cct | College Dublin
Computing • IT • Business

Teaching Scheme

Programme Name & Stage:	MSc in Data Analytics
Module Title:	Big Data Storage and Processing
Semester:	2 (Aug – Nov 2024)
Lecturer:	Dr. Muhammad Iqbal

Assessment Weighting:

Integrated CA1 Individual	40%
Integrated CA2 Individual	60%

NOTE: This Teaching Scheme is intended as a GUIDE ONLY. It is possible that the topics/areas covered may be changed from time-to-time.

Week No.	Date Commencing	Major Topic(s) / Subject Area(s)	References
1	26th Aug 2024	Legacy Approaches for Relational and Non-relational databases and Big data storage, Practical tutorial	NoSQL Database for Storage and Retrieval of Data in Cloud, Ganesh Chandra Deka, CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, 2017.
2	2nd Sept 2024	Distributed Systems and Data management Hadoop framework practical, Practical tutorial	Tom White 2012, Hadoop: The Definitive Guide, O'Reilly Media/Yahoo Press [ISBN: 1449311520]
3	9th Sept 2024	Apache Hadoop and HDFS, MapReduce and YARN (resource management) MapReduce practical programming, Practical tutorial	
4	16th Sept 2024	Apache Spark Streaming and practical implementation, Practical tutorial	Learning Spark, Jules Damji, Denny Lee, Brooke Wenig, Tathagata Das, 2nd Edition, O'Reilly Media, Inc., July 2020.
5	23rd Sept 2024	Spark MLib and further expertise, Practical tutorial	
6	30th Sept 2024	NoSQL database HBase, Practical tutorial	
7	7th Oct 2024	Apache Pig (High level scripts on Apache Hadoop) and practical implementation, Practical tutorial	
8	14th Oct 2024	Apache Hive (SQL-like queries) and practical implementation, Practical tutorial	Apache Hive Essentials, Dayong Du, Packt Publishing, June 2018.

9	21st Oct 2024	READING WEEK (No classes from Tuesday 22nd Oct – Friday 25th Oct)	
10	28th Oct 2024	Apache Mahout (ML algorithms on hadoop), Practical tutorial	
11	4th Nov 2024	Columnar Databases (Cassandra) and practical implementation, Practical tutorial	
12	11th Nov 2024	Kafka understanding and practical implementation, Practical tutorial	
13	18th Nov 2024	Apache Graph and practical implementation, Practical tutorial	
14	25th Nov 2024	**Exams**	

Examination Period: Exams will be held from Monday 25th Nov 2024 to Friday 29th Nov 2024 (inclusive). An exact Examination Timetable will be issued closer to the time.

Although there are no classes scheduled during Reading Week, please note that this is not a Holiday period and it is possible that additional classes may be scheduled during a Reading Week, if necessary.

*** During Revision Periods your lecturer may schedule an additional class, if this is necessary.*