



# Introduction to Footballmetrics

Who will win the 2006 world cup?





# Data: FIFA world cups



- Samples are goals scored by the country
- Variables are goals against the country

FIFA2002-real												
	England	France	Germany	Ireland	Italy	Japan	Mexico	Nigeria	Paraguay	Poland	Portugal	Rus
	11	12	13	14	15	16	17	18	19	20	21	22
ARG	1	0.0000	m	m	m	m	m	1.0000	m	m	m	m
BEL	2	m	m	m	m	m	2.0000	m	m	m	m	m
BRA	3	2.0000	m	2.0000	m	m	m	m	m	m	m	3.0000
CAM	4	m	m	0.0000	0.0000	m	m	m	m	m	m	m
CHI	5	m	m	m	m	m	m	m	m	m	m	m
COS	6	m	m	m	m	m	m	m	m	m	m	m
CRO	7	m	m	m	m	2.0000	m	0.0000	m	m	m	m
DEN	8	0.0000	2.0000	m	m	m	m	m	m	m	m	m
ECU	9	m	m	m	m	0.0000	m	1.0000	m	m	m	m
ENG	10	m	m	m	m	m	m	0.0000	m	m	3.0000	m
FRA	11	m	m	m	m	m	m	m	m	m	m	m
GER	12	m	m	m	1.0000	m	m	m	1.0000	m	m	m
IRE	13	m	m	1.0000	m	m	m	m	m	m	m	m
ITA	14	m	m	m	m	m	m	1.0000	m	m	m	m
JAP	15	m	m	m	m	m	m	m	m	m	m	1.0000
MEX	16	m	m	m	m	1.0000	m	m	m	m	m	m
NIG	17	0.0000	m	m	m	m	m	m	m	m	m	m
PAR	18	m	m	0.0000	m	m	m	m	m	m	m	m
POL	19	m	m	m	m	m	m	m	m	m	0.0000	m
POR	20	2.0000	m	m	m	m	m	m	m	4.0000	m	m

Problem: Data table suffers from severe sparsity





# Data filling by PCA

FIFA2002-real

		England	France	Germany	Ireland	Italy	Japan	Mexico	Nigeria	Paraguay	Poland	Portugal	Rus
		11	12	13	14	15	16	17	18	19	20	21	22
ARG	1	0.0000	m	m	m	m	m	m	1.0000	m	m	m	
BEL	2	m	m	m	m	m	2.0000	m	m	m	m	m	
BRA	3	2.0000	m	2.0000	m	m	m	m	m	m	m	m	3.0000
CAM	4	4.7828e-02	0.2663	0.0000	0.0000	0.6426	0.1708	0.4090	0.3051	m	m	m	
CHI	5	5.2349e-03	0.2258	-0.3007	4.4728e-02	0.6043	0.1359	0.3729	0.2669	m	m	m	
COS	6	1.0172	1.1881	0.6183	0.6761	1.5125	0.9662	1.2318	1.1741	m	m	m	
CRO	7	0.2507	0.4592	-7.7760e-02	0.1979	2.0000	0.3373	0.0000	0.4869	m	m	m	
DEN	8	0.0000	2.0000	0.4554	0.5642	1.3515	0.8190	1.0795	1.0133	m	m	m	1.0000
ECU	9	0.3556	0.5590	1.7505e-02	0.2633	0.0000	0.4233	1.0000	0.5810	m	m	m	
ENG	10	1.1263	1.2918	0.7174	0.7442	1.6104	1.0557	1.3244	0.0000	m	m	m	
FRA	11	-0.2682	-3.4214e-02	-0.5490	-0.1259	0.3590	-8.8483e-02	0.1408	2.1784e-02	m	m	m	
GER	12	0.9755	1.1484	0.5804	1.0000	1.4750	0.9319	1.1963	1.1366	m	m	m	
IRE	13	0.9041	1.0805	1.0000	0.6055	1.4109	0.8733	1.1357	1.0726	m	m	m	
ITA	14	0.9697	1.1429	0.5752	0.6465	1.4699	0.9272	1.0000	1.1315	m	m	m	
JAP	15	0.8088	0.9899	0.4290	0.5461	1.3255	0.7952	1.0549	0.9872	m	m	m	
MEX	16	0.6235	0.8136	0.2607	0.4304	1.0000	0.6431	0.8975	0.8211	m	m	m	
NIG	17	0.0000	2.3325e-02	-0.4940	-8.8113e-02	0.4133	-3.8836e-02	0.1921	7.6027e-02	m	m	m	
PAR	18	0.7373	0.9219	0.0000	0.5015	1.2613	0.7365	0.9942	0.9232	m	m	m	
POL	19	0.7213	0.9067	0.3496	0.4915	1.2469	0.7234	0.9806	0.9088	m	m	m	
POR	20	2.0000	1.4719	0.8894	0.8623	1.7803	1.2111	1.4851	1.4416	m	m	m	

