**Load data in emp\_data:**

emp\_data = LOAD 'sample\_data.txt' USING PigStorage(',') AS (emp\_id:int, emp\_name:chararray,emp\_gender:chararray,emp\_age:int);

dump emp\_data;

**1) Concat:**

emp\_data\_concat = foreach emp\_data Generate CONCAT (emp\_name,emp\_gender);

dump emp\_data\_concat;

**2) Tokenize:**

emp\_data\_tokenize = foreach emp\_data Generate TOKENIZE(emp\_name);

dump emp\_data\_tokenize;

**3) Sum:**

emp\_group = Group emp\_data all;

emp\_age\_sum = FOREACH emp\_group GENERATE SUM(emp\_data.emp\_age);

dump emp\_age\_sum;

**4) Min:**

emp\_group = Group emp\_data all;

emp\_age\_min = FOREACH emp\_group GENERATE MIN(emp\_data.emp\_age);

dump emp\_age\_min;

**5) Max:**

emp\_group = Group emp\_data all;

emp\_age\_max = FOREACH emp\_group GENERATE MAX(emp\_data.emp\_age);

dump emp\_age\_max;

**6) Limit:**

limit\_emp\_data = LIMIT emp\_data 4;

dump limit\_emp\_data;

**7) Store:**

STORE emp\_data INTO 'stored/' using PigStorage (',');

**8) Distinct:**

distinct\_emp\_data = DISTINCT emp\_data;

dump distinct\_emp\_data;

**9) Flatten:**

lines = LOAD 'sample\_data.txt' AS (line:chararray);

words = FOREACH lines GENERATE FLATTEN(TOKENIZE(line)) as word;

dump words;

**10) IsEmpty:**

emp\_data\_group = GROUP emp\_data BY emp\_age;

emp\_data\_filter = filter emp\_data\_group by not IsEmpty(emp\_data);

dump emp\_data\_filter;