# Analysis

I do not follow American Football and I could not understand the dataset. Also, I could not locate any documentation on what the fields meant. So I have made the following assumptions :

* That the score of the 'TeamName' team is ScoreOff while that of the 'Opponent' is ScoreDef.
* home vs. away games is coded in H vs V field (after 2002).
* Data issues: Inconsistent number of columns per row. Something changed in the data after 2008. Needed to do some manual cleanup.

## Home vs. away hypothesis

To test this I tabulated home vs. away victories for all teams (and cumulatively) and computed the p-values of the home-victory differential under a null hypothesis of fair odds.

This calculation was done using a Gaussian approximation to the Bernouilli. This is justified as almost all teams played 88 games with H vs V recorded which is sufficient for a decent normal approximation, and the power of the test is acceptably high. Also, cumulatively approximately 1400 games were played - giving very good power to this test.

Conclusion: home vs. away advantage is statistically significant for almost all cases.

-----------------------------------------------------------------------

Statistics for team New Orleans Saints

-----------------------------------------------------------------------

home games 86 wins 48 , win-pct 55.814

away games 89 wins 48 , win-pct 53.9326

Under a fair odds null hypothesis approximate p-value is 0.140444

-----------------------------------------------------------------------

Statistics for team Pittsburgh Steelers

-----------------------------------------------------------------------

home games 88 wins 64 , win-pct 72.7273

away games 88 wins 50 , win-pct 56.8182

Under a fair odds null hypothesis approximate p-value is 1.00393e-05

-----------------------------------------------------------------------

Statistics for team New England Patriots

-----------------------------------------------------------------------

home games 88 wins 73 , win-pct 82.9545

away games 87 wins 61 , win-pct 70.1149

Under a fair odds null hypothesis approximate p-value is 3.14831e-10

-----------------------------------------------------------------------

Statistics for team Tampa Bay Buccaneers

-----------------------------------------------------------------------

home games 86 wins 45 , win-pct 52.3256

away games 88 wins 36 , win-pct 40.9091

Under a fair odds null hypothesis approximate p-value is 0.333114

-----------------------------------------------------------------------

Statistics for team Philadelphia Eagles

-----------------------------------------------------------------------

home games 88 wins 52 , win-pct 59.0909

away games 88 wins 52 , win-pct 59.0909

Under a fair odds null hypothesis approximate p-value is 0.0440408

-----------------------------------------------------------------------

Statistics for team Atlanta Falcons

-----------------------------------------------------------------------

home games 88 wins 57 , win-pct 64.7727

away games 88 wins 44 , win-pct 50

Under a fair odds null hypothesis approximate p-value is 0.002789

-----------------------------------------------------------------------

Statistics for team Cleveland Browns

-----------------------------------------------------------------------

home games 88 wins 35 , win-pct 39.7727

away games 88 wins 26 , win-pct 29.5455

Under a fair odds null hypothesis approximate p-value is 0.0275044

-----------------------------------------------------------------------

Statistics for team Cincinnati Bengals

-----------------------------------------------------------------------

home games 88 wins 46 , win-pct 52.2727

away games 88 wins 36 , win-pct 40.9091

Under a fair odds null hypothesis approximate p-value is 0.334908

-----------------------------------------------------------------------

Statistics for team Oakland Raiders

-----------------------------------------------------------------------

home games 88 wins 34 , win-pct 38.6364

away games 88 wins 26 , win-pct 29.5455

Under a fair odds null hypothesis approximate p-value is 0.0165031

-----------------------------------------------------------------------

