UK

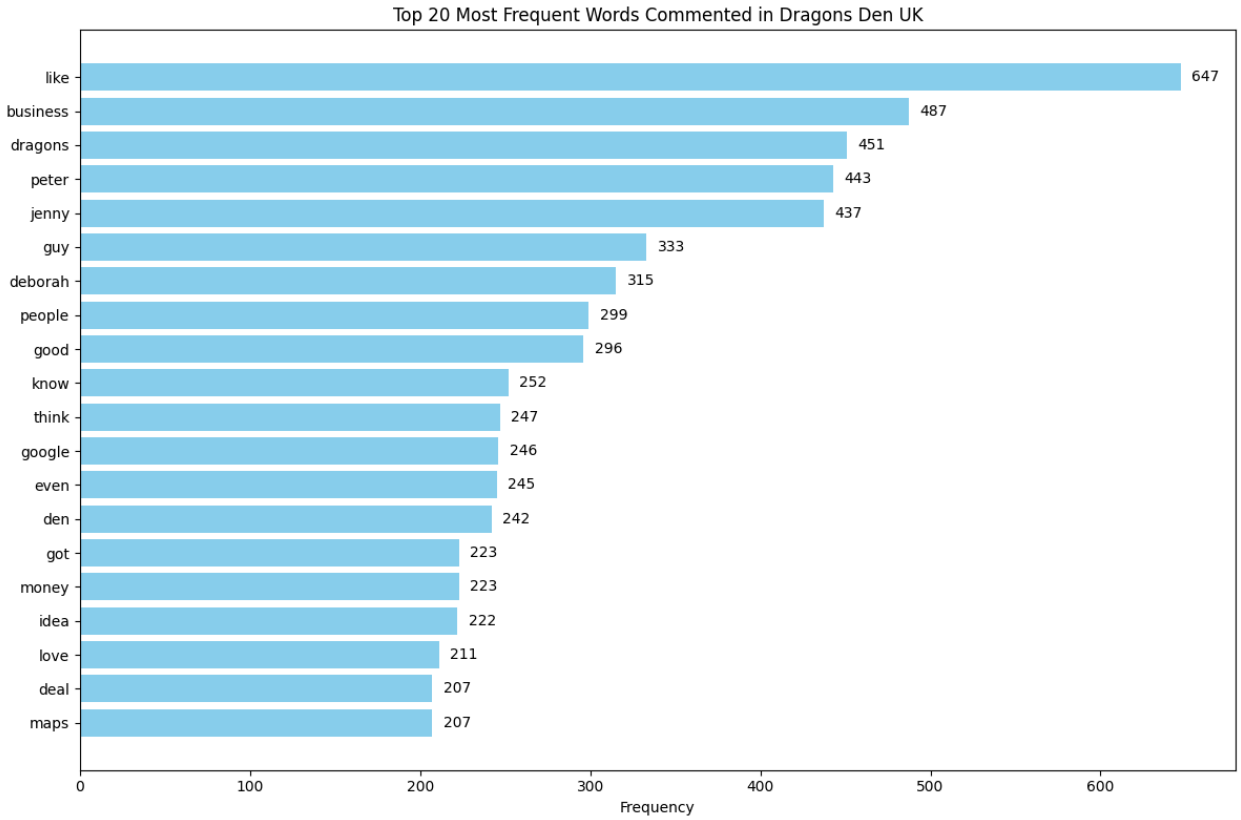
Dataset: 7105 documents

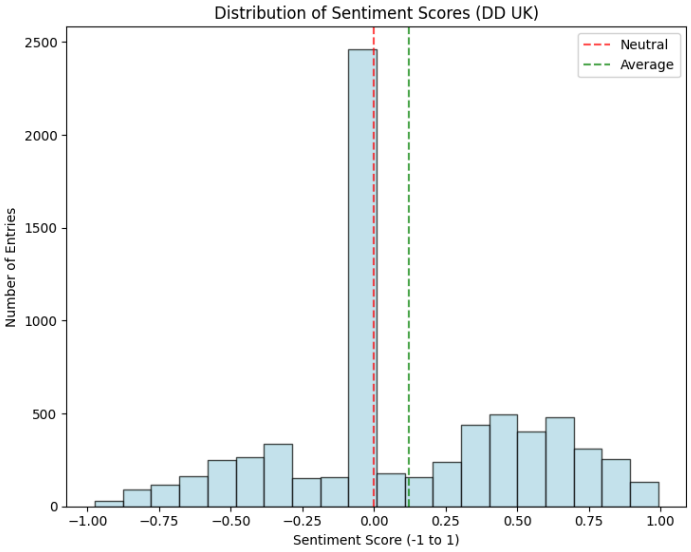
1 TERM FREQUENCY (from HW4-1):
Vocabulary size: 8006 unique words
Top 10 words in POSITIVE entries (2937 entries):
like (510 times)
good (244 times)
business (238 times)
dragons (231 times)
peter (215 times)
love (199 times)
guy (179 times)
deborah (164 times)
lol (159 times)
people (139 times)

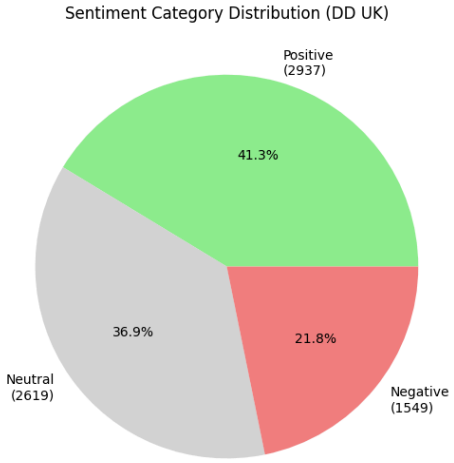
Top 10 words in NEGATIVE entries (1549 entries):
business (150 times)
dragons (127 times)
people (113 times)
peter (111 times)
like (109 times)
even (83 times)
deborah (76 times)
idea (75 times)
guy (70 times)
jenny (68 times)

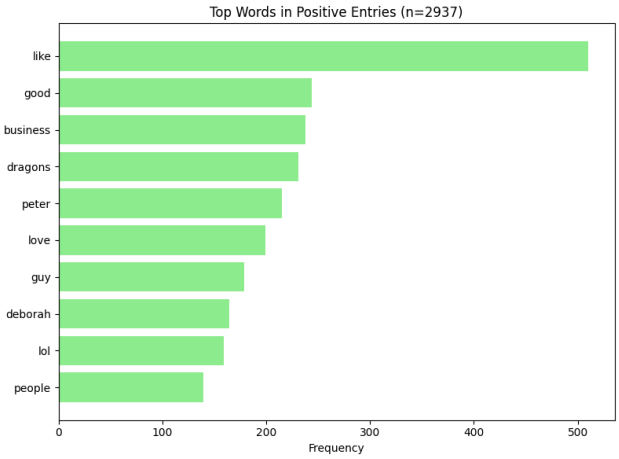
2 SENTIMENT ANALYSIS (from HW4-1):
Average sentiment: 0.123
Positive: 2937 (41.3%)
Neutral: 2619 (36.9%)
Negative: 1549 (21.8%)

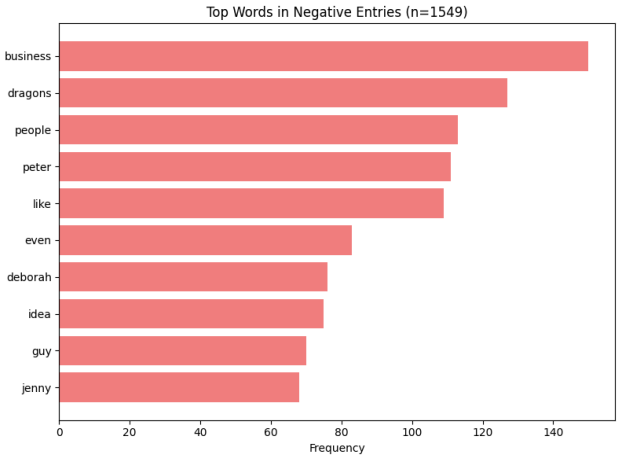
3 TOPIC MODELING (HW4-2):
Number of topics: 3
Topic 0 (1213 docs, avg sentiment 0.08): google, map, phone
Topic 1 (3629 docs, avg sentiment 0.15): dragon, peter, guy
Topic 2 (2263 docs, avg sentiment 0.11): like, jenny, look

