**1)**Create a sample dataset and implement the below Pig commands on the same dataset.

1)Concat

2) Tokenize

3) Sum

4) Min

5) Max

6) Limit

7) Store

8) Distinct

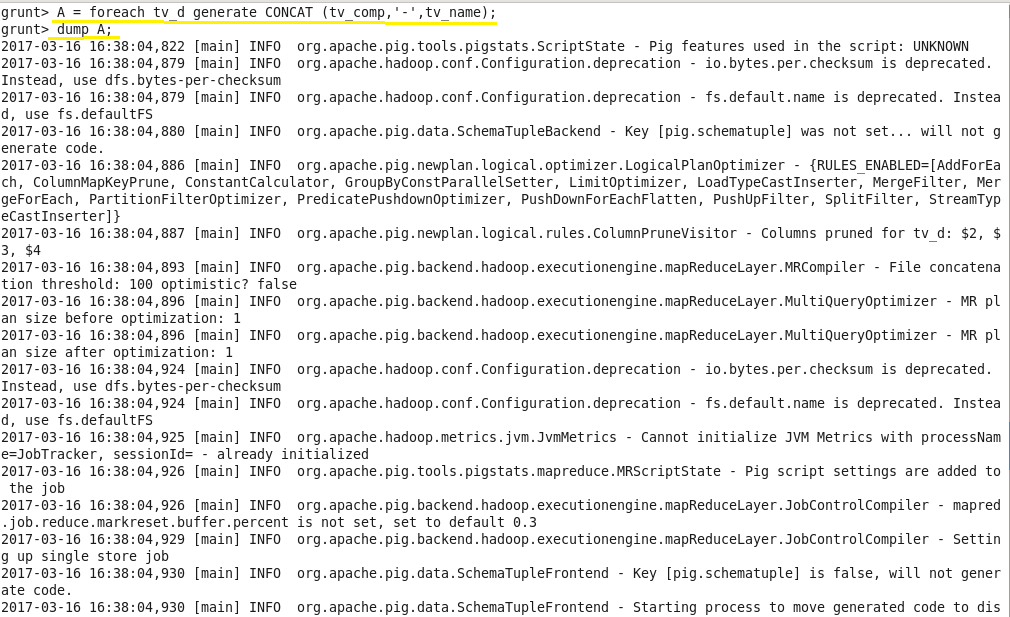
9) Flatten

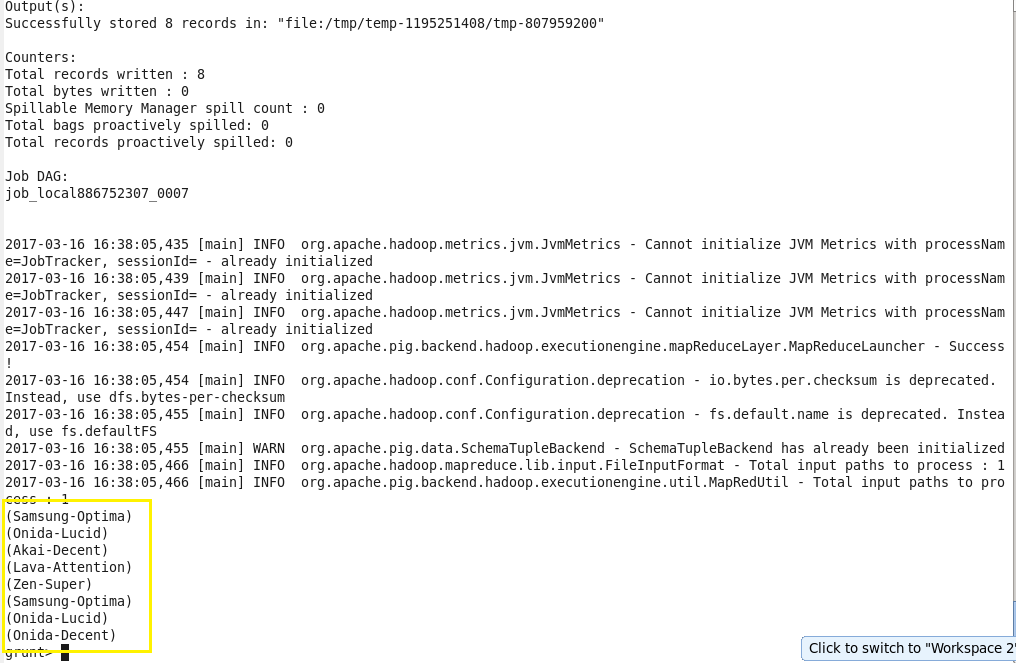
10) IsEmpty

**Answers:**

1. **concat:**

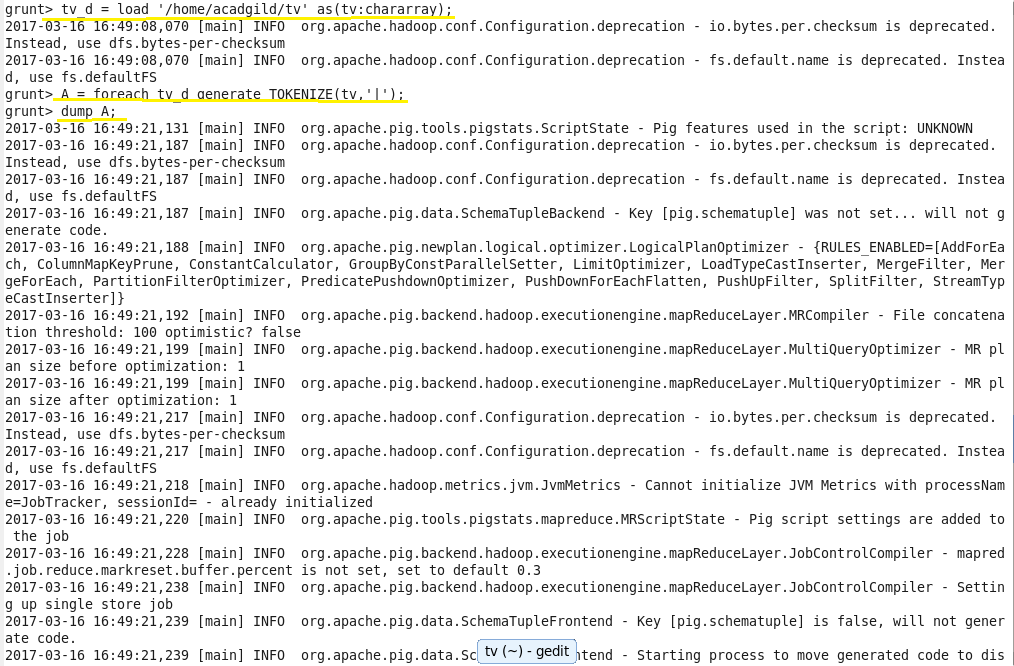
Concatenates two fields of type chararray or two fields of type bytearray.

