|  |  |
| --- | --- |
| نام و نام خانوادگی | **میثاق محقق** |
| شماره دانشجویی | **810199484** |
| تاریخ ارسال گزارش | **1402.08.13** |

|  |  |  |
| --- | --- | --- |
|  | **به نام خدا**  **دانشگاه تهران**  **دانشکده‌ مهندسی برق و کامپیوتر** |  |
| **درس آزمایشگاه پایگاه‌داده**  **دستورکار سوم** | | |

**فهرست**

[**پایگاه‌داده نمونه Twitter** 2](#_Toc150032090)

[1-0. Connect 2](#_Toc150032091)

[1-1. Play 2](#_Toc150032092)

[1-2. Graph your Application 3](#_Toc150032093)

[1-3. Your Mentions 3](#_Toc150032094)

[2-3. As Table 4](#_Toc150032095)

[1-4. Most Influential Followers 4](#_Toc150032096)

[1-5. Most Tagged 5](#_Toc150032097)

[1-6. Followback Rate 5](#_Toc150032098)

[1-7. Follower Recommendations 6](#_Toc150032099)

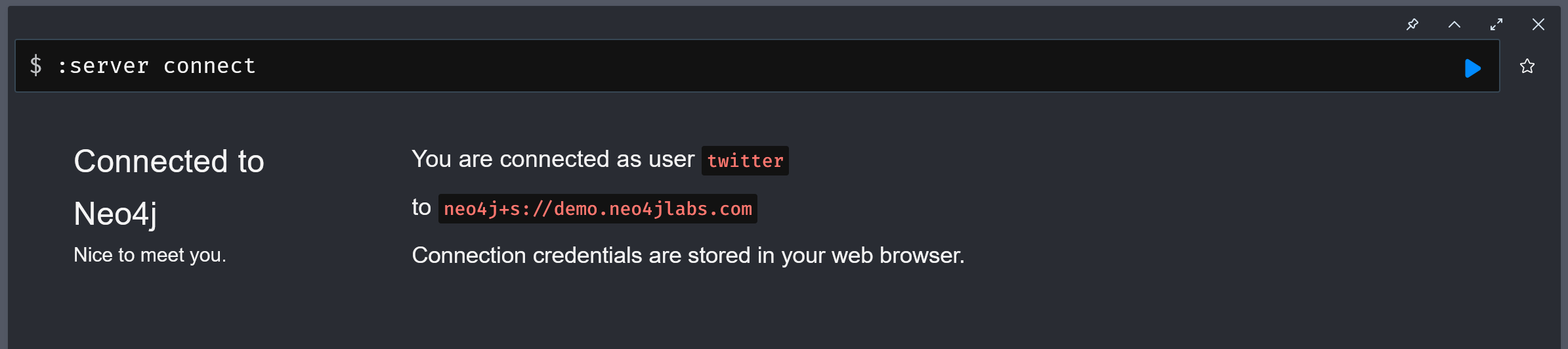
[1-8. Links from Interesting Retweets 7](#_Toc150032100)

[1-9. Common Hashtags 7](#_Toc150032101)

