wrangle_act

September 21, 2022

1 Project: Wrangling and Analyze Data

1.1 Table of Content

1.2 Data Gathering

In the cell below, gather **all** three pieces of data for this project and load them in the notebook.
Note: the methods required to gather each data are different. 1. Directly download the WeRate-Dogs Twitter archive data (twitter_archive_enhanced.csv)

```python
In [1]: # import neccesary library

        import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        %matplotlib inline

In [2]: # import neccesary library

        import tweepy
        from tweepy import OAuthHandler
        import json
        from timeit import default_timer as timer

        pd.options.display.max_colwidth = 600

In [3]: #collect twitter archive data

        df_1 = pd.read_csv('twitter-archive-enhanced.csv')
        df_1.head(2)
```

```
Out[3]:            tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
        0  892420643555336193                    NaN                  NaN
        1  892177421306343426                    NaN                  NaN

                       timestamp  \
        0  2017-08-01 16:23:56 +0000
        1  2017-08-01 00:17:27 +0000

                                                                   source  \
        0  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>
        1  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>

        0                                          This is Phineas. He's a mystical
        1  This is Tilly. She's just checking pup on you. Hopes you're doing ok. If not, she's a

           retweeted_status_id  retweeted_status_user_id retweeted_status_timestamp  \
        0                  NaN                       NaN                        NaN
        1                  NaN                       NaN                        NaN

                                              expanded_urls  \
        0  https://twitter.com/dog_rates/status/892420643555336193/photo/1
        1  https://twitter.com/dog_rates/status/892177421306343426/photo/1

           rating_numerator  rating_denominator     name doggo floofer pupper puppo
        0                13                  10  Phineas  None    None   None  None
        1                13                  10    Tilly  None    None   None  None
```

In [4]: # Use the Requests library to download the tweet image prediction (image_predictions.tsu

```python
import re
import requests as req
url = "https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predictio
```

In [5]: data = req.get(url)
        data.headers['content-type']

Out[5]: 'text/tab-separated-values; charset=utf-8'

In [6]: data.raise_for_status()
        open('samplefile.txt', 'wb').write(data.content)

Out[6]: 335079

In [7]: df_2 = pd.read_table('samplefile.txt', sep='\t')
        df_2.head(2)

```
Out[7]:            tweet_id                                        jpg_url  \
        0  666020888022790149  https://pbs.twimg.com/media/CT4udnOWwAAOaMy.jpg
```

```
      1  666029285002620928   https://pbs.twimg.com/media/CT42GRgUYAA5iDo.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

        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
```

In [8]:
```python
# Query Twitter API for each tweet in the Twitter archive and save JSON in a text file
# These are hidden to comply with Twitter's API terms and conditions
consumer_key = ''
consumer_secret = ''
access_token = ''
access_secret = ''
```

In [9]:
```python
auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)
```

In [10]:
```python
# Use the Tweepy library to query additional data via the Twitter API (tweet_json.txt)

api = tweepy.API(auth, wait_on_rate_limit=True)
```

In [11]:
```python
# NOTE TO REVIEWER: this student had mobile verification issues so the following
# Twitter API code was sent to this student from a Udacity instructor
# Tweet IDs for which to gather additional data via Twitter's API
tweet_ids = df_1.tweet_id.values
len(tweet_ids)
```

Out[11]: 2356

In [12]:
```python
# Query Twitter's API for JSON data for each tweet ID in the Twitter archive
count = 0
fails_dict = {}
start = timer()
```

In [13]:
```python
# Save each tweet's returned JSON as a new line in a .txt file
with open('tweet_json.txt', 'w') as outfile:
    # This loop will likely take 20-30 minutes to run because of Twitter's rate limit
    for tweet_id in tweet_ids:
        count += 1
        print(str(count) + ": " + str(tweet_id))
        try:
            tweet = api.get_status(tweet_id, tweet_mode='extended')
            print("Success")
            json.dump(tweet._json, outfile)
            outfile.write('\n')
        except tweepy.TweepError as e:
```

```python
                print("Fail")
                fails_dict[tweet_id] = e
                pass
        end = timer()
        #api = tweepy.API(auth, wait_on_rate_limit=True, sleep_on_rate_limit=False)

        print(end - start)
        print(fails_dict)
```

```
1: 892420643555336193
Success
2: 892177421306343426
Success
3: 891815181378084864
Success
4: 891689557279858688
Success
5: 891327558926688256
Success
6: 891087950875897856
Success
7: 890971913173991426
Success
8: 890729181411237888
Success
9: 890609185150312448
Success
10: 890240255349198849
Success
11: 890006608113172480
Success
12: 889880896479866881
Success
13: 889665388333682689
Success
14: 889638837579907072
Success
15: 889531135344209921
Success
16: 889278841981685760
Success
17: 888917238123831296
Success
18: 888804989199671297
Success
19: 888554962724278272
Success
20: 888202515573088257
```

```
Fail
21: 888078434458587136
Success
22: 887705289381826560
Success
23: 887517139158093824
Success
24: 887473957103951883
Success
25: 887343217045368832
Success
26: 887101392804085760
Success
27: 886983233522544640
Success
28: 886736880519319552
Success
29: 886680336477933568
Success
30: 886366144734445568
Success
31: 886267009285017600
Success
32: 886258384151887873
Success
33: 886054160059072513
Success
34: 885984800019947520
Success
35: 885528943205470208
Success
36: 885518971528720385
Success
37: 885311592912609280
Success
38: 885167619883638784
Success
39: 884925521741709313
Success
40: 884876753390489601
Success
41: 884562892145688576
Success
42: 884441805382717440
Success
43: 884247878851493888
Success
44: 884162670584377345
```

```
Success
45: 883838122936631299
Success
46: 883482846933004288
Success
47: 883360690899218434
Success
48: 883117836046086144
Success
49: 882992080364220416
Success
50: 882762694511734784
Success
51: 882627270321602560
Success
52: 882268110199369728
Success
53: 882045870035918850
Success
54: 881906580714921986
Success
55: 881666595344535552
Success
56: 881633300179243008
Success
57: 881536004380872706
Success
58: 881268444196462592
Success
59: 880935762899988482
Success
60: 880872448815771648
Success
61: 880465832366813184
Success
62: 880221127280381952
Success
63: 880095782870896641
Success
64: 879862464715927552
Success
65: 879674319642796034
Success
66: 879492040517615616
Success
67: 879415818425184262
Success
68: 879376492567855104
```

```
Success
69: 8791305795764755649
Success
70: 8790507492626555488
Success
71: 8790082295310295606
Success
72: 8787760934230876188
Success
73: 8786047072117268552
Success
74: 8784047773481369664
Success
75: 8783161107680870411
Success
76: 8782815110064783366
Success
77: 8780576130401157122
Success
78: 8777364723291914244
Success
79: 8776111728322273288
Success
80: 8775562467312148488
Success
81: 8773168213214289993
Success
82: 8772018374259261444
Success
83: 8768381206285393922
Success
84: 8765376660612218899
Success
85: 8764840539098726440
Success
86: 8761202751961702400
Success
87: 8757477678675230722
Success
88: 8751442898561146888
Success
89: 8750971926120775688
Success
90: 8750212112515973122
Success
91: 8746800970551787522
Success
92: 8744348182595256344
```

Success
93: 874296783580663808
Success
94: 874057562936811520
Success
95: 874012996292530176
Success
96: 873697596434513921
Fail
97: 873580283840344065
Success
98: 873337748698140672
Success
99: 873213775632977920
Success
100: 872967104147763200
Success
101: 872820683541237760
Success
102: 872668790621863937
Fail
103: 872620804844003328
Success
104: 872486979161796608
Success
105: 872261713294495745
Fail
106: 872122724285648897
Success
107: 871879754684805121
Success
108: 871762521631449091
Success
109: 871515927908634625
Success
110: 871166179821445120
Success
111: 871102520638267392
Success
112: 871032628920680449
Success
113: 870804317367881728
Success
114: 870726314365509632
Success
115: 870656317836468226
Success
116: 870374049280663552

```
Success
117: 870308999962521604
Success
118: 870063196459192321
Success
119: 869988702071779329
Fail
120: 869772420881756160
Success
121: 869702957897576449
Success
122: 869596645499047938
Success
123: 869227993411051520
Success
124: 868880397819494401
Success
125: 868639477480148993
Success
126: 868622495443632128
Success
127: 868552278524837888
Success
128: 867900495410671616
Success
129: 867774946302451713
Success
130: 867421006826221569
Success
131: 867072653475098625
Success
132: 867051520902168576
Success
133: 866816280283807744
Fail
134: 866720684873056260
Success
135: 866686824827068416
Success
136: 866450705531457537
Success
137: 866334964761202691
Success
138: 866094527597207552
Success
139: 865718153858494464
Success
140: 865359393868664832
```

```
Success
141: 865006731092295680
Success
142: 864873206498414592
Success
143: 864279568663928832
Success
144: 864197398364647424
Success
145: 863907417377173506
Success
146: 863553081350529029
Success
147: 863471782782697472
Success
148: 863432100342583297
Success
149: 863427515083354112
Success
150: 863079547188785154
Success
151: 863062471531167744
Success
152: 862831371563274240
Success
153: 862722525377298433
Success
154: 862457590147678208
Success
155: 862096992088072192
Success
156: 861769973181624320
Fail
157: 861383897657036800
Success
158: 861288531465048066
Success
159: 861005113778896900
Success
160: 860981674716409858
Success
161: 860924035999428608
Success
162: 860563773140209665
Success
163: 860524505164394496
Success
164: 860276583193509888
```

```
Success
165: 860184849394610176
Success
166: 860177593139703809
Success
167: 859924526012018688
Success
168: 859851578198683649
Success
169: 859607811541651456
Success
170: 859196978902773760
Success
171: 859074603037188101
Success
172: 858860390427611136
Success
173: 858843525470990336
Success
174: 858471635011153920
Success
175: 858107933456039936
Success
176: 857989990357356544
Success
177: 857746408056729600
Success
178: 857393404942143489
Success
179: 857263160327368704
Success
180: 857214891891077121
Success
181: 857062103051644929
Success
182: 857029823797047296
Success
183: 856602993587888130
Fail
184: 856543823941562368
Success
185: 856526610513747968
Success
186: 856330835276025856
Fail
187: 856288084350160898
Success
188: 856282028240666624
```

```
Success
189: 855862651834028034
Success
190: 855860136149123072
Success
191: 855857698524602368
Success
192: 855851453814013952
Success
193: 855818117272018944
Success
194: 855459453768019968
Success
195: 855245323840757760
Success
196: 855138241867124737
Success
197: 854732716440526848
Success
198: 854482394044301312
Success
199: 854365224396361728
Success
200: 854120357044912130
Success
201: 854010172552949760
Success
202: 853760880890318849
Success
203: 853639147608842240
Success
204: 853299958564483072
Success
205: 852936405516943360
Success
206: 852912242202992640
Success
207: 852672615818899456
Success
208: 852553447878664193
Success
209: 852311364735569921
Success
210: 852226086759018497
Success
211: 852189679701164033
Success
212: 851953902622658560
```

Fail
213: 851861385021730816
Fail
214: 851591660324737024
Success
215: 851464819735769094
Success
216: 851224888060895234
Success
217: 850753642995093505
Success
218: 850380195714523136
Success
219: 850333567704068097
Success
220: 850145622816686080
Success
221: 850019790995546112
Success
222: 849776966551130114
Success
223: 849668094696017920
Success
224: 849412302885593088
Success
225: 849336543269576704
Success
226: 849051919805034497
Success
227: 848690551926992896
Success
228: 848324959059550208
Success
229: 848213670039564288
Success
230: 848212111729840128
Success
231: 847978865427394560
Success
232: 847971574464610304
Success
233: 847962785489326080
Success
234: 847842811428974592
Success
235: 847617282490613760
Success
236: 847606175596138505

