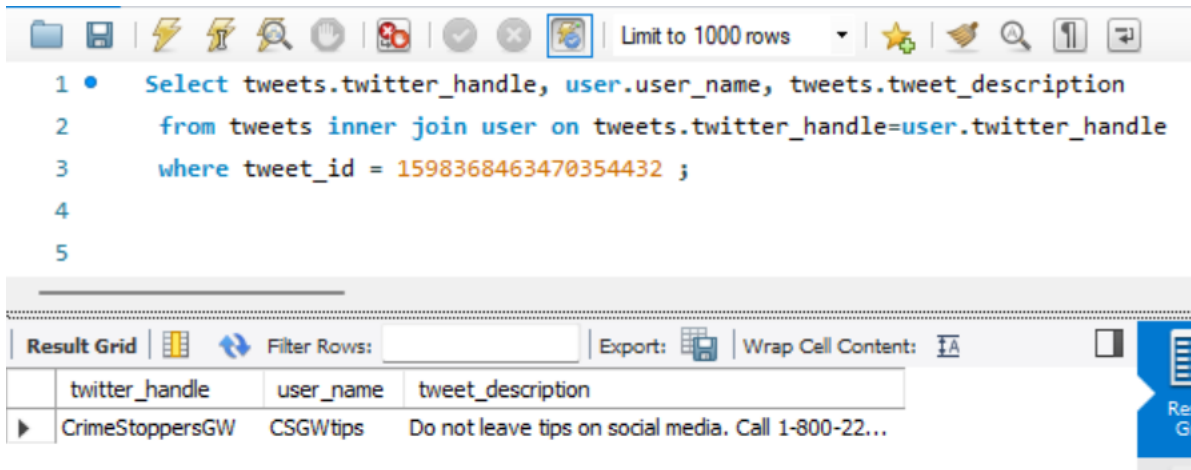


## SQL Queries and Relational Algebra answering physical model

1. What user posted this tweet?

Query1:

```
Select tweets.twitter_handle, user.user_name, tweets.tweet_description
from tweets inner join user on tweets.twitter_handle=user.twitter_handle
where tweet_id = 1598368463470354432;
```

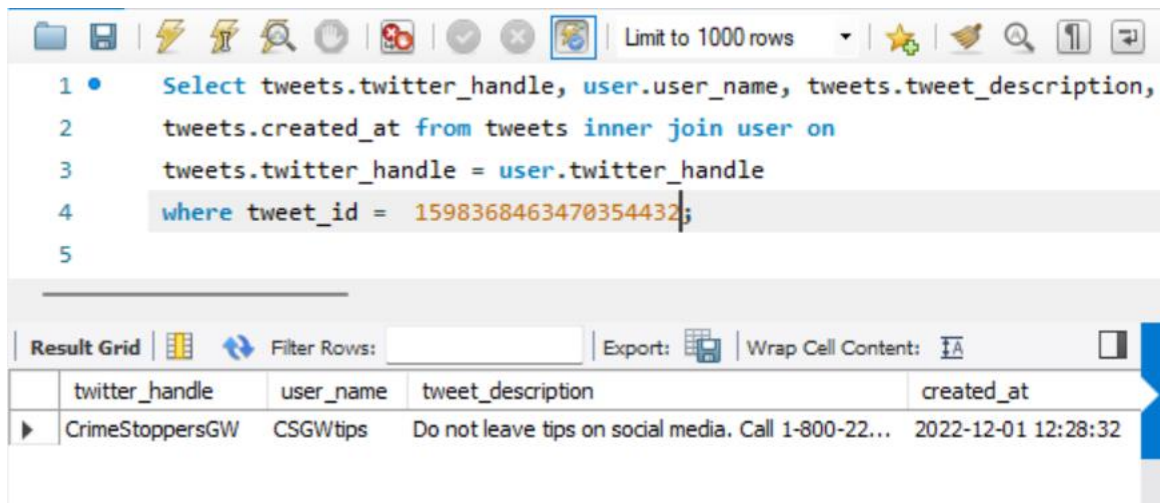


Relational Notation:  $\Pi_{\text{twitter\_handle, tweet\_text}}(\sigma_{\text{tweet\_id} = '1596997211048808448'}(\text{tweets}))$

