CASE-STUDY (NUMBER - 5) – SOLUTION SUBMISSION

ON

AZURE ANALYTICS

BY

NAME –SUPRIYA BHARATHA ROLL NUMBER DXCAB1211

BATCH:DXC-262-ANALYTICS-B12-AZURE COMPANY – DXC TECHNOLOGY

TRAINING UNDER – MANIPAL PRO LEARN TRAINER NAME – MR. AJAY KUMAR DATE OF SUBMISSION – 3rd JUNE 2022 TOTAL NUMBER OFQUESTIONS: 10 EMPLOYEE DOMAIN - AZURE ANALYTICS

* THE QUESTIONS:

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 ?

1. Explain various Difference between SQL & NoSQL DBs ?

SOLUTION:

|  |  |
| --- | --- |
| SQL | NOSQL |
| Databases are belong to the relational database management system(RDBMS). | NOSQL databases are belong to non-relational or distributed databse system. |
| SQL databases have the fixed or static or predefined schema. | They have dynamic schema. |
| SQL is called as table based database because it is displays data in the form of tables. | They display databases as collection of key-value pair,graph,documents. |
| These are vertically scalable. | They are horizontally stable |
| This databases can be useful for complex queries. | These databases are not suitable for the complex queries due to powerfulas SQL queries. |
| Eg: MySQL,Oracle,MS-SQL etc. | Eg: MongoDB, RavenDB ,Redis , Bigtables and etc; |

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

SOLUTION:

* When compared to relational databases, NoSQL databases are often more scalable and provide superior performance.
* In addition, the flexibility and ease of use of their data models can speed development in comparison to the relational model, 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.

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

SOLUTION:

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

SOLUTION: 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.

1. Explain major difference between databases & datawarehouses ?

SOLUTION:

|  |  |
| --- | --- |
| **Database** | **Data Warehouse** |
| Is designed to record | Is designed to analyze |
| The database uses the Online Transactional Processing (OLTP) | Data warehouse uses Online Analytical Processing (OLAP). |
| The database helps to perform fundamental operations for your business | Data warehouse allows you to analyze your business. |
| 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. |
| Is an application-oriented collection of data | It is a subject-oriented collection of data |
| Generally limited to a single application | Stores data from any number of applications |
| Data is available real-time | Data is refreshed from source systems as and when needed |

1. Explain the architecture of datawarehouses ?

SOLUTION:

DATA BASE

SALES

USER GROUP1

AGGREGATE DATA

USER GROUP2

DATA WAREHOUSE

STAGING AREA

META DATA

PURHASE

STAGING

DATABASE

RAW DATA

USER GROUP3

STOCK

FLAT FILES

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

SOLUTION:

DATAMARTS: datamart is a smaller version of the data warehouse which deals with a single subject.

Datamarts are focused on one area. Hence, they draw from a limited number of sources.

Time taken to build marts is very less compared to time taken to build a warehouse.

A database is a transactional data repository (OLTP). A data mart is an analytical data repository (OLAP). A database captures all the aspects and activities of one subject in particular. A data mart will house data from multiple subjects.

TYPES OF DATAMARTS:

* Dependent data mart: the data first extracted from the OLTP systems and then populated in the central DWH. From the DWH, the data travels to the data marts.
* Independent data mart: The data is directly received from the source system. This is suitable for small organizations or small groups within an organization.
* Hybrid data mart: The data is fed both from OLTP systems as wll as data warehouse.

1. Explain OLAP & OLTP with examples ?

SOLUTION:

|  |  |
| --- | --- |
| 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. |

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

SOLUTION:

* BI is the act of transforming raw/ operational data into useful information for business analysis.

1. BI based on Data Warehouse technology extracts information from a company's operational systems.

2. The data is transformed (cleaned and integrated), and loaded into Data Warehouses.

3. Since this data is credible, it is used for business insights.

* The role of business intelligence is to improve an organization's business operations through the use of relevant data.
* Companies that effectively employ BI tools and techniques can translate their collected data into valuable insights about their business processes and strategies.

