

Course 1 (Week 1)

# Introduction to Data Analytics, Python Programming, and GitHub Copilot

Lecture 3: Data Analysis with Python





What's your favorite way to unwind after a long day?

# Skills Covered

- Mastering data manipulation using NumPy arrays and Pandas DataFrames.
- Loading and cleaning datasets using Pandas.
- Performing filtering, sorting, and transforming operations on datasets.
- Applying group-by operations and aggregations to summarize data.
- Merging and joining datasets for more complex data manipulation.

# Objectives for today

1. Introduction to Data Analysis
2. Python Libraries
3. NumPy Basics
4. Introduction to Pandas for Data Manipulation
5. Data Manipulation with Pandas
6. Advanced Pandas: Merging and Joining DataFrames

# Learning Outcomes

- Ability to create and manipulate NumPy arrays for data analysis.
- Proficiency in loading, exploring, and transforming data using Pandas.
- Understanding how to clean and prepare datasets for analysis.
- Applying advanced Pandas techniques like merging and grouping data.
- Capability to export processed datasets and derive meaningful insights from data.



# Introduction to Data Analysis

## Sections

**Section 1**

Section 2

Section 3

Section 4

Section 5

# Data Analysis

- XX

# Python Libraries

- XX



# Numpy

- XX

# Pandas

- XX

# Installing and Setting Up the Libraries

- XX

**Let's go to the JN**

# NumPy



## Sections

Section 1

**Section 2**

Section 3

Section 4

Section 5

# NumPy

- Introduction to NumPy Arrays: Understanding the difference between NumPy arrays and Python lists.
- Creating Arrays: Various methods (from lists, using `arange()`, `linspace()`, etc.).
- Array Operations: Basic arithmetic, broadcasting, and aggregation functions (`sum`, `mean`, etc.).
- Indexing and Slicing: Accessing and modifying data in arrays.
- Hands-on Exercise: Create and manipulate NumPy arrays to perform basic data manipulation.



# Data Manipulation with Pandas

## Sections

Section 1

Section 2

**Section 3**

Section 4

Section 5

# Pandas

- Filtering and Sorting Data: Applying conditions, sorting by columns, and working with indexes.
- Data Cleaning Techniques: Removing duplicates, dealing with NaNs, renaming columns, etc.
- Data Transformation: Using functions like `apply()`, `map()`, and vectorized operations for efficient transformations.
- GroupBy and Aggregation: Summarizing data using grouping and aggregate functions.
- Hands-on Exercise: Apply filtering, sorting, and grouping operations on a sample dataset.





# Advanced Pandas: Merging and Joining DataFrames

## Sections

Section 1

Section 2

Section 3

**Section 4**

Section 5

# Data Manipulation with Pandas

- Combining DataFrames: Concatenation, merging, and joining datasets.
- Pivot Tables: Creating and using pivot tables to summarize data.
- Exporting Data: Saving cleaned and processed data to CSV or Excel formats.
- Hands-on Exercise: Merge multiple datasets and create pivot tables for data insights.

# Key highlights

That's a wrap

**Any Questions?**

# Bibliography

- XXX