#Preprocessing, EDA and Evaluation

```
!pip install transformers
!pip install datasets
```

```
Looking in indexes: https://pypi.org/simple, https://us-
python.pkg.dev/colab-wheels/public/simple/
Collecting transformers
  Downloading transformers-4.28.1-py3-none-any.whl (7.0 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 7.0/7.0 MB 75.0 MB/s eta
0:00:00
ent already satisfied: tqdm>=4.27 in /usr/local/lib/python3.9/dist-
packages (from transformers) (4.65.0)
Requirement already satisfied: requests in
/usr/local/lib/python3.9/dist-packages (from transformers) (2.27.1)
Collecting huggingface-hub<1.0,>=0.11.0
  Downloading huggingface_hub-0.13.4-py3-none-any.whl (200 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 200.1/200.1 kB 22.6 MB/s eta
0:00:00
ent already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.9/dist-
packages (from transformers) (6.0)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.9/dist-packages (from transformers) (1.22.4)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.9/dist-packages (from transformers)
(2022.10.31)
Requirement already satisfied: filelock in
/usr/local/lib/python3.9/dist-packages (from transformers) (3.11.0)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.9/dist-packages (from transformers) (23.0)
Collecting tokenizers!=0.11.3,<0.14,>=0.11.1
  Downloading tokenizers-0.13.3-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (7.8 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 7.8/7.8 MB 91.6 MB/s eta
0:00:00
ent already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.9/dist-packages (from huggingface-
hub<1.0,>=0.11.0->transformers) (4.5.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.9/dist-packages (from requests->transformers)
(2022.12.7)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.9/dist-packages (from requests->transformers)
(3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/usr/local/lib/python3.9/dist-packages (from requests->transformers)
(1.26.15)
Requirement already satisfied: charset-normalizer~=2.0.0 in
/usr/local/lib/python3.9/dist-packages (from requests->transformers)
(2.0.12)
```

```
Installing collected packages: tokenizers, huggingface-hub,
transformers
Successfully installed huggingface-hub-0.13.4 tokenizers-0.13.3
transformers-4.28.1
Looking in indexes: https://pypi.org/simple, https://us-
python.pkg.dev/colab-wheels/public/simple/
Collecting datasets
  Downloading datasets-2.11.0-py3-none-any.whl (468 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 468.7/468.7 kB 30.6 MB/s eta
0:00:00
ent already satisfied: numpy>=1.17 in /usr/local/lib/python3.9/dist-
packages (from datasets) (1.22.4)
Collecting dill<0.3.7,>=0.3.0
  Downloading dill-0.3.6-py3-none-any.whl (110 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 110.5/110.5 kB 14.8 MB/s eta
0:00:00
ent already satisfied: tqdm>=4.62.1 in /usr/local/lib/python3.9/dist-
packages (from datasets) (4.65.0)
Requirement already satisfied: fsspec[http]>=2021.11.1 in
/usr/local/lib/python3.9/dist-packages (from datasets) (2023.4.0)
Collecting multiprocess
  Downloading multiprocess-0.70.14-py39-none-any.whl (132 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 132.9/132.9 kB 17.9 MB/s eta
0:00:00
ent already satisfied: pyarrow>=8.0.0 in
/usr/local/lib/python3.9/dist-packages (from datasets) (9.0.0)
Requirement already satisfied: huggingface-hub<1.0.0,>=0.11.0 in
/usr/local/lib/python3.9/dist-packages (from datasets) (0.13.4)
Requirement already satisfied: packaging in
/usr/local/lib/python3.9/dist-packages (from datasets) (23.0)
Requirement already satisfied: pandas in
/usr/local/lib/python3.9/dist-packages (from datasets) (1.5.3)
Collecting responses<0.19
  Downloading responses-0.18.0-py3-none-any.whl (38 kB)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.9/dist-packages (from datasets) (6.0)
Collecting aiohttp
  Downloading aiohttp-3.8.4-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.0 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 1.0/1.0 MB 75.0 MB/s eta
0:00:00
ent already satisfied: requests>=2.19.0 in
/usr/local/lib/python3.9/dist-packages (from datasets) (2.27.1)
Collecting xxhash
  Downloading xxhash-3.2.0-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (212 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 212.2/212.2 kB 25.2 MB/s eta
0:00:00
ent already satisfied: charset-normalizer<4.0,>=2.0 in
/usr/local/lib/python3.9/dist-packages (from aiohttp->datasets)
```

```
(2.0.12)
Collecting frozenlist>=1.1.1
  Downloading frozenlist-1.3.3-cp39-cp39-
manylinux_2_5_x86_64.manylinux1_x86_64.manylinux_2_17_x86_64.manylinux
2014_x86_64.whl (158 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 158.8/158.8 kB 21.3 MB/s eta
0:00:00
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (264 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 264.6/264.6 kB 24.5 MB/s eta
0:00:00
ultidict<7.0,>=4.5
  Downloading multidict-6.0.4-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (114 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 114.2/114.2 kB 15.3 MB/s eta
0:00:00
ent already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.9/dist-
packages (from aiohttp->datasets) (22.2.0)
Collecting async-timeout<5.0,>=4.0.0a3
  Downloading async_timeout-4.0.2-py3-none-any.whl (5.8 kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.9/dist-packages (from huggingface-
hub<1.0.0,>=0.11.0->datasets) (3.11.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.9/dist-packages (from huggingface-
hub<1.0.0,>=0.11.0->datasets) (4.5.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.9/dist-packages (from requests>=2.19.0-
