ČASOVNE VRSTE

1. UVOD

Številsko spremenljivko Y opazujemo v času. Podatki se nanašajo na zaporedna **časovna obdobja** $t_1, t_2, \ldots, t_T$.

Statistično vrsto $y_1, y_2, ..., y_T$ imenujemo **časovna vrsta**.
T dolžina časovne vrste

Časovne vrste so:
- **trenutne**: vrednosti se nanašajo na trenutek

Primer: temperatura zraka ob 7h vsak dan

- **intervalne**: vrednosti se nanašajo na časovni interval

Primer: število rojstev, smrti, prometnih nesreč, itd. na leto

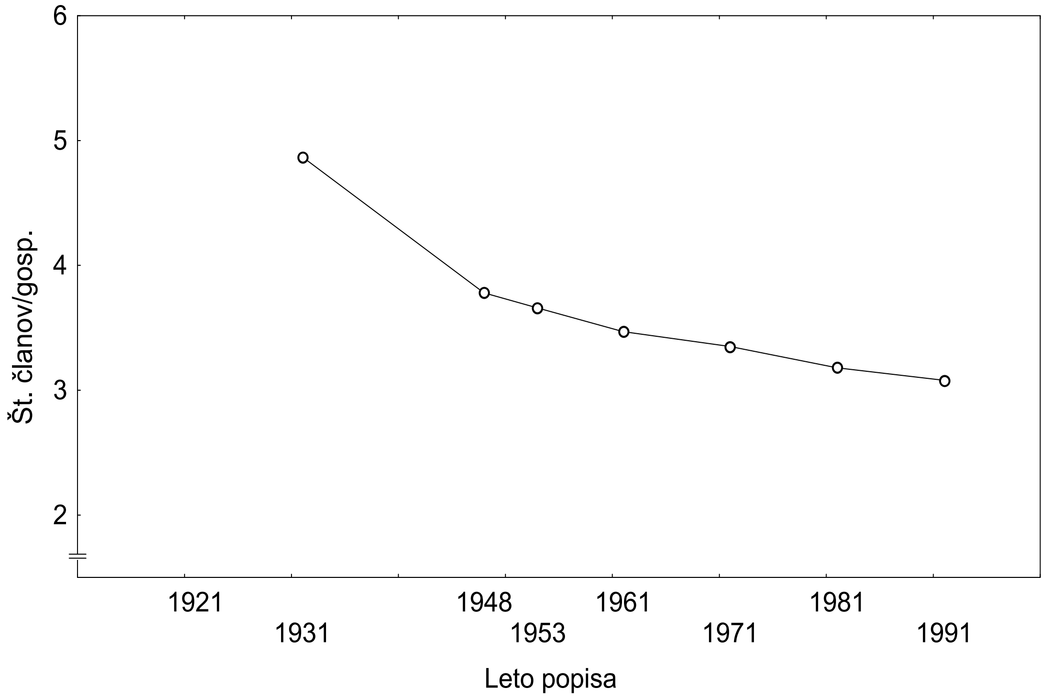
- **izvedene:** vrednosti so izračunane

Primer: letna stopnja inflacije, povprečna dnevna temperatura

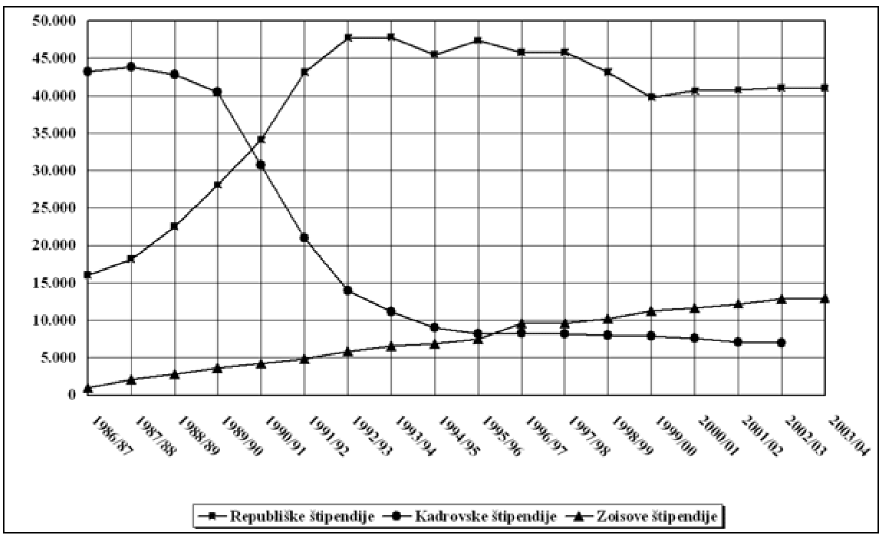
Osnovni namen analize časovnih vrst je
- **proučevanje dinamike** pojava, ki ga časovna vrsta opisuje;
- **napovedovanje** tega pojava.

