PREMIER LEAGUE ANALYSIS & PREDICTION USING BAYESIAN

DATA SOURCE – ENGLISH PREMIER LEAGUE OFFICIAL WEBSITE

PREMIER LEAGUE PARTICIPATION SINCE 1992–93

Seasons in the Premier League

PREMIER LEAGUE POINTS TOTAL SINCE 1992–93
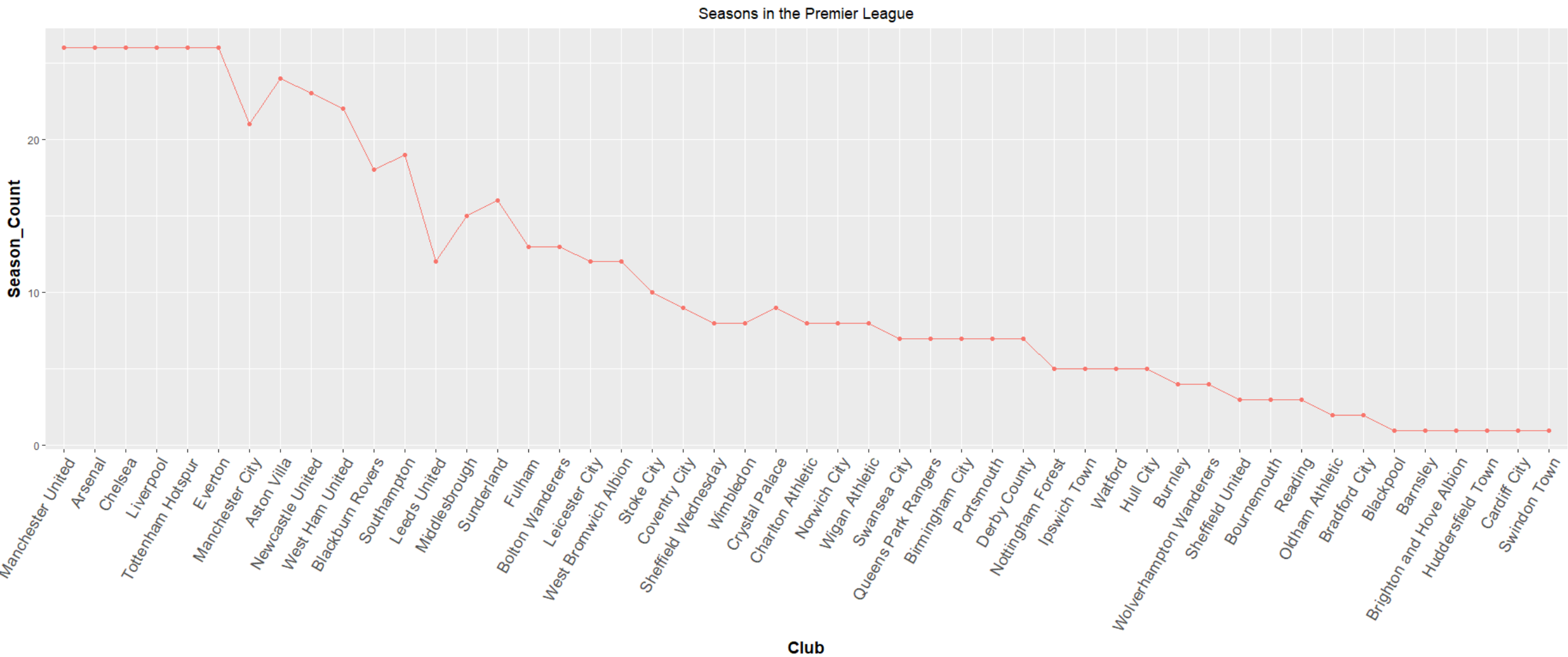
Points Distribution
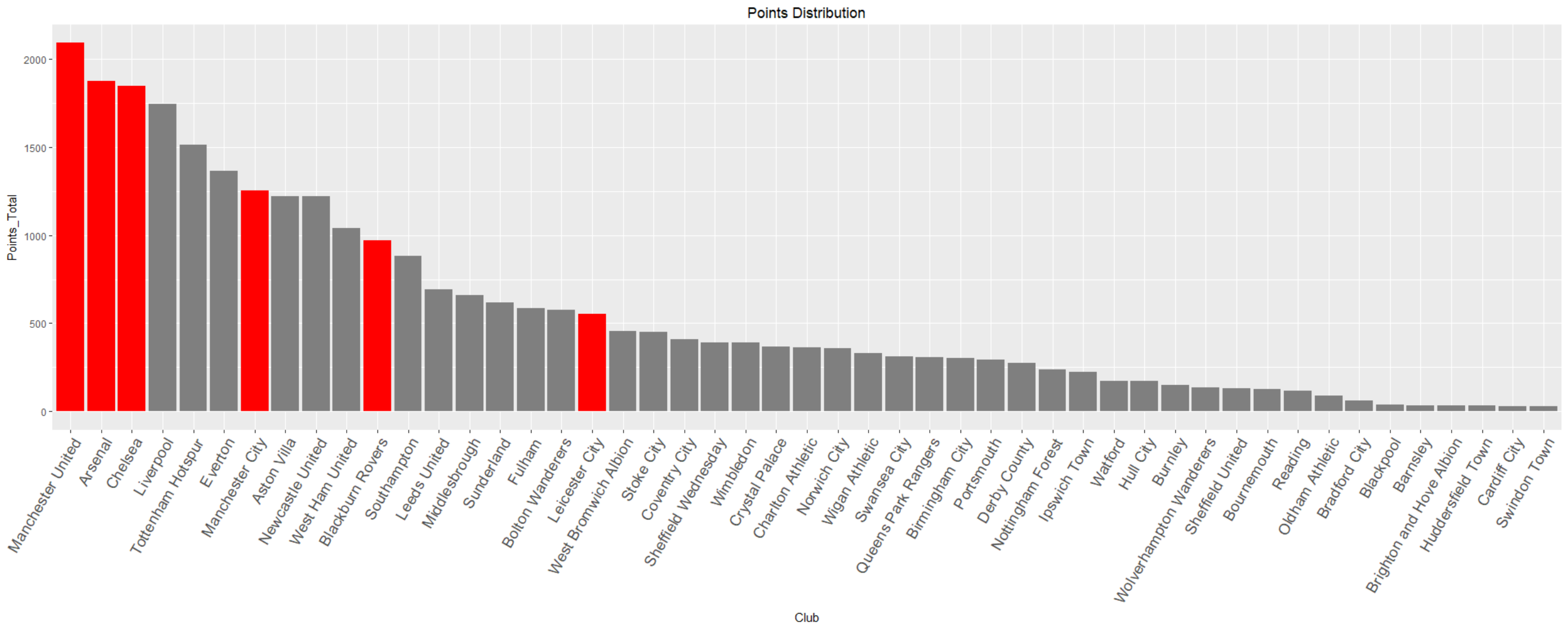
PREMIER LEAGUE TOP 6 GOALS DISTRIBUTION

