Custom Bootcamp

29-08-2023 to 01-09-2023

Sushant Patil

• database:

it helps to store the data in easy manner and process the data using sql.

Types of data

- 1.Structured data --> relational database
- 2.Semi structured data --> json data, objects
- 3.Unstructured data --> video data and audio data

Dataware house

It is a central hub where database related to the organization will be stored.

Dataware house usually contains the history data which will be used to analyze the data.

Normalization

Normalization means breaking down the tables to the simple level possible

1NF there should not be multi valued column in the table.

2NF there should not be partial dependency in the table

3NF there should not be a transitive dependency in the table.

Slowly changing dimension

Scd1 will not store the previous data.

Scd2 will store history data about any changes made to the database it will add new row for each change made in the table.

Scd3 will helps to store the previous data but it adds new column to table whenever the changes will be made to the table.

- 4V's of bigdata
- Volume
- Velocity
- Veracity
- Variety
- Types of data processing
- Batch processing it will be performed on chunk of data after specific time interval
- Stream processing means it is used to process the real time data
- Huge data will be performed in distributed manner or parallel processing
- Data ware house will be used to store the structured data
- It will have the fixed schema
- Data ware house has acid properties of database

- Datalake concept
- Datalake can be used to store all kinds of data
- Structured data
- Unstructured data
- Semi structured data
- It has flexible schema
- It won't has acid properties of database
- It is used to diverse amount of data
- Lake house is combination of both dataware house and data lake

- Cloud computing
- Accessing the resource through the internet
- Types of cloud computing
- Private cloud means organization will have its server and storage options within the organization
- Public cloud means server and storage will maintained by the third party vendors
- Hybrid cloud means it will have both private and public cloud
- Benefits of cloud
- Scalability
- Availability
- Elasticity

- Infrastructure as a service cloud service provider will give the server, storage and network required but user can decide about operating system and application installed on the server
- Platform as a service in this type of model the cloud vendor will provide server, storage and operating system for the platform required for the developers.
- Software as a service in this type the cloud provider will give the software and storage and all other things will be the control of service provider.

- Azure is cloud platform which provides a lot of services.
- Azure sql database will be used to store database on cloud and security and scalability will be taken by the cloud provider.
- Subscription will be used to billing purpose
- Resource group is logical entity that is used to store the resources.
- Resource could be anything for example server, sql database etc



















