

COPYING DATA FROM "LOCAL FILE SYSTEM" OF LINUX TO "HDFS"

In order to work in pig -x mapreduce mode

Step 1: Check directory structure of root in hdfs using below command:

```
[acadgild@localhost ~]$ hadoop fs -ls /
17/08/12 01:30:33 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
Found 5 items
-rw-r--r--    1 acadgild supergroup          87 2017-06-28 14:28
/four
drwxr-xr-x    - acadgild supergroup           0 2015-11-09 19:21
/hbasestorage
drwxrwxr-x    - acadgild supergroup           0 2015-11-05 13:46
/tmp
drwxr-xr-x    - acadgild supergroup           0 2015-11-17 01:56
/user
drwxr-xr-x    - acadgild supergroup           0 2015-11-05 12:56
/zookeeper
```

Step 2: Check whether "my_pig_stuff" directory already exists inside "/user" or not in hdfs

```
[acadgild@localhost ~]$ hadoop fs -ls /user
17/08/12 01:30:52 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
Found 3 items
drwxr-xr-x    - acadgild supergroup           0 2017-07-08 19:53
/user/acadgild
drwxr-xr-x    - acadgild supergroup           0 2015-11-05 12:52
/user/hive
drwxr-xr-x    - acadgild supergroup           0 2015-11-08 17:35
/user/prateek
"my_pig_stuff" directory does not exist inside "/user" in hdfs
```

Step 3: Create "my_pig_stuff" directory inside "/user" using following command

```
[acadgild@localhost ~]$ hadoop fs -mkdir /user/my_pig_stuff
17/08/12 01:34:02 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
```

Step 4: List "/user" directory to check whether "my_pig_stuff" directory is created or not

```
[acadgild@localhost ~]$ hadoop fs -ls /user
17/08/12 01:34:18 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
```

```

Found 4 items
drwxr-xr-x  - acadgild supergroup          0 2017-07-08 19:53
/user/acadgild
drwxr-xr-x  - acadgild supergroup          0 2015-11-05 12:52
/user/hive
drwxr-xr-x  - acadgild supergroup          0 2017-08-12 01:34
/user/my_pig_stuff
drwxr-xr-x  - acadgild supergroup          0 2015-11-08 17:35
/user/prateek
"my_pig_stuff" directory is created inside "/user" in hdfs

```

******* COPYING "students_data.txt" FROM
"/home/acadgild/Documents/pig" in LOCAL FILE SYSTEM OF
LINUX TO "/user/my_pig_stuff" in HDFS *******

**Step 5: Copy students_data.txt from local file system of linux
to hdfs using following command**

```

[acadgild@localhost ~]$ hadoop fs -put  

/home/acadgild/Documents/pig/students_data.txt  

/user/my_pig_stuff  

17/08/12 01:35:56 WARN util.NativeCodeLoader: Unable to load  

native-hadoop library for your platform... using builtin-java  

classes where applicable

```

**Step 6: Check whether data is copied successfully or not in
hdfs using following command**

```

[acadgild@localhost ~]$ hadoop fs -ls /user/my_pig_stuff  

17/08/12 01:36:20 WARN util.NativeCodeLoader: Unable to load  

native-hadoop library for your platform... using builtin-java  

classes where applicable  

Found 1 items  

-rw-r--r--  1 acadgild supergroup          141 2017-08-12 01:36  

/user/my_pig_stuff/students_data.txt  

Data is copied successfully

```

Step 7: Check contents of copied file in hdfs

```

[acadgild@localhost ~]$ hadoop fs -cat  

/user/my_pig_stuff/students_data.txt  

17/08/12 01:48:49 WARN util.NativeCodeLoader: Unable to load  

native-hadoop library for your platform... using builtin-java  

classes where applicable  

1,Amit,Sharma,86,12  

2,Sumit,Kumar,70,13  

3,Ramesh,Verma,65,11  

4,Suraj,Gupta,69,12  

5,Ravi,Verma,80,13  

6,Manav,Gupta,79,11  

7,Rajesh,Sharma,50,12  

8,Sameer,Kumar,78,13  

9,Rajan,Gupta,90,12  

10,Aman,Sharma,87,11

```

Contents are exactly copied

```
***** COPYING "duplicate_data.txt" FROM
"/home/acadgild/Documents/pig" in LOCAL FILE SYSTEM OF
LINUX TO "/user/my_pig_stuff" in HDFS *****
```

