CELEBRATING 14 YEARS

QualityThought®

Transforming Dreams! Redefining Future!

snowflake

Snowflake Introduction

Snowflake is a cloud-based data warehousing platform that allows organizations to store, manage, and analyze their data in a scalable and flexible manner. It is designed to address the challenges of traditional on-premises data warehouses by leveraging the power and flexibility of cloud computing.

What are job opportunities in Snowflake?

Snowflake Developer/Engineer

Responsibilities may include designing and developing data pipelines, working on data modeling, and ensuring optimal performance in a Snowflake environment.

Snowflake Architect

Architects are involved in designing the overall data architecture, including the configuration and optimization of Snowflake databases for specific business needs.

Data Engineer (with Snowflake Skills)

Data engineers working with Snowflake focus on building, maintaining, and optimizing data pipelines for effective data processing.

Data Analyst (with Snowflake Experience)

Data analysts use Snowflake to analyze and interpret data, creating insights and reports for decision-making within organizations.

BI Developer (with Snowflake)

Business Intelligence (BI) developers leverage Snowflake to create and maintain dashboards, reports, and visualizations for business users.

Cloud Data Warehouse Consultant

Consultants specializing in Snowflake often work with organizations to implement and optimize Snowflake solutions, providing expertise in best practices and performance tuning.

Database Administrator (Snowflake)

Database administrators focused on Snowflake are responsible for managing and maintaining Snowflake databases, ensuring data integrity, security, and optimal performance.

Data Scientist (with Snowflake Skills)

Data scientists may use Snowflake as part of their toolkit for accessing and analyzing data, especially in scenarios where a cloud data warehouse is preferred.

DevOps Engineer (with Snowflake)

DevOps professionals with Snowflake skills may be responsible for managing the deployment, monitoring, and automation of Snowflake environments.

qSolution Architect (with Snowflake)

Solution architects design end-to-end data solutions, including integration with Snowflake, to meet business requirements.

CELEBRATING 14 YEARS

QualityThought®

Transforming Dreams! Redefining Future!

snowflake

Course Curriculum

- Introduction to Cloud
- Data Warehousing Concepts
- Introduction to Snowflake
- Connecting to Snowflake
- Loading / Unloading Data into/from Snowflake
- Using Snowflake
- Snowflake Continuous Data Pipelines
- Sharing Data in Snowflake
- Managing Your Snowflake Account
- Understanding Snowflake Data Transfer Billing
- Managing Security in Snowflake
- Performance Tuning
- Interview Questions

Introduction to Cloud

- What is Cloud
- Different Cloud Vendors
- Advantages of Cloud over On-Premises

Data WareHousing Concepts

- What is a Data Warehouse
- Data Base Vs Data Warehouse
- Data Warehouse Architecture
- Why do we need Data Warehouse
- OLTP Vs OLAP
- What is ETL

Introduction to Snowflake

- Key Concepts & Architecture
- Cloud Platforms
- Snowflake Editions
- Overview of Key Features
- Overview of the Data Lifecycle
- Continuous Data Protection

Connecting to Snowflake

- Overview of the Ecosystem
- Snowflake Partner Connect
- SnowSQL (CLI Client) - Installation, Configuring and Usage

Loading / Unloading Data into/from Snowflake

- Overview of Data Loading/Unloading
- Data Loading/Unloading Considerations
- File Formats
- Preparing to Load/Unload Data
- Bulk Loading/Unloading from a Local File System
- Bulk Loading/Unloading from Amazon S3
- Loading Continuously Using Snowpipe
- Loading Using the Web Interface
- Querying Data in Staged Files
- Querying Metadata for Staged Files
- Transforming Data During a Load
- Troubleshooting Bulk Data Loads

Using Snowflake

- Web Interface
- Virtual Warehouses
- Databases, Tables & Views
- Queries
- Date & Time Data
- Semi-structured Data
- Snowflake Time Travel & Fail-safe

Snowflake Continuous Data Pipelines

- Streams,
- Tasks
- Snow pipe

Sharing Data in Snowflake

- Introduction to Data Sharing
- Data Providers
- Data Consumers
- Working with Shares
- Using Secure Objects to Control Data Access
- Configuring a Reader Account
- Managing Reader Accounts

Managing Your Snowflake Account

- System Usage & Billing
- Understanding Snowflake Credit and Storage Usage

Understanding Snowflake Data Transfer Billing

- Monitoring Credit and Storage Usage
- Resource Monitors
- Parameter Management
- User Management

Managing Security in Snowflake

- Summary of Security Features
- Network Policies
- AWS Private Link & Snowflake
- Multi-Factor Authentication (MFA)
- Federated Authentication & SSO
- Access Control in Snowflake (DAC & RBAC)
- Data Encryption

Performance Tuning

- Query Tuning
- Clustering
- Views
- Materialized Views
- Secure Views
- Performance Considerations

Prerequisites

SQL (Structured Query Language)

SQL is a standardized programming language designed for managing and manipulating relational databases. It provides a set of commands for querying databases to retrieve, insert, update, or delete data. SQL is widely used for data management and is essential for working with relational database management systems (RDBMS) like MySQL, PostgreSQL, Oracle, SQL Server, and SQLite

Database

A database is an organized collection of structured data stored electronically. It can consist of tables, rows, and columns, with each table representing a specific type of data entity. Databases are designed to efficiently store, retrieve, and manage large volumes of information. There are various types of databases, including relational databases, NoSQL databases, and in-memory databases, each serving different purposes based on the nature of the data and the requirements of applications.

