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

**crime = load '/home/cloudera/Crimes\_-\_2001\_to\_present.csv' using PigStorage(',') as (ID:int,CaseNumber:chararray,Date:datetime,Block:chararray,IUCR:int,PrimaryType:chararray,Description:chararray,LocationDescription:chararray,Arrest:chararray,Domestic:chararray,Beat:int,District:int,Ward:int,CommunityArea:int,FBICode:chararray,XCoordinate:int,YCoordinate:int,Year:int,UpdatedOn:int,Latitude:float,Longitude:float,Location:map[]);**

**grouped = group crime by FBICode;**

**totalfbi = foreach grouped generate group, COUNT(crime.FBICode);**

**dump totalfbi;**

**(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)**

**(,0)**

1. **Write a mapreduce/pig program to calculate the number of cases investigated under FBI code 32.**

**crime = load '/home/cloudera/Crimes\_-\_2001\_to\_present.csv' using PigStorage(',') as (ID:int,CaseNumber:long,Date:datetime,Block:chararray,IUCR:int,PrimaryType:chararray,Description:chararray,LocationDescription:chararray,Arrest:chararray,Domestic:chararray,Beat:int,District:int,Ward:int,CommunityArea:int,FBICode:int,XCoordinate:int,YCoordinate:int,Year:int,UpdatedOn:int,Latitude:float,Longitude:float,Location:map[]);**

**filtered = filter crime by FBICode==32;**

**grouped = group filtered by FBICode;**

**totalfbi32 = foreach grouped generate group, COUNT(filtered.FBICode);**

**dump totalfbi32;**

**(32,76)**

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

**crime = load '/home/cloudera/Crimes\_-\_2001\_to\_present.csv' using PigStorage(',') as (ID:int,CaseNumber:long,Date:datetime,Block:chararray,IUCR:int,PrimaryType:chararray,Description:chararray,LocationDescription:chararray,Arrest:chararray,Domestic:chararray,Beat:int,District:int,Ward:int,CommunityArea:int,FBICode:int,XCoordinate:int,YCoordinate:int,Year:int,UpdatedOn:int,Latitude:float,Longitude:float,Location:map[]);**

**crimedist = DISTINCT crime;**

**grouped = group crimedist by District;**

**final = foreach grouped generate group, COUNT(crimedist.District);**

**dump final;**

**(1,12448)**

**(2,11308)**

**(3,14288)**

**(4,17531)**

**(5,12539)**

**(6,17277)**

**(7,16706)**

**(8,18628)**

**(9,13813)**

**(10,12686)**

**(11,20798)**

**(12,13006)**

**(14,9549)**

**(15,13141)**

**(16,9892)**

**(17,7940)**

**(18,11969)**

**(19,12526)**

**(20,4529)**

**(22,9267)**

**(24,7589)**

**(25,16395)**

**(31,1)**

**(111,3)**

**(121,3)**

**(122,4)**

**(131,1)**

**(133,2)**

**(314,1)**

**(331,1)**

**(332,7)**

**(412,1)**

**(413,8)**

**(622,2)**

**(712,2)**

**(813,3)**

**(814,12)**

**(815,1)**

**(821,1)**

**(823,1)**

**(825,1)**

**(832,5)**

**(911,1)**

**(922,1)**

**(924,3)**

**(925,2)**

**(1034,1)**

**(1113,1)**

**(1223,3)**

**(1225,1)**

**(1235,2)**

**(1413,1)**

**(1432,1)**

**(1531,1)**

**(1612,1)**

**(1624,1)**

**(1651,5)**

**(1654,46)**

**(1722,5)**

**(1731,1)**

**(1821,1)**

**(1822,2)**

**(1834,3)**

**(1933,1)**

**(2431,1)**

**(2521,1)**

**(2524,3)**

**(2533,7)**

**(2534,2)**

**(,0)**