

**Workflow Management with Apache Airflow**

**Course Number:** PYTH-248WA  
**Duration:** 2 days

**Overview**

Apache Airflow is an open-source, Python-based solution that allows developers to programmatically author, schedule, and monitor complex workflows. This Workflow Management with Apache Airflow training course teaches attendees how to manage workflows successfully and incorporate sophisticated automation techniques into their processes.

**Prerequisites**

Participants must have some familiarity with Python or have a programming background.

**Materials**

All Apache Airflow training students receive comprehensive courseware.

**Software Needed on Each Student PC**

* Python 3.5 or later
* Airflow 2.1 or later

**Objectives**

* Work in the Airflow environment
* Understand Airflow DAG
* Schedule Jobs
* Work with backfilling
* Pass Parameters
* Perform XCom messaging
* Understand task branching
* Understand re-tries
* Use SimpleHttpOperator

**Outline**

* Apache Airflow Introduction
  + A Traditional ETL Approach
  + Apache Airflow Defined
  + Airflow Core Components
  + The Component Collaboration Diagram
  + Workflow Building Blocks and Concepts
  + Airflow CLI
  + Main Configuration File
  + Extending Airflow
  + Jinja Templates
  + Variables and Macros
* Apache Airflow Web UI
  + Web UI - the Landing (DAGs) Page
  + Web UI - the DAG Graph View
  + Run Status Legends
  + The Pause Button (Trigger Latch)
  + The DAG Triggering/Job Checking Sequence
  + The Control Panel for a Task
  + Sample Log File Messages (Abridged for Space)
* Anatomy of a DAG and Scheduling
  + What is a DAG?
  + Scheduled and Manually Triggered DAG Runs
  + The DAG Object
  + Tasks
  + Task Lifecycle
  + Operators
  + Idempotent Operators
  + Operator Types
  + Airflow Common Operators
  + Specifying Dependencies
  + Associating Operators with a DAG
  + Associating Operators Using the "With DAG" Statement Example
  + Associating Operators with DAG Using the Operator's Constructor
  + The default\_args Parameter
  + Passing DAG Parameters Through Web UI
  + DAG Run Scheduling
  + Examples of the schedule\_interval Parameter
  + DAG Scheduling Nuances
  + Understanding The Backfill Process
  + Killing/Stopping DAG Runs
  + An XCom Messaging Example
* Conclusion