Proseminar Datenbanksysteme

Universität Innsbruck - Institut für Informatik

Y. Ipek, M. Mayerl, M. Vötter, E. Zangerle



31.10.2019

Übungsblatt 4 - Lösungsvorschlag

Diskussionsteil (im PS zu lösen; keine Abgabe nötig)

a) degeben seien folgende Relationen, die einen Blogpost (Post) und dazugehörige Tags (Tag) beinhalten, wobei Tags dazu verwendet werden, um Blogposts verschiedenen Themen zuzuordnen:

Post(PostID, Headline, Author)

PostID	Headline	Author
1	My favourite recipes	Mary Potter
2	Setting up a Linux VM	Jane Doe
3	First Steps in Python	Bob Smith
4	Travels 2019	John Doe
5	Introduction to SEO	Anne Johnson
6	Knitting a scarf	William Gold
7	Getting ready for my first marathon	Alicia Silverstone

Tag(PostID, Tag)

•	_
PostID	Tag
1	cooking
1	recipe
1	diy
3	coding
3	hacking
4	usa
4	roadtrip
4	travels
6	wool
6	diy

Lösung

~

RelaX zu dieser Datenbank: http://dbis-uibk.github.io/relax/calc.htm?data=gist: 65141c587bdf3dd124e5ae27a93c85b8

Berechnen Sie das Ergebnis folgender Abfragen:

a) $\sigma_{\text{Headline}=\text{"Travels 2019"}} \text{Post}$



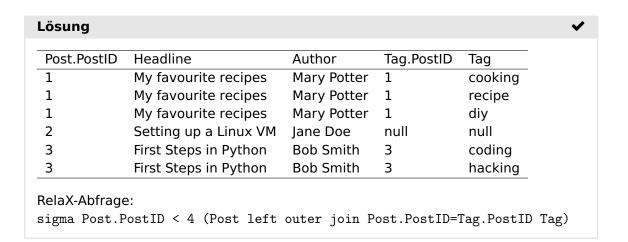
b) $Post \bowtie_{Post.ID=Tag.PostID} Tag$

Post.PostID	Post.Headline	Post.Author	Tag.PostID	Tag.Tag
1	My favourite recipes	Mary Potter	1	cooking
1	My favourite recipes	Mary Potter	1	recipe
1	My favourite recipes	Mary Potter	1	diy
3	First Steps in Python	Bob Smith	3	coding
3	First Steps in Python	Bob Smith	3	hacking
4	Travels 2019	John Doe	4	usa
4	Travels 2019	John Doe	4	roadtrip
4	Travels 2019	John Doe	4	travels
6	Knitting a scarf	William Gold	6	wool
6	Knitting a scarf	William Gold	6	diy

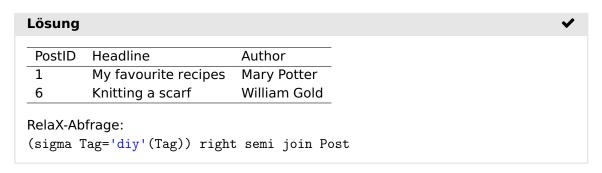
c) $\pi_{Headline,Tag}((\sigma_{Author="MaryPotter"}Post) \times Tag)$



d) $\sigma_{Post.PostID < 4}(Post \bowtie_{Post.PostID = Tag.PostID} Tag)$



e) $(\sigma_{Tag=\text{"diy"}}Tag) \rtimes Post$



b) \bigstar Für die Ausführung der Operationen \cup , -, \cap , \div müssen die Schemata der beiden Relationen ident sein - wieso gilt dies nicht für z.B. Joins?

Lösung



Die Schemata müssen ident sein, da die Operationen \cup , -, \cap , \div auf Zeilenbasis berechnet werden (es werden einzelne Zeilen vergleichen; dies ist nur bei gleichem Schema möglich). Bei Joins hingegen wird der Join nur auf zwei Spalten berechnet, das Schema ist unerheblich für diese Operation.

Lösung



Dies kann durch Projektionen erreicht werden, die so auf die Relationen angewendet werden, dass diese danach das gleiche Schema aufweisen.

Hausaufgabenteil (Zuhause zu lösen; Abgabe nötig)

Aufgabe 1 (Relationale Algebra: Basics)

[4 Punkte]

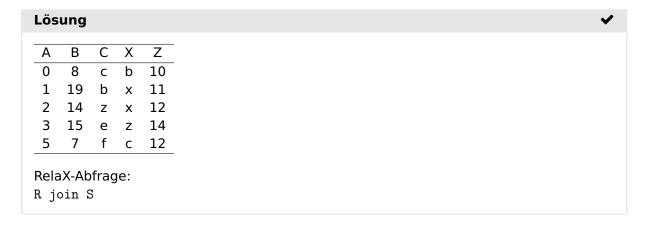
Gegeben sei das folgende relationale Modell:

R(A,B,C)					
Α	В	С			
0	8	С			
1	19	b			
2	14	z			
3	15	е			
4	15	t			
5	7	f			
6	9	Х			

Lösung

RelaX zu dieser Datenbank: http://dbis-uibk.github.io/relax/calc.htm?data=gist: 98e1f0d0710f119c3fcb1113371ae885

a) $\boxed{\textit{0.5 Punkte}}$ Berechnen Sie $R\bowtie S$.