Success
237: 847251039262605312
Success
238: 847157206088847362
Success
239: 847116187444137987
Success
240: 846874817362120707
Success
241: 846514051647705089
Success
242: 846505985330044928
Success
243: 846153765933735936
Success
244: 846139713627017216
Success
245: 846042936437604353
Success
246: 845812042753855489
Success
247: 845677943972139009
Success
248: 845459076796616705
Fail
249: 845397057150107648
Success
250: 845306882940190720
Success
251: 845098359547420673
Success
252: 844979544864018432
Success
253: 844973813909606400
Success
254: 844704788403113984
Fail
255: 844580511645339650
Success
256: 844223788422217728
Success
257: 843981021012017153
Success
258: 843856843873095681
Success
259: 843604394117681152
Success
260: 843235543001513987

```
Success
261: 842892208864923648
Fail
262: 842846295480000512
Success
263: 842765311967449089
Success
264: 842535590457499648
Success
265: 842163532590374912
Success
266: 842115215311396866
Success
267: 841833993020538882
Success
268: 841680585030541313
Success
269: 841439858740625411
Success
270: 841320156043304961
Success
271: 841314665196081154
Success
272: 841077006473256960
Success
273: 840761248237133825
Success
274: 840728873075638272
Success
275: 840698636975636481
Success
276: 840696689258311684
Success
277: 840632337062862849
Success
278: 840370681858686976
Success
279: 840268004936019968
Success
280: 839990271299457024
Success
281: 839549326359670784
Success
282: 839290600511926273
Success
283: 839239871831150596
Success
284: 838952994649550848
```

```
Success
285: 838921590096166913
Success
286: 838916489579200512
Success
287: 838831947270979586
Success
288: 838561493054533637
Success
289: 838476387338051585
Success
290: 838201503651401729
Success
291: 838150277551247360
Success
292: 838085839343206401
Success
293: 838083903487373313
Success
294: 837820167694528512
Success
295: 837482249356513284
Success
296: 837471256429613056
Success
297: 837366284874571778
Fail
298: 837110210464448512
Success
299: 837012587749474308
Fail
300: 836989968035819520
Success
301: 836753516572119041
Success
302: 836677758902222849
Success
303: 836648853927522308
Success
304: 836397794269200385
Success
305: 836380477523124226
Success
306: 836260088725786625
Success
307: 836001077879255040
Success
308: 835685285446955009
```

```
Success
309: 835574547218894849
Success
310: 835536468978302976
Success
311: 835309094223372289
Success
312: 835297930240217089
Success
313: 835264098648616962
Success
314: 835246439529840640
Success
315: 835172783151792128
Success
316: 835152434251116546
Success
317: 834931633769889797
Success
318: 834786237630337024
Success
319: 834574053763584002
Success
320: 834477809192075265
Success
321: 834458053273591808
Success
322: 834209720923721728
Success
323: 834167344700198914
Success
324: 834089966724603904
Success
325: 834086379323871233
Success
326: 833863086058651648
Success
327: 833826103416520705
Success
328: 833732339549220864
Success
329: 833722901757046785
Success
330: 833479644947025920
Success
331: 833124694597443584
Success
332: 832998151111966721
```

```
Success
333: 832769181346996225
Success
334: 832757312314028032
Success
335: 832682457690300417
Success
336: 832645525019123713
Success
337: 832636094638288896
Success
338: 832397543355072512
Success
339: 832369877331693569
Success
340: 832273440279240704
Success
341: 832215909146226688
Success
342: 832215726631055365
Success
343: 832088576586297345
Success
344: 832040443403784192
Success
345: 832032802820481025
Success
346: 831939777352105988
Success
347: 831926988323639298
Success
348: 831911600680497154
Success
349: 831670449226514432
Success
350: 831650051525054464
Success
351: 831552930092285952
Success
352: 831322785565769729
Success
353: 831315979191906304
Success
354: 831309418084069378
Success
355: 831262627380748289
Success
356: 830956169170665475
```

```
Success
357: 830583320585068544
Success
358: 830173239259324417
Success
359: 830097400375152640
Success
360: 829878982036299777
Success
361: 829861396166877184
Success
362: 829501995190984704
Success
363: 829449946868879360
Success
364: 829374341691346946
Fail
365: 829141528400556032
Success
366: 829011960981237760
Success
367: 828801551087042563
Success
368: 828770345708580865
Success
369: 828708714936930305
Success
370: 828650029636317184
Success
371: 828409743546925057
Success
372: 828408677031882754
Success
373: 828381636999917570
Success
374: 828376505180889089
Success
375: 828372645993398273
Success
376: 828361771580813312
Success
377: 828046555563323392
Success
378: 828011680017821696
Success
379: 827933404142436356
Success
380: 827653905312006145
```

```
Success
381: 827600520311402496
Success
382: 827324948884643840
Success
383: 827228250799742977
Fail
384: 827199976799354881
Success
385: 826958653328592898
Success
386: 826848821049180160
Success
387: 826615380357632002
Success
388: 826598799820865537
Success
389: 826598365270007810
Success
390: 826476773533745153
Success
391: 826240494070030336
Success
392: 826204788643753985
Success
393: 826115272272650244
Success
394: 825876512159186944
Success
395: 825829644528148480
Success
396: 825535076884762624
Success
397: 825147591692263424
Success
398: 825120256414846976
Success
399: 825026590719483904
Success
400: 824796380199809024
Success
401: 824775126675836928
Success
402: 824663926340194305
Success
403: 824325613288833024
Success
404: 824297048279236611
```

```
Success
405: 824025158776213504
Success
406: 823939628516474880
Success
407: 823719002937630720
Success
408: 823699002998870016
Success
409: 823581115634085888
Success
410: 823333489516937216
Success
411: 823322678127919110
Success
412: 823269594223824897
Success
413: 822975315408461824
Success
414: 822872901745569793
Success
415: 822859134160621569
Success
416: 822647212903690241
Success
417: 822610361945911296
Success
418: 822489057087389700
Success
419: 822462944365645825
Success
420: 822244816520155136
Success
421: 822163064745328640
Success
422: 821886076407029760
Success
423: 821813639212650496
Success
424: 821765923262631936
Success
425: 821522889702862852
Success
426: 821421320206483457
Success
427: 821407182352777218
Success
428: 821153421864615936
```

```
Success
429: 821149554670182400
Success
430: 821107785811234820
Success
431: 821044531881721856
Success
432: 820837357901512704
Success
433: 820749716845686786
Success
434: 820690176645140481
Success
435: 820494788566847489
Success
436: 820446719150292993
Success
437: 820314633777061888
Success
438: 820078625395449857
Success
439: 820013781606658049
Success
440: 819952236453363712
Success
441: 819924195358416896
Success
442: 819711362133872643
Success
443: 819588359383371776
Success
444: 819347104292290561
Success
445: 819238181065359361
Success
446: 819227688460238848
Success
447: 819015337530290176
Success
448: 819015331746349057
Success
449: 819006400881917954
Success
450: 819004803107983360
Success
451: 818646164899774465
Success
452: 818627210458333184
```

```
Success
453: 818614493328580609
Success
454: 818588835076603904
Success
455: 818536468981415936
Success
456: 818307523543449600
Success
457: 818259473185828864
Success
458: 818145370475810820
Success
459: 817908911860748288
Success
460: 817827839487737858
Success
461: 817777686764523521
Success
462: 817536400337801217
Success
463: 817502432452313088
Success
464: 817423860136083457
Success
465: 817415592588222464
Success
466: 817181837579653120
Success
467: 817171292965273600
Success
468: 817120970343411712
Success
469: 817056546584727552
Success
470: 816829038950027264
Success
471: 816816676327063552
Success
472: 816697700272001025
Success
473: 816450570814898180
Success
474: 816336735214911488
Success
475: 816091915477250048
Success
476: 816062466425819140
```

```
Success
477: 816014286006976512
Success
478: 815990720817401858
Success
479: 815966073409433600
Success
480: 815745968457060357
Success
481: 815736392542261248
Success
482: 815639385530101762
Success
483: 815390420867969024
Success
484: 814986499976527872
Success
485: 814638523311648768
Success
486: 814578408554463233
Success
487: 814530161257443328
Success
488: 814153002265309185
Success
489: 813944609378369540
Success
490: 813910438903693312
Success
491: 813812741911748608
Success
492: 813800681631023104
Success
493: 813217897535406080
Success
494: 813202720496779264
Success
495: 813187593374461952
Success
496: 813172488309972993
Success
497: 813157409116065792
Success
498: 813142292504645637
Success
499: 813130366689148928
Success
500: 813127251579564032
```

```
Success
501: 813112105746448384
Success
502: 813096984823349248
Success
503: 813081950185472002
Success
504: 813066809284972545
Success
505: 813051746834595840
Success
506: 812781120811126785
Success
507: 812747805718642688
Fail
508: 812709060537683968
Success
509: 812503143955202048
Success
510: 812466873996607488
Success
511: 812372279581671427
Success
512: 811985624773361665
Success
513: 811744202451197953
Success
514: 811647686436880384
Success
515: 811627233043480576
Success
516: 811386762094317568
Success
517: 810984652412424192
Success
518: 810896069567610880
Success
519: 810657578271330305
Success
520: 810284430598270976
Success
521: 810254108431155201
Success
522: 809920764300447744
Success
523: 809808892968534016
Success
524: 809448704142938112
```

```
Success
525: 8092200512116039969
Success
526: 809084759137812480
Success
527: 808838249661788160
Success
528: 808733504066486276
Success
529: 808501579447930884
Success
530: 808344865868283904
Success
531: 808134635716833280
Success
532: 808106460588765185
Success
533: 808001312164028416
Success
534: 807621403335917568
Success
535: 807106840509214720
Success
536: 807059379405148160
Success
537: 807010152071229440
Success
538: 806629075125202948
Success
539: 806620845233815552
Success
540: 806576416489959424
Success
541: 806542213899489280
Success
542: 806242860592926720
Success
543: 806219024703037440
Success
544: 805958939288408065
Success
545: 805932879469572096
Success
546: 805826884734976000
Success
547: 805823200554876929
Success
548: 805520635690676224
```

```
Success
549: 805487436403003392
Success
550: 805207613751304193
Success
551: 804738756058218496
Success
552: 804475857670639616
Success
553: 804413760345620481
Success
554: 804026241225523202
Success
555: 803773340896923648
Success
556: 803692223237865472
Success
557: 803638050916102144
Success
558: 803380650405482500
Success
559: 803321560782307329
Success
560: 803276597545603072
Success
561: 802952499103731712
Success
562: 802624713319034886
Success
563: 802600418706604034
Success
564: 802572683846291456
Success
565: 802323869084381190
Success
566: 802265048156610565
Success
567: 802247111496568832
Fail
568: 802239329049477120
Success
569: 802185808107208704
Success
570: 801958328846974976
Success
571: 801854953262350336
Success
572: 801538201127157760
```

```
Success
573: 801285448605831168
Success
574: 801167903437357056
Success
575: 801127390143516673
Success
576: 801115127852503040
Success
577: 800859414831898624
Success
578: 800855607700029440
Success
579: 800751577355128832
Success
580: 800513324630806528
Success
581: 800459316964663297
Success
582: 800443802682937345
Success
583: 800388270626521089
Success
584: 800188575492947969
Success
585: 800141422401830912
Success
586: 800018252395122689
Success
587: 799774291445383169
Success
588: 799757965289017345
Success
589: 799422933579902976
Success
590: 799308762079035393
Success
591: 799297110730567681
Success
592: 799063482566066176
Success
593: 798933969379225600
Success
594: 798925684722855936
Success
595: 798705661114773508
Success
596: 798701998996647937
```

```
Success
597: 798697898615730177
Success
598: 798694562394996736
Success
599: 798686750113755136
Success
600: 798682547630837760
Success
601: 798673117451325440
Success
602: 798665375516884993
Success
603: 798644042770751489
Success
604: 798628517273620480
Success
605: 798585098161549313
Success
606: 798576900688019456
Success
607: 798340744599797760
Success
608: 798209839306514432
Success
609: 797971864723324932
Success
610: 797545162159308800
Success
611: 797236660651966464
Success
612: 797165961484890113
Success
613: 796904159865868288
Success
614: 796865951799083009
Success
615: 796759840936919040
Success
616: 796563435802726400
Success
617: 796484825502875648
Success
618: 796387464403357696
Success
619: 796177847564038144
Success
620: 796149749086875649
```

```
Success
621: 796125600683540480
Success
622: 796116448414461957
Success
623: 796080075804475393
Success
624: 796031486298386433
Success
625: 795464331001561088
Success
626: 795400264262053889
Success
627: 795076730285391872
Success
628: 794983741416415232
Success
629: 794926597468000259
Success
630: 794355576146903043
Success
631: 794332329137291264
Success
632: 794205286408003585
Success
633: 793962221541933056
Success
634: 793845145112371200
Success
635: 793614319594401792
Success
636: 793601777308463104
Success
637: 793500921481273345
Success
638: 793286476301799424
Success
639: 793271401113350145
Success
640: 793256262322548741
Success
641: 793241302385262592
Success
642: 793226087023144960
Success
643: 793210959003287553
Success
644: 793195938047070209
```

```
Success
645: 793180763617361921
Success
646: 793165685325201412
Success
647: 793150605191548928
Success
648: 793135492858580992
Success
649: 793120401413079041
Success
650: 792913359805018113
Success
651: 792883833364439040
Success
652: 792773781206999040
Success
653: 792394556390137856
Success
654: 792050063153438720
Success
655: 791821351946420224
Success
656: 791784077045166082
Success
657: 791780927877898241
Success
658: 791774931465953280
Success
659: 791672322847637504
Success
660: 791406955684368384
Success
661: 791312159183634433
Success
662: 791026214425268224
Success
663: 790987426131050500
Success
664: 790946055508652032
Success
665: 790723298204217344
Success
666: 790698755171364864
Success
667: 790581949425475584
Success
668: 790337589677002753
```

