# Final Project
(ID2221)

Chuting Zhu
chuting@kth.se

Suhas Sheshadri
suhass@kth.se

November 4, 2018

## 1 Topic

We implement sentiment analysis on ”Sweden’s football journey during the FIFA World Cup-2018” based on the tweets collected from Twitter. Our project shows the sentiments of people during the whole world cup vs when Sweden lost matches. And also, the Term Frequency (tf) and Document Frequency (df) for all the words form the tweets collected.

## 2 Dataset

We collect data from Twitter for a given search query and date range. Search queries would be relevant to Sweden and Football and we use ”Sweden football” as the query to get related tweets. The date range would be the duration of the world cup tournament, which was from 14 June to 15 July 2018.

We also collected data separately for the days of 23 June and 7 July when Sweden lost the matches on these two days and combine them into a single file.

There are two files, one is $AllTweets.txt$ corresponding to the tweets relevant to ”Sweden football” during the whole world cup, the other is $LossTweets.txt$ which are tweets related when Sweden lost.

## 3 Methods

- Collect relevant data from Twitter.

  Since TwitteR API allows us to collect the tweets only from last 7 days, we had to use a different scrapping method. We have used the GetOldTweets implemented by Jefferson-Henrique for scrapping old tweets. We have modified it to suit our need. [2]

  The above obtained data are in .CSV format. We clean the data and convert the files to .TXT format.

- Calculate the sentiment score of the collected tweets. The sentiment score of a tweet is equivalent to the sum of the sentiment scores for each term in the tweet. And we check all the tweets for words which are present in $AFINN - 111$ file. [1]

  AFINN is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive). The words have been manually labeled by Finn rup Nielsen in 2009-2011.

  If the word is present in this file, we will add the word’s score obtained from the file to the sentiment score of the corresponding tweet. Moreover, a positive sentiment has a positive score and vice versa.

- Utilize MapReduce Framework to calculate the sentiment score for each tweet.

- Utilize MapReduce Framework to calculate the term frequency and document frequency of all the words from the collected tweets.

# 4 Run the application

1. Analyze the sentiment of all tweets during the whole world cup.

```
python tweets_sentiment.py AFINN-111.txt AllTweets.txt
```

2. Analyze the TF-DF for AllTweets.txt file

```
python tfdf.py AllTweets.txt
```

3. Analyze the sentiment when Sweden lost.

```
python tweets_sentiment.py AFINN-111.txt LossTweets.txt
```

4. Analyze the TF-DF for LossTweets.txt file

```
python tfdf.py LossTweets.txt
```

# 5 Results

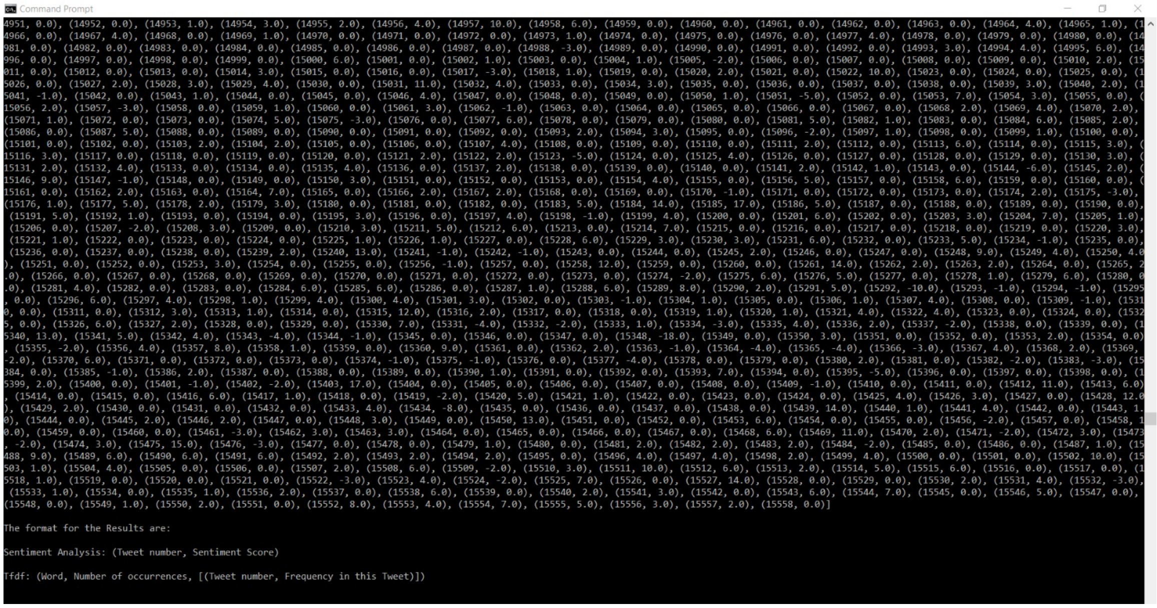
The Results are in the format:

Sentiment Analysis: (Tweet number, Sentiment Score)

For reference:
[1, 2.0] Sentiment score of tweet 1 is 2
[5, 0] Sentiment score of tweet 5 is 0

Tfdf: (Word, Number of occurrences, [(Tweet number, Frequency in this Tweet)])

For reference:
[”hello”, 4, [[1, 1], [2, 3], [6, 1], [9, 2]]] This shows that term hello has df 4 and occurs in tweets 1, 2, 6, 9 with tf 1, 3, 1, 2 respectively.
[”world”, 1, [[3, 1]]] This shows that term world has df 1 and occurs in tweet 3 with tf 1.



