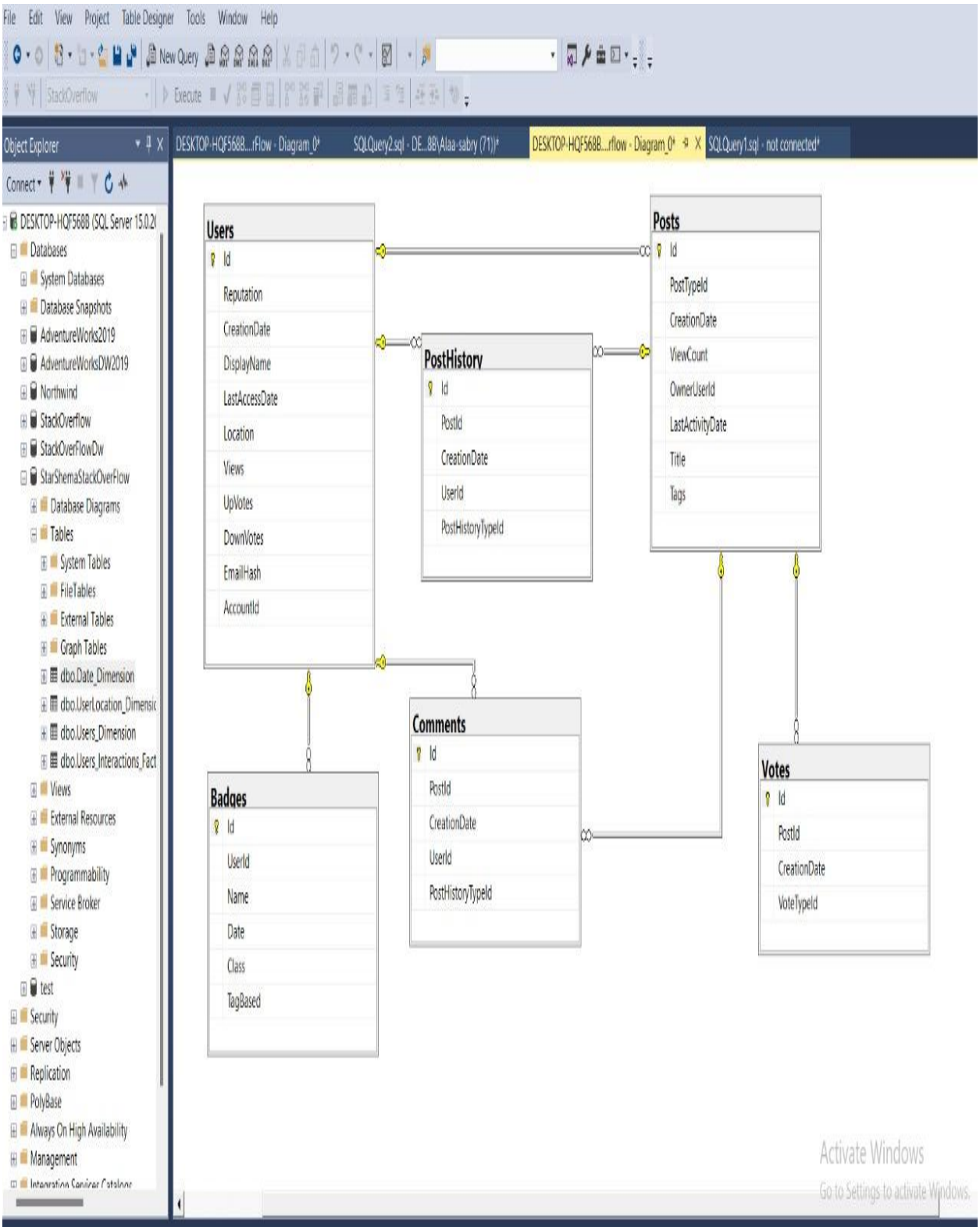
names	id	emails	Group
Hassan Ragab	20190177	hassanragab776@gmail.com	5
Alaa Sabry	20190335	am5814476@gmail.com	5
Zainab medhat	20170454		
Nouran ahmed	20190596		
Shammakh salah	20190793		

