

Master's Program Curriculum

Data Engineer & AI Training

The definitive Master's programme taking you from foundational Python and Data Engineering to building advanced Generative AI and Autonomous Agent systems. Become the AI-native professional the industry is desperately seeking.

Section 1

Python for AI & Data

Section 2

SQL for AI & Data

Section 3

Power BI for Data Analysis

Section 4

Data Engineering with MS Fabric

Section 5

Generative AI & Agentic AI

Digital Edify

India's First AI-Native Training Institute

Learn AI. Build Agents. Lead Future.

About Digital Edify

India's #1 Training Institute for the AI Era

Established: 2016

Headquarters: Hyderabad, Telangana

Reach: Global (Online + Offline)

The Transformation Narrative

Digital Edify has evolved from a premium training institute in the Automation Era to an AI-first organisation leading the Agentic AI revolution. Since 2016, we've transformed over 100,000 professionals and built partnerships with more than 1,000 industry leaders. Our journey reflects the technological evolution of our time—from traditional job placement to career transformation, and now to building AI-native professionals who will shape the future of work.



Automation Era (2016-2023)

Premium Training Institute focused on job placement with 100K+ students trained

AI Revolution (2024-2025)

AI-Powered Training with industry-AI integration and career transformation focus

Agentic AI Leadership (2026+)

AI First Institute building AI-Native Professionals with 1 Million AI-Native Vision

"We started in the Automation Era. We evolved through the AI Revolution. Now, we're leading the Agentic AI Future—with 100,000+ professionals already transformed and 1,000+ industry partners trusting our graduates."

Vision & Mission

Vision

"To Create 1 Million AI-Native Professionals Who Will Build the Agentic Future of Work"

Mission

"We transform learners into AI-native professionals through industry-aligned programmes that integrate Agentic AI into every discipline—from development to data science to enterprise platforms."

Course Highlights

Section 1: Python for AI & Data

Build a strong Python foundation for data handling, automation, and AI-ready programming.

Section 2: SQL for AI & Data

Learn to design, query, and optimize databases to support analytics, AI workflows, and data-driven applications.

Section 3: Power BI for Data Analysis

Transform raw data into interactive dashboards and business insights using Power BI and DAX.

Section 4: Data Engineering with Microsoft Fabric

Design scalable, end-to-end data engineering pipelines using Microsoft Fabric, OneLake, Spark, and real-time analytics.

Section 5: Generative AI & Agentic AI

Build intelligent AI systems using LLMs, RAG pipelines, and autonomous agent-based architectures.

Python for AI & Data

Module 1: Python Fundamentals

Core Concepts

- Python Interpreter Installation (Windows/Mac)
- IDE Setup with Visual Studio Code
- Python Syntax & 35 Keywords
- Identifiers & Naming Conventions
- Variables & Memory Management
- Data Types (Simple & Complex)

Operations & Control

- Type Conversion & Type Casting
- Operators (Arithmetic, Comparison, Logical)
- Conditional Statements (if, elif, else, match-case)
- Loops (while, for) & range() function
- Control Flow (break, continue, pass)
- User Input with input() function

Module 2: String Manipulation

Master comprehensive string operations essential for data processing and text analysis in AI applications. Learn indexing, slicing, formatting, and advanced manipulation techniques.



String Basics

Definition, rules, indexing (positive & negative), slicing with start:end:step syntax, and immutability concepts.



Formatting & Operations

String formatting with f-strings and format(), concatenation, repetition, and case conversion methods.



Search & Validation

Search methods (find, index, count), checking methods (isalpha, isdigit), and trimming operations.



Advanced Methods

Replacement techniques, split/join operations, and string alignment methods for professional formatting.

Module 3: Data Structures

Lists & Tuples

Lists: Dynamic Collections

Master Python's most versatile data structure with comprehensive operations for data manipulation.

- Creation, indexing, and slicing techniques
- Adding elements: append, insert, extend
- Removing elements: remove, pop, clear
- Searching and counting: index, count
- Sorting and reversing operations
- Powerful list comprehensions

There was an error generating this image

Tuples: Immutable Sequences

- Creation and operations
- Understanding immutability benefits
- Tuple packing and unpacking
- Performance comparisons with lists

Module 4: Data Structures

Dictionaries & Sets

1 Dictionaries

Key-value pair mastery for efficient data storage and retrieval.

- Creation and access patterns
- Operations and methods
- Keys, values, items methods
- Dictionary comprehensions
- Nested dictionary structures

2 Sets

Unique, unordered collections with powerful mathematical operations.

- Creation and UUU properties
- Union, intersection, difference
- Subset and superset checks
- Frozen sets for immutability
- Practical applications



Module 5: Advanced Collections

Iterators & Generators

Unlock memory-efficient data processing with advanced Python collections, custom iterators, and generator functions essential for handling large datasets in AI applications.



