

Dygraphs_{package}

Wstęp

Do swojej pracy domowej „na warsztat” wybrałem paczkę dygraphs. Decyzja ta zapadła z trzech powodów:

1. Narzędzie to jest równie proste w działaniu jak chociażby poznane już przez nas ggplot czy plotly, jednak cechuje się (moim zdaniem) większą intuicyjnością.
2. Dygraphs jest wystarczająco rozpowszechniony i opisany w Internecie, przez co prosto wychodziło mi szukanie rozwiązań danych problemów pojawiających się podczas kodowania wizualizacji danych.
3. Mocno zaciekała mnie opcja predykcji zachowań danych, której dotąd nie spotykałem w innych paczkach. (O tym nieco później)

Dane

Do zabawy wybrałem proste dane opisujące ilość urodzeń danego dnia na przestrzeni lat 1994-2014 w Stanach Zjednoczonych.

Dane te pobrałem ze strony:

<https://www.kaggle.com/code/avinashvudatala/us-births-1994-2014/notebook>

Kod

```
install.packages("dygraphs")
library(dygraphs)
library(dplyr)
library(xts)

#Konwertowanie naszej ramki danych
births <- read.csv("C:\\Users\\zareb\\OneDrive\\Desktop\\ZarebaAdrian\\US_births_1994-2014.csv") %>%
  group_by(year, month) %>%
  summarise_each(funs(sum), -date_of_month, -day_of_week) %>%
  mutate(year = ifelse(month < 10,
    paste(year, month, sep = "0"),
    paste(year, month, sep = ""))) %>%
  mutate(year = as.yearmon(as.character(year),
    format = "%Y%m")) %>% select(year, births)

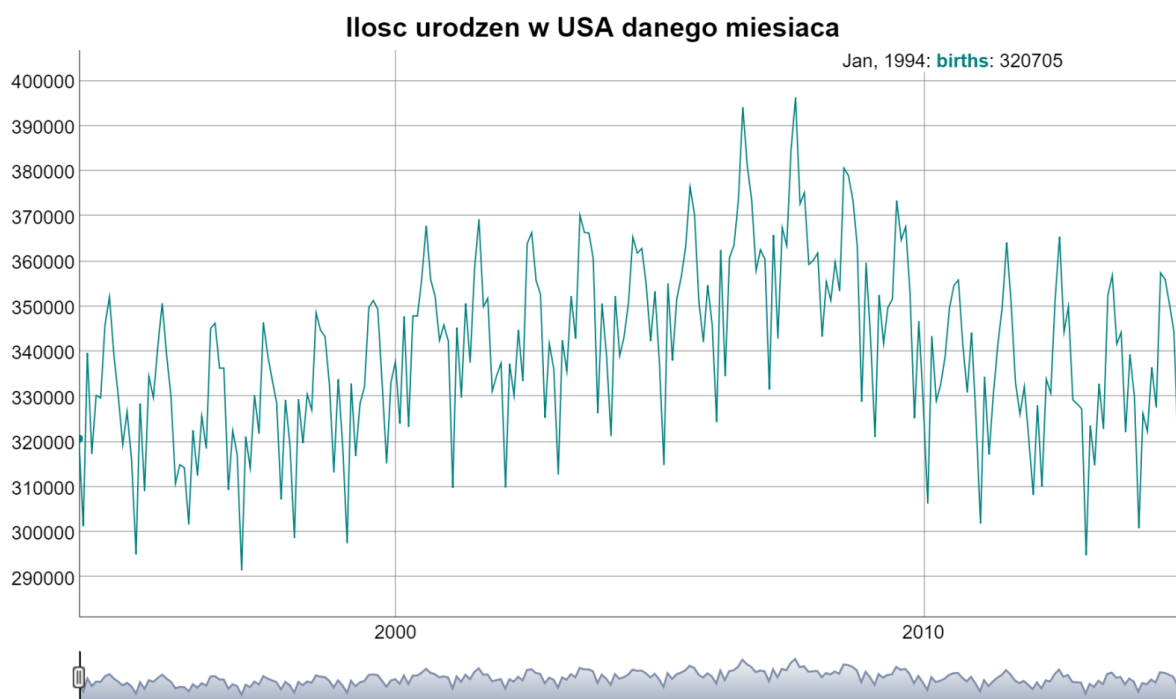
#Zapisanie naszej ramki danych w xts (inaczej paczka nie kompiluje)
births_xts <- xts(x = births[,1], order.by=births$year)

#Tworzenie wizualizacji naszych danych
visual <- dygraph(births_xts, main = "Ilosc urodzen w USA danego miesiaca") %>%
  dyRangeSelector()

visual
```