Goals Distribution
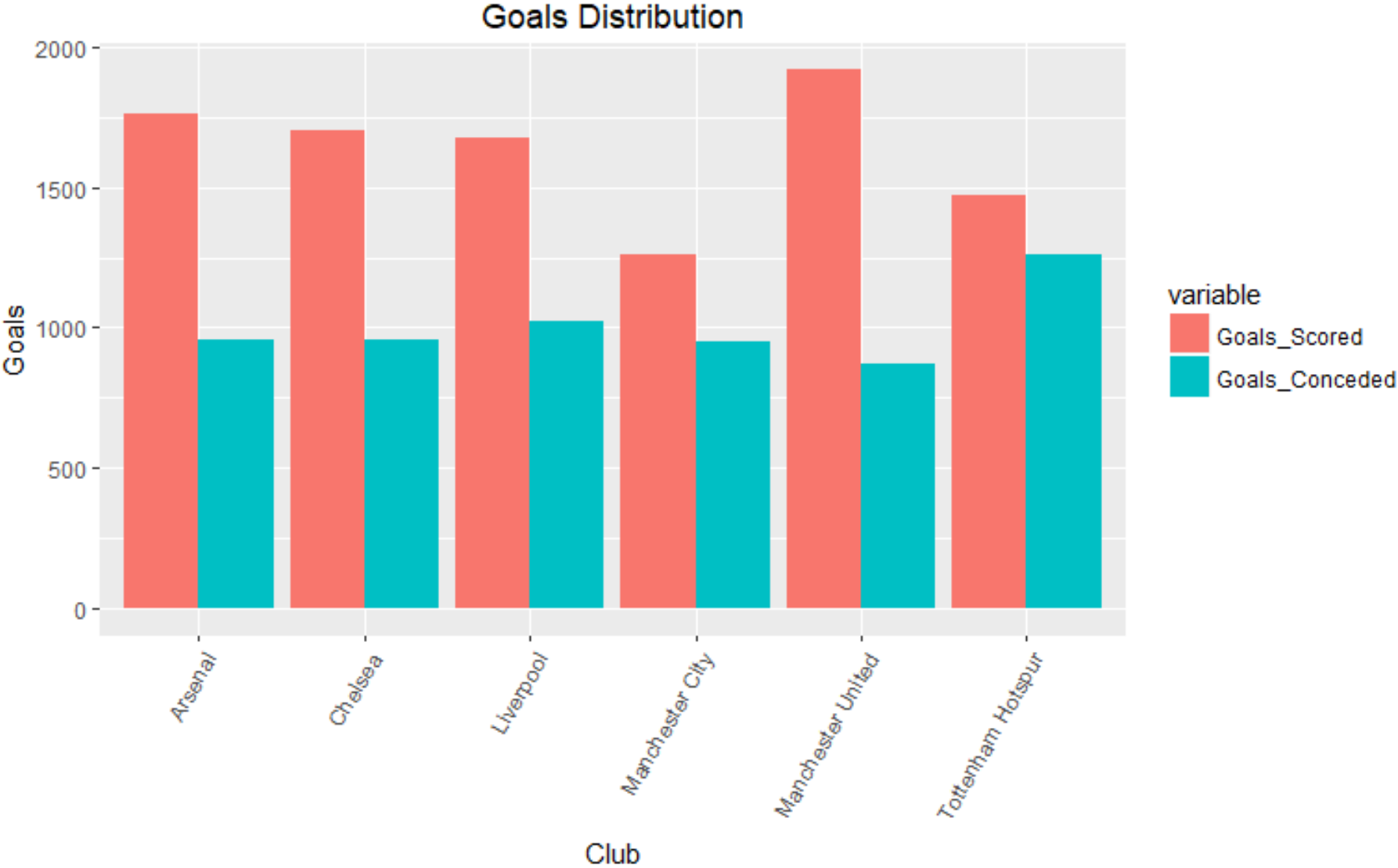
GOALS DIFFERENCE (GOALS FOR – GOALS CONCEDED)