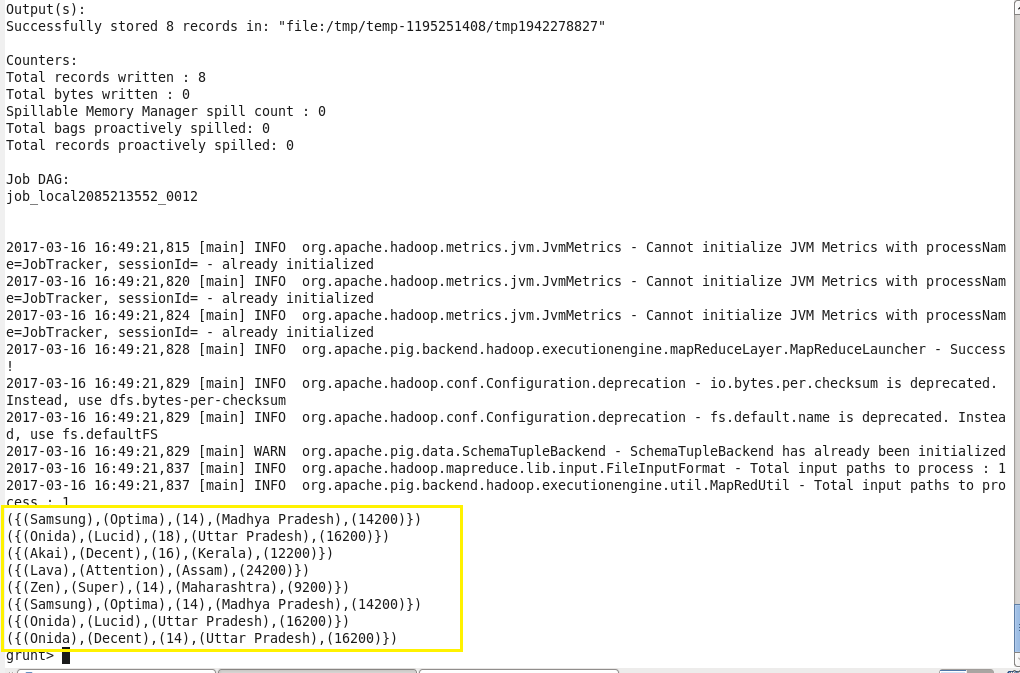




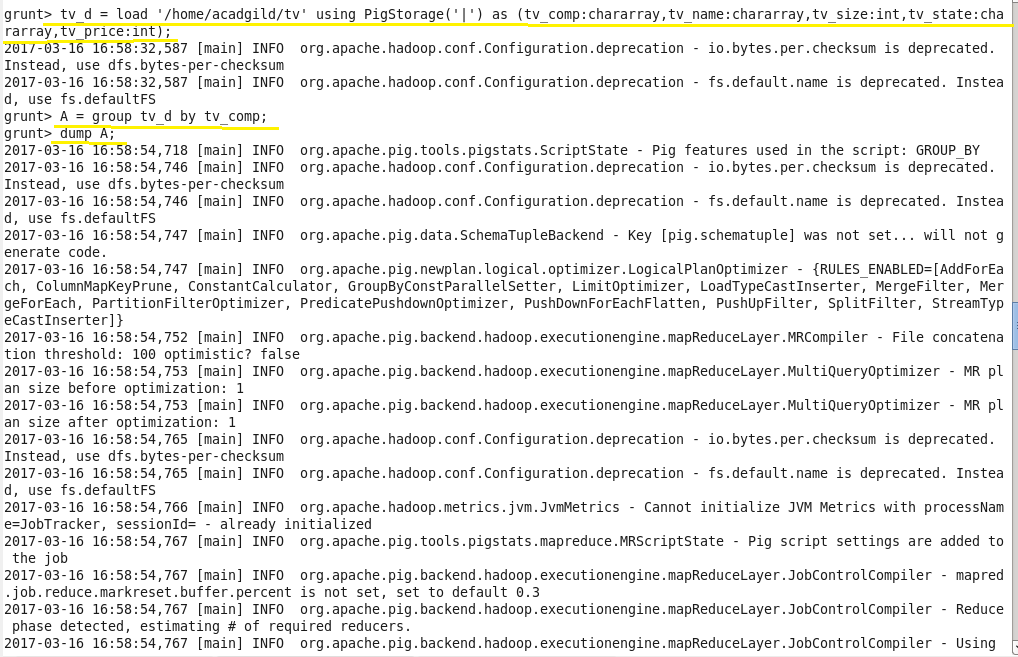
1. **Tokenize:**

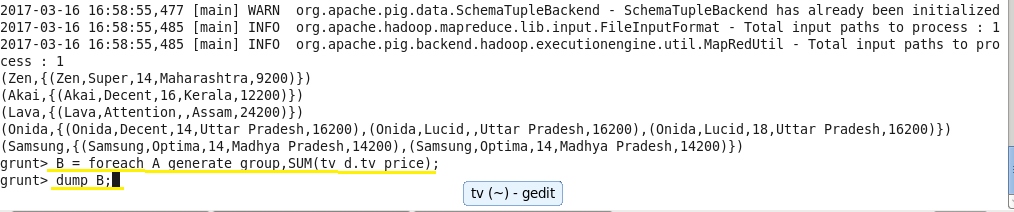
Splits a string and outputs a bag of words.

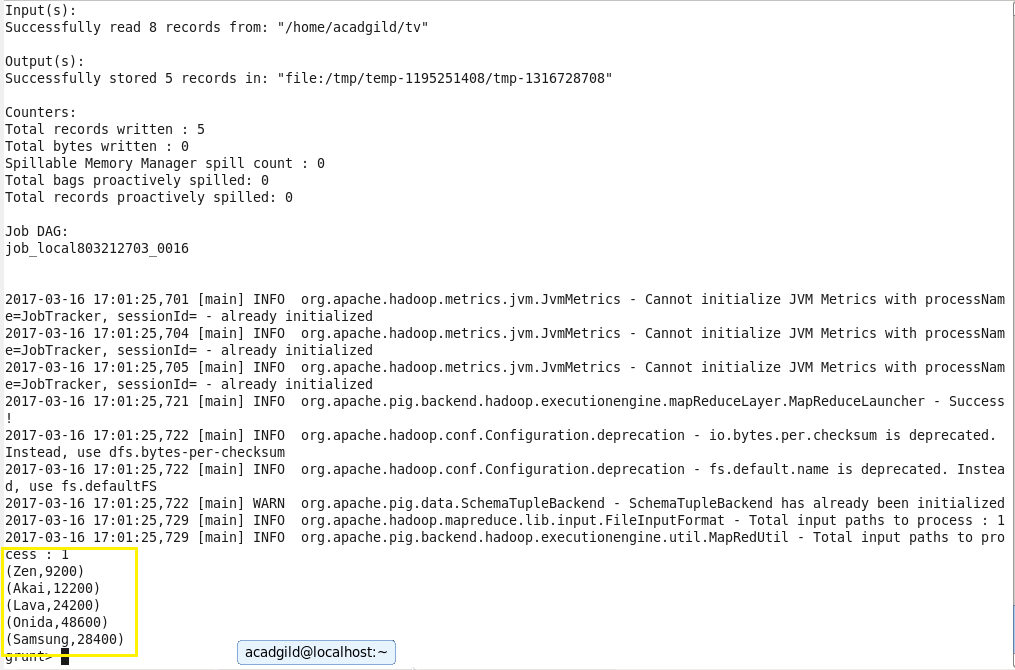




