Python for ETL Training Course

Professional Development Training Course Outline

Category	Business Intelligence	**Duration**	5 Days
Base Fee	$1,500 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

Python for ETL (Extract, Transform, Load) is a highly practical and industry-relevant course designed to empower professionals with the skills to automate, optimize, and manage data pipelines effectively. In today's data-driven world, organizations need to process vast amounts of data efficiently, and Python's robust libraries and frameworks make it an ideal choice for ETL processes. Python for ETL Training Course provides a hands-on learning experience to develop advanced data engineering skills, enabling participants to handle complex data transformation, integration, and automation projects with confidence.

The course emphasizes real-world applications, combining theory with practical exercises, case studies, and project-based learning. Participants will gain expertise in Python scripting, data extraction from multiple sources, data transformation using Pandas, SQL, and other libraries, and loading data into data warehouses or analytics platforms. With a focus on best practices, performance optimization, and error handling, this training equips professionals to deliver scalable ETL solutions and contribute significantly to organizational data strategy and business intelligence initiatives.

Programme Curriculum

Python for ETL Training Course

Introduction

Python for ETL (Extract, Transform, Load) is a highly practical and industry-relevant course designed to empower professionals with the skills to automate, optimize, and manage data pipelines effectively. In today's data-driven world, organizations need to process vast amounts of data efficiently, and Python's robust libraries and frameworks make it an ideal choice for ETL processes. Python for ETL Training Course provides a hands-on learning experience to develop advanced data engineering skills, enabling participants to handle complex data transformation, integration, and automation projects with confidence.

The course emphasizes real-world applications, combining theory with practical exercises, case studies, and project-based learning. Participants will gain expertise in Python scripting, data extraction from multiple sources, data transformation using Pandas, SQL, and other libraries, and loading data into data warehouses or analytics platforms. With a focus on best practices, performance optimization, and error handling, this training equips professionals to deliver scalable ETL solutions and contribute significantly to organizational data strategy and business intelligence initiatives.

Course Objectives

By the end of this course, participants will be able to:

1. Understand Python fundamentals and advanced programming concepts for ETL processes.
2. Extract data from multiple sources including APIs, databases, and cloud storage.
3. Transform and clean large datasets efficiently using Python libraries.
4. Load data into structured databases, data warehouses, or cloud platforms.
5. Automate ETL pipelines for real-time and batch processing scenarios.
6. Integrate Python with SQL, Spark, and other ETL frameworks.
7. Implement error handling, logging, and monitoring in ETL workflows.
8. Optimize Python ETL scripts for performance and scalability.
9. Apply best practices for data quality, validation, and governance.
10. Develop reusable Python modules for standardized ETL tasks.
11. Handle unstructured and semi-structured data formats effectively.
12. Execute project-based ETL solutions for business intelligence use cases.
13. Analyze and visualize ETL process outcomes for reporting and insights.

Organizational Benefits

1. Streamlined data pipelines and automated workflows.
2. Improved data quality and consistency across systems.
3. Faster processing and transformation of large datasets.
4. Reduced manual intervention and human error in ETL tasks.
5. Enhanced reporting, analytics, and decision-making capabilities.
6. Scalability of ETL processes to handle growing data volumes.
7. Increased team efficiency with standardized Python modules.
8. Real-time monitoring and issue resolution of ETL pipelines.
9. Cost optimization through automation and process efficiency.
10. Strengthened data governance and compliance adherence.

Target Audiences

1. Data Engineers seeking Python ETL expertise.
2. Business Analysts wanting automated data workflows.
3. Data Scientists handling large datasets and integration.
4. IT Professionals involved in database management.
5. BI Developers requiring Python scripting skills.
6. Software Developers transitioning to data engineering roles.
7. Cloud Engineers managing ETL pipelines in cloud environments.
8. Professionals in analytics teams focused on data transformation.

Course Duration: 5 days

Course Modules

Module 1: Python Basics for ETL

- Introduction to Python programming for ETL
- Variables, data types, and operators
- Control structures and loops for data workflows
- Functions and modules for reusable ETL code
- Working with Python libraries: Pandas and NumPy
- Case Study: Automating CSV data extraction and transformation

Module 2: Data Extraction Techniques

- Reading data from CSV, Excel, JSON, and XML
- Connecting to SQL and NoSQL databases
- Accessing APIs for real-time data extraction
- Extracting data from cloud platforms like AWS and Azure
- Handling large-scale data extraction efficiently
- Case Study: Extracting sales data from multiple sources

Module 3: Data Transformation with Python

- Data cleaning and preprocessing techniques
- Using Pandas for transformation operations
- Handling missing values and duplicates
- Data aggregation, merging, and reshaping
- Applying business rules to datasets
- Case Study: Transforming raw customer data for analysis

Module 4: Data Loading Strategies

- Writing data to SQL databases and data warehouses
- Loading data to cloud storage and analytics platforms
- Managing incremental and full-load processes
- Implementing transaction control and rollback mechanisms
- Optimizing data loading for large datasets
- Case Study: Loading transformed sales data to PostgreSQL

Module 5: ETL Automation and Scheduling

- Introduction to workflow automation with Python
- Scheduling jobs using cron, Airflow, or Prefect
- Logging, monitoring, and alerting ETL pipelines
- Handling failures and retry mechanisms
- Best practices for automation scripts
- Case Study: Automating daily ETL pipeline for e-commerce data

Module 6: Error Handling and Data Validation

- Python exception handling in ETL workflows
- Validating data against schema and rules
- Detecting anomalies and inconsistencies

- Implementing retry and alert mechanisms
- Creating reports for failed data validation
- Case Study: Data quality validation for banking transactions

Module 7: Performance Optimization

- Efficient data processing with Pandas and NumPy
- Reducing memory footprint and improving speed
- Parallel processing and multithreading techniques
- Query optimization for database interactions
- Profiling ETL scripts for bottlenecks
- Case Study: Optimizing ETL workflow for large retail dataset

Module 8: Project Implementation and Case Studies

- End-to-end ETL pipeline development
- Applying Python for real business scenarios
- Integrating multiple data sources and formats
- Best practices for maintainable ETL projects
- Performance monitoring and reporting of ETL pipelines
- Case Study: End-to-end ETL project for healthcare data integration

Training Methodology

- Interactive lectures with live coding demonstrations
- Hands-on exercises for each ETL stage
- Case studies based on real-world business scenarios
- Group discussions and problem-solving sessions
- Step-by-step guidance on building end-to-end ETL pipelines
- Continuous feedback and performance assessment

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Online	$1,000
23 Nov 2026	27 Nov 2026	Physical	$1,500
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$4,500
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com