

BASES DE DATOS II

MEMORIA DESCRIPTIVA: PRÁCTICA 3 DE BASES DE DATOS II

Cristina Abdul Massih
Juan José Coronado
José Manuel Martínez

2º ISI
Escuela Politécnica de Montepíncipe
CEU San Pablo

Madrid, Marzo 2020

INTRODUCCIÓN

La siguiente práctica busca ofrecer una aplicación desarrollada en Java, en donde se realiza una conexión con las bases de datos “Sakila” y “World.” De esta forma, la aplicación se conecta a una instancia de MySQL para acceder a las bases de datos de ejemplo. La BBDD Sakila cuenta con diversa información de películas mientras que la BBDD World contiene información sobre países del mundo. Se realizarán seis consultas a ambas bases de datos, pudiendo ser consultas con funciones, simples o complejas. La aplicación desarrollada mide el tiempo de ejecución de cada consulta en nanosegundos. Esta práctica busca comparar tiempos de ejecución de consultas sobre las dos bases de datos, y analizar por qué tardarán más o menos. Para ello, se demostrará la consulta ejecutada en la aplicación desarrollada, su duración calculada en dos ejecuciones y la duración que ofrece el terminal (que es menos exacta) para la misma consulta.

ANÁLISIS DE CONSULTAS

Consultas con funciones

Sakila

La siguiente consulta sobre Sakila utiliza las funciones max, min, avg y sum (máximo, mínimo, promedio y suma) sobre las ventas por categoría de película. Por ende, debe devolver 4 números.

Imagen 1. Consulta 1 con funciones sobre Sakila.

Aplicación Consultas

SELECT

FROM

WHERE

GROUP BY

ORDER BY

LIMIT

☒ sakila ☐ world

Tiempo de Ejecución: 84911645 nanosegundos

El tiempo de ejecución en segundos es 0.084911645. Esta consulta puede llegar a demorarse mucho tiempo para ejecutarse ya que debe leer todos los datos de cada columna y realizar la operación sobre ella. Mientras más datos tenga la tabla, más tiempo demorará en ejecutarse. Sin embargo, en Sakila se hacen las funciones especificadas solo sobre 16 categorías existentes de películas, por lo que el tiempo esperado de respuesta es 0.08 segundos. Este es un tiempo de ejecución corto, ya que solo debe leer 16 filas y ejecutar las funciones.

Imagen 2. Segunda consulta 1 con funciones sobre Sakila.

Aplicación Consultas

SELECT

FROM

WHERE

GROUP BY

ORDER BY

LIMIT

☒ sakila ☐ world

Tiempo de Ejecución: 75087215 nanosegundos

Se puede observar como el tiempo de respuesta es muy similar al obtenido al ejecutar la consulta por primera vez, disminuyendo un poco únicamente.

```
mysql> select max(total_sales), min(total_sales), avg(total_sales), sum(total_sales) from sales_by_film_category;
+-----+-----+-----+-----+
| max(total_sales) | min(total_sales) | avg(total_sales) | sum(total_sales) |
+-----+-----+-----+-----+
|          5314.21 |          3417.72 |         4212.910000 |          67406.56 |
+-----+-----+-----+-----+
1 row in set (0.08 sec)
```

La siguiente consulta, ejecutada nuevamente sobre Sakila, pretende contar todas las direcciones almacenadas en la base de datos

Imagen 3. Consulta 2 con funciones sobre Sakila.

The screenshot shows a web application titled 'Aplicación Consultas'. It has a form with the following fields: 'SELECT' containing 'count(*)', 'FROM' containing 'address', 'WHERE' (empty), 'GROUP BY' (empty), 'ORDER BY' (empty), and 'LIMIT' (empty). To the right of the form are two radio buttons: 'sakila' (selected) and 'world'. Below the form, it displays 'Tiempo de Ejecución: 1291468 nanosegundos' and a button labeled 'Realizar Consulta'.