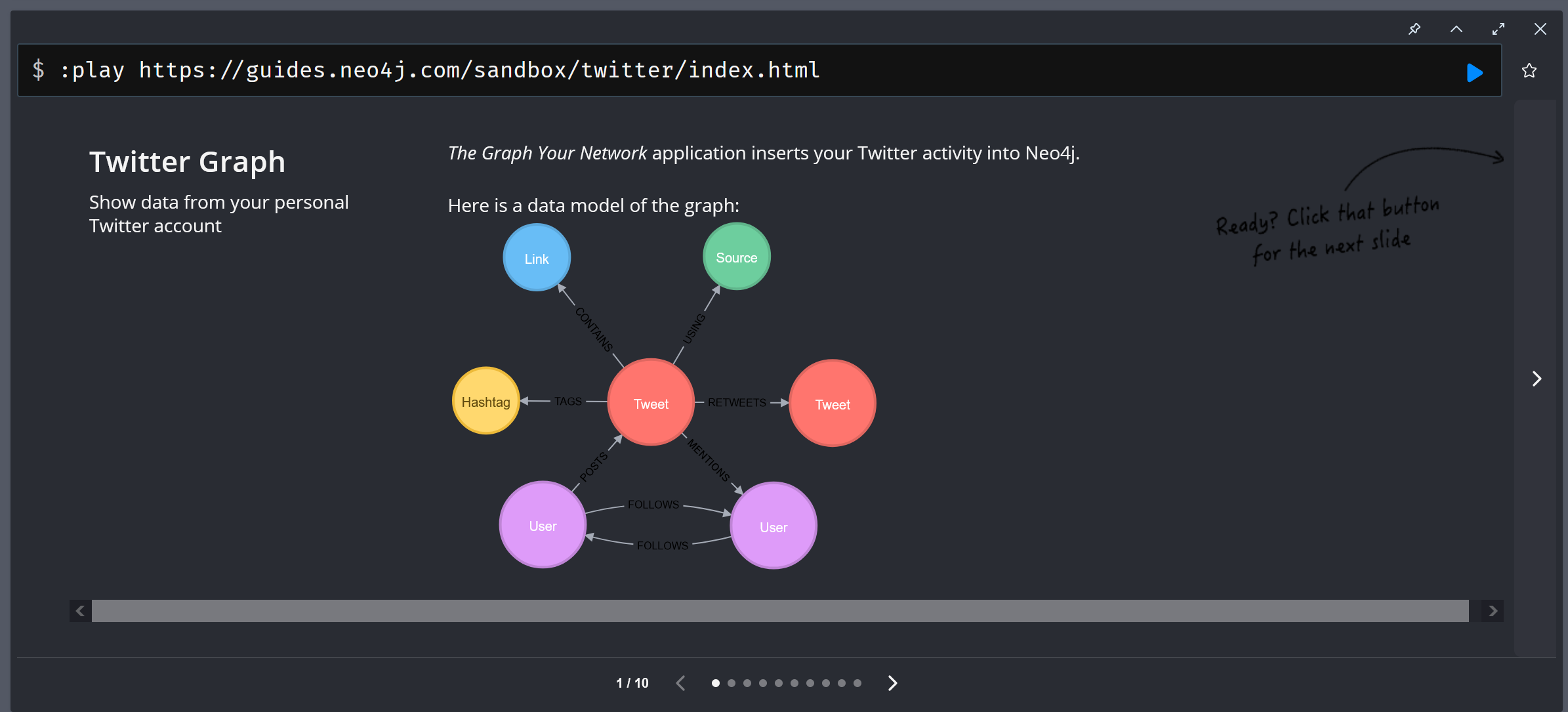
# **پایگاه‌داده نمونه Twitter**

## **1-0. Co**nne**ct**

به لینک داده شده متصل می‌شویم و رمز twitter را وارد می‌کنیم:

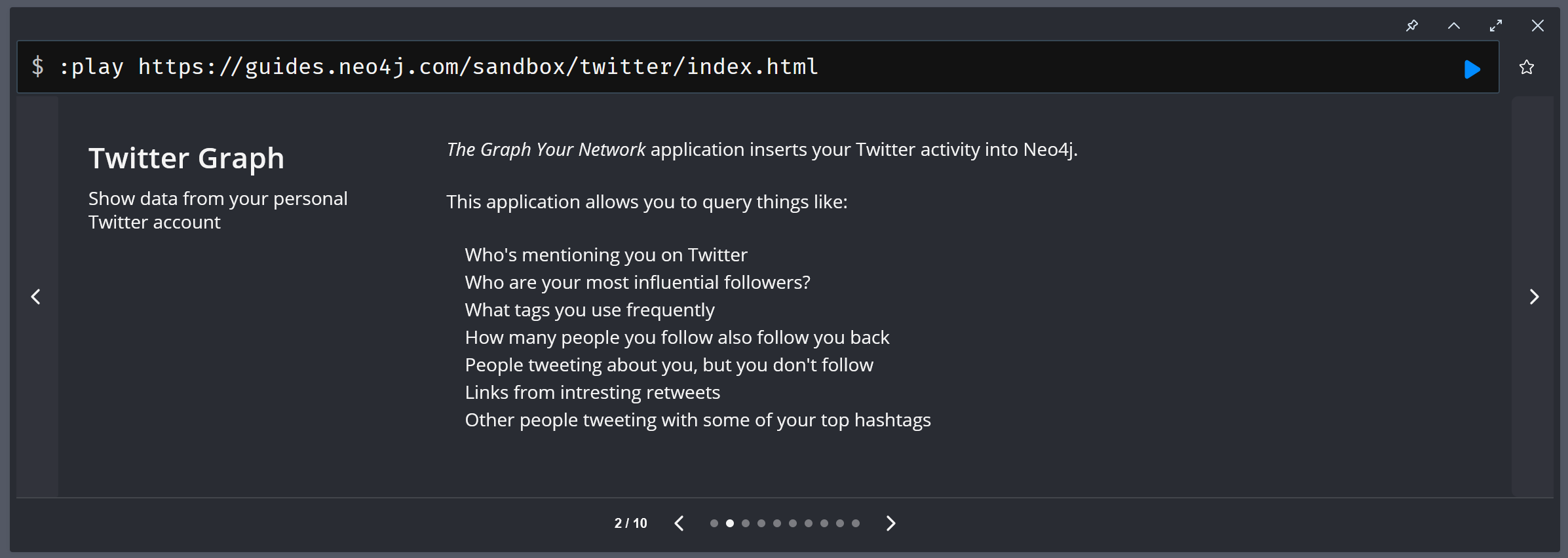


## 1-1. Play



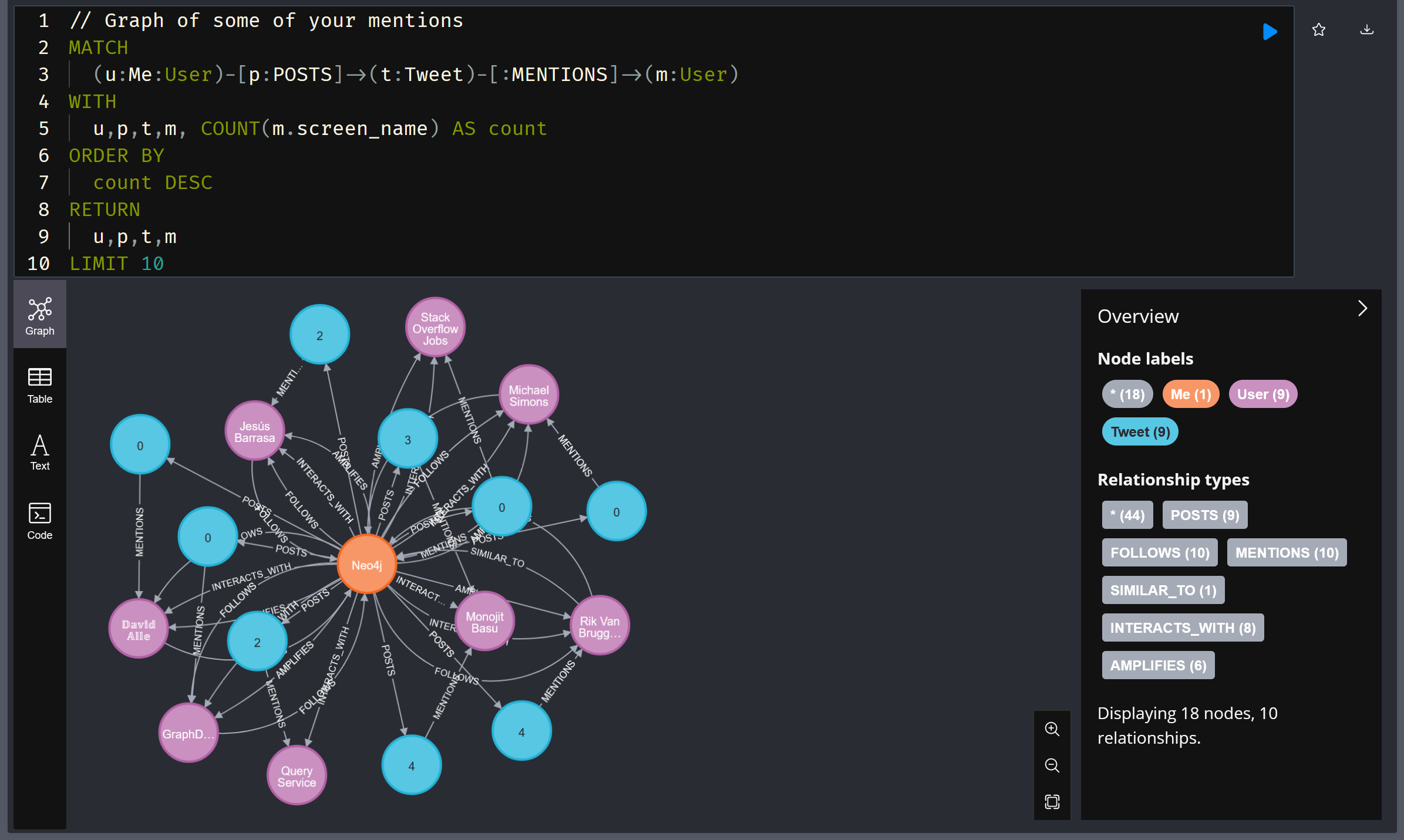
در اینجا می‌توانیم مدل گرافی دیتاها را مشاهده کنیم. نودهای این دیتابیس از چند label تشکیل شده اند که شامل توییت، لینک، هشتگ، کاربر و منبع ارسال است. طبق گراف، یک توییت می‌تواند ری‌توییت باشد، حاوی لینک بوده و هشتگ‌هایی داشته باشد. همچنان می‌توان در یک توییت یک کاربر را mention کرد و کاربرها می‌توانند همدیگر را follow کنند. پست‌های کاربران شامل توییت‌های آنها است.

## 1-2. Graph your Application



در این صفحه مثال‌هایی از کوئری‌هایی که می‌توان انجام داد را نوشته است.

## 1-3. Your Mentions



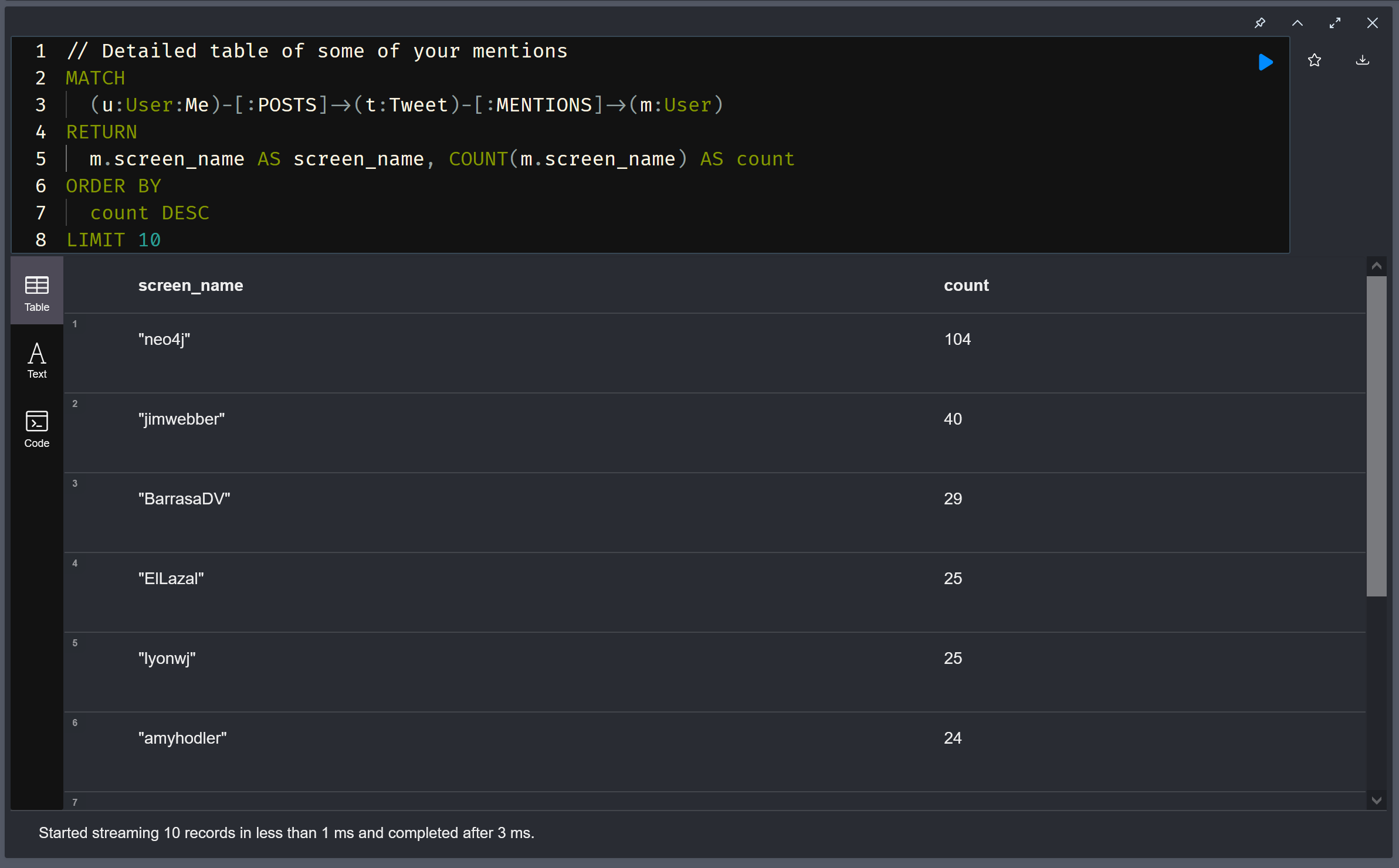
این Cypher ابتدا u که نودی است که هم لیبل Me دارد و هم لیبل User (در این دیتابیس فقط یکی هست) را گرفته و همه توییت‌هایی که Me زده و یک کاربر دیگری را داخل آن mention می‌کند را دریافت می‌کند. سپس group by کرده و توییت‌ها را بر حسب تعداد منشن‌های داخلش سورت می‌کند.

