

## Bases de Datos

# Laboratorio 7

Acceso Programático

Integrantes: Jean Cherubini  
Profesores: Aidan Hogan  
Auxiliares: Sebastián Ferrada A.  
Ayudantes: Henry Rosales M.  
Marco Caballero G.  
Pablo Miranda Á.

Fecha de entrega: 10 de julio de 2018  
Santiago, Chile

## P1

Se utilizaron las consultas mostradas en el código de abajo para obtener la cantidad de tuplas por dataset. Lo obtenido se muestra en la Tabla 1. Se observa además al analizar los índices de la tabla **Lab7\_index**. Dichos índices tienen un formato *btree*, es decir, como un árbol binario.

```
SELECT COUNT(*) FROM lab7.actor100;
SELECT COUNT(*) FROM lab7.actor1k;
SELECT COUNT(*) FROM lab7.actor10k;
SELECT COUNT(*) FROM lab7.pelicula100;
SELECT COUNT(*) FROM lab7.pelicula1k;
SELECT COUNT(*) FROM lab7.pelicula10k;
SELECT COUNT(*) FROM lab7.personaje100;
SELECT COUNT(*) FROM lab7.personaje1k;
SELECT COUNT(*) FROM lab7.personaje10k;
```

Tabla 1: Cantidad de Tuplas por Dataset

X	100 Votos	1k Votos	10K Votos
actor	856421	440234	197219
pelicula	72696	22490	6401
personaje	2170454	944936	372367

## P2

Las consultas utilizadas para obtener los resultados pedidos corresponden al código que se muestra debajo, en conjunto con su respectiva planificación. Los resultados de los tiempos de ejecución se muestran en la Tabla 2 y en el grafico 1.

Tabla 2: Tiempos de ejecución para consultas en distintas tablas.

X	Lab7 [ms]	Lab7_index [ms]
personaje100	413.661	1.031
personaje1k	178.434	0.548
personaje10k	69.533	0.480

