Pour le TP MapReduce en python, il faut apporter les modifications suivantes  
   pour exécuter le mapper.py et le reducer.py  
   a. changer le fichier de configuration mapred-site.xml  
      a.1 cd $HADOOP\_CONF\_DIR  
      a.2 gedit mapred-site.xml  
      a.3 le nouveau contenu est :  
       
<configuration>  
<!-- Configurations for MapReduce Applications: -->  
 <property>  
      <name>[mapreduce.framework.name](http://mapreduce.framework.name/)</name>  
      <value>yarn</value>  
  </property>  
<property>  
 <name>yarn.app.mapreduce.am.env</name>  
 <value>HADOOP\_MAPRED\_HOME=/usr/local/hadoop</value>  
</property>  
<property>  
 <name>mapreduce.map.env</name>  
 <value>HADOOP\_MAPRED\_HOME=/usr/local/hadoop</value>  
</property>  
<property>  
 <name>mapreduce.reduce.env</name>  
 <value>HADOOP\_MAPRED\_HOME=/usr/local/hadoop</value>  
</property>  
  
</configuration>  
  
  
    b. changer le contenu de yarn-site.xml  
  
    b.1 cd $HADOOP\_CONF\_DIR  
b.2 gedit yarn-site.xml  
b.3 le nouveau contenu est yarn-site est :  
<configuration>  
<property>  
<name>yarn.nodemanager.aux-services</name>  
<value>mapreduce\_shuffle</value>  
</property>  
</configuration>  
  
3. créer un répertoire TPMAPRED  
4. Enregister mapper.py et reducer.py dans TPMAPRED  
5. Donner le droit d'exécution de mapper et de reducer  
**chmod a+x mapper.py  
chmod a+x reducer.py**  
6. Le hadoop-streaming-3.3..1.jar se trouve dans le répertoire  
**$HADOOP\_HOME/share/hadoop/tools/lib**  
7. Lancer la commande d'exécution de mapper et reducer  
  -  cd TPMAPRED  
   **hadoop jar $HADOOP\_HOME/share/hadoop/tools/lib/hadoop-streaming-3.3.1.jar -mapper mapper.py  
   -reducer reducer.py -file mapper.py -file reducer.py -input /input -output /joboutput**