>datasets) (2022.12.7)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.9/dist-packages (from requests>=2.19.0-
>datasets) (3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/usr/local/lib/python3.9/dist-packages (from requests>=2.19.0-
>datasets) (1.26.15)
Requirement already satisfied: python-dateutil>=2.8.1 in
/usr/local/lib/python3.9/dist-packages (from pandas->datasets) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.9/dist-packages (from pandas->datasets)
(2022.7.1)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.9/dist-packages (from python-dateutil>=2.8.1-
>pandas->datasets) (1.16.0)
Installing collected packages: xxhash, multidict, frozenlist, dill,
async-timeout, yarl, responses, multiprocess, aiosignal, aiohttp,
datasets
Successfully installed aiohttp-3.8.4 aiosignal-1.3.1 async-timeout-
4.0.2 datasets-2.11.0 dill-0.3.6 frozenlist-1.3.3 multidict-6.0.4
multiprocess-0.70.14 responses-0.18.0 xxhash-3.2.0 yarl-1.8.2
```

```
from datasets import load_dataset

#Importing the dataset from hugging face
dataset =  load_dataset('jeremyf/fanfiction_z')
```

```
{"model_id":"f3e4e73d73324b86b270c723286bc84b","version_major":2,"vers
ion_minor":0}
```

```
Downloading and preparing dataset json/jeremyf--fanfiction_z to
/root/.cache/huggingface/datasets/jeremyf___json/jeremyf--
fanfiction_z-ee1a1447c757b7ff/0.0.0/
fe5dd6ea2639a6df622901539cb550cf8797e5a6b2dd7af1cf934bed8e233e6e...
```

```
{"model_id":"15051b5c076c42fc93d12322b9794ebd","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"b339064bdac041ef878c8633a8cb7caf","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"1ee6b5b479d9415f9a564e3269d316c1","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"2cd37f3097e54fd8ba1ccea69bf60ddb","version_major":2,"vers
ion_minor":0}
```

```
Dataset json downloaded and prepared to
/root/.cache/huggingface/datasets/jeremyf___json/jeremyf--
fanfiction_z-ee1a1447c757b7ff/0.0.0/
fe5dd6ea2639a6df622901539cb550cf8797e5a6b2dd7af1cf934bed8e233e6e.
Subsequent calls will reuse this data.
```

```
{"model_id":"dd54e990ede14e08af71b3abaaff6468","version_major":2,"vers
ion_minor":0}
```

```
dataset
```

```
DatasetDict({
    train: Dataset({
        features: ['story', 'title', 'category'],
        num_rows: 943
    })
})
```

```
import pandas as pd
import re

#Fetch the store list
story_list = [[i] for i in list(dataset['train']['story'])]
df = pd.DataFrame(story_list,columns=['story'])

#Extract the titles and the categories of the stories
df['title'] = dataset['train']['title']
df['category'] = dataset['train']['category']
```

```
#Clean up line breaks, extra punctuations etc
df['story_cleaned'] = df['story'].apply(lambda x: x.replace('**',''))
df['story_cleaned'] = df['story_cleaned'].apply(lambda x: x.replace("\
n\n", '\n'))
df['story_cleaned'] = df['story_cleaned'].apply(lambda x:
re.sub(r'^([a-z])\1*$', 'WORD', x))
# for i in range(df.shape[0]):
#     title = df.iloc[i]['title'].partition('by')[0]
#     df.iloc[i]['story_cleaned'].replace(title,'')
#     print(df.iloc[i]['story_cleaned'])

df
```

```
                                                  story  \
0    just a question\n\nWhat I thought when I first...
1    pain\n\nIt may be sad but it can be better\n\n...
2    Zombie Mayhem\n\nZombie Mayhem\n\nIn 2050, the...
3    Zero Wing 2: AYBASBTU Game Script\n\nZERO WING...
4    Not prosaic\n\nDisclaimer, I do not claim to o...
..                                                 ...
938  More\n\n**I've watched this movie 3 times this...
939  Change Comes From Within\n\n**As usual, I own ...
940  Changes\n\nYou know that feeling where people ...
941  Undying Hero\n\nZero percent.\n\nVictory canno...
942  Unlosing Haunted Mansion\n\n**Disclaimer**: I ...

                                                  title  \
0                       just a question by multifics123
1                                pain by Isaiah Thomas
2                              Zombie Mayhem by kelbey342
3    Zero Wing 2: AYBASBTU Game Script by Rodrigo Shin
4                                 Not prosaic by Fault
..                                                 ...
938                                 More by Lizwontcry
939         Change Comes From Within by Dominican Girl
940                               Changes by michieexx3
941                             Undying Hero by Arukoir
942                  Unlosing Haunted Mansion by Smak64

                       category  \
0                          Zom-B
1                          Zenda
2                 Zoombie Blondes
3                       Zero Wing
4                       Zero Wing
..                            ...
938  Zack and Miri Make a Porno
939  Zack and Miri Make a Porno
940  Zack and Miri Make a Porno
941                       Z.H.P.
```

```
942                    Z.H.P.

                                                story_cleaned
0    just a question\nWhat I thought when I first r...
1    pain\nIt may be sad but it can be better\nThe ...
2    Zombie Mayhem\nZombie Mayhem\nIn 2050, the mil...
3    Zero Wing 2: AYBASBTU Game Script\nZERO WING 2...
4    Not prosaic\nDisclaimer, I do not claim to own...
..                                                 ...
938  More\nI've watched this movie 3 times this wee...
939  Change Comes From Within\nAs usual, I own noth...
940  Changes\nYou know that feeling where people be...
941  Undying Hero\nZero percent.\nVictory cannot be...
942  Unlosing Haunted Mansion\nDisclaimer: I do not...

[943 rows x 4 columns]
```

