Neo4j and Hadoop video

<https://www.youtube.com/watch?v=RGCuSBmvtNo>

Hadoop Tutorial: Hadoop Ecosystem

<https://www.edureka.co/blog/hadoop-tutorial/#lastFMCaseStudy>

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Check Hadoop and java in cmd:

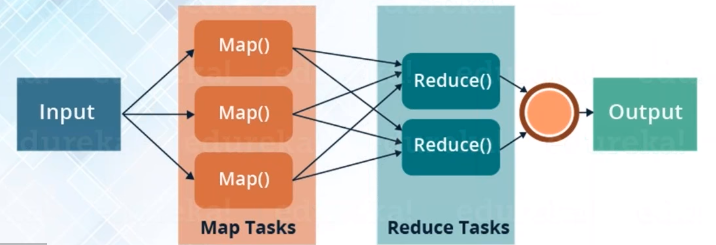
* which Hadoop
* java -version

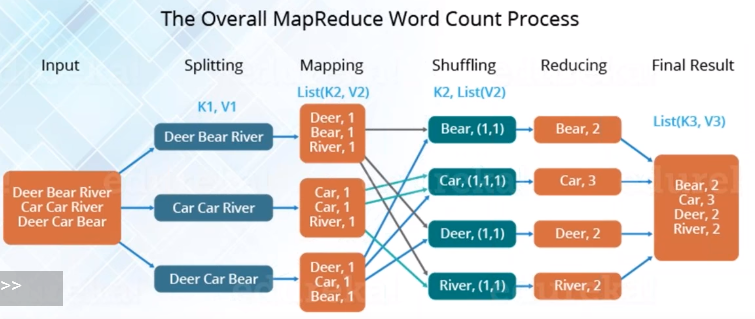
Install ubuntu command

Use command sudo apt-get install neo4j=3.1.4

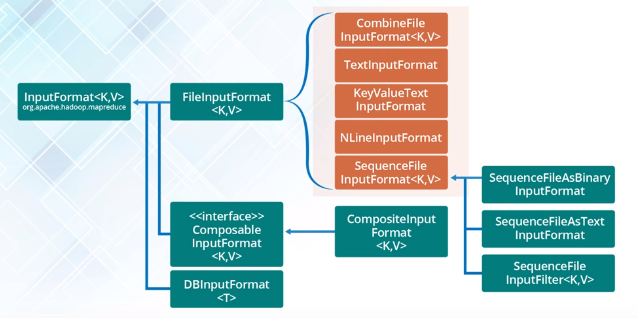
after

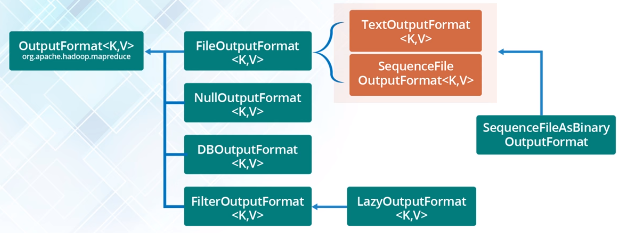
MapReduce Example on word Count Process





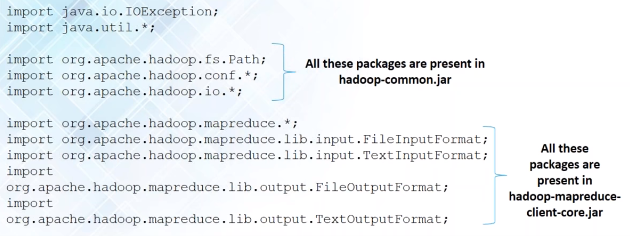
Input/Output Classes in MapReduce



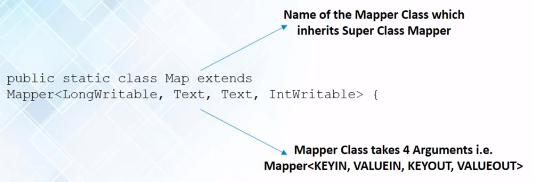


Packages and Classes in Word Count Map Reduce Example

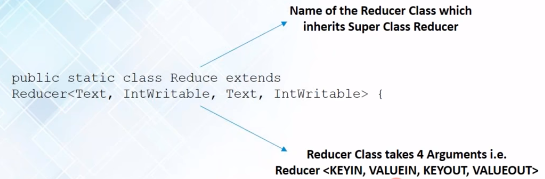
Packages to Import in Map programe



Mapper Class

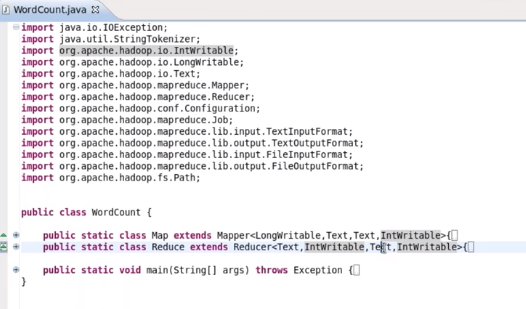


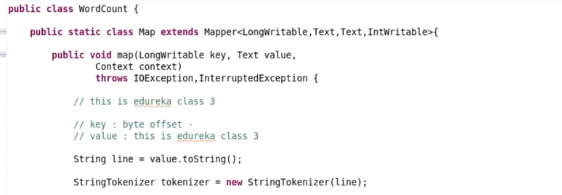
Reducer Class

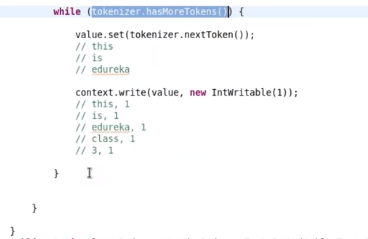


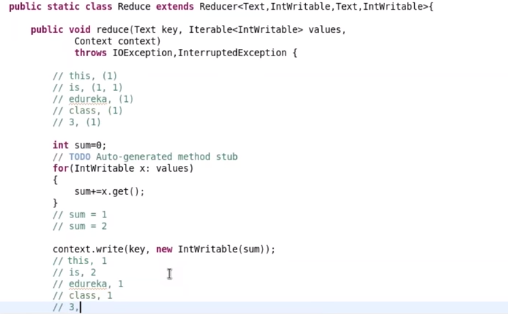
Save the code below as the .jar file

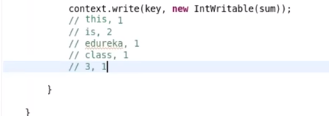
<https://www.youtube.com/watch?v=x-PCNX4prLA>

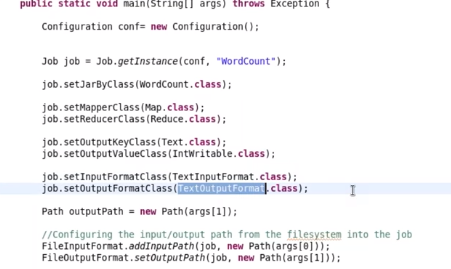


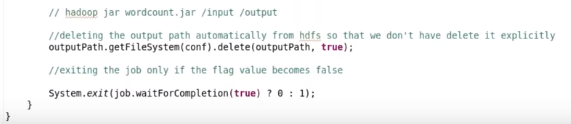












\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Run bash cmd / liunx:

./< dir of scrape.py>

hadoop dfs –put <dir of hadoopApache> /

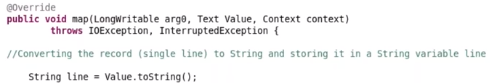
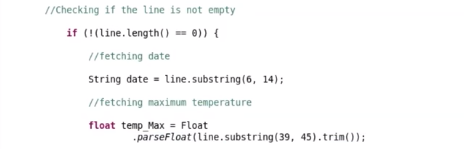
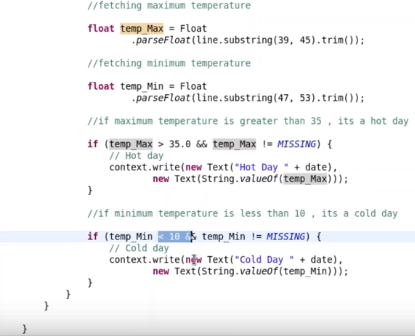
hadoop jar <dir of .jar> /hadoopApache /myWordCount

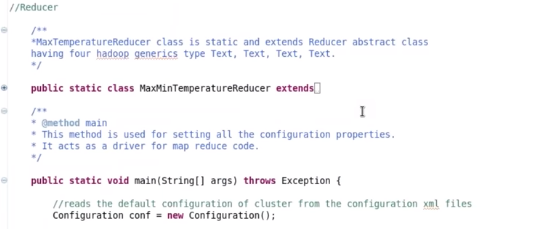
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

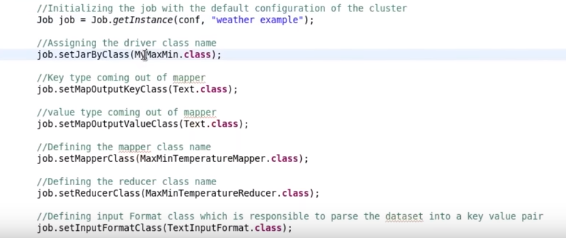
MapReduce is majorly used for:

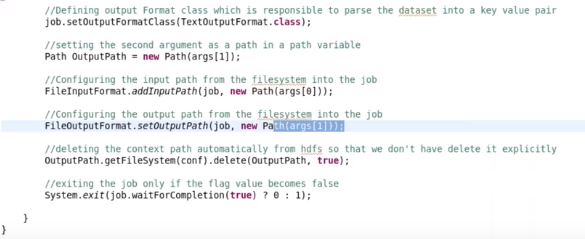
* Analytics: where is can process the data and give the desired results
* Testing: Perform few test cases using MRUnit

Whether Forecasting:







\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Run bash cmd / liunx:

./< dir of scrape.py>

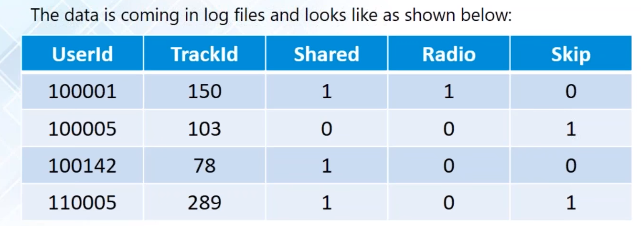
hadoop dfs –put <dir of weather\_dataset> /

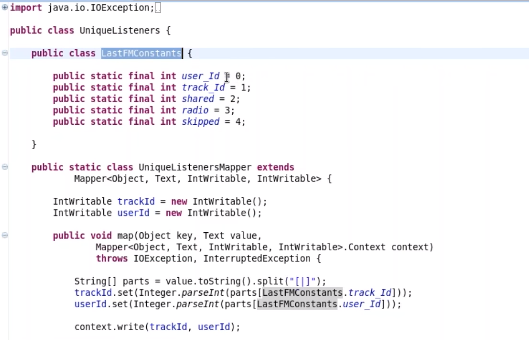
hadoop jar <dir of .jar> / weather\_dataset /my weather\_dataset

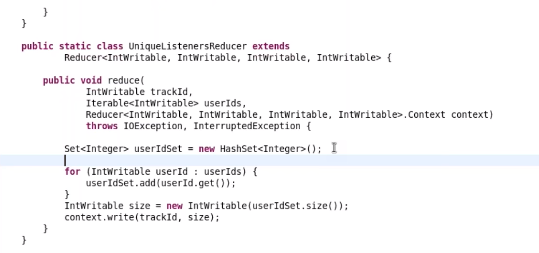
go to VM VirtualBox to check the result

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

MapReduce last.fm Example – an online music website: every song every user listen –csv file









\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Run bash cmd / liunx:

hadoop dfs –put <dir of lastFM> /

hadoop jar <dir of .jar> / lastfm\_sample /MyLastFM

Map Reduce Testing

