

190905514
SEM : 6th

MOHAMMAD TOFIK
ROLL NO : 62

BATCH : C2
SECTION : C

-: DS-LAB-WEEK-5 :-

1. mapper.py

```
import sys
import pandas as pd
df = pd.read_csv("in.csv")
l = df['Country/Region'].tolist()
for x in l:
    print("%s\t%s"%(x,1))
```

reducer.py

```
from operator import itemgetter
import sys
currentWord = None
currentCount = 0
word = None
for line in sys.stdin:
    line = line.strip()
    word, count = line.split('\t', 1)
    try:
        count = int(count)
    except ValueError:
        continue
    if(currentWord == word):
        currentCount += count
    else:
        if(currentWord):
            print('%s\t%s'%(currentWord, currentCount))
            currentCount = count
            currentWord = word
        if (currentWord == word):
            print('%s\t%s'%(currentWord, currentCount))
```

```
$ python3 mapper.py | sort | python3 reducer.py
Afghanistan      213
Albania 199
Algeria 212
Andorra 206
Angola 188
Antigua and Barbuda 195
Argentina 205
Armenia 207
Aruba 7
Australia 1804
Austria 212
Azerbaijan 208
Bahamas 186
Bahamas, The 3
Bahrain 213
Bangladesh 200
Barbados 191
```

2. freqmap1.py

```
from __future__ import print_function
import sys
import pandas as pd
df = pd.read_csv('covid_19_data.csv')
country = df['Deaths']
for word in country:
    print(word, '\t', 1)
```

freqred.py

```
from __future__ import print_function
import sys
lastWord = None
sum = 0
for line in sys.stdin:
    word, count = line.strip().split('\t', 1)
    count = int(count)
    if(lastWord == None):
        lastWord = word
        sum = count
        continue
    if(word == lastWord):
        sum += count
    else:
        print(lastWord, '\t', sum)
        sum = count
        lastWord = word
```

```
if(lastWord == word):  
    print(lastWord, '\t', sum)
```

freqmap2.py

```
from __future__ import print_function  
import sys  
for line in sys.stdin:  
    word, count = line.strip().split('\t', 1)  
    count = int(count)  
    print(count, '\t', word)
```

freqred2.py

```
from __future__ import print_function  
import sys  
mostFreq = []  
currentMax = -1  
for line in sys.stdin:  
    count, word = line.strip().split('\t', 1)  
    count = int(count)  
    if(count > currentMax):  
        currentMax = count  
        mostFreq = [word]  
    elif(count == currentMax):  
        mostFreq.append(word)  
    for word in mostFreq:  
        print(word, '\t', currentMax)
```

```
(kali㉿kali)-[~/Desktop]  
$ cat 1.txt | python3 freqmap1.py | sort | python3 freqred1.py | python3 freqmap2.py | sort | python3  
freqred2.py  
Amex    13
```

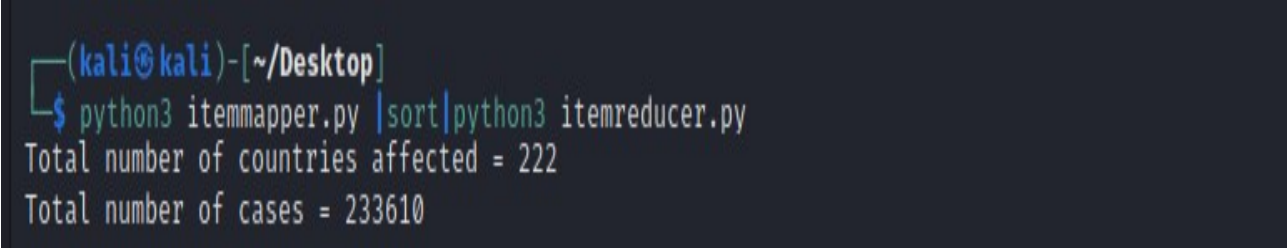
```
(kali㉿kali)-[~/Desktop]  
$ cat in.csv | python3 freqmap1.py | sort | python3 freqred1.py | python3 freqmap2.py | sort | python  
3 freqred2.py  
and     2675
```

3. mapper.py

```
import pandas as pd
df = pd.read_csv("c19.csv")
l = df['Country/Region'].tolist() for x in l:
print('%s\t%s'%(x, 1))
```

reducer.py

```
from operator import itemgetter import sys
current_word = None current_count = 0
country_count = 0 word = None
for line in sys.stdin:
line = line.strip()
word, count = line.split('\t', 1) try:
count = int(count) except ValueError:
continue current_count+=count
if current_word != word:
country_count+=1 current_word = word
if current_word == word: country_count+=count
print("Total number of Countries affected: %d\nTotal number of cases: %d"%
(state_count,current_count))
```



```
(kali@kali)-[~/Desktop]
└─$ python3 itemmapper.py |sort|python3 itemreducer.py
Total number of countries affected = 222
Total number of cases = 233610
```

