PURPLE GRIFFON
The IT Service Management Training & Consultancy Specialists

Course Title: Data Wrangling With Python	**Course Duration**: 3.0 Days
Exam: Not Included	**Exam Type**: Proctored Exam
Qualification: Data Wrangling With Python Certificate	

Course Syllabus

Our Data Wrangling With Python training course covers the following Modules:

Module 1: Introduction to Data Structure Using Python

- Python for Data Wrangling
- Lists, Sets, Strings, Tuples, and Dictionaries

Module 2: Advanced Operations on Built-In Data Structure

- Advanced Data Structures
- Basic File Operations in Python

Module 3: Introduction to NumPy, Pandas, and Matplotlib

- NumPy Arrays
- Pandas Data Frames
- Statistics and Visualisation with NumPy and Pandas
- Using NumPy and Pandas to Calculate Basic Descriptive Statistics on the Data Frame

Module 4: Deep Dive into Data Wrangling with Python

- Sub-setting, Filtering, and Grouping
- Detecting Outliers and Handling Missing Values
- Concatenating, Merging, and Joining
- Useful Methods of Pandas

Module 5: Getting Comfortable with Different Data Sources

- Reading Data from Different Text-Based (and Non-Text-Based) Sources
- Introduction to BeautifulSoup4 and Web Page Parsing

Module 6: Learning the Hidden Secrets of Data Wrangling

- Advanced List Comprehension and the zip function
- Data Formatting

Module 7: Advanced Web Scraping and Data Gathering

- Basics of Web Scraping and BeautifulSoup libraries

- Reading Data from XML

Module 8: RDBMS and SQL

- Refresher of RDBMS and SQL
- Using an RDBMS (MySQL/PostgreSQL/SQLite)

Module 9: Application in Real Life and Conclusion of Course

- Applying Your Knowledge to a Real-life Data Wrangling Task
- An Extension to Data Wrangling

Course Overview

Our Data three-day Wrangling With Python training course will teach how to use Python to extract/transform data from various sources, including large database vaults and Excel financial tables. You will also explore insights into why you should avoid traditional data cleaning methods, as done in other languages, and take advantage of the specialised functions from NumPy and Pandas.

Course Learning Outcomes

Our Data Wrangling With Python training course will teach you to become proficient in the following:

- Extract and parse data from various sources.
- Transform and clean data using Numpy and Pandas.
- Summarise and visualise data with Matplotlib.
- Read HTML, XML, and JSON data from internet resources.
- Search and filter data sets.
- Apply Python tools and techniques to process data sets efficiently.
- Continue learning and face new challenges with after-course one-on-one instructor coaching.

Audience

Our Data Wrangling With Python training course will benefit several individuals and organisations including but not limited to:

- Data Analysts
- Data Scientists
- Data Engineers
- Business Analysts
- Researchers
- Software Engineers
- Business Intelligence Professionals
- Anyone working with data

Entry-Level Requirements

Our Data Wrangling With Python training course requires attendees to know Python basics, including data structures, importing and using modules, creating functions, and using the Jupyter Notebook platform.

Recommended Reading

There is no recommended reading for our Data Wrangling With Python training course.

What's Included

Our Data Wrangling With Python training course contains the following:

- 3-day instructor-led training course
- After-course coaching available
- Pre-reading
- Course Manual
- Quizzes
- Exercises

Exam Information

Data Wrangling With Exam:

- Format: Multiple Choice
- Questions: 40
- Pass Mark: 70%

What's Next

Attendees may enjoy our Applied Data Science With Python And Jupyter training course.

Our Applied Data Science With Python And Jupyter training course you will:

- Learn about some of the most commonly used libraries that are part of the Anaconda distribution and then explore machine learning models with real datasets.
- learn about creating reproducible data processing pipelines, visualisations, and prediction models, all with the goal of giving you the skills and exposure you'll need for the real world.

Data Science is one of the fastest growing professions across all industries. Open source tools like Python have become increasingly popular, and when paired with Jupyter Notebooks, can provide a variety of data-science applications. Attend this one-day hands-on course and learn to leverage all that these powerful tools have to offer.

Additional Information

Our Data Wrangling With Python training course offers several benefits to individuals and organisations including but not limited to:

- **Data Cleaning and Pre-processing**: Python provides powerful libraries like Pandas that make it easier to clean and pre-process data. You can handle missing values, remove duplicates, standardise formats, and perform other data cleaning tasks efficiently, ensuring data quality and reliability.

- **Data Transformation and Reshaping**: Python's data manipulation libraries offer versatile tools for transforming and reshaping data.
- Efficiency and Productivity: Python's expressive and concise syntax, combined with its extensive ecosystem of data manipulation libraries, allows you to perform complex data wrangling tasks efficiently.
- **Reproducibility and Documentation**: Python encourages good coding practices, including documentation and code organisation.
- **Integration with Data Analysis and Machine Learning**: Python is widely used in the data analysis and machine learning communities.
- **Handling Large Datasets**: Python offers libraries and techniques for efficiently handling and processing large datasets, such as memory optimisation, chunking, and parallel processing.
- **Web Scraping and Data Collection**: Python provides libraries like Beautiful Soup and Scrapy for web scraping and data collection.
- **Problem Solving and Critical Thinking**: Data wrangling often involves working with messy and complex data.