Success
669: 790277117346975746
Success
670: 790227638568808452
Success
671: 789986466051088384
Success
672: 789960241177853952
Success
673: 789903600034189313
Success
674: 789628658055020548
Success
675: 789599242079838210
Success
676: 789530877013393408
Success
677: 789314372632018944
Success
678: 789280767834746880
Success
679: 789268448748703744
Success
680: 789137962068021249
Success
681: 788908386943430656
Success
682: 788765914992902144
Success
683: 788552643979468800
Success
684: 788412144018661376
Success
685: 788178268662984705
Success
686: 788150585577050112
Success
687: 788070120937619456
Success
688: 788039637453406209
Success
689: 787810552592695296
Success
690: 787717603741622272
Success
691: 787397959788929025
Success
692: 787322443945877504

```
Success
693: 787111942498508800
Success
694: 786963064373534720
Success
695: 786729988674449408
Success
696: 786709082849828864
Success
697: 786664955043049472
Success
698: 786595970293370880
Success
699: 786363235746385920
Success
700: 786286427768250368
Success
701: 786233965241827333
Success
702: 786051337297522688
Success
703: 786036967502913536
Success
704: 785927819176054784
Success
705: 785872687017132033
Success
706: 785639753186217984
Success
707: 785533386513321988
Success
708: 785515384317313025
Success
709: 785264754247995392
Success
710: 785170936622350336
Success
711: 784826020293709826
Success
712: 784517518371221505
Success
713: 784431430411685888
Success
714: 784183165795655680
Success
715: 784057939640352768
Success
716: 783839966405230592
```

```
Success
717: 783821107061198850
Success
718: 783695101801398276
Success
719: 783466772167098368
Success
720: 783391753726550016
Success
721: 783347506784731136
Success
722: 783334639985389568
Success
723: 783085703974514689
Success
724: 782969140009107456
Success
725: 782747134529531904
Success
726: 782722598790725632
Success
727: 782598640137187329
Success
728: 782305867769217024
Success
729: 782021823840026624
Success
730: 781955203444699136
Success
731: 781661882474196992
Success
732: 781655249211752448
Success
733: 781524693396357120
Success
734: 781308096455073793
Success
735: 781251288990355457
Success
736: 781163403222056960
Success
737: 780931614150983680
Success
738: 780858289093574656
Success
739: 780800785462489090
Success
740: 780601303617732608
```

```
Success
741: 780543529827336192
Success
742: 780496263422808064
Success
743: 780476555013349377
Success
744: 780459368902959104
Success
745: 780192070812196864
Success
746: 780092040432480260
Success
747: 780074436359819264
Success
748: 779834332596887552
Success
749: 779377524342161408
Success
750: 779124354206535695
Success
751: 779123168116150273
Fail
752: 779056095788752897
Success
753: 778990705243029504
Success
754: 778774459159379968
Success
755: 778764940568104960
Success
756: 778748913645780993
Success
757: 778650543019483137
Success
758: 778624900596654080
Success
759: 778408200802557953
Success
760: 778396591732486144
Success
761: 778383385161035776
Success
762: 778286810187399168
Success
763: 778039087836069888
Success
764: 778027034220126208
```

```
Success
765: 777953400541634568
Success
766: 777885040357281792
Success
767: 777684233540206592
Success
768: 777641927919427584
Success
769: 777621514455814149
Success
770: 777189768882946048
Success
771: 776819012571455488
Success
772: 776813020089548800
Success
773: 776477788987613185
Success
774: 776249906839351296
Success
775: 776218204058357768
Success
776: 776201521193218049
Success
777: 776113305656188928
Success
778: 776088319444877312
Success
779: 775898661951791106
Success
780: 775842724423557120
Success
781: 775733305207554048
Success
782: 775729183532220416
Success
783: 775364825476165632
Success
784: 775350846108426240
Success
785: 775096608509886464
Fail
786: 775085132600442880
Success
787: 774757898236878852
Success
788: 774639387460112384
```

```
Success
789: 774314403806253056
Success
790: 773985732834758656
Success
791: 773922284943896577
Success
792: 773704687002451968
Success
793: 773670353721753600
Success
794: 773547596996571136
Success
795: 773336787167145985
Success
796: 773308824254029826
Success
797: 773247561583001600
Success
798: 773191612633579521
Success
799: 772877495989305348
Success
800: 772826264096874500
Success
801: 772615324260794368
Success
802: 772581559778025472
Success
803: 772193107915964416
Success
804: 772152991789019136
Success
805: 772117678702071809
Success
806: 772114945936949249
Success
807: 772102971039580160
Success
808: 771908950375665664
Success
809: 771770456517009408
Success
810: 771500966810099713
Success
811: 771380798096281600
Success
812: 771171053431250945
```

```
Success
813: 771136648247640064
Success
814: 771102124360998913
Success
815: 771014301343748096
Success
816: 771004394259247104
Fail
817: 770787852854652928
Success
818: 770772759874076672
Success
819: 770743923962707968
Fail
820: 770655142660169732
Success
821: 770414278348247044
Success
822: 770293558247038976
Success
823: 770093767776997377
Success
824: 770069151037685760
Success
825: 769940425801170949
Success
826: 769695466921623552
Success
827: 769335591808995329
Success
828: 769212283578875904
Success
829: 768970937022709760
Success
830: 768909767477751808
Success
831: 768855141948723200
Success
832: 768609597686943744
Success
833: 768596291618299904
Success
834: 768554158521745409
Success
835: 768473857036525572
Success
836: 768193404517830656
```

```
Success
837: 767884188863397888
Success
838: 767754930266464257
Success
839: 767500508068192258
Success
840: 767191397493538821
Success
841: 767122157629476866
Success
842: 766864461642756096
Fail
843: 766793450729734144
Success
844: 766714921925144576
Success
845: 766693177336135680
Success
846: 766423258543644672
Success
847: 766313316352462849
Success
848: 766078092750233600
Success
849: 766069199026450432
Success
850: 766008592277377025
Success
851: 765719909049503744
Success
852: 765669560888528897
Success
853: 765395769549590528
Success
854: 765371061932261376
Success
855: 765222098633691136
Success
856: 764857477905154048
Success
857: 764259802650378240
Success
858: 763956972077010945
Success
859: 763837565564780549
Success
860: 763183847194451968
```

```
Success
861: 763167063695355904
Success
862: 763103485927849985
Success
863: 762699858130116608
Success
864: 762471784394268675
Success
865: 762464539388485633
Success
866: 762316489655476224
Success
867: 762035686371364864
Success
868: 761976711479193600
Success
869: 761750502866649088
Success
870: 761745352076779520
Success
871: 761672994376806400
Success
872: 761599872357261312
Success
873: 761371037149827077
Success
874: 761334018830917632
Success
875: 761292947749015552
Success
876: 761227390836215808
Success
877: 761004547850530816
Success
878: 760893934457552897
Success
879: 760656994973933572
Success
880: 760641137271070720
Success
881: 760539183865880579
Success
882: 760521673607086080
Success
883: 760290219849637889
Success
884: 760252756032651264
```

```
Success
885: 760190180481531904
Success
886: 760153949710192640
Success
887: 759943073749200896
Success
888: 759923798737051648
Fail
889: 759846353224826880
Success
890: 759793422261743616
Success
891: 759566828574212096
Fail
892: 759557299618865152
Success
893: 759447681597108224
Success
894: 759446261539934208
Success
895: 759197388317847553
Success
896: 759159934323924993
Success
897: 759099523532779520
Success
898: 759047813560868866
Success
899: 758854675097526272
Success
900: 758828659922702336
Success
901: 758740312047005698
Success
902: 758474966123810816
Success
903: 758467244762497024
Success
904: 758405701903519748
Success
905: 758355060040593408
Success
906: 758099635764359168
Success
907: 758041019896193024
Success
908: 757741869644341248
```

```
Success
909: 757729163776290825
Success
910: 757725642876129280
Success
911: 757611664640446465
Success
912: 757597904299253760
Success
913: 757596066325864448
Success
914: 757400162377592832
Success
915: 757393109802180609
Success
916: 757354760399941633
Success
917: 756998049151549440
Success
918: 756939218950160384
Success
919: 756651752796094464
Success
920: 756526248105566208
Success
921: 756303284449767430
Success
922: 756288534030475264
Success
923: 756275833623502848
Success
924: 755955933503782912
Success
925: 755206590534418437
Success
926: 755110668769038337
Success
927: 754874841593970688
Success
928: 754856583969079297
Success
929: 754747087846248448
Success
930: 754482103782404096
Success
931: 754449512966619136
Success
932: 754120377874386944
```

```
Success
933: 754011816964026368
Fail
934: 753655901052166144
Success
935: 753420520834629632
Success
936: 753398408988139520
Success
937: 753375668877008896
Success
938: 753298634498793472
Success
939: 753294487569522689
Success
940: 753039830821511168
Success
941: 753026973505581056
Success
942: 752932432744185856
Success
943: 752917284578922496
Success
944: 752701944171524096
Success
945: 752682090207055872
Success
946: 752660715232722944
Success
947: 752568224206688256
Success
948: 752519690950500352
Success
949: 752334515931054080
Success
950: 752309394570878976
Success
951: 752173152931807232
Success
952: 751950017322246144
Success
953: 751937170840121344
Success
954: 751830394383790080
Success
955: 751793661361422336
Success
956: 751598357617971201
```

```
Success
957: 751583847268179968
Success
958: 751538714308972544
Success
959: 751456908746354688
Success
960: 751251247299190784
Success
961: 751205363882532864
Success
962: 751132876104687617
Success
963: 750868782890057730
Success
964: 750719632563142656
Success
965: 750506206503038976
Success
966: 750429297815552001
Success
967: 750383411068534784
Success
968: 750381685133418496
Success
969: 750147208377409536
Success
970: 750132105863102464
Success
971: 750117059602808832
Success
972: 750101899009982464
Success
973: 750086836815486976
Success
974: 750071704093859840
Success
975: 750056684286914561
Success
976: 750041628174217216
Success
977: 750026558547456000
Success
978: 750011400160841729
Success
979: 749996283729883136
Success
980: 749981277374128128
```

```
Success
981: 749774190421639168
Success
982: 749417653287129088
Success
983: 749403093750648834
Success
984: 749395845976588288
Success
985: 749317047558017024
Success
986: 749075273010798592
Success
987: 749064354620928000
Success
988: 749036806121881602
Success
989: 748977405889503236
Success
990: 748932637671223296
Success
991: 748705597323898880
Success
992: 748699167502000129
Success
993: 748692773788876800
Success
994: 748575535303884801
Success
995: 748568946752774144
Success
996: 748346686624440324
Success
997: 748337862848962560
Success
998: 748324050481647620
Success
999: 748307329658011649
Success
1000: 748220828303695873
Success
1001: 747963614829678593
Success
1002: 747933425676525569
Success
1003: 747885874273214464
Success
1004: 747844099428986880
```

```
Success
1005: 747816857231626240
Success
1006: 747651430853525504
Success
1007: 747648653817413632
Success
1008: 747600769478692864
Success
1009: 747594051852075008
Success
1010: 747512671126323200
Success
1011: 747461612269887489
Success
1012: 747439450712596480
Success
1013: 747242308580548608
Success
1014: 747219827526344708
Success
1015: 747204161125646336
Success
1016: 747103485104099331
Success
1017: 746906459439529985
Success
1018: 746872823977771008
Success
1019: 746818907684614144
Success
1020: 746790600704425984
Success
1021: 746757706116112384
Success
1022: 746726898085036033
Success
1023: 746542875601690625
Success
1024: 746521445350707200
Success
1025: 746507379341139972
Success
1026: 746369468511756288
Success
1027: 746131877086527488
Success
1028: 746056683365994496
```

```
Success
1029: 745789745784041472
Success
1030: 745712589599014916
Success
1031: 745433870967832576
Success
1032: 745422732645535745
Success
1033: 745314880350101504
Success
1034: 745074613265149952
Success
1035: 745057283344719872
Success
1036: 744995568523612160
Success
1037: 744971049620602880
Success
1038: 744709971296780288
Success
1039: 744334592493166593
Success
1040: 744234799360020481
Success
1041: 744223424764059648
Success
1042: 743980027717509120
Success
1043: 743895849529389061
Success
1044: 743835915802583040
Success
1045: 743609206067040256
Success
1046: 743595368194129920
Success
1047: 743545585370791937
Success
1048: 743510151680958465
Success
1049: 743253157753532416
Success
1050: 743222593470234624
Success
1051: 743210557239623680
Success
1052: 742534281772302336
```

```
Success
1053: 742528092657332225
Success
1054: 742465774154047488
Success
1055: 742423170473463808
Success
1056: 742385895052087300
Success
1057: 742161199639494656
Success
1058: 742150209887731712
Success
1059: 741793263812808706
Success
1060: 741743634094141440
Success
1061: 741438259667034112
Success
1062: 741303864243200000
Success
1063: 741099773336379392
Success
1064: 741067306818797568
Success
1065: 740995100998766593
Success
1066: 740711788199743490
Success
1067: 740699697422163968
Success
1068: 740676976021798912
Success
1069: 740373189193256964
Success
1070: 740365076218183684
Success
1071: 740359016048689152
Success
1072: 740214038584557568
Success
1073: 739979191639244800
Success
1074: 739932936087216128
Success
1075: 739844404073074688
Success
1076: 739623569819336705
```

```
Success
1077: 739606147276148736
Success
1078: 739544079319588864
Success
1079: 739485634323156992
Success
1080: 739238157791694849
Success
1081: 738891149612572673
Success
1082: 738885046782832640
Success
1083: 738883359779196928
Success
1084: 738537504001953792
Success
1085: 738402415918125056
Success
1086: 738184450748633089
Success
1087: 738166403467907072
Success
1088: 738156290900254721
Success
1089: 737826014890496000
Success
1090: 737800304142471168
Success
1091: 737678689543020544
Success
1092: 737445876994609152
Success
1093: 737322739594330112
Success
1094: 737310737551491075
Success
1095: 736736130620620800
Success
1096: 736392552031657984
Success
1097: 736365877722001409
Success
1098: 736225175608430592
Success
1099: 736010884653420544
Success
1100: 735991953473572864
```

```
Success
1101: 735648611367784448
Success
1102: 735635087207878657
Success
1103: 735274964362878976
Success
1104: 735256018284875776
Success
1105: 735137028879360001
Success
1106: 734912297295085568
Success
1107: 734787690684657664
Success
1108: 734776360183431168
Success
1109: 734559631394082816
Success
1110: 733828123016450049
Success
1111: 733822306246479872
Success
1112: 733482008106668032
Success
1113: 733460102733135873
Success
1114: 733109485275860992
Success
1115: 732732193018155009
Success
1116: 732726085725589504
Success
1117: 732585889486888962
Success
1118: 732375214819057664
Success
1119: 732005617171337216
Success
1120: 731285275100512256
Success
1121: 731156023742988288
Success
1122: 730924654643314689
Success
1123: 730573383004487680
Success
1124: 730427201120833536
```

```
Success
1125: 730211855403241472
Success
1126: 730196704625098752
Success
1127: 729854734790754305
Success
1128: 729838605770891264
Success
1129: 729823566028484608
Success
1130: 729463711119904772
Success
1131: 729113531270991872
Success
1132: 728986383096946689
Success
1133: 728760639972315136
Success
1134: 728751179681943552
Success
1135: 728653952833728512
Success
1136: 728409960103686147
Success
1137: 728387165835677696
Success
1138: 728046963732717569
Success
1139: 728035342121635841
Success
1140: 728015554473250816
Success
1141: 727685679342333952
Success
1142: 727644517743104000
Success
1143: 727524757080539137
Success
1144: 727314416056803329
Success
1145: 727286334147182592
Success
1146: 727175381690781696
Success
1147: 727155742655025152
Success
1148: 726935089318363137
```

```
Success
1149: 726887082820554753
Success
1150: 726828223124897792
Success
1151: 726224900189511680
Success
1152: 725842289046749185
Success
1153: 725786712245440512
Success
1154: 725729321944506368
Success
1155: 725458796924002305
Success
1156: 724983749226668032
Success
1157: 724771698126512129
Success
1158: 724405726123311104
Success
1159: 724049859469295616
Success
1160: 724046343203856385
Success
1161: 724004602748780546
Success
1162: 723912936180330496
Success
1163: 723688335806480385
Success
1164: 723673163800948736
Success
1165: 723179728551723008
Success
1166: 722974582966214656
Success
1167: 722613351520608256
Success
1168: 721503162398597120
Success
1169: 721001180231503872
Success
1170: 720785406564900865
Success
1171: 720775346191278080
Success
1172: 720415127506415616
```

```
Success
1173: 720389942216527872
Success
1174: 720340705894408192
Success
1175: 720059472081784833
Success
1176: 720043174954147842
Success
1177: 719991154352222208
Success
1178: 719704490224398336
Success
1179: 719551379208073216
Success
1180: 719367763014393856
Success
1181: 719339463458033665
Success
1182: 719332531645071360
Success
1183: 718971898235854848
Success
1184: 718939241951195136
Success
1185: 718631497683582976
Success
1186: 718613305783398402
Success
1187: 718540630683709445
Success
1188: 718460005985447936
Success
1189: 718454725339934721
Success
1190: 718246886998687744
Success
1191: 718234618122661888
Success
1192: 717841801130979328
Success
1193: 717790033953034240
Success
1194: 717537687239008257
Success
1195: 717428917016076293
Success
1196: 717421804990701568
```

```
Success
1197: 717047459982213120
Success
1198: 717009362452090881
Success
1199: 716802964044845056
Success
1200: 716791146589110272
Success
1201: 716730379797970944
Success
1202: 716447146686459905
Success
1203: 716439118184652801
Success
1204: 716285507865542656
Success
1205: 716080869887381504
Success
1206: 715928423106027520
Success
1207: 715758151270801409
Success
1208: 715733265223708672
Success
1209: 715704790270025728
Success
1210: 715696743237730304
Success
1211: 715680795826982913
Success
1212: 715360349751484417
Success
1213: 715342466308784130
Success
1214: 715220193576927233
Success
1215: 715200624753819648
Success
1216: 715009755312439296
Success
1217: 714982300363173890
Success
1218: 714962719905021952
Success
1219: 714957620017307648
Success
1220: 714631576617938945
```

```
Success
1221: 714606013974974464
Success
1222: 714485234495041536
Success
1223: 714258258790387713
Success
1224: 714251586676113411
Success
1225: 714214115368108032
Success
1226: 714141408463036416
Success
1227: 713919462244790272
Success
1228: 713909862279876608
Success
1229: 713900603437621249
Success
1230: 713761197720473600
Success
1231: 713411074226274305
Success
1232: 713177543487135744
Success
1233: 713175907180089344
Success
1234: 712809025985978368
Success
1235: 712717840512598017
Success
1236: 712668654853337088
Success
1237: 712438159032893441
Success
1238: 712309440758808576
Success
1239: 712097430750289920
Success
1240: 712092745624633345
Success
1241: 712085617388212225
Success
1242: 712065007010385924
Success
1243: 711998809858043904
Success
1244: 711968124745228288
```

```
Success
1245: 711743778164514816
Success
1246: 711732680602345472
Success
1247: 711694788429553666
Success
1248: 711652651650457602
Success
1249: 711363825979756544
Success
1250: 711306686208872448
Success
1251: 711008018775851008
Success
1252: 710997087345876993
Success
1253: 710844581445812225
Success
1254: 710833117892898816
Success
1255: 710658690886586372
Success
1256: 710609963652087808
Success
1257: 710588934686908417
Success
1258: 710296729921429505
Success
1259: 710283270106132480
Success
1260: 710272297844797440
Success
1261: 710269109699739648
Success
1262: 710153181850935296
Success
1263: 710140971284037632
Success
1264: 710117014656950272
Success
1265: 709918798883774466
Success
1266: 709901256215666688
Success
1267: 709852847387627521
Success
1268: 709566166965075968
```

```
Success
1269: 709556954897764353
Success
1270: 709519240576036864
Success
1271: 709449600415961088
Success
1272: 709409458133323776
Success
1273: 709225125749587968
Success
1274: 709207347839836162
Success
1275: 709198395643068416
Success
1276: 709179584944730112
Success
1277: 709158332880297985
Success
1278: 709042156699303936
Success
1279: 708853462201716736
Success
1280: 708845821941387268
Success
1281: 708834316713893888
Success
1282: 708810915978854401
Success
1283: 708738143638450176
Success
1284: 708711088997666817
Success
1285: 708479650088034305
Success
1286: 708469915515297792
Success
1287: 708400866336894977
Success
1288: 708356463048204288
Success
1289: 708349470027751425
Success
1290: 708149363256774660
Success
1291: 708130923141795840
Success
1292: 708119489313951744
```

```
Success
1293: 708109389455101952
Success
1294: 708026248782585858
Success
1295: 707995814724026368
Success
1296: 707983188426153984
Success
1297: 707969809498152960
Success
1298: 707776935007539200
Success
1299: 707741517457260545
Success
1300: 707738799544082433
Success
1301: 707693576495472641
Success
1302: 707629649552134146
Success
1303: 707610948723478529
Success
1304: 707420581654872064
Success
1305: 707411934438625280
Success
1306: 707387676719185920
Success
1307: 707377100785885184
Success
1308: 707315916783140866
Success
1309: 707297311098011648
Success
1310: 707059547140169728
Success
1311: 707038192327901184
Success
1312: 707021089608753152
Success
1313: 707014260413456384
Success
1314: 706904523814649856
Success
1315: 706901761596989440
Success
1316: 706681918348251136
```

```
Success
1317: 706644897839910912
Success
1318: 706593038911545345
Success
1319: 706538006853918722
Success
1320: 706516534877929472
Success
1321: 706346369204748288
Success
1322: 706310011488698368
Success
1323: 706291001778950144
Success
1324: 706265994973601792
Success
1325: 706169069255446529
Success
1326: 706166467411222528
Success
1327: 706153300320784384
Success
1328: 705975130514706432
Success
1329: 705970349788291072
Success
1330: 705898680587526145
Success
1331: 705786532653883392
Success
1332: 705591895322394625
Success
1333: 705475953783398401
Success
1334: 705442520700944385
Success
1335: 705428427625635840
Success
1336: 705239209544720384
Success
1337: 705223444686888960
Success
1338: 705102439679201280
Success
1339: 705066031337840642
Success
1340: 704871453724954624
```

```
Success
1341: 704859558691414016
Success
1342: 704847917308362754
Success
1343: 704819833553219584
Success
1344: 704761120771465216
Success
1345: 704499785726889984
Success
1346: 704491224099647488
Success
1347: 704480331685040129
Success
1348: 704364645503647744
Success
1349: 704347321748819968
Success
1350: 704134088924532736
Success
1351: 704113298707505153
Success
1352: 704054845121142784
Success
1353: 703774238772166656
Success
1354: 703769065844768768
Success
1355: 703631701117943808
Success
1356: 703611486317502464
Success
1357: 703425003149250560
Success
1358: 703407252292673536
Success
1359: 703382836347330562
Success
1360: 703356393781329922
Success
1361: 703268521220972544
Success
1362: 703079050210877440
Success
1363: 703041949650034688
Success
1364: 702932127499816960
```

```
Success
1365: 702899151802126337
Success
1366: 702684942141153280
Success
1367: 702671118226825216
Success
1368: 702598099714314240
Success
1369: 702539513671897089
Success
1370: 702332542343577600
Success
1371: 702321140488925184
Success
1372: 702276748847800320
Success
1373: 702217446468493312
Success
1374: 701981390485725185
Success
1375: 701952816642965504
Success
1376: 701889187134500865
Success
1377: 701805642395348998
Success
1378: 701601587219795968
Success
1379: 701570477911896070
Success
1380: 701545186879471618
Success
1381: 701214700881756160
Success
1382: 700890391244103680
Success
1383: 700864154249383937
Success
1384: 700847567345688576
Success
1385: 700796979434098688
Success
1386: 700747788515020802
Success
1387: 700518061187723268
Success
1388: 700505138482569216
```

```
Success
1389: 700462010979500032
Success
1390: 700167517596164096
Success
1391: 700151421916807169
Success
1392: 700143752053182464
Success
1393: 700062718104104960
Success
1394: 700029284593901568
Success
1395: 700002074055016451
Success
1396: 699801817392291840
Success
1397: 699788877217865730
Success
1398: 699779630832685056
Success
1399: 699775878809702401
Success
1400: 699691744225525762
Success
1401: 699446877801091073
Success
1402: 699434518667751424
Success
1403: 699423671849451520
Success
1404: 699413908797464576
Success
1405: 693370870310113280
Success
1406: 699323444782047232
Success
1407: 699088579889332224
Success
1408: 699079609774645248
Success
1409: 699072405256409088
Success
1410: 699060279947165696
Success
1411: 699036661657767936
Success
1412: 698989035503689728
```

```
Success
1413: 698953797952008193
Success
1414: 698907974262222848
Success
1415: 698710712454139905
Success
1416: 698703483621523456
Success
1417: 698635131305795584
Success
1418: 698549713696649216
Success
1419: 698355670425473025
Success
1420: 698342080612007937
Success
1421: 698262614669991936
Success
1422: 698195409219559425
Success
1423: 698178924120031232
Success
1424: 697995514407682048
Success
1425: 697990423684476929
Success
1426: 697943111201378304
Success
1427: 697881462549430272
Success
1428: 697630435728322560
Success
1429: 697616773278015490
Success
1430: 697596423848730625
Success
1431: 697575480820686848
Success
1432: 697516214579523584
Success
1433: 697482927769255936
Success
1434: 697463031882764288
Success
1435: 697270446429966336
Success
1436: 697259378236399616
```

```
Success
1437: 697255105972801536
Success
1438: 697242256848379904
Success
1439: 696900204696625153
Success
1440: 696894894812565505
Success
1441: 696886256886657024
Success
1442: 696877980375769088
Success
1443: 696754882863349760
Success
1444: 696744641916489729
Success
1445: 696713835009417216
Success
1446: 696518437233913856
Success
1447: 696490539101908992
Success
1448: 696488710901260288
Success
1449: 696405997980676096
Success
1450: 696100768806522880
Success
1451: 695816827381944320
Success
1452: 695794761660297217
Success
1453: 695767669421768709
Success
1454: 695629776980148225
Success
1455: 695446424020918272
Success
1456: 695409464418041856
Success
1457: 695314793360662529
Success
1458: 695095422348574720
Success
1459: 695074328191332352
Success
1460: 695064344191721472
```

```
Success
1461: 695051054296211456
Success
1462: 694925794720792577
Success
1463: 694905863685980160
Success
1464: 694669722378485760
Success
1465: 694356675654983680
Success
1466: 694352839993344000
Success
1467: 694342028726001664
Success
1468: 694329668942569472
Success
1469: 694206574471057408
Success
1470: 694183373896572928
Success
1471: 694001791655137281
Success
1472: 693993230313091072
Success
1473: 693942351086120961
Success
1474: 693647888581312512
Success
1475: 693644216740769793
Success
1476: 693642232151285760
Success
1477: 693629975228977152
Success
1478: 693622659251335168
Success
1479: 693590843962331137
Success
1480: 693582294167244802
Success
1481: 693486665285931008
Success
1482: 693280720173801472
Success
1483: 693267061318012928
Success
1484: 693262851218264065
```

```
Success
1485: 693231807727280129
Success
1486: 693155686491000832
Success
1487: 693109034023534592
Success
1488: 693095443459342336
Success
1489: 692919143163629568
Success
1490: 692905862751522816
Success
1491: 692901601640583168
Success
1492: 692894228850999298
Success
1493: 692828166163931137
Success
1494: 692752401762250755
Success
1495: 692568918515392513
Success
1496: 692535307825213440
Success
1497: 692530551048294401
Success
1498: 692423280028966913
Success
1499: 692417313023332352
Success
1500: 692187005137076224
Success
1501: 692158366030913536
Success
1502: 692142790915014657
Success
1503: 692041934689402880
Success
1504: 692017291282812928
Success
1505: 691820333922455552
Success
1506: 691793053716221953
Success
1507: 691756958957883396
Success
1508: 691675652215414786
```

```
Success
1509: 691483041324204033
Success
1510: 691459709405118465
Success
1511: 691444869282295808
Success
1512: 691416866452082688
Success
1513: 691321916024623104
Success
1514: 691096613310316544
Success
1515: 691090071332753408
Success
1516: 690989312272396288
Success
1517: 690959652130045952
Success
1518: 690938899477221376
Success
1519: 690932576555528194
Success
1520: 690735892932222976
Success
1521: 690728923253055490
Success
1522: 690690673629138944
Success
1523: 690649993829576704
Success
1524: 690607260360429569
