The Actuarial Profession
making financial sense of the future

Life conference and exhibition 2010
Greg Becker

Who will win the Premier League?

7-9 November 2010

http://www.youtube.com/watch?v=u1RtEQ6ly3Q&feature=related

Team talk – setting expectations

- Practical statistics
  - Showing how a complex problem that can’t be solved using traditional methods, can be solved using Monte Carlo methods
  - Touching on Bayesian statistics, Monte Carlo methods
- Model development process
  - Theoretical foundation
  - Data & data problems
  - Testing and refining a model – an iterative process
- Practical application in other areas of actuarial work

Fixture List – a tale of two halves

Theory

- What could be in a model?
- What data could be used?
- Lessons learnt from the World Cup
  - Article Written in The Actuary
  - Compared to reality
- Half time – with half-time entertainment

Model in practice

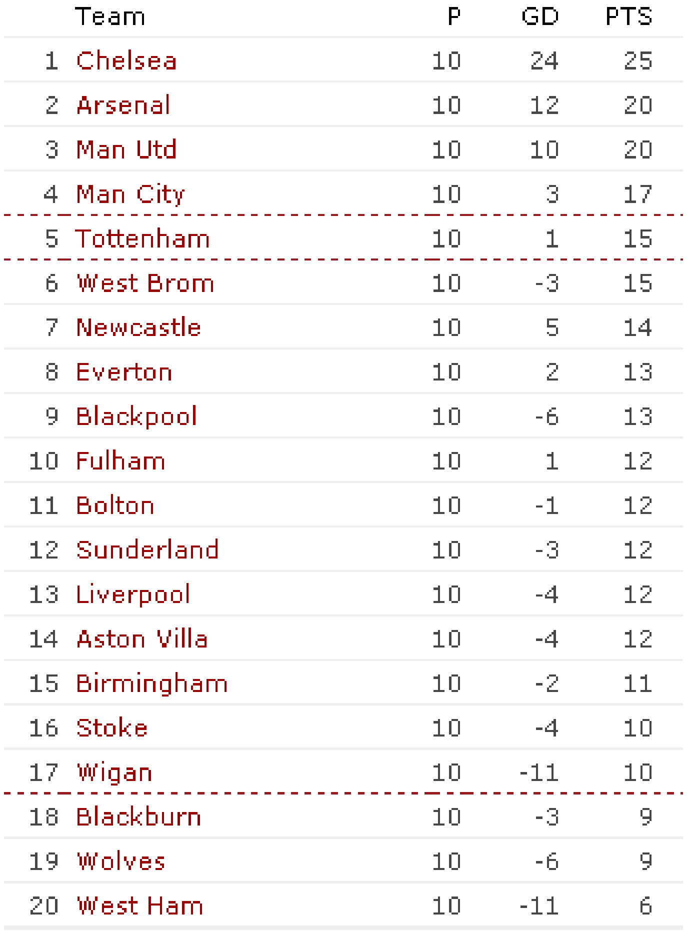
- Model proposed
  - Why and how?
  - How would it have done in 2009-2010?
  - How would it have done in 2008-2009?
- Betting stats
- Actuarial lessons

They say you should know your audience: Please clap or cheer when your team logo comes up

http://www.premierleague.com/page/FixturesResults/0,,12306,00.html

After 100 matches so far this season, this is the points table:

- In 2007/8, Arsenal and Manchester United were leading, and while Manchester United went on to win, but both Arsenal and Chelsea had led the table later in the season
- In 2008/9, Chelsea and Liverpool were leading at this stage, and neither went on to win!
- In 2009/10, Chelsea was already leading by 2 points, although Manchester United was leading the table as late as 2/4/2010



	Team	P	GD	PTS
1	Chelsea	10	24	25
2	Arsenal	10	12	20
3	Man Utd	10	10	20
4	Man City	10	3	17
5	Tottenham	10	1	15
6	West Brom	10	-3	15
7	Newcastle	10	5	14
8	Everton	10	2	13
9	Blackpool	10	-6	13
10	Fulham	10	1	12
11	Bolton	10	-1	12
12	Sunderland	10	-3	12
13	Liverpool	10	-4	12
14	Aston Villa	10	-4	12
15	Birmingham	10	-2	11
16	Stoke	10	-4	10
17	Wigan	10	-11	10
18	Blackburn	10	-3	9
19	Wolves	10	-6	9
20	West Ham	10	-11	6

Betting on the leader at this stage of the season is not a sure bet!

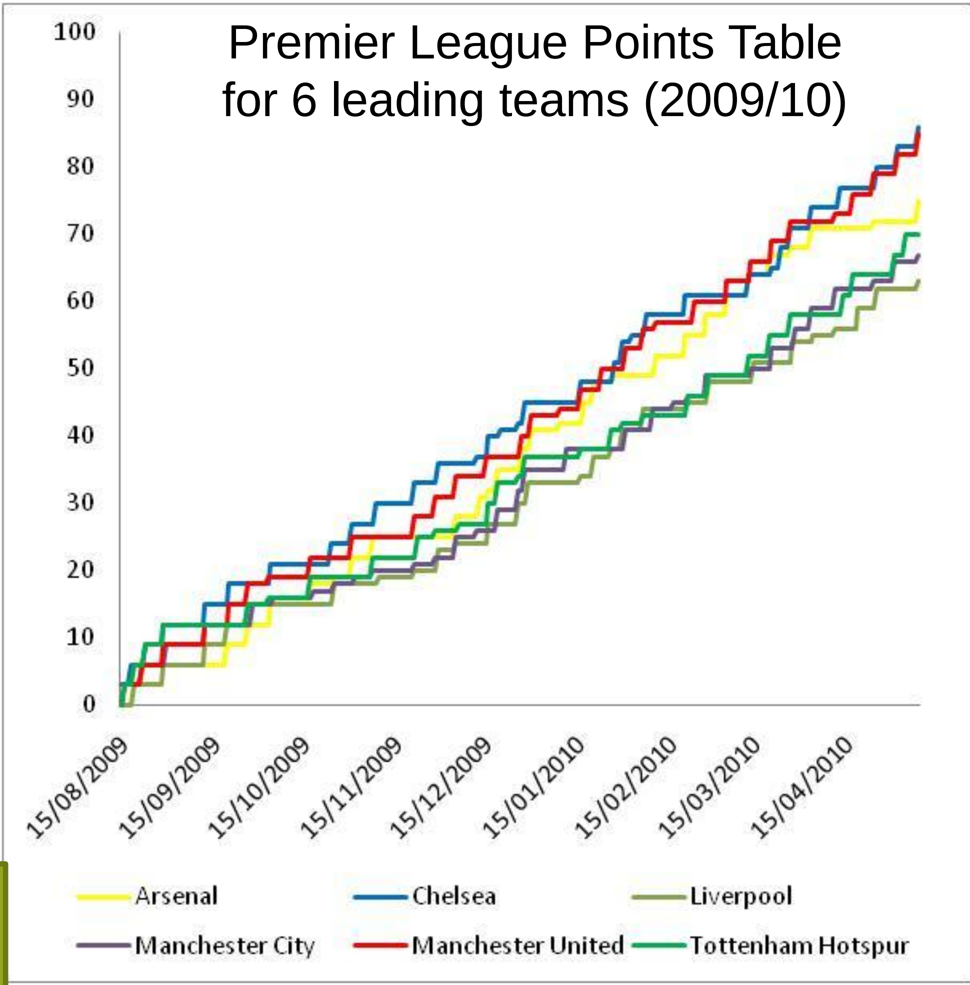
http://news.bbc.co.uk/sport1/hi/football/eng_prem/default.stm

Simplest prediction tool: The team leading is almost certainly the best team to bet on…

But many things should be taken into account

- games in hand
- the opposition to come
- the injury list
- involvement in other competitions and
- playing home or away

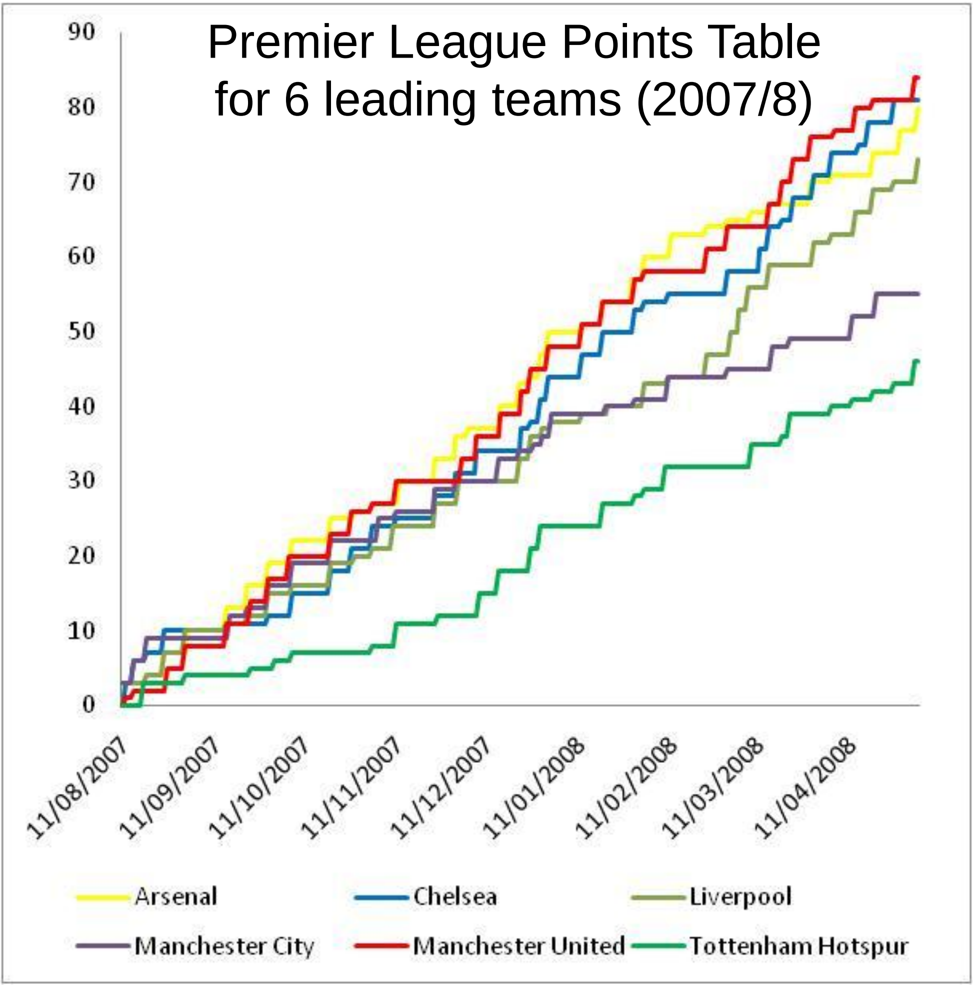
which all can influence the outcome

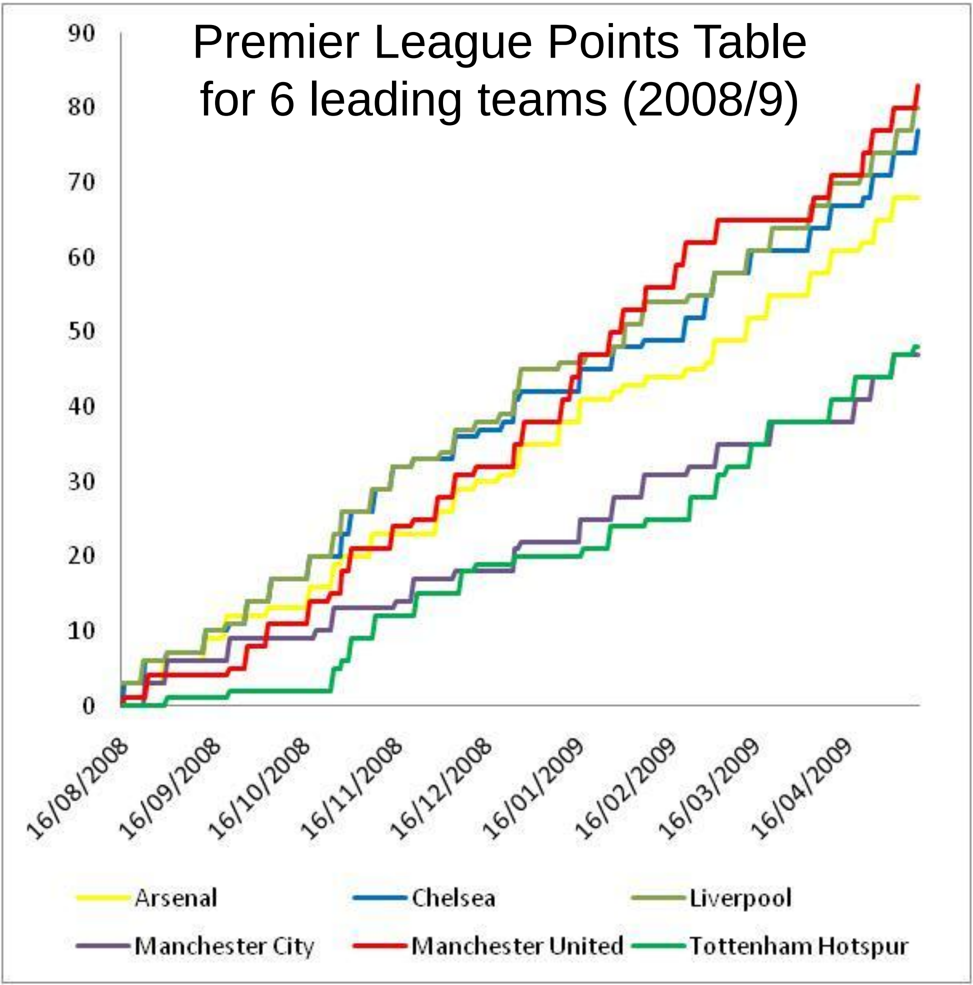


Premier League Points Table for 6 leading teams (2009/10)

Manchester United had more points than Chelsea on 24/4/2010, although Chelsea won the league by a point 16 days later!

