**Name:** Ahsan Bilal

## Can a database be used as Data Warehouse?

Yes, a database can be used as a data warehouse. In fact, many data warehouses are built using databases. However, a data warehouse typically requires some additional features and capabilities beyond what is offered by a traditional database management system (DBMS).

A data warehouse is a large, centralized repository that stores data from multiple sources and is optimized for reporting and analysis. It typically requires support for complex queries, data transformation and integration, and large-scale data storage and retrieval. Additionally, a data warehouse may also include features such as data partitioning, indexing, and compression to improve performance.

Some modern DBMSs include features that are specifically designed for data warehousing, such as columnar storage, compression, and partitioning. For example, Amazon Redshift, Google BigQuery, and Microsoft Azure SQL Data Warehouse are all cloud-based data warehouse solutions that use a relational database engine as their underlying technology.

So, while a traditional database can be used as a data warehouse, it may not be optimized for the specific requirements of a data warehousing environment. It is important to carefully evaluate the features and capabilities of any database management system being considered for use as a data warehouse to ensure that it meets the specific needs of the organization.

## Major differences between structured and Un-structured data.

*Definition:* Structured data refers to data that has a defined format or schema, making it easily searchable and sortable. Unstructured data, on the other hand, refers to data that does not have a predefined structure, making it more difficult to search and analyze.

*Examples:* Examples of structured data include data stored in a relational database, such as customer information, transaction records, or inventory levels. Examples of unstructured data include text documents, images, videos, social media posts, and email messages.

*Organization:* Structured data is typically organized in rows and columns or in a specific data model, such as a hierarchical or network model. Unstructured data is typically stored as files or objects without a predefined structure.

**Date:** 23/3/2023

**Name:** Ahsan Bilal

*Analysis:* Structured data is well-suited for quantitative analysis, such as statistical analysis or data mining, because it can be easily queried and aggregated. Unstructured data is often analyzed using natural language processing (NLP) or machine learning techniques, which can extract insights from unstructured text or image data.

*Volume:* Structured data tends to be smaller in volume and more easily manageable, while unstructured data can be much larger in volume and more difficult to store, manage, and process.

## What are the duties of a data engineer?

The duties of a data engineer can vary depending on the specific organization and industry, but some high-level responsibilities may include:

- Data architecture design: Data engineers design and implement data architectures, including data storage solutions, data pipelines, and data integration processes.
- Data modeling and schema design: Data engineers design and maintain data models, schemas, and data dictionaries to ensure consistency, accuracy, and data quality.
- ETL development: Data engineers build and maintain Extract, Transform, and Load (ETL) processes to move data between systems, transform data to fit business requirements, and load data into data storage solutions.
- Data quality assurance: Data engineers develop and implement data quality assurance processes, such as data profiling and data validation, to ensure data accuracy and completeness.
- Database administration: Data engineers manage and maintain databases, including monitoring database performance, optimizing database configurations, and troubleshooting issues.

**Date:** 23/3/2023