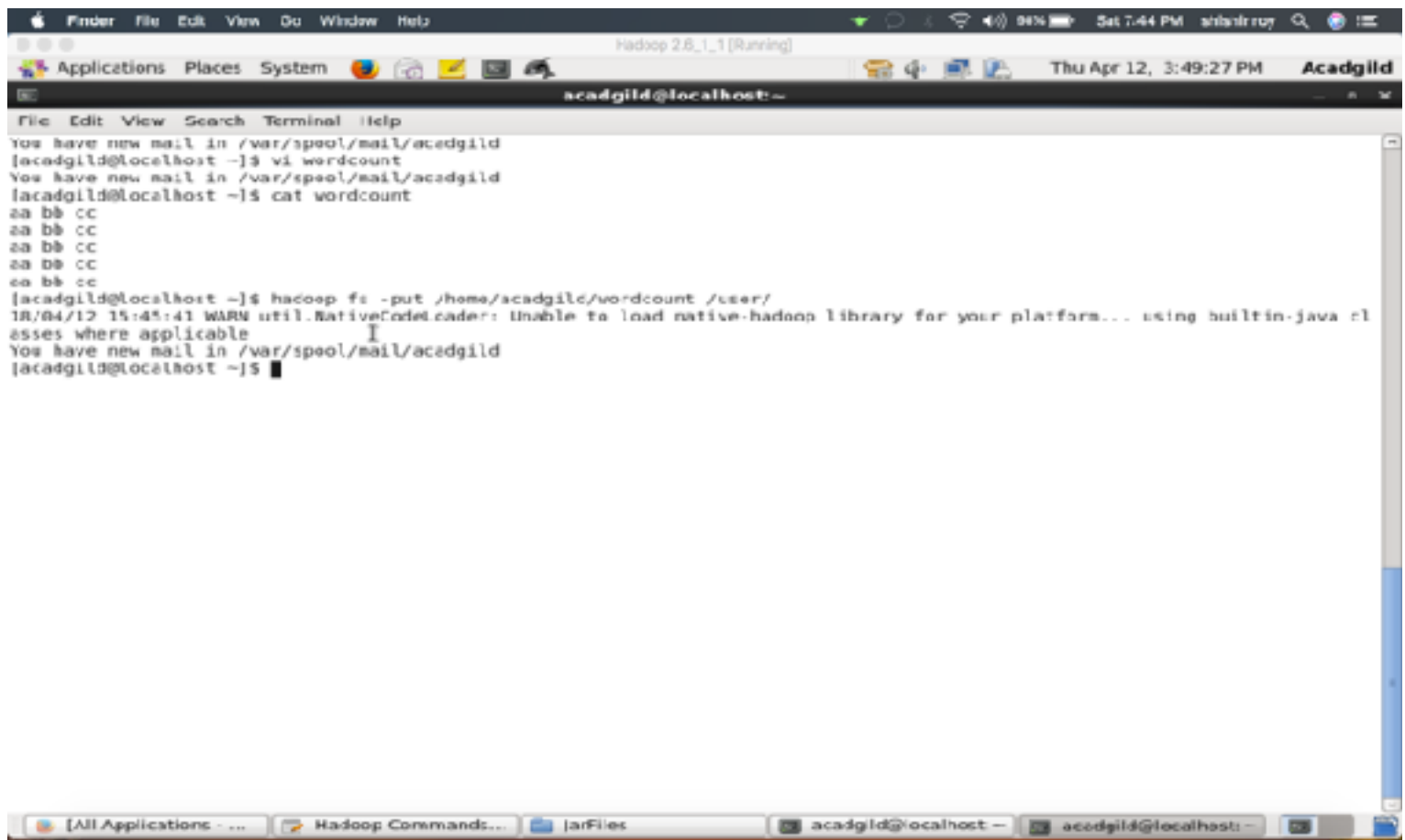
