Mr.G.Muruganandam, Group Project Manager, HCL Technologies, Chennai	Dr.Muthu, Professor, Loyola College, Chennai	*Mrs. Vasavi J , CSH, SRM IST*
Mr.M. Hemachandar, Tech Lead, Wipro Limited, Chennai	Dr.Vincent, Associate Professor, VIT	

Course Code	***PAD21D02T***	**Course Name**	***DATA ARCHITECTURE AND BIG DATA***	**Course Category**	**D**	**Discipline Specific Elective**	**L**	**T**	**P**	**C**
							4	**0**	**0**	**4**

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

Course Learning Rationale (CLR):	The purpose of learning this course is to,
CLR-1 :	To learn how to Data architects help companies manage, store and secure their data.
CLR-2 :	To learn about SQL and NoSQL databases
CLR-3 :	To get a clear understanding about Hadoop
CLR-4 :	To get in-depth knowledge of the Big Data framework using Hadoop
CLR-5 :	To analyze large data sets to find trends, correlations or other insights not visible with smaller data sets or traditional processing methods.
CLR-6 :	To observe various customer related patterns and trends.

Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:	Learning (Bloom)	Expected Proficiency (%)	Expected Attainment (%)
CLO-1 :	To gather extensive knowledge in Big Data and its architecture	2	85	80
CLO-2 :	To improve and understand ETL process and relevant ETL tools	3	85	80
CLO-3 :	To differentiate between a batch layer for large volumes of data and a speed layer for real time processing of data streams	3	85	80
CLO-4 :	To understand Distributed systems	3	85	80
CLO-5 :	To understand Data modelling techniques	3	85	80
CLO-6 :	To incorporate data from all sources is key to optimizing the insights gained with Big Data.	3	85	80

Program Learning Outcomes (PLO)														
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Disciplinary Knowledged	Critical Thinking	Problem Solving	Analytical Reasoning	Research Skills	Team Work	Scientific Reasoning	Reflective Thinking	Self-Directed Learning	Multicultural Competence	Ethical Reasoning	Community Engagement	ICT Skills	Leadership Skills	Life Long Learning
L	*H*	*H*	*H*	*H*	*M*	-	*H*	*M*	*H*	-	*H*	-	-	-
L	*H*	*H*	*H*	*H*	-	-	*M*	*M*	*L*	-	*H*	-	-	-
L	*H*	*H*	*H*	*H*	-	-	*M*	*M*	*L*	-	*H*	-	-	-
L	*H*	*H*	*H*	*H*	-	-	*M*	*M*	*L*	-	*H*	-	-	-
L	*H*	*H*	*H*	*H*	-	-	*M*	*M*	*L*	-	*H*	-	-	-
L	*H*	*H*	*H*	*H*	-	-	*M*	*M*	*L*	-	*H*	-	-	-

Duration (hour)		**12**	**12**	**12**	**12**	**12**
S-1	SLO-1	Introduction to Data and Data lifecycle principles	Introduction to Big data and Hadoop	Four V's in Big Data	Distributed computing and ETL	Data lake and data warehouse

S-2	SLO-1	Data and data lifecycle	Introduction to Bigdata	Big Data - reading data from CSV	Data Driven Organizations	Data lakes
S-3	SLO-1	Databases	Characteristics of Big Data	Overview of the Four Vs	Decision Making	Characteristics of data lakes
S-4	SLO-1	Database types	Getting started with Hadoop	The Importance of Volume	Distributed systems	Data lake Architecture
S-5	SLO-1	SQL	Building systems to scale with data	The Importance of Variety - The Importance of Velocity	Batch vs in-memory processing	Data warehouse
S-6	SLO-1	NoSQL	A quick overview of Hadoop	The Relationship	Tools for data management	Characteristics of Data Warehouse
S-7	SLO-1	Creating ERD (Entity relationship diagram)	MapReduce overview	Between the Four Vs	Understanding ETL	Data warehouse Architecture
S-8	SLO-1	Implementing SQL with AWS	Map	Variety and Data Structure	ETL with Talend open studio	Data lakes Vs Data Streams
S-9	SLO-1	Implementing NoSQL with AWS	Phases of MapReduce	Validity and Volatility	ETL pipeline in python	Data streams
S-10	SLO-1	Create NoSQL	Shuffle phase	Finding Balance in the – Four Vs	AI and machine learning	Migrate data to AWS
S-11	SLO-1	NoSQL DB with python	Reduce phase	**Use Cases**	Data modelling	Data lakes on AWS
S-12	SLO-1	Create SQL DB with python	Data Silos	Extracting Value from the Four V's	Data partitioning/engineering/ reporting	Working with data lakes on AWS

Learning Resources	1.Big Data Fundamentals: Concepts, Drivers & Techniques, By Wajid Khattak, Paul Buhler and Thomas Erl, January 2016 2.The Enterprise Big Data Lake, Alex Gorelik, March 2019	1.Practical Enterprise Data Lake Insights: Handle Data-Driven Challenges in an Enterprise Big Data Lake, Saurabh Gupta and Venkata Giri, June 2018

Learning Assessment

Level	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 Understand	40%	-	40%	-	40%	-	40%	-	40%	-
Level 2	Apply Analyze	40%	-	40%	-	40%	-	40%	-	40%	-
Level 3	Evaluate Create	20%	-	20%	-	20%	-	20%	-	20%	-
	Total	100 %		100 %		100 %		100 %		100 %	