How are countries doing in accommodating their soccer players? We look at sankey-network diagram to find out how players are moving across to play in clubs worldwide.

```r
setwd("D:/msc-ds/course-resource/data-visualization/project")
rm(list=ls())

library(RColorBrewer)
library(dplyr)
library(networkD3)
library(htmlwidgets)

soccer.preprocessed <- read.csv(
  "soccer-preprocessed-with-club-nationality.csv",
  encoding = "UTF-8"
)

nw.data <- soccer.preprocessed[
  !is.na(soccer.preprocessed$Club.Nationality),
]

top_countries_having_most_player <- sort(table(nw.data[, 3]), decreasing=T)[1:10]

# nw.data <- nw.data[
#   nw.data$Nationality %in% names(top_countries_having_most_player),
#   c(3, 7)
# ]

nw.data <- nw.data[, c(3, 7)]

nw.data$Nationality <- as.character(nw.data$Nationality)
nw.data$Club.Nationality <- as.character(nw.data$Club.Nationality)


nw.data$Nationality <- ifelse(
  nw.data$Nationality %in% names(top_countries_having_most_player), nw.data$Nationality, "Others"
)

nw.data$Club.Nationality <- ifelse(
  nw.data$Club.Nationality %in% names(top_countries_having_most_player), nw.data$Club.Nationality, "Oth
)

# View(nw.data)

links <- data.frame(
  source=character(),
  target=character(),
  value=integer(),
  IDsource=integer(),
  IDtarget=integer()
)

nationalities <- unique(c(as.character(nw.data$Nationality)))
for(n in nationalities) {
  club.nationalities <- table(nw.data$Club.Nationality[nw.data$Nationality==n])
```

```r
  for(cn in names(club.nationalities)) {
    srdf <- data.frame(
      source=n,
      target=cn,
      value=club.nationalities[cn],
      IDsource=NA,
      IDtarget=NA
    )
    links <- bind_rows(links, srdf)
  }
}

rownames(links) <- row.names(c(1:nrow(links)))
links$target <- paste(links$target, " ", sep="")

nodes <- data.frame(
  name=c(
    as.character(links$source),
    as.character(links$target)
  ) %>% unique()
)

nodes$name <- sort(nodes$name)

links$IDsource=match(links$source, nodes$name)-1
links$IDtarget=match(links$target, nodes$name)-1

# View(links)

pallete.paired <- brewer.pal(n=12, name="Paired")
# pallete.paired

ColourScal ='d3.scaleOrdinal()
              .range([
                "#A6CEE3",
                "#1F78B4",
                "#B2DF8A",
                "#33A02C",
                "#FB9A99",
                "#E31A1C",
                "#FDBF6F",
                "#FF7F00",
                "#CAB2D6",
                "#6A3D9A",
                "#FFFF99",
                "#B15928"
            ])'

sankey.net <- sankeyNetwork(
  Links=links,
  Nodes=nodes,
  Source="IDsource",
  Target="IDtarget",
```
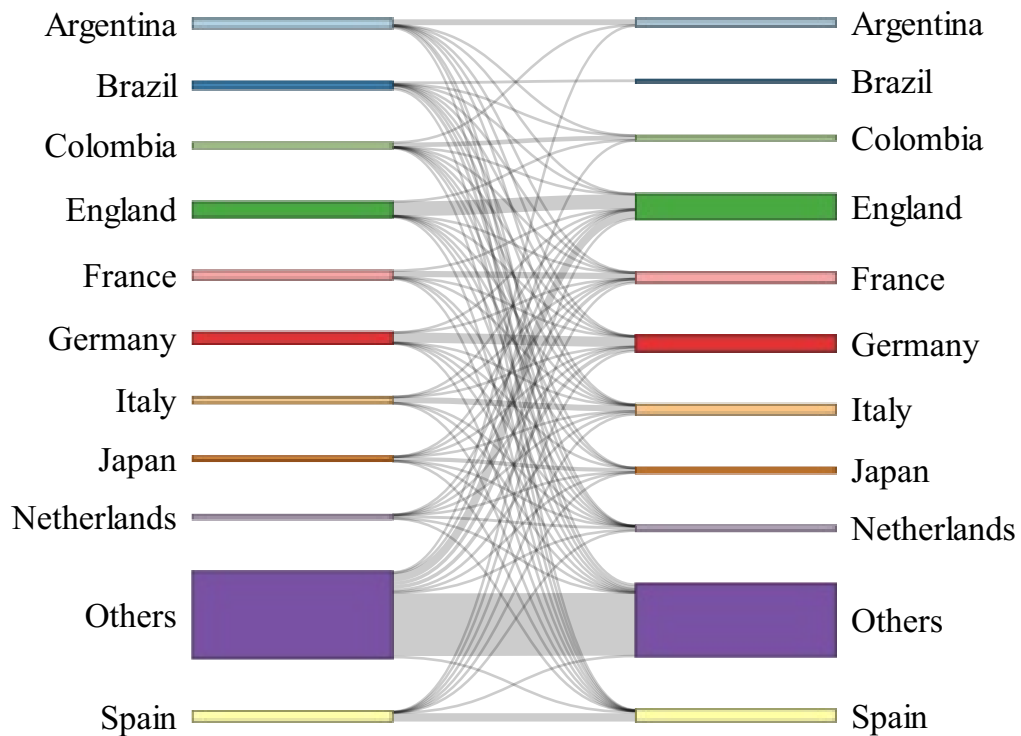
```r
    Value="value",
    NodeID="name",
    sinksRight=FALSE,
    colourScale=ColourScal,
    nodeWidth=75,
    fontSize=13,
    nodePadding=20,
    iterations=0
)

sankey.net <- onRender(
    sankey.net,
    '
    function(el,x){
    d3.select(el)
    .selectAll(".node text")
    .filter(function(d) { return d.name.lastIndexOf(" ") != d.name.length-1 })
    .attr("x", x.options.nodeWidth - (x.options.nodeWidth + 6))
    .attr("text-anchor", "end");
    }
    '
)

sankey.net
```

Observations:

1. Out of 17918 players in total, the top 10 countries with most players have almost half of the players, 8839, combined whereas the rest of the world, apparently 153 other countries, have 9079 players.

2. Major European countries, with most popular leagues and clubs, are quite successful in retaining their home players, e.g. England (93%), Germany (90%), Italy (91%), Spain (82%), France (72%) etc.

3. South American countries, Argentina with 67%, Columbia with 80% and one of the most popular destination of soccer in Asia, Japan with 91% are also doing a good job accommodating their talents in their homeland.

4. One sad exception seems to be Brazil. Despite being 5 times title winner of FIFA World Cup, most of her players are playing outside of Brazil. Only 36% of the players playing in home clubs either indicates poor infrastructure of their club football ecosystem or the clubs are unable to pay their talents by their demands.

5. England draws the most influx of players from other countries, 879, indicating the most popular soccer career destination among the players.