



# Hadoop

```
# 1. 사용자 홈 디렉토리 생성
hdfs dfs -mkdir -p .

# 2. 다른 사용자가 내 디렉토리에 접근 못하게 막음
hdfs dfs -chmod 0700 .

# 3. DB에서 data를 DFS에 업로드
sqoop import --connect jdbc:mysql://3.34.43.90:3306/percast?useSSL=false --table 'table이름' --target-dir 'hdfs 데이터 입력할 경로' --usern

# 4. Input File 업로드 확인
hdfs dfs -cat 'hdfs 데이터 입력할 경로'/part-m-000000

# 5. 환경 변수 편집
vim bashrc

# 6. 환경 변수 편집
export HADOOP_CLASSPATH=$(hadoop classpath)

# 7. WordCount.java 컴파일
mkdir WordCountCompile

javac -classpath ${HADOOP_CLASSPATH} -d 'WordCountCompile 경로' 'WordCount.java 경로'

# 8. WordCount jar 파일 만들기
jar -cvf wordcount.jar -C 'WordCountCompile 경로' 'jar 파일 생성할 경로'

# 9. WordCount 실행
hadoop jar 'jar 파일 경로' 'WordCount 클래스 이름' 'input 데이터 경로' 'output 저장할 경로'

# 10. 결과 확인
hdfs dfs -cat 'output 저장할 경로'/part-r-000000

# 11. hdfs에서 DB로 데이터 전송
sqoop export --connect jdbc:mysql://3.34.43.90:3306/percast?useSSL=false --table ttt --export-dir '전송할 데이터 경로' --username '해당 use

# 12. 다른 프로그램을 사용하기 위해서는 output을 지워야 한다.
hdfs dfs -rm -r output
```