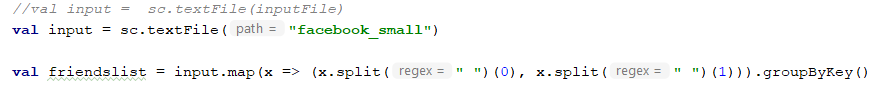
# Facebook Friends Suggestions Problem

This problem has you starting with a list of users and the people that they are friends to, and then you must generate another list that has all user combinations and the friends that they have in common.

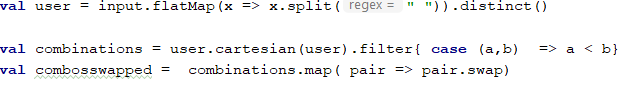
**Method:**

I started by importing the dataset



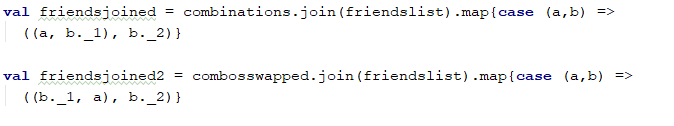
I then generate a key value pair of all the users and then group by the key so that you get one key value and then the list of all the users that user is friends with.

After that I get a list of all the users into a single rdd by rerunning the input and taking the distinct key from that.



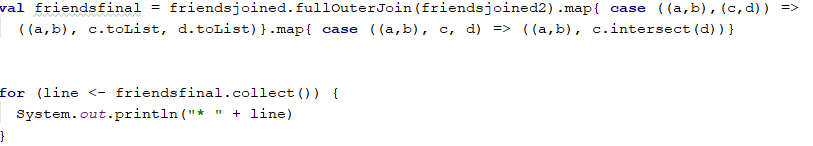
After that I generate a list of every possible combination of users only allowing x to be less than y. I then generate a second list of the inverse of that.

Then I join the value of the keys friends to an rdd.



Also do the inverse to get a list of all the friends for both the x and the y in the dataset.

Finally, you do a join of the two datasets keeping the values in order, and then run an intersect on the two lists to get a list back of just the values that are in both of the lists.



And then I output the results to the console.

(note) I had to run on a shrunken data set due to memory restrictions.

The output looks something like this:

