**Assignment 9.4**

**Question:** Practice the relational operators in pig by following the steps in the below blog

**Relational Operators:**

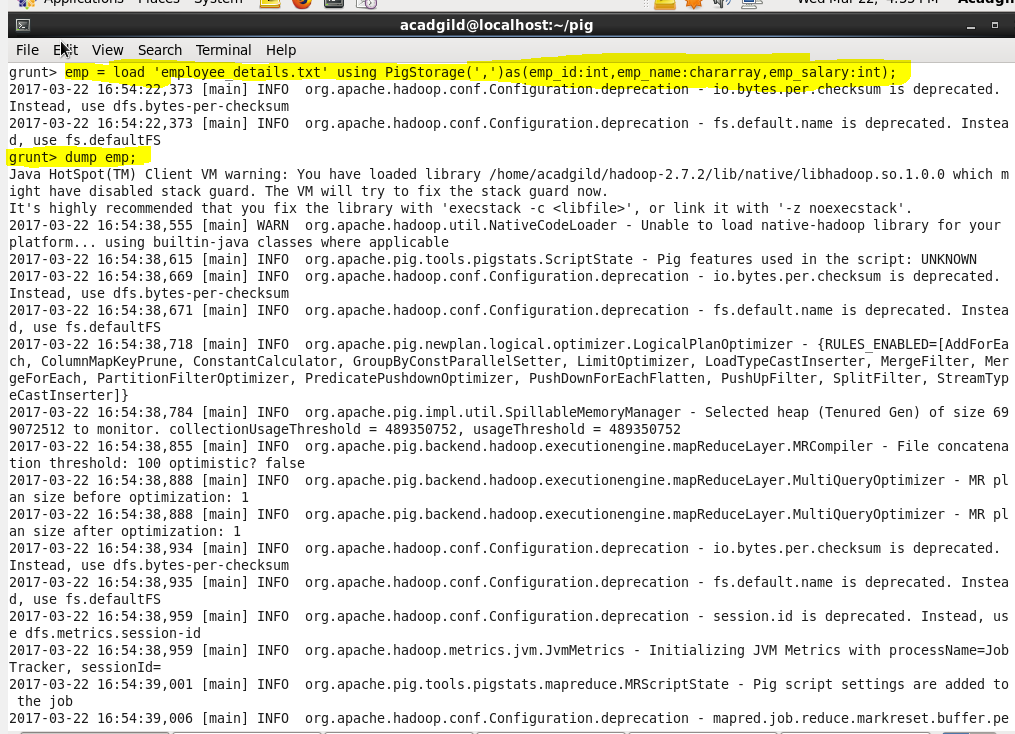
***LOAD:***

* We can load data into Pig from the file system (HDFS/ Local) using LOAD operator of Pig Latin.

**Syntax:**

***Relation\_name = LOAD ‘Input file Path’ USING function as schema;***

The load statement consists of two parts divided by the “=” operator. On the left-hand side, we need to mention the name of the relation where we want to store the data, and on the right-hand side, we have to define how we store the data. Given below is the syntax of the Load operator.

