

Data Science Workshop

Module 2

Python Basics: Variables, Data Types, and Operators

- **Variables:** Containers for storing data values.
- **Data Types:**
 - **Numeric:** int, float
 - **String:** str
 - **Boolean:** bool
 - **None:** NoneType
- **Operators:**
 - **Arithmetic:** +, -, *, /, %, ** (exponentiation)
 - **Comparison:** ==, !=, <, >, <=, >=
 - **Logical:** and, or, not

Control Flow: Conditional Statements and Loops

- **Conditional Statements:**
 - **If Else:** Used to make decisions.
- **Loops:**
 - **For Loop:** Iterates over a sequence. Example:
 - **While Loop:** Continues as long as the condition is True.

Functions and Modules

- **Functions:** Reusable block of code.
 - **Function Parameters:** Accept inputs.
 - **Return Values:** Provide output.
- **Modules:** Collection of functions and variables.
 - **Standard Library Modules:** os, sys, math, datetime, etc.
 - **Third-party Libraries:** NumPy, Pandas, etc.

Data Structures: Lists, Tuples, and Dictionaries

- **Lists:** Ordered, mutable collection.
- **Tuples:** Ordered, immutable collection.
- **Dictionaries:** Key-value pairs.

NumPy: Numerical Operations on Arrays

- **NumPy Arrays:** Multi-dimensional, homogeneous data structures for numerical computation.
- **Operations:**
 - **Element-wise operations:** Add, subtract, multiply, divide arrays.
 - **Array manipulation:** reshaping, slicing, indexing.
 - **Statistical operations:** mean, median, standard deviation

Pandas: Data Manipulation and Analysis

- **Pandas DataFrames:** Tabular data structure.
 - **DataFrame Operations**
 - **Selecting Columns**
 - **Filtering Data**
 - **Handling Missing Data**
 - **Data Grouping**