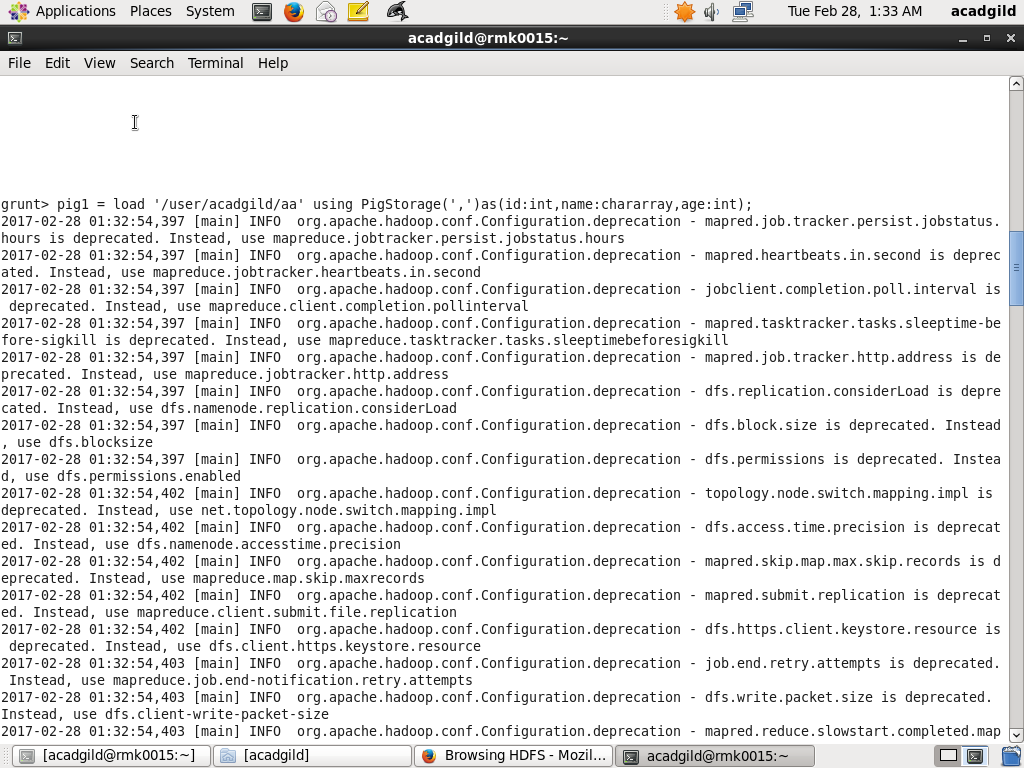
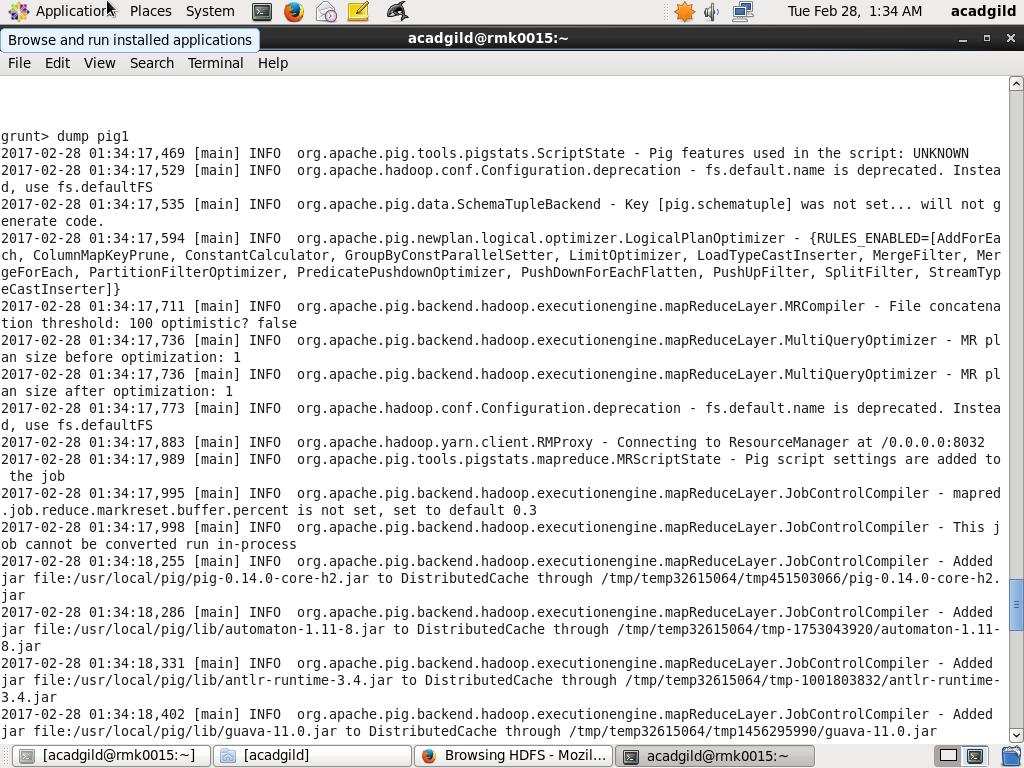
Explain the working of below pig script commands with an example:

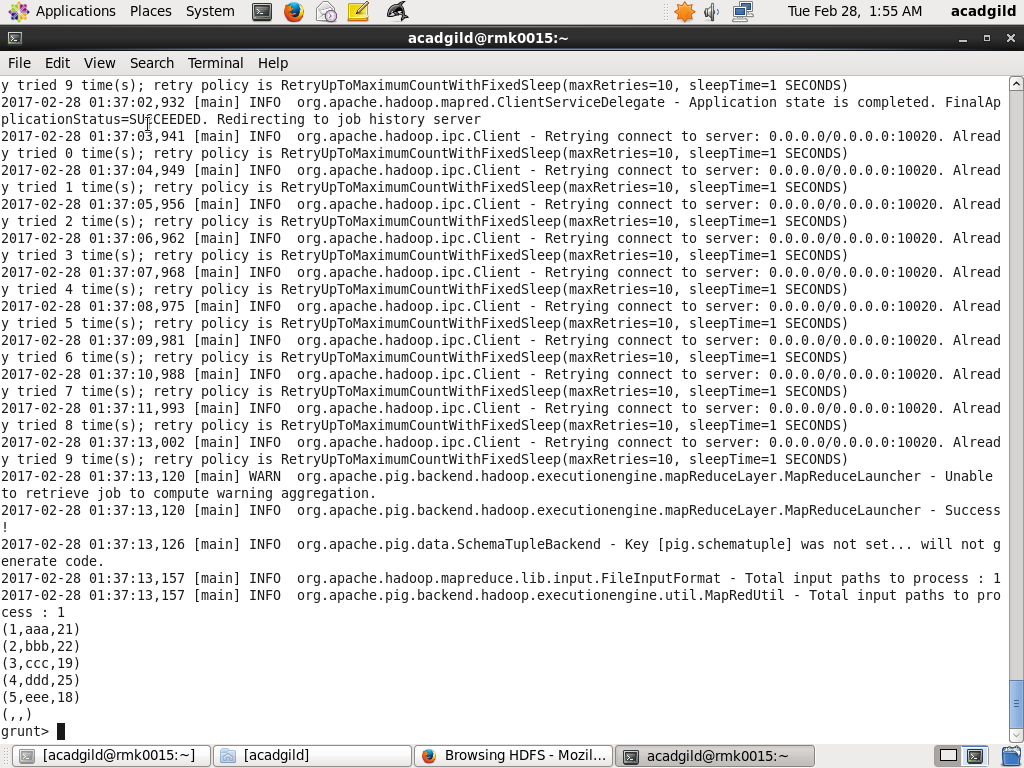
1. LOAD:

* This command is used to Load the data from the file system.
* Syntax : relation\_name = load 'file\_name' [USING function] [AS schema];
* Suppose we have a data file called aa. The fields are separated by commas.
* In this example the default load function, PigStorage, loads data from aa to form relation pig1.
* The two LOAD statements are equivalent,because no schema is specified, the fields are not named and all fields default to type bytearray.
* pig 1 = load 'a.txt';
* pig 1 = load 'aa’ using PigStorage(',') as (id:int,name:chararray,age:int);

EXAMPLE:







1. STORE:

* Stores or saves results to the file system.
* Syntax : STORE alias INTO 'directory' [USING function];

pig1 = LOAD 'aa' AS (id:int,name:chararray,age:int);

* DUMP pig1;

(1,aaa,21)

(2,bbb,19)

(3,ccc,22)

(4,ddd,25)

(5,eee,18)

STORE pig1 INTO 'myoutput' USING PigStorage ('\*');

CAT myoutput;

