



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΞΑΜΗΝΙΑΙΑ ΕΡΓΑΣΙΑ ΣΤΙΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

6ο εξάμηνο, Ακαδημαϊκή περίοδος : 2019-2020, ΣΗΜΜΥ

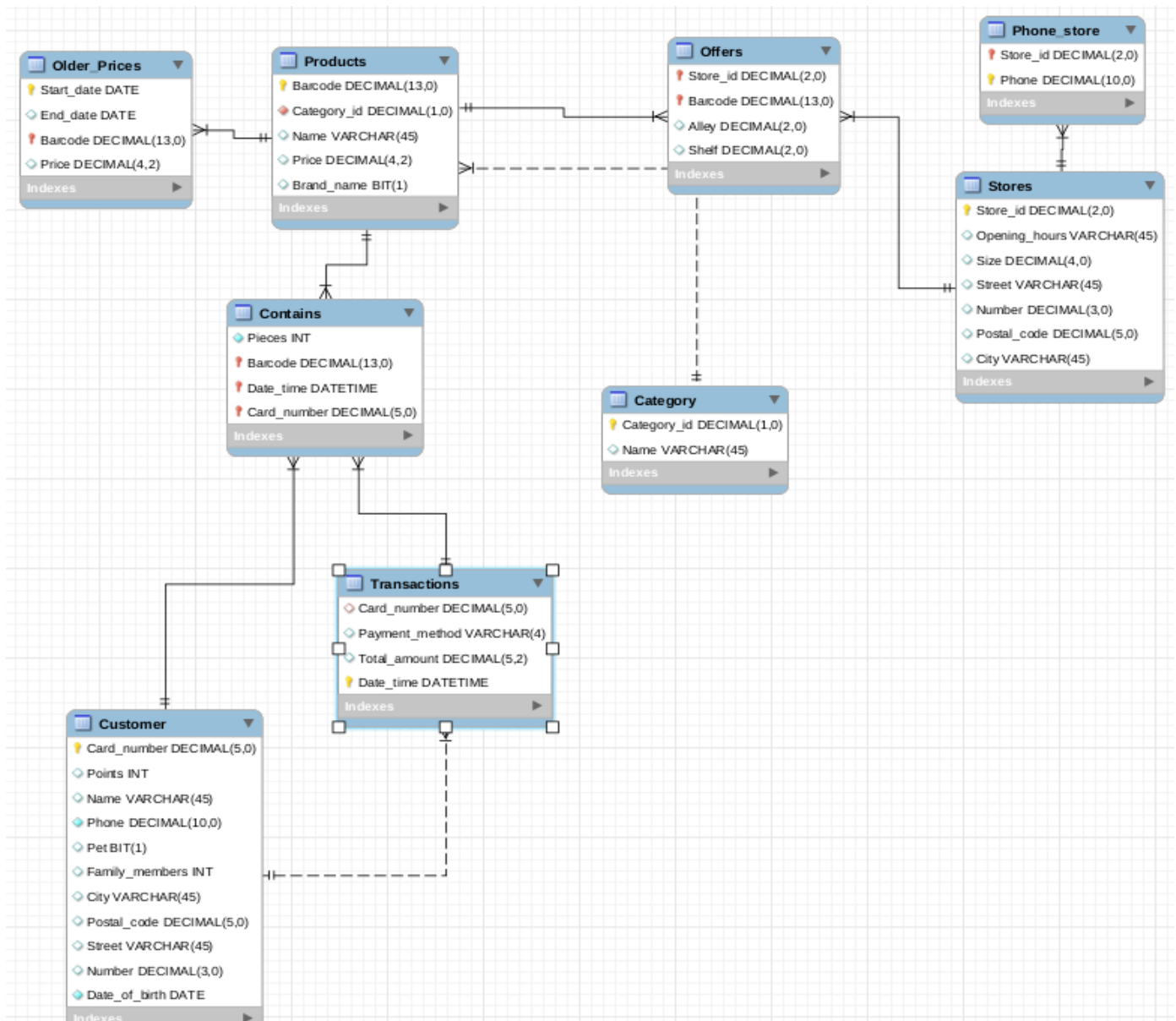
ΑΝΑΛΥΣΗ ΤΟΥ ΚΑΛΑΘΙΟΥ ΑΓΟΡΩΝ - SHOPPING BASKET ANALYSIS

Ομάδα 29

Μέλη ομάδας : Βασίλης Αμοιρίδης , AM : 03117846

& Αλέξανδρος Καλαϊτζής, AM : 03117017

1. ΣΧΕΣΙΑΚΟ ΔΙΑΓΡΑΜΜΑ ΤΗΣ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ



α) Οι περιορισμοί που έχουν οριστεί ανά κατηγορία είναι οι εξής:

Category:

- Category id numeric(1,0): Κάθε κατηγορία χαρακτηρίζεται από ακριβώς έναν μονοψήφιο αριθμό διότι οι κατηγορίες είναι 6 και το γνώρισμα αυτό αποτελεί Primary key της σχέσης Category
- Name VARCHAR(45): Το όνομα της κατηγορίας μπορεί να πάρει μέχρι 45 χαρακτήρες

Products:

- Barcode numeric(13,0) : Κάθε προϊόν χαρακτηρίζεται από έναν αριθμό ακριβώς 13 ψηφίων (Primary Key)
- Category_id numeric(1,0) : Foreign key από την οντότητα Category
- Name VARCHAR(45) : Το όνομα του προϊόντος μπορεί να πάρει μέχρι 45 χαρακτήρες
- Price numeric(4,2) check(Price > 0) : Η τιμή του προϊόντος πρέπει να έχει 2 δεκαδικά ψηφία και 2 ακέραια και να είναι θετική
- Brand_name BIT : Τιμή 1 → ετικέτα καταστήματος, Τιμή 0 → προϊόν άλλης εταιρείας

Stores:

- Store_id numeric(2,0) : Κάθε κατάστημα χαρακτηρίζεται από έναν αριθμό ακριβώς 2 ψηφίων και αποτελεί το Primary key
- Opening hours VARCHAR(45) : Το ωράριο γράφεται σαν μια συμβολοσειρά μεγέθους μέχρι 45 χαρακτήρων
- Size numeric(3,0) : Τα τετραγωνικά μέτρα του καταστήματος είναι τριψήφιος αριθμός
- Street VARCHAR(45) : Το όνομα της οδού έχει μέχρι 45 χαρακτήρες
- Number numeric(3,0) : Το νούμερο της οδού είναι τριψήφιος αριθμός
- Postal_code numeric(5,0) : Ο ταχυδρομικός κώδικας είναι πενταψήφιος αριθμός
- City VARCHAR(45) : Το όνομα της πόλης έχει μέχρι 45 χαρακτήρες

Customer:

- Card_number numeric(5,0) : Κάθε πελάτης χαρακτηρίζεται από έναν αριθμό ακριβώς 5 ψηφίων. Έχουμε θεωρήσει ότι από τα 5 ψηφία τα 3 πρώτα αποτελούν το ΑΦΜ του πελάτη και τα 2 τελευταία ψηφία έχουν το store_id του καταστήματος από το οποίο ψωνίζουν (Primary key)
- Points INT : Πόντοι σε ακέραιο αριθμό χωρίς περιορισμό

- Name VARCHAR(45) NOT NULL: Το ονοματεπώνυμο του πελάτη έχει μέχρι 45 χαρακτήρες και δεν μπορεί να είναι null
- Phone numeric(10,0): Ο αριθμός τηλεφώνου πελάτη έχει ακριβώς 10 ψηφία
- Pet BIT: Τιμή 1 → έχει κατοικίδιο, Τιμή 0 → δεν έχει
- Family_members INT: Αριθμός μελών οικογένειας χωρίς περιορισμούς
- City VARCHAR(45): Το όνομα της πόλης που μένει έχει μέχρι 45 χαρακτήρες
- Street VARCHAR(45): Το όνομα της οδού που μένει έχει μέχρι 45 χαρακτήρες
- Number numeric(3,0): Το νούμερο της οδού που μένει είναι τριψήφιος αριθμός
- Postal_code numeric(5,0): Ο ταχυδρομικός κώδικας της περιοχής που μένει είναι πενταψήφιος αριθμός
- Date of birth DATE NOT NULL: Η ημερομηνία γέννησης της μορφής "YYYY-MM-DD" που δεν μπορεί να πάρει τιμή null.

Phone_store:

- Phone numeric(10,0): Δεκαψήφιος αριθμός τηλεφώνου καταστήματος που αποτελεί μέρος του Primary key
- Store_id numeric(2,0): Foreign key από την οντότητα stores και συμπληρώνει το PK του Phone store.

Older_Prices:

- Start_date DATE: Ημερομηνία αλλαγής τιμής της μορφής "YYYY-MM-DD" που αποτελεί μέρος του Primary key
- End_date DATE: Ημερομηνία λήξης ισχύος αυτής της τιμής του προϊόντος . Παίρνει τιμή null αν η τελευταία τιμή του προϊόντος συνεχίζει να ισχύει μέχρι σήμερα
- Barcode numeric(13,0): Foreign key από την οντότητα Products, συναποτελεί το PK διότι το Start_date από μόνο του δεν προσδιορίζει μοναδικά την ημερομηνία αλλαγής της τιμής του προϊόντος
- Price numeric(4,2) check(Price>0): Τιμή προϊόντος με 2 δεκαδικά ψηφία και 2 ακέραια, πάντα θετική

Transactions:

- Date time DATETIME: Ημερομηνία συναλλαγής της μορφής "YYYY-MM-DD HH:MI:SS " (μέρος του Primary key)
- Card number numeric(5,0): Foreign key από την οντότητα Customer, διότι το Date_time από μόνο του δεν αρκεί για να προσδιοριστεί μοναδικά μία συναλλαγή
- Payment method varchar(4): Μέθοδος πληρωμής 'cash' ή 'card'
- Total amount numeric(5,2): Συνολικό ποσό πληρωμής για την συγκεκριμένη συναλλαγή που περιγράφεται από 3 ακέραια ψηφία και δύο δεκαδικά

Offers:

- Store id numeric(2,0): Παίρνει ως μέρος του primary key το κλειδί της οντότητας stores
- Barcode numeric(13,0): Παίρνει ως μέρος του primary key το κλειδί της οντότητας products
- Alley numeric(2,0): Διψήφιος αριθμός για τον διάδρομο που βρίσκεται το προϊόν στο κατάστημα
- Shelf numeric(2,0): Διψήφιος αριθμός για το ράφι που βρίσκεται το προϊόν στο κατάστημα

Contains:

- Pieces INT NOT NULL: Τεμάχια από ένα προϊόν που βρίσκεται στην συναλλαγή
- Barcode numeric(13,0): Κομμάτι του Primary key το Barcode του προϊόντος
- Date time DATETIME: Κομμάτι του Primary key η ημερομηνία της συναλλαγής
- Card number numeric(5,0): Συναποτελεί το Primary key του πίνακα Contains μαζί με τα Barcode και Date_time

b) Τα ευρετήρια που έχουν οριστεί είναι τα εξής:

create index transaction_date_time_idx on Transactions(Date_time);

create index contains_card_number_idx on Contains(Card_number);

create index contains_barcode_idx on Contains(Barcode);

Έχουμε ορίσει τα παραπάνω τρία ευρετήρια γιατί η ιδιότητα Date_time του πίνακα Transactions και τα χαρακτηριστικά Barcode και Card_number του πίνακα Contains είναι αυτά που χρησιμοποιούνται περισσότερο για τον υπολογισμό των ζητούμενων queries. Για τα δύο χαρακτηριστικά του πίνακα Contains θα μπορούσαμε να είχαμε φτιάξει έναν index, αλλά επειδή ήταν μόνο ένα το ερώτημα στο οποίο χρησιμοποιήσαμε και τα δύο (ενώ πολύ περισσότερα αυτά στα οποία χρησιμοποιούνται ξεχωριστά) επιλέξαμε να ορίσουμε δύο ξεχωριστά ευρετήρια στον ίδιο πίνακα.

c) Το σύστημα και η γλώσσα προγραμματισμού που χρησιμοποιήθηκαν για την ανάπτυξη της εφαρμογής είναι το Mysql Workbench και η SQL, αντίστοιχα.

d) Για να εγκαταστήσει κάποιος την εφαρμογή στο επίπεδο που την έχουμε αναπτύξει αρκεί να προχωρήσει στην εγκατάσταση της Mysql και του Mysql Workbench.

3. Όλος ο sql κώδικας που έχουμε γράψει παραδίδεται σε ξεχωριστό αρχείο μαζί με την παρούσα αναφορά.

5. Ο παρακάτω πίνακας περιλαμβάνει για κάθε ερώτημα το αρχικό και το τελικό timestamp στο video μας :

Ερώτημα	Αρχή	Τέλος
1	00:00	00:08
2	00:08	00:27
3	00:27	01:46
4	01:46	02:15
5	02:15	03:00
6	03:00	05:23
7	05:23	06:17
8	06:17	07:32
9	07:32	08:06
10	08:06	09:50
11	09:50	10:10
12	10:10	10:56
13	10:56	11:36
14	11:36	12:19
15	-	-

Σημείωση: Η υποβολή του βίντεο στο MS Teams έγινε από τον Αλέξανδρο Καλαϊτζή (el17017).