

Introduction to SQL Training Course

Course Contents - DAY 1

Course Introduction

- Administration and Course Materials
- Course Structure and Agenda
- Delegate and Trainer Introductions

Session 1: RELATIONAL DATABASE CONCEPTS

- What is a Relational Database?
- Relational Database Structures
- Tables, Rows and Columns
- Indexes, Primary Keys and Foreign Keys
- Data Types
- The Data Dictionary

Session 2: RETRIEVING DATA WITH THE SELECT STATEMENT

- The SELECT Statement
- The SELECT and FROM Clauses
- Conditions and the WHERE Clause
- Other Conditional Operators
- Logical Operators
- The ORDER BY Clause
- Column Aliases
- Arithmetic Operations
- Precedence of Operators

Session 3: AGGREGATE FUNCTIONS

- Overview of Built in Aggregate Functions
- The GROUP BY Clause
- The HAVING Clause

Session 4: JOINING TABLES

- Overview of Table Joins
- Inner Joins
- Table Aliases
- Outer Joins
- Self Joins

- ANSI Standard Joins

Session 5: BASIC SUBQUERIES AND SET OPERATORS

- Overview of Subqueries
- Basic Subqueries
- Set Operators

Excel

Section 1: Foundation

Unit 1.1 – Excel Introduction

A description of the purpose and application of Excel – what it can do, why it's useful.

Unit 1.2 – The Excel Interface

Understanding and working with the Excel interface – getting to know your way around the application, how to access functionality and how to customize it. How to create, save and open Workbooks. Under Save As, a brief explanation of backwards compatibility and other formats (e.g. PDF) save options. Skydrive, alternative online storage, and 'local' saving. Basic printing.

Unit 1.3 – Basic Navigation & Editing

Selection and navigation. Moving around a spreadsheet efficiently. Entering information into cells, types of data (text, numbers, dates). Basic formatting. Working with multiple Worksheets.

Unit 1.4 – Getting going

Takes a quick look at a first 'real' spreadsheet, in order to introduce the concept of functions and formulas, and give a feel for how Excel can be useful.

Section 2: Orientation and efficiency

Unit 2.1 – Editing

Copying, pasting, inserting and deleting cells, ranges of cells, rows, columns and Worksheets. 'Special' pasting – for example to copy just formulas, or just values. Common problems with pasting. The differences between copying and 'cutting'. Excel aids to copying – Fill Handle and Flash Fill.

Unit 2.2 – Viewing

The various 'Views' or modes of spreadsheet display available in Excel, and how to choose the right View for the task in hand. Techniques for making working with a spreadsheet easier – Freezing Panes, splitting windows, and using multiple windows.

Unit 2.3 – Spreadsheet structure

Changing the structure of your spreadsheet to make it clearer. Grouping rows and columns; changing cells size and shape.

Unit 2.4 – Cell references

How (and why) to connect (or link) cells and ranges, and how connected cells behave. Fixed or absolute references (dollar signs) for cells; partially fixed references. The "F4" shortcut.

Unit 2.5 – Named Ranges

Making formulas clearer and more easily understood by replacing cell references with meaningful names.

Unit 2.6 – Basic Macros

Using Macros to automate repetitive tasks in Excel. (This Unit covers recorded Macros – the later unit 5.8 covers writing and editing Macros using VBA in detail.)

Unit 2.7 – Design

Why spreadsheet design important in increasing efficiency, reducing errors and aiding transparency. Principles of good spreadsheet design.

Section 3: Administration

Unit 3.1 – Customising Excel

Changing Excel settings to suit the way you use Excel. Includes: font face, number of worksheets, frequency of autorecovery, default view, Autocorrect and Ribbon alterations. Excel Add-Ins – Solver, Analysis ToolPak. PowerPivot and Inquire.

Unit 3.2 – Housekeeping

Naming and storing Excel files. Version control. Making backups.

Unit 3.3 – Connecting Workbooks

- Setting up links between Workbooks. The dangers and pitfalls of doing so.
- Managing links between Workbooks and updating external references.
- Troubleshooting – finding ‘missing links’ and using auditing tools.

