TOY_1

Paula Moreno Blazquez

Enero 2022

```r
library(tidyr)
library(dplyr)
```

PARTE 1: Crear Stints

DATA

```r
home_player <- paste0("home_player",1:10)
away_player <- paste0("away_player",1:10)


toy <- data.frame(
  "id_play" = 1:15,
  "season" = c(rep("S2017",15)),
  "game_code_code" = c(rep("G1",10), rep("G2",5)),
  "quarter" = c(rep(1,3), rep(2,2),rep(3,2),rep(4,3),rep(1,4),rep(2,1)),
  "points_home" = c(0,0,0,2,2,2,5,5,7,7,7,7,10,10,12),
  "points_awayway" = c(0,2,2,2,4,4,7,7,9,9,10,10,12,12,12),
  "home_player1" = c(rep(home_player[1],4),rep(home_player[6],5),rep(home_player[1],6)),
  "home_player2" = home_player[2],
  "home_player3" = c(rep(home_player[3],7), rep(home_player[7],2),rep(home_player[3],6)),
  "away_player1" = away_player[1],
  "away_player2" = c(rep(away_player[2],8),away_player[7], rep(away_player[2],6)),
  "away_player3" = away_player[3]
  )

toy2 <- data.frame(                          #DF2 para comprovar que funciona
  "id_play" = 1:20,
  "season" = c(rep("S2017",20)),
  "game_code" = c(rep("G1",10), rep("G2",10)),
  "quarter" = c(rep(1,3), rep(2,2),rep(3,2),rep(4,3),rep(1,4),rep(2,6)),
  "points_home" = c(0,0,0,2,2,2,5,5,7,7,0,2,2,5,5,7,7,7,9,10),
  "points_awayway" = c(0,2,2,2,4,4,7,7,9,9,0,0,2,2,4,4,6,6,6,6),
  "home_player1" = c(rep(home_player[1],4),rep(home_player[6],5),rep(home_player[1],8), rep(home_player
  "home_player2" = home_player[2],
  "home_player3" = c(rep(home_player[3],7), rep(home_player[7],2),rep(home_player[3],9), rep(home_player
  "away_player1" = away_player[1],
  "away_player2" = c(rep(away_player[2],8),away_player[7], rep(away_player[2],9), rep(away_player[5],2))
  "away_player3" = away_player[3]
  )
```

1

```r
toy_backup <- toy

df <- toy2

df <- df %>% unite("lineups", home_player1:away_player3, remove = TRUE)

#Se crean dos PM porque depende de si eres local o visitante sera positivo o negativo
df$PM_home <- df$points_home-df$points_away
df$PM_away <- -df$PM_home

names(df)
```

```
## [1] "id_play"        "season"         "game_code"      "quarter"
## [5] "points_home"    "points_awayway" "lineups"        "PM_home"
## [9] "PM_away"
```

```r
df
```

```
##    id_play season game_code quarter points_home points_awayway
## 1        1  S2017        G1       1           0              0
## 2        2  S2017        G1       1           0              2
## 3        3  S2017        G1       1           0              2
## 4        4  S2017        G1       2           2              2
## 5        5  S2017        G1       2           2              4
## 6        6  S2017        G1       3           2              4
## 7        7  S2017        G1       3           5              7
## 8        8  S2017        G1       4           5              7
## 9        9  S2017        G1       4           7              9
## 10      10  S2017        G1       4           7              9
## 11      11  S2017        G2       1           0              0
## 12      12  S2017        G2       1           2              0
## 13      13  S2017        G2       1           2              2
## 14      14  S2017        G2       1           5              2
## 15      15  S2017        G2       2           5              4
## 16      16  S2017        G2       2           7              4
## 17      17  S2017        G2       2           7              6
## 18      18  S2017        G2       2           7              6
## 19      19  S2017        G2       2           9              6
## 20      20  S2017        G2       2          10              6
##                                                               lineups
## 1  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 2  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 3  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 4  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 5  home_player6_home_player-2_home_player3_away_player1_away_player2_away_player3
## 6  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 7  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 8  home_player6_home_player2_home_player7_away_player1_away_player2_away_player3
## 9  home_player6_home_player2_home_player7_away_player1_away_player7_away_player3
## 10 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 11 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 12 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 13 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 14 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
```

```
## 15 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 16 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 17 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 18 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 19 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
## 20 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
##     PM_home PM_away
## 1         0       0
## 2        -2       2
## 3        -2       2
## 4         0       0
## 5        -2       2
## 6        -2       2
## 7        -2       2
## 8        -2       2
## 9        -2       2
## 10       -2       2
## 11        0       0
## 12        2      -2
## 13        0       0
## 14        3      -3
## 15        1      -1
## 16        3      -3
## 17        1      -1
## 18        1      -1
## 19        3      -3
## 20        4      -4
```

