his is a class in using a database system to work with data. So we need to have some agreement on a few terms. For this first unit of class, we really do not need a deep understanding of these concepts. We will develop these over the next few units. But I want you to connect to your database system and get started, so some terminology is required. Remember these are the quick-and-dirty explanations.

Data– values that have meaning to us for some purpose. For example, your name, your student ID, your date of birth, the number of units for this class- these are all data. Some data values are numeric, other data values are dates, and other data values are text. To be useful, we need to organize the data and store it in a way that we can retrieve it efficiently. If we have a lot of data, we will generally store it using a computer system and one of the more useful ways to organize and store data is to use a database.

Database – a collection of data that we can use for more than one purpose. For example, CCSF has a student database that we use for registration, production of transcripts, billing students, and sending reports to the state for funding. Databases are everywhere these days.

Table – the data in the database is stored in structures that we can think of as tables.

Student ID	*LastName*	*FirstName*
W0000003	*Babbage*	*Charles*
W0000001	*Hopper*	*Grace*
W0000002	*Teape*	*Betsy*

A table consists of a set of rows, this table has three rows of data. The shaded row at the top contains the table header. A table stores data of a specific type defined by its columns. In the student database, we might have a table for basic student data; each row would contain the data values for a specific student; the columns might have names such as StudentLastName, StudentFirstName, StudentResidencyState, StudentID, and StudentDOB. Each column stores one type of information that we need to keep for a student. We might have another table, ClassesTaken, for the classes a student has taken; in that table the columns might have names such as StudentID, CourseID, Semester, and FinalGrade.

Identifier - If we are storing data about students in a table, we would want to be able to identify a specific student and his data. Most tables will have a column that is used to identify each row of data. With the student table described in the previous paragraph, we would probably use the StudentID as the identifier. What do you think we could use as an identifier for the ClassTaken table? Consider that a student might take a class more than once.

Query – a question that we ask about the data- such as "for which classes is the student with ID W00000009 registered?" or "show me how many students we have registered for classes this semester who do not live in California".

Schema – this is the collection of tables that you have for a particular purpose. In a company we might have a schema for our customer data and another schema for the products that we sell and another schema for our employees. In MySQL database systems, the terms schema and database mean the same thing; in other database systems, the schema is a subcategory of a database.

If we want to store our data in tables in a database, we will need to have computer programs that help us store and retrieve the data. The main programs are the RDBMS, clients, and application programs.

RDBMS –Relational Database Management System – this is a collection of programs that handles the physical files that make up the database. When we create a table, add rows of data to that table, or look at the data in that table, we issue a command that is transferred to one or more of the programs in the RDBMS which does the actual reading or writing of the data. The RDBMS also handles security and other issues.

Client/Server – a client/server system has two types of software- the server software and the client software. The server software is also called the database engine. These two types of software can be installed on the same physical computer or on different computers that are linked via a network.

Database server/engine – this is the set of programs in the RDBMS that work with the data. When you run a query, you are actually passing the query to the database server. In many systems the terms server and engine mean the same thing. MySQL uses these terms in a somewhat different way that we will discuss. In this class, we are using a MySQL database server.

Client – to communicate with the database engine, you need a program to enter your SQL statements and see the results; that is a client program. We are using the mysql client to issue queries to the dbms and receive the results. You can use several different clients to work with the same DBMS.

Application program- this would be the programs that a typical end user would work with. A typical end-user does not interact directly with the SQL language. These application programs would often be written in languages such as Java, Visual Basic or C#. We do not work at this level in this class.

SQL – a language which is commonly used to communicate with a database. Much of this semester is devoted to learning how to use SQL effectively.

Set-based operations and closure – These are basic concepts of the relational system. When we work with a database using the SQL language our commands are applied to the entire table as a unit; we can think of the table as a set of rows and we think of the query as working on all of the rows at the same time. The result of a command is another table; this is the feature of closure.

The first question listed above was "What classes is the student with ID W00000009 registered for?" The result of that query might be a list of classes showing the ID and the CRN of each class and its meeting days. That looks like a table.

Student ID	*LastName*	*Class CRN*	*Class ID*	*Days*
W00000009	*Jones*	*78956*	*CS 151A*	*Online*
W00000009	*Jones*	*78512*	*CS 159A*	*Tues*
W00000009	*Jones*	*70236*	*CS 112A*	*Thurs*

The result of the second query "Show me how many students we have registered for classes this semester who do not live in California" is also a table- even if it has only one row and one column.

OutOfStateStudents
845

With some queries, the result set might be empty. This is still considered a table- it might be called an empty table. Depending on the client, you might not see a column header row such as indicated here.

StudentsUnder12YearsOfAgeEnrolledInCSClasses

Instead you might get a response such as

```
   Empty set
```