KU3DSCCSC203: PYTHON FOR DATA ANALYTICS

Semester	Course Type	Course Level	Course Code	Credits	Total Hours
3	DSC	200-299	KU3DSCCSC203	4 (3+1)	75

Learning Approach (Hours/ Week)			Marks Distribution			Duration of ESE (Hours)
Lecture	Practical/ Internship	Tutorial	CE	ESE	Total	
3	2		35	65	100	1.5hrs.

Course Description:

The course introduces students to the Python programming language using an embedded programming environment. The modules explore if-statements and loops to illustrate concepts of flow-control and iteration and also some data visualization tools.

Course Prerequisite: NIL

Course Outcomes:

CO No.	Expected Outcome	Learning Domains
1	Identify the basics of Python and identify the data types in Python	U
2	Design programs using various python features like operators, control structures and other python objects.	U, A
3	Differentiate the usage of functions, modules and file handling in python from other languages	U, An
4	Design programs using built-in modules	U, C
5	Apply various visualization tools in python to draw graphs and plots.	U, A

Remember (R), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C)

	PSO 1	PSO 2	PSO 3	PSO 4	PSO 5	PSO 6	PSO 7
CO 1	3	2		2			2
CO 2	2	3		2			
CO 3	3	3	2				2
CO 4	2	3	2				
CO 5	2	3	3	2			3

Course Description:

This course introduces students to the fundamental concepts and techniques of data analytics. Students will learn how to collect, clean, analyze, and visualize data using various tools and techniques. The course will cover data collection, data cleaning, data analysis, and data visualization

Course Prerequisite: NIL

Course Outcomes:

CO No.	Expected Outcome	Learning Domains
1	Identify the basics of data analytics.	U
2	Organise the methods to collect and clean data.	An
3	Examine different data analysis techniques.	A
4	Apply data visualization techniques in handling real life data effectively.	A/An
5	Judge on various data analytics concepts in real-world scenarios.	E

Remember (R), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C)

Mapping of Course Outcomes to PSOs

	PSO 1	PSO 2	PSO 3	PSO 4	PSO 5	PSO 6	PSO 7
CO 1	3			3			
CO 2	3						2
CO 3	2	2	3	3		2	2
CO 4	2		3	3			2
CO5	3		3			2	3

COURSE CONTENTS

Contents for Classroom Transaction:

MODULE	UNIT	DESCRIPTION	HOURS
1		**MODULE TITLE:** INTRODUCTION TO DATA ANALYTICS	**15**
	1	Data analytics definition-steps in data analytics-types of data analytics- Career Opportunities in Data Analytics	
	2	Importance of data analytics	
	3	Overview of data analytics process	
	4	Types of data analysis	

MODULE	UNIT	DESCRIPTION	HOURS
2		**MODULE TITLE: PROCESS OF DATA ANALYSIS**	**15**
	1	2. Define the Problem 3. Collect Data 3. Data Cleaning 4. Analyzing the Data 5. Data Visualization 6. Presenting Data	
	2	Sources of data -primary and secondary data- Data types -structured and unstructured data	
	3	Data collection methods -- Methods to collect primary and secondary data	
	4	Data storage and management -what is data management-need of data management-importance and types of data management	

MODULE	UNIT	DESCRIPTION	HOURS
3		**MODULE TITLE:** DATA CLEANING	**15**
	1	Data preprocessing techniques - Data cleaning -Dimensionality reduction -Feature engineering: Handling a large amount of data (sampling data)-Data transformation:	
	2	Handling missing data - Effective Strategies for Handling Missing Values in Data Analysis	
	3	Dealing with outliers - Trimming/Remove the outliers- Quantile Based Flooring and Capping- Mean/Median Imputation- Visualizing the Data after Treating the Outlier	

	4	Features of Python for data analysis- Syntax-Keywords in Python-Comments in Python-Python Variables	

4		**MODULE TITLE :** PANDA FOR DATA ANALYSIS	15
	1	Python Data Types-Strings-Numbers-Booleans-Python List-Python Tuples-Python Sets-Python Dictionary	
	2	Operators in Python -Arithmetic operators Comparison Operators-Logical Operators-Bitwise Operators-Assignment Operators- Control structures -branching and looping.	
	3	Pandas Library in Python-Why Python Pandas used for-data analysis Panda series.-how to create series. Data Frames- create data frame using DataFrame () pandas Read CSV-pandas analyzing data- head() and tail() functions.	
	4	Data set cleaning- removing Empty cells-dopna() -removing Data in wrong format- to_datetime()-removing Wrong data-loc() and drop(). removing Duplicates-duplicated(). Data Visualization using pandas-panda plot().	

5		**Teacher Specific Module**	
		Teacher can suggest activities on following titles	
		1. write a python code for arithmetic operations 2. write a python code for sorting a list of numbers. 3. Write a program to check whether a given number is even or odd. 4. Create a program to calculate the factorial of a given number. 5. Write a program to check if a given number is a prime number. 5. Write a Pandas program to create the mean and standard deviation of the data of a given Series. 6. Write a Pandas program to detect duplicates using duplicated() method. 7. Write a Pandas program to create a Pivot table and find the region wise total sale. 8. Write a Pandas program to split the following data frame into groups based on school code. Also check the type of GroupBy object. **Test Data:**	15

school	class		name	date_Of_Birth	age	height	weight	address	
S1	s001	V	Alberto Franco	15/05/2002	12	173	35	street1	
S2	s002	V	Gino Mcneill	17/05/2002	12		19232	street2	S3
	s003	VI	Ryan Parkes	16/02/1999	13		18633	street3	S4
	s001	VI	Eesha Hinton	25/09/1998	13	167	30	street1	
S5	s002	V	Gino Mcneill	11/05/2002	14		15131	street2	S6
	s004	VI	David Parkes	15/09/1997	12	159			

Essential Readings:

1. **Data Science for Business" by Foster Provost and Tom Fawcett**

2. **"Python for Data Analysis" by Wes McKinney**

3. **Online resources and tutorials**

Assessment Rubrics:

Evaluation Type			Marks	Evaluation Type			Marks	Total
Lecture			75	**Practical**			25	
a)	**ESE**		50	a)	**ESE**		15	
					Program code and execution		8	**100**
					Output		3	
					Viva		2	
					Modification		2	
b)	**CCA**		25	b)	**CCA**		10	
	i	Test Paper	5		i	Punctuality	3	
		Model exam	10					

	ii	Assignment/ Book-Article review /field report	5		ii	Model exam	4
	iii	Seminar/ Viva-Voce	5		iii	Record	3

KU3DSCCSC204: BASICS OF OBJECT ORIENTED PROGRAMMING USING C++

Semester	Course Type	Course Level	Course Code	Credits	Total Hours
3	DSC	200-299	KU3DSCCSC204	4	75

Learning Approach (Hours/ Week)			Marks Distribution			Duration of ESE (Hours)
Lecture	Practical/ Internship	Tutorial	CE	ESE	Total	
3	2		35	65	100	1.5hrs.

Course Description:

This course introduces the fundamentals of C++ programming with a focus on Object-Oriented Programming (OOP) concepts.

Prerequisite: NIL

Course Outcomes:

CO No.	Expected Outcome	Learning Domains