4.mapper.py

```
import sys
def read_input(file):
for line in file:
yield line.split()

def main(separator='\t'):
data = read_input(sys.stdin) for words in data:
for word in words:
print ('%s%s%d' % (word, separator, 1))

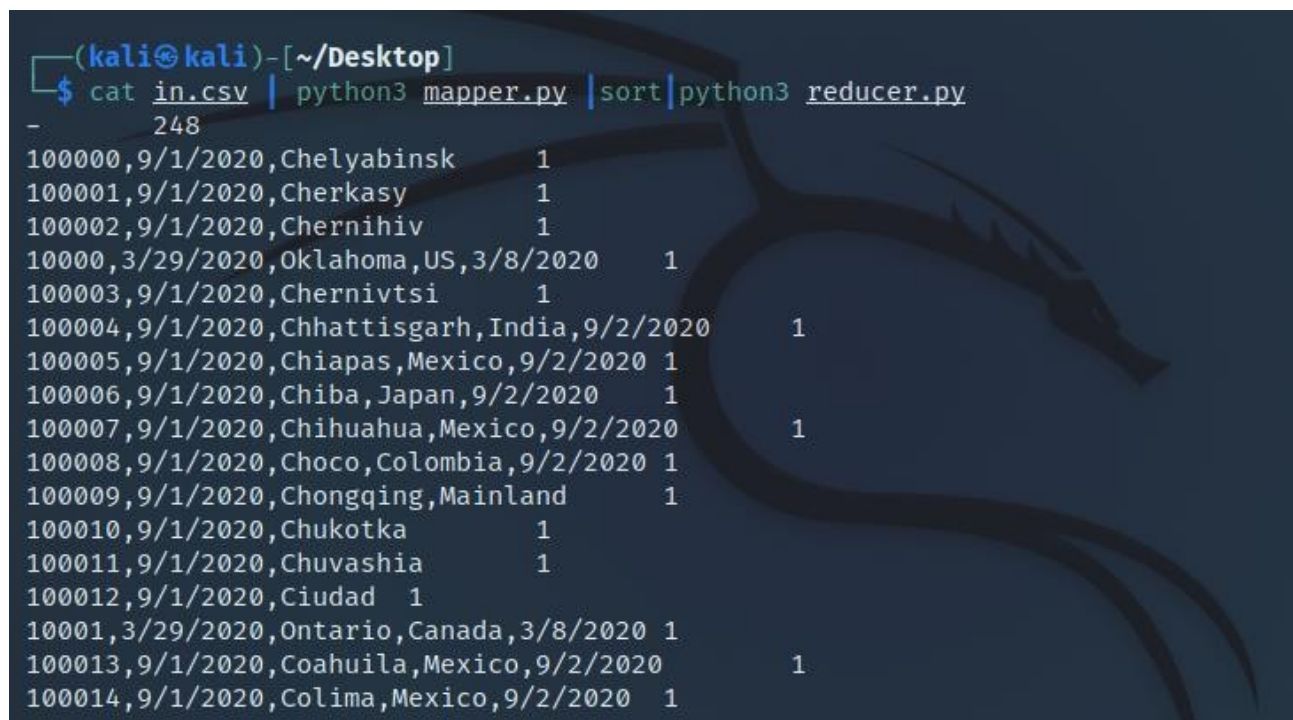
if name == " main ": main()
```

reducer.py

```
from itertools import groupby from operator import itemgetter import sys
def read_mapper_output(file, separator='\t'): for line in file:
yield line.rstrip().split(separator, 1)

def main(separator='\t'):
data = read_mapper_output(sys.stdin, separator=separator) for current_word,
group in groupby(data, itemgetter(0)):
try:
total_count = sum(int(count) for current_word, count in group) print ("%s%s%d"
% (current_word, separator, total_count))
except ValueError: pass

if name == " main ": main()
```



