Create a sample dataset and implement the below Pig commands on the same dataset.

1. Concat 2) Tokenize 3) Sum 4) Min 5) Max 6) Limit 7) Store 8) Distinct 9) Flatten 10) IsEmpty
2. CONCAT:

B = foreach A generate CONCAT(f1,f3);

1. TOKENIZE:

A = load '/home/cloudera/tokenize.txt' using PigStorage(',') as (id:int,name:chararray,age:int,city:chararray);

B = foreach A generate TOKENIZE(name);

1. SUM:

A = load ‘/home/cloudera/ex.txt’ using PigStorage(‘,’) as (id:int,name:chararray,salary:int);

B = group A by name;

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

1. MIN:

A = load ‘/home/cloudera/ex.txt’ using PigStorage(‘,’) as (name:chararray,age:int,city:chararray);

B = group A by name;

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

1. MAX:

A = load ‘/home/cloudera/ex.txt’ using PigStorage(‘,’) as (name:chararray,age:int,city:chararray);

B = group A by name;

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

1. LIMIT:

Lim = limit A 2;

1. Store:

Stores or saves results to the file system.

A = load ‘/home/cloudera/ex.txt’ using PigStorage(‘,’) as (name:chararray,age:int,city:chararray);

Store A into ‘myoutput’ using PigStorage(‘\*’);

1. Distinct:

A = load ‘/home/cloudera/ex.txt’ using PigStorage(‘,’) as (name:chararray,age:int,city:chararray);

Uniq = DISTINCT A;

1. Flatten:

X = FOREACH C GENERATE group, FLATTEN(A);

1. IsEmpty:

Y = filter X by IsEmpty(SSN\_NAME);