2. When did the user post this tweet?

Query2:

```
Select tweets.twitter_handle, user.user_name, tweets.tweet_description,
tweets.created_at from tweets inner join user on
tweets.twitter_handle = user.twitter_handle
where tweet_id = 1596997211048808448;
```



1 • Select tweets.twitter\_handle, user.user\_name, tweets.tweet\_description,  
 2 tweets.created\_at from tweets inner join user on  
 3 tweets.twitter\_handle = user.twitter\_handle  
 4 where tweet\_id = 1598368463470354432;  
 5

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

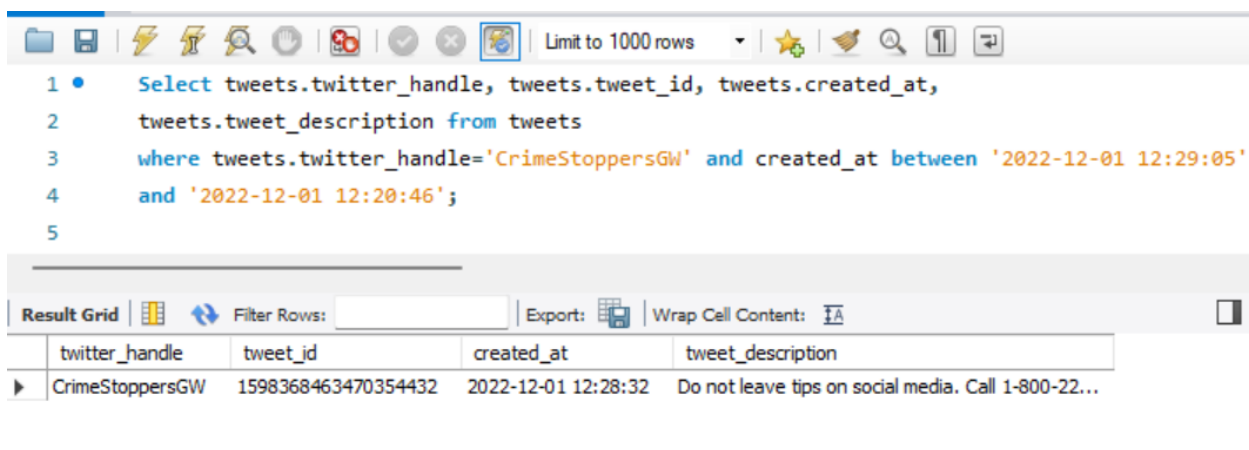
	twitter_handle	user_name	tweet_description	created_at
▶	CrimeStoppersGW	CSGWtips	Do not leave tips on social media. Call 1-800-22...	2022-12-01 12:28:32

Relational Notation:  $\Pi_{\text{twitter\_handle, tweet\_text, created\_at}}(\sigma_{\text{tweet\_id} = '1596997211048808448'}(\text{tweets}))$

3. What tweets have this user posted in the past 24 hours?

Query3:

Select tweets.twitter\_handle, tweets.tweet\_id, tweets.created\_at,  
 tweets.tweet\_description from tweets  
 where tweets.twitter\_handle='Patricia' and created\_at between '2022-11-11'  
 and '2022-11-12';



1 • Select tweets.twitter\_handle, tweets.tweet\_id, tweets.created\_at,  
 2 tweets.tweet\_description from tweets  
 3 where tweets.twitter\_handle='CrimeStoppersGW' and created\_at between '2022-12-01 12:29:05'  
 4 and '2022-12-01 12:20:46';  
 5

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	twitter_handle	tweet_id	created_at	tweet_description
▶	CrimeStoppersGW	1598368463470354432	2022-12-01 12:28:32	Do not leave tips on social media. Call 1-800-22...

