

1^η εργασία – PL/SQL, Ευρετήρια και Optimizer

Στόχος των εργασιών είναι η εξοικείωση με θεωρητικά και πρακτικά προβλήματα των Βάσεων Δεδομένων, μέσα από χρηστικά παραδείγματα. Στην πρώτη εργασία θα ασχοληθούμε α) με τη χρήση PL/SQL για το χειρισμό εγγραφών της ΒΔ, β) με τις εντολές εμφάνισης του πλάνου εκτέλεσης, τις πληροφορίες που μας δίνουν και τον τρόπο με τον οποίο μπορούμε να επιταχύνουμε την εκτέλεση ενός ερωτήματος.

Άσκηση 1^η – Δημιουργία και γέμισμα σχήματος

Στη ΒΔ σας θα πρέπει να έχετε δύο σχέσεις **Customers** και **Orders**, δεδομένα για τις οποίες μπορείτε να αντιγράψετε από τη ΒΔ **XSALES** και συγκεκριμένα από τους πίνακες **customers**, **orders** και **order_items**.

Ανεξάρτητα από το αρχικό σχήμα, οι σχέσεις θα πρέπει να έχουν την ακόλουθη δομή:

Η customers:

ID	NAME	GENDER	AGE	MARITAL_STATUS	INCOME_LEVEL	CREDIT_LIMIT
1	Abigail Kessel	Male	71 (null)		G: 130,000 - 149,999	9000
2	Anne Koch	Female	60 (null)		I: 170,000 - 189,999	10000
3	Buick Emmerson	Male	78 (null)		B: 30,000 - 49,999	1500
4	Frank Hardy	Male	83 (null)		B: 30,000 - 49,999	1500
5	Macklin Gowen	Male	48 (null)		B: 30,000 - 49,999	1500
6	Harriett Charles	Female	92 single		A: Below 30,000	1500
7	Linette Ingram	Female	31 single		B: 30,000 - 49,999	1500
8	Vida Puleo	Female	53 married		C: 50,000 - 69,999	3000
9	Rosamond Krider	Female	81 divorced		C: 50,000 - 69,999	3000
10	Gertrude Atkins	Female	70 widow		D: 70,000 - 89,999	3000

Το πεδίο age θα προκύψει από το birth_date και την τρέχουσα ημερομηνία.

Η orders:

ORDER_ID	CUSTOMER_ID	PRICE	COST	PROFIT	PROFITLEVEL
15126	9163	13176.98	10328.06	2848.92	Medium
3812	1824	8590.36	6917.98	1672.38	Medium
7837	3473	9871.83	8029.11	1842.72	Medium
507	207	4878.62	4001.96	876.66	Low
7587	3363	8908.68	7465.94	1442.74	Medium
14925	8878	14548.59	11533.47	3015.12	Medium
11742	5873	5060.77	4212.08	848.69	Low
3809	1823	7368.21	5828.61	1539.6	Medium
10612	5028	12689.28	10338.63	2350.65	Medium
2139	1072	5927.88	4543.62	1384.26	Medium

Τα πεδία price, cost, profit έχουν προκύψει από τη συνένωση του order με τον order_items (με τα προϊόντα κάθε παραγγελίας) και περιέχουν αντίστοιχα τα συνολικά έσοδα, το συνολικό κόστος και το καθαρό κέρδος (έσοδα-κόστος) για κάθε διαφορετική παραγγελία.

Το πεδίο profitlevel έχει τιμές Low, Medium, High για profit<1000, profit<5000 και profit>=5000 αντίστοιχα.

Τα πιο πάνω θα πρέπει να γίνουν με SQL ή PL-SQL και θα πρέπει να δώσετε το script με το οποίο τα δημιουργήσατε.

Άσκηση 2^η – Στατιστικά στοιχεία

Υπολογίστε:

- α) τον αριθμό των πλειάδων κάθε πίνακα,
- β) τον καταμερισμό των πλειάδων στις διαφορετικές τιμές των πεδίων age, credit_limit και income_level του πίνακα CUSTOMERS, το εύρος τιμών του πεδίου profit του πίνακα ORDERS. Δώστε τις εντολές που τρέξατε για κάθε υποερώτημα και τα αποτελέσματα.
- γ) Προσθέστε στην customers τρεις επιπλέον στήλες total_money_spend, total_orders_made, total_profit που να περιέχει τα συνολικά χρήματα που ξόδεψε σε παραγγελίες κάθε πελάτης, το πλήθος των παραγγελιών που έκανε και το συνολικό profit για την εταιρία από αυτές. Αν δεν υπάρχουν παραγγελίες από κάποιο πελάτη τότε οι τρεις στήλες θα περιέχουν 0.
- δ) Με ποιο ερώτημα θα βρείτε αν υπάρχουν πελάτες που έχουν δαπανήσει κατά μέσο όρο πάνω από 10000 ευρώ στις παραγγελίες τους;
- ε) Προσθέστε στον customers μια επιπλέον στήλη coupon που θα υπολογίζει ένα εκπτωτικό κουπόνι για επόμενη αγορά κάθε πελάτη, ανάλογα με την κατανάλωση που έχει κάνει και το profit που είχαμε από αυτόν. Σε πελάτες με πάνω από 3 αγορές (>3) ή πάνω από 30 χιλιάδες συνολικές δαπάνες το κουπόνι θα είναι 10% του total_profit που είχαμε από αυτόν. Σε πελάτες με μέσο όρο αγορών που ξεπερνά τις 5000 ευρώ το κουπόνι θα είναι 5% του total_profit.
- ζ) Ποιο το συνολικό τελικό profit για την εταιρία αν όλοι οι πελάτες χρησιμοποιήσουν τα κουπόνια τους;

