

Step 1) Data is downloaded from:

https://data.cityofnewyork.us/Transportation/VZV_Speed-Limits/7n5j-865y

The dataset is a total of 350 MBs.

Step 2) Run the following command in the directory downloaded dataset is kept:

```
scp ./speedlimit.csv bm2515@dumbo.es.its.nyu.edu:/home/bm2515
speedlimit.csv          100% 22MB 7.5MB/s 00:02
```

Step 3) Make a directory to move dataset in HDFS

Run Command:

```
Hdfs dfs -mkdir /user/bm2515/myHDFSFolder
```

Step 4) Move data from Local File system to HDFS by using CopyFromLocal Command:

```
Hdfs dfs -CopyFromLocal /user/bm2515/speedlimit.csv /user/bm2515/myHDFSFolder
```

Finally, data has been moved from local to the HDFS and can be used to run MR frameworks.