

What is a Subquery

22 February 2023 14:25

In SQL, a subquery is a query within another query. It is a SELECT statement that is nested inside another SELECT, INSERT, UPDATE, or DELETE statement. The subquery is executed first, and its result is then used as a parameter or condition for the outer query.

Note - The topic is slightly difficult and needs a lot of practice

Example - Find the movie with highest rating

name	rating	genre	year	released	score	votes	director	writer
The Shining	R	Drama	1980	June 13, 1980 (United States)	8.4	927000.0	Stanley Kubrick	Stephen King
The Blue Lagoon	R	Adventure	1980	July 2, 1980 (United States)	5.8	65000.0	Randal Kleiser	Henry De Vere Stacpoole
Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	June 20, 1980 (United States)	8.7	1200000.0	Irvin Kershner	Leigh Brackett
Airplane!	PG	Comedy	1980	July 2, 1980 (United States)	7.7	221000.0	Jim Abrahams	Jim Abrahams
Caddyshack	R	Comedy	1980	July 25, 1980 (United States)	7.3	108000.0	Harold Ramis	Brian Doyle-Murray
Friday the 13th	R	Horror	1980	May 9, 1980 (United States)	6.4	123000.0	Sean S. Cunningham	Victor Miller
The Blues Brothers	R	Action	1980	June 20, 1980 (United States)	7.9	188000.0	John Landis	Dan Aykroyd
Raging Bull	R	Biography	1980	December 19, 1980 (United States)	8.2	330000.0	Martin Scorsese	Jake LaMotta
Superman II	PG	Action	1980	June 19, 1981 (United States)	6.8	101000.0	Richard Lester	Jerry Siegel
The Long Riders	R	Biography	1980	May 16, 1980 (United States)	7.0	10000.0	Walter Hill	Bill Bryden
Any Which Way You Can	PG	Action	1980	December 17, 1980 (United States)	6.1	18000.0	Buddy Van Horn	Stanford Sherman

```
SELECT * FROM sql_cx_live.movies
WHERE score = (SELECT MAX(score) FROM sql_cx_live.movies)
```

What is the scope of inner query?

Types of Subqueries

22 February 2023 14:26

Based on:

1. The result it returns ✓
2. Based on working

Returned Data

Scalar ✓
subquery
(9.3)
horror

Row ✓
subquery

genres
horror
romance
action

Table ✓
subquery

genre	avg rating
horror	6.5
romance	7
action	6.8

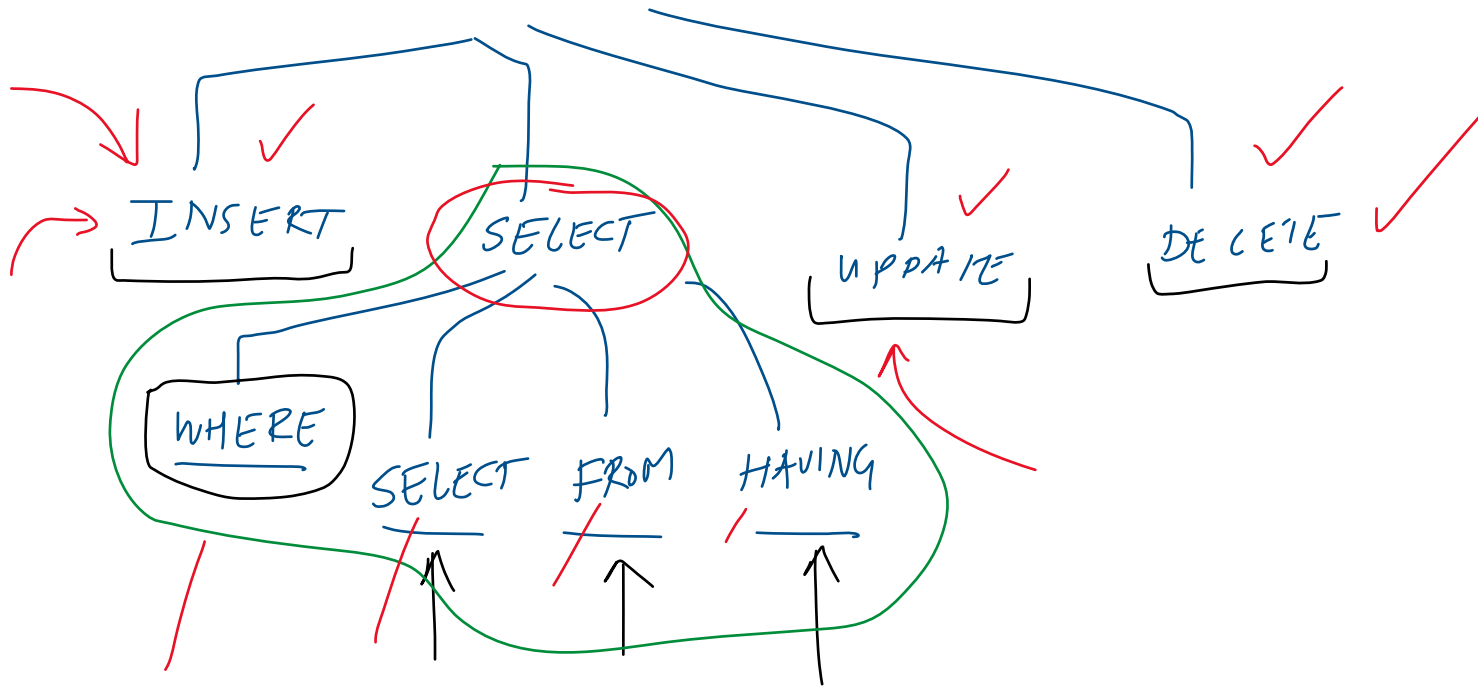
Working

Independent

Correlated

Where can subqueries be used?

22 February 2023 14:27



Independent Subquery - Scalar Subquery

22 February 2023 14:28

1. Find the movie with highest profit(vs order by)
2. Find how many movies have a rating > the avg of all the movie ratings(Find the count of above average movies) ✓
3. Find the highest rated movie of 2000
4. Find the highest rated movie among all movies whose number of votes are > the dataset avg votes



$$O(n) + O(n) = O(2n) \checkmark$$

$$O(n \log n) + O(1)$$

$$\nearrow \quad \underline{O(n \log n)} \checkmark$$

index-m-x

Independent Subquery - Row Subquery(One Col Multi Rows)

22 February 2023 14:29

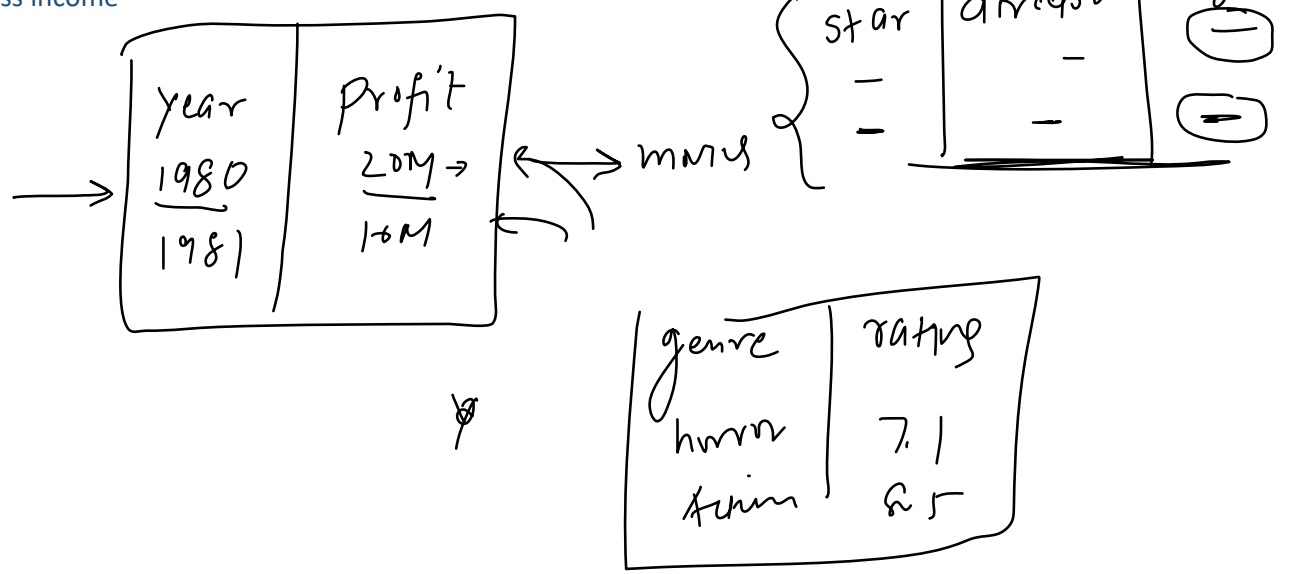
1. Find all users who never ordered
2. Find all the movies made by top 3 directors(in terms of total gross income)
3. Find all movies of all those actors whose filmography's avg rating > 8.5(take 25000 votes as cutoff)

NOT
IN

Independent Subquery - Table Subquery(Multi Col Multi Row)

22 February 2023 14:29

- 1. Find the most profitable movie of each year
- 2. Find the highest rated movie of each genre votes cutoff of 25000
- 3. Find the highest grossing movies of top 5 actor/director combo in terms of total gross income



Correlated Subquery

22 February 2023 14:29

1. Find all the movies that have a rating higher than the average rating of movies in the same genre.[Animation]
2. Find the favorite food of each customer.

Usage with SELECT ←

22 February 2023 14:29

1. Get the percentage of votes for each movie compared to the total number of votes.
2. Display all movie names ,genre, score and avg(score) of genre

-> Why this is inefficient?

20 + 20

name | percent of votes

Usage with FROM

22 February 2023 14:30

1. Display average rating of all the restaurants

Usage with HAVING

22 February 2023 14:30

1. Find genres having avg score > avg score of all the movies

Subquery In INSERT

22 February 2023 14:30

Populate a already created `loyal_customers` table with records of only those customers who have ordered food more than 3 times.

Subquery in UPDATE

22 February 2023 14:31

Populate the money col of loyal_customer table using the orders table. Provide a 10% app money to all customers based on their order value.

Subquery in DELETE

22 February 2023 14:31

Delete all the customers record who have never ordered.