twitter-questions

April 15, 2019

```python
In [1]: import pandas as pd
        import json
        import ast
        import operator
        import re
        from wordcloud import WordCloud, STOPWORDS
        import matplotlib.pyplot as plt
        import warnings
        warnings.filterwarnings('ignore')

In [2]: def dict_sort_des(diction):
            sorted_d = sorted(diction.items(), key=operator.itemgetter(1),reverse=True)
            return sorted_d

        def clean_tweet(tweet):
            return ' '.join(re.sub(r"(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])|(\w+:\/\/\S+)", " ", twee

        def wordcloud_plot(text, colour):
            wordcloud = WordCloud(
                width = 3000,
                height = 2000,
                background_color = colour,
                stopwords = STOPWORDS).generate(str(text))

            fig = plt.figure(
                figsize = (40, 30),
                facecolor = 'k',
                edgecolor = 'k')

            plt.imshow(wordcloud, interpolation = 'bilinear')
            plt.axis('off')
            plt.tight_layout(pad=0)
            return plt.show()

In [3]: df = pd.read_csv('BJP4India_tweets.csv')
        # Filtering the retweets by text.
        df =  df[~df.text.str.contains("RT")]
        df
```

```
Out[3]:                        id          created_at  \
        0     1116322959956697093  2019-04-11 12:51:48
        1     1116320552761749505  2019-04-11 12:42:14
        2     1116318097034469376  2019-04-11 12:32:29
        3     1116312258169016320  2019-04-11 12:09:17
        4     1116296458439761920  2019-04-11 11:06:30
        5     1116296194647400449  2019-04-11 11:05:27
        6     1116296089244557312  2019-04-11 11:05:02
        7     1116295959065980930  2019-04-11 11:04:31
        8     1116295774218833926  2019-04-11 11:03:47
        9     1116294230052184064  2019-04-11 10:57:38
        10    1116293614009638912  2019-04-11 10:55:11
        11    1116293188443000832  2019-04-11 10:53:30
        12    1116291919812775937  2019-04-11 10:48:28
        13    1116291571169652736  2019-04-11 10:47:04
        14    1116290865385787393  2019-04-11 10:44:16
        15    1116289434448564224  2019-04-11 10:38:35
        16    1116284892659064833  2019-04-11 10:20:32
        18    1116282980106817537  2019-04-11 10:12:56
        19    1116280596840345601  2019-04-11 10:03:28
        20    1116277167149555712  2019-04-11 09:49:50
        21    1116276694459932672  2019-04-11 09:47:58
        22    1116276166489337856  2019-04-11 09:45:52
        23    1116275545036050438  2019-04-11 09:43:23
        24    1116275049055449093  2019-04-11 09:41:25
        25    1116274730422554625  2019-04-11 09:40:09
        26    1116273757553713152  2019-04-11 09:36:17
        27    1116273270150447105  2019-04-11 09:34:21
        28    1116272717387288576  2019-04-11 09:32:09
        29    1116272051650584576  2019-04-11 09:29:31
        30    1116269847057354752  2019-04-11 09:20:45
        ...                   ...                  ...
        2954  1100371764939575296  2019-02-26 12:27:27
        2955  1100369614528106496  2019-02-26 12:18:54
        2956  1100364952957739009  2019-02-26 12:00:23
        2957  1100364753858260997  2019-02-26 11:59:35
        2958  1100363788723712005  2019-02-26 11:55:45
        2959  1100363301521743872  2019-02-26 11:53:49
        2960  1100362753405968384  2019-02-26 11:51:38
        2961  1100362449302089728  2019-02-26 11:50:26
        2962  1100361739181338624  2019-02-26 11:47:37
        2963  1100357006534176768  2019-02-26 11:28:48
        2964  1100355427244552192  2019-02-26 11:22:32
        2965  1100350220053835776  2019-02-26 11:01:50
        2966  1100346265563353088  2019-02-26 10:46:07
        2967  1100344354068094976  2019-02-26 10:38:32
        2968  1100328342035681280  2019-02-26 09:34:54
        2969  1100321245474639872  2019-02-26 09:06:42
```

```
2970   1100320087297257472   2019-02-26 09:02:06
2971   1100319590947536898   2019-02-26 09:00:08
2972   1100319037290999808   2019-02-26 08:57:56
2973   1100318239676985344   2019-02-26 08:54:45
2974   1100317848612655105   2019-02-26 08:53:12
2975   1100316748861956099   2019-02-26 08:48:50
2976   1100316311962378245   2019-02-26 08:47:06
2977   1100316140457185281   2019-02-26 08:46:25
2978   1100315560854704128   2019-02-26 08:44:07
2979   1100314269105582080   2019-02-26 08:38:59
2980   1100313771497508864   2019-02-26 08:37:00
2981   1100312402761895936   2019-02-26 08:31:34
2984   1100280586575335424   2019-02-26 06:25:08
2985   1100278248615936000   2019-02-26 06:15:51
```

```
                                                text  retweet_count  \
0      Mamata Didi wants to impose Urdu over Bengali ...            250
1      Mamata Banerjees TMC stands for \n\nT : 'Tush...            239
2      Congress stalled the Assam Accord for three de...            258
3      A new scam, Tughlak road election scam, has be...            458
4                  ,    ...            221
5      We are committed to balanced development of As...            162
6                        ...            342
7                      ...            422
8                      ...            360
9      We are committed to implementing Citizen Amend...            185