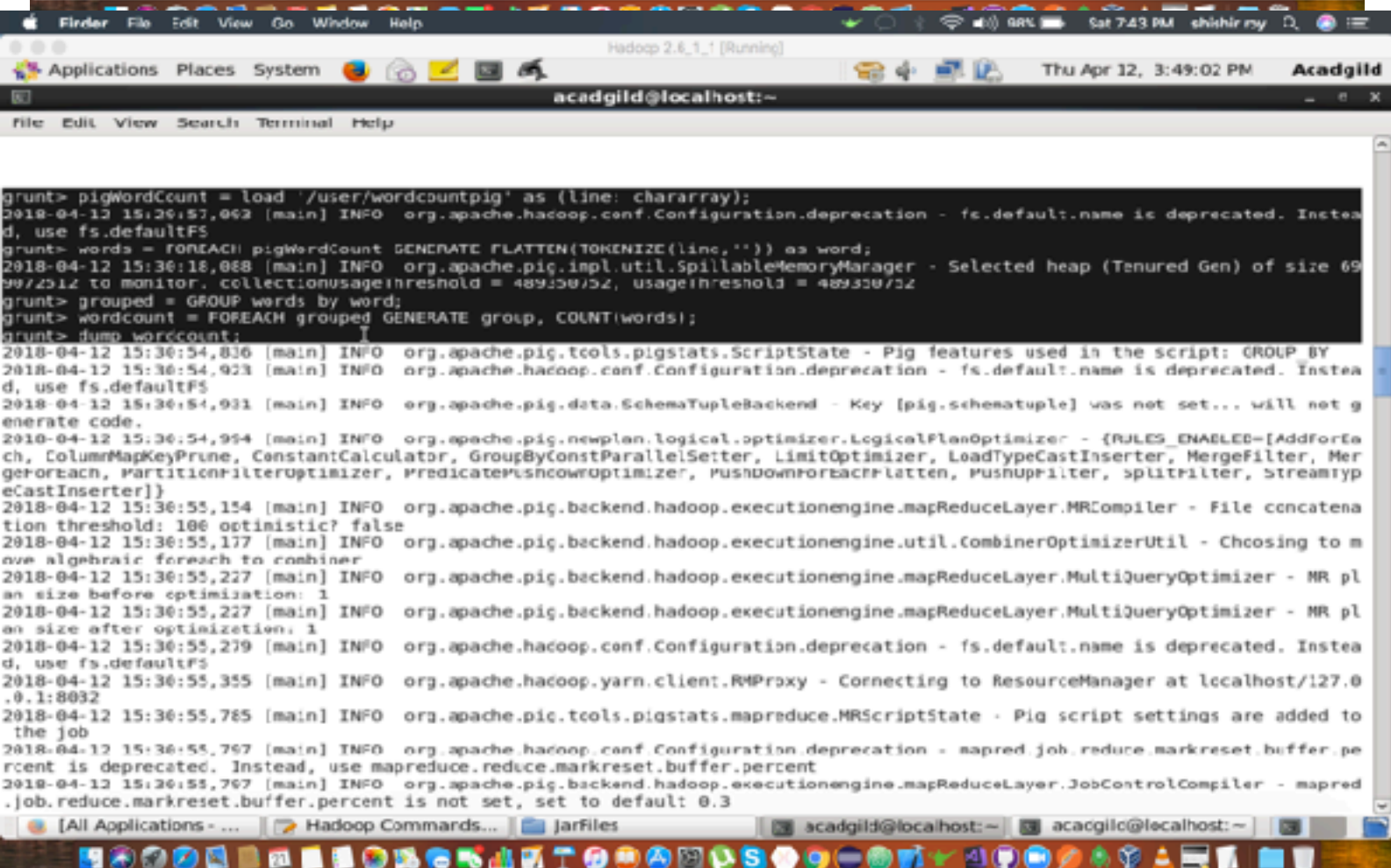


