**Graph Analytics**

**Modeling Chat Data Using a Graph Data Model**

The graph model is a network based on chat interactions between users. A chat session can be initiated by a user, other users on the same team are able to join and leave the session. Interactions between users begins when a user create a post. It’s possible for a user, mention another user. All relationship between entities are logged with a timestamp.

**Creation of the Graph Database for Chats**

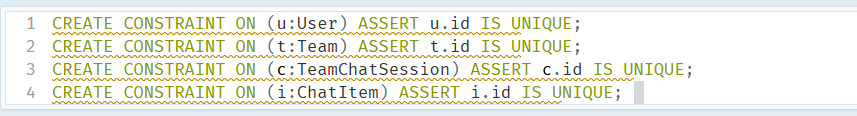
Describe the steps you took for creating the graph database

Write the schema of the 6 CSV files

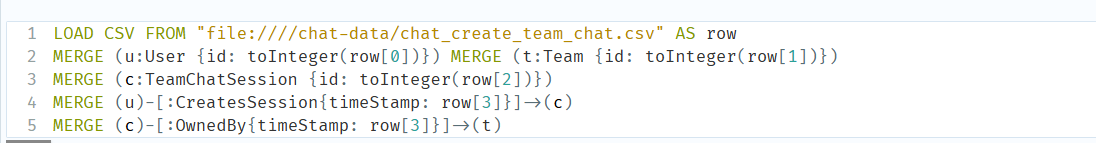
|  |  |
| --- | --- |
| **File** | **Entity** |
| Chat\_create\_team\_chat.csv | userID  teamID  teamChatSessionID  timestamp |
| Chat\_join\_team\_chat.csv | userID  teamChatSessionID  timestamp |
| Chat\_leave\_team\_chat.csv | userID  teamChatSessionID  timestamp |
| Chat\_item\_team\_chat.csv | userID  teamChatSessionID  chatItemID  timestamp |
| Chat\_mention\_team\_chat.csv | ChatItemID  UseID  timestamp |
| Chat\_respons\_team\_chat.csv | ChatItemID\_1  ChatItemID\_2  timestamp |

