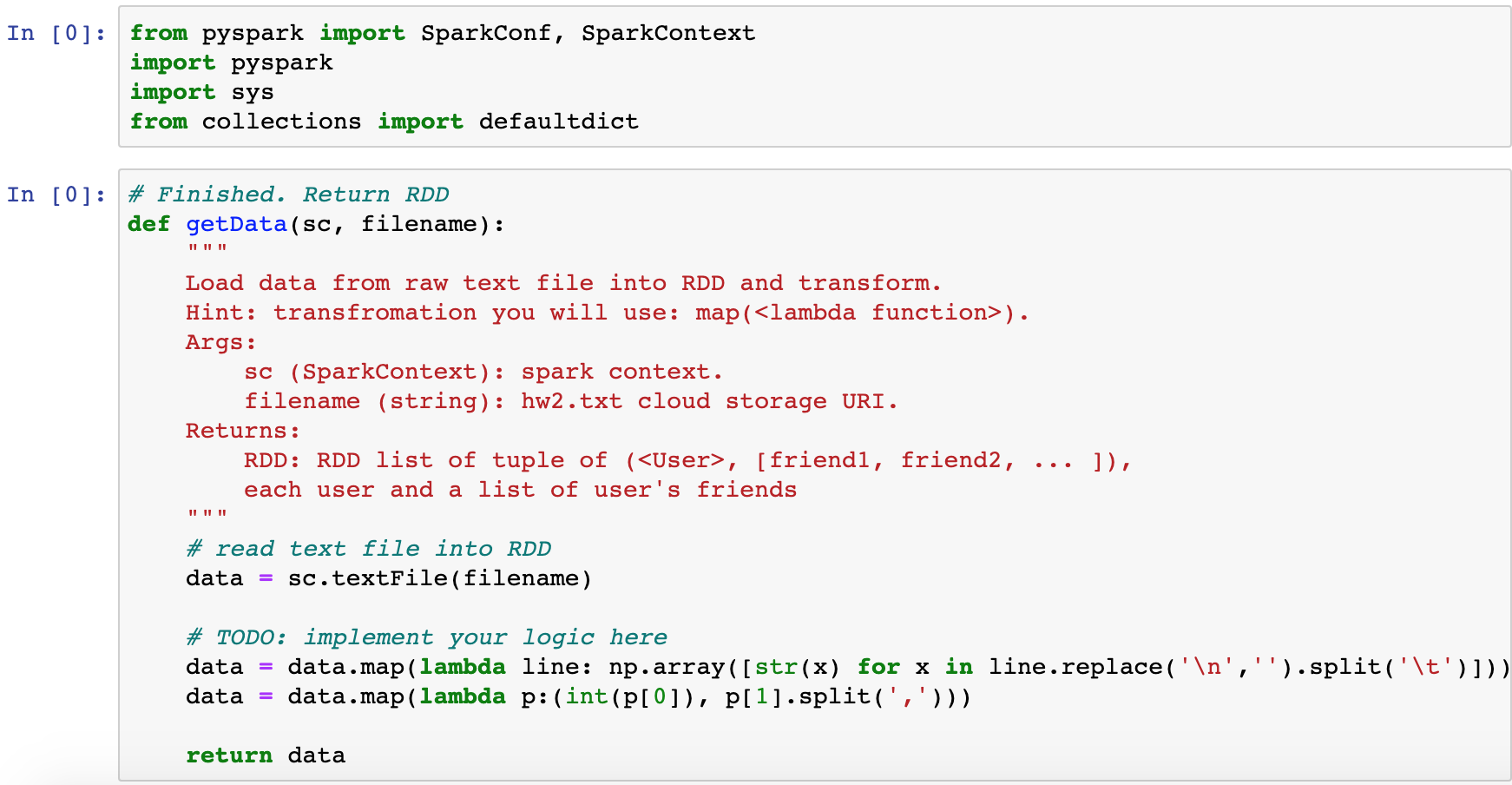
