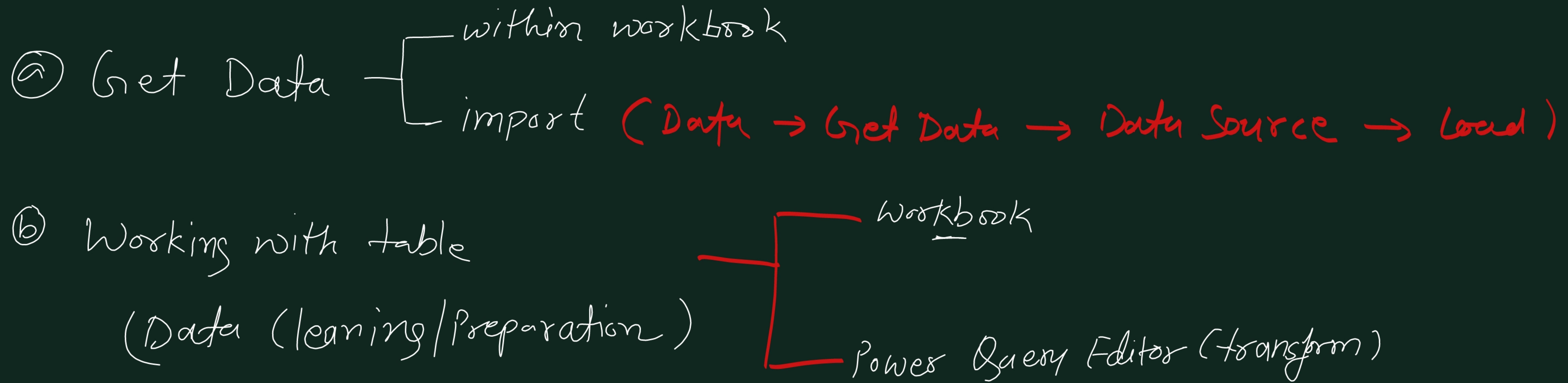


Data Science Roadmap :

- ① Excel , SQL
 - ② Business Intelligence (Power BI)
 - ③ Statistics
 - ④ Python Programming
 - ⑤ Python for DS + (EDA , ML , DL , NLP)
- ✓ ★
- Data Analyst

Excel



Workbook (Table)

- ③ Ctrl + T
- ④ Functions / Formulae
- ⑤ Conditional Formatting

- ⑥ Sorting / Filtering
- ⑦ Rename / Datatype

Function / Formulae

- ① Aggregate - { max(), min(), count
maxif(), countif() }
- ② Text - { lower(), proper()
len(), left(), concat(), trim() }
- ③ Date - date(), year(), datedif(), text()
- ④ Additional - vlookup(), index-match, if-else
-) Excel Feature
- ① Text to Column
- ② Data Validation
- ③ Paste Special
- ④ Clear

Large Data

- Ⓐ Pivot table - [Group
Aggregate]
- Ⓑ Slicer → Connect pivots
- Ⓒ Pivot chart & Formatting
- Ⓓ Dashboard

Power BI

- Ⓐ Excel Vs BI
 - Data Size
 - Data Source
 - Sharing
- Ⓑ Power BI Versions :
 - + Desktop, Service, mobile
- Ⓒ Power BI Component :
 - Report
 - Table
 - Model (Power Pivot)
 - DAX Query
 - TMDL
 - Power Query Editor

① Visualization

- Basic : Card, Column
- AI : Q&A, Smart Narrative
- Custom : Python Visual

② Data Modeling (Model View)

+ Star



① Fact

+ Snowflake

② Dimension

③ DAX Language (Top 10 Formulas)

- Measure - Aggregate Calc
- Column - Row Level Calculation
- Table

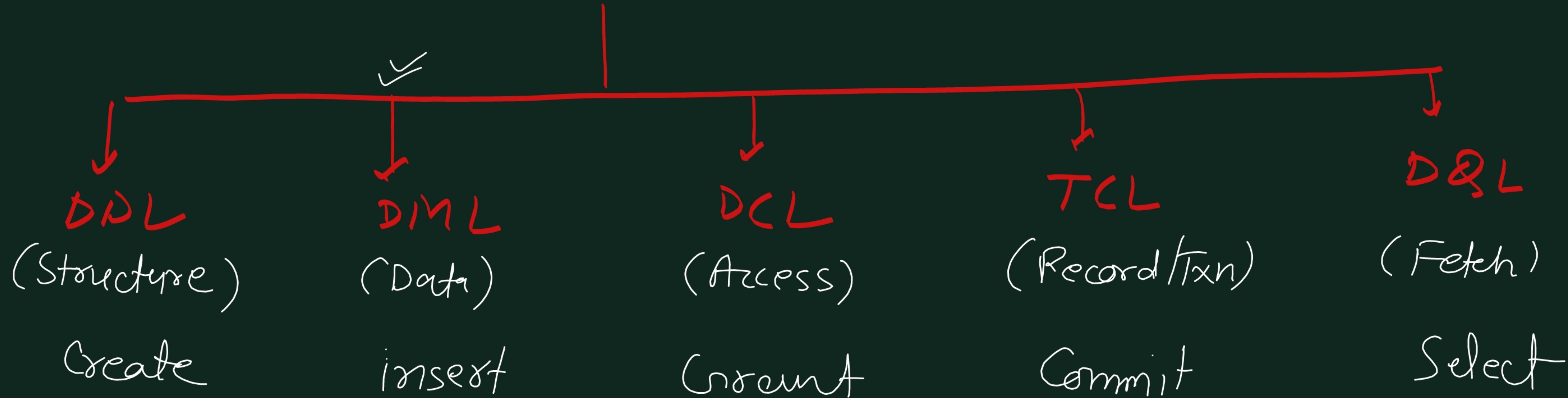
Advanced

=
+ Tooltip

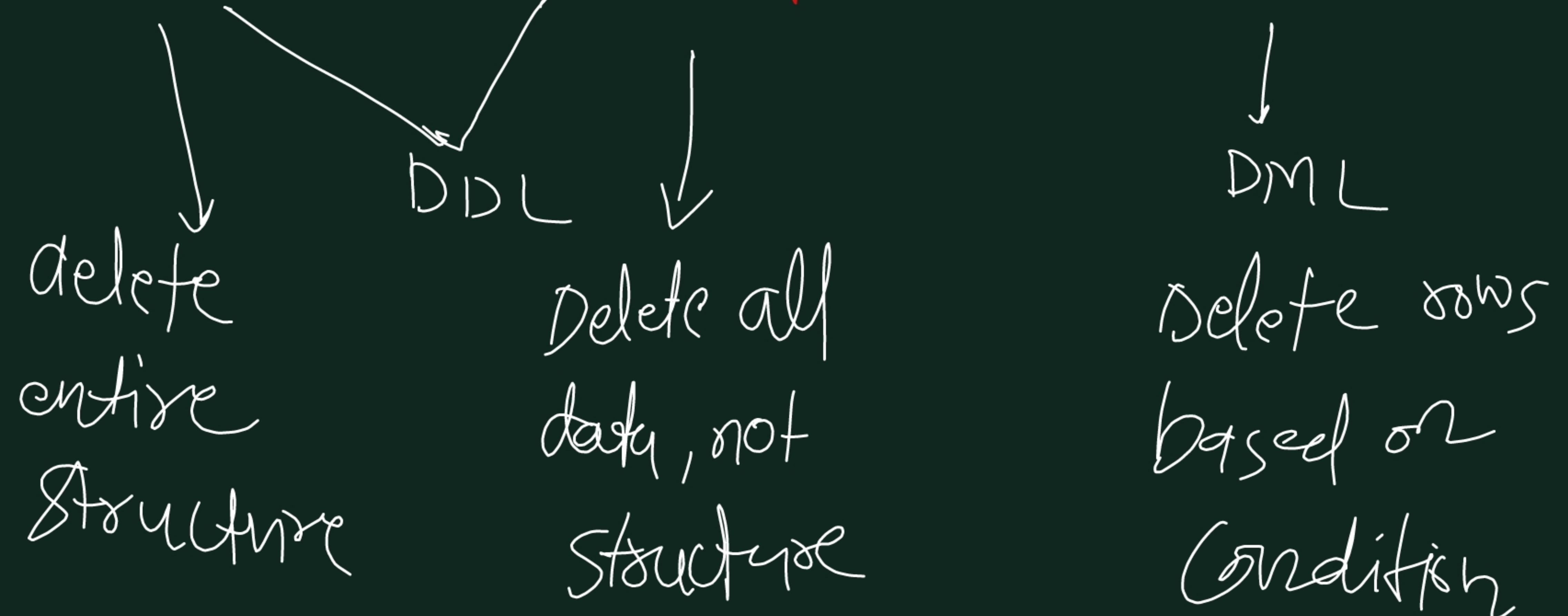
+ Page Navigation

+ Row Level Security

SQL (Structured Query Language)



Drop Vs Truncate Vs Delete



SQL Constraints (Rules)

(a) Unique key

(b) Primary key

(c) Foreign key

(d) Not null

(e) Default

(f) Check

* Datatypes : int, double, varchar(), float -

Statements

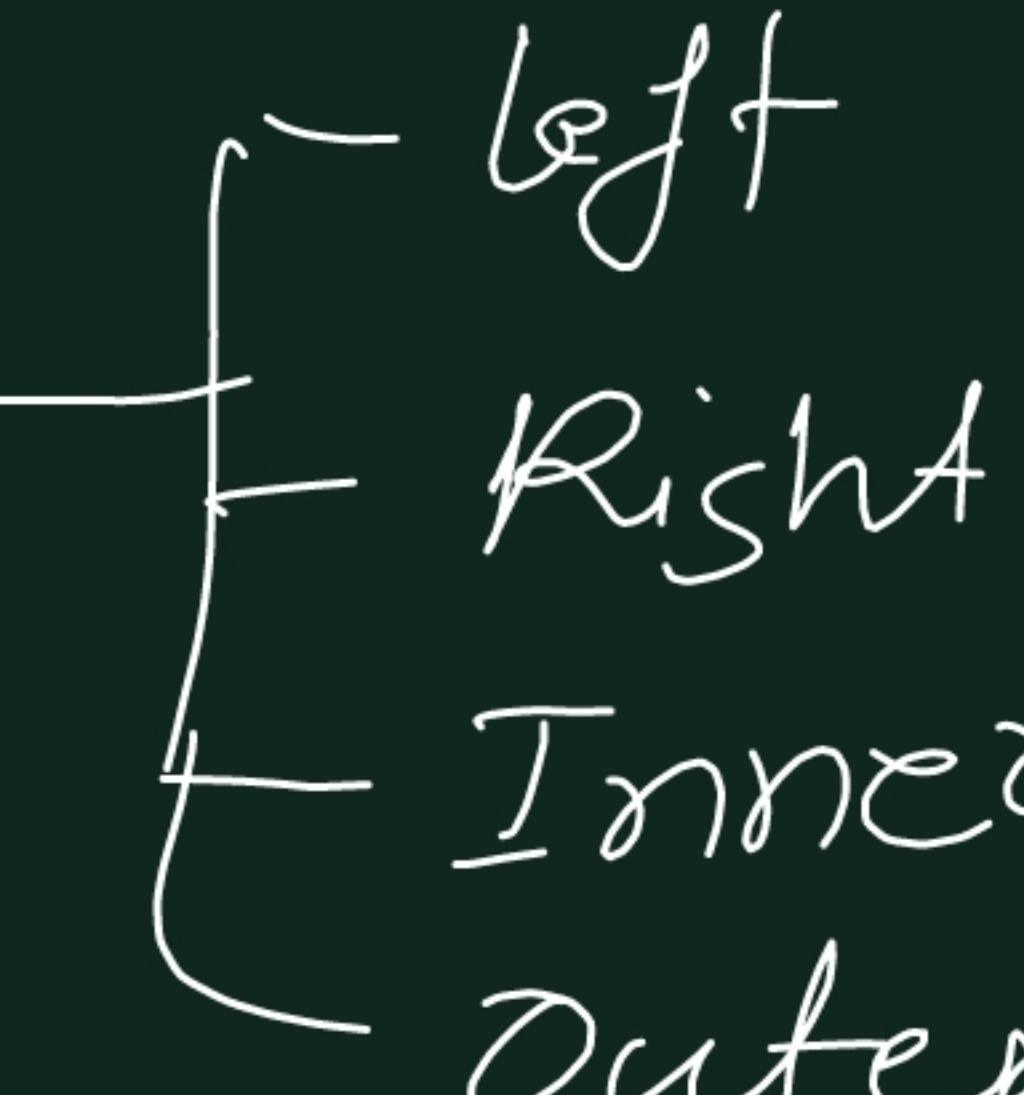
- ① Select → Selects Fields , AS , Distinct
- ② From → Selects table
- ③ Where → Filter table data , \geq , like , between
IN
- ④ Group by → Group & Aggregate
- ⑤ Having → Filter grouped data
- ⑥ Order By → Sort the data
- ⑦ Limit → Fix rows count

Working with multiple tables

① Nested Query

Select -  Fetching

- -
Where cond (Select -
- -) ;
 Condition

② Joins 
left
Right
Inner
Outer
cross , self

$\Rightarrow \begin{cases} \text{Select } T1.c, T2.c \\ \text{From } T1 \text{ Join } T2 \\ \text{On } T1.c = T2.c \end{cases}$

Advanced

① Windows Functions / Analytical Function

↳ Row level Calculation

row_number(), rank(), dense_rank()

② CTEs (Common Table Expression)

↳ Saves output table temporarily.

} rank() over (order by col asc)

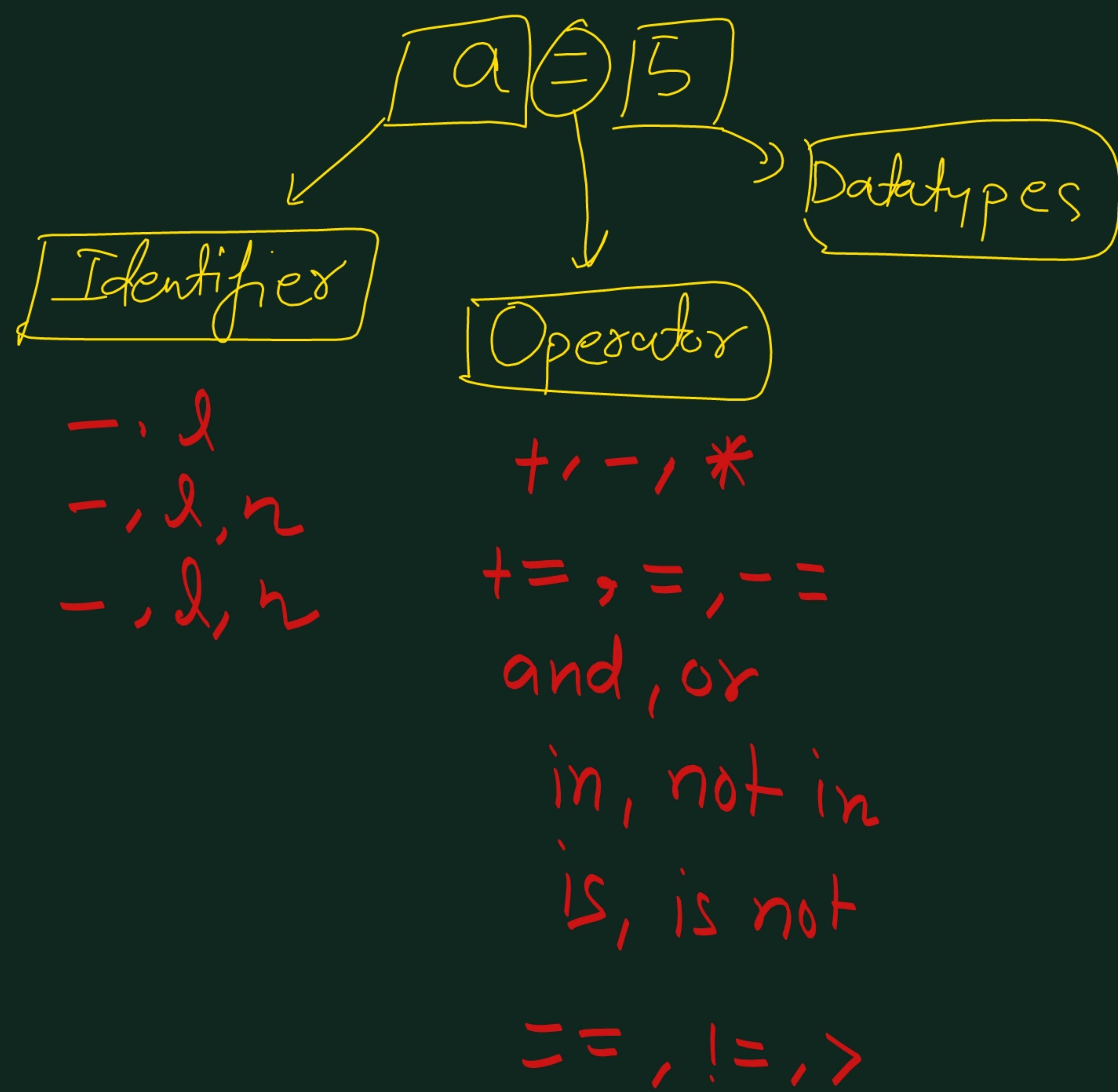


with CTEname as

(
Select
From
)

Select * From CTEname
Where cond. ↗

Core Python



Basic : int, float, string, bool

Collection : list, tuple, set, dict

✳️ **Basic :** sequential, conditional

Collection : loop → sequential/conditional

Functions : def ↗ anywhere , lambda { operation
Binary output }

OOP

- ① class
- ② object
- ③ attribute
- ④ method

OOP properties:

- ① Inheritance
- ② Polymorphism
- ③ Encapsulation
- ④ Abstraction

EDA (Data Analyst)

- ① Numpy → Array - [1D
2D
3D] + indexing + Op
- ② Pandas - [Series - 1D
Datframe - 2D] + indexing + Function + Group by + Pivot table
- ③ Matplotlib - Basic → plotname() Cross tab hist()
- ④ Seaborn → Advanced plotnameplot() histplot()