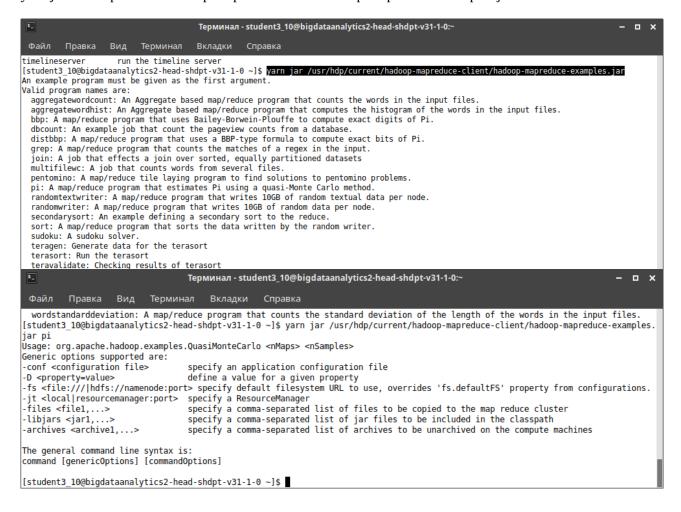
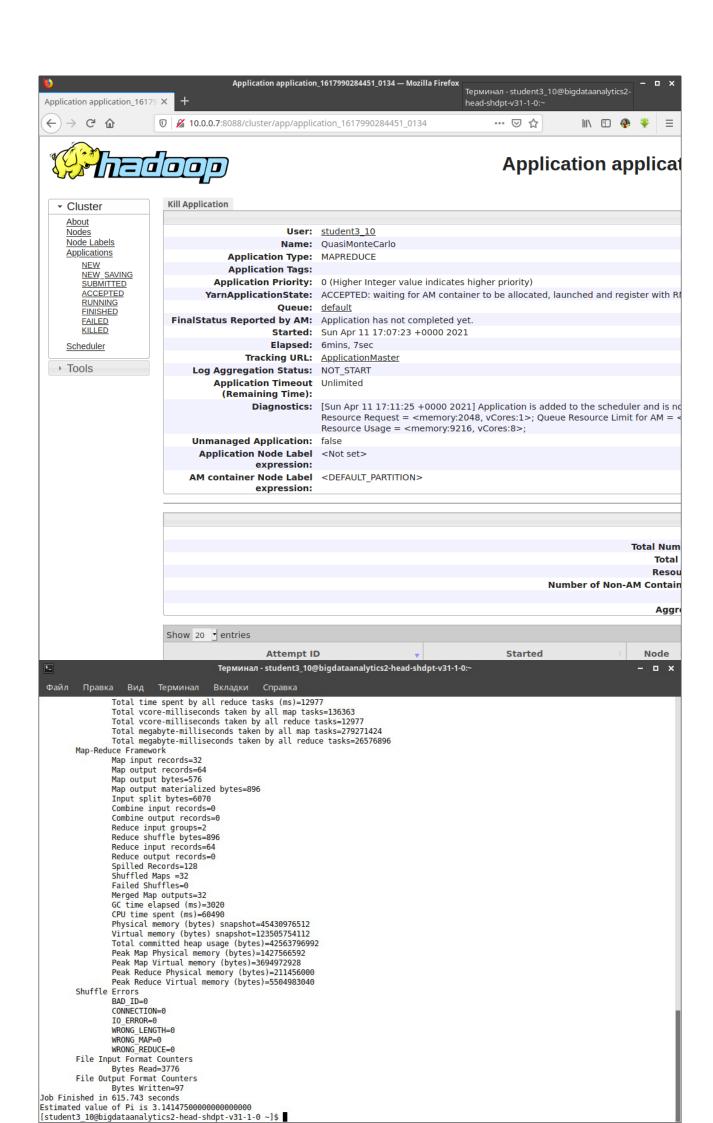
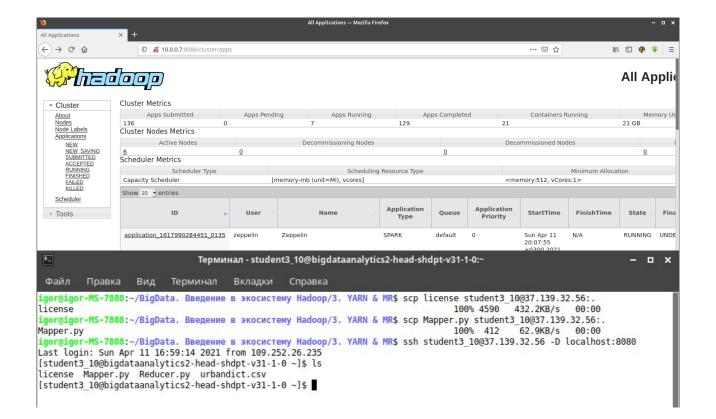
ssh student3 10@37.139.32.56

