DATABASE VS DATA WAREHOUSE

Databases and Data Warehouses are used to generate different type of information.

Information generated by both are used for different purposes.

These are generated by Database Administrator or Database Engineer

NOTE:

`Row of a Table is also called Record`

`Column of a Table is aslo called Attribute`

What is Database?

A database is a collection of data that is organized so that it can be easily accessed, managed, and updated. Databases are used to store all sorts of information, such as customer records, product inventory, financial data, and medical records.

Some of the most common types of databases include:

- **Relational databases:** Relational databases are the most common type of database. They are organized into tables, which are made up of rows and columns. Each row represents a single record, and each column represents a single piece of data about that record.

- **Non-relational databases:** Non-relational databases are not organized into tables. Instead, they store data in a more flexible way, which can be useful for storing large amounts of unstructured data.
- **Object-oriented databases:** Object-oriented databases store data in objects, which are similar to classes in programming languages. This makes them well-suited for storing data that is related to objects, such as product catalogs and customer profiles.

Here are some of the benefits of using databases:

- **Efficiency:** Databases can help organizations to be more efficient by storing data in a centralized location. This makes it easier to access and use data, and it can also help to reduce the amount of data duplication.
- **Accuracy:** Databases can help to improve the accuracy of data by ensuring that it is stored in a consistent format. This can help to prevent errors and omissions, and it can also make it easier to find and use data.
- **Security:** Databases can help to protect data by providing security features such as authentication, authorization, and encryption. This can help to prevent unauthorized access to data, and it can also help to protect data from unauthorized modification or destruction.

Transactional Database:

A **transactional database** is a database management system (DBMS) that supports ACID (atomicity, consistency, isolation, and durability) transactions. A transaction is a set of database read and write operations where either all or none of the operations succeed.

Transactional databases are used in applications where data integrity is critical, such as banking, e-commerce, and healthcare. They provide a number of features that help to ensure data integrity, including:

- **Atomicity:** A transaction is either committed or rolled back as a whole. This means that if any part of a transaction fails, the entire transaction is rolled back and the database is restored to its previous state.
- **Consistency:** Once a transaction is committed, the database is guaranteed to be in a consistent state. This means that the data in the database will always be accurate and reflect the results of all committed transactions.

- **Isolation:** Transactions are isolated from each other, which means that they cannot see changes made by other transactions until those changes are committed. This prevents one transaction from seeing the intermediate state of another transaction.
- **Durability:** Once a transaction is committed, the changes made by the transaction are guaranteed to be permanent. This means that the data in the database will not be lost even if there is a system failure.

Here are some of the benefits of using transactional databases:

- **Data integrity:** Transactional databases help to ensure data integrity by providing ACID transaction support. This means that data is always consistent and accurate, even in the event of a system failure.
- **Performance:** Transactional databases are designed to provide high performance for read and write operations. This makes them ideal for applications that require high throughput, such as e-commerce and online banking.
- **Scalability:** Transactional databases can be scaled to meet the needs of growing applications. This can be done by adding more servers or by using a distributed database architecture.
- **Security:** Transactional databases provide a number of security features, such as authentication, authorization, and encryption. This helps to protect data from unauthorized access, modification, or destruction.

```
If you are looking for a database that can provide high performance, scalability, and security, then a transactional database is a good solution.
```

Operational Database:

An operational database, also known as an online transaction processing (OLTP) database, is a database that is designed to support high-volume, low-latency transactions. Operational databases are used to store data that is used in day-to-day business operations, such as customer orders, product inventory, and financial transactions.

Operational databases are different from data warehouses in a few key ways. Data warehouses are designed for analytical queries, which typically involve large amounts of data and complex calculations. Operational databases, on the other hand, are designed for transactional queries, which typically involve small amounts of data and simple calculations.

Low Latency Transaction: It refers to the time it takes to complete a transaction, such as buying a product online or transferring money from one bank account to another. Low latency transactions are essential for applications that require real-time updates, such as stock trading and online gaming. Here an example:

- **A gamer is playing an online game.**
- The gamer makes a move, and the game server receives the move in milliseconds.
- The game server updates the game state, and the other players in the game see the gamer's move in real time.

Here are some of the benefits of using operational databases:

- **Performance:** Operational databases are designed to provide high performance for transactional queries. This makes them ideal for applications that require high throughput, such as e-commerce and online banking.
- **Scalability:** Operational databases can be scaled to meet the needs of growing applications. This can be done by adding more servers or by using a distributed database architecture.
- **Security:** Operational databases provide a number of security features, such as authentication, authorization, and encryption. This helps to protect data from unauthorized access, modification, or destruction.

