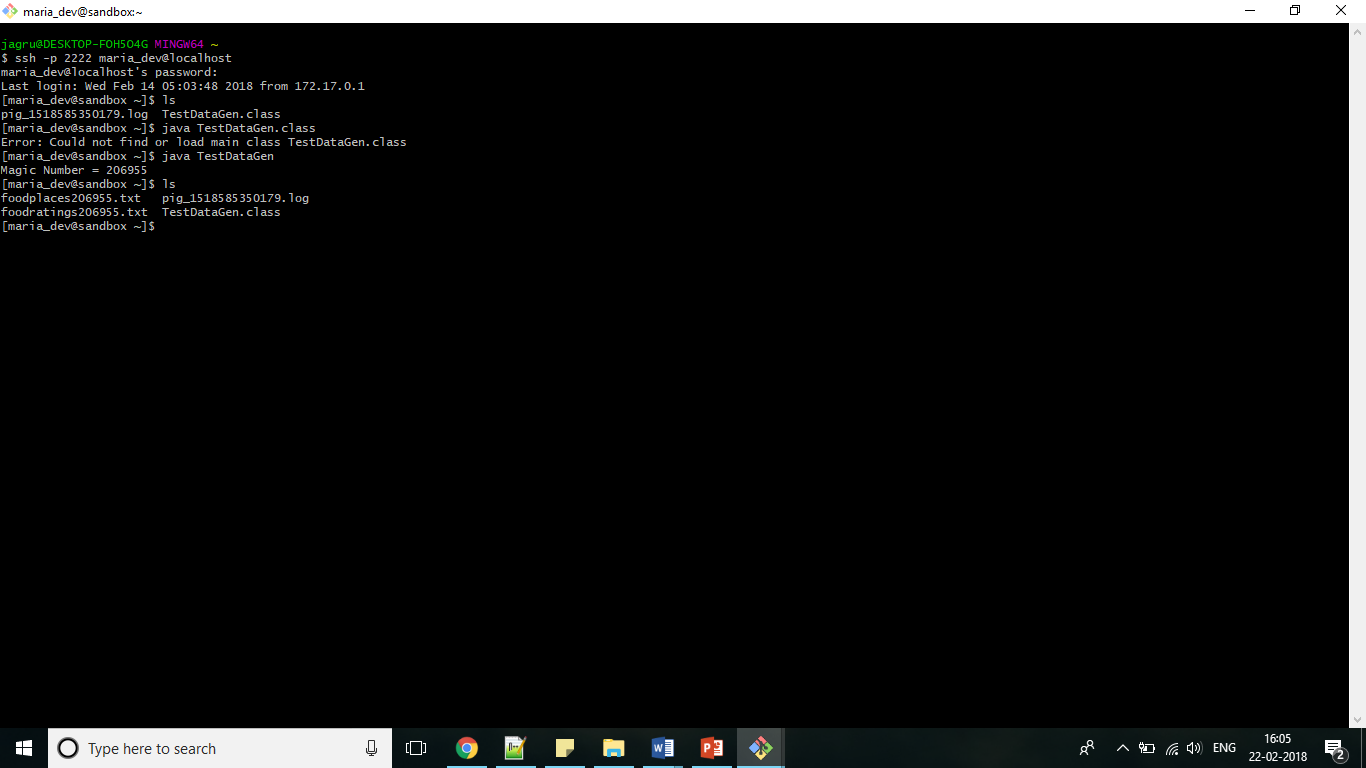
**CSP 595 - Assignment 6**

Name: Jagruti Vichare

Email ID: [jvichare@hawk.iit.edu](mailto:jvichare@hawk.iit.edu)

