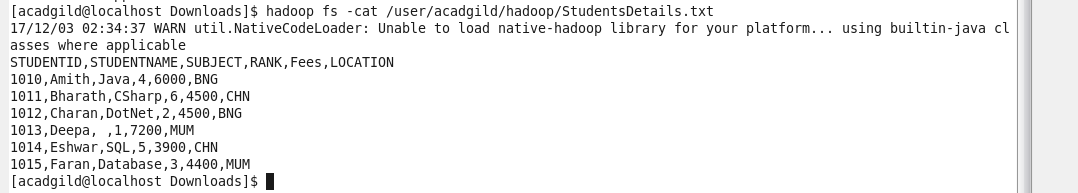
**Assignment 4.2**

Sample dataset to implement pig commands:



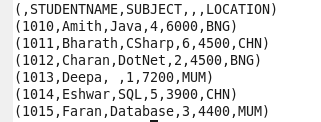
Copying Sample dataset from local to hdfs:



Loading sample dataset:



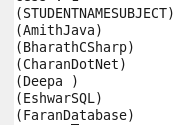




1. **Concat**

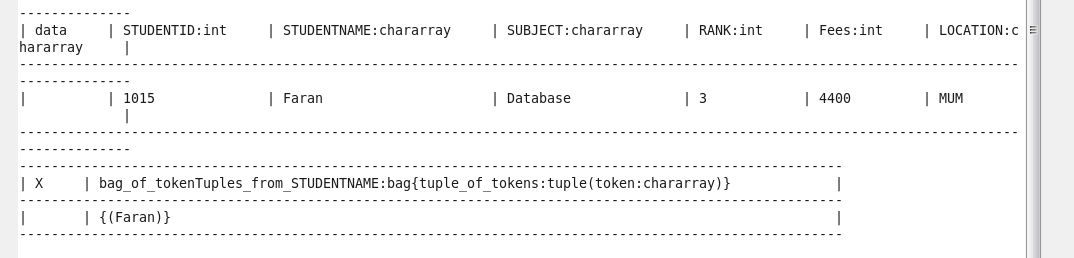
Concatenates two fields of type chararray or bytearray

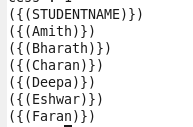




1. **Tokenize**



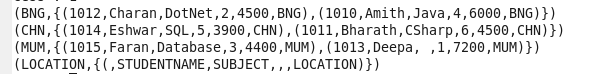




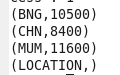
1. **SUM**

Computes the sum of the numeric values in a single-column bag. SUM requires a preceding GROUP ALL statement for global sums and a GROUP BY statement for group sums.







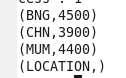


1. **MIN**

Computes the minimum of the numeric values or chararrays in a single-column bag.

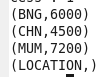
MIN requires a preceding GROUP ALL statement for global minimums and a GROUP BY statement for group minimums.





1. **MAX**





1. **LIMIT**

****

****

1. **STORE**

Stores or saves results to the file system



1. **Distinct**

Added few duplicate records in to a dataset

Using DISTINCT removes duplicate tuples in a relation





Only distinct tuples in a relation are displayed