ssh student3 10@37.139.32.56 -D localhost:8080

ssh -i /home/igor/Загрузки//id_rsa_student3_10 student3_10@37.139.32.56 -D localhost:8080 yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-mapreduce-examples.jar







Файл Правка Вид Терминал Вкладки Справка

[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]\$ cat license

All of the documentation and software included in the 4.4BSD and 4.4BSD-Lite Releases is copyrighted by The Regents of the Un iversity of California.

Copyright 1979, 1980, 1983, 1986, 1988, 1989, 1991, 1992, 1993, 1994 The Regents of the University of California. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following co nditions are met:

Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaime r in the documentation and/or other materials provided with the distribution.

All advertising materials mentioning features or use of this software must display the following acknowledgement:

This product includes software developed by the University of California, Berkeley and its contributors.

Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL T HER REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTI ON) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OT HERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

The Institute of Electrical and Electronics Engineers and the American National Standards Committee X3, on Information Proces sing Systems have given us permission to reprint portions of their documentation.

In the following statement, the phrase "this text" refers to portions of the system documentation.

Portions of this text are reprinted and reproduced in electronic form in the second BSD Networking Software Release, from IEE E Std 1003.1-1988, IEEE Standard Portable Operating System Interface for Computer Environments (POSIX), copyright C 1988 by the Institute of Electrical and Electronics Engineers, Inc. In the event of any discrepancy between these versions and the ori ginal IEEE Standard, the original IEEE Standard is the referee document.

In the following statement, the phrase "This material" refers to portions of the system documentation.

This material is reproduced with permission from American National Standards Committee X3, on Information Processing Systems. Computer and Business Equipment Manufacturers Association (CBEMA), 311 First St., NW, Suite 500, Washington, DC 20001-2178. The developmental work of Programming Language C was completed by the X3J11 Technical Committee.

The views and conclusions contained in the software and documentation are those of the authors and should not be interpreted as representing official policies, either expressed or implied, of the Regents of the University of California.

NOTE: The copyright of UC Berkeley's Berkeley Software Distribution ("BSD") source has been updated. The copyright addendum m ay be found at ftp://ftp.cs.berkeley.edu/pub/4bsd/README.Impt.License.Change and is included below.

July 22, 1999

To All Licensees, Distributors of Any Version of BSD:

As you know, certain of the Berkeley Software Distribution ("BSD") source code files require that further distributions of products containing all or portions of the software, acknowledge within their advertising materials that such products contain software developed by UC Berkeley and its contributors.

