

Міністерство освіти і науки України
НТУУ «Київський політехнічний інститут»
Фізико-технічний інститут

Системи та засоби інтерактивної аналітики

Лабораторна робота №4

Advanced SQL

Варіант №6

Виконав:
Студент 4 курсу ФТІ
групи ФБ-01

Сахній Назар

Перевірив:

Лабораторна робота №4

Advanced SQL

Тема: Запити розширеного SQL (advanced SQL).

Завдання: Навчитися створювати розширені SQL запити.

№ варіанта	Завдання
6	<p><u>Використати існуючу БД (лаб. роб. 2,3)</u></p> <p>Запити:</p> <ol style="list-style-type: none"> 1. Вивести повні адреси JS скриптів (розширення js) і їх розмір, які запитали з комп'ютера з IP адресою 83.227.29.211, для яких запит завершився вдало. 2. Придумайте будь який запит з використанням UNION. 3. Знайдіть розмір всього скачаного комп'ютером з IP адресою 83.227.29.211

Робоче завдання

1. До БД створеної у попередніх роботах імпортувати таблицю з даними.

Імпортування до бази даних "law_firm_pro"

Файл для імпорту:

Файл може бути стиснений (gzip, bzip2) або нестиснений.
Ім'я стисненого файлу повинно закінчуватися на `[format].[compression]`. Приклад: `.sql.zip`

Переглянути Ваш комп'ютер: (Максимум: 40МБ)

Вибрати файл

tblaccesslog.sql

Ви також можете перетягнути файл на будь-якій сторінці.

Кодування файлу:

utf-8

```
✔ Імпорти завершився ефекно. 31 запит виконано. (tblaccesslog.sql)
```

```
✔ MySQL повернула пустий результат (тобто нуль рядків). (Запит виконувався 0.0003 секунди.)
```

```
CREATE TABLE `tblaccesslog` ( `ID` int(11) NOT NULL, `Line` text NOT NULL ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4;
```

[Порядкові редагування] [Редaguвати] [Сторінки PHP код]

```
✔ 160 рядків додано. (Запит виконувався 0.0004 секунди.)
```

```
-- -- Dumping data for table `tblaccesslog`: -- INSERT INTO `tblaccesslog` (`ID`, `Line`) VALUES (1381, '13.66.139.0 - - [19/Dec/2020:13:57:26 +0100] "GET /index.php?option=com_phocagallery&view=category&id=1&layout=raith&Itemid=53 HTTP/1.1" 200 32053 "-" \"/Mozilla/5.0 (compatible; bingbot/2.0; http://www.bing.com/bingbot.htm)" "-"', (1382, '157.40.153.185 - - [19/Dec/2020:14:09:06 +0100] "\GET /apache-log/access.log HTTP/1.1" 200 233 "-" \"/Mozilla/5.0 (Windows NT 6.3; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/87.0.4280.88 Safari/537.36)" "-"', (1383, '157.48.193.185 - - [19/Dec/2020:14:08:08 +0100] "\GET /favicon.ico HTTP/1.1" 404 217 "\http://mm.alshuette-raith.at/apache-log/access.log)" \"/Mozilla/5.0 (Windows NT 6.3; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/87.0.4280.88 Safari/537.36)" "-"', (1384, '216.244.66.230 - - [19/Dec/2020:14:14:26 +0100] "\GET /robots.txt HTTP/1.1" 200 304 "-" \"/Mozilla/5.0 (compatible; Dotbot/1.1; [...]
```

[Редaguвати]

```
✔ 186 рядків додано. (Запит виконувався 0.0003 секунди.)
```

```
INSERT INTO `tblaccesslog` (`ID`, `Line`) VALUES (1546, '54.36.149.8 - - [19/Dec/2020:18:07:36 +0100] "\GET /index.php?option=com_phocagallery&view=category&id=1&layout=raith&Itemid=53&limitstart=20 HTTP/1.1" 200 15462 "-" \"/Mozilla/5.0 (compatible; AhrefsBot/7.0; http://ahrefs.com/robot)" "-"', (1550, '45.132.207.154 - - [19/Dec/2020:18:07:50 +0100] "\GET /index.php?option=com_contact&view=contact&id=1 HTTP/1.1" 200 9873 "-" \"/Mozilla/5.0 (Linux; Android8.1.0;zh-CN;EML-AL00Build/HUAWEIFH-AL00) AppleWebKit/537.36 (KHTML, like Gecko) Version/4.0 Chrome/57.0.2987.108baidu.sogo.uc.UCBrowser/11.9.4.974UI/5.2.15.1.48MobileSafari/537.36AliTalk/4.5.11) com.alibaba.a)" "-"', (1551, '45.132.207.154 - - [19/Dec/2020:18:07:50 +0100] "\POST /index.php?option=com_contact&view=contact&id=1 HTTP/1.1" 200 188 "-" \"/Mozilla/5.0 (Linux; Android8.1.0;zh-CN;EML-AL00Build/HUAWEIFH-AL00) AppleWebKit/537.36 (KHTML, like Gecko) Version/4.0 Chrome/57.0.2987.108baidu.sogo.uc.UCBrowser/11.9.4.4[...]
```

[Редaguвати]

2. Написання відповідних SQL-запитів, які необхідно для виконання завдання.

