DXC AZURE ANALYTICS
ASSIGNMENT-4

Name : Nandi Vomkara Aditya Mohan   DXC_ID : DXCAB1203

Date of submission: 03-06-2022

1) Explain various difference between sql & Nosql DBs?

Ans)

SQL	NO SQL
Tables with fixed rows and columns	JSON documents, Key-value: key-value pairs, Wide-column: tables with rows and dynamic columns, Graph: nodes and edges
General purpose	general purpose, Key-value: large amounts of data with simple lookup queries, Wide-column: large amounts of data with predictable query patterns, analysing and traversing relationships between connected data
Rigid schemas	Flexible Schemas
Vertical (scale-up with a larger server)	Horizontal (scale-out across commodity servers)
Requires ORM (object-relational mapping)	Many do not require ORMs.
Popular relational (SQL) databases include: IBM DB2, Oracle Database, Microsoft SQL Server, MySQL	Popular non-relational (NoSQL) databases include Apache ,Cassandra. Apache HBase, mongo DB/
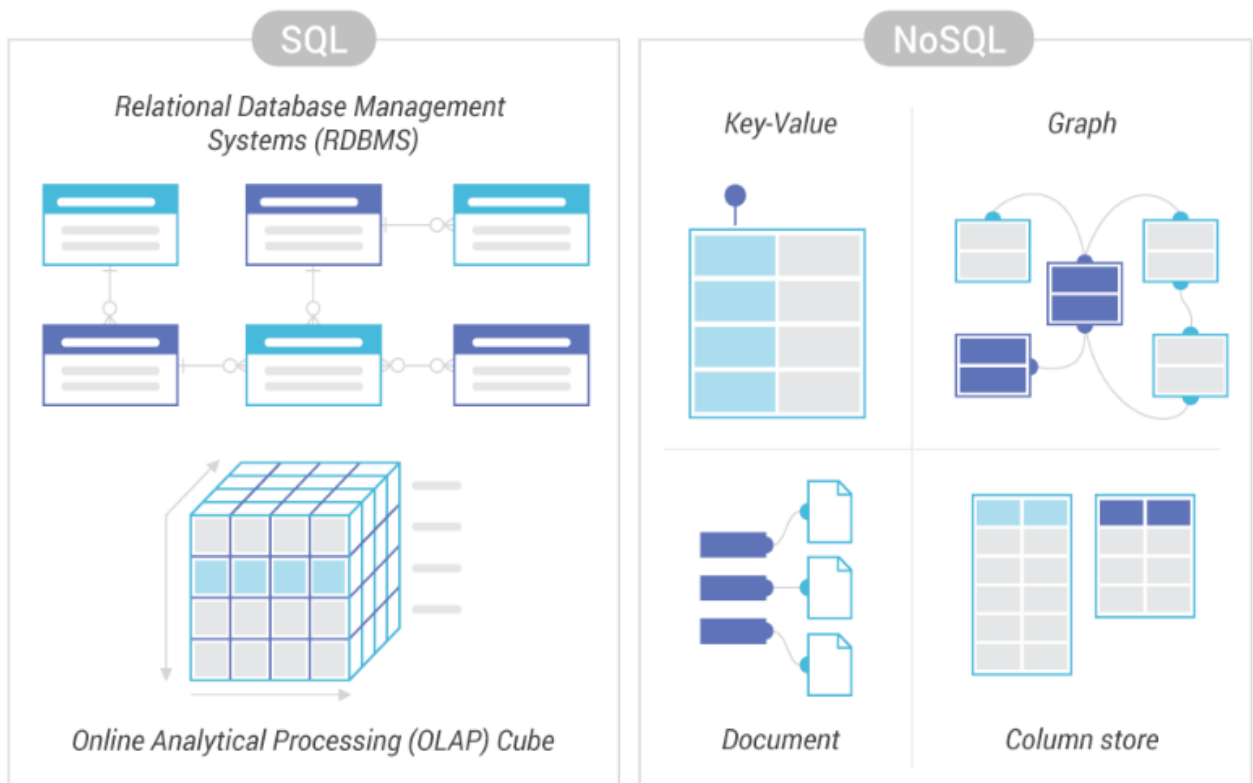
2.explain the advantages of NoSql DBs? Explain how MongoDB data will be inserted?