The points table leader (at this stage of the season) did not remain on top of the league until the end of the season

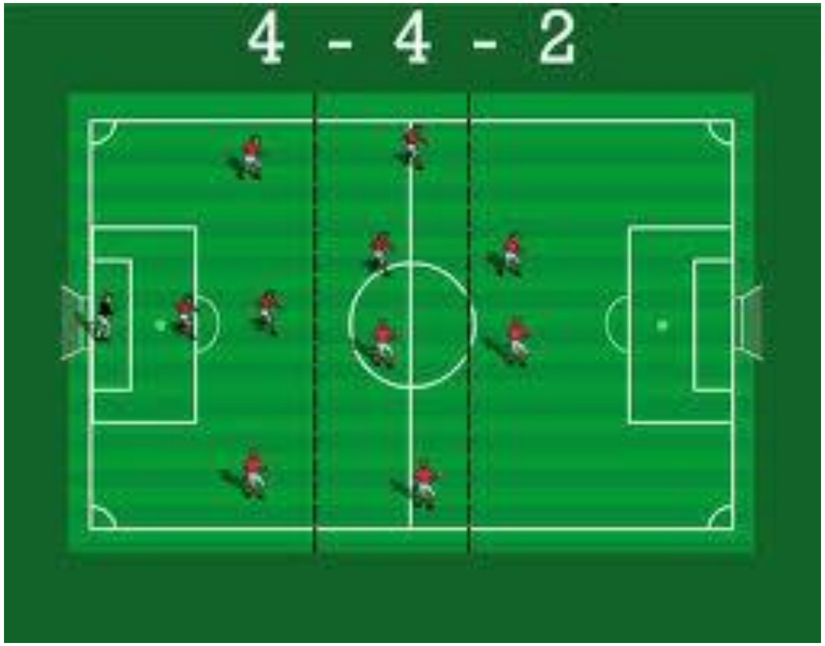


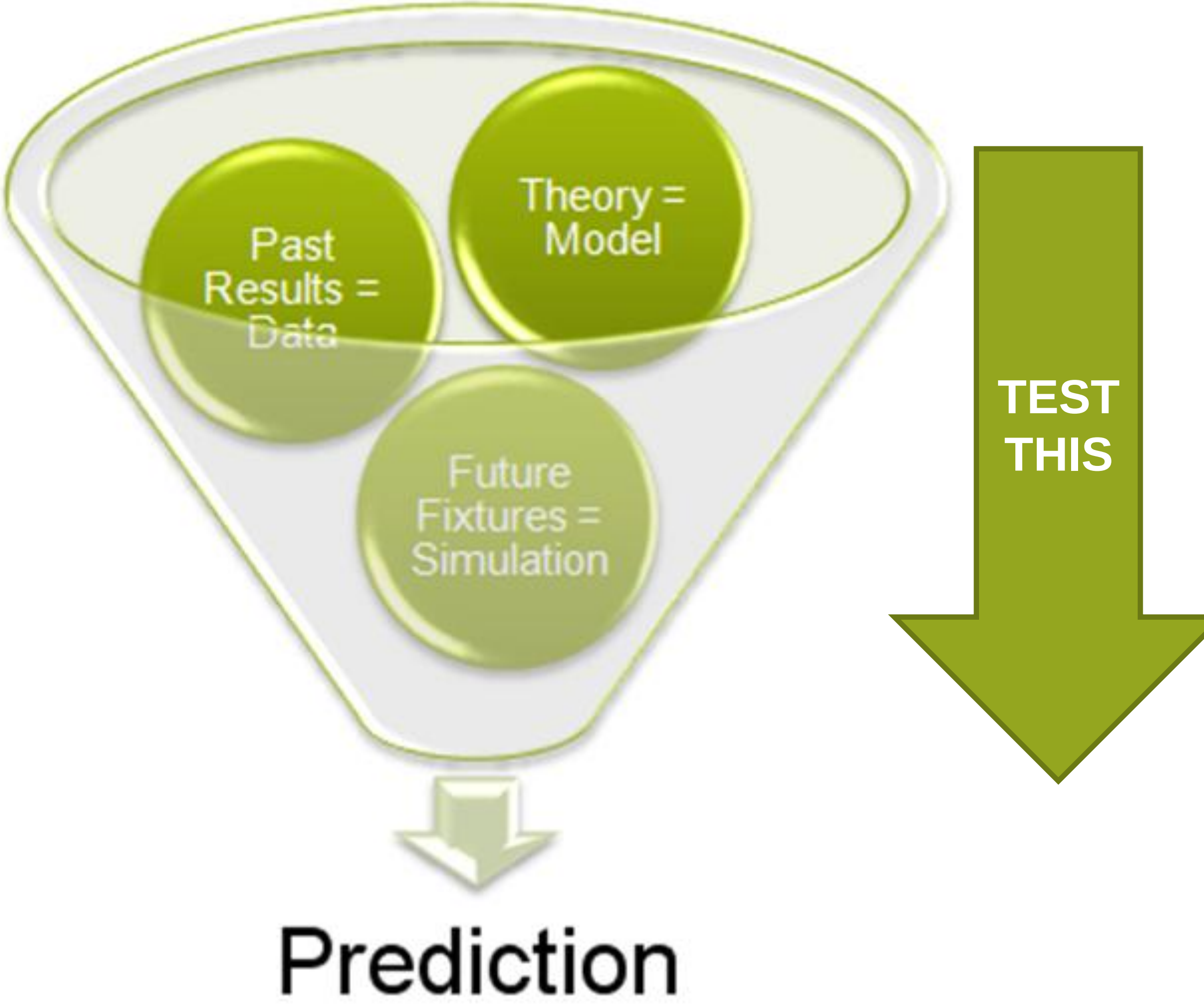


In 2007/8 Manchester United remained on top of the table from 15/3/2008 onwards

In 2008/9 Manchester United remained on top of the table from 7/2/2009 onwards

How does everything fit together?





Past Results = Data

Theory = Model

Future Fixtures = Simulation

Prediction

TEST THIS

What could be in a model to predict the season’s league winner? Some ground rules…

(Simplifying) Assumptions

- Teams don’t change over the season
- Results in other competitions do not affect the premier league outcome
- Each match is independent of each other
- Teams perform the same way independent of the competition
- No “bankruptcy” point penalties

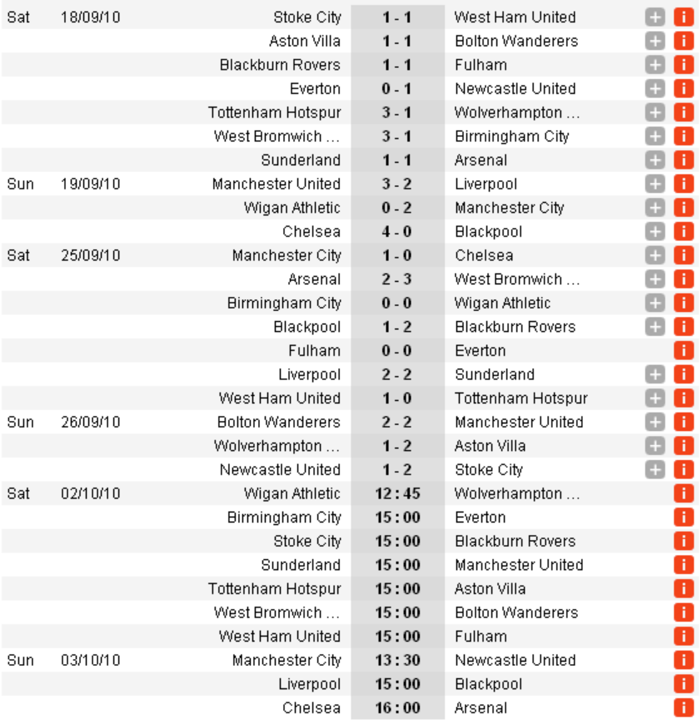
Complications overlooked

- Players get traded, players get injured, or go into and out of form
- Players getting over-played by the end of the season due to success in other competitions
- Teams have streaks of form that affects their confidence
- Teams have derby’s and particular rivalries

We are going to use the results so far to predict the results of the remaining matches, and thus discern who will win

Why a Bayesian statistical approach? If we regard 2010/11 results as the only relevant data then…

- At the start of the season we know nothing
  - Every team has the same chance of winning
- As the season progresses we gather more data
  - Goals scored and goals conceded by every team, home and away
- Our prediction needs to be continuously refined to fit the latest data



Sat	18/09/10	Stoke City	1 - 1	West Ham United	+	i
		Aston Villa	1 - 1	Bolton Wanderers	+	i
		Blackburn Rovers	1 - 1	Fulham	+	i
		Everton	0 - 1	Newcastle United	+	i
		Tottenham Hotspur	3 - 1	Wolverhampton …	+	i
		West Bromwich …	3 - 1	Birmingham City	+	i
		Sunderland	1 - 1	Arsenal	+	i
Sun	19/09/10	Manchester United	3 - 2	Liverpool	+	i
		Wigan Athletic	0 - 2	Manchester City	+	i
		Chelsea	4 - 0	Blackpool	+	i
Sat	25/09/10	Manchester City	1 - 0	Chelsea	+	i
		Arsenal	2 - 3	West Bromwich …	+	i
		Birmingham City	0 - 0	Wigan Athletic	+	i
		Blackpool	1 - 2	Blackburn Rovers	+	i
		Fulham	0 - 0	Everton		i
		Liverpool	2 - 2	Sunderland	+	i
		West Ham United	1 - 0	Tottenham Hotspur	+	i
Sun	26/09/10	Bolton Wanderers	2 - 2	Manchester United	+	i
		Wolverhampton …	1 - 2	Aston Villa	+	i
		Newcastle United	1 - 2	Stoke City	+	i
Sat	02/10/10	Wigan Athletic	12 : 45	Wolverhampton …		i
		Birmingham City	15 : 00	Everton		i
		Stoke City	15 : 00	Blackburn Rovers		i
		Sunderland	15 : 00	Manchester United		i
		Tottenham Hotspur	15 : 00	Aston Villa		i
		West Bromwich …	15 : 00	Bolton Wanderers		i
		West Ham United	15 : 00	Fulham		i
Sun	03/10/10	Manchester City	13 : 30	Newcastle United		i
		Liverpool	15 : 00	Blackpool		i
		Chelsea	16 : 00	Arsenal		i

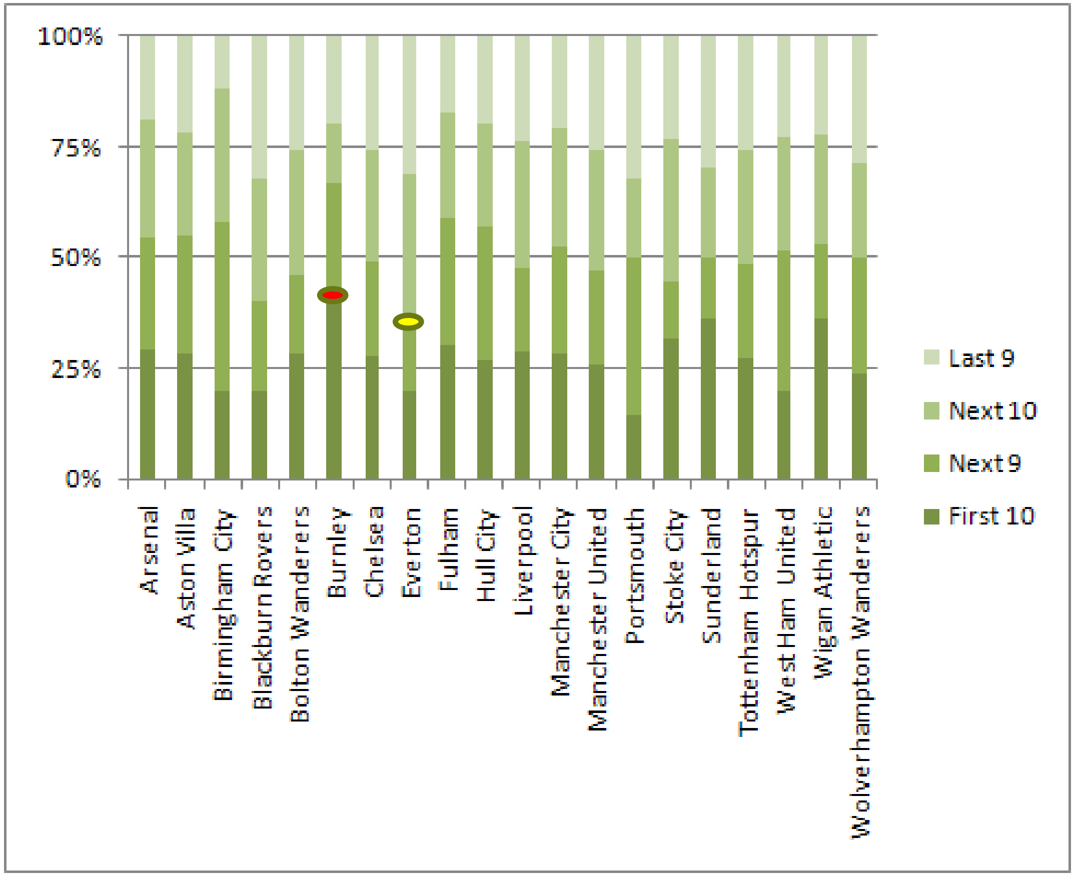
$$P(H|E) = \frac{P(E|H)\,P(H)}{P(E)}$$

http://en.wikipedia.org/wiki/Bayesian_inference
http://uk.soccerway.com/national/england/premier-league/2010-2011/regular-season/matches/

This model relies on teams being consistent. Is this a valid assumption?

As we can see, in the 2009/2010 season, there were some outliers

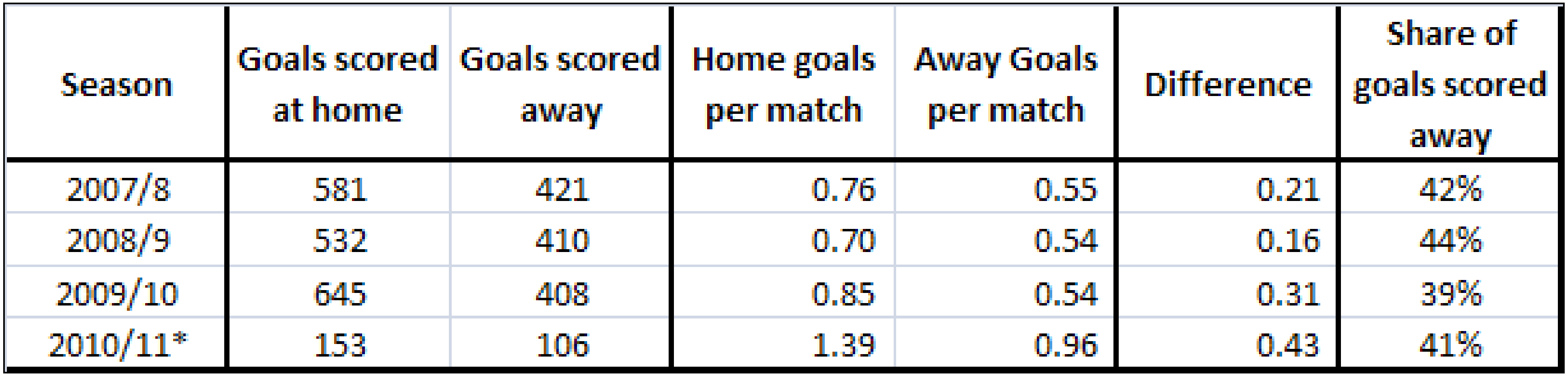
- Burley scored 40% of their seasons points in their first 10 matches
- Everton scored 64% of their points in the second half of the season



What about home advantage?