Se puede observar como el tiempo de ejecución de la consulta son 1291468 nanosegundos, 0,001291468 segundos. Este tiempo comparado con el obtenido en la consulta anterior es más bajo ya que en este caso no es necesario para la consulta realizar ningún tipo de función si no que únicamente debe contar los registros almacenados por la base de datos.

Imagen 4. Segunda consulta 2 con funciones sobre Sakila

This screenshot is identical to the one in Image 3, showing the 'Aplicación Consultas' interface with the same SQL query and selected 'sakila' database. The only difference is the execution time displayed at the bottom, which is now 'Tiempo de Ejecución: 1432204 nanosegundos'.

Se puede volver a apreciar que el tiempo de ejecución es prácticamente igual al anterior y no se ha producido demasiado cambio en estos tiempos de ejecución a pesar de ejecutar la consulta por segunda vez

Database changed
mysql> select count(*) from address;

| count(*) |
|----------|
| 603 |

1 row in set (0.04 sec)

mysql>

World

La siguiente consulta muestra por pantalla el nombre de un país, el mismo en comillas y la longitud de la palabra del país.

Imagen 5. Consulta 1 con funciones sobre World.

Aplicación Consultas

SELECT: Name, quote(Name), length(Name) as "Longitud nombre"

FROM: country

WHERE:

GROUP BY:

ORDER BY:

LIMIT:

☐ sakila ☒ world

Tiempo de Ejecución: 51626303 nanosegundos

Realizar Consulta

El tiempo de ejecución en segundos es 0.0051626303 y tiene como resultado 239 filas. Es una consulta sencilla mientras hayan pocos datos. La consulta se repitió varias veces y el resultado obtenido fue el mismo.

Imagen 6. Segunda consulta 1 con funciones en World

Aplicación Consultas

SELECT: Name, quote(Name), length(Name) as "Longitud nombre"

FROM: country

WHERE:

GROUP BY:

ORDER BY:

LIMIT:

☐ sakila ☒ world

Tiempo de Ejecución: 5826423 nanosegundos

Realizar Consulta

El tiempo de ejecución nuevamente ha sido muy similar al obtenido al ejecutar la consulta por primera vez.

| | | |
|----------------------------|----------------|----|
| Yugoslavia | 'Yugoslavia' | 10 |
| South Africa | 'South Africa' | 12 |
| Zambia | 'Zambia' | 6 |
| Zimbabwe | 'Zimbabwe' | 8 |
| +-----+-----+ | | |
| 239 rows in set (0.02 sec) | | |

En esta consulta se busca obtener los distintos idiomas además de que se escriban en mayúsculas y en minúsculas.

Imagen 7. Consulta 2 con funciones sobre World.

The screenshot shows a web application titled 'Aplicación Consultas'. It contains a form for writing SQL queries with fields for SELECT, FROM, WHERE, GROUP BY, ORDER BY, and LIMIT. The SELECT field contains 'Language, upper(Language), lower(Language)' and the FROM field contains 'countrylanguage'. To the right of the form are two radio buttons: 'sakila' (unselected) and 'world' (selected). Below the form, it displays 'Tiempo de Ejecución: 6065751 nanosegundos' and a button labeled 'Realizar Consulta'.

El tiempo de ejecución obtenido ha sido de 0,006065751 segundos, tiempo bastante pequeño ya que es una función simple que únicamente recorre la base de datos imprimiendo el valor de los registros que contiene, además de cambiarlos a mayúsculas o a minúsculas.

Imagen 8: Segunda consulta 2 con funciones sobre World

This screenshot is identical to the one in Imagen 7, showing the same SQL query and form. However, the execution time displayed at the bottom is 'Tiempo de Ejecución: 5206449 nanosegundos'.