PREMIER LEAGUE POINTS TOTAL SINCE 1992–93

Points distribution across 26 PL Seasons
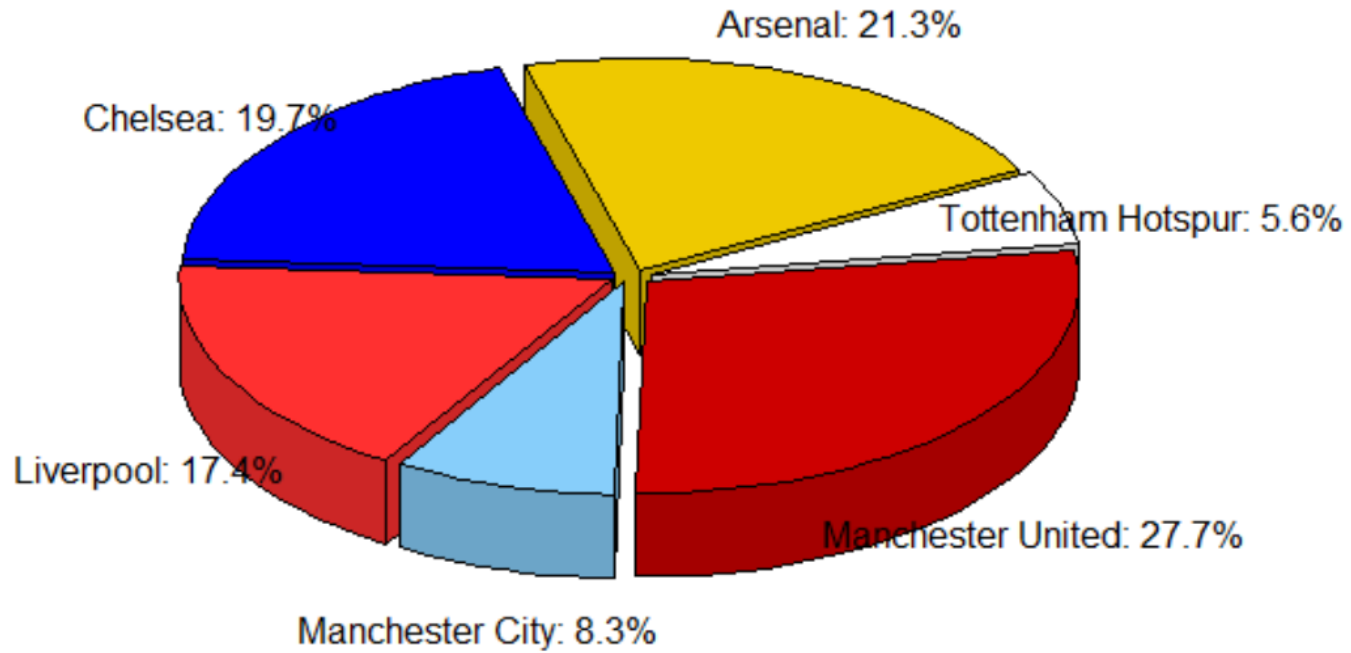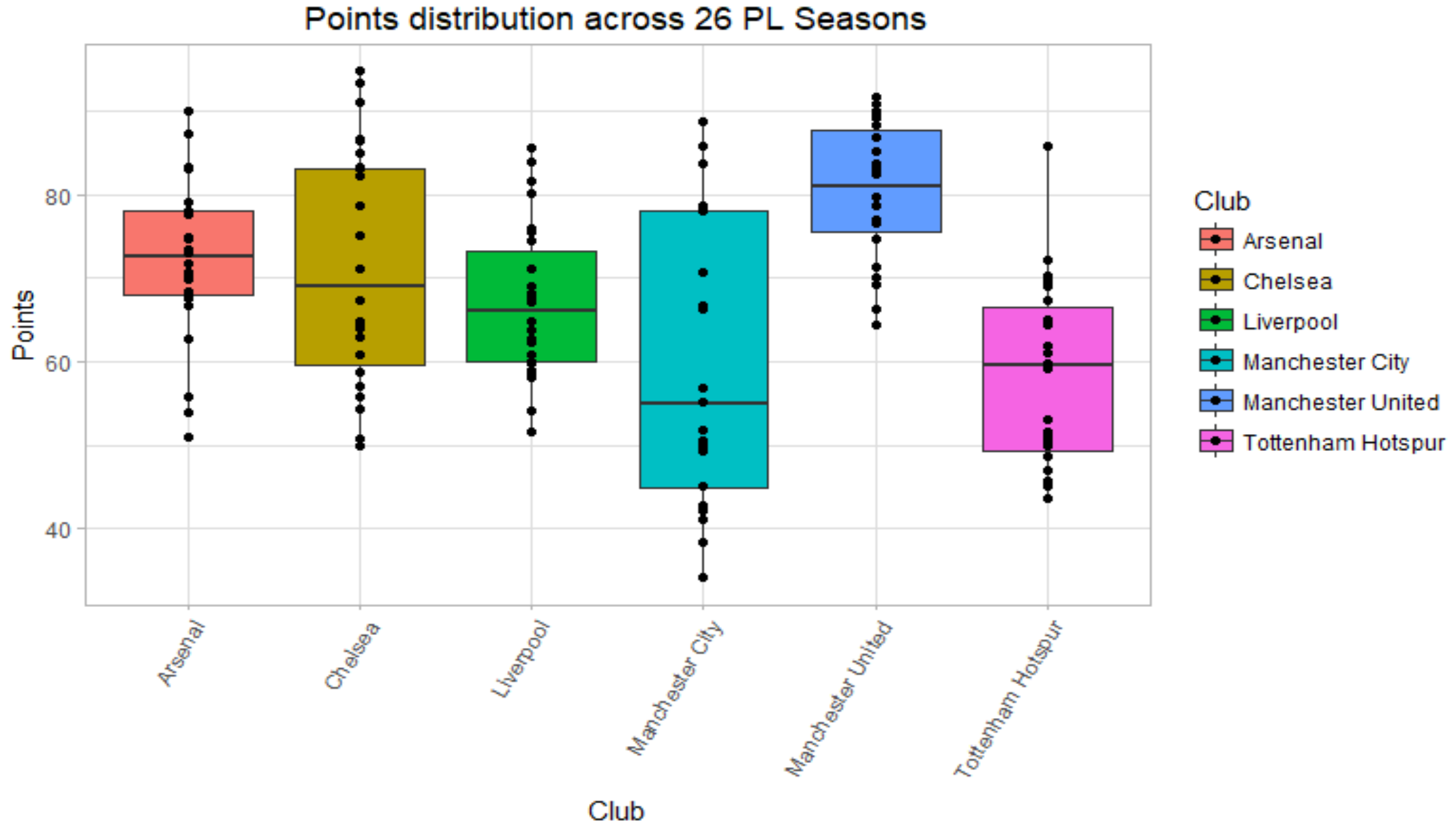
BRADLEY–TERRY MODEL