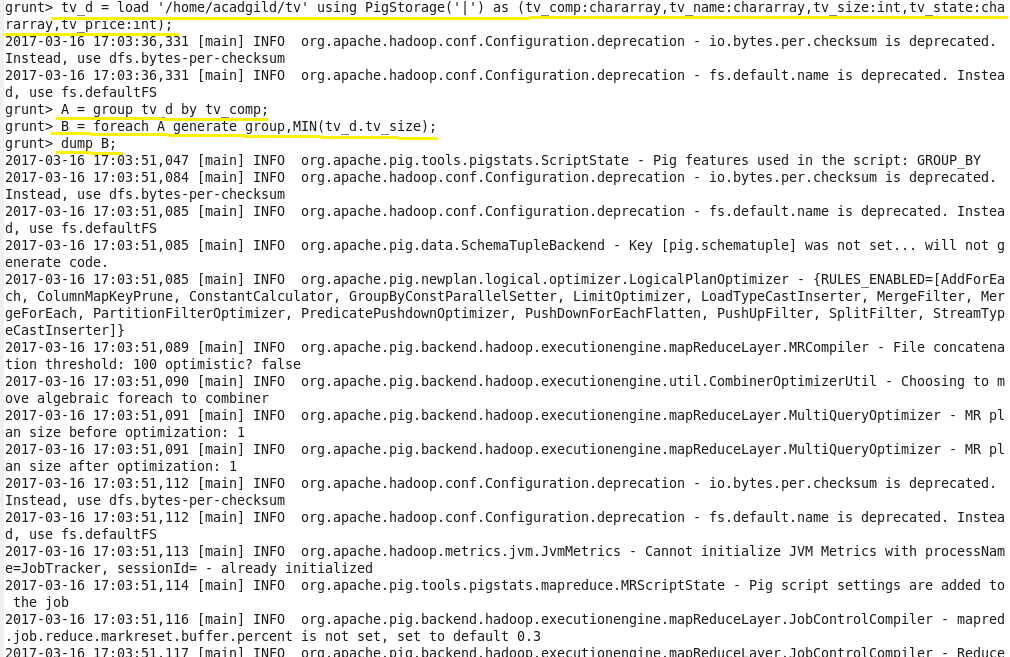
1. **Sum:** Computes the sum of the numeric values in a single-column bag. SUM requires a preceding GROUP ALL statement for global sums and a GROUP BY statement for group sums.