Success
1525: 690597161306841088
Success
1526: 690400367696297985
Success
1527: 690374419777196032
Success
1528: 690360449368465409
Success
1529: 690348396616552449
Success
1530: 690248561355657216
Success
1531: 690021994562220032
Success
1532: 690015576308211712
```

```
Success
1533: 690005060500217858
Success
1534: 689999384604450816
Success
1535: 689993469801164801
Success
1536: 689977555533848577
Success
1537: 689905486972461056
Success
1538: 689877686181715968
Success
1539: 689835978131935233
Success
1540: 689661964914655233
Success
1541: 689659372465688576
Success
1542: 689623661272240129
Success
1543: 689599056876867584
Success
1544: 689557536375177216
Success
1545: 689517482558820352
Success
1546: 689289219123089408
Success
1547: 689283819090870273
Success
1548: 689280876073582592
Success
1549: 689275259254616065
Success
1550: 689255633275777024
Success
1551: 689154315265683456
Success
1552: 689143371370250240
Success
1553: 688916208532455424
Success
1554: 688908934925697024
Success
1555: 688898160958271489
Success
1556: 688894073864884227
```

```
Success
1557: 688828561667567616
Success
1558: 688804835492233216
Success
1559: 688789766343622656
Success
1560: 688547210804498433
Success
1561: 688519176466644993
Success
1562: 688385280030670848
Success
1563: 688211956440801280
Success
1564: 688179443353796608
Success
1565: 688116655151435777
Success
1566: 688064179421470721
Success
1567: 687841446767013888
Success
1568: 687826841265172480
Success
1569: 687818504314159109
Success
1570: 687807801670897665
Success
1571: 687732144991551489
Success
1572: 687704180304273409
Success
1573: 687664829264453632
Success
1574: 687494652870668288
Success
1575: 687480748861947905
Success
1576: 687476254459715584
Success
1577: 687460506001633280
Success
1578: 687399393394311168
Success
1579: 687317306314240000
Success
1580: 687312378585812992
```

```
Success
1581: 687127927494963200
Success
1582: 687124485711986689
Success
1583: 687109925361856513
Success
1584: 687102708889812993
Success
1585: 687096057537363968
Success
1586: 686947101016735744
Success
1587: 686760001961103360
Success
1588: 686749460672679938
Success
1589: 686730991906516992
Success
1590: 686683045143953408
Success
1591: 686618349602762752
Success
1592: 686606069955735556
Success
1593: 686394059078897668
Success
1594: 686386521809772549
Success
1595: 686377065986265092
Success
1596: 686358356425093120
Success
1597: 686286779679375361
Success
1598: 686050296934563840
Success
1599: 686035780142297088
Success
1600: 686034024800862208
Success
1601: 686007916130873345
Success
1602: 686003207160610816
Success
1603: 685973236358713344
Success
1604: 685943807276412928
```

```
Success
1605: 685906723014619143
Success
1606: 685681090388975616
Success
1607: 685667379192414208
Success
1608: 685663452032069632
Success
1609: 685641971164143616
Success
1610: 685547936038666240
Success
1611: 685532292383666176
Success
1612: 685325112850124800
Success
1613: 685321586178670592
Success
1614: 685315239903100929
Success
1615: 685307451701334016
Success
1616: 685268753634967552
Success
1617: 685198997565345792
Success
1618: 685169283572338688
Success
1619: 684969860808454144
Success
1620: 684959798585110529
Success
1621: 684940049151070208
Success
1622: 684926975086034944
Success
1623: 684914660081053696
Success
1624: 684902183876321280
Success
1625: 684880619965411328
Success
1626: 684830982659280897
Success
1627: 684800227459624960
Success
1628: 684594889858887680
```

```
Success
1629: 684588130326986752
Success
1630: 684567543613382656
Success
1631: 684538444857667585
Success
1632: 684481074559381504
Success
1633: 684460069371654144
Success
1634: 684241637099323392
Success
1635: 684225744407494656
Success
1636: 684222868335505415
Success
1637: 684200372118904832
Success
1638: 684195085588783105
Success
1639: 684188786104872960
Success
1640: 684177701129875456
Success
1641: 684147889187209216
Success
1642: 684122891630342144
Success
1643: 684097758874210310
Success
1644: 683857920510050305
Success
1645: 683852578183077888
Success
1646: 683849932751646720
Success
1647: 683834909291606017
Success
1648: 683828599284170753
Success
1649: 683773439333797890
Success
1650: 683742671509258241
Success
1651: 683515932363329536
Success
1652: 683498322573824003
```

```
Success
1653: 683481228088049664
Success
1654: 683462770029932544
Success
1655: 683449695444799489
Success
1656: 683391852557561860
Success
1657: 683357973142474752
Success
1658: 683142553609318400
Success
1659: 683111407806746624
Success
1660: 683098815881154561
Success
1661: 683078886620553216
Success
1662: 683030066213818368
Success
1663: 682962037429899265
Success
1664: 682808988178739200
Success
1665: 682788441537560576
Success
1666: 682750546109968385
Success
1667: 682697186228989953
Success
1668: 682662431982772225
Success
1669: 682638830361513985
Success
1670: 682429480204398592
Success
1671: 682406705142087680
Success
1672: 682393905736888321
Success
1673: 682389078323662849
Success
1674: 682303737705140231
Success
1675: 682259524040966145
Success
1676: 682242692827447297
```

```
Success
1677: 682088079302213632
Success
1678: 682059653698686977
Success
1679: 682047327939461121
Success
1680: 682032003584274432
Success
1681: 682003177596559360
Success
1682: 681981167097122816
Success
1683: 681891461017812993
Success
1684: 681694085539872773
Success
1685: 681679526984871937
Success
1686: 681654059175129088
Success
1687: 681610798867845120
Success
1688: 681579835668455424
Success
1689: 681523177663676416
Success
1690: 681340665377193984
Success
1691: 681339448655802368
Success
1692: 681320187870711809
Success
1693: 681302363064414209
Success
1694: 681297372102656000
Success
1695: 681281657291280384
Success
1696: 681261549936340994
Success
1697: 681242418453299201
Success
1698: 681231109724700672
Success
1699: 681193455364796417
Success
1700: 680970795137544192
```

```
Success
1701: 680959110691590145
Success
1702: 680940246314430465
Success
1703: 680934982542561280
Success
1704: 680913438424612864
Success
1705: 680889648562991104
Success
1706: 680836378243002368
Success
1707: 680805554198020098
Success
1708: 680801747103793152
Success
1709: 680798457301471234
Success
1710: 680609293079592961
Success
1711: 680583894916304897
Success
1712: 680497766108381184
Success
1713: 680494726643068929
Success
1714: 680473011644985345
Success
1715: 680440374763077632
Success
1716: 680221482581123072
Success
1717: 680206703334408192
Success
1718: 680191257256136705
Success
1719: 680176173301628928
Success
1720: 680161097740095489
Success
1721: 680145970311643136
Success
1722: 680130881361686529
Success
1723: 680115823365742593
Success
1724: 680100725817409536
```

```
Success
1725: 680085611152338944
Success
1726: 680070545539371008
Success
1727: 680055455951884288
Fail
1728: 679877062409191424
Success
1729: 679872969355714560
Success
1730: 679862121895714818
Success
1731: 679854723806179328
Success
1732: 679844490799091713
Success
1733: 679828447187857408
Success
1734: 679777920601223168
Success
1735: 679736210798047232
Success
1736: 679729593985699840
Success
1737: 679722016581222400
Success
1738: 679530280114372609
Success
1739: 679527802031484928
Success
1740: 679511351870550016
Success
1741: 679503373272485890
Success
1742: 679475951516934144
Success
1743: 679462823135686656
Success
1744: 679405845277462528
Success
1745: 679158373988876288
Success
1746: 679148763231985668
Success
1747: 679132435750195208
Success
1748: 679111216690831360
```

```
Success
1749: 6790062614270468097
Success
1750: 6790047485189439488
Success
1751: 6790001094530465792
Success
1752: 6789991772295516161
Success
1753: 6789969228704284672
Success
1754: 6788800283649069056
Success
1755: 6787798276842360832
Success
1756: 6787774928607469569
Success
1757: 6787767140346941444
Success
1758: 6787764513869611008
Success
1759: 6787755239630127104
Success
1760: 6787400035362037760
Success
1761: 6787708137298427904
Success
1762: 6786675843183484930
Success
1763: 6786643457146150913
Success
1764: 6784461515704427904
Success
1765: 6784243121063936100
Success
1766: 6784102103152476160
Success
1767: 6783996521993093120
Success
1768: 6783967962599751680
Success
1769: 6783890286144880640
Success
1770: 6783802368625786880
Success
1771: 6783410753759477760
Success
1772: 6783344973608591360
```

```
Success
1773: 678278586130948096
Success
1774: 678255464182861824
Success
1775: 680023323247357953
Success
1776: 678021115718029313
Success
1777: 677961670166224897
Success
1778: 677918531514703872
Success
1779: 677895101218201600
Success
1780: 677716515794329600
Success
1781: 677700003327029250
Success
1782: 677698403548192770
Success
1783: 677687604918272002
Success
1784: 677673981332312066
Success
1785: 677662372920729601
Success
1786: 677644091929329666
Success
1787: 677573743309385728
Success
1788: 677565715327688705
Success
1789: 677557565589463040
Success
1790: 677547928504967168
Success
1791: 677530072887205888
Success
1792: 677335745548390400
Success
1793: 677334615166730240
Success
1794: 677331501395156992
Success
1795: 677328882937298944
Success
1796: 677314812125323265
```

```
Success
1797:  6773010331697888928
Success
1798:  677269281705472000
Success
1799:  677228873407442944
Success
1800:  6771187300187611136
Success
1801:  676975532580409345
Success
1802:  676957860086095872
Success
1803:  676949632774234114
Success
1804:  676948236477857792
Success
1805:  676946864479084545
Success
1806:  676942428000112642
Success
1807:  676936541936185344
Success
1808:  676916996760600576
Success
1809:  676897532954456065
Success
1810:  676864501615042560
Success
1811:  676821958043033607
Success
1812:  676819651066732545
Success
1813:  676811746707918848
Success
1814:  676776431406465024
Success
1815:  676617503762681856
Success
1816:  676613908052996102
Success
1817:  676606785097199616
Success
1818:  676603393314578432
Success
1819:  676593408224403456
Success
1820:  676590572941893632
```

```
Success
1821: 676588346097852417
Success
1822: 676582956622721024
Success
1823: 676575501977128964
Success
1824: 676533798876651520
Success
1825: 676496375194980353
Success
1826: 676470639084101634
Success
1827: 676440007570247681
Success
1828: 676430933382295552
Success
1829: 676263575653122048
Success
1830: 676237365392908289
Success
1831: 676219687039057920
Success
1832: 676215927814406144
Success
1833: 676191832485810177
Success
1834: 676146341966438401
Success
1835: 676121918416756736
Success
1836: 676101918813499392
Success
1837: 676098748976615425
Success
1838: 676089483918516224
Success
1839: 675898130735476737
Success
1840: 675891555769696257
Success
1841: 675888385639251968
Success
1842: 675878199931371520
Success
1843: 675870721063669760
Success
1844: 675853064436391936
```

```
Success
1845: 6758490184447167488
Success
1846: 6758456573554215424
Success
1847: 6758227674735051008
Success
1848: 6758209296667219457
Success
1849: 6757984427703122432
Success
1850: 6757815629965868544
Success
1851: 6757403607753160193
Success
1852: 6757108909956750848
Success
1853: 6757073302066547968
Success
1854: 6757066394771788032
Success
1855: 6755344944439489536
Success
1856: 6755314759455709568
Success
1857: 6755224035882218240
Success
1858: 6755178289099424640
Success
1859: 6755010759657489664
Success
1860: 6754971033222386432
Success
1861: 6754899716171296384
Success
1862: 6754834309022214656
Success
1863: 6754327465175426176
Success
1864: 6753722404484454658
Success
1865: 6753626097392066656
Success
1866: 6753544359215775936
Success
1867: 6753493843395420016
Success
1868: 6753340601563013112
```

```
Success
1869: 675166823650848770
Success
1870: 675153376133427200
Success
1871: 675149409102012420
Success
1872: 675147105808306176
Success
1873: 675146535592706048
Success
1874: 675145476954566656
Success
1875: 675135153782571009
Success
1876: 675113801096802304
Success
1877: 675111688094527488
Success
1878: 675109292475830276
Success
1879: 675047298674663426
Success
1880: 675015141583413248
Success
1881: 675006312288268288
Success
1882: 675003128568291329
Success
1883: 674999807681908736
Success
1884: 674805413498527744
Success
1885: 674800520222154752
Success
1886: 674793399141146624
Success
1887: 674790488185167872
Success
1888: 674788554665512960
Success
1889: 674781762103414784
Success
1890: 674774481756377088
Success
1891: 674767892831932416
Success
1892: 674764817387900928
```

```
Success
1893: 674754018082705410
Success
1894: 674752233200820224
Success
1895: 674743008475090944
Success
1896: 674742531037511680
Success
1897: 674739953134403584
Success
1898: 674737130913071104
Success
1899: 674690135443775488
Success
1900: 674670581682434048
Success
1901: 674664755118911488
Success
1902: 674646392044941312
Success
1903: 674644256330530816
Success
1904: 674638615994089473
Success
1905: 674632714662858753
Success
1906: 674606911342424069
Success
1907: 674468880899788800
Success
1908: 674447403907457024
Success
1909: 674436901579923456
Success
1910: 674422304705744896
Success
1911: 674416750885273600
Success
1912: 674410619106390016
Success
1913: 674394782723014656
Success
1914: 674372068062928900
Success
1915: 674330906434379776
Success
1916: 674318007229923329
```

```
Success
1917: 674307341513269249
Success
1918: 674291837063053312
Success
1919: 674271431610523648
Success
1920: 674269164442398721
Success
1921: 674265582246694913
Success
1922: 674262580978937856
Success
1923: 674255168825880576
Success
1924: 674082852460433408
Success
1925: 674075285688614912
Success
1926: 674063288070742018
Success
1927: 674053186244734976
Success
1928: 674051556661161984
Success
1929: 674045139690631169
Success
1930: 674042553264685056
Success
1931: 674038233588723717
Success
1932: 674036086168010753
Success
1933: 674024893172875264
Success
1934: 674019345211760640
Success
1935: 674014384960745472
Success
1936: 674008982932058114
Success
1937: 673956914389192708
Success
1938: 673919437611909120
Success
1939: 673906403526995968
Success
1940: 673887867907739649
```

```
Success
1941: 673716320723169284
Success
1942: 673715861853720576
Success
1943: 673711475735838725
Success
1944: 673709992831262724
Success
1945: 673708611235921920
Success
1946: 673707060090052608
Success
1947: 673705679337693185
Success
1948: 673700254269775872
Success
1949: 673697980713705472
Success
1950: 673689733134946305
Success
1951: 673688752737402881
Success
1952: 673686845050527744
Success
1953: 673680198160809984
Success
1954: 673662677122719744
Success
1955: 673656262056419329
Success
1956: 673636718965334016
Success
1957: 673612854080196609
Success
1958: 673583129559498752
Success
1959: 673580926094458881
Success
1960: 673576835670777856
Success
1961: 673363615379013632
Success
1962: 673359818736984064
Success
1963: 673355879178194945
Success
1964: 673352124999274496
```

```
Success
1965: 673350198937153538
Success
1966: 673345638550134785
Success
1967: 673343217010679808
Success
1968: 673342308415348736
Success
1969: 673320132811366400
Success
1970: 673317986296586240
Success
1971: 673295268553605120
Success
1972: 673270968295534593
Success
1973: 673240798075449344
Success
1974: 673213039743795200
Success
1975: 673148804208660480
Success
1976: 672997845381865473
Success
1977: 672995267319328768
Success
1978: 672988786805112832
Success
1979: 672984142909456390
Success
1980: 672980819271634944
Success
1981: 672975131468300288
Success
1982: 672970152493887488
Success
1983: 672968025906282496
Success
1984: 672964561327235073
Success
1985: 672902681409806336
Success
1986: 672898206762672129
Success
1987: 672884426393653248
Success
1988: 672877615439593473
```

```
Success
1989: 672834301050937345
Success
1990: 672828477930868736
Success
1991: 672640509974827008
Success
1992: 672622327801233409
Success
1993: 672614745925664768
Success
1994: 672609152938721280
Success
1995: 672604026190569472
Success
1996: 672594978741354496
Success
1997: 672591762242805761
Success
1998: 672591271085670400
Success
1999: 672538107540070400
Success
2000: 672523490734551040
Success
2001: 672488522314567680
Success
2002: 672482722825261057
Success
2003: 672481316919734272
Success
2004: 672475084225949696
Success
2005: 672466075045466113
Success
2006: 672272411274932228
Success
2007: 672267570918129665
Success
2008: 672264251789176834
Success
2009: 672256522047614977
Success
2010: 672254177670729728
Success
2011: 672248013293752320
Success
2012: 672245253877968896
```

```
Success
2013: 672239279297454080
Success
2014: 672231046314901505
Success
2015: 672222792075620352
Success
2016: 672205392827572224
Success
2017: 672169685991993344
Success
2018: 672160042234327040
Success
2019: 672139350159835138
Success
2020: 672125275208069120
Success
2021: 672095186491711488
Success
2022: 672082170312290304
Success
2023: 672068090318987265
Success
2024: 671896809300709376
Success
2025: 671891728106971137
Success
2026: 671882082306625538
Success
2027: 671879137494245376
Success
2028: 671874878652489728
Success
2029: 671866342182637568
Success
2030: 671855973984772097
Success
2031: 671789708968640512
Success
2032: 671768281401958400
Success
2033: 671763349865160704
Success
2034: 671744970634719232
Success
2035: 671743150407421952
Success
2036: 671735591348891648
```

```
Success
2037: 671729906628341761
Success
2038: 671561002136281088
Success
2039: 671550332464455680
Success
2040: 671547767500775424
Success
2041: 671544874165002241
Success
2042: 671542985629241344
Success
2043: 671538301157904385
Success
2044: 671536543010570240
Success
2045: 671533943490011136
Success
2046: 671528761649688577
Success
2047: 671520732782923777
Success
2048: 671518598289059840
Success
2049: 671511350426865664
Success
2050: 671504605491109889
Success
2051: 671497587707535361
Success
2052: 671488513339211776
Success
2053: 671486386088865792
Success
2054: 671485057807351808
Success
2055: 671390180817915904
Success
2056: 671362598324076544
Success
2057: 671357843010908160
Success
2058: 671355857343524864
Success
2059: 671347597085433856
Success
2060: 671186162933985280
```

```
Success
2061: 671182547775299584
Success
2062: 671166507850801152
Success
2063: 671163268581498880
Success
2064: 671159727754231808
Success
2065: 671154572044468225
Success
2066: 671151324042559489
Success
2067: 671147085991960577
Success
2068: 671141549288370177
Success
2069: 671138694582165504
Success
2070: 671134062904504320
Success
2071: 671122204919246848
Success
2072: 671115716440031232
Success
2073: 671109016219725825
Success
2074: 670995969505435648
Success
2075: 670842764863651840
Success
2076: 670840546554966016
Success
2077: 670838202509447168
Success
2078: 670833812859932673
Success
2079: 670832455012716544
Success
2080: 670826280409919488
Success
2081: 670823764196741120
Success
2082: 670822709593571328
Success
2083: 670815497391357952
Success
2084: 670811965569282048
```

```
Success
2085: 670807719151067136
Success
2086: 670804601705242624
Success
2087: 670803562457407488
Success
2088: 670797304698376195
Success
2089: 670792680469889025
Success
2090: 670789397210615808
Success
2091: 670786190031921152
Success
2092: 670783437142401025
Success
2093: 670782429121134593
Success
2094: 670780561024270336
Success
2095: 670778058496974848
Success
2096: 670764103623966721
Success
2097: 670755717859713024
Success
2098: 670733412878163972
Success
2099: 670727704916926465
Success
2100: 670717338665226240
Success
2101: 670704688707301377
Success
2102: 670691627984359425
Success
2103: 670679630144274432
Success
2104: 670676092097810432
Success
2105: 670668383499735048
Success
2106: 670474236058800128
Success
2107: 670468609693655041
Success
2108: 670465786746662913
```

```
Success
2109: 670452855871037440
Success
2110: 670449342516494336
Success
2111: 670444955656130560
Success
2112: 670442337873600512
Success
2113: 670435821946826752
Success
2114: 670434127938719744
Success
2115: 670433248821026816
Success
2116: 670428280563085312
Success
2117: 670427002554466305
Success
2118: 670421925039075328
Success
2119: 670420569653809152
Success
2120: 670417414769758208
Success
2121: 670411370698022913
Success
2122: 670408998013820928
Success
2123: 670403879788544000
Success
2124: 670385711116361728
Success
2125: 670374371102445568
Success
2126: 670361874861563904
Success
2127: 670338931251150849
Success
2128: 670319130621435904
Success
2129: 670303360680108032
Success
2130: 670290420111441920
Success
2131: 670093938074779648
Success
2132: 670086499208155136
```

```
Success
2133: 670079681849372674
Success
2134: 670073503555706880
Success
2135: 670069087419133954
Success
2136: 670061506722140161
Success
2137: 670055038660800512
Success
2138: 670046952931721218
Success
2139: 670040295598354432
Success
2140: 670037189829525505
Success
2141: 670003130994700288
Success
2142: 669993076832759809
Success
2143: 669972011175813120
Success
2144: 669970042633789440
Success
2145: 669942763794931712
Success
2146: 669926384437997569
Success
2147: 669923323644657664
Success
2148: 669753178989142016
Success
2149: 669749430875258880
Success
2150: 669684865554620416
Success
2151: 669683899023405056
Success
2152: 669682095984410625
Success
2153: 669680153564442624
Success
2154: 669661792646373376
Success
2155: 669625907762618368
Success
2156: 669603084620980224
```

```
Success
2157: 669597912108789760
Success
2158: 669583744538451968
Success
2159: 669573570759163904
Success
2160: 669571471778410496
Success
2161: 669567591774625800
Success
2162: 669564461267722241
Success
2163: 669393256313184256
Success
2164: 669375718304980992
Success
2165: 669371483794317312
Success
2166: 669367896104181761
Success
2167: 669363888236994561
Success
2168: 669359674819481600
Success
2169: 669354382627049472
Success
2170: 669353438988365824
Success
2171: 669351434509529089
Success
2172: 669328503091937280
Success
2173: 669327207240699904
Success
2174: 669324657376567296
Success
2175: 669216679721873412
Success
2176: 669214165781868544
Success
2177: 669203728096960512
Success
2178: 669037058363662336
Success
2179: 669015743032369152
Success
2180: 669006782128353280
```

```
Success
2181: 669000397445533696
Success
2182: 668994913074286592
Success
2183: 668992363537309700
Success
2184: 668989615043424256
Success
2185: 668988183816871936
Success
2186: 668986018524233728
Success
2187: 668981893510119424
Success
2188: 668979806671884288
Success
2189: 668975677807423489
Success
2190: 668967877119254528
Success
2191: 668960084974809088
Success
2192: 668955713004314625
Success
2193: 668932921458302977
Success
2194: 668902994700836864
Success
2195: 668892474547511297
Success
2196: 668872652652679168
Success
2197: 668852170888998912
Success
2198: 668826086256599040
Success
2199: 668815180734689280
Success
2200: 668779399630725120
Success
2201: 668655139528511488
Success
2202: 668645506898350081
Success
2203: 668643542311546881
Success
2204: 668641109086707712
```

```
Success
2205:  668636665813057536
Success
2206:  668633411083464705
Success
2207:  668631377374486528
Success
2208:  668627278264475648
Success
2209:  668625577880875008
Success
2210:  668623201287675904
Success
2211:  668620235289837568
Success
2212:  668614819948453888
Success
2213:  668587383441514497
Success
2214:  668567822092664832
Success
2215:  668544745690562560
Success
2216:  668542336805281792
Success
2217:  668537837512433665
Success
2218:  668528771708952576
Success
2219:  668507509523615744
Success
2220:  668496999348633600
Success
2221:  668484198282485761
Success
2222:  668480044826800133
Success
2223:  668466899341221888
Success
2224:  668297328638447616
Success
2225:  668291999406125056
Success
2226:  668286279830867968
Success
2227:  668274247790391296
Success
2228:  668268907921326080
```

```
Success
2229: 668256321989451776
Success
2230: 668248472370458624
Success
2231: 668237644992782336
Success
2232: 668226093875376128
Success
2233: 668221241640230912
Success
2234: 668204964695683073
Success
2235: 668190681446379520
Success
2236: 668171859951755264
Success
2237: 668154635664932864
Success
2238: 668142349051129856
Success
2239: 668113020489474048
Success
2240: 667937095915278337
Success
2241: 667924896115245057
Success
2242: 667915453470232577
Success
2243: 667911425562669056
Success
2244: 667902449697558528
Success
2245: 667886921285246976
Success
2246: 667885044254572545
Success
2247: 667878741721415682
Success
2248: 667873844930215936
Success
2249: 667866724293877760
Success
2250: 667861340749471744
Success
2251: 667832474953625600
Success
2252: 667806454573760512
```

```
Success
2253: 667801013445750784
Success
2254: 667793409583771648
Success
2255: 667782464991965184
Success
2256: 667773195014021121
Success
2257: 667766675769573376
Success
2258: 667728196545200128
Success
2259: 667724302356258817
Success
2260: 667550904950915073
Success
2261: 667550882905632768
Success
2262: 667549055577362432
Success
2263: 667546741521195010
Success
2264: 667544320556335104
Success
2265: 667538891197542400
Success
2266: 667534815156183040
Success
2267: 667530908589760512
Success
2268: 667524857454854144
Success
2269: 667517642048163840
Success
2270: 667509364010450944
Success
2271: 667502640335572993
Success
2272: 667495797102141441
Success
2273: 667491009379606528
Success
2274: 667470559035432960
Success
2275: 667455448082227200
Success
2276: 667453023279554560
```

```
Success
2277: 667443425659232256
Success
2278: 667437278097252352
Success
2279: 667435689202614272
Success
2280: 667405339315146752
Success
2281: 667393430834667520
Success
2282: 667369227918143488
Success
2283: 667211855547486208
Success
2284: 667200525029539841
Success
2285: 667192066997374976
Success
2286: 667188689915760640
Success
2287: 667182792070062081
Success
2288: 667177989038297088
Success
2289: 667176164155375616
Success
2290: 667174963120574464
Success
2291: 667171260800061440
Success
2292: 667165590075940865
Success
2293: 667160273090932737
Success
2294: 667152164079423490
Success
2295: 667138269671505920
Success
2296: 667119796878725120
Success
2297: 667090893657276420
Success
2298: 667073648344346624
Success
2299: 667070482143944705
Success
2300: 667065535570550784
```

```
Success
2301: 667062181243039745
Success
2302: 667044094246576128
Success
2303: 667012601033924608
Success
2304: 666996132027977728
Success
2305: 666983947667116034
Success
2306: 666837028449972224
Success
2307: 666835007768551424
Success
2308: 666826780179869698
Success
2309: 666817836334096384
Success
2310: 666804364988780544
Success
2311: 666786068205871104
Success
2312: 666781792255496192
Success
2313: 666776908487630848
Success
2314: 666739327293083650
Success
2315: 666701168228331520
Success
2316: 666691418707132416
Success
2317: 666649482315059201
Success
2318: 666644823164719104
Success
2319: 666454714377183233
Success
2320: 666447344410484738
Success
2321: 666437273139982337
Success
2322: 666435652385423360
Success
2323: 666430724426358785
Success
2324: 666428276349472768
```

```
Success
2325: 666421158376562688
Success
2326: 666418789513326592
Success
2327: 666411507551481857
Success
2328: 666407126856765440
Success
2329: 666396247373291520
Success
2330: 666373753744588802
Success
2331: 666362758909284353
Success
2332: 666353288456101888
Success
2333: 666345417576210432
Success
2334: 666337882303524864
Success
2335: 666293911632134144
Success
2336: 666287406224695296
Success
2337: 666273097616637952
Success
2338: 666268910803644416
Success
2339: 666104133288665088
Success
2340: 666102155909144576
Success
2341: 666099513787052032
Success
2342: 666094000022159362
Success
2343: 666082916733198337
Success
2344: 666073100786774016
Success
2345: 666071193221509120
Success
2346: 666063827256086533
Success
2347: 666058600524156928
Success
2348: 666057090499244032
```

```
Success
2349: 666055525042405380
Success
2350: 666051853826850816
Success
2351: 666050758794694657
Success
2352: 666049248165822465
Success
2353: 666044226329800704
Success
2354: 666033412701032449
Success
2355: 666029285002620928
Success
2356: 666020888022790149
Success
1939.1978713939989
{888202515573088257: TweepError([{'code': 144, 'message': 'No status found with that ID.'}],), 8
```

```
In [14]: df_3 = pd.read_json('tweet_json.txt', lines=True)
```

```
In [15]: data.raise_for_status()
         open('tweet_json.txt', 'wb').write(data.content)
```

```
Out[15]: 335079
```

