

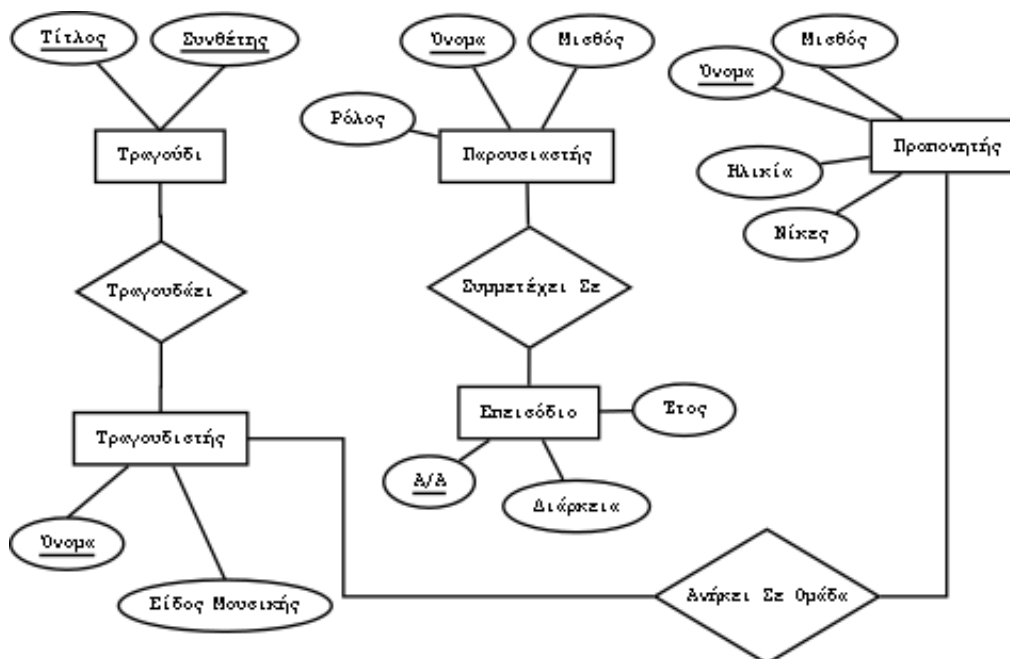
**ΔΕΝ ΕΠΙΤΡΕΠΕΤΑΙ Η ΧΡΗΣΗ ΚΙΝΗΤΩΝ ΤΗΛΕΦΩΝΩΝ - ΚΑΛΗ ΕΠΙΤΥΧΙΑ!**

**1. ER διάγραμμα (20 μονάδες)**

Σας ζητείται να σχεδιάσετε το ER διάγραμμα για την ακόλουθη εφαρμογή καταγραφής στοιχείων για ένα τηλεπαιχνίδι για την ανακάλυψη του νέου μουσικού ταλέντου στην Ελλάδα:

- **Μέλη Παραγωγής:** Κάθε μέλος της παραγωγής χαρακτηρίζεται μοναδικά από το όνομά του. Για κάθε μέλος της παραγωγής κρατάμε επιπλέον πληροφορία για το μισθό του. Τα μέλη της παραγωγής διακρίνονται σε **Παρουσιαστές** και σε **Προπονητές**.
- Για κάθε προπονητή κρατάμε πληροφορία και για την ηλικία του, καθώς και για το συνολικό αριθμό φορών που παίκτης του έχει κερδίσει το τηλεπαιχνίδι σε προηγούμενες χρονιές.
- Για κάθε παρουσιαστή κρατάμε πληροφορία και για το ρόλο του στο τηλεπαιχνίδι (πχ, βασικός παρουσιαστής, παρασκήνια, συνεντεύξεις κτλ).
- Κάθε παρουσιαστής συμμετέχει σε 0 ή περισσότερα επεισόδια του τηλεπαιχνιδιού. Σε κάθε επεισόδιο συμμετέχουν 1 ή περισσότεροι παρουσιαστές.
- **Επεισόδια:** Κάθε επεισόδιο του τηλεπαιχνιδιού χαρακτηρίζεται μοναδικά από το συνδυασμό του αύξοντα αριθμού (Α/Α) του και του έτους προβολής του, Για κάθε επεισόδιο κρατάμε πληροφορία και για τη διάρκειά του.
- **Τραγουδιστής:** Κάθε Τραγουδιστής χαρακτηρίζεται μοναδικά από το όνομά του. Για κάθε τραγουδιστή κρατάμε πληροφορία και για το αγαπημένο του είδος μουσικής.
- **Τραγούδι:** Κάθε Τραγούδι που ακούγεται στο τηλεπαιχνίδι χαρακτηρίζεται μοναδικά από το συνδυασμό του ονόματός του και του ονόματος του συνθέτη του.
- Ο κάθε τραγουδιστής μπορεί να τραγουδάει 0 ή περισσότερα τραγούδια σε κάθε επεισόδιο. Σε κάθε επεισόδιο τραγουδάνε 1 ή περισσότεροι τραγουδιστές. Ένα τραγούδι μπορεί να έχει ακουστεί σε 0 ή περισσότερα επεισόδια.
- Ο κάθε Προπονητής είναι υπεύθυνος για μία ακριβώς ομάδα 16 το πολύ Τραγουδιστών. Κάθε τραγουδιστής ανήκει στην ομάδα ακριβώς ενός προπονητή.