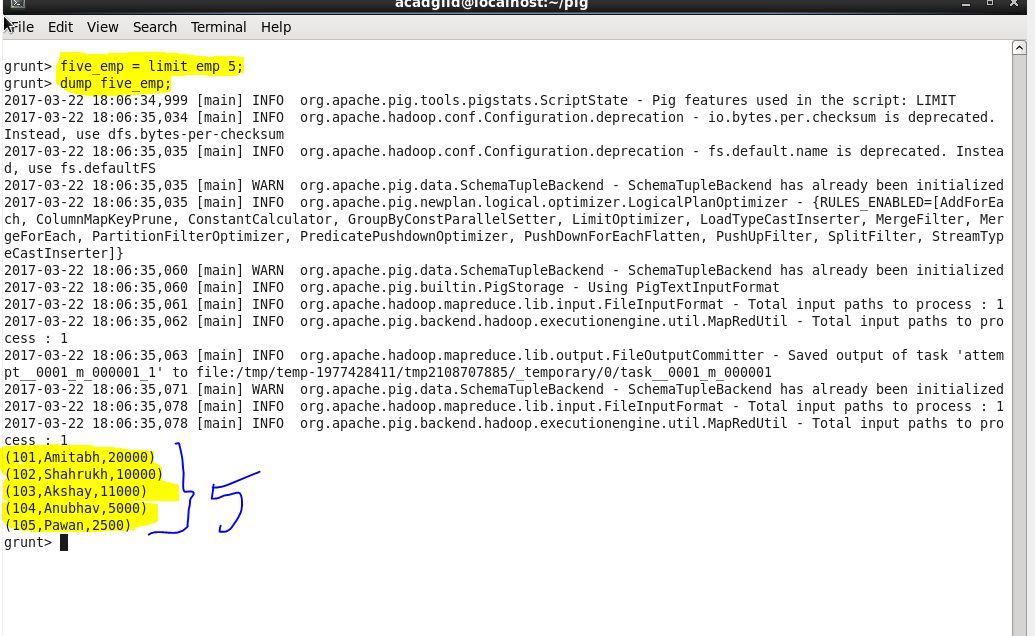


***Limit:***

* Used to limit the number of outputs to the desired number.