Relational Notation:  $\Pi_{\text{twitter\_id, tweet\_text}}(\sigma_{\text{twitter\_handle} = '@smith' \cap \text{created\_at between ('2022-11-11) and ('2022-11-12))}(\text{Tweets}))$

4. How many tweets have this user posted in the past 24 hours?

Query4:

Select count(tweet\_id) as NoOfTweets from tweets where  
 twitter\_handle= 'Patricia' and created\_at between '2022-11-11'  
 and '2022-11-12';

```

1 • Select count(tweet_id) as NoOfTweets from tweets where
2   twitter_handle= 'CrimeStoppersGW' and created_at between '2022-11-11'
3   and '2022-11-12';
4
5

```

Result Grid

NoOfTweets
1

Relational Notation:  $\rho$  NoOfTweets ( $\Pi$  count(tweet\_id) ( $\sigma$   
 twitter\_handle='Patricia'  $\cap$  created\_at between (SYSDATE-1) and  
 SYSDATE ())) (Tweets))

5. When did this user join twitter?

Query5:

Select u.twitter\_handle, u.user\_name, u.join\_date from tweets t, user u  
 where u.twitter\_handle=t.twitter\_handle and u.twitter\_handle= 'Patricia';

```

1 • Select u.twitter_handle, u.user_name, u.join_date from tweets t, user u
2   where u.twitter_handle=t.twitter_handle and u.twitter_handle= 'CrimeStoppersGW';
3
4
5

```

Result Grid

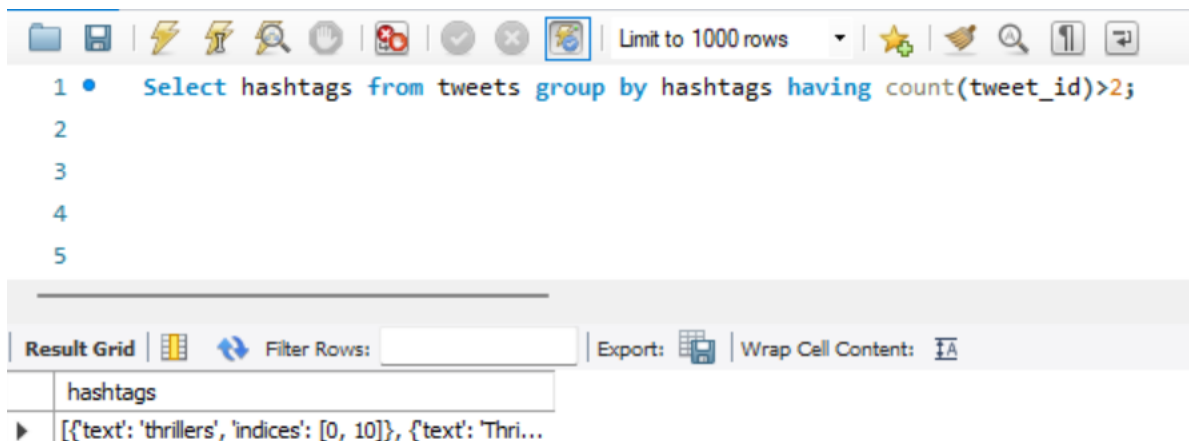
twitter_handle	user_name	join_date
CrimeStoppersGW	CSGWtips	2014-09-03 09:15:37

Relational Notation:  $\Pi$  (u.twitter\_handle, u.join\_date)  $\sigma$  (u.twitter\_handle =  
 "smith") (User  $\bowtie$  Tweets  $u$ .twitter\_handle=t.twitter\_handle )

6. What keywords/hashtags are popular?

Query6:

Select hashtags from tweets group by hashtags having count(tweet\_id)>2;