DB Source Name: Stack Over Flow

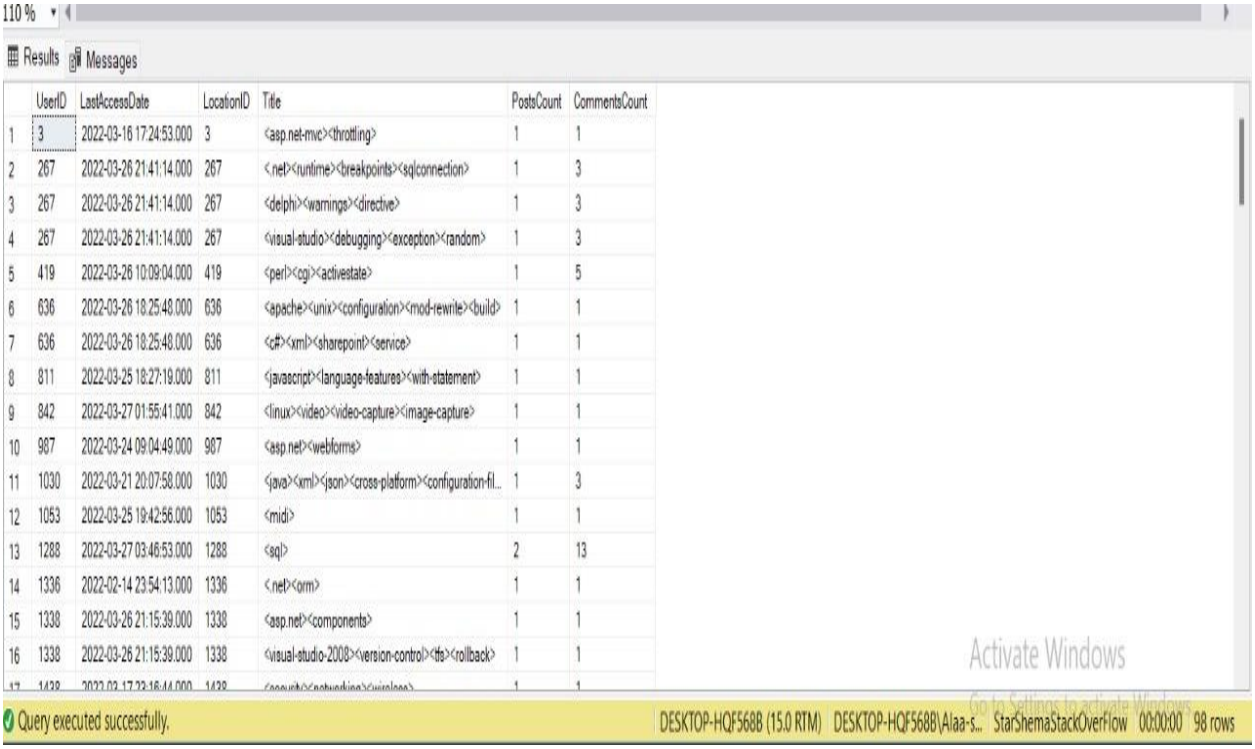
DB Source Link : https://data.stackexchange.com/stackoverflow/query/new

Stack OverFlow ERD



The motivation:

We're making the Stack Over Flow star schema in order to analyse and count the amount of posts and comments that a user made at a certain time and in a specific location on a specific topic. We'd be able to examine how users interact with one another in a published challenge over a period of time. The fact table shown below allows us to examine each user's actions over time.



