

Analiza Szeregów Czasowych

Prognozowanie Szeregów Czasowych

Natalia Nehrebecka

1

Cel analizy szeregów czasowych

- ▶ Podstawowym celem analizy szeregów czasowych jest:
 - **zbudowanie modelu**, który będzie dobrze opisywał dynamikę czasowa obserwowanego zjawiska
 - i który może być następnie wykorzystany do **prognozowania przyszłych (nieznanych) wartości**

Zadania analizy szeregów czasowych

- ▶ Dwa główne zadania ASC:
 - identyfikacja regularnych tendencji (tzw. dekompozycja szeregu czasowego),
 - prognozowanie.

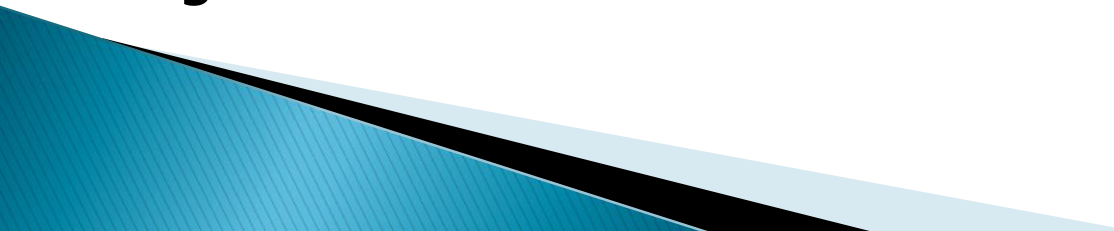
Agenda

- ▶ Zanim zaczniemy
 - Zasady zaliczenia
- ▶ Szeregi czasowe
 - Dane statystyczne
 - Podstawowe definicje
 - Praktyka

Agenda

- ▶ Zanim zaczniemy
 - Zasady zaliczenia
- ▶ Szeregi czasowe
 - Dane statystyczne
 - Podstawowe definicje
 - Praktyka

Sprawy organizacyjne

- adres mailowy: **nnehrebecka@wne.uw.edu.pl**
 - strona internetowa: **www.ekonometria.wne.uw.edu.pl**
 - dyżur: **po uzgodnieniu mailowym**
 - **3 nieobecności**
- 

Zasady zaliczenia

- ▶ Zajęcia mają formę konwersatorium.
- ▶ Forma zaliczenia:
 - raport z badania - **100 punktów**,
- ▶ warunek dodatkowy: należy uzyskać co najmniej **50 punktów**.
- ▶ Obecność podczas zajęć jest obowiązkowa.

Projekt zaliczeniowy

- ▶ Praca zaliczeniowa jest pracą samodzielną.
- ▶ Polega na modelowaniu i prognozowaniu dwóch szeregów czasowych:
 - niesezonowego
 - sezonowego.
- ▶ Celem projektu jest dokonanie:
 - **dekompozycji szeregu czasowego,**
 - zbadanie stacjonarności/sezonowości szeregu,
 - **dopasowanie odpowiedniego modelu z klasy ARIMA/SARIMA,**
 - dokonanie prognozy z modelu klasy ARIMA (prognoza krótkookresowa: 2-3 okresy),
 - od pewnego momentu prognozy w modelach ARIMA zaczynają zbiegać do średniej i najlepiej prognozować za pomocą tych modeli na tyle okresów do przodu ile wynosi rząd modelu
 - dokonanie prognozy z modelu klasy SARIMA (prognoza obejmująca minimum 1 pełen cykl: 4 obserwacje dla danych kwartalnych, oraz 12 obserwacji dla danych miesięcznych).
 - **oraz prognozy za pomocą modeli ekstrapolacyjnych.**

Projekt zaliczeniowy

- ▶ Publikacje online mogą stanowić wartość dodaną do profesjonalnego portfolio!
 - Rpubs
 - https://rmarkdown.rstudio.com/authoring_basics.html

Ramowy plan na semestr

I. Definicja szeregu czasowego

II. Dekompozycja szeregu czasowego

III. Modele ekstrapolacyjne

IV. Prognozy i miary precyzji prognoz

V. Definicja stacjonarności, badanie stacjonarności, integracja

VI. Modele ARMA, ARIMA, SARIMA



Literatura

- ▶ Enders W. (2014), *Applied Econometric Time Series*, Wiley.
- ▶ Ghysels E., Osborn D.R. (2001), *The econometric analysis of seasonal time series*, Cambridge University Press.
- ▶ Pawełek B., Wanat S., Zeliaś A. (2003/2013), *Prognozowanie ekonomiczne – Teoria, przykłady, zadania*, PWN.

