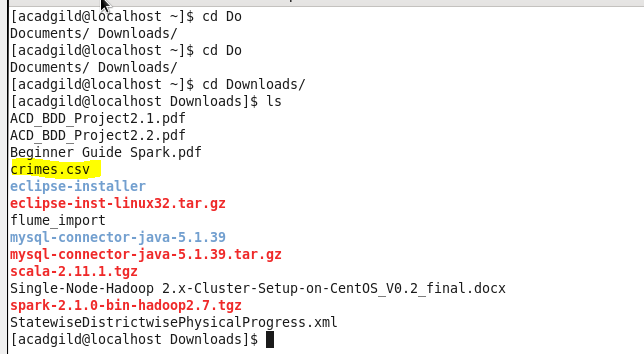
# **USA Crime Analysis**

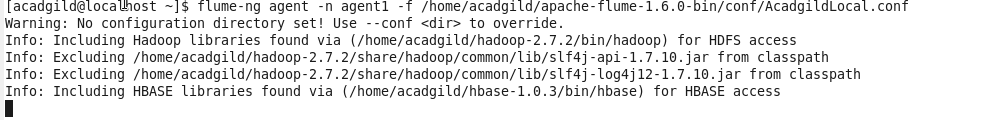
**This dataset contains attributes related to crimes taking place in various areas like type of crime, FBI code related to that criminal case, arrest frequency, location of crime etc.**

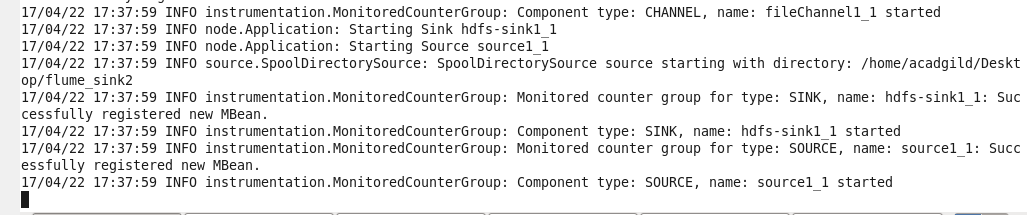
**Using Flume to load the data in HDFS**