$$P(\text{TEAM A BEATS TEAM B}) = \frac{\text{ABILITY OF TEAM A}}{\text{ABILITY OF TEAM A } + \text{ ABILITY OF TEAM B}}$$

ASSUMPTIONS –

1) IF A GOOD TEAM PLAYS A BAD TEAM, THE PROBABILITY OF THE GOOD TEAM WINNING IS HIGHER.

2) IF A GOOD TEAM PLAYS A GOOD TEAM, THE PROBABILITY OF THE EACH TEAM WINNING IS AROUND 50%.

BRADLEY–TERRY MODEL

*PRIOR DISTRIBUTION

LOG ABILITY[i] ~ **DNORM(0,1)**

ABILITY[i] <- EXP(LOG ABILITY)

*i – NUMBER OF TEAMS

*LIKELIHOOD/ SAMPLING DISTRIBUTION

$$PROB[i] = \frac{ABILITY(HOME\ CODE\ [i])}{ABILITY(HOME\ CODE\ [i]) + ABILITY(AWAY\ CODE\ [i])}$$

HOME WIN [i] ~ DBIN(PROB [i],1)

*i – NUMBER OF GAMES PLAYED

* HOME TEAM'S WINNING PROBABILITY

PROB NEXT <- ABILITY HOME /(ABILITY HOME + ABILITY AWAY)

HOME WIN ~ DBIN(PROB NEXT,1)

