

### Άσκηση 6 (ΑΠΟΘΗΚΕΣ ΔΕΔΟΜΕΝΩΝ)

Η διοίκηση μιας ακαδημαϊκής βιβλιοθήκης ενδιαφέρεται να αναπτύξει μια αποθήκη δεδομένων, ώστε να είναι σε θέση να αναλύει τα δεδομένα και να παράγει στατιστικές αναφορές σχετικά με την πορεία των δανεισμών και την εξυπηρέτηση των χρηστών. Τα δεδομένα που θα τροφοδοτούν την αποθήκη είναι διαθέσιμα υπό την μορφή ενός γραμμογραφημένου αρχείου κάθε εγγραφή του οποίου αποτελείται από 14 πεδία.

Οι απαιτήσεις της διοίκησης εστιάζουν στην ανάλυση του αριθμού των δανεισμών ανά τοποθεσία (copyloc), είδος τεκμηρίου (material), τμήμα (department) και φύλο (sex) δανειζομένων, καθώς και οποιονδήποτε συνδυασμό αυτών. Εξυπακούεται ότι στην ανάλυση των δεδομένων θα πρέπει να ληφθεί υπόψη ο παράγοντας του χρόνου έτσι ώστε, η διοίκηση της βιβλιοθήκης να είναι σε θέση να παράγει στατιστικές αναφορές με τα στοιχεία των δανεισμών ανα ημέρα, εβδομάδα, μήνα και τρίμηνο για το διάστημα ενός ή περισσότερων ετών.

Καλείστε να σχεδιάσετε και να υλοποιήσετε την αποθήκη δεδομένων της βιβλιοθήκης προκειμένου να αυξήσετε την αποτελεσματικότητα διεξαγωγής χρήσιμων στατιστικών στοιχείων, μειώνοντας ταυτόχρονα τον χρόνο εκτέλεσης των επερωτήσεων. Στην συνέχεια να τροφοδοτήσετε την αποθήκη με δεδομένα και να εκτελέσετε ορισμένες επερωτήσεις για την παραγωγή χρήσιμων στατιστικών αναφορών. Την αποθήκη δεδομένων θα την υλοποιήσετε με την χρήση του DBMS SQL Server. Για την παρουσίαση των αποτελεσμάτων θα αξιοποιήσετε τις δυνατότητες του Power BI.

Ακολουθεί αναλυτική περιγραφή των δεδομένων και των ζητούμενων της άσκησης.

Το αρχείο **libdata.txt** περιέχει 91.640 εγγραφές. Κάθε εγγραφή αποτελείται από 14 πεδία τα οποία διαχωρίζονται με τον χαρακτήρα "|". Ακολουθεί η περιγραφή των πεδίων.

Libdata.txt	
<b>bibno</b>	Κωδικός βιβλιογραφικής εγγραφής
<b>title</b>	Τίτλος του τεκμηρίου που περιγράφει η βιβλιογραφική εγγραφή
<b>material</b>	Είδος τεκμηρίου (book, serial, audiovisual, paper κ.λπ.)
<b>lang</b>	Γλώσσα τεκμηρίου (gre, eng, fre, κλπ)
<b>place</b>	Τόπος έκδοσης
<b>copyno</b>	Κωδικός αντιτύπου.
<b>copyloc</b>	Τριψήφιος Κωδικός που δηλώνει την τοποθεσία του αντιτύπου. Π.χ. "ANA" σημαίνει "Αναγνωστήριο".
<b>bid</b>	Κωδικός μέλους
<b>bname</b>	Ονοματεπώνυμο μέλους
<b>sex</b>	Φύλο (F=female, M=Male)
<b>depcode</b>	Κωδικός τμήματος
<b>depname</b>	Όνομα τμήματος
<b>lid</b>	Κωδικός συναλλαγής δανεισμού.
<b>loandate</b>	Ημερομηνία δανεισμού

### Ζήτημα Πρώτο

1. Να εκτελέσετε το SQL script **createLIBDW.sql** το οποίο:
  - a. Θα δημιουργήσει την βάση δεδομένων LIBDW (Library Data Warehouse).
  - b. Θα δημιουργήσει τον πίνακα libdata
  - c. Θα φορτώσει στον πίνακα libdata τα δεδομένα του αρχείου libdata.txt.
2. Να υλοποιήσετε το λογικό σχήμα της αποθήκης δεδομένων το οποίο θα πρέπει να έχει την μορφή αστερά (Star Schema).
3. Να γράψετε κατάλληλες εντολές σε γλώσσα SQL, οι οποίες θα τροφοδοτούν το σχήμα της αποθήκης με τα απαραίτητα στοιχεία από τον πίνακα libdata.
4. Να αναπαραστήσετε διαγραμματικά το σχήμα της αποθήκης χρησιμοποιώντας την επιλογή "Database Diagrams" του Sql Server Management Studio.

