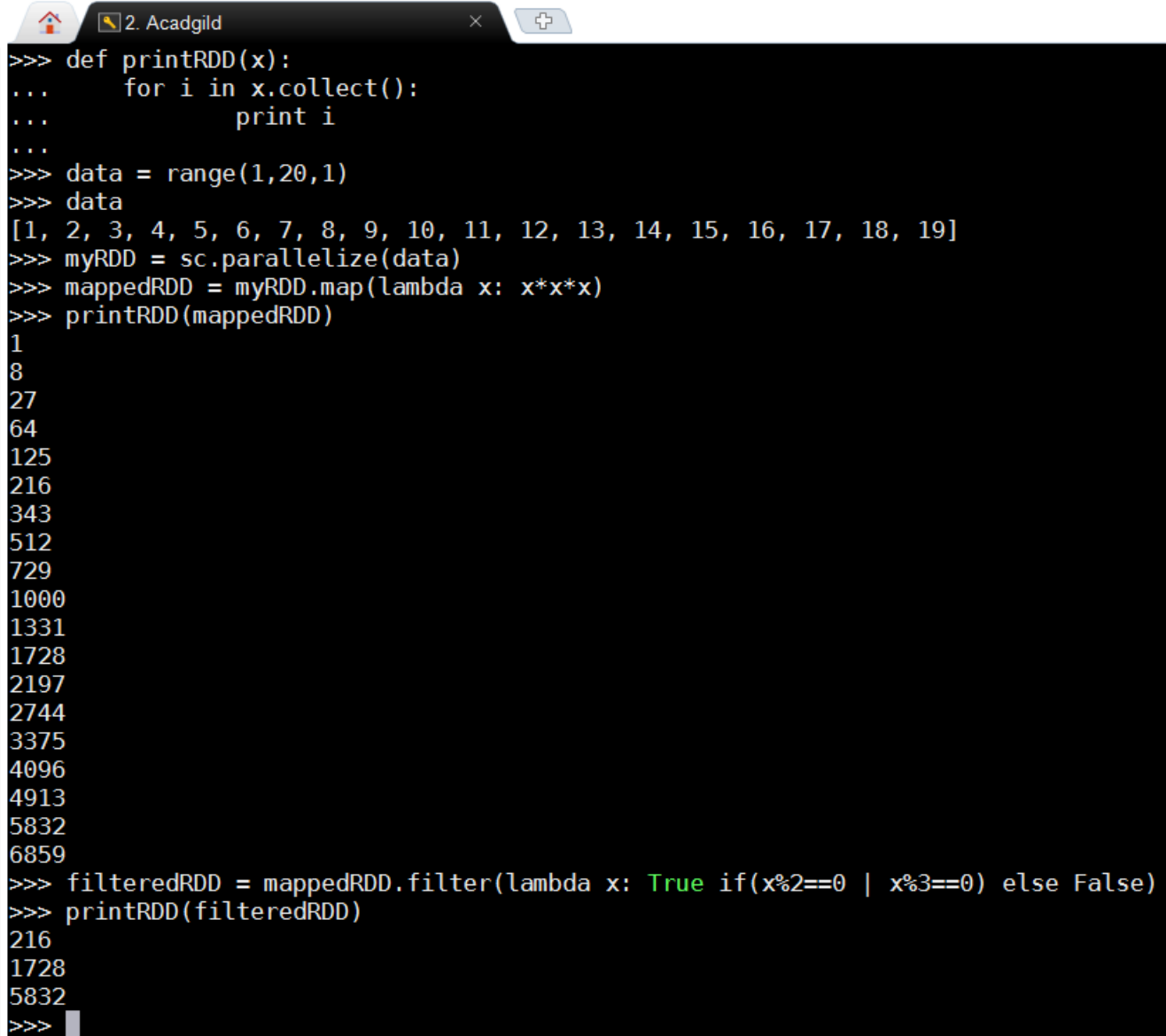


Assignment 13.3

Write the code to Turn a collection into a RDD and perform map operation on it to cube every number and filter the number which are divided by two and three.



```
>>> def printRDD(x):
...     for i in x.collect():
...         print i
...
>>> data = range(1,20,1)
>>> data
[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]
>>> myRDD = sc.parallelize(data)
>>> mappedRDD = myRDD.map(lambda x: x*x*x)
>>> printRDD(mappedRDD)
1
8
27
64
125
216
343
512
729
1000
1331
1728
2197
2744
3375
4096
4913
5832
6859
>>> filteredRDD = mappedRDD.filter(lambda x: True if (x%2==0 | x%3==0) else False)
>>> printRDD(filteredRDD)
216
1728
5832
>>> █
```

```
def printRDD(x):                #created function to print RDD

    for i in x.collect():

        print i
```

```
data = range(1,20,1)          # created list of integers

data

myRDD = sc.parallelize(data)   # created RDD with the data list

mappedRDD = myRDD.map(lambda x: x*x*x)      # performed cube of each element and stored in mappedRDD

printRDD(mappedRDD)           # printed mappedRDD

filteredRDD = mappedRDD.filter(lambda x: True if(x%2==0 | x%3==0) else False)  #filetered RDD with number divisible by 2 & 3

printRDD(filteredRDD)         # printed fileteredRDD
```