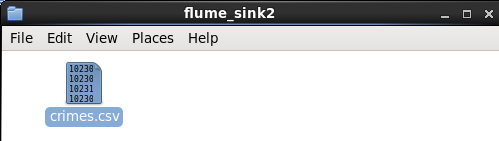
**Data file in below path**



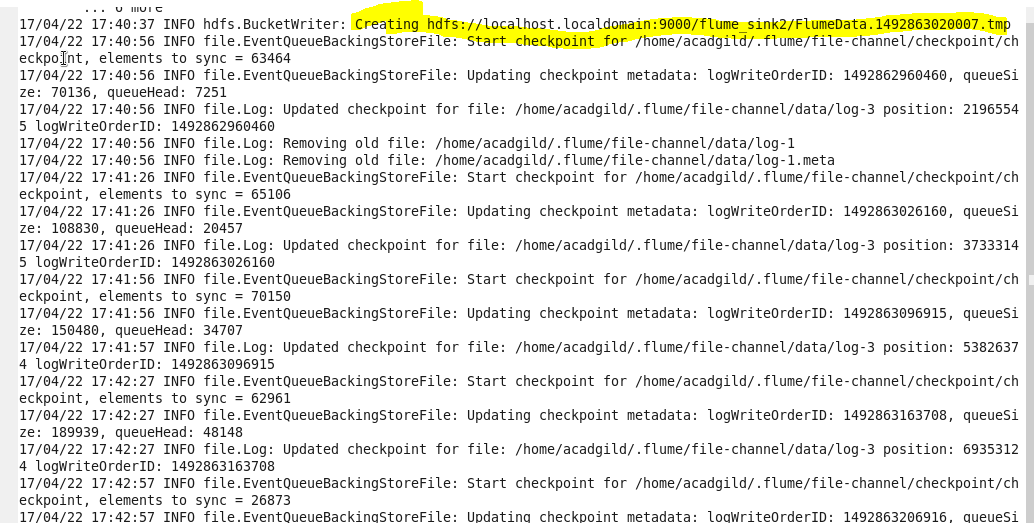
**Run the flume agent**

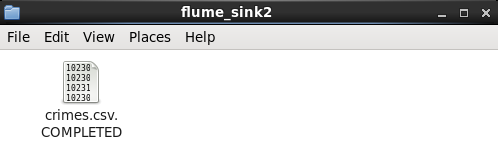


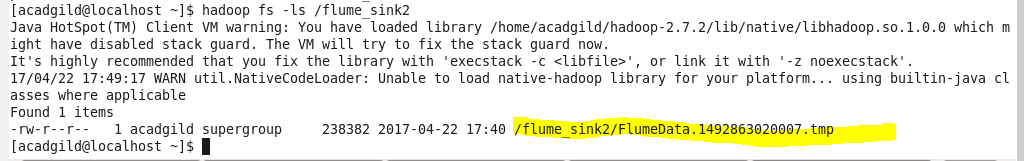




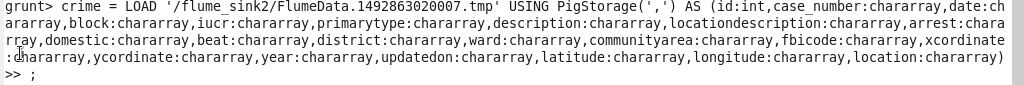
**After complete loading**



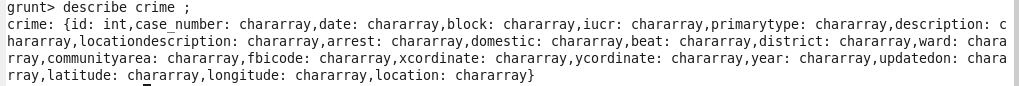




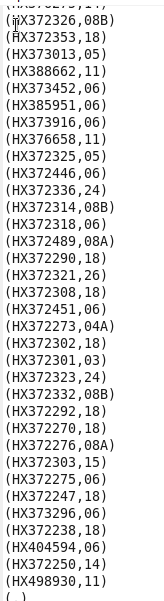
**Creating pig relation**



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

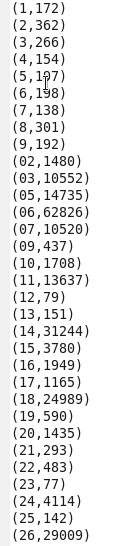


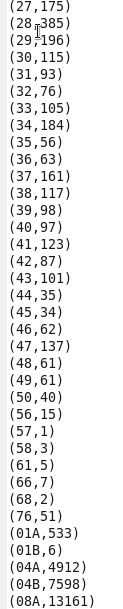








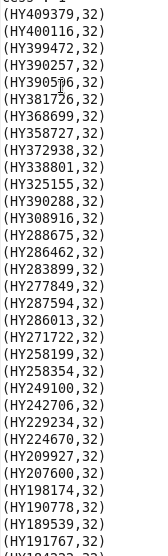




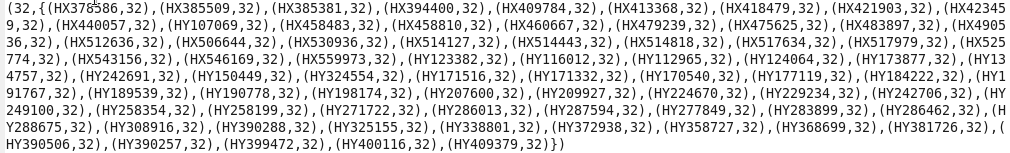


# **2.Write a mapreduce/pig program to calculate the number of cases investigated under FBI code 32.**







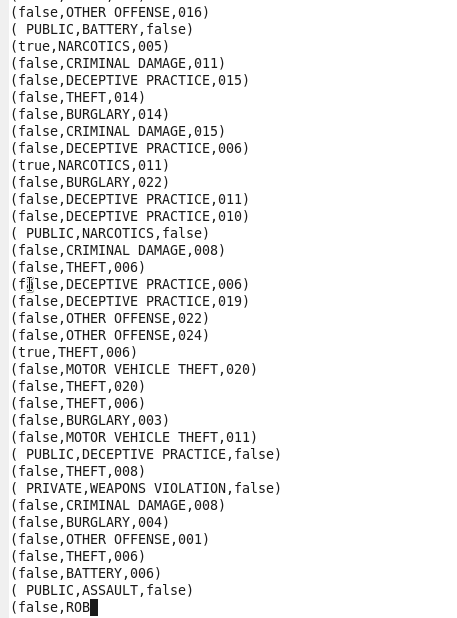


