

# Dekompozycja szeregów czasowych

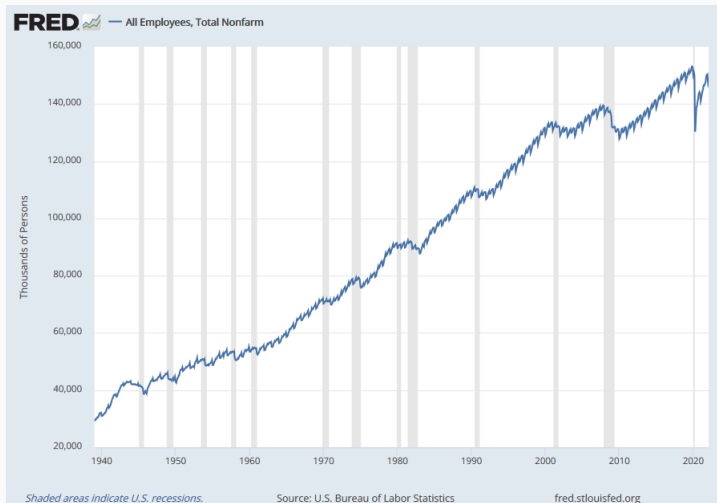
ASC 2023

---

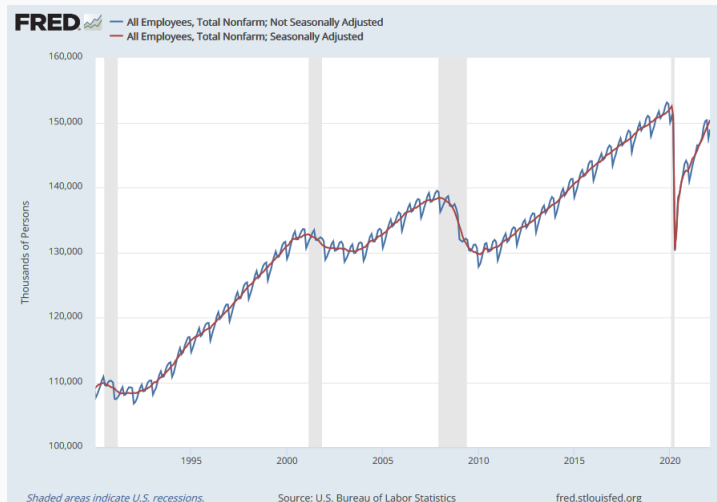
Piotr Żoch

- Składowe szeregi czasowych.
- Metody wyrównywania sezonowego.

# Składowe szeregów czasowych



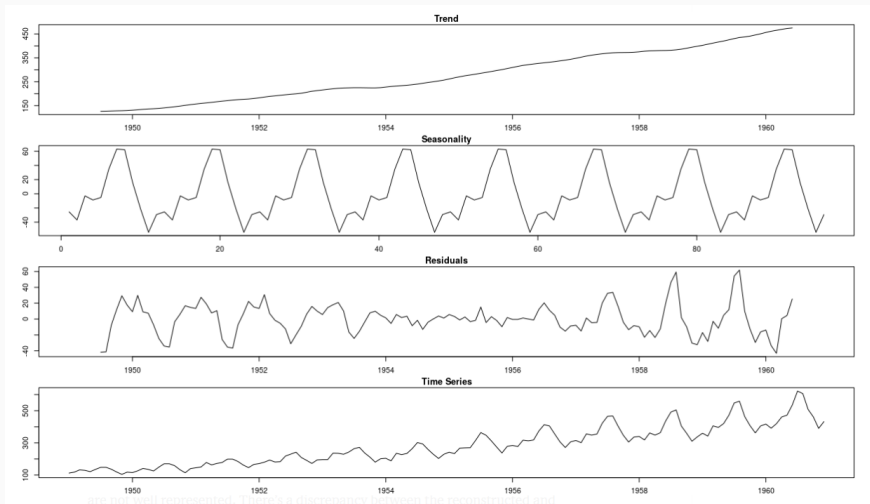
# Składowe szeregi czasowych



- Składowa przypadkowa.
- Składowa systematyczna (efekt oddziaływań stałego zestawu czynników na zmienną prognozowaną):
  - Trend (tendencja rozwojowa).
  - Stały poziom (przeciętny długookresowy poziom zmiennej).
  - Składowa okresowa:
    - Wahania sezonowe.
    - Wahania cykliczne.

