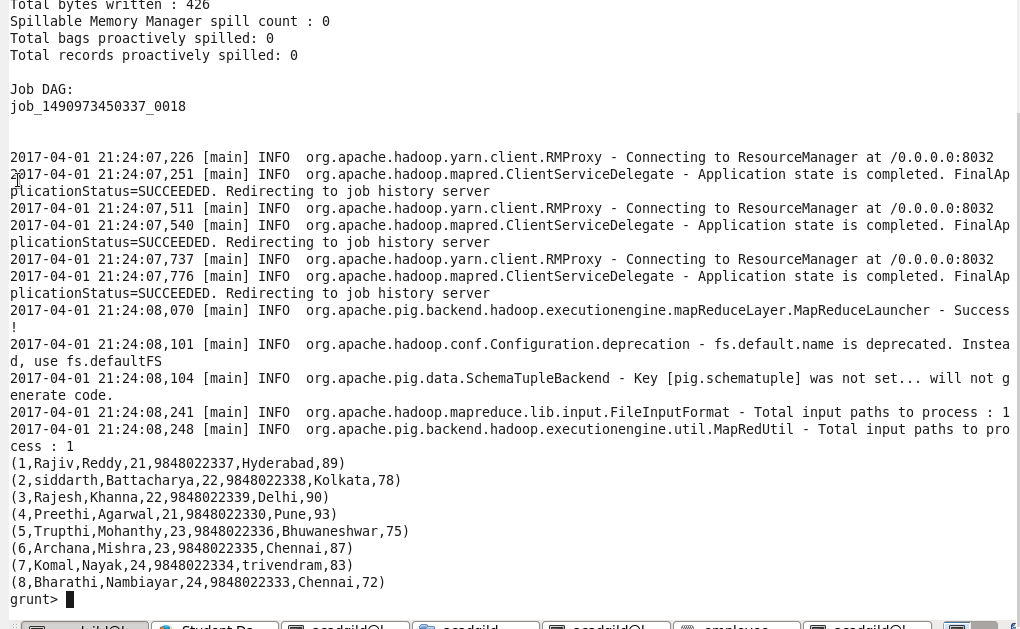
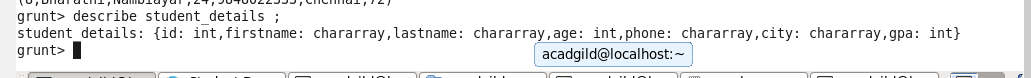


Dump Student\_details

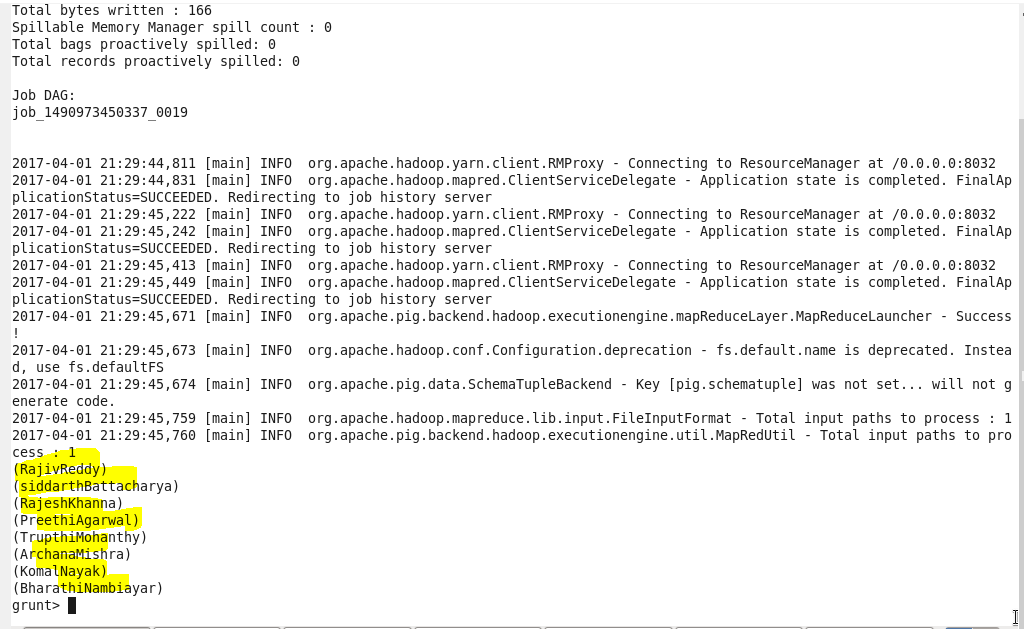


Describe student\_details

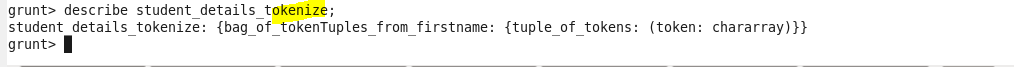


# Concat





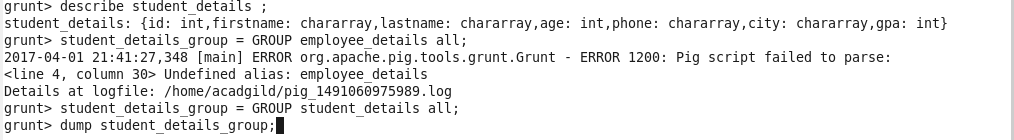