1\*aaa\*21

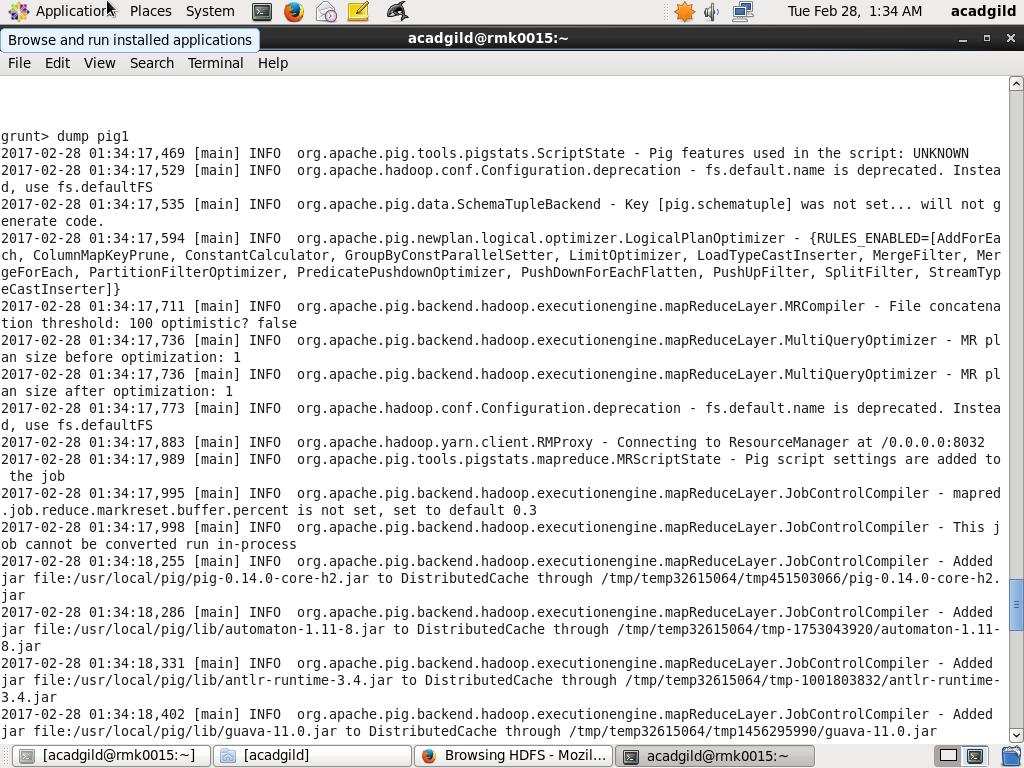
2\*bbb\*22

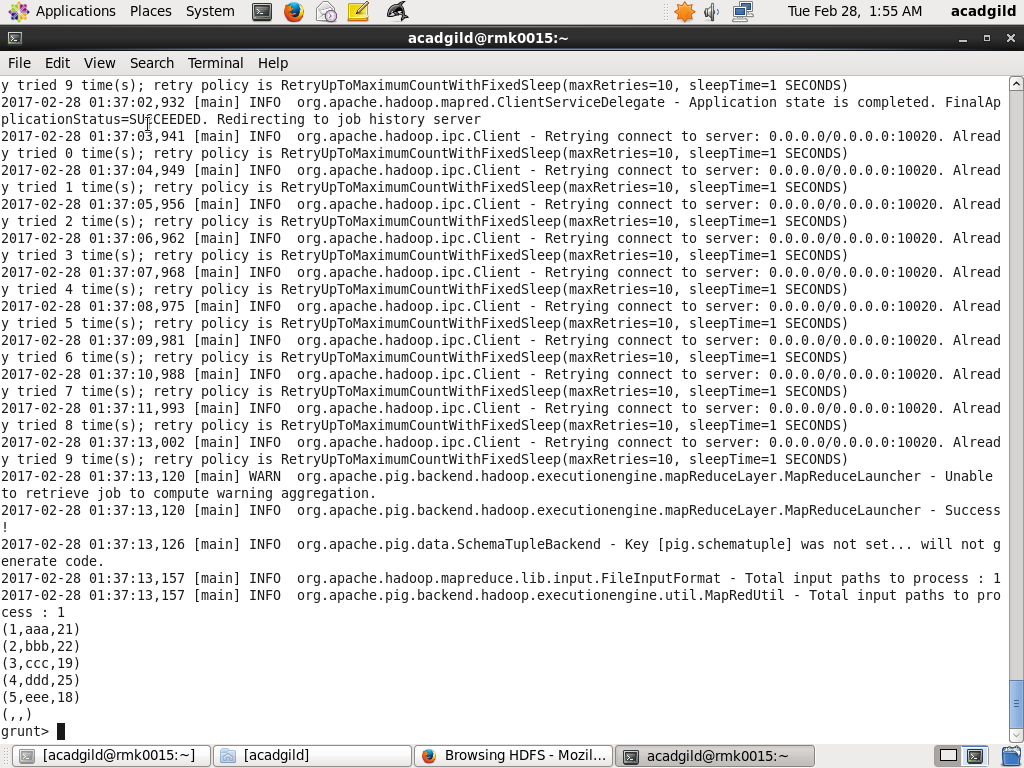
3\*ccc\*19

4\*ddd\*25

5\*eee\*18

EXAMPLE:





1. DUMP:

* Dumps or displays results to screen.
* Syntax DUMP alias;
* Here,

DUMP pig1;

(1,aaa,21)

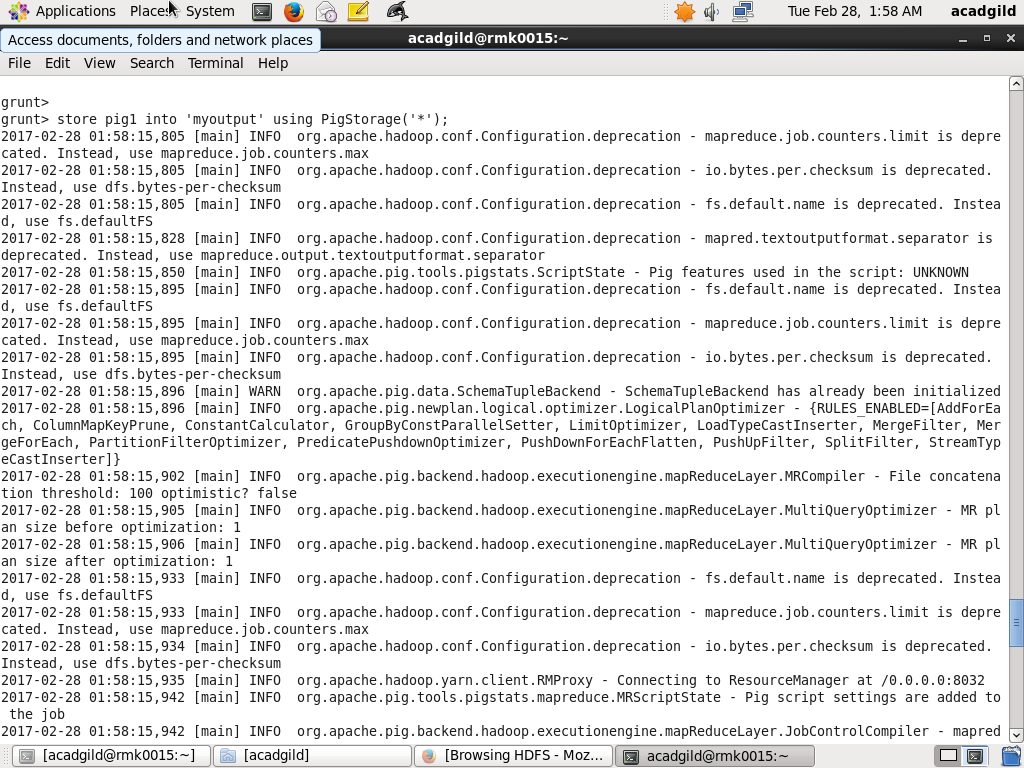
(2,bbb,22)