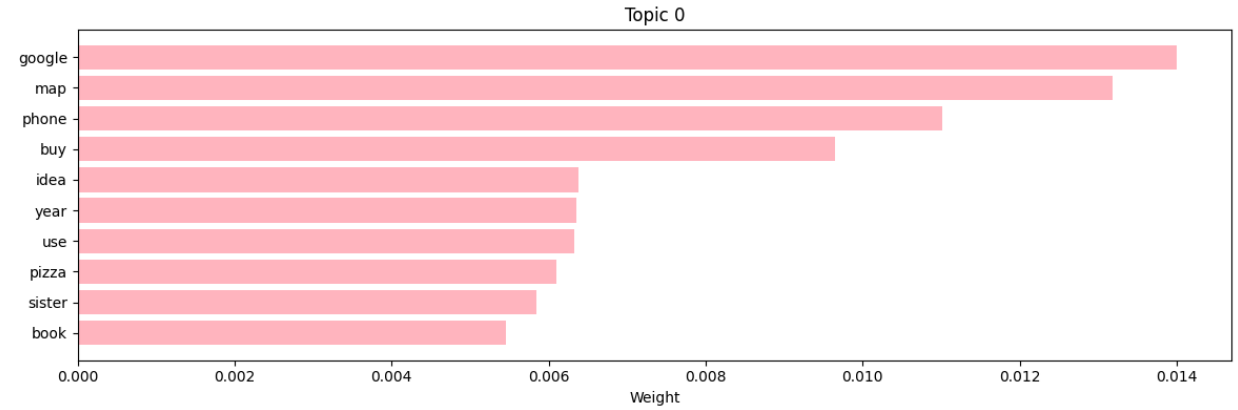

Top 20 Most Frequent Words Commented in Dragons Den UK
like 647
business 487
dragons 451
peter 443
jenny 437
guy 333
deborah 315
people 299
good 296
know 252
think 247
google 246
even 245
den 242
got 223
money 223
idea 222
love 211
deal 207
maps 207
0
100
200
300
400
500
600
Frequency

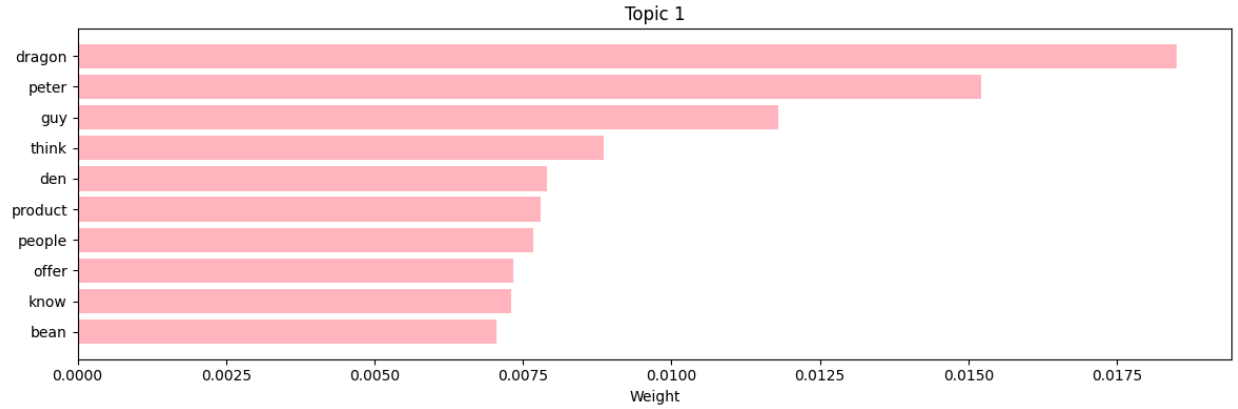

Distribution of Sentiment Scores (DD UK)
Neutral
Average
2500
2000
1500
1000
500
0
Number of Entries
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Sentiment Score (-1 to 1)

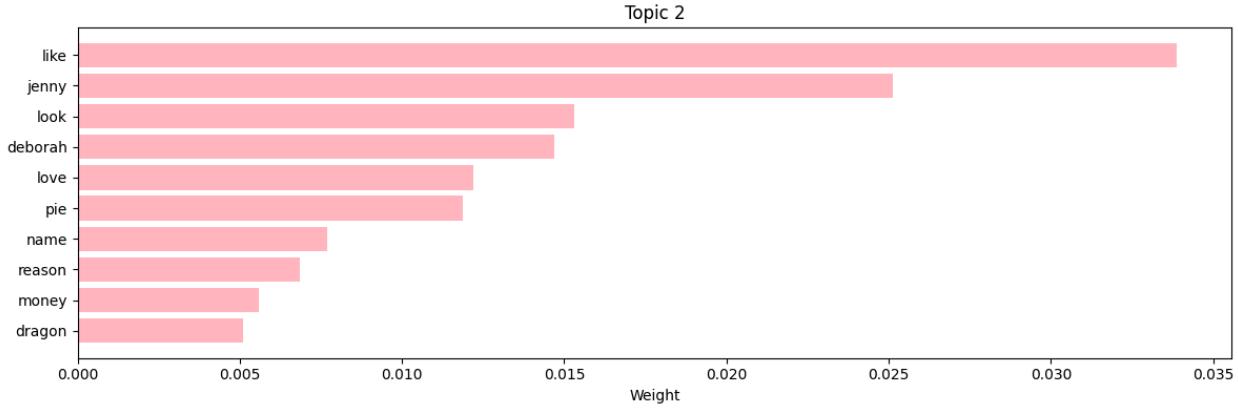

Sentiment Category Distribution (DD UK)
Positive
(2937)
41.3%
36.9%
21.8%
Neutral
(2619)
Negative
(1549)

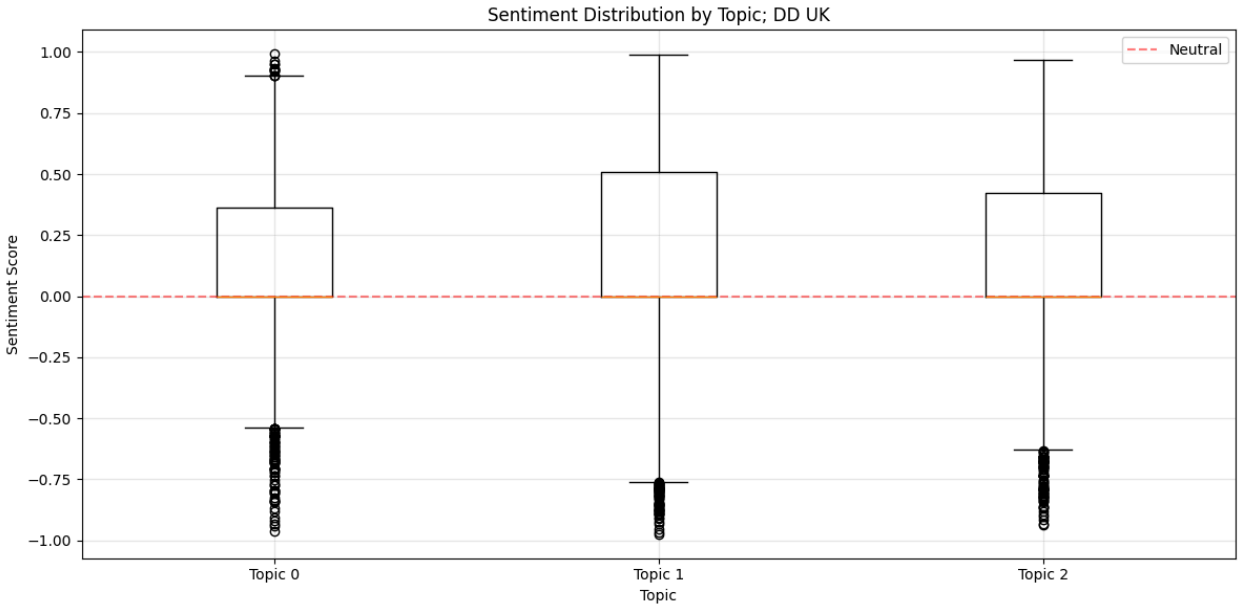

Top Words in Positive Entries (n=2937)
like
good
business
dragons
peter
love
guy
deborah
lol
people
0
100
200
300
400
500
Frequency


Top Words in Negative Entries (n=1549)
business
dragons
people
peter
like
even
deborah
idea
guy
jenny
0
20
40
60
80
100
120
140
Frequency


Topic 0
google
map
phone
buy
idea
year
use
pizza
sister
book
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
Weight


Topic 1
dragon
peter
guy
think
den
product
people
offer
know
bean
0.0000
0.0025
0.0050
0.0075
0.0100
0.0125
0.0150
0.0175
Weight


Topic 2
like
jenny
look
deborah
love
pie
name
reason
money
dragon
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
Weight


Sentiment Distribution by Topic; DD UK
Neutral
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Sentiment Score
Topic 0
Topic 1
Topic 2
Topic