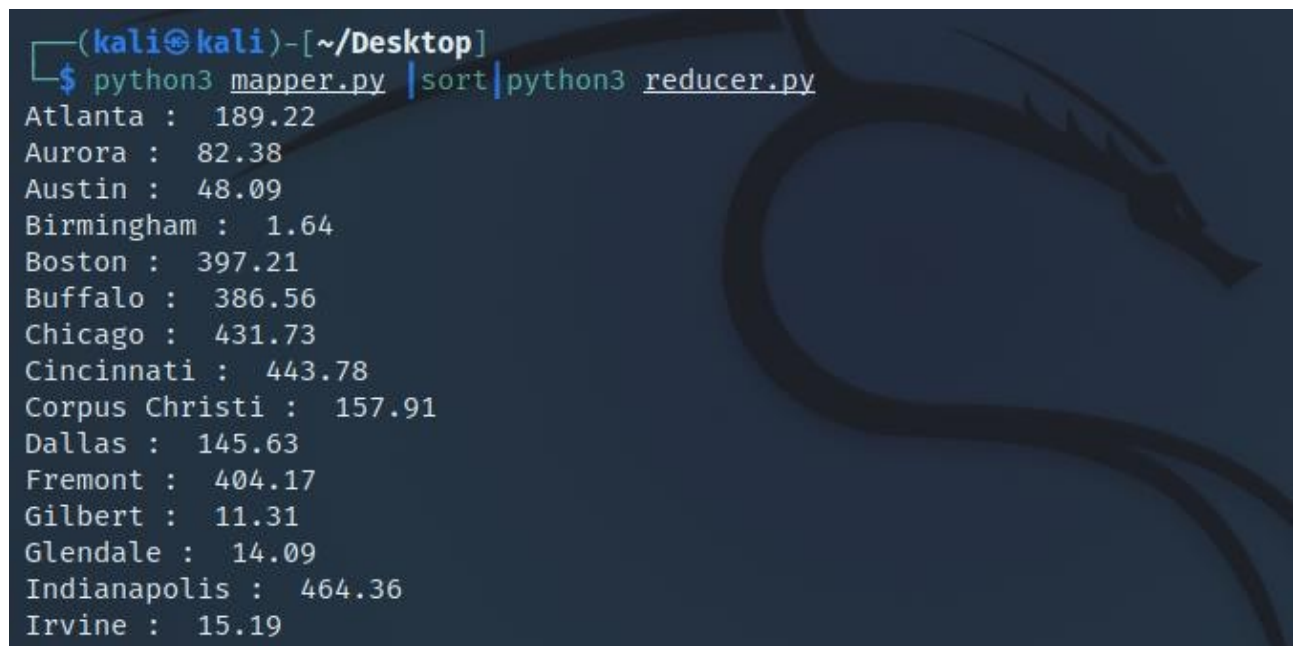
```
(kali@kali)-[~/Desktop]
$ cat in.csv | python3 mapper.py | sort | python3 reducer.py
-
248
100000,9/1/2020,Chelyabinsk 1
100001,9/1/2020,Cherkasy 1
100002,9/1/2020,Chernihiv 1
10000,3/29/2020,Oklahoma,US,3/8/2020 1
100003,9/1/2020,Chernivtsi 1
100004,9/1/2020,Chhattisgarh,India,9/2/2020 1
100005,9/1/2020,Chiapas,Mexico,9/2/2020 1
100006,9/1/2020,Chiba,Japan,9/2/2020 1
100007,9/1/2020,Chihuahua,Mexico,9/2/2020 1
100008,9/1/2020,Choco,Colombia,9/2/2020 1
100009,9/1/2020,Chongqing,Mainland 1
100010,9/1/2020,Chukotka 1
100011,9/1/2020,Chuvashia 1
100012,9/1/2020,Ciudad 1
10001,3/29/2020,Ontario,Canada,3/8/2020 1
100013,9/1/2020,Coahuila,Mexico,9/2/2020 1
100014,9/1/2020,Colima,Mexico,9/2/2020 1
```

5.mapper.py

```
import sys
import pandas as pd
df = pd.read_table("example.txt",index_col=False)
df.columns=['Date', 'Time', 'Location', 'Item', 'Cost', 'Payment']
df.dropna(axis=0,subset=['Location'],inplace=True)
df = df.iloc[:,2:5:2]
d = df.to_records(index=False) for x in d:
print("%s\t%s"%(x[0],x[1]))
```

reducer.py

```
import sys
res = dict()
for line in sys.stdin:
    location, price = line.strip().split('\t', 2)
    try:
        price = float(price)
    except ValueError:
        continue
    if res.get(location, price) == price:
        res[location] = price
    d = res.get(location)
    if price > d:
        res[location] = price
for key, val in res.items():
    print(key, ": ", str(val))
```



```
(kali㉿kali)-[~/Desktop]
$ python3 mapper.py | sort | python3 reducer.py
Atlanta : 189.22
Aurora : 82.38
Austin : 48.09
Birmingham : 1.64
Boston : 397.21
Buffalo : 386.56
Chicago : 431.73
Cincinnati : 443.78
Corpus Christi : 157.91
Dallas : 145.63
Fremont : 404.17
Gilbert : 11.31
Glendale : 14.09
Indianapolis : 464.36
Irvine : 15.19
```

6.mapper.py

```
import sys
def f(x):
    return 4.0/(1.0+x*x)
```

```
N = int(input("Enter number of intervals for integration: "))
dx = 1.0/N
for i in range(N):
    print("%1.10f"%(f(i*dx)*dx))
```

reducer.py

```
from operator import itemgetter
import sys
current_year = None
year_tot = [0]
year_male = [0]
year_fem = [0]
month_tot = [0]
```

```

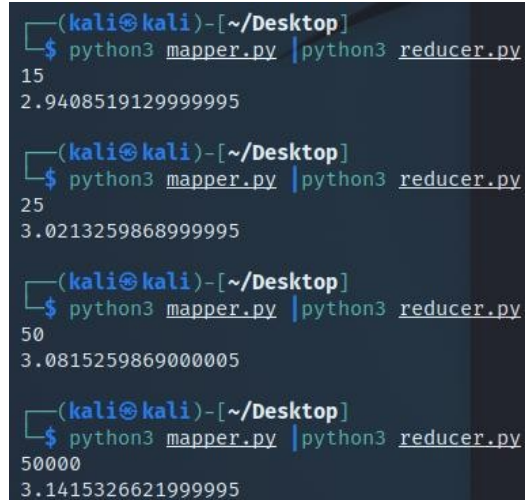
month_male = [0]
month_fem = [0]
specyear = 2000
i = 1
while i<40:
    year_tot.append(0)
    year_male.append(0)
    year_fem.append(0)
    i += 1
    i = 0
    while i<12:
        month_tot.append(0)
        month_male.append(0)
        month_fem.append(0)
        i += 1
        year = None
        for line in sys.stdin:
            line = line.strip()
            sex, month, year = line.split(' ')
            sex = int(sex)
            month = int(month)
            year = int(year)
            year_tot[year-1980] += 1
            if sex == 0:
                year_male[year-1980] += 1
            else:
                year_fem[year-1980] += 1
            if year == specyear:
                month_tot[month-1] += 1
                if sex == 0:
                    month_male[month-1] += 1
                else:
                    month_fem[month-1] += 1
            i = 0
        while i<40:
            if year_tot[i] == 0:
                i += 1
                continue
            print('Year %d Total: %d' %(i+1980, year_tot[i]))
            print('Males: %d' %(year_male[i]))
            print('Females: %d' %(year_fem[i]))
            print('\n')
            i += 1
            print('Year %d'% (specyear))
            i = 0
        while i<12:
            if month_tot[i] == 0:
                i += 1
                continue
            print('Month %d Total: %d' %(i+1, month_tot[i]))
            print('Males: %d' %(month_male[i]))

```

```

print('Females: %d' %(month_fem[i]))
print('\n')
i += 1

```



```

(kali@kali)~/Desktop
$ python3 mapper.py | python3 reducer.py
15
2.9408519129999995

(kali@kali)~/Desktop
$ python3 mapper.py | python3 reducer.py
25
3.0213259868999995

(kali@kali)~/Desktop
$ python3 mapper.py | python3 reducer.py
50
3.0815259869000005

(kali@kali)~/Desktop
$ python3 mapper.py | python3 reducer.py
50000
3.1415326621999995

```

7.mapper.py

```

import numpy as np import sys
l = list(np.random.randint(low=1,high=1000,size=150)) for i in l:
print("%d"%(i))

```

reducer.py

```

from operator import itemgetter
import sys
odd_count = 0
even_count = 0
for line in sys.stdin:
line = line.strip()
num = line.split()
for currnum in num:
try:
odd_count = int(odd_count)
even_count = int(even_count)
currnum = int(currnum)
except ValueError:
continue
if currnum%2 == 0:
even_count += 1
else:
odd_count += 1
print ('%s odd and %s even'%(odd_count, even_count))

```



```
(kali㉿kali)-[~/Desktop]  
└─$ python3 mapper.py |python3 reducer.py  
Even Count: 82  
Odd Count: 68
```

```
(kali㉿kali)-[~/Desktop]  
└─$ python3 mapper.py |python3 reducer.py  
Even Count: 74  
Odd Count: 76
```