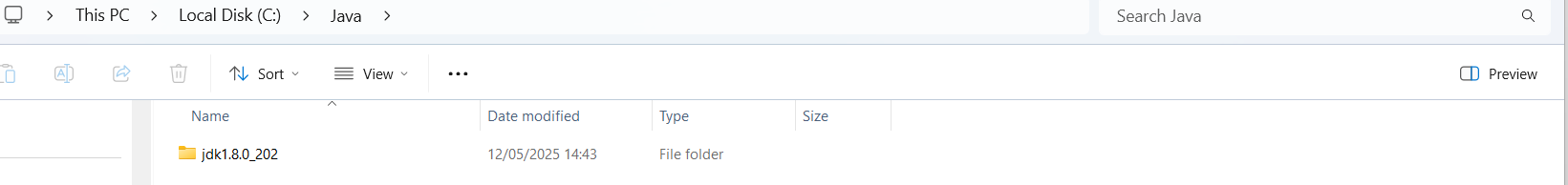
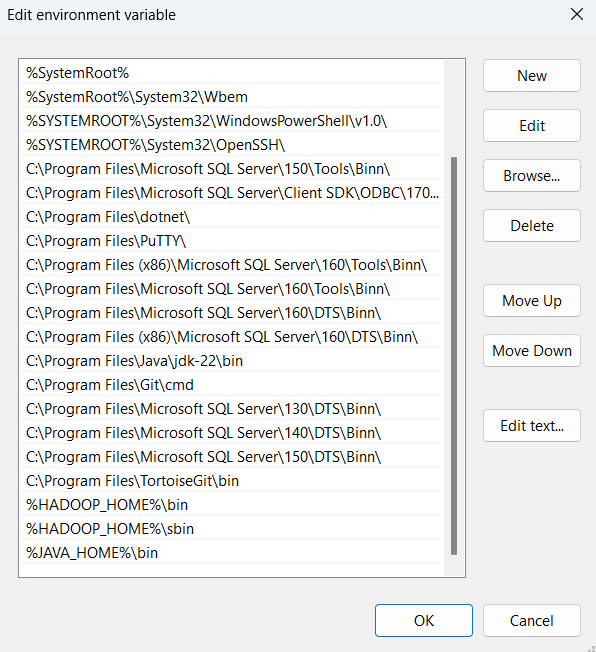
CÀI ĐẶT HADOOP ECOSYSTEM TRÊN WINDOW

