

# wrangle\_act

July 21, 2020

## 1 Project : Wrangle and Analyze Data

\*\*\*Data Source: WeRateDogs Twitter data\*\*\*

### 1.0.1 Introduction

The WeRateDogs Twitter archive(`twitter_archive_enhanced.csv`). I Download this file manually from the Udacity Resources.

The tweet image predictions, i.e., what breed of dog (or other object, animal, etc.) is present in each tweet according to a neural network. This file (`image_predictions.tsv`) is hosted on Udacity's servers and I download it programmatically using the Requests library and the following URL: [https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad\\_image-predictions/image-predictions.tsv](https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predictions/image-predictions.tsv)

Each tweet's retweet count and favorite ("like") count at minimum, and any additional data that can be interesting. Using the tweet IDs in the WeRateDogs Twitter archive, the Twitter API for each tweet's JSON data should be query using Python's Tweepy library and store each tweet's entire set of JSON data in a file called `tweet_json.txt` file. Each tweet's JSON data should be written to its own line. Then I read this .txt file line by line into a pandas DataFrame with (at minimum) tweet ID, retweet count, and favorite count.

### 1.0.2 Gathering Data

- The WeRateDogs Twitter archive (on-hand file)
- The tweet image predictions (download using the requestslibrary)
- Each tweet's retweet count and favorite ("like") count (query Twitter API using Tweepy library)

```
In [1]: #Import the packages needed for this project
import pandas as pd
import numpy as np
import requests
import os
import tweepy
import json
import time
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

## Dataset 1: The WeRateDogs Twitter archive (on-hand file provided by udacity.com)

In [2]: *#Load the csv file*

```
archive = pd.read_csv('twitter-archive-enhanced.csv')
```

*#Display the 5 first lines of the dataset*

```
archive.head()
```

```
Out[2]:
```

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	\
0	892420643555336193	NaN	NaN	
1	892177421306343426	NaN	NaN	
2	891815181378084864	NaN	NaN	
3	891689557279858688	NaN	NaN	
4	891327558926688256	NaN	NaN	

	timestamp	\
0	2017-08-01 16:23:56 +0000	
1	2017-08-01 00:17:27 +0000	
2	2017-07-31 00:18:03 +0000	
3	2017-07-30 15:58:51 +0000	
4	2017-07-29 16:00:24 +0000	

	source	\
0	<a href="http://twitter.com/download/iphone" r...	
1	<a href="http://twitter.com/download/iphone" r...	
2	<a href="http://twitter.com/download/iphone" r...	
3	<a href="http://twitter.com/download/iphone" r...	
4	<a href="http://twitter.com/download/iphone" r...	

	text	retweeted_status_id	\
0	This is Phineas. He's a mystical boy. Only eve...	NaN	
1	This is Tilly. She's just checking pup on you...	NaN	
2	This is Archie. He is a rare Norwegian Pouncin...	NaN	
3	This is Darla. She commenced a snooze mid meal...	NaN	
4	This is Franklin. He would like you to stop ca...	NaN	

	retweeted_status_user_id	retweeted_status_timestamp	\
0	NaN	NaN	
1	NaN	NaN	
2	NaN	NaN	
3	NaN	NaN	
4	NaN	NaN	

	expanded_urls	rating_numerator	\
0	https://twitter.com/dog_rates/status/892420643...	13	
1	https://twitter.com/dog_rates/status/892177421...	13	
2	https://twitter.com/dog_rates/status/891815181...	12	
3	https://twitter.com/dog_rates/status/891689557...	13	
4	https://twitter.com/dog_rates/status/891327558...	12	

	rating_denominator	name	doggo	floofer	pupper	puppo
0	10	Phineas	None	None	None	None
1	10	Tilly	None	None	None	None
2	10	Archie	None	None	None	None
3	10	Darla	None	None	None	None
4	10	Franklin	None	None	None	None

## Dataset 2: The tweet image predictions

```
In [3]: # Download tweet image predictions from Udacity server
url = 'https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predicti
response = requests.get(url)

#Write the file with the name image_predictions in tsv format
with open('image_predictions.tsv', 'wb') as file:
    file.write(response.content)

# read the tsv file using pandas
image_prediction = pd.read_csv('image_predictions.tsv', sep='\t')
image_prediction.head()
```

```
Out[3]:
```

	tweet_id	jpg_url	\
0	666020888022790149	https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg	
1	666029285002620928	https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg	
2	666033412701032449	https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg	
3	666044226329800704	https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg	
4	666049248165822465	https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg	

	img_num	p1	p1_conf	p1_dog	p2	\
0	1	Welsh_springer_spaniel	0.465074	True	collie	
1	1	redbone	0.506826	True	miniature_pinscher	
2	1	German_shepherd	0.596461	True	malinois	
3	1	Rhodesian_ridgeback	0.408143	True	redbone	
4	1	miniature_pinscher	0.560311	True	Rottweiler	

	p2_conf	p2_dog	p3	p3_conf	p3_dog
0	0.156665	True	Shetland_sheepdog	0.061428	True
1	0.074192	True	Rhodesian_ridgeback	0.072010	True
2	0.138584	True	bloodhound	0.116197	True
3	0.360687	True	miniature_pinscher	0.222752	True
4	0.243682	True	Doberman	0.154629	True

## Dataset 3: Each tweet's retweet count and favorite ("like") count

```
In [4]: # Twitter API authentication
consumer_key = 'TPc4fkzA3kzkUNRlWTeisVATk'
consumer_secret = 'men0sNd6IbNjpXuvD11KJQVMQdLEUIy8kjjz0uLSFq8N9Sm09ka'
```

```
access_token = '706471922901479424-qa0Ju4lEipTEoIx8NSEurUsxkst28hd'
access_secret = '0QgB0Lx5Lir9kcX0jABmdFiHfXNEADi5IsY1EtflZVi6X'
```

```
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)
```

```
api = tweepy.API(auth)
```

```
In [5]: # Twitter API query
```

```
id_list = archive.tweet_id.tolist()
for tweet_id in id_list:
    try:
        start = time.time()
        status = api.get_status(tweet_id, tweet_mode='extended')
        json_str = json.dumps(status._json)
        with open("tweet_json.txt", "a") as file:
            file.write(json_str)
            file.write('\n')
        end = time.time()
        print(tweet_id, "succeeded in", end-start)
    except:
        print(tweet_id, "error")
```

```
892420643555336193 succeeded in 0.19655418395996094
892177421306343426 succeeded in 0.17999672889709473
891815181378084864 succeeded in 0.1556699275970459
891689557279858688 succeeded in 0.1507861614227295
891327558926688256 succeeded in 0.22413897514343262
891087950875897856 succeeded in 0.2794520854949951
890971913173991426 succeeded in 0.18677330017089844
890729181411237888 succeeded in 0.178450345993042
890609185150312448 succeeded in 0.17640018463134766
890240255349198849 succeeded in 0.14935636520385742
890006608113172480 succeeded in 0.15752243995666504
889880896479866881 succeeded in 0.15045905113220215
889665388333682689 succeeded in 0.15186810493469238
889638837579907072 succeeded in 0.18098855018615723
889531135344209921 succeeded in 0.1784510612487793
889278841981685760 succeeded in 0.19187426567077637
888917238123831296 succeeded in 0.2090442180633545
888804989199671297 succeeded in 0.17664313316345215
888554962724278272 succeeded in 0.16559362411499023
888202515573088257 error
888078434458587136 succeeded in 0.21283984184265137
887705289381826560 succeeded in 0.1658153533935547
887517139158093824 succeeded in 0.15332436561584473
887473957103951883 succeeded in 0.20612287521362305
887343217045368832 succeeded in 0.1824970245361328
```

887101392804085760 succeeded in 0.15956354141235352  
886983233522544640 succeeded in 0.20634126663208008  
886736880519319552 succeeded in 0.17993855476379395  
886680336477933568 succeeded in 0.14924192428588867  
886366144734445568 succeeded in 0.18349885940551758  
886267009285017600 succeeded in 0.15352845191955566  
886258384151887873 succeeded in 0.1885225772857666  
886054160059072513 succeeded in 0.16664481163024902  
885984800019947520 succeeded in 0.18913698196411133  
885528943205470208 succeeded in 0.15644025802612305  
885518971528720385 succeeded in 0.20757031440734863  
885311592912609280 succeeded in 0.19087433815002441  
885167619883638784 succeeded in 0.21589255332946777  
884925521741709313 succeeded in 0.17712664604187012  
884876753390489601 succeeded in 0.18174457550048828  
884562892145688576 succeeded in 0.1846013069152832  
884441805382717440 succeeded in 0.15421557426452637  
884247878851493888 succeeded in 0.15934228897094727  
884162670584377345 succeeded in 0.1804213523864746  
883838122936631299 succeeded in 0.1856982707977295  
883482846933004288 succeeded in 0.1843576431274414  
883360690899218434 succeeded in 0.19221997261047363  
883117836046086144 succeeded in 0.20423245429992676  
882992080364220416 succeeded in 0.1506955623626709  
882762694511734784 succeeded in 0.18434381484985352  
882627270321602560 succeeded in 0.18597936630249023  
882268110199369728 succeeded in 0.1515212059020996  
882045870035918850 succeeded in 0.2069845199584961  
881906580714921986 succeeded in 0.14809012413024902  
881666595344535552 succeeded in 0.19165754318237305  
881633300179243008 succeeded in 0.16114068031311035  
881536004380872706 succeeded in 0.17913126945495605  
881268444196462592 succeeded in 0.18312835693359375  
880935762899988482 succeeded in 0.14879179000854492  
880872448815771648 succeeded in 0.15070128440856934  
880465832366813184 succeeded in 0.19035911560058594  
880221127280381952 succeeded in 0.18358516693115234  
880095782870896641 succeeded in 0.1798722743988037  
879862464715927552 succeeded in 0.15911102294921875  
879674319642796034 succeeded in 0.18000197410583496  
879492040517615616 succeeded in 0.21718788146972656  
879415818425184262 succeeded in 0.21947121620178223  
879376492567855104 succeeded in 0.18320846557617188  
879130579576475649 succeeded in 0.20911669731140137  
879050749262655488 succeeded in 0.15411972999572754  
879008229531029506 succeeded in 0.20512795448303223  
878776093423087618 succeeded in 0.2066638469696045  
878604707211726852 succeeded in 0.16277503967285156

878404777348136964 succeeded in 0.15473318099975586  
878316110768087041 succeeded in 0.15352654457092285  
878281511006478336 succeeded in 0.15962934494018555  
878057613040115712 succeeded in 0.20669174194335938  
877736472329191424 succeeded in 0.15195798873901367  
877611172832227328 succeeded in 0.1539762020111084  
877556246731214848 succeeded in 0.1812272071838379  
877316821321428993 succeeded in 0.1856858730316162  
877201837425926144 succeeded in 0.15021538734436035  
876838120628539392 succeeded in 0.20592093467712402  
876537666061221889 succeeded in 0.17961549758911133  
876484053909872640 succeeded in 0.17778873443603516  
876120275196170240 succeeded in 0.15836501121520996  
875747767867523072 succeeded in 0.15387535095214844  
875144289856114688 succeeded in 0.18700623512268066  
875097192612077568 succeeded in 0.1676480770111084  
875021211251597312 succeeded in 0.18723320960998535  
874680097055178752 succeeded in 0.15838336944580078  
874434818259525634 succeeded in 0.15911436080932617  
874296783580663808 succeeded in 0.18268704414367676  
874057562936811520 succeeded in 0.1506798267364502  
874012996292530176 succeeded in 0.18485069274902344  
873697596434513921 error  
873580283840344065 succeeded in 0.15891361236572266  
873337748698140672 succeeded in 0.15801048278808594  
873213775632977920 succeeded in 0.15099740028381348  
872967104147763200 succeeded in 0.17800593376159668  
872820683541237760 succeeded in 0.16084861755371094  
872668790621863937 error  
872620804844003328 succeeded in 0.18076491355895996  
872486979161796608 succeeded in 0.20605087280273438  
872261713294495745 error  
872122724285648897 succeeded in 0.17197871208190918  
871879754684805121 succeeded in 0.21135163307189941  
871762521631449091 succeeded in 0.15015149116516113  
871515927908634625 succeeded in 0.14588427543640137  
871166179821445120 succeeded in 0.15375375747680664  
871102520638267392 succeeded in 0.2288062572479248  
871032628920680449 succeeded in 0.1930387020111084  
870804317367881728 succeeded in 0.18057680130004883  
870726314365509632 succeeded in 0.20397472381591797  
870656317836468226 succeeded in 0.18388962745666504  
870374049280663552 succeeded in 0.2499687671661377  
870308999962521604 succeeded in 0.2522127628326416  
870063196459192321 succeeded in 0.18251514434814453  
869988702071779329 error  
869772420881756160 succeeded in 0.2185518741607666  
869702957897576449 succeeded in 0.1839146614074707

869596645499047938 succeeded in 0.16982460021972656  
 869227993411051520 succeeded in 0.18504881858825684  
 868880397819494401 succeeded in 0.15430688858032227  
 868639477480148993 succeeded in 0.22976088523864746  
 868622495443632128 succeeded in 0.19039130210876465  
 868552278524837888 succeeded in 0.15549612045288086  
 867900495410671616 succeeded in 0.18381357192993164  
 867774946302451713 succeeded in 0.1847386360168457  
 867421006826221569 succeeded in 0.1811363697052002  
 867072653475098625 succeeded in 0.24001693725585938  
 867051520902168576 succeeded in 0.15139532089233398  
 866816280283807744 error  
 866720684873056260 succeeded in 0.16588950157165527  
 866686824827068416 succeeded in 0.20328235626220703  
 866450705531457537 succeeded in 0.1902461051940918  
 866334964761202691 succeeded in 0.18218755722045898  
 866094527597207552 succeeded in 0.2099306583404541  
 865718153858494464 succeeded in 0.15185284614562988  
 865359393868664832 succeeded in 0.15197181701660156  
 865006731092295680 succeeded in 0.18473219871520996  
 864873206498414592 succeeded in 0.1536557674407959  
 864279568663928832 succeeded in 0.2063744068145752  
 864197398364647424 succeeded in 0.17987561225891113  
 863907417377173506 succeeded in 0.21004128456115723  
 863553081350529029 succeeded in 0.23229503631591797  
 863471782782697472 succeeded in 0.1899280548095703  
 863432100342583297 succeeded in 0.15284013748168945  
 863427515083354112 succeeded in 0.20920610427856445  
 863079547188785154 succeeded in 0.17756199836730957  
 863062471531167744 succeeded in 0.22957897186279297  
 862831371563274240 succeeded in 0.20629405975341797  
 862722525377298433 succeeded in 0.15234088897705078  
 862457590147678208 succeeded in 0.18533039093017578  
 862096992088072192 succeeded in 0.1806049346923828  
 861769973181624320 error  
 861383897657036800 succeeded in 0.15369725227355957  
 861288531465048066 succeeded in 0.15289521217346191  
 861005113778896900 succeeded in 0.18546628952026367  
 860981674716409858 succeeded in 0.1521446704864502  
 860924035999428608 succeeded in 0.22155332565307617  
 860563773140209665 succeeded in 0.17940425872802734  
 860524505164394496 succeeded in 0.20415425300598145  
 860276583193509888 succeeded in 0.15607428550720215  
 860184849394610176 succeeded in 0.15348434448242188  
 860177593139703809 succeeded in 0.17439818382263184  
 859924526012018688 succeeded in 0.18406462669372559  
 859851578198683649 succeeded in 0.15129709243774414  
 859607811541651456 succeeded in 0.20927739143371582

859196978902773760 succeeded in 0.2394247055053711  
859074603037188101 succeeded in 0.1908400058746338  
858860390427611136 succeeded in 0.18897438049316406  
858843525470990336 succeeded in 0.19669008255004883  
858471635011153920 succeeded in 0.22457003593444824  
858107933456039936 succeeded in 0.1630384922027588  
857989990357356544 succeeded in 0.21541047096252441  
857746408056729600 succeeded in 0.22332286834716797  
857393404942143489 succeeded in 0.20614004135131836  
857263160327368704 succeeded in 0.15105175971984863  
857214891891077121 succeeded in 0.19761347770690918  
857062103051644929 succeeded in 0.19543671607971191  
857029823797047296 succeeded in 0.20872759819030762  
856602993587888130 error  
856543823941562368 succeeded in 0.19260072708129883  
856526610513747968 succeeded in 0.15149140357971191  
856330835276025856 succeeded in 0.22125911712646484  
856288084350160898 succeeded in 0.20669317245483398  
856282028240666624 succeeded in 0.15167832374572754  
855862651834028034 succeeded in 0.18189239501953125  
855860136149123072 succeeded in 0.20675420761108398  
855857698524602368 succeeded in 0.1599564552307129  
855851453814013952 succeeded in 0.15549087524414062  
855818117272018944 succeeded in 0.16237354278564453  
855459453768019968 succeeded in 0.15554308891296387  
855245323840757760 succeeded in 0.15622496604919434  
855138241867124737 succeeded in 0.2025766372680664  
854732716440526848 succeeded in 0.19339799880981445  
854482394044301312 succeeded in 0.15603041648864746  
854365224396361728 succeeded in 0.14988946914672852  
854120357044912130 succeeded in 0.15253901481628418  
854010172552949760 succeeded in 0.19157123565673828  
853760880890318849 succeeded in 0.20712924003601074  
853639147608842240 succeeded in 0.18476247787475586  
853299958564483072 succeeded in 0.16744041442871094  
852936405516943360 succeeded in 0.16520237922668457  
852912242202992640 succeeded in 0.16300344467163086  
852672615818899456 succeeded in 0.18723106384277344  
852553447878664193 succeeded in 0.14924907684326172  
852311364735569921 succeeded in 0.15364837646484375  
852226086759018497 succeeded in 0.16464614868164062  
852189679701164033 succeeded in 0.22040319442749023  
851953902622658560 error  
851861385021730816 succeeded in 0.1842813491821289  
851591660324737024 succeeded in 0.18678498268127441  
851464819735769094 succeeded in 0.18806886672973633  
851224888060895234 succeeded in 0.17424345016479492  
850753642995093505 succeeded in 0.14952301979064941



850380195714523136 succeeded in 0.17906451225280762  
850333567704068097 succeeded in 0.1569993495941162  
850145622816686080 succeeded in 0.1675553321838379  
850019790995546112 succeeded in 0.19141650199890137  
849776966551130114 succeeded in 0.15481305122375488  
849668094696017920 succeeded in 0.16032004356384277  
849412302885593088 succeeded in 0.18635845184326172  
849336543269576704 succeeded in 0.18192267417907715  
849051919805034497 succeeded in 0.16020727157592773  
848690551926992896 succeeded in 0.18834447860717773  
848324959059550208 succeeded in 0.18245434761047363  
848213670039564288 succeeded in 0.1604290008544922  
848212111729840128 succeeded in 0.15343761444091797  
847978865427394560 succeeded in 0.1666722297668457  
847971574464610304 succeeded in 0.21088504791259766  
847962785489326080 succeeded in 0.20621490478515625  
847842811428974592 succeeded in 0.15208840370178223  
847617282490613760 succeeded in 0.14963221549987793  
847606175596138505 succeeded in 0.18596386909484863  
847251039262605312 succeeded in 0.15788006782531738  
847157206088847362 succeeded in 0.15527558326721191  
847116187444137987 succeeded in 0.19782257080078125  
846874817362120707 succeeded in 0.15393471717834473  
846514051647705089 succeeded in 0.17505598068237305  
846505985330044928 succeeded in 0.2190723419189453  
846153765933735936 succeeded in 0.20655179023742676  
846139713627017216 succeeded in 0.22208189964294434  
846042936437604353 succeeded in 0.20624279975891113  
845812042753855489 succeeded in 0.1578226089477539  
845677943972139009 succeeded in 0.14588069915771484  
845459076796616705 error  
845397057150107648 succeeded in 0.15508437156677246  
845306882940190720 succeeded in 0.17839288711547852  
845098359547420673 succeeded in 0.22110819816589355  
844979544864018432 succeeded in 0.2078416347503662  
844973813909606400 succeeded in 0.14851117134094238  
844704788403113984 error  
844580511645339650 succeeded in 0.14998650550842285  
844223788422217728 succeeded in 0.15568804740905762  
843981021012017153 succeeded in 0.16010284423828125  
843856843873095681 succeeded in 0.15012097358703613  
843604394117681152 succeeded in 0.15079808235168457  
843235543001513987 succeeded in 0.15196847915649414  
842892208864923648 error  
842846295480000512 succeeded in 0.16293549537658691  
842765311967449089 succeeded in 0.15234732627868652  
842535590457499648 succeeded in 0.14835453033447266  
842163532590374912 succeeded in 0.1513381004333496

842115215311396866 succeeded in 0.16012287139892578  
841833993020538882 succeeded in 0.21019554138183594  
841680585030541313 succeeded in 0.18428874015808105  
841439858740625411 succeeded in 0.151902437210083  
841320156043304961 succeeded in 0.17103981971740723  
841314665196081154 succeeded in 0.1889784336090088  
841077006473256960 succeeded in 0.2086782455444336  
840761248237133825 succeeded in 0.1874842643737793  
840728873075638272 succeeded in 0.18598389625549316  
840698636975636481 succeeded in 0.18366336822509766  
840696689258311684 succeeded in 0.21266913414001465  
840632337062862849 succeeded in 0.16093158721923828  
840370681858686976 succeeded in 0.15653324127197266  
840268004936019968 succeeded in 0.18096256256103516  
839990271299457024 succeeded in 0.20429706573486328  
839549326359670784 succeeded in 0.15863966941833496  
839290600511926273 succeeded in 0.19318175315856934  
839239871831150596 succeeded in 0.15170955657958984  
838952994649550848 succeeded in 0.291644811630249  
838921590096166913 succeeded in 0.18442726135253906  
838916489579200512 succeeded in 0.20079517364501953  
838831947270979586 succeeded in 0.15700483322143555  
838561493054533637 succeeded in 0.20466256141662598  
838476387338051585 succeeded in 0.20638084411621094  
838201503651401729 succeeded in 0.15790390968322754  
838150277551247360 succeeded in 0.15573763847351074  
838085839343206401 succeeded in 0.18297314643859863  
838083903487373313 succeeded in 0.1521918773651123  
837820167694528512 succeeded in 0.15841221809387207  
837482249356513284 succeeded in 0.21612262725830078  
837471256429613056 succeeded in 0.15977835655212402  
837366284874571778 error  
837110210464448512 succeeded in 0.20704388618469238  
837012587749474308 error  
836989968035819520 succeeded in 0.22302722930908203  
836753516572119041 succeeded in 0.1922314167022705  
836677758902222849 succeeded in 0.21650958061218262  
836648853927522308 succeeded in 0.20331501960754395  
836397794269200385 succeeded in 0.15766692161560059  
836380477523124226 succeeded in 0.1701056957244873  
836260088725786625 succeeded in 0.16167616844177246  
836001077879255040 succeeded in 0.19490313529968262  
835685285446955009 succeeded in 0.21071791648864746  
835574547218894849 succeeded in 0.15830183029174805  
835536468978302976 succeeded in 0.15567755699157715  
835309094223372289 succeeded in 0.21893572807312012  
835297930240217089 succeeded in 0.18522214889526367  
835264098648616962 succeeded in 0.20528483390808105

835246439529840640 succeeded in 0.20673084259033203  
835172783151792128 succeeded in 0.20724034309387207  
835152434251116546 succeeded in 0.17722105979919434  
834931633769889797 succeeded in 0.15251851081848145  
834786237630337024 succeeded in 0.18717598915100098  
834574053763584002 succeeded in 0.15219378471374512  
834477809192075265 succeeded in 0.1553335189819336  
834458053273591808 succeeded in 0.2200610637664795  
834209720923721728 succeeded in 0.15839934349060059  
834167344700198914 succeeded in 0.1553947925567627  
834089966724603904 succeeded in 0.218458890914917  
834086379323871233 succeeded in 0.15300679206848145  
833863086058651648 succeeded in 0.22031283378601074  
833826103416520705 succeeded in 0.1553652286529541  
833732339549220864 succeeded in 0.21312999725341797  
833722901757046785 succeeded in 0.18480753898620605  
833479644947025920 succeeded in 0.15343618392944336  
833124694597443584 succeeded in 0.21306705474853516  
832998151111966721 succeeded in 0.15732765197753906  
832769181346996225 succeeded in 0.15967082977294922  
832757312314028032 succeeded in 0.1801161766052246  
832682457690300417 succeeded in 0.2143397331237793  
832645525019123713 succeeded in 0.16104960441589355  
832636094638288896 succeeded in 0.1520369052886963  
832397543355072512 succeeded in 0.1570417881011963  
832369877331693569 succeeded in 0.22831487655639648  
832273440279240704 succeeded in 0.15596938133239746  
832215909146226688 succeeded in 0.16622471809387207  
832215726631055365 succeeded in 0.15796995162963867  
832088576586297345 succeeded in 0.18267369270324707  
832040443403784192 succeeded in 0.16765618324279785  
832032802820481025 succeeded in 0.15489792823791504  
831939777352105988 succeeded in 0.18460297584533691  
831926988323639298 succeeded in 0.18088984489440918  
831911600680497154 succeeded in 0.1830155849456787  
831670449226514432 succeeded in 0.1829066276550293  
831650051525054464 succeeded in 0.18500852584838867  
831552930092285952 succeeded in 0.19003033638000488  
831322785565769729 succeeded in 0.1537611484527588  
831315979191906304 succeeded in 0.1526329517364502  
831309418084069378 succeeded in 0.1863083839416504  
831262627380748289 succeeded in 0.1852715015411377  
830956169170665475 succeeded in 0.1634836196899414  
830583320585068544 succeeded in 0.20406794548034668  
830173239259324417 succeeded in 0.17529296875  
830097400375152640 succeeded in 0.17804646492004395  
829878982036299777 succeeded in 0.18361544609069824  
829861396166877184 succeeded in 0.19901514053344727

829501995190984704 succeeded in 0.17920565605163574  
829449946868879360 succeeded in 0.1554868221282959  
829374341691346946 error  
829141528400556032 succeeded in 0.20516371726989746  
829011960981237760 succeeded in 0.23304486274719238  
828801551087042563 succeeded in 0.2114253044128418  
828770345708580865 succeeded in 0.18367218971252441  
828708714936930305 succeeded in 0.18814516067504883  
828650029636317184 succeeded in 0.18668293952941895  
828409743546925057 succeeded in 0.21517205238342285  
828408677031882754 succeeded in 0.153764009475708  
828381636999917570 succeeded in 0.1477499008178711  
828376505180889089 succeeded in 0.14992141723632812  
828372645993398273 succeeded in 0.15289974212646484  
828361771580813312 succeeded in 0.15797972679138184  
828046555563323392 succeeded in 0.1954634189605713  
828011680017821696 succeeded in 0.1527101993560791  
827933404142436356 succeeded in 0.17797541618347168  
827653905312006145 succeeded in 0.14936137199401855  
827600520311402496 succeeded in 0.19086217880249023  
827324948884643840 succeeded in 0.2097029685974121  
827228250799742977 error  
827199976799354881 succeeded in 0.18513751029968262  
826958653328592898 succeeded in 0.20734953880310059  
826848821049180160 succeeded in 0.18599176406860352  
826615380357632002 succeeded in 0.1918485164642334  
826598799820865537 succeeded in 0.23319602012634277  
826598365270007810 succeeded in 0.1489267349243164  
826476773533745153 succeeded in 0.15375828742980957  
826240494070030336 succeeded in 0.18535995483398438  
826204788643753985 succeeded in 0.20610284805297852  
826115272272650244 succeeded in 0.15254735946655273  
825876512159186944 succeeded in 0.18429136276245117  
825829644528148480 succeeded in 0.1942605972290039  
825535076884762624 succeeded in 0.24609017372131348  
825147591692263424 succeeded in 0.20657086372375488  
825120256414846976 succeeded in 0.1543889045715332  
825026590719483904 succeeded in 0.1613161563873291  
824796380199809024 succeeded in 0.18210387229919434  
824775126675836928 succeeded in 0.18778562545776367  
824663926340194305 succeeded in 0.15809226036071777  
824325613288833024 succeeded in 0.15373969078063965  
824297048279236611 succeeded in 0.18348288536071777  
824025158776213504 succeeded in 0.2043755054473877  
823939628516474880 succeeded in 0.15529346466064453  
823719002937630720 succeeded in 0.18645715713500977  
823699002998870016 succeeded in 0.183089017868042  
823581115634085888 succeeded in 0.17875957489013672

823333489516937216 succeeded in 0.21022796630859375  
823322678127919110 succeeded in 0.18561577796936035  
823269594223824897 succeeded in 0.20941400527954102  
822975315408461824 succeeded in 0.14801406860351562  
822872901745569793 succeeded in 0.17866754531860352  
822859134160621569 succeeded in 0.14643120765686035  
822647212903690241 succeeded in 0.17498779296875  
822610361945911296 succeeded in 0.15159249305725098  
822489057087389700 succeeded in 0.19565320014953613  
822462944365645825 succeeded in 0.16212797164916992  
822244816520155136 succeeded in 0.18311595916748047  
822163064745328640 succeeded in 0.16249656677246094  
821886076407029760 succeeded in 0.18648576736450195  
821813639212650496 succeeded in 0.18831086158752441  
821765923262631936 succeeded in 0.17171740531921387  
821522889702862852 succeeded in 0.15148568153381348  
821421320206483457 succeeded in 0.164276123046875  
821407182352777218 succeeded in 0.18572282791137695  
821153421864615936 succeeded in 0.1504220962524414  
821149554670182400 succeeded in 0.1853034496307373  
821107785811234820 succeeded in 0.21071076393127441  
821044531881721856 succeeded in 0.21567797660827637  
820837357901512704 succeeded in 0.1562802791595459  
820749716845686786 succeeded in 0.15624403953552246  
820690176645140481 succeeded in 0.183060884475708  
820494788566847489 succeeded in 0.15907883644104004  
820446719150292993 succeeded in 0.16300702095031738  
820314633777061888 succeeded in 0.20539641380310059  
820078625395449857 succeeded in 0.2110133171081543  
820013781606658049 succeeded in 0.18962740898132324  
819952236453363712 succeeded in 0.17964911460876465  
819924195358416896 succeeded in 0.20610284805297852  
819711362133872643 succeeded in 0.18709945678710938  
819588359383371776 succeeded in 0.18343019485473633  
819347104292290561 succeeded in 0.1556100845336914  
819238181065359361 succeeded in 0.1535789966583252  
819227688460238848 succeeded in 0.15236115455627441  
819015337530290176 succeeded in 0.1929759979248047  
819015331746349057 succeeded in 0.21015238761901855  
819006400881917954 succeeded in 0.15362119674682617  
819004803107983360 succeeded in 0.18472838401794434  
818646164899774465 succeeded in 0.1869962215423584  
818627210458333184 succeeded in 0.19043684005737305  
818614493328580609 succeeded in 0.2099299430847168  
818588835076603904 succeeded in 0.1918337345123291  
818536468981415936 succeeded in 0.15087628364562988  
818307523543449600 succeeded in 0.1865394115447998  
818259473185828864 succeeded in 0.17892241477966309

818145370475810820 succeeded in 0.18486952781677246  
817908911860748288 succeeded in 0.16106319427490234  
817827839487737858 succeeded in 0.15567255020141602  
817777686764523521 succeeded in 0.1795344352722168  
817536400337801217 succeeded in 0.1617424488067627  
817502432452313088 succeeded in 0.20593738555908203  
817423860136083457 succeeded in 0.15258073806762695  
817415592588222464 succeeded in 0.18682360649108887  
817181837579653120 succeeded in 0.19946956634521484  
817171292965273600 succeeded in 0.16673827171325684  
817120970343411712 succeeded in 0.18209099769592285  
817056546584727552 succeeded in 0.15235161781311035  
816829038950027264 succeeded in 0.21467185020446777  
816816676327063552 succeeded in 0.18257546424865723  
816697700272001025 succeeded in 0.2142024040222168  
816450570814898180 succeeded in 0.20888686180114746  
816336735214911488 succeeded in 0.1779336929321289  
816091915477250048 succeeded in 0.15692448616027832  
816062466425819140 succeeded in 0.15757203102111816  
816014286006976512 succeeded in 0.1967780590057373  
815990720817401858 succeeded in 0.18572735786437988  
815966073409433600 succeeded in 0.20615625381469727  
815745968457060357 succeeded in 0.18338942527770996  
815736392542261248 succeeded in 0.19895696640014648  
815639385530101762 succeeded in 0.2093212604522705  
815390420867969024 succeeded in 0.17901396751403809  
814986499976527872 succeeded in 0.20472145080566406  
814638523311648768 succeeded in 0.2101120948791504  
814578408554463233 succeeded in 0.16094255447387695  
814530161257443328 succeeded in 0.1909337043762207  
814153002265309185 succeeded in 0.15215802192687988  
813944609378369540 succeeded in 0.19214868545532227  
813910438903693312 succeeded in 0.27085304260253906  
813812741911748608 succeeded in 0.24168753623962402  
813800681631023104 succeeded in 0.19037222862243652  
813217897535406080 succeeded in 0.15621519088745117  
813202720496779264 succeeded in 0.18010377883911133  
813187593374461952 succeeded in 0.1772470474243164  
813172488309972993 succeeded in 0.17675232887268066  
813157409116065792 succeeded in 0.1596062183380127  
813142292504645637 succeeded in 0.15498113632202148  
813130366689148928 succeeded in 0.1578354835510254  
813127251579564032 succeeded in 0.15164995193481445  
813112105746448384 succeeded in 0.16274285316467285  
813096984823349248 succeeded in 0.21100759506225586  
813081950185472002 succeeded in 0.17803645133972168  
813066809284972545 succeeded in 0.1547245979309082  
813051746834595840 succeeded in 0.15521693229675293

812781120811126785 succeeded in 0.15020298957824707  
812747805718642688 error  
812709060537683968 succeeded in 0.1485457420349121  
812503143955202048 succeeded in 0.15134382247924805  
812466873996607488 succeeded in 0.1955699920654297  
812372279581671427 succeeded in 0.19640851020812988  
811985624773361665 succeeded in 0.20693016052246094  
811744202451197953 succeeded in 0.17841506004333496  
811647686436880384 succeeded in 0.18773269653320312  
811627233043480576 succeeded in 0.17456388473510742  
811386762094317568 succeeded in 0.17998385429382324  
810984652412424192 succeeded in 0.23275494575500488  
810896069567610880 succeeded in 0.19411659240722656  
810657578271330305 succeeded in 0.18016910552978516  
810284430598270976 succeeded in 0.18293070793151855  
810254108431155201 succeeded in 0.16698789596557617  
809920764300447744 succeeded in 0.16227006912231445  
809808892968534016 succeeded in 0.18751263618469238  
809448704142938112 succeeded in 0.1758720874786377  
809220051211603969 succeeded in 0.18341422080993652  
809084759137812480 succeeded in 0.1901686191558838  
808838249661788160 succeeded in 0.19432520866394043  
808733504066486276 succeeded in 0.20816397666931152  
808501579447930884 succeeded in 0.20734715461730957  
808344865868283904 succeeded in 0.2077944278717041  
808134635716833280 succeeded in 0.15477895736694336  
808106460588765185 succeeded in 0.19310259819030762  
808001312164028416 succeeded in 0.18450593948364258  
807621403335917568 succeeded in 0.20531129837036133  
807106840509214720 succeeded in 0.20715665817260742  
807059379405148160 succeeded in 0.18812203407287598  
807010152071229440 succeeded in 0.22408127784729004  
806629075125202948 succeeded in 0.15154194831848145  
806620845233815552 succeeded in 0.22936558723449707  
806576416489959424 succeeded in 0.20658516883850098  
806542213899489280 succeeded in 0.18213748931884766  
806242860592926720 succeeded in 0.16404485702514648  
806219024703037440 succeeded in 0.1627962589263916  
805958939288408065 succeeded in 0.18892121315002441  
805932879469572096 succeeded in 0.18625092506408691  
805826884734976000 succeeded in 0.2066497802734375  
805823200554876929 succeeded in 0.19124913215637207  
805520635690676224 succeeded in 0.2169480323791504  
805487436403003392 succeeded in 0.15645098686218262  
805207613751304193 succeeded in 0.18808579444885254  
804738756058218496 succeeded in 0.18053555488586426  
804475857670639616 succeeded in 0.20272064208984375  
804413760345620481 succeeded in 0.15538835525512695

804026241225523202 succeeded in 0.18668580055236816  
803773340896923648 succeeded in 0.15592527389526367  
803692223237865472 succeeded in 0.19052982330322266  
803638050916102144 succeeded in 0.16632580757141113  
803380650405482500 succeeded in 0.1881730556488037  
803321560782307329 succeeded in 0.18779206275939941  
803276597545603072 succeeded in 0.15141916275024414  
802952499103731712 succeeded in 0.14954018592834473  
802624713319034886 succeeded in 0.19797682762145996  
802600418706604034 succeeded in 0.14818787574768066  
802572683846291456 succeeded in 0.15985751152038574  
802323869084381190 succeeded in 0.18046903610229492  
802265048156610565 succeeded in 0.16034674644470215  
802247111496568832 error  
802239329049477120 succeeded in 0.1566298007965088  
802185808107208704 succeeded in 0.16154098510742188  
801958328846974976 succeeded in 0.19351863861083984  
801854953262350336 succeeded in 0.20621538162231445  
801538201127157760 succeeded in 0.1860826015472412  
801285448605831168 succeeded in 0.15319347381591797  
801167903437357056 succeeded in 0.14920854568481445  
801127390143516673 succeeded in 0.1913919448852539  
801115127852503040 succeeded in 0.14918160438537598  
800859414831898624 succeeded in 0.19580984115600586  
800855607700029440 succeeded in 0.1887648105621338  
800751577355128832 succeeded in 0.20415925979614258  
800513324630806528 succeeded in 0.1516284942626953  
800459316964663297 succeeded in 0.15421509742736816  
800443802682937345 succeeded in 0.2165536880493164  
800388270626521089 succeeded in 0.20649981498718262  
800188575492947969 succeeded in 0.152876615524292  
800141422401830912 succeeded in 0.21294593811035156  
800018252395122689 succeeded in 0.1815805435180664  
799774291445383169 succeeded in 0.2184741497039795  
799757965289017345 succeeded in 0.16175150871276855  
799422933579902976 succeeded in 0.20538806915283203  
799308762079035393 succeeded in 0.19024991989135742  
799297110730567681 succeeded in 0.1572740077972412  
799063482566066176 succeeded in 0.1858968734741211  
798933969379225600 succeeded in 0.18491816520690918  
798925684722855936 succeeded in 0.15247797966003418  
798705661114773508 succeeded in 0.1587080955505371  
798701998996647937 succeeded in 0.18473219871520996  
798697898615730177 succeeded in 0.19790124893188477  
798694562394996736 succeeded in 0.21553683280944824  
798686750113755136 succeeded in 0.18053507804870605  
798682547630837760 succeeded in 0.15368437767028809  
798673117451325440 succeeded in 0.1562178134918213



798665375516884993 succeeded in 0.17234063148498535  
798644042770751489 succeeded in 0.19123387336730957  
798628517273620480 succeeded in 0.15171337127685547  
798585098161549313 succeeded in 0.15768885612487793  
798576900688019456 succeeded in 0.15534162521362305  
798340744599797760 succeeded in 0.16598916053771973  
798209839306514432 succeeded in 0.15273380279541016  
797971864723324932 succeeded in 0.1550760269165039  
797545162159308800 succeeded in 0.21588635444641113  
797236660651966464 succeeded in 0.20641803741455078  
797165961484890113 succeeded in 0.158461332321167  
796904159865868288 succeeded in 0.18660950660705566  
796865951799083009 succeeded in 0.15494537353515625  
796759840936919040 succeeded in 0.15076661109924316  
796563435802726400 succeeded in 0.18453669548034668  
796484825502875648 succeeded in 0.19175004959106445  
796387464403357696 succeeded in 0.15295720100402832  
796177847564038144 succeeded in 0.1891796588897705  
796149749086875649 succeeded in 0.15468716621398926  
796125600683540480 succeeded in 0.18191003799438477  
796116448414461957 succeeded in 0.18943500518798828  
796080075804475393 succeeded in 0.1513216495513916  
796031486298386433 succeeded in 0.15403103828430176  
795464331001561088 succeeded in 0.1527109146118164  
795400264262053889 succeeded in 0.18667960166931152  
795076730285391872 succeeded in 0.21532535552978516  
794983741416415232 succeeded in 0.18439650535583496  
794926597468000259 succeeded in 0.1651914119720459  
794355576146903043 succeeded in 0.1815955638885498  
794332329137291264 succeeded in 0.1858985424041748  
794205286408003585 succeeded in 0.19045615196228027  
793962221541933056 succeeded in 0.15344643592834473  
793845145112371200 succeeded in 0.1530160903930664  
793614319594401792 succeeded in 0.18206143379211426  
793601777308463104 succeeded in 0.18381977081298828  
793500921481273345 succeeded in 0.22340774536132812  
793286476301799424 succeeded in 0.1540687084197998  
793271401113350145 succeeded in 0.1559305191040039  
793256262322548741 succeeded in 0.2066059112548828  
793241302385262592 succeeded in 0.20666241645812988  
793226087023144960 succeeded in 0.1603238582611084  
793210959003287553 succeeded in 0.15335440635681152  
793195938047070209 succeeded in 0.18526172637939453  
793180763617361921 succeeded in 0.21865439414978027  
793165685325201412 succeeded in 0.18239498138427734  
793150605191548928 succeeded in 0.19277167320251465  
793135492858580992 succeeded in 0.2046647071838379  
793120401413079041 succeeded in 0.1503000259399414

792913359805018113 succeeded in 0.19118595123291016  
792883833364439040 succeeded in 0.21331119537353516  
792773781206999040 succeeded in 0.1645526885986328  
792394556390137856 succeeded in 0.16767668724060059  
792050063153438720 succeeded in 0.17853426933288574  
791821351946420224 succeeded in 0.20334434509277344  
791784077045166082 succeeded in 0.15837359428405762  
791780927877898241 succeeded in 0.20304322242736816  
791774931465953280 succeeded in 0.20933151245117188  
791672322847637504 succeeded in 0.21683883666992188  
791406955684368384 succeeded in 0.18033099174499512  
791312159183634433 succeeded in 0.15628552436828613  
791026214425268224 succeeded in 0.19935154914855957  
790987426131050500 succeeded in 0.15150737762451172  
790946055508652032 succeeded in 0.1565837860107422  
790723298204217344 succeeded in 0.18905901908874512  
790698755171364864 succeeded in 0.15180635452270508  
790581949425475584 succeeded in 0.1655125617980957  
790337589677002753 succeeded in 0.18212485313415527  
790277117346975746 succeeded in 0.2913501262664795  
790227638568808452 succeeded in 0.1592850685119629  
789986466051088384 succeeded in 0.18236303329467773  
789960241177853952 succeeded in 0.20354914665222168  
789903600034189313 succeeded in 0.20714855194091797  
789628658055020548 succeeded in 0.15064167976379395  
789599242079838210 succeeded in 0.15871596336364746  
789530877013393408 succeeded in 0.15088105201721191  
789314372632018944 succeeded in 0.23058247566223145  
789280767834746880 succeeded in 0.16323089599609375  
789268448748703744 succeeded in 0.20858001708984375  
789137962068021249 succeeded in 0.15337705612182617  
788908386943430656 succeeded in 0.15234827995300293  
788765914992902144 succeeded in 0.15999674797058105  
788552643979468800 succeeded in 0.20042967796325684  
788412144018661376 succeeded in 0.1528785228729248  
788178268662984705 succeeded in 0.15330886840820312  
788150585577050112 succeeded in 0.18321895599365234  
788070120937619456 succeeded in 0.16525530815124512  
788039637453406209 succeeded in 0.2041788101196289  
787810552592695296 succeeded in 0.18367242813110352  
787717603741622272 succeeded in 0.15734648704528809  
787397959788929025 succeeded in 0.17731881141662598  
787322443945877504 succeeded in 0.15968632698059082  
787111942498508800 succeeded in 0.156508207321167  
786963064373534720 succeeded in 0.18509173393249512  
786729988674449408 succeeded in 0.15370726585388184  
786709082849828864 succeeded in 0.15032100677490234  
786664955043049472 succeeded in 0.1818375587463379

786595970293370880 succeeded in 0.20516681671142578  
786363235746385920 succeeded in 0.14843463897705078  
786286427768250368 succeeded in 0.15326142311096191  
786233965241827333 succeeded in 0.15059518814086914  
786051337297522688 succeeded in 0.1858663558959961  
786036967502913536 succeeded in 0.2107384204864502  
785927819176054784 succeeded in 0.1573038101196289  
785872687017132033 succeeded in 0.14925408363342285  
785639753186217984 succeeded in 0.15015077590942383  
785533386513321988 succeeded in 0.22995829582214355  
785515384317313025 succeeded in 0.20659136772155762  
785264754247995392 succeeded in 0.15356826782226562  
785170936622350336 succeeded in 0.16020750999450684  
784826020293709826 succeeded in 0.17592740058898926  
784517518371221505 succeeded in 0.16427278518676758  
784431430411685888 succeeded in 0.1787877082824707  
784183165795655680 succeeded in 0.17766571044921875  
784057939640352768 succeeded in 0.15401268005371094  
783839966405230592 succeeded in 0.1908128261566162  
783821107061198850 succeeded in 0.16142940521240234  
783695101801398276 succeeded in 0.1853957176208496  
783466772167098368 succeeded in 0.1477506160736084  
783391753726550016 succeeded in 0.1566619873046875  
783347506784731136 succeeded in 0.15827012062072754  
783334639985389568 succeeded in 0.18847274780273438  
783085703974514689 succeeded in 0.3035151958465576  
782969140009107456 succeeded in 0.1522049903869629  
782747134529531904 succeeded in 0.15413403511047363  
782722598790725632 succeeded in 0.15836548805236816  
782598640137187329 succeeded in 0.19512605667114258  
782305867769217024 succeeded in 0.15272808074951172  
782021823840026624 succeeded in 0.21095061302185059  
781955203444699136 succeeded in 0.15057754516601562  
781661882474196992 succeeded in 0.2055366039276123  
781655249211752448 succeeded in 0.1608564853668213  
781524693396357120 succeeded in 0.1509242057800293  
781308096455073793 succeeded in 0.14921832084655762  
781251288990355457 succeeded in 0.17908143997192383  
781163403222056960 succeeded in 0.15262985229492188  
780931614150983680 succeeded in 0.219696044921875  
780858289093574656 succeeded in 0.15662288665771484  
780800785462489090 succeeded in 0.165985107421875  
780601303617732608 succeeded in 0.14907264709472656  
780543529827336192 succeeded in 0.15348505973815918  
780496263422808064 succeeded in 0.1921250820159912  
780476555013349377 succeeded in 0.18449878692626953  
780459368902959104 succeeded in 0.18418431282043457  
780192070812196864 succeeded in 0.18088698387145996

780092040432480260 succeeded in 0.1639711856842041  
780074436359819264 succeeded in 0.16497325897216797  
779834332596887552 succeeded in 0.1878516674041748  
779377524342161408 succeeded in 0.22178196907043457  
779124354206535695 succeeded in 0.21018481254577637  
779123168116150273 error  
779056095788752897 succeeded in 0.1677110195159912  
778990705243029504 succeeded in 0.152679443359375  
778774459159379968 succeeded in 0.1552567481994629  
778764940568104960 succeeded in 0.18887543678283691  
778748913645780993 succeeded in 0.18668532371520996  
778650543019483137 succeeded in 0.21511149406433105  
778624900596654080 succeeded in 0.17709684371948242  
778408200802557953 succeeded in 0.15303444862365723  
778396591732486144 succeeded in 0.2228226661682129  
778383385161035776 succeeded in 0.19055414199829102  
778286810187399168 succeeded in 0.19502496719360352  
778039087836069888 succeeded in 0.17764902114868164  
778027034220126208 succeeded in 0.17848420143127441  
777953400541634568 succeeded in 0.18799471855163574  
777885040357281792 succeeded in 0.1650679111480713  
777684233540206592 succeeded in 0.20476102828979492  
777641927919427584 succeeded in 0.19049334526062012  
777621514455814149 succeeded in 0.18598508834838867  
777189768882946048 succeeded in 0.17644238471984863  
776819012571455488 succeeded in 0.1581554412841797  
776813020089548800 succeeded in 0.19668841361999512  
776477788987613185 succeeded in 0.18590044975280762  
776249906839351296 succeeded in 0.18845129013061523  
776218204058357768 succeeded in 0.15947413444519043  
776201521193218049 succeeded in 0.14863371849060059  
776113305656188928 succeeded in 0.18928742408752441  
776088319444877312 succeeded in 0.20694518089294434  
775898661951791106 succeeded in 0.1568284034729004  
775842724423557120 succeeded in 0.15038824081420898  
775733305207554048 succeeded in 0.15404844284057617  
775729183532220416 succeeded in 0.21015691757202148  
775364825476165632 succeeded in 0.1793515682220459  
775350846108426240 succeeded in 0.2101125717163086  
775096608509886464 error  
775085132600442880 succeeded in 0.15280628204345703  
774757898236878852 succeeded in 0.1940155029296875  
774639387460112384 succeeded in 0.19160008430480957  
774314403806253056 succeeded in 0.17813920974731445  
773985732834758656 succeeded in 0.18622159957885742  
773922284943896577 succeeded in 0.18200039863586426  
773704687002451968 succeeded in 0.1972520351409912  
773670353721753600 succeeded in 0.18229937553405762

773547596996571136 succeeded in 0.1856677532196045  
773336787167145985 succeeded in 0.15618538856506348  
773308824254029826 succeeded in 0.15237903594970703  
773247561583001600 succeeded in 0.31195974349975586  
773191612633579521 succeeded in 0.19027209281921387  
772877495989305348 succeeded in 0.1652064323425293  
772826264096874500 succeeded in 0.14986371994018555  
772615324260794368 succeeded in 0.15773963928222656  
772581559778025472 succeeded in 0.18409299850463867  
772193107915964416 succeeded in 0.1860039234161377  
772152991789019136 succeeded in 0.7555782794952393  
772117678702071809 succeeded in 0.18793749809265137  
772114945936949249 succeeded in 0.21391868591308594  
772102971039580160 succeeded in 0.16357898712158203  
771908950375665664 succeeded in 0.16828584671020508  
771770456517009408 succeeded in 0.18099117279052734  
771500966810099713 succeeded in 0.1858513355255127  
771380798096281600 succeeded in 0.15234923362731934  
771171053431250945 succeeded in 0.19140291213989258  
771136648247640064 succeeded in 0.1522517204284668  