در خروجی، از هر label حداکثر 10 تا آمده که همانطور که می‌بینیم، مثلا از Me به یک توییت رابطه posts هست که آن پست یک کاربر را mention می‌کند.

(اعداد روی توییت‌ها تعداد favorite-ها اند)

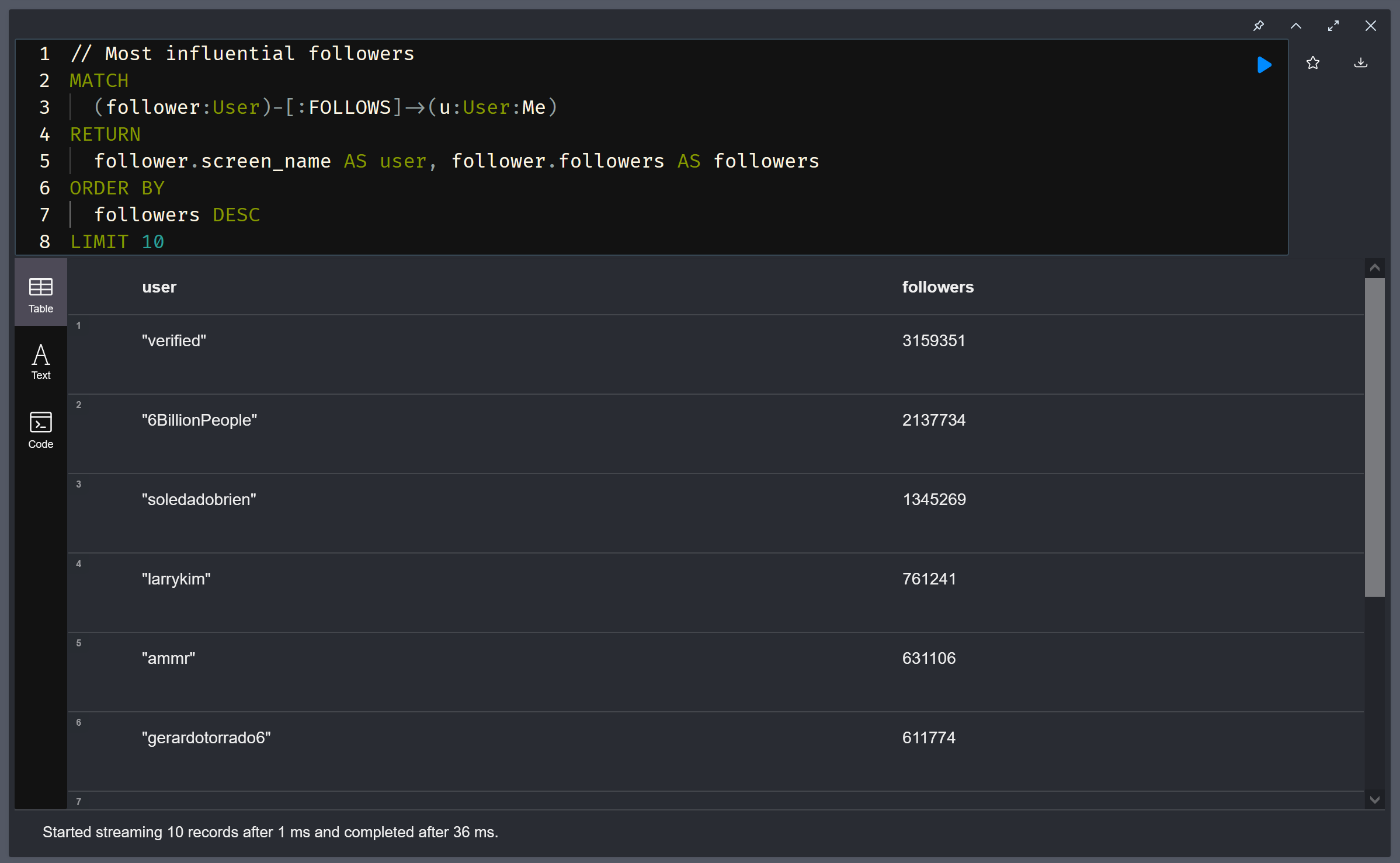
## 2-3. As Table

در نمایش جدولی، می‌بینیم که مثلا کاربر jimwebber را 40 بار در میان توییت‌هایمان منشن کرده‌ایم.