**Syntax**:

***Relation\_name2 = LIMIT Relation\_name1 n;***

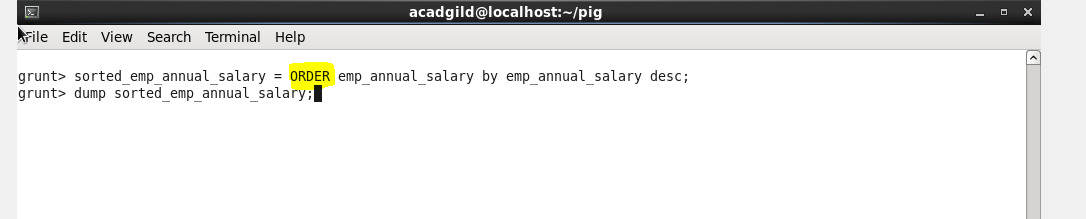
****

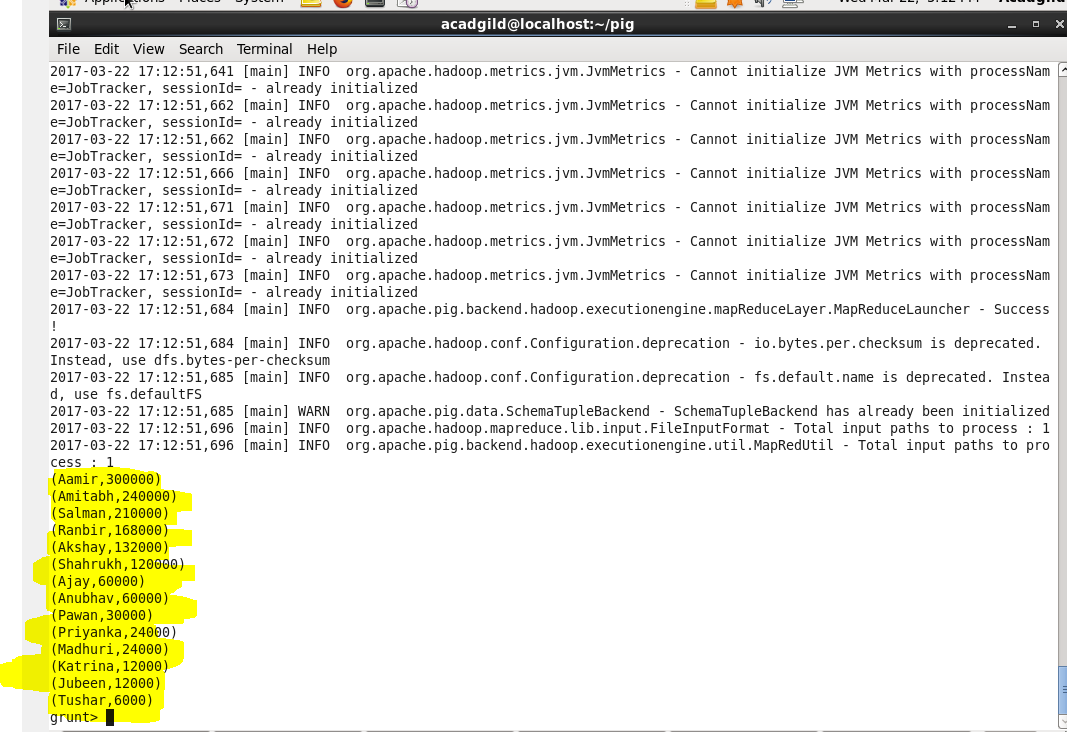
***ORDER BY:***

* The *ORDER BY* operator is used to display the contents of a relation in a sorted order based on one or more fields.