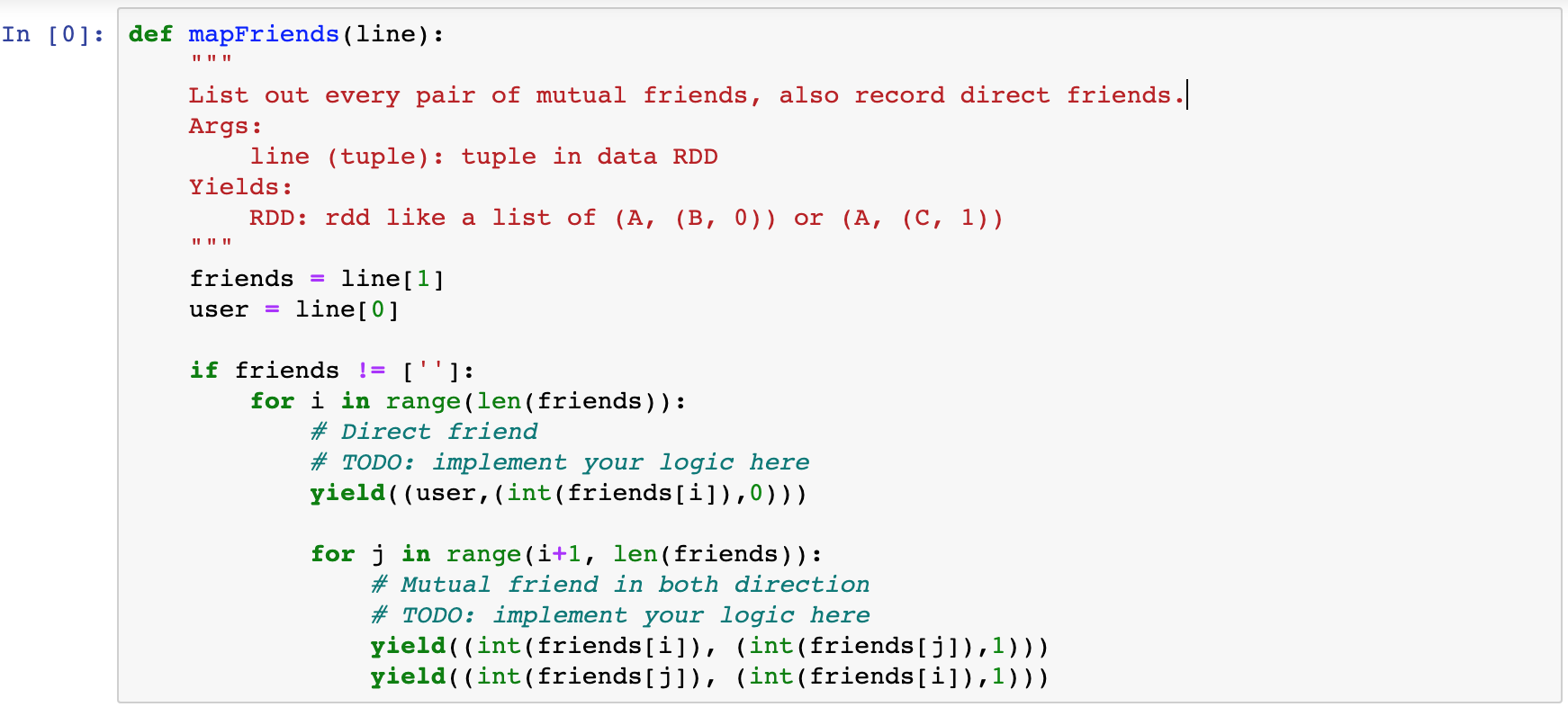
EECS E6893 HW2

