| | | |
|---|---|---|
| **Course Code** | **:** | **MCSL-065** |
| **Course Title** | **:** | **Data Science Lab.** |
| **Assignment Number** | **:** | **MSCDSA(I)/L-065/Lab_Assignment/2026** |
| **Maximum Marks** | **:** | **100** |
| **Weightage** | **:** | **30%** |
| **Last Date of Submission** | **:** | **31stMay, 2026 (for January session)**<br>**31stOctober, 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 aCSV 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 = (Total / 400) * 100
Find:
- Topper (max Total)
- Lowest scorer (min Total)

d) Plot a scatter diagram and histogram of Maths scores andinterpret 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, andIdentify performance variations (ups/downs).

**R-PROGRAMMING BASED PROBLEM:** **(10-Marks)**

5. Use R programming language to perform the tasks given below:

**Note:**Save the above data given in Dataset: Students' Exam Scores in aCSV file (students_scores.csv), save it, and Now Write a program in R to perform following tasks:

a) Read the CSV file into a data frame and Display:
- First 5 rows
- Structure of dataset
- Summary of dataset

b) Compute descriptive statistics (subject-wise),For each subject (Physics, Chemistry, Maths, English), calculate:
- Mean
- Median
- Mode
- Minimum
- Maximum
- Range
- Variance
- Standard Deviation

c) Create a new column Total = sum of 4 subjects; also
Also, Create Percentage = (Total / 400) * 100
Find:
- Topper (max Total)
- Lowest scorer (min Total)

d) Plot a scatter diagram and histogram of Maths scores andinterpret 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, andIdentify performance variations (ups/downs).

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

**Answer**

The Students' Exam Scores data are first entered in an Excel worksheet with Student_ID and Student_Name as text fields and Physics, Chemistry, Maths and English as numeric fields. The following calculations and charts use exactly the ten records supplied in the assignment.

### a. Create a Histogram for Maths marks

The Maths values are 85, 90, 70, 94, 58, 88, 75, 89, 63 and 90. In Excel, select the Maths column and use Insert > Statistical Chart > Histogram. With class intervals 50–59, 60–69, 70–79, 80–89 and 90–99, the frequencies are 1, 1, 2, 3 and 3 respectively. The scores are therefore concentrated mainly in the upper intervals, although two lower scores increase the spread.

### b. Plot a Line Chart for Maths scores against Student_ID.

Student_ID is kept on the horizontal axis and Maths marks on the vertical axis. The line rises from S01 to S02, falls through S03, reaches the highest point at S04 (94), drops sharply at S05 (58), and then recovers for the later students. This makes individual changes easy to compare.

### c. Calculate average marks for each subject, and Create a Bar Chart comparing subject-wise average scores.

The averages can be obtained with AVERAGE() in Excel. The calculated subject means are shown below.

| Subject | Average marks |
|---|---|
| Physics | 76.4 |
| Chemistry | 77.9 |
| Maths | 80.2 |
| English | 79.2 |

Maths has the highest average at 80.2, followed by English at 79.2, Chemistry at 77.9 and Physics at 76.4.

### d. Using average marks, create a Pie Chart showing percentage contribution of each subject.

The four subject averages total 313.7. Therefore, each contribution is calculated as subject average divided by 313.7 and multiplied by 100.

| Subject | Average | Contribution |
|---|---|---|