**Syntax:**

***Relation\_name2 = ORDER Relation\_name1 BY (ASC/DESC);***



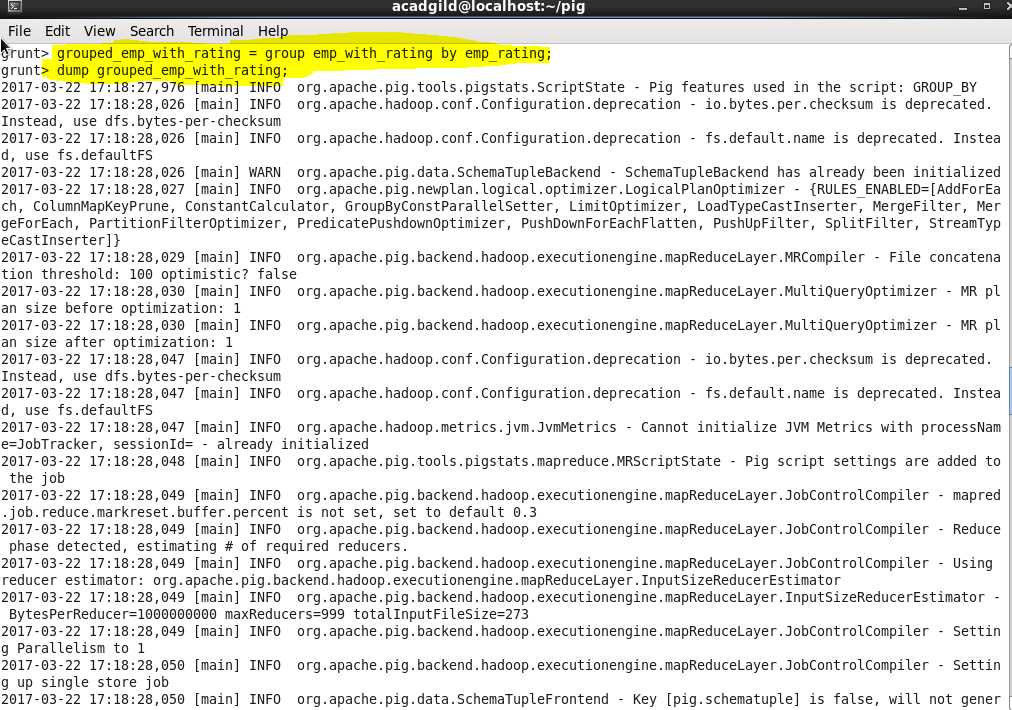


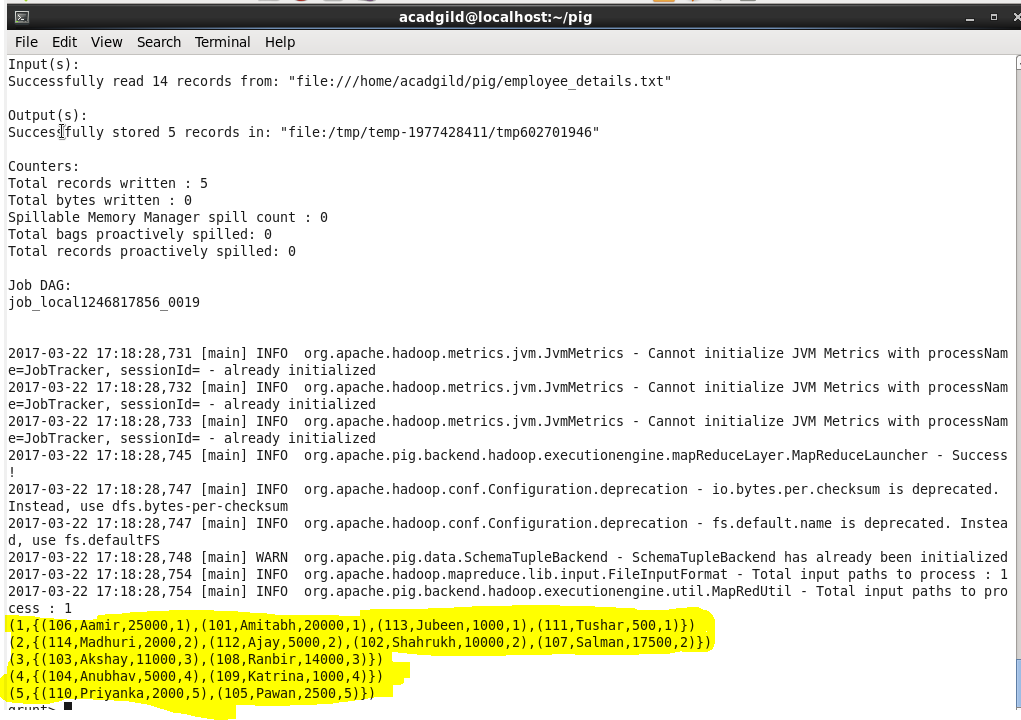
***GROUP BY:***

* The GROUP operator is used to group the data in one or more relations. It collects the data having the same key.

