**USA CRIME ANALYSIS USE CASE 1**

Download the dataset from below URL:

<https://drive.google.com/file/d/0B1QaXx7tpw3SaUJHOHBZclBXWG8/view?usp=sharing>

This dataset contains attributes related to crimes taking place in various areas like type of

crime, FBI code related to that criminal case, arrest frequency, location of crime etc.

**Dataset Description:**

ID,Case Number,Date,Block,IUCR,Primary Type,Description,Location

Description,Arrest,Domestic,Beat,District,Ward,Community Area,FBICode,X Coordinate,Y

Coordinate,Year,Updated On,Latitude,Longitude,Location

1. Write a MapReduce/Pig program to calculate the number of cases investigated under each FBI code.

**REGISTER '/home/acadgild/pig/ProjectFiles/piggybank.jar';**

**US\_crime\_data = load '/home/acadgild/pig/ProjectFiles/Crimes\_-\_2001\_to\_present.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',','NO\_MULTILINE','UNIX','SKIP\_INPUT\_HEADER');**

**FBICODE\_CASE\_DATA = FOREACH US\_crime\_data GENERATE $14 as FBICODE,$1 as CASENUMBER;**

**GROUPBY\_FBICODE = group FBICODE\_CASE\_DATA by FBICODE;**

**Result = FOREACH GROUPBY\_FBICODE GENERATE group,COUNT(FBICODE\_CASE\_DATA.CASENUMBER);**

**dump Result;**