FIFA2002-real

England France Germany Ireland Italy Japan Mexico Nigeria Par

11 12 13 14 15 16 17 18

ARG 1 0.0000 0.5446 3.8001e-03 0.2539 0.9052 0.4110 0.6574 1.0000 1

BEL 2 0.5543 0.7479 0.1980 0.3873 1.0971 2.0000 0.8389 0.7591 1

BRA 3 2.0000 2.1445 2.0000 1.3036 2.4150 1.7914 2.0854 2.0756 2

CAM 4 4.7828e-02 0.2663 0.0000 0.0000 0.6426 0.1708 0.4090 0.3051 1

CHI 5 5.2349e-03 0.2258 -0.3007 4.4728e-02 0.6043 0.1359 0.3729 0.2669 0

COS 6 1.0172 1.1881 0.6183 0.6761 1.5125 0.9662 1.2318 1.1741 1

CRO 7 0.2507 0.4592 -7.7760e-02 0.1979 2.0000 0.3373 0.0000 0.4869 1

DEN 8 0.0000 2.0000 0.4554 0.5642 1.3515 0.8190 1.0795 1.0133 1

ECU 9 0.3556 0.5590 1.7505e-02 0.2633 0.0000 0.4233 1.0000 0.5810 1

ENG 10 1.1263 1.2918 0.7174 0.7442 1.6104 1.0557 1.3244 0.0000 2

FRA 11 -0.2682 -3.4214e-02 -0.5490 -0.1259 0.3590 -8.8483e-02 0.1408 2.1784e-02 0

GER 12 0.9755 1.1484 0.5804 1.0000 1.4750 0.9319 1.1963 1.1366 1

IRE 13 0.9041 1.0805 1.0000 0.6055 1.4109 0.8733 1.1357 1.0726 1

ITA 14 0.9697 1.1429 0.5752 0.6465 1.4699 0.9272 1.0000 1.1315 1

JAP 15 0.8088 0.9899 0.4290 0.5461 1.3255 0.7952 1.0549 0.9872 1

MEX 16 0.6235 0.8136 0.2607 0.4304 1.0000 0.6431 0.8975 0.8211 1

NIG 17 0.0000 2.3325e-02 -0.4940 -8.8113e-02 0.4133 -3.8836e-02 0.1921 7.6027e-02 0

