

Visit the link archive.apache.org/dist/pig/pig-0.15.0/

Click on pig-0.15.0.tar.gz

Extract the file to /home/bigdata/pig

Open the terminal by pressing the Ctrl+Alt+T keys and run the following commands:

```
$ sudo gedit /etc/profile
```

Add the following lines in the file:

```
export PIG_HOME=/home/bigdata/pig
```

```
export PATH=$PATH:$PIG_HOME/bin
```

Save and close the file

```
$ sudo gedit /home/bigdata/.bashrc
```

Add the following lines in the file:

```
export PIG_HOME=/home/bigdata/pig
```

```
export PATH=$PATH:$PIG_HOME/bin
```

Save and close the file.

```
$ touch /home/bigdata/hadoopMyFiles/user_data.csv
```

Add the following lines in the file:

```
15624510, "Male", 19, 19000, 0
```

```
15810944, "Male", 35, 20000, 0
```

```
15668575, "Female", 26, 57000, 0
```

```
15603246, "Female", 27, 43000, 0
```

```
15804002, "Male", 19, 76000, 0
```

Save and close the file

Delete the contents of /home/bigdata/hadoop/hadoop-2.4.1/temp/dfs/data

```
$ . /etc/profile
```

```
$ source /home/bigdata/.bashrc
```

```
$ hadoop namenode -format
```

```
$ start-all.sh
```

```
$ jps
```

```
$ hdfs dfs -mkdir /user
```

```
$ hdfs dfs -mkdir /hadoopMyFiles
```

```
$ hdfs dfs -put /home/bigdata/hadoopMyFiles/user_data.csv /hadoopMyFiles/user_data.csv
```

```
$ pig -version
```

```
$ pig
```

```
> user = load '/hadoopMyFiles/user_data.csv' using PigStorage(',') as (user_id:long,  
gender:chararray, age:int, estimatedsalary:long, pruchased:int);
```

```
> describe user;
```

```
> dump user;
```

```
> store user into '/user/pig_output' using PigStorage(',');
```

```
> quit;
```

```
$ hdfs dfs -cat /user/pig_output/part-m-*
```