Unit 3.4 – Documentation

What to document, and how, in order to produce clear and robust spreadsheets.

Unit 3.5 – Protecting and Sharing

Protecting Workbooks and when it’s useful or necessary locking cells, Worksheet protection, Workbook protections, allowing users to edit ranges.

Permissions and granting access to folders & files, how to work with others on the same files and different files (both in process terms and using Excel functionality; when not to use it, avoiding conflicts). Sharing via the Cloud.

Unit 3.6 – Googledocs

What is Googledocs, and why use Google Spreadsheets for spreadsheeting?

Unit 3.7 – Excel troubleshooting

Common problems and how to avoid or overcome them. For example building in checks into spreadsheets, techniques for finding errors. Advanced techniques for diagnosing more complex problems.

Section 4: Data handling

Unit 4.1 – Sorting & Filtering

Techniques for sorting and filtering data, including controlling the order of precedence in a sort, advanced filters, and an introduction to PivotTables. Using sorting and filtering to check and ‘clean’ data.

Unit 4.2 – Controlling user input

Controlling the way users can enter data into a spreadsheet to reduce risk of error and increase efficiency. Covers Data Validation and using IS- functions to trap errors.

Unit 4.3 – Working with Dates and Times

Date and time formatting, performing basic maths with dates and times, and date and time functions.

Unit 4.4 – Working with Text

Entering and formatting text, extracting or combining parts of text. ‘Cleaning’ data for typos and bugs.

Unit 4.5 – Lookup & Reference

Looking up information in a basic table, and more flexibly. VLOOKUP() and the more flexible INDEX()/MATCH() combination. Advanced lookups.

Unit 4.6 – Logical Functions

Automatic decision making and error handling. Advanced logical tests. Matrix user of IF as an array function.

Section 5: Data analysis

Unit 5.1 – Working with Numbers

Maths: basic addition, subtraction, multiplication and division. How to use brackets to get the correct result. Powers. Rounding.

Unit 5.2 – Summarising data

Understanding a data set by summarizing it. Totals and counts – SUM(), COUNT() and COUNTA(). Other statistics. Conditional totals and counts. Advanced conditional sums, SUMPRODUCT and array formulas.

Unit 5.3 – PivotTables 1 – Simple Summaries

Uses of PivotTables. Summarising data (presenting database in terms of different subcategories). Understanding data (understanding categories, spotting typos, understanding range of values). Interpreting data (analysing to work out trends, areas of growth / profit / loss etc). Filtering.

Sorting and grouping

Unit 5.4 – PivotTables 2 – Manipulating Data

Changing the layout and format of data in a PivotTable. Changing Value Field Settings to show counted rather than 'summed' data. Setting up references to a PivotTable.

Unit 5.5 – PivotTables 3 – Interpreting Data

Using PivotTables to understand and interpret data. Includes Slicers and Timelines, PivotCharts and Conditional Formatting of PivotTables.

Unit 5.6 – PowerPivot – Handling Big Data

Why PowerPivot is different to a normal PivotTable: size, data model (inclusion in 2013 standard Excel), DAX, KPIs. PowerPivot data tools. Connecting to data, managing multiple tables and connections, including data from within the Excel spreadsheet

Unit 5.7 – Formula Auditing

Understanding and debugging spreadsheets. Tracing calculations and monitoring the effect of changes. Formula evaluator, Watch window and the INQUIRE Add-In.

Unit 5.8 – Advanced Macros and VBA

Troubleshooting macros. Writing code directly. User defined functions. Program structures. Using Macros to interact with users.

Unit 5.9 – Modelling

Principles. What modelling is; modelling broken down into distinct stages; important principles - simplification, built-in checks; including a presentation-ready sheet

Techniques. Calculating growth with data points over several years; extrapolating that into the future; scenario modelling and optimisation; prioritisation (RANK); build process.

Section 6: Presentation

Unit 6.1 – Cell Formatting

Basic font formatting, alignment – Including horizontal and vertical alignment, wrapping, merging, orientation, Alt-Return; consistent, sensible row heights & column

Styles – to make spreadsheets more and consistent. Formatting tools

Unit 6.2 – Number Formatting

Types of number (E.g., currency, %, decimal, negative numbers) – and Excel formats best employed. Custom number formats and how to maximise impact and clarity.

