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

Звіт № 1 з
дисципліни
«Програмування інтелектуальних
інформаційних систем»

Виконав студент ІП-13 Романюк Діана Олексіївна
(шифр, прізвище, ім'я, по батькові)

Перевірів Баришич Лука Маріянович
(прізвище, ім'я, по батькові)

Лабораторна робота 1

Постановка задачі:

1. Створити стовпчикову і звичайну бд
 2. Розрахувати сумарну затримку по містах
 3. Порахувати кількість польотів по містах
 4. Знайти місто з найменшою і найбільшою затримкою
 5. Знайти всі польоти з затримкою більше за середній час затримки
 6. Заміряти вбудованими методами об'єм БД та швидкість виконання запитів.
- Порівняти звичайну і стовпчикову

Завдання 1

Створення звичайної та стовпчикової бд

```
drop database if exists `innodb_bts`;

create database `innodb_bts`;

use `innodb_bts`;

create table `airlines` (
  `iata_code` varchar(2) not null,
  `airline` varchar(30) default null,
  primary key (`iata_code`),
  key `airline` (`airline`)
) engine=innodb default charset=utf8mb3 collate=utf8mb3_general_ci;

create table `airports` (
  `iata_code` varchar(3) not null,
  `airport` varchar(80) default null,
  `city` varchar(30) default null,
  `state` varchar(2) default null,
  `country` varchar(30) default null,
  `latitude` decimal(11,4) default null,
  `longitude` decimal(11,4) default null,
  primary key (`iata_code`),
  key `state` (`state`)
) engine=innodb default charset=utf8mb3 collate=utf8mb3_general_ci;

create table `flights` (
  `year` smallint(6) default null,
  `month` tinyint(4) default null,
  `day` tinyint(4) default null,
  `day_of_week` tinyint(4) default null,
  `fl_date` date default null,
  `carrier` varchar(2) default null,
  `tail_num` varchar(6) default null,
  `fl_num` smallint(6) default null,
  `origin` varchar(5) default null,
  `dest` varchar(5) not null,
  `crs_dep_time` varchar(4) default null,
  `dep_time` varchar(4) default null,
```

```

`dep_delay` decimal(13,2) default null,
`taxi_out` decimal(13,2) default null,
`wheels_off` varchar(4) default null,
`wheels_on` varchar(4) default null,
`taxi_in` decimal(13,2) default null,
`crs_arr_time` varchar(4) default null,
`arr_time` varchar(4) default null,
`arr_delay` decimal(13,2) default null,
`cancelled` decimal(13,2) default null,
`cancellation_code` varchar(20) default null,
`diverted` decimal(13,2) default null,
`crs_elapsed_time` decimal(13,2) default null,
`actual_elapsed_time` decimal(13,2) default null,
`air_time` decimal(13,2) default null,
`distance` decimal(13,2) default null,
`carrier_delay` decimal(13,2) default null,
`weather_delay` decimal(13,2) default null,
`nas_delay` decimal(13,2) default null,
`security_delay` decimal(13,2) default null,
`late_aircraft_delay` decimal(13,2) default null,
key `carrier` (`carrier`),
key `year` (`year`),
key `carrier_delay` (`carrier_delay`),
key `weather_delay` (`weather_delay`),
key `nas_delay` (`nas_delay`),
key `security_delay` (`security_delay`),
key `late_aircraft_delay` (`late_aircraft_delay`),
key `arr_delay` (`arr_delay`),
key `month` (`month`),
key `dest` (`dest`)
) engine=innodb default charset=utf8mb3 collate=utf8mb3_general_ci;

drop database if exists `columnstore_bts`;

create database `columnstore_bts`;

use `columnstore_bts`;

create table `airlines` (
  `iata_code` varchar(2) default null,
  `airline` varchar(30) default null
) engine=columnstore default charset=utf8mb3 collate=utf8mb3_general_ci;

create table `airports` (
  `iata_code` varchar(3) default null,
  `airport` varchar(80) default null,
  `city` varchar(30) default null,
  `state` varchar(2) default null,
  `country` varchar(30) default null,
  `latitude` decimal(11,4) default null,
  `longitude` decimal(11,4) default null
) engine=columnstore default charset=utf8mb3 collate=utf8mb3_general_ci;

create table `flights` (
  `year` smallint(6) default null,
  `month` tinyint(4) default null,
  `day` tinyint(4) default null,
  `day_of_week` tinyint(4) default null,
  `fl_date` date default null,
  `carrier` varchar(2) default null,
  `tail_num` varchar(6) default null,
  `fl_num` smallint(6) default null,
  `origin` varchar(5) default null,
  `dest` varchar(5) default null,
  `crs_dep_time` varchar(4) default null,
  `dep_time` varchar(4) default null,

```

```
`dep_delay` decimal(13,2) default null,  
`taxi_out` decimal(13,2) default null,  
`wheels_off` varchar(4) default null,  
`wheels_on` varchar(4) default null,  
`taxi_in` decimal(13,2) default null,  
`crs_arr_time` varchar(4) default null,  
`arr_time` varchar(4) default null,  
`arr_delay` decimal(13,2) default null,  
`cancelled` decimal(13,2) default null,  
`cancellation_code` varchar(20) default null,  
`diverted` decimal(13,2) default null,  
`crs_elapsed_time` decimal(13,2) default null,  
`actual_elapsed_time` decimal(13,2) default null,  
`air_time` decimal(13,2) default null,  
`distance` decimal(13,2) default null,  
`carrier_delay` decimal(13,2) default null,  
`weather_delay` decimal(13,2) default null,  
`nas_delay` decimal(13,2) default null,  
`security_delay` decimal(13,2) default null,  
`late_aircraft_delay` decimal(13,2) default null  
) engine=columnstore default charset=utf8mb3 collate=utf8mb3_general_ci;
```