Statistics for team Buffalo Bills

-----------------------------------------------------------------------

home games 85 wins 42 , win-pct 49.4118

away games 88 wins 28 , win-pct 31.8182

Under a fair odds null hypothesis approximate p-value is 0.456813

-----------------------------------------------------------------------

Statistics for team New York Giants

-----------------------------------------------------------------------

home games 89 wins 49 , win-pct 55.0562

away games 85 wins 46 , win-pct 54.1176

Under a fair odds null hypothesis approximate p-value is 0.170042

-----------------------------------------------------------------------

Statistics for team Detroit Lions

-----------------------------------------------------------------------

home games 88 wins 34 , win-pct 38.6364

away games 88 wins 17 , win-pct 19.3182

Under a fair odds null hypothesis approximate p-value is 0.0165031

-----------------------------------------------------------------------

Statistics for team Carolina Panthers

-----------------------------------------------------------------------

home games 88 wins 45 , win-pct 51.1364

away games 88 wins 41 , win-pct 46.5909

Under a fair odds null hypothesis approximate p-value is 0.415585

-----------------------------------------------------------------------

Statistics for team San Francisco 49ers

-----------------------------------------------------------------------

home games 87 wins 50 , win-pct 57.4713

away games 88 wins 30 , win-pct 34.0909

Under a fair odds null hypothesis approximate p-value is 0.0816971

-----------------------------------------------------------------------

Statistics for team Indianapolis Colts

-----------------------------------------------------------------------

home games 88 wins 66 , win-pct 75

away games 88 wins 56 , win-pct 63.6364

Under a fair odds null hypothesis approximate p-value is 1.36325e-06

-----------------------------------------------------------------------

Statistics for team Seattle Seahawks

-----------------------------------------------------------------------

home games 88 wins 59 , win-pct 67.0455

away games 88 wins 33 , win-pct 37.5

Under a fair odds null hypothesis approximate p-value is 0.000691896

-----------------------------------------------------------------------

Statistics for team Arizona Cardinals

-----------------------------------------------------------------------

home games 88 wins 48 , win-pct 54.5455

away games 88 wins 22 , win-pct 25

Under a fair odds null hypothesis approximate p-value is 0.196884

-----------------------------------------------------------------------

Statistics for team Houston Texans

-----------------------------------------------------------------------

home games 88 wins 45 , win-pct 51.1364

away games 88 wins 32 , win-pct 36.3636

Under a fair odds null hypothesis approximate p-value is 0.415585

-----------------------------------------------------------------------

Statistics for team Tennessee Titans

-----------------------------------------------------------------------

home games 88 wins 51 , win-pct 57.9545

away games 88 wins 41 , win-pct 46.5909

Under a fair odds null hypothesis approximate p-value is 0.0677965

-----------------------------------------------------------------------

Statistics for team Jacksonville Jaguars

-----------------------------------------------------------------------

home games 88 wins 47 , win-pct 53.4091

away games 88 wins 31 , win-pct 35.2273

Under a fair odds null hypothesis approximate p-value is 0.261216

-----------------------------------------------------------------------

Statistics for team Chicago Bears

-----------------------------------------------------------------------

home games 88 wins 54 , win-pct 61.3636

away games 86 wins 36 , win-pct 41.8605

Under a fair odds null hypothesis approximate p-value is 0.0165031

-----------------------------------------------------------------------

Statistics for team San Diego Chargers

-----------------------------------------------------------------------

home games 87 wins 58 , win-pct 66.6667

away games 87 wins 45 , win-pct 51.7241

