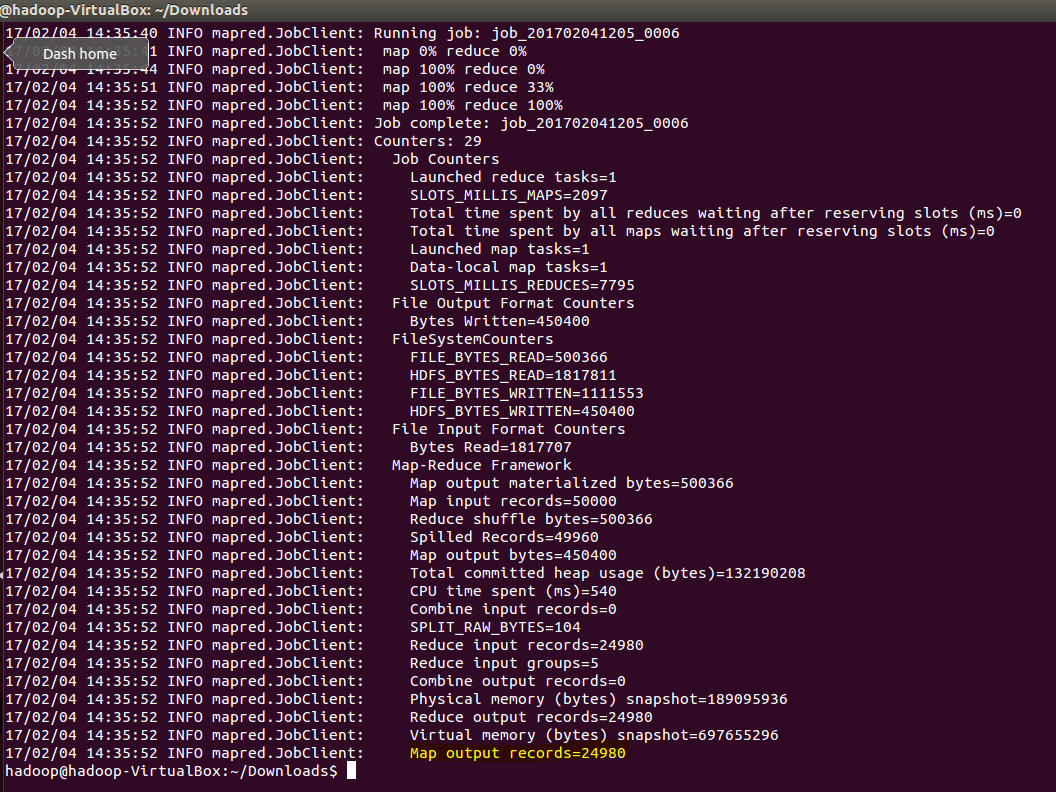
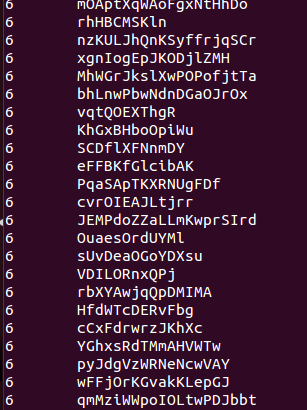
3.1 Query 1

Total 24980 record of 50,000 customer with country code between 2-6



Part of filter result, country code and customer name

3.2

4.1

/\*Write an Apache Pig query that reports the customer names that have the least number of transactions\*/



See the customer\_id and transaction number, min 51

(16523, 68) 
(21530 , 68) 
(15239,67) 
(14043,67) 
(38564, 67) 
(45607 , 67) 
(45095, 66) 
(44646 , 64) 
(47126, 63) 
(47871, 63) 
(29002 , 59) 
(24608, 51) 
2017-02-05 [main] ERROR org.apache. 
pig. tools.grunt.Grunt - 
ERROR 10 
00: Error during parsing. Encountered " <PATH> 
"D3-ORDERED DESC 
at line 2, c 
lumn 6. 
Was expecting one of: 
< EOF> 

Maxium 145

(17506,140) 
(28883 , 140) 
(37328 , 140) 
(5707 , 141) 
(2839 , 142) 
(18398, 143) 
(6346 , 145) 

 Retrieve the lowest customer name

228 [main] INFO org.apache . pig. backend. hadoop. 
2017-02-05 12:09:14, 
ne . util . MapRedUttl 
- Total input paths to process 
(ZQubROFZQmuo,51) 
grunt> 

File: /tmp/PIG_Q1.csv/part-r-OOOOO 
Goto : Vtmp/PlG Qi.csv 
Go back to dir listin 
Advanced view/download options 
ZQubRoFzQmJ',No , 51 

4.2

/\*

Write an Apache Pig query that join Customers and Transactions using Broadcast (replicated) join. The query reports for each customer the following info:

CustomerID, Name, Salary, NumOf Transactions, TotalSum, MinItem

number of transaction total number of transaction,

totalsum is sum of transtotal

minum number of item in transaction

\*/



Code using 'replicated'

