

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

#### Assessment Rubrics:

- Quiz / Assignment/ Discussion / Seminar
- Trivial content creation assignments
- Midterm Exam
- Final Exam

#### Mapping of COs to Assessment Rubrics:

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

### 7. PYTHON FOR DATA ANALYTICS

Discipline	COMPUTER SCIENCE
Course Code	UK2DSCCSC106
Course Title	PYTHON FOR DATA ANALYTICS

Type of Course	DSC/ VOCATIONAL				
Semester	II				
Academic Level	1				
Course Details	Credit	Lecture per week	Tutorial per week	Practical per week	Total Hours/Week
	4	3 hours	-	2 hours	5 hours
Pre-requisites	-				
Course Summary	This course is designed to enable students to get familiar with the features of python, its libraries, module creation, implementing various data structures and data visualization .				

### Detailed Syllabus:

Module	Unit	Content	Hrs (L + P)
<b>I</b>	<b>Introduction</b>		<b>15</b>
	1	Data Analytics Lifecycle overview – Discovery,Data Preparation,Model Planning, Model Building,communicate results, operationalize.	
	2	Features of Python, Variables, output, input in Python, Operators ,Control flow statements: Decision making structures, Loops, Nesting of conditional statements and loops, abnormal loop termination	
	3	Functions: uses, syntax, Types – built in and user-defined functions, String functions in python. Recursive function	
	4	Errors and Exception handling	
<b>II</b>	<b>Data Structures</b>		<b>15</b>
	3	Data Types in Python- Numeric, Dictionary,Boolean, Set,Sequence type	
	4	Modules: In-built modules and user defined modules, import statement, from .... import statement.	
	5	Numpy library for arrays: One-dimensional and multi-dimensional	
<b>III</b>	<b>Data Processing</b>		<b>15</b>
	6	Pandas library for data processing	

	7	Basics of data frame, import of data, functions of data frame	
	8	Data extraction, Group by functionality	
	9	Creating charts for dataframe, missing values	
<b>IV</b>	<b>Data Visualization</b>		<b>15</b>
	10	Matplotlib library for visualization: Visualization for categorical variable, visualization of continuous variable.	
	11	Seaborn library for visualization: Visualization for categorical variable, visualization of continuous variable.	
<b>V</b>	<b>Additional Core Libraries ( Not for end semester Examination )</b>		<b>15</b>
	12	SciPy Library for Statistics	
	13	SQLAlchemy Library for SQL	
	14	StatsModels Library for time series models - Introduction	

### **TEXT BOOK**

1. Bharti Motwani, Data Analytics using Python,Wiley,2022
2. Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data,EMC Education services Wiley Publication

### **REFERENCES**

1. Joel Grus,Data Science from Scratch: First Principles with Python,O'Reilly Media,2015
2. Wes McKinney,Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython,O'Reilly Media,2017
3. Jake VanderPlas,Python Data Science Handbook: Essential Tools for Working with Data,O'Reilly Media,2016
4. Aurélien Géron,Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow,O'Reilly Media,2019
5. Python for Data Analysis: 3rd Edition, Wes McKinney ,Publisher(s): O'Reilly Media, Inc.

### **LAB EXERCISES**

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.
11. Read and write data from/to files in Python.
12. Programs to demonstrate creating and handling of modules and packages
13. Programs involving regular expressions
14. Programs to draw simple bar chart, pie chart, histogram and scatter plot
15. Create a python program to draw a Histogram, Column Chart, Box plot chart, Pie Chart, and Scatter plot using pandas and mat plot lib.
16. Create a python program to export data (store Data Frame in CSV Format)
17. Create a python program to handle the missing data from a dataset using numpy and pandas.
18. Create a python program to import data from any .csv file and analyze using the statistical functions of pandas tools
  - (a) Create a python program to draw a Histogram, Column Chart, Box plot chart, Pie Chart, and Scatter plot using pandas and mat plot lib for the following data. The categorical data on 1997 U.S. Health Care Expenditures. The data are in file healthexpendituresdata.csv.
  - (b) The monthly data on the total return from the Standard and Poor 500 stock index (with reinvestment of dividends) from 1970 to 2018. The data are in file SandP500stockpricedata.csv. Create a python program to import data from any .csv file and analyze using the statistical functions of pandas tools. Also create a python program to draw different charts.
  - (c) If at the end of each month, a saver deposited \$100 into a savings account that paid 6% compounded monthly, how much would he have at the end of 10 years? Create a python program to calculate it?

	A	B
1	Category	Expenditures
2	Hospital	371
3	Physician	218
4	Drugs and Supplies	109
5	Other Personal	92
6	Nursing Home	83
7	Dental	51
8	Admin & Insurance	50
9	Public Health	39
10	Home Health	32
11	Research	18
12	Construction	17
13	Eye and Equipment	14

(d) Draw a pie chart and other charts that shows the amount of subscription generated for Indian Bonds from different categories of Investors. Create a python program for the above problem Use pandas and mat plot lib to draw charts

(e) The share holding pattern of a company WIPRO is given. Create a python program for the above problem. Use pandas and matplotlib to draw charts

### Course Outcomes

No.	Upon completion of the course the graduate will be able to	Cognitive Level	PSO addressed
CO-1	Familiarize Data Analytics Lifecycle and Python basics	Ap	PSO-1,2,3
CO-2	Experiment with Python Data Structures and Modules	Ap	PSO-1,2,3

CO-3	Use Pandas library, data frames, and data extraction methods.	Ap	PSO-1,2, 3
CO-4	Experiment with Python libraries Matplotlib and Seaborn for data visualization of both categorical and continuous variables.	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 ANALYTICS USING PYTHON**

**Credits: 4:0:0 (Lecture:Tutorial:Practical)**

CO No.	CO	PO/PSO	Cognitive Level	Knowledge Category	Lecture (L)/Tutorial (T)	Practical (P)
1	Familiarize Data Analytics Lifecycle and Python basics	PO- 6,7 PSO-1, 2,3	Ap	F, C	T	p
2	Experiment with Python Data Structures and Modules	PO- 6,7 PSO-1, 2,3	Ap	F,C,P	L	P
3	Use Pandas library, data frames, and data extraction methods.	PO-6, 7 PSO-1, 2, 3	Ap	F,C,P	L	P
4	Experiment with Python libraries Matplotlib and Seaborn for data visualization of both	PO-6,7 PSO-1, 2,3	Ap	F,C,P	L	p

	categorical and continuous variables.					
--	---------------------------------------	--	--	--	--	--

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

**Mapping of COs with PSOs and POs :**

	PO1	PO2	PO3	PO4	PO 5	PO6	PO7	PO8	PSO 1	PSO 2	PSO 3	PSO 4
<b>CO 1</b>	-	-	-	-	-	2	2	-	2	2	2	-
<b>CO 2</b>	-	-	-	-	-	2	2	-	2	2	2	-
<b>CO 3</b>	-	-	-	-	2	2	2	-	2	2	2	-
<b>CO 4</b>	-	-	-	-	2	2	2	-	2	2	2	-

**Correlation Levels:**

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

**Assessment Rubrics:**

- Quiz / Assignment/ Quiz/ Discussion / Seminar