2.1. Вивести повні адреси JS скриптів (розширення **js**) і їх розмір, які запросили з комп'ютера з IP адресою **83.227.29.211**, для яких запит завершився вдало:

Виконати SQL запит(и) у базі даних law_firm_pro.tblaccesslog: ⓘ

```
1 SELECT
2   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 1), ' ', -1) AS IPAddress,
3   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, '[', -1), ' ', 1) AS LogDate,
4   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 7), ' ', -1) AS FilePath,
5   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 9), ' ', -1) AS HttpStatusCode,
6   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 10), ' ', -1) AS ByteSize
7 FROM
8   tblaccesslog
9 WHERE
10  Line LIKE ('83.227.29.211%.js%')
11  AND SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 9), ' ', -1) BETWEEN '200' AND '206'
12
```

☐ Показати все | Число рядків: 25 | Фільтрувати рядки: Шукати в таблиці | Сортувати за ключем: Жодного

Екстра параметри

IPAddress	LogDate	FilePath	HttpStatusCode	ByteSize
83.227.29.211	19/Dec/2020:19:54:09	/media/system/js/modal.js	200	10588
83.227.29.211	19/Dec/2020:19:54:10	/components/com_phocagallery/assets/js/shadowbox/s...	200	27272
83.227.29.211	19/Dec/2020:19:54:10	/components/com_phocagallery/assets/js/shadowbox/s...	200	2337
83.227.29.211	19/Dec/2020:19:54:10	/components/com_phocagallery/assets/js/shadowbox/s...	200	3495
83.227.29.211	19/Dec/2020:19:54:10	/components/com_phocagallery/assets/js/shadowbox/s...	200	8324

2.2. Придумайте будь який запит з використанням UNION:

Виконати SQL запит(и) у базі даних law_firm_pro.tblaccesslog: ⓘ

```
1 SELECT
2   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 1), ' ', -1) AS IPAddress,
3   SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ']', -1), ' ', 5) AS LogRecord
4 FROM
5   tblaccesslog
6 WHERE
7   (Line LIKE '%GET%' AND Line LIKE '%HTTP/1.1\" 301%')
8   OR (Line LIKE '%POST%' AND Line LIKE '%HTTP/1.1\" 404%')
9
10 UNION ALL
11
12 SELECT
13   ResponsibleLawyer,
14   Note
15 FROM
16   responsibility
17 WHERE
18   Note IS NOT NULL
```

<input type="checkbox"/> Показати все	Число рядків: 25	Фільтрувати рядки: Шукати в таблиці	Сортувати за ключем: Жодного
Екстра параметри			
IPAddress	LogRecord		
193.106.30.100	"POST /sites/default/settings.php HTTP/1.1" 404		
193.106.30.100	"POST /sites/all/stats.php HTTP/1.1" 404		
193.106.30.100	"POST /stats.php HTTP/1.1" 404		
194.54.183.74	"GET /apache-log HTTP/1.1" 301		
5.61.253.37	"GET /cache/content-post.php.suspected HTTP/1.1" ...		
Луцик М. В.	Кримінальний юрист з багаторічним досвідом. Спеціа...		
Падик О. В.	Адвокат із спеціалізацією в адміністративних справ...		
Іваненко І. І.	Спеціаліст з господарського права та корпоративних...		
Ковальчук В. І.	Юрист із фаховою експертизою у господарському та к...		
Харенко О. Н.	Спеціаліст з права нерухомості, який допомагає клі...		
Мельник О. О.	Корпоративний юрист зі спеціалізацією на оподаткув...		
Бондаренко О. В.	Спеціалізується на інтелектуальній власності та ав...		
Березенко А. В.	Спеціалізується на адміністративних справах та роб...		
Тур К. Б.	Юрист із судової практики та апеляційних справ. До...		
Івасюк В. М.	Один із основоположників української естрадної муз...		

