



Python for Data Analysis: A Multi-Disciplinary Approach

Course Description:

This 12-week, intensive live online course introduces undergraduate students from all academic backgrounds to the power of programming for data analysis. Using Python, one of the world's most popular and accessible languages, students will learn to collect, clean, analyze, and visualize data. The course takes a practical, project-based approach, using real-world datasets from various fields (scientific, social, business, and artistic) to demonstrate universal data analysis concepts. No prior programming experience is required.

Target Audience: Undergraduate students from Science, Arts, Commerce, and other faculties.

Duration: 3 Months (12 weeks)

Format: 2 live interactive classes per week (90-120 minutes each), with one practical assignment after every two classes (bi-weekly).

Key Outcome: Students will complete the course with a portfolio of data analysis projects and a final capstone project analyzing a dataset of their choice from their field of interest.

Weekly Breakdown & Class Schedule

Month 1: Python Foundations & Data Wrangling

Week 1: Why Python? Kickstarting Your Coding Journey

- Live Class 1: Introduction to Data Analysis & Python Setup
 - The data analysis workflow: Ask, Prepare, Process, Analyze, Share, Act.
 - Why Python? Introduction to Jupyter Notebooks.
 - Setting up your environment (Google Colab recommended for ease).
 - Your first program: Print, comments, and basic operations.
- Live Class 2: Python Basics - Variables and Data Structures
 - Variables and data types (Integers, Floats, Strings, Booleans).
 - Introduction to essential data structures: Lists and Dictionaries.
 - Basic operations on lists (indexing, appending).
- Assignment 1 (Due Week 2): "Data Biography." Write a Jupyter Notebook that uses variables and lists to store and display information about yourself (e.g., name, courses, hobbies). Perform basic operations like calculating an average or concatenating strings.



Research Innovation & Thesis Center

Dhanmondi 9/A, Dhaka-1209, Bangladesh

Week 2: Data Structures & Control Flow

- Live Class 3: Controlling Your Code - Loops and Conditionals
 - Making decisions: if, elif, else statements.
 - Automating repetition: for loops to iterate over lists.
 - Practical examples: filtering lists, simple data categorization.
- Live Class 4: Organizing Data - More Data Structures & Functions
 - Introduction to Tuples and Sets.
 - What are functions? Using built-in functions (len(), max(), type()).
 - Writing your own simple functions with def.
- Assignment 2 (Due Week 3): "Data Analyzer." Create a notebook that takes a list of numbers (e.g., daily screen time in minutes) and uses loops and conditionals to calculate statistics like total, average, and identify high/low usage days.

Week 3: The Data Analyst's Toolkit: Pandas I

- Live Class 5: Introduction to Pandas & DataFrames
 - What is the Pandas library? The DataFrame object (the core of data analysis in Python).
 - Loading data from CSV and Excel files into a DataFrame.
 - Viewing your data: head(), tail(), info(), describe().
- Live Class 6: Inspecting and Selecting Data
 - Selecting columns and rows (slicing, filtering with conditions).
 - Creating new columns.
 - Hands-on with a dataset (e.g., a simple sales record or survey data).
- Assignment 3 (Due Week 4): "Dataset Explorer." Load a provided dataset (e.g., movie ratings, product inventory, simple weather data). Write code to inspect its shape, display summary statistics, and answer specific questions by selecting/filtering rows.



Research Innovation & Thesis Center

Dhanmondi 9/A, Dhaka-1209, Bangladesh

Week 4: Taming Your Data: Pandas II

- Live Class 7: Data Cleaning I - Handling Missing Data & Duplicates
 - The reality of messy data.
 - Finding and handling missing values (isnull(), dropna(), fillna()).
 - Finding and removing duplicates.
- Live Class 8: Data Cleaning II - String Manipulation & Data Types
 - Basic string operations: changing case, replacing substrings.
 - Converting data types (e.g., string to numeric).
 - A complete data cleaning workflow on a messy dataset.
- Assignment 4 (Due Week 5): "The Data Cleaner." Given a deliberately messy dataset (with missing values, duplicates, and formatting issues), students will write a script to clean it and prepare it for analysis.

Month 2: Statistics, Visualization & Deeper Analysis

Week 5: The Language of Data: Descriptive Statistics

- Live Class 9: Statistics for Data Analysis I
 - Measures of Central Tendency: Mean, Median, Mode (and when to use which).
 - Measures of Spread: Range, Variance, Standard Deviation.
 - Calculating these using Pandas and NumPy.
- Live Class 10: Statistics for Data Analysis II
 - Understanding distributions.
 - Introduction to correlation: What does it mean for two variables to be related?
 - Calculating correlation coefficients in Pandas.
- Assignment 5 (Due Week 6): "Statistical Profile." Perform a full descriptive statistical analysis on a dataset. Calculate and interpret mean, median, standard deviation, and correlations between two key variables.

Week 6: Telling Stories with Data: Visualization I

- Live Class 11: Introduction to Visualization with Matplotlib
 - The importance of data visualization.
 - Creating basic plots: Line charts, Scatter plots, Bar charts.
 - Customizing plots: adding titles, labels, and legends.
- Live Class 12: Advanced Visualization with Seaborn
 - Introduction to the Seaborn library for statistical graphics.
 - Creating elegant histograms and boxplots to understand distributions.
 - Creating heatmaps for correlation matrices.
- Assignment 6 (Due Week 7): "The Data Storyteller." For a given dataset, create a series of 3-4 different plots (e.g., bar chart, histogram, scatter plot) that reveal key insights about the data. Write a brief interpretation for each plot.

Week 7: Grouping and Aggregating Data

- Live Class 13: The GroupBy Operation
 - The "Split-Apply-Combine" paradigm.
 - Using `groupby()` to segment your data (e.g., sales by region, grades by department).
 - Performing aggregations: `.sum()`, `.mean()`, `.count()` on groups.
- Live Class 14: Pivot Tables in Pandas
 - Creating pivot tables for multi-dimensional analysis.
 - Practical examples: analyzing survey data, sales performance.
- Assignment 7 (Due Week 8): "Comparative Analysis." Use `groupby` and/or pivot tables to analyze a dataset by categories. For example, compare average income by field of study, or product sales by month and category.



Research Innovation & Thesis Center

Dhanmondi 9/A, Dhaka-1209, Bangladesh

Week 8: Connecting to Data Sources & Capstone Kick-off

- Live Class 15: Getting Data from the Web
 - Introduction to APIs (Application Programming Interfaces).
 - Reading data directly from a URL into a Pandas DataFrame.
 - A simple example (e.g., fetching stock data or public COVID-19 data).
- Live Class 16: Capstone Project Introduction & Ideation
 - Review of the full data analysis pipeline.
 - Brainstorming project ideas: Finding a dataset from your own field.
 - Q&A and project proposal guidance.
- Assignment 8 (Due Week 9): Capstone Project Proposal. Submit a one-page proposal outlining the dataset you will use, the research questions you want to answer, and the analysis/visualizations you plan to perform.

Month 3: Advanced Topics & Capstone Project

Week 9: Introduction to Data Collection & Web Scraping

- Live Class 17: Introduction to Web Scraping
 - The ethics of web scraping.
 - Introduction to the requests and BeautifulSoup libraries.
 - Scraping data from a simple HTML table.
- Live Class 18: Working with Text Data
 - Introduction to text as data (highly relevant for Arts/Commerce).
 - Simple text analysis: word counts, finding frequent words.
 - Applying this to a column of text in a DataFrame (e.g., product reviews).
- Assignment 9 (Due Week 10): "Mini-Scraper or Text Analyzer." Option 1: Scrape a simple table from a webpage (e.g., Wikipedia) and load it into a DataFrame. Option 2: Perform a basic word frequency analysis on a set of product reviews or news headlines.



Research Innovation & Thesis Center

Dhanmondi 9/A, Dhaka-1209, Bangladesh

Week 10: Time Series Analysis Basics

- Live Class 19: Working with Dates and Times
 - Handling date/time data in Pandas.
 - Converting to datetime objects.
 - Setting a datetime index.
- Live Class 20: Analyzing Trends Over Time
 - Resampling time series data (e.g., converting daily to monthly data).
 - Plotting time series trends.
 - Practical example: stock prices, weather data, sales over time.
- Assignment 10 (Due Week 11): "Trend Spotter." Analyze a time-series dataset (e.g., stock prices, monthly sales). Plot the trend and use resampling to visualize the data at different time granularities (e.g., daily vs. weekly).

Week 11: Capstone Project Work Sessions

- Live Class 21: Capstone Workshop I - Data Wrangling & EDA
 - Guided work session. Instructor available for Q&A.
 - Focus on loading, cleaning, and performing exploratory data analysis (EDA) on capstone data.
- Live Class 22: Capstone Workshop II - Visualization & Storytelling
 - Guided work session.
 - Focus on creating meaningful visualizations and refining the narrative for the final presentation.
- Assignment: Dedicated time for Capstone Project work. No separate assignment.



Research Innovation & Thesis Center

Dhanmondi 9/A, Dhaka-1209, Bangladesh

Week 12: Final Presentations & Course Wrap-Up

- Live Class 23: Student Capstone Presentations (Part 1)
 - Students present their final projects (5-minute presentations + Q&A).
- Live Class 24: Student Capstone Presentations (Part 2) & Next Steps
 - Students present their final projects.
 - Course wrap-up: How to continue your data analysis journey.
 - Introduction to potential next steps (SQL, Machine Learning, etc.).
- Final Assignment: Capstone Project Submission. Due at the end of Week 12. Includes the Jupyter Notebook, a summary report, and the presentation slides.

Course Materials & Tools

- Primary Tool: Google Colab (No setup required, runs in the browser) or Jupyter Notebook with Anaconda.
- Libraries: Python, Pandas, NumPy, Matplotlib, Seaborn, BeautifulSoup.
- Datasets: A variety of curated datasets from Kaggle, Google Dataset Search, and other public repositories will be provided, covering topics like economics, social sciences, entertainment, and science.

Assessment & Certification

- Certificate of Completion: Awarded to students who submit at least 9 out of 11 assignments.
- Certificate of Excellence with Distinction: Awarded to students who demonstrate outstanding work in all assignments and the final capstone project.
- Assessment is based on the understanding of concepts and the ability to apply them, rather than perfect code syntax.

This syllabus is designed to build confidence and practical skills week-by-week, ensuring that every student, regardless of their major, can finish the course feeling empowered to use data to answer questions in their own field. Good luck with the course