771102124360998913 succeeded in 0.15871524810791016  
771014301343748096 succeeded in 0.2063140869140625  
771004394259247104 error  
770787852854652928 succeeded in 0.17838406562805176  
770772759874076672 succeeded in 0.1712193489074707  
770743923962707968 error  
770655142660169732 succeeded in 0.18001317977905273  
770414278348247044 succeeded in 0.149627685546875  
770293558247038976 succeeded in 0.15981841087341309  
770093767776997377 succeeded in 0.19324326515197754  
770069151037685760 succeeded in 0.15905404090881348  
769940425801170949 succeeded in 0.18602371215820312  
769695466921623552 succeeded in 0.15120482444763184  
769335591808995329 succeeded in 0.16403770446777344  
769212283578875904 succeeded in 0.18758320808410645  
768970937022709760 succeeded in 0.18346762657165527  
768909767477751808 succeeded in 0.2384190559387207  
768855141948723200 succeeded in 0.1479175090789795  
768609597686943744 succeeded in 0.20965576171875  
768596291618299904 succeeded in 0.15247750282287598  
768554158521745409 succeeded in 0.16002774238586426  
768473857036525572 succeeded in 0.15532851219177246  
768193404517830656 succeeded in 0.15185141563415527  
767884188863397888 succeeded in 0.2743651866912842  
767754930266464257 succeeded in 0.18896222114562988  
767500508068192258 succeeded in 0.15578913688659668  
767191397493538821 succeeded in 0.15573644638061523  
767122157629476866 succeeded in 0.21918797492980957

766864461642756096 succeeded in 0.23073172569274902  
766793450729734144 succeeded in 0.1801896095275879  
766714921925144576 succeeded in 0.18609023094177246  
766693177336135680 succeeded in 0.20975947380065918  
766423258543644672 succeeded in 0.1520555019378662  
766313316352462849 succeeded in 0.2125847339630127  
766078092750233600 succeeded in 0.15760207176208496  
766069199026450432 succeeded in 0.15213632583618164  
766008592277377025 succeeded in 0.21417951583862305  
765719909049503744 succeeded in 0.2109389305114746  
765669560888528897 succeeded in 0.21238923072814941  
765395769549590528 succeeded in 0.15301299095153809  
765371061932261376 succeeded in 0.17848515510559082  
765222098633691136 succeeded in 0.15386152267456055  
764857477905154048 succeeded in 0.17596125602722168  
764259802650378240 succeeded in 0.216233491897583  
763956972077010945 succeeded in 0.32237792015075684  
763837565564780549 succeeded in 0.15363121032714844  
763183847194451968 succeeded in 0.15507221221923828  
763167063695355904 succeeded in 0.15961909294128418  
763103485927849985 succeeded in 0.19018769264221191  
762699858130116608 succeeded in 0.19621968269348145  
762471784394268675 succeeded in 0.222517728805542  
762464539388485633 succeeded in 0.2448892593383789  
762316489655476224 succeeded in 0.18789958953857422  
762035686371364864 succeeded in 0.1991434097290039  
761976711479193600 succeeded in 0.1526355743408203  
761750502866649088 succeeded in 0.15741181373596191  
761745352076779520 succeeded in 0.17812013626098633  
761672994376806400 succeeded in 0.1793215274810791  
761599872357261312 succeeded in 0.18927907943725586  
761371037149827077 succeeded in 0.18382000923156738  
761334018830917632 succeeded in 0.2115473747253418  
761292947749015552 succeeded in 0.1844186782836914  
761227390836215808 succeeded in 0.1858358383178711  
761004547850530816 succeeded in 0.20830941200256348  
760893934457552897 succeeded in 0.15377378463745117  
760656994973933572 succeeded in 0.2076551914215088  
760641137271070720 succeeded in 0.2062361240386963  
760539183865880579 succeeded in 0.1500086784362793  
760521673607086080 succeeded in 0.1606442928314209  
760290219849637889 succeeded in 0.1522808074951172  
760252756032651264 succeeded in 0.15236592292785645  
760190180481531904 succeeded in 0.15180277824401855  
760153949710192640 succeeded in 0.15455412864685059  
759943073749200896 succeeded in 0.15834522247314453  
759923798737051648 succeeded in 0.19166851043701172  
759846353224826880 succeeded in 0.1545400619506836

759793422261743616 succeeded in 0.21073174476623535  
759566828574212096 error  
759557299618865152 succeeded in 0.21598410606384277  
759447681597108224 succeeded in 0.18930840492248535  
759446261539934208 succeeded in 0.20190715789794922  
759197388317847553 succeeded in 0.15298700332641602  
759159934323924993 succeeded in 0.1892712116241455  
759099523532779520 succeeded in 0.20846843719482422  
759047813560868866 succeeded in 0.15811538696289062  
758854675097526272 succeeded in 0.15303540229797363  
758828659922702336 succeeded in 0.16010093688964844  
758740312047005698 error  
758474966123810816 error  
758467244762497024 error  
758405701903519748 error  
758355060040593408 error  
758099635764359168 error  
758041019896193024 error  
757741869644341248 error  
757729163776290825 error  
757725642876129280 error  
757611664640446465 error  
757597904299253760 error  
757596066325864448 error  
757400162377592832 error  
757393109802180609 error  
757354760399941633 error  
756998049151549440 error  
756939218950160384 error  
756651752796094464 error  
756526248105566208 error  
756303284449767430 error  
756288534030475264 error  
756275833623502848 error  
755955933503782912 error  
755206590534418437 error  
755110668769038337 error  
754874841593970688 error  
754856583969079297 error  
754747087846248448 error  
754482103782404096 error  
754449512966619136 error  
754120377874386944 error  
754011816964026368 error  
753655901052166144 error  
753420520834629632 error  
753398408988139520 error  
753375668877008896 error

753298634498793472 error  
753294487569522689 error  
753039830821511168 error  
753026973505581056 error  
752932432744185856 error  
752917284578922496 error  
752701944171524096 error  
752682090207055872 error  
752660715232722944 error  
752568224206688256 error  
752519690950500352 error  
752334515931054080 error  
752309394570878976 error  
752173152931807232 error  
751950017322246144 error  
751937170840121344 error  
751830394383790080 error  
751793661361422336 error  
751598357617971201 error  
751583847268179968 error  
751538714308972544 error  
751456908746354688 error  
751251247299190784 error  
751205363882532864 error  
751132876104687617 error  
750868782890057730 error  
750719632563142656 error  
750506206503038976 error  
750429297815552001 error  
750383411068534784 error  
750381685133418496 error  
750147208377409536 error  
750132105863102464 error  
750117059602808832 error  
750101899009982464 error  
750086836815486976 error  
750071704093859840 error  
750056684286914561 error  
750041628174217216 error  
750026558547456000 error  
750011400160841729 error  
749996283729883136 error  
749981277374128128 error  
749774190421639168 error  
749417653287129088 error  
749403093750648834 error  
749395845976588288 error  
749317047558017024 error



749075273010798592 error  
749064354620928000 error  
749036806121881602 error  
748977405889503236 error  
748932637671223296 error  
748705597323898880 error  
748699167502000129 error  
748692773788876800 error  
748575535303884801 error  
748568946752774144 error  
748346686624440324 error  
748337862848962560 error  
748324050481647620 error  
748307329658011649 error  
748220828303695873 error  
747963614829678593 error  
747933425676525569 error  
747885874273214464 error  
747844099428986880 error  
747816857231626240 error  
747651430853525504 error  
747648653817413632 error  
747600769478692864 error  
747594051852075008 error  
747512671126323200 error  
747461612269887489 error  
747439450712596480 error  
747242308580548608 error  
747219827526344708 error  
747204161125646336 error  
747103485104099331 error  
746906459439529985 error  
746872823977771008 error  
746818907684614144 error  
746790600704425984 error  
746757706116112384 error  
746726898085036033 error  
746542875601690625 error  
746521445350707200 error  
746507379341139972 error  
746369468511756288 error  
746131877086527488 error  
746056683365994496 error  
745789745784041472 error  
745712589599014916 error  
745433870967832576 error  
745422732645535745 error  
745314880350101504 error

745074613265149952 error  
745057283344719872 error  
744995568523612160 error  
744971049620602880 error  
744709971296780288 error  
744334592493166593 error  
744234799360020481 error  
744223424764059648 error  
743980027717509120 error  
743895849529389061 error  
743835915802583040 error  
743609206067040256 error  
743595368194129920 error  
743545585370791937 error  
743510151680958465 error  
743253157753532416 error  
743222593470234624 error  
743210557239623680 error  
742534281772302336 error  
742528092657332225 error  
742465774154047488 error  
742423170473463808 error  
742385895052087300 error  
742161199639494656 error  
742150209887731712 error  
741793263812808706 error  
741743634094141440 error  
741438259667034112 error  
741303864243200000 error  
741099773336379392 error  
741067306818797568 error  
740995100998766593 error  
740711788199743490 error  
740699697422163968 error  
740676976021798912 error  
740373189193256964 error  
740365076218183684 error  
740359016048689152 error  
740214038584557568 error  
739979191639244800 error  
739932936087216128 error  
739844404073074688 error  
739623569819336705 error  
739606147276148736 error  
739544079319588864 error  
739485634323156992 error  
739238157791694849 error  
738891149612572673 error

738885046782832640 error  
738883359779196928 error  
738537504001953792 error  
738402415918125056 error  
738184450748633089 error  
738166403467907072 error  
738156290900254721 error  
737826014890496000 error  
737800304142471168 error  
737678689543020544 error  
737445876994609152 error  
737322739594330112 error  
737310737551491075 error  
736736130620620800 error  
736392552031657984 error  
736365877722001409 error  
736225175608430592 error  
736010884653420544 error  
735991953473572864 error  
735648611367784448 error  
735635087207878657 error  
735274964362878976 error  
735256018284875776 error  
735137028879360001 error  
734912297295085568 error  
734787690684657664 error  
734776360183431168 error  
734559631394082816 error  
733828123016450049 error  
733822306246479872 error  
733482008106668032 error  
733460102733135873 error  
733109485275860992 error  
732732193018155009 error  
732726085725589504 error  
732585889486888962 error  
732375214819057664 error  
732005617171337216 error  
731285275100512256 error  
731156023742988288 error  
730924654643314689 error  
730573383004487680 error  
730427201120833536 error  
730211855403241472 error  
730196704625098752 error  
729854734790754305 error  
729838605770891264 error  
729823566028484608 error

729463711119904772 error  
729113531270991872 error  
728986383096946689 error  
728760639972315136 error  
728751179681943552 error  
728653952833728512 error  
728409960103686147 error  
728387165835677696 error  
728046963732717569 error  
728035342121635841 error  
728015554473250816 error  
727685679342333952 error  
727644517743104000 error  
727524757080539137 error  
727314416056803329 error  
727286334147182592 error  
727175381690781696 error  
727155742655025152 error  
726935089318363137 error  
726887082820554753 error  
726828223124897792 error  
726224900189511680 error  
725842289046749185 error  
725786712245440512 error  
725729321944506368 error  
725458796924002305 error  
724983749226668032 error  
724771698126512129 error  
724405726123311104 error  
724049859469295616 error  
724046343203856385 error  
724004602748780546 error  
723912936180330496 error  
723688335806480385 error  
723673163800948736 error  
723179728551723008 error  
722974582966214656 error  
722613351520608256 error  
721503162398597120 error  
721001180231503872 error  
720785406564900865 error  
720775346191278080 error  
720415127506415616 error  
720389942216527872 error  
720340705894408192 error  
720059472081784833 error  
720043174954147842 error  
71999115435222208 error

719704490224398336 error  
719551379208073216 error  
719367763014393856 error  
719339463458033665 error  
719332531645071360 error  
718971898235854848 error  
718939241951195136 error  
718631497683582976 error  
718613305783398402 error  
718540630683709445 error  
718460005985447936 error  
718454725339934721 error  
718246886998687744 error  
718234618122661888 error  
717841801130979328 error  
717790033953034240 error  
717537687239008257 error  
717428917016076293 error  
717421804990701568 error  
717047459982213120 error  
717009362452090881 error  
716802964044845056 error  
716791146589110272 error  
716730379797970944 error  
716447146686459905 error  
716439118184652801 error  
716285507865542656 error  
716080869887381504 error  
715928423106027520 error  
715758151270801409 error  
715733265223708672 error  
715704790270025728 error  
715696743237730304 error  
715680795826982913 error  
715360349751484417 error  
715342466308784130 error  
715220193576927233 error  
715200624753819648 error  
715009755312439296 error  
714982300363173890 error  
714962719905021952 error  
714957620017307648 error  
714631576617938945 error  
714606013974974464 error  
714485234495041536 error  
714258258790387713 error  
714251586676113411 error  
714214115368108032 error

714141408463036416 error  
713919462244790272 error  
713909862279876608 error  
713900603437621249 error  
713761197720473600 error  
713411074226274305 error  
713177543487135744 error  
713175907180089344 error  
712809025985978368 error  
712717840512598017 error  
712668654853337088 error  
712438159032893441 error  
712309440758808576 error  
712097430750289920 error  
712092745624633345 error  
712085617388212225 error  
712065007010385924 error  
711998809858043904 error  
711968124745228288 error  
711743778164514816 error  
711732680602345472 error  
711694788429553666 error  
711652651650457602 error  
711363825979756544 error  
711306686208872448 error  
711008018775851008 error  
710997087345876993 error  
710844581445812225 error  
710833117892898816 error  
710658690886586372 error  
710609963652087808 error  
710588934686908417 error  
710296729921429505 error  
710283270106132480 error  
710272297844797440 error  
710269109699739648 error  
710153181850935296 error  
710140971284037632 error  
710117014656950272 error  
709918798883774466 error  
709901256215666688 error  
709852847387627521 error  
709566166965075968 error  
709556954897764353 error  
709519240576036864 error  
709449600415961088 error  
709409458133323776 error  
709225125749587968 error

709207347839836162 error  
709198395643068416 error  
709179584944730112 error  
709158332880297985 error  
709042156699303936 error  
708853462201716736 error  
708845821941387268 error  
708834316713893888 error  
708810915978854401 error  
708738143638450176 error  
708711088997666817 error  
708479650088034305 error  
708469915515297792 error  
708400866336894977 error  
708356463048204288 error  
708349470027751425 error  
708149363256774660 error  
708130923141795840 error  
708119489313951744 error  
708109389455101952 error  
708026248782585858 error  
707995814724026368 error  
707983188426153984 error  
707969809498152960 error  
707776935007539200 error  
707741517457260545 error  
707738799544082433 error  
707693576495472641 error  
707629649552134146 error  
707610948723478529 error  
707420581654872064 error  
707411934438625280 error  
707387676719185920 error  
707377100785885184 error  
707315916783140866 error  
707297311098011648 error  
707059547140169728 error  
707038192327901184 error  
707021089608753152 error  
707014260413456384 error  
706904523814649856 error  
706901761596989440 error  
706681918348251136 error  
706644897839910912 error  
706593038911545345 error  
706538006853918722 error  
706516534877929472 error  
706346369204748288 error

706310011488698368 error  
706291001778950144 error  
706265994973601792 error  
706169069255446529 error  
706166467411222528 error  
706153300320784384 error  
705975130514706432 error  
705970349788291072 error  
705898680587526145 error  
705786532653883392 error  
705591895322394625 error  
705475953783398401 error  
705442520700944385 error  
705428427625635840 error  
705239209544720384 error  
705223444686888960 error  
705102439679201280 error  
705066031337840642 error  
704871453724954624 error  
704859558691414016 error  
704847917308362754 error  
704819833553219584 error  
704761120771465216 error  
704499785726889984 error  
704491224099647488 error  
704480331685040129 error  
704364645503647744 error  
704347321748819968 error  
704134088924532736 error  
704113298707505153 error  
704054845121142784 error  
703774238772166656 error  
703769065844768768 error  
703631701117943808 error  
703611486317502464 error  
703425003149250560 error  
703407252292673536 error  
703382836347330562 error  
703356393781329922 error  
703268521220972544 error  
703079050210877440 error  
703041949650034688 error  
702932127499816960 error  
702899151802126337 error  
702684942141153280 error  
702671118226825216 error  
702598099714314240 error  
702539513671897089 error



702332542343577600 error  
702321140488925184 error  
702276748847800320 error  
702217446468493312 error  
701981390485725185 error  
701952816642965504 error  
701889187134500865 error  
701805642395348998 error  
701601587219795968 error  
701570477911896070 error  
701545186879471618 error  
701214700881756160 error  
700890391244103680 error  
700864154249383937 error  
700847567345688576 error  
700796979434098688 error  
700747788515020802 error  
700518061187723268 error  
700505138482569216 error  
700462010979500032 error  
700167517596164096 error  
700151421916807169 error  
700143752053182464 error  
700062718104104960 error  
700029284593901568 error  
700002074055016451 error  
699801817392291840 error  
699788877217865730 error  
699779630832685056 error  
699775878809702401 error  
699691744225525762 error  
699446877801091073 error  
699434518667751424 error  
699423671849451520 error  
699413908797464576 error  
699370870310113280 error  
699323444782047232 error  
699088579889332224 error  
699079609774645248 error  
699072405256409088 error  
699060279947165696 error  
699036661657767936 error  
698989035503689728 error  
698953797952008193 error  
698907974262222848 error  
698710712454139905 error  
698703483621523456 error  
698635131305795584 error

698549713696649216 error  
698355670425473025 error  
698342080612007937 error  
698262614669991936 error  
698195409219559425 error  
698178924120031232 error  
697995514407682048 error  
697990423684476929 error  
697943111201378304 error  
697881462549430272 error  
697630435728322560 error  
697616773278015490 error  
697596423848730625 error  
697575480820686848 error  
697516214579523584 error  
697482927769255936 error  
697463031882764288 error  
697270446429966336 error  
697259378236399616 error  
697255105972801536 error  
697242256848379904 error  
696900204696625153 error  
696894894812565505 error  
696886256886657024 error  
696877980375769088 error  
696754882863349760 error  
696744641916489729 error  
696713835009417216 error  
696518437233913856 error  
696490539101908992 error  
696488710901260288 error  
696405997980676096 error  
696100768806522880 error  
695816827381944320 error  
695794761660297217 error  
695767669421768709 error  
695629776980148225 error  
695446424020918272 error  
695409464418041856 error  
695314793360662529 error  
695095422348574720 error  
695074328191332352 error  
695064344191721472 error  
695051054296211456 error  
694925794720792577 error  
694905863685980160 error  
694669722378485760 error  
694356675654983680 error

694352839993344000 error  
694342028726001664 error  
694329668942569472 error  
694206574471057408 error  
694183373896572928 error  
694001791655137281 error  
693993230313091072 error  
693942351086120961 error  
693647888581312512 error  
693644216740769793 error  
693642232151285760 error  
693629975228977152 error  
693622659251335168 error  
693590843962331137 error  
693582294167244802 error  
693486665285931008 error  
693280720173801472 error  
693267061318012928 error  
693262851218264065 error  
693231807727280129 error  
693155686491000832 error  
693109034023534592 error  
693095443459342336 error  
692919143163629568 error  
692905862751522816 error  
692901601640583168 error  
692894228850999298 error  
692828166163931137 error  
692752401762250755 error  
692568918515392513 error  
692535307825213440 error  
692530551048294401 error  
692423280028966913 error  
69241731302332352 error  
692187005137076224 error  
692158366030913536 error  
692142790915014657 error  
692041934689402880 error  
692017291282812928 error  
691820333922455552 error  
691793053716221953 error  
691756958957883396 error  
691675652215414786 error  
691483041324204033 error  
691459709405118465 error  
691444869282295808 error  
691416866452082688 error  
691321916024623104 error

691096613310316544 error  
691090071332753408 error  
690989312272396288 error  
690959652130045952 error  
690938899477221376 error  
690932576555528194 error  
690735892932222976 error  
690728923253055490 error  
690690673629138944 error  
690649993829576704 error  
690607260360429569 error  
690597161306841088 error  
690400367696297985 error  
690374419777196032 error  
690360449368465409 error  
690348396616552449 error  
690248561355657216 error  
690021994562220032 error  
690015576308211712 error  
690005060500217858 error  
689999384604450816 error  
689993469801164801 error  
689977555533848577 error  
689905486972461056 error  
689877686181715968 error  
689835978131935233 error  
689661964914655233 error  
689659372465688576 error  
689623661272240129 error  
689599056876867584 error  
689557536375177216 error  
689517482558820352 error  
689289219123089408 error  
689283819090870273 error  
689280876073582592 error  
689275259254616065 error  
689255633275777024 error  
689154315265683456 error  
689143371370250240 error  
688916208532455424 error  
688908934925697024 error  
688898160958271489 error  
688894073864884227 error  
688828561667567616 error  
688804835492233216 error  
688789766343622656 error  
688547210804498433 error  
688519176466644993 error

688385280030670848 error  
688211956440801280 error  
688179443353796608 error  
688116655151435777 error  
688064179421470721 error  
687841446767013888 error  
687826841265172480 error  
687818504314159109 error  
687807801670897665 error  
687732144991551489 error  
687704180304273409 error  
687664829264453632 error  
687494652870668288 error  
687480748861947905 error  
687476254459715584 error  
687460506001633280 error  
687399393394311168 error  
687317306314240000 error  
687312378585812992 error  
687127927494963200 error  
687124485711986689 error  
687109925361856513 error  
687102708889812993 error  
687096057537363968 error  
686947101016735744 error  
686760001961103360 error  
686749460672679938 error  
686730991906516992 error  
686683045143953408 error  
686618349602762752 error  
686606069955735556 error  
686394059078897668 error  
686386521809772549 error  
686377065986265092 error  
686358356425093120 error  
686286779679375361 error  
686050296934563840 error  
686035780142297088 error  
686034024800862208 error  
686007916130873345 error  
686003207160610816 error  
685973236358713344 error  
685943807276412928 error  
685906723014619143 error  
685681090388975616 error  
685667379192414208 error  
685663452032069632 error  
685641971164143616 error

685547936038666240 error  
685532292383666176 error  
685325112850124800 error  
685321586178670592 error  
685315239903100929 error  
685307451701334016 error  
685268753634967552 error  
685198997565345792 error  
685169283572338688 error  
684969860808454144 error  
684959798585110529 error  
684940049151070208 error  
684926975086034944 error  
684914660081053696 error  
684902183876321280 error  
684880619965411328 error  
684830982659280897 error  
684800227459624960 error  
684594889858887680 error  
684588130326986752 error  
684567543613382656 error  
684538444857667585 error  
684481074559381504 error  
684460069371654144 error  
684241637099323392 error  
684225744407494656 error  
684222868335505415 error  
684200372118904832 error  
684195085588783105 error  
684188786104872960 error  
684177701129875456 error  
684147889187209216 error  
684122891630342144 error  
684097758874210310 error  
683857920510050305 error  
683852578183077888 error  
683849932751646720 error  
683834909291606017 error  
683828599284170753 error  
683773439333797890 error  
683742671509258241 error  
683515932363329536 error  
683498322573824003 error  
683481228088049664 error  
683462770029932544 error  
683449695444799489 error  
683391852557561860 error  
683357973142474752 error

683142553609318400 error  
683111407806746624 error  
683098815881154561 error  
683078886620553216 error  
683030066213818368 error  
682962037429899265 error  
682808988178739200 error  
682788441537560576 error  
682750546109968385 error  
682697186228989953 error  
682662431982772225 error  
682638830361513985 error  
682429480204398592 error  
682406705142087680 error  
682393905736888321 error  
682389078323662849 error  
682303737705140231 error  
682259524040966145 error  
682242692827447297 error  
682088079302213632 error  
682059653698686977 error  
682047327939461121 error  
682032003584274432 error  
682003177596559360 error  
681981167097122816 error  
681891461017812993 error  
681694085539872773 error  
681679526984871937 error  
681654059175129088 error  
681610798867845120 error  
681579835668455424 error  
681523177663676416 error  
681340665377193984 error  
681339448655802368 error  
681320187870711809 error  
681302363064414209 error  
681297372102656000 error  
681281657291280384 error  
681261549936340994 error  
681242418453299201 error  
681231109724700672 error  
681193455364796417 error  
680970795137544192 error  
680959110691590145 error  
680940246314430465 error  
680934982542561280 error  
680913438424612864 error  
680889648562991104 error

680836378243002368 error  
680805554198020098 error  
680801747103793152 error  
680798457301471234 error  
680609293079592961 error  
680583894916304897 error  
680497766108381184 error  
680494726643068929 error  
680473011644985345 error  
680440374763077632 error  
680221482581123072 error  
680206703334408192 error  
680191257256136705 error  
680176173301628928 error  
680161097740095489 error  
680145970311643136 error  
680130881361686529 error  
680115823365742593 error  
680100725817409536 error  
680085611152338944 error  
680070545539371008 error  
680055455951884288 error  
679877062409191424 error  
679872969355714560 error  
679862121895714818 error  
679854723806179328 error  
679844490799091713 error  
679828447187857408 error  
679777920601223168 error  
679736210798047232 error  
679729593985699840 error  
679722016581222400 error  
679530280114372609 error  
679527802031484928 error  
679511351870550016 error  
679503373272485890 error  
679475951516934144 error  
679462823135686656 error  
679405845277462528 error  
679158373988876288 error  
679148763231985668 error  
679132435750195208 error  
679111216690831360 error  
679062614270468097 error  
679047485189439488 error  
679001094530465792 error  
678991772295516161 error  
678969228704284672 error



678800283649069056 error  
678798276842360832 error  
678774928607469569 error  
678767140346941444 error  
678764513869611008 error  
678755239630127104 error  
678740035362037760 error  
678708137298427904 error  
678675843183484930 error  
678643457146150913 error  
678446151570427904 error  
678424312106393600 error  
678410210315247616 error  
678399652199309312 error  
678396796259975168 error  
678389028614488064 error  
678380236862578688 error  
678341075375947776 error  
678334497360859136 error  
678278586130948096 error  
678255464182861824 error  
678023323247357953 error  
678021115718029313 error  
677961670166224897 error  
677918531514703872 error  
677895101218201600 error  
677716515794329600 error  
677700003327029250 error  
677698403548192770 error  
677687604918272002 error  
677673981332312066 error  
677662372920729601 error  
677644091929329666 error  
677573743309385728 error  
677565715327688705 error  
677557565589463040 error  
677547928504967168 error  
677530072887205888 error  
677335745548390400 error  
677334615166730240 error  
677331501395156992 error  
677328882937298944 error  
677314812125323265 error  
677301033169788928 error  
677269281705472000 error  
677228873407442944 error  
677187300187611136 error  
676975532580409345 error

676957860086095872 error  
676949632774234114 error  
676948236477857792 error  
676946864479084545 error  
676942428000112642 error  
676936541936185344 error  
676916996760600576 error  
676897532954456065 error  
676864501615042560 error  
676821958043033607 error  
676819651066732545 error  
676811746707918848 error  
676776431406465024 error  
676617503762681856 error  
676613908052996102 error  
676606785097199616 error  
676603393314578432 error  
676593408224403456 error  
676590572941893632 error  
676588346097852417 error  
676582956622721024 error  
676575501977128964 error  
676533798876651520 error  
676496375194980353 error  
676470639084101634 error  
676440007570247681 error  
676430933382295552 error  
676263575653122048 error  
676237365392908289 error  
676219687039057920 error  
676215927814406144 error  
676191832485810177 error  
676146341966438401 error  
676121918416756736 error  
676101918813499392 error  
676098748976615425 error  
676089483918516224 error  
675898130735476737 error  
675891555769696257 error  
675888385639251968 error  
675878199931371520 error  
675870721063669760 error  
675853064436391936 error  
675849018447167488 error  
675845657354215424 error  
675822767435051008 error  
675820929667219457 error  
675798442703122432 error

675781562965868544 error  
675740360753160193 error  
675710890956750848 error  
675707330206547968 error  
675706639471788032 error  
675534494439489536 error  
675531475945709568 error  
675522403582218240 error  
675517828909424640 error  
675501075957489664 error  
675497103322386432 error  
675489971617296384 error  
675483430902214656 error  
675432746517426176 error  
675372240448454658 error  
675362609739206656 error  
675354435921575936 error  
675349384339542016 error  
675334060156301312 error  
675166823650848770 error  
675153376133427200 error  
675149409102012420 error  
675147105808306176 error  
675146535592706048 error  
675145476954566656 error  
675135153782571009 error  
675113801096802304 error  
675111688094527488 error  
675109292475830276 error  
675047298674663426 error  
675015141583413248 error  
675006312288268288 error  
675003128568291329 error  
674999807681908736 error  
674805413498527744 error  
674800520222154752 error  
674793399141146624 error  
674790488185167872 error  
674788554665512960 error  
674781762103414784 error  
674774481756377088 error  
674767892831932416 error  
674764817387900928 error  
674754018082705410 error  
674752233200820224 error  
674743008475090944 error  
674742531037511680 error  
674739953134403584 error

674737130913071104 error  
674690135443775488 error  
674670581682434048 error  
674664755118911488 error  
674646392044941312 error  
674644256330530816 error  
674638615994089473 error  
674632714662858753 error  
674606911342424069 error  
674468880899788800 error  
674447403907457024 error  
674436901579923456 error  
674422304705744896 error  
674416750885273600 error  
674410619106390016 error  
674394782723014656 error  
674372068062928900 error  
674330906434379776 error  
674318007229923329 error  
674307341513269249 error  
674291837063053312 error  
674271431610523648 error  
674269164442398721 error  
674265582246694913 error  
674262580978937856 error  
674255168825880576 error  
674082852460433408 error  
674075285688614912 error  
674063288070742018 error  
674053186244734976 error  
674051556661161984 error  
674045139690631169 error  
674042553264685056 error  
674038233588723717 error  
674036086168010753 error  
674024893172875264 error  
674019345211760640 error  
674014384960745472 error  
674008982932058114 error  
673956914389192708 error  
673919437611909120 error  
673906403526995968 error  
673887867907739649 error  
673716320723169284 error  
673715861853720576 error  
673711475735838725 error  
673709992831262724 error  
673708611235921920 error

673707060090052608 error  
673705679337693185 error  
673700254269775872 error  
673697980713705472 error  
673689733134946305 error  
673688752737402881 error  
673686845050527744 error  
673680198160809984 error  
673662677122719744 error  
673656262056419329 error  
673636718965334016 error  
673612854080196609 error  
673583129559498752 error  
673580926094458881 error  
673576835670777856 error  
673363615379013632 error  
673359818736984064 error  
673355879178194945 error  
673352124999274496 error  
673350198937153538 error  
673345638550134785 error  
673343217010679808 error  
673342308415348736 error  
673320132811366400 error  
673317986296586240 error  
673295268553605120 error  
673270968295534593 error  
673240798075449344 error  
673213039743795200 error  
673148804208660480 error  
672997845381865473 error  
672995267319328768 error  
672988786805112832 error  
672984142909456390 error  
672980819271634944 error  
672975131468300288 error  
672970152493887488 error  
672968025906282496 error  
672964561327235073 error  
672902681409806336 error  
672898206762672129 error  
672884426393653248 error  
672877615439593473 error  
672834301050937345 error  
672828477930868736 error  
672640509974827008 error  
672622327801233409 error  
672614745925664768 error

672609152938721280 error  
672604026190569472 error  
672594978741354496 error  
672591762242805761 error  
672591271085670400 error  
672538107540070400 error  
672523490734551040 error  
672488522314567680 error  
672482722825261057 error  
672481316919734272 error  
672475084225949696 error  
672466075045466113 error  
672272411274932228 error  
672267570918129665 error  
672264251789176834 error  
672256522047614977 error  
672254177670729728 error  
672248013293752320 error  
672245253877968896 error  
672239279297454080 error  
672231046314901505 error  
672222792075620352 error  
672205392827572224 error  
672169685991993344 error  
672160042234327040 error  
672139350159835138 error  
672125275208069120 error  
672095186491711488 error  
672082170312290304 error  
672068090318987265 error  
671896809300709376 error  
671891728106971137 error  
671882082306625538 error  
671879137494245376 error  
671874878652489728 error  
671866342182637568 error  
671855973984772097 error  
671789708968640512 error  
671768281401958400 error  
671763349865160704 error  
671744970634719232 error  
671743150407421952 error  
671735591348891648 error  
671729906628341761 error  
671561002136281088 error  
671550332464455680 error  
671547767500775424 error  
671544874165002241 error

671542985629241344 error  
671538301157904385 error  
671536543010570240 error  
671533943490011136 error  
671528761649688577 error  
671520732782923777 error  
671518598289059840 error  
671511350426865664 error  
671504605491109889 error  
671497587707535361 error  
671488513339211776 error  
671486386088865792 error  
671485057807351808 error  
671390180817915904 error  
671362598324076544 error  
671357843010908160 error  
671355857343524864 error  
671347597085433856 error  
671186162933985280 error  
671182547775299584 error  
671166507850801152 error  
671163268581498880 error  
671159727754231808 error  
671154572044468225 error  
671151324042559489 error  
671147085991960577 error  
671141549288370177 error  
671138694582165504 error  
671134062904504320 error  
671122204919246848 error  
671115716440031232 error  
671109016219725825 error  
670995969505435648 error  
670842764863651840 error  
670840546554966016 error  
670838202509447168 error  
670833812859932673 error  
670832455012716544 error  
670826280409919488 error  
670823764196741120 error  
670822709593571328 error  
670815497391357952 error  
670811965569282048 error  
670807719151067136 error  
670804601705242624 error  
670803562457407488 error  
670797304698376195 error  
670792680469889025 error

670789397210615808 error  
670786190031921152 error  
670783437142401025 error  
670782429121134593 error  
670780561024270336 error  
670778058496974848 error  
670764103623966721 error  
670755717859713024 error  
670733412878163972 error  
670727704916926465 error  
670717338665226240 error  
670704688707301377 error  
670691627984359425 error  
670679630144274432 error  
670676092097810432 error  
670668383499735048 error  
670474236058800128 error  
670468609693655041 error  
670465786746662913 error  
670452855871037440 error  
670449342516494336 error  
670444955656130560 error  
670442337873600512 error  
670435821946826752 error  
670434127938719744 error  
670433248821026816 error  
670428280563085312 error  
670427002554466305 error  
670421925039075328 error  
670420569653809152 error  
670417414769758208 error  
670411370698022913 error  
670408998013820928 error  
670403879788544000 error  
670385711116361728 error  
670374371102445568 error  
670361874861563904 error  
670338931251150849 error  
670319130621435904 error  
670303360680108032 error  
670290420111441920 error  
670093938074779648 error  
670086499208155136 error  
670079681849372674 error  
670073503555706880 error  
670069087419133954 error  
670061506722140161 error  
670055038660800512 error



670046952931721218 error  
670040295598354432 error  
670037189829525505 error  
670003130994700288 error  
669993076832759809 error  
669972011175813120 error  
669970042633789440 error  
669942763794931712 error  
669926384437997569 error  
669923323644657664 error  
669753178989142016 error  
669749430875258880 error  
669684865554620416 error  
669683899023405056 error  
669682095984410625 error  
669680153564442624 error  
669661792646373376 error  
669625907762618368 error  
669603084620980224 error  
669597912108789760 error  
669583744538451968 error  
669573570759163904 error  
669571471778410496 error  
669567591774625800 error  
669564461267722241 error  
669393256313184256 error  
669375718304980992 error  
669371483794317312 error  
669367896104181761 error  
669363888236994561 error  
669359674819481600 error  
669354382627049472 error  
669353438988365824 error  
669351434509529089 error  
669328503091937280 error  
669327207240699904 error  
669324657376567296 error  
669216679721873412 error  
669214165781868544 error  
669203728096960512 error  
669037058363662336 error  
669015743032369152 error  
669006782128353280 error  
669000397445533696 error  
668994913074286592 error  
668992363537309700 error  
668989615043424256 error  
668988183816871936 error

668986018524233728 error  
668981893510119424 error  
668979806671884288 error  
668975677807423489 error  
668967877119254528 error  
668960084974809088 error  
668955713004314625 error  
668932921458302977 error  
668902994700836864 error  
668892474547511297 error  
668872652652679168 error  
668852170888998912 error  
668826086256599040 error  
668815180734689280 error  
668779399630725120 error  
668655139528511488 error  
668645506898350081 error  
668643542311546881 error  
668641109086707712 error  
668636665813057536 error  
668633411083464705 error  
668631377374486528 error  
668627278264475648 error  
668625577880875008 error  
668623201287675904 error  
668620235289837568 error  
668614819948453888 error  
668587383441514497 error  
668567822092664832 error  
668544745690562560 error  
668542336805281792 error  
668537837512433665 error  
668528771708952576 error  
668507509523615744 error  
668496999348633600 error  
668484198282485761 error  
668480044826800133 error  
668466899341221888 error  
668297328638447616 error  
668291999406125056 error  
668286279830867968 error  
668274247790391296 error  
668268907921326080 error  
668256321989451776 error  
668248472370458624 error  
668237644992782336 error  
668226093875376128 error  
668221241640230912 error

668204964695683073 error  
668190681446379520 error  
668171859951755264 error  
668154635664932864 error  
668142349051129856 error  
668113020489474048 error  
667937095915278337 error  
667924896115245057 error  
667915453470232577 error  
667911425562669056 error  
667902449697558528 error  
667886921285246976 error  
667885044254572545 error  
667878741721415682 error  
667873844930215936 error  
667866724293877760 error  
667861340749471744 error  
667832474953625600 error  
667806454573760512 error  
667801013445750784 error  
667793409583771648 error  
667782464991965184 error  
667773195014021121 error  
667766675769573376 error  
667728196545200128 error  
667724302356258817 error  
667550904950915073 error  
667550882905632768 error  
667549055577362432 error  
667546741521195010 error  
667544320556335104 error  
667538891197542400 error  
667534815156183040 error  
667530908589760512 error  
667524857454854144 error  
667517642048163840 error  
667509364010450944 error  
667502640335572993 error  
667495797102141441 error  
667491009379606528 error  
667470559035432960 error  
667455448082227200 error  
667453023279554560 error  
667443425659232256 error  
667437278097252352 error  
667435689202614272 error  
667405339315146752 error  
667393430834667520 error

667369227918143488 error  
667211855547486208 error  
667200525029539841 error  
667192066997374976 error  
667188689915760640 error  
667182792070062081 error  
667177989038297088 error  
667176164155375616 error  
667174963120574464 error  
667171260800061440 error  
667165590075940865 error  
667160273090932737 error  
667152164079423490 error  
667138269671505920 error  
667119796878725120 error  
667090893657276420 error  
667073648344346624 error  
667070482143944705 error  
667065535570550784 error  
667062181243039745 error  
667044094246576128 error  
667012601033924608 error  
666996132027977728 error  
666983947667116034 error  
666837028449972224 error  
666835007768551424 error  
666826780179869698 error  
666817836334096384 error  
666804364988780544 error  
666786068205871104 error  
666781792255496192 error  
666776908487630848 error  
666739327293083650 error  
666701168228331520 error  
666691418707132416 error  
666649482315059201 error  
666644823164719104 error  
666454714377183233 error  
666447344410484738 error  
666437273139982337 error  
666435652385423360 error  
666430724426358785 error  
666428276349472768 error  
666421158376562688 error  
666418789513326592 error  
666411507551481857 error  
666407126856765440 error  
666396247373291520 error

