Project Milestone 2
COL362

Team name : BotsAlot
Siddhesh Kalekar, Deepansh Jindal, Prakhar Aggarwal

Queries are as follows :

A) Database creation

```sql
drop table if exists users;

drop table if exists post;

drop table if exists reply;

drop table if exists tags;



CREATE TABLE Users(

    user_id bigint,

    user_name text,

    bio text,

    age bigint,

    creation_date timestamptz,

    last_active timestamptz,

    location text,

    rating bigint,

    likes bigint,

    dislikes bigint,
```

```sql
    num_views bigint,

    user_icon_url text,

    site_url text,

    PRIMARY KEY (user_id)

);


CREATE TABLE Post (

    post_id bigint,

    creator_id bigint,

    creation_date timestamptz,

    rating bigint,

    heading text,

    content text,

    CONSTRAINT questions_pkey PRIMARY KEY (post_id),

    CONSTRAINT questions_fkey FOREIGN KEY (creator_id) REFERENCES Users(user_id)

);


CREATE table Reply(

    reply_id bigint,

    creator_id bigint,

    creation_date timestamptz,

    parent_id bigint,

    rating bigint,

    content text,
```

```
    CONSTRAINT answers_pkey PRIMARY KEY (reply_id),

    CONSTRAINT answers_user_fkey FOREIGN KEY (creator_id) REFERENCES Users(user_id),

    CONSTRAINT answers_question_fkey FOREIGN KEY (parent_id) REFERENCES Post(post_id)

);



CREATE TABLE Tags(

    tag_id bigint,

    tag_name text,

    CONSTRAINT tags_question_fkey FOREIGN KEY (tag_id) REFERENCES Post(post_id)

);
```

B) Queries are as follows :

```
WITH tag_counts AS (

  SELECT tag_name, COUNT(tag_name) AS count

    FROM tags

    GROUP BY tag_name

)

SELECT tag_name

FROM tag_counts

ORDER BY tag_name;




-- q2 --
```

```sql
SELECT heading, rating, post_id

FROM post

ORDER BY rating DESC;



-- q3 optimised --

WITH answered_post as (

    SELECT parent_id

    FROM reply

    WHERE parent_id IS NOT NULL

    GROUP BY parent_id

)

SELECT q.heading, q.post_id

FROM post q

LEFT JOIN answered_post aq ON q.post_id = aq.parent_id

WHERE aq.parent_id IS NULL

ORDER BY q.post_id;




-- q4 --

SELECT heading, creation_date, post_id

FROM post

ORDER BY creation_date DESC;
```

```sql
-- q5 --

SELECT *

FROM users

ORDER BY rating DESC;




-- q6 --

WITH unique_tags AS (

    SELECT DISTINCT tag_name

    FROM tags

)

SELECT COUNT(tag_name)

FROM unique_tags;




-- q7 --

WITH AnswerInfo AS (

 SELECT

    a.content,

    a.rating,

    a.creation_date,

    u.user_name,

    a.reply_id

  FROM
```

```sql
    reply a

  JOIN

    users u

  ON

    a.creator_id = u.user_id

  WHERE

    a.parent_id = 90

)

SELECT

 content,

 rating,

 creation_date,

 user_name,

 reply_id

FROM

 AnswerInfo

ORDER BY

 rating DESC;




-- q8 --

SELECT distinct creator_id

FROM post

WHERE post_id > 1000
```

```
ORDER BY creator_id;



-- q9 --

SELECT distinct creator_id

FROM reply

WHERE reply_id > 2000000

ORDER BY creator_id;



-- q10 --

WITH high_rating_reply AS (

SELECT parent_id, creator_id, creation_date, rating

FROM reply

WHERE rating > 1000

)

SELECT parent_id, creator_id, creation_date, rating

FROM high_rating_reply

ORDER BY creation_date DESC;
```

Indexing Done on :

```
CREATE INDEX idx1 ON reply(parent_id);

CREATE INDEX idx2 ON tags(tag_name);
```

Performance :

Time Without indexing

```
Time: 51.554 ms
Time: 47.757 ms
Time: 104.225 ms
Time: 58.533 ms
Time: 138.206 ms
 count
-------
  4329
(1 row)

Time: 20.710 ms
Time: 17.346 ms
Time: 26.283 ms
Time: 25.294 ms
Time: 22.856 ms
```

Time with indexing

```
Time: 52.292 ms
Time: 35.704 ms
Time: 73.937 ms
Time: 37.901 ms
Time: 189.662 ms
 count
-------
  4329
(1 row)

Time: 20.934 ms
Time: 0.766 ms
Time: 25.182 ms
Time: 12.100 ms
Time: 28.514 ms
```

Database size :

Tags — 29270 entries — 457 KB

Posts —10000 entries —- 9.3 MB

Reply —-----41288 entries —- 28 MB

Users —- 17724 entries —-   3MB