110 %

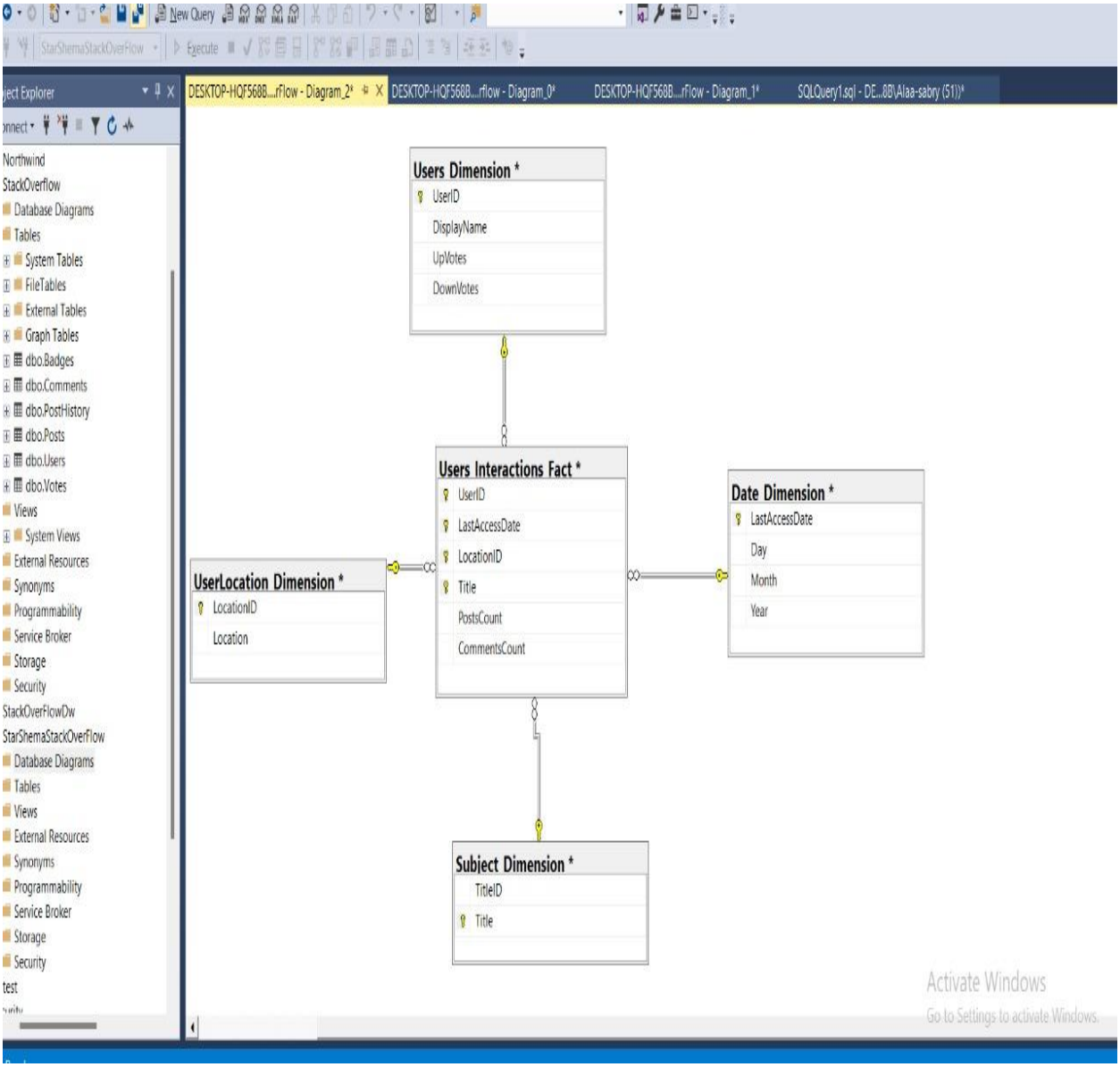
Results | Messages

	UserID	LastAccessDate	LocationID	Title	PostsCount	CommentsCount
1	3	2022-03-16 17:24:53.000	3	<asp.net-mvc><throttling>	1	1
2	267	2022-03-26 21:41:14.000	267	<.net><runtime><breakpoints><sqlconnection>	1	3
3	267	2022-03-26 21:41:14.000	267	<delphi><warnings><directive>	1	3
4	267	2022-03-26 21:41:14.000	267	<visual-studio><debugging><exception><random>	1	3
5	419	2022-03-26 10:09:04.000	419	<perl><cgi><activestate>	1	5
6	636	2022-03-26 18:25:48.000	636	<apache><unix><configuration><mod-rewrite><build>	1	1
7	636	2022-03-26 18:25:48.000	636	<c#><xml><sharepoint><service>	1	1
8	811	2022-03-25 18:27:19.000	811	<javascript><language-features><with-statement>	1	1
9	842	2022-03-27 01:55:41.000	842	<linux><video><video-capture><image-capture>	1	1
10	987	2022-03-24 09:04:49.000	987	<asp.net><webforms>	1	1
11	1030	2022-03-21 20:07:58.000	1030	<java><xml><json><cross-platform><configuration-fil...	1	3
12	1053	2022-03-25 19:42:56.000	1053	<midi>	1	1
13	1288	2022-03-27 03:46:53.000	1288	<sql>	2	13
14	1336	2022-02-14 23:54:13.000	1336	<.net><orm>	1	1
15	1338	2022-03-26 21:15:39.000	1338	<asp.net><components>	1	1
16	1338	2022-03-26 21:15:39.000	1338	<visual-studio-2008><version-control><tfs><rollback>	1	1

Activate Windows

Query executed successfully. DESKTOP-HQF568B (15.0 RTM) DESKTOP-HQF568B\Alaa-s... StarShemaStackOverFlow 00:00:00 98 rows

Star Schema Model



Schema description

1) Dimensions

We decided to make for dimensions_ Date dimension, Users dimensions, User Location dimension and Subject dimension, each one of them has attributes related to the schema`s measures as shown above.

a) Subject_Dimension

select Id as TitleID , Tags as Title from posts

Transform posts data (Id, Tags) to Subject dimension (TitleId, Title)

