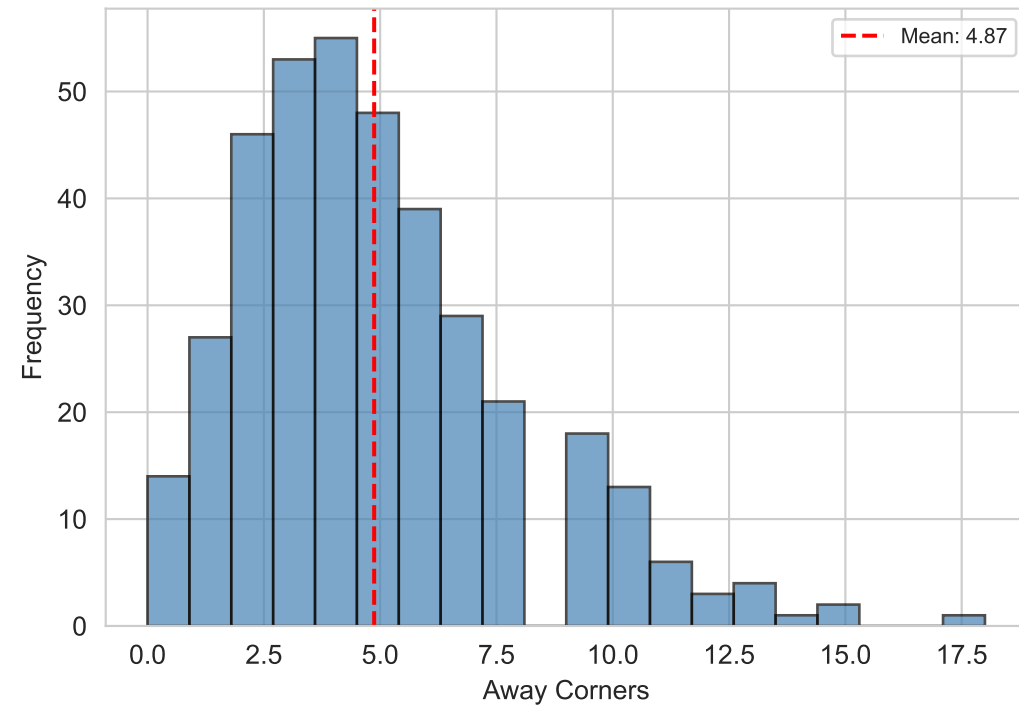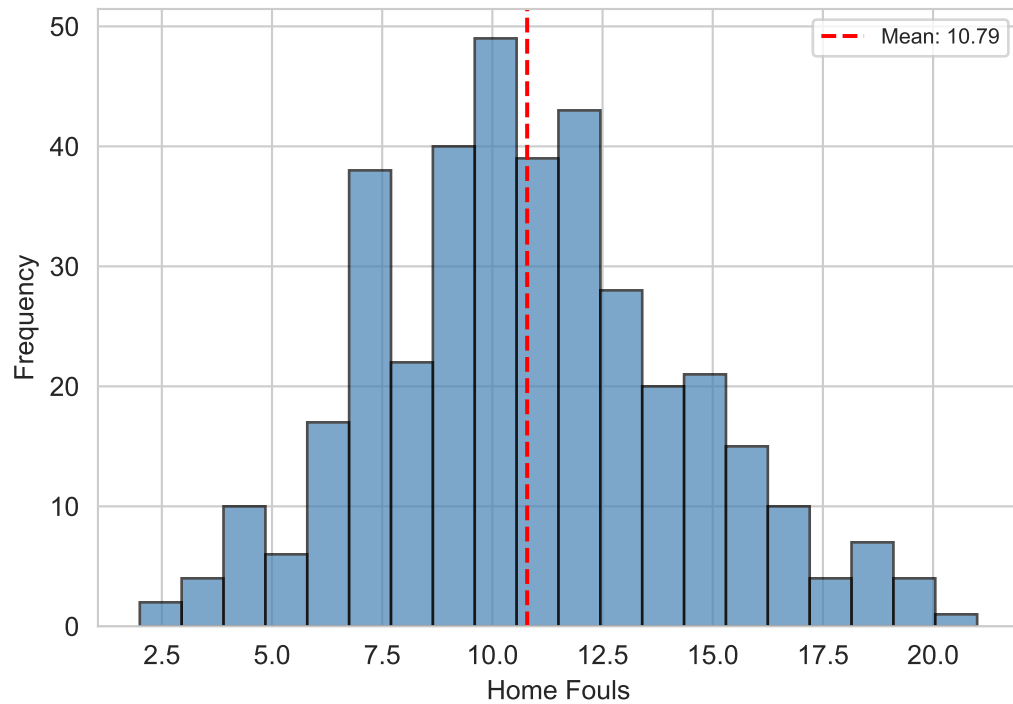