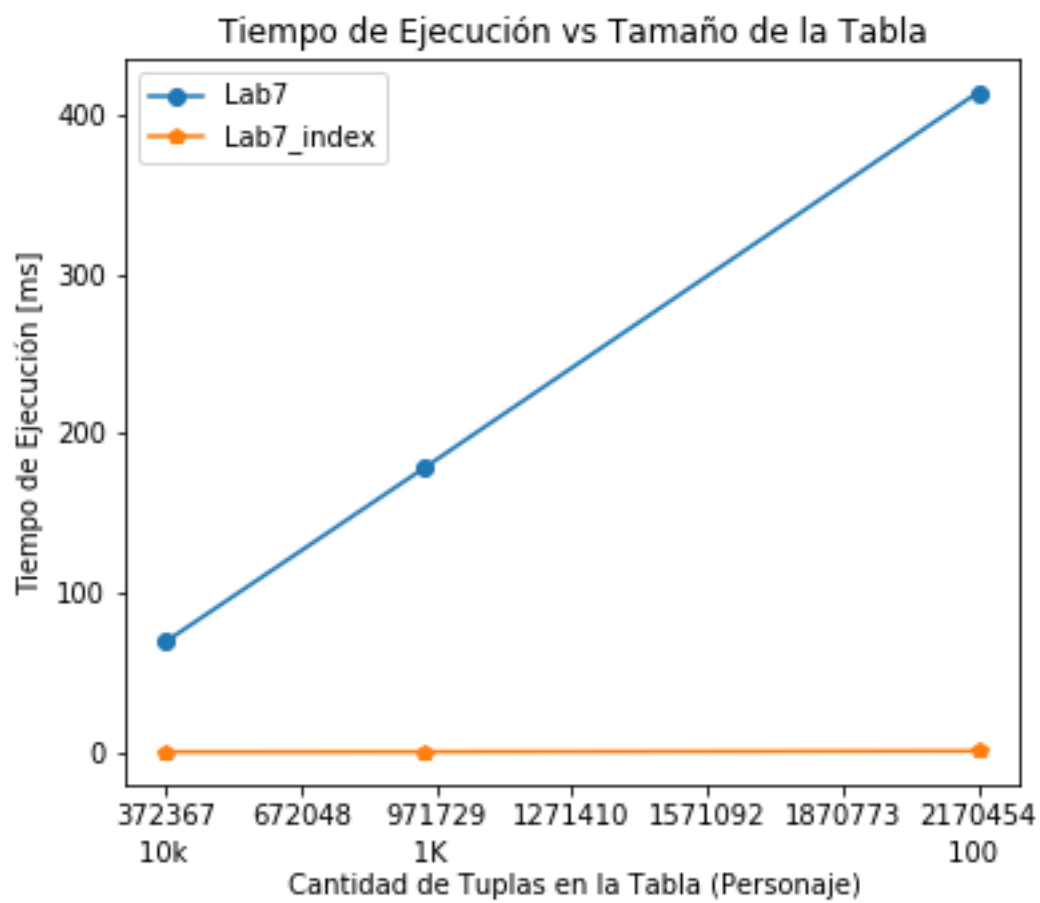


Figura 1: Tiempos de ejecución para consultas en distintas tablas.

```

EXPLAIN ANALYZE SELECT *
FROM lab7.personaje100
WHERE p_nombre='Pulp Fiction';
Query Plan:
    Seq Scan on personaje100
    (cost=0.00..49359.68 rows=46 width=49)
    (actual time=21.638..413.587 rows=58 loops=1)
        Filter: ((p_nombre)::text = 'Pulp Fiction'::text)
        Rows Removed by Filter: 2170396
    Planning time: 0.642 ms
    Execution time: 413.661 ms)

EXPLAIN ANALYZE SELECT *
FROM lab7_index.personaje100
WHERE p_nombre='Pulp Fiction';
Query Plan:
    Bitmap Heap Scan on personaje100
    (cost=4.78..179.27 rows=45 width=49)
    (actual time=0.467..0.972 rows=58 loops=1)
        Recheck Cond: ((p_nombre)::text = 'Pulp Fiction'::text)
        Heap Blocks: exact=53
        -> Bitmap Index Scan on personaje100_pnombreanho
        (cost=0.00..4.77 rows=45 width=0) (actual time=0.445..0.445
        rows=58 loops=1)
            Index Cond: ((p_nombre)::text = 'Pulp Fiction'::text)
    Planning time: 0.395 ms
    Execution time: 1.031 ms)

EXPLAIN ANALYZE SELECT *
FROM lab7.personaje1k
WHERE p_nombre='Pulp Fiction';
Query Plan:
    Seq Scan on personaje1k
    (cost=0.00..21495.04 rows=51 width=49)
    (actual time=1.881..178.360 rows=58 loops=1)
        Filter: ((p_nombre)::text = 'Pulp
        Fiction'::text)
        Rows Removed by Filter: 944905
    Planning time: 0.359 ms
    Execution time: 178.434 ms

```

```

EXPLAIN ANALYZE SELECT *
FROM lab7_index.personaje1k
WHERE p_nombre='Pulp Fiction';
Query Plan:
  Bitmap Heap Scan on personaje1k
    (cost=4.82..198.35 rows=51 width=48)
    (actual time=0.096..0.478 rows=58 loops=1)
      Recheck Cond: ((p_nombre)::text = 'Pulp Fiction'::text)
      Heap Blocks: exact=53
      -> Bitmap Index Scan on personaje1k_pnombreanho
        (cost=0.00..4.81 rows=51 width=0) (actual time=0.073..0.073
        rows=58 loops=1)
          Index Cond: ((p_nombre)::text = 'Pulp Fiction'::text)
    Planning time: 0.444 ms
    Execution time: 0.548 ms

```

```

EXPLAIN ANALYZE SELECT *
FROM lab7.personaje10k
WHERE p_nombre='Pulp Fiction';
Query Plan:
  Seq Scan on personaje10k
    (cost=0.00..8477.59 rows=58 width=49)
    (actual time=0.419..69.469 rows=58 loops=1)
      Filter: ((p_nombre)::text = 'Pulp Fiction'::text)
      Rows Removed by Filter: 372309
    Planning time: 0.282 ms
    Execution time: 69.533 ms

```

```

EXPLAIN ANALYZE SELECT *
FROM lab7_index.personaje10k
WHERE p_nombre='Pulp Fiction';
Query Plan:
  Bitmap Heap Scan on personaje10k
    (cost=4.87..216.17 rows=58 width=49)
    (actual time=0.101..0.407 rows=58 loops=1)
      Recheck Cond: ((p_nombre)::text = 'Pulp Fiction'::text)
      Heap Blocks: exact=52
      -> Bitmap Index Scan on personaje10k_pnombreanho
        (cost=0.00..4.86 rows=58 width=0)
        (actual time=0.077..0.077 rows=58 loops=1)
          Index Cond: ((p_nombre)::text = 'Pulp Fiction'::text)
    Planning time: 0.535 ms
    Execution time: 0.480 ms

```

## P3

Se plantearon 3 consultas según lo pedido. Cada una fue ejecutada en cada tabla pedida.

### Consulta de Rango

Se utilizaron las consultas mostradas debajo en conjunto con sus planificaciones. Los tiempos de ejecución resultantes se muestran en la tabla 3 y la figura 2.

Tabla 3: Tiempos de ejecución para consulta de Rango.