Unit 6.3 – Conditional Formatting

Changing the format of cells depending on their value. Graphical conditional formats. Writing conditional format formulas.

Unit 6.4 – Graphs and Charts

Creating simple charts and editing them to control and improve formatting. Choosing the right chart

Principles and guidelines for communicating well with charts. Sparklines and Maps (when and how to use).

Unit 6.5 – Page & Print Setup

Displaying spreadsheets as pages. Page layout view, page breaks, print area,. Printing from multiple Worksheets (and that default will be to print only from active Worksheet). Headers and footers. Adjusting page setup. Printing very large sheets of data

Tableau

Session - 1

- GETTING STARTED
- What is Tableau? What does the Tableau product suite comprise of? How Does Tableau Work?
- Tableau Architecture
- What is My Tableau Repository?

- Connecting to Data & Introduction to data source concepts
- Understanding the Tableau workspace
- Dimensions and Measures
- Data Types & Default Properties
- Tour of Shelves & Marks Card
- Using Show Me!
- Building basic views
- Saving and Sharing your work-overview
- BUILDING VIEWS (REPORTS) – BASICS
- Date Aggregations and Date parts
- Cross tab & Tabular charts
- Totals & Subtotals
- Bar Charts & Stacked Bars
- Line Graphs with Date & Without Date
- Tree maps
- Scatter Plots
- Individual Axes, Blended Axes, Dual Axes & Combination chart
- Edit axis
- Parts of Views
- Sorting
- Trend lines
- Reference Lines
- Forecasting
- Filters
- Context filters
- Sets
- In/Out Sets
- Combined Sets
- Grouping
- Bins/Histograms
- Drilling up/down – drill through
- Hierarchies
- View data

- Actions (across sheets)

Session - 2

- BUILDING VIEWS (REPORTS) – ADVANCED MAPS
- Explain latitude and longitude
- Default location/Edit locations
- Symbol Map & Filled Map
- Custom Geo Coding
- CALCULATED FIELDS
- Working with aggregate versus disaggregate data
- Explain - #Number of Rows
- Basic Functions (String, Date, Numbers etc)
- Usage of Logical conditions
- TABLE CALCULATIONS
- Explain scope and direction
- Percent of Total, Running / Cumulative calculations
- PARAMETERS
- Create What-If analysis
- Using Parameters in
- Calculated fields
- Bins
- Reference Lines
- Filters/Sets
- Display Options (Dynamic Dimension/Measure Selection)

Session - 3

- BUILDING INTERACTIVE DASHBOARDS- (BUILDING & CUSTOMIZING)
- Combining multiple visualizations into a dashboard (overview)
- Making your worksheet interactive by using actions
- Filter
- URL
- Highlight

- FORMATTING
- Options in Formatting your Visualization
- Working with Labels and Annotations
- Effective Use of Titles and Captions
- WORKING WITH DATA
- Multiple Table Join
- Data Blending
- Difference between joining and blending data, and when we should do each
- Working with the Data Engine / Extracts
- Working with Custom SQL
- Toggle between to Direct Connection and Extracts
- WORKING WITH TABLEAU SERVER – ACCESSING REPORTS THROUGH WEB
- Publishing to Tableau Server – Overview of publishing
- Server Administration - Managing Users, Projects & Object level and Data Security as per Users
- User Filters

Power BI

Power BI Desktop:

1. Introduction (2 hr)
 - Introduction to business intelligence
 - Introduction to data analysis
 - Introduction to data visualization
 - Overview of self-service BI
 - Considerations of self-service BI
 - Microsoft tools for self-service BI
 - Power BI vs. SSRS Report
 - Installation – Power BI Desktop
 - Configure tools

2. **Getting and Shaping Data** (3 hr)
 - Using databases as a Power BI data source

1. SQL Server Database
 2. Microsoft Access Database
- Using Excel as a Power BI data source
 - Using databases as a Power BI data source
 - The Power BI service
 - Import data into Power BI
 - Data Clean up
 - Transforming Data
 - Append Operation
 - Merge Operations
 - Lab :
 1. Import data into Power BI
 2. Data Clean up
 3. Transforming Data
 4. Append Operation
 5. Merge Operations

3. **M Language:** (2 hr)

The Power Query M formula language is optimized for building highly flexible data mashup queries. It's a functional, case sensitive language similar to F#.