Ans)

NoSQL databases offer many benefits over relational databases. NoSQL databases have flexible data models, scale horizontally, have incredibly fast queries, and are easy for developers to work with.

- Flexible data models NoSQL databases typically have very flexible schemas. A flexible schema allows you to easily make changes to your database as requirements change. You can iterate quickly and continuously integrate new application features to provide value to your users faster.

- Horizontal scaling : Most SQL databases require you to scale-up vertically (migrate to a larger, more expensive server) when you exceed the capacity requirements of your current server. Conversely, most NoSQL databases allow you to scale-out horizontally, meaning you can add cheaper, commodity servers whenever you need to.
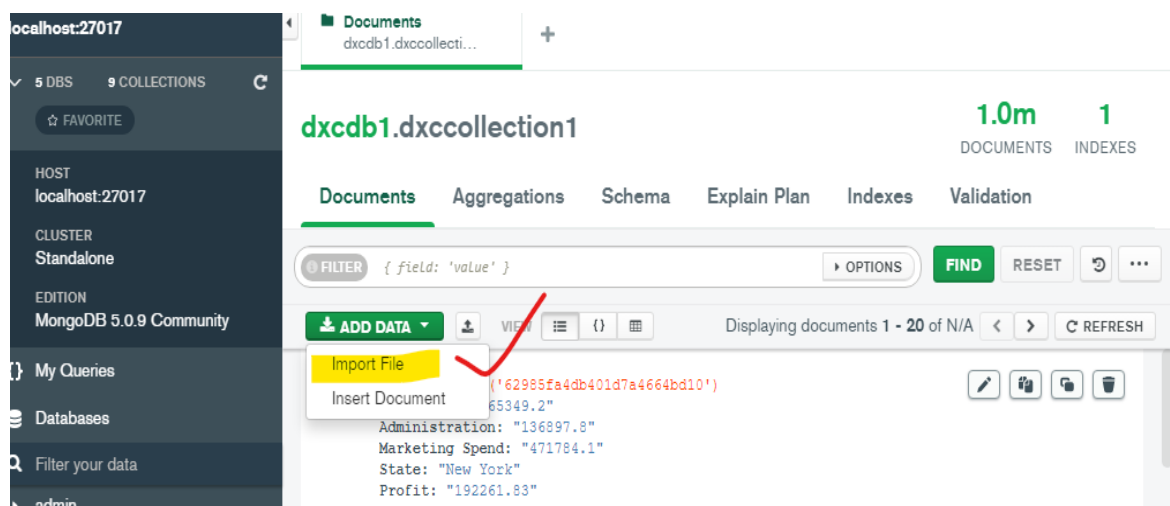
- Fast queries

