

Steps to extract tweets and then store in HDFS:

1. Run program tweet\_extraction.py
2. This will create saved\_tweets.csv file.
3. Download 'Fugu'
4. Follow the instruction <https://wikis.nyu.edu/display/NYUHPC/Fugu+-+SFTP+for+Mac+OS>
5. Then drag and drop the file to Dumbo.
6. Follow the below commands:  
**hdfs dfs -ls /**  
**hdfs dfs -ls /user**  
**hdfs dfs -ls /user/sb6632**  
**hdfs dfs -mkdir /user/sb6632/project**  
**hdfs dfs -put saved\_tweets.csv /user/sb6632/project**