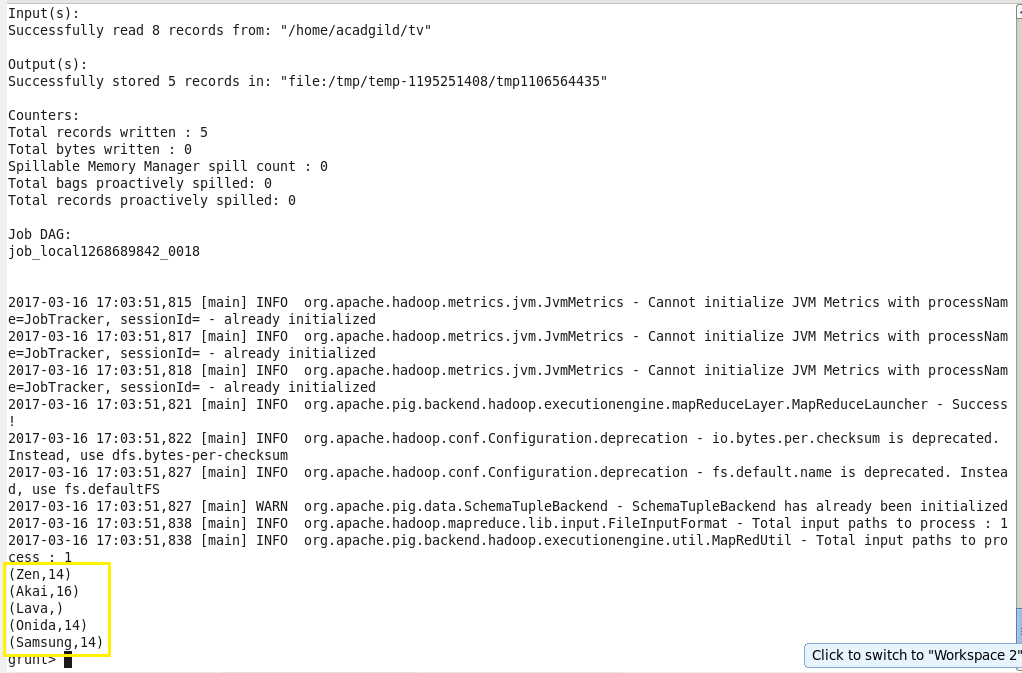






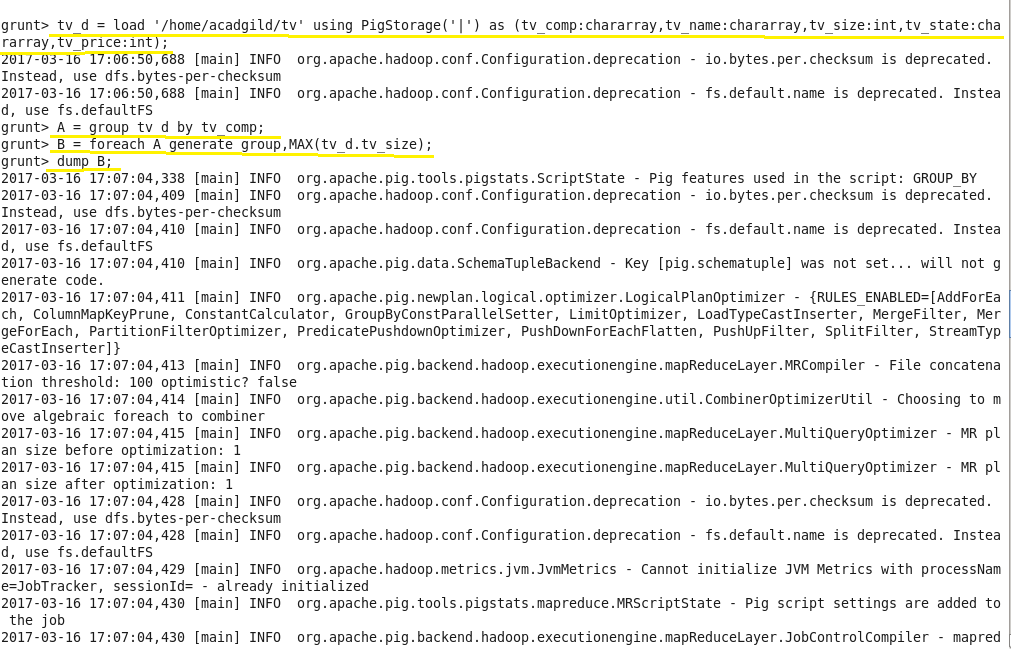
1. **Min**: Computes the minimum of the numeric values or chararrays in a single-column bag. MIN requires a preceding GROUP… ALL statement for global minimums and a GROUP … BY statement for group minimums.