EDA

```
from matplotlib import pyplot as plt
import matplotlib

#Find and print different types of categories
df.describe()
df['category'].unique()
```

```
array(['Zom-B', 'Zenda', 'Zoombie Blondes', 'Zero Wing', 'Zoom',
       'Zettai Karen Children', 'Zero Day', 'Zombie Prom',
       'Zombie Survival Guide', 'Zeke and Luther', 'Zen', 'Zoop',
       "Zanna, Don't!", 'Zenonia', 'Zetman', 'Z Nation', 'Zatch Bell',
       'Zombieland', 'Zone of The Enders', "Zeke's Pad", 'Zoey 101',
       'Zinda/ज़िंदा', 'Zoo Tycoon', 'Zenon', 'Zombie Fallout',
       'Zombies, Run!', 'Z for Zachariah', 'Zapped', 'Zoombinis',
       'Zombie Powder', 'Zorro', 'Zoey 101, iCarly', 'Zoolander',
'Zoids',
       'Zevo-3', 'Zero Dark Thirty', 'Zombie-Loan', 'Zodiac P.I.',
       'Zeta Project', 'Zathura', 'Zack and Miri Make a Porno',
'Z.H.P.'],
      dtype=object)
```

```
#Find out dominant categories
category_count = df['category'].value_counts()
category_count
```

