**Introduction to Data Science** 

Duration: 3 Days

Course Code: NTVPY-230

Description

This workshop will introduce participants to the field of data science using Python. We’ll cover key concepts such as data manipulation, data analysis, and data visualization using popular Python libraries. Participants will also learn basic machine learning techniques and how to apply them to real-world datasets.

Objectives

* Understand the fundamentals of data science and its importance in industry.
* Use key Python libraries like Pandas, NumPy, and Matplotlib for data manipulation and analysis.
* Apply basic machine learning techniques to real-world datasets.
* Gain proficiency in data cleaning and visualization using Python.
* Develop a foundational knowledge of machine learning algorithms and model evaluation.

Prerequisites

* Basic knowledge of Python programming
* Familiarity with concepts of data types, variables, loops, and functions in Python
* No prior knowledge of data science is required

Outline

* Chapter One – Getting Started
	+ What is Data Science?
	+ Importance of Data Science in Industry
	+ The Wave
	+ Overview of key Python Libraries
		- Pandas
		- NumPy
		- Matplotlib
* Chapter Two – NumPy
	+ NumPy Basics
	+ Array Operations
	+ Handling Multidimensional Data
	+ Array Indexing and Slicing
	+ Broadcasting and Vectorization
	+ Data Aggregation
* Chapter Three – Pandas Mechanics
	+ Introduction to Pandas
	+ Data Structures
	+ Data Importing and Exporting
	+ Data Cleaning
* Chapter Four – Data Analysis with Pandas
	+ Grouping and Aggregating Data
	+ Time Series Analysis
	+ Date Transformation and Reshaping
* Chapter Five – Data Visualization with Matplotlib
	+ Understanding Matplotlib
	+ Creating Basic Plots (line plots, bar plots, scatter plots)
	+ Customizing Plots and Adding Annotations
* Chapter Six – Overview of Machine Learning
	+ What is Machine Learning
	+ Types of Machine Learning Algorithms
		- Supervised Learning
		- Unsupervised Learning
		- Reinforcement Learning
* Chapter Seven – scikit-learn
	+ Exploring the scikit-learn library
	+ Building a Machine Learning Model
	+ Model Evaluation and Validation