Literatura

- ▶ Zagdanski & Suchwałko (2016) Analiza i prognozowanie szeregów czasowych: praktyczne wprowadzenie na podstawie środowiska R, Wydawnictwo Naukowe PWN
- ▶ Cowpertwait & Metcalfe (2009) Introductory Time Series With R, Springer
- ▶ Biecek (2016) Przewodnik po pakiecie R
- ▶ Gagolewski (2014) Programowanie w języku R
- ▶ Cryer & Chan (2008) Time Series Analysis: With Applications in R, Springer
- ▶ Kirchgassner & Wolters & Hassler (2007) Introduction to Modern Time Series Analysis, Springer
- ▶ Shumway & Stoffer (2011) Time Series Analysis and Its Applications: With R Examples, Springer

Agenda

- ▶ Zanim zaczniemy
 - Zasady zaliczenia
- ▶ Szeregi czasowe
 - Dane statystyczne
 - Podstawowe definicje
 - Praktyka

Formy danych statystycznych

- ▶ Dane przekrojowe (***cross sectional data***)
 - wiele obiektów obserwowanych w jednej jednostce czasu.
- ▶ Szeregi czasowe (***time series data***)
 - jeden obiektów obserwowany w wielu jednostkach czasu.
- ▶ Dane panelowe (***panel data, time series cross sectional data***)

Dane panelowe

- ▶ Przykładem może być PKB per capita w poszczególnych krajach UE, w kolejnych latach.
- ▶ Każda obserwacja w zbiorze panelowym jest indeksowana podwójnie:
 - Po jednostkach
 - Po czasie

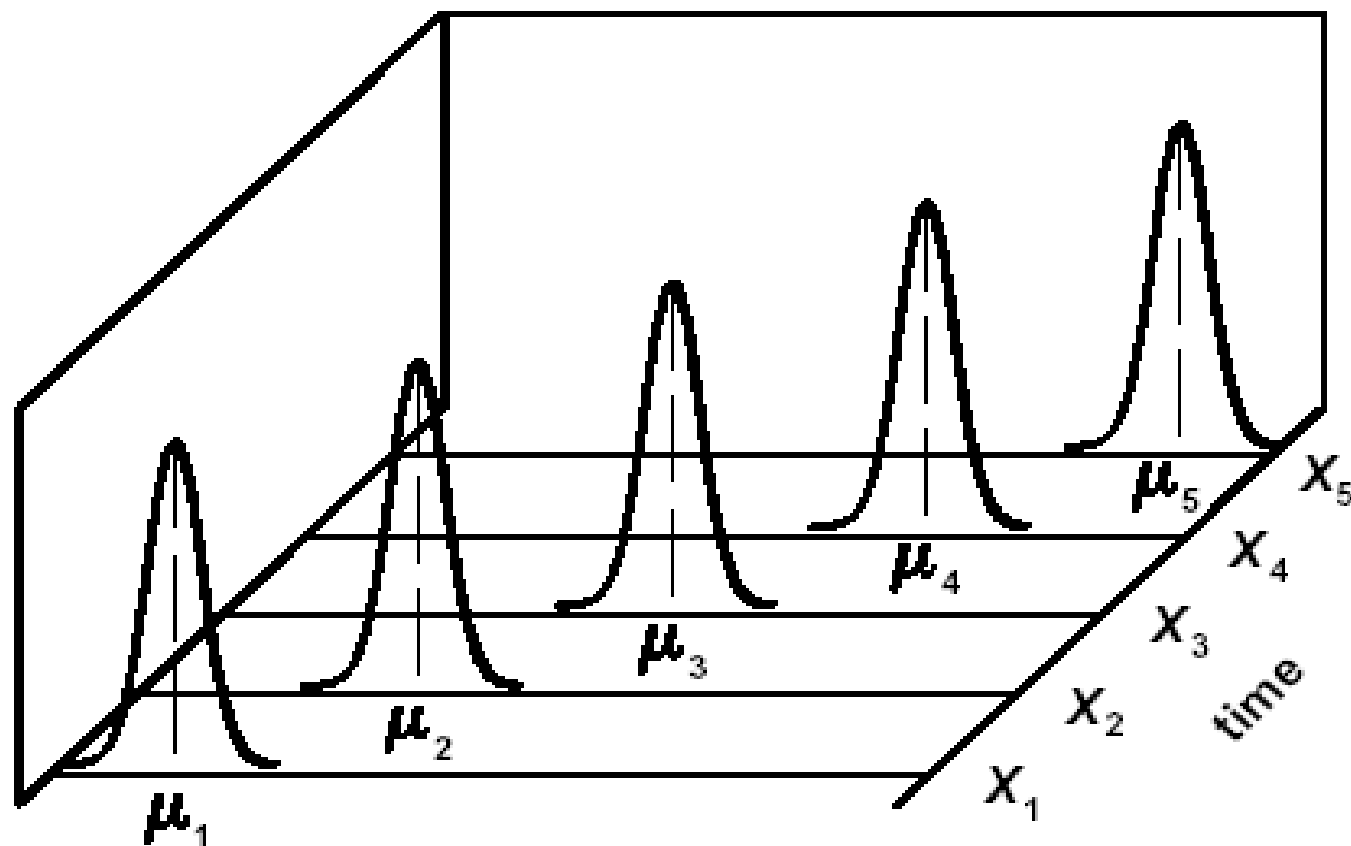
kraj	rok	y	x
Austria	2000	y	x
Austria	2001	y	x
Austria
Austria	2005	y	x
Belgia	2000	y	x
Belgia	2001	y	x
...
Belgia	2005	y	x
...
W. Brytania	2000	y	x
W. Brytania	2001	y	x
...
W. Brytania	2005	y	x

