papergrid
Date: 3 / 8 / 2024

Day 9

SQL

→ Fundamentals of Database
- →RDBMS
  - MySQL
- →NOSQL
  - MongoDB
  - Cassandra
  - GraphDB
  - HDFS

→ CRUD

→ Datatypes & Constraints
- Primary key,
- Not Null,
- Unique key,
- Foreign key

→ Schema of table
- DDL
  - Create
  - Alter
- DML
  - Update
  - Delete
  - Insert

→ WHERE (filteration)

papergrid

Date: / /

→ Aggregation & Grouping

- Count
- Sum
- max
- min
- avg

- Group by
- having

→ Order of Execution

→ Joins
- Inner
- Left
- Right
- Self
- Cross

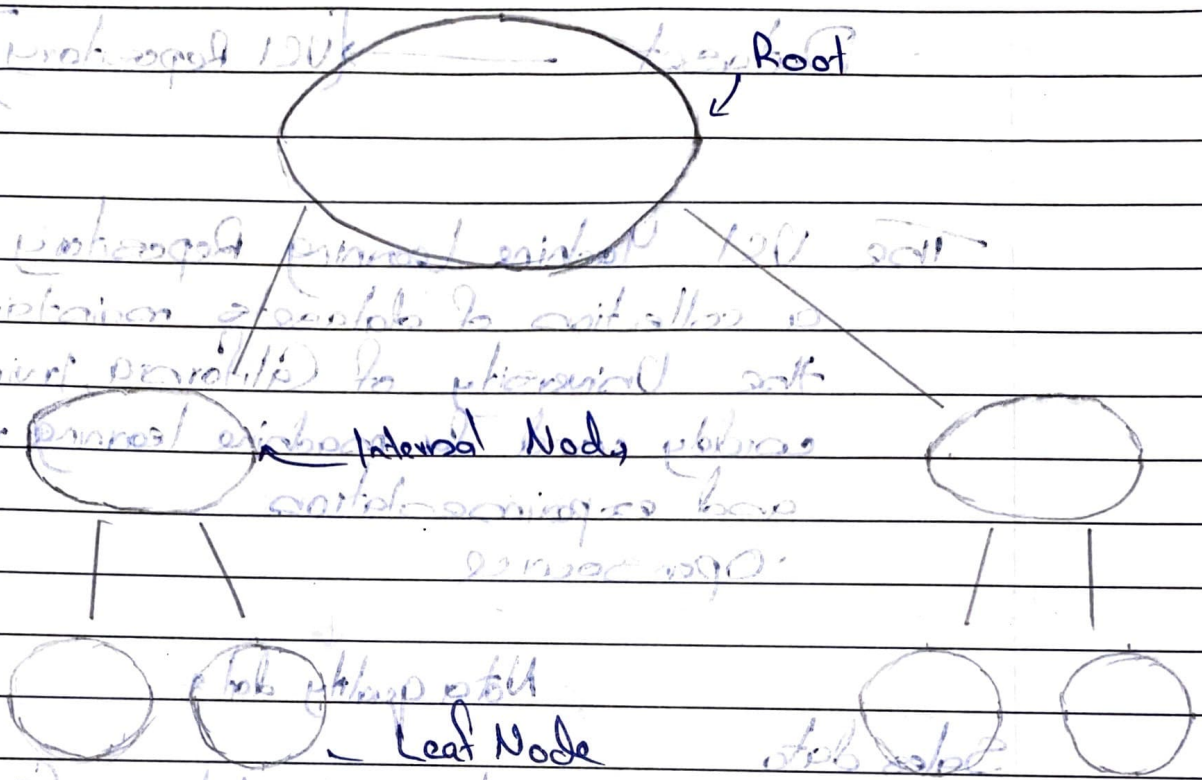
→ Window Function
- OVER()
- PARTITION BY
- RANK vs DENSE RANK vs ROW

→ Case statements vs Case expression

→ Common Table Expression
- WITH

→ Interview Questions.

n-ary tree



Tree Data structure (Binary)

B-Tree &
B+ Tree

- File indexing

Day 10

Dataset → [UCI Repository]

The UCI Machine Learning Repository is a collection of datasets maintained by the University of California, Irvine (UCI) widely used for machine learning research and experimentation
· Open source

Sales data → Data quality data

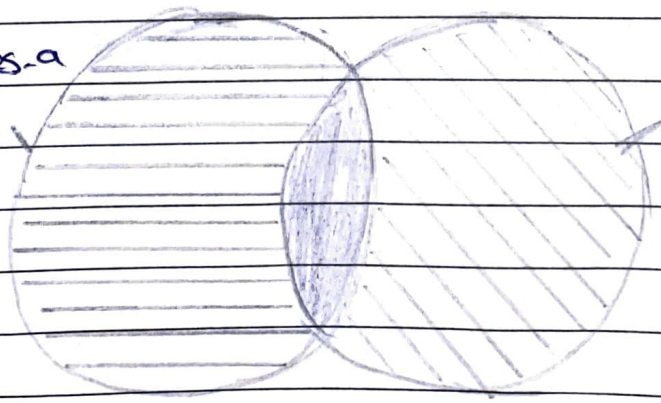
Exploratory Data Analysis (EDA)

Quality Data

SQL Queries

Data Analysis

Statistics knowledge



Employees_a:

id	name
1	Alice
2	Bob

Employees_b:

id	name
2	Bob
3	Carol

```
SELECT id, name FROM employees_a
UNION
SELECT id, name employees_b
```

id	name
1	Alice
2	Bob
3	Carol

Union All

id	name
1	Alice
2	Bob
2	Bob
3	Carol

UNION ALL

- faster Operation
- No removal of duplicates