**What is data mart**?  
A data mart is **a subset of a data warehouse focused** on a particular line of business, department, or subject area. Data marts make specific data available to a defined group of users, which allows those users to quickly access critical insights without wasting time searching through an entire data warehouse

**Data warehouse tools**.

* Oracle
* Mark Logic
* Amazon Redshift

**Advantage of Data warehouse in cloud**

Companies are increasingly moving towards cloud-based data warehouses instead of traditional on-premises systems. Cloud-based data warehouses differ from traditional warehouses in the following ways

There is no need to purchase physical hardware.

It’s quicker and cheaper to set up and scale cloud data warehouses.

Cloud-based data warehouse architectures can typically perform complex analytical queries much faster because they use massively parallel processing (MPP).

**Diagram representation of Data warehouse.**

Diagram

Description automatically generated

**What is Databricks used for ?**  
Databricks provides a **unified, open platform for all your data**. It empowers data scientists, data engineers, and data analysts with a simple collaborative environment to run interactive, and scheduled data analysis workloads.  
Azure Databricks, is a **fully managed service** which provides powerful ETL, analytics, and machine learning capabilities. Unlike other vendors, it is a first party service on Azure which integrates seamlessly with other Azure services such as event hubs and Cosmos DB.

In short It can be defined as

Databricks SQL provides an easy-to-use platform for analysts who want to run SQL queries on their data lake, **create multiple visualization types to explore query results from different perspectives**, and build and share dashboards.

**Data Lake :**  
A data lake is **a centralized repository that allows you to store all your structured and unstructured data at any scale**.

**Complete Architecture of Data lake and Data warehouse**

Diagram

Description automatically generated

**To understand better about Data-Lake, Data warehouse and Data mart**

Here is the link given : <https://www.holistics.io/blog/data-lake-vs-data-warehouse-vs-data-mart/>

Below are the topics to get a reference from YouTube and website for better understanding of the topics.

Topics that got covered

1. Overview of Data Lake , Data warehouse and Data bricks
2. How Data warehouse works and additional information about AWS redshift
3. Databricks and how python script and SQL are used in data bricks
4. Insertion and deletion of data into data bricks.
5. User guide for Data mart in AWS

Below are the links given for the reference.

* <https://youtu.be/n-yt_3HvkOI> ( Introduction to data bricks )
* <https://youtu.be/TScSdIJ-_Oo> ( Azure Data bricks tutorial )
* <https://youtu.be/xZ2D34T5dUo> ( Working with data bricks using python script )
* <https://youtu.be/7bfOllAyxlg> ( Amazon redshift tutorial by simpli learn )
* Aws.amazon.com ( reference for AWS and user guide )

A picture containing text, outdoor

Description automatically generated

**Compression Encoding for the columns:**

**There are 7 Types:**A picture containing chart

Description automatically generated

**A Complete Basic Tutorial for reference purpose is given below:**

It is having a playlist of 10 videos from scratch regarding the Data warehousing and a deep dive to AWS redshift.

<https://www.youtube.com/watch?v=xSQNIk4btCI&t=202s&ab_channel=DataandAnalytics>

NetApp from scratch :

<https://www.youtube.com/watch?v=PKmM3K89ccc&ab_channel=MindMajix>

A picture containing graphical user interface

Description automatically generated