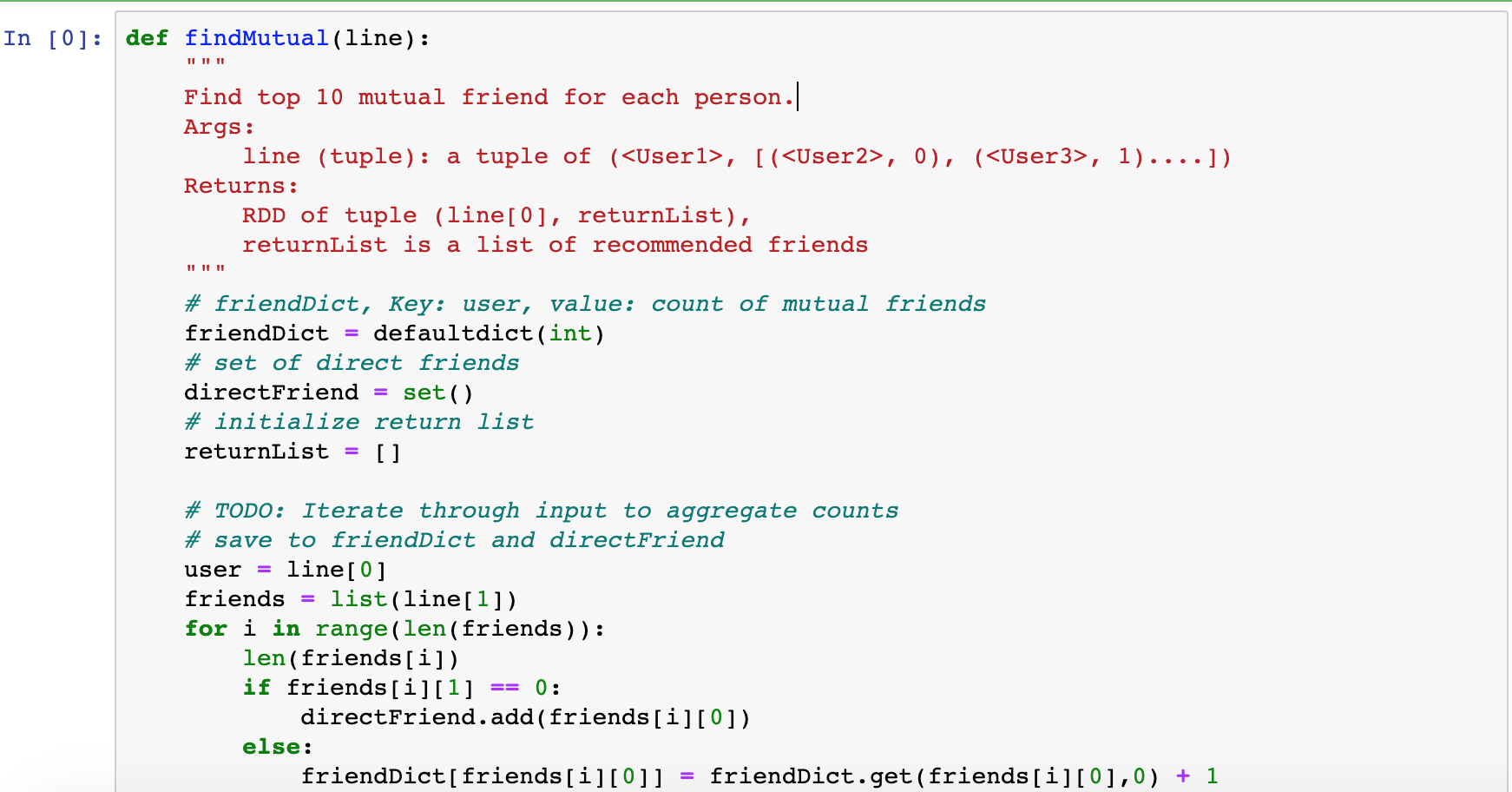
Jing Qian (jq2282)

