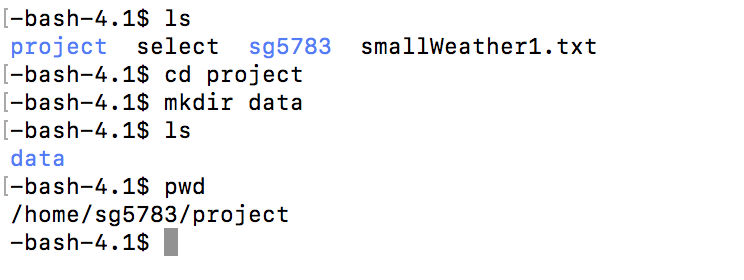
**(I) Dataset Name: Fire Department Calls for Service**

**Dataset source:** <https://data.sfgov.org/Public-Safety/Fire-Department-Calls-for-Service/nuek-vuh3/data>

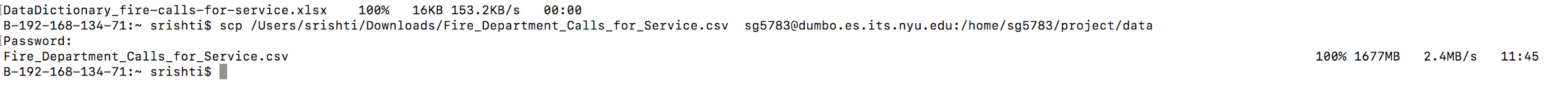
**Steps For Data Ingestion:**

1. Creation of a project data folder in remote machine (HPC server) where data needs to be transferred :



2. Transferring data from local machine to remote machine (HPC server) using scp :

scp /Users/srishti/Downloads/Fire\_Department\_Calls\_for\_Service.csv sg5783@dumbo.es.its.nyu.edu:/home/sg5783/project/data

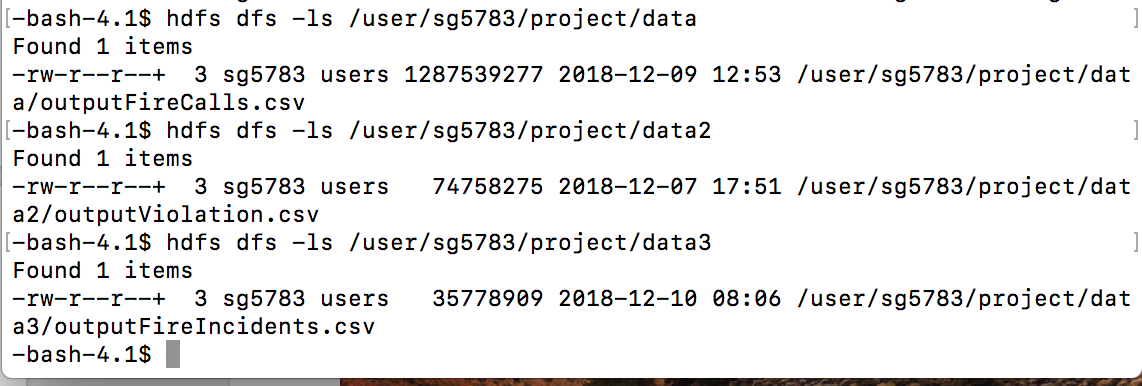




3. Storing data on HDFS using -put command:



Final project data folders on HDFS look like :



**(II) Dataset Name: Notices of Violation issued by the Department of Building Inspection**

**Dataset source:** <https://data.sfgov.org/Housing-and-Buildings/Notices-of-Violation-issued-by-the-Department-of-B/nbtm-fbw5/data>

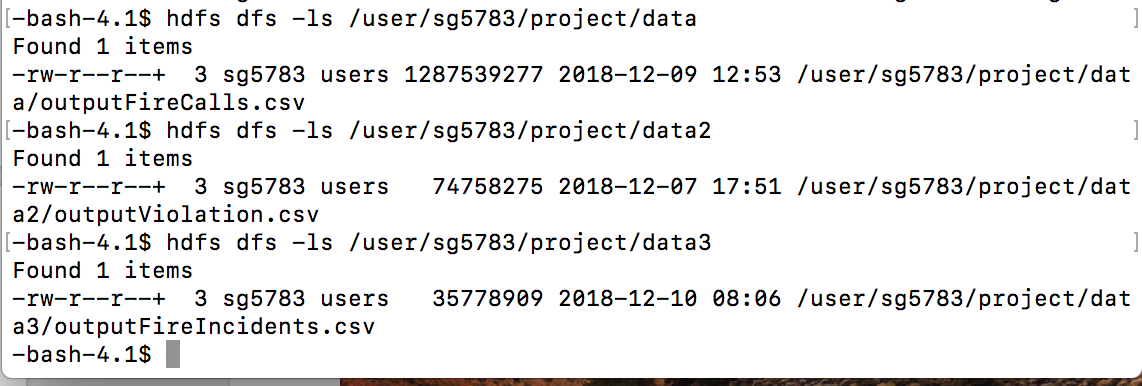
**Steps For Data Ingestion:**

1. Transferring data from local machine to remote machine (HPC server) using scp :

scp /Users/srishti/Downloads/outputViolations.csv sg5783@dumbo.es.its.nyu.edu:/home/sg5783/project/data



2. Storing data on HDFS using -put command:



**(III) Dataset Name: Fire Incidents**

**Dataset source:**

<https://data.sfgov.org/Public-Safety/Fire-Incidents/wr8u-xric/data>

**Steps For Data Ingestion:**

1. Transferring data from local machine to remote machine (HPC server) using scp :

scp /Users/srishti/Downloads/outputFireIncidents.csv sg5783@dumbo.es.its.nyu.edu:/home/sg5783/project/data



2. Storing data on HDFS using -put command:

