**3rd JUNE Assignment**

**1)**

|  |  |
| --- | --- |
| **SQL** | **NoSQL** |
| It stores in tabular information | It provides mechanism for storage and retrieval of data other than tabular relation |
| Used to store normal less data | Used to store big data and real time applications |
| It is slow | It is faster |
| Low scalability | High scalability |

**2)Advantages of NoSQL:**

It supports query language

It is faster performance

It provides horizontal scalability

It supports real time data storage

**Data inserting in MongoDB:**

In mongodb we can insert data in 2 ways

**By cmd prompt:**

db.table\_name.insert(fields) here it stores like key value pairs

**By using Mongodb software:**

We can directly go to collections and then we insert data in the format of json or csv format

**3)** At first login to microsoft . >After successful login then search azure cosomosdb >create collections>write the data what we have to insert.

**4)** json query takes the values like key : value pairs like dictionary in python but we write all information inside the square braces along with curly braces

**5)**

|  |  |
| --- | --- |
| **Data base** | **Data Ware House** |
| A database is any collection of data organized for storage, accessibility, and retrieval. | A data warehouse is a type of database the integrates copies of transaction data from disparate source systems and provisions them for analytical use. |
| Database is designed to record data | Data warehouse is designed to analyze data |
| Database is application-oriented-collection of data | whereas Data Warehouse is the subject-oriented collection of data. |
| It maintains the current information | It maintains the history |
| Easy to maintain | Difficulty to maintain |
| Efficiency is high | Low efficiency |

**6)** Data source

Staggin data

users

Data mart

**7)** Data is smaller version of the DWH which deals with single subject and it has less information. Data marts are focused on one area. Hence they draw data from a limited number of sources.

There are 3 types

Dependent data mart

Independent data mart

Hybrid data mart

**8)**

|  |  |
| --- | --- |
| **OLTP** | **OLAP** |
| It has less information | If has high information |
| It is very fast | It is too slow |
| It stores less amount of data 100MB to 1GB | It maintains large data 1GB to 1TB |
| It maintains the current information | It maintains the history |
| Easy to maintain | Difficulty to maintain |
| Efficiency is high | Low efficiency |

**9)** It is the activity of growth of the company. Act of transferring operational data into useful information. Here the data is transferred. Since this data is credible.

Planning Data gathering Data analysis Business action

**10)**

**ETL** is the process extracting the data from various sources transforming data to meet requirement and then loading it into a target DWH