

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ – ΣΧΟΛΗ ΗΜΜΥ
ΜΑΘΗΜΑ: ΠΛΗ303 - ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ
Διδάσκων: Καθ. Αντώνιος Δεληγιαννάκης
Εμβόλιμη Εξεταστική Σεπτεμβρίου 2023
Διάρκεια: 2 ώρες και 20 λεπτά

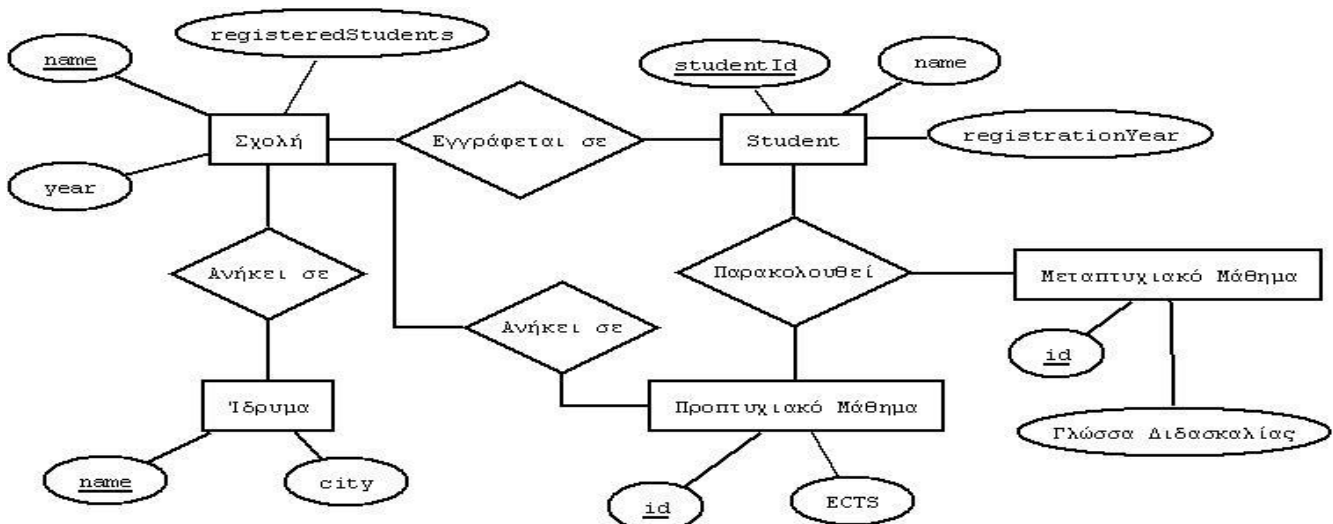
ΔΕΝ ΕΠΙΤΡΕΠΕΤΑΙ Η ΧΡΗΣΗ ΚΙΝΗΤΩΝ ΤΗΛΕΦΩΝΩΝ – ΚΑΛΗ ΕΠΙΤΥΧΙΑ!

1. ER διάγραμμα (28 μονάδες)

Σας ζητείται να σχεδιάσετε το ER διάγραμμα για μία εφαρμογή φοιτητολογίου. Στην εφαρμογή καταγράφονται στοιχεία σχετικά με Ιδρύματα (ΕΜΠ, Πολυτεχνείο Κρήτης, κτλ), Σχολές, Φοιτητές, καθώς και για τα μαθήματα που έχουν παρακολουθήσει οι φοιτητές:

- **Ίδρυμα:** Κάθε ίδρυμα χαρακτηρίζεται μοναδικά από το όνομά του (name). Για κάθε ίδρυμα διατηρούμε επιπλέον πληροφορία για την πόλη (city) στην οποία είναι η πρυτανεία του.
- **Σχολή:** Κάθε Σχολή χαρακτηρίζεται μοναδικά από το συνδυασμό του ονόματός της (name) και του ιδρύματος στο οποίο ανήκει. Για κάθε Σχολή διατηρούμε επιπλέον πληροφορία για το έτος ίδρυσής της (year), καθώς και για το συνολικό αριθμό φοιτητών (registeredStudents) που έχουν εγγραφεί σε αυτήν (σε οποιοδήποτε έτος). Όλες οι εγγραφές φοιτητών σε Σχολές είναι καταγεγραμμένες στη Βάση Δεδομένων της εφαρμογής.
- Κάθε Ίδρυμα περιέχει 1 ή περισσότερες Σχολές. Κάθε Σχολή ανήκει σε ακριβώς 1 Ίδρυμα.
- **Φοιτητής (Student):** Κάθε φοιτητής χαρακτηρίζεται μοναδικά από το συνδυασμό του κωδικού φοιτητή (studentId) και του ονόματός του (name). Κάθε φοιτητής έχει εγγραφεί σε 1 ή περισσότερες Σχολές στη ζωή του. Σε κάθε Σχολή έχουν εγγραφεί 0 ή περισσότεροι φοιτητές. Για κάθε φοιτητή διατηρούμε και πληροφορία (registrationYear) για το έτος κατά το οποίο έκανε την 1^η του εγγραφή σε κάθε Σχολή στην οποία έχει φοιτήσει.
- Κατά τη διάρκεια των σπουδών του, κάθε φοιτητής παρακολουθεί 1 ή περισσότερα προπτυχιακά μαθήματα και 0 ή 1 μεταπτυχιακά μαθήματα. Κάθε προπτυχιακό ή μεταπτυχιακό μάθημα παρακολουθείται από 0 ή περισσότερους φοιτητές.
- Κάθε **προπτυχιακό μάθημα** χαρακτηρίζεται μοναδικά από το αναγνωριστικό του (id). Για κάθε προπτυχιακό μάθημα διατηρούμε και πληροφορία για τον αριθμό των ECTS του.
- Κάθε **μεταπτυχιακό μάθημα** χαρακτηρίζεται μοναδικά από το αναγνωριστικό του (id). Για κάθε μεταπτυχιακό μάθημα διατηρούμε και πληροφορία για τη γλώσσα διδασκαλίας του.
- Σε κάθε Σχολή ανήκουν 0 ή περισσότερα μαθήματα. Κάθε μάθημα (προπτυχιακό ή μεταπτυχιακό) ανήκει σε 1 ακριβώς Σχολή.

