

## PL Fixtures - (Betfair) Odds & (Implied) Probabilities

**Skip to pages 3 and 4 to see the results.**

```
library(XML);
library(xtable);
library(knitr);
```

Reading Data from <http://www.betfair.com/exchange/football/competition?id=31>

```
matches_URL <- "https://www.betfair.com/exchange/football/competition?id=31"

# For some reason, Betfair has changed its design recently.
# It's not a table anymore
# matches_Table <- readHTMLTable(matches_URL)

# ReadLine and Parse the HTML page.

# download.file(matches_URL, "aa.html")
matches_html <- readLines("aa.html")

# matches_html <- readLines(matches_URL)

matches_parse <- htmlTreeParse(matches_html,useInternal=TRUE)
```

## Cleaning Data:

```
# Extract the relevant bits.
data_odds_back <- xpathSApply(matches_parse,
                               "//*[button[@class = 'bet-button back cta cta-back i13n-ltxt-FltBetSlpB i13n-SlpB i13n-SlpL i13n-SlpR]]",
                               , xmlValue)

data_odds_lay <- xpathSApply(matches_parse,
                              "//*[button[@class = 'bet-button lay cta cta-lay i13n-ltxt-FltBetSlpL i13n-SlpB i13n-SlpR]]",
                              , xmlValue)

data_home <- xpathSApply(matches_parse, "//*[span[@class = 'home-team']]", xmlValue)
data_away <- xpathSApply(matches_parse, "//*[span[@class = 'away-team']]", xmlValue)

# Otherwise team names would be interpreted as factors.
options(stringsAsFactors = FALSE)

# make "odds" numeric

data_back <- data.frame(
  apply(
    matrix(data_odds_back, ncol = 3, byrow = TRUE)
    , 2, as.numeric))
```

```

data_lay <- data.frame(
  apply(
    matrix(data_odds_lay, ncol = 3, byrow = TRUE)
    , 2, as.numeric))

# Matches data.frame
all_matches <- cbind(data_home, data_away, data_back, data_lay)
colnames(all_matches) <-
  c("Home", "Away", "H_B", "D_B", "A_B", "H_L", "D_L", "A_L")

```

Creating probabilities data.frame (a rough estimate + normalisation). The results are reported with 0 decimal points.

```

# Output data.frames

H <-
  round((100/all_matches[,3]+ 100/all_matches[,6])/rowSums(1/all_matches[,3:8])
    , digits = 0)
D <-
  round((100/all_matches[,4]+ 100/all_matches[,7])/rowSums(1/all_matches[,3:8])
    , digits = 0)
A <-
  round((100/all_matches[,5]+ 100/all_matches[,8])/rowSums(1/all_matches[,3:8])
    , digits = 0)

prob_output <- data.frame(
  "Home" = all_matches[,1], H, D, A, "Away" = all_matches[,2])

odds_output <- data.frame(cbind(
  "Home" = all_matches[,1],
  H = paste(all_matches[,3], all_matches[,6], sep = "/"),
  D = paste(all_matches[,4], all_matches[,7], sep = "/"),
  A = paste(all_matches[,5], all_matches[,8], sep = "/"),
  "Away" = all_matches[,2])
)

odds_output <- odds_output[1:10, ]

prob_output <- prob_output[1:10, ]

prob_output <-
  prob_output[order(apply(prob_output[,2:4],1, max)),]

```

Home	H	D	A	Away
Everton	35	30	35	Southampton
Sunderland	38	31	31	Newcastle
Leicester	45	27	28	West Ham
Burnley	28	27	45	Tottenham
Swansea	51	29	21	Hull
Arsenal	53	26	21	Liverpool
West Brom	56	27	18	QPR
C Palace	18	24	59	Man City
Man Utd	74	17	8	Aston Villa
Chelsea	77	16	7	Stoke

Table 1: Coming Fixtures (Implied) Probabilities

	Home	H	D	A	Away
1	Arsenal	1.88/1.9	3.85/3.9	4.7/4.8	Liverpool
2	Everton	2.82/2.84	3.3/3.35	2.86/2.9	Southampton
3	Leicester	2.22/2.24	3.65/3.7	3.6/3.65	West Ham
4	Man Utd	1.34/1.35	5.7/5.8	12/12.5	Aston Villa
5	Swansea	1.97/1.98	3.45/3.5	4.8/4.9	Hull
6	West Brom	1.8/1.81	3.7/3.8	5.6/5.7	QPR
7	Chelsea	1.29/1.3	6.2/6.4	14.5/15	Stoke
8	Burnley	3.55/3.6	3.65/3.7	2.24/2.26	Tottenham
9	Sunderland	2.62/2.64	3.25/3.3	3.2/3.25	Newcastle
10	C Palace	5.6/5.8	4.2/4.3	1.7/1.71	Man City

Table 2: Coming Fixtures Odds