

Correlation Levels:

Level	Correlation
-	Nil
1	Slightly / Low
2	Moderate / Medium
3	Substantial / High

Assessment Rubrics:

- Quiz / Assignment/ Quiz/ Discussion / Seminar
- Midterm Exam
- Programming Assignments
- Final Exam

Mapping of COs to Assessment Rubrics :

	Internal Exam	Assignment	Seminar	End Semester Examinations
CO 1	✓	✓		✓
CO 2				✓
CO 3	✓		✓	✓
CO 4		✓	✓	✓

15. DATA SCIENCE USING PYTHON

Discipline	COMPUTER SCIENCE
Course Code	UK2MDCCSC104
Course Title	DATA SCIENCE USING PYTHON

Type of Course	MDC				
Semester	II				
Academic Level	1 -				
Course Details	Credit	Lecture per week	Tutorial per week	Practical per week	Total Hours/Week
	3	2 hours	-	2	4 hours
Pre-requisites	Basic Knowledge about Programming and Computer Technologies				
Course Summary	This course will help to learn the basics of Python along with different techniques in data science.				

Detailed Syllabus:

Module	Unit	Content	Hrs (L+P)
I	Basics of Python Programming		12
	1	Introduction to Python, Python tokens, Literal constants, Type conversion	
	2	Variables and identifiers, Creating variables	
	3	Data types of identifiers, Input operation	
	4	Comments, Reserved words, Indentation	
	5	Operators and expressions, Expressions in Python, Decision control statements	
	6	Selection/conditional branching statements, Iterative statements, The range () function, Selecting an appropriate loop, Nested loops	
	7	Break statement, The continue statement, The pass statement, The else statement used with loops	
II	Basic Data structures in Python		12
	8	Revisiting Data Structures in Python, Introduction to Python strings, String indexing, Finding the number of characters in a string, Traversing a string, Concatenating, appending and multiplying strings, The str() function, Strings are immutable, String formatting operator, The format() function	

	9	Built-in string methods and functions, Comparing strings, ord() and chr() functions, In and not in operator	
	10	Lists, Accessing values in lists, The eval() function, Updating values in lists, Relational operations on lists, Nested lists, List aliasing and cloning, Deleting elements, Deep copies and shallow copies in Python, Basic list operations, List methods	
	11	Tuple, Creating tuple. Utility of tuples, Accessing values in a tuple, Updating tuple, Deleting elements in tuple, Joining tuples, Unpacking tuples, Basic tuple operations, Tuple assignment, Accessing using index, Tuples for returning multiple values, Nested tuples, The count() method, The zip() function, Advantages of tuple over list	
III		Dictionaries and functions in Python	12
	12	Dictionaries, Creating dictionary, Accessing values in a dictionary, Adding an item in a dictionary, Modifying an item in a dictionary, Deleting items, Traversing a dictionary, Nested dictionaries, The copy() method	
	13	Built-in dictionary functions and methods, Difference between a list and a dictionary.	
	14	Function declaration and function definition, Function definition, Function call, Function parameters, Parameter passing mutable/immutable properties, The return statement, Types of function parameters, Passing strings, lists, tuples, dictionaries to functions, Modules	
IV		Data Handling Using Numpy and Python Pandas	12
	15	Data and its purpose, Data science and its applications, The numpy module, Creating numpy arrays, Array attributes, Converting 2D numpy array into 1D array	
	16	Array slicing: Accessing subarrays, Reshaping of arrays, Array concatenation (joining) and splitting	
	17	How numpy broadcasting works, Performing mathematical operations on numpy arrays, Transposing arrays, Inserting and deleting array elements	
	18	Find the index of a value, Sorting a numpy array, Normalize array, Array subsets	
	19	Python Pandas, Data frame, Pandas data frame functions and attributes, Pivoting data frame, Sorting, Missing data, Combining data frames	

	20	Descriptive statistics, Summarizing or describing data, Function application, Aggregation (group by), Transform function in Python, Reindexing in Pandas dataframe, Altering column labels, Data wrangling, Time series data structures	
V	Flexi Module: Not included for End Semester Exams		12
	21	Plotting Graphs, Importance of data visualization, Bar chart, Plotting histograms, Frequency polygon, Box plot, Scatter plot, Correlation matrix plot, The Seaborn library, The color palette, Plotting univariate distribution, Plotting bivariate distribution	
	22	Visualizing pairwise relationship, Box Plot in Seaborn, Violin plots, Statistical estimation, Plotting categorical data, Facet grid and facetgridmap(), Pair grid, Linear relationships, Heatmap, Bubble chart, Plotting time series data, Visualizing sparse matrix	

References

1. Dr. Reema Thareja, Data Science and Machine Learning using Python, McGraw Hill Education (India) Private Limited

Lab Exercises

A Questions on Python based on the syllabus related to Data Science

1. Programs using Python strings, lists, tuples, and dictionaries.
2. Read and write data from/to files in Python.
3. Programs to demonstrate creating and handling of modules and packages
4. Programs involving regular expressions
5. Programs to draw simple bar chart, pie chart, histogram and scatter plot
6. Create a python program to draw a Histogram, Column Chart, Box plot chart, Pie Chart, and Scatter plot using pandas and mat plot lib.
7. Create a python program to export data (store Data Frame in CSV Format)
8. Create a python program to handle the missing data from a dataset using numpy and pandas.
9. Create a python program to import data from any .csv file and analyze using the statistical functions of pandas tools
10. Programs using Python strings, lists, tuples, and dictionaries.

Course Outcomes

No.	Upon completion of the course the graduate will be able to	Cognitive Level	PSO addressed
-----	--	-----------------	---------------

CO-1	Develop knowledge on Python	Ap	PSO-1,2,3
CO-2	Identify basic Data structures in Python	Ap	PSO-1,2,3
CO-3	Use Dictionaries and functions in Python	Ap	PSO-1,2,3
CO-4	Manipulate Data Using Numpy and Python Pandas	Ap	PSO-1,2,3

R-Remember, U-Understand, Ap-Apply, An-Analyse, E-Evaluate, C-Create

Note: 1 or 2 COs/module

Name of the Course: DATA SCIENCE USING PYTHON

Credits: 2:0:1 (Lecture: Tutorial: Practical)

CO No.	CO	PO/PSO	Cognitive Level	Knowledge Category	Lecture(L)/ Tutorial (T)	Practical (P)
CO-1	Develop knowledge on Python	PO-1, 2, 6,7 PSO-1,2, 3	Ap	F, C, P	L	P
CO-2	Identify basic Data structures in Python	PO-1, 2, 6, 7 PSO-1,2, 3	Ap	F, C, P	L	P
CO-3	Use Dictionaries and functions in Python	PO-1, 2, 6, 7 PSO-1, 2, 3	Ap	F, C, P	L	P
CO-4	Manipulate Data Using Numpy and Python Pandas	PO-1, 2, 6, 7 PSO-1,2, 3	Ap	F, C, P	L	P

F-Factual, C- Conceptual, P-Procedural, M-Metacognitive

Mapping of COs with PSOs and POs :

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PSO 1	PSO 2	PSO 3	PSO4
CO 1	1	1	-	-	-	2	2	-	3	1	2	-
CO 2	2	1	-	-	-	2	2	-	3	2	2	-
CO 3	2	1	-	-	-	2	2	-	3	2	2	-
CO 4	2	2	-	-	-	2	2	-	3	2	2	-

Correlation Levels:

Level	Correlation
-	Nil
1	Slightly / Low
2	Moderate / Medium
3	Substantial / High

Assessment Rubrics:

- Quiz / Assignment/ Quiz/ Discussion / Seminar
- Midterm Exam

- Programming Assignments
- Final Exam

Mapping of COs to Assessment Rubrics :

	Internal Exam	Assignment	Lab Assignment	End Semester Examinations
CO 1	✓	✓	✓	✓
CO 2	✓		✓	✓
CO 3	✓		✓	✓
CO 4	✓	✓	✓	✓