India

Dataset: 598 comments

1 TERM FREQUENCY (from HW4-1):

Top 10 words in POSITIVE entries (203 entries):

ad	(31 times)
like	(25 times)
best	(22 times)
aman	(20 times)
good	(18 times)
shark	(18 times)
h	(16 times)
ki	(15 times)
namita	(14 times)
nahi	(12 times)

Top 10 words in NEGATIVE entries (65 entries):

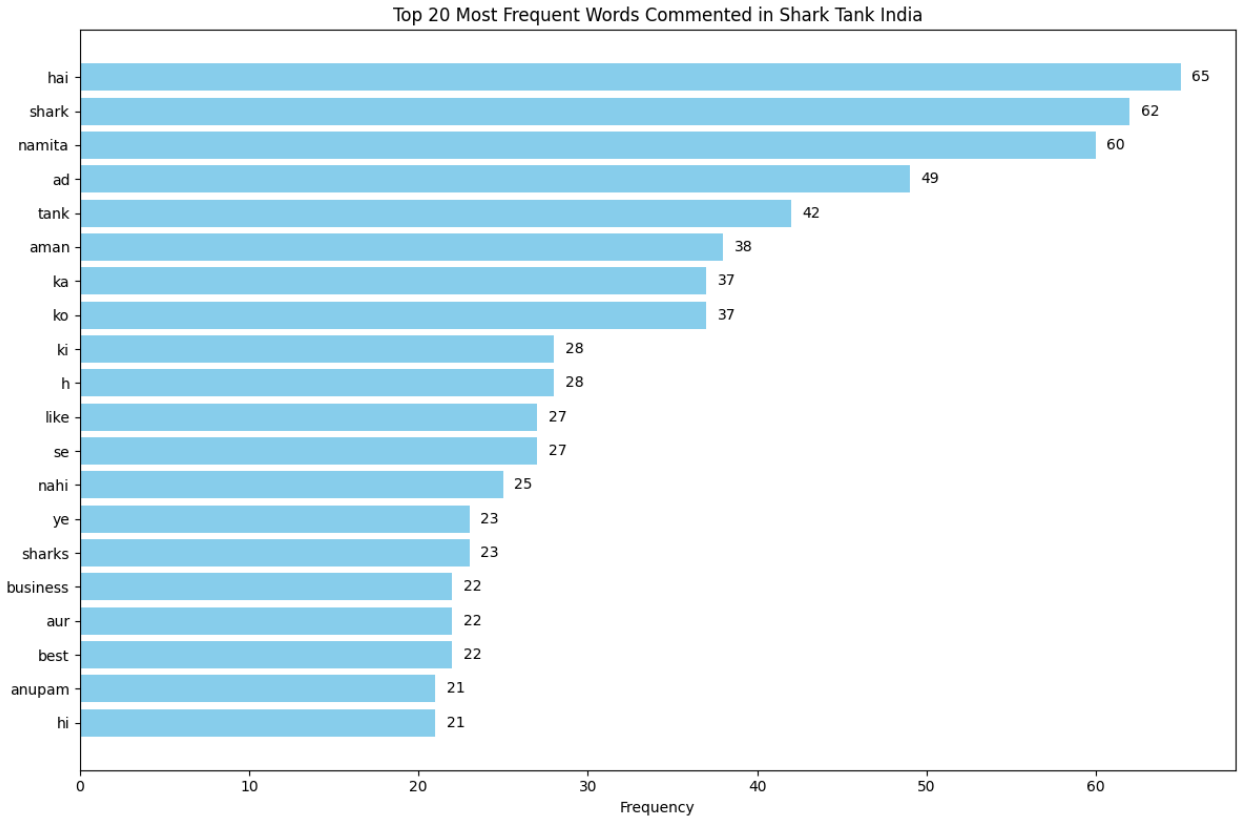
shark	(10 times)
hai	(9 times)
ka	(7 times)
tank	(5 times)
baat	(4 times)
show	(4 times)
namita	(4 times)
se	(4 times)
ko	(4 times)
rha	(4 times)

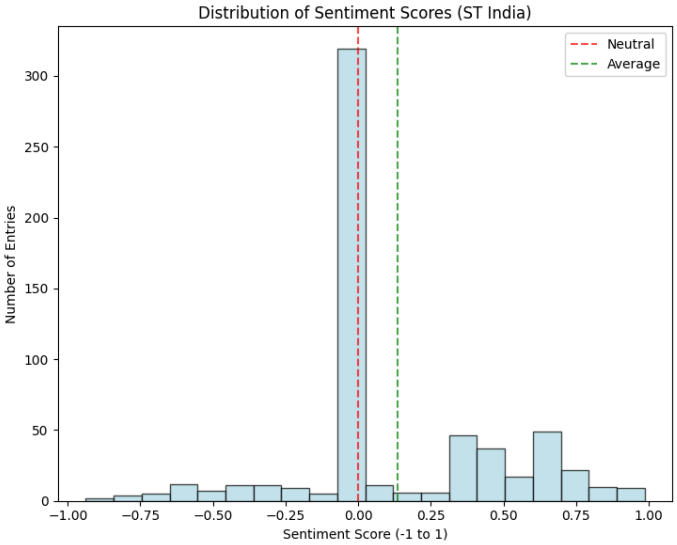
2 SENTIMENT ANALYSIS (from HW4-1):

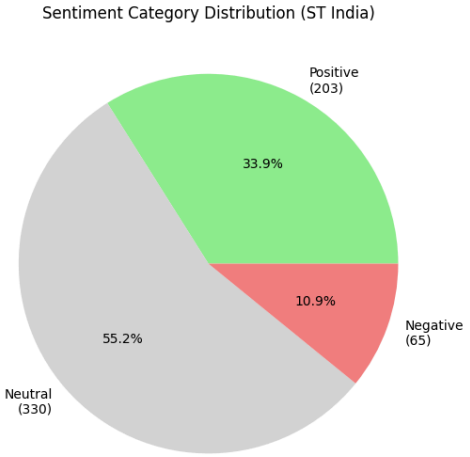
Average sentiment: 0.136
Positive: 203 (33.9%)
Neutral: 330 (55.2%)
Negative: 65 (10.9%)

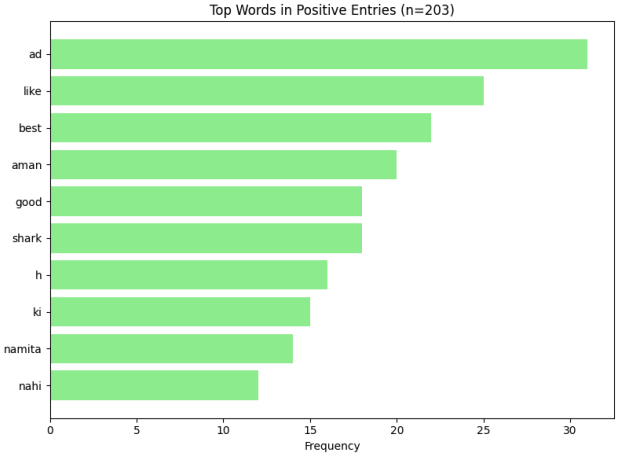
3 TOPIC MODELING (HW4-2):

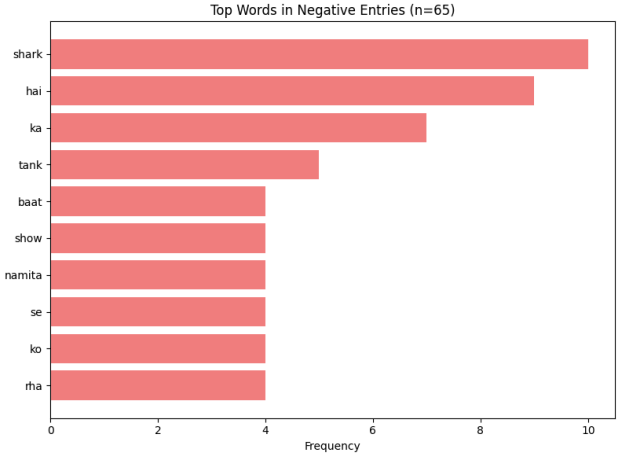
Number of topics: 3
Topic 0 (177 docs, avg sentiment 0.10): namita, hai, nahi
Topic 1 (164 docs, avg sentiment 0.18): hai, best, namita
Topic 2 (257 docs, avg sentiment 0.13): tank, aman, like

