



SELEZIONE PUBBLICA, PER TITOLI ED ESAMI, PER IL RECLUTAMENTO DI N. 1 UNITÀ DI TECNOLOGO DI PRIMO LIVELLO CON RAPPORTO DI LAVORO SUBORDINATO A TEMPO DETERMINATO DELLA DURATA DI 36 MESI, PRESSO L'UNIVERSITÀ DEGLI STUDI DI MILANO - DIPARTIMENTO DI ECONOMIA, MANAGEMENT E METODI QUANTITATIVI, PER L'ATTUAZIONE DEL PROGRAMMA DI RICERCA E INNOVAZIONE DAL TITOLO "MUSA - MULTILAYERED URBAN SUSTAINABILITY ACTION" (CUP G43C22001370007) NELL'AMBITO DEL PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) - CODICE 22289

La Commissione giudicatrice della selezione, nominata con Determina Direttoriale n. 7241 del 9.5.2023, composta da:

Prof.ssa Silvia Salini	Presidente
Prof. Giancarlo Manzi	Componente
Dott. Stefano Trancossi	Componente
Sig.ra Cristina Crippa	Segretaria

comunica le tracce relative alla prova scritta.

Traccia 1

1. Qual è il comando in R per creare una matrice 2x3 con i valori 1, 2, 3, 4, 5, 6?

- a) `matrix(1:6, nrow = 2, ncol = 3)`
- b) `matrix(1:6, nrow = 3, ncol = 2)`
- c) `matrix(1:6, nrow = 2)`
- d) `matrix(1:6, ncol = 3)`

2. Quale funzione in R viene utilizzata per calcolare il prodotto scalare tra due vettori?

- a) `crossprod()`
- b) `%*%`
- c) `%o%`
- d) `outer()`

3. Quale comando in Python viene utilizzato per importare la libreria Matplotlib?

- a) `import mpl`
- b) `import matplotlib.pyplot as plt`
- c) `import plt from matplotlib`
- d) `import matplotlib as mpl`

4. Quale libreria in Python viene utilizzata per l'elaborazione numerica efficiente di array multidimensionali?

- a) Matplotlib
- b) NumPy
- c) Pandas
- d) SciPy

5. Cos'è una lista in programmazione?

6. Cos'è una funzione in programmazione?



Esercizio 1

Supponiamo di avere un dataset che contiene informazioni sui voti degli studenti in tre esami (Matematica, Scienze e Storia). Il dataset è strutturato con le seguenti colonne:

- Studente: il nome dello studente
- Matematica: il voto ottenuto nell'esame di Matematica
- Scienze: il voto ottenuto nell'esame di Scienze
- Storia: il voto ottenuto nell'esame di Storia

Scrivi come svolgeresti le seguenti attività utilizzando il pacchetto che preferisci (R, Python, STATA, Matlab, ecc.):

1. Carica il dataset nel tuo ambiente
2. Visualizza le prime righe del dataset per verificarne la corretta lettura
3. Calcola la media dei voti di ciascun esame
4. Calcola la media dei voti per ogni studente
5. Calcola la varianza dei voti di ciascun esame
6. Crea un grafico a barre che mostri la media dei voti per ogni esame
7. Utilizzando l'ANOVA, verifica se ci sono differenze significative tra le medie dei voti degli esami.

Esercizio 2

Supponi di avere il seguente schema di database relazionale per una tabella "Clienti":

Clienti (id_cliente, nome, cognome, data_nascita, indirizzo, città, provincia, cap, email)

Si assuma che ci sia un'altra tabella "Ordini" contenente le informazioni sugli ordini effettuati dai clienti, con almeno i campi "id_cliente" e "importo".

Scrivi le seguenti query SQL:

1. Seleziona il nome e il cognome dei clienti con età superiore a 40 anni.
2. Seleziona il numero di clienti per ogni città, in ordine decrescente per numero di clienti.
3. Seleziona l'indirizzo email dei clienti che vivono nella città di Milano.
4. Seleziona il nome, il cognome e la data di nascita dei clienti che hanno effettuato almeno un ordine.
5. Seleziona il nome, il cognome e l'indirizzo dei clienti che hanno effettuato almeno un ordine di importo superiore a 1000 euro.



Traccia 2

1. Quale libreria in Python offre strumenti per l'apprendimento automatico (machine learning)?
 - a) Matplotlib
 - b) NumPy
 - c) Pandas
 - d) Scikit-learn
2. Quale metodo di Pandas viene utilizzato per caricare un file CSV in un DataFrame?
 - a) read_csv()
 - b) load_csv()
 - c) import_csv()
 - d) open_csv()
3. Quale pacchetto in R viene utilizzato per la manipolazione dei dati mancanti?
 - a) dplyr
 - b) ggplot2
 - c) tidyr
 - d) mice
4. Quale comando in R viene utilizzato per creare una nuova variabile in un data frame con dplyr?
 - a) mutate()
 - b) transform()
 - c) create()
 - d) new_variable()
5. Cosa significa "compilare" un programma?
6. Qual è la differenza tra una lista e un'array?

