

**ΑΝΩΤΑΤΟ ΣΥΜΒΟΥΛΙΟ ΕΠΙΛΟΓΗΣ ΠΡΟΣΩΠΙΚΟΥ
ΚΕΝΤΡΙΚΗ ΕΠΙΤΡΟΠΗ ΔΙΑΓΩΝΙΣΜΟΥ**

ΠΡΟΚΗΡΥΞΗ 1Γ/2022
ΟΜΑΔΑ Δ
Κλάδοι ΠΕ –ΤΕ ΠΛΗΡΟΦΟΡΙΚΗΣ

Εξέταση στο μάθημα **Σχεδιασμός και Διαχείριση Βάσεων Δεδομένων και Επιχειρησιακή Νοημοσύνη**
Κυριακή 23 Οκτωβρίου 2022

Το **ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ** που ακολουθεί αφορά στο μάθημα Σχεδιασμός και Διαχείριση Βάσεων Δεδομένων και Επιχειρησιακή Νοημοσύνη και αποτελείται από εξήντα (60) ερωτήσεις πολλαπλών επιλογών, **βαθμολογικά ισοδύναμες**, με τέσσερις (4) εναλλακτικές απαντήσεις (Α, Β, Γ, Δ), από τις οποίες **μία (1) μόνο είναι η ορθή**.

Το μάθημα βαθμολογείται με άριστα τις εκατό (**100**) μονάδες, ενώ ως βαθμολογία **βάσης** του μαθήματος ορίζονται οι **πενήντα πέντε (55)** μονάδες. Για κάθε λανθασμένη απάντηση εφαρμόζεται **αρνητική βαθμολόγηση** με συντελεστή μείον 35%.

Για τις απαντήσεις σας να χρησιμοποιήσετε το **Απαντητικό Φύλλο** σύμφωνα με τις οδηγίες που σας έχουν διανεμηθεί.

ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

1. Ένα Σύστημα Διαχείρισης Βάσης Δεδομένων είναι:

- A. Ένα σύνολο από αρχεία δεδομένων
- B. Ένας εξυπηρετητής δεδομένων
- Γ. Το λογισμικό μέσω του οποίου γίνεται η δημιουργία, η διαχείριση, η συντήρηση και η χρήση μιας ηλεκτρονικής βάσης δεδομένων
- Δ. Ένα πληροφοριακό σύστημα που διαχειρίζεται δεδομένα

2. Ποιο από τα παρακάτω επίπεδα αφαίρεσης δεδομένων (abstraction levels) αναφέρεται στον τρόπο που είναι αποθηκευμένα τα δεδομένα σε μια βάση δεδομένων;

- A. Εννοιολογικό Επίπεδο
- B. Φυσικό Επίπεδο
- Γ. Λογικό Επίπεδο
- Δ. Λειτουργικό Επίπεδο

3. Το σχεσιακό μοντέλο δεδομένων αποτελεί ένα _____ αφαίρεσης δεδομένων.

- A. Εννοιολογικό Επίπεδο
- B. Φυσικό Επίπεδο
- Γ. Λογικό Επίπεδο
- Δ. Λειτουργικό Επίπεδο

4. Σε έναν σχεσιακό πίνακα, ισχύει ότι η σειρά με την οποία είναι αποθηκευμένες οι εγγραφές:

- A. διαφοροποιεί τα αποτελέσματα (τις εγγραφές που θα φέρει) ενός ερωτήματος
- B. δεν διαφοροποιεί τα αποτελέσματα (τις εγγραφές που θα φέρει) ενός ερωτήματος
- Γ. είναι αύξουσα με βάση κάποιο από τα πεδία του πίνακα
- Δ. κανένα από τα παραπάνω

5. Κατά τον σχεδιασμό μιας σχεσιακής βάσης δεδομένων, ποιον από τους παρακάτω πίνακες θα δημιουργήσετε με τα αντίστοιχα κλειδιά, για να αποτυπώσετε τη σχέση ότι ένας εργαζόμενος μπορεί να δουλεύει σε πολλά έργα;

- A. Employee (emp_id, ...), Project(proj_id,...)
- B. Employee (emp_id,..., proj_id), Project(proj_id,...)
- Γ. Employee (emp_id,...), Project(proj_id,..., worksFor)
- Δ. Employee (emp_id,...), Project(proj_id,...), WorksFor (emp_id, proj_id)

6. Κατά τον σχεδιασμό ενός πίνακα που διατηρεί στοιχεία εργαζομένων σε μια σχεσιακή βάση δεδομένων, θέλετε να ορίσετε ένα νέο πεδίο (π.χ. email επικοινωνίας) που μπορεί να διατηρεί πολλαπλές τιμές (π.χ. email 1, email 2, ...) και ένα εκ των οποίων είναι το βασικό για την επικοινωνία (π.χ. ένα πεδίο is primary που παίρνει τιμές yes/no). Ποιος από τους παρακάτω τρόπους σχεδιασμού θεωρείται ότι είναι ο πιο αποδοτικός;

