

Using Group By instead of Distinct

Diego Rosales , Josué Carrera

Using Group By instead of Distinct

Definición: Este proceso de refactorización busca optimizar las consultas que utilizan la cláusula DISTINCT mediante la aplicación y el reemplazo por GROUP BY. Esta transformación no afecta la funcionalidad de la consulta original, pero contribuye significativamente a mejorar su rendimiento y eficiencia.

```
SELECT DISTINCT product  
FROM sales
```



```
SELECT product  
FROM sales  
GROUP BY product
```

Using Group By instead of Distinct

- Precondiciones(entrada)

- Aplicable únicamente a “consultas”.
- Es una consulta válida (parsea ok).
- Que la consulta no contenga la Cláusula **GROUP BY**.
- Que la consulta contenga la Cláusula **DISTINCT**
- Que la consulta no incluya funciones de agregación.
- Que la consulta no incluya wildcard asterisco.

- Postcondiciones(salida)

- Es una consulta válida (parsea ok).
- Que el resultado contenga la Cláusula **GROUP BY**.
- Que el resultado no contenga la Cláusula **DISTINCT**.

Using Group By instead of Distinct

- Implementación

- PrePostConditionsVisitor

- Redefinimos el método "visitSelect_core(SQLiteParser.Select_coreContext ctx)" para incorporar las validaciones asociadas a las precondiciones y poscondiciones.

- TransformerVisitor

- Redefinimos el método "visitSelect_core(SQLiteParser.Select_coreContext ctx)" para llevar a cabo las transformaciones necesarias.

- TerminalNodesToTextVisitor

- Redefinimos el método "visitTerminal(TerminalNode node)" para recorrer todos los nodos terminales y retornarlos convertidos a texto separados por un espacio en blanco.

- ## Conclusión

- Using Group By instead of Distinct implementa `vistSelect_core(...)` para cumplir que solo aplica a consultas. No tomamos en cuenta las subconsultas.
- Es importante recalcar que este refactoring puede no mantener el orden de las tuplas resultantes en comparación a la consulta original.
- Si consideramos que cambiar el orden de las tuplas resultantes es no mantener el comportamiento entonces esto no es un refactoring.