Analiza časovnih vrst se začne s tem, da časovno vrsto **najprej narišemo!!!!!!!!!!!!** Grafični prikaz časovne vrste: **LINIJSKI GRAFIKON:** na abscisni osi je časovna skala. Pri grafičnem prikazu moramo ustrezno prikazati časovno zaporedje vrednosti.

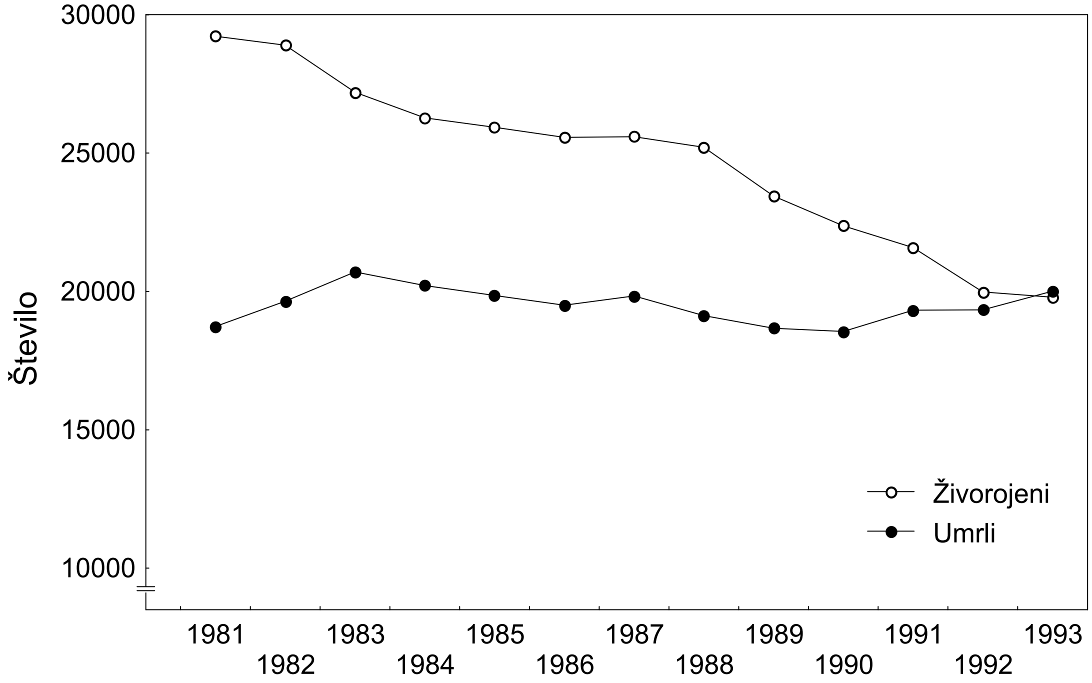
Če časovna vrsta ni **ekvidistantna**, torej med vrednostmi ni enak časovni razmik, moramo na sliki to upoštevati.



Slika: Število članov na gospodinjstvo po popisih prebivalstva v Sloveniji v obdobju 1931 do 1991



Slika: Število štipendistov po letih v obdobju 1986/87 do 2003/04 v Sloveniji glede vrsto štipendije



Slika: Število umrlih in število živorojenih po letih v obdobju 1981 - 1993

Osnovno orodje za analizo časovnih vrst so **indeksi.**

INDEKSI S STALNO OSNOVO

INDEKSI S PREMIČNO OSNOVO:

- **VERIŽNI INDEKSI,**
- **STOPNJA RASTI**

Ponovite, kaj pomeni in kako se izračuna povprečna stopnja rasti!!!!

