

Centrované klouzavé průměry – řešený příklad

Máme k dispozici údaje o počtu krádeží vloupáním v Libereckém kraji v jednotlivých čtvrtletích let 2016-2018. Vyrovnajte tyto údaje pomocí vhodného typu klouzavých průměrů.

Čtvrtletí/Rok	Počet krádeží
I/2016	586
II/2016	1688
III/2016	2526
IV/2016	3365
I/2017	491
II/2017	1361
III/2017	2252
IV/2017	3032
I/2018	488
II/2018	1304
III/2018	2190
IV/2018	3131

Zdroj: <https://www.policie.cz/clanek/statisticke-prehledy-kriminality-za-rok-2018.aspx>

Řešení:

Klouzavé průměry používáme k vyrovnávání hodnot časových řad, a to zejména takových, které se vyznačují většími nepravidelnostmi v trendu. Důležitou otázkou hned na počátku je volba rozsahu klouzavé části období interpolace, tedy m , kde $m = 2p+1$. Pokud pracujeme s časovými řadami neperiodickými, volíme rozsah klouzavé části číslo liché a menší, často 3, 5, 7 či 9. Pokud pracujeme s časovými řadami periodickými, volíme m podle délky dané periody. Zde máme časovou řadu periodickou. Údaje jsou zde dány za jednotlivá čtvrtletí, proto zvolíme $m = 4$, neboť dává smysl součet 4 po sobě jdoucích hodnot (4 čtvrtletí v rámci roku). Pokud ovšem m je sudé číslo, nejsou pak střední body jednotlivých klouzavých částí čísla celá a pak nelze hodnoty klouzavých průměrů přímo přiřadit jednotlivým pozorováním. Proto hodnoty klouzavých průměrů ještě poté musíme „vycentrovat“, tj. ze dvou prostých čtyřčlenných klouzavých průměrů vypočítáme jeden centrovaný klouzavý průměr jako jejich aritmetický průměr. Zbývá ještě určit, kolik hodnot na začátku a na konci dané časové řady zůstane nevyrovnáno.

$$4 = 2p+1$$

$$p = 1,5$$

Vidíme, že by na začátku a na konci časové řady mělo zůstat nevyrovnáno 1,5 hodnoty. To jasně indikuje už zmiňovanou skutečnost, že při výpočtu prostých čtyřčlenných klouzavých průměrů nedokážeme tyto průměry přiřadit konkrétním pozorováním, protože jsou pomyslně mezi danými pozorováními. Při výpočtu centrovaných čtyřčlenných klouzavých průměrů pak zůstanou nevyrovnané 2 hodnoty na začátku a na konci časové řady. Ukažme si nejprve ruční výpočet alespoň prvních pár hodnot centrovaných klouzavých průměrů.

Hodnotu prvního prostého čtyřčlenného klouzavého průměru vypočítáme jako součet prvních 4 hodnot časové řady a vydělíme jejich počtem, tj. $\frac{586+1688+2526+3365}{4} = 2041,25$. Další hodnotu získáme jako součet druhé, třetí, čtvrté a páté hodnoty časové řady a opět dělíme čtyřmi, tj. $\frac{1688+2526+3365+491}{4} = 2017,5$. Další hodnoty získáme postupným klouzáním o jednu hodnotu časové řady směrem do budoucnosti – hodnoty klouzavých průměrů pro přehlednost uvedme do následující tabulky.

Čtvrtletí/Rok	Počet krádeží	Prosté 4-členné klouz. průměry
I/2016	586	•
II/2016	1688	- 2041,25
III/2016	2526	2017,5
IV/2016	3365	1935,75
I/2017	491	1867,25
II/2017	1361	1784
III/2017	2252	1783,25
IV/2017	3032	1769
I/2018	488	1753,5
II/2018	1304	1778,25
III/2018	2190	-
IV/2018	3131	•

Hodnoty centrováných klouzavých průměrů dostaneme jako aritmetický průměr sousedních prostých klouzavých průměrů. První z nich bude $\frac{2041,25+2017,5}{2} = 2029,375$, další pak $\frac{2017,5+1935,75}{2} = 1976,625$. Třetí hodnotu dostaneme jako $\frac{1935,75+1867,25}{2} = 1901,5$.

Když to shrneme a uvedeme od začátku, postup výpočtu první hodnoty čtyřčlenného centrováného klouzavého průměru, dostaneme jako

