Q1. Creating a sample dataset and implementing the below Pig commands on the same dataset.

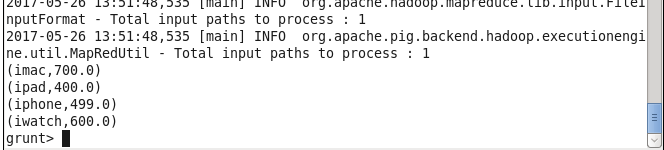
1. Min:

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = group A by $1;

C = foreach B generate group, MIN(A.$2);

dump C;

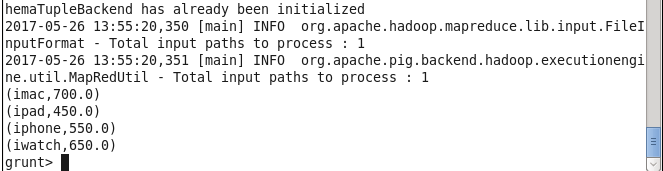


1. Max:

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = group A by $1;

C = foreach B generate group, MAX(A.$2);



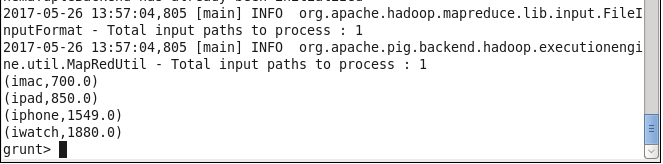
1. Sum:

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = group A by $1;

C = foreach B generate group, SUM(A.$2);

dump C;

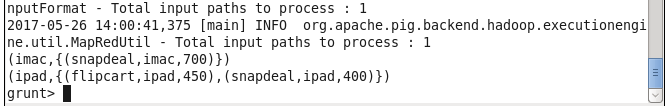


1. Limit:

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = group A by $1;

C = limit B 2;

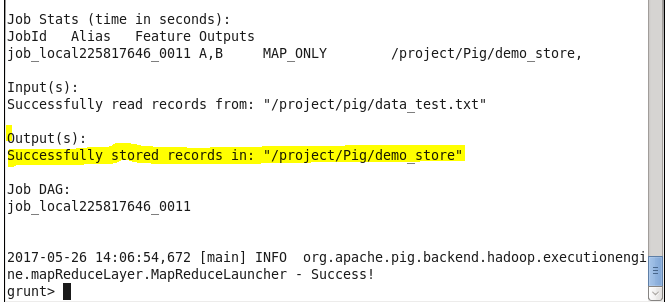


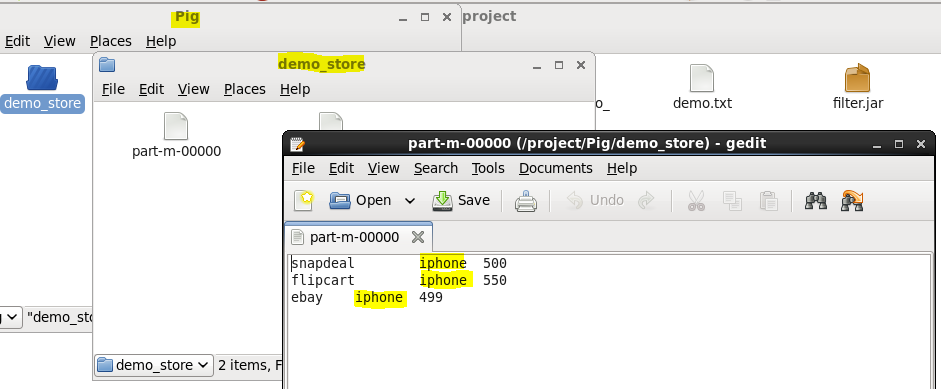
1. Store:

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = filter A by $1 == 'iphone' ;

store B INTO '/project/demo\_store';





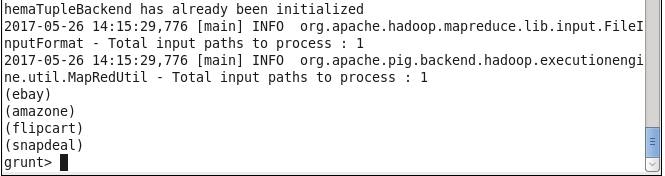
1. Distinct:

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = foreach A generate $0;

C = Distinct B;

dump C;

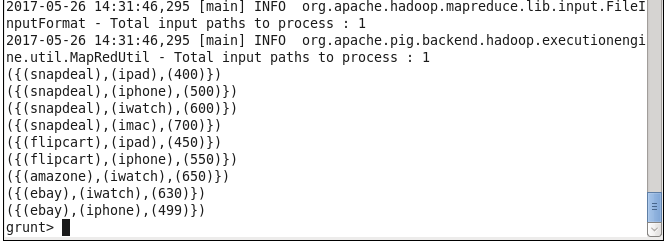


1. Tokenizer:

A = LOAD '/project/pig/data\_test.txt' as (line) ;

B = foreach A generate TOKENIZE(line) as word;

dump B;



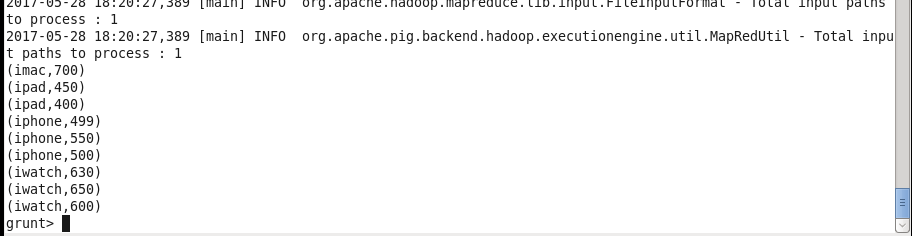
1. Flatten

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = group A by $1;

C = foreach B generate $0, flatten(A.$0);

dump C;



1. IsEmpty

A = LOAD '/project/pig/data\_test.txt' USING PigStorage(',') ;

B = group A by $0;

C = filter B by IsEmpty (A) ;

dump C;