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



**2. SQL (40 μονάδες: 8 + 10 + 10 + 12)**

Για μία εφαρμογή παρόμοια με το Ερώτημα 1, θεωρήστε το ακόλουθο σχεσιακό σχήμα που περιγράφει Τραγουδιστές, Τραγούδια, καθώς και τραγούδια που έχουν τραγουδίσει οι τραγουδιστές:

**Song(title, duration)**

// Τίτλος, Διάρκεια σε δευτερόλεπτα

**Sings(singerName, songName)**

// Όνομα Τραγουδιστή, Τίτλος Τραγουδιού

**Singer(name, musicType, age)**

// Όνομα Τραγουδιστή, Είδος Μουσικής, Ηλικία

Προσέξτε ότι ένας τραγουδιστής μπορεί να μην έχει τραγουδίσει κανένα τραγούδι, και αντίστοιχα μπορεί κάποιο τραγούδι να μην έχει τραγουδιστεί από κανένα τραγουδιστή.

Δώστε τα ακόλουθα ερωτήματα σε SQL:

1. Τυπώστε για κάθε μουσικό τη μέγιστη διάρκεια τραγουδιού που έχει τραγουδίσει.
2. Τυπώστε τη μέση ηλικία όλων των τραγουδιστών που έχουν τραγουδίσει είτε το τραγούδι «Δυνατά», είτε το τραγούδι «Αγριολούλουδο».
3. Τυπώστε όλα τα στοιχεία των τραγουδιών που δεν έχει τραγουδήσει κανένας τραγουδιστής μικρότερος των 20 ετών.
4. Τυπώστε όλα τα στοιχεία των τραγουδιστών που το **είδος της μουσικής τους** είναι το Λαϊκό Τραγούδι και που τα τραγούδια που έχουν τραγουδήσει **έχουν (αθροιστικά) τη μέγιστη διάρκεια χρόνου**.  
Προσέξτε ότι η διάρκεια ενός τραγουδιού δίνεται από το γνώρισμα duration και πρέπει για κάθε τραγουδιστή να συνυπολογίσετε όλα τα τραγούδια που έχει τραγουδίσει.

### 3. Βελτιστοποίηση Ερωτήσεων (20 μονάδες – 6 + 14)

Θα υπολογίσουμε το κόστος του πλάνου στα δεξιά, πάνω στους πίνακες:

$R(A,B,E)$ ,  $S(\Gamma,\Delta)$ .

Θεωρήστε ότι έχετε τα ακόλουθα στατιστικά:

$T(R) = 3000$ ,  $B(R) = 300$

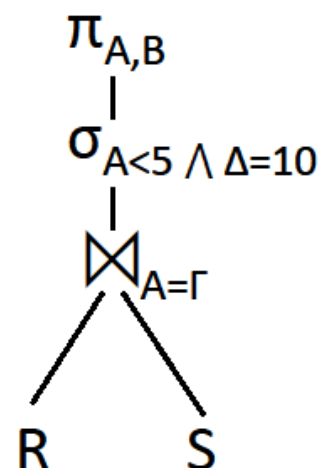
$T(S) = 1000$ ,  $B(S) = 200$

και γνωρίζετε ότι:

- το γνώρισμα  $R.A$  λαμβάνει ακέραιες τιμές από το 1 έως και το 10, ενώ
- το γνώρισμα  $S.\Delta$  λαμβάνει ακέραιες τιμές από το 0 έως και το 20.
- το γνώρισμα  $R.A$  είναι **ξένο κλειδί** στο γνώρισμα  $S.\Gamma$
- όλα τα γνώρισμα 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  $R.A$ .



A. Εφαρμόστε τους **Ευριστικούς Κανόνες Βελτιστοποίησης** για να μετατρέψετε το πλάνο εκτέλεσης σε ένα πιο αποδοτικό πλάνο.

B. Υπολογίστε το κόστος όλου του πλάνου που δημιουργήσατε στο προηγούμενο υποερώτημα αν για το Join χρησιμοποιήσετε **block nested loop join** και γνωρίζετε ότι έχετε διαθέσιμη μνήμη ίση με 20 σελίδες για το join.

### 4. Κανονικοποίηση (20 μονάδες)

Έστω η σχέση  $V(A, B, C, D)$  για την οποία σας δίνεται ότι ισχύουν οι ακόλουθες συναρτησιακές εξαρτήσεις:  
 $F = \{ A \rightarrow C, B \rightarrow D, D \rightarrow B, C \rightarrow A \}$ .

- 1) **(5 μονάδες)** Βρείτε τα κλειδιά της σχέσης  $V$
- 2) **(3 μονάδες)** Ανήκει η  $V$  στην 3NF; Εξηγήστε αναλυτικά την απάντησή σας.
- 3) **(12 μονάδες)** Εφαρμόστε τον αλγόριθμο αποσύνθεσης σε BCNF για να λάβετε μία χωρίς απώλειες αποσύνθεση της  $V$  σε ένα σύνολο από σχέσεις που είναι σε BCNF. Δείξτε αναλυτικά τη διαδικασία που ακολουθήσατε.

Η απάντησή σας πρέπει να περιλαμβάνει τους πίνακες, με τα χαρακτηριστικά τους και μία ένδειξη για τα κλειδιά κάθε πίνακα (με υπογράμμιση), καθώς και αιτιολόγηση για το πώς προέκυψε το κλειδί του κάθε πίνακα.

**ΣΗΜΕΙΩΣΗ:** Ζητείται συγκεκριμένος αλγόριθμος παραπάνω. Η εφαρμογή άλλου αλγόριθμου (πχ, σύνθεσης σε 3NF) οδηγεί αυτόματα σε μηδενισμό του συγκεκριμένου κομματιού.