# Tokenize

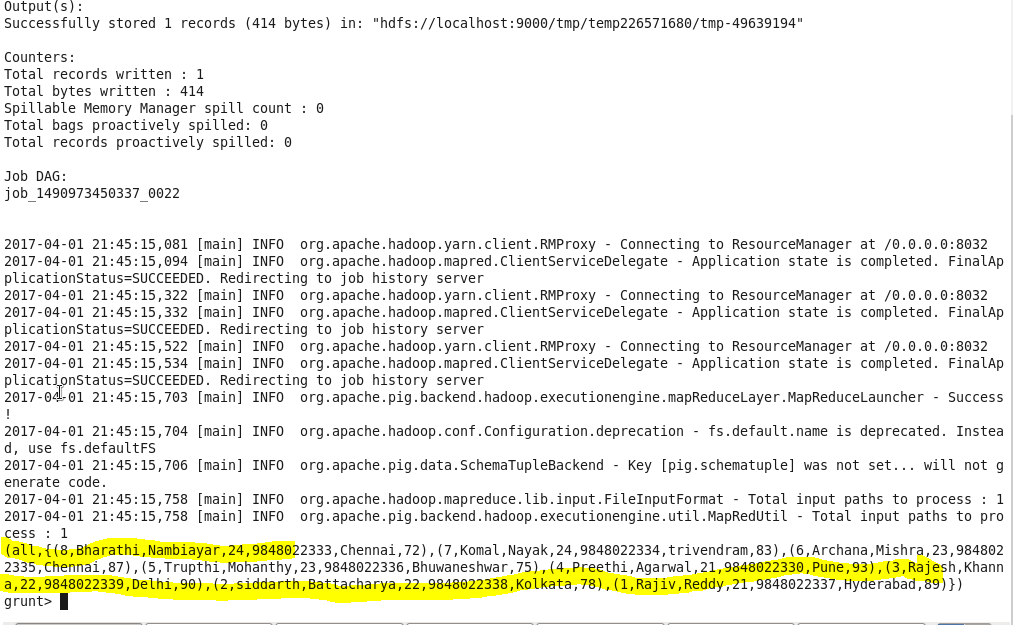




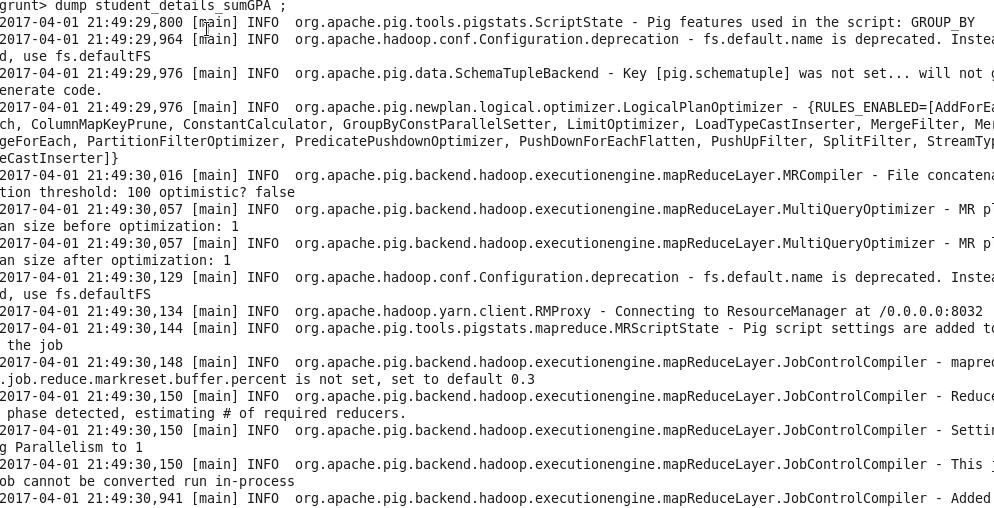
# Sum

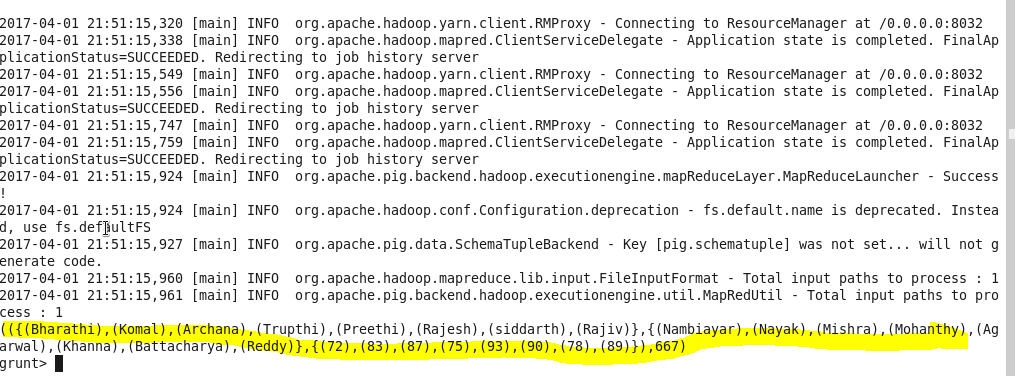
Calculating sum of GPA





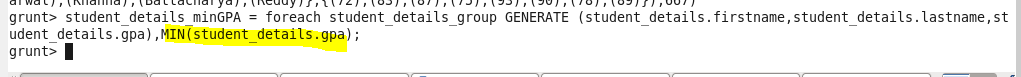


