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

Skip to pages 3 and 4 to see the results.

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

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

```r
matches_URL <- "http://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.
matches_html  <- readLines(matches_URL)
matches_parse <- htmlTreeParse(matches_html,useInternal=TRUE)
```

Cleaning Data:

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

data_odds_lay <- xpathSApply(matches_parse,
                            "//button[@class = 'bet-button lay cta cta-lay i13n-ltxt-FltBetSlpL i13n-S
                            , 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))
```

```r
# 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.

```r
# 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(cbind(
    "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])
    )

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

	Home	H	D	A	Away
2	Burnley	38	30	32	C Palace
3	Leicester	38	30	32	Stoke
7	Newcastle	31	29	40	Southampton
1	Aston Villa	21	28	52	Liverpool
10	Everton	53	27	20	West Brom
9	Man City	54	24	21	Arsenal
8	West Ham	55	27	18	Hull
4	QPR	16	24	60	Man Utd
6	Tottenham	61	23	15	Sunderland
5	Swansea	14	23	63	Chelsea

Table 1: Coming Fixtures Odds

	Home	H	D	A	Away
1	Aston Villa	4.8/4.9	3.6/3.65	1.93/1.94	Liverpool
2	Burnley	2.62/2.64	3.3/3.35	3.1/3.15	C Palace
3	Leicester	2.62/2.64	3.3/3.35	3.1/3.15	Stoke
4	QPR	6.2/6.4	4.2/4.3	1.65/1.67	Man Utd
5	Swansea	7.2/7.6	4.3/4.4	1.57/1.59	Chelsea
6	Tottenham	1.62/1.63	4.2/4.3	6.4/6.6	Sunderland
7	Newcastle	3.2/3.25	3.4/3.45	2.48/2.5	Southampton
8	West Ham	1.8/1.82	3.7/3.75	5.4/5.6	Hull
9	Man City	1.83/1.84	4.1/4.2	4.6/4.7	Arsenal
10	Everton	1.87/1.88	3.7/3.75	4.9/5	West Brom

Table 2: Coming Fixtures (Implied) Probabilities