WEEK 12. Data Warehouses – Chapter 16, 17 Self-Check Questions

1. Which one of the following applications/activities is typical of a data warehouse?
   a. Updating of the inventory data as the sales occur in the supermarket
   b. **Comparing last years sales with this year sales to identify the most promising product**
   c. Looking at the sales this month to identify the most valuable salesperson of the month
   d. Reviewing last month customer accounts to prepare overdue notices

*B will make the most sense when utilizing a data warehouse to answer a decision making question.* (pg560)
2. Which of the following may occur in a star schema?
   a. Many to many relationships
   b. **One to many relationships**
   c. M-way relationships
   d. Recursive relationships

   *Recursive relationships are those that relate two or more entities of the same type.*
   *The star schema and its variations require 1-M relationship from dimensions table to fact table.* (pg608). *In some cases, it may seem necessary to use M-N relationship.*
3. Which clause of SQL SELECT statement may use the CUBE operations?
   a. GROUP BY
   b. HAVING
   c. Both of the above

   *Only by GROUP BY clause. The cube operator clause produces all possible subtotal combinations in addition to the normal totals show in a GROUP BY clause.* (pg 649)
4. Which of the following statements is not true about populating a data warehouse:
   a. It involves matching decision support needs with available data
   b. **It is a matter of simply copying data from various sources**
   c. Sources of data may be internal or external or both
   d. All of the above

   *Data warehouse has information that can be critical to decision making from either internal or external sources.*
5. In maintaining a data warehouse, which of the following is **not** a task in the PREPARATION phase:
   a. Extraction
   b. **Merging**
   c. Transportation
   d. Cleaning

Preparation phase can be perform multiple times and in any order to transform ‘dirty’ data into some kind of standardized data across users. (pg 636)
Discussion Questions
6. Why are operational databases not particularly suited for decision support applications?

Operational databases are mostly intended to be used for real-time data manipulations such as tracking and retrieving activities/data of a company therefore it cannot be used to manipulate data from the past or predict future data.