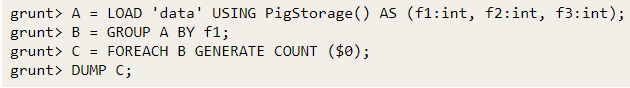
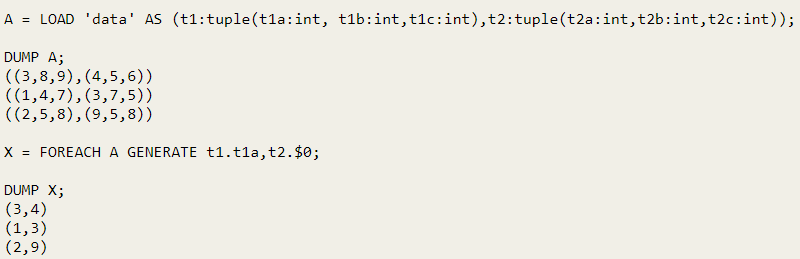
PIG

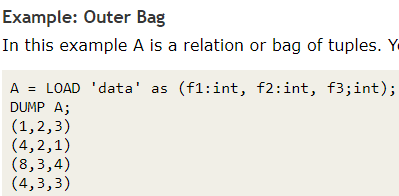
Case Senstivity

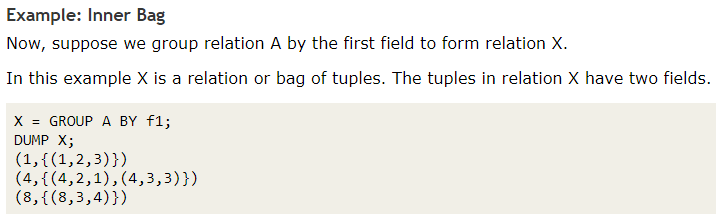
* The names (aliases) of relations A, B, and C are case sensitive.
* The names (aliases) of fields f1, f2, and f3 are case sensitive.
* Function names PigStorage and COUNT are case sensitive.
* Keywords LOAD, USING, AS, GROUP, BY, FOREACH, GENERATE, and DUMP are case insensitive. They can also be written as load, using, as, group, by, etc.