Power Query statements are made up of individual language elements such as functions, variables, expressions, and primitive and structured values that together define the logic necessary to shape the data. In this module, we will share

- Quick tour of the Power Query M formula language
- Commonly used functions
 1. let expression
 2. RemoveColumns function
 3. PromoteHeaders function
 4. RenameColumns function
 5. SelectRows Function
 6. ReplaceValue function
 7. TransformColumns function
 8. Adding a Calculated column

4. **Power BI Model** (3 hr)

- Creating Relationships
 1. Append
 2. Merge
- Joins
 1. Left, Inner, Join, Full Outer Join, Left Anti Joins, Right Anti Join
- Extracting Columns From XML
- Setting Datatypes
- Transforming Columns
- Splitting Columns
- Defining New Columns
- Concatenating Columns
- Performing Lookup
- Translating
- Adding Date Table
- Lab:
 1. Creating Relationships
 2. Defining New Columns
 3. Concatenating Columns
 4. Performing Lookup
 5. Translating
 6. Adding Date Table

5. **Enhancing Power BI Model** (3 hrs)

- Defining Hierarchies
- Configuring Properties e.g. Data Categories , Formatting
- Creating Measures
- Creating Parameters
- Import from Reports
- Lab
 1. Creating Measures
 2. Creating Parameters
 3. Import from Reports

6. **Data Flows** (1 hr)

- What is Data Flow
- Why /When to use Data Flow?
- Lab:
 1. Implementing Data Flow

Here-in the data will be uploaded to Power BI Service, and different visualizations /reports will be created, and

7. Visualizing (3hrs)

- Visualizing Data
- Adding Visuals
- Filtering
- Slicer Filtering
- Highlighting
- Setting Properties
- Drilling within the Hierarchy
- Creating Bookmarks
- Power BI Mobile
- Custom Visuals

8. **Security** (1 hr)

- Data Storage Security
- Row Level Security
- Demo: Security

9. **Refreshing Data** (2 hrs)

- Publishing DataSets
- Refreshing Data
- Power BI Mobile Apps
- Using Power BI Gateway - Personal
- Understanding Power BI Gateway – Enterprise

- Power BI Pro vs. Free
- Publishing To Power BI
- Setting Up Auto-Refresh
- Content Packs
- Sharing Within the team /with organization
- Creating Dashboards
- Sharing Reports /Dashboards

DAX

1. **Introducing DAX** (1 hr)
 1. Understanding DAX Calculations
 2. Understanding Calculated Columns And Measures
 3. Variables
 4. Handling Errors In DAX Expressions
 5. Formatting DAX Code
 6. Common DAX Functions
2. **Using Basic Table Functions** (2 hrs)
 1. Introducing Table Functions
 2. Evaluate Syntax
 3. Using Table Expressions
 4. Understanding Filter
 5. Understanding All, Allexcept, And Allnoblankrow
 6. UNDERSTANDING VALUES AND DISTINCT
3. **Understanding Evaluation Contexts** (2 hrs)
 1. Introduction To Evaluation Contexts
 2. Testing Your Evaluation Context Understanding
 3. Creating A Row Context With Iterators
 4. Understanding Filter, All, And Context Interactions
 5. Working With Many Tables
 6. Evaluation Contexts Recap

7. Creating A Parameter Table

4. **Understanding CALCULATE And CALCULATETABLE** (2 hrs)

1. Understanding Calculate
2. Calculate Examples
3. Understanding Context Transition
4. Variables And Evaluation Contexts
5. Understanding Circular Dependencies
6. Calculate Rules
7. Introducing Allselected

5. **DAX Examples** (2 hrs)

1. Computing Ratios And Percentages
2. Computing Cumulative Totals

6. **Optimizing Data Models and Strategy** (1 hr)

1. Gathering Information About The Data Model
2. Denormalization
3. Columns Cardinality
4. Calculated Columns
5. Choosing The Right Columns To Store
6. Optimizing Column Storage

Python Programming

1. INTRODUCTION TO PYTHON: why learn python, feature of python, importance of learning python, application of python programming.

