Faysal Aziz

Data Wrangling & Exploratory Data Analysis with Python

Unleashing the Power of Your Data in 1 day

FA-0158

Software Development · Data & Analytics

Introduction

Welcome to your toolkit for taming the data wilderness! This hands-on, full-day workshop will transform how you wrangle unruly datasets and uncover their hidden stories.

Designed for Python-savvy professionals, this intensive experience puts theory into immediate practice with real-world datasets you'll actually encounter in your career. No academic exercises here—just practical skills you can deploy on Monday morning.

Your guide on this journey brings over three decades of industry battle scars and more than 10,000 hours leading professionals just like you through these very challenges. The techniques you'll master have been refined through years of frontline data work across multiple industries.

By day's end, you'll walk away with powerful, immediately applicable approaches to preprocessing and exploratory data analysis that will distinguish your work and accelerate your projects.

Learning outcomes

- Master the use of Python libraries such as Pandas, NumPy, and Matplotlib for data manipulation and visualization.
- Develop techniques to clean, transform, and merge datasets from diverse sources.
- Gain proficiency in identifying and handling missing data and outliers.
- Perform comprehensive exploratory data analysis to uncover patterns and insights.
- Apply statistical methods to summarize and interpret data effectively.

Prerequisites

- Strong proficiency in Python programming, including experience with data structures, functions, and modules.
- Familiarity with Jupyter Notebook or similar interactive Python environments.
- Basic understanding of statistical concepts and data analysis methodologies.

Training outline

1. Introduction to Data Wrangling and EDA

1. Understanding the data analysis pipeline.
2. The significance of data wrangling and EDA in the data science workflow.
3. Overview of Python's ecosystem for data analysis.

2. Data Structures and Manipulation with Pandas

1. Deep dive into Pandas Series and DataFrame objects.
2. Techniques for data selection, filtering, and indexing.
3. Implementing group operations and pivot tables for data summarization.

3. Data Cleaning Techniques

1. Strategies for detecting and handling missing data.
2. Identifying and addressing outliers to ensure data integrity.
3. Standardizing and normalizing data for consistency.

4. Data Transformation and Feature Engineering

1. Converting data types and formatting for analysis readiness.
2. Creating new features through mathematical transformations and domain-specific knowledge.
3. Encoding categorical variables for analytical modeling.

5. Merging and Joining Datasets

1. Techniques for combining datasets using joins and concatenations.
2. Resolving common issues in merging data from multiple sources.

6. Exploratory Data Analysis Techniques

1. Utilizing descriptive statistics to summarize data distributions.
2. Visual exploration of data using Matplotlib and Seaborn.
3. Assessing relationships between variables through correlation analysis.

7. Case Study: Real-World Data Analysis

1. Applying learned techniques to a real-world dataset.
2. Conducting a full cycle of data wrangling and EDA.
3. Interpreting results and deriving actionable insights.

8. Best Practices and Industry Insights

1. Leveraging the instructor's extensive industry experience to discuss common challenges and solutions.
2. Adopting best practices for efficient and reproducible data analysis workflows.

© 2026 Faysal Aziz · Kuala Lumpur