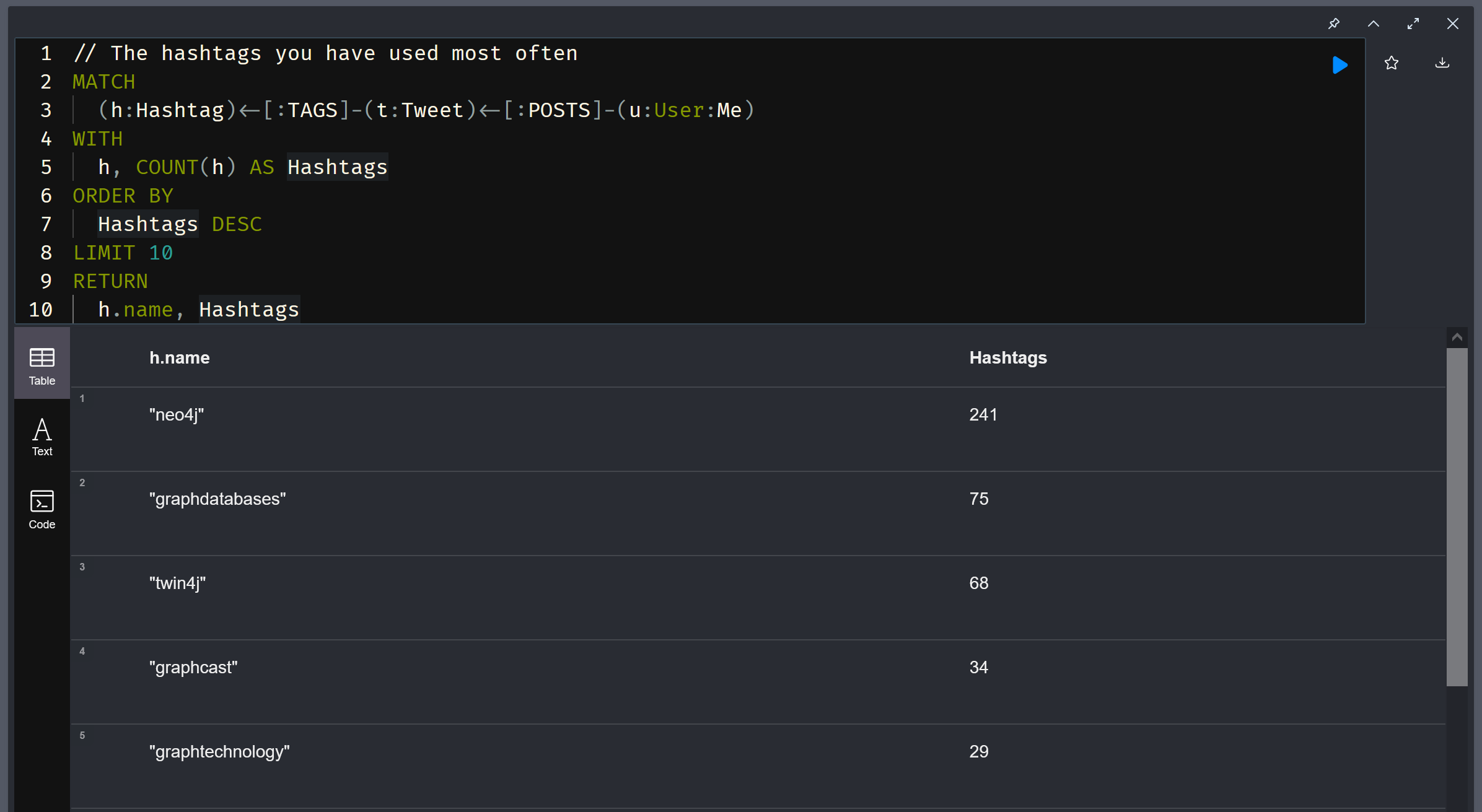
## 1-4. Most Influential Followers

در این بخش دنبال کسانی می‌گردیم که ما را follow کرده‌اند و تعداد follower-های آنها زیاد است. یعنی بین افرادی که ما را follow کرده اند، 10 نفری را انتخاب می‌کنیم که بیشترین تعداد follower را دارند.