Task 1

Write a program to implement wordcount using Pig.



```
Finder File Edit View Go Window Help
hadoop 2.6.1_1 [Running]
Applications Places System
acacgild@localhost:~
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$ vi wordcount
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$ cat wordcount
aa bb cc
aa bb cc
aa bb cc
aa bb cc
aa bb cc
[acacgild@localhost ~]$ hadoop fs -put /home/acacgild/wordcount /user/
18/04/12 15:45:41 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$
```



```
Finder File Edit View Go Window Help
hadoop 2.6.1_1 [Running]
Applications Places System
acacgild@localhost:~
File Edit View Search Terminal Help
grunt> pigWordCount = load '/user/wordcount/pig' as (line: chararray);
2018-04-12 15:36:57,962 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> words = FOREACH pigWordCount GENERATE FLATTEN(TOKENIZE(line, '')) as word;
2018-04-12 15:36:18,068 [main] INFO org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (Tenured Gen) of size 69
997/212 to monitor. collectionUsageThreshold = 489338/52, usageThreshold = 489338/52
grunt> grouped = GROUP words by word;
grunt> wordcount = FOREACH grouped GENERATE group, COUNT(words);
grunt> dump wordcount;
2018-04-12 15:36:54,836 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP BY
2018-04-12 15:36:54,923 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2018-04-12 15:36:54,931 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not g
enerate code.
2018-04-12 15:36:54,994 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEa
ch, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, Mer
geForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTyp
eCastInserter]}
2018-04-12 15:36:55,154 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatena
tion threshold: 100 optimistic? false
2018-04-12 15:36:55,177 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to m
ove algebraic foreach to combiner
2018-04-12 15:36:55,227 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR pl
an size before optimization: 1
2018-04-12 15:36:55,227 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR pl
an size after optimization: 1
2018-04-12 15:36:55,279 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instea
d, use fs.defaultFS
2018-04-12 15:36:55,355 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0
.0.1:8032
2018-04-12 15:36:55,785 [main] INFO org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-04-12 15:36:55,797 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.reduce.markreset.buffer.p
ercent is deprecated. Instead, use mapreduce.reduce.markreset.buffer.percent
2018-04-12 15:36:55,797 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - mapred
.job.reduce.markreset.buffer.percent is not set, set to default: 0.3
```

A screenshot of a Mac terminal window. The title bar shows 'Finder File Edit View Go Window Help' and 'Hadoop 2.6.1_1 [Running]'. The menu bar includes 'Applications Places System' and system status icons. The terminal prompt is 'acadgild@localhost:~'. The output shows Hadoop job results for 'wordcount', including input/output records, counters, and job DAG. Below this, there are several log messages from the Hadoop ecosystem (YARN, MapReduce, Pig) indicating successful completion and some deprecation warnings. The bottom of the image shows a dock with various application icons.

```
Input(s):
Successfully read 5 records (408 bytes) from: "/user/wordcount"

Output(s):
Successfully stored 1 records (17 bytes) in: "hdfs://localhost:8020/tmp/temp2063247015/tmp-975112111"

Counters:
Total records written : 1
Total bytes written : 17
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_1521868459622_0034

2018-04-12 15:47:44,420 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-12 15:47:44,438 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-12 15:47:44,537 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-12 15:47:44,546 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-12 15:47:44,664 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-12 15:47:44,675 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-12 15:47:44,752 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success
2018-04-12 15:47:44,798 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2018-04-12 15:47:44,798 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
2018-04-12 15:47:44,816 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-04-12 15:47:44,817 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(aa bb cc,5)
```

Task 2

We have employee_details and employee_expenses files. Use local mode while running Pig and write Pig Latin script to get below results:

employee_details (EmpID,Name,Salary,DepartmentID) https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/employee_details.txt