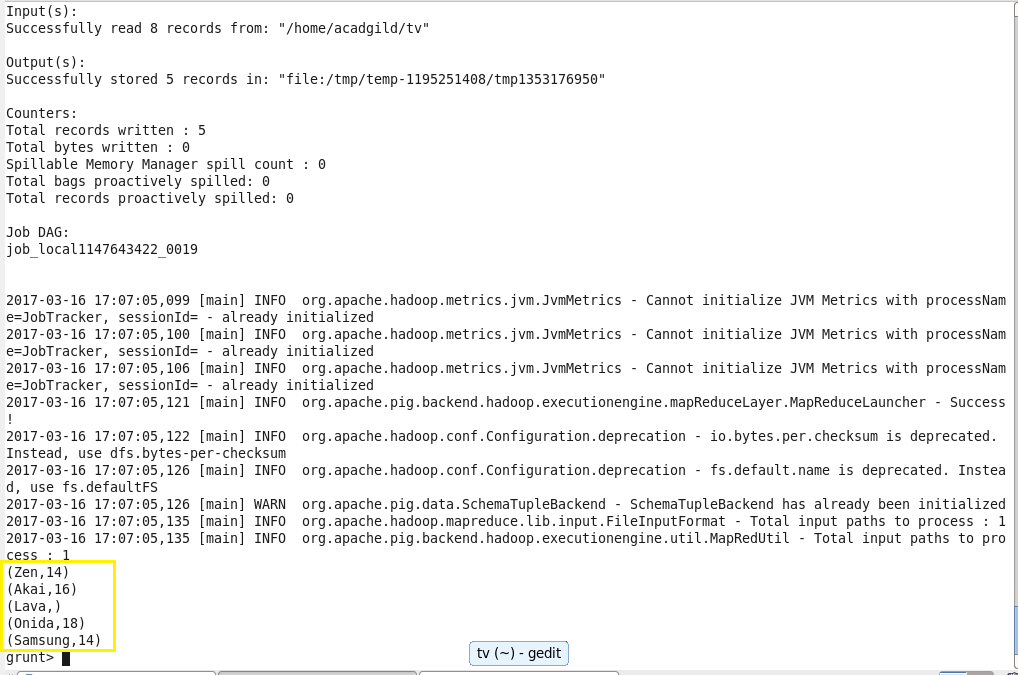




1. **Max:**

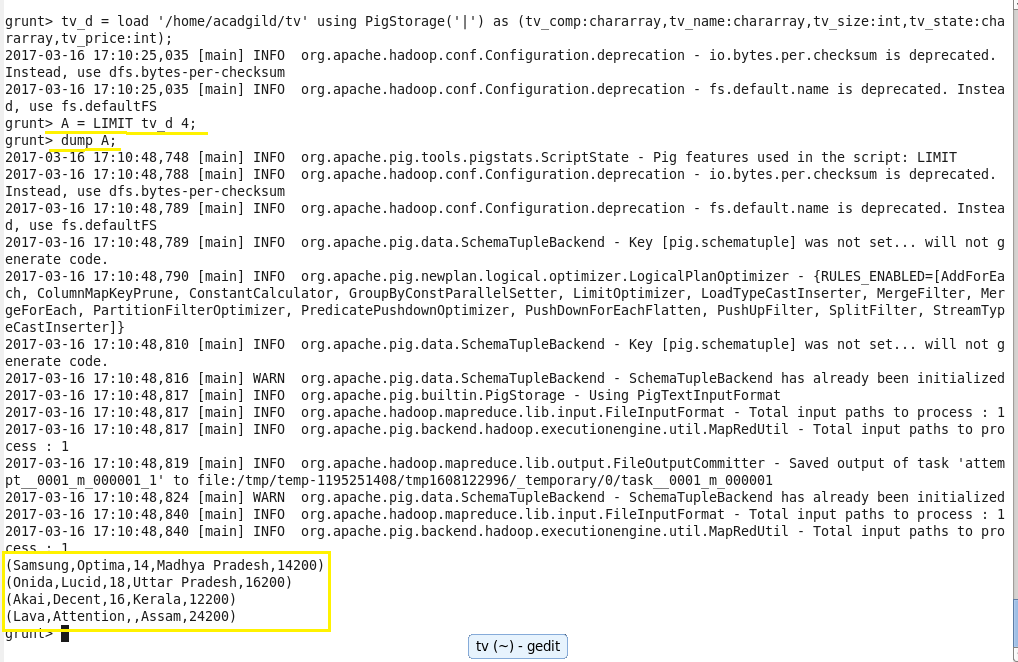
Computes the maximum of the numeric values or chararrays in a single-column bag. MAX requires a preceding GROUP ALL statement for global maximums and a GROUP BY statement for group maximums.