1.3   Assessing Data

In this section, detect and document at least **eight (8) quality issues and two (2) tidiness issue**. You must use **both** visual assessment programmatic assessement to assess the data.

   Note: pay attention to the following key points when you access the data.

- You only want original ratings (no retweets) that have images. Though there are 5000+ tweets in the dataset, not all are dog ratings and some are retweets.
- Assessing and cleaning the entire dataset completely would require a lot of time, and is not necessary to practice and demonstrate your skills in data wrangling. Therefore, the requirements of this project are only to assess and clean at least 8 quality issues and at least 2 tidiness issues in this dataset.
- The fact that the rating numerators are greater than the denominators does not need to be cleaned. This unique rating system is a big part of the popularity of WeRateDogs.
- You do not need to gather the tweets beyond August 1st, 2017. You can, but note that you won't be able to gather the image predictions for these tweets since you don't have access to the algorithm used.

```
In [16]: #Assess the enhanced twitter archive data

         df_1.sample(5)
```

```
Out[16]:                 tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
       1893  674752233200820224                    NaN                  NaN
       442   819588359383371776                    NaN                  NaN
       2193  668902994700836864                    NaN                  NaN
       661   791026214425268224                    NaN                  NaN
       2248  667866724293877760                    NaN                  NaN


                        timestamp  \
       1893   2015-12-10 00:47:23 +0000
       442    2017-01-12 16:54:26 +0000
       2193   2015-11-23 21:24:36 +0000
       661    2016-10-25 21:18:40 +0000
       2248   2015-11-21 00:46:50 +0000


                                                                    sourc
       1893  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
       442   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
       2193  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
       661   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
       2248  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a


       1893                                    Everyone please just appreciate how perfect t
       442       This is Jazzy. She just found out that sandwich wasn't for her. Shocked and
       2193                 Meet Lambeau. He's a Whistling Haiku from the plains of southern G
       661                                        RT @dog_rates: This is Alfie. He's touching
       2248  This is Shaggy. He knows exactly how to solve the puzzle but can't talk. All he w


             retweeted_status_id  retweeted_status_user_id  \
       1893                  NaN                       NaN
       442                   NaN                       NaN
       2193                  NaN                       NaN
       661          7.638376e+17              4.196984e+09
       2248                  NaN                       NaN


           retweeted_status_timestamp  \
       1893                        NaN
       442                         NaN
       2193                        NaN
       661    2016-08-11 20:40:41 +0000
       2248                        NaN


       1893  https://twitter.com/dog_rates/status/674752233200820224/photo/1,https://twitter.c
       442                                                              https://twitter.c
       2193                                                             https://twitter.c
       661   https://twitter.com/dog_rates/status/763837565564780549/photo/1,https://twitter.c
       2248                                                             https://twitter.c
```

	rating_numerator	rating_denominator	name	doggo	floofer	pupper	puppo
1893	12	10	None	None	None	None	None
442	13	10	Jazzy	None	None	None	None
2193	11	10	Lambeau	None	None	None	None
661	11	10	Alfie	None	None	None	None
2248	10	10	Shaggy	None	None	None	None

In [17]: df_1.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 [18]: sum(df_1['tweet_id'].value_counts())

Out[18]: 2356

In [19]: # Assess the image prediction data

df_2.sample(5)

Out[19]:

	tweet_id	jpg_url \
1267	749403093750648834	https://pbs.twimg.com/media/CmZqIslWIAQFiqe.jpg
1868	844223788422217728	https://pbs.twimg.com/media/C7dJCnqU4AAswat.jpg
1924	857263160327368704	https://pbs.twimg.com/media/C-WcS4MXoAADrBU.jpg
1104	722613351520608256	https://pbs.twimg.com/media/Cgc9AjMVIAERdUA.jpg
504	675888385639251968	https://pbs.twimg.com/media/CWE85snWIAEG5ES.jpg

```
            img_num                          p1    p1_conf  p1_dog  \
      1267        1         Chesapeake_Bay_retriever  0.694541    True
      1868        1             Labrador_retriever  0.719510    True
      1924        1                       Samoyed  0.998021    True
      1104        1             Labrador_retriever  0.530915    True
      504         1   West_Highland_white_terrier  0.672117    True


                            p2    p2_conf  p2_dog                p3  \
      1267     curly-coated_retriever  0.076335    True  Irish_water_spaniel
      1868   Chesapeake_Bay_retriever  0.122019    True          Newfoundland
      1924               Pomeranian  0.000922    True              keeshond
      1104           golden_retriever  0.288230    True                  chow
      504        Old_English_sheepdog  0.146147    True              komondor


            p3_conf  p3_dog
      1267  0.048550    True
      1868  0.038828    True
      1924  0.000311    True
      1104  0.044854    True
      504   0.023141    True
```

In [20]: df_2.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 [21]: sum(df_2['tweet_id'].value_counts())

Out[21]: 2075

In [22]: # Assess the twitter API data

        df_3.sample(5)

```
Out[22]:       contributors  coordinates          created_at display_text_range  \
        1579           NaN          NaN 2016-01-09 03:24:40          [0, 109]
        421            NaN          NaN 2017-01-13 17:00:21          [0, 139]
        1950           NaN          NaN 2015-12-05 03:28:25          [0, 121]
        951            NaN          NaN 2016-07-04 15:00:45           [0, 54]
        646            NaN          NaN 2016-10-24 15:53:19          [0, 115]


        1579  {'hashtags': [], 'symbols': [], 'user_mentions': [{'screen_name': 'PDolan37', 'na
        421   {'hashtags': [], 'symbols': [], 'user_mentions': [], 'urls': [{'url': 'https://t.
        1950  {'hashtags': [], 'symbols': [], 'user_mentions': [], 'urls': [], 'media': [{'id':
        951   {'hashtags': [], 'symbols': [], 'user_mentions': [], 'urls': [], 'media': [{'id':
        646   {'hashtags': [], 'symbols': [], 'user_mentions': [], 'urls': [], 'media': [{'id':


        1579  {'media': [{'id': 685663358637486080, 'id_str': '685663358637486080', 'indices':
        421   {'media': [{'id': 819952225594261509, 'id_str': '819952225594261509', 'indices':
        1950  {'media': [{'id': 672980779580829696, 'id_str': '672980779580829696', 'indices':
        951   {'media': [{'id': 749850882955702272, 'id_str': '749850882955702272', 'indices':
        646   {'media': [{'id': 790581940780924928, 'id_str': '790581940780924928', 'indices':


              favorite_count  favorited  \
        1579            2962      False
        421             4997      False
        1950            1623      False
        951             4765      False
        646            19441      False


        1579                                           Meet Brooks. He's confused
        421   This is Oliver. He has dreams of being a service puppo so he can help his owner.
        1950                                          Extraordinary dog here. Looks large.
        951
        646                           This is Duke. He sneaks into the fridge sometimes. It'

              geo  \
        1579  NaN
        421   NaN
        1950  NaN
        951   NaN
        646   NaN


        1579
        421
        1950
        951
        646
```

```
             quoted_status   quoted_status_id quoted_status_id_str  \
      1579             NaN                NaN                  NaN
      421              NaN                NaN                  NaN
      1950             NaN                NaN                  NaN
      951              NaN                NaN                  NaN
      646              NaN                NaN                  NaN


             quoted_status_permalink   retweet_count   retweeted   retweeted_status  \
      1579                       NaN            1328       False                NaN
      421                        NaN            1077       False                NaN
      1950                       NaN             884       False                NaN
      951                        NaN            2263       False                NaN
      646                        NaN            6605       False                NaN


                                                                                   sour
      1579   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</
      421    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</
      1950   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</
      951    <a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">TweetDeck</
      646    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</


           truncated  \
      1579     False
      421      False
      1950     False
      951      False
      646      False



      1579  {'id': 4196983835, 'id_str': '4196983835', 'name': 'WeRateDogső', 'screen_name':
      421   {'id': 4196983835, 'id_str': '4196983835', 'name': 'WeRateDogső', 'screen_name':
      1950  {'id': 4196983835, 'id_str': '4196983835', 'name': 'WeRateDogső', 'screen_name':
      951   {'id': 4196983835, 'id_str': '4196983835', 'name': 'WeRateDogső', 'screen_name':
      646   {'id': 4196983835, 'id_str': '4196983835', 'name': 'WeRateDogső', 'screen_name':


      [5 rows x 32 columns]

In [23]: df_3.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2327 entries, 0 to 2326
Data columns (total 32 columns):
contributors               0 non-null float64
coordinates                0 non-null float64
created_at                 2327 non-null datetime64[ns]
display_text_range         2327 non-null object
entities                   2327 non-null object
```

```
extended_entities              2057 non-null object
favorite_count                 2327 non-null int64
favorited                      2327 non-null bool
full_text                      2327 non-null object
geo                            0 non-null float64
id                             2327 non-null int64
id_str                         2327 non-null int64
in_reply_to_screen_name        77 non-null object
in_reply_to_status_id          77 non-null float64
in_reply_to_status_id_str      77 non-null float64
in_reply_to_user_id            77 non-null float64
in_reply_to_user_id_str        77 non-null float64
is_quote_status                2327 non-null bool
lang                           2327 non-null object
place                          1 non-null object
possibly_sensitive             2195 non-null float64
possibly_sensitive_appealable  2195 non-null float64
quoted_status                  24 non-null object
quoted_status_id               26 non-null float64
quoted_status_id_str           26 non-null float64
quoted_status_permalink        26 non-null object
retweet_count                  2327 non-null int64
retweeted                      2327 non-null bool
retweeted_status               160 non-null object
source                         2327 non-null object
truncated                      2327 non-null bool
user                           2327 non-null object
dtypes: bool(4), datetime64[ns](1), float64(11), int64(4), object(12)
memory usage: 518.2+ KB
```

