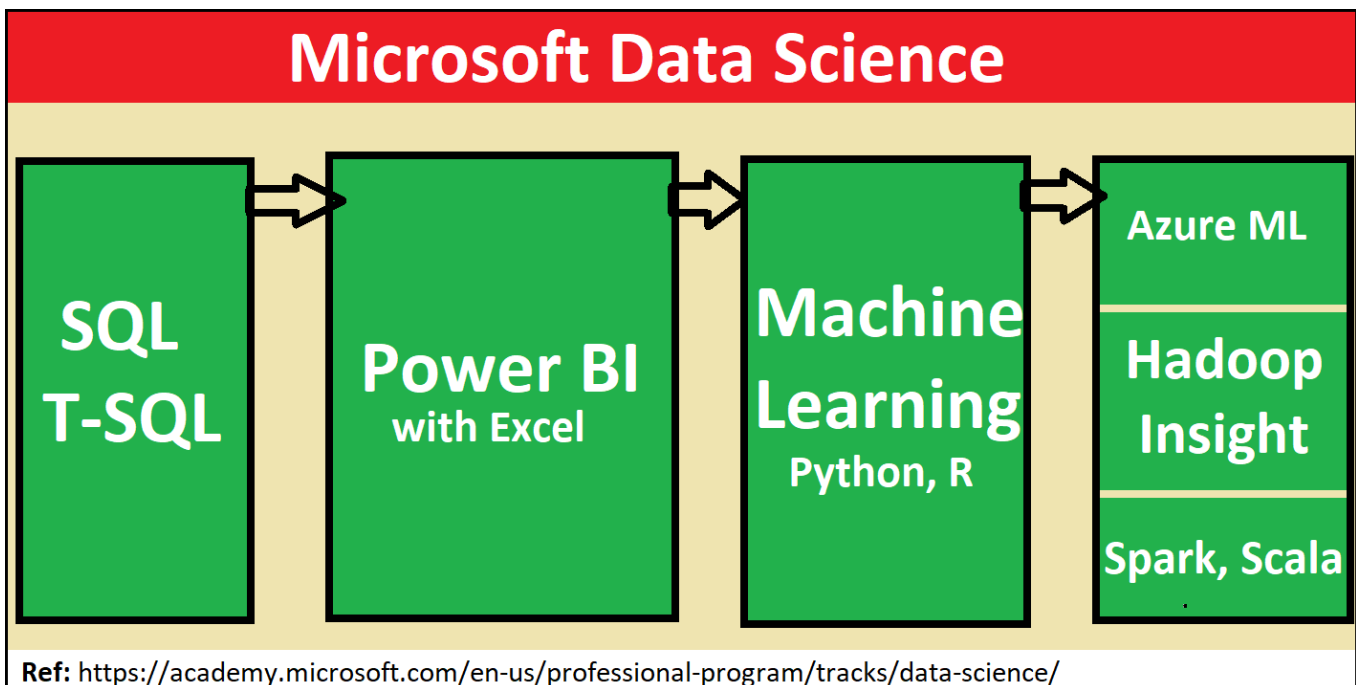


# SQL School <sup>TM</sup>

## Quality Training Assured

### MICROSOFT DATA SCIENCE TRAINING

Complete Practical, Real-time Job Oriented Training



### Course Plan

Training Module		Duration	Plan A	Plan B	Plan B
Module 1	SQL Basics, SQL Server, T-SQL Queries	2.5 W	✓	✓	✓
Module 2	Power BI & Big Data Analytics, Excel	3.5 W	✓	✓	✓
Module 3	Machine Learning, AI with Python, R	7 W	X	✓	✓
Module 4	Azure Machine Learning, HD Insight	3 W	X	X	✓
<b>Total Duration [Including Project, Resume]</b>		<b>16 W</b>	<b>6 W</b>	<b>13 W</b>	<b>16 W</b>

# Data Science Training Plans

DATA SCIENCE TRAINING - Plans			
	Plan A	Plan B	Plan C
<b>COURSES INCLUDES :</b>	T-SQL, Power BI	T-SQL, Power BI, Python, R for Machine Learning	T-SQL, Power BI, Python, R for ML, HDFS, Spark, Azure Machine Learning
Data Scientists - Basics, Job Roles	✓	✓	✓
SQL and Database Basics	✓	✓	✓
SQL Server & T-SQL Basics	✓	✓	✓
Queries, Joins, Data Access	✓	✓	✓
Power BI For Big Data Analysis	✓	✓	✓
Power BI For Statistical Analysis	✓	✓	✓
Power BI with Excel - Data Analysis	✓	✓	✓
Power BI with Report Server, Cloud	✓	✓	✓
Power Query, DAX & Data Modelling	✓	✓	✓
Python for Machine Learning	X	✓	✓
R for Machine Learning	X	✓	✓
Artificial Intelligence with R, Python	X	✓	✓
NLP and Deep Learning with R, Python	X	✓	✓
Unsupervised Learning with R, Python	X	✓	✓
Artificial Neural Networks Algorithms	X	✓	✓
Statistical Application, Big Data Analysis	X	✓	✓
Text Mining, Graphics systems in R	X	✓	✓
Azure Machine Learning (AML) Tools	X	X	✓
Internet Of Things (IoT) Implementation	X	X	
Hadoop Insight For Big Data Analytics	X	X	✓
Spark and Scala with Hadoop Insight	X	X	✓
Apache Zeppelin and HIVE ODBC	X	X	✓
Big Data with Azure Machine Learning	X	X	✓
<b>Course Duration</b>	<b>6 Weeks</b>	<b>13 Weeks</b>	<b>16 Weeks</b>

