

# DOMANDE CONCORSO PUBBLICO PER ESAMI PER L'ASSUNZIONE A TEMPO INDETERMINATO DI N. 2 UNITA' DI PERSONALE - AREA DEI FUNZIONARI E DELL'ELEVATA QUALIFICAZIONE, PROFILO PROFESSIONALE "FUNZIONARIO SISTEMI INFORMATIVI E TECNOLOGIE" - REGIONE TOSCANA - PROVA SCRITTA - BUSTA 2

- 
- 1) **Nei sistemi di controllo versione, quale pratica è fondamentale per mantenere una storia dei commit chiara e gestibile?**
- A Evitare di documentare le modifiche nei commit usando invece i tag
  - B Accumulare numerosi cambiamenti in un unico commit generico
  - C Effettuare commit frequenti con messaggi descrittivi e significativi
- 
- 2) **Quale dei seguenti è un esempio di analisi supervisionata nel contesto del machine learning?**
- A Segmentare gli utenti di un sito web sulla base delle loro abitudini di navigazione o di acquisto, individuando pattern di comportamento.
  - B Analizzare un insieme di immagini per raggrupparle sulla base delle loro caratteristiche o dei loro contenuti, individuando categorie.
  - C Utilizzare un modello per prevedere il prezzo di una casa sulla base di caratteristiche come la superficie e il numero di stanze, usando un dataset con prezzi noti.
- 
- 3) **Quale delle seguenti affermazioni è errata riguardo al concetto di correlazione statistica?**
- A La correlazione indica sempre una relazione causa-effetto tra due variabili.
  - B Una correlazione positiva indica che all'aumentare di una variabile, l'altra aumenta.
  - C Una correlazione negativa indica che all'aumentare di una variabile, l'altra diminuisce.
- 
- 4) **Quale modello di servizio cloud consente agli sviluppatori di concentrarsi esclusivamente sullo sviluppo del codice e delle funzionalità, senza preoccuparsi della gestione delle infrastrutture sottostanti?**
- A Software as a Service (SaaS)
  - B Infrastructure as a Service (IaaS)
  - C Platform as a Service (PaaS)
- 
- 5) **Quale delle seguenti affermazioni descrive correttamente il processo di data ingestion?**
- A La data ingestion è il processo di pulizia e trasformazione dei dati raccolti in un data lake, per renderli pronti per l'analisi.
  - B La data ingestion è il processo di acquisizione e trasferimento dei dati da diverse fonti verso un data lake.
  - C La data ingestion è il processo di creazione di report a partire dai dati raccolti in un data lake.
- 
- 6) **Quale delle seguenti affermazioni descrive correttamente il concetto di metadati?**
- A I metadati forniscono informazioni strutturate su altre risorse, come dati, file e database, descrivendo la loro origine, formato, e struttura.
  - B I metadati sono dati trasformati e ottimizzati per essere utilizzati direttamente nell'ambito di analisi e reportistica.
  - C I metadati sono i dati grezzi che vengono raccolti da fonti esterne per le successive analisi.
- 
- 7) **Quale dei seguenti è un elemento chiave della data governance?**
- A La definizione e la creazione di report periodici e certificati sui dati raccolti.
  - B La definizione di ruoli, responsabilità e politiche per la gestione dei dati all'interno dell'organizzazione.
  - C La disponibilità di risorse dedicate all'analisi dei dati per prevedere tendenze e comportamenti futuri.
- 
- 8) **Quale dei seguenti ruoli è principale responsabilità di un Data Owner in un'organizzazione?**

- A Garantire che i dati siano utilizzati in conformità alle politiche aziendali e alle normative legali.
  - B Gestire l'implementazione delle soluzioni di archiviazione dei dati.
  - C Eseguire analisi statistiche sui dati per ottenere insight.
- 

9) Cosa si ottiene con una operazione di filtro in ambito di analisi dei dati?

- A L'applicazione di una modifica ai dati che ne aumenta la dimensione e ne riduce l'omogeneità.
  - B L'esclusione dei dati che non soddisfano determinati criteri, consentendo di concentrarsi solo su un sottoinsieme specifico di informazioni.
  - C La trasformazione dei dati in un formato facilmente leggibile da tutti gli utenti, indipendentemente dal loro livello di competenza.
- 

10) In quale situazione è più appropriato utilizzare un grafico a torta?

- A Quando si vuole rappresentare la distribuzione percentuale di una grandezza numerica sulla base di una variabile categorica qualitativa.
  - B Quando si vuole mostrare la relazione esistente tra due variabili numeriche, per esaminarne il livello di interdipendenza.
  - C Quando si desidera esaminare l'andamento e le tendenze nel tempo di un fenomeno rappresentato da una grandezza quantitativa.
- 

11) Quale delle seguenti definizioni descrive meglio i Linked Data?

- A Un metodo per organizzare e memorizzare dati in formato strutturato tabellare, per una facile consultazione e analisi.
  - B Una tecnica che permette di connettere e integrare dati provenienti da diverse fonti tramite l'uso di identificatori e risorse collegate tra loro.
  - C Un sistema di analisi dei dati che progettato per l'elaborazione di grandi volumi di informazioni non strutturate.
- 

12) Which of the following definitions is closest to the concept of a report?

- A A type of algorithm that allows data cleansing.
  - B A software used to create and manage relational databases.
  - C A document that presents data, analysis, and conclusions on a specific topic.
- 

13) Quali sono le quattro tipologie principali di database NoSQL?

- A Document-Based, Key-Value, Relational-Based, Dictionary-Based
  - B Document-Based, Key-Value, Column-Based, Graph-Based
  - C Document-Based, Key-Value, Relational-Based, Designer-Based
- 

14) Come gestiscono i database NoSQL la scalabilità orizzontale?

- A I database NoSQL non gestiscono la scalabilità orizzontale ma solo quella verticale.
  - B I database NoSQL gestiscono la scalabilità orizzontale distribuendo i dati su più nodi o server.
  - C I database NoSQL gestiscono la scalabilità orizzontale aumentando la capacità del singolo server.
- 

15) Quale delle seguenti definizioni meglio descrive una applicazione "responsive"?

- A E' ottimizzata per essere utilizzata su dispositivi mobili.
  - B Risponde più rapidamente alle interazioni dell'utente.
  - C E' progettata principalmente per desktop
- 

16) Cos'è un Data Lake e come differisce da un Data Warehouse?

- A Un Data Lake è un tipo di database relazionale per l'archiviazione di grandi moli di dati, mentre un Data Warehouse è un tipo di database NoSQL dedicato alla reportistica.
- B Un Data Lake è una piattaforma per l'archiviazione di grandi volumi di dati non strutturati o semi-strutturati, mentre un Data Warehouse è progettato per l'archiviazione dei dati in una struttura schematica.
- C Un Data Lake è uno storage di grandi dimensioni per dati strutturati, mentre un Data Warehouse è uno storage di medie dimensioni per dati non strutturati.

- 
- 17) **In un RDBMS, quale delle seguenti affermazioni circa l'integrità referenziale è corretta?**
- A La funzione principale della chiave esterna in un relazione tra tabelle è quella di migliorare le prestazioni del database
  - B L'integrità referenziale permette di modificare una chiave primaria senza modificare la chiave esterna.
  - C L'integrità referenziale previene inconsistenze nei dati mantenendo correttamente le relazioni tra tabelle
- 
- 18) **In un RDBMS, se si cerca di inserire un valore duplicato in una colonna definita come chiave primaria cosa succede?**
- A Il record viene inserito
  - B Viene restituito un errore e l'inserimento viene rifiutato
  - C Viene restituito un alert e il record viene inserito
- 
- 19) **Cosa significa "scalabilità" in un contesto cloud?**
- A Capacità di aumentare la sicurezza
  - B Capacità di ridurre i costi
  - C Maggiore capacità di aumentare o diminuire le risorse computazionali.
- 
- 20) **Qual è il ruolo di una CDN (Content Delivery Network)?**
- A Aumentare la latenza
  - B Aumentare la larghezza di banda
  - C Distribuire contenuti in modo efficiente
- 
- 21) **Cosa significa "latency" in un contesto di rete?**
- A Capacità di archiviazione
  - B Tempo di risposta
  - C Sicurezza dei dati
- 
- 22) **Quale dei seguenti è la principale funzione delle API?**
- A Gestire la cancellazione di record.
  - B Integrare sistemi diversi
  - C Gestire il backup di database.
- 
- 23) **What is the correct approach for creating an effective dashboard?**
- A Create a dashboard with as much information as possible, without worrying about simplicity and clarity.
  - B Use only complex and advanced charts to display all available data.
  - C Focus on the user's goals, select the most relevant data, and present information in a clear and visually understandable way.
- 
- 24) **What are the main components of a dashboard?**
- A Navigation bar, menu and advanced settings.
  - B Widgets, filters and charts.
  - C Log files, configuration files.
- 
- 25) **What is a data warehouse?**
- A A system for storing real-time data, optimized for frequent update operations.
  - B A software for managing relational databases.
  - C A centralized system for collecting, storing, and analyzing large volumes of data from various sources.
- 
- 26) **What characterizes a fully denormalized data model?**
- A Data is organized into multiple tables with many relationships between them to minimize redundancy.
  - B Data is stored in an unstructured format to facilitate complex analysis.
  - C Data is organized into a single table with some duplicate information to improve performance in read

**27) Quale delle seguenti attività descrive correttamente il data profiling?**

- A** Il data profiling implica l'analisi dei dati allo scopo di identificare e correggere errori, come valori mancanti, duplicati o incoerenze, e garantire la qualità dei dati prima di utilizzarli per analisi o decisioni.
  - B** Il data profiling implica la creazione di report mensili a supporto di decisioni strategiche sui dati raccolti.
  - C** Il data profiling si concentra sull'esplorazione dei dati per comprenderne la struttura, la distribuzione, i valori e le relazioni, allo scopo di fornire una base per supportare le successive azioni dell'analista.
- 

**28) Quale delle seguenti affermazioni descrive correttamente il funzionamento dell'architettura Lambda per l'elaborazione dei Big Data?**

- A** L'architettura Lambda si basa su un solo livello di elaborazione, gestendo sia i dati in tempo reale che i dati storici in un'unica fase di elaborazione per semplificare il processo e ridurre la complessità, uniformando scopo e metodo di elaborazione.
  - B** L'architettura Lambda gestisce i dati in tempo reale, evitando l'elaborazione dei dati batch, in modo da garantire che tutte le analisi siano basate su dati freschi e immediati, eliminando la necessità di memorizzare i dati storici e riducendo i costi di storage.
  - C** L'architettura Lambda prevede l'elaborazione dei dati in tempo reale nello "Speed Layer", mentre i dati storici vengono elaborati nel "Batch Layer"; entrambi i risultati vengono poi uniti nel "Serving Layer", integrando così i dati in tempo reale e i dati storici in un unico ambiente di restituzione finale.
- 

**29) Cos'è un catalogo dati?**

- A** Un sistema di archiviazione fisica dei dati e dei relativi metadati.
  - B** Un elenco centralizzato che descrive i dati disponibili in un'organizzazione, incluse le loro fonti e metadati.
  - C** Un software per l'analisi statistica dei dati basato sull'uso di ontologie.
- 

**30) Quali sono le caratteristiche desiderabili di un dato aperto?**

- A** I dati devono essere disponibili tramite una licenza aperta ed inclusi in documenti leggibili e interpretabili facilmente da un essere umano.
- B** I dati devono essere esposti usando standard W3C, quali RDF e SPARQL, ed essere descritti semanticamente tramite metadati e ontologie.
- C** I dati devono essere disponibili in forma strutturata con licenza aperta e utilizzare un formato proprietario.