

Elaborazione dati e visualizzazione grafica dell'informazione in campo ambientale e sanitario

*Tecniche statistiche, geostatistiche e chemiometriche
per la rappresentazione grafica dell'informazione contenuta in grandi data set
ai fini del controllo ambientale e sanitario*

MODULO 1/3

Destinatari: studenti del I anno del CdL in *Tecniche della Prevenzione
nell'ambiente e nei luoghi di lavoro* dell'Università di Trieste

Durata: 12ore; A.A. 2021-2022 II semestre (marzo)

Tutor: **Andrea Mistaro** (ARPA FVG – Sede di Trieste)
email: andrea.mistaro@arpa.fvg.it

Programma indicativo:

1. Dalla progettazione del campionamento al *data analysis*: la “ruota” di Vandeginste.
2. Estrazione dati: il *data mining* e i problemi connessi; formattazione del database.
3. Chemiometria: estrazione dell'informazione rilevante dai dati:
 - *Exploratory Data Analysis* (EDA) vs *Confirmatory Statistics* (statistica descrittiva)
4. Reporting techniques: visualizzazione grafica di un *data set*;
 - Rappresentazione grafica e tabellare a confronto;
 - *Graphical perception* (cenni teorici); *Graphical integrity* vs *distorsion*;
 - Integrazione dei grafici nel testo; infografica (cenni).
5. Data Maps: cenni di geostatistica:
 - l'interpolazione spaziale dei dati e relative tecniche di *reporting* grafico.
 - Cenni ed esempi di modellistica previsionale.
6. Analisi univariata: cenni di statistica descrittiva e relative tecniche di *reporting* grafico:
 - La distribuzione di frequenza e le sue tecniche di rappresentazione (istogrammi; curve di distribuzione; *Box & Whisker plot*);
 - Serie temporali (introduzione e peculiarità).

Programma delle esercitazioni al PC

lo stesso giorno della rispettiva parte teorica, gli studenti saranno invitati a svolgere delle esercitazioni al PC (in coppie o terne di studenti) su alcuni degli argomenti trattati, con la supervisione del docente

- 6a. Esercitazioni di analisi statistica descrittiva classica (univariata) e relativo *reporting* grafico:
 - distribuzioni di frequenza: istogrammi; curve di distribuzione; *Box & Whisker plot*; *normal probability plot*.