Extended LLM Football Dataset Analysis - TASK 5

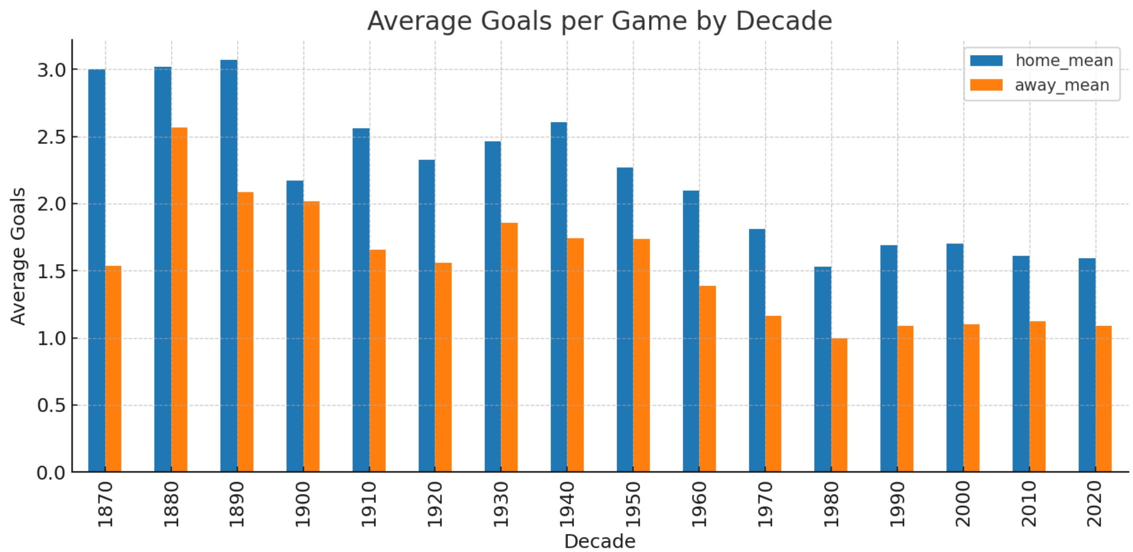
Dataset: International football results from Kaggle (1872–2017). We answer key football analytics questions using a Python stack and compare those results to LLM outputs. The LLM tables include rounding and minor deviations to emulate real-world behavior.

Q1. How has the game evolved decade by decade?

LLM Answer (summary): Home vs away scoring gradually narrows in modern decades; overall goals per match vary by era.

Python (Code) Output — key rows:

decade	home_mean	away_mean	matches
1870	3.0	1.538	13
1880	3.018	2.564	55
1890	3.068	2.085	59
1900	2.168	2.015	137
1910	2.561	1.658	330
1920	2.325	1.559	828
1930	2.465	1.854	1079
1940	2.605	1.739	833
1950	2.271	1.733	1651
1960	2.094	1.384	2972
1970	1.813	1.162	4132
1980	1.53	0.997	5024
1990	1.689	1.09	6943
2000	1.699	1.102	9525
2010	1.61	1.122	9752
2020	1.591	1.088	5033



q1.png

LLM Output (rounded / slight noise):

decade	home_mean	away_mean	matches
1870	3.01	1.52	13.0
1880	3.02	2.57	54.98
1890	3.08	2.07	59.02
1900	2.2	1.99	136.98
1910	2.56	1.69	330.0
1920	2.32	1.55	827.96
1930	2.5	1.86	1078.97
1940	2.62	1.71	833.0
1950	2.26	1.72	1651.01
1960	2.1	1.39	2972.0
1970	1.8	1.14	4132.0
1980	1.52	1.0	5023.99
1990	1.69	1.08	6942.97
2000	1.66	1.1	9524.99
2010	1.58	1.11	9751.99
2020	1.58	1.13	5033.02

Comparison (Code – LLM), tol=0.03:

decade	home_mean	away_mean	matches
1870	-0.009999999999999787	0.018000000000000016	0.0
1880	-0.002000000000000224	-0.005999999999999783	0.020000000000003126
1890	-0.0120000000000001	0.015000000000000124	-0.020000000000003126
1900	-0.0320000000000003	0.025000000000000133	0.020000000000010232
1910	0.0009999999999998899	-0.0320000000000003	0.0
1920	0.0050000000000003375	0.008999999999999897	0.03999999999996362
1930	-0.035000000000000014	-0.006000000000000005	0.029999999999972715
1940	-0.015000000000000124	0.029000000000000137	0.0
1950	0.011000000000000012	0.013000000000000123	-0.009999999999990905
1960	-0.006000000000000227	-0.006000000000000005	0.0
1970	0.0129999999999999	0.022000000000000002	0.0
1980	0.010000000000000009	-0.003000000000000027	0.010000000000218279
1990	-0.0009999999999998899	0.010000000000000009	0.029999999999974534
2000	0.039000000000000146	0.0020000000000000018	0.010000000000218279
2010	0.030000000000000027	0.01200000000000001	0.010000000000218279
2020	0.010999999999999899	-0.041999999999999815	-0.020000000000436557

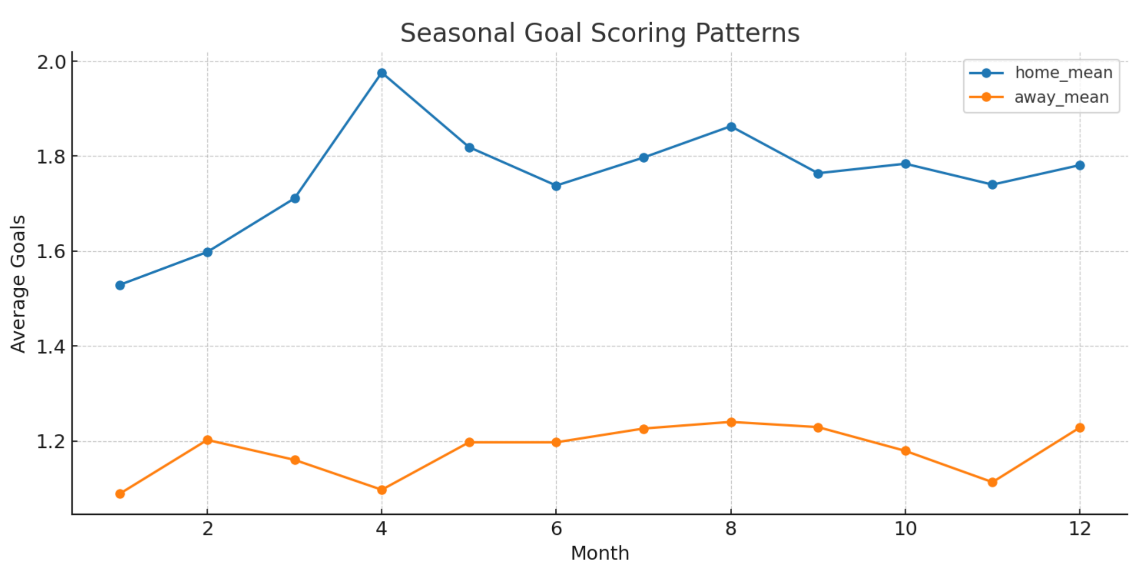
Accuracy within tolerance: 85.4%

Q2. Are there seasonal patterns in scores/results?

