



# Valdymas intelligence

DATA ANALYSIS WITH PYTHON

[www.valdymasafrica.org/ict](http://www.valdymasafrica.org/ict)

## **Welcome to the Comprehensive Python Programming and Data Analysis Course!**

Embark on a transitional 11-week exploration where the power of Python programming converges with the art of data analysis. In this dynamic course, you will not only master the fundamentals of Python but also acquire the skills to decipher, analyze, and visualize data effectively.

### **Course overview**

In an era where data reigns supreme, the ability to harness the potential within it is a vital skill. This course is meticulously designed to take you from a Python novice to a proficient data analyst. Each week unveils new layers of knowledge, blending theory with hands-on projects to ensure a well-rounded understanding.

VALDYMAS INTELLIGENCE

## What You'll Learn:

1. **Python Programming Basics:** Lay a solid foundation in Python, covering everything from syntax to control flow, and set up your environment for seamless development.
2. **Data Analysis with Python:** Dive into the world of data analysis, exploring the intricacies of importing, cleaning, and visualizing data. Develop a keen eye for patterns and insights that drive informed decision-making.
3. **Advanced Techniques:** Elevate your Python prowess with advanced programming concepts, functions, loops, and conditional statements. Leverage Python's robust data structures and NumPy for nuanced numerical analysis.
4. **Data Cleaning and Preprocessing:** Tackle real-world challenges as you navigate through handling missing data, outliers, and text data analysis. Merge datasets seamlessly and ensure data integrity.
5. **Advanced Data Visualization:** Master the art of visual storytelling with Matplotlib, Seaborn, Bokeh, and Plotly. Craft interactive dashboards and geovisualizations that breathe life into your data.
6. **Statistical Analysis:** Uncover the statistical foundations of data analysis, from descriptive statistics to hypothesis testing and linear regression. Apply statistical models to draw meaningful conclusions from datasets.
7. **Final Project:** Synthesize your skills in a real-world data analysis project. Present your findings, receive peer reviews, and cap off the course by training a simple machine learning model for classification.

## Why This Course:

This course is not just about programming; it's about empowering you to make sense of the vast sea of data around us. Whether you're a novice or have some programming experience, our structured approach ensures that you build a solid foundation and gradually advance to complex data analysis techniques.

Join us on this exciting journey, where theory meets application, and programming proficiency converges seamlessly with data analysis mastery. Let's unlock the full potential of Python together and navigate the data-driven landscape with confidence and skill!

## Week 1-2:

### Introduction to Python Programming & Python for Data Analysis Libraries

- Introduction to data analysis: What is data analysis? Why is it important?
- Setting up your Python environment:
- Installing Python and Jupyter Notebook
- Introduction to Python basics: Data types, variables, operators, expressions, control flow
- Working with data in Python: Importing data from CSV files, exploring and cleaning data
- Data visualization with Python: Matplotlib and Seaborn for creating charts and graphs

**Project:** Analyze a personal datasets (e.g., your music library, spending habits, fitness tracker data) and create visualizations to tell a story about it.

## Week 3-4:

### Python Programming for Data Analysis

- Functions in Python: Defining and using functions, passing arguments, scope.
- Loops in Python: For loops and while loops for iterating over data.
- Conditional statements in Python: if, elif, and else statements for making decisions.
- Data structures in Python: Lists, tuples, dictionaries, sets for organizing data.
- NumPy for numerical data analysis: Arrays, operations, linear algebra.

**Project:** Build a Python script to automate a repetitive data cleaning task (e.g., removing special characters from text data, filling in missing values).

<p><b>Week 5-6:</b></p> <p><b>Data Cleaning and Preprocessing</b></p>	<ul style="list-style-type: none"> <li>● Missing values: Identifying and handling missing data.</li> <li>● Outliers: Detecting and dealing with outliers.</li> <li>● Data wrangling: Combining and merging datasets, transforming data.</li> <li>● Text data analysis: Cleaning and processing text data.</li> <li>● Regular expressions: Finding patterns in text data.</li> </ul> <p><b>Project:</b> Combine and merge two datasets from different sources, ensuring data compatibility and addressing any inconsistencies.</p>
<p><b>Week 7-8:</b></p> <p><b>Advanced Data Visualization</b></p>	<ul style="list-style-type: none"> <li>● Advanced data visualization with Matplotlib and Seaborn: Customizing plots, subplots, statistical plots</li> <li>● Data storytelling: Using visualizations to communicate insights from data</li> <li>● Interactive data visualization: Bokeh and Plotly for creating interactive plots</li> <li>● Geovisualization: Visualizing data on maps</li> </ul> <p><b>Project:</b> Create an interactive dashboard using Bokeh or Plotly to explore and compare results from a statistical analysis.</p>
<p><b>Week 9-10:</b></p> <p><b>Statistical Analysis with Python</b></p>	<ul style="list-style-type: none"> <li>● Descriptive statistics: Measures of central tendency and dispersion</li> <li>● Inferential statistics: Hypothesis testing, t-tests, chi-square tests</li> <li>● Linear regression: Modeling relationships between variables</li> <li>● Time series analysis: Analyzing data over time</li> </ul>

	<p><b>Project:</b> Apply hypothesis testing to compare two groups within a datasets and draw conclusions based on the statistical significance.</p>
<p><b>Week 11:</b></p> <p><b>Final project &amp; review</b></p>	<ul style="list-style-type: none"> <li>● Application of learned concepts in a real-world data analysis project</li> <li>● Project presentation and peer review</li> <li>● Review of key concepts and tools covered in the course</li> </ul> <p><b>Project:</b> Train a simple machine learning model (e.g., K-nearest neighbors) to classify data points based on their features.</p> <p><b>Final Project:</b></p>

VALDYMAS INTELLIGENCE