Proces stochastyczny

- ▶ Przez proces stochastyczny rozumiemy rodzinę zmiennych losowych o wartościach rzeczywistych, indeksowaną przez t , gdzie t oznacza czas.
- ▶ Proces stochastyczny oznaczamy jako zbiór $\{X_t\}$.
- ▶ Każdy element X_1, X_2, \dots, X_t procesu stochastycznego $\{X_t\}$ jest zmienną losową.

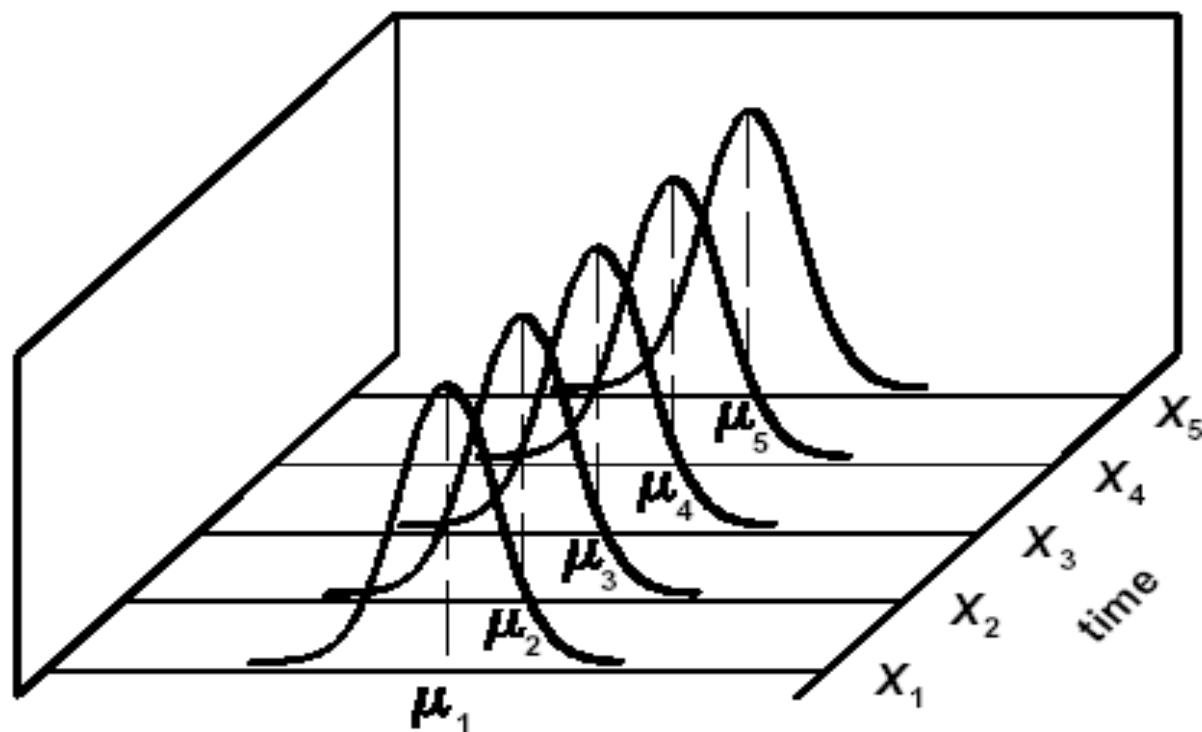
Proces stochastyczny

- ▶ Proces stochastyczny z rosnącą średnią:



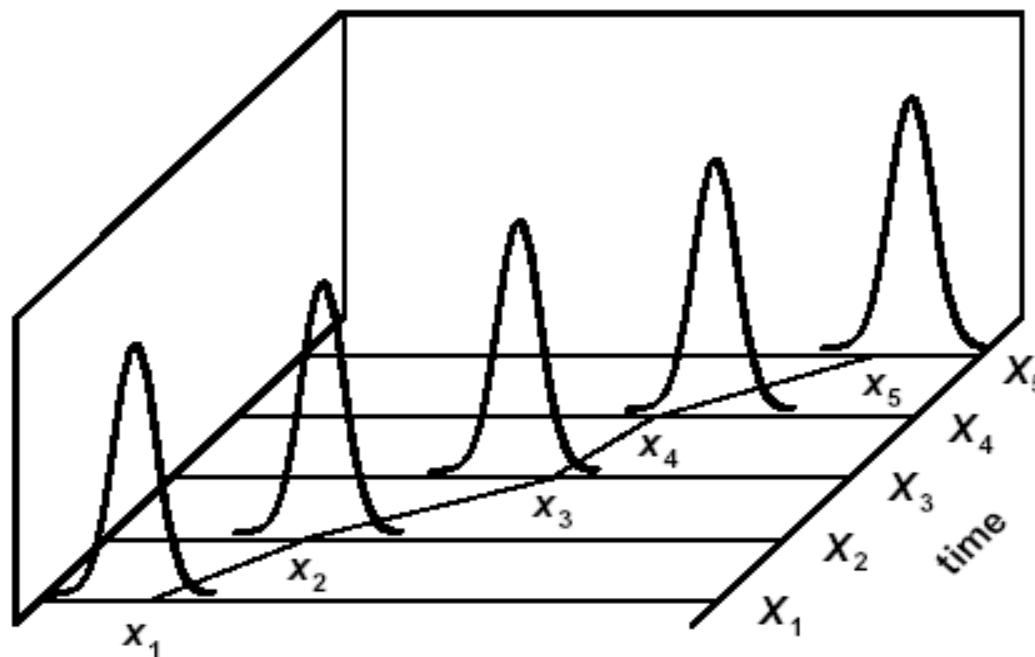
Proces stochastyczny

- ▶ Proces stochastyczny ze stałą średnią:



Realizacja procesu stochastycznego

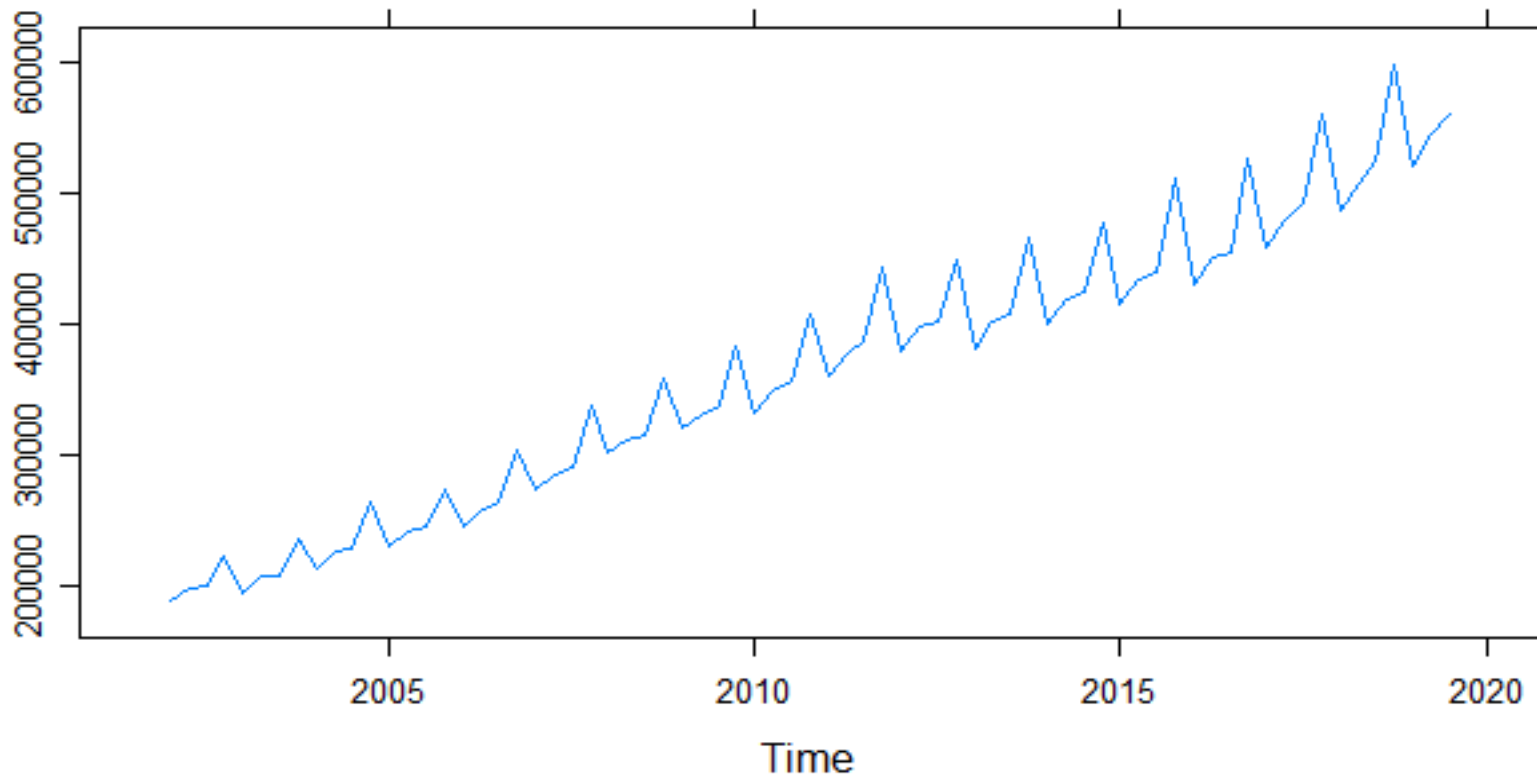
- Szereg czasowy – pojedyncza realizacja procesu stochastycznego



- Dla uproszczenia często utożsamia się szereg czasowy z procesem stochastycznym.

Szereg czasowy

- ▶ PKB w Polsce w mln PLN (dane kwartalne)