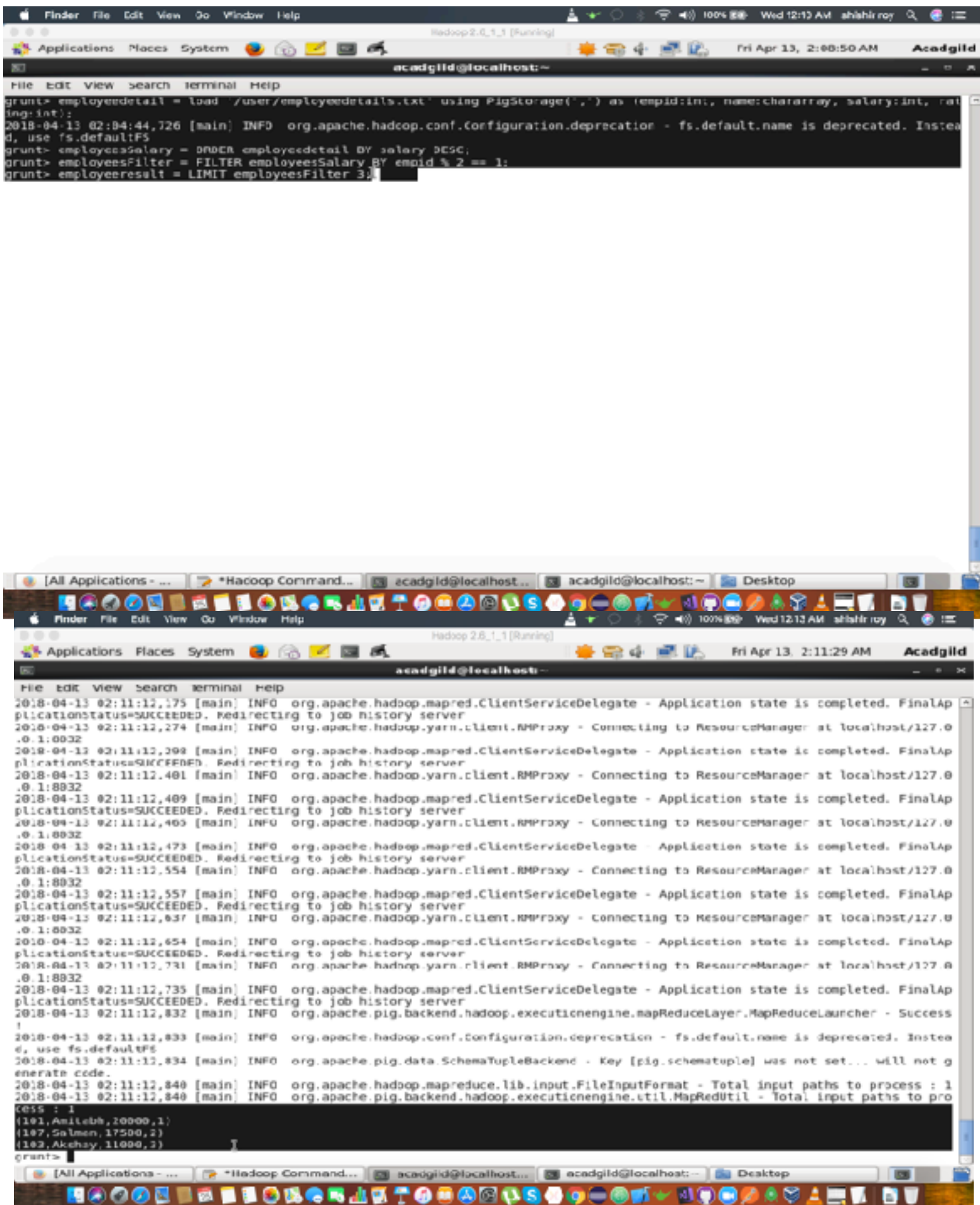
employee_expenses(EmpID,Expense) https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/employee_expenses.txt

(a) Top 5 employees (employee id and employee name) with highest rating.

```
Finder  File  Edit  View  Go  Window  Help
Hadoop 2.8.1_1 [Running]
Applications  Places  System
acdgild@localhost:~
File Edit View Search Terminal Help
grunt> employeeDetail = load '/user/employeeDetails.txt' using PigStorage(',') as (empid:int, name:chararray, salary:int, rating:int);
2018-04-13 01:59:41,547 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> employeeOrder = ORDER employeeDetail BY rating DESC;
grunt> employeeResult = LIMIT employeeOrder 5;
grunt> dump employeeResult;
```

```
Finder  File  Edit  View  Go  Window  Help
Hadoop 2.8.1_1 [Running]
Applications  Places  System
acdgild@localhost:~
File Edit View Search Terminal Help
2018-04-13 02:02:55,214 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-13 02:02:55,229 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 02:02:55,337 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-13 02:02:55,343 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 02:02:55,457 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-13 02:02:55,475 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 02:02:55,619 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-13 02:02:55,647 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 02:02:55,797 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-13 02:02:55,893 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 02:02:55,993 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-04-13 02:02:55,939 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 02:02:56,058 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success
2018-04-13 02:02:56,062 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2018-04-13 02:02:56,064 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
2018-04-13 02:02:56,072 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-04-13 02:02:56,072 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(110,Priyanka,2000,5)
(105,Ravan,2500,5)
(100,Katrina,1000,4)
(104,Anubhav,5000,4)
(108,Rashir,14000,3)
grunt>
```


(b) Top 3 employees (employee id and employee name) with highest salary, whose employee id is an odd number.



The image shows two screenshots of a Mac terminal window. The top screenshot shows the execution of Pig commands to load employee data, sort it by salary, filter for odd employee IDs, and limit the results to 3. The bottom screenshot shows the output of the Pig job, which includes status messages from Hadoop and the final result of the query.

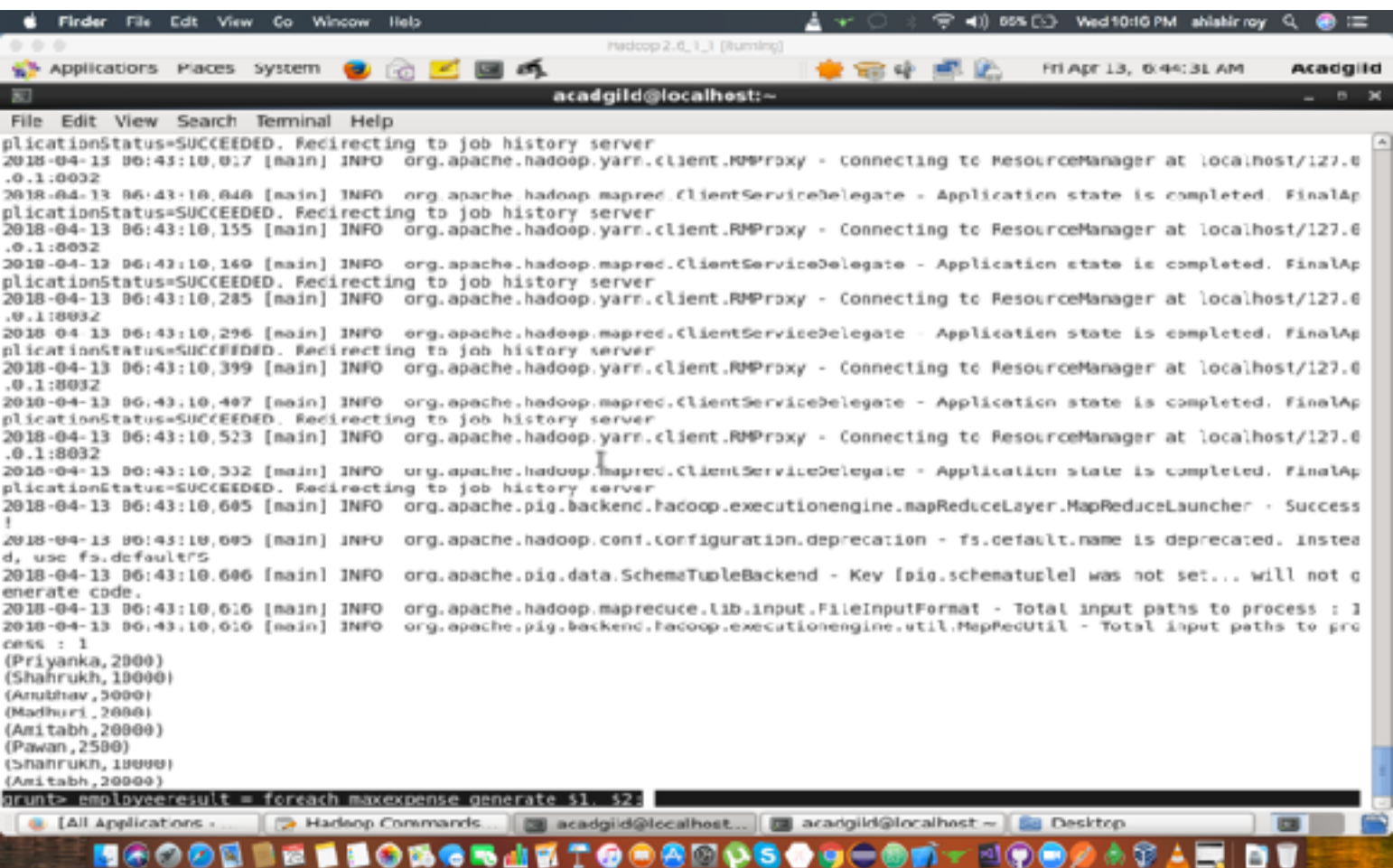
```
acadmild@localhost:~  
grunt> employeeDetail = load '/user/employeeDetails.txt' using PigStorage(',') as (empid:int, name:chararray, salary:int, rating:int);  
2018-04-13 02:04:44,726 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS  
grunt> employeesSalary = ORDER employeeDetail BY salary DESC;  
grunt> employeesFilter = FILTER employeesSalary BY empid % 2 == 1;  
grunt> employeesResult = LIMIT employeesFilter 3;
```

```
2018-04-13 02:11:12,175 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,274 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032  
2018-04-13 02:11:12,398 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,401 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032  
2018-04-13 02:11:12,409 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,409 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032  
2018-04-13 02:11:12,473 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,554 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032  
2018-04-13 02:11:12,557 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,637 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032  
2018-04-13 02:11:12,654 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,731 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.0.1:8032  
2018-04-13 02:11:12,735 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server  
2018-04-13 02:11:12,832 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success  
1  
2018-04-13 02:11:12,833 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS  
2018-04-13 02:11:12,834 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.  
2018-04-13 02:11:12,840 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1  
2018-04-13 02:11:12,840 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1  
(101,Amitebh,20000,1)  
(107,Salmen,17500,2)  
(102,Akshay,11000,2)  
grunt>
```

(c) Employee (employee id and employee name) with maximum expense



```
Finder File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
acadmild@localhost:~
File Edit View Search Terminal Help
grunt> employeeDetail = load '/user/employeeDetails.txt' using PigStorage(',') as (empid:int, name:chararray, salary:int, rat
lag:int);
2018-04-13 06:39:14,761 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instea
d, use fs.DefaultFS
grunt> employeeExpense = load '/user/employeeExpenses.txt' using PigStorage('\t') as (empid:int, expense:int);
2018-04-13 06:39:20,291 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instea
d, use fs.DefaultFS
grunt> employeeData = JOIN employeeDetail BY empid, employeeExpense BY empid;
grunt> maxExpense = GROUP employeeData BY expense DESC;
```



```
Finder File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
acadmild@localhost:~
File Edit View Search Terminal Help
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 06:43:10,017 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0
.0.1:8032
2018-04-13 06:43:10,040 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 06:43:10,155 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0
.0.1:8032
2018-04-13 06:43:10,169 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 06:43:10,285 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0
.0.1:8032
2018-04-13 06:43:10,296 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 06:43:10,399 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0
.0.1:8032
2018-04-13 06:43:10,407 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 06:43:10,523 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0
.0.1:8032
2018-04-13 06:43:10,532 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-13 06:43:10,605 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success
!
2018-04-13 06:43:10,605 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instea
d, use fs.DefaultFS
2018-04-13 06:43:10,606 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not g
enerate code.
2018-04-13 06:43:10,616 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-04-13 06:43:10,616 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapReduceUtil - Total input paths to pro
cess : 1
(Priyanka,2000)
(Shahrukh,10000)
(Aruthav,5000)
(Madhuri,2000)
(Amitabh,20000)
(Pawan,2500)
(Shahrukh,10000)
(Amitabh,20000)
grunt> employeeResult = foreach maxExpense generate $1, $2;
```

(d) List of employees (employee id and employee name) having entries in employee_expenses file.

```
File Edit View Go Window Help
Hadoop 2.6.1 [Shanley]
Applications Places Systems
acsdgild@localhost:~$
File Edit View Search Terminal Help
2018-04-14 03:32:53,606 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-14 03:32:53,718 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost:127.0
.0.1:8032
2018-04-14 03:32:54,722 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-14 03:32:54,848 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost:127.0
.0.1:8032
2018-04-14 03:32:54,848 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-14 03:32:54,948 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost:127.0
.0.1:8032
2018-04-14 03:32:53,940 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-14 03:32:54,912 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost:127.0
.0.1:8032
2018-04-14 03:32:54,920 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-14 03:32:54,110 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost:127.0
.0.1:8032
2018-04-14 03:32:54,120 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2018-04-14 03:32:54,101 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MapReduceLauncher - Success
2018-04-14 03:32:54,192 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instea
d, use fs.defaultFS
2018-04-14 03:32:54,192 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not g
enerate code.
2018-04-14 03:32:54,202 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-04-14 03:32:54,202 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to pro
cess : 1
101. Amitha)
102. Shanley)
104. Amitha)
105. Pema)
110. Priyanka)
114. Madhuri)
gumail> result = DISTINCT detail;You have new mail in /var/spool/mail/acsdgild
acsdgild@localhost:~$
```

(e) List of employees (employee id and employee name) having no entry in employee_expenses file.

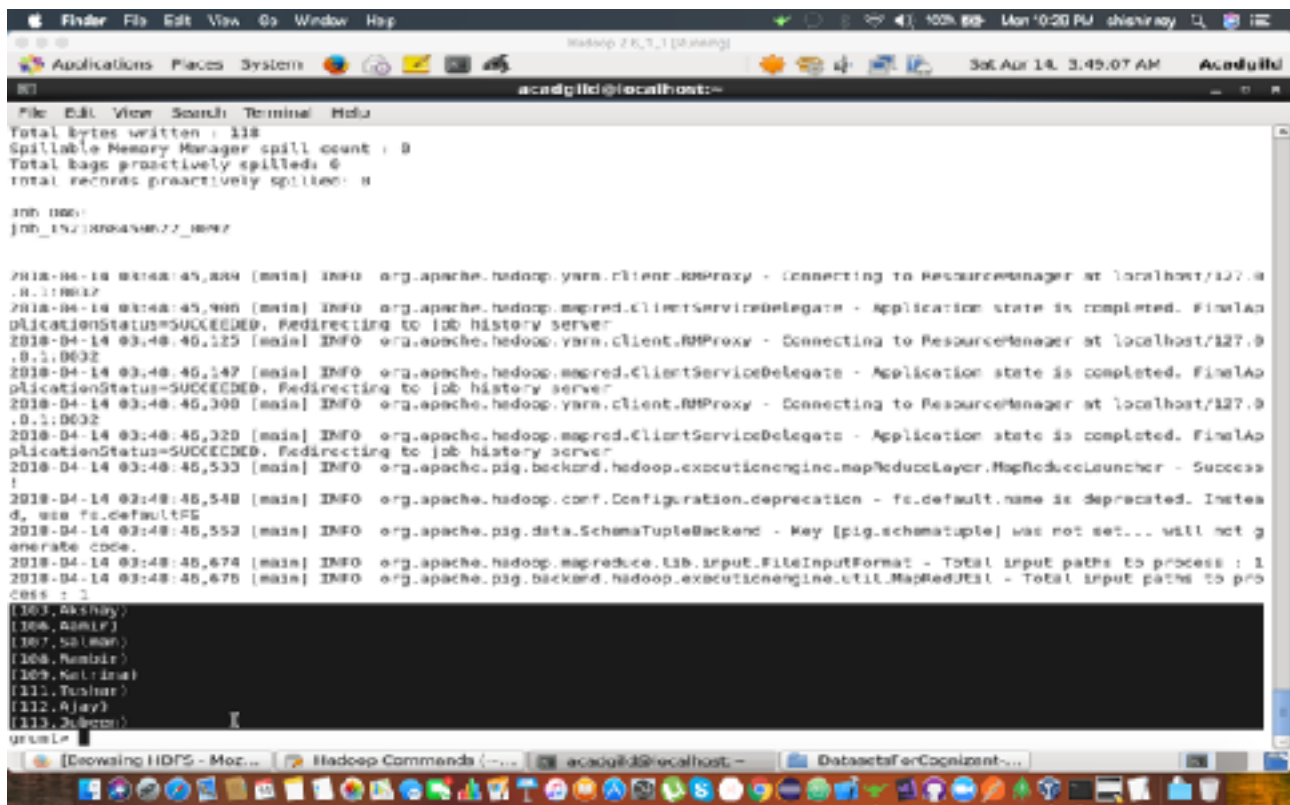
The screenshot shows a Mac desktop environment. At the top, there is a menu bar with 'Finder', 'File', 'Edit', 'View', 'Go', 'Window', and 'Help' menus. Below the menu bar is a dock containing several application icons, including Finder, Safari, Mail, and others. The main area of the screen is occupied by a terminal window titled 'Hadoop 2.6.1 [Running]'. The terminal shows the following commands and output:

```

acacgild@localhost:~$
File Edit View Search Terminal Help
gruata employeeDetail = load '/user/employeeDetails.txt' using PigStorage(',') as (empid:int, name:chararray, salary:int, rating:int);
2018-04-14 03:27:24,096 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
gruata employeeExpense = load '/user/employeeexpenses.txt' using PigStorage(',') as (empid:int, expense:int);
2018-04-14 03:28:26,047 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
gruata empWithOutExpense = join employeeDetail by empid (left outer), employeeExpense by empid;
2018-04-14 03:28:26,047 [main] INFO org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (reserved heap) of size 64 MB/256 MB to monitor. collectionUsageThreshold = 64MB/256MB, usageThreshold = 64MB/256MB
gruata empFilter = filter empWithOutExpense by employeeExpense.empid is null;
gruata empData = FORWRITE empFilter (WRITE employeeDetail::empid, employeeDetail::name);
gruata>

```

At the bottom of the screen, there is a dock with various application icons, including Finder, Safari, Mail, and others. The dock is partially visible, showing icons for Finder, Safari, Mail, and others.



Task 3

Implement the use case present in below blog link and share the complete steps along with screenshot(s) from your end.

<https://acadgild.com/blog/aviation-data-analysis-using-apache-pig/>

I am facing some issue while implementing this use case and raised with support.

Find below screenshot for error:



Once this issue is fixed then i will update this use case