10     Congress has a history of creating trouble. In...            437
11                 ,    ...            320
12                 ,    ...            245
13     NDA will win all five seats of Assam on which ...            236
14                 ,    ...            240
15     LIVE: PM Modi addresses public meeting at Silc...            359
16                     ...            560
18     LIVE: Shri @naqvimukhtar is addressing a press...            347
19     In Bengal, the goons of TMC have killed more t...            831
20           ,\n\n   ...            302
21           , 40    ...            188
22             ,    ...            319
23     BJP is committed to bring NRC in West Bengal l...            230
24       4.24         ...            187
25         -       ...            282
26     Mamata Banerjees TMC stands for \nT : Tushtik...            551
27     The people of the country have made up their m...            184
28                 ,...            205
29     Mamata Didi wants to impose Urdu over Bengali ...            254
30     LIVE: Shri @AmitShah addresses public meeting ...            727
...                                          ...            ...
2954                 ...            589
```

```
2955  Shri @AmitShah is speaking at Chaupal Pe Charc...        1713
2956  Look at the affection and love for PM Modi in ...         702
2957              ...              929
2958              ...              524
2959              ...              575
2960              ...              826
2961              ...              570
2962              ...              548
2963  PM Shri @narendramodi unveils world's largest ...         845
2964  LIVE : PM Modi unveils world's largest Bhagwad...        1075
2965      ,    0...              946
2966      , ,   ...             2429
2967      1    ...              701
2968      \n    ...             4779
2969              ...             3356
2970      ,       ...              514
2971       50  ...              625
2972              ...              571
2973  :        ,  ...              624
2974    ,        ...              753
2975              ...              477
2976       OROP ...              679
2977        \n\n ...              432
2978      ,       ...              978
2979      ,       ...              551
2980        \n\n ...              608
2981  :         ,  ...             1112
2984  Let's crack some job numbers for the naysayers...         535
2985  Let's crack some job numbers for the naysayers...         716

       favorite_count  photos  videos  url  \
0                 743       0       1    1
1                 705       0       1    1
2                 766       0       1    1
3                1101       0       1    1
4                 794       0       0    0
5                 504       0       0    0
6                1173       0       0    0
7                1826       0       0    0
8                1390       0       0    0
9                 524       2       0    1
10               1141       0       0    0
11               1017       0       0    0
12                676       2       0    1
13                691       2       0    1
14                629       2       0    1
15               1079       0       0    1
16               1530       0       0    0
```

```
18            992        0        0      1
19           2363        0        1      1
20            863        2        0      1
21            599        2        0      1
22           1120        0        0      0
23            626        2        0      1
24            529        2        0      1
25            844        2        0      1
26           1605        0        0      0
27            516        2        0      1
28            638        2        0      1
29            701        2        0      1
30           1965        0        0      1
...            ...      ...      ...    ...
2954         1331        2        0      1
2955         4940        0        0      1
2956         2019        2        0      1
2957         3161        0        0      0
2958         1331        2        0      1
2959         1619        0        0      0
2960         3044        0        0      0
2961         1324        2        0      1
2962         1364        2        0      1
2963         2491        2        0      1
2964         3340        0        0      1
2965         3071        0        1      1
2966         9691        0        1      1
2967         1959        0        1      1
2968        18331        0        1      1
2969        13796        0        1      1
2970         1303        2        0      1
2971         2170        0        0      0
2972         1768        2        0      1
2973         2235        0        0      0
2974         2675        2        0      1
2975         1247        2        0      1
2976         2419        0        0      0
2977         1102        2        0      1
2978         2747        2        0      1
2979         1559        2        0      1
2980         1846        2        0      1
2981         3475        0        0      1
2984         1312        0        1      1
2985         2139        0        1      1

                                             hashes
0     [{u'indices': [238, 254], u'text': u'NaMoForNe...