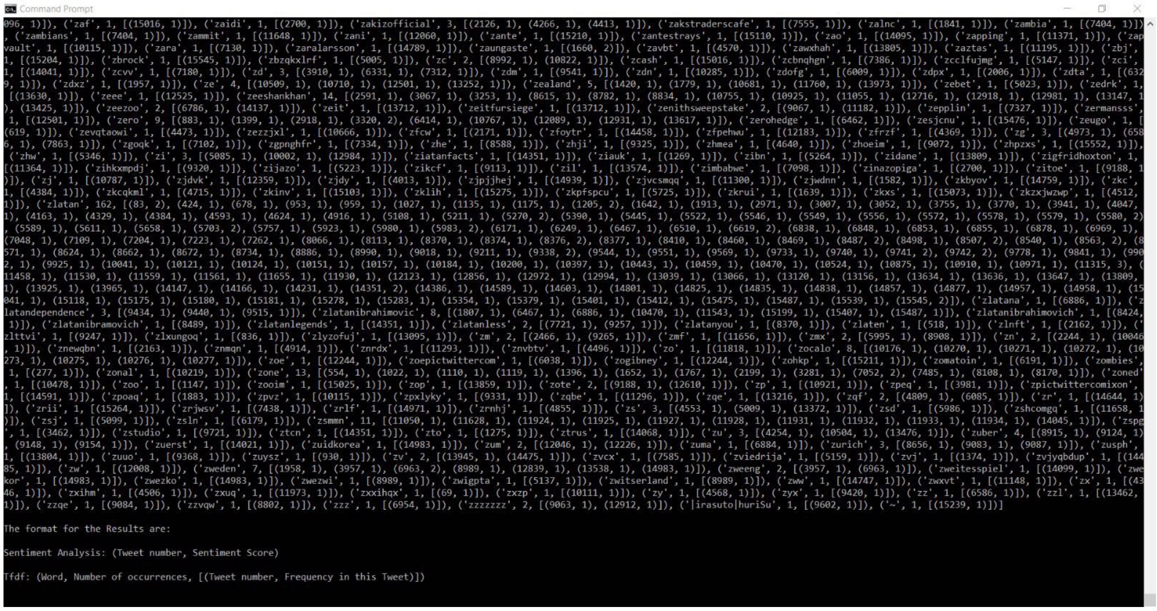
```
4951, 0.0), (14952, 0.0), (14953, 1.0), (14954, 3.0), (14955, 2.0), (14956, 4.0), (14957, 10.0), (14958, 6.0), (14959, 0.0), (14960, 0.0), (14961, 0.0), (14962, 0.0), (14963, 0.0), (14964, 4.0), (14965, 1.0), (1
4966, 0.0), (14967, 4.0), (14968, 0.0), (14969, 1.0), (14970, 0.0), (14971, 0.0), (14972, 0.0), (14973, 1.0), (14974, 0.0), (14975, 0.0), (14976, 0.0), (14977, 4.0), (14978, 0.0), (14979, 0.0), (14980, 0.0), (14
981, 0.0), (14982, 0.0), (14983, 0.0), (14984, 0.0), (14985, 0.0), (14986, 0.0), (14987, 0.0), (14988, -3.0), (14989, 0.0), (14990, 0.0), (14991, 0.0), (14992, 0.0), (14993, 3.0), (14994, 4.0), (14995, 6.0), (14
996, 0.0), (14997, 0.0), (14998, 0.0), (14999, 0.0), (15000, 6.0), (15001, 0.0), (15002, 1.0), (15003, 0.0), (15004, 1.0), (15005, -2.0), (15006, 0.0), (15007, 0.0), (15008, 0.0), (15009, 0.0), (15010, 2.0), (15
011, 0.0), (15012, 0.0), (15013, 0.0), (15014, 3.0), (15015, 0.0), (15016, 0.0), (15017, -3.0), (15018, 1.0), (15019, 0.0), (15020, 2.0), (15021, 0.0), (15022, 10.0), (15023, 0.0), (15024, 0.0), (15025, 0.0), (1
5026, 0.0), (15027, 2.0), (15028, 3.0), (15029, 4.0), (15030, 0.0), (15031, 11.0), (15032, 4.0), (15033, 0.0), (15034, 3.0), (15035, 0.0), (15036, 0.0), (15037, 0.0), (15038, 0.0), (15039, 3.0), (15040, 2.0), (1
5041, -1.0), (15042, 0.0), (15043, 1.0), (15044, 0.0), (15045, 0.0), (15046, 4.0), (15047, 0.0), (15048, 0.0), (15049, 0.0), (15050, 1.0), (15051, -5.0), (15052, 0.0), (15053, 7.0), (15054, 3.0), (15055, 0.0), (
15056, 2.0), (15057, -3.0), (15058, 0.0), (15059, 1.0), (15060, 0.0), (15061, 3.0), (15062, -1.0), (15063, 0.0), (15064, 0.0), (15065, 0.0), (15066, 0.0), (15067, 0.0), (15068, 2.0), (15069, 4.0), (15070, 2.0),
(15071, 1.0), (15072, 0.0), (15073, 0.0), (15074, 5.0), (15075, -3.0), (15076, 0.0), (15077, 6.0), (15078, 0.0), (15079, 0.0), (15080, 0.0), (15081, 5.0), (15082, 1.0), (15083, 0.0), (15084, 6.0), (15085, 2.0),
(15086, 0.0), (15087, 5.0), (15088, 0.0), (15089, 0.0), (15090, 0.0), (15091, 0.0), (15092, 0.0), (15093, 2.0), (15094, 3.0), (15095, 0.0), (15096, -2.0), (15097, 1.0), (15098, 0.0), (15099, 1.0), (15100, 0.0),
(15101, 0.0), (15102, 0.0), (15103, 2.0), (15104, 2.0), (15105, 0.0), (15106, 0.0), (15107, 4.0), (15108, 0.0), (15109, 0.0), (15110, 0.0), (15111, 2.0), (15112, 0.0), (15113, 6.0), (15114, 0.0), (15115, 3.0), (
15116, 3.0), (15117, 0.0), (15118, 0.0), (15119, 0.0), (15120, 0.0), (15121, 2.0), (15122, 2.0), (15123, -5.0), (15124, 0.0), (15125, 4.0), (15126, 0.0), (15127, 0.0), (15128, 0.0), (15129, 0.0), (15130, 3.0), (
15131, 2.0), (15132, 4.0), (15133, 0.0), (15134, 0.0), (15135, 4.0), (15136, 0.0), (15137, 2.0), (15138, 0.0), (15139, 0.0), (15140, 0.0), (15141, 2.0), (15142, 1.0), (15143, 0.0), (15144, -6.0), (15145, 2.0), (
15146, 9.0), (15147, -1.0), (15148, 0.0), (15149, 0.0), (15150, 3.0), (15151, 0.0), (15152, 0.0), (15153, 0.0), (15154, 4.0), (15155, 0.0), (15156, 5.0), (15157, 0.0), (15158, 6.0), (15159, 0.0), (15160, 0.0), (
15161, 0.0), (15162, 2.0), (15163, 0.0), (15164, 7.0), (15165, 0.0), (15166, 2.0), (15167, 2.0), (15168, 0.0), (15169, 0.0), (15170, -1.0), (15171, 0.0), (15172, 0.0), (15173, 0.0), (15174, 2.0), (15175, -3.0),
(15176, 1.0), (15177, 5.0), (15178, 2.0), (15179, 3.0), (15180, 0.0), (15181, 0.0), (15182, 0.0), (15183, 5.0), (15184, 14.0), (15185, 17.0), (15186, 5.0), (15187, 0.0), (15188, 0.0), (15189, 0.0), (15190, 0.0),
 (15191, 5.0), (15192, 1.0), (15193, 0.0), (15194, 0.0), (15195, 3.0), (15196, 0.0), (15197, 4.0), (15198, -1.0), (15199, 4.0), (15200, 0.0), (15201, 6.0), (15202, 0.0), (15203, 3.0), (15204, 7.0), (15205, 1.0),
 (15206, 0.0), (15207, -2.0), (15208, 3.0), (15209, 0.0), (15210, 3.0), (15211, 5.0), (15212, 6.0), (15213, 0.0), (15214, 7.0), (15215, 0.0), (15216, 0.0), (15217, 0.0), (15218, 0.0), (15219, 0.0), (15220, 3.0),
 (15221, 1.0), (15222, 0.0), (15223, 0.0), (15224, 0.0), (15225, 1.0), (15226, 1.0), (15227, 0.0), (15228, 6.0), (15229, 3.0), (15230, 3.0), (15231, 6.0), (15232, 0.0), (15233, 5.0), (15234, -1.0), (15235, 0.0),
 (15236, 0.0), (15237, 0.0), (15238, 0.0), (15239, 2.0), (15240, 13.0), (15241, -1.0), (15242, -1.0), (15243, 0.0), (15244, 0.0), (15245, 2.0), (15246, 0.0), (15247, 0.0), (15248, 9.0), (15249, 4.0), (15250, 4.0
), (15251, 0.0), (15252, 0.0), (15253, 3.0), (15254, 0.0), (15255, 0.0), (15256, -1.0), (15257, 0.0), (15258, 12.0), (15259, 0.0), (15260, 0.0), (15261, 14.0), (15262, 2.0), (15263, 2.0), (15264, 0.0), (15265, 2
