**1. FILTERING STRATEGIES.**  **Quite often you have to "down-select" from a large set of data items to then do further processing (including further filtering).**

**a) What kind of tools, techniques, or strategies can you apply to do this?**

The SELECT statement is a tool allows the user to down-select from a large set of data items for further processing. Some of the techniques include predicates and operators, such as WHERE and HAVING that allow the user to do further processing.

**b) Do they all work equally well, or do some have shortcomings or "hidden side-effects" you should consider when choosing which way to filter?**

All of the tools and techniques have some shortcomings. For example, HAVING must be applied after GROUP BY.

**2. DATA TYPES.**

**What is a "data type?"**

Data types tell the computer how much memory to allocate and what type of operations can be performed on the data. Regular data types use 1 byte of storage for each character and include CHAR and VARCHAR. Unicode data types use 2 bytes per character and include NCHAR and NVARCHAR.

**Why do we need to "type" a field?**

Efficient memory management.

**How does our use of data types in SQL database queries and statements differ from the way we use them in programming (in C#, for example)?**

C# data types are more varied in their functionality than SQL data types. For example, in C# you can use string, int, double, bool, etc depending on the operation. SQL does not involve as much programming so the data types are not as robust.

**What special considerations do you need when mixing types?**

Some data types have variable-length which means that SQL uses as much storage space in the row as required to store the characters that appear in the character string, plus two extra bytes for offset data. Other data types are fixed-length which means that SQL preserves space in the row based on the column’s defined size.

**What other "oddities" might occur?**

Because storage consumption for variable-length data types is less than that for fixed-length types, read operations are faster. However, updates might result in row expansion, which might result in data movement outside the current page. Therefore, updates of data having variable-length data types are less efficient than updates of data having fixed-length data types.

**3. CONSTRAINTS.**

**What are they, and how do we use them?**

Data integrity is achieved through rules called constraints that are defined in the data model and enforced by the RDBMS. The simplest methods of enforcing integrity are assigning an attribute type with its attendant “nullability.” Candidate keys provide entity integrity. Foreign keys enforce referential integrity.

**What purposes do they serve?**

Data integrity, uniqueness, and restricted values of only the ones needed.

**If you're working in a database or file system environment that does not implement constraints as SQL does, can you "do it yourself?" How?  And would that work as effectively as a "proper SQL" approach would? Why or why not?**

R Studio provides a dataset manipulation library called Dplyr that allows users to normalize, denormalize, and fetch data without primary or foreign key constraints. It works just as effectively as “proper SQL” and is much easier to type and read.

**4. FUNCTIONS.  In algebra, we talk a lot about functions, which are defined ways of expressing a concept (like taking a square root of a number, or a cosine of an angle). Why do we need functions in SQL? Isn't that what we have programming languages for?**

T-SQL provides several functions that operate on character strings, such as CONCAT for concatenation and SUBSTRING that extracts a substring from a string. We need functions in SQL for data manipulation and cleaning so that we receive only the information we want.

**Compare how you use built-in functions in SQL with how you use functions in programming languages such as C#. What's different?**

In C# creating methods (functions) is an essential component for building an application. In SQL using the predefined functions to retrieve data is more important than creating methods that perform an action.

**If you need a function but it's not built in to either SQL or a programming language, what can you do?**

You can use CREATE FUNCTION to create your own function.