# Module 1: SQL Server & T-SQL Concepts

Applicable for Plans A, B, C

## Chapter 1 : INTRODUCTION TO DATA SCIENCE

Need for Big Data and BI Technologies; Purpose of Data Science Techniques and Tools. Microsoft Data Science Track: Curriculum and Implementation Models. Data Scientist Job Roles and Responsibilities. Big Data Analytics, Machine Learning, Artificial Intelligence, Internet Of Things (IOT), Natural Language Processing (NLP) and Deep Learning with Microsoft Data Science. Need for SQL Server, Excel, Power BI and Azure Machine Learning (AML) for Data Scientists. Course Plan and Lab Setup;

## Chapter 2: SERVER INSTALLATION and TOOLS INSTALLATION

SQL Server 2016 Installation, Guidance; SQL Server 2017 Installation, Guidance; Instance Types: Default, Named Instances; SQL Server Features; SQL Server Database Engine For OLTP; File Stream and Collation; SQL Server Authentication Modes; Windows Login & SQL Server sa Login; System Databases: Master, Model, MSDB, TempDB, Resource Database; SQL Server Management Studio. SSMS; Test Connection to Local Servers and Remote (Online) Servers;

## Chapter 3: DATABASE & SQL BASICS - Level 1

SQL : Purpose and Real-time Usage; DDL, DML, SELECT, DCL and TCL; SSMS Tool : Connections and Usage; SQL Versus T-SQL : Basic Difference; Server Connections, Session Creations; Creating Databases and DB Connections; Creating Tables. Int, Char Data Types; Single Row Inserts, Multi Row Inserts; INSERT and INSERT INTO Statements; SELECT Statement for Table Retrieval; WHERE Conditions with =, OR, IN; AND, OR, NOT, IN, NOT IN;

## Chapter 4: DATABASE & SQL BASICS - Level 2

Creating Databases: Files [MDF, LDF]; Single Row Inserts, Multi Row Inserts; SELECT... WHERE Conditions, Operators; AND, OR, NOT, Mathematical Operators; IN, NOT IN, BETWEEN, NOT BETWEEN; IS NULL, LIKE, NOT LIKE. % and \_; CHAR Vs VARCHAR Data Types; GO Statement, SQL BATCH; DISTINCT, TOP, FETCH, ORDER BY; Basic Sub Queries; UPDATE, DELETE, TRUNCATE, ALTER, ADD and DROP, UNION ALL & UNION;

## Chapter 5: SQL BASICS 3, EXCEL IMPORTS

Schemas : Real-time Usage, Security; Schemas Creation, Usage with Tables; Schema - Table Transfer. 2P, 3P Naming; Temporary Tables : Real-time Use; Local and Global Temporary Tables; Temp Tables and Session Scope; Temp Tables and Connection Scope; Excel File Imports into SQL Database; Using Import / Export Wizard in SSMS; Import Data @ Multiple Excel Sheets; Working with Microsoft JET Driver; SQL Native Client (SNAC) Connections; GUI for Database and Tables Creation; GUI Limitations for Big Data Inserts;

## Chapter 6 : CONSTRAINTS, INDEXES - BASICS

Constraints and Keys - Data Integrity; NULL, NOT NULL Property on Tables; UNIQUE KEY Constraints: Importance; PRIMARY KEY Constraint: Importance; FOREIGN KEY Constraint:

Importance; REFERENCES, CHECK and DEFAULT; Candidate Keys and Identity Property; Diagrams and ER Models; Indexes : Basic Types and Creation; Index Search Advantages; Clustered & NonClustered Indexes; Primary Key and Unique Key Indexes;

### **Chapter 7: JOINS, T-SQL QUERIES - Level 1**

JOINS - Table Comparisons Queries; INNER JOIN - Examples, WHERE, ON; OUTER JOIN - Examples, WHERE, ON; Left Outer Joins with Example Queries; Right Outer Joins with Example Queries; FULL Outer Joins - Real-time Scenarios; MERGE, LOOP, HASH Join Options; Big Table Versus Small Table Joins; Join Types Versus Join Options in T-SQL; CROSS JOIN Versus CROSS APPLY; Joining Unrelated Tables, Options;

### **Chapter 8: JOINS, T-SQL QUERIES - Level 2**

GROUP BY Queries and Aggregations; GROUP BY Queries with Having Clause; Group By Queries - Query Design Rules; ROLLUP( ) & CUBE( ) Summary Values; GROUPING() Function for Row Status; Comparing WHERE, HAVING Conditions; Using UNION and UNION ALL in Queries; Date and Time Functions in T-SQL; DATEADD, DATEDIFF, Year, Month, Day; CAST, CONVERT, STRING Functions; PIVOT Functions and Normalization; RANK, DENSE\_RANK, ROW\_NUMBER; PARTITION BY and MERGE Statement; MERGE with IIF and CASE Statement;

### **Chapter 9: Views, Functions, Procedures - Basics**

Views :Usage in Real-time; System Predefined Views, Audits; Listing Databases, Tables, Indexes; Functions : Types, Usage in Real-time; System Predefined Functions, Audits; Variables & Parameters in SQL Server; Procedures :Usage in Real-time; Working with Parameters and Dynamic Joins in Real-time; Working with Data Access using Functions and Stored Procedures;

### **Chapter 10: Triggers and Transactions - Basics**

Triggers - Purpose, Types Of Usage; DML Triggers; Transactions; Auto Commit Transaction, ACID Options; Open Transactions & Query Blocking Scenarios @ Real-time; NOLOCK and READPAST Lock Hints; Schemas Creation, Use, Table; Real-time Usage; Synonyms : Creation, Real-time use;