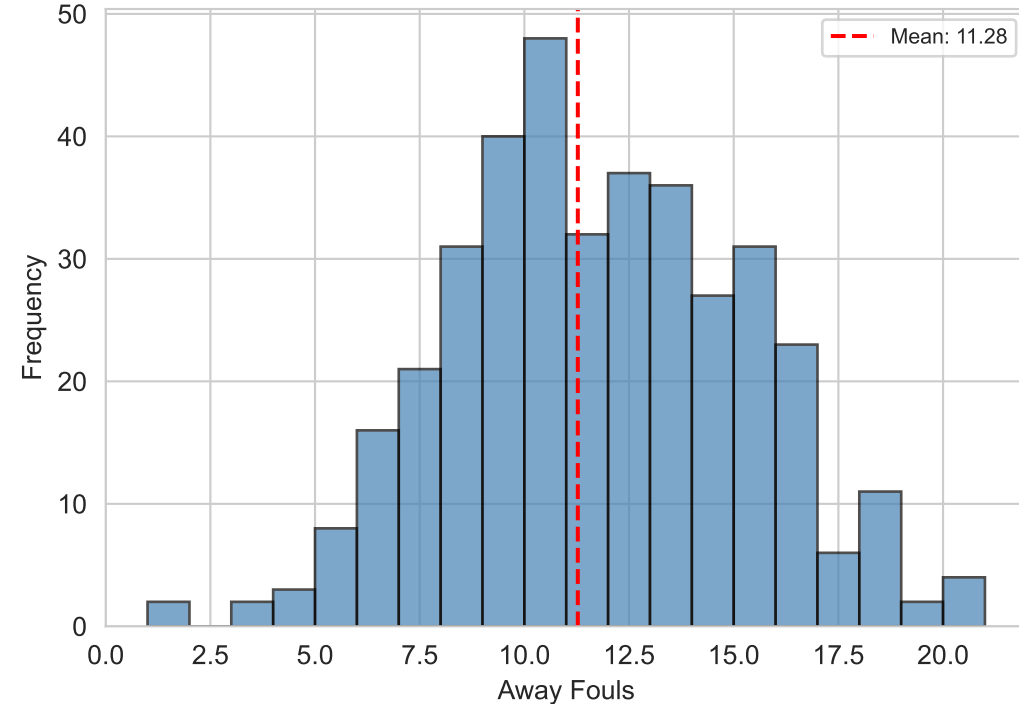
4.3 Home Advantage Analysis: Home vs Away Statistics