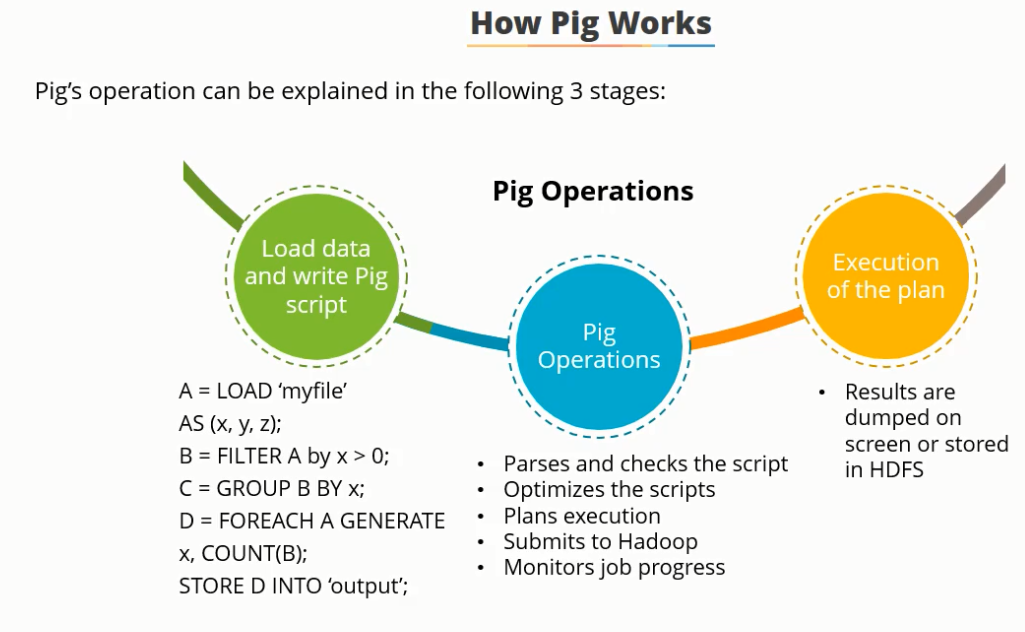


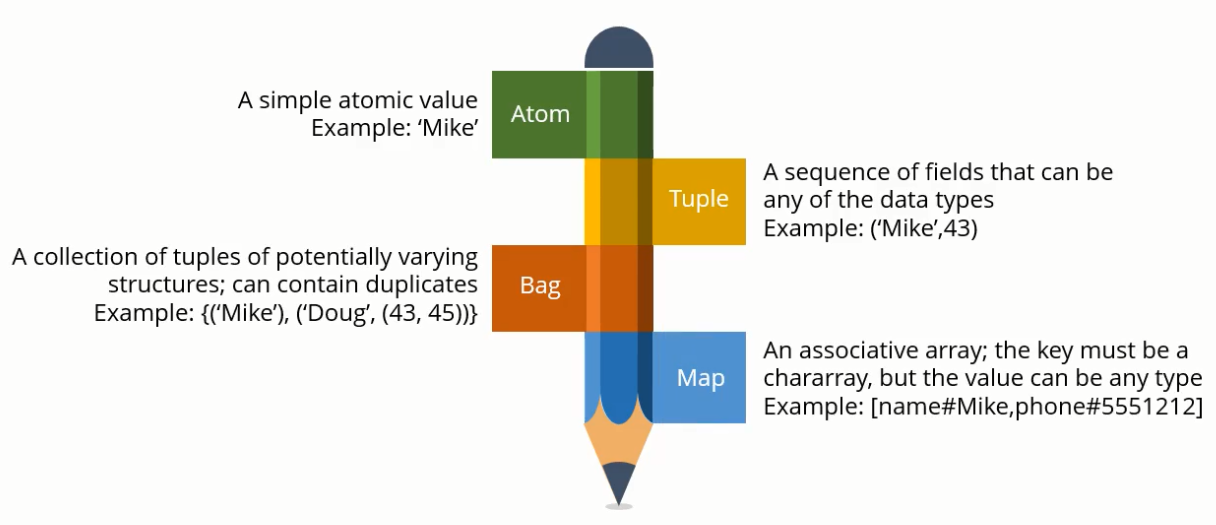
DataTypes

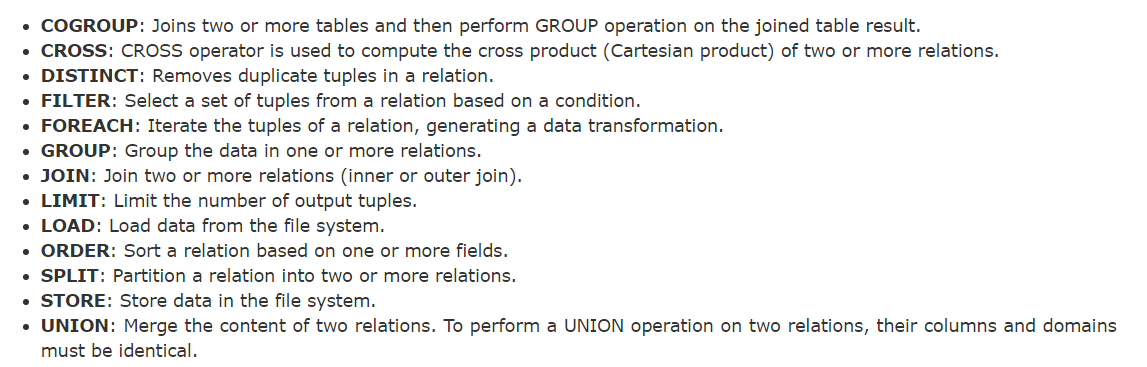


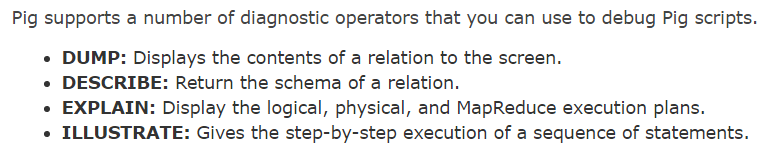








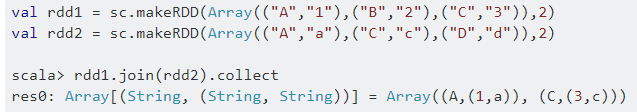


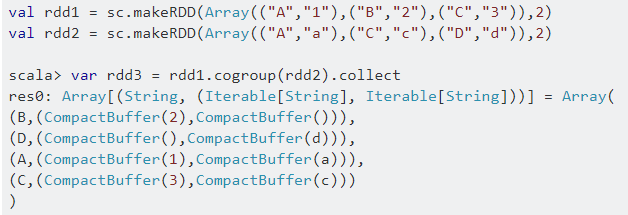


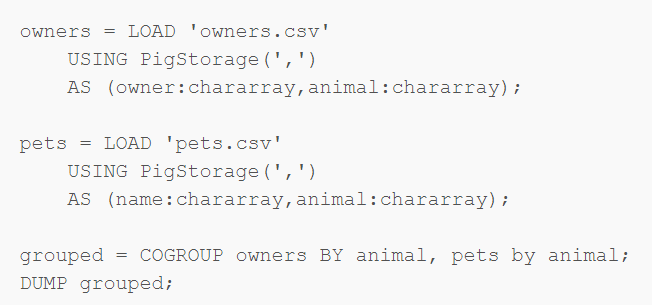
## ****The difference between GROUP and COGROUP operators in Pig?****

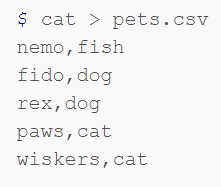
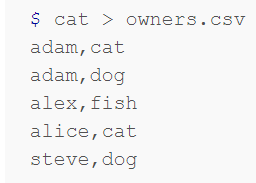
Group and Cogroup operators are identical. For readability, GROUP is used in statements involving one relation and COGROUP is used in statements involving two or more relations. Group operator collects all records with the same key. Cogroup is a combination of group and join, it is a generalization of a group instead of collecting records of one input depends on a key, it collects records of n inputs based on a key. At a time, we can Cogroup up to 127 relations.

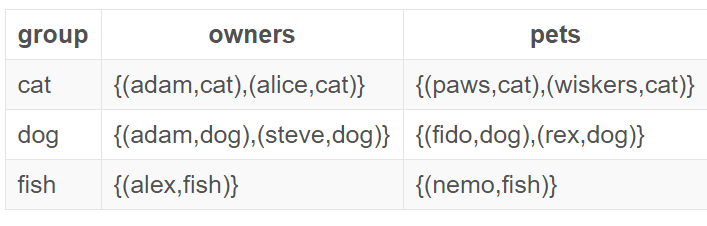
In spark a cogroup is similar to a full outer join











Interactive mode: Line by line execution and batch mode - .pig file

