

**MASTER OF SCIENCE  
DATA SCIENCE AND ANALYTICS  
(MSCDSA)**

**MSCDSA/ASSIGN/SEMESTER-I**

**ASSIGNMENTS**

**(January – 2026 & July – 2026)**

**MCS-061, MCS-062, MCS-063, MCS-207, MCSL-064,  
MCSL-065**



**SCHOOL OF COMPUTER AND INFORMATION SCIENCES  
INDIRA GANDHI NATIONAL OPEN UNIVERSITY  
MAIDAN GARHI, NEW DELHI – 110 068**

## CONTENTS

Course Code	Assignment No.	Submission-Schedule		Page No.
		For January-June Session	For July-December Session	
MCS-061	MSCDSA(I)/061/Assign/26	30 <sup>th</sup> April, 2026	31 <sup>st</sup> October, 2026	3
MCS-062	MSCDSA(I)/062/Assign/26	30 <sup>th</sup> April, 2026	31 <sup>st</sup> October, 2026	5
MCS-063	MSCDSA(I)/063/Assign/26	30 <sup>th</sup> April, 2026	31 <sup>st</sup> October, 2026	8
MCS-207	MSCDSA(I)/207/Assign/26	30 <sup>th</sup> April, 2026	31 <sup>st</sup> October, 2026	9
MCSL-064	MSCDSA(I)/L-064/Assign/26	30 <sup>th</sup> April, 2026	31 <sup>st</sup> October, 2026	10
MCSL-065	MSCDSA(I)/L-065/Assign/26	30 <sup>th</sup> April, 2026	31 <sup>st</sup> October, 2026	12

### Important Notes

1. Submit your assignments to the Coordinator of your Study Centre on or before the due date.
2. Assignment submission before due dates is compulsory to become eligible for appearing in corresponding Term End Examinations. For further details, please refer to MSCDSA Programme Guide.
3. To become eligible for appearing the Term End Practical Examination for the lab courses, it is essential to fulfill the minimum attendance requirements as well as submission of assignments (on or before the due date). For further details, please refer to the MSCDSA Programme Guide.
4. The viva voce is compulsory for the assignments. For any course, if a student submitted the assignment and not attended the viva-voce, then the assignment is treated as not successfully completed and would be marked as ZERO.

**Course Code** : **MCSL-065**  
**Course Title** : **Data Science Lab**  
**Assignment Number** : **MSCDSA(I)/L-065/Lab\_Assign/26**  
**Maximum Marks** : **100**  
**Weightage** : **30%**  
**Last Date of Submission** : **30<sup>th</sup> April, 2026 (for January session)**  
**31<sup>st</sup> October, 2026 (for July session)**

There are two Sections in this assignment carrying a total of 40 marks. Your Lab Record will carry 40 Marks. 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. Submit the screenshots along with the coding and documentation.

**Dataset :** To attempt the problems given in this assignment, refer to the Dataset: Students' Exam Scores given below:

Student_ID	Student_Name	Physics	Chemistry	Maths	English
S01	Rahul	78	72	85	80
S02	Ananya	88	91	90	86
S03	Amit	65	69	70	72
S04	Neha	92	89	94	90
S05	Karan	55	60	58	62
S06	Priya	81	85	88	84
S07	Rohit	70	68	75	73
S08	Simran	90	92	89	91
S09	Arjun	60	65	63	67
S10	Pooja	85	88	90	87

**SECTION-1 DATA SCIENCE LAB –GUI TOOLS (20 Marks)**

**EXCEL BASED PROBLEM:**

**1. Use Excel to perform following tasks: (6 Marks)**

- a. Create a Histogram for Maths marks
- b. Plot a Line Chart for Maths scores against Student\_ID.
- c. Calculate average marks for each subject, and Create a Bar Chart comparing subject-wise average scores.
- d. Using average marks, create a Pie Chart showing percentage contribution of each subject.
- e. Use data analysis toolpak to perform following:
  - i. Generate descriptive statistics
  - ii. Perform a two-sample t-test to check whether there is a significant difference between the marks of Physics and Chemistry.
  - iii. Perform an F-test to compare the variance of marks in Maths&English
  - iv. Calculate Pearson Correlation Coefficient between
    - Maths and Physics
    - Maths and Chemistry

**TABLEAU BASED PROBLEM:**

**(7 Marks)**

**2. Use Tableau to perform following tasks:**

- a. Save the above data given in Dataset: Students' Exam Scores in an excel file or CSV file, save it, and perform following:
  - Import the Students' Exam Scores dataset into Tableau.
  - Verify data types for each subject.
  - Rename fields appropriately (if required).
- b. Create a Bar Chart showing average marks for:
  - Physics
  - Chemistry
  - Maths
  - English

Label axes and add chart title.
- c. Create a Line Chart showing Maths scores across students. (Use Student\_ID on X-axis.)
- d. Create a Pie Chart showing percentage contribution of each subject based on average marks. (Display percentage labels.)
- e. Create a Table View displaying: Average, Minimum, Maximum, and Standard Deviation. (For all four subjects).
- f. Now, create a single dashboard that includes:
  - i. Bar Chart (Subject-wise Average)
  - ii. Line Chart (Maths Trend)
  - iii. Pie Chart (Subject Contribution)
  - iv. Summary Statistics Table

**POWER BI BASED PROBLEM:**

**(7 Marks)**

**3. Use Power BI to perform following tasks:**

- a. Save the above data given in Dataset: Students' Exam Scores in an excel file or CSV file, save it, and perform following:
  - Import the Students' Exam Scores dataset into Power BI.
  - Verify data types for each subject.
  - Rename fields appropriately (if required).
- b. Create a Bar Chart showing average marks for:
  - Physics
  - Chemistry
  - Maths
  - English

Identify the subject with the highest and lowest average score.
- c. Create a Line Chart showing Maths scores across students. (Use Student\_ID on X-axis.)
- d. Create a Pie Chart showing percentage contribution of each subject based on average marks. (Display percentage labels.)

- e. Create the following **DAX measures**: Average, Minimum, Maximum, and Standard Deviation. (For all four subjects).
- f. Create a Table View displaying: Average, Minimum, Maximum, and Standard Deviation. (For all four subjects).
- g. Now, create a single dashboard that includes:
  - i. Bar Chart (Subject-wise Average)
  - ii. Line Chart (Maths Trend)
  - iii. Pie Chart (Subject Contribution)
  - iv. Summary Statistics Table

## **SECTION-2 DATA SCIENCE LAB – PROGRAMMING BASED**

**(20 Marks)**

### **PYTHON PROGRAMMING BASED PROBLEM:**

**(10 Marks)**

4. Use Python programming language to perform the tasks given below:

**Note:** Save the above data given in Dataset: Students' Exam Scores in a CSV file (students\_scores.csv), save it, and Now Write a program in Python to perform following tasks:

- a) Read students\_scores.csv into a panda DataFrame. and Display:
  - First 5 rows
  - Column names
  - Data types
- b) Generate descriptive statistics, For each subject (Physics, Chemistry, Maths, English), compute:
  - Mean
  - Median
  - Mode
  - Minimum
  - Maximum
  - Range
  - Variance
  - Standard Deviation
- c) Create a new column Total = sum of 4 subjects; also  
Also, Create Percentage =  $(\text{Total} / 400) * 100$   
Find:
  - Topper (max Total)
  - Lowest scorer (min Total)
- d) Plot a scatter diagram and histogram of Maths scores and interpret whether scores are concentrated or spread out.
- e) Compute average marks of each subject, and plot a bar chart of subject-wise averages
- f) Plot a line chart of Total marks vs Student\_ID, and Identify performance variations (ups/downs).