Collections Module

namedtuple, Counter, defaultdict, deque for specialized data handling

Iterators

Iteration protocol, custom iterators, and memory efficiency concepts

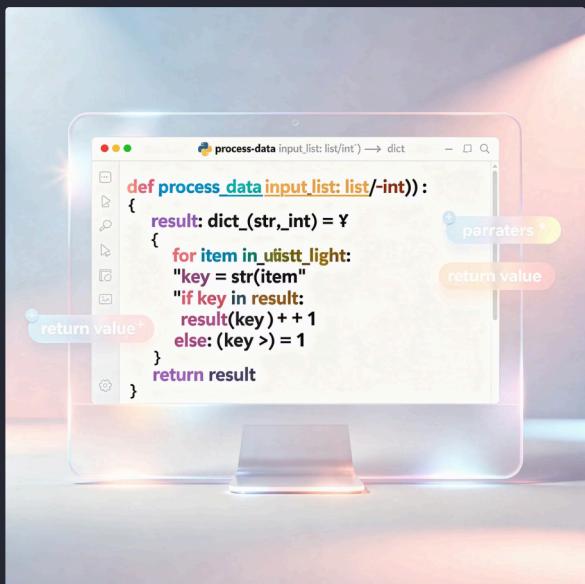
Generators

yield statement, generator expressions, and pipeline patterns

Functional Programming

Lambda functions, map, filter, reduce for elegant data transformations

Module 6: Functions & Scope



Function Fundamentals

- Function definition and calling
- Parameters and arguments
- Positional and keyword arguments
- Default arguments
- Arbitrary positional arguments (*args)
- Arbitrary keyword arguments (**kwargs)
- Return statements and multiple values

1 Scope Management

Local and global scope, global keyword usage, and variable visibility

2 Function Types

Built-in, user-defined, lambda functions (IIFE), and recursive functions

3 Documentation

Docstrings and best practices for professional code documentation

Module 7: Modules & Packages

Master Python's modular architecture for building scalable, maintainable applications. Learn to leverage built-in modules, create custom packages, and manage external dependencies efficiently.

01

Module Types

Built-in, user-defined, and external modules with importing techniques

02

Common Built-ins

math, random, datetime, os, sys for essential operations

03

Package Structure

Creating packages, __init__.py purpose, and nested packages

04

Dependency Management

pip package manager, installing external packages, requirements.txt

05

Popular Packages

requests, pandas, numpy, and module best practices

Module 8: Working with Data Formats

File Operations & Data Serialization

File Operations

- CRUD operations basics
- open() function and file modes
- Reading: read, readline, readlines
- Writing: write, writelines
- Append mode operations
- File path operations
- Directory management (os, shutil)

Structured Data Formats

- CSV file operations
- csv.reader and csv.writer
- csv.DictReader and csv.DictWriter
- JSON file operations
- JSON methods: dump, dumps, load, loads
- Data serialization and deserialization
- File handling best practices

Module 9: Advanced Python Concepts



Exception Handling

try-except-else-finally blocks, catching specific exceptions, raising and re-raising exceptions, custom exception classes, and built-in exception types.



Decorators

Function decorators, decorators with arguments, multiple decorators, class decorators, and practical applications for code enhancement.



Generators Deep Dive

Generator expressions, infinite generators, memory efficiency, and advanced iteration patterns for data processing.



Context Managers

Built-in context managers, custom context manager creation, and resource management best practices.

Module 10: Object-Oriented Programming

Master OOP fundamentals and the four pillars essential for building robust, scalable AI and data applications. Learn class design, inheritance patterns, and advanced Python OOP features.

1 OOP Fundamentals

Classes, objects, attributes (instance & class variables), `__init__` constructor, understanding `self`, instance/class/static methods.

2 Encapsulation

Access modifiers: public, protected, private attributes and methods for data hiding and security.

3 Inheritance

Single, multi-level, and multiple inheritance patterns. Method overriding and `super()` function usage.

4 Abstraction

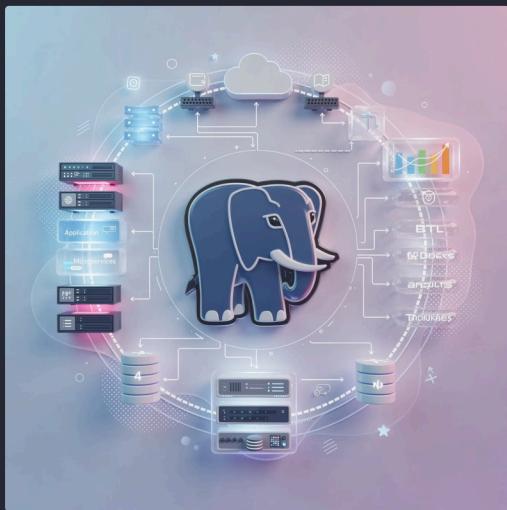
Abstract classes, abstract methods, ABC module for defining interfaces and contracts.

5 Polymorphism

Method overriding, duck typing, special/magic methods (`__str__`, `__repr__`, `__len__`), and real-world applications.

