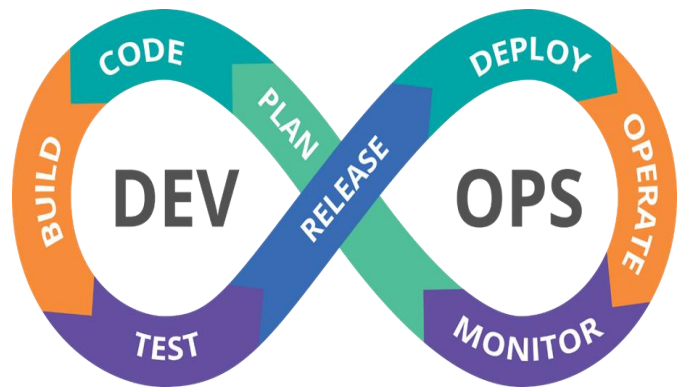




**SAL CONNECTS**  
MODERN TAKSHASILA FOR TECHIES

## OFFERING COURSES



**Red Hat**  
Enterprise Linux



+91 9390190337 | +91 7416088742



salconnectss@gmail.com



**SAL CONNECTS**  
MODERN TAKSHASILA FOR TECHIES

# Data Analytics



## Program Details

Name of the Course: DATA ANALYTICS

Duration: 3 months

If any one is interested



+91 9390190337 | +91 7416088742



salconnectss@gmail.com

# SQL Learning Roadmap

## 1. Introduction to SQL

- Understand what SQL is and its importance in data management.
- Learn about relational databases and SQL's role in querying data.
- Install a database (e.g., MySQL, PostgreSQL, SQLite).
- Set up a simple database and connect to it.

## 2. Basic SQL Commands

- Learn DDL commands: CREATE, ALTER, DROP.
- Understand DML commands: SELECT, INSERT, UPDATE, DELETE.
- Use DCL commands: GRANT, REVOKE.

## 3. SELECT Statement Basics

- Writing simple SELECT queries to retrieve data.
- Filtering data with WHERE clause.
- Sorting data with ORDER BY.
- Using DISTINCT to remove duplicates.

## 4. Working with Functions

- Using aggregate functions: COUNT, SUM, AVG, MIN, MAX.
- Using string functions: CONCAT, LENGTH, SUBSTRING.
- Date and time functions: NOW, CURDATE, DATEDIFF.

## 5. Joining Tables

- Understanding primary and foreign keys.

- Performing INNER JOIN, LEFT JOIN, RIGHT JOIN, FULL JOIN.
- Working with self-joins and cross joins.

## **6. Subqueries and Nested Queries**

- Writing subqueries in SELECT, FROM, and WHERE clauses.
- Using correlated subqueries.
- Understanding subquery performance implications.

## **7. Advanced SQL Concepts**

- Using CASE statements for conditional logic.
- Understanding and implementing indexes for performance.
- Writing recursive queries with CTEs (Common Table Expressions).

## **8. Data Manipulation and Transactions**

- Inserting, updating, and deleting data.
- Using COMMIT and ROLLBACK for transactions.
- Understanding ACID properties of transactions.

## **9. Database Design and Normalization**

- Understanding normalization (1NF, 2NF, 3NF).
- Designing efficient relational database schemas.
- Understanding denormalization and when to use it.

## **10. SQL Performance Optimization**

- Identifying and optimizing slow queries.
- Using EXPLAIN and query execution plans.
- Best practices for indexing.

## **11. Working with Views**

- Creating and managing views.
- Using views for simplified data access.
- Understanding materialized views.

## **12. Stored Procedures and Triggers**

- Creating and executing stored procedures.
- Understanding triggers and their use cases.
- Best practices for writing procedures and triggers.

## **13. SQL for Data Analysis**

- Using SQL to perform exploratory data analysis.
- Working with window functions (ROW\_NUMBER, RANK, NTILE).
- Grouping and aggregating data with GROUP BY and HAVING.

## **14. Working with Big Data and NoSQL**

- Introduction to SQL on big data (e.g., Hive, Spark SQL).
- Understanding differences between SQL and NoSQL.
- Using SQL-like queries in NoSQL databases.

## **15. Mini-Projects**

- Building a library database to manage books and users.
- Creating a sales and inventory tracking system.
- Developing an employee database with role-based access.



# Excel Learning Roadmap

## 1. Fundamentals of Excel

- Introduction to Excel and its interface.
- Understanding rows, columns, cells, and ranges.
- Performing basic operations (data entry, copy, paste, undo, redo).
- Saving and managing Excel files.

## 2. Basic Formatting

- Formatting cells: font, color, borders, and alignment.
- Using number formats (currency, date, percentage).
- Adjusting column width and row height.
- Applying conditional formatting.

## 3. Basic Formulas and Functions

- Writing basic formulas (addition, subtraction, multiplication, division).
- Understanding relative, absolute, and mixed cell references.
- Common functions: SUM, AVERAGE, MIN, MAX, COUNT.

## 4. Data Management

- Sorting and filtering data.
- Using find and replace.
- Removing duplicates.
- Freezing panes and splitting windows.

## 5. Charts and Graphs

- Creating basic charts (line, bar, pie, column).
- Customizing chart elements (titles, legends, colors).
- Working with trendlines and data labels.

## **6. Intermediate Formulas and Functions**

- Logical functions: IF, AND, OR.
- Lookup functions: VLOOKUP, HLOOKUP, INDEX, MATCH.
- Text functions: LEFT, RIGHT, MID, CONCATENATE, LEN.
- Date functions: TODAY, NOW, DATEDIF.

## **7. Working with Tables**

- Creating and formatting Excel tables.
- Structured references in tables.
- Sorting, filtering, and summarizing table data.

## **8. Data Analysis**

- Using PivotTables and PivotCharts.
- Grouping and ungrouping data.
- Performing data summarization.
- Creating slicers for interactive analysis.

## **9. Advanced Excel Techniques**

- Using array formulas.
- Advanced conditional formatting with formulas.
- Data validation and drop-down lists.
- Using macros for automation.

## **10. Power Query and Power Pivot**

- Importing and transforming data with Power Query.
- Creating data models with Power Pivot.
- Understanding relationships between tables.
- Writing DAX formulas for advanced calculations.

## **11. Excel for Business**

- Financial modeling basics.
- Budget planning templates.
- Inventory and sales tracking.
- Creating dashboards for visual reporting.

## **12. Mini-Projects**

- Sales and revenue analysis.
- Employee attendance tracker.
- Dynamic project timeline.
- Personal expense tracker.

## **13. Tips and Tricks**

- Keyboard shortcuts for efficiency.
- Using the Quick Access Toolbar.
- Troubleshooting common errors.
- Protecting and sharing workbooks.



# PySpark Learning Roadmap

## 1. Fundamentals of PySpark

- Introduction to Apache Spark and PySpark
- Understand Spark's architecture (driver, executors, cluster manager).
- Install PySpark locally or set up a notebook (e.g., Jupyter).
- Write a basic PySpark script.

## 2. SparkSession and SparkContext

- Learn the difference between SparkSession and SparkContext.
- Create and configure a SparkSession.

## 3. RDDs (Resilient Distributed Datasets)

- Understand RDD operations: transformations (e.g., map, filter) and actions (e.g., collect, count).

## 4. RDD Operations: Advanced

- Learn flatMap, reduceByKey, groupByKey.
- Practice basic word count programs.

## 5. DataFrames Basics

- Understand DataFrames and their advantages over RDDs.
- Create DataFrames from structured data (CSV, JSON).

## 6. DataFrame Operations

- Explore actions (show, count) and transformations (select, filter, groupBy).

## 7. DataFrame APIs

- Dive into sorting, joins, and aggregations.
- Practice querying data with PySpark DataFrames.

## **8. PySpark SQL**

- Learn to run SQL queries on DataFrames.
- Practice converting DataFrames to temporary views for SQL queries.

## **9. Schema and Data Types**

- Define and enforce schemas in DataFrames.
- Understand common PySpark data types.

## **10. Data Sources**

- Read and write data to/from various sources (e.g., Parquet, ORC, HDFS, S3).

## **11. PySpark UDFs (User-Defined Functions)**

- Write and use UDFs to extend PySpark functionality.

## **12. PySpark MLlib: Basics**

- Understand the concept of Spark MLlib for machine learning.
- Perform basic feature engineering.

## **13. PySpark MLlib: Classification**

- Train and evaluate a simple classifier (e.g., logistic regression).

## **14. PySpark MLlib: Clustering**

- Implement clustering algorithms (e.g., k-means).

## **15. PySpark Streaming Basics**

- Understand Spark Streaming and structured streaming.
- Implement a basic streaming job.

## **16. Structured Streaming Advanced**

- Process real-time data using watermarks and window operations.

## **17. Optimizations: Spark Configurations**

- Explore Spark configurations for performance optimization (e.g., partitions, caching).