CWID: A20378092

**Question** 1



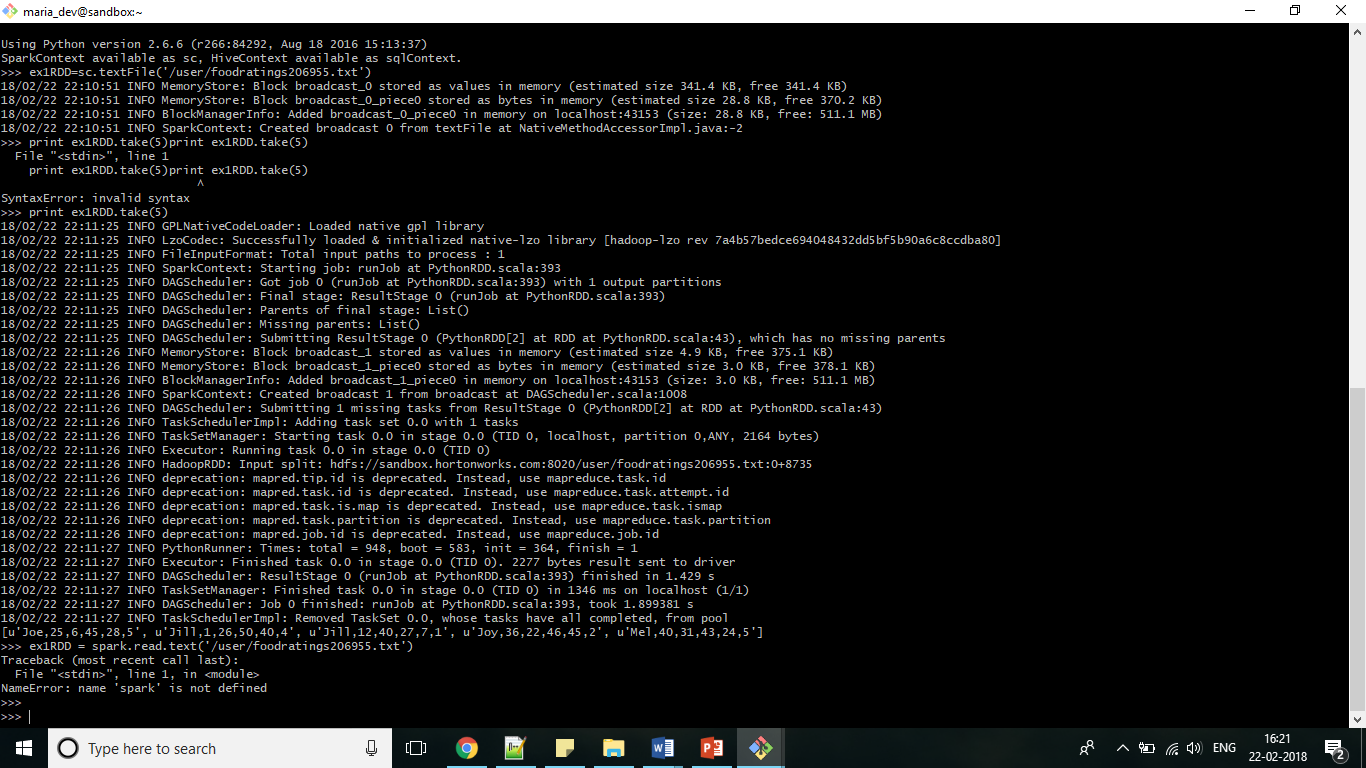
Commands:

ex1RDD=sc.textFile('/user/foodratings206955.txt')

print ex1RDD.take(5)

Output:

[u'Joe,25,6,45,28,5', u'Jill,1,26,50,40,4', u'Jill,12,40,27,7,1', u'Joy,36,22,46,45,2', u'Mel,40,31,43,24,5']



