**Project 1.1 - USA Crime Analysis**

**Project Objectives**

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

FBI code

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

code 32.

3. Write a MapReduce/Pig program to calculate the number of arrests in theft district wise.

4. Write a MapReduce/Pig program to calculate the number of arrests done between October

2014 and October 2015.

Solution

* Given data file is loaded is loaded in to hdfs using the put command

hdfs dfs -put Crimes\_Project1.csv /crime\_data

* MapReduce code is used to accomplish the objectives given
* Maven is used to build jar
* Eclipse IDE is used as Editor
* To run the code jar is exported from the eclipse and it executed in the command line by passing the input file path the output path:

Example: yarn jar Arrests\_dist\_Count.jar /crime\_csv outputArrestdis

* Output is view in the hdfs shell using the cat command :

Example: hdfs dfs -cat /outputArrest/part-r-00000