499100

Assignment#2

1.

```
[baoxin.l@ip-172-31-93-164 ~]$ wget http://54.161.173.97/dataset/youtube.csv.bz2
--2021-11-11 16:36:40--  http://54.161.173.97/dataset/youtube.csv.bz2
Connecting to 54.161.173.97:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 2683136 (2.6M) [application/x-bzip2]
Saving to: 'youtube.csv.bz2'

100%[==============================================================================================>] 2,683,136   --.-K/s   in 0.009s

2021-11-11 16:36:40 (294 MB/s) - 'youtube.csv.bz2' saved [2683136/2683136]

[baoxin.l@ip-172-31-93-164 ~]$ bzip2 -d youtube.csv.bz2
[baoxin.l@ip-172-31-93-164 ~]$ head youtube.csv
video_id,trending_date,title,channel_title,category_id,publish_time,tags,views,likes,dislikes,comment_count,thumbnail_link,comments_disabled,ratings_disabled,video_error_or_removed
2kyS6SvSYSE,17.14.11,WE WANT TO TALK ABOUT OUR MARRIAGE,CaseyNeistat,22,2017-11-13T17:13:01.000Z,SHANtell martin,748374,57527,2966,15954,https://i.ytimg.com/vi/2kyS6SvSYSE/default.jpg,FALSE,FALSE,FALSE
1ZAPwfrtAFY,17.14.11,The Trump Presidency: Last Week Tonight with John Oliver (HBO),LastWeekTonight,24,2017-11-13T07:30:00.000Z, last week tonight trump presidency  last week tonight donald trump   john oliv
er trump   donald trump  ,2418783,97185,6146,12703,https://i.ytimg.com/vi/1ZAPwfrtAFY/default.jpg,FALSE,FALSE,FALSE
5qpjK5DgCt4,17.14.11,Racist Superman | Rudy MancusoKing Bach & Lele Pons,Rudy Mancuso,23,2017-11-12T19:05:24.000Z, racist superman   rudy   mancuso   king   bach   racist   superman   love   rudy mancuso poo
bear black white official music video   iphone x by pineapple   lelepons   hannahstocking   rudymancuso   inanna   anwar   sarkis   shots   shotsstudios   alesso   anitta   brazil   Getting My Driver's Licen
se | Lele Pons  ,3191434,146033,5339,8181,https://i.ytimg.com/vi/5qpjK5DgCt4/default.jpg,FALSE,FALSE,FALSE
puqaWrEC7tY,17.14.11,Nickelback Lyrics: Real or Fake?,Good Mythical Morning,24,2017-11-13T11:00:04.000Z, rhett and link  gmm   good mythical morning   rhett and link good mythical morning   good mythical mor
ning rhett and link   mythical morning   Season 12   nickelback lyrics   nickelback lyrics real or fake   nickelback   nickelback songs   nickelback song   rhett link nickelback   gmm nickelback   lyrics (we
bsite category)   nickelback (musical group)   rock   music   lyrics   chad kroeger   canada   music (industry)   mythical   gmm challenge   comedy   funny   challenge  ,343168,10172,666,2146,https://i.ytimg
.com/vi/puqaWrEC7tY/default.jpg,FALSE,FALSE,FALSE
d380meD0W0M,17.14.11,I Dare You: GOING BALD!?,nigahiga,24,2017-11-12T18:01:41.000Z, ryan  higa   higatv   nigahiga   i dare you   idy   rhpc   dares   no truth   comments   comedy   funny   stupid   fail  ,2
095731,132235,1989,17518,https://i.ytimg.com/vi/d380meD0W0M/default.jpg,FALSE,FALSE,FALSE
gHZ1Qz0KiKM,17.14.11,2 Weeks with iPhone X,iJustine,28,2017-11-13T19:07:23.000Z, ijustine  week with iPhone X   iphone x   apple   iphone   iphone x review   iphone x unboxing  ,119180,9763,511,1434,https://
i.ytimg.com/vi/gHZ1Qz0KiKM/default.jpg,FALSE,FALSE,FALSE
39idvpPF7NQ,17.14.11,Roy Moore & Jeff Sessions Cold Open - SNL,Saturday Night Live,24,2017-11-12T05:37:17.000Z, SNL   Saturday Night Live   SNL Season 43   Episode 1730   Tiffany Haddish   Roy Moore   Mikey D
ay   Mike Pence   Beck Bennett   Jeff Sessions   Kate McKinnon   s43   s43e5   episode 5   live   new york   comedy   sketch   funny   hilarious   late night   host   music   guest   laugh   impersonation
actor   improv   musician   comedian   actress   If Loving You Is Wrong   Oprah Winfrey   OWN   Girls Trip   The Carmichael Show   Keanu   Taylor Swift   Reputation   Look What You Made Me Do   ready for it?
cold open  ,2103417,15993,2445,1970,https://i.ytimg.com/vi/39idvpPF7NQ/default.jpg,FALSE,FALSE,FALSE
nc99ccSXST0,17.14.11,5 Ice Cream Gadgets put to the Test,CrazyRussianHacker,28,2017-11-12T21:50:37.000Z, 5 Ice Cream Gadgets  Ice Cream   Cream Sandwich Maker   gadgets   gadget review   review   unboxing
kitchen gadgets   Gadgets put to the Test   testing   10 Kitchen Gadgets   7 Camping Coffee Gadgets   10 Kitchen Gadgets put to the Test  ,817732,23663,778,3432,https://i.ytimg.com/vi/nc99ccSXST0/default.jpg
,FALSE,FALSE,FALSE
jr9QtXwC9vc,17.14.11,The Greatest Showman | Official Trailer 2 [HD] | 20th Century FOX,20th Century Fox,1,2017-11-13T14:00:23.000Z, Trailer  Hugh Jackman   Michelle Williams   Zac Efron   Zendaya   Rebecca F
erguson   pasek and paul   la la land   moulin rouge   high school musical   hugh jackman musical   zac efron musical   musical   the greatest showman   greatest showman   Michael Gracey   P.T. Barnum   Barn
um and Bailey   Barnum Circus   Barnum and Bailey Circus   20th century fox   greatest showman trailer   trailer   official trailer   the greatest showman trailer   logan   Benj Pasek   Justin Paul  ,826059,
3543,119,340,https://i.ytimg.com/vi/jr9QtXwC9vc/default.jpg,FALSE,FALSE,FALSE
[baoxin.l@ip-172-31-93-164 ~]$
```