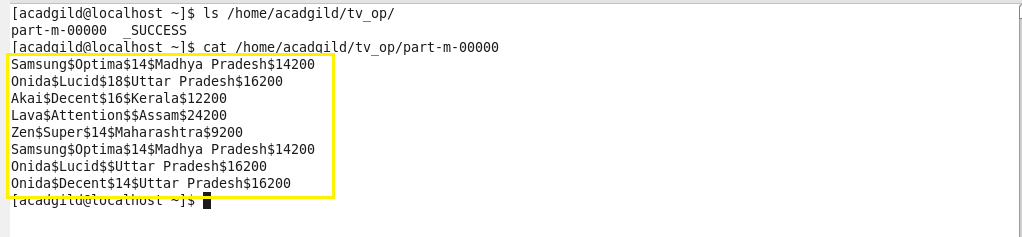
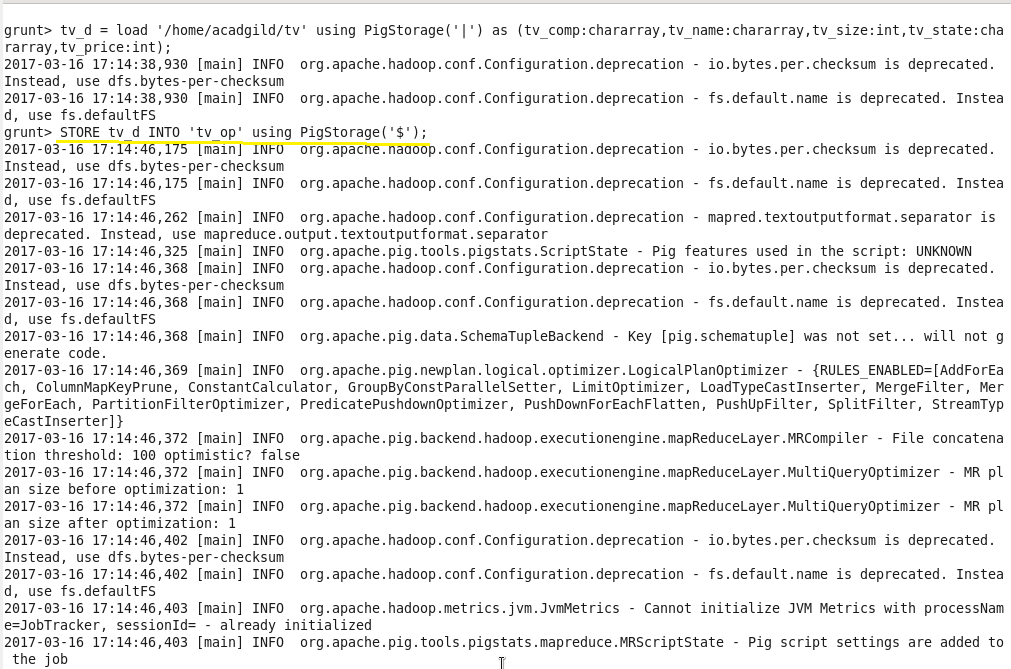
1. **Limit:**

Limits the number of output tuples.



