```r
library(tidyverse)
library(rethinking)
library(jtcr)
library(lubridate)

theme_set(theme_jtc())

box_scores <- read_csv(here::here("content/data/nba-boxscores-2020.csv")) %>%
    rename(game_id = X1)
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   X1 = col_character(),
##   date = col_character(),
##   location = col_character(),
##   losing_abbr = col_character(),
##   losing_name = col_character(),
##   winner = col_character(),
##   winning_abbr = col_character(),
##   winning_name = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```r
box_scores_processed <- box_scores %>%
    mutate(
        winner = if_else(home_points > away_points, "home", "away"),
        home_team = if_else(home_wins == 1, winning_abbr, losing_abbr),
        away_team = if_else(home_wins == 1, losing_abbr, winning_abbr),
        home_net_rtg = home_offensive_rating - home_defensive_rating,
        away_net_rtg = away_offensive_rating - away_defensive_rating,
    )

fix_names <- function(dat, home_away) {

    type <- home_away
    dat$win <- ifelse(dat$winner == type, 1L, 0L)
    dat <- select(dat, game_id, win, starts_with(type))
    orig <- names(dat)
    new <- str_remove(orig, "home_|away_")
    names(dat) <- new
    dat$type <- type
    dat$home <- as.integer(as.factor(dat$type))
    dat

}

box_scores_tidy <- bind_rows(
    fix_names(box_scores_processed, "home"),
```

```
    fix_names(box_scores_processed, "away")
) %>%
    mutate(team_int = as.integer(as.factor(team)))
```

Looking at expected win percentage

Formula is:

$P = \mathcal{N}\frac{PPG - DPPG}{\surd}$