## **18. Performance Tuning**

- Learn about job monitoring, DAG visualization, and performance tuning.

## **19. Broadcast Variables and Accumulators**

- Understand and use broadcast variables and accumulators.

## **20. DataFrame Optimization**

- Use Catalyst Optimizer and Tungsten execution for better performance.

## **21. PySpark with Databricks**

- Learn about Databricks as a cloud-based Spark platform.

## **22. Real-World Dataset Exploration**

- Perform ETL operations on large datasets.

## **23. Mini-Project 1: Batch Processing**

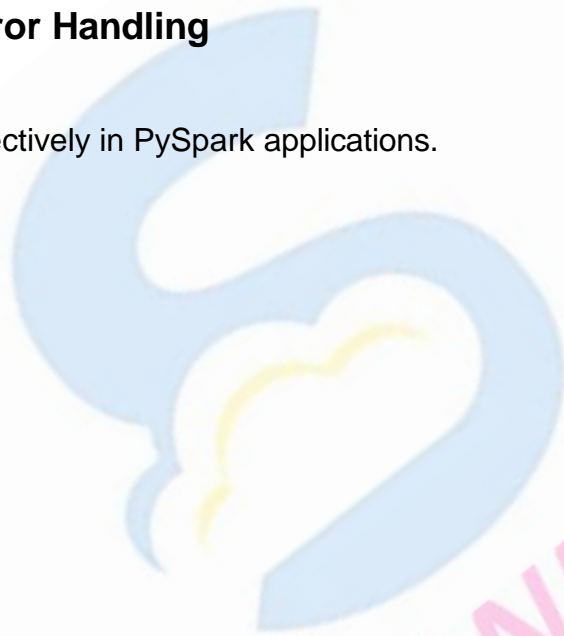
- Process and analyze a dataset (e.g., analyzing sales data).

## **24. Mini-Project 2: Streaming**

- Build a real-time data pipeline using structured streaming.

## **25. Debugging and Error Handling**

- Learn to handle errors effectively in PySpark applications.



**SAL CONNECTS**  
MODERN TAKSHASILA FOR TECHIES

# Power BI Learning Roadmap

## 1. Introduction to Power BI

- Understand what Power BI is and its key features.
- Explore the Power BI ecosystem (Power BI Desktop, Service, and Mobile).
- Install Power BI Desktop and get familiar with its interface.

## 2. Connecting to Data Sources

- Learn how to connect to various data sources (Excel, SQL, web, etc.).
- Understand data import vs. direct query.
- Managing data connections and credentials.

## 3. Data Transformation with Power Query

- Understanding Power Query Editor.
- Performing data cleaning (removing duplicates, handling missing values).
- Transforming data: splitting, merging, and pivoting columns.

## 4. Data Modeling

- Understanding relationships between tables.
- Creating and managing relationships.
- Using calculated columns and measures.

## 5. DAX (Data Analysis Expressions)

- Introduction to DAX and its use in Power BI.
- Writing basic DAX formulas for calculated columns and measures.
- Using common DAX functions: SUM, COUNT, AVERAGE, CALCULATE.

## **6. Visualizations Basics**

- Creating basic visuals: bar chart, line chart, pie chart, and table.
- Customizing visual properties (titles, colors, legends).
- Adding filters and slicers to visuals.

## **7. Advanced Visualizations**

- Using maps for geographic data.
- Creating hierarchical visuals (drill-downs).
- Using KPI cards and gauges for performance metrics.

## **8. Reports and Dashboards**

- Designing interactive reports.
- Using bookmarks and navigation buttons.
- Publishing reports to Power BI Service and creating dashboards.

## **9. Power BI Service**

- Uploading reports to Power BI Service.
- Sharing dashboards with others.
- Understanding roles and access control.

## **10. Data Refresh and Gateways**

- Scheduling data refresh for imported data.
- Setting up and managing Power BI gateways.
- Troubleshooting refresh issues.

## **11. Power BI Advanced Features**



- Using R and Python visuals in Power BI.
- Working with Power BI apps and templates.
- Embedding Power BI reports into other applications.

## **12. Power BI and AI**

- Using AI visuals like Q&A and Key Influencers.
- Integrating machine learning models.
- Performing predictive analysis in Power BI.

## **13. Mini-Projects**

- Sales performance dashboard.
- Customer segmentation and analysis.
- Real-time data monitoring dashboard.

## **14. Tips and Best Practices**

- Optimizing Power BI performance (data model and visuals).
- Using themes and templates for consistency.
- Staying updated with new Power BI features.