Unit 3 Research Project

Eddie S. Jackson

Kaplan University

IT526: SQL Query Design

Jon Walter McKeeby, DSc, MBA

07/08/2014

Unit 3 Research Project

Part 1

a. Which aggregate function would you use to determine the latest orderdate value? MAX()

b. Which of the following clauses is evaluated third: SELECT, FROM, WHERE, GROUP BY, HAVING, ORDER BY? GROUP BY is evaluated third.

c. When creating a table, what are the two properties that must be specified for each column? Column name and column data type

What property should be specified for each column? The default constraints such as NULL, NOT NULL, NOT NULL DEFAULT

d. What is the purpose of a database schema? A database schema provides a place to group database objects together. These objects could include tables, views, stored procedures, and other database objects.

e. How does the textbook define the "most efficient data type?" The most efficient data type is one that utilizes the least amount of disk storage, while at the same time capturing data that will not need future modification as data fills the table (Ben-Gan, Sarka, Talmage, 2013).

f. What is the purpose of a Primary Key constraint? The purpose of the primary key constraint is to provide a way to uniquely identify each row in a table.

g. State two differences between a Primary Key constraint and a Unique constraint. One main difference between primary keys and the unique constraint is primary keys cannot be NULL and UNIQUE can have one NULL value. A second difference between the two is that each table can contain only one primary key, however can have multiple unique constraints.

h. Why is it redundant to add a Unique constraint to a column defined as the Primary Key? The very nature of primary keys is that primary keys are unique in themselves, thus adding the UNIQUE constraint to an already unique primary key is redundant.

i. What may a Foreign Key reference? A foreign may reference a parent table, specifically a primary key in another table.

j. In what namespace (database, schema, or table) must a constraint name be unique? Constraint names must be unique in the schema namespace.

Part 2

a. Describe the characteristics a column name must have in order for it to be a regular identifier. There are three main rule sets when considering the naming convention for columns. One, the first character of a column name must be in the alphabetic range of A-Z, a-z, an at sign (@), an underscore (_), or a number sign (#). Two, all subsequent characters can include an at sign (@), an underscore (_), or a number sign (#), decimal numbers, and letters. And finally, the column identifier cannot have spaces, not include extra characters, and cannot be a reserved T-SQL keyword (Ben-Gan, Sarka, & Talmage, 2013).

b. List all the ways that data values in a table column can be constrained. There are five table constraints; they are: Primary key constraint, Unique constraint, Foreign key constraint, Check constraint, and Default constraint.

c. What is the purpose of a Foreign Key constraint? The purpose of the Foreign Key constraint is to provide a link to lookup data in another table, which commonly referred to as a lookup table. The foreign key points to the Primary Key in another table.

d. Your company has a coding policy that Foreign Key constraints must be included in the CREATE TABLE statements. In what order would you create the following tables in the TSQL2012 database: Categories, Customers, Employees, OrderDetails, Orders, Products, Shippers, Suppliers? Due to the design specifications, the order would be Shippers, Employees, Customers, Orders, Suppliers, Categories, Products, and OrderDetails.