.0), (15266, 0.0), (15267, 0.0), (15268, 0.0), (15269, 0.0), (15270, 0.0), (15271, 0.0), (15272, 0.0), (15273, 0.0), (15274, -2.0), (15275, 6.0), (15276, 5.0), (15277, 0.0), (15278, 1.0), (15279, 6.0), (15280, 0
.0), (15281, 4.0), (15282, 0.0), (15283, 0.0), (15284, 6.0), (15285, 6.0), (15286, 0.0), (15287, 1.0), (15288, 6.0), (15289, 8.0), (15290, 2.0), (15291, 5.0), (15292, -10.0), (15293, -1.0), (15294, -1.0), (15295
, 0.0), (15296, 6.0), (15297, 4.0), (15298, 1.0), (15299, 4.0), (15300, 4.0), (15301, 3.0), (15302, 0.0), (15303, -1.0), (15304, 1.0), (15305, 0.0), (15306, 1.0), (15307, 4.0), (15308, 0.0), (15309, -1.0), (1531
0, 0.0), (15311, 0.0), (15312, 3.0), (15313, 1.0), (15314, 0.0), (15315, 12.0), (15316, 2.0), (15317, 0.0), (15318, 0.0), (15319, 1.0), (15320, 1.0), (15321, 4.0), (15322, 4.0), (15323, 0.0), (15324, 0.0), (1532
5, 0.0), (15326, 6.0), (15327, 2.0), (15328, 0.0), (15329, 0.0), (15330, 7.0), (15331, -4.0), (15332, -2.0), (15333, 1.0), (15334, -3.0), (15335, 4.0), (15336, 2.0), (15337, -2.0), (15338, 0.0), (15339, 0.0), (1
5340, 13.0), (15341, 5.0), (15342, 4.0), (15343, -4.0), (15344, -1.0), (15345, 0.0), (15346, 0.0), (15347, 0.0), (15348, -18.0), (15349, 0.0), (15350, 3.0), (15351, 0.0), (15352, 0.0), (15353, 2.0), (15354, 0.0)
, (15355, -2.0), (15356, 4.0), (15357, 8.0), (15358, 1.0), (15359, 0.0), (15360, 9.0), (15361, 0.0), (15362, 2.0), (15363, -1.0), (15364, -4.0), (15365, -4.0), (15366, -3.0), (15367, 4.0), (15368, 2.0), (15369,
-2.0), (15370, 6.0), (15371, 0.0), (15372, 0.0), (15373, 0.0), (15374, -1.0), (15375, -1.0), (15376, 0.0), (15377, -4.0), (15378, 0.0), (15379, 0.0), (15380, 2.0), (15381, 0.0), (15382, -2.0), (15383, -3.0), (15
384, 0.0), (15385, -1.0), (15386, 2.0), (15387, 0.0), (15388, 0.0), (15389, 0.0), (15390, 1.0), (15391, 0.0), (15392, 0.0), (15393, 7.0), (15394, 0.0), (15395, -5.0), (15396, 0.0), (15397, 0.0), (15398, 0.0), (1
5399, 2.0), (15400, 0.0), (15401, -1.0), (15402, -2.0), (15403, 17.0), (15404, 0.0), (15405, 0.0), (15406, 0.0), (15407, 0.0), (15408, 0.0), (15409, -1.0), (15410, 0.0), (15411, 0.0), (15412, 11.0), (15413, 6.0)
, (15414, 0.0), (15415, 0.0), (15416, 6.0), (15417, 1.0), (15418, 0.0), (15419, -2.0), (15420, 5.0), (15421, 1.0), (15422, 0.0), (15423, 0.0), (15424, 0.0), (15425, 4.0), (15426, 3.0), (15427, 0.0), (15428, 12.0
), (15429, 2.0), (15430, 0.0), (15431, 0.0), (15432, 0.0), (15433, 4.0), (15434, -8.0), (15435, 0.0), (15436, 0.0), (15437, 0.0), (15438, 0.0), (15439, 14.0), (15440, 1.0), (15441, 4.0), (15442, 0.0), (15443, 1.
0), (15444, 0.0), (15445, 2.0), (15446, 2.0), (15447, 0.0), (15448, 3.0), (15449, 0.0), (15450, 13.0), (15451, 0.0), (15452, 0.0), (15453, 6.0), (15454, 0.0), (15455, 0.0), (15456, -2.0), (15457, 0.0), (15458, 1
0.0), (15459, 0.0), (15460, 0.0), (15461, -3.0), (15462, 3.0), (15463, 3.0), (15464, 0.0), (15465, 0.0), (15466, 0.0), (15467, 0.0), (15468, 6.0), (15469, 11.0), (15470, 2.0), (15471, -2.0), (15472, 3.0), (15473
, -2.0), (15474, 3.0), (15475, 15.0), (15476, -3.0), (15477, 0.0), (15478, 0.0), (15479, 1.0), (15480, 0.0), (15481, 2.0), (15482, 2.0), (15483, 2.0), (15484, -2.0), (15485, 0.0), (15486, 0.0), (15487, 1.0), (15
488, 9.0), (15489, 6.0), (15490, 6.0), (15491, 6.0), (15492, 2.0), (15493, 2.0), (15494, 2.0), (15495, 0.0), (15496, 4.0), (15497, 4.0), (15498, 2.0), (15499, 4.0), (15500, 0.0), (15501, 0.0), (15502, 10.0), (15
503, 1.0), (15504, 4.0), (15505, 0.0), (15506, 0.0), (15507, 2.0), (15508, 6.0), (15509, -2.0), (15510, 3.0), (15511, 10.0), (15512, 6.0), (15513, 2.0), (15514, 5.0), (15515, 6.0), (15516, 0.0), (15517, 0.0), (1
5518, 1.0), (15519, 0.0), (15520, 0.0), (15521, 0.0), (15522, -3.0), (15523, 4.0), (15524, -2.0), (15525, 7.0), (15526, 0.0), (15527, 14.0), (15528, 0.0), (15529, 0.0), (15530, 2.0), (15531, 4.0), (15532, -3.0),
 (15533, 1.0), (15534, 0.0), (15535, 1.0), (15536, 2.0), (15537, 0.0), (15538, 6.0), (15539, 0.0), (15540, 2.0), (15541, 3.0), (15542, 0.0), (15543, 6.0), (15544, 7.0), (15545, 0.0), (15546, 5.0), (15547, 0.0),
(15548, 0.0), (15549, 1.0), (15550, 2.0), (15551, 0.0), (15552, 8.0), (15553, 4.0), (15554, 7.0), (15555, 5.0), (15556, 3.0), (15557, 2.0), (15558, 0.0)]

The format for the Results are:

Sentiment Analysis: (Tweet number, Sentiment Score)

Tfdf: (Word, Number of occurrences, [(Tweet number, Frequency in this Tweet)])
```