```

666373753744588802 error
666362758909284353 error
666353288456101888 error
666345417576210432 error
666337882303524864 error
666293911632134144 error
666287406224695296 error
666273097616637952 error
666268910803644416 error
666104133288665088 error
666102155909144576 error
666099513787052032 error
666094000022159362 error
666082916733198337 error
666073100786774016 error
666071193221509120 error
666063827256086533 error
666058600524156928 error
666057090499244032 error
666055525042405380 error
666051853826850816 error
666050758794694657 error
666049248165822465 error
666044226329800704 error
666033412701032449 error
666029285002620928 error
666020888022790149 error

```

```

In [7]: # Load data from json_tweet.txt
        data = []
        with open('tweet_json.txt') as f:
            for line in f:
                data.append(json.loads(line))

        # Count number of line in this file
        count = len(open('tweet_json.txt').readlines())

        # Retrieve the id, favorite_count, and retweet_count and store the data into a new dataframe
        df_list = []
        for index in range(count-1):
            tweet_id = data[index]['id']
            fav_count = data[index]['favorite_count']
            ret_count = data[index]['retweet_count']
            df_list.append({'tweet_id':tweet_id,
                           'favorite_count':fav_count,
                           'retweet_count': ret_count})
        tweet_count = pd.DataFrame(df_list,columns = ['tweet_id','favorite_count','retweet_c

```

## 1.1 Assessing Data

### Quality archive table

- Useless retweeted info (we only want the original ratings according to instructions)
- Inconsistency with rating numerator and denominator columns
- Rating denominator equals to "0" for tweet\_id:835246439529840640
- Redundant information in the "text" column (repeated rating and short link)
- Ambiguous information in the "source" column

### Tidiness

- One variable (dog\_stage) in four columns in archive table
- Three sub-tables should be merged into one table as tweet\_id duplicated in all datasets

**Note:** concerning the missing values in tweet\_count and image tables, I chose the minimum number (2075, same as the total number in imagetable) since we target to the original ratings that have pictures.

In [8]: *#archive table*

