Training 4 Programmers

PostgreSQL for Python Programmers

Duration

2 days

Description

This comprehensive course is designed for computer programming professionals seeking to master PostgreSQL, a powerful open-source object-relational database system. The course covers a wide range of topics, from basic SQL syntax to advanced database design and performance tuning. Participants will also learn how to use Python for programming with PostgreSQL, handle transactions, and work with large datasets using Pandas. The course also delves into extending PostgreSQL using procedural languages and popular extensions. With a mix of theoretical concepts and practical exercises, this course equips participants with the skills to effectively use PostgreSQL in their professional roles.

Objectives

- Understand the key features, benefits, and use cases of PostgreSQL in the industry.
- Learn how to install and configure PostgreSQL on different operating systems.
- Master basic and advanced SQL techniques using PostgreSQL.
- Gain knowledge in database design, modeling, and performance tuning.
- Grasp the concepts of concurrency, transactions, and security in PostgreSQL.
- Learn how to use Python for programming with PostgreSQL and handling transactions.
- Develop skills to load and write data from PostgreSQL into Pandas efficiently.
- Discover how to extend PostgreSQL using procedural languages and popular extensions.

Prerequisites

Experience wiht Python programming is required. No prior experience with PostgreSQL is necessary. Students should have a basic understanding of SQL and relational databases.

Training Materials

All students receive comprehensive courseware covering all topics in the course. Courseware is distributed via GitHub in the form of documentation and extensive code samples. Students practice the topics covered through challenging hands-on lab exercises.

Software Requirements

Students will need a free, personal GitHub account to access the courseware. Student will need permission to install Docker Desktop, Visual Studio Code, and Visual Studio Code Extensions on their computers. If students are unable to configure a local environment, a cloud-based environment can be provided.

Outline

- Introduction to PostgreSQL for Python Programming
    - Overview of PostgreSQL
    - Key Features and Benefits
    - Use Cases in the Industry
- Setting Up PostgreSQL
    - Installing PostgreSQL on Different Operating Systems
    - Basic Configuration and Post-Installation Setup
- PostgreSQL Architecture
    - System Architecture and Process Model
    - Data Storage and Tablespaces
- Basic SQL with PostgreSQL
    - Basic SQL Syntax and Data Types
    - Creating, Altering, and Dropping Tables
    - CRUD Operations (Create, Read, Update, Delete)
- Advanced SQL Techniques
    - Complex Queries: Joins, Subqueries, and Common Table Expressions (CTEs)
    - Functions and Operators: Built-in and User-Defined
    - Aggregate Functions and Grouping
- Database Design and Modeling
    - Normalization and Denormalization
    - Entity-Relationship Modeling and Schema Design
- Performance Tuning and Optimization
    - Using EXPLAIN and ANALYZE for Query Optimization
    - Indexing Strategies and Index Maintenance
    - Performance Tips and Configuration Tuning
- Concurrency and Transactions
    - Transactions, ACID Properties, and Isolation Levels
    - Locking Mechanisms and Concurrency Control
    - Multi-Version Concurrency Control (MVCC)

- Security and Authentication
    - User Management and Role-Based Access Control
    - Authentication Methods and Configuring pg_hba.conf
    - Data Encryption Techniques and SSL/TLS Configuration
- Python Programming with PostgreSQL
    - Setting Up Psycopg2
    - Connecting to PostgreSQL from Python
    - Performing CRUD Operations with Python
    - Using SQLAlchemy with PostgreSQL
    - Handling Transactions in Python
    - Working with JSON Data
    - Working with Vectors
- Loading Data from PostgreSQL into Pandas
    - Writing SQL Queries to Retrieve Data
    - Using pandas.read_sql() to Load Data
    - Handling Large Datasets Efficiently
- Writing Data from Pandas to PostgreSQL
    - Creating New Tables from DataFrames
    - Inserting and Updating Data in Existing Tables
    - Using pandas.to_sql() for Efficient Data Transfers
- Handling Data Types and Indexes
    - Mapping Data Types Between Pandas and PostgreSQL
    - Ensuring Data Integrity and Consistency
    - Managing Indexes for Optimized Performance
- Extending PostgreSQL
    - Using Procedural Languages: PL/pgSQL, PL/Python, PL/Perl
    - Working with Extensions and Modules
    - Popular Extensions (PostGIS, hstore, etc.)
- Conclusion
    - Summary of Key Concepts
    - Q&A
    - Further Resources and Next Steps