1     [{u'indices': [176, 192], u'text': u'NaMoForNe...
```

5

```
2      [{u'indices': [229, 245], u'text': u'NaMoForNe...
3      [{u'indices': [265, 281], u'text': u'NaMoForNe...
4      [{u'indices': [151, 167], u'text': u'NaMoForNe...
5      [{u'indices': [186, 202], u'text': u'NaMoForNe...
6      [{u'indices': [197, 213], u'text': u'NaMoForNe...
7      [{u'indices': [150, 166], u'text': u'NaMoForNe...
8      [{u'indices': [206, 222], u'text': u'NaMoForNe...
9      [{u'indices': [130, 146], u'text': u'NaMoForNe...
10     [{u'indices': [263, 279], u'text': u'NaMoForNe...
11     [{u'indices': [171, 187], u'text': u'NaMoForNe...
12     [{u'indices': [240, 256], u'text': u'NaMoForNe...
13     [{u'indices': [145, 161], u'text': u'NaMoForNe...
14     [{u'indices': [215, 231], u'text': u'NaMoForNe...
15     [{u'indices': [58, 74], u'text': u'NaMoForNewI...
16     [{u'indices': [185, 201], u'text': u'NaMoForNe...
18                                                   []
19     [{u'indices': [245, 261], u'text': u'NaMoForNe...
20     [{u'indices': [53, 69], u'text': u'NaMoForNewI...
21     [{u'indices': [193, 209], u'text': u'NaMoForNe...
22     [{u'indices': [158, 174], u'text': u'NaMoForNe...
23     [{u'indices': [181, 197], u'text': u'NaMoForNe...
24     [{u'indices': [121, 137], u'text': u'NaMoForNe...
25     [{u'indices': [239, 255], u'text': u'NaMoForNe...
26     [{u'indices': [137, 153], u'text': u'NaMoForNe...
27     [{u'indices': [251, 267], u'text': u'NaMoForNe...
28     [{u'indices': [207, 223], u'text': u'NaMoForNe...
29     [{u'indices': [168, 184], u'text': u'NaMoForNe...
30     [{u'indices': [78, 94], u'text': u'NaMoForNewI...
...                                                 ...
2954                                                 []
2955                                                 []
2956                                                 []
2957                                                 []
2958                                                 []
2959                                                 []
2960                                                 []
2961                                                 []
2962                                                 []
2963                                                 []
2964                                                 []
2965                                                 []
2966                                                 []
2967                                                 []
2968                                                 []
2969                                                 []
2970                                                 []
2971                                                 []
2972                                                 []
```

```
         2973                                                  []
         2974                                                  []
         2975                                                  []
         2976                                                  []
         2977                                                  []
         2978                                                  []
         2979                                                  []
         2980                                                  []
         2981                                                  []
         2984  [{u'indices': [261, 276], u'text': u'HereAreTh...
         2985  [{u'indices': [203, 218], u'text': u'HereAreTh...

         [2827 rows x 9 columns]
```

```
In [4]: k = df[df['favorite_count']==df['favorite_count'].max()]
        pd.set_option('display.max_colwidth', -1)
        k
```

```
Out[4]:                        id           created_at  \
        1804   1110886268408233984  2019-03-27 12:48:20


        1804  Congress led UPA\nSurgical Strike : Don't do it\nAir Strike: Don't do it\nA-SAT Mi


              retweet_count  favorite_count  photos  videos  url  \
        1804  13821          43525           0       0       0


                                                              hashes
        1804  [{u'indices': [210, 224], u'text': u'MissionShakti'}]
```

```
In [5]: k = df[df['retweet_count']==df['retweet_count'].max()]
        pd.set_option('display.max_colwidth', -1)
        k
```

```
Out[5]:                        id           created_at  \
        705   1114495595421376512  2019-04-06 11:50:30


        705                  \n\n ? https://t.co/U2KZlMFf7b


              retweet_count  favorite_count  photos  videos  url hashes
        705   18116          37735           0       1       1   []
```

0.1 Algorithm for popularity -> retweet_count(0.7) + favourite_count(0.3) ->score

```
In [13]: j = df.assign(f = df['retweet_count']*0.7 + df['favorite_count']*0.3).sort_values('f').
