UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Statistica per la Sicurezza e Analisi dei Dati

2627-1-F8805N030

Obiettivi formativi

Il corso ha un taglio principalmente pratico-applicativo e si propone di fornire agli studenti e alle studentesse conoscenze e competenze di base per l'analisi statistica dei dati sulla sicurezza, dall'acquisizione dei dati fino all'analisi predittiva.

Conoscenza e comprensione
Al termine del corso, gli studenti e le studentesse avranno acquisito:

- la conoscenza dei principi fondamentali della statistica esplorativa, inferenziale e predittiva applicata all'analisi dei dati sulla sicurezza;
- la comprensione delle principali metodologie per la raccolta, l'organizzazione e l'analisi di dati provenienti da fonti ufficiali;
- la comprensione delle assunzioni, delle potenzialità e dei limiti degli strumenti statistici utilizzati.

Capacità di applicare conoscenza e comprensione
Al termine del corso, gli studenti e le studentesse saranno in grado di:

- acquisire, organizzare e gestire dataset provenienti da fonti ufficiali;
- utilizzare il software R e l'ambiente RStudio per la manipolazione e l'analisi dei dati;
- descrivere e sintetizzare i dati mediante indicatori statistici e rappresentazioni grafiche;
- applicare metodi inferenziali allo studio di fenomeni legati alla sicurezza;
- costruire e interpretare modelli di regressione e machine learning;
- utilizzare procedure appropriate per l'analisi di dati reali.

Autonomia di giudizio
Al termine del corso, gli studenti e le studentesse saranno capaci di:

- valutare criticamente la qualità e l'affidabilità dei dati sulla sicurezza;
- selezionare gli strumenti statistici più appropriati in funzione del problema;

- interpretare criticamente i risultati ottenuti.

Abilità comunicative
Al termine del corso, gli studenti e le studentesse saranno in grado di:

- presentare e comunicare i risultati delle analisi statistiche attraverso tabelle, grafici e indicatori sintetici;
- descrivere in modo chiaro le metodologie adottate e i risultati ottenuti utilizzando una terminologia statistica appropriata.

Capacità di apprendimento
Al termine del corso, gli studenti e le studentesse svilupperanno la capacità di:

- reperire, valutare e utilizzare criticamente nuovi dati relativi a fenomeni legati alla sicurezza;
- affrontare in autonomia nuovi problemi analitici applicando le conoscenze e le competenze acquisite durante il corso;
- aggiornare e sviluppare le proprie competenze nell'ambito dell'analisi dei dati.

Contenuti sintetici

Il corso si propone di trasmettere conoscenze e competenze fondamentali per l'analisi statistica per la sicurezza. L'approccio combina aspetti teorici e applicativi attraverso l'analisi di dati reali sulla sicurezza, con il supporto del software statistico R e dell'ambiente RStudio.
Il corso è organizzato in quattro macro-argomenti.
Il primo argomento introduce i concetti fondamentali della statistica applicata alla sicurezza, con particolare attenzione alle fonti dei dati (come ISTAT e Ministero della Giustizia), alla qualità dell'informazione statistica e alla struttura dei dati. Viene inoltre introdotto l'utilizzo di R per l'importazione e la gestione dei dati.
Il secondo argomento è dedicato all'analisi esplorativa dei dati e introduce concetti chiave come distribuzioni di frequenza, misure di sintesi, rappresentazioni grafiche e analisi delle relazioni tra variabili.
Il terzo argomento riguarda l'inferenza statistica e i principali strumenti di analisi probabilistica e inferenziale. Sono trattati variabili casuali, distribuzioni di probabilità, stima puntuale e intervallare, test d'ipotesi, confronto tra gruppi e analisi della varianza.
Il quarto argomento è dedicato all'analisi predittiva, con introduzione alla regressione lineare e logistica e ai principali metodi di machine learning.
Tutte le tecniche e gli strumenti presentati durante il corso sono implementati in R e RStudio, soffermandosi in particolare sull'interpretazione critica dei risultati ottenuti.

Programma esteso

Argomento 1 Cenni introduttivi alla statistica per la sicurezza e al software R

- Introduzione alla statistica per la sicurezza: obiettivi e strumenti
- Sicurezza oggettiva e soggettiva, fonti ufficiali di dati
- Popolazioni, campioni e unità statistiche
- Qualità del dato: errori e affidabilità delle misure
- Introduzione al software statistico R

Argomento 2 Analisi esplorativa

- Distribuzioni di frequenza

- Misure di tendenza centrale e dispersione
- Tabelle di contingenza
- Relazioni tra coppie di variabili
- Rappresentazioni grafiche

Argomento 3 Inferenza statistica

- Variabili casuali discrete e continue
- Distribuzioni di probabilità
- Stima puntuale e intervallare
- Test statistici univariati
- Test statistici a due campioni
- Confronto tra gruppi e analisi della varianza

Argomento 4 Analisi predittiva

- Regressione lineare semplice e multipla
- Regressione logistica
- Introduzione al machine learning

Prerequisiti

Competenze base di matematica e calcolo.

Metodi didattici

Il corso si compone di 56 ore svolte in presenza, prevalentemente in didattica erogativa (DE). Le lezioni combinano la presentazione dei concetti teorici mediante slide con pratiche svolte utilizzando R e RStudio.
Durante il corso sono previsti momenti di didattica interattiva (DI) finalizzati a favorire la partecipazione attiva degli studenti e delle studentesse attraverso l'applicazione in tempo reale delle procedure illustrate a lezione utilizzando i propri dispositivi personali. Le attività di didattica interattiva sono integrate nelle lezioni e consistono principalmente nello svolgimento guidato di esercitazioni in R, nella discussione dei risultati e nella risoluzione di problemi applicativi.
Gli studenti e le studentesse potranno inoltre replicare autonomamente le analisi presentate durante il corso utilizzando i codici messi a disposizione dal docente.
La piattaforma e-learning del corso (https://elearning.unimib.it) è regolarmente aggiornata con il materiale didattico, incluse le slide delle lezioni, i codici R, i dataset utilizzati e ulteriori risorse di approfondimento.

Modalità di verifica dell'apprendimento

La valutazione dell'apprendimento avviene mediante una prova scritta finalizzata ad accertare il raggiungimento degli obiettivi formativi.
La prova comprende domande a scelta multipla e domande aperte nelle quali viene richiesto di riportare i codici sviluppati in R utilizzati per l'individuazione della risposta corretta e di fornire un'interpretazione critica dei risultati ottenuti.
Ulteriori informazioni verranno fornite durante il corso e rese disponibili sulla pagina e-learning.

Criteri di valutazione:

- Completezza e correttezza delle conoscenze acquisite;
- Capacità di utilizzare R/RStudio per l’analisi di dati sulla sicurezza;
- Capacità di interpretare criticamente i risultati delle analisi statistiche;
- Capacità di applicare gli strumenti statistici appropriati alla soluzione di problemi specifici;
- Capacità di sintesi;
- Uso appropriato del linguaggio e della terminologia statistica.

Testi di riferimento

Slide delle lezioni, dispense e altro materiale didattico sarà reso disponibile sulla pagina e-learning e integrato durante il corso.

Testi di utile consultazione:

- F. Mecatti “Statistica di Base. Come, Quando e Perchè”. McGraw Hill, III Edizione 2022
- A. Wooditch, N.J. Johnson, R. Solymosi, J. Medina Ariza, S. Langton “A beginner's guide to statistics for criminology and criminal justice using R”. Springer International Publishing, 2021.
- D. Weisburd, D.B. Wilson, A. Wooditch, C. Britt “Advanced statistics in criminology and criminal justice”. Springer, 2021.

Sustainable Development Goals

RIDURRE LE DISUGUAGLIANZE | CITTÀ E COMUNITÀ SOSTENIBILI | PACE, GIUSTIZIA E ISTITUZIONI SOLIDE