

Computer Science COMP3420 (2009) – Tutorial Two

Question 1. A data warehouse can be modeled by either a *star schema* or a *snowflake schema*. Briefly describe the similarities and the differences of the two models, and then analyze their advantages and disadvantage with regard to one another. Give your opinion of which might be more empirically useful and state the reasons behind your answers.

Question 2. Suppose a group of 12 *sales price* records has been sorted as follows.

5, 10, 11, 13, 15, 35, 50, 55, 72, 92, 204, 215

Partition them into three bins by each of the following methods.

- (a) equal-frequency (equidepth) partitioning
- (b) equal-width partitioning

Question 3. What are the distinctions between OLAP and OLAM?

Question 4. What is the multidimensional data model? What are the OLAP operations?

Question 5. What is the three-tier data warehouse architecture? Can you describe the basic components of the architecture and the connection of these components?

Question 6. What is the meta-data? Can you list three common types of meta-data?

Question 7. Why most data warehouses stored lots materialized views? what is the benefit by doing so? What are the possible challenges to maintain these materialized views?

Question 8. What are the differences between the three main types of data warehouse usages: *information processing*, *analytical processing* and *data mining*? Discuss the motivation behind OLAP mining (OLAM).