# Wprowadzenie

ASC 2025

Piotr Żoch

# Plan

- Sprawy organizacyjne
- Szereg czasowy: sezonowość, stacjonarność, integracja.

# Sprawy organizacyjne

## Sprawy organizacyjne

- **Email:** p.zoch@uw.edu.pl.
- **Dyżur:** uzgadniany indywidualnie.
- **Forma zajęć:** konwersatorium
- **Slajdy i inne materiały:** Github
- **Zaliczenie:** projekt zaliczeniowy (50%) + egzamin (50%)

# Projekt zaliczeniowy

- Dokładne wytyczne i przykładowe projekty: http://www.ekonometria.wne.uw.edu.pl
- Indywidualna analiza dwóch szeregów czasowych (dane miesięczne/kwartalne/roczne)
  - Szereg sezonowy.
  - Szereg niesezonowy.
- Dekompozycja szeregu, dobranie modelu ARIMA/SARIMA, prognozy z modelu.
- Termin: **8 czerwca, 23:59** (termin poprawkowy: 1 września 23:59)
- Wybrane szeregi należy ze mną skonsultować do **6 kwietnia**.
- Kto pierwszy, ten lepszy!

# Literatura

- Z USOSa:
  - **Enders W. (2014), Applied Econometric Time Series, Wiley.**
  - Pawełek B., Wanat S., Zeliaś A. (2003/2013), Prognozowanie ekonomiczne – Teoria, przykłady, zadania, PWN.
- Klasyk:
  - **Hamilton, J. (1994), Time Series Analysis, Princeton University Press.**

# Szeregi czasowe

# Dane

- **Dane przekrojowe** *(cross sectional data)* - wiele jednostek $n = 1, 2, ..., N$ obserwowanych w jednym okresie
- **Szeregi czasowe** *(time series data)* - jedna jednostka obserwowana w wielu okresach $t = 1, 2, ..., T$
- **Dane panelowe** *(panel data)* - wiele jednostek $n = 1, 2, ..., N$ obserwowanych w wielu okresach $t = 1, 2, ..., T$

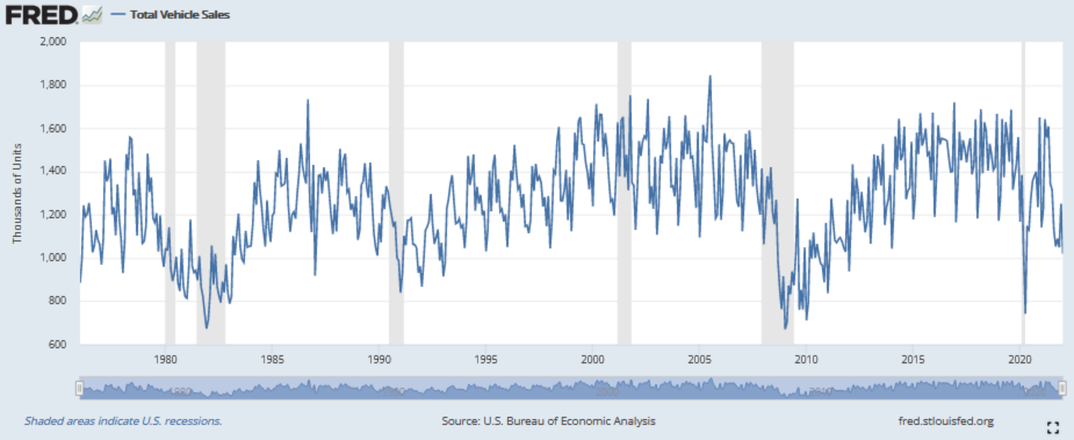
# Szereg czasowy

- **Proces stochastyczny:** zbiór zmiennych losowych $\{Y_t\}_{t \in T}$, które przyjmują wartości w przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ i są indeksowane przez zbiór $T$.
  - Przykład: $Y_t \sim N\left(0, \sigma^2\right)$
  - Przykład: $Y_t \sim N\left(t, \sigma_t^2\right)$
- **Szereg czasowy:** pojedyncza realizacja pewnego procesu stochastycznego.

$$\{y_t\}_{t \in T}$$

  - Ciąg obserwacji pokazujący kształtowanie się badanego zjawiska w kolejnych okresach (dniach, miesiącach, kwartałach, latach, itp.)