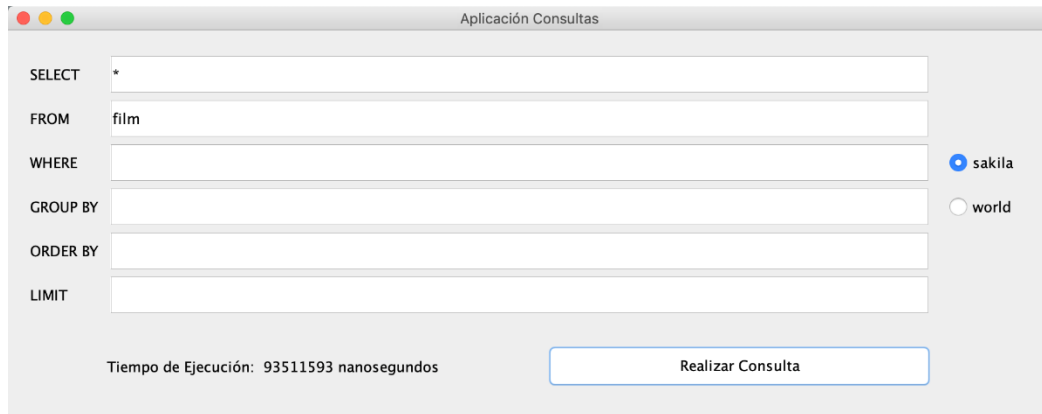
El tiempo de ejecución ha disminuido con respecto a la primera ejecución, pero al estar tratándose de nanosegundos son casi inapreciables a la hora del usuario ejecutar dicha consulta.

```
| Luvale | LUVALE | luvale |
| Mbundu | MBUNDU | mbundu |
| Nyaneka-nkhumbi | NYANEKA-NKHUMBI | nyaneka-nkhumbi |
| Ovimbundu | OVIMBUNDU | ovimbundu |
| English | ENGLISH | english |
| Albaniana | ALBANIANA | albaniana |
+-----+-----+-----+
20 rows in set (0.00 sec)
```

Consultas simples

Sakila

Imagen 9. Consulta simple 1 sobre Sakila.



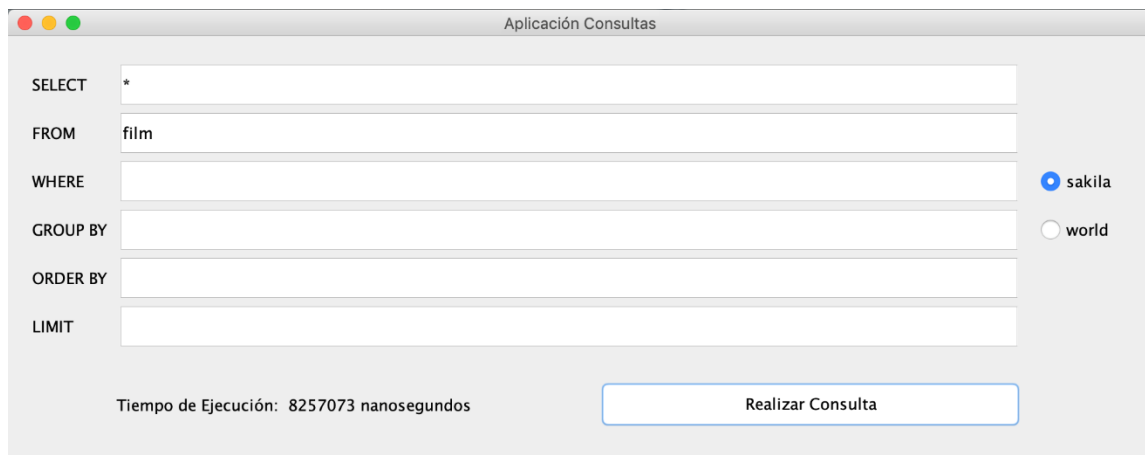
The screenshot shows a window titled "Aplicación Consultas". It contains a form for writing SQL queries with the following fields:

- SELECT:** *
- FROM:** film
- WHERE:** (empty)
- GROUP BY:** (empty)
- ORDER BY:** (empty)
- LIMIT:** (empty)

On the right side, there are two radio buttons: "sakila" (selected) and "world". At the bottom left, it says "Tiempo de Ejecución: 93511593 nanosegundos". At the bottom right, there is a button labeled "Realizar Consulta".