**Question 1**

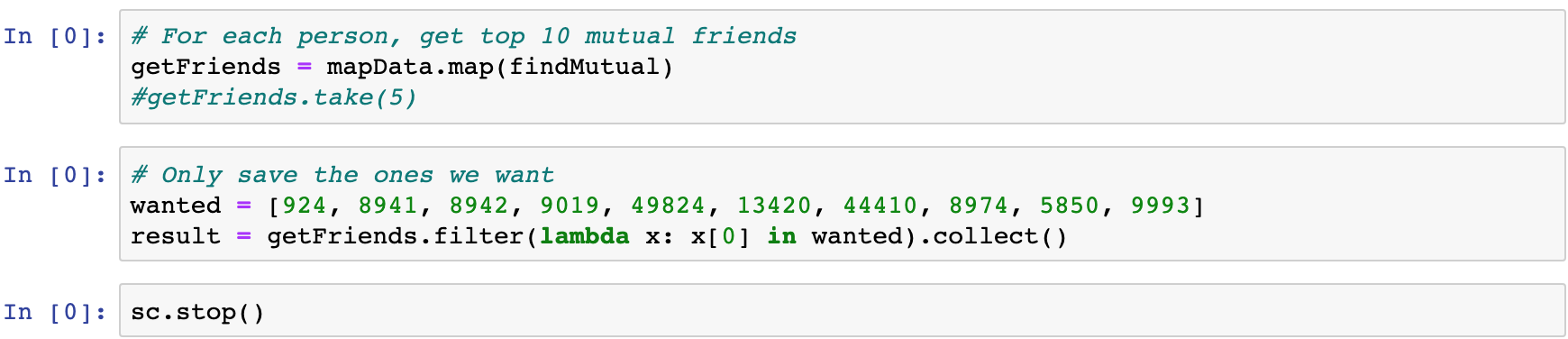
(1) Screenshots of the code



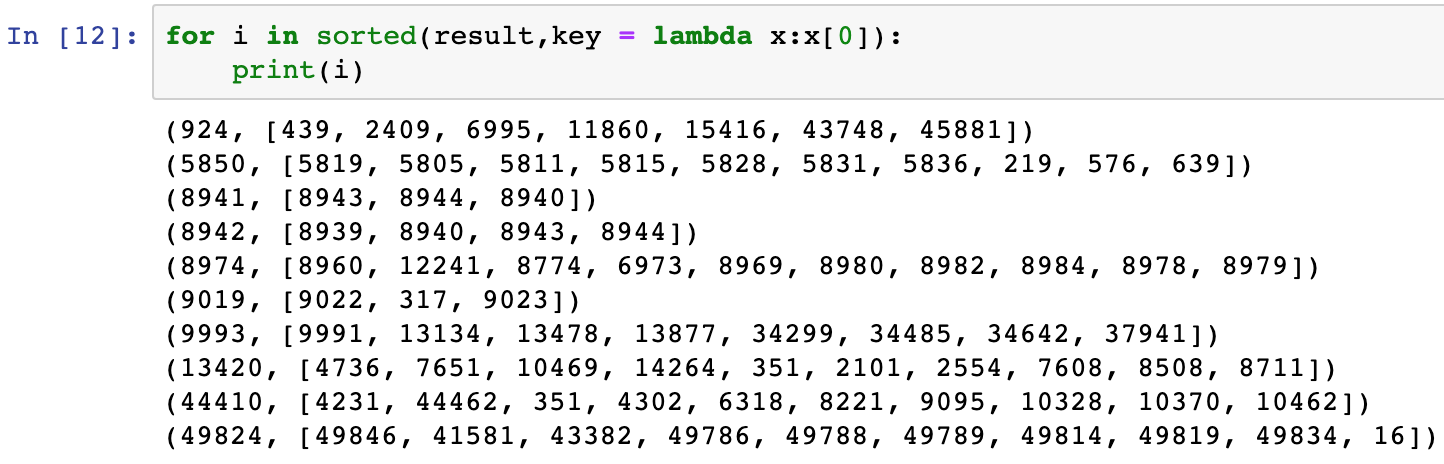




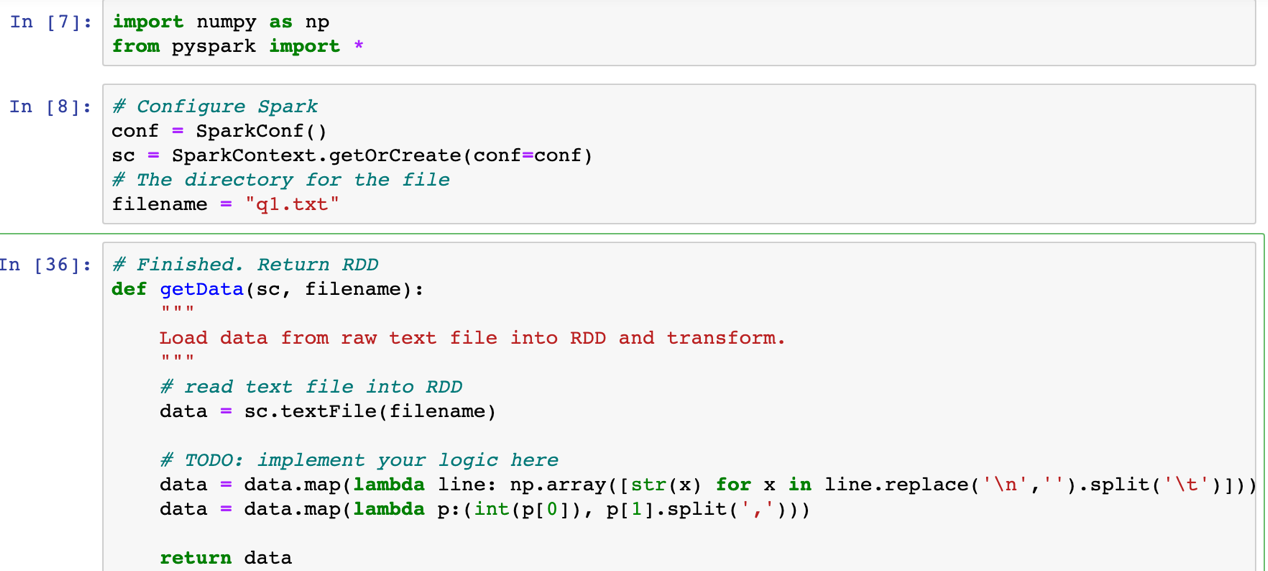