**Explain the loading process and include a sample LOAD command**

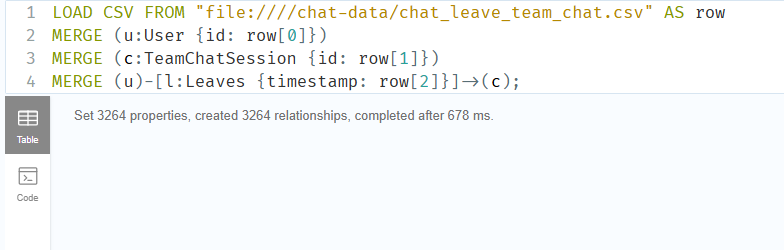
Creating constraints

****

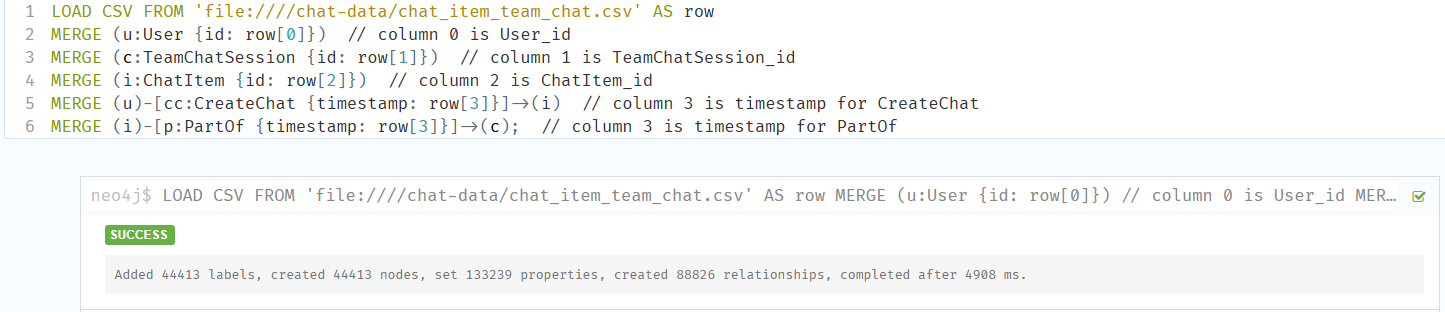
Loading chat\_create\_team\_chat

****

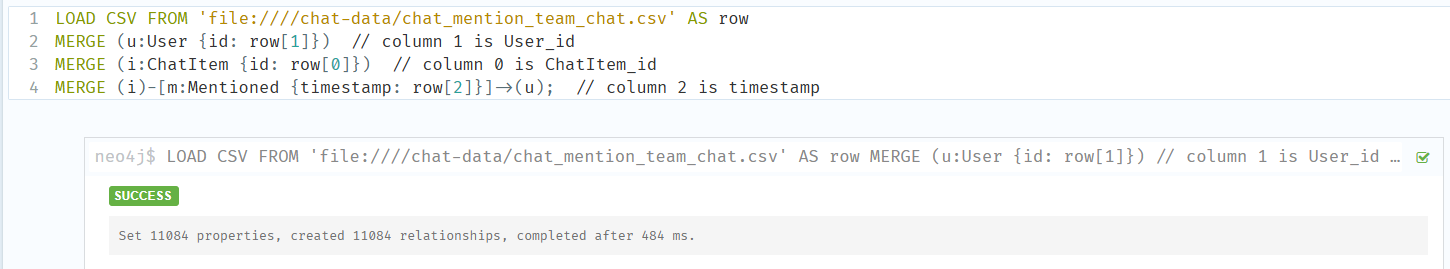
Loading chat\_leave\_team\_chat

****

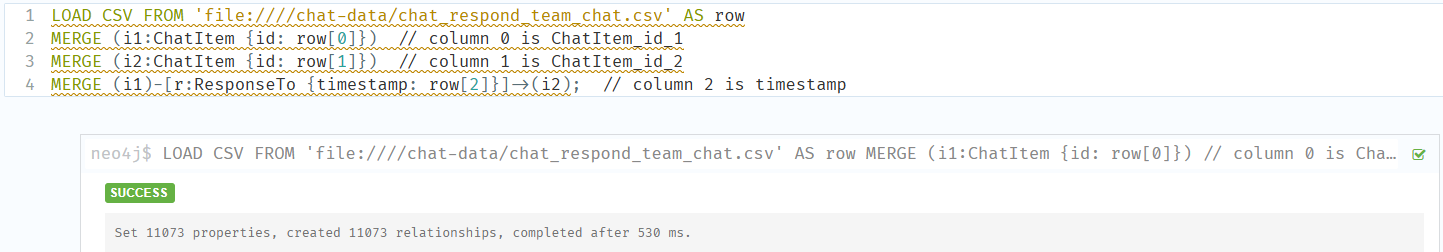