2. KOMPONENTE ČASOVNE VRSTE

Pri časovni vrsti skušamo identificirati naslednje komponente:

TREND: T
predstavlja **dolgoročno** gibanje pojava, podaja pa osnovno smer razvoja.
Trend je lahko naraščajoč, padajoč ali stacionaren.
Dolgoročne spremembe nastanejo npr. zaradi gospodarskih sprememb, sprememb v okolju, bioloških dejavnikov

SEZONSKA KOMPONENTA: S
se nanaša na **periodične spremembe**. Sezonske variacije nastanejo npr. zaradi letnih časov, praznikov, vremena itd.
Sezonska komponenta ima svojo dolžino, dolžina sezonske komponente je konstantna.
Zato lahko sezonsko komponento predvidimo. Ponavadi se veže na eno koledarsko leto.

CIKLIČNA KOMPONENTA: C
se nanaša na **neperiodične spremembe**. Predstavlja nihanje okoli trenda in je očitna v zelo dolgih obdobjih. Ciklične spremembe nastanejo npr. zaradi sprememb v gospodarstvu, okolju.
Dolžina cikla ni določena.

SLUČAJNA KOMPONENTA (IREGULARNI VPLIVI, ŠUM): I
se nanaša na ostale nam neznane dejavnike, ki jih ne moremo razložiti z trendom, sezonsko in ciklično komponento.

Časovna vrsta ni nujno rezultat delovanja vseh teh komponent. V časovni vrsti vedno obstoja slučajna komponenta.

Teorija časovnih vrst skuša identificirati te komponente, jih razumeti in uporabiti pri napovedovanju.

Za vsako časovno vrsto skušamo identificirati njen **model**. Modelov je veliko vrst. Najsplošnejša sta dva:

ADITIVNI MODEL

$$Y = T + S + C + I$$

Časovna vrsta je vsota komponent, ki sestavljajo časovno vrsto.

MULTIPLIKATIVNI MODEL

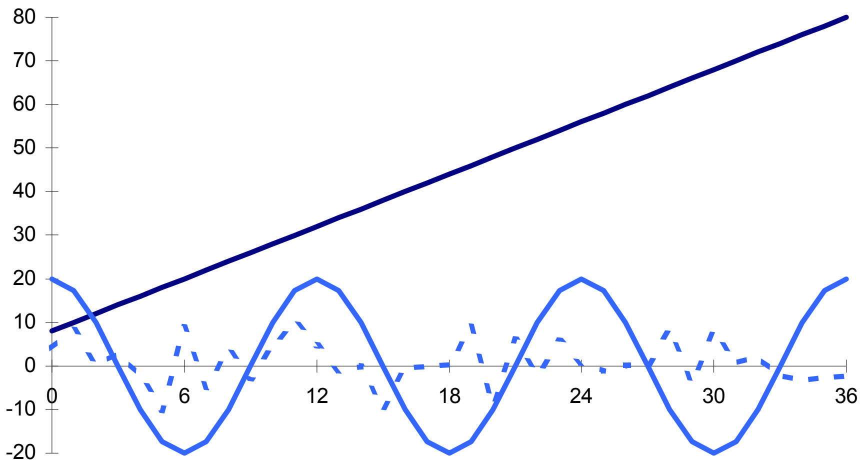
$$Y = T \cdot S \cdot C \cdot I$$

drugače zapisano:

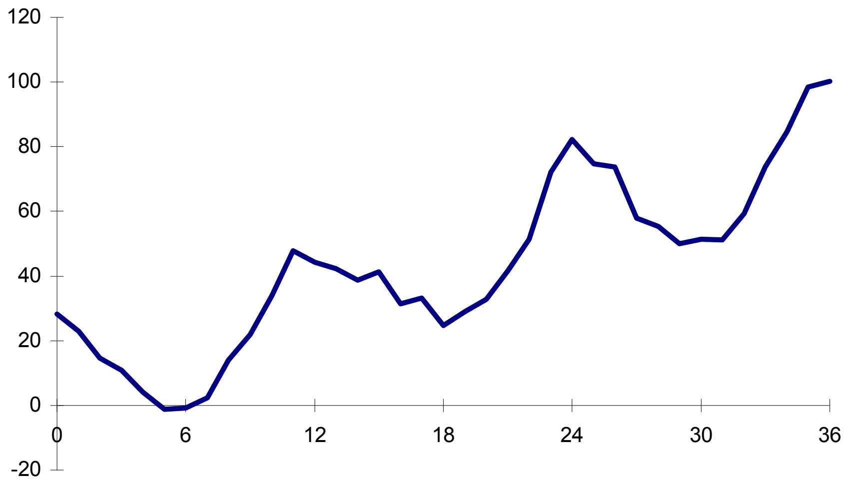
$$\log Y = \log T + \log S + \log C + \log I$$

Logaritem časovne vrste je vsota logaritmiranih komponent.

Primer:

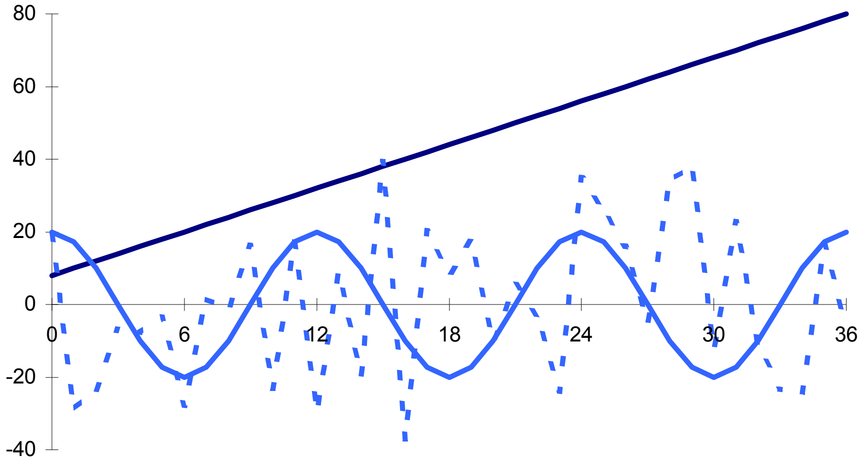


Slika : Linearni trend, sezonska komponenta in slučajni vplivi

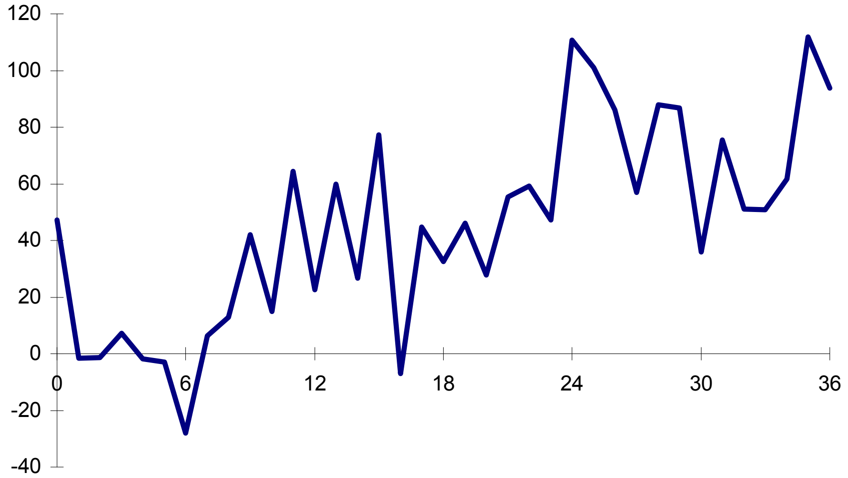


Slika: Aditivni model za časovno vrsto

Primer 2
Slučajna komponenta ima večjo variabilnost:



Slika : Linearni trend, sezonska komponenta in slučajni vplivi



Slika: Aditivni model za časovno vrsto

Slučajna komponenta zabriše vpliv trenda in sezonske komponente.

TEHNIKE GLAJENJA
so postopki, s katerimi poskušamo izločiti vpliv slučajne komponente, da bi odkrili trend, sezonsko in ciklično komponento. Najbolj uporabljeni metodi glajenja sta:

- metoda drsečih sredin
- eksponentno glajenje (ne bomo obravnavali)

METODA DRSEČIH SREDIN

Originalni časovni vrsti $Y = y_1, y_2, ..., y_T$ priredimo **časovno vrsto drsečih sredin.**
Odločimo se za **red drsečih sredin** r.
Vsaka drseča sredina je na sredini intervala dolžine r.

Oznake:
r – red drsečih sredin
Y^r, npr. drseče sredine reda 3 označimo Y^3

Način izračunavanja drsečih sredin

r je liho število, $r = 2i + 1$

$$y_k^r = \frac{1}{r}(y_{k-i} + y_{k-i+1} + ... + y_k + ... + y_{k+i-1} + y_{k+i})$$

b) r je sodo število, $r = 2i$

$$y_k^r = \frac{1}{r}(\frac{1}{2} y_{k-i} + y_{k-i+1} + ... + y_k + ... + y_{k+i-1} + \frac{1}{2} y_{k+i})$$

$$Y^r = y_{i+1}^r, y_{i+1}^r, ..., y_{T-i}^r$$

Časovna vrsta drsečih sredin je krajša od originalne časovne vrste, manjka ji *i*- členov na začetku in *i*-na koncu.

$$Y^r = y_{i+1}^r, y_{i+1}^r, ..., y_{T-i}^r$$

Primer: prirast smreke po letih, dolžina časovne vrste =15

Leto	Prirast (mm^2)	Drseče sredine r=3	Drseče sredine r=4	Drseče sredine r=6
1	830			
2	1026	903,0		
3	853	843,3	804,4	
4	651	683,0	742,8	808,8
5	545	671,3	746,3	815,6
6	818	817,3	800,5	810,4
7	1089	918,7	878,1	794,0
8	849	968,7	871,1	791,0
9	968	718,7	777,8	824,3
10	339	709,3	755,8	844,3
11	821	700,3	797,1	857,3
12	941	989,7	895,4	894,8
13	1207	1011,3	1033,8	
14	886	1158,0		
15	1381			

$r = 3 = 2 \cdot 1 + 1$ (i=1)

$$y_2^3 = \frac{1}{3}(830 + 1026 + 853) = 903{,}0$$

$$y_3^3 = \frac{1}{3}(1026 + 853 + 651) = 843{,}3$$

...

$r = 4 = 2 \cdot 2$ (i=2)

$$y_3^4 = \frac{1}{4}(\frac{1}{2} \cdot 830 + 1026 + 853 + 651 + \frac{1}{2} \cdot 545) = 804{,}375$$

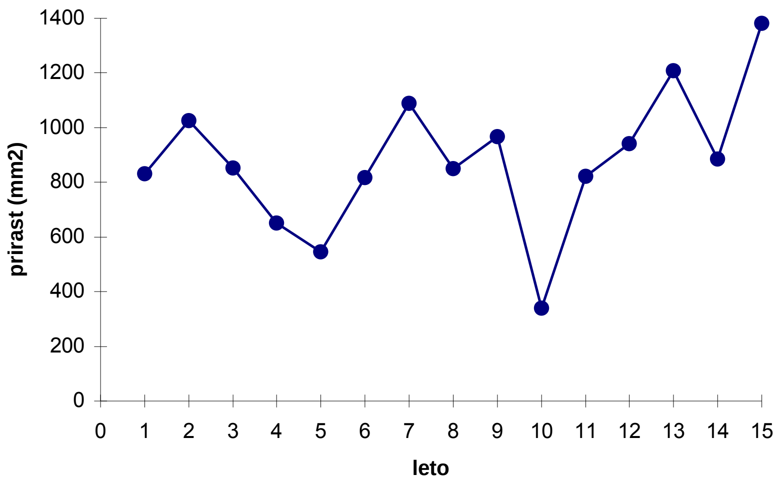
$$y_4^4 = \frac{1}{4}(\frac{1}{2} \cdot 1026 + 853 + 651 + 545 + \frac{1}{2} \cdot 818) = 742{,}750$$

...

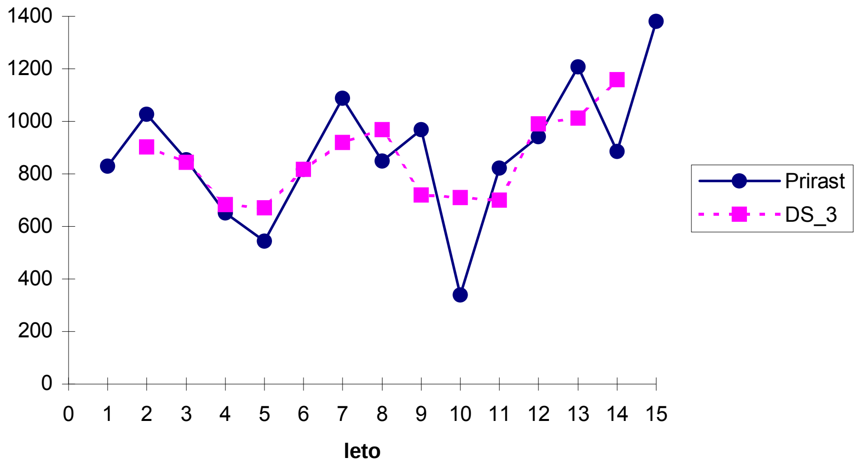
$r = 6 = 2 \cdot 3$ (i=3)

$$y_4^6 = \frac{1}{6}(\frac{1}{2} \cdot 830 + 1026 + 853 + 651 + 545 + 818 + \frac{1}{2} \cdot 1089) = 808{,}750$$

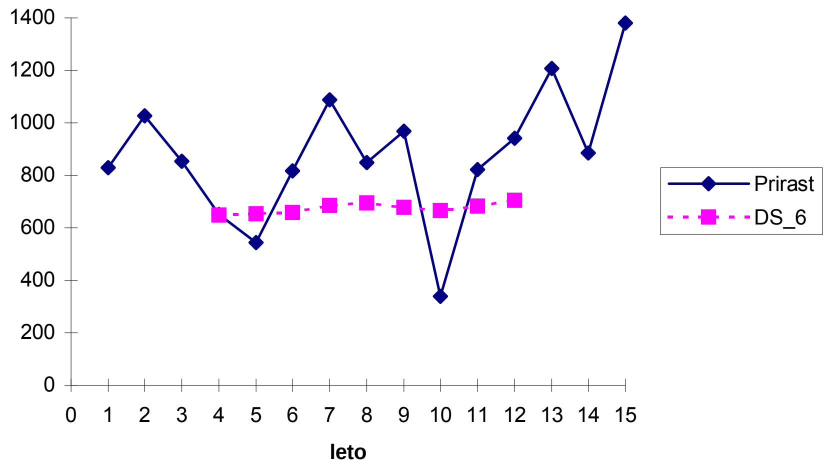
$$y_5^6 = \frac{1}{6}(\frac{1}{2} \cdot 1026 + 853 + 651 + 545 + 818 + 1089 + \frac{1}{2} \cdot 849) = 815{,}583$$



Slika 1: Prirast lesa



Slika 2: Prirast in drseče sredine reda 3



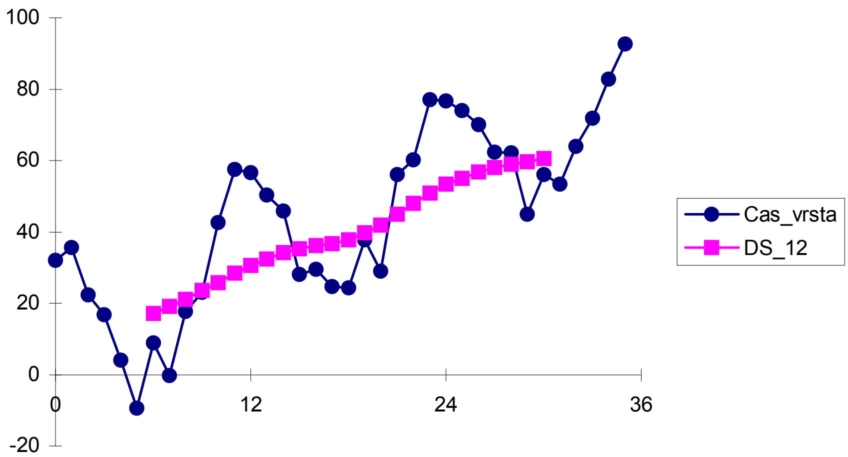
Slika 3: Prirast in drseče sredine reda 6

Efekt drsečih sredin je **glajenje podatkov**.Večji red r povzroči večje glajenje osnovne časovne vrste.

Slabost časovne vrste drsečih sredin je, da je krajša od osnovne časovne vrste.

Prednosti časovne vrste drsečih sredin: **odkrivanje trenda**
- odkrivanje trenda: če je trend osnovne časovne vrste **linearen**, časovna vrsta drsečih sredin prikazuje trend.
- če ima časovna vrsta periodično oz. ciklično komponento, časovna vrsta drsečih sredin prikazuje trend, če je njen red r enak dolžini periode.

Primer:



Slika: Časovna vrsta s periodo 12 in drseče sredine reda 12

METODE ZA DOLOČANJE TRENDA

- prostoročno določanje trenda na grafičnem prikazu: OK za začetek analize
- metoda drsečih sredin (v določenih primerih)

- analitične metode. Najenostavnejša metoda: regresija po metodi najmanjših kvadratov

Za lažje računanje ocen parametrov regresijskega modela transformiramo časovno skalo tako, da uvedemo t.i. **tehnični čas** t.

Postopek:
Ordinatno os premaknemo v sredino časovnega intervala.
S tem dosežemo, da je vsota vrednosti na abscisi enaka 0.

a) dolžina časovne vrste T je **liho število**: $T = 2k + 1$
Tehnični čas: $t = -k, -k+1, ..., 0, ..., k-1, k$

Primer: $T = 11$
tehnični čas: $t = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5$

b) dolžina časovne vrste T je **sodo število**: $T = 2k$
tehnični čas: $-k + \frac{1}{2}, -k + \frac{3}{2}, ..., k - \frac{3}{2}, k - \frac{1}{2}$

Primer: $T = 10$
tehnični čas: $-4\frac{1}{2}, -3\frac{1}{2}, -2\frac{1}{2}, -1\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, 1\frac{1}{2}, 2\frac{1}{2}, 3\frac{1}{2}, 4\frac{1}{2}$

LINEARNI REGRESIJKI MODEL ZA DOLOČANJE TRENDA

Predpostavka: iz slike ugotovimo, da je linearni trend sprejemljiv.

Model:

$$Y = \alpha + \beta t + \varepsilon$$

Izračun ocen parametrov linearnega regresijskega modela:

$$b = \frac{\sum_{i=1}^{n} y_i t_i}{\sum_{i=1}^{n} t_i^2}$$

$$a = \overline{y}$$

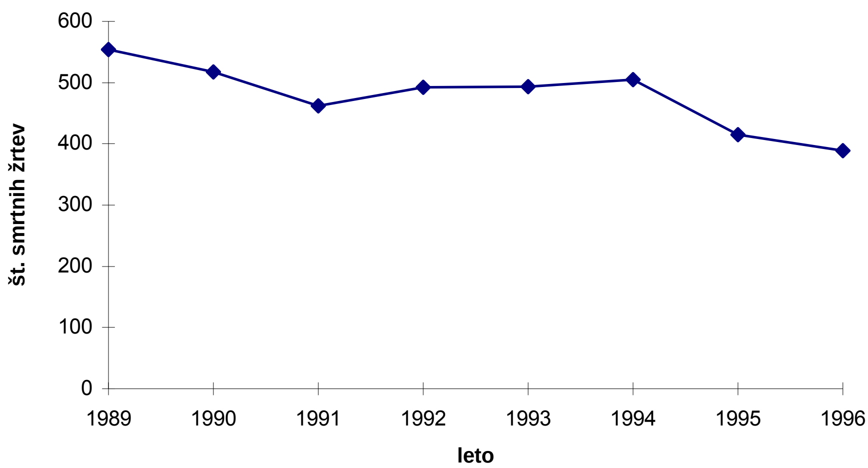
Pomen parametrov:

- parameter a je povprečje pojava v opazovanem obdobju;
- parameter b pove, za koliko se spremeni pojav v povprečju, če se čas spremeni za eno enoto.

Primer:

Tabela: število smrtno ponesrečenih oseb v cestnem prometu po letih v obdobju 1989-1996 (Vir: SL-95 in SL-97)

Leto	Število smrtnih žrtev
1989	554
1990	517
1991	462
1992	492
1993	493
1994	505
1995	415
1996	389



Slika 1: Število smrtnih žrtev v prometnih nesrečah po letih v obdobju 1989-1996

Časovna vrsta kaže trendno komponento in slučajno komponento. Privzeli bomo model:

$$Y = T + I$$

Slika kaže, da je linearni trend sprejemljiv.

$$Y = \alpha + \beta \cdot t + \varepsilon$$

Pomožni računi:

Leto	t	t^2	y	t•y
1989	-3,5	12,25	554	-1939,0
1990	-2,5	6,25	517	-1292,5
1991	-1,5	2,25	462	-693,0
1992	-0,5	0,25	492	-246,0
1993	0,5	0,25	493	246,5
1994	1,5	2,25	505	757,5
1995	2,5	6,25	415	1037,5
1996	3,5	12,25	389	1361,5
Vsota	**0**	**42**	**3827**	**-767,5**

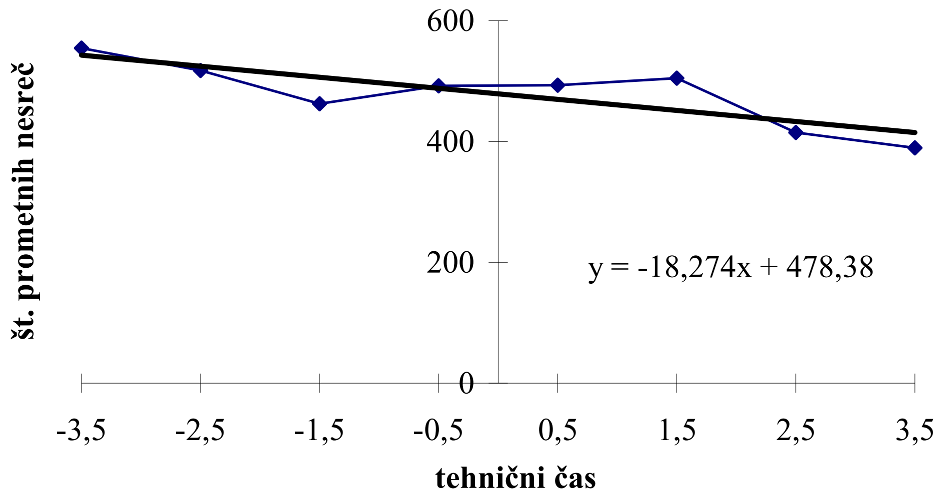
$$a = \frac{3827}{8} = 478{,}375$$

$$b = \frac{-767{,}5}{42} = -18{,}1369$$

Linearni regresijski model:
$\hat{Y} = 478{,}375 - 18{,}1369 \cdot t$

Obrazložitev:

- **povprečno letno število smrtnih žrtev v obravnavanem obdobju je 478**.
- **v povprečju se je v obravnavanem obdobju število smrtih žrtev vsako leto zmanjšalo za 18.**



Napoved za leto 1997:
$\hat{y}(t=4,5)=396,2$

Dodatni podatek: število smrtno ponesrečenih nesreč v letu 1997 je bilo 355.
Napoved je za 12,6% večja od dejanske vrednosti.