`In [24]: sum(df_3['id'].value_counts())`

`Out[24]: 2327`

1.4 Quality issues

`Twitter Image Prediction Table` 1. Column header (p1, p1_config, p1_dog, p2, p2_config, p2_dog, p3, p3_config, p3_dog) not descriptive enough. 2. Invalid data representation for percentage in p1_config, p2_config, p3_config

`Twitter Archive Enhanced Table` 3. 745 None values as dog name 4. Some dog names are not feasibly correct like 'a', 'an'. 5. Text column contains url apart from the text column 6. Time stamp is contians string instead of date time format 7. Rating denominator can't be zero

`Twitter API data Table` 8. Nondescriptive column header "id"

1.5 Tidiness issues

`Twitter Archieved Enhanced Table` 1. A single observation unit is stored in multiple column (doggo, floofer, pupper, puppo) shoild all be in dog stages. 2. Delete rows that are not original

tweets

1.6 Cleaning Data

In this section, clean **all** of the issues you documented while assessing.

Note: Make a copy of the original data before cleaning. Cleaning includes merging individual pieces of data according to the rules of tidy data. The result should be a high-quality and tidy master pandas DataFrame (or DataFrames, if appropriate).

```
In [25]:  # Make copies of original pieces of data
          df1_clean = df_1.copy()
          df2_clean = df_2.copy()
          df3_clean = df_3.copy()
```

1.7 Quality Issues

```
Twitter image prediction table
```

1.7.1 Issue #1: Column header (p1, p1_config, p1_dog, p2, p2_config, p2_dog, p3, p3_config, p3_dog) not descriptive enough.

Define: Rename nondescriptive column headers to be more readeble. "img_num, p1, p1_config, p1_dog, p2, p2_config, p2_dog, p3, p3_config, p3_dog" should be "img_number, 1st_prediction, 1st_confidence, 1st_DogBreed, 2nd_prediction, 2nd_confidence, 2nd_DogBreed, 3rd_prediction, 3rd_confidence, 3rd_DogBreed"

```
In [26]:  #View inappropriate column heads

          df_2.head(2)
```

```
Out[26]:               tweet_id                                        jpg_url  \
          0  666020888022790149  https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg
          1  666029285002620928  https://pbs.twimg.com/media/CT42GRgUYAA5iDo.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

              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
```

Code

```
In [27]:  #rename inappropriate column head

          df2_clean.rename(columns={'img_num':'img_number', 'p1':'1st_prediction', 'p1_conf':'1st
```

109

Test

```
In [28]: df2_clean.head(2)
```

```
Out[28]:              tweet_id                                         jpg_url  \
         0   666020888022790149   https://pbs.twimg.com/media/CT4udnOWwAAOaMy.jpg
         1   666029285002620928   https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg

            img_number          1st_prediction  1st_confidence  1st_DogBreed  \
         0           1   Welsh_springer_spaniel        0.465074          True
         1           1                  redbone        0.506826          True

              2nd_prediction  2nd_confidence  2nd_DogBreed       3rd_prediction  \
         0             collie        0.156665          True    Shetland_sheepdog
         1  miniature_pinscher        0.074192          True  Rhodesian_ridgeback

            3rd_confidence  3rd_DogBreed
         0        0.061428          True
         1        0.072010          True
```

1.7.2 Issue #2: Invalid data representation for percentage in p1_conf, p2_conf, p3_conf

Define: Reformat each prediction's confidence level (1st, 2nd, 3rd) to be represented in percentage.

```
In [29]: #View inappropriate confidence level format

         df_2[['p1_conf', 'p2_conf', 'p3_conf']].head(2)
```

```
Out[29]:     p1_conf   p2_conf   p3_conf
         0  0.465074  0.156665  0.061428
         1  0.506826  0.074192  0.072010
```

Code

```
In [30]: #correct confidence level format

         df2_clean['1st_confidence'] = df2_clean['1st_confidence'].map(lambda n: '{:,.2%}'.forma
         df2_clean['2nd_confidence'] = df2_clean['2nd_confidence'].map(lambda n: '{:,.2%}'.forma
         df2_clean['3rd_confidence'] = df2_clean['3rd_confidence'].map(lambda n: '{:,.2%}'.forma
```

Test

```
In [31]: df2_clean.sample(3)
```

```
Out[31]:               tweet_id                                         jpg_url  \
         391   673576835670777856   https://pbs.twimg.com/media/CVkGjsxU8AA5OYX.jpg
         531   676821958043033607   https://pbs.twimg.com/media/CWSN-vaXAAA8Ehr.jpg
         64    667171260800061440   https://pbs.twimg.com/media/CUJEuRIXIAAPDLt.jpg
```

```
        img_number    1st_prediction 1st_confidence  1st_DogBreed  \
391               1              teddy         25.52%         False
531               2      Great_Pyrenees         86.98%          True
64                1     giant_schnauzer        84.13%          True


              2nd_prediction 2nd_confidence  2nd_DogBreed        3rd_prediction  \
391   Christmas_stocking              9.83%         False                pajama
531               kuvasz              7.98%          True       standard_poodle
64        Lakeland_terrier            5.27%          True   Irish_water_spaniel


        3rd_confidence  3rd_DogBreed
391              7.27%         False
531              1.33%          True
64               3.44%          True

Twitter Archive Enhanced Table
```

1.7.3   Issue #3: 745 None values as dog name

Define: A significant number of dogs (745) has None as their name in the data

```
In [32]: # Count the number of names with None value
         df_1['name'].value_counts()

Out[32]: None       745
         a           55
         Charlie     12
         Oliver      11
         Lucy        11
         Cooper      11
         Penny       10
         Tucker      10
         Lola        10
         Bo           9
         Winston      9
         Sadie        8
         the          8
         Bailey       7
         Toby         7
         an           7
         Buddy        7
         Daisy        7
         Koda         6
         Leo          6
         Jack         6
         Dave         6
         Milo         6
         Oscar        6
```

```
      Stanley      6
      Rusty        6
      Bella        6
      Scout        6
      Jax          6
      Alfie        5
                 ...
      Jordy        1
      Kota         1
      Raphael      1
      Gert         1
      Antony       1
      by           1
      Strider      1
      Rose         1
      Kevon        1
      JD           1
      Lillie       1
      Cannon       1
      Ridley       1
      Tess         1
      Lizzie       1
      Brockly      1
      Brat         1
      Nugget       1
      Rey          1
      Sid          1
      Marlee       1
      Peanut       1
      Stuart       1
      Cora         1
      Josep        1
      Akumi        1
      Traviss      1
      Asher        1
      Roscoe       1
      Ralf         1
      Name: name, Length: 957, dtype: int64
```

Code   No programmatic solution for this issue as we can't make up names afor the dogs. We can only observe and take note of them as they can affect our analysis.

Test

```
In [33]: df1_clean_name_is_none = df_1.query('name == "None"')
         df1_clean_name_is_none[['tweet_id', 'name']].sample(3)

Out[33]:                tweet_id  name
```

```
234    847617282490613760   None
1832   676191832485810177   None
912    757596066325864448   None
```

1.7.4 Issue #4: Text column contains url apart from the text column

Define: Split the text colunmn into two with the keyword 'https' and delete the newly newly formed column.

```
In [34]: # View the text column containing url path
         df_1['text'].sample(5)
```

```
Out[34]: 2331                          Here we have a mixed Asiago from the Galápagos I
         166       Meet Milky. She has no idea what happened. Just as pupset as you. Perhaps a s
         2138
         702                              RT @dog_rates: This is Scout. He really wan
         276     Say hello to Maddie and Gunner. They are considerably pupset about bath time. B
         Name: text, dtype: object
```

Code

```
In [35]: #split the text column using https:\\

         df1_clean['newtext'] = df1_clean['text'].apply(lambda x: re.split('https:\/\/.*', str(x
```

```
In [36]: #drop column containing url path

         df1_clean = df1_clean.drop(columns='text')
```

1.7.5 Test

```
In [37]: df1_clean['newtext'].sample(5)
```

```
Out[37]: 380     Meet Tucker. It's his birthday. He's pupset with you because you're too busy pl
         1263                            This pupper got her hair chalked for her birthday.
         39                        This is Lola. It's her first time outside. Must test the e
         1032
         884                              Meet Sadie. She's addicted to balloons. It's tearing
         Name: newtext, dtype: object
```

1.7.6 Issue #5: Some dog names are not feasibly correct like 'a', 'an'.

Define: Not all the dog names are realistically correct. Here, I will be correct the names that have 'a' and 'an' in the column name.

```
In [38]: #Count the number of names that has 'a' and 'an' as value

         df1_clean['name'].value_counts()
```

```
Out[38]:  None      745
          a          55
          Charlie    12
          Oliver     11
          Lucy       11
          Cooper     11
          Penny      10
          Tucker     10
          Lola       10
          Bo          9
          Winston     9
          Sadie       8
          the         8
          Bailey      7
          Toby        7
          an          7
          Buddy       7
          Daisy       7
          Koda        6
          Leo         6
          Jack        6
          Dave        6
          Milo        6
          Oscar       6
          Stanley     6
          Rusty       6
          Bella       6
          Scout       6
          Jax         6
          Alfie       5
                     ...
          Jordy       1
          Kota        1
          Raphael     1
          Gert        1
          Antony      1
          by          1
          Strider     1
          Rose        1
          Kevon       1
          JD          1
          Lillie      1
          Cannon      1
          Ridley      1
          Tess        1
          Lizzie      1
          Brockly     1
          Brat        1
```

```
Nugget        1
Rey           1
Sid           1
Marlee        1
Peanut        1
Stuart        1
Cora          1
Josep         1
Akumi         1
Traviss       1
Asher         1
Roscoe        1
Ralf          1
Name: name, Length: 957, dtype: int64
```

Code

In [40]: *#select text that has 'a' and 'an' in them*

```python
name = ['a', 'an']
df1_name = df1_clean.loc[df1_clean['name'].isin(name)]
df1_name[['newtext', 'name']].sample(20)
```

Out[40]:

```
2128    This is a Speckled Cauliflower Yosemite named Hemry. He's terrified of intruder
1853  This is a Sizzlin Menorah spaniel from Brooklyn named Wylie. Lovable eyes. Chille
1017         This is a carrot. We only rate dogs. Please only send in dogs. You all real
1854                              Seriously guys?! Only send in dogs. I only rate dogs.
1955   This is a Lofted Aphrodisiac Terrier named Kip. Big fan of bed n breakfasts. Fit
2350  This is a truly beautiful English Wilson Staff retriever. Has a nice phone. Privi
2235             This is a Trans Siberian Kellogg named Alfonso. Huge ass eyeballs. Act
1785                               This is a dog swinging. I really enjoyed it I
2287        This is a Dasani Kingfisher from Maine. His name is Daryl. Daryl doesn't li
1368      This is a Wild Tuscan Poofwiggle. Careful not to startle. Rare tongue slip. C
2264                  This is a southwest Coriander named Klint. Hat looks expens
1207      This is a taco. We only rate dogs. Please only send in dogs. Dogs are what we
2352      This is a purebred Piers Morgan. Loves to Netflix and chill. Always looks like
2116   This is a Deciduous Trimester mix named Spork. Only 1 ear works. No seat belt. I
1340                       Here is a heartbreaking scene of an incredible pupper b
2218    This is a Birmingham Quagmire named Chuk. Loves to relax and watch the game whi
1877      C'mon guys. We've been over this. We only rate dogs. This is a cow. Please onl
2314    This is a golden Buckminsterfullerene named Johm. Drives trucks. Lumberjack (?).
2204    This is an Irish Rigatoni terrier named Berta. Completely made of rope. No eyes.
2146     This is a spotted Lipitor Rumpelstiltskin named Alphred. He can't wait for the

        name
2128       a
1853       a
```

```
      1017      a
      1854      a
      1955      a
      2350      a
      2235      a
      1785      a
      2287      a
      1368      a
      2264      a
      1207      a
      2352      a
      2116      a
      1340      a
      2218      a
      1877      a
      2314      a
      2204      an
      2146      a
```

In [42]: *#split text string using 'named' keyword*

```python
name = df1_name['newtext'].str.split('named ', n=1, expand=True)
```

In [43]: `named = name[1].str.split('.', n=1, expand=True)`

In [44]:
```python
df1_name['new_name'] = named[0]
df1_named = df1_name[['tweet_id', 'new_name']]
```

```
/opt/conda/lib/python3.6/site-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
  """Entry point for launching an IPython kernel.
```

In [45]: *#merge new name to the entire data*

```python
df1_cleaned = pd.merge(df1_clean, df1_named, on='tweet_id', how='left')
```

In [46]:
```python
df1_cleaned['correct_name'] = df1_cleaned['name']
df1_cleaned['correct_name'][(df1_cleaned['new_name'] != 'None') & (df1_cleaned['new_nam
```

```
/opt/conda/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
```

Test

```
In [47]: df1_cleaned[(df1_cleaned['new_name'] != 'None') & (df1_cleaned['new_name'].notna())][['
```

```
Out[47]:         name   correct_name        new_name
          1853    a          Wylie            Wylie
          1955    a            Kip              Kip
          2034    a   Jacob (Yacb)     Jacob (Yacb)
          2066    a          Rufus            Rufus
          2116    a          Spork            Spork
          2125    a       Cherokee         Cherokee
          2128    a          Hemry            Hemry
          2146    a        Alphred          Alphred
          2161    a        Alfredo          Alfredo
          2191    a          Leroi            Leroi
          2204   an          Berta            Berta
          2218    a           Chuk             Chuk
          2235    a        Alfonso          Alfonso
          2249    a         Cheryl           Cheryl
          2255    a        Jessiga          Jessiga
          2264    a          Klint            Klint
          2273    a           Kohl             Kohl
          2304    a           Pepe             Pepe
          2311    a      Octaviath        Octaviath
          2314    a           Johm             Johm
```

1.7.7 Issue #6: Time stamp is contians string instead of date time format

Define: Convert column format of 'timestamp' to datrtimestamp

```
In [48]: #view timestamp data type

         df_1.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
```

117

```
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
```

Code

In [49]: *#change timestamp to datetime format*

        df1_clean['timestamp'] = pd.to_datetime(df1_clean['timestamp'])

Test

In [50]: df1_clean.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 datetime64[ns]
source                       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
newtext                      2356 non-null object
dtypes: datetime64[ns](1), float64(4), int64(3), object(9)
memory usage: 313.0+ KB
```

1.7.8 Issue #7: Rating denominator can't be zero

Define: Review the text value corresponding to the denominator reading zero and extract the correct value for the numerator.

118

```
In [51]:  # view row that has 0 as denominator

          df_1.query('rating_denominator == 0')[['tweet_id', 'text', 'rating_denominator']]

Out[51]:                tweet_id  \
          313   835246439529840640


          313   @jonnysun @Lin_Manuel ok jomny I know you're excited but 960/00 isn't a valid rati

                rating_denominator
          313                    0
```

Code

```
In [52]:  #select row with value zero

          df1_sn = df_1.iloc[[313]]
          df1_sn

Out[52]:                tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
          313   835246439529840640           8.352460e+17           26259576.0

                               timestamp  \
          313   2017-02-24 21:54:03 +0000


                                                                                   source
          313   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>


          313   @jonnysun @Lin_Manuel ok jomny I know you're excited but 960/00 isn't a valid rati

                retweeted_status_id  retweeted_status_user_id retweeted_status_timestamp  \
          313                  NaN                       NaN                        NaN

                expanded_urls  rating_numerator  rating_denominator  name doggo floofer  \
          313            NaN               960                   0  None  None    None

                pupper puppo
          313   None  None
```

```
In [53]:  #split text with the '/' key to get the correct denominator

          dn = df1_sn['text'].str.split('/', n=1, expand=True)
          dn1 = dn[1].str.split('/', n=1, expand=True)
          dn1

Out[53]:                            0          1
          313   00 isn't a valid rating, 13   10 is tho
```

119

```
In [54]: df1_sn["denominator"]= dn1[1].str.extract(r'(\d+)', expand=True).astype(float)
         df1_sn = df1_sn[['tweet_id','denominator']]
         df1_sn
```

```
/opt/conda/lib/python3.6/site-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
  """Entry point for launching an IPython kernel.
```

```
Out[54]:              tweet_id  denominator
         313  835246439529840640         10.0
```

```
In [55]: #merge correct denominator into data

         df1_clean = pd.merge(df1_clean, df1_sn, on='tweet_id', how='left')
```

```
In [56]: df1_clean['correct_denominator'] = df1_clean['rating_denominator']
         df1_clean['correct_denominator'][df1_clean.denominator.notnull()] = df1_clean['denomina
```

```
/opt/conda/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
```

```
In [57]: df1_clean = df1_clean.drop(columns='denominator')
```

Test

```
In [59]: df1_clean.iloc[[313]][['newtext', 'correct_denominator']]
```

```
Out[59]:
         313  @jonnysun @Lin_Manuel ok jomny I know you're excited but 960/00 isn't a valid rati

              correct_denominator
         313                   10
```

Twitter API data table

1.7.9   Issue #8: Nondescriptive column header "id"

Define: Rename the column name 'id' to 'tweet_id' to make the column descriptive and the data relational.

