

## INDEX

<b>I&amp;D (Advance SQL + Teradata + Snowflake) LoT Course Structure .....</b>	<b>2</b>
<b>Data Warehouse Concepts .....</b>	<b>3</b>
<b>Data Modeling for Business Intelligence .....</b>	<b>4</b>
<b>ETL Basics .....</b>	<b>5</b>
<b>Software Testing for BI.....</b>	<b>6</b>
<b>Python Programming .....</b>	<b>6</b>
<b>Oracle SQL Advance.....</b>	<b>8</b>
<b>Teradata Basics.....</b>	<b>11</b>
<b>Online training with Badge 1 &amp; Badge 2.....</b>	<b>13</b>
<b>Snowflake Lab Guide .....</b>	<b>14</b>
<b>Snowflake Lab Assignments .....</b>	<b>15</b>

## I&D (ADVANCE SQL + TERADATA + SNOWFLAKE) LoT COURSE STRUCTURE

I&D (Advance SQL + Teradata + Snowflake) LoT provides exposure to a band of data warehousing technologies. It focuses on application development for data warehouses. The following table lists the course structure for I&D LoT.

Sr. No.	Course	Duration (In Days)	Remarks
1	Discover	0	Online
2	Soft Skills Foundation – Part 1	1	Soft Skills Part 1 (Saturday)
2	Data Warehouse Concepts	1	
3	Data Modelling for Business Intelligence	1.5	
4	ETL Basics	0.25	
5	Software Testing for BI	0.25	
6	Python Programming	4	Project Kick Off
7	Oracle SQL Advance	8	
8	Soft Skills Foundation – Part 2	1	Soft Skills Part 2 (Saturday)
9	Module 1 Assessment	1	Coding and MCQ
10	Teradata Basic	5	Sprint 1 Implementation, MCQ
11	Soft Skills Foundation – Part 3	1	Soft Skills Part 3
12	Sprint 1 Evaluation	2	Sprint 1 Evaluation
13	Online training with Badge 1 & Badge 2 (SME support Query clarification)	10	Sprint 2 Implementation, MCQ
14	Soft Skills Foundation – Part 4	1	Soft Skills Part 4 (Evaluation)
15	Snowflake Lab guide	4	Sprint 2 Implementation, MCQ
16	Snowflake Assignment	5	
17	Sprint 2 Evaluation	1	Sprint 2 Evaluation
18	L1 Test	1	MCQ
Total Training Duration		48	

# **I&D (Advance SQL + Teradata + Snowflake)**

## **Curriculum**

### **Data Warehouse Concepts**

**Program Duration:** 1 day.

#### **Contents:**

##### **Business Intelligence**

- Business Intelligence
- Need for Business Intelligence
- Terms used in BI
- Components of BI

##### **General concept of Data Warehouse**

- Data Warehouse
- History of Data Warehousing
- Need for Data Warehouse
- Data Warehouse Architecture
- Data Mining Works with DWH
- Features of Data warehouse
- Data Mart
- Application Areas

##### **Dimensional modeling**

- Dimension modeling
- Fact and Dimension tables
- Database schema
- Schema Design for Modeling
- Star
- Snow Flake
- Fact Constellation schema

##### **ETL and Metadata**

- ETL process
- Metadata used in ETL
- Metadata in Data Warehousing
- Simple Data warehouse model

##### **Online Analytical Processing (OLAP)**

- Online Analytical Processing (OLAP)
- Nature of OLAP analysis
- Types of OLAP
- OLAP Tools
- OLTP and OLAP

- OLAP Functional requirements
- OLAP Fast and Selective
- Operational versus Informational System
- Data Mining
  - Data mining
  - The Knowledge Discovery process
  - Need of Data Mining
  - Use of Data mining
  - Data mining and Business Intelligence
  - Types of data used in Data mining
  - Data Mining applications
  - Data Mining products
  - Data Mining market
- Best Practices for Building Data Warehouse
  - Recipe for a Successful data warehouse
  - Data warehouse pitfalls
  - Popular BI DW tools and suits
  - Trends in BIDW