**Loading chat\_item\_team\_chat**

****

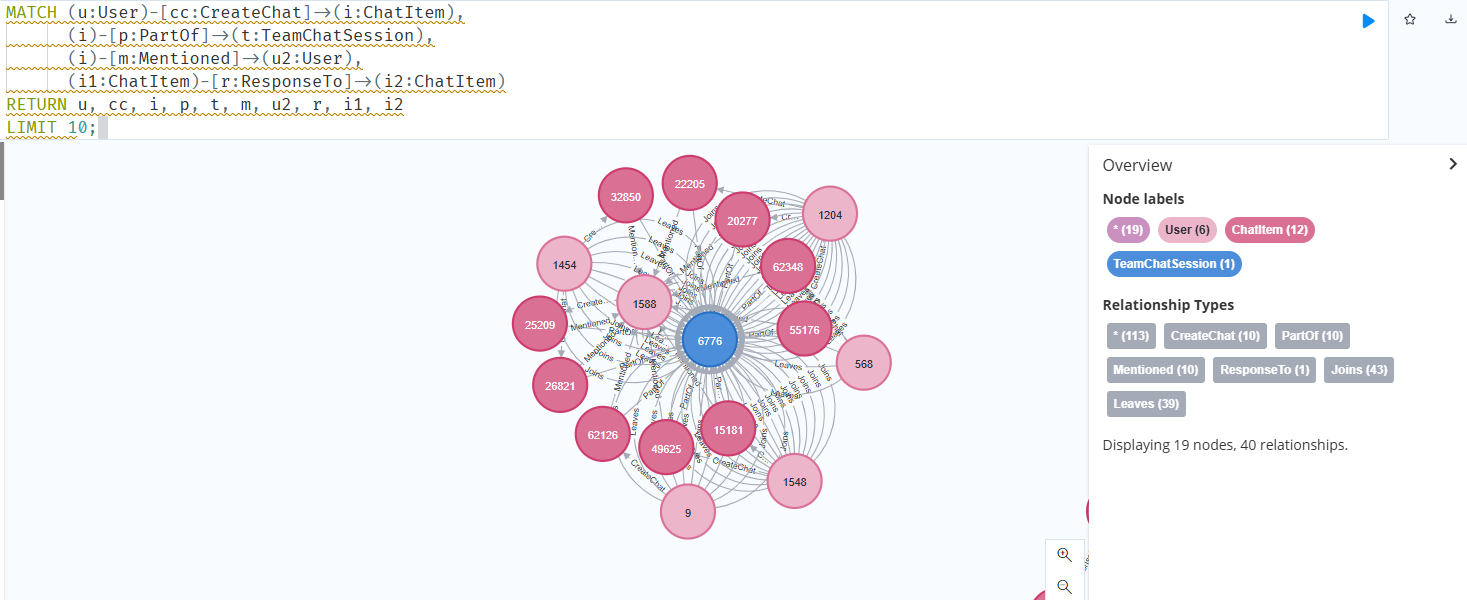
Loading chat\_mention\_team\_chat

****

Loading Chat\_resoibd\_team\_chat

****

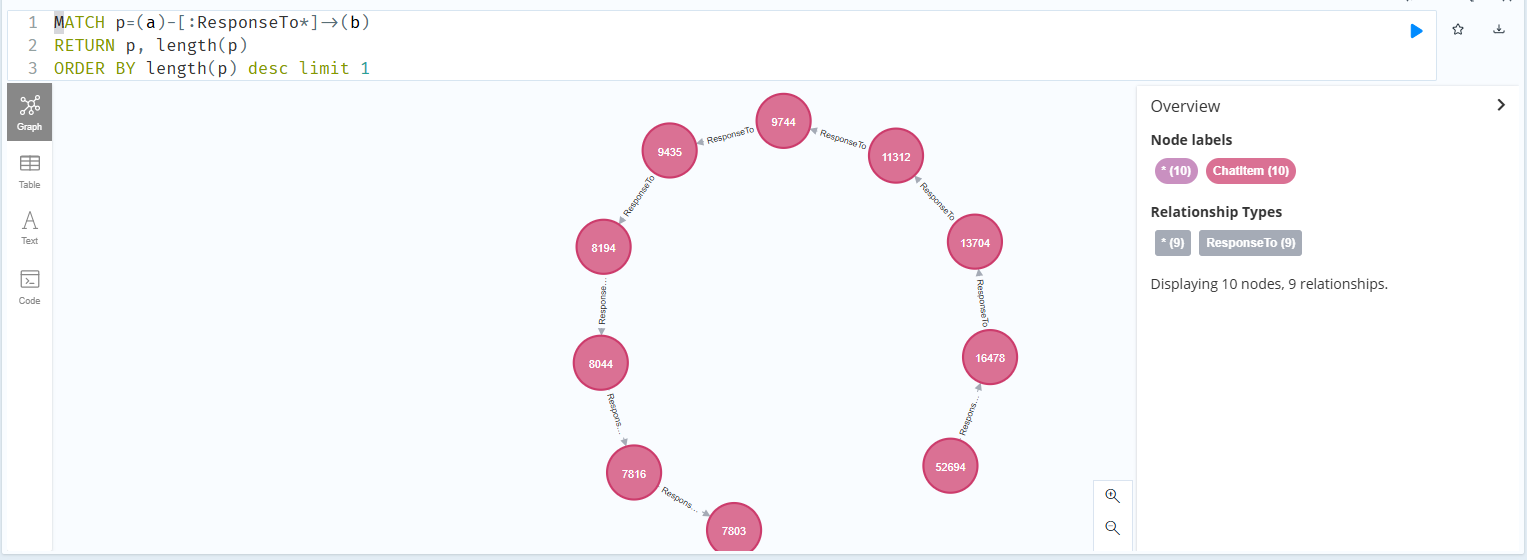
**Present a screenshot of some part of the graph you have generated. The graphs must include clearly visible examples of most node and edge types.**