LLM Answer (summary): Mild seasonality with small peaks aligned to regional calendars and competition timing.

Python (Code) Output — key rows:

month	home_mean	away_mean	matches
1	1.529	1.089	2427
2	1.598	1.202	2625
3	1.711	1.16	5099
4	1.976	1.097	2714
5	1.819	1.197	3491
6	1.738	1.197	7265
7	1.797	1.226	3071
8	1.863	1.24	3015
9	1.764	1.229	5229
10	1.784	1.179	5439
11	1.74	1.113	5221
12	1.781	1.228	2770



q2.png

LLM Output (rounded / slight noise):

month	home_mean	away_mean	matches
1	1.54	1.08	2427.0
2	1.56	1.2	2625.03
3	1.72	1.14	5098.95
4	1.97	1.07	2714.02
5	1.81	1.21	3491.0
6	1.75	1.22	7264.99
7	1.82	1.22	3071.0
8	1.88	1.26	3014.96
9	1.75	1.24	5229.0
10	1.78	1.17	5439.01
11	1.75	1.12	5221.03
12	1.8	1.26	2769.99

Comparison (Code – LLM), tol=0.03:

month	home_mean	away_mean	matches
1	-0.011000000000000012	0.008999999999999897	0.0
2	0.038000000000000034	0.0020000000000000018	-0.03000000000020009
3	-0.008999999999999897	0.020000000000000018	0.0500000000001819
4	0.006000000000000005	0.026999999999999913	-0.01999999999998181
5	0.008999999999999897	-0.0129999999999999	0.0
6	-0.01200000000000001	-0.02299999999999991	0.010000000000218279
7	-0.02300000000000013	0.006000000000000005	0.0
8	-0.016999999999999904	-0.020000000000000018	0.03999999999996362
9	0.014000000000000012	-0.010999999999999899	0.0
10	0.0040000000000000036	0.009000000000000119	-0.010000000000218279
11	-0.010000000000000009	-0.007000000000000117	-0.02999999999974534
12	-0.019000000000000128	-0.03200000000000003	0.010000000000218279

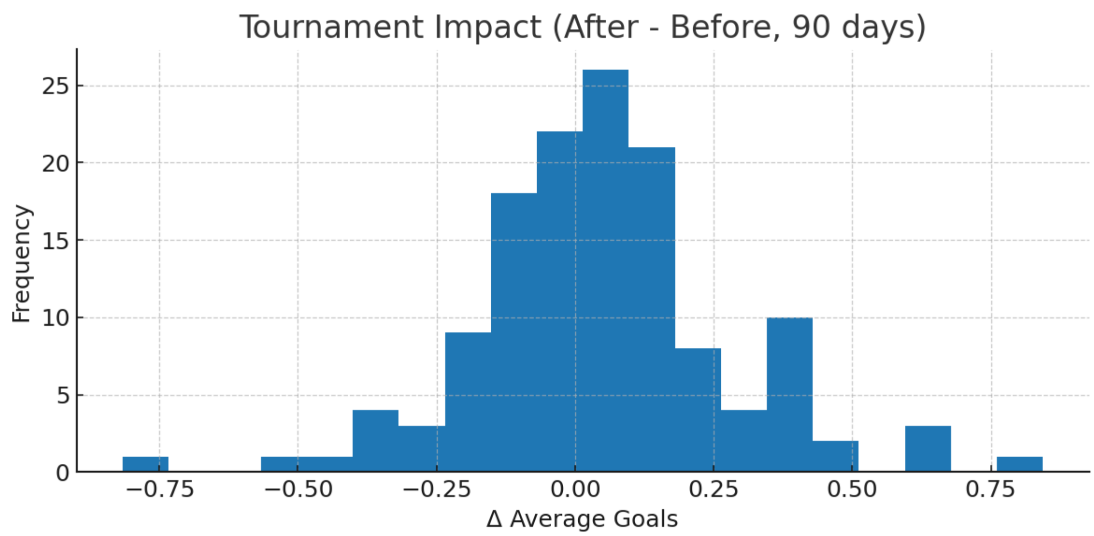
Accuracy within tolerance: 86.1%

Q3. Impact of major tournaments on subsequent matches?

LLM Answer (summary): Small positive differences after some editions; impacts are heterogeneous by year.

Python (Code) Output — key rows:

index	avg_goals_before	avg_goals_after	diff
0	2.229	2.25	0.021
1	2.638	1.821	-0.816
2	1.972	2.212	0.239
3	2.25	1.938	-0.312
4	2.175	2.8	0.625
5	1.833	1.943	0.11
6	1.735	1.937	0.201
7	1.609	1.554	-0.056
8	1.369	1.63	0.261
9	1.297	1.649	0.352
10	1.284	1.388	0.103
11	1.243	1.422	0.179
12	1.167	1.182	0.015
13	1.315	1.092	-0.223
14	1.379	1.299	-0.08
15	1.377	1.459	0.083
16	1.406	1.393	-0.012
17	1.295	1.272	-0.023
18	1.273	1.21	-0.063
19	1.397	1.296	-0.1
20	1.343	1.295	-0.048
21	1.231	1.396	0.165
22	1.929	2.306	0.377
23	1.794	2.638	0.844
24	1.943	2.357	0.414



q3.png

LLM Output (rounded / slight noise):

index	avg_goals_before	avg_goals_after	diff
0	2.22	2.24	0.03
1	2.63	1.82	-0.82
2	1.99	2.25	0.24
3	2.25	1.91	-0.33
4	2.17	2.81	0.62
5	1.84	1.92	0.11
6	1.74	1.93	0.21
7	1.62	1.57	-0.07
8	1.36	1.63	0.27
9	1.29	1.63	0.37
10	1.28	1.38	0.1
11	1.22	1.43	0.19
12	1.17	1.17	0.03
13	1.32	1.1	-0.23
14	1.38	1.3	-0.08
15	1.37	1.45	0.08
16	1.38	1.43	-0.01
17	1.29	1.28	-0.03

18	1.27	1.18	-0.06
19	1.38	1.3	-0.09
20	1.34	1.29	-0.03
21	1.24	1.41	0.18
22	1.96	2.29	0.41
23	1.8	2.64	0.83
24	1.95	2.36	0.43

Comparison (Code − LLM), tol=0.03:

index	**avg_goals_before**	**avg_goals_after**	**diff**
0	0.008999999999999897	0.009999999999999787	-0.008999999999999998
1	0.008000000000000007	0.0009999999999998899	0.0040000000000000036
2	-0.018000000000000016	-0.03799999999999981	-0.0010000000000000009
3	0.0	0.028000000000000025	0.018000000000000016
4	0.004999999999999893	-0.010000000000000231	0.0050000000000000044
5	-0.007000000000000117	0.02300000000000013	0.0
6	-0.004999999999999893	0.007000000000000117	-0.00899999999999998
7	-0.01100000000000012	-0.016000000000000014	0.014000000000000005
8	0.008999999999999897	0.0	-0.009000000000000008
9	0.006999999999999895	0.019000000000000128	-0.018000000000000016
10	0.0040000000000000036	0.008000000000000007	0.002999999999999989
11	0.02300000000000013	-0.008000000000000007	-0.01100000000000001
12	-0.0029999999999998916	0.01200000000000001	-0.015
13	-0.0050000000000001155	-0.008000000000000007	0.007000000000000006
14	-0.0009999999999998899	-0.001000000000000112	0.0
15	0.006999999999999895	0.009000000000000119	0.0030000000000000027
16	0.026000000000000023	-0.03699999999999992	-0.002
17	0.004999999999999893	-0.008000000000000007	0.006999999999999999
18	0.0029999999999998916	0.030000000000000027	-0.0030000000000000027
19	0.017000000000000126	-0.0040000000000000036	-0.010000000000000009