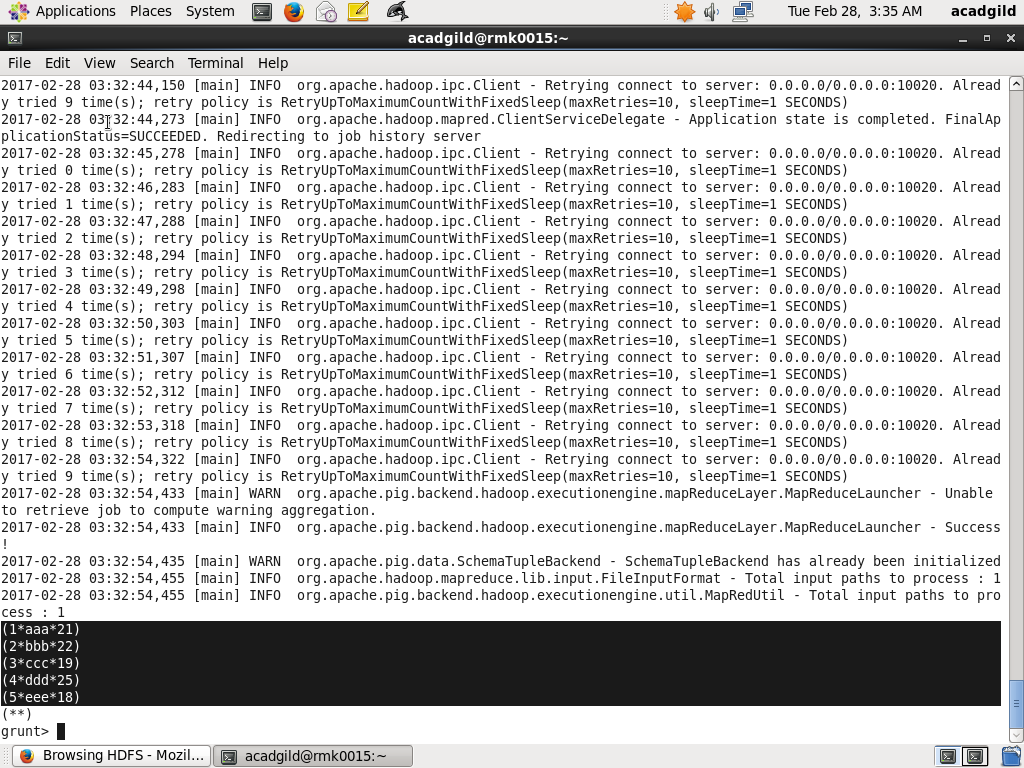
(3,ccc,19)

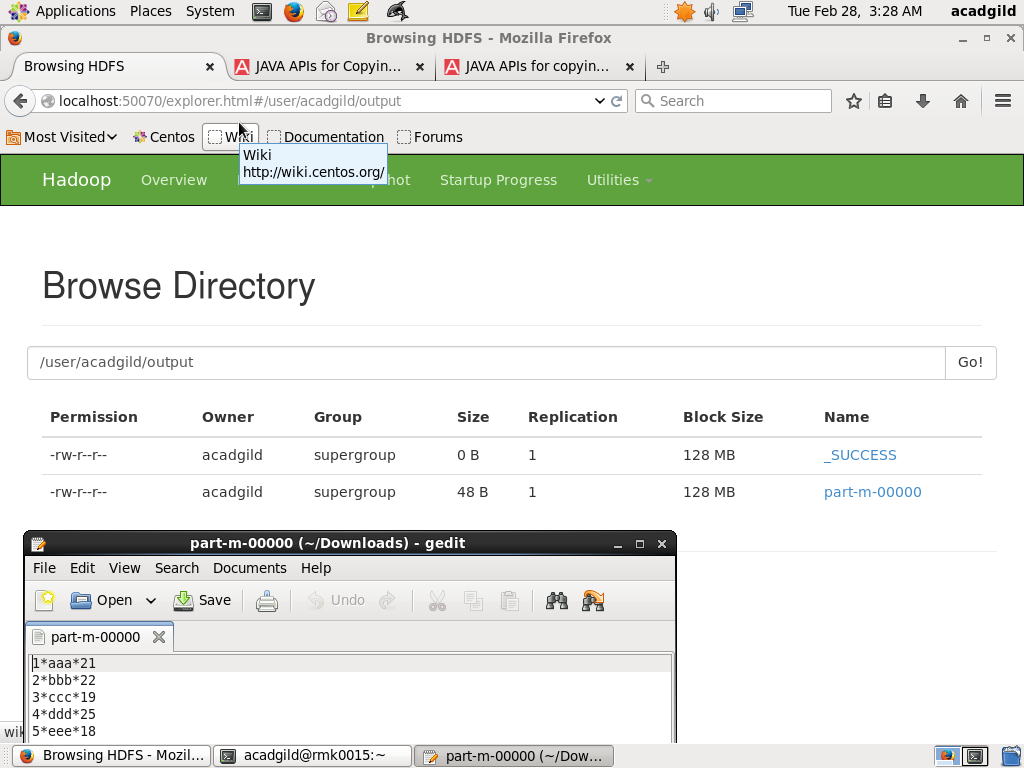
(4,ddd,25)

(5,eee,18)

EXAMPLE:





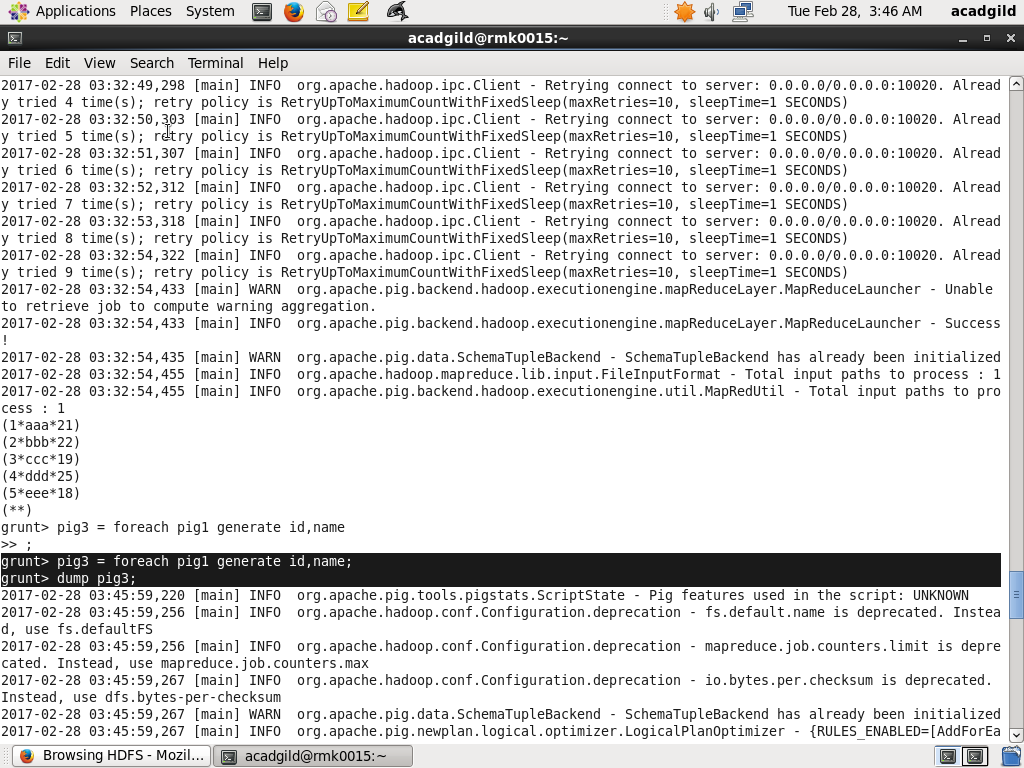


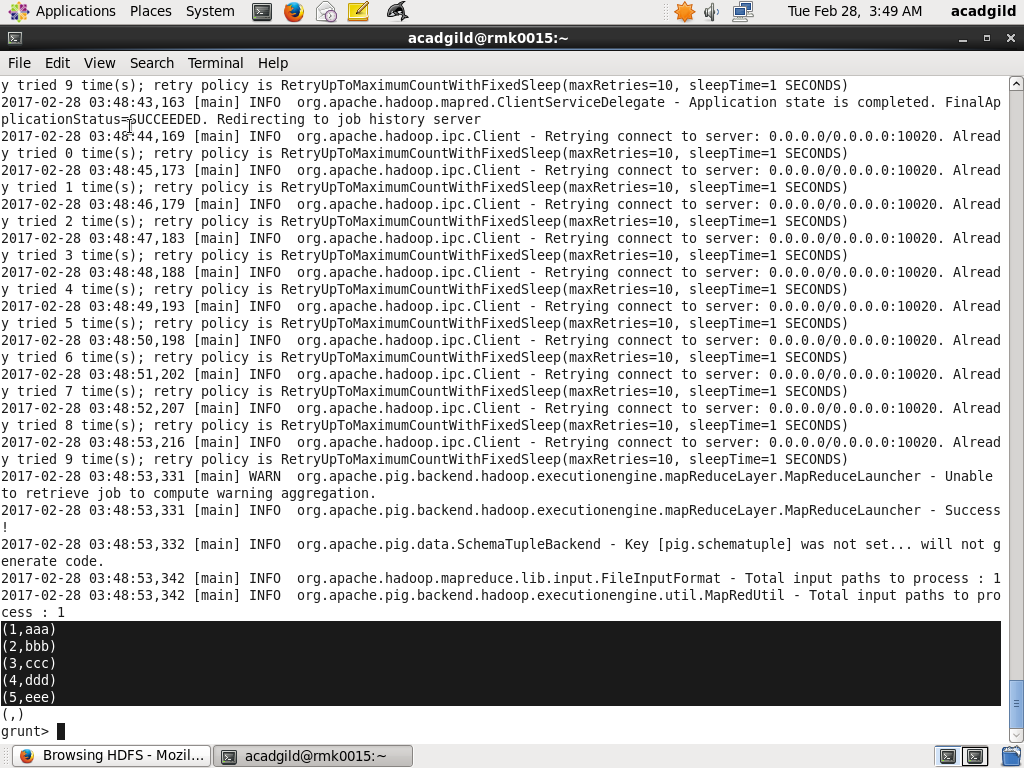
1. FOREACH:

* Generates data transformations based on columns of data i.e it can be used to display selective coloums.
* Syntax alias = FOREACH generate\_operations [AS schema];
* Here,

pig3=foreach pig1 generate id,name;

Example:



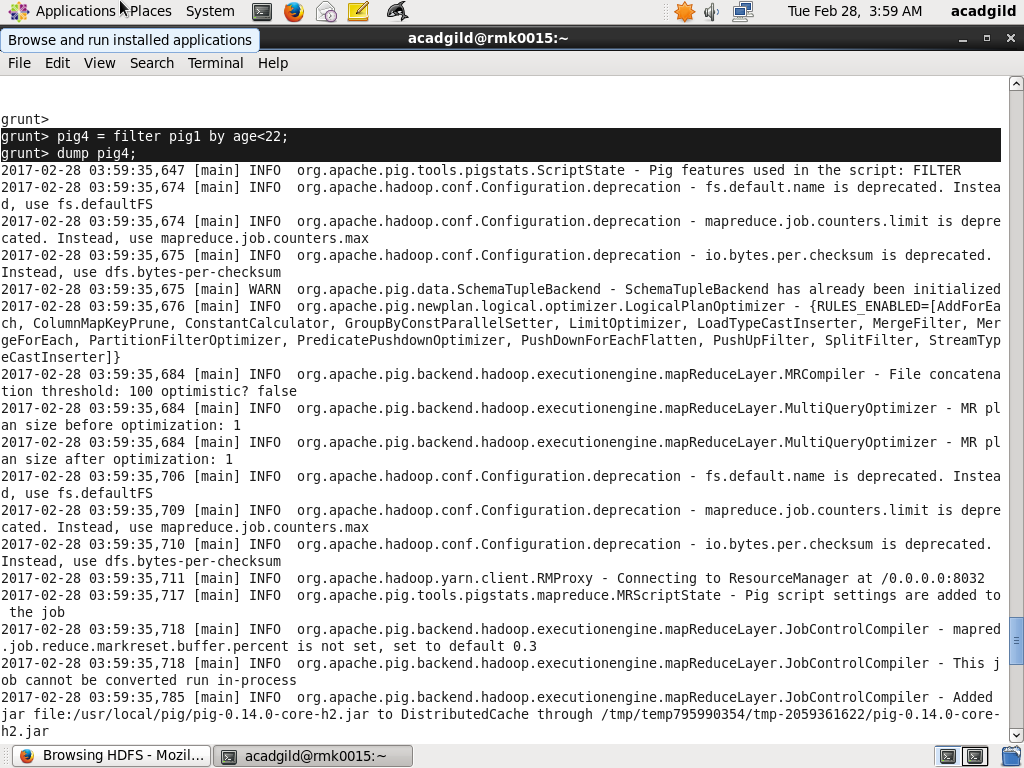


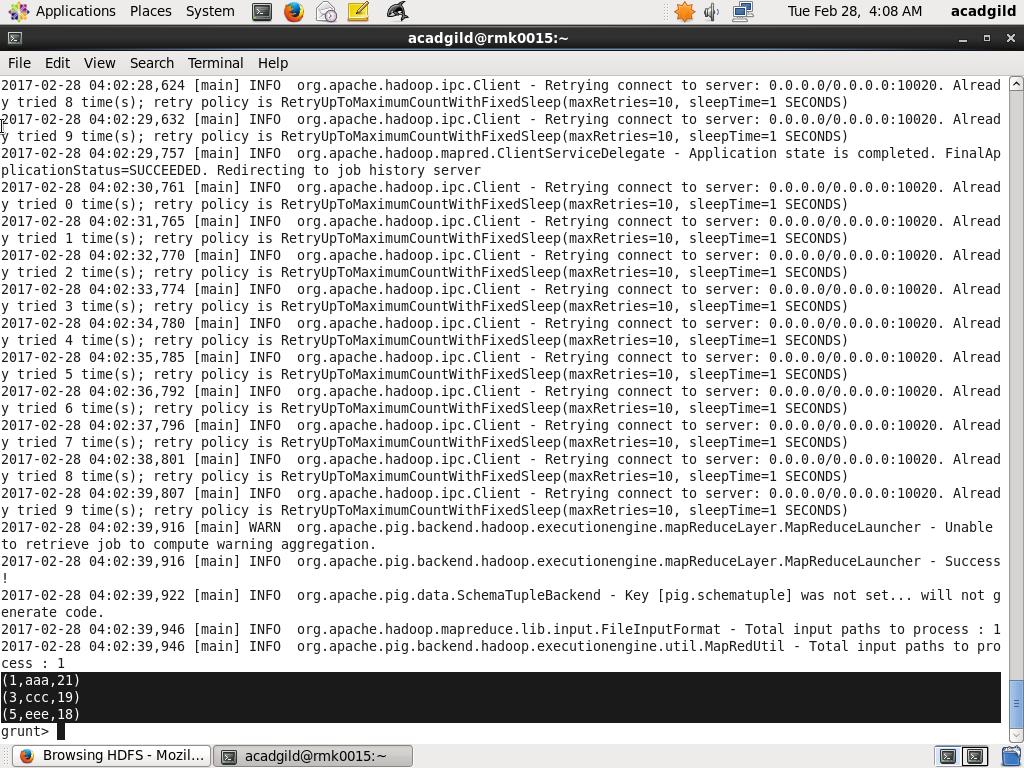
1. FILTER:

* Selects tuples from a relation based on some condition.
* Syntax alias = FILTER alias BY expression;
* Here,

Pig4=filter pig1 by age < 22;

Example:



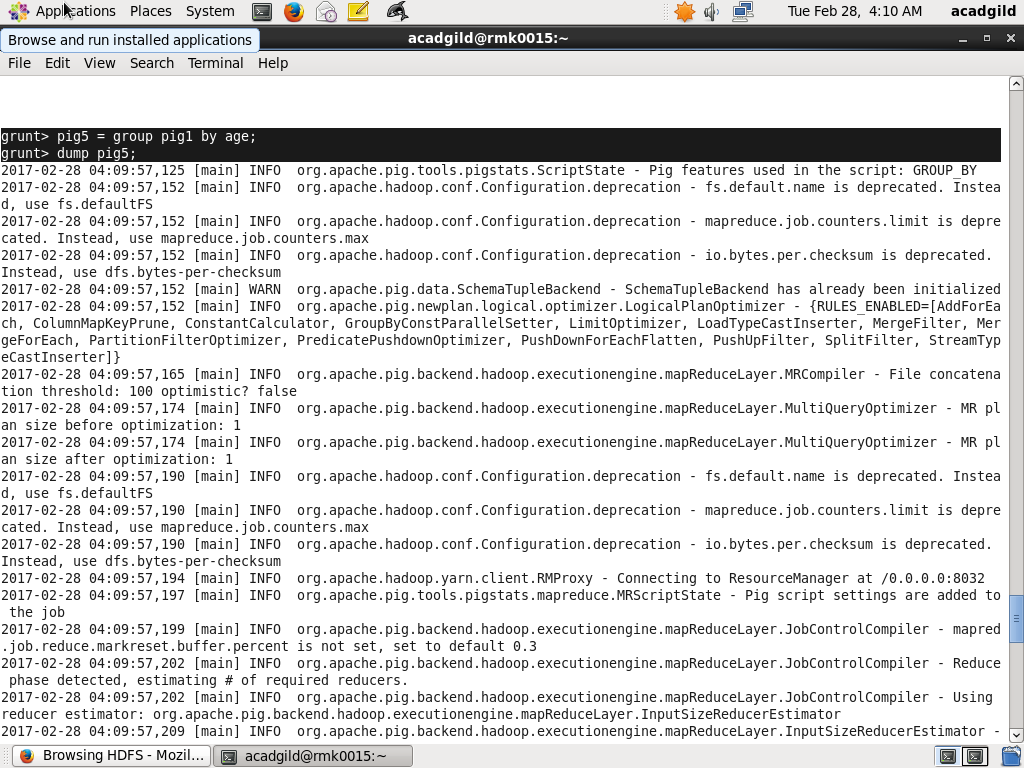


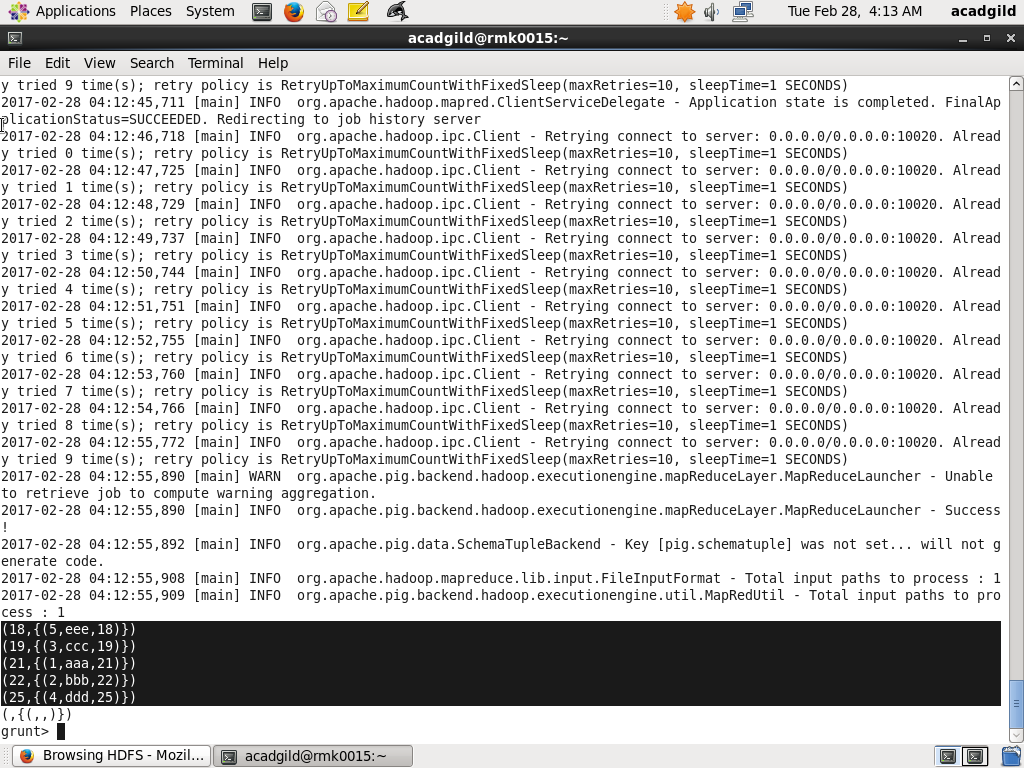
1. GROUP BY:

* Groups the data in one or multiple relations.
* Syntax alias = GROUP alias { ALL | BY expression} [, alias ALL | BY expression …] [PARALLEL n];
* The GROUP operator groups together tuples that have the same group key .
* The key field will be a tuple if the group key has more than one field, otherwise it will be the same type as that of the group key.
* The result of a GROUP operation is a relation that includes one tuple per group.
* Here,

Pig5 = group pig1 by age;

Example:



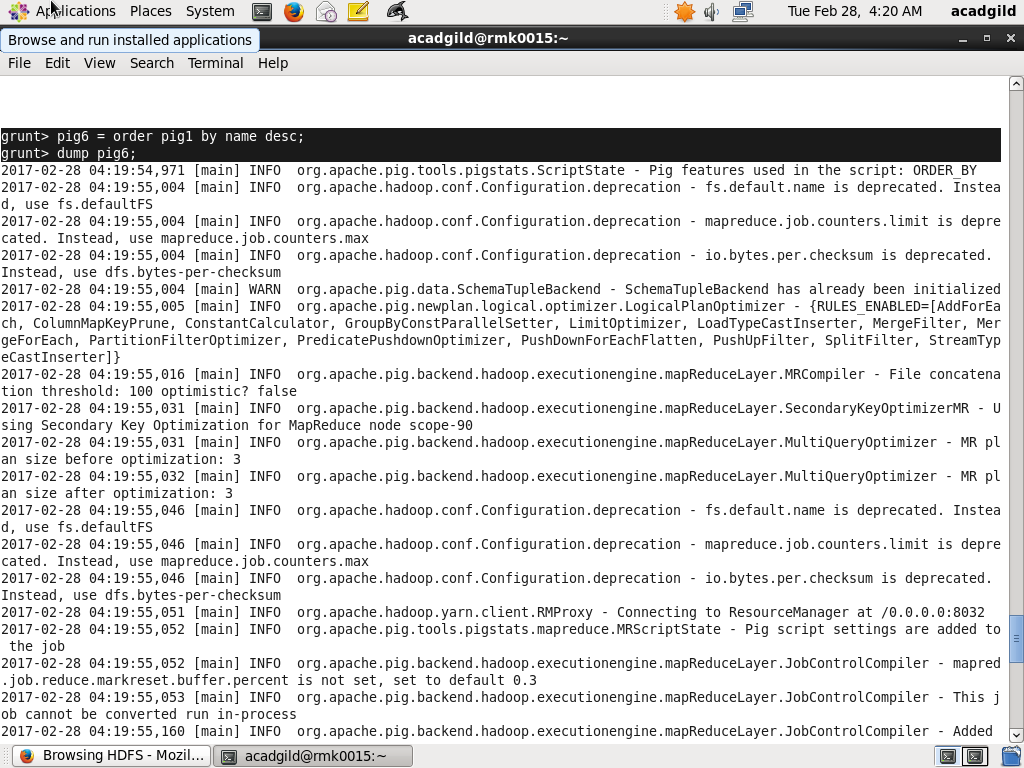


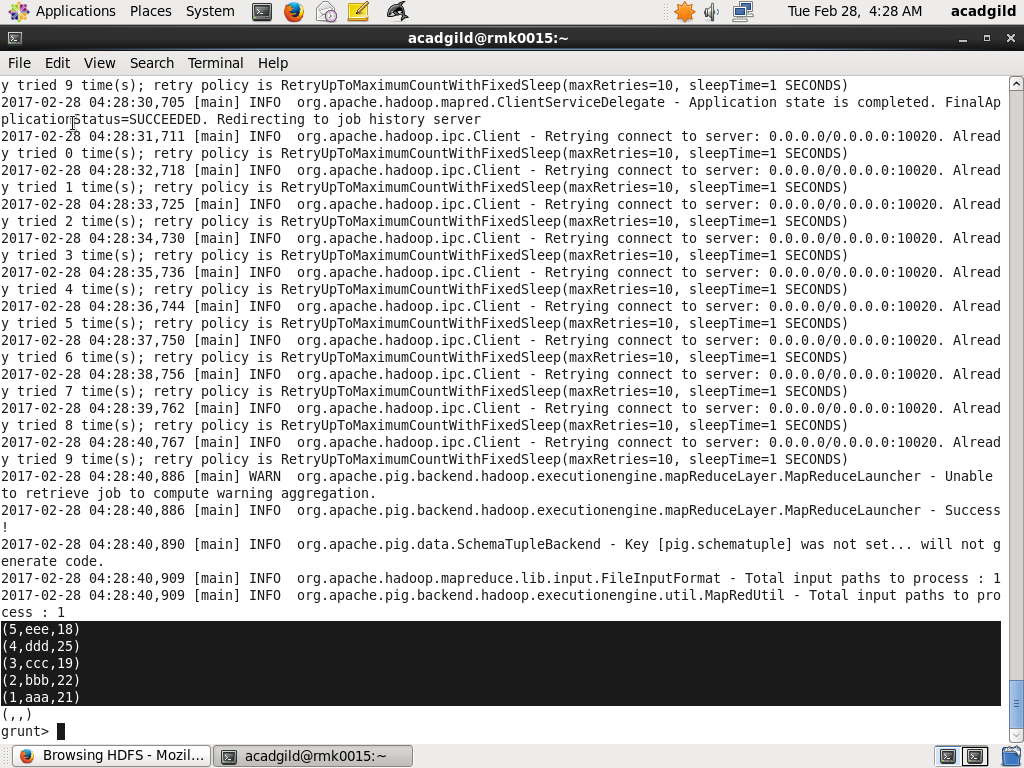
1. ORDER BY:

* Sorts a relation based on one or more fields.
* Syntax alias = ORDER alias BY { \* [ASC|DESC] | field\_alias [ASC|DESC] [, field\_alias [ASC|DESC] …] } [PARALLEL n];
* ASC - Sort in ascending order.
* DESC - Sort in descending order.
* Increase the parallelism of a job by specifying the number of reduce tasks, n. The default value for n is 1 .
* Here,

Pig6=order pig1 by name desc;

Example:





1. DESCRIBE:

* Returns the schema of an alias.
* Syntax DESCRIBE alias;
* Here,

Describe pig1;

Example:

