

Міністерство освіти і науки України
Національний технічний університет України «Київський політехнічний
інститут імені Ігоря Сікорського»
Факультет інформатики та обчислювальної техніки

Кафедра інформатики та програмної інженерії

Звіт

з лабораторної роботи №1 з дисципліни
«Програмування інтелектуальних інформаційних систем»

„Стовпчикові бази даних”

Виконав(ла) _____

ІП-11 Дякунчак Ілона Віталіївна
(шифр, прізвище, ім'я, по батькові)

Перевірів _____

Баришич Л. М.
(прізвище, ім'я, по батькові)

Київ 2023

1 РОЗРАХУВАТИ СУМАРНУ ЗАТРИМКУ ПО МІСТАХ

Код SQL для звичайної БД:

```
SELECT a.city, SUM(f.arr_delay) AS total_delay
FROM innodb_bts.flights f
JOIN innodb_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city
ASC LIMIT 5;
```

Код SQL для стовпчикової БД:

```
SELECT a.city, SUM(f.arr_delay) AS total_delay
FROM columnstore_bts.flights f
JOIN columnstore_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city
ASC LIMIT 5;
```

Скріншот виконання:

```
-> GROUP BY a.city
-> ASC LIMIT 5;
+-----+-----+
| city      | total_delay |
+-----+-----+
| Aberdeen  | -525.00     |
| Abilene    | 5952.00     |
| Adak       | -293.00     |
| Agana      | 994.00      |
| Aguadilla  | 2634.00     |
+-----+-----+
5 rows in set (7.564 sec)

MariaDB [(none)]> SELECT a.city, SUM(f.arr_delay) AS total_delay
-> FROM columnstore_bts.flights f
-> JOIN columnstore_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city
-> ASC LIMIT 5;
+-----+-----+
| city          | total_delay |
+-----+-----+
| West Palm Beach | 46876.00    |
| Raleigh        | 39343.00    |
| Barrow         | 1197.00     |
| Ketchikan      | 1552.00     |
| Knoxville      | 20697.00    |
+-----+-----+
5 rows in set (0.528 sec)

MariaDB [(none)]>
```

Швидкість виконання на стовпчиковій базі даних: 0,528 секунди

Швидкість виконання на звичайній базі даних: 7,564 секунди

2 ПОРАХУВАТИ КІЛЬКІСТЬ ПОЛЬОТІВ ПО МІСТАХ

Код SQL для звичайної БД:

```
SELECT a.city, COUNT(*) AS total_flights
FROM innodb_bts.flights f
JOIN innodb_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city LIMIT 5;
```

Код SQL для стовпчикової БД:

```
SELECT a.city, COUNT(*) AS total_flights
FROM columnstore_bts.flights f
JOIN columnstore_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city LIMIT 5;
```

Скріншот виконання:

```
-> GROUP BY a.city
-> LIMIT 5;

+-----+-----+
| city      | total_flights |
+-----+-----+
| Aberdeen  |          121 |
| Abilene   |          568 |
| Adak      |           21 |
| Agana     |           65 |
| Aguadilla |          224 |
+-----+-----+
5 rows in set (1.601 sec)
```

```
MariaDB [(none)]> SELECT a.city, COUNT(*) AS total_flights
-> FROM columnstore_bts.flights f
-> JOIN columnstore_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city
-> LIMIT 5;

+-----+-----+
| city              | total_flights |
+-----+-----+
| West Palm Beach  |          5431 |
| Raleigh          |          5947 |
| Barrow           |           176 |
| Ketchikan        |           416 |
| Knoxville        |          1410 |
+-----+-----+
5 rows in set (0.352 sec)
```

Швидкість виконання на стовпчиковій базі даних: 0,352 секунди

Швидкість виконання на звичайній базі даних: 1,601 секунди

3 ЗНАЙТИ МІСТО З НАЙМЕНШОЮ І НАЙБІЛЬШОЮ ЗАТРИМКОЮ

Код SQL для звичайної БД:

```
SELECT a.city, MIN(f.arr_delay) as min_delay
FROM innodb_bts.flights f
JOIN innodb_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city ORDER BY min_delay LIMIT 1;

SELECT a.city, MAX(f.arr_delay) as max_delay
FROM innodb_bts.flights f
JOIN innodb_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city ORDER BY max_delay DESC LIMIT 1;
```

Код SQL для стовпчикової БД:

```
SELECT a.city, MIN(f.arr_delay) as min_delay
FROM columnstore_bts.flights f
JOIN columnstore_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city ORDER BY min_delay LIMIT 1;

SELECT a.city, MAX(f.arr_delay) as max_delay
FROM columnstore_bts.flights f
JOIN columnstore_bts.airports a ON f.dest = a.iata_code
GROUP BY a.city ORDER BY max_delay DESC LIMIT 1;
```

