GGlobNet
**HARNESSING SKILLS**

# Diploma in Data Science

| | |
|---|---|
| Topic - 1 | Advanced Excel With VBA |
| Topic - 2 | SQL (Structured Query Language) |
| Topic - 3 | Data Analytics Using R Programing |
| Topic - 4 | Python Programming |
| Topic - 5 | Advanced Python Programing |
| Topic - 6 | SPSS (Statistical Package for Social Science) |
| Topic - 7 | Tableau |

# Advanced Excel with VBA

## Module – 01 : Basic Course – Basic Formula, Functions & Charts

- Formulas and Functions
- Formula Tab
- Examples of Formulas
- Auto Sum Features
- Formula
- Charts
- Smart Art Graphics

## Module – 02 : Advanced Functions in Excel

- Introduction
- If Statements
- Nesting IF()'s
- Using Other Functions with IF's
- Concatenation
- LookUp Functions(VLOOKUP & HLOOKUP)
- Subtotals
- Hiding & Displaying Data

## Module – 03 : Data Validations

- Introduction
- Setting Data Validations
- Data Validation
- Checking for Invalid Data

## Module – 04 : Excels Analytical Tools

- Goal Seek
- Scenarios
- Summarizing Scenarios
- Creating a Scenario Pivot Table Report

## Module – 05 :  Pivot Table

- Using Pivot Tables to analyze data
- Pivoting
- Pivot Table Example
- Guidelines for creating a Pivot Table in Excel
- Creating a Pivot Table

**Module – 06 : Introduction to VBA**

- Introduction to VBA
- Work with VBA objects, properties, methods
- Working with the Visual Basic Editor
- Sub procedure, function procedure, property procedure
- Referring to Objects
- Concepts – Containers or Collections, properties, methods, events,
- Working with Workbook
- Referring to Objects
- Applying Methods
- Working with Variables and Values

**Module – 07 : Working with Variables in Excel VBA**

- Concept of Variables
- Valid and invalid variable names
- Variables - Numeric Data Types
- Variables – Non - Numeric Data Types

**Module – 08 : Message Box & Loop**

- Style Values and Command Buttons
- Return Values and Command Buttons
- Looping
- For...Next loop
- Do.......Loop While
- Do until.............Loop
- Do while.......... Loop
- Do...........Loop until

**Module – 09 : Array in Excel VBA**

- What is an Array?
- Declaring Arrays in Excel VBA
- One Dimensional Array
- Two Dimensional Array

**Module – 10 : Developing Macros in Excel**

- Creating a Macro
- Excel Macro Recording facility
- Modifying the existing Macro in VB editor
- Understanding the Macro and saving a workbook with Macro contents
- Exporting files to different applications

**Module – 11 : UserForm**

- Requirement of UserForm
- Working with objects like textboxes, buttons, check boxes, spin buttons etc.
- Filling up UserForm with pre-defined values
- Macro Coding for different buttons
- Creating Connectivity between UserForm and Excel Worksheet

**Module – 12 : UserForm (Cont'd)**

- Designing UserForm with Validations
- Filling up UserForm with pre-defined values
- Creating Connectivity between UserForm and Excel Worksheet using offset
- Transfer of data from excel file to a word document
- Transfer of data from txt file to an excel document

# SQL (Structured Query Language)

**Introduction to DBMS**

- File Management System And Its Drawbacks
- Database Management System (DBMS) and Data Models
- Physical Data Models
- Logical Data Models
- Hierarchical Data Model (HDBMS)
- Network Data Model (NDBMS)
- Relational Data Model (RDBMS)
- Object Data Model (ODBMS)
- Object Relational Data Model (ORDBMS)
- Conceptual Data Models
- Entity – Relationship (E-R) Model

**Introduction to SQL Server**

- Advantages and Drawbacks Of SQL Server
- Connecting To Server
- Server Type
- Server Name
- Authentication Modes
- Sql Server Authentication Mode
- Windows Authentication Mode
- Login and Password
- Sql Server Management Studio and Tools In Management Studio
- Object Explorer
- Object Explorer Details
- Query Editor

**TSQL (Transact-Structured Query Language)**

- Introduction to TSQL
- History and Features of TSQL
- Types Of TSQL Commands
- Data Definition Language (DDL)
- Data Manipulation Language (DML)
- Data Query Language (DQL)
- Data Control Language (DCL)
- Transaction Control Language (TCL)