- A. Employee (emp_id, ..., email1, is_email1_primary, email2, is_email2_primary, ..., email10, is_email10_primary)
- B. Employee (emp_id,...), Emp_Email (emp_id, email, is_primary)
- Γ. Employee (emp_id, email, is_primary)
- Δ. Employee (emp_id,...), Email (emailID, email), Emp_Email (emp_id, emailID, is_primary)

7. Ποιο από τα παρακάτω ερωτήματα SQL θα επιστρέψει το όνομα των υπαλλήλων που δουλεύουν στο τμήμα με όνομα «Λογιστήριο»;

- A. SELECT emp.name FROM emp, dept WHERE dept.name='Λογιστήριο';
- B. SELECT emp.name FROM emp INNER JOIN dept ON emp.deptId=dept.id WHERE dept.name='Λογιστήριο';
- Γ. SELECT emp.name FROM emp INNER JOIN dept ON dept.name='Λογιστήριο';
- Δ. SELECT emp.name FROM emp OUTER JOIN dept ON dept.name='Λογιστήριο';

8. Ποιο από τα παρακάτω ερωτήματα SQL θα επιστρέψει λίστα εγγραφών με τον κωδικό τμήματος, το όνομα και τον μισθό όλων των υπαλλήλων ταξινομημένη πρώτα με αύξουσα σειρά βάσει του κωδικού τμήματος και του ονόματος και στη συνέχεια με φθίνουσα σειρά βάσει του μισθού;

- A. SELECT deptId ASC, name ASC, salary DESC FROM emp;
- B. SELECT deptId, name, salary FROM emp ORDER BY deptId ASC, salary, name DESC;
- Γ. SELECT deptId, name, salary FROM emp GROUP BY deptId, name, salary DESC;
- Δ. SELECT deptId, name, salary FROM emp ORDER BY deptId, name, salary DESC;

9. Εξηγήστε τι επιστρέφει το ακόλουθο ερώτημα SQL:

```
SELECT count(emp.Name), emp.deptID  
FROM emp  
GROUP BY emp.deptID;
```

- A. Επιστρέφει τον αριθμό των υπαλλήλων ανά κωδικό τμήματος που δουλεύουν.
- B. Επιστρέφει τον αριθμό των μοναδικών ονομάτων των υπαλλήλων ανά κωδικό τμήματος που δουλεύουν.
- Γ. Επιστρέφει τον συνολικό αριθμό των υπαλλήλων που δουλεύουν σε όλα τα τμήματα.
- Δ. Επιστρέφει τον αριθμό των υπαλλήλων ανά κωδικό τμήματος που δουλεύουν, εξαιρώντας τους υπαλλήλους που δεν ανήκουν σε κάποιο τμήμα.

10. Ποια από τις παρακάτω εντολές SQL χρησιμοποιείται για τη διαγραφή εγγραφών από έναν πίνακα σε μια Βάση Δεδομένων;

- A. DELETE
- B. DROP
- Γ. REMOVE
- Δ. ALTER

11. Ποια από τις παρακάτω εντολές SQL χρησιμοποιείται για τη διαγραφή ενός πίνακα από μια Βάση Δεδομένων;

- A. DELETE
- B. DROP
- Γ. REMOVE
- Δ. ALTER

12. Εξηγήστε τι κάνει η παρακάτω εντολή:

```
INSERT INTO emp_copy (name, salary)  
SELECT name, salary FROM emp WHERE emp.deptID=20;
```

- A. Αντιγράφει όλες τις εγγραφές του πίνακα emp στον πίνακα emp_copy
- B. Μεταφέρει στον πίνακα emp_copy, το όνομα και τον μισθό όλων των υπαλλήλων που δουλεύουν στο τμήμα με κωδικό 20 και στη συνέχεια τις διαγράφει από τον πίνακα emp.
- Γ. Εισάγει στον πίνακα emp_copy, το όνομα και τον μισθό των υπαλλήλων από τον πίνακα emp που δουλεύουν στο τμήμα με κωδικό 20
- Δ. Ενημερώνει το όνομα και τον μισθό των υπαλλήλων που υπάρχουν στον πίνακα emp_copy με το όνομα και τον μισθό των υπαλλήλων από τον πίνακα emp που δουλεύουν στο τμήμα με κωδικό 20

13. Εξηγήστε τι κάνει η παρακάτω εντολή:

```
DELETE FROM employees  
WHERE department_id in  
(SELECT department_id  
FROM departments  
WHERE department_name LIKE '%Accounting%');
```

