**Index**

[BI (DataStage + Teradata + Qlik view) LoT Course Structure 2](#_Toc512501337)

[Introduction to Python 3](#_Toc512501338)

[Oracle 5](#_Toc512501339)

[UNIX 8](#_Toc512501340)

[Java Programming Basics 9](#_Toc512501341)

[Data Warehouse Concepts 11](#_Toc512501342)

[Data Modeling for Business Intelligence 11](#_Toc512501343)

[ETL Basics 12](#_Toc512501344)

[Software Testing for BI 13](#_Toc512501345)

[Teradata Basics 13](#_Toc512501346)

[Datastage 15](#_Toc512501347)

[Qlik View 16](#_Toc512501348)

[Qlik Sense 18](#_Toc512501349)

[Big Data Overview 19](#_Toc512501350)

[Introduction to Software Engineering 19](#_Toc512501351)

[Quality Process Awareness 20](#_Toc512501352)

[Pseudo Live Project (PLP) 20](#_Toc512501353)

BI (Big Data) LoT-Regular Talent Course Structure

BI (BIG DATA) 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 BI LoT.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Course** | **Duration** | **Including Saturdays** |
|  | **(In Days)** |
|  | Discover (Induction) | 2 | **4 Saturdays-->** Assignments:14 Hrs Mini Project Work:4 Hrs **Mandatory org. Elearnings: 4 Hrs** Soft Skills : 4 Hrs EF SOLO Pretest: 2 Hrs EF Solo Learning: 4 Hrs |
|  | Campus to Corporate | 2 |
| **Module 1** | Introduction to Python | 2 |
| Oracle | 3 |
| UNIX | 2.5 |
| Introduction to Software Engineering | 0.5 |
| Java Programming Basics | 8 |
| Module 1 Test + Soft Skills Part 1 | 1 |
| **Module 2** | Data Warehouse Concepts | 1 | **4 Saturdays-->** Assignments:6 Hrs Mini Project Work:16 Hrs 101 Modules: 2 Hrs **Emerging Technologies: 2 Hrs** EF Solo Learning: 6 Hrs |
| Data Modeling for Business Intelligence | 1 |
| ETL Basics | 0.5 |
| Software Testing for BI | 0.5 |
| Big Data Platform | 7 |
| Soft Skills Part 2 + Module 2 Test | 1 |
| **Module 3** | Spark in memory | 4 | **1 Saturday-->** Mini Project Work: 3 Hrs 101 Modules: 2 Hrs **Emerging Technologies: 2 Hrs** EF Solo Learning: 1 Hr |
| No SQL Databases-HBASE / Cassandra | 2 |
| Big Data Testing | 2 |
| Soft Skills Part 3 + Module 3 Test | 1 |
| **Module 4** | Talend | 5 |
| Cloud Introduction and cloud native big data (SaaS, PaaS, AaaS) models | 2 |
| Module 4 Test | 1 |
| **Module 5** | Mini Project presentation | 1 | **2 Saturdays-->** Assignments:0.5 Hr Mini Project Work:9 Hrs 101 Modules: 2 Hrs **Emerging Technologies: 2 Hrs** EF Solo Learning: 2.5 Hrs |
| **Quality Process Awareness + PLP + PLP Presentation** | 8 |
| L1 Preparation + Test | 2 |
|  | **Total Training Duration** | **60** | **2 Saturdays-->** 101 Modules: 2 Hrs **Emerging Technologies: 2 Hrs** EF SOLO PostTest: 2 Hrs PLP+L1 Prep: 12 Hrs |
|  |  |  | **13** |

