scp /Users/emmayang/Desktop/Weather.csv yy1316@dumbo.es.its.nyu.edu:/home/yy1316/Project

hdfs dfs -put Weather.csv /user/yy1316/Project

scp -r /Users/emmayang/Citibike yy1316@dumbo.es.its.nyu.edu:/home/yy1316/Project

hdfs dfs -put Citibike /user/yy1316/Project

scp -r /Users/emmayang/Downloads/yello\_taxi yy1316@dumbo.es.its.nyu.edu:/home/yy1316/Project

hdfs dfs -put yello\_taxi /user/yy1316/Project

Using scp to upload the data to server and then ingests the data into hdfs through a command:

Upload the data to server:

scp PROJECT xl1638@dumbo.es.its.nyu.edu

Ingests the data into hdfs:

Hfs dfs -put PROJECT

Please note that all the data is under the PROJECT folder.

I have performed a cleaning using local cleaning tasks, please check data cleaning and data schema for details. Files has been ingested to HDFS via the following steps:

1. scp the entire project file to dumbo.

2. Use dumbo to put the project file on HDFS.

command:

> scp PROJECT jl8456@dumbo.es.its.nyu.edu

> hdfs dfs -put PROJECT