X	Lab7 [ms]	Lab7_index [ms]
rango100	15.565	0.044
rango1k	5.791	0.067
rango10k	1.435	0.054

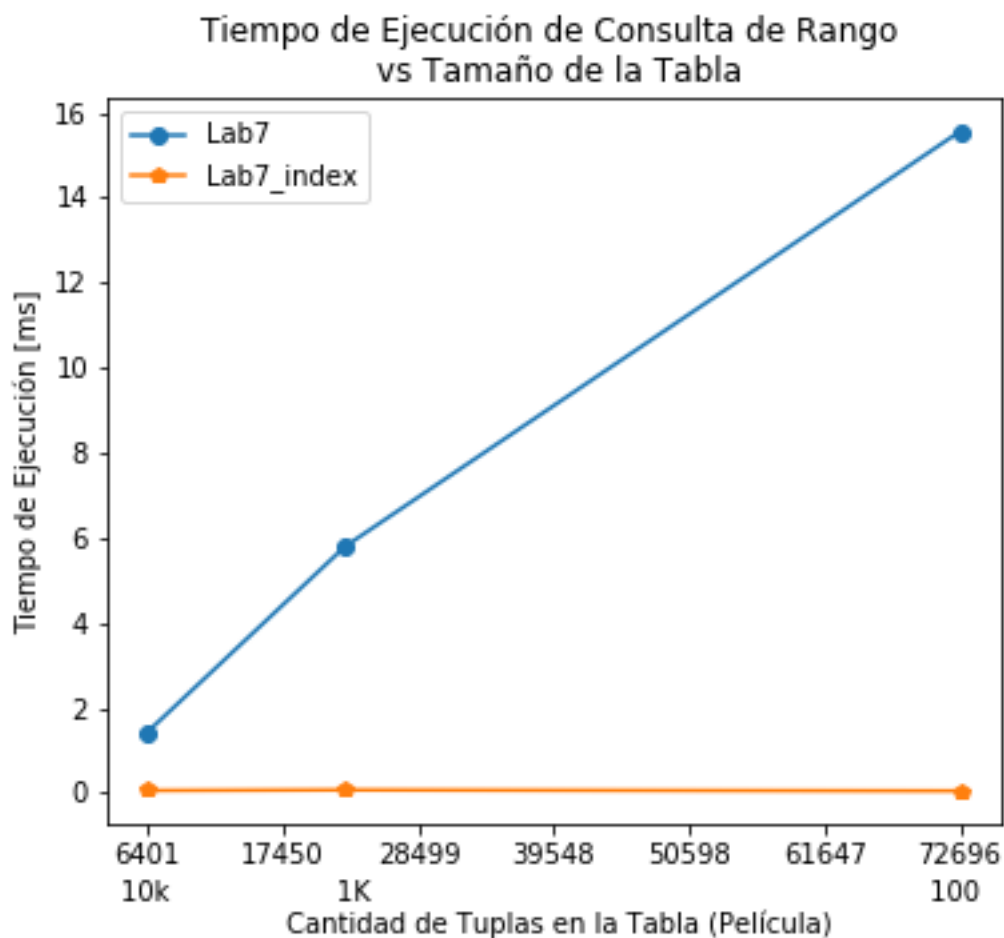


Figura 2: Tiempos de ejecución para consulta de Rango

```
EXPLAIN ANALYZE SELECT *
FROM lab7.pelicula100
WHERE calificacion >7
AND calificacion < 7.1;
```

Query Plan:

```
Seq Scan on pelicula100
(cost=0.00..1706.44 rows=5 width=33)
(actual time=15.542..15.542 rows=0 loops=1)
  Filter: ((calificacion > '7'::double precision)
    AND (calificacion < '7.1'::double precision))
  Rows Removed by Filter: 72696
Planning time: 0.091 ms
Execution time: 15.565 ms
```

```
EXPLAIN ANALYZE SELECT *
FROM lab7_index.pelicula100
WHERE calificacion >7
AND calificacion < 7.1;
```

Query Plan:

```
Bitmap Heap Scan on pelicula100
(cost=4.35..26.67 rows=6 width=33)
(actual time=0.018..0.018 rows=0 loops=1)
  Recheck Cond: ((calificacion > '7'::double precision)
    AND (calificacion < '7.1'::double precision))
  -> Bitmap Index Scan on pelicula100_calificacion
    (cost=0.00..4.35 rows=6 width=0)
    (actual time=0.015..0.015 rows=0 loops=1)
    Index Cond: ((calificacion > '7'::double precision)
      AND (calificacion < '7.1'::double precision))
Planning time: 0.125 ms
Execution time: 0.044 ms
```

```
EXPLAIN ANALYZE SELECT *
FROM lab7.pelicula1k
WHERE calificacion >7
AND calificacion < 7.1;
```

