

Eksploracji danych

Lista 1

1. Sprawdź dostępne pliki danych w pakiecie datasets (`data(package="datasets")`) i w pakietach forecast i expsmooth. Sprawdź dostępne funkcje w pakiecie forecast (`library(help="forecast")`)
Zobacz dane na stronie
<http://www-eio.upc.edu/~pau/cms/rdata/datasets.html>.
2. Wczytaj dane z danego pakietu (`data()`), sprawdź ich format (`class()`), długość (`length()`), częstotliwość (`frequency()`), liczba obserwacji na jednostkę czasu), początek (`start()`), koniec danych (`end()`) i obejrż całe dane. Sprawdź funkcje `deltat()` (odstęp pomiędzy obserwacjami), `time()` (wektor punktów czasowych), `cycle()` (pozycja poszczególnych obserwacji w cyklu).
3. Zrób wykres danych, podpisz osie i dodaj legendę (`plot()`).
4. Utwórz plik tekstowy z danymi (lub skopiuj dane). Wczytaj plik do R (`scan()`) i sprawdź jego rodzaj (`class()`). Przekształć plik w plik ts o częstotliwości 4 i zaczynający się od I kwartału 2023. Przekształć w plik numeryczny i tekstowy. (`ts()`, `as.numeric()`, `as.character()`).
5. Ze strony <http://stat.gov.pl/bdl/> ściągnij plik z danymi w formacie .csv i wczytaj go do R (`read.csv()`) lub (`read.table()`). Następnie zapisz plik w formacie .ts (`as.ts()`). Zapisz go jako plik o częstotliwości 4 lub 12 (`ts()`).
6. Znajdź plik z danymi w formacie .txt, przekształć do formatu .ts i sprawdź jego podstawowe własności.
7. Wczytaj dany plik .ts i sprawdź czy jest to plik .ts (`class()`) lub (`is.ts()`) jego podstawowe własności (`tsp()`) lub (`attributes()`).
8. Utwórz plik numeryczny za pomocą funkcji (`rnorm()`) o długości 72 (liczby losowe z rozkładu normalnego) i utwórz pliki .ts o częstotliwości odpowiednio 12, 4 i 1 zaczynające się odpowiednio od stycznia 2010, I kwartału 2010 i 2010 za pomocą funkcji (`ts()`). Jaka jest różnica pomiędzy funkcją (`ts()`) i (`as.ts()`)?
9. Za pomocą funkcji (`data.frame()`) stwórz dane klasy ramka danych (data frame). Sprawdź klasę danych i klasę poszczególnych kolumn

- ($\$$). Dodaj nowy wiersz. Dodaj nową kolumnę. Sprawdź liczbę kolumn, wymiar ramki, wyświetl 2 pierwsze wiersze i wyświetl 2 ostatnie wiersze. Wyświetl skrócony opis typów danych `str{utils}`.
10. Wczytaj dane `Airpassengers` i za pomocą funkcji (`window()`), utwórz dane zaczynające się od stycznia 1950 i kończące się na grudniu 1957 i narysuj ich wykres.
 11. Wczytaj plik `usgdp` z biblioteki `expsmooth` i za pomocą funkcji (`window()`) podziel go na tzw. część uczącą i testową (na ogół w proporcji 70% i 30%). Narysuj wykresy części testowej i uczącej na jednym rysunku innym typem linii i opisz (zobacz `ts.plot{stats}`).
 12. Za pomocą funkcji (`getSymbols(Symbols, src, from, to){quantmod}`) ściągnij notowania indeksu DJIA (funkcja znajduje się w bibliotece (`quantmod`), (`install.packages("", dependencies=TRUE)`)) np. ze strony YAHOO. Podobnie pobierz kurs EURO/USD (`EURUSD=X`) i kurs bitcoina w USD (`BTC-USD`). Wyświetl 4 pierwsze wiersze. Narysuj wykresy np. kursów zamknięcia. Sprawdź klasę (typ) danych.
 13. Wykonaj to samo co w poprzednim zadaniu dla funkcji (`get.hist.quote()/tseries`) (np. `get.hist.quote(instrument="IBM", quote=c("Open", "High"), provider="yahoo", start="2015-07-01", end="2020-08-06")`) z pakietu `tseries`.
 14. Załaduj dane `Airpassengers` i narysuj dla nich wykresy rozrzutu z przesunięciem $lag = 1, 2, \dots, 12$ `lag.plot{stats}[R]`. Zrób to samo dla obciętych danych `Airpassengers`.
 15. Wygeneruj biały szum (np. za pomocą funkcji `rnorm`) i narysuj jego wykres i wykresy rozrzutu z przesunięciem $lag = 1, 2, 3, 4$. Porównaj wykresy z wykresami z poprzedniego zadania.
 16. Zróżnicuj dane `Airpassengers` z opóźnieniem 1 i 12 i narysuj dla nich wykresy rozrzutu z przesunięciem $lag = 1, 2, \dots, 12$.
 17. Dla danych `Nile` z pakietu `datasets` narysuj wykres za pomocą funkcji `xyplot{lattice}[R]` z różną proporcją obrazu (`aspect`).
 18. Dla danych `Nile` z pakietu `datasets` narysuj wykres z trzema panelami (`cut`) i częścią wspólną 50% (`overlap`, dane `Nile` podzielone na trzy części z częścią wspólną 50%).