ETL Testing & Development

About CDPL

In today's fast-paced world, IT industry professionals are in demand but aren't employable because of lack of experiential training. CDPL believes in bridging the gap between institutional education and industry demands.

Through our Software certification programs, we aspire to fill the supply needs of IT employers by preparing graduates and working professionals with these skills.

Our certification training courses bundle a comprehensive curriculum with practical skills that will equip our potential students with skills they can apply to various industries.

The training is provided by working professionals with years of experience working with renowned organizations in India and abroad. The courses are offered to meet current industry standards and equip learners with advanced IT skills.

Why Join Us

- Learn Real-Time Software Solutions inspired by years of experience we have in satisfying our corporate clients.
- Learn Software skills in an industrial environment with practical demonstration and live scenarios.
- High-Quality Standard Maintenance in each session delivery.
- Train on real projects to enhance your skills and make yourself industry-ready.
- Learn corporate ethics and moral code of conduct.
- Extensive interview preparation.
- Resume-building and grooming Sessions.
- Guarantee Job Interviews and Placement assistance.

Why ETL Testing & Development?

The demand for professionals skilled in ETL testing, development, and automation has skyrocketed across industries like retail, telecom, and financial services. This program equips learners with the skills to design, test, and optimize ETL/ELT processes, automate testing workflows, and build comprehensive data integration solutions using leading tools and programming languages such as Python, Selenium, Talend, Informatica, and Snowflake.

With a focus on both manual and automated ETL testing, relational database concepts, and data transformation techniques, this course is designed to prepare learners for diverse roles in the ETL and data pipeline ecosystem.

Expectations and Goals

- This course will make you a Proficient ETL Test Engineer and ETL developer.
- At the end of this course, you will gain a strong understanding of ETL testing, data integration, and data warehousing concepts.
- Master tools like Talend, Informatica, and Selenium for ETL workflows.
- Learn Python programming for ETL automation and troubleshooting.
- Acquire hands-on experience in testing relational databases (Oracle, MySQL, MSSQL).
- Prepare for agile and fast-paced environments with real-world projects.

Who This Course is for?

- Undergraduates or job seekers seeking to launch their careers in the IT domain.
- Working professionals with non-IT experience who want to transition to the IT field.
- Fresh graduates or postgraduates aiming to establish their careers in the IT domain.
- Professionals aspiring for roles in ETL Testing, ETL Development, Data Science, or Software Engineering.
- Software Test Engineers looking to expand their expertise into ETL and data testing.
- Data Engineers who want to learn automated and manual ETL validation techniques.
- IT Professionals interested in working with modern tools like Snowflake, SnapLogic, and Power BI.

Course Curriculum   **Total Duration: 100 Hours**

Sr. No	Topics	Duration
1	**Introduction to ETL** Overview of ETL/ELT processes, importance of ETL in data pipelines, roles of testers/developers	2 Hours
2	**Data Warehousing Concepts** Data warehouse architecture, star vs. snowflake schema, relational vs. dimensional models	4 Hours
3	**SQL for ETL Testing** Writing basic and advanced SQL queries, data validation with SQL, debugging SQL-based ETL jobs	16 Hours
4	**Manual ETL Testing** Creating test plans, designing test cases for completeness and accuracy, defect management	10 Hours
5	**Automation in ETL Testing** Automating tests with Python and Selenium, parameterization, reporting	20 Hours

cdpl

6	**Talend for ETL Development** Working with Talend Studio: tMap, tJoin, dynamic schemas, integrating data from multiple sources	10 Hours
7	**Informatica for ETL Development** Creating workflows, handling transformations, debugging, and optimization	10 Hours
8	**Data Transformation and Validation** Handling null values, deduplication, applying business logic, real-time data validation	10 Hours
9	**ETL Tools Overview** Introduction to Snowflake, SnapLogic, Power BI for ETL workflows	10 Hours
10	**Capstone Projects** Hands-on projects in ETL testing, automation, and development using multiple tools	8 Hours

Tools & Technology Covered

- **ETL Tools**: Talend, Informatica, SnapLogic.
- **Programming and Automation**: Python, Selenium.
- **Databases**: Oracle, MySQL, MSSQL, Snowflake.
- **Reporting and Visualization**: Power BI.

Capstone Project

Project 1: ETL Testing for a Retail Sales Data Pipeline

Project Overview:
 This project focuses on testing a simulated ETL pipeline that processes retail sales data. The pipeline extracts data from multiple sources, transforms it based on specific business rules, and loads it into a data warehouse for business intelligence (BI) reporting. The objective is to ensure the ETL process delivers accurate, complete, high-performance, and high-quality results.

Project Deliverables:

1. **Test Plan**: Comprehensive document defining the scope, objectives, and testing strategies for the ETL pipeline.
2. **Test Cases**: Detailed scenarios covering validation of extracted data, transformation logic, and load phase.
3. **Test Scripts**: Automated SQL scripts or Python programs for testing the ETL phases.
4. **Test Results**: A report summarizing the outcomes, issues, and resolutions of the testing process.

Learning Outcome:

- Expertise in validating each phase of the ETL process, including data extraction, transformation, and loading.
- Hands-on experience creating test plans, test cases, and automated scripts for ETL validation.
- Knowledge of troubleshooting issues in data pipelines and ensuring data integrity.

Project 2: Building and Validating an ETL Pipeline for an E-Commerce Business

Project Overview:
 This project involves designing, developing, and validating a fully functional ETL pipeline tailored for an e-commerce business. The pipeline integrates customer data, sales data, and inventory information. Manual and automated testing ensures accuracy and efficiency, while data visualization provides actionable insights.

Project Deliverables:

1. **ETL Pipeline**: Fully operational pipeline integrating data from multiple sources.
2. **Test Cases**: Comprehensive manual and automated test cases for each phase of the ETL process.
3. **Reports**: Insights on customer behavior, sales trends, and inventory status, visualized using Power BI.

Learning Outcome:

- Mastery in designing and implementing ETL pipelines.
- Practical experience with tools and technologies such as Talend, Informatica, Python, Selenium, and Snowflake.
- Proficiency in manual and automated ETL testing.
- Data visualization and reporting expertise using Power BI.

How This Course Links to Job Roles

cdpl

1. **Python with ETL Tester**

   - Learn Python for scripting and automating ETL tests.
   - Work on data transformation processes in Snowflake and relational databases.
   - Gain expertise in building ELT pipelines and validating data.

2. **Selenium Testing + Data/ETL Testing**

   - Automate ETL validation using Selenium.
   - Test data pipelines for integrity, performance, and scalability.

3. **ETL Developer**

   - Build ETL pipelines using Talend and Informatica.
   - Migrate ETL processes to tools like SnapLogic.
   - Handle real-time and batch data workflows.

4. **ETL Tester**

   - Perform manual testing for ETL processes.
   - Validate data transformation rules and referential integrity.
   - Learn agile testing methodologies.

5. **Software Engineer (ETL Informatica)**

   - Develop and support ETL workflows in Informatica.
   - Resolve production issues and optimize ETL pipelines.
   - Work on post-production support and enhancement projects.

cdpl