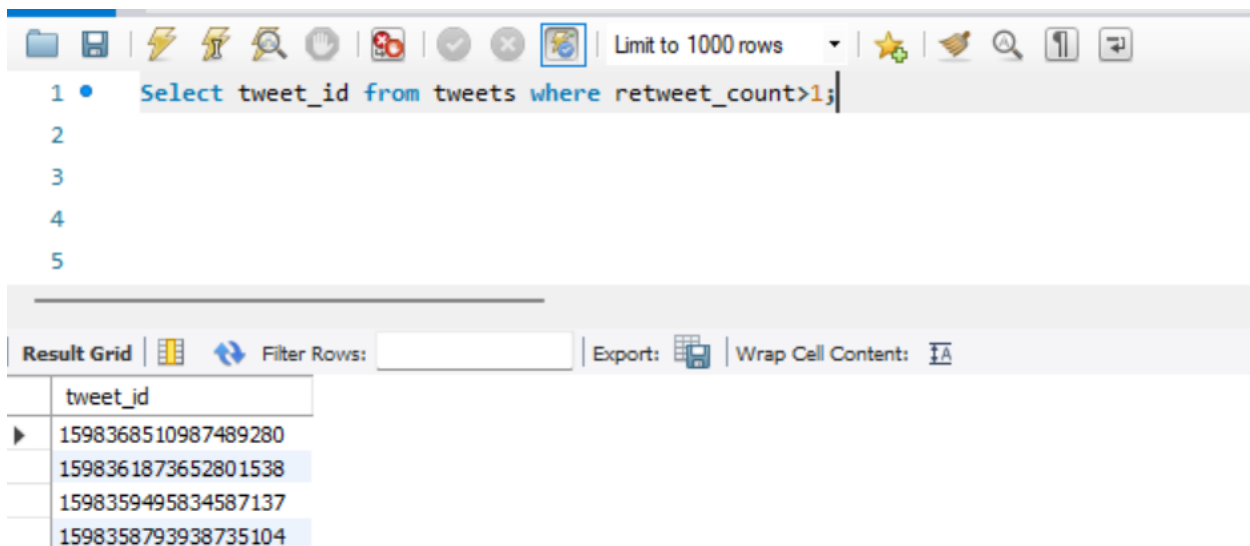
The screenshot shows a SQL query editor with a toolbar at the top containing icons for file operations, execution, and search. The query text is: `1 • Select hashtags from tweets group by hashtags having count(tweet_id)>2;`. Below the query editor, the 'Result Grid' is displayed with a single column header 'hashtags'. The first row of data is truncated: `[{'text': 'thrillers', 'indices': [0, 10]}, {'text': 'Thri...`. The interface includes a 'Filter Rows' input field, an 'Export' button, and a 'Wrap Cell Content' toggle.

Relational Notation: Not possible as Group By function does not exist in relational algebra.

7. What tweets are popular?

Query7:

Select tweet\_id from tweets where retweet\_count>1;



The screenshot shows a SQL query editor with a toolbar at the top. The query text is: `1 • Select tweet_id from tweets where retweet_count>1;`. Below the query editor, the 'Result Grid' is displayed with a single column header 'tweet\_id'. The first four rows of data are: `1598368510987489280`, `1598361873652801538`, `1598359495834587137`, and `1598358793938735104`. The interface includes a 'Filter Rows' input field, an 'Export' button, and a 'Wrap Cell Content' toggle.

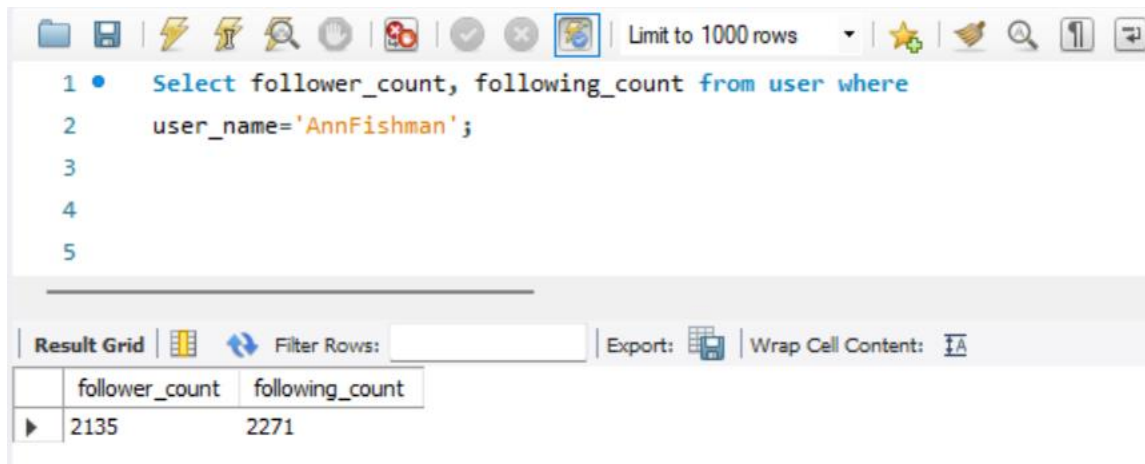
Relational Notation:  $\pi$  (tweet\_id)(  $\sigma$  retweet > 60 ) (tweets)

8. What is the followers and following count for this user?

Query8:

Select follower\_count, following\_count from user where

user\_name='Patricia';



The screenshot shows a SQL query editor interface. The query is: `Select follower_count, following_count from user where user_name='AnnFishman';`. The result grid below shows one row of data.

