

Assignment - 2

Fundamentals of Data Collection and Cleaning

Programme :- BCA (AI & DS) Hon. with Research

Semester :- I

Section :- 'C'

Course :- Foundations of Data Driven Decision Making

Name :- Om Mathur

Roll No.:- 2501201026

Task-1

Design a short survey (5-10 questions) to collect feedback about student satisfaction at your college.

Student Satisfaction Survey

1. How satisfied are you with the overall academic environment of the college?

- Very Satisfied
- Satisfied
- Neutral
- Dissatisfied
- Very Dissatisfied

2. How would you rate the teaching quality of your faculty?

- Excellent
- Good
- Average
- Poor

3. Are you satisfied with the availability of classroom facilities (projectors, seating, ventilation, etc.)?

- Yes
- Partially
- No

4. How satisfied are you with the cleanliness of the campus?

- Very Satisfied
- Satisfied
- Neutral

5. How helpful is the administrative staff when you need support?

- Very Helpful
- Helpful
- Neutral
- Not Helpful

Task-2

Collect 10-15 responses (real or simulated) and store them in a spreadsheet

Response	Academic Satisfaction	Teaching Quality	Classroom Facilities	Cleanliness	Admin Helpfulness
1	Satisfied	Good	Yes	Satisfied	Helpful
2	Very Satisfied	Excellent	Yes	Very Satisfied	Very Helpful
3	Neutral	Average	Partially	Neutral	Neutral
4	Satisfied	Good	Yes	Satisfied	Helpful
5	Dissatisfied	Poor	No	Dissatisfied	Not Helpful
6	Satisfied	Good	Yes	Satisfied	Helpful
7	Very Satisfied	Excellent	Yes	Very Satisfied	Very Helpful
8	Neutral	Average	Partially	Neutral	Neutral
9	Satisfied	Good	Yes	Satisfied	Helpful
10	Very Satisfied	Excellent	Yes	Very Satisfied	Very Helpful

Task-3

Identify and fix issues such as missing data, duplicate entries or inconsistent formatting.

1 Handling Missing Values

If any entries has a value of 'NA'. Then we can change 'NA' with "No suggestion provided".

2. Removing Duplicate Entries

Check to see if we have two same response IDs.

3. Fixing Inconsistent formatting

- Standardized text (e.g., "Very Satisfied" instead of "Very satisfied" or "vsatisfied").
- Ensured consistency in Yes/No responses.
- Ensured proper capitalization across dataset.
- Remove extra spaces (e.g., "Good" → "Good").

4. Verifying Data Types

- Categorical values stored as text.
- Response ID stored as numeric.

Task-4

Write a short note on ethical considerations while collecting personnel data.

Ethical handling of personnel data is essential in any data-driven project.

During this survey, the following ethical principles must be followed:

1. Informed Consent

Students should be told why the data is being collected and how it will be used.

2. Voluntary Participation :- Participation must be optional without any pressure or consequence.

3. Privacy & Anonymity - Names and personal identifiers should not be collected unless necessary.

4. Data Security - Data must be used only for academic or research purposes, not for unrelated decisions.

5. Avoiding Bias - Question should be neutral and not force specific opinions.