Database Systems (Solution)
Assignment #2
Spring 2018

Due Date: Before the start of the class (14th February, 2018)

Instructions:

Please ignore the errors of spellings, brackets if found any.

There can be multiple ways of solving the same problem

Please ignore the Relation Algebra Signs mistakes as well and do follow as per convention/book.

⋈ = joining operation

⟕ = left outer join

F= Aggregate Function

Question 1:

An emerging mobile company wants to develop an online database of mobile apps (applications) available for download to its users. You can assume that they want to develop a simpler version of Android PlayStore or IOs AppStore.

You are given the following relational database schema

App(app-id, app-name, developer-id, cost, category)

User(user-id, user-name, country)

Developer(developer-id, developer-name, age, country)

Downloads(app-id, user-id, rating, review)

For each App they record the app id, app name, the app developer's id, cost of the app (some app may be free of cost) and category. The category can be games, productivity, kids, entertainment etc. The system maintains the download details such as which app is downloaded by which particular user. In addition to this it also records the rating (on a scale of 1-5) and review (if any) given by each user to a particular app.

Write down relational expressions for the following queries:

1. Print the names of developers who belong to Pakistan and are under the age of 18.

$$\pi_{\text{developer-name}}(\sigma_{\text{age} < 18 \text{ and } \text{country='pakistan'}}(\text{Developer}))$$

2. Print the names of the apps developed by 'Ali-Shah'

$$R1 \leftarrow \pi_{\text{developer-id}}(\sigma_{\text{developer-name='ali shah'}}(\text{Developer}))$$

$$\pi_{\text{app-name}}(R1 \bowtie_{\text{R1.developer-id = App.developer-id}} \text{App})$$

3. List the ids of the users who bought 'kids' apps developed by Pakistanis. Note kids is a category of apps.

$$R1 \leftarrow \pi_{\text{developer-id}}(\sigma_{\text{developer-country='pakistan'}}(\text{Developer}))$$

$$R2 \leftarrow \pi_{\text{app-id}}(R1 \bowtie_{\text{R1.developer-id = App.developer-id}} (\sigma_{\text{category='kids'}} \text{App}))$$

$$\pi_{\text{user-id}}(R2 \bowtie_{\text{R2.appid = downloads.appid}} \text{Downloads})$$

4. Find the names and ids of the users who purchased ‘productivity’ apps but did not buy ‘entertainment’ apps.

$$Q \leftarrow (\text{Downloads} \bowtie_{\text{downloads.app-id = App.app-id}} (\sigma_{\text{category= ‘productivity’}} \text{App}))$$

$$R1 \leftarrow \pi_{\text{user-id, username}} (\text{User} \bowtie_{\text{user.user-id = q.user-id}} Q)$$

$$Y \leftarrow (\text{Downloads} \bowtie_{\text{downloads.app-id = App.app-id}} (\sigma_{\text{category= ‘entertainment’}} \text{App}))$$

$$R2 \leftarrow \pi_{\text{user-id, username}} (\text{User} \bowtie_{\text{user.user-id = y.user-id}} Y)$$

$$\text{Result} = R1 - R2$$

5. Find the ids of users who purchased the apps named ‘The Zone’ and ‘Fighters’.

$$A1 \leftarrow \pi_{\text{app-id}} (\sigma_{\text{app-name=’The Zone’ or app-name=’Fighters’}} \text{App})$$

$$A2 \leftarrow \pi_{\text{user-id, app-id}} (\text{Downloads})$$

$$\text{Users} \leftarrow A2 \div A1$$

Alternate Method: Through Intersection (division is expensive operation try solving through intersection method)

$$A1 \leftarrow \pi_{\text{app-id}} (\sigma_{\text{app-name=’Fighters’}} \text{App})$$

$$P1 = \pi_{\text{user-id}} (A1 \bowtie_{\text{A1.appid = downloads.appid}} \text{Downloads})$$

$$A2 \leftarrow \pi_{\text{app-id}} (\sigma_{\text{app-name=’The Zone’}} \text{App})$$

$$P2 = \pi_{\text{user-id}} (A2 \bowtie_{\text{A2.appid = downloads.appid}} \text{Downloads})$$

$$\text{Result} = P1 \cap P2.$$

6. Find id of the users who purchased apps developed by ‘Ahmed’.

$$R1 \leftarrow \pi_{\text{developer-id}} (\sigma_{\text{developer-name='ahmad'}}(\text{Developer})$$

$$R2 \leftarrow \pi_{\text{app-id}} (R1 \bowtie_{\text{R1.developer-id = App.developer-id}} \text{App})$$

$$\pi_{\text{user-id}} (R2 \bowtie_{\text{R2.appid = downloads.appid}} \text{Downloads})$$

