### **Itinerary for 2-Day Training for Freshers - Python**

### **Day 1: Python Fundamentals (5 hours)**

#### **Session 1: Introduction to Python (2.5 hours)**

**09:00 AM - 09:15 AM**: **Welcome and Overview**

* Introduction to the training session
* Importance and applications of Python in various domains

**09:15 AM - 10:00 AM**: **Python Basics**

* Python installation and setting up the development environment
* Python syntax and semantics
* Basic data types (integers, floats, strings, and booleans)

**10:00 AM - 10:15 AM**: **Break**

**10:15 AM - 11:00 AM**: **Control Structures**

* Conditional statements (if, elif, else)
* Loops (for and while loops)
* Practical examples and exercises

**11:00 AM - 11:30 PM**: **Hands-On Practice**

* Coding exercises covering the topics discussed
* Group activity: Write a simple script using control structures

**11:30 PM - 12:00 PM**: **Lunch Break**

#### **Session 2: Working with Data in Python (2.5 hours)**

**12:00 PM - 12:45 PM**: **Data Structures**

* Lists, tuples, sets, and dictionaries
* Operations and methods on data structures

**12:45 PM - 01:30 PM**: **Functions and Modules**

* Defining and using functions
* Scope and lifetime of variables
* Introduction to modules and the standard library

**01:30 PM - 01:45 PM**: **Break**

**01:45 PM - 02:20 PM**: **File Handling**

* Reading from and writing to files
* File handling techniques and best practices

**02:20 PM - 02:30 PM**: **Recap and Q&A**

* Summary of Day 1 topics
* Q&A session to clarify doubts and review concepts

### **Day 2: Advanced Python and Applications (5 hours)**

#### **Session 1: Python for Data Handling (2.5 hours)**

**09:00 AM - 09:15 AM**: **Recap and Introduction**

* Quick review of Day 1
* Introduction to Day 2 topics and goals

**09:15 AM - 10:00 AM**: **Introduction to Libraries for Data Analysis**

* Overview of NumPy and pandas
* Working with arrays and data frames

**10:00 AM - 10:15 AM**: **Break**

**10:15 AM - 11:00 AM**: **Data Manipulation with pandas**

* Loading data into pandas dataframes
* Data cleaning and preprocessing
* Practical examples: Filtering, grouping, and merging data

**11:00 AM - 11:30 PM**: **Hands-On Practice**

* Exercises on data manipulation using NumPy and pandas
* Group activity: Perform basic analysis on a sample dataset

**11:30 PM - 12:00 PM**: **Lunch Break**

#### **Session 2: Advanced Concepts and Final Project (2.5 hours)**

**12:00 PM - 12:45 PM**: **Error Handling and Debugging**

* Understanding exceptions
* Using try, except, and finally blocks
* Best practices for debugging Python code

**12:45 PM - 01:30 PM**: **Introduction to Object-Oriented Programming (OOP)**

* Basic concepts: Classes, objects, and methods
* Creating and using classes
* Practical examples: Implementing simple OOP concepts

**01:30 PM - 01:45 PM**: **Break**

**01:45 PM - 02:30 PM**: **Hands-On Practice**

* Practical exercises on the topics covered
* Q&A and troubleshooting common issues

### **Itinerary for 2-Day Training for Freshers - Databricks**

### **Day 1: Introduction to Databricks and Basic Operations**

**Total Duration: 5 hours**

#### **Session 1: Getting Started with Databricks (1.5 hours)**

**09:00 AM - 09:15 AM**: **Welcome and Overview**

* Introduction to the training agenda and objectives
* Brief overview of Databricks and its importance in data engineering

**09:15 AM - 09:45 AM**: **Databricks Basics**

* What is Databricks? Overview of its architecture and use cases
* Understanding the Databricks workspace and its components

**09:45 AM - 10:15 AM**: **Setting Up Databricks Environment**

* How to create a Databricks account and workspace setup
* Introduction to clusters: Creating, configuring, and managing clusters