         j = j.tail(5)
         j
```

```
Out[13]:                        id         created_at   \
         171   1113779196499611649   2019-04-04 12:23:48
         1397  1096443075872858117   2019-02-15 16:16:14
         1169  1100681833451204608   2019-02-27 08:59:33
         69    1115449252837122049   2019-04-09 03:00:00
         2790  1072472676260872193   2018-12-11 12:46:26


         171   The streets of Wayanad, Kerala are filled with cheers for Congress President @Rah
         1397  Congress President @RahulGandhi pays his respects to the fallen soldiers in the t
         1169  With the growing hysteria it is easy to succumb to hate propaganda &amp; fake new
         69    Bollywood has one message to the people of India: Vote karo, magar soch samajh ke
         2790  Democracy has won!\n\nThank you India, you have chosen love over hate, peace over

               retweet_count  favorite_count  photos  videos  url  \
         171    4192           14562           0       1       1
         1397   3700           16606           2       0       1
         1169   4904           15761           0       0       0
         69     5635           16427           0       1       1
         2790   5115           23821           2       0       1


         171   [{u'indices': [200, 219], u'text': u'RahulGandhiWayanad'}, {u'indices': [220, 235
         1397  []
         1169  []
         69    []
         2790  [{u'indices': [139, 155], u'text': u'CongressWinsBIG'}]

In [7]: hash_dict ={}
        for k in df['hashes']:
            if 'text' in k:
                l = k.strip('[')
                l = l.strip(']')
                m = ast.literal_eval(l)
                if (type(m))  == dict:
                    j = (m['text'])
                    if j not in hash_dict.keys():
                        hash_dict[j] = 1
                    else:
                        hash_dict[j] += 1
                else:
                    for h in m:
                        if (type(m))  == dict:
                            j = (m['text'])
                            if j not in hash_dict.keys():
                                hash_dict[j] = 1
                            else:
                                hash_dict[j] += 1
```

```
        print(dict_sort_des(hash_dict)[:5])
```

[('BJPSankalpPatr2019', 139), ('IsBaarNaMoPhirSe', 134), ('IsBaarPhirModi', 121), ('IndiaWithNaM

0.2   Indian National Congress

```
In [8]: df = pd.read_csv('INCIndia_tweets.csv')
        # Filtering the retweets by text.
        df =  df[~df.text.str.contains("RT")]
        df
```

```
Out[8]:                        id            created_at  \
        1     1116282117405626370   2019-04-11 10:09:30
        2     1116269243891191808   2019-04-11 09:18:21
        3     1116265788271259648   2019-04-11 09:04:37
        4     1116259069608599553   2019-04-11 08:37:55
        5     1116249506691969024   2019-04-11 07:59:55
        6     1116247776763867136   2019-04-11 07:53:03
        7     1116244919956332545   2019-04-11 07:41:42
        8     1116241950938546176   2019-04-11 07:29:54
        9     1116240536212348933   2019-04-11 07:24:17
        10    1116233896612392965   2019-04-11 06:57:54
        11    1116170252965093376   2019-04-11 02:45:00
        12    1116166477420187648   2019-04-11 02:30:00
        13    1116030075827179521   2019-04-10 17:27:59
        14    1116006352860610560   2019-04-10 15:53:43
        15    1116000384454090752   2019-04-10 15:30:00
        16    1115991320693108736   2019-04-10 14:53:59
        17    1115977972157046784   2019-04-10 14:00:57
        19    1115957956325990400   2019-04-10 12:41:24
        20    1115957952119107584   2019-04-10 12:41:23
        21    1115957948612665344   2019-04-10 12:41:23
        22    1115957945089417219   2019-04-10 12:41:22
        23    1115957941989806080   2019-04-10 12:41:21
        24    1115957936201711621   2019-04-10 12:41:20
        25    1115957932833656832   2019-04-10 12:41:19
        26    1115957930203828226   2019-04-10 12:41:18
        27    1115943441714651137   2019-04-10 11:43:44
        28    1115941904074539008   2019-04-10 11:37:37
        29    1115941898739384325   2019-04-10 11:37:36
        30    1115941893303603201   2019-04-10 11:37:35
        31    1115941887465152514   2019-04-10 11:37:33
        ...                   ...                   ...