Query Plan:

```
Seq Scan on pelicula1k
(cost=0.00..526.35 rows=2 width=32)
(actual time=5.772..5.772 rows=0 loops=1)
  Filter: ((calificacion > '7'::double precision) AND
    (calificacion < '7.1'::double precision))
  Rows Removed by Filter: 22490
Planning time: 0.091 ms
Execution time: 5.791 ms
```

```
EXPLAIN ANALYZE SELECT *
FROM lab7_index.pelicula1k
WHERE calificacion >7
AND calificacion < 7.1;
```

Query Plan:

```
Bitmap Heap Scan on pelicula1k
(cost=4.31..11.72 rows=2 width=32)
(actual time=0.043..0.043 rows=0 loops=1)
  Recheck Cond: ((calificacion > '7'::double precision)
AND (calificacion < '7.1'::double precision))
  -> Bitmap Index Scan on pelicula1k_calificacion
      (cost=0.00..4.31 rows=2 width=0)
      (actual time=0.039..0.039 rows=0 loops=1)
      Index Cond: ((calificacion > '7'::double precision)
AND (calificacion < '7.1'::double precision))
Planning time: 0.121 ms
Execution time: 0.067 ms
```

```
EXPLAIN ANALYZE SELECT *
FROM lab7.pelicula10k
WHERE calificacion >7
AND calificacion < 7.1;
```

Query Plan:

```
Seq Scan on pelicula10k
(cost=0.00..150.01 rows=1 width=32)
(actual time=1.415..1.415 rows=0 loops=1)
  Filter: ((calificacion > '7'::double precision) AND
(calificacion < '7.1'::double precision))
  Rows Removed by Filter: 6401
Planning time: 0.089 ms
Execution time: 1.435 ms
```

```
EXPLAIN ANALYZE SELECT *
FROM lab7_index.pelicula10k
WHERE calificacion >7
AND calificacion < 7.1;
```

Query Plan:

```
Index Scan using pelicula10k_calificacion on pelicula10k
(cost=0.28..8.30 rows=1 width=32)
(actual time=0.029..0.029 rows=0 loops=1)
  Index Cond: ((calificacion > '7'::double precision)
AND (calificacion < '7.1'::double precision))
Planning time: 64.612 ms
Execution time: 0.054 ms
```



## Consulta con Join

Se realizaron las consultas mostradas abajo. Los tiempos de ejecución obtenidos se muestran en la Tabla 4 y en la Figura 3.

Tabla 4: Tiempos de ejecución para consulta con Join.

X	Lab7 [ms]	Lab7_index [ms]
join100	15.565	0.044
join1k	5.791	0.067
join10k	1.435	0.054

