

Timeseries

studio delle sequenze temporali

- [Introduzione](#)
- [Dimestificazione delle serie temporali \[1\]](#)
- [Configurazione enviroment \[2\]](#)

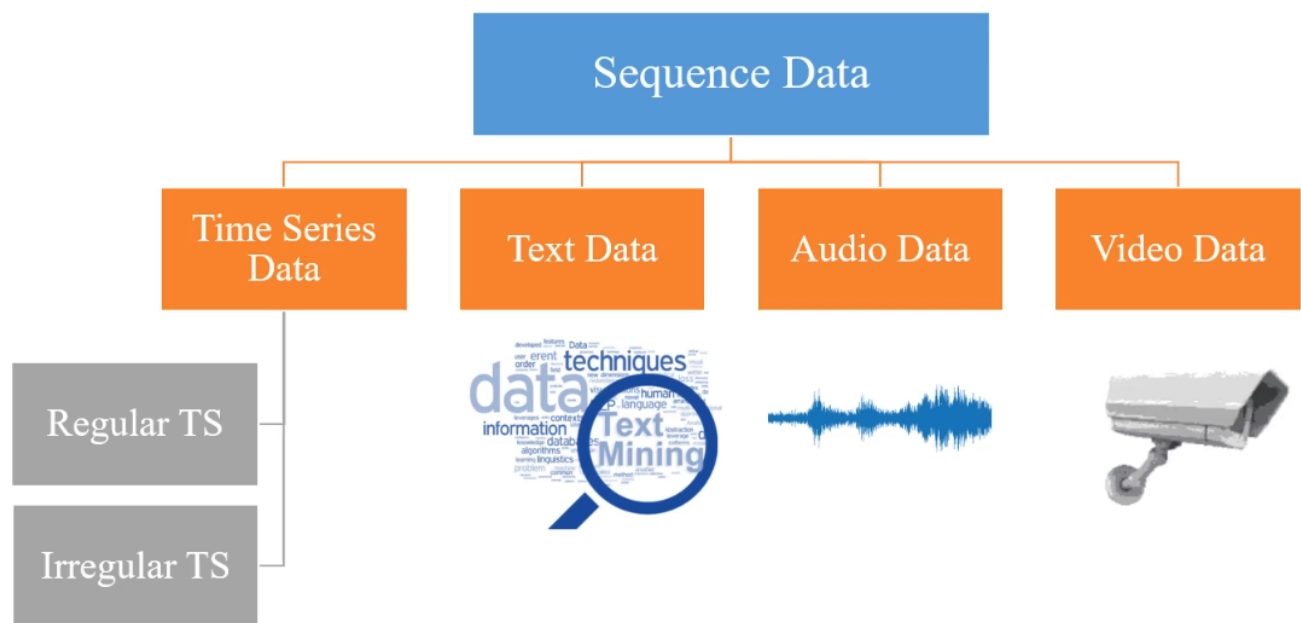
Introduzione

Dimestificazione delle serie temporali [1]

Tipologia di dati in sequenza

Dati in sequenza temporale (TS) significano dati che presentano uno specifico ordine o una sequenza nella quale vengono presentati, es:

Lo studio che verte in particolare sulle sequenze temporali che si suddividono in 1) serie TS regolari e 2) serie TS irregolari. Quelle regolari sono registrate con un timing regolare (es. ogni secondo) mentre quelle irregolari - come dice la parola - vengono registrate quando si verificano, quindi non necessariamente con cadenza regolare.



TS Tasks

I TS possono essere utilizzati per varie tipologie di task, es:

1) previsioni, che possono essere qualitative o quantitative

- Le previsioni qualitative vengono utilizzate quando non si hanno dati storici perchè per es. viene lanciato un nuovo prodotto
- Le previsioni quantitative invece fanno affidamento su dati storici, l'assunzione è che ciò che è avvenuto in passato dovrebbe ripetersi almeno in parte..

2) classificazioni

Le classificazioni invece vengono utilizzate quando vogliamo assegnare delle "label" alle sequenze temporali. Un esempio potrebbe essere quello dello smartwatch che, sulla base della frequenza cardiaca classifica il tuo stato che potrebbe essere, in corsa, camminata, oppure a riposo. questa modalità è anche detta "*supervised learning*")

3) clusterizzazioni

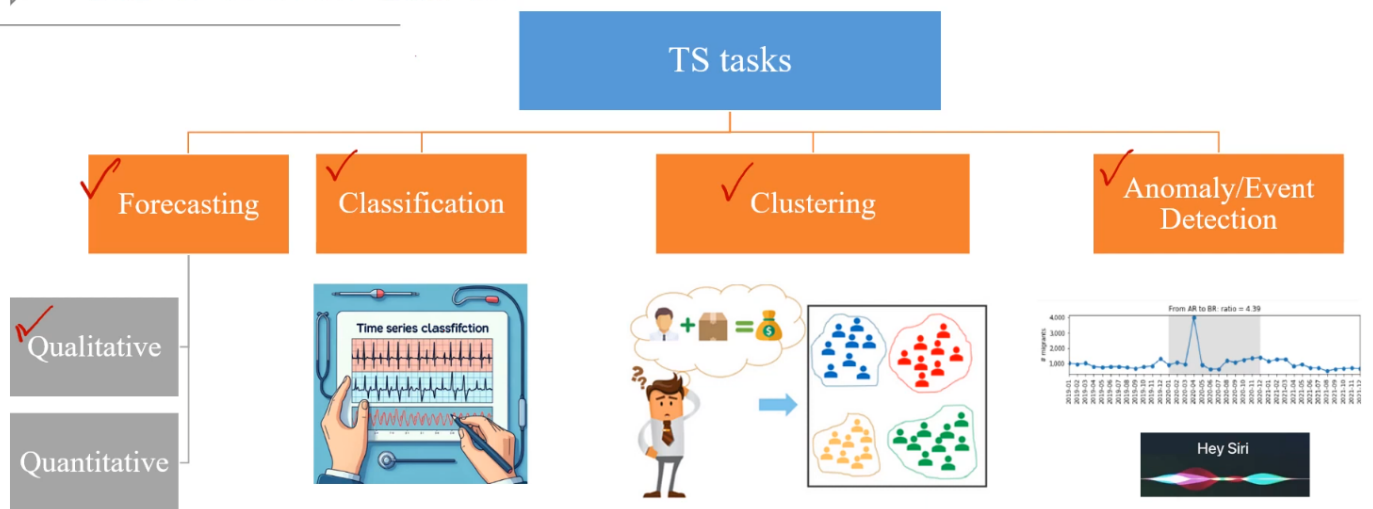
Le clusterizzazioni funzionano come le classificazioni ma senza le label, è il sistema che automaticamente raggruppa i dati sulla base di caratteristiche che identifica autonomamente. (questa modalità è anche detta "*un-supervised learning*")

Un utilizzato nella classificazione dei clienti o per capire eventuali trend.

4) rilevazione di anomalie / eventi

Viene utilizzato per identificare dei comportamenti fuori dal trend o inaspettati

➔ Time series Tasks



Scomposizione di una serie temporale

Un serie temporale può essere scomposta in:

- 1) dati della serie
- 2) tendenze
- 3) stagionalità

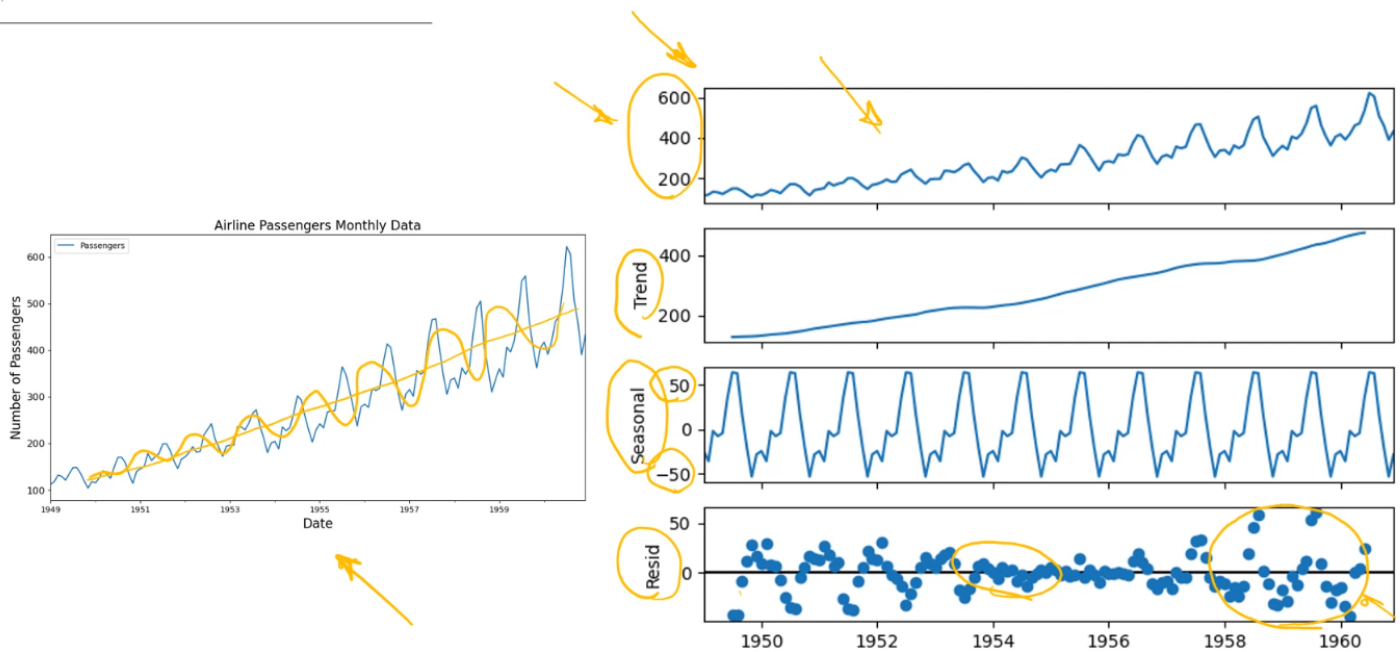
4) dati residui

Per esempio nel grafico si vedono il numero di passeggeri nell'arco degli anni. Nel primo grafico a dx si vedono tutti dati "**grezzi**" in una scala da 0 a 600. Nel **trend** invece, gli stessi dati vengono "puliti" per mostrare un andamento con una scala diversa da quella dei dati grezzi. Invece nel terzo grafico a dx gli stessi dati grezzi vengono filtrati per **stagionalità**, anche qui notare come la scala cambia.

Nell'ultimo grafico vengono invece evidenziati i dati "**residui**" in pratica si tratta di "rumore" che in qualche modo vanno filtrati per rendere l'analisi più precisa.

1. Timeseries I

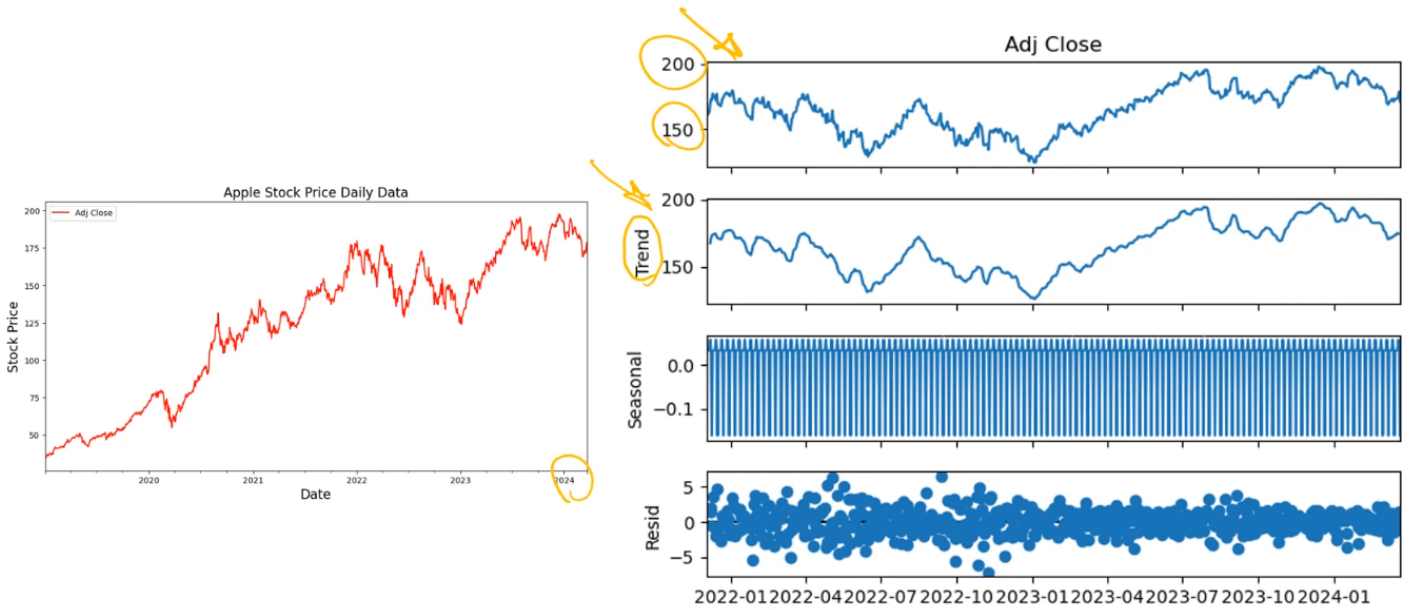
TS Decomposition: Trend, Seasonality, Residual



Altro esempio con il prezzo di Apple.

In questo caso si può notare come in particolare nella "stagionalità" non ha una particolare incidenza nell'arco temporale, quindi rende la predizione del dato quasi impossibile. (almeno basandosi su questa tipologia di dato)

TS Decomposition: Trend, Seasonality, Residual

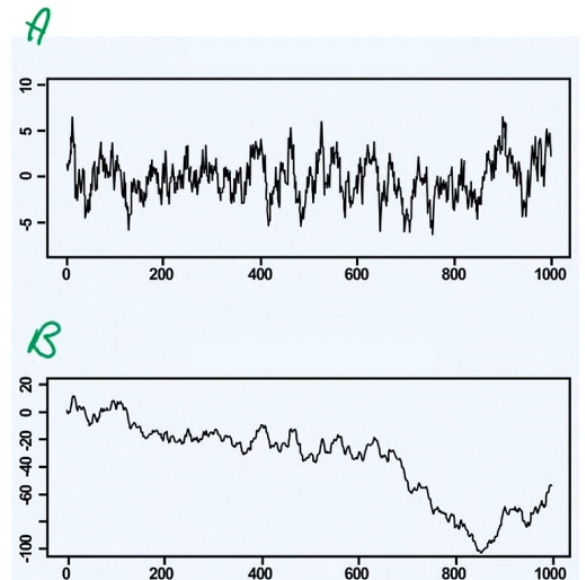


Dati stazionari vs non stazionari

Stationarity

- ✓ Stationary vs Non-Stationary Data. What makes a data set **Stationary**?
 - In a stationary timeseries, the statistical properties do **not depend on the time**
 - **Predictability (long-horizon)**: Stationary time series are easier to forecast because you can assume that future statistical properties will not change.
 - **This doesn't mean we cannot forecast non-stationary data!**

1. Timeseries Basics



I dati stazionari si identificano come dati che non dipendano dal tempo e sono più stabili nel tempo. (e quindi più facili da prevedere)

I dati non stazionari invece variano molto nel tempo e non una ciclicità a differenza dei dati stazionari

Configurazione enviroment

[2]