Under a fair odds null hypothesis approximate p-value is 0.000938207

-----------------------------------------------------------------------

Statistics for team Miami Dolphins

-----------------------------------------------------------------------

home games 87 wins 43 , win-pct 49.4253

away games 87 wins 33 , win-pct 37.931

Under a fair odds null hypothesis approximate p-value is 0.457311

-----------------------------------------------------------------------

Statistics for team New York Jets

-----------------------------------------------------------------------

home games 88 wins 49 , win-pct 55.6818

away games 87 wins 36 , win-pct 41.3793

Under a fair odds null hypothesis approximate p-value is 0.143211

-----------------------------------------------------------------------

Statistics for team St Louis Rams

-----------------------------------------------------------------------

home games 88 wins 39 , win-pct 44.3182

away games 88 wins 25 , win-pct 28.4091

Under a fair odds null hypothesis approximate p-value is 0.143211

-----------------------------------------------------------------------

Statistics for team Baltimore Ravens

-----------------------------------------------------------------------

home games 88 wins 67 , win-pct 76.1364

away games 88 wins 37 , win-pct 42.0455

Under a fair odds null hypothesis approximate p-value is 4.70441e-07

-----------------------------------------------------------------------

Statistics for team Kansas City Chiefs

-----------------------------------------------------------------------

home games 88 wins 46 , win-pct 52.2727

away games 88 wins 30 , win-pct 34.0909

Under a fair odds null hypothesis approximate p-value is 0.334908

-----------------------------------------------------------------------

Statistics for team Denver Broncos

-----------------------------------------------------------------------

home games 88 wins 55 , win-pct 62.5

away games 87 wins 44 , win-pct 50.5747

Under a fair odds null hypothesis approximate p-value is 0.00950824

-----------------------------------------------------------------------

Statistics for team Washington Redskins

-----------------------------------------------------------------------

home games 88 wins 41 , win-pct 46.5909

away games 87 wins 34 , win-pct 39.0805

Under a fair odds null hypothesis approximate p-value is 0.261216

-----------------------------------------------------------------------

Statistics for team Green Bay Packers

-----------------------------------------------------------------------

home games 88 wins 62 , win-pct 70.4545

away games 88 wins 48 , win-pct 54.5455

Under a fair odds null hypothesis approximate p-value is 6.21181e-05

-----------------------------------------------------------------------

Statistics for team Minnesota Vikings

-----------------------------------------------------------------------

home games 87 wins 55 , win-pct 63.2184

away games 88 wins 32 , win-pct 36.3636

Under a fair odds null hypothesis approximate p-value is 0.00683426

-----------------------------------------------------------------------

Statistics for team Dallas Cowboys

-----------------------------------------------------------------------

home games 88 wins 52 , win-pct 59.0909

away games 88 wins 42 , win-pct 47.7273

Under a fair odds null hypothesis approximate p-value is 0.0440408

-----------------------------------------------------------------------

CUMULATIVE STATISTICS

-----------------------------------------------------------------------

home games 1403 wins 805 , win-pct 57.4127

away games 1403 wins 599 , win-pct 42.6942

Under a fair odds null hypothesis approximate p-value is 1.63444e-08

## Early vs. late hypothesis

For this I defined early games to be those played before and including October and Late as those played after October. Here the picture is definitely not as clear - either way. Except for a few teams, nothing statistically significant.

For the cumulative statistic - I divide the season into a good half and a bad half, and tried to see cumulatively across teams if there is a siginificant change in good vs bad half. It has a pvalue of 0.011