**10:15 AM - 10:30 AM**: **Break**

#### **Session 2: Working with Databricks Notebooks (1.5 hours)**

**10:30 AM - 11:00 AM**: **Introduction to Databricks Notebooks**

* Creating and organizing notebooks
* Basics of using notebooks: Cells, markdown, and code execution

**11:00 AM - 11:30 AM**: **Basic Data Operations in Databricks**

* Loading data into Databricks: Different methods and sources
* Simple data manipulations: Filtering, sorting, and aggregating data

**11:30 AM - 12:00 PM**: **Hands-On Practice**

* Guided exercise: Load a sample dataset and perform basic operations
* Q&A and troubleshooting common issues

**12:00 PM - 12:30 PM**: **Lunch Break**

#### **Session 3: Introduction to Apache Spark in Databricks (1 hour)**

**12:30 PM - 12:50 PM**: **Understanding Apache Spark**

* Overview of Apache Spark and its integration with Databricks
* Key concepts: RDDs, DataFrames, and Spark SQL

**12:50 PM - 01:10 PM**: **Data Processing with Spark in Databricks**

* Using Spark DataFrames for data manipulation
* Practical example: Loading and processing a dataset with Spark

**01:10 PM - 01:30 PM**: **Exploratory Data Analysis (EDA) with Spark**

* Techniques for data exploration and analysis using Spark
* Practical example: Perform EDA on a dataset in Databricks

**01:30 PM - 01:40 PM**: **Break**

#### **Session 4: Data Visualization in Databricks (1 hour)**

**01:40 PM - 02:00 PM**: **Introduction to Data Visualization**

* Overview of data visualization capabilities in Databricks
* Creating simple visualizations directly within Databricks notebooks

**02:00 PM - 02:20 PM**: **Hands-On Practice and Q&A**

* Exercise: Create visualizations for the datasets used in previous sessions
* Open floor for questions and discussion

**02:20 PM - 02:40 PM**: **Recap and Wrap-Up**

* Summary of Day 1 learnings
* Brief introduction to the topics for Day 2

### **Day 2: Advanced Databricks and Data Engineering**

**Total Duration: 5 hours**

#### **Session 1: Advanced Data Operations in Databricks (3 hours)**

**09:00 AM - 09:10 AM**: **Recap and Introduction to Day 2**

* Quick review of Day 1 highlights and overview of Day 2 agenda

**09:10 AM - 10:00 AM**: **Advanced Spark DataFrame Operations**

* Joins, aggregations, and window functions in Spark DataFrames
* Practical example: Advanced data transformation tasks

**10:00 AM - 10:30 AM**: **Optimizing Spark Jobs**

* Understanding and optimizing Spark job performance
* Techniques for improving efficiency: Caching, partitioning, and tuning

**10:30 AM - 10:40 AM**: **Break**

**10:40 AM - 11:10 AM**: **Working with Structured Streaming**

* Introduction to streaming data processing with Spark
* Setting up and managing streaming jobs in Databricks

**11:10 AM - 12:00 PM**: **Hands-On Practice**

* Guided exercise: Implementing advanced data operations and streaming
* Q&A and troubleshooting common issues

**12:00 PM - 12:30 PM**: **Lunch Break**

#### **Session 2: Integration and Advanced Features (2 hours)**

**12:30 PM - 01:00 PM**: **Integrating Databricks with External Systems**

* Connecting Databricks to external data sources (cloud storage, databases)
* Using Databricks for ETL processes

**01:00 PM - 01:30 PM**: **Databricks SQL and Delta Lake**

* Introduction to Databricks SQL for data querying and analytics
* Overview of Delta Lake for data reliability and performance

**01:30 PM - 02:00 PM**: **Building and Managing Workflows**

* Using Databricks Jobs to automate and schedule workflows
* Practical example: Setting up a data pipeline in Databricks

**02:00 PM - 02:10 PM**: **Break**

#### **Session 3: Security, Collaboration, and Final Project (1 hour)**

**02:10 PM - 02:30 PM**: **Security and Collaboration in Databricks**

* Managing user access and permissions
* Collaborative features in Databricks for team projects

### **Itinerary for 1-Day Training for Freshers - Apache Spark**

### **Day 1: Introduction to Apache Spark**

**Total Duration: 5 hours**

#### **Session 1: Introduction to Apache Spark (2 hours)**

**09:00 AM - 09:15 AM**: **Welcome and Overview**

* Introduction to the training objectives and schedule
* Overview of Apache Spark and its role in big data processing

**09:15 AM - 09:45 AM**: **Spark Fundamentals**

* What is Apache Spark? Key features and architecture
* Understanding the components: Spark Core, Spark SQL, Spark Streaming, MLlib, and GraphX
* How Spark handles big data: RDDs, DataFrames, and Datasets

**09:45 AM - 10:15 AM**: **Setting Up Spark Environment**

* Setting up a local Spark environment (using PySpark or Databricks)
* Introduction to Spark interfaces: Spark Shell, PySpark, and notebooks

**10:15 AM - 10:25 AM**: **Break**

**10:25 AM - 11:00 AM**: **Basic Spark Operations**

* Loading and inspecting data in Spark
* Basic transformations and actions in Spark RDDs and DataFrames
* Practical example: Load a sample dataset and perform basic operations

#### **Session 2: Spark DataFrames and Spark SQL (1.5 hours)**

**11:00 AM - 11:45 AM**: **Introduction to Spark DataFrames**

* What are DataFrames? Benefits over RDDs
* Creating and manipulating DataFrames in Spark
* Practical example: Load a dataset and perform transformations using DataFrames

**11:45 AM - 12:15 PM**: **Lunch Break**

**12:15 PM - 01:00 PM**: **Spark SQL for Data Analysis**

* Overview of Spark SQL and its use cases
* Running SQL queries on Spark DataFrames
* Practical example: Analyze data using Spark SQL queries

**01:00 PM - 01:30 PM**: **Hands-On Practice**

* Exercises: Implementing data transformations and running SQL queries on datasets
* Q&A and troubleshooting common issues

#### **Session 3: Advanced Features and Practical Applications (1.5 hours)**

**01:30 PM - 02:00 PM**: **Performance Optimization in Spark**

* Techniques for optimizing Spark jobs: Caching, partitioning, and tuning
* Understanding Spark's execution plan and performance metrics
* Practical example: Optimize a Spark job for better performance

**02:00 PM - 02:10 PM**: **Break**

**02:10 PM - 02:30 PM**: **Introduction to Spark Streaming**

* Overview of Spark Streaming and real-time data processing
* Setting up a basic Spark Streaming job
* Practical example: Stream data from a source and perform real-time processing

**02:30 PM - 02:50 PM**: **Wrap-Up**

* Summary of key learnings and open floor for questions

**02:50 PM - 03:00 PM**: **Recap and Q&A**

* Recap of the day's topics
* Final Q&A and feedback session

### **Itinerary for 2-Day Training for Freshers - Power BI**

### **Day 1: Introduction to Power BI and Basic Data Analysis**

**Total Duration: 5 hours**

#### **Session 1: Getting Started with Power BI (1.5 hours)**

**09:00 AM - 09:15 AM**: **Welcome and Overview**

* Introduction to the training objectives and agenda
* Overview of Power BI: What it is and its role in data analysis and business intelligence

**09:15 AM - 09:45 AM**: **Power BI Desktop Basics**

* Installing and setting up Power BI Desktop
* Overview of the Power BI interface: Canvas, fields, and visualizations panes
* Basic navigation and functionalities

**09:45 AM - 10:15 AM**: **Connecting to Data Sources**

* Connecting to various data sources (Excel, databases, online services)
* Overview of supported data types and formats
* Practical example: Connect to a sample dataset