7. Find id of the users who purchased **every** app developed by 'Ali-Shah'.

$$R1 \leftarrow \pi_{\text{developer-id}} (\sigma_{\text{developer-name='Ali-Shah'}}(\text{Developer}))$$

$$R2 \leftarrow \pi_{\text{app-id}} (R1 \bowtie_{\text{R1.developer-id = App.developer-id}} \text{App})$$

$$A \leftarrow (\pi_{\text{user-id,app-id}} (\text{Downloads}))$$

$$\pi_{\text{user-id}} (A \div R2)$$

8. Locate the apps that are bought by at least two different users.

$$R1 \leftarrow {}_{\text{app-id}} F_{\text{count(user-id)}} (\text{Downlaods})$$

$$R2 \leftarrow (\sigma_{\text{count\_user-id=>2}}(R1))$$

$$\pi_{\text{app-id, app-name}} (R2 \bowtie_{\text{r2.app-id = App.app-id}} \text{APP})$$

9. Find the most popular apps (that is apps with maximum rating).

$$R1 \leftarrow F_{\text{max(rating)}} (\text{Downlaods})$$

$$R2 \leftarrow \pi_{\text{app-id}} (R1 \bowtie_{\text{R1.max\_rating = downloads.rating}} \text{Downlaods})$$

$\pi_{\text{app-id, app-name}}$ (R2 $\bowtie_{\text{r2.app-id = App.app-id}}$ APP)

10. Find the most expensive apps (that is apps with highest cost)

$R1 \leftarrow F_{\max(\text{cost})}(\text{App})$

$\pi_{\text{app-id, app-name}}$ (R1 $\bowtie_{\text{R1.max\_cost = app.cost}}$ App)

11. Find the most expensive apps made by the developer 'Reeta'

$R1 \leftarrow \pi_{\text{developer-id}}(\sigma_{\text{developer-name='reeta'}}(\text{Developer}))$

$R2 \leftarrow (R1 \bowtie_{\text{R1.developer-id = App.developer-id}} \text{App})$

$R3 \leftarrow F_{\max(\text{cost})}(R2)$

$\pi_{\text{app-id, app-name}}$ (R3 $\bowtie_{\text{R3.max\_cost = R2.cost}}$ R2)

12. List the app-id, app-name and developer-name of all the free apps.

$R1 \leftarrow (\sigma_{\text{cost=0}} \text{App})$

$R2 \leftarrow \pi_{\text{app-id, app-name, developer-name}}(\text{Developer} \bowtie_{\text{Developer.developer-id = R1.developer-id}} R1)$

Free apps mean that their cost is '0'. Null Value has three different meanings (mentioned in your book) therefore (cost=NULL is not valid condition for this)

13. List the app-id of all free apps with average rating greater than 3.

$$R1 \leftarrow (\sigma_{cost=0} App)$$

$$R1 \leftarrow \pi_{app\text{-}id,rating} (Downloads \bowtie_{downloads.app\text{-}id = r1.app\text{-}id} R1)$$

$$\sigma_{avg\_rating>3} ({}_{app\text{-}id} F_{avg(rating)}(R1))$$

14. Find the total number of apps made by each developer.

$$R1 \leftarrow \pi_{app\text{-}id,\ app\text{-}name,\ developer\text{-}name} (Developer \bowtie_{Developer.developer\text{-}id = App.developer\text{-}id} App)$$

$${}_{developer\text{-}id,\ developer\text{-}name} F_{count(app\text{-}id)}(R1)$$

15. List the total income earned by each developer.

$$R1 \leftarrow (Developer \bowtie_{Developer.developer\text{-}id = App.developer\text{-}id} App)$$

$${}_{developer\text{-}Id,\ developer\text{-}name} F_{sum(cost)}(Downloads \bowtie_{downloads.app\text{-}id = app.app\text{-}id} R1)$$

16. Find the total revenue generated for each country.

$$R1 \leftarrow (Developer \bowtie_{Developer.developer\text{-}id = App.developer\text{-}id} App)$$

$${}_{country} F_{sum(cost)}(Downloads \bowtie_{downloads.app\text{-}id = app.app\text{-}id} R1)$$

17. Find the id and name of the most expensive app in each category.

$$R1 \leftarrow {}_{category} F_{max(cost)}(App)$$

$$\pi_{app.app\text{-}id,\ app.app\text{-}name} (R1 \bowtie_{r1.category = app.category \text{ and } r1.max\_cost=qpp.cost} App)$$

18. Find the number of apps made by each developer.

$${}_{developer\text{-}Id,\ developer\text{-}name} F_{count(app\text{-}id)}((Developer \bowtie_{Developer.developer\text{-}id = App.developer\text{-}id} App))$$

Also include the developer who have not released any app yet?

19. Find the developers who have made exactly three ‘games’.

$$Q \leftarrow \sigma_{\text{category='games'}}(\text{App})$$

$$R1 \leftarrow {}_{\text{developer-Id, developer-name}}F_{\text{count(app-id)}}(\text{Developer} \bowtie_{\text{Developer.developer-id = Q.developer-id}} Q)$$

$$\sigma_{\text{count\_app-id=3}}(R1)$$

20. List the name of the developers and the app they created. If they have not developed any app then indicate this with a null value.

$$\pi_{\text{developer-name, app-name}}(\text{Developer} \ltimes_{\text{Developer.developer-id = App.developer-id}} \text{App}))$$

Question 2:

Consider the following relational database for the Baseball League. It keeps track of teams in the league, coaches and players on the teams, work experience of the coaches and which players have played on which teams.

