

Data Warehousing

CS-452

Spring 2023

Activity 9

Question 1:

Download and import [AdventureWorksDW2012](#) in the MS SQL Server. If you want you can use this [script](#) to create AdventureWorksDW2012.

Question 2: The AdventureWorksDW2012 Internet Sales Schema

Create database diagram using MS Server Management Studio. Use the following Tables to create diagram.

DimCustomer
DimDate
DimGeography
DimProduct
DimProductCategory
DimProductSubcategory
FactInternetSales

Question 3:

Review the AdventureWorksDW2012 schema and answer the following facts:

- The DimDate dimension has _____ additional lookup tables associated with it and therefore uses the _____ schema.
- The DimProduct table is _____ it uses the _____ lookup table, which further uses the DimProductCategory lookup table.
- What tables are connected to DimCustomer dimension. Draw the diagram.
- List the name of tables that stores lineage information.

Question 4: The AdventureWorksDW2012 Internet Sales Schema

Create database diagram using MS Server Management Studio. Use the following Tables to create diagram.

DimProduct
DimProductCategory
DimProductSubcategory

Question 5:

Draw the database diagram using MS Server Management Studio by adding the DimSalesReason table to the previously created diagram in last question.

Question 6:

Using the previously created diagram, answer the following

- The DimProduct dimension has a natural hierarchy:
ProductCategory → _____ → Product.

- b. The DimProduct dimension has many additional attributes that are useful for _____ but that are not a part of any natural hierarchy. For example, Color and Size are such attributes.
- c. Some columns in the DimProduct dimension, such as the LargePhoto and Description columns, are _____ (respect to dimension column type).
- d. DimSalesReason uses a _____ schema.
- e. In the DimSalesReason dimension, it seems that there is a natural hierarchy: SalesReasonType → _____.

Question 7: The AdventureWorksDW2012 Internet Sales Schema

Create database diagram using MS Server Management Studio. Use the following Tables to create diagram.

DimProduct

DimDate

FactProductInventory

Question 8: Analyze Fact Table Columns

Learn more details about the fact table in the schema you created in the previous question. Note that you have to conclude these details from the names of the measure columns; in a real-life project, you should check the content of the columns as well.

- a. Knowing how an inventory works, you can conclude that the UnitsIn and UnitsOut are additive measures. Using the SUM aggregate function for these two columns is reasonable for aggregations over any dimension.
- b. The UnitCost measure is a non-additive measure. Summing it over any dimension does not make sense.
- c. The UnitsBalance measure is a semi-additive measure. You can use the SUM aggregate function over any dimension but time.

Question 9: Case Scenario 1: A Quick POC Project

You are hired to implement a quick proof of concept (POC) data warehousing project. You have to prepare the schema for sales data. Your customer's subject matter experts (SME) would like to analyze sales data over customers, products, and time. Before creating a DW and tables, you need to make a couple of decisions and answer a couple of questions:

- i. What kind of schema would you use?
- ii. What would the dimensions of your schema be?
- iii. Do you expect additive measures only?

Question 10: Case Scenario 2: Extending the POC Project

After you implemented the POC sales data warehouse in Case Scenario 1, your customer was very satisfied. In fact, the business would like to extend the project to a real, long-term data warehouse. However, when interviewing analysts, you also discovered some points of dissatisfaction.

Interviews

Here's a list of company personnel who expressed some dissatisfaction during their interviews, along with their statements:

- Sales SME "I don't see correct aggregates over regions for historical data."
- DBA Who Creates Reports "My queries are still complicated, with many joins."

You need to solve these issues by answering the following questions

- i. How would you address the Sales SME issue?
- ii. What kind of schema would you implement for a long-term DW?
- iii. How would you address the DBA's issue?

Note:

Prepare a word document to upload using snapshot and written explanation.
MS Word file name should be as follows (Don't change the format).

19B-001-SE_Your_Name_Activity09_Group#

If you use any website, book, or research paper provide the link(s).

Make sure to upload prepared activity Word Document, script, package and dataset files or any other supporting files.