****

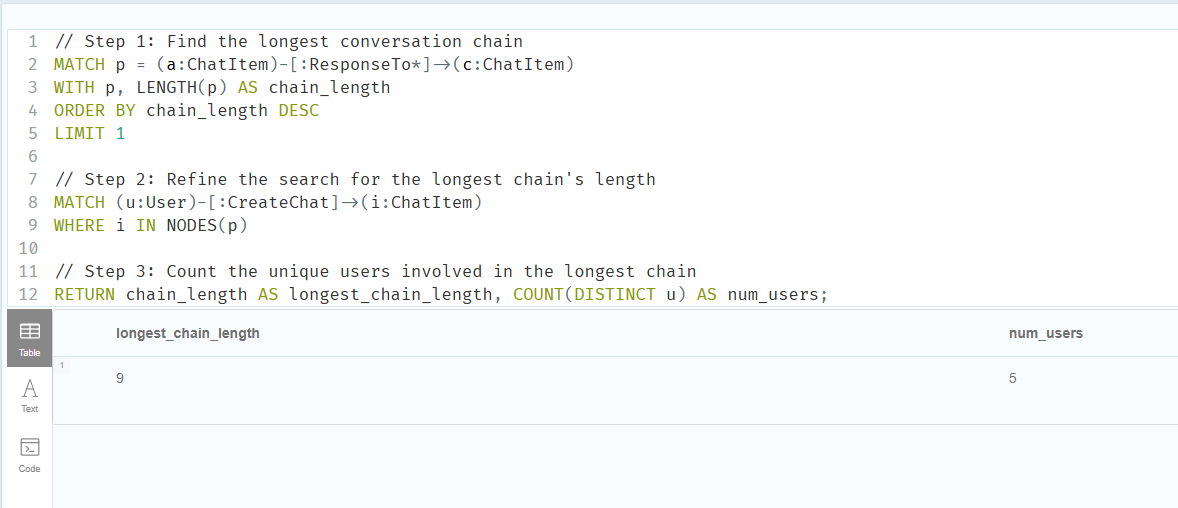
**Finding the longest conversation chain and its participants**

Report the results including the length of the conversation and how many unique users were part of the conversation chain. Describe the steps. Write the query produces the correct answer

**How many chats are involved in it?**

****

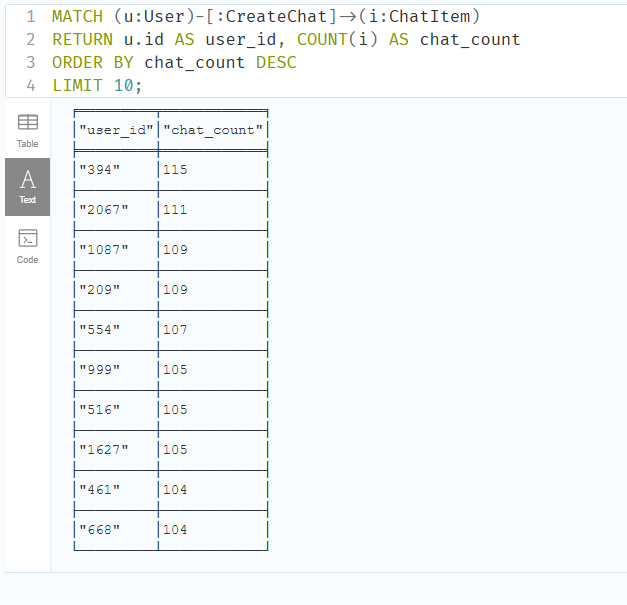
**How many users participated in this chain?**