2. BASIC ELEMENTS OF PYTHON: keywords, datatypes, identifiers, operators, statements.

3. INSTALLATION OF PYTHON.

4. INPUT AND OUTPUT STATEMENTS IN PYTHON.

5. DATA TYPES IN PYTHON: int, float, strings, lists, tuples, dictionaries.

6. OPERATORS IN PYTHON: arithmetic, logical, assignment, relational, bitwise, membership operators.
7. TYPE CONVERSION IN PYTHON: implicit and explicit conversion in python.
8. BASIC PROGRAMMING IN PYTHON.
9. INTRODUCTION TO MATH MODULE IN PYTHON.
10. CONDITIONAL STATEMENTS IN PYTHON: simple if, simple if-else, multiple if-else, nested if-else.
11. LOOPING IN PYTHON: for loop, while loop in python, break and continue keywords in python, nested for loop in python.
12. FUNCTIONS IN PYTHON: declaration and definition of functions in python, passing arguments in functions, return keyword, function calling, arbitrary arguments, keywords arguments, default arguments in python.
13. STRINGS IN PYTHON: basics of strings, string functions.
14. LISTS IN PYTHON: introduction, list functions in python, list programs.
15. TUPLES IN PYTHON: basics, tuple functions.
16. SETS IN PYTHON: basics, set operations.
17. DICTIONARIES IN PYTHON: basics, functions in dictionaries, examples.
18. BUILT-IN METHODS IN PYTHON.
19. LAMBDA FUNCTIONS IN PYTHON.
20. FILE HANDLING IN PYTHON.
21. EXCEPTION HANDLING IN PYTHON.
22. PYTHON PROGRAM EXAMPLES.
23. OPERATOR PRECEDENCE IN PYTHON.
24. RECURSION IN PYTHON.
25. PROGRAMMING EXAMPLES IN PYTHON.
26. OBJECT ORIENTED PROGRAMMING IN PYTHON: class, objects, inheritance, polymorphism, terminologies, class attributes, object attributes, examples.
27. TIME MODULE IN PYTHON.
28. BOOLEAN IN PYTHON.
29. MISCELLANEOUS TOPICS.

R Programming

Fundamentals of R Language

- Introduction to R
- History of R
- Why R programming Language
- Comparison between R and Python
- Application of R

Setup of R Language

- Local Environment setup
- Installing R on Windows
- Installing R on Linux
- RStudio
- What is Rstudio?
- Installation of Rstudio
- First Program - Hello World

Variables and Data Types

- Variables in R
- Declaration of variable
- Variable assignment
- Finding variable
- Data types in R
- Data type conversion
- R programs for Variables and Data types in Rstudio

Input-Output Features in R

- scan() function

- readline() function
- paste() function
- paste0() function
- cat() function
- R Programs for implementing these functions in Rstudio

Operators in R

- Arithmetic Operators
- Relational Operators
- Logical Operators
- Assignment Operators
- Miscellaneous Operators
- R Programs to perform various operations using operators in Rstudio

Data Structure in R (part-I)

- What is data structure?
- Types of data structure
- Vector
 - What is a vector in R?
 - Creating a vector
 - Accessing element of vector
 - Some more operations on vectors
 - R Programs for vectors in Rstudio
- Application of Vector in R
- List
 - What is a list in R?
 - Creating a list
 - Accessing element of list
 - Modifying element of list

- Some more operations on list
- R Programs for list in Rstudio

Data Structure in R (part-II)

- Matrix or Matrices
 - What is matrix in R?
 - Creating a matrix
 - Accessing element of matrix
 - Modifying element of matrix
 - Matrix Operations
 - R Programs for matrices in Rstudio
- Application of Matrices in R
- Arrays
 - What are arrays in R?
 - Creating an array
 - Naming rows and columns
 - Accessing element of an array
 - Some more operations on arrays
 - R Programs for arrays in Rstudio

