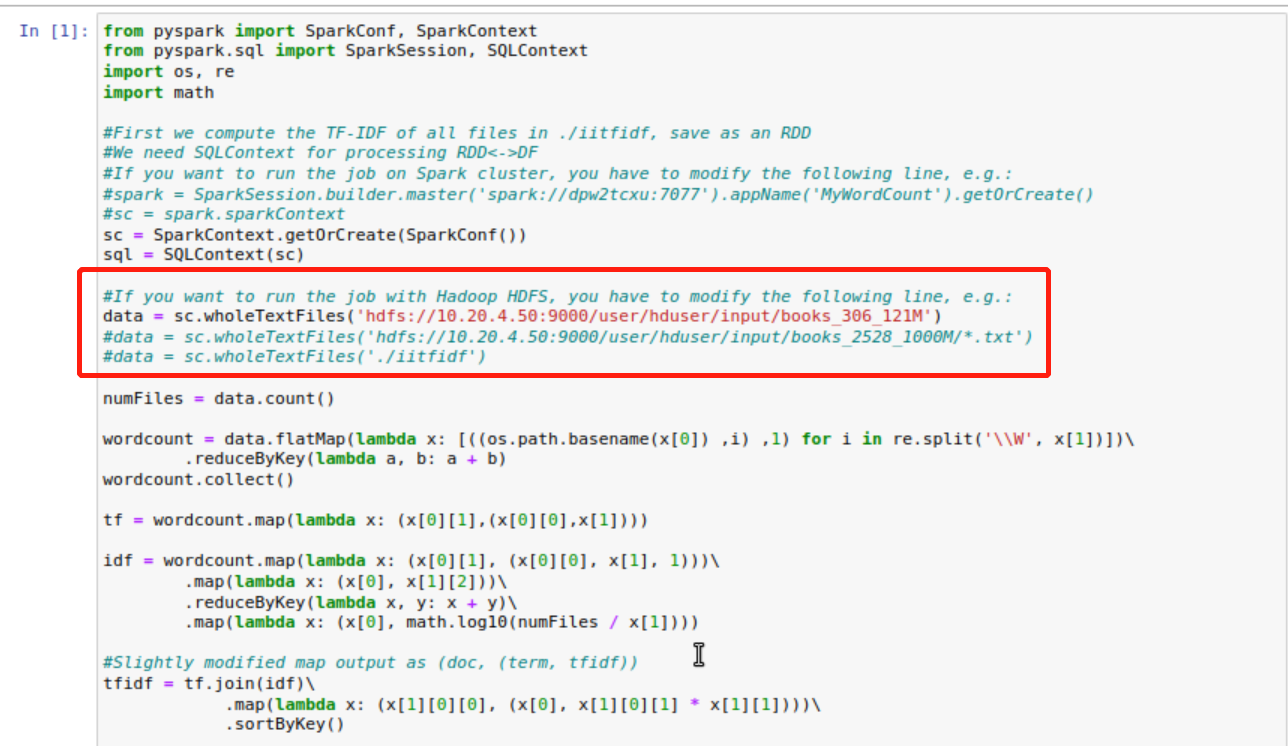
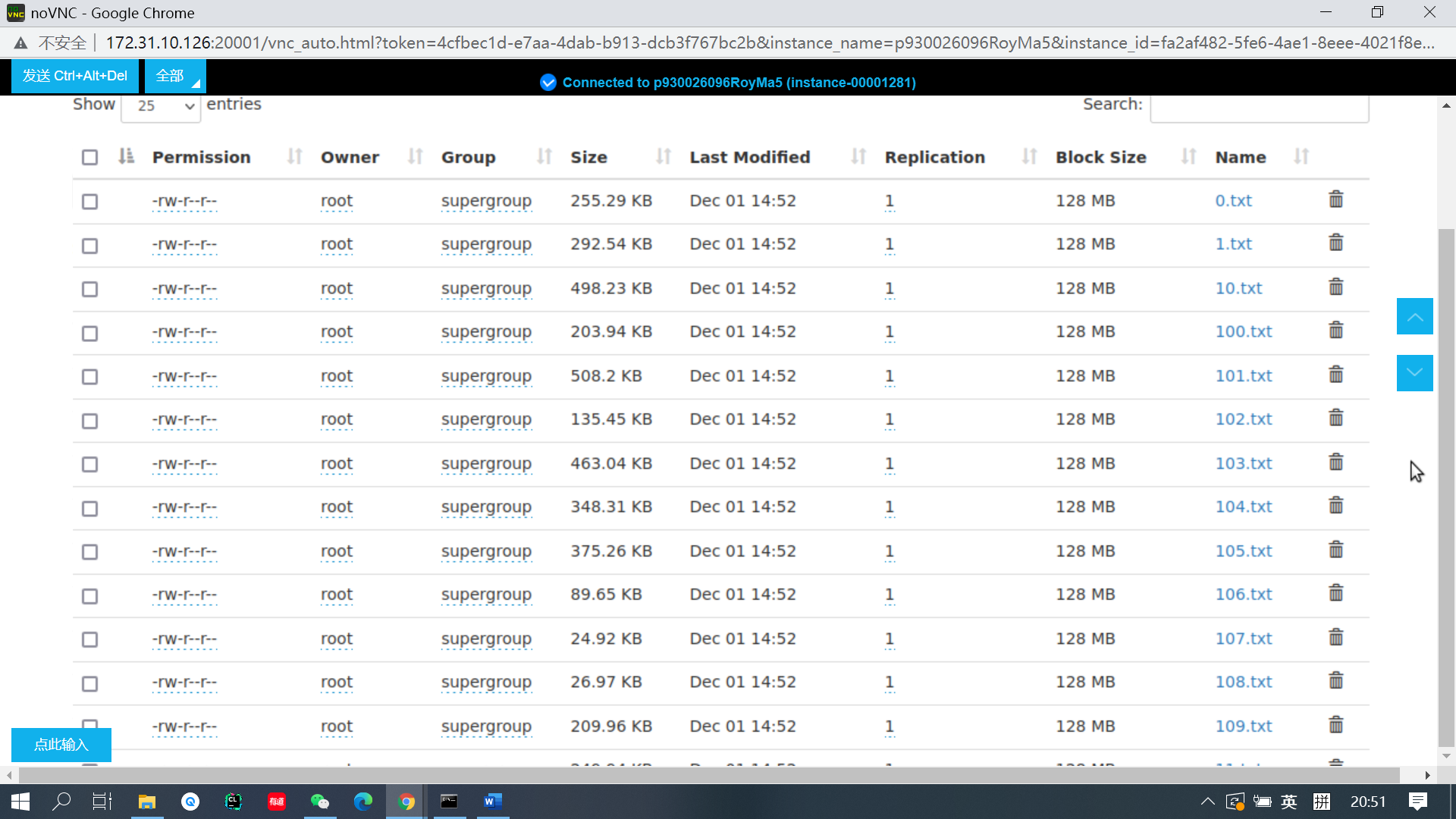
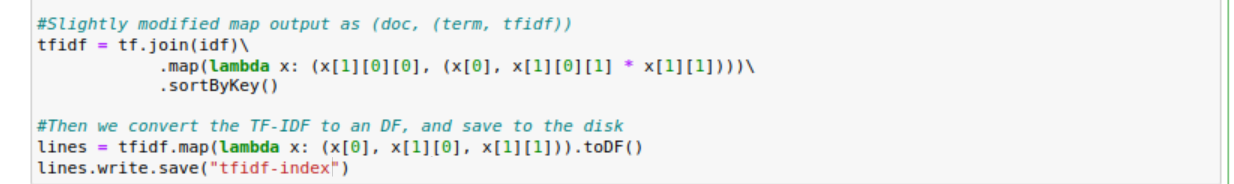
**Intermediate**