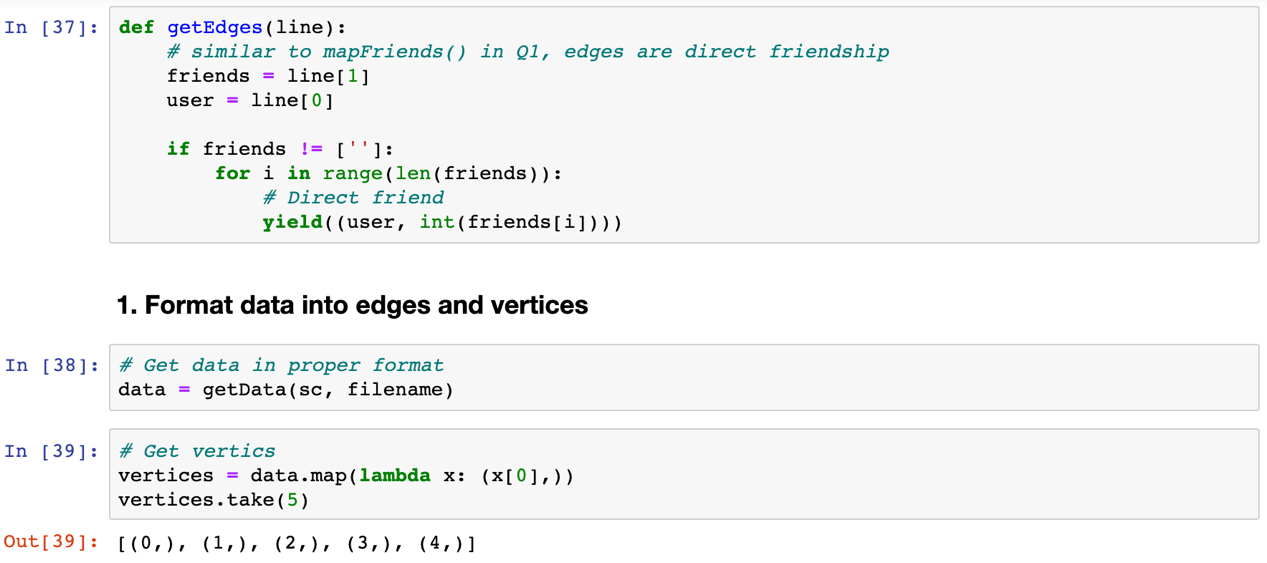


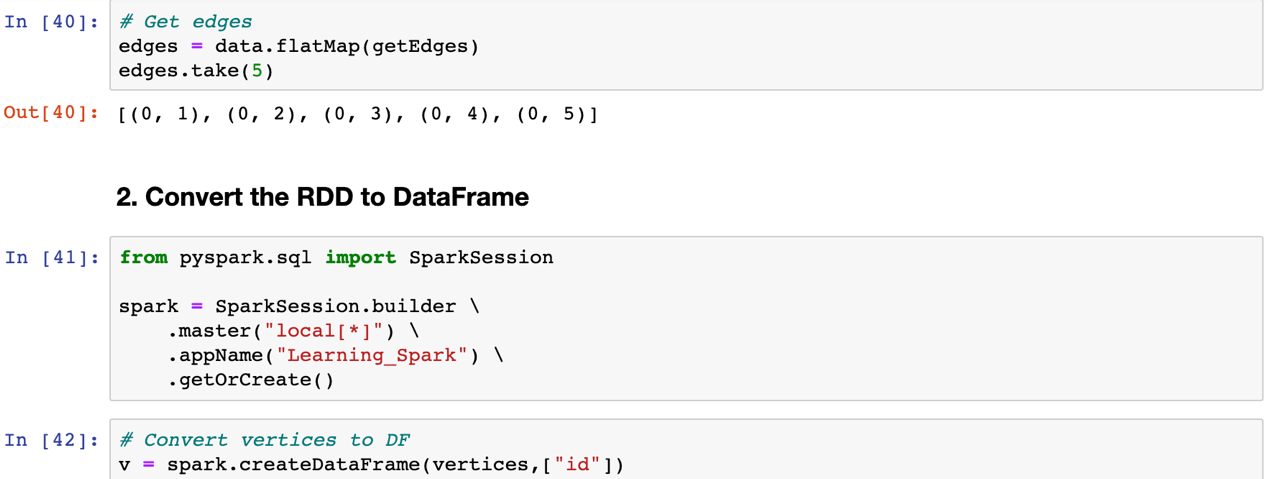
(2) Screenshots of the recommendation results.