-----------------------------------------------------------------------

Statistics for team New Orleans Saints

-----------------------------------------------------------------------

early games 103 wins 55 , win-pct 53.3981

early games 107 wins 60 , win-pct 56.0748

Under a fair odds null hypothesis approximate p-value is 0.245182

-----------------------------------------------------------------------

Statistics for team Pittsburgh Steelers

-----------------------------------------------------------------------

early games 98 wins 65 , win-pct 66.3265

early games 110 wins 71 , win-pct 64.5455

Under a fair odds null hypothesis approximate p-value is 0.000613586

-----------------------------------------------------------------------

Statistics for team New England Patriots

-----------------------------------------------------------------------

early games 101 wins 68 , win-pct 67.3267

early games 107 wins 83 , win-pct 77.5701

Under a fair odds null hypothesis approximate p-value is 0.000248257

-----------------------------------------------------------------------

Statistics for team Tampa Bay Buccaneers

-----------------------------------------------------------------------

early games 100 wins 49 , win-pct 49

early games 109 wins 51 , win-pct 46.789

Under a fair odds null hypothesis approximate p-value is 0.42074

-----------------------------------------------------------------------

Statistics for team Philadelphia Eagles

-----------------------------------------------------------------------

early games 99 wins 55 , win-pct 55.5556

early games 111 wins 72 , win-pct 64.8649

Under a fair odds null hypothesis approximate p-value is 0.134462

-----------------------------------------------------------------------

Statistics for team Atlanta Falcons

-----------------------------------------------------------------------

early games 98 wins 55 , win-pct 56.1224

early games 110 wins 57 , win-pct 51.8182

Under a fair odds null hypothesis approximate p-value is 0.112721

-----------------------------------------------------------------------

Statistics for team Cleveland Browns

-----------------------------------------------------------------------

early games 101 wins 38 , win-pct 37.6238

early games 107 wins 33 , win-pct 30.8411

Under a fair odds null hypothesis approximate p-value is 0.00643054

-----------------------------------------------------------------------

Statistics for team Cincinnati Bengals

-----------------------------------------------------------------------

early games 100 wins 40 , win-pct 40

early games 108 wins 52 , win-pct 48.1481

Under a fair odds null hypothesis approximate p-value is 0.0227501

-----------------------------------------------------------------------

Statistics for team Oakland Raiders

-----------------------------------------------------------------------

early games 100 wins 44 , win-pct 44

early games 110 wins 39 , win-pct 35.4545

Under a fair odds null hypothesis approximate p-value is 0.11507

-----------------------------------------------------------------------

Statistics for team Buffalo Bills

-----------------------------------------------------------------------

early games 100 wins 41 , win-pct 41

early games 108 wins 41 , win-pct 37.963

Under a fair odds null hypothesis approximate p-value is 0.0359303

-----------------------------------------------------------------------

Statistics for team New York Giants

-----------------------------------------------------------------------

early games 101 wins 68 , win-pct 67.3267

early games 109 wins 50 , win-pct 45.8716

Under a fair odds null hypothesis approximate p-value is 0.000248257

-----------------------------------------------------------------------

Statistics for team Detroit Lions

-----------------------------------------------------------------------

early games 97 wins 35 , win-pct 36.0825

early games 111 wins 27 , win-pct 24.3243

Under a fair odds null hypothesis approximate p-value is 0.00305858

-----------------------------------------------------------------------

Statistics for team Carolina Panthers

-----------------------------------------------------------------------

early games 101 wins 41 , win-pct 40.5941

early games 107 wins 53 , win-pct 49.5327

Under a fair odds null hypothesis approximate p-value is 0.0293408

-----------------------------------------------------------------------

Statistics for team San Francisco 49ers

-----------------------------------------------------------------------

early games 101 wins 45 , win-pct 44.5545

early games 107 wins 54 , win-pct 50.4673

Under a fair odds null hypothesis approximate p-value is 0.136859

-----------------------------------------------------------------------

Statistics for team Indianapolis Colts

-----------------------------------------------------------------------

early games 96 wins 64 , win-pct 66.6667

early games 113 wins 74 , win-pct 65.4867

Under a fair odds null hypothesis approximate p-value is 0.000545418

-----------------------------------------------------------------------

Statistics for team Seattle Seahawks

-----------------------------------------------------------------------

early games 98 wins 46 , win-pct 46.9388

early games 110 wins 61 , win-pct 55.4545

Under a fair odds null hypothesis approximate p-value is 0.272227

-----------------------------------------------------------------------

Statistics for team Arizona Cardinals

-----------------------------------------------------------------------

early games 98 wins 36 , win-pct 36.7347

early games 110 wins 44 , win-pct 40