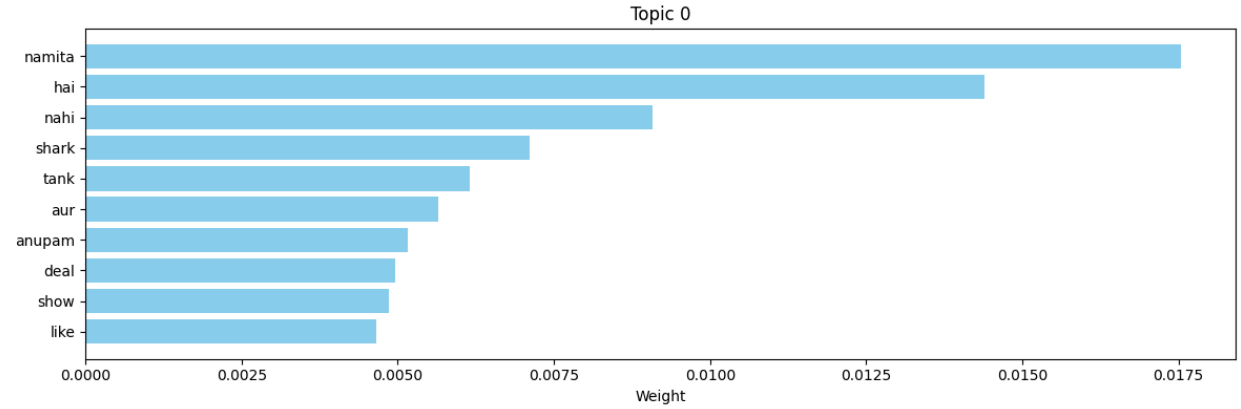

Top 20 Most Frequent Words Commented in Shark Tank India
hai
shark
namita
ad
tank
aman
ka
ko
ki
h
like
se
nahi
ye
sharks
business
aur
best
anupam
hi
65
62
60
49
42
38
37
37
28
28
27
27
25
23
23
22
22
22
21
21
0
10
20
30
40
50
60
Frequency

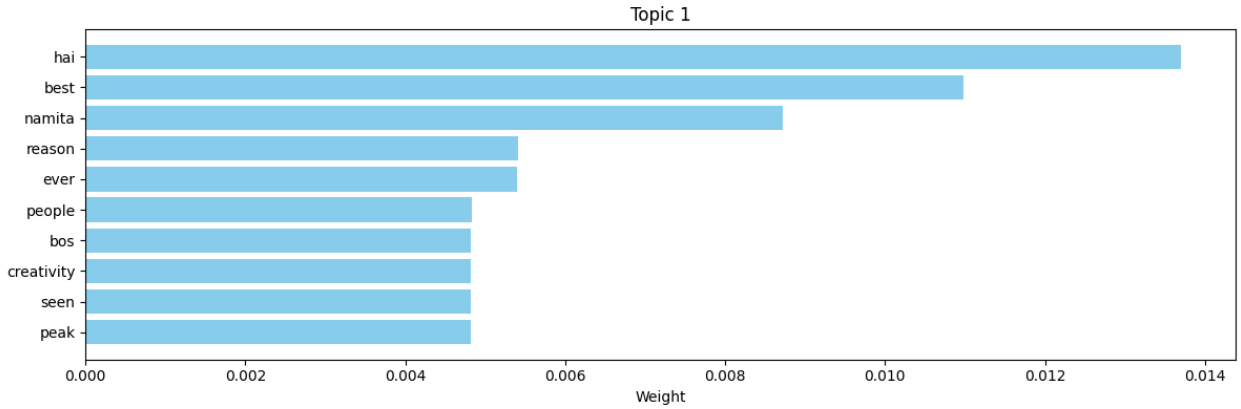

Distribution of Sentiment Scores (ST India)
Neutral
Average
Number of Entries
0
50
100
150
200
250
300
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Sentiment Score (-1 to 1)

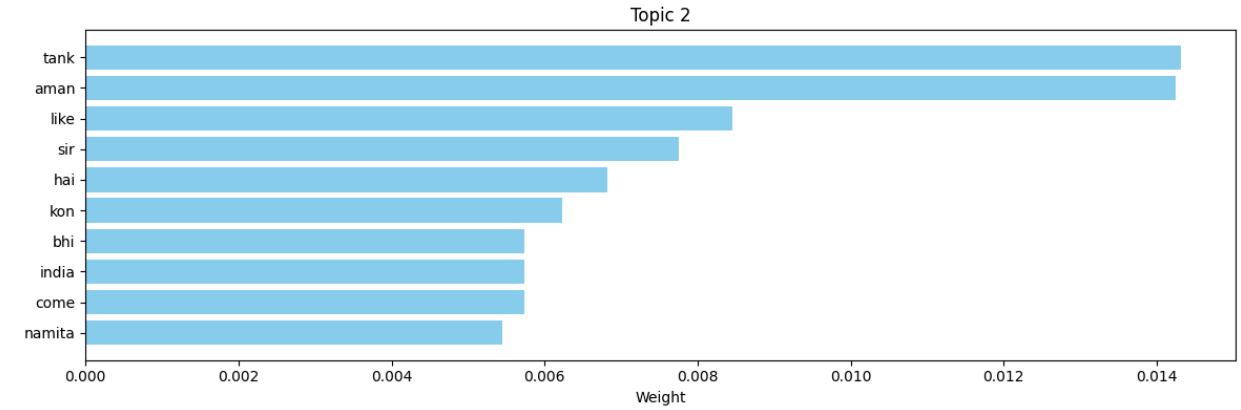

Sentiment Category Distribution (ST India)
Positive
(203)
33.9%
10.9%
Negative
(65)
55.2%
Neutral
(330)

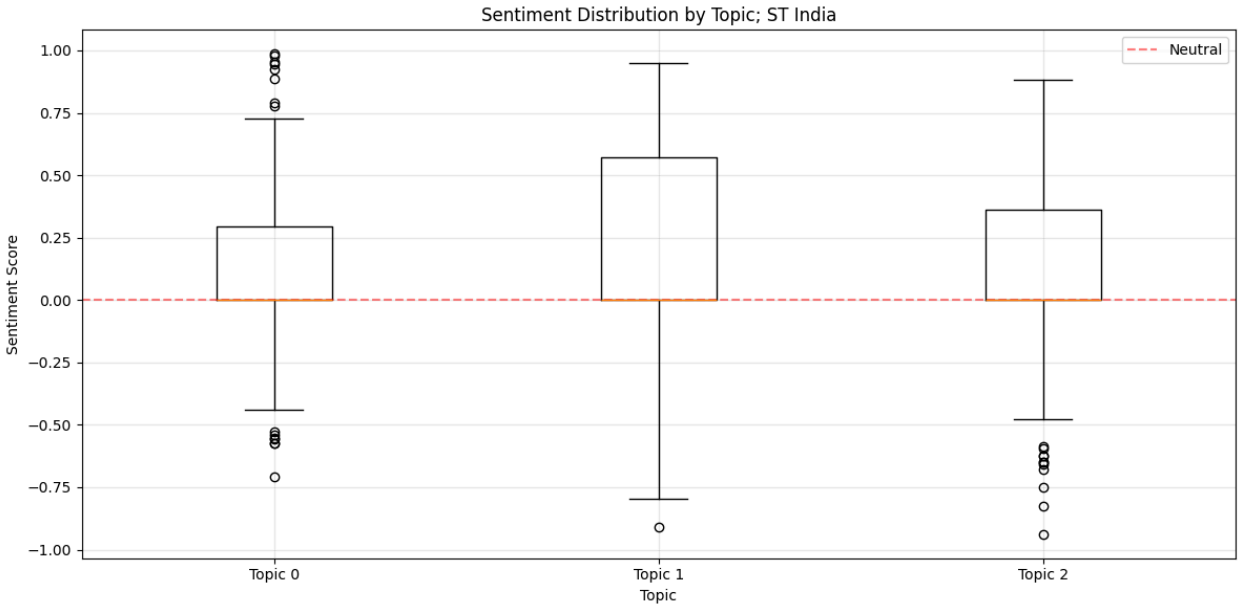

Top Words in Positive Entries (n=203)
ad
like
best
aman
good
shark
h
ki
namita
nahi
0
5
10
15
20
25
30
Frequency


Top Words in Negative Entries (n=65)
shark
hai
ka
tank
baat
show
namita
se
ko
rha
0
2
4
6
8
10
Frequency


Topic 0
namita
hai
nahi
shark
tank
aur
anupam
deal
show
like
0.0000
0.0025
0.0050
0.0075
0.0100
0.0125
0.0150
0.0175
Weight


Topic 1
hai
best
namita
reason
ever
people
bos
creativity
seen
peak
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
Weight


Topic 2
tank
aman
like
sir
hai
kon
bhi
india
come
namita
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
Weight


Sentiment Distribution by Topic; ST India
Neutral
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Sentiment Score
Topic 0
Topic 1
Topic 2
Topic

