1 Model

Based on all matches from the three previous world cups, the number of goals a team score against a specific opponent were modeled. The model evaluates the influence of the following covariates on the number of scored goals:

- **Economic Factors:**
  GDP per capita, population

- **Sportive Factors:**
  ODDSET odds, FIFA rank

- **Home advantage:**
  host of the world cup, same continent as host

- **Factors describing the team's structure**
  Maximum number of teammates, Second maximum number of teammates, average age, number of CL players, number of Europa League players, number of players abroad

- **Factors describing the team's coach**
  age, nationality, tenure

Additionally, for each team specific attacking and defending parameters were included. Therefore, for a match team i vs. team j at world cup k, the following model is applied:

$$y_{ijk}|\boldsymbol{x}_{ik}, \boldsymbol{x}_{jk} \quad \sim \quad Pois(\lambda_{ijk})$$

$$log(\lambda_{ijk}) \quad = \quad (\boldsymbol{x}_{ik} - \boldsymbol{x}_{jk})^T \boldsymbol{\beta} + \mathrm{att}_i - \mathrm{def}_j .$$

- team $i, i = 1, \ldots, n$, n number of teams

- opponent $j, j = 1, \ldots, n$

- tournament $k, k = 1, 2, 3$

- att_i, def_i: attacking and defending parameter of team i

- $\boldsymbol{x}_{ik}, \boldsymbol{x}_{jk}$: covariates of teams i and j, varying over tournaments

- y_{ijk} number of goals scored by team i against team j in tournament k

The model is basically a Poisson model for the number of scored goals, but both the covariates effects and the team specific attacking and defending parameters were regularized by a Lasso penalty.

The resulting model contained all covariates except for the maximum number of teammates, the number of players abroad and the dummy variable for the host of the world cup.

Based on the model, each match (with the exact number of goals for both teams) of the FIFA World Cup 2014 was simulated 10000 times. The following table presents the probilities for each team to become the new world champion:

2 Probabilities for FIFA World Cup 2014 Winner

team	$\hat{p}_{\text{Lasso}}$	$\hat{p}_{\text{Oddset}}$	team	$\hat{p}_{\text{Lasso}}$	$\hat{p}_{\text{Oddset}}$
1. GER	0.2880	0.1420	17. GHA	0.0022	0.0071
2. BRA	0.2765	0.2028	18. KOR	0.0019	0.0024
3. ESP	0.0900	0.1092	19. ALG	0.0018	0.0071
4. BEL	0.0819	0.0592	20. ECU	0.0017	0.0071
5. ARG	0.0582	0.1420	21. USA	0.0016	0.0071
6. POR	0.0522	0.0237	22. MEX	0.0012	0.0071
7. SUI	0.0413	0.0071	23. JPN	0.0010	0.0047
8. CRO	0.0210	0.0071	24. BIH	0.0008	0.0047
9. ENG	0.0193	0.0355	25. GRE	0.0005	0.0071
10. FRA	0.0135	0.0355	26. RUS	0.0004	0.0118
11. NED	0.0129	0.0355	27. NGA	0.0004	0.0035
12. ITA	0.0094	0.0355	28. AUS	0.0003	0.0024
13. URU	0.0071	0.0284	29. HON	0.0002	0.0005
14. CHI	0.0063	0.0203	30. CRC	0	0.0071
15. COL	0.0052	0.0394	31. CMR	0	0.0024
16. CIV	0.0032	0.0071	32. IRN	0	0.0005

3 Most probable tournament outcome

Group A 44%		Group B 24%		Group C 16%		Group D 18%		Group E 22%		Group F 36%		Group G 37%		Group H 26%	
1.	BRA	1.	ESP	1.	COL	1.	ENG	1.	SUI	1.	ARG	1.	GER	1.	BEL
2.	CRO	2.	NED	2.	CIV	2.	ITA	2.	FRA	2.	BIH	2.	POR	2.	KOR
	MEX		CHI		JPN		URU		ECU		NGA		GHA		RUS
	CMR		AUS		GRE		CRC		HON		IRN		USA		ALG

```
BRA - NED ●——80%——
                        BRA - ITA
COL - ITA ●  54%              ——84%——
                                        BRA - GER
SUI - BIH ●——77%——                            ——50%——
                        SUI - GER  73%                GER - ESP ——66%—— GER
GER - KOR ●  90%                                            ——52%——
ESP - CRO ●——66%——
                        ESP - ENG
ENG - CIV ●  63%              ——69%——
                                        ESP - BEL  55%
ARG - FRA ●——63%——
                        ARG - BEL
BEL - POR ●  55%
```