



REGIONE EMILIA-ROMAGNA

**Procedura selettiva pubblica
per la copertura di n. 3 posti dell'area
"funzionari ed elevata qualificazione"
c.c.n.l. funzioni locali – profilo professionale
"specialista in analisi dei dati e statistica"**

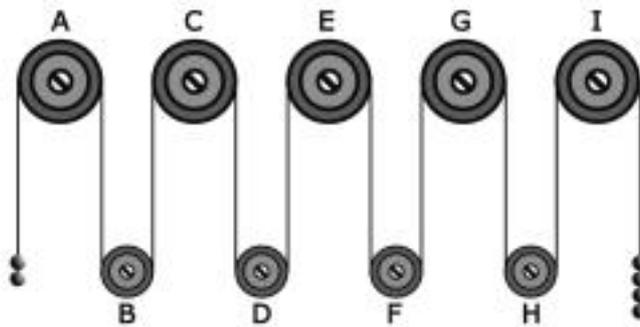
Prova 01

- 1 "In base ai dati in suo possesso, il ricercatore ha dedotto che non è sbagliata la tesi secondo cui quel materiale non è biodegradabile". Basandosi sulla precedente affermazione, quale delle seguenti alternative è vera?
- A** Quel materiale non è biodegradabile in base ai dati in possesso del ricercatore
 - B** Quel materiale è biodegradabile in base ai dati in possesso del ricercatore
 - C** Non si può dire che quel materiale non è biodegradabile

- 2 Quattro amici (A, B, C, D) hanno vinto 450 euro a una lotteria. Ad A spetta il doppio di B, a C una volta e mezzo la cifra di A e a D il triplo di B. Quanto riceve D?
- A** 150 euro
 - B** 50 euro
 - C** 160 euro

FIGURA SU 14

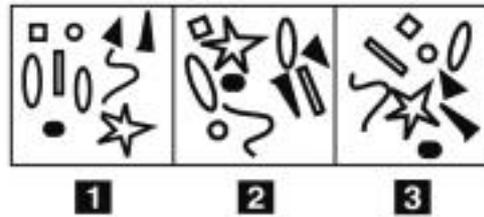
Il sistema indicato in figura si presenta nella sua fase iniziale con sfere di metallo tutte della stessa massa e fune di massa trascurabile.



- 3 Rispondere al seguente quesito facendo riferimento alla FIGURA SU 14
- Se il sistema è sottoposto soltanto al campo gravitazionale con assenza di attrito, allora si può concludere che:
- A** la carrucola G gira in senso orario
 - B** la carrucola B gira in senso orario
 - C** la carrucola A gira in senso antiorario

- 4 Una gioielleria ha venduto 400 anelli in dicembre e 580 in gennaio. A quanto ammonta l'incremento percentuale delle vendite da un mese all'altro?
- A** 45%
 - B** 50%
 - C** 18%

- 5 Maria corre più veloce di Paolo, che è più lento di Federico. Anna, che è più lenta di Federico, è più veloce di Laura. Dalle precedenti informazioni si può dedurre con certezza che:
- A** Laura è più lenta di Federico
 - B** Paolo è più lento di Anna
 - C** Anna è più veloce di Paolo



- 6 Rispondere al seguente quesito facendo riferimento alla FIGURA SZ 56 A00006
Individuare la figura da scartare.
A Figura 3
B Figura 2
C Figura 1

- 7 A00007
Quale tipologia di grafico è consigliabile per rappresentare l'andamento del tasso di inflazione nel tempo?
A Grafico a linee
B Grafico a torta
C Istogramma

- 8 A00008
Una variabile categoriale (o misurata su "scala nominale", secondo la classificazione delle variabili proposta da Steven) è rappresentata da:
A caratteri qualitativi sconnessi, detti *modalità*, che possono essere o dicotomiche o politomiche
B caratteri qualitativi ordinati, detti *modalità*, che possono essere o dicotomiche o politomiche
C caratteri quantitativi ordinati, detti *valori*, che possono essere o dicotomiche o politomiche

- 9 A00009
Un panel è un campione che contiene osservazioni:
A su N individui per T periodi
B su N individui per un periodo
C su un individuo per T periodi