2.

```
[baoxin.l@ip-172-31-93-164 ~]$ cut -d "," -f 3 youtube.csv | sort | uniq -c| sort -rn| head -1
     30 WE MADE OUR MOM CRY...HER DREAM CAME TRUE!
[baoxin.l@ip-172-31-93-164 ~]$
```

3.

```
[baoxin.l@ip-172-31-93-164 ~]$ sort -rnk9 -t "," youtube.csv | cut -d ","  -f 3,9 | head -50 | cut -d "," -f 1| sort|
uniq -c
      5 Ariana Grande - No Tears Left To Cry
     13 BTS (방탄소년단) 'FAKE LOVE' Official MV
     20 Childish Gambino - This Is America (Official Video)
      6 So Sorry.
      6 YouTube Rewind: The Shape of 2017 | #YouTubeRewind
[baoxin.l@ip-172-31-93-164 ~]$
```

4.

```
[baoxin.l@ip-172-31-93-164 ~]$ sort -k 13 -t "," youtube.csv | cut -d "," -f 1,3,13 | uniq | cut -d "," -f 3 | uniq -
c
      1 comments_disabled
   6688 FALSE
    113 TRUE
[baoxin.l@ip-172-31-93-164 ~]$
```

5.

```
[baoxin.l@ip-172-31-93-164 ~]$ head -1 youtube.csv | sed 's/[^,]//g' | wc -c
15
```

6.

```
[baoxin.l@ip-172-31-93-164 ~]$ sort -rnk 1 -t "," youtube.csv | cut -d "," -f 1 | uniq -c | head -1
      8 8857-yQOZ2o
[baoxin.l@ip-172-31-93-164 ~]$ grep "8857-yQOZ2o" youtube.csv | cut -d "," -f 1,3,6 | uniq
8857-yQOZ2o,2018 Bugatti Chiron -- Now in Forza Motorsport 7!,2018-02-15T20:29:34.000Z
[baoxin.l@ip-172-31-93-164 ~]$
```

7.

```
[baoxin.l@ip-172-31-93-164 ~]$ grep "8857-yQOZ2o" youtube.csv | head -8| cut -d "," -f 11| paste -s -d + -| bc
5052
```

8.

```
[baoxin.l@ip-172-31-93-164 ~]$ grep -e "android" youtube.csv | wc -l
133
[baoxin.l@ip-172-31-93-164 ~]$
```

9.

```
[baoxin.l@ip-172-31-93-164 ~]$ cut -d "," -f 1,2,3,4,5,6,7,8,9,10,11,12 youtube.csv| head -1 >> youtube_sample.csv
[baoxin.l@ip-172-31-93-164 ~]$ wc -l youtube.csv
40935 youtube.csv
[baoxin.l@ip-172-31-93-164 ~]$ tail -n 40934 youtube.csv | shuf -n 20 >> youtube_sample.csv
[baoxin.l@ip-172-31-93-164 ~]$
```

10.

```
[baoxin.l@ip-172-31-93-164 ~]$ ls -l youtube.csv
-rw-rw-r-- 1 baoxin.l baoxin.l 19701845 Oct 15 19:45 youtube.csv
[baoxin.l@ip-172-31-93-164 ~]$ chmod go-rw youtube.csv
[baoxin.l@ip-172-31-93-164 ~]$ ls -l youtube.csv
-rw------- 1 baoxin.l baoxin.l 19701845 Oct 15 19:45 youtube.csv
[baoxin.l@ip-172-31-93-164 ~]$
```

```
[baoxin.l@ip-172-31-93-164 ~]$ history | grep -i nano
     6  nano test.txt
    25  nano lab1.txt
    41  nano 499100.txt
    78  nano /tmp/
   113  sort -rnk9 -t "," youtube.csv | cut -d "," -f 1| uniq -c | grep -i nano
   139  ls -l youtube.csv | grep -i nano
   140  history | grep -i nano
[baoxin.l@ip-172-31-93-164 ~]$
```