Figure 1: Sentiment Analysis for AllTweets.txt



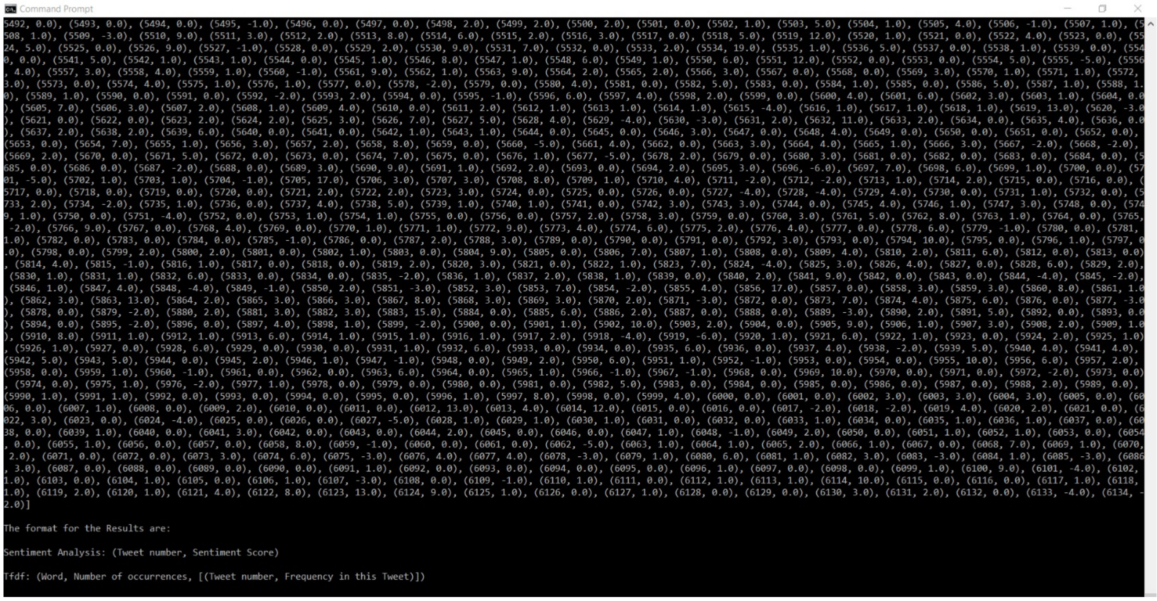
```
096, 1)]), ('zaf', 1, [(15016, 1)]), ('zaidi', 1, [(2700, 1)]), ('zakizofficial', 3, [(2126, 1), (4266, 1), (4413, 1)]), ('zakstraderscafe', 1, [(7555, 1)]), ('zalnc', 1, [(1841, 1)]), ('zambia', 1, [(7404, 1)])
, ('zambians', 1, [(7404, 1)]), ('zammit', 1, [(11648, 1)]), ('zani', 1, [(12060, 1)]), ('zante', 1, [(15210, 1)]), ('zantestrays', 1, [(15110, 1)]), ('zao', 1, [(14095, 1)]), ('zapping', 1, [(11371, 1)]), ('zap
vault', 1, [(10115, 1)]), ('zara', 1, [(7130, 1)]), ('zaralarsson', 1, [(14789, 1)]), ('zaungaste', 1, [(1660, 2)]), ('zavbt', 1, [(4570, 1)]), ('zawxhah', 1, [(13805, 1)]), ('zaztas', 1, [(11195, 1)]), ('zbj',
1, [(15204, 1)]), ('zbrock', 1, [(15545, 1)]), ('zbzqkxlrf', 1, [(5005, 1)]), ('zc', 2, [(8992, 1), (10822, 1)]), ('zcash', 1, [(15016, 1)]), ('zcbnqhgn', 1, [(7386, 1)]), ('zcclfujmg', 1, [(5147, 1)]), ('zci',
1, [(14041, 1)]), ('zcvv', 1, [(7180, 1)]), ('zd', 3, [(3910, 1), (6331, 1), (7312, 1)]), ('zdm', 1, [(9541, 1)]), ('zdn', 1, [(10285, 1)]), ('zdofg', 1, [(6009, 1)]), ('zdpx', 1, [(2006, 1)]), ('zdta', 1, [(632
9, 1)]), ('zdxz', 1, [(1957, 1)]), ('ze', 4, [(10509, 1), (10710, 1), (12501, 1), (13252, 1)]), ('zealand', 5, [(1420, 1), (1779, 1), (10681, 1), (11760, 1), (13973, 1)]), ('zebet', 1, [(5023, 1)]), ('zedrk', 1,
 [(13630, 1)]), ('zeee', 1, [(12525, 1)]), ('zeeshankhan', 14, [(2591, 1), (3067, 1), (3253, 1), (8615, 1), (8782, 1), (8834, 1), (10755, 1), (10925, 1), (11055, 1), (12716, 1), (12918, 1), (12981, 1), (13147, 1
), (13425, 1)]), ('zeezoo', 2, [(6786, 1), (14137, 1)]), ('zeit', 1, [(13712, 1)]), ('zeitfursiege', 1, [(13712, 1)]), ('zenithsweepstake', 2, [(9067, 1), (11182, 1)]), ('zepplin', 1, [(7327, 1)]), ('zermansss',
 1, [(12501, 1)]), ('zero', 9, [(883, 1), (1399, 1), (2918, 1), (3320, 2), (6414, 1), (10767, 1), (12089, 1), (12931, 1), (13617, 1)]), ('zerohedge', 1, [(6462, 1)]), ('zesjcnu', 1, [(15476, 1)]), ('zeugo', 1, [
(619, 1)]), ('zevqtaowi', 1, [(4473, 1)]), ('zezzjxl', 1, [(10666, 1)]), ('zfcw', 1, [(2171, 1)]), ('zfoytr', 1, [(14458, 1)]), ('zfpehwu', 1, [(12183, 1)]), ('zfrzf', 1, [(4369, 1)]), ('zg', 3, [(4973, 1), (658
6, 1), (7863, 1)]), ('zgoqk', 1, [(7102, 1)]), ('zgpnghfr', 1, [(7334, 1)]), ('zhe', 1, [(8588, 1)]), ('zhji', 1, [(9325, 1)]), ('zhmea', 1, [(4640, 1)]), ('zhoeim', 1, [(9072, 1)]), ('zhpzxs', 1, [(15552, 1)]),
 ('zhw', 1, [(5346, 1)]), ('zi', 3, [(5085, 1), (10002, 1), (12984, 1)]), ('ziatanfacts', 1, [(14351, 1)]), ('ziauk', 1, [(1269, 1)]), ('zibn', 1, [(5264, 1)]), ('zidane', 1, [(13809, 1)]), ('zigfridhoxton', 1,
[(11364, 1)]), ('zihkxmpdj', 1, [(9320, 1)]), ('zijazo', 1, [(5223, 1)]), ('zikcf', 1, [(9113, 1)]), ('zil', 1, [(13574, 1)]), ('zimbabwe', 1, [(7098, 1)]), ('zinazopiga', 1, [(2700, 1)]), ('zitoe', 1, [(9188, 1
)]), ('zj', 1, [(10787, 1)]), ('zjdvk', 1, [(12359, 1)]), ('zjdy', 1, [(4013, 1)]), ('zjpjjhej', 1, [(14939, 1)]), ('zjvcsmqq', 1, [(11300, 1)]), ('zjwdnn', 1, [(1582, 1)]), ('zkbyov', 1, [(14759, 1)]), ('zkc',
1, [(4384, 1)]), ('zkcqkml', 1, [(4715, 1)]), ('zkinv', 1, [(15103, 1)]), ('zklih', 1, [(15275, 1)]), ('zkpfspcu', 1, [(5725, 1)]), ('zkrui', 1, [(1639, 1)]), ('zkxs', 1, [(15073, 1)]), ('zkzxjwzwp', 1, [(4512,
1)]), ('zlatan', 162, [(83, 2), (424, 1), (678, 1), (953, 1), (959, 1), (1027, 1), (1135, 1), (1175, 1), (1205, 2), (1642, 1), (1913, 1), (2971, 1), (3007, 1), (3052, 1), (3755, 1), (3770, 1), (3941, 1), (4047,
1), (4163, 1), (4329, 1), (4384, 1), (4593, 1), (4624, 1), (4916, 1), (5108, 1), (5211, 1), (5270, 2), (5390, 1), (5445, 1), (5522, 1), (5546, 1), (5549, 1), (5556, 1), (5572, 1), (5578, 1), (5579, 1), (5580, 2)
, (5589, 1), (5611, 1), (5658, 1), (5703, 2), (5757, 1), (5923, 1), (5980, 1), (5983, 2), (6171, 1), (6249, 1), (6467, 1), (6510, 1), (6619, 2), (6838, 1), (6848, 1), (6853, 1), (6855, 1), (6878, 1), (6969, 1),
(7048, 1), (7109, 1), (7204, 1), (7223, 1), (7262, 1), (8066, 1), (8113, 1), (8370, 1), (8374, 1), (8376, 2), (8377, 1), (8410, 1), (8460, 1), (8469, 1), (8487, 2), (8498, 1), (8507, 2), (8540, 1), (8563, 2), (8
571, 1), (8624, 1), (8662, 1), (8672, 1), (8734, 1), (8886, 1), (8990, 1), (9018, 1), (9211, 1), (9338, 2), (9544, 1), (9551, 1), (9569, 1), (9733, 1), (9740, 1), (9741, 2), (9742, 2), (9778, 1), (9841, 1), (990
2, 1), (9925, 1), (10041, 1), (10121, 1), (10124, 1), (10151, 1), (10157, 1), (10184, 1), (10200, 1), (10397, 1), (10443, 1), (10459, 1), (10470, 1), (10524, 1), (10875, 1), (10910, 1), (10971, 1), (11315, 3), (
11458, 1), (11530, 1), (11559, 1), (11561, 1), (11655, 1), (11930, 1), (12123, 1), (12856, 1), (12972, 1), (12994, 1), (13039, 1), (13066, 1), (13120, 1), (13156, 1), (13634, 1), (13636, 1), (13647, 1), (13809,
1), (13925, 1), (13965, 1), (14147, 1), (14166, 1), (14231, 1), (14351, 2), (14386, 1), (14589, 1), (14603, 1), (14801, 1), (14825, 1), (14835, 1), (14838, 1), (14857, 1), (14877, 1), (14957, 1), (14958, 1), (15
041, 1), (15118, 1), (15175, 1), (15180, 1), (15181, 1), (15278, 1), (15283, 1), (15354, 1), (15379, 1), (15401, 1), (15412, 1), (15475, 1), (15487, 1), (15539, 1), (15545, 2)]), ('zlatana', 1, [(6886, 1)]), ('z
latandependence', 3, [(9434, 1), (9440, 1), (9515, 1)]), ('zlatanibrahimovic', 8, [(1807, 1), (6467, 1), (6886, 1), (10470, 1), (11543, 1), (15199, 1), (15407, 1), (15487, 1)]), ('zlatanibrahimovich', 1, [(8424,
 1)]), ('zlatanibramovich', 1, [(8489, 1)]), ('zlatanlegends', 1, [(14351, 1)]), ('zlatanless', 2, [(7721, 1), (9257, 1)]), ('zlatanyou', 1, [(8370, 1)]), ('zlaten', 1, [(518, 1)]), ('zlnft', 1, [(2162, 1)]), ('
zlttvi', 1, [(9247, 1)]), ('zlxungoq', 1, [(836, 1)]), ('zlyzofuj', 1, [(13095, 1)]), ('zm', 2, [(2466, 1), (9265, 1)]), ('zmf', 1, [(11656, 1)]), ('zmx', 2, [(5995, 1), (8908, 1)]), ('zn', 2, [(2244, 1), (10046
, 1)]), ('znewqbn', 1, [(2163, 1)]), ('znmqn', 1, [(4914, 1)]), ('znrdx', 1, [(11293, 1)]), ('znvbtv', 1, [(4496, 1)]), ('zo', 1, [(11818, 1)]), ('zocalo', 8, [(10176, 1), (10270, 1), (10271, 1), (10272, 1), (10
273, 1), (10275, 1), (10276, 1), (10277, 1)]), ('zoe', 1, [(12244, 1)]), ('zoepictwittercom', 1, [(6038, 1)]), ('zogibney', 1, [(12244, 1)]), ('zohkp', 1, [(15211, 1)]), ('zomatoin', 1, [(6191, 1)]), ('zombies',
 1, [(277, 1)]), ('zonal', 1, [(10219, 1)]), ('zone', 13, [(554, 1), (1022, 1), (1110, 1), (1119, 1), (1396, 1), (1652, 1), (1767, 1), (2199, 1), (3281, 1), (7052, 2), (7485, 1), (8108, 1), (8170, 1)]), ('zoned'
, 1, [(10478, 1)]), ('zoo', 1, [(1147, 1)]), ('zooim', 1, [(15025, 1)]), ('zop', 1, [(13859, 1)]), ('zote', 2, [(9188, 1), (12610, 1)]), ('zp', 1, [(10921, 1)]), ('zpeq', 1, [(3981, 1)]), ('zpictwittercomixon',
1, [(14591, 1)]), ('zpoaq', 1, [(1883, 1)]), ('zpvz', 1, [(10115, 1)]), ('zpxlyky', 1, [(9331, 1)]), ('zqbe', 1, [(11296, 1)]), ('zqe', 1, [(13216, 1)]), ('zqf', 2, [(4809, 1), (6085, 1)]), ('zr', 1, [(14644, 1)
]), ('zrii', 1, [(15264, 1)]), ('zrjwsv', 1, [(7438, 1)]), ('zrlf', 1, [(14971, 1)]), ('zrnhj', 1, [(4855, 1)]), ('zs', 3, [(4553, 1), (5009, 1), (13372, 1)]), ('zsd', 1, [(5986, 1)]), ('zshcomgq', 1, [(11658, 1
)]), ('zsj', 1, [(5099, 1)]), ('zsln', 1, [(6179, 1)]), ('zsmmn', 11, [(11050, 1), (11628, 1), (11924, 1), (11925, 1), (11927, 1), (11928, 1), (11931, 1), (11932, 1), (11933, 1), (11934, 1), (14045, 1)]), ('zspg
', 1, [(3462, 1)]), ('zstudio', 1, [(9721, 1)]), ('ztcn', 1, [(14351, 1)]), ('zto', 1, [(1275, 1)]), ('ztrus', 1, [(14068, 1)]), ('zu', 3, [(4254, 1), (10504, 1), (13476, 1)]), ('zuber', 4, [(8915, 1), (9124, 1)
, (9148, 1), (9154, 1)]), ('zuerst', 1, [(14021, 1)]), ('zuidkorea', 1, [(14983, 1)]), ('zum', 2, [(12046, 1), (12226, 1)]), ('zuma', 1, [(6884, 1)]), ('zurich', 3, [(8656, 1), (9083, 1), (9087, 1)]), ('zusph',
1, [(13804, 1)]), ('zuuo', 1, [(9368, 1)]), ('zuysz', 1, [(930, 1)]), ('zv', 2, [(13945, 1), (14475, 1)]), ('zvcx', 1, [(7585, 1)]), ('zviedrija', 1, [(5159, 1)]), ('zvj', 1, [(1374, 1)]), ('zvjyqbdup', 1, [(144
85, 1)]), ('zw', 1, [(12008, 1)]), ('zweden', 7, [(1958, 1), (3957, 1), (6963, 2), (8989, 1), (12839, 1), (13538, 1), (14983, 1)]), ('zweeng', 2, [(3957, 1), (6963, 1)]), ('zweitesspiel', 1, [(14099, 1)]), ('zwe
kor', 1, [(14983, 1)]), ('zwezko', 1, [(14983, 1)]), ('zwezwi', 1, [(8989, 1)]), ('zwigpta', 1, [(5137, 1)]), ('zwitserland', 1, [(8989, 1)]), ('zww', 1, [(14747, 1)]), ('zwxvt', 1, [(11148, 1)]), ('zx', 1, [(43
46, 1)]), ('zxihm', 1, [(4506, 1)]), ('zxuq', 1, [(11973, 1)]), ('zxxihqx', 1, [(69, 1)]), ('zxzp', 1, [(10111, 1)]), ('zy', 1, [(4568, 1)]), ('zyx', 1, [(9420, 1)]), ('zz', 1, [(6586, 1)]), ('zzl', 1, [(13462,
1)]), ('zzqe', 1, [(9084, 1)]), ('zzvqw', 1, [(8802, 1)]), ('zzz', 1, [(6954, 1)]), ('zzzzzzz', 2, [(9063, 1), (12912, 1)]), ('|irasuto|huriSu', 1, [(9602, 1)]), ('~', 1, [(15239, 1)])]

The format for the Results are:

Sentiment Analysis: (Tweet number, Sentiment Score)

Tfdf: (Word, Number of occurrences, [(Tweet number, Frequency in this Tweet)])
```