- **Tendencja rozwojowa (trend)**
  - długookresowa skłonność do jednokierunkowych zmian (wzrostu lub spadku) wartości badanej zmiennej.
- **Stały (przeciętny) poziom**
  - występuje, gdy w szeregu czasowym nie ma tendencji rozwojowej, wartości zmiennej prognozowanej oscylują zaś wokół pewnego (stałego poziomu).
- **Wahania cykliczne**
  - długookresowe rytmiczne wahania wartości wokół tendencji rozwojowej lub stałego (przeciętnego) poziomu.
- **Wahania sezonowe**
  - wahania wartości obserwowanej zmiennej wokół tendencji rozwojowej lub stałego (przeciętnego) poziomu tej zmiennej. Mają skłonność do powtarzania się w określonym czasie.

# Składowe szeregów czasowych



- Wyodrębnienie poszczególnych składników nie zawsze jest łatwe z uwagi na rozmaite interakcje, które zachodzą pomiędzy nimi.
- Różne metody dekompozycji (średnie ruchome, wydzielanie komponentów o różnej częstotliwości, procedury TRAMO/SEATS itd).



- Trend
  - deterministyczny

$$y_t = \alpha + \beta t + \epsilon_t$$

$$y_t = \alpha + \beta t + \gamma t^2 + \epsilon_t$$

$$y_t = \kappa e^{\delta t} e^{\epsilon_t}$$

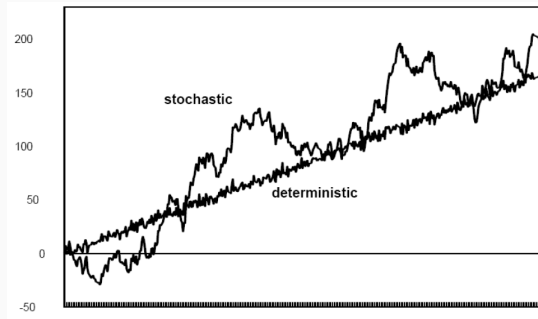
$$\epsilon_t \sim iid(0, \sigma_\epsilon^2)$$

- stochastyczny

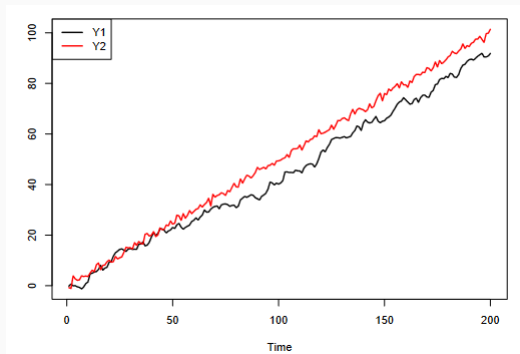
$$y_t = \alpha + y_{t-1} + \epsilon_t$$

$$\epsilon_t \sim iid(0, \sigma_\epsilon^2)$$

# Trend



- W praktyce rozróżnienie może nie być łatwe...



- Dwa typy sezonowości:
  - **Deterministyczna**
    - Wartości obserwacji szeregu cechującego się sezonowością deterministyczną wahają się z amplitudą względnie stałą.
    - Proces, którego bezwarunkowa średnia zależy od podokresu roku (np. miesiąca, kwartału)
    - Ten rodzaj sezonowości jest modelowany za pomocą zmiennych zerojedynkowych, ich liczba zależna od liczby okresów w roku.
  - **Stochastyczna**
    - Charakteryzuje się zmiennym w czasie wzorcem sezonowości.

- Deterministyczna sezonowość (przykład):

$$y_t = \sum_{s=1}^S \delta_{t,s} m_s + \epsilon_t$$

$\delta_{t,s} = 1$  jeśli okres  $t$  przypada na sezon  $s$  (np.  $\delta_{t,12} = 1$  jeśli okres  $t$  to grudzień w przypadku danych miesięcznych).

- Stochastyczna sezonowość (przykład – sezonowy pierwiastek jednostkowy):

$$y_t = y_{t-S} + \epsilon_t$$

- Punktem wyjścia jest założenie, że szereg czasowy można zdekomponować na elementy:
  - **czynnik sezonowy**  $S_t$  – regularne fluktuacje sezonowe;
  - **cykl i trend**  $C_t$  – długookresowy trend, cykl koniunkturalny i inne składniki cykliczne;
  - **trading-day**  $TD_t$  – kształt kalendarza – korekta ze względu na różną długość miesięcy, liczbę dni roboczych itp.; nie liczony dla danych kwartalnych
  - **składnik losowy**  $I_t$
- Chcemy uzyskać dane wyrównane sezonowo  $SA_t$

- Postać addytywna

$$Y_t = S_t + C_t + TD_t + I_t$$

$$SA_t = C_t + I_t$$

czyli amplituda wahań sezonowych i losowych jest stała.