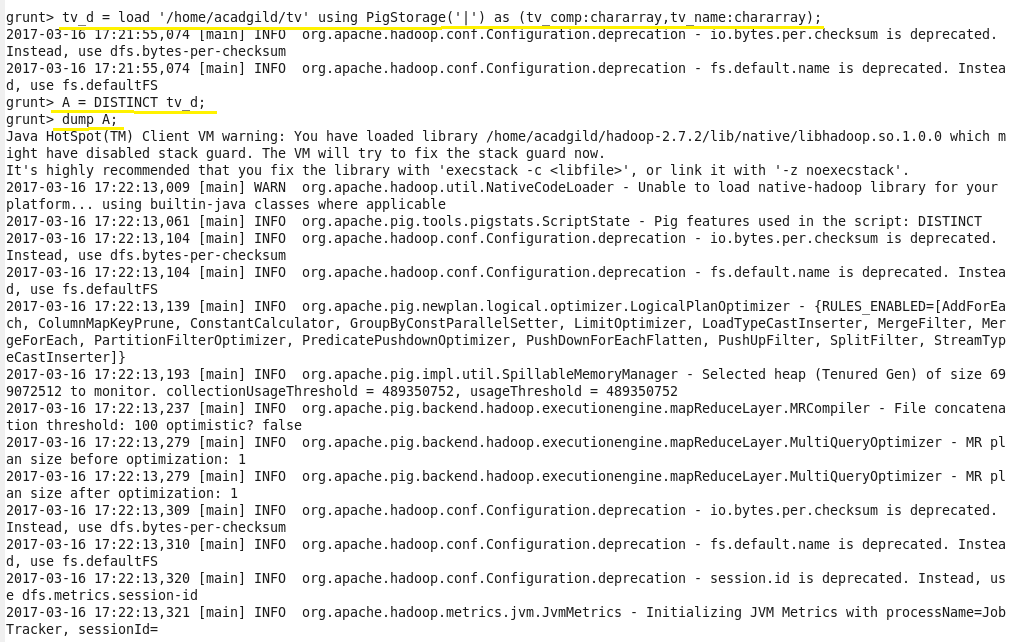
1. **Store:**

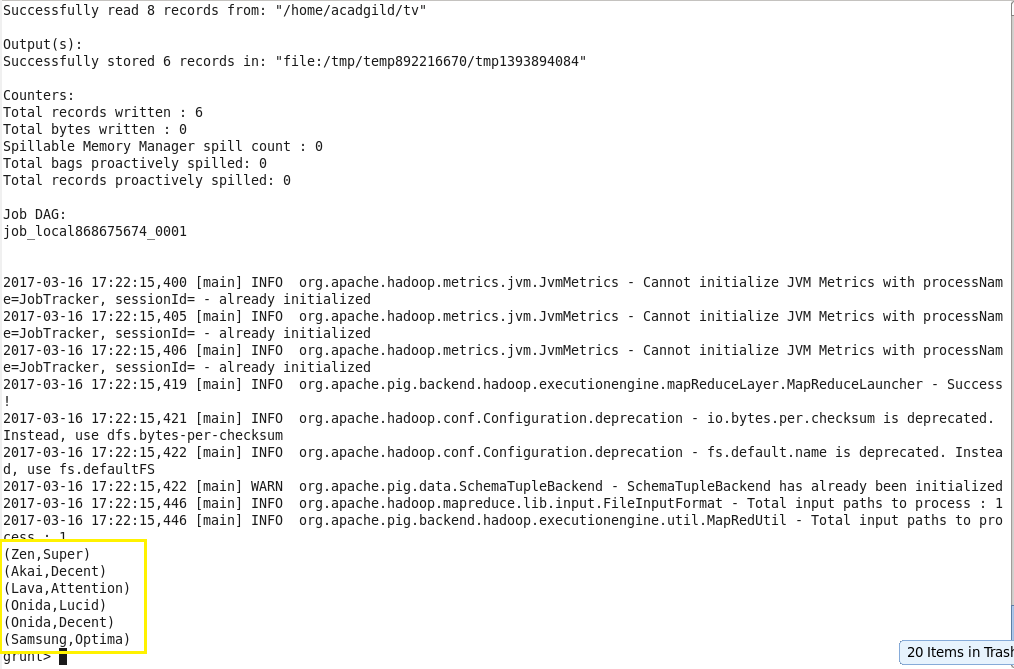
Stores or saves results to the file system.



1. **Distinct:**

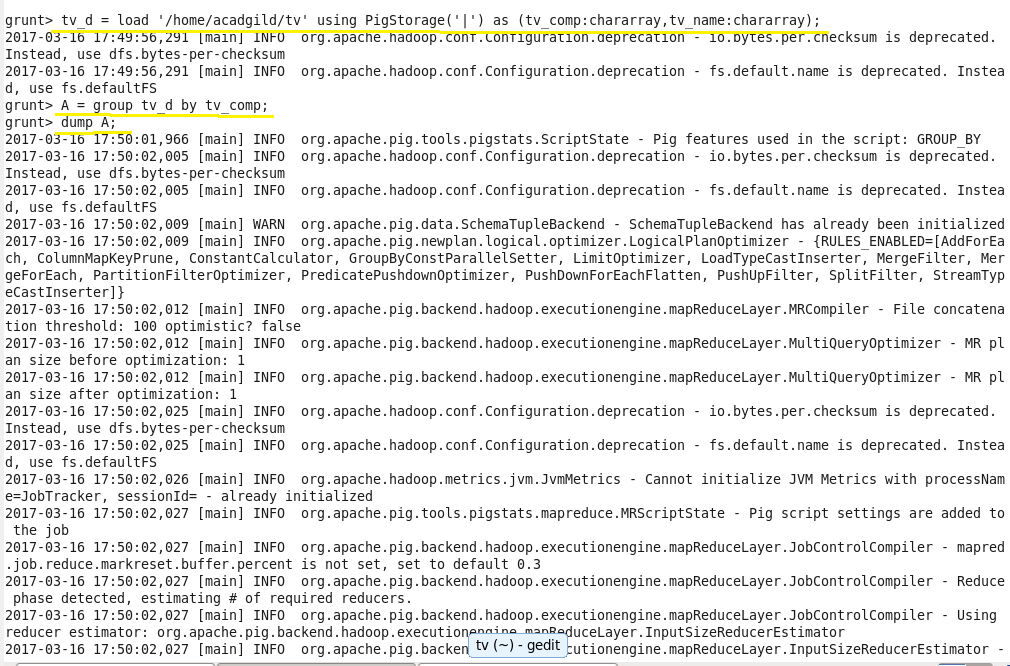
Removes duplicate tuples in a relation.