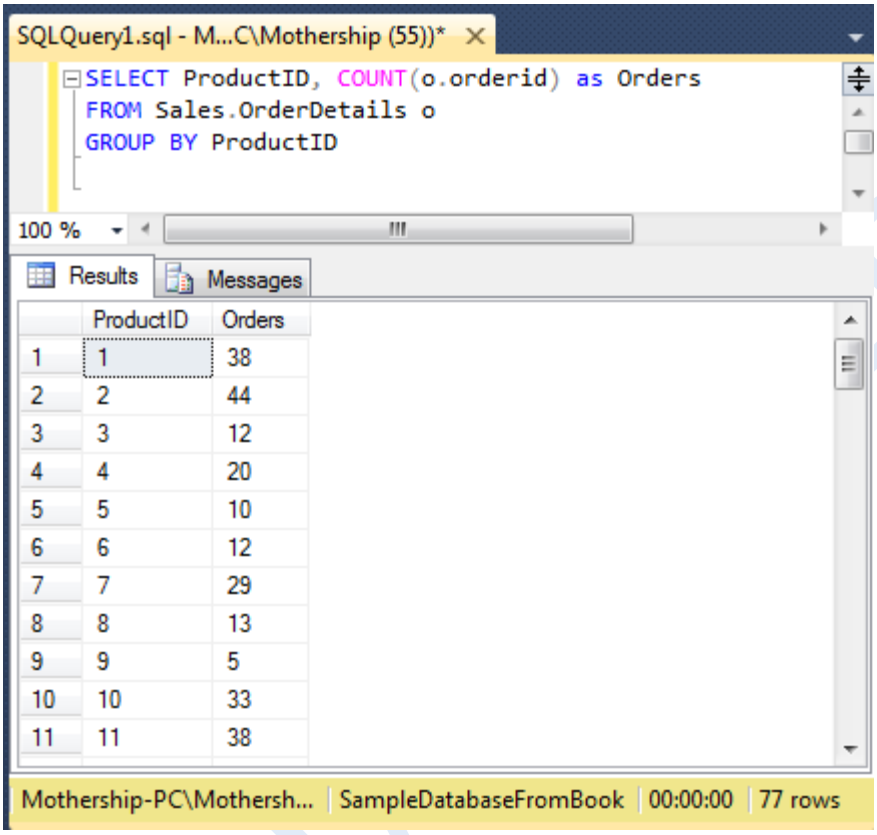
e. Define "data integrity." Cite at least two references, using in-text citation and a reference list. A citation of material from the textbook must give a page number or page range. Write approximately 100–200 words. At least 75% of what you write must be in your own words.

Data integrity is a concept that acts as a form of validation before data gets stored in tables. When considering data integrity, it is important to review the main types of data integrity that can be used to enforce the integrity of data. As Microsoft (n.d.) suggests, the four categories are entity integrity, domain integrity, referential integrity, and user-defined integrity (Microsoft, n.d.). Entity integrity is about ensuring that table rows can be uniquely identified. Domain integrity in the checking of entries in columns. Referential integrity is associated with maintaining the relationships between one table and the next as rows get modified, either through new entries or deletions. And finally, user-defined integrity is a customized category that allows the user to create rules outside of the other categories. It is important to note that each of these categories use certain constraints to enforce the integrity of data. For example, entity integrity can use non-NULL primary keys; referential integrity uses foreign keys (Mimer Developer, n.d.). Other common constraints include UNIQUE, CHECK, and Default (Ben-Gan, Sarka, & Talmage, 2013).

Part 3

For each query, give the problem number, copy the query from the query window into your Word Assignment document, then copy the results with headers into your Word document. If the result set has more than 10 rows, copy the first five rows with headers and state how many rows were returned.

1. Query the OrderDetails table to answer the following question: For each product, how many orders are there?



2a. Query the OrderDetails table to answer the following question: For each order, how many different products were ordered?

SQLQuery1.sql - M...C\Mothership (55))*

```
SELECT o.OrderID, count(o.productID) as Different_Products
FROM Sales.OrderDetails o
GROUP BY o.OrderID
```

100 %

Results | Messages

	OrderID	Different_Products
1	10248	3
2	10249	2
3	10250	3
4	10251	3
5	10252	3
6	10253	3
7	10254	3
8	10255	4
9	10256	2
10	10257	3
11	10258	3

(1... | Mothership-PC\Mothersh... | SampleDatabaseFromBook | 00:00:00 | 830 rows

2b. Which orders had more than one product ordered?

SQLQuery1.sql - M...C\Mothership (55))*

```
SELECT o.OrderID as OrdersGreaterThanOne
FROM Sales.OrderDetails o
GROUP BY o.OrderID
HAVING COUNT(o.OrderID) > 1;
```

100 %

Results | Messages

	OrdersGreaterThanOne
1	10248
2	10249
3	10250
4	10251
5	10252
6	10253
7	10254
8	10255
9	10256
10	10257
11	10258