- Postać multiplikatywna

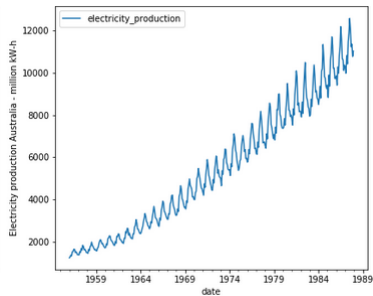
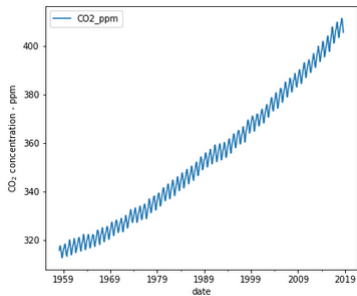
$$Y_t = S_t \times C_t \times TD_t \times I_t$$

$$SA_t = C_t \times I_t$$

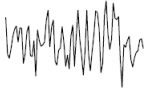
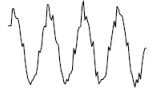

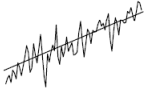
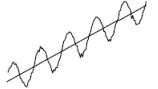
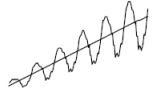
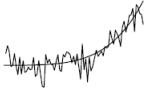
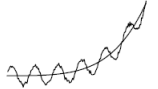
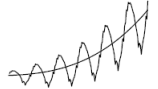
czyli amplituda wahań sezonowych i losowych zależy od trendu.



# Postać addytywna vs postać multiplikatywna



# Postać addytywna vs postać multiplikatywna

Sezonowość → Trend ↓	Brak	Sezonowość addytywna	Sezonowość multiplikatywna
Brak			
Trend addytywny			
Trend multiplikatywny			

- O charakterze jednorazowym (*additive outlier (AO)*) powodujące zmianę wartości zmiennej zależnej tylko w jednym okresie

$$AO_t^{(t_0)} = \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}$$

- O charakterze przejściowym (*temporary change (TC)*), powodujące tymczasowe przesunięcie poziomu zmiennej zależnej, przy czym powrót ze skokowej zmiany wartości zmiennej zależnej w do poziomu pierwotnego następuje zgodnie z funkcją wykładniczą w postaci

$$TC_t^{(t_0)} = \begin{cases} 1 & t = t_0 \\ \alpha^{t-t_0} & t \neq t_0 \end{cases}$$

- O charakterze długotrwałym (*level shift (LS)*), powodujące długotrwałe przesunięcie poziomu zmiennej zależnej

$$LS_t^{(t_0)} = \begin{cases} -1 & t < t_0 \\ 0 & t \geq t_0 \end{cases}$$

- Dwie popularne metody to:
  - **X-12-ARIMA** (obecnie X-13-ARIMA-SEATS)
  - **TRAMO/SEATS**
- Obie dostępne w programie **Demetra+** (opracowany przez Eurostat, darmowy).

- Metoda ad hoc, dobór filtrów nie zależy od właściwości statystycznych szeregu czasowego.
- Nieco myląca nazwa, model ARIMA jest w niej wykorzystywany jedynie do estymacji wartości prognozowanych.
  1. Trend obliczany za pomocą średnich ruchomych.
  2. Odjąć trend od szeregu, pozostaje komponent sezonowy i nieregularny.
  3. Komponent sezonowy obliczany za pomocą średnich ruchomych.
  4. Odjąć komponent sezonowy od szeregu.
  5. Ponowić (1) i (2) z nieco innymi wagami.
  6. Ponowić (3) i (4) z nieco innymi wagami.
  7. Otrzymać trend i komponent nieregularny odejmując uzyskany w (6) komponent sezonowy od szeregu.

- **TRAMO** ("Time Series Regression with ARIMA Noise, Missing Observations and Outliers"): dopasowuje proces SARIMAX do szeregu i oczyszcza szereg z wpływu wybranych czynników X (kalendarzowych, nietypowych...) za pomocą tego modelu.
- **SEATS** ("Seasonal Extraction in ARIMA Time Series"): korzysta ze specyfikacji SARIMA, by zdekomponować szereg na elementy: trendu, sezonowości, nieregularny.