

Jupyter for Data Science Teams Training

COURSE CONTENT

GET IN TOUCH



Multisoft Systems
B - 125, Sector - 2, Noida



(+91) 9810-306-956



info@multisoftsystems.com



www.multisoftsystems.com

About Multisoft

Train yourself with the best and develop valuable in-demand skills with Multisoft Systems. A leading certification training provider, Multisoft collaborates with top technologies to bring world-class one-on-one and certification trainings. With the goal to empower professionals and business across the globe, we offer more than 1500 training courses, which are delivered by Multisoft's global subject matter experts. We offer tailored corporate training; project Based Training, comprehensive learning solution with lifetime e-learning access, after training support and globally recognized training certificates.

About Course

The Jupyter for Data Science Teams training by Multisoft Systems is designed to equip professionals with the skills required to collaborate effectively in data-driven environments using Jupyter notebooks. As one of the most widely adopted open-source platforms, Jupyter provides an interactive workspace where teams can write code, visualize data, document workflows, and share results seamlessly.

Module 1: Introduction to Jupyter

1.1 Overview of Jupyter and its ecosystem

- ✓ Introduction to Jupyter Notebook, JupyterLab, and JupyterHub
- ✓ Explanation of Jupyter's role in data science workflows

1.2 Installation and setup

- ✓ Step-by-step guide to installing Jupyter on various platforms (Windows, macOS, Linux)
- ✓ Configuring Jupyter settings for optimal performance and customization

1.3 Configuring Jupyter for team collaboration

- ✓ Setting up JupyterHub for multi-user collaboration
- ✓ Managing user permissions and access control in JupyterHub environments

Module 2: Collaborative Features

2.1 Using Git for version control

- ✓ Introduction to version control concepts and Git
- ✓ Integrating Git with enhanced functionality and productivity

Module 3: Creating and Managing Notebooks

3.1 Notebook structure and functionality

- ✓ Understanding the components of a Jupyter notebook: cells, kernels, and markdown
- ✓ Exploring different cell types and their usage (code, markdown, raw)

3.2 Sharing and organizing notebooks

- ✓ Methods for sharing notebooks with team members and external stakeholders

- ✓ Organizing notebooks into projects and directories for efficient management and retrieval

Module 4: Programming with Jupyter

4.1 Choosing and using programming languages (Python, R, Scala)

- ✓ Overview of supported programming languages in Jupyter and their respective kernels
- ✓ Best practices for selecting the appropriate language for specific data science tasks

4.2 Writing and executing code

- ✓ Writing code in Jupyter cells and executing them interactively
- ✓ Understanding code execution order and kernel interruptions

4.3 Integrating with big data systems (Apache Spark)

- ✓ Overview of Apache Spark integration with Jupyter for big data processing
- ✓ Running Spark jobs and analyzing large datasets within Jupyter notebooks

Module 5: Advanced Jupyter Features

5.1 Customizing Jupyter environment

- ✓ Personalizing Jupyter interface and themes for improved user experience
- ✓ Installing and managing Jupyter extensions for additional functionality

5.2 Automating workflows with Jupyter

- ✓ Leveraging Jupyter for automating repetitive tasks and data processing workflows
- ✓ Creating custom scripts and extensions to streamline complex workflows within Jupyter