SQL for AI & Data

Module 1: Foundations of Databases & PostgreSQL



PostgreSQL Tools & Objects

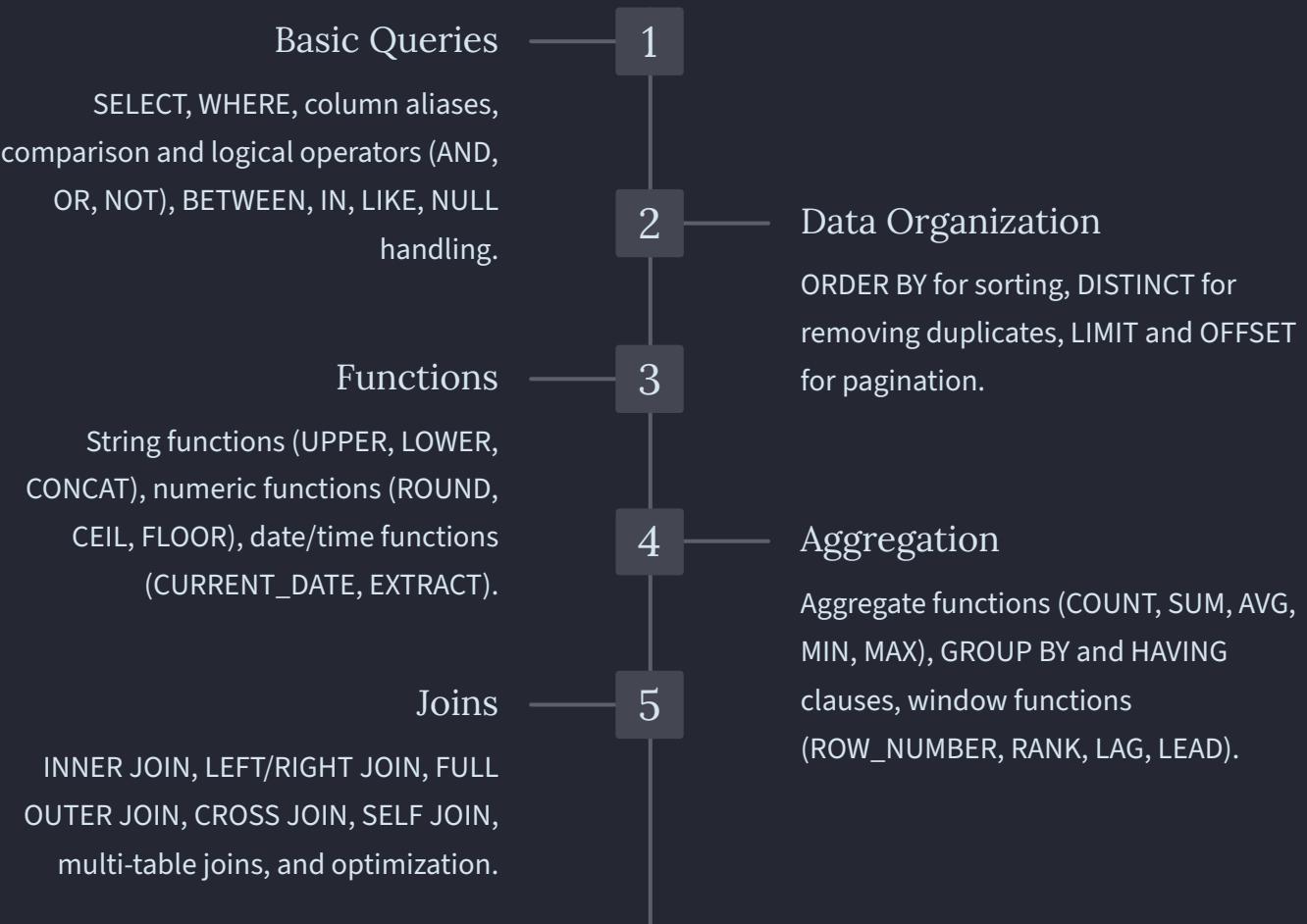
- psql command-line interface
- pgAdmin 4 graphical interface
- Database objects: databases, schemas, tables
- Data types: numeric, character, date/time, Boolean, special types
- Constraints: PRIMARY KEY, FOREIGN KEY, UNIQUE, NOT NULL, CHECK, DEFAULT
- Creating databases and tables
- INSERT operations and data population
- Referential integrity management

Database Fundamentals

- Introduction to DBMS and RDBMS
- ACID Properties (Atomicity, Consistency, Isolation, Durability)
- PostgreSQL installation and setup

Module 2: Querying and Analyzing Data

Master comprehensive SQL querying techniques from basic SELECT statements to advanced window functions and multi-table joins for sophisticated data analysis.





Module 3: Advanced Queries

Data Manipulation & Transactions

Subqueries & CTEs

- Subqueries in WHERE, SELECT, FROM clauses
- Correlated subqueries
- EXISTS and NOT EXISTS
- IN and NOT IN with subqueries
- Common Table Expressions (CTEs)
- Recursive CTEs for hierarchical data
- Multiple CTEs

Set Operations

- UNION and UNION ALL
- INTERSECT
- EXCEPT

Data Modification

- UPDATE statements with expressions
- UPDATE with JOIN
- DELETE statements with subqueries
- TRUNCATE vs DELETE

Transaction Management

- BEGIN, COMMIT, ROLLBACK
- Savepoints
- Transaction isolation levels
- Concurrency control

Module 4: Database Programming

Automation & Advanced Features

1 Schema Management

ALTER TABLE operations, adding/modifying/dropping columns, managing constraints, indexes and performance optimization.

2 Indexes

Index types (B-tree, Hash, GIN, GiST), creating and managing indexes, performance considerations, when to use indexes.

3 Views

Creating views for abstraction, updatable views, materialized views, refreshing materialized views.

4 Stored Functions

PL/pgSQL programming language, function parameters and return types, control structures (IF, CASE, LOOP), functions returning tables.

5 Procedures & Triggers

Stored procedures, procedure vs function differences, exception handling, BEFORE/AFTER/INSTEAD OF triggers, audit logging, data validation.

Module 5: Database Design

Optimization & Best Practices

Master database design principles, normalization techniques, and performance optimization strategies for building efficient, scalable data systems.