Specifically, the provision reads:

- * 3. All advertising materials mentioning features or use of this software
- must display the following acknowledgement:
- This product includes software developed by the University of

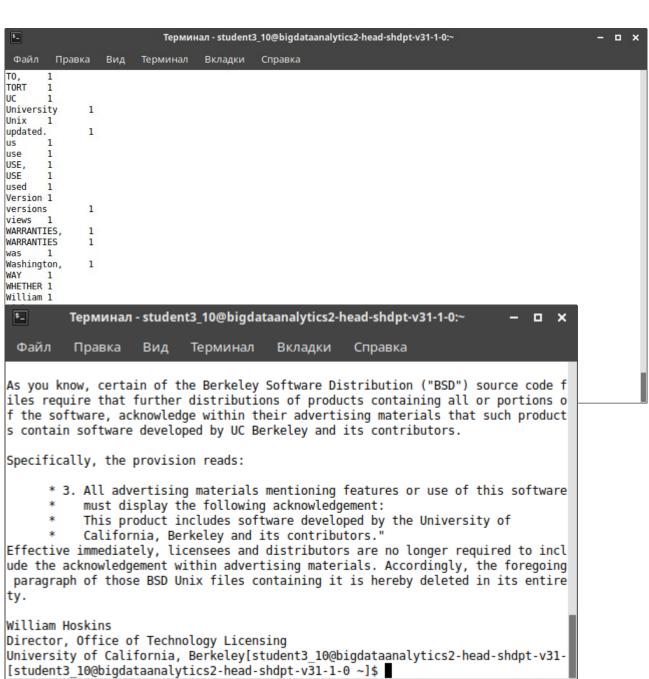
California, Berkeley and its contributors.

Effective immediately, licensees and distributors are no longer required to include the acknowledgement within advertising ma terials. Accordingly, the foregoing paragraph of those BSD Unix files containing it is hereby deleted in its entirety.

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:-

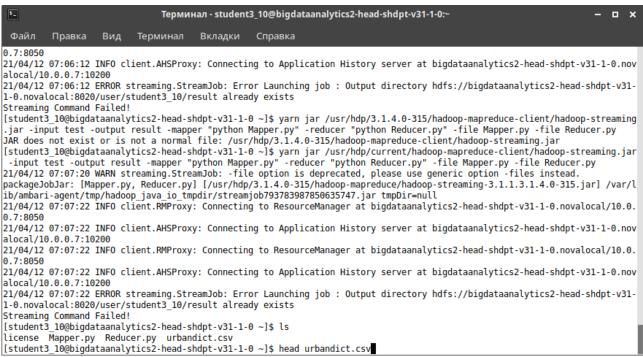
				терминал	students_roesigual	danaly (1632-116au	-311upt-v31-1-0.		^
Файл І	Правка	Вид	Терминал	Вкладки	Справка				
those 1									
BSD 1									
Unix 1									
files 1									
containing	1								
it 1 is 1									
hereby 1									
deleted 1									
in 1 its 1									
entirety.	1								
William 1	1								
Hoskins 1									
Director,	1								
Office 1	1								
of 1									
Technology	1								
Licensing	1								
University									
of 1	1								
California	, 1								
Berkeley	1								
	10@bigda	taanaly	tics2-head-s	hdpt-v31-1-	0 ~]\$				

\$_	Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~									
Файл	Правка	Вид	Терминал	Вкладки	Справка					
with	1									
with	1									
with	1									
within	1									
within	1									
without	1									
without	1									
work	1									
written	1									
ХЗ,	1									
ХЗ,	1									
X3J11	1									
you	1									
) ~]\$ vim Reducer.py					
			ics2-head-sh:) ~]\$ ls			- 1		
			er.py urband							
[student	:3_10@bigdat	taanalyt	ics2-head-sh:	dpt-v31- <u></u> 1-0) ~]\$					



yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming.jar -input test -output result -mapper "python3.7 Mapper.py" -reducer "python3.7 Reducer.py" -file Mapper.py -file Reducer.py