# Szereg czasowy

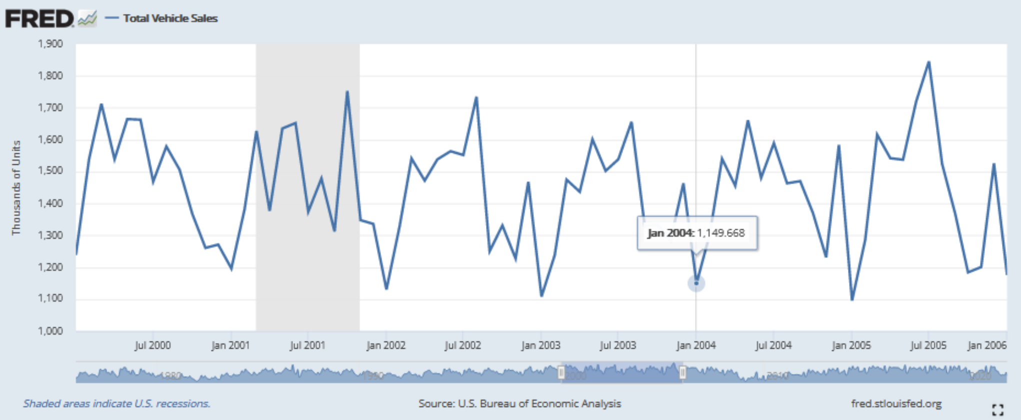


# Praca z szeregami czasowymi

- Po co?
  - Prognozowanie przyszłych wartosci
  - Badanie dynamiki szeregów
  - Dodatkowo: pewne właściwości szeregów czasowych ważne nawet przy zwykłej regresji liniowej
- Co robić?
  - Analiza czasowa (ten kurs)
  - Analiza częstotliwościowa (może o tym wspomnimy)

# Sezonowość

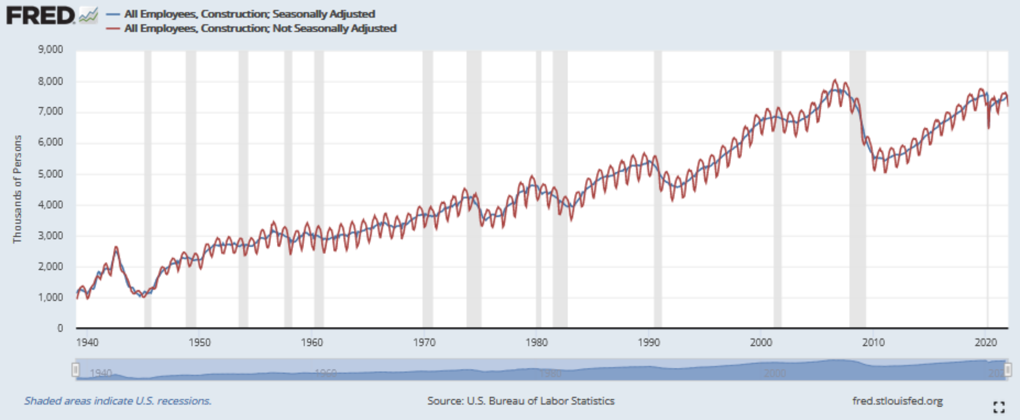
- **Sezonowość:** zmienność związana z cyklem kalendarzowym:
  - Przykład: dla zmiennych miesięcznych — sezonowość miesięczna.



## Sezonowość

- Jeśli sezonowość ma wpływ na związek między zmienną objaśniającą a objaśnianą i nie zostanie uwzględniona w modelu, to pojawi się w resztach (nie będą spełniały założeń klasycznego modelu regresji liniowej).
- Co robić?
  - Pracować z danymi wyrównanymi sezonowo (różne metody, które poznamy na tym kursie).
  - Potraktować sezonowość poważnie i uwzględnić ją w modelu.

# Sezonowość



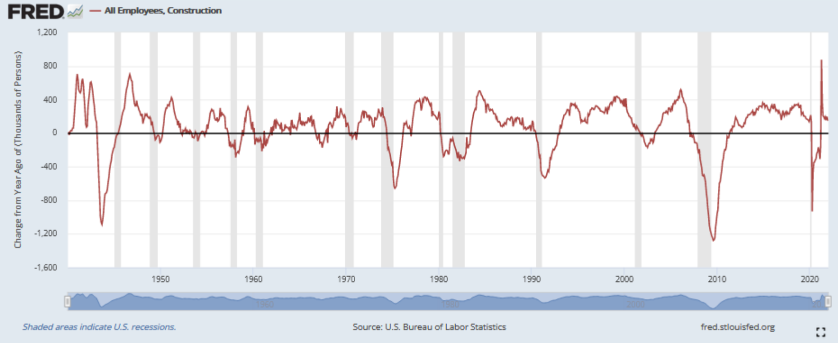
FRED
— All Employees, Construction; Seasonally Adjusted
— All Employees, Construction; Not Seasonally Adjusted
Thousands of Persons
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
1940
1950
1960
1970
1980
1990
2000
2010
2020
Shaded areas indicate U.S. recessions.
Source: U.S. Bureau of Labor Statistics
fred.stlouisfed.org

# Sezonowość

- **Różnicowanie sezonowe:** prosty sposób na pozbycie się sezonowości — zamiast pierwotnych wartości zmiennych możemy wykorzystać

$$\Delta_s y_t := y_t - y_{t-s}$$

gdzie $s = 4$ dla zmiennych kwartalnych, $s = 12$ dla zmiennych miesięcznych.



# Stacjonarność

- Zmienna jest **stacjonarna** w sensie **ścisłym** jeśli łączne rozkłady $\{Y_t, Y_{t+1}, \cdots, Y_T\}$ i $\{Y_{t+\tau}, Y_{t+1+\tau}, \cdots, Y_{T+\tau}\}$ są takie same dla wszystkich $T \geq t \geq 1, \tau \geq 0$.
- Inaczej: $F(\{Y_t, Y_{t+1}, \cdots, Y_T\}) = F(\{Y_{t+\tau}, Y_{t+1+\tau}, \cdots, Y_{T+\tau}\})$
- Rozkład nie zależy od czasu.

# Stacjonarność

- Zmienna jest **stacjonarna** w sensie **słabym** *(stacjonarność kowariancyjna)* jeśli spełnione są trzy warunki:
  - $E[Y_t] = \mu < \infty$
  - $Var[Y_t] = \sigma^2 < \infty$
  - $Cov(Y_t, Y_{t+h}) = Cov(Y_s, Y_{s+h}) = \gamma_h$
- Intuicyjnie: właściwości zmiennej nie zmieniają się w czasie.
- Jeśli którykolwiek warunek nie jest spełniony, zmienna jest **niestacjonarna**.
- W praktyce będziemy badać stacjonarność w sensie słabym.

# Stacjonarność

- Przykład: biały szum (*white noise*): $Y_t$ jest i.i.d. z wartością oczekiwaną równą 0 i wariancją $\sigma^2$.
  - *independent and identically distributed*
- Stacjonarność
  - $E[Y_t] = 0 < \infty$
  - $Var[Y_t] = \sigma^2 < \infty$
  - $Cov(Y_t, Y_s) = 0$ dla $t \neq s$.
- Tak.

# Stacjonarność

- Przykład: model AR(1)

$$y_t = \rho y_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$$
$$|\rho| < 1$$

- Stacjonarność? Sprawdźmy trzy warunki (słabej) stacjonarności.

