

8. SELECT mit GROUP BY und HAVING

Die Gruppierung erzeugt eine Relation, in der Attribute auftreten, die für jede Gruppe einen Wert haben. Bis auf die Gruppierungsattribute selbst, deren Werte sich in der Originalrelation befinden, werden die Werte zur Laufzeit der Abfrage erst berechnet. Mit der zusätzlichen HAVING-Klausel kann nun eine Selektion nach der Gruppenbildung vorgenommen werden. Das bedeutet, dass eine Selektion unter den Gruppen erfolgt.

Für jeden Artikel hatten wird die Anzahl aller Bestellungen so ermittelt:

```
SELECT F_Artikel_Nr, SUM (Bestellmenge)
FROM T_Positionen
GROUP BY F_Artikel_Nr;
```

Die Auswertung zeigt, dass es Artikel gibt, die mehr als zehnmal bestellt wurden, und solche, die weniger oft geordert wurden. Wollen wir nur Artikel mit Bestellmengen von über zehn berücksichtigen, so muss die obige SQL-Anweisung um eine HAVING- Klausel ergänzt werden, nämlich:

```
... HAVING SUM(Bestellmenge) > 10
```

Die Grundform der HAVING-Klausel in Verbindung mit einer GROUP BY-Klausel lautet:

```
GROUP BY attributliste
HAVING bedingung
```

Da die HAVING-Klausel die Ergebnisse der Gruppierung selektiert, enthält sie meist selbst eine Aggregatfunktion.

Beispiele

```
... HAVING COUNT(*) > 1
... HAVING SUM (Liefermenge) < SUM(Bestellmenge)
```

Die WHERE-Klausel und die HAVING-Klausel können auch in Kombination eingesetzt werden.

Dabei gilt die syntaktische Reihenfolge:

```
WHERE
GROUP BY
HAVING
```

Dies entspricht der logischen Abfolge:

1. Zunächst wird mit der WHERE-Klausel eine Selektion auf der Basisrelation ausgeführt.
2. Die verbleibenden Tupel werden gruppiert.
3. Zum Schluss werden die Gruppen selektiert.

Die Kombination von WHERE und HAVING zeigt das folgende Beispiel, in dem die Anzahl der Bestellpositionen ermittelt wird, in denen ein Artikel auftritt. Vor der Gruppierung werden aber die Tupel selektiert, bei denen die Artikelnummer mit „K“ anfängt.

```
SELECT F_Artikel_Nr, COUNT(*)
FROM T_Positionen
WHERE F_Artikel_Nr LIKE 'K%'
GROUP BY F_Artikel_Nr
HAVING COUNT (*) > 1
```

A08-1

Wie viele Ihrer Kunden haben jeweils die gleiche Postleitzahl?

A08-2

Unter welcher Postleitzahl haben Sie mehr als einen Kunden?

A08-3

Zu welcher Zahlungsart haben Sie mehr als zwei Kunden?

Reihenfolge der Komponenten der SELECT-Anweisung

Da nunmehr alle Klauseln der SELECT-Anweisung vorgestellt worden sind, kann ein Resümee gezogen werden.

Syntaktisch gesehen sind die einzelnen Klauseln der SELECT-Anweisung zwingend in folgender Reihenfolge anzuordnen:

1. SELECT
2. FROM
3. WHERE
4. GROUP BY
5. HAVING
6. ORDER BY

Jede andere Reihenfolge führt zu einer Fehlermeldung.

Von der logischen Bearbeitung her gilt allerdings etwas anderes:

Die FROM-Klausel wird zuerst ausgewertet. Damit stehen die Relationen fest, auf die sich die Abfrage richtet. Als nächstes folgt die Selektionsklausel WHERE. Es schließt sich die Gruppierung aufgrund der GROUP BY-Klausel an, gefolgt von der Gruppenselektion mit HAVING. Zum Schluss werden die Attribute der Ergebnisrelation mit der SELECT-Klausel festgelegt. Etwaige Berechnungen von virtuellen Spalten werden also auch erst dann vorgenommen, wenn feststeht, welche Tupel die Ergebnisrelation umfasst.

Die technische Abarbeitung einer SELECT-Anweisung kann durchaus von der hier gezeigten logischen Abfolge abweichen.