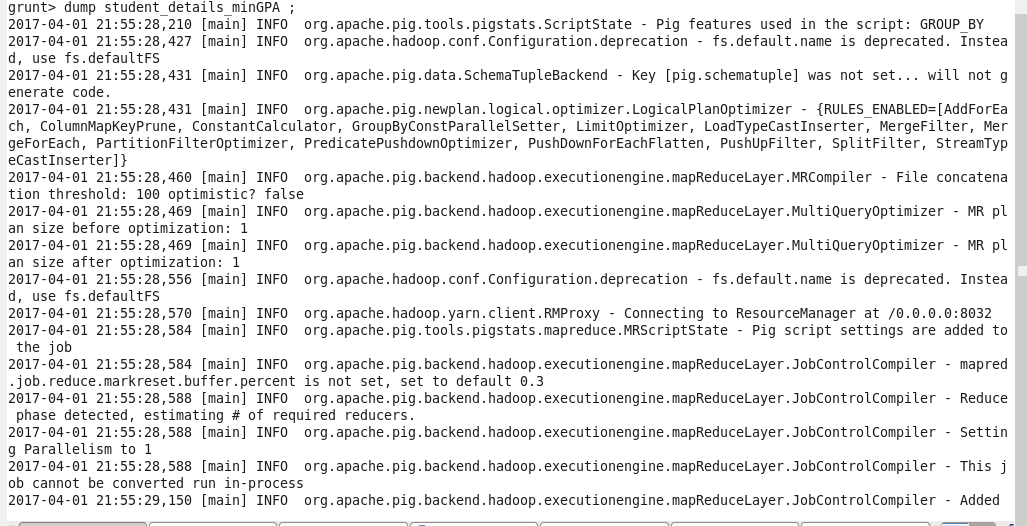


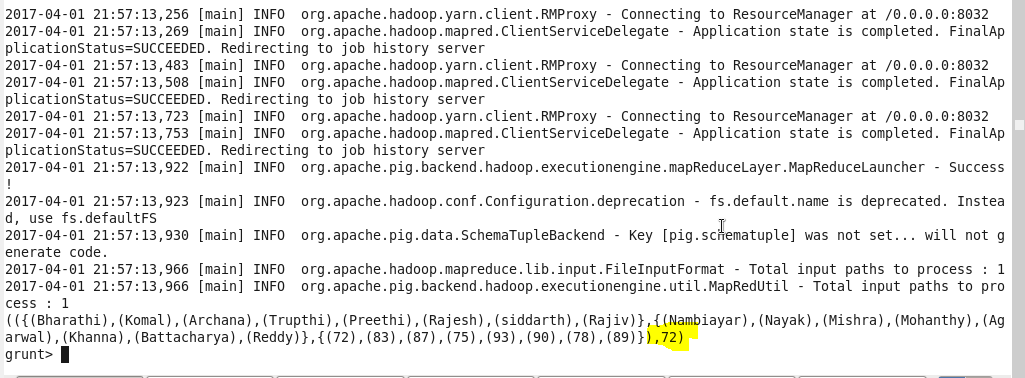


# Min

Find minimum of GPA



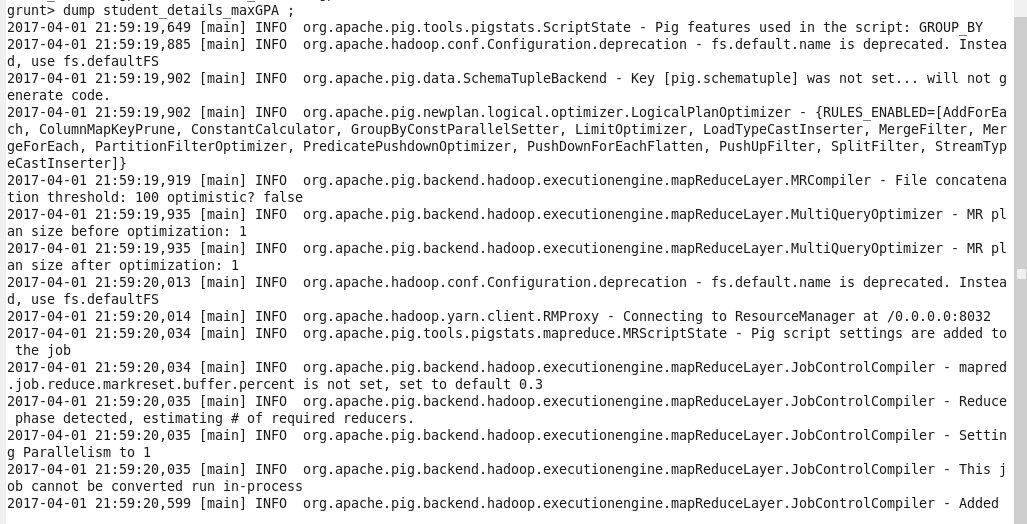


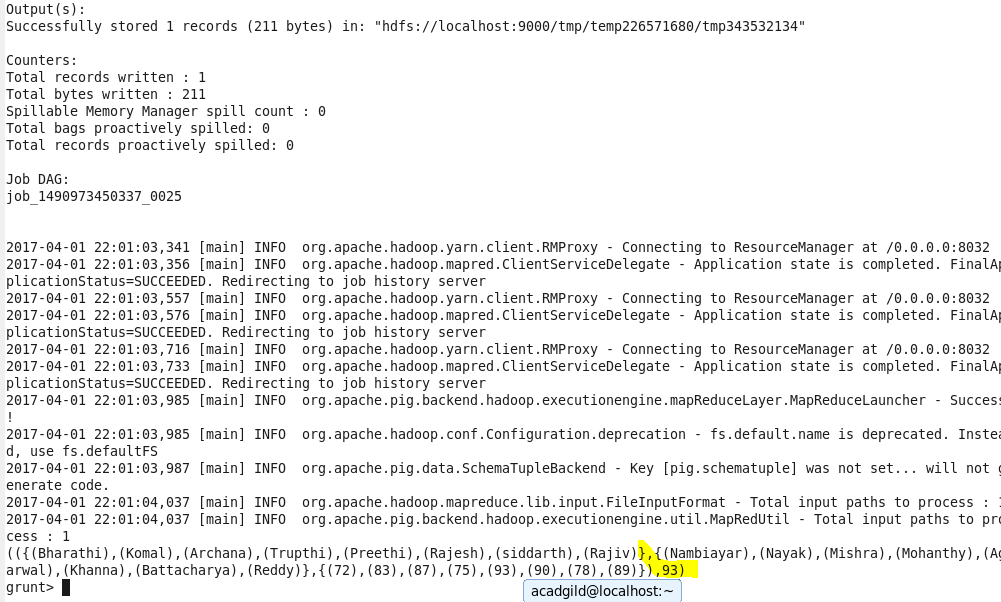