**Question 2:**

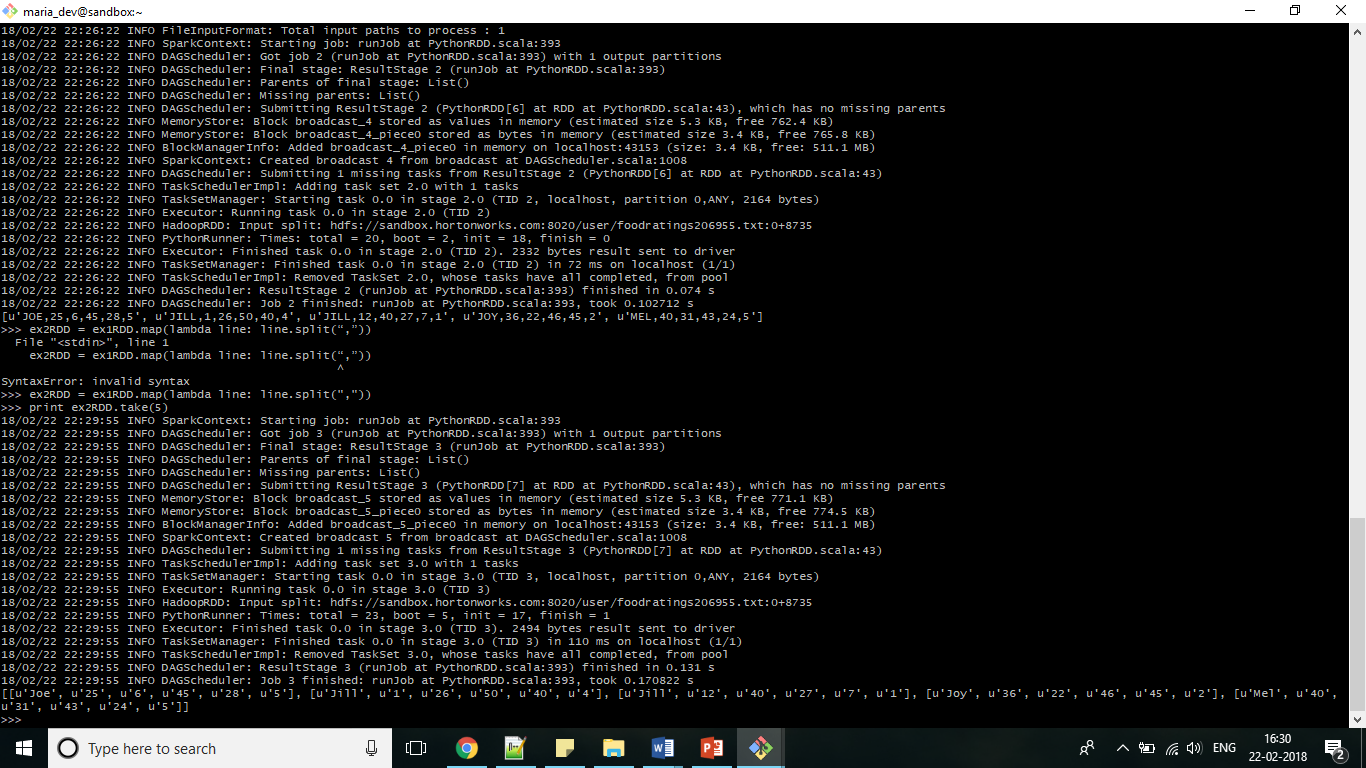
Commands:

ex2RDD = ex1RDD.map(lambda line: line.split(“,”))

print ex2RDD.take(5)

Output:

[[u'Joe', u'25', u'6', u'45', u'28', u'5'], [u'Jill', u'1', u'26', u'50', u'40', u'4'], [u'Jill', u'12', u'40', u'27', u'7', u'1'], [u'Joy', u'36', u'22', u'46', u'45', u'2'], [u'Mel', u'40', u'31', u'43', u'24', u'5']]



**Question 3:**

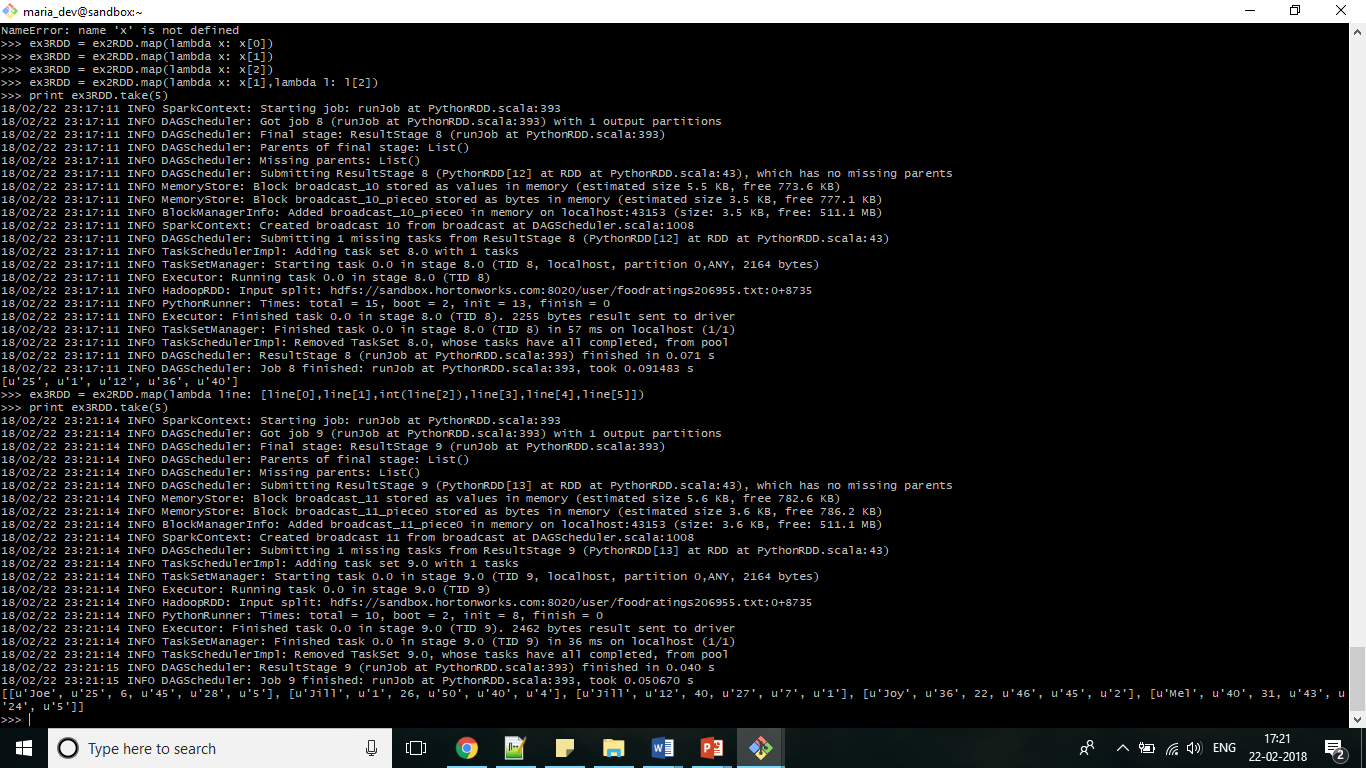
Commands:

ex3RDD = ex2RDD.map(lambda line: [line[0],line[1],int(line[2]),line[3],line[4],line[5]])

print ex3RDD.take(5)

Output:

[[u'Joe', u'25', 6, u'45', u'28', u'5'], [u'Jill', u'1', 26, u'50', u'40', u'4'], [u'Jill', u'12', 40, u'27', u'7', u'1'], [u'Joy', u'36', 22, u'46', u'45', u'2'], [u'Mel', u'40', 31, u'43', u'24', u'5']]



**Question 4**

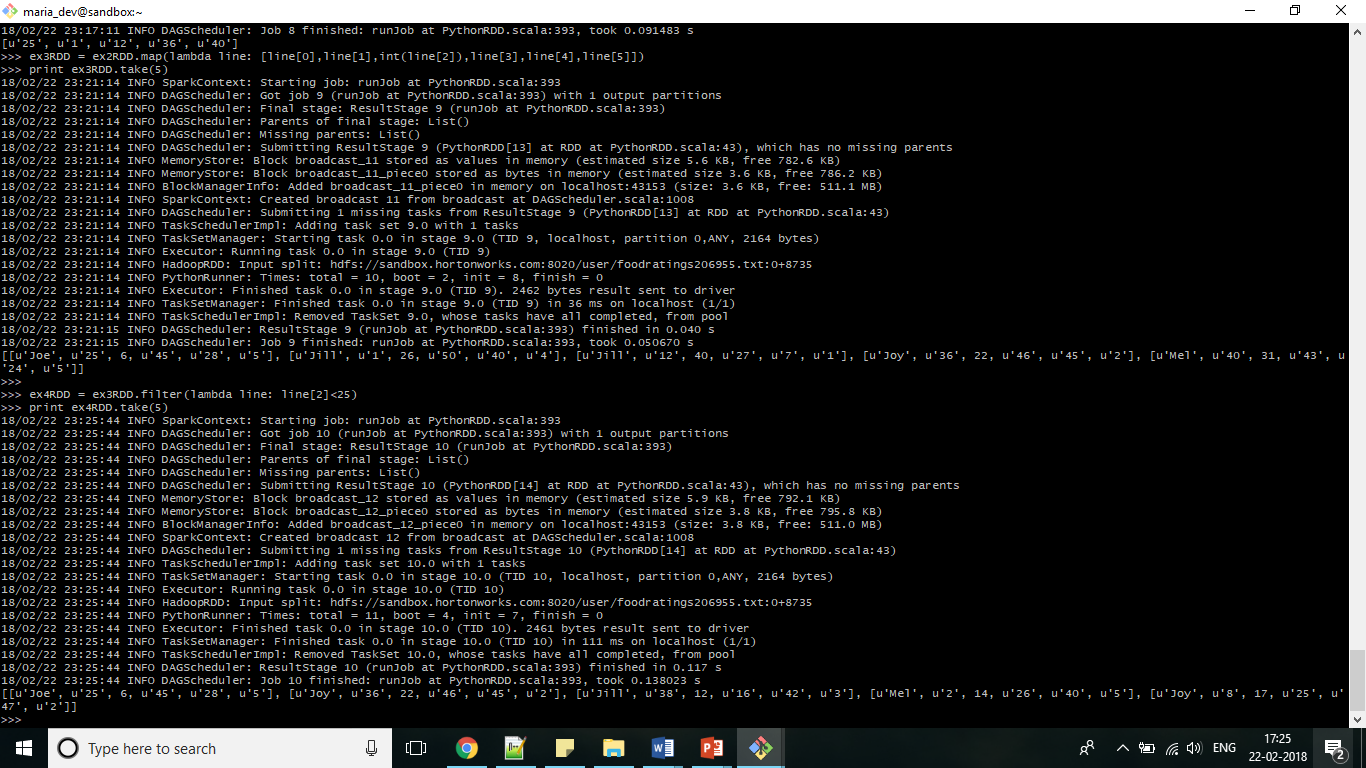
Commands:

ex4RDD = ex3RDD.filter(lambda line: line[2]<25)

print ex4RDD.take(5)

Output:

[[u'Joe', u'25', 6, u'45', u'28', u'5'], [u'Joy', u'36', 22, u'46', u'45', u'2'], [u'Jill', u'38', 12, u'16', u'42', u'3'], [u'Mel', u'2', 14, u'26', u'40', u'5'], [u'Joy', u'8', 17, u'25', u'47', u'2']]