        2786  1072494699938607104   2018-12-11 14:13:57
        2788  1072484262128091136   2018-12-11 13:32:28
        2789  1072478466015141888   2018-12-11 13:09:26
        2790  1072472676260872193   2018-12-11 12:46:26
```

9

```
2791   1072471362818768896   2018-12-11 12:41:13
2792   1072465942104416257   2018-12-11 12:19:40
2793   1072424659398856705   2018-12-11 09:35:38
2794   1072374123052052480   2018-12-11 06:14:49
2795   1072350613390938112   2018-12-11 04:41:24
2796   1072317545909121024   2018-12-11 02:30:00
2797   1072128806444507137   2018-12-10 14:00:01
2798   1072127544269266944   2018-12-10 13:55:00
2799   1072126285940314113   2018-12-10 13:50:00
2800   1072125819751813120   2018-12-10 13:48:09
2801   1072125027632373760   2018-12-10 13:45:00
2802   1072122511905775616   2018-12-10 13:35:00
2803   1072121504777818115   2018-12-10 13:31:00
2804   1072107516517990400   2018-12-10 12:35:25
2805   1072099195266449408   2018-12-10 12:02:21
2806   1072092066019078144   2018-12-10 11:34:01
2807   1072076906433036289   2018-12-10 10:33:47
2808   1072076301178032128   2018-12-10 10:31:23
2809   1072069489473114112   2018-12-10 10:04:19
2810   1072046542486269952   2018-12-10 08:33:08
2811   1072044256972419072   2018-12-10 08:24:03
2812   1072044148721541121   2018-12-10 08:23:37
2813   1072044016248606721   2018-12-10 08:23:05
2814   1072043807674265600   2018-12-10 08:22:16
2815   1072043679924203520   2018-12-10 08:21:45
2816   1072043438084771840   2018-12-10 08:20:47
```

```
1      Congress President @RahulGandhi speaks to the media in Rae Bareli. \n\n#SoniaGandh
2      Is Modi invincible?\nUPA Chairperson Smt. Sonia Gandhi gives a fitting reply to th
3      Four time MP from Rae Bareli, UPA Chairperson Smt. Sonia Gandhi files her nominati
4      UPA Chairperson Smt. Sonia Gandhi at her road show in Rae Bareli before filing her
5      The streets of Rae Bareli are filled with love &amp; support for Smt. Sonia Gandhi
6      Shared moments of a Puja ahead of Smt. Sonia Gandhi's nomination today. \n\n#Sonia
7      UPA Chairperson Smt. Sonia Gandhi, Congress President @RahulGandhi &amp; GS I/c UP
8      Hundreds of people gather in Rae Bareli today to share their love &amp; support fc
9      Listen to the song by @RaghavMeattle that perfectly captures the question on every
10     Voting is the right of every citizen, no matter creed or caste. We strongly conden
11         -              ,     ,      ,          \n#MyVoteForCongress #AbHogaNyay https://
12     Today you decide\nLove over hate\nJobs over pakoda\nPolicies over propoganda\nOne
13     The triumph of truth begins.\nThe time for jumlas are over.\nExhibit A - https://t
14         - ,    ,\n   ,  \n    ,   ,\n    , ""   \n#AbHogaNYAY https://t.co/JOQDZjDtlN
15     Thank you India for sending in your entries for NYAY. Now it's your time to vote f
16     "           ,  ,     ,   " :   @RahulGandhi\n#JanSankalpRally #ChowkidarChorHai h
17     "I challenge Narendra Modi to debate me on corruption": Congress President @RahulG
19     .@narendramodi wants two Hindustans: one for the rich &amp; one for the poor. I wi
20               ,   -                       ' '     : @RahulGandhi\n#BengalWithRahulGandhi
21     In the last 70 years, Congress party has never gotten into an alliance with the BJ
```

```
22     I said in Rajasthan, MP and Chhattisgarh that if Modi can waive lakhs of crores of
23        22    ,          10      ,    ,    :   @RahulGandhi\n#BengalWithRahulG
24                      ,                  :   @RahulGandhi\n#BengalWithRahulGandhi
25     Did anyone here receive 15 lakhs in your account? We have discussed with leading e
26        ""    12,000     20%    72,000    :   @RahulGandhi\n#BengalWithRahulGa
27     LIVE: Congress President @RahulGandhi addresses public meeting in Raiganj, West Be
28        ,                     :   @RahulGandhi\n#JanSankalpRally
29           ,            ,    , ,  ,    :   @RahulGandhi\n#JanSankalpRally
30     CBI Director was removed at midnight after he started an inquiry into the Rafale d
31              -       :   @RahulGandhi\n#JanSankalpRally
...