- 10 A00010
Una serie storica è:
A una serie ordinata di numeri reali che misurano un certo fenomeno con riguardo alla sua evoluzione rispetto alla variabile tempo
B una serie ordinata di valori passati di una certa variabile, da cui è sempre possibile estrapolare i valori futuri di X
C una serie stocastica di valori passati di una certa variabile, da cui è sempre possibile estrapolare i valori futuri di X

-
- 11 A00011
Nel caso in cui tramite il "test del chi quadrato" si dimostri che la differenza tra due proporzioni:
A non è dovuta al caso, allora si può affermare che essa è statisticamente significativa
B non è dovuta al caso, allora si può affermare che essa è statisticamente non significativa
C è dovuta al caso, allora si può affermare che essa è statisticamente significativa
-
- 12 A00012
In una popolazione campione il 20% degli intervistati mangia frutta ogni giorno, il 10% mangia formaggio tutti i giorni e il 5% mangia ogni giorno sia formaggio sia frutta. Se, scegliendo a caso un individuo, questi ha dichiarato di mangiare frutta, che probabilità c'è che mangi ogni giorno anche formaggio?
A 25%
B 5%
C 20%
-
- 13 A00013
Il coefficiente di variazione di una distribuzione, rispetto all'unità di misura della variabile, è espresso:
A senza unità di misura in quanto si tratta di un numero puro
B in unità al quadrato
C in unità al cubo
-
- 14 A00014
Quale tra i seguenti è un database relazionale?
A MySQL
B Lotus Notes
C OrientDB
-
- 15 A00015
Qual è la differenza di fondo tra un "database" e "data warehouse"?
A Un database viene utilizzato per registrare dati; un data warehouse viene utilizzato per aggregare set di dati a fini analitici
B Un database contiene dati ordinati in tabelle; un data warehouse contiene dati non strutturati
C Un database consente l'estrazione e l'analisi dei dati contenuti; un data warehouse è utilizzato come archivio storico di dati
-
- 16 A00016
Nel linguaggio R, la struttura dati DataFrame rispetto alla struttura dati Matrix:
A può avere numero variabile di righe e colonne
B è una struttura omogenea
C può contenere solo caratteri
-
- 17 A00017
Se due variabili sono perfettamente e linearmente correlate, i loro valori:
A si dispongono tutti lungo una retta
B sono sempre positivi
C tendono a coincidere
-

-
- 18 **Che cos'è un "data mart"?** A00018
- A** Un database strutturato in base all'argomento, che spesso rappresenta una sezione del data warehouse aziendale
 - B** Un insieme di dati poco significativi che i data warehouse aziendali ignorano
 - C** Un altro modo per chiamare il cloud
-
- 19 **Quale tra le seguenti NON è una fase di un processo ETL?** A00019
- A** Compressione
 - B** Estrazione
 - C** Trasformazione
-
- 20 **Un buon tool di data preparation, in un progetto di Data Mining, deve:** A00020
- A** consentire l'accesso ai dati di qualunque provenienza (file xls, cvs, data warehouse e applicazioni cloud)
 - B** fornire un ambiente di sviluppo per l'implementazione di procedure ETL
 - C** fornire interfacce video per l'eliminazione dei dati
-
- 21 **Quale dei seguenti NON è un tipo di campionamento probabilistico?** A00021
- A** Campionamento per quote
 - B** Campionamento casuale semplice
 - C** Campionamento per *clusters*
-
- 22 **Quale delle seguenti tecniche di autenticazione è la più forte?** A00022
- A** Autenticazione a due fattori
 - B** Autenticazione basata su user-id e password
 - C** Autenticazione basata su PIN
-
- 23 **Quale dei seguenti strumenti permette di recuperare dati corrotti?** A00023
- A** Backup
 - B** Antivirus
 - C** Videosorveglianza
-
- 24 **La Direzione di un'azienda può comunicare la propria attenzione alla sicurezza informatica in un documento detto:** A00024
- A** politica per la sicurezza delle informazioni
 - B** regolamento per l'uso dei pc
 - C** bollettino dei vendor
-