Figure 2: TF-DF for AllTweets.txt



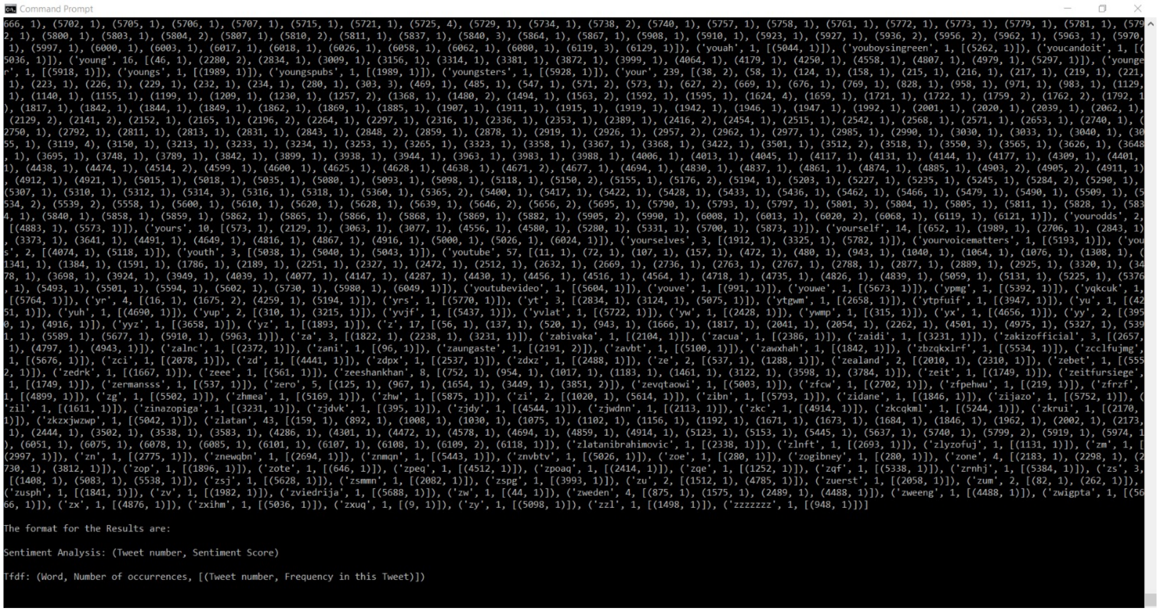
```
5492, 0.0), (5493, 0.0), (5494, 0.0), (5495, -1.0), (5496, 0.0), (5497, 0.0), (5498, 2.0), (5499, 2.0), (5500, 2.0), (5501, 0.0), (5502, 1.0), (5503, 5.0), (5504, 1.0), (5505, 4.0), (5506, -1.0), (5507, 1.0), (5
508, 1.0), (5509, -3.0), (5510, 9.0), (5511, 3.0), (5512, 2.0), (5513, 8.0), (5514, 6.0), (5515, 2.0), (5516, 3.0), (5517, 0.0), (5518, 5.0), (5519, 12.0), (5520, 1.0), (5521, 0.0), (5522, 4.0), (5523, 0.0), (55
24, 5.0), (5525, 0.0), (5526, 9.0), (5527, -1.0), (5528, 0.0), (5529, 2.0), (5530, 9.0), (5531, 7.0), (5532, 0.0), (5533, 2.0), (5534, 19.0), (5535, 1.0), (5536, 5.0), (5537, 0.0), (5538, 1.0), (5539, 0.0), (554
0, 0.0), (5541, 5.0), (5542, 1.0), (5543, 1.0), (5544, 0.0), (5545, 1.0), (5546, 8.0), (5547, 1.0), (5548, 6.0), (5549, 1.0), (5550, 6.0), (5551, 12.0), (5552, 0.0), (5553, 0.0), (5554, 5.0), (5555, -5.0), (5556
, 4.0), (5557, 3.0), (5558, 4.0), (5559, 1.0), (5560, -1.0), (5561, 9.0), (5562, 1.0), (5563, 9.0), (5564, 2.0), (5565, 2.0), (5566, 3.0), (5567, 0.0), (5568, 0.0), (5569, 3.0), (5570, 1.0), (5571, 1.0), (5572,
3.0), (5573, 0.0), (5574, 4.0), (5575, 1.0), (5576, 1.0), (5577, 0.0), (5578, -2.0), (5579, 0.0), (5580, 4.0), (5581, 0.0), (5582, 5.0), (5583, 0.0), (5584, 1.0), (5585, 0.0), (5586, 5.0), (5587, 1.0), (5588, 1.
0), (5589, 1.0), (5590, 0.0), (5591, 0.0), (5592, -2.0), (5593, 2.0), (5594, 0.0), (5595, -1.0), (5596, 6.0), (5597, 4.0), (5598, 2.0), (5599, 0.0), (5600, 4.0), (5601, 6.0), (5602, 3.0), (5603, 1.0), (5604, 0.0
), (5605, 7.0), (5606, 3.0), (5607, 2.0), (5608, 1.0), (5609, 4.0), (5610, 0.0), (5611, 2.0), (5612, 1.0), (5613, 1.0), (5614, 1.0), (5615, -4.0), (5616, 1.0), (5617, 1.0), (5618, 1.0), (5619, 13.0), (5620, -3.0
), (5621, 0.0), (5622, 0.0), (5623, 2.0), (5624, 2.0), (5625, 3.0), (5626, 7.0), (5627, 5.0), (5628, 4.0), (5629, -4.0), (5630, -3.0), (5631, 2.0), (5632, 11.0), (5633, 2.0), (5634, 0.0), (5635, 4.0), (5636, 0.0
), (5637, 2.0), (5638, 2.0), (5639, 6.0), (5640, 0.0), (5641, 0.0), (5642, 1.0), (5643, 1.0), (5644, 0.0), (5645, 0.0), (5646, 3.0), (5647, 0.0), (5648, 4.0), (5649, 0.0), (5650, 0.0), (5651, 0.0), (5652, 0.0),
(5653, 0.0), (5654, 7.0), (5655, 1.0), (5656, 3.0), (5657, 2.0), (5658, 8.0), (5659, 0.0), (5660, -5.0), (5661, 4.0), (5662, 0.0), (5663, 3.0), (5664, 4.0), (5665, 1.0), (5666, 3.0), (5667, -2.0), (5668, -2.0),
(5669, 2.0), (5670, 0.0), (5671, 5.0), (5672, 0.0), (5673, 0.0), (5674, 7.0), (5675, 0.0), (5676, 1.0), (5677, -5.0), (5678, 2.0), (5679, 0.0), (5680, 3.0), (5681, 0.0), (5682, 0.0), (5683, 0.0), (5684, 0.0), (5
685, 0.0), (5686, 0.0), (5687, -2.0), (5688, 0.0), (5689, 3.0), (5690, 9.0), (5691, 1.0), (5692, 2.0), (5693, 0.0), (5694, 2.0), (5695, 3.0), (5696, -6.0), (5697, 7.0), (5698, 6.0), (5699, 1.0), (5700, 0.0), (57
01, -5.0), (5702, 1.0), (5703, 1.0), (5704, -1.0), (5705, 17.0), (5706, 3.0), (5707, 3.0), (5708, 8.0), (5709, 1.0), (5710, 4.0), (5711, -2.0), (5712, -2.0), (5713, 1.0), (5714, 2.0), (5715, 0.0), (5716, 0.0), (
5717, 0.0), (5718, 0.0), (5719, 0.0), (5720, 0.0), (5721, 2.0), (5722, 2.0), (5723, 3.0), (5724, 0.0), (5725, 0.0), (5726, 0.0), (5727, -4.0), (5728, -4.0), (5729, 4.0), (5730, 0.0), (5731, 1.0), (5732, 0.0), (5
733, 2.0), (5734, -2.0), (5735, 1.0), (5736, 0.0), (5737, 4.0), (5738, 5.0), (5739, 1.0), (5740, 1.0), (5741, 0.0), (5742, 3.0), (5743, 3.0), (5744, 0.0), (5745, 4.0), (5746, 1.0), (5747, 3.0), (5748, 0.0), (574
9, 1.0), (5750, 0.0), (5751, -4.0), (5752, 0.0), (5753, 1.0), (5754, 1.0), (5755, 0.0), (5756, 0.0), (5757, 2.0), (5758, 3.0), (5759, 0.0), (5760, 3.0), (5761, 5.0), (5762, 8.0), (5763, 1.0), (5764, 0.0), (5765,
 -2.0), (5766, 9.0), (5767, 0.0), (5768, 4.0), (5769, 0.0), (5770, 1.0), (5771, 1.0), (5772, 9.0), (5773, 4.0), (5774, 6.0), (5775, 2.0), (5776, 4.0), (5777, 0.0), (5778, 6.0), (5779, -1.0), (5780, 0.0), (5781,
1.0), (5782, 0.0), (5783, 0.0), (5784, 0.0), (5785, -1.0), (5786, 0.0), (5787, 2.0), (5788, 3.0), (5789, 0.0), (5790, 0.0), (5791, 0.0), (5792, 3.0), (5793, 0.0), (5794, 10.0), (5795, 0.0), (5796, 1.0), (5797, 0
.0), (5798, 0.0), (5799, 2.0), (5800, 2.0), (5801, 0.0), (5802, 1.0), (5803, 0.0), (5804, 9.0), (5805, 0.0), (5806, 7.0), (5807, 1.0), (5808, 0.0), (5809, 4.0), (5810, 2.0), (5811, 6.0), (5812, 0.0), (5813, 0.0)
