

Assignment No 3

Develop a MapReduce program to find the grades of students.

gedit student.csv

```
Name,Score,
kaustubh,90
Srushti,100
Anuja,98
Sammrudhi,90
Rutuja,98
Tejashree,89
sanskriti,80
nikita,70
sneha,70
shreyash,95
kartik,90
```

gedit reducer.py

```
#!/usr/bin/env python
import sys
```

```
for line in sys.stdin:
```

```
    line = line.strip()
    if not line:
        continue
```

```
    parts = line.split("\t")
    if len(parts) != 2:
        continue
```

```
    name, grade = parts
    print("{0},{1}".format(name, grade))
```

gedit mapper.py

```
#!/usr/bin/env python
import sys
import csv
```

```
reader = csv.reader(sys.stdin)
next(reader) # Skip header
```

```
for row in reader:
    if len(row) != 2:
        continue

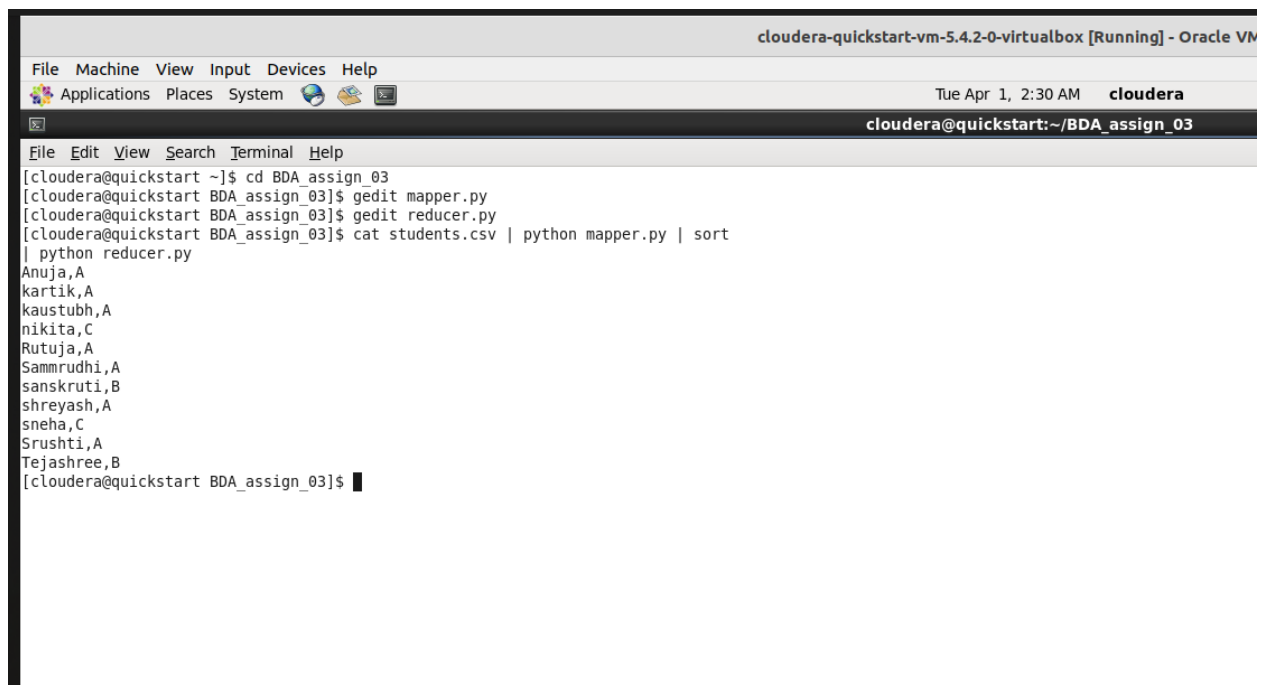
    name, score = row
    try:
        score = int(score)
    except ValueError:
        continue

    if score >= 90:
        grade = "A"
    elif score >= 80:
        grade = "B"
    elif score >= 70:
        grade = "C"
    elif score >= 60:
        grade = "D"
    else:
        grade = "F"

    print("{0}\t{1}".format(name, grade))
```

run this command to get output

cat students.csv | python mapper.py | sort | python reducer.py



```
cloudera-quickstart-vm-5.4.2-0-virtualbox [Running] - Oracle VM
File Machine View Input Devices Help
Applications Places System
Tue Apr 1, 2:30 AM cloudera
cloudera@quickstart:~/BDA_assign_03
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ cd BDA_assign_03
[cloudera@quickstart BDA_assign_03]$ gedit mapper.py
[cloudera@quickstart BDA_assign_03]$ gedit reducer.py
[cloudera@quickstart BDA_assign_03]$ cat students.csv | python mapper.py | sort
| python reducer.py
Anuja,A
Kartik,A
kaustubh,A
nikita,C
Rutuja,A
Sammrudhi,A
sanskriti,B
shreyash,A
sneha,C
Srushti,A
Tejashree,B
[cloudera@quickstart BDA_assign_03]$
```