SQL 1 Data Analyst Fellowship

Saksham Arora

Software Engineer 2 @ Intuit

Ex - Software Engineer @ Microsoft

[Saksham Arora - Intuit | LinkedIn]

1. Introduction to DML and SQL Queries

- **Topics:**
  - SELECT, FROM, WHERE Clauses
  - Basic Anatomy of a Query

2. Using Operators and Filtering Data

- **Topics:**
  - Relational/Comparison Operators
  - Logical Operators
  - LIKE, BETWEEN, DISTINCT
- Write queries using different operators to filter data.

3. Advanced SQL Queries

- **Topics:**
  - Aliasing, Sorting
  - Grouping Aggregation
  - HAVING Clause, LIMIT
- Write queries

4. Exploring Data with SQL

- **Topics:**
  - Exploratory Data Analysis using SQL
  - Functions (MIN, MAX, AVG, SUM, ROUND)

What is SQL?

Structured Query Language

SQL is a programming language used to interact with relational databases.

SQL used CRUD operations to communicate with DB.

1. CREATE - execute INSERT statements to insert new rows into the tables.
2. READ - Read data already in the tables.
3. UPDATE - Modify already inserted data in the tables.
4. DELETE - Delete specific data point/tuple/row or multiple rows.

SQL is not DB, it is a query language.

NextLeap SQL Compiler - SQL Online Editor

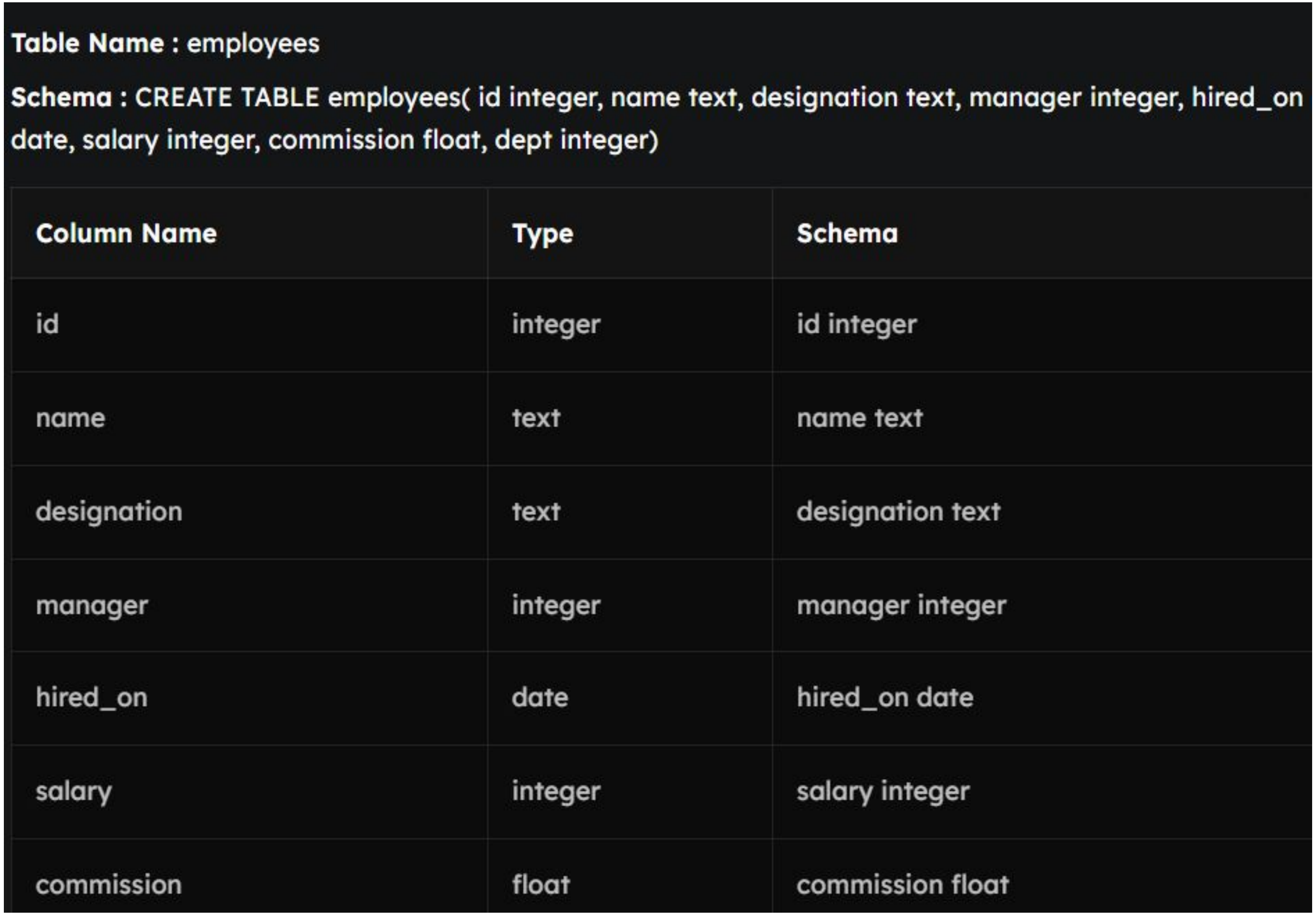
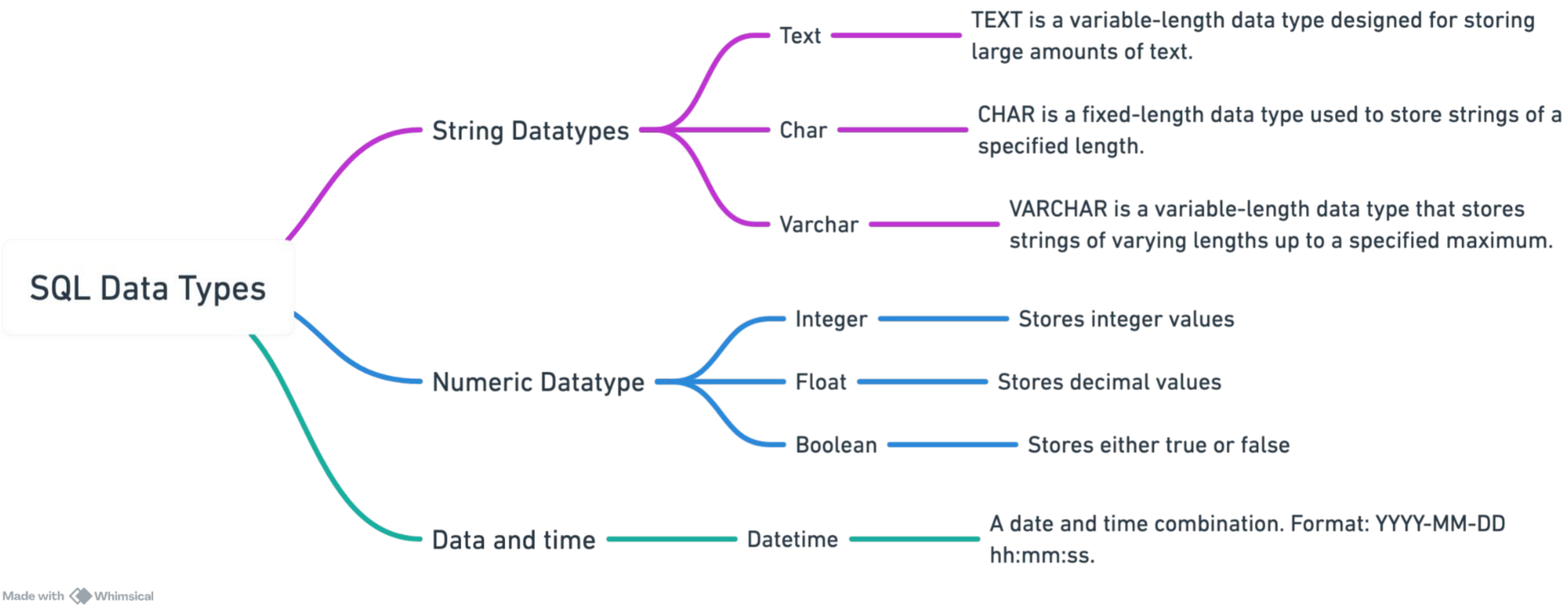


Table Name : employees

Schema : CREATE TABLE employees(id integer, name text, designation text, manager integer, hired_on date, salary integer, commission float, dept integer)

Column Name	Type	Schema
id	integer	id integer
name	text	name text
designation	text	designation text
manager	integer	manager integer
hired_on	date	hired_on date
salary	integer	salary integer
commission	float	commission float

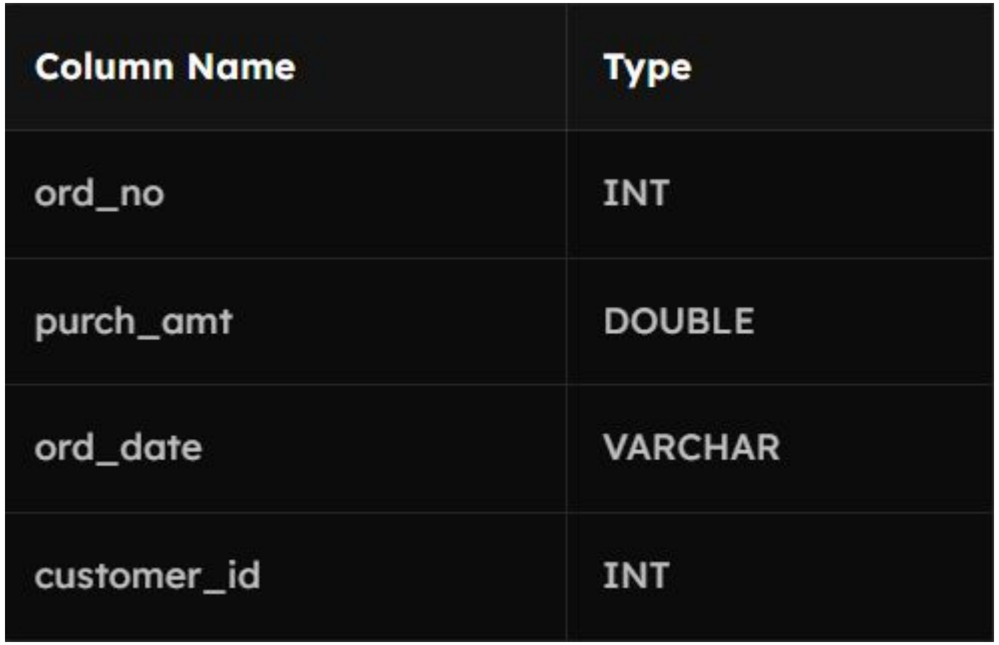
Datatypes in SQL



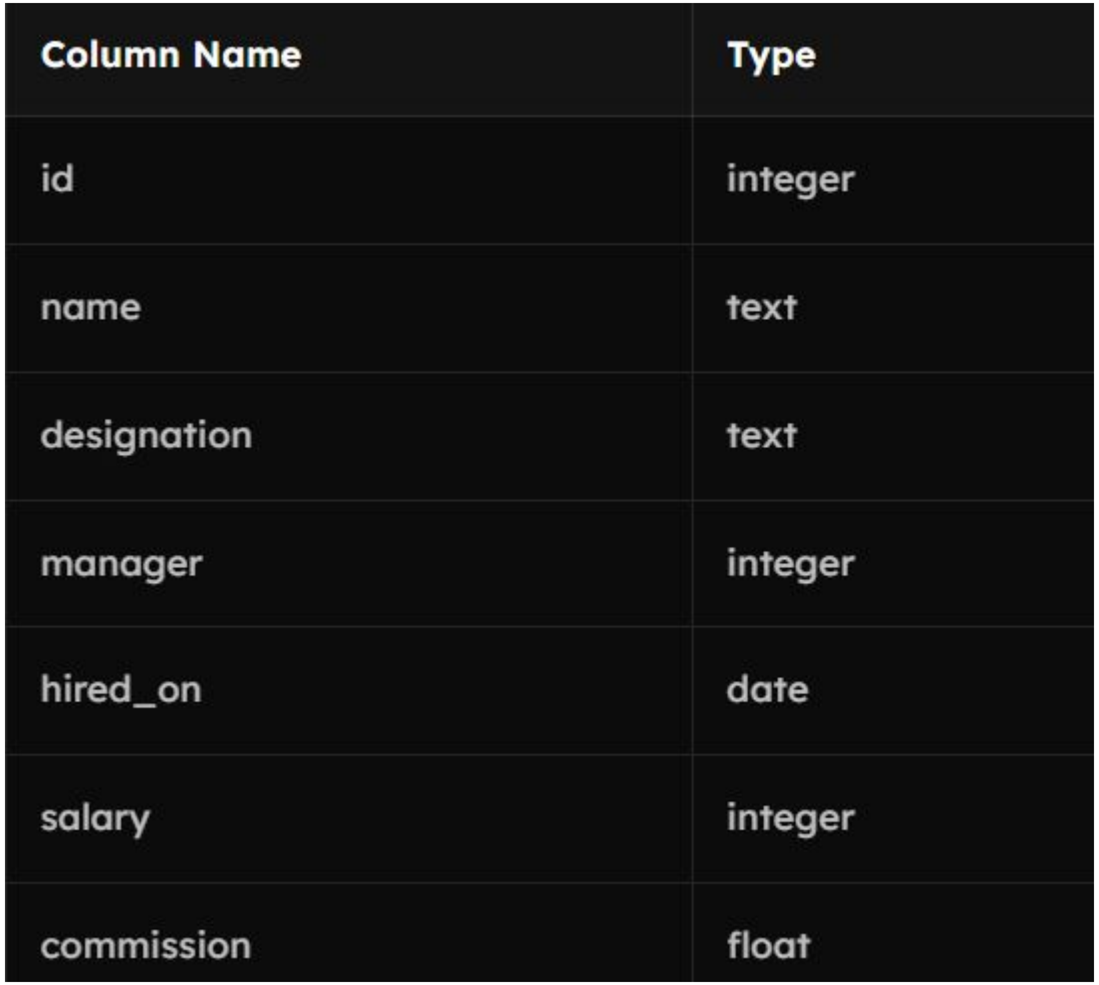
SQL Data Types

- String Datatypes
  - Text — TEXT is a variable-length data type designed for storing large amounts of text.
  - Char — CHAR is a fixed-length data type used to store strings of a specified length.
  - Varchar — VARCHAR is a variable-length data type that stores strings of varying lengths up to a specified maximum.
- Numeric Datatype
  - Integer — Stores integer values
  - Float — Stores decimal values
  - Boolean — Stores either true or false
- Data and time
  - Datetime — A date and time combination. Format: YYYY-MM-DD hh:mm:ss.

Examples of Datatypes in SQL



Column Name	Type
ord_no	INT
purch_amt	DOUBLE
ord_date	VARCHAR
customer_id	INT



Column Name	Type
id	integer
name	text
designation	text
manager	integer
hired_on	date
salary	integer
commission	float

Writing Your First SQL Query

Creating Your First Database Query

SELECT

1. Syntax: SELECT columnName FROM tableName ;
2. Q. Can we use the SELECT keyword without using FROM clause?

e.g., SELECT 55 + 11;

WHERE

1. Reduce rows based on given conditions.
2. E.g., WHERE id > 10;

BETWEEN

1. SELECT * FROM employees WHERE id between 0 AND 5;
2. In the above e.g. 0 and 5 are inclusive.

IN

1. Reduces OR conditions;

AND/OR/NOT

1. AND: WHERE cond1 AND cond2
2. OR: WHERE cond1 OR cond2 3. NOT: WHERE col_name NOT IN (1,2,3,4);

IS NULL

1. e.g., SELECT * FROM customer WHERE prime_status is NULL;

Pattern Searching / Wildcard (‘%’, ‘_’)

1. ‘%’, any number of characters from 0 to n. Similar to ‘*’ asterisk in regex.
2. ‘_’, only one character.
3. SELECT * FROM customer WHERE name LIKE ‘%p_’;

nextleap

💻

Hands-On Practice

Applying What You've Learned

Practice Questions -

NextLeap | Food Delivery Drivers - Earnings Filter

NextLeap | Salaries for Associate Professors

NextLeap | High Protein and Non-Perishable Items

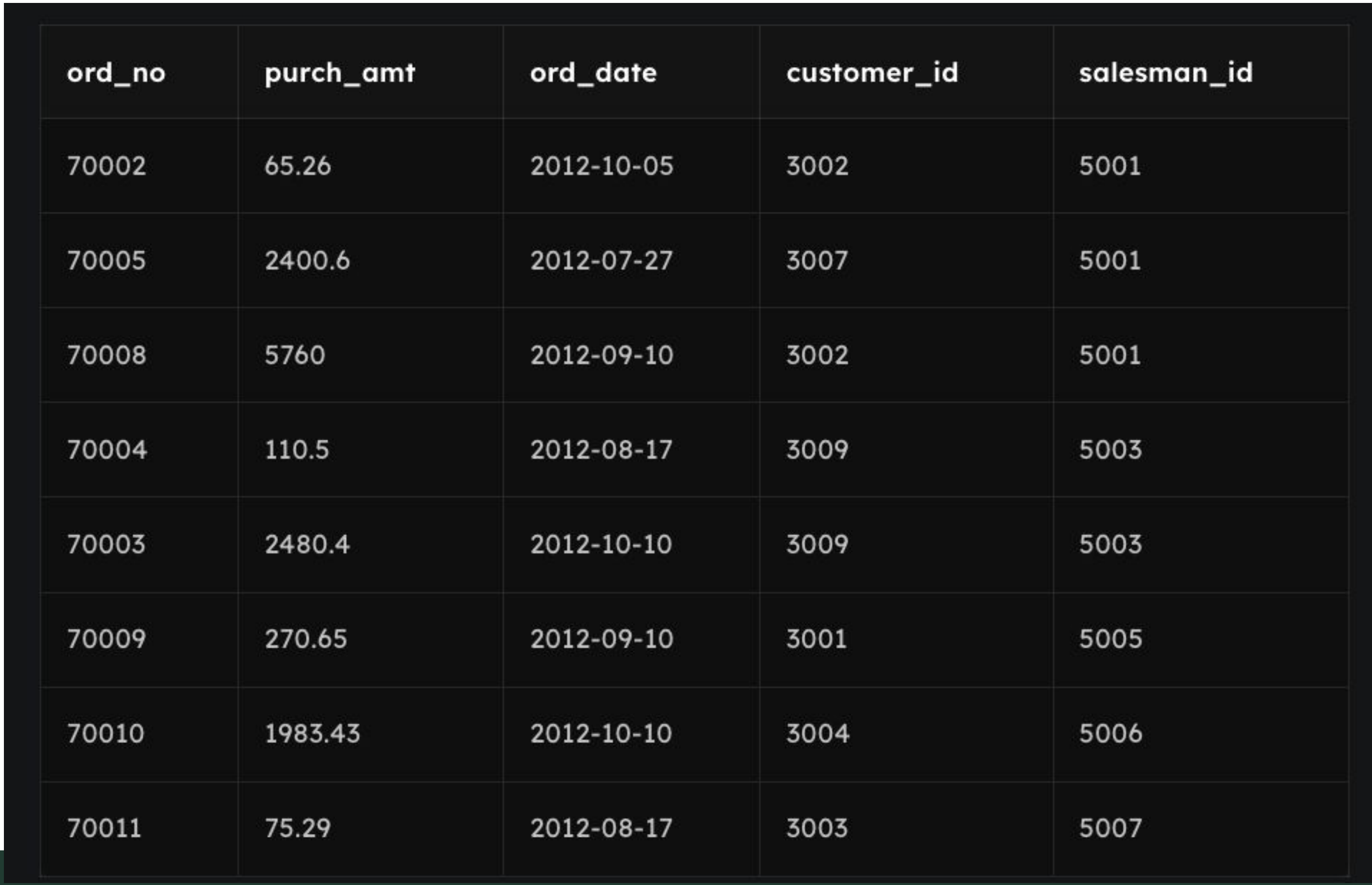
NextLeap | Unfinished Assembly Parts

Missing Phone Numbers

ORDER BY

ORDER BY

1. Sorting the data retrieved using WHERE clause.
2. ORDER BY DESC;
3. DESC = Descending and ASC = Ascending
4. e.g., SELECT * from Orders
   ORDER BY salesman_id



ord_no	purch_amt	ord_date	customer_id	salesman_id
70002	65.26	2012-10-05	3002	5001
70005	2400.6	2012-07-27	3007	5001
70008	5760	2012-09-10	3002	5001
70004	110.5	2012-08-17	3009	5003
70003	2480.4	2012-10-10	3009	5003
70009	270.65	2012-09-10	3001	5005
70010	1983.43	2012-10-10	3004	5006
70011	75.29	2012-08-17	3003	5007

LIMIT

- The LIMIT clause is used to set an upper limit on the number of tuples returned by SQL.
- Example if we write LIMIT 2, we will see only Top 2 results.

DISTINCT

- Find distinct values in the table.
- SELECT DISTINCT col_name FROM table_name;

Group By

1. GROUP BY Clause is used to collect data from multiple records and group the result by one or more columns.
2. Groups into categories based on the column given.
3. SELECT c1, c2, c3 FROM sample_table WHERE cond GROUP BY c1, c2, c3.
4. Used with aggregation functions to perform various actions.
    1. COUNT() - It gives us count of data in each group
    2. SUM() - It gives us sum of data in each group
    3. AVG() - It gives us average of data in each group
    4. MIN() - It gives smallest number in each group
    5. MAX() - It gives the largest number in each group

5. Let’s discuss examples with the tables given on NextLeap compiler - SQL Online Editor

Group By HAVING

1. Out of the categories made by GROUP BY, we would like to know only a particular thing (cond).
2. Similar to WHERE.
3. WHERE vs HAVING
    - Both have the same function of filtering the row based on certain conditions.
    - WHERE clause is used to filter the rows from the table based on specified condition
    - HAVING clause is used to filter the rows from the groups based on the specified condition.
    - HAVING is used after GROUP BY while WHERE is used before GROUP BY clause.
    - If you are using HAVING, GROUP BY is necessary.

4. Let's discuss examples with the tables given on NextLeap compiler - https://nextleap.app/online-compiler/sql-programming

Problem statements -

1. Unique Subjects by Instructor

2. Duplicate Phone Numbers

3. Popular Courses

Homework Problem

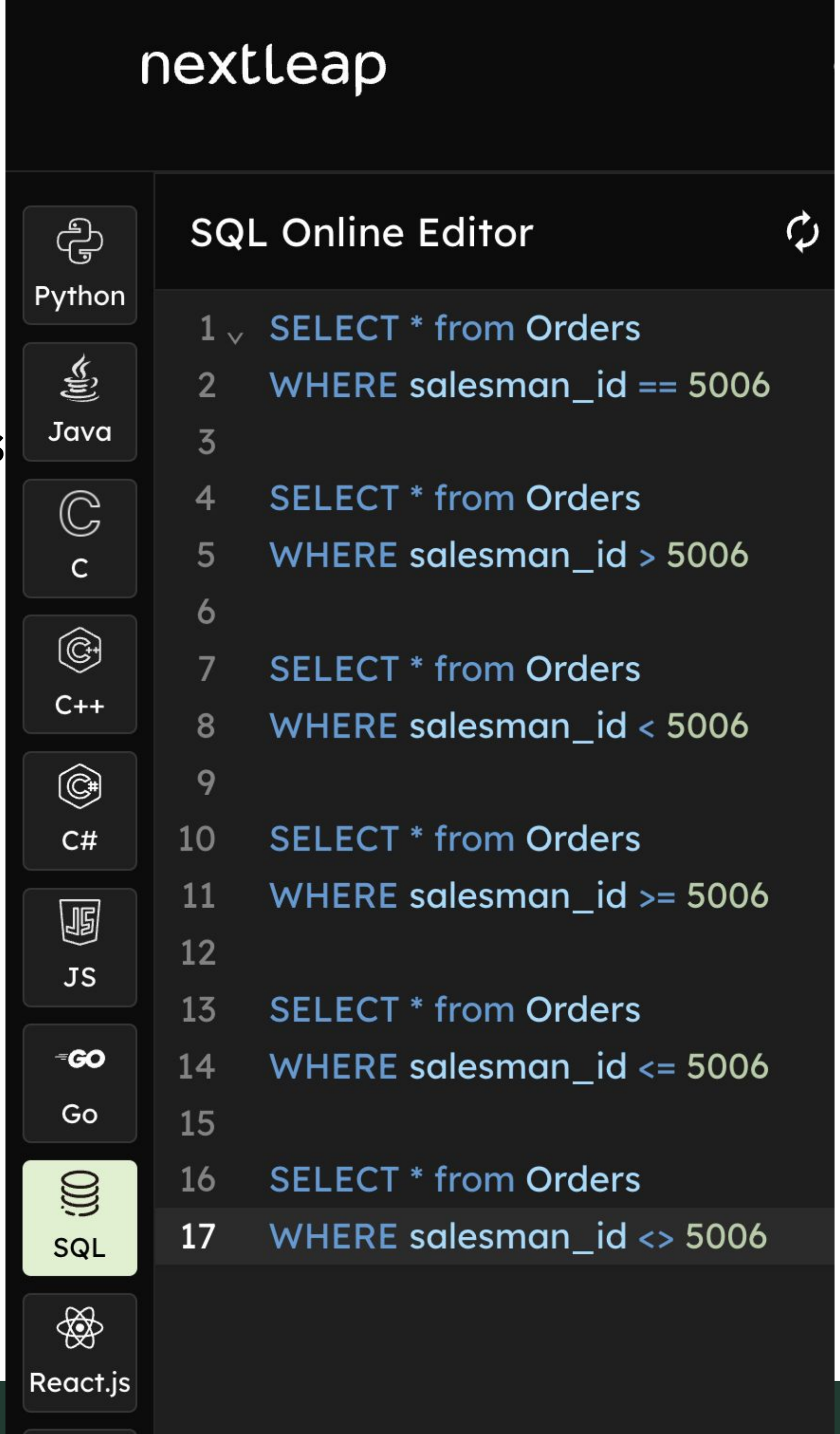
4. NextLeap | Count Followers for Users

Operators in SQL

Comparison Operator

1. == (Equal to Equal to), = (this works
2. > (Greater Than)
3. < (Less Than)
4. >= (Greater Than)
5. <= (Less Than)
6. <> (Not Equal to)

Let’s discuss the example to see how these operators work with datetime.



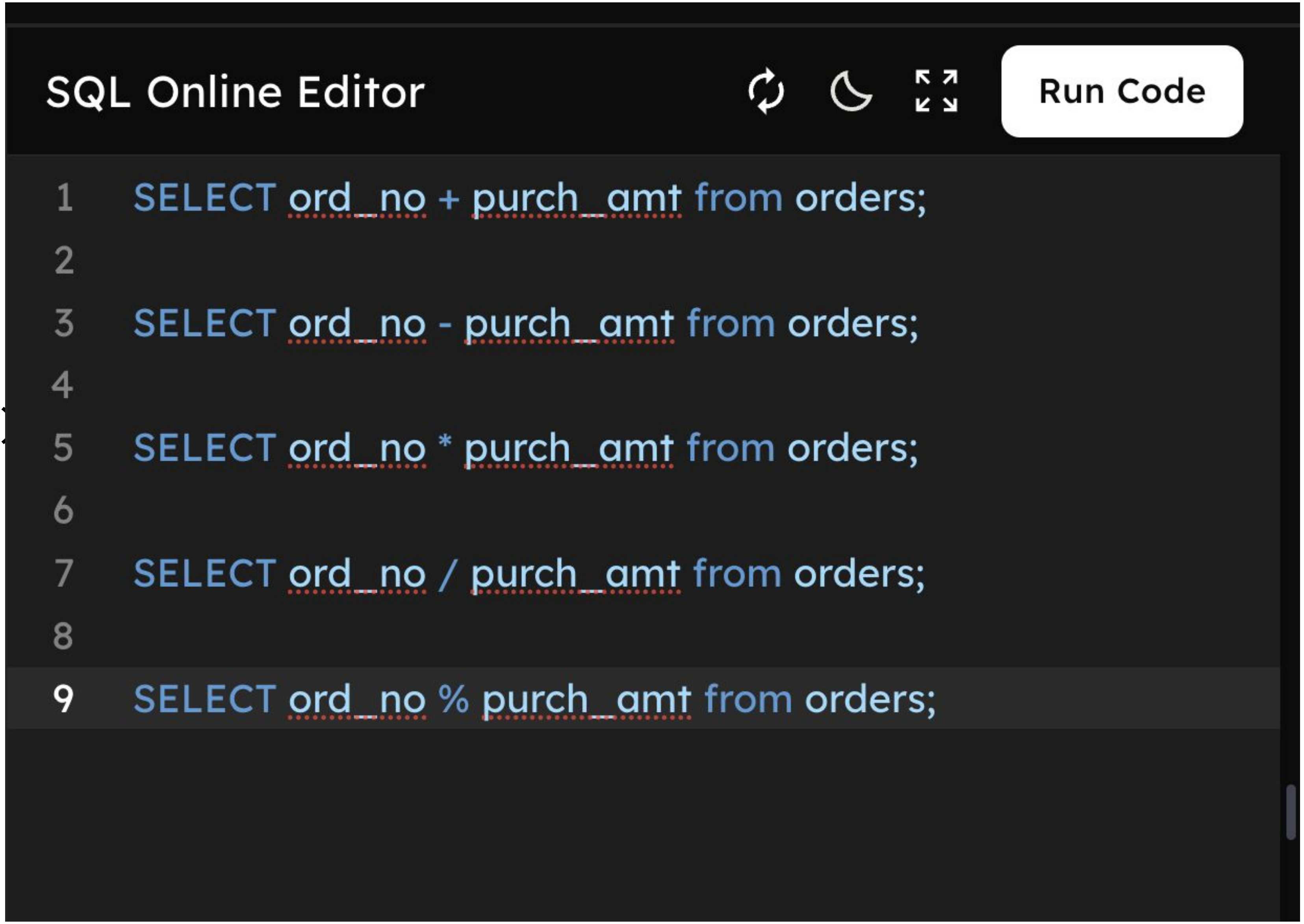
SQL Online Editor

```
SELECT * from Orders
WHERE salesman_id == 5006

SELECT * from Orders
WHERE salesman_id > 5006

SELECT * from Orders
WHERE salesman_id < 5006

SELECT * from Orders
WHERE salesman_id >= 5006

SELECT * from Orders
WHERE salesman_id <= 5006

SELECT * from Orders
WHERE salesman_id <> 5006
```

Operators in SQL

Arithmetic Operator

1. + (Add)
2. - (Subtract)
3. * (Multiply)
4. / (Divide)
5. % (Modulo) (Remainder)

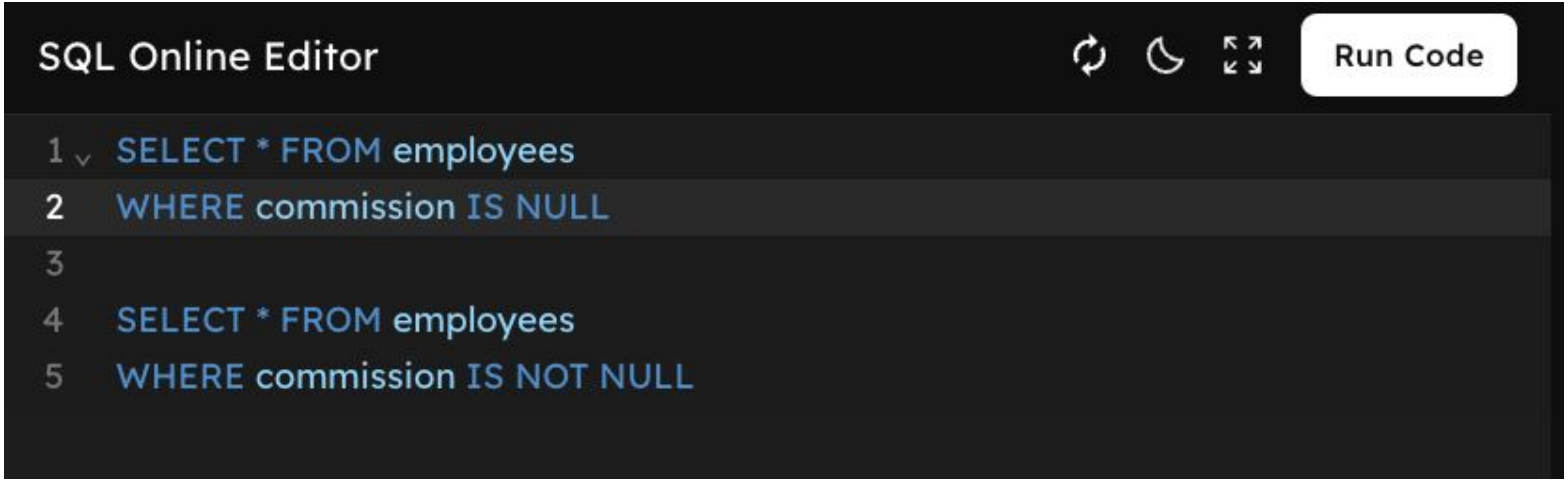


```
SELECT ord_no + purch_amt from orders;

SELECT ord_no - purch_amt from orders;

SELECT ord_no * purch_amt from orders;

SELECT ord_no / purch_amt from orders;

SELECT ord_no % purch_amt from orders;
```

Operators in SQL

NULL Operator

1. IS NULL
2. IS NOT NULL



```sql
SELECT * FROM employees
WHERE commission IS NULL

SELECT * FROM employees
WHERE commission IS NOT NULL
```

Problem Statements

1. NextLeap | New Marketing Hires in Q3

2. Engaging Movies

3. NextLeap | Consistent Communicators

4. Patients with Type I Diabetes

5. Teams Super Chatters (Saksham to fix test cases)

Q&A

What’s on your mind?

I have several questions.

Feed us back!