Serie Storiche

Introduzione all'analisi delle serie storiche

Corso di Laurea in Statistica per le Tecnologie e le Scienze
Corso di Laurea in Statistica per l'Economia e l'Impresa
matricole dispari

Francesco Lisi

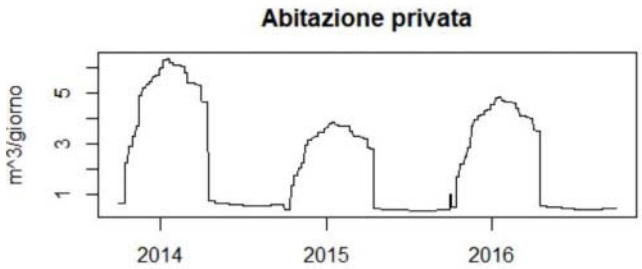
a.a. 2026/2027

Le serie storiche o temporali

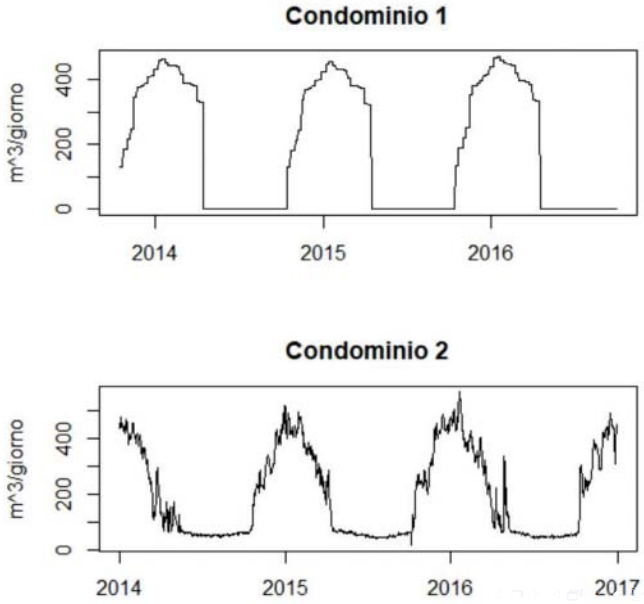
- **Serie statistica**
  Collezione di osservazioni ciascuna caratterizzata dalla modalità di una o più variabili qualitative o quantitative.
- **Serie storica (temporale)**
  Successione di osservazioni su una variabile numerica fatte sequenzialmente nel tempo.

- L'ordine dei dati non è accidentale: elementi ordinati in modo naturale dal valore dell'indice temporale.
- Osservazioni dipendenti: i dati presentano "regolarità" o persistenze legate alla posizione dell'osservazione nella sequenza.

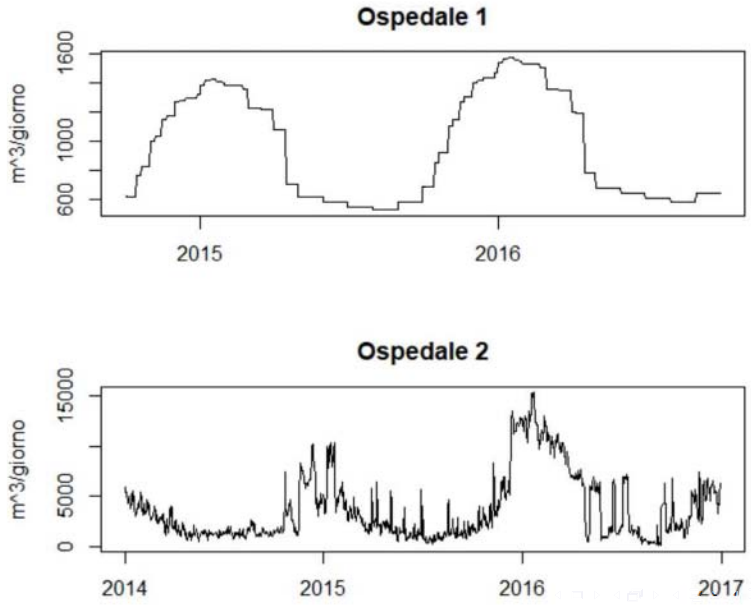
Consumo giornaliero di gas in alcune realtà dell'Alto Adige (01.10.2013 - 31.12.2016)



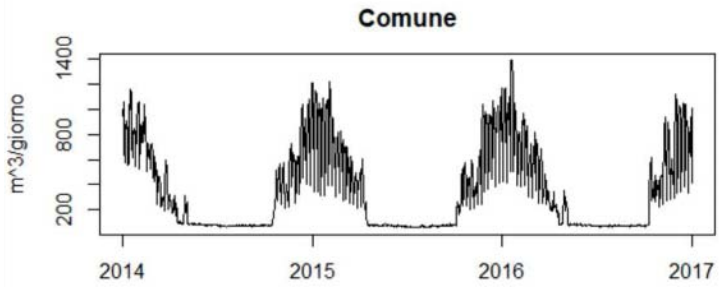
Consumo giornaliero di gas in alcune realtà dell'Alto Adige (01.10.2013 - 31.12.2016)



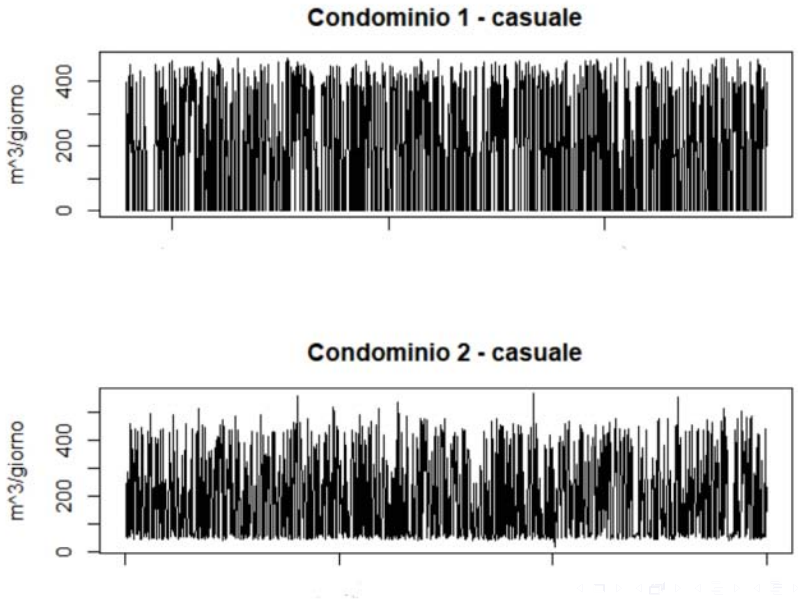
Consumo giornaliero di gas in alcune realtà dell'Alto Adige (01.10.2013 - 31.12.2016)



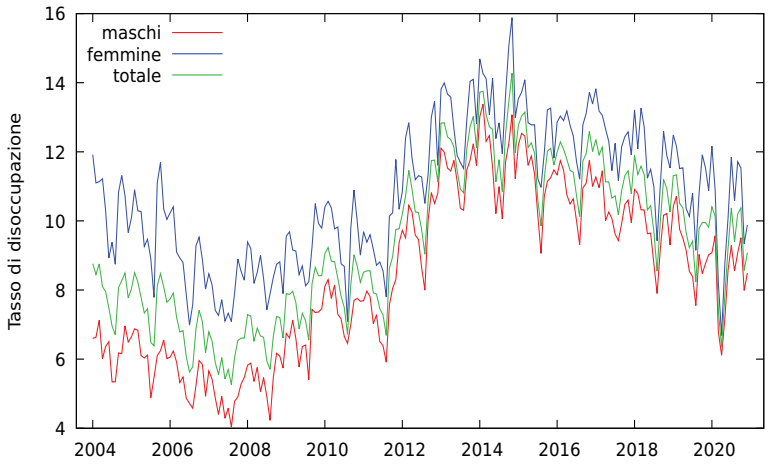
Consumo giornaliero di gas in alcune realtà dell'Alto Adige (01.10.2013 - 31.12.2016)



Se non tenessimo conto dell'ordinamento temporale...

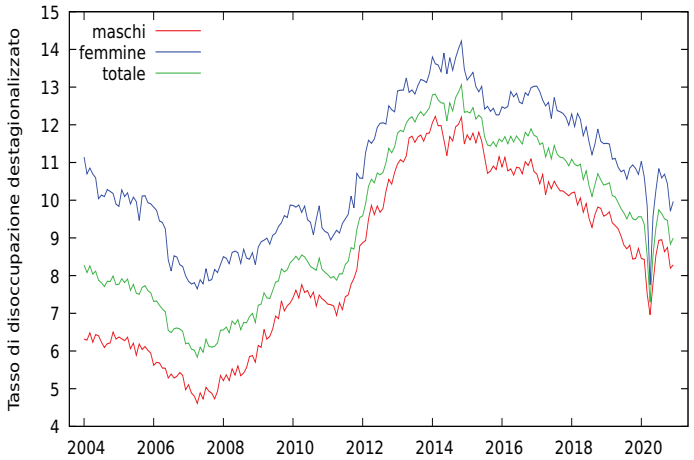


Tasso di disoccupazione mensile - Italia, 2004:1 - 2020:07, dati grezzi, fonte: ISTAT

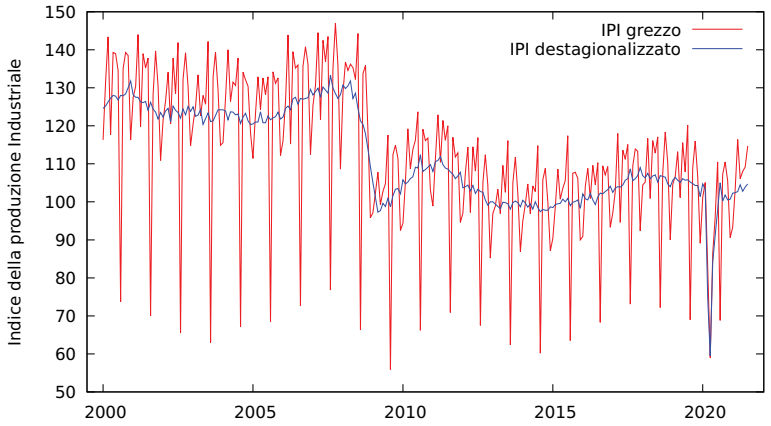


Tasso di disoccupazione mensile - Italia, 2004:1 - 2020:07, dati destagionalizzati,

fonte: ISTAT



Indice mensile della Produzione Industriale - Italia, dati grezzi e destagionalizzati, base 2015=100, 2000:1 - 2020:07, fonte: ISTAT



Prodotto Interno Lordo ai prezzi di mercato - Italia, dati grezzi e destagionalizzati

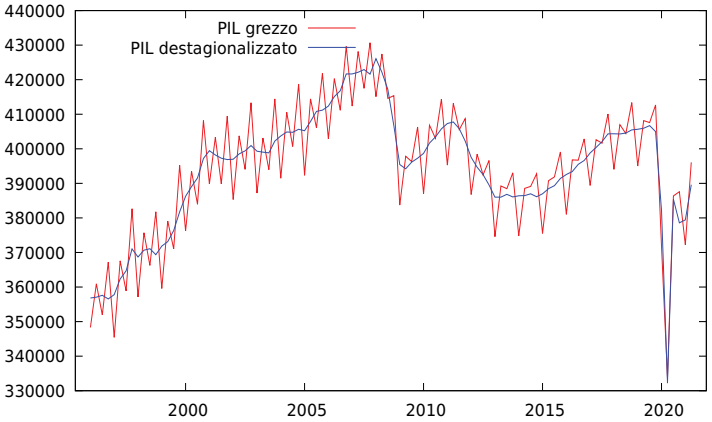
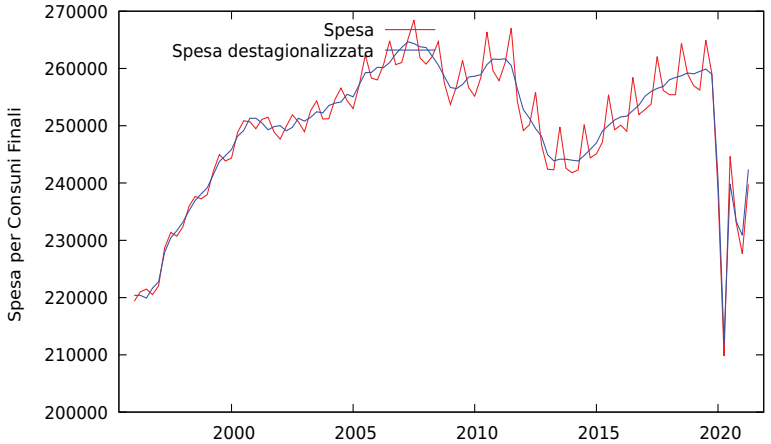


Figure: Milioni di euro, valori concatenati con anno di riferimento 2015. Serie trimestrale. 1996:1 - 2020:2. Fonte: FRED

Spesa per Consumi finali nazionali - Italia, dati grezzi e destagionalizzati



Milioni di euro, valori concatenati con anno di riferimento 2015. Serie trimestrale. 1996:1 - 2020:2. Fonte: Istat

Serie giornaliera (5 giorni) del tasso di cambio euro/dollaro

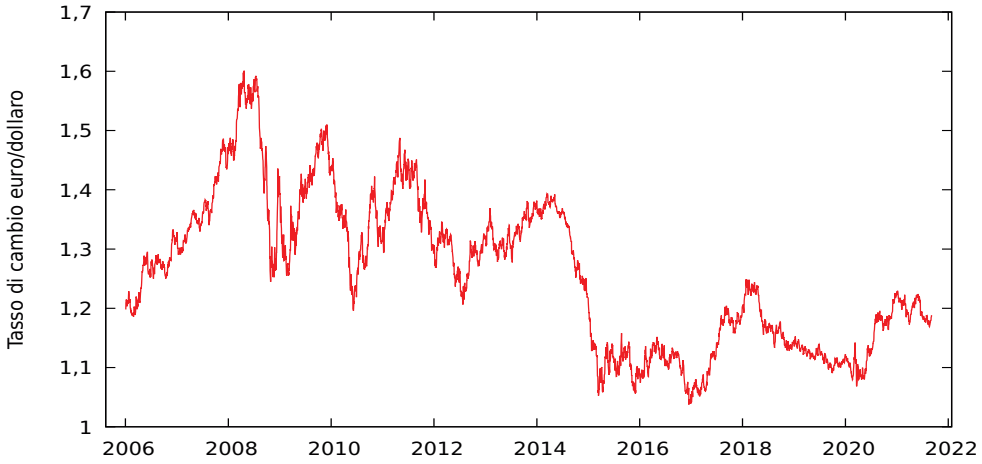


Figure: Dati giornalieri (5 giorni). (8 settembre 2009 - 22 settembre 2020). Fonte: Yahoo Finance

Serie giornaliera (5 giorni) dell'indice azionario Dow Jones

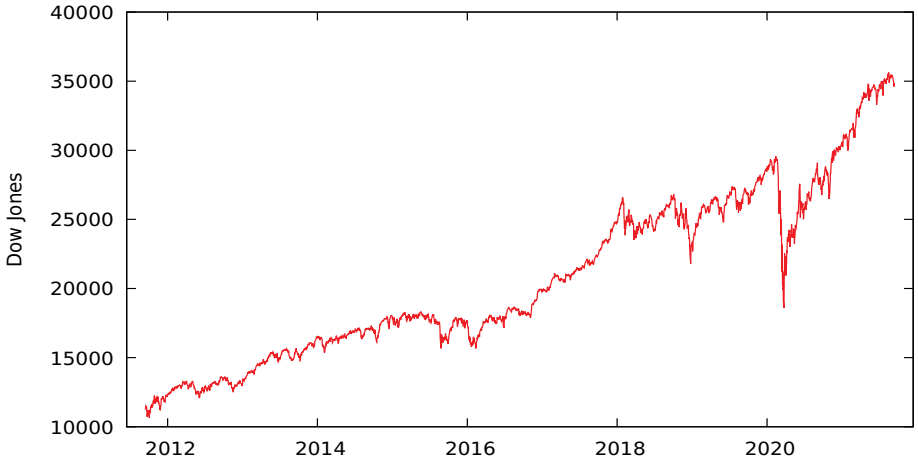


Figure: Dati giornalieri (5 giorni). (2 gennaio 2001 - 21 settembre 2020). Fonte: Yahoo Finance

Indice azionario Dow Jones: rendimenti

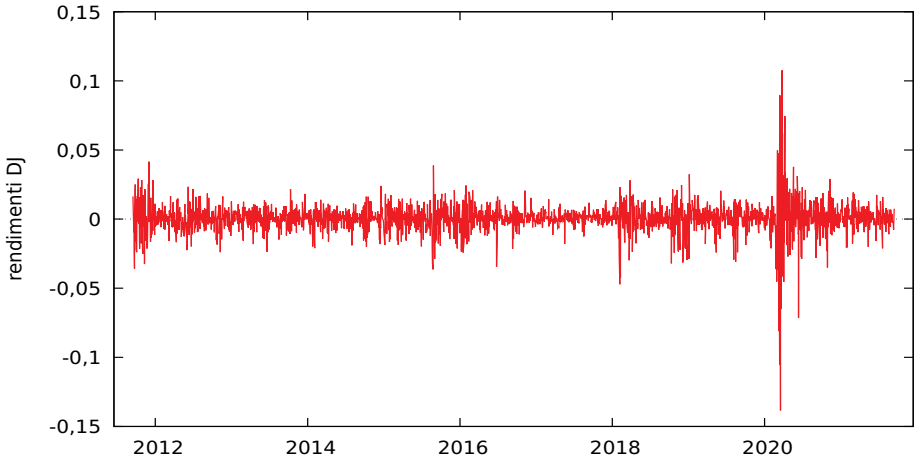


Figure: Dati giornalieri (5 giorni). (2 gennaio 2001 - 21 settembre 2020). Fonte: Yahoo Finance

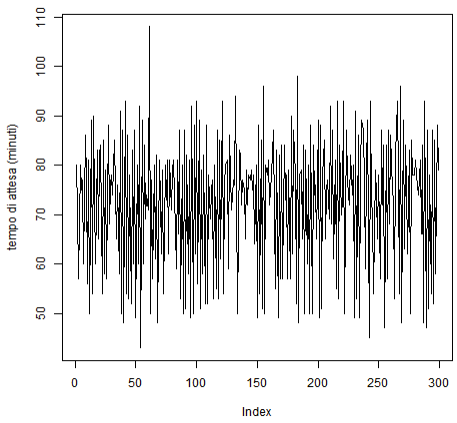
Non solo serie storiche economiche...

- Serie storiche fisiche
  - livello settimanale delle precipitazioni piovose; massimi e minimi giornalieri delle temperature
  - concentrazione di ozono presente nell'aria (media giornaliera, massimi giornalieri delle medie orarie, massimi giornalieri delle medie su 8 ore)
- Serie storiche demografiche
  - popolazione residente annuale; nascite mensili; numero di matrimoni per trimestre
- Serie storiche derivanti dal controllo di processo
- Serie storiche binarie: possono assumere 2 soli valori (es. scelta fra due marche di uno stesso prodotto)
- Serie storiche dati di conteggio (numero accessi ad un Pronto Soccorso)

The Old Faithful gesyer

- Serie dei tempi di attesa tra due eruzioni consecutive (272 eruzioni consecutive)





Serie storica degli accessi orari al PS di Padova tra l'1 marzo 2023 e il 30 aprile 2023

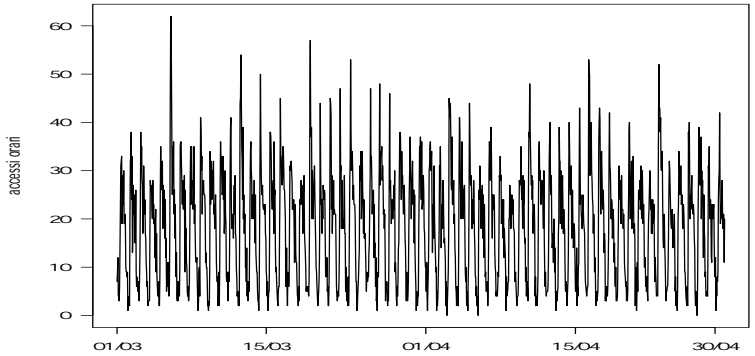


Figure: Accessi al PS: dati orari 1/3/2023-30/04/2023

Accessi orari al PS di Padova Dettaglio per due settimane

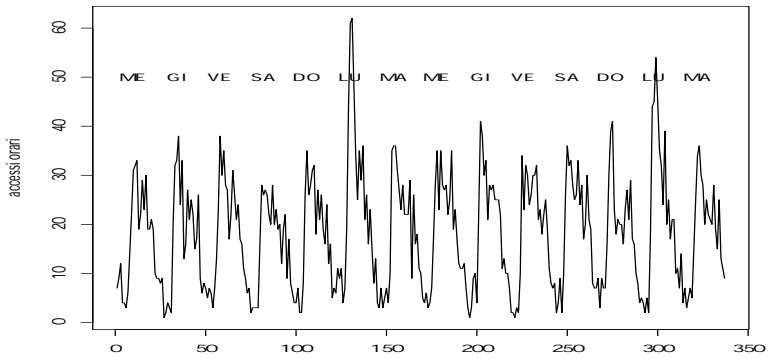


Figure: Accessi al PS: dati orari per due settimane

Serie mensile delle precipitazioni piovose a Padova - stazione Orto Botanico

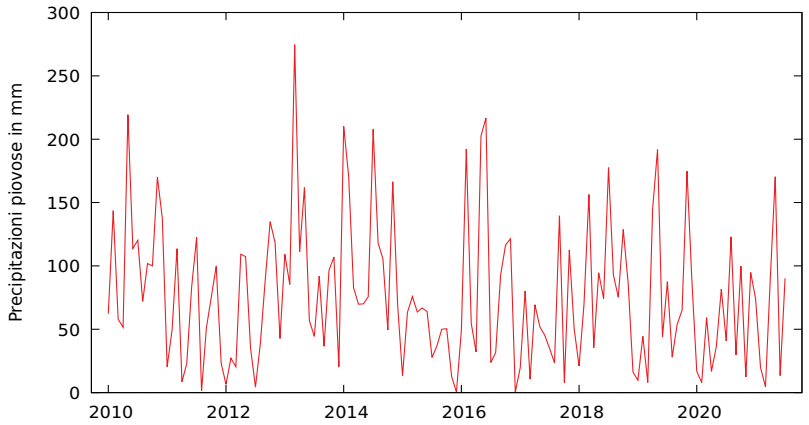


Figure: Dati mensili (gennaio 2010 - luglio 2020). Fonte: ARPAV

Tasso mensile di natalità - Italia

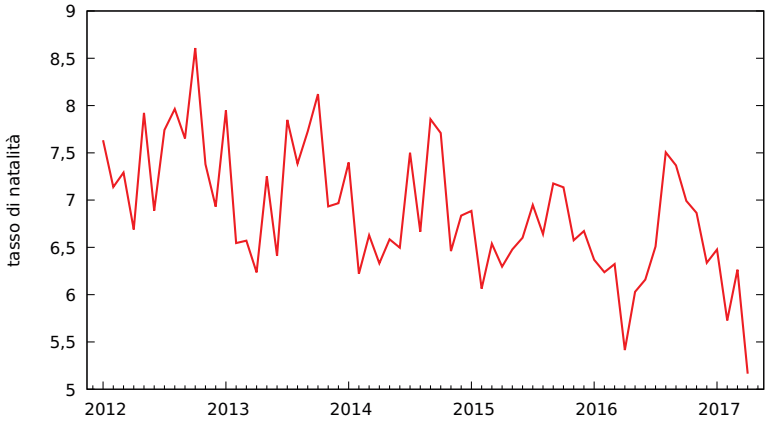


Figure: Dati mensili (gennaio 2012 - aprile 2017). Fonte: ISTAT

Serie annuale dei movimenti sismici di magnitudo ≥ 4.0

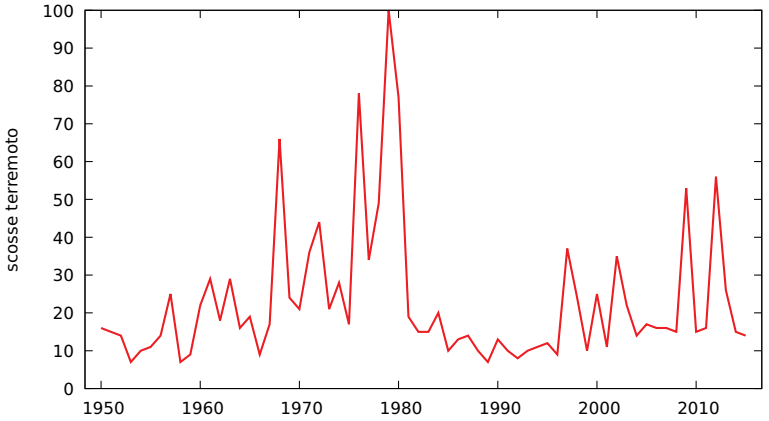


Figure: Dati annuali, 1950-2015. Fonte: ISTAT

Serie mensile delle anomalie della temperatura globale terrestre

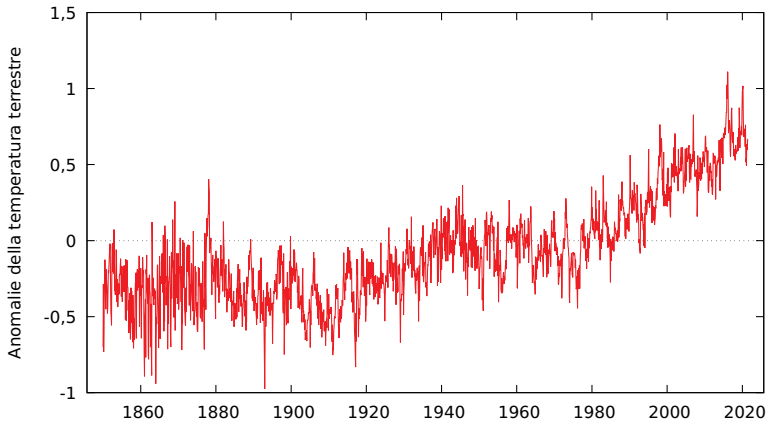


Figure: Dati mensili, 1950:1-2020:7. Fonte: HudCRUT4

Terminologia

- **Serie storica:**
  univariata: dinamica temporale di una sola variabile.
  multivariata: dinamica congiunta di più variabili.
  equispaziata: osservata ad istanti di tempo t equispaziati.
- t : indice temporale, che definisce l'ordinamento dei dati:
  $t \in T$
  T : insieme parametrico continuo o discreto
- Serie storica a parametro continuo o continua (pur potendo il fenomeno osservato essere non continuo)
- Serie storica a parametro discreto o discreta (pur potendo il fenomeno osservato essere non discreto)

Terminologia

- Caratteristiche delle serie storiche di cui ci occuperemo:

  Univariate, a parametro discreto, riferite a tempi (intervalli o istanti) equispaziati:

$$y_1, \ldots, y_t, \ldots, y_n \quad \Leftrightarrow \quad \{y_t\}_{t=1}^{n}$$

- Scelta della cadenza temporale con cui osservare il fenomeno (frequenza)
- **Serie storica di dati di flusso** (Pil, consumi, reddito, ecc.): si ottiene aggregando o cumulando

  Serie storica di dati di stock (base monetaria, popolazione di un paese, ecc.)

Obiettivi dell'analisi

- **Descrizione**: prime informazioni sulle caratteristiche della serie; individuazione di valori anomali.
- **Spiegazione**: individuazione del meccanismo generatore dei dati.
- **Previsione**: inferenza sull'andamento futuro del fenomeno.
- **Filtraggio**: stima di componenti non osservabili.
- **Controllo**: di un sistema, di un sistema produttivo etc.

Processo generatore dei dati

- Processo generatore della serie storica osservata

$$Y_t = f(t) + u_t$$

- $f(t)$: componente *sistematica*, puramente deterministica
- "Legge" di evoluzione temporale del processo
- u_t : componente *stocastica*
- Entrambe le componenti sono **non osservabili**

Approccio classico e moderno

- Introduciamo innanzitutto i disturbi puramente casuali (processo a componenti incorrelate) o *white noise* (WN), ϵ_t :

$$\begin{aligned} E(\epsilon_t) &= 0 \quad \forall t \\ Var(\epsilon_t) &= \sigma_\epsilon^2 \quad \forall t \\ E[\epsilon_r \epsilon_s] &= 0 \quad r \neq s \end{aligned}$$

- **Approccio classico**: $u_t \equiv \epsilon_t$
- **Approccio moderno**: u_t è un processo a componenti correlate

$$u_t = g(Y_{t-1}, Y_{t-2}, \ldots, \epsilon_{t-1}, \epsilon_{t-2}, \ldots) + \epsilon_t \qquad (1)$$

Le componenti di una serie storica

$$f(t) = f(T_t, C_t, S_t)$$

- **Trend**: T_t
  - Tendenza di fondo, riferita ad un lungo periodo di tempo
  - Dinamica regolare, legata all'evoluzione strutturale del sistema
- **Ciclo**: C_t
  - Componente legata all'alternarsi di fasi di espansione e di recessione del sistema economico
  - Fluttuazioni di periodo ampio (in genere, qualche anno)

Le componenti di una serie storica

- **Stagionalità**: S_t
  - descrive effetti che si ripetono ciclicamente entro un certo periodo (ad es., l'anno) e che si esauriscono nell'arco del periodo stesso;
  - influenza di fattori climatici, sociali, ecc., che si ripetono nello stesso periodo dell'anno.
  - periodi più comuni: annuale, settimanale, giornaliero (ma non solo). La lunghezza del ciclo dipende dalla frequenza di osservazione del fenomeno;
  - osservazioni distanti tra loro 12 mesi, o 4 trimestri, sono *collegate* tra loro in qualche modo, tuttavia
  - rimane ancora variazione fra i valori corrispondenti allo stesso periodo (stesso mese, stesso trimestre, . . .)
- **Altri effetti deterministici** dipendenti dal tempo.

Modelli di combinazione delle componenti

- **Modello additivo**: $Y_t = T_t + C_t + S_t + \epsilon_t$
  - componenti indipendenti
  - stessa unità di misura
- **Modello moltiplicativo**: $Y_t = T_t \cdot C_t \cdot S_t \cdot \epsilon_t$
  - T_t espressa nella stessa unità di misura di Y_t, le altre componenti assumono la forma di numeri indici
  - dipendenza tra le componenti
  - l'espressione diventa linearizzabile mediante trasformata logaritmica

$$\log Y_t = \log T_t + \log C_t + LogS_t + \log \epsilon_t$$

- **Modello misto**: vari tipi, p.e.: $Y_t = T_t \cdot C_t \cdot S_t + \epsilon_t$

Decomposizione di una serie storica - modello additivo

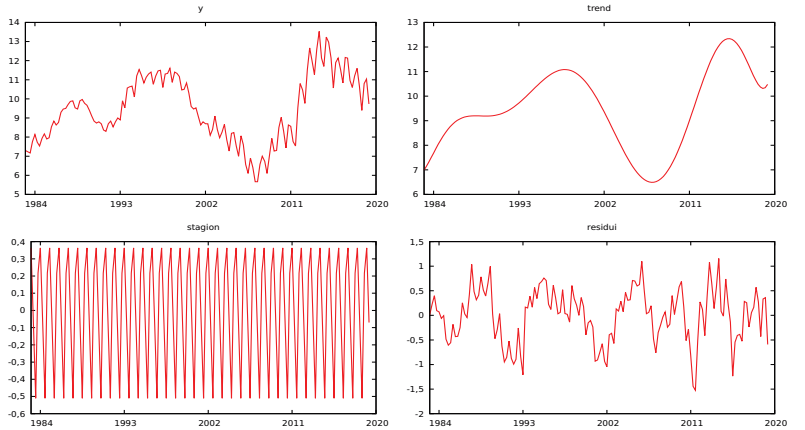


Figure: Tasso di disoccupazione armonizzato, periodo 1983:1 - 2019:2. Fonte FRED