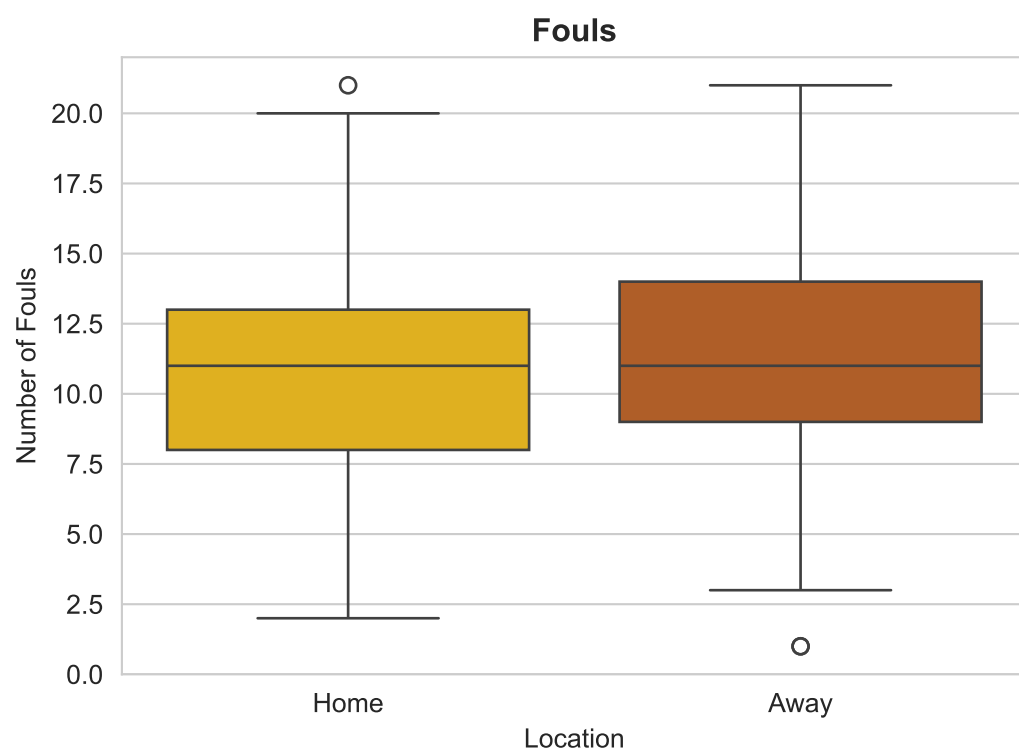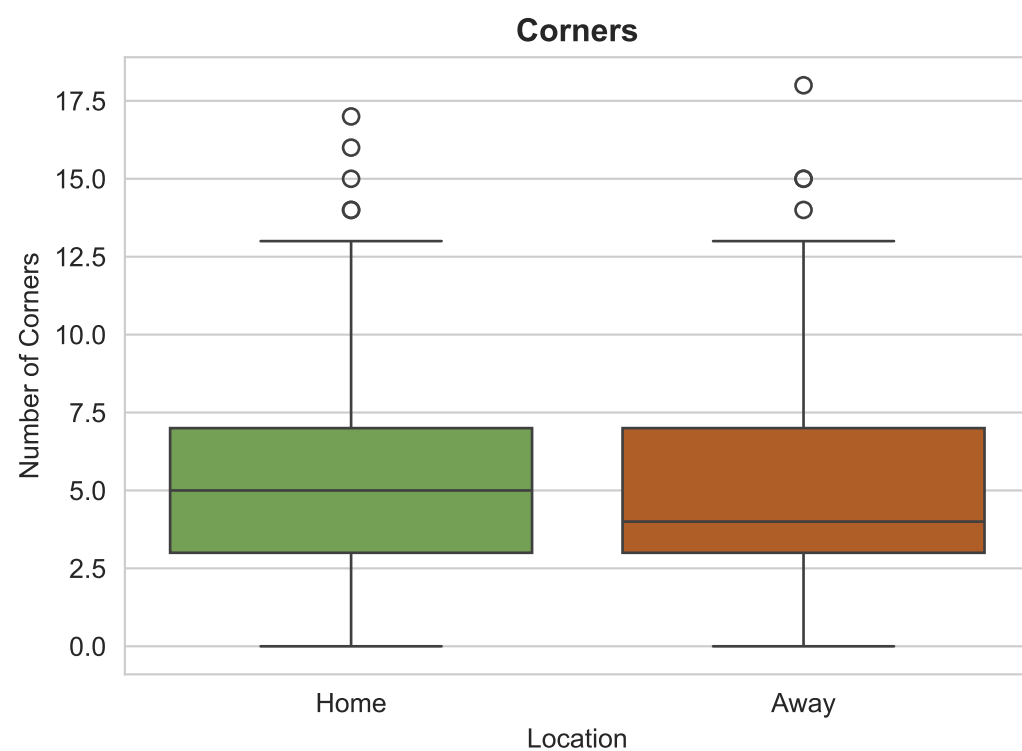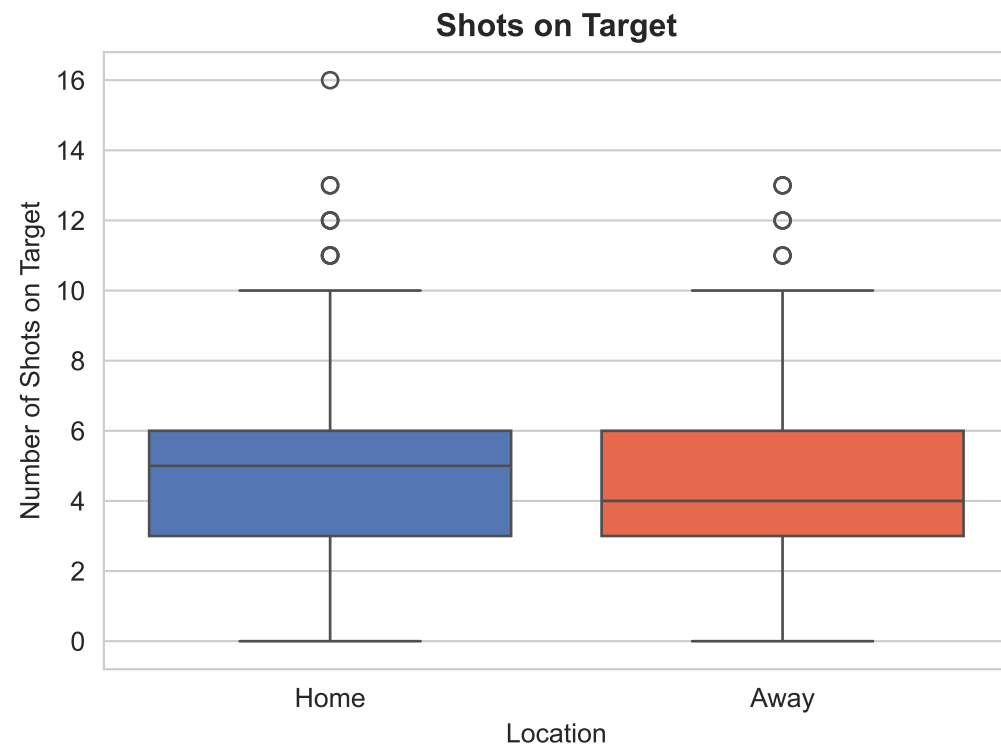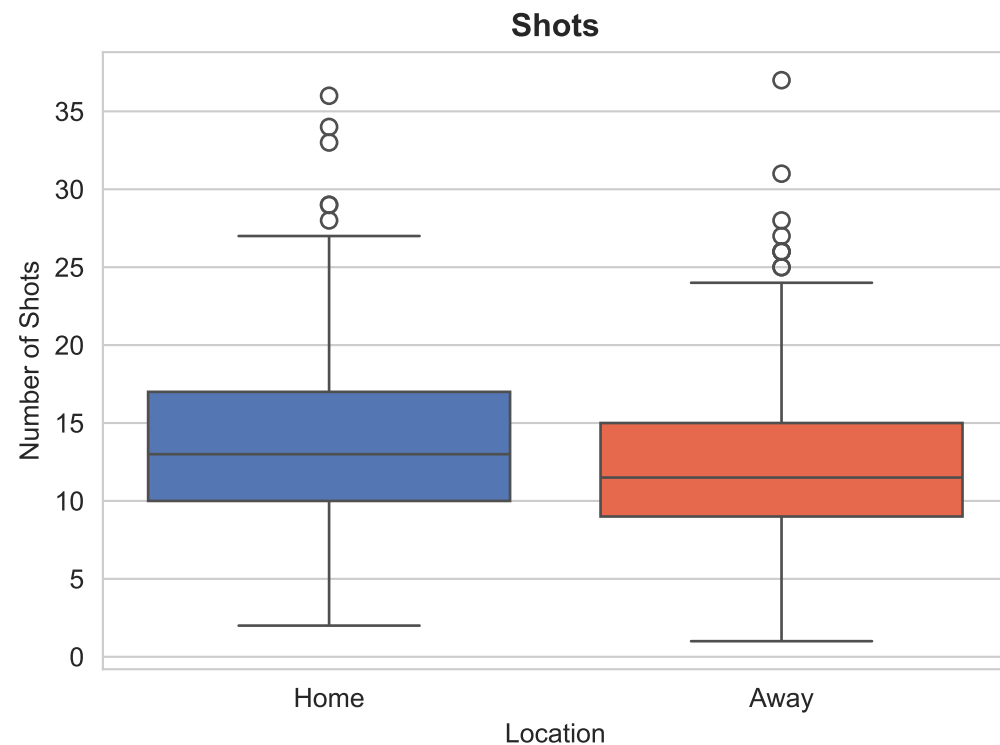
4.4 Disciplinary Analysis
