

Data Cleaning & Essential Functions

1. What is data cleaning, and why is it important in data analysis ?

- What are the potential consequences of analyzing unclean or messy data?
- Explain the common steps involved in cleaning and organizing data.

Answer: -

- Data Cleaning:-
 - Data cleaning is a process of removing errors , duplicates and inconsistencies from the data .
 - It is essential as it ensures that the data is accurate and reliable for analysis.
 - It ensures that the data is unbiased.
- Why it is important :-
 - Ensures accurate analysis and results
 - Improves data quality and reliability
 - Saves time during analysis
 - Helps in making correct business decisions
- Consequences of analyzing unclean data :-
 - Incorrect conclusions
 - Misleading reports
 - Poor decision-making
 - Loss of trust in data
- Common steps in data cleaning :-
 - Remove duplicate records
 - Handle missing values
 - Correct spelling and formatting errors
 - Standardize data (dates, text case, units)
 - Remove irrelevant or incorrect data
 - Validate data accuracy

2. How would you sort the following dataset first by "Department" (A-Z) and then by "Salary" (Largest to Smallest)? Write a step-by-step approach.

Answer: -

- Select the entire table (Employee, Department, Salary).
- Go to Data → Sort.

Data Cleaning & Essential Functions

- In Sort by, select Department.
- Choose A to Z.
- Click Add Level.
- In Then by, select Salary.
- Choose Largest to Smallest.
- Click OK.

3. Explain the use of text functions such as TRIM , LEFT, RIGHT, MID, and CONCAT in data Cleaning.

Answer: -

- TRIM – Removes extra spaces from text =TRIM(A1)
- LEFT – Extracts characters from the left =LEFT(A1,5)
- RIGHT – Extracts characters from the right =RIGHT(A1,4)
- MID – Extracts text from the middle =MID(A1,2,4)
- CONCAT – Combines multiple text values =CONCAT(A1," ",B1)

This function help clean and standardize text data.

4. What is the role of date functions like TODAY in managing datasets?

Answer: -

- The TODAY() function returns the current date.
- Importance :-
 - Tracks real-time data updates
 - Calculates age, deadlines, or durations
 - Helps in dynamic reports
 - Automatically updates when the file opens
- Example:-
=TODAY()

5. Apply Data Validation to restrict Quantity values to only whole numbers between 1 and 10.

- Configure an input message that appears when a user selects a cell in the "Quantity" column explaining: "Please enter a whole number between 1 and 10."
- Set up an error alert message that triggers if the user enters a number less than 1 or greater than 10, showing: "Invalid input! The quantity must be a whole number between 1 and 10."

Write a step-by-step approach for this question

Data Cleaning & Essential Functions

Customer Name	Product Name	Category	Quantity	Unit Price (\$)
Jane Smith	Shoes	Electronics		81
Isabella Moore	Laptop	Electronics		121
Daniel Davis	Sofa	Clothing		239
Alex Moore	Shoes	Electronics		500
Michael Johnson	Table Lamp	Home Decor		423
Daniel Johnson	Backpack	Electronics		160
Isabella Davis	Headphones	Electronics		348
Jane Davis	Headphones	Electronics		152
Alex Wilson	T-shirt	Home Decor		369

Answer: -

Step 1:- Select Quantity column

- Select all cells under Quantity.

Step 2:- Open Data Validation

- Go to Data – Data Validation.

Step 3:- Set validation rule

- Allow: Whole Number
- Data: Between
- Minimum: 1
- Maximum: 10

Step 4:-Input Message

- Title: Quantity Input
- Message:

- “Please enter a whole number between 1 and 10”.

Step 5:- Error Alert

- Style: Stop
- Title: Invalid Input
- Message:

- “Invalid input! The quantity must be a whole number between 1 Or 10”.

Data Cleaning & Essential Functions

Step 6:-

- Quantity input is now restricted correctly.

6. Understand and apply fundamental text functions like LEFT, RIGHT, MID, and LEN.

- Extract the first 5 characters from the string "ExcelTipsAreGreat" using the LEFT function.
- Extract the last 4 characters from "DataAnalysis.xlsx" using the RIGHT function.
- Extract the substring "Tips" from "ExcelTipsAreGreat" using the MID function.
- Count the total number of characters in the string "Hello World!" using the LEN function.
- Create a formula to extract the middle 6 characters from "12345-67890-ABCDE".

Answer: -

A. Extract first 5 characters :-

=LEFT("ExcelTipsAreGreat",5)

Result: Excel

B. Extract last 4 characters :-

=RIGHT("DataAnalysis.xlsx",4)

Result : xlsx

C. Extract "Tips" :-

=MID("ExcelTipsAreGreat",6,4)

Result : Tips

D. Count characters

=LEN("Hello World")

Result : 11

E. Extract middle 6 characters :-

=MID("12345-67890-ABCDE",7,6)

Result : 67890-

7. Understand how to combine text using CONCAT, TEXTJOIN, and the & operator.

- a. Use CONCAT to combine "Hello" and "World" with a space in between.
- b. Combine "Apple", "Banana", and "Cherry" into a single string separated by commas using TEXTJOIN.
- c. Use the & operator to create the string "2025: Excel Functions" by combining "2025", ":", " and " Excel Functions".
- d. Create a comma-separated list from the range A1:A5 using TEXTJOIN

Data Cleaning & Essential Functions

E. Combine first names in column A with last names in column B to create full names in Column C.

Answer: -

a. Use CONCAT to combine "Hello" and "World" with a space :-

=CONCAT("Hello"," "; "World")

Result : Hello World

b. Combine "Apple", "Banana", and "Cherry" using TEXTJOIN:-

=TEXTJOIN(", ", TRUE, "Apple","Banana","Cherry")

Result : Apple, Banana, cherry

c. Use the & operator to create “2025:Excel Functions”:-

= "2025" & ": " & "Excel Functions"

Result : 2025:Excel Functions

d. Create a comma-separated list from range A1:A5:-

=TEXTJOIN(", ", TRUE, A1:A5)

E. Combine first names (Column A) and last names (Column B) Into Column C

=A1 & " " & B1

8. Understanding TODAY() and NOW()

a. What is the difference between TODAY() and NOW() in Excel? Provide an example of when you would use each function.

b. If cell A1 contains the date 2025-06-10, write a formula using TODAY() to determine how many days are left until that date

c. Write an Excel formula using NOW() to display the current date and time in the format MM/DD/YYYY HH:MM AM/PM.

d. If a cell contains =TODAY(), what will happen when the worksheet is reopened the next day ?
Explain

e. You want to store a static date (today's date) in a cell without it changing every day. What keyboard shortcut should you use?

Answer: -

- a. Understanding between TODAY() And NOW()
- TODAY() returns the current date only.

Data Cleaning & Essential Functions

- NOW() returns the current date and time

Example use :-

- Use TODAY() to calculate age or due dates.
 - Use NOW() to timestamp entries or track exact time of activity.
- b. If cell A1 contains 2025-06-01, formula to determine how many days are left until that date :-
$$=A1 - \text{Today}()$$
 - c. formula using NOW() to display current date and time in format MM/DD/YYYY HH:MM AM/PM.
$$=\text{TEXT}(\text{NOW()}, "MM/DD/YYYY HH:MM AM/PM")$$
 - d. If a cell contains = TODAY(), what happens when reopened the next day ?
 - The date will automatically update to the new current date because TODAY() is a dynamic (volatile) function.
 - e. Keyboard shortcut to store a static date (that does not change):
 - Ctrl+ ; - Insert current date
 - Ctrl + Shift +; - Insert current time

Data Cleaning & Essential Functions