ETL (Extract, Transform, Load)

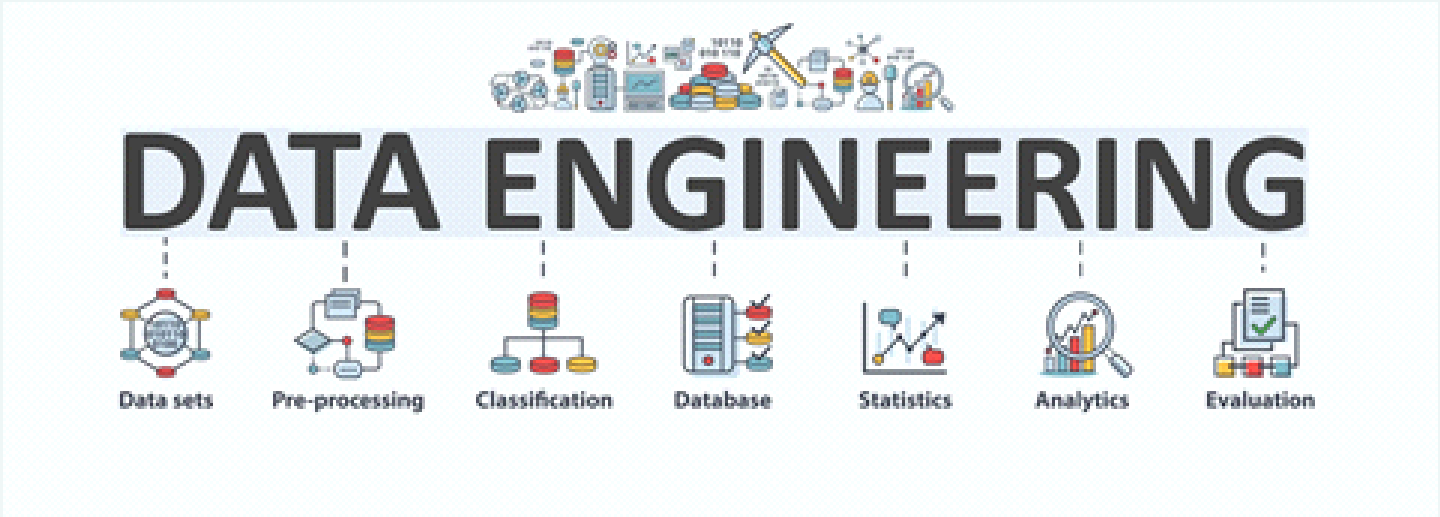
ETL is a data integration process that involves extracting data from source systems, transforming it to meet business requirements, and loading it into a target data warehouse or database. The ETL process typically includes data cleansing, enrichment, aggregation, and formatting. ETL is commonly used in traditional data warehousing scenarios, where data is extracted from operational databases, transformed into a suitable format, and then loaded into a central data repository for analysis and reporting

