1.What is Zero Copy Cloning in SnowFlake?

Zero-Copy Cloning is a Snowflake feature that makes a copy of a database without duplicating the data it contains. The clone operation takes a snapshot of the source data when the clone is created and makes this data available to the cloned object.

2.What is Time Travel in SnowFlake?

Snowflake Time Travel enables accessing historical data at any point within a defined period. It serves as a powerful tool for performing the following tasks:

Restoring data-related objects that might have been accidentally or intentionally deleted.

Duplicating and backing up data from key points in the past.

Analyzing data usage/manipulation over specified periods of time.

3.Temporary Vs Transient Vs External vs Permanent tables?  
Temporary tables: This are used for storing intermediate or temporary data during a session. They are created and exist only within the session where they are created. Once the session ends or the user disconnects, the temporary table is automatically dropped, and the data is deleted.

Transient Table:

Transient tables are similar to regular permanent tables but with a time-to-live (TTL) defined. They are useful for storing data that is not critical and can be removed after a certain period to manage costs and storage resources.

External Table:

External tables in Snowflake are references to data stored outside the Snowflake environment, such as in cloud storage services like Amazon S3 or Azure Blob Storage. They provide a way to query external data without actually loading it into Snowflake, making it a more cost-effective solution for certain use cases.

Permanent Tables:

Permanent tables are the standard tables in Snowflake, which persist data indefinitely until explicitly dropped by the user or administrator.Permanent tables store data within Snowflake, either in the default cloud-based storage provided by Snowflake or in a user-defined storage location (e.g., using Snowflake's external stage).

4.FailSafe and Retention implications in all tables?

Failsafe continuously protects your data by making and keeping copies of your tables and data. Failsafe maintains historical versions of your tables allowing you to quickly restore data in the event of an error.

Retention (Time Travel)

Enables to access previously saved historical data in ,Without the need for additional database copies or separate backups, Time Travel enables you to query your tables as of a given point in the past and evaluate historical data.

5.What will happen in zero copy cloning when you change the data on source?

When you change the data on the source table after creating a zero-copy clone, the changes made on the source will not be reflected in the clone. The clone maintains a point-in-time snapshot of the source table at the moment it was cloned. Any subsequent changes to the source table do not affect the data in the clone, and the two tables become independent of each other.

6.What will happen to source when there is change in Zero Copy Cloning?

When a clone is made, the clone process creates a snapshot of the source data and makes it accessible to the copied object. Once the clone has reached this stage, it is no longer dependent on the source, thus any modifications made do not affect the other.

7.What is a Task in snowflake?

A Task refers to a scheduled unit of work or a set of actions that you want to execute at specified intervals. It can have dependencies, notifications, and resource allocations, making them useful for automating ETL processes, data backups, and other routine database maintenance tasks.

8.How to load 10k records from salesforce to snowflake?

1. Set up Salesforce and Snowflake accounts.
2. Obtain Salesforce API credentials for authentication.
3. Configure Snowflake connection settings.
4. Prepare the data in Salesforce and identify the records to extract.
5. Use Salesforce API or tools like Data Loader to extract the data.
6. Optionally, perform data transformations if required.
7. Load the data into Snowflake using COPY INTO or ETL tools.
8. Monitor the process for errors or issues.
9. Validate the data in Snowflake for accuracy.
10. Consider scheduling regular updates for keeping the data up to date.

9. How do you perform performance tuning in your project?

1. Optimize data streaming and data ingestion processes to handle large volumes of real-time data from monitoring devices.
2. Implement redundancy and failover mechanisms to ensure system availability in case of hardware failures.
3. Apply appropriate data preprocessing and filtering to remove noise, outliers, or irrelevant data.
4. Divide the monitoring tasks into smaller parallelizable units and process them concurrently.

10.How do you handle duplicates in Kafka?

1. Set up monitoring and alerting systems to quickly identify any anomalies, such as sudden spikes in duplicate messages, and take corrective actions.
2. Consider using a unique identifier for the messages and handling out-of-order delivery at the consumer end to avoid processing duplicates.
3. Kafka ensures that each message is delivered to only one consumer within a consumer group, reducing the chances of duplicate processing.
4. Configure the Kafka producer to be idempotent. With idempotence enabled, the producer ensures that messages with the same key and value are not sent multiple times, even in case of retries or network failures.

11.How do you read data from external API in Json format and send notification to specific busimnees unit?

1. Obtain the necessary API credentials and endpoints from the external API provider.
2. Store the credentials securely in Azure Key Vault.
3. Set up a trigger for the Logic App, e.g., Recurrence or HTTP trigger, to start the workflow.
4. Add an action in the Logic App to retrieve data from the external API.
5. Use the API credentials stored in Azure Key Vault to authenticate the API request.
6. Use the Azure Notification Hubs connector within the Logic App to send notifications by using send push notification action to the specific business unit.
7. Configure the notification with the formatted data retrieved from the external API.
8. Select the appropriate notification channel based on the business unit's preferences (e.g., email, SMS, push notifications).