Name

UIN

Homework Assignment 1

Due Date: Monday, Feb 16, 2026

CS480 - Database Systems Results

Please leave this empty!

1.1 1.2 1.3 1.4 1.5 1.6

1.7 1.8 1.9 1.10 1.11 Sum

Instructions

- Try to answer all the questions using what you have learned in class
- **When writing a query, write the query in a way that it would work over all possible database instances and not just for the given example instance!**
- Some questions are marked as bonus. You do not have to answer these questions to get full points for the assignment. However, you can get bonus points for these questions!
- **Please submit the homework electronically using gradescope**
- **Submissions will be automatically graded. You can submit multiple times.**

Solutions for this question have to be written in the provided text files (e.g., `q1.1.1.ra` for question `1.1.1`. Relational algebra expressions are written in the format of the `radb` python package (`https://users.cs.duke.edu/~junyang/radb/start.html`). To test your code, install this package and use it to connect to a database (the autograder uses Postgres: `https://www.postgresql.org/`) that has the provide example schema and data (load `homework_1.sql`). For example, $\pi_{title}(songs)$ would we written as follows in `radb`'s dialect:

```
\project_{title}(songs)
```

Consider the following database schema and example instance:

songs

songid	title	genre	duration	releasedate	bandid
101	Bohemian Rhapsody	Rock	354	1975-10-31	1
102	Under Pressure	Rock	248	1981-10-26	1
103	Under Pressure	Rock	248	1981-10-26	2
104	Numb/Encore	Hip-Hop	212	2004-11-30	3

platforms

platformid	platformname	type	unitrevenue
10	Spotify	audio	1.00
20	YouTube	video	1.00
30	Apple Music	audio	2.00
40	SoundCloud	audio	3.00

bands

id	name	label
1	Queen	EMI
2	David Bowie	RCA
3	Linkin Park	Warner
4	Jay-Z	Roc-A-Fella

streams_daily

songid	bandid	platformid	streamdate	streamscount
101	1	10	2025-01-01	1500
102	1	10	2025-01-01	1100
103	2	10	2025-01-01	1050
104	3	20	2025-01-01	1800

Hints:

- Attributes shown with grey background are the **primary key** of a relation.
- Attribute *songid* of relation *streams_daily* is a foreign key to relation *songs*.
- Attribute *bandid* of relation *streams_daily* is a foreign key to relation *bands*.
- Attribute *platformid* of relation *streams_daily* is a foreign key to relation *platforms*.
- Attribute *bandid* of relation *songs* is a foreign key to relation *bands*.

Part 1.1 Relational Algebra (Total: 100 + 10 bonus Points)

Question 1.1.1 (6 Points)

Write a relational algebra expression that returns the `title`, `releasedate`, and `duration` of all *Rock* genre songs.

Solution

$$\pi_{title,releasedate,duration}(\sigma_{genre=Rock}(songs))$$

Question 1.1.2 (6 Points)

Write a relational algebra expression that returns the songs longer than 4 minutes. Note that in the song table, durations are specifcied in seconds.

Solution

$$\pi_{songid,title}\big(\sigma_{duration>240}(songs)\big)$$

Question 1.1.3 (8 Points)

Write a relational algebra expression that returns `title`, `platformname`, `bandname`, `streamdate`, and `streamcount` of all the songs released by the *EMI label*.

Solution

$$
\begin{aligned}
Q_{emiBands} &\leftarrow \pi_{id,name}\big(\sigma_{label=EMI}(bands)\big)\\
Q_{emiSongs} &\leftarrow \pi_{songid,title,bandid}\big(songs \bowtie_{bandid=id} Q_{emiBands}\big)\\
Q_{emiStr} &\leftarrow \pi_{songid,bandid,platformid,streamdate,streamscount}\big(streams\_daily \rtimes_{bandid=bandid} Q_{emiSongs}\big)\\
Q_{join} &\leftarrow (Q_{emiStr} \bowtie_{songid=songid \wedge bandid=bandid} Q_{emiSongs}) \bowtie_{platformid=platformid} platforms\\
Q &\leftarrow \pi_{title,platformname,name,streamdate,streamscount}(Q_{join})
\end{aligned}
$$

Question 1.1.4 (10 Points)

Write a relational algebra expression that returns the **genre** with the highest total number of streams across all platforms and dates.

