|  |
| --- |
|  |
| |  | | --- | |  | | A.Load customer records | |  | cust = LOAD '/user/hduser/custs' USING PigStorage(',') AS ( custid, firstname, lastname, age:long, profession); | |  |  | |  | describe cust; | |  |  | |  | dump cust; | |  | teacher\_bag = filter cust by profession=='Teacher'; | |  |  | |  | B. Select only 10 records | |  | amt = LIMIT cust 10; | |  | dump amt; | |  | describe amt; | |  |  | |  | C. Group customer records by profession | |  | groupbyprofession = GROUP cust BY profession; | |  | describe groupbyprofession; | |  | dump groupbyprofession; | |  |  | |  | D. Count no of customers by profession | |  | countbyprofession = FOREACH groupbyprofession GENERATE group as profession, COUNT (cust) as headcount; | |  |  | |  | describe countbyprofession; | |  | dump countbyprofession; | |  |  | |  | D 2. Sorting the output by profession | |  |  | |  | orderbyprofession = order countbyprofession by $0; | |  |  | |  | orderbycount = order countbyprofession by $1 desc; | |  |  | |  | dump orderbyprofession; | |  |  | |  | store orderbyprofession into '/home/hduser/cust\_count'; | |  |  | |  | illustrate orderbyprofession; | |  |  | |  | topprof = limit orderbycount 10; | |  |  | |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |
|  | |  |