, (5814, 4.0), (5815, -1.0), (5816, 1.0), (5817, 0.0), (5818, 0.0), (5819, 2.0), (5820, 3.0), (5821, 0.0), (5822, 1.0), (5823, 7.0), (5824, -4.0), (5825, 3.0), (5826, 4.0), (5827, 0.0), (5828, 6.0), (5829, 2.0),
 (5830, 1.0), (5831, 1.0), (5832, 6.0), (5833, 0.0), (5834, 0.0), (5835, -2.0), (5836, 1.0), (5837, 2.0), (5838, 1.0), (5839, 0.0), (5840, 2.0), (5841, 9.0), (5842, 0.0), (5843, 0.0), (5844, -4.0), (5845, -2.0),
 (5846, 1.0), (5847, 4.0), (5848, -4.0), (5849, -1.0), (5850, 2.0), (5851, -3.0), (5852, 3.0), (5853, 7.0), (5854, -2.0), (5855, 4.0), (5856, 17.0), (5857, 0.0), (5858, 3.0), (5859, 3.0), (5860, 8.0), (5861, 1.0
), (5862, 3.0), (5863, 13.0), (5864, 2.0), (5865, 3.0), (5866, 3.0), (5867, 8.0), (5868, 3.0), (5869, 3.0), (5870, 2.0), (5871, -3.0), (5872, 0.0), (5873, 7.0), (5874, 4.0), (5875, 6.0), (5876, 0.0), (5877, -3.0
), (5878, 0.0), (5879, -2.0), (5880, 2.0), (5881, 3.0), (5882, 3.0), (5883, 15.0), (5884, 0.0), (5885, 6.0), (5886, 2.0), (5887, 0.0), (5888, 0.0), (5889, -3.0), (5890, 2.0), (5891, 5.0), (5892, 0.0), (5893, 0.0
), (5894, 0.0), (5895, -2.0), (5896, 0.0), (5897, 4.0), (5898, 1.0), (5899, -2.0), (5900, 0.0), (5901, 1.0), (5902, 10.0), (5903, 2.0), (5904, 0.0), (5905, 9.0), (5906, 1.0), (5907, 3.0), (5908, 2.0), (5909, 1.0
), (5910, 8.0), (5911, 1.0), (5912, 1.0), (5913, 6.0), (5914, 1.0), (5915, 1.0), (5916, 1.0), (5917, 2.0), (5918, -4.0), (5919, -6.0), (5920, 1.0), (5921, 6.0), (5922, 1.0), (5923, 0.0), (5924, 2.0), (5925, 1.0)
, (5926, 1.0), (5927, 0.0), (5928, 6.0), (5929, 0.0), (5930, 0.0), (5931, 1.0), (5932, 6.0), (5933, 0.0), (5934, 0.0), (5935, 6.0), (5936, 0.0), (5937, 4.0), (5938, -2.0), (5939, 5.0), (5940, 4.0), (5941, 4.0),
(5942, 5.0), (5943, 5.0), (5944, 0.0), (5945, 2.0), (5946, 1.0), (5947, -1.0), (5948, 0.0), (5949, 2.0), (5950, 6.0), (5951, 1.0), (5952, -1.0), (5953, 0.0), (5954, 0.0), (5955, 10.0), (5956, 6.0), (5957, 2.0),
(5958, 0.0), (5959, 1.0), (5960, -1.0), (5961, 0.0), (5962, 0.0), (5963, 6.0), (5964, 0.0), (5965, 1.0), (5966, -1.0), (5967, -1.0), (5968, 0.0), (5969, 10.0), (5970, 0.0), (5971, 0.0), (5972, -2.0), (5973, 0.0)
, (5974, 0.0), (5975, 1.0), (5976, -2.0), (5977, 1.0), (5978, 0.0), (5979, 0.0), (5980, 0.0), (5981, 0.0), (5982, 5.0), (5983, 0.0), (5984, 0.0), (5985, 0.0), (5986, 0.0), (5987, 0.0), (5988, 2.0), (5989, 0.0),
(5990, 1.0), (5991, 1.0), (5992, 0.0), (5993, 0.0), (5994, 0.0), (5995, 0.0), (5996, 1.0), (5997, 8.0), (5998, 0.0), (5999, 4.0), (6000, 0.0), (6001, 0.0), (6002, 3.0), (6003, 3.0), (6004, 3.0), (6005, 0.0), (60
06, 0.0), (6007, 1.0), (6008, 0.0), (6009, 2.0), (6010, 0.0), (6011, 0.0), (6012, 13.0), (6013, 4.0), (6014, 12.0), (6015, 0.0), (6016, 0.0), (6017, -2.0), (6018, -2.0), (6019, 4.0), (6020, 2.0), (6021, 0.0), (6
022, 3.0), (6023, 0.0), (6024, -4.0), (6025, 0.0), (6026, 0.0), (6027, -5.0), (6028, 1.0), (6029, 1.0), (6030, 1.0), (6031, 0.0), (6032, 0.0), (6033, 1.0), (6034, 0.0), (6035, 1.0), (6036, 1.0), (6037, 0.0), (60
38, 0.0), (6039, 1.0), (6040, 0.0), (6041, 3.0), (6042, 0.0), (6043, 0.0), (6044, 2.0), (6045, 0.0), (6046, 0.0), (6047, 1.0), (6048, -1.0), (6049, 2.0), (6050, 0.0), (6051, 1.0), (6052, 1.0), (6053, 0.0), (6054
, 0.0), (6055, 1.0), (6056, 0.0), (6057, 0.0), (6058, 8.0), (6059, -1.0), (6060, 0.0), (6061, 0.0), (6062, -5.0), (6063, 1.0), (6064, 1.0), (6065, 2.0), (6066, 1.0), (6067, 0.0), (6068, 7.0), (6069, 1.0), (6070,
 2.0), (6071, 0.0), (6072, 0.0), (6073, 3.0), (6074, 6.0), (6075, -3.0), (6076, 4.0), (6077, 4.0), (6078, -3.0), (6079, 1.0), (6080, 6.0), (6081, 1.0), (6082, 3.0), (6083, -3.0), (6084, 1.0), (6085, -3.0), (6086
, 3.0), (6087, 0.0), (6088, 0.0), (6089, 0.0), (6090, 0.0), (6091, 1.0), (6092, 0.0), (6093, 0.0), (6094, 0.0), (6095, 0.0), (6096, 1.0), (6097, 0.0), (6098, 0.0), (6099, 1.0), (6100, 9.0), (6101, -4.0), (6102,
1.0), (6103, 0.0), (6104, 1.0), (6105, 0.0), (6106, 1.0), (6107, -3.0), (6108, 0.0), (6109, -1.0), (6110, 1.0), (6111, 0.0), (6112, 1.0), (6113, 1.0), (6114, 10.0), (6115, 0.0), (6116, 0.0), (6117, 1.0), (6118,
1.0), (6119, 2.0), (6120, 1.0), (6121, 4.0), (6122, 8.0), (6123, 13.0), (6124, 9.0), (6125, 1.0), (6126, 0.0), (6127, 1.0), (6128, 0.0), (6129, 0.0), (6130, 3.0), (6131, 2.0), (6132, 0.0), (6133, -4.0), (6134, -
2.0)]

The format for the Results are:

Sentiment Analysis: (Tweet number, Sentiment Score)

Tfdf: (Word, Number of occurrences, [(Tweet number, Frequency in this Tweet)])
```