	follower_count	following_count
▶	2135	2271

Relational Notation:  $\Pi$  (tweet\_id) (  $\sigma$  retweet > 60 ) (tweets)

9. Who are the influential users and what are the tweets posted by them?

Query9:

Select distinct t.twitter\_handle, t.tweet\_description from user u, tweets t  
where user\_type = 'True';


Limit to 1000 rows

```

1 • Select distinct t.twitter_handle, t.tweet_description from user u, tweets t
2   where user_type = 'True';
3
4
5
6
7

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [iA](#)

twitter_handle	tweet_description
CJ Bowen	If it has anything to do with people and commu...
The Daily Star	Bangladesh's largest English Daily. The latest & ...
Ann Fishman	Ann A. Fishman is an expert in providing insight...
Gurbaksh Singh Chahal	CEO, @Procurenent, @Medriva, @BNNBreaking ...
CrimeStoppersGW	Do not leave tips on social media. Call 1-800-22...
DR.	Lifelong Buckeye, former hotelier, University Ad...
CrimeSafetyTips	
Real Talk B	~ Bringing you the latest news & entertainment...
Jason S  us	Houstonian. Local Intermodal Truck Driver. Mod...
PartnersInCrimeTours	#PICT Tours is a #VBT company who helps #cri...
Dee	Here for laughs
Self-Publishers' Showc...	The best books, from the best #Indieauthors, s...
Anne Williams	Happily retired, enjoying reading + sharing my l...
nigel maxwell	Senior reporter paNOW, wedding dj, CWE ring ...
IdeallyaNews	#WorldNews, compiled from public broadcaster...
Christopher Oldcorn (...)	Journalist and Opinion Writer @WSOnlineNews ...
Mr. L	Live and work in the southern part of the U.S.A...
True Crime Lists	Informative and interesting true crime facts! Tr...

Result 77 x

10. List the top 5 locations based on the tweets posted ?

Query10:

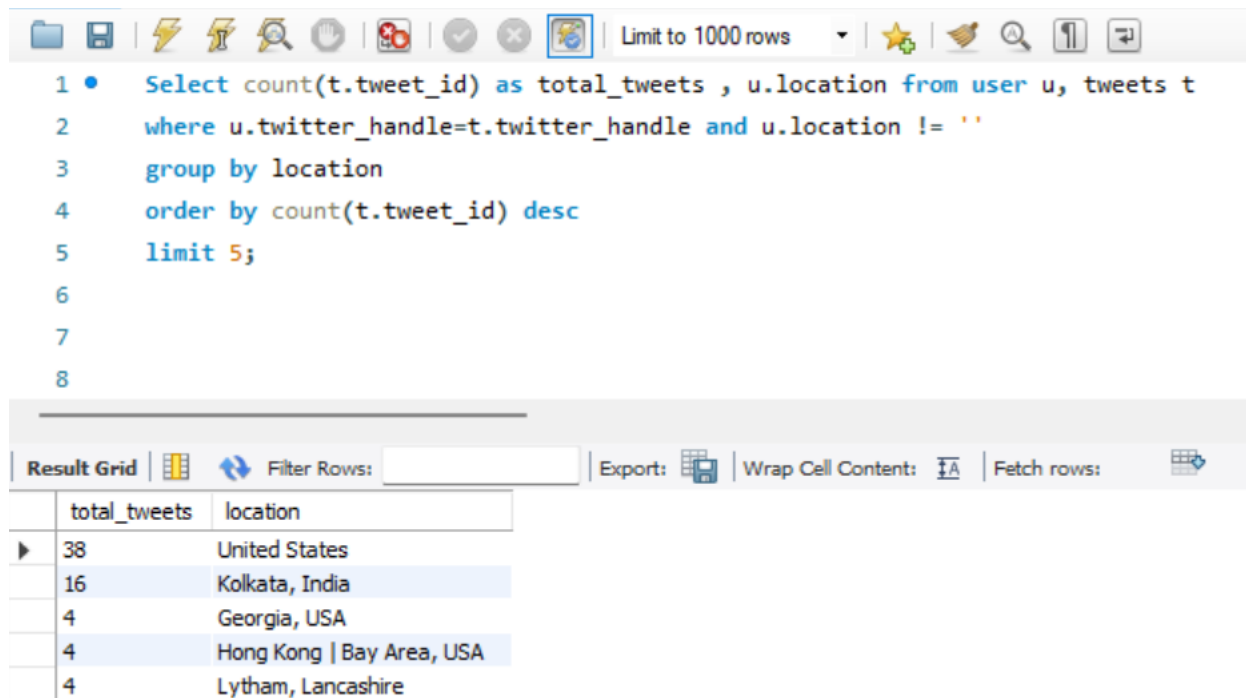
Select count(t.tweet\_id) as total\_tweets , u.location from user u, tweets t

where u.twitter\_handle=t.twitter\_handle and u.location != ''

group by location

order by count(t.tweet\_id) desc

limit 5;



```

1 • Select count(t.tweet_id) as total_tweets, u.location from user u, tweets t
2   where u.twitter_handle=t.twitter_handle and u.location != ''
3   group by location
4   order by count(t.tweet_id) desc
5   limit 5;
6
7
8

```

	total_tweets	location
▶	38	United States
	16	Kolkata, India
	4	Georgia, USA
	4	Hong Kong   Bay Area, USA
	4	Lytham, Lancashire

Relational Notation: Not possible as Group By function does not exist in relational algebra.

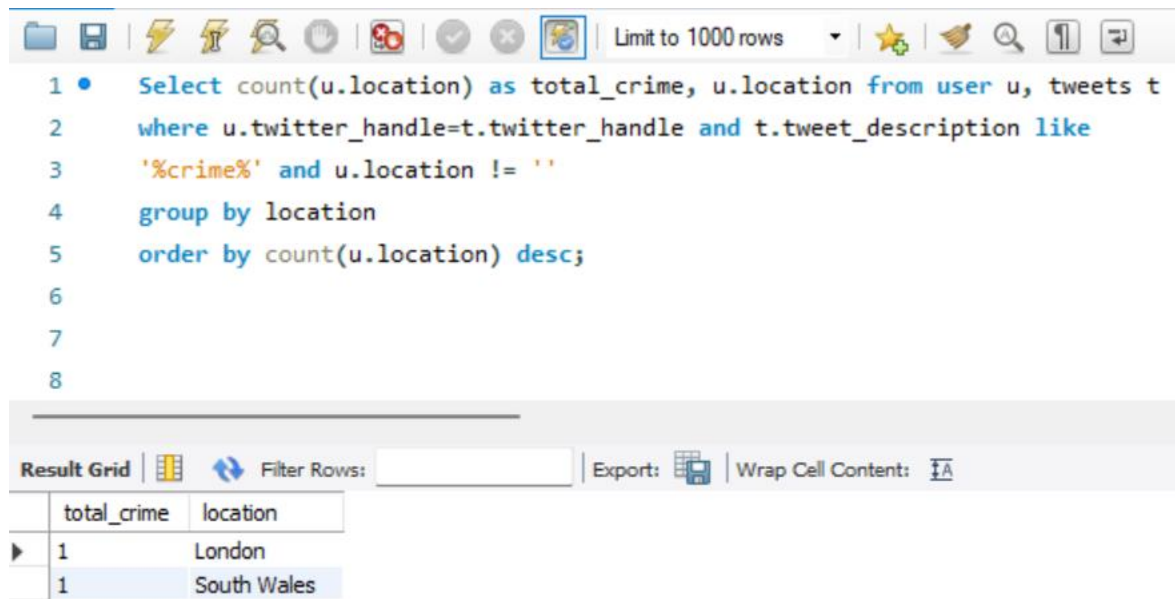
11. List the locations with number of crime incidents reported?

Query11:

```

Select count(u.location) as total_crime, u.location from user u, tweets t
where u.twitter_handle=t.twitter_handle and t.tweet_description like
'%crime%' and u.location != ''
group by location
order by count(u.location) desc;

```



1 • Select count(u.location) as total\_crime, u.location from user u, tweets t  
2 where u.twitter\_handle=t.twitter\_handle and t.tweet\_description like  
3 '%crime%' and u.location != ''  
4 group by location  
5 order by count(u.location) desc;  
6  
7  
8

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	total_crime	location
▶ 1	1	London
1	1	South Wales

Relational Notation: Not possible as Group By function does not exist in relational algebra.