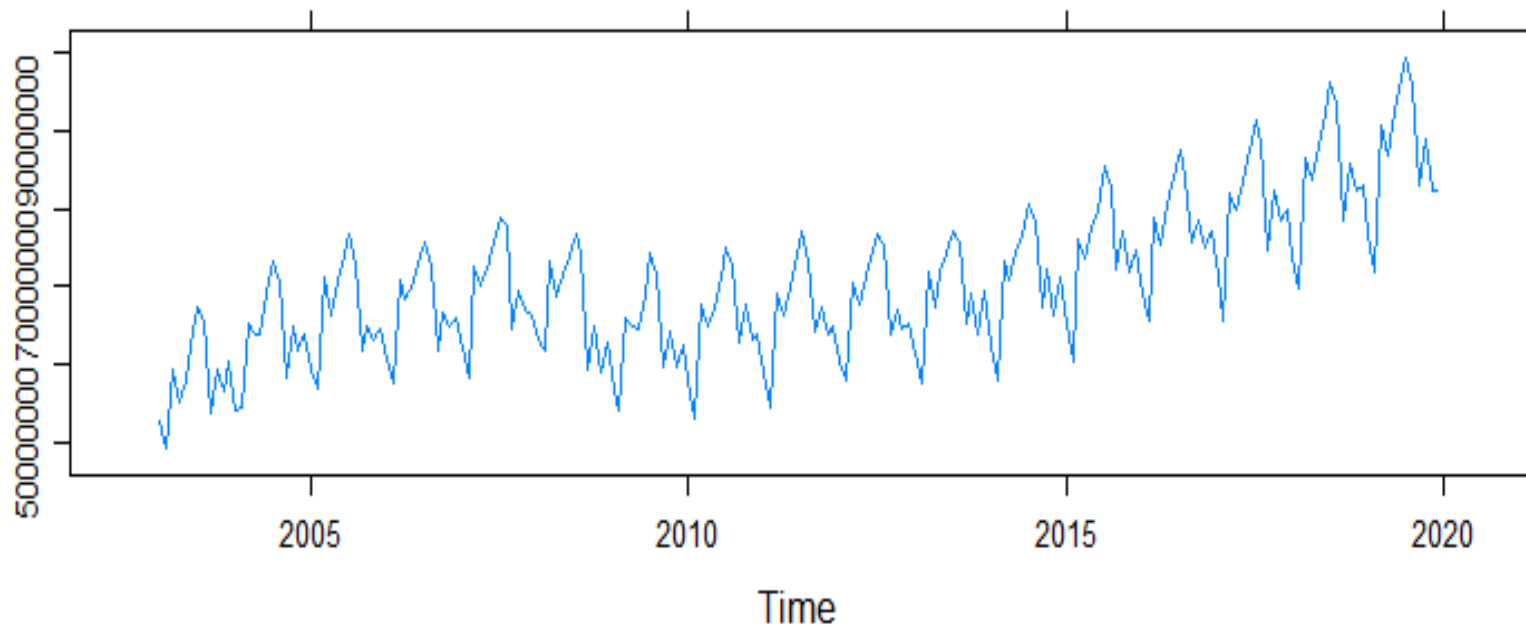
Szereg czasowy

- ▶ WIG20 – cena zamknięcia (dane dzienne)



Szereg czasowy

- ▶ Liczba pasażerów linii lotniczych w USA (dane miesięczne)



RStudio

- ▶ Po uruchomieniu systemu R kolejne **pakiety** można zainstalować funkcją **install.packages()**.
- ▶ Poniższe polecenie instaluje pakiet Rcmdr wraz z pakietami zależnymi, wymaganymi do jego działania:
 - `install.packages("Rcmdr", dependencies = TRUE)`

RStudio

- ▶ Wszystkie **pakiety** są wgrywane jako podkatalogi do katalogu **library** na dysku twardym komputera.
 - Po uruchomieniu platformy R ładowane są pakiety podstawowe takie jak: base, graphics, stats, itp.
- ▶ Aby skorzystać z dodatkowych funkcji, należy załadować **(włączyć) pakiet**, w którym się one znajdują (po jego zainstalowaniu).
 - Pakiety włącza się poleceniem **library()**. Poniższa instrukcja włącza pakiet Rcmdr:
 - library(Rcmdr)
 - Gdyby ten pakiet nie był zainstalowany, to pojawiłby się poniżej przedstawiony komentarz:
 - **Error in library(Rcmdr) : there is no package called 'Rcmdr'**

RStudio

- ▶ Wyświetlamy pomoc dotyczącą funkcji plot():
 - `?plot`
- ▶ Wyświetlamy przykłady użycia funkcji plot():
 - `example(plot)`
- ▶ Wyświetlamy funkcje ze słowem "test" w nazwie:
 - `apropos("test")`
- ▶ Wyświetlamy nazwy funkcji ze zwrotem "normality test" w opisie:
 - `help.search("normality test")`

Rstudio – ts objects

- ▶ Dane o regularnych odstępach czasu, które są oddzielone stałym przedziałem czasu
 - Takie dane są zwykle obserwowane miesięcznie, kwartalnie lub corocznie,
- ▶ **ważne ograniczenie obiektów *ts***
 - Klasa *ts* jest raczej ograniczona, szczególnie w przypadku reprezentacji danych finansowych, które nie są regularnie rozmieszczone
 - Na przykład klasa *ts* nie może być używana do reprezentowania dziennych danych finansowych, ponieważ takie dane są obserwowane tylko w dni robocze
 - działa od poniedziałku do piątku, z pominięciem weekendów
 - Zatem dane są równomiernie rozłożone w czasie w ciągu tygodnia, ale odstęp między piątkiem a poniedziałkiem jest inny
 - Tego typu nieregularnych odstępów nie można przedstawić za pomocą klasy *ts*

Rstudio – xts objects

- ▶ Bardzo elastyczna klasa dla szeregów czasowy
- ▶ Klasa została zaprojektowana do obsługi danych szeregów czasowych z dowolnie uporządkowanym indeksem czasowym
- ▶ Indeks ten może być regularnie rozmieszczoną sekwencją dat, nieregularnie rozmieszczoną sekwencją dat lub indeksem liczbowym

Dziękuję za uwagę