Η δημιουργία του λογικού σχήματος (2) και η τροφοδότηση της αποθήκης με τα δεδομένα (3) θα γίνουν με την εκτέλεση ενός SQL script το οποίο θα πρέπει να γράψετε.

### Ζήτημα Δεύτερο

Χρησιμοποιώντας την αποθήκη δεδομένων που δημιουργήσατε στο προηγούμενο ζήτημα, να γράψετε και να εκτελέσετε επερωτήσεις σε γλώσσα SQL, οι οποίες να απαντούν στα ακόλουθα ερωτήματα (απαιτήσεις) της διοίκησης της βιβλιοθήκης:

1. Εμφανίστε έναν κατάλογο με τον αριθμό των δανεισμών ανα τμήμα (department) και φύλο (sex).
2. Εμφανίστε έναν κατάλογο με τον αριθμό των δανεισμών ανα τοποθεσία (copyloc) και είδος τεκμηρίου (material).
3. Εμφανίστε έναν κατάλογο με τον αριθμό των δανεισμών ανα μήνα και φύλο για το έτος 2000.
4. Η διοίκηση της βιβλιοθήκης θέλει μια αναφορά που θα περιέχει **μόνο** τις ακόλουθες πληροφορίες.
  - a. Τον συνολικό αριθμό των δανεισμών
  - b. Το αριθμό των δανεισμών ανα έτος
  - c. Τον αριθμό των δανεισμών ανα έτος και τρίμηνο
  - d. Τον αριθμό των δανεισμών ανα έτος, τρίμηνο και μήνα.Γράψτε **μία μόνο** επερώτηση σε γλώσσα SQL η οποία να παράγει την παραπάνω αναφορά.
5. Γράψτε μια επερώτηση σε γλώσσα SQL το αποτέλεσμα της οποίας είναι η δημιουργία ενός κύβου (data cube), κάθε κελί του οποίου περιέχει τον αριθμό των δανεισμών ανα έτος και τμήμα (department).

**ΠΡΟΣΟΧΗ:** Για τα ερωτήματα **1 έως και 5** δεν επιτρέπεται η δημιουργία και χρήση όψεων.

## **Ζήτημα Τρίτο**

### **Ζήτημα 3**

Δημιουργήστε μια αναφορά (report) με το POWER BI στην οποία θα απεικονίσετε με κατάλληλα γραφήματα (ή και πίνακες) τα αποτελέσματα των ερωτήσεων που γράψατε στο προηγούμενο ζήτημα. Για κάθε μία ερώτηση να επιλέξετε έναν τύπο γραφήματος το οποίο θεωρείτε κατάλληλο. Μπορείτε εφόσον το επιθυμείτε να απεικονίσετε τα στοιχεία κάποιας ερώτησης με περισσότερα από ένα γραφήματα (ή και πίνακες).

**Σημείωση:** Εξυπακούεται ότι θα χρησιμοποιήσετε το start σχήμα που δημιουργήσατε στον SQL SERVER.

### **ΠΑΡΑΔΟΤΕΑ**

Τα παραδοτέα της εργασίας σας θα είναι τα εξής:

1. Ένα αρχείο pdf με όνομα ΕΠΩΝΥΜΟ.pdf το οποίο θα περιέχει:

#### **Παραδοτέα πρώτου ζητήματος**

- Τον κώδικα (εντολές SQL) για την δημιουργία του λογικού σχήματος της αποθήκης δεδομένων και την εισαγωγή των εγγραφών στους αντίστοιχους πίνακες.
- Το διάγραμμα του σχήματος αστέρα της αποθήκης δεδομένων.

#### **Παραδοτέα δεύτερου ζητήματος**

- Τον κώδικα με τις ερωτήσεις SQL του δεύτερου ζητήματος.

#### **Παραδοτέα τρίτου ζητήματος**

- Την αναφορά (report) που δημιουργήσατε στο PowerBI. Μπορείτε να κάνετε print screen τα γραφήματα που δημιουργήσατε και να τα παραθέσετε στο αρχείο pdf.

2. Το αρχείο .pbix που δημιουργήσατε με το power bi. Το όνομα του αρχείου να είναι της μορφής: ΕΠΩΝΥΜΟ.pbix

Τα δύο παραπάνω αρχεία (PDF, και pbix) θα τα τοποθετήσετε σε έναν φάκελο τον οποίο στην συνέχεια θα τον συμπιέσετε σε ένα αρχείο της μορφής επώνυμο.zip και θα το ανεβάσετε στο moodle.

**Προσοχή:** Μην ξεχάσετε στην αρχή του αρχείου pdf να γράψετε το ονοματεπώνυμό σας.

**ΚΑΛΗ ΕΠΙΤΥΧΙΑ!!!**