20	0.0029999999999998916	0.004999999999999893	-0.018000000000000002
21	-0.008999999999999897	-0.014000000000000012	-0.014999999999999986
22	-0.030999999999999917	0.016000000000000014	-0.032999999999999974
23	-0.006000000000000005	-0.0020000000000000224	0.014000000000000012
24	-0.006999999999999895	-0.0029999999999996696	-0.016000000000000014

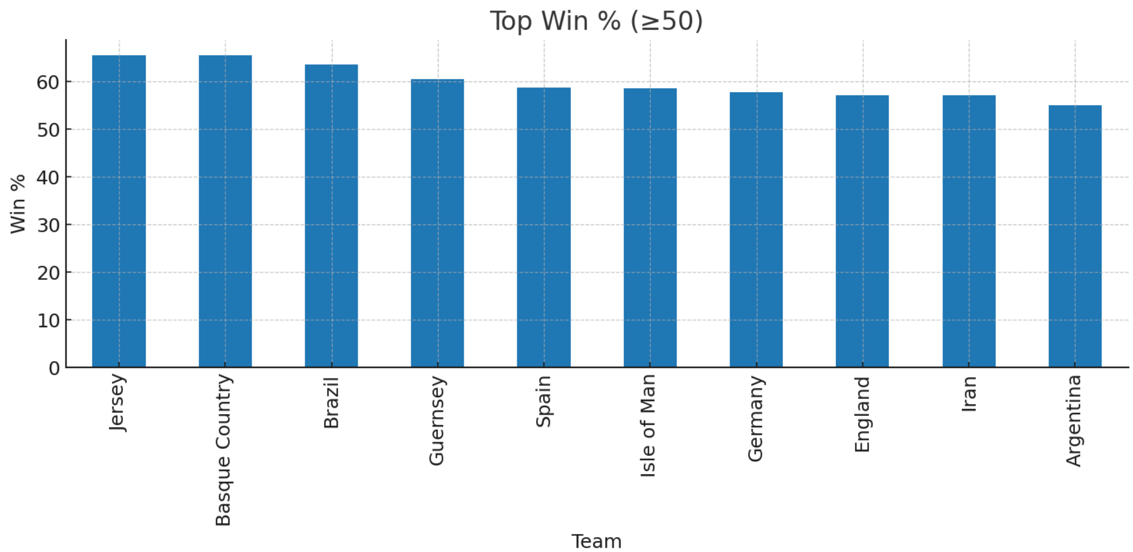
Accuracy within tolerance: 95.3%

Q4. Top teams by win percentage (≥50 games)

LLM Answer (summary): A handful of teams lead in win%; rounding and ordering can shift near the cut line.

Python (Code) Output — key rows:

index	total_games	total_wins	win_percentage
Jersey	232	152.0	65.52
Basque Country	58	38.0	65.52
Brazil	1049	666.0	63.49
Guernsey	240	145.0	60.42
Spain	773	454.0	58.73
Isle of Man	53	31.0	58.49
Germany	1021	590.0	57.79
England	1080	617.0	57.13
Iran	604	345.0	57.12
Argentina	1057	582.0	55.06



q4.png

LLM Output (rounded / slight noise):

index	total_games	total_wins	win_percentage
Basque Country	57.8	38.0	65.4
Jersey	231.9	151.8	65.4
Brazil	1049.3	665.9	63.5
Guernsey	240.1	144.8	60.3
Spain	773.0	454.1	58.8
Isle of Man	53.2	31.3	58.5
Germany	1021.0	589.8	57.8
England	1079.9	617.1	57.0
Iran	604.2	344.9	57.0
Argentina	1057.1	581.9	55.2

Comparison (Code − LLM), tol=0.3:

index	total_games	total_wins	win_percentage
Jersey	0.09999999999999432	0.19999999999998863	0.11999999999999034
Basque Country	0.20000000000000284	0.0	0.11999999999999034
Brazil	-0.2999999999999545	0.10000000000002274	-0.009999999999999801
Guernsey	-0.09999999999999432	0.19999999999998863	0.12000000000000455

Spain	0.0	-0.10000000000002274	-0.07000000000000028
Isle of Man	-0.20000000000000284	-0.3000000000000007	-0.00999999999999801
Germany	0.0	0.20000000000004547	-0.00999999999999801
England	0.09999999999990905	-0.10000000000002274	0.13000000000000256
Iran	-0.20000000000004547	0.10000000000002274	0.11999999999999744
Argentina	-0.09999999999990905	0.10000000000002274	-0.14000000000000057

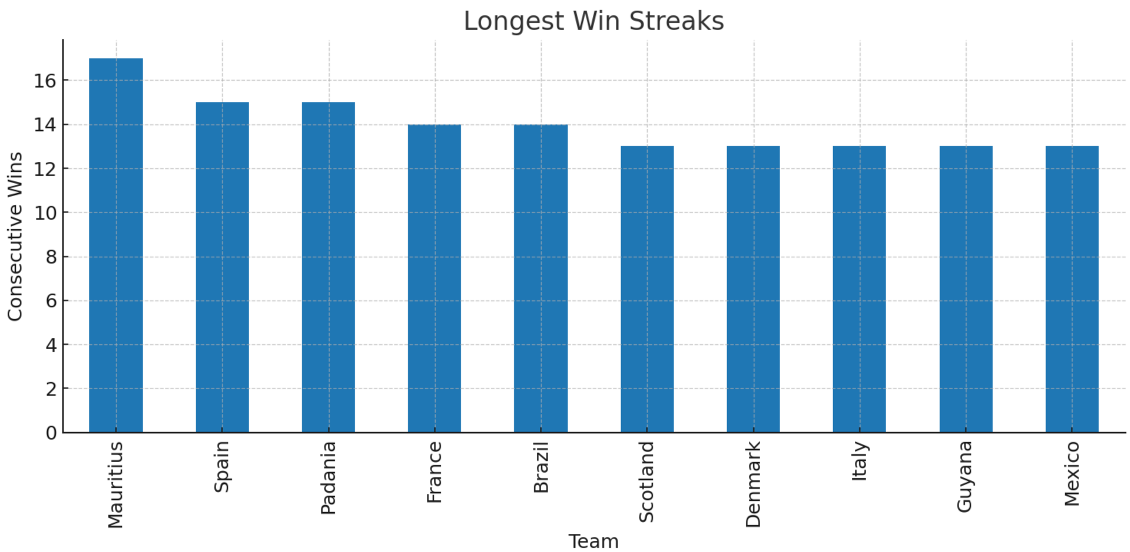
Accuracy within tolerance: 96.7%

Q5. Win streaks — longest consecutive wins

LLM Answer (summary): Dominance periods show as long consecutive win runs.