- A. Διαγραφή όλων των υπαλλήλων που δουλεύουν στο τμήμα Accounting
- B. Διαγραφή όλων των υπαλλήλων του τμήματος που περιέχει τη λέξη Accounting στο όνομά του
- Γ. Διαγραφή όλων των τμημάτων που περιέχει τη λέξη Accounting στο όνομά του
- Δ. Διαγραφή όλων των υπαλλήλων του τμήματος που το όνομα του αρχίζει από τη λέξη Accounting

14. Εξηγήστε τι επιστρέφει το ακόλουθο ερώτημα SQL:

```
SELECT emp.name, emp.salary
FROM emp LEFT OUTER JOIN dept on emp.deptId=dept.id
WHERE dept.id is null;
```

- A. Επιστρέφει το όνομα και τον μισθό των υπαλλήλων που δουλεύουν σε κάποιο τμήμα χωρίς όνομα.
- B. Επιστρέφει το όνομα και τον μισθό των υπαλλήλων που δουλεύουν στο τμήμα με κωδικό = null;
- Γ. Επιστρέφει το όνομα και τον μισθό των υπαλλήλων που δε δουλεύουν σε κάποιο τμήμα.
- Δ. Επιστρέφει το όνομα και τον μισθό όλων των υπαλλήλων που δουλεύουν ανεξαρτήτως τμήματος.

15. Ποιο από τα ερωτήματα θα επιστρέψει τα ίδια αποτελέσματα με το παρακάτω:

```
SELECT name, course_id
FROM instructor, teaches
WHERE instructor_id= teaches_id;
```

- A. SELECT name, course_id FROM instructor RIGHT OUTER JOIN teaches ON instructor_id= teaches_id;
- B. SELECT name, course_id FROM instructor, teaches WHERE instructor_id= course_id;
- Γ. SELECT name, course_id FROM instructor;
- Δ. SELECT name, course_id FROM instructor INNER JOIN teaches ON instructor_id= teaches_id;

16. Ποιο από τα ερωτήματα θα χρησιμοποιούσατε στη διαδικασία εύρεσης διπλότυπων εγγραφών (με το ίδιο όνομα, επώνυμο και ΑΔΤ) από τον πίνακα employees;

- A. SELECT name, surname, id_card FROM employees;
- B. SELECT DISTINCT name, surname, id_card FROM employees;
- Γ. SELECT name, surname, id_card FROM employees GROUP BY name, surname, id_card HAVING count(*)>1;
- Δ. SELECT name, surname, id_card FROM employees ORDER BY name, surname, id_card DESC;

17. Ποιος είναι ο λόγος που θέλετε να δημιουργήσετε μια όψη (view) σε μια Βάση Δεδομένων (ΒΔ);

- A. Για απλούστευση πολύπλοκων ερωτημάτων.
- B. Συντελεί στη μεγαλύτερη ασφάλεια των βάσεων δεδομένων, καθώς απομονώνει από τους χρήστες των εφαρμογών το σχήμα της ΒΔ.
- Γ. Για βελτίωση της ταχύτητας εκτέλεσης πολύπλοκων ερωτημάτων.
- Δ. Όλα τα παραπάνω.

18. Τί είναι μια όψη (view) σε μια Βάση Δεδομένων (ΒΔ);

- A. ένας εικονικός πίνακας που προκύπτει από άλλους πίνακες μέσω ενός ερωτήματος SQL, για την οποία δεν απαιτείται να είναι αποθηκευμένη σε φυσική μορφή
- B. μια συνάρτηση που εκτελείται στους πίνακες της ΒΔ
- Γ. μηχανισμός εξουσιοδότησης πρόσβασης σε πίνακες σε μια ΒΔ
- Δ. ένα παραμετροποιημένο ερώτημα που χρησιμοποιείται στην παραγωγή αναφορών

19. Ποια είναι η λειτουργία ενός δείκτη (index) σε μια σχεσιακή βάση δεδομένων;

- A. Να βελτιώσει την ταχύτητα εκτέλεσης όλων των ερωτημάτων DDL.
- B. Να βελτιώσει την ταχύτητα εκτέλεσης συγκεκριμένων ερωτημάτων DDL.
- Γ. Να βελτιώσει την ταχύτητα εκτέλεσης συγκεκριμένων ερωτημάτων SQL.
- Δ. Να βελτιώσει την απόδοση συνολικά της βάσης δεδομένων.