Figure 3: Sentiment Analysis for LossTweets.txt



```
666, 1), (5702, 1), (5705, 1), (5706, 1), (5707, 1), (5715, 1), (5721, 1), (5725, 4), (5729, 1), (5734, 1), (5738, 2), (5740, 1), (5757, 1), (5758, 1), (5761, 1), (5772, 1), (5773, 1), (5779, 1), (5781, 1), (579
2, 1), (5800, 1), (5803, 1), (5804, 2), (5807, 1), (5810, 2), (5811, 1), (5837, 1), (5840, 3), (5864, 1), (5867, 1), (5908, 1), (5910, 1), (5923, 1), (5927, 1), (5936, 2), (5956, 2), (5962, 1), (5963, 1), (5970,
 1), (5997, 1), (6000, 1), (6003, 1), (6017, 1), (6018, 1), (6026, 1), (6058, 1), (6062, 1), (6080, 1), (6119, 3), (6129, 1)]), ('youah', 1, [(5044, 1)]), ('youboysingreen', 1, [(5262, 1)]), ('youcandoit', 1, [(
5036, 1)]), ('young', 16, [(46, 1), (2280, 2), (2834, 1), (3009, 1), (3156, 1), (3314, 1), (3381, 1), (3872, 1), (3999, 1), (4064, 1), (4179, 1), (4250, 1), (4558, 1), (4807, 1), (4979, 1), (5297, 1)]), ('younge
r', 1, [(5918, 1)]), ('youngs', 1, [(1989, 1)]), ('youngspubs', 1, [(1989, 1)]), ('youngsters', 1, [(5928, 1)]), ('your', 239, [(38, 2), (58, 1), (124, 1), (158, 1), (215, 1), (216, 1), (217, 1), (219, 1), (221,
 1), (223, 1), (226, 1), (229, 1), (232, 1), (234, 1), (280, 1), (303, 3), (469, 1), (485, 1), (547, 1), (571, 2), (573, 1), (627, 2), (669, 1), (676, 1), (769, 1), (828, 1), (958, 1), (971, 1), (983, 1), (1129,
 1), (1140, 1), (1155, 1), (1199, 1), (1209, 1), (1230, 1), (1257, 2), (1368, 1), (1480, 2), (1494, 1), (1563, 2), (1592, 1), (1595, 1), (1624, 4), (1659, 1), (1721, 1), (1722, 1), (1759, 2), (1762, 2), (1792, 1
), (1817, 1), (1842, 1), (1844, 1), (1849, 1), (1862, 1), (1869, 1), (1885, 1), (1907, 1), (1911, 1), (1915, 1), (1919, 1), (1942, 1), (1946, 1), (1947, 1), (1992, 1), (2001, 1), (2020, 1), (2039, 1), (2062, 1),
 (2129, 2), (2141, 2), (2152, 1), (2165, 1), (2196, 2), (2264, 1), (2297, 1), (2316, 1), (2336, 1), (2353, 1), (2389, 1), (2416, 2), (2454, 1), (2515, 1), (2542, 1), (2568, 1), (2571, 1), (2653, 1), (2740, 1), (
2750, 1), (2792, 1), (2811, 1), (2813, 1), (2831, 1), (2843, 1), (2848, 2), (2859, 1), (2878, 1), (2919, 1), (2926, 1), (2957, 2), (2962, 1), (2977, 1), (2985, 1), (2990, 1), (3030, 1), (3033, 1), (3040, 1), (30
55, 1), (3119, 4), (3150, 1), (3213, 1), (3233, 1), (3234, 1), (3253, 1), (3265, 1), (3323, 1), (3358, 1), (3367, 1), (3368, 1), (3422, 1), (3501, 1), (3512, 2), (3518, 1), (3550, 3), (3565, 1), (3626, 1), (3648
, 1), (3695, 1), (3748, 1), (3789, 1), (3842, 1), (3899, 1), (3938, 1), (3944, 1), (3963, 1), (3983, 1), (3988, 1), (4006, 1), (4013, 1), (4045, 1), (4117, 1), (4131, 1), (4144, 1), (4177, 1), (4309, 1), (4401,
1), (4438, 1), (4474, 1), (4514, 2), (4599, 1), (4600, 1), (4625, 1), (4628, 1), (4638, 1), (4671, 2), (4677, 1), (4694, 1), (4830, 1), (4837, 1), (4861, 1), (4874, 1), (4885, 1), (4903, 2), (4905, 2), (4911, 1)
, (4912, 1), (4921, 1), (5015, 1), (5018, 1), (5035, 1), (5080, 1), (5093, 1), (5098, 1), (5118, 1), (5150, 2), (5155, 1), (5176, 2), (5194, 1), (5203, 1), (5227, 1), (5235, 1), (5245, 1), (5284, 2), (5290, 1),
(5307, 1), (5310, 1), (5312, 1), (5314, 3), (5316, 1), (5318, 1), (5360, 1), (5365, 2), (5400, 1), (5417, 1), (5422, 1), (5428, 1), (5433, 1), (5436, 1), (5462, 1), (5466, 1), (5479, 1), (5490, 1), (5509, 1), (5
534, 2), (5539, 2), (5558, 1), (5600, 1), (5610, 1), (5620, 1), (5628, 1), (5639, 1), (5646, 2), (5656, 2), (5695, 1), (5790, 1), (5793, 1), (5797, 1), (5801, 3), (5804, 1), (5805, 1), (5811, 1), (5828, 1), (583
4, 1), (5840, 1), (5858, 1), (5859, 1), (5862, 1), (5865, 1), (5866, 1), (5868, 1), (5869, 1), (5882, 1), (5905, 2), (5990, 1), (6008, 1), (6013, 1), (6020, 2), (6068, 1), (6119, 1), (6121, 1)]), ('yourodds', 2,
 [(4883, 1), (5573, 1)]), ('yours', 10, [(573, 1), (2129, 1), (3063, 1), (3077, 1), (4556, 1), (4580, 1), (5280, 1), (5331, 1), (5700, 1), (5873, 1)]), ('yourself', 14, [(652, 1), (1989, 1), (2706, 1), (2843, 1)
, (3373, 1), (3641, 1), (4491, 1), (4649, 1), (4816, 1), (4867, 1), (4916, 1), (5000, 1), (5026, 1), (6024, 1)]), ('yourselves', 3, [(1912, 1), (3325, 1), (5782, 1)]), ('yourvoicematters', 1, [(5193, 1)]), ('you
s', 2, [(4074, 1), (5118, 1)]), ('youth', 3, [(5038, 1), (5040, 1), (5043, 1)]), ('youtube', 57, [(11, 1), (72, 1), (107, 1), (157, 1), (472, 1), (480, 1), (943, 1), (1040, 1), (1064, 1), (1076, 1), (1308, 1), (
1341, 1), (1384, 1), (1591, 1), (1786, 1), (2189, 1), (2251, 1), (2327, 1), (2472, 1), (2512, 1), (2632, 1), (2669, 1), (2736, 1), (2763, 1), (2767, 1), (2788, 1), (2877, 1), (2889, 1), (2925, 1), (3320, 1), (34
78, 1), (3698, 1), (3924, 1), (3949, 1), (4039, 1), (4077, 1), (4147, 1), (4287, 1), (4430, 1), (4456, 1), (4516, 1), (4564, 1), (4718, 1), (4735, 1), (4826, 1), (4839, 1), (5059, 1), (5131, 1), (5225, 1), (5376
, 1), (5493, 1), (5501, 1), (5594, 1), (5602, 1), (5730, 1), (5980, 1), (6049, 1)]), ('youtubevideo', 1, [(5604, 1)]), ('youve', 1, [(991, 1)]), ('youwe', 1, [(5673, 1)]), ('ypmg', 1, [(5392, 1)]), ('yqkcuk', 1,
 [(5764, 1)]), ('yr', 4, [(16, 1), (1675, 2), (4259, 1), (5194, 1)]), ('yrs', 1, [(5770, 1)]), ('yt', 3, [(2834, 1), (3124, 1), (5075, 1)]), ('ytgwm', 1, [(2658, 1)]), ('ytpfuif', 1, [(3947, 1)]), ('yu', 1, [(42
51, 1)]), ('yuh', 1, [(4690, 1)]), ('yup', 2, [(310, 1), (3215, 1)]), ('yvjf', 1, [(5437, 1)]), ('yvlat', 1, [(5722, 1)]), ('yw', 1, [(2428, 1)]), ('ywmp', 1, [(315, 1)]), ('yx', 1, [(4656, 1)]), ('yy', 2, [(395
0, 1), (4916, 1)]), ('yyz', 1, [(3658, 1)]), ('yz', 1, [(1893, 1)]), ('z', 17, [(56, 1), (137, 1), (520, 1), (943, 1), (1666, 1), (1817, 1), (2041, 1), (2054, 1), (2262, 1), (4501, 1), (4975, 1), (5327, 1), (539
1, 1), (5589, 1), (5677, 1), (5910, 1), (5963, 1)]), ('za', 3, [(1822, 1), (2238, 1), (3231, 1)]), ('zabivaka', 1, [(2104, 1)]), ('zacua', 1, [(2386, 1)]), ('zaidi', 1, [(3231, 1)]), ('zakizofficial', 3, [(2657,
 1), (4797, 1), (4943, 1)]), ('zalnc', 1, [(2372, 1)]), ('zani', 1, [(96, 1)]), ('zaungaste', 1, [(2191, 2)]), ('zavbt', 1, [(5100, 1)]), ('zawxhah', 1, [(1842, 1)]), ('zbzqkxlrf', 1, [(5534, 1)]), ('zcclfujmg',
 1, [(5676, 1)]), ('zci', 1, [(2078, 1)]), ('zd', 1, [(4441, 1)]), ('zdpx', 1, [(2537, 1)]), ('zdxz', 1, [(2488, 1)]), ('ze', 2, [(537, 1), (1288, 1)]), ('zealand', 2, [(2010, 1), (2310, 1)]), ('zebet', 1, [(555
2, 1)]), ('zedrk', 1, [(1667, 1)]), ('zeee', 1, [(561, 1)]), ('zeeshankhan', 8, [(752, 1), (954, 1), (1017, 1), (1183, 1), (1461, 1), (3122, 1), (3598, 1), (3784, 1)]), ('zeit', 1, [(1749, 1)]), ('zeitfursiege',
 1, [(1749, 1)]), ('zermansss', 1, [(537, 1)]), ('zero', 5, [(125, 1), (967, 1), (1654, 1), (3449, 1), (3851, 2)]), ('zevqtaowi', 1, [(5003, 1)]), ('zfcw', 1, [(2702, 1)]), ('zfpehwu', 1, [(219, 1)]), ('zfrzf',
1, [(4899, 1)]), ('zg', 1, [(5502, 1)]), ('zhmea', 1, [(5169, 1)]), ('zhw', 1, [(5875, 1)]), ('zi', 2, [(1020, 1), (5614, 1)]), ('zibn', 1, [(5793, 1)]), ('zidane', 1, [(1846, 1)]), ('zijazo', 1, [(5752, 1)]), (
'zil', 1, [(1611, 1)]), ('zinazopiga', 1, [(3231, 1)]), ('zjdvk', 1, [(395, 1)]), ('zjdy', 1, [(4544, 1)]), ('zjwdnn', 1, [(2113, 1)]), ('zkc', 1, [(4914, 1)]), ('zkcqkml', 1, [(5244, 1)]), ('zkrui', 1, [(2170,
1)]), ('zkzxjwzwp', 1, [(5042, 1)]), ('zlatan', 43, [(159, 1), (892, 1), (1008, 1), (1030, 1), (1075, 1), (1102, 1), (1156, 1), (1192, 1), (1671, 1), (1673, 1), (1684, 1), (1846, 1), (1962, 1), (2002, 1), (2173,
 1), (2444, 1), (3502, 1), (3538, 1), (3583, 1), (4286, 1), (4301, 1), (4472, 1), (4578, 1), (4694, 1), (4859, 1), (4914, 1), (5123, 1), (5153, 1), (5445, 1), (5637, 1), (5740, 1), (5799, 2), (5919, 1), (5974, 1
), (6051, 1), (6075, 1), (6078, 1), (6085, 1), (6101, 1), (6107, 1), (6108, 1), (6109, 2), (6118, 1)]), ('zlatanibrahimovic', 1, [(2338, 1)]), ('zlnft', 1, [(2693, 1)]), ('zlyzofuj', 1, [(1131, 1)]), ('zm', 1, [
(2997, 1)]), ('zn', 1, [(2775, 1)]), ('znewqbn', 1, [(2694, 1)]), ('znmqn', 1, [(5443, 1)]), ('znvbtv', 1, [(5026, 1)]), ('zoe', 1, [(280, 1)]), ('zogibney', 1, [(280, 1)]), ('zone', 4, [(2183, 1), (2298, 1), (2
730, 1), (3812, 1)]), ('zop', 1, [(1896, 1)]), ('zote', 1, [(646, 1)]), ('zpeq', 1, [(4512, 1)]), ('zpoaq', 1, [(2414, 1)]), ('zqe', 1, [(1252, 1)]), ('zqf', 1, [(5338, 1)]), ('zrnhj', 1, [(5384, 1)]), ('zs', 3,
 [(1408, 1), (5083, 1), (5538, 1)]), ('zsj', 1, [(5628, 1)]), ('zsmmn', 1, [(2082, 1)]), ('zspg', 1, [(3993, 1)]), ('zu', 2, [(1512, 1), (4785, 1)]), ('zuerst', 1, [(2058, 1)]), ('zum', 2, [(82, 1), (262, 1)]),
('zusph', 1, [(1841, 1)]), ('zv', 1, [(1982, 1)]), ('zviedrija', 1, [(5688, 1)]), ('zw', 1, [(44, 1)]), ('zweden', 4, [(875, 1), (1575, 1), (2489, 1), (4488, 1)]), ('zweeng', 1, [(4488, 1)]), ('zwigpta', 1, [(56
66, 1)]), ('zx', 1, [(4876, 1)]), ('zxihm', 1, [(5036, 1)]), ('zxuq', 1, [(9, 1)]), ('zy', 1, [(5098, 1)]), ('zzl', 1, [(1498, 1)]), ('zzzzzzz', 1, [(948, 1)])]

The format for the Results are:

Sentiment Analysis: (Tweet number, Sentiment Score)

Tfdf: (Word, Number of occurrences, [(Tweet number, Frequency in this Tweet)])
```

Figure 4: TF-DF for LossTweets.txt

## References

[1] Afinn. http://www2.imm.dtu.dk/pubdb/views/publication$_{d}etails.php?id = 6010$.

[2] Github. https://github.com/Jefferson-Henrique/GetOldTweets-python.