T A PAI MANAGEMENT INSTITUTE(TAPMI), MANIPAL

DATA SCIENCE FUNDAMENTALS

Programme: MBA-AIDS

Batch: 2025 - 2027

Term: 1

Course Name: Data Science Fundamentals – ADS 5703

Credits: 3 (24 Sessions)

Course Instructor: Mr.Devi Prasad M (MSIS)

PART 1

INTRODUCTION

This course provides MBA students with a foundational understanding of data science, emphasizing the practical application of Python programming in a business context. Students will gain proficiency in core libraries such as NumPy and Pandas, enabling efficient data manipulation and analysis.

The curriculum covers essential topics including data cleaning, exploratory data analysis, data visualization, and introductory predictive modeling. Instruction is delivered through a combination of lectures and hands-on exercises utilizing real-world business datasets, allowing students to develop the skills necessary to extract meaningful insights.

By the conclusion of the course, students will be equipped to employ Python-based analytical methods to support data-driven decision-making and contribute to strategic business initiatives.

COURSE OBJECTIVES (CO)

1. **Introduction to Data Science Tools**: Learn to use Python libraries like NumPy and Pandas Data Frames for data manipulation and analysis.

2. **Understand Key Data Science Concepts**: Gain proficiency in data cleaning, exploration, visualization, and predictive modelling techniques.

3. **Apply Data Science to Business**: Develop the ability to analyse real-world business datasets and extract actionable insights for strategic decision-making.

COURSE LEARNING OUTCOMES (CLO)

By virtue of the learning gained in this course, the student should be able to:

CLO 1: **Proficiently Use Python for Data Analysis**: Apply Python libraries such as NumPy and Pandas to clean, manipulate, and analyse business data. **(PLG 1)**

CLO 2: **Execute Data Science Techniques**: Demonstrate the ability to perform data exploration, visualization, and build predictive models using real-world datasets. **(PLG 2)**

CLO 3: **Make Data-Driven Decisions**: Utilize data science insights to inform strategic decisions and solve business problems effectively. **(PLG 3)**

COURSE CONTENT

- Programming in Python language.
- Using NumPy and pandas for data processing.
- Introduction to Matplotlib and Seaborn for data visualization.
- Using Scikit for applying basic machine learning techniques.
- Using Python for automation.

PRESCRIBED TEXTBOOK

1. Python for Data Analysis (3rd edition): https://wesmckinney.com/book/

2. Automate the Boring Stuff with Python: https://automatetheboringstuff.com/

   - Free online version is available at the link provided above

OTHER READINGS AND REFERENCES

1. NumPy Learn (Official website): https://numpy.org/learn/

2. Pandas (Official website): pandas - Python Data Analysis Library

3. Kaggle – Learn Pandas: https://www.kaggle.com/learn/pandas/

4. Data Cleaning: Learn Data Cleaning | Kaggle

5. Numerical Programming with Python: https:/python-course.eu/numerical-programming

6. https://wesmckinney.com/book/data-aggregation#groupby_aggregation)

7. Automate the Boring Stuff with Python: https://automatetheboringstuff.com/

POLICY ON PLAGIARISM:

A similarity index scores up to 10% in all evaluation components, where applicable, is permitted. Any score beyond the permitted range will invoke section 8.4 (Academic integrity) of the PGP handbook.

ASSESSMENT SCHEME AND WEIGHTAGE:

{a} MID-TERM & END-TERM

Evaluation Component	Evaluation	Weightage (%)	Duration (in minutes)	Open/Closed Book	CLO Tested
Mid-Term	Mid-Term	20	75	Lab-based, Closed Book examination	CLO 1 & CLO 2
End-Term	End-Term	40	75	Lab-based, Closed Book examination	CLO 2 & CLO 3

(b) OTHER ASSESSMENTS

Sl. No.	Evaluation Component	Evaluation Item*	Unit of Evaluation	Weight	Time	CLO
1	Quiz 1	Quiz 1	Individual	10%	After Session 7	CLO 1 & CLO 2
2	Quiz 2	Quiz 2	Individual	10%	After session 16	CLO 1 & CLO 2
3	Assignment	Assignment 1	Group of 3	10%	After session 10	NA
4	Assignment	Assignment 2	Group of 3	10%	After session 16	NA

PART 2: SESSION PLAN

Pedagogy: Every session is an *active learning* session. The students work with shared code snippets, programming and discussing as the lecture evolv**es.**

Session.	Topic(s)	Problem/Case
1	Installation and environment setup Introduction to Python programming language Data model – objects, values, and types. Execution model. Using read-eval-print loop (REPL), command-line interface (CLI), and Google Collab.	Group Messaging System. Double Entry Ledger.
2	Tuples, lists, sets, dictionaries, defaultdict. Standard data objects and property access. Using standard libraries – os, sys, datetime, math, decimal, secrets, json, csv	Group Messaging System. Double Entry Ledger.
3-4	Statement-level Structured Programming. Statements, Compound statements, Expressions. Building a larger program using lists, sets, and dictionaries. Logging program execution.	Group Messaging System. Double Entry Ledger
5-6	More expressions for clean coding. Building a larger program using more powerful expressions.	Group Messaging System. Double Entry Ledger.
7	Program organization and modularity – Part 1. Function-level Structured Programming.	Group Messaging System. Double Entry Ledger.
8-9	Program organization and modularity – Part 2 (Object-oriented Programming). Hiding state and implementation details Defining good interfaces.	Group Messaging System. Double Entry Ledger.
10-13	Project definition – 10 groups, each of three students. NumPy and Matplotlib and seaborn. Visualizing time series through aggregation. Arrays, ndarrays, shape of arrays and reshaping. Vector arithmetic – distance, similarity. Tensors – broadcasting and vectorized operations. Device detection and selection.	

Session.	Topic(s)	Problem/Case
14-17	Pandas and Matplotlib Data Cleaning and Preparation. Join, combine and reshape. Aggregation and grouping. Pivot tables.	
18-21	Machine learning - Introducing Scikit-Learn. Naive Bayes Classification. Decision Trees and Random Forests. Project presentations – last three hours.	Data classification and prediction.
22-24	Project presentations – last three hours	

* Please refer to 'Other readings & references' as per the indicators a, b, c & d listed above the plagiarism policy note.

PLG MAPPING FOR THE COURSE

PLG#	Program Level Learning Goal	Addressed by Course? (Yes / No)	CLO No.
PLG1	**Application of Fundamentals** *Traits:* Demonstrate application of functional / conceptual knowledge to business situations	**Yes**	**CLO 1**
PLG2	**Problem Identification and Solution** *Traits:* Demonstrate ability to identify a problem, critically assess various alternatives and suggest appropriate solution	**Yes**	**CLO 2**
PLG3	**Integrative Thinking** *Traits:* Demonstrate ability to identify inter-linkages among functional areas within an enterprise and assess the impact of external environment on its performance	**Yes**	**CLO 3**
PLG4	**Effective Communication** *Traits:* Demonstrate proficiency in Oral and Written Communication	**No**	**NA**
PLG5	**Ethical Responsibility** *Traits:* Demonstrate awareness and assess impact of ethical behavior on business	**No**	**NA**
PLG6	**Leadership** *Traits:* Demonstrate capability to take leadership role in a business situation	**No**	**NA**

*****

Name of Faculty: Mr. Devi Prasad M

Signature & Date: _____

Endorsed by: Prof. Gaurav Sarin

Signature & Date: _____