**Question 2 Graph Analysis**









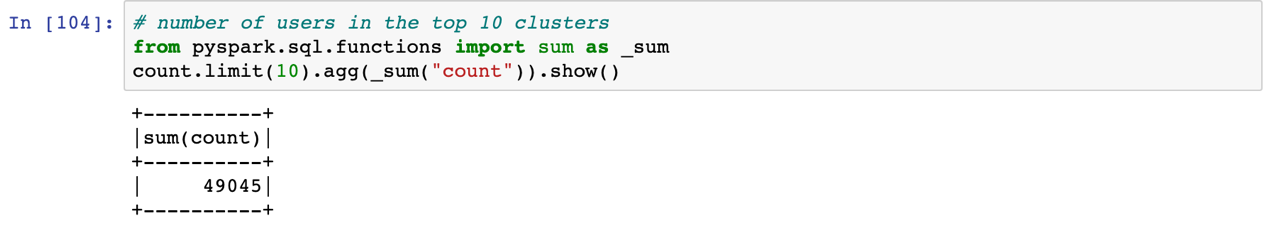


(1) There are 917 clusters/connected components in total for this dataset.



(2) There are 49045 users in the top 10 clusters.



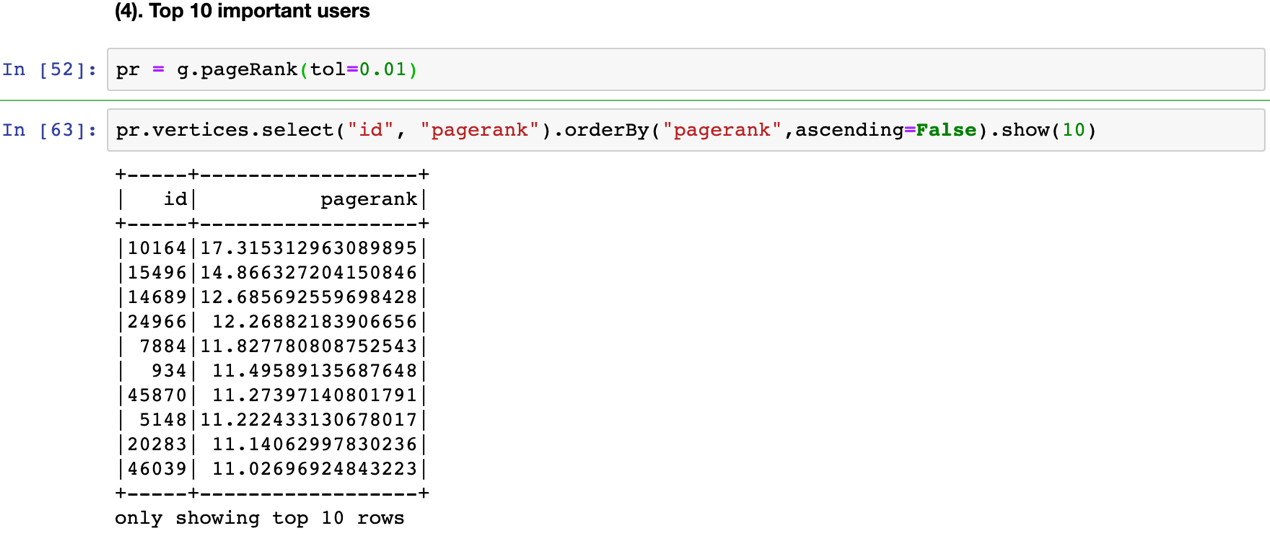


(3) The user ids for the cluster which has 25 users are: 18233 - 18257.





(4) The most important user is the one with User ID 10164.

