**1.Write a mapreduce/pig program to calculate the number of cases investigated**

**under each FBI code**

whole\_data =  load '/home/acadgild/pig/crimes\_2001.csv' using PigStorage(',');  
  
A =  foreach whole\_data generate $1 as case\_no, $14 as fbi\_code;  
C =  group A by fbi\_code;  
D = foreach C generate group,COUNT(A.case\_no);  
E = FILTER D BY $0 is not null;  
store  E into 'solution1' using PigStorage(',');  
  
OUTPUT:  
  
(1,172)  
(2,362)  
(3,266)  
(4,154)  
(5,197)  
(6,198)  
(7,138)  
(8,301)  
(9,192)  
(02,1480)  
(03,10552)  
(05,14735)  
(06,62826)  
(07,10520)  
(09,437)  
(10,1708)  
(11,13637)  
(12,79)  
(13,151)  
(14,31244)  
(15,3780)  
(16,1949)  
(17,1165)  
(18,24989)  
(19,590)  
(20,1435)  
(21,293)  
(22,483)  
(23,77)  
(24,4114)  
(25,142)  
(26,29009)  
(27,175)  
(28,385)  
(29,196)  
(30,115)  
(31,93)  
(32,76)  
(33,105)  
(34,184)  
(35,56)  
(36,63)  
(37,161)  
(38,117)  
(39,98)  
(40,97)  
(41,123)  
(42,87)  
(43,101)  
(44,35)  
(45,34)  
(46,62)  
(47,137)  
(48,61)  
(49,61)  
(50,40)  
(56,15)  
(57,1)  
(58,3)  
(61,5)  
(66,7)  
(68,2)  
(76,51)  
(01A,533)  
(01B,6)  
(04A,4912)  
(04B,7598)  
(08A,13161)  
(08B,44935)  
(1923,1)

**2.Write a mapreduce/pig program to calculate the number of cases investigated**

**under FBI code 32.**

whole\_data =  load '/home/acadgild/pig/crimes\_2001.csv' using PigStorage(',');  
  
A = filter whole\_data by $14==32;  
B = group A all;  
C =foreach B generate COUNT($1);  
DUMP C;  
  
OUTPUT:  
(76)

**3.Write a mapreduce/pig program to calculate the number of arrests in theft district**

**wise.**

A = filter whole\_data by ($5=='THEFT') AND ($8==TRUE);  
B = foreach A generate $5 as theft,$8 as arrest,$11 as district;   
C = group B by district;  
D =foreach C generate group,COUNT(B.arrest);  
  
  
SAMPLE OUTPUT:  
  
(001,1119)  
(002,220)  
(003,157)  
(004,221)  
(005,273)  
(006,649)  
(007,172)  
(008,458)  
(009,318)  
(010,166)  
(011,174)  
(012,353)  
(014,227)  
(015,111)  
(016,171)  
(017,227)  
(018,732)  
(019,499)  
(020,241)  
(022,207)  
(024,224)  
(025,591)

**4.Write a mapreduce/pig program to calculate the number of arrests done**

**between october 2014 and october 2015.**

A = filter whole\_data by ($8==TRUE);  
B =foreach A generate $0 as id,$2 as time;  
  
C = foreach B generate id,(int)SUBSTRING(time,0,2) as month, (int)SUBSTRING(time,6,10) as year ;  
  
E= filter C by ($1 >= 09 AND $2 == 2014) OR ($1 <= 09 AND $2 == 2015);  
  
  
final =foreach (group E by $2)  generate group,COUNT(E.id) as count;  
  
G = group final all;  
  
final\_result = foreach  G generate SUM(final.count);  
  
OUTPUT:  
(69652)