



To discuss this course and customizations:  
Call: 434-509-5680 or Email: [sales@cloudcontraptions.com](mailto:sales@cloudcontraptions.com)

## **Apache Airflow Programming: Developing, Configuring, and Automating Workflows**

### **Duration**

28 hours

### **Target Audience**

- Practical experience with Python.
- Familiarity with Containerization and Container Orchestration.
- Basic Linux command line skills.

### **Executive Summary**

Over three days, this course immerses participants in Apache Airflow's architecture and configuration, guiding them through setting up environments, choosing executors, and developing robust DAGs with Python. Through hands-on exercises—ranging from dynamic task mapping and templating to cloud integrations and custom plugin development—attendees will master best practices for automating, monitoring, and optimizing production-ready workflows.

### **Description**

This course provides a comprehensive introduction to Apache Airflow, covering its architecture, configuration, and workflow automation capabilities. Participants will learn how to set up and manage Airflow environments, configure executors, and develop DAGs using Python. The course explores essential components like tasks, operators, variables, and connections, as well as advanced topics such as dynamic DAGs, templating, and custom plugins. Hands-on exercises include running DAGs, scheduling tasks, integrating cloud providers, and monitoring workflows through logs and the Airflow UI. By the end of the course, participants will be equipped to build, automate, and optimize data pipelines using Airflow.

### **Objectives**

- Understand Apache Airflow's architecture and how it automates distributed workflows.



To discuss this course and customizations:  
Call: 434-509-5680 or Email: [sales@cloudcontraptions.com](mailto:sales@cloudcontraptions.com)

- Set up and configure Airflow using different execution modes and database backends.
- Learn key Airflow components, including DAGs, tasks, operators, variables, and connections.
- Develop and run DAGs using the Operator API, TaskFlow API, and dynamic task mapping.
- Integrate Airflow with cloud providers such as AWS and Azure.
- Utilize built-in operators and sensors to automate task execution and monitoring.
- Extend Airflow by creating custom operators, providers, and plugins.
- Apply best practices for scheduling, logging, debugging, and optimizing workflows.

## Training Materials

Students receive comprehensive courseware, including reference documents, code samples, and lab guides.

## Software Requirements

Students will need a free, personal GitHub account to access the courseware. Students will need permission to install Python and Visual Studio Code on their computers. Also, students will need permission to install Python Packages and Visual Studio Extensions. If students are unable to configure a local environment, a cloud-based environment can be provided.

## Training Topics

### What is Apache Airflow?

- Distributed Task Automation
- Compared to Cron Jobs
- Compared to Celery
- Scalability and Reliability
- Directed Acyclic Graphs (DAGs)
- Workflows as Code

### Workflows as Code (no programming)

- Anatomy of a DAG
- Directed Acyclic Graphs
- Operators



To discuss this course and customizations:  
Call: 434-509-5680 or Email: [sales@cloudcontraptions.com](mailto:sales@cloudcontraptions.com)

- Tasks
- Variables
- XComs
- Providers
- Connections
- Explore how DAG parts connect to the UI
- DAG Serialization
- Listeners
- Schedulers
- Pools

#### Installation and Configuration

- Python Virtual Environment
- Install Airflow
- Airflow Constraints File
- Standalone Mode
- Run the Webserver and Scheduler Independently
- SQLite vs PostgreSQL
- Configure with PostgreSQL
- Airflow and Kubernetes (with Minikube)
- Airflow and AWS Elastic Kubernetes Service (EKS)
- Airflow Helm Chart

#### Hands-On Kubernetes (K8s)

- Containerization and Orchestration
- Kubectl
- Helm
- Nodes
- Namespaces
- Pods, Containers, and Services
- Connect to the Internet (EKS)
- Keda Autoscaler
- Pod Logs
- SSH into Pods/Containers
- Live Upgrading Airflow



To discuss this course and customizations:  
Call: 434-509-5680 or Email: [sales@cloudcontraptions.com](mailto:sales@cloudcontraptions.com)

### Airflow Configuration

- Airflow Configuration File Location
- Airflow Executor Configuration
- Airflow Log Levels
- Helm Chart Configuration
- Learn How to Configure Airflow and K8s Pods
- Local Executor
- Celery Executor
- K8s Pod Executor

### Airflow Custom Image

- Airflow Container Image
- Why Create a Custom Image?
- Create a Custom Image
- Install Software with Apt
- Install Software with PyPi
- Install Providers and Custom Software
- Use the Custom Container Image

### Monitoring

- Logging
- Log File Structure
- Log Levels
- Review Task Logs in the Web UI
- External Log Storage
- Metrics Configuration
- Monitor with Grafana
- Notifications