**Database**

- Creating Database
- Altering Database
- Deleting Database
- Constrains
- Procedural Integrity Constraints
- Declarative Integrity Constraints
- Not Null, Unique, Default and Check constraints
- Primary Key and Referential Integrity or foreign key constraints
- Data Types In TSQL
- Table
- Creating Table
- Altering Table
- Deleting Table

**Data Manipulation Language**

- Insert
- Identity
- Creating a Table from another Table
- Inserting Rows from One Table to Another
- Update
- Computed Columns
- Delete
- Truncate
- Differences Between Delete and Truncate

**Data Query Language (DQL)**

- Select
- Where clause
- Order By Clause
- Distinct Keyword
- Isnull() function
- Column aliases
- Predicates
- Between … And
- In
- Like
- Is Null

**Top n Clause**

- Set Operators
- Union
- Intersect
- Except
- Joins
- Inner Join
- Equi Join
- Natural Join
- Non-Equi Join
- Self Join
- Outer Join
- Left Outer Join
- Right Outer Join
- Full Outer Join
- Cross Join

**Sub Queries**

- Single Row Sub Queries
- Multi Row Sub Queries
- Any or Some
- ALL
- Nested Sub Queries
- Co-Related Sub Queries
- Exists and Not Exists
- Indexes
- Clustered Index
- Non Clustered Index
- Create, Alter and Drop Indexes
- Using Indexes

**Stored Sub Programs**

- Advantages Of Stored Sub Programs compared to Independent SQL Statements
- Stored Procedures
- Creating , Altering and Dropping
- Optional Parameters
- Input and Output Parameters
- Permissions on Stored Procedures
- User Defined Functions
- Creating, Altering and Dropping
- Types Of User Defined Functions

- Scalar Functions
- Table Valued Functions
- Inline Table Valued Functions
- Multi Statement Table Valued Functions
- Permissions On User Defined Functions
- Triggers
- Purpose of Triggers
- Differences Between Stored Procedures and User Defined Functions and Triggers
- Creating, Altering and Dropping Triggers
- Magic Tables
- Instead Of Triggers
- Exception Handling
- Implementing Exception Handling
- Adding and removing User Defined Error Messages To And From SQL Server Error Messages List

# Data Analytics Using R Programming

**Exploring R**

- Installing R
- Working with Scripts
- Navigating the Workspace

**Reading Datasets into R, Exporting Data from R**

- Using C() command to create Data
- Using scan() command for getting Data in R
- Reading Bigger Data files
- Getting data out of R
- Saving your work in R

**Manipulating and Processing Data in R**

- Deciding most appropriate data structure
- Creating subset of data
- Adding calculated fields to data
- Combining and merging datasets in R
- Sorting and ordering Data
- Introduction to the formula interface
- Putting your data into Shape

**Using Functions and Packages in R**

- Moving from Scripts to Functions
- Using Argument the smart way
- Scope of the function
- Dispatching to a Method
- Packages
- Using Packages

**Descriptive Statistics in R**

- Summary Commands
- Name Commands
- Summarizing Samples
- Cumulative Statistics
- Summary Statistics for Data Frames
- Summary Statistics for Matrix Objects
- Summary Statistics for Lists
- Contingency Tables
- Cross Tabulation

**Analyzing Data Using Functions, Loops, and Data Frames**

- Matrices, Lists, and Data Frames
- Indexing vectors, Matrices, and Lists
- Programming in R

**Graphical Analysis in R**

- Plots for single variable
- Plots with two variables
- Plots with multiple Comparisons
- Plots with multiple Variables
- Special plots
- Saving Graphs to External Files

**Hypotheses Testing in R**

- Introduction to Statistical Hypotheses
- Using the student's t-test
- U-test
- Paired t- and u-test
- Tests for Association
- Goodness of Fit Tests

**Linear Regression in R**

- Basics of Linear Regression Analysis
- Working with Linear Regression
- Simple Linear Regression in R
- Linear Model result Objects

**Using R Commander Package**

# Python Programming

**OVERVIEW**

- History of Python
- Python Features
- Environment Setup
- Getting Python installing Python
- Setting up PATH
- Setting path at Windows
- Python Environment Variables
- Running Python

**BASIC SYNTAX**

- Python Program
- Python Identifiers
- Python Keywords
- Lines and Indentation
- Multi-Line Statements
- Quotation in Python
- Python Comments in Python
- Multiple Statements on a Single Line
- Command Line Arguments
- Accessing Command-Line Arguments
- Parsing Command-Line Arguments

