Codebook

All variables are from the home team's perspective xunless specified otherwise.

Response Variables

`result` (nominal): Match result (win, draw or loss)

`gd` (numeric): Goal differential, calculated by goals scored by home team minus goals scored by away team.

Predictor Variables

`win_odds` (numeric): Pre-match result odds from Bet365

`draw_odds` (numeric): Pre-match result odds from Bet365

`loss_odds` (numeric): Pre-match result odds from Bet365

`h_prev_result1` (nominal): Home team's most recent result

`h_prev_result2` (nominal): Home team's second most recent result

`h_prev_result3` (nominal): Home team's third most recent result

`a_prev_result1` (nominal): Away team's most recent result

`a_prev_result2` (nominal): Away team's second most recent result

`a_prev_result3` (nominal): Away team's third most recent result

`form_diff` (numeric): Home team's goal differential over the last 5 games minus the away team's goal differential over the last 5 games. A large positive value indicates that the home team has been playing better in recent games.

`home_yellows` (numeric): Number of yellow cards the home team has gotten in their last 5 matches. Many yellow cards can lead to defenders being suspensed and are an indicator of how close recent games have been.

`away_yellows` (numeric): Number of yellow cards the away team has gotten in their last 5 matches.

`home_reds` (numeric): Number of red cards the home team has gotten in the last 3 matches. If a player gets a red card, they automatically miss two games.

`away_reds` (numeric): Number of red cards the away team has gotten in their last 3 matches.

`significance` (norminal): Indicates whether a match is more significant for one team than the other. For example, if the away team has already been relegated and the home team is in contention to win the league, the match is more significant for the home team. May be `N`, meaning the match is significant for both teams or insignificant for both teams.

`pts_diff` (numeric): Points earned at home minus points earned away from home in the previous season. Metric for home advantage strength.

`gdd` (numeric): Goal differential at home minus goal differential away from home in the previous season.

`attendance_home` (numeric): Average home attendance of the home team in the previous season. Metric for team size/underdog status.

`attendance_away` (numeric): Average home attendance of the away team in the previous season.

`result_prev` (nominal): Result of this match from the previous season, if it occured in the Premier League.