Australia

Dataset: 2336 documents

1 TERM FREQUENCY (from HW4-1):

Top 10 words in POSITIVE entries (1017 entries):

like	(239 times)
sharks	(156 times)
product	(132 times)
lol	(104 times)
good	(97 times)
business	(93 times)
shark	(88 times)
make	(78 times)
idea	(68 times)
think	(65 times)

Top 10 words in NEGATIVE entries (538 entries):

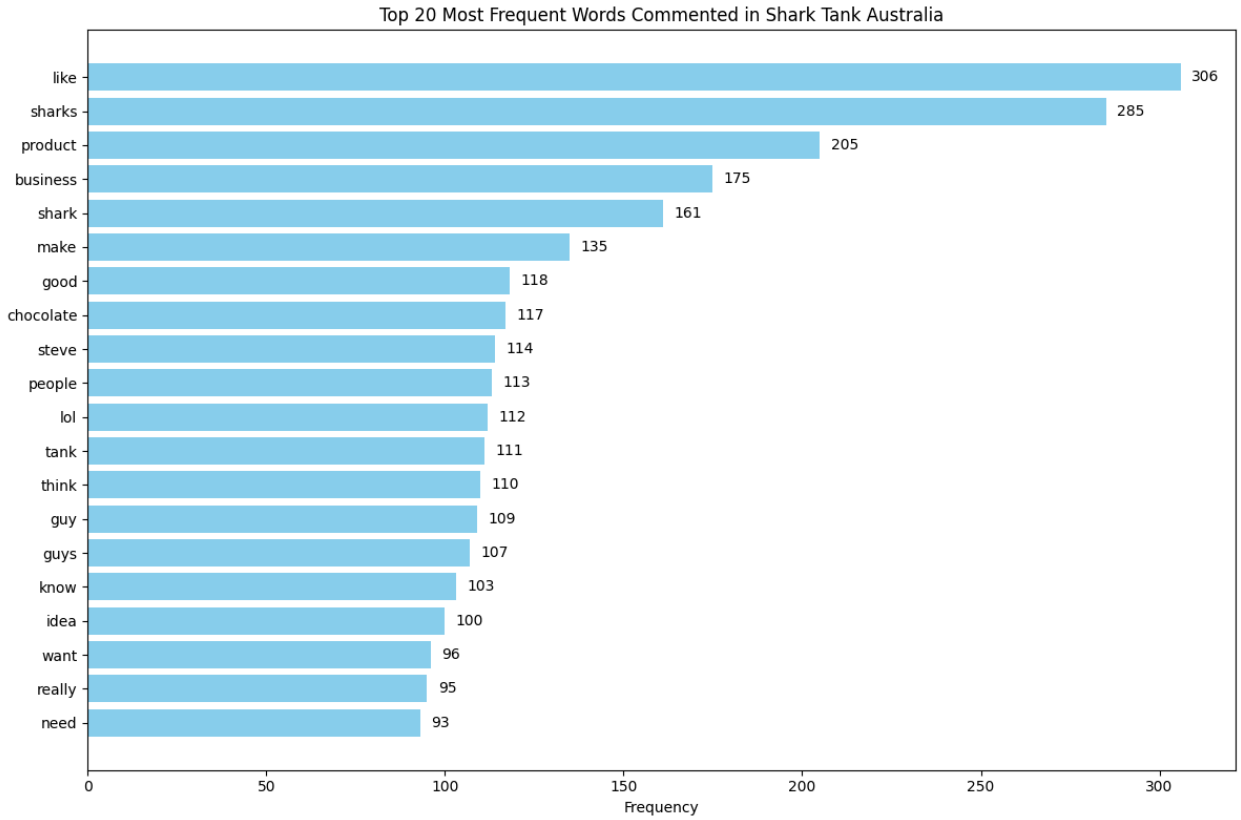
sharks	(79 times)
like	(50 times)
business	(48 times)
product	(47 times)
people	(32 times)
guys	(31 times)
india	(30 times)
make	(30 times)
shark	(29 times)
family	(28 times)

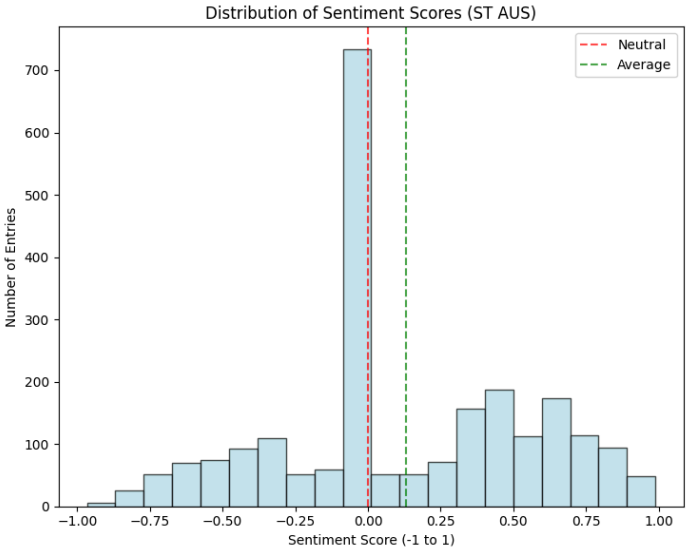
2 SENTIMENT ANALYSIS (from HW4-1):

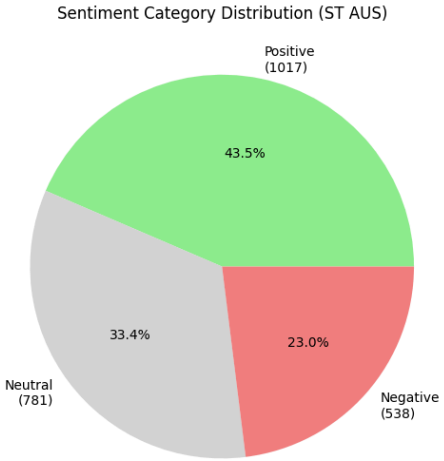
Average sentiment: 0.132
Positive: 1017 (43.5%)
Neutral: 781 (33.4%)
Negative: 538 (23.0%)

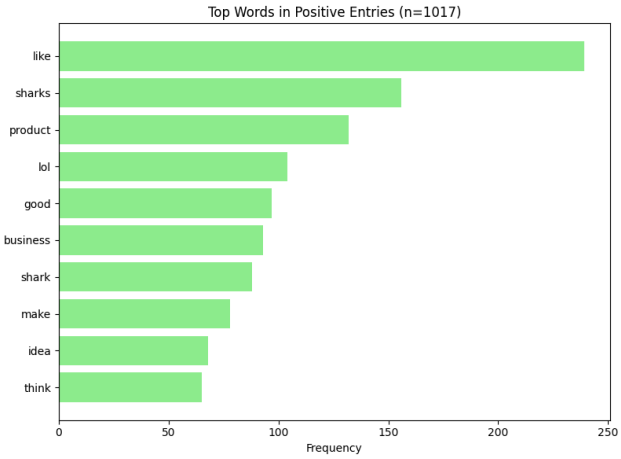
3 TOPIC MODELING (HW4-2):

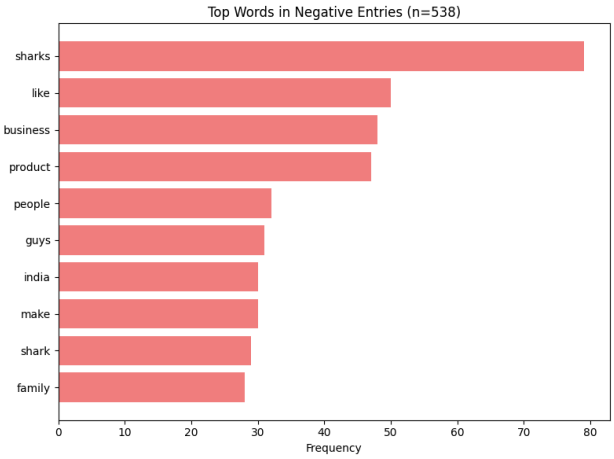
Number of topics: 3
Topic 0 (1197 docs, avg sentiment 0.14): shark, like, product
Topic 1 (567 docs, avg sentiment 0.10): guy, proven, neck
Topic 2 (572 docs, avg sentiment 0.15): like, shark, janine