Скріншот виконання:

```
-> FROM innodb_bts.flights f
-> JOIN innodb_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city ORDER BY min_delay LIMIT 1;

+-----+-----+
| city   | min_delay |
+-----+-----+
| Seattle |    -82.00 |
+-----+-----+
1 row in set (6.083 sec)

MariaDB [(none)]> SELECT a.city, MAX(f.arr_delay) as max_delay
-> FROM innodb_bts.flights f
-> JOIN innodb_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city ORDER BY max_delay DESC LIMIT 1;

+-----+-----+
| city               | max_delay |
+-----+-----+
| Dallas-Fort Worth |    1971.00 |
+-----+-----+
1 row in set (6.491 sec)

MariaDB [(none)]> SELECT a.city, MIN(f.arr_delay) as min_delay
-> FROM columnstore_bts.flights f
-> JOIN columnstore_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city ORDER BY min_delay LIMIT 1;

+-----+-----+
| city   | min_delay |
+-----+-----+
| Seattle |    -82.00 |
+-----+-----+
1 row in set (0.337 sec)
```

```

MariaDB [(none)]> SELECT a.city, MIN(f.arr_delay) as min_delay
-> FROM columnstore_bts.flights f
-> JOIN columnstore_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city ORDER BY min_delay LIMIT 1;

```

```

+-----+-----+
| city   | min_delay |
+-----+-----+
| Seattle |    -82.00 |
+-----+-----+
1 row in set (0.337 sec)

```

```

MariaDB [(none)]> SELECT a.city, MAX(f.arr_delay) as max_delay
-> FROM columnstore_bts.flights f
-> JOIN columnstore_bts.airports a ON f.dest = a.iata_code
-> GROUP BY a.city ORDER BY max_delay DESC LIMIT 1;

```

```

+-----+-----+
| city           | max_delay |
+-----+-----+
| Dallas-Fort Worth |    1971.00 |
+-----+-----+
1 row in set (0.328 sec)

```

■

Швидкість виконання на стовпчиковій базі даних: 0,337 та 0,328 секунди

Швидкість виконання на звичайній базі даних: 6,083 та 6,491 секунди

4 ЗНАЙТИ ВСІ ПОЛЬОТИ З ЗАТРИМКОЮ БІЛЬШЕ ЗА СЕРЕДНІЙ ЧАС ЗАТРИМКИ

Код SQL для звичайної БД:

```
SELECT f.*
FROM innodb_bts.flights f
WHERE f.arr_delay > (SELECT AVG(f.arr_delay) FROM
innodb_bts.flights f);
```

Код SQL для стовпчикової БД:

```
SELECT f.*
FROM columnstore_bts.flights f
WHERE f.arr_delay > (SELECT AVG(f.arr_delay) FROM
columnstore_bts.flights f);
```

Скріншот виконання:

```
| 2015 | 1 | 23 | 5 | 2015-01-23 | AA | N598AA | 1322 | BHM | DFW | 700 |
5.00 | 914 | 1805 | 1971.00 | 0.00 | 0.00 | 0.00 | 134.00 |
0.00 | 0.00 | 0.00 | 0.00 |
+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+
309557 rows in set (41.593 sec)
```

MariaDB [(none)]>

```
| 2015 | 1 | 23 | 5 | 2015-01-23 | AA | N598AA | 1322 | BHM | DFW | 700 |
0.00 | 32.00 | 0.00 | 0.00 |
+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+
309557 rows in set (5.171 sec)
```

MariaDB [(none)]>

Швидкість виконання на стовпчиковій базі даних: 5,171 секунди

Швидкість виконання на звичайній базі даних: 41,593 секунди

5 ЗАМІРЯТИ ВБУДОВАНИМИ МЕТОДАМИ ОБ'ЄМ БД ТА ШВИДКІСТЬ ВИКОНАННЯ ЗАПИТІВ. ПОРІВНЯТИ ЗВИЧАЙНУ І СТОВПЧИКОВУ

5.1 Порівняння об'ємів БД

Для звичайної БД:

```
du /var /lib/mysql /innodb_bts -h
```

Для стовпчикової БД:

```
CALL columnstore_info.table_usage('columnstore_bts', NULL);
```

Скріншоти результатів:

```
sh-4.4# du /var/lib/mysql/innodb_bts -h
413M    /var/lib/mysql/innodb_bts
sh-4.4# mariadb
Welcome to the MariaDB monitor.  Commands end with ; or \g.
Your MariaDB connection id is 5
Server version: 11.1.1-MariaDB-log MariaDB Server

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> CALL columnstore_info.table_usage('columnstore_bts', NULL);
+-----+-----+-----+-----+-----+
| TABLE_SCHEMA | TABLE_NAME | DATA_DISK_USAGE | DICT_DATA_USAGE | TOTAL_USAGE |
+-----+-----+-----+-----+-----+
| columnstore_bts | airlines   | 240.00 KB        | 216.00 KB        | 456.00 KB   |
| columnstore_bts | airports   | 840.00 KB        | 872.00 KB        | 1.67 MB     |
| columnstore_bts | flights    | 52.23 MB         | 6.38 MB          | 58.61 MB    |
+-----+-----+-----+-----+-----+
3 rows in set (0.577 sec)
```

Об'єм звичайної БД: 413 МВ

Об'єм стовпчикової БД: 58,61 МВ

5.2 Порівняння швидкості виконання запитів

	звичайна	стовпчикова	порівняння
Сумарна затримка по містах	7,564	0,528	14
Кількість польотів по містах	1,601	0,352	4
Місто з найменшою і найбільшою затримкою	6,083	0,337 тата 0,328	18
Польоти з затримкою більше за середній час затримки	41,593	5,171	8

Висновок

За результатами виконання лабораторної роботи стовпчикова БД виявилася в рази ефективнішою при виконанні запитів та займає менше обсягу пам'яті.