El tiempo de ejecución en segundos según la aplicación es 0.093511593. Sin embargo, hemos medido el tiempo de ejecución al ejecutar dicha consulta sobre Sakila por el terminal, y ha resultado un tiempo promedio de 1min 1seg (61 segundos) tras efectuar la consulta tres veces y promediar sus tiempos. Debido a que no se muestran los minutos sino solo los segundos al pasar cada minuto, se observa el resultado de 0.01 segundos. Este es un tiempo de ejecución un tanto elevado, debido a que muestra por pantalla todos los datos guardados en películas (siendo 1000 filas). Se puede concluir que cada fila se lee en 0.061 segundos.

Imagen 10. Segunda consulta simple 1 sobre Sakila



This screenshot is identical to the one in Image 9, showing the "Aplicación Consultas" window with the same SQL query: SELECT * FROM film. The "sakila" database is selected. The execution time displayed is "Tiempo de Ejecución: 8257073 nanosegundos". The "Realizar Consulta" button is also present.

El tiempo de ejecución vuelve a verse reducido con respecto a la primera ejecución pero nuevamente vuelve a ser muy despreciable el cambio para el usuario percibirlo.


```

| 2006-02-15 05:03:42 |
| 999 | ZOOLANDER FICTION | A Fateful Reflection of a Waitress And a Boat who must Discover a Sumo Wrestler in Ancient China
| 2006 | 1 | NULL | 5 | 2.99 | 101 | 28.99 | R | Trailers,Deleted Scenes
| 2006-02-15 05:03:42 |
| 1000 | ZORRO ARK | A Intrepid Panorama of a Mad Scientist And a Boy who must Redeem a Boy in A Monastery
| 2006 | 1 | NULL | 3 | 4.99 | 50 | 18.99 | NC-17 | Trailers,Commentaries,Behind the Sc
nes
| 2006-02-15 05:03:42 |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1000 rows in set (0.01 sec)

```

La siguiente consulta simple demuestra el id de la película, su título y duración mientras que la duración de renta sea mayor o igual a 4 y el límite de edad sea 13 años. Se ordena de manera descendiente.

Imagen 11. Consulta simple 2 sobre Sakila.

Aplicación Consultas

SELECT

film_id as "id", title as "título", length as "duración"

FROM

film

WHERE

rental_duration >= 4 and rating = "PG-13"

GROUP BY

ORDER BY

length desc

LIMIT

☒ sakila
☐ world

Tiempo de Ejecución: 3321777 nanosegundos

Realizar Consulta

El tiempo de ejecución en segundos es 0.003321777. Esta consulta se ha ejecutado rápidamente por ser una consulta simple, donde solo debe leer una tabla y mostrar unos datos específicos. Mientras más condiciones le agreguemos a la consulta, menos datos deberán devolverse y más rápido se ejecutará. En este caso, solo cumplían con las condiciones 184 películas, de modo que cada una se leía en 0.00001793 segundos.

Imagen 12. Segunda consulta simple 2 sobre Sakila

Aplicación Consultas

SELECT

film_id as "id", title as "título", length as "duración"

FROM

film

WHERE

rental_duration >= 4 and rating = "PG-13"

GROUP BY

ORDER BY

length desc

LIMIT

☒ sakila
☐ world

Tiempo de Ejecución: 2035854 nanosegundos

Realizar Consulta

El tiempo de ejecución vuelve a ser inferior pero no existe una reducción demasiado brusca sobre dicho tiempo de ejecución.

| | | |
|----------------------------|---------------------|----|
| 630 | NOTTING SPEAKEASY | 48 |
| 657 | PARADISE SABRINA | 48 |
| 931 | VALENTINE VANISHING | 48 |
| 393 | HALLOWEEN NUTS | 47 |
| 407 | HAWK CHILL | 47 |
| 504 | KWAI HOMEWARD | 46 |
| 505 | LABYRINTH LEAGUE | 46 |
| +-----+-----+-----+ | | |
| 184 rows in set (0.00 sec) | | |

World

Mediante esta consulta buscamos obtener todos los registros presentes en la tabla countrylanguage

Imagen 13. Consulta simple 1 sobre World.