Home and Away: do we need to take this into account?

- Different for different teams?
- Home ground has bigger or smaller impact/difference?
- Which home fans are the best? Or is it teams not travelling well?



Season	Goals scored at home	Goals scored away	Home goals per match	Away Goals per match	Difference	Share of goals scored away
2007/8	581	421	0.76	0.55	0.21	42%
2008/9	532	410	0.70	0.54	0.16	44%
2009/10	645	408	0.85	0.54	0.31	39%
2010/11*	153	106	1.39	0.96	0.43	41%

We could estimate that just over 40% of goals are scored by the away team

Why are we using the Monte Carlo method?

We can’t solve the problem analytically!

- The winner of the premier league will be the result of the remaining 280 matches
- Since each match can have one of three outcomes

Win home + loss away

Draw

Loss home + Win away

modeling the rest of the season deterministically would result in 3^{280} different possible outcomes being calculated – which is a number that has **133** digits!

We know that the best answer should reflect our uncertainty, and the Monte Carlo method reflects this, generating as a result, a distribution of the relative likelihood of the alternate outcomes

How can we set the model up to be run using the Monte Carlo method?

- Since the remaining matches all happen independently, we can model each independently
- Since each match has a home team and an away team we can reflect that too
- Since each team has played a series of matches, and has scored and conceded goals, we can model the probability of all possible results, where for instance the result

Home scores H and Away scores A

can be reflected as

P(Home=H and Away=A)=P(Home score H)P(Away concede H)P(Home concede A)P(Away score A)

All we need now is estimates of these probabilities… DATA

Does it matter who plays who in each individual match?

Since our measure of how good the model is, is the likelihood estimate, and the likelihood estimate is of the following form:

$$\prod_{i\ matches}(\text{Home team scoring } x_i)(\text{Away team conceding } x_i)(\text{Home team conceding } y_i)(\text{Away team scoring } y_i)$$

it can be shown that the estimates are independent of who played who, but rather dependent on how many goals were scored or conceded by the home and away team each game:

$$\prod_{i\ matches}(\text{Home team scoring } x_i)\prod_{i\ matches}(\text{Away team conceding } x_i)\prod_{i\ matches}(\text{Home team conceding } y_i)\prod_{i\ matches}(\text{Away team scoring } y_i)$$

This may be counter intuitive, but reflects our underlying assumption that a team has a constant ‘average scoring rate’ and ‘average conceding rate’ – which is constant across the season irrespective of the opposition

What do I mean by data?

Retrospective

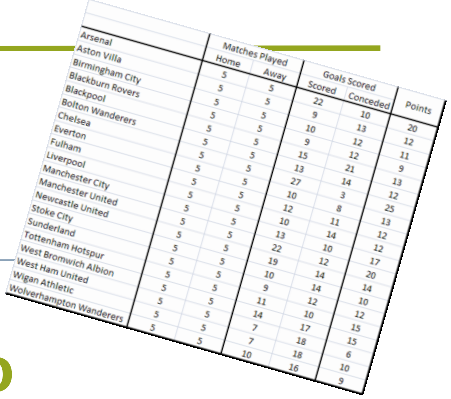
- Past results
  - Goals scored
  - Goals conceded
  - Who played
  - Current league points

Prospective

- Fixture list
  - Home and away
  - Playing against whom

Clean the data, validate the data.
Luckily there are no reporting delays…
but one is almost immediately out of date…

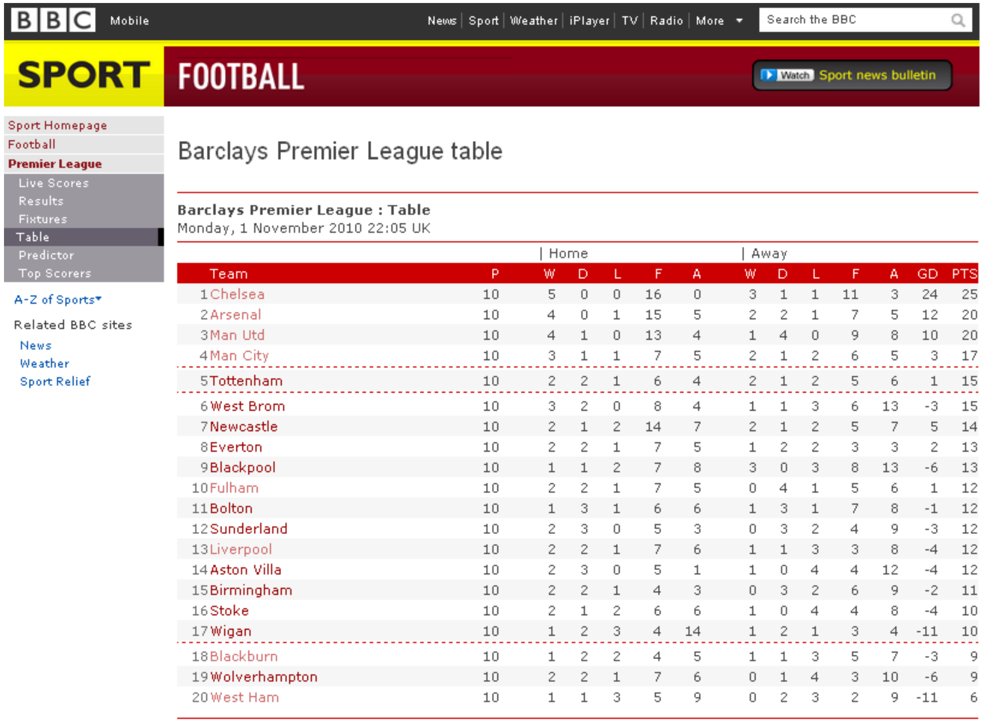
So what is our data?



	Matches Played		Goals Scored		Points
	Home	Away	Scored	Conceded	
Arsenal	5	5	22	10	20
Aston Villa	5	5	9	13	12
Birmingham City	5	5	10	12	11
Blackburn Rovers	5	5	9	12	9
Blackpool	5	5	15	21	13
Bolton Wanderers	5	5	13	14	12
Chelsea	5	5	27	3	25
Everton	5	5	10	8	13
Fulham	5	5	12	11	12
Liverpool	5	5	10	14	12
Manchester City	5	5	13	10	17
Manchester United	5	5	22	12	20
Newcastle United	5	5	19	14	14
Stoke City	5	5	10	14	10
Sunderland	5	5	9	12	12
Tottenham Hotspur	5	5	11	10	15
West Bromwich Albion	5	5	14	17	15
West Ham United	5	5	7	18	6
Wigan Athletic	5	5	7	18	10
Wolverhampton Wanderers	5	5	10	16	9

100 Matches completed

- We know where we've been



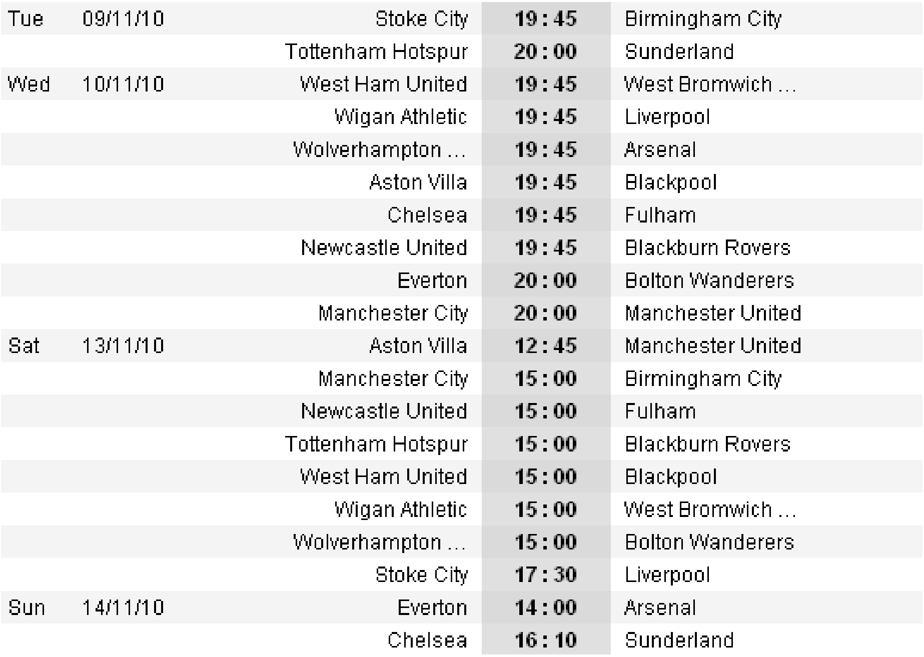
Barclays Premier League table

Barclays Premier League : Table
Monday, 1 November 2010 22:05 UK

		Home					Away						
Team	P	W	D	L	F	A	W	D	L	F	A	GD	PTS
1 Chelsea	10	5	0	0	16	0	3	1	1	11	3	24	25
2 Arsenal	10	4	0	1	15	5	2	2	1	7	5	12	20
3 Man Utd	10	4	1	0	13	4	1	4	0	9	8	10	20
4 Man City	10	3	1	1	7	5	2	1	2	6	5	3	17
5 Tottenham	10	2	2	1	6	4	2	1	2	5	6	1	15
6 West Brom	10	3	2	0	8	4	1	1	3	6	13	-3	15
7 Newcastle	10	2	1	2	14	7	2	1	2	5	7	5	14
8 Everton	10	2	2	1	7	5	1	2	2	3	3	2	13
9 Blackpool	10	1	1	2	7	8	3	0	3	8	13	-6	13
10 Fulham	10	2	2	1	7	5	0	4	1	5	6	1	12
11 Bolton	10	1	3	1	6	6	1	3	1	7	8	-1	12
12 Sunderland	10	2	3	0	5	3	0	3	2	4	9	-3	12
13 Liverpool	10	2	2	1	7	6	1	1	3	3	8	-4	12
14 Aston Villa	10	2	3	0	5	1	1	0	4	4	12	-4	12
15 Birmingham	10	2	2	1	4	3	0	3	2	6	9	-2	11
16 Stoke	10	2	1	2	6	6	1	0	4	4	8	-4	10
17 Wigan	10	1	2	3	4	14	1	2	1	3	4	-11	10
18 Blackburn	10	1	2	2	4	5	1	1	3	5	7	-3	9
19 Wolverhampton	10	2	2	1	7	6	0	1	4	3	10	-6	9
20 West Ham	10	1	1	3	5	9	0	2	3	2	9	-11	6

280 matches to go

- We know where we're going



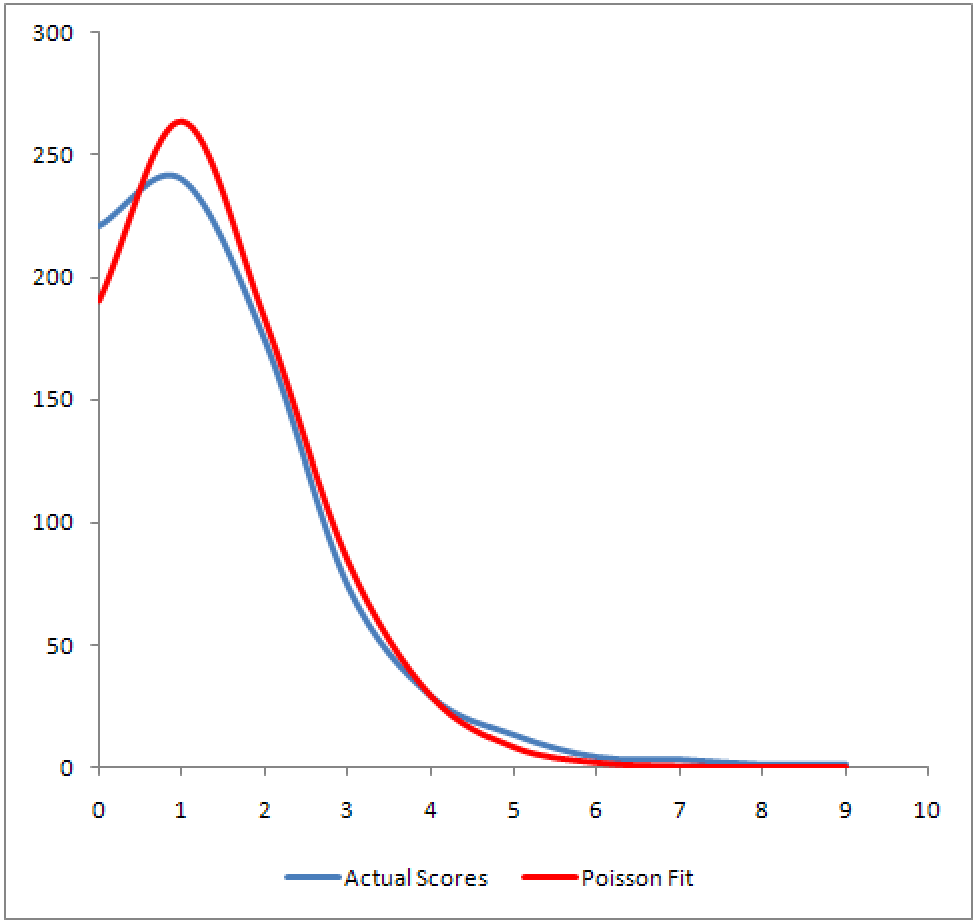
Tue	09/11/10	Stoke City	19 : 45	Birmingham City
		Tottenham Hotspur	20 : 00	Sunderland
Wed	10/11/10	West Ham United	19 : 45	West Bromwich ...
		Wigan Athletic	19 : 45	Liverpool
		Wolverhampton ...	19 : 45	Arsenal
		Aston Villa	19 : 45	Blackpool
		Chelsea	19 : 45	Fulham
		Newcastle United	19 : 45	Blackburn Rovers
		Everton	20 : 00	Bolton Wanderers
		Manchester City	20 : 00	Manchester United
Sat	13/11/10	Aston Villa	12 : 45	Manchester United
		Manchester City	15 : 00	Birmingham City
		Newcastle United	15 : 00	Fulham
		Tottenham Hotspur	15 : 00	Blackburn Rovers
		West Ham United	15 : 00	Blackpool
		Wigan Athletic	15 : 00	West Bromwich ...
		Wolverhampton ...	15 : 00	Bolton Wanderers
		Stoke City	17 : 30	Liverpool
Sun	14/11/10	Everton	14 : 00	Arsenal
		Chelsea	16 : 10	Sunderland

