**Excel**

**To create a random number *[rand(), randbetween()]***

1. **Rand() :**- To create a random number between 0 to 1 we need to use rand function.

**Syntax**: =rand()

1. **Rand\_between():**

**Syntax:** =Rand\_between[bottom, top]

To create a random numbers according to user arguments (number between lowest number and largest number).

**HOW TO REMOVE THE "BACKGROUND FORMULA"**

**Mathematical Functions**

1. **Sum:** It is used to add the values.
   1. **Syntax:** =sum(num1,num2………………………….)
2. **Count**: It counts only numerical data.
   1. **Syntax**: =Count(value1, value2 ………………….)
      1. **Ex**: =Count(1, 2, 4) 🡪 Ans: 3
   2. **Syntax**: =CountA(value1, value2 ………………..)
      1. **Ex**: =CountA(Rohan, smith, death) 🡪 Ans: 3
3. **Min:** It displays Min value.
   1. **Syntax:** =Min(value1, value2 ……………………….)
4. Max: It displays Max value.
   1. **Syntax:** =Max(value1, value2 ……………………….)

**Selection of data:**

1. **Single Cell Selection:**
2. Shift +
3. Shift +
4. Shift +
5. Shift +
6. **Multiple Cell Selection:**
7. Shit + Ctrl +
8. Shift + Ctrl +
9. Shift + Ctrl +
10. Shift + Ctrl +

**CELL REFERENCING**

|  |  |  |
| --- | --- | --- |
| **SAL** | **Annual sal logic** | **TOTAL** |
| **100** | **A1\*12** | **1200** |
| **200** | **A2\*12** | **2400** |

Press F4

|  |  |  |
| --- | --- | --- |
| **A(SAL)** | **B** |  |
| **100** | **12** | **A2 \* $B$2 = 1200** |
| **200** |  |  |

**SUMif 🡪 Summation based on condition**

1. **Syntax:-**

**sumif(range, criteria, sum range)**

1. **Range** 🡪 **some set of cell/column/columns.**

🡪**select 1st column/Multiple columns.**

1. **Criteria** 🡪 **Condition**
2. **Sum range** 🡪 **Select only one column (numerical)**
3. **Avg range** 🡪 **Select only one column (numerical)**

**AVERAGEIF:-**

**Finding Average based on condition**

**Syntax:-**

**=Averageif(range, criteria, avg\_range)**

**Range 🡪 some set of cell/column/columns.**

* + **select 1st column/Multiple columns**

**Criteria 🡪 Condition**

**Avg\_range 🡪 select one column**

**Sumifs:-**

**Summation based on condition.**

**Syntax:-**

**Sumifs(sum\_range, criteria\_range 1 , criteria1, criteria\_range2 , criteria 2)**

**Sum\_range=**1cloumn (number).

**criteria\_range1=**1st condition col (where your 1st condition is present).

**criteria1=**1st condition.

**criteria\_range2=**2nd condition col (where your 2nd condition is present).

**criteria2=**2nd condition.

Ex:-

Total sal of Sneha in “ Analyst”[ sal is sum,Sneha is if, “Analyst” is also if]

If we have 2 ifs then we should use sumifs

=sumifs[salary col, namecol, “Sneha”, dept col, “Analyst”]

Assessment:-

Find the total profit of chairs?

Find the average profit of south region?

How many orders where shipped to new York city?

Find the total profit of technology in south region?

Find an average sales value of tables in east region?

How many orders where shipped to Florida using 1st class mode?

**Date and time functions:-**

Date, month , year, day, days , text, today, now, edate, emonth ..

DATE FUNCTION:-date()

It converts number into a proper date format.

Syntax:-

Date[year, month, date].

Date[2004,10,12].

12-10-2004

DAY FUNCTION:-

It use to extract the day from date format.

Syntax:-

=Day[serial number]

Day[“12-10-2004”]

12.

MONTH FUNCTION:-

It extract month from the date format.

Syntax:-

=month[serial number]

Month[“12-10-2004”]

10.

YEAR FUNCTION:-

Year function extract the year from the date format.

Syntax:-

=year[serial number]

Year[“12-10-2004”]

2004.

DAYS FUNCTION:-

It is used to find the difference between 2 dates.

Syntax:-

=days[end date, start date]

Days[“today (), “12-10-2004”]

TODAY FUNCTION:-

It gives you the current date.

Syntax:-

=Today()

NOW FUNCTION:-

It gives you both the current date and time according to your system.

Syntax:-

=Now()

TEXT FUNCTION:-

Syntax:-

=text(value, format text)

“12-10-2004”

D-12

dd-12

ddd-teu

dddd-teusday.

m-10

mm-10

mmm-oct

mmmm-october.

y/yy-04.

yyy/yyyy-2004.

EDATE():-

It returns the serial number of the date , that is indicated number of months before.

Syntax:-

=edate(start date, months)

Edate(“12-10-2004”, 0)

12-10-2004.

Edate(“12-10-2004”,1)

12-11-2004.

EMONTHS():-

Emonths returns the serial number of the last day of the month, before or after specified number of months.

Syntax:-

=emonth(starting date ,month)

emonth(“12-10-2004”, 0)

31-10-2004.

emonth(“12-10-2004”, 1)

30-11-2004.

TEXT FUNCTION:-

Lower, upper, right, left, trim, mid, length, concatenate, proper, find, search, substitute, replace.

LOWER ():-

It converts the text into lower case.

Syntax:-

=lower(text)

Ex:-

=lower(“SHAN”) //shan

NOTE:--

TEXT DATA SHOULD BE PROVIDED IN THE DOUBLE COTETIONS.

NUMARICAL DATA CANOT BE PROVIDED IN THE DOUBLE COTETIONS.

UPPER ():-

It converts the text data into upper case.

Syntax:-

=upper(text)

Ex:-

=upper(“Shan”) //SHAN

TRIM():-

It removes the extra spaces from the text.

Syntax:-

=trim(text)

Ex:-

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | S | H | A | N |  |

=trim(“ SHAN ”) // SHAN

PROPER():-

It converts the data or a text into the proper case.

Syntax();-

=proper(text)

|  |
| --- |
| YaSeen rAShad |

Ex:-

=proper(“YaSeen rAShad ”) //Yaseen Rashad.

USING MULTIPLE FUNCTIONS AT A TIME: (SUB\_QUERY):-

Query written inside the query.

It has two quires that is inner and the outer query.

WORKING PRINCIPLE:-

In sub\_query the inner query executes 1st and it provides the partial output

This partial output is sent as an input for outer query

Outer query executes and then provides a complete output.

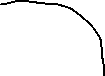
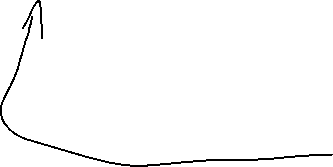
|  |
| --- |
| YASeen rAShAd |

Formula:--=trim(proper(text))/// =proper(trim(text)

=trim(proper(“ YASeen rAShAd ”)) // Yaseen Rashad

INPUT TO OUTER QUERY

1ST EXECUTES



PARTIAL OUTPUT

INNER QUERY

OUTER QUERY

RIGHT():-

Its used to execute the right most characters from the selected text.

Syntax:-

=right(text,number of characters)

Ex:-

Rashad

=right(“Rashad,1) // d

=right(“Rashad,2) //ad

NOTE:--

SPACE IS ALSO CONSIDERED AS A CHARACTER.

IF YOUR NOT PROVIDING THE NUMBER OF CHARACTERS THEN BY DEFUALT IT WILL TAKE ONE CHARACTER AND DISPLAY.

LEFT():-

Its used to execute the left most characters from the selected text.

Syntax:-

=left(text,number of characters)

Ex:-

Rashad

=left(“Rashad,1) // R

=left(“Rashad,2) //Ra

Length():-

It is used to count the number extract the length of string or text.

Syntax:-

Len(text)

Ex:-

=Len(“Rohith”) 🡪 6

Concatenate it joins or combines 2 strings or text.

Syntax:-

=Concatenate(text 1, text2,text3,………………)

Ex:-

=Concatenate(“Rohith”, “ ”, “Yash”) 🡪 Rohith Yash

Mid():-

It is used to extract the middle characters but, based upon the start position and number of characters provided by the users.

Syntax:-

=Mid(text, start\_no, no of char)

Find():-

It returns the starting position of one text within another text.

Note :-

Find is case sensitive.

Syntax:-

=FIND(find\_text,within\_text,start\_no)

EX:-

=FIND(“a”, “Yaseen”, 1-3) 🡪 1=Find(“A”, “ ”, 3) 🡪 Value Error.

Search():-

It returns no of char in which a specific char or text is first found  
Note :-

Search is not case sensitive

Syntax: -

=Search(Find\_text, within text, start no)

Ex:-

=Search(“R”, “Rohith”)

Replace()-

It replaces the part of text with some another text

Syntax:-

=Replace(old text, start\_no, no of char, new text)

LOOKUP():-

Lookup;-looking up for a value or a data .

TYPES :-

* V LOOKUP(VERTICAL LOOKUP):-

Looking up for a value in vertical direction.

Syntax:-

=vlookup(lookup value, table array, column index number, range lookup)

1. Lookup value:-condition or criteria.
2. Table array:- it starts from where your look value is present.
3. Column index number:- result column, it also start from where your lookup value is present.
4. Range lookup:-select false or 0 .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | NAME | LOCATION | DEPT | SAL |
| 001 | Santu | Bangalore | DS | 1000 |
| 002 | Kavya | Hyderabad | DA | 500 |
| 003 | Nikhil | Che | HR | 100 |
| 004 | Santu | Bangladesh | DS | 1000 |
| 005 | Ramu | Bangalore | DA | 1500 |
| 006 | Raju | Hyderabad | DS | 1000 |

NOTE:-

It works from left to right direction.

Note:-

When ever we have duplicate values it gives up most value.

Find the sal of santu find the dept of Nikhil.

=VLOOKUP(“Santu”,b1,e7,4,0)

* H LOOKUP(HORIZONTAL LOOKUP):-

It looks up for a value in a horizontal values.(Row wise).

It works both from left to right and right to left.

Syntax:-= =hlookup(lookup value, table array, row index number, range lookup)

1. Lookup value:-column name is considered as lookup value.
2. Table array:- select entire table.(look the table array using f4 or fn+f4).
3. Row index number:- row number.
4. Range lookup:-select false or 0 .

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E |
| 1 | ID | NAME | LOCATION | DEPT | SAL |
| 2 | 001 | Nikil | Bangalore | DS | 1000 |
| 3 | 002 | Nitin | Hyderabad | DA | 500 |
| 4 | 003 | Nayana | Che | HR | 100 |
| 5 | 004 | Vidya | Bangladesh | DS | 1000 |
| 6 | 005 | Sneha | Bangalore | DA | 1500 |
| 7 | 006 | Vidya | Hyderabad | DS | 1000 |

Find 3rd row data like SAL, ID, DEPT.

=hlookup(“sal”, a1:e7,3,0)

500, 002, DA.

XLOOKUP():-

It’s the combination of v and h lookup.

Syntax:-

=x(lookup value, lookup array, return array)e6

* Lookup value:- condition.
* Lookup array:- where your condition is present that column is lookup array.
* Return array:- result column.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | A | B | C | D |
| 1 | ID | NAME | SAL | DEPT |
| 2 | 001 | Raju | 500 | Dev |
| 3 | 002 | Ram | 500 | Dev |
| 4 | 003 | Rahul | 1000 | DS |

=(“Rahul”, B1:B4,C1:C4)

500.

NOTE:-

Vlookup, xlookup ,match(), index(match) gives you the top most values when duplicate are present.

INDEX():-

Syntax:-

=index(array, row number)

* Array:- selects 1 column\ result column.
* Row number :- number of rows will be selected by the user.

|  |  |  |
| --- | --- | --- |
|  | A | B |
| 1 | Product ID | Product Name |
| 2 | CA-001 | Blue pen |
| 3 | CA-002 | Black pen |
| 4 | CA-003 | Pencil |
| 5 | CA-004 | Sharpener |
| 6 | CA-005 | Pencil |

PIVOT TABLE AND PIVOT CHARTS:-

Pivot tables represents the data in the Tabler format(table format), numerical format and which allows you to aggregate and summarise the columns and rows.

PIVOT CHARTS:-

Pivot charts it represent your data using graphs or charts.

HOW TO INSERT TABLE / CHARTS

Select entire data

Go to the insert tab , from that select pivot table / chart .

It will ask for new sheet, at first it will be in new sheet.

Then click okay.

Select entire data{shift+ ctrl+🡪,down arrow}

Go to the insert tab , from that select pivot table / chart .

It will ask for new sheet, at first it will be in new sheet.

Then click okay.

Existing sheet(for 2nd time)

It will ask for location.

Go for sheet or initialy give sheet name as PT.

Select empty cell and click okay.

CONDITIONS:-

* Drag numerical columns.
* Drag id columns.
* It should not be empty.

Rows field and column field accepts only text formate.

NOTE:-

If you want to filter the numerical data or a text data then drag the columns into rows or column field.

**STEPS INVOLVED IN PROJECT:-**

* Gather the data and requirements.
* Data cleaning and data handling.
* Data transformation.
* Data modling and data analysis.
* Data visualization.
* Report creation.
* Publishing the report into services.
* Creation of dash board.
* Sharing with the end users / stake holders/ clients.

**DATA CLEANING:-**

Removing the Duplicate rows, fixing errors, correcting the datatypes and handling the missing or null values.

DATA TRANSFERMATION:-

It can be as simple as removing the column or filtering the rows, or as common as using the first row as table headers.

There are also an advance transformation options such as merge, append, group by, pivot and un pivot.

NOTE:-

* Replacing the values can be done using mean, median, mode()s.

Mean and median uses to numeric values..

Mode uses to text values.

* Deleting or removing can be done whenever primary key having null values

Ex:-

Eid, cid, etc…

* Removing the extra spaces from the data can be done using trim function.

**DATA CLEAMIMG STEPS FOR THE DATA SET(DATA CLEANING).**

STEP1:- Align the column properly.

STEP2:-Using count blank() check for any empty values or null values are present in your data set.

STEP3:-To verify that in which we have blank values , we need to apply the filters.

**Client column:-**

* Insert new column in that new column use lower () and convert the data into lower case
* Remove the data present inside the brackets(use ctrl+h)
* In find what type (\*).
* In replace “keep it empty”
* Copy the data which is converted into the lower case then paste it as values ( to remove background formulas).

**Contact column:-**

* using a trim and proper () we need to remove the extra spaces and convert the camel case data into a proper data.

**Department column:-**

* insert a new column.
* Select the department column 🡪 go to DATA TAB 🡪 TEXT TO COLUMN 🡪 use DELIMETER 🡪 then tap NEXT 🡪 choose other option and give “\_” 🡪 then tap NEXT 🡪 and tap FINISH.
* It will divide the column into two 1st is department itself and 2nd is state.

**Payment:-**

* check if the values are empty if yes the replace with NA using replace function using(ctrl+H).

**Revenue:-**

* check if the values are empty if yes the replace with NA using replace function using(ctrl+H).

**Profit margin:-**

* type = in the next column and use IFERROR() and select ‘I2,”NA” ‘
* Once you get the answer select it and go to numbers group on the home row, select “%” and later increase decimal to .00.

**Data Transformation:** It is a process of cleaning, shaping and modifying the data before loading it into a data model for reporting and visualization.

Note: It can be performed in Power Query editor.

Qs: Why Data Transformation is needed?

Ans: Raw Data/ Cleaned Data contains errors, duplicate values, missing values or irrelevant feeds.

Different sources may use different formats(Dates, Numbers, Currencies, Date & Time).

Business reporting requires structured, cleaned and optimized data.

Common Data Transformation Steps:

Step 1: Use first row as headers. :- It is used to convert the row data into headers.

Step 2: Data Type/ Detect Data Type :- To change the data type.

Step 3: Choose column’s :- it is used to display only the columns whichever the user wants to display

Step 4: Remove Column :- It is used to remove the selected column.

Step 5: Keep Rows :- It is used to display the number of rows according to the users selection.

Step 6: Remove Rows :- It is used to remove the number of rows according to the users selection.

Step 7: Sorting the data :- It is used to sort the data from ascending to descending or A-Z or Z-A.

Step 8: Split the Column :- It is used to split the column into 2 or more columns based on the following:

* + - 1. By delimiter
      2. By number of characters
      3. By position
      4. By lower case to Upper case
      5. By Upper case to lower case
      6. Digits to Non-Digits

Step 9: Replace values :- It is used to replace the values with some another value.

Step 10: Merging Query as new :- It is used to merge or combine 2 or more tables into a single table using a primary key.

Note : It acts like Joins in SQL.

Step 11: Append Queries as New :- It is used to append the data from 2 or more tables into a single table.

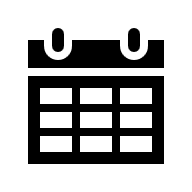
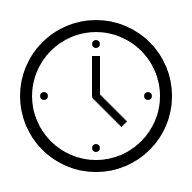
Step 12: Rename :- It is used to rename the columns.

Step 13: Conditional Column/ Custom Column :- It is used to add the new columns based on the conditions (Without using DAX functions).

Step 14: Column Profile/ Column Distribution/ Column quality :- These are used to validate the data.

Step 15: Fill (Up, Down) :- These are used to fill the empty values based on upper value or lower value.  
Step 16: Group By :- It is used to group the records based on the conditions.

Data Types:

* Number(123)
* Text(ABC)
* Date(  )
* Time(  )
* Percentage(%)
* Decimals(1.2)

Steps to Open Power Query:

Step 1: Open a Blank Workbook.  
Step 2: Go to Data Tab  
Step 3: Click Get Data  
Step 4: From File  
Step 5: From Excel  
Step 6: Browse the file from pc  
Step 7: Import

Define power-bi

what are the components of power bi

what are different types of views in power bi

what are the filters in power bi

what is the role of get data in power bi

   role of power query editor.

define data modelling, data transformation, data cleaning and data filtering

working principle of power bi

1. Data Integration
2. Extract and integrate the data from different sources. After integration the data is converted into standard format and stored in common area called as stagging area.
3. Data Processing
4. Once the data is assembled and integrated, it requires cleaning up. So, we perform few transformation and cleaning operation on the raw data to remove redundant values.
5. Data Presentation
6. Once the data is transformed and cleaned, it is visually represented on power bi desktop as reports, dashboards or score cards. These reports can be shared via mobile apps or websites to the various business users.

Star schema: It is a data modelling technique that arrange the data into a centre “Fact table” surrounded by dimensional table, reassembling a star shape.

Why POWER BI use star schema?

1. Performance
2. Simplicity
3. Useability
4. Scalability