Lok Jagruti Kendra University
University with a Difference

Masters of Business Administration (MBA) – Semester - 3
Course Teaching Plan

Course Title: **Data Processing and Analysis using Python**
Course Code: **340039311**
Course Credit: **4.0**

1. **Course Introduction:**
   Data analytics is the science of analyzing raw data to draw insights for business decision making. It relies on a variety of software tools and open-source languages for data manipulation. Python is an object-oriented open-source programming language which has many in-built mathematical functions and libraries for data manipulation, data visualization, and data modelling, making it easier to calculate mathematical problems and to perform advance data analysis. In current times, there is a growing need for hands on orientation to an advanced analytical application so that the students can be better equipped to work in corporate world.

2. **Course Objective:**

- Describe Python Basics including Data Types, Expressions, Variables, and Data Structures.
- Apply Python programming logic using Loops, Functions
- Develop Python code for cleaning and preparing data for analysis - including handling missing values, formatting, normalizing, binning data and data smoothing
- Perform exploratory data analysis and apply analytical techniques to real-word datasets using libraries such as Pandas, Numpy and Matplotlib

3. **Course Learning Outcome:**

CLO1: Understand Python Basics including Data Types, Expressions, Variables, and Data Structures.
CLO2: Apply Python programming logic using Branching, Loops, Functions
CLO3: Develop Python code for cleaning and preparing data for analysis - including handling missing values, formatting, normalizing, binning data and data smoothing
CLO4: Demonstrate proficiency in using Python libraries such as Pandas, Numpy and Matplotlib

4. **CLO –PO Mapping Matrix**

	PO1	PO2	PO3	PO4	PO5
CLO1	3	2	3	-	2
CLO2	2	2	3	-	2
CLO3	3	3	3	-	2
CLO4	3	3	3	2	2

Correlation levels: 3= ‘High’, 2=‘Medium’, 1=‘Low’ and ‘-‘ = No correlation

5. Course Contents & Session Plan:

Session No.	Topic
Unit I - Python fundamentals	
1-2	Python Basics including Expressions, Variables, input and output operations
3-4	Python data structures - use of lists, tuples, dictionaries and sets and their inbuilt methods
5-9	Python programming fundamentals- - conditions - loops - functions - Python libraries such as Pandas and Numpy.
10-12	Reading and writing text and binary files. Loading Data with Pandas, creating and manipulating Dataframes using Pandas, Mathematical and statistical operations with Numpy
Unit II – Data Cleaning	
13-15	Data Extraction - from xlsx, csv, pdf, HTML
16-24	Data Cleaning – handling missing data, wrong data type, mixed data type in same variable, cleaning junk characters, wrong spellings, abbreviations, outlier analysis
Unit II – Data transformation and Visualization	
25-30	Data transformation - data encodings, normalisation and standardisation of data
31-33	Data Visualization using Matplotlib
34-40	Project: Analysis of dataset provided

6. **Assessment Scheme :**

Exam: 50% **Continuous Assessment 50 %**

Specific assessment method	% Weightage	Theory	Practical
Exam	50%	✓	✓
Case Analysis/Class Participation	10%		✓
LAB Assignment	30%		✓
Quiz	10%	✓	✓

7. Educational Resources

Educational Resources	Description
i. Text Book	
ii. Reference Book	1. Introduction to Computation And Programming Using Python, John V. Guttag, PHI Publication, Latest edition,

	2. Python Data Science Handbook, Jake VanderPlas, 2nd Edition, O'Reilly Media, Inc., ISBN: 9781098121228
iii. Blogs/ Magazine/periodical	**Data cleaning :** -https://www.analyticsvidhya.com/blog/2021/06/data-cleaning-using-pandas/ -https://www.mygreatlearning.com/blog/data-cleaning-in-python/#what-is-data-cleaning **Data transformation and visualization:** - https://byuidatascience.github.io/python4ds/transform.html - https://www.analyticsvidhya.com/blog/2020/03/understanding-transform-function-python/
iv. Video lecture	
v. Course related important Web links	- IBM- Python for Data Science, AI & Development https://www.coursera.org/learn/python-for-applied-data-science-ai#modules - IBM Data Analysis with Python https://www.coursera.org/learn/data-analysis-with-python