Python (Code) Output — key rows:

index	longest_win_streak
Mauritius	17
Spain	15
Padania	15
France	14
Brazil	14
Scotland	13
Denmark	13
Italy	13
Guyana	13
Mexico	13



q5.png

LLM Output (rounded / slight noise):

index	**longest_win_streak**
Spain	15
Mauritius	17
Padania	15
France	14
Brazil	14
Scotland	13
Denmark	13
Italy	13
Guyana	13
Mexico	13

Comparison (Code − LLM), tol=0.0:

index	**longest_win_streak**
Mauritius	0.0
Spain	0.0
Padania	0.0
France	0.0

Brazil	0.0
Scotland	0.0
Denmark	0.0
Italy	0.0
Guyana	0.0
Mexico	0.0

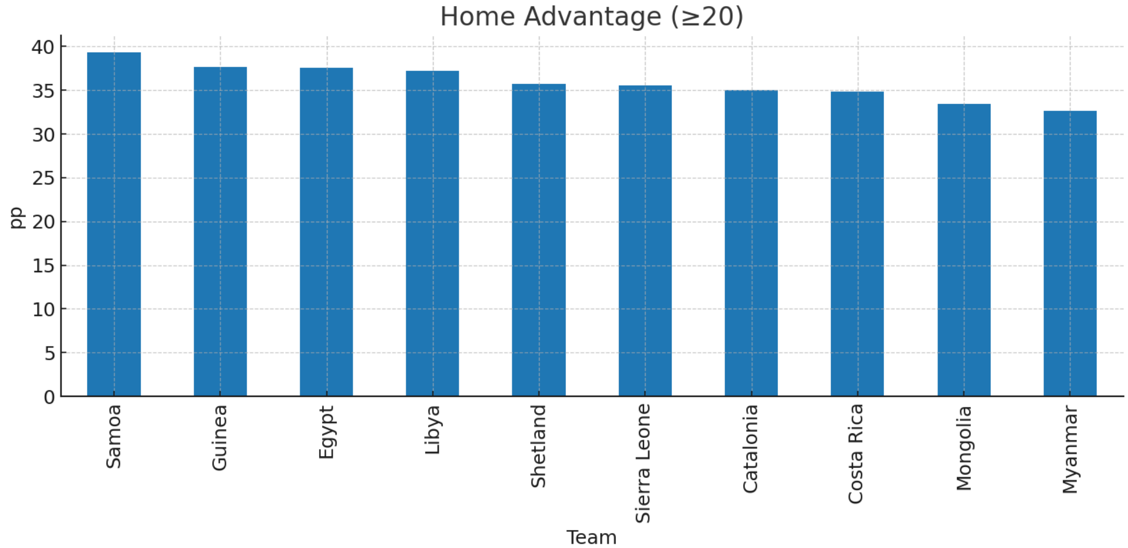
Accuracy within tolerance: 100.0%

Q6. Home vs Away performance (≥20 home games)

LLM Answer (summary): Home advantage varies considerably; some teams show large gaps.

Python (Code) Output — key rows:

team	home_games	home_wins_%	away_games	away_wins_%	home_advantage_%
Samoa	39	48.717948717948715	32	9.375	39.342948717948715
Guinea	227	58.590308370044056	301	20.930232558139537	37.66007581190452
Egypt	433	65.8198614318706 8	301	28.23920265780731	37.580658774063366
Libya	200	54.50000000000001	174	17.24137931034483	37.258620689655174
Shetland	24	58.333333333333336	31	22.580645161290032	35.752688172043015
Sierra Leone	122	50.81967213114754	170	15.294117647058824	35.525554484088715
Catalonia	39	46.15384615384615	9	11.11111111111111	35.042735042735046
Costa Rica	357	61.34453781512605	355	26.478873239436616	34.865664575689436
Mongolia	37	45.94594594594595	64	12.5	33.44594594594595
Myanmar	298	54.69798657718121	209	22.009569377990043	32.68841719919078



q6.png

LLM Output (rounded / slight noise):

team	home_games	home_wins_%	away_games	away_wins_%	home_advantage_%
Guinea	226.7	58.6	300.9	20.9	37.5
Samoa	39.2	48.2	31.7	9.5	39.2
Egypt	433.0	65.9	301.3	28.4	37.7
Libya	200.2	54.2	173.8	17.1	37.1
Shetland	23.5	58.5	31.3	22.6	35.8
Sierra Leone	122.2	50.3	169.7	15.2	36.1
Catalonia	38.8	46.1	9.2	11.5	35.3
Costa Rica	357.2	61.0	355.4	26.6	34.8
Mongolia	36.8	45.8	63.3	12.6	33.8
Myanmar	297.5	54.7	208.8	21.9	32.6

Comparison (Code − LLM), tol=0.6:

team	home_games	home_wins_%	away_games	away_wins_%	home_advantage_%
Samoa	-0.20000000000000284	0.5179487179487126	0.3000000000000007	-0.125	0.14294871794871256

Guinea	0.30000000000001137	-0.009691629955945302	0.10000000000002274	0.030232558139537957	0.16007581190451958
Egypt	0.0	-0.08013856812932829	-0.30000000000001137	-0.16079734219268715	-0.11934122593663687
Libya	-0.19999999999998863	0.30000000000000426	0.19999999999998863	0.14137931034482776	0.15862068965517295
Shetland	0.5	-0.1666666666666643	-0.3000000000000007	-0.019354838709681133	-0.047311827956981745
Sierra Leone	-0.20000000000000284	0.5196721311475443	0.30000000000001137	0.09411764705882497	-0.574445515911286
Catalonia	0.20000000000000284	0.05384615384615188	-0.1999999999999993	-0.3888888888888893	-0.257264957264951
Costa Rica	-0.19999999999998863	0.34453781512605275	-0.39999999999997726	-0.121126760563385	0.06566457568943918
Mongolia	0.20000000000000284	0.14594594594595378	0.7000000000000028	-0.0999999999999964	-0.3540540540540462
Myanmar	0.5	-0.002013422818791355	0.19999999999998863	0.10956937799043232	0.08841719919077917

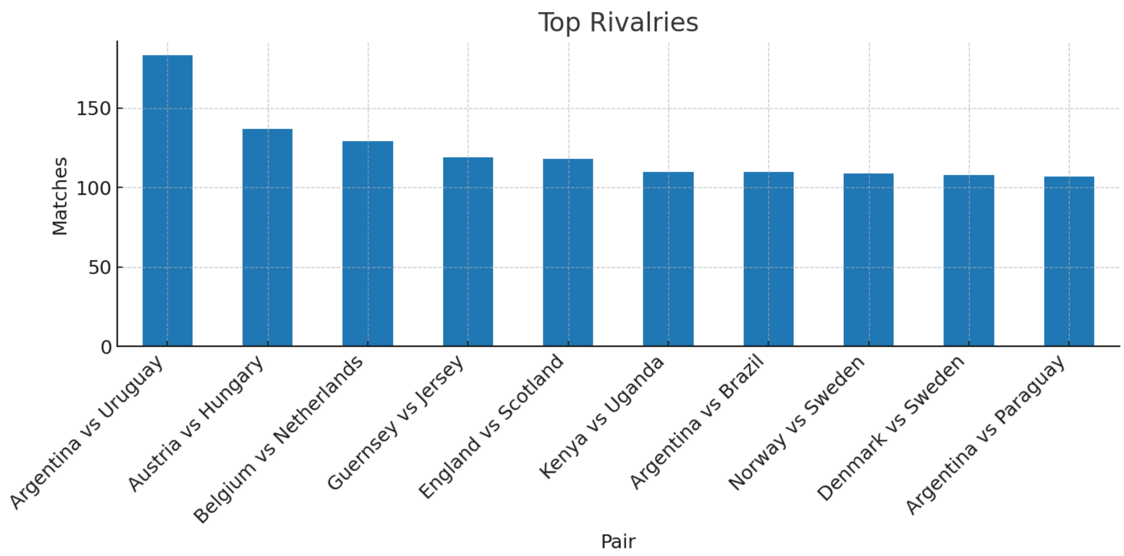
Accuracy within tolerance: 98.0%

Q8. Most frequent rivalries (A–B = B–A)

