```
# job MapReduce / Word Count e Top K
# Prof. Jairson Rodrigues
# Especialização em Ciência de Dados e Analytics
# Soluções para Processamento de Big Data


------------------------ job wc_v1_map / reduce ----------------------------
#!/bin/bash

export STREAM_JAR=/home/vagrant/hadoop/share/hadoop/tools/lib/hadoop-streaming-2.7.4.jar

cd ~/code/wc

hadoop fs -rm -r /poli/output/wc1

hadoop jar $STREAM_JAR     \
 -files wc_map_v1.py,wc_reduce.py     \
 -mapper wc_map_v1.py     \
 -reducer wc_reduce.py     \
 -input /poli/gutenberg     \
 -output /poli/output/wc1

hadoop fs -text /poli/output/wc1/part* | more


------------------------ job wc_v2_map / reduce ----------------------------

#!/bin/bash

export STREAM_JAR=/home/vagrant/hadoop/share/hadoop/tools/lib/hadoop-streaming-2.7.4.jar

cd ~/code/wc

hadoop fs -rm -r /poli/output/wc2

hadoop jar  $STREAM_JAR     \
 -files wc_map_v2.py,wc_reduce.py     \
 -mapper wc_map_v2.py     \
 -reducer wc_reduce.py     \
 -input /poli/gutenberg     \
 -output /poli/output/wc2

hadoop fs -text /poli/output/wc2/part* | more


------------------------ job wc_v3_map / reduce ----------------------------
#!/bin/bash

export STREAM_JAR=/home/vagrant/hadoop/share/hadoop/tools/lib/hadoop-streaming-2.7.4.jar

cd ~/code/wc

hadoop fs -rm -r /poli/output/wc3

hadoop jar  $STREAM_JAR     \
 -files wc_map_v3.py,wc_reduce.py  \
 -mapper wc_map_v3.py     \
 -reducer wc_reduce.py     \
 -input /poli/gutenberg     \
 -output /poli/output/wc3
```

```bash
hadoop fs -text /poli/output/wc3/part* | more


------------------------ job wc_v4_map / reduce ----------------------------
#!/bin/bash

export STREAM_JAR=/home/vagrant/hadoop/share/hadoop/tools/lib/hadoop-streaming-2.7.4.jar

cd ~/code/wc

hadoop fs -rm -r /poli/output/wc4

hadoop jar  $STREAM_JAR     \
 -files wc_map_v4.py,wc_reduce.py     \
 -mapper wc_map_v4.py     \
 -reducer wc_reduce.py     \
 -input /poli/gutenberg     \
 -output /poli/output/wc4

hadoop fs -text /poli/output/wc4/part* | more



------------------------ job topk map / reduce ----------------------------
#!/bin/bash

export STREAM_JAR=/home/vagrant/hadoop/share/hadoop/tools/lib/hadoop-streaming-2.7.4.jar

cd ~/code/topk

hadoop fs -rm -r /poli/output/topk

hadoop jar  $STREAM_JAR     \
 -files topk_map.py,topk_reduce.py     \
 -mapper topk_map.py     \
 -reducer topk_reduce.py     \
 -input /poli/gutenberg     \
 -output /poli/output/topk

hadoop fs -text /poli/output/topk/part* | more
```