ER Modeling

Entities, attributes, relationships, ER diagrams, relationship types (1:1, 1:M, M:N)



Normalization

1NF, 2NF, 3NF principles, benefits and trade-offs, when to denormalize



Design Best Practices

Naming conventions, data type selection, primary/foreign key strategies



Query Optimization

EXPLAIN and EXPLAIN ANALYZE, execution plans, index strategies, query rewriting



Performance Tuning

Database statistics (ANALYZE), VACUUM maintenance, connection pooling, table partitioning

Module 1: Power BI for Data Analysis

1 BI Fundamentals

Business intelligence concepts, Power BI architecture, Desktop vs Service capabilities, interface navigation.

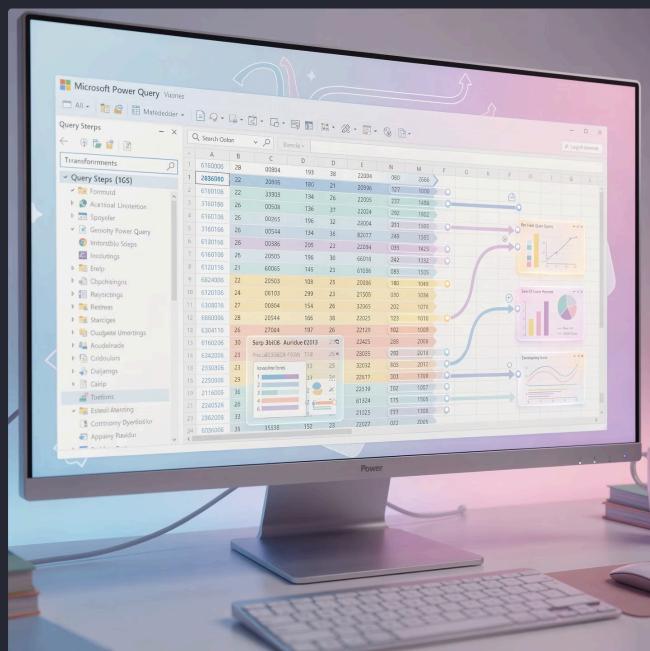
2 Data Connectivity

File, database, cloud, and web sources. Import vs DirectQuery vs Live Connection modes, credential management.

Data Transformation & Modeling

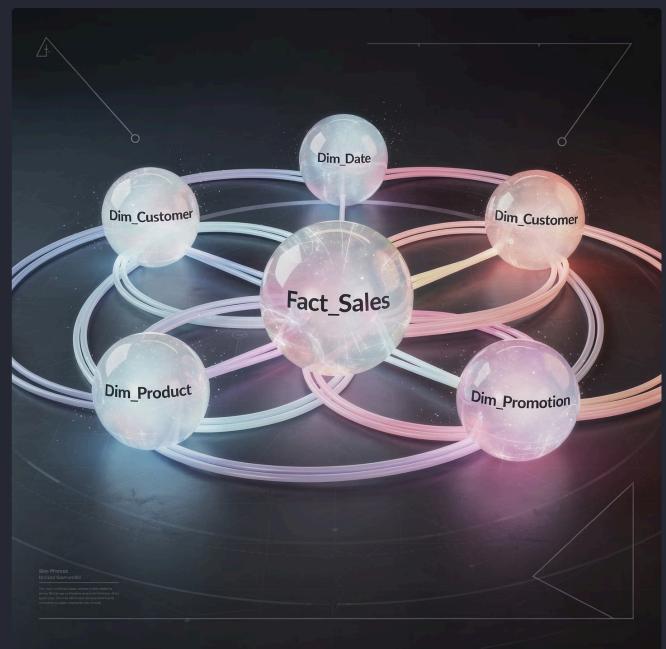
Module 3: Power Query

- Power Query interface and applied steps
- Data profiling and quality assessment
- Essential transformations: filtering, splitting, merging
- Reshaping: pivot, unpivot, grouping
- Combining queries: append and merge operations



Module 4: Data Modeling

- Star schema vs snowflake schema design
- Creating and managing table relationships
- Primary and foreign keys
- Hierarchies and date dimension tables
- Data model optimization strategies



Visualizations & DAX Fundamentals

Module 5: Building Visual Reports & Dashboards



Core Visualizations

Charts, tables, maps, KPIs, and selecting the right visualization for your data.



Interactivity

Slicers, filters, bookmarks, drill-through for dynamic user experiences.



Layout & Design

Dashboard layout principles, mobile optimization, and data storytelling techniques.

Module 6: Introduction to DAX

1

2

3

DAX Basics

Syntax, structure, calculated columns vs measures

Essential Functions

Aggregation, logical, text, date/time functions

Advanced DAX

CALCULATE, FILTER functions, creating KPIs

Advanced Analytics & Custom Visuals

Module 7: Time Intelligence & Advanced DAX

1

Time Intelligence

YTD, MTD, QTD functions, prior period comparisons, growth rates

2

Context Mastery

Filter context vs row context, variables, iterator functions

3

Performance

DAX optimization techniques for faster calculations

Module 8: Advanced Visualizations & Custom Analytics

- Custom visuals from AppSource
- Advanced chart types: waterfall, funnel, decomposition tree
- R and Python integration for advanced analytics
- AI visuals: Key Influencers, Q&A, Smart Narratives
- Dynamic visuals with parameters