### **Chapter 11: SQL SERVER ARCHITECTURE**

SQL Server Architecture Components; Protocols, SQL Native Client (SNAC); Parser, Compiler, SQL Query Validations; Query Optimizer (QO) and SQL Manager; Storage Engine, File and DB Manager; Transaction Manager and Lock Manager; Buffer Manager, SQL OS and IO Buffer; Synchronization and Thread Scheduler; MDAC and CLR; Local & Global Temporary Tables;

### **Chapter 12 : EXCEL INTEGRATIONS, ER MODELS**

Self Referencing Keys and Self Joins; Composite Keys and Composite Indexes; 1:1, 1:M, M:1, M:M Relationship Types; ER Diagrams for Table Relations, Joins; Normal Forms for Entity Relationships; First, Second, Third Normal Forms Usage; Boycee-Codd Normal Form : BNCF : Usage; Functional Dependency and Candidate Keys; Multi-Valued and Transitive Dependencies; 4 NF, EKNF and ETNF Differences, Usage;

## Module 2: Power BI Report Design For Big Data Analytics

Applicable for Plans A, B, C

### CHAPTER 13 : POWER BI ARCHITECTURE

Power BI Architecture, Tools; Power BI - Integrations in Real-time; Need for Power BI Service & Cloud; Need for Power BI Report Server, SSOT; Power BI Report Server & Report Builder; Mobile Report Publisher; Report Types: Interactive; Analytical, Paginated & Mobile Reports; Cortana - Q&A; Configure Power BI Cloud; Install Power BI Desktop, Gateways;

### CHAPTER 14 : USING POWER BI DESKTOP

Install and Configure Power BI Desktop; Understanding Power BI Desktop Tool; In-Memory, SMDL and Vertipaq Options; Visuals, Fields, Pages and Filters; Report, Data and Relationship Options; Concept & Use of PBIX and PBIT Files; Designing Simple / Basics Reports in PBI; Get Data from DAT Files, Basic Reports; Enter Data from Excel Files, Basic Reports; Creating PBIX Files and Re-Use; Creating PBIT Files; Table Reports; Interactive Visuals and Focus Mode;

### CHAPTER 15 : SLICER & VISUAL SYNC

Visual Interactions in Power BI; Edit Interactions - Format Options; Filter and Slice Control Options; Interactions with Related Data; Slicer : Number, Text, Date Data; Slicer Orientation, List, Dropdowns; Visual Sync Property - Limitations; Grouping and Binning with Fields; Grouping Static / Fixed Data Values; Grouping Dynamic / Changing Data; Bin Size and Biz Limits (Max, Min); Bin Count and Grouping; Min and Max Value; Value Distributions and Ranges;

### CHAPTER 16 : HIERARCHIES&DRILLDOWN ACTIONS

Creating Hierarchies in Power BI; Hierarchies with Field Groups; Hierarchies in Table Visuals; Hierarchies in Matrix, Charts; Hierarchies with Date, Time; Built-In Hierarchy Labels in Visuals; Inline Hierarchy Labels in Visuals; Hierarchies: Issues and Solutions; Independent Drill-Down; Dependant Drill-Down; Conditional Drilldowns, Data Points; Drill Up, Include & Exclude; Drill Thru Reports, Conditional Filters; Show Data, Export Data, See Records;

### CHAPTER 17 : FILTERS, DRILLTHRU, BOOKMARKS

**Filters** : Types and Usage in Real-time; Visual Filter, Page Filter, Report Filter; Basic, Advanced and TOP N Filters; Category Level, Summary Level Filters; Slicer Versus Filters - Comparisons; DrillThru Filters and Drill Thru Reports; Keep All Filters" Options in **DrillThru**; Drill-thru Options with Page Navigations; **Bookmarks** : Navigations, Real-time Use; Using Bookmarks for Visual Filters; Using Bookmarks for Page Navigations; Using Selection Pane with Bookmarks; Using Buttons, Images with **Actions**; Buttons, Actions and URLs; Image Actions, **Selection** Pane;

### CHAPTER 18 : BIG DATA ACCESS, MODELING

Power BI Reports with OLTP Databases; Power BI Reports with Big Data Sources; Azure Cloud Database Access, Reports; Import and Direct Query with Power BI; Using SQL Queries in Power BI Desktop; Using Views, Tables, Functions in PBI; Data Modeling : Summary, Format; Data

Modeling : Currency, Relations; Relationship Management with OLTP;1:1, 1:M, M:1 Relations in Power BI; Synonyms : Creation and Usage; Connection with MS Access Databases;

### **Chapter 19 : POWER BI VISUAL PROPERTIES – 1**

Fields, Formats & Analytics; Table Visuals & Properties, Filters; Data Bar, Data Scaling; Divergent Property, Data Labels; Matrix : Sub Totals, Grand Totals; Drilldown: Row, Column; Stacked Bar Chart, Clustered Colum Chart, Stacked, Clustered Colum Chart; 100% Stacked Bar, Column Charts; Area, Stacked Area Chart. Ribbon Chart; Waterfall Chart, Scatter Chart, Pie Chart; Line, Doughnut, TreeMap, Funnel, Gauge; Labels, Values, Legend; Axis, Title, Filters.