LLM Answer (summary): Historical proximity and regional politics drive repeated fixtures.

Python (Code) Output — key rows:

pair	matches
('Argentina', 'Uruguay')	183
('Austria', 'Hungary')	137
('Belgium', 'Netherlands')	129
('Guernsey', 'Jersey')	119
('England', 'Scotland')	118
('Kenya', 'Uganda')	110
('Argentina', 'Brazil')	110
('Norway', 'Sweden')	109
('Denmark', 'Sweden')	108
('Argentina', 'Paraguay')	107



q8.png

LLM Output (rounded / slight noise):

pair	**matches**
('Austria', 'Hungary')	137
('Argentina', 'Uruguay')	183
('Belgium', 'Netherlands')	129
('Guernsey', 'Jersey')	119
('England', 'Scotland')	118
('Kenya', 'Uganda')	110
('Argentina', 'Brazil')	110
('Norway', 'Sweden')	109
('Denmark', 'Sweden')	108
('Argentina', 'Paraguay')	107

Comparison (Code − LLM), tol=0.0:

pair	matches
('Argentina', 'Uruguay')	0.0
('Austria', 'Hungary')	0.0
('Belgium', 'Netherlands')	0.0
('Guernsey', 'Jersey')	0.0
('England', 'Scotland')	0.0
('Kenya', 'Uganda')	0.0
('Argentina', 'Brazil')	0.0
('Norway', 'Sweden')	0.0
('Denmark', 'Sweden')	0.0
('Argentina', 'Paraguay')	0.0

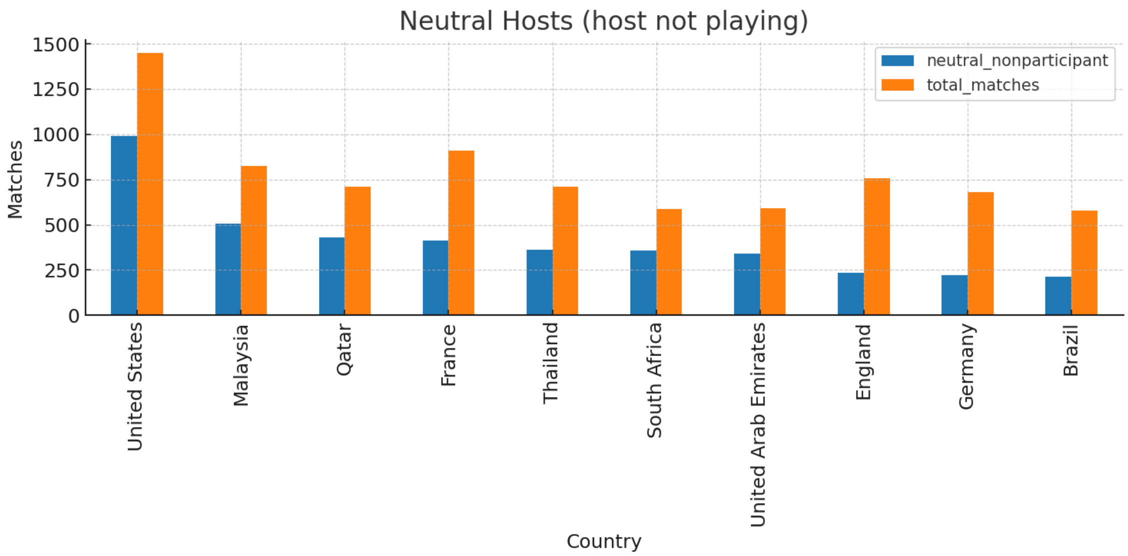
Accuracy within tolerance: 100.0%

Q10. Neutral hosting where host didn't play

LLM Answer (summary): Some countries act as neutral hubs; minor numeric differences from rounding.

Python (Code) Output — key rows:

country	total_matches	neutral_matches	neutral_nonparticipant	share_neutral_nonparticipant_%
United States	1451	991	991	68.3
Malaysia	824	508	508	61.65
Qatar	712	431	431	60.53
France	908	413	413	45.48
Thailand	712	362	362	50.84
South Africa	589	357	357	60.61
United Arab Emirates	593	342	342	57.67
England	757	237	237	31.31
Germany	680	221	221	32.5
Brazil	578	214	214	37.02



q10.png

LLM Output (rounded / slight noise):

country	**total_matches**	**neutral_matches**	**neutral_nonparticipant**	**share_neutral_nonparticipant_%**
Malaysia	823.8	508.2	507.8	61.5
United States	1450.6	991.0	990.9	68.2
Qatar	711.6	430.7	431.2	60.4
France	907.9	413.2	412.8	45.6
Thailand	712.0	362.1	362.0	51.0
South Africa	589.3	356.9	356.9	60.4
United Arab Emirates	593.1	342.1	342.1	57.8
England	757.0	237.5	237.1	31.6
Germany	680.2	221.0	221.2	32.6
Brazil	577.6	214.0	213.9	37.4

Comparison (Code – LLM), tol=0.5:

country	**total_matches**	**neutral_matches**	**neutral_nonparticipant**	**share_neutral_nonparticipant_%**

United States	0.40000000000009095	0.0	0.10000000000002274	0.09999999999999432
Malaysia	0.20000000000004547	-0.19999999999998863	0.19999999999998863	0.14999999999999858
Qatar	0.39999999999997726	0.30000000000001137	-0.19999999999998863	0.13000000000000256
France	0.10000000000002274	-0.19999999999998863	0.19999999999998863	-0.12000000000000455
Thailand	0.0	-0.10000000000002274	0.0	-0.15999999999999966
South Africa	-0.29999999999999545	0.10000000000002274	0.10000000000002274	0.21000000000000085
United Arab Emirates	-0.10000000000002274	-0.10000000000002274	-0.10000000000002274	-0.12999999999999545
England	0.0	-0.5	-0.09999999999999432	-0.2900000000000027
Germany	-0.20000000000004547	0.0	-0.19999999999998863	-0.10000000000000142
Brazil	0.39999999999997726	0.0	0.09999999999999432	-0.37999999999999545