Mothership-PC\Mothersh... | SampleDatabaseFromBook | 00:00:00 | 693 rows

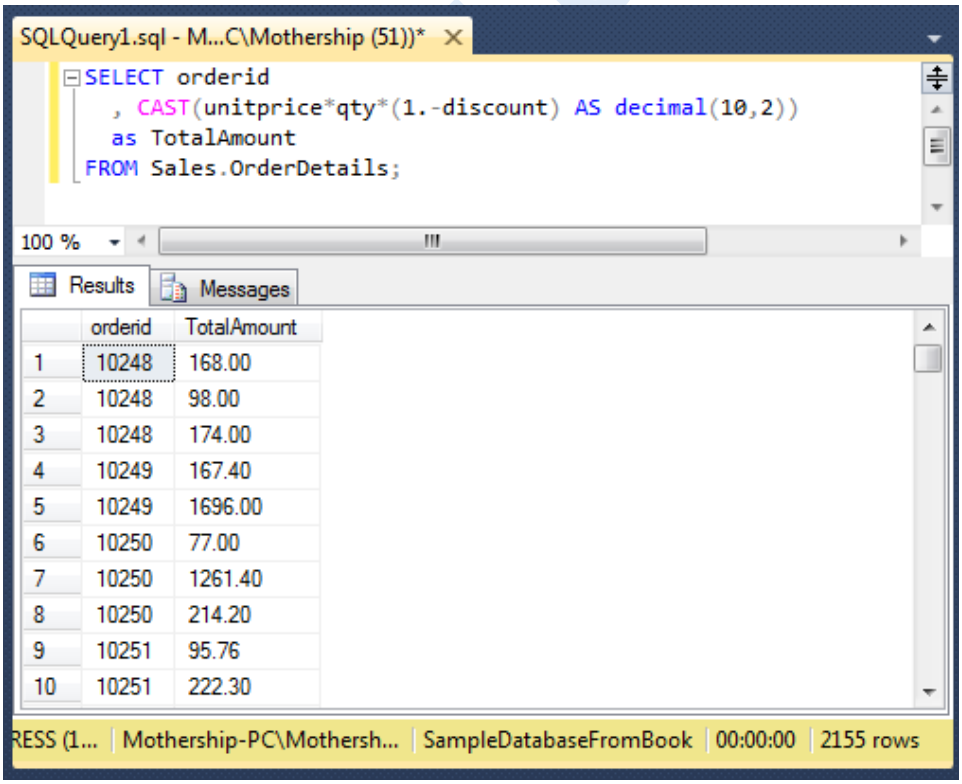
3. Write the T-SQL code that will create a table that will track horses. The table will be in database Farm, in schema Animal. The desired attributes are HorseID (an auto-generated integer), Name, Birthdate, Height, and Weight. The Height and Weight are for an adult horse; they might not be known for a young horse. Height will always be in inches and Weight will always be in pounds, so the data will be numbers.

```
USE Farm;
Go
CREATE TABLE Animal.Horses(
   HorseID   INT          NOT NULL IDENTITY,
   NAME      VARCHAR(30)  NOT NULL,
   Birthdate DATE         NOT NULL,
   Height    decimal(4,2) NULL,
   Weight    decimal(5,2) NULL,
   CONSTRAINT PK_Horse PRIMARY KEY(HorseID)
);
```