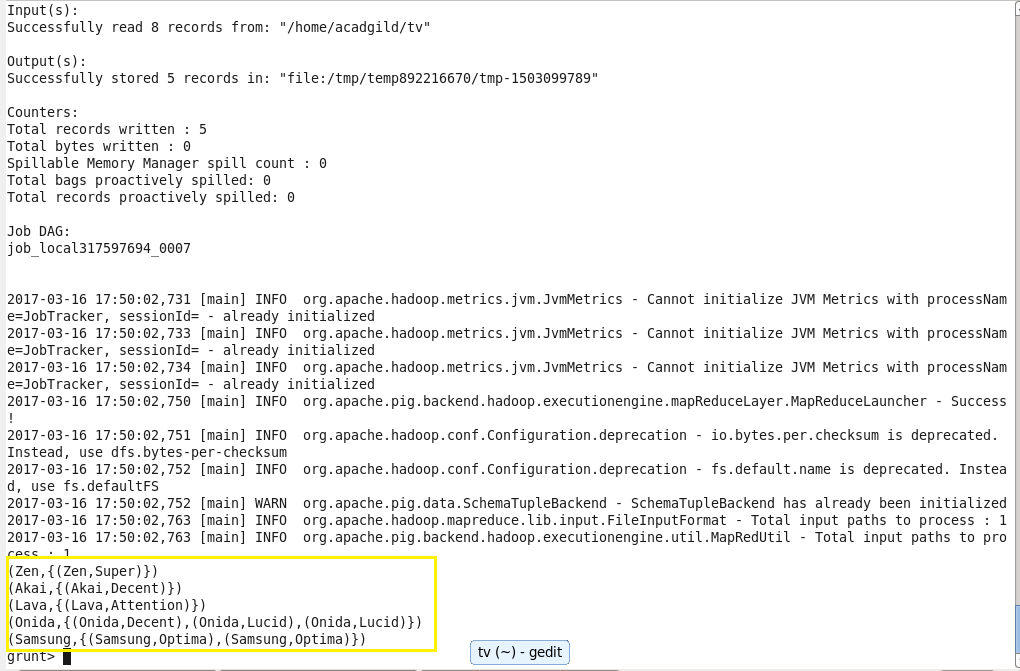


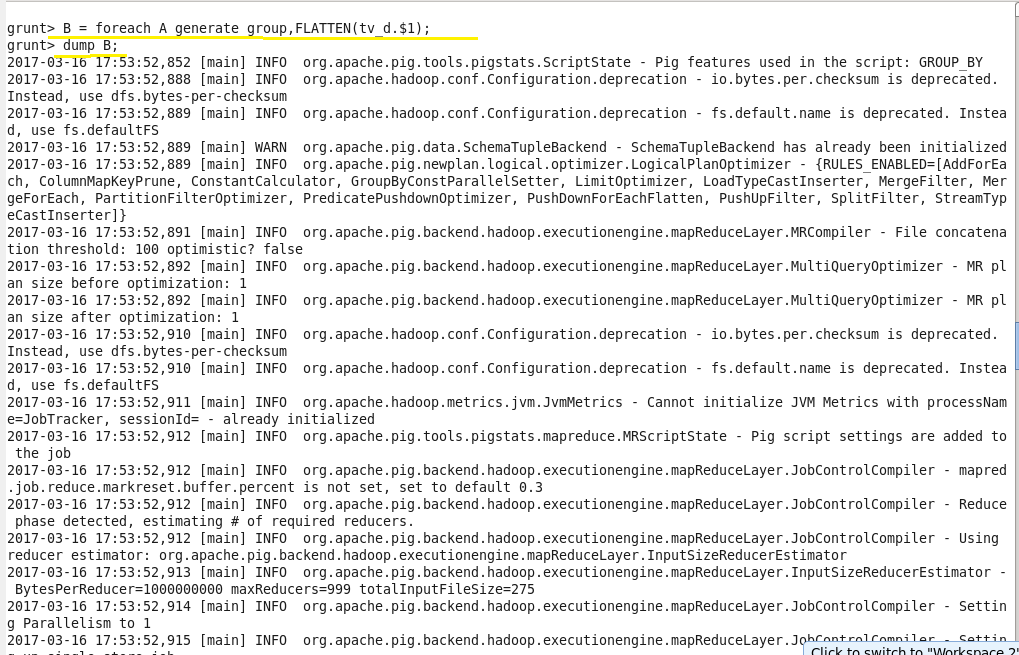


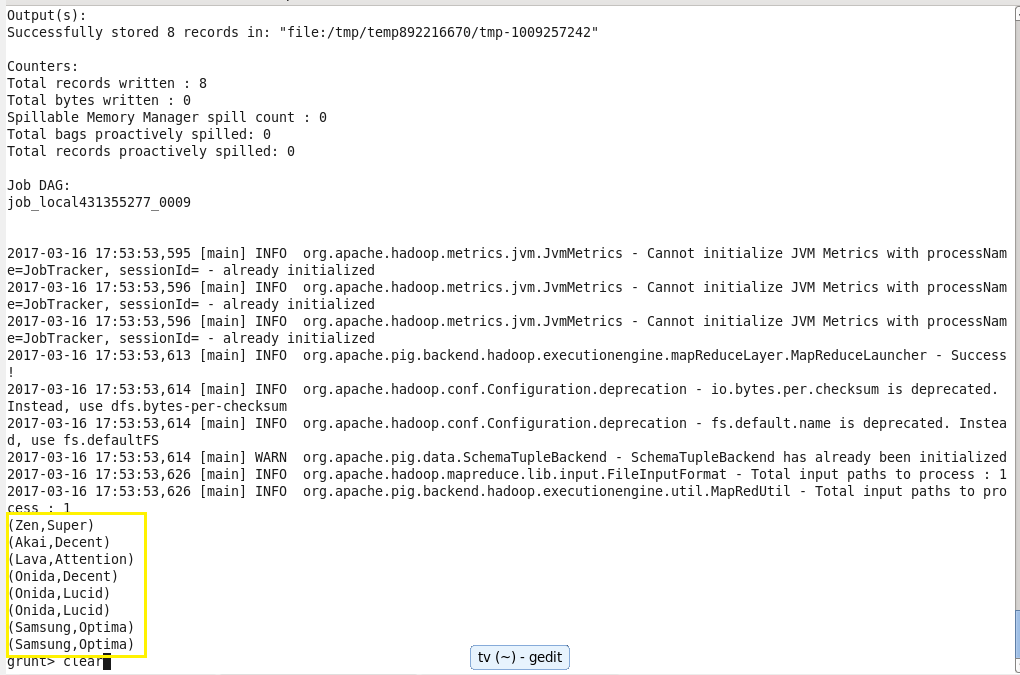
1. **Flatten:**

The FLATTEN operator looks like a UDF syntactically, but it is actually an operator that changes the structure of tuples and bags in a way that a UDF cannot. Flatten un-nests tuples as well as bags.









1. **IsEmpty :**

Checks if a bag or map is empty.