2786   LIVE: Congress President @RahulGandhi addresses the media on Assembly Elections. #
2788      \n ...   ;          \n    ...\n\n#CongressWinsBIG https://t.co/mCQJcobOLk
2789      ,    ,        \n\n#CongressWinsBIG https://t.co/DvDTS5R467
2790   Democracy has won!\n\nThank you India, you have chosen love over hate, peace over
2791                          \n\n#CongressWinsBIG https://t.co/dNJOX1Bf7R
2792                  \n\n#CongressWinsBIG https://t.co/vGelK3s3Bi
2793   Very unfortunate and is a severe blow to the nations economy: Former PM Dr. Manmoh
2794   During 2008's global recession, India was one of the only countries to stay afloat
2795   We wish Former President, Shri. Pranab Mukherjee a Happy Birthday! \nWith a politi
2796   Subramania Bharati was a poet and a writer from Tamil Nadu. He was a social reform
2797   It is clear to the entire country, the resignation of RBI Governor Urjit Patel is
2798   There was consensus in our room that BJPs corruption on Rafale, Demo &amp; other a
2799   The govt. is getting more &amp; more desperate. The govt. is taking measures that
2800   Taking away the reserves from the RBII to save your skin to atone for the sins you
2801   Former RBI Governor Raghuram Rajan is right, we should all be concerned by the gov
2802   Were not going to tolerate what is happening in the CBI, RBI &amp; every other ins
2803   We have to stop the BJPs assault on our institutions, stop the assault on the CBI,
2804   ,     ""      ...      | https://t.co/dUCOkYpmQ3
2805   Another one bites the dust. \n\nThis is the result of our 'chowkidar's' assault on
2806   Change is coming. \n\nUPA Chairperson Smt. Sonia Gandhi, Former PM Dr. Manmohan Si
2807      \n      https://t.co/qQ4rxzHU17
2808      |\n      || https://t.co/7j14T8XLk3
2809   Former Prime Minister Dr Manmohan Singh and Congress President @RahulGandhi and ot
2810   In a political climate where citizens are arrested, lynched or humiliated for expr
2811                  2019    :   @RahulGandhi\n#__ https://t.co/Y72dSP9yu4
2812   Institutions are encroached by Modi govt; Supreme Court is not allowed to function
2813            :   @RahulGandhi \n#__ https://t.co/mypsv4sK7P
2814   Defeating BJP in elections is not our ultimate goal. Our ultimate goal is solve fa
2815   Anger of our nation is rising. This anger is used by BJP-RSS for their political g
2816            ,     -      :   @RahulGandhi\n#__ https://t.co/ISfCYZ4Plo


    retweet_count  favorite_count  photos  videos  url  \
1   791            2518            0       1       1
2   727            2085            0       1       1
3   495            1737            2       0       1
4   527            1814            2       0       1
5   955            2749            0       1       1
```

6	987	4526	2	0	1
7	1522	5702	0	1	1
8	476	1465	2	0	1
9	417	954	0	1	1
10	376	831	2	0	1
11	901	2531	2	0	1
12	2192	6702	2	0	1
13	1685	4292	0	1	1
14	758	2048	0	1	1
15	887	2564	0	0	1
16	1910	6586	0	1	1
17	1701	5847	0	1	1
19	550	1957	0	0	0
20	640	2419	0	0	0
21	212	402	0	0	0
22	172	274	0	0	0
23	178	264	0	0	0
24	162	244	0	0	0
25	170	267	0	0	0
26	162	248	0	0	0
27	911	2936	0	0	1
28	733	2922	0	0	0
29	836	3292	0	0	0
30	140	254	0	0	0
31	126	236	0	0	0
...	...	...	..	..	..