**BI (I&D) Curriculum**

**Introduction to Python**

**Program Duration:** 2 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

Data Types and Operations

• Numbers

• Strings

• List

• Tuple

• Dictionary

• Other Core Types

Statements and Syntax in Python

• Assignments, Expressions and prints

• If tests and Syntax Rules

• While and For Loops

• Iterations and Comprehensions

Functions in Python

• Function definition and call

• Function Scope

• Arguments

• Function Objects

• Anonymous Functions

Modules and Packages-Basic

• Module Creations and Usage

• Package Creation and Importing

Classes in Python

• Classes and instances

• Classes method calls

File Operations

• Opening a file

• Using Files

• Other File tools

Libraries

* Importing a library
* Math
* Numpy

**Oracle**

**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

**UNIX**

**Program Duration:** 2.5 days.

**Contents:**

Introduction to UNIX Operating System and Basic UNIX commands

Introduction to UNIX Operating System and Basic UNIX commands

Operating System

Basic UNIX Commands

UNIX File System

UNIX File System

File Types

File Permissions

File Related Commands

Filters

Simple Filters

Advanced Filters

Vi Editor

Vi Editor

Input Mode Commands

Vi Editor – Save & Quit

Cursor Movement Commands

Shell Programming

Shell Variables

Environmental Variables

Shell script Commands

Arithmetic Operations

Command Substitution

Command Line Arguments

Conditional Execution

if Statement Format

Test - String Comparison

The Case Statement

While Statement

Break & Continue Statement

Until Statement

Shell functions

Using arrays

**Java Programming Basics**

**Program Duration:** 3 day.

**Contents:**

§Lesson 1: OOP Concepts

§1.1: Introduction to OOP

§1.2: Characteristics of OOP

§1.3: Classes and Objects

§1.4: Abstraction

§1.5: Encapsulation

§1.6: Inheritance

§1.7: Polymorphism

§Lesson 2: Getting Started

§2.1: Introduction to Java

§2.2: Writing, compiling and Running a program

§2.3: Platform Independency in Java

§2.4: Integrated Development Environment

§2.5: Some Important Terms in Java

§2.6: JVM Basic Architecture

§Lesson 3: Basic Language Constructs

§3.1: Naming Conventions in Java

§3.2: Variable and Data Types

§3.3: Operators (arithmetic, assignment, relational, logical and bitwise)

§3.4: Promotion and Demotion Rules for Operators

§3.5: Looping (while, do…while, for loop)

§3.6: Conditional Statements (if else, switch)

§3.7: break and continue statements

§3.8: Reference Variables

§3.9 : Arrays

§Lesson 4: Classes and Objects

§4.1: Classes and Objects

§4.2: Access Controls

§4.3: Constructor

§4.4: Overloading

§4.5: Static methods and fields

§4.6: Garbage Collection –finalize() method

§4.7: Extended Parameters for JVM –xmx,-xms

§4.8: Memory Leakage, Overflow and Out of Memory

§Lesson 5: Extending Classes

§5.1: Inheritance

§5.2: Protected Keyword

§5.3: Constructors in Extended Classes

§5.4: Overriding Methods

§5.5: Polymorphism

§5.6: Making Methods and Classes Final

§Lesson 6: Abstract Classes

§6.1: Abstract classes and methods

§6.2: Extending abstract class

§6.3: Abstract class and Polymorphism

§Lesson 7: Interfaces and Loose Coupling

§7.1: Declaring interfaces

§7.2: Implementing interfaces

§7.3: Extending interfaces

§7.4: Loose coupling using interfaces

§Lesson 8: Packages and Static Imports

§8.1: Creating packages

§8.2: Naming packages

§8.3: Package Access

§8.4: Packages and class path

§8.4: Importing packages

§8.5: Static imports

§Lesson 9: Exception Handling

§9.1: Checked exceptions

§9.2: Unchecked exceptions

§9.3: The “try-throw-catch” structure

§9.4: The “finally” clause

§9.5: Custom Exception

§9.6: Exception chaining

§9.7: New Features of Java 7

•Try with Resources

• AutoCloseable

• Handling Suppresed Exceptions

• Catch Block Handling Multiple Exceptions

§Lesson 10: The Input and Output Classes

§Types of Input and Output Streams

§Byte-based stream

§Character-based stream

§Reader and Writers

§PrintWriter

§NIO Overview

**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 day

**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:** 0.5 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:** 0.5 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

o   Possible Areas of Automation

**Teradata Basics**

**Program Duration:** 4 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

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

**Datastage**

**Program Duration:** 8 days

**Contents:**

Introduction to IBM IIS and Datastage

Introduction to IBM IIS

Setting Up Your DataStage Environment

Architecture of IBM IIS

Creating Parallel Jobs

Assessing Sequential Data

Datastage runtime Architecture and various stages.

Platform Architecture

Combining Data

Sorting and Aggregating

Standard data transformation Techniques

Transforming Data

Standards and Techniques

Accessing Relational Data

Metadata in the Parallel Framework

Explain and Create schemas

Runtime Column Propogation (RCP)

Shared Container

Special Topic: Job Control and Case Study

Change Capture Stage

Job Control

Error Handling

Restart

Advance Topics

What’s New in IBM InfoSphere DataStage latest version

Execution of DS jobs through UNIX

"Advance stages

1. MQ,

2. XML

3. Real time stages "

Advance Topics and Case Study

Case Study

Solution Development Case Study.