### **Chapter 20 : POWER BI VISUAL PROPERTIES –2**

Map Reports and Filled Map Reports; ArcGIS Maps and Real-time Usage. Options; Longitudes and Latitudes. Bing Maps; Color Analytical Maps, Styles, Options; KPI Reports - Target, Value and Trend; Plot Area & Responsiveness Properties; Sentiment Colors, Color Saturations; Breakdowns, Title, Axis Properties; Category & Data Labels, Constant Lines; Working with SLICER Visual : Value Filters; Horizontal and Vertical Slicer Filters; Single Select and Multi Select Options; Frown (Errors), Details. Lock Aspects; Custom Visuals and PBIVIZ File; PBI & VM;

### **CHAPTER 21 : POWER QUERY &M LANGUAGE – 1**

Power Query [M Language] - Purpose; Power Query Usage & Operation Types; Power Query Architecture and Usage; QUERY Concept, Properties, Validations; Power Query for Data Mashup Operations; Basic Data Types, Literals and Values; Expressions and Primitives in M Language; LIST : Syntax, Examples and Usage; RECORD : Syntax, Examples and Usage; TABLE : Syntax, Examples and Usage; Power Query Connection Formats, Settings; let, source & in M Lang; Power BI Canvas: Edits, Applied Steps; Queries and Applied Steps, Edit Queries;

### **CHAPTER 22 : POWER QUERY &M LANGUAGE – 2**

Data Sources with Excel, File Formats; Data Sources with Database Connections; BLANK Data Sources Creation & Scope; Creating LIST, RECORD& TABLE Queries; Functions in Power Query and Usage; Defining Functions, Invoke Options; Mashup Operations in Power Query Editor; Row Filters, Column Filters, Renames; Promoting Headers and Query Settings; MERGE Queries For Combining Queries; Inner Join and Left/Right Outer Joins; Left Anti Join, Right Anti Join, Full Join; UNION All, Group By and Aggregations; Close and Apply Options. Data Imports;

### **CHAPTER 23 : POWER QUERY &M LANGUAGE – 3**

Creating Parameters in Power Query; Intrinsic & Query Parameters, Usage; Parameter Data Types, Default Value Lists;; Static Lists, Dynamic Lists For Parameters; Reports with Range Values - Usage; Parameters with PBIT Reports, Prompts ; New Queries from Dataset Fields, Usage; Removing Duplicate Rows and Columns; Column Delimiters and Field Formatting; Change Type and Transform Operations; Replace Values & Errors, Transpose; Group By Aggregations, Fill, Reverse Rows; Format Options in M Language; Pivot and Unpivot Options in Power Query; Data Type Conversions & Format Options; Visual Sync Features and Bookmarks;

## **CHAPTER 24 : POWER QUERY & M LANGUAGE – 4**

Data Modeling Options with Power Query; Modeling Operations - Custom Columns; Query Transforms, Sort Direction, Statistics; Enabling / Disabling Loads. ToList Options; Creating Parameters & Variables. Functions; Date and Time Columns. String Functions; Advanced Edit Options, Custom Queries; Custom Functions in Power Query M Lang; Grouping and Sub Groups with Queries; Binning with Groups and Query Formats; SubString Functions, Splits (Row, Column); Parsing XML and JSON Files. Formatting;

## **CHAPTER 25 : DAX FUNCTIONS, EXPRESSIONS – 1**

DAX : Importance and Real-time Usage; DAX Data Types and DAX Calculations; Syntax, Functions, Context Options; ROW Context and Filter Context in DAX; DAX Functions. Aggregates and Usage; Creating and Using Measures with DAX; Creating and Using Columns with DAX; Vertipaq Engine & Special Characters; DAX Real-time Usage; Quick Measures in DAX;

## **CHAPTER 26 : DAX FUNCTIONS, EXPRESSIONS – 2**

In-Memory Processing: DAX Performance; Date and Time & Text Functions; Time Intelligence Functions in DAX; Logical & Mathematical Functions; PowerPivot xVelocity, Vertipaq Store; Data Modeling with DAX. Creating Roles; Currency Conversions, Field Formatting; SUM and SUMX Functions: Differences; CALCULATE(), CALCULATEX() Functions; CALCULATETABLE Functions;

## **CHAPTER 27 : DAX FUNCTIONS, EXPRESSIONS – 3**

VAR, VARP; DATESMTD, DATESQTD, DATESYTD; ENDOF(), FirstDay(), LastDay(); SAMEPERIODLASTYEAR, STARTOF(); Other DAX Functions and Examples; DAX Expressions with Quick Measure; DAX Usage for Row Level Security; Creating Roles with Power BI Desktop; DAX Filters & Multi Value Conditions; Manage Roles and Membership Options; VIEW AS ROLE;

## **CHAPTER 28 : DAX FUNCTIONS, EXPRESSIONS – 4**

DAX Variables, Joins and Usage; Value Assignments with DAX; CONCATENATE Function in DAX; HOUR, NOW, SWITCH, TRUE; RETURN Values with DAX Functions; Numeric and Other DAX Options; DAX Modeling Components - use; TABLES, COLUMNS, RELATIONS; MEASURES and HIERARCHIES; INTELLISENSE with DAX Editors; Comments in DAX - Syntax Options; DISTINCTCOUNT and DIVIDE; SAMEPERIODLASTYEAR and IF; Revenue and Budget with DAX; VARIANCE Calculations in DAX; TOTALYTD, YTD, QTD and MTD;

## **CHAPTER 29 : POWER BI SERVICE (CLOUD) – 1**

Power BI Cloud & Power BI Workspace; App Workspace Creation in Real-time; Publish Reports from Power BI Desktop; Reports and Datasets in Power BI Cloud; Pin Visuals and Pin LIVE Report Pages; Dashboard Creation and Tiles, Media; Images, Web Content, Videos, Q& A; Mobile View, Web View, QR Codes, Shares; Enabling Interactive Visuals. Embed Codes; Report Shares and Dashboard Shares; App Publish, Updates; Power BI Report Edits, Visuals; Download PBIX;

## **CHAPTER 30 : POWER BI SERVICE (CLOUD) – 2**

Dashboard Properties and Security; Report properties and Security; Dataset properties and Security; Report Insights, Related Items & Metrics; Quick Insights. Publishing App Workspace;

Power BI LIVE Report Edits, Downloads; Content Packs : Creation; Organizational and Service Level; Content Packs; Personal and Enterprise Gateways; NT SERVICE\PBIEgwSERVICE Account; Dataset Refresh Options and Schedules; Azure Databus Integration, ODG Logs

