

Práctica 4. CONSULTAS COMBINADAS Y AGRUPADAS

0. Introducción

En esta práctica se continúa con el estudio de la sentencia SELECT, presentando dos tipos de consultas nuevas: las agrupadas y las que se combinan utilizando los operadores de conjuntos. A partir de esta práctica utilizaremos para los ejemplos las tablas de la base de datos de Tráfico.

MARCAS	
COD_MARCA	NOM_MARCA
1	AUDI
2	BMW
3	CITROEN
4	FORD
5	MERCEDES
6	NISSAN
7	OPEL
8	RENAULT
9	VOLVO

MODELOS			
COD_MARCA	COD_MODELO	NOM_MODELO	POTENCIA
1	1	A3 1.8	110
1	2	A4 1.8	110
1	3	A4 2.0	130
1	4	A6 2.0	130
1	5	A8 3.0	240
7	1	VECTRA 2.0	150

PERSONAS							
NIF	APELLIDO1	APELLIDO2	NOMBRE	DOMICILIO	COD_POSTAL	MUNICIPIO	COD_PROV
07345128H	PEREZ	PEREZ	JUAN	CL/ PARIS, 20	10005	CACERES	10
27365138A	GARCIA	SANCHEZ	PAULA	CL/ PARRAS, 6	10001	CACERES	10
10134562B	LARRA	ALVAREZ	JUAN PEDRO	PL/ MAYOR, 15	10001	CACERES	10
57176355K	VAZQUEZ	PEDROSA	IRENE	PL/ MAYOR, 4	10001	CACERES	10
08124356C	TENA	TENA	ALBERTO	CL/ DINAMARCA, 8	6001	BADAJOS	6

VEHICULOS					
MATRICULA	FEC_MATRICULACION	PROPIETARIO	BASTIDOR	COD_MARCA	COD_MODELO
9744BFN	03/11/01	08124356C		1	2
7715BFP	01/12/01	08124356C		1	5
0001BBB	01/04/01	27365138A		7	1

AGENTES		
NUM_AGENTE	NIF	UNIDAD
1	07345128H	1
2	57176355K	1

MULTAS										
EXP	INFRACTOR	NUM_AGENTE	FECHA	ARTICULO	IMPORTE	MATRICULA	LUGAR	KM	DIRECCION	ESTADO
1	08124356C	1	01/11/02	144-B	100	7715BFP	N-521	13	TRUJILLO-CACERES	1
2	08124356C	2	30/11/02	144-B	100	0001BBB	N-521	34	CACERES-TRUJILLO	2
3	27365138A	2	02/11/02	130	300	0001BBB	N-630	545	GIJON-SEVILLA	1

*El nombre de la columna EXP es en realidad EXPEDIENTE y de KM es kilómetro

1. Combinación de consultas mediante operadores de conjuntos

Combinan los resultados de múltiples consultas en un simple resultado. Existen tres operadores:

- UNION (con la variante UNION ALL)
- INTERSECT
- MINUS

1.1 UNION

El operador **UNION** devuelve las tuplas de cada una de las consultas.

Ejemplo: Seleccionar los dni de los propietarios de vehículos matriculados en el 2001 y el de los infractores del artículo 144-B

```
/*      ( $\Pi_{\text{propietario}}(\sigma_{\text{fec\_mat} \geq '01/01/01' \text{ and } \text{fec\_mat} \leq '31/12/01' (\text{Vehiculos})) \cup$   
  
 $\Pi_{\text{infractor}}(\sigma_{\text{articulo} = '144-B' (\text{multas}))$       */
```

```
SELECT DISTINCT propietario  
FROM practbd.vehiculos  
WHERE fec_matriculacion between '01/01/01' AND '31/12/01'  
UNION  
SELECT infractor  
FROM practbd.multas  
WHERE articulo = '144-B';
```

PROPIETAR

08124356C

27365138A

2 filas seleccionadas.

Si se utiliza el operador **UNION ALL** se devuelve el resultado sin eliminar los duplicados.

PROPIETAR

08124356C

27365138A

08124356C

08124356C

4 filas seleccionadas.

1.2 INTERSECT

El operador **INTERSECT** devuelve la intersección de los conjuntos seleccionados en cada consulta. MySQL no soporta este operador.

Ejemplo: Seleccionar los dni de los infractores por multas de importe superior a 200 euros que residan en Cáceres.

```
/* ( $\Pi_{\text{nif}}(\sigma_{\text{municipio} = 'CACERES' (\text{Personas})) \cap (\Pi_{\text{infractor}}(\sigma_{\text{importe} > 200 (\text{Multas}))$       */
```

```
SELECT nif  
FROM practbd.personas  
WHERE municipio = 'CACERES'  
INTERSECT  
SELECT infractor  
FROM practbd.multas  
WHERE importe > 200;
```

```
NIF
-----
27365138A
```

1 fila seleccionada.

1.3 MINUS

El operador **MINUS** devuelve la diferencia de los conjuntos seleccionados en cada consulta. MySQL no soporta este operador.

Ejemplo: Seleccionar los dni de las personas que no son propietarias de vehículos.

```
/* (Πnif(personas)) - (Πpropietario(vehiculos)) */
```

```
SELECT nif
FROM practbd.personas
MINUS
SELECT propietario
FROM practbd.vehiculos;
```

```
NIF
-----
07345128H
10134562B
57176355K
```

3 filas seleccionadas.

2. Consultas agrupadas

Las consultas agrupadas en SQL, nos permiten llevar a cabo agrupaciones de tuplas que tienen el mismo valor para un atributo o atributos especificado. Los atributos por los que se va a agrupar se ponen en la cláusula GROUP BY separados por comas. **Sólo podremos seleccionar los atributos agrupados y funciones de grupo.**

Sintaxis

```
SELECT lista_atributos_agrupados/funciones_grupo
FROM esquema.nombre_tabla
[WHERE condicion]
GROUP BY lista_atributos_agrupados [HAVING condicion_grupo]
[ORDER BY lista_atributos [ASC/DESC]]
```

Ejemplo: Seleccionar para cada infractor, el numero de multas y el importe total de las mismas

```
/* Σ (infractor) ⊕ sum(importe) total ⊕ count(*) multas (multas) */
```

```
SELECT infractor , COUNT(*) numero, SUM(importe) total
FROM practbd.multas
GROUP BY infractor;
```

INFRACITOR	NUMERO	TOTAL
08124356C	2	180
27365138A	1	300

2 filas seleccionadas.

Para seleccionar grupos pondremos el predicado de selección del grupo en la cláusula HAVING.

Ejemplo: Seleccionar los códigos de marca que tienen mas de 2 modelos con una potencia superior a 125.

```
/*  $\Pi_{\text{cod\_marca}} (\sigma_{\text{numero} > 2} (\sum(\text{cod\_marca}) \oplus_{\text{count}(*)} \text{numero} (\sigma_{\text{potencia} > 125} (\text{modelos}))))$  */
```

```
SELECT cod_marca
FROM practbd.modelos
WHERE potencia > 125
GROUP BY cod_marca
HAVING COUNT(*) > 2
```

```
COD_MARCA
-----
1
```

1 fila seleccionada.

Anexo I: Esquema sakila

Para el desarrollo de las prácticas se ha incluido dentro del esquema sakila de Oracle, una base de datos con información acerca de una tienda de alquiler de dvds de películas. A continuación se presentan varias tablas pertenecientes a esta base de datos:

FILM

ATRIBUTO	TIPO	DESCRIPCION	NULO
FILM_ID	NUMBER(10)	Identificador de la película.	NO
TITLE	VARCHAR2(255)	Título.	NO
DESCRIPTION	CLOB	Trama.	SI
RELEASE_YEAR	NUMBER(4)	Año de lanzamiento.	SI
LANGUAGE_ID	NUMBER(10)	Código del idioma en la que se encuentra el dvd de la película.	NO
ORIGINAL_LANGUAGE_ID	NUMBER(10)	Código del idioma original de la película.	SI
RENTAL_DURATION	NUMBER(10)	Duración del alquiler del dvd en días.	NO
RENTAL_RATE	NUMBER(6,2)	Precio del alquiler del dvd en dólares.	NO
LENGTH	NUMBER(10)	Duración de la película.	SI
REPLACEMENT_COST	NUMBER(6,2)	Coste de reemplazo del dvd.	NO
RATING	VARCHAR2(5)	Clasificación parental,	SI
SPECIAL_FEATURES	VARCHAR2(54)	Características especiales del dvd.	SI
LAST_UPDATE	DATE	Ultima actualización de los datos del dvd.	NO

FILM_ACTOR

ATRIBUTO	TIPO	DESCRIPCION	NULO
FILM_ID	NUMBER(5)	Identificador de la película.	NO
ACTOR_ID	NUMBER(5)	Identificador del actor	NO
LAST_UPDATE	DATE	Ultima actualización de datos del actor relacionado con el film	NO

ACTOR

ATRIBUTO	TIPO	DESCRIPCION	NULO
ACTOR_ID	NUMBER(5)	Identificador del actor	NO
FIRST_NAME	VARCHAR2(45)	Nombre del actor	NO
LAST_NAME	VARCHAR2(45)	Apellido del actor	NO
LAST_UPDATE	DATE	Ultima actualización de datos del actor	NO

Anexo II: Ejercicios

1. Seleccionar los códigos de clasificación (rating) que tienen al menos una película de más de 180 minutos de duración.

```
RATING
-----
NC-17
PG-13
PG
R
G
```

2. Seleccionar los códigos de clasificación que tienen al menos una película cuyo título empieza por 'BIRD'.

```
RATING
-----
G
NC-17
```

3. Seleccionar los códigos de clasificación que tienen al menos una película de más de 180 minutos de duración y otra cuyo título empieza por 'BIRD'. Razonar por qué esta consulta no puede hacerse usando el operador lógico AND.

```
RATING
-----
G
NC-17
```

4. Obtener, por cada periodo distinto de alquiler, el número de películas y el importe mínimo y máximo de alquiler para cada uno considerando sólo aquellas películas clasificadas como 'G'. Ordenar de mayor a menor por el periodo de alquiler.

PERIODO	PELICULAS	MINIMO	MAXIMO
-----	-----	-----	-----
7	29	,99	4,99
6	39	,99	4,99
5	33	,99	4,99
4	28	,99	4,99
3	49	,99	4,99

5. Seleccionar el número de películas existentes por cada valor del código de clasificación (rating), considerando únicamente las películas que tienen un periodo de alquiler de 5 días y siempre que dicho número de películas sea mayor de 40.

RATING	PELICULAS
-----	-----
NC-17	43
PG-13	42

6. Seleccionar, por cada periodo distinto de alquiler, el importe total del alquiler para todas las películas que corresponden a dicho periodo, siempre que tengan al menos 20 películas y el importe total de alquiler sea superior a 600 dólares. Ordenar mayor a menor por el importe total del alquiler.

DURACION	IMPORTE
6	613,88
5	611,09
4	602,97

7. Seleccionar los identificativos de las películas que cumplen las siguientes características, ordenados descendentemente:

- Tienen la clasificación 'G'
- Interviene un sólo actor
- La duración del alquiler es de 3 días

FILM_ID
848
575
472
50

8. Seleccionar los identificativos de las películas en las que no intervienen actores, ordenando de menor a mayor por el identificador de película.

FILM_ID
102
257
264
323
352
580
581
681
701
789
803
995

9. Seleccionar el importe total de reemplazo de las películas para cada clasificación (rating), siempre que dicha clasificación tenga películas con al menos cuatro periodos distintos de alquiler.

RATING	COSTE_REEMPLAZO
G	3582,22
NC-17	4228,9
PG	3678,06
PG-13	4549,77
R	3945,05

10. Seleccionar, para los actores cuyo identificador está entre 150 y 160, el identificador del actor y el número de películas en las que ha intervenido cada actor, ordenando la salida por el identificador del actor (incluir los actores que tienen 0 películas).

ACTOR_ID	PELICULAS
150	34
151	26
152	23

153	20
154	28
155	31
156	22
157	17
158	0
159	0
160	0