  Queries in NoSQL databases can be faster than SQL databases. Why? Data in SQL databases is typically normalized, so queries for a single object or entity require you to join data from multiple tables. As your tables grow in size, the joins can become expensive. However, data in NoSQL databases is typically stored in a way that is optimized for queries. The rule of thumb when you use MongoDB is Data that is accessed together should be stored together. Queries typically do not require joins, so the queries are very fast.
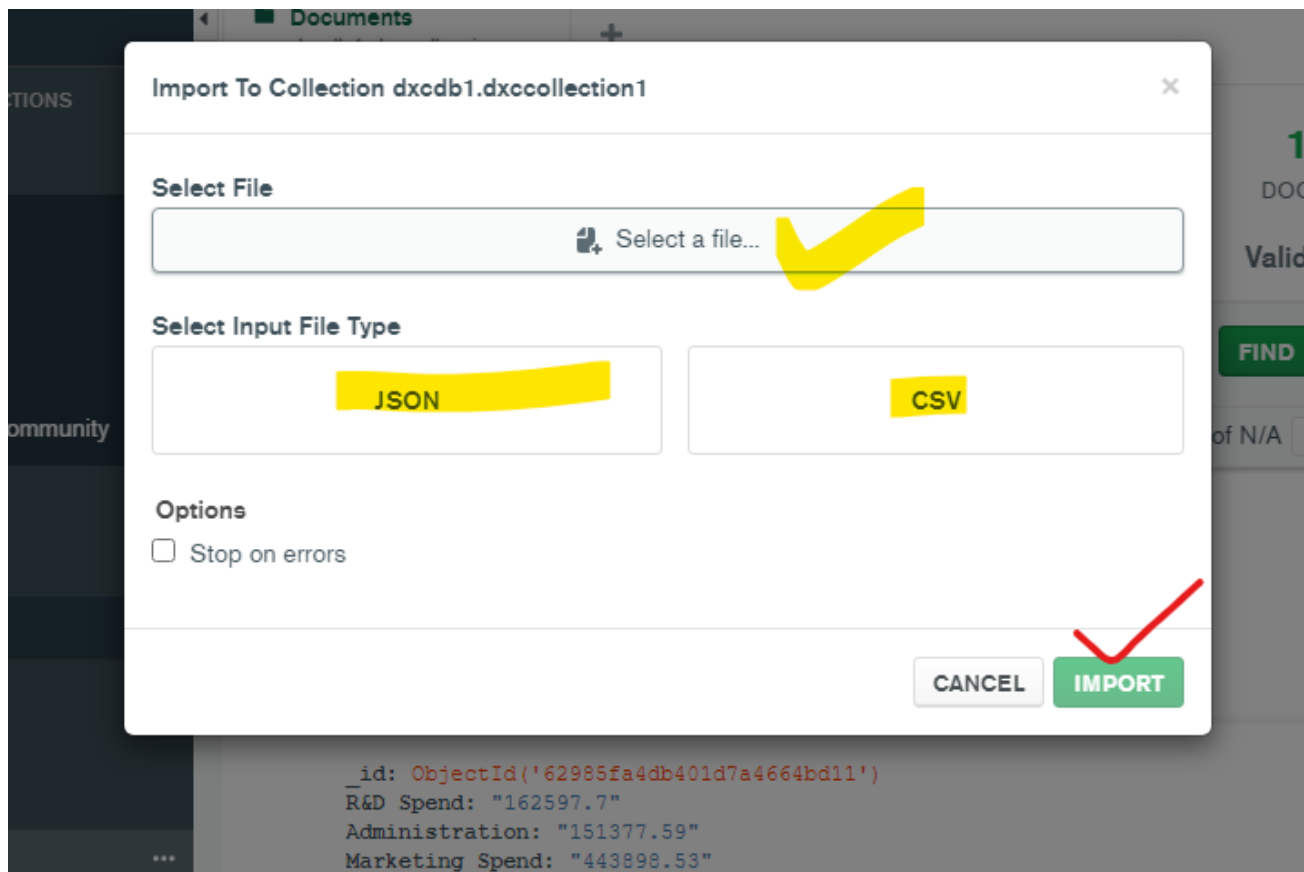
- Easy for developers

  Some NoSQL databases like MongoDB map their data structures to those of popular programming languages. This mapping allows developers to store their data in the same way that they use it in their application code. While it may seem like a trivial advantage, this mapping can allow developers to write less code, leading to faster development time and fewer bugs.

  In MongoDB the data will be inserted in the following ways:

  1. We can directly import the data through a csv or json file forma through compass

2. We can enter manually through command prompt

3. We can add document in the compass

3.Explain the steps – how COSMOS DB is created?

