

Datenbanksysteme I  
Anfragebearbeitung und  
-optimierung

29.6.-1.7.2009

Jana Bauckmann

# VL: Kommutativität und Assoziativität

2

- $\times$  ist kommutativ und assoziativ
  - $R \times S = S \times R$
  - $(R \times S) \times T = R \times (S \times T)$
- $\cup$  ist kommutativ und assoziativ
  - $R \cup S = S \cup R$
  - $(R \cup S) \cup T = R \cup (S \cup T)$
- $\cap$  ist kommutativ und assoziativ
  - $R \cap S = S \cap R$
  - $(R \cap S) \cap T = R \cap (S \cap T)$
- $\bowtie$  ist kommutativ und assoziativ
  - $R \bowtie S = S \bowtie R$
  - $(R \bowtie S) \bowtie T = R \bowtie (S \bowtie T)$

Gilt jeweils für Mengen  
und Multimengen

Ausdrücke können in beide  
Richtungen verwendet werden.

## Selektion

$$\sigma_{c_1 \text{ AND } c_2}(R) = \sigma_{c_1}(\sigma_{c_2}(R))$$

$$\sigma_{c_1 \text{ OR } c_2}(R) = \sigma_{c_1}(R) \cup \sigma_{c_2}(R)$$

- Nicht bei Multimengen

$$\sigma_{c_1}(\sigma_{c_2}(R)) = \sigma_{c_2}(\sigma_{c_1}(R))$$

$$\sigma_c(R \Phi S) \equiv (\sigma_c(R)) \Phi (\sigma_c(S))$$

- $\Phi \in \{\cup, \cap, -, \bowtie\}$

$$\sigma_c(R \Phi S) \equiv (\sigma_c(R)) \Phi S$$

- $\Phi \in \{\cup, \cap, -, \bowtie\}$

- Falls sich c nur auf Attribute in R bezieht.

## Projektion

$$\pi_L(R \bowtie S) = \pi_L(\pi_M(R) \bowtie \pi_N(S))$$

$$\pi_L(R \bowtie_C S) = \pi_L(\pi_M(R) \bowtie_C \pi_N(S))$$

$$\pi_L(R \times S) = \pi_L(\pi_M(R) \times \pi_N(S))$$

$$\pi_L \sigma_C(R) = \pi_L(\sigma_C(\pi_M(R)))$$

# Aufgabe: Projektionen pushen

4

Untersuchen Sie

- $\pi_L(R \cup S) = (\pi_L(R)) \cup (\pi_L(S))$ 
  - Für Mengen-Vereinigung
  - Für Multimengen-Vereinigung
  
- $\pi_L(R - S) = (\pi_L(R)) - (\pi_L(S))$ 
  - Für Mengen-Differenz
  - Für Multimengen-Differenz

Selektivität schätzt Anzahl der qualifizierenden Tupel relativ zur Gesamtanzahl der Tupel in der Relation.

Projektion:

- $sf = |R| / |R| = 1$

Selektion:

- $sf = |\sigma_C(R)| / |R|$

Join:

- $sf = |R \bowtie S| / |R \times S| = |R \bowtie S| / (|R| \cdot |S|)$

## Selektion:

- Selektion auf einen Schlüssel:
  - $sf = 1 / |R|$
- Selektion auf einen Attribut A mit m verschiedenen Werten:
  - $sf = (|R| / m) / |R| = 1/m$
  - Dies ist nur geschätzt!

