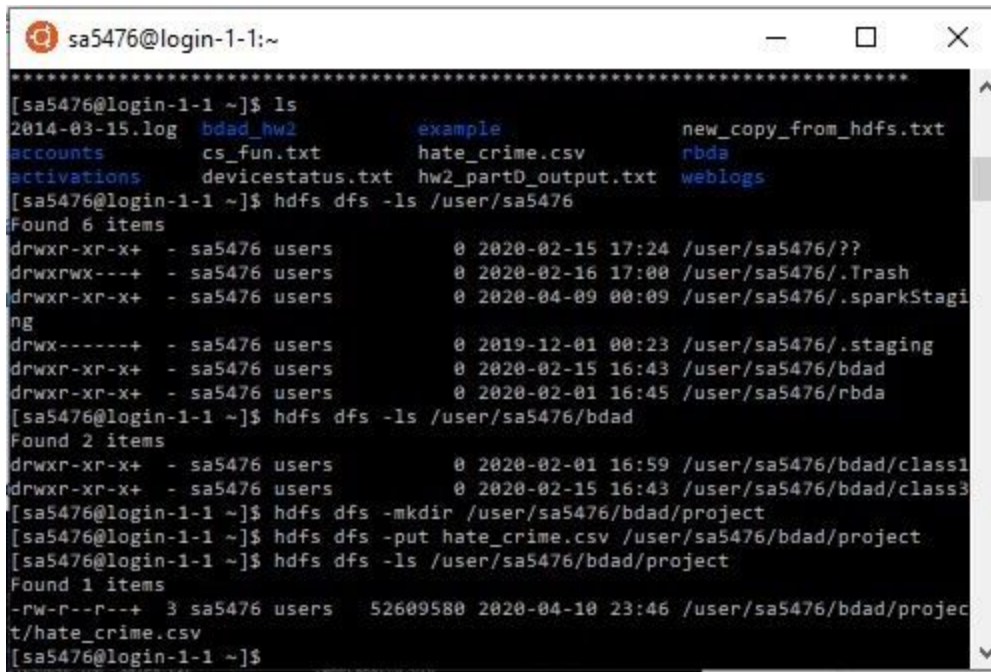


Data Ingestion

Steps to ingest data:

1. Downloaded WinSCP Software from NYU Data Services Lab
 - <https://guides.nyu.edu/software>
 - <https://winscp.net/eng/download.php>
2. Install and Open installed WinSCP Application
3. Use IP address of dumbo to login to WinSCP
4. Drag and drop data from local to dumbo.
5. Check in the dumbo terminal for the required folder using -ls command.



```
sa5476@login-1-1:~  
[sa5476@login-1-1 ~]$ ls  
2014-03-15.log bdad_hw2 example new_copy_from_hdfs.txt  
accounts cs_fun.txt hate_crime.csv rbda  
activations devicestatus.txt hw2_partD_output.txt weblogs  
[sa5476@login-1-1 ~]$ hdfs dfs -ls /user/sa5476  
Found 6 items  
drwxr-xr-x+ - sa5476 users 0 2020-02-15 17:24 /user/sa5476/??  
drwxrwx---+ - sa5476 users 0 2020-02-16 17:00 /user/sa5476/.Trash  
drwxr-xr-x+ - sa5476 users 0 2020-04-09 00:09 /user/sa5476/.sparkStagi  
ng  
drwx-----+ - sa5476 users 0 2019-12-01 00:23 /user/sa5476/.staging  
drwxr-xr-x+ - sa5476 users 0 2020-02-15 16:43 /user/sa5476/bdad  
drwxr-xr-x+ - sa5476 users 0 2020-02-01 16:45 /user/sa5476/rbda  
[sa5476@login-1-1 ~]$ hdfs dfs -ls /user/sa5476/bdad  
Found 2 items  
drwxr-xr-x+ - sa5476 users 0 2020-02-01 16:59 /user/sa5476/bdad/class1  
drwxr-xr-x+ - sa5476 users 0 2020-02-15 16:43 /user/sa5476/bdad/class3  
[sa5476@login-1-1 ~]$ hdfs dfs -mkdir /user/sa5476/bdad/project  
[sa5476@login-1-1 ~]$ hdfs dfs -put hate_crime.csv /user/sa5476/bdad/project  
[sa5476@login-1-1 ~]$ hdfs dfs -ls /user/sa5476/bdad/project  
Found 1 items  
-rw-r--r--+ 3 sa5476 users 52609580 2020-04-10 23:46 /user/sa5476/bdad/projec  
t/hate_crime.csv  
[sa5476@login-1-1 ~]$
```

6. Folder "hate_crime.csv" is present.
7. Use the following commands to make folder project_01 in hdfs and ingest this data.:
hdfs dfs -mkdir /user/sa5476/bdad/project
hdfs dfs -put hate_crime.csv /user/sa5476/bdad/project
hdfs dfs -cat /user/sa5476/bdad/project/hate_crime.csv

8. Snippet of the dataset.

```
sa5476@login-1-1:~  
,"South","6","Cities from 2,500 thru 9,999",05-OCT-18,0,0,2,2,0,"White","Unknown",  
,"1","Drug/Narcotic Violations",0,"Highway/Road/Alley/Street/Sidewalk","Anti-Lesb  
ian (Female)","Society/Public","S","S"  
466130,2018,"WV0540100","Parkersburg","",,"City","WV","West Virginia","South Atla  
ntic","South","4","Cities from 25,000 thru 49,999",14-DEC-18,1,0,0,,,"Unknown",  
Unknown",1,"Burglary/Breaking & Entering",1,"Residence/Home","Anti-Black or Afri  
can American","Individual","S","S"  
466159,2018,"WV0540100","Parkersburg","",,"City","WV","West Virginia","South Atla  
ntic","South","4","Cities from 25,000 thru 49,999",14-NOV-18,1,0,1,0,1,"White",  
Unknown",1,"Simple Assault",1,"School/College","Anti-Black or African American",  
,"Individual","S","S"  
508677,2018,"WV0540200","Vienna","",,"City","WV","West Virginia","South Atlantic",  
,"South","5","Cities from 10,000 thru 24,999",26-NOV-18,1,0,0,,,"Unknown","Unkno  
wn",1,"Intimidation",1,"Highway/Road/Alley/Street/Sidewalk","Anti-Asian","Indivi  
dual","S","S"  
463503,2018,"WVWSP2400","State Police:",,"Madison","State Police","WV","West Virg  
inia","South Atlantic","South","9D","MSA counties under 10,000",31-JAN-18,2,0,1,  
1,0,"White","Unknown",1,"Intimidation",2,"Residence/Home","Anti-White","Law Enfo  
rcement Officer","S","S"  
544359,2018,"WVWSP3300","State Police:",,"Parkersburg","State Police","WV","West  
Virginia","South Atlantic","South","9D","MSA counties under 10,000",21-JAN-18,0,  
0,0,,,"Unknown","Unknown",2,"Burglary/Breaking & Entering;Destruction/Damage/Van  
dalism of Property",0,"Church/Synagogue/Temple/Mosque","Anti-Other Religion","Re  
ligious Organization","M","S"  
[sa5476@login-1-1 ~]$
```