

HBase

Overview:

Download amazon video game review data from the link https://s3.amazonaws.com/amazon-reviews-pds/tsv/amazon_reviews_us_Video_Games_v1_00.tsv.gz (Links to an external site.) . After extracting you will get a “.tsv” file where the columns are ‘tab’ separated. The description of the columns can be found in the link <https://s3.amazonaws.com/amazon-reviews-pds/tsv/index.txt> (Links to an external site.) . For this assignment, you must run different queries on **HBase** using this dataset.

Data File: amazon_reviews_us_Video_Games_v1_00.tsv

- Removed the first line of the given tsv file (in the question) which contains the column names
- Added only those columns which were required in the question

Part 1:

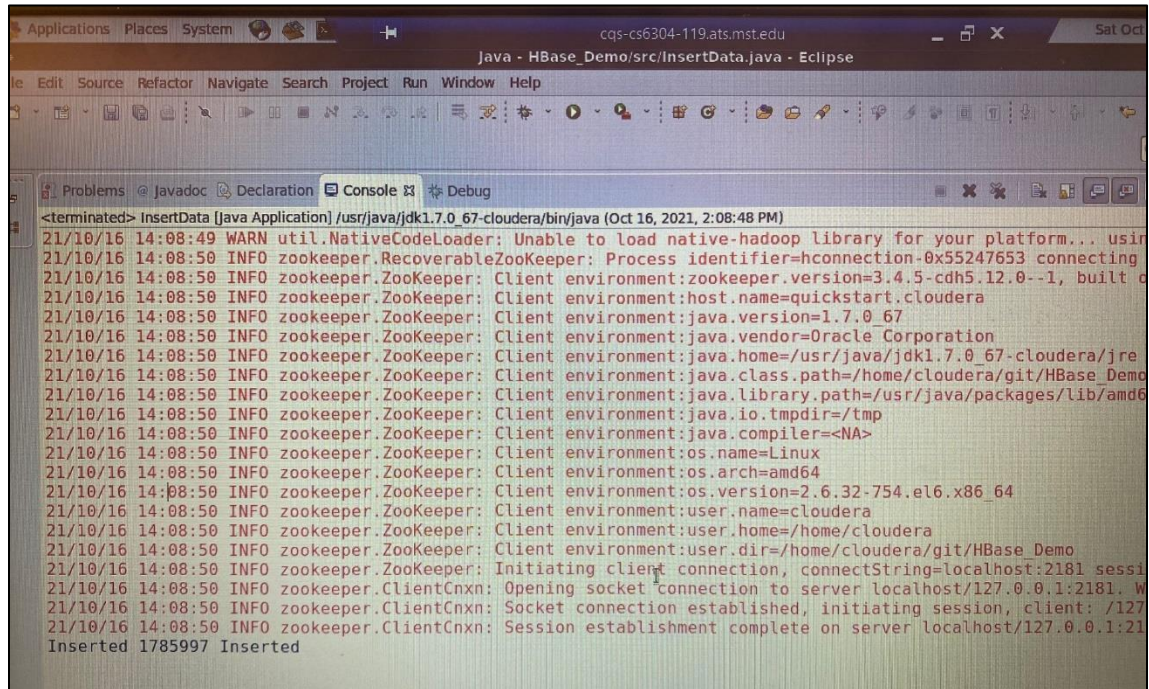
Download and Insert this dataset to a HBase table named "GameTable":

