



Dataset

Gruppo 1

Il gruppo di ricerca del Dipartimento di Scienze Sociali vuole esaminare come variabili caratterizzanti le transazioni di vendita influenzino i comportamenti dei clienti e le performance di vendita.

Si richiede di:

1. formulare ipotesi,
2. scegliere variabili da analizzare,
3. proporre analisi statistiche e interpretare potenziali risultati.

Esplicitare i possibili risultati a cui può giungere il gruppo di lavoro.

Gruppo 2

Il gruppo di ricerca del Dipartimento di Scienze Economiche vuole esaminare come variabili di mercato e caratteristiche delle azioni siano correlate tra loro.

Si richiede di analizzare le relazioni tra variabili chiave per comprendere pattern di mercato e:

- determinare se esistono correlazioni tra la volatilità delle azioni e il settore industriale;
- valutare la relazione tra età dell'azione sul mercato e performance di rendimento annuale.

Esplicitare i possibili risultati a cui può giungere il gruppo di lavoro.





Gruppo 3

Il gruppo di ricerca del ministero del lavoro vuole analizzare come caratteristiche demografiche e lavorative degli impiegati influenzino aspetti come performance e livello salariale.

Si richiede di analizzare le relazioni tra variabili chiave del personale per comprendere dinamiche organizzative e:

- determinare se esistono associazioni significative tra lo stato occupazionale dei dipendenti (es. attivo, inattivo) e il livello di performance;
- valutare la relazione tra l'età dei dipendenti e il livello salariale, distinguendo tra differenti fasce di età.

Esplicitare i possibili risultati a cui può giungere il gruppo di lavoro.

Gruppo 4

Il gruppo di ricerca del Dipartimento di Economia e Management sta conducendo un'indagine statistica per analizzare le dinamiche di vendita nel commercio elettronico e comprendere come le caratteristiche dei prodotti influenzino il comportamento dei consumatori sulla piattaforma Amazon.

Si richiede di analizzare le relazioni tra variabili chiave del commercio elettronico per comprendere pattern di vendita e comportamento dei clienti e:

- ANOVA**
- determinare se esistono correlazioni tra la **categoria del prodotto** e la valutazione media (**rating**) ricevuta dai clienti;
- Sperman**
- valutare la relazione tra sconti applicati sui prodotti (ad esempio **percentuale di sconto**) e numero di valutazioni (**rating_count**), come indicatore di engagement dell'utente.

Esplicitare i possibili risultati a cui può giungere il gruppo di lavoro.





Attività

Output atteso

- Realizzare una presentazione con Power BI e PowerPoint, dove mostrare frammenti di codice Python, la struttura delle tabelle MySQL, dettagli sul backend
- Pubblicare un report con Power BI

Primo step (Ardigò / Gulinelli)

- Recuperare i dataset necessari
- Importare i dataset in pandas e "ripulirli" (check dei dati mancanti, anomali, ecc.)
- Effettuare un'analisi esplorativa dei dati con determinazione degli indici statistici richiesti
- Effettuare la correlazione e regressione usando le opportune librerie
- Presentare le conclusioni a cui il gruppo è giunto e generare opportuni grafici

4-8 h per l'analisi statistica

2 gg per sviluppo Python

Secondo step (Gulinelli / Parisi)

- Memorizzare i dataset (originali e ripuliti) in un DB MySQL

4 h

Terzo step (Gulinelli)

- Realizzare un backend in flask per consentire la consultazione "al pubblico" dei dati via browser web attraverso RESTful API.

1 h

Quarto step (Ardigò)

- Importare i medesimi dati attraverso i seguenti due metodi
 - a) via API dal web service creato in Flask
 - b) dalle tabelle del database MySQLPoi usare nel report le tabelle prese soltanto dal MySQL e disattivare l'"enable load" dell'altra sorgente.
- Creare un report di presentazione
- Pubblicare il report in Power BI Service e mostrarlo con link pubblico
- (Opzionale) Installare e settare un Gateway Personale e dimostrare che, al variare del dato, dopo un refresh, anche il report cambia.

4-8 h