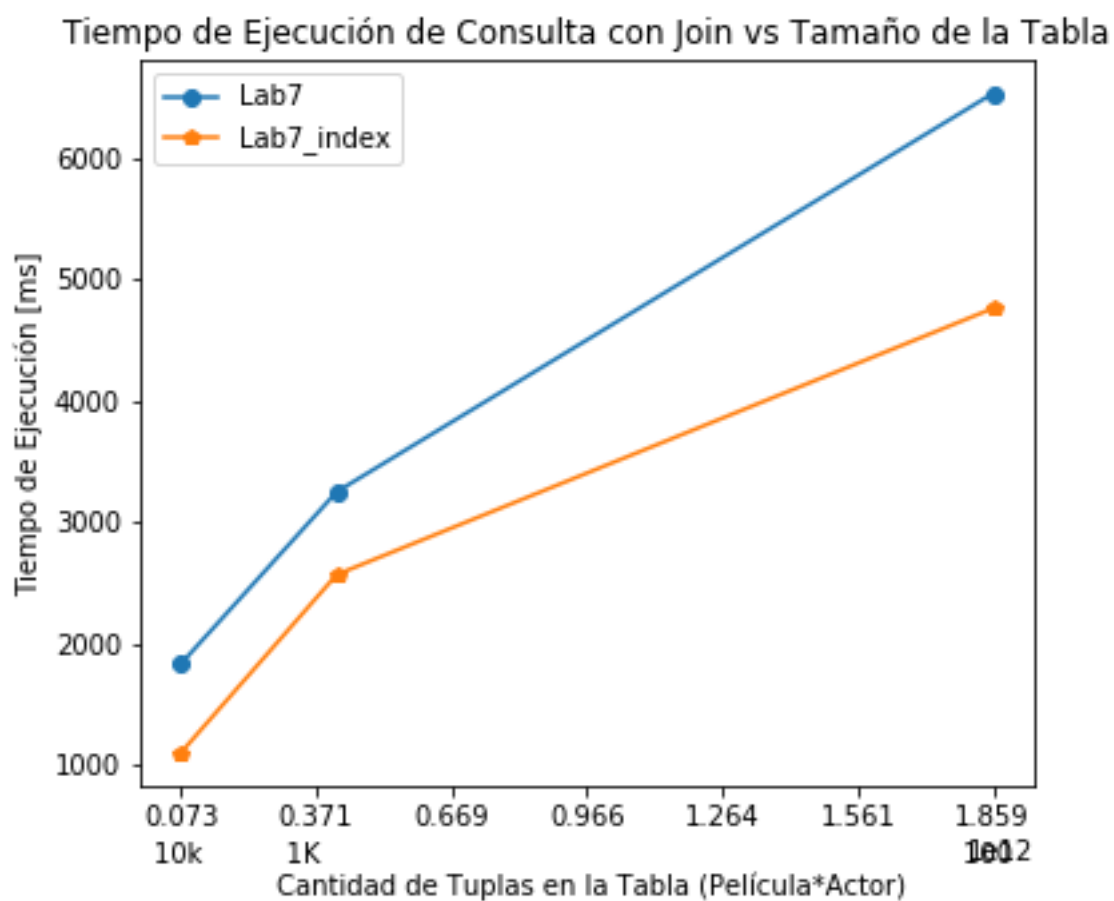


Figura 3: Tiempos de ejecución para consulta de Rango

```
EXPLAIN ANALYZE SELECT a.nombre, p.personaje
FROM lab7.actor100 a, lab7.personaje100 p
WHERE a.genero='F'
AND p.a_nombre=a.nombre;
```

Query Plan:

```
Hash Join (cost=21339.66..112401.99 rows=780858 width=28)
(actual time=2631.231..6172.130 rows=673818 loops=1)
  Hash Cond: ((p.a_nombre)::text = (a.nombre)::text)
  -> Seq Scan on personaje100 p
    (cost=0.00..43933.54 rows=2170454 width=28)
    (actual time=0.034..2889.811 rows=2170454 loops=1)
  -> Hash (cost=15983.26..15983.26 rows=308112 width=16)
    (actual time=570.448..570.448 rows=306428 loops=1)
    Buckets: 131072 Batches: 8 Memory Usage: 2850kB
  -> Seq Scan on actor100 a
    (cost=0.00..15983.26 rows=308112 width=16)
    (actual time=0.024..332.079 rows=306428 loops=1)
  Filter: (genero = 'F'::bpchar)
  Rows Removed by Filter: 549993
Planning time: 0.428 ms
Execution time: 6532.759 ms
```

```
EXPLAIN ANALYZE SELECT a.nombre, p.personaje
FROM lab7_index.actor100 a, lab7_index.personaje100 p
WHERE a.genero='F'
AND p.a_nombre=a.nombre;
```

Query Plan:

```
Hash Join (cost=15930.26..106929.77 rows=775576 width=28)
(actual time=471.934..4443.332 rows=673818 loops=1)
  Hash Cond: ((p.a_nombre)::text = (a.nombre)::text)
  -> Seq Scan on personaje100 p
    (cost=0.00..43933.54 rows=2170454 width=28)
    (actual time=0.036..1430.975 rows=2170454 loops=1)
  -> Hash (cost=10609.91..10609.91 rows=306028 width=16)
    (actual time=471.678..471.678 rows=306428 loops=1)
    Buckets: 131072 Batches: 8 Memory Usage: 2850kB
  -> Index Scan using actor100_genero on actor100 a
    (cost=0.42..10609.91 rows=306028 width=16)
    (actual time=0.050..228.449 rows=306428 loops=1)
  Index Cond: (genero = 'F'::bpchar)
Planning time: 0.699 ms
Execution time: 4763.284 ms
```

```
EXPLAIN ANALYZE SELECT a.nombre, p.personaje
FROM lab7.actor1k a, lab7.personaje1k p
WHERE a.genero='F'
AND p.a_nombre=a.nombre;
```

Query Plan:

```
Hash Join (cost=10910.88..50592.97 rows=332785 width=29)
(actual time=769.346..3072.596 rows=290411 loops=1)
  Hash Cond: ((p.a_nombre)::text = (a.nombre)::text)
  -> Seq Scan on personaje1k p
      (cost=0.00..19132.63 rows=944963 width=29)
      (actual time=0.016..834.442 rows=944963 loops=1)
  -> Hash (cost=8214.92..8214.92 rows=155036 width=16)
      (actual time=768.620..768.620 rows=153380 loops=1)
      Buckets: 131072 Batches: 4 Memory Usage: 2839kB
      -> Seq Scan on actor1k a
          (cost=0.00..8214.92 rows=155036 width=16)
          (actual time=0.025..627.444 rows=153380 loops=1)
      Filter: (genero = 'F'::bpchar)
      Rows Removed by Filter: 286854
Planning time: 96.623 ms
Execution time: 3252.081 ms
```

```
EXPLAIN ANALYZE SELECT a.nombre, p.personaje
FROM lab7_index.actor1k a, lab7_index.personaje1k p
WHERE a.genero='F'
AND p.a_nombre=a.nombre;
```

