

Databasesystemer, Forår 2005

Rasmus Pagh

April 19, 2005

Øvelser d. 21/4

Øvelserne er maskinøvelser, og falder i to dele: Første del af øvelserne handler om OLAP. Der skal arbejdes med salgsdata, der findes i to formater: En Excel fil (link i skema) og en SQL fil, der opretter en tabel med data i Oracle. Sidstnævnte findes i `data` kataloget under kursushjemmesiden (kopier ind i SQL*Plus eller TOra):

<http://www.itu.dk/people/pagh/DBS05/data/OLAP-table.sql>

1. I Excel er der skabt en såkaldt pivottabel ud fra data. Pivottabellen i filen viser salg fordelt på lokationer. Kun salg hvor værdien af HOW er "Store" medtages. Dette svarer i datacube terminologi til, at vi kigger på en slice af HOW dimensionen svarende til værdien "Store". I denne slice er der lavet drill down i WHAT og WHERE dimensionerne, således at vi får aggregatet (antal solgte enheder) for hver vare og lokation.
2. Ved hjælp af "drag and drop" kan pivottabellen laves om. Få først vist en oversigt over salget af kamelsadler, hvor der er lavet "drill down" på HOW og WHERE dimensionerne. Undersøg hvad der sker, hvis der laves drill down på en tredje dimension (WHEN).
3. Data i tabellen OLAP er rå "fact" data, dvs. ikke i den anbefalede star schema form. Din opgave er at oprette WHAT, HOW og WHERE som dimensionstabeller, og lave en ny version af OLAP tabellen, der refererer til disse dimensionstabeller. For at få en kompakt, unik nøgle for hver tupel i dimensionstabellerne, kan de første 3 tegn i strengen benyttes. (Generelt bør man benytte en "dum" nøgle, der er uafhængig af data, men denne simple metode rækker i dette tilfælde.)
4. Lav SQL forespørgsler på stjerneschemaet, der beregner de samme data som pivottabellerne skabt i Excel.

Anden del er opsamling af øvelserne fra 7/4, der som bekendt løb ind i tekniske problemer. Problemerne skulle være løst nu, således at øvelserne kan udføres.