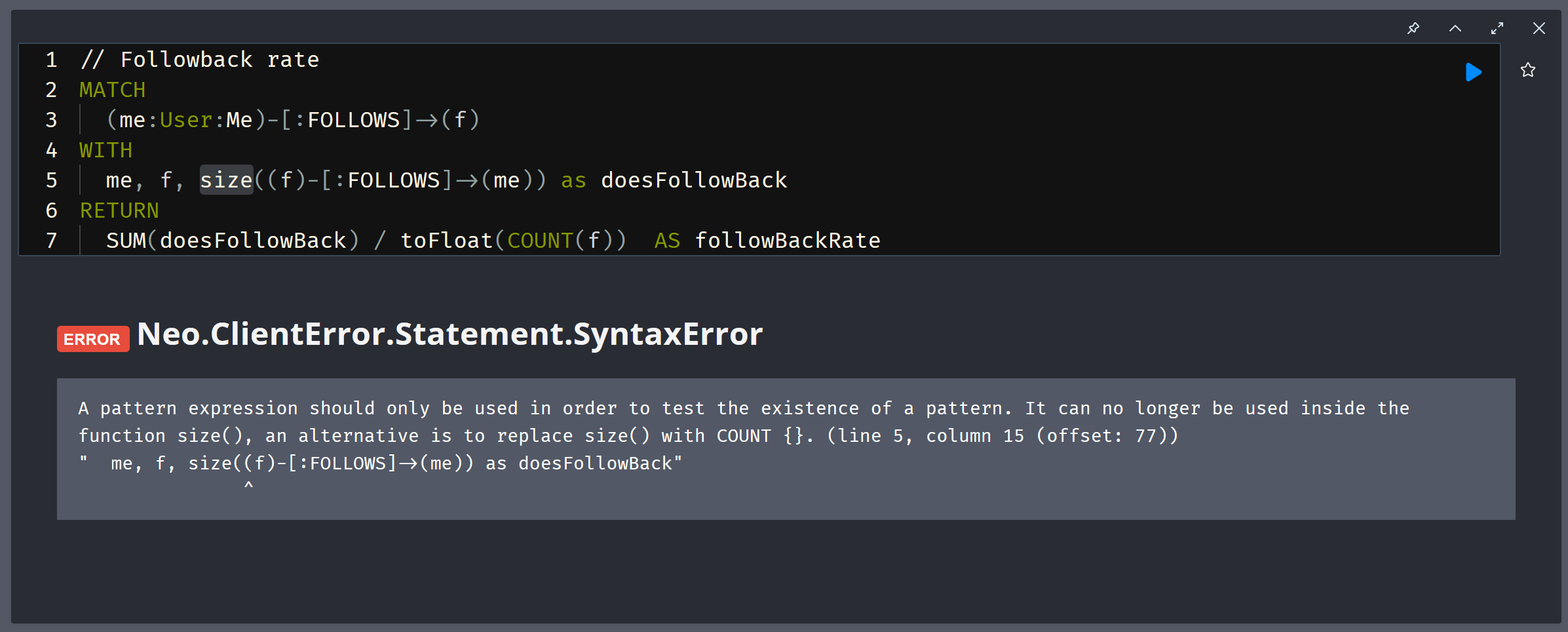
## 1-5. Most Tagged

در اینجا MATCH به صورت برعکس نوشته شده است و در نهایت تگ‌هایی که بیشتر از همه در توییت‌ها استفاده کرده‌ایم را (به همراه تعداد استفاده) برمی‌گرداند.



## 1-6. Followback Rate

کد این بخش با سینتکس ارور برخورد می‌کند:



با توجه به نکته ارور، با جایگزینی size با COUNT به دستور صحیح می‌رسیم:

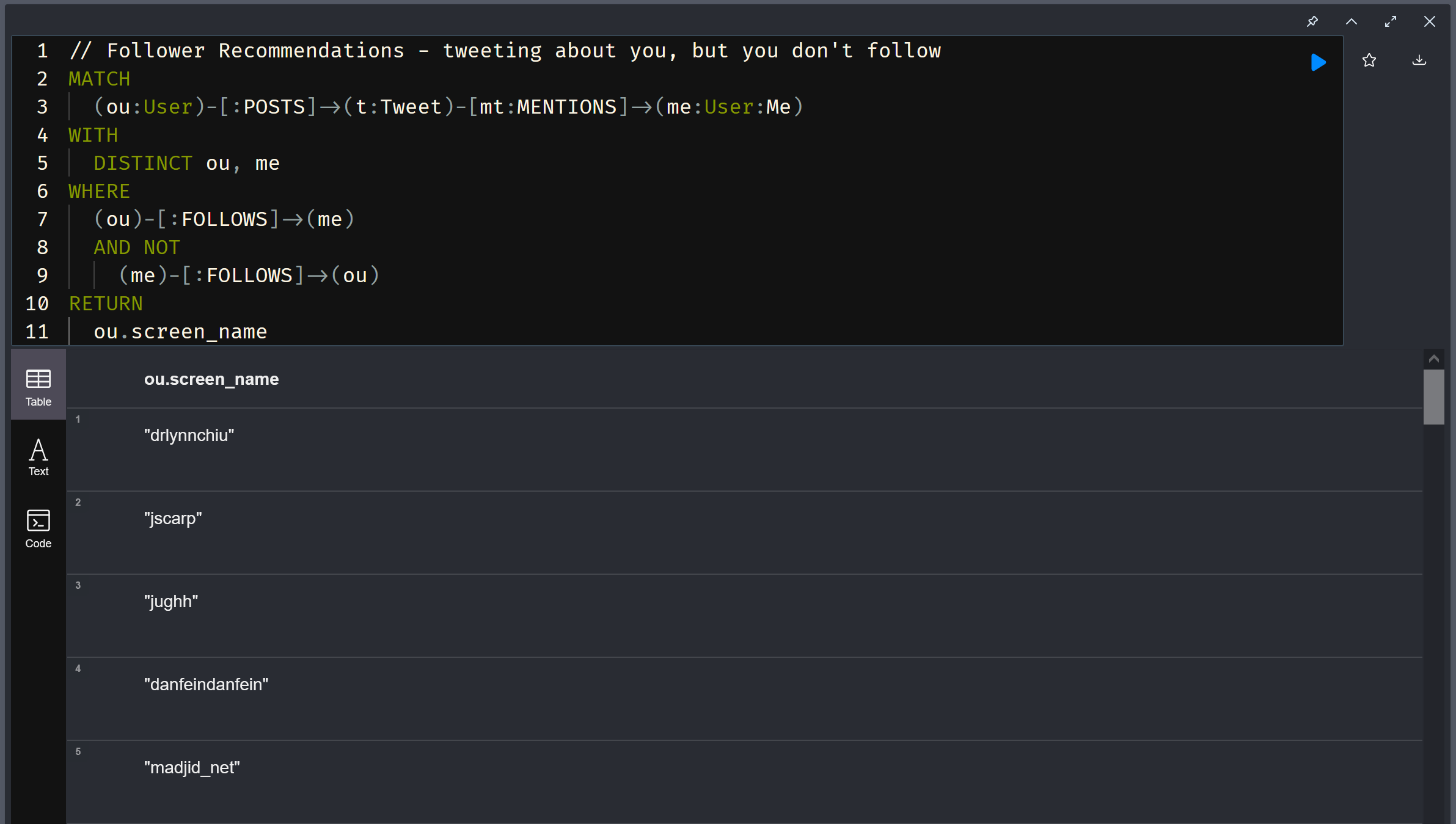


همانطور که می‌بینیم، ابتدا همه کسانی که ما follow می‌کنیم را می‌گیریم. سپس از بین آنها، آنهایی که ما را follow می‌کنند را پیدا کرده و نسبت آن به کل تعداد following-های خود را حساب می‌کنیم.

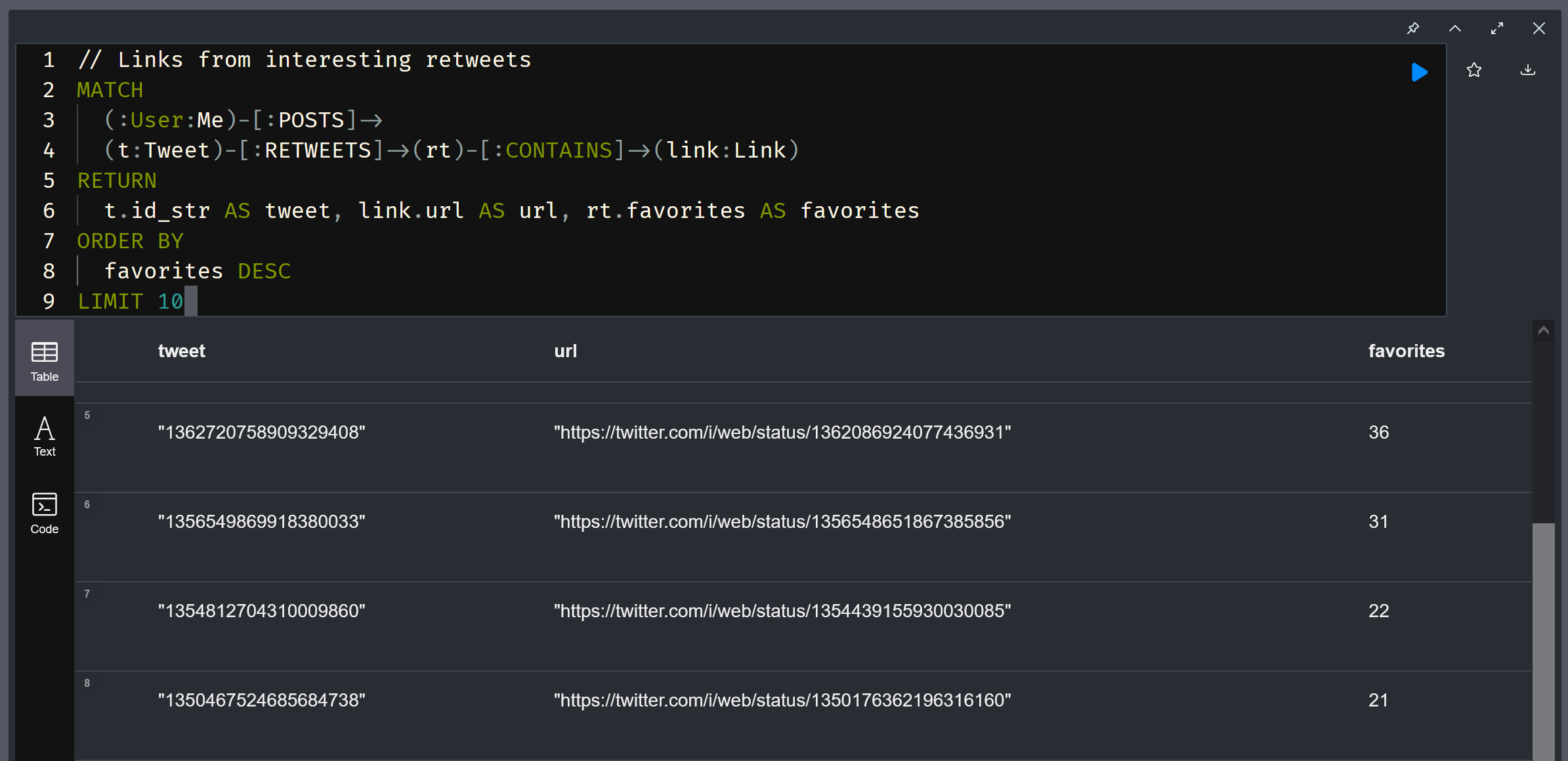
## 1-7. Follower Recommendations

با این کوئری، لیست کاربرانی که ما را در توییتی mention کرده اند را دریافت می‌کنیم که ما او را follow نمی‌کنیم.

استفاده از distinct برای این است که اگر کاربری چند بار mention شده باشد، یک رای بیشتر حساب نشود.



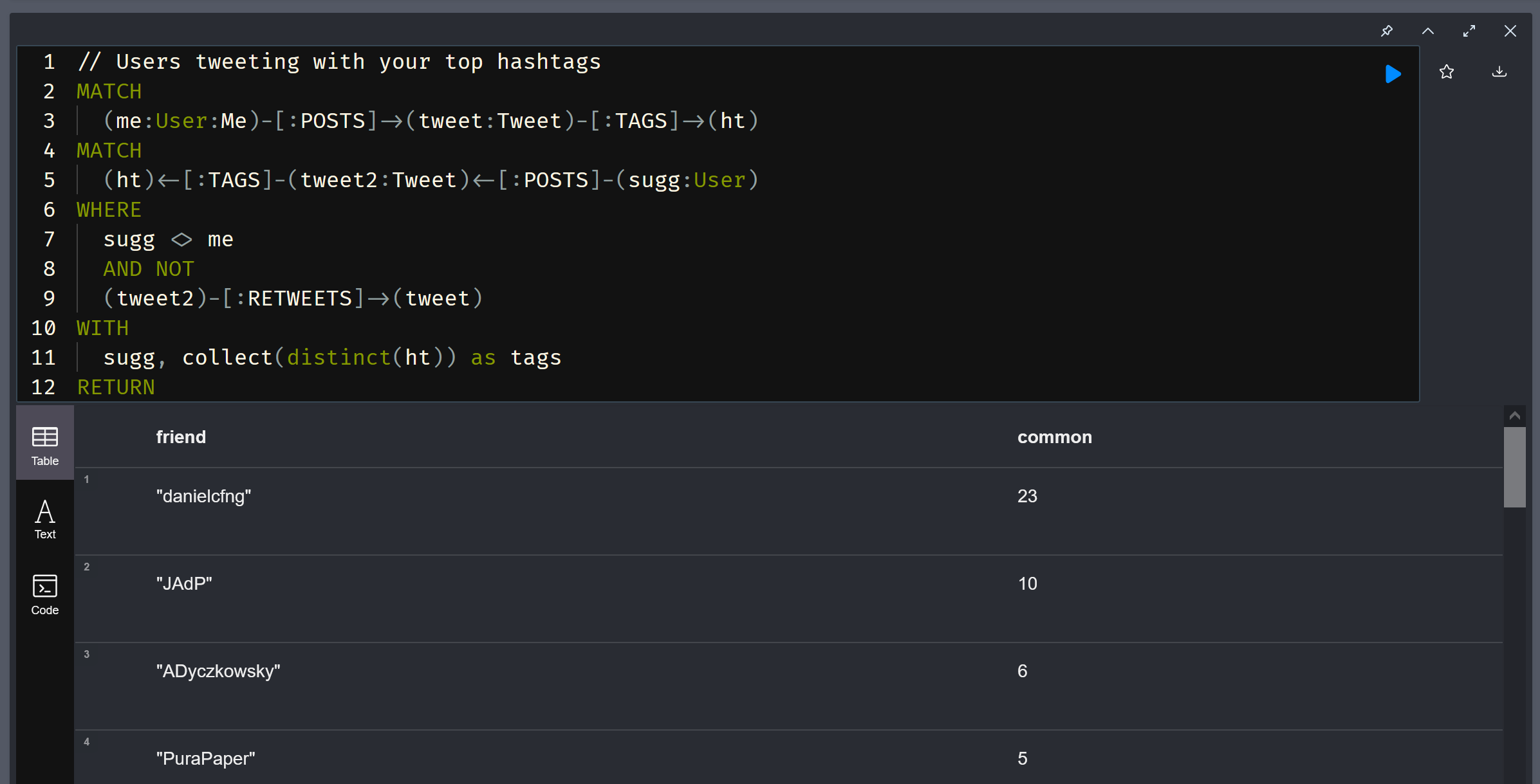
## 1-8. Links from Interesting Retweets

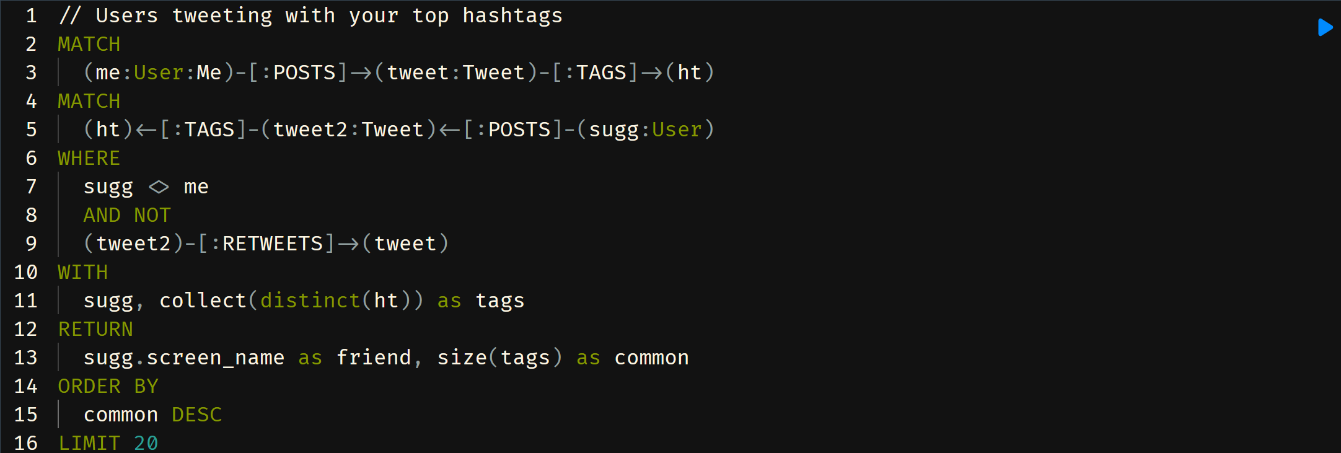


با این کوئری، پست‌هایی از خود را دریافت می‌کنیم که یک retweet هستند و در آنها لینک وجود دارد.

سپس بر حسب سورت تعداد favorite-های آن، 10 تای اول (بیشترین fav) را می‌گیریم.

## 1-9. Common Hashtags





این کوئری، ابتدا همه توییت‌های ما که هشتگ دارند را مچ می‌کند. سپس همه توییت‌های بقیه کاربرها که همان هشتگ‌ها را دارند را مچ می‌کند.

سپس با اعمال یک فیلتر WHERE، مطمئن می‌شویم که کاربرهای دیگر شامل خود ما نمی‌شود، و پست آنها یک retweet از توییت ما نیست.

حال هشتگ‌های متمایز استفاده شده توسط هر کاربر را گرفته و در نهایت 20 سطر، شامل نام کاربری که از هشتگ ما استفاده کرده، و تعداد هشتگ‌های مشترکی که استفاده کرده را به صورت سورت نزولی ریترن می‌کنیم.