-
- 25 **Riguardo il volume dei dati, che cosa si intende per Big Data?** A00025
- A** Un insieme di dati di quantità superiore a 100 TB
 - B** Un insieme di dati di quantità compresa entro 100 TB
 - C** Un qualunque insieme di dati
-
- 26 **L'interfaccia per l'utilizzatore di strumenti di Business intelligence è nota come:** A00026
- A** cruscotto o dashboard
 - B** tabella
 - C** schermata
-
- 27 **La missione di Eurostat è:** A00027
- A** fornire all'Unione europea un servizio informativo statistico di elevata qualità, con dati comparabili tra Paesi e Regioni
 - B** fornire all'Unione europea un servizio informativo statistico di elevata qualità, con dati assoluti e non localizzabili
 - C** fornire agli Stati nazionali dell'UE un servizio informativo statistico di elevata qualità, a finalità esclusivamente interna
-
- 28 **Nella teoria dei test d'ipotesi, l'errore del primo tipo si commette quando:** A00028
- A** si rifiuta H_0 quando è vera
 - B** si "accetta" H_1 quando è vera
 - C** si "accetta" H_0 quando è falsa
-
- 29 **L'intervista in profondità è:** A00029
- A** un tipo di intervista in cui il grado di strutturazione è prossimo allo zero, l'intervistatore non dispone di una traccia ma si limiterà ad introdurre il tema lasciando interamente al soggetto la conduzione dell'intervista
 - B** un tipo di intervista in cui l'intervistatore dispone di una lista di temi fissati in precedenza che deve sottoporre a tutti gli intervistati
 - C** un tipo di intervista in cui le domande a la sequenza in cui devono essere sottoposte sono fissate in anticipo
-
- 30 **Come vengono calcolate le componenti principali nell'Analisi delle Componenti Principali (PCA)?** A00030
- A** Le componenti principali sono calcolate come combinazioni lineari delle variabili originali
 - B** Le componenti principali sono la media dei dati nel set
 - C** Le componenti principali sono determinate tramite un algoritmo di clustering chiamato DBSCAN
-
- 31 **Qual è il vantaggio principale nell'utilizzo di una vista rispetto a una tabella?** A00031
- A** Le viste possono nascondere la complessità dei dati sottostanti e semplificare le interrogazioni
 - B** Le viste sono più efficienti dal punto di vista delle prestazioni rispetto alle tabelle
 - C** Le viste consentono di memorizzare i dati in modo permanente, mentre le tabelle no
-

-
- 32 A00032
Che cos'è Power BI?
A Power BI è una piattaforma unificata e scalabile per business intelligence (BI) aziendale e in modalità self-service
B Power BI è una piattaforma volta a recuperare da istituti terzi e provider esterni le Business Information relative alle aziende
C Power BI è un software di Machine Learning in grado di accedere ai dati aziendali e prevedere il loro andamento
-
- 33 A00033
Cosa si intende per "indicizzazione" nel contesto di un database relazionale?
A Un processo di ottimizzazione delle prestazioni nella selezione dei dati presenti in tabella
B Un processo di archiviazione dei dati
C Un processo di crittografia dei dati
-
- 34 A00034
Quale di questi strumenti NON può essere utilizzato per visualizzare reportistica basata sull'analisi dei dati?
A ElasticSearch
B Tableau
C Power BI
-
- 35 A00035
Quale dei seguenti è uno dei principali vantaggi dell'utilizzo di Power BI per la reportistica?
A Possibilità di creare report interattivi e dashboard dinamici con dati provenienti da varie fonti
B Capacità di integrare solo dati provenienti da database relazionali
C Possibilità di costruire report in PDF usando direttamente SQL
-
- 36 A00036
Qual è lo scopo di utilizzare un modello di sentence embedding nelle applicazioni di NLP (Natural Language Processing)?
A La possibilità di catturare il significato di frasi differenti e calcolarne la somiglianza
B La possibilità di salvare frasi differenti in maniera cifrata non riconoscibile senza una chiave privata
C La capacità di salvare molteplici frasi all'interno di un indice ElasticSearch
-
- 37 A00037
Cos'è il "masked language modeling"?
A Una tipologia di obiettivo nei modelli linguistici basato sulla previsione di una parola occultata all'interno di una frase
B Una tipologia di obiettivo nei modelli linguistici basato sulla previsione di una parola occultata alla fine di una frase
C Una tipologia di obiettivo nei modelli linguistici basato sulla previsione di una parola all'interno di una frase che è stata sostituita con un sinonimo
-
- 38 A00038
Che cosa è l'effetto "acquiescent set"?
A La tendenza degli intervistati a rispondere, alle domande su temi sui quali non hanno opinioni precise, secondo quello che credono essere gradito all'intervistatore
B La tendenza a scegliere da parte dell'intervistato, in una batteria di domande, sempre la stessa risposta
C Il rifiuto da parte dell'intervistato a rispondere a determinate domande
-

-
- 39 A00039
- Come avviene un campionamento stratificato non proporzionale?**
- A** La popolazione viene divisa in strati omogenei rispetto alla variabile da stimare ma il numero di casi per ogni strato viene scelto sulla base di esigenze di analisi e quindi non saranno proporzionali rispetto la popolazione
 - B** La popolazione è divisa in strati omogenei rispetto alla variabile e si estrae un campione casuale semplice da ciascuno strato; in seguito si uniscono i campioni dei singoli strati per ottenere il campione finale. Questa procedura richiede che per tutte le unità della popolazione sia nota la variabile posta alla base della stratificazione
 - C** Le unità campionarie vengono scelte in modo razionale sulla base di alcune loro caratteristiche, e si utilizza quando l'ampiezza del campione è limitata

-
- 40 A00040
- Che cos'è il sistema CAWI?**
- A** È un sistema computerizzato di rilevazione che prevede l'invio di una e-mail a tutte le unità statistiche campionate contenente l'invito a partecipare alla ricerca e il sito dove risiede fisicamente il questionario da compilare
 - B** Una modalità di rilevazione diretta di unità statistiche realizzata attraverso interviste telefoniche, dove l'intervistatore legge le domande all'intervistato e registra le risposte su un computer, tramite un apposito software
 - C** Un insieme di tecniche per analizza le reti sociali

-
- 41 A00041
- Cos'è uno shapefile?**
- A** È un formato vettoriale utilizzato nell'ambito dei sistemi informativi geografici per archiviare la posizione, la forma e gli attributi delle identità geografiche
 - B** È un formato di file utilizzato in informatica per memorizzare immagini
 - C** È un formato dati utilizzato per la rappresentazione di immagini raster

-
- 42 A00042
- Qual è il grafico più appropriato per mostrare la relazione tra due variabili quantitative?**
- A** Il grafico di dispersione
 - B** Il grafico ad area
 - C** Il grafico a torta

-
- 43 A00043
- Quale dei seguenti termini è quello meglio descritto dalla definizione: "Un repository centralizzato in grado di archiviare, elaborare e proteggere grandi quantità di dati strutturati, semistrutturati e non strutturati nel loro formato nativo."?**
- A** Data lake
 - B** Data warehouse
 - C** ETL

-
- 44 A00044
- In statistica, l'errore di secondo tipo consiste:**
- A** nell'accettare l'ipotesi nulla quando è falsa
 - B** nel rifiutare l'ipotesi nulla quando è vera
 - C** nell'accettare l'ipotesi nulla quando è vera

-
- 45 A00045
In statistica, quale di queste affermazioni sulla validazione dei dati è vera?
A La validazione dei dati è un processo iterativo che si ripete più volte durante il ciclo di vita dell'analisi
B La validazione dei dati è un processo una tantum che si effettua all'inizio del progetto di analisi
C La validazione dei dati è necessaria solo per i dataset di grandi dimensioni o complessi
-
- 46 A00046
Quale misura di occorrenza è appropriata per effettuare un confronto della mortalità tra differenti aree geografiche (es. regioni)?
A Tasso di mortalità standardizzato con metodo diretto (con popolazione standard)
B Il rapporto di mortalità standardizzato con metodo indiretto (con tassi standard)
C Sia tasso di mortalità standardizzato con metodo diretto che rapporto di mortalità standardizzato con metodo indiretto
-
- 47 A00047
Cos'è SAS?
A Sistema software completo che offre una vasta gamma di funzionalità per la gestione, l'analisi e la visualizzazione dei dati
B Pacchetto statistico specializzato nella modellistica per i trasporti
C Pacchetto statistico specializzato per l'analisi in ambito finanziario
-
- 48 A00048
Il modello delle 5 V dei big data, per definire nuovi dati generati dall'aumento delle fonti informative, NON COMPRENDE:
A verosimiglianza
B varietà
C veridicità
-
- 49 A00049
L'indagine multiscopo ISTAT sulle famiglie – aspetti della vita quotidiana – viene svolta:
A ogni anno
B ogni 5 anni
C ogni 2 anni
-
- 50 A00050
Secondo il Regolamento UE 2016/679 (GDPR), i dati personali oggetto di trattamento sono, tra l'altro:
A conservati in una forma che consenta l'identificazione dell'interessato
B conservati sempre su supporto cartaceo
C consultabili anche da terzi, previa autorizzazione del Garante, senza necessità di consenso dell'interessato
-
- 51 A00051
Una tra le seguenti affermazioni NON è espressamente conforme al Regolamento UE 2016/679 (GDPR) per quanto riguarda l'informativa preventiva sulla raccolta dei dati personali: quale?
A L'informativa deve indicare il diritto al risarcimento del danno derivante da un eventuale uso improprio dei dati
B L'informativa deve indicare le finalità del trattamento cui sono destinati i dati
C L'informativa deve indicare gli eventuali destinatari dei dati personali
-

-
- 52 A00052
In base al Regolamento UE 2016/679 (GDPR):
A la cancellazione e la distruzione dei dati costituiscono trattamento degli stessi
B la cancellazione dei dati non costituisce trattamento degli stessi, a differenza dalla distruzione
C la distruzione dei dati non costituisce trattamento degli stessi, a differenza dalla cancellazione
-
- 53 A00053
In base all'art. 4 del Regolamento UE 2016/679, il consenso dell'interessato a che i dati personali che lo riguardano siano oggetto di trattamento deve essere espresso:
A mediante dichiarazione o azione positiva inequivocabile
B solo mediante dichiarazione
C solo mediante dichiarazione scritta
-
- 54 A00054
In base all'art. 72 dello Statuto della Regione Emilia-Romagna, l'attività amministrativa è soggetta al controllo:
A di gestione
B di qualità
C di revisione
-
- 55 A00055
In base allo Statuto della Regione Emilia-Romagna, individuare il completamento ERRATO: il Consiglio delle Autonomie locali è organo ...
A di indirizzo
B di rappresentanza
C di consultazione
-
- 56 A00056
In base al Codice di comportamento della Regione Emilia-Romagna, così come da ultimo modificato dalla delibera della Giunta Regionale 905/2018, il Codice si applica ai titolari di contratti di collaborazione coordinata e continuativa da svolgersi a favore della Regione?
A Sì, per quanto compatibile
B Sì, ma solo se la collaborazione viene espletata all'interno di strutture, sedi o uffici regionali
C Sì, per quanto compatibile, ma solo se la collaborazione ha durata superiore a 30 giorni lavorativi
-
- 57 A00057
In base all'art. 4 del Codice di comportamento della Regione Emilia-Romagna, così come da ultimo modificato dalla delibera della Giunta Regionale 905/2018, il collaboratore regionale può sollecitare regali o utilità?
A No, mai
B Sì, ma solo se di modico valore e nell'ambito delle normali relazioni di cortesia istituzionali o in occasione di festività
C Sì, ma solo per altri e se di modico valore e nell'ambito delle normali relazioni di cortesia istituzionali o in occasione di festività
-

58

In Microsoft Excel, la notazione "\$C2" inserita in una formula indica un riferimento:

- A** misto
- B** assoluto
- C** incrociato

59

"Central banks differ from commercial banks in two main ways. Firstly, they are public bodies and therefore non-profit organisations. They are entirely owned by governments and implement their policies. Secondly, they do not deal with the general public. Individuals cannot open an account with them or ask for a loan". According to the text, who entirely owns central banks?

- A** Governments
- B** Private investors
- C** Overseas entrepreneurs

60

"Companies are investing in technological innovation to find ways of reducing the cost of delivery, while at the same time meeting their customers' demands for fast and reliable delivery services. Two of the newest methods that they are experimenting with are drone delivery and self-driving cars. Amazon recently started drone delivery through its Prime Air delivery service. Amazon is a trendsetter, so if the company starts relying heavily on drones, then it's probable that a lot of other companies will start doing the same". What has Amazon recently started to use for their deliveries?

- A** Drones
- B** Self-driving cars
- C** Trains