```
In [60]: #view all column heads for non descriptive entry

         df_3.info()
```

120

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2327 entries, 0 to 2326
Data columns (total 32 columns):
contributors                  0 non-null float64
coordinates                   0 non-null float64
created_at                    2327 non-null datetime64[ns]
display_text_range            2327 non-null object
entities                      2327 non-null object
extended_entities             2057 non-null object
favorite_count                2327 non-null int64
favorited                     2327 non-null bool
full_text                     2327 non-null object
geo                           0 non-null float64
id                            2327 non-null int64
id_str                        2327 non-null int64
in_reply_to_screen_name       77 non-null object
in_reply_to_status_id         77 non-null float64
in_reply_to_status_id_str     77 non-null float64
in_reply_to_user_id           77 non-null float64
in_reply_to_user_id_str       77 non-null float64
is_quote_status               2327 non-null bool
lang                          2327 non-null object
place                         1 non-null object
possibly_sensitive            2195 non-null float64
possibly_sensitive_appealable 2195 non-null float64
quoted_status                 24 non-null object
quoted_status_id              26 non-null float64
quoted_status_id_str          26 non-null float64
quoted_status_permalink       26 non-null object
retweet_count                 2327 non-null int64
retweeted                     2327 non-null bool
retweeted_status              160 non-null object
source                        2327 non-null object
truncated                     2327 non-null bool
user                          2327 non-null object
dtypes: bool(4), datetime64[ns](1), float64(11), int64(4), object(12)
memory usage: 518.2+ KB
```

Code

```python
In [61]: # rename column head 'id' into 'tweet_id'

        df3_clean.rename(columns={'id':'tweet_id'}, inplace=True)

In [62]: df3_clean = df3_clean[['tweet_id', 'favorite_count', 'retweet_count']]
```

Test

```
In [63]: df3_clean.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2327 entries, 0 to 2326
Data columns (total 3 columns):
tweet_id          2327 non-null int64
favorite_count    2327 non-null int64
retweet_count     2327 non-null int64
dtypes: int64(3)
memory usage: 54.6 KB
```

1.8 Tidiness Issues

```
Twitter Archieved Enhanced Table
```

1.8.1 Issue #1: A single observation unit is stored in multiple column (doggo, floofer, pupper, puppo) shoild all be in dog stages.

Define: Create a new column named 'dog_stage' and assign corresponding to the dog stages according to doggo, floofer, pupper and puppo form the table in the new single column.

```
In [64]: #view column holding dog stages

         df_1.sample(5)

Out[64]:              tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
         504    813051746834595840                    NaN                  NaN
         1550   689154315265683456                    NaN                  NaN
         241    846505985330044928                    NaN                  NaN
         1593   686386521809772549                    NaN                  NaN
         1721   680130881361686529                    NaN                  NaN


                             timestamp  \
         504    2016-12-25 16:00:16 +0000
         1550   2016-01-18 18:36:07 +0000
         241    2017-03-27 23:35:28 +0000
         1593   2016-01-11 03:17:53 +0000
         1721   2015-12-24 21:00:12 +0000


                                                                                    sourc
         504    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         1550   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         241    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         1593   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         1721   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a


         504                                    This is Bauer. He had nothing to do with the cookies
```

```
1550       We normally don't rate birds but I feel bad cos this one forgot to fly south
241                                             THIS WAS NOT HIS FAULT HE HA
1593  Say hello to Crimson. He's a Speckled Winnebago. Main passions are air hockey &am
1721                                        This is Reggie. His Santa hat is a l

      retweeted_status_id  retweeted_status_user_id  \
504                   NaN                       NaN
1550                  NaN                       NaN
241                   NaN                       NaN
1593                  NaN                       NaN
1721                  NaN                       NaN

      retweeted_status_timestamp  \
504                          NaN
1550                         NaN
241                          NaN
1593                         NaN
1721                         NaN

                                               expanded_urls  \
504   https://twitter.com/dog_rates/status/813051746834595840/photo/1
1550  https://twitter.com/dog_rates/status/689154315265683456/photo/1
241          https://twitter.com/shomaristone/status/846484798663245829
1593  https://twitter.com/dog_rates/status/686386521809772549/photo/1
1721  https://twitter.com/dog_rates/status/680130881361686529/photo/1

      rating_numerator  rating_denominator     name doggo floofer pupper puppo
504                 13                  10    Bauer  None    None   None  None
1550                 9                  10     None  None    None   None  None
241                 11                  10     None  None    None   None  None
1593                11                  10  Crimson  None    None   None  None
1721                10                  10   Reggie  None    None   None  None
```

Code

```
In [65]: # Combine all four columns into one and select each stoge as entry

         df1_clean['stage'] = df1_clean['doggo'] + '_' + df1_clean['floofer'] + '_' + df1_clean[

In [66]: df1_clean['stage'] = df1_clean['stage'].str.replace('None', '')

In [67]: df1_clean['dog_stage'] = df1_clean['stage'].str.replace('_', '')

In [68]: df1_clean = df1_clean.drop(columns=['doggo', 'floofer', 'pupper', 'puppo', 'stage'])
```

Test

```
In [69]: df1_clean.sample(5)
```

```
Out[69]:                    tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
         662   790987426131050500                    NaN                  NaN
         794   773336787167145985                    NaN                  NaN
         125   868622495443632128                    NaN                  NaN
         1207  715733265223708672                    NaN                  NaN
         1720  680145970311643136                    NaN                  NaN


                        timestamp  \
         662   2016-10-25 18:44:32
         794   2016-09-07 01:47:12
         125   2017-05-28 00:18:35
         1207  2016-04-01 02:51:22
         1720  2015-12-24 22:00:10


                                                                         sourc
         662   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         794   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         125   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         1207  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
         1720  <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a


               retweeted_status_id  retweeted_status_user_id  \
         662                   NaN                       NaN
         794          7.713808e+17              4.196984e+09
         125                   NaN                       NaN
         1207                  NaN                       NaN
         1720                  NaN                       NaN


               retweeted_status_timestamp  \
         662                          NaN
         794       2016-09-01 16:14:48 +0000
         125                          NaN
         1207                         NaN
         1720                         NaN



         662
         794   https://twitter.com/dog_rates/status/771380798096281600/photo/1,https://twitter.c
         125
         1207
         1720                                               https://twitter.c


               rating_numerator  rating_denominator   name  \
         662                 12                  10  Misty
         794                 11                  10   Fizz
         125                 13                  10   None
         1207                10                  10      a
         1720                10                  10  Kawhi
```

```
662
794    RT @dog_rates: Meet Fizz. She thinks love is a social construct consisting solely
125                                    Here's a h*ckin peaceful boy. Unbothered by
1207                          This is a taco. We only rate dogs. Please only send in
1720                             Say hello to Kawhi. He was doing fine until his hat

        correct_denominator dog_stage
662                      10
794                      10
125                      10
1207                     10
1720                     10      pupper
```

1.8.2 Issue #2: Delete rows that are not original tweets

Define: According to the project motivation, we only want original tweets and not retweet. We
will drop every text row entry that start with `RT @` as this stands for retweet

In [70]: *#View row samples that are retweets*

```
df_1[df_1['text'].str.contains('RT @')]['text'].sample(5)
```

Out[70]: 185   RT @Jenna_Marbles: @dog_rates Thanks for rat
         506   RT @dog_rates: Meet Sammy. At first I was like "that's a snowflake. we only rate
         589   RT @dog_rates: I WAS SENT THE ACTUAL DOG IN THE PROFILE PIC BY HIS OWNER THIS IS
         36   RT @dog_rates: This is Lilly. She just parallel barked. Kindly requests a
         285   RT @KibaDva: I collected all the go
         Name: text, dtype: object

code

In [71]: *#Delete row that have the text start with 'RT @'*

```
df1_text = df1_clean[df1_clean['newtext'].str.contains('RT @')]
df1_text
```

Out[71]:   tweet_id in_reply_to_status_id in_reply_to_user_id \
         19   888202515573088257   NaN   NaN
         32   886054160059072513   NaN   NaN
         36   885311592912609280   NaN   NaN
         68   879130579576475649   NaN   NaN
         73   878404777348136964   NaN   NaN
         74   878316110768087041   NaN   NaN
         78   877611172832227328   NaN   NaN
         91   874434818259525634   NaN   NaN
         95   873697596434513921   NaN   NaN
         97   873337748698140672   NaN   NaN

101	872668790621863937	NaN	NaN
109	871166179821445120	NaN	NaN
118	869988702071779329	NaN	NaN
124	868639477480148993	NaN	NaN
130	867072653475098625	NaN	NaN
132	866816280283807744	NaN	NaN
137	866094527597207552	NaN	NaN
146	863471782782697472	NaN	NaN
155	861769973181624320	NaN	NaN
159	860981674716409858	NaN	NaN
160	860924035999428608	NaN	NaN
165	860177593139703809	NaN	NaN
171	858860390427611136	NaN	NaN
180	857062103051644929	NaN	NaN
182	856602993587888130	NaN	NaN
185	856330835276025856	NaN	NaN
194	855245323840757760	NaN	NaN
195	855138241867124737	NaN	NaN
204	852936405516943360	NaN	NaN
211	851953902622658560	NaN	NaN
...	...	...	...
784	775096608509886464	NaN	NaN
794	773336787167145985	NaN	NaN
800	772615324260794368	NaN	NaN
811	771171053431250945	NaN	NaN
815	771004394259247104	NaN	NaN
818	770743923962707968	NaN	NaN
822	770093767776997377	NaN	NaN
826	769335591808995329	NaN	NaN
829	768909767477751808	NaN	NaN
833	768554158521745409	NaN	NaN
841	766864461642756096	NaN	NaN
847	766078092750233600	NaN	NaN
860	763167063695355904	NaN	NaN
868	761750502866649088	NaN	NaN
872	761371037149827077	NaN	NaN
885	760153949710192640	NaN	NaN
890	759566828574212096	NaN	NaN
895	759159934323924993	NaN	NaN
908	757729163776290825	NaN	NaN
911	757597904299253760	NaN	NaN
926	754874841593970688	NaN	NaN
937	753298634498793472	NaN	NaN
943	752701944171524096	NaN	NaN
949	752309394570878976	NaN	NaN
1012	747242308580548608	NaN	NaN
1023	746521445350707200	NaN	NaN
1043	743835915802583040	NaN	NaN

```
1242   7119988098580443904                              NaN                     NaN
2259   6675509049550915073                             NaN                     NaN
2260   6675508829055632768                             NaN                     NaN

                   timestamp   \
19     2017-07-21 01:02:36
32     2017-07-15 02:45:48
36     2017-07-13 01:35:06
68     2017-06-26 00:13:58
73     2017-06-24 00:09:53
74     2017-06-23 18:17:33
78     2017-06-21 19:36:23
91     2017-06-13 01:14:41
95     2017-06-11 00:25:14
97     2017-06-10 00:35:19
101    2017-06-08 04:17:07
109    2017-06-04 00:46:17
118    2017-05-31 18:47:24
124    2017-05-28 01:26:04
130    2017-05-23 17:40:04
132    2017-05-23 00:41:20
137    2017-05-21 00:53:21
146    2017-05-13 19:11:30
155    2017-05-09 02:29:07
159    2017-05-06 22:16:42
160    2017-05-06 18:27:40
165    2017-05-04 17:01:34
171    2017-05-01 01:47:28
180    2017-04-26 02:41:43
182    2017-04-24 20:17:23
185    2017-04-24 02:15:55
194    2017-04-21 02:22:29
195    2017-04-20 19:16:59
204    2017-04-14 17:27:40
211    2017-04-12 00:23:33
...                    ...
784    2016-09-11 22:20:06
794    2016-09-07 01:47:12
800    2016-09-05 02:00:22
811    2016-09-01 02:21:21
815    2016-08-31 15:19:06
818    2016-08-30 22:04:05
822    2016-08-29 03:00:36
826    2016-08-27 00:47:53
829    2016-08-25 20:35:48
833    2016-08-24 21:02:45
841    2016-08-20 05:08:29
847    2016-08-18 01:03:45
```

127

```
860  2016-08-10 00:16:21
868  2016-08-06 02:27:27
872  2016-08-05 01:19:35
885  2016-08-01 16:43:19
890  2016-07-31 01:50:18
895  2016-07-29 22:53:27
908  2016-07-26 00:08:05
911  2016-07-25 15:26:30
926  2016-07-18 03:06:01
937  2016-07-13 18:42:44
943  2016-07-12 03:11:42
949  2016-07-11 01:11:51
1012 2016-06-27 01:37:04
1023 2016-06-25 01:52:36
1043 2016-06-17 16:01:16
1242 2016-03-21 19:31:59
2259 2015-11-20 03:51:52
2260 2015-11-20 03:51:47
```

```
                                                                    sourc
19    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
32    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
36    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
68    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
73    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
74    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
78    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
91    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
95    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
97    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
101   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
109   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
118   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
124   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
130   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
132   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
137   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
146   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
155   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
159   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
160   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
165   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
171   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
180   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
182   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
185   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
194   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
195   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
```

```
204    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
211    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
...                                                                                    ..
784    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
794    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
800    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
811    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
815    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
818    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
822    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
826    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
829    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
833    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
841    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
847    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
860    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
868    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
872    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
885    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
890    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
895    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
908    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
911    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
926    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
937    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
943    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
949    <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
1012   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
1023   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
1043   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
1242   <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a
2259              <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a
2260              <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a

       retweeted_status_id  retweeted_status_user_id  \
19            8.874740e+17              4.196984e+09
32            8.860537e+17              1.960740e+07
36            8.305833e+17              4.196984e+09
68            8.780576e+17              4.196984e+09
73            8.782815e+17              4.196984e+09
74            6.690004e+17              4.196984e+09
78            8.768508e+17              5.128045e+08
91            8.663350e+17              4.196984e+09
95            8.688804e+17              4.196984e+09
97            8.732138e+17              4.196984e+09
101           8.726576e+17              1.547674e+08
109           8.410770e+17              4.196984e+09
118           8.591970e+17              4.196984e+09
```

129

124	8.685523e+17	4.196984e+09
130	8.650134e+17	7.874618e+17
132	8.664507e+17	4.196984e+09
137	8.378202e+17	4.196984e+09
146	8.630625e+17	4.196984e+09
155	8.066291e+17	4.196984e+09
159	8.605638e+17	4.196984e+09
160	8.609145e+17	3.638908e+08
165	7.616730e+17	4.196984e+09
171	8.395493e+17	4.196984e+09
180	8.570611e+17	5.870972e+07
182	8.447048e+17	4.196984e+09
185	8.563302e+17	6.669901e+07
194	8.421635e+17	4.196984e+09
195	8.551225e+17	7.475543e+17
204	8.316501e+17	4.196984e+09
211	8.293743e+17	4.196984e+09
...	...	...
784	7.403732e+17	4.196984e+09
794	7.713808e+17	4.196984e+09
800	7.652221e+17	4.196984e+09
811	6.733201e+17	4.196984e+09
815	7.710021e+17	1.732729e+09
818	7.392382e+17	4.196984e+09
822	7.410673e+17	4.196984e+09
826	7.069045e+17	4.196984e+09
829	7.001438e+17	4.196984e+09
833	7.399792e+17	4.196984e+09
841	7.599238e+17	4.196984e+09
847	7.258423e+17	4.196984e+09
860	6.732953e+17	4.196984e+09
868	6.853251e+17	4.196984e+09
872	7.116948e+17	4.196984e+09
885	7.601538e+17	1.950368e+08
890	7.395441e+17	4.196984e+09
895	6.703191e+17	4.196984e+09
908	6.790626e+17	4.196984e+09
911	7.575971e+17	2.804798e+08
926	6.791584e+17	4.196984e+09
937	6.815232e+17	4.196984e+09
943	6.835159e+17	4.196984e+09
949	6.753544e+17	4.196984e+09
1012	7.047611e+17	4.196984e+09
1023	6.678667e+17	4.196984e+09
1043	6.671383e+17	4.196984e+09
1242	7.119983e+17	7.832140e+05
2259	6.675487e+17	4.296832e+09
2260	6.675484e+17	4.296832e+09

```
     retweeted_status_timestamp  \
19     2017-07-19 00:47:34 +0000
32     2017-07-15 02:44:07 +0000
36     2017-02-12 01:04:29 +0000
68     2017-06-23 01:10:23 +0000
73     2017-06-23 16:00:04 +0000
74     2015-11-24 03:51:38 +0000
78     2017-06-19 17:14:49 +0000
91     2017-05-21 16:48:45 +0000
95     2017-05-28 17:23:24 +0000
97     2017-06-09 16:22:42 +0000
101    2017-06-08 03:32:35 +0000
109    2017-03-13 00:02:39 +0000
118    2017-05-02 00:04:57 +0000
124    2017-05-27 19:39:34 +0000
130    2017-05-18 01:17:25 +0000
132    2017-05-22 00:28:40 +0000
137    2017-03-04 00:21:08 +0000
146    2017-05-12 16:05:02 +0000
155    2016-12-07 22:38:52 +0000
159    2017-05-05 18:36:06 +0000
160    2017-05-06 17:49:42 +0000
165    2016-08-05 21:19:27 +0000
171    2017-03-08 18:52:12 +0000
180    2017-04-26 02:37:47 +0000
182    2017-03-23 00:18:10 +0000
185    2017-04-24 02:13:14 +0000
194    2017-03-16 00:00:07 +0000
195    2017-04-20 18:14:33 +0000
204    2017-02-14 23:43:18 +0000
211    2017-02-08 17:00:26 +0000
...                          ...
784    2016-06-08 02:41:38 +0000
794    2016-09-01 16:14:48 +0000
800    2016-08-15 16:22:20 +0000
811    2015-12-06 01:56:44 +0000
815    2016-08-31 15:10:07 +0000
818    2016-06-04 23:31:25 +0000
822    2016-06-10 00:39:48 +0000
826    2016-03-07 18:09:06 +0000
829    2016-02-18 02:24:13 +0000
833    2016-06-07 00:36:02 +0000
841    2016-08-01 01:28:46 +0000
847    2016-04-29 00:21:01 +0000
860    2015-12-06 00:17:55 +0000
868    2016-01-08 05:00:14 +0000
872    2016-03-20 23:23:54 +0000
```

```
885     2016-08-01 16:42:51 +0000
890     2016-06-05 19:47:03 +0000
895     2015-11-27 19:11:49 +0000
908     2015-12-21 22:15:18 +0000
911     2016-07-25 15:23:28 +0000
926     2015-12-22 04:35:49 +0000
937     2015-12-28 17:12:42 +0000
943     2016-01-03 05:11:12 +0000
949     2015-12-11 16:40:19 +0000
1012    2016-03-01 20:11:59 +0000
1023    2015-11-21 00:46:50 +0000
1043    2015-11-19 00:32:12 +0000
1242    2016-03-21 19:29:52 +0000
2259    2015-11-20 03:43:06 +0000
2260    2015-11-20 03:41:59 +0000


19
32
36
68
73
74
78
91
95
97
101
109
118
124
130
132
137
146
155
159
160
165
171
180
182
185
194
195
204
211
...
```

```
784
794
800
811    https://twitter.com/dog_rates/status/673320132811366400/photo/1,https://twitter.c
815
818
822
826
829
833
841
847
860
868
872
885
890
895
908
911
926
937
943
949
1012
1023
1043
1242
2259
2260
```