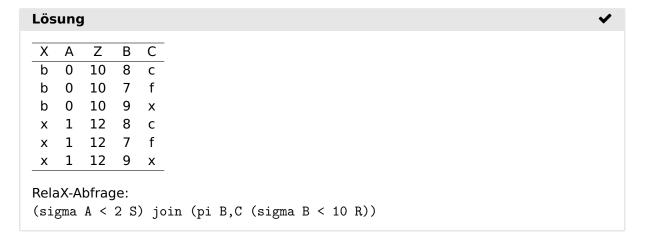


b) 0.5 Punkte Berechnen Sie $R\bowtie_{R.B=S.Z} S$.

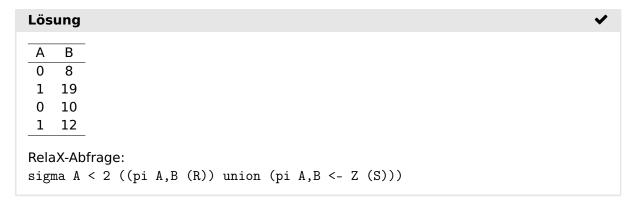


c) 0.5 Punkte Berechnen Sie $(\sigma_{A>0}S)\bowtie (\sigma_{B<10}R).$

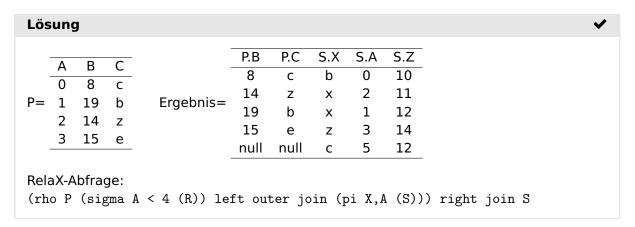
Lös	Lösung							
X	Α	Z	В	С	-			
С	5	12	7	f	- -			
	<pre>RelaX-Abfrage: (sigma A > 0 (S)) join (sigma B < 10 (R))</pre>							



e) 1 Punkt Berechnen Sie $\sigma_{A<2}\left((\pi_{A,B}R)\cup(\pi_{A,B\leftarrow Z}S)\right)$.



f) 1 Punkt Berechnen Sie $(\rho_P(\sigma_{A<4}(R))\bowtie(\pi_{X,A}S))\bowtie S$.



Aufgabe 2 (Relationale Algebra)

[6 Punkte]

A

Gegeben sei das folgende Relationenschema:

```
Customer (<u>CustomerId</u>, FirstName, LastName, Address, Email)
InvoiceParts (<u>InvoicePartId</u>, InvoiceId, TrackId, UnitPrice, Quantity)
Invoice (<u>InvoiceId</u>, CustomerId, InvoiceDate, Total)
Genre (<u>GenreId</u>, Name)
Playlist (<u>PlaylistId</u>, Name)
PlaylistContent (<u>PlaylistId</u>, TrackId)
Artist (<u>ArtistId</u>, Name)
Album (<u>AlbumId</u>, Title, ArtistId)
Track (<u>TrackId</u>, Name, AlbumId, GenreId, Miliseconds, Bytes, UnitPrice)
```

Erstellen Sie auf Basis dieses Relationenschemas die folgenden Anfragen in relationaler Algebra. Sie können dazu RelaX (einen Rechner für relationale Algebra) verwenden. Überlegen Sie sich jedoch trotzdem, wie man die Operationen "händisch" berechnen würde. Mit folgendem Link ist RelaX inklusive des für diese Aufgabe benötigten Schemas und den enthaltenen Daten erreichbar: https://dbis-uibk.github.io/relax/calc.htm?data=gist:e8628d74e467b945a564d27d4d74b83e. Geben Sie für die folgenden Aufgaben **sowohl die Abfrage als auch das Ergebnis** ab. Bei sehr großen Ergebnismengen geben Sie bitte die ersten zehn Zeilen und die Gesamtanzahl der Tupel in der Ergebnismenge an.

Hinweis

Sie können im RelaX-Tool Zwischenergebnisse einer Variable zuweisen und später auf diese zugreifen:

Result1 = Artist join (Artist.ArtistId = Album.AlbumId) Album pi Name, Title Result1

a) 1 Punkt Geben Sie bitte die Id, die Customerld, das InvoiceDate und die Gesamtsumme (Total) aller Rechnungen aus, die eine Gesamtsumme von über 10 Euro aufweisen.

