

Dokumentácia k zadaniu 2

Autor: Dominik Miškovič
Predmet: Databázové systémy
Cvičiaci: Ing. William Brach
Obdobie: Letný semester 2024

Obsah

1	Endpoint 1	3
1.1	Query	3
1.2	Opis fungovania	3
1.3	Príklad volania	3
2	Endpoint 2	4
2.1	Query	4
2.2	Opis fungovania	4
2.3	Príklad volania	5
3	Endpoint 3	6
3.1	Query	6
3.2	Opis fungovania	6
3.3	Príklad volania	6
4	Endpoint 4	7
4.1	Query	7
4.2	Opis fungovania	7
4.3	Príklad volania	7
5	Endpoint 5	8
5.1	Query	8
5.2	Opis fungovania	8
5.3	Príklad volania	8

1 Endpoint 1

1.1 Query

```
1  SELECT
2  u.id, u.reputation, TO_CHAR(u.creationdate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF'),
3  u.displayname, TO_CHAR(u.lastaccessdate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF'),
4  u.websiteurl, u.location, u.aboutme, u.views, u.upvotes, u.downvotes, u.profileimageurl,
5  u.age, u.accountid
6  FROM users u
7  JOIN comments c ON u.id = c.userid
8  WHERE c.postid = '{post_id}'
9  ORDER BY c.creationdate DESC
10
```

1.2 Opis fungovania

Najskôr sa vyberú relevantné informácie o používateľoch. Dátumy sa formátujú podľa štandardu ISO8601 vo funkcii TO_CHAR. Podmienka WHERE c.postid = 'post_id' zabezpečuje, že sa vyberajú len komentáre, ktoré majú priradené ID príspevku rovnaké ako post_id. Tabuľka users sa pripojí na tabuľku comments na základe rovnakého ID používateľa. Výsledky sú nakoniec usporiadané podľa dátumu vytvorenia komentárov, od najnovších k najstarším.

1.3 Príklad volania

http://localhost:8000/v2/posts/1819157/users

```
{
  "items": [
    {
      "id": 1866388,
      "reputation": 1,
      "creationdate": "2023-12-01T00:05:24.337+01",
      "displayname": "TomR.",
      "lastaccessdate": "2023-12-03T06:18:19.607+01",
      "websiteurl": null,
      "location": null,
      "aboutme": null,
      "views": 1,
      "upvotes": 0,
      "downvotes": 0,
      "profileimageurl": null,
      "age": null,
      "accountid": 30035903
    }
  ]
}
```

2 Endpoint 2

2.1 Query

```
1  WITH filtered_posts AS (  
2  SELECT p.id  
3  FROM posts p  
4  WHERE p.owneruserid = '{user_id}'  
5  
6  UNION  
7  
8  SELECT p.id  
9  FROM comments c  
10 JOIN posts p ON p.id = c.postid  
11 WHERE c.userid = '{user_id}'  
12 )  
13 SELECT DISTINCT  
14 u.id, u.reputation, TO_CHAR(u.creationdate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF') AS cd,  
15 u.displayname, TO_CHAR(u.lastaccessdate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF'),  
16 u.websiteurl, u.location, u.aboutme, u.views, u.upvotes, u.downvotes,  
17 u.profileimageurl, u.age, u.accountid  
18 FROM users u  
19 JOIN comments c ON u.id = c.userid  
20 WHERE c.postid IN (SELECT id FROM filtered_posts)  
21 ORDER BY cd ASC  
22
```

2.2 Opis fungovania

Najskôr sa vytvorí pomocná tabuľka `filtered_posts`. Vyberú sa príspevky, ktoré boli vytvorené používateľom s ID `user_id`. Potom sa vyberú príspevky na ktorých používateľ komentoval. Nakoniec sa tieto výbery spoja pomocou `UNION`. Potom nasleduje hlavná časť query. Vyberú sa relevantné informácie o používateľoch a za použitia `DISTINCT` zabránime duplikátom. Pripoja sa tabuľky `users` a `comments` pomocou podmienky, že ID používateľa sa rovná ID používateľa v komentároch. Výber sa obmedzí na príspevky, ktoré boli identifikované v dočasnej tabuľke `filtered_posts` pomocou `WHERE`. Nakoniec sa výsledky usporiadajú podľa dátumu registrácie používateľov začínajúc od najstarších registrovaných používateľov.

