



Academic Year 2025

Semester: 1

Assignment Work- 2nd

Full Name: Vikash Singh

Program: B.C.A(AI and D.S)

Roll No: 2501060151

Section:(C)

Course: Foundation of Data
Driven Decision making

Faculty: Dr. Satinder

Module 2: Fundamental of Data

Collection and cleaning

Student Sign:

Vikash Singh

Professor Sign:

Task-1: Design a short Survey (5-10 questions) to collect feedback about student satisfaction at your college.

The 5-10 questions are :-

- Q1. How satisfied are you with the quality of teaching in your program?
 - very satisfied • Neutral • Satisfied • Dissatisfied
- Q2. How clear and helpful do you find the course materials (notes, presentations, assignments)?
 - very clear and helpful • Clear • Neutral • Unclear
- Q3. How satisfied are you with the classroom and laboratory facilities?
 - very satisfied • Satisfied • Neutral • Dissatisfied
- Q4. How would you rate the availability and helpfulness of faculty and staff when you need support?
 - Excellent • Good • Average • Poor • very poor
- Q5. How satisfied are you with campus facilities (library, wi-Fi, cleanliness, etc.)?
 - very satisfied • Satisfied • Neutral • Dissatisfied
- Q6. Do you feel the campus provides a safe and comfortable environment?
 - Yes • mostly • Neutral • Not much • No
- Q7. How satisfied are you with your overall experience?
 - very satisfied • Satisfied • Neutral • Dissatisfied
- Q8. Would you recommend this college to other students?
 - Definitely Yes • Probably Yes • Not sure • Probably no

Task-2: collect 10-15 responses (real or simulated) and store them in a spreadsheet.

Student Id	Teaching quality	course materials clarity	Faculty support	campus safety	campus facilities	overall satisfaction
1	satisfied	clear	good	satisfied yes	satisfied	satisfied
2	very satisfied	very clear	Excellent	yes	very satisfied	very satisfied
3	Neutral	Neutral	Average	mostly	Neutral	Neutral
4	dissatisfied	unclear	Poor	Not much	dissatisfied	dissatisfied
5	satisfied	clear	Good	yes	satisfied	satisfied
6	very satisfied	very clear	Excellent	yes	satisfied	very satisfied
7	Neutral	Neutral	Average	Neutral	Neutral	Neutral
8	satisfied	clear	Good	mostly	satisfied	satisfied
9.	Very satisfied	very clear	Excellent	yes	very satisfied	very satisfied
10.	dissatisfied	unclear	Poor	Neutral	Neutral	Neutral
11.	satisfied	clear	Good	yes	satisfied	satisfied
12	very satisfied	very clear	Excellent	yes	very satisfied	very satisfied
13	Neutral	Neutral	Average	Neutral	Neutral	Neutral

Task-3. Identify and fix issues such as missing data, duplicate entries, or inconsistent formatting.

Common issues identified are :-

1. Missing Data:

How to identify

- Look for blank cells/twells
- Use functions: • Excel

How to Fix

option 1: Remove rows with missing data

option 2: Replace / impute missing values.

option 3: Forward/Backward fill.

2. Duplicates Entries

How to identify

- Rows that appear more than once
- Same ID or email repeated
- Same timestamp and value repeated

How to check

• Excel: Data → Remove Duplicates

How to fix

• Remove exact duplicates

• Handle partial duplicates

3. Fixing inconsistent Formatting

Common formatting issues

- Uppercase vs Lowercase (Vikash, vikash, VIKASH)
- Date formats (15/11/2025, 11-15-25, 2025/11/15)
- Extra spaces ("Delhi" vs "Delhi ")

How to fix

- Text standardization
- Date standardization
- Number Formatting
- Categorical Standardization

Task 4: Write a short note on ethical considerations while collecting personal Data.

Collecting personal data requires careful attention to ethics to ensure privacy, trust and responsible use. First, informed consent must be obtained, meaning individuals should clearly understand what data is being collected, why it is needed, and how it will be used. Data collectors should follow the principle of data minimization, collecting only the information that is genuinely necessary. Ensuring data privacy and security is crucial - personal data must be protected from unauthorized access, misuse, or breaches. Individuals should also be given the right to access, update, or delete their personal information. Data must be used fairly and transparently, without deception, or exploitation and organizations should follow relevant laws and guidelines such as GDPR or local privacy regulations. Overall, ethical data collection builds trust and respects the dignity and rights of individuals.