



---

# TEORIE, MODELLI E TECNICHE INFORMATICHE E DI ANALISI DEI DATI

6 CFU - 2° Semester

## Teaching Staff

**CESARE GAROFALO** - Module Monovariate and multivariate analysis - SPS/07 - 3 CFU

**Email:** cesaregarofalo@yahoo.com

**Office:** Da concordare

**Phone:** 339 2984739

**Office Hours:** Per appuntamento

**GIOVANNI GIUFFRIDA** - Module From data to information - INF/01 - 3 CFU

**Email:** ggiuffrida@dmi.unict.it

**Office:** Palazzo Reburdone, Viale Vittorio Emanuele 8 I Piano

**Phone:** 095 70305265

**Office Hours:** Mercoledì 10-13

---

## DETAILED COURSE CONTENT

### ▪ Monovariate and multivariate analysis

The course will focus on the study of univariate, bivariate and multivariate analysis using the R language, an open source environment for data management, statistical analysis, graphing and, more generally, for the use of a variety of formal methods (Networks Analysis, Time Series Analysis, Differential Equations, Machine Learning, Multivariate Statistics, etc.).

The course covers the following:

- 1) basic mathematical notions and logical propedeutics to computer programming;
- 2) operations on vectors, matrices, factors, lists, tables, data frames, using the R language;
- 3) read and write operations on external files using the R language;
- 4) graphic representations of the data using the R language;
- 5) programming with R: definitions of new functions, control constructs, conditional constructs and iterative constructs (if, ifelse, for, while, break, repeat, next);
- 6) univariate and bivariate descriptive statistics using the R language;
- 7) linear correlation and regression using the R language;

8) main component analysis using the R language;

9) cluster analysis using the R language;

10) network analysis using the R language;

▪ **From data to information**

We start discussing about the data and knowledge and differences between those. We then move on on the relational techniques to manage large amount of data. We discuss in quite details about the data management systems and the transactions which guarantee data consistency. We also discuss about relational algebra which is the founding pillar for the information retrieval languages and in particular is the basis of all SQL based language widely used in today's data base management systems.

---

## TEXTBOOK INFORMATION

▪ **Monovariate and multivariate analysis**

Lecturer's assignments

▪ **From data to information**

- Optional: Introduction to Computational Social Science, Principle and Applications. Claudio Cioffi-Revilla (In inglese)

- Optional: Big data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà. Viktor Mayer-Schönberger, Kenneth N. Cukier e R. Merlini

---