

HW6

Krzysztof Sawicki

2022-12-28

Wprowadzenie

Zadanie polegało na opisie narzędzia/pakietu służącego do wizualizacji danych, którego nie poznaliśmy na zajęciach. Zdecydowałem się na pakiet Lattice.

Opis pakietu

Paczka **Lattice** służy wizualizacji danych w języku R, została ona zaprojektowana z naciskiem na dane wielowymiarowe i umożliwia łatwe tworzenie wielu małych wykresów. Dzięki lattice bez większych trudności jesteśmy w stanie tworzyć panelowe rysunki do porównywania różnych wartości albo podgrup danych zmiennej. W lattice możemy przygotować większość ze znanych nam wykresów oraz dowolnie je customizować, paczka ta funkcjonalnością nie odbiega od ggplot oraz posiada niezbyt skomplikowaną składnię.

Wizualizacja

Użyjemy lattice do porównania liczby rozegranych minut przez dwóch znakomych zawodników NBA Lebrona Jamesa oraz Stephena Currego. W sezonie 2015/2016 po fantastycznie rozegranym sezonie owi zawodnicy spotkali się w finale. Lebron grał świetnie, lecz Curry wyglądał na zmęczonego i momentami był cieniem samego siebie w wyniku czego jego drużyna przegrała. Sprawdźmy czy lattice pomoże nam ustalić przyczynę jego porażki.

Warto wiedzieć:

- sezon zasadniczy składa się z 82 spotkań
- mecz trwa 48 minut

Wczytanie potrzebnych bibliotek

```
library(dplyr)
library(lattice)
```

Obróbka danych

```

curry <- read.csv("curry.csv", fileEncoding = "UTF-8")
lebron <- read.csv("lebron.csv", fileEncoding = "UTF-8")

lebron %>%
  select(Rk,MP) %>%
  mutate(MP = as.numeric(sub(":", ".", lebron$MP, fixed = TRUE))) %>%
  mutate(PartsOfSeason = case_when(
    0 < Rk & Rk <= 20 ~ "First part of a season",
    20 < Rk & Rk <= 40 ~ "Second part of a season",
    40 < Rk & Rk <= 60 ~ "Third part of a season",
    60 < Rk & Rk <= 82 ~ "Fourth part of a season"
  )) %>%
  mutate(Player = "Lebron James") %>%
  mutate(PartsOfSeason = factor(PartsOfSeason,
                                levels = c("First part of a season",
                                             "Second part of a season",
                                             "Third part of a season",
                                             "Fourth part of a season"))) -> lebron

curry %>%
  select(Rk,MP) %>%
  mutate(MP = as.numeric(sub(":", ".", curry$MP, fixed = TRUE))) %>%
  mutate(PartsOfSeason = case_when(
    0 < Rk & Rk <= 20 ~ "First part of a season",
    20 < Rk & Rk <= 40 ~ "Second part of a season",
    40 < Rk & Rk <= 60 ~ "Third part of a season",
    60 < Rk & Rk <= 82 ~ "Fourth part of a season"
  )) %>%
  mutate(Player = "Stephen Curry") %>%
  mutate(PartsOfSeason = factor(PartsOfSeason,
                                levels = c("First part of a season",
                                             "Second part of a season",
                                             "Third part of a season",
                                             "Fourth part of a season")))-> curry

df <- rbind(lebron,curry)
df[is.na(df)] = 0

```

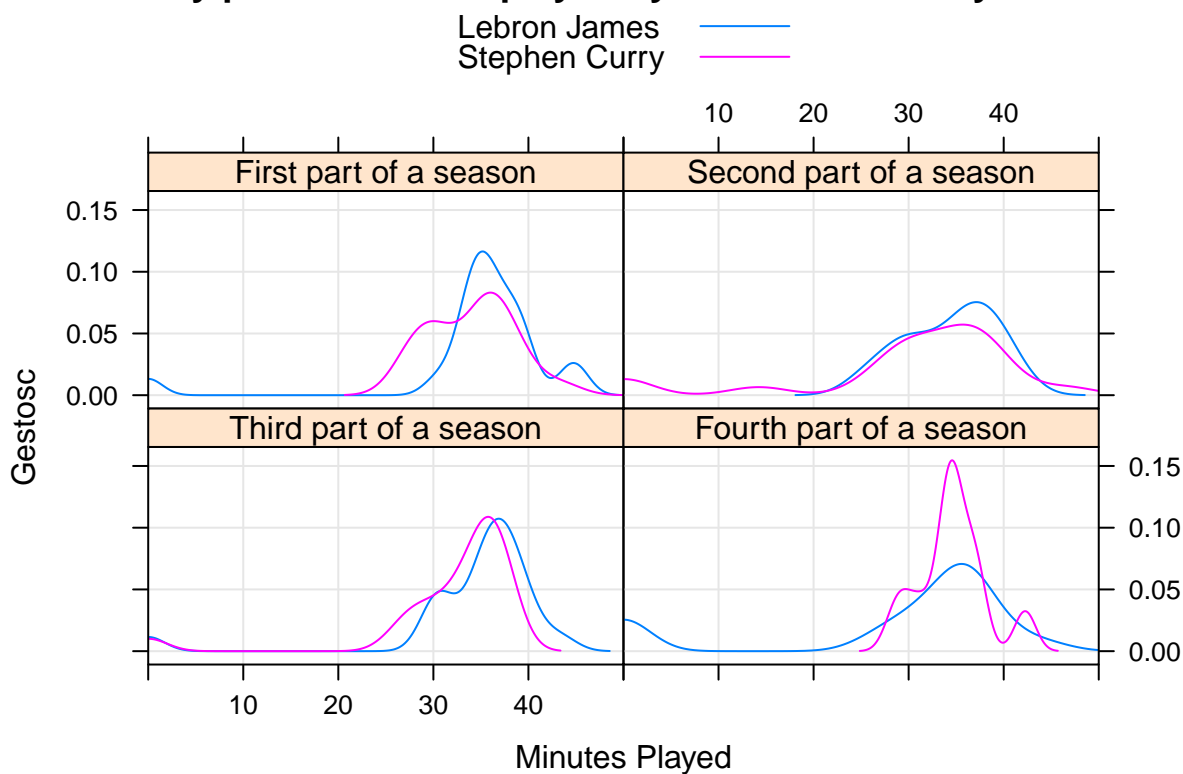
Wykres

```

densityplot(~MP | PartsOfSeason, data= df,
            groups = Player,type=c("p","g"),auto.key = TRUE,
            xlab="Minutes Played",
            main="Density plot of minutes played by Lebron and Curry in 2015",
            plot.points=FALSE,
            as.table = T,
            xlim = c(0, 50))

```

Density plot of minutes played by LeBron and Curry in 2015



Źródło danych: <https://www.basketball-reference.com/>

Podsumowanie

Jak możemy odczytać z wykresu, w ostatniej fazie sezonu LeBron grał zdecydowanie rzadziej, podczas kilku spotkań nie wchodził nawet na boisko (zagrał 0 minut). Curry z kolei cały czas grał na wysokiej intensywności co mogło spowodować, że na spotkania finałowe brakło mu sił.