20. Στον πίνακα των φορολογουμένων, θέλετε να ορίσετε έναν δείκτη στο πεδίο «Οικογενειακή κατάσταση», που λαμβάνει προκαθορισμένες τιμές (π.χ. έγγαμος/η, άγαμος/η, σε σύμφωνο κ.λπ.). Τι τύπου δείκτη θα χρησιμοποιήσετε;

- A. B-Tree
- B. Bitmap
- Γ. Binary
- Δ. Primary

21. Στον πίνακα των φορολογουμένων, θέλετε να βελτιστοποιήσετε το παρακάτω ερώτημα. Τι τύπου δείκτη θα χρησιμοποιήσετε;

```
SELECT taxpayerID, taxpayerName, income  
FROM taxPayer  
WHERE income between 10K and 15K;
```

- A. B-Tree στο πεδίο taxpayerID
- B. B-Tree στα πεδία taxpayerID, taxpayerName, income
- Γ. B-Tree στο πεδίο income
- Δ. Clustered index στο πεδίο income

22. Ποια είναι η λειτουργία της εντολής MERGE....INTO.... στην SQL;

- A. Εισάγει εγγραφές από έναν πίνακα σε έναν άλλο.
- B. Ελέγχει αν οι εγγραφές από έναν πίνακα υπάρχουν στον άλλον, αν υπάρχουν τις ενημερώνει, αλλιώς τις εισάγει.
- Γ. Ενημερώνει τις εγγραφές από έναν πίνακα με βάση τις εγγραφές ενός άλλου.

Δ. Τίποτα από τα παραπάνω. Χρησιμοποιείται για την αρχικοποίηση ενός πίνακα με δεδομένα από αρχεία.

23. Κατά την εκτέλεση μιας δοσοληψίας (transaction) που δεν έχει ολοκληρωθεί, ποιες από τις παρακάτω εντολές εκτελείται αυτόματα στην περίπτωση ενός αναπάντεχου shutdown;

- A. ROLLBACK
- B. COMMIT
- Γ. DELETE
- Δ. SAVE

24. Μια μεταφορά χρημάτων μεταξύ δύο λογαριασμών διακόπτεται. Τα χρήματα αφαιρούνται από τον πρώτο λογαριασμό, αλλά δεν προστίθενται στον δεύτερο λογαριασμό. Ποια από τις παρακάτω αρχές παραβιάζεται;

- A. Ατομικότητα (Atomicity)
- B. Συνέπεια (Consistency)
- Γ. Απομόνωση (Isolation)
- Δ. Ανθεκτικότητα (Durability)

25. Εάν ένας δείκτης _____, τα μεταδεδομένα και τα στατιστικά παραμένουν ως έχουν.

- A. μεταβληθεί ως προς τον ορισμό του
- B. απενεργοποιηθεί
- Γ. διαγραφεί
- Δ. απενεργοποιηθεί και μεταβληθεί

26. Ποια είναι η πολυπλοκότητα ενός αλγορίθμου γραμμικής εύρεσης (linear search) με βάση τον αριθμό των εγγραφών n;

- A. $O(n)$
- B. $O(\log n)$
- Γ. $O(n^2)$
- Δ. $O(n \log n)$

27. Οι συνήθεις εργασίες του Διαχειριστή Βάσεων Δεδομένων (Database Administrator - DBA) περιλαμβάνουν:

- I. Τη λήψη αντιγράφων ασφαλείας κάθε βάσης δεδομένων που διατηρεί ο οργανισμός
- II. Τη διαχείριση των χρηστών και των δικαιωμάτων πρόσβασης στα αντικείμενα της βάσης δεδομένων
- III. Τη βελτιστοποίηση της βάσης δεδομένων μέσω κατάλληλης αναδιοργάνωσης των αρχείων στον δίσκο
- IV. Τη δημιουργία πινάκων σε μια βάση δεδομένων.

- A. Το I και το II
- B. Όλα τα παραπάνω
- Γ. Το I , II, III
- Δ. Το II, III

28. Τι δικαιώματα δίνει η παρακάτω εντολή;

GRANT select, insert ON taxpayers TO bob, alice;

- A. Επιλογής και εισαγωγής στα δεδομένα του πίνακα taxpayers στους χρήστες (με username) bob και alice
- B. Επιλογής στα δεδομένα του πίνακα taxpayers στον χρήστη bob και εισαγωγής στον χρήστη alice.
- Γ. Εισαγωγής μόνο στα δεδομένα του πίνακα taxpayers στους χρήστες bob και alice.
- Δ. Επιλογής και εισαγωγής από τα οποία προκύπτει και η τροποποίηση και διαγραφή στα δεδομένα του πίνακα taxpayers στους χρήστες bob και alice.

29. Τι είναι οι ρόλοι χρηστών σε μια βάση δεδομένων;

- A. Μια ομαδοποίηση από δικαιώματα που μπορούν να ανατεθούν σε χρήστες.
- B. Ένας ειδικός τύπος χρηστών με αυξημένα δικαιώματα.
- Γ. Ψευδώνυμα που μπορούν να ανατεθούν σε χρήστες της βάσης δεδομένων.
- Δ. Τίποτα από τα παραπάνω.

