

<<<<----- Creation of DATASET ----->>>> IN LOCAL FILE SYSTEM OF LINUX

Step 1:- CD to /home/acadgild/Documents which is equivalent to /\$HOME/Documents

```
[acadgild@localhost ~]$ cd /$HOME/Documents
```

Step 2:- Check present working directory

```
[acadgild@localhost Documents]$ pwd  
/home/acadgild/Documents
```

Step 3:- List present working directory to check whether "pig" directory is present inside /home/acadgild/Documents or not

```
[acadgild@localhost Documents]$ ls -lrt  
total 84  
-rw-r--r--. 1 acadgild acadgild 87 Jun 28 14:45 four_copied  
-rw-rw-r--. 1 acadgild acadgild 24709 Jun 28 14:50  
first_access_to_commands  
-rw-rw-r--. 1 acadgild acadgild 71 Jun 28 14:56  
students.txt  
-rw-rw-r--. 1 acadgild acadgild 69 Jun 28 15:04  
employees.txt  
-rw-rw-r--. 1 acadgild acadgild 4054 Jun 28 16:58 Session2  
_Assignment1  
-rw-rw-r--. 1 acadgild acadgild 2359 Jun 28 17:24 Session2  
_Assignment2  
-rw-rw-r--. 1 acadgild acadgild 24138 Jun 29 19:36  
map_reduce_execution1  
drwxrwxr-x. 2 acadgild acadgild 4096 Jul 4 17:17  
MapReduceWork  
drwxrwxr-x. 4 acadgild acadgild 4096 Jul 6 15:23 mydata  
drwxrwxr-x. 9 acadgild acadgild 4096 Jul 8 19:23  
SharedDatafromWindows
```

pig directory is not present inside current working directory

Step 4:- Make directory "pig" inside /home/acadgild/Documents

```
[acadgild@localhost Documents]$ mkdir pig
```

Step 5:- List present working directory to check whether "pig" directory is created inside /home/acadgild/Documents or not

```
[acadgild@localhost Documents]$ ls -lrt  
total 88  
-rw-r--r--. 1 acadgild acadgild 87 Jun 28 14:45 four_copied  
-rw-rw-r--. 1 acadgild acadgild 24709 Jun 28 14:50  
first_access_to_commands  
-rw-rw-r--. 1 acadgild acadgild 71 Jun 28 14:56  
students.txt  
-rw-rw-r--. 1 acadgild acadgild 69 Jun 28 15:04  
employees.txt
```

```
-rw-rw-r--. 1 acadgild acadgild 4054 Jun 28 16:58 Session2_Assignment1
-rw-rw-r--. 1 acadgild acadgild 2359 Jun 28 17:24 Session2_Assignment2
-rw-rw-r--. 1 acadgild acadgild 24138 Jun 29 19:36 map_reduce_execution1
drwxrwxr-x. 2 acadgild acadgild 4096 Jul 4 17:17 MapReduceWork
drwxrwxr-x. 4 acadgild acadgild 4096 Jul 6 15:23 mydata
drwxrwxr-x. 9 acadgild acadgild 4096 Jul 8 19:23 SharedDatafromWindows
drwxrwxr-x. 2 acadgild acadgild 4096 Aug 12 01:23 pig
pig directory is created
```

Step 6:- CD to "pig" directory

```
[acadgild@localhost Documents]$ cd pig
```

Step 7:- Check present working directory

```
[acadgild@localhost pig]$ pwd
/home/acadgild/Documents/pig
```

<<<<<<----- CREATION OF "students_data.txt" ----->>>>>>
USING vi editor

Step 8:- To create students_data.txt in vi type following command

```
[acadgild@localhost pig]$ vi students_data.txt
```

Step 9:- After saving file in vi, check contents of file using cat command as follows:

```
[acadgild@localhost pig]$ cat students_data.txt
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
```

<<<<<<----- CREATION OF "duplicate_data.txt" ----->>>>>>
USING vi editor

Step 10:- To create duplicate_data.txt in vi type following command

```
[acadgild@localhost pig]$ vi duplicate_data.txt
```

Step 11:- After saving file in vi, check contents of file

using cat command as follows:

```
[acadgild@localhost pig]$ cat duplicate_data.txt
1,Aman,50
2,Suraj,60
3,Ravi,70
1,Aman,50
2,Suraj,60
```

**<<<<<----- CREATION OF "emp1_data.txt" ----->>>>>
USING vi editor**

Step 12:- To create emp1_data.txt in vi type following command

```
[acadgild@localhost pig]$ vi emp1.txt
```

Step 13:- After saving file in vi, check contents of file using cat command as follows:

```
[acadgild@localhost pig]$ cat emp1.txt
101,Amit,2000,sales,30
102,Suraj,3000,sales,35
103,Raman,4000,sales,30
104,Ravi,3000,sales,30
```

**<<<<<----- CREATION OF "emp2_data.txt" ----->>>>>
USING vi editor**

Step 14:- To create emp2_data.txt in vi type following command

```
[acadgild@localhost pig]$ vi emp2.txt
```

Step 15:- After saving file in vi, check contents of file using cat command as follows:

```
[acadgild@localhost pig]$ cat emp2.txt
101,Amit,4000,35,sales
102,Rajeev,3500,30,admin
103,Ravijot,4500,40,sales
104,Ratan,5000,42,admin
[acadgild@localhost pig]$
```