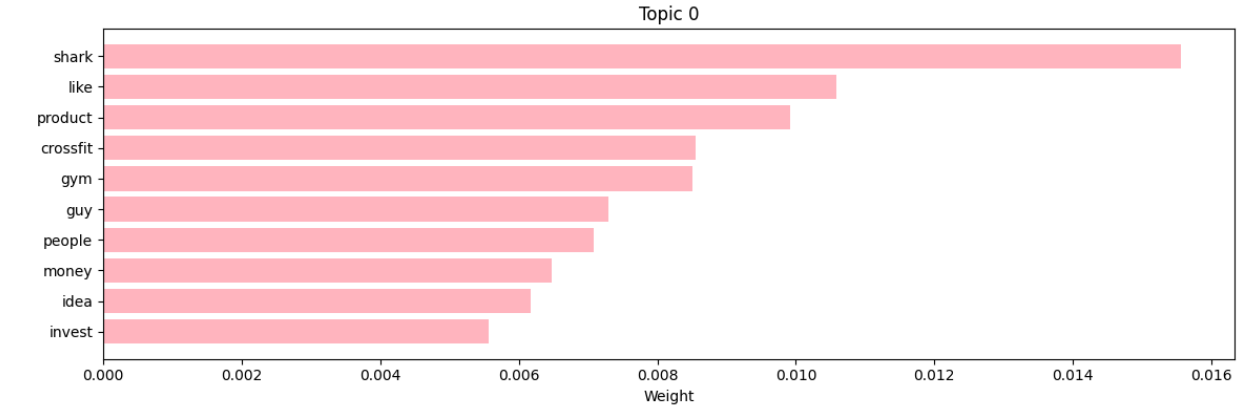

Top 20 Most Frequent Words Commented in Shark Tank Australia
like
sharks
product
business
shark
make
good
chocolate
steve
people
lol
tank
think
guy
guys
know
idea
want
really
need
306
285
205
175
161
135
118
117
114
113
112
111
110
109
107
103
100
96
95
93
0
50
100
150
200
250
300
Frequency

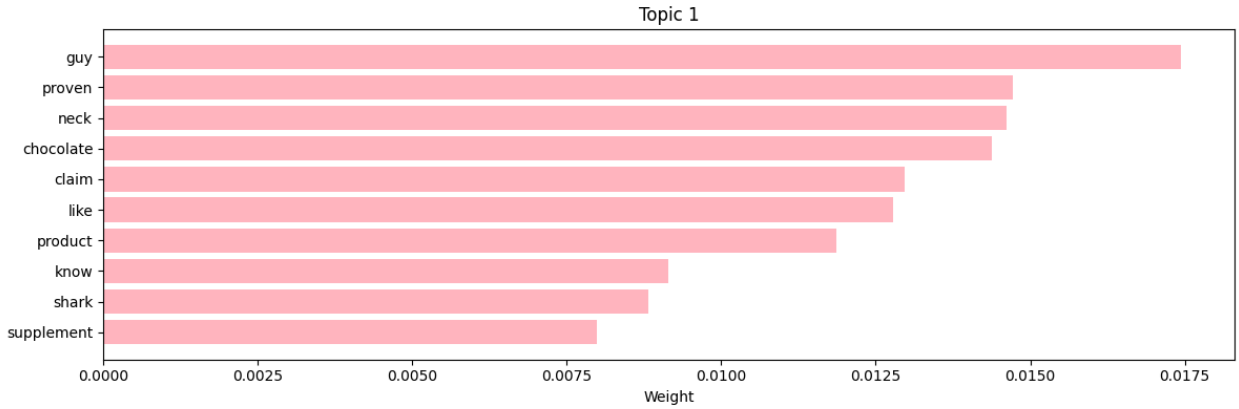

Distribution of Sentiment Scores (ST AUS)
Neutral
Average
Number of Entries
0
100
200
300
400
500
600
700
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Sentiment Score (-1 to 1)

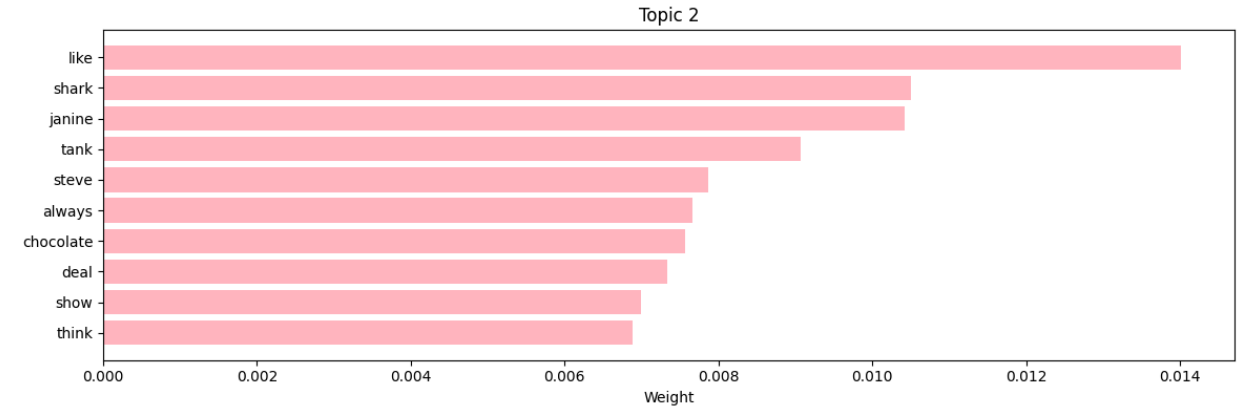

Sentiment Category Distribution (ST AUS)
Positive
(1017)
43.5%
33.4%
23.0%
Neutral
(781)
Negative
(538)

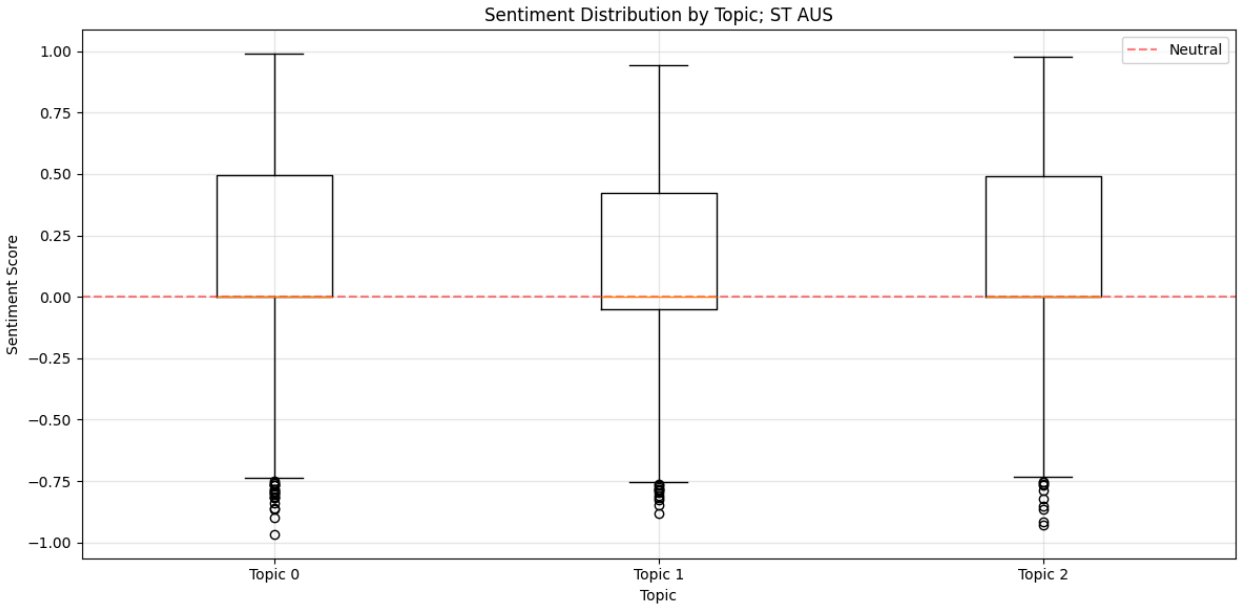

Top Words in Positive Entries (n=1017)
like
sharks
product
lol
good
business
shark
make
idea
think
0
50
100
150
200
250
Frequency


Top Words in Negative Entries (n=538)
sharks
like
business
product
people
guys
india
make
shark
family
0
10
20
30
40
50
60
70
80
Frequency


Topic 0
shark
like
product
crossfit
gym
guy
people
money
idea
invest
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
0.016
Weight


Topic 1
guy
proven
neck
chocolate
claim
like
product
know
shark
supplement
0.0000
0.0025
0.0050
0.0075
0.0100
0.0125
0.0150
0.0175
Weight


Topic 2
like
shark
janine
tank
steve
always
chocolate
deal
show
think
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
Weight


Sentiment Distribution by Topic; ST AUS
Neutral
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Sentiment Score
Topic 0
Topic 1
Topic 2
Topic

U.S.