Query Plan:

```
Hash Join (cost=8076.50..47758.59 rows=332785 width=28)
(actual time=311.582..2420.140 rows=290411 loops=1)
  Hash Cond: ((p.a_nombre)::text = (a.nombre)::text)
  -> Seq Scan on personaje1k p
      (cost=0.00..19132.63 rows=944963 width=28)
      (actual time=0.096..802.462 rows=944963 loops=1)
  -> Hash (cost=5380.55..5380.55 rows=155036 width=16)
      (actual time=311.269..311.269 rows=153380 loops=1)
      Buckets: 131072 Batches: 4 Memory Usage: 2839kB
      -> Index Scan using actor1k_genero on actor1k a
          (cost=0.42..5380.55 rows=155036 width=16)
          (actual time=0.033..154.981 rows=153380 loops=1)
      Index Cond: (genero = 'F'::bpchar)
Planning time: 0.530 ms
Execution time: 2571.060 ms
```

```
EXPLAIN ANALYZE SELECT a.nombre, p.personaje
FROM lab7.actor10k a, lab7.personaje10k p
WHERE a.genero='F'
AND p.a_nombre=a.nombre;
```

Query Plan:

```
Hash Join (cost=4853.93..20493.34 rows=127437 width=29)
(actual time=1145.048..1779.777 rows=114210 loops=1)
  Hash Cond: ((p.a_nombre)::text = (a.nombre)::text)
  -> Seq Scan on personaje10k p
      (cost=0.00..7546.67 rows=372367 width=29)
      (actual time=0.020..244.233 rows=372367 loops=1)
  -> Hash (cost=3680.24..3680.24 rows=67495 width=16)
      (actual time=1144.426..1144.426 rows=66928 loops=1)
      Buckets: 131072 Batches: 2 Memory Usage: 2621kB
      -> Seq Scan on actor10k a
          (cost=0.00..3680.24 rows=67495 width=16)
          (actual time=0.023..1089.020 rows=66928 loops=1)
      Filter: (genero = 'F'::bpchar)
      Rows Removed by Filter: 130291
Planning time: 99.502 ms
Execution time: 1836.043 ms
```

```
EXPLAIN ANALYZE SELECT a.nombre, p.personaje
FROM lab7_index.actor10k a, lab7_index.personaje10k p
WHERE a.genero='F'
AND p.a_nombre=a.nombre;
```

Query Plan:

```
Hash Join (cost=3520.26..19160.30 rows=127499 width=29)
(actual time=181.707..1034.435 rows=114210 loops=1)
  Hash Cond: ((p.a_nombre)::text = (a.nombre)::text)
  -> Seq Scan on personaje10k p
      (cost=0.00..7546.67 rows=372367 width=29)
      (actual time=0.016..324.834 rows=372367 loops=1)
  -> Hash (cost=2346.16..2346.16 rows=67528 width=16)
      (actual time=180.966..180.966 rows=66928 loops=1)
      Buckets: 131072 Batches: 2 Memory Usage: 2621kB
      -> Index Scan using actor10k_genero on actor10k a
          (cost=0.42..2346.16 rows=67528 width=16)
          (actual time=15.063..127.974 rows=66928 loops=1)
      Index Cond: (genero = 'F'::bpchar)
Planning time: 299.041 ms
Execution time: 1104.358 ms
```

## Consulta Anidada

Se realizaron las consultas mostradas abajo en conjunto con sus planificaciones respectivas. En la Tabla 5 y la Figure 4 se pueden observar los tiempos de ejecución obtenidos.

Tabla 5: Tiempos de ejecución para consulta con consulta anidada.

X	Lab7 [ms]	Lab7_index [ms]
nested100	111.273	103.558
nested1k	33.047	32.726
nested10k	9.426	9.925

