***>>> ip = sc.textFile('Documents/EMP\_MGR')  
>>> ip.collect()***  
[u'100:Alan:400', u'200:Bob:400', u'300:Jane:500', u'400:Jeff:500', u'500:Martha:500']

***>>> lines = ip.map(lambda x:(x.split(':')))  
>>> lines.collect()***[[u'100', u'Alan', u'400'], [u'200', u'Bob', u'400'], [u'300', u'Jane', u'500'], [u'400', u'Jeff', u'500'], [u'500', u'Martha', u'500']]

***>>> cartRDD = lines.cartesian(lines)  
>>> cartRDD.collect()***[([u'100', u'Alan', u'400'], [u'100', u'Alan', u'400']), ([u'100', u'Alan', u'400'], [u'200', u'Bob', u'400']), ([u'100', u'Alan', u'400'], [u'300', u'Jane', u'500']), ([u'100', u'Alan', u'400'], [u'400', u'Jeff', u'500']), ([u'100', u'Alan', u'400'], [u'500', u'Martha', u'500']), ([u'200', u'Bob', u'400'], [u'100', u'Alan', u'400']), ([u'300', u'Jane', u'500'], [u'100', u'Alan', u'400']), ([u'200', u'Bob', u'400'], [u'200', u'Bob', u'400']), ([u'200', u'Bob', u'400'], [u'300', u'Jane', u'500']), ([u'300', u'Jane', u'500'], [u'200', u'Bob', u'400']), ([u'300', u'Jane', u'500'], [u'300', u'Jane', u'500']), ([u'200', u'Bob', u'400'], [u'400', u'Jeff', u'500']), ([u'200', u'Bob', u'400'], [u'500', u'Martha', u'500']), ([u'300', u'Jane', u'500'], [u'400', u'Jeff', u'500']), ([u'300', u'Jane', u'500'], [u'500', u'Martha', u'500']), ([u'400', u'Jeff', u'500'], [u'100', u'Alan', u'400']), ([u'500', u'Martha', u'500'], [u'100', u'Alan', u'400']), ([u'400', u'Jeff', u'500'], [u'200', u'Bob', u'400']), ([u'400', u'Jeff', u'500'], [u'300', u'Jane', u'500']), ([u'500', u'Martha', u'500'], [u'200', u'Bob', u'400']), ([u'500', u'Martha', u'500'], [u'300', u'Jane', u'500']), ([u'400', u'Jeff', u'500'], [u'400', u'Jeff', u'500']), ([u'400', u'Jeff', u'500'], [u'500', u'Martha', u'500']), ([u'500', u'Martha', u'500'], [u'400', u'Jeff', u'500']), ([u'500', u'Martha', u'500'], [u'500', u'Martha', u'500'])]

***>>> filteredRDD = cartRDD.filter(lambda x:x[0][2]==x[1][0])  
>>> filteredRDD.collect()***[([u'100', u'Alan', u'400'], [u'400', u'Jeff', u'500']), ([u'200', u'Bob', u'400'], [u'400', u'Jeff', u'500']), ([u'300', u'Jane', u'500'], [u'500', u'Martha', u'500']), ([u'400', u'Jeff', u'500'], [u'500', u'Martha', u'500']), ([u'500', u'Martha', u'500'], [u'500', u'Martha', u'500'])]

***>>> output = filteredRDD.map(lambda x:(x[0][0],x[0][1],x[1][1]))  
>>> output.collect()***[(u'100', u'Alan', u'Jeff'), (u'200', u'Bob', u'Jeff'), (u'300', u'Jane', u'Martha'), (u'400', u'Jeff', u'Martha'), (u'500', u'Martha', u'Martha')]  
>>>