2.3. Знайдіть розмір всього скачаного комп'ютером з IP адресою 83.227.29.211:

Виконати SQL запит(и) у базі даних law_firm_pro.tblaccesslog: ?	
1	SELECT
2	IPAddress, SUM(ByteSize) AS TotalSize
3	FROM (
4	SELECT
5	SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 1), ' ', -1) AS IPAddress,
6	SUBSTRING_INDEX(SUBSTRING_INDEX(Line, 'GET ', -1), ' HTTP', 1) AS FilePath,
7	SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 9), ' ', -1) AS HttpStatusCode,
8	SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 10), ' ', -1) AS ByteSize
9	FROM
10	tblaccesslog
11	WHERE Line REGEXP ('^83\.227\.29\.211*')
12	AS DownloadedByIP

<input type="checkbox"/> Показати все	Число рядків: 25	Фільтрувати рядки: Шукати в таблиці
Екстра параметри		
IPAddress	TotalSize	
83.227.29.211	462376	

3. SQL-запити для створення всієї структури БД.

```
--
-- Вивести повні адреси JS скриптів (розширення js) і їх розмір, які запросили
-- з комп'ютера з IP адресою 83.227.29.211, для яких запит завершився вдало.
--
SELECT
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 1), ' ', -1) AS IPAddress,
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, '[', -1), ' ', 1) AS LogDate,
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 7), ' ', -1) AS FilePath,
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 9), ' ', -1) AS HttpStatusCode,
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 10), ' ', -1) AS ByteSize
FROM
    tblaccesslog
WHERE
    Line LIKE ('83.227.29.211%.js%')
    AND SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 9), ' ', -1) BETWEEN '200' AND '206'

--
-- Придумайте будь який запит з використанням UNION
--
SELECT
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 1), ' ', -1) AS IPAddress,
    SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ']', -1), ' ', 5) AS LogRecord
FROM tblaccesslog
WHERE
    (Line LIKE '%GET%' AND Line LIKE '%HTTP/1.1\" 301%')
    OR
    (Line LIKE '%POST%' AND Line LIKE '%HTTP/1.1\" 404%')

UNION ALL

SELECT
    ResponsibleLawyer,
    Note
FROM
    responsibility
WHERE
    Note IS NOT NULL

--
-- Знайдіть розмір всього скачаного комп'ютером з IP адресою 83.227.29.211
--
SELECT
    IPAddress, SUM(ByteSize) AS TotalSize
FROM (
    SELECT
        SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 1), ' ', -1) AS IPAddress,
        SUBSTRING_INDEX(SUBSTRING_INDEX(Line, 'GET ', -1), ' HTTP', 1) AS FilePath,
        SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 9), ' ', -1) AS HttpStatusCode,
        SUBSTRING_INDEX(SUBSTRING_INDEX(Line, ' ', 10), ' ', -1) AS ByteSize
    FROM
        tblaccesslog
    WHERE Line REGEXP ('^83\.227\.29\.211*'))
AS DownloadedByIP
```

Контрольні питання

1. Що таке запити розширеного SQL (advanced SQL)?

Запити розширеного SQL – це високорівневі SQL-операції, які виходять за рамки базового синтаксису SQL. Вони включають у себе різноманітні функції, операції та можливості, які дозволяють більш гнучко та ефективно взаємодіяти з базами даних. До них можуть входити агрегатні функції, віконні функції, операції над множинами, аналітичні функції та інші.

2. Що таке регулярні вирази?

Регулярні вирази – це шаблони для пошуку та обробки тексту. Вони дозволяють визначати певні правила для пошуку, вилучення або заміни рядків в текстових даних. Регулярні вирази широко використовуються в програмуванні, редакторах тексту, операціях обробки даних та запитах до баз даних для більш гнучкого та потужного пошуку та аналізу текстової інформації.

3. Для чого використовуються запити з командою UNION?

Запити з командою UNION використовуються для об'єднання результатів двох або більше SELECT-запитів у єдиний результат. UNION видаляє дублікати, об'єднуючи унікальні рядки. Це дозволяє комбінувати дані з різних таблиць чи умов в один результат.

4. Що таке агрегатні функції?

Агрегатні функції – це функції в SQL, які виконують обчислення на підсумкових даних, таких як сума, середнє значення, мінімум, максимум тощо. Ці функції застосовуються до груп рядків та повертають одне значення для кожної групи.

5. Для чого використовуються запити з командою ORDER?

Запити з командою ORDER використовуються для сортування результатів запиту. ORDER BY дозволяє вказати порядок сортування для одного чи кількох стовпців у вибірці даних. Можна сортувати дані в порядку зростання (ASC) або спадання (DESC).

6. Для чого використовуються запити з командою HAVING?

Запити з командою HAVING використовуються для фільтрації результатів групованих даних на основі агрегатних функцій. HAVING використовується після GROUP BY і дозволяє вибирати групи, які відповідають певним умовам. Це дає можливість фільтрувати дані після групування та агрегації.

Висновки

У ході виконання лабораторної роботи було освоєно написання **Advanced SQL**-запитів.

Було використано можливості функції SUBSTRING_INDEX для вилучення та обробки інформації з лог-файлів. Команда UNION була необхідна для комбінування результатів різних запитів, що є зручно для об'єднання даних із різних джерел чи умов. Також, використання агрегатних функцій (SUM) було необхідне для обчислення підсумкових значень та виведення загального обсягу завантажень для конкретної IP-адреси.