	rating_numerator	rating_denominator	name \
19	13	10	Canela
32	12	10	None
36	13	10	Lilly
68	14	10	Emmy
73	13	10	Shadow
74	11	10	Terrance
78	14	10	None
91	12	10	Coco
95	14	10	Walter
97	12	10	Sierra
101	14	10	None
109	12	10	Dawn
118	12	10	quite
124	12	10	Cooper
130	13	10	None
132	13	10	Jamesy

137	12	10	None
146	13	10	Quinn
155	13	10	None
159	13	10	Lorenzo
160	13	10	None
165	10	10	None
171	12	10	Winston
180	13	10	None
182	13	10	Luna
185	14	10	None
194	12	10	George
195	14	10	None
204	13	10	None
211	13	10	Astrid
...	...	...	...
784	9	11	None
794	11	10	Fizz
800	10	10	Gromit
811	11	10	Frankie
815	12	10	None
818	13	10	None
822	12	10	just
826	13	10	None
829	10	10	None
833	12	10	Nollie
841	10	10	None
847	12	10	Colby
860	8	10	Eve
868	10	10	None
872	13	10	None
885	11	10	None
890	10	10	None
895	11	10	None
908	11	10	Chompsky
911	11	10	None
926	11	10	Rubio
937	12	10	Carly
943	11	10	None
949	13	10	None
1012	13	10	None
1023	10	10	Shaggy
1043	10	10	None
1242	12	10	None
2259	12	10	None
2260	5	10	None

19   RT @dog_rates: This is Canela. She attempted some fanc

```
32
36                                                    RT @dog_rates: This is Lilly. She just parallel b
68                             RT @dog_rates: This is Emmy. She was adopted today. Massive round of p
73      RT @dog_rates: Meet Shadow. In an attempt to reach maximum zooming borkdrive, he
74                            RT @dog_rates: Meet Terrance. He's being yelled at because he st
78                                        RT @rachel2195: @dog_rates the boyfrie
91                   RT @dog_rates: This is Coco. At first I thought she was a cloud but
95                                    RT @dog_rates: This is Walter. He won't start hydroth
97      RT @dog_rates: This is Sierra. She's one precious pupper. Absolute 12/10. Been in
101       RT @loganamnosis: Penelope here is doing me quite a divertir. Well done, @dog_r
109                          RT @dog_rates: This is Dawn. She's just checking pup on you. Mak
118     RT @dog_rates: We only rate dogs. This is quite clearly a smol broken polar bea
124             RT @dog_rates: Say hello to Cooper. His expression is the same wet or dr
130                                                                           RT @
132                  RT @dog_rates: This is Jamesy. He gives a kiss to every other
137                  RT @dog_rates: Here's a pupper before and after being asked "
146             RT @dog_rates: Say hello to Quinn. She's quite the goofball. Not even a
155                                 RT @dog_rates: "Good afternoon class tod
159       RT @dog_rates: Meet Lorenzo. He's an avid nifty hat wearer and absolute
160
165            RT @dog_rates: Ohboyohboyohboyohboyohboyohboyohboyohboyohboyohboyoh
171     RT @dog_rates: Meet Winston. He knows he's a little too big for the swing, but
180                   RT @AaronChewning: First time wearing my @dog_rates hat on a flig
182      RT @dog_rates: This is Luna. It's her first time outside and a bee stung her n
185                                 RT @Jenna_Marbles: @dog_rates Thanks for ra
194     RT @dog_rates: Meet George. He looks slightly deflated but overall quite powerf
195     RT @frasercampbell_: oh my... what's that... beautiful scarf around your neck..
204           RT @dog_rates: I usually only share these on Friday's, but this is Blue
211                                       RT @dog_rates: This is Astrid.
...
784     RT @dog_rates: After so many requests, this is Bretagne. She was the last survi
794     RT @dog_rates: Meet Fizz. She thinks love is a social construct consisting sole
800                        RT @dog_rates: This is Gromit. He's pupset because
811                                     RT @dog_rates: This is Frank
815                      RT @katieornah: @dog_rates learning a lot at
818        RT @dog_rates: Here's a doggo blowing bubbles. It's downright legenda
822                                       RT @dog_rates: This is
826              RT @dog_rates: Ever seen a dog pet another dog? Bot
829     RT @dog_rates: When it's Janet from accounting's birthday but you can't eat the
833               RT @dog_rates: This is Nollie. She's waving at you. If you don't wa
841     RT @dog_rates: We only rate dogs... this is a Taiwanese Guide Walrus. Im gettin
847                   RT @dog_rates: This is Colby. He's currently regretting
860     RT @dog_rates: Meet Eve. She's a raging alcoholic 8/10 (would b 11/10 but puppe
868                         RT @dog_rates: "Tristan do not speak to me
872
885                            RT @hownottodraw: The story/person behind
890                     RT @dog_rates: This... is a Tyrannosaurus rex. We
895
```

908	RT
911	RT @jon_hill987: @dog_rates Ther
926	
937	RT @dog_rates: This is Carly. She's actually 2 dogs fused together. Very innova
943	RT @dog_rates: HEY PUP WHAT'S THE PART OF THE H
949	
1012	RT @dog_rates: This pupper killed this great white in an epic sea batt
1023	RT @dog_rates: This is Shaggy. He knows exactly how to solve the puzzle but can
1043	RT @dog_rates: Extremely intelligent dog here. Has learned to
1242	RT @twitter:
2259	RT @dogratingrating: Exceptional talent
2260	RT @dogratingrating:

	correct_denominator	dog_stage
19	10	
32	10	
36	10	
68	10	
73	10	
74	10	
78	10	pupper
91	10	
95	10	
97	10	pupper
101	10	
109	10	
118	10	
124	10	
130	10	
132	10	pupper
137	10	pupper
146	10	
155	10	
159	10	
160	10	
165	10	
171	10	
180	10	
182	10	
185	10	
194	10	
195	10	
204	10	
211	10	doggo
...	...	...
784	11	
794	10	
800	10	

```
811                    10
815                    10          pupper
818                    10           doggo
822                    10  doggopupper
826                    10
829                    10          pupper
833                    10
841                    10
847                    10
860                    10          pupper
868                    10
872                    10
885                    10
890                    10
895                    10
908                    10
911                    10          pupper
926                    10
937                    10
943                    10
949                    10
1012                   10          pupper
1023                   10
1043                   10
1242                   10
2259                   10
2260                   10

[181 rows x 15 columns]
```

In [72]: df1_clean = df1_clean[~df1_clean.newtext.isin(df1_text.newtext)]

Test

In [73]: # count row that start with 'RT @' in the cleaned data

df1_clean[df1_clean['newtext'].str.contains('RT @')].newtext.count()

Out[73]: 0

1.9   Data Storing

Store the clean master data frame in a CSV `twitter_archive_master.csv`

In [74]: df_master = df1_clean.merge(df2_clean, on='tweet_id').merge(df3_clean, on='tweet_id')
         df_master.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1986 entries, 0 to 1985
```

```
Data columns (total 28 columns):
tweet_id                    1986 non-null int64
in_reply_to_status_id       23 non-null float64
in_reply_to_user_id         23 non-null float64
timestamp                   1986 non-null datetime64[ns]
source                      1986 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               1986 non-null object
rating_numerator            1986 non-null int64
rating_denominator          1986 non-null int64
name                        1986 non-null object
newtext                     1986 non-null object
correct_denominator         1986 non-null int64
dog_stage                   1986 non-null object
jpg_url                     1986 non-null object
img_number                  1986 non-null int64
1st_prediction              1986 non-null object
1st_confidence              1986 non-null object
1st_DogBreed                1986 non-null bool
2nd_prediction              1986 non-null object
2nd_confidence              1986 non-null object
2nd_DogBreed                1986 non-null bool
3rd_prediction              1986 non-null object
3rd_confidence              1986 non-null object
3rd_DogBreed                1986 non-null bool
favorite_count              1986 non-null int64
retweet_count               1986 non-null int64
dtypes: bool(3), datetime64[ns](1), float64(4), int64(7), object(13)
memory usage: 409.2+ KB
```

In [75]: *# store all three data as CSV*

```python
df_master.to_csv('twitter_archive_master.csv', index=False)
```

Additionally, Store the clean master data in a SQL database

In [76]: *# import sqlalchemy library*

```python
from sqlalchemy import create_engine
engine = create_engine('sqlite://', echo=False)
sql_connection = engine.connect()
```

In [77]: *# save all three cleaned data as sql lite*

```python
df_master.to_sql('twitter_archive_master', sql_connection, if_exists='replace')
```

1.10   Analyzing and Visualizing Data

```
In [78]: df1_clean.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2175 entries, 0 to 2355
Data columns (total 15 columns):
tweet_id                       2175 non-null int64
in_reply_to_status_id          78 non-null float64
in_reply_to_user_id            78 non-null float64
timestamp                      2175 non-null datetime64[ns]
source                         2175 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                  2117 non-null object
rating_numerator               2175 non-null int64
rating_denominator             2175 non-null int64
name                           2175 non-null object
newtext                        2175 non-null object
correct_denominator            2175 non-null int64
dog_stage                      2175 non-null object
dtypes: datetime64[ns](1), float64(4), int64(4), object(6)
memory usage: 271.9+ KB
```

1.10.1   Insights:

1. The most common dog name   We can get the number of dogs with common names and the highest occuring name will be the most popular name. From the twitter archived data, after cleaning and accessing. We observe that there is a significant number of dogs that have their name registered as NONE, and some with non feasible dog names such as 'A' and 'AN'.

Some of the non-feasible dog names with a readible pattern were corrected and cleaned.

```
In [79]: #group dogs by dog name

         dv1 = df1_clean.groupby('name')['name'].count().sort_values().tail(10)
         dv1

Out[79]: name
         Lola          8
         Sadie         8
         Tucker        9
         Penny         9
         Cooper       10
         Oliver       10
         Lucy         11
         Charlie      11
         a            55
```

```
None        680
Name: name, dtype: int64
```

Ingoring the `None` and `a`, The most common name of dog from this project is `Charlie`and `Lucy` with 11 counts, followed by `Oliver`, and `Cooper` who have 10 dogs each answering them.

2. The most popular dog breed Another interesting insight we can get from the data is the most popular dog breed. Here we use the twitter image prediction data and the prediction of the first confidence.

```
In [80]: df3_clean.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2327 entries, 0 to 2326
Data columns (total 3 columns):
tweet_id         2327 non-null int64
favorite_count   2327 non-null int64
retweet_count    2327 non-null int64
dtypes: int64(3)
memory usage: 54.6 KB
```

```
In [81]: #select favorite count from the cleaned data

         dv2 = df3_clean[['tweet_id', 'favorite_count']]
```

```
In [82]: #merge favorite count with the image prediction data

         dv2 = pd.merge(df2_clean, dv2, on='tweet_id', how='inner')
```

```
In [83]: # Find the most popular dog breed
         popular_breeds = dv2['1st_prediction'].value_counts().head(5)
         popular_breeds
```

```
Out[83]: golden_retriever       150
         Labrador_retriever      96
         Pembroke                88
         Chihuahua               81
         pug                     57
         Name: 1st_prediction, dtype: int64
```

```
In [84]: breeds = ['golden_retriever','Labrador_retriever','Pembroke','Chihuahua','pug']
         top_breeds = dv2.loc[dv2['1st_prediction'].isin(breeds)]
```

```
In [85]: #statistical summary of the popular dog breeds

         top_breeds.groupby('1st_prediction').favorite_count.describe()
```

```
Out[85]:                        count         mean            std      min      25%      50%  \
         1st_prediction
         Chihuahua            81.0  7347.419753  14337.055160      0.0  1008.00  2393.0
         Labrador_retriever   96.0  9432.125000  16884.542013      0.0  1732.25  4515.0
         Pembroke             88.0  9806.556818  10692.361425    139.0  2676.50  5513.0
         golden_retriever    150.0  9783.793333  11361.647695      0.0  2636.25  6165.5
         pug                  57.0  4732.596491   6960.793458      0.0  1244.00  2178.0

                                  75%        max
         1st_prediction
         Chihuahua             8419.00   111035.0
         Labrador_retriever   12674.50   143948.0
         Pembroke             13057.00    59900.0
         golden_retriever     13515.25    73602.0
         pug                   4486.00    31424.0
```

Correlation between the favorite_count and retweet_count   We can measure the correlation between the favorite counts an retweet count to see how they vary together if it's positive or negative.

```
In [86]: #select the favorite count and retweet count columns

         dv3 = df3_clean[['favorite_count', 'retweet_count']]

In [87]: #statistical summary of the favourite and retweet count

         dv3.describe()

Out[87]:        favorite_count   retweet_count
         count     2327.000000     2327.000000
         mean      7015.309411     2455.496777
         std      10899.541160     4157.858444
         min          0.000000        1.000000
         25%       1218.500000      492.500000
         50%       3033.000000     1143.000000
         75%       8551.500000     2842.500000
         max     143948.000000    70213.000000
```
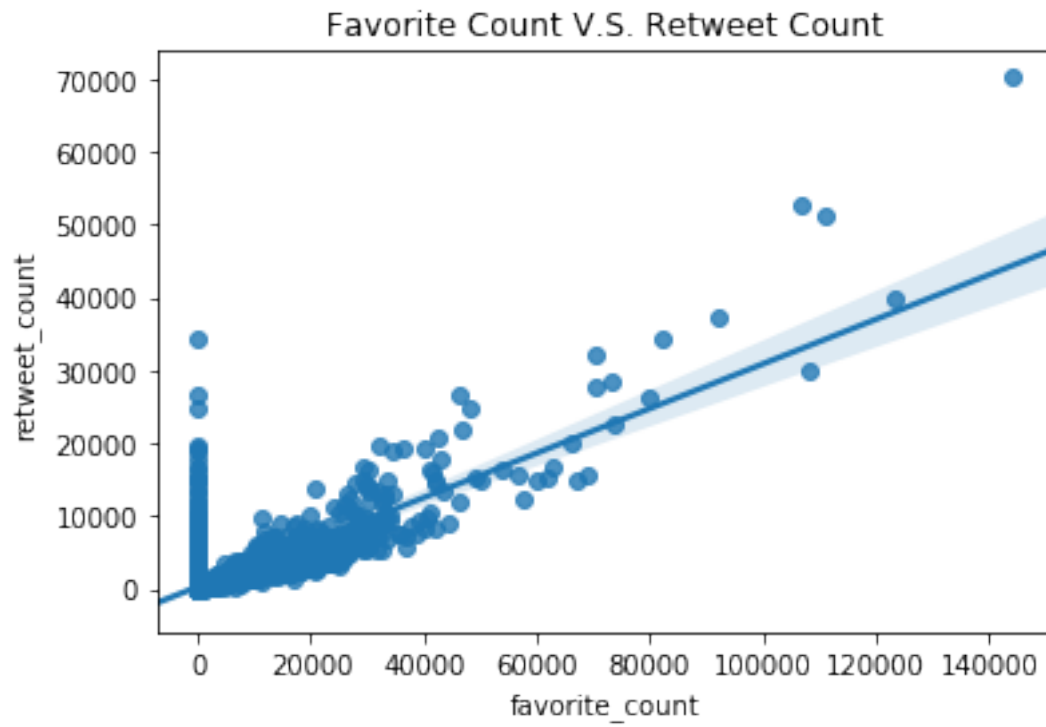
1.10.2   Visualization

```
In [88]: #Correlation plot of the favorite and retweet count

         sns.regplot(x="favorite_count", y="retweet_count", data=dv3).set_title('Favorite Count
         plt.savefig('correlation.png');
```

Favorite Count V.S. Retweet Count

In [ ]: