

Progetto di gruppo SLPD

Direzione e Consulenza d'Impresa – a.a. 2024/25

Linee guida per l'elaborato

Informazioni generali

Obiettivi del progetto. Gli obiettivi del progetto di gruppo includono: (i) la conduzione di un'analisi statistica esplorativa (EDA) utilizzando software statistici; (ii) modellare i dati per rispondere a una domanda di ricerca; (iii) interpretare e comunicare i risultati con rigore scientifico.

Scelta dell'argomento. Ciascun gruppo dovrà analizzare **un solo dataset**, selezionabile dalla pagina Moodle del corso. Uno schema riassuntivo degli argomenti è riportato in Tabella 1. La scelta dell'argomento **non è esclusiva**: più gruppi potranno lavorare sullo stesso dataset. È possibile utilizzare il dataset in modo integrale o parziale, in base alla domanda di ricerca e/o alle tecniche di analisi scelte.

Ricevimento. Per garantire pari opportunità a tutti i gruppi, ciascun gruppo potrà richiedere al massimo un incontro settimanale con il docente, da concordare preventivamente.

Analisi dei dati

Prima di condurre l'analisi, è fondamentale valutare le caratteristiche del dataset: selezione delle variabili, distribuzione delle variabili, trattamento di eventuali valori mancanti e/o outlier, ecc.

Ciascun gruppo dovrà implementare **almeno tre metodi a scelta** tra quelli trattati nella seconda parte del corso, scegliendo al massimo una tecnica per argomento:

- 1) Modelli predittivi: regressione lineare semplice; regressione lineare multipla.
- 2) Modelli di classificazione: regressione logistica; regressione logistica multinomiale.
- 3) Riduzione della dimensionalità: analisi fattoriale; Analisi delle Componenti Principali (PCA).
- 4) Tecniche di clustering: clustering gerarchico (hierarchical clustering); *k-means* clustering.

A titolo esemplificativo, una possibile combinazione sarebbe: (2) regressione logistica; (3) analisi fattoriale; (4) *k-means* clustering. Ad integrazione delle tecniche sopraelencate, ciascun gruppo è libero di arricchire l'analisi statistica utilizzando le tecniche presentate nella prima parte del corso. Questo punto è facoltativo e a discrezione del gruppo.

Software. Il software di riferimento del corso è SPSS, per il quale sono disponibili tutorial specifici per ciascun argomento nelle slides su Moodle. È possibile usufruire di una prova gratuita di 30 giorni. In alternativa, è consentito l'utilizzo di software alternativi open-source come PSPP, Jamovi, Gretl, ecc.. Sono invece esclusi i modelli di intelligenza artificiale (es. ChatGPT).

Struttura dell'elaborato

L'elaborato dovrà articolarsi nei seguenti punti:

- 1) **Introduzione.** Breve introduzione all'argomento scelto: informazioni di contesto sull'industria/tematica di riferimento e motivazione della scelta.

- 2) **Domanda/e di ricerca.** Definizione di una o più domande di ricerca.
- 3) **Dati utilizzati.** Descrizione del campione: origine dei dati, tipologia delle variabili, eventuali operazioni di pulizia del dato, statistiche descrittive in formato tabellare e grafico.
- 4) **Metodologie statistiche.** Presentazione delle tecniche statistiche impiegate, con illustrazione delle finalità e breve descrizione dei metodi anche dal punto di vista statistico/matematico.
- 5) **Risultati empirici.** Presentazione dei risultati ottenuti mediante tabelle e grafici.
- 6) **Discussione e conclusioni.** Considerazioni sui risultati: implicazioni teoriche, tecniche e/o manageriali.

Ogni progetto dovrà avere un titolo originale, in base alla tematica scelta.

Per quanto riguarda la lunghezza dell'elaborato, non c'è un numero minimo o massimo di pagine/caratteri. Ogni elaborato sarà sottoposto a controlli antiplagio e anti-AI.

Consegna e presentazione dell'elaborato

L'elaborato dovrà essere consegnato al docente tramite email entro la mezzanotte del giorno precedente alla prova finale. Non saranno ammessi ritardi o integrazioni future.

Presentazione. È prevista una presentazione supportata da slide della durata di 15 minuti, alla quale dovranno partecipare obbligatoriamente **tutti i membri** del gruppo. La presentazione dell'elaborato si svolgerà il giorno della prova finale.

Valutazione finale

La valutazione (in 30esimi) prenderà in considerazione il lavoro del gruppo nel suo complesso (elaborato e presentazione), nonché la capacità dei singoli membri di rispondere alle domande poste durante la presentazione. Saranno premiati il rigore nell'applicazione dei metodi statistici, la capacità di interpretare e commentare i risultati, e l'originalità dell'analisi. Il giudizio finale sarà **unico per tutti i membri del gruppo** e si sommerà all'esito individuale dei quiz svolti in aula.

Tabella 1. Argomenti.

Nome dataset	Tematica	Descrizione
EWCTS	Social	L'European Working Conditions Telephone Survey (EWCTS) è uno studio che indaga la condizione di 70,000+ lavoratori in 36 paesi Europei. Dati relativi al 2021 per i paesi EU12.
Housing	Economic	Dati su 1,400+ immobili derivanti da un'indagine relativa al mercato immobiliare di Melbourne (Australia).
ISSP	Social	L'International Social Survey Programme (ISSP) si concentra su questioni relative all'ambiente, al cambiamento climatico e alla protezione dell'ambiente. Dati relativi al 2020 per 4 paesi: Italia, Francia, Germania, Spagna.
Soccer	Sport	Dati sulla performance individuale (e di club) di 2,600+ calciatori in Europa per la stagione 2022/2023.
Spotify	Music	Caratteristiche principali di 3,000+ canzoni di Spotify.
Supply_chain	Economic	Dati di bilancio delle aziende della supply chain di un'azienda ceramica (distretto di Sassuolo). I dati fanno riferimento all'anno 2021.