SQL Queries & Relational Algebra

Note: More than required fields are selected in SQL queries to provide context to the information retrieved.

-- 1. what user posted this tweet ?

```
SELECT
users.user_id,
users.username,
hashtag.text
FROM hashtag
JOIN users ON hashtag.user_id = users.user_id
WHERE hashtag.text = 'ballot drop off 📹 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj';
```

Relational Algebra

$\pi_{\text{username}}(\sigma_{\text{hashtag.text = " ballot drop off 📹 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj "}}(\text{hashtag.user\_id} \bowtie \text{users.user\_id}))$

-- 2. when did this user post this tweet?

```
SELECT
users.user_id,
users.username,
hashtag.text,
hashtag.time
FROM hashtag
JOIN users ON hashtag.user_id = users.user_id
WHERE hashtag.text = 'ballot drop off 📹 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj';
```

Relational Algebra

$\pi_{\text{hashtag.time}}(\sigma_{\text{hashtag.text = " ballot drop off 📹 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj "}}(\text{hashtag.user\_id} \bowtie \text{users.user\_id}))$

-- 3. what tweets have this user posted in past 24h ?

SELECT

user_id,

text,

created_at

FROM timeline

WHERE created_at >= NOW() - INTERVAL 1 DAY AND user_id IN

(SELECT users.user_id FROM hashtag JOIN users

ON hashtag.user_id = users.user_id

WHERE hashtag.text = 'ballot drop off 📥 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj');

Relational Algebra

$R \leftarrow (\pi_{\text{users.user\_id}} (\sigma_{\text{hashtag.text =" ballot drop off 📥 (@ Mayfair Park) on \#Yelp https://t.co/HNnC5QjJpj"}} (\text{hashtag.user\_id} \bowtie \text{users.user\_id})))$

$\pi_{\text{timeline.text}} (\sigma_{\text{timeline.created\_at >= NOW()-INTERVAL 1 DAY ^ timeline.user\_id = R}} (\text{hashtag.user\_id} \bowtie \text{users.user\_id})$

-- 4. How many tweets have this user posted in past 24h

SELECT

COUNT(text)

FROM timeline

WHERE created_at >= NOW() - INTERVAL 1 DAY AND user_id IN

(SELECT users.user_id FROM hashtag JOIN users

ON hashtag.user_id = users.user_id

WHERE hashtag.text = 'ballot drop off 📥 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj');

Relational Algebra

$R \leftarrow (\pi_{\text{users.user\_id}} (\sigma_{\text{hashtag.text =" ballot drop off 📥 (@ Mayfair Park) on \#Yelp https://t.co/HNnC5QjJpj"}} (\text{hashtag.user\_id} \bowtie \text{users.user\_id})))$

$\pi_{\text{COUNT(timeline.text)}} (\sigma_{\text{timeline.created\_at >= NOW()-INTERVAL 1 DAY ^ timeline.user\_id = R}} (\text{hashtag.user\_id} \bowtie \text{users.user\_id})$

-- 5. when did this user join twitter?

SELECT

users.username,

users.created_at

FROM hashtag

JOIN users ON hashtag.user_id = users.user_id

WHERE hashtag.text = 'ballot drop off 📩 (@ Mayfair Park) on #Yelp https://t.co/HNnC5QjJpj';

Relational Algebra

$\pi_{\text{users.created\_at}}(\sigma_{\text{hashtag.text = '' ballot drop off 📩 (@ Mayfair Park) on \#Yelp https://t.co/HNnC5QjJpj ''}}(\text{hashtag.user\_id} \bowtie \text{users.user\_id}))$

-- 6. What tweets are popular?

SELECT

username,

text,

url,

likes,

retweets

FROM hashtag

JOIN users ON hashtag.user_id = users.user_id

WHERE likes >= 5 or retweets >= 5;

$\pi_{\text{hashtag.text}}(\sigma_{\text{hashtag.likes >= 5 v hashtag.retweets >=5}}(\text{hashtag.user\_id} \bowtie \text{users.user\_id}))$