Different kinds of databases:

- **Relational Database:**

  A relational database is made up of a set of tables with data that fits into a predefined category.

- **Distributed Database:**

  A distributed database is a database in which portions of the database are stored in multiple physical locations, and in which processing is dispersed or replicated among different points in a network.

What is Data Warehouse?

- A data warehouse is a place where you store all of your data.
- It is a big database that is designed for analysis and reporting.
- It can help you to make better decisions, improve efficiency, and identify new opportunities.
- In data warehouse, a large amount of heterogeneous data is collected and transformed according to decision making system for generating analytical reports.
- `The Data warehouse is not only used in present decision making but also contributes toward future decision making.`
- Warehouses can be started and stopped at any time. They can also be resized at any time, even while running, to accommodate the need for more or less compute resources, based on the type of operations being performed by the warehouse.
- A warehouse **provides the required resources, such as CPU, memory, and temporary storage**, to perform the following operations in a Snowflake.

A data warehouse is different from a traditional database in several ways.

1. A data warehouse is designed for analysis and reporting, not for transaction processing. This means that it is optimized for querying large amounts of data, not for inserting, updating, and deleting data.
2. A data warehouse typically contains historical data, not just current data. This allows users to analyze trends over time and to identify patterns that would not be visible if they only had access to current data.

Here is an example of how a data warehouse can be used. Let's say you are the manager of a retail store. You want to know which products are selling well and which products are not. You could use a data warehouse to analyze sales data from the past year. This would allow you to identify the products that are most popular with your customers and to make changes to your inventory accordingly.

Here are some of the benefits of using a data warehouse:

- Improved decision-making: A data warehouse can help you to make better decisions by providing you with access to historical data and trends.
- Increased efficiency: A data warehouse can help you to improve efficiency by automating tasks such as data integration and reporting.

- Better customer service: A data warehouse can help you to provide better customer service by giving you insights into customer behavior and preferences.
- Increased revenue: A data warehouse can help you to increase revenue by identifying new opportunities for growth and by optimizing your marketing campaigns.

What is d/f between Database and Data Warehouse? And why do we need it?

~	DATABASE	DATA WAREHOUSE
DEFINITION	A database is a collection of data organized for easy access and retrieval.	A data warehouse is a repository of historical data that is used for analysis and reporting.
DESIGN	Design of Operational database mainly observes data accuracy when updating real-time data	The design of data warehouse ensures vast range of data which is used overtime for analysis purpose.
FOCUS	Databases are focused on storing and managing data for a specific application or department.	Data warehouses are focused on storing and managing data for analysis and reporting.
TYPE OF INFORMATION	Databases typically store structured data, such as customer records, product inventory, and financial transactions.	Data warehouses typically store both structured and unstructured data, such as customer emails, social media data, and sensor data.
TYPES	There are many types of databases. The examples are OLTP, CSV, excel spreadsheets and XML files etc.	It is an OLAP type of database which exist on the top layer of other database and perform analysis.

~	DATABASE	DATA WAREHOUSE
OPTIMIZATION	Databases are optimized for high performance, with users expecting to be able to run complex queries without any performance degradation.	Data warehouses are optimized for good performance, with users expecting to be able to run complex queries within a reasonable amount of time.
REPORTING	Databases can be used to generate reports, but they are not specifically designed for this purpose. Example, you may need data of a specific patient to check the history of disease. The result can be in PDF formats. The data may be combined from several tables and actual queries may be complex which requires expertise in this field.	Data warehouses are specifically designed for generating reports, and they can be used to create a wide variety of reports, including dashboards, scorecards, and ad hoc reports.
DATA DUPLICATION	Databases can contain duplicate data, but this is typically not a problem.	Data warehouses are designed to avoid data duplication, as this can lead to performance problems and data integrity issues.

Difference between OLTP and OLAP? When is each used?

OLTP	OLAP
OLTP(online transaction processing) is defined by several online transaction to obtain the information.	OLAP(online analytical processing). In this system, very few transactions are performed.
The query processing of this system is very fast.	It uses large amount of current as well as historical data.
Its effectiveness depends upon the transactions carried out per second.	
It has original data source.	The data comes from several sources
It is highly normalized system which makes it efficient.	It is Denormalized and reports are generated from vast range of data.

OLTP	OLAP
It is used to run and control the basic task of a business.	It is used for improving the performace of business and in making decisions and planning.

Similarities

The similarity between data warehouse and database is that both the systems maintain data in form of table, indexes, columns, views and keys. Also data is retrieved in both by using SQL queries.