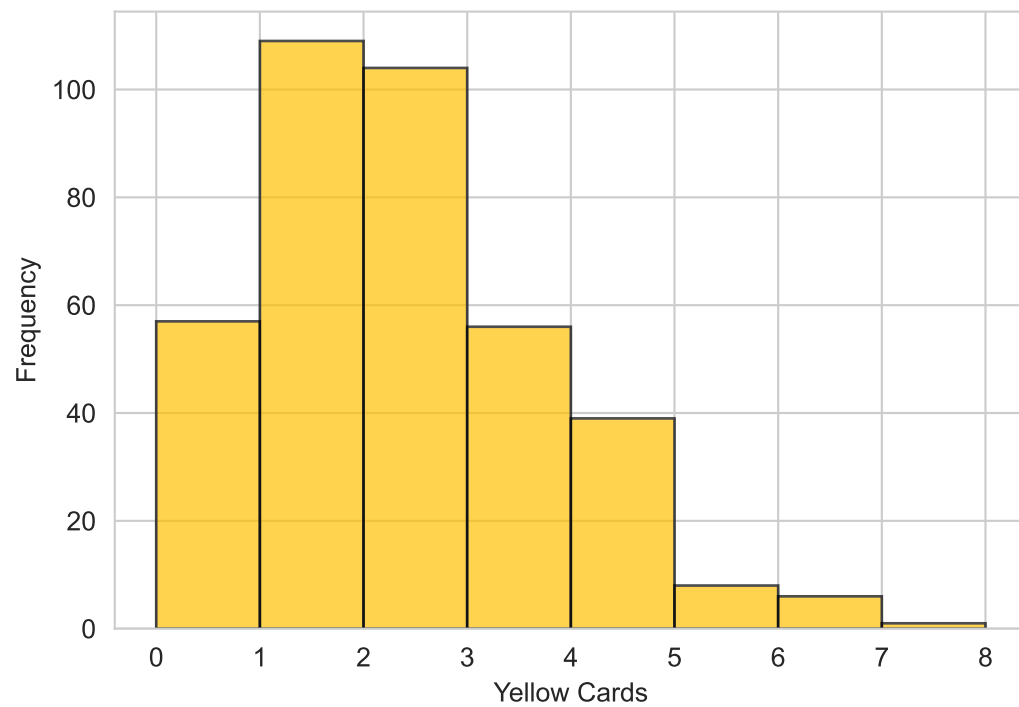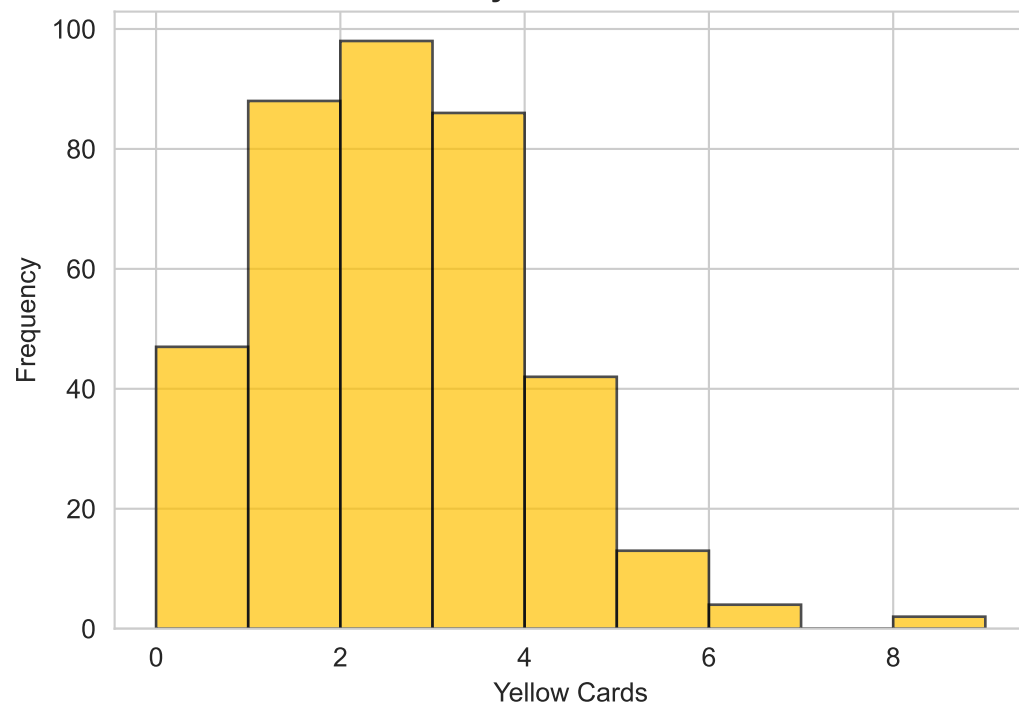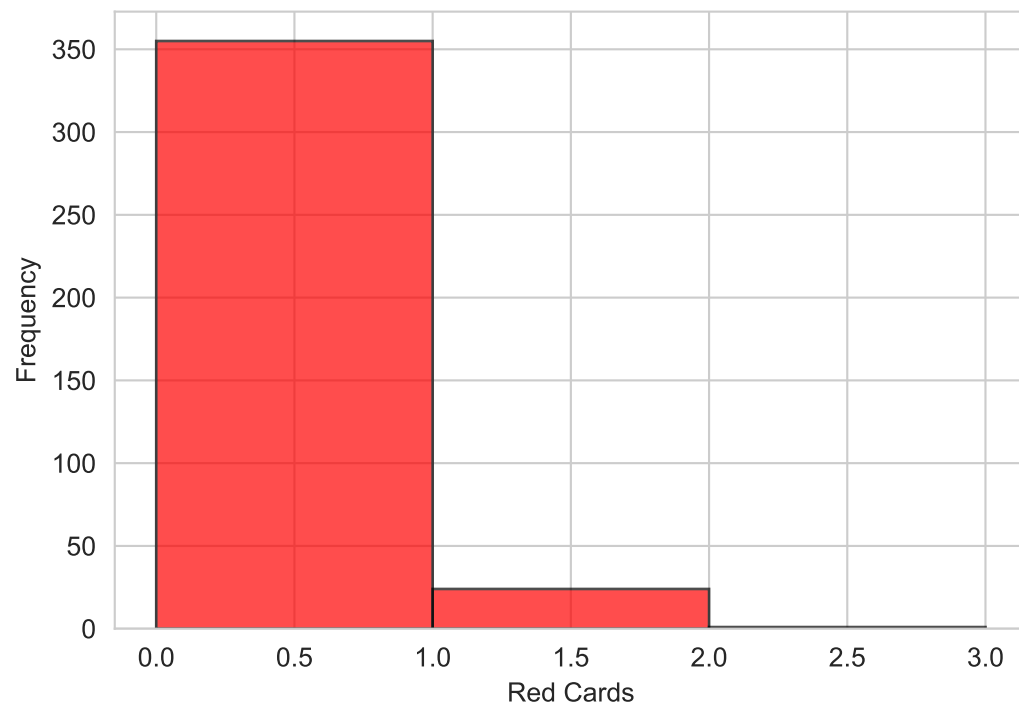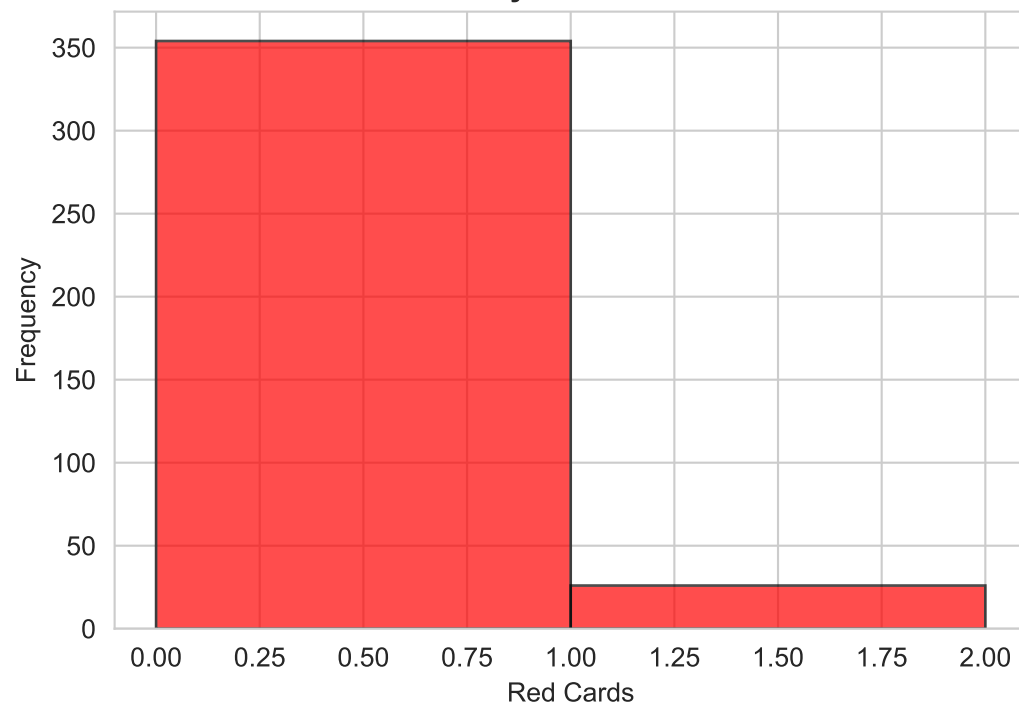
4.5 Total Red Cards Issued per Team
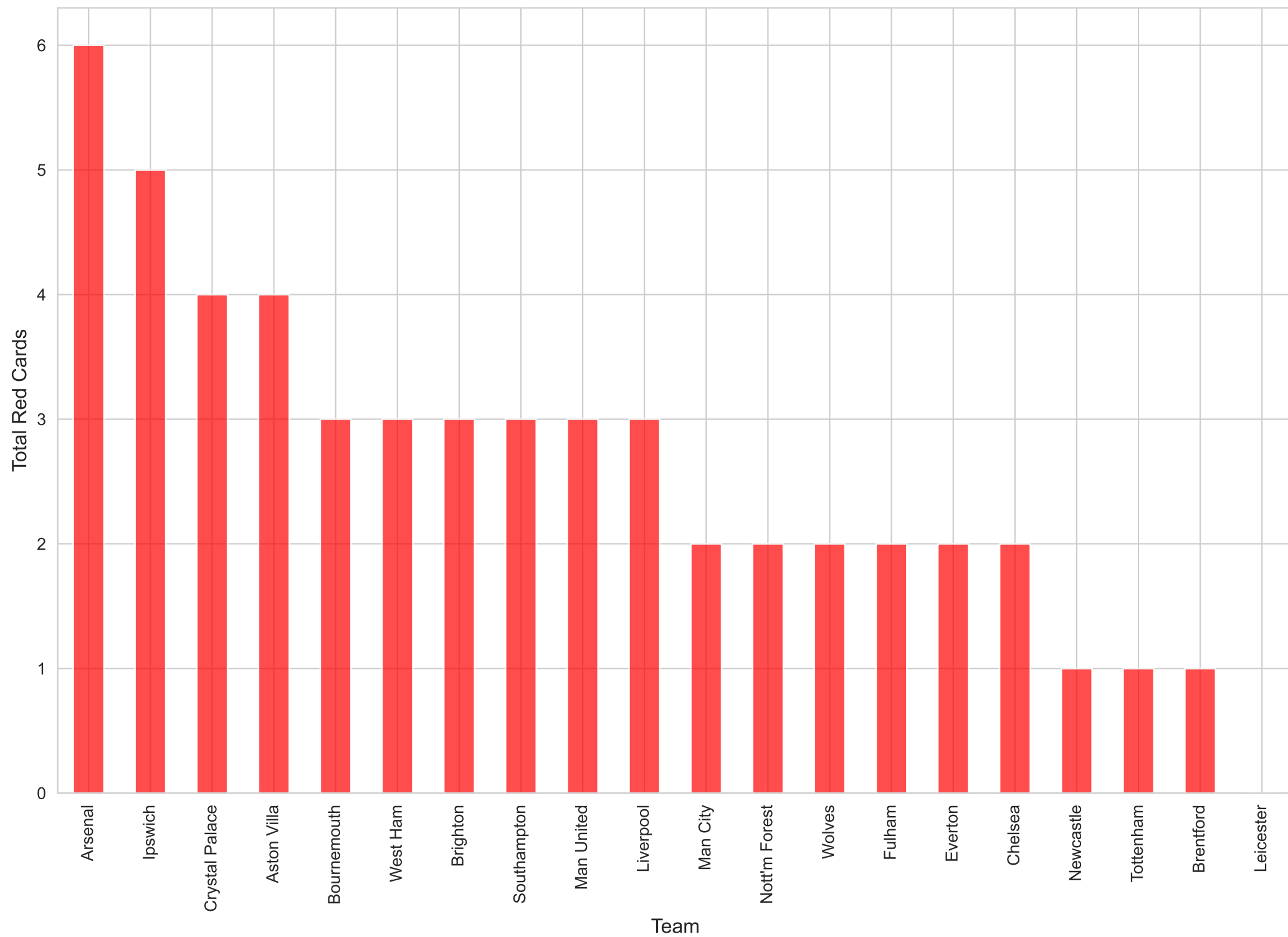
5.1 Match Statistics