Usual actuarial checklist here: data capture

http://uk.soccerway.com/national/england/premier-league/2010-2011/regular-season/matches/
http://news.bbc.co.uk/sport1/hi/football/eng_prem/table/default.stm

The number of goals scored/conceded by a team can be fit using a Poisson distribution

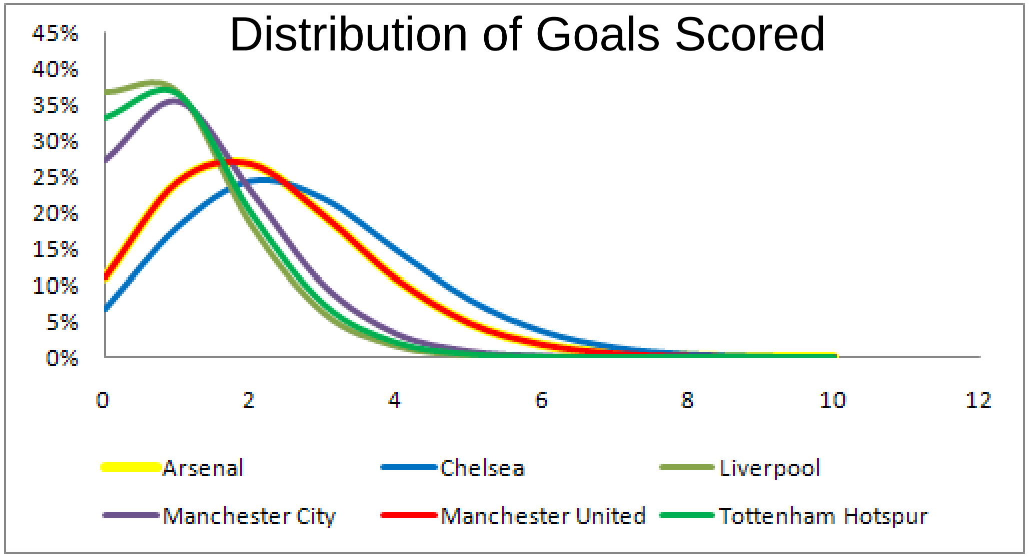
- In the 2009/10 season, there were 20 teams, and with each playing everyone else twice, there were 380 matches
- In each match, two teams ‘scored goals’ – making 760 data points, illustrated here
- A Poisson fits this distribution very well

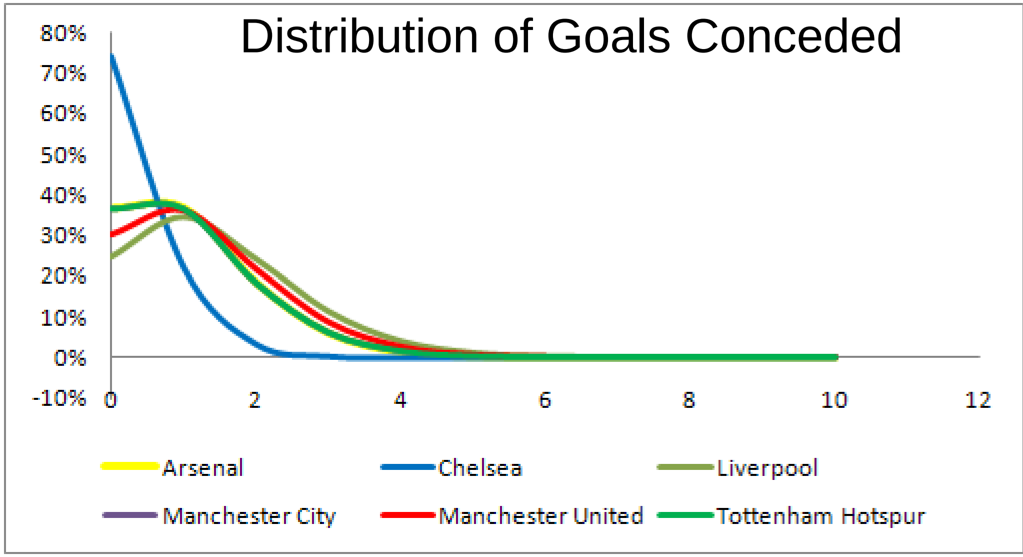


The goals scored and goals conceded results have been used to estimate a Poisson parameter for each team

- We are using a Poisson distribution assuming that the chance of scoring / conceding in a match some time in the future can be estimated using results from earlier in the season

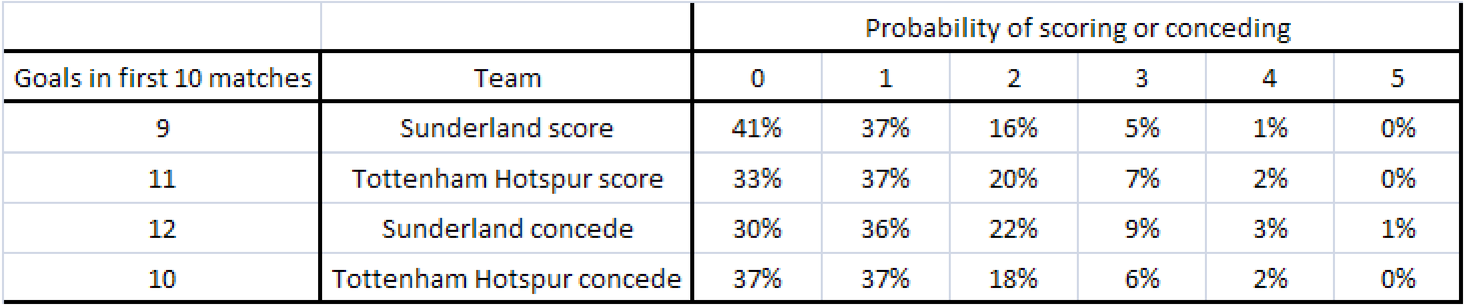
$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!},$$


Distribution of Goals Scored


Distribution of Goals Conceded

We are then able to model the outcome of a match using these goal “scoring” and “conceding” estimates

- A match coming up soon between Sunderland and Tottenham Hotspur can be modeled as follows
- Generate the goal scoring and goal conceding probabilities for each team based on it’s record (in this case their involvement in the first 100 matches of the season)

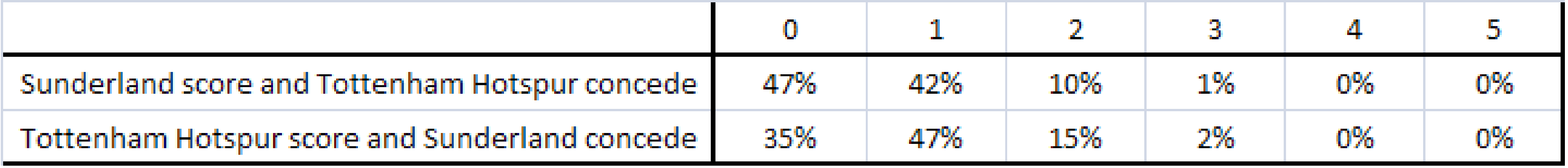


		Probability of scoring or conceding					
Goals in first 10 matches	Team	0	1	2	3	4	5
9	Sunderland score	41%	37%	16%	5%	1%	0%
11	Tottenham Hotspur score	33%	37%	20%	7%	2%	0%
12	Sunderland concede	30%	36%	22%	9%	3%	1%
10	Tottenham Hotspur concede	37%	37%	18%	6%	2%	0%

- Use this to develop an estimate of the number of goals scored by Sunderland & conceded by Tottenham
- Use this to develop an estimate of the number of goals conceded by Sunderland & scored by Tottenham
- Randomly simulate the match and calculate the result

We are then able to model the outcome of a match using these goal “scoring” and “conceding” estimates

- A match coming up soon between Sunderland and Tottenham Hotspur can be modeled as follows
- Generate the goal scoring and goal conceding probabilities for each team based on it’s record (in this case their involvement in the first 100 matches of the season)
- Use this to develop an estimate of the number of goals scored by Sunderland & conceded by Tottenham
- Use this to develop an estimate of the number of goals conceded by Sunderland & scored by Tottenham



	0	1	2	3	4	5
Sunderland score and Tottenham Hotspur concede	47%	42%	10%	1%	0%	0%
Tottenham Hotspur score and Sunderland concede	35%	47%	15%	2%	0%	0%

- Randomly simulate the match and calculate the result i.e.
  - 3 points for Sunderland and 0 for Tottenham
  - 1 each, or
  - 3 for Tottenham and 0 for Sunderland

Once you have a parameter set, what next? Runs and runs and runs and runs and runs and runs and runs…

- Now that we have agreed on the parameter set, we can 'randomly simulate the results of the rest of the season'
- We have done this 10,000 times

As an aside, the actual likelihood for our best set of parameters (which is the best estimate for the model we have developed, or the maximum likelihood estimate of the parameters) can be used to show that the score lines so far this season have about a 1 in 10^{-225} probability of having occurred

This is a small number, but since there are an infinite number of possibilities…

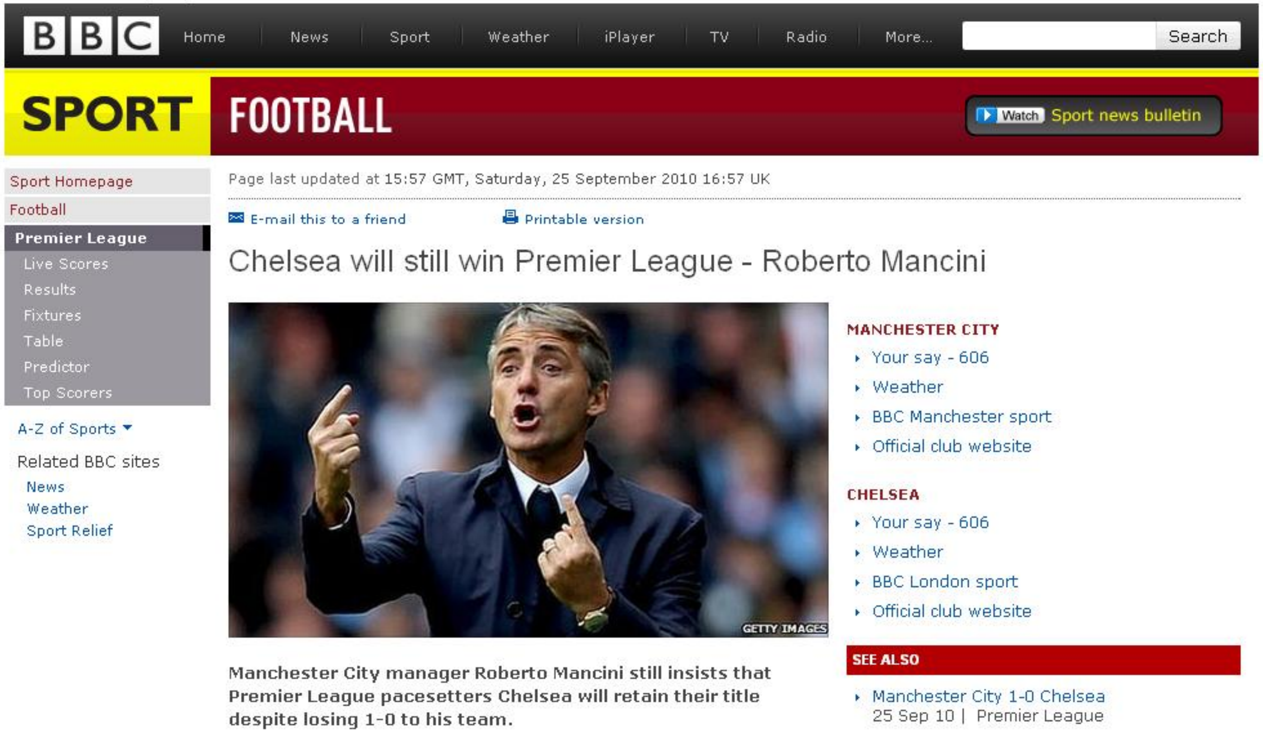
Testing the model – sometimes the model answer isn't what you expect!