2.3 Príklad volania

`http://localhost:8000/v2/users/1076348/friends`

```
{
  "items": [
    {
      "id": 482362,
      "reputation": 10581,
      "creationdate": "2015-08-11T17:42:36.267+02",
      "displayname": "DrZoo",
      "lastaccessdate": "2023-12-03T06:41:11.750+01",
      "websiteurl": null,
      "location": null,
      "aboutme": null,
      "views": 1442,
      "upvotes": 555,
      "downvotes": 46,
      "profileimageurl": null,
      "age": null,
      "accountid": 2968677
    },
    {
      "id": 1076348,
      "reputation": 1,
      "creationdate": "2019-08-15T16:00:28.473+02",
      "displayname": "Richard",
      "lastaccessdate": "2019-09-10T16:57:48.527+02",
      "websiteurl": null,
      "location": null,
      "aboutme": null,
      "views": 0,
      "upvotes": 0,
      "downvotes": 0,
      "profileimageurl": null,
      "age": null,
      "accountid": 16514661
    }
  ]
}
```

3 Endpoint 3

3.1 Query

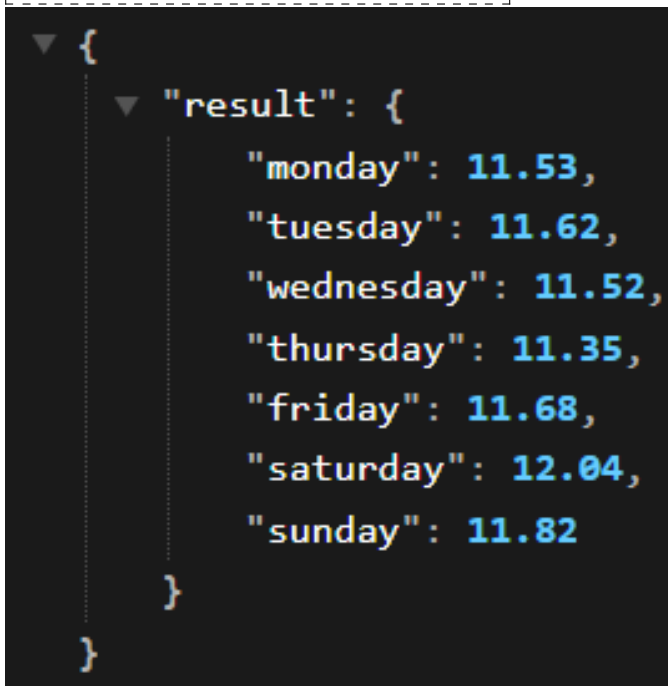
```
1  SELECT
2  TO_CHAR(p.creationdate, 'day') AS day,
3  ROUND((((COUNT(pt.post_id)
4  FILTER (WHERE t.tagname = '{tagname}'))::numeric) /
5  COUNT(DISTINCT pt.post_id)::numeric) * 100.0, 2)
6  FROM posts p
7  JOIN post_tags pt ON p.id = pt.post_id
8  JOIN tags t ON pt.tag_id = t.id
9  GROUP BY EXTRACT(ISODOW FROM creationdate), day
10 ORDER BY EXTRACT(ISODOW FROM creationdate)
11
```

3.2 Opis fungovania

Najskôr sa vyberie čas vytvorenia príspevkov preformátovaný na tvar konkrétneho mena dňa v týždni. Potom sa vypočíta percentuálna časť príspevkov s konkrétnym tagom vzhľadom na celkový počet príspevkov. Použije sa FILTER na vyfiltrovanie príspevkov s konkrétnym tagom. Výsledok sa zaokrúhli na dve desatinné miesta pomocou funkcie ROUND. Tabuľky posts, post_tags, a tags sa spoja na základe vzájomných identifikátorov. Skupinuje sa podľa dňa v týždni (EXTRACT(ISODOW FROM creationdate)) a reprezentácie dňa (day). Je potrebné použiť GROUP BY lebo sa používa agregácia. Agreguje sa počet príspevkov s daným tagom pre každý deň týždňa. Výsledky sú nakoniec usporiadané podľa dňa v týždni od pondelka do nedele a to je dosiahnuté vďaka ISODOW, ktoré radí dni od pondelka do nedele narozdiel od DOW ktoré radí dni od nedele do pondelka.

3.3 Príklad volania

`http://localhost:8000/v2/tags/linux/stats`



```
{
  "result": {
    "monday": 11.53,
    "tuesday": 11.62,
    "wednesday": 11.52,
    "thursday": 11.35,
    "friday": 11.68,
    "saturday": 12.04,
    "sunday": 11.82
  }
}
```

4 Endpoint 4

4.1 Query

```
1  SELECT
2  p.id, TO_CHAR(p.creationdate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF'),
3  p.viewcount, TO_CHAR(p.lasteditdate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF'),
4  TO_CHAR(p.lastactivitydate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF'), p.title,
5  TO_CHAR(p.closeddate, 'YYYY-MM-DD"T"HH24:MI:SS.MSOF') AS cd,
6  ROUND(EXTRACT(EPOCH FROM (p.closeddate - p.creationdate)) / 60.0, 2)
7  FROM posts p
8  WHERE p.closeddate IS NOT NULL
9  AND EXTRACT(EPOCH FROM (p.closeddate - p.creationdate)) / 60.0 <= '{duration}'
10 ORDER BY cd DESC
11 LIMIT '{limit}';
12
```