Α. Διορθώστε το παρακάτω ER διάγραμμα, το οποίο ίσως περιέχει κάποια λάθη.



2. SQL + Σχεσιακή Άλγεβρα (51 μονάδες: 13 + 13 + 13 + 12)

Θεωρήστε το ακόλουθο σχεσιακό σχήμα που περιγράφει Φοιτητές, Μαθήματα, καθώς και εγγραφές φοιτητών σε μαθήματα:

Course(id, ects) // Αναγνωριστικό Μαθήματος, Αριθμός ECTS

Takes(studentid, courseid) // Αναγνωριστικό Φοιτητή, Αναγνωριστικό Μαθήματος

Student(id, name, age) // Αναγνωριστικό Φοιτητή, Όνομα Φοιτητή, Ηλικία

Προσέξτε ότι σε ένα μάθημα μπορεί να μην έχει εγγραφεί κάποιος φοιτητής, αλλά κάθε φοιτητής έχει εγγραφεί σε τουλάχιστον ένα μάθημα.

A. Δώστε τα ακόλουθα ερωτήματα σε SQL:

- Υπολογίστε τον αριθμό των εγγεγραμμένων φοιτητών ανά μάθημα. Προσέξτε ότι πρέπει να συμπεριλάβετε και μαθήματα στα οποία δεν έχει εγγραφεί κάποιος φοιτητής.
- Υπολογίστε όλα τα στοιχεία των μαθημάτων στα οποία δεν έχει εγγραφεί κάποιος φοιτητής με ηλικία μεγαλύτερη των 20 ετών.
- Τυπώστε το αναγνωριστικό όλων των φοιτητών οι οποίοι έχουν το μεγαλύτερο (ανάμεσα σε όλους τους φοιτητές) άθροισμα ects στα μαθήματα στα οποία έχουν εγγραφεί.

B. Δώστε το ακόλουθο ερώτημα σε έκφραση σχεσιακής άλγεβρας:

- Τυπώστε τα στοιχεία όλων των μαθημάτων στα οποία έχει εγγραφεί τουλάχιστον 1 φοιτητής ηλικίας μικρότερης των 20 ετών και τουλάχιστον 1 φοιτητής ηλικίας μεγαλύτερης των 24 ετών (προσέξτε ότι πρέπει να ικανοποιούνται και οι 2 αυτές συνθήκες για το μάθημα).

3. Βελτιστοποίηση Επερωτήσεων (21 μονάδες: 12 + 9)

Θα υπολογίσουμε το κόστος του πλάνου στα δεξιά, πάνω στους πίνακες: $R(A, B, Y)$, $S(\Gamma, \Delta)$.

Θεωρήστε ότι έχετε τα ακόλουθα στατιστικά:

$T(R) = 20000$, $B(R) = 1000$

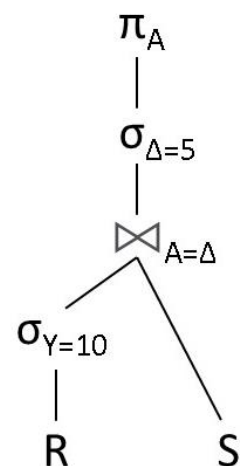
$T(S) = 600$, $B(S) = 20$

και γνωρίζετε ότι:

- το γνώρισμα $R.Y$ λαμβάνει 50 ακέραιες τιμές,
- το γνώρισμα $R.A$ λαμβάνει 10 ακέραιες τιμές,
- το γνώρισμα $S.\Delta$ λαμβάνει 10 ακέραιες τιμές,
- όλα τα γνωρίσματα των σχέσεων R και S έχουν το ίδιο μέγεθος.

Θεωρήστε ότι υπάρχει μη-συσταδοποιημένο hash ευρετήριο στο $S.\Delta$.

Θεωρήστε ότι υπάρχει συσταδοποιημένο B+-tree ευρετήριο στο $R.Y$.



A. Υπολογίστε το κόστος του πλάνου της εικόνας αν για το Join χρησιμοποιήσετε index nested loop join.

B. Υπολογίστε το κόστος του πλάνου της εικόνας αν για το Join χρησιμοποιήσετε block nested loop join και γνωρίζετε ότι έχετε διαθέσιμη μνήμη ίση με 6 σελίδες για το join.