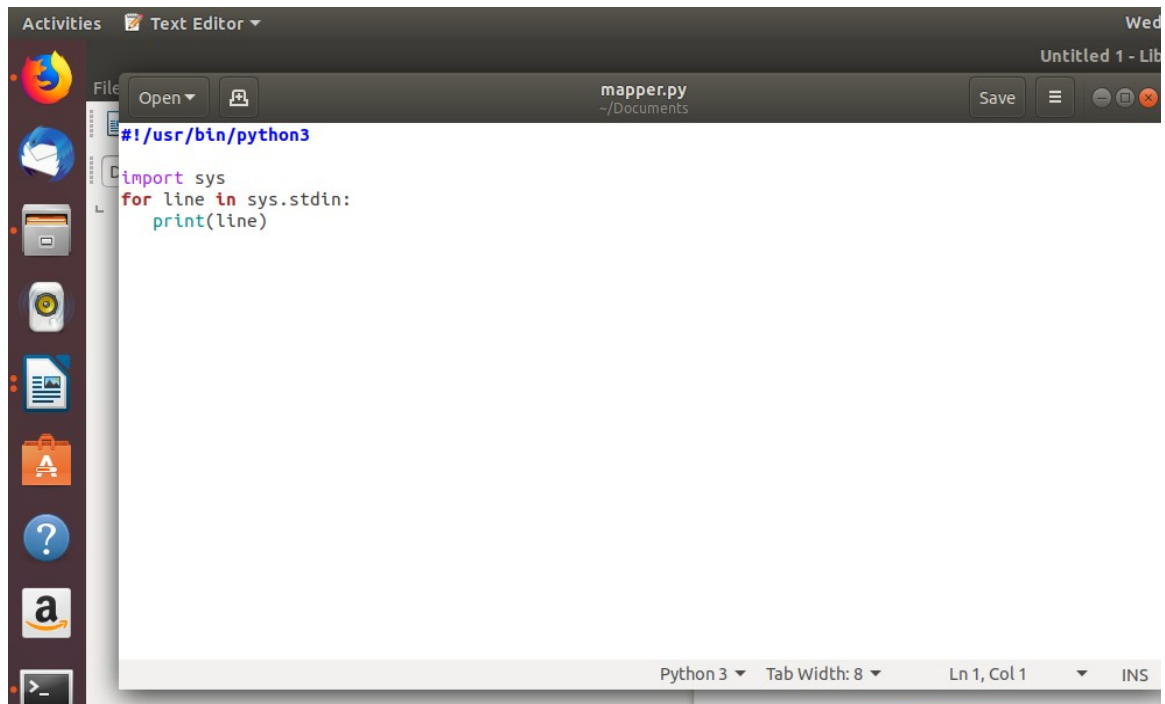


ASSIGNMENT-4

JOINING

MAPPER code:

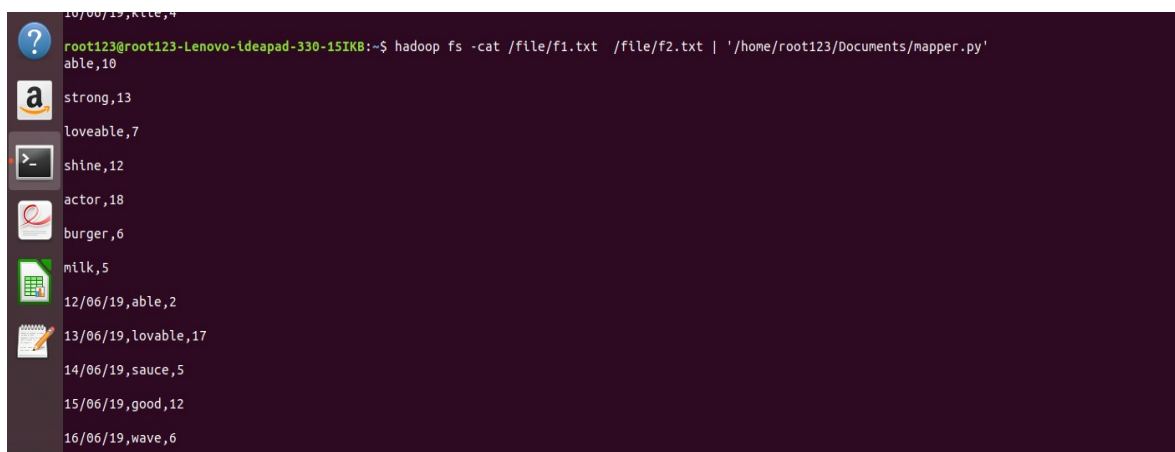


A screenshot of a Linux desktop environment. The 'Activities' window is open, showing a list of applications on the left sidebar. The 'Text Editor' window is active, displaying a file named 'mapper.py' located in the '~/.Documents' directory. The code in the file is as follows:

```
#!/usr/bin/python3
import sys
for line in sys.stdin:
    print(line)
```

The status bar at the bottom of the text editor shows 'Python 3', 'Tab Width: 8', 'Ln 1, Col 1', and 'INS'.

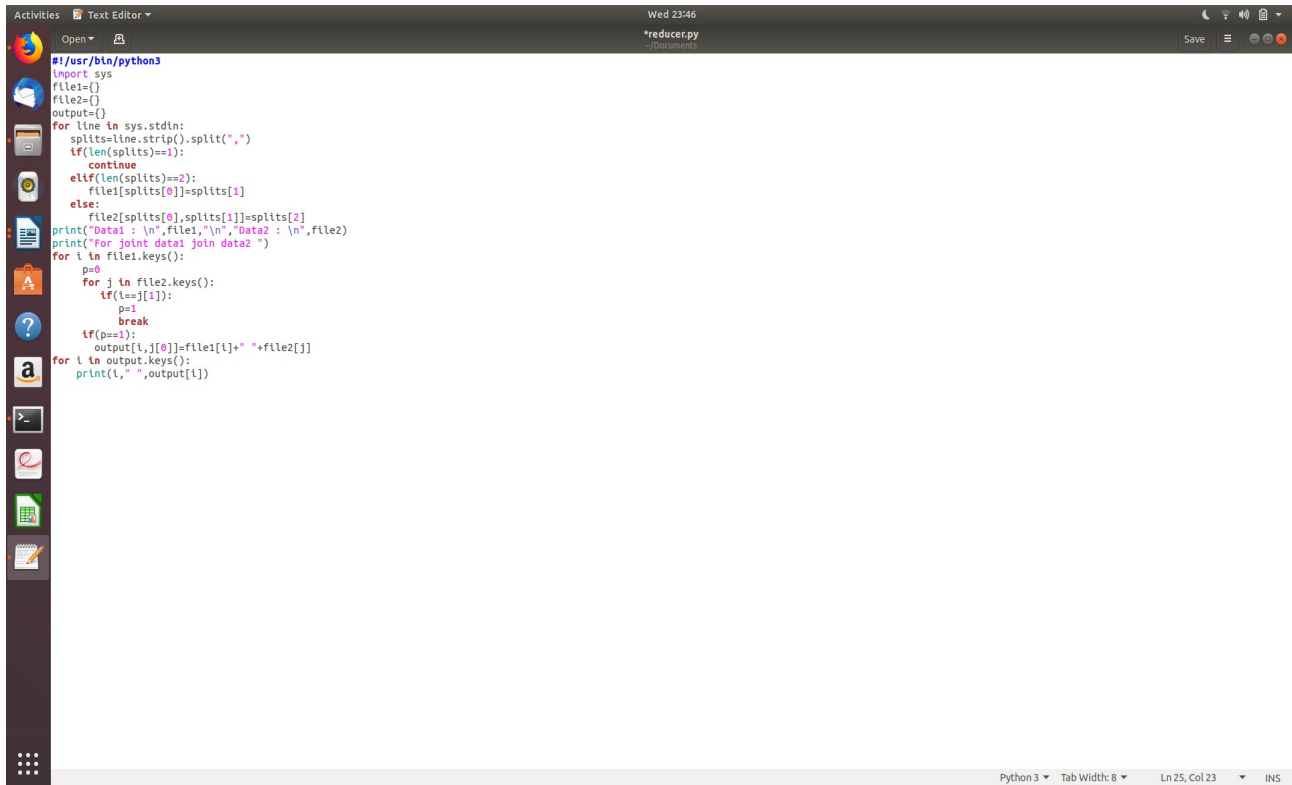
MAPPER OUTPUT:



A screenshot of a Linux terminal window. The prompt is 'root123@root123-Lenovo-ideapad-330-15IK8:~\$'. The command executed is 'hadoop fs -cat /file/f1.txt /file/f2.txt | '/home/root123/Documents/mapper.py''. The output of the command is as follows:

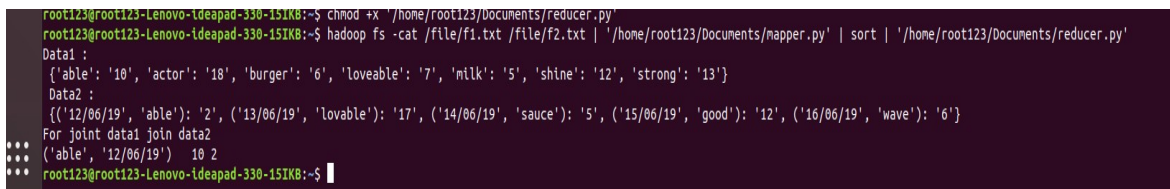
```
10/06/19,kite,4
able,10
strong,13
loveable,7
shine,12
actor,18
burger,6
milk,5
12/06/19,able,2
13/06/19,loveable,17
14/06/19,sauce,5
15/06/19,good,12
16/06/19,wave,6
```

REDUCER CODE:



```
#!/usr/bin/python3
import sys
file1={}
file2={}
output={}
for line in sys.stdin:
    splits=line.strip().split(",")
    if(len(splits)==1):
        continue
    elif(len(splits)==2):
        file1[splits[0]]=splits[1]
    else:
        file2[splits[0],splits[1]]=splits[2]
print("Data1 : \n",file1,"\n","Data2 : \n",file2)
print("For joint data1 join data2 ")
for i in file1.keys():
    p=0
    for j in file2.keys():
        if(i==j[1]):
            p=i
            break
    if(p==1):
        output[i,j[0]]=file1[i]+" "+file2[j]
for i in output.keys():
    print(i," ",output[i])
```

REDUCER OUTPUT:



```
root123@root123-Lenovo-Ideapad-330-15IK8:~$ chmod +x /home/root123/Documents/reducer.py
root123@root123-Lenovo-Ideapad-330-15IK8:~$ hadoop fs -cat /file/f1.txt /file/f2.txt | '/home/root123/Documents/mapper.py' | sort | '/home/root123/Documents/reducer.py'
Data1 :
{'able': '10', 'actor': '18', 'burger': '6', 'loveable': '7', 'milk': '5', 'shine': '12', 'strong': '13'}
Data2 :
{('12/06/19', 'able'): '2', ('13/06/19', 'loveable'): '17', ('14/06/19', 'sauce'): '5', ('15/06/19', 'good'): '12', ('16/06/19', 'wave'): '6'}
For joint data1 join data2
('able', '12/06/19')    10 2
root123@root123-Lenovo-Ideapad-330-15IK8:~$
```