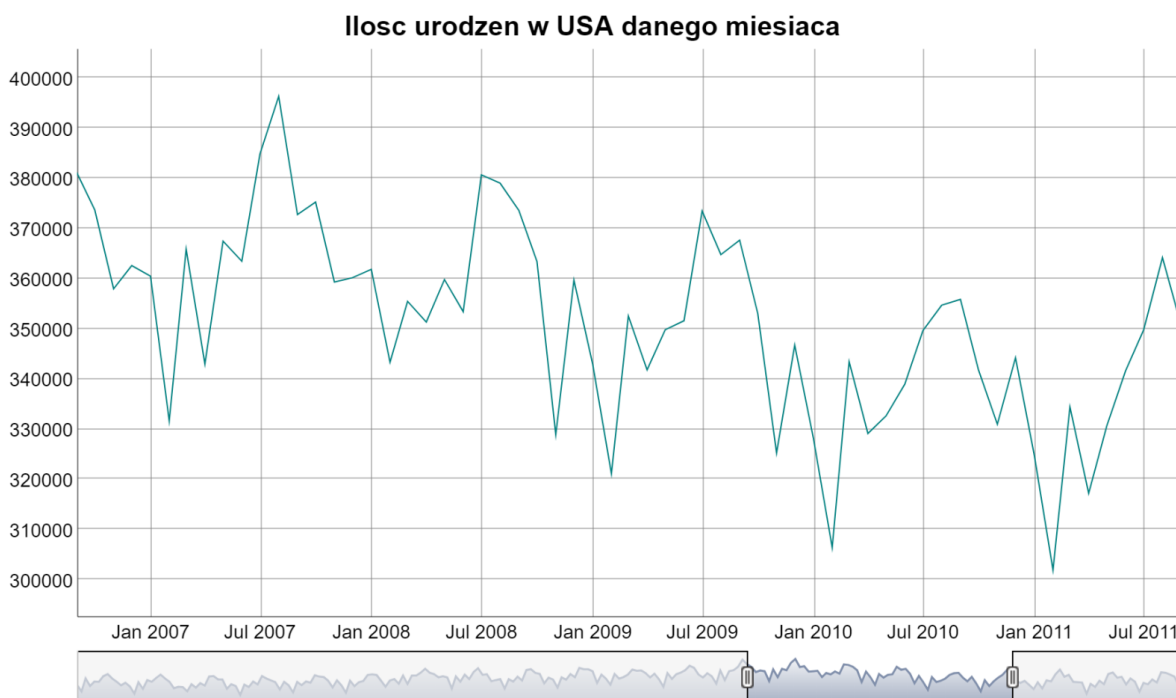
Wizualizacja

Paczka ta w bardzo przejrzysty sposób prezentuje podane informacje.
Oto wspomniany wykres urodzeń:



Wykres jest interaktywny. Możemy, po najechaniu na krzywą, odczytać dokładną wartość zaprezentowaną w prawym górnym rogu.

Ponadto dostajemy ciekawą opcję wyboru pożądanego przez nas przedziału czasowego za pomocą suwaka na dole:



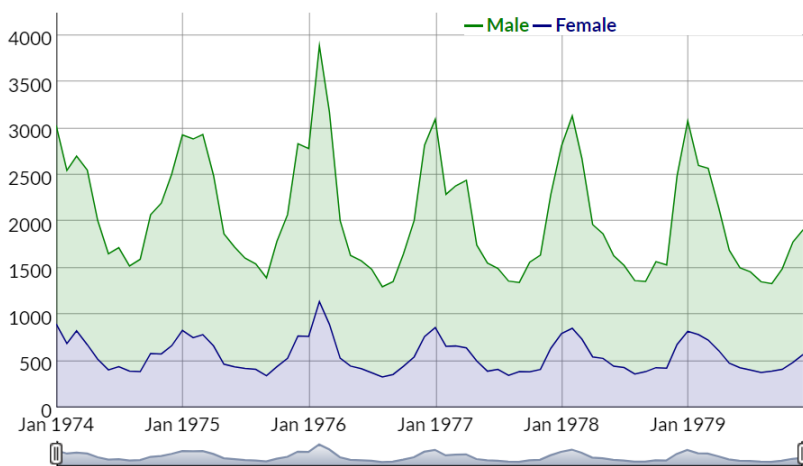
Opis użytej paczki

Dygraphs jest to szybka, elastyczna biblioteka wykresów. Pozwala ona eksplorować i interpretować gęste zbiory danych w łatwy i przejrzysty sposób. Najlepiej spisuje się w prezentowaniu pewnych danych od czasu.

Paczka ta posiada bardzo ciekawą możliwość predykcji zachowań pojawiających się w przyszłości, polegającą na znajdowaniu pojawiających się schematów we wprowadzonych danych z przeszłości lub/i teraźniejszości.

Oto wizualizacja tego mechanizmu (**NIE mojego autorstwa**), predykcji ilości śmierci chorobami płuc od czasu:

```
dygraph(lungDeaths) %>%  
  dySeries("mdeaths", label = "Male") %>%  
  dySeries("fdeaths", label = "Female") %>%  
  dyOptions(stackedGraph = TRUE) %>%  
  dyRangeSelector(height = 20)
```



```
hw <- HoltWinters(ldeaths)  
predicted <- predict(hw, n.ahead = 72, prediction.interval = TRUE)  
  
dygraph(predicted, main = "Predicted Lung Deaths (UK)") %>%  
  dyAxis("x", drawGrid = FALSE) %>%  
  dySeries(c("lwr", "fit", "upr"), label = "Deaths") %>%  
  dyOptions(colors = RColorBrewer::brewer.pal(3, "Set1"))
```

Predicted Lung Deaths (UK)