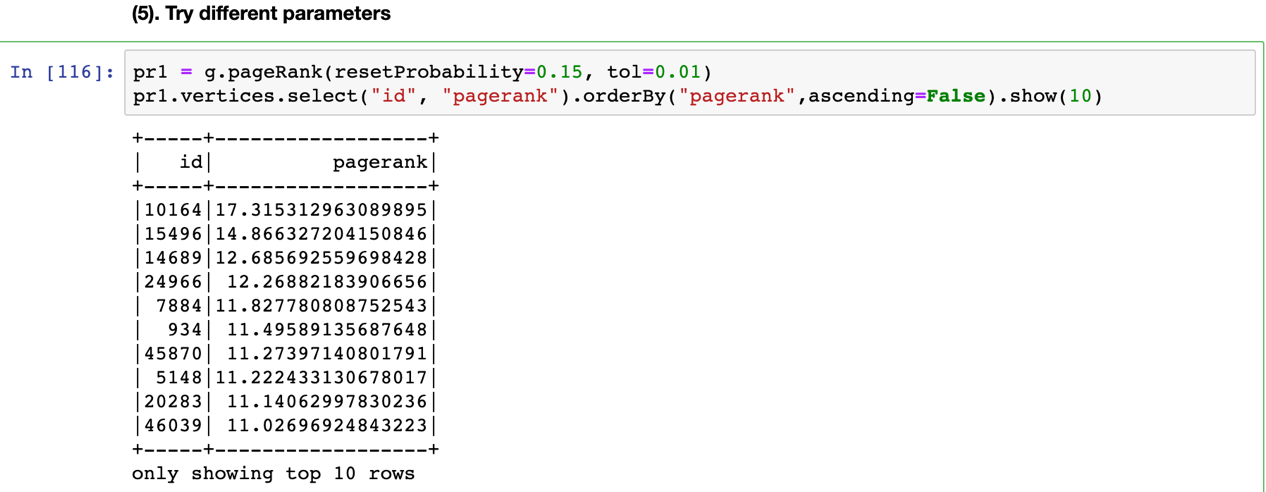


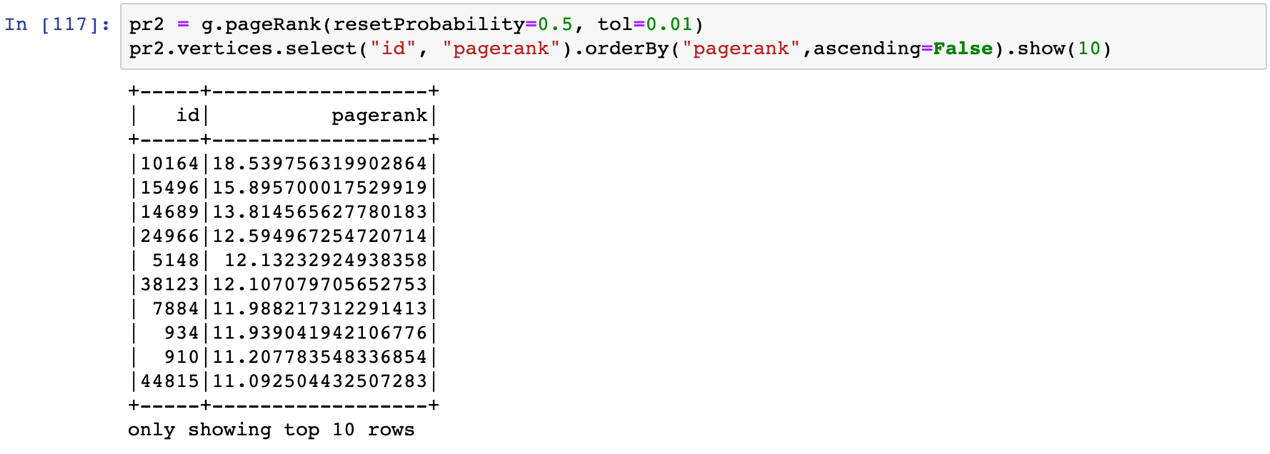
(5) Using different parameters setting for PageRank would lead to differences in the result.

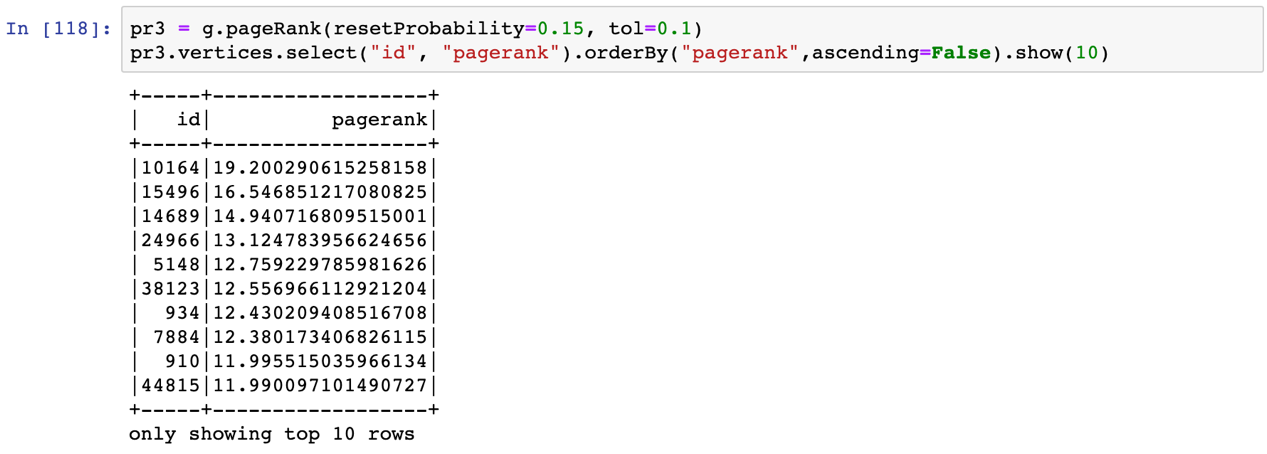
\* Increase the **"resetProbability"** from 0.15 (pr1) to 0.5 (pr2). We could see that the top four important users remain the same and the rest are different. "resetProbability" is the parameter that defines probability of resetting to a random vertex.

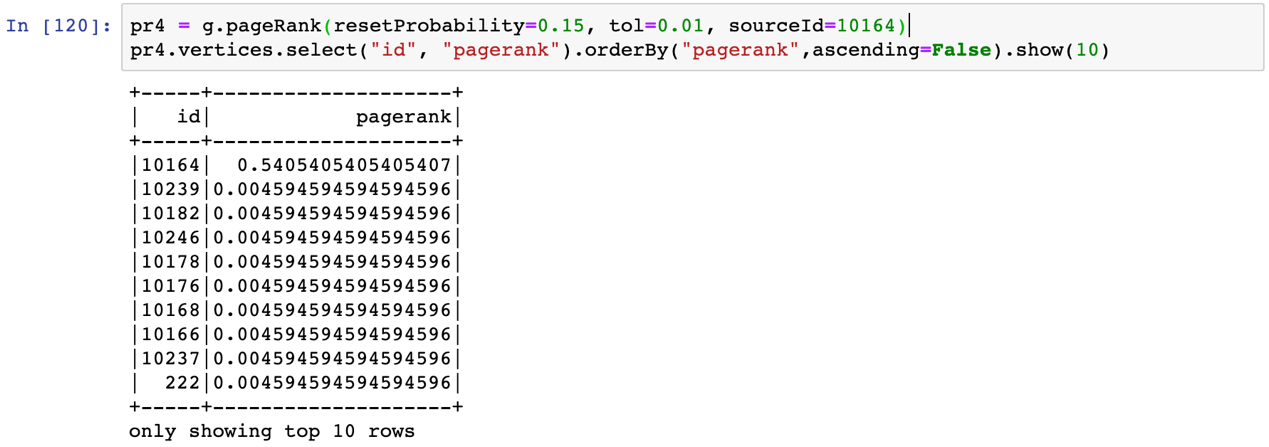
\* Increase the **"tol"** from 0.01 (pr1) to 0.1 (pr3). We could see that the top four important users remain the same and the rest are different. "tol" is the parameter that defines the convergence tolerance that algorithm runs. Increasing "tol" tends to decrease the algorithm iteration numbers.

\* Set the **"sourceId"** to 10164 (pr4). We could see that only the top user 10164 remains its position while others all change. Also, the pagerank values are quite different from original ones. "sourceId" is the parameter that assigns the source vertex for a personalized PageRank.



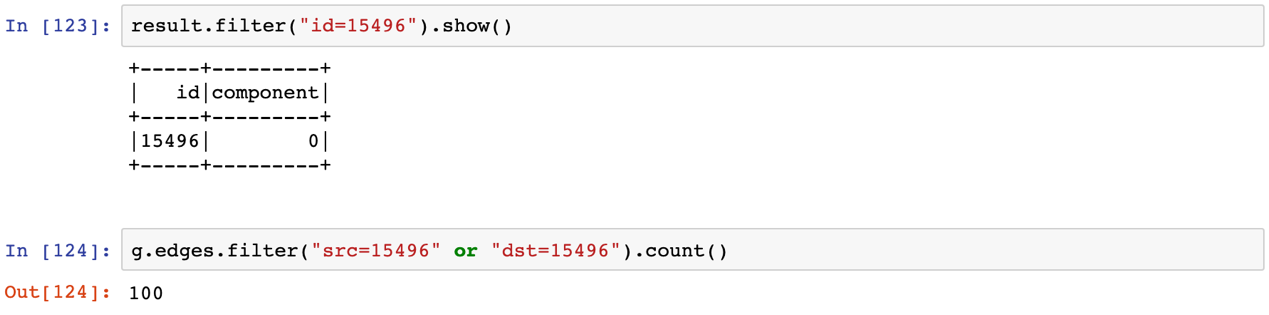






(6) The user with user id 10164 is the most important user. It belongs to the main cluster -- component 0, which is much larger than the rest clusters. Moreover, our graph is a bidirection graph. User 10164 has 100 edges, which means s/he is friends with many people and is able to share their importance. The second important user has the similar properties. These properties, being in the main cluster and having many edges, make 10164 the most important user.





(7) PageRank Calculation

For this question, I used self-defined function to do the calculation and the result is as following along with the code.