MERGE Temporada+game_code

```r
df_merged <- df %>%  unite("SeasonGame", c("season", "game_code"))
df_merged
```

```
##    id_play SeasonGame quarter points_home points_awayway
## 1        1    S2017_G1       1           0              0
## 2        2    S2017_G1       1           0              2
## 3        3    S2017_G1       1           0              2
## 4        4    S2017_G1       2           2              2
## 5        5    S2017_G1       2           2              4
## 6        6    S2017_G1       3           2              4
## 7        7    S2017_G1       3           5              7
## 8        8    S2017_G1       4           5              7
## 9        9    S2017_G1       4           7              9
## 10      10    S2017_G1       4           7              9
## 11      11    S2017_G2       1           0              0
## 12      12    S2017_G2       1           2              0
## 13      13    S2017_G2       1           2              2
## 14      14    S2017_G2       1           5              2
## 15      15    S2017_G2       2           5              4
## 16      16    S2017_G2       2           7              4
## 17      17    S2017_G2       2           7              6
## 18      18    S2017_G2       2           7              6
## 19      19    S2017_G2       2           9              6
## 20      20    S2017_G2       2          10              6
```

```
##                                                                          lineups
## 1   home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 2   home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 3   home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 4   home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 5   home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 6   home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 7   home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 8   home_player6_home_player2_home_player7_away_player1_away_player2_away_player3
## 9   home_player6_home_player2_home_player7_away_player1_away_player7_away_player3
## 10  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 11  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 12  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 13  home_player1_home_player2_home_player3_away_merged1_away_player2_away_player3
## 14  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 15  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 16  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 17  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 18  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 19  home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
## 20  home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
##     PM_home PM_away
## 1         0       0
## 2        -2       2
## 3        -2       2
## 4         0       0
## 5        -2       2
## 6        -2       2
## 7        -2       2
## 8        -2       2
## 9        -2       2
## 10       -2       2
## 11        0       0
## 12        2      -2
## 13        0       0
## 14        3      -3
## 15        1      -1
## 16        3      -3
## 17        1      -1
## 18        1      -1
## 19        3      -3
## 20        4      -4
```