```
archive.head()
```

```
Out[8]:
```

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	\
0	892420643555336193	NaN	NaN	
1	892177421306343426	NaN	NaN	
2	891815181378084864	NaN	NaN	
3	891689557279858688	NaN	NaN	
4	891327558926688256	NaN	NaN	

	timestamp	\
0	2017-08-01 16:23:56 +0000	
1	2017-08-01 00:17:27 +0000	
2	2017-07-31 00:18:03 +0000	
3	2017-07-30 15:58:51 +0000	
4	2017-07-29 16:00:24 +0000	

	source	\
0	<a href="http://twitter.com/download/iphone" r...	
1	<a href="http://twitter.com/download/iphone" r...	
2	<a href="http://twitter.com/download/iphone" r...	
3	<a href="http://twitter.com/download/iphone" r...	
4	<a href="http://twitter.com/download/iphone" r...	

	text	retweeted_status_id	\
0	This is Phineas. He's a mystical boy. Only eve...	NaN	
1	This is Tilly. She's just checking pup on you...	NaN	
2	This is Archie. He is a rare Norwegian Pouncin...	NaN	
3	This is Darla. She commenced a snooze mid meal...	NaN	
4	This is Franklin. He would like you to stop ca...	NaN	

	retweeted_status_user_id	retweeted_status_timestamp	\
0	NaN	NaN	
1	NaN	NaN	
2	NaN	NaN	
3	NaN	NaN	
4	NaN	NaN	

	expanded_urls	rating_numerator	\
0	https://twitter.com/dog_rates/status/892420643...	13	
1	https://twitter.com/dog_rates/status/892177421...	13	
2	https://twitter.com/dog_rates/status/891815181...	12	
3	https://twitter.com/dog_rates/status/891689557...	13	
4	https://twitter.com/dog_rates/status/891327558...	12	

	rating_denominator	name	doggo	floofer	pupper	puppo
0	10	Phineas	None	None	None	None
1	10	Tilly	None	None	None	None
2	10	Archie	None	None	None	None
3	10	Darla	None	None	None	None
4	10	Franklin	None	None	None	None

```
In [9]: archive.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2356 entries, 0 to 2355
Data columns (total 17 columns):
tweet_id                2356 non-null int64
in_reply_to_status_id    78 non-null float64
in_reply_to_user_id      78 non-null float64
timestamp                2356 non-null object
source                  2356 non-null object
text                    2356 non-null object
retweeted_status_id      181 non-null float64
retweeted_status_user_id 181 non-null float64
retweeted_status_timestamp 181 non-null object
expanded_urls            2297 non-null object
rating_numerator         2356 non-null int64
rating_denominator       2356 non-null int64
name                    2356 non-null object
doggo                   2356 non-null object
floofer                 2356 non-null object
pupper                  2356 non-null object
puppo                   2356 non-null object
dtypes: float64(4), int64(3), object(10)
memory usage: 313.0+ KB
```

```
In [10]: archive[['rating_numerator', 'rating_denominator']].describe()
```

```
Out[10]:
```

	rating_numerator	rating_denominator
count	2356.000000	2356.000000
mean	13.126486	10.455433
std	45.876648	6.745237
min	0.000000	0.000000
25%	10.000000	10.000000
50%	11.000000	10.000000
75%	12.000000	10.000000
max	1776.000000	170.000000

```
In [11]: archive.query('rating_numerator == 0')
```

```
Out[11]:
```

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	\
315	835152434251116546	NaN	NaN	
1016	746906459439529985	7.468859e+17	4.196984e+09	

	timestamp	\
315	2017-02-24 15:40:31 +0000	
1016	2016-06-26 03:22:31 +0000	

	source	\
315	<a href="http://twitter.com/download/iphone" r...	
1016	<a href="http://twitter.com/download/iphone" r...	

	text	retweeted_status_id	\
315	When you're so blinded by your systematic plag...	NaN	
1016	PUPDATE: can't see any. Even if I could, I cou...	NaN	

	retweeted_status_user_id	retweeted_status_timestamp	\
315	NaN	NaN	
1016	NaN	NaN	

	expanded_urls	rating_numerator	\
315	https://twitter.com/dog_rates/status/835152434...	0	
1016	https://twitter.com/dog_rates/status/746906459...	0	

	rating_denominator	name	doggo	floofer	pupper	puppo
315	10	None	None	None	None	None
1016	10	None	None	None	None	None

```
In [12]: # Check the info about dogs rated "0"
archive.expanded_urls[315]
```

```
Out[12]: 'https://twitter.com/dog_rates/status/835152434251116546/photo/1,https://twitter.com/do
```

```
In [13]: archive.expanded_urls[1016]
```

```
Out[13]: 'https://twitter.com/dog_rates/status/746906459439529985/photo/1'
```



```

In [14]: # Check duplicated id
         archive[archive.tweet_id.duplicated()]

Out[14]: Empty DataFrame
         Columns: [tweet_id, in_reply_to_status_id, in_reply_to_user_id, timestamp, source, text]
         Index: []

In [15]: # Check the information about "source" column
         archive.source.value_counts()

Out[15]: <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>
         <a href="http://vine.co" rel="nofollow">Vine - Make a Scene</a>
         <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
         <a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">TweetDeck</a>
         Name: source, dtype: int64

In [16]: # Look at
         archive.expanded_urls[0]

Out[16]: 'https://twitter.com/dog_rates/status/892420643555336193/photo/1'

In [17]: archive.text[0]

Out[17]: "This is Phineas. He's a mystical boy. Only ever appears in the hole of a donut. 13/10

In [18]: #tweet_count table
         tweet_count

Out[18]:
         tweet_id  favorite_count  retweet_count
0      892420643555336193          35916          7627
1      892177421306343426          30993          5645
2      891815181378084864          23314          3739
3      891689557279858688          39204          7795
4      891327558926688256          37448          8393
5      891087950875897856          18871          2817
6      890971913173991426          10961          1833
7      890729181411237888          60492         17026
8      890609185150312448          25968          3876
9      890240255349198849          29653          6626
10     890006608113172480          28559          6613
11     889880896479866881          25981          4495
12     889665388333682689          44708          9021
13     889638837579907072          25168          4044
14     889531135344209921          14100          2037
15     889278841981685760          23447          4815
16     888917238123831296          27097          4055
17     888804989199671297          23793          3823
18     888554962724278272          18370          3139
19     888078434458587136          20277          3136

```

20	887705289381826560	28166	4878
21	887517139158093824	43185	10637
22	887473957103951883	63973	16254
23	887343217045368832	31338	9464
24	887101392804085760	28512	5394
25	886983233522544640	32484	6929
26	886736880519319552	11152	2899
27	886680336477933568	20928	4043
28	886366144734445568	19675	2860
29	886267009285017600	111	4
...	...	...	...
1723	761750502866649088	0	3911
1724	761745352076779520	4271	861
1725	761672994376806400	49476	29060
1726	761599872357261312	4121	1187
1727	761371037149827077	0	18040
1728	761334018830917632	5193	1449
1729	761292947749015552	4489	1084
1730	761227390836215808	5323	1562
1731	761004547850530816	11323	3467
1732	760893934457552897	3814	975
1733	760656994973933572	6584	1906
1734	760641137271070720	4977	1293
1735	760539183865880579	7557	3613
1736	760521673607086080	4215	1384
1737	760290219849637889	26745	11462
1738	760252756032651264	3944	866
1739	760190180481531904	5688	1799
1740	760153949710192640	0	31
1741	759943073749200896	5943	2076
1742	759923798737051648	14658	5633
1743	759846353224826880	6730	1964
1744	759793422261743616	5952	1904
1745	759557299618865152	4687	1181
1746	759447681597108224	8471	2452
1747	759446261539934208	1672	485
1748	759197388317847553	6069	1942
1749	759159934323924993	0	1153
1750	759099523532779520	14479	4159
1751	759047813560868866	6471	2004
1752	758854675097526272	3544	893

[1753 rows x 3 columns]

```
In [19]: # Before accessing the image table, I first cleaned the useless retweeted tweet to get t
# that we need for this data wrangling.
archive = archive[archive.retweeted_status_id.isnull()]
archive = archive[archive.in_reply_to_status_id.isnull()]
```

```
archive.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 17 columns):
tweet_id                2097 non-null int64
in_reply_to_status_id   0 non-null float64
in_reply_to_user_id     0 non-null float64
timestamp               2097 non-null object
source                  2097 non-null object
text                    2097 non-null object
retweeted_status_id     0 non-null float64
retweeted_status_user_id 0 non-null float64
retweeted_status_timestamp 0 non-null object
expanded_urls           2094 non-null object
rating_numerator        2097 non-null int64
rating_denominator      2097 non-null int64
name                    2097 non-null object
doggo                   2097 non-null object
floofer                 2097 non-null object
pupper                  2097 non-null object
puppo                   2097 non-null object
dtypes: float64(4), int64(3), object(10)
memory usage: 294.9+ KB
```

```
In [23]: #image table
```

```
image = pd.read_csv('image_predictions.tsv', sep='\t')
image
```

```
Out[23]:
```

	tweet_id	jpg_url \
0	666020888022790149	<a href="https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg</a>
1	666029285002620928	<a href="https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg">https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg</a>
2	666033412701032449	<a href="https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg">https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg</a>
3	666044226329800704	<a href="https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg">https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg</a>
4	666049248165822465	<a href="https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg">https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg</a>
5	666050758794694657	<a href="https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg">https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg</a>
6	666051853826850816	<a href="https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg">https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg</a>
7	666055525042405380	<a href="https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg">https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg</a>
8	666057090499244032	<a href="https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg">https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg</a>
9	666058600524156928	<a href="https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg">https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg</a>
10	666063827256086533	<a href="https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg">https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg</a>
11	666071193221509120	<a href="https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg">https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg</a>
12	666073100786774016	<a href="https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg">https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg</a>
13	666082916733198337	<a href="https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg">https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg</a>
14	666094000022159362	<a href="https://pbs.twimg.com/media/CT5w9gUW4AAAsBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAAsBNN.jpg</a>
15	666099513787052032	<a href="https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg">https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg</a>
16	666102155909144576	<a href="https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg">https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg</a>

17	666104133288665088	<a href="https://pbs.twimg.com/media/CT56LSZW0AA1Jj2.jpg">https://pbs.twimg.com/media/CT56LSZW0AA1Jj2.jpg</a>
18	666268910803644416	<a href="https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg">https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg</a>
19	666273097616637952	<a href="https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg">https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg</a>
20	666287406224695296	<a href="https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg">https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg</a>
21	666293911632134144	<a href="https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg">https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg</a>
22	666337882303524864	<a href="https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg">https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg</a>
23	666345417576210432	<a href="https://pbs.twimg.com/media/CT9Vn7PW0AA_ZCM.jpg">https://pbs.twimg.com/media/CT9Vn7PW0AA_ZCM.jpg</a>
24	666353288456101888	<a href="https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg">https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg</a>
25	666362758909284353	<a href="https://pbs.twimg.com/media/CT9lXGsUcAAyUft.jpg">https://pbs.twimg.com/media/CT9lXGsUcAAyUft.jpg</a>
26	666373753744588802	<a href="https://pbs.twimg.com/media/CT9vZEYUAA1Z05.jpg">https://pbs.twimg.com/media/CT9vZEYUAA1Z05.jpg</a>
27	666396247373291520	<a href="https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg">https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg</a>
28	666407126856765440	<a href="https://pbs.twimg.com/media/CT-NvwmW4AAuGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAuGZ.jpg</a>
29	666411507551481857	<a href="https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg">https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg</a>
...	...	...
2045	886366144734445568	<a href="https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg">https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg</a>
2046	886680336477933568	<a href="https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg">https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg</a>
2047	886736880519319552	<a href="https://pbs.twimg.com/media/DE5Se8FXcAAJF4.jpg">https://pbs.twimg.com/media/DE5Se8FXcAAJF4.jpg</a>
2048	886983233522544640	<a href="https://pbs.twimg.com/media/DE8yicJW0AAAABJ.jpg">https://pbs.twimg.com/media/DE8yicJW0AAAABJ.jpg</a>
2049	887101392804085760	<a href="https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg">https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg</a>
2050	887343217045368832	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88734...">https://pbs.twimg.com/ext_tw_video_thumb/88734...</a>
2051	887473957103951883	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2052	887517139158093824	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88751...">https://pbs.twimg.com/ext_tw_video_thumb/88751...</a>
2053	887705289381826560	<a href="https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg">https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg</a>
2054	888078434458587136	<a href="https://pbs.twimg.com/media/DFMWn56WsAAkA7B.jpg">https://pbs.twimg.com/media/DFMWn56WsAAkA7B.jpg</a>
2055	888202515573088257	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2056	888554962724278272	<a href="https://pbs.twimg.com/media/DFTH_0-UQAAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAAACu20.jpg</a>
2057	888804989199671297	<a href="https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg">https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg</a>
2058	888917238123831296	<a href="https://pbs.twimg.com/media/DFYRgsOUQAARGh0.jpg">https://pbs.twimg.com/media/DFYRgsOUQAARGh0.jpg</a>
2059	889278841981685760	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88927...">https://pbs.twimg.com/ext_tw_video_thumb/88927...</a>
2060	889531135344209921	<a href="https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg">https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg</a>
2061	889638837579907072	<a href="https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg">https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg</a>
2062	889665388333682689	<a href="https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg">https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg</a>
2063	889880896479866881	<a href="https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg">https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg</a>
2064	890006608113172480	<a href="https://pbs.twimg.com/media/DFnwSY4WAAAMliS.jpg">https://pbs.twimg.com/media/DFnwSY4WAAAMliS.jpg</a>
2065	890240255349198849	<a href="https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg">https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg</a>
2066	890609185150312448	<a href="https://pbs.twimg.com/media/DFwUU_-XcAEpyXI.jpg">https://pbs.twimg.com/media/DFwUU_-XcAEpyXI.jpg</a>
2067	890729181411237888	<a href="https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg">https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg</a>
2068	890971913173991426	<a href="https://pbs.twimg.com/media/DF1eOmZXUAAALucq.jpg">https://pbs.twimg.com/media/DF1eOmZXUAAALucq.jpg</a>
2069	891087950875897856	<a href="https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg">https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg</a>
2070	891327558926688256	<a href="https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg</a>
2071	891689557279858688	<a href="https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg">https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg</a>
2072	891815181378084864	<a href="https://pbs.twimg.com/media/DGBdLU1WsAANxJ9.jpg">https://pbs.twimg.com/media/DGBdLU1WsAANxJ9.jpg</a>
2073	892177421306343426	<a href="https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg">https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg</a>
2074	892420643555336193	<a href="https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg">https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg</a>

	img_num		p1	p1_conf	p1_dog	\
0	1	Welsh_springer_spaniel	0.465074		True	
1	1	redbone	0.506826		True	

2	1	German_shepherd	0.596461	True
3	1	Rhodesian_ridgeback	0.408143	True
4	1	miniature_pinscher	0.560311	True
5	1	Bernese_mountain_dog	0.651137	True
6	1	box_turtle	0.933012	False
7	1	chow	0.692517	True
8	1	shopping_cart	0.962465	False
9	1	miniature_poodle	0.201493	True
10	1	golden_retriever	0.775930	True
11	1	Gordon_setter	0.503672	True
12	1	Walker_hound	0.260857	True
13	1	pug	0.489814	True
14	1	bloodhound	0.195217	True
15	1	Lhasa	0.582330	True
16	1	English_setter	0.298617	True
17	1	hen	0.965932	False
18	1	desktop_computer	0.086502	False
19	1	Italian_greyhound	0.176053	True
20	1	Maltese_dog	0.857531	True
21	1	three-toed_sloth	0.914671	False
22	1	ox	0.416669	False
23	1	golden_retriever	0.858744	True
24	1	malamute	0.336874	True
25	1	guinea_pig	0.996496	False
26	1	soft-coated_wheaten_terrier	0.326467	True
27	1	Chihuahua	0.978108	True
28	1	black-and-tan_coonhound	0.529139	True
29	1	coho	0.404640	False
...	...	...	...	...
2045	1	French_bulldog	0.999201	True
2046	1	convertible	0.738995	False
2047	1	kuvasz	0.309706	True
2048	2	Chihuahua	0.793469	True
2049	1	Samoyed	0.733942	True
2050	1	Mexican_hairless	0.330741	True
2051	2	Pembroke	0.809197	True
2052	1	limousine	0.130432	False
2053	1	basset	0.821664	True
2054	1	French_bulldog	0.995026	True
2055	2	Pembroke	0.809197	True
2056	3	Siberian_husky	0.700377	True
2057	1	golden_retriever	0.469760	True
2058	1	golden_retriever	0.714719	True
2059	1	whippet	0.626152	True
2060	1	golden_retriever	0.953442	True
2061	1	French_bulldog	0.991650	True
2062	1	Pembroke	0.966327	True
2063	1	French_bulldog	0.377417	True

2064	1	Samoyed	0.957979	True
2065	1	Pembroke	0.511319	True
2066	1	Irish_terrier	0.487574	True
2067	2	Pomeranian	0.566142	True
2068	1	Appenzeller	0.341703	True
2069	1	Chesapeake_Bay_retriever	0.425595	True
2070	2	basset	0.555712	True
2071	1	paper_towel	0.170278	False
2072	1	Chihuahua	0.716012	True
2073	1	Chihuahua	0.323581	True
2074	1	orange	0.097049	False

		p2	p2_conf	p2_dog		p3 \
0		collie	0.156665	True		Shetland_sheepdog
1	miniature_pinscher		0.074192	True		Rhodesian_ridgeback
2	malinois		0.138584	True		bloodhound
3	redbone		0.360687	True		miniature_pinscher
4	Rottweiler		0.243682	True		Doberman
5	English_springer		0.263788	True	Greater_Swiss_Mountain_dog	
6	mud_turtle		0.045885	False		terrapin
7	Tibetan_mastiff		0.058279	True		fur_coat
8	shopping_basket		0.014594	False		golden_retriever
9	komondor		0.192305	True	soft-coated_wheaten_terrier	
10	Tibetan_mastiff		0.093718	True		Labrador_retriever
11	Yorkshire_terrier		0.174201	True		Pekinese
12	English_foxhound		0.175382	True		Ibizan_hound
13	bull_mastiff		0.404722	True		French_bulldog
14	German_shepherd		0.078260	True		malinois
15	Shih-Tzu		0.166192	True		Dandie_Dinmont
16	Newfoundland		0.149842	True		borzoi
17	cock		0.033919	False		partridge
18	desk		0.085547	False		bookcase
19	toy_terrier		0.111884	True		basenji
20	toy_poodle		0.063064	True		miniature_poodle
21	otter		0.015250	False		great_grey_owl
22	Newfoundland		0.278407	True		groenendael
23	Chesapeake_Bay_retriever		0.054787	True		Labrador_retriever
24	Siberian_husky		0.147655	True		Eskimo_dog
25	skunk		0.002402	False		hamster
26	Afghan_hound		0.259551	True		briard
27	toy_terrier		0.009397	True		papillon
28	bloodhound		0.244220	True	flat-coated_retriever	
29	barracouta		0.271485	False		gar
...	...		...	...		...
2045	Chihuahua		0.000361	True		Boston_bull
2046	sports_car		0.139952	False		car_wheel
2047	Great_Pyrenees		0.186136	True		Dandie_Dinmont
2048	toy_terrier		0.143528	True		can_opener

2049	Eskimo_dog	0.035029	True	Staffordshire_bullterrier
2050	sea_lion	0.275645	False	Weimaraner
2051	Rhodesian_ridgeback	0.054950	True	beagle
2052	tow_truck	0.029175	False	shopping_cart
2053	redbone	0.087582	True	Weimaraner
2054	pug	0.000932	True	bull_mastiff
2055	Rhodesian_ridgeback	0.054950	True	beagle
2056	Eskimo_dog	0.166511	True	malamute
2057	Labrador_retriever	0.184172	True	English_setter
2058	Tibetan_mastiff	0.120184	True	Labrador_retriever
2059	borzoi	0.194742	True	Saluki
2060	Labrador_retriever	0.013834	True	redbone
2061	boxer	0.002129	True	Staffordshire_bullterrier
2062	Cardigan	0.027356	True	basenji
2063	Labrador_retriever	0.151317	True	muzzle
2064	Pomeranian	0.013884	True	chow
2065	Cardigan	0.451038	True	Chihuahua
2066	Irish_setter	0.193054	True	Chesapeake_Bay_retriever
2067	Eskimo_dog	0.178406	True	Pembroke
2068	Border_collie	0.199287	True	ice_lolly
2069	Irish_terrier	0.116317	True	Indian_elephant
2070	English_springer	0.225770	True	German_short-haired_pointer
2071	Labrador_retriever	0.168086	True	spatula
2072	malamute	0.078253	True	kelpie
2073	Pekinese	0.090647	True	papillon
2074	bagel	0.085851	False	banana

	p3_conf	p3_dog
0	0.061428	True
1	0.072010	True
2	0.116197	True
3	0.222752	True
4	0.154629	True
5	0.016199	True
6	0.017885	False
7	0.054449	False
8	0.007959	True
9	0.082086	True
10	0.072427	True
11	0.109454	True
12	0.097471	True
13	0.048960	True
14	0.075628	True
15	0.089688	True
16	0.133649	True
17	0.000052	False
18	0.079480	False
19	0.111152	True

20	0.025581	True
21	0.013207	False
22	0.102643	True
23	0.014241	True
24	0.093412	True
25	0.000461	False
26	0.206803	True
27	0.004577	True
28	0.173810	True
29	0.189945	False
...	...	...
2045	0.000076	True
2046	0.044173	False
2047	0.086346	True
2048	0.032253	False
2049	0.029705	True
2050	0.134203	True
2051	0.038915	True
2052	0.026321	False
2053	0.026236	True
2054	0.000903	True
2055	0.038915	True
2056	0.111411	True
2057	0.073482	True
2058	0.105506	True
2059	0.027351	True
2060	0.007958	True
2061	0.001498	True
2062	0.004633	True
2063	0.082981	False
2064	0.008167	True
2065	0.029248	True
2066	0.118184	True
2067	0.076507	True
2068	0.193548	False
2069	0.076902	False
2070	0.175219	True
2071	0.040836	False
2072	0.031379	True
2073	0.068957	True
2074	0.076110	False

[2075 rows x 12 columns]

In [24]: image.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2075 entries, 0 to 2074
```



```
Data columns (total 12 columns):
tweet_id      2075 non-null int64
jpg_url       2075 non-null object
img_num       2075 non-null int64
p1            2075 non-null object
p1_conf       2075 non-null float64
p1_dog        2075 non-null bool
p2            2075 non-null object
p2_conf       2075 non-null float64
p2_dog        2075 non-null bool
p3            2075 non-null object
p3_conf       2075 non-null float64
p3_dog        2075 non-null bool
dtypes: bool(3), float64(3), int64(2), object(4)
memory usage: 152.1+ KB
```

```
In [25]: image.describe()
```

```
Out[25]:
```

	tweet_id	img_num	p1_conf	p2_conf	p3_conf
count	2.075000e+03	2075.000000	2075.000000	2.075000e+03	2.075000e+03
mean	7.384514e+17	1.203855	0.594548	1.345886e-01	6.032417e-02
std	6.785203e+16	0.561875	0.271174	1.006657e-01	5.090593e-02
min	6.660209e+17	1.000000	0.044333	1.011300e-08	1.740170e-10
25%	6.764835e+17	1.000000	0.364412	5.388625e-02	1.622240e-02
50%	7.119988e+17	1.000000	0.588230	1.181810e-01	4.944380e-02
75%	7.932034e+17	1.000000	0.843855	1.955655e-01	9.180755e-02
max	8.924206e+17	4.000000	1.000000	4.880140e-01	2.734190e-01

```
In [26]: image.iloc[2070]
```

```
Out[26]:
```

tweet_id	891327558926688256
jpg_url	https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg
img_num	2
p1	basset
p1_conf	0.555712
p1_dog	True
p2	English_springer
p2_conf	0.22577
p2_dog	True
p3	German_short-haired_pointer
p3_conf	0.175219
p3_dog	True
Name:	2070, dtype: object

```
In [27]: image.query('img_num >= 2')
```

```
Out[27]:
```

	tweet_id	jpg_url
144	668623201287675904	https://pbs.twimg.com/media/CUdtP1xUYAIeBnE.jpg