30. Πώς ονομάζεται το κλειδί κρυπτογράφησης που χρησιμοποιείται τόσο για την κρυπτογράφηση όσο και για την αποκρυπτογράφηση των δεδομένων;

- A. Δημόσιο κλειδί
- B. Ιδιωτικό κλειδί
- Γ. Συμμετρικό κλειδί
- Δ. Ασύμμετρο κλειδί

31. Ένα μήνυμα, που είναι κρυπτογραφημένο με ένα δημόσιο κλειδί, μπορεί να αποκρυπτογραφηθεί:

- A. μόνο με το ιδιωτικό κλειδί του δέκτη.
- B. με το ίδιο δημόσιο κλειδί.
- Γ. με τον συνδυασμό ιδιωτικού και δημόσιου κλειδιού.
- Δ. με ένα password που στέλνει ο αποστολέας στον δέκτη.

32. Πώς ονομάζονται οι επιθέσεις, όπου ο επιτιθέμενος επιδιώκει να εκτελέσει ένα κακόβουλο λογισμικό μέσω της εκτέλεσης ενός ερωτήματος SQL;

- A. SQL attack
- B. DDOS attack
- Γ. SQL injection
- Δ. SQL bombing

33. Ποια θεωρείται μια καλή στρατηγική για τη λήψη αντιγράφων ασφαλείας σε μια παραγωγική βάση δεδομένων με καθημερινές δοσοληψίες;

- A. Λήψη διαφορετικών αντιγράφων ασφαλείας καθημερινά και ένα ολικό μια φορά την εβδομάδα.
- B. Λήψη ολικών αντιγράφων ασφαλείας καθημερινά.

Γ. Λήψη διαφορικών αντιγράφων ασφαλείας καθημερινά και ένα ολικό backup μια φορά το εξάμηνο.

Δ. Λήψη ολικών αντιγράφων ασφαλείας καθημερινά και διαφορικών αντιγράφων μηνιαία.

34. Ποια θεωρείται σημαντική εργασία ενός Διαχειριστή Βάσεων Δεδομένων (Database Administration-DBA) για τη σωστή λήψη αντιγράφων ασφαλείας (backup);

A. Να λαμβάνει χειροκίνητα backup.

B. Να δοκιμάζει ανά τακτά διαστήματα και να επικυρώνει την επαναφορά ενός αντιγράφου ασφαλείας σε μια δοκιμαστική βάση δεδομένων.

Γ. Να μεταφέρει τα αντίγραφα ασφαλείας σε κρυπτογραφημένο χώρο.

Δ. Να διαγράφει τα παλιά αντίγραφα ασφαλείας.

35. Η ενοποίηση πληροφορίας (information integration) σε έναν οργανισμό επιτρέπει:

A. Την αντιγραφή (replication) δεδομένων από ένα πληροφοριακό σύστημα σε ένα άλλο.

B. Την αξιοποίηση δεδομένων από διαφορετικά πληροφοριακά συστήματα του οργανισμού.

Γ. Την απομακρυσμένη πρόσβαση σε δεδομένα.

Δ. Την πρόσβαση σε δεδομένα ίδιου μόνο μορφότυπου (format).

36. Ποια/ες από τις ακόλουθες επιλογές θεωρείτε ότι είναι τα πιο σημαντικά βήματα για τη βελτίωση της ποιότητας σε διαδικασίες ενοποίησης δεδομένων ενός οργανισμού;

I. Βεβαιωθείτε ότι τα δεδομένα είναι καθαρά

II. Βεβαιωθείτε ότι τα δεδομένα είναι συνεπή σε όλο τον οργανισμό

III. Αναπτύξτε το δικό σας εξειδικευμένο ETL εργαλείο για την ενσωμάτωση των δεδομένων

IV. Επιβεβαιώστε ότι τα δεδομένα που εξάγονται από τα συστήματα προέλευσης είναι τα ίδια δεδομένα που εισάγονται στην αποθήκη δεδομένων (data warehouse).

A. I

B. II και III

Γ. II, III και IV

Δ. I, II και IV

37. Ποιο από τα παρακάτω χαρακτηρίζει καλύτερα τη διαδικασία συνεχούς ολοκλήρωσης (Continuous Integration);

A. Ενσωμάτωση δεδομένων ή λειτουργιών από μια εφαρμογή σε μια άλλη.

B. Χρήση καταμεμημένων ερωτημάτων με στόχο τα δεδομένα που συλλέγονται από διάφορες εφαρμογές ή συστήματα να είναι προσβάσιμα σε πραγματικό χρόνο.

Γ. Η πρακτική της συχνής (π.χ. καθημερινά, ή πολλές φορές την ημέρα) και αυτοματοποιημένης ενσωμάτωσης δεδομένων και λειτουργιών σε μια εφαρμογή ή σύστημα με στόχο τη μείωση των σφαλμάτων και των επαναλαμβανόμενων χειροκίνητων διαδικασιών.

