APPLICATION OF MAPREDUCE

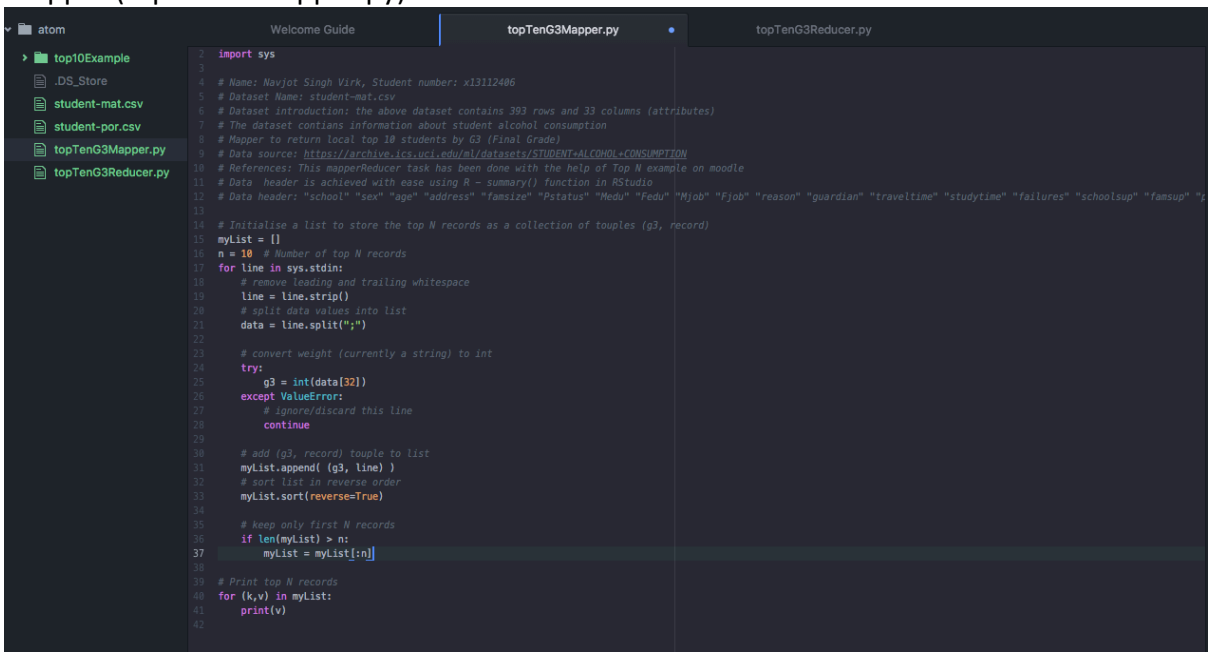
Programming Paradigm for Processing

Students Alcohol Consumption

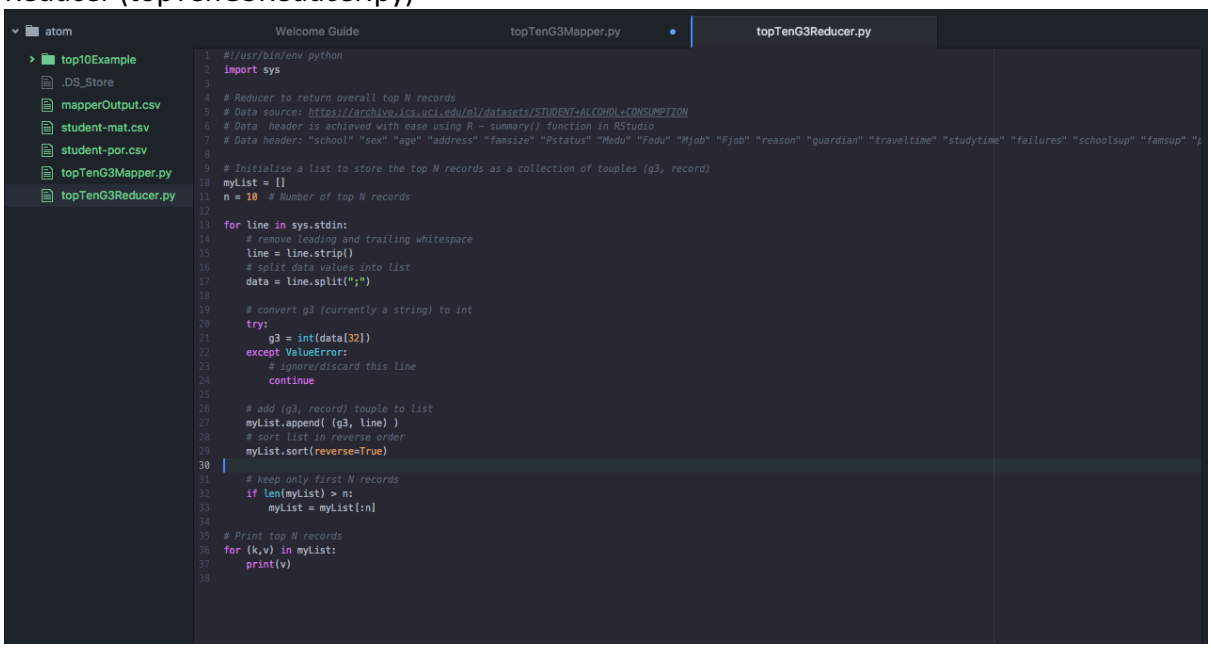
DATASET 1: student-mat.csv
(Mathematics Students)

Implement and Present the algorithms to process the dataset.
The idea is to mapReduce and get top 10 student records with best G3 (Final Grade) and output a csv file and then use that csv to draw some meaningful plots/predictions.