**CHAPTER 31 : EXCEL WITH POWER BI**

Using Excel with Power BI Reports; Using Excel Analyser in Power BI; Using Excel Publisher in Cloud; Creating Datasets with Excel, Office 365; Excel Uploads & Cloud Imports, Office 365; Cell Selection and PINS in Power BI; Excel Online Service - Edits and Pins; Excel ODC Connections and Real-time Use; Excel Power Pivot and Power BI Cloud; Excel Drilldowns and Drillthru Reports; KPI Reports; Excel for Big Data Analysis, Reporting; Worksheets and Dashboards with Excel; Power Pivot Reports in with Excel; Power View Report Options in Excel;

**CHAPTER 32 : POWER BI REPORT SERVER**

Settings and Power BI Admin Console; Configuring Power BI Report Server; Report Server Database TempDB Database; Webservice URL, Webportal URL - Usage; Report Builder Installation & Usage; Designing Paginated Reports (RDL), Tests; Deploy to Power BI Report Server, Settings; Data Source Connections, Report Options; Register Power BI Report Server to Cloud; Tenant IDs Generation and Real-time Usage; Publish RDL Reports to Power BI Cloud;

**CHAPTER 33 : REST APIS & POWER BI ADMIN**

REST APIs with Power BI Cloud Service; Streaming Datasets Creation in Power BI; API : Usage. SubKeys and Channels; Push and Pubnub Datasets in Power BI; Azure Stream for Real-time Data Reports; Real-time Data Tracking Visuals (Interactive); Power BI Mobile Publisher and Mobile Reports; Power BI Vs Tableau, SSRS, Share Point; Licensing & Pricing Plans in Power BI; MSBI SSAS and SSRS with Power BI; Custom Visuals in Power BI; Alerts in Power BI; App Publish;

**REAL-TIME PROJECT(ERP)**

✓ End to End Implementation - ERP	✓ Project based Interview Questions
✓ Real-time Methodologies, Approaches	✓ Project Explanation with Solutions
✓ Big Data Access in Azure Cloud	✓ Line of Actions, Implementations
✓ Power BI Cloud Integration	✓ Resume Support with Project Work
✓ Power BI Report Server	✓ Face - to - Face Mock Interviews
✓ Power BI with SQL_AS	✓ Complete Solution to Project

**Module 3.1: Python For Machine Learning**

Applicable for Plans B, C

**CHAPTER 34 : INTRODUCTION TO SCRIPT**

What is Script in Python?; What is a program in Python?; Types of Scripts in Python?; Difference between Script; programming languages list; main features of scripting Lang.; limitation of client side scripting; Programming Language Paradigms; Basic understanding of Python; Is Python a compiled language?; where is python used in real life?; Why is Python called Python?;

**CHAPTER 35 : INTRODUCTION TO PYTHON**



What is Python Programming?; Why Python is used in DS?; Where is python Mostly used?; Characteristics of Python Programming; History of Python Programming Language; What is PSF Python Programming?; Python Versions - Python Application; How to Download Python, print to the screen; How to Install Python , Creating Program; Install Python with Diff IDEs; Features of Python Programming; Limitations of Python Programming;

### **CHAPTER 36: Data Analytics**

Introduction to data Big Data?; Introduction to NumPY and SciPY; Introduction to Pandas and **Matplotlib**; Data Science; What is Data Science in Python; Data Science Life Cycle in python; what is data analysis using python; what is Data Mining in Python; Analytics vs Data Science in python; How Python is used in big data?; Is Python or R better for data science?; Why is Python used in data science?

### **CHAPTER 37 : String Handling for Data Science**

what is String ? - String Operations - String indices; String Functions - len , upper, lower, join, Split; SwapCase(), Title(), find(), isupper(), islower(); Delete a string - Python Keywords; String Multiplication and concatenation; Python Identifiers - Python Literals; string formatting operator in python; Built-in String Methods - Data Structures; Structuring with indentation in Python; Define Data Structure in Python Language; Reverse words - Reverse Characters Examples; How do you split a string in Python?;

### **CHAPTER 38: Python Conditional for Data Science**

Control Structures - Sequential Control Structure; Selective and Repetative Control Structure; How to use “if condition” in conditional; control Structures in python; if statement (One-Way Decisions); if ..else statement (Two-way Decisions); How to use “else condition”; if .. elif .. else statement (Multi-way); “else condition” working; How to use “elif” condition; Execute conditional statements; minimal code - Nested IF Statement; Nested IF Statement in python;

### **CHAPTER 39: Python LOOPS for Data Science**

How to use While loop and For loop; Break and Continue Statements in For loop; Python Enumerate function for For Loop; **Sequence or Collections and Lists for Data Science**; Strings - Unicode Strings; Lists - Tuples - Sets - Dictionary - Xrange; Lists are mutable - Accessing the List; Updating a List - Deleting a List; List indices - Traversing a list; List operations - List Slices - List Methods; Map, filter and reduce - Deleting elements; What is list of list in python?; What is Python list function?; How do you add to a list in Python?

### **CHAPTER 40: Python TUPLE for Machine Learning**

Advantages of Tuple over List; Packing and Unpacking - Tuples; Creating Nested tuple - Examples; Deleting Tuples - Slicing of Tuples; Comparing Tuple Membership Test; Built in Functions ,Dotted Charts; **Python Sets for Machine Learning**; how to create/declare a set in python; Iteration Over Sets - Python Methods; Python Set Operations - Union of Sets; Built-in Functions with Set; python frozenset get element;

### **CHAPTER 41 : Python Dictionary for Machine Learning**

How to create a dictionary?; PYTHON HASHING - Dictionary Methods; Copying dictionary -

Updating Dictionary; Delete Keys from the dictionary; Sorting the Dictionary - Dictionary len(); Python Dictionary in-built Functions; Variable Types - python List Cmp(); Python List cmp() Method; Python Dictionary Str(dict); How do you create a dictionary in Python?; Can Python dictionary have multiple values?; How do you add to a dictionary in python?