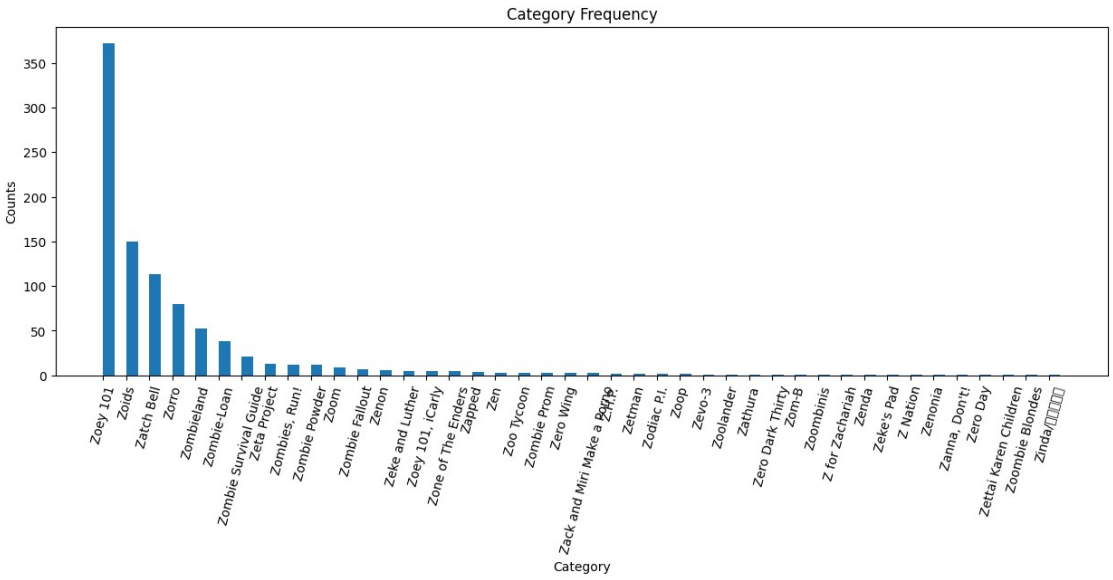
```
Zoey 101                    372
Zoids                       150
Zatch Bell                  113
Zorro                        80
Zombieland                   52
Zombie-Loan                  38
Zombie Survival Guide        21
```

```
Zeta Project                    13
Zombies, Run!                   12
Zombie Powder                   12
Zoom                             9
Zombie Fallout                   7
Zenon                            6
Zeke and Luther                  5
Zoey 101, iCarly                 5
Zone of The Enders               5
Zapped                           4
Zen                              3
Zoo Tycoon                       3
Zombie Prom                      3
Zero Wing                        3
Zack and Miri Make a Porno       3
Z.H.P.                           2
Zetman                           2
Zodiac P.I.                      2
Zoop                             2
Zevo-3                           1
Zoolander                        1
Zathura                          1
Zero Dark Thirty                 1
Zom-B                            1
Zoombinis                        1
Z for Zachariah                  1
Zenda                            1
Zeke's Pad                       1
Z Nation                         1
Zenonia                          1
Zanna, Don't!                    1
Zero Day                         1
Zettai Karen Children            1
Zoombie Blondes                  1
Zinda/ज़ंदा            1
Name: category, dtype: int64
```

```
#Plot counts of stories per category
plt.figure(figsize=(15,5))
plt.xticks(rotation = 75)
plt.bar(category_count.index,category_count,align='edge', width= 0.5)
plt.title('Category Frequency')
plt.xlabel('Category')
plt.ylabel('Counts')
plt.show()
```

```
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2332 (\N{DEVANAGARI LETTER JA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
```

```
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Matplotlib currently does not support Devanagari
natively.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2364 (\N{DEVANAGARI SIGN NUKTA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2306 (\N{DEVANAGARI SIGN ANUSVARA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2342 (\N{DEVANAGARI LETTER DA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2366 (\N{DEVANAGARI VOWEL SIGN AA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
```