Δ. Η πρακτική της ενσωμάτωσης όλων των εταιρικών εφαρμογών κάθε φορά που μια νέα εφαρμογή ή σύστημα εισάγεται ή προστίθεται σε ένα δίκτυο.

38. Ποια είναι η διαφορά μεταξύ των τεχνικών ενοποίησης δεδομένων (data integration) και εικονικοποίησης δεδομένων (data virtualization);

A. Δεν υπάρχει διαφορά, είναι ταυτόσημες τεχνικές.

B. Η ενοποίηση δεδομένων μεταφέρει και μετασχηματίζει τα δεδομένα από πολλαπλές πηγές σε μια ενιαία βάση δεδομένων, ενώ η εικονικοποίηση δημιουργεί όψεις εικονικής πρόσβασης στις πρωτογενείς πηγές δεδομένων.

Γ. Η ενοποίηση δεδομένων μεταφέρει και μετασχηματίζει τα δεδομένα από πολλαπλές πηγές σε μια ενιαία βάση δεδομένων, ενώ η εικονικοποίηση αφορά την εικονική μεταφορά δεδομένων από μια βάση δεδομένων σε μια άλλη.

Δ. Δεν έχουν σχέση. Η εικονικοποίηση δεδομένων δημιουργεί οπτικές αναπαραστάσεις (visualization) από τα δεδομένα, ενώ η ενοποίηση μεταφέρει και μετασχηματίζει τα δεδομένα από πολλαπλές πηγές σε μια ενιαία βάση δεδομένων.

39. Όλα τα δεδομένα σε μια αποθήκη δεδομένων (data warehouse):

A. είναι σε έναν ενιαίο πίνακα (flat file format)

B. μπορεί να είναι κανονικοποιημένα, αλλά συνήθως δεν είναι

Γ. πρέπει να είναι τουλάχιστον σε 3η κανονική μορφή

Δ. πρέπει να είναι τουλάχιστον σε 4η κανονική μορφή

40. Πριν το φόρτωμα ενός μεγάλου όγκου δεδομένων σε μια αποθήκη δεδομένων καλό είναι να:

A. απενεργοποιούνται οι δείκτες στους πίνακες που θα φορτωθούν τα δεδομένα

B. ενεργοποιούνται οι δείκτες στους πίνακες που θα φορτωθούν τα δεδομένα

Γ. σβήσουν/διαγραφούν όλα τα παλιά δεδομένα από τους πίνακες που θα φορτωθούν τα δεδομένα

Δ. Τίποτα από τα παραπάνω

41. Τι εξυπηρετεί, μεταξύ άλλων, η ενοποίηση των δεδομένων από τα επιμέρους συστήματα ενός οργανισμού σε μια αποθήκη δεδομένων;

A. Τη συνέπεια δεδομένων μεταξύ των διαφορετικών συστημάτων.

B. Τη συμπλήρωση δεδομένων μεταξύ των διαφορετικών συστημάτων.

Γ. Τη δημιουργία μια ολοκληρωμένης και αληθούς εικόνας όλων των δεδομένων ενός οργανισμού με στόχο την απλοποίηση παραγωγής στατιστικών και αναφορών προς τη διοίκηση.

Δ. Την ανταλλαγή δεδομένων μεταξύ των διαφορετικών συστημάτων.

42. _____ ονομάζεται η περιοχή όπου προσωρινά δεδομένα αποθηκεύονται κατά την εκτέλεση μιας ροής ETL (Extract-Transform-Load) σε μια αποθήκη δεδομένων.

A. Data Bus

B. Data Staging Area

Γ. Temporary Database

Δ. Temporary Data Warehouse

43. Σε μια ροή ETL που φορτώνει δεδομένα σε μια αποθήκη δεδομένων, ποια είναι η πρακτική που ακολουθείται στην περίπτωση που οι κωδικοί (τα πρωτεύοντα κλειδιά) για τις ίδιες οντότητες από 2 διαφορετικά συστήματα δεν ταυτίζονται;

- A. Χρήση των κλειδιών από το ένα από τα 2 συστήματα με τυχαίο τρόπο.
- B. Χρήση των κλειδιών από το σύστημα που περιέχει τις πιο πρόσφατες εγγραφές.
- Γ. Χρήση των κλειδιών από το σύστημα που περιέχει τις πιο παλιές εγγραφές.
- Δ. Χρήση νέων κλειδιών αντικατάστασης που αντιστοιχίζονται στα κλειδιά των επιμέρους συστημάτων.

