KOENIG
step forward

KOENIG SOLUTIONS · COURSE OUTLINE

Python Automation: Web Scraping, Excel & PDF

Learn to automate file, web, Excel and PDF tasks with Python, from core syntax to scheduling and parallel execution.

15
MODULES

70
TOPICS

COURSE OVERVIEW

This course begins with Python fundamentals, covering data types, variables, data structures, functions, modules and flow control. It then moves into automation, including setting up a development environment, handling files and directories, scraping web data with BeautifulSoup and Requests, and automating GUI, database and scheduled tasks. It also addresses error handling, logging and parallel execution using threading and multiprocessing. The final modules cover automating Excel and PDF files, along with image processing tasks such as resizing, cropping and OCR.

WHO THIS COURSE IS FOR

Python Developers | **Automation Engineers** | **Data Analysts** | **QA Engineers** | **IT Administrators**

LEARNING OBJECTIVES

- Explain core Python syntax including data types and data structures
- Apply functions, modules and flow control in Python scripts
- Describe the benefits of automation and available Python libraries
- Implement file handling operations such as reading, writing and appending
- Apply web scraping techniques using BeautifulSoup and Requests
- Configure GUI, database and task scheduling automation
- Implement error handling and logging in automation scripts
- Develop Excel, PDF and image processing automation solutions

COURSE SYLLABUS

15 Modules, 70 Topics

01

Python Basics

TOPICS

Introduction to Python · Data Types · Variables · Data Structure · Operations on Data Structure · Inbuilt methods

02

Data Types and Variables

TOPICS

- Data Types
- Understanding Variables
- Getting User Input and Print statements
- Use quotes and escape character
- Lists
- Tuples
- Dictionaries
- Sets
- Different Methods on data structures
- String Indexing and Slicing Strings
- Concatenation and Repetition
- Common String Methods
- String Formatting
- Namespaces
- Formatted String Literals (f-strings and .format() method)
- Built-in String Functions

03

Functions and Modules

TOPICS

Defining Functions · Using Parameters and Return Values · Using Arguments and Default Parameters · Scope of Variables · Modules and Packages · Writing and Importing Modules

04

Flow Control

TOPICS

Python Operators · If, else, elif clauses · Loops - for and while loops · Iteration

05

Introduction to Python Automation

TOPICS

- Overview of Automation and its Benefits
- Introduction to Python Libraries for Automation
- Setting up the development environment

06

File Handling and Manipulation

TOPICS

Working with files and directories · Automating file operations · File searching and filtering · Opening Files · The os and os.path modules · Reading files · Writing into a file · Appending data into a file

07

Web Scraping and Data Extraction

TOPICS

- Introduction to web scraping
- Using BeautifulSoup and Requests libraries for web scraping

08

GUI Automation

TOPICS

Introduction to GUI automation · Automating mouse and keyboard actions · Interacting with desktop applications

09

Database Automation

TOPICS

- Introduction to database automation
- Connecting to databases using Python
- Automating database queries and data manipulation tasks

10

Task Scheduling and Automation

TOPICS

Introduction to task scheduling · Using cron jobs and task scheduler libraries · Automating recurring tasks and processes

11

Error Handling and Logging

TOPICS

- Handling errors and exceptions in automation scripts
- Implementing logging for debugging and monitoring

12

Parallel Execution in Python

TOPICS

Threading · Multiprocessing for parallel execution · Implementing automation best practices · Performance optimization techniques

13

Excel Automation

TOPICS

- Introduction to Excel automation
- Reading and writing Excel files using pandas library
- Handling different Excel formats (e.g., .xls, .xlsx)
- Extracting data from Excel worksheets and ranges
- Automating data manipulation and analysis tasks with Excel files

14

PDF Automation

TOPICS

- Reading and extracting data from PDF files using PyPDF2 or pdf plumber library
- Automating PDF generation and manipulation tasks
- Converting PDF files to other formats (e.g., text, images)

15

Image Processing with Python

TOPICS

- Automating image manipulation tasks (e.g., resizing, cropping, filtering)
- Working with image recognition and OCR (Optical Character Recognition) for automation

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training | Live Online Classes (ILO) | Fly-Me-A-Trainer (FMAT) | Flexi (Self-Paced) | 1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com | info@koenig-solutions.com