Just because you have a good looking model, doesn't mean you have the answer!

http://www.youtube.com/watch?v=WALIARHHLII&feature=related

Predictions are difficult things to make… it's easy to make a blooper…

Man City are title rivals - Ancelotti

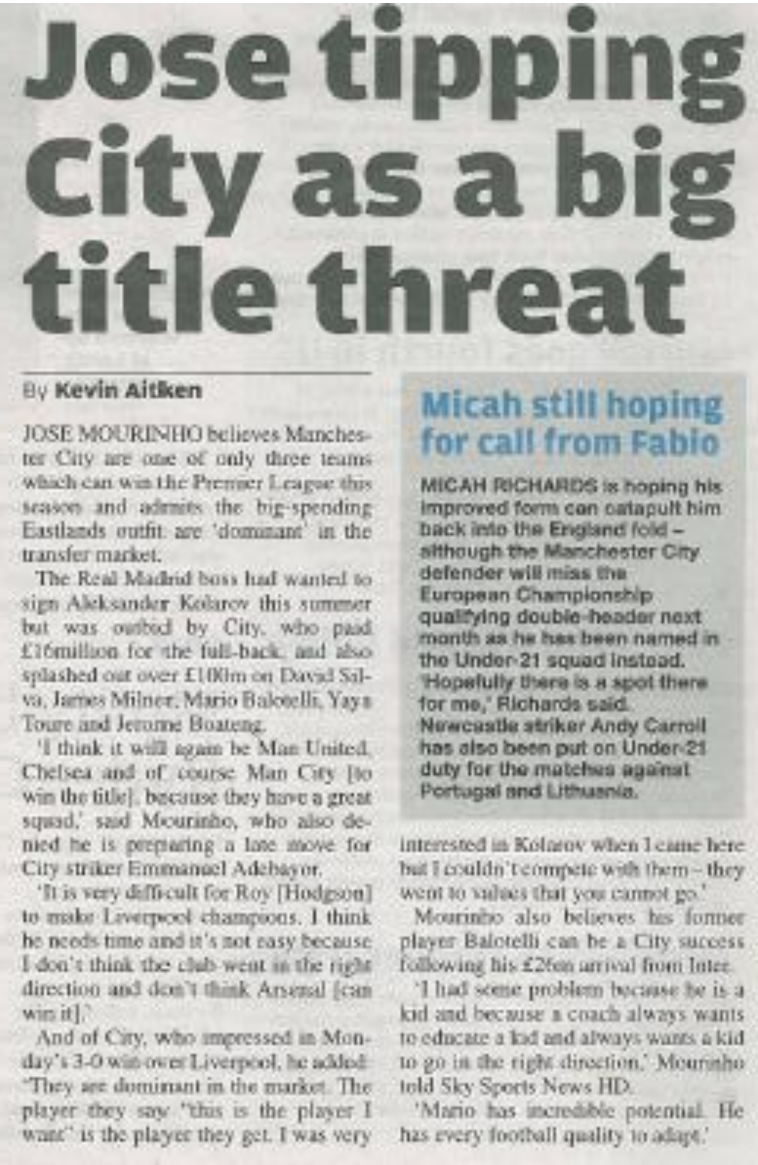


Page last updated at 15:57 GMT, Saturday, 25 September 2010 16:57 UK

Chelsea will still win Premier League - Roberto Mancini

Manchester City manager Roberto Mancini still insists that Premier League pacesetters Chelsea will retain their title despite losing 1-0 to his team.

Manchester City 1-0 Chelsea
25 Sep 10 | Premier League



Jose tipping City as a big title threat

By **Kevin Aitken**

JOSE MOURINHO believes Manchester City are one of only three teams which can win the Premier League this season and admits the big-spending Eastlands outfit are 'dominant' in the transfer market.

The Real Madrid boss had wanted to sign Aleksander Kolarov this summer but was outbid by City, who paid £16million for the full-back, and also splashed out over £100m on David Silva, James Milner, Mario Balotelli, Yaya Toure and Jerome Boateng.

'I think it will again be Man United, Chelsea and of course Man City [to win the title], because they have a great squad,' said Mourinho, who also denied he is preparing a late move for City striker Emmanuel Adebayor.

'It is very difficult for Roy [Hodgson] to make Liverpool champions. I think he needs time and it's not easy because I don't think the club went in the right direction and don't think Arsenal [can win it].'

And of City, who impressed in Monday's 3-0 win over Liverpool, he added: 'They are dominant in the market. The player they say "this is the player I want" is the player they get. I was very interested in Kolarov when I came here but I couldn't compete with them – they went to values that you cannot go.'

Mourinho also believes his former player Balotelli can be a City success following his £26m arrival from Inter.

'I had some problem because he is a kid and because a coach always wants to educate a kid and always wants a kid to go in the right direction,' Mourinho told Sky Sports News HD.

'Mario has incredible potential. He has every football quality to adapt.'

Micah still hoping for call from Fabio

MICAH RICHARDS is hoping his improved form can catapult him back into the England fold – although the Manchester City defender will miss the European Championship qualifying double-header next month as he has been named in the Under-21 squad instead. 'Hopefully there is a spot there for me,' Richards said. Newcastle striker Andy Carroll has also been put on Under-21 duty for the matches against Portugal and Lithuania.

http://news.bbc.co.uk/sport1/hi/football/eng_prem/9033272.stm

Half Time Entertainment: Clip of some great goals from the premier league season so far

http://www.youtube.com/watch?v=7fzxvSMdw3E

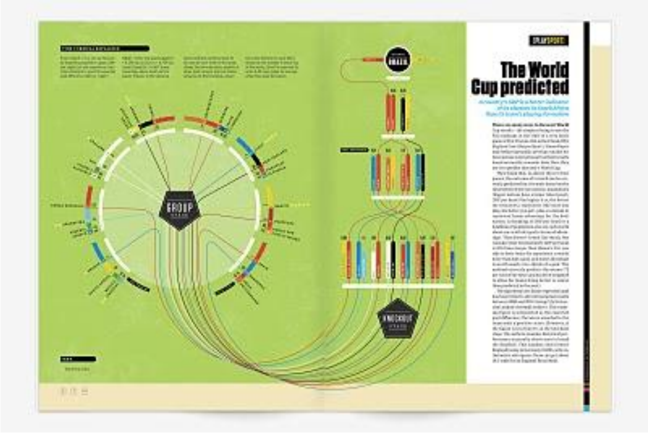
Lessons Learnt from the 2010 World Cup Prediction Model: retrospective data, prospective gamblers

Statistics World Cup

World Cup fever

Greg Becker and Arminder Kainth ask who is likely to win the 2010 FIFA World Cup in South Africa? Can you get an answer from a model?

Model	Predictive Ability
Baseline (Actual Results)	100%
UBS	66%
JP Morgan	59%
Goldman Sachs	53%
Simon Kuper and Stefan Szymanski (Wired Mag)	41%


The World Cup predicted

The World Cup and Economics 2010

Goldman Sachs

J.P.Morgan

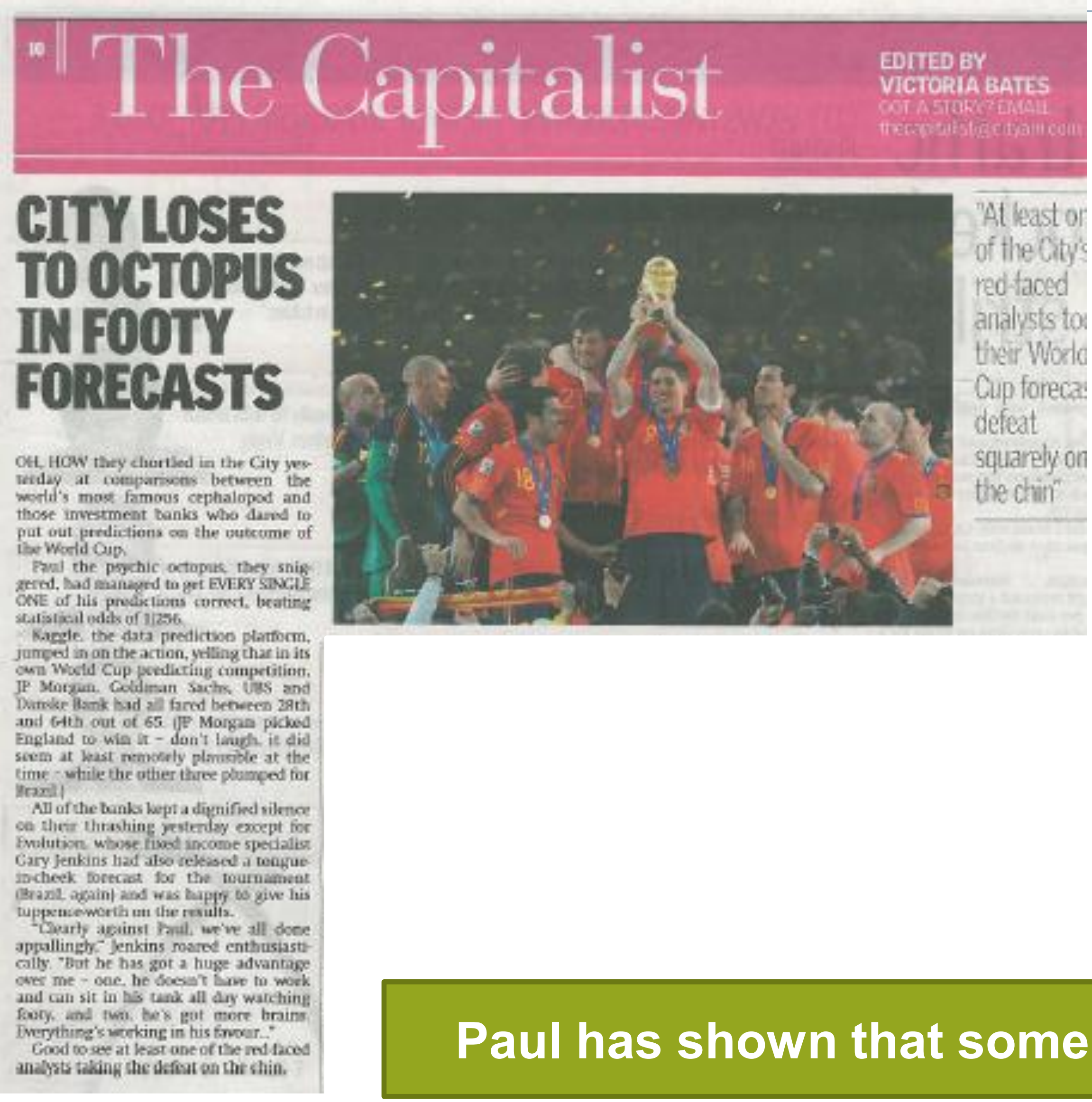
England to Win the World Cup!

A Quantitative Guide to the 2010 World Cup

Betfair was better than Frontier Economic, Goldman or JP Morgan!

http://db.riskwaters.com/global/actuary/digital/Actuary_0610%20BeckerKainth.pdf http://worldcupeconomics.blogspot.com/

There is only downside potential… unless you cover your bets…



10 The Capitalist

EDITED BY VICTORIA BATES
GOT A STORY? EMAIL
thecapitalist@cityam.com

CITY LOSES TO OCTOPUS IN FOOTY FORECASTS

"At least or
of the City's
red-faced
analysts to
their World
Cup forecas
defeat
squarely on
the chin"

OH, HOW they chortled in the City yesterday at comparisons between the world's most famous cephalopod and those investment banks who dared to put out predictions on the outcome of the World Cup.

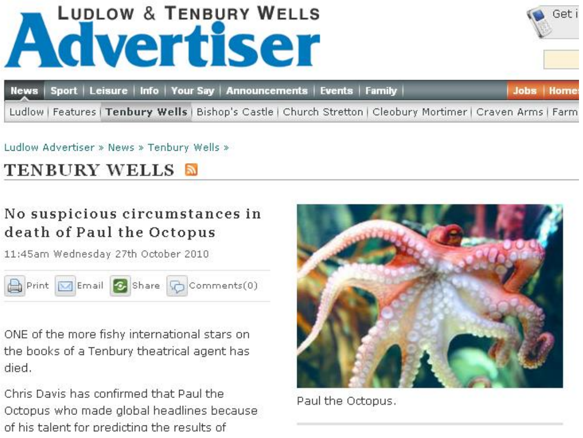
Paul the psychic octopus, they sniggered, had managed to get EVERY SINGLE ONE of his predictions correct, beating statistical odds of 1/256.

Kaggle, the data prediction platform, jumped in on the action, yelling that in its own World Cup predicting competition, JP Morgan, Goldman Sachs, UBS and Danske Bank had all fared between 28th and 64th out of 65. (JP Morgan picked England to win it – don't laugh, it did seem at least remotely plausible at the time – while the other three plumped for Brazil.)

All of the banks kept a dignified silence on their thrashing yesterday except for Evolution, whose fixed income specialist Gary Jenkins had also released a tongue-in-cheek forecast for the tournament (Brazil, again) and was happy to give his tuppence-worth on the results.

"Clearly against Paul, we've all done appallingly," Jenkins roared enthusiastically. "But he has got a huge advantage over me – one, he doesn't have to work and can sit in his tank all day watching footy, and two, he's got more brains. Everything's working in his favour..."

Good to see at least one of the red-faced analysts taking the defeat on the chin.



Get i
News | Sport | Leisure | Info | Your Say | Announcements | Events | Family | Jobs | Home
Ludlow | Features | Tenbury Wells | Bishop's Castle | Church Stretton | Cleobury Mortimer | Craven Arms | Farm
Print
Email
Share
Comments(0)

LUDLOW & TENBURY WELLS **Advertiser**

Ludlow Advertiser » News » Tenbury Wells »

TENBURY WELLS

No suspicious circumstances in death of Paul the Octopus

11:45am Wednesday 27th October 2010

ONE of the more fishy international stars on the books of a Tenbury theatrical agent has died.

Chris Davis has confirmed that Paul the Octopus who made global headlines because of his talent for predicting the results of

Paul the Octopus.

Paul has shown that some things are still certain though…

Analysis by David Forrest and Robert Simmons into tipsters, betting odds and statistical models shows…

- Statistical models do better than tipsters
- Combining tipster estimates can be better than using an individual tipster
- Betting odds are the best – maybe because the bookies pay the biggest salaries?

feature

Read all about it?

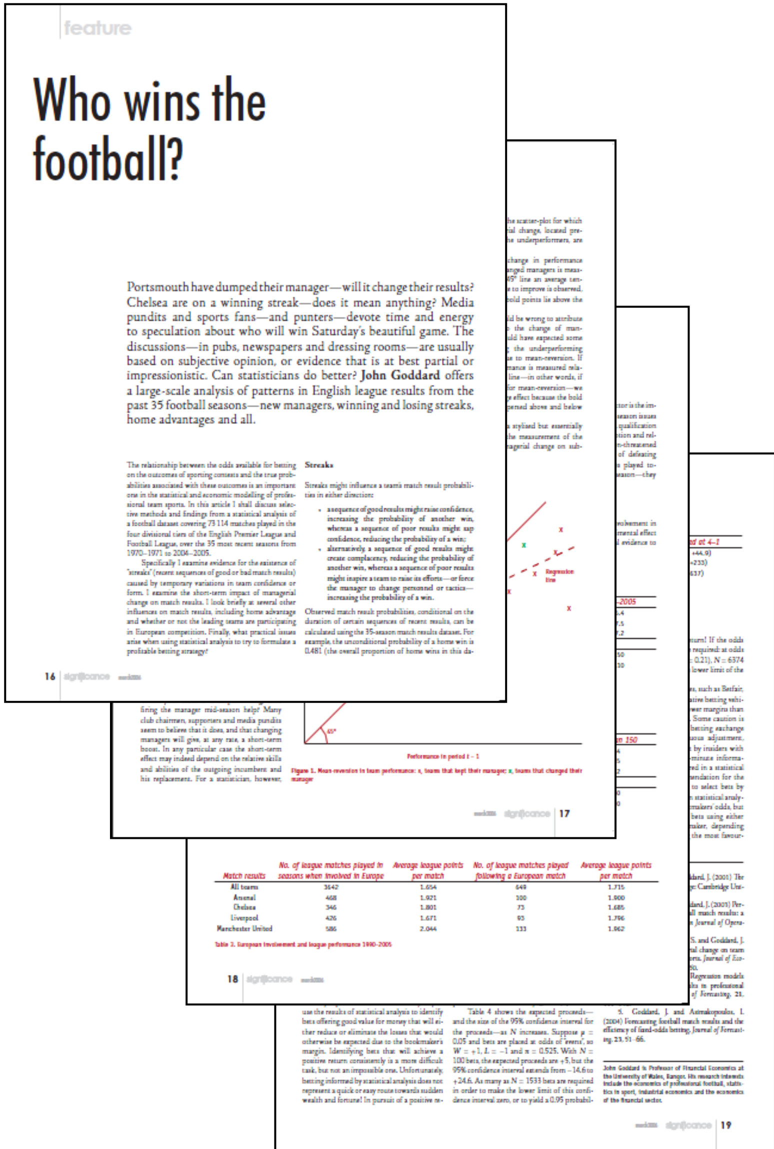
If statistics are of little help in predicting individual football results, are the professional followers of the sport any better? **David Forrest** and **Robert Simmons** analyse the performance of newspaper tipsters in helping the gambler to beat the bookies.