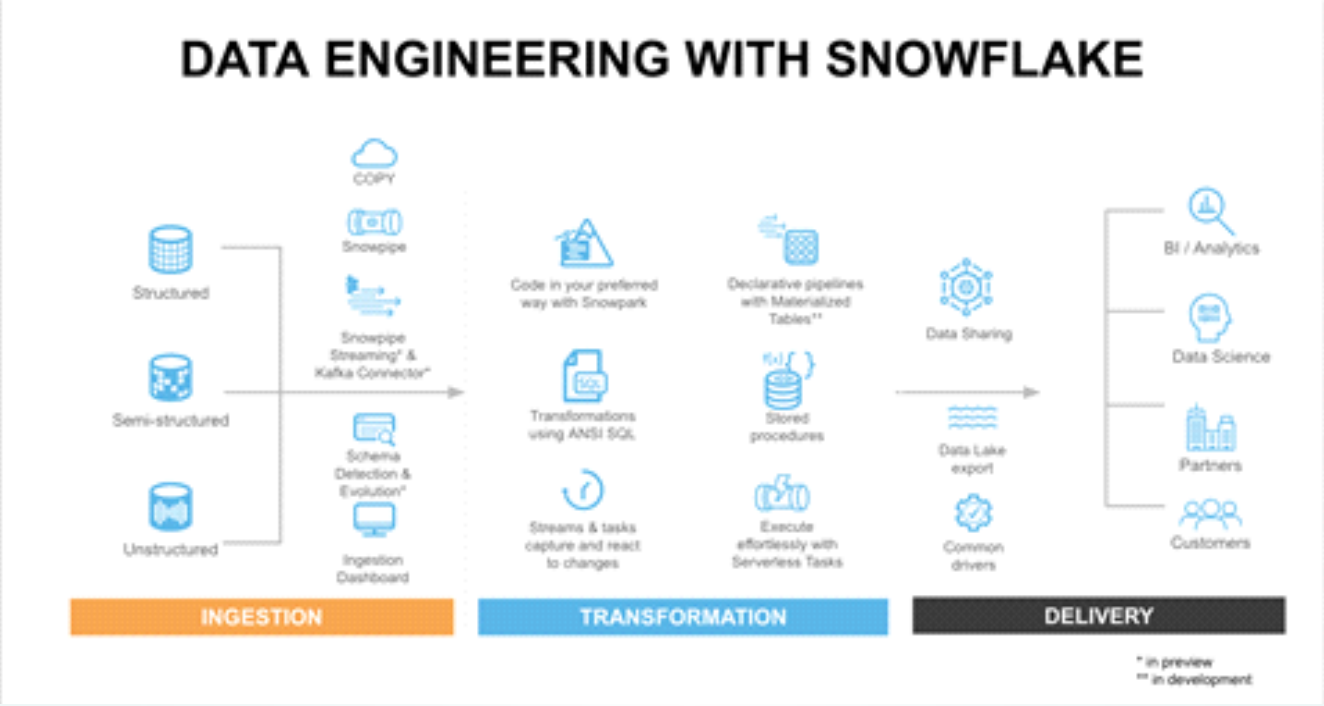
Python (Optional not Mandatory)

Python is a high-level, interpreted programming language known for its simplicity, readability, and versatility. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming. Python is widely used for web development, data science, artificial intelligence, automation, and scripting. It has a large and active community, extensive libraries, and frameworks, making it a popular choice for developers across various domains.

Java (Optional not Mandatory).

Java is a versatile, object-oriented programming language developed by Sun Microsystems. It is designed to be platform-independent, allowing Java applications to run on any device with the Java Virtual Machine (JVM). Java is known for its "write once, run anywhere" philosophy, making it suitable for building cross-platform applications. It is widely used for developing enterprise-level applications, mobile applications (Android), web applications, and large-scale systems. Java's strong typing, reliability, and extensive libraries contribute to its popularity in the software development industry




DATA ENGINEERING WITH SNOWFLAKE
COPY
Snowpipe
Snowpipe Streaming* & Kafka Connector*
Schema Detection & Evolution*
Ingestion Dashboard
Structured
Semi-structured
Unstructured
Code in your preferred way with Snowpark
Transformations using ANSI SQL
Streams & tasks capture and react to changes
Declarative pipelines with Materialized Tables**
Stored procedures
Execute effortlessly with Serverless Tasks
Data Sharing
Data Lake export
Common drivers
BI / Analytics
Data Science
Partners
Customers
INGESTION
TRANSFORMATION
DELIVERY
* in preview
** in development


Data Sources
OLTP Databases
Enterprise Applications
Third-Party
Web/Log Data
IoT
ETL, Streaming
Data Warehouse
Data Lake
Data Engineering
Data Exchange
Data Applications
Data Science
snowflake
Google Cloud
aws
Azure
Data Consumers
Data Monetization
Operational Reporting
Ad Hoc Analysis
Real-time Analytics

Our Students Are Placed In

TCS · Capgemini · LTIMindtree · accenture · 66degrees · AFFINE · High Noon Consulting
AGILISIUM · ALTIMETRIK · brillio · CGI · cisco · Cloud4C · amazon
CLOUD9 · CloudCover · Coforge · Cognizant · Convergytics · CSS CORP · Medtronic
databricks · dataflex · DXC TECHNOLOGY · Deloitte. · EMIDS · Foray · STIC Soft
genpact · HCL · HSBC · I2O · IBM · GAPRUS IT
Infosys · Legato · MINDGRAPH · Mindtree · motivitylabs · photon · meslova
pwc · quantiphi · RENAULT · SIGMASOFT · SOFTILITY · syren cloud · splashBI
TIGER ANALYTICS · wipro · ENVISION · virtusa · ValueLabs · zensor · SOLIX
JRD SYSTEMS · ValueLabs · archents · magnasoft · Innominds · techforce.ai · techolution
SOLUGENIX · Nendrasys · S&P Global · Aadvi · iPrism Technologies · ENVISION · Sunseaz Technologies Pvt Ltd