Course Code	21CSE295T	Course Name	Statistical Data Analytics and Integration	Course Category	E	Professional Elective	L	T	P	C
							2	1	0	3

Pre-requisite Courses	*Nil*	Co-requisite Courses	*Nil*	Progressive Courses	*Nil*
Course Offering Department	*Data Science and Business Systems*	Data Book / Codes/Standards		*Nil*	

Course Learning Rationale (CLR):	*The purpose of learning this course is to:*
CLR-1 :	*To equip learners with foundational skills in SAS programming, libraries and handle data*
CLR-2 :	*To facilitate with essential skills in data manipulation, reporting, and programming in SAS*
CLR-3 :	*To analyze complex data sets for meaningful insights and informed decision-making*
CLR-4 :	*To analyze data effectively and present their findings by visualizing it*
CLR-5 :	*To represent data using various advanced visualization techniques for data analysis and presentation*

Course Learning Outcomes (CO):	*At the end of this course, learners will be able to:*	Program Learning Outcomes (PO)(1- Low, 2 - Medium, or High-3)														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		Engineering Knowledge	Problem Analysis	Design & Development	Analysis, Design, Research	Modern Tool Usage	Society & Culture	Environment & Sustainability	Ethics	Individual & Team Work	Communication	Project Mgt. & Finance	Life Long Learning	PSO - 1	PSO - 2	PSO - 3
CO-1 :	*Understand with foundational skills in SAS programming, libraries and handle data*	*3*	*1*	-	*3*	-	-	-	-	-	-	-	-	-	-	-
CO-2 :	*Demonstrate with essential skills in data manipulation, reporting, and programming in SAS*	*1*	*2*	-	*3*	-	-	-	-	-	-	-	-	-	-	-
CO-3 :	*Analyze complex data sets for meaningful insights and informed decision-making*	*3*	-	*2*	*3*	-	-	-	-	-	-	-	-	-	-	-
CO-4 :	*Create and analyze data effectively and visualize it*	*3*	-	-	*3*	*3*	-	-	-	-	-	-	-	-	-	-
CO-5 :	*Apply SAS programming for advanced visualization techniques for data analysis and presentation*	*2*	*3*	-	-	*3*	-	-	-	-	-	-	-	-	-	-

Unit 1 Introduction to Statistical Analysis **9 hours**

Computational Statistics- Statistical Pattern Recognition, Nonparametric Regression, Exploratory Data Analysis, Monte Carlo Methods for Inferential Statistics-Basic concepts of statistics, Types of data (categorical, numerical), Populations and samples, Measures of central tendency (mean, median, mode), Measures of variability (range, variance, standard deviation), Descriptive vs. inferential statistics.

Tutorial:

- SAS Statements, Structure, Libraries, Rules, Datasets, Representation of missing values, Assigning Libraries, PROC CONTENTS, visualize the descriptor and datasets attributes

Unit 2 Data Analysis Using SAS **9 hours**

Basic Reports – Selecting Variables – Identifying observations – Subsetting the Data – Limiting the observations-SORT Procedure – Generating Column totals – Specifying titles and footnotes in procedure output – Assigning Descriptive labels and permanent label -How SAS Processes Programs – Compilation Phase – PDV – Descriptor Portion-Execution Phase – Data Portion – Debugging a Data Step – Testing Programs-BY-Group Processing Definitions – Preprocessing Data – Determine Whether the Data Requires Preprocessing – Sorting Observations for BY-Group Processing-FIRST. and LAST. DATA Step Variables – How the DATA Step Identifies BY Groups – Assignment Statements – Operators in SAS Expressions

Tutorial:

- Introduction of report and PRINT Procedure-Hands on practice of VAR, ID, WHERE statement, where expression[eq, ne, gt, lt, ge, le, AND, OR, CONTAINS]

Unit 3 Advanced Statistical Techniques in SAS **9 hours**

Determining the Structure and Contents of Data Sets – Testing Program – Methods of Combining SAS Data Sets-One-to-One Reading – Concatenating – Match-Merging – Renaming Variables – Excluding Unmatched Observations -Selecting Matching Observations -Basics of DO Loops – Executing DO Loops -OUTPUT Statements – Nesting DO Loops-Iteratively Processing Observations from a Data Set -DO UNTIL Statement – DO WHILE Statement Defining and Processing Arrays –Arrays -Applying SAS Formats and Informats – Specifying SAS Formats

Tutorial:

- One-to-One Reading to Combine Data Sets-Concatenating to Combine Data Sets-Match-Merging to Combine Data Sets-Renaming Variables-Processing Iterative DO Loops-Indenting and Nesting DO Groups

Unit 4 SAS Programming and Automation	**9 hours**
Definitions – Reading Date and Time with Informats – Displaying Date and Time Values with Formats-Basics of SAS Functions - SAS Functions Syntax – Converting Data with Functions- The MEANS Procedure – Specifying Descriptive Statistics – Limiting Decimal Places – Specifying Variable Using VAR Statement – Summarized Data Set Using OUTPUT Statement – The FREQ Procedure-Specifying Variable Using TABLES Statements - Two-Way and N-Way Tables – LIST Options-The Output Delivery System (ODS) – HTML Outputs – PDF Outputs-RTF Outputs – EXCEL Outputs – The EXPORT Procedure	
Tutorial: • Reading Date and Time with Formats and Informats- Execution of Data Values Using Numeric Functions-Creating a Descriptive Statistics by Using PROC MEANS-Creating a Descriptive Statistics by Using PROC FREQ-Creating Outputs with HTML, PDF, RTF, EXCEL Formats-Exporting a External File	
Unit 5 Advanced SAS Topics	**9 hours**
Creating a Default List Report – Windowing/Non Windowing Mode - Define Statement-Defining Variables – Column Statement – Selecting Observations – Defining the Variables-Group Processing the Variables by Using Different Options – Computing the Variable-Bar Chart – Box Plot – Mean Measurement Over the Time using Line Plot – Spaghetti Plot – Survival Plot – Waterfall Plot – Forest Plot-Introduction to Macro Facility – SAS Programs and Macro Processing – Macro Variables – Macro Processing – Scopes of Macro Variables – Macro Expressions-Scopes of Macro Variables – Macro Expressions-Macro Quoting – Interfaces with Macro Facility -Storing and Reusing Macros.	
Tutorial: • Creating Graphs by Using Bar Chart -Creating Graphs Using Various Types of Plots-Generating SAS Code Using Macros-Defining User-Defined Macro Variables-Creating a Tables Using PROC SQL with different Clauses- Generating a Cartesian Product -Combining Tables Using Multiple Joins	

Learning Resources	1. "The Little SAS Book: A Primer", Lora D. Delwiche and Susan J. Slaughter, SAS Institute 2. "Statistics and Data Analysis for Financial Engineering", David Ruppert, David S. Matteson, and James E. Gentle, Springer, Third Edition. 3. "Applied Multivariate Statistical Analysis", Richard A. Johnson and Dean W. Wichern, Pearson, Seventh Edition	4. "SAS Certification Prep Guide: Advanced Programming for SAS 9", SAS Institute, Fourth Edition 5. "SAS for Data Science", Brett Wujek, Robert A. Muenchen, and Arthur L. Carpenter, SAS Institute, First Edition

Learning Assessment

	Bloom's Level of Thinking	*Continuous Learning Assessment (CLA)*				*Summative Final Examination (40% weightage)*	
		Formative CLA-1 Average of unit test (50%)		*Life Long Learning CLA-2 – (10%)*			
		Theory	*Practice*	*Theory*	*Practice*	*Theory*	*Practice*
Level 1	*Remember*	*15%*	-	*15%*	-	*15%*	-
Level 2	*Understand*	*25%*	-	*20%*	-	*25%*	-
Level 3	*Apply*	*30%*	-	*25%*	-	*30%*	-
Level 4	*Analyse*	*30%*	-	*25%*	-	*30%*	-
Level 5	*Evaluate*	-	-	*10%*	-	-	-
Level 6	*Create*	-	-	*5%*	-	-	-
	Total	*100 %*		*100 %*		*100 %*	

Course Designers

Experts from Industry	**Experts from Higher Technical Institutions**	**Internal Experts**
Mr. Nishanth Nalan, Technical Head, ACL Digital India		Dr. Priyadarsini .K, AP/DSBS,SRMIST
		Dr. Jeba Sonia J, AP, DSBS, SRMIST