Завантаження даних

```
USE `innodb_bts`;  
USE `columnstore_bts`;  
  
TRUNCATE TABLE airlines;  
LOAD DATA INFILE '../.../.../data/airlines.csv'  
    INTO TABLE airlines  
    FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"'  
    LINES TERMINATED BY '\n'  
    IGNORE 1 LINES;  
  
TRUNCATE TABLE airports;  
LOAD DATA INFILE '../.../.../data/airports.csv'  
    INTO TABLE airports  
    FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"'  
    LINES TERMINATED BY '\n'  
    IGNORE 1 LINES;  
  
TRUNCATE TABLE flights;  
SET SESSION sql_mode = '';  
LOAD DATA INFILE '../.../.../data/flights.csv'  
    INTO TABLE flights  
    FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"'  
    LINES TERMINATED BY '\n';
```

Завдання 2

Розрахунок сумарної затримки по містах

```
-- task2  
-- total delay for each city  
SELECT a.city, SUM(f.arr_delay) AS 'total delay'  
FROM flights AS f  
JOIN airports AS a ON f.dest = a.iata_code  
GROUP BY a.city;
```

	city	`total delay`
1	Aberdeen	-525.00
2	Abilene	5952.00
3	Adak	-293.00
4	Agana	994.00
5	Aguadilla	2634.00
6	Akron	6390.00
7	Albany	12775.00
8	Albuquerque	18043.00

	city	`total delay`
1	Boston	241473.00
2	Chicago	804105.00
3	Ft. Myers	36346.00
4	Phoenix	100955.00

Завдання 3

Розрахунок кількості польотів по містах

```
-- task3
-- Count the number of flights for each city
SELECT a.city, COUNT(*) AS 'amount of flights'
FROM flights f
JOIN airports a ON f.dest = a.iata_code
GROUP BY a.city;
```

	city	`amount of flights`
1	Aberdeen	121
2	Abilene	568
3	Adak	21
4	Agana	65
5	Aguadilla	224
6	Akron	1282
7	Albany	1523
8	Albuquerque	3462

	city	`amount of flights`
1	Chicago	69097
2	Baltimore	15169
3	Ft. Lauderdale	16717
4	Raleigh	5947
5	New Orleans	7765

Завдання 4

Знайти місто з найменшою і найбільшою затримкою

```
-- task4
-- city with the highest and lowest delays

-- city with the lowest delay
SELECT a.city, f.arr_delay
FROM airports a
JOIN flights f ON a.iata_code = f.dest
WHERE f.arr_delay = (SELECT MIN(arr_delay) FROM flights)
```

	city	arr_delay
1	Seattle	-82.00

```
ORDER BY f.arr_delay;

-- city with the highest delay
SELECT a.city, f.arr_delay
FROM airports a
JOIN flights f ON a.iata_code = f.dest
WHERE f.arr_delay = (SELECT MAX(arr_delay) FROM flights)
ORDER BY f.arr_delay;
```

	city	arr_delay
1	Dallas-Fort Worth	1971.00

Завдання 5

Знайти всі польоти з затримкою більше за середній час затримки

```
-- task5
-- find flights with a delay greater than the average delay

-- calculate and display the average delay
SELECT AVG(arr_delay) AS 'Average Delay'
FROM flights;

SELECT fl_date, tail_num, fl_num, dest, arr_delay
FROM flights
WHERE arr_delay > (SELECT AVG(arr_delay) FROM flights);
```

SELECT AVG(arr_delay) AS 'Average Delay'	
FROM flights;	
SELECT fl_date, tail_num, fl_num, dest, arr_delay	
FROM flights	
WHERE arr_delay > (SELECT AVG(arr_delay) FROM flights);	
-- task6	
Output Average Delay:decimal calculate and display the average delay	
1 row	
Average Delay	
1	6.612437

Завдання 6

Заміряти вбудованими методами об'єм БД та швидкість виконання запитів.

```
-- task6
-- database size

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

SELECT
    table_schema AS 'Database',
    SUM(data_length + index_length) / 1024 / 1024 AS 'Database Size (MB)'
FROM information_schema.tables
WHERE table_schema = 'innodb_bts'
GROUP BY table_schema;
```

Висновок:

Порівняємо звичайну і стовпчикову бд

Об'єм

Стовпчикова бд columnstore_bts:

	TABLE_SCHEMA	TABLE_NAME	DATA_DISK_USAGE	DICT_DATA_USAGE	TOTAL_USAGE
1	columnstore_bts	airlines	240.00 KB	216.00 KB	456.00 KB
2	columnstore_bts	airports	840.00 KB	872.00 KB	1.67 MB
3	columnstore_bts	flights	50.23 MB	6.38 MB	56.61 MB

Innodb_bts:

	`Database`	`Database Size (MB)`
1	innodb_bts	0.23437500

Швидкість виконання запитів

Стовпчикова бд columnstore_bts:

 lab1.sql 73 ms

Innodb_bts:

 lab1.sql 82 ms