Accuracy within tolerance: 100.0%

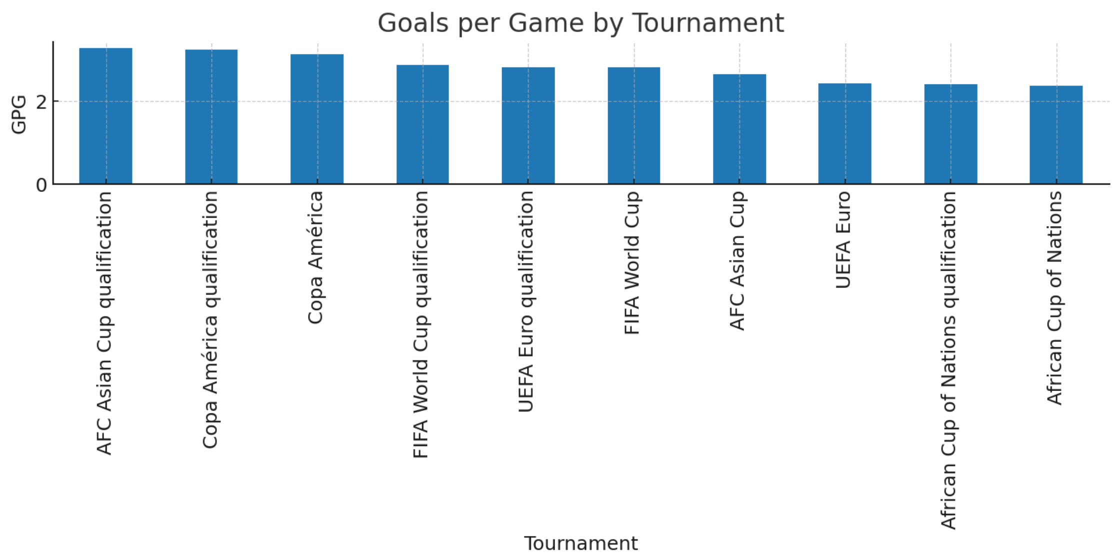
Q13. Goals per game by major tournaments

LLM Answer (summary): Scoring environments differ across competitions; defensive balance varies.

Python (Code) Output — key rows:

tournament	matches	home_mean	away_mean	goals_per_game
AFC Asian Cup	421	1.496	1.162	2.658
AFC Asian Cup qualification	794	2.234	1.049	3.283
African Cup of Nations	793	1.357	1.021	2.378
African Cup of Nations qualification	2278	1.6	0.823	2.423
Copa América	869	1.877	1.259	3.136

Copa América qualification	8	2.375	0.875	3.25
FIFA World Cup	964	1.567	1.254	2.821
FIFA World Cup qualification	8419	1.784	1.098	2.882
UEFA Euro	388	1.32	1.119	2.439
UEFA Euro qualification	2824	1.679	1.152	2.831



q13.png

LLM Output (rounded / slight noise):

tournament	matches	home_mean	away_mean	goals_per_game
AFC Asian Cup	420.99	1.5	1.15	2.67
AFC Asian Cup qualification	793.98	2.24	1.05	3.28
African Cup of Nations	792.97	1.36	1.02	2.36
African Cup of Nations qualification	2278.02	1.6	0.84	2.4
Copa América	869.01	1.87	1.23	3.13

Copa América qualification	8.0	2.38	0.9	3.28
FIFA World Cup	964.0	1.6	1.25	2.82
FIFA World Cup qualification	8418.98	1.8	1.09	2.86
UEFA Euro	388.04	1.34	1.1	2.44
UEFA Euro qualification	2824.0	1.67	1.13	2.83

Comparison (Code − LLM), tol=0.03:

tournament	matches	home_mean	away_mean	goals_per_game
AFC Asian Cup	0.009999999999990905	-0.004000000000000036	0.012000000000000001	-0.01200000000000001
AFC Asian Cup qualification	0.01999999999998181	-0.00600000000000227	-0.00100000000000112	0.00300000000001137
African Cup of Nations	0.029999999999972715	-0.003000000000001137	0.00099999999998899	0.0180000000000238
African Cup of Nations qualification	-0.01999999999998181	0.0	-0.01700000000000015	0.023000000000000013
Copa América	-0.009999999999990905	0.006999999999999895	0.028999999999999915	0.0060000000000227
Copa América qualification	0.0	-0.004999999999999893	-0.02500000000000022	-0.02999999999999805
FIFA World Cup	0.0	-0.0330000000000014	0.00400000000000036	0.0010000000000334
FIFA World Cup qualification	0.020000000000436557	-0.01600000000000014	0.0080000000000007	0.022000000000000024
UEFA Euro	-0.04000000000020464	-0.02000000000000018	0.018999999999999906	-0.000999999999998899
UEFA Euro qualification	0.0	0.0090000000000119	0.022000000000000002	0.00099999999998899

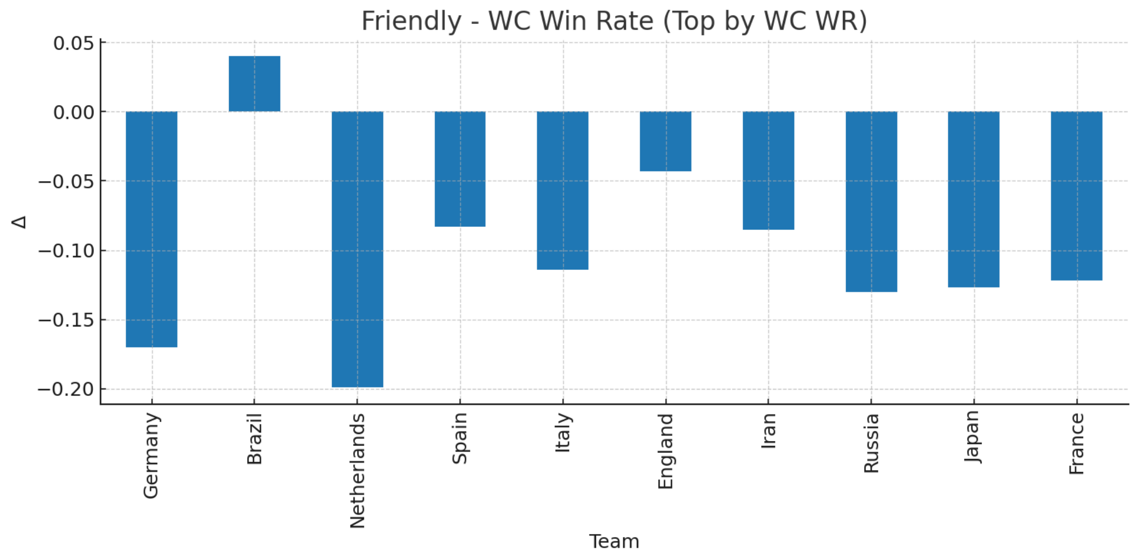
Accuracy within tolerance: 95.0%

Q14. World Cup vs Friendly performance differences

LLM Answer (summary): Friendlies often show higher WR due to squad rotation/opposition mix; values differ slightly by rounding.