- Run this file to insert the data from .tsv file into the table



InsertData.java

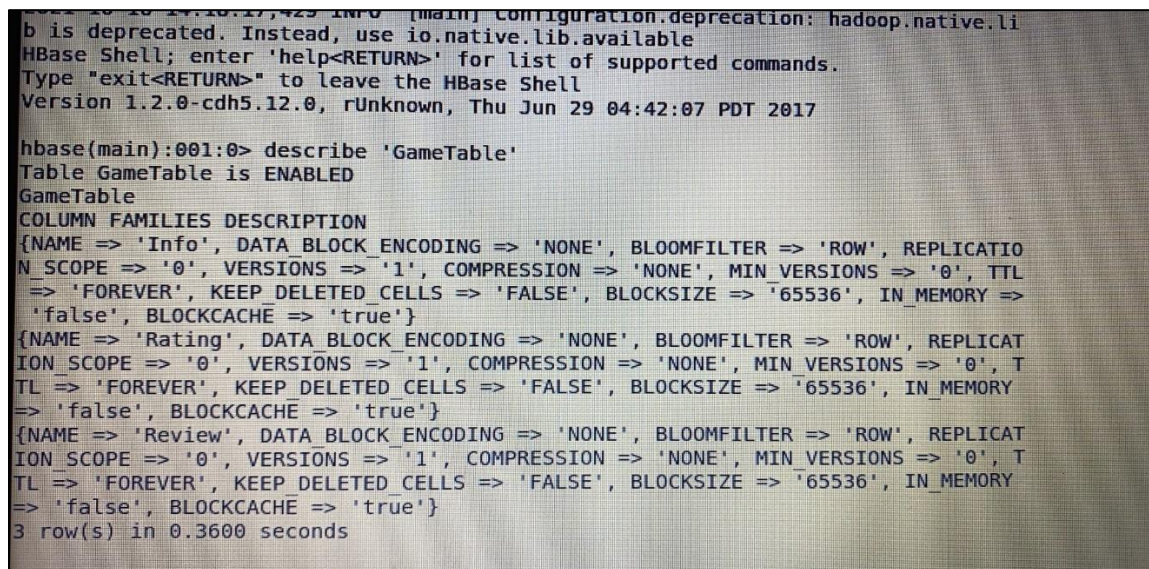
Data Inserted: 1785997 rows



```
<terminated> InsertData [Java Application] /usr/java/jdk1.7.0_67-cloudera/bin/java (Oct 16, 2021, 2:08:48 PM)
21/10/16 14:08:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java-only ones to communicate
21/10/16 14:08:50 INFO zookeeper.RecoverableZooKeeper: Process identifier=hconnection-0x55247653 connecting
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:zookeeper.version=3.4.5-cdh5.12.0--1, built w
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:host.name=quickstart.cloudera
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.version=1.7.0_67
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.vendor=Oracle Corporation
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.home=/usr/java/jdk1.7.0_67-cloudera/jre
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.class.path=/home/cloudera/git/HBase_Demo
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.library.path=/usr/java/packages/lib/amd64
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.io.tmpdir=/tmp
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:java.compiler=<NA>
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:os.name=Linux
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:os.arch=amd64
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:os.version=2.6.32-754.el6.x86_64
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:user.name=cloudera
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:user.home=/home/cloudera
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Client environment:user.dir=/home/cloudera/git/HBase_Demo
21/10/16 14:08:50 INFO zookeeper.ZooKeeper: Initiating client connection, connectString=localhost:2181 sessi
21/10/16 14:08:50 INFO zookeeper.ClientCnxn: Opening socket connection to server localhost/127.0.0.1:2181. W
21/10/16 14:08:50 INFO zookeeper.ClientCnxn: Socket connection established, initiating session, client: /127
21/10/16 14:08:50 INFO zookeeper.ClientCnxn: Session establishment complete on server localhost/127.0.0.1:21
Inserted 1785997 Inserted
```

Command: Describe

hbase(main):001:0> describe 'GameTable'



```
hbase(main):001:0> describe 'GameTable'
Table GameTable is ENABLED
GameTable
COLUMN FAMILIES DESCRIPTION
{NAME => 'Info', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATIO
N_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL
=> 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY =>
'false', BLOCKCACHE => 'true'}
{NAME => 'Rating', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICAT
ION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', T
TL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
=> 'false', BLOCKCACHE => 'true'}
{NAME => 'Review', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICAT
ION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', T
TL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
=> 'false', BLOCKCACHE => 'true'}
3 row(s) in 0.3600 seconds
```