44. Για ποιον λόγο χρησιμοποιούμε τεχνικές καθαρισμού δεδομένων;

- A. Για την αφαίρεση θορύβου και μη αποδεκτών τιμών
- B. Για τη διόρθωση ασυνεπειών
- Γ. Για τη βελτίωση της ποιότητάς τους
- Δ. Όλα τα παραπάνω.

45. Ποια από τα παρακάτω δεν ανήκει σε τεχνικές καθαρισμού δεδομένων;

- A. Ανάκτηση Δεδομένων
- B. Διαχείριση ελλιπών τιμών
- Γ. Εύρεση διπλότυπων εγγραφών
- Δ. Αποσαφήνιση διαφορούμενων τιμών

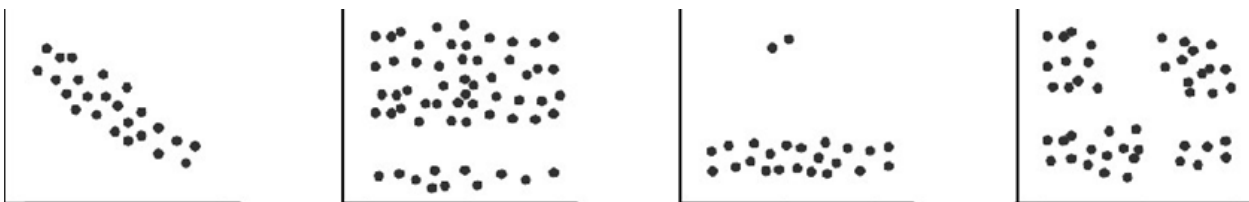
46. Ποιες από τις παρακάτω τεχνικές εφαρμόζονται στην ανάλυση δεδομένων με ελλιπείς τιμές;

- A. Imputation (Απόδοση)
- B. Interpolation (Παρεμβολή)
- Γ. Listwise Deletion (Κατά Λίστα Διαγραφή)
- Δ. Όλες οι παραπάνω

47. Σε ένα σύνολο δεδομένων με 100 εγγραφές και 10 πεδία, το 10% κάθε πεδίου έχει τιμές που λείπουν. Ποιος είναι ο μεγαλύτερος και μικρότερος αριθμός εγγραφών που μπορούν να διαγραφούν αν ακολουθηθεί η τεχνική Listwise Deletion;

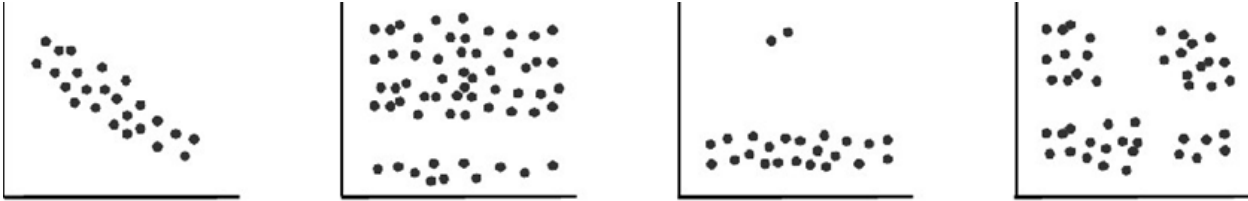
- A. 100 και 10 εγγραφές, αντίστοιχα
- B. 10 και 10 εγγραφές, αντίστοιχα
- Γ. 100 και 0 εγγραφές, αντίστοιχα
- Δ. 100 και 100 εγγραφές, αντίστοιχα

48 Έχετε στη διάθεσή σας τα ακόλουθα σκόρπια (scatter) διαγράμματα. Ποιο από τα παρακάτω οπτικοποιεί μια τάση στα δεδομένα;



- A. Το πρώτο από αριστερά
- B. Το δεύτερο από αριστερά
- Γ. Το τρίτο από αριστερά
- Δ. Το τέταρτο από αριστερά

49. Έχετε στη διάθεσή σας τα ακόλουθα σκόρπια (scatter) διαγράμματα. Ποιο από τα παρακάτω οπτικοποιεί συστάδες (clusters) στα δεδομένα;



- A. Το πρώτο και το δεύτερο από αριστερά
- B. Το δεύτερο, τρίτο και τέταρτο από αριστερά
- Γ. Το τρίτο και το τέταρτο από αριστερά
- Δ. Το τέταρτο από αριστερά

50. Θέλετε να βρείτε με οπτικό τρόπο το είδος της συσχέτισης μεταξύ 2 αριθμητικών χαρακτηριστικών (π.χ. εισόδημα φορολογούμενου σε σχέση με την ηλικία του) σε ένα σύνολο δεδομένων. Ποιο είδος διαγράμματος είναι το πιο κατάλληλο να χρησιμοποιήσετε;