Solution

$$
\begin{aligned}
Q_T &\leftarrow streams\_daily \bowtie_{songid=songid} songs\\
Q_G &\leftarrow \rho_{genre,tot}\Big(\gamma_{sum(streamscount);genre}(Q_T)\Big)\\
Q_M &\leftarrow \rho_{maxtot}\Big(\gamma_{\max(tot);}(Q_G)\Big)\\
Q &\leftarrow \pi_{genre}\Big(Q_G \bowtie_{tot=maxtot} Q_M\Big)
\end{aligned}
$$

Question 1.1.5 (10 Points)

Write a relational algebra expression that returns the `songid`, `title`, and `releasedate` of the **oldest released song** that was streamed at least once in the year `2025`.

Solution

$$
\begin{aligned}
Q_{S2025} &\leftarrow \sigma_{\text{2025-01-01} \le streamdate \wedge streamdate < \text{2026-01-01}}(streams\_daily)\\
Q_C &\leftarrow songs \bowtie_{songid=songid} \pi_{songid}(Q_{S2025})\\
Q_{Min} &\leftarrow \rho_{minrd}\Big(\gamma_{\min(releasedate);}(Q_C)\Big)\\
Q &\leftarrow \pi_{songid,title,releasedate}\Big(Q_C \bowtie_{releasedate=minrd} Q_{Min}\Big)
\end{aligned}
$$

Question 1.1.6 (10 Points)

Write a relational algebra expression that returns pairs of bands `(band1, band2)` such that the two bands have **collaborated on at least one song**.

Solution

$$
\begin{aligned}
Q_{S1} &\leftarrow \rho_{t,rd,b1}\big(\pi_{title,releasedate,bandid}(songs)\big)\\
Q_{S2} &\leftarrow \rho_{t,rd,b2}\big(\pi_{title,releasedate,bandid}(songs)\big)\\
Q_{C} &\leftarrow \pi_{b1,b2}\big(\sigma_{b1\neq b2}(Q_{S1} \bowtie_{t=t\wedge rd=rd} Q_{S2})\big)\\
Q_{B1} &\leftarrow \rho_{b1,band1}\big(\pi_{id,name}(bands)\big)\\
Q_{B2} &\leftarrow \rho_{b2,band2}\big(\pi_{id,name}(bands)\big)\\
Q &\leftarrow \pi_{band1,band2}\big((Q_C \bowtie_{b1=b1} Q_{B1}) \bowtie_{b2=b2} Q_{B2}\big)
\end{aligned}
$$

Question 1.1.7 (10 Points)

Write a relational algebra expression that returns the `songid` and `title` of songs that have never been streamed on Spotify or YouTube.

Solution

$$
\begin{aligned}
Q_P &\leftarrow \pi_{platformid}\big(\sigma_{platformname=Spotify \vee platformname=YouTube}(platforms)\big) \\
Q_{Bad} &\leftarrow \pi_{songid}\big(streams\_daily \bowtie_{platformid=platformid} Q_P\big) \\
Q_{All} &\leftarrow \pi_{songid}(songs) \\
Q &\leftarrow \pi_{songid,title}\Big(songs \bowtie_{songid=songid} (Q_{All} - Q_{Bad})\Big)
\end{aligned}
$$

Question 1.1.8 (10 Points)

Write a relational algebra expression that returns the most profitable label per platform type. The profit can be calculated as streamcount times unitrevenue.

Solution

$$
\begin{aligned}
Q_J &\leftarrow streams\_daily \bowtie_{platformid=platformid} platforms \bowtie_{bandid=id} bands \\
Q_R &\leftarrow \rho_{type,label,rev}\Big(\gamma_{sum(streamscount \cdot unitrevenue);type,label}(Q_J)\Big) \\
Q_M &\leftarrow \rho_{type,maxrev}\Big(\gamma_{\max(rev);type}(Q_R)\Big) \\
Q &\leftarrow \pi_{type,label}\Big(Q_R \bowtie_{type=type \wedge rev=maxrev} Q_M\Big)
\end{aligned}
$$

Question 1.1.9 (10 Points)

Write a relational algebra expression that returns the song and band with the above-average total streams.