Ans:

```
open portel.azure.com
        ↓
sign in and search for
COSMOS DB in the search
        bar
        ↓
select the CORE
sql(recommended one)
        ↓
choose the name & required
parameters and and wait for
the deployment
        ↓
once deployment is done
the DB is ready to work
```
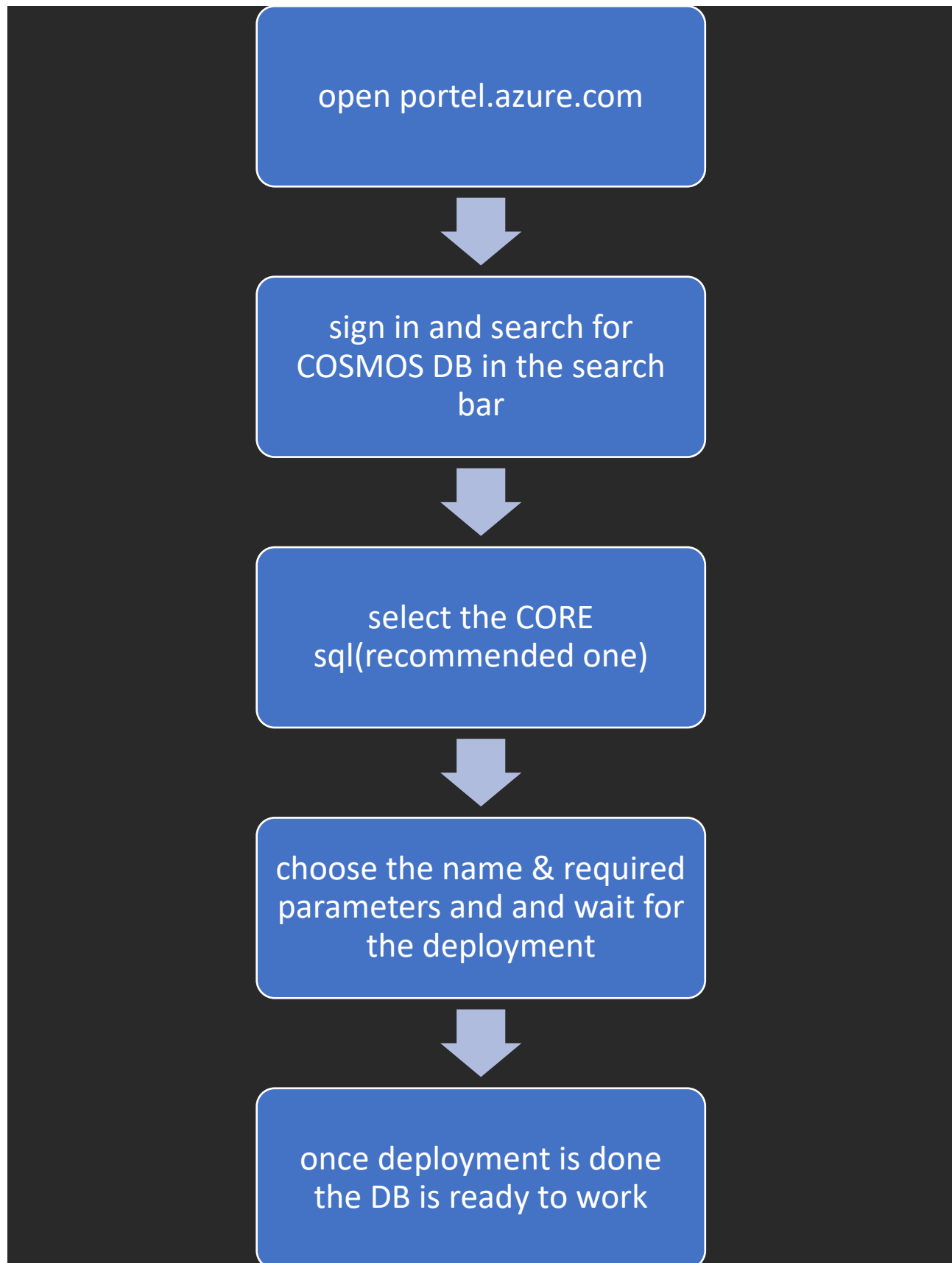
4.Explain how to write JSON query in COSMOS DB?

Ans)

In COSMOS DB the data is entered in the form of KEY and value pairs

So, each and every statement that we are adding into the database is in the from of key and value pairs only

Example
{
"id" : "12s3",
"Name" : "Aditya",
"Course" : "AZURE",
"Company" : "DXC",
"Branch" : "MECH"
}

- Once the data is inserted the data will be automatically updated and there will be system generated key
- This system generated key is used to keep the versions of data that we are inserted time to time
- The data that we enter in the to do list is always stoed in the form key-value pairs

{
"ID" : '23214uimgjfeiwriu6u5' #which is automatically generated.
"id" : "12s3",
"Name" : "Aditya",
"Course" : "AZURE",
"Company" : "DXC",
"Branch" : "MECH"
}

5. Explain the major difference between data bases and data ware houses?

Ans)

Parameter	Database	Data Warehouse
Purpose	Is designed to record	Is designed to analyze
Processing Method	The database uses the Online Transactional Processing (OLTP)	Data warehouse uses Online Analytical Processing (OLAP).
Usage	The database helps to perform fundamental operations for your business	Data warehouse allows you to analyze your business.
Tables and Joins	Tables and joins of a database are complex as they are normalized.	Table and joins are simple in a data warehouse because they are denormalized.
Orientation	Is an application-oriented collection of data	It is a subject-oriented collection of data
Storage limit	Generally limited to a single application	Stores data from any number of applications
Availability	Data is available real-time	Data is refreshed from source systems as and when needed
Usage	ER modeling techniques are used for designing.	Data modeling techniques are used for designing.
Technique	Capture data	Analyze data
Data Type	Data stored in the Database is up to date.	Current and Historical Data is stored in Data Warehouse. May not be up to date.
Storage of data	Flat Relational Approach method is used for data storage.	Data Ware House uses dimensional and normalized approach for the data structure. Example: Star and snowflake schema.
Query Type	Simple transaction queries are used.	Complex queries are used for analysis purpose.
Data Summary	Detailed Data is stored in a database.	It stores highly summarized data.

6. Explain architecture of data ware houses?
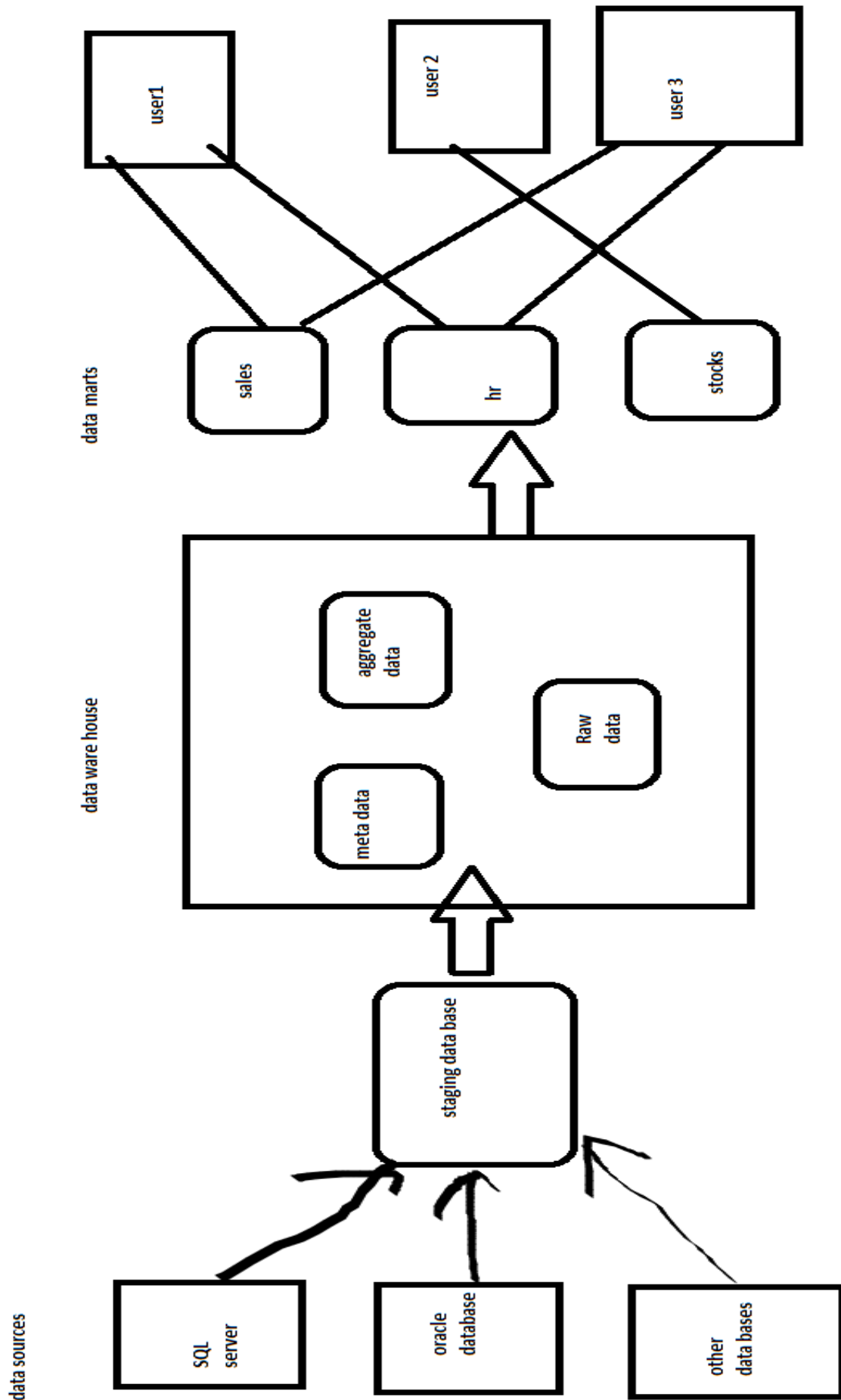
7.Explain what are Datamarts & how different from DATA Bases? And mention the types of data marts too

Ans)

Data marts : data marts are the smaller versions of the databases which deals with a single subject

These are different from the data bases because they don't have the complete access which data bases have. And the data marts are constrained to the focused area.

Types of data marts:

1. Dependant mart
2. Independent mart
3. hybrid data mart

8. Explain OLAP & OLTP with exmaples

Ans)

OLAP : online analytics processing all data ware houses use this information systems

Example : The historical data of a citizen who travelled many countries are recorded in a data ware house to analyse the money he spent for travel and by using the data airlines or any travel agencies provide the discounts or any business strategies can be planed

OLTP : online Transaction processing which is used by the relation data base information system

Example: The payments that we are doing is stored in a database for records

9. explain what is BI & how BI helps Business to take intelligent decisions?

Ans)

BI means Business Intelligence, this works completely on the data which is in the data ware houses. which is used to improve the insights of the business and improve their productivity and value of the product by gathering the data from the customer, integrate the data , visualize the data and analysis is done to take the decision for according to the demand.

10. explain how ETL works with data ware house?

Ans) ETL stands for extract, transform and load. It is a process in which an ETL tool extract data from various data from various source systems, transforms into the staging area ,and then finally loads into the data warehouse.