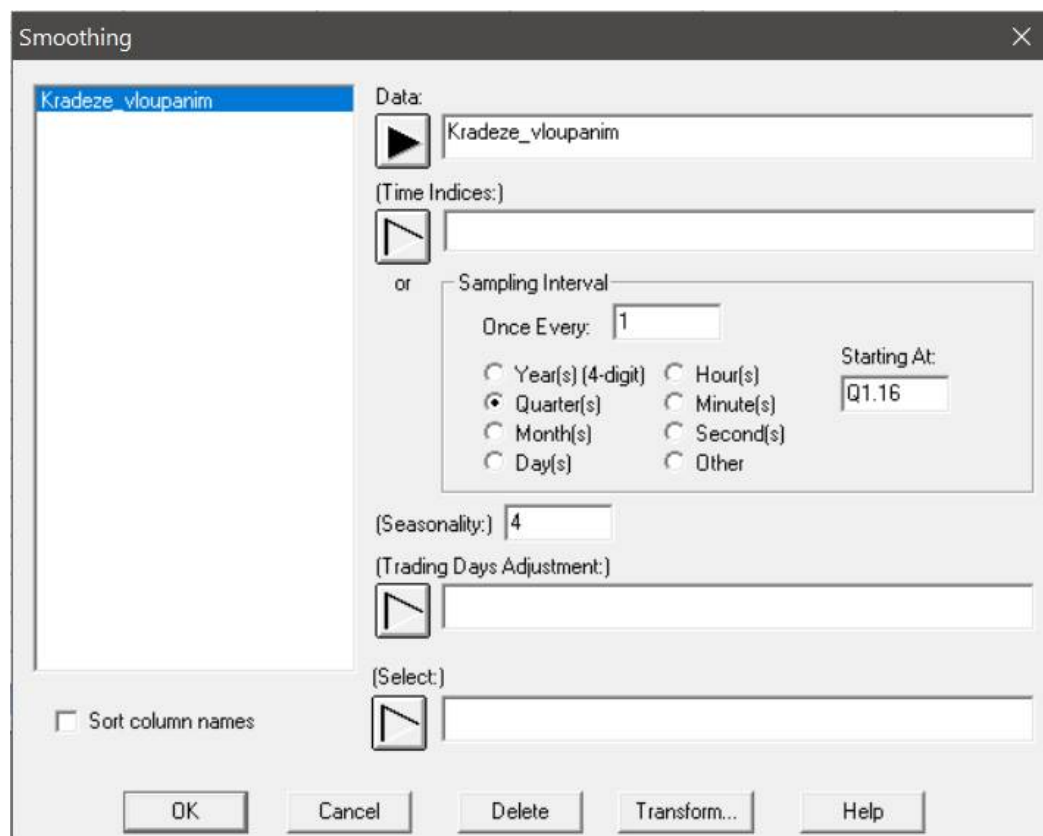
$$\bar{y}_3 = \frac{\frac{586+1688+2526+3365}{4} + \frac{1688+2526+3365+491}{4}}{2} = 2029,375.$$

Obdobně další bude $\bar{y}_4 = \frac{\frac{1688+2526+3365+491}{4} + \frac{2526+3365+491+1361}{4}}{2} = 1976,625$. Atd.

Čtvrtletí/Rok	Počet krádeží	4-členné centrování klouz. průměry
I/2016	586	•
II/2016	1688	•
III/2016	2526	2029,375
IV/2016	3365	1976,625
I/2017	491	1901,5
II/2017	1361	1825,625
III/2017	2252	1783,625
IV/2017	3032	1776,125
I/2018	488	1761,25
II/2018	1304	1765,875
III/2018	2190	•
IV/2018	3131	•

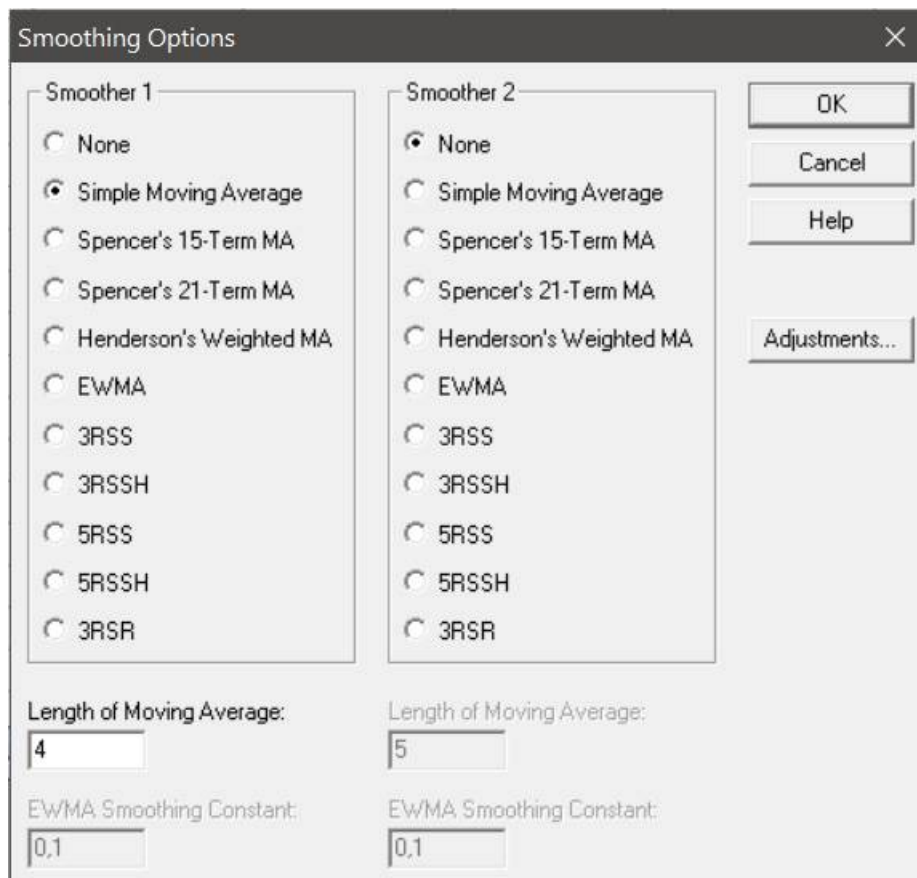
Řešení v SGP:

Data o počtu krádeží vloupáním vložíme do prvního sloupce a zvolíme posloupnost procedur **Describe – Time Series – Smoothing ...** Do řádku *Data* zadáme uloženou proměnnou, v dalším poli *Sampling Interval* označíme *Quarter(s)* a můžeme vyplnit i *Starting At* hodnotou Q1.16, což značí, že máme první údaj z prvního čtvrtletí roku 2016. Ještě je pak potřeba vyplnit pole *Seasonality* hodnotou 4 – je to perioda, kterou daná časová řada má – viz Obr. 1.



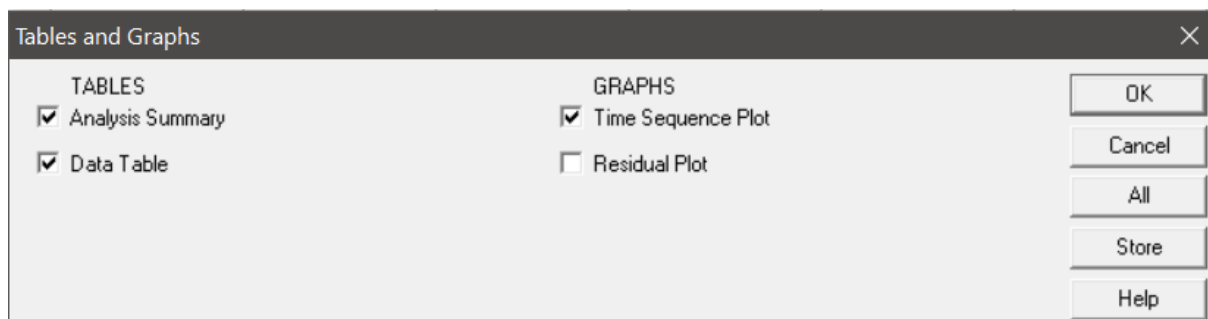
Obrázek 1 – Vstupní panel procedury Smoothing

Když údaje ve vstupním panelu potvrdíme tlačítkem OK, objeví se nabídka výstupů, které je možné vypočítat – viz Obr. 2. Ačkoli počítáme hodnoty centrovaných klouzavých průměrů, necháme puntík u položky *Simple Moving Average*. Je to trochu matoucí, ale nic se s tím nedá dělat. To, že chceme vypočítat čtyřčlenné centrované klouzavé průměry, ještě dáme programu na srozuměnou tím, že v poli *Length of Moving Average* zadáme hodnotu 4. Potvrdíme tlačítkem OK.



Obrázek 2 – Zadání délky klouzavé části období interpolace ve Smoothing Options

Z nabídky *Tables and Graphs* vybereme v levé části ještě výstup *Data Table*, jak ukazuje Obr. 3.



Obrázek 3 – Nabídka Tables and Graphs v proceduře Smoothing

Právě tento výstup obsahuje hodnoty čtyřčlenných centrovaných klouzavých průměrů. Jsou ve třetím sloupci tabulky, označené jako *Smooth*, jak vidíme na Obr. 4. Pomocné výpočty, tj. hodnoty prostých čtyřčlenných klouzavých průměrů už zde nenajdeme.

Smoothing - Kradeze_vloupanim			
Data Table for Kradeze_vloupanim			
First smoother: simple moving average of 4 terms			
Second smoother: none			
Period	Data	Smooth	Rough
Q1.16	586,0		
Q2.16	1688,0		
Q3.16	2526,0	2029,38	496,625
Q4.16	3365,0	1976,63	1388,38
Q1.17	491,0	1901,5	-1410,5
Q2.17	1361,0	1825,63	-464,625
Q3.17	2252,0	1783,63	468,375
Q4.17	3032,0	1776,13	1255,88
Q1.18	488,0	1761,25	-1273,25
Q2.18	1304,0	1765,88	-461,875
Q3.18	2190,0		
Q4.18	3131,0		

Obrázek 4 – Výstup Data Table

Grafické znázornění vyrovnání časové řady krádeží vloupáním pomocí čtyřčlenných centrovaných klouzavých průměrů najdeme v pravé části výstupů (grafické výstupy), nazvané jako *Time Sequence Plot* – najdeme ho na Obr. 5.