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
                                                                                                                                     Файл
          Правка
                     Вид Терминал
                                           Вкладки Справка
9288 0028 m 000001
Job failed as tasks failed. failedMaps:1 failedReduces:0 killedMaps:0 killedReduces: 0
21/04/12 06:55:25 INFO mapreduce.Job: Counters: 12
         Job Counters
                  Failed map tasks=7
                  Killed map tasks=1
                  Killed reduce tasks=1
                  Launched map tasks=7
                  Other local map tasks=5
                  Data-local map tasks=1
                  Rack-local map tasks=1
                  Total time spent by all maps in occupied slots (ms)=116692
                  Total time spent by all reduces in occupied slots (ms)=0
                  Total time spent by all map tasks (ms)=29173
                  Total vcore-milliseconds taken by all map tasks=29173
                  Total megabyte-milliseconds taken by all map tasks=59746304
21/04/12 06:55:25 ERROR streaming.StreamJob: Job not successful!
Streaming Command Failed!
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming.jar
-input test -output result -mapper "python Mapper.py" -reducer "python Reducer.py" -file Mapper.py -file Reducer.py
```



head -n 1 urbandict.csv

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
                                                                                                                                                         Файл
            Правка
                        Вид Терминал
                                                 Вкладки Справка
21/04/12 07:07:22 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics2-head-shdpt-v31-1-0.novalocal/10.0.
0.7:8050
21/04/12 07:07:22 INFO client.AHSProxy: Connecting to Application History server at bigdataanalytics2-head-shdpt-v31-1-0.nov
alocal/10.0.0.7:10200
21/04/12 07:07:22 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics2-head-shdpt-v31-1-0.novalocal/10.0.
0.7:8050
21/04/12 07:07:22 INFO client.AHSProxy: Connecting to Application History server at bigdataanalytics2-head-shdpt-v31-1-0.nov
alocal/10.0.0.7:10200
21/04/12 07:07:22 ERROR streaming.StreamJob: Error Launching job : Output directory hdfs://bigdataanalytics2-head-shdpt-v31-
1-0.novalocal:8020/user/student3_10/result already exists
Streaming Command Failed!
[student3 10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ls
license Mapper.py Reducer.py urbandict.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ head urbandict.csv
word_id,word,up_votes,down_votes,author,definition
0000007,Janky,296,255,dc397b2f,"Undesirable; less-than optimum."
0000008,slumpin',16,37,dc397b2f,"low down and funky, but [knee deep] enough to ride to."
0000009,yayeeyay,19,27,dc397b2f,"affirmation; suggestion of encouragement, approval, or interest."
.0000012,hard-core,162,96,d1610749,"anything out of our league that can be good or bad.
0000013,brutal,12,45,40ecelef,"anything out of our league that can be good of bad.
0000013,brutal,12,45,40ecelef, "anything that makes you sweat"
0000014,skanky,9,48,485e4db7,"Anything of or pertaining to a $10,000 hooker."
0000015,ho-bag,26,27,b37fba05,"a term of endearment, used affectionately for your roommate. First used in the schools' parking lot after an incident with the hall moniters."
0000016,massive,36,45,b9dcf126,"really really good. excellently good."
0000017,wtf,183,99,a6c97ba3,"what the fuck? ;; use it <u>in place of expletives.</u> a more polite alternative."
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ head -n 1 urbandict.csv
word id,word,up votes,down votes,author,definition
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

head urbandict.csv > urbandict_test.csv cat vehicles.csv | python Mapper.py Удалем резалт: hdfs dfs -rm -r result

Просмотр: cat vehicles.csv | python3.7 mapper.py



cat vehicles.csv | python3.7 mapper.py | sort | python3.7 reducer.py

```
Герминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:
                                                                                                                                                                                                                               •
File "Reducer.py", line 11
SyntaxError: Non-ASCII character '\xd1' in file Reducer.py on line 11, but no encoding declared; see http://www.python.org/peps/pep-0263.html for details
SyntaxError: Non-ASCII character '\xd1' in file Reducer.py on line 11, but no encoding declared; see http://www.python.org/peps/pep-0263.html for details [student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cat vehicles.csv | python3.7 Mapper.py | sort | python3.7 Reducer.py

Traceback (most recent call last):
File "Reducer.py", line 18, in <module>
word, count = line.split('\t', 1)
ValueError: not enough values to unpack (expected 2, got 1)
[student3 10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
[students] 10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cat vehicles.csv | python3.7 mapper.py | sort | python3.7 reducer.py acura 12143246
alfa-romeo
                       894208
aston-martin 915150
audi 27500345
            53041036
buick 9
           9414316
                        24267276
chevrolet
chrysler
datsun 280918
                        1452593447
dodge 44587781
ferrari 4808024
            2931321
 fiat
           275376101
2873750514
ford
gmc 28/3/20214
harley-davidson 328464
honda 59867277
hyundai 21931470
 infiniti
                       13550397
           6137143
69589094
 jeep
kia
            19559407
land rover 9
lexus 30034843
lincoln 7796232
                        94400
mazda 14377882
mercedes-benz
mercury 1162558
mini 7507965
mitsubishi
                       65436510
                        3672210
morgan 74977
nissan 67049248
pontiac 3417008
porsche 17043208
ram 92764848
 rover
            9843502
 saturn 885346
subaru 30836333
tesla 10943555
toyota 134519824
volkswagen 2
volvo 8615208
                        26791675
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

scp mapper.py student3_10@37.139.32.56:.

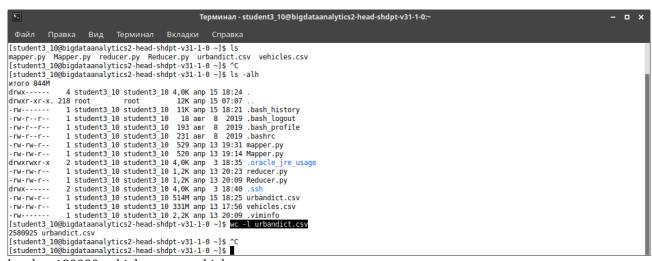
wc -l vehicles.csv

head -n 100000 vehicles.csv > small.csv

Посмотреть что хранится: ls -alh Почистить/удалить файлы: rm

Посмотреть сколько строк в файле: wc -l <имя.файла>

wc -l urbandict.csv



head -n 100000 vehicles.csv > vehicles.csv

head -n 1290462 urbandict.csv > urbandict_new.csv

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:-
  -rw----- 1 student3_10 student3_10 2,2K anp 13 20:09 .viminfo
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ wc -l urbandict.csv
  2580925 urbandict.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ^C
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ |
[student3_10@bigdataanalytics
     итого 1,1G
                                                                               4 student3_10 student3_10 4,0K anp 15 18:53
                                                                           18 root root 12K anp 15 07:07 .
1 student3_10 student3_10 11K anp 15 07:07 .
1 student3_10 student3_10 11K anp 15 18:21 .bash_history
1 student3_10 student3_10 18 abr 8 2019 .bash_logout
1 student3_10 student3_10 193 abr 8 2019 .bash_profile
 drwxr-xr-x. 218 root
     -rw-r--r--
     -rw-r--r--
                                                                           1 student3_10 student3_10 193 aBr 8 2019 .bashrc 1 student3_10 student3_10 231 aBr 8 2019 .bashrc 1 student3_10 student3_10 529 anp 13 19:31 mapper.py 1 student3_10 student3_10 520 anp 13 19:14 Mapper.py 2 student3_10 student3_10 4,0K anp 3 18:35 .oracle_jrc 1 student3_10 student3_10 1,2K anp 13 20:23 reducer.py
     - rw- rw- r- -
     - rw- rw- r- -
     - rw- rw- r- -
                                                                           1 student3_10 student3_10 1,2K anp 13 20:09 Reducer.py
2 student3_10 student3_10 1,2K anp 13 20:09 Reducer.py
2 student3_10 student3_10 4,0K anp 3 18:40 .ssh
1 student3_10 student3_10 514M anp 15 18:25 urbandict.csv
1 student3_10 student3_10 24M anp 13 17:56 vehicles.csv
1 student3_10 student3_10 331M anp 13 17:56 vehicles.csv
1 student3_10 student3_10 2,2K anp 13 20:09 .viminfo
     -rw-rw-r--
  drwx-----
-rw-rw-r--
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

rm urbandict.csv mv small.csv vihicles.csv Переименовываю файл mv urbandict new.csv urbandict.csv

Загружаю: hdfs dfs -put vehicles.csv . Загружаю: hdfs dfs -put urbandict.csv .

Правильный просмотр размера файла: hdfs dfs -ls -h

```
Терминал - student3 10@biqdataanalytics2-head-shdpt-v31-1-0:
                                                                                                                                                                                                   0
  Файл Правка Вид Терминал Вкладки Справка
                  1 student3_10 student3_10 520 anp 13 19:14 Mapper.py
 - rw- rw- r - -
                   2 student3 10 student3 10 4,0K anp 3 18:35 .oracle jr
1 student3 10 student3 10 1,2K anp 13 20:23 reducer.py
drwxrwxr-x
 - rw - rw - r - -
                  1 student3 10 student3 10 1,2K anp 13 20:09 Reducer.py
 - rw- rw- r - -
                  2 student3 10 student3 10 4,0K anp 3 18:40 .ssh
1 student3 10 student3 10 246M anp 15 18:53 urbandict.csv
 - rw- rw- r--
-rw-rw-r-- 1 student3 10 student3 10 331M ann 13 17:56 vehicles.csv
-rw-rw--- 1 student3 10 student3 10 2,2K ann 13 20:09 .viminfo
[student3 10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -put urbandict.csv .
[student3 10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
                 - student3_10 student3_10
                                                              0 2021-04-15 18:00 .Trash
drwx-----
Found 5 items
                 - student3_10 student3_10
- student3_10 student3_10
- student3_10 student3_10
drwx-----
                                                              0 2021-04-15 18:00 .Trash
                                                              0 2021-04-13 20:35 .staging
0 2021-04-15 18:20 lesson
drwxr-xr-x
drwxr-xr-x
                 - student3_10 student3_10
2 student3_10 student3_10
                                                              0 2021-04-15 17:36 test
```

Удаляем директорию .Trash со всем содержимым: hdfs dfs -rm -r .Trash

```
Терминал - student3 10@bigdataanalytics2-head-shdpt-v31-1-0:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  о x
    .bash_history
                                                                        .bash_profile
                                                                                                                                               mapper.py
                                                                                                                                                                                                                       .oracle_jre_usage/ Reducer.py
                                                                                                                                                                                                                                                                                                                                                                     urbandict.csv
                                                                                                                                                                                                                                                                                                                                                                                                                                             .viminfo
.bash_logout bashrc Mapper.py reducer.py .

[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -rm -r .Trash Deleted .Trash [student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
                                                                                                                                                                                                                                                                                              .ssh/
                                                                                                                                                                                                                                                                                                                                                                     vehicles.csv
    Found 3 items
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
| Total Control | Total Control Contro
 Deleted .Trash
  [student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
     ound 2 items
drwxr.xr.x - student3_10 student3_10 0 2021-04-15 19:19 less drwxr.xr.x - student3_10 student3_10 0 2021-04-15 17:36 test [student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
                                                                                                                                                                                    0 2021-04-15 19:19 lesson
```

Проверка перемещённых файлов hdfs dfs -ls lesson

hdfs dfs -ls test

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:-
                                                                                                                                                                                                                                                    о x
                   - student3_10 student3_10
- student3_10 student3_10
- student3_10 student3_10
                                                                              0 2021-04-13 20:35 .staging
drwx-----
                                                                              0 2021-04-15 19:19 lesson
0 2021-04-15 17:36 test
drwxr-xr-x
[student3_10@bigdataanalytics2-head-shdpt-v31-10-0]$ hdfs offs://bigdataanalytics2-head-shdpt-v31-10-0]$ hdfs offs://bigdataanalytics2-head-shdpt-v31-10-0.novalocal:8020/user/student3_10/.staging to trash at: hd fs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/.staging
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
Found 3 items
                  - student3_10 student3_10
- student3_10 student3_10
- student3_10 student3_10
                                                                              0 2021-04-15 19:24 .Trash
drwx----
                                                                              0 2021-04-15 19:19 lesson
0 2021-04-15 17:36 test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -rm -r .Trash
Deleted .Trash
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
Found 2 items
drwxr-xr-x - student3_10 student3_10
drwxr-xr-x - student3_10 student3_10
student3_10-student3_10-student3_10
                                                                              0 2021-04-15 19:19 lesson
0 2021-04-15 17:36 test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls lesson
 |Students_regulations.com, ----|
Found 1 items
-rw-r--r-- 2 student3_10 student3_10 257081603 2021-04-15 19:07 | | lesson/urbandict.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls tes
Found 1 items
-rw-r--r- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ |
```

Проверка размера моей папки: hdfs dfs -du -h

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -rm -r .staging
21/04/15 19:24:53 INFO fs.TrashPolicyDefault: Moved: 'hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/.staging' to trash at: hd
fs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/.Trash/Current/user/student3_10/.staging
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
Found 3 items
                  - student3_10 student3_10
- student3_10 student3_10
- student3_10 student3_10
                                                                             0 2021-04-15 19:24 .Trash
0 2021-04-15 19:19 lesson
0 2021-04-15 17:36 test
drwxr-xr-x
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -rm -r .Trash
Deleted .Trash
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
Found 2 items
drwxr-xr-x - student3_10 student3_10
drwxr-xr-x - student3_10 student3_10
                                                                             0 2021-04-15 19:19 lesson
                                                                             0 2021-04-15 17:36 test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls lesson
Found 1 items
round 11cems
-rw-r--r- 2 student3 10 student3 10 257081603 2021-04-15 19:07 lesson/urbandict.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test
Found 1 items
-rw-r--r- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -du -h
245.2 M 490.3 M lesson
330.3 M 660.6 M test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

cat vehicles.csv | python3.7 mapper.py

cat vehicles.csv | python3.7 mapper.py | sort | python3.7 reducer.py



current

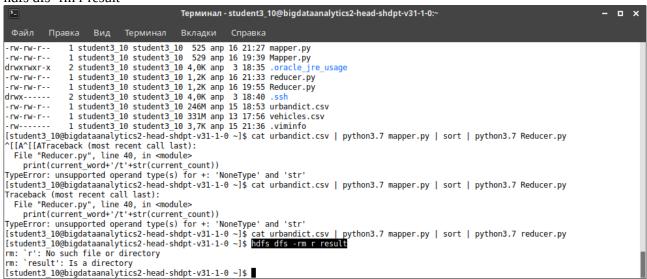
Удалить: hdfs dfs -rm -r result

Посмотреть результат: hdfs dfs -cat result/part*

cat urbandict.csv | python3.7 mapper.py



hdfs dfs -rm r result



Для запуска распределенного вычисления воспользуемся командой:

yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming.jar -input lesson -output result -mapper "python3.7 mapper.py" -reducer "python3.7 reducer.py" -file mapper.py -file reducer.py

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
 Файл
        Правка Вид Терминал
                                      Вкладки Справка
21/04/16 21:56:02 INFO mapreduce.Job: Job job 1618415779628_0247 running in uber mode : false
21/04/16 21:56:02 INFO mapreduce.Job: map 0% reduce 0%
21/04/16 21:56:12 INFO mapreduce.Job: Task Id : attempt_1618415779628_0247_m_0000000_0, Status : FAILED
[2021-04-16 21:56:10.417]Container killed on request. Exit code is 137
[2021-04-16 21:56:10.431]Container exited with a non-zero exit code 137.
[2021-04-16 21:56:10.434]Killed by external signal
21/θ4/16 21:56:15 INFO mapreduce.Job: map 50% reduce 0%
21/04/16 21:56:23 INFO mapreduce.Job: map 100% reduce 0% 21/04/16 21:56:26 INFO mapreduce.Job: map 100% reduce 100%
21/04/16 21:56:26 INFO mapreduce.Job: Job job 1618415779628 0247 completed successfully
21/04/16 21:56:26 INFO mapreduce.Job: Counters: 56
        File System Counters
                FILE: Number of bytes read=15399937
                FILE: Number of bytes written=31517869
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=257122730
                HDFS: Number of bytes written=0
                HDFS: Number of read operations=11
HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Failed map tasks=1
                Launched map tasks=3
                Launched reduce tasks=1
                Other local map tasks=1
                Data-local map tasks=1
                Rack-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=112752
Total time spent by all reduces in occupied slots (ms)=31328
                Total time spent by all map tasks (ms)=28188
Total time spent by all reduce tasks (ms)=7832
                Total vcore-milliseconds taken by all map tasks=28188
Total vcore-milliseconds taken by all reduce tasks=7832
                Total megabyte-milliseconds taken by all map tasks=57729024
                Total megabyte-milliseconds taken by all reduce tasks=16039936
        Map-Reduce Framework
                Map input records=1290462
                Map output records=1290462
 PC
             PycharmProjects – ~/BigData. Введение в экосистему Hadoop/4. Hive & HUE/mapper.py
                                                                                                                     - 0 X
File Edit View Navigate Code Refactor Run Tools Git Window Help
                                                                          # □ Git: V / 7 ○ 5 Q $
doop ) 4. Hive & HUE ) 🐌 mapper.py
                                                         runner *
      mapper.py X
                          is reducer.py
 Project
               #!/usr/bin/env python
                                                                                                                     A1 ^ v
Ba
      2
               """mapper.py"""
 Commit
               import sys
      5
 -0
              for line in sys.stdin:
                    # удаляем лишние пробелы
                    line = line.strip()
                    # делим строки на слова
                    words = line.split(',')
                    try:
                          word = words[4]
                          # price = words[5]
                          # добавляем значение для счетчика
 Structure
    15
                          # for word in words:
                          # выводим в output пару ключ и значение
    17
                          print(word.lower()+'\t')
 .
                     except Exception:
     19
                          continue
 Favorites
 *
                                                                                                               C Event Log
              I TODO

    Problems

                                        Python Packages

■ Terminal

                                                                                Python Console
```

20:1 LF UTF-8 4 spaces Python 3.7 12 master 1

Nim toolchain: Cannot find Nim co... (30 minutes ago)

```
PC
         PycharmProjects – ~/BigData. Введение в экосистему Hadoop/4. Hive & HUE/reducer.py
<u>File Edit View Navigate Code Refactor Run Tools Git Window Help</u>
doop 〉 4. Hive & HUE 〉 🐉 reducer.py 🎩 🔻 📗 👂 runner 🔻 🕨 🐞 🖫 📕 🕒 Git: 🗹 🗸 🧷 🧐 🔾 🜣
    #!/usr/bin/env python
                                                                           A1 A1 2 A V
"""reduser.py"""
    2
Commit
         from operator import itemgetter
         import sys
    6
    7
          current_word = None
          current_count = 0
   9
          word = None
          # читаем строки из STDIN
         for line in sys.stdin:
   12
              line = line.strip()
   13
             # line = line.lower()
              # удаляем пробелы в конце и начале строки
            # разделяем пары ключ и значение
              try:
   17
                  word, count = line.split('\t', 1)
              except Exception:
   19
                  continue
   21
              # преобразуем значение в число
              trv:
   23
                  count = int(count)
   24
              except ValueError:
                  # игнорируем ошибки
                  continue
              # на Reduce приходят данные после фазы Sort
              # поэтому этот код будет работать
              if current_word == word:
   31
                  current_count += count
         else:
   32
   33
                  if current word:
                      # пишем результат в STDOUT
   34
Structure
                      print(current_word+'\t'+str(current_count))
                  current_count = count
   37
                  current_word = word
...
   38
          # не забываем про последнее слово
Favorites
   39
          if current_word == word:
              print(current_word+'\t'+str(current_count))
   40
*
           for line in sys.stdin > else

    ▼ Terminal  Python Console

                                                                              C Event Log
                                              32:7 LF UTF-8 4 spaces Python 3.7 12 master 🐿
Nim toolchain: Cannot find Nim co... (31 minutes ago)
```