Δώστε τις εντολές με τις οποίες υλοποιήσατε το κάθε ερώτημα και ένα screenshot με τα αποτελέσματα.

Άσκηση 3^η – Βελτιστοποίηση ερωτήματος

ΕΡΩΤΗΜΑ 1: Επερώτηση ακριβούς ταιριάσματος με ευρετήριο

Χρησιμοποιώντας την εντολή EXPLAIN ελέγξτε πώς λειτουργεί ο optimizer για το ακόλουθο ερώτημα

```
SELECT c.name, o.price, o.profit
FROM customers c JOIN orders o ON c.id=o.customer_id
WHERE profitlevel='High' AND age=50 and credit_limit=5000;
```

- 1) Σύμφωνα με την EXPLAIN ποιο είναι το εκτιμώμενο συνολικό κόστος για την εκτέλεση του καλύτερου πλάνου για το ερώτημα αυτό; Ποια τα CPU_COST και IO_COST; Ποια είναι η πιο χρονοβόρα ενέργεια.
- 2) Ποιο είναι το εκτιμώμενο πλήθος αποτελεσμάτων για το ερώτημα; Πόσες πλειάδες επιστρέφει πραγματικά το ερώτημα; Χρησιμοποιώντας συμβολισμούς σχεσιακής άλγεβρας, σχεδιάστε το πλάνο εκτέλεσης που επέλεξε ο optimizer.

Επιχειρήστε να βελτιστοποιήσετε το ερώτημα δημιουργώντας τα κατάλληλα ευρετήρια.

- 3) Ποιο το τελικό κόστος μετά τη βελτιστοποίηση της σχεδίασης;

Δώστε τις εντολές που χρησιμοποιήσατε σε κάθε βήμα

Άσκηση 4^η – Βελτιστοποίηση ερωτήματος

ΕΡΩΤΗΜΑ 1: Επερώτηση εύρους τιμών με ευρετήριο

Τι θα αλλάξει αν κάνετε το ακόλουθο ερώτημα

```
SELECT c.name, o.price, o.profit
FROM customers c JOIN orders o ON c.id=o.customer_id
WHERE profitlevel='High' AND age=50 and credit_limit>5000;
```

- 1) Σύμφωνα με την EXPLAIN ποιο είναι το εκτιμώμενο συνολικό κόστος για την εκτέλεση του καλύτερου πλάνου χωρίς ευρετήρια για το ερώτημα αυτό; Ποια τα CPU_COST και IO_COST; Ποια είναι η πιο χρονοβόρα ενέργεια.
- 2) Επιχειρήστε να βελτιστοποιήσετε το ερώτημα δημιουργώντας τα κατάλληλα ευρετήρια. Ποιο το τελικό κόστος μετά τη βελτιστοποίηση της σχεδίασης;

Δώστε τις εντολές που χρησιμοποιήσατε σε κάθε βήμα

B – Οδηγίες Παράδοσης

Η εργασία θα υλοποιηθεί από ομάδες των **3 ατόμων (το πολύ)**, αν και επιτρέπονται μικρότερες ομάδες. Θα πρέπει τελικά να ανεβάσετε ένα zip αρχείο με ονομασία τους ΑΜ των μελών της ομάδας: π.χ. **AM1-AM2-AM3.zip**

- Το zip θα περιλαμβάνει:
 - ένα αρχείο readme.txt
 - με τα ονοματεπώνυμα και τους ΑΜ των φοιτητών της ομάδας
 - το αρχείο word με την τελική εργασία. Τυχόν συνοδευτικά αρχεία sql,rpt κλπ.

Γ – Άλλες Οδηγίες

Όσες εργασίες δεν τηρούν τις οδηγίες παράδοσης, θα έχουν επίπτωση στο βαθμό.

Όσες εργασίες κριθούν ότι είναι **αντιγραφές θα μηδενίζονται**.

Ημερομηνία παράδοσης: **Στο e-class με οριστική τελική ημερομηνία 18-11-2016**

Όσες εργασίες παραδοθούν μετά το πέρας της ημερομηνίας και μέχρι τις 20-11-2016 θα έχουν μείωση 2 μονάδων στο βαθμό.