312	671547767500775424	<a href="https://pbs.twimg.com/media/CVHRliqWEAAj98K.jpg">https://pbs.twimg.com/media/CVHRliqWEAAj98K.jpg</a>
315	671735591348891648	<a href="https://pbs.twimg.com/media/CVJ79MzW4AEpTom.jpg">https://pbs.twimg.com/media/CVJ79MzW4AEpTom.jpg</a>
319	671768281401958400	<a href="https://pbs.twimg.com/media/CVKZsHtWwAA6gPj.jpg">https://pbs.twimg.com/media/CVKZsHtWwAA6gPj.jpg</a>
345	672272411274932228	<a href="https://pbs.twimg.com/media/CVRkLuJWUAAhhYp.jpg">https://pbs.twimg.com/media/CVRkLuJWUAAhhYp.jpg</a>
381	673317986296586240	<a href="https://pbs.twimg.com/media/CVgbIobUYAEaeI3.jpg">https://pbs.twimg.com/media/CVgbIobUYAEaeI3.jpg</a>
382	673320132811366400	<a href="https://pbs.twimg.com/media/CVgdfjNWEAAxmbq.jpg">https://pbs.twimg.com/media/CVgdfjNWEAAxmbq.jpg</a>
410	673887867907739649	<a href="https://pbs.twimg.com/media/CVoHa_IU4AAZ7vi.jpg">https://pbs.twimg.com/media/CVoHa_IU4AAZ7vi.jpg</a>
441	674468880899788800	<a href="https://pbs.twimg.com/media/CVwx3dQXAAAOKsL.jpg">https://pbs.twimg.com/media/CVwx3dQXAAAOKsL.jpg</a>
452	674752233200820224	<a href="https://pbs.twimg.com/media/CV0zkzEU4AAZLc5.jpg">https://pbs.twimg.com/media/CV0zkzEU4AAZLc5.jpg</a>
454	674764817387900928	<a href="https://pbs.twimg.com/media/CV0_BSuWIAIvE9k.jpg">https://pbs.twimg.com/media/CV0_BSuWIAIvE9k.jpg</a>
478	675334060156301312	<a href="https://pbs.twimg.com/media/CV9EvZNUwAAgLCK.jpg">https://pbs.twimg.com/media/CV9EvZNUwAAgLCK.jpg</a>
479	675349384339542016	<a href="https://pbs.twimg.com/media/CV9SrABU4AQI46z.jpg">https://pbs.twimg.com/media/CV9SrABU4AQI46z.jpg</a>
494	675710890956750848	<a href="https://pbs.twimg.com/media/CWCbd8ZW0AAatqoH.jpg">https://pbs.twimg.com/media/CWCbd8ZW0AAatqoH.jpg</a>
511	676191832485810177	<a href="https://pbs.twimg.com/media/CWJQ4UmWoAIJ29t.jpg">https://pbs.twimg.com/media/CWJQ4UmWoAIJ29t.jpg</a>
517	676440007570247681	<a href="https://pbs.twimg.com/media/CWMy19EUWAAanZJO.jpg">https://pbs.twimg.com/media/CWMy19EUWAAanZJO.jpg</a>
530	676819651066732545	<a href="https://pbs.twimg.com/media/CWSL4W8WsAAE4KU.jpg">https://pbs.twimg.com/media/CWSL4W8WsAAE4KU.jpg</a>
531	676821958043033607	<a href="https://pbs.twimg.com/media/CWSN-vaXAAA8Ehr.jpg">https://pbs.twimg.com/media/CWSN-vaXAAA8Ehr.jpg</a>
545	677314812125323265	<a href="https://pbs.twimg.com/media/CWZ00IUW4AAQrX_.jpg">https://pbs.twimg.com/media/CWZ00IUW4AAQrX_.jpg</a>
548	677334615166730240	<a href="https://pbs.twimg.com/media/CWZgPPUWUAAU0vu.jpg">https://pbs.twimg.com/media/CWZgPPUWUAAU0vu.jpg</a>
553	677573743309385728	<a href="https://pbs.twimg.com/media/CWc5uVPXIAErLYr.jpg">https://pbs.twimg.com/media/CWc5uVPXIAErLYr.jpg</a>
570	678396796259975168	<a href="https://pbs.twimg.com/media/CWomSU_XIAAUyIK.jpg">https://pbs.twimg.com/media/CWomSU_XIAAUyIK.jpg</a>
587	679062614270468097	<a href="https://pbs.twimg.com/media/CWyD2HGUYAQIxa7.jpg">https://pbs.twimg.com/media/CWyD2HGUYAQIxa7.jpg</a>
602	679828447187857408	<a href="https://pbs.twimg.com/media/CW88XN4WsAAlo8r.jpg">https://pbs.twimg.com/media/CW88XN4WsAAlo8r.jpg</a>
609	680085611152338944	<a href="https://pbs.twimg.com/media/CXAiiHUVkAIN_28.jpg">https://pbs.twimg.com/media/CXAiiHUVkAIN_28.jpg</a>
627	680836378243002368	<a href="https://pbs.twimg.com/media/CXLREjOW8AEIfk6.jpg">https://pbs.twimg.com/media/CXLREjOW8AEIfk6.jpg</a>
632	680959110691590145	<a href="https://pbs.twimg.com/media/CXNAsm6WsAAEST9R.jpg">https://pbs.twimg.com/media/CXNAsm6WsAAEST9R.jpg</a>
654	682059653698686977	<a href="https://pbs.twimg.com/media/CXcpovWMAAMcfv.jpg">https://pbs.twimg.com/media/CXcpovWMAAMcfv.jpg</a>
693	684225744407494656	<a href="https://pbs.twimg.com/media/CX7br3HWsAAQ9L1.jpg">https://pbs.twimg.com/media/CX7br3HWsAAQ9L1.jpg</a>
705	684940049151070208	<a href="https://pbs.twimg.com/media/CYF1VUFWwAAEsWX.jpg">https://pbs.twimg.com/media/CYF1VUFWwAAEsWX.jpg</a>
...	...	...
1940	860924035999428608	<a href="https://pbs.twimg.com/media/C_KVJjDXsAEUCWn.jpg">https://pbs.twimg.com/media/C_KVJjDXsAEUCWn.jpg</a>
1944	861769973181624320	<a href="https://pbs.twimg.com/media/CzG425nWgAAanP7P.jpg">https://pbs.twimg.com/media/CzG425nWgAAanP7P.jpg</a>
1945	862096992088072192	<a href="https://pbs.twimg.com/media/C_bIo7QXYAAGfPu.jpg">https://pbs.twimg.com/media/C_bIo7QXYAAGfPu.jpg</a>
1948	862831371563274240	<a href="https://pbs.twimg.com/media/C_lkieeVwAAmOL4.jpg">https://pbs.twimg.com/media/C_lkieeVwAAmOL4.jpg</a>
1949	863062471531167744	<a href="https://pbs.twimg.com/media/C_o2vKCUwAAgtOp.jpg">https://pbs.twimg.com/media/C_o2vKCUwAAgtOp.jpg</a>
1954	864197398364647424	<a href="https://pbs.twimg.com/media/C_4-8iPVOAA1Twg.jpg">https://pbs.twimg.com/media/C_4-8iPVOAA1Twg.jpg</a>
1956	864873206498414592	<a href="https://pbs.twimg.com/media/DAClmHkXcAA1kSv.jpg">https://pbs.twimg.com/media/DAClmHkXcAA1kSv.jpg</a>
1958	865359393868664832	<a href="https://pbs.twimg.com/media/DAJfxqGV0AAAnvQt.jpg">https://pbs.twimg.com/media/DAJfxqGV0AAAnvQt.jpg</a>
1961	866450705531457537	<a href="https://pbs.twimg.com/media/DAZAUfBXcAAG_Nn.jpg">https://pbs.twimg.com/media/DAZAUfBXcAAG_Nn.jpg</a>
1966	867774946302451713	<a href="https://pbs.twimg.com/media/DAR0tDZXUAEmvdu.jpg">https://pbs.twimg.com/media/DAR0tDZXUAEmvdu.jpg</a>
1976	870308999962521604	<a href="https://pbs.twimg.com/media/DBP1asiUAAEKZI5.jpg">https://pbs.twimg.com/media/DBP1asiUAAEKZI5.jpg</a>
1978	870656317836468226	<a href="https://pbs.twimg.com/media/DBUxSSTXsAA-Jn1.jpg">https://pbs.twimg.com/media/DBUxSSTXsAA-Jn1.jpg</a>
1981	871515927908634625	<a href="https://pbs.twimg.com/media/DBG_HT9WAAEeIMM.jpg">https://pbs.twimg.com/media/DBG_HT9WAAEeIMM.jpg</a>
1982	871762521631449091	<a href="https://pbs.twimg.com/media/DBkfY58XcAEdzZy.jpg">https://pbs.twimg.com/media/DBkfY58XcAEdzZy.jpg</a>
1985	872261713294495745	<a href="https://pbs.twimg.com/media/DBrlZk2UQAafAkd.jpg">https://pbs.twimg.com/media/DBrlZk2UQAafAkd.jpg</a>
1988	872820683541237760	<a href="https://pbs.twimg.com/media/DBzhxOPWAAEh1OE.jpg">https://pbs.twimg.com/media/DBzhxOPWAAEh1OE.jpg</a>
1989	872967104147763200	<a href="https://pbs.twimg.com/media/DB1m871XUAAw5vZ.jpg">https://pbs.twimg.com/media/DB1m871XUAAw5vZ.jpg</a>
1993	874012996292530176	<a href="https://pbs.twimg.com/media/DCEeLxjXsAAvNSM.jpg">https://pbs.twimg.com/media/DCEeLxjXsAAvNSM.jpg</a>

1997	875021211251597312	<a href="https://pbs.twimg.com/media/DCSzF3NV AAPzT4.jpg">https://pbs.twimg.com/media/DCSzF3NV AAPzT4.jpg</a>
2007	877736472329191424	<a href="https://pbs.twimg.com/media/DC5YqoQWOAArOLH.jpg">https://pbs.twimg.com/media/DC5YqoQWOAArOLH.jpg</a>
2010	878776093423087618	<a href="https://pbs.twimg.com/media/DDIKMXzWOAEibje.jpg">https://pbs.twimg.com/media/DDIKMXzWOAEibje.jpg</a>
2016	879862464715927552	<a href="https://pbs.twimg.com/media/DDXmPrbWAAEKMy.jpg">https://pbs.twimg.com/media/DDXmPrbWAAEKMy.jpg</a>
2031	883117836046086144	<a href="https://pbs.twimg.com/media/DEF2-_hXoAA62q.jpg">https://pbs.twimg.com/media/DEF2-_hXoAA62q.jpg</a>
2040	885167619883638784	<a href="https://pbs.twimg.com/media/DEi_N9qXYAAgEEw.jpg">https://pbs.twimg.com/media/DEi_N9qXYAAgEEw.jpg</a>
2048	886983233522544640	<a href="https://pbs.twimg.com/media/DE8yicJWOAAAvBJ.jpg">https://pbs.twimg.com/media/DE8yicJWOAAAvBJ.jpg</a>
2051	887473957103951883	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2055	888202515573088257	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2056	888554962724278272	<a href="https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg</a>
2067	890729181411237888	<a href="https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg">https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg</a>
2070	891327558926688256	<a href="https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg</a>

	img_num	p1	p1_conf	p1_dog	\
144	4	Chihuahua	0.708163	True	
312	2	Loafer	0.255088	False	
315	2	stone_wall	0.271121	False	
319	2	Chihuahua	0.500373	True	
345	2	pug	0.914685	True	
381	2	miniature_pinscher	0.384099	True	
382	3	Samoyed	0.978833	True	
410	2	Brabancon_griffon	0.216767	True	
441	2	chow	0.526230	True	
452	2	vizsla	0.665516	True	
454	2	Samoyed	0.634695	True	
478	2	Pembroke	0.773135	True	
479	3	borzoi	0.866367	True	
494	2	standard_schnauzer	0.441427	True	
511	2	Chihuahua	0.376741	True	
517	2	Maltese_dog	0.579472	True	
530	2	rain_barrel	0.625555	False	
531	2	Great_Pyrenees	0.869804	True	
545	2	Blenheim_spaniel	0.924127	True	
548	2	Lakeland_terrier	0.859392	True	
553	2	patio	0.535070	False	
570	2	Pembroke	0.956180	True	
587	2	cash_machine	0.802333	False	
602	3	Chihuahua	0.346545	True	
609	3	pillow	0.778113	False	
627	3	Pembroke	0.427781	True	
632	2	carousel	0.500992	False	
654	2	jigsaw_puzzle	0.995873	False	
693	2	golden_retriever	0.203249	True	
705	2	Border_collie	0.665578	True	
...	...	...	...	...	
1940	2	envelope	0.933016	False	
1944	2	Arabian_camel	0.366248	False	
1945	2	chow	0.677589	True	

1948	2	Australian_terrier	0.207281	True
1949	2	French_bulldog	0.935804	True
1954	4	golden_retriever	0.945905	True
1956	2	pole	0.478616	False
1958	2	Chesapeake_Bay_retriever	0.832435	True
1961	2	French_bulldog	0.905334	True
1966	2	Border_collie	0.661953	True
1976	2	Greater_Swiss_Mountain_dog	0.622752	True
1978	4	Pembroke	0.945495	True
1981	2	komondor	0.974781	True
1982	2	Labrador_retriever	0.921393	True
1985	2	Labrador_retriever	0.972019	True
1988	3	pug	0.999120	True
1989	2	Labrador_retriever	0.476913	True
1993	2	Cardigan	0.806674	True
1997	2	West_Highland_white_terrier	0.714319	True
2007	2	Chesapeake_Bay_retriever	0.837956	True
2010	2	Italian_greyhound	0.734684	True
2016	3	basset	0.813507	True
2031	2	golden_retriever	0.949562	True
2040	4	malamute	0.812482	True
2048	2	Chihuahua	0.793469	True
2051	2	Pembroke	0.809197	True
2055	2	Pembroke	0.809197	True
2056	3	Siberian_husky	0.700377	True
2067	2	Pomeranian	0.566142	True
2070	2	basset	0.555712	True

	p2	p2_conf	p2_dog	p3 \
144	Pomeranian	0.091372	True	titi
312	platypus	0.090019	False	cowboy_boot
315	Irish_wolfhound	0.063078	True	poncho
319	French_bulldog	0.112796	True	Italian_greyhound
345	Norwegian_elkhound	0.014982	True	Siamese_cat
381	bloodhound	0.079923	True	Rottweiler
382	Pomeranian	0.012763	True	Eskimo_dog
410	Chihuahua	0.190958	True	golden_retriever
441	Pomeranian	0.283647	True	toy_poodle
452	redbone	0.173366	True	basset
454	Arctic_fox	0.309853	False	kuvasz
478	Cardigan	0.116810	True	chow
479	Saluki	0.122079	True	Irish_wolfhound
494	miniature_schnauzer	0.248885	True	Sealyham_terrier
511	Italian_greyhound	0.173114	True	muzzle
517	toy_poodle	0.133446	True	Shih-Tzu
530	barrel	0.138383	False	Labrador_retriever
531	kuvasz	0.079814	True	standard_poodle
545	Japanese_spaniel	0.054790	True	Chihuahua

548	Airedale	0.067292	True	Irish_water_spaniel
553	folding_chair	0.080419	False	parallel_bars
570	Cardigan	0.031803	True	Chihuahua
587	schipperke	0.045519	True	German_shepherd
602	dalmatian	0.166246	True	toy_terrier
609	apron	0.095023	False	wallet
627	Shetland_sheepdog	0.160669	True	Pomeranian
632	feather_boa	0.064390	False	pug
654	Siamese_cat	0.000781	False	pizza
693	Samoyed	0.067958	True	Great_Pyrenees
705	collie	0.176846	True	Old_English_sheepdog
...	...	...	...	...
1940	oscilloscope	0.012591	False	paper_towel
1944	house_finch	0.209852	False	cocker_spaniel
1945	Pomeranian	0.270648	True	Pekinese
1948	Irish_terrier	0.156296	True	German_shepherd
1949	pug	0.059576	True	boxer
1954	Labrador_retriever	0.021264	True	Tibetan_mastiff
1956	lakeside	0.114182	False	wreck
1958	Labrador_retriever	0.163551	True	Weimaraner
1961	Boston_bull	0.078060	True	pug
1966	Cardigan	0.175718	True	collie
1976	Appenzeller	0.158463	True	EntleBucher
1978	Cardigan	0.045875	True	beagle
1981	briard	0.020041	True	swab
1982	golden_retriever	0.064608	True	bloodhound
1985	flat-coated_retriever	0.008178	True	Chesapeake_Bay_retriever
1988	French_bulldog	0.000552	True	bull_mastiff
1989	Chesapeake_Bay_retriever	0.174145	True	German_short-haired_pointer
1993	Pembroke	0.116622	True	kelpie
1997	Siberian_husky	0.091913	True	Great_Pyrenees
2007	Labrador_retriever	0.062034	True	Weimaraner
2010	whippet	0.150487	True	Ibizan_hound
2016	beagle	0.146654	True	cocker_spaniel
2031	Labrador_retriever	0.045948	True	kuvasz
2040	Siberian_husky	0.071712	True	Eskimo_dog
2048	toy_terrier	0.143528	True	can_opener
2051	Rhodesian_ridgeback	0.054950	True	beagle
2055	Rhodesian_ridgeback	0.054950	True	beagle
2056	Eskimo_dog	0.166511	True	malamute
2067	Eskimo_dog	0.178406	True	Pembroke
2070	English_springer	0.225770	True	German_short-haired_pointer

	p3_conf	p3_dog
144	0.067325	False
312	0.066536	False
315	0.048226	False
319	0.062893	True

345	0.009221	False
381	0.068594	True
382	0.001853	True
410	0.163288	True
441	0.067665	True
452	0.134783	True
454	0.019641	True
478	0.039036	True
479	0.004020	True
494	0.164967	True
511	0.071485	False
517	0.094397	True
530	0.039465	True
531	0.013263	True
545	0.008204	True
548	0.049531	True
553	0.034796	False
570	0.006276	True
587	0.023353	True
602	0.117502	True
609	0.049326	False
627	0.111250	True
632	0.044357	True
654	0.000432	False
693	0.065327	True
705	0.065175	True
...	...	...
1940	0.011178	False
1944	0.046403	True
1945	0.038110	True
1948	0.123536	True
1949	0.001412	True
1954	0.020493	True
1956	0.055926	False
1958	0.002770	True
1961	0.001771	True
1966	0.087142	True
1976	0.148115	True
1978	0.004329	True
1981	0.003228	False
1982	0.003383	True
1985	0.007359	True
1988	0.000073	True
1989	0.092861	True
1993	0.049182	True
1997	0.046038	True
2007	0.040599	True
2010	0.039725	True

2016	0.009485	True
2031	0.002471	True
2040	0.055770	True
2048	0.032253	False
2051	0.038915	True
2055	0.038915	True
2056	0.111411	True
2067	0.076507	True
2070	0.175219	True

[295 rows x 12 columns]

## 1.2 Cleaning Data

**Note1:** concerning the missing values in tweet\_count and image tables, I chose the minimum number (2075, same as the total number in imagetable) since we target to the original ratings that have pictures. **Note2:** the tweet (id:835246439529840640) was a retweeted tweet and was cleaned in the previous section.