Solution

$$
\begin{aligned}
Q_T &\leftarrow \rho_{songid,bandid,tot}\Big(\gamma_{sum(streamscount);songid,bandid}(streams\_daily)\Big)\\
Q_A &\leftarrow \rho_{avgtot}\Big(\gamma_{avg(tot);}(Q_T)\Big)\\
Q_W &\leftarrow \pi_{songid,bandid,tot}\Big(\sigma_{tot>avgtot}(Q_T \times Q_A)\Big)\\
Q_J &\leftarrow Q_W \bowtie_{songid=songid \wedge bandid=bandid} songs \bowtie_{bandid=id} bands\\
Q &\leftarrow \pi_{title,name,tot}(Q_J)
\end{aligned}
$$

Question 1.1.10 (10 Points)

Write a relational algebra expression that returns the top 2 songs per platform. The popularity of a song on a platform is measured as its total stream count.

Solution

$$
\begin{aligned}
Q_T &\leftarrow \rho_{platformid,songid,tot}\Big(\gamma_{sum(streamscount);platformid,songid}(streams\_daily)\Big)\\
Q_{T1} &\leftarrow \rho_{p,s,t}(Q_T)\\
Q_{T2} &\leftarrow \rho_{p,s2,t2}(Q_T)\\
Q_{Better} &\leftarrow \pi_{p,s,s2}\Big(\sigma_{p=p \wedge t2>t}(Q_{T1} \times Q_{T2})\Big)\\
Q_{Cnt} &\leftarrow \rho_{platformid,songid,numBetter}\Big(\gamma_{count(s2);p,s}(Q_{Better})\Big)\\
Q_{Top2} &\leftarrow \pi_{platformid,songid}\big(\sigma_{numBetter<2}(Q_{Cnt})\big)\\
Q_J &\leftarrow Q_{Top2} \bowtie_{songid=songid} songs \bowtie_{platformid=platformid} platforms\\
Q &\leftarrow \pi_{platformname,title}(Q_J)
\end{aligned}
$$

Question 1.1.11 (10 Points)

Write a relational algebra expression that returns all the bands with above-average profits. The profit of a band is calculated as follows. For each stream of one of the band's songs on a platform X the band receives the `unitrevenue` for this platform.

Solution

$$
\begin{aligned}
Q_J &\leftarrow streams\_daily \bowtie_{platformid=platformid} platforms \\
Q_P &\leftarrow \rho_{bandid,profit}\Big(\gamma_{sum(streamscount \cdot unitrevenue);bandid}(Q_J)\Big) \\
Q_A &\leftarrow \rho_{avgprofit}\Big(\gamma_{avg(profit);}(Q_P)\Big) \\
Q_W &\leftarrow \pi_{bandid,profit}\Big(\sigma_{profit>avgprofit}(Q_P \times Q_A)\Big) \\
Q &\leftarrow \pi_{name,profit}\Big(bands \bowtie_{id=bandid} Q_W\Big)
\end{aligned}
$$

Question 1.1.12 BONUS QUESTION (10 Points)

Write a relational algebra expression that returns, for each band, a **timeline of streaming activity**.

The result schema should be:

$$(\texttt{bandid}, \texttt{startdate}, \texttt{enddate}, \texttt{total\_streams})$$

Each returned tuple should represent the **largest possible interval of consecutive dates (inclusive)** during which the total number of streams for the band remains unchanged.

For example, if a band has total streams 100 on Jan 1 and Jan 2, and 150 on Jan 3, then the interval (Jan 1, Jan 2, 100) should be returned.

Hint: You may find it useful to first identify all dates on which the streaming activity of a band changes, and then construct maximal intervals between such dates.

Solution

$$
\begin{aligned}
Q_T &\leftarrow \rho_{bandid,date,tot}\Big(\gamma_{\sum(streamscount);bandid,streamdate}(streams\_daily)\Big)\\
Q_D &\leftarrow \pi_{bandid,date}(Q_T)\\
Q_{pairs} &\leftarrow \pi_{b,sd,ed}\Big(\rho_{b,sd}(Q_D) \bowtie_{b=b2 \wedge sd<ed} \rho_{b2,ed}(Q_D)\Big)\\
Q_{inbetween} &\leftarrow Q_{pairs} \rtimes_{b=bandid \wedge sd<date \wedge date<ed} Q_D\\
Q_{adjacent} &\leftarrow Q_{pairs} - Q_{inbetween}\\
Q_{Ts} &\leftarrow \rho_{bandid,sd,tot}\big(\pi_{bandid,date,tot}(Q_T)\big)\\
Q &\leftarrow \rho_{bandid,startdate,enddate,total\_streams}\Big(\pi_{b,sd,ed,tot}\big(Q_{adjacent} \bowtie_{b=bandid \wedge sd=sd} Q_{Ts}\big)\Big)
\end{aligned}
$$