# Max

Find max of GPA

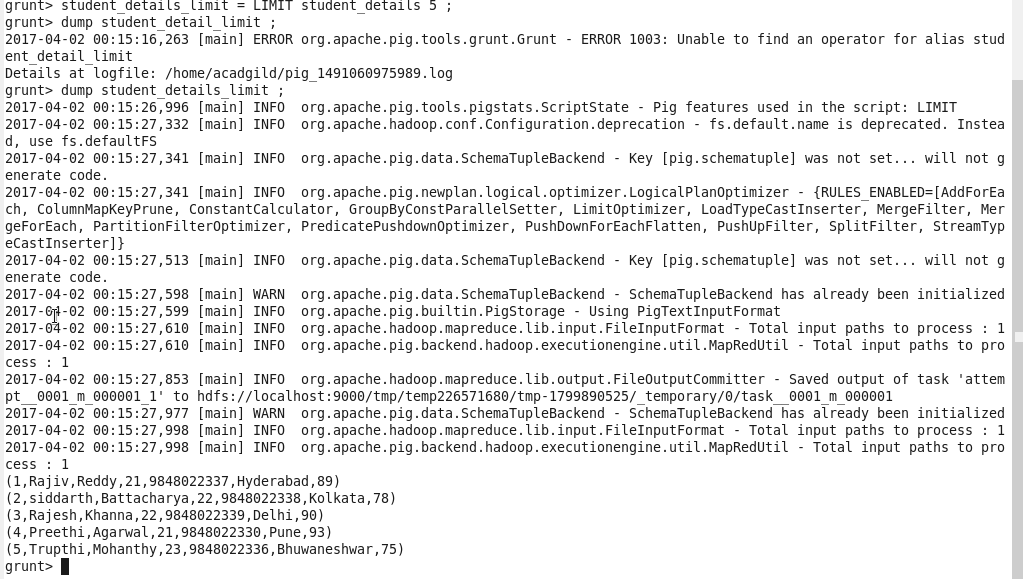






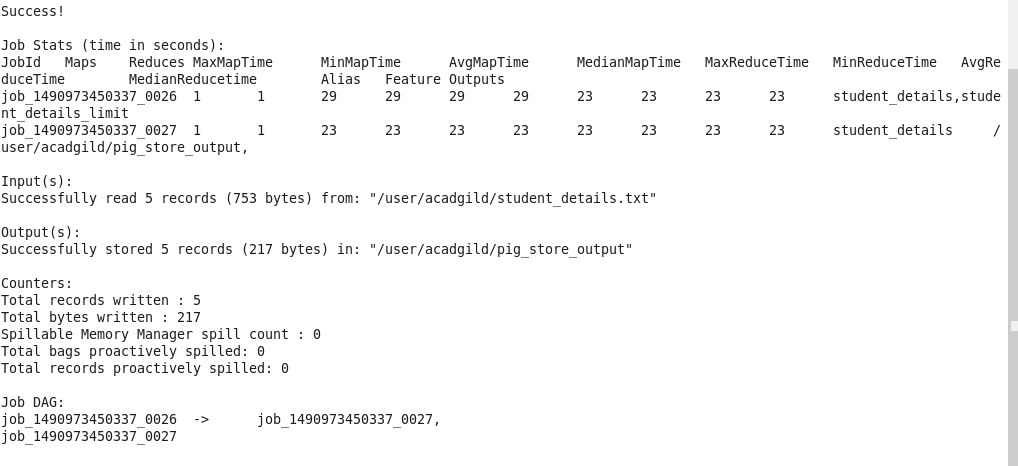
# Limit

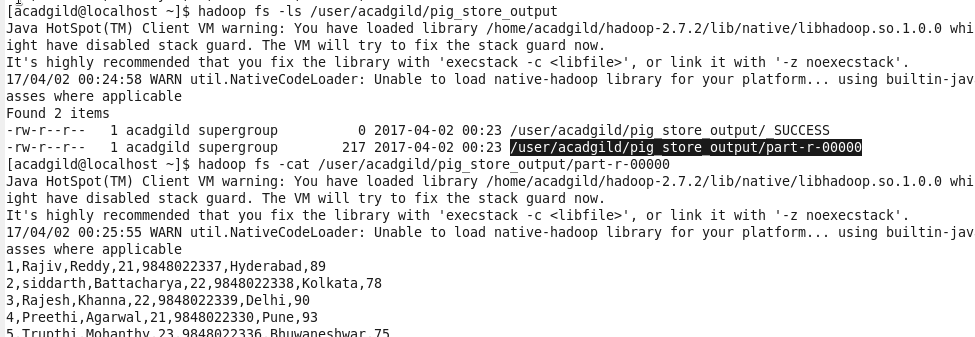
Student\_details\_limit = LIMIT student\_details 5 ;



# Store



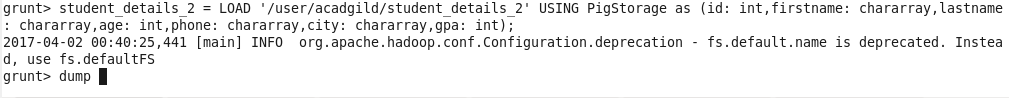


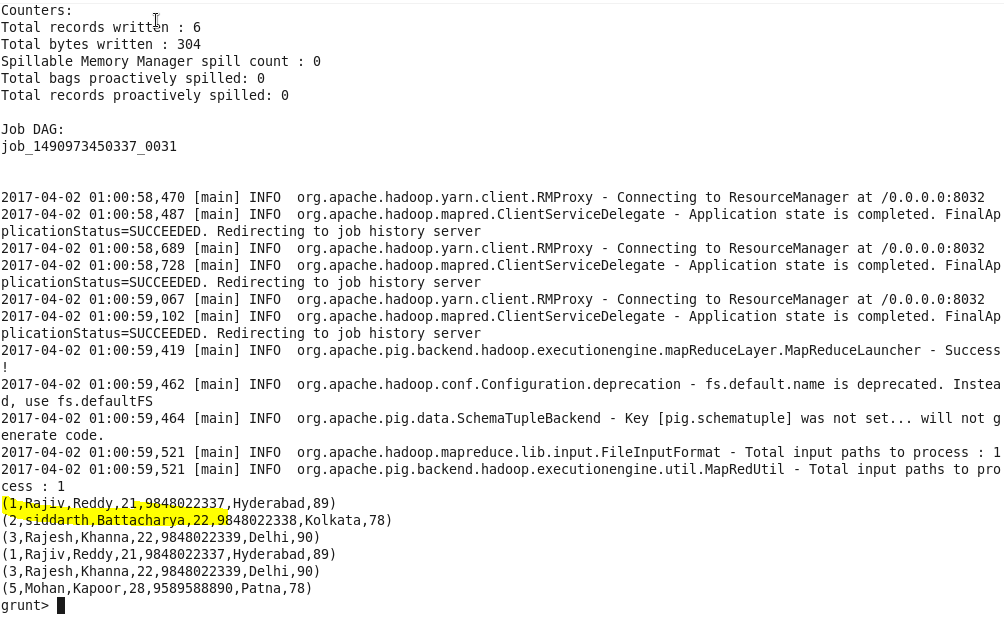


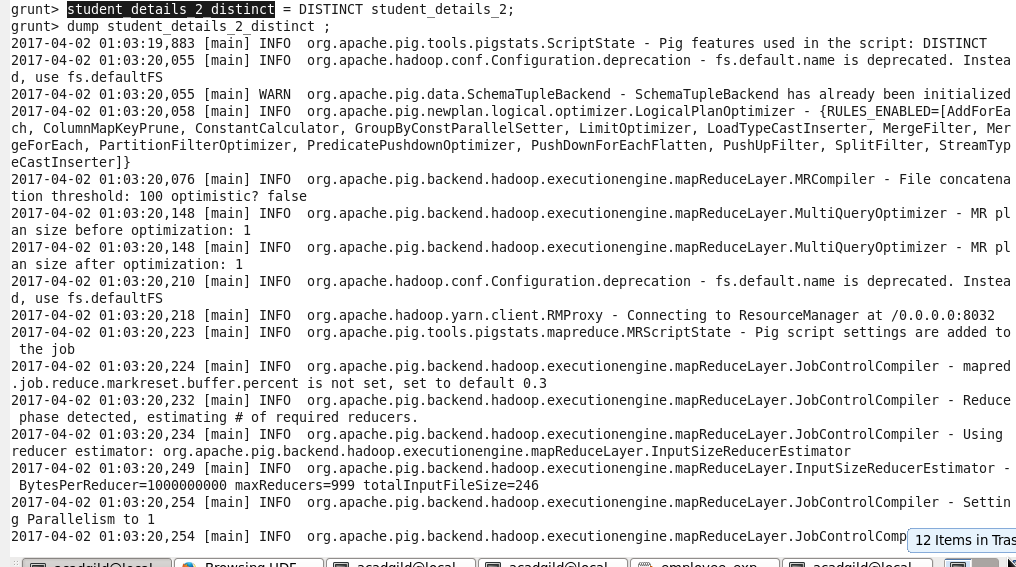
# Distinct

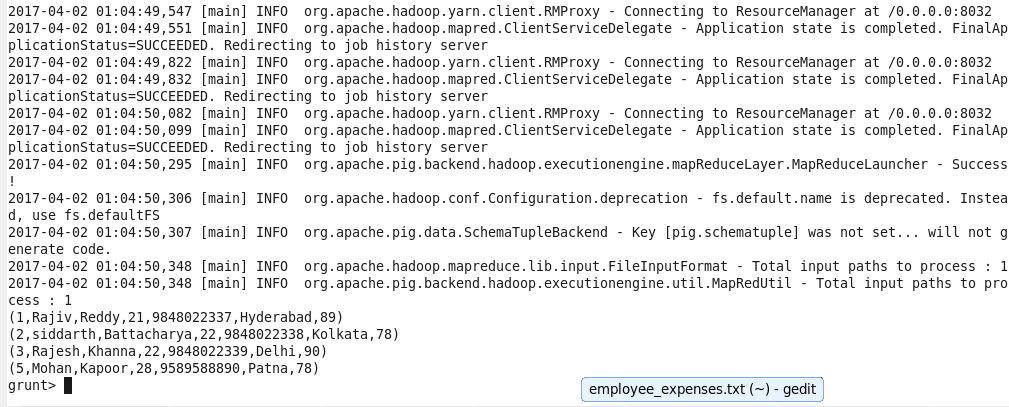


Loading a different file into Pig with few duplicate data





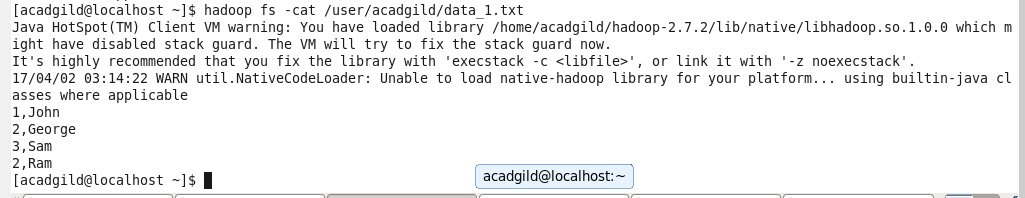