Mapper (topTenG3Mapper.py)



Reducer (topTenG3Reducer.py)



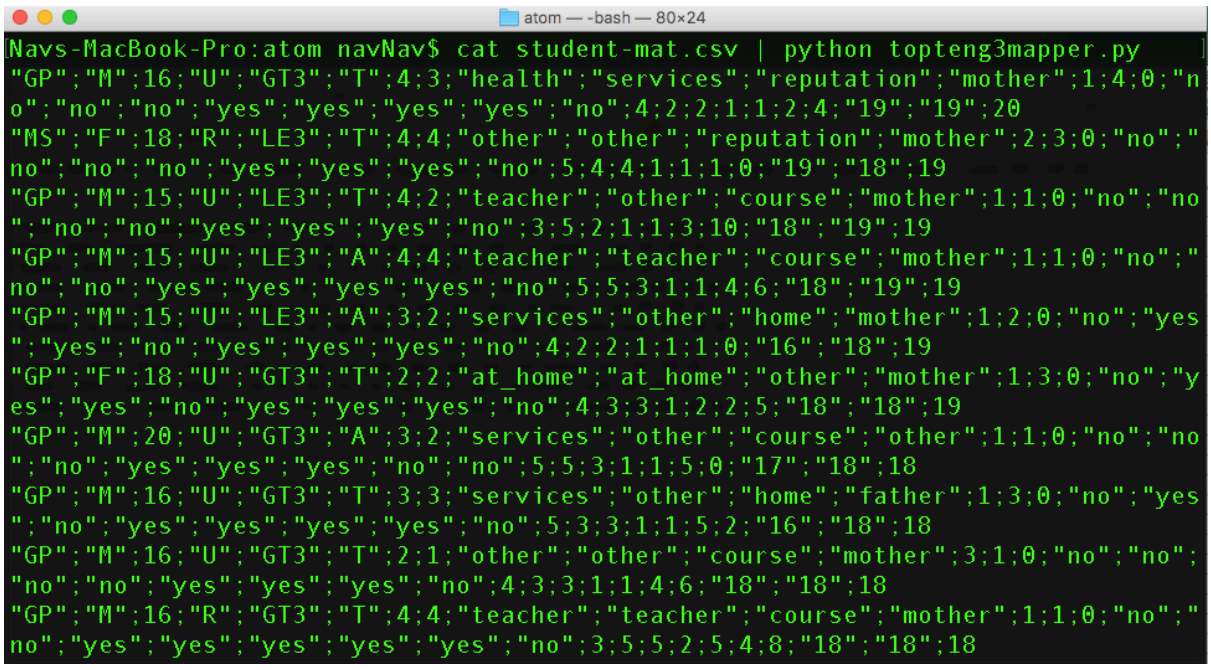
Reference: I am using Top N example available on moodle for this mapReduce processing

Configuration Details

For mapReduce processing, I used Atom editor for editing my .py files, and used mac terminal to run mapper and reducer on csv files and output csv result files and then I took final results csv file and used it in RStudio to plot some graphs.

Present and Discuss your results

1 – Run mapper on the dataset



```
cat student-mat.csv |
topteng3mapper.py
```

2 – Outputting the mapper's result into a csv file

```
Navs-MacBook-Pro:atom navNav$ cat student-mat.csv | python topteng3mapper.py > m
apperOutput.csv
```

```
cat student-mat.csv |
topteng3mapper.py > mapperOutput.csv
```