The screenshot shows a window titled 'Aplicación Consultas'. It contains a form for writing SQL queries. The fields are: SELECT (containing '*'), FROM (containing 'countrylanguage'), WHERE (empty), GROUP BY (empty), ORDER BY (empty), and LIMIT (empty). To the right of the WHERE, GROUP BY, and ORDER BY fields are radio buttons for 'sakila' and 'world'. The 'world' radio button is selected. At the bottom left, it says 'Tiempo de Ejecución: 76870418 nanosegundos'. At the bottom right is a button labeled 'Realizar Consulta'.

El tiempo de ejecución obtenido ha sido de 0,076870418 segundos (76870418 nanosegundos). Esta es una consulta que debe imprimir todos los registros presentes de la tabla countrylanguage de World por lo que es previsible que se demore más que algunas más simples, aun así sigue siendo un valor bastante bajo de tiempo.

Imagen 14. Segunda consulta simple 1 sobre World

This screenshot is identical to the one in Image 13, showing the 'Aplicación Consultas' window with the same SQL query and settings. The execution time at the bottom left is now 'Tiempo de Ejecución: 11751209 nanosegundos'.

Al ejecutar dicha consulta por segunda vez hemos podido apreciar como el valor ha aumentado, algo que no había ocurrido hasta este ejemplo. Esto puede ocurrir por cambios de tareas en el procesador en el momento de ejecución, tardándose más o menos en una consulta.

```
| ZMB | Nsenga | F | 4.3 |
| ZMB | Nyanja | F | 7.8 |
| ZMB | Tongan | F | 11.0 |
| ZWE | English | T | 2.2 |
| ZWE | Ndebele | F | 16.2 |
| ZWE | Nyanja | F | 2.2 |
| ZWE | Shona | F | 72.1 |
+-----+-----+-----+-----+
984 rows in set (0.01 sec)
```

En esta consulta se busca obtener el CountryCode así como el language de la tabla countrylanguage siempre y cuando no sea oficial dicho idioma y se encuentre en un porcentaje inferior al 15%.

Imagen 15. Consulta simple 2 sobre World.

Aplicación Consultas

SELECT: CountryCode, Language

FROM: countrylanguage

WHERE: IsOfficial = "F" and Percentage > "15"

GROUP BY:

ORDER BY: Language

LIMIT:

radio sakila radio world (selected)

Tiempo de Ejecución: 1866152 nanosegundos

Realizar Consulta

El tiempo obtenido tras la ejecución de la consulta es de 1866152 nanosegundos, que convertido a segundos supone 0,001866152. Estos valores de tiempo son muy inferiores al segundo y en este caso bastante inferior al de la otra consulta ejecutada en la misma base de datos.

Imagen 16. Segunda consulta simple 2 sobre World

Aplicación Consultas

SELECT: CountryCode, Language

FROM: countrylanguage

WHERE: IsOfficial = "F" and Percentage > "15"

GROUP BY:

ORDER BY: Language

LIMIT:

radio sakila radio world (selected)

Tiempo de Ejecución: 1014335 nanosegundos

Realizar Consulta

El tiempo de ejecución nuevamente se ha reducido con respecto a la ejecución de la misma consulta.

```
TCD | Sara
THA | Lao
TJK | Uzbek
TTO | English
TUN | Arabic-French
TWN | Min
TZA | Nyamwesi
UGA | Ganda
UKR | Russian
VCT | Creole English
WSM | Samoan-English
YUG | Albaniana
ZMB | Bemba
ZWE | Ndebele
ZWE | Shona
+-----+
121 rows in set (0.09 sec)
```

Consultas complejas

Sakila

Imagen 17. Consulta compleja 1 sobre Sakila.

