ASSIGNMENT------SOLUTION SUBMISSION
ON
AZURE ANALYTICS
BY

NAME : SAI KIRAN ANCHE
ROLL NO: DXC262AB12021
BATCH:DXC-262-ANALYTICS-B12-AZURE
COMPANY – DXC TECHNOLOGY
TRAINING UNDER : MANIPAL PRO LEARN
TRAINER NAME – MR. AJAY KUMAR
DATE OF SUBMISSION : 3-06-2022
NO OF QUESTIONS :10
EMPLOYEE DOMAIN - AZURE ANALYTICS

1. Explain various Difference between SQL & NoSQL DBs ?

2. Explain advantages of NoSQL DBs ? Explain how MongoDB data will be inserted ?

3. Explain the steps - how COSMOS DB can be creatd with screens ?

4. Explain how to write JSON query in COSMOS DB ?

5. Explain major difference between databases & datawarehouses ?

6. Explain the architecture of datawarehouses ?

7. Explain what are Datamarts & how different from DATABASES ?
& mention the types of Datamarts too.

8. Explain OLAP & OLTP with examples ?

9. Explain what is BI & how BI helps business to take intelligent decisions ?

10.Explain how ETL works with Datawarehouses ?

Please create a word / pdf document, and send it to : avyuktitraining1@gmail.com

INTRODUCTION

This Assignment is given by manipal pro learn team on the basis of the training done in the forenoon session of this morning. The main objective behind this assignment is to master the theory and enhance knowledge over database, cosmos db , mongo db etc...

There are 10 questions and they are of easy to moderately difficult level. All the questions have been focused on what the trainer taught in the earlier sessions.Some questions have been answered partially due to unavailability of access.

This assignment gave me immense confidence in mastering the domain that has been assigned to me.

ANSWERS

1.

SQL	NOSQL
SQL databases are relational	NoSQL databases are non-relational.
These have a predefined schema.	These have dynamic schemas.
These are vertically scalable	These are horizontally scalable
These databases are table_based	These are of document,key_value , graph, or wide_column stores.
Better for multi-row transactions	Better for unstructured data like documents or JSON

2.

The limitations of traditional relational database technologies prompted the development of NoSQL databases. NoSQL databases are frequently more scalable and give better performance than relational databases. Furthermore, in comparison to the relational model, the flexibility and ease of use of their data models can speed up development, especially in the cloud computing environment.

Data insertion in mongo db:

To use MongoDB Compass to insert a single document, follow these steps: Go to the collection where you want to put the document: Click the database to which your target collection belongs in the left-hand MongoDB Compass navigation window. Select the target collection name from the database view.

3.

4.

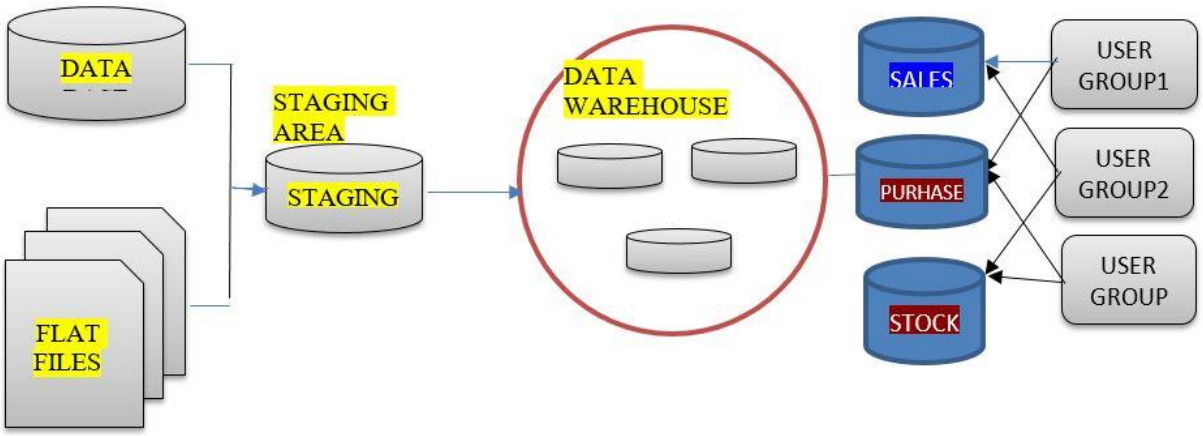
To Place Json Documents in Azure Cosmos DB

- Select "Data Explorer" from the drop-down menu.
- After that, go to the database and collection.
- Choose "Items" from the drop-down menu.
- You should see a button that says "Upload Item," which you may click.
- On the right, a new panel will open where you may choose a file to upload.
- Done.

5.

Database Management System	**Warehouse of Data**
• It assists with operating procedures. • Capture and keep track of the data. • Data that is currently available. • Within the confines of this single system, data is balanced. • When a transaction happens, the data is updated. • When data entry is completed, it is verified.	• It aids in the analysis and reporting of results. • Investigate the data. • Hundreds of years of history. • Data from many systems must be combined and balanced. • Scheduled procedures are used to update data. • After the data has been collected, it is verified.

6.



7.

DATAMARTS: A datamart is a simplified form of a data warehouse that focuses on a specific topic.

Datamarts concentrate on a single topic. As a result, they rely on a small number of sources. When compared to the time it takes to construct a warehouse, the time it takes to construct a mart is quite short.

A database is a system for storing transactional data (OLTP). A data mart is a repository for analytical data (OLAP). A database is a collection of information about a single topic's various features and activities. Data from many subjects will be stored in a data mart.

DATAMARTS ARE CLASSIFIED INTO THE FOLLOWING TYPES:

• Dependent data mart: data from OLTP systems is extracted and subsequently populated in the central DWH. The data is sent from the DWH to the data marts.

• Data is received directly from the source system by an independent data mart. This is appropriate for small businesses or small groups within a larger corporation.

• Hybrid data mart: Data is provided from both OLTP and data warehouse systems.

8.

Relational database(OLTP)	**Analytical data warehouse(OLAP)**

Useful for running the business.	Useful in the analyzing of business.
It is based on the entity relationship model.	It is based on star,snowflake and Fact constellation schema.
This provides High performance and fast.	It is highly flexible but not fast.
Database size ranges from 100MB to 1GB.	Data warehouse size ranges from 100GB to 1TB.
Provides primitive and highly detailed data.	Provides summarized and constellation schema.
This can be used for writing the data into database.	This can be used for reading the data from the database.
This contains the current data.	This contains historical data.
EX: All bank transactions made by the cutomers.	EX: Bank transactions made by a customer at a particular time.

9.

Business intelligence's only objective is to aid and promote better business decisions. BI gives businesses access to data that is crucial to their performance in a variety of areas, including sales, finance, marketing, and a slew of other departments.

BI software may be used to gain market intelligence and data can be analysed more effectively using BI tools, just as it can be used to improve competitive intelligence. Identifying consumer insights, buying habits, tracking customer behaviour, and projecting market trends are all part of this.

10.

ETL stands for Extract, Transform, and Load, and it is a Data Warehousing procedure. An ETL tool collects data from numerous data source systems, transforms it in the staging area, and then loads it into the Data Warehouse system.

RESULT

Almost all the test questions have been solved and presented successfully in the present document except few due to lack of data .

CONCLUSIONS

All the questions have been solved successfully with all the concepts that have been covered in the training session. It's really a great experience of learning while solving the cases. This assignment gave me immense confidence regarding my ability to upskill in new technologies.