Step 1: Download **Pig tar** file.

Command: wget https://downloads.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz

Step 2: Extract the **tar** file using tar command. In below tar command, **x** means extract an archive file, **z** means filter an archive through gzip, **f** means filename of an archive file.

Command: tar -xzf pig-0.16.0.tar.gz

Command: ls

Step 3: Edit the "**.bashrc**" file to update the environment variables of Apache Pig. We are setting it so that we can access pig from any directory, we need not go to pig directory to execute pig commands. Also, if any other application is looking for Pig, it will get to know the path of Apache Pig from this file.

Command: sudo vi .bashrc

Add the following at the end of the file:

 # Set PIG_HOME

export PIG_HOME=/home/edureka/pig-0.16.0

export PATH=$PATH:/home/edureka/pig-0.16.0/bin

export PIG_CLASSPATH=$HADOOP_CONF_DIR

Also, make sure that hadoop path is also set.

Run below command to make the changes get updated in same terminal.

Command: source .bashrc

Step 4: Check pig version. This is to test that Apache Pig got installed correctly. In case, you don't get the Apache Pig version, you need to verify if you have followed the above steps correctly.

Command: pig -version

Step 5: Check pig help to see all the pig command options.

Command: pig -help

Step 6: Run Pig to start the grunt shell. Grunt shell is used to run Pig Latin scripts.

Command: pig

If you look at the above image correctly, Apache Pig has two modes in which it can run, by default it chooses MapReduce mode. The other mode in which you can run Pig is Local mode. Let me tell you more about this.

Execution modes in Apache Pig:

- *MapReduce Mode* – This is the default mode, which requires access to a Hadoop cluster and HDFS installation. Since, this is a default mode, it is not necessary to specify -x flag (you can execute *pig* OR *pig -x mapreduce*). The input and output in this mode are present on HDFS.
- *Local Mode* – With access to a single machine, all files are installed and run using a local host and file system. Here the local mode is specified using '-x flag' (*pig -x local*). The input and output in this mode are present on local file system.

Command: pig -x local