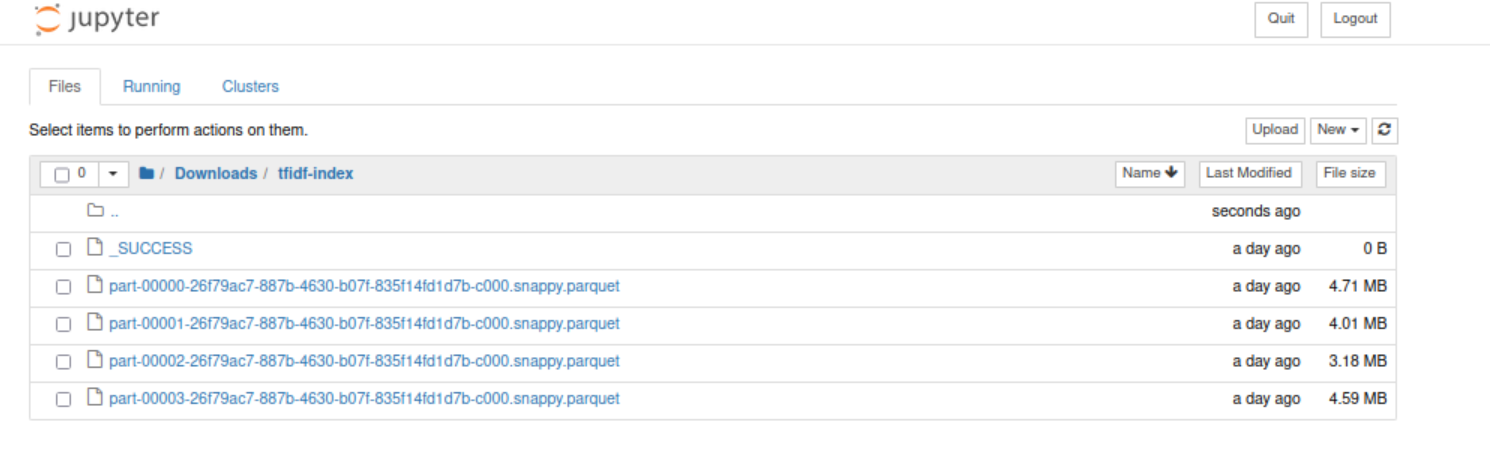
**Change the directory to own HDFS server and based on the documents in HDFS.**



**Here is the files in our own HDFS**



**Calculate the tf-idf and save the files in a file directory**



**Here is the files that generated by the codes**