Expand Contractions

```
import re
contractions_dict = { "ain't": "are not","'s":" is","aren't": "are
not",
                     "can't": "cannot","can't've": "cannot have",
                     "'cause": "because","could've": "could
have","couldn't": "could not",
                     "couldn't've": "could not have", "didn't": "did
not","doesn't": "does not",
                     "don't": "do not","hadn't": "had
not","hadn't've": "had not have",
```

```
                    "hasn't": "has not","haven't": "have not","he'd":
"he would",
                    "he'd've": "he would have","he'll": "he will",
"he'll've": "he will have",
                    "how'd": "how did","how'd'y": "how do
you","how'll": "how will",
                    "I'd": "I would", "I'd've": "I would
have","I'll": "I will",
                    "I'll've": "I will have","I'm": "I am","I've": "I
have", "isn't": "is not",
                    "it'd": "it would","it'd've": "it would
have","it'll": "it will",
                    "it'll've": "it will have", "let's": "let
us","ma'am": "madam",
                    "mayn't": "may not","might've": "might
have","mightn't": "might not",
                    "mightn't've": "might not have","must've": "must
have","mustn't": "must not",
                    "mustn't've": "must not have", "needn't": "need
not",
                    "needn't've": "need not have","o'clock": "of the
clock","oughtn't": "ought not",
                    "oughtn't've": "ought not have","shan't": "shall
not","sha'n't": "shall not",
                    "shan't've": "shall not have","she'd": "she
would","she'd've": "she would have",
                    "she'll": "she will", "she'll've": "she will
have","should've": "should have",
                    "shouldn't": "should not", "shouldn't've":
"should not have","so've": "so have",
                    "that'd": "that would","that'd've": "that would
have", "there'd": "there would",
                    "there'd've": "there would have", "they'd": "they
would",
                    "they'd've": "they would have","they'll": "they
will",
                    "they'll've": "they will have", "they're": "they
are","they've": "they have",
                    "to've": "to have","wasn't": "was not","we'd":
"we would",
                    "we'd've": "we would have","we'll": "we
will","we'll've": "we will have",
                    "we're": "we are","we've": "we have", "weren't":
"were not","what'll": "what will",
                    "what'll've": "what will have","what're": "what
are", "what've": "what have",
                    "when've": "when have","where'd": "where did",
"where've": "where have",
                    "who'll": "who will","who'll've": "who will
have","who've": "who have",
```

```
                    "why've": "why have","will've": "will
have","won't": "will not",
                    "won't've": "will not have", "would've": "would
have","wouldn't": "would not",
                    "wouldn't've": "would not have","y'all": "you
all", "y'all'd": "you all would",
                    "y'all'd've": "you all would have","y'all're":
"you all are",
                    "y'all've": "you all have", "you'd": "you
would","you'd've": "you would have",
                    "you'll": "you will","you'll've": "you will
have", "you're": "you are",
                    "you've": "you have"}

# Regular expression for finding contractions
contractions_re=re.compile('(%s)' %
'|'.join(contractions_dict.keys()))

# Function for expanding contractions
def expand_contractions(text,contractions_dict=contractions_dict):
  def replace(match):
    return contractions_dict[match.group(0)]
  return contractions_re.sub(replace, text)


# Expanding Contractions in the reviews
df['story_cleaned']=df['story_cleaned'].apply(lambda
x:expand_contractions(x))
```