**Define**

**archive table**

- 1.Drop the retweeted rows and the "in\_reply\_to\_status\_id", "in\_reply\_to\_user\_id", "retweeted\_status\_id","retweeted\_status\_user\_id", "retweeted\_status\_timestamp" columns
- 2.Add a new column "rating" with the calculated rating (formula: rating\_numerator/rating\_denominator)
- 3.Move the short link into a new created column "short\_url" and drop the long and redudant "expanded\_urls" column (which also has some missing values - 2094 instead of 2097)
- 4.Delete the short link and keep the original text as that in tweet in the "text" column
- 5.Only keep the essential information in the middle in the "source" column(iPhone, Vine, Web, and TweetDeck)
- 6.Convert the datatype of "timestamp" column to datetime
- 7.Create a new column "dog\_stage" containing the information in the "doggo","floofer","pupper", and "puppo" columns and drop the intermediate columns. (Note: some tweets contain two dog stages. Some of them mention two ages for one dog; others have two dogs in one image. Use "/" to label two dog stage)
- 8.Arrange the order of columns

### Others

- Only keep "tweet\_id" and "jpg.url" columns in the image table (as other information will be used for simulating the confusion matrix)
- Merge thearchive, tweet\_count, and image tables into one final table

```
In [28]: # Make a copy of the archive table
a_clean = archive.copy()
```

1. Drop the retweeted rows and the "in\_reply\_to\_status\_id", "in\_reply\_to\_user\_id", "retweeted\_status\_id","retweeted\_status\_user\_id", "retweeted\_status\_timestamp" columns

```
In [29]: # Drop the retweeted rows and the "in_reply_to_status_id", "in_reply_to_user_id", "retweeted_status_id", "retweeted_status_timestamp" columns
a_clean = a_clean.drop(['in_reply_to_status_id', 'in_reply_to_user_id', 'retweeted_status_id', 'retweeted_status_timestamp'])

# Test
a_clean.head(1)
```

```
Out[29]:
```

	tweet_id	timestamp	source	text	expanded_urls	rating_numerator	rating_denominator	name	doggo	floofer	pupper	puppo
0	892420643555336193	2017-08-01 16:23:56 +0000	<a href="http://twitter.com/download/iphone" r...	This is Phineas. He's a mystical boy. Only eve...	https://twitter.com/dog_rates/status/892420643...	13	10	Phineas	None	None	None	None

1. Add a new column "rating" with the calculated rating (formula: rating\_numerator/rating\_denominator)

```
In [30]: #Add a new column "rating" with the calculated rating (formula: rating_numerator/rating_denominator)
a_clean['rating'] = a_clean['rating_numerator']/a_clean['rating_denominator']

# Test
a_clean.head(1)
```

```
Out[30]:
```

	tweet_id	timestamp	source	text	expanded_urls	rating_numerator	rating_denominator	name	doggo	floofer	pupper	puppo	rating
0	892420643555336193	2017-08-01 16:23:56 +0000	<a href="http://twitter.com/download/iphone" r...	This is Phineas. He's a mystical boy. Only eve...	https://twitter.com/dog_rates/status/892420643...	13	10	Phineas	None	None	None	None	1.3

1. Move the short link into a new created column "short\_url" and drop the long and redundant "expanded\_urls" column (which also has some missing values - 2094 instead of 2097)
2. Delete the short link and keep the original text as that in tweet in the "text" column



```

In [31]: # Display the original "text" column
a_clean.text[0]

Out[31]: "This is Phineas. He's a mystical boy. Only ever appears in the hole of a donut. 13/10"

In [32]: # Count the length of the short link as it keeps the same length in the dataset
len(" https://t.co/MgUWQ76dJU")

Out[32]: 24

In [33]: # Move the short link into a new created column and keep the original text as that in t
a_clean['short_url'] = a_clean['text'].str[-23:]
a_clean['text'] = a_clean['text'].str[:-24]

In [34]: # Test
a_clean.text[0]

Out[34]: "This is Phineas. He's a mystical boy. Only ever appears in the hole of a donut. 13/10"

In [35]: # Test
a_clean.short_url[0]

Out[35]: 'https://t.co/MgUWQ76dJU'

In [36]: # Drop the long and redundant "expanded_urls" column
a_clean = a_clean.drop(['expanded_urls'], axis=1)

In [37]: # Test
a_clean.head(3)

Out[37]:
      tweet_id      timestamp \
0  892420643555336193  2017-08-01 16:23:56 +0000
1  892177421306343426  2017-08-01 00:17:27 +0000
2  891815181378084864  2017-07-31 00:18:03 +0000

      source \
0  <a href="http://twitter.com/download/iphone" r...
1  <a href="http://twitter.com/download/iphone" r...
2  <a href="http://twitter.com/download/iphone" r...

      text  rating_numerator \
0  This is Phineas. He's a mystical boy. Only eve...      13
1  This is Tilly. She's just checking pup on you...      13
2  This is Archie. He is a rare Norwegian Pouncin...      12

      rating_denominator  name  doggo  floofer  pupper  puppo  rating \
0              10  Phineas  None      None      None      None      1.3
1              10    Tilly  None      None      None      None      1.3
2              10   Archie  None      None      None      None      1.2

```

```

short_url
0 https://t.co/MgUWQ76dJU
1 https://t.co/OXxu71qeIV
2 https://t.co/wUnZnhtVJB

```

1. Only keep the essential information in the middle in the "source" column(iPhone, Vine, Web, and TweetDeck)

```
In [38]: a_clean.source.value_counts()
```

```

Out[38]: <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>
<a href="http://vine.co" rel="nofollow">Vine - Make a Scene</a>
<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
<a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">TweetDeck</a>
Name: source, dtype: int64

```

```

In [39]: # Only keep the essential information in the middle in the "source" column(iPhone, Vine, Web, and TweetDeck)
a_clean['source'] = a_clean['source'].replace({'<a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>': 'iPhone',
                                                '<a href="http://vine.co" rel="nofollow">Vine - Make a Scene</a>': 'Vine',
                                                '<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>': 'Web',
                                                '<a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">TweetDeck</a>': 'TweetDeck'})

# Test
a_clean.source.value_counts()

```

```

Out[39]: iPhone      1964
Vine                91
Web                 31
TweetDeck          11
Name: source, dtype: int64

```

1. Convert the datatype of "timestamp" column to datetime

```

In [40]: # Convert the datatype of "timestamp" column to datetime.
a_clean.timestamp = pd.to_datetime(a_clean.timestamp)

# Test
a_clean.info()

```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 13 columns):
tweet_id      2097 non-null int64
timestamp     2097 non-null datetime64[ns]
source        2097 non-null object
text          2097 non-null object
rating_numerator  2097 non-null int64
rating_denominator  2097 non-null int64
name          2097 non-null object
doggo         2097 non-null object

```

```

floofer          2097 non-null object
pupper           2097 non-null object
puppo            2097 non-null object
rating           2097 non-null float64
short_url        2097 non-null object
dtypes: datetime64[ns](1), float64(1), int64(3), object(8)
memory usage: 309.4+ KB

```

1. Create a new column "dog\_stage" containing the information in the "doggo","floofer","pupper", and "puppo" columns and drop the intermediate columns. (Note: some tweets contain two dog stages. Some of them mention two ages for one dog; others have two dogs in one image.)

```

In [41]: def find_dog_stage(a_clean):
        dog_stage = []
        for i, row in a_clean.iterrows():
            if row['doggo'] == "doggo" and row['floofer'] == "floofer":
                dog_stage.append('doggo/floofer')
            elif row['doggo'] == "doggo" and row['puppo'] == "puppo":
                dog_stage.append('doggo/puppo')
            elif row['doggo'] == "doggo" and row['pupper'] == "pupper":
                dog_stage.append('doggo/pupper')
            elif row['doggo'] == "doggo":
                dog_stage.append('doggo')
            elif row['floofer'] == "floofer":
                dog_stage.append('floofer')
            elif row['puppo'] == "puppo":
                dog_stage.append('puppo')
            elif row['pupper'] == "pupper":
                dog_stage.append('pupper')
            else:
                dog_stage.append('None')
        return dog_stage

```

```

a_clean['dog_stage'] = find_dog_stage(a_clean)
a_clean.dog_stage.value_counts()

```

```

Out[41]: None          1761
        pupper         221
        doggo          72
        puppo          23
        floofer         9
        doggo/pupper    9
        doggo/puppo     1
        doggo/floofer   1
        Name: dog_stage, dtype: int64

```

```

In [42]: # Test
        doggo_floofer = a_clean.query('doggo == "doggo" and floofer == "floofer"').shape[0]

```

```

doggo_puppo = a_clean.query('doggo == "doggo" and puppo == "puppo"]').shape[0]
doggo_pupper = a_clean.query('doggo == "doggo" and pupper == "pupper"]').shape[0]
print('pupper:', a_clean.query('pupper == "pupper"]').shape[0] - doggo_pupper)
print('doggo:', a_clean.query('doggo == "doggo"]').shape[0] - doggo_floofer - doggo_puppo)
print('puppo:', a_clean.query('puppo == "puppo"]').shape[0] - doggo_puppo)
print('floofer:', a_clean.query('floofer == "floofer"]').shape[0] - doggo_floofer)
print('doggo/pupper:', doggo_pupper)
print('doggo/floofer:', doggo_floofer)
print('doggo/puppo:', doggo_puppo)

```

```

pupper: 221
doggo: 72
puppo: 23
floofer: 9
doggo/pupper: 9
doggo/floofer: 1
doggo/puppo: 1

```

```

In [43]: # Drop the intermediate columns
a_clean.drop(['doggo', 'floofer', 'pupper', 'puppo'], axis=1, inplace=True)

# Test
a_clean.head(1)

```

```

Out[43]:
      tweet_id      timestamp  source \
0  892420643555336193  2017-08-01 16:23:56  iPhone

      text  rating_numerator \
0  This is Phineas. He's a mystical boy. Only eve...      13

      rating_denominator  name  rating      short_url  dog_stage
0              10  Phineas      1.3  https://t.co/MgUWQ76dJU      None

```

```

In [44]: # Arrange the order of columns
a_clean = a_clean[['tweet_id', 'timestamp', 'source', 'text', 'name', 'dog_stage', 'rating_nu

# Test
a_clean.head(1)

```

```

Out[44]:
      tweet_id      timestamp  source \
0  892420643555336193  2017-08-01 16:23:56  iPhone

      text      name  dog_stage \
0  This is Phineas. He's a mystical boy. Only eve...  Phineas      None

      rating_numerator  rating_denominator  rating      short_url
0              13              10      1.3  https://t.co/MgUWQ76dJU

```

```
In [45]: # Save the cleaned dataset into "a_clean.csv"
a_clean.to_csv('a_clean.csv', index=False)
```

```
In [46]: # Make a copy of "image" table
i_clean = image.copy()
i_clean
```

```
Out[46]:
```

	tweet_id	jpg_url \
0	666020888022790149	<a href="https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg</a>
1	666029285002620928	<a href="https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg">https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg</a>
2	666033412701032449	<a href="https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg">https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg</a>
3	666044226329800704	<a href="https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg">https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg</a>
4	666049248165822465	<a href="https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg">https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg</a>
5	666050758794694657	<a href="https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg">https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg</a>
6	666051853826850816	<a href="https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg">https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg</a>
7	666055525042405380	<a href="https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg">https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg</a>
8	666057090499244032	<a href="https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg">https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg</a>
9	666058600524156928	<a href="https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg">https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg</a>
10	666063827256086533	<a href="https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg">https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg</a>
11	666071193221509120	<a href="https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg">https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg</a>
12	666073100786774016	<a href="https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg">https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg</a>
13	666082916733198337	<a href="https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg">https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg</a>
14	666094000022159362	<a href="https://pbs.twimg.com/media/CT5w9gUW4AAsBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAsBNN.jpg</a>
15	666099513787052032	<a href="https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg">https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg</a>
16	666102155909144576	<a href="https://pbs.twimg.com/media/CT54YGiwUAENZnoK.jpg">https://pbs.twimg.com/media/CT54YGiwUAENZnoK.jpg</a>
17	666104133288665088	<a href="https://pbs.twimg.com/media/CT56LSZWAA1Jj2.jpg">https://pbs.twimg.com/media/CT56LSZWAA1Jj2.jpg</a>
18	666268910803644416	<a href="https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg">https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg</a>
19	666273097616637952	<a href="https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg">https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg</a>
20	666287406224695296	<a href="https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg">https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg</a>
21	666293911632134144	<a href="https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg">https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg</a>
22	666337882303524864	<a href="https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg">https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg</a>
23	666345417576210432	<a href="https://pbs.twimg.com/media/CT9Vn7PWAAA_ZCM.jpg">https://pbs.twimg.com/media/CT9Vn7PWAAA_ZCM.jpg</a>
24	666353288456101888	<a href="https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg">https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg</a>
25	666362758909284353	<a href="https://pbs.twimg.com/media/CT9lXGsUcAAyUFt.jpg">https://pbs.twimg.com/media/CT9lXGsUcAAyUFt.jpg</a>
26	666373753744588802	<a href="https://pbs.twimg.com/media/CT9vZEYUAA1Z05.jpg">https://pbs.twimg.com/media/CT9vZEYUAA1Z05.jpg</a>
27	666396247373291520	<a href="https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg">https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg</a>
28	666407126856765440	<a href="https://pbs.twimg.com/media/CT-NvwmW4AAAugGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAAugGZ.jpg</a>
29	666411507551481857	<a href="https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg">https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg</a>
...	...	...
2045	886366144734445568	<a href="https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg">https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg</a>
2046	886680336477933568	<a href="https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg">https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg</a>
2047	886736880519319552	<a href="https://pbs.twimg.com/media/DE5Se8FXcAAJF4.jpg">https://pbs.twimg.com/media/DE5Se8FXcAAJF4.jpg</a>
2048	886983233522544640	<a href="https://pbs.twimg.com/media/DE8yicJWOAAAABJ.jpg">https://pbs.twimg.com/media/DE8yicJWOAAAABJ.jpg</a>
2049	887101392804085760	<a href="https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg">https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg</a>
2050	887343217045368832	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88734...">https://pbs.twimg.com/ext_tw_video_thumb/88734...</a>
2051	887473957103951883	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2052	887517139158093824	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88751...">https://pbs.twimg.com/ext_tw_video_thumb/88751...</a>
2053	887705289381826560	<a href="https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg">https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg</a>

2054	888078434458587136	<a href="https://pbs.twimg.com/media/DFMWn56WsAAkA7B.jpg">https://pbs.twimg.com/media/DFMWn56WsAAkA7B.jpg</a>
2055	888202515573088257	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2056	888554962724278272	<a href="https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg</a>
2057	888804989199671297	<a href="https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg">https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg</a>
2058	888917238123831296	<a href="https://pbs.twimg.com/media/DFYRgsOUQAARgh0.jpg">https://pbs.twimg.com/media/DFYRgsOUQAARgh0.jpg</a>
2059	889278841981685760	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88927...">https://pbs.twimg.com/ext_tw_video_thumb/88927...</a>
2060	889531135344209921	<a href="https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg">https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg</a>
2061	889638837579907072	<a href="https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg">https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg</a>
2062	889665388333682689	<a href="https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg">https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg</a>
2063	889880896479866881	<a href="https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg">https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg</a>
2064	890006608113172480	<a href="https://pbs.twimg.com/media/DFnwsY4WAAAMliS.jpg">https://pbs.twimg.com/media/DFnwsY4WAAAMliS.jpg</a>
2065	890240255349198849	<a href="https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg">https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg</a>
2066	890609185150312448	<a href="https://pbs.twimg.com/media/DFwUU__XcAEpyXI.jpg">https://pbs.twimg.com/media/DFwUU__XcAEpyXI.jpg</a>
2067	890729181411237888	<a href="https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg">https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg</a>
2068	890971913173991426	<a href="https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg">https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg</a>
2069	891087950875897856	<a href="https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg">https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg</a>
2070	891327558926688256	<a href="https://pbs.twimg.com/media/DF6hr6BUMAAZzgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAZzgT.jpg</a>
2071	891689557279858688	<a href="https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg">https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg</a>
2072	891815181378084864	<a href="https://pbs.twimg.com/media/DGBdLU1WsAAANxJ9.jpg">https://pbs.twimg.com/media/DGBdLU1WsAAANxJ9.jpg</a>
2073	892177421306343426	<a href="https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg">https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg</a>
2074	892420643555336193	<a href="https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg">https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg</a>

	img_num		p1	p1_conf	p1_dog	\
0	1	Welsh_springer_spaniel	0.465074		True	
1	1	redbone	0.506826		True	
2	1	German_shepherd	0.596461		True	
3	1	Rhodesian_ridgeback	0.408143		True	
4	1	miniature_pinscher	0.560311		True	
5	1	Bernese_mountain_dog	0.651137		True	
6	1	box_turtle	0.933012		False	
7	1	chow	0.692517		True	
8	1	shopping_cart	0.962465		False	
9	1	miniature_poodle	0.201493		True	
10	1	golden_retriever	0.775930		True	
11	1	Gordon_setter	0.503672		True	
12	1	Walker_hound	0.260857		True	
13	1	pug	0.489814		True	
14	1	bloodhound	0.195217		True	
15	1	Lhasa	0.582330		True	
16	1	English_setter	0.298617		True	
17	1	hen	0.965932		False	
18	1	desktop_computer	0.086502		False	
19	1	Italian_greyhound	0.176053		True	
20	1	Maltese_dog	0.857531		True	
21	1	three-toed_sloth	0.914671		False	
22	1	ox	0.416669		False	
23	1	golden_retriever	0.858744		True	
24	1	malamute	0.336874		True	

25	1	guinea_pig	0.996496	False
26	1	soft-coated_wheaten_terrier	0.326467	True
27	1	Chihuahua	0.978108	True
28	1	black-and-tan_coonhound	0.529139	True
29	1	coho	0.404640	False
...	...	...	...	...
2045	1	French_bulldog	0.999201	True
2046	1	convertible	0.738995	False
2047	1	kuvasz	0.309706	True
2048	2	Chihuahua	0.793469	True
2049	1	Samoyed	0.733942	True
2050	1	Mexican_hairless	0.330741	True
2051	2	Pembroke	0.809197	True
2052	1	limousine	0.130432	False
2053	1	basset	0.821664	True
2054	1	French_bulldog	0.995026	True
2055	2	Pembroke	0.809197	True
2056	3	Siberian_husky	0.700377	True
2057	1	golden_retriever	0.469760	True
2058	1	golden_retriever	0.714719	True
2059	1	whippet	0.626152	True
2060	1	golden_retriever	0.953442	True
2061	1	French_bulldog	0.991650	True
2062	1	Pembroke	0.966327	True
2063	1	French_bulldog	0.377417	True
2064	1	Samoyed	0.957979	True
2065	1	Pembroke	0.511319	True
2066	1	Irish_terrier	0.487574	True
2067	2	Pomeranian	0.566142	True
2068	1	Appenzeller	0.341703	True
2069	1	Chesapeake_Bay_retriever	0.425595	True
2070	2	basset	0.555712	True
2071	1	paper_towel	0.170278	False
2072	1	Chihuahua	0.716012	True
2073	1	Chihuahua	0.323581	True
2074	1	orange	0.097049	False

	p2	p2_conf	p2_dog	p3 \
0	collie	0.156665	True	Shetland_sheepdog
1	miniature_pinscher	0.074192	True	Rhodesian_ridgeback
2	malinois	0.138584	True	bloodhound
3	redbone	0.360687	True	miniature_pinscher
4	Rottweiler	0.243682	True	Doberman
5	English_springer	0.263788	True	Greater_Swiss_Mountain_dog
6	mud_turtle	0.045885	False	terrapin
7	Tibetan_mastiff	0.058279	True	fur_coat
8	shopping_basket	0.014594	False	golden_retriever
9	komondor	0.192305	True	soft-coated_wheaten_terrier

10	Tibetan_mastiff	0.093718	True	Labrador_retriever
11	Yorkshire_terrier	0.174201	True	Pekinese
12	English_foxhound	0.175382	True	Ibizan_hound
13	bull_mastiff	0.404722	True	French_bulldog
14	German_shepherd	0.078260	True	malinois
15	Shih-Tzu	0.166192	True	Dandie_Dinmont
16	Newfoundland	0.149842	True	borzoi
17	cock	0.033919	False	partridge
18	desk	0.085547	False	bookcase
19	toy_terrier	0.111884	True	basenji
20	toy_poodle	0.063064	True	miniature_poodle
21	otter	0.015250	False	great_grey_owl
22	Newfoundland	0.278407	True	groenendael
23	Chesapeake_Bay_retriever	0.054787	True	Labrador_retriever
24	Siberian_husky	0.147655	True	Eskimo_dog
25	skunk	0.002402	False	hamster
26	Afghan_hound	0.259551	True	briard
27	toy_terrier	0.009397	True	papillon
28	bloodhound	0.244220	True	flat-coated_retriever
29	barracouta	0.271485	False	gar
...	...	...	...	...
2045	Chihuahua	0.000361	True	Boston_bull
2046	sports_car	0.139952	False	car_wheel
2047	Great_Pyrenees	0.186136	True	Dandie_Dinmont
2048	toy_terrier	0.143528	True	can_opener
2049	Eskimo_dog	0.035029	True	Staffordshire_bullterrier
2050	sea_lion	0.275645	False	Weimaraner
2051	Rhodesian_ridgeback	0.054950	True	beagle
2052	tow_truck	0.029175	False	shopping_cart
2053	redbone	0.087582	True	Weimaraner
2054	pug	0.000932	True	bull_mastiff
2055	Rhodesian_ridgeback	0.054950	True	beagle
2056	Eskimo_dog	0.166511	True	malamute
2057	Labrador_retriever	0.184172	True	English_setter
2058	Tibetan_mastiff	0.120184	True	Labrador_retriever
2059	borzoi	0.194742	True	Saluki
2060	Labrador_retriever	0.013834	True	redbone
2061	boxer	0.002129	True	Staffordshire_bullterrier
2062	Cardigan	0.027356	True	basenji
2063	Labrador_retriever	0.151317	True	muzzle
2064	Pomeranian	0.013884	True	chow
2065	Cardigan	0.451038	True	Chihuahua
2066	Irish_setter	0.193054	True	Chesapeake_Bay_retriever
2067	Eskimo_dog	0.178406	True	Pembroke
2068	Border_collie	0.199287	True	ice_lolly
2069	Irish_terrier	0.116317	True	Indian_elephant
2070	English_springer	0.225770	True	German_short-haired_pointer
2071	Labrador_retriever	0.168086	True	spatula



2072	malamute	0.078253	True	kelpie
2073	Pekinese	0.090647	True	papillon
2074	bagel	0.085851	False	banana

	p3_conf	p3_dog
0	0.061428	True
1	0.072010	True
2	0.116197	True
3	0.222752	True
4	0.154629	True
5	0.016199	True
6	0.017885	False
7	0.054449	False
8	0.007959	True
9	0.082086	True
10	0.072427	True
11	0.109454	True
12	0.097471	True
13	0.048960	True
14	0.075628	True
15	0.089688	True
16	0.133649	True
17	0.000052	False
18	0.079480	False
19	0.111152	True
20	0.025581	True
21	0.013207	False
22	0.102643	True
23	0.014241	True
24	0.093412	True
25	0.000461	False
26	0.206803	True
27	0.004577	True
28	0.173810	True
29	0.189945	False
...	...	...
2045	0.000076	True
2046	0.044173	False
2047	0.086346	True
2048	0.032253	False
2049	0.029705	True
2050	0.134203	True
2051	0.038915	True
2052	0.026321	False
2053	0.026236	True
2054	0.000903	True
2055	0.038915	True
2056	0.111411	True

```

2057  0.073482    True
2058  0.105506    True
2059  0.027351    True
2060  0.007958    True
2061  0.001498    True
2062  0.004633    True
2063  0.082981    False
2064  0.008167    True
2065  0.029248    True
2066  0.118184    True
2067  0.076507    True
2068  0.193548    False
2069  0.076902    False
2070  0.175219    True
2071  0.040836    False
2072  0.031379    True
2073  0.068957    True
2074  0.076110    False

```

```
[2075 rows x 12 columns]
```

```

In [47]: # Only keep "tweet_id" and "jpg.url" columns in the `image` table (as other information
i_clean = i_clean[['tweet_id', 'jpg_url']]

#Test
i_clean.head(3)

```

```

Out[47]:
      tweet_id      jpg_url
0  666020888022790149  https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg
1  666029285002620928  https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg
2  666033412701032449  https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg

```

```

In [48]: # Save the cleaned dataset into "a_clean.csv"
i_clean.to_csv('i_clean.csv', index=False)

```

```
In [49]: tweet_count.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1753 entries, 0 to 1752
Data columns (total 3 columns):
tweet_id      1753 non-null int64
favorite_count 1753 non-null int64
retweet_count 1753 non-null int64
dtypes: int64(3)
memory usage: 41.2 KB

```

```

In [50]: # Use inner join to merge the dataframes to ensure the data completeness
df = pd.merge(i_clean, a_clean, on=['tweet_id'], how='inner')

```