Aplicación Consultas

SELECT customer.first_name AS "Nombre", customer.last_name AS "Apellido", payment.amount AS "Cantidad pagada"

FROM customer JOIN payment ON customer.customer_id = payment.customer_id

WHERE (SELECT avg(payment.amount) FROM payment) < amount

GROUP BY

ORDER BY

LIMIT 5

☒ sakila ☐ world

Tiempo de Ejecución: 6659995 nanosegundos

Realizar Consulta

El tiempo de ejecución en segundos es 0.006659995. Aunque sea una consulta compleja, donde debe hacer un join entre dos tablas, filtrar sus contenidos y hasta hacer una subconsulta, es una consulta que se hace rápidamente. En el terminal, obtuvimos un resultado de ejecución promedio de 0.01 segundos, cercano al obtenido en la aplicación. La rapidez de la consulta se debe a que se limita la búsqueda a solo los 5 primeros resultados. Si este límite lo elimináramos, la consulta tomaría unos 14 segundos según hemos medido en el terminal y contaría con 7746 datos.

Imagen 18. Segunda consulta compleja 1 sobre Sakila

Aplicación Consultas

SELECT customer.first_name AS "Nombre", customer.last_name AS "Apellido", payment.amount AS "Cantidad pagada"

FROM customer JOIN payment ON customer.customer_id = payment.customer_id

WHERE (SELECT avg(payment.amount) FROM payment) < amount LIMIT 5

GROUP BY

ORDER BY

LIMIT

☒ sakila ☐ world

Tiempo de Ejecución: 8282760 nanosegundos

Realizar Consulta

Se puede apreciar cómo se ha producido un gran aumento de tiempo al ejecutar esta consulta compleja por segunda vez.

| | | |
|--------------------------|-------|------|
| MARY | SMITH | 9.99 |
| MARY | SMITH | 4.99 |
| MARY | SMITH | 4.99 |
| MARY | SMITH | 5.99 |
| +-----+-----+-----+ | | |
| 5 rows in set (0.01 sec) | | |

Imagen 19. Consulta compleja 2 sobre Sakila.

Aplicación Consultas

SELECT: rental.rental_date, film.title, customer.first_name, customer.last_name

FROM: ry_id = inventory.inventory_id AND inventory.film_id = film.film_id AND rental.customer_id = customer.customer_id

WHERE:

GROUP BY:

ORDER BY:

LIMIT:

Radio buttons: ☒ sakila, ☐ world

Tiempo de Ejecución: 82770419 nanosegundos

Realizar Consulta

El tiempo de ejecución obtenido es de 82770419 nanosegundos, 0,082770419 segundos. Esta es una de las consultas que tarda más tiempo, esto se debe y se puede explicar de manera que al ser una consulta compleja tiene que acceder a varias tablas y obtener de ellas varias columnas por lo que el acceso a los datos toma más tiempo y por consiguiente el tiempo de ejecución también se verá afectado.

Imagen 20. Segunda consulta compleja 2 sobre Sakila

Aplicación Consultas

SELECT: rental.rental_date, film.title, customer.first_name, customer.last_name

FROM: ry_id = inventory.inventory_id AND inventory.film_id = film.film_id AND rental.customer_id = customer.customer_id

WHERE:

GROUP BY:

ORDER BY:

LIMIT:

Radio buttons: ☒ sakila, ☐ world

Tiempo de Ejecución: 53757107 nanosegundos

Realizar Consulta

En este caso el tiempo de ejecución se ha reducido notablemente y podemos entender que es gracias a que algunos de los registros se han podido almacenar en memoria de manera que al ejecutarla nuevamente y seguida de la anterior ya los posee.

World

Imagen 21. Consulta compleja 1 sobre World.