## **Data Modeling for Business Intelligence**

**Program Duration:** 1.5 days

**Contents:**

- Introduction to Data Modeling
  - Importance of data modeling
  - Features of a good data model
  - Who should be involved in data modeling
  - Database design stages and deliverables
  - Classification of information
- Understanding Business Requirements
  - Need of Requirement Analysis
  - Characteristics of a Good Requirement
  - The Data Life cycle
  - Methods of Collecting requirement
  - Business Requirement Specification (BRS)
- Conceptual Model
  - Define conceptual model
  - Objectives of conceptual model
  - Components of Conceptual Model
  - Types of Modeling
  - Entity-Relationship (ER) model

- Types of Attributes
- Join Problems
- Steps of dimension modeling
- Star Schema
- Snowflake Schema
- Bill Inmon Vs Ralph Kimball Approach
- Logical Model
  - Define logical model
  - List features of a logical model
  - Transformations required to be done while converting a conceptual model into a Logical model
  - Activities in table specification
  - Activities in column specification
  - Activities in Primary key specification

## **ETL Basics**

**Program Duration:** 0.25 day

**Contents:**

- Basic Concepts
  - Data warehouse
  - Data warehousing strategies
  - Data warehouse architecture
  - ETL Meaning
  - Need for ETL
  - ETL Process
  - Operational Considerations
- ETL Process
  - Data extraction
  - Data transformation
  - Data Loading
- Operational Considerations
  - Exceptional Handling
  - Alerts and Notification
  - Process restart-ability
  - Job Scheduling and Monitoring

**ETL Tools**

- Leading ETL tool vendors
- ETL tool strengths / weaknesses
- Choosing the correct ETL tool

# Software Testing for BI

**Program Duration:** 0.25 day

## **Contents:**

- Introduction to Software testing for BI
  - o Business requirements
  - o BI Project versus BI Program
  - o How is BI testing different from traditional code based testing?
  - o BI SDLC
- Testing concepts
  - o What is testing? Testing – Why? Testing – How?
  - o Principles of Testing
  - o Test Case and Test Suite
  - o Testing scope
  - o Test Strategy
  - o Verification and Validation
- Types of Testing
  - o Static Testing, Dynamic Testing, Automated testing
  - o V Model for BI Testing
- Testing for BI
  - o Testing document purpose (Test documentation)
  - o General BI Testing Principles
  - o BI Testing Mission
  - o Production Verification Testing
  - Possible Areas of Automation

# Python Programming

**Program Duration:** 4 days.

## **Contents:**

Introduction to Python Programming

- Why do we need Python?
- Program structure in Python

Execution steps

- Interactive Shell
- Executable or script files.
- User Interface or IDE

Flow Control

Boolean Operators

- Comparison Operators
- Binary Boolean Operators
- The not Operator
- Data Types and Operations
  - Numbers
  - Strings
  - List
  - Tuple
  - Dictionary
  - Other Core Types
- Changing Values in a List with Indexes
- List Concatenation and List Replication
- Using for Loops with Lists
- Removing Values from Lists with del Statements
- Pattern Matching with Regular Expressions
  - Regular Expression Matching
  - Finding Patterns of Text with Regular Expressions
  - Grouping with Parentheses
  - Matching Multiple Groups with the Pipe
  - Matching Zero or More with the Star
  - Matching Specific Repetitions with Curly Brackets
  - Case-Insensitive Matching
- Statements and Syntax in Python
  - Assignments, Expressions and prints
  - If tests and Syntax Rules
  - While and For Loops
  - Iterations and Comprehensions
- Break/Continue Statements
- Functions in Python
  - Function definition and call
  - Function Scope
  - Return Values and return Statements
  - Local and Global Scope
  - Arguments
  - Function Objects
  - Anonymous Functions
  - Exception Handling
- Modules and Packages-Basic
  - Module Creations and Usage
  - Package Creation and Importing
- Classes in Python
  - Classes and instances
  - Classes method calls
- File Operations
  - Backslash on Windows and Forward Slash on OS X and Linux
  - Absolute vs. Relative Paths
  - Finding File Sizes and Folder Contents