Lower case

```
df['story_cleaned']=df['story_cleaned'].apply(lambda x: x.lower())

df
```

```
                                                  story  \
0    just a question\n\nWhat I thought when I first...
1    pain\n\nIt may be sad but it can be better\n\n...
2    Zombie Mayhem\n\nZombie Mayhem\n\nIn 2050, the...
3    Zero Wing 2: AYBASBTU Game Script\n\nZERO WING...
4    Not prosaic\n\nDisclaimer, I do not claim to o...
..                                                 ...
938  More\n\n**I've watched this movie 3 times this...
939  Change Comes From Within\n\n**As usual, I own ...
940  Changes\n\nYou know that feeling where people ...
941  Undying Hero\n\nZero percent.\n\nVictory canno...
942  Unlosing Haunted Mansion\n\n**Disclaimer**: I ...

                                                  title  \
0                    just a question by multifics123
1                              pain by Isaiah Thomas
```

```
2                            Zombie Mayhem by kelbey342
3    Zero Wing 2: AYBASBTU Game Script by Rodrigo Shin
4                                 Not prosaic by Fault
..                                                 ...
938                                 More by Lizwontcry
939         Change Comes From Within by Dominican Girl
940                               Changes by michieexx3
941                             Undying Hero by Arukoir
942                  Unlosing Haunted Mansion by Smak64

                       category  \
0                         Zom-B
1                         Zenda
2                 Zoombie Blondes
3                       Zero Wing
4                       Zero Wing
..                            ...
938  Zack and Miri Make a Porno
939  Zack and Miri Make a Porno
940  Zack and Miri Make a Porno
941                        Z.H.P.
942                        Z.H.P.

                                            story_cleaned
0    just a question\nwhat i thought when i first r...
1    pain\nit may be sad but it can be better\nthe ...
2    zombie mayhem\nzombie mayhem\nin 2050, the mil...
3    zero wing 2: aybasbtu game script\nzero wing 2...
4    not prosaic\ndisclaimer, i do not claim to own...
..                                                 ...
938  more\ni have watched this movie 3 times this w...
939  change comes from within\nas usual, i own noth...
940  changes\nyou know that feeling where people be...
941  undying hero\nzero percent.\nvictory cannot be...
942  unlosing haunted mansion\ndisclaimer: i do not...

[943 rows x 4 columns]
```

Remove digits and words containing digits

```
df['story_cleaned']=df['story_cleaned'].apply(lambda x: re.sub('\w*\d\
w*','', x))
```

Remove Punctuations

```
import string
df['story_cleaned']=df['story_cleaned'].apply(lambda x: re.sub('[%s]'
% re.escape(string.punctuation), '', x))

df
```

```
                                                  story  \
0    just a question\n\nWhat I thought when I first...
1    pain\n\nIt may be sad but it can be better\n\n...
2    Zombie Mayhem\n\nZombie Mayhem\n\nIn 2050, the...
3    Zero Wing 2: AYBASBTU Game Script\n\nZERO WING...
4    Not prosaic\n\nDisclaimer, I do not claim to o...
..                                                 ...
938  More\n\n**I've watched this movie 3 times this...
939  Change Comes From Within\n\n**As usual, I own ...
940  Changes\n\nYou know that feeling where people ...
941  Undying Hero\n\nZero percent.\n\nVictory canno...
942  Unlosing Haunted Mansion\n\n**Disclaimer**: I ...

                                                  title  \
0                       just a question by multifics123
1                                pain by Isaiah Thomas
2                             Zombie Mayhem by kelbey342
3    Zero Wing 2: AYBASBTU Game Script by Rodrigo Shin
4                                  Not prosaic by Fault
..                                                  ...
938                                More by Lizwontcry
939         Change Comes From Within by Dominican Girl
940                              Changes by michieexx3
941                            Undying Hero by Arukoir
942                 Unlosing Haunted Mansion by Smak64

                      category  \
0                         Zom-B
1                         Zenda
2                Zoombie Blondes
3                     Zero Wing
4                     Zero Wing
..                          ...
938  Zack and Miri Make a Porno
939  Zack and Miri Make a Porno
940  Zack and Miri Make a Porno
941                       Z.H.P.
942                       Z.H.P.

                                          story_cleaned
0    just a question\nwhat i thought when i first r...
1    pain\nit may be sad but it can be better\nthe ...
2    zombie mayhem\nzombie mayhem\nin  the military...
3    zero wing  aybasbtu game script\nzero wing   \...
4    not prosaic\ndisclaimer i do not claim to own ...
..                                                 ...
938  more\ni have watched this movie  times this we...
939  change comes from within\nas usual i own nothi...
940  changes\nyou know that feeling where people be...
941  undying hero\nzero percent\nvictory cannot be ...
```

```
942  unlosing haunted mansion\ndisclaimer i do not ...

[943 rows x 4 columns]
```