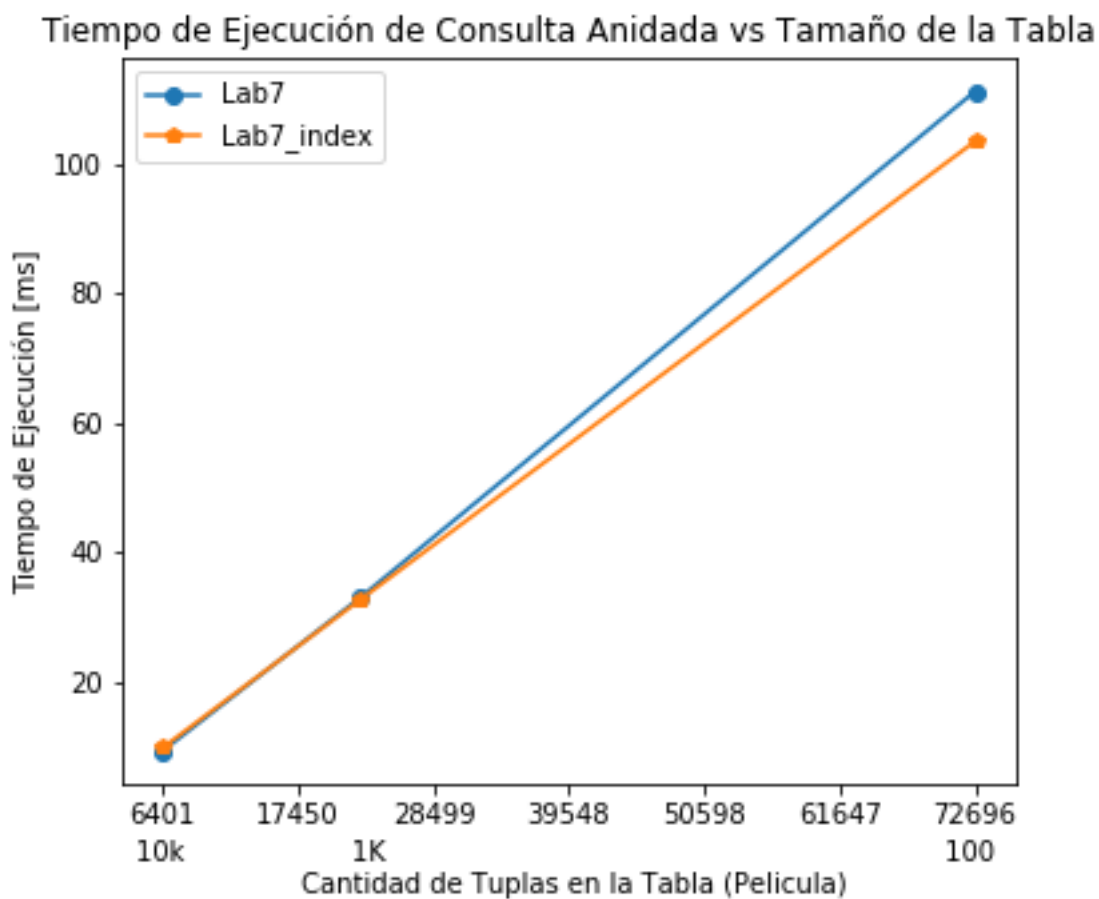


Figura 4: Tiempos de ejecución para consulta con consulta anidada

```

EXPLAIN ANALYZE SELECT anho
FROM (SELECT anho,avg(calificacion) AS promedio
      FROM lab7.pelicula100
      GROUP BY anho
      HAVING EVERY(calificacion>7.0)) x
ORDER BY x.promedio DESC;
Query Plan:
  Sort (cost=2077.09..2077.40 rows=125 width=12)
    (actual time=111.189..111.190 rows=1 loops=1)
    Sort Key: x.promedio DESC
    Sort Method: quicksort  Memory: 25kB
    -> Subquery Scan on x
      (cost=2069.92..2072.73 rows=125 width=12)
      (actual time=111.148..111.161 rows=1 loops=1)
      -> HashAggregate (cost=2069.92..2071.48 rows=125 width=12)
        (actual time=111.146..111.158 rows=1 loops=1)
        Group Key: pelicula100.anho
        Filter: every((pelicula100.calificacion > '7'::double precision))
        Rows Removed by Filter: 128
        -> Seq Scan on pelicula100 (cost=0.00..1342.96 rows=72696 width=12)
          (actual time=0.017..40.815 rows=72696 loops=1)
  Planning time: 0.153 ms
  Execution time: 111.273 ms

EXPLAIN ANALYZE SELECT anho
FROM (SELECT anho,avg(calificacion) AS promedio
      FROM lab7_index.pelicula100
      GROUP BY anho
      HAVING EVERY(calificacion>7.0)) x
ORDER BY x.promedio DESC;
Query Plan:
  Sort (cost=2077.15..2077.47 rows=126 width=12)
    (actual time=103.485..103.486 rows=1 loops=1)
    Sort Key: x.promedio DESC
    Sort Method: quicksort  Memory: 25kB
    -> Subquery Scan on x
      (cost=2069.92..2072.76 rows=126 width=12)
      (actual time=103.459..103.472 rows=1 loops=1)
      -> HashAggregate (cost=2069.92..2071.50 rows=126 width=12)
        (actual time=103.457..103.469 rows=1 loops=1)
        Group Key: pelicula100.anho
        Filter: every((pelicula100.calificacion > '7'::double precision))
        Rows Removed by Filter: 128
        -> Seq Scan on pelicula100 (cost=0.00..1342.96 rows=72696 width=12)
          (actual time=0.018..39.573 rows=72696 loops=1)
  Planning time: 0.163 ms
  Execution time: 103.558 ms