(49988 , vejHDNwmtweTEQYDcn , 9671.999 , 77 , 42907.84966182709 , 1) 
(49989 , sww0PhcqguhzxjE1 , 5999.9507 , 95 , 48367.03768348694, 1) 
(49990, 
axNcrqLRdqzjGFeHzw , 6904.114 , 99 , 50608.32879924774, 1) 
(49991 , DygyjQFXYVEALLdOKWZ , 4405.227 , 109 , 52607.33077812195 , 1) 
(49992 , AEWBhrAYqpHKZeE , 5267.631 , 106 , 55242.20034599304, 1) 
(49993, 
kxnwzmhsmwttB , 1989.1237 , 114 , 53473.36391830444 , 1) 
(49994 , 
eabMIYSHsKubXYgzvv , 5709.443 , 117 , 59660.40106201172, 1) 
(49995 , otwRJtw1nAmZKMf1 , 6963.8047 , 115 , 63341.01471328735 , 1) 
(49996, 
ABcgcnwNxgLVADPNL , 5244.646 , 87 , 39914.57150268555 , 1) 
(49997, 
tTFe1CKzUFY0j0NYh , 284.10858 , 100 , 50321.8217048645 , 1) 
(49998, 
cVkDzzfSQSBNtnck , 1424.8522 , 90 , 40804.66194152832 , 1) 
(49999, 
dbuqysJoudmR , 9311.197 , 100 , 50218.591426849365 , 1) 
( 50000 , twtcoxZk1YBQheyp , 2496.2346 , 101 ,47782.20666885376 , 1) 
grunt> 

 Show CustomerID, Name, Salary, NumOf Transactions, TotalSum, MinItem

File: /tmp/PIG_Q2.csv/part-r-OOOOO 
. Vtmp/PlG Q2.csv 
Goto • 
Go back to dir listin 
Advanced view/download options 
View Next chunk 
1, dczVfNzxyoHi , 9383 589, 92, 43034 40139389038, 
2, RbbyNNVcxptfzG, 3561 0757 , 79, 37820 205266952515, 
3, UGFtNmkjNnCeTzsHoJSO, 2262 9424, 112, 55001 48632621765, 
4, NydJQN0vQAyqcy , 9176 407 , 112, 59234 990156173706, 
6, Fhq1QccUBcMC, 4540 4175, 108, 54330 246509552, 
7, ksn80b0VNnKyHDJ , 2198 9424, 95, 46040 71289539337 , 1 
, Rwh8SaFPoMbGNYKwMC, 8678 553, 111 , 54854 305086135864, 
g, ivPcMmHGj ttE, 5722 621 , 107 , 52010 08278656006, 
10, PQtdwgf8LkKEG, 5578 2817, 85, 40245 088148117065, 1 
11 , X8JvxTfDQidezj , 319 8501 , 95, 43882 56953620911 , 1 
12, iySa1YKZrFLcNcRO, 9138 897 , 93, 46818 79751968384, 
13, XEJRSFtqXtfM, 5174 7583, 104, 54540 527338027954, 
14, zpWoGcxhcNNtUOyk , 4713 1416, 110, 55666 08066558838, 
15, MzYJtMvcCHednnv , 5025 0244, 103, 55025 11685466766, 
16, j , 742 9388, 107 , 53380 211141586304, 
17, aznxrLyxbJN1pxFR-x, 227 79965, 95, 48307 497371673584, 
18, amfbxTYuUf1gMoxxqk , 7853 1714, 105, 52289 17593193054, 
19, ebAUcbHm rDpvvKk , 8022 gog, 101 , 48325 555203437805, 
20, fiTgRXLj1G, 6525 8857 , 108, 60290 024965286255, 
21 , mKxvwgPMybagEv1BeF, 6762 5645, 100, 49793 27274131775, 
22, XjRQBK1nNYbD, 6832 72, 98, 52933 705921173096, 1 
23, , 7743 1235, 95, 50793 06426048279, 
24, oseRhgopbaTDKEZ, 6055 5767 , 85, 44320 136865615845, 
25, LpNSAkaPzMYTG, 9987 636, 100, 48944 0419960022, 
26, KaLDtSSssaZQmkLkos , 2374 6895, 95, 47549 40232658386, 

4.3

/\*

Write an Apache Pig query that reports the Country Codes having number of customers greater than 5000 or less than 2000

\*/



 Country code and customer number

adoop-VirtualBox: — 
30b DAG: 
2017-02-05 14:20:29 , 405 [main] INFO org.apache.pig.backend.hadoop . executionengt 
ne .mapReduceLayer .MapReduceLauncher - 
Success ! 
2017-02-05 14:20:29 , 406 [main] INFO org.apache.ptg.data.SchemaTupleBackend - 
y [pig. schematuple] was not set... 
will not generate code. 
2017-02-05 14:20:29 , 409 [main] INFO org.apache.hadoop.mapreduce.lib.input.Filel 
nputFormat - Total input paths to process 
1 
2017-02-05 14:20: 29,409 [main] INFO org.apache.pig.backend.hadoop.executtonengt 
ne.uttl.MapRedUttl - Total input paths to process 
(1, 5032) 
(2,4977) 
(3,4969) 
(4, 5039) 
(5, 5092) 
(6,4903) 
(7,4975) 
(8 , 5024) 
(9,4985) 
(10, 5004) 
grunt> 

Dump show

"e .mapReduceLayer .MapReduceLauncher - 
Success ! 
'017-02-05 14:22:19, 320 [main] INFO org.apache.ptg.data.SchemaTupleBackend 
[pig. schematuple] was not set. 
will not generate code. 
'017-02-05 14:22:19, 326 [main] INFO org.apache.hadoop.mapreduce.lib.input.F 
IputFormat - Total input paths to process 
.1 
'017-02-05 14:22:19, 326 [main] INFO org.apache.pig.backend.hadoop.executton 
le.uttl.MapRedUttl - Total input paths to process 
a, 5032) 
g, 5039) 
, 5024) 
, 5004) 
Jrunt> 

Filter >5000 or < 2000

File: /tmp/PIG_Q3.csv/part-r-OOOOO 
Goto : 
/tmp/PlG Q3.csv 
go 
Go back to dir listin 
Advanced view/download options 
1, 5032 
4, 5039 
s, 5092 
8, 5024 
10, 5004 