3– Sorting

```
atom — -bash — 80×24
mapperOutput.csv        top10Example            topTenG3Reducer.py
student-mat.csv         topTenG3Mapper.py
Navs-MacBook-Pro:atom navNav$ cat mapperOutput.csv | sort
"GP";"F";18;"U";"GT3";"T";2;2;"at_home";"at_home";"other";"mother";1;3;0;"no";"y
es";"yes";"no";"yes";"yes";"yes";"no";4;3;3;1;2;2;5;"18";"18";19
"GP";"M";15;"U";"LE3";"A";3;2;"services";"other";"home";"mother";1;2;0;"no";"yes
";"yes";"no";"yes";"yes";"yes";"no";4;2;2;1;1;1;0;"16";"18";19
"GP";"M";15;"U";"LE3";"A";4;4;"teacher";"teacher";"course";"mother";1;1;0;"no";"
no";"no";"yes";"yes";"yes";"yes";"no";5;5;3;1;1;4;6;"18";"19";19
"GP";"M";15;"U";"LE3";"T";4;2;"teacher";"other";"course";"mother";1;1;0;"no";"no
";"no";"no";"yes";"yes";"yes";"no";3;5;2;1;1;3;10;"18";"19";19
"GP";"M";16;"R";"GT3";"T";4;4;"teacher";"teacher";"course";"mother";1;1;0;"no";"
no";"yes";"yes";"yes";"yes";"yes";"no";3;5;5;2;5;4;8;"18";"18";18
"GP";"M";16;"U";"GT3";"T";2;1;"other";"other";"course";"mother";3;1;0;"no";"no";
"no";"no";"yes";"yes";"yes";"no";4;3;3;1;1;4;6;"18";"18";18
"GP";"M";16;"U";"GT3";"T";3;3;"services";"other";"home";"father";1;3;0;"no";"yes
";"no";"yes";"yes";"yes";"yes";"no";5;3;3;1;1;5;2;"16";"18";18
"GP";"M";16;"U";"GT3";"T";4;3;"health";"services";"reputation";"mother";1;4;0;"n
o";"no";"no";"yes";"yes";"yes";"yes";"no";4;2;2;1;1;2;4;"19";"19";20
"GP";"M";20;"U";"GT3";"A";3;2;"services";"other";"course";"other";1;1;0;"no";"no
";"no";"yes";"yes";"yes";"no";"no";5;5;3;1;1;5;0;"17";"18";18
"MS";"F";18;"R";"LE3";"T";4;4;"other";"other";"reputation";"mother";2;3;0;"no";"
no";"no";"no";"yes";"yes";"yes";"no";5;4;4;1;1;1;0;"19";"18";19
Navs-MacBook-Pro:atom navNav$ cat mapperOutput.csv | sort
```

```
cat mapperOutput.csv | sort
```

4 – Run reducer on the mapper results csv file

```
Navs-MacBook-Pro:atom navNav$ cat mapperoutput.csv | sort | python topteng3reduc
er.py
"GP";"M";16;"U";"GT3";"T";4;3;"health";"services";"reputation";"mother";1;4;0;"n
o";"no";"no";"yes";"yes";"yes";"yes";"no";4;2;2;1;1;2;4;"19";"19";20
"MS";"F";18;"R";"LE3";"T";4;4;"other";"other";"reputation";"mother";2;3;0;"no";"
no";"no";"no";"yes";"yes";"yes";"no";5;4;4;1;1;1;0;"19";"18";19
"GP";"M";15;"U";"LE3";"T";4;2;"teacher";"other";"course";"mother";1;1;0;"no";"no
";"no";"no";"yes";"yes";"yes";"no";3;5;2;1;1;3;10;"18";"19";19
"GP";"M";15;"U";"LE3";"A";4;4;"teacher";"teacher";"course";"mother";1;1;0;"no";"
no";"no";"yes";"yes";"yes";"yes";"no";5;5;3;1;1;4;6;"18";"19";19
"GP";"M";15;"U";"LE3";"A";3;2;"services";"other";"home";"mother";1;2;0;"no";"yes
";"yes";"no";"yes";"yes";"yes";"no";4;2;2;1;1;1;0;"16";"18";19
"GP";"F";18;"U";"GT3";"T";2;2;"at_home";"at_home";"other";"mother";1;3;0;"no";"y
es";"yes";"no";"yes";"yes";"yes";"no";4;3;3;1;2;2;5;"18";"18";19
"GP";"M";20;"U";"GT3";"A";3;2;"services";"other";"course";"other";1;1;0;"no";"no
";"no";"yes";"yes";"yes";"no";"no";5;5;3;1;1;5;0;"17";"18";18
"GP";"M";16;"U";"GT3";"T";3;3;"services";"other";"home";"father";1;3;0;"no";"yes
";"no";"yes";"yes";"yes";"yes";"no";5;3;3;1;1;5;2;"16";"18";18
"GP";"M";16;"U";"GT3";"T";2;1;"other";"other";"course";"mother";3;1;0;"no";"no";
"no";"no";"yes";"yes";"yes";"no";4;3;3;1;1;4;6;"18";"18";18
"GP";"M";16;"R";"GT3";"T";4;4;"teacher";"teacher";"course";"mother";1;1;0;"no";"
no";"yes";"yes";"yes";"yes";"yes";"no";3;5;5;2;5;4;8;"18";"18";18
```

```
cat mapperoutput.csv | sort | python
topteng3reducer.py
```

5 – Output the mapper results

```
Navs-MacBook-Pro:atom navNav$ ls
mapperOutput.csv        student-por.csv         topTenG3Reducer.py
outputReducer.csv       top10Example
student-mat.csv         topTenG3Mapper.py
Navs-MacBook-Pro:atom navNav$
```

```
cat mapperoutput.csv | sort | python
topteng3reducer.py >
outputReducer.csv
```