Cài đặt Java 8

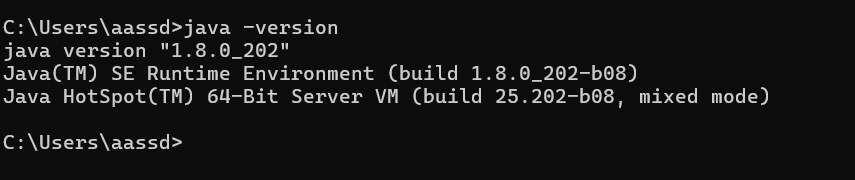


Thiết lập biến môi trường



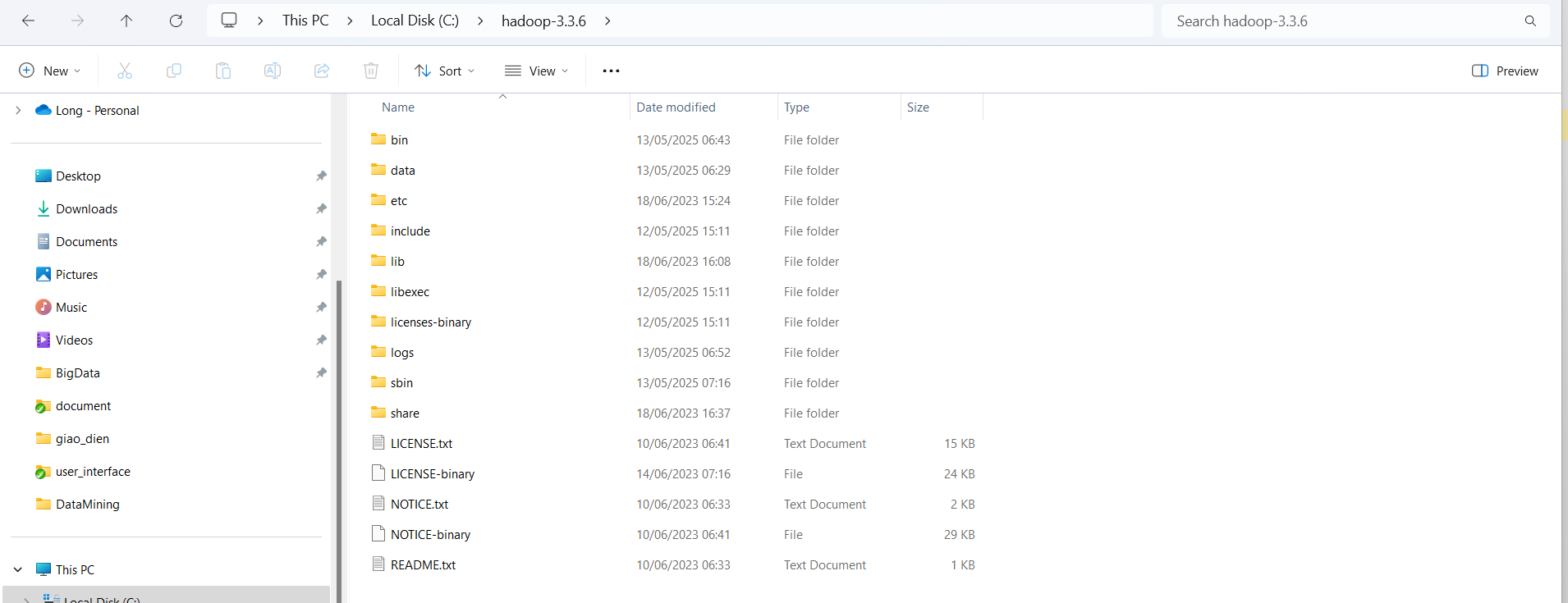


Kiểm tra phiên bản



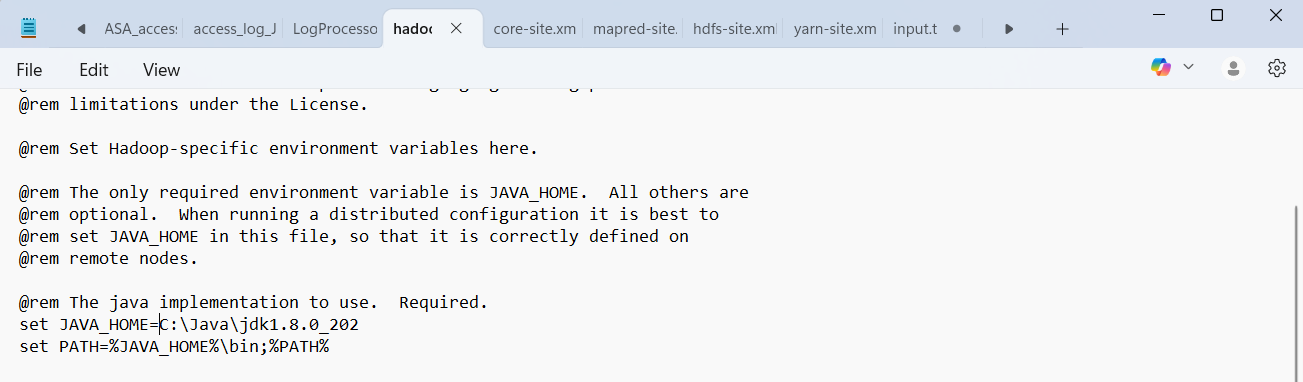
Cài đặt hadoop

Cài gói hadoop-3.3.6 về và giải nén



Cấu hình các file quan trọng.