## Join

- Equijoin zwischen R und S über Fremdschlüssel in S
  - $sf = 1/ |R|$
  - „Beweis“:  $sf = |R \bowtie S| / (|R| \times |S|) = |S| / (|R| \cdot |S|)$

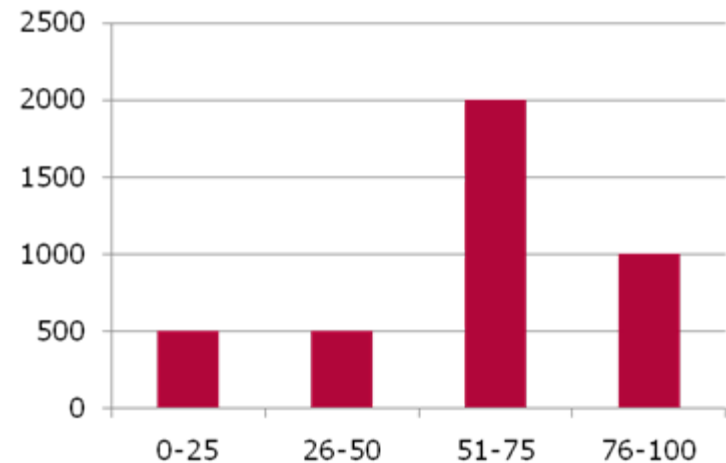
# Aufgabe: Ergebniskardinalität abschätzen

7

- Zulieferer (zid, name, adresse) 10 Tupel
- Teile (tid, bezeichnung, farbe) 1000 Tupel
- Katalog (zid, tid, kosten) 4000 Tupel

- farbe in {schwarz, rot, blau, weiß} gleichverteilt
- kosten in [1,100] gleichverteilt

- $\sigma_{\text{farbe}=\text{'schwarz'} \vee \text{farbe}=\text{'blau'}}((\text{Zulieferer} \bowtie \sigma_{\text{kosten}<25}(\text{Katalog})) \bowtie \text{Teile})$

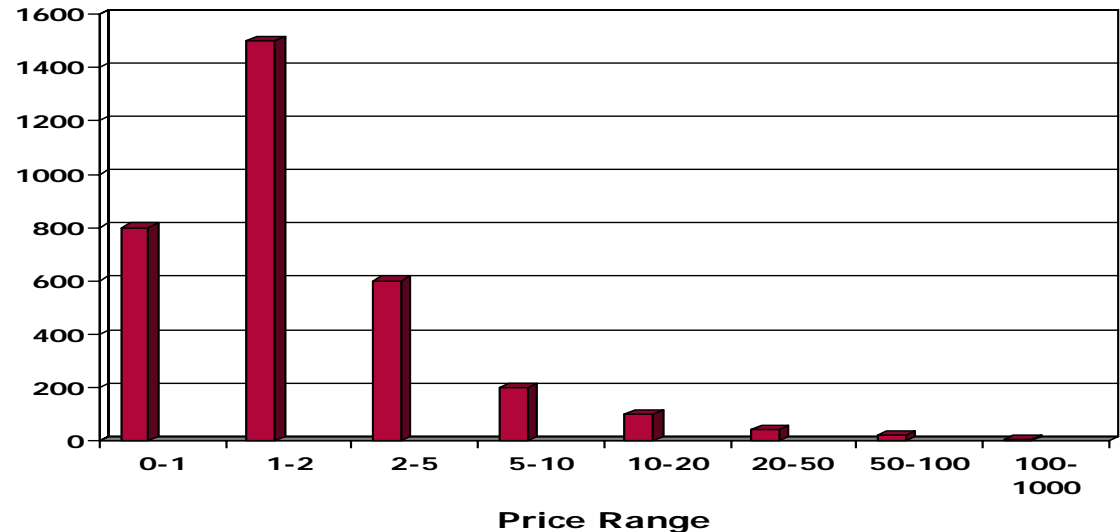


# VL: Beispiel zu Histogrammen

8

Frequency

```
SELECT *
FROM product p, sales S
WHERE p.id=s.p_id and
      p.price > 100
```



Gegeben 3300 products, 1M sales

Gleichverteilung

- Preisspanne ist 0-1000 => Selektivität der Bedingung ist 9/10
  - Erwartet:  $9/10 * 3300 \approx 3000$  Produkte

Histogramm-basiert

- Angenommen 10 equi-width buckets
- Selektivität der Bedingung ist  $5/3300 \approx 0,0015$  also 5 Produkte