50 Snowflake Cortex Interview Questions

Production-Grade Data Engineering. No Fluff.

Comprehensive guide covering Cortex Analyst, Agent, Document AI, SiS compatibility, and production patterns.

Snowbrix Academy | snowbrix.academy/cortex-masterclass

Section 1: Snowflake Cortex Fundamentals

Q1: What is Snowflake Cortex and its three main components?

Snowflake Cortex is Snowflake's AI/ML service with three components: (1) Cortex Analyst for natural language to SQL, (2) Cortex Agent for multi-step investigations, (3) Cortex Search for semantic search and RAG pipelines.

Q2: How does Cortex Analyst differ from traditional BI tools?

Traditional BI requires dashboard navigation. Cortex Analyst enables conversational analytics where users ask questions in natural language. It uses semantic YAML models to generate SQL dynamically.

Q3: What is a semantic YAML model?

A structured metadata file defining tables, columns, relationships, and verified queries. Required for Cortex Analyst to generate correct SQL.

Q4: How to enable Cortex in Snowflake?

Use supported regions (us-east-1, us-west-2) or enable cross-region: ALTER ACCOUNT SET CORTEX_ENABLED_CROSS_REGION = 'ANY_REGION';

Q5: Cortex pricing model?

Consumption-based: per API call, tokens processed, and warehouse compute. Optimize with AUTO_SUSPEND, caching, and appropriate TARGET_LAG.

Section 2: Cortex Agent & Tool Orchestration

Q11: What is Cortex Agent?

Autonomous reasoning system performing multi-step investigations using tools. Unlike Analyst (single-turn), Agent plans, executes tools, analyzes, and iterates for complex workflows.

Q12: What is a tool in Cortex Agent?

A function the agent calls to gather information. Includes name, description, parameters. Agent decides which tools to call based on descriptions.

Q13: How does Agent decide which tools to call?

Uses ReAct pattern: Reasoning (analyze question), Acting (select and execute tool), Observation (analyze result), repeat until done.

Q14: Tool description best practices?

Be specific about purpose, include when to use, specify output format, mention constraints, provide examples.

Section 3: Document AI & Cortex Search

Q21: What is Cortex Search?

Hybrid search combining keyword (BM25) and semantic (vector embeddings) search. Matches both exact terms and semantic meaning.

Q22: How to create Cortex Search service?

CREATE CORTEX SEARCH SERVICE ON document_text ATTRIBUTES ... AS (SELECT ...);

Q23: What is TARGET_LAG?

Controls index freshness: 1 minute (near real-time, high cost) to 24 hours (daily refresh, low cost). Tradeoff between freshness and cost.

Q25: What is RAG pipeline?

Retrieval-Augmented Generation: Search relevant docs (Cortex Search), augment LLM context with retrieved docs, generate answer citing sources.

Section 4: Streamlit in Snowflake Compatibility

Q31: SiS vs local Streamlit?

SiS limitations: No st.chat_input (use st.text_input+button), uppercase Pandas columns (normalize), use _snowflake.send_snow_api_request().

Q32: How to handle chat in SiS?

Use st.text_input + st.button instead of st.chat_input. Store history in st.session_state.

Q33: Why uppercase column names?

Snowflake stores identifiers as UPPERCASE. Always normalize: df.columns = [c.lower() for c in df.columns]

Q34: How to call Cortex APIs in SiS?

Use _snowflake.send_snow_api_request(method='POST', url=endpoint, body=body)

Section 5: Production Patterns

Q41: Error handling for Cortex APIs?

Multi-layer: (1) Retry with exponential backoff, (2) Validate responses, (3) Log errors to table, (4) Provide user fallback.

Q42: Semantic YAML version control?

Use Git, semantic versioning, automated validation, test suite for verified queries, rollback plan.

Q43: Optimize warehouse costs?

Right-size warehouses, auto-suspend=60, multi-cluster with ECONOMY policy, cache semantic models, monitor with cost dashboard.

Q44: Implement RBAC?

Create roles per user type, grant table/app access, limit semantic model scope by role, use row-level security.

Q50: Building production Cortex app?

Requirements (1 day), data modeling (2 days), semantic model (2 days), Streamlit app (2 days), testing (3 days). Total: 10 days MVP.

Want to Master Snowflake Cortex?

Join the **Snowflake Cortex Masterclass** launching February 20, 2026.

Build 3 production AI apps: Cortex Analyst, Data Agent, Document AI + RAG

Early Bird: Rs.4,999 (Save Rs.2,000, first 50 students)

snowbrix.academy/cortex-masterclass