Command: Scan, Limit

hbase(main):004:0> scan 'GameTable', {'LIMIT'=>5}

```
10000099_R38GX1BSE6I column=Rating:star_rating, timestamp=1634418631509, value=
3LM \x00\x00\x00\x05
10000099_R38GX1BSE6I column=Rating:total_votes, timestamp=1634418631509, value=
3LM \x00\x00\x00\x00
10000099_R38GX1BSE6I column=Review:review_body, timestamp=1634418631509, value=
3LM I love this charger. Works perfect and charges way faster
than the use method
10000099_R38GX1BSE6I column=Review:review_headline, timestamp=1634418631509, va
3LM lue=Five Stars
10000124_R312DHUSFZP column=Info:marketplace, timestamp=1634418967941, value=US
Q0P
10000124_R312DHUSFZP column=Info:verified_purchase, timestamp=1634418967941, va
Q0P lue=Y
10000124_R312DHUSFZP column=Rating:helpful_votes, timestamp=1634418967941, valu
Q0P e=\x00\x00\x00\x00
10000124_R312DHUSFZP column=Rating:star_rating, timestamp=1634418967941, value=
Q0P \x00\x00\x00\x02
10000124_R312DHUSFZP column=Rating:total_votes, timestamp=1634418967941, value=
Q0P \x00\x00\x00\x00
10000124_R312DHUSFZP column=Review:review_body, timestamp=1634418967941, value=
Q0P Sadly i was one of the people who got a bad unit. The char
ging port wasnt even charging my controller without me wat
ching it and holding it down just so the contact needles h
it the battery :\x5C
10000124_R312DHUSFZP column=Review:review_headline, timestamp=1634418967941, va
Q0P lue=Not worth the buy
5 row(s) in 0.2470 seconds
```

Part 2:

Command: Alter to change versions

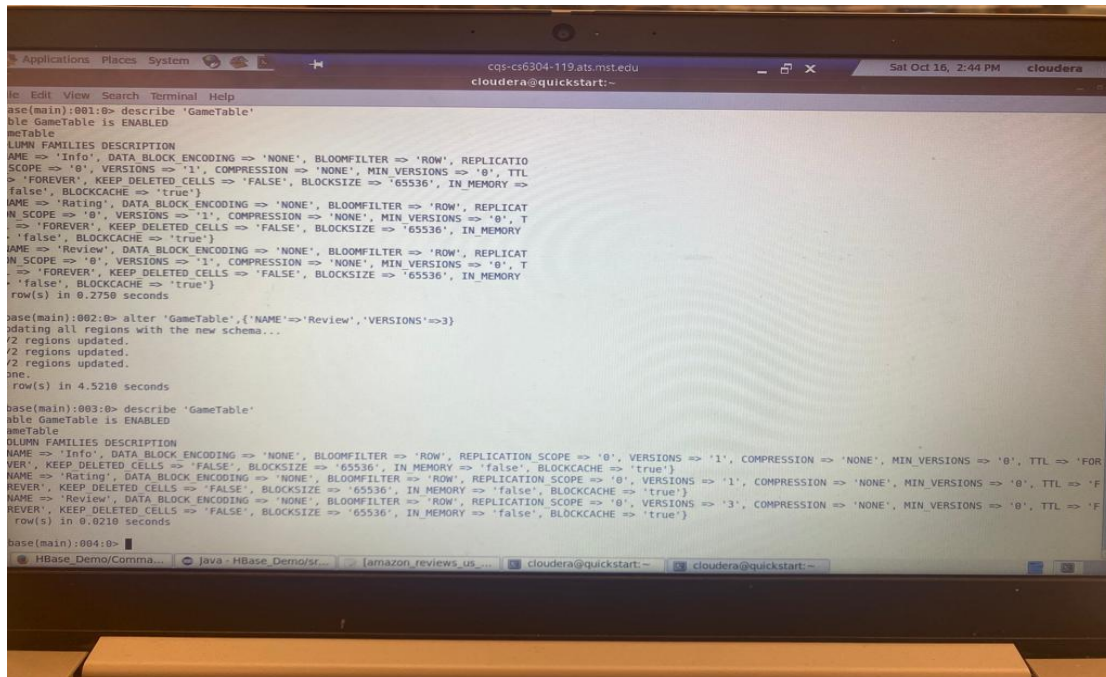
hbase(main):017:0> alter 'GameTable', {'NAME'=>'Review', 'VERSIONS'=>3}

```
File Edit View Search Terminal Help
cloudera@quickstart:~$ hbase shell
2021-10-16 14:40:20,161 INFO [main] Configuration.deprecation: hadoop.native.li
b is deprecated. Instead, use io.native.lib.available
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type 'exit<RETURN>' to leave the HBase Shell
Version 1.2.0-cdh5.12.0, rUnKnown, Thu Jun 29 04:42:07 PDT 2017

hbase(main):001:0> describe 'GameTable'
Table GameTable is ENABLED
GameTable
COLUMN FAMILIES DESCRIPTION
(NAME => 'Info', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATIO
N_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL
=> 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY =>
'false', BLOCKCACHE => 'true')
(NAME => 'Rating', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICAT
ION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', T
TL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
=> 'false', BLOCKCACHE => 'true')
(NAME => 'Review', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICAT
ION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', T
TL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
=> 'false', BLOCKCACHE => 'true')
3 row(s) in 0.2750 seconds

hbase(main):002:0> alter 'GameTable', {'NAME'=>'Review', 'VERSIONS'=>3}
Updating all regions with the new schema...
0/2 regions updated.
1/2 regions updated.
2/2 regions updated.
Done.
0 row(s) in 4.5210 seconds
```