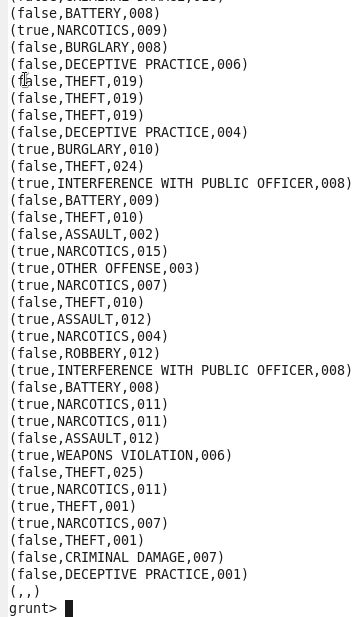




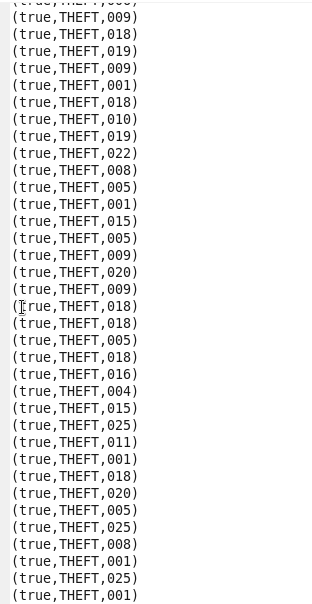
# **3.Write a mapreduce/pig program to calculate the number of arrests in theft district wise.**









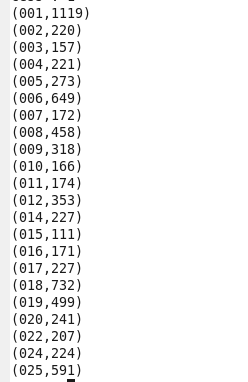








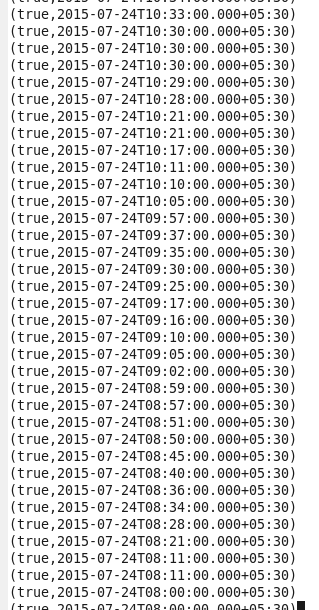
**Final output: district and total cases as required for theft**

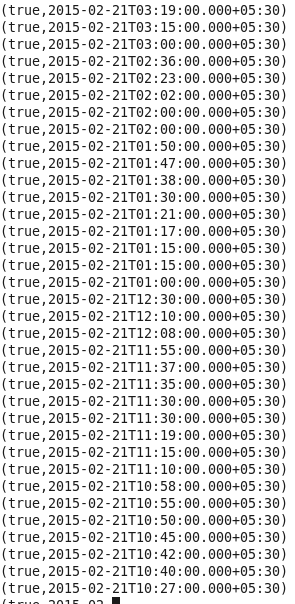


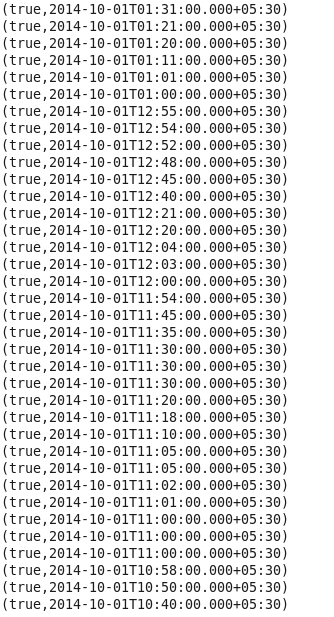
# **4.Write a mapreduce/pig program to calculate the number of arrests done between october 2014 and october 2015.**

















**True here means arrested**

**So total number of arrests between oct 2014 and oct 2015 = 63173**