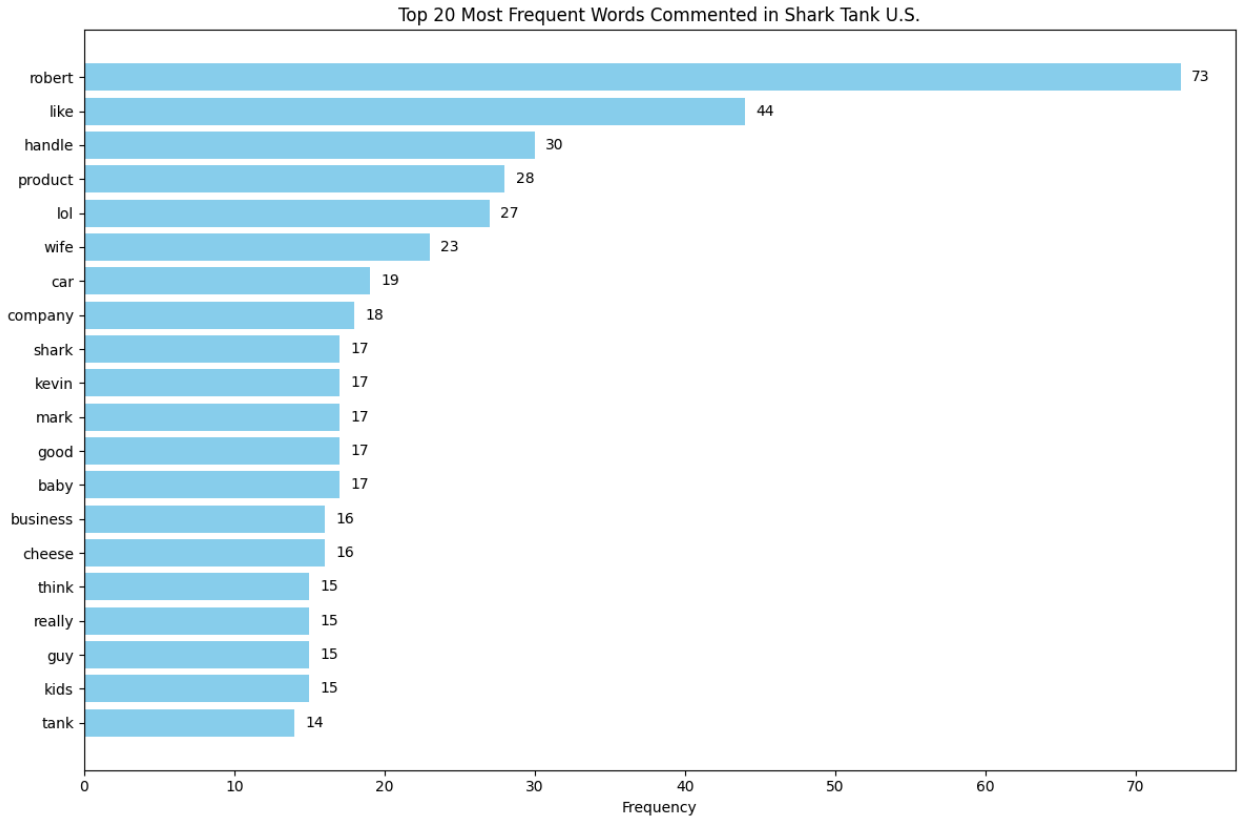
Dataset: 446 documents

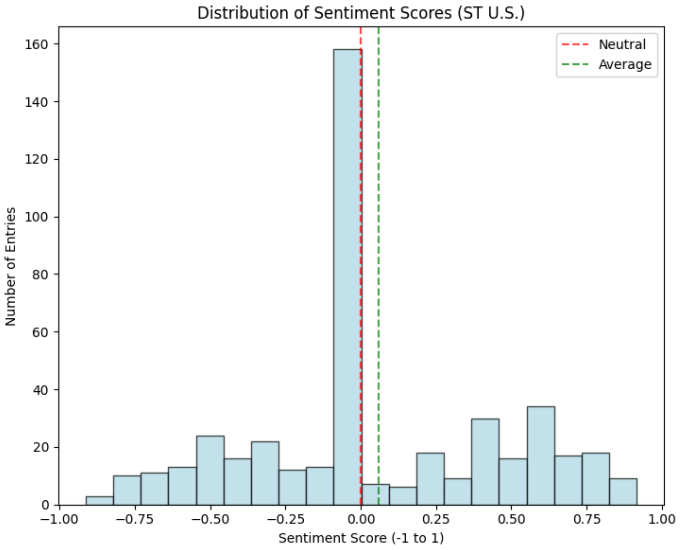
1 TERM FREQUENCY (from HW4-1):
Top 10 words in POSITIVE entries (157 entries):
like (33 times)
robert (23 times)
lol (20 times)
handle (15 times)
kevin (14 times)
think (11 times)
good (11 times)
product (11 times)
car (11 times)
great (10 times)

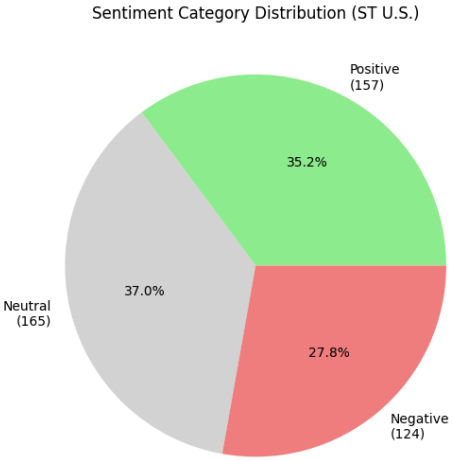
Top 10 words in NEGATIVE entries (124 entries):
robert (20 times)
product (10 times)
like (9 times)
business (8 times)
ever (7 times)
shark (7 times)
bad (7 times)
company (7 times)
people (6 times)
toys (6 times)

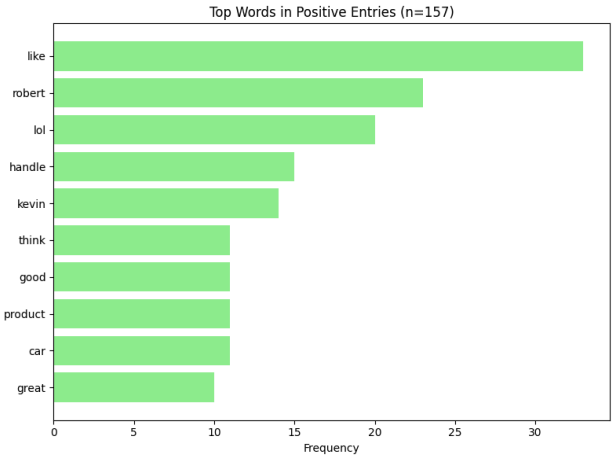
2 SENTIMENT ANALYSIS (from HW4-1):
Average sentiment: 0.060
Positive: 157 (35.2%)
Neutral: 165 (37.0%)
Negative: 124 (27.8%)

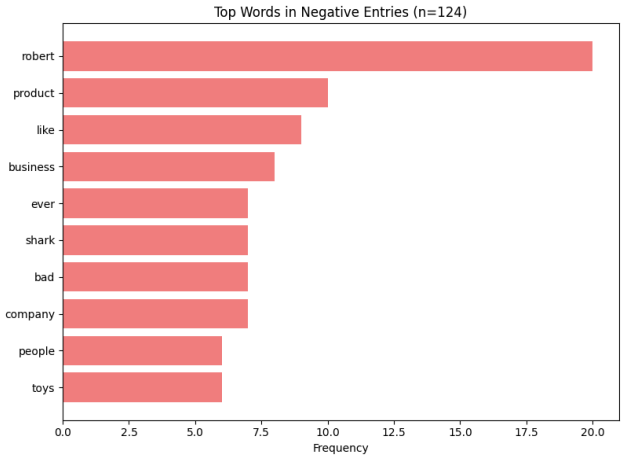
3 TOPIC MODELING (HW4-2):
Number of topics: 3
Topic 0 (184 docs, avg sentiment 0.09): robert, like, product
Topic 1 (116 docs, avg sentiment 0.09): like, toy, lol
Topic 2 (146 docs, avg sentiment -0.01): robert, handle, car

