Snowflake Interview Questions & Answers - 2025 Edition

Q: What is Snowflake and how is it different from traditional databases?

A: Snowflake is a cloud-based data warehouse that separates compute, storage, and services layers. It scales automatically, supports semi-structured data, and is cloud-agnostic.

Q: What is a virtual warehouse in Snowflake?

A: A virtual warehouse provides compute resources for loading, querying, and transforming data. It can scale up, scale out, or auto-suspend when idle.

Q: What are micro-partitions in Snowflake?

A: Micro-partitions are internal storage units (around 16MB each) that enable automatic data clustering and pruning for faster queries.

Q: What is Time Travel in Snowflake?

A: Time Travel allows querying, restoring, or cloning data from previous points in time. Useful for recovery and auditing.

Q: What is a stage in Snowflake?

A: A stage is a storage location for files before loading them into tables. It can be internal (Snowflake-managed) or external (e.g., AWS S3, Azure Blob).

Q: How do you load data into Snowflake?

A: Use the COPY INTO command to load data from a stage into a table, specifying file formats and options for error handling.

Q: What is a Snowpipe?

A: Snowpipe is a continuous data ingestion service that automatically loads new files from cloud storage into Snowflake as they arrive.

Q: Explain the MERGE command in Snowflake.

A: MERGE allows you to perform INSERT, UPDATE, and DELETE in one statement based on matching conditions. Commonly used for upserts or data synchronization.

Q: What is the VARIANT data type?

A: VARIANT stores semi-structured data like JSON, XML, or Avro. It allows flexible schema and supports querying using dot notation.

Q: What is a materialized view in Snowflake?

A: A materialized view stores precomputed query results for performance optimization. It automatically refreshes when underlying data changes.

Q: What are transient and temporary tables?

A: Transient tables persist beyond sessions but lack fail-safe recovery. Temporary tables last only for the session that created them.

Q: What is zero-copy cloning in Snowflake?

A: Zero-copy cloning creates a new table, schema, or database instantly without copying data. It references existing micro-partitions until changes occur.

Q: What are caching mechanisms in Snowflake?

A: Snowflake uses result caching, metadata caching, and data caching to speed up queries by reusing prior computations.

Q: How does Snowflake ensure data security?

A: Snowflake encrypts data end-to-end, supports multi-factor authentication, role-based access control (RBAC), and network policies.

Q: How do you optimize query performance in Snowflake?

A: Use clustering keys, pruning, caching, correct warehouse sizing, and limit unnecessary transformations.

Q: What is a resource monitor in Snowflake?

A: A resource monitor tracks credit usage for warehouses and can automatically suspend them when limits are reached.

Q: What is automatic scaling in Snowflake?

A: Automatic scaling adds or removes compute clusters based on query load, ensuring consistent performance.

Q: What are tasks in Snowflake?

A: Tasks are scheduled SQL statements used for automation (e.g., running ETL jobs or refreshing materialized views).

Q: What is data sharing in Snowflake?

A: Data sharing allows secure, real-time access to datasets across different Snowflake accounts without copying or moving data.

Q: How do you integrate Snowflake with Power BI?

A: Use the native Snowflake connector in Power BI. Connect with credentials and select either Import or DirectQuery mode for reporting.