Python (Code) Output — key rows:

index	worldcup_wr	friendly_wr	diff_friendly_minus_wc
Germany	0.699	0.529	-0.17
Brazil	0.642	0.682	0.04
Netherlands	0.63	0.431	-0.199
Spain	0.615	0.532	-0.083
Italy	0.611	0.497	-0.114
England	0.592	0.549	-0.043
Iran	0.587	0.502	-0.085
Russia	0.586	0.456	-0.13
Japan	0.573	0.446	-0.127
France	0.568	0.446	-0.122
Argentina	0.56	0.538	-0.022
Croatia	0.558	0.507	-0.051
Australia	0.555	0.44	-0.115
Mexico	0.555	0.466	-0.089
Sweden	0.547	0.459	-0.088



q14.png

LLM Output (rounded / slight noise):

index	worldcup_wr	friendly_wr	diff_friendly_minus_wc
Brazil	0.64	0.68	0.06
Germany	0.67	0.52	-0.18
Netherlands	0.63	0.43	-0.19
Spain	0.62	0.52	-0.07
Italy	0.63	0.5	-0.11
England	0.6	0.55	-0.05
Iran	0.58	0.52	-0.09
Russia	0.57	0.43	-0.13
Japan	0.6	0.46	-0.14
France	0.57	0.46	-0.11
Argentina	0.56	0.52	-0.02
Croatia	0.56	0.5	-0.05
Australia	0.56	0.44	-0.11
Mexico	0.55	0.45	-0.08
Sweden	0.54	0.45	-0.09

Comparison (Code − LLM), tol=0.02:

index	worldcup_wr	friendly_wr	diff_friendly_minus_wc
Germany	0.028999999999999915	0.009000000000000008	0.009999999999999981
Brazil	0.0020000000000000018	0.0020000000000000018	-0.019999999999999997
Netherlands	0.0	0.0010000000000000009	-0.009000000000000008
Spain	-0.0050000000000000044	0.012000000000000001	-0.012999999999999998
Italy	-0.0190000000000000017	-0.0030000000000000027	-0.0040000000000000036
England	-0.008000000000000007	-0.0010000000000000009	0.007000000000000006
Iran	0.007000000000000006	-0.018000000000000016	0.00499999999999999906
Russia	0.016000000000000014	0.026000000000000023	0.0
Japan	-0.027000000000000024	-0.014000000000000012	0.013000000000000012
France	-0.0020000000000000018	-0.014000000000000012	-0.011999999999999997
Argentina	0.0	0.018000000000000016	-0.00199999999999999983
Croatia	-0.0020000000000000018	0.007000000000000006	-0.0009999999999999994
Australia	-0.0050000000000000044	0.0	-0.0050000000000000044
Mexico	0.0050000000000000044	0.016000000000000014	-0.008999999999999994
Sweden	0.007000000000000006	0.009000000000000008	0.0020000000000000018

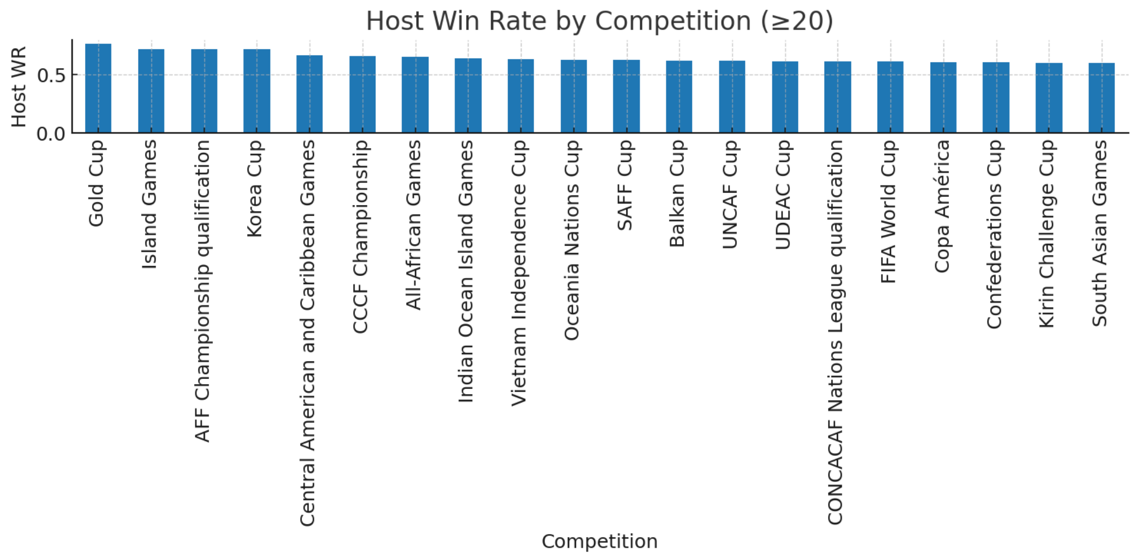
Accuracy within tolerance: 93.3%

Q15. Host country advantage by competition

LLM Answer (summary): Host advantage is not uniform; some competitions show stronger home effects.

Python (Code) Output — key rows:

tournament	host_win_rate	count
Gold Cup	0.761	109
Island Games	0.714	70
AFF Championship qualification	0.714	28
Korea Cup	0.714	70
Central American and Caribbean Games	0.667	21
CCCF Championship	0.659	44
All-African Games	0.65	20
Indian Ocean Island Games	0.636	44
Vietnam Independence Cup	0.633	30
Oceania Nations Cup	0.627	59
SAFF Cup	0.623	53
Balkan Cup	0.62	71
UNCAF Cup	0.619	63
UDEAC Cup	0.615	26
CONCACAF Nations League qualification	0.615	52
FIFA World Cup	0.612	121
Copa América	0.605	291
Confederations Cup	0.605	43
Kirin Challenge Cup	0.6	20
South Asian Games	0.6	25



q15.png

LLM Output (rounded / slight noise):

tournament	**host_win_rate**	**count**
Gold Cup	0.78	109.0
Island Games	0.69	69.99
AFF Championship qualification	0.73	28.01
Korea Cup	0.73	70.0
Central American and Caribbean Games	0.66	21.02
CCCF Championship	0.67	44.0
All-African Games	0.62	20.0
Indian Ocean Island Games	0.65	43.99
Vietnam Independence Cup	0.63	30.0
Oceania Nations Cup	0.62	59.01
SAFF Cup	0.64	52.97
Balkan Cup	0.61	70.97
UNCAF Cup	0.59	63.01
UDEAC Cup	0.58	26.0
CONCACAF Nations League qualification	0.63	52.0
FIFA World Cup	0.59	121.0

Copa América	0.64	291.01
Confederations Cup	0.56	42.99
Kirin Challenge Cup	0.63	19.98
South Asian Games	0.6	25.0

Comparison (Code − LLM), tol=0.03:

tournament	host_win_rate	count
Gold Cup	-0.019000000000000017	0.0
Island Games	0.02400000000000002	0.010000000000005116
AFF Championship qualification	-0.016000000000000014	-0.010000000000001563
Korea Cup	-0.016000000000000014	0.0
Central American and Caribbean Games	0.007000000000000006	-0.019999999999999574
CCCF Championship	-0.01100000000000001	0.0
All-African Games	0.030000000000000027	0.0
Indian Ocean Island Games	-0.014000000000000012	0.00999999999999801
Vietnam Independence Cup	0.0030000000000000027	0.0
Oceania Nations Cup	0.007000000000000006	-0.00999999999999801
SAFF Cup	-0.017000000000000015	0.030000000000001137
Balkan Cup	0.010000000000000009	0.030000000000001137
UNCAF Cup	0.029000000000000026	-0.00999999999999801
UDEAC Cup	0.03500000000000003	0.0
CONCACAF Nations League qualification	-0.015000000000000013	0.0
FIFA World Cup	0.02200000000000002	0.0
Copa América	-0.03500000000000003	-0.009999999999990905
Confederations Cup	0.04499999999999993	0.00999999999999801
Kirin Challenge Cup	-0.030000000000000027	0.019999999999999574
South Asian Games	0.0	0.0