**10:15 AM - 10:30 AM**: **Break**

#### **Session 2: Data Transformation and Modeling (1.5 hours)**

**10:30 AM - 11:10 AM**: **Introduction to Power Query**

* Overview of Power Query and its role in data preparation
* Basic data transformation operations (filtering, merging, and appending data)
* Practical example: Clean and transform data using Power Query

**11:10 AM - 11:40 AM**: **Data Modeling in Power BI**

* Creating relationships between tables
* Introduction to calculated columns and measures
* Practical example: Build a simple data model and create basic measures

**11:40 AM - 12:00 PM**: **Hands-On Practice**

* Guided exercise: Transform and model a dataset using Power Query and data modeling tools
* Q&A and troubleshooting common issues

**12:00 PM - 12:30 PM**: **Lunch Break**

#### **Session 3: Creating Basic Visualizations (1.5 hours)**

**12:30 PM - 01:00 PM**: **Building Basic Visualizations**

* Introduction to Power BI visualizations: Charts, tables, and maps
* How to create and customize visualizations
* Practical example: Create simple visualizations from the modeled data

**01:00 PM - 01:30 PM**: **Enhancing Visualizations**

* Using filters, slicers, and drill-through to enhance reports
* Mobile layout
* Practical example: Add interactivity to visualizations

**01:30 PM - 02:00 PM**: **Visualizations Best Practices**

* Tips for effective data visualization
* Common pitfalls and how to avoid them

**02:00 PM - 02:10 PM**: **Break**

### **Day 2: Advanced Power BI Features and Publishing Reports**

**Total Duration: 5 hours**

#### **Session 1: Advanced Data Modeling and DAX (1.5 hours)**

**09:00 AM - 09:10 AM**: **Recap and Introduction to Day 2**

* Brief review of Day 1 highlights
* Overview of Day 2 topics and goals

**09:40 AM - 10:10 AM**: **Introduction to DAX (Data Analysis Expressions)**

* Basics of DAX and its syntax
* Creating calculated columns and measures with DAX
* Practical example: Implement common DAX functions and formulas

**10:10 AM - 10:30 AM**: **Advanced DAX for Data Analysis**

* Using DAX for complex calculations and business logic

Practical example: Apply advanced DAX functions to solve real-world problems

**09:10 AM - 09:40 AM**: **Advanced Report Creation**

* Field Parameters in Power Bi
* Co-pilot for report creation

**10:30 AM - 10:40 AM**: **Break**

#### **Session 2: Advanced Visualizations and Report Design (1.5 hours)**

**10:40 AM - 11:10 AM**: **Creating Advanced Visualizations**

* Introduction to custom visuals and where to find them
* Using advanced visual types for more insightful reporting
* Practical example: Implement custom visuals in reports

**11:10 AM - 11:40 AM**: **Licenses** **Designing Effective Reports**

* Different Power BI Licenses

**11:40 AM - 12:00 PM**: **Hands-On Practice**

* Guided exercise: Enhance and redesign an existing report using advanced visualizations and design techniques
* Q&A and discussion

**12:00 PM - 12:30 PM**: **Lunch Break**

#### **Session 3: Publishing and Sharing Reports (1.5 hours)**

**12:30 PM - 01:00 PM**: **Publishing to Power BI Service**

* Overview of Power BI Service and its features
* Steps to publish reports from Power BI Desktop to Power BI Service
* Practical example: Publish a report to Power BI Service

**01:00 PM - 01:30 PM**: **Sharing and Collaborating on Reports**

* Managing workspaces and sharing reports with others
* Collaboration features: Comments, subscriptions, and report sharing options
* Practical example: Share and collaborate on a report in Power BI Service

**01:30 PM - 02:00 PM**: **Data Refresh and Security**

* Configuring data refresh schedules
* Overview of data security and managing user access
* Practical example: Set up data refresh and manage report access

**02:00 PM - 02:10 PM**: **Break**

#### **Session 4: Wrap-up (1 hour)**

**03:00 PM - 03:30 PM**: **Wrap-Up and Q&A**

* Recap of the two-day training
* Final Q&A and discussion on next steps and resources for further learning