PAR 18 0.7373 0.9219 0.0000 0.5015 1.2613 0.7365 0.9942 0.9232 1

POL 19 0.7213 0.9067 0.3496 0.4915 1.2469 0.7234 0.9806 0.9088 1

POR 20 2.0000 1.4719 0.8894 0.8623 1.7803 1.2111 1.4851 1.4416 2

Luckily, Unscrambler can now fill in missing data





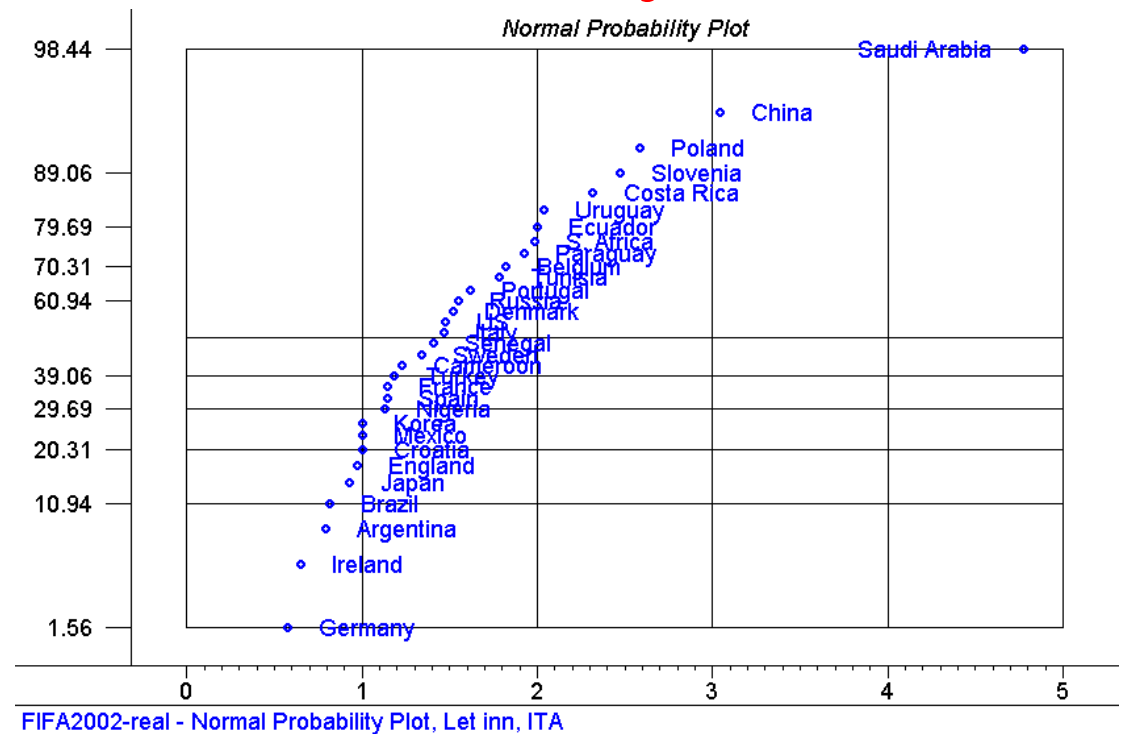
# Goal score statistics



Sample statistics are now estimates of goals scored

Italy would then score 4.7 and 0.6 goals against Saudi Arabia and Germany, respectively

Italy



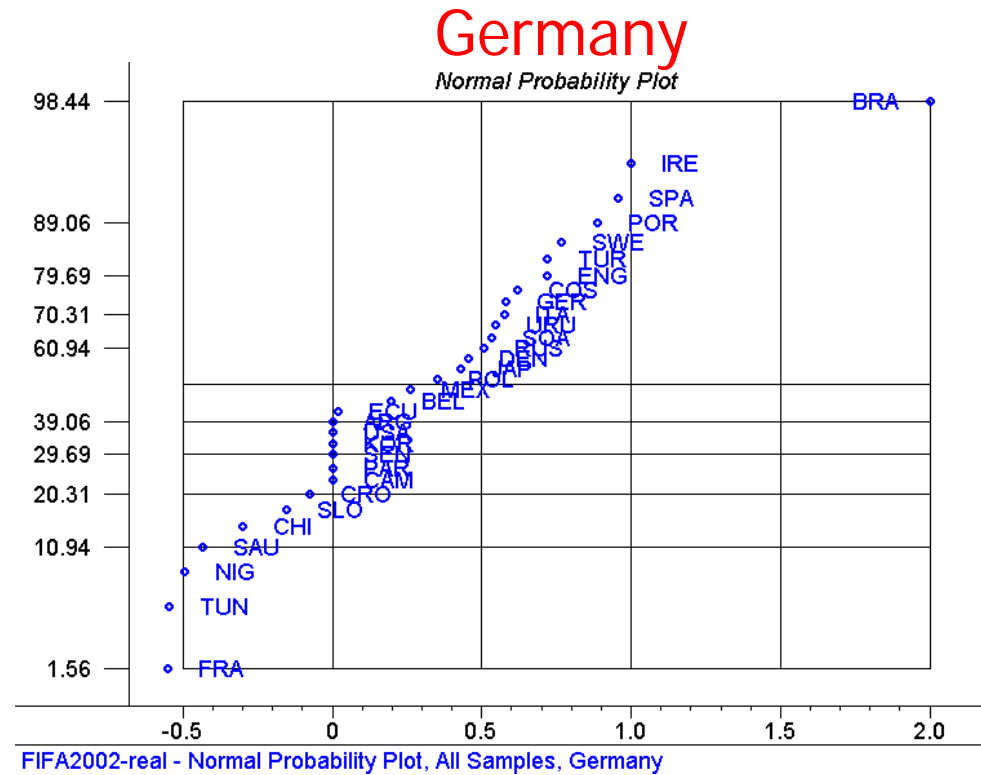


# Goal score statistics



Variable statistics are now estimates of goals against

Germany would then let in 2 and -0.6(!) goals against Brazil and France, respectively

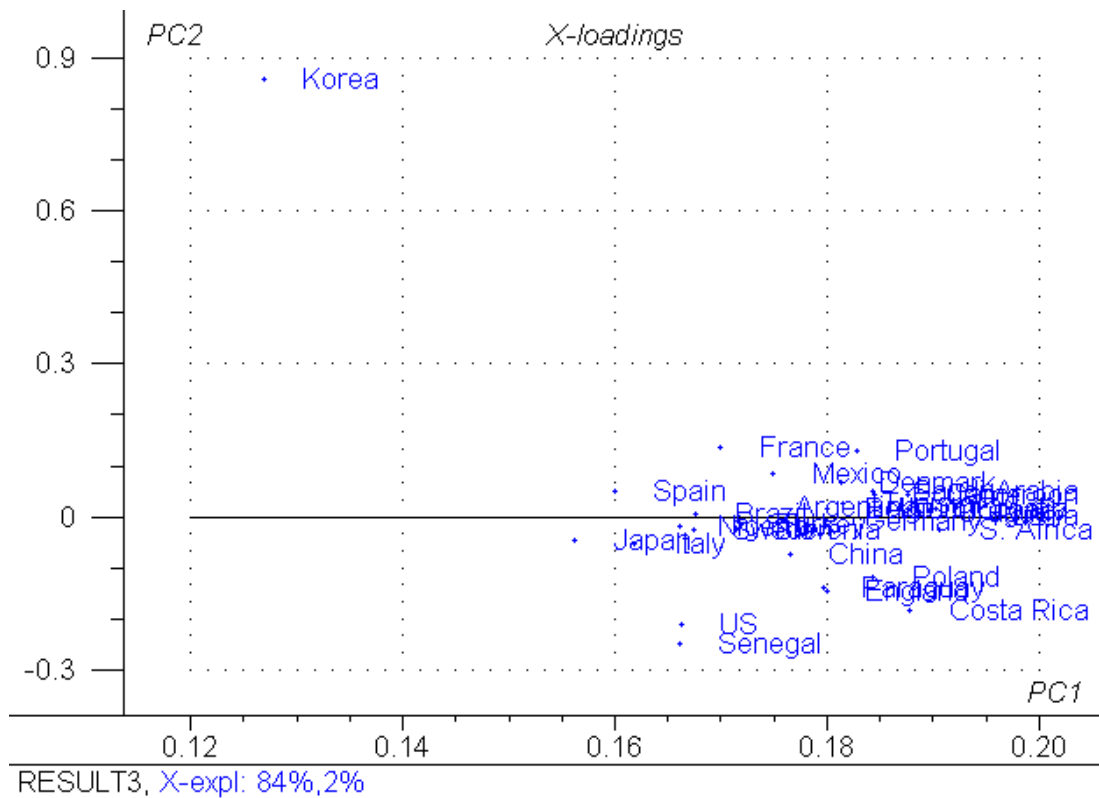




# PCA



Loading plot shows positive scoring pattern  
Korea stands out, as the 2002 surprise semi finalists.







# Second round (1/8-Finals)



Germany	2	1	Sweden
Argentina	1	1	Mexico
England	3	0	Ecuador
Portugal	2	1	Netherlands
Italy	3	1	Australia
Switzerland	3	2	Ukraine
Brazil	3	1	Ghana
Spain	2	1	France







# Quarter Finals



Germany	1	0	Argentina
Italy	1	0	Switzerland
England	3	2	Portugal
Brazil	2	1	Spain



Portugal penalized for having half the team with Brazilian passports.

Michael Ballack scores for Germany with "the hand of God"





# Semi Finals



Germany	2	1	Italy
England	1	0	Brazil



De Rossi sent off for a second elbow offence



Own goal by Ronaldinho in the 92. minute





# Final



Germany **1** **1** England

David Beckham misses in the penalty shoot out.





# Conclusion



We have confirmed the 1986 hypothesis of Gary Lineker:

*22 men play for 90 minutes...  
... and in the end Germany wins*

