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

CONTENTS

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 : **MCS-061**
Course Title : **Mathematical Foundations-I**
Assignment Number : **MSCDSA (I)/061/Assign/2026**
Maximum Marks : **100**
Weightage : **30%**
Last Date of Submission : **30th April, 2026 (for January session)**
31st October, 2026 (for July session)

Note: There are four questions in this assignment, carrying a total of 80 marks. The remaining 20 marks are for viva-voce. Answer all the questions.

Q1 a) What is set? Explain the Finite set and the Infinite set with example. **(3 Marks)**

b) What is Power Set? Find P(A) for A = { a,b,c,e,f,g) **(2 Marks)**

c) What is a function? Explain the following types of functions with examples. **(3 Marks)**

i) Surgective
ii) Injective
iii) Bijective

d) Explain how to draw a graph of a linear function. Draw graph of (x) = 5x +2 **(3 Marks)**
e) Find the inverse of the following function: **(3 Marks)**

$$f(x) = \frac{x^2 + 9}{x - 5}, x \neq 5$$

f) What is a relation? Explain an equivalence relation with the help of an example. **(3 Marks)**

g) Draw Venn diagram to represent the followings: **(3 Marks)**

i) $A \subseteq B$
ii) $(A \subset B$
iii) $(A \cap B \cap C) \cap (A \cup B \cap C)$

h) Explain whether the function $f(x) = x^2 + 2$ is one-one or not. **(2 Marks)**
i) Let f and g be the two functions such that $f(x) = x^2 + 5$ and $g(x) = 2x+5$. Define fof, fog, gof and gog. **(3 Marks)**

Q2 a) Show that: $\begin{vmatrix} b+c & c+a & a+b \\ c+a & a+b & b+c \\ a+b & b+c & c+a \end{vmatrix} = 2 \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$ **(2 Marks)**

b) If A= $\begin{bmatrix} -1 & 2 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, show that $A^2 = A^{-1}$ **(2 Marks)**

c) What is a matrix? Explain Upper Triangular Matrix and Lower Triangular Matrix with examples. **(2 Marks)**

d) Find the inverse of the matrix $A = \begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{pmatrix}$, if it exists. **(3 Marks)**

e) Solve the following system of equations by using the Matrix Method. **(3 Marks)**

$3x + 4y + 7z = 14$
$2x - y + 3z = 4$
$2x + 2y - 3z = 0$

f) Explain how to find sum of n terms of a G.P. **(2 Marks)**

g) If m times the mth term of an A.P. is n times its nth term, show that $(m + n)^{th}$ term of the A.P. is zero. **(3 Marks)**

h) Explain what minors and cofactors are in a determinant? Find minor of each element of the following matrix. **(3 Marks)**

$$\begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{pmatrix}$$

Q3 a) What is vector? Explain how to find the inner product of vectors with an example. **(3 Marks)**

b) What is Eigen Vector? What is Eigen Value? Explain the characteristic equation. **(3 Marks)**

c) Explain the fundamental principle of multiplication. **(3 Marks)**

d) Prove $^{n+1}C_r = {^nC_r} + {^nC_r}n$ **(3 Marks)**

e) Expend $\left(x + \frac{1}{x}\right)^4$ by binomial theorem. **(3 Marks)**

Q4 a) Explain the concept of limit. **(2 Marks)**

b) Show that if $y = ae^{mx} + be^{-mx}$, prove that $d^2y/dx^2 = m^2 y$ **(2 Marks)**

c) Evaluate the integrals: **(3 Marks)**

i) $\int e^x(e^x + 7)^5\ dx$.

ii) $\int \frac{1}{x\ log\ x}dx$

d) Evaluate $\int(x + 1)e^x(xe^x + 5)^4\ dx$. **(2 Marks)**

e) Evaluate the followings: **(3 Marks)**

i) $\int(x + 1)e^x(xe^x + 5)^4\ dx$.   ii) $\int \frac{x^6 - x^3 + 1}{x^2}dx$

f) Evaluate the followings: **(3 Marks)**

i) $\lim\limits_{n\to 0} \frac{|x|}{x}$does not exist   ii) $f(x) = |x|$ is continuous at $x = 0$.

iii) $\lim\limits_{x\to 0} f(x)$, where $f(x) = \begin{cases} 2 - x^2, x \neq 0 \\ 2, x = 0 \end{cases}$

g) What is a derivative? Explain its meaning and geometrical interpretation with an example. **(3 Marks)**

h) What is definite integration? Give a geometrical interpretation of it with the help of an example. **(2 Marks)**

4

MASTER OF SCIENCE
DATA SCIENCE AND ANALYTICS
(MSCDSA)

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

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	:	MCS-062
Course Title	:	**Introduction to Data Science**
Assignment Number	:	**MSCDSA(I)/062/Assign/2026**
Maximum Marks	:	**100**
Weightage	:	**30%**
Last Dates for Submission	:	**30th April, 2026 (for January session)**
		31st October, 2026 (for July session)

There are 25 questions in this assignment which carries 80 marks. Question No. 1 to 20 are of 3 marks each, remaining question no. 21 to 25 are of 4 marks each. Rest 20 marks are for viva-voce.

Answer all the questions. Questions are application oriented. Some questions require numerical calculations. You may use illustrations and diagrams to enhance the explanations. Include the screen layouts also along with your assignment responses. Please go through the guidelines regarding assignments given in the Programme Guide for the format of presentation.

Q1: Identify a real-life problem from healthcare or banking and explain how the Data Science lifecycle can be applied to solve it.

Q2: A company wants to predict future sales instead of analyzing past trends. Should they use Data Analytics or Data Science? Justify.

Q3: Classify the following data: (a) Customer feedback text, (b) Monthly temperature readings, (c) Employee ID numbers. Mention data type and scale.

Q4: A dataset contains values: 20, 25, 30, 35, 40. Identify whether the data is discrete or continuous and calculate the range.

Q5: Suggest suitable data acquisition tools for collecting air quality data in a smart city project and justify your choice.

Q6: Explain how sampling and signal conditioning are used in a temperature monitoring system.

Q7: List the steps you would follow to clean and integrate sales data collected from two different e-commerce platforms.

Q8: Normalize the values 200, 400, 600, 800 using Min–Max normalization.

Q9: Perform Exploratory Data Analysis (EDA) for a student marks dataset. What insights can be obtained?

Q10: Calculate mean and standard deviation for the dataset: 10, 12, 15, 18, 20.

Q11: A bank wants to predict loan default. Is this inferential or predictive analysis? Explain the reasoning.

Q12: A sample mean is 52, population mean is 50, and standard deviation is 4 for n = 16. Compute the Z-value.

Q13: Which visualization technique will you use to compare monthly sales across two years and why?

Q14: Draw a histogram for the following data: 5, 7, 8, 10, 12, 15, 18, 20 and interpret it.

Q15: Using Excel, explain how frequency distribution and central tendency help in summarizing student performance.

Q16: A dataset shows X = [1,2,3,4,5] and Y = [2,4,6,8,10]. Calculate the correlation coefficient and interpret.

Q17: Differentiate between RDBMS and NoSQL databases with an example from an e-commerce system.

Q18: Explain the 5V characteristics of Big Data with one real-world example.

Q19: Explain data stream mining and its importance in real-time traffic monitoring systems.

Q20: Compare Python, R, Tableau, and Power BI for data analysis and visualization use cases.

Q21: Given the data set: 2, 4, 4, 6, 8, 10

    a. Calculate the Mean, Median, and Mode.
    b. Find the Variance and Standard Deviation of the data.
    c. Determine the Interquartile Range (IQR).
    d. Briefly explain what each measure tells you about the data.

Q22: suppose we have a sample of 25 students with a sample mean score of 75 and a sample standard deviation of 10. We wish to test whether the mean score differs from 70. State the Null Hypothesis and the alternate hypothesis. Use t-test to test whether the null hypothesis will be accepted or not. Given, 24 degrees of freedom, the critical t-value at a 5% level is approximately 2.064.

Q23: Explain Logistic regression, and use it to Predict pass (1) or fail (0) based on study hours. The relevant data related to the study hours and result(pass/fail) is given below:
                Study Hours (X): 2, 4, 6, 8, 10;
                  Pass/Fail (Y): 0, 0, 1, 1, 1

Q24: Describe the utility of time series forecasting in data science. Also, compare the Moving Average and Exponential Smoothing techniques for time series forecasting. Also, perform the following tasks:

    a. Given the Sales for 5 months as: [100, 120, 130, 110, 140] compute the 3-month moving average for month 6.
    b. Write and explain the Formula for Exponential smoothing

Q25: A company wants to compare the variability in salaries between two departments: Department A and Department B. The company randomly samples 10 employees from each department and records their salaries (in $1000s):
        · Department A Salaries (Sample 1): 45, 48, 50, 52, 49, 46, 51, 47, 50, 48
        · Department B Salaries (Sample 2): 55, 58, 54, 60, 57, 53, 59, 56, 55, 58

The company wants to know if there is a significant difference in the variances of salaries between the two departments. Use F-test to compare the variances and test the null hypothesis.
Given the Critical values ($\alpha = 0.05$, two-tailed)

$$F_{upper} = F_{0.975}(9,9) = 4.026 \&$$
$$F_{lower} = F_{0.025}(9,9) = 0.248.$$

Decision rule: Reject H₀ if $F < 0.248$ or $F > 4.026$.

MASTER OF SCIENCE
DATA SCIENCE AND ANALYTICS
(MSCDSA)

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

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	:	**MCS-063**
Course Title	:	**Data Structures using Python**
Assignment Number	:	**MSCDSA (I)/063/Assign/2026**
Maximum Marks	:	**100**
Weightage	:	**25%**
Last Date of Submission	:	**30th April, 2026 (for January session)**
		31st October, 2026 (for July session)

There are four questions in this assignment, which carry80 marks. Each question carries 20 marks. Rest 20 marks are for viva voce. All algorithms should be written nearer to Python programming language. You may use illustrations and diagrams to enhance the explanations, if necessary.

Q1: Explain the process of exception handling in Python with an example. **(20 Marks)**

Q2: What is a Tree ? How does it differ from a Binary Tree ? Explain the process of converting a Tree into a Binary Tree. **(20 Marks)**

Q3: Explain built in sorting functions of Python with examples. **(20 Marks)**

Q4: What is a Minimum Cost Spanning Tree? Explain with an example. **(20 Marks)**

MASTER OF SCIENCE
DATA SCIENCE AND ANALYTICS
(MSCDSA)

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

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	:	MCS-207
Course Title	:	**Database Management Systems**
Assignment Number	:	**MSCDSA (I)/207/Assignment/2026**
Maximum Marks	:	**100**
Weightage	:	**25%**
Last Dates for Submission	:	**30th April, 2026 (for January session)**
		31st October, 2026 (for July session)

There are four questions in this assignment, which carries 80 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 the presentation. The answer to each part of the question should be confined to about 300 words. Make suitable assumption, if any.

Q1: Foundations of Data-Centric Database Systems   **(5 × 4 = 20 Marks)**

    a. Data scientists often work with large, evolving datasets. Explain the **limitations of traditional file-based systems** in the context of analytics and machine learning workflows. How does a DBMS address these limitations?

    b. Explain the following concepts with respect to the **relational data model**, giving one data-science–oriented example for each:

- Candidate Key

- Functional Dependency

- Referential Integrity

- Selection Operation

- Projection Operation

    c. A **health analytics platform** maintains data about patients, diagnostic tests, doctors, and test results. Analysts want to query trends across diseases, age groups, and regions. Design an **ER diagram** for this system. Clearly identify entities, relationships, key attributes, and constraints. State assumptions made.

    d. Convert the ER diagram designed in part (c) into **normalized relations up to 3NF**, clearly indicating primary and foreign keys.

    e. Explain the importance of **indexes** in analytical databases. Differentiate between primary index, secondary index, and clustering index with suitable examples.

Q2: Data Normalization, Dependencies, and SQL for Analytics   **(4+4+4 = 20 Marks)**

    a. Consider the relation:
Dataset(DatasetID, DatasetName, Source, CollectionDate, Domain, OwnerName, OwnerEmail, UpdateFrequency)

- Identify the **primary key**
- List meaningful **functional dependencies**

b. Populate the relation with **8–10 sample records** and highlight potential **data redundancy and anomalies**.
c. Decompose the above relation into **2NF and 3NF**, explaining each step clearly.
d. Consider the following relations used in a **research publication analytics system**:

- Researcher(ResearcherID, Name, Email, Affiliation)

- Publication(PubID, Title, Year, Venue)

- Authorship(ResearcherID, PubID, AuthorOrder)

Perform the following using SQL:

- Create tables with appropriate primary and foreign keys

- Insert sample data (minimum 5 researchers, 6 publications)

- List publications by a given researcher

- Find researchers who have not authored any publication

- Find the publication with the highest number of authors

- List venues that have more than two publications

Q3: Transactions, Concurrency, and Consistency in Data Systems   (4+8+8 = 20 Marks)

a. Explain the **ACID properties** of transactions with a **data science pipeline example**, such as feature store updates or model versioning.
b. Consider the following schedule involving two transactions T1 and T2 operating on a dataset table storing aggregate metrics:

Time	T1	T2
t1	READ(X)	
t2	X = X + 50	
t3		READ(X)
t4		X = X * 1.2
t5		WRITE(X)
Time	T1	T2
t6	WRITE(X)	

     i. Compute the **final value of X** (assume initial X = 100).
    ii. Determine whether the schedule is **serializable**.
   iii. Identify the concurrency issue involved and explain its impact on analytics accuracy.

c. Explain the **Two-Phase Locking (2PL) protocol**. Discuss how it ensures consistency in multi-user analytical environments. Can deadlocks still occur? Explain with a suitable example.

Q4: Advanced Topics and Case-Based Understanding   (5 × 4 = 20 Marks)

Write short notes on the following. Support your answers with examples relevant to data science or analytics.

a. Centralized vs Distributed Databases in Large-Scale Analytics
b. Star Schema and Snowflake Schema in Data Warehousing
c. NoSQL Databases for Data Science (Explain one type with use cases)
d. Query Optimization Techniques for Analytical Queries
e. Log-Based Recovery and Checkpointing in Data-Intensive Systems

MASTER OF SCIENCE
DATA SCIENCE AND ANALYTICS
(MSCDSA)

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

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-064**
Course Title	:	**Data Structures using Python Lab**
Assignment Number	:	**MSCDSA(I)/L-064/Lab_Assign/26**
Maximum Marks	:	**100**
Weightage	:	**30%**
Last Date of Submission	:	**30th April, 2026 (for January session)**
		31st October, 2026 (for July session)

There are two questions in this assignment carrying a total of 40 marks. Each question carries 20 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.

Q1: Write a program in Python for addition of two matrices. **(20 Marks)**

Q2: Write a program in Python for implementation of a singly linked list. **(20 Marks)**

MASTER OF SCIENCE
DATA SCIENCE AND ANALYTICS
(MSCDSA)

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

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.

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 = (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, andIdentify performance variations (ups/downs).