**VARIABLE TYPES**

- Assigning Values to Variables
- Multiple Assignment
- Standard Data Types
- Python Numbers
- Python Strings
- Python Lists
- Data Type Conversion

**BASIC OPERATORS**

- Types of Operators
- Arithmetic Operators
- Comparison Operators
- Logical Operators
- DECISION MAKING
- If…else Statement
- The elif Statement
- Single Statement Suites
- LOOPS While Loop
- The Infinite Loop
- Using else Statement with Loops
- Single Statement Suites
- For Loop
- Break Statement
- Continue Statement

**FUNCTIONS**

- Defining a Function
- Calling a Function
- Passing by Reference Vs Passing by Value
- Global vs. Local variables
- Locating Modules
- Packages in Python

**FILES I/O**

- Printing to the Screen
- Reading Keyboard Input
- The raw_input Function
- The input Function
- Opening and Closing Files
- The open Function
- The file Object Attributes
- The close() Method
- Reading and Writing Files
- The write() Method
- The read() Method

## EXCEPTIONS

- What is Exception
- Handling an Exception
- The except Clause with No Exceptions
- The except Clause with Multiple Exceptions
- The try-finally Clause
- Argument of an Exception
- Raising an Exception
- User-Defined Exceptions
- CLASSES AND OBJECTS
- Overview of OOP Terminology
- Creating Classes
- Creating Instance
- Objects
- Built-In Class Attributes
- Class Inheritance
- Overriding Methods
- Base Overloading Methods

# Advanced Python Programming

- Operating on Data in Pandas & Missing Values
- Hierarchical Indexing
- Combining Datasets – Join, Merge, append etc
- Aggregation & Grouping
- Vectorized String Operations
- Visualisation with Matplotlib
- Analysing Data Through Advanced Visualisation
- Inferential Statistics
- Designing Models with Linear Regression
- Designing Models with Logistic Regression
- Hypothesis Checking
- K-Means Clustering

# SPSS (Statistical Package for Social Science)

**Introducing to SPSS**

- SPSS - Introduction
- The interface
- The variable view
- The data view
- The output view
- The syntax view

**Reading Data from various Sources**

- Reading Data from a Text file
- Reading Data from a Database
- Reading Data from Excel Sheets
- Reading SPSS Data Files

**Variables & Dataset Creation**

- Naming of Variable names
- Creating Labels for Variables
- Variables Type
- Binary Variables
- New data set Creation

**Data Transformations**

- Transformation of Data
- Expressions creation with more than one variable
- Various Conditional Expressions

**Modifying Data Values**

- Data Values Modifications
- New Variable Computation

**Crosstab Report**

- Crosstab Statistics
- Creating of Crosstab Report
- Crosstab cells
- Adding various layers to crosstabs

**Sorting and Selecting Data**

- Data Selection and Sorting
- Split-File
- Creation of Subsets of Cases

**Working with Output**

- Working with Output
- Pivot Table Editor
- Using SPSS Results in Other Applications
- Exporting SPSS Results to Microsoft Excel, Microsoft Word and PDF Files

**Descriptive Statistics**

- Descriptive statistics: Descriptive (univariate)
- Frequencies
- Categorical Data Measurement
- Categorical Data - Charts
- Scale Variables – Their measurements
- Recoding the existing variables

**Univariate Analysis**

- Line Graphs
- Bar Graphs
- Pie Graph
- Graphs for cumulative frequency
- Histograms and frequency statistics – with variables
- Determining the nature of the distribution of continuous variables
- Boxplot
- T-Test and error bar

**Multivariate Analysis**

- Bar Graph for Means
- Line graph for comparing median
- Scatters
- Correlations
- Bivariate correlations
- Partial correlations
- Plotting scatters of several variables against one other
- Execute the analyses for means comparison: t test, between-subjects ANOVA
- Perform the regression analysis (simple and multiple regression)
- Time Series

# Tableau

- Introduction to Data Visualization
- Introduction of Data Visualization using Tableau
- Tableau Basics
- Working with Sorting and Filters
- Creating Dual Axis and Combo Charts
- Table Calculations
- Calculated Field
- Logical Calculations
- Date Calculations
- Parameters
- Using Actions to Create Interactive Dashboards
- Advanced Charts
- Working with data
- Sets
- Drilling Up/Down using Hierarchies
- Grouping
- Bins/Histograms
- Analytics using Tableau
- Dashboards
- Story Telling with Data
- Data Interpreter