Bhasheyam Krishnan

A20380078

Magic number used = 63666

Que1 :

Command used:

```
lines=sc.textFile('/user/maria_dev/foodratings63666.txt')
print lines.take(5)
```

Ouput:

```
18/02/21 04:49:41 INFO DAGScheduler: Job 0 finished: runJob at PythonRDD.scala:393, took 1.415764 s
[u'Sam,16,9,16,28,4', u'Jill,10,24,43,31,4', u'Jill,34,40,6,19,5', u'Joe,25,38,35,24,5', u'Joe,15,48,34,39,1']
>>>
```

Que2:

Command

```
lines=sc.textFile('/user/maria_dev/foodratings63666.txt')
ex2RDD = lines.map(lambda line: line.split(","))
print ex2RDD.take(5)
```

Output:

```
[[u'Sam', u'16', u'9', u'16', u'28', u'4'], [u'Jill', u'10', u'24', u'43', u'31', u'4'], [u'Jill', u'34', u'40', u'6',
19', u'5'], [u'Joe', u'25', u'38', u'35', u'24', u'5'], [u'Joe', u'15', u'48', u'34', u'39', u'1']]
>>>
```

Que 3:

Command :

```
lines=sc.textFile('/user/maria_dev/foodratings63666.txt')
ex2RDD = lines.map(lambda line: line.split(","))
ex3RDD = ex2RDD.map(lambda line : [line[0], line[1], int(line[2]), line[3], line[4], line[5]])
print ex3RDD.take(5)
```

Output:

```
[[u'Sam', u'16', 9, u'16', u'28', u'4'], [u'Jill', u'10', 24, u'43', u'31', u'4'], [u'Jill', u'34', 40, u'6', u'19
'], [u'Joe', u'25', 38, u'35', u'24', u'5'], [u'Joe', u'15', 48, u'34', u'39', u'1']]
>>>
```

Que 4:

Command:

```
lines=sc.textFile('/user/maria_dev/foodratings63666.txt')
ex2RDD = lines.map(lambda line: line.split(","))
ex3RDD = ex2RDD.map(lambda line : [line[0], line[1], int(line[2]), line[3], line[4], line[5]])
ex4RDD = ex3RDD.filter(lambda line :int(line[2]) < 25)
print ex4RDD.take(5)
```

Output:

```
18/02/21 05:01:58 INFO DAGScheduler: Job 0 Finished: runJob at PythonRDD.scala:395, took 0.198011 s
[[u'Sam', u'16', 9, u'16', u'28', u'4'], [u'Jill', u'10', 24, u'43', u'31', u'4'], [u'Jill', u'11', 22, u'50', u'34'
4'], [u'Joe', u'24', 7, u'17', u'6', u'3'], [u'Mel', u'42', 5, u'26', u'3', u'2']]
>>>
```

Fro Que3 we can see the three record is not present in this output as it has 40 in the Third field.

Que 5:

Command:

```
lines=sc.textFile('/user/maria_dev/foodratings63666.txt')
ex2RDD = lines.map(lambda line: line.split(","))
ex3RDD = ex2RDD.map(lambda line : [line[0], line[1], int(line[2]), line[3], line[4], line[5]])
ex4RDD = ex3RDD.filter(lambda line :int(line[2]) < 25)
ex5RDD = ex4RDD.map(lambda x: (x[0], x))
print ex5RDD.take(5)
```

Output:

```
[(u'Sam', [u'Sam', u'16', 9, u'16', u'28', u'4']), (u'Jill', [u'Jill', u'10', 24, u'43', u'31', u'4']), (u'Jill', [u'Ji
l', u'11', 22, u'50', u'34', u'4']), (u'Joe', [u'Joe', u'24', 7, u'17', u'6', u'3']), (u'Mel', [u'Mel', u'42', 5, u'26'
u'3', u'2'])]
```

Que 6:

Command:

```
lines=sc.textFile('/user/maria_dev/foodratings63666.txt')
ex2RDD = lines.map(lambda line: line.split(","))
ex3RDD = ex2RDD.map(lambda line : [line[0], line[1], int(line[2]), line[3], line[4], line[5]])
ex4RDD = ex3RDD.filter(lambda line :int(line[2]) < 25)
ex5RDD = ex4RDD.map(lambda x: (x[0], x))
ex6RDD = ex5RDD.sortByKey(True)
print ex6RDD.take(5)
```

Output:

```
[(u'Jill', [u'Jill', u'10', 24, u'43', u'31', u'4']), (u'Jill', [u'Jill', u'11', 22, u'50', u'34', u'4']), (u'Jill',
Jill', u'1', 7, u'48', u'18', u'3']), (u'Jill', [u'Jill', u'18', 13, u'20', u'49', u'4']), (u'Jill', [u'Jill', u'31'
u'46', u'18', u'1'])]
```