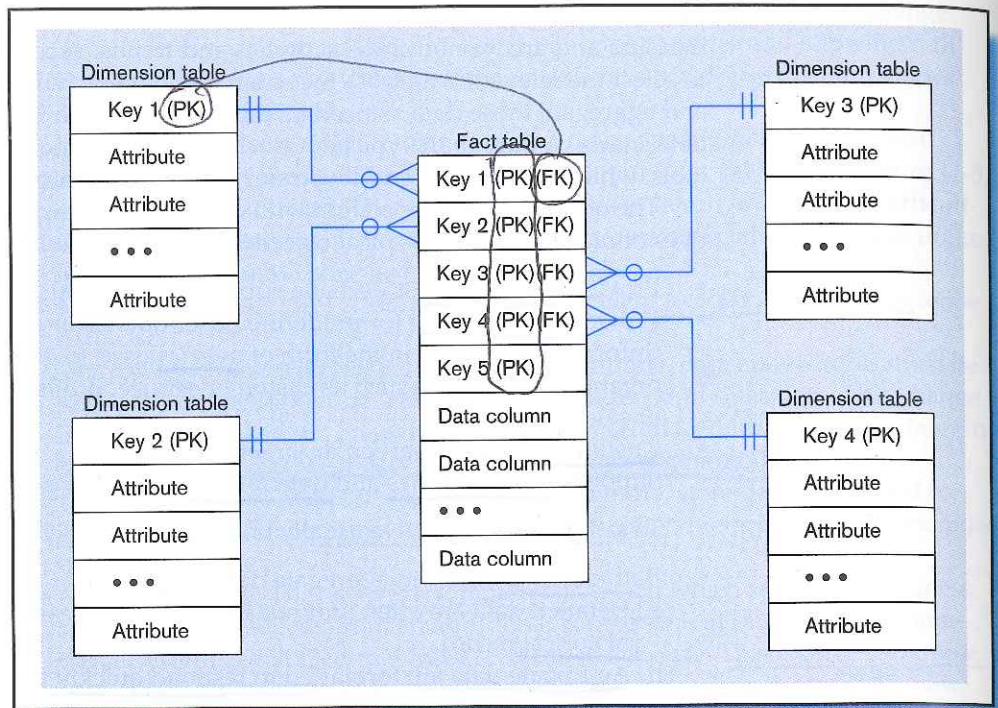Figure 11-9

Components of a star
schema

Each dimension table has a one-to-many relationship to the central fact table. Each dimension table generally has a simple primary key, as well as several nonkey attributes. The primary key, in turn, is a foreign key in the fact table (as shown in Figure 11-9). The primary key of the fact table is a composite key that consists of the concatenation of all of the foreign keys (four keys in Figure 11-9), plus possibly other components that do not correspond to dimensions. The relationship between each dimension table and the fact table provides a join path that allows users to query the database easily, using SQL statements for either predefined or ad hoc queries.

By now you have probably recognized that the star schema is not a new data model, but instead a denormalized implementation of the relational data model. The fact table plays the role of a normalized n-ary associative entity that links the instances of the various dimensions, which are in second, but possibly not third, normal form. To review associative entities see Chapter 3, and for an example of the use of an associative entity, see Figure 3-11. The dimension tables are denormalized. Most experts view this denormalization as acceptable because dimensions are not updated and avoid costly joins; thus the star is optimized around certain facts and business objects to respond to specific information needs. Relationships between dimensions are not allowed; although such a relationship might exist in the organization (e.g., between employees and departments), such relationships are outside the scope of a star schema. As we will see later, there may be other tables related to dimensions, but these tables are never related directly to the fact table.

Example Star Schema

A star schema provides answers to a domain of business questions. For example, consider the following questions:

1. Which cities have the highest sales of large products?
2. What is the average monthly sales for each store manager?
3. In which stores are we losing money on which products? Does this vary by quarter?

A simple example of a star schema that could provide answers to such questions is shown in Figure 11-10. This example has three dimension tables: PRODUCT,