hadoop-env.cmd sửa JAVA\_HOME



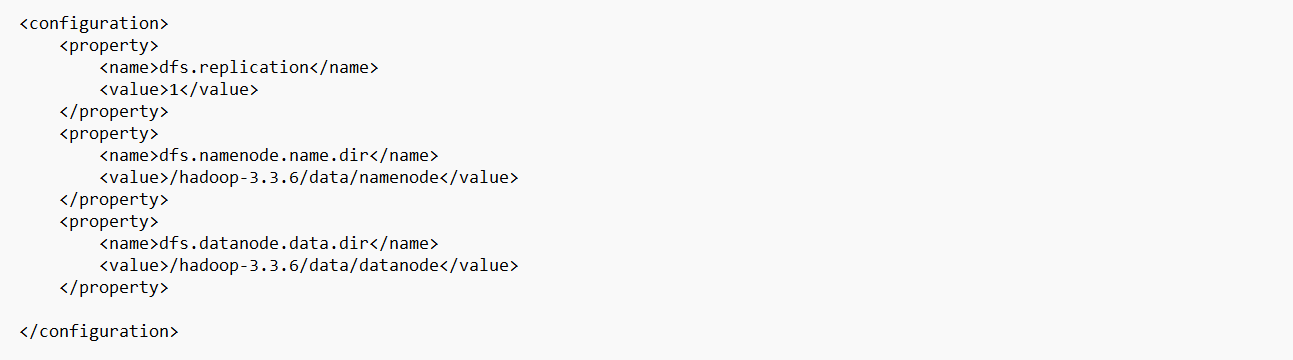
core-site.xml



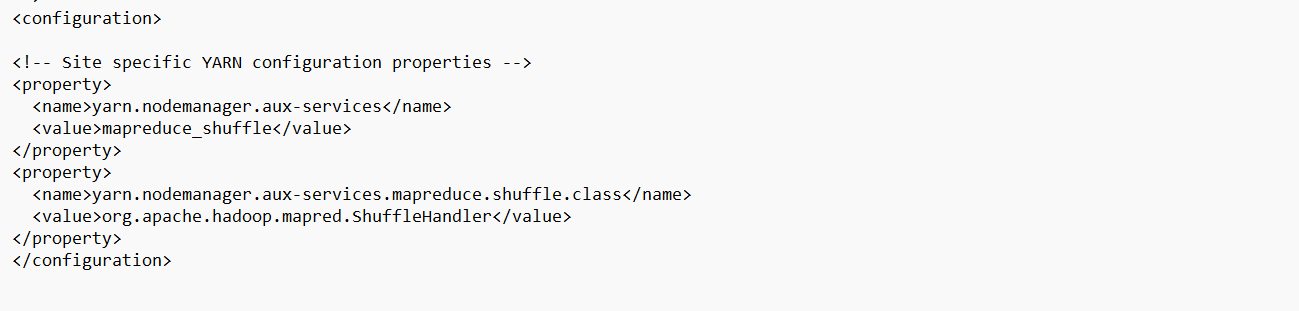
mapre-site.xml



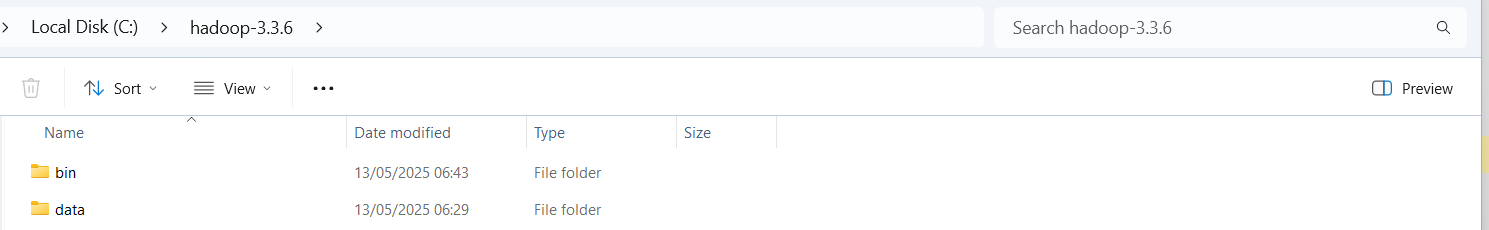
hdfs-site.xml



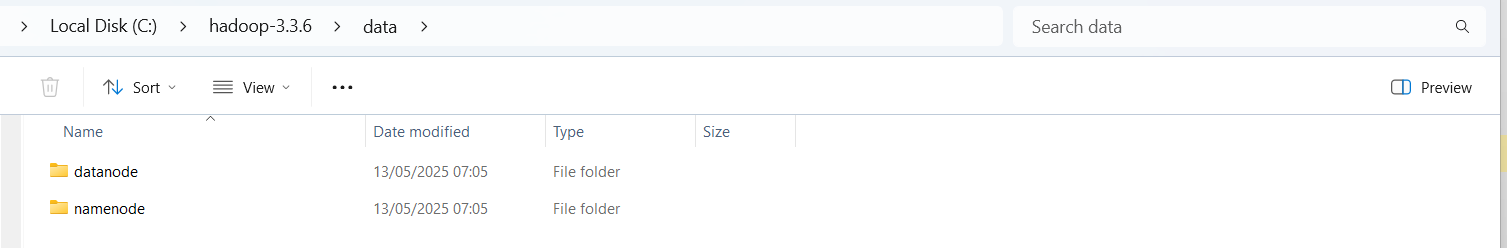
yarn-site.xml



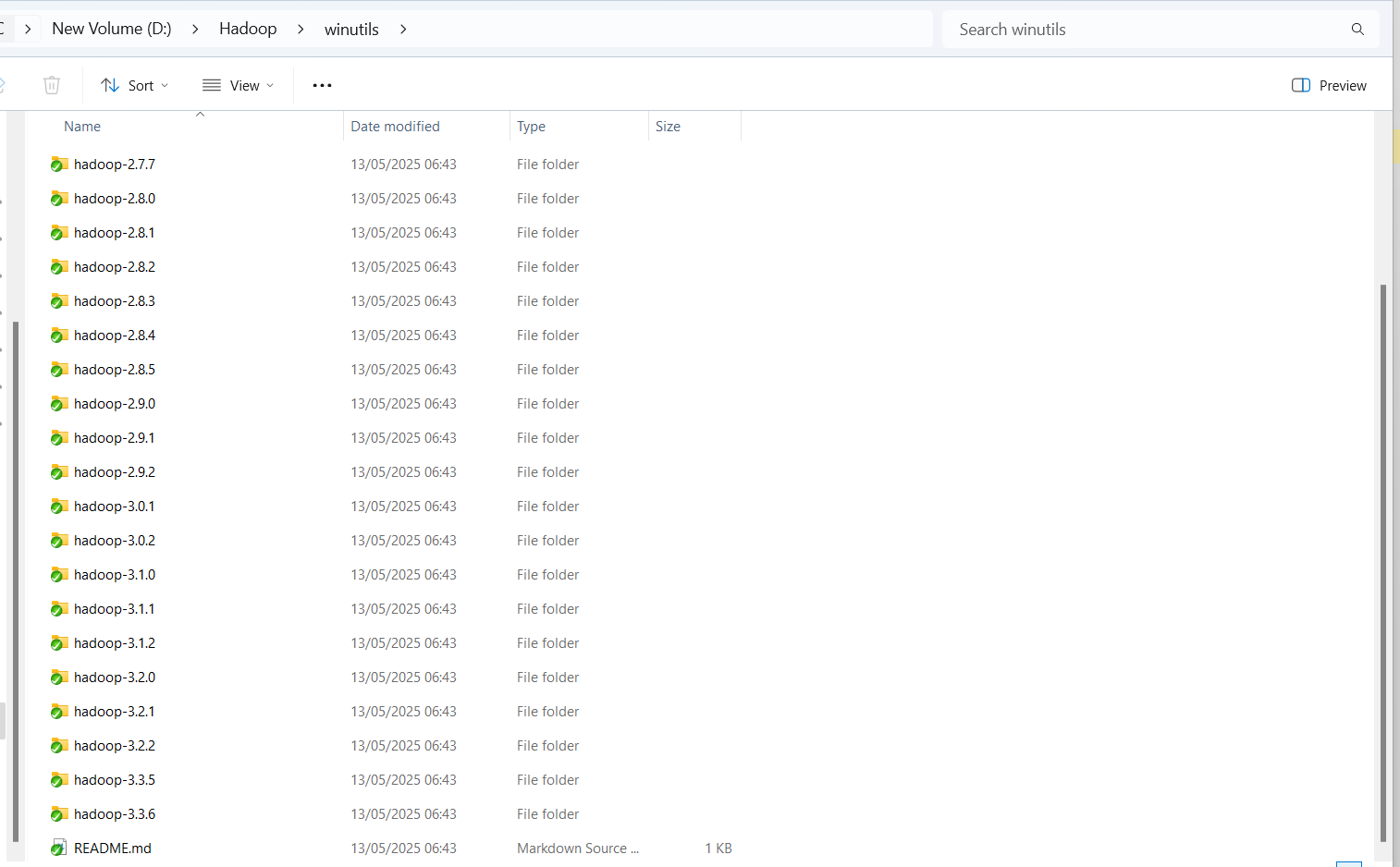
Tạo mới thư mục data



Tạo 2 thư mục trống là datanode và namenode



Tải gói winutils về

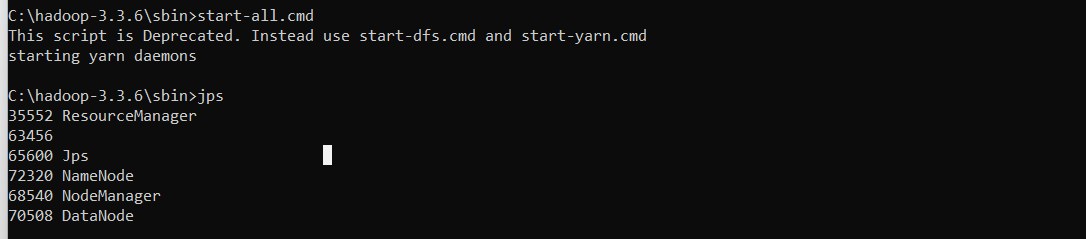


Vào hadoop-3.3.6 copy thư mục bin và dán vào thư mục bin của hadoop-3.3.6 ta tải khi nãy

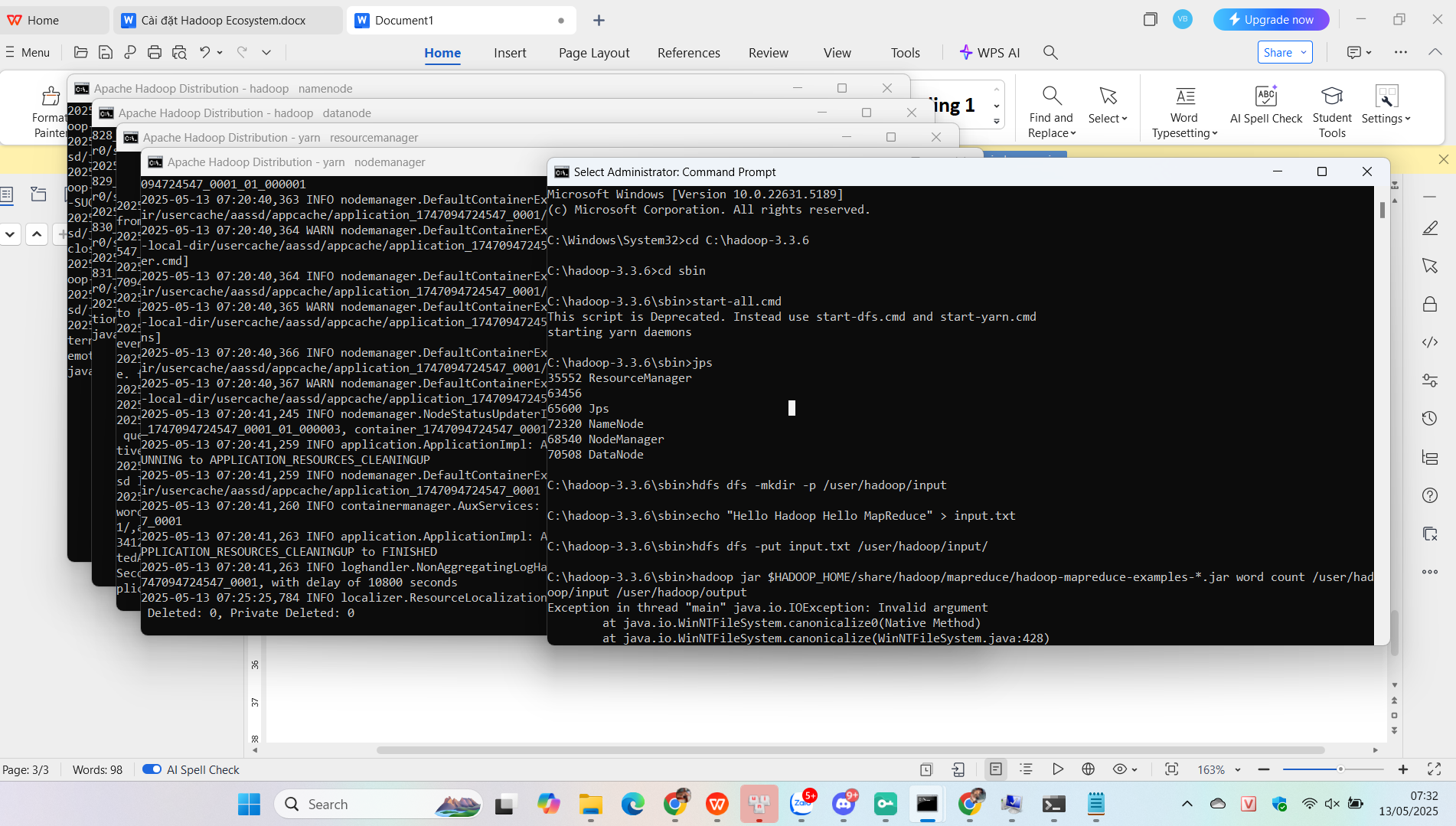
Start hadoop, vào thư mục sbin của hadoop-3.3.6



Chạy start-all.cmd, sau đó jps để kiểm tra start thành công chưa



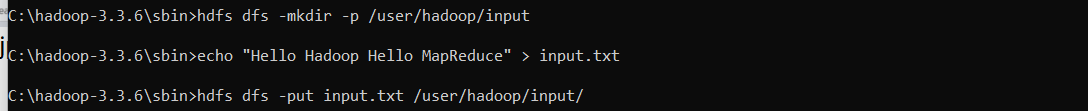
Khi start có 4 cửa sổ hiện ra để chạy các tác vụ



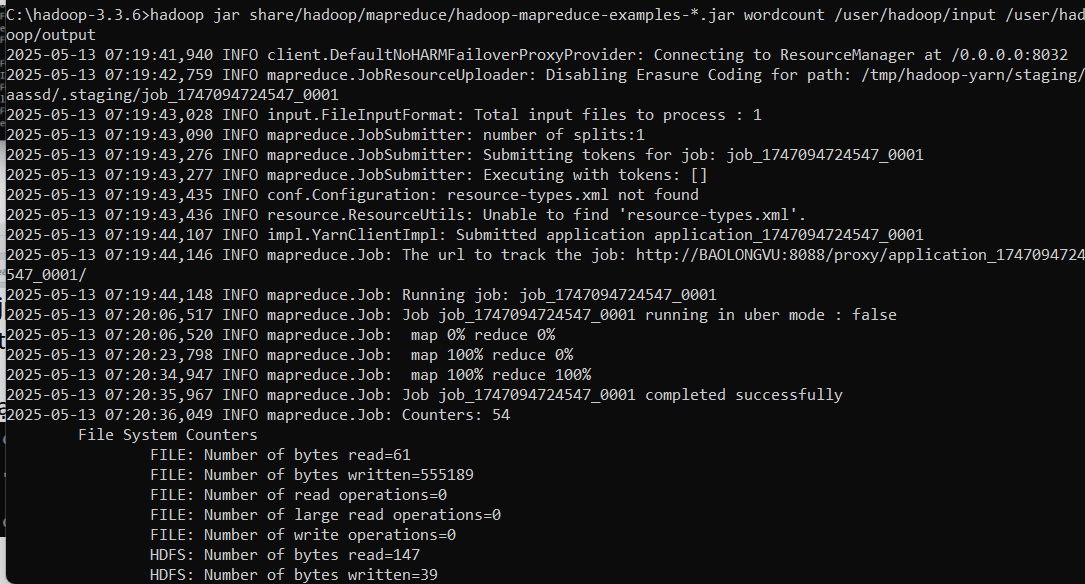
Thông qua đó nếu khi jps không đủ các tiến trình thì sẽ check xem lỗi ở phần nào để khắc phục

Vậy là cài đặt Hadoop trên window thành công!

Test thử mapreduce, tạo thư mục input trên hdfs, tạo 1 file text ở local sau đó đưa lên hdfs



Chạy mapreduce đếm số từ



Sau khi chạy thành công đọc thử file kết quả