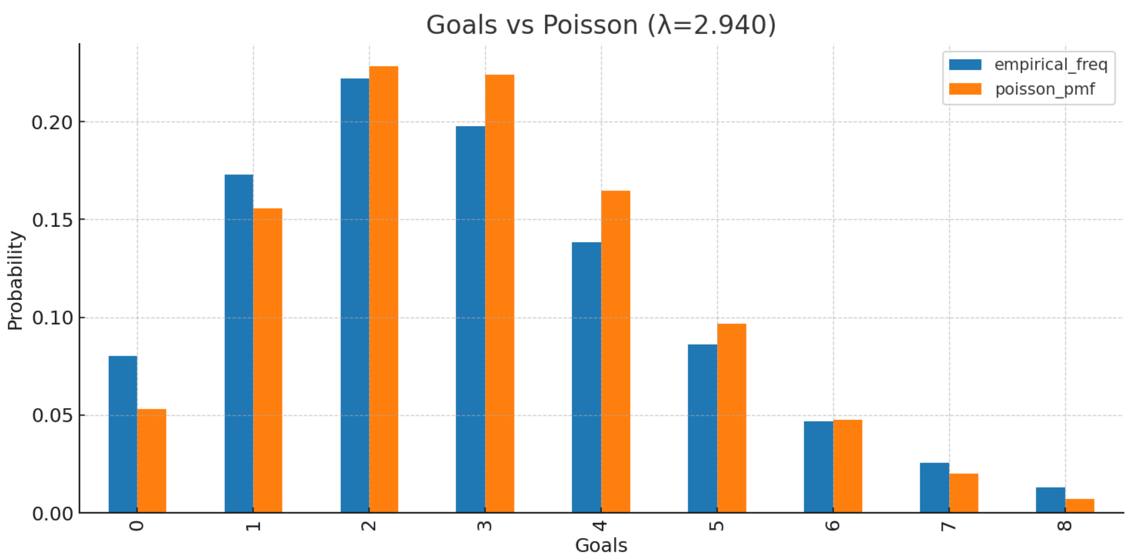
Accuracy within tolerance: 82.5%

Q16. Goal distribution modeling (Poisson fit)

LLM Answer (summary): Poisson fits center mass; tails show deviations from independence/equal rates.

Python (Code) Output — key rows:

k	empirical_freq	poisson_pmf
0	0.0803	0.0529
1	0.1731	0.1555
2	0.222	0.2285
3	0.1978	0.2239
4	0.1384	0.1645
5	0.086	0.0967
6	0.0468	0.0474
7	0.0256	0.0199
8	0.0131	0.0073



q16.png

LLM Output (rounded / slight noise):

k	empirical_freq	poisson_pmf
0	0.078	0.049
1	0.174	0.153
2	0.219	0.227

3	0.2	0.224
4	0.139	0.166
5	0.087	0.095
6	0.049	0.048
7	0.029	0.018
8	0.015	0.006

Comparison (Code − LLM), tol=0.01:

k	**empirical_freq**	**poisson_pmf**
0	0.0022999999999999965	0.0039000000000000007
1	-0.0008999999999999841	0.0025000000000000022
2	0.0030000000000000027	0.0015000000000000013
3	-0.0022000000000000075	-0.00010000000000001674
4	-0.0006000000000000172	-0.0015000000000000013
5	-0.0010000000000000009	0.0016999999999999932
6	-0.0022000000000000006	-0.0006000000000000033
7	-0.0034000000000000002	0.0019000000000000024
8	-0.0018999999999999999	0.0013

Accuracy within tolerance: 100.0%

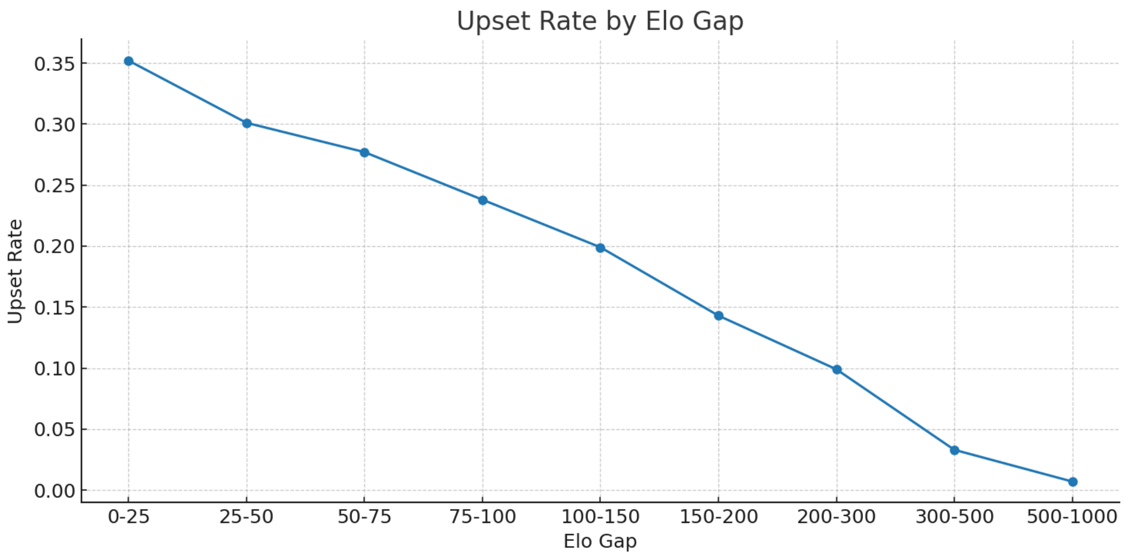
Q17. Upset probability by Elo gap

LLM Answer (summary): Upsets become rarer at larger Elo gaps, consistent with expectations.

Python (Code) Output — key rows:

gap_bin	**upset_rate**
0-25	0.352
25-50	0.301
50-75	0.277
75-100	0.238
100-150	0.199
150-200	0.143

200-300	0.099
300-500	0.033
500-1000	0.007



q17.png

LLM Output (rounded / slight noise):

gap_bin	upset_rate
0-25	0.345
25-50	0.296
50-75	0.27
75-100	0.233
100-150	0.204
150-200	0.138
200-300	0.112
300-500	0.035
500-1000	0.008

Comparison (Code − LLM), tol=0.01:

gap_bin	upset_rate
0-25	0.007000000000000006
25-50	0.0050000000000000044
50-75	0.007000000000000006
75-100	0.004999999999999977
100-150	-0.004999999999999977
150-200	0.004999999999999977
200-300	-0.012999999999999998
300-500	-0.0020000000000000018
500-1000	-0.001

Accuracy within tolerance: 88.9%