hbase(main):001:0> describe 'GameTable' => to show the change in versions



```
hbase(main):001:0> describe 'GameTable'
describe GameTable is ENABLED
GameTable
COLUMN FAMILIES DESCRIPTION
NAME => 'Info', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'
NAME => 'Rating', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'
NAME => 'Review', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'
row(s) in 0.2750 seconds

hbase(main):002:0> alter 'GameTable',{ 'NAME'=>'Review','VERSIONS'=>3}
Updating all regions with the new schema...
2 regions updated.
2 regions updated.
2 regions updated.
one.
row(s) in 4.5210 seconds

hbase(main):003:0> describe 'GameTable'
describe GameTable is ENABLED
GameTable
COLUMN FAMILIES DESCRIPTION
NAME => 'Info', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'
NAME => 'Rating', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'
NAME => 'Review', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '3', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'
row(s) in 0.0210 seconds

hbase(main):004:0>
```

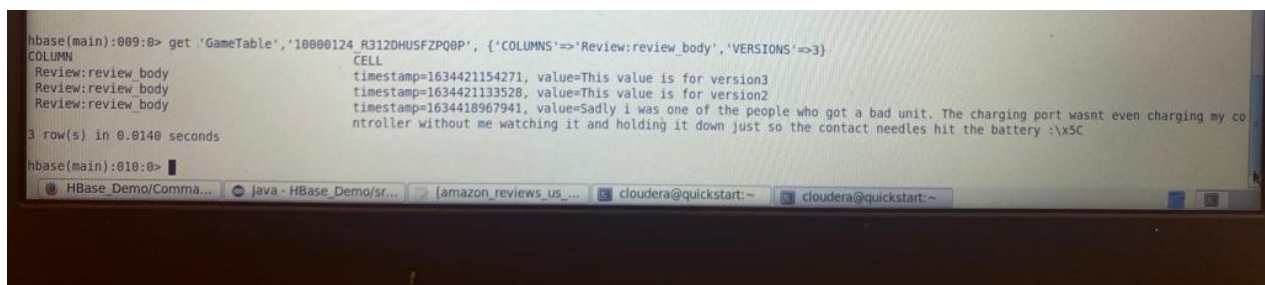
Random row to put additional 2 different 'review_body':

hbase(main):017:0> put 'GameTable','10000124_R312DHUSFZPQ0P','Review:review_body','This value is for version2'

hbase(main):017:0> put 'GameTable','10000124_R312DHUSFZPQ0P','Review:review_body','This value is for version3'

hbase(main):017:0>

get 'GameTable','10000124_R312DHUSFZPQ0P',{ 'COLUMNS'=>'Review:review_body','VERSIONS'=>3}



```
hbase(main):009:0> get 'GameTable','10000124_R312DHUSFZPQ0P',{ 'COLUMNS'=>'Review:review_body','VERSIONS'=>3}
COLUMN CELL
Review:review_body timestamp=1634421154271, value=This value is for version3
Review:review_body timestamp=1634421133528, value=This value is for version2
Review:review_body timestamp=1634418967941, value=Sadly I was one of the people who got a bad unit. The charging port wasnt even charging my controller without me watching it and holding it down just so the contact needles hit the battery :\x5C
3 row(s) in 0.0140 seconds

hbase(main):010:0>
```