ORDENAR DF -> Ya no hace falta

UNIQUE quintetos

```r
library(dplyr)
players_Merged <- df_merged$lineups == lag(df_merged$lineups) ## col_TrueFalse si son igual que lag row
SeasonGame_DIF <- df_merged$SeasonGame == lag(df_merged$SeasonGame)

df_merged$TF_Cambios <- ifelse((((players_Merged == FALSE)|(SeasonGame_DIF == FALSE )), "C", "NC")
df_merged
```

```
##    id_play SeasonGame quarter points_home points_awayway
## 1        1    S2017_G1       1           0              0
## 2        2    S2017_G1       1           0              2
## 3        3    S2017_G1       1           0              2
## 4        4    S2017_G1       2           2              2
## 5        5    S2017_G1       2           2              4
## 6        6    S2017_G1       3           2              4
## 7        7    S2017_G1       3           5              7
## 8        8    S2017_G1       4           5              7
## 9        9    S2017_G1       4           7              9
## 10      10    S2017_G1       4           7              9
## 11      11    S2017_G2       1           0              0
## 12      12    S2017_G2       1           2              0
## 13      13    S2017_G2       1           2              2
## 14      14    S2017_G2       1           5              2
## 15      15    S2017_G2       2           5              4
## 16      16    S2017_G2       2           7              4
## 17      17    S2017_G2       2           7              6
## 18      18    S2017_G2       2           7              6
## 19      19    S2017_G2       2           9              6
## 20      20    S2017_G2       2          10              6
##                                                                               lineups
## 1  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 2  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 3  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 4  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 5  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 6  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 7  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 8  home_player6_home_player2_home_player7_away_player1_away_player2_away_player3
## 9  home_player6_home_player2_home_player7_away_player1_away_player7_away_player3
## 10 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 11 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 12 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 13 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 14 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 15 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 16 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 17 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 18 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 19 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
## 20 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
##    PM_home PM_away TF_Cambios
## 1        0       0       <NA>
## 2       -2       2         NC
## 3       -2       2         NC
## 4        0       0         NC
## 5       -2       2          C
## 6       -2       2         NC
## 7       -2       2         NC
## 8       -2       2          C
## 9       -2       2          C
## 10      -2       2          C
## 11       0       0          C
```

```
## 12        2        -2          NC
## 13        0         0          NC
## 14        3        -3          NC
## 15        1        -1          NC
## 16        3        -3          NC
## 17        1        -1          NC
## 18        1        -1          NC
## 19        3        -3           C
## 20        4        -4          NC
```

Hasta aquí tenemos detectados cuando hay cambios

STINTS

```r
df_merged$stint <- ifelse(lead(df_merged$TF)=="C", df_merged$TF, NA)  #Filas que tenemos que conservar

last_row <- tail(df_merged, n =1)              #Tendremos que añadirla al final.
last_row$stint <- "LastRow"

df_merged
```

```
##     id_play SeasonGame quarter points_home points_awayway
## 1         1    S2017_G1       1           0              0
## 2         2    S2017_G1       1           0              2
## 3         3    S2017_G1       1           0              2
## 4         4    S2017_G1       2           2              2
## 5         5    S2017_G1       2           2              4
## 6         6    S2017_G1       3           2              4
## 7         7    S2017_G1       3           5              7
## 8         8    S2017_G1       4           5              7
## 9         9    S2017_G1       4           7              9
## 10       10    S2017_G1       4           7              9
## 11       11    S2017_G2       1           0              0
## 12       12    S2017_G2       1           2              0
## 13       13    S2017_G2       1           2              2
## 14       14    S2017_G2       1           5              2
## 15       15    S2017_G2       2           5              4
## 16       16    S2017_G2       2           7              4
## 17       17    S2017_G2       2           7              6
## 18       18    S2017_G2       2           7              6
## 19       19    S2017_G2       2           9              6
## 20       20    S2017_G2       2          10              6
##                                                                        lineups
## 1  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 2  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 3  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 4  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 5  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 6  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 7  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 8  home_player6_home_player2_home_player7_away_player1_away_player2_away_player3
## 9  home_player6_home_player2_home_player7_away_player1_away_player7_away_player3
## 10 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 11 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
```

```
## 12 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 13 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 14 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 15 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 16 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 17 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 18 home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 19 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
## 20 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
##    PM_home PM_away TF_Cambios stint
## 1        0       0       <NA>  <NA>
## 2       -2       2         NC  <NA>
## 3       -2       2         NC  <NA>
## 4        0       0         NC    NC
## 5       -2       2          C  <NA>
## 6       -2       2         NC  <NA>
## 7       -2       2         NC    NC
## 8       -2       2          C     C
## 9       -2       2          C     C
## 10      -2       2          C     C
## 11       0       0          C  <NA>
## 12       2      -2         NC  <NA>
## 13       0       0         NC  <NA>
## 14       3      -3         NC  <NA>
## 15       1      -1         NC  <NA>
## 16       3      -3         NC  <NA>
## 17       1      -1         NC  <NA>
## 18       1      -1         NC    NC
## 19       3      -3          C  <NA>
## 20       4      -4         NC  <NA>
```

```r
STINTS <- df_merged %>% drop_na(stint)        #Eliminar NA rows
STINTS <- rbind(STINTS, last_row)
STINTS$ID_Stint <- 1:nrow(STINTS)

STINTS_PM <- STINTS %>% select(c(ID_Stint, SeasonGame, quarter, lineups,
                                 PM_home, PM_away))

STINTS_PM
```

```
##    ID_Stint SeasonGame quarter
## 1         1    S2017_G1       2
## 2         2    S2017_G1       3
## 3         3    S2017_G1       4
## 4         4    S2017_G1       4
## 5         5    S2017_G1       4
## 6         6    S2017_G2       2
## 20        7    S2017_G2       2
##                                                                   lineups
## 1  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 2  home_player6_home_player2_home_player3_away_player1_away_player2_away_player3
## 3  home_player6_home_player2_home_player7_away_player1_away_player2_away_player3
## 4  home_player6_home_player2_home_player7_away_player1_away_player7_away_player3
## 5  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
## 6  home_player1_home_player2_home_player3_away_player1_away_player2_away_player3
```

```
## 20 home_player1_home_player2_home_player9_away_player1_away_player5_away_player3
##    PM_home PM_away
## 1        0        0
## 2       -2        2
## 3       -2        2
## 4       -2        2
## 5       -2        2
## 6        1       -1
## 20       4       -4
```

```
dim(STINTS_PM)
```

```
## [1] 7 6
```

```
names(STINTS_PM)
```

```
## [1] "ID_Stint"   "SeasonGame" "quarter"    "lineups"    "PM_home"
## [6] "PM_away"
```

