FINAL: QUESTION 2

Please use the Enron dataset you imported for the previous problem. For this question you will use the aggregation framework to figure out pairs of people that tend to communicate a lot. To do this, you will need to unwind the To list for each message.

This problem is a little tricky because a recipient may appear more than once in the To list for a message. You will need to fix that in a stage of the aggregation before doing your grouping and counting of (sender, recipient) pairs.

Which pair of people have the greatest number of messages in the dataset?

- ○ susan.mara@enron.com to jeff.dasovich@enron.com
- ● susan.mara@enron.com to richard.shapiro@enron.com
- ○ soblander@carrfut.com to soblander@carrfut.com
- ○ susan.mara@enron.com to james.steffes@enron.com
- ○ evelyn.metoyer@enron.com to kate.symes@enron.com
- ○ susan.mara@enron.com to alan.comnes@enron.com

SUBMIT

◀ ▶