In this post, I will talk about *Apache Pig installation on Linux*. Let’s start off with the basic definition of Apache Pig and Pig Latin.

*Apache Pig* is a tool/platform for creating and executing Map Reduce program used with Hadoop. It is a tool/platform for analyzing large sets of data. You can say, Apache Pig is an abstraction over MapReduce. Programmers who are not so good at Java used to struggle working on Hadoop, majorly while writing MapReduce jobs. Apache Pig has its own language *Pig Latin* which is boon for poor programmers.

**A basic introduction to Pig Latin will help you understand better:**

The high-level procedural language used in Apache Pig platform is called *Pig Latin*. Apache Pig features ‘Pig Latin’ which is a relatively simpler language which can run over distributed datasets on Hadoop File System (HDFS). In Apache Pig, you need to write Pig scripts using Pig Latin language, which gets converted to MapReduce job when your run you Pig script. Apache Pig has various operators which are used to perform the tasks like reading, writing , processing the data. To learn about Apache Pig operators, go to our blog “[***Operators in Apache Pig: Part 1- Relational Operators***](http://www.edureka.co/blog/operators-in-apache-pig/?utm_source=blog&utm_medium=content-link&utm_campaign=apache-pig-installation)”.

Now that you have basic understanding of Apache Pig, let us start with Apache Pig Installation on Linux.

**Apache Pig Installation on Linux:**

Below are the steps for Apache Pig Installation on Linux (ubuntu/centos/windows using Linux VM). I am using Ubuntu 16.04 in below setup.

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

**Command:** wget http://www-us.apache.org/dist/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 gedit .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

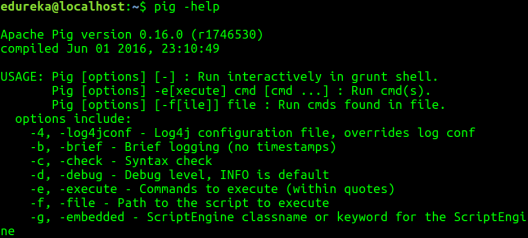
**Step 4:** Check pig version. This is to test that Apache Pig got installed correctly. In case, you don’t get the Apache Pig version, you need to verify if you have followed the above steps correctly.

**Command:** pig -version

Pig Version - Pig Installation - Edureka

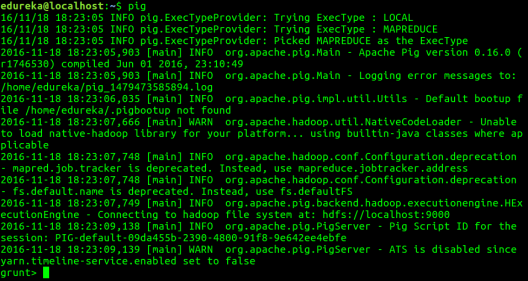
**Step 5**: Check pig help to see all the pig command options.

**Command:** pig -help



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

**Command:** pig

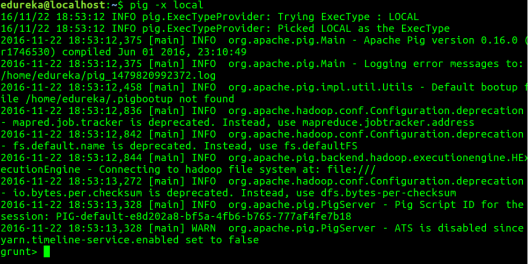


If you look at the above image correctly, Apache Pig has two modes in which it can run, by default it chooses MapReduce mode. The other mode in which you can run Pig is Local mode. Let me tell you more about this.

**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



You can go through below video to watch Apache Pig Installation on Linux:

**Apache Pig Installation | Pig Installation on Linux | Edureka**

Now that you are done with Apache Pig Installation on Linux, the next step forward is to try out some relational Pig operators on Pig Grunt shell. Hence, the next blog “[***Operators in Apache Pig: Part 1- Relational Operators***](http://www.edureka.co/blog/operators-in-apache-pig/?utm_source=blog&utm_medium=content-link&utm_campaign=apache-pig-installation)” will help you to master Pig operators.

[<< Previous Blog](https://www.edureka.co/blog/introduction-to-pig?utm_source=blog&utm_medium=blog-cta&utm_campaign=apache-pig-installation-p)