PARTE 2: Dummys Jugadores

```r
#PRUEBA1

vec_players <- unique(c(home_player, away_player))               #Vector con los nombres de los jugadores
#length(vec_players)

NA_players <- matrix(data=NA, nrow = dim(STINTS_PM)[1], ncol = length(vec_players)) #Solo jugadores

df_DummyPlayers <- cbind(STINTS_PM, NA_players)                      #DF con jugadores como columnas
names(df_DummyPlayers) <- c(names(STINTS_PM), vec_players)

## TRUE/FALSE si aparecen en la alineacion
df_DummyPlayers[vec_players] <- grepl(vec_players, df_DummyPlayers$lineups, fixed=TRUE)

#Me las rellena todas igual al primer jugador que aparece en vec_players (home_player1). Como hacer
#para que rellene con todos los jugadores??


#Como deberian ser:
d1 <- grepl(vec_players[1], df_DummyPlayers$lineups, fixed=TRUE) #home_player1
d2 <- grepl(vec_players[2], df_DummyPlayers$lineups, fixed=TRUE) #home_player2
d3 <- grepl(vec_players[3], df_DummyPlayers$lineups, fixed=TRUE) #home_player3
d4 <- grepl(vec_players[4], df_DummyPlayers$lineups, fixed=TRUE) #home_player4

df_Dum <- data.frame(
  "home_player1" <- d1,
  "home_player2" <- d2,
  "home_player3" <- d3,
  "home_player4" <- d4
)

names(df_Dum) <- paste0("home_player", 1:4)
df_Dum
```

```r
#Lo que sale
df_DummyPlayers
```

```r
#PRUEBA2 - fastDummies
vec_players <- unique(c(home_player, away_player)) #Vector con los nombres de los jugadores
#length(vec_players)

#dummy_cols(STINTS_PM, select_columns = ??? )
```