**Syntax:**

***Group\_data = GROUP Relation\_name by column\_name;***



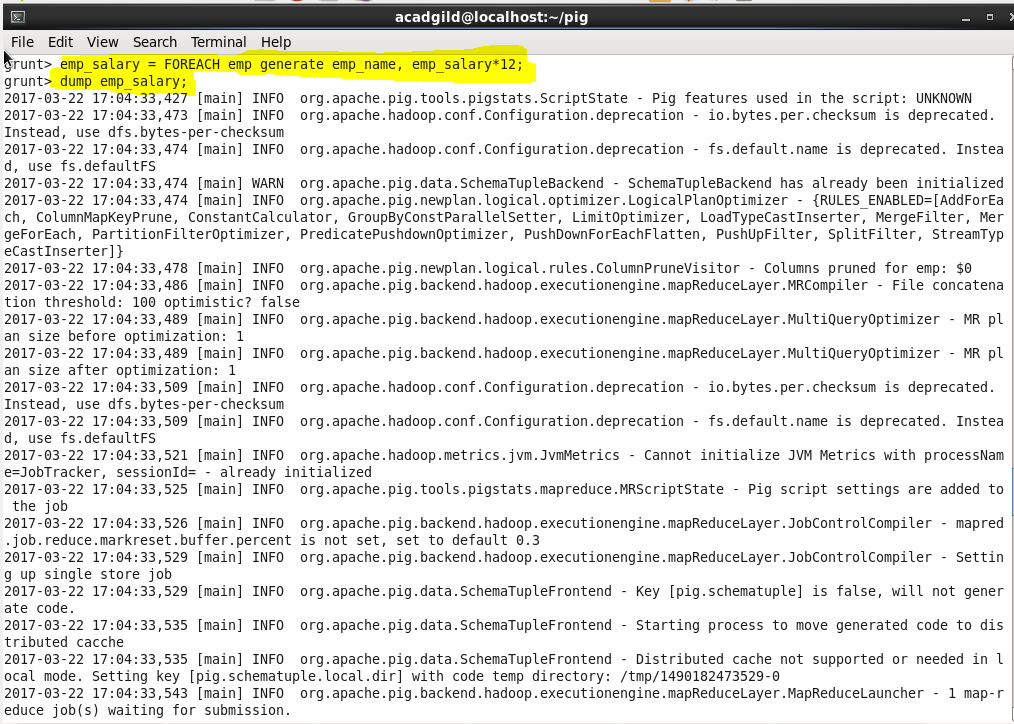


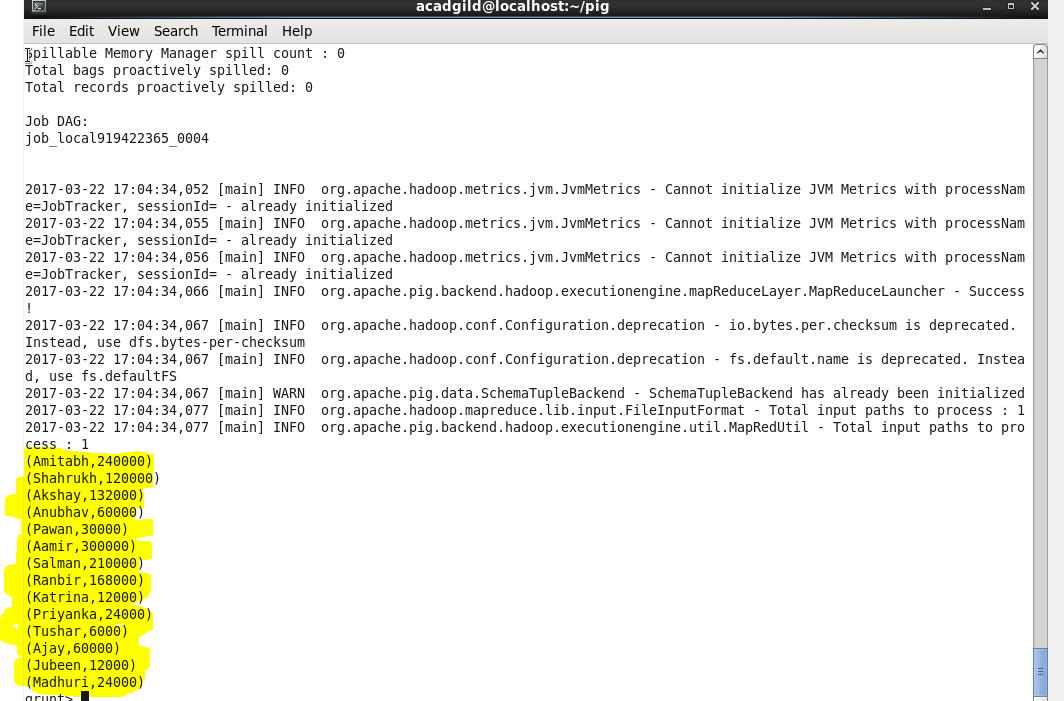
***FOREACH:***

* The *FOREACH* operator is used to generate specified data transformations based on the column data.