**Question 5**

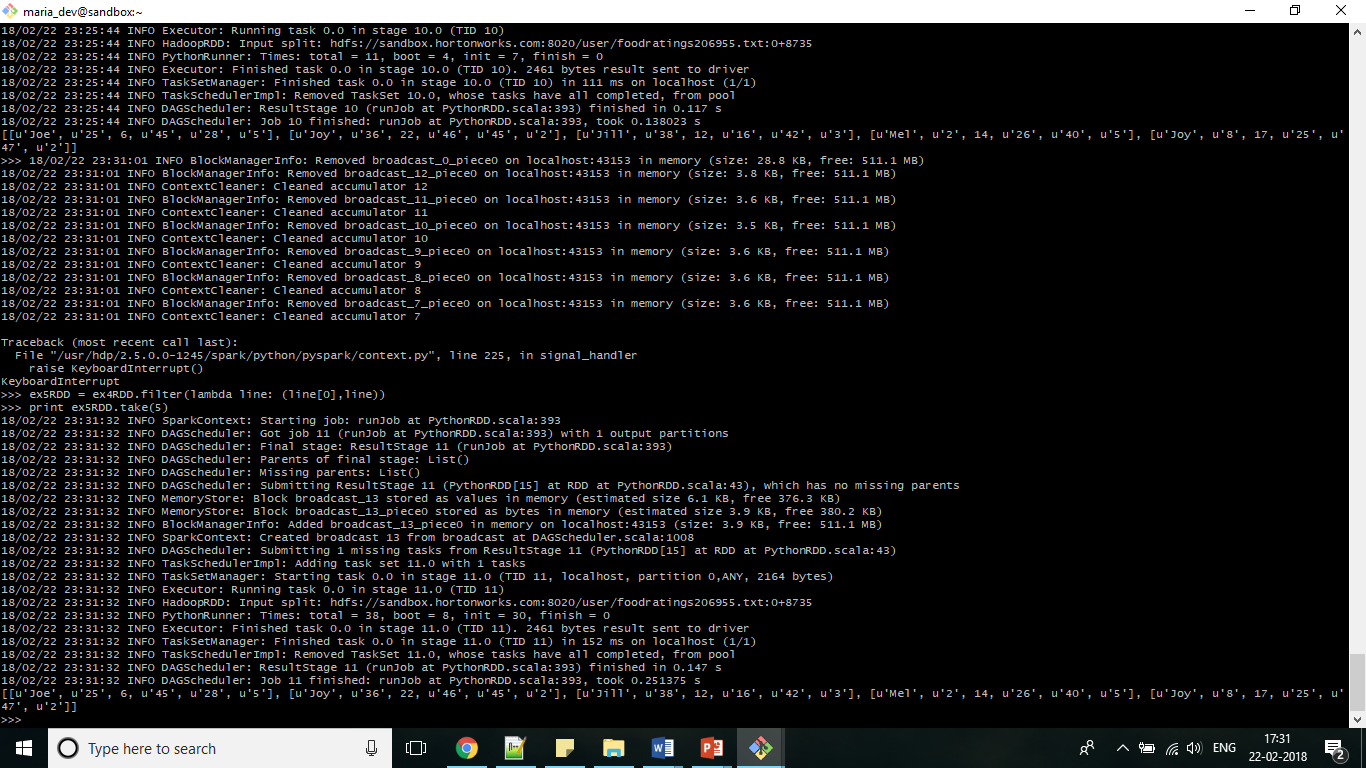
Commands:

ex5RDD = ex4RDD.filter(lambda line: (line[0],line))

print ex5RDD.take(5)

Output:

[[u'Joe', u'25', 6, u'45', u'28', u'5'], [u'Joy', u'36', 22, u'46', u'45', u'2'], [u'Jill', u'38', 12, u'16', u'42', u'3'], [u'Mel', u'2', 14, u'26', u'40', u'5'], [u'Joy', u'8', 17, u'25', u'47', u'2']]



**Question 6**

Commands:

ex6RDD = ex5RDD.sortByKey(True,1)

print ex6RDD.take(5)

Output:

[[u'Jill', u'38', 12, u'16', u'42', u'3'], [u'Jill', u'31', 18, u'18', u'25', u'4'], [u'Jill', u'38', 5, u'11', u'33', u'2'], [u'Jill', u'36', 10, u'26', u'18', u'3'], [u'Jill', u'4', 13, u'6', u'47', u'4']]

