analyze

August 9, 2020

[77]:
```
class rawData:
    def __init__(self, aid, bvid, createdTime, title, length, description,␣
↪play, comment, videoReview,coin,danmaku,favorite,like, reply, share):
        self.aid = aid
        self.bvid = bvid
        self.createdTime = createdTime
        self.title = title
        self.length = length
        self.description = description
        self.play = play
        self.comment = comment
        self.videoReview = videoReview
        self.coin = coin
        self.danmaku = danmaku
        self.favorite = favorite
        self.like = like
        self.reply = reply
        self.share = share
    def __str__(self):
        return str(self.aid) + '\t' + self.bvid + '\t' + str(self.createdTime)␣
↪+ '\t' + self.title + '\t' \
        + self.length + '\t' + self.description + '\t' + str(self.play) + '\t'␣
↪+ str(self.comment) + '\t' \
        + str(self.videoReview) + str(self.coin) + str(self.danmaku) + str(self.
↪favorite )+ '\t' \
        + str(self.like) + '\t' + str(self.reply) + '\t' + str(self.share) +␣
↪'\n'
    def __hash__(self):
        return hash(self.title)
    def __eq__(self, value):
        return self.bvid == value.bvid
class total:
    def __init__(self,key,video,length, play, comment, coin, danmaku, favorite,␣
↪like, reply, share):
        self.key = key
        self.video = video
        self.length = length
```

```
        self.play = play
        self.comment = comment
        self.coin = coin
        self.danmaku = danmaku
        self.favorite = favorite
        self.like = like
        self.reply = reply
        self.share = share
    def __str__(self):
        return "{0}\t{1}\t{2}\t{3}\t{4}\t{5}\t{6}\t{7}\t{8}\t{9}\t{10}\n".
↪format(self.key,self.video,self.length,self.play,self.comment,self.coin,self.
↪danmaku,self.favorite,self.like,self.reply,self.share)



class _time:
    def __init__(self,converted):
        self.h = converted[0]
        self.m = converted[1]
        self.s = converted[2]
        self.m += self.s // 60
        self.s = self.s % 60
        self.h += self.m // 60
        self.m = self.m % 60
    def __str__(self):
        return "{:02d} : {:02d} : {:02d}".format(self.h, self.m, self.s)
    def __repr__(self):
        return str(self)
    def __add__(self, other):
        return _time((self.h + other.h, self.m + other.m, self.s + other.s))
    def __sub__(self, other):
        return _time((self.h - other.h, self.m - other.m, self.s - other.s))
def convert(time_str):
    if time_str == None or time_str == "":
        return "Error"
    else:
        temp = time_str.split(":")
        if len(temp) == 2:
            return _time((0,int(temp[0]),int(temp[1])))
        elif len(temp) == 3:
            return _time((int(temp[0]),int(temp[1]), int(temp[2])))
        else:
            return "error"
```

[78]:
```
print(convert("70:61") + convert("70:61"))
```

```
02 : 22 : 02
```

[79]:
```
import csv

def load(fileDir):
    fp = open(fileDir)
    reader = csv.reader(fp, delimiter='\t')
    header = next(reader)
    temp = []
    for row in reader:
        temp +=␣
 ↪[rawData(row[0],row[1],row[2],row[3],row[4],row[5],row[6],row[7],row[8],row[9],row[10],row[
    return temp
data = load('searched.csv')
```

[80]:
```
key_dict = [[' ', ' ', '  '], [' ', '  '], [' ', '  ', ' '], [' ', '  '],␣
 ↪[' '], [' ', ' ', ' '], [' '], [' '], [' ', '  '], [' ', ' '], [' ',␣
 ↪'  ', ' '], [' ', '  '], [' ', '  '], [' ', '  ', ' '], [' ', '  '],␣
 ↪[' ', '  ',' '], [' ', '  ']]

result = {' ': [], ' ': [], ' ': [], ' ': [], ' ': [], ' ': [], ' ': [],␣
 ↪' ': [], ' ': [], ' ': [], '  ': [], ' ': [], ' ': [], ' ': [] , ' ':␣
 ↪[],'  ': [], ' ': [], ' ': [] }
```

[81]:
```
for row in data:
    for i in key_dict:
        key = i[0]
        for j in i:
            if j in row.title or j in row.description:
                try:
                    result[key] += [row]
                except:
                    pass
```

[82]:
```
for i in result:
    result[i] = list(set(result[i]))
```

[83]:
```
data_set = {' ': None, ' ': None, ' ': None, ' ': None, ' ': None, '  ':␣
 ↪None, ' ': None, ' ': None, ' ': None, ' ': None, '  ': None, ' ': None,␣
 ↪' ': None,' ': None , ' ': None,'  ': None, ' ': None, ' ': None  }
for key in result:
    temp = [key,0,convert('00:00'),0,0,0,0,0,0,0,0]
    for video in result[key]:
        temp[1]+=1
        temp[2]+=convert(video.length)
        temp[3]+=int(video.play)
        temp[4]+=int(video.comment)
        temp[5]+=int(video.coin)
        temp[6]+=int(video.danmaku)
```

```
        temp[7]+=int(video.favorite)
        temp[8]+=int(video.like)
        temp[9]+=int(video.reply)
        temp[10]+=int(video.share)
    data_set[key] =␣
↪total(temp[0],temp[1],temp[2],temp[3],temp[4],temp[5],temp[6],temp[7],temp[8],temp[9],temp[
```

[84]:
```
print(data_set[' '])
```

```
       88      15 : 35 : 31    88102561        322709  3414746 1194938 1220925
6580266 322764  581382
```

[85]:
```
output = open("./final.csv",'w')
output.
↪write('name\tvideo\tlength\tplay\tcomment\tcoin\tdanmaku\tfavorite\tlike\treply\tshare\n')
for host in data_set:
    output.write(str(data_set[host]))
output.close()
```

[ ]:
```

```