```

WeRateDog_tweet_cleaned = pd.merge(df, tweet_count, on=['tweet_id'], how='inner')
df.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1971 entries, 0 to 1970
Data columns (total 11 columns):
tweet_id          1971 non-null int64
jpg_url           1971 non-null object
timestamp         1971 non-null datetime64[ns]
source            1971 non-null object
text              1971 non-null object
name              1971 non-null object
dog_stage         1971 non-null object
rating_numerator  1971 non-null int64
rating_denominator 1971 non-null int64
rating            1971 non-null float64
short_url         1971 non-null object
dtypes: datetime64[ns](1), float64(1), int64(3), object(6)
memory usage: 184.8+ KB

In [51]: # Arrange the order of columns
WeRateDog_tweet_cleaned = WeRateDog_tweet_cleaned[['tweet_id', 'timestamp', 'source', 'text',
                                                    'name', 'dog_stage', 'rating', 'favorite_count', 'retweet_count']]
# Test
WeRateDog_tweet_cleaned.head(1)

Out[51]:
   tweet_id      timestamp  source \
0  758828659922702336  2016-07-29 00:57:05  iPhone

   text \
0  This doggo is just waiting for someone to be p...

   jpg_url      short_url \
0  https://pbs.twimg.com/media/Cofmom_VUAA4dR0.jpg  https://t.co/9T2h14yn4Q

   name dog_stage  rating  favorite_count  retweet_count
0  None    doggo    1.3          11200          3819

```

## 1.3 Storing, Analyzing, and Visualizing Data

### 1.3.1 Storing Data

```

In [52]: # Save the cleaned dataset into twitter_archive_master.csv
WeRateDog_tweet_cleaned.to_csv('twitter_archive_master.csv', index=False)

```

### 1.3.2 Analyzing Data

```

In [53]: df = pd.read_csv('twitter_archive_master.csv')
df.head()

```

```

Out [53]:          tweet_id          timestamp source \
0  758828659922702336  2016-07-29 00:57:05  iPhone
1  758854675097526272  2016-07-29 02:40:28  iPhone
2  758854675097526272  2016-07-29 02:40:28  iPhone
3  759047813560868866  2016-07-29 15:27:55  iPhone
4  759047813560868866  2016-07-29 15:27:55  iPhone

          text \
0  This doggo is just waiting for someone to be p...
1  This is Lilli Bee & Honey Bear. Unfortunat...
2  This is Lilli Bee & Honey Bear. Unfortunat...
3  This is Spencer. He's part of the Queen's Guar...
4  This is Spencer. He's part of the Queen's Guar...

          jpg_url          short_url \
0  https://pbs.twimg.com/media/Cofmom_VUAA4dR0.jpg  https://t.co/9T2h14yn4Q
1  https://pbs.twimg.com/media/Cof-SuqVYAAs4kZ.jpg  https://t.co/4Urf0ZhztW
2  https://pbs.twimg.com/media/Cof-SuqVYAAs4kZ.jpg  https://t.co/4Urf0ZhztW
3  https://pbs.twimg.com/media/Coit84_VYAEMtLi.jpg  https://t.co/8W5iS0gXfx
4  https://pbs.twimg.com/media/Coit84_VYAEMtLi.jpg  https://t.co/8W5iS0gXfx

          name dog_stage  rating  favorite_count  retweet_count
0      None      doggo      1.3           11200           3819
1     Lilli      None      1.1            3544            893
2     Lilli      None      1.1            3544            893
3   Spencer      None      1.1            6471           2004
4   Spencer      None      1.1            6471           2004

```

**The tweet that was retweeted the most (also the one that get the most "favorite")**

```
In [54]: df.favorite_count.max()
```

```
Out [54]: 132080
```

```
In [55]: df.query('favorite_count == 156364')
```

```
Out [55]: Empty DataFrame
```

```

Columns: [tweet_id, timestamp, source, text, jpg_url, short_url, name, dog_stage, rating]
Index: []

```

```
In [56]: df.text[1199]
```

```
Out [56]: "Meet Elliot. He's a Canadian Forrest Pup. Unusual number of antlers for a dog. Sneaky"
```

```
In [57]: df.jpg_url[1199]
```

```
Out [57]: 'https://pbs.twimg.com/media/DDrk-f9WAAI-WQv.jpg'
```

**The tweet scored highest**

```
In [58]: df.rating.max()
```

```
Out[58]: 7.5
```

```
In [61]: df.query('rating == 7.5')
```

```
Out[61]:
```

	tweet_id	timestamp	source	\		text	\		jpg_url	short_url	\
309	786709082849828864	2016-10-13 23:23:56	iPhone		309	This is Logan, the Chow who lived. He solemnly...			https://pbs.twimg.com/media/CurzvFTXgAA2_AP.jpg	https://t.co/yB05wuqaPS	
310	786709082849828864	2016-10-13 23:23:56	iPhone		310	This is Logan, the Chow who lived. He solemnly...			https://pbs.twimg.com/media/CurzvFTXgAA2_AP.jpg	https://t.co/yB05wuqaPS	

	name	dog_stage	rating	favorite_count	retweet_count
309	Logan	None	7.5	18430	6182
310	Logan	None	7.5	18430	6182

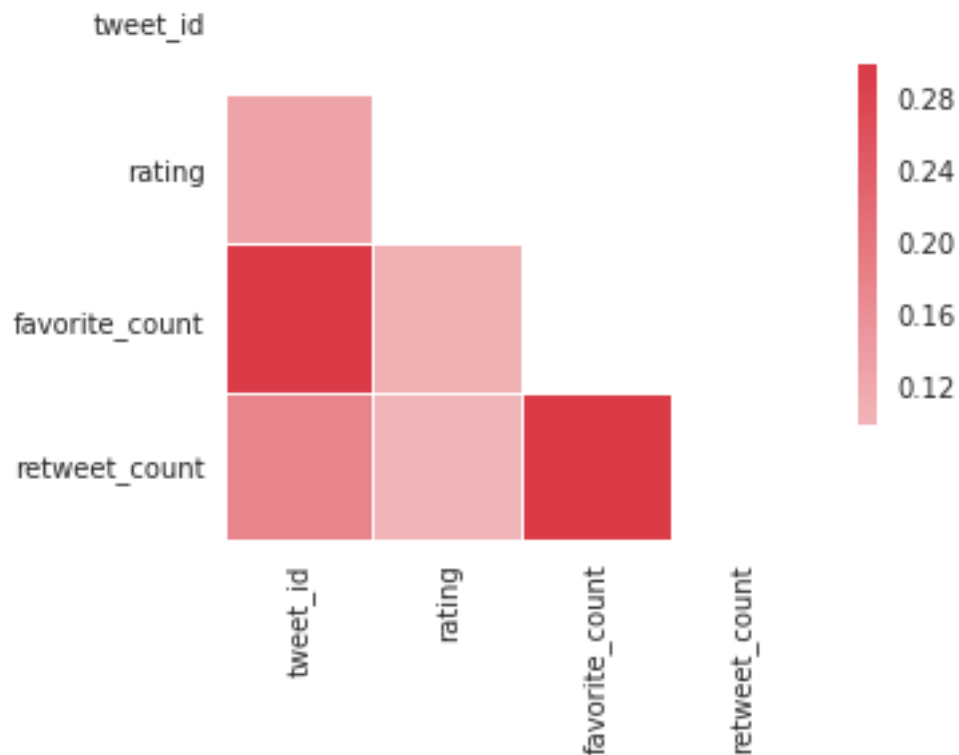
```
In [62]: df.text[1246]
```

```
Out[62]: "This is Derek. He's late for a dog meeting. 13/10 pet...al to the metal"
```

**Insight 1: The highest rating score belongs to a dog called “Derek” with an scarf in its picture. The tweet was posted on, 13 October 2016, for which I suppose the rating “13/10” refers to this memorable day.**

**The correlation heatmap**

```
In [63]: # Correlation Heatmap
sns.set(style="white")
corr = df.corr()
mask = np.triu(np.ones_like(corr, dtype=np.bool))
f, ax = plt.subplots(figsize=(5, 5))
cmap = sns.diverging_palette(220, 10, as_cmap=True)
sns.heatmap(corr, mask=mask, cmap=cmap, vmax=.3, center=0,
            square=True, linewidths=.5, cbar_kws={"shrink": .5});
```



```
In [64]: # choose the data for analyzing
df_analyze = df[['source', 'dog_stage', 'rating', 'favorite_count', 'retweet_count']]
df_analyze.head(3)
```

```
Out[64]:
```

	source	dog_stage	rating	favorite_count	retweet_count
0	iPhone	doggo	1.3	11200	3819
1	iPhone	None	1.1	3544	893
2	iPhone	None	1.1	3544	893

```
In [65]: # Group by "source"
df_source = df_analyze.groupby('source')
df_source['rating'].describe()
```

```
Out[65]:
```

	count	mean	std	min	25%	50%	75%	max
source								
Web	12.0	1.266667	0.098473	1.1	1.2	1.3	1.3	1.4
iPhone	1289.0	1.214086	0.299595	0.0	1.1	1.2	1.3	7.5

**Insight 2:** Most of tweets come from iPhone and the descriptive data and does not show any evidence about bias based on source given the fact that the sample size of data from web is too small compared to that from iPhone.

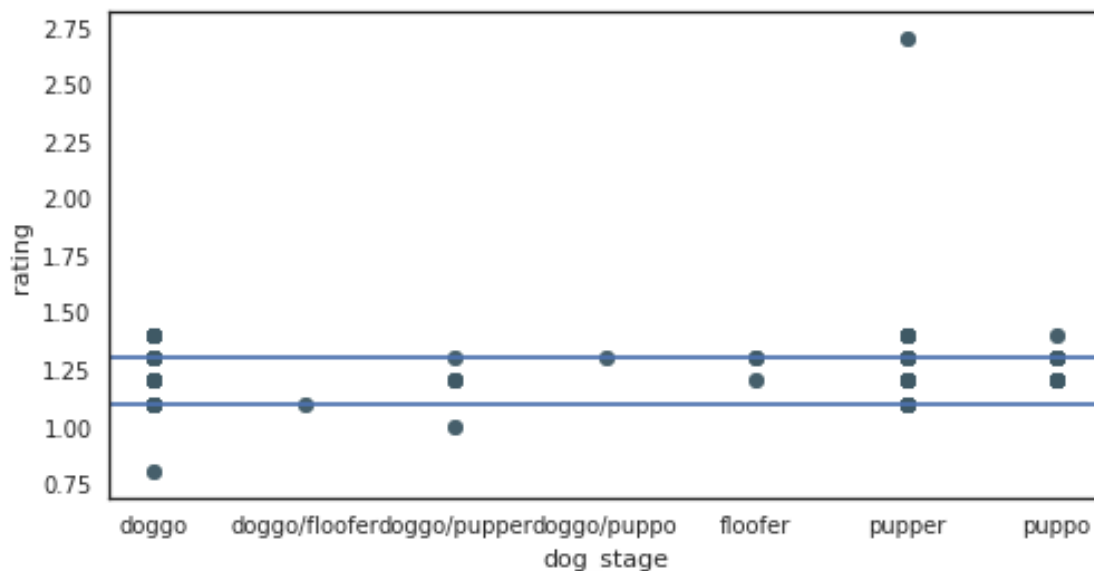
```
In [66]: # Group by "dog_stage"
df_stage = df_analyze.groupby('dog_stage')
df_stage.rating.describe()
```

```
Out[66]:
```

	count	mean	std	min	25%	50%	75%	max
dog_stage								
None	1088.0	1.209060	0.316698	0.0	1.100	1.2	1.3	7.5
doggo	87.0	1.219540	0.120887	0.8	1.100	1.2	1.3	1.4
doggo/floofer	2.0	1.100000	0.000000	1.1	1.100	1.1	1.1	1.1
doggo/pupper	10.0	1.180000	0.103280	1.0	1.200	1.2	1.2	1.3
doggo/puppo	2.0	1.300000	0.000000	1.3	1.300	1.3	1.3	1.3
floofer	6.0	1.266667	0.051640	1.2	1.225	1.3	1.3	1.3
pupper	74.0	1.267568	0.257548	1.1	1.200	1.2	1.3	2.7
puppo	32.0	1.268750	0.059229	1.2	1.200	1.3	1.3	1.4

### 1.3.3 Visualizing Data

```
In [74]: # Plot the rating and the correspondong dog stages data
plt.figure(figsize=(8,4))
plt_stage = sns.regplot(x='dog_stage', y='rating', fit_reg=False, data=df.query('dog_st
plt.axhline(y=np.percentile(df.rating,10));
plt.axhline(y=np.percentile(df.rating,90));
```



**Insight 3:** Regarding the dog stages (if mentioned in the tweet), the ratings of tweets that mention the stages, including pupper, floofer, doggo, puppo, or two of them, distribute in the middle of the whole ratings. This statistic rarely shows the evidence about rating bias based on the dog stages.

The "outliers" in the dataset

```
In [76]: df.query('rating >= 7.5')
```

```
Out[76]:
```

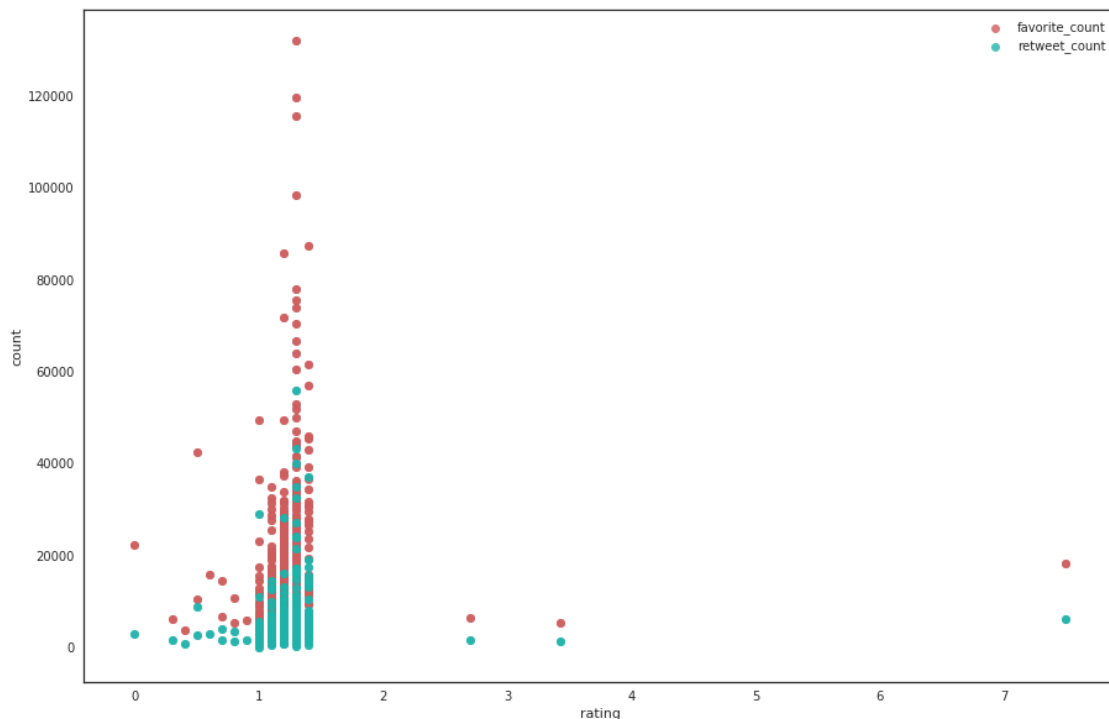
	tweet_id	timestamp	source	text	jpg_url	short_url
309	786709082849828864	2016-10-13 23:23:56	iPhone	This is Logan, the Chow who lived. He solemnly...	https://pbs.twimg.com/media/CurzvFTXgAA2_AP.jpg	https://t.co/yB05wuqaPS
310	786709082849828864	2016-10-13 23:23:56	iPhone	This is Logan, the Chow who lived. He solemnly...	https://pbs.twimg.com/media/CurzvFTXgAA2_AP.jpg	https://t.co/yB05wuqaPS

	name	dog_stage	rating	favorite_count	retweet_count
309	Logan	None	7.5	18430	6182
310	Logan	None	7.5	18430	6182

```
In [79]: # Plot the rating(<25) and the corresponding favorite and retweet counts
plt.figure(figsize=(15,10))
plt_rating = sns.regplot(x='rating', y='favorite_count', fit_reg=False, data=df.query('r
plt_rating = sns.regplot(x='rating', y='retweet_count', fit_reg=False, data=df.query('r
plt_rating.set(xlabel="rating", ylabel="count");
plt.legend(('favorite_count', 'retweet_count'))
```

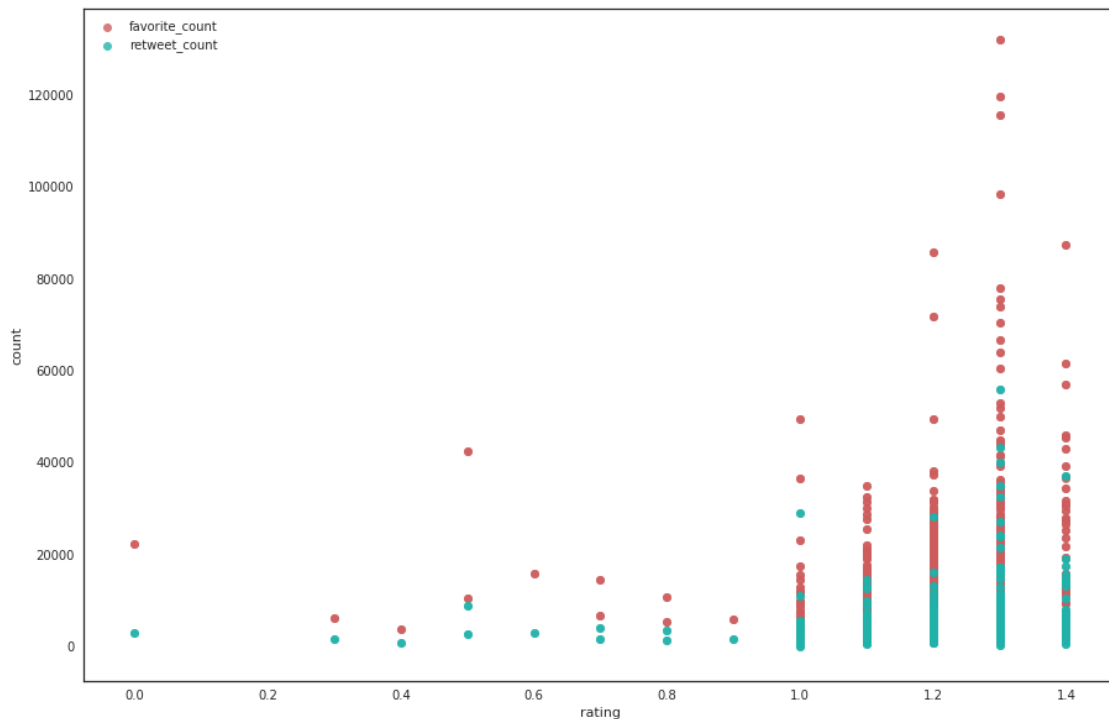
```
Out[79]: <matplotlib.legend.Legend at 0x7f27ba402d68>
```





```
In [80]: # Plot the rating(<2) and the corresponding favorite and retweet counts
plt.figure(figsize=(15,10))
plt_rating = sns.regplot(x='rating', y='favorite_count', fit_reg=False, data=df.query('r
plt_rating = sns.regplot(x='rating', y='retweet_count', fit_reg=False, data=df.query('r
plt_rating.set(xlabel="rating", ylabel="count");
plt.legend(('favorite_count', 'retweet_count'))
```

```
Out[80]: <matplotlib.legend.Legend at 0x7f27b9da69b0>
```



**Insight 4:** With the rating range from 0 to 1.4, the higher rating tends to gain more “likes”. The relationship between the retweet counts and the rating is not as strong as that between favorite counts and ratings.

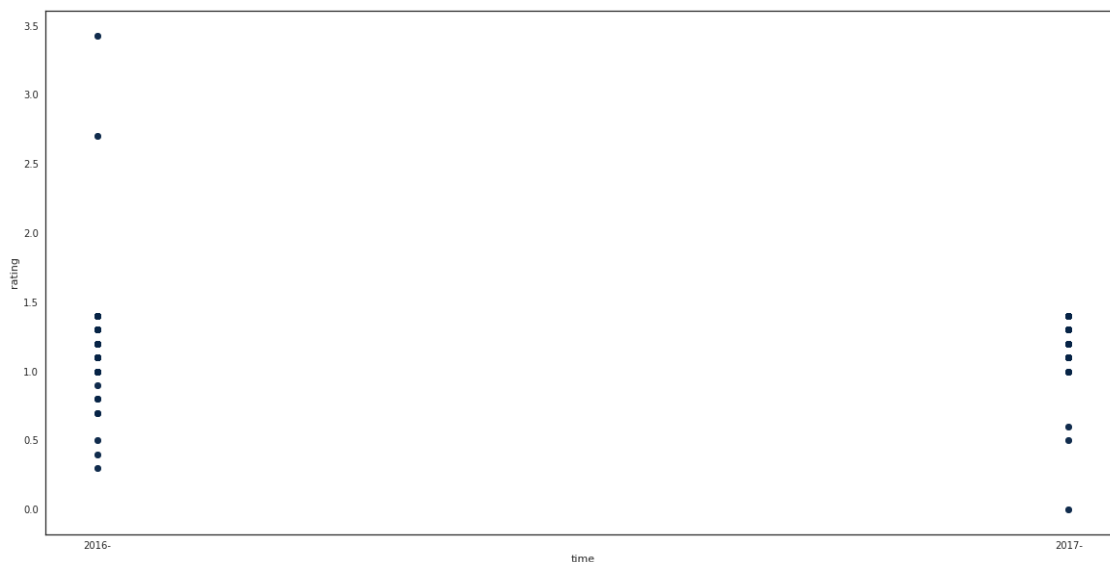
```
In [86]: # Plot the time by year and month and the corresponding favorite and retweet counts
df['year_month'] = df['timestamp'].str[: -len('23:05:30+00:00')]
plt.figure(figsize=(20,10))
plt_rating = sns.regplot(x='year_month', y='favorite_count', fit_reg=False, data=df, co
plt_rating = sns.regplot(x='year_month', y='retweet_count', fit_reg=False, data=df, col
plt_rating.set(xlabel="time", ylabel="count");
plt.legend(('favorite_count', 'retweet_count'))
```

```
Out[86]: <matplotlib.legend.Legend at 0x7f27b835ce10>
```



**Insight 5:** This visualization reiterates the history of this twitter account: in 2017 and 2016, it gained its most “likes” count, the time does not seem to be a strong factor leading the retweet counts.

```
In [88]: # Plot the time by year and month and the corresponding ratings
plt.figure(figsize=(20,10))
plt_ratedtime = sns.regplot(x='year_month', y='rating', fit_reg=False, data=df.query('r
plt_ratedtime.set(xlabel="time", ylabel="rating");
```



**Insight 6:** as the “rating right” belongs to the owner of the account, I would like to know whether the rating habit has changed with time, so I made another scatterplot which shows clearly that the range of rating was narrowed according to the time.

```
In [ ]:
```