

# ANALISI DEI DATI MULTIVARIATA: ANALISI DELLE COMPONENTI PRINCIPALI, FATTORIALE E CLUSTER

## OBIETTIVO DEL CORSO

Grazie alla significativa quantità di dati ormai disponibile in ogni settore commerciale, industriale e sanitario, le tecniche di analisi statistica multivariata ricoprono oggi più che mai un ruolo fondamentale per l'estrazione di utili informazioni dai dati stessi. L'obiettivo del corso pertanto, è di illustrare le principali metodologie per l'analisi dei dati multivariata (analisi dei cluster, analisi delle componenti principali, analisi fattoriale) attraverso una serie di esempi e casi applicati. Dopo aver mostrato nella prima sezione come si possono sintetizzare le informazioni contenute in una banca dati multivariata tramite una rappresentazione grafica dei dati, sull'individuazione di *outlier* e sulla gestione dei dati mancanti, il corso si concentra su: l'analisi delle componenti principali, l'analisi fattoriale e l'analisi dei cluster e multidimensional scaling.

In linea con la nostra filosofia di formazione, le sessioni teoriche sono affiancate da illustrazioni pratiche ed esempi provenienti da diverse discipline sia sociali che biomediche, in cui il docente chiarisce le limitazioni e i punti di forza di ogni metodologia, nonché i criteri per la scelta e l'implementazione dello strumento di analisi statistica più appropriato per il problema oggetto di studio. Alla fine del corso, i partecipanti saranno in grado, con l'aiuto delle routine di Stata sviluppato durante le sessioni, di implementare autonomamente le metodologie utilizzate durante il corso nel proprio specifico contesto di ricerca.

## DESTINATARI

Il corso è di particolare interesse per ricercatori e professionisti che si occupano di economia, finanza, sanità

pubblica e di scienze sociali, e che desiderano acquisire gli strumenti necessari per condurre autonomamente condurre ricerche empiriche utilizzando dati multivariati.

## REQUISITI RICHIESTI

Conoscenze di base di Statistica e del Software Stata.

## PROGRAMMA

### SESSIONE I: I DATI MULTIVARIATI: PRIMI INDICATORI DI SINTESI E RAPPRESENTAZIONI GRAFICHE

1. Tipi di variabili e il problema dei dati mancanti
2. Covarianze, correlazioni e misure di distanza
3. La distribuzione normale multivariata
4. Grafici per la visualizzazione di dati multivariati

### SESSIONE II: ANALISI DELLE COMPONENTI PRINCIPALI E ANALISI FATTORIALE

1. Analisi delle componenti principali
  - Introduzione
  - Calcolo delle componenti principali
  - Calcolo degli *scores* delle componenti principali
2. Analisi dei fattori
  - Introduzione
  - Stima dei fattori
  - Scelta del numero di fattori
  - Rotazione dei fattori

### SESSIONE III: ANALISI DEI CLUSTER

1. Introduzione agli algoritmi di *clustering*
2. Principali metodi agglomerativi di clustering
  - Il dendrogramma
  - *Single linkage*
  - *Complete linkage*
  - *Average linkage*
  - Metodo di *Ward*
3. Principali metodi divisivi di clustering
  - *K-means*
4. Profilazione dei cluster

## SESSIONE IV: ALTRI METODI DI ANALISI DEI DATI MULTIVARIATI

1. Estensioni degli approcci presentati
2. Altre tecniche di analisi multivariata (analisi delle corrispondenze, scaling multidimensionale, etc.)

## TESTI UTILI

- Anderson, T. W. (2003), *An Introduction to Multivariate Statistical Analysis*, 3rd edition, Wiley.
- Everitt, B. S., Landau, S., Leese, M. and Stahl, D. (2011), *Cluster Analysis*, 5th edition, Wiley.
- Flury, B. (1997), *A First Course in Multivariate Statistics*, Springer
- Hastie, T., Tibshirani, R. and Friedman, J. (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edition, Springer
- Izenman, A. J. (2008), *Modern Multivariate Statistical Techniques*, Springer.
- Johnson, R. A. and Wichern, D. W. (2007), *Applied Multivariate Statistical Analysis*, 6th edition, Pearson/Prentice Hall.
- Mardia, K. V., Kent, J. T. and Bibby J. M. (1980), *Multivariate Analysis*, Academic Press.
- Timm, N. H. (2002), *Applied Multivariate Analysis*, Springer

## DATA E ISCRIZIONE

Il corso è in programma il 25-26 Giugno a Milano.

La partecipazione al corso è soggetta al pagamento della seguente quota di iscrizione:

Studente\*: € 530.00

Assegnista / Specializzando: € 700.00

Università: € 880.00

Commerciale: € 1095.00

\*Per usufruire dello status "studente" è necessario presentare copia del libretto universitario o un certificato di iscrizione (in carta semplice) all'Università ed essere *studenti a tempo pieno*. Studenti lavoratori dovranno considerare la tariffa Assegnisti / Specializzandi.

I prezzi si intendono IVA 22% esclusa. L'aliquota IVA non sarà applicata per Enti Pubblici soggetti ad esenzione a norma dell'art. 14 c. 10 della L. 537/93 per la partecipazione a corsi di formazione dei propri dipendenti.

La quota di iscrizione include il pranzo, il materiale didattico e una licenza temporanea del software Stata (si consiglia di venire muniti del proprio computer o di chiedere informazioni alla segreteria per l'eventuale noleggio, al momento dell'iscrizione). Dà inoltre diritto ad uno sconto sull'acquisto di una nuova licenza per singolo utente del Software Stata (ad esclusione della versione per Studenti) e sull'acquisto di testi in catalogo Stata Press.

L'iscrizione al corso dovrà avvenire tramite lo specifico modulo di registrazione e pervenire a TStat S.r.l. entro il 10 Giugno 2019. Lo svolgimento è condizionato dal raggiungimento di un numero minimo di 5 partecipanti ed un numero massimo di 8.

### PER ULTERIORI INFORMAZIONI RIVOLGERSI A:

**Monica Gianni**

TStat S.r.l. | Via Rettangolo, 12-14  
I-67039 Sulmona (AQ)  
T. +39 0864 210101

TStat Training | Kleebergstraße, 8  
D-60322 Frankfurt am Main

[www.tstat.it](http://www.tstat.it) | [www.tstattraining.eu](http://www.tstattraining.eu) | [formazione@tstat.it](mailto:formazione@tstat.it)



The logo for TStat, featuring the word 'TStat' in a stylized, handwritten font. The 'T' is red and the 'Stat' is black. The logo is set against a large, light gray circular background.