

Updated on 12/04/2024

Sign up

## DuckDB training

2 days (14 hours)

### Presentation

DuckDB is an open source relational database designed for fast, efficient data analysis. Unlike traditional systems, DuckDB stands out for its ability to handle large volumes of data while offering high performance and compatibility with standard SQL.

This solution is renowned for its speed, reliability and advanced features, making it an ideal choice for applications requiring real-time data analysis or the processing of large datasets.

Our DuckDB training course will immerse you in the world of DuckDB, enabling you to discover its features, master its use and explore its advanced capabilities.

Our program will give you advanced skills in SQL with DuckDB, including joins, data aggregation and analytical functions.

As with all our training courses, we'll be using the latest stable version of the program and the latest resources: [DuckDB 0.10](#).

### Objectives

- Understanding MariaDB SGDB
- Installing and configuring MariaDB
- Know how to administer databases
- Managing MariaDB security
- Set up backup, restore and maintenance operations

### Target audience

- Data analysts
- Developers
- Database administrators
- Data science professionals

## Prerequisites

- Knowledge of SQL
- Operating system knowledge
- Familiarity with a programming language, preferably Python
- Notions in relational databases

## OUR DUCKDB TRAINING PROGRAM

### INTRODUCTION TO DUCKDB

- Introducing DuckDB: origin, objectives and key features
- DuckDB benefits and typical use cases
- Comparing DuckDB with other database systems
- Exploring the growing DuckDB ecosystem and community
- Discussion of the importance of DuckDB in the processing of massive data

### INSTALLATION AND CONFIGURATION

- Detailed instructions for installing DuckDB on different operating systems
- Optimum DuckDB configuration for best performance
- Checking installation and solving common problems
- Introducing the DuckDB working environment
- Introduction to DuckDB development tools and interfaces

### SQL FUNDAMENTALS WITH DUCKDB

- Basic SQL syntax and query creation
- Create, modify and manage tables and data schemas
- Data selection, filtering and sorting techniques
- Updating and deleting data in DuckDB
- Best practices for writing clean, efficient SQL code

### ADVANCED DUCKDB FEATURES

- Use of complex joins and subqueries
- Data aggregation and use of GROUP BY and aggregation functions

- Introduction to analytical windows and functions
- Indexing techniques and query optimization for better performance
- Strategies for managing and analyzing large volumes of data

## DUCKDB INTEGRATION WITH PYTHON

- Introducing the Python interface for DuckDB
- Establishing a connection between a Python script and DuckDB
- Executing SQL queries via Python and manipulating results
- DuckDB integration into Python data analysis workflows
- Practical examples of data manipulation between DuckDB and Pandas

## MAINTENANCE, SAFETY AND OPTIMIZATION

- Introduction to transactions and sustainability in DuckDB
- Security mechanisms and access authorization management
- Backup and restore techniques for DuckDB databases
- Performance monitoring and troubleshooting
- Advanced use of statistics for query optimization and load management

## CONCLUSION AND COMMUNITY RESOURCES

- Summary of key points covered during the course
- Discussion of DuckDB's involvement in community development
- Presentation of additional resources and advanced documentation
- Tips for continuing to learn and improve your DuckDB skills
- Question-and-answer session to clarify doubts and plan next steps

## Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

## Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

## Teaching methods

Practical training: 60% Practical, 40% Theory. Training material distributed in

to all participants.

## Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

## Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

## Sanction

A certificate will be issued to each trainee who completes the course.