Homework 2 (Due Feb. 17th, 2014)

1. Read paper: “One Size Fits All”: An idea whose time has come and gone, by M. Stonebraker et al. ICDE’05, and answer the following questions:
	1. What are the key implementation/optimization techniques for OLTP and OLAP?
	2. What are the challenges for stream processing? Why existing relational DBMS cannot handle it well?
	3. (Bonus question) Why **high availability** requirement has conflicts with **ACID transaction management**?
2. Describe multi-dimensional data model. Explain the distinction between measures and dimensions and between fact tables and dimension tables. What is a star schema?
3. Common OLAP operations have received special names: roll-up, drill-down, pivoting, slicing and dicing. Describe each of these operations and illustrate them using examples (from the slides).
4. Bonus Question: What is cube operator in OLAP?

Paper: Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Table, and Sub-Total, by Jim Gray, et al. 1995.

1. Considering we have the following schema:



* 1. Show the result of pivoting the relation on prodid and date (with measure qty)
	2. Show the result of roll-up on storied.
	3. Show the result of roll-up on storid followed by drill-down on prodid.
	4. Write a collection of SQL queries for a, b, & c.
1. Given a fact table (i.e. sale table format from Question 5):
	1. Write a SQL query with group-by on storid and prodid for total of quality.
	2. Implement a Java program to do the computation.

(input: the sale table as a file where each row is on a line, separated by “,” or “ “).

Please use two data structures to implement this: one is Hash-table, and another is Array (vector).

* 1. Write a program to generate a bitmap-index for storid and prodid.
	2. Explain how bitmap-index can help speedup the OLAP computation in general?