# Stacjonarność

- Warunek pierwszy: $E\left[y_t\right] = \mu < \infty$.
- Zauważmy, że

$$\begin{aligned} E\left[y_t\right] &= E\left[\sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i}\right] \\ &= \sum_{i=0}^{\infty} \rho^i E\left[\varepsilon_{t-i}\right] \\ &= 0 < \infty \end{aligned}$$

# Stacjonarność

- Warunek drugi: $Var\left[y_t\right] = \sigma^2 < \infty$.
- Zauważmy, że

$$\begin{aligned} Var\left[y_t\right] &= Var\left[\sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i}\right] \\ &= \sum_{i=0}^{\infty} \rho^{2i} Var\left[\varepsilon_{t-i}\right] \\ &= \frac{\sigma_\varepsilon^2}{1-\rho^2} < \infty \end{aligned}$$

## Stacjonarność

- Warunek drugi: $Cov\left(Y_t, Y_{t+h}\right) = Cov\left(Y_s, Y_{s+h}\right) = \gamma_h$.

$$
\begin{aligned}
Cov\left[y_t, y_{t-h}\right] &= Cov\left[\sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i}, \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-h-i}\right] \\
&= Cov\left[\sum_{i=0}^{h-1} \rho^i \varepsilon_{t-i} + \rho^h \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i-h}, \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-h-i}\right] \\
&= \rho^h \sum_{i=0}^{\infty} \rho^{2i} Var\left[\varepsilon_{t-i-h,}\right] \\
&= \rho^h \frac{\sigma_\varepsilon^2}{1-\rho^2}
\end{aligned}
$$

- Wniosek?

# Stacjonarność

- Przykład: błądzenie losowe (*random walk*)

$$y_t = y_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim iid\left(0, \sigma^2_\varepsilon\right)$$

- Podobne do AR(1), ale z $\rho = 1$.
- Stacjonarność?

# Stacjonarność

- Zauważmy, że:

$$
\begin{aligned}
E\left[y_t\right] &= 0 \\
Var\left[y_t\right] &= \sum_{s=1}^{t} Var\left[\varepsilon_s\right] \\
&= t\sigma_\varepsilon^2 \\
Cov\left[y_t, y_{t-h}\right] &= \sum_{s=1}^{t-h} Var\left[\varepsilon_s\right] \\
&= \left(t-h\right)\sigma_\varepsilon^2
\end{aligned}
$$

- Stacjonarność?

# Stacjonarność

- Przykład: błądzenie losowe z dryfem

$$
\begin{aligned}
y_t &= \mu + y_{t-1} + \varepsilon_t \\
\varepsilon_t &\sim iid\left(0, \sigma_\varepsilon^2\right)
\end{aligned}
$$

- Stacjonarność?

# Stacjonarność wokół trendu

- Zmienna jest **stacjonarna wokoł trendu** *(trendostacjonarna)* jeśli odchylenie od trendu

$$y_t - E\left[y_t\right]$$

  jest stacjonarne

- Przykład: trend liniowy

$$\begin{aligned} y_t &= \alpha + \beta t + \varepsilon_t \\ E\left[y_t\right] &= \alpha + \beta t \\ y_t - E\left[y_t\right] &= \varepsilon_t \end{aligned}$$

# Zmienna zintegrowana

- Zmienna jest **zintegrowana**, jeśli jest zmienną niestacjonarną, której $d$-te różnice są stacjonarne.
    - $d$-te różnice:

$$\begin{aligned}\Delta_1 y_t &:= y_t - y_{t-1} \\ \Delta_2 y_t &:= \Delta_1 \left(\Delta_1 y_t\right) \\ \Delta_d y_t &:= \Delta_1 \left(\Delta_{d-1} y_t\right)\end{aligned}$$

- Mówimy, że zmienna $y_t$ jest zintegrowana stopnia $d$.

$$y_t \sim I\left(d\right)$$

- Zmienne stacjonarne są zintegrowane stopnia 0.

$$y_t \sim I\left(0\right)$$

# Zmienna zintegrowana

- Przykład: błądzenie losowe (*random walk*)

$$y_t = y_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$$

- Wiemy już, że $y_t$ jest zmienną niestacjonarną.
- Czy pierwsze różnice są stacjonarne?

## Zmienna zintegrowana

- W tym przypadku mamy

$$\begin{aligned}\Delta_1 y_t &= y_t - y_{t-1} \\ &= \varepsilon_t\end{aligned}$$

  czyli $\Delta_1 y_t$ to biały szum, zmienna stacjonarna.
- Konkluzja: $y_t \sim I(1)$.