

Sesión 5: OLAP Relacional (ROLAP)

En esta sesión vamos realizar algunas consultas SQL para resolver problemas OLAP, utilizando operadores avanzados de agrupación (ROLLUP y CUBE) y funciones de ventana (Window Functions). Toda la sintaxis y funcionalidad está bien descrita en las páginas de la documentación de PostgreSQL.

- ROLLUP y CUBE: <https://www.postgresql.org/docs/current/queries-table-expressions.html#QUERIES-GROUPING-SETS>
- Tutorial de funciones de ventana: <https://www.postgresql.org/docs/current/tutorial-window.html>
- Sintaxis de las funciones de ventana: <https://www.postgresql.org/docs/current/sql-expressions.html#SYNTAX-WINDOW-FUNCTIONS>
- Lista de funciones de ventana del sistema: <https://www.postgresql.org/docs/current/functions-window.html>

Group By

1.- Obtener para las productoras 'New Line Cinema' y 'Paramount Pictures' la evolución mensual de los beneficios entre principios de 2014 y finales de 2016

```
select p.nombre as productora, t.ano as ano, t.mes as mes, sum(ingresos-coste) as beneficio
from dw.finanzas f, dw.productoras p, dw.tiempo t
where f.productora = p.id
    and f.tiempo = t.id
    and p.nombre in ('Paramount Pictures', 'New Line Cinema')
    and t.ano between 2014 and 2016
group by p.id, t.ano, t.mes
order by productora, ano, mes
```

Roll-up

2.- Usamos ahora la clausula rollup. Vemos que se añaden filas para representar los totales por año y por productora, y el total global.

```
select p.nombre as productora, t.ano as ano, t.mes as mes, sum(ingresos-coste) as beneficio
from dw.finanzas f, dw.productoras p, dw.tiempo t
where f.productora = p.id
    and f.tiempo = t.id
    and p.nombre in ('Paramount Pictures', 'New Line Cinema')
    and t.ano between 2014 and 2016
group by rollup (p.nombre, t.ano, t.mes)
order by productora, ano, mes
```

Cube

3.- Usamos ahora la clausula cube. Vemos que se añaden filas para ver datos mensuales y anuales.

```
select p.nombre as productora, t.ano as ano, t.mes as mes, sum(ingresos-coste) as beneficio
from dw.finanzas f, dw.productoras p, dw.tiempo t
where f.productora = p.id
    and f.tiempo = t.id
    and p.nombre in ('Paramount Pictures', 'New Line Cinema')
    and t.ano between 2014 and 2016
group by cube (p.nombre, t.ano, t.mes)
order by productora, ano, mes
```

Funciones de ventana: Agregados sobre particiones

4.- Para cada una de las productoras obtener el porcentaje que representa el beneficio de cada mes en el beneficio anual. Obtener también el porcentaje que representa cada mes en el beneficio total de cada productora

```
select *,
    sum(beneficio) over (partition by productora, ano) as beneficio_anual,
    (beneficio / sum(beneficio) over (partition by productora, ano))*100 as porcentaje_beneficio_anual,
    (beneficio / sum(beneficio) over (partition by productora))*100 as porcentaje_beneficio_productora
from beneficio_mensual
```

5.- [Usamos ahora la sintaxis para construir una ventana de un tamaño concreto alrededor de cada fila] Suaviza los valores del beneficio mensual realizando la media de cada valor con el valor anterior y siguiente

```
select *,
    avg(beneficio) over (partition by productora order by ano, mes
                        rows between 1 preceding and 1 following) as beneficio_suavizado
from beneficio_mensual1
```

Funciones de ventana: Funciones de ranking

6.- Obtener para cada productora y mes, la posición que ocupa ese mes (ordenando por beneficio) en todos los meses de ambas productoras y en la productora actual, el percentil de ese mes en la distribución de beneficios de la productora y el ratio de incremento de ese mes respecto al mes anterior

```
select *,
    rank() over (order by beneficio desc) as ranking_total,
    rank() over (partition by productora order by beneficio desc) as ranking_productora,
    cume_dist() over (partition by productora order by ano, mes) as percentil,
    beneficio / lag (beneficio, 1, null) over (partition by productora order by ano, mes) as ratio_incremento_beneficio
from beneficio_mensual1
order by productora, ano, mes
```

ENTREGA INDIVIDUAL 5: Define 5 nuevas consultas de interés que utilicen la extensión de OLAP de SQL (Rollup, Cube, funciones de ventana). Describe para cada consultas el interés para la productora de cine y entrega el enunciado, la descripción y el código SQL.

Última modificación: martes, 18 de octubre de 2022, 14:12