Find the 'review_bodys' that have the word 'awesome'. => 102690 rows

hbase(main):017:0>


```
scan 'GameTable', {'COLUMNS'=>'Review:review_body', 'FILTER'=>"ValueFilter(=, 'substring:awesome')"}}
```

```

column=Review:review_body, timestamp=163441947692, value=This is the best Call of duty game so far, since Modern Warfare 1.
The game is better than MW2, graphics wise its the same, a little bit better tho. Multiplayer online its awesome, and this
is coming from a gamer that played Call of Duty 4 on promod leagues competitively for 3 years. My PC consist on: Intel i5
750 Cpu, Sapphire HD 5850, 4 GB G.Skill ram, Samsung 2232r2 lcd monitor, Coolermaster HAF 922 case. Game running at 1680
x 1050, 4x aa, extra settings, getting 70-80 FPS on single player campaign, and 60 to 125 FPS(capped by console command)dur
ing online gameplay. Forgive all the lag and freezes madness, that was because they configured wrong the settings on the
servers the first day, problem solved since day 2. In liking the game more and more every time i play, just waiting to the m
od tools to be released, for some promod action and the game will be perfect!! PD: Do not take into account the noob hate
rs reviewers of the game, if u like FPS and Call of Duty, buy the game, you wont regret it.
column=Review:review_body, timestamp=1634419157414, value=Got it properly packaged no problems put it to use right away work
s great does exactly what it advertised all and a great item for price and quality is awesome
column=Review:review_body, timestamp=1634418824104, value=I will begin this review saying that i normally don't write review
. I do read alot of them and this is my first time writing one. I have played every forza game up to this point, starting wi
th Forza Motorsport on the original xbox. The games were definitely fun in but i always wondered what would forza would be l
ike if there was an open world. My dream became a reality when Forza horizon was released on the 360. I loved that game, it
was definitely new and fresh for the forza series. However for me the game was short lived because they wasn't many things
to do. As time went on and the xbox one was released i wondered again if a Forza horizon would be released and here we are t
oday.<br /><br />Initial Impression: Upon starting the game, they give you this preview that for me sent chills down my spin
e. You can literally close your eyes and the narrator will send you off into a deep thought about cars, racing and sportin
g. After that is done they throw you in a huracan and off you go to the horizon festival. Let me tell you, the graphics on thi
s game are very nice. You can see reflections off the windshield, dirt as you drive off road and scratches off the things yo
u hit. Interiors are called are wonderful to look at as well. In addition to that, the sounds of the cars are really close t
o how they are in reality.<br /><br />Differences between it's younger brother: Forza horizon had a totally different scenar
y, considering the map was based off of Colorado. It was on the smaller side and there wasn't much free roam off the actual
road. Here comes Horizon 2 with a map that is probably double the size of the original and based off of Italy and France. Th
is map is a lot larger with plenty of different landscape that will blow your mind away. Also the car selection, in my opini
on, is a lot better this time around. They have something for everyone. Plus you can fine tune your car, which was impossib
le in the last games.<br /><br />Gameplay: Gameplay is great. Lots of different races including sprint, circuit, cross country
, trail races and showcases. Modifying the cars in game is awesome. Painting them are fun as well. The road trips add a diff
erent sense into the game in where it doesn't feel like a video game anymore.<br /><br />The bad: There are some sound glitch
es in the game relating to the cars. One car doesn't have any sound at all and i have came across like one or two graphical e
rrors. Other then that i don't think anything else is bad.<br /><br />Overall: The game is amazing. From the introduction, t
o the cars, to the scenery, the developers did extremely well on this. It actually made me re-buy the Xbox one. If you can g
et past the slight glitches in the game play, this game will keep you entertained for hours. If you want an even better time
grab your friends and go for a road trip. Give the game a shot. If your a car guy like you, there is no reason why you shou
ldn't like it.
182690 row(s) in 71.5370 seconds
hbase(main):011:0>

```

Find the 'review_headlines' that have any characters apart from alphanumerical characters (use regex)