Aplicación Consultas

SELECT Name, Capital, Population

FROM country

WHERE Code in (SELECT CountryCode FROM countrylanguage WHERE Language = "Spanish" and IsOfficial = "T")

GROUP BY

ORDER BY

LIMIT

☐ sakila ☒ world

Tiempo de Ejecución: 1857720 nanosegundos

Realizar Consulta

El tiempo de ejecución en segundos es 0.00185772. La consulta obtenida tiene como resultado 20 países en donde el lenguaje oficial es el español. Es una consulta compleja porque realiza una subconsulta, de manera que el tiempo de ejecución podría haber sido mucho más elevado si más países cumplieran con la condición. No hay variación significativa entre los tiempos de ejecución al repetir la consulta.

Imagen 22. Segunda consulta compleja 1 sobre World

Aplicación Consultas

SELECT Name, Capital, Population

FROM country

WHERE Code in (SELECT CountryCode FROM countrylanguage WHERE Language = "Spanish" and IsOfficial = "T")

GROUP BY

ORDER BY

LIMIT

☐ sakila ☒ world

Tiempo de Ejecución: 1780480 nanosegundos

Realizar Consulta

El tiempo de ejecución de esta consulta por segunda vez es muy similar al de la primera vez que se ejecuta dicha consulta.

```
| Paraguay | 2885 | 5496000 |
| El Salvador | 645 | 6276000 |
| Uruguay | 3492 | 3337000 |
| Venezuela | 3539 | 24170000 |
+-----+-----+-----+
20 rows in set (0.00 sec)
```

Imagen 23. Consulta compleja 2 sobre World.

Aplicación Consultas

SELECT: country.Name, country.Continent, country.Population , city.Name, city.Population, countrylanguage.Language

FROM: city JOIN countrylanguage ON country.Code = city.CountryCode AND country.Code = countrylanguage.CountryCode

WHERE: countrylanguage.IsOfficial = "T"

GROUP BY:

ORDER BY:

LIMIT:

Radio buttons: ☐ sakila, ☒ world

Tiempo de Ejecución: 14333730 nanosegundos

Realizar Consulta

El tiempo de ejecución de esta consulta es de 14333730 nanosegundos, 0,01433373 segundos, siendo unas de las que más tiempo han tardado y esto puede ser debido a que debe seleccionar una gran cantidad de registros de distintas columnas y además debe acceder a varias tablas distintas.

Imagen 23. Segunda consulta compleja 2 sobre World

Aplicación Consultas

SELECT: country.Name, country.Continent, country.Population , city.Name, city.Population, countrylanguage.Language

FROM: city JOIN countrylanguage ON country.Code = city.CountryCode AND country.Code = countrylanguage.CountryCode

WHERE: countrylanguage.IsOfficial = "T"

GROUP BY:

ORDER BY:

LIMIT:

Radio buttons: ☐ sakila, ☒ world

Tiempo de Ejecución: 9650265 nanosegundos

Realizar Consulta

CONCLUSIÓN

Tras realizar consultas de distintos tipos, sobre distintas bases y con distintas complejidades y posteriormente analizar los tiempos de ejecución devueltos, podemos concluir que a pesar de que todos ellos son muy pequeños y prácticamente incomparables e inapreciables por el usuario final a la hora de ejecutarlos, si que pueden variar ligeramente, aumentando según la complejidad de la consulta (número de tablas a las que acceder, número de columnas a imprimir o número de registros a tener en cuenta). Se puede además llegar a la conclusión que la variación de la duración de ejecución de una misma consulta ejecutada varias veces será a causa del procesador, ya que no siempre realizará la consulta con la misma rapidez.

De esta manera se puede concluir que las consultas más rápidas serán aquellas que deban mostrar por pantalla la menor cantidad de datos, y acceder a la menor cantidad de tablas para formular un resultado. A veces las consultas complejas (con JOINS o subconsultas) serán más rápidas que una simple donde se haga un select *, ya que en esta se deben imprimir todos los datos contenidos en la tabla definida, sin filtrarlos. Mientras más condiciones se le agreguen que reduzcan los resultados a imprimir, menos durará la ejecución porque menos datos deberá leer.