**Syntax:**

***Relation\_name2 = FOREACH Relation\_name1 GENERATE (required data);***



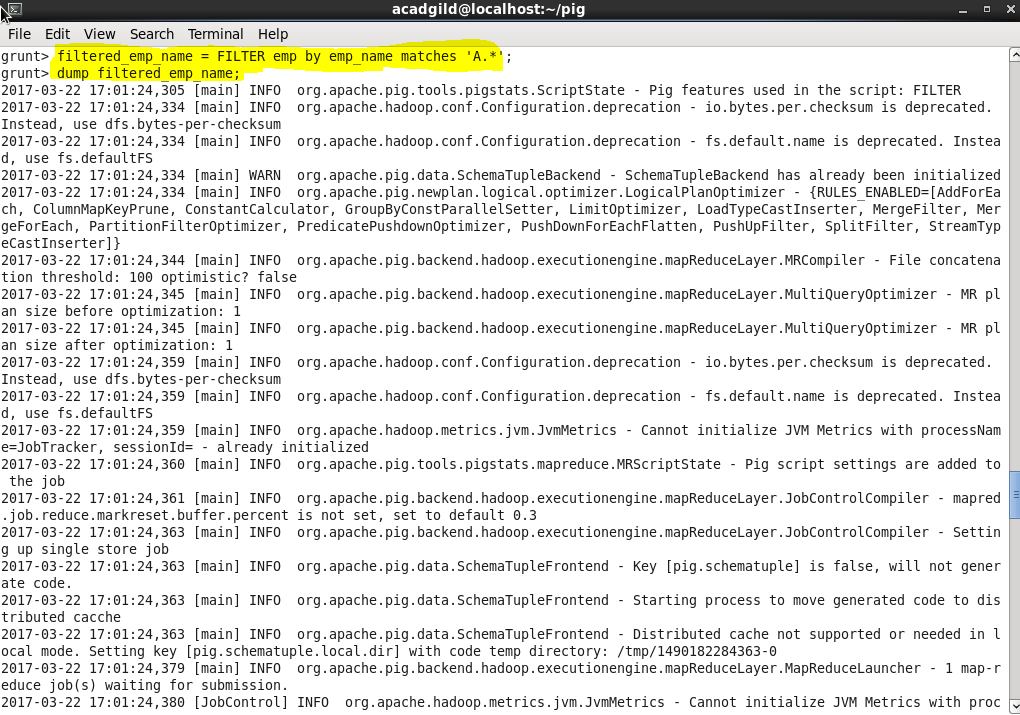


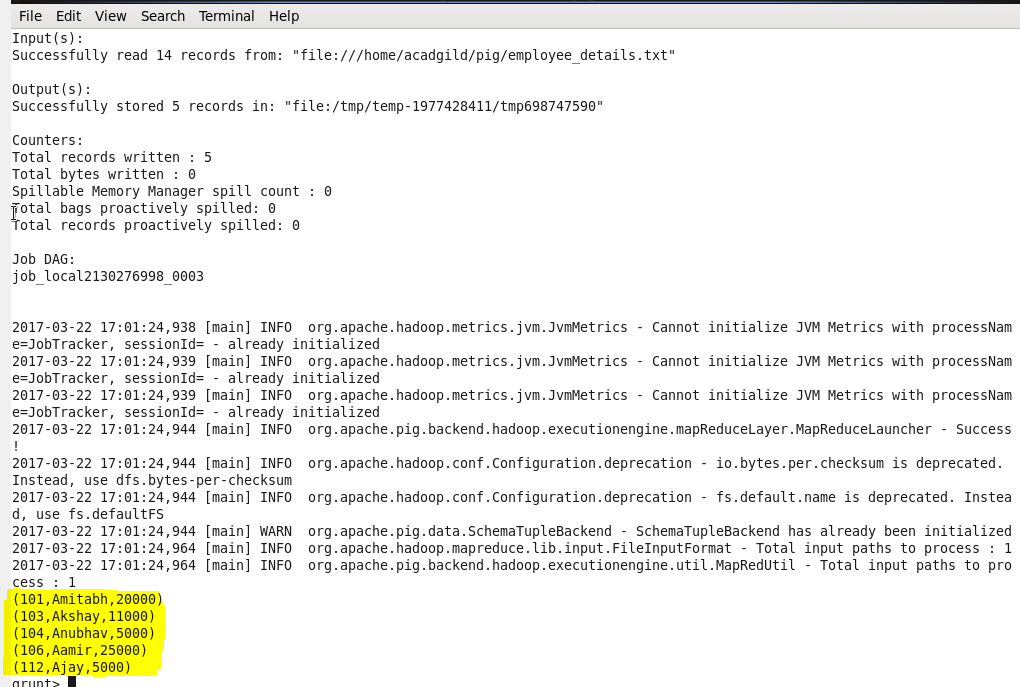
***FILTER:***

* The *FILTER* operator is used to select the required tuples from a relation based on a condition.

**Syntax:**

***Relation\_name2 = FILTER Relation\_name1 BY (condition);***





***STORE:***

* We can store the loaded data in the file system using the *store* operator.

**Syntax:**

***STORE Relation\_name INTO ‘required\_directory\_path’ [USING Function];***