Enterprise Deployment & Governance

Module 9: Publishing, Sharing & Collaboration

01

Publishing

Workspace management, dashboards vs reports

02

Data Refresh

Scheduled refresh, gateway configuration

03

Sharing

Sharing strategies, Power BI apps

04

Integration

Teams, SharePoint, Excel, PowerPoint integration

Module 10: Enterprise Deployment, Governance & Administration

- Power BI admin portal and tenant settings
- Row-Level Security (RLS) and Object-Level Security (OLS)
- Incremental refresh and aggregations
- Dataflows and deployment pipelines
- Performance optimization and capacity management
- Enterprise licensing models
- APIs and embedded analytics

Data Engineering with MS Fabric

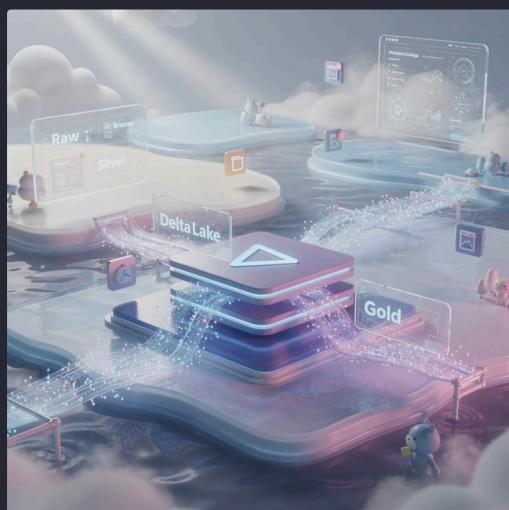
Module 1: Introduction to Microsoft Fabric



Licensing spans F2-F2048 capacity units. Migration paths available from Azure Synapse, Databricks, and Snowflake.

Module 2: OneLake

The Unified Data Lake



OneLake Fundamentals

OneLake is Fabric's unified data lake built on Azure Data Lake Storage Gen2, providing a single storage layer for all workloads.

- OneLake architecture and concepts
- Delta Lake and Parquet file formats
- ACID transactions and versioning
- OneLake shortcuts to external data sources
- OneLake Catalog and data discovery
- Security and access control

Module 3: Lakehouse Architecture

Master the Lakehouse paradigm that combines the best of data lakes and data warehouses, enabling both structured and unstructured data analytics with ACID guarantees.



1 Lakehouse Components

Creating and managing Lakehouses, SQL Analytics Endpoint, Delta table operations and optimization.

2 Time Travel

Query historical versions of data, rollback changes, audit data lineage with Delta Lake versioning.

Module 4: Data Factory

Data Integration & Orchestration



Data Pipelines

Creation, configuration, orchestration, and error handling for automated data workflows.



Dataflows Gen2

Power Query integration, M language fundamentals for data transformation at scale.



Database Mirroring

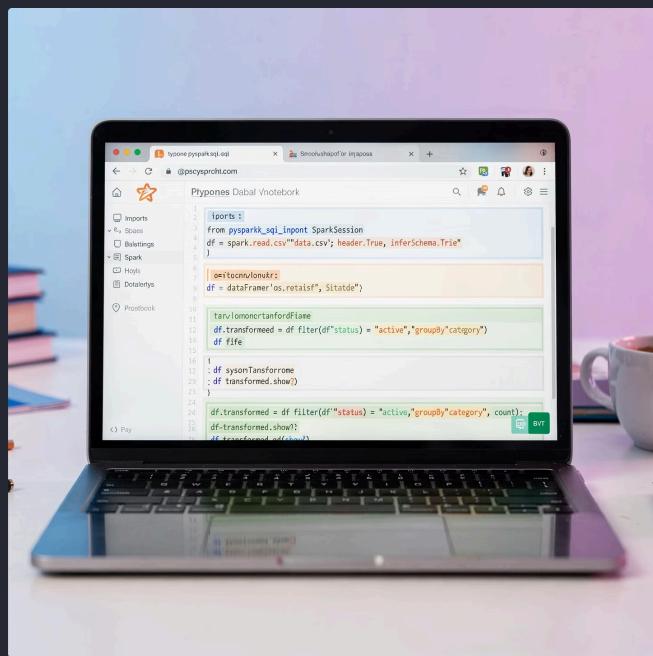
Real-time replication from Azure SQL, Cosmos DB, and PostgreSQL into Fabric.

Comprehensive connector library, CI/CD support for enterprise deployment, and extensive capabilities for data integration across cloud and on-premises sources.

Module 5: Data Engineering with Apache Spark

Spark in Fabric

- Apache Spark integration and architecture
- Fabric Notebooks with Copilot assistance
- PySpark DataFrames and transformations
- Spark SQL queries and optimization
- Spark job definitions and scheduling



AI Functions

Leverage built-in AI capabilities directly within Spark for intelligent data processing:

- Text summarization for document processing
- Classification for categorizing data
- PII obfuscation for data privacy
- Sentiment analysis
- Entity extraction

Module 6: Data Warehouse

SQL Analytics at Scale

Build enterprise-grade data warehouses with comprehensive T-SQL support, dimensional modeling, and performance optimization for analytical workloads.

1 Warehouse Fundamentals

Fabric Data Warehouse overview, T-SQL support, user-defined functions, schema design and table management.