Lösung 🗸

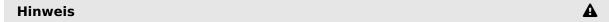
pi InvoiceId, CustomerId, InvoiceDate, Total (sigma Total > 10 (Invoice))

InvoiceId	CustomerId	InvoiceDate	Total
82	28	2009-12-18	13.86
138	37	2010-08-23	13.86
180	29	2011-02-25	13.86
187	8	2011-03-28	13.86
243	17	2011-12-01	13.86
285	9	2012-06-04	13.86
292	47	2012-07-05	13.86
348	56	2013-03-10	13.86
390	48	2013-09-12	13.86
397	27	2013-10-13	13.86

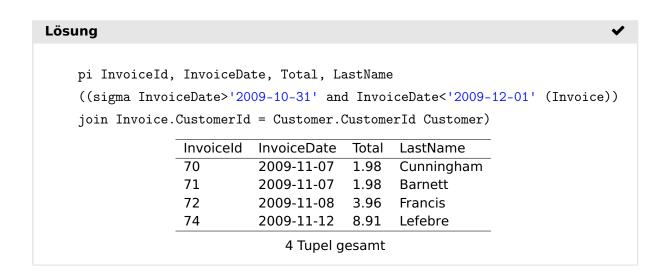
10 Tupel gesamt^a

^aBeachten Sie bitte, dass bei den Musterlösungen die Spaltennamen der angeführten Lösungstabellen aus Gründen der Übersichtlichkeit abgekürzt sind (so wird z.B. aus Invoice.Invoiceld lediglich Invoiceld).

b) 1 Punkt Geben Sie bitte alle Rechnungen aus, die mit November 2009 datiert sind. Dabei sollten die Id und das Datum der Rechnung, die Gesamtsumme und der Nachname des Kunden ausgegeben werden.



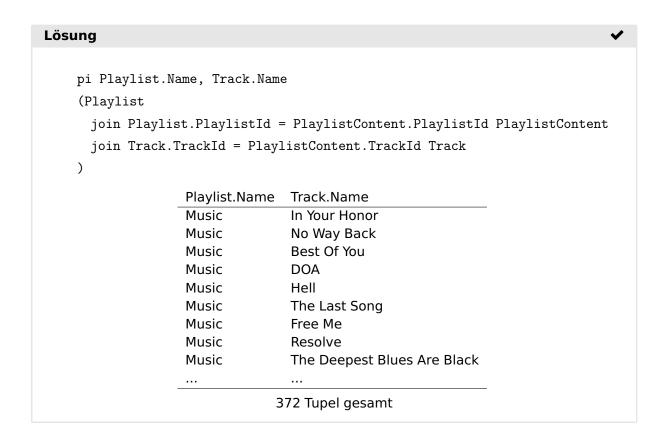
Datumseinträge werden im Format YYYY-MM-DD gespeichert und angegeben und können mit <, >, etc. verglichen werden. (z.B. birthday > '1934-01-01')



c) 1 Punkt Finden Sie alle Tracks des Genres Rock, die auch tatsächlich gekauft wurden. Geben Sie dazu den Namen des Tracks und dessen Id aus.



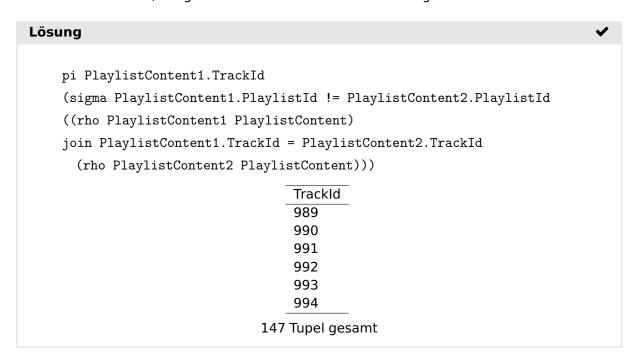
d) 1 Punkt Finden Sie für die vorhandenen Playlists die enthaltenen Tracks. Geben Sie dazu bitte den Namen des Tracks und der Playlist aus.



e) 1 Punkt Geben Sie für alle Kunden, deren Nachname mit 'A' oder 'B' beginnt, die von ihnen gekauften Lieder (Track Name, Artist Name, Album Name) aus und führen Sie auch den Vorund Nachnamen des Kunden an.



f) 1 Punkt Finden Sie alle Tracks, die in zwei verschiedenen Playlists vorkommen. In einem ersten Schritt reicht es, lediglich die Trackld dieser Tracks auszugeben.



Wichtig: Laden Sie bitte Ihre Lösung in OLAT hoch und geben Sie mittels der Ankreuzliste auch unbedingt an, welche Aufgaben Sie gelöst haben. Die Deadline dafür läuft am Vortag des Proseminars um 23:59 (Mitternacht) ab.