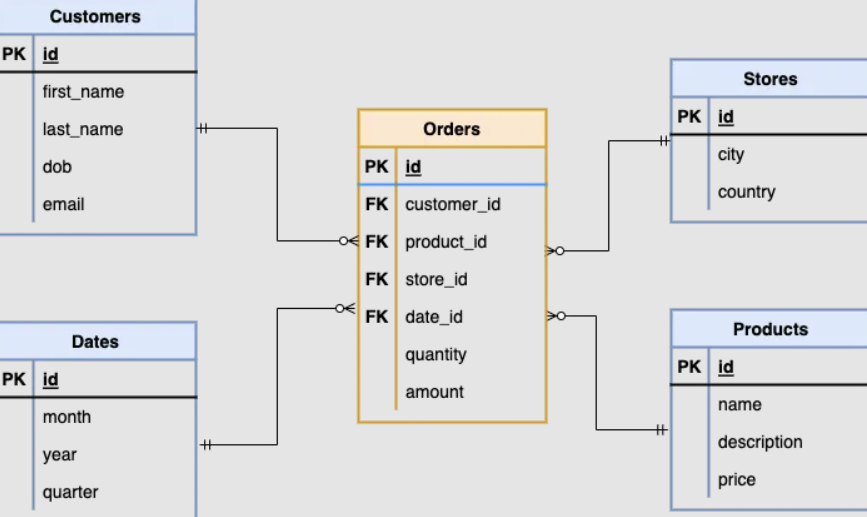
<https://medium.com/@mohamad.ashour203/fact-table-definition-and-its-types-in-data-warehousing-with-examples-4fa89cc53dee>

Afact table stores quantitative information for analysis and is often denormalized. A fact table works with dimension tables, and it holds the data to be analyzed and a dimension table stores data about the ways in which the data can be analyzed.

A fact table or a fact entity is a table or entity in a star or snowflake schema that **stores measures that measure the business**, such as sales, cost of goods, or profit.

The fact table is a central table in the data schemas.

A fact table **consists of two types of columns**. The f**oreign keys column** allows to join with dimension tables and the **measure columns**contain the data that is being analyzed.



We can notice that the orders table is the fact table which contain the foreign key as you see and the measurements columns like quantity and amount to calculate any information across all tables if needed.

A fact table is a database table in a data warehouse or Business Intelligence (BI) application that stores quantitative data and foreign keys for analysis. Fact tables are a central component of a dimensional model's star or snowflake schema, surrounded by dimension tables.

A diagram of a table

Description automatically generated

**What is a Fact Table?**

A fact table stores quantitative information for analysis and is often denormalized. A fact table works with dimension tables, and it holds the data to be analyzed and a dimension table stores data about the ways in which the data can be analyzed.

A fact table or a fact entity is a table or entity in a star or snowflake schema that **stores measures that measure the business**, such as sales, cost of goods, or profit.

The fact table is a central table in the data schemas.

A fact table **consists of two types of columns**. The f**oreign keys column** allows to join with dimension tables and the **measure columns**contain the data that is being analyzed.

A diagram of a product

Description automatically generated with medium confidence

We can notice that the orders table is the fact table which contain the foreign key as you see and the measurements columns like quantity and amount to calculate any information across all tables if needed.

<https://www.simplilearn.com/snowflake-schema-in-data-warehouse-model-article#:~:text=The%20snowflake%20schema%20is%20a%20type%20of%20database%20design%20that,the%20points%20on%20a%20snowflake>.

One of the many variations of the star schema is the **snowflake** schema. The centralized fact table is linked to multiple dimensions in this case. Dimensions are present in a normalized form in multiple related tables in the snowflake schema.

When the dimensions of a star schema are detailed and highly structured, with several levels of relationship, and the child tables have multiple parent tables, the snowflake structure appears. The snowflake effect only affects dimension tables and has no effect on fact tables.

In a snowflake schema, a central fact table is surrounded by several dimensions that are represented by their tables. The dimensions are connected to the central fact table through foreign keys, which allow the fact table to reference the attributes of the dimensions.