Removing extra spaces

```
df['story_cleaned']=df['story_cleaned'].apply(lambda x: re.sub(' +','
',x))
```

```
df['story_cleaned']
```

```
0      just a question\nwhat i thought when i first r...
1      pain\nit may be sad but it can be better\nthe ...
2      zombie mayhem\nzombie mayhem\nin the military ...
3      zero wing aybasbtu game script\nzero wing \nal...
4      not prosaic\ndisclaimer i do not claim to own ...
                            ...
938    more\ni have watched this movie times this wee...
939    change comes from within\nas usual i own nothi...
940    changes\nyou know that feeling where people be...
941    undying hero\nzero percent\nvictory cannot be ...
942    unlosing haunted mansion\ndisclaimer i do not ...
Name: story_cleaned, Length: 943, dtype: object
```

```
import spacy

#Loading the model and then lemmatization
nlp = spacy.load('en_core_web_sm', disable = ['parser','ner'])
df['story_lemmatized'] = df['story_cleaned'].apply(lambda x:'
'.join([token.lemma_ for token in list(nlp(x)) if (token.is_stop ==
False)]))
```

```
df
```

```
                                                  story  \
0    just a question\n\nWhat I thought when I first...
1    pain\n\nIt may be sad but it can be better\n\n...
2    Zombie Mayhem\n\nZombie Mayhem\n\nIn 2050, the...
3    Zero Wing 2: AYBASBTU Game Script\n\nZERO WING...
4    Not prosaic\n\nDisclaimer, I do not claim to o...
..                                                 ...
938  More\n\n**I've watched this movie 3 times this...
939  Change Comes From Within\n\n**As usual, I own ...
940  Changes\n\nYou know that feeling where people ...
941  Undying Hero\n\nZero percent.\n\nVictory canno...
942  Unlosing Haunted Mansion\n\n**Disclaimer**: I ...

                                                  title  \
0                      just a question by multifics123
1                                pain by Isaiah Thomas
2                           Zombie Mayhem by kelbey342
```

```
3    Zero Wing 2: AYBASBTU Game Script by Rodrigo Shin
4                                  Not prosaic by Fault
..                                                   ...
938                                  More by Lizwontcry
939         Change Comes From Within by Dominican Girl
940                               Changes by michieexx3
941                            Undying Hero by Arukoir
942                 Unlosing Haunted Mansion by Smak64

                       category  \
0                         Zom-B
1                         Zenda
2                Zoombie Blondes
3                      Zero Wing
4                      Zero Wing
..                           ...
938  Zack and Miri Make a Porno
939  Zack and Miri Make a Porno
940  Zack and Miri Make a Porno
941                       Z.H.P.
942                       Z.H.P.

                                          story_cleaned  \
0    just a question\nwhat i thought when i first r...
1    pain\nit may be sad but it can be better\nthe ...
2    zombie mayhem\nzombie mayhem\nin the military ...
3    zero wing aybasbtu game script\nzero wing \nal...
4    not prosaic\ndisclaimer i do not claim to own ...
..                                                 ...
938  more\ni have watched this movie times this wee...
939  change comes from within\nas usual i own nothi...
940  changes\nyou know that feeling where people be...
941  undying hero\nzero percent\nvictory cannot be ...
942  unlosing haunted mansion\ndisclaimer i do not ...

                                       story_lemmatized
0    question \n think read b guy think plz leave r...
1    pain \n sad well \n yellow sun glisten face \n...
2    zombie mayhem \n zombie mayhem \n military tes...
3    zero wing aybasbtu game script \n zero wing \n...
4    prosaic \n disclaimer claim zero wing make mon...
..                                                 ...
938  \n watch movie time week love urge write littl...
939  change come \n usual character belong \n \n ex...
940  change \n know feel people believe impacting e...
941  undying hero \n zero percent \n victory gain \...
942  unlose haunt mansion \n disclaimer zettai hero...

[943 rows x 5 columns]
```

```
#Grouping data according to the category
df_grouped =
df[['category','story_lemmatized']].groupby(by='category').agg(lambda
x:' '.join(x))
df_grouped
```

```
story_lemmatized
category

Z Nation                    violent delight \n violent delight \n
desire c...
Z for Zachariah             get save \n hey guy hope like god bless
check ...
Z.H.P.                      undying hero \n zero percent \n victory
gain \...
Zack and Miri Make a Porno  \n watch movie time week love urge write
littl...
Zanna, Don't!               story \n p styletextalign startnot story
try b...
Zapped                      trust \n hello people \n see barely fan
fic th...
Zatch Bell                  healing tear \n sherryyou fight hardplease
for...
Zathura                       dominique jakowec \n test type \n\n end
file \n
Zeke and Luther             wanna skateboard \n helllo know upload new
cha...
Zeke's Pad                  meaning friendship \n write month ago see
fina...
Zen                         informer \n informer \n zen arianna \n
second ...
Zenda                       pain \n sad well \n yellow sun glisten
face \n...
Zenon                         chapter \n flame red soil \n disclaimer
zeno...
Zenonia                       chapter \n yeah yeah update schedule hm
anyw...
Zero Dark Thirty            zero dark thirty alternate footage \n
member s...
Zero Day                    mad world \n late christmas present good
frien...
Zero Wing                   zero wing aybasbtu game script \n zero
wing \n...
Zeta Project                old thing \n old thing \n pass trough door
hea...
Zetman                      poem \n test \n eye \n test story \n party
\n ...
Zettai Karen Children       tea time \n tea time \n set postunlimite
andy ...
```

```
Zevo-3                       memory \n memory \n \n unfortunately
zevo \n j...
Zinda/ज़िंदा          \n planet p \n disclaimer zinda character \n \...
Zodiac P.I.                  christmas shoe \n disclaimer zodiac pi
base st...
Zoey 101                       song \n author note story want song
think go...
Zoey 101, iCarly             cotton swab christmas tree \n \n cotton
swab c...
Zoids                        christmas proposal \n author note wow know
day...
Zom-B                        question \n think read b guy think plz
leave r...
Zombie Fallout               end world end apolla \n end end apolla \n
song...
Zombie Powder                wait \n title wait \n characterspairing
elwood...
Zombie Prom                  scream \n scream \n cry room \n scream bad
doo...
Zombie Survival Guide        everstone \n monday run right scream kevin
run...
Zombie-Loan                  natural phenomenon \n person death natural
phe...
Zombieland                   zombie hunter \n zombie hunter \n season
final...
Zombies, Run!                home \n procrastinate update fanfiction
write ...
Zone of The Enders           leo play hotline miami \n leo play hotline
mia...
Zoo Tycoon                     fanfiction \n late night late usual dark
wet...
Zoolander                    hansel \n derek pout lip hansel think
hanstupi...
Zoom                         nightmare \n anjust quick oneshot \n
disclaime...
Zoombie Blondes              zombie mayhem \n zombie mayhem \n military
tes...
Zoombinis                    put bubble machine inside mountain \n
wonder b...
Zoop                         jacksepticeye \n second day america
birthday d...
Zorro                        unconsciously conscious \n author note
word dr...
```

```
from sklearn.feature_extraction.text import CountVectorizer

#Create a Document term matrix for plotting word clouds
cv=CountVectorizer(analyzer='word')
data=cv.fit_transform(df_grouped['story_lemmatized'])
df_dtm = pd.DataFrame(data.toarray(),
```

```
columns=cv.get_feature_names_out())
df_dtm.index=df_grouped.index
df_dtm.head(10)

                            aa  aaa  aaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaa  \
category

Z Nation                     0    0              0
0
Z for Zachariah              0    0              0
0
Z.H.P.                       0    0              0
0
Zack and Miri Make a Porno   0    0              0
0
Zanna, Don't!                0    0              0
0
Zapped                       0    0              0
0
Zatch Bell                   1    0              0
0
Zathura                      0    0              0
0
Zeke and Luther              1    0              0
0
Zeke's Pad                   0    0              0
0


aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaahhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
h  \
category

Z Nation
0

Z for Zachariah
0

Z.H.P.
0

Zack and Miri Make a Porno
0

Zanna, Don't!
0
```