2786	2176	7609	0	0	1
2788	635	2351	2	0	1
2789	3131	10644	0	1	1
2790	5115	23821	2	0	1
2791	696	2098	0	1	1
2792	846	3300	0	1	1
2793	1206	4764	0	0	1
2794	831	2265	0	1	1
2795	530	2362	2	0	1
2796	293	906	2	0	1
2797	548	1415	2	0	1
2798	819	2625	2	0	1
2799	610	1544	2	0	1
2800	496	1242	2	0	1
2801	569	1566	2	0	1
2802	2862	9807	0	1	1
2803	1204	3433	0	1	1
2804	562	1375	2	0	1
2805	877	2236	0	0	1
2806	718	2227	2	0	1
2807	245	700	2	0	1
2808	542	1719	2	0	1

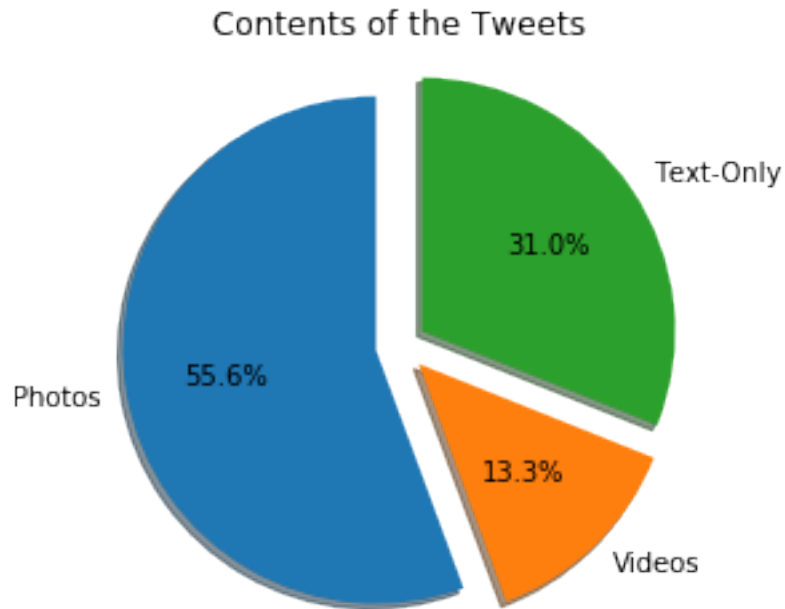|					
2809	423	1292	2	0	1
2810	421	1217	2	0	1
2811	359	943	2	0	1
2812	265	681	0	0	0
2813	184	383	2	0	1
2814	183	311	2	0	1
2815	154	260	0	0	0
2816	234	457	2	0	1

```
1       [{u'indices': [69, 90], u'text': u'SoniaGandhiRaeBareli'}]
2       [{u'indices': [134, 155], u'text': u'SoniaGandhiRaeBareli'}]
3       [{u'indices': [121, 142], u'text': u'SoniaGandhiRaeBareli'}]
4       [{u'indices': [96, 117], u'text': u'SoniaGandhiRaeBareli'}]
5       [{u'indices': [99, 120], u'text': u'SoniaGandhiRaeBareli'}]
6       [{u'indices': [74, 95], u'text': u'SoniaGandhiRaeBareli'}]
7       [{u'indices': [178, 199], u'text': u'SoniaGandhiRaeBareli'}]
8       []
9       []
10      []
11      [{u'indices': [190, 208], u'text': u'MyVoteForCongress'}, {u'indices': [209, 220],
12      [{u'indices': [151, 169], u'text': u'MyVoteForCongress'}, {u'indices': [170, 181],
13      []
14      [{u'indices': [170, 181], u'text': u'AbHogaNYAY'}]
15      [{u'indices': [131, 142], u'text': u'AbHogaNYAY'}]
16      [{u'indices': [164, 180], u'text': u'JanSankalpRally'}, {u'indices': [181, 198], u
17      [{u'indices': [88, 104], u'text': u'JanSankalpRally'}, {u'indices': [105, 119], u'
19      [{u'indices': [262, 284], u'text': u'BengalWithRahulGandhi'}]
20      [{u'indices': [226, 248], u'text': u'BengalWithRahulGandhi'}]
21      [{u'indices': [245, 267], u'text': u'BengalWithRahulGandhi'}]
22      [{u'indices': [254, 276], u'text': u'BengalWithRahulGandhi'}]
23      [{u'indices': [218, 240], u'text': u'BengalWithRahulGandhi'}]
24      [{u'indices': [218, 240], u'text': u'BengalWithRahulGandhi'}]
25      [{u'indices': [230, 252], u'text': u'BengalWithRahulGandhi'}]
26      [{u'indices': [175, 197], u'text': u'BengalWithRahulGandhi'}]
27      [{u'indices': [88, 110], u'text': u'BengalWithRahulGandhi'}]
28      [{u'indices': [170, 186], u'text': u'JanSankalpRally'}]
29      [{u'indices': [247, 263], u'text': u'JanSankalpRally'}]
30      [{u'indices': [197, 213], u'text': u'JanSankalpRally'}]
31      [{u'indices': [137, 153], u'text': u'JanSankalpRally'}]
...                                                      ...
