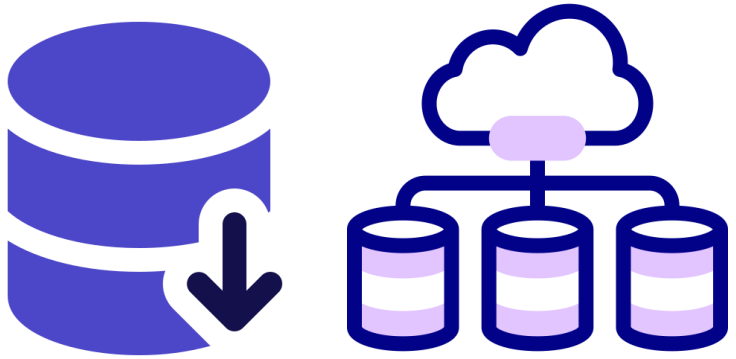
RDBMS Presentation



Topic :

“Concept of Indexing”

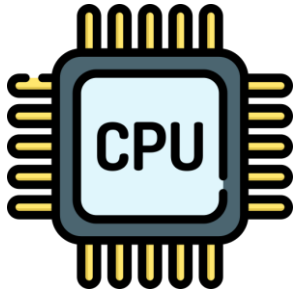
Rajneesh Rai

INDEX IN DATABASE

WHAT IS AN INDEX IN DATABASE?

- A database index is a DATA STRUCTURE that improves the SPEED OF DATA RETRIEVAL operations on a database table.
- It is used to LOCATE AND ACCESS the data in a database table QUICKLY, this process is called as INDEXING.

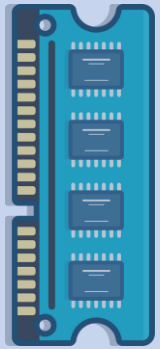
Why Indexing is used ?



Processor

Speed : Very High (MIPS)

Capacity : Low (Hz)



Volatile & Temporary

Speed : High (TIPS)

Capacity : Low (GB)

PRIMARY MEMORY

Non_Volatile & Permanent

Speed : Low

Capacity : High (TB)

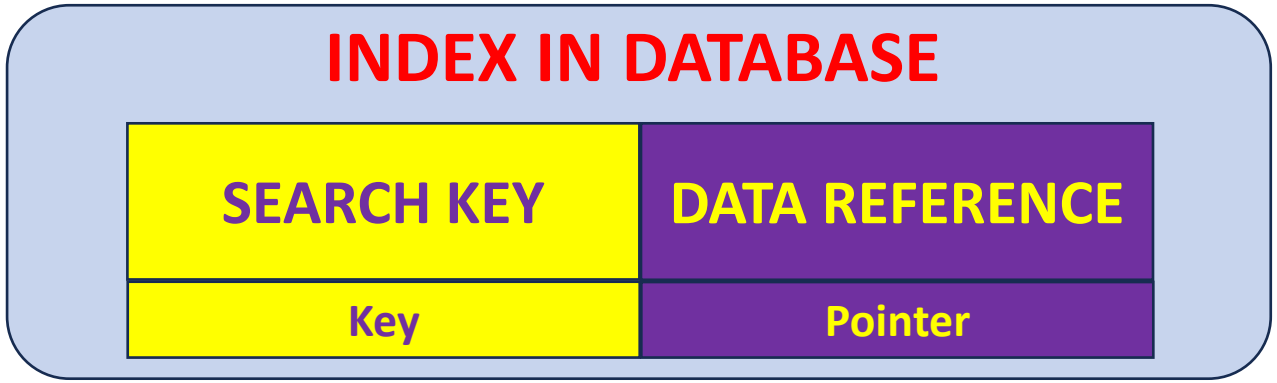
SECONDARY MEMORY

Index in Books

	SEARCH KEY	DATA POINTER
S.NO.	TOPIC	PAGE NO.
1	Tables in Database	1
2	Keys in Database	6
3	Dependencies in Database	11
4	Normalization in Database	15
5	Indexing in Database	28
6	Data Dictionary in Database	35
7	E-R Model	42

Index in DATABASE

- A database index is a DATA STRUCTURE that improves the SPEED OF DATA RETRIEVAL operations on a database table.
- It is used to LOCATE AND ACCESS the data in a database table QUICKLY, this process is called as INDEXING.
- Indexing is used to OPTIMIZE the performance of a database by MINIMIZING THE NUMBER OF DISK ACCESSES REQUIRED when a QUERY IS PROCESSED.
- Indexes are created using a FEW DATABASE COLUMNS.
- An INDEX is a small table having ONLY TWO COLUMNS.
- The FIRST COLUMN comprises a copy of the PRIMARY or CANDIDATE KEY of a table.
- Its SECOND COLUMN contains a SET OF POINTERS for holding the ADDRESS of the DISK BLOCK where that SPECIFIC KEY value stored.



Index in DATABASE

Indexing is defined based on indexing attributes, they are:

1) **ACCESS TYPES** : This refers to the type of access such as value based search, range access, etc.

2) **ACCESS TIME** : It refers to the time needed to find particular data element or set of elements.

3) **INSERTION TIME** : It refers to the time taken to find the appropriate space and insert a new data.

4) **DELETION TIME** : Time taken to find an item and delete it as well as update the index structure.

5) **SPACE OVERHEAD** : It refers to the additional space required by the index.

Index in DATABASE

PRIMARY INDEX

Defined on an Ordered data file, which is based on a PRIMARY KEY field of the relation.

- DENSE INDEX → Stores ALL SEARCH-KEY values, needs MORE SPACE, make SEARCHING FASTER.
- SPARSE INDEX → Stores only SOME SEARCH-KEY values.

CLUSTER INDEX

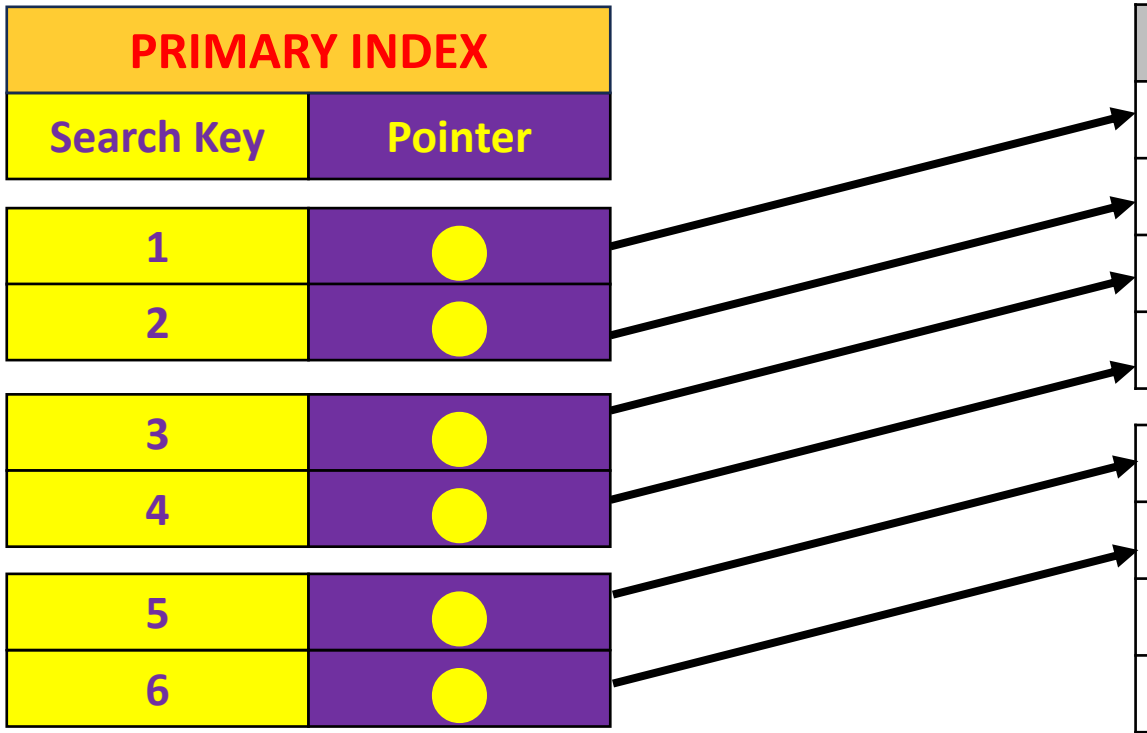
Defined on an Ordered data file, which is based on a NON-KEY field.

SECONDARY INDEX

Defined on an un-ordered data file, generated from a CANDIDATE KEY of the relation.

PRIMARY INDEX – DENSE INDEX

- Data is SORTED based on Primary Key.
- Dense Index Stores ALL SEARCH-KEY VALUES.
- Needs MORE SPACE and MAKES SEARCHING FASTER.
- Index records contain SEARCH KEY VALUE and a POINTER to the ACTUAL RECORD on the DISK.



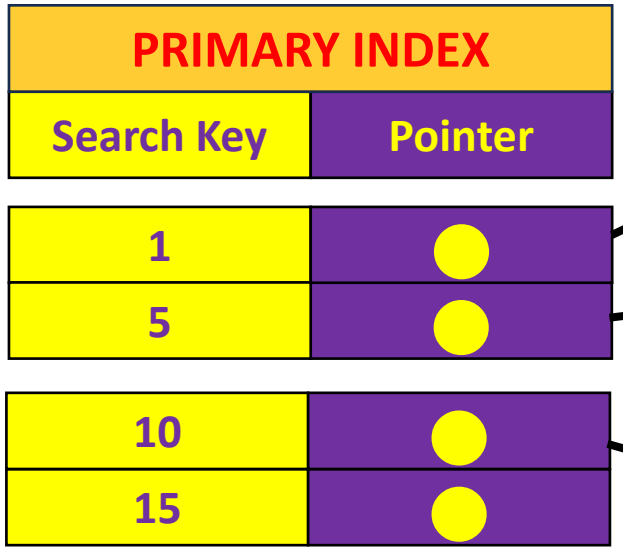
Search key = Data inside each block

HARD DISK or HARD DRIVE

E_ID	E_NAME	ADHR_No
1	RAJNEESH	568
2	SHOBHIT	896
3	D.PRAVEEN	456
4	ANJALI	125
5	SONIYA	157
6	ANU	865
7	SHIVANI	325
8	MANALI	421
9	SAFIYA	874

PRIMARY INDEX – SPARSE INDEX

- Data is SORTED based on Primary Key.
- Stores ONLY SOME SEARCH-KEY values.
- Needs LESS SPACE, LESS MAINTENANCE overhead.
- In Sparse indexing technique, a RANGE OF INDEX COLUMNS stores the same DATA BLOCK ADDRESS, and when data needs to be RETRIEVED, the block ADDRESS will be FETCHED.



HARD DISK or HARD DRIVE

E_ID	E_NAME	ADHR_No	
1	RAJNEESH	568	BLOCK - 1
2	SHOBHIT	896	
3	D.PRAVEEN	456	
4	ANJALI	125	
5	SONIYA	157	BLOCK - 2
6	ANU	865	
7	SHIVANI	325	
8	MANALI	421	
9	SAFIYA	874	
10	SANJU	652	BLOCK - 3
11	BHAGIRATH	124	
12	SHIVANSHI	128	
13	SHEETAL	475	
14	POORNIMA	280	

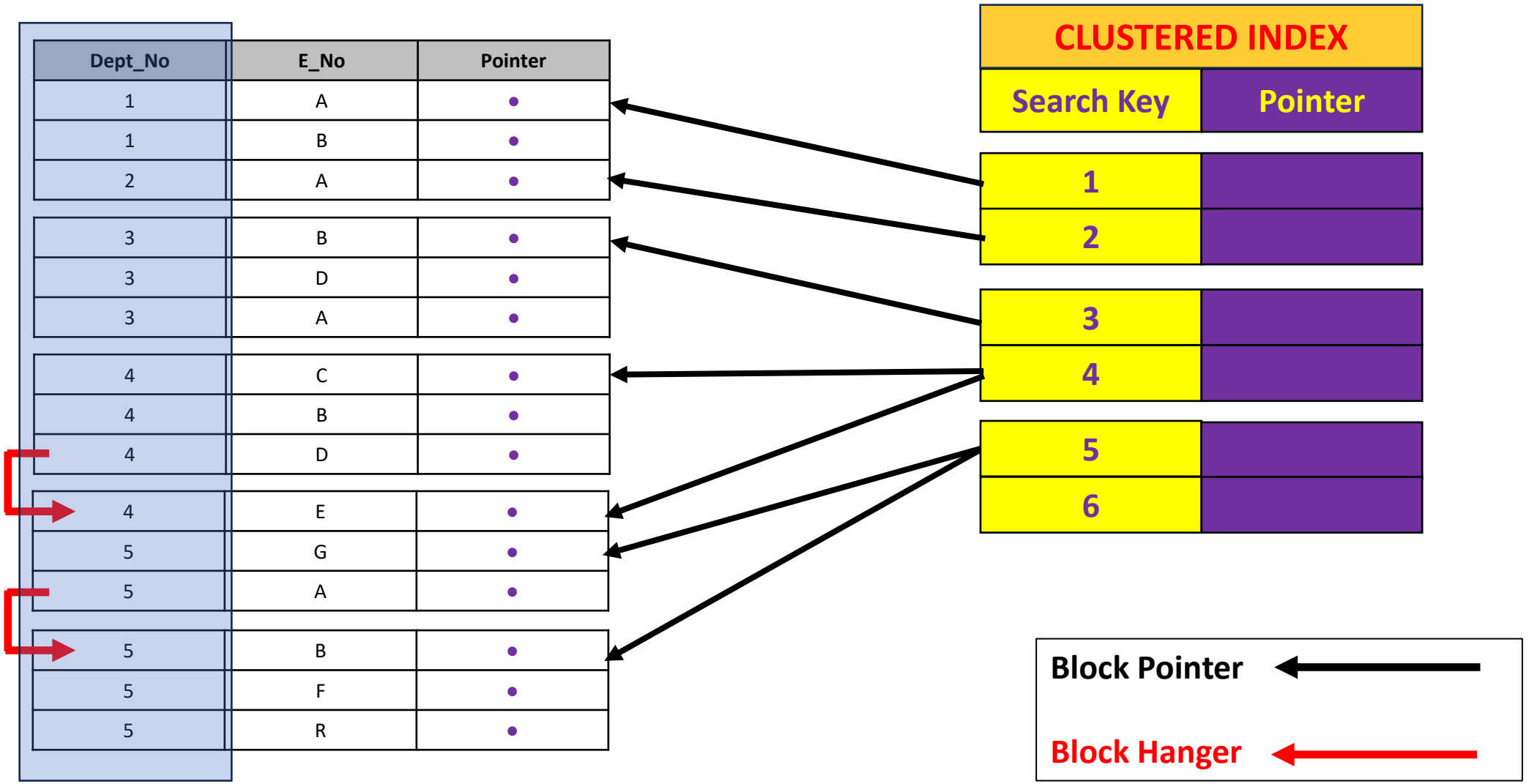
INDEX – CLUSTERED INDEX

INDEX CLUSTERED INDEXWhen to use Clustered Index ?

1) **Ordered Data.**
2) **Data already have Primary Index**
3) **Non-Key based search required.**

- Mostly clustered Index will have COMBINATION OF NON KEY ATTRIBUTE ALONG WITH CANDIDATE KEY of the tables.
- Clustered Index is a SPARSE INDEX TYPE, where only ONE SEARCH KEY VALUE per block.

INDEX – CLUSTERED INDEX



INDEX – SECONDARY INDEX

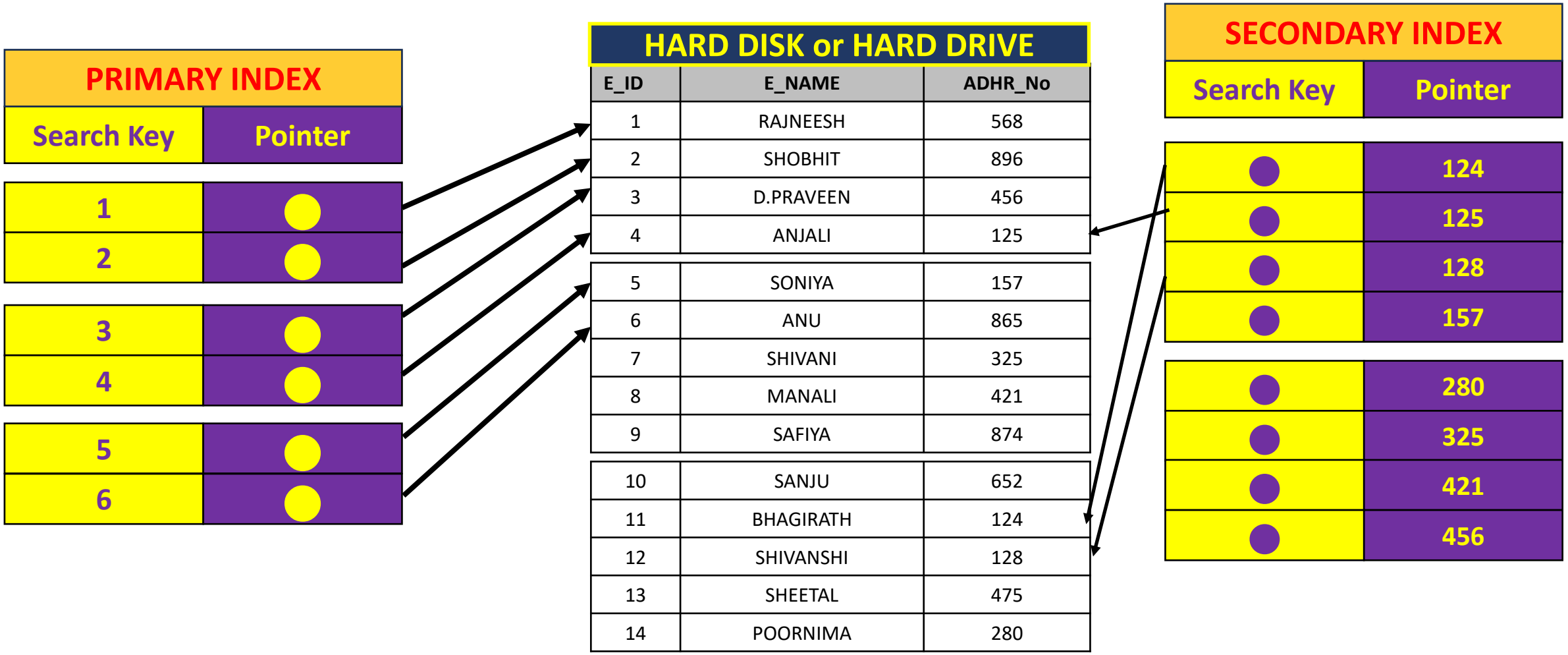
When to use Secondary Index?

1) Unordered Data
2) Non-Key or Key based search required.
3) Data already have Primary Index

- Data searching requires MORE TIME as compared to the clustered index
- Secondary Index is of DENSE INDEX type, where ALL SEARCH KEY VALUE pointing to specific block.

INDEX – SECONDARY INDEX

- When you need to search data OTHER THAN KEY OF PRIMARY INDEX, such as Aadhar Card No, which is also a UNIQUE KEY IN DATABASE.



HARD DISK or HARD DRIVE

E_ID	E_NAME	ADHR_No
1	RAJNEESH	568
2	SHOBHIT	896
3	D.PRAVEEN	456
4	ANJALI	125
5	SONIYA	157
6	ANU	865
7	SHIVANI	325
8	MANALI	421
9	SAFIYA	874
5	SANJU	652
6	BHAGIRATH	124
7	SHIVANSHI	128
8	SHEETAL	475
9	POORNIMA	280

QUICK RECAP

WHAT IS AN INDEX IN DATABASE?

- A database index is a DATA STRUCTURE that improves the SPEED OF DATA RETRIEVAL operations on a database table.
- It is used to LOCATE AND ACCESS the data in a database table QUICKLY, this process is called as INDEXING.

Index in DATABASE

PRIMARY INDEX	CLUSTER INDEX	SECONDARY INDEX
Defined on an Ordered data file, which is based on a PRIMARY KEY field of the relation.	Defined on an Ordered data file, which is based on a NON-KEY field.	Defined on an un-ordered data file, generated from a CANDIDATE KEY of the relation.
TYPES : **DENSE INDEX** **SPARSE INDEX** **TIME : LESS**	**Clustered index type is of Sparse type index** **TIME : More than Primary Index**	**Secondary index type is of Dense type index** **TIME : More than Clustered Index**