b) UserLocation_Dimension

select Users.Id as LocationID , Location from Users

Transform user location data (Users.Id, Location) to UserLocation dimension (LocationID, Location)

c) Users_Dimension

select Users.Id as UserID , DisplayName ,UpVotes , DownVotes from Users

Transform user data (Users.Id, DisplayName, UpVotes, DownVotes) to Users dimension (UserId, Name , UpVotes, DownVotes)

d) Date_Dimension

Select distinct (LastAccessDate), DAY(LastAccessDate) AS Day, MONTH(LastAccessDate) AS Month, YEAR(LastAccessDate) AS Year FROM Users Order by LastAccessDate

Transform Users date(LastAccessDate DAY(LastAccessDate), MONTH(LastAccessDate) YEAR(LastAccessDate)) to the Date dimension(LastAsccessDate, Day, Month, Year). The dimension levels are day, month year.

2) Fact Table

- Users_Interaction_Fact

```
select UserID1 AS UserID ,Date1 as LastAccessDate , LocationID ,Subject
,PostsCount,CommentsCount from

(select Posts.OwnerUserId as UserID1 , Posts.Tags as Subject ,Users.Location as
location1 , Users.LastAccessDate as Date1 ,count(Posts.Tags) as PostsCount

from Posts , Users

where Posts.Tags != 'NULL'

Group by OwnerUserId , Posts.Tags,Users.Location ,Users.LastAccessDate) x1

INNER JOIN

(select Users.Id as UserID2 , Users.Id as LocationID , Users.Location as
location2, Users.LastAccessDate as Date2 , COUNT(Comments.CreationDate)
AS CommentsCount

from Users INNER JOIN Comments ON Comments.UserId = Users.Id

Group by Users.Id ,Users.LastAccessDate, Users.Location) x2

ON x1.UserID1 = x2.UserID2 and x1.location1 = x2.location2 and x1.Date1 =
x2.Date2
```

In the fact table, we have a foreign key for each dimension table, and we have two measures (postCount and commentCount) to be able to analyze the user actions.