Zapped
0

Zatch Bell
1

Zathura
0

Zeke and Luther
0

Zeke's Pad
0

aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaahhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
hhhhhhhhhhhhhhhhhhhhhhh \
category

Z Nation
0

Z for Zachariah
0

Z.H.P.
0

Zack and Miri Make a Porno
0

Zanna, Don't!
0

Zapped
0

Zatch Bell
1

Zathura
0

Zeke and Luther

0

Zeke's Pad
0

category	aaaaaaaaaaaaaaaaaaaaahhhhhhhhhhhh	\
Z Nation		0
Z for Zachariah		0
Z.H.P.		0
Zack and Miri Make a Porno		0
Zanna, Don't!		0
Zapped		0
Zatch Bell		0
Zathura		0
Zeke and Luther		0
Zeke's Pad		0

category	aaaaaaaaaaaaanyway	aaaaaaaaahhh	aaaaaaaahhhhhhh	\
Z Nation	0	0	0	
Z for Zachariah	0	0	0	
Z.H.P.	0	0	0	
Zack and Miri Make a Porno	0	0	0	
Zanna, Don't!	0	0	0	
Zapped	0	0	0	
Zatch Bell	0	0	0	
Zathura	0	0	0	
Zeke and Luther	0	0	0	
Zeke's Pad	0	0	0	

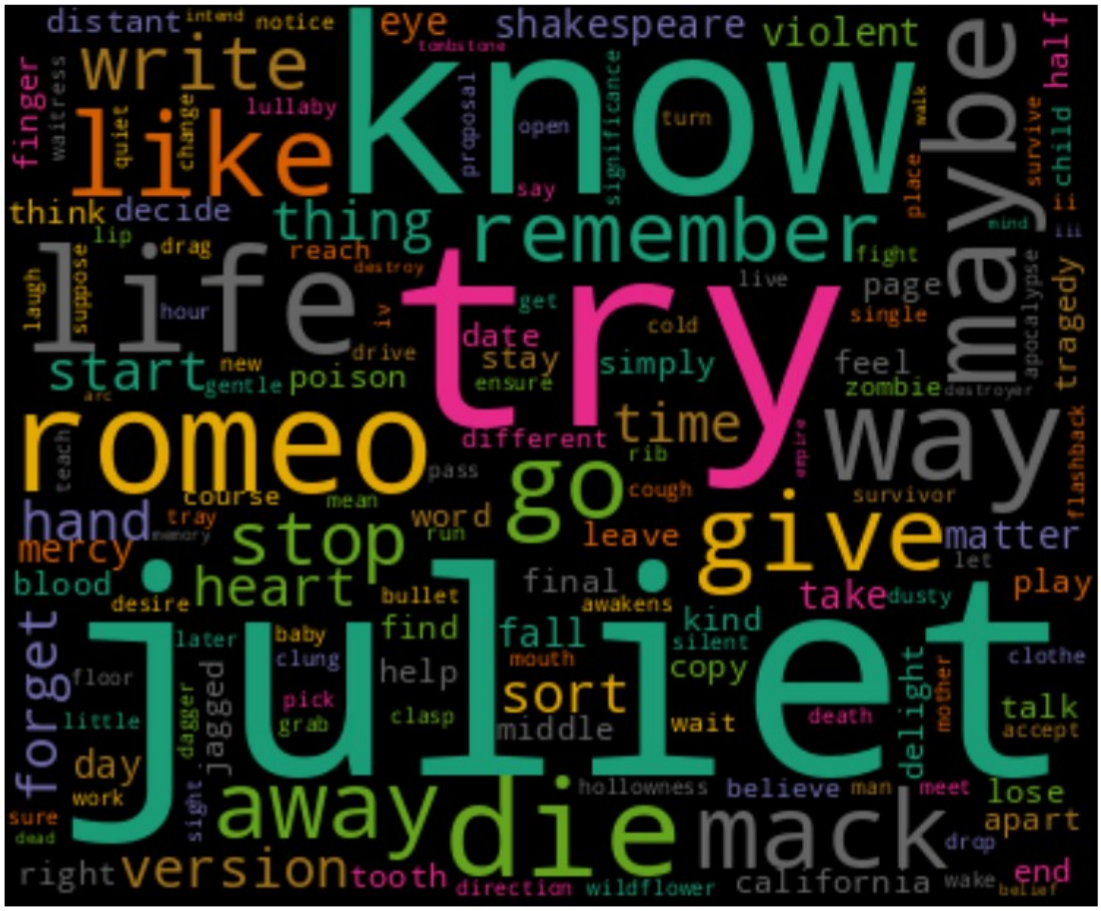
category	...	zzz	zzzaker	zzzoa	zzzoid	zzzzzzzzzzz	\
Z Nation	...	0	0	0	0		

```
0
Z for Zachariah             ...    0        0      0       0
0
Z.H.P.                      ...    0        0      0       0
0
Zