

Elementos Fundamentales SQL

Tercera parte

Funciones agrupadoras

Contar registros (count)

En SQL tendremos funciones que nos realizan operaciones con conjuntos de registros. La más básica es la que nos permite contar el número de registros.

Formato:

```
select count(*) from tabla
```

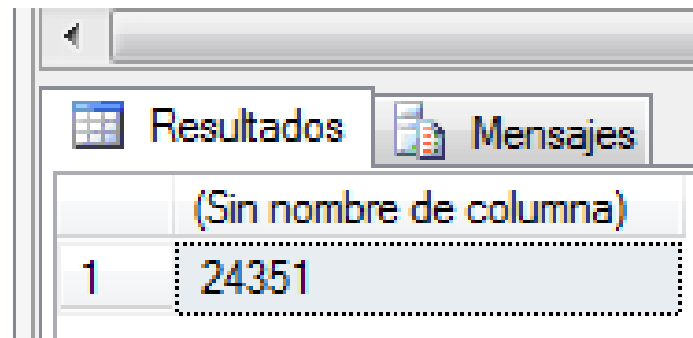
En este caso el * será válido, ya que representa que contaremos **todos** los registros de la tabla.

Usando la base de datos parobasicas veremos algunos ejemplos.
Esta base de datos contiene una única tabla, la DatosCompletoTabla, con los datos del nº de habitantes (padron) y el nº de parados (TotalParoRegistrado) de cada municipio de España en los tres primeros meses de 2013. Se especifica también la Provincia y la Comunidad Autónoma (CA) del municipio.

```
--select completo  
SELECT id,CA,Provincia,ISLA,Municipio,Fecha  
        ,Padron,TotalParoRegistrado  
FROM DatosCompletoTabla;  
GO
```

Count(*) contará todos los registros de la tabla

```
-- contar los registros totales  
select count (*)  
from DatosCompletoTabla;  
go
```

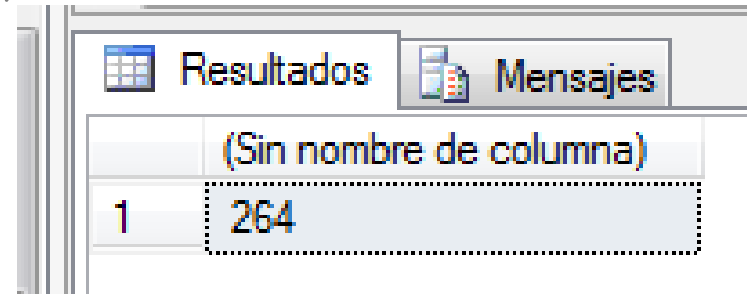


The screenshot shows a window with two tabs: 'Resultados' (Results) and 'Mensajes' (Messages). The 'Resultados' tab is active, displaying a single row of data. The column header is '(Sin nombre de columna)' (No column name). The row contains the value '24351'.

	(Sin nombre de columna)
1	24351

Contar los registros que cumplen una condición especificada en el Where.

```
--contar los registros con filtro  
--contar los registros de Canarias  
select count(*)  
from DatosCompletoTabla  
where CA='Canarias';  
go
```



The screenshot shows a SQL query results window with two tabs: 'Resultados' (Results) and 'Mensajes' (Messages). The 'Resultados' tab is active, displaying a table with one column and one row. The column header is '(Sin nombre de columna)' (No column name) and the row value is 264.

	(Sin nombre de columna)
1	264

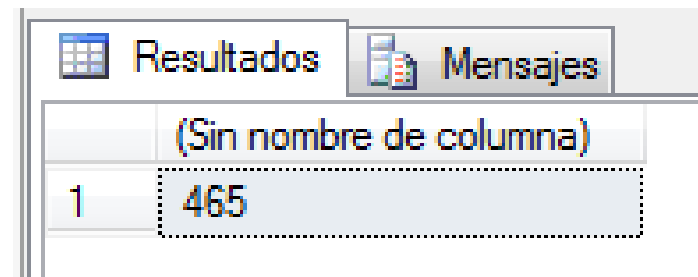
```
--contar los registros del mes de marzo  
select count(*)  
from DatosCompletoTabla  
where datename(month, fecha) = 'Marzo';  
go
```

Resultados		Mensajes
(Sin nombre de columna)		
1	8117	

También podemos contar los registros para los que un campo no tome el valor NULL
Sólo tenemos que colocarlo dentro del count :

count(campo)

```
--contar los registros con isla no nulo  
select count(isla)  
from DatosCompletosTabla  
go
```



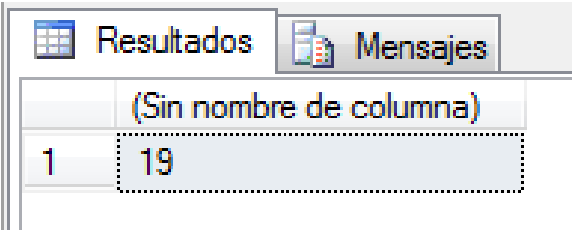
Resultados		Mensajes
(Sin nombre de columna)		
1	465	

La última posibilidad es que contemos cuántos valores distintos toma un campo, sin tener en cuenta repeticiones.

count(distinct campo)

Por ejemplo, contamos las comunidades autónomas diferentes que aparecen en nuestra tabla, aunque aparezcan muchas veces, las contaremos una única vez.

```
--contar las Comunidades Autónomas  
select count(distinct CA)  
from DatosCompletoTabla  
go
```



Resultados		Mensajes
(Sin nombre de columna)		
1	19	

En el caso en que el nº de registros sea muy grande podremos usar **count_big** para contarlos.

Funcionarán igual:

```
select count_big(*) from tabla
```

nº de registros

```
select count_big(campo) from tabla
```

nº de registros con campo no NULL

```
select count_big(distinct campo) from tabla
```

nº de valores distintos de campo

Ejercicio:

- Contar nº de registros con provincia Santa Cruz de Tenerife.
- Contar el número de Islas diferentes.
- Contar el nº de Municipios de la provincia de "Palmas, Las".
- Contar el número de registros del segundo mes del año.
- Igual que el anterior si el número fuera muy grande.
- Contar los registros con menos de 1000 habitantes con islas.
- Contar las comunidades autónomas que contengan una E.

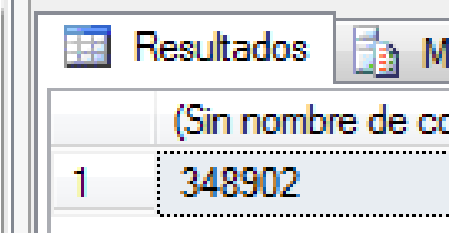
Sumar campos (sum)

Otra función agrupadora es la que nos permite sumar los valores de un campo numérico

Su formato es:

select sum(campo) from tabla



```
--Sumar el número de habitantes de la provincia de Lugo
--según los datos de fecha 01/01/2013
set dateformat dmy
select sum(padron)
from DatosCompletosTabla
where (fecha='01/01/2013') and (provincia='Lugo');
go
```



Resultados

	(Sin nombre de co
1	348902

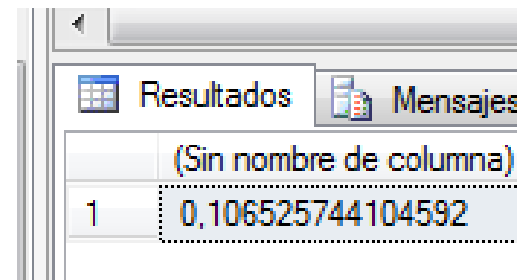
```
--Sumar el número de parados y el número de habitantes
--según los datos de fecha 01/02/2013
set dateformat dmy
select sum(TotalParoRegistrado), sum(Padron)
from DatosCompletoTabla
where (fecha='01/02/2013');
go
```

<div> <div> Resultados</div> <div> Mensajes</div> </div>		
	(Sin nombre de columna)	(Sin nombre de columna)
1	5040222	47267851

Podemos usar varios campos SUM en operaciones.

Hay que tener en cuenta que devuelven un dato del tipo del campo que sumamos, por lo que para algunas operaciones hay que realizar un cast previo que permita su ejecución.

```
--calcular el cociente entre en número de parados  
--entre el número de habitantes para toda España  
--según los datos de fecha 01/03/2013  
set dateformat dmy  
select cast(sum(TotalParoRegistrado) as float) /  
        cast(sum(padron) as float)  
from DatosCompletoTabla  
where (fecha='01/03/2013');  
go
```



Resultados		Mensajes
	(Sin nombre de columna)	
1	0,106525744104592	

Ejercicios:

- Sumar el padrón a fecha '01/03/2013' de la provincia de Santa Cruz de Tenerife
- Sumar el paro a fecha '01/02/2013' de los municipios de la comunidad autónoma de "Madrid, Comunidad de"
- Sumar los habitantes a fecha '01/02/2013' de las Comunidades autónomas de Extremadura, Andalucía, Aragón y Canarias.

Resto de funciones agrupadoras: max, min,
avg, stdev y var.

Las funciones agregadoras como el sum y count , junto a las que veremos en este apartado devuelven un único valor (salvo agrupamientos que veremos después). Por tanto si las acompañamos de un campo nos darán un error.

```
select provincia, SUM(padron)
from DatosCompletoTabla
where (fecha='01/03/2013');
go
```

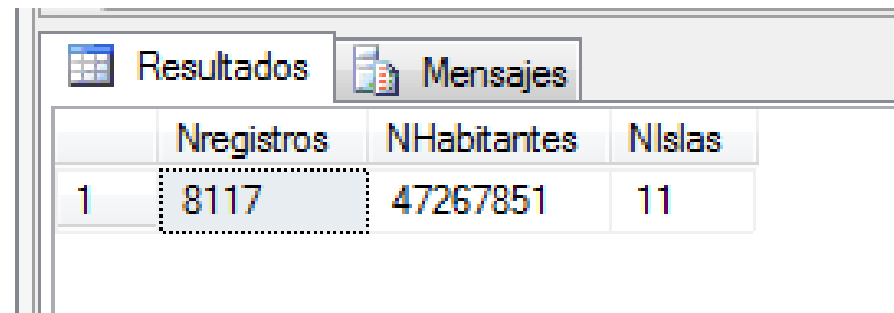
Mens. 8120, Nivel 16, Estado 1, Línea 1
La columna 'DatosCompletoTabla.Provincia' de la lista de selección no es válida, porque no está contenida en una función de agregado ni en la cláusula GROUP BY.

Lo que sí estará permitido es sacar varios valores de funciones agrupadoras en el mismo select.

Pudiéndoles dar nombre con el alias.

Por ejemplo:

```
select count(*) as Nregistros,  
       SUM(padron) as NHabitantes,  
       count(distinct isla) as NIslas  
from DatosCompletoTabla  
where (fecha='01/03/2013');  
go
```



The screenshot shows a SQL query results window with two tabs: 'Resultados' (active) and 'Mensajes'. The 'Resultados' tab displays a table with the following data:

	Nregistros	NHabitantes	NIslas
1	8117	47267851	11

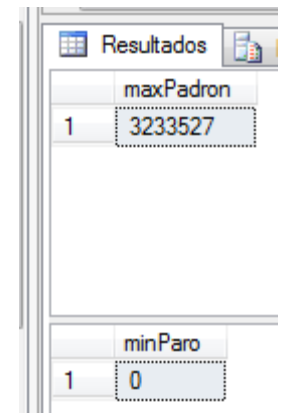
Min y max nos devuelve el valor máximo o mínimo de una expresión.

En este caso puede ser numérica o de texto. En este último caso el máximo o mínimo hace referencia a la ordenación alfanumérica de textos.

--Mostrar el valor máximo del padrón

```
select max(padron) as maxPadron  
from DatosCompletoTabla;
```

go



maxPadron	
1	3233527

minParo	
1	0

--Mostrar el valor mínimo del paro

```
select min(TotalParoRegistrado) as minParo  
from DatosCompletoTabla;
```

go

--Mostrar la CA de menor valor alfanumérico

```
select min(ca)
from DatosCompletoTabla;
go
```

--Mostrar el municipio de mayor valor alfanumérico

```
select max(municipio)
from DatosCompletoTabla;
go
```

Resultados		Mensajes
(Sin nombre de columna)		
1	Andalucía	
(Sin nombre de columna)		
1	Zurgena	

Funciones agregadas estadísticas:

- AVG: media
- STDEV: desviación típica
- VAR: varianza

Datos: 2, 3, 6, 8, 11.

Media

$$\bar{x} = \frac{2 + 3 + 6 + 8 + 11}{5} = 6$$

Varianza

$$\sigma^2 = \frac{2^2 + 3^2 + 6^2 + 8^2 + 11^2}{5} - 6^2 = 10.8$$

Desviación típica

$$\sigma = \sqrt{10.8} = 3.286$$

```

--media, varianza y desviación típica
--de los datos de fecha 01/03/2013
select avg(padron) as media, var(padron) as varianza
      , stdev(padron) as desviacionTipica
from DatosCompletosTabla
where (fecha='01/03/2013');
go

```

Resultados		Mensajes	
	media	varianza	desviacionTipica
1	5823	2227689097,19549	47198,4014262718

Ejercicios:

- Calcular nº de registros, nº de registros con isla, media de paro, suma de padrón, varianza de padrón de los datos de Canarias y Andalucía de fecha 01/02/2013.
- Calcular la media de los datos de paro de los registros con isla.
- Calcular la media y la desviación típica de los datos de padrón de los municipios que contengan Villa.

Cláusula distinct

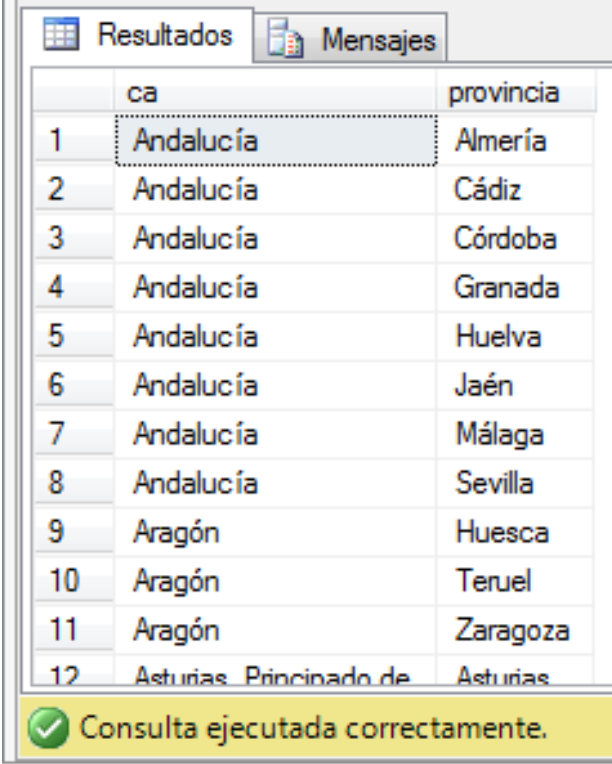
Se puede especificar la cláusula `distinct` al principio de la lista de campos del `select` para que nos muestre sólo los valores diferentes de la lista de campos, aunque aparezcan muchos registros con ese conjunto de valores.

Por ejemplo podemos ver los nombres de las Comunidades Autónomas.

```
--Mostrar las diferentes comunidades
--autónomas ordenadas
select distinct ca
from DatosCompletoTabla
order by ca;
go
```

Resultados		Mensajes
	ca	
1	Andalucía	
2	Aragón	
3	Asturias, Principado de	
4	Baleares, Illes	
5	Canarias	
6	Cantabria	
7	Castilla - La Mancha	
8	Castilla y León	
9	Cataluña	
10	Ceuta	
11	Comunitat Valenciana	
12	Extremadura	
13	Galicia	
14	Madrid, Comunidad de	
15	Melilla	
16	Murcia, Región de	
17	Navarra, Comunidad ...	
18	País Vasco	
19	Rioja, La	

```
--Mostrar las diferentes provincias con su comunidad autónoma
--ordenadas por comunidad autónoma y provincia
select distinct ca, provincia
from DatosCompletoTabla
order by ca, provincia;
go
```



	ca	provincia
1	Andalucía	Almería
2	Andalucía	Cádiz
3	Andalucía	Córdoba
4	Andalucía	Granada
5	Andalucía	Huelva
6	Andalucía	Jaén
7	Andalucía	Málaga
8	Andalucía	Sevilla
9	Aragón	Huesca
10	Aragón	Teruel
11	Aragón	Zaragoza
12	Asturias Principado de	Asturias

✓ Consulta ejecutada correctamente.

Ejercicios:

Mostrar los nombres de las provincias (sólo una vez), ordenados.

Mostrar los nombres de los municipios de Aragón (sólo una vez), ordenados.

Cláusula TOP

La cláusula TOP se utiliza al inicio de la lista de campos del select para indicar el número de registros que queremos sacar.

select TOP nº lista de campos...

```
select top 10 ca, provincia, municipio, fecha, padron
from DatosCompletoTabla;
go
```

Resultados		Mensajes			
	ca	provincia	municipio	fecha	padron
1	Castilla y León	Ávila	Hija de Dios, La	2013-01-01 00:00:00.000	85
2	Castilla y León	Ávila	Hija de Dios, La	2013-02-01 00:00:00.000	85
3	Castilla y León	Ávila	Hija de Dios, La	2013-03-01 00:00:00.000	85
4	Castilla y León	Ávila	Horcajada, La	2013-01-01 00:00:00.000	556
5	Castilla y León	Ávila	Horcajada, La	2013-02-01 00:00:00.000	556
6	Castilla y León	Ávila	Horcajada, La	2013-03-01 00:00:00.000	556
7	Castilla y León	Ávila	Horcajo de las Torres	2013-01-01 00:00:00.000	603
8	Castilla y León	Ávila	Horcajo de las Torres	2013-02-01 00:00:00.000	603
9	Castilla y León	Ávila	Horcajo de las Torres	2013-03-01 00:00:00.000	603
10	Castilla y León	Ávila	Homillo, El	2013-01-01 00:00:00.000	342

Suele ser más útil cuando ordenamos los registros.

```
--Mostrar los 5 primeros municipios en nº de habitantes  
--para el dato de fecha 01/03/2013.
```

```
select top 5 ca, provincia, municipio, fecha, padron  
from DatosCompletosTabla  
where fecha='01/03/2013'  
order by padron desc;  
go
```

	ca	provincia	municipio	fecha	padron
1	Madrid, Comunidad de	Madrid	Madrid	2013-03-01 00:00:00.000	3233527
2	Cataluña	Barcelona	Barcelona	2013-03-01 00:00:00.000	1620943
3	Comunitat Valenciana	Valencia/València	Valencia	2013-03-01 00:00:00.000	797028
4	Andalucía	Sevilla	Sevilla	2013-03-01 00:00:00.000	702355
5	Aragón	Zaragoza	Zaragoza	2013-03-01 00:00:00.000	679624

Se puede especificar en vez de un número exacto, un porcentaje de registros.

TOP nº percent

```
--Mostrar el 2% de los municipios en nº de habitantes  
--para el dato de fecha 01/03/2013 de Extremadura  
select top 2 percent ca, provincia, municipio, fecha, padron  
from DatosCompletosTabla  
where (fecha='01/03/2013') and (CA='Extremadura')  
order by padron desc;  
go
```

Resultados		Mensajes			
	ca	provincia	municipio	fecha	padron
1	Extremadura	Badajoz	Badajoz	2013-03-01 00:00:00.000	152270
2	Extremadura	Cáceres	Cáceres	2013-03-01 00:00:00.000	95668
3	Extremadura	Badajoz	Mérida	2013-03-01 00:00:00.000	58164
4	Extremadura	Cáceres	Plasencia	2013-03-01 00:00:00.000	41002
5	Extremadura	Badajoz	Don Benito	2013-03-01 00:00:00.000	37048
6	Extremadura	Badajoz	Almendralejo	2013-03-01 00:00:00.000	34694
7	Extremadura	Badajoz	Villanueva de la Serena	2013-03-01 00:00:00.000	26071
8	Extremadura	Cáceres	Navalmoral de la Mata	2013-03-01 00:00:00.000	17401

Podremos tener en cuenta los empates. Para ello tenemos la cláusula:

TOP n° WITH TIES

que nos mostrará en número de registros especificado, añadiendo los que empaten con el último en criterio de ordenación.

--Mostrar los 7 municipios con menos habitantes

--para el dato de fecha 01/03/2013

--mostrando los municipios que empaten con el último

```
select top 7 with ties municipio, padron
```

```
from DatosCompletoTabla
```

```
where (fecha='01/03/2013')
```

```
order by padron asc
```

Resultados		Mensajes
	municipio	padron
1	Illán de Vacas	2
2	Jaramillo Quemado	4
3	Castilnuevo	8
4	Estepa de San Juan	8
5	Valdemadera	9
6	Villaroya	9
7	Villamedianilla	10
8	Villanueva de Gormaz	10
9	Salcedillo	10

Ejercicio:

- Mostrar los 10 municipios con más habitantes dentro de los de menos de 1000. Datos de fecha 01/03/2013
- Mostrar los 10 municipios con más parados dentro de los de menos de 100 habitantes, sacando los que empaten con el último. Datos de fecha 01/01/2013.
- Mostrar los 10 municipios de Canarias con menor cociente paro partido por habitantes. Datos de fecha 01/01/2013