****

**Analyzing the relationship between top 10 chattiest users and top 10 chattiest teams**

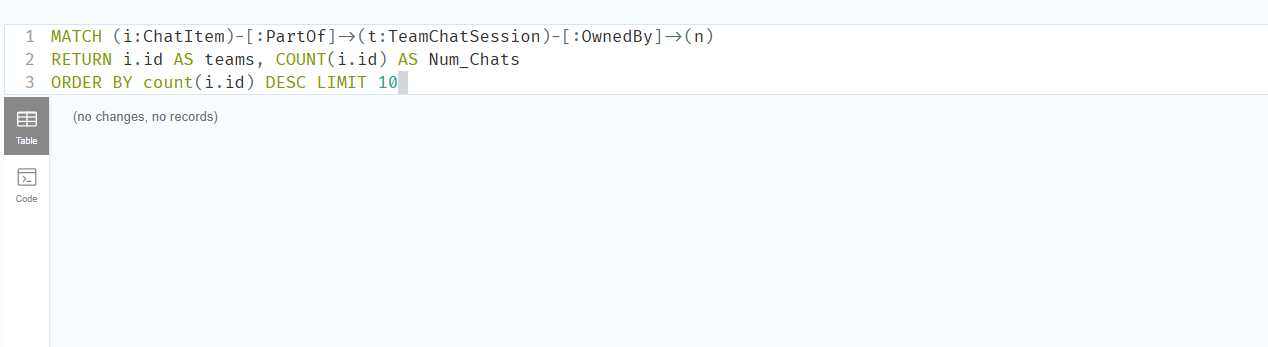
Describe your steps from Question 2. In the process, create the following two tables. You only need to include the top 3 for each table. Identify and report whether any of the chattiest users were part of any of the chattiest teams.

**Chattiest Users**Determine the number of chats created by a user from the CreateChat edge

****

**Chattiest Teams**

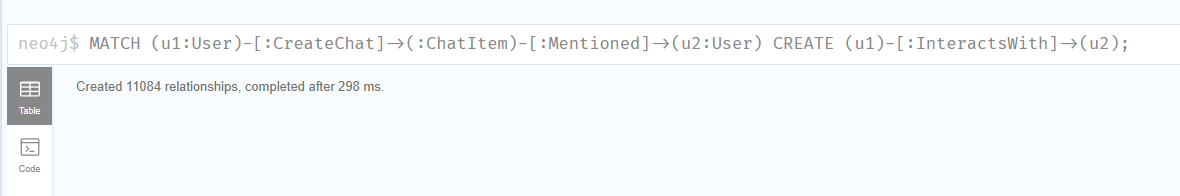
Match all ChatItem with a PartOd edge and connect them with a TeamChatSession node that have an OwnedBy edge connection them with any other node.

**  
**

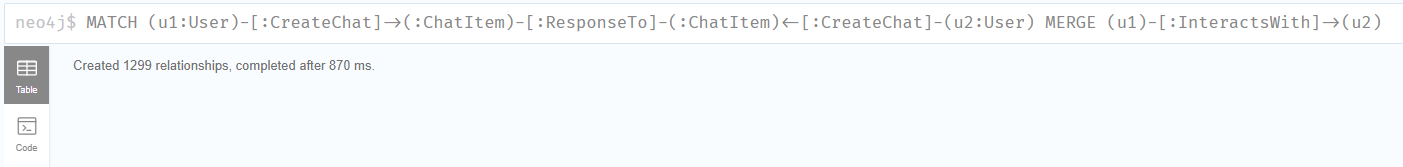
**How Active Are Groups of Users?**

In this question, we will compute an estimate of how “dense” the neighborhood of a node is. In the context of chat that translates to how mutually interactive a certain group of users are. If we can identify these highly interactive neighborhoods, we can potentially target some members of the neighborhood for direct advertising. We will do this in a series of steps.

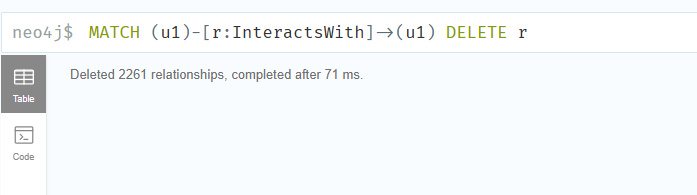
**Connect mentioned users**



**Connect user’s responses with the chat creator**

****

**Eliminate all self-interaction**

****

**Calculate the cluster coefficient.**

****

**Most Active Users**

|  |  |
| --- | --- |
| **User ID** | **Coefficient** |
| **394** | **0.9167** |
| **2067** | **0.7679** |
| **209** | **0.9524** |