4.2 Opis fungovania

Najskôr sa vyberú relevantné informácie z tabuľky posts. Dátumy sa formátujú pomocou funkcie TO_CHAR na formát ISO8601. Vyberajú sa len príspevky, ktoré majú nastavený dátum uzavretia, nakoľko potrebujeme zistiť duration v tvare počtu minút medzi creationdate a closeddate. Ďalej sa filtrujú príspevky, ktoré boli otvorené maximálne po dobu určenú v minútach (duration). Toto sa robí pomocou EPOCH. To vráti časový úsek medzi dvomi obdobiami v sekundách, preto je potrebné ešte výsledok vydeliť 60 aby sme získali minúty. Výsledky sú potom usporiadané podľa dátumu uzavretia zostupne, od najnovších k najstarším a počet výsledkov je limitovaný na zadaný výsledok pomocou LIMIT.

4.3 Príklad volania

http://localhost:8000/v2/posts/?duration=5&limit=2

```
{
  "items": [
    {
      "id": 1818849,
      "creationdate": "2023-11-30T16:55:32.137+01",
      "viewcount": 22924,
      "lasteditdate": null,
      "lastactivitydate": "2023-11-30T16:55:32.137+01",
      "title": "Why is my home router address is 10.x.x.x and not 100.x.x.x which is properly reserved and widely accepted for CGNAT?",
      "closeddate": "2023-11-30T16:59:23.560+01",
      "duration": 3.86
    },
    {
      "id": 1818386,
      "creationdate": "2023-11-27T18:26:57.617+01",
      "viewcount": 19,
      "lasteditdate": null,
      "lastactivitydate": "2023-11-27T18:26:57.617+01",
      "title": "Are there any libraries for parsing DWG files with LGPL, MIT, Apache, BSD?",
      "closeddate": "2023-11-27T18:29:18.947+01",
      "duration": 2.36
    }
  ]
}
```

5 Endpoint 5

5.1 Query

```
1  SELECT
2  p.id, TO_CHAR(p.creationdate, 'YYYY-MM-DD"HH24:MI:SS.MSOF') AS cd,
3  p.viewcount, TO_CHAR(p.lasteditdate, 'YYYY-MM-DD"HH24:MI:SS.MSOF'),
4  TO_CHAR(p.lastactivitydate, 'YYYY-MM-DD"HH24:MI:SS.MSOF'),
5  p.title, p.body, p.answercount,
6  TO_CHAR(p.closeddate, 'YYYY-MM-DD"HH24:MI:SS.MSOF'), ARRAY_AGG(t.tagname)
7  FROM posts p
8  JOIN post_tags pt ON p.id = pt.post_id
9  JOIN tags t ON pt.tag_id = t.id
10 WHERE UNACCENT(LOWER(p.title)) LIKE UNACCENT(LOWER('%{_query}%')) OR
11 UNACCENT(LOWER(p.body)) LIKE UNACCENT(LOWER('%{_query}%'))
12 GROUP BY p.id, p.creationdate, p.viewcount, p.lasteditdate, p.lastactivitydate,
13 p.title, p.body, p.answercount, p.closeddate
14 ORDER BY cd DESC
15 LIMIT {limit};
16
```

5.2 Opis fungovania

Najskôr sa vyberú relevantné informácie z tabuľky posts. Dátumy sa formátujú pomocou funkcie TO_CHAR na formát ISO8601. Tabuľky posts, post_tags a tags sa spoja na základe vzájomných identifikátorov. Keďže vyhľadávanie nie je citlivé na diakritiku a malé/veľké písmená, najskôr sa použije LOWER na všetky porovnávané výrazy. To ich preformátuje na malé písmená. Ďalej sa použije UNACCENT, čo odstráni všetku diakritiku z porovnávaných výrazov. Všetky tagy, ktoré sú priradené danému príspevku, sú agregované do poľa pomocou funkcie ARRAY_AGG. Výsledky sú potom usporiadané podľa dátumu vytvorenia zostupne, od najnovších po najstaršie. Počet výsledkov je nakoniec obmedzený na zadaný limit pomocou LIMIT.

5.3 Príklad volania

```
http://localhost:8000/v2/posts/?limit=1&query=linux
```

```
{
  "items": [
    {
      "id": 1819160,
      "creationdate": "2023-12-03T05:22:43.587+01",
      "viewcount": 7,
      "lasteditdate": null,
      "lastactivitydate": "2023-12-03T05:22:43.587+01",
      "title": "keyboard not working on kali linux",
      "body": "<p>i have recently installed virtualbox on my windows 10 and trying to run Linux Ubuntu and Kali. Everything working on Ubuntu without any issue but when I am running kali it is not taking keyboard(Samsung bluetooth 500) input. Please can anyone help me out here.\nMany thanks in advance!!</p>\n",
      "answercount": 0,
      "closeddate": null,
      "tags": [
        "virtual-machine"
      ]
    }
  ]
}
```