

# ČASOVÉ ŘADY ZÁKLADNÍ POJMY

**PŘF UJEP, 8.3.2024**

Ing. Ondřej Vozár, Ph.D. ČSÚ Praha,  
Katedra statistiky a pravděpodobnosti, FIS, VŠE Praha  
[ondrej.vozar@czso.cz](mailto:ondrej.vozar@czso.cz)

# Obsah přednášky

- Pojem časové řady
- Dekompozice (ekonomických) časových řad
  - Trend, cyklus, sezónnost, kalendářní variace, náhodná složka
  - Aditivní a multiplikativní model dekompozice časové řady
- Základní pojmy intervenční analýzy
- Proč a oč sezónně očišťovat
- Metody k očištění kalendářních variací
- Využití výsledků dekompozice časových řad v praxi
- Přehled metod sezónního očišťování
- Krátký kvíz

Děkuji kolegyni RNDr. Radce Lechnerové, Ph.D. za souhlas s poskytnutím částí společných prezentací pro interní potřeby ČSÚ.

# Pojem časové řady

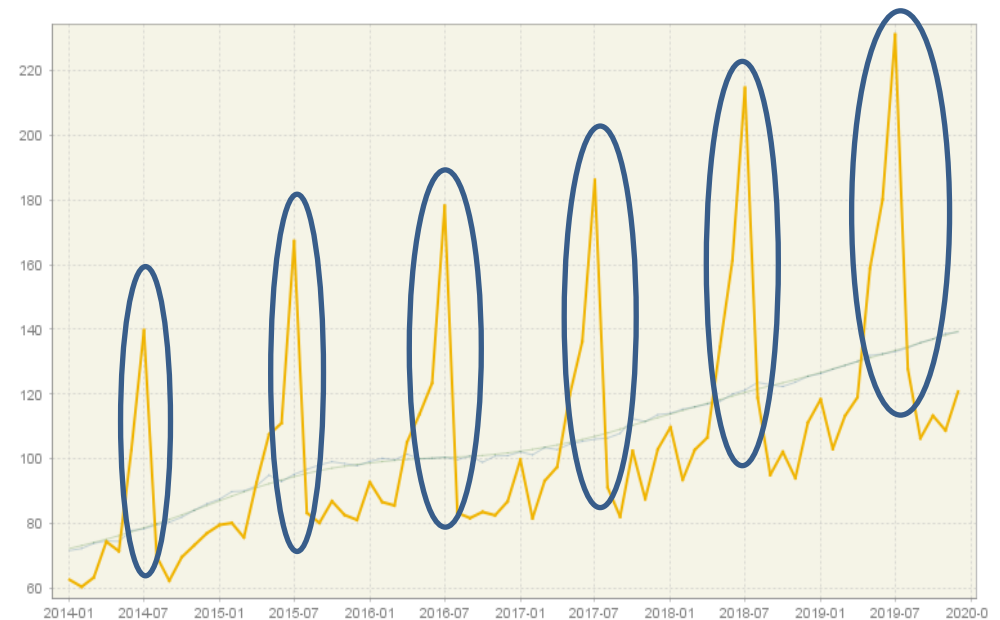
**Časová řada** je posloupnost měření nějakého jevu zpravidla prováděného v pravidelných časových intervalech

- např. denní, týdenní, měsíční, čtvrtletní, roční

**Krátkodobá (short-term) časová řada** má periodicitu nižší než 1 rok, nejčastěji **čtvrtletí, měsíc, pololetí, týden (covid-19)**...

## Ekonomické časové řady

- Náhodné chování
- Podléhají různým vnějším vlivům



# Co a oč očišťovat?

## Očišťování o kalendářní vlivy

- odstranění efektů, který souvisí s kalendářem
- Rozdílný počet pracovních dní v měsíci, státní svátky, Velikonoce a jiné pohyblivé svátky

## Sezónní očišťování

- proces odhadu a odstranění sezónních efektů z časové řady
- sezónnost = účinek, který se koná každý rok ve stejné části roku a se stejnou velikostí a směrem
- součástí sezónního očišťování je i očištění o kalendářní vlivy

## Výhoda očištěných dat

- zvyšuje srovnatelnost dat: v čase, prostoru, mezi odvětvími
- součástí je rozložení časovou řadu na několik různých složek
- užitečnost pro analýzu krátkodobého i dlouhodobého vývoje řady
- Pozn.: důležité je zajistit srovnatelnost různých měsíců/čtvrtletí v roce

# Dekompozice časových řad

Ekonomické časové řady je vhodné rozložit na tyto složky:

$$Y_t = TC_t \circ Sez_t \circ Cal_t \circ I_t,$$

kde

- $Y_t$ : pozorované (neочиštěná) časová řada
- $TC_t$ : „trendcyklus“
- $Sez_t$ : sezónní složka
- $Cal_t$ : kalendářní složka
- $I_t$ : náhodná (iregulární) složka

Je-li operace  $\circ$

a) sčítání (+)

⇒ **aditivní model**

b) násobení ( $\cdot$ )

⇒ **multiplikativní model**

# Aditivní vs. multiplikativní model

## Aditivní model

$$Y_t = TC_t + Sez_t + Cal_t + I_t,$$

- **Velikost složek** (např. sezónnost, vliv 1 pracovního dne v měsíci navíc) časové řady **nezávisí** úrovni (trendcyklu) časové řady v sledovaném období.
- Časové řady s konstantní úrovní/technické aplikace, přírodní vědy.

## Multiplikativní model

$$Y_t = TC_t \cdot Sez_t \cdot Cal_t \cdot I_t,$$

- Velikost složek (např. sezónnost, vliv 1 pracovního dne v měsíci navíc) časové řady **závisí** úrovni (trendcyklu) časové řady v sledovaném období.
- Časové řady s proměnlivou úrovní/ekonomické časové řady.

# Složky krátkodobých časových řad

## Trend-cyklus $TC_t$

### Trend

- zahrnuje dlouhodobé změny v průměrném chování řady
- zlomy/posuny úrovně se modelují v rámci intervenční analýzy pomocí tzv. posunu úrovně (level shift)

### Cyklická složka (cyklus)

- zahrnuje fluktuaci kolem trendu, ve které se střídají se fáze růstu a poklesu
- je zde možná nestejná délka fluktuací nebo periodické změny s periodou delší než 1 rok
- samostatně se obtížně odhaduje

Trend a cyklus tvoří při dekompozici společnou složku, která se nazývá trendcyklus ( $TC_t$ ). Tu publikují statistické úřady.

# Složky ekonomických časových řad

## Sezónní složka $Sez_t$

= každoročně se opakující periodické změny během jednoho kalendářního roku

### Příčiny sezónnosti

- podnebí – střídání ročních období, změny počasí  
příklady: zemědělství, spotřeba elektřiny (vytápění),...
- chování institucí – sociální zvyklosti a praktiky nebo správní pravidla  
příklady: vliv Vánoc na maloobchod, vliv fiskálního roku na některé finanční proměnné, vliv školního roku



# Složky krátkodobých časových řad

## Kalendářní složka $Cal_t$

= vlivy související s různým počtem pracovních/obchodních dní v jednotlivých měsících/čtvrtletích a přítomností státních svátků (pevně stanovených i pohyblivých).

- Produkce, tržby obvykle rostou s vyšším počtem pracovních dnů v měsíci
- V praxi se využívají přístupy založené na regresi s regresory:
  - pro národní (český) kalendář, tj. národní svátky (např. 28. října ...)
  - regresor pro pracovní dny (Working days): rozlišují se všední a víkendové dny
  - regresor pro „Trading days“: rozlišují se jednotlivé dny v týdnu (Po-So)
  - regresory pro vliv přestupného roku a Velikonoc a jiných pohyblivých svátků
  - uživatelsky definované jiné proměnné (regresory, např. dopad celozávodních dovolených v automobilovém průmyslu)

# Složky ekonomických časových řad

## Iregulární složka $I_t$

= zbývající součást časové řady po odstranění kalendářní, sezónní složky a trendcyklu z původních dat.

- Někdy se též nazývá reziduální složka, obsahuje náhodnou složku (tzv. bílý šum).
- Má zachycovat zbývající krátkodobé výkyvy v řadě, které nejsou ani systematické, ani předvídatelné
- V praxi se do iregulární složky zahrnují tyto typy výkyvů v časové řadě (outliery), které se odhadují separátně:
  - **aditivní outlier (AO):** jednorázový výkyv v časové řady
  - **transitory change (TC):** postupně odeznívající výkyv v časové řadě.

## Pojem bílý šum (white noise)

Definice (Cipra, T., Finanční ekonometrie, str. 233)

- Bílý šum je posloupnost  $\{\varepsilon_t\}$  nekorelovaných náhodných veličin s nulovou střední hodnotou a konstantním konečných rozptylem  $\sigma^2 > 0$ :

$$E(\varepsilon_t) = 0, \text{var}(\varepsilon_t) = \sigma^2 > 0$$
$$\text{cov}(\varepsilon_s, \varepsilon_t) = 0 \text{ pro } s \neq t.$$

Pozn.: Obvykle se předpokládá, že náhodné veličiny  $\varepsilon_t$  mají normální rozdělení.

# Intervenční analýza

## Vychýlené hodnoty (Outliers)

### Slouží

- k modelování externích šoků působících na ekonomické časové řady
- k posouzení ekonomického dopadu a kvantifikaci událostí
- mohou upozornit na též chyby v datech (překlepy, metodické změny)
- často je lze identifikovat z grafu časové řady  $Y_t$  nebo jejich prvních diferencí
- modelují se pomocí 0-1 (dummy) proměnných
- obvykle se přičítají k modelu (zlogaritmované) řady

### Nejsou-li ošetřeny

- zkreslují odhady jednotlivých složek
- znemožňují správnou identifikaci modelu

Všechny outliery je třeba věcně vysvětlit!

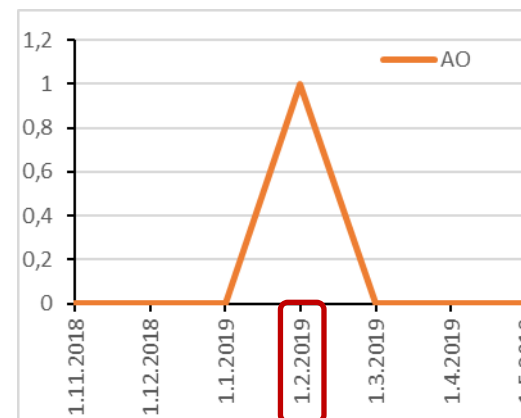
# Typy outlierů

## Odlehlé pozorování (AO = Additive outlier)

- ovlivněna je hodnota pouze jediného pozorování
- může být způsobeno náhodnými faktory nebo důsledkem identifikovatelné příčiny
- např. stávka, špatné počasí, odstávka provozu, ...
- standardně součást náhodné složky
- v odůvodněných případech lze přiřadit do sezónní složky (SAO = seasonal additive outlier)

## Charakter regresní proměnné

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

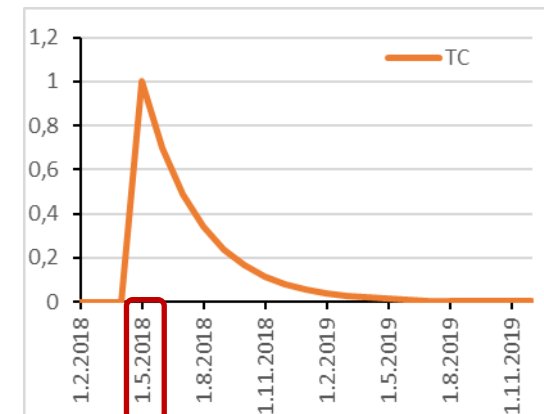


# Typy outlierů

## Dočasná změna (TC = Transitory change)

- hodnota jednoho pozorování je extrémně vysoká/nízká, u následných pozorování se velikost odchylky postupně (exponenciálně) snižuje až na počáteční úroveň
- např. vliv zaváděcí ceny na prodej, uvedení nového produktu
- součást náhodné složky
- Charakter regresní proměnné pro  $\alpha \in (0; 1)$

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



# Typy outlierů

## Posun úrovně (LS = Level shift)

- počínaje daným časovým obdobím prochází úroveň časové řady trvalou změnou
- např. při změně definicích pojmů v zjišťování, způsobu sběru dat nebo metodice zpracování, při změně v ekonomickém chování populace, v právních předpisech nebo v sociálních tradicích
- součást trendu/trendcyklu
- charakter regresní proměnné

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



# Co a oč očišťovat? (připomenutí)

## Očišťování o kalendářní vlivy

- odstranění efektu, který souvisí s kalendářem
- počet pracovních resp. obchodních dní v měsíci, státní svátky, ...

## Sezónní očišťování

- proces odhadu a odstranění sezónních efektů z časové řady
- sezónnost = účinek, který se koná každý rok ve stejnou dobu a se stejnou velikostí a směrem
- zahrnuje i očištění o kalendářní vlivy

## Výhoda očištěných dat

- zvyšování srovnatelnosti dat – v čase, prostoru, mezi odvětvími
- součástí je rozložení časovou řadu na několik různých složek
- užitečnost pro analýzu krátkodobého i dlouhodobého vývoje řady
- Pozn.: důležité je zajištění srovnatelnosti různých měsíců/čtvrtletí v roce



# Očištění o sezónní a kalendářní vlivy

## Očištění o kalendářní vlivy

- multiplikativní model:  $Y_{Cal_t} = \frac{Y_t}{Cal_t} \quad (= TC_t \cdot Sez_t \cdot I_t)$

- aditivní model:  $Y_{Cal_t} = Y_t - Cal_t \quad (= TC_t + Sez_t + I_t)$

## Očištění o sezónní vlivy

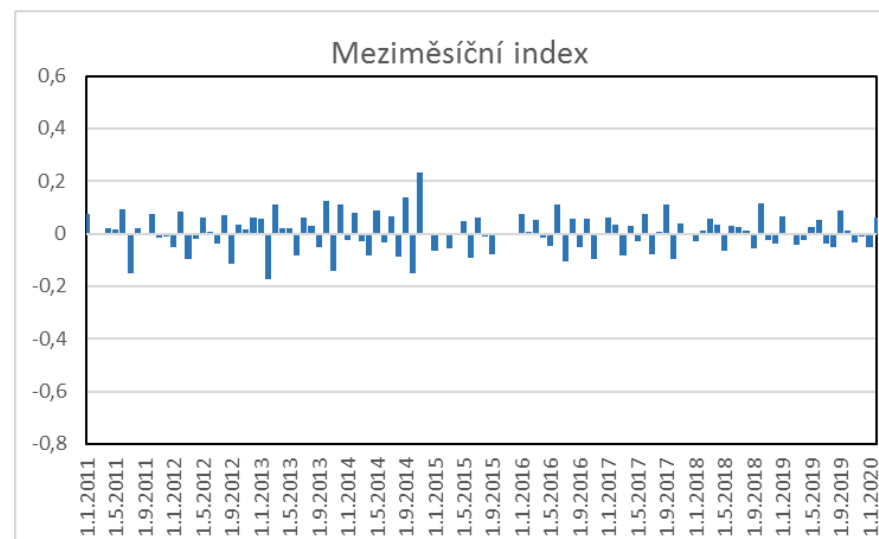
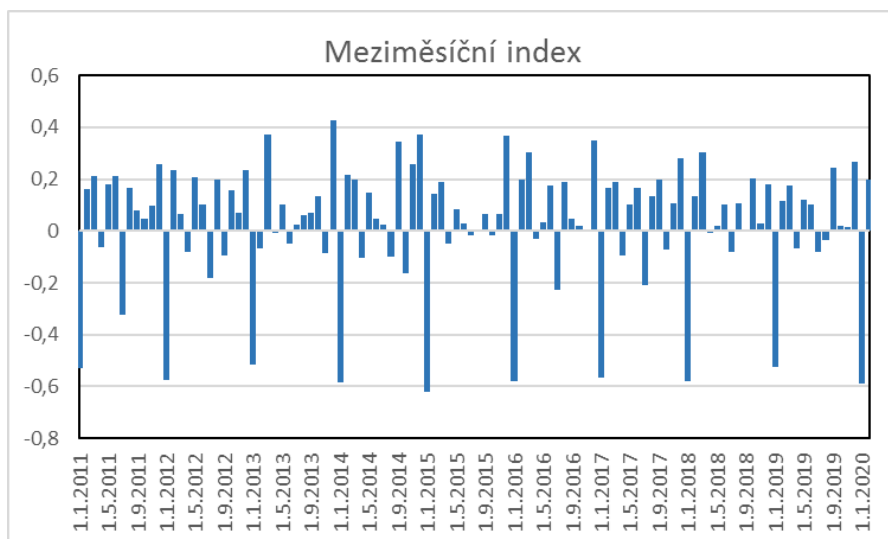
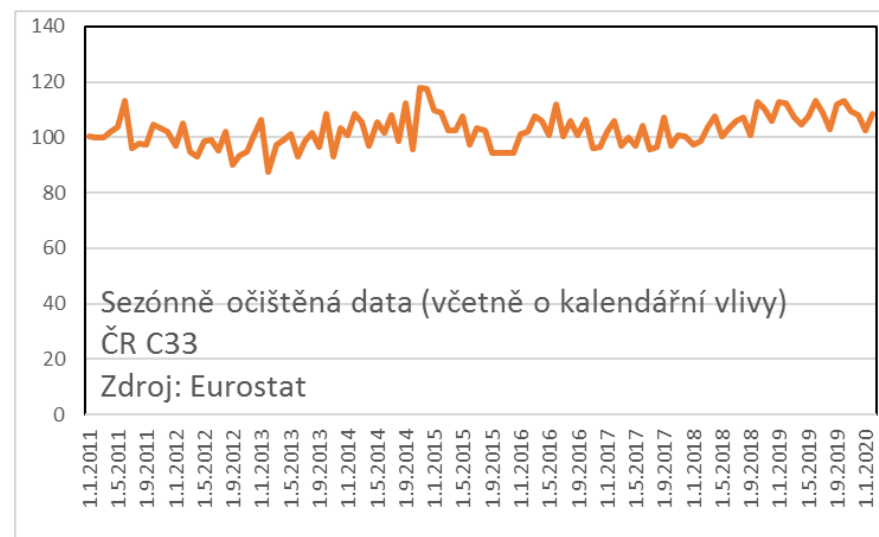
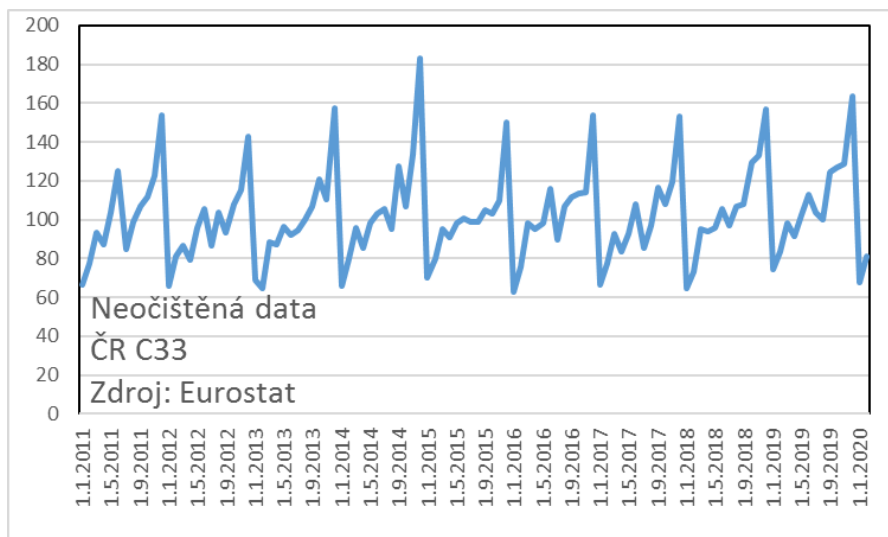
- multiplikativní model:  $SA_t = \frac{Y_t}{Cal_t \cdot Sez_t} \quad (= TC_t \cdot I_t)$

- aditivní model:  $Y_{Cal_t} = Y_t - Cal_t - Sez_t \quad (= TC_t + I_t)$

Pozn. 1: Očištění o sezónní vlivy (sezónně očištěné údaje) zahrnují též očištění o kalendářní vlivy.

Pozn. 2: Sezónně očištěná data zahrnují iregulární sloku ( $I_t$ ), která zahrnuje přechodně působící outliery (AO, TC).

# Srovnatelnost dat v čase: Index průmyslové produkce v ČR, NACE 33: Opravy a instalace strojů



# Metody očištění o kalendářní dny

## ■ Metoda měsíce průměrné délky („trojčlenka“)

- V každém období se hodnota přepočte na měsíc s obvyklým/průměrným počtem pracovních dní (např. 20 dní,  $WD_t$  je počet pracovních dní v daném měsíc)

$$Y_{Cal_t} = \frac{Y_t}{Cal_t} = 20 \cdot \frac{Y_t}{WD_t}$$

- Předpokládá multiplikativní model dekompozice časové řady s  $Cal_t = \frac{WD_t}{20}$

- Výhoda: jednoduché, přímočaré

- Nevýhody:

- Ad hoc metoda,
- Apriori silný předpoklad, že tržby/produkce jsou přímo úměrné počtu pracovních dní
- Nelze statisticky testovat přítomnost kalendářních efektů
- Obtížně zapracovat vliv klouzavých svátků, jiných kalendářních efektů
- Nadhodnocuje vliv jednoho pracovního dne navíc pro meziroční vývoj (index průmyslové produkce ČR: cca 2,7 proc. bodu oproti 5 p.b. (21/20))

# Metody očištění o kalendářní dny II

## ■ Regresní metoda (state-of-art)

1. Working days: 1 regresor: počet pracovních dní -  $\frac{5}{2}$ (počet víkendových dní v měsíci): průmysl, stavebnictví
2. 6 regresorů ( $x_1$ : počet pondělků-počet nedělí v neděli, ...,  $x_6$ : počet sobot-počet nedělí neděli): např. maloobchodní tržby potravinářského zboží, služby
3. Regresor na přestupný rok (v praxi ČSÚ se neosvědčil)
4. Regresory pro pohyblivé svátky (Velikonoce, Ramadán, Lunární nový rok)
5. Uživatelsky definované regresory (např. dopad celozávodních dovolených v automobilkách)
  - Pozn.: Pevně dané státní svátky: počítají se jako neděle.
  - Výhody: flexibilita, univerzálnost, lze testovat statistická významnost jednotlivých efektů, lze je separátně kvantifikovat.

# Využití sezónně očištěných údajů v praxi ČSÚ

Rychlá informace ČSÚ Průmysl v prosinec 2023 (z 6.2.2024):

<https://www.czso.cz/csu/czso/cri/prumysl-prosinec-2023>

**Průmyslová produkce v prosinci meziročně reálně klesla o 0,7 %. Meziměsíčně byla vyšší o 2,8 %.**

**Pozn.: Meziroční vývoj všech ukazatelů, není-li uvedeno jinak, je publikován po očištění o vliv počtu pracovních dnů.**

**Meziměsíční, popř. mezičtvrtletní tempa jsou očištěna také o vliv sezónnosti. Příspěvky k růstu, nebo poklesu, jsou počítány z dat očištěných o vliv počtu pracovních dnů.**

# Využití sezónně očištěných údajů v praxi ČSÚ

Meziroční vývoj:

- Meziroční indexy/tempa z údajů očištěných o kalendářní variace  $YCal_t$

Meziměsíční/mezičtvrtletní vývoj:

- Meziměsíční/mezičtvrtletní indexy/tempa růstu z sezónně očištěných údajů (včetně o kalendářní variace)  $YCal_t$

# Srovnání meziročních údajů multiplikativní model měsíčních řad

- Meziroční vývoj neočištěných (původních) údajů:  $Y_t/Y_{t-12}$
- Meziroční vývoj údajů očištěných o kalendářní variace:

$$Y_{cal_t}/Y_{cal_{t-12}} = Y_t/Y_{t-12} / (Cal_t/Cal_{t-12})$$

$$WD_t < WD_{t-12} \rightarrow Cal_t < Cal_{t-12} \rightarrow Y_t/Y_{t-12} < Y_{cal_t}/Y_{cal_{t-12}}$$

- Meziroční vývoj sezónně očištěných údajů  $SA_t/SA_{t-12}$
- Meziroční indexy sezónně očištěných údajů a údajů očištěných o kalendářní variace se mohou oproti očekávání lišit, protože pokročilé metody (např. TRAMO/SEATS) identifikují model s měnící se sezónností

$$SA_t/SA_{t-12} = Y_{cal_t}/Y_{cal_{t-12}} / (Sez_t/Sez_{t-12})$$

# Přehled metod k odstranění sezónnosti

- Jednoduché klouzavé průměry
  - Délky 4 pro čtvrtletní data, délka 12 pro měsíční data
- Regresní metody
  - Aditivní model sezónnosti: lineární (polynomický) trend + 3/11 dummy (0-1) sezonních regresních proměnných pro jednotlivá čtvrtletní/měsíce
- Adaptivní metody založené na exponenciálním vyrovnávání (aditivní/multiplikativní model: Winters-Holtova metoda)
- Metody využívané statistickými úřady
  - TRAMO/SEATS
  - X-13-ARIMA-SEATS
  - Pozn.: Programové vybavení sponzoruje Eurostat, JDemetra+ nebo R knihovny.



# Volba jednoduchého klouzavého průměru pro sezónní očišťování kvartálních dat

- Jednoduché klouzavé průměry délky  $2l+1$  ( $l=1,2,\dots$ ) v čase  $t$

$$\hat{y}_t = \frac{1}{2l+1} (y_{t-l} + \dots + y_{t-1} + y_t + y_{t+1} + \dots + y_{t+l})$$

- Jak na to pro sudý počet čtvrtletí ( $S=4$ , tj.  $2l$  pro  $l=2$ ) v roce?
- Jednoduchý klouzavý průměr délky pět?

$$\hat{y}_t = \frac{1}{5} (y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2})$$

- Nevyhovující ze dvou důvodů:
  - Klouzavý průměr pro dané období zahrnuje více než 1 ucelený rok
  - Vždy se jeden typ čtvrtletí (např. 1. čtvrtletí) promítne dvakrát celkovou dvojnásobnou vahou.

# Volba jednoduchého klouzavého průměru pro sezónní očišťování kvartálních dat

■ Jednoduché klouzavý průměr délky 4 (2l pro l=2) v čase t lze konstruovat dvěma způsoby:

1. Využijí se více data z minulosti („dva kvartály před, jeden kvartál po“)

$$\hat{y}_t = \frac{1}{4} (y_{t-2} + y_{t-1} + y_t + y_{t+1})$$

2. Využijí se více data z současnosti („jeden kvartál před, dva kvartály po“)

$$\hat{y}_t = \frac{1}{4} (y_{t-1} + y_t + y_{t+1} + y_{t+2})$$

■ Který klouzavý průměr volit? Tak zvolíme jejich průměr

$$\hat{y}_t = \frac{1}{8} (y_{t-2} + 2y_{t-1} + 2y_t + 2y_{t+1} + y_{t+2})$$

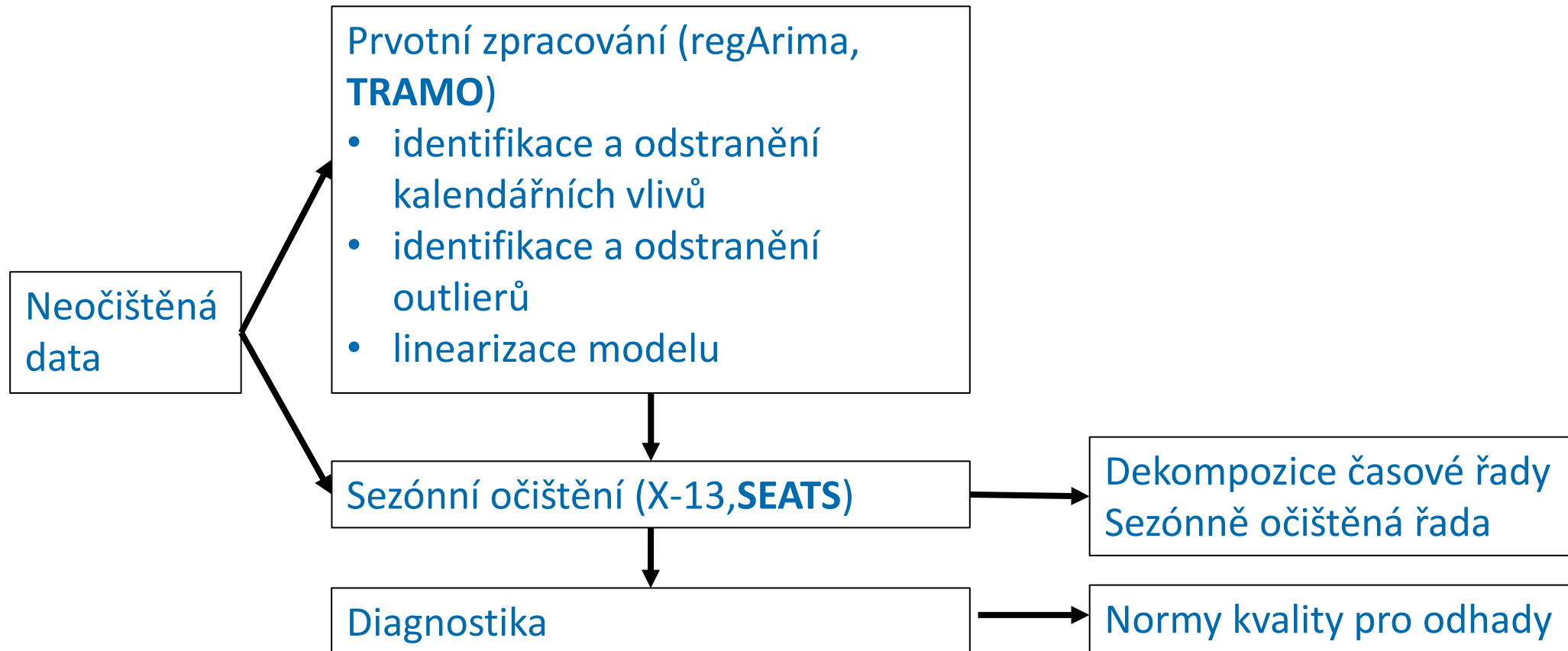
# Vlastnosti jednoduchého klouzavého průměru pro sezónní očištění kvartálních dat

- Vlastnosti klouzavého průměru pro čtvrtletní data

$$\hat{y}_t = \frac{1}{8} (y_{t-2} + 2y_{t-1} + 2y_t + 2y_{t+1} + y_{t+2})$$

1. Je symetrický (využívá vždy časové okno v rozmezí předchozích dvou čtvrtletí a následujících dvou čtvrtletí)
2. Každý typ čtvrtletí se promítne do vyhlazené hodnoty stejnou vahou (1/4).

# Postup sezónního očištění v softwaru **JDemetra+**



# Používané metody – charakteristika

## TRAMO/SEATS

- TRAMO je metoda sloužící k odstranění deterministických efektů, linearizaci modelu a nalezení vhodný SARIMA model (Box-Jenkinsova metodologie)
- SEATS je metoda použitá k vlastní dekompozici časové řady, využívá tzv. spektrální dekompozici

## X-13ARIMA-SEATS

- tato metoda nabízí dva algoritmy pro dekompozici: SEATS a X-13
- X-13 je založena na lineárních filtrech (klouzavých průměrech)

ČSÚ (vlastní výzkum) užívá pouze metodou TRAMO/SEATS.

# Více o sezonním očišťování

## ČSÚ:

- Tereza Košťáková, O složitém jednoduše, ČSÚ Praha

<https://www.statistikaamy.cz/o-slozitem-jednoduse>

<https://www.statistikaamy.cz/o-slozitem-jednoduse/sezonnost-a-sezonni-ocistenii/>

- Holý, D., Vozár, O., Proč sezónně očišťovat

<https://www.statistikaamy.cz/2015/03/30/proc-sezonne-ocistovat/>

## Eurostat:

- Rozcestník k tématu (např. praktické postupy Guidelines on Seasonal Adjustment)

<https://ec.europa.eu/eurostat/web/research-methodology/seasonal-adjustment>

- Handbook on Seasonal Adjustment

<https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-001>

- Software JDemetra+

[https://ec.europa.eu/eurostat/cros/content/software-jdemetra\\_en](https://ec.europa.eu/eurostat/cros/content/software-jdemetra_en)



**Děkuji za pozornost 😊**

# Krátký kvíz

Zařadte následující jevy do jednotlivých složek

	Událost	Složka
1.	Vyšší produkce v důsledku vyššího počtu pracovních dní v daném měsíci oproti předchozím rokům	
2.	Nižší výroba v daném průmyslovém odvětví v zimě	
3.	Vyšší tržby v důsledku každoroční významné výstavy v květnu	
4.	Vyšší tržby v důsledku konání olympijských her	
5.	Dlouhodobě nižší produkce související s hospodářským cyklem	
6.	Významný pokles tržeb v důsledku dopadu opatření ve stavu nouze	
7.	Vyšší tržby v období Vánoc	
8.	Vyšší tržby v období Velikonoc	
9.	Nižší produkce v daném odvětví související s ukončením činnosti významné firmy na našem trhu	



# Krátký kvíz

Zařadte následující jevy do jednotlivých složek

	Událost	Složka
1.	Vyšší produkce v důsledku vyššího počtu pracovních dní v daném měsíci oproti předchozím rokům	$Cal_t$
2.	Nižší výroba v daném průmyslovém odvětví v zimě	$S_t$
3.	Vyšší tržby v důsledku každoroční významné výstavy v květnu	$S_t$
4.	Vyšší tržby v důsledku konání olympijských her	$I_t$
5.	Dlouhodobě nižší produkce související s hospodářským cyklem	$T_t$
6.	Významný pokles tržeb v důsledku dopadu opatření ve stavu nouze	$I_t$
7.	Vyšší tržby v období Vánoc	$S_t$
8.	Vyšší tržby v období Velikonoc	$Cal_t$
9.	Nižší produkce v daném odvětví související s ukončením činnosti významné firmy na našem trhu	$T_t$