

Assignment-2, Data Cleaning & Essential Functions

Prob-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.

- Data cleaning is the process of identifying, correcting, or removing errors, and irrelevant data from a dataset to make it accurate, complete, and reliable for analysis.

Importance of Data Cleaning:

Ensures accurate results, improves data consistency, helps in better decision-making, reduces errors in calculations, charts, and reports, makes data ready to use for tools like Excel formulas, Pivot Tables, and charts.

Potential consequences of analyzing unclean or messy data: -

- Incorrect results, Misleading insights, Errors in Excel formula, Duplicate records, Time wastage, Loss of credibility

Common steps involved in cleaning and organizing data: -

- Remove Duplicate Data, fill missing data, Correct Data Types, Fix Inconsistent Data, Remove Errors, Trim Extra Spaces, Organize Data Structure with proper headings, Validate Data.

Prob-2

Employee	Department	Salary
Sonu	IT	4000
Pranav	HR	5000
Rahul	IT	2500

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.

- 1. Select the entire dataset
2. Go to the Data tab & click Sort.
3. In Sort by-Department, In Sort On- Cell Values, In Order- A to Z (Department)
4. Click on Add Level, then by- select Salary, In Order, choose Largest to Smallest (Salary)
5. Ensure “My data has headers” is checked & click “OK”

Employee	Department	Salary
Pranav	HR	5000
Sonu	IT	4000
Rahul	IT	2500

(Answer)

For Question 8 (B, C—Please visit to the Excel File)

Assignment-2, Data Cleaning & Essential Functions

Prob-3

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

→

Functions	USE	Importance	Example
TRIM	Removes extra spaces from text	Fixes inconsistent spacing. Helps avoid errors in lookup, sorting, and matching.	=TRIM(A1)
LEFT	Extracts characters from the left side of a text string.	Helps separate codes, prefixes, or categories. Useful when data is not properly structured.	=LEFT(A1,4) "IT-4001" → "IT-4"
RIGHT	Extracts characters from the right side of a text string.	Used to extract suffixes, years, IDs, or numbers	=RIGHT(A1,4) "EMP2025" → "2025"
MID	Extracts text from the middle of a string	Helps break combined data into meaningful parts	=MID(A1,3,2) "HR-5500" → "55"
CONCAT	Joins two or more text values into one combined text.	Combines cleaned data into a usable format	=CONCAT(A1, " ", B1) "Sonu" + "Paul" → "Sonu Paul"

Prob-4

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

- The TODAY() function returns the current date and is used to track, update, and analyze time-based data automatically.

Assignment-2, Data Cleaning & Essential Functions

Prob-5

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

- a) 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."
- b) 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."

→ **Applying Data Validation on the Quantity column: -**

1. Select the Quantity column (exclude the header).
2. Open **Data Validation**, go to the **Data tab**, Click on **Data Validation**.
3. Validation Criteria (**Allow:** Whole number, **Data:** Between, **Minimum:** 1, **Maximum:** 10)
4. Go to the **Input Message** tab, Check **Show input message when cell is selected**, **Title:** Quantity Entry; **Input Message:** Please enter a whole number between 1 and 10.
5. Go to the **Error Alert** tab, **Style:** Stop, **Title:** Invalid Input, **Error Message:** Invalid input! The quantity must be a whole number between 1 and 10.
6. Click OK & Data validation will be applied.

Prob-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".



1. String: "ExcelTipsAreGreat"
Formula: =LEFT(H2,5) [H2- ExcelTipsAreGreat]

2. String: "DataAnalysis.xlsx",
Formula: =RIGHT(H3,4) [H3- DataAnalysis.xlsx]

For Question 8 (B, C—Please visit to the Excel File)

Assignment-2, Data Cleaning & Essential Functions

3. String: "ExcelTipsAreGreat"

Formula: =MID(H2,6,4) [H2- ExcelTipsAreGreat]

4. String: "Hello World!"

Formula: =LEN(H4) [H4- Hello World!] Output= 12

5. String: "12345-67890-ABCDE"

Formula: =MID(H5,7,6) [H5-12345-67890-ABCDE] Output= 67890-

Prob-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?



A.

=TODAY()	=NOW()
Used to find current date only.	Used to find current date & time.
Calculating due dates, ageing, deadlines	Time-stamping entries, tracking login or update time
=TODAY() → 01/02/2026	=NOW() → 01/02/2026 5:37 PM.

D.

When you reopen the worksheet the next day, TODAY() automatically updates to the new current date.

Example: If you saved the sheet on Jan 2, 2026, it will show Jan 3, 2026 when reopened on the next day.

E.

Keyboard shortcut for a static date without changing everyday= **[ctrl+ shift(+)+ ;]**

For Question 8 (B, C—Please visit to the Excel File)