2 Dimensional Modeling

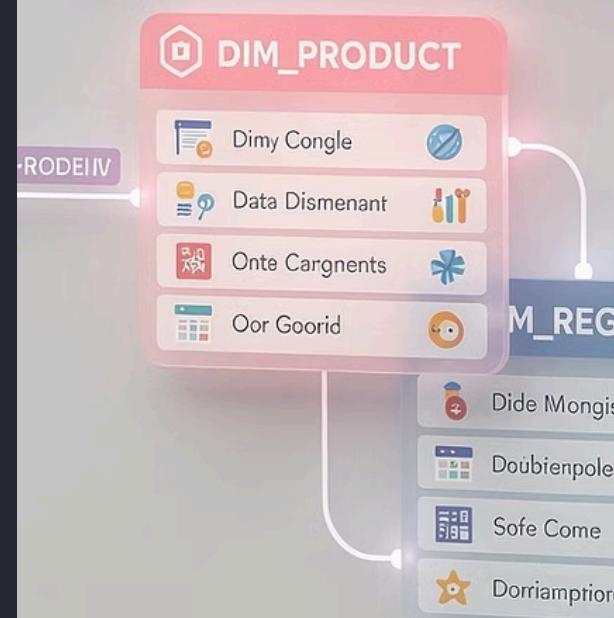
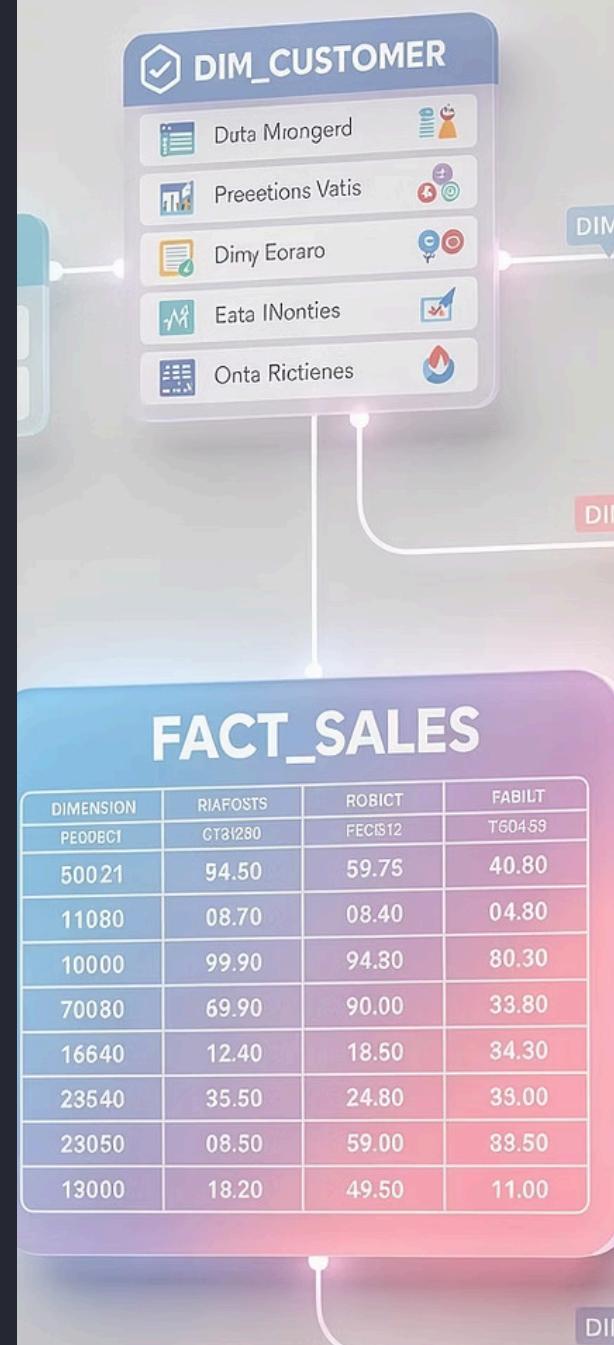
Star schema design, fact and dimension tables, slowly changing dimensions (SCD) implementation.

3 SQL Database

SQL Database in Fabric for transactional workloads, integration with warehouse for hybrid scenarios.

4 Performance Optimization

Query tuning, indexing strategies, partitioning, and materialized views for fast analytics.



Module 7: Real-Time Intelligence



Module 8: Power BI & Semantic Models

Integrated Analytics & Reporting

- Power BI integration with Fabric
- Direct Lake mode for zero-copy analytics
- Storage modes and performance
- Semantic models and relationships
- DAX fundamentals and syntax
- Row-level security (RLS) and object-level security (OLS)
- Incremental refresh and aggregations
- Report development with Copilot
- Embedding and sharing strategies

Module 9: Data Science in Fabric

ML Model Development & Deployment

Build, train, and deploy machine learning models within Fabric's integrated data science experience, leveraging MLflow for experiment tracking and model versioning.

Exploratory Analysis

Data exploration, visualization, and statistical analysis in notebooks.

Model Training

ML model development, hyperparameter tuning, and cross-validation.

Experiment Tracking

MLflow integration for versioning, metrics, and model registry.

Semantic Link

SemPy for connecting Power BI semantic models with Python.

Deployment

Batch scoring, real-time predictions, and model monitoring.

Module 10: AI & Copilot in Fabric

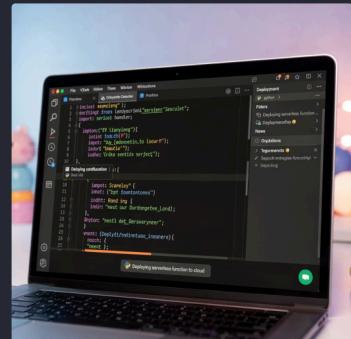
 Copilot Integration	 Data Agents
AI assistance across all workloads: Notebooks, SQL, KQL, Pipelines, and Reports for accelerated development.	Conversational AI for natural language queries, data exploration, and automated insights generation.
 Operations Agents	 Fabric IQ
Intelligent monitoring, anomaly detection, and automated troubleshooting for operational excellence.	Ontology models for semantic understanding, AI Functions integration, and Azure AI Foundry connectivity.