- Open/Read/Write/Append into file
- Using Files
- Copying Files and Folders

#### Libraries

Importing a library using PIP, CONDO etc

Math

Numpy

#### Working with RDBMS

Connection to Database

Cursor Creation

Fire Query & Collect results from Tables/Queries

Insert Data into Tables

Bulk Insert into Tables

Procedure Calls with Arguments & Collect Result Values

#### Debugging

Raising Exceptions

Getting the Traceback as a String

Assertions

Logging Module

Logging to a File

IDLE's Debugger

Breakpoints

#### Working with CSV Files and JSON Data

CSV Module

Reading Data from Reader Objects in a for Loop

Writer Objects

The delimiter and line terminator Keyword Arguments

JSON Module

Reading JSON with the loads() Function

Writing JSON with the dumps() Function

#### Multithreading

Running Other Python Scripts

## Oracle SQL Advance

**Program Duration:** 8 days.

#### **Contents:**

Introduction to Database

Introduction to DBMS

Characteristics of DBMS

DBMS Models

Relational DBMS

Data Integrity

Normalization & Codd's Rules for "FULLY" Functional System



- First Normal Form
- Second Normal Form
- Third Normal Form
- Relational DBMS
- Data Integrity
- Structured Query Language
  - Interacting SQL using SQL \*Plus
  - Using SQL \*Plus
  - What is SQL?
  - Rules for SQL statements
  - Standard SQL Statement Groups
  - Basic DataTypes
  - Rules for naming a Table
  - Specifying Integrity Constraints
  - DDL Statements: Create, Alter, Drop
  - Regular vs Temporary tables
- Data Manipulation Language
  - Inserting Rows Into a Table
  - Deleting Rows from a Table
  - Updating Rows in a Table
- Database Objects
  - Index
  - Synonym
  - Sequence
  - Views
- Data Query Language (Select Statement)
  - Select Statement
  - Distinct Clause
  - Comparison, arithmetic & Logical Operators SQL Operators
  - The ORDER BY Clause
  - Tips and Tricks
- Aggregate Functions, Group By and Having Clause
  - Aggregate Functions
  - The GROUP BY Clause
  - HAVING Clause
  - ROLLUP Operation
  - CUBE Operation
  - Tips and Tricks
- SQL (Single Row) Functions
  - Character Functions
  - Number Functions
  - Data Conversion Function

- Formats for Date functions
- Date Functions
- Miscellaneous Functions
- Tips and Tricks
- Transactions
  - Transaction
  - Commit Command
  - Rollback and Savepoints
- Joins and Subqueries
  - Inner/Equi Join
  - Outer Join
  - Self Join
  - Subquery
  - SUBQUERIES Using Comparison Operators Co-related Subquery
  - Exists / Not Exists Operator
  - Connect By and Start with clauses
  - Tips and Tricks
- Set Operations
  - The UNION Operator
  - The INTERSECT Operator
  - The MINUS Operator
  - The UNION Operator
  - The INTERSECT Operator
  - Tips and Tricks
- Data Control Language
- Introduction to Oracle Architecture
- Introduction to Data Dictionary
- PL/SQL Basics
  - Introduction to PL/SQL
  - PL/SQL Block Structure
  - Handling Variables in PL/SQL
  - SQL in PL/SQL
  - Programmatic Constructs
- Introduction to Cursors
  - Introduction to Cursors
  - Implicit Cursors and Explicit Cursors
  - Cursor with Parameters
  - Usage of Cursor Variables
- Exception Handling and Dynamic SQL
  - Error Handling (Exception Handling)
  - Predefined Exception
  - Numbered Exceptions