But they conclude that the betting exchanges involve people working with more complex models and more data – so don't bet against them!

Forrest D, Simmons R 2006 "Read all about it" RSS Significance p151-153 Volume 3 Issue 1 March 2006

John Goddard has looked at ‘streaks’,‘firing a manager’, ‘fighting for survival’, the effect of ‘playing in Europe’…

- Winning teams keep on winning
  - Confidence?
- Firing a manager doesn’t seem to help a team if one controls for ‘mean reversion’
- Home advantage has an effect, and while it has been decreasing over the most recent 35 seasons, the home ground advantage is larger when the away team has to travel further
- Relegation threatened teams “fighting for survival” are more likely to beat their mid-table rivals at the end of the season than before
- The “playing in Europe” effect has not been shown to be a hindrance, although it has been for some top teams



feature

Who wins the football?

Portsmouth have dumped their manager—will it change their results? Chelsea are on a winning streak—does it mean anything? Media pundits and sports fans—and punters—devote time and energy to speculation about who will win Saturday’s beautiful game. The discussions—in pubs, newspapers and dressing rooms—are usually based on subjective opinion, or evidence that is at best partial or impressionistic. Can statisticians do better? **John Goddard** offers a large-scale analysis of patterns in English league results from the past 35 football seasons—new managers, winning and losing streaks, home advantages and all.

Goddard J 2006 “Who wins the football” RSS Significance p16-19 Volume 3 Issue 1 March 2006

David Spiegelhalter and Yin-Lam Ng have shown that statistical models can outperform sports commentators

- They modelled a single round of matches
- Their model had similarities
  - They also used a Poisson approach to estimate ‘number of goals scored/conceded’
  - They also looked at a result as being a combination of scoring/conceding
  - They developed a way for teams to interact

feature

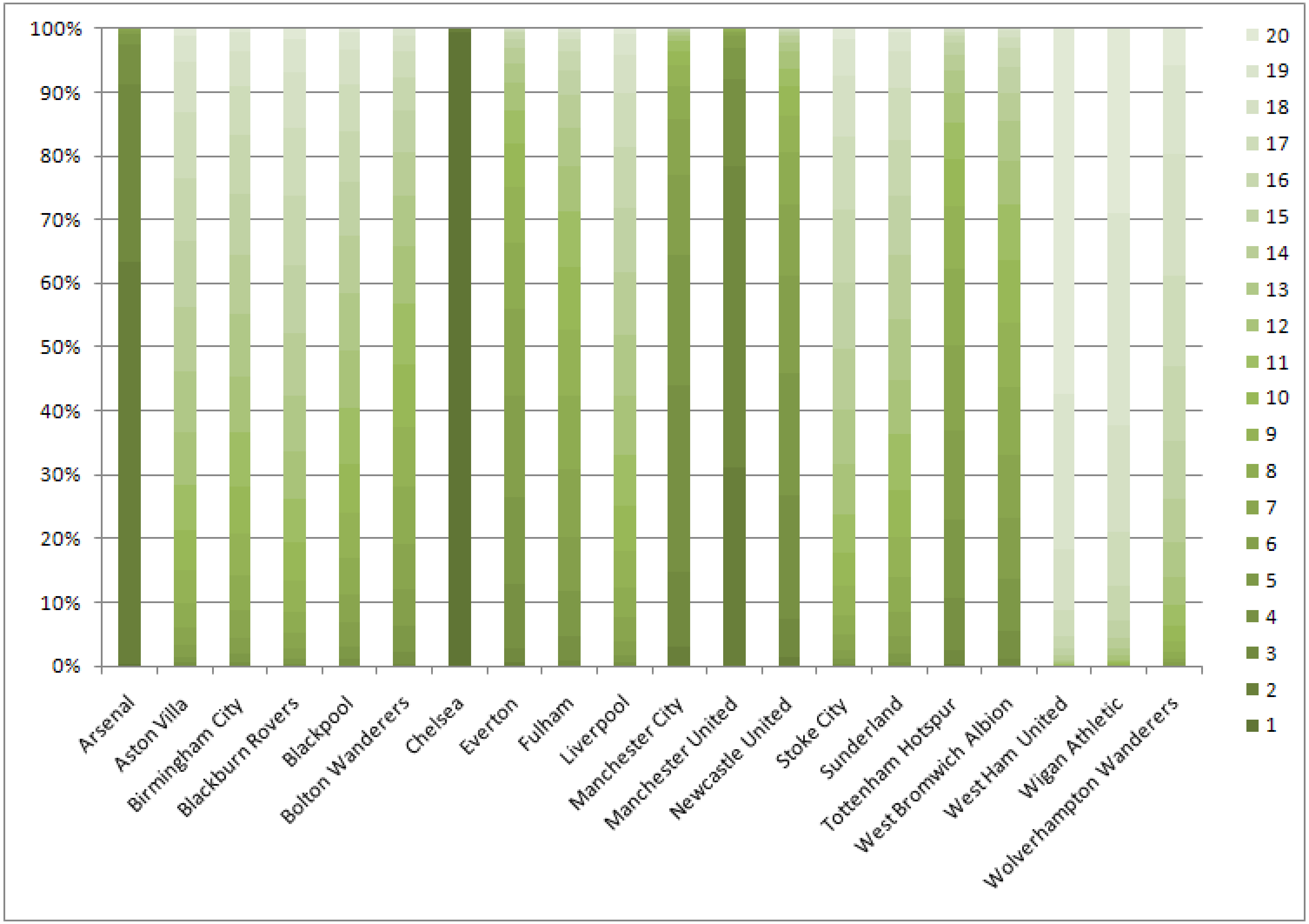
One match to go!

Can statistics really predict the results of Premier League football matches? **David Spiegelhalter** and **Yin-Lam Ng** put their skills to the test, and their reputations on the line—and scored!

Spiegelhalter D, Ng Y-L 2009 “One match to go” - RSS Significance p151-153 Volume 6 Issue 4 December 2009

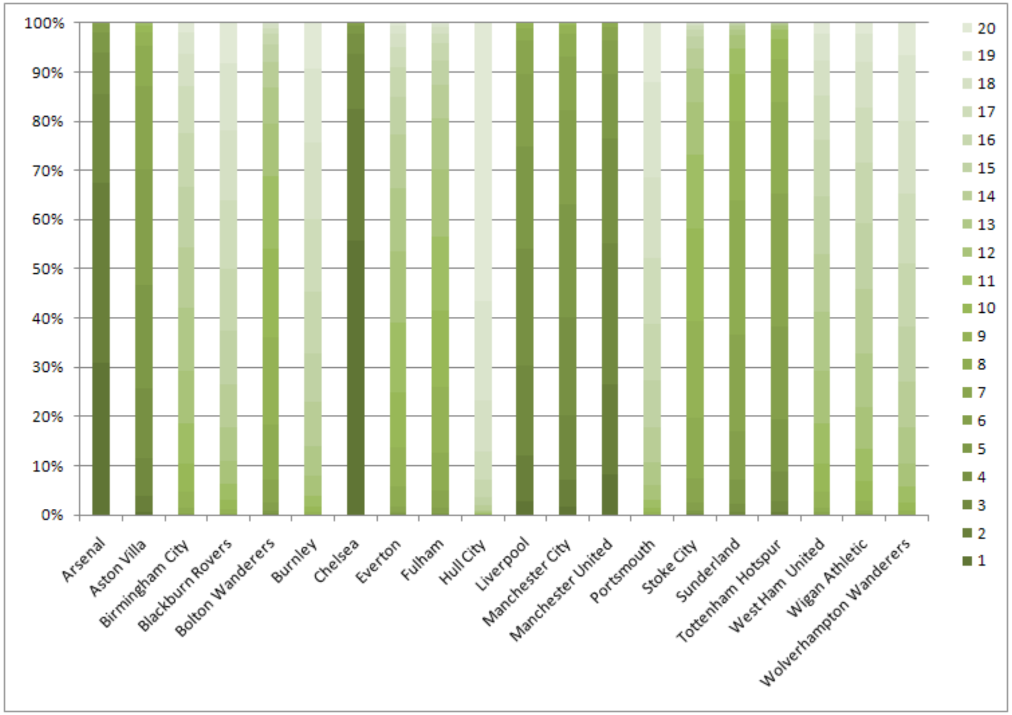
Finally, we get to the model predictions!

The slide you have been waiting for:
The darker, the more likely the team will win (2010-11)



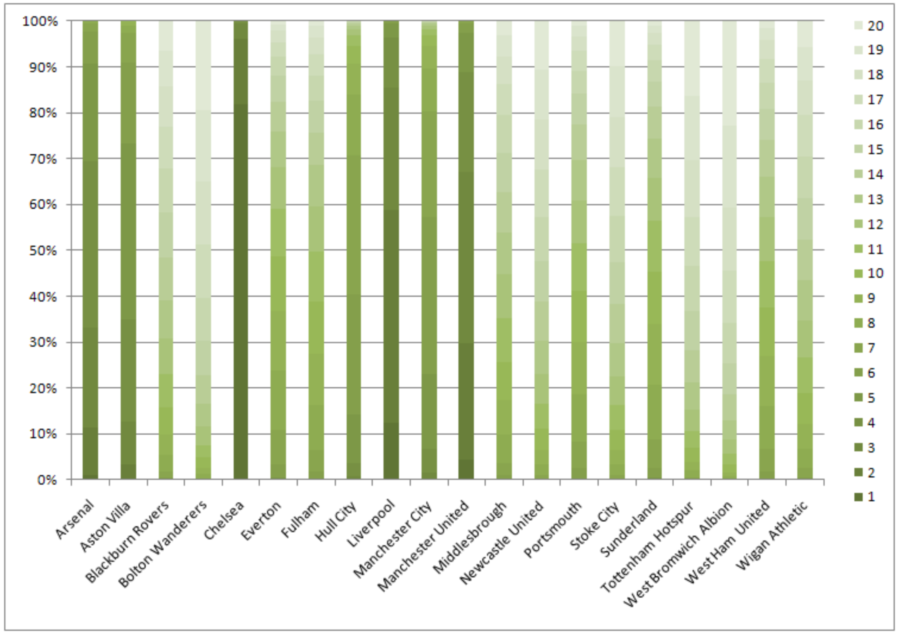
http://www.premierleague.com/page/FixturesResults/0,,12306,00.html http://www.abc.net.au/sport/stories/2010/08/11/2979870.htm

The 2009-10 predictions after 100 games predicted that Chelsea was the most likely team to win



The model predicts that Chelsea is the most likely team to win – and they did!

The predictions at this stage of the 2008-2009 season had Chelsea as likely winners



BARCLAYS PREMIER LEAGUE CHAMPIONS
BARCLAYS PREMIER LEAGUE CHAMPIONS 2008-09

Both Chelsea and Liverpool were in a better position. Manchester United was forecast to win with a 4% probability, and finish in the top 3 more than 2 out of 3 times

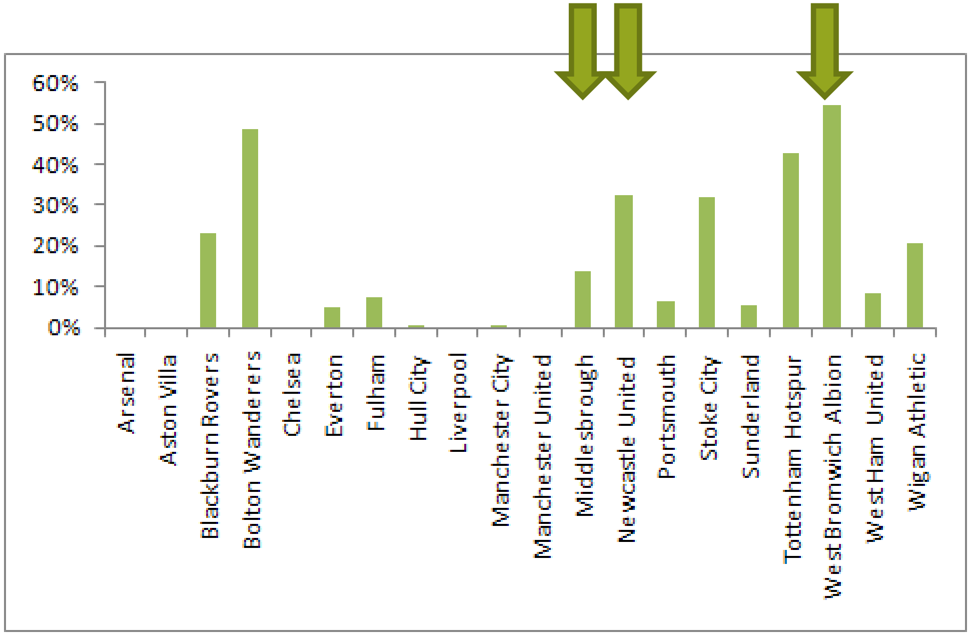
http://tkey.co.uk/man-united-premier-league-champions-200809 http://www.premierleague.com/page/FixturesResults/0,,12306,00.html http://www.abc.net.au/sport/stories/2010/08/11/2979870.htm

This season, the premier league winner is getting almost as much press as the relegation zone

How good is the model at predicting the teams to be relegated at this stage of the season?