Module 11: User Data Functions

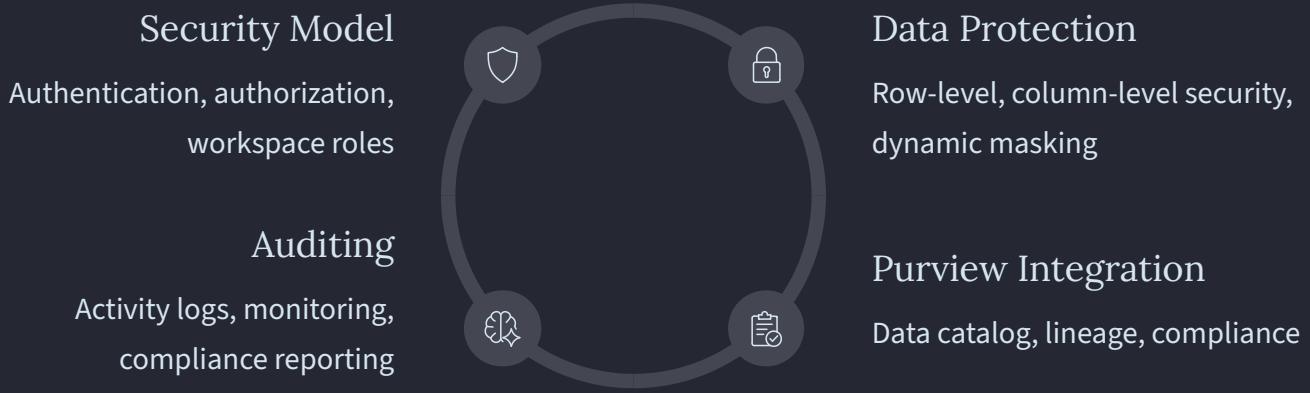
Serverless Python Functions

Function Development

- User Data Functions overview and architecture
- Python-based serverless function development
- VS Code extension for local development
- Integration with Notebooks, Pipelines, and SQL
- Testing, debugging, and deployment workflows



Module 12: Security & Governance





Module 13:

Administration &

Monitoring

Enterprise Management

1 Admin Portal

Fabric admin portal, tenant settings, capacity management, SKU selection (F2-F2048)

2 Monitoring Hub

Performance dashboards, query monitoring, pipeline tracking, resource utilization

3 DevOps

Git integration, deployment pipelines, CI/CD patterns for automated deployments

Module 14: Advanced Topics

Best Practices & Enterprise Patterns

Performance Optimization

- Query optimization techniques
- Partition strategies for large datasets
- Caching mechanisms
- Resource allocation and tuning

Architecture Patterns

- Enterprise architecture design
- Data mesh implementation
- Multi-tenant strategies

Migration & Integration

- Migration strategies from legacy systems
- Azure service integration
- Third-party connector development

Developer Tools

- Fabric CLI for automation
- REST APIs and extensibility
- Custom connector development





Module 15: Certification Preparation

DP-600 & DP-700 Exam Readiness

1

DP-600 Structure

Exam format, skills measured, domain breakdown, and scoring methodology.

2

DP-700 Focus

Fabric Data Engineer certification path, specialized topics, and advanced scenarios.

3

Practice Tests

Mock exams, question strategies, time management, and performance analysis.