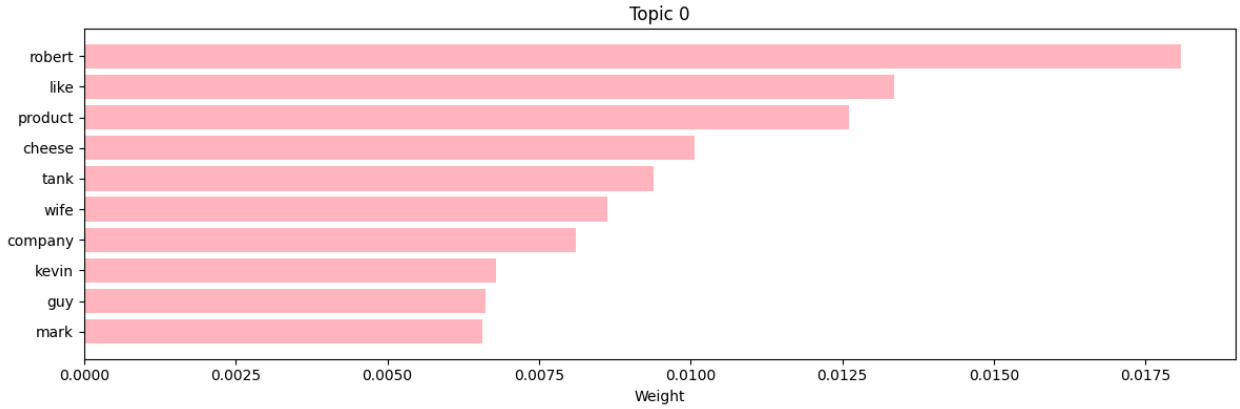

Top 20 Most Frequent Words Commented in Shark Tank U.S.
robert
like
handle
product
lol
wife
car
company
shark
kevin
mark
good
baby
business
cheese
think
really
guy
kids
tank
73
44
30
28
27
23
19
18
17
17
17
17
17
16
16
15
15
15
15
14
0
10
20
30
40
50
60
70
Frequency

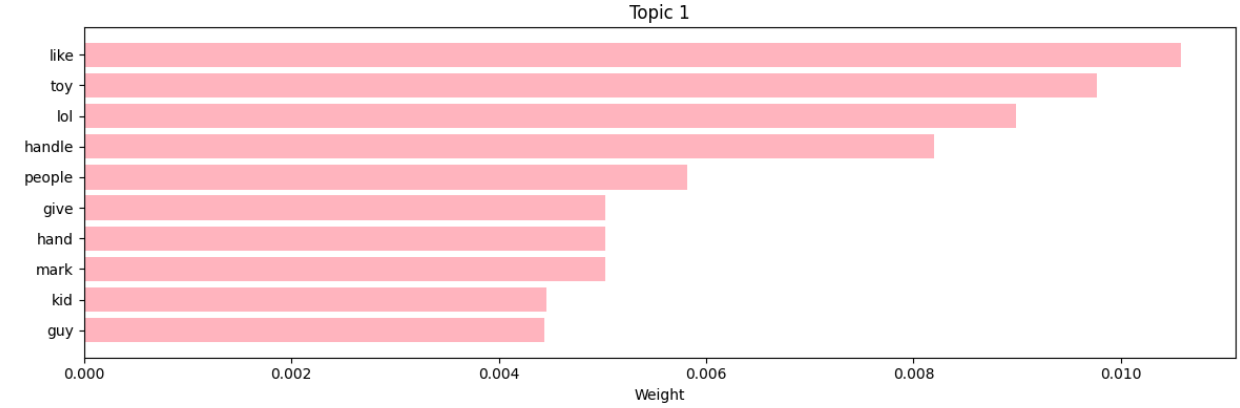

Distribution of Sentiment Scores (ST U.S.)
Neutral
Average
Number of Entries
0
20
40
60
80
100
120
140
160
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Sentiment Score (-1 to 1)

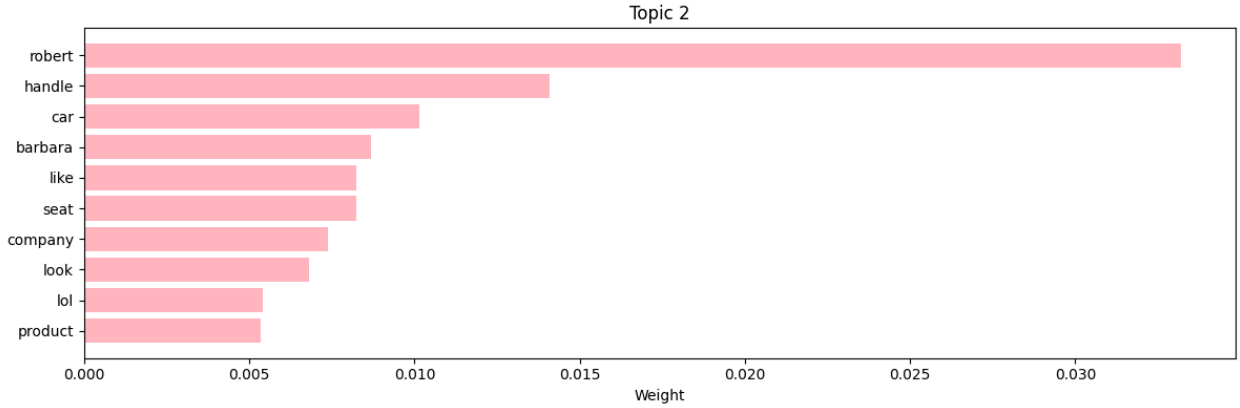

Sentiment Category Distribution (ST U.S.)
Positive
(157)
35.2%
Neutral
(165)
37.0%
Negative
(124)
27.8%

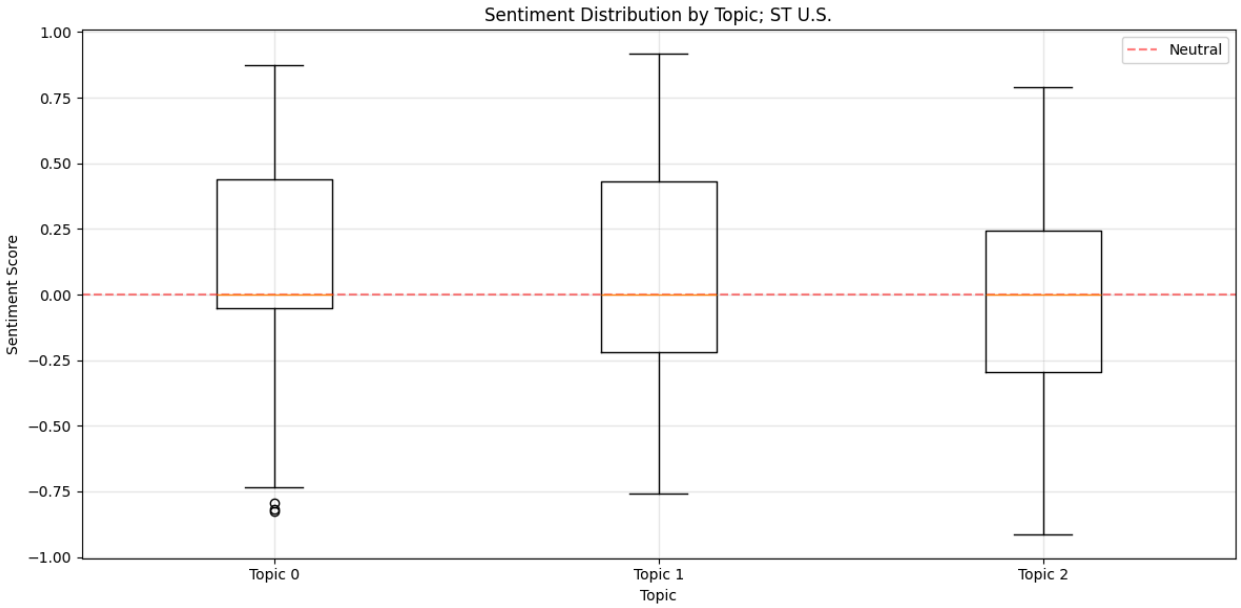

Top Words in Positive Entries (n=157)
like
robert
lol
handle
kevin
think
good
product
car
great
0
5
10
15
20
25
30
Frequency


Top Words in Negative Entries (n=124)
robert
product
like
business
ever
shark
bad
company
people
toys
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Frequency


Topic 0
robert
like
product
cheese
tank
wife
company
kevin
guy
mark
0.0000
0.0025
0.0050
0.0075
0.0100
0.0125
0.0150
0.0175
Weight


Topic 1
like
toy
lol
handle
people
give
hand
mark
kid
guy
0.000
0.002
0.004
0.006
0.008
0.010
Weight


Topic 2
robert
handle
car
barbara
like
seat
company
look
lol
product
0.000
0.005
0.010
0.015
0.020
0.025
0.030
Weight


Sentiment Distribution by Topic; ST U.S.
Neutral
Sentiment Score
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Topic 0
Topic 1
Topic 2
Topic