EX1

```
ex1RDD = sc.textFile("hdfs:/user/maria_dev/foodratings204245.txt")
ex1RDD.take(5):
```

```
93, took 0.499158 s
[u'Joy,45,23,3,1,5', u'Joe,1,40,50,9,4', u'Joe,5,49,40,29,3', u'Sam,30,7,1,46,4'
, u'Mel,30,42,21,37,5']
```

Magic number is 204245

EX2

```
ex2RDD = ex1RDD.map(lambda line:line.split(","))
ex2RDD.take(5)
```

```
93, took 0.042014 s
[[u'Joy', u'45', u'23', u'3', u'1', u'5'], [u'Joe', u'1', u'40', u'50', u'9', u'
4'], [u'Joe', u'5', u'49', u'40', u'29', u'3'], [u'Sam', u'30', u'7', u'1', u'46
', u'4'], [u'Mel', u'30', u'42', u'21', u'37', u'5']]
>>> 
```

EX3

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

```
ex3RDD.take(5)
```

```
[[u'Joy', u'45', 23, u'3', u'1', u'5'], [u'Joe', u'1', 40, u'50', u'9', u'4'], [
u'Joe', u'5', 49, u'40', u'29', u'3'], [u'Sam', u'30', 7, u'1', u'46', u'4'], [u
'Mel', u'30', 42, u'21', u'37', u'5']]
```

Ex4

```
>>> ex4RDD = ex3RDD.filter(lambda line:line[2]<25)
for x in ex4RDD.take(5):
    print(x)
```

```
[[u'Joy', u'45', 23, u'3', u'1', u'5'], [u'Sam', u'30', 7, u'1', u'46', u'4'], [
u'Joy', u'31', 3, u'16', u'6', u'5'], [u'Joe', u'38', 14, u'30', u'41', u'1'], [
u'Jill', u'48', 10, u'6', u'21', u'3']]
>>>
```

Ex5

```
ex5RDD = ex4RDD.map(lambda line:(line[0],line))
ex5RDD.take(5)
```

```
[(u'Joy', [u'Joy', u'45', 23, u'3', u'1', u'5']), (u'Sam', [u'Sam', u'30', 7, u'
1', u'46', u'4']), (u'Joy', [u'Joy', u'31', 3, u'16', u'6', u'5']), (u'Joe', [u'
Joe', u'38', 14, u'30', u'41', u'1']), (u'Jill', [u'Jill', u'48', 10, u'6', u'21
', u'3'])]
>>>
```

Ex6

```
ex6RDD = ex5RDD.sortByKey(True)
ex6RDD.take(5)
```

```
93, took 0.228421 s
[(u'Jill', [u'Jill', u'48', 10, u'6', u'21', u'3']), (u'Jill', [u'Jill', u'49',
7, u'43', u'2', u'3']), (u'Jill', [u'Jill', u'17', 22, u'23', u'44', u'3']), (u'
Jill', [u'Jill', u'1', 11, u'35', u'1', u'4']), (u'Jill', [u'Jill', u'2', 8, u'1
5', u'24', u'3'])]
>>>
```