Sangy Academy

www.sangyacademy.cQom

📚 Course Title: Introduction to Data Analysis with Python

✅ Course Objectives

By the end of this course, learners will:

- Understand core data analysis workflows.
- Be proficient with NumPy, pandas, Matplotlib, and seaborn.
- Know how to clean, transform, and visualize data.
- Perform exploratory data analysis (EDA) and basic statistical analysis.
- Work on real-world datasets and present insights.

🗓️ Weekly Breakdown

Month 1: Foundations of Data Analysis

Week 1: Course Overview + Working with Data in Python

- Lesson 1: Course intro, the data analysis process, types of data
- Lesson 2: Jupyter Notebooks, file types (CSV, Excel), basic I/O with pandas

Week 2: Introduction to NumPy

- Lesson 3: NumPy arrays, broadcasting, indexing, slicing
- Lesson 4: Vectorized operations, reshaping, aggregations

Week 3: Introduction to pandas

- Lesson 5: Series and DataFrames, loading datasets, basic exploration
- Lesson 6: Indexing, selecting data, filtering rows, basic statistics

Week 4: Data Cleaning Basics

- Lesson 7: Handling missing data, replacing values, type conversion
- Lesson 8: Removing duplicates, renaming columns, combining DataFrames

---

Month 2: Intermediate Analysis and Visualization

Week 5: Data Manipulation with pandas

- Lesson 9: Sorting, grouping, and aggregation
- Lesson 10: Applying functions with `.apply()`, `.map()`, and `.lambda`

Week 6: Data Visualization with Matplotlib

- Lesson 11: Plotting basics: line, bar, histogram, scatter
- Lesson 12: Plot customization: labels, titles, subplots, styles

Week 7: Data Visualization with seaborn

- Lesson 13: seaborn basics: countplot, boxplot, violinplot
- Lesson 14: Pairplot, heatmaps, correlation plots, styling

Week 8: Exploratory Data Analysis (EDA)

- Lesson 15: What is EDA, univariate analysis
- Lesson 16: Multivariate analysis, patterns, and correlations

---

Month 3: Real-World Projects and Statistics

Week 9: Time Series and Date Data

- Lesson 17: Working with datetime in pandas, time-based indexing
- Lesson 18: Resampling, rolling windows, time series plots

Week 10: Intro to Statistical Analysis

- Lesson 19: Descriptive statistics (mean, median, std, etc.)
- Lesson 20: Inferential statistics (distributions, z-scores, basic hypothesis testing)

Week 11: Final Project – Part 1

- Lesson 21: Project dataset intro, planning analysis, group feedback
- Lesson 22: Guided work on EDA, visualizations, story development

Week 12: Final Project – Part 2

- Lesson 23: Final polishing, creating summary reports or slides
- Lesson 24: Student presentations, feedback, and wrap-up

🧰 Tools and Libraries

- Jupyter Notebook
- pandas
- NumPy
- Matplotlib
- seaborn
- (Optional: SciPy, statsmodels)

📘 Assessment

- Weekly exercises (10%)
- Mid-course quiz (10%)
- Mini project (20%)
- Final project (40%)
- Participation & attendance (20%)
