BIT2008A/ITEC2000A: Multimedia Data Management
Fall 2022

Assignment 1

Relational Algebra

Question 1. Generally, if a relation R has n attributes, namely $(A_1, A_2, ..., A_n)$, and we have k number of those attributes taken as candidate keys, how many superkeys does the relation R have? (5 points)

Question 2. Please specify the truth of each statement and include your reason. (15 points)

1) $\sigma_\theta(R \cup S) = \sigma_\theta(R) \cup \sigma_\theta(S)$ is always valid.
2) $\Pi_L(\sigma_\theta(S)) = \sigma_\theta(\Pi_L(S))$ is always valid.
3) $\sigma_\theta(R) - S = \sigma_\theta(R - S)$ is always valid.

Question 3. Assume that we have manged to store the data of a transportation system as the relations below:

- Company($\underline{\text{comp\_id}}$, name)
- Trip($\underline{\text{trip\_id}}$, date, origin, dest, comp_id, duration)
- Passenger($\underline{\text{pass\_id}}$, name)
- Pass_in_Trip($\underline{\text{trip\_id}}$, $\underline{\text{pass\_id}}$,

Write each of these queries using relational algebra. (15 points)

1) The date of every trip starting from ”Ottawa”.
2) The list of every city that a passenger named ”Alex” had a trip to.
3) The ID of every passenger who has a trip after 2022-08-08, which takes less than 12 minutes.

Question 4. Having the relations below for a retail management system:

- Producer($\underline{\text{s\_id}}$, s_name, s_city)
- Product($\underline{\text{p\_id}}$, p_name, p_color)
- Produce($\underline{\text{s\_id}}$, $\underline{\text{p\_id}}$)

Please interpret the result of the following query. (5 points)

$\Pi_{s\_name}(\sigma_{Producer.s\_city=P2.s\_city}(Producer \times \rho_{P2}(\Pi_{s\_city}(\sigma_{s\_id=8}(Producer)))))$

Implementation

In this part, we assume that we are responsible for managing the database of an online store. This database stores information from customers, products, and shopping bags. Additionally, customers are allowed to rate the products. Basically, you are expected to do these objectives using PostgreSQL.

1) Creating the schema (25 points)
2) Importing the instances (5 points)
3) Executing the queries (30 points)

Creating the schema. Please create a database with the schema described below. It is up to you to define the superkeys, foreign keys, and the datatype of the attributes for each table.

User Table:

- user_id: this field is a unique numerical ID assigned to each user.
- name: this filed consists the name of a user.
- address: this field only includes the street address of a user.
- phone: this field include the phone number of a user.

Product Table:

- product_id: this field is a unique numerical ID assigned to each product.
- name: this field consists the name of a product.

- category_id: this field holds the category ID of each product.
- price: this field shows the unit price of a product.

Order Table:

- order_id: this field is a numerical ID assigned to each order.
- user_id: this field holds the user ID of the person who has submitted an order.
- product_id: this field holds the ID of each product included in an order.
- quantity: this field shows how many units of a certain product there are in an order.
- shipping: this field shows the shipping status of a record.
- timestamp: this field stores the time and date when an order is placed.

Category Table:

- category_id: this field holds a unique ID assigned to each category.
- name: this field contains the name of a category.

Rating Table:

- user_id: this field specifies the user who has done the rating.
- product_id: this field specifies the product to be rated.
- rate: this field shows the score given to a product from 0 to 5.
- timestamp: this field stores the time and date when the rating is submitted.

Importing the instances. It is up to you to import enough records to make sure that the queries will make sense and show meaningful results.

Executing the queries.

1) The list of every product without a rating assigned to them.
2) The list of user names who have ordered more than three times from the products in the ”Sport” category.
3) The list of every product with an average rating above 4.
4) The list of every product that has been sold above $100 in total.
5) Number of users who have spent more than $100 in 2021-11-25.
6) The name and total spending of the user with the ID of 5.
7) The details of the best-selling product.
8) The name of the best-selling category.
9) The name of the second user with the most number of submitted ratings.
10) The name of the user with the longest gap between their first and last ratings.

Description

1) The due date of this assignment is on September 25th, 11:55 PM. Late submission policy can be found on the course outline.
2) You are expected to submit your solution for all of the assignments, then the maximum three scores will be calculated.
3) Please include your answer to the first part of the assignment (Q1 to Q4) in a PDF file. It can be handwritten, but make sure it is easily readable.
4) Please create a file named Table.sql for the first part of the implementation (Creating the schema).
5) Please create a file named Data.sql for the second part of the implementation (Importing the instances).
6) Please create a file named Query.sql for the third part of the implementation (Executing the queries).
7) Please upload your submission as Lastname_Firstname_StudentID.zip on Brightspace.