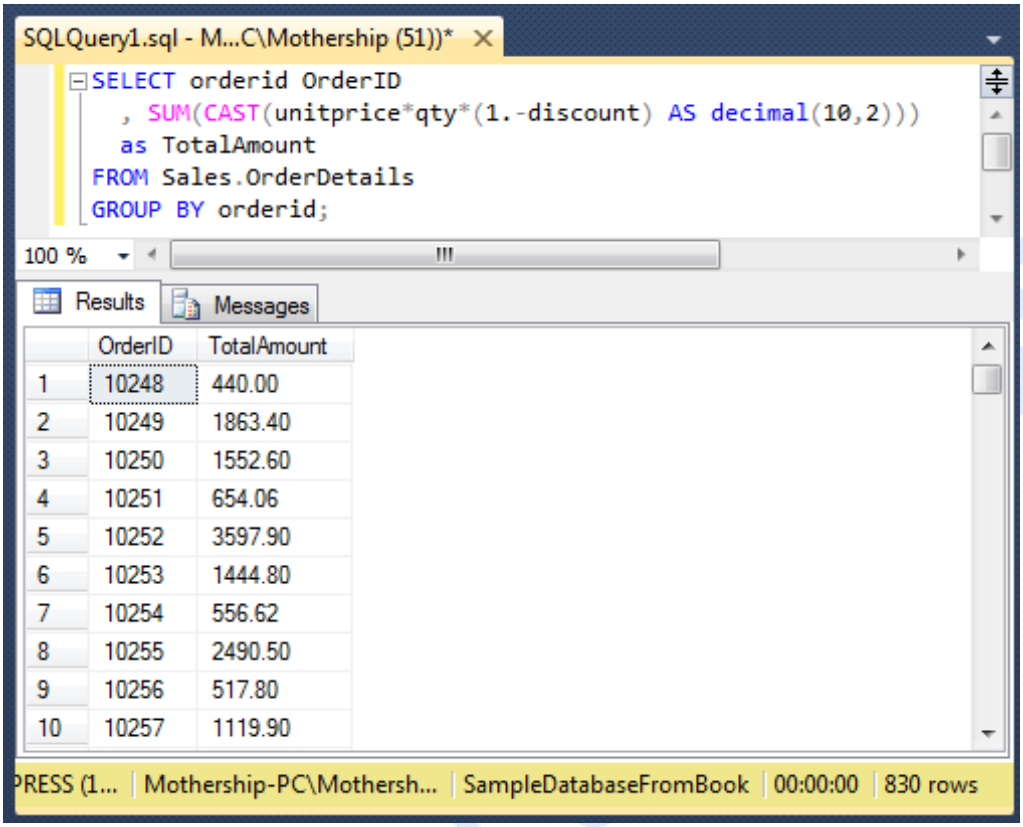
4. Write the T-SQL code that will add a column to the table in query 3 to store the name of the horse's sire.

```
ALTER TABLE Animal.Horses ADD Sire_Name VARCHAR(30) NULL;
```

5a. Join Orders to OrderDetails to return orderid and the total amount charged for each product. TotalAmount is unit price multiplied by qty multiplied by (1.-discount). TotalAmount should be cast as decimal(10,2). Do not perform any grouping.



5b. Modify the query in 5a to calculate, for each order, the total amount charged for the entire order. Result set columns should be titled OrderID and TotalAmount.



5c. Join Employees, Orders, and OrderDetails to return, for each employee, the total amount of sales. Return employee lastname, firstname, and total sales in descending order of total sales. Result set columns should be titled LastName, FirstName, and TotalSales.

SQLQuery1.sql - M...C\Mothership (51))*

```
SELECT LastName, FirstName
  , SUM(CAST(unitprice*qty*(1.-discount) AS decimal(10,2)))
  as TotalSales
FROM Sales.OrderDetails od
JOIN Sales.Orders ord ON od.orderid = ord.orderid
JOIN HR.Employees emp ON emp.empid = ord.empid
GROUP BY LastName, FirstName
ORDER BY TotalSales DESC;
```

100 %

Results | Messages

	LastName	FirstName	TotalSales
1	Peled	Yael	232890.89
2	Lew	Judy	202812.88
3	Davis	Sara	192107.67
4	Funk	Don	166537.76
5	Cameron	Maria	126862.30
6	King	Russell	124568.24
7	Dolgopyatova	Zoya	77308.09
8	Suurs	Paul	73913.15
9	Buck	Sven	68792.31

EXPRESS (1... | Mothership-PC\Mothersh... | SampleDatabaseFromBook | 00:00:00 | 9 rows

References

Ben-Gan, I., Sarka, D., & Talmage, R. (2013). *Training Kit (Exam 70-461): Querying Microsoft® SQL Server® 2012*. Microsoft Press.

Microsoft. (n.d.). Data Integrity. Retrieved from http://technet.microsoft.com/en-us/library/ms184276(v=sql.105).aspx

Mimer Developer. (n.d.). Data Integrity. Retrieved from http://developer.mimer.com/documentation/html_92/Mimer_SQL_Mobile_DocSet/DB_environment14.html