

Course Code	PCS21C06J	Course Name	BIG DATA ANALYTICS	Course Category	C	Professional Core Course	L	T	P	C
							3	0	4	5

Pre-requisite Courses	Nil	Co-requisite Courses	Nil	Progressive Courses	Nil
Course Offering Department	Computer Science		Data Book / Codes/Standards	Nil	

Course Learning Rationale (CLR):		Learning			Program Learning Outcomes (PLO)																
The purpose of learning this course is to:		1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
CLR-1 :	Understand the evolution of computer networks using the layered network architecture	Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Fundamental Knowledge	Application of Concepts	Link with Related Disciplines	Procedural Knowledge	Skills in Specialization	Ability to Utilize Knowledge	Skills in Modeling	Analyze, Interpret Data	Investigative Skills	Problem Solving Skills	Communication Skills	Analytical Skills	PSO 1	PSO 2	PSO 3		
CLR-2 :	Understand the addressing concepts and learn networks devices				L	H	-	H	L	-	-	-	-	-	-	-	-	-	-	-	-
CLR-3 :	Design computer networks using subnetting and routing concepts				M	H	M	H	L	-	-	-	-	-	-	-	-	-	-	-	-
CLR-4 :	Understand the error types, framing, flow control				M	H	M	H	L	-	-	-	-	-	-	-	-	-	-	-	-
CLR-5 :	Understand the various Medium Access Control techniques and also the characteristics of physical layer functionalities				H	H	M	H	L	-	-	-	-	-	-	-	-	-	-	-	-
CLR-6 :	Know the algorithms behind the protocols that helps data transfer				L	H	-	H	L	-	-	-	-	-	-	-	-	-	-	-	-
Course Learning Outcomes (CLO):		At the end of this course, learners will be able to:			L	H	-	H	L	-	-	-	-	-	-	-	-	-	-		
CLO-1 :	Acquire the basics of computer network and its architecture	3	80	70	L	H	-	H	L	-	-	-	-	-	-	-	-	-	-		
CLO-2 :	Acquire the knowledge of various networks devices and addressing methods	3	85	75	M	H	L	M	L	-	-	-	-	-	-	-	-	-	-		
CLO-3 :	Design the network routing methods	3	75	70	M	H	M	H	L	-	-	-	-	-	-	-	-	-	-		
CLO-4 :	Find the error type that may happen during data transportation	3	85	80	M	H	M	H	L	-	-	-	-	-	-	-	-	-	-		
CLO-5 :	Understand the physical layer functions and components	3	85	75	H	H	M	H	L	-	-	-	-	-	-	-	-	-	-		
CLO-6 :	Speak on the topology chosen for a architecting a network that an organization demands	3	80	70	L	H	-	H	L	-	-	-	-	-	-	-	-	-	-		

Duration (hour)	21	21	21	21	21
S-1	SLO-1 SLO-2	Introduction to BigData platform	Null and Alternative Hypotheses. Type-I and Type-II ErrorCriticalRegion and Level of Significance	History of Hadoop	Setting up a Hadoop Cluster
S-2	SLO-1 SLO-2	Challenges of Conventional Systems	One tailed and two tailed tests	The Hadoop Distributed File System	Cluster specification -
S-3	SLO-1 SLO-2	Intelligent data Analysis	Critical values of significant values	Components of Hadoop	Cluster Setup and Installation
S4-7	SLO-1 SLO-2	Laboratory 1 : Practice elementary mathematical operations and control statements	Laboratory 4 : Creating Various types of plots /charts from various data source	Laboratory 7 : Implementation of Linear regression with multiple regression	Laboratory 10: Implementation of classifier problem
S-8	SLO-1 SLO-2	Nature of data	Tests of Significance for Large Samples	Analyzing the Data with Hadoop Scaling Out- Hadoop Streaming	Hadoop Configuration
S-9	SLO-1 SLO-2	Analytic process and tools	- Test of Significance for Single Proportion Test of Significance for Difference of Proportions	Java interfaces to HDFS Java interfaces to HDFS	Security in Hadoop
S-10	SLO-1	Analysis Vs Reporting	Test of Significance for Single Mean	How Map Reduce Works	Administering Hadoop



Duration (hour)		21	21	21	21	21
	SLO-2	Modern Data analytic tools	Test of Significance for Difference of Means. Chi-Square Distribution	Anatomy of a Map Reduce Job run		Database
S11-14	SLO-1	Laboratory 2 : Operations on Matrices and Vectors	Laboratory 5 : Create subplots and color plots	Laboratory 8 : Implementation of Data preprocessing methods , Correlation matrix	Laboratory 11 : Implementation of K-Mean Clustering	Example 14 : Implementation of Random Forest
	SLO-2					
S-15	SLO-1	Parameter and Statistic	To test the goodness of fit. - To test the independence of Attributes. Student's "t" - Distribution	Failures	Administering Hadoop -HDFS	Structure of Zoo Keeper
	SLO-2					
S-16	SLO-1	Sampling Distribution-	Definition- Applications of Student's "t" – Distribution- To test for Single Mean- To test for Difference of Means	Task Execution –Map read and Map write anatomy	Monitoring	The Zoo keeper services
	SLO-2					
S-17	SLO-1	Meaning-Standard Error and its uses. Tests of Significance	F-Distribution- Definition- To Test for Equality of Two Population variances. Meaning of Resampling and its uses	Map reduce features	Maintenance	Case study
	SLO-2					
S18-21	SLO-1	Laboratory 3 : Vectorized operation on simple matrix operations	Laboratory 6 : Implement Linear regression problem	Laboratory 9 : Implementation of spam and non-spam classification problem.	Laboratory 12 : Implementation of K-Mean Clustering	Laboratory 15 : Implementation of CART
	SLO-2					

Learning Resources	1. Michael Berthold, David J. Hand, (2007), "Intelligent Data Analysis", Springer.	4. AnandRajaraman and Jeffrey David Ullman, (2012) "Mining of Massive Datasets", Cambridge University Press.
	2. RSN Pillai, Bagavathi, "Statistics Theory and Practice", S.Chand 3. Tom White (2012), " Hadoop: The Definitive Guide" Third Edition, O'reilly Media	5. Viktor Mayer,Schonberger,KennethCukier , "Big Data : A Revolution That Will Transform How We Live, Work and Think".

Learning Assessment											
Bloom's Level of Thinking		Continuous Learning Assessment (50% weightage)								Final Examination (50% weightage)	
		CLA – 1 (10%)		CLA – 2 (10%)		CLA – 3 (20%)		CLA – 4 (10%)#			
		Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice
Level 1	Remember	20%	20%	15%	15%	15%	15%	15%	15%	15%	15%
	Understand										
Level 2	Apply	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
	Analyze										
Level 3	Evaluate	10%	10%	15%	15%	15%	15%	15%	15%	15%	15%
	Create										
	Total	100 %		100 %		100 %		100 %		100%	

# CLA – 4 can be from any combination of these: Assignments, Seminars, Short Talks, Mini-Projects, Case-Studies, Self-Study, MOOCs, Certifications, Conf. Paper etc.,

Course Designers		
Experts from Industry	Experts from Higher Technical Institutions	Internal Experts
Mr. S. Karthik, Assistant Consultant, Tata Consultancy Services	Dr.S.Sasikala, Associate Professor and Head, Dept. of Computer Science, University of Madras	Mrs. P. Yogalakshmi
		Dr. P.Muthulakshmi