Correlation Heatmap
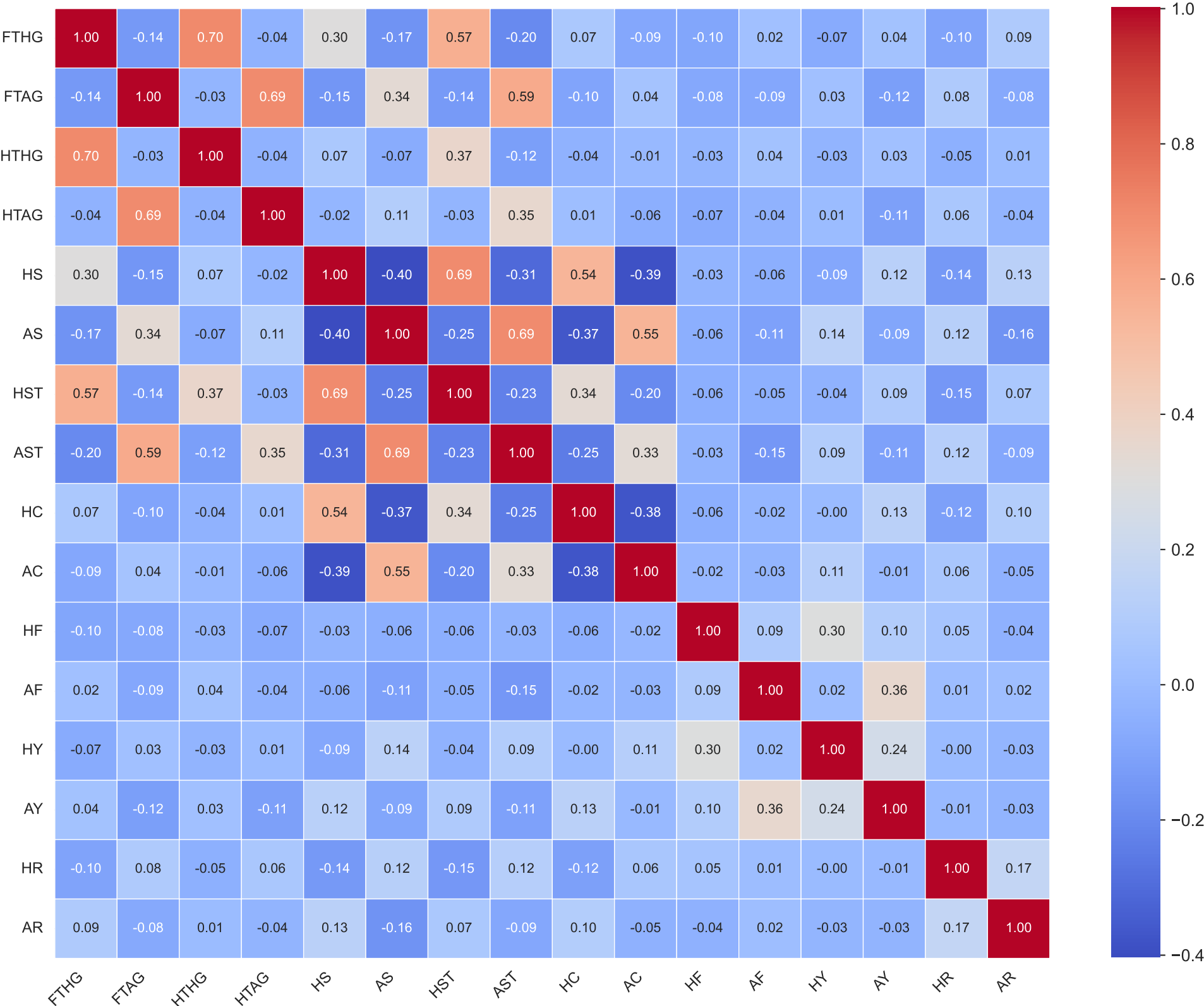
5.2 Shots on Target vs. Goals