MASTER OF SCIENCE (INFORMATION SECURITY) (MSCIS)

ASSIGNMENTS
JULY 2025

Semester-III

MSE-029, MSE-030, MSE-031, MCS-226 and MSEL-032

SCHOOL OF VOCATIONAL EDUCATION AND TRAINING
INDIRA GANDHI NATIONAL OPEN UNIVERSITY
MAIDAN GARHI, NEW DELHI – 110 068

CONTENTS

Guidelines regarding submission of assignments

1. It is compulsory for the students to submit the prescribed assignments. They will not be allowed to appear for the term-end examination of a course if they do not submit the assignments in time for that course.

2. The assignment responses should be complete in all respects. Before submission, the students should ensure that they have answered all the questions in all assignments. Incomplete answer sheets bring poor grades.

3. The University/ Regional Centre have the right to reject the assignments received after the due date. Therefore, the students are advised to submit their assignments before the due date.

4. Students should submit before the last dates prescribed for submission of assignments.

5. In case the students have already done some assignments prescribed in a course, they are required to do the **left-over assignments before taking the Term-end Examination**. If they have qualified in a course on the basis of lesser number of assignments and Term-end Examination, they will **not be eligible to re-do the assignments** with a view to improve the overall qualifying score of that course.

6. In case any student fails to submit the assignments or fails to score minimum qualifying marks, s/he has to wait for fresh assignments meant for the current batch of students.

7. For their own record, students should retain a copy of all the assignment responses, which they submit.

8. Once the students get the pass marks in an assignment, they can not re-submit it for improvement of marks. If the student secured requisite marks in Term-End Examination and Assignments, but did not get requisite overall percentage, then the student can either resubmit the assignment or reappear for the term-end examination for securing the requisite overall percentage.

9. Assignments are not subject to re-evaluation.

Instructions for Doing Assignments

While answering Assignments, the following guidelines are required to be observed:

1. The student should write their Complete correct Enrolment Number, Name, Full Address, Signature and Date on the top right hand corner of the first page of the response sheet.
2. The students should write the Programme Title, Course Code and Course Title on the left hand corner of the first page of their response sheet. Course code may be reproduced from the assignment.

The top of the first page of your response sheet for each assignment should look like this:

PROGRAMME TITLE: ....   ENROLMENT No.: ……………….

COURSE CODE: ....   NAME : ……………………………

COURSE TITLE: ....

ADDRESS:……………………………………………………………………………………..

SIGNATURE: ……………………

DATE: …………………………..

---

3.   The students should write the responses in their own hand. **They should not print or type the answers. They should not reproduce their answers from the units sent to them by the University. If they reproduce from units, they will get poor marks for the respective question**.

4.   The students should write each assignment separately. All the assignments should not be written in continuity.

5.   The students should write the question number with each answer.

6.   The students should use only foolscap size paper for their response and tie all the pages carefully. Avoid using very thin paper. They should allow a 4 cm. margin on the left and at least 4 lines in between each answer. This may facilitate the evaluator to write useful comments on the margins at appropriate places.

7.   The students should not copy from the response sheets of other students. If copying is noticed, the assignments of such students will be rejected, and disciplinary action can be taken against the students as per rules of the University.

8.   Please remember that it is compulsory to send scanned copies of handwritten assignments through email to the respective Regional Centre email or upload on the link provided on the respective Regional centre's website or the hard copy of handwritten assignments to your concerned Regional centre/Study centre before you can take the Term End Exams or else you will not be issued hall tickets.

   Under no circumstances should they be sent to the (SED) for evaluation.

Course Code	:	MCS-226
Course Title	:	Data Science and Big Data
Maximum Marks	:	100
Weightage	:	30%
Last date of Submission	:	31st October, 2025 (For latest update, Pl. check IGNOU's Website)

This assignment has ten questions of 8 Marks each, answer all questions. Rest 20 marks are for viva voce. You may use illustrations and diagrams to enhance the explanations. Please go through the guidelines regarding assignments given in the Programme Guide for the format of presentation.

Q1: Define the term data science. Describe its applications in two industries of your choice (e.g., healthcare, finance, e-commerce). What role does the data science lifecycle play in managing data projects?

Q2: Explain Exploratory Data Analysis (EDA) and its importance. What are the main steps in performing EDA on a new dataset? Describe two methods for detecting outliers and how handling outliers impacts data analysis.

Q3: Describe the role of statistical hypothesis testing in data analysis. What are Type I and Type II errors, and how do they affect decision-making? Provide an example of hypothesis testing in a real-world scenario.

Q4: Discuss the 4 Vs of big data (Volume, Velocity, Variety, and Veracity). Provide a real-world example of each, explaining how these characteristics create challenges in big data management.

Q5: Explain the Hadoop architecture with a focus on HDFS and the master/slave architecture. How do NameNode and DataNodes work together to store and manage large datasets? Provide a basic example of this storage process.

Q6: Compare Apache Spark, Hive, and HBase in terms of functionality, data processing methods, and use cases. When would Spark be preferred over traditional MapReduce, and why?

Q7: Describe the purpose and functionality of a *Bloom filter* in data stream processing. How does the Bloom filter efficiently check for element presence? Describe the Flajolet-Martin algorithm for cardinality estimation in data streams.

Q8: What is the PageRank algorithm, and how is it used in link analysis? Describe the concept of "flow of rank" in PageRank. Explain how the PageRank of a web page is calculated using the flow model.

Q9: Discuss the challenges of online advertising and recommendation systems. Explain the concept of collaborative filtering with an example, and discuss the role of clustering

in social network analysis.

Q10: What is the Random Forest algorithm? Explain how it can be applied to classification problems. Write a program in R to implement a Random Forest classifier on a sample dataset and explain its output.