Shell Scripting and Data Analysis for Biologists Syllabus and Evaluation Scheme

SHELL SCRIPTING AND DATA ANALYSIS FOR BIOLOGISTS			35 HRS
COURSE OBJECTIVES: 1. To understand the concept of the Operating System and Command Line Handling. 2. To study the basic Programming Constructs and Algorithms. 3. To learn the handling of data using a Spreadsheet (Excel). 4. To familiarize and visualize the data using Power BI.			
COURSE OUTCOMES: The students will be able to- 1. Acquire the knowledge of file and directory handling on the Linux Command Line. 2. Develop simple shell scripts and Programming skills. 3. Analyze and visualize the features of the Spreadsheet Application (Excel). 4. Apply MS-Excel skills and Power BI in handling Big Data and its analysis.			
WEEK 1	**Day 1:**	**OPERATING SYSTEM**	**HOURS: 02**
		Concept of operating system; Kernel, Shell, Command line on Linux; Basic Shell commands for Files and Directories handling, Process handling.	
	Day 2:	**BASICS OF SHELL SCRIPTING**	**HOURS: 02**
		Shell scripting, Concept of Shell files; Variables, Decisions and loops; Input and Output; Command line input with options.	
	Day 3:	**FUNCTIONS**	**HOURS: 02**
		Functions and its Types; Passing and Parsing Arguments and Parameters; Reading Command Line Arguments, Single and Multiple.	
	Day 4:	**FILE HANDLING**	**HOURS: 02**
		File Reading; File Test Operators; Interactive and Non-interactive Shell	
	Day 5:	**REGEX HANDLING**	**HOURS: 03**
		Concept of regular expressions and pattern matching; Symbols, anchors and quantifiers for regular expressions;	

		Assignment	
WEEK 2	**Day 6:**	**ERROR HANDLING AND DEBUGGING**	**HOURS: 02**
		Error Handling; Debugging and Automation.	
	Day 7:	**SMALL PROJECT (P1) USING SHELL SCRIPTING**	**HOURS: 03**
		A Small Project Using Concepts Of Shell Scripting.	
WEEK 2	**Day 8:**	**INTRODUCTION TO SPREADSHEET (EXCEL)**	**HOURS: 02**
		Overview and Basics; Features of formatting, Basic and advanced filtration; Sorting.	
	Day 9:	**FUNCTIONS IN SPREADSHEET (EXCEL) PART 1**	**HOURS: 02**
		Text to column and column to text conversion; Basic and Advanced Formula usage.	
	Day 10:	**FUNCTIONS IN SPREADSHEET (EXCEL) PART 2**	**HOURS: 03**
		Data analysis with Formula; VLOOKUP, HLOOKUP.	
WEEK 3	**Day 11:**	**DATA ANALYSIS AND VISUALIZATION**	**HOURS: 02**
		Goal Seek; Subtotals and Pivot, Pivot Charts, Other types of Charts.	
	Day 12:	**MACRO DEVELOPMENT**	**HOURS: 02**
		Concept of Macros; Writing and Using Macros in Spreadsheet (Excel).	
	Day 13:	**SMALL PROJECT (P2) FOR DATA ANALYSIS**	**HOURS: 03**
		Small project for handling biological data using a Spreadsheet (Excel).	
	Day 14:	**POWER BUSINESS INTELLIGENCE (BI)**	**HOURS: 03**
		Dashboard Basics; Power BI service Concepts; Data Handling and Analysis.	
	Day 15:	**ASSESSMENT AND FEEBACK (P1& P2)**	**HOURS: 02**
		Assessment and Feedback for the P1 and P2.	
References			

1. Mallett, A. (2015). *Mastering Linux Shell Scripting*. Packt Publishing Ltd.
2. Parker, S. (2011). *Shell scripting: expert recipes for Linux, Bash, and more*. John Wiley & Sons.
3. Harvey, G. (2018). *Excel 2019 for dummies*. John Wiley & Sons.
4. Guerrero, H., Guerrero, R., & Rauscher. (2019). *Excel data analysis*. Springer International Publishing.

Evaluation Scheme

Course Name	Internal Assessment (IA) [40 Marks]				Project [60 Marks]		Total [100 Marks]
	Day 1-Day 7		Day 8-Day 15				
	Class Participation	Attendance	Class Participation	Attendance	P1	P2	IA+P1+P2
Shell Scripting and Data Analysis for Biologists	**10**	**10**	**10**	**10**	**30**	**30**	**100**

Requirement to qualify for Certificate:

1. Students are required to finish the coursework with a minimum of 80% attendance.
2. Students need to score a minimum of 60% marks to qualify for the Certificate.