

Question 1

The data provided [here](#) contains feedback from a particular class regarding a newly launched platform by the tech team. Note that only 60% of the class responded. The feedback is rated as follows:

- Highly Useful/Highly Unuseful: Ratings of 6 or 7
 - Useful/Unuseful: Ratings of 4 or 5
 - Maybe Useful/Unuseful: Ratings of 1 to 3
-
- A. Using the feedback data provided, create charts that simplify the interpretation of the feedback for decision-makers, keeping in mind that they may not have technical backgrounds. Add short explanations (1-3 sentences) for the charts where necessary.
 - B. If Moringa School wanted to gain deeper insights into student sentiment about the new platform, what changes would you recommend for the feedback form? Consider improvements such as adding, removing, or modifying questions to enhance the quality and clarity of responses.
 - C. How could Moringa School determine the likelihood of students recommending the platform to their peers using the existing feedback data? Additionally, what changes could be made to the feedback form to gather more actionable insights regarding student recommendations?
 - D. Create a concise presentation summarizing your findings, recommendations, and proposed next steps based on the feedback analysis

Question 2

Moringa is looking to improve its module-to-module retention rates for its Software Development program. Students must complete five modules in the following order:

- Intro to Software Development
- Angular
- Java/Python
- Android/Django
- Professional Development

To understand and address retention challenges, we have provided [data](#) for a randomly selected class, including the following metrics for each student in each module:

Metric	Description	Type
Attendance	Percentage of times present for each attendance taken	Percentage
IP Grade Average	Average of the combined set of Independent Project grades a student has obtained	Percentage
Total Activity Time	Total time (in seconds) a student spends engaging with courseware	Integer
Participation Level	Number (0–3) representing relative participation in learning activities	Integer
Page View Level	Number (0–3) representing the amount of content viewed relative to others	Integer

Your task is to analyze this data to uncover insights and propose actionable recommendations.

A. Clean the data with these assumptions

- a. Students with no metrics for a module are presumed to have dropped out.
- b. Students scoring below **50%** in a module and not enrolling in the next are presumed to have failed.
- c. Module progression rules
 - i. Angular graduates **can** proceed to Java or Python, but **not** both.
 - ii. Java leads to Android, **and** Python leads to Django.
 - iii. Graduates of Android/Django progress to Professional Development.

Provide the cleaned dataset and a summary of your cleaning process.

B. Calculate the retention rate between each module (e.g., Intro → Angular) and the overall retention. Present your results with **clear** visualizations (e.g., bar charts or flow diagrams).

C. Identify the top two metrics most correlated with retention.

Question 3

Moringa school collects data on students, courses, and grades across multiple modules. The database has the following relevant tables

1. Students
 - a. student_id

- b. first_name
 - c. last_name
 - d. grade_level
 - e. Enrollment_date
- 2. Courses
 - a. course_id
 - b. course_name
 - c. mentor_name
- 3. Enrollments
 - a. student_id
 - b. course_id
 - c. enrollment_date
- 4. Grades
 - a. student_id
 - b. course_id
 - c. grade
 - d. Module
- A. Generate a report showing the following
 - a. Student Performance by Course - For each student, list their courses, the grade they received for each course in the latest module, and the average grade for each student across all their courses in the most recent module.
 - b. Student Information - Include the student's first name, last name, and grade level.
- B. Write an SQL query to generate a report that includes
 - a. Student's first name, last name, and grade level.
 - b. Each course the student is enrolled in.
 - c. The student's grade for that course in the most recent term.
 - d. The student's average grade across all their courses in the most recent term.
- C. Assume the report you created above will be visualized in Power BI(or any other similar tool).
 - a. How would you load this data into Power BI(or any other similar tool)?
 - b. What kind of visualizations would you use to best display the following
 - i. A list of students and their grades for each course.
 - ii. A summary visualization showing the average grade of students in each grade level across all courses.

- c. What steps would you take to ensure the data model in Power BI is optimized for performance when dealing with large datasets, such as students and grades across multiple years?

Submission Guidelines

Complete the tasks and submit the following

- A written document (Word or PDF) with detailed explanations, pseudocode, API requests, and any supporting materials.
- SQL queries, code snippets, or API calls as public github repos, plain text or embedded in the document.
- Any flow charts or diagrams as image files or embedded in the document.
- Include any assumptions you made while completing the tasks.
- Share any other necessary material and provide access