**Qlik View**

**Program Duration:** 4 days

**Contents:**

|  |  |
| --- | --- |
| Introduction To Business Intelligence & Visualization | |
| What is Business Intelligence? | |
|  | What is data visualization? |
|  | Need For Visualization |
|  | Uses of visualization |
|  |  |
| Introduction to Qlik (Qlikview) | |
| What is Qlikview | |
|  | Qlik vs Traditional BI Tools |
|  | Installing Qlik Desktop |
|  | Qlikview Installation and Licensing Overview |
|  | Benefits of Qlik |
|  | Qlik Project Architecture |
|  |  |
| Dive into Qlik | |
| Explore Qlik Interface | |
|  | Understand various Qlik Terminologies |
|  | Case Study |
|  |  |
|  |  |
| Fueling More Data - Connecting Data Sources | |
| Connection Options | |
|  | Different Types of Connectors |
|  | Data Loading |
|  | Identifying Dimension and Measures |
|  |  |
|  |  |
| Data Modeling in Qlik | |
| Basic data modeling concepts | |
|  | Structuring the script |
|  | Loading Data and Data transformations |
|  | Resolving data issues |
|  | Generating data with the script |
|  | Debugging a data load |
|  | Scripting and data model challenges |
|  | Different Types of Load |
|  | Advanced Calculations and Functions |
|  | Performance Consideration |
|  |  |
|  |  |
| Visualizations in Qlik | |
| Getting started with Qlik App | |
|  | Visualizations and Chart Types |
|  | Enhancing Charts and Properties customization |
|  | Sharing your insights |
|  |  |
| Advanced Development and Design Fundamentals | |
| QlikView data storage (Working with QVD) | |
|  | Incremental loads |
|  | Set analysis and its components |
|  | Additional functions |
|  | Calculated Fields and Dimensions |
|  | Direct Discovery |
|  | Security and Section Access |
|  |  |
|  |  |
| Qlik Deployment Overview\* | |
| Server Environment Overview | |
|  | Server Components |
|  | Deploying Apps |
|  | Accessing Apps |
|  | Make Estimates using Parameters |
|  |  |
| Dashboards Best Practices | |
| Creating an Interactive Dashboard | |
|  | Adding Actions to a Dashboard |
|  | Best Practices for Dashboard Design |
|  | Case Study |

**Qlik Sense**

**Program Duration:** 2 days

**Contents:**

|  |
| --- |
| **Introduction to QlikSense** |
| Why Analytics? |
| Understanding Signal and Noise |
| Understanding Data |
| Creating a Simple App |
|          Step by Step Process to Create an App |
|          Navigation within Qlik Sense |
|          Data Load Basics |
|          Introduction to the Data Model Viewer |
|          Introduction to the Data Manager |
|          Introduction to the Data Load Editor |
|          Reload Data for Analysis |
| Considerations when Loading Data from Excel |
| Overview of different Types of Data Connectors |
| Using the Qlik DataMarket Package |
| **Data Load Fundamentals** |
| Introduction to QVDs |
|          Understanding QVD’s |
|          What is QVD file? |
|          Advantages of using QVDs |
|          How to create a QVD |
|          How to load data from QVDs |
| Data Load Challenges |
| Working with Variables |
| Debugging in Qlik Sense |
| Understanding Dimensions and Measures |
|  |
|  |
|  |
| **Qlik Sense Visualizations and Chart Types** |
| Overview of all the Visualizations in Qlik Sense and Demo |
| Editing and setting up Chart Properties |
|          Sorting Items in Visualizations |
|          Limit Dimensions Displayed |
|          Chart Expressions Overview |
|          Control Visualization Calculation |
|          Color in Visualizations |
| **Advanced Fundamentals of Qlik Sense** |
| Incremental loads |
|          Types of Incremental loads |
|          How to do incremental loads |
| Set analysis and set expressions fundamentals |
| Understanding Extensions |
| Introduction to Developer Hub and its Features |
| **Qlik Sense Administration and Story Telling** |
| Build and Play Stories |
|  |

**Big Data Overview**

**Program Duration:** 2 days

**Contents:**

Big Data overview

Hadoop insight

Pig Latin

Hive query language

MapReduce

Introduction to Real-time Processing and brief on Kafka

Introduction to In-memory Processing

Introduction to Big data Analytics

One Big Data use case briefing

**Introduction to Software Engineering**

**Program Duration:** 0.5 day

**Contents:**

* To Understand the following :
  + What is Software Engineering (SE)
  + Common life cycle models
  + Phases in SE
  + Familiarizing Requirements Phase
  + Familiarizing Design Phase
  + Familiarizing Construction Phase
  + Familiarizing Testing and acceptance Phase
  + Review and Configuration Management Process

**Quality Process Awareness**

**Program Duration:** 0.5 day

**Contents:**

* Understand the following :
  + Quality – What and Why
  + Introduction to Quality Management System
  + QMS support to Software Methodology
  + Metrics
  + Defect Prevention

**Pseudo Live Project (PLP)**

**Program Duration:** 7.5 days

**Contents:**

* Pseudo Live Project (PLP) program is primarily to handhold participants who are fresh into the IT stream & newly recruited from college.PLP project is executed to orient the trainees towards Quality processes followed in the organization. Participants have to understand the value & usage of the various forms, templates & review mechanisms. In PLP, more importance given to “Process Adherence”
* The following SDLC activities are carried out during PLP
  + Requirement Analysis
  + Design ( High Level Design and Low Level Design)
  + Design of UTP(Unit Test Plan) with test cases
  + Coding
  + Code Review
  + Testing
  + Deployment
  + Configuration Management
  + Final Presentation