#### **CHAPTER 42 : Python Functions for Data Science**

What is a function? - Types of Function; How to define and call a function in Python; Significance of Indentation (Space) in Python; How Function Return Value?; Types of Arguments in Functions; Default Arguments - Non Default Arg.; Keyword Arguments - Non Keyword Arg.; Arbitrary Arguments in python; Various Forms of Function Arguments; Scope and Lifetime of variables - Nested Fun; Call By Value, Call by Reference in python; Anonymous Functions/Lambda functions;

#### **CHAPTER 43 : Python Modules for Data Science**

What is a Module? - Types of Modules; The import Statement - The from... import st; ..import \* Statement - Underscores in python; The Dir() Function in python; Creating User defined Modules; Command line Arguments in python; Getting Python Module Search Path; What are modules and packages in Python?; What is Python import statement?; How do you import random in Python?; import <module\_name> string python; from <module\_name> import <name(s)>; from <module> import <name> as <name>;

#### **CHAPTER 44 : Packages in Python for Data Science**

What is a Package in Python?; Introduction to Packages?; py file - Creating a package; Importing module from a package; Creating Sub Package in Python; Importing from Sub-Packages; Most Popular Python Packages; How many libraries are there in Python?; What are libraries in Python?; What is the difference between NumPy&SciPy?; Why is SciPy and NumPy used?; Python what is Seaborn? - Examples; Is NumPy a Python framework?;

#### **CHAPTER 45 : Python Date and Time for Machine Learning**

How to Use Date &Date Time Class; How to Format Time Output; How to use TimedeltaObjects; Calendar in Python; datetime classes in Python; How to Format Time Output?; Python Calendar Module,Time Module; Python Text Calendar; Python HTML Calendar Class; Unix Date and Time Commands; Python strftime(); How strftime() works?;

#### **CHAPTER 46 : Machine Learning Basics**

Gain an in-depth understanding of data structure and data manipulation; What Converting business problems to data problems; Understanding supervised and unsupervised learning with examples; Understanding biases associated with any machine learning algorithm; Ways of reducing bias and increasing generalization capabilities; Drivers of machine learning algorithms; Cost functions; Brief introduction to gradient descent; Importance of model validation; Methods of model validation; Cross validation & average error; Understand and use linear and non-linear regression models and classification techniques for data analysis; Gain expertise in mathematical computing using the NumPy and Scikit-Learn packages; Learn to analyze data using Tableau and become proficient in building interactive dashboards;

## **CHAPTER 47 : Generalized Linear Models in Python**

Linear Regression; What is linear regression used for?; How does a linear regression work?; What is a simple linear regression model?; What is linear regression for dummies?; Regularization of Generalized Linear Models; Ridge and Lasso Regression; Logistic Regression; Methods of threshold determination and; performance measures for classification score models; What is logistic regression in Python?; What is penalty in logistic regression?; What is solver in logistic regression?; What is logistic regression in ML?;

**Tree Models using Python;** Introduction to decision trees; Tuning tree size with cross validation; Introduction to bagging algorithm; Random Forests; Grid search and randomized grid search; Extra Trees (Extremely Randomized Trees); Partial dependence plots;

## **CHAPTER 48 : Boosting Algorithms using Python**

Concept of weak learners; Introduction to boosting algorithms; Adaptive Boosting; Extreme Gradient Boosting (XGBoost); What is a strong learner?; What is boosting in ML?; What is bagging and boosting in machine learning?; What is learner in machine learning?; What is the difference between GBM and XGBoost?; What is linear regression used for?; How does a linear regression work?; What is a simple linear regression model?;

### **Support Vector Machines (SVM) &kNN in Python**

Introduction to idea of observation based learning; Distances and similarities; k Nearest Neighbors (kNN) for classification; Brief mathematical background on SVM/|i>; Regression with kNN& SVM; Can Knn be used for regression?; Which is better KNN or SVM?; Is Knn generative or discriminative?; Is SVM better than random forest?; What is SVM Python?; How SVM is used for classification?; Can SVM be used for multiclass classification?; What is SVM in machine learning?; SVM for regression?; Sklearn in Python?; How is SVM trained?; What is SVM score?;

## **CHAPTER 49 : Unsupervised learning in Python**

Need for dimensionality reduction; Principal Component Analysis (PCA); Difference between PCAs and Latent Factors; Factor Analysis; Hierarchical, K-means & DBSCAN Clustering; What is PCA algorithm?; What is PCA used for?; What is PCA in neural network?; How does a support vector machine work?; What is the purpose of factor analysis?; What is factor analysis with example?; Why is factor analysis done?; What are the types of factor analysis?;

## **CHAPTER 50: Artificial Neural Networks in Python**

Introduction to Neural Networks; Single layer neural network; Multiple layer Neural network; Back propagation Algorithm; Neural Networks Implementation in Python; Are neural networks machine learning?; What is neural network explain in detail?; What is neural networks in deep learning?; How do artificial neural networks work?; What is the difference between neural networks and machine learning?; Does machine learning require neural networks?; What is the difference between deep learning and neural networks?; Is artificial neural network and neural network same?; What are the advantages of neural networks in machine learning?; What is back propagation algorithm in neural network?; What is the back propagation algorithm?; What is back propagation algorithm in machine learning?;

## **CHAPTER 51 : Text Mining in Python**

Gathering text data using web scraping with urllib; Processing raw web data with BeautifulSoup; Interacting with Google search using urllib with custom user agent; Collecting twitter data with Twitter API; Naive Bayes Algorithm; Feature Engineering with text data; Sentiment analysis; What is text mining in python?; What is preprocessing in text mining?; How do you do text mining?; What is text mining in data science?; What is naive Bayes classification algorithm?; What is naive Bayes used for?; What is naive Bayes algorithm in data mining?; Is naive Bayes supervised or unsupervised?;

## Module 3. 2 : R-Language For Machine Learning

Applicable for Plans B, C

### Chapter 52: Essential to R programming

Introduction to the R language; Programming statistical graphics; Programming with R; Simulation; Computational linear algebra; Numerical optimization;

### Chapter 53: Data Manipulation Techniques using R programming

Data in R; Reading and Writing Data; R and Databases; Dates, Factors, Subscribing; Character Manipulation; Data Aggregation, Reshaping Data;

### Chapter 54: Statistical Applications using R programming

The R Environment; Probability and distributions; Descriptive statistics and graphics; One-and two-sample tests; Regression and correlation; Analysis of variance and the Kruskal–Wallis test; Tabular data; Power and the computation of sample size; Advanced data handling; Multiple Regression; Linear models; Logistic regression; Survival analysis; Rates and Poisson regression; Nonlinear curve fitting;

### Chapter 55: Data Science with R

Introduction; About S, About R, About CRAN; Installation of R; About working directory; Changing working directory temporarily; Changing working directory Permanently; Installation of R studio; Atomic Datatypes in R;

### Chapter 56: Vectors

Creating vector; Naming Vector; Vector selections; Adding elements to vector; Update elements of vector; Delete elements of vector; Functions (c(), names() ...etc);

**Matrices** : What is matrix; Create matrix; Naming a matrix; Arithmetic with matrix; Adding row; Adding column; Selection of matrix elements; Insert /delete/update matrix elements; Transpose matrix; Combine rows of matrix; Combine columns of matrix;

### Chapter 57: Factors

Categorical variables; Continuous variables; What is factor; Factors in Data Frame; Factor Levels in customized format; Nominal factors; Ordinal factors; components of a factor; How to modify a factor?; Updating Factors; methods for handling factors;

**Chapter 58: Functions**

Introduction to functions; Function documentation; Use a function; Create own function; Nested functions; Function scoping;

**Chapter 59: Packages**

Available R Packages; Packages installation; Default packages; Create package; Attach package...etc; Load Package to Library;

**Chapter 60: Graphics systems in R**

Base graphics, Plot; Histogram, Scatter; Bar plot, Qqplot; Sunflowerplot, Boxplot; Add more detail to graphs; Grid graphics; Lattice graphics; ggplot2 graphics; Data layer; Aesthetics layer; Geometries layer; Facets layer; Statistics layer; Coordinates layer; Themes layer;

**Chapter 61: Cleaning data( equal to ETL work)**

gather function; spread() function; separate() function; unite() function; Working with lubridate package; Working with stringr package; Working with Missing values; Working with Special values;

**Chapter 62: Machine Learning & Artificial Intelligence**

What is machine learning?; What is ETP?; Types of machine learning; Supervised learning; Unsupervised learning; Semi-supervised Learning; Reinforcement learning; Algorithms or Model for Machine Learning; Linear Regression, Logistic Regression; Jackknife Regression; Density Estimation, Confidence Interval; Test of Hypotheses; Pattern Recognition; Supervised Learning; Time Series & Decision Trees; Random Numbers; Monte-Carlo Simulation; Bayesian Statistics; Naive Bayes, Principal Component;

**Chapter 63: Analysis - (PCA)**

Ensembles, Neural Networks; Support Vector Machine - (SVM); Nearest Neighbours - (k-NN); Feature Selection - (aka Variable Reduction); Indexation / Cataloguing and Collaborative Filtering; (Geo-) Spatial Modelling, Graphs; Recommendation Engine; Search Engine and Attribution Modeling; Rule System, Linkage Analysis; Association Rules, Scoring Engine; Segmentation, Predictive Modelling;

**Chapter 64: Deep Learning**

Reinforced Learning, Reinforcement learning Process Flow, Reinforced Learning Use cases, Deep Learning, Biological Neural Networks, Understand Artificial Neural Networks, Building an Artificial Neural Network, How ANN works, Important Terminologies of ANN's;

**Chapter 65: Text Mining : The concepts of text-mining**

Use cases; Text Mining Algorithms; Quantifying text; TF-IDF; Beyond TF-IDF;

# Module 4 : Azure Machine Learning, HD Insight, Scala & Spark

Applicable for Data Science Training : Plan C

## **Chapter 66: Azure Machine Learning (AML)**

Familiarity with Azure Machine Learning; Azure Machine Learning Techniques; Deep Learning Models; AI Models in Azure Machine Learning; AML Studio Installation and Real-time usage;

## **Chapter 67: Basics of Machine Learning**

Basics of Machine Learning; What You Will Learn in This Section; The course slides for all sections; Important Message About Udemy Reviews; Why Machine Learning is the Future?; What is Machine Learning?; Understanding various aspects of data - Type, Variables, Category; Common Machine Learning Terms - Probability, Mean, Mode, Median, Range; Types of Machine Learning Models - Classification, Regression, Clustering etc

## **Chapter 68: Started with Azure ML**

Getting Started with Azure ML; What You Will Learn in This Section?; What is Azure ML and high level architecture.; Creating a Free Azure ML Account; Azure ML Studio Overview and walk-through; Azure ML Experiment Workflow; Azure ML Cheat Sheet for Model Selection

