

# DOMANDE SELEZIONE PUBBLICA PER L'ASSUNZIONE A TEMPO INDETERMINATO DI 100 UNITÀ PER LA TERZA AREA FUNZIONALE, FASCIA RETRIBUTIVA F1, PROFILO PROFESSIONALE FUNZIONARIO INFORMATICO - DS50 - PROVA SCRITTA - BUSTA 3

---

1) Nel linguaggio "R", un vettore (vector) può contenere elementi ...

- A solo di tipo numerico.
  - B di tipo e classe anche diversi per ogni elemento.
  - C che appartengono allo stesso tipo.
- 

2) Nel linguaggio di programmazione "R" un data frame viene utilizzato per memorizzare...

- A dati in formato tabellare, con colonne di qualsiasi tipo.
  - B dati non strutturati.
  - C dati in formato tabellare, con colonne di tipo numerico.
- 

3) Come si chiama il "package" che contiene le funzioni fondamentali del linguaggio "R"?

- A Start-Up.
  - B Base.
  - C Core.
- 

4) Quale dei seguenti frammenti di codice è sintatticamente corretto in Python?

- A 

```
for x in vettore {  
    somma = somma + x  
}
```
  - B 

```
for x in vettore:  
    somma = somma + x
```
  - C 

```
for x in vettore:  
begin  
    somma = somma + x  
end
```
- 

5) In Python, la numerazione delle liste (lists) parte da ...

- A 1
  - B poiché Python usa la notazione polacca inversa, la numerazione delle liste parte dalla dimensione della lista, così come specificata in fase di definizione e si muove a ritroso.
  - C 0
- 

6) In quale linguaggio di programmazione la seguente espressione è sintatticamente corretta?  
`a=1:100`

- A Sia in R che in Python.
  - B Solo in R.
  - C Solo in Python.
- 

7) Quale tra i seguenti è un tipico approccio nella progettazione di un database?

- A Right – Left.
  - B Top – Down.
  - C Left – Right .
- 

8) In SQL, l'operatore UNION tra due result set A e B restituisce...

- A l'insieme delle righe di A e B che hanno una chiave in comune.
  - B l'insieme delle colonne di A e B.
  - C l'insieme delle righe di A e B, se A e B hanno le stesse colonne.
- 

9) **La seguente espressione SQL**

```
Select sum(volume_affari)
from IVA
group by codice_fiscale
```

restituisce...

- A un result set con numero di righe sempre uguale a quello della tabella IVA.
  - B un result set con numero di righe pari al numero di valori distinti della colonna codice\_fiscale.
  - C un result set con numero di righe pari al numero di codici fiscali aventi un valore distinto di volume d'affari.
- 

10) **Date due tabelle A e B di un database relazionale, l'espressione**

```
SELECT * FROM
A LEFT JOIN B ON A.KEY=B.KEY
```

restituisce...

- A Solo le righe di A e B che hanno il valore della colonna KEY in comune.
  - B Solo le colonne di A.
  - C Tutte le righe di A.
- 

11) **Quale delle seguenti espressioni in SQL viene usata per inserire una nuova riga (con un solo attributo di tipo carattere) in una tabella T?**

- A INSERT ('ABC') INTO T
  - B LOAD T WITH ('ABC')
  - C INSERT INTO T VALUES ('ABC')
- 

12) **In SQL, qual è la cardinalità della tabella creata con la seguente istruzione?**

```
CREATE TABLE Indirizzi (
    Codice integer,
    Cognome char(60),
    Nome char(60),
    Telefono varchar(60)
)
```

- A Non si può dire, la cardinalità dipende dal numero di righe che verranno inserite.
  - B 4
  - C 184
- 

13) **In un database relazionale, una relazione "multi-a-molti" tra due entità di un modello concettuale viene implementata attraverso...**

- A una chiave primaria.
  - B una tabella.
  - C una riga.
- 

14) **Nel campo dell'inferenza statistica cosa si intende per ipotesi nulla?**

- A Un'ipotesi non verificabile.
  - B L'ipotesi la cui probabilità è nulla.
  - C Una ipotesi su una popolazione statistica, la quale viene sottoposta a test sulla base di un campione.
- 

15) **Nel campo dell'inferenza statistica cosa rappresenta l'errore di prima specie?**

- A L'errore per cui si rifiuta l'ipotesi alternativa quando invece questa è vera.
- B L'errore che si commette nel calcolo della probabilità di un evento E.
- C L'errore per cui si rifiuta l'ipotesi nulla quando invece questa è vera.

- 
- 16) Nel campo dell'inferenza statistica quale delle seguenti NON è una proprietà degli stimatori?
- A Variabilità.
  - B Correttezza.
  - C Consistenza.
- 

- 17) Nell'ambito dell'inferenza statistica il test t di Student si applica...
- A solo a variabili continue.
  - B sia a variabili continue che discrete.
  - C solo a variabili discrete.
- 

- 18) Quale caratteristica di una distribuzione denota la seguente espressione:

$$\mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

- A Media.
  - B Indice di concentrazione.
  - C Varianza.
- 

- 19) Quale è lo scopo dell'analisi predittiva?
- A Analizzare soltanto i dati attuali allo scopo di effettuare previsioni su dati ed eventi futuri.
  - B Analizzare i dati attuali e storici allo scopo di effettuare previsioni su dati ed eventi futuri.
  - C Analizzare i dati attuali e storici allo scopo di condizionare dati ed eventi futuri.
- 

- 20) Nel machine learning non supervisionato...
- A vengono identificati pattern presenti nei dati, in assenza di dati storici etichettati.
  - B il modello viene addestrato sulla base di dati storici etichettati.
  - C vengono generati algoritmicamente nuovi dati.
- 

- 21) Nell'apprendimento con rinforzo...
- A gli algoritmi scelgono un'azione in un ambiente e poi l'agente viene ricompensato o penalizzato rispetto all'azione scelta.
  - B gli algoritmi apprendono da un set di dati etichettato e fornito in input.
  - C gli algoritmi suddividono i dati in cluster.
- 

- 22) Quale dei seguenti è un tipico utilizzo di metodi di apprendimento supervisionato?
- A Etichettare i dati di input.
  - B Dividere i dati in cluster.
  - C Eseguire previsioni.
- 

- 23) Per quale applicazione è inadatto un metodo di apprendimento non supervisionato?

- A Effettuare previsioni a partire da serie storiche.
  - B Riduzione dimensionale dei dati.
  - C Clustering dei dati.
- 

24) **Nella Cluster Analysis un elemento può appartenere a due diversi cluster?**

- A Sì, a patto che i due cluster abbiano almeno una proprietà in comune.
  - B Sì.
  - C No. La Cluster Analysis è mutuamente esclusiva.
- 

25) **Quale forma di analisi dei dati stima i coefficienti di un'equazione lineare di una o più variabili indipendenti al fine di predire il valore della variabile dipendente?**

- A Regressione lineare.
  - B Alberi di regressione.
  - C Regressione inversa.
- 

26) **Nell'ambito del metodo di regressione lineare cosa sono i residui?**

- A Sono la somma dei coefficienti dell'equazione lineare.
  - B Sono i coefficienti non considerati nel modello lineare.
  - C Sono la differenza tra i valori osservati e predetti.
- 

27) **Secondo il metodo dei minimi quadrati in un algoritmo di regressione lineare...**

- A la retta che approssima meglio i dati è quella che minimizza il quadrato della somma dei residui.
  - B la retta che approssima meglio i dati è quella che minimizza la somma dei residui.
  - C la retta che approssima meglio i dati è quella che minimizza la somma dei quadrati dei residui.
- 

28) **In cosa consiste l'operazione di data cleaning nella fase di pre-elaborazione dei dati?**

- A Nel portare tutti i dati allo stesso ordine di grandezza.
  - B Nell'eliminare tutti i dati prima di un nuovo caricamento.
  - C Nel riempire i campi con i valori mancanti e rimuovere i valori non realistici.
- 

29) **Il Mean Square Error (MSE) può essere calcolato come...**

- A la somma della distorsione e della deviazione standard.
  - B la somma della varianza e del quadrato della distorsione.
  - C la radice quadrata del coefficiente di determinazione R<sup>2</sup>.
- 

30) **Il K-Means è...**

- A un algoritmo di clusterizzazione.
  - B un algoritmo di calcolo della varianza.
  - C un algoritmo di apprendimento supervisionato.
- 

31) **In quale ambito viene utilizzato l'algoritmo di backpropagation?**

- A Nell'ambito dell'inferenza statistica.
  - B Nell'ambito delle reti neurali.
  - C Nell'ambito dell'ottimizzazione degli alberi di classificazione.
- 

32) **Nell'ambito del text mining, qual è il termine per indicare le parole non utilizzate nell'analisi?**

- A Tokens.
  - B Bad words.
  - C Stopwords.
- 

33) **Come si chiama la metodologia che utilizza tecniche di teoria dei grafi per analizzare i dati in termini delle relazioni esistenti tra diverse entità?**

- A Social Network Analysis.
  - B Entity-Relationship Analysis.
  - C Connection Analysis.
- 

34) **Nell'ambito del text mining quale delle seguenti operazioni NON viene svolta nel pre-processing dei testi?**

- A Controllo ortografico.
  - B Lemmatizzazione.
  - C Trasformazione del testo in minuscolo.
- 

35) **Quale delle seguenti è una tipica applicazione del text-mining?**

- A Riconoscimento automatico del testo da un'immagine (OCR).
  - B Sentiment Analysis.
  - C Creazione di sommari automatici in Microsoft Word.
- 

36) **La matrice delle adiacenze è ...**

- A un costrutto del linguaggio R.
  - B una struttura dati utilizzata per rappresentare grafi.
  - C una struttura dati specifica del text mining.
- 

37) **Quale delle seguenti affermazioni sulla qualità dei dati è vera?**

- A Non esistono standard che definiscono caratteristiche di qualità.
  - B Gli aspetti relativi alla qualità non sono mai misurabili in termini quantitativi.
  - C Le dimensioni di qualità sono definite in appositi standard.
- 

38) **Nell'ambito della qualità dei dati, la dimensione di completezza di un dataset indica...**

- A la capacità del dato di ricostruire pienamente la semantica dell'entità del mondo reale.
  - B la quantità di dati necessaria per svolgere una determinata attività.
  - C la percentuale di dati valorizzati rispetto alla popolazione di riferimento.
- 

39) **Quali delle seguenti caratteristiche NON attiene alla qualità dei dati?**

- A Compattezza.
  - B Accuratezza.
  - C Completezza.
- 

40) **In un dataset relativo ad un insieme di imprese, quale dimensione di qualità viene violata nel caso in cui ad un numero di dipendenti nullo corrisponde un costo del lavoro maggiore di zero?**

- A Tempestività.
- B Accuratezza.
- C Consistenza.