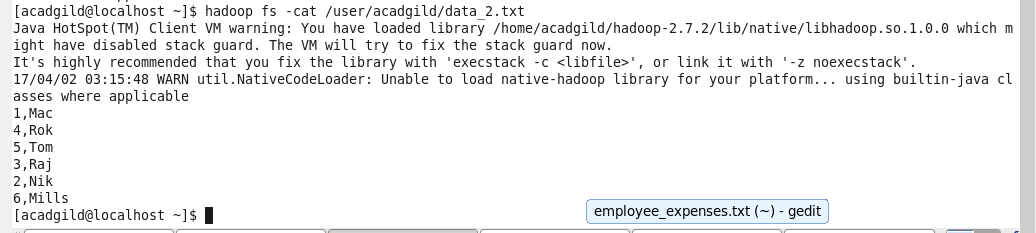
# Flatten

Loaded 2 data files in HDFS

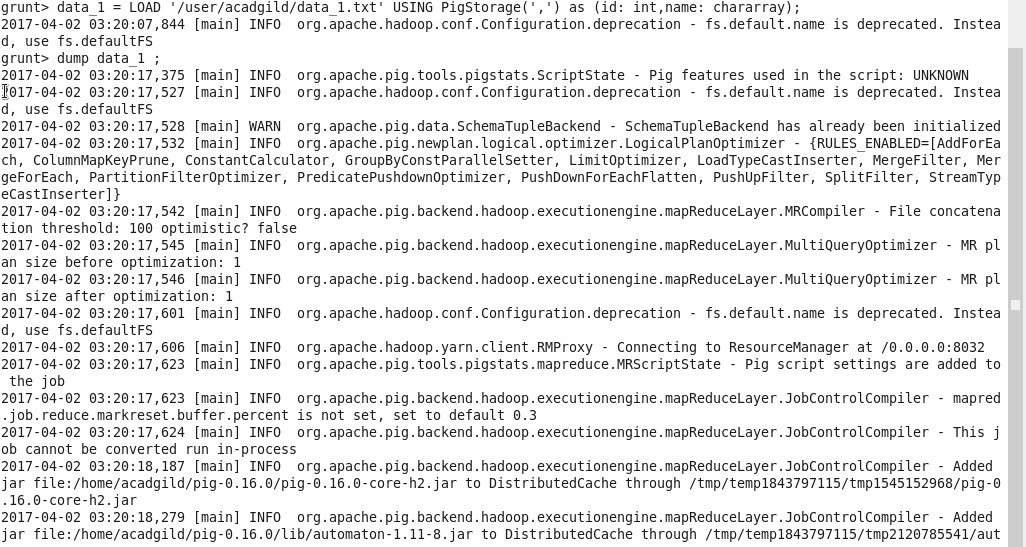
Data\_1.txt



Data\_2.txt

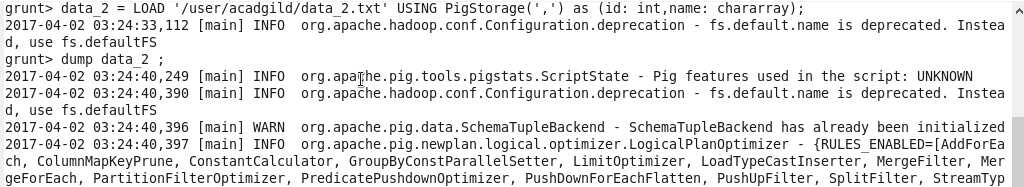


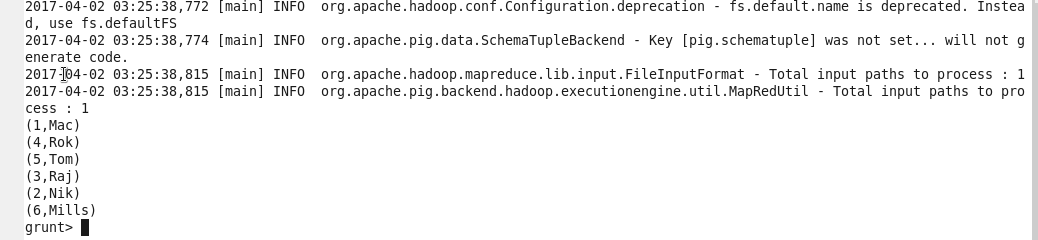
Loading data\_1 in pig

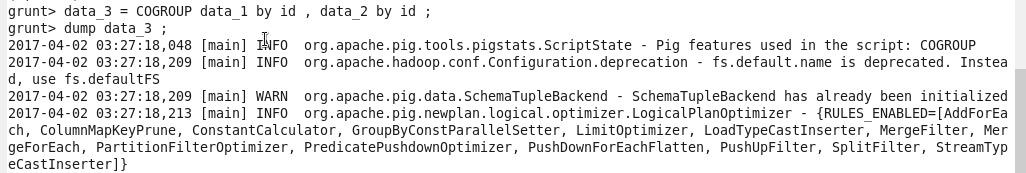


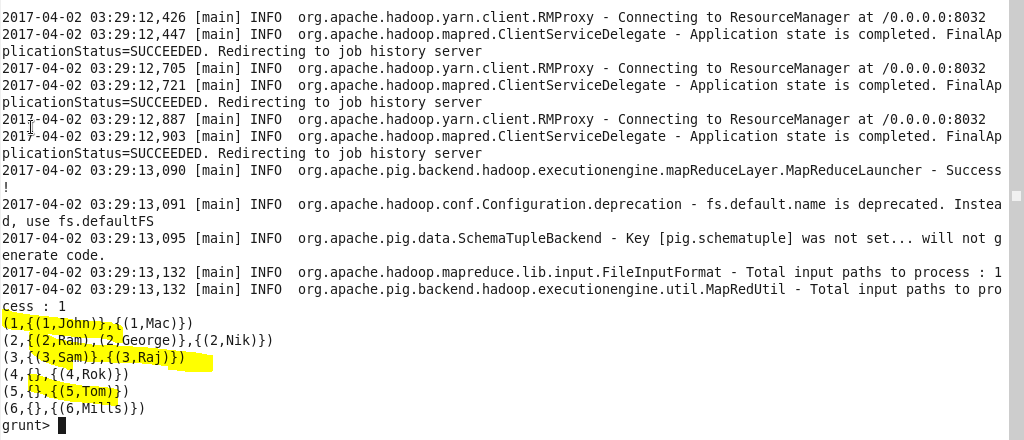


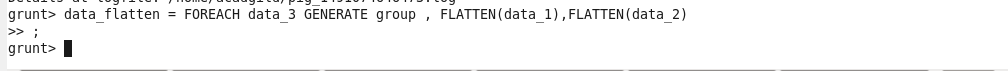
Loading data\_2 in Pig



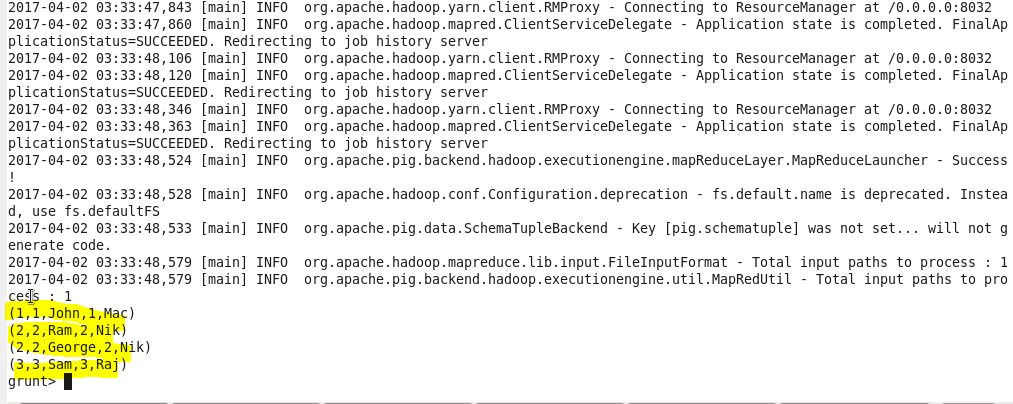




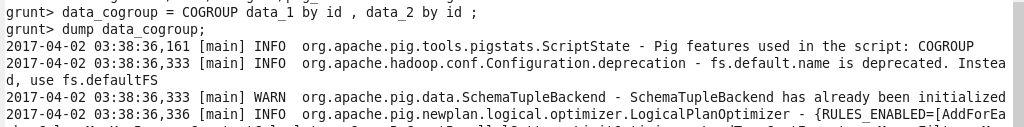


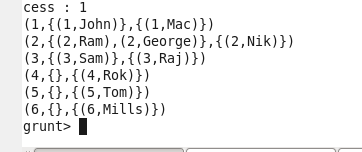


Flatten output



# IsEmpty





IsEmpty code

