

**Maulana Abul Kalam Azad University of Technology, West Bengal**  
*(Formerly West Bengal University of Technology)*  
**Syllabus for M. Sc. In Data Science**  
**(Effective for Academic Session 2019-2020)**

**Detailed Syllabus**  
**Semester I**

**Paper: STATISTICAL METHODS**  
**Code: MDS 101**  
**Contacts Hours / Week: 3L+1T**  
**Credits: 4**

**Objectives:**

To enable the students to:  
 provide an understanding for the student on statistical concepts to include measurements of location and dispersion, probability, probability distributions, sampling, estimation, hypothesis testing, regression, and correlation analysis, multiple regression and business/economic forecasting

UNITS	COURSE CONTENT
1	<b>Descriptive statistics</b>
2	<b>Classical and axiomatic definitions of Probability and consequences. Law of total probability, Conditional probability, Bayes' theorem and applications. Discrete and continuous random variable</b>
3	<b>Distribution functions and their properties. Standard discrete and continuous probability distributions - Bernoulli, Uniform, Binomial, Poisson, Geometric, Rectangular, Exponential, Normal. Random vectors, Joint and marginal distributions, Conditional distributions, Distributions of functions of random variables</b>
4	<b>Mathematical expectation and conditional expectation. Central limit theorem</b>
5	<b>Hypothesis testing, Estimation and sampling techniques</b>
6	<b>Sampling distributions of sample mean, sample variance, t, chi-square and F tests of significance based on them, Small sample tests</b>

**References:**

1. Outline of statistics 1 and 2 by Goon-Gupta-Dasgupta
2. Fundamentals of mathematical statistics by Gupta and Kapoor
3. Complete Business Statistics Book By Amir Aczel

**Maulana Abul Kalam Azad University of Technology, West Bengal**  
*(Formerly West Bengal University of Technology)*  
**Syllabus for M. Sc. In Data Science**  
**(Effective for Academic Session 2019-2020)**

**Paper: Data Base Management System**

**Code: MDS 102**

**Contacts Hours / Week: 3L+1T**

**Credits: 4**

**Objectives:**

To enable the Students to:  
 Understand the basics of Database Management System

UNITS	COURSE CONTENT
1	Overview of Database Management, Conceptual Database Design, Logical Database Design, Physical Database Design.
2	Introduction to Relational Database : Relation, Optimization, The Catalog, Base Relvars and Views, Transactions, The Suppliers and Parts Database
3	Relational Model Concepts, Relational Model, Constraining, Referential Integrity Constraints, Defining Referential Integrity Constraints, Update Operations on Relations, Structured Query Language (SQL), Data Definition Language Commands, Data Manipulation Language Commands, Transaction Control Commands, SQL Command Syntax and Usage, The Basic Query Block, Querying Data with Multiple Conditions, Basic Relational Algebra Operations, The Select Operation, Additional Relational Operations
4	ER- and EER-to-Relational Mapping: ER- to Relational Mapping Algorithm, Summary of Mapping for Model Constructs and Constraints Mapping EER Model Concepts to Relations, Query, Processing and Optimization: Query Processing, Query Optimization, Database Tuning
5	Object Oriented Database Systems: Characteristics of an Object-relation Database Management System (ORDBMS), Complex Objects, Inheritance, Function Overloading, Rules. Distributed Database : Distributed Database System, Distributed Database Design, Data Fragmentation, Data Replication, Data Allocation, Query Processing in Distributed Databases.
6	Recovery: Transactions, Transaction Recovery, System Recovery, Media Recovery, Two-phase Commit. Database Models, Introduction to HDFS, Graph based Modeling

**References:**

1. Database System Concepts – 6th Edition by Silberschatz, Korth and Sudarshan
2. Fundamentals of Database Systems – 5th Edition by R.Elmasri, S. Navathe
3. Database Design and Relational Theory: Normal Forms and All That Jazz by C.J. Date

**Maulana Abul Kalam Azad University of Technology, West Bengal**  
(Formerly West Bengal University of Technology)  
**Syllabus for M. Sc. In Data Science**  
(Effective for Academic Session 2019-2020)

**Paper: Programming Using Python**

**Code: MDS 103**

**Contacts Hours / Week: 3L+1T**

**Credits: 4**

**Objectives:**

To enable the students to:

- Understand basics of binary computation
- Understand the programming basics (operations, control structures, data types, etc.)
- Readily use the Python programming language
- Apply various data types and control structure
- Understand class inheritance and polymorphism
- Understand the object-oriented program design and development
- Understand and begin to implement code

UNITS	COURSE CONTENT
1	<b>Python Basics</b> Introduction, your first program, type, expressions and variables, string operations, packages
2	<b>Python Data Structure</b> Lists and Truples, Sets, Directories
3	<b>Programming Fundamentals</b> Conditions and Branching, Loops, Functions, Object and Classes, Strings
4	<b>Python Database</b> Working with database
5	<b>Working with Data</b> Open/Reading/Writing files, Loading data with Pandas, Saving data with Pandas

**References:**

1. Introduction to Programming in Python: An Interdisciplinary Approach, By Sedgewick, Wayne and Dondero
2. An Introduction to Python, By Guido Van Rossum.
3. Programming in Python 3: A Complete Introduction to Python Language, By Mark Summerfield
4. Core Python Programming, By Dr R Nageshwar Rao

**Maulana Abul Kalam Azad University of Technology, West Bengal**  
*(Formerly West Bengal University of Technology)*  
**Syllabus for M. Sc. In Data Science**  
**(Effective for Academic Session 2019-2020)**

**Paper: Fundamentals of Analytics**  
**Code: MDS 104**  
**Contacts Hours / Week: 3L+1T**  
**Credits: 4**

**Objectives:**

To enable the Students to:  
Understand the fundamentals of Analytics and how it helps in the business process

UNITS	COURSE CONTENT
1	Introduction to Data Structure and Algorithm: Use of Big O and Small o notations, Use of Big Omega and small omega notations. Efficiency of algorithms. Analysis of recursive programs. Solving recurrence equation, Implementation of Abstract Data Types(ADT), list, stack, queue, hashing. Tree Structure binary trees, AVL trees, B and B+ trees, Red-Black Trees, heap, Tree-Traversal Algorithms, Graphs and algorithms, Prim's and Kruskal's algorithms, Dijkstra's method, backtracking, minimum spanning trees, Sorting and searching algorithms
2	Introduction to Data modeling: Conceptual, Logical and physical modeling, Top down and Bottom Up modeling, ER Modeling with different methodologies, Generic data modeling, Semantic data modeling, Data modeling with different techniques (finite state machine, UML, state charts, Markov Chain, Decision table and decision tree), Data Dictionary

**References:**

1. Data modeling essentials – Graeme Simsion
2. The Data model resource book- Len Silverston

**Maulana Abul Kalam Azad University of Technology, West Bengal**  
*(Formerly West Bengal University of Technology)*  
**Syllabus for M. Sc. In Data Science**  
**(Effective for Academic Session 2019-2020)**

**Paper: Analytics Lab 1**

**Code: MDS 191**

**Contacts Hours / Week: 1L+1T+2P**

**Credits: 4**

Installation, Basic Data type, functions for reading and writing data, control structures, functions, loops, debugging tools, simulation and profiling, str function, R environment, Descriptive statistics and graphics, probability and distribution.

**Paper: DBMS Lab**

**Code: MDS 192**

**Contacts Hours / Week: 1L+1T+2P**

**Credits: 4**

Introduction to SQL constructs. Review of Basic SQL statements Select, Project, Join, Describing  
Oracle tables, Restricting row returns Creating basic reports, Using the set commands, Adding prompts to queries  
Joining Oracle tables -Equi-join, Outer join Hiding joins by creating views,UsingIN, NOT IN, EXISTS and NOTEXISTS, Subqueries, Exercise – write a subquery,Correlated subquery, Non-correlated subqueries  
Advanced SQL operators -Between operator ,IN and NOT IN operators, Sub-queries- EXISTS  
clause, Using wildcards in queries (LIKE operator),Aggregation in SQL -Count(\*),Sum, Avg, Min and max. Using the group by clause, SQL access methods ,Review of Basic joining methods-Merge join, Hash Join, Nested Loop join.

**Paper: Python Lab**

**Code: MDS 193**

**Contacts Hours / Week: 2L+2P**

**Credits: 4**

Introduction, your first program, type, expressions and variables, string operations, packages, Lists and Tuples, Sets, Directories, Conditions and Branching, Loops, Functions, Object and Classes, Strings, Open/Reading/Writing files, Loading data with Pandas, Saving data with Pandas