MODEL 1 – TRACE & DENSITY PLOT
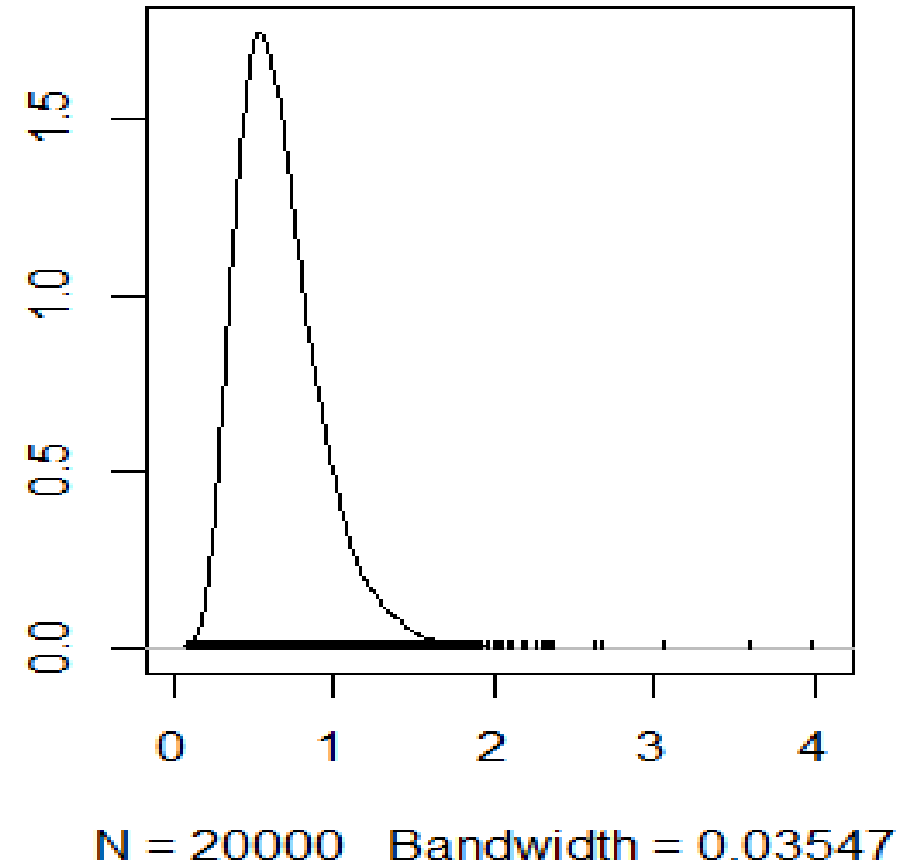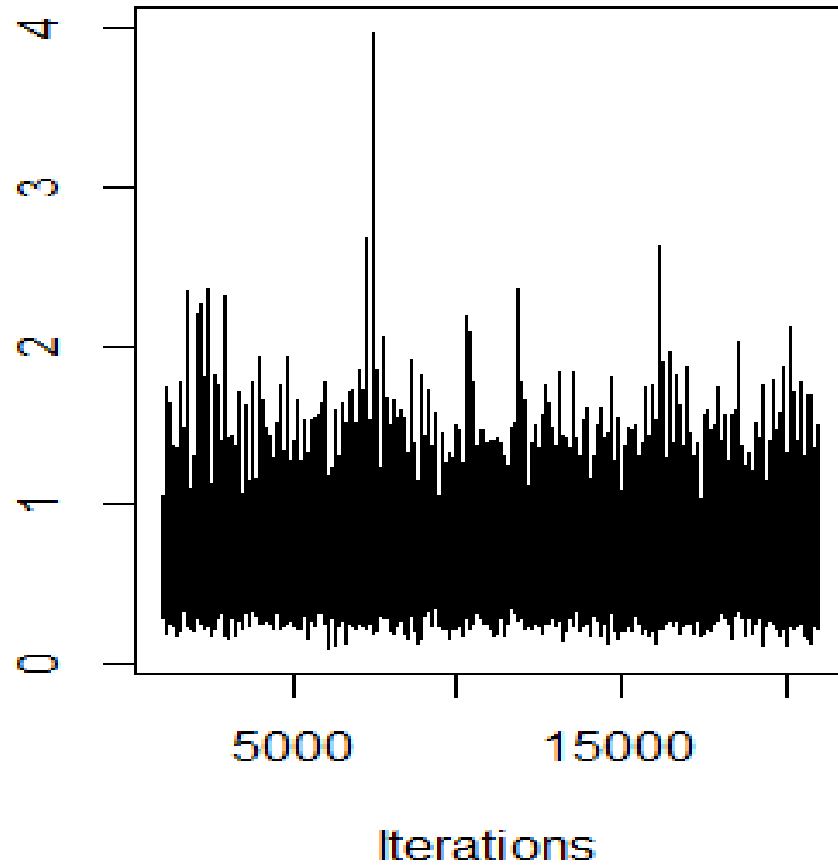
HIERARCHICAL MODEL

*PRIOR DISTRIBUTION (HYPERPARAMETER)

DIVERSITY ~ DUNIF(0,2)

LOG ABILITY[i] ~ DNORM(0,1/DIVERSITY ^{2})

ABILITY[i] <− EXP(LOG ABILITY)

*i – NUMBER OF TEAMS

*LIKELIHOOD/ SAMPLING DISTRIBUTION

$$\text{PROB}[i] = \frac{\text{ABILITY(HOME CODE } [i])}{\text{ABILITY(HOME CODE } [i]) + \text{ABILITY(AWAY CODE } [i])}$$

HOME WIN [i] ~ DBIN(PROB [i],1)

*i – NUMBER OF GAMES PLAYED

* HOME TEAM'S WINNING PROBABILITY

PROB NEXT <− ABILITY HOME /(ABILITY HOME + ABILITY AWAY)

HOME WIN ~ DBIN(PROB NEXT,1)

MODEL 2 − TRACE & DENSITY PLOT

COMPARING TRACE & DENSITY PLOTS OF BOTH MODELS

MODEL 1 – BRADLEY–TERRY MODEL

MODEL 2 – BRADLEY–TERRY (HIERARCHICAL MODEL)

HISTOGRAM FOR HYPERPARAMETER (DIVERSITY)

PREDICTIONS

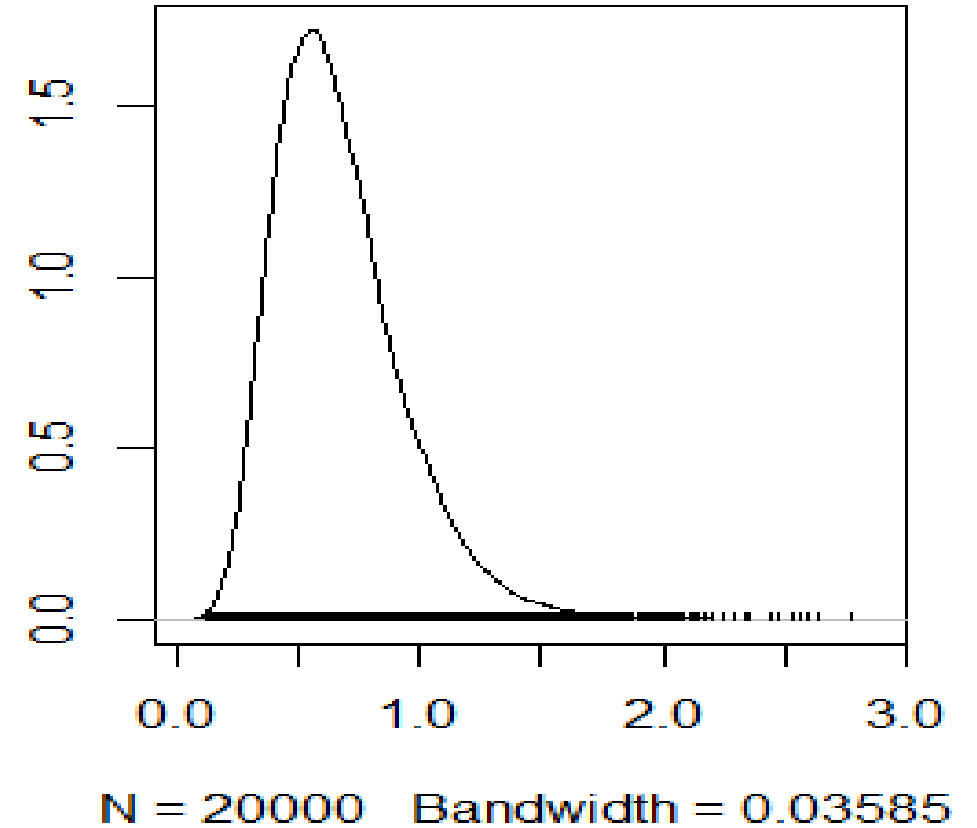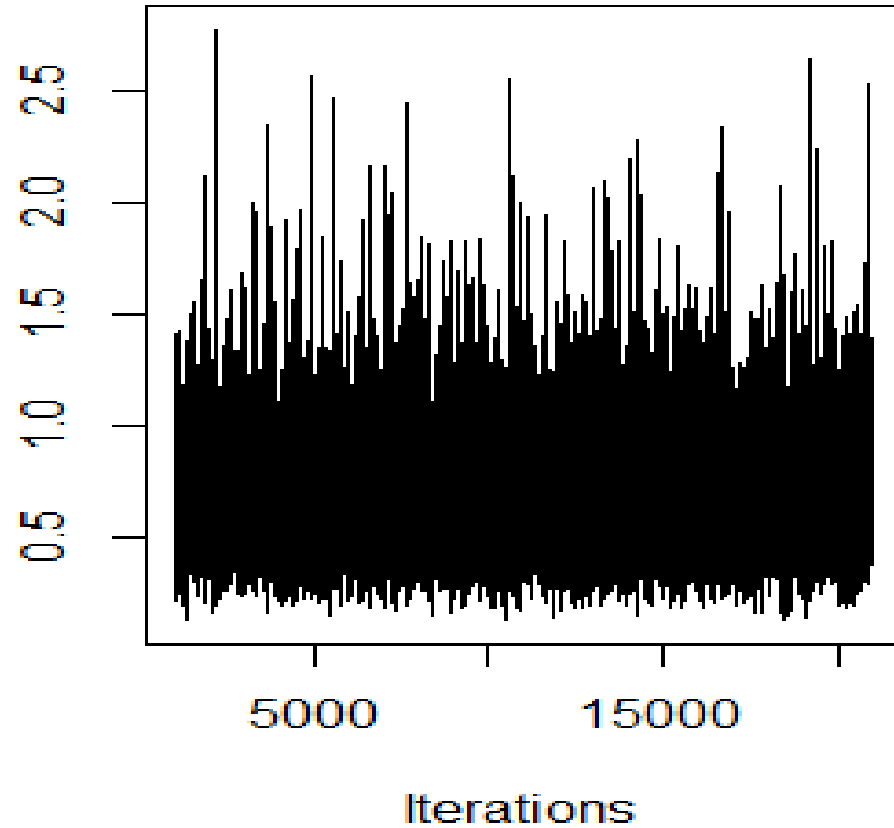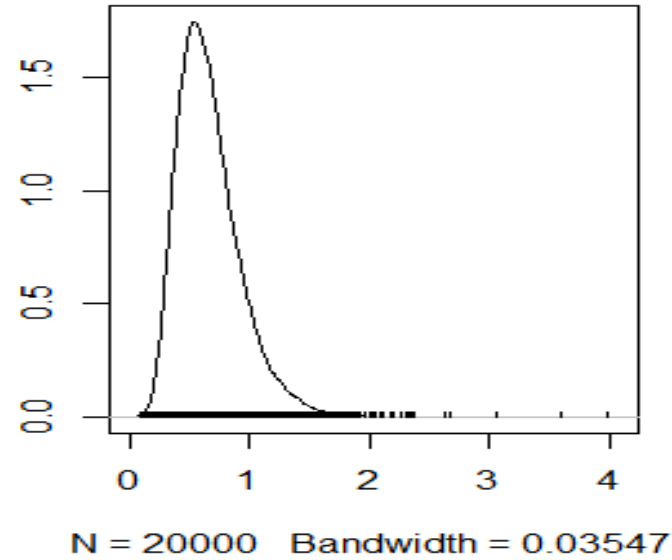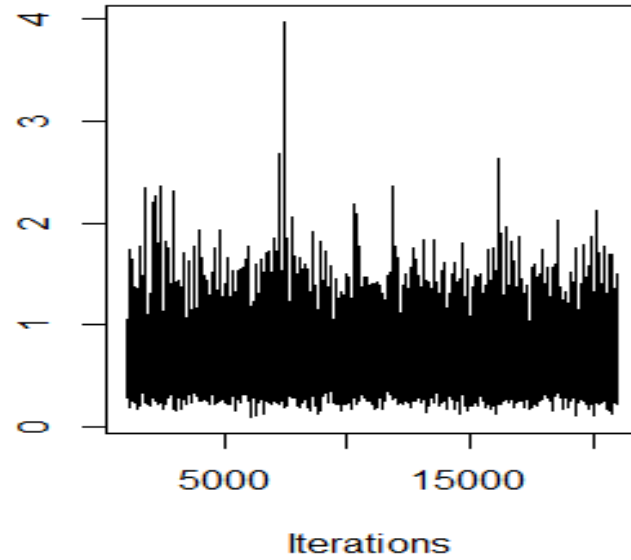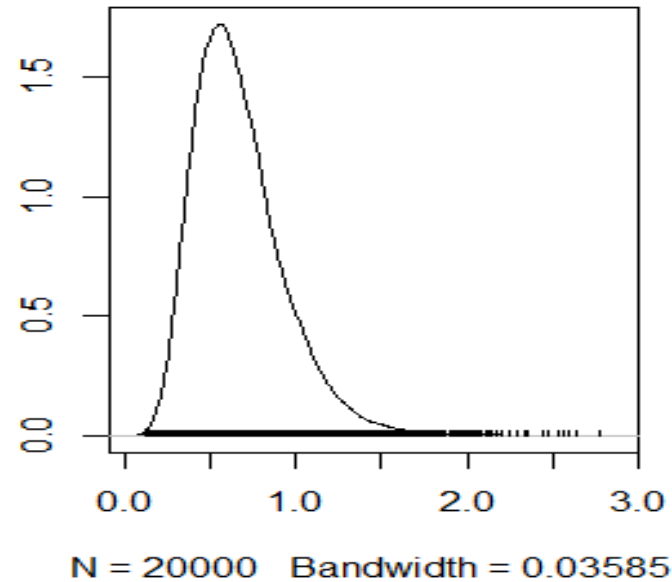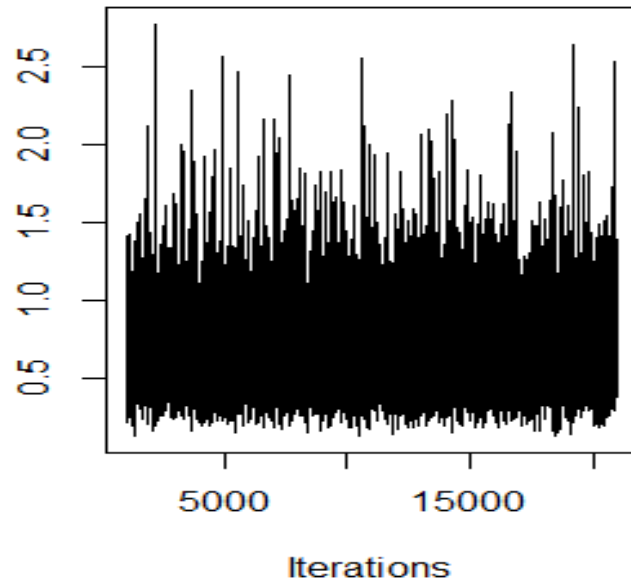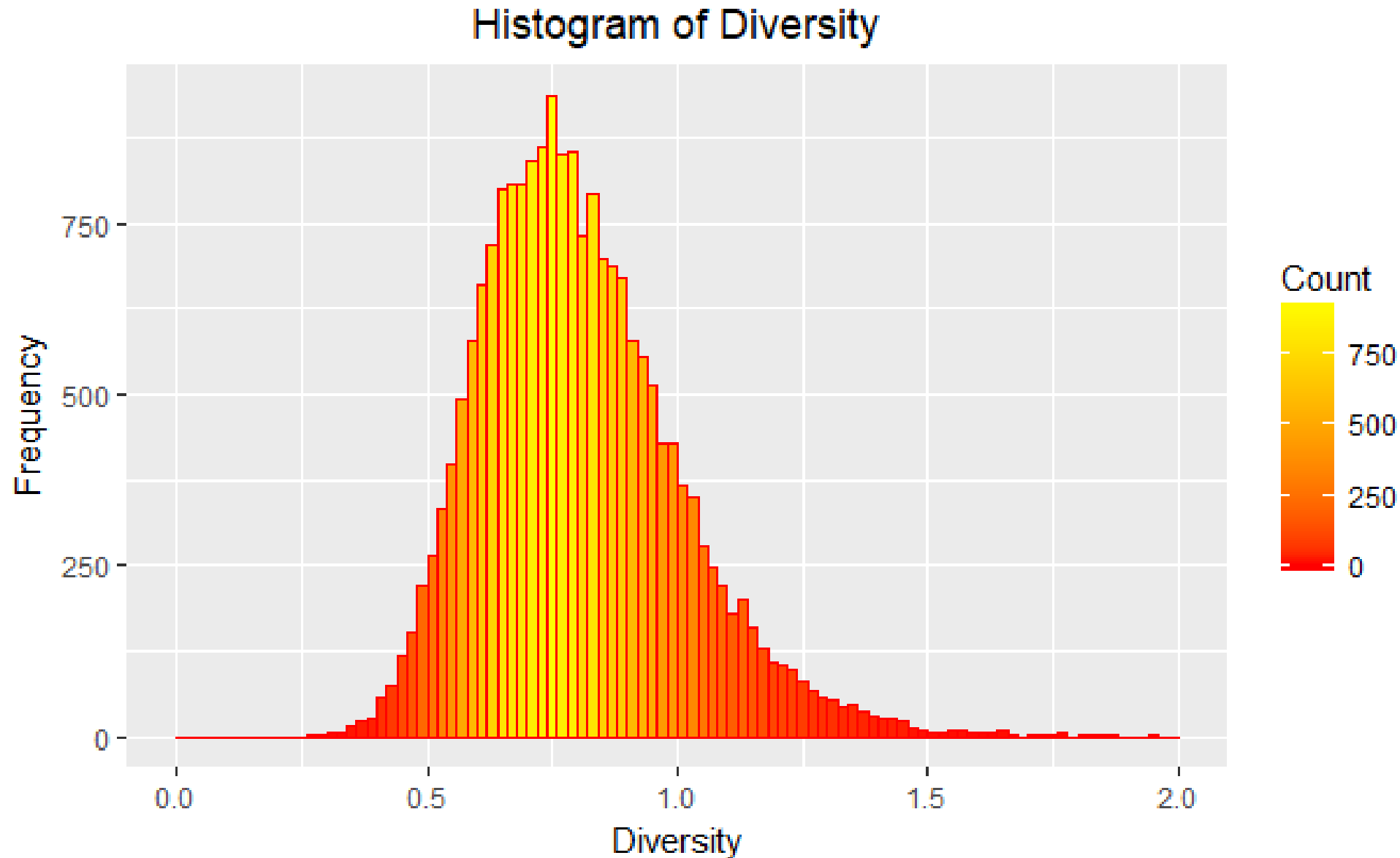| Fixture Date	Home	Away	Home_Code	Away_Code	MODEL 1 Home_Win_%(Bradley-Terry Model)	MODEL 2 Home_Win_%(Hierarchical)	Actual_Outcome	Correct Pred?
March 31 2018	Crystal Palace	Liverpool	6	10	23%	25%	Crystal Palace lose 2-1	Y
March 31 2018	Brighton	Leicester	3	9	38%	40%	Brighton lose 1-0	Y
March 31 2018	Man Utd	Swansea	12	17	79%	76%	Man Utd win 2-0	Y
March 31 2018	Newcastle	Huddersfield	13	8	59%	58%	Newcastle win 1-0	Y
March 31 2018	Watford	Bournemouth	18	2	44%	45%	Draw	
March 31 2018	West Brom	Burnley	19	4	20%	23%	West Brom lose 2-1	Y
March 31 2018	West Ham	Southampton	20	14	52%	53%	West Ham win 3-0	Y
March 31 2018	Everton	Man City	7	11	14%	17%	Man City win 3-1	Y
April 1 2018	Arsenal	Stoke	1	16	72%	71%	Arsenal win 3-0	Y
April 1 2018	Chelsea	Spurs	5	15	47%	47%	Spurs win 3-1	Y
April 7 2018	Everton	Liverpool	7	10	34%	36%	Draw	
April 7 2018	Bournemouth	Crystal Palace	2	6	57%	57%	Draw	
April 7 2018	Brighton	Huddersfield	3	8	57%	55%	Draw	
April 7 2018	Leicester	Newcastle	9	13	60%	58%	Leicester lose 2-1	N
April 7 2018	Stoke	Spurs	16	15	21%	23%	Stoke lose 2-1	Y
April 7 2018	Watford	Burnley	18	4	33%	34%	Watford lose 2-1	Y
April 7 2018	West Brom	Swansea	19	17	30%	32%	Draw	
April 7 2018	Man City	Man Utd	11	12	69%	67%	Man City lose 3-2	N
April 8 2018	Arsenal	Southampton	1	14	72%	71%	Arsenal win 3-2	Y
April 8 2018	Chelsea	West Ham	5	20	74%	72%	Draw	
April 14 2018	Southampton	Chelsea	14	5	23%	26%	Southampton lose 3-2	Y
April 14 2018	Burnley	Leicester	4	9	56%	57%	Burnley win 2-1	Y
April 14 2018	Crystal Palace	Brighton	6	3	49%	49%	Crystal Palace win 3-2	N
April 14 2018	Huddersfield	Watford	8	18	43%	44%	Huddersfield win 1-0	N
April 14 2018	Swansea	Everton	17	7	41%	43%	Draw	
April 14 2018	Liverpool	Bournemouth	10	2	71%	69%	Liverpool win 3-0	Y
April 14 2018	Spurs	Man City	15	11	24%	28%	Spurs lose 1-3	Y
April 15 2018	Newcastle	Arsenal	13	1	31%	34%	Arsenal lose 1-0	N
April 15 2018	Man Utd	West Brom	12	19	89%	87%	Man Utd lose 1-0	N
April 16 2018	West Ham	Stoke	20	16	52%	53%	Draw	
April 17 2018	Brighton	Spurs	3	15	22%	25%	Draw	

Hit Rate 73%