Under a fair odds null hypothesis approximate p-value is 0.00431471

-----------------------------------------------------------------------

Statistics for team Houston Texans

-----------------------------------------------------------------------

early games 83 wins 38 , win-pct 45.7831

early games 93 wins 39 , win-pct 41.9355

Under a fair odds null hypothesis approximate p-value is 0.22114

-----------------------------------------------------------------------

Statistics for team Tennessee Titans

-----------------------------------------------------------------------

early games 102 wins 52 , win-pct 50.9804

early games 107 wins 59 , win-pct 55.1402

Under a fair odds null hypothesis approximate p-value is 0.421511

-----------------------------------------------------------------------

Statistics for team Jacksonville Jaguars

-----------------------------------------------------------------------

early games 100 wins 43 , win-pct 43

early games 108 wins 48 , win-pct 44.4444

Under a fair odds null hypothesis approximate p-value is 0.0807567

-----------------------------------------------------------------------

Statistics for team Chicago Bears

-----------------------------------------------------------------------

early games 97 wins 50 , win-pct 51.5464

early games 111 wins 60 , win-pct 54.0541

Under a fair odds null hypothesis approximate p-value is 0.380334

-----------------------------------------------------------------------

Statistics for team San Diego Chargers

-----------------------------------------------------------------------

early games 99 wins 50 , win-pct 50.5051

early games 109 wins 59 , win-pct 54.1284

Under a fair odds null hypothesis approximate p-value is 0.459972

-----------------------------------------------------------------------

Statistics for team Miami Dolphins

-----------------------------------------------------------------------

early games 98 wins 41 , win-pct 41.8367

early games 112 wins 59 , win-pct 52.6786

Under a fair odds null hypothesis approximate p-value is 0.0530208

-----------------------------------------------------------------------

Statistics for team New York Jets

-----------------------------------------------------------------------

early games 100 wins 50 , win-pct 50

early games 108 wins 55 , win-pct 50.9259

Under a fair odds null hypothesis approximate p-value is 0.5

-----------------------------------------------------------------------

Statistics for team St Louis Rams

-----------------------------------------------------------------------

early games 102 wins 45 , win-pct 44.1176

early games 107 wins 43 , win-pct 40.1869

Under a fair odds null hypothesis approximate p-value is 0.117382

-----------------------------------------------------------------------

Statistics for team Baltimore Ravens

-----------------------------------------------------------------------

early games 100 wins 60 , win-pct 60

early games 111 wins 69 , win-pct 62.1622

Under a fair odds null hypothesis approximate p-value is 0.0227501

-----------------------------------------------------------------------

Statistics for team Kansas City Chiefs

-----------------------------------------------------------------------

early games 100 wins 48 , win-pct 48

early games 108 wins 41 , win-pct 37.963

Under a fair odds null hypothesis approximate p-value is 0.344578

-----------------------------------------------------------------------

Statistics for team Denver Broncos

-----------------------------------------------------------------------

early games 101 wins 57 , win-pct 56.4356

early games 108 wins 61 , win-pct 56.4815

Under a fair odds null hypothesis approximate p-value is 0.0979107

-----------------------------------------------------------------------

Statistics for team Washington Redskins

-----------------------------------------------------------------------

early games 102 wins 47 , win-pct 46.0784

early games 106 wins 44 , win-pct 41.5094

Under a fair odds null hypothesis approximate p-value is 0.214146

-----------------------------------------------------------------------

Statistics for team Green Bay Packers

-----------------------------------------------------------------------

early games 99 wins 61 , win-pct 61.6162

early games 109 wins 70 , win-pct 64.2202

Under a fair odds null hypothesis approximate p-value is 0.0104002

-----------------------------------------------------------------------

Statistics for team Minnesota Vikings

-----------------------------------------------------------------------

early games 102 wins 52 , win-pct 50.9804

early games 108 wins 52 , win-pct 48.1481

Under a fair odds null hypothesis approximate p-value is 0.421511

-----------------------------------------------------------------------

Statistics for team Dallas Cowboys

-----------------------------------------------------------------------

early games 99 wins 49 , win-pct 49.4949

early games 109 wins 55 , win-pct 50.4587

Under a fair odds null hypothesis approximate p-value is 0.459972

# -------------------------------------------------------------------------------------------------------CUMULATIVE STATISTICS-------------------------------------------------------------------------------------------------------

# good-half games 1652 wins 780 , win-pct 47.2315

# poor-half games 1669 wins 881 , win-pct 52.8002

# Under a fair odds null hypothesis approximate p-value is 0.0118021

# To do:

see if we can fit linear models to the data to predict scoring / or classification (e.g. logistic) models to predict winning from other stats like possession, rushing yards, etc.