



Übung zur Vorlesung *Einsatz und Realisierung von Datenbanksystemen* im SoSe18

Alexander van Renen, Maximilian E. Schüle (i3erdb@in.tum.de)
<http://db.in.tum.de/teaching/ss18/impldb/>

Blatt Nr. 06

Hausaufgabe 1 Fensterln mit SQL

Analysieren wir die Gehälter von Professoren mittels Windowfunctions und führen Sie die Abfragen unter hyper-db.de aus. Dazu orientieren wir uns an der Relation *Professoren* des erweiterten Universitätschemas:

```
with Professoren (persnr, name, rang, raum, gehalt, steuerklasse) as (  
  values (2125, 'Sokrates', 'C4', 226, 85000, 1),  
         (2126, 'Russel', 'C4', 232, 80000, 3),  
         (2127, 'Kopernikus', 'C3', 310, 65000, 5),  
         (2128, 'Aristoteles', 'C4', 250, 85000, 1),  
         (2133, 'Popper', 'C3', 52, 68000, 1),  
         (2134, 'Augustinus', 'C3', 309, 55000, 5),  
         (2136, 'Curie', 'C4', 36, 95000, 3),  
         (2137, 'Kant', 'C4', 7, 98000, 1)  
)
```

1. Ermitteln Sie zu jedem Professor das Durchschnittsgehalt aller Professoren.
2. Ermitteln Sie zu jedem Professor das Durchschnittsgehalt aller Professoren partitioniert nach Rang.
3. Ermitteln Sie nun die wachsende Summe (das Quantil) des Gehaltes aller Professoren partitioniert nach Rang und absteigend sortiert nach ihrem Gehalt. Gleich verdienende Professoren sind im selben Quartil.
4. Ermitteln Sie nun die wachsende Summe des Gehaltes aller Professoren partitioniert nach Rang und absteigend (total) sortiert nach ihrem Gehalt (reihenweise, nicht als Range-Query).
5. Ermitteln Sie nun das gleitende Durchschnittsgehalt aus genau zwei mehr bzw. weniger verdienenden Professoren sortiert nach Gehalt und partitioniert nach Rang.
6. Ermitteln Sie nun das gleitende Durchschnittsgehalt aus den 500 Einheiten mehr bzw. weniger verdienenden Professoren sortiert nach Gehalt und partitioniert nach Rang.
7. Geben sie zu jedem Professor das Gehalt des eins besser wie eins schlechter verdienenden.
8. Ermitteln Sie die drei bestverdienendsten Professoren einmal mit und einmal ohne Windowfunctions.

Hausaufgabe 2 Lösen Sie folgende Anfrage mit SQL basierend auf dem bekannten Universitätschema.

1. Bestimmen Sie die Durchschnittsnote für jeden Studenten.
2. Basierend auf dieser Durchschnittsnote, bestimmen Sie für alle Studenten ihren Rangplatz innerhalb ihrer Kohorte (Studenten desselben Semesters).
3. Berechnen Sie zusätzlich für jeden Studenten auch noch die **Abweichung** seiner Durchschnittsnote von der Durchschnittsnote der Kohorte (also vom Durchschnitt der Durchschnittsnote der Studenten der Kohorte) ausgegeben werden.

Lösen Sie Teilaufgaben 2 und 3 jeweils einmal mit und einmal ohne Nutzung von Windowfunktionen. Ihre Anfragen können Sie auf hyper-db.de testen. Nutzen Sie folgende erweiterte *pruefen* Relation:

```
with mehr_pruefen(MatrnNr, VorlNr, PersNr, Note) as (
  select * from pruefen
  union
  values (29120,0,0,3.0), (29555,0,0,2.0), (29555,0,0,1.3), (29555,0,0,1.0)
)
```

Hausaufgabe 3 Größer Fensterln mit SQL in TPC-H

Wenden wir nun Fensterfunktionen an dem fiktiven TPC-H Schema an.

1. Erstellen Sie in SQL eine Abfrage nach dem jährlichen Exportvolumen pro Jahr und Land. Verwenden Sie diese Abfrage als Hilfstabelle (**with**-Statement) in den folgenden Abfragen, orientieren Sie sich an TPC-H Anfrage 7.
2. Ranken Sie Länder anhand ihres jährlichen Exportvolumens, auf Platz eins ist das Land mit dem höchsten Volumen, das je in einem Jahr getätigt worden ist.
3. Küren Sie nun die Jahressieger. Ranken Sie dazu die Länder partitioniert nach Jahr.
4. Ermitteln Sie nun das laufende Exportmittel der Länge drei (Vorjahr, Nachjahr, falls erfasst).

Hausaufgabe 4

Die in Abbildung 1 dargestellten Relationen Mietspiegel und Kindergarten dienen der Bewertung von Wohngebieten im Großraum München. Für eine junge Familie ist ausschlaggebend, wie hoch die Lebenshaltungskosten gemessen an zu zahlender Miete und zu entrichtender Gebühr für den Kindergarten im jeweiligen Wohnort ausfallen. Illustrieren Sie die Ausführung einer Top-1-Berechnung (zur Bestimmung des günstigsten Wohnorts) für eine junge Familie mit zwei Kindern. Zeigen Sie die phasenweise Berechnung des Ergebnisses jeweils mit dem Threshold- und dem NRA-Algorithmus.

Hausaufgabe 5

Geben die Relation Klausur:

MatrnNr	Vorbereitungszeit	Note
1	150	1.7
2	70	2.7
3	450	2.0
4	180	1.7
5	2500	1.3

Mietspiegel		Kindergarten		WohnLage	
Ort	Miete	Ort	Beitrag	Ort	Lage
Garching	800	Grünwald	-100	Grünwald	München-Süd
Ismaning	900	Unterföhring	0	Unterföhring	München-Nord
Unterföhring	1000	Bogenhausen	100	Ismaning	München-Nord
Nymphenburg	1500	Ismaning	200	Garching	München-Nord
Bogenhausen	1600	Garching	250	Bogenhausen	München-City
Grünwald	1700	Nymphenburg	300	Nymphenburg	München-City

Abbildung 1: Münchner Wohnlagen zur Berechnung der monatlichen Kosten für eine Familie.

- Formulieren Sie die Anfrage, die die MatrNr in der Skyline für die Attribute Vorbereitungszeit und Note erzeugt (kleiner ist jeweils besser) in SQL mit Hilfe des Skyline Operators.
- Formulieren Sie die Anfrage in SQL ohne Skyline Operator.
- Bestimmen Sie das Ergebnis der Anfrage.

Hausaufgabe (wird nicht in der Übung besprochen) Cubes mit Studenten

Erstellen Sie anhand des Universitätsschemas einen Cube mit der Faktentabelle *mehr_pruefen* und den Dimensionstabellen *Vorlesungen*, *Professoren*, *Studenten*. Zählen Sie das Auftreten.

```
with mehr_pruefen(MatrnNr,VorlNr,PersNr,Note) as (
  select * from pruefen
  union
  values (28106,5041,2137,3.0)
)
```