=> 1798 rows

hbase(main):017:0>

```
scan 'GameTable', {'COLUMNS'=>'Review:review_headline', 'FILTER'=>"ValueFilter(=, 'regexstring:^(^a-zA-Z0-9)*$')"}}
```

```

column=Review:review_headline, timestamp=1634419132773, value=
column=Review:review_headline, timestamp=1634418628696, value=
column=Review:review_headline, timestamp=1634418678074, value=
column=Review:review_headline, timestamp=1634418687931, value=
column=Review:review_headline, timestamp=1634419861466, value=
column=Review:review_headline, timestamp=1634419865989, value=
column=Review:review_headline, timestamp=1634419874688, value=
column=Review:review_headline, timestamp=1634419933662, value=
column=Review:review_headline, timestamp=1634419933884, value=
column=Review:review_headline, timestamp=1634419218805, value=
column=Review:review_headline, timestamp=1634418932549, value=
column=Review:review_headline, timestamp=1634419596429, value=
column=Review:review_headline, timestamp=1634418804170, value=
column=Review:review_headline, timestamp=1634419160609, value=
column=Review:review_headline, timestamp=1634419167504, value=
column=Review:review_headline, timestamp=1634419140608, value=
column=Review:review_headline, timestamp=1634419197980, value=
column=Review:review_headline, timestamp=1634419195105, value=
column=Review:review_headline, timestamp=1634418623796, value=
column=Review:review_headline, timestamp=1634419891229, value=
column=Review:review_headline, timestamp=1634418731347, value=
column=Review:review_headline, timestamp=1634418648773, value=
column=Review:review_headline, timestamp=1634419077316, value=
column=Review:review_headline, timestamp=1634419198506, value=
column=Review:review_headline, timestamp=1634419198509, value=
column=Review:review_headline, timestamp=1634419198509, value=
column=Review:review_headline, timestamp=1634419614841, value=
column=Review:review_headline, timestamp=1634419248005, value=
column=Review:review_headline, timestamp=1634419248034, value=
column=Review:review_headline, timestamp=1634419084197, value=
column=Review:review_headline, timestamp=1634419082150, value=
column=Review:review_headline, timestamp=1634418595059, value=
column=Review:review_headline, timestamp=163441944250, value=
798 row(s) in 7.6760 seconds
hbase(main):012:0>

```

Part 3:

- i) Find reviews have 'star_rating' equal to 5: 1026924 rows

```
hbase(main):017:0> scan 'GameTable', {'COLUMNS'=>'Rating:star_rating',  
'FILTER'=>"SingleColumnValueFilter('Rating', 'star_rating', '=', 'binary:\x00\x00\x00\x05')"} }
```

```
Applications Places System
File Edit View Search Terminal Help
cslucloudera@quickstart:~
9998786 R2FH0P57F04PZQ column=Rating:star rating, timestamp=1634418991012, value=\x00\x00\x00\x05
9998726 R3GV7MIH1JR2Y8 column=Rating:star rating, timestamp=1634419599898, value=\x00\x00\x00\x05
9998761 RTLRA4HUZY5KZSO column=Rating:star rating, timestamp=1634418767935, value=\x00\x00\x00\x05
9998943 RIAT0VTACSN176Q column=Rating:star rating, timestamp=1634419442359, value=\x00\x00\x00\x05
9998943 R8GVY3H1USRUY column=Rating:star rating, timestamp=1634419564964, value=\x00\x00\x00\x05
9998954 RJ3KJN4F8JROAL column=Rating:star rating, timestamp=1634419495426, value=\x00\x00\x00\x05
9998961 R1C90S5191D0 column=Rating:star rating, timestamp=1634419567773, value=\x00\x00\x00\x05
9998961 RJ38H4JWTSYAN column=Rating:star rating, timestamp=1634419340044, value=\x00\x00\x00\x05
9998961 RY5KQDV0TSP5S1 column=Rating:star rating, timestamp=1634419417692, value=\x00\x00\x00\x05
9998994 R21A8I63RS0045 column=Rating:star rating, timestamp=1634419335352, value=\x00\x00\x00\x05
9998912 R3I8IAB5DM08052 column=Rating:star rating, timestamp=1634418782285, value=\x00\x00\x00\x05
9998934 R2M7YRDZ0NL9N5 column=Rating:star rating, timestamp=1634418563224, value=\x00\x00\x00\x05
9999110 R169F653TCTM0Z column=Rating:star rating, timestamp=1634418808085, value=\x00\x00\x00\x05
9999185 R6V03JZ7281LG column=Rating:star rating, timestamp=1634419007369, value=\x00\x00\x00\x05
9999311 R127I6X6LM0G6Z5 column=Rating:star rating, timestamp=1634418964925, value=\x00\x00\x00\x05
9999368 R1VEKHPW10A719 column=Rating:star rating, timestamp=1634418553727, value=\x00\x00\x00\x05
9999368 R3A5S0535K8KRS column=Rating:star rating, timestamp=1634418662472, value=\x00\x00\x00\x05
9999415 R18I103P4033AH column=Rating:star rating, timestamp=1634418918066, value=\x00\x00\x00\x05
9999427 RC1NTU0823K70 column=Rating:star rating, timestamp=1634419862210, value=\x00\x00\x00\x05
9999443 R2ZB2RNZ4UZZTV column=Rating:star rating, timestamp=1634419134725, value=\x00\x00\x00\x05
9999585 R27MPF3ZWHJ5Z0 column=Rating:star rating, timestamp=1634418714738, value=\x00\x00\x00\x05
9999512 R12YUCL44500C column=Rating:star rating, timestamp=1634419339722, value=\x00\x00\x00\x05
9999529 RYCU4T1V1511 column=Rating:star rating, timestamp=1634418991641, value=\x00\x00\x00\x05
9999577 R2VY164A0F3RCL column=Rating:star rating, timestamp=1634418791449, value=\x00\x00\x00\x05
9999577 R3FMO25778Y9 column=Rating:star rating, timestamp=1634418790812, value=\x00\x00\x00\x05
9999628 RLTH0PFPYU01YM column=Rating:star rating, timestamp=1634419019536, value=\x00\x00\x00\x05
9999645 R2UG0AG4LTLAW8 column=Rating:star rating, timestamp=1634418718923, value=\x00\x00\x00\x05
9999668 R2EZ27003234Y column=Rating:star rating, timestamp=1634418632667, value=\x00\x00\x00\x05
9999796 R30F8AG3J4YAP0 column=Rating:star rating, timestamp=1634419157414, value=\x00\x00\x00\x05
9999834 R3J8E3R30355CSC column=Rating:star rating, timestamp=1634418824104, value=\x00\x00\x00\x05
9999872 R26HJ337VM6C69 column=Rating:star rating, timestamp=1634418898621, value=\x00\x00\x00\x05
9999897 R1Q0CNDM1H5UBA9 column=Rating:star rating, timestamp=1634418863963, value=\x00\x00\x00\x05
9999933 R1F5MEPEFG0T25 column=Rating:star rating, timestamp=1634418549196, value=\x00\x00\x00\x05
9999933 R24LKG119U0ZDE column=Rating:star rating, timestamp=1634418550772, value=\x00\x00\x00\x05
9999981 R167Y5X48E5ORV column=Rating:star rating, timestamp=1634419489923, value=\x00\x00\x00\x05
1026924 row(s) in 5.3880 seconds
hbase(main):017:0> scan 'GameTable', {'COLUMNS'=>['Rating:star rating'], 'FILTER'=>['SingleColumnValueFilter('Rating', 'star rating',=, 'binary:\x00\x00\x00\x05')']}
```

- ii) Find the average 'helpful_vote' in the dataset.