Data Structure in R (part-III)

- Data frame
 - What is a data frame in R?
 - Creating a data frame
 - Accessing element of data frame
 - Modifying element of data frame
 - Add the new element or component in data frame
 - Deleting element of data frame
 - Some more operations on data frame
 - R Programs for data frame in Rstudio

- Factors
- Factors in R
- Creating a factor
- Accessing element of factor
- Modifying element of factor
- R Programs for factors in Rstudio
- Application of Factors in R

Decision Making in R

- Introduction to Decision making
- Types of decision-making statements
- Introduction, syntax, flowchart and programs for
- if statement
- if...else statement
- if...else if...else statement
- switch statement

Loop control in R

- Introduction to loops in R
- Types of loops in R
- for loop
- while loop
- repeat loop
- nested loop
- break and next statement in R
- Introduction, syntax, flowchart and programs for
- for loop
- while loop
- repeat loop
- nested loop

Functions in R

- Introduction to function in R
- Built-in Function
- User-defined Function
- Creating a Function
- Function Components
- Calling a Function
- Recursive Function
- Various programs for functions in Rstudio

Strings in R

- Introduction to string in R
- Rules to write R Strings
- Concatenate two or more strings in R
- Find length of String in R
- Extract Substring from a String in R
- Changing the case i.e. Upper to lower case and lower to upper case
- Various programs for String in Rstudio

Packages in R

- Introduction to Packages in R
- Get the list of all the packages installed in Rstudio
- Installation of the packages
- How to use the packages in R
- Useful R Packages for Data Science
- R program for package in Rstudio

Data and file management in R

- Getting and Setting the Working Directory
- Input as CSV File
- Analysing the CSV File
- Writing into a CSV File
- R prog

Data Visualization in Python

Data Visualization in Python – Course Syllabus

Introduction to Data Visualisation

- What is data visualization
- Benefits of data visualization
- Importance of data visualization
- Top Python Libraries for Data Visualization

Matplotlib

- Introduction to Matplotlib
- Install Matplotlib with pip
- Basic Plotting with Matplotlib
- Plotting two or more lines on the same plot

Numpy and Pandas

- What is numpy?
- Why use numpy?
- Installation of numpy
- Example of numpy
- What is a panda?
- Key features of pandas
- Python Pandas - Environment Setup

- Pandas – Data Structure with example

Data Visualisation tools

- Bar chart
- Histogram
- Pie Chart

More Data Visualisation tools

- Scatter Plot
- Area Plot
- Stacked Area Plot
- Box Plot

Advanced data Visualisation tools

- Waffle Chart
- Word Cloud
- HEAT MAP

Specialized data Visualisation tools (Part-I)

- Bubble charts
- Contour plots
- Quiver Plot

Specialized data Visualisation tools (Part-II)

- Three-Dimensional Plotting in Matplotlib
- 3D Line Plot
- 3D Scatter Plot
- 3D Contour Plot
- 3D Wireframe Plot
- 3D Surface Plot

Seaborn

- Introduction to seaborn
- Seaborn Functionalities
- Installing seaborn
- Different categories of plot in Seaborn
- Some basic plots using seaborn

Data Visualisation using Seaborn

- Strip Plot
- Swarm Plot
- Plotting Bivariate Distribution
- Scatter plot, Hexbin plot, KDE, Regplot
- Visualizing Pairwise Relationship
- Box plot, Violin Plots, Point Plot

Project on Data Visualisation using Python

Data Visualization in R

Data Visualisation in R

- What is data visualization?
- Selecting right chart type
- Importance of data visualization & its benefits
- Applications of DATA Visualization
- R Programs for Scatterplot, Histogram, Bar & Stacked bar chart, boxplot, heatmap, line chart, density plot, pie chart

Data Visualisation with ggplot2 package

- What is ggplot2

- Plotting with ggplot2
- Building your plots iteratively
- Univariate distributions & bar plot
- Annotation with ggplot2
- Axis manipulation
- Density plot

More Data Visualisation tools in R

- Text mining and word cloud
- Radar chart
- Waffle chart
- Area Chart
- Correlogram

Project on Data Visualisation