Esercizio 1

Supponiamo di avere un dataset che contiene informazioni sulle vendite mensili di tre prodotti (A, B e C) per un periodo di un anno. Il dataset è strutturato con le seguenti colonne:

- Mese: il mese di riferimento (gennaio, febbraio, marzo, ecc.)
- Prodotto: il nome del prodotto (A, B o C)
- Vendite: il numero di unità vendute nel mese

Scrivi come svolgeresti le seguenti attività utilizzando il pacchetto che preferisci (R, Python, STATA, Matlab, ecc.):

1. Carica il dataset nel tuo ambiente
2. Visualizza le prime righe del dataset per verificarne la corretta lettura.
3. Calcola la somma delle vendite totali per ogni prodotto.
4. Calcola la media delle vendite mensili per ogni prodotto
5. Calcola la varianza delle vendite mensili per ogni prodotto
6. Crea un grafico a linee che mostri le vendite mensili per ogni prodotto nel corso dell'anno



7. Utilizzando il test ANOVA, verifica se ci sono differenze significative tra le medie delle vendite mensili dei prodotti.

Esercizio 2

Supponi di avere il seguente schema di database relazionale per due tabelle: "Studenti" e "Corsi":

Studenti (id_studente, nome, cognome, data_nascita)

Corsi (id_corso, nome_corso, docente, crediti)

Si assume che ci sia una tabella "Iscrizioni" che tenga traccia delle iscrizioni degli studenti ai corsi, con almeno i campi "id_studente" e "id_corso".

Scrivi le seguenti query SQL:

1. Seleziona il nome e il cognome degli studenti iscritti al corso con ID 1.
2. Seleziona il nome e il cognome degli studenti che non sono iscritti a nessun corso.
3. Seleziona il nome del corso e il nome del docente per tutti i corsi con un numero di crediti superiore a 3.
4. Seleziona il nome del corso e il numero di studenti iscritti per tutti i corsi.
5. Seleziona il nome del corso e il nome del docente per tutti i corsi in cui sono iscritti almeno 3 studenti.



Traccia 3

1. Quale libreria in Python offre strumenti per l'elaborazione delle immagini e la visione artificiale?
 - a) Matplotlib
 - b) OpenCV
 - c) TensorFlow
 - d) Keras
2. Quale libreria in Python offre funzionalità per la creazione di grafici statistici avanzati?
 - a) Matplotlib
 - b) Seaborn
 - c) Plotly
 - d) Bokeh
3. Quale comando in R viene utilizzato per specificare la posizione di un elemento in un'applicazione Shiny?
 - a) locate()
 - b) position()
 - c) place()
 - d) render()
4. Quale funzione in R viene utilizzata per calcolare la trasposta di una matrice?
 - a) transpose()
 - b) t()
 - c) matrix_transpose()
 - d) matrix_t()
5. Qual è la differenza tra un ciclo "for" e un ciclo "while"?
6. Cosa significa "variabile" nel contesto della programmazione?



Esercizio 1

Supponiamo di avere un dataset che contiene informazioni sulle vendite giornaliere di un negozio in diversi mesi. Il dataset è strutturato con le seguenti colonne:

- Data: la data della vendita
- Vendite: il valore delle vendite giornaliere

Scrivi come svolgeresti le seguenti attività utilizzando il pacchetto che preferisci (R, Python, STATA, Matlab, ecc.):

1. Carica il dataset nel tuo ambiente
2. Visualizza le prime righe del dataset per verificarne la corretta lettura.
3. Converti la colonna "Data" in un oggetto di tipo data.
4. Calcola il numero totale di vendite registrate nel dataset.
5. Calcola il valore medio delle vendite giornaliere.
6. Trova la data con il valore di vendita massimo e la relativa vendita.
7. Trova la data con il valore di vendita minimo e la relativa vendita.
8. Crea un grafico a linee che mostri l'andamento delle vendite giornaliere nel corso del tempo.

Esercizio 2

Supponi di avere il seguente schema di database relazionale per due tabelle: "Prodotti" e "Vendite":

Prodotti (id_prodotto, nome, prezzo)

Vendite (id_vendita, id_prodotto, data, quantita)

Si assume che la tabella "Vendite" abbia una chiave esterna "id_prodotto" che fa riferimento alla tabella "Prodotti".

Scrivi le seguenti query SQL:

1. Seleziona il nome e il prezzo di tutti i prodotti.
2. Seleziona il numero totale di vendite registrate.
3. Seleziona il nome del prodotto e la quantità venduta per tutte le vendite effettuate nel mese di gennaio.
4. Seleziona il nome del prodotto e il ricavo totale (prezzo x quantità) per ogni vendita.
5. Seleziona il nome del prodotto più venduto (con la quantità massima) e la relativa quantità.



UNIVERSITÀ DEGLI STUDI DI MILANO

Milano, 17 maggio 2023

La Commissione

Prof.ssa Silvia Salini - Presidente

Prof. Giancarlo Manzi - Componente

Dott. Stefano Trancossi - Componente

Sig.ra Cristina Crippa - Segretaria