Note the following facts about this environment:

• The database keeps track of the history of all the teams that each player has played on and all the players who have played on each team.

• The database only keeps track of the current team that a coach works for.

• Team number, team name, and player number are each unique attributes across the league.

• Coach name is unique only within a team (and we assume that a team cannot have two coaches of the same name).

• In the Affiliation table, the Years attribute indicates the number of years that a player played on a team; the batting average is for the years that a player played on a team.

Team (TeamNum, TeamName, City, Manager)

Coach (TeamNum, CoachName, Address)

WorkExperience (TeamNum, CoachName, ExperienceType, YearsExperience)

Player (PlayerNum, PlayerName, Age)

Affiliation (PlayerNum, TeamNum, Years, BattingAvg)

Construct the following RA queries for the above relational database.

1. Display the names of the players who are under 18.

$\pi_{\text{PlayerName}}(\sigma_{\text{age}< 18}(\text{Player}))$

2. List all the coaches who have between 5 and 10 years of experience as college coaches (see YEARSEXPERIENCE and EXPERIENCETYPE).

$\pi_{\text{coachName}}(\sigma_{\text{ExperienceType='college' and (YearsExperience >= 5 and YearsExperience <= 10)}} (\text{WorkExperience}))$

3. Find the total number of years of experience of each coach on team number 23.

$_{\text{courseName}}\text{F}_{\text{yearsExperience}}(\sigma_{\text{TeamNum=23}} (\text{WorkExperience}))$

4. Assume that team names are unique. Find the names of the players who have played for the Dodgers for at least five years (see YEARS in the AFFILIATION Table.)

$\text{R2} <- (\sigma_{\text{TeamName='dodgers'}} (\text{Team}))$

$\text{R3} <- \text{Affiliation} \bowtie_{\text{affiliation.TeamNum = R2.TeamNum}} \text{R2}$

$\text{R1} <- \sigma_{\text{sum\_years >= 5}} (_{\text{PlayerNum}}\text{F}_{\text{sum(years)}} (\text{R3}))$ OR $(\sigma_{\text{sum\_years >= 5}} (\text{R3}))$

$\pi_{\text{PlayerName}} (\text{R1} \bowtie_{\text{R1.TeamNum = Player.TeamNum}} (\text{Player}))$

5. Assume that team names are unique. Find the total number of years of work experience of each coach on the Dodgers, but include in the result only those coaches who have more than eight years of experience.

$\text{R1} <- (\text{WorkExperience} \bowtie_{\text{WorkExperience.TeamNum = Team.TeamNum}} (\sigma_{\text{TeamName='dodgers'}} (\text{Team})))$

$\text{R2} <- _{\text{TeamNum, CoachName}}\text{F}_{\text{sum(yearsExperience)}} (\text{R1})$

$(\sigma_{\text{sum\_YearsExperience} > 8} (\text{R2}))$

6. Find the names of the league's youngest players.

$$R1 \leftarrow \left({}_{\min(age)}F(Player) \right)$$

$$\pi_{PlayerName} \left(R1 \bowtie_{R1.min\_age = Player.Age} (Player) \right)$$

7. Find the players who have been affiliated with more than two teams.

$$R1 \leftarrow \left({}_{PlayerNum}F_{count(TeamNum)}(Affiliation) \right)$$

$$R2 \leftarrow \left(\sigma_{count\_TeamNum > 2}(Affiliation) \right)$$

$$\pi_{Player.PlayerNum,\ Player.PlayerName} \left(R2 \bowtie_{R2.PlayerNum = Player.PlayerNum} (Player) \right)$$

8. Find the players who are not affiliated with any team.

$$R1 \leftarrow \pi_{PlayerNum}(Player)$$

$$R2 \leftarrow \pi_{PlayerNum}(\sigma_{count\_TeamNum > (Affiliation)})$$

$$R3 \leftarrow R1 - R2$$

$$\pi_{Player.PlayerNum,\ Player.PlayerName} \left(R3 \bowtie_{R3.PlayerNum = Player.PlayerNum} (Player) \right)$$

9. For each player, find the number of teams they have been affiliated with in their career.

$$R1 \leftarrow \left({}_{PlayerNum,PlayerName}F_{count(TeamNum)}(Affiliation) \right)$$

$$\pi_{PlayerNum,\ PlayerName}(Player \bowtie_{Player.PlayerNum = R1.PlayerNum} R1)$$

10. For each team, list the average batting average.

$$R1 \leftarrow {}_{TeamNum}F_{average(battingaverage)}(Affiliation)$$

$$\pi_{\text{Team.TeamNum, Team.TeamName, R1.avg\_BattingAvg}} (\text{Team} \bowtie_{\text{Team.TeamNum = R1.TeamNum}} (\text{R1}))$$