Secure Coding for Databricks – SQL & Python

Overview:
This hands-on training equips data teams with best practices to develop secure Databricks workflows by addressing critical risks in SQL, Python, and platform configurations. Participants will learn to mitigate injection attacks, protect sensitive data, enforce least-privilege access, and implement auditing—while aligning with compliance standards (GDPR, CCPA).

Key Outcomes:
✓ Build injection-resistant SQL queries and Python UDFs.
✓ Implement secrets management and encryption in Databricks.
✓ Configure Unity Catalog for row/column-level security.
✓ Harden clusters and networks against misconfigurations.
✓ Detect vulnerabilities in notebooks and dependencies.

Duration : 2 days

Training Outline

1. Introduction to Databricks Security

- Shared Responsibility Model
- Databricks Security Architecture
- Top Security Risks in Data Engineering
- Security Best Practices Lifecycle

2. Platform Security Fundamentals

- **Workspace Security**
    - Access Controls (Users/Service Principals)
    - Cluster Policies & ACLs
    - Notebook/Job Permissions
- **Unity Catalog Security**
    - Data Governance Framework
    - Object-Level Permissions (GRANT/REVOKE)
    - Row/Column-Level Masking

3. Secure SQL Development

- SQL Injection Prevention (Parameterized Queries, Dynamic SQL Safeguards)
- Data Exposure Controls (Column Encryption, Dynamic Views)
- Least-Privilege Query Design

4. Secure Python Development

- Secrets Management (Databricks Secrets API, Environment Variables)

- Input Validation & Sanitization (UDFs, Notebook Parameters)
- Safe DataFrame Operations (PII Filtering, Encryption)

5. Infrastructure Hardening

- Cluster Security (Isolation, Library Whitelisting, Init Scripts)
- Network Protection (VPC/PrivateLink, Firewall Rules)

6. Data Protection Techniques

- Encryption (TLS, Managed Keys)
- Anonymization (Pseudonymization, Differential Privacy)

7. Vulnerability Management

- Dependency Scanning (OSS Tools like Snyk)
- Code Scanning (SAST for Notebooks, Secrets Detection)

8. Audit & Compliance

- Monitoring (Audit Logs, Anomaly Detection)
- Compliance (GDPR/CCPA, Unity Catalog Tags)

9. Anti-Patterns & Case Studies

- Common Pitfalls (Hardcoded Credentials, Over-Permissioned Clusters)
- Real-World Breach Examples

10. Hands-On Labs

- SQL Injection Mitigation
- Secrets Rotation
- Row-Level Security Setup
- Vulnerability Scanning Demo