



# La modellizzazione matematica della realtà

un percorso di elaborazione applicativa dei contenuti didattici della matematica

INCONTRO II 10/06/2020





# Tipologie di modelli matematici stocastici

Nei modelli matematici rappresentati da una funzione occorre chiarire bene quali sono le variabili in gioco.

1

Analisi di una sola variabile

Questa analisi cerca di associare un modello matematico capace di descrivere le frequenze di un carattere nella popolazione. Particolarmente importante per lo studio delle popolazioni normali.

1+

$$y = f(t)$$

Analisi di una variabile nel tempo

Questa analisi considera il movimento di un carattere nel tempo. Particolarmente importante per lo studio delle popolazioni ad esempio. Se il carattere ha natura economica questo tipo di studio prende il nome di analisi delle serie storiche ed ha un suo sviluppo metodologico molto importante.

2

$$y = f(x)$$

Analisi di due variabili

Questo tipo di studio è legato al concetto di dipendenza di due variabili e di come si può spiegare una rispetto al valore dell'altra, in ambiente statistico prende il nome di regressione.

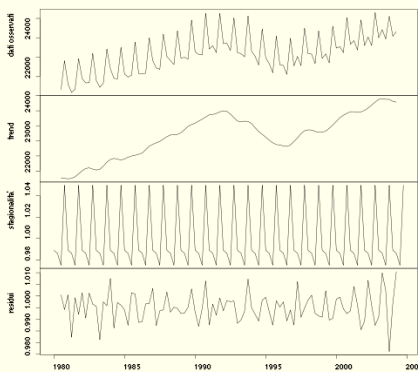
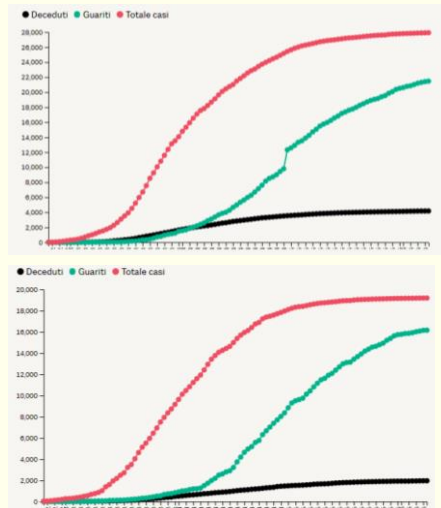
+

$$y = f(x_i)$$

Analisi di più variabili

Questi modelli sono molto complicati da gestire e anche da rappresentare oltre al fatto che richiederebbero una introduzione allo studio di funzione in più variabili. Tra i più famosi, in ambito economico, ci sono i modelli che spiegano la produzione con le variabili capitale e lavoro (modelli di Cobb Douglas).

# Obiettivi dello studio delle serie storiche



## DESCRIZIONE

- Individuazione delle caratteristiche generali di un fenomeno

## SPIEGAZIONE

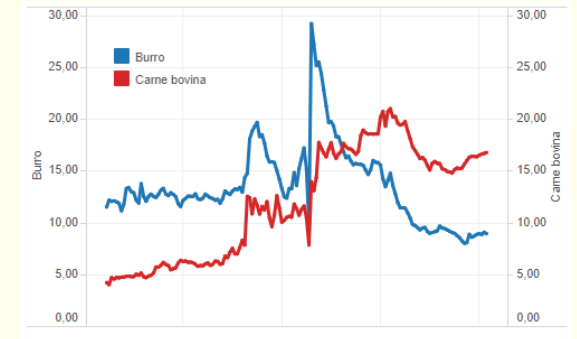
- Individuazione del meccanismo che genera la serie e le relazioni che legano la variabile ad altri fenomeni

## SCOMPOSIZIONE

- Individuazioni del valore o del peso delle diverse componenti

## CONTROLLO

- Di un sistema di produzione, di una rete di vendita, di irregolarità osservate





# Analisi delle serie storiche

Si tratta di osservare dei dati di un fenomeno nel tempo e di comprenderne le diverse componenti. Esistono diversi approcci, quello considerato in questa analisi è l'approccio classico. Questo approccio prevede la disaggregazione di ogni dato in quattro componenti: trend (T), ciclicità (C), stagionalità (S) e accidentalità (A).

A sua volta l'approccio classico prevede due modalità distinte del calcolo e disaggregazione delle componenti: uno additivo e uno moltiplicativo.

Con l'approccio additivo ogni dato viene ricavato come somma delle quattro componenti.

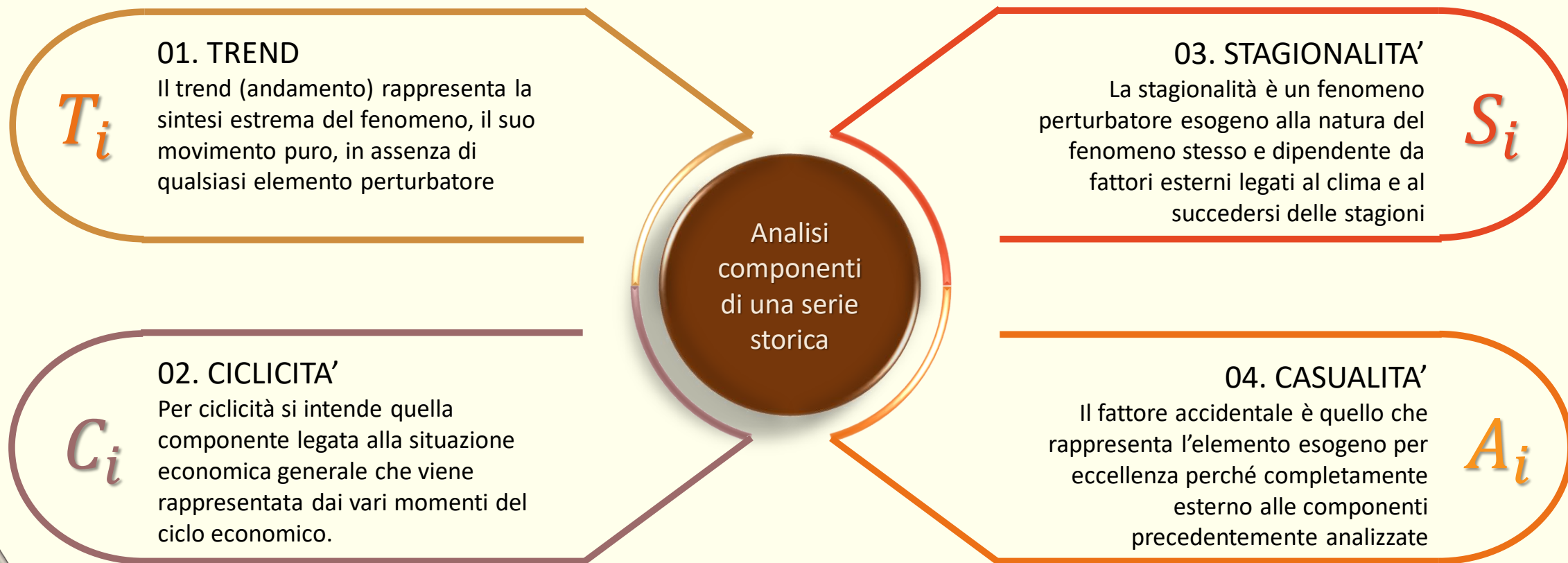
$$y_i = T_i + C_i + S_i + A_i$$

Con l'approccio moltiplicativo invece ogni valore si ottiene dal prodotto delle componenti.

$$y_i = T_i \cdot C_i \cdot S_i \cdot A_i$$

Queste componenti sono indipendenti tra di loro, cioè il variare di una influenza il valore complessivo, ma non modifica in nessun modo le altre.

# Le 4 componenti in sintesi



# Il trend



Il valore del trend rappresenta il valore del fenomeno in assenza di tutti gli elementi perturbatori esogeni e endogeni che possono modificarlo. Rappresenta quel movimento che il fenomeno ha nel tempo ed è l'obiettivo principale dell'analisi.

Il primo passo quindi dell'analisi è quello di ISOLARE il trend come componente principale del fenomeno.

Per fare questo si possono usare due metodi:

Attraverso il metodo delle medie mobili

Attraverso il metodo dei minimi quadrati

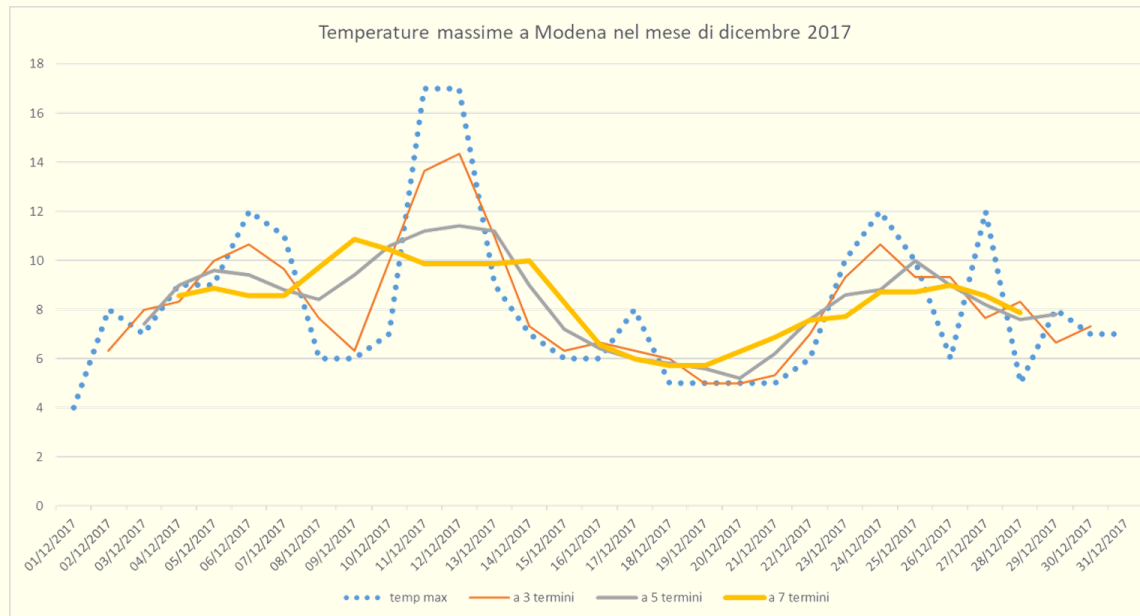
$y_i$  È il valore osservato

$T_i = \hat{y}_i = f(x_i)$  È il valore teorico corrispondente all'i-esimo termine reale

## Metodo delle medie mobili

Il metodo delle medie mobili consiste nel creare una nuova serie storica che sostituisce ad un valore di un periodo la media aritmetica di un certo numero di valori (chiamati termini della media mobile): metà che lo precedono con altrettanti che lo seguono.

A seconda della quantità di dati disponibili si può scegliere il numero di termini della media mobile (preferibilmente dispari per motivi di centralità)



data	temp max	a 3 termini	a 5 termini	a 7 termini
01/12/2017	4			
02/12/2017	8	6,3		
03/12/2017	7	8,0	7,4	
04/12/2017	9	8,3	9,0	8,6
05/12/2017	9	10,0	9,6	8,9
06/12/2017	12	10,7	9,4	8,6
07/12/2017	11	9,7	8,8	8,6
08/12/2017	6	7,7	8,4	9,7
09/12/2017	6	6,3	9,4	10,9
10/12/2017	7	10,0	10,6	10,4
11/12/2017	17	13,7	11,2	9,9
12/12/2017	17	14,3	11,4	9,9
13/12/2017	9	11,0	11,2	9,9
14/12/2017	7	7,3	9,0	10,0
15/12/2017	6	6,3	7,2	8,3
16/12/2017	6	6,7	6,4	6,6
17/12/2017	8	6,3	6,0	6,0
18/12/2017	5	6,0	5,8	5,7
19/12/2017	5	5,0	5,6	5,7
20/12/2017	5	5,0	5,2	6,3
21/12/2017	5	5,3	6,2	6,9
22/12/2017	6	7,0	7,6	7,6
23/12/2017	10	9,3	8,6	7,7
24/12/2017	12	10,7	8,8	8,7
25/12/2017	10	9,3	10,0	8,7
26/12/2017	6	9,3	9,0	9,0
27/12/2017	12	7,7	8,2	8,6
28/12/2017	5	8,3	7,6	7,9
29/12/2017	8	6,7	7,8	
30/12/2017	7	7,3		
31/12/2017	7			



## Metodo delle medie mobili

Qui a fianco sono le quotazioni giornaliere del Ftse Mib degli ultimi 5 anni, cioè circa 1300 osservazioni, (grafico in alto) comparate con le medie mobili 21 termini (grafico al centro) e comparate con medie mobili a 200 termini (grafico in basso)

Più si allarga l'intervallo di termini considerati, più si appiattisce la curva rappresentante la media mobile, questo per l'effetto livellatore che ha l'operatore media aritmetica.

Con questo metodo la descrizione del trend può essere solamente grafica.



## Metodo dei minimi quadrati

Questo metodo ha un approccio formale molto più elegante e ha come obiettivo quello di individuare il valore dei parametri di una funzione che descrive il trend del fenomeno.

Questi parametri devono essere tali da minimizzare le distanze tra i valori reali e quelli teorici immagini della funzione.

Per ogni osservazione quindi costruisce il quadrato della distanza e deve essere minima la somma di quei quadrati.

Se considerassimo il modello di trend come un modello lineare del tipo  $y=mx+q$  l'obiettivo è quello di determinare quei valori di  $m$  e  $q$  che rendono minima la sommatoria.

Si dimostra che sviluppando al quadrato l'argomento della sommatoria si ottiene una funzione  $S(m,q)$ , ponendo nulle le derivate parziali e risolvendo si ottiene:

$$S = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \hat{y}_i = mx_i + q$$

$$S = \sum_{i=1}^n (y_i - mx_i - q)^2$$

$$\begin{cases} \sum_{i=1}^n y_i - m \sum_{i=1}^n x_i - nq = 0 \\ \sum_{i=1}^n y_i x_i - m \sum_{i=1}^n x_i^2 - q \sum_{i=1}^n x_i = 0 \end{cases}$$

$$m = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}$$

$$q = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}$$

## Metodo dei minimi quadrati

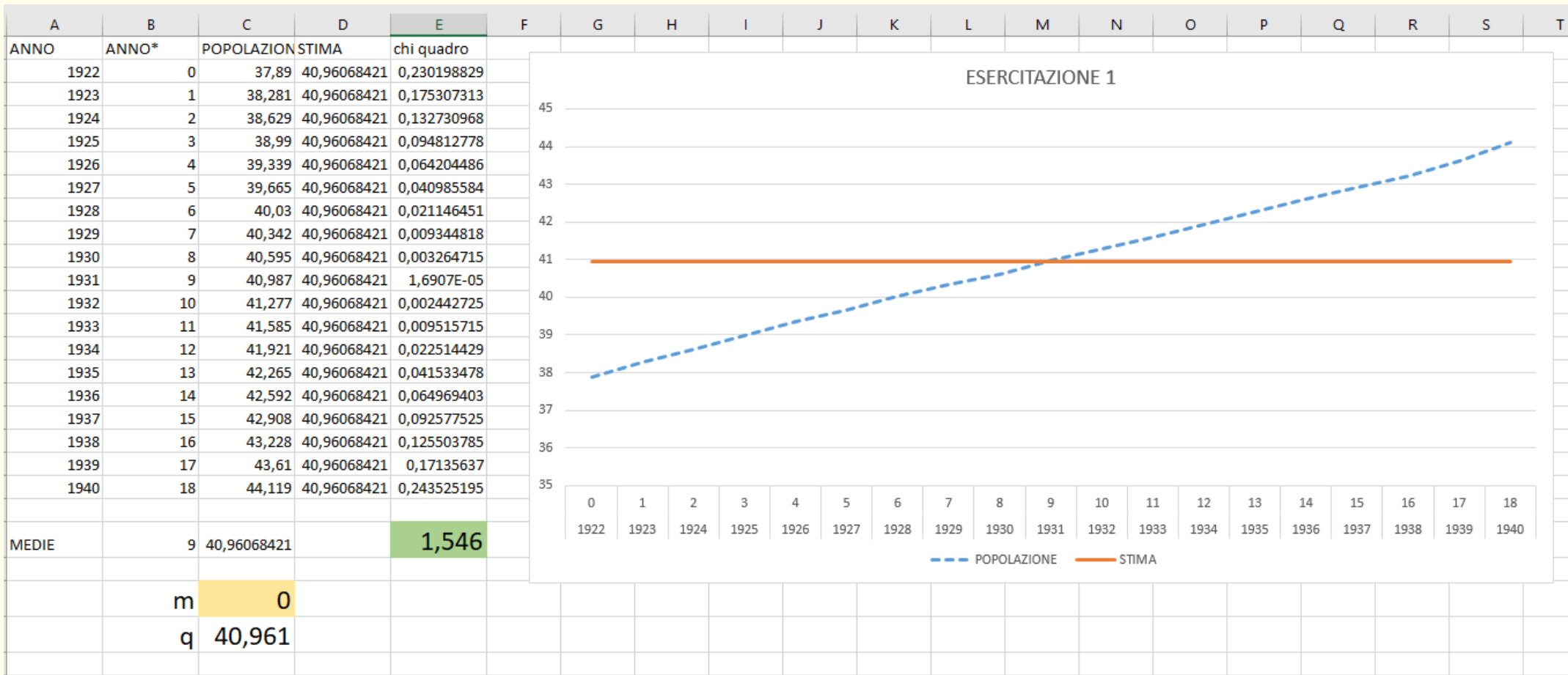
In generale potremmo schematizzare il principio in questo modo: occorre trovare il minimo della sommatoria delle distanza fra valori reali e teorici i quali presentano dei parametri il cui valore incide direttamente sul totale. Questo totale quindi è una funzione che ha all'interno l'operatore sommatoria e che presenta tante variabili quanti sono i parametri necessari al modello scelto per rappresentare il trend.

È evidente come questa spiegazione si inserisca a fatica nella didattica per la complessità già presente in un modello teorico che presenta solo due parametri.

Pertanto dovremmo tener presente un altro tipo di approccio didattico che può svilupparsi attraverso due strade: la prima «per tentativi» creando una tabella che presenta dati reali e teorici sulla base di una funzione scelta, calcoli automaticamente le differenze fra modello e realtà e porti a riflettere sulle conseguenze del valore modificato; la seconda invece facendo calcolare ad un software i parametri opportuni.

$$S = \sum_{i=1}^n (y_i - f(x_i, a_1, a_2, a_3 \dots))^2$$

# Riflettere sui valori per tentativi



## Utilizzare un'applicazione

Formato linea di tendenza

Opzioni linea di tendenza

Opzioni linea di tendenza

Esponenziale

Lineare

Logaritmica

Polinomiale Ordine 2

Potenza

Media mobile Periodo 2

Nome linea di tendenza

Automatica Lineare (Serie1)

Personalizza

Previsione

Futura 0,0 periodi

Verifica 0,0 periodi

Imposta intercetta 0,0

Visualizza l'equazione sul grafico

Visualizza il valore R quadrato sul grafico

## EXCEL

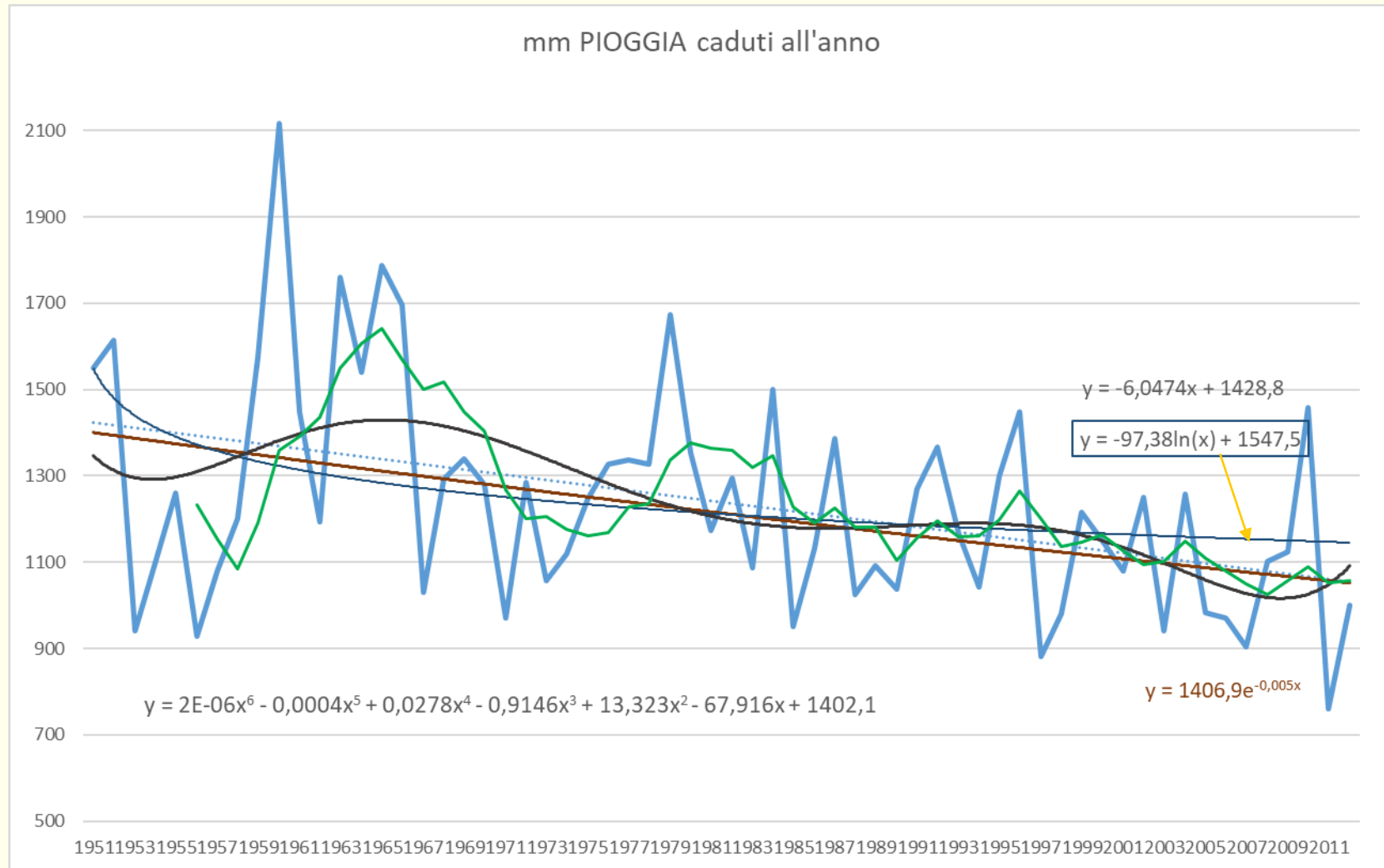
Posizionandosi sul grafico delle osservazioni e cliccando il pulsante destro appare un menu a tendina dove si deve selezionare «Aggiungi linea di tendenza»

Apparirà quindi questa finestra di dialogo dove l'utente può scegliere tra diversi modelli matematici proposti ed anche la modalità delle medie mobili specificando il numero di termini.

È possibile poi che nel grafico venga visualizzata l'equazione.

Il confronto fra le equazioni può portare a riflessioni sulla qualità della approssimazione e comprensione dei valori dei parametri presenti.

# Utilizzare un'applicazione

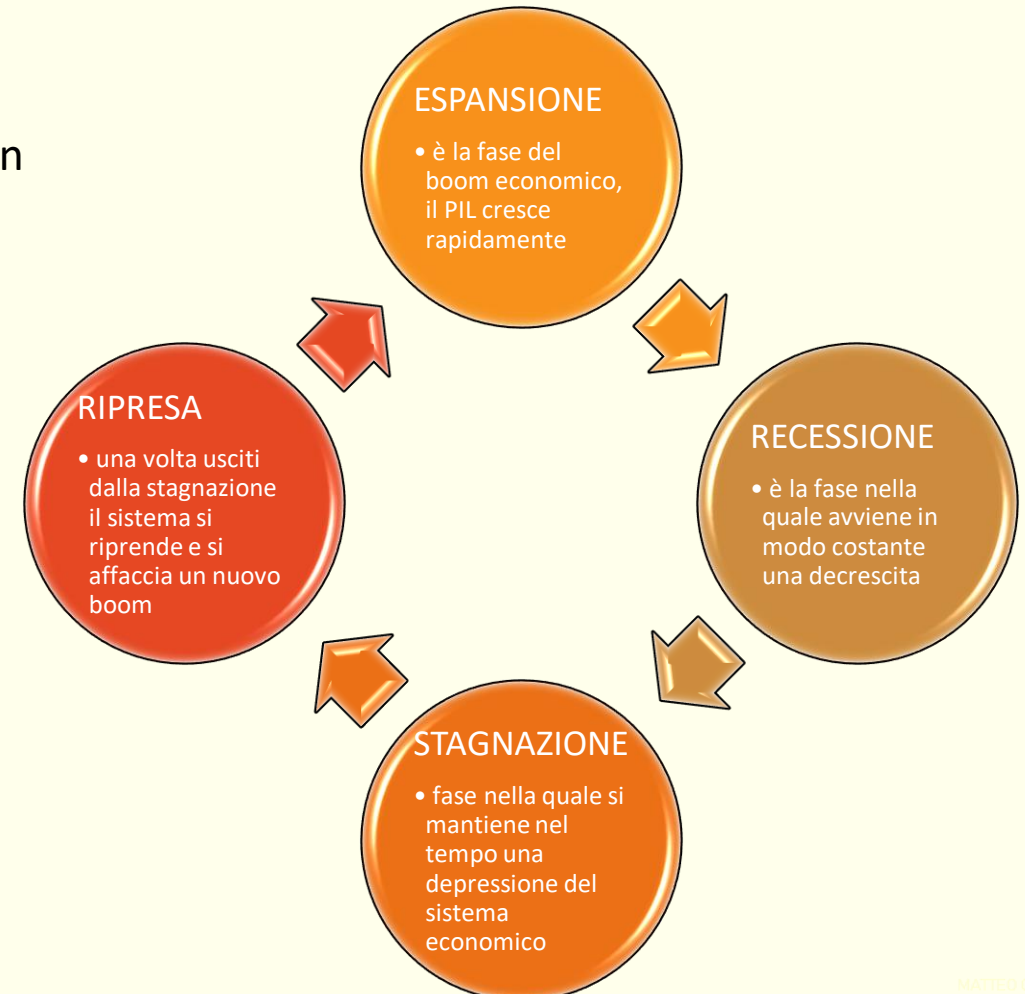
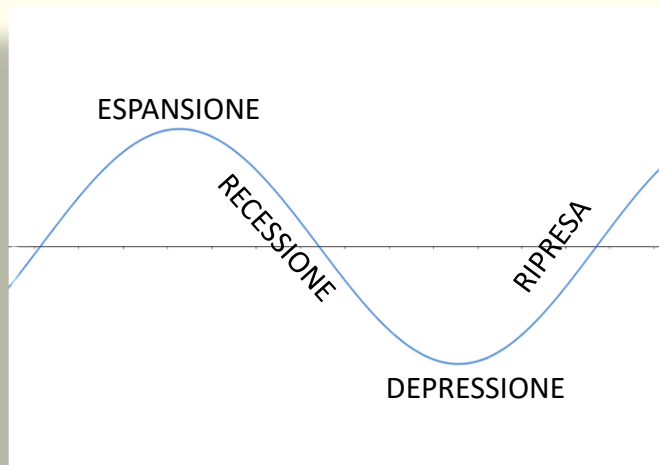


# La ciclicità



Per ciclo economico si intende quella fluttuazione economica di medio lungo periodo che alterna in maniera ordinata, ma non regolare, quattro fasi che caratterizzano l'attività economica di un paese il cui indicatore principale risulta essere il PIL.

Le fasi sono le seguenti



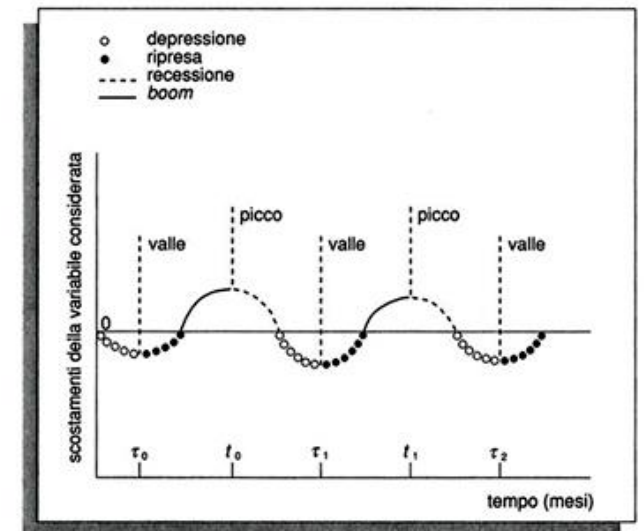
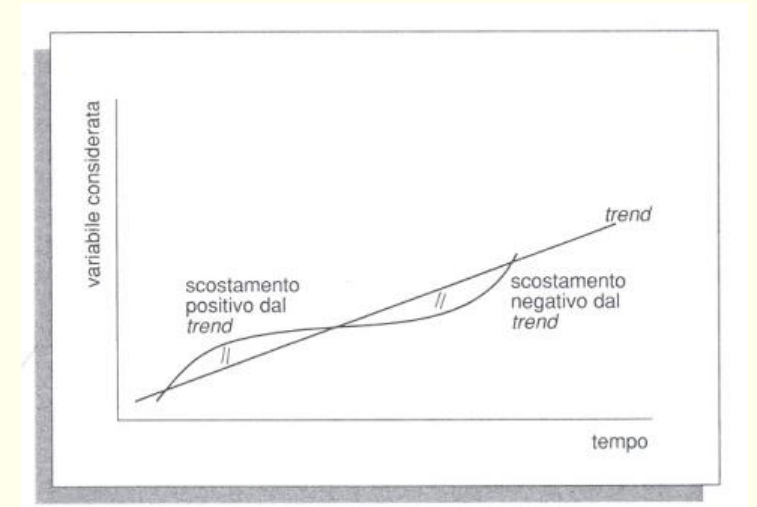


## Come influisce la ciclicità

Considerando quindi un particolare valore osservato, quel valore risulterà al di sopra del trend in fase ciclica positiva mentre inferiore al trend perché il sistema economico è in un momento di recessione.

La componente ciclica informa di quanto dal dato atteso ci si sposta per effetto della congiuntura economica.

Ci sono beni economici che risentono direttamente del ciclo ed altri che risentono in maniera indiretta, in questo caso la componente ciclica avrà segno opposto rispetto a quello estratto dal PIL.





## Durata dei cicli economici

La durata di un ciclo economico può riguardare intervalli di tempo molto diversi:

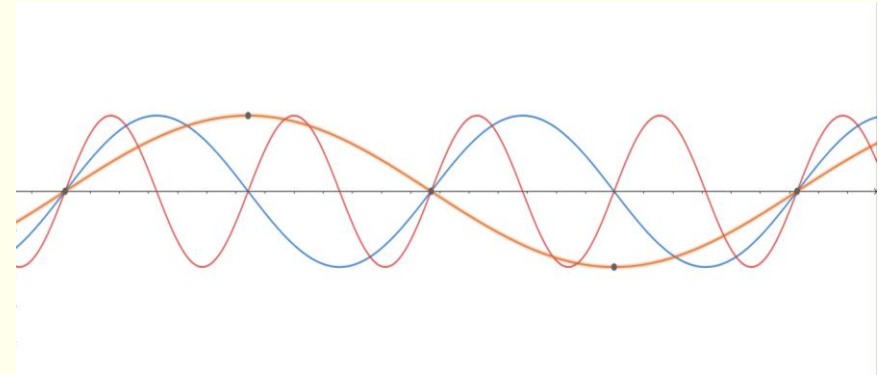
Legati alle variazioni delle scorte (cicli brevi 2-4 anni)

Legati al credito e alle riserve bancarie (cicli medi 4-10 anni) (ciclo biblico)

Legati alle infrastrutture (cicli lunghi 10-25 anni)

Legati al cambio di sistema (50 - 60 anni)

L'analisi del ciclo economico e le teorie relative sono molto collegati alle scelte di politica economica di uno stato soprattutto in termini di tassi di sconto (più in generale di politiche monetarie), investimenti pubblici, politiche del lavoro.



Molte analisi delle serie storiche inseriscono la componente ciclica nel trend parlando di componente trend-ciclo. Per semplicità quindi adotteremo questa scelta ragionando soprattutto sulla componente stagionale.

# La stagionalità



Per stagionalità si intende l'identificazione e isolamento dell'effetto stagionale sul fenomeno.

Le abitudini sui consumi ad esempio riflettono in misura più o meno marcata la componente stagionale.

Fenomeni naturali come la temperatura media, i millimetri di pioggia, la portata di un fiume, il raccolto di un determinato frutto, sono fortemente influenzati dal ciclo delle stagioni.

**Questa componente si mantiene costante annualmente** e agisce con lo stesso peso in corrispondenza dello stesso periodo in anni successivi.

Questo tipo di componente la si può analizzare solo se le osservazioni sono in frazione di anno.



## Come isolare la componente stagionale

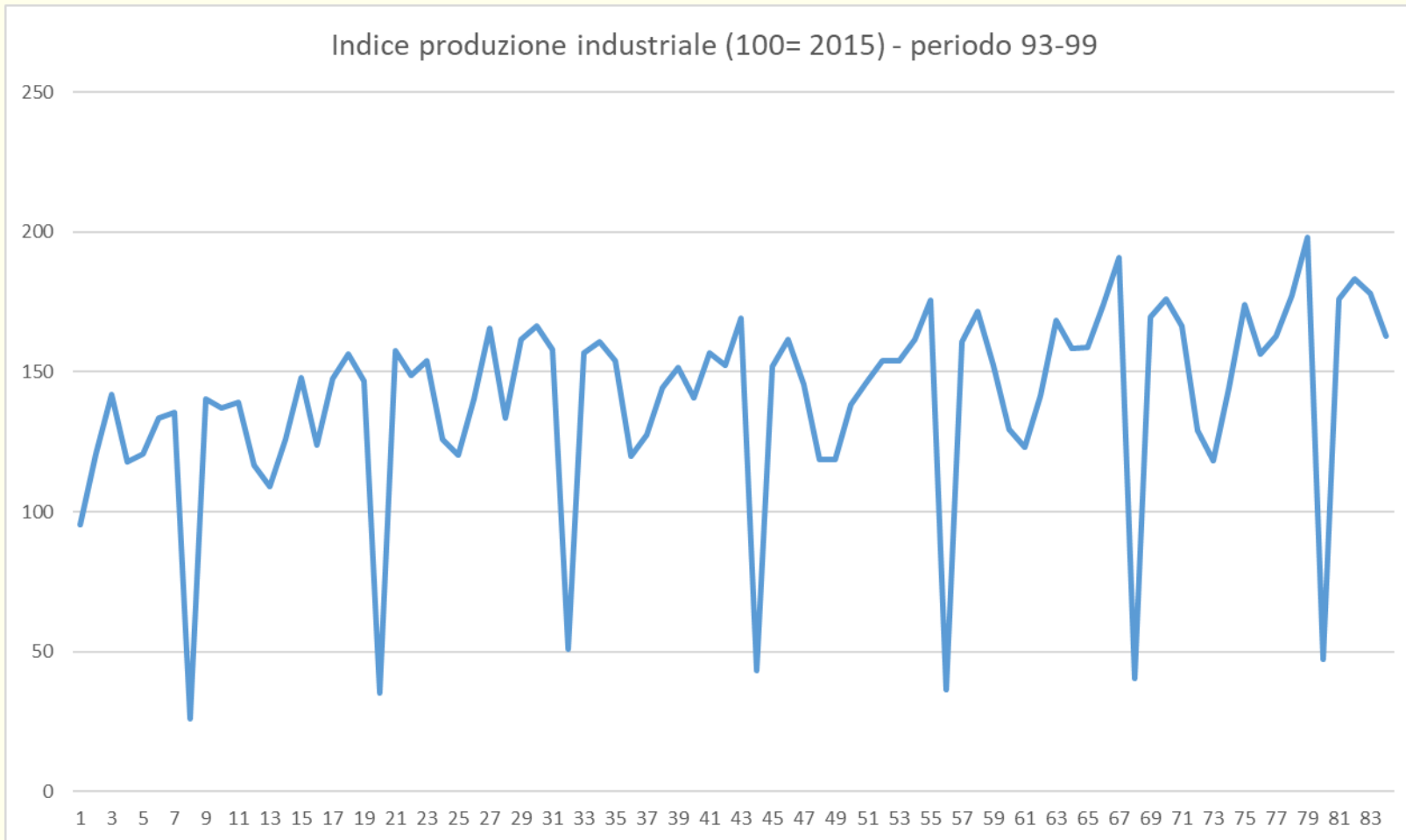
Per isolare la componente stagionale occorre disporre già del dato di trend, confrontare lo scarto tra osservazione reale e dato di trend per più anni e isolare da quelle differenze la componente stagionale.

L'ipotesi alla base è che l'influenza del periodo dell'anno agisca sempre con lo stesso peso.

Prendiamo come esempio la serie storica dei valori mensili dell'indice della produzione industriale in Italia nel settennio 1993-1999.

Anno	Mesi											
	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novembre	Dicembre
1993	95,4	120,5	141,8	117,8	120,6	133,4	135,4	25,8	140,3	137,1	139,3	116,6
1994	109,0	125,9	147,8	124,0	147,7	156,5	146,9	35,1	157,7	148,6	154,1	125,8
1995	120,4	140,3	165,5	133,7	161,5	166,5	158,0	51,0	156,8	160,9	154,1	120,0
1996	127,3	144,2	151,5	140,7	156,8	152,2	169,1	43,2	152,1	161,8	145,6	118,8
1997	118,7	138,4	146,5	154,1	153,9	161,6	175,5	36,5	160,9	171,8	152,2	129,5
1998	123,0	141,7	168,4	158,3	158,6	174,1	190,8	40,5	169,6	175,9	166,4	129,1
1999	118,2	144,5	174,1	156,4	162,9	177,4	198,0	47,1	176,2	183,2	178,2	162,7

## Prima analisi: esiste una componente stagionale?



Osservando l'andamento del grafico che riporta i dati grezzi, notiamo come i valori sembrano tendere verso l'alto anche se in presenza di forti oscillazioni. Queste si manifestano con regolarità periodiche di picchi sia verso il basso che verso l'alto.

## Individuiamo la componente del trend / ciclo

L'idea di crescita ci fa pensare ad una linearità.

Individuiamo pertanto il modello matematico lineare che gestisce il fenomeno.

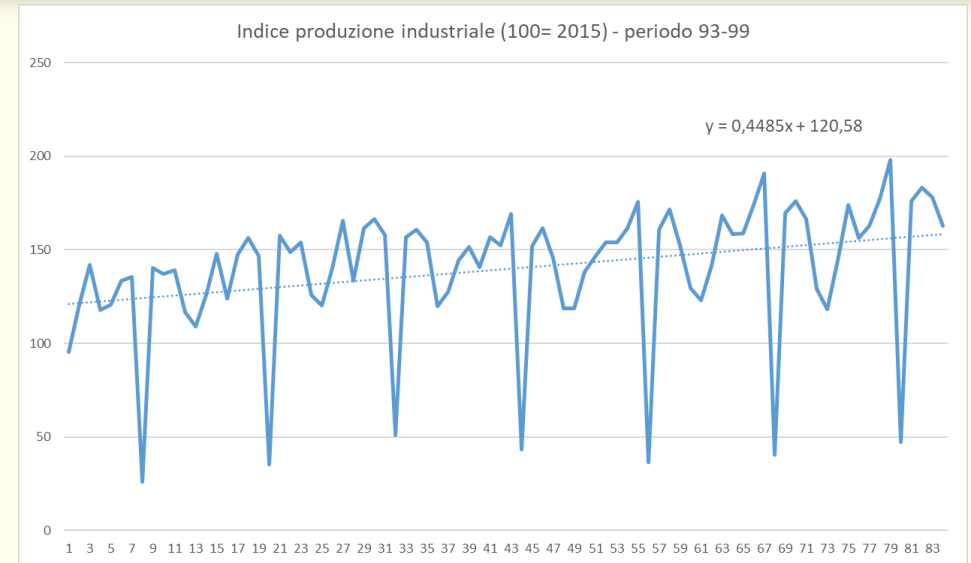
$$\hat{y} = 0,4485x + 120,58$$

Possiamo interpretarlo in questo modo: nel periodo considerato si è registrato un aumento medio mensile di 0,4485 partendo da un valore iniziale di 120,58 (dicembre 1992).

In questo modo possiamo creare a fianco della serie storica dei dati osservati, quella dei dati puri.

La differenza tra i due valori risente della componente stagionale e di quella accidentale.

Creiamo la serie storica delle differenze.



periodo	indice	T+C
Gen-1993	95,4	121,0285
Feb-1993	120,5	121,477
Mar-1993	141,8	121,9255
Apr-1993	117,8	122,374
Mag-1993	120,6	122,8225
Giu-1993	133,4	123,271
Lug-1993	135,4	123,7195
Ago-1993	25,8	124,168
Set-1993	140,3	124,6165
Ott-1993	137,1	125,065
Nov-1993	139,3	125,5135
Dic-1993	116,6	125,962
Gen-1994	109	126,4105
Feb-1994	125,9	126,859
Mar-1994	147,8	127,3075

## Definiamo la componente stagionale

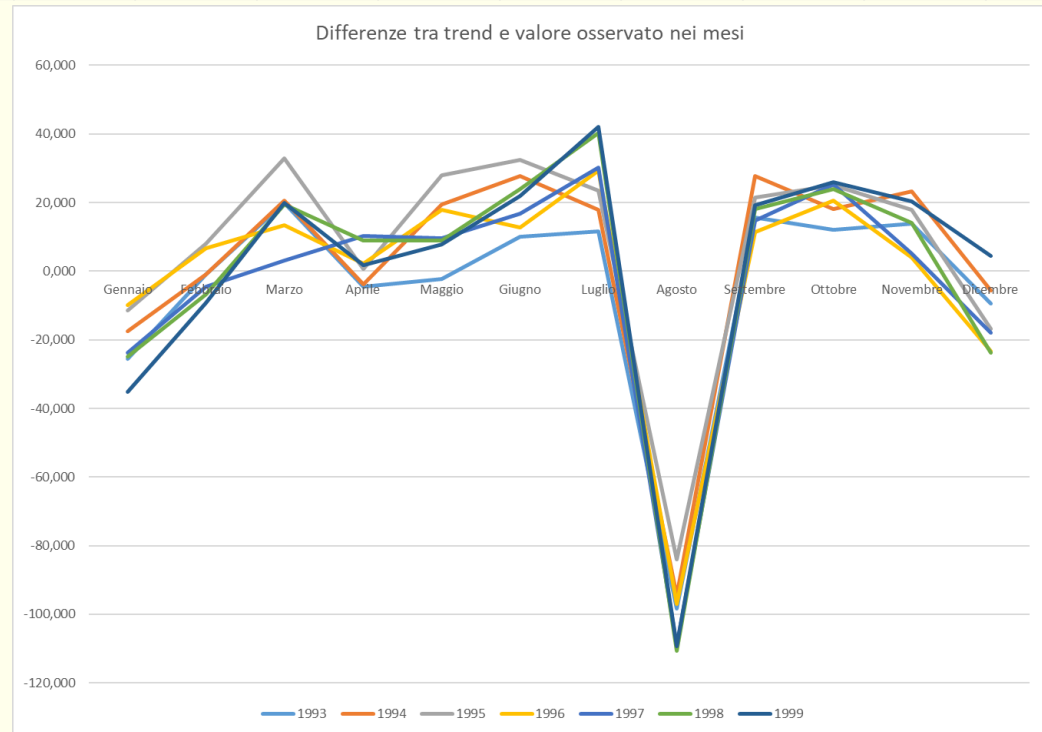
Nella tabella riportiamo queste differenze. Se la stagionalità è presente il loro evolversi sarà regolare.

ANNO	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novembre	Dicembre
1993	-25,629	-0,977	19,875	-4,574	-2,223	10,129	11,681	-98,368	15,684	12,035	13,787	-9,362
1994	-17,411	-0,959	20,493	-3,756	19,496	27,847	17,799	-94,450	27,702	18,153	23,205	-5,544
1995	-11,393	8,059	32,811	0,562	27,914	32,465	23,517	-83,932	21,420	25,071	17,823	-16,726
1996	-9,875	6,577	13,429	2,180	17,832	12,783	29,235	-97,114	11,338	20,589	3,940	-23,308
1997	-23,857	-4,605	3,047	10,198	9,549	16,801	30,253	-109,196	14,756	25,207	5,159	-17,990
1998	-24,939	-6,687	19,565	9,016	8,868	23,919	40,171	-110,578	18,074	23,925	13,977	-23,772
1999	-35,121	-9,269	19,883	1,734	7,786	21,837	41,989	-109,360	19,292	25,843	20,395	4,446

Vediamo graficamente le differenze anno per anno e notiamo come il comportamento sia diverso (ovvio) ma abbia caratteristiche simili.

Calcolando la media delle differenze per ogni mese associamo a quel valore la componente stagionale del fenomeno.

Nella tabella in fondo vengono individuati gli addendi della componente stagionale per il periodo considerato.



	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novembre	Dicembre
media	-21,175	-1,123	18,443	2,194	12,746	20,826	27,806	-100,428	18,323	21,546	14,041	-13,179



## L'analisi dei residui

Riconsideriamo l'equazione iniziale

$$y_i = T_i + C_i + S_i + A_i$$

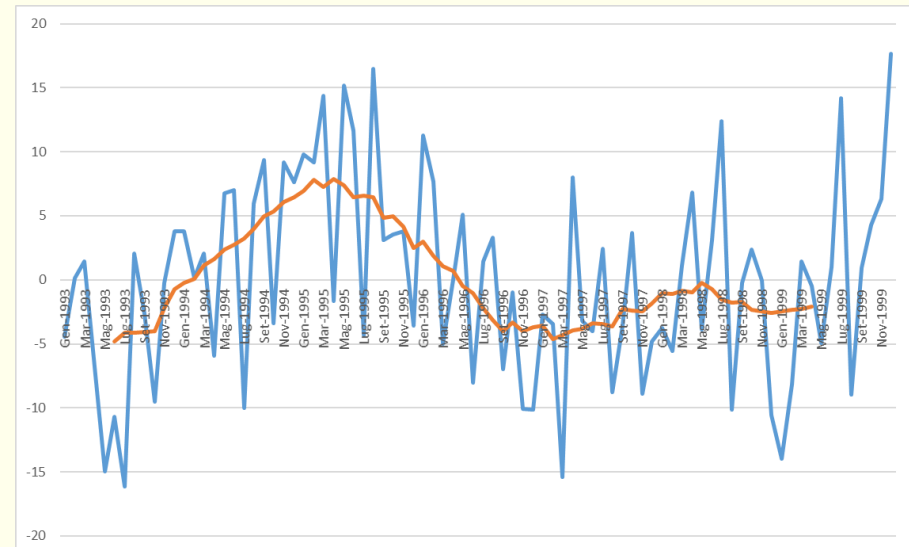
In questo momento abbiamo assunto che i primi due addendi siano un valore unico espresso dal modello lineare, quindi abbiamo isolato la stagionalità.

Possiamo perciò creare la serie destagionalizzata (ovvero che non risente della componente ciclica mensile)

L'oscillazione di questa serie storica considererò l'aspetto accidentale (e ciclico)

Se la forma grafica della serie storica è perfettamente irregolare allora non c'è componente ciclica altrimenti è possibile cercare di individuare l'incidenza del ciclo economico sui dati.

periodo	indice	T+C	stagionali	comp st	T+C+S	A
Gen-1993	95,4	121,0285	-25,6285	-21,1745	99,854	-4,454
Feb-1993	120,5	121,477	-0,977	-1,123	120,354	0,146
Mar-1993	141,8	121,9255	19,8745	18,44279	140,3683	1,431714
Apr-1993	117,8	122,374	-4,574	2,194286	124,5683	-6,76829
Mag-1993	120,6	122,8225	-2,2225	12,74579	135,5683	-14,9683
Giu-1993	133,4	123,271	10,129	20,82586	144,0969	-10,6969
Lug-1993	135,4	123,7195	11,6805	27,80593	151,5254	-16,1254
Ago-1993	25,8	124,168	-98,368	-100,428	23,73971	2,060286
Set-1993	140,3	124,6165	15,6835	18,32321	142,9397	-2,63971
Ott-1993	137,1	125,065	12,035	21,54614	146,6111	-9,51114
Nov-1993	139,3	125,5135	13,7865	14,0405	139,554	-0,254
Dic-1993	116,6	125,962	-9,362	-13,1794	112,7826	3,817429
Gen-1994	109	126,4105	-17,4105	-21,1745	105,236	3,764
Feb-1994	125,9	126,859	-0,959	-1,123	125,736	0,164

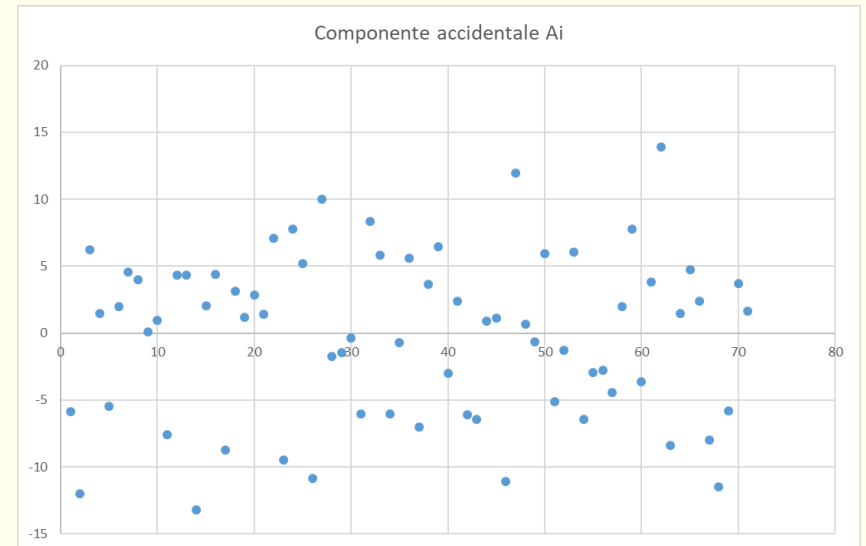
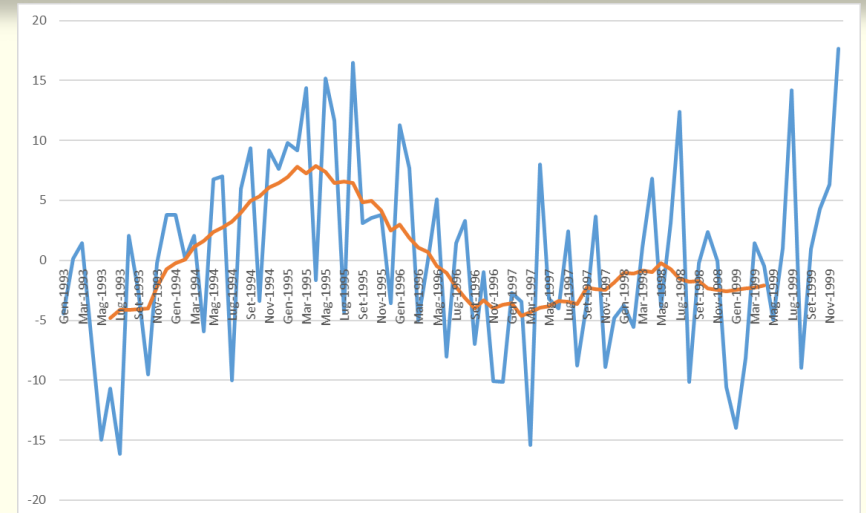


## La componente accidentale

Attraverso la comparazione dei valori dei residui (A) e calcolandone la media mobile a 12 termini si è evidenziata la presenza di un ciclo economico.

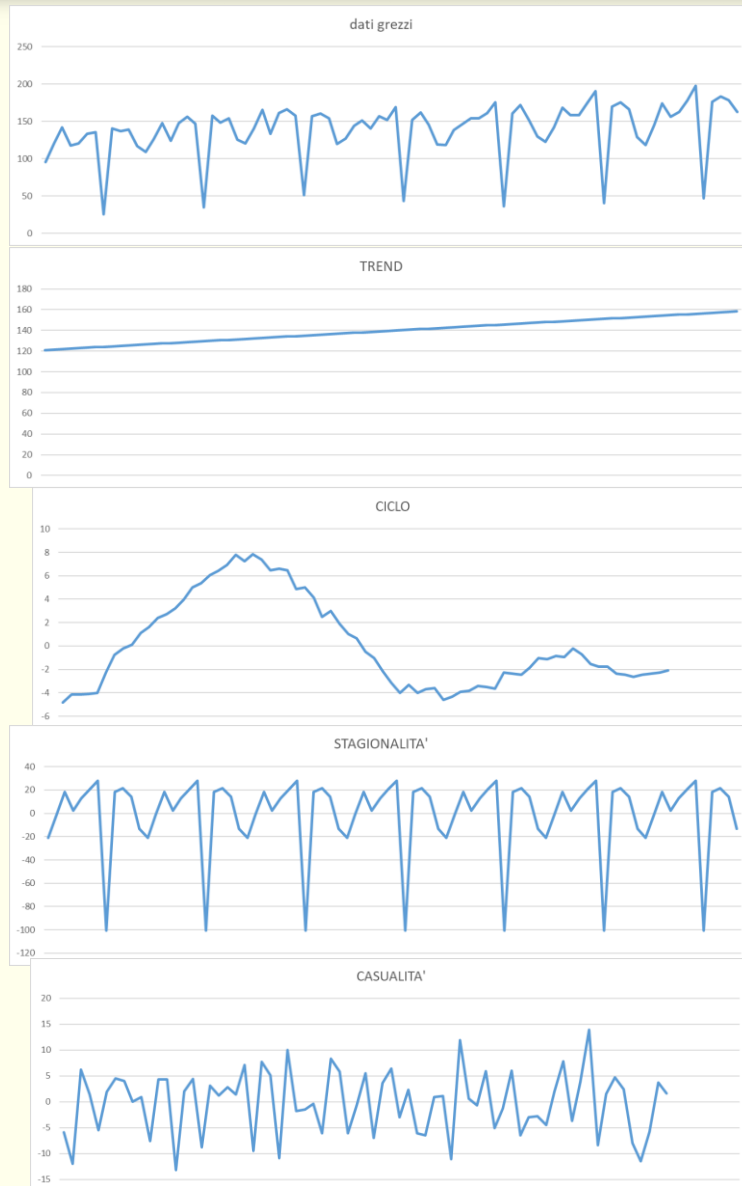
La differenza del residuo con la media mobile che rappresenta il ciclo economico individua la componente accidentale che graficamente non deve minimamente dare delle impressioni di regolarità

Come criterio di controllo, si assume che la media dei valori della componente accidentale dovrebbe essere 0 (perché non deve rappresentare nel lungo periodo né un beneficio né una limitazione), nel nostro caso la media dei valori della componente accidentale è -0,15 che è un valore decisamente accettabile.





## Conclusioni



Ogni valore reale risulta dalla somma di 4 componenti:

Il trend ovvero l'aspetto puro del fenomeno

+

La stagionalità presente nel fenomeno

+

La presenza di un ciclo economico

+

Una componente accidentale

periodo	indice	T	S	C	A
Mag-1997	153,9	144,3505	12,74579	-3,83469	0,638405