Step 8: Copy duplicate_data.txt from local file system of linux to hdfs using following command

```
[acadgild@localhost ~]$ hadoop fs -put
/home/acadgild/Documents/pig/duplicate_data.txt
/user/my_pig_stuff
17/08/12 02:40:45 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
```

Step 9: Check whether data is copied successfully or not in hdfs using following command

```
[acadgild@localhost ~]$ hadoop fs -ls /user/my_pig_stuff
17/08/12 02:41:09 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
Found 4 items
-rw-r--r-- 1 acadgild supergroup          52 2017-08-12 02:40
/user/my_pig_stuff/duplicate_data.txt
-rw-r--r-- 1 acadgild supergroup        204 2017-08-12 01:48
/user/my_pig_stuff/students_data.txt
Data is copied successfully
```

Step 10: Check contents of copied file in hdfs

```
[acadgild@localhost ~]$ hadoop fs -cat
/user/my_pig_stuff/duplicate_data.txt
17/08/12 02:41:31 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
1,Aman,50
2,Suraj,60
3,Ravi,70
1,Aman,50
2,Suraj,60
Contents are copied correctly
```

```
***** COPYING "empl_data.txt" FROM
"/home/acadgild/Documents/pig" in LOCAL FILE SYSTEM OF
LINUX TO "/user/my_pig_stuff" in HDFS *****
```

Step 11: Copy empl_data.txt from local file system of linux to hdfs using following command

```
[acadgild@localhost ~]$ hadoop fs -put
/home/acadgild/Documents/pig/empl.txt /user/my_pig_stuff
17/08/12 03:18:40 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
```

Step 12: Check whether data is copied successfully or not in hdfs using following command

```
[acadgild@localhost ~]$ hadoop fs -ls /user/my_pig_stuff
17/08/12 03:19:48 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
Found 6 items
-rw-r--r--   1 acadgild supergroup          52 2017-08-12 02:40
/user/my_pig_stuff/duplicate_data.txt
-rw-r--r--   1 acadgild supergroup          94 2017-08-12 03:18
/user/my_pig_stuff/empl.txt
-rw-r--r--   1 acadgild supergroup        204 2017-08-12 01:48
/user/my_pig_stuff/students_data.txt
Data is copied successfully
```

Step 13: Check contents of copied file in hdfs

```
[acadgild@localhost ~]$ hadoop fs -cat
/user/my_pig_stuff/empl.txt17
/08/12 03:18:59 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
101,Amit,2000,sales,30
102,Suraj,3000,sales,35
103,Raman,4000,sales,30
104,Ravi,3000,sales,30
Contents are correct
```

******* COPYING "emp2_data.txt" FROM
"/home/acadgild/Documents/pig" in LOCAL FILE SYSTEM OF
LINUX TO "/user/my_pig_stuff" in HDFS *******

Step 14: Copy emp2_data.txt from local file system of linux to hdfs using following command

```
[acadgild@localhost ~]$ hadoop fs -put
/home/acadgild/Documents/pig/emp2.txt /user/my_pig_stuff
17/08/12 03:19:20 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
```

Step 15: Check whether data is copied successfully or not in hdfs using following command

```
[acadgild@localhost ~]$ hadoop fs -ls /user/my_pig_stuff
17/08/12 03:19:48 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
```

```
classes where applicable
Found 6 items
-rw-r--r--  1 acadgild supergroup          52 2017-08-12 02:40
/user/my_pig_stuff/duplicate_data.txt
-rw-r--r--  1 acadgild supergroup          94 2017-08-12 03:18
/user/my_pig_stuff/empl.txt
-rw-r--r--  1 acadgild supergroup          98 2017-08-12 03:19
/user/my_pig_stuff/emp2.txt
-rw-r--r--  1 acadgild supergroup        204 2017-08-12 01:48
/user/my_pig_stuff/students_data.txt
Data is copied successfully
```

Step 16: Check contents of copied file in hdfs

```
[acadgild@localhost ~]$ hadoop fs -cat
/user/my_pig_stuff/emp2.txt
17/08/12 03:19:34 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java
classes where applicable
101,Amit,4000,35,sales
102,Rajeev,3500,30,admin
103,Ravijot,4500,40,sales
104,Ratan,5000,42,admin
Contents are correct
```