4

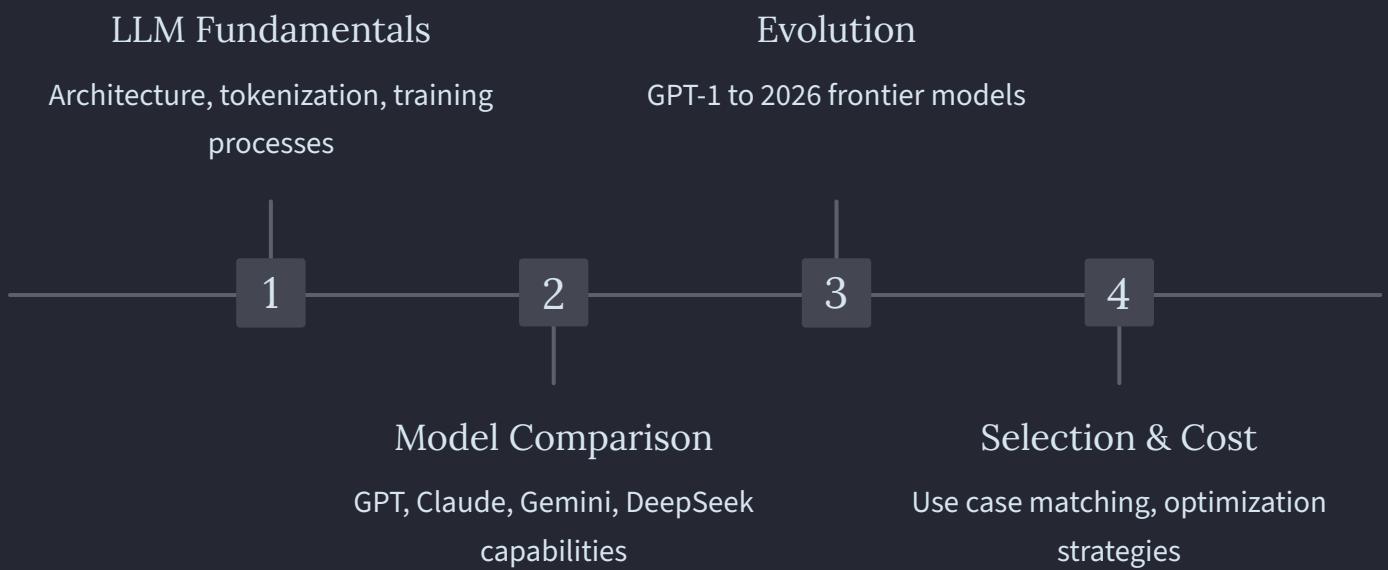
Hands-on Labs

Practical exercises aligned with exam objectives, real-world scenarios, and skill validation.

Generative AI & Agentic AI

Module 1: Foundations of Generative AI

Master the fundamentals of Large Language Models, transformer architecture, and the evolving landscape of frontier AI models from GPT to DeepSeek.



Module 2: Prompt Engineering Context Design & Optimization

Advanced Techniques

- Advanced prompt engineering strategies
- Context engineering and design principles
- Reasoning mode optimization
- Reducing hallucinations effectively

Prompting Strategies

- Zero-shot prompting for novel tasks
- Few-shot learning with examples
- Chain-of-thought reasoning
- Multimodal prompting (text, image, audio)
- Domain-specific prompt design

Module 3: LLM APIs & LangChain 1.0

Build production-ready AI applications with LangChain 1.0's `create_agent` abstraction, multi-provider integration, and cost-optimized pipelines.

1 API Integration

OpenAI, Anthropic, Google, and DeepSeek APIs with unified interfaces and authentication.

2 LangChain 1.0

Fundamentals, `create_agent` abstraction, middleware systems for customization and control.

3 Multi-Provider

Seamless integration across providers, fallback strategies, and load balancing.

4 Advanced Features

Streaming responses, batching, function calling, structured outputs, cost optimization.

Module 6: Introduction to Agentic AI Autonomous AI Systems

Discover the paradigm shift from simple LLM calls to autonomous agents that can plan, reason, and act to accomplish complex tasks.



Plan

Breaking down complex tasks into actionable steps



Reason

Logical decision-making and problem-solving



Act

Executing actions and using tools effectively

1 LangChain 1.0 Agents

Agent creation with middleware, tool integration patterns, and customization

2 Model Context Protocol

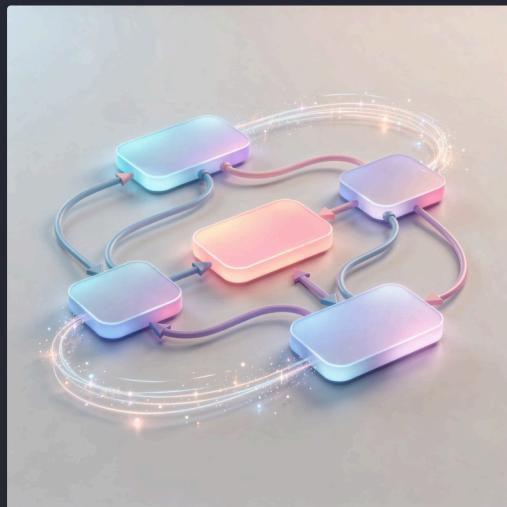
MCP for standardized tool access, security, and interoperability

3 Enterprise Adoption

Use cases, design patterns, and architectural considerations

Module 7: LangGraph 1.0 Fundamentals

Graph-Based AI Workflows



Core Concepts

- LangGraph 1.0 architecture and design philosophy
- State management and graph-based logic
- Node caching for efficient development
- Pre/Post hooks for implementing guardrails
- Building complex AI workflows
- Production use cases and patterns

LangGraph enables sophisticated AI workflows with explicit state management, conditional routing, and robust error handling for enterprise applications.

Module 8: Advanced Workflow Patterns



Real-World Applications

- Essay evaluation systems with rubrics
- Customer feedback routing and analysis
- Multi-stage approval workflows
- Quality-gated content generation

Module 9: Persistence & Human-in-the-Loop Durable AI Workflows

Build resilient, long-running AI systems with persistent state management, human oversight, and enterprise-grade reliability for mission-critical applications.

1 Durable State

State management across sessions, built-in persistence with PostgreSQL and Redis.

2 Human-in-the-Loop

HITL implementations for approval workflows, quality control, and compliance requirements.

3 Multi-Day Workflows

Support for long-running processes, checkpoint management, and resumption capabilities.

4 Enterprise Features

Compliance and audit trails, restart and failure recovery, transaction management.



Module 10: Production Agentic Systems

Enterprise-Grade AI Deployment



LangGraph Platform

Production deployment, scaling, and infrastructure management for enterprise AI systems.



Multi-Agent Systems

Design patterns for coordinating multiple agents, Google A2A Protocol for agent-to-agent communication.



Observability

LangSmith for monitoring, debugging, and performance analysis of AI workflows.



Security & Compliance

MCP security model, prompt injection prevention, agent guardrails, compliance and audit trails.

Build production-ready agentic systems with comprehensive security, monitoring, and governance for enterprise deployment at scale.