DB 1

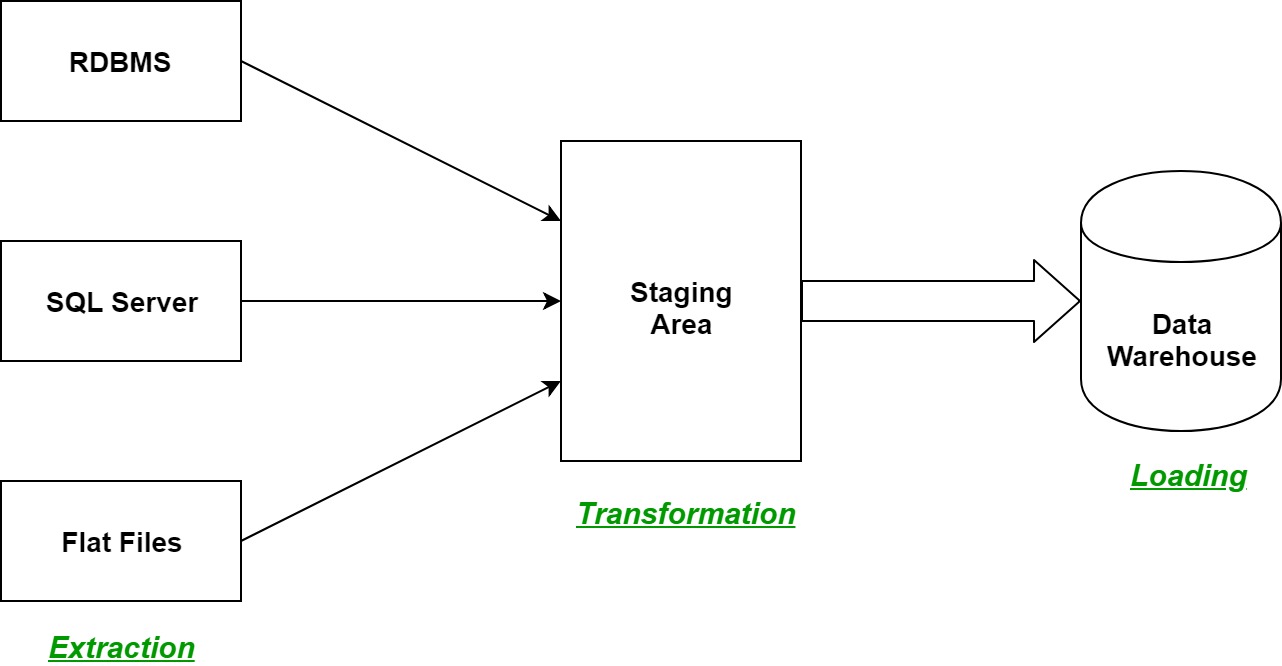
DATA WAREHOUSE

DB 2

END USER

1. Explain how ETL works with Datawarehouses ?

SOLUTION: ETL 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.



It is becoming increasingly common for data to be extracted from its source locations, then loaded into a target data warehouse to be transformed into actionable business intelligence. This process consists of three steps:

Extract — This step works similarly in both ETL and ELT data management approaches. Raw streams of data from virtual infrastructure, software, and applications are ingested either in their entirety or according to predefined rules.

Load — Here is where ELT branches off from its ETL cousin. Rather than deliver this mass of raw data and load it to an interim processing server for transformation, ELT delivers it directly to the target storage location. This shortens the cycle between extraction and delivery.

Transform — The database or data warehouse sorts and normalizes the data, keeping part or all of it on hand and accessible for customized reporting. The overhead for storing this much data is higher, but it offers more opportunities to mine it for relevant business intelligence in near real-time.

RESULTS:

WE HAVE SUCCESSFULLY ANSWERED ALLTHE QUESTIONS AS PER ASSIGNMENT REQUIREMENT.

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