## **Chapter 69: Data Processing**

Data Processing; Data Input-Output - Upload Data; Data Input-Output - Convert and Unpack; Data Input-Output - Import Data; Data Transform - Add Rows/Columns, Remove Duplicates, Select Columns; Apply SQL Transformation, Clean Missing Data, Edit Metadata; Sample and Split Data - Partition or Sample, Train and Test Data

## **Chapter 70: Classification**

Classification; Logistic Regression - What is Logistic Regression?; Logistic Regression - Build Two-Class Loan Approval Prediction Model; Logistic Regression - Understand Parameters and Their Impact; Understanding the Confusion Matrix, AUC, Accuracy, Precision, Recall and F1Score; Logistic Regression - Model Selection and Impact Analysis; [Hands On] Logistic Regression - Build Multi-Class Wine Quality Prediction Model; Decision Tree - What is Decision Tree?; Decision Tree - Ensemble Learning - Bagging and Boosting; Decision Tree - Parameters - Two Class Boosted Decision Tree; Two-Class Boosted Decision Tree - Build Bank Telemarketing Prediction; Decision Forest - Parameters Explained; Two Class Decision Forest - Adult Census Income Prediction; Preview; Decision Tree - Multi Class Decision Forest IRIS Data; SVM - What is Support Vector Machine?; SVM - Adult Census Income Prediction

## **Chapter 71: Hyperparameter Tuning**

Hyperparameter Tuning; [Hands On] - Tune Hyperparameter for Best Parameter Selection; Hyperparameter Tuning

## **Chapter 72: Deploy Webservice**

Azure ML Webservice - Prepare the experiment for webservice; Deploy Machine Learning Model As a Web Service; Use the Web Service - Example of Excel; AzureML Web Service

### **Chapter 73: Regression Analysis**

What is Linear Regression?; Regression Analysis - Common Metrics; Linear Regression model using OLS; Linear Regression - R Squared; Gradient Descent; Linear Regression: Online Gradient Descent; Experiment Online Gradient; What is Regression Tree?; What is Boosted Decision Tree Regression?; Decision Tree - Experiment Boosted Decision Tree; Regression Analysis

### **Chapter 74: Clustering**

What is Cluster Analysis?; Cluster Analysis Experiment 1; Cluster Analysis Experiment 2 - Score and Evaluate; Clustering or Cluster Analysis

### **Chapter 75: Data Processing - Solving Data Processing Challenges**

How to Summarize Data?; Summarize Data – Experiment; Outliers Treatment - Clip Values; Outliers Treatment - Clip Values; Clean Missing Data with MICE; Clean Missing Data with MICE; SMOTE - Create New Synthetic Observations; Preview; [Hands On] – SMOTE; Data Normalization - Scale and Reduce; Data Normalization; PCA - What is PCA and Curse of Dimensionality?; Principal Component Analysis; Join Data - Join Multiple Datasets based on common keys; Join Data - Experiment

### **Chapter 76: Data Correlations**

Pearson Correlation Coefficient; Chi Square Test of Independence; Kendall Correlation Coefficient; Spearman's Rank Correlation; Comparison Experiment for Correlation Coefficients; Filter Based Selection - AzureML Experiment; Fisher Based LDA – Intuition; Fisher Based LDA - Experiment

### **Chapter 77: HD Insight with Power BI**

Microsoft Power BI Desktop has a native connector to perform direct query against HDInsight Interactive Query cluster. You can explore and visualize the data in interactive manner. To learn more see Visualize Interactive Query Hive data with Power BI in Azure HDInsight and Visualize big data with Power BI in Azure HDInsight.

### **Chapter 78: Apache Zeppelin**

Apache Zeppelin interpreter concept allows any language/data-processing-backend to be plugged into Zeppelin. Interactive Query Access from Apache Zeppelin using a JDBC interpreter.

### **Chapter 79: Hadoop Insight (HD Insight)**

Familiarity with Azure HDInsight, Familiarity with databases and SQL, introduction to Data Science with Sparki, Get started with Spark clusters in Azure HDInsight, and use Spark to run Python or Scala code to work with data, Getting Started with Machine Learning, Learn how to build classification and regression models using the Spark ML library., Evaluating Machine Learning Models, Learn how to evaluate supervised learning models, and how to optimize model parameters, Recommenders and Unsupervised Models, Learn how to build recommenders and clustering models using Spark ML

### **Chapter 80: Apache Hadoop**

Apache Hadoop, Apache Spark, Apache Kafka, Apache HBase, Apache Hive LLAP, Apache Storm, Machine Learning

### **Chapter 81: Apache Modules with HDINSIGHT**

#### **Apache Spark**

Run interactive queries | Visualize data | Machine learning

#### **Apache Kafka**

Structured streaming with Kafka | Use with Storm on HDInsight | Use Kafka Producer and Consumer APIs

#### **Apache HBase**

Create HBase clusters in a VNET | Use Apache Phoenix | Connect to Spark

#### **Apache Storm**

Create Storm topology in Java | Deploy Storm topologies on HDInsight | Write from Storm to Data Lake Storage

### **Chapter 82: ML Services**

Use R Tools for Visual Studio and Connect with Power BI using Direct Query

### **Chapter 83: Hive ODBC**

Open Database Connectivity (ODBC) API, a standard for the Hive database management system, enables ODBC compliant applications to interact seamlessly with Hive through a standard interface.

### **Chapter 84: HD Insight with Excel**

Microsoft Excel is the most popular data analysis tool and connecting it with big data is even more interesting for our customers. Azure HDInsight Interactive Query cluster can be integrated with Excel with ODBC connectivity

**Real-time Project for Complete Data Science**

**Resume Guidance and Placement Assistance**

\* All Concepts as per **Microsoft Data Science Curriculum** at:

<https://academy.microsoft.com/en-us/professional-program/tracks/data-science/>