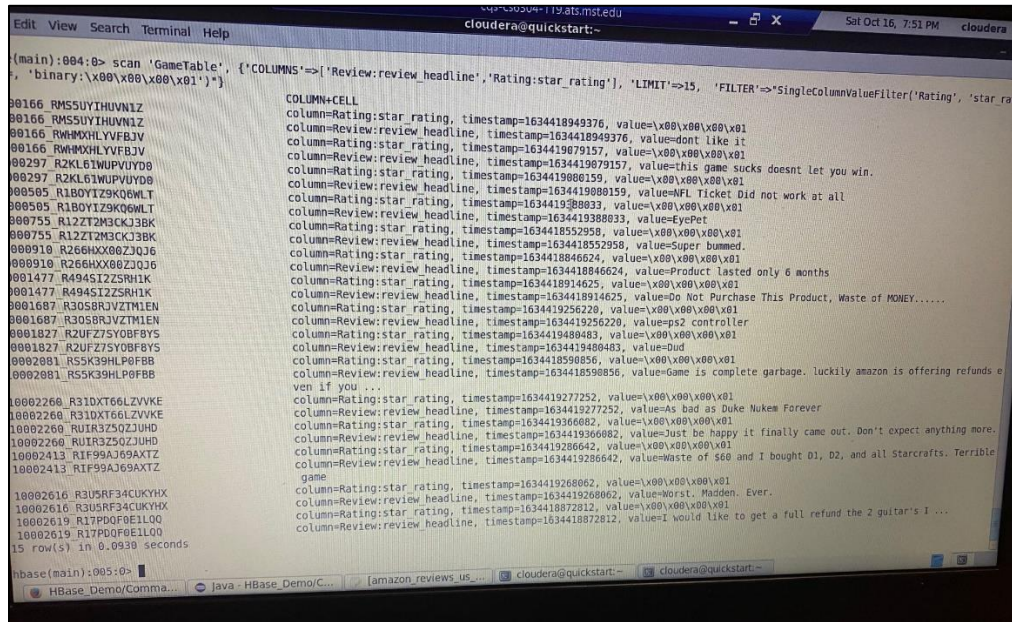


Summation.java

[illegible]

- iii) Show the 'review_headlines' that got 1 'star_rating'


```
hbase(main):004:0> scan 'GameTable',  
{'COLUMNS'=>['Review:review_headline','Rating:star_rating'],'LIMIT'=>15,  
'FILTER'=>"SingleColumnValueFilter('Rating', 'star_rating', =, 'binary:\x00\x00\x00\x01')"}  
15 row(s) in 0.0930 seconds
```



The screenshot shows a terminal window titled "cloudera@quickstart:~" with a timestamp of "Sat Oct 16, 7:51 PM". The terminal displays the execution of an HBase scan command and its output. The command is: `{'COLUMNS'=>['Review:review_headline','Rating:star_rating'],'LIMIT'=>15, 'FILTER'=>"SingleColumnValueFilter('Rating', 'star_rating', =, 'binary:\x00\x00\x00\x01')"}`. The output shows 15 rows of data, each with a key (e.g., "0002260_R31DX766LVVKE") and a value (e.g., "column=Rating:star rating, timestamp=1634418949376, value=\x00\x00\x00\x01"). The terminal also shows the command prompt "hbase(main):005:0>" and the file "HBase_Demo/Comma..." in the background.

```
{'COLUMNS'=>['Review:review_headline','Rating:star_rating'],'LIMIT'=>15, 'FILTER'=>"SingleColumnValueFilter('Rating', 'star_rating', =, 'binary:\x00\x00\x00\x01')"}  
COLUMN+CELL  
column=Rating:star rating, timestamp=1634418949376, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634418949376, value=dont like it  
column=Rating:star_rating, timestamp=1634419079157, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419079157, value=this game sucks doesnt let you win.  
column=Rating:star rating, timestamp=1634419080159, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419080159, value=AFL Ticket Did not work at all  
column=Rating:star rating, timestamp=1634419380833, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419380833, value=EyePet  
column=Rating:star rating, timestamp=1634418552958, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634418552958, value=Super bunned.  
column=Rating:star rating, timestamp=1634418846624, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634418846624, value=Product lasted only 6 months  
column=Rating:star rating, timestamp=1634418914625, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634418914625, value=Do Not Purchase This Product, Waste of MONEY.....  
column=Rating:star rating, timestamp=1634419256220, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419256220, value=ps2 controller  
column=Rating:star rating, timestamp=1634419480483, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419480483, value=Dud  
column=Rating:star rating, timestamp=1634418590856, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634418590856, value=Game is complete garbage. Luckily amazon is offering refunds e  
ven if you ...  
column=Rating:star rating, timestamp=1634419277252, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419277252, value=As bad as Duke Nukem Forever  
column=Rating:star rating, timestamp=1634419366082, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419366082, value=Just be happy it finally came out. Don't expect anything more.  
column=Rating:star rating, timestamp=1634419286642, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419286642, value=Waste of $60 and I bought D1, D2, and all Starcrafts. Terrible  
game  
column=Rating:star rating, timestamp=1634419268062, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634419368062, value=Worst Madden. Ever.  
column=Rating:star rating, timestamp=1634418872812, value=\x00\x00\x00\x01  
column=Review:review_headline, timestamp=1634418872812, value=I would like to get a full refund the 2 guitar's I ...  
hbase(main):005:0>  
HBase_Demo/Comma... Java - HBase_Demo/C... [amazon_reviews_US... cloudera@quickstart:~ cloudera@quickstart:~
```