2786    [{u'indices': [81, 97], u'text': u'CongressWinsBIG'}]
2788    [{u'indices': [132, 148], u'text': u'CongressWinsBIG'}]
2789    [{u'indices': [65, 81], u'text': u'CongressWinsBIG'}]
2790    [{u'indices': [139, 155], u'text': u'CongressWinsBIG'}]
2791    [{u'indices': [182, 198], u'text': u'CongressWinsBIG'}]
2792    [{u'indices': [100, 116], u'text': u'CongressWinsBIG'}]
2793    [{u'indices': [125, 136], u'text': u'UrjitPatel'}]
```

```
2794  []
2795  []
2796  []
2797  [{u'indices': [134, 150], u'text': u'ModiDestroysRBI'}]
2798  [{u'indices': [209, 225], u'text': u'ModiDestroysRBI'}]
2799  [{u'indices': [262, 278], u'text': u'ModiDestroysRBI'}]
2800  [{u'indices': [213, 229], u'text': u'ModiDestroysRBI'}]
2801  [{u'indices': [139, 155], u'text': u'ModiDestroysRBI'}]
2802  [{u'indices': [181, 197], u'text': u'ModiDestroysRBI'}]
2803  [{u'indices': [173, 189], u'text': u'ModiDestroysRBI'}]
2804  []
2805  []
2806  []
2807  []
2808  []
2809  []
2810  [{u'indices': [215, 230], u'text': u'HumanRightsDay'}]
2811  [{u'indices': [147, 163], u'text': u'\u0917\u093e\u0902\u0927\u0940_\u0915\u093e_\
2812  [{u'indices': [232, 248], u'text': u'\u0917\u093e\u0902\u0927\u0940_\u0915\u093e_\
2813  [{u'indices': [98, 114], u'text': u'\u0917\u093e\u0902\u0927\u0940_\u0915\u093e_\u
2814  []
2815  []
2816  [{u'indices': [179, 195], u'text': u'\u0917\u093e\u0902\u0927\u0940_\u0915\u093e_\

[2661 rows x 9 columns]
```

```python
In [9]: p_count = 0
        for k in df['photos']:
            if k>0:
                p_count += 1
        print(str(p_count) + ' tweets consist of images.')

        v_count = 0
        for k in df['videos']:
            if k>0:
                v_count += 1

        k = df[(df['photos']==0)][(df['videos']==0)]
        ot_count =  len(k)
        labels = ['Photos','Videos','Text-Only' ]
        sizes = [p_count, v_count, ot_count]
        explode = (0.1, 0.1, 0.1)

        fig1, ax1 = plt.subplots()
        ax1.pie(sizes, explode=explode, labels=labels, autopct='%1.1f%%',
                shadow=True, startangle=90)
        ax1.axis('equal')
        ax1.set_title('Contents of the Tweets')
```

14

```
    plt.show()
```

1480 tweets consist of images.

Contents of the Tweets


```
In [10]: # k = df[(df['retweet_count']*0.7 + df['favorite_count']*0.3)==(df['retweet_count']*0.7
         j = df.assign(f = df['retweet_count']*0.7 + df['favorite_count']*0.3).sort_values('f').
         j = j.tail(1000)

In [11]: tweet_words = []
         for k in j['text']:
             tweet_words.append(clean_tweet(k))
         tweet_words = list(filter(None, tweet_words))
         wordcloud_plot(tweet_words, 'black')
```

The words that occur most are: 1. Congress President - Rahul Gandhi, who is an important Prime Ministerial candidate and the president of Congress party. 2. Modi - The popular potential and ex Prime Minister Narendra Modi. 3. Rafale - The deal which is a contraversial aspect of the ruling party.

```
In [12]: hash_dict ={}
         for k in df['hashes']:
            if 'text' in k:
                l = k.strip('[')
                l = l.strip(']')
                m = ast.literal_eval(l)
                if (type(m))  == dict:
                    j = (m['text'])
                    if j not in hash_dict.keys():
                        hash_dict[j] = 1
                    else:
                        hash_dict[j] += 1
                else:
                    for h in m:
                        if (type(m))  == dict:
                            j = (m['text'])
                            if j not in hash_dict.keys():
                                hash_dict[j] = 1
                            else:
```

16

```
                            hash_dict[j] += 1
            print(dict_sort_des(hash_dict)[:5])
```

[('NammaRahulGandhi', 56), ('TelanganasProgressWithCongress', 45), ('YoungIndiaRising', 43), ('V