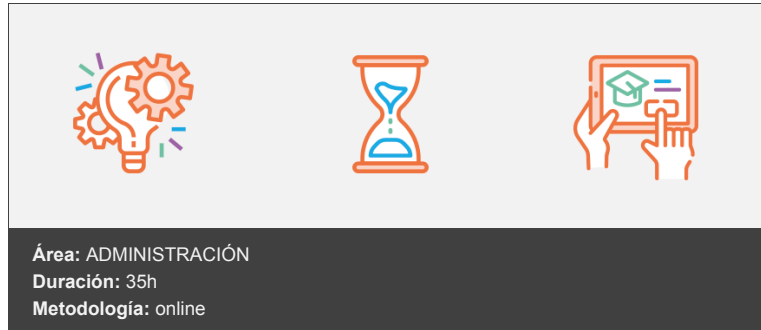


Snowflake. Data warehouse en la nube



Objetivos

Entender la arquitectura de Snowflake para entender por qué es la plataforma de gestión y almacenamiento de datos en la nube más popular. Cargar datos en Snowflake desde diferentes tipos de fuente tanto estructuradas como no estructuradas. Realizar y configurar la carga de datos en modo masivo (bulk) o bien en tiempo real con Snowpipe. Conseguir cargar desde plataformas en la nube como AWS, Azure o GCP. Proteger nuestros datos en Snowflake y compartirlos de manera segura garantizando un ciclo de vida de los datos confiable. Automatizar los procesos de ingesta, transformación y carga con la definición de tareas y streams en Snowflake. Conectar Snowflake con plataformas analíticas que consuman los datos como Power BI. Administrar los roles y permisos apropiados para garantizar la seguridad de nuestros datos.

Contenidos y estructura del curso

Introducción y arquitectura de Snowflake

¿Qué es un data warehouse?

Arquitectura de Snowflake y sus ventajas

Ejemplo de escalado computacional

Ejemplo de almacenamiento y consulta

Ediciones de Snowflake y precios

Descripción de las ediciones de Snowflake

¿Cómo calcular el coste de Snowflake?

Registro en Snowflake y primeros pasos

Crear un nuevo warehouse mediante interfaz o usando SQL

Creación de un Warehouse mediante la interfaz gráfica de usuario (UI)

Creación de un Warehouse usando SQL

¿Cómo explorar bases de datos y tablas en Snowflake?

Accediendo a la información de esquema en Snowflake

Monitorizar el uso de Snowflake

Resumen

Carga de datos en Snowflake

Adquisición e ingesta de datos en Snowflake (stages)

Carga de datos desde la interfaz fichero local

Cómo crear un stage

Carga de datos utilizando el comando COPY

Transformación de datos en Snowflake

Transformaciones durante la carga (ELT)

Manejo de errores - Opción COPY: "ON_ERROR"

Reutilizar formato ficheros con objeto file_format

Modo de validación y tratamiento de errores

Resumen

Carga de datos no estructurados

¿Qué son los datos NO estructurados?

Creación del stage y el fichero bruto con datos no estructurados (JSON)

Parseo y análisis del fichero no estructurado JSON
Manejo de datos nested y arrays
Consulta de datos nested y arrays en Snowflake
Manipulación de arrays en Snowflake
Jerarquías y flatten del fichero no estructurado JSON
Inserción de datos no estructurados JSON en la tabla objetivo (estructurados)
Resumen

Carga de datos desde cloud (AWS, Azure, GCP)
¿Cómo cargamos datos desde la nube?
Carga desde AWS S3
Carga desde Azure Blob Storage
Carga desde Google Cloud Storage (GCS)
Creación del bucket y carga de ficheros en la nube
Crear rol y política en la plataforma en la nube
Políticas en AWS
Políticas en Azure
Políticas en GCP
Crear objeto de integración en Snowflake
Configuración para Azure Blob Storage
Configuración para Google Cloud Storage
Cargar datos a la tabla objetivo de Snowflake
Optimización de la carga de datos desde la nube en Snowflake
Elegir el método de carga correcto
Dimensionar correctamente los File Sizes
Utilizar el formato de archivo eficiente
Comprimir archivos de datos
Uso de Staging Areas
Paralelizar las cargas de datos
Monitoreo y ajuste
Resumen

Snowpipe (ingesta en tiempo real)
¿Qué es Snowpipe y los pasos para la creación de pipes?
Creación de stage para Snowpipe
Definición de un External Stage
Configuración de la File Format
Validación y Pruebas del Stage
Creación del pipe
Configurar el pipe y las notificaciones
Gestión de pipes
Modificación y gestión de pipes existentes
Mostrar los pipes creados
Eliminación de un pipe
Migración y Carga desde una BBDD SQL
Exportación e importación manual
Utilización de Snowpipe
Herramientas de ETL y ELT
AWS Database Migration Service (AWS DMS) o Azure Data Factory
Scripting y Automatización
Uso de Snowpipe con Kafka
Configuración de Snowpipe
Conexión con Kafka
Monitoreo y Gestión de Errores
Opciones Avanzadas
Resumen

Ciclo de vida, protección y compartición de datos en Snowflake
Time Travel y Fail Safe
Tipos de tabla en Snowflake
Compartición de datos con otra cuenta de Snowflake
Creación de un Share
Asignar permisos a la cuenta del consumidor
Acceso a los datos por parte del consumidor
Actualizar y revocar un Share
¿Cómo conectar Snowflake y Power BI?
¿Cómo conectar Snowflake y Python?

- Instalación del Snowflake Connector para Python
- Configuración de la conexión
- Ejecución de consultas SQL
- Gestión de excepciones
- Uso de pandas para análisis de datos
- Protección - Recomendaciones para evitar brechas de seguridad
- Recomendaciones para usuarios
- Recomendaciones para empresas
- Resumen

- Creación de tareas y streams en Snowflake
 - ¿Qué son las tareas en Snowflake?
 - Creación de tareas
 - Definición de la tarea
 - Activación de la tarea
 - Monitorización de ejecuciones
 - Uso de CRON para crear tareas con periodicidad flexible
 - Ejemplo de creación de una tarea con CRON
 - Ejemplo de ejecución semanal
 - Consideraciones adicionales
 - Creación de Árboles de Tareas y verificar Histórico de Programación
 - ¿Qué son los Streams en Snowflake?
 - Funcionamiento de los Streams
 - Creación de un Stream
 - Ventajas de Utilizar Streams
 - Creación de Stream y operación INSERT
 - Uso de Stream para operación UPDATE
 - Uso de Stream para operación DELETE
 - Procesamiento de todos los cambios de datos en Streams
 - Combinar Tareas y Streams en Snowflake
 - Consideraciones y mejores prácticas
 - Resumen

- Administración de roles y permisos en Snowflake
 - Conceptos clave para el control de acceso en Snowflake
 - Rol ACCOUNT ADMIN
 - Rol SECURITY ADMIN
 - Gestión de usuarios y roles
 - Configuración de políticas de seguridad
 - Rol SYS ADMIN
 - Privilegios del rol SYS ADMIN
 - Administración de permisos
 - Rol USER ADMIN
 - Rol PUBLIC y roles personalizados
 - Roles, máscaras y cifrado de datos
 - Cifrado de datos a roles
 - Resumen

Metodología

En Critería creemos que para que la formación e-Learning sea realmente exitosa, tiene que estar basada en contenidos 100% multimedia (imágenes, sonidos, vídeos, etc.) diseñados con criterio pedagógico y soportados en una plataforma que ofrezca recursos de comunicación como chats, foros y conferencias...Esto se logra gracias al trabajo coordinado de nuestro equipo e-Learning integrado por profesionales en pedagogía, diseño multimedia y docentes con mucha experiencia en las diferentes áreas temáticas de nuestro catálogo.

Perfil persona formadora

Esta acción formativa será impartida por un/a experto/a en el área homologado/a por Critería, en cumplimiento con los procedimientos de calidad, con experiencia y formación pedagógica.

*En Critería queremos estar bien cerca de ti, ayúdanos a hacerlo posible:
¡Suscríbete a nuestro blog y síguenos en redes sociales!*

Blog de Criteria

