# Data Cleaning

# Learning Objectives

- Concatenate
- Left & Right
- Mid & Trim
- Uppercase & Lowercase
- Find & Search
- Substitute & Replace
- Len & Char
- And &  Or
- Round
- Floor
- Conditional formatting

# Concatenate in Excel

Concatenate function is used to Join the strings.

- Using Concatenate Function
  - Example : = CONCATENATE(B8," ",C8," ",D8)
  - " " is used to add space in between.

# Concatenate in Excel

- Using '&' operator
  - Example : =B8 & " " & C8 & " " & D8

- Concatenate String and numerical
  - Example :  ="Your bill is " & B16

# Left, Right in Excel

- Left function returns the number of characters defined from the string from Left.
  - Syntax : =LEFT(Text, number of characters)
  - Example 1 : =LEFT([@Column name],3)
  - Example 2 :  =LEFT("Text", 3)

- Right function returns the number of characters defined from the string from Right.
  - Syntax : =RIGHT(Text, number of characters)
  - Example 1 : =RIGHT([@Column name],3)
  - Example 2 :  =RIGHT("Text", 3)

# Mid, in Excel

- The Mid function returns a specified number of characters from the middle of a supplied text string.
  - Syntax : =MID(Text, Starting Position, number of characters)
  - Example 1 : =MID([@City],2,3)
  - Example 2 :=MID("Aniruddha",2,3)

# Trim in Excel

- TRIM function returns a text value with the leading and trailing spaces removed. It is also used to remove unnecessary spaces between words in a string.

    - Syntax : =TRIM("Text")
    - Example 1 : =TRIM("   Hello! How are you ??   ")
    - Example 2 : =TRIM("Hello!     How    are you???")

# Lower & Upper Case in Excel

- Lower Function coverts string in lower case.
    - Syntax : =LOWER("Text")
    - Example 1 : =LOWER("HELLO! How are You?")  ➜ hello! How are you?

- Upper Function coverts string in Upper case.
    - Syntax : =UPPER("Text")
    - Example 1 : =UPPER("hello! How are You?") ➜ HELLO! HOW ARE YOU?

- Proper Function coverts string in Upper case.
    - Syntax : =PROPER("Text")
    - Example 1 : =PROPER("hello! HOW are You?") ➜ Hello! How Are You?

# Find & Search in Excel

- Find function is used to find position of the particular substring within a string.
  - Syntax : =FIND("Find text","within text")
  - Example : = FIND("Hello",''Hello!,How are you?") ➔ 1
  - Example : = FIND("hello",''Hello!,How are you?") ➔ ERROR
  - *Find is CASE SENSITIVE*

- Search Function converts string in Upper case.
  - Syntax : =SEARCH("Find text","within text")
  - Example : =SEARCH("hello",''Hello!,How are you?")➔ 1
  - Example  : =SEARCH("l?o ","Hello!,How are you?") ➔ 3
  - ? Is used as any single character and * can be used as any series of characters
  - *Search is INCASE SENSITIVE*

# Substitute & Replace in Excel

- When we know the text to be replaced we use Substitute function.
  - Syntax : =SUBSTITUTE(Cell name , "Old text", "New text", Instance)
  - Example : = SUBSTITUTE(C1,"Hello","Hi") ➔ Hello will be replaced by Hi
  - Example : = SUBSTITUTE(C1,"Hello","Hi",2) ➔ 2nd Hello will be replaced by Hi
  - *Substitute is CASE SENSITIVE*

- When we know the position of the text to be replaced we use Replace Function.
  - Syntax : =REPLACE(Cell name , start number, number of characters , "New text")
  - Example : = REPLACE(C1,4,3,"Hi")➔ Starting from 4th position & length 3 will be replaced by Hi.

# Len & Char in Excel

- Len Function is used to calculate the length of the string
  - Syntax : =LEN("text")
  - Example : = LEN("Bangalore") ➜ 9
  - Example : = LEN("Goa") ➜ 3

- We can retrieve the character from its ASCII value using char function.
  - Syntax : =CHAR(ASCII Code)
  - Example : = CHAR(65)➜ a
  - Example : = CHAR(64)➜ @

# Round in Excel

- It returns a number rounded to a given number of digits. The ROUND function can round to either left or right of the decimal point.

- Syntax : =ROUND (number, number digit)

  – Example : = ROUND(6.7845,1) ➔ 6.8 Rounding to nearest 1 decimal place
  – Example : = ROUND(6.7845,2) ➔ 6.78 Rounding to nearest 2 decimal place
  – Example : = ROUND(6.7845,3) ➔ 6.785 Rounding to nearest 3 decimal place
  – Example : = ROUND(6.7845,0) ➔ 7 Rounding to nearest whole number
  – Example : = ROUND(9518.5,-1) ➔ 9520 Rounding to nearest 10
  – Example : = ROUND(9518.5,-2) ➔ 9500 Rounding to nearest 100
  – Example : = ROUND(9518.5,-3) ➔ 10000 Rounding to nearest 1000

# Floor in Excel

- It rounds a given number down to the nearest specified multiple.
- Syntax : =FLOOR (number, multiple)
  - Example : = FLOOR (232,3) ➜ 231 Rounding to nearest multiple of 3
  - Example : = FLOOR (324,2) ➜ 324 Rounding to nearest multiple of 2
  - Example : = FLOOR (441,5) ➜ 440 Rounding to nearest multiple of 5
  - Example : = FLOOR (654,100) ➜ 700 Rounding to nearest multiple of 100
  - Example : = FLOOR (-7.8,-2) ➜ -6 Rounding to nearest nearest multiple of 2 in right.
  - Example : = FLOOR (-7.8,2) ➜ -8 Rounding to nearest multiple of 2 in left.
  - Note : This works only with the negative numbers

# AND & OR in Excel

- AND means when **both** of the conditions are True then only it is True other wise it is False.

- OR means when **any** one of the condition is True then its True or when both of them are False then its False.

Below is the example of AND and OR operators

- Type 3 : Using AND Operator
  - Example = =IF(AND([@Maths]>40,[@Physics]>40),"Pass","Fail")
- Type 4 : Using OR Operator
  - Example = =IF(OR([@Maths]>40,[@Physics]>40),"Pass","Fail")

# Conditional Formatting

- It enables you to highlight cells with a certain color, depending on the cell's value.

  – With Highlight Cells Rules  we can highlight cells greater/less/equal/not equal than certain value.

  – With Top/Bottom Rules we can highlight top/bottom 10% cells , etc.

  *Conditional Formatting button is present in home tab of Excel.*

# The End