- User Defined Exceptions
- OTHERS Exception Handler
- Procedures, Functions, and Packages
  - Subprograms in PL/SQL
  - Anonymous Blocks versus Stored Subprograms
  - Procedures, Functions, Packages
- Database Triggers
- Locks
- Built-in Packages in Oracle
  - DBMS\_OUTPUT
  - UTL\_FILE
  - DBMS\_LOB
- SQL \* Plus Reports
  - SQL \* Plus Reporting
  - SQL \* Plus Commands
- SQL \* Loader
  - What is SQL \* Loader?
  - SQL \* Loader as a Utility
  - SQL \* Loader Environment
  - The Bad File and Discard File
  - Invoking SQL \* Loader
  - SQL \* Loader Examples
- Oracle Tools

## Teradata Basics

**Program Duration:** 5 days

**Contents:**

- An Overview of Teradata
  - RDBMS Concepts
  - Teradata Overview
  - Teradata and Data warehouse
  - Components and Architecture
  - Teradata Training, NA BI, Capgemini India
  - Teradata Utilities
- Teradata Utilities
  - Introduction about Teradata Utility.
  - Introduction to BTEQ.
  - Use of BTEQ
  - Transaction Mode in BTEQ

Conditional Logic in BTEQ  
Teradata Training to BTEQ  
BTEQ Return Codes  
Using BTEQ to Export Data  
Using BTEQ to Import Data  
BTEQ Commands  
TPT ( Teradata Parallel Transporter)

#### OLAP Functionalities

To be familiar with popular OLAP functions.  
To be familiar with the PARTITION By concept.  
To be familiar with RANK() ,ROW\_NUMBER(), QUALIFY functions

#### TD SQL

Aggregation Function  
Basic SQL Function  
Collect Statistics  
Data Manipulation Language (DML)  
Date Functions  
Distinct Vs Group By Functions  
Explain  
Format Functions  
Help and Show  
Join Functions  
Join Indexes  
Math Functions  
OLAP Functions  
Substrings and Positioning Functions  
Temporal Tables Create function  
Temporary Tables  
Teradata Parallel Transport  
The Quantile Function  
Top SQL Command Cheat Sheet  
View Functions  
The Where Clause  
Sample  
Set Operators functions  
Statistical Aggregate Functions  
Stored Procedure Functions  
Sub Query Functions  
Distinct Vs Group By Functions  
Explain  
Format Functions

- Help and Show
- Join Functions
- Join Indexes
- OLAP Functions
- Substrings and Positioning Functions
- Temporal Tables Create function
- Temporary Tables
- The Quantile Function
- Top SQL Command Cheat Sheet
- View Functions

## **Online training with Badge 1 & Badge 2**

**Program Duration:** 10 days

**Contents:**

Essential Data Warehousing (Formerly Web UI Essentials)

- Intro & User Roles
- DB Navigation
- Data Exploration
- Warehouses
- Creating DB Objects
- Transformations
- WH Concepts
- Staging Data
- Data Storage Structures
- Intro to Semi-Structured Data
- Semi-Structured Nested Data
- Module Exam

## Data Sharing, Marketplace & Exchanges Workshop (SMEW)

- Orientation
- Inbound Shares
- Using Shared Data
- Outbound Shares
- Reader Accounts
- Readers vs Full Accounts
- Shop the Marketplace
- List on the Marketplace
- Data Exchanges
- Submit Project

## Snowflake Lab Guide

**Program Duration:** 4 days

**Contents:**

- Prepare Your Lab Environment
- The Snowflake User Interface & Lab “Story”
- Preparing to Load Data
- Loading Data
- Analytical Queries, Results Cache, Cloning
- Working With Semi-Structured Data, Views, JOIN
- Using Time Travel
- Roles Based Access Controls and Account Admin
- Data Sharing

# Snowflake Lab Assignments

**Program Duration:** 5 days

**Contents:**

- Data Loading Assignments
- Data Processing Assignments
- Data Sharing Assignments
- Data Recovery Assignments
- Stored Procedure Assignments
- Security Assignments
- Cloud Integration Assignments