Star Schema vs Snowflake Schema

A star schema and a snowflake schema are two different types of database schemas that are used to organize data in a structured manner.

Both types of schema involve a central fact table surrounded by dimension tables, but there are some key differences between the two:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Feature | Star Schema | | Snowflake Schema | |
| Structure | A central fact table surrounded by a few dimension tables | | A central fact table surrounded by multiple dimension tables, which may be further divided into sub-dimension tables | |
| Data Modeling | | Simple and easy to understand | | More complex and detailed |
| Query Performance | | Faster query performance due to fewer join | | Slower query performance due to more joins |
| Data Redundancy | | Low data redundancy due to denormalized structure | | Higher data redundancy due to normalized structure |
| Data Integrity | | Lower data integrity due to denormalized structure | | Higher data integrity due to normalized structure |
| Data Storage | | Takes up more space due to the denormalized structure | | Takes up less space due to the normalized structure |

Overall, the choice between a star schema and a snowflake schema will depend on the specific needs of your database and the type of analysis you are conducting. The star schema is generally simpler and faster, but the snowflake schema provides more detailed and normalized data that may be necessary for certain types of analysis.

To understand snowflake schema in a better way, let us look at some real-world examples.

## **Examples of Snowflake Schema**

The snowflake schema is a type of database design that is used in [data warehousing](https://www.simplilearn.com/data-warehouse-article). It is called a snowflake schema because the diagram of the schema resembles a snowflake, with the central fact table at the center and the dimension tables branching out from it like the points on a snowflake.

Here are a few of the real-world use cases of the snowflake schema:

1. A [database](https://www.simplilearn.com/what-is-database-management-article) for a hospital that stores patient medical records: The central fact table might be the "Patient Visits" table, which contains data on the visits that patients have made to the hospital. Dimension tables might include the "Patients" table, which contains data on the patients themselves, the "Doctors" table, which contains data on the doctors who treated the patients, and the "Procedures" table, which contains data on the medical procedures that were performed.
2. A database for a university that stores student enrollment and grades: The central fact table might be the "Enrollment" table, which contains data on the courses that students are enrolled in. Dimension tables might include the "Students" table, which contains data on the students themselves, the "Courses" table, which contains data on the courses being offered, and the "Teachers" table, which contains data on the teachers who are teaching the courses.
3. A database for a social media platform that stores user activity data: The central fact table might be the "Activity" table, which contains data on the actions that users take on the platform, such as liking a post or commenting on a photo. Dimension tables might include the "Users" table, which contains data on the users themselves, the "Posts" table, which contains data on the posts that are made on the platform, and the "Groups" table, which contains data on the groups that users belong to on the platform.

In order to get a sense of whether the snowflake schema is the right fit for our use case, we need to get an overview of its various advantages and disadvantages.

## **Advantages of Snowflake Schema**

A snowflake schema is a type of database schema that is designed to improve the organization and efficiency of a database by providing a more structured and normalized data model.

Some of the main advantages of the snowflake schema are given below:

* Data integrity: It helps to improve data integrity by reducing redundancy and ensuring that data is stored in a more organized and consistent way. This can make it easier to maintain the database and ensure that the data is accurate and up-to-date.
* Lesser disk space required: It uses small disk space because the data is highly structured and normalized. This can be especially useful for large databases that need to be stored on disk, as it can help to save space and reduce the overall size of the database.
* Eases database organization: Useful tool for organizing and managing complex databases, particularly in data warehousing and business intelligence applications where the focus is on analyzing and reporting on large amounts of data.
* Easier to update and maintain: It is easier to maintain and update compared to other schemas like the star schema. In a star schema, all of the dimension tables are directly connected to the central fact table, which can make it more difficult to add or modify dimensions. In a snowflake schema, the dimensions are organized into separate tables, which can make it easier to modify the schema as the data warehousing requirements change.

We learned about the main advantages of the snowflake schema, and now let us understand its various disadvantages.

## **Disadvantages of Snowflake Schema**

While the snowflake schema can provide a number of benefits in terms of query performance and [data modeling](https://www.simplilearn.com/what-is-data-modeling-article), it also has a number of potential drawbacks. Some of the main disadvantages of using a snowflake schema include the following:

* Performance issues: While a snowflake schema can improve query performance in certain situations, it can also introduce performance issues in others. For example, if the schema is not well-optimized or if the data is not distributed evenly, queries may be slow to execute or may require a large number of resources.
* Reduced flexibility: The hierarchical structure of a snowflake schema can make it more difficult to add or modify data, as changes to one dimension table may require corresponding changes to other tables in the hierarchy. This can be particularly problematic if the schema is not well-designed or if the data requirements change over time.
* Data redundancy: A snowflake schema may result in data redundancy, as the same data may be stored in multiple tables and in multiple locations within each table. This can lead to issues with data consistency and integrity, as well as increased storage requirements.
* Increased complexity: Because a snowflake schema involves multiple dimension tables and foreign key relationships, it can be more complex to design and maintain than a simpler, flatter schema. This complexity can make it more difficult for users to understand and work with the data, and it can also increase the risk of errors and inconsistencies.

Enroll in the[Professional Certificate Program in Data Analytics](https://www.simplilearn.com/pgp-data-analytics-certification-training-course?source=GhPreviewCTAText) to learn over a dozen of data analytics tools and skills, and gain access to masterclasses by Purdue faculty and IBM experts, exclusive hackathons, Ask Me Anything sessions by IBM.

## **Advance Your Career in Data Analytics With Simplilearn’s PCP Data Analytics Certification Training Course**

In summary, a snowflake schema is a database design that organizes data into a hierarchical structure with multiple levels of dimension tables. It can offer improved query performance and efficient data modeling, but it can also be complex and inflexible. It is important to carefully consider the specific needs and goals of an organization before deciding if a snowflake schema is the right choice.

If you're interested in learning more about database design and [data analytics](https://www.simplilearn.com/tutorials/data-analytics-tutorial/what-is-data-analytics), then consider enrolling in [Simplilearn's Professional Certificate Program in Data Analytics Course](https://www.simplilearn.com/pgp-data-analytics-certification-training-course?source=GhPreviewCoursepages). This comprehensive course will teach you the skills and knowledge you need to design and implement effective data analytics solutions using a variety of tools and techniques.

With a focus on real-world applications and hands-on practice, this course is designed to give you the skills and confidence you need to succeed in the field of data analytics. Don't miss this opportunity to take your career to the next level – enroll in [Simplilearn's Professional Certificate Program in Data Analytics course](https://www.simplilearn.com/pgp-data-analytics-certification-training-course?source=GhPreviewCoursepages) today!

## **FAQs**

### **1. What is a snowflake schema?**

A snowflake schema is an expansion of a star schema that includes subdivided dimension tables that follows a multi-dimensional data model.

### **2. What is the difference between star and snowflake schema?**

A star schema contains dimension tables as well as fact tables. A snowflake schema includes all three types of tables: dimension tables, fact tables, and sub-dimension tables.

### **3. Why is it called a snowflake schema?**

It is known as a snowflake schema because the diagram resembles a snowflake. Snowflaking is a technique used to normalize the dimension tables in star schemas.

### **4. Is the snowflake and snowflake schema the same?**

No, "snowflake" and "snowflake schema" refer to two different things.

Snowflake is a specific type of data warehouse software that is designed to work with large amounts of [structured](https://www.simplilearn.com/structured-vs-unstructured-data-article) and [semi-structured data](https://www.simplilearn.com/semi-structured-data-article).

A snowflake schema, on the other hand, is a specific type of database schema that is used to organize data in a star schema.

### **5. What are the four types of snowflake tables?**

In a snowflake schema, there are typically four types of tables:

* Fact tables
* Dimension tables
* Hierarchy tables
* Bridge tables

### **6. Can we join 2 fact tables?**

Joining the two fact tables is much easier now that we have this bridge table with the row identifiers for both fact tables.

### **7. What are the 3 types of schema?**

The 3 types of schema are:

* Logical Schema
* Physical Schema
* View Schema