Activity Overview

In this activity, you will examine a schema and evaluate it using a schema validation checklist. You will complete this step in your role as a business intelligence professional when you need to ensure that data is properly transformed and moved into a database.

Scenario

Review the following scenario. Then complete the step-by-step instructions.

Your coworker moved data from several tables into your database using a data pipeline. You notice that there are some inconsistencies between the source data and the data in its destination (your database). In this activity, you will use a schema validation checklist to evaluate the schema and figure out where the inconsistencies are. Then, you will answer some quiz questions.

Step-By-Step Instructions

Follow the instructions to complete the activity and answer the quiz questions that follow. After you complete the quiz, you can compare your answers to the feedback provided.

Step 1: Access the schema

To use the schema for this course item, click the following link and select *Use Template*.

Link to schema: Database schema

OR

If you don't have a Google account, you can download the template directly from the following attachment.

Activity Template_ Database schema
PPTX File

Step 2: Review the schema

Review the schema's tables and column titles. This database schema contains eight tables: **Sales Fact, Shipments, Billing, Order Items, Product, Product Price, Order Details,** and **Customer**, which are connected via keys.

The central table is **Sales Fact**. The foreign keys in the **Sales Fact** table link to the other tables as follows:

- "order_sid" key links to the **Order Items, Order Details, Shipments**, and **Billing** tables
- "customer_sid" links to **Order Details**; "order_item_sid" links to **Order Items, Shipments**, and **Billing**
- "shipment_sid" links to **Shipments**; and "billing_sid" links to **Billing**
- "product_id" from the **Product** table links to **Order Items** and **Product Price**

The **Customer** table currently doesn't have any links to other tables. It contains the following columns: "customer_sid," "customer_name," and "customer_type."

Step 3: Evaluate the schema

Go through each item on the following checklist to verify that the schema can be properly validated. If the schema fails one of the checks, determine the reason it failed. If you need help determining the source of a problem, pay close attention to the names of the dimensions/columns within each table.

- **The keys are still valid:** Primary and foreign keys build relationships between tables in relational databases. These keys should continue to function after you have moved data from one system into another.
- **The conventions are consistent:** The conventions for incoming data must be consistent with the target database's schema. Data from outside sources might use different conventions for naming columns in tables—it's important to align these before they're added to the target system.
- **The table relationships have been preserved:** The keys help preserve the relationships used to connect the tables so that keys can still be used to connect tables. It's important to make sure that these relationships are preserved or that they are transformed to match the target schema.

What to Include in Your Response

Be sure to address the following criteria in your completed activity:

- Verify that the schema's keys are still valid or determine the reason they aren't.
- Verify that the schema's table relationships have been preserved or determine the reason they weren't.
- Verify that the schema's conventions are consistent or determine the reason they aren't.