Database Query Optimization

# Task

Based on the given SQL query we need to optimize this for better performance.

## Original Query

***SELECT***

***c.Name,f.Role,***

***SUM(c.hours) AS Total\_Tracked\_Hours***

***SUM(f.Estimed Hours) AS Total\_Allocated\_Hours,***

***Date***

***FROM***

***ClickUp c***

***JOIN***

***Float f on c.Name = f.Name***

***GROUP BY***

***c.Name, f.Role***

***HAVING***

***SUM(c.hours) > 100***

***ORDER BY***

***Total\_Allocated\_Hours DESC;***

## Query Optimization Process

From the above query, we can observe the following issue and where to improve.

### Origins Query Issues:

* *Absence of appropriate table joins and aliases*
* *Uncertain date column reference*
* *Inadequate join conditions*
* *No indexes were mentioned.*
* *Ineffective grouping*

### Optimized Query

Using multiple CTEs to improve the readability of queries for easy understanding and better. For larger data, we added the filter where the date range in case of larger data. Note this is an option depending on the range of data you want to see.



#### Breakdown

**Common Table Expressions (CTEs)**

The query uses two Common Table Expressions (CTEs) to organize data before the final selection:

1. **First CTE (time\_tracking)**

* Calculates total tracked hours per employee, role, and date.
* Filters data for the year 2024.
* Joins tables: ***fact\_time\_tracking, dim\_employee, dim\_date, fact\_allocation, and dim\_role***.

1. **Second CTE (allocation\_hours)**

* Calculates total allocated hours per employee and role.
* Joins tables: ***fact\_allocation, dim\_employee, and dim\_role***.

1. **Final Query**

* Combines data from the two CTEs using ***employee\_id and role\_id.***
* Filters for employees with tracked hours > 100.
* Outputs employee name, role name, date, tracked hours, and allocated hours.
* Sorts results by allocated hours in descending order.

### Improve Table Structure by Indexing

To improve query performance by adding indexes on frequently joined and sorted columns.

#### Breakdown

**Index Creation Statements**

1. **Index on fact\_time\_tracking**

* Name: idx\_fact\_time\_tracking\_employee\_date
* Columns: employee\_id, date\_id

1. **Index on fact\_allocation**

* Name: idx\_fact\_allocation\_employee\_role
* Columns: employee\_id, role\_id

1. **Index for Sorting on fact\_allocation**

* Name: idx\_fact\_allocation\_hours
* Column: estimated\_hours