```

```

EXPLAIN ANALYZE SELECT anho
FROM (SELECT anho,avg(calificacion) AS promedio
      FROM lab7.pelicula1k
      GROUP BY anho
      HAVING EVERY(calificacion>7.0)) x
ORDER BY x.promedio DESC;
Query Plan:
  Sort (cost=645.71..646.01 rows=121 width=12)
    (actual time=32.954..32.955 rows=9 loops=1)
    Sort Key: x.promedio DESC
    Sort Method: quicksort  Memory: 25kB
    -> Subquery Scan on x (cost=638.80..641.52 rows=121 width=12)
      (actual time=32.912..32.937 rows=9 loops=1)
      -> HashAggregate (cost=638.80..640.31 rows=121 width=12)
        (actual time=32.911..32.929 rows=9 loops=1)
        Group Key: pelicula1k.anho
        Filter: every((pelicula1k.calificacion > '7'::double precision))
        Rows Removed by Filter: 112
      -> Seq Scan on pelicula1k (cost=0.00..413.90 rows=22490 width=12)
        (actual time=0.015..12.433 rows=22490 loops=1)
Planning time: 0.113 ms
Execution time: 33.047 ms

```

```

EXPLAIN ANALYZE SELECT anho
FROM (SELECT anho,avg(calificacion) AS promedio
      FROM lab7_index.pelicula1k
      GROUP BY anho
      HAVING EVERY(calificacion>7.0)) x
ORDER BY x.promedio DESC;
Query Plan:
  Sort (cost=645.71..646.01 rows=121 width=12) (actual time=32.661..32.662 rows=9 loops=1)
    Sort Key: x.promedio DESC
    Sort Method: quicksort  Memory: 25kB
    -> Subquery Scan on x (cost=638.80..641.52 rows=121 width=12)
      (actual time=32.618..32.645 rows=9 loops=1)
      -> HashAggregate (cost=638.80..640.31 rows=121 width=12)
        (actual time=32.617..32.631 rows=9 loops=1)
        Group Key: pelicula1k.anho
        Filter: every((pelicula1k.calificacion > '7'::double precision))
        Rows Removed by Filter: 112
      -> Seq Scan on pelicula1k (cost=0.00..413.90 rows=22490 width=12)
        (actual time=0.014..12.550 rows=22490 loops=1)
Planning time: 0.135 ms
Execution time: 32.726 ms

```

```

EXPLAIN ANALYZE SELECT anho
FROM (SELECT anho,avg(calificacion) AS promedio
      FROM lab7.pelicula10k
      GROUP BY anho
      HAVING EVERY(calificacion>7.0)) x
ORDER BY x.promedio DESC;
Query Plan:
  Sort (cost=187.65..187.91 rows=101 width=12)
    (actual time=9.301..9.302 rows=38 loops=1)
      Sort Key: x.promedio DESC
      Sort Method: quicksort Memory: 26kB
      -> Subquery Scan on x (cost=182.02..184.29 rows=101 width=12)
        (actual time=9.170..9.254 rows=38 loops=1)
          -> HashAggregate (cost=182.02..183.28 rows=101 width=12)
            (actual time=9.169..9.207 rows=38 loops=1)
              Group Key: pelicula10k.anho
              Filter: every((pelicula10k.calificacion > '7'::double precision))
              Rows Removed by Filter: 63
            -> Seq Scan on pelicula10k (cost=0.00..118.01 rows=6401 width=12)
              (actual time=0.021..3.596 rows=6401 loops=1)
  Planning time: 0.154 ms
  Execution time: 9.426 ms

```

```

EXPLAIN ANALYZE SELECT anho
FROM (SELECT anho,avg(calificacion) AS promedio
      FROM lab7_index.pelicula10k
      GROUP BY anho
      HAVING EVERY(calificacion>7.0)) x
ORDER BY x.promedio DESC;
Query Plan:
  Sort (cost=187.65..187.91 rows=101 width=12)
    (actual time=9.798..9.836 rows=38 loops=1)
      Sort Key: x.promedio DESC
      Sort Method: quicksort Memory: 26kB
      -> Subquery Scan on x (cost=182.02..184.29 rows=101 width=12)
        (actual time=9.687..9.757 rows=38 loops=1)
          -> HashAggregate (cost=182.02..183.28 rows=101 width=12)
            (actual time=9.685..9.715 rows=38 loops=1)
              Group Key: pelicula10k.anho
              Filter: every((pelicula10k.calificacion > '7'::double precision))
              Rows Removed by Filter: 63
            -> Seq Scan on pelicula10k (cost=0.00..118.01 rows=6401 width=12)
              (actual time=0.022..3.665 rows=6401 loops=1)
  Planning time: 0.229 ms
  Execution time: 9.925 ms

```