2008/2009

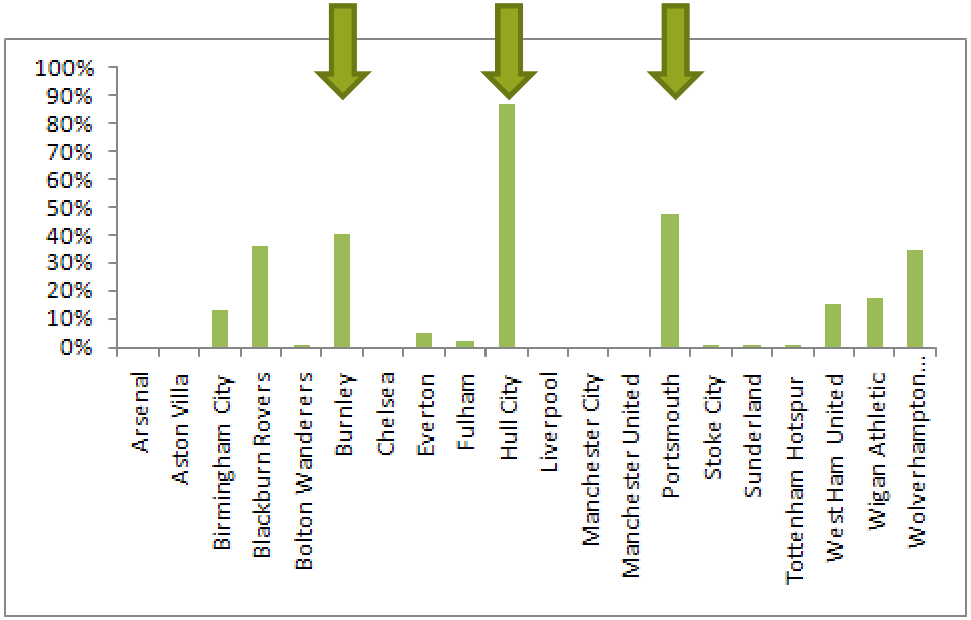
- Who was relegated? (Estimated probability of going down at this stage of the season in brackets)
  - Middlesborough (14%)
  - Newcastle United (32%)
  - West Bromwich Albion (54%)
- Who was predicted to be relegated after 100 games?
  - The model predicted that many teams had a good chance of going down



- The three that did, were all in the 8 most likely to go down

This season, the premier league winner is getting almost as much press as the relegation zone

How good is the model at predicting the teams to be relegated at this stage of the season?



2009/2010

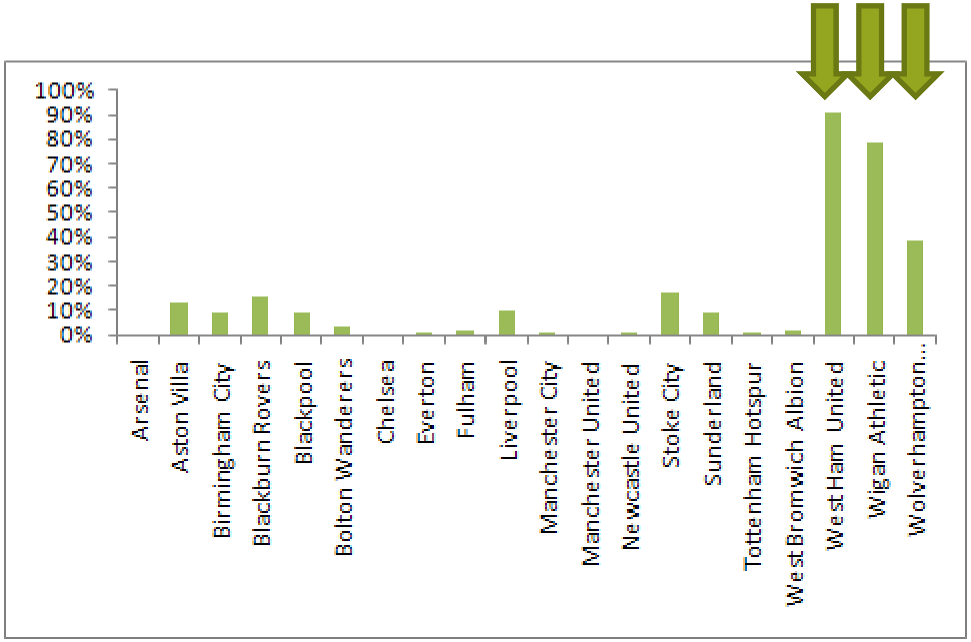
- Who was relegated? (Estimated probability of going down at this stage of the season in brackets)
  - Burnley (40%)
  - Hull (87%)
  - Portsmouth (48%)
- Who was predicted to be relegated after 100 games?
  - The three forecast to have the greatest probability of going down, went down!

So the model worked very well for the last season, but not so well for the one before. How about 2010/11?

Who is forecast to get relegated? Something beginning with… W

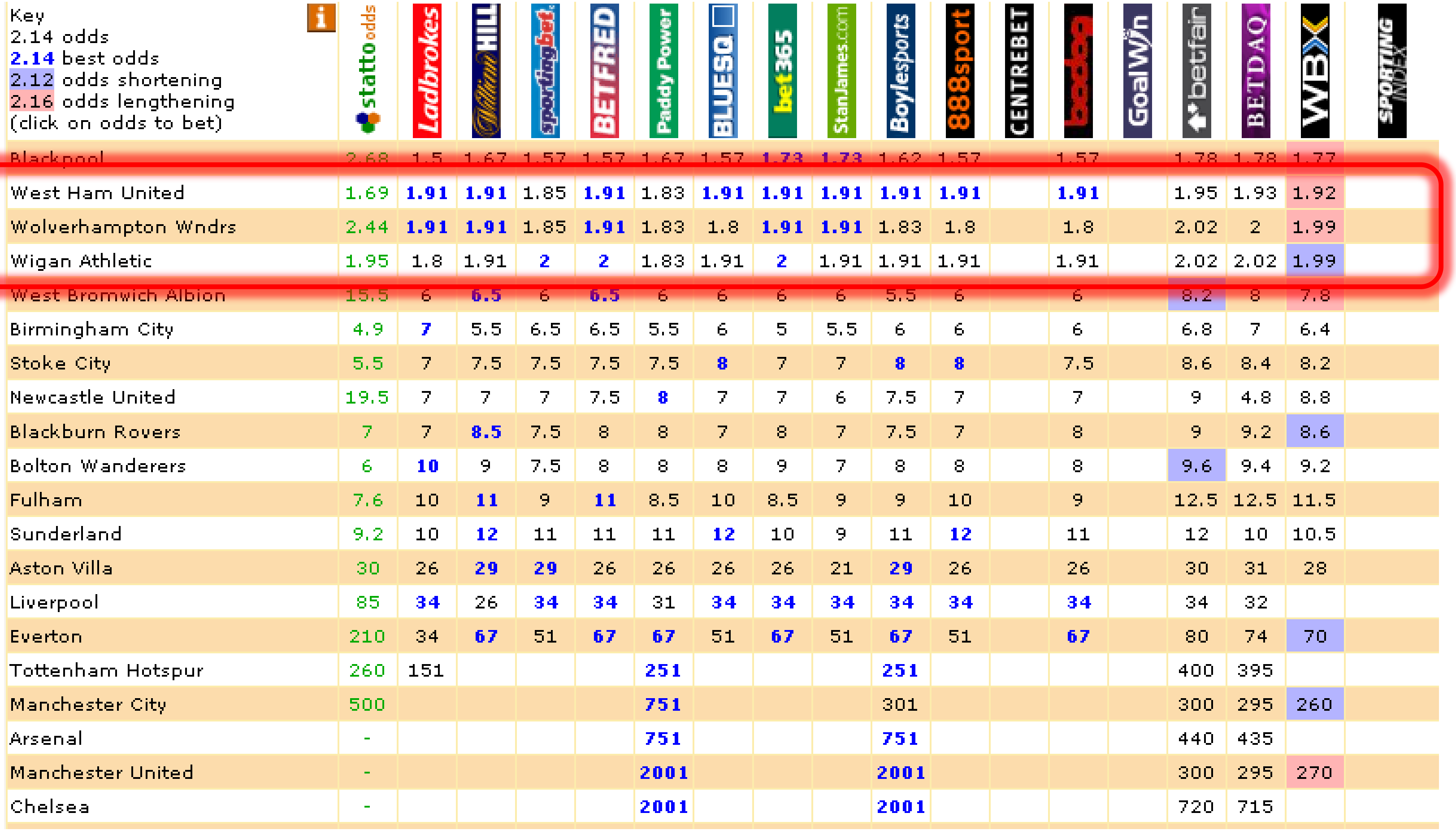
2010-11

- Who is being predicted to be relegated?
  - West Ham United (91%)
  - Wigan Athletic (79%)
  - Wolverhampton Wanderers (39%)



http://www.premierleague.com/page/FixturesResults/0,,12306,00.html http://www.abc.net.au/sport/stories/2010/08/11/2979870.htm

Interestingly, the betting stats agree – the three W's are in trouble!

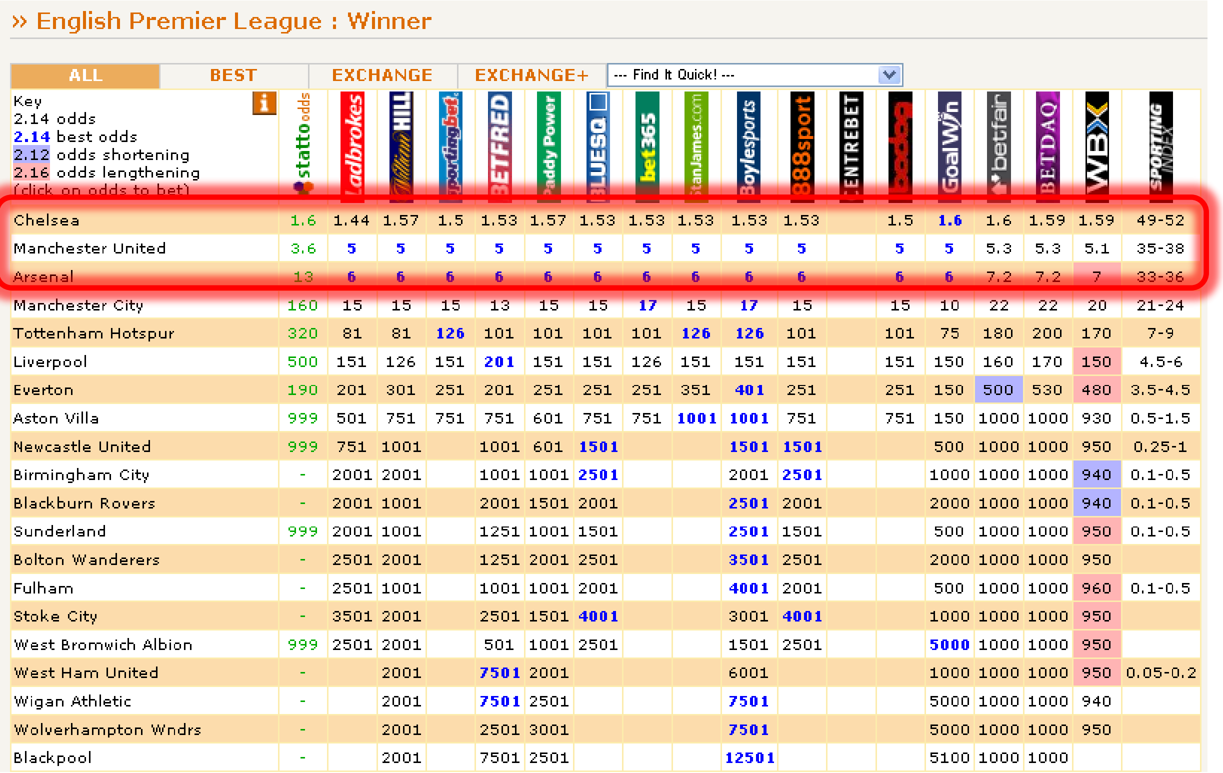


Key
2.14 odds
2.14 best odds
2.12 odds shortening
2.16 odds lengthening
(click on odds to bet)

	statto odds	Ladbrokes	William Hill	sportingbet	BETFRED	Paddy Power	BLUESQ	bet365	StanJames.com	Boylesports	888sport	CENTREBET	bodog	GoalWin	betfair	BETDAQ	WBX	SPORTING INDEX
Blackpool	2.68	1.5	1.67	1.57	1.57	1.67	1.57	**1.73**	**1.73**	1.62	1.57		1.57		1.78	1.78	1.77	
West Ham United	1.69	**1.91**	**1.91**	1.85	**1.91**	1.83	**1.91**	**1.91**	**1.91**	**1.91**	**1.91**		**1.91**		1.95	1.93	1.92	
Wolverhampton Wndrs	2.44	**1.91**	**1.91**	1.85	**1.91**	1.83	1.8	**1.91**	**1.91**	1.83	1.8		1.8		2.02	2	1.99	
Wigan Athletic	1.95	1.8	1.91	**2**	**2**	1.83	1.91	**2**	1.91	1.91	1.91		1.91		2.02	2.02	1.99	
West Bromwich Albion	15.5	6	**6.5**	6	**6.5**	6	6	6	6	5.5	6		6		8.2	8	7.8	
Birmingham City	4.9	**7**	5.5	6.5	6.5	5.5	6	5	5.5	6	6		6		6.8	7	6.4	
Stoke City	5.5	7	7.5	7.5	7.5	7.5	**8**	7	7	**8**	**8**		7.5		8.6	8.4	8.2	
Newcastle United	19.5	7	7	7	7.5	**8**	7	7	6	7.5	7		7		9	4.8	8.8	
Blackburn Rovers	7	7	**8.5**	7.5	8	8	7	8	7	7.5	7		8		9	9.2	8.6	
Bolton Wanderers	6	**10**	9	7.5	8	8	8	9	7	8	8		8		9.6	9.4	9.2	
Fulham	7.6	10	**11**	9	**11**	8.5	10	8.5	9	9	10		9		12.5	12.5	11.5	
Sunderland	9.2	10	**12**	11	11	11	**12**	10	9	11	**12**		11		12	10	10.5	
Aston Villa	30	26	**29**	**29**	26	26	26	26	21	**29**	26		26		30	31	28	
Liverpool	85	**34**	26	**34**	**34**	31	**34**	**34**	**34**	**34**	**34**		**34**		34	32		
Everton	210	34	**67**	51	**67**	**67**	51	**67**	51	**67**	51		**67**		80	74	70	
Tottenham Hotspur	260	151				**251**				**251**					400	395		
Manchester City	500					**751**				301					300	295	260	
Arsenal	-					**751**				**751**					440	435		
Manchester United	-					**2001**				**2001**					300	295	270	
Chelsea	-					**2001**				**2001**					720	715		

http://www.statto.com/football/odds/england/premier-league/relegation

Betting odds have Chelsea as the 2010/11 favourite – with the odd's reflecting a 63-69% chance of winning the league



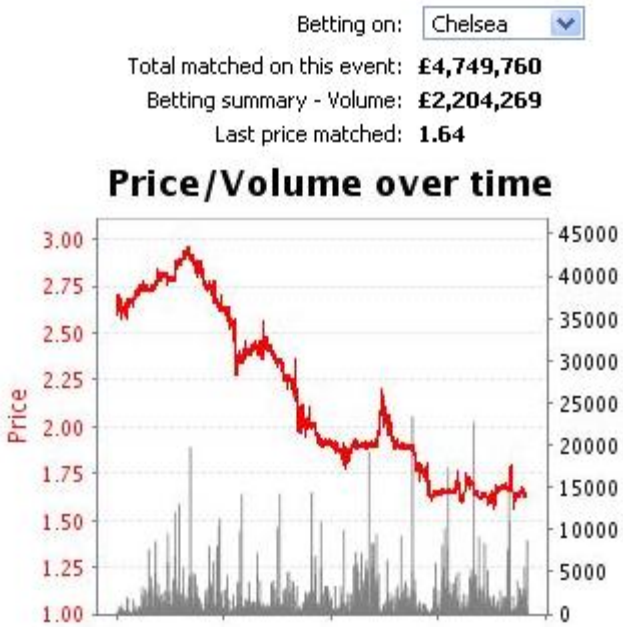
» English Premier League : Winner

Key
2.14 odds
2.14 best odds
2.12 odds shortening
2.16 odds lengthening
(click on odds to bet)

	statto odds	Ladbrokes	William Hill	sportingbet	BETFRED	Paddy Power	BLUESQ	bet365	StanJames.com	Boylesports	888sport	CENTREBET	bodog	GoalWin	betfair	BETDAQ	WBX	SPORTING INDEX
Chelsea	1.6	1.44	1.57	1.5	1.53	1.57	1.53	1.53	1.53	1.53	1.53		1.5	**1.6**	1.6	1.59	1.59	49-52
Manchester United	3.6	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**		**5**	**5**	5.3	5.3	5.1	35-38
Arsenal	13	**6**	**6**	**6**	**6**	**6**	**6**	**6**	**6**	**6**	**6**		**6**	**6**	7.2	7.2	7	33-36
Manchester City	160	15	15	15	13	15	15	**17**	15	**17**	15		15	10	22	22	20	21-24
Tottenham Hotspur	320	81	81	**126**	101	101	101	101	**126**	**126**	101		101	75	180	200	170	7-9
Liverpool	500	151	126	151	**201**	151	151	126	151	151	151		151	150	160	170	150	4.5-6
Everton	190	201	301	251	201	251	251	251	351	**401**	251		251	150	500	530	480	3.5-4.5
Aston Villa	999	501	751	751	751	601	751	751	**1001**	**1001**	751		751	150	1000	1000	930	0.5-1.5
Newcastle United	999	751	1001		1001	601	**1501**			**1501**	**1501**			500	1000	1000	950	0.25-1
Birmingham City	-	2001	2001		1001	1001	**2501**			2001	**2501**			1000	1000	1000	940	0.1-0.5
Blackburn Rovers	-	2001	1001		2001	1501	2001			**2501**	2001			2000	1000	1000	940	0.1-0.5
Sunderland	999	2001	1001		1251	1001	1501			**2501**	1501			500	1000	1000	950	0.1-0.5
Bolton Wanderers	-	2501	2001		1251	2001	2501			**3501**	2501			2000	1000	1000	950	
Fulham	-	2501	1001		1001	1001	2001			**4001**	2001			500	1000	1000	960	0.1-0.5
Stoke City	-	3501	2001		2501	1501	**4001**			3001	**4001**			1000	1000	1000	950	
West Bromwich Albion	999	2501	2001		501	1001	2501			1501	2501			**5000**	1000	1000	950	
West Ham United	-		2001		**7501**	2001				6001				1000	1000	1000	950	0.05-0.2
Wigan Athletic	-		2001		**7501**	2501				**7501**				5000	1000	1000	940	
Wolverhampton Wndrs	-		2001		2501	3001				**7501**				5000	1000	1000	950	
Blackpool	-		2001		7501	2501				**12501**				5100	1000	1000		

Are those betting taking into account more factors? Are those betting less rational?

- Manchester United is second most likely to win, and Arsenal third
- Our model has the same top 3, but reflects a belief that Arsenal has a better chance of finishing top
- The odds reflected by the betting stats are far less supportive of a Chelsea victory than our model

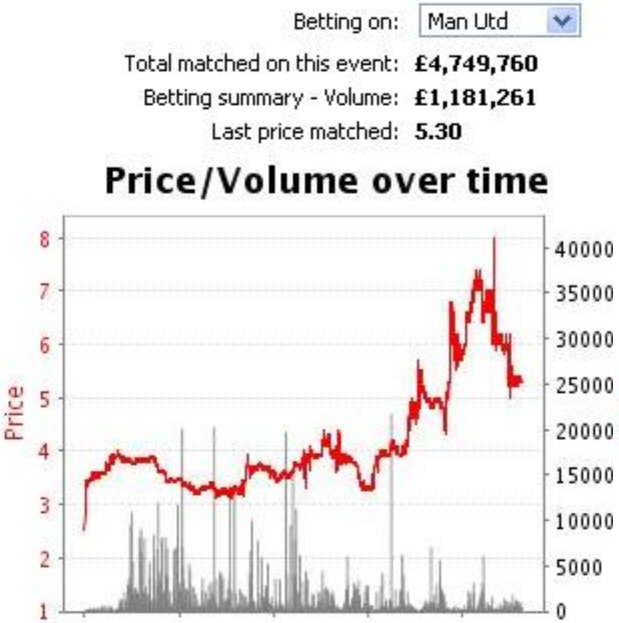


Betting on: Chelsea

Total matched on this event: **£4,749,760**

Betting summary - Volume: **£2,204,269**

Last price matched: **1.64**

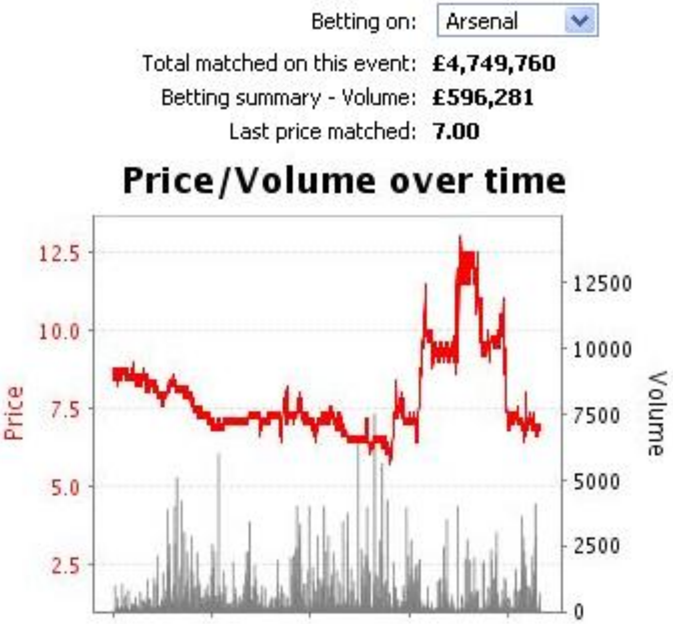


Betting on: Man Utd

Total matched on this event: **£4,749,760**

Betting summary - Volume: **£1,181,261**

Last price matched: **5.30**

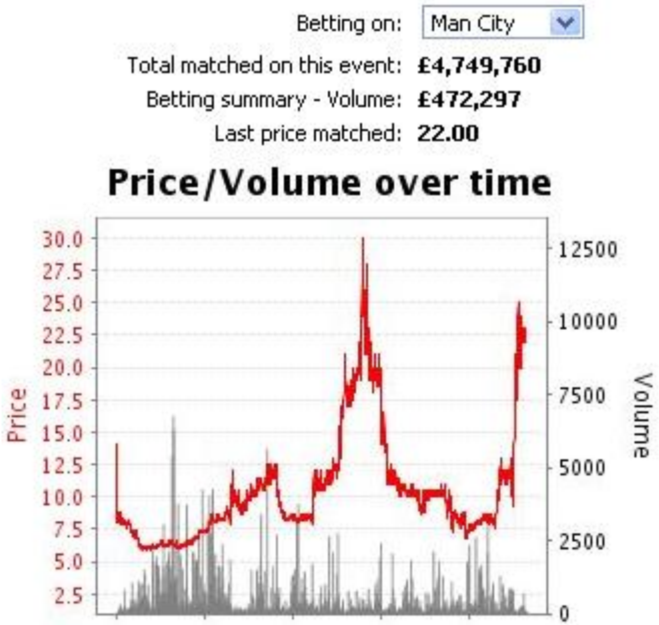


Betting on: Arsenal

Total matched on this event: **£4,749,760**

Betting summary - Volume: **£596,281**

Last price matched: **7.00**

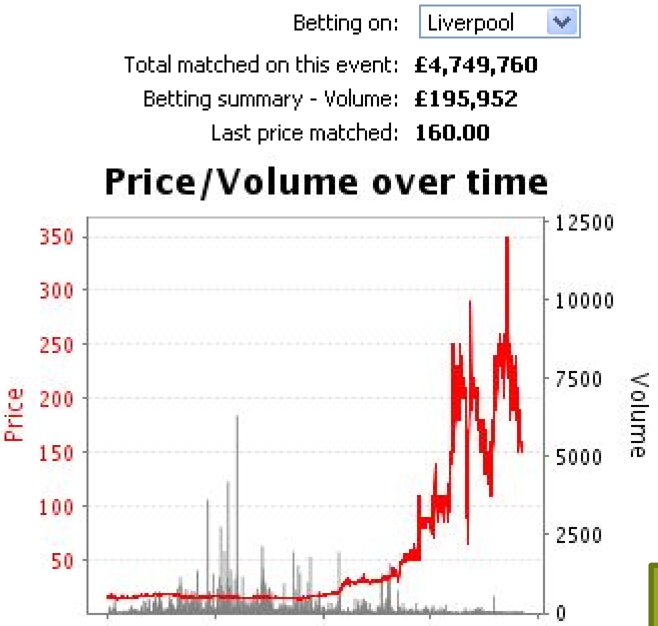


Betting on: Man City

Total matched on this event: **£4,749,760**

Betting summary - Volume: **£472,297**

Last price matched: **22.00**

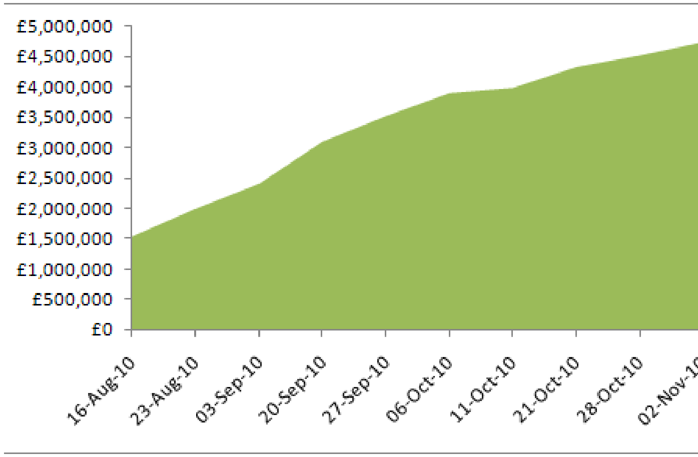


Betting on: Liverpool

Total matched on this event: **£4,749,760**

Betting summary - Volume: **£195,952**

Last price matched: **160.00**

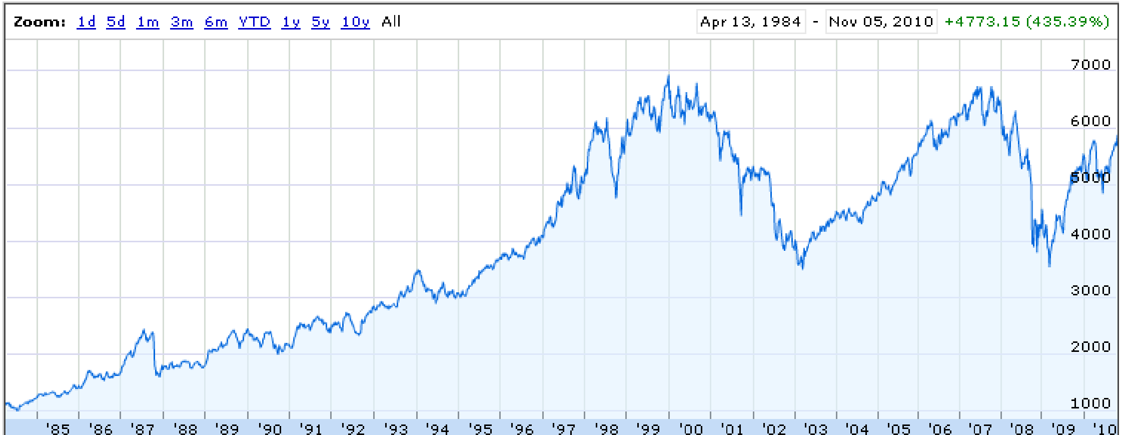


The odds are continuously changing

http://www.betfair.com

Bringing it back to our day jobs…

- Be aware of your model limitations
- Just because it works for the past – doesn't mean it'll work for the future
- The known unknowns, and the unknown unknowns
  - Otherwise, someone could get injured?
- Model predictions should be interpreted using actuarial judgement



Is the efficient market hypothesis strong or weak?

Questions or comments? (No curve balls please…)

Expressions of individual views by members of The Actuarial Profession and its staff are encouraged.

The views expressed in this presentation are those of the presenter.

http://www.youtube.com/watch?v=aC-f9IL7aA4&feature=related