

Rakshith Churchagundi Amarnath
A20424771
Assignment 5

1) Magic Number = 158096

```
hadoop@ip-172-31-92-18:~$ ls
TestDataGen.class
hadoop@ip-172-31-92-18:~$ java TestDataGen
Magic Number = 158096
hadoop@ip-172-31-92-18:~$ ls
foodplaces158096.txt foodratings158096.txt TestDataGen.class
hadoop@ip-172-31-92-18:~$ hadoop fs -copyFromLocal foodratings158096.txt /user/hadoop
hadoop@ip-172-31-92-18:~$ hadoop fs -copyFromLocal foodplaces158096.txt /user/hadoop
hadoop@ip-172-31-92-18:~$ hadoop fs -ls /user/hadoop
Found 2 items
-rw-r--r-- 1 hadoop hadoop      59 2020-02-18 04:05 /user/hadoop/foodplaces158096.txt
-rw-r--r-- 1 hadoop hadoop  17446 2020-02-18 04:04 /user/hadoop/foodratings158096.txt
hadoop@ip-172-31-92-18:~$
```

```
food_ratings = load 'foodratings158096.txt' USING PigStorage(',') as (name:chararray, f1:int, f2:int, f3:int, f4:int, placeid:int);
DESCRIBE food_ratings;
```

```
grunt> food_ratings = load 'foodratings158096.txt' USING PigStorage(',') as (name:chararray, f1:int, f2:int, f3:int, f4:int, placeid:int);
164697 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line 1, column 112> mismatched input 'f4' expecting RIGHT_PAREN
20/02/18 04:29:33 ERROR grunt.Grunt: ERROR 1200: <line 1, column 112> mismatched input 'f4' expecting RIGHT_PAREN
Details at logfile: /mnt/var/log/pig/pig_1582000009087.log
grunt> food_ratings = load 'foodratings158096.txt' USING PigStorage(',') as (name:chararray, f1:int, f2:int, f3:int, f4:int, placeid:int);
179857 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line 1, column 117> mismatched input 'placeid' expecting RIGHT_PAREN
20/02/18 04:29:48 ERROR grunt.Grunt: ERROR 1200: <line 1, column 117> mismatched input 'placeid' expecting RIGHT_PAREN
Details at logfile: /mnt/var/log/pig/pig_1582000009087.log
grunt> food_ratings = load 'foodratings158096.txt' USING PigStorage(',') as (name:chararray, f1:int, f2:int, f3:int, f4:int, placeid:int);
grunt> D
DEFINE DESC DESCRIBE DIFF DISTINCT DOUBLE DUMP
grunt> DESCRIBE food_ratings;
food_ratings: {name: chararray,f1: int,f2: int,f3: int,f4: int,placeid: int}
grunt>
```

2)

```

grunt> food_ratings_subset = FOREACH food_ratings GENERATE name, f4;
grunt> store food_ratings_subset into 'food_ratings_subset';
X = LIMIT food_ratings_subset 6;
DUMP x;

```

```

grunt> food_ratings_subset = FOREACH food_ratings GENERATE name, f4;
grunt> store food_ratings_subset into 'food_ratings_subset';

```

```

Input(s):
Successfully read 1000 records (17446 bytes) from: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/foodratings158096.txt"

Output(s):
Successfully stored 1000 records (7017 bytes) in: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/food_ratings_subset"

```

```

grunt> X = LIMIT food_ratings_subset 6;
grunt> DUMP X;

```

```

hadoop@ip-172-31-92-18:~
, usageThreshold = 489580128
20/02/18 04:53:59 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
20/02/18 04:53:59 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
20/02/18 04:53:59 INFO output.DirectFileOutputCommitter: Direct Write: DISABLED
266502 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
20/02/18 04:53:59 INFO data.SchemaTupleBackend: Key [pig.schematuple] was not set... will not generate code.
266578 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
20/02/18 04:53:59 WARN data.SchemaTupleBackend: SchemaTupleBackend has already been initialized
266582 [main] INFO org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
20/02/18 04:53:59 INFO builtin.PigStorage: Using PigTextInputFormat
20/02/18 04:53:59 INFO input.FileInputFormat: Total input files to process : 1
266590 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 04:53:59 INFO util.MapRedUtil: Total input paths to process : 1
20/02/18 04:53:59 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library
20/02/18 04:53:59 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo library [hadoop-lzo rev 5f788d5e8f90539ee331702c753fa250727128f4]
20/02/18 04:54:00 INFO output.FileOutputCommitter: Saved output of task 'attempt_0001_m_000001_1' to hdfs://ip-172-31-92-18.ec2.internal:8020/tmp-1334520956/tmp-282973547/_temporary/0/task_0001_m_000001
267279 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
20/02/18 04:54:00 WARN data.SchemaTupleBackend: SchemaTupleBackend has already been initialized
20/02/18 04:54:00 INFO input.FileInputFormat: Total input files to process : 1
267286 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 04:54:00 INFO util.MapRedUtil: Total input paths to process : 1
(Sam,10)
(Sam,3)
(Jill,50)
(Mel,17)
(Jill,33)
(Joy,33)
grunt>
grunt> clear

```

3)

```
grouped = GROUP food_ratings ALL;
```

```

food_ratings_profile = FOREACH grouped GENERATE MIN(food_ratings.f2),
MAX(food_ratings.f2), AVG(food_ratings.f2), MIN(food_ratings.f3),
MAX(food_ratings.f3), AVG(food_ratings.f3);

```

```
DUMP food_ratings_profile
```

```

hadoop@ip-172-31-92-18:~
grunts> grouped = GROUP food_ratings ALL;
grunts> food_ratings_profile = FOREACH grouped GENERATE MIN(food_ratings.f2), MAX(food_ratings.f2), AVG(food_ratings.f2), MIN(food_ratings.f3), MAX(food_ratings.f3), AVG(food_ratings.f3);
grunts> DUMP food_ratings_profile;
749817 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP_BY
20/02/18 05:02:02 INFO pigstats.ScriptState: Pig features used in the script: GROUP_BY
749855 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
20/02/18 05:02:02 INFO data.SchemaTupleBackend: Key [pig.schematuple] was not set... will not generate code.
749856 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
20/02/18 05:02:02 INFO optimizer.LogicalPlanOptimizer: {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
749923 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezLauncher - Tez staging directory is /tmp/temp1067925336 and resources directory is /tmp/temp1067925336
20/02/18 05:02:03 INFO tez.TezLauncher: Tez staging directory is /tmp/temp1067925336 and resources directory is /tmp/temp1067925336
749955 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.plan.TezCompiler - File concatenation threshold: 100 optimistic? false
20/02/18 05:02:03 INFO plan.TezCompiler: File concatenation threshold: 100 optimistic? false
749983 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to move algebraic foreach to combiner
20/02/18 05:02:03 INFO util.CombinerOptimizerUtil: Choosing to move algebraic foreach to combiner
750076 [main] INFO org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
20/02/18 05:02:03 INFO builtin.PigStorage: Using PigTextInputFormat
20/02/18 05:02:03 INFO input.FileInputFormat: Total input files to process : 1
750079 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:02:03 INFO util.MapRedUtil: Total input paths to process : 1
750082 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths (combined) to process : 1
20/02/18 05:02:03 INFO util.MapRedUtil: Total input paths (combined) to process : 1
20/02/18 05:02:03 INFO hadoop.MRInputHelpers: NumSplits: 1, SerializedSize: 408
750776 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezJobCompiler - Local resource: joda-time-2.9.4.jar
20/02/18 05:02:03 INFO tez.TezJobCompiler: Local resource: joda-time-2.9.4.jar
750776 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezJobCompiler - Local resource: pig-0.17.0-core-h2.jar

```

```

hadoop@ip-172-31-92-18:~
ApplicationId: job_1581997964630_0002
TotalLaunchedTasks: 151
FileBytesWritten: 87
HdfsBytesRead: 17446
HdfsBytesWritten: 28
SpillableMemoryManager spill count: 0
Bags proactively spilled: 0
Records proactively spilled: 0

DAG Plan:
Tez vertex scope-60 -> Tez vertex scope-61,
Tez vertex scope-61

Vertex Stats:
VertexId Parallelism TotalTasks InputRecords ReduceInputRecords OutputRecords FileBytesRead FileBytesWritten HdfsBytesRead HdfsBytesWritten Alias Fe
ature Outputs
scope-60 1 1 1000 0 1000 64 87 17446 0 food_ratings,fo
od_ratings_profile.grouped
scope-61 1 0 1 1 87 0 0 28 food_ratings_pr
ofile GROUP_BY hdfs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp237022981,

Input(s):
Successfully read 1000 records (17446 bytes) from: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/foodratings158096.txt"

Output(s):
Successfully stored 1 records (28 bytes) in: "hdfs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp237022981"

20/02/18 05:02:25 INFO input.FileInputFormat: Total input files to process : 1
772333 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:02:25 INFO util.MapRedUtil: Total input paths to process : 1
(1,50,25,253,1,50,24,958)
grunts>

```

4)

food_ratings_filtered = FILTER food_ratings BY f1 < 20 and f3 > 5;

X = LIMIT food_ratings_filtered 6;
DUMP X;

```

hadoop@ip-172-31-92-18:~
grunts> food_ratings_filtered = FILTER food_ratings BY f1 < 20 and f3 > 5;
grunts> X = LIMIT food_ratings_filtered 6;
993355 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line 14, column 22> Syntax error, unexpected symbol at or near ')'
20/02/18 05:06:06 ERROR grunt.Grunt: ERROR 1200: <line 14, column 22> Syntax error, unexpected symbol at or near ')'
Details at logfile: /mnt/var/log/pig/pig-1582001373192.log
grunts> X = LIMIT food_ratings_filtered 6;
1006784 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line 14, column 32> Syntax error, unexpected symbol at or near '6'
20/02/18 05:06:19 ERROR grunt.Grunt: ERROR 1200: <line 14, column 32> Syntax error, unexpected symbol at or near '6'
Details at logfile: /mnt/var/log/pig/pig-1582001273102.log
grunts> X = LIMIT food_ratings_filtered 6;
grunts> DUMP X;
1021008 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: FILTER,LIMIT
20/02/18 05:06:34 INFO pigstats.ScriptState: Pig features used in the script: FILTER,LIMIT
1021044 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
20/02/18 05:06:34 INFO data.SchemaTupleBackend: Key [pig.schematuple] was not set... will not generate code.
1021044 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
20/02/18 05:06:34 INFO optimizer.LogicalPlanOptimizer: {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
1021083 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezLauncher - Tez staging directory is /tmp/temp1067925336 and resources directory is /tmp/temp1067925336
20/02/18 05:06:34 INFO tez.TezLauncher: Tez staging directory is /tmp/temp1067925336 and resources directory is /tmp/temp1067925336
1021084 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.plan.TezCompiler - File concatenation threshold: 100 optimistic? false
20/02/18 05:06:34 INFO plan.TezCompiler: File concatenation threshold: 100 optimistic? false
1021105 [main] INFO org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
20/02/18 05:06:34 INFO builtin.PigStorage: Using PigTextInputFormat
20/02/18 05:06:34 INFO input.FileInputFormat: Total input files to process : 1
1021106 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:06:34 INFO util.MapRedUtil: Total input paths to process : 1

```

```

HdfsBytesWritten: 119
SpillableMemoryManager spill count: 0
Bags proactively spilled: 0
Records proactively spilled: 0

DAG Plan:
Tez vertex scope-141 -> Tez vertex scope-143,
Tez vertex scope-143

Vertex Stats:
VertexId Parallelism TotalTasks InputRecords ReduceInputRecords OutputRecords FileBytesRead FileBytesWritten HdfsBytesRead HdfsBytesWritten Alias Fe
ature Outputs
scope-141 1 1 9 0 6 0 141 17446 0 x,food_ratings,
food_ratings_filtered
scope-143 1 1 6 0 141 0 0 119 x LIMIT hd
fs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp2065546895,

Input(s):
Successfully read 9 records (17446 bytes) from: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/foodratings158096.txt"

Output(s):
Successfully stored 6 records (119 bytes) in: "hdfs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp2065546895"

20/02/18 05:06:46 INFO input.FileInputFormat: Total input files to process : 1
1033320 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:06:46 INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil: Total input paths to process : 1
(Sam,4,43,46,3,1)
(Jill,13,12,34,50,3)
(Jill,13,40,19,33,1)
(Jill,9,19,42,13,2)
(Joe,1,47,42,49,4)
(Joe,8,37,11,2,1)
grunt>

```

5.

```

food_ratings_2percent = SAMPLE food_ratings 0.02;
X = LIMIT food_ratings_2percent 10;
DUMP X;

```

```

grunt> food_ratings_2percent = SAMPLE food_ratings 0.02;
grunt> X = LIMIT food_ratings_2percent 10;
grunt> DUMP X;
1205100 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: FILTER,LIMIT
20/02/18 05:09:38 INFO pigstats.ScriptState: Pig features used in the script: FILTER,LIMIT
1205145 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
20/02/18 05:09:38 INFO data.SchemaTupleBackend: Key [pig.schematuple] was not set... will not generate code.
1205146 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED:[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
20/02/18 05:09:38 INFO optimizer.LogicalPlanOptimizer: {RULES_ENABLED:[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
1205188 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezLauncher - Tez staging directory is /tmp/temp1067925336 and resources directory is /tmp/temp1067925336
20/02/18 05:09:38 INFO tez.TezLauncher: Tez staging directory is /tmp/temp1067925336 and resources directory is /tmp/temp1067925336
1205189 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.plan.TezCompiler - File concatenation threshold: 100 optimistic? false
20/02/18 05:09:38 INFO plan.TezCompiler: File concatenation threshold: 100 optimistic? false
1205208 [main] INFO org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
20/02/18 05:09:38 INFO builtin.PigStorage: Using PigTextInputFormat
20/02/18 05:09:38 INFO input.FileInputFormat: Total input files to process : 1
1205210 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:09:38 INFO util.MapRedUtil: Total input paths to process : 1
1205212 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths (combined) to process : 1
20/02/18 05:09:38 INFO util.MapRedUtil: Total input paths (combined) to process : 1
20/02/18 05:09:38 INFO hadoop.MRInputParsers: NumSplits: 1 SerializedSize: 408
1205232 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezJobCompiler - Local resource: joda-time-2.9.4.jar
20/02/18 05:09:38 INFO tez.TezJobCompiler: Local resource: joda-time-2.9.4.jar
1205233 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezJobCompiler - Local resource: pig-0.17.0-core-h2.jar
20/02/18 05:09:38 INFO tez.TezJobCompiler: Local resource: pig-0.17.0-core-h2.jar
1205233 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezJobCompiler - Local resource: antlr-runtime-3.4.jar
20/02/18 05:09:38 INFO tez.TezJobCompiler: Local resource: antlr-runtime-3.4.jar
1205233 [main] INFO org.apache.pig.backend.hadoop.executionengine.tez.TezJobCompiler - Local resource: automaton-1.11-8.jar
20/02/18 05:09:38 INFO tez.TezJobCompiler: Local resource: automaton-1.11-8.jar

```

```

DAG Plan:
Tez vertex scope-172 -> Tez vertex scope-174,
Tez vertex scope-174

Vertex Stats:
VertexId Parallelism TotalTasks InputRecords ReduceInputRecords OutputRecords FileBytesRead FileBytesWritten HdfsBytesRead HdfsBytesWritten Alias Fe
scope-172 1 1 407 0 10 0 206 17446 0 X,food_ratings,
food_ratings_2percent
scope-174 1 1 10 0 10 206 0 0 200 X LIMIT hd
fs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp-1833187669,

Input(s):
Successfully read 407 records (17446 bytes) from: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/foodratings158096.txt"

Output(s):
Successfully stored 10 records (200 bytes) in: "hdfs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp-1833187669"

20/02/18 05:09:50 INFO input.FileInputFormat: Total input files to process : 1
1217437 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:09:50 INFO util.MapRedUtil: Total input paths to process : 1
(Wel,38,9,6,42,4)
(Sam,47,31,7,27,1)
(Sam,42,40,31,9,5)
(Joy,9,2,18,24,4)
(Wel,26,49,6,70,3)
(Ciil,27,5,44,42,2)
(Joe,34,9,47,46,3)
(Joe,23,22,4,24,5)
(Ciil,13,49,19,1,1)
(Joe,20,47,12,49,5)
grunt>

```

6)

```

food_places = load 'foodplaces158096.txt' USING PigStorage(',') as (placeid: int,
placename:chararray);

DESCRIBE food_places;

food_ratings_w_place_names = JOIN food_ratings BY placeid, food_places BY placeid;

v = LIMIT food_ratings_w_place_names 6;

```

DUMP V;

```

DAG Plan:
Tez vertex scope-603 -> Tez vertex scope-604,
Tez vertex scope-604 -> Tez vertex scope-606,
Tez vertex scope-606

Vertex Stats:
VertexId Parallelism TotalTasks InputRecords ReduceInputRecords OutputRecords FileBytesRead FileBytesWritten HdfsBytesRead HdfsBytesWritten Alias Fe
scope-602 1 1 1000 0 1000 112 9819 17446 0 food_rating_w_p
place_names,food_ratings
scope-603 1 1 5 0 5 112 200 59 0 food_places,foo
d_rating_w_place_names
scope-604 2 1 0 208 6 16148 150 0 0 V,food_rating_w
_place_names
scope-606 1 1 6 0 6 150 0 0 210 V LIMIT hd
fs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp-63855786,

Input(s):
Successfully read 1000 records (17446 bytes) from: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/foodratings158096.txt"
Successfully read 5 records (59 bytes) from: "hdfs://ip-172-31-92-18.ec2.internal:8020/user/hadoop/foodplaces158096.txt"

Output(s):
Successfully stored 6 records (210 bytes) in: "hdfs://ip-172-31-92-18.ec2.internal:8020/tmp/temp-1334520956/tmp-63855786"

20/02/18 05:36:26 INFO input.FileInputFormat: Total input files to process : 1
2813793 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
20/02/18 05:36:26 INFO util.MapRedUtil: Total input paths to process : 1
(Joe,13,23,46,42,1,1,China Bistro)
(Joy,27,17,43,48,1,1,China Bistro)
(Joy,45,46,25,24,1,1,China Bistro)
(Joe,24,30,36,19,1,1,China Bistro)
(Joy,11,19,38,48,1,1,China Bistro)
(Joy,4,20,6,20,1,1,China Bistro)
grunt>

```

7)

I. (A) LIMIT

II. (C) DISTINCT

III. (B) (f1: STRING, f2: INT, f3: INT, f4: INT)

IV. (B) relB = FOREACH relA GENERATE \$0, f3;

V. (B) data flow

VI. : (A) relB = FILTER relA by \$0 < 20