

## ASSIGNMENT 5

NAME: Aishee Bhattacharya

BATCH: DXC-262-Analytics-B12-Azure

DATE: 03/06/2022

### 1. Explain various Difference between SQL & NoSQL DBs ?

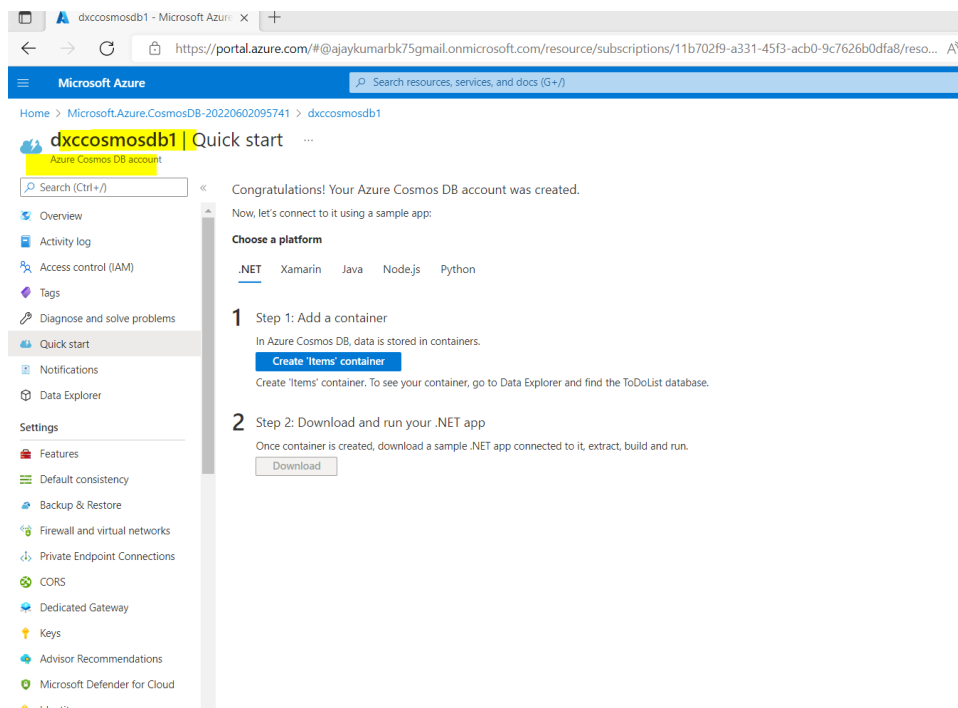
SQL Database	NoSQL Database
<ul style="list-style-type: none"><li>• Databases are categorized as Relational Database Management System (RDBMS).</li><li>• SQL databases have fixed or static or predefined schema.</li><li>• SQL databases display data in form of table so it is known as table-based database.</li><li>• SQL databases are vertical scalable.</li><li>• SQL databases are best suited for complex queries.</li></ul> <p>Eg: MySQL, Oracle, Sqlite, PostgreSQL and MS-SQL etc. are the example of SQL database.</p>	<ul style="list-style-type: none"><li>• NoSQL databases are categorized as Non-relational or distributed database system.</li><li>• NoSQL databases have dynamic schema.</li><li>• NoSQL databases display are horizontally scalable.</li><li>• NoSQL databases are not so good for complex queries because these are not as powerful as SQL queries.</li></ul> <p>Eg: MongoDB, BigTable, Redis, RavenDB, Cassandra, Hbase, CouchDB etc. are the example of nosql database.</p>

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

### Advantages of NoSQL Database:

- It supports query language.
- It provides faster performance.
- It provides horizontal scalability.
- It supports high scalable Data application (Big Data) and realtime data storage.

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



dxccosmosdb1 - Microsoft Azure

Home > Microsoft Azure > CosmosDB-20220602095741 > dxcosmosdb1

dxccosmosdb1 | Data Explorer


Azure Cosmos DB account


Search (Ctrl+/)


New Container Enable Azure Synapse Link New Notebook Connect to GitHub

# Welcome to Cosmos DB

Globally distributed, multi-model database service for any scale

**Launch quick start**  
Launch a quick start tutorial to get started with sample data

**New Container**  
Create a new container for storage and throughput

**Connect**  
Prefer using your own choice of tooling? Find the connection string you need to connect

**Recents**

**Top 3 things you need to know**

- [Advanced Modeling Patterns](#)  
Learn advanced strategies to optimize your database.
- [Partitioning Best Practices](#)  
Learn to apply data model and partitioning strategies.
- [Plan Your Resource Requirements](#)  
Get to know the different configuration choices.

**Learning Resources**

- [Get Started using an SDK](#)  
Learn about the Azure Cosmos DB SDK.
- [Master Complex Queries](#)  
Learn how to author complex queries.
- [Migrate Your Data](#)  
Migrate data using Azure services and open-source solutions.

SQL API

DATA

NOTEBOOKS

Notebooks is currently not available. We are working on it.

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Quick start

Notifications

**Data Explorer**

Settings

- Features
- Default consistency
- Backup & Restore
- Firewall and virtual networks
- Private Endpoint Connections
- CORS
- Dedicated Gateway
- Keys
- Advisor Recommendations
- Microsoft Defender for Cloud
- Identity

dxccosmosdb1 - Microsoft Azure

Home > Microsoft Azure > CosmosDB-20220602095741 > dxcosmosdb1

dxccosmosdb1 | Data Explorer


Azure Cosmos DB account


Search (Ctrl+/)

New Container Enable Azure Synapse Link New Notebook Connect to GitHub

# Welcome to Cosmos DB

Globally distributed, multi-model database service for any scale

**Launch quick start**  
Launch a quick start tutorial to get started with sample data

**New Container**  
Create a new container for storage and throughput

**Recents**

**Top 3 things you need to know**

- [Advanced Modeling Patterns](#)  
Learn advanced strategies to optimize your database.
- [Partitioning Best Practices](#)  
Learn to apply data model and partitioning strategies.
- [Plan Your Resource Requirements](#)  
Get to know the different configuration choices.

**New Container**

**Database id**

Create new Use existing

Music

**Container id**

Music

**Partition key**

Partition key

For small workloads, the item ID is a suitable choice for the partition key.

**Unique keys**

Add unique key

**Analytical store**

On Off

Azure Synapse Link is required for creating an analytical store container. Enable Synapse Link for this Cosmos DB account. [Learn more](#)

Enable

Advanced

OK

dxccosmosdb1 - Microsoft Azure

https://portal.azure.com/#@ajaykumarbk75gmail.onmicrosoft.com/resource/subscriptions/11b702f9-a331-45f3-acb0-9c7626b0dfa...

Microsoft Azure

Search resources, services, and docs (G+)

Home > Microsoft.Azure.CosmosDB-20220602095741 > dxcosmosdb1

dxccosmosdb1 | Data Explorer

Azure Cosmos DB account

Search (Ctrl+/)

New Container

Enable Azure Synapse Link

New Notebook

Connect to GitHub

New SQL Query

Open Q...

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Quick start

Notifications

Data Explorer

Settings

Features

Default consistency

Backup & Restore

Firewall and virtual networks

Private Endpoint Connections

CORS

Dedicated Gateway

SQL API

DATA

Music

Musicians

Items

Settings

Stored Procedures

User Defined Functions

Triggers

NOTEBOOKS

Notebooks is currently not available. We are working on it.

Welcome to

Globally distributed, multi-mode

Launch quick start

Launch a quick start tutorial to get started with sample data

New Conta

Create a new cc and throughpu

Recents

Top 3 things you ne

Advanced Modeling Patte

Learn advanced strategie: database.

Partitioning Best Practices:

Learn to apply data mode

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

https://portal.azure.com/#@ajaykumarbk75gmail.onmicrosoft.com/resource/subscriptions/11b702f9-a331-45f3-acb0-9c7626b0dfa...

Microsoft Azure

Search resources, services, and docs (G+)

Home > Microsoft.Azure.CosmosDB-20220602095741 > dxcosmosdb1

dxccosmosdb1 | Data Explorer

Azure Cosmos DB account

New Item

Save

Discard

Upload Item

SQL API

DATA

Music

Musicians

Items

Settings

Stored Procedures

User Defined Functions

Triggers

NOTEBOOKS

Notebooks is currently not available. We are working on it.

SELECT \* FROM c

Edit Filter

id

/Albums

a3965f71-e...

9bfe61f9-0...

a3728cc4-c...

AndersenFa...

Load more

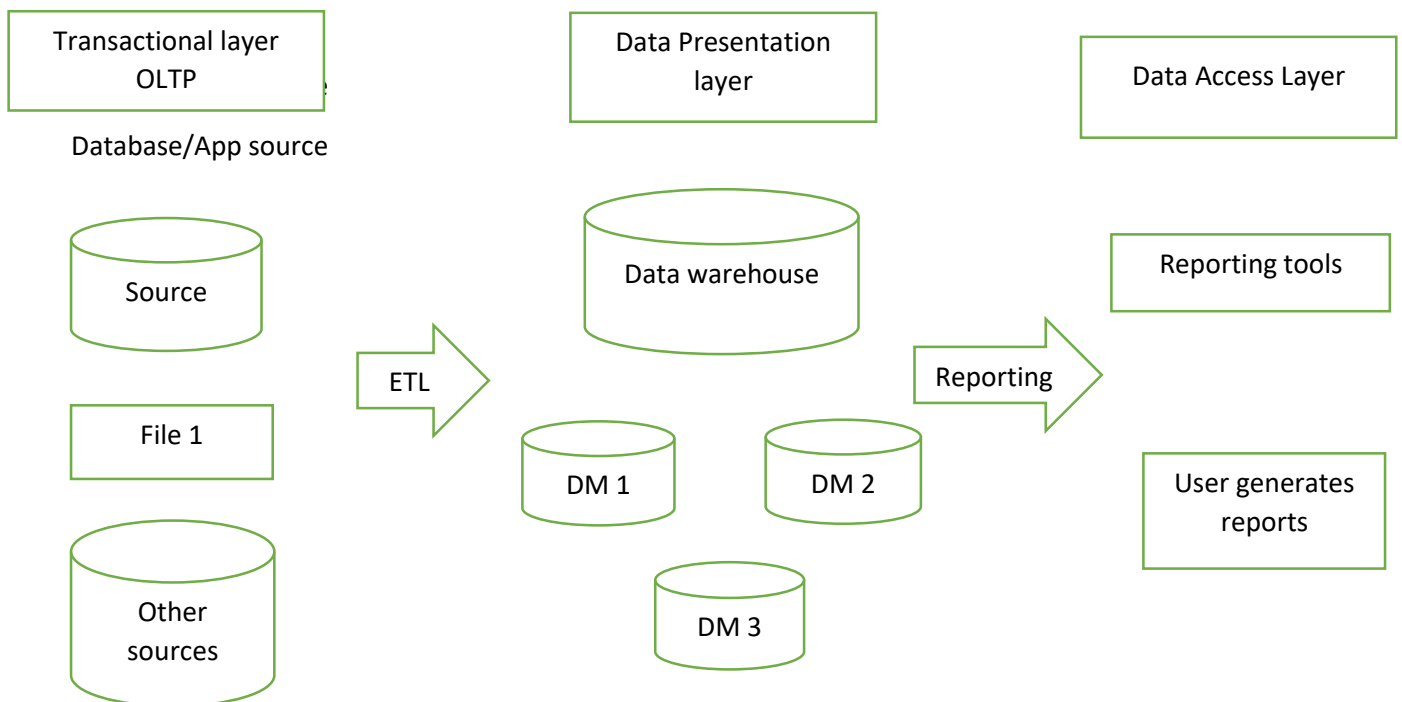
```
1 {
2   "id": "WakefieldFamily",
3   "children": [
4     {
5       "familyName": "Merriam",
6       "givenName": "Jesse",
7       "gender": "female",
8       "grade": 1,
9     },
10    {
11      "familyName": "Miller",
12      "givenName": "Lisa",
13      "gender": "female",
14      "grade": 8
15    }
16  ]
17 }
```

The screenshot shows the Microsoft Azure Data Explorer interface. The left sidebar displays a tree view with 'DATA' expanded, showing 'Music' and 'Musicians'. The 'Musicians' folder is selected, showing a list of items with IDs. The main pane displays a JSON document for a musician's account, with fields like 'id', 'balance', 'checkingAccount', and 'savingsAccount' highlighted in yellow. The document is a JSON array with one object representing a musician's account details.

5. Explain major difference between databases & datawarehouses ?

Databases	Datawarehouses
<ul style="list-style-type: none"> <li>A database is a collection of data organized for storage accessibility and retrieval.</li> <li>The data stored in various databases cannot be visualised. Hence it is transformed(cleaned and integrated) and loaded into data warehouse for analysis.</li> </ul>	<ul style="list-style-type: none"> <li>A central location where consolidated data from multiple locations(databases) are stored and analysed in order to make business decisions.</li> <li>DWH is maintained separately from an organization's operational database. End users access it when any information is needed</li> </ul>

6. Explain the architecture of datawarehouses ?



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

- Data marts are smaller version of data warehouse which deals with a single subject. They are focused on one area. Hence they draw data from a limited no. of sources.
- Time taken to build a data mart is less compared to the time taken to build a data warehouse.
- Data marts occupies less memory compared to Data warehouse.
- Types of Data mart:
  1. Dependent Data Mart
  2. Independent Data Mart
  3. Hybrid Data Mart

8. Explain OLAP & OLTP with examples ?

OLTP :

- Contains current data.
- Used for writing into the database
- Useful in running the business
- Based on entity relationship model

- Provides primitive and highly detailed data.
- Database size ranges from 100 MB to 1GB
- Provides high performance
- No. of records accessed is in tens
- Example :A supermarket server which records every single product purchased at the market.

#### OLAP:

- Contains historical data
- Used for reading data from data warehouse
- Useful in analysing the business
- Based on star, snowflake and Fact constellation schema
- Provides summarized and consolidated data.
- Data warehousing ranges from 100 GB to 1 TB
- Highly flexible but not fast
- No. of records accessed is in millions
- Example: An insurance company wants to know the no. of policies each agent has sold. This will help in better performance management of the agents.

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

BI (Business Intelligence) is the act of transforming raw or operational data into useful information for business analysis.

BI helps business to take intelligent decision:

- BI based on data warehouse technology extracts information from a company's operational system.
- The data is transformed(cleaned and integrated) and loaded into Data warehouses.
- Since this data is credible, it's used for business insights.

#### 10. Explain how ETL works with Datawarehouses ?

ETL(Extract,Transform & Load) is the process of extracting the data from various sources, transforming this data to meet your requirement and then loading it into a target data warehouse.