- A. Scatter Diagram
- B. Bar Chart
- Γ. Pie Chart
- Δ. Line Chart

51. Θέλετε να συγκρίνετε με οπτικό τρόπο τον μέσο όρο της τιμής ενοικίων σε πολλές (>10) διαφορετικές περιοχές της Αττικής. Ποιο είδος διαγράμματος είναι το πιο κατάλληλο να χρησιμοποιήσετε;

- A. Scatter Diagram
- B. Bar Chart
- Γ. Pie Chart
- Δ. Line Chart

52. Θέλετε να εξερευνήσετε με οπτικό τρόπο τη συσχέτιση μεταξύ 2 κατηγορικών χαρακτηριστικών (π.χ. το εισόδημα φορολογουμένων ανά δήμο της Αττικής και ανά μορφωτικό επίπεδο). Ποιο είδος διαγράμματος είναι το πιο κατάλληλο να χρησιμοποιήσετε;

- A. Scatter Diagram
- B. Heat Map
- Γ. Bar Chart
- Δ. Line Chart

53. Η ομαδοποίηση k-μέσων είναι δημοφιλής στην ανάλυση συστάδων (κλάδος της εξόρυξης δεδομένων). Η ομαδοποίηση αυτή έχει ως στόχο να διαχωρίσει ένα σύνολο δεδομένων (παρατηρήσεις) σε συστάδες, έτσι ώστε:

- A. κάθε παρατήρηση να ανήκει σε πολλαπλές κοντινές συστάδες

Β. κάθε παρατήρηση να ανήκει σε μια συστάδα με το κοντινότερο μέσο, το οποίο χρησιμεύει ως ένα χαρακτηριστικό δείγμα της συστάδας.

Γ. ο αριθμός των συστάδων να υπολογίζεται αυτόματα από τον αλγόριθμο

Δ. ο αριθμός των συστάδων και των παρατηρήσεων να μην επηρεάζουν την ταχύτητα του αλγορίθμου.

54. Η διαδικασία «_____» στοχεύει στον εντοπισμό στοιχείων μέσα σε ένα σύνολο δεδομένων που αποκλίνουν από τη στατιστική συμπεριφορά της πλειοψηφίας των δεδομένων.

A. Ανάλυση ακραίων τιμών (Outlier Analysis)

B. Πρόβλεψη (Prediction)

Γ. Κατηγοριοποίηση (Classification)

Δ. Εύρεση Μοτίβων (Pattern Extraction)

55. Τι είναι ένα δέντρο αποφάσεων (Decision Tree);

A. Ένα διάγραμμα ροής για την ενσωμάτωση δεδομένων.

B. Μια δενδρική δομή που κάθε εσωτερικός κόμβος αντιστοιχεί σε μια λογική συνθήκη σε ένα γνώρισμα των δεδομένων και κάθε φύλλο αντιστοιχεί σε μια κατηγοριοποίηση των δεδομένων.

Γ. Μια δενδρική δομή που κάθε εσωτερικός κόμβος αντιστοιχεί σε μια κατηγοριοποίηση των δεδομένων και κάθε φύλλο σε μια λογική συνθήκη πάνω σε ένα γνώρισμα των δεδομένων.

Δ. Ένας γράφος για την κατηγοριοποίηση δεδομένων.

56. Που χρησιμοποιείται ένα δέντρο αποφάσεων (Decision Tree);

A. Στην ανίχνευση ελλιπών τιμών στα δεδομένα.

B. Στην ενοποίηση δεδομένων.

Γ. Στην παλινδρόμηση δεδομένων.

Δ. Στην κατηγοριοποίηση δεδομένων.

57. Σε ποια από τις παρακάτω εργασίες θα χρησιμοποιούσατε διασταυρούμενη επικύρωση (cross validation);

A. Επιλογή χαρακτηριστικών που θα συμπεριληφθούν σε ένα μοντέλο

B. Αξιολόγηση και Σύγκριση προγνωστικών μοντέλων

Γ. Επιλογή παραμέτρων στη συνάρτηση πρόβλεψης

Δ. Όλα τα παραπάνω

58. Ποιο από τα παρακάτω δεν είναι αλγόριθμος μηχανικής μάθησης;

A. SVG

B. SVM

Γ. Random forest

Δ. Decision trees

59. Η εκμάθηση ενός μοντέλου που πραγματοποιείται από επισημασμένα δεδομένα (labelled data) ονομάζεται

A. Supervised learning

- B. Unsupervised learning
- Γ. Reinforcement learning
- Δ. Missing data imputation

60. Η ανάλυση συσχέτισης (correlation analysis) σε χαρακτηριστικά των δεδομένων χρησιμοποιείται για:

- A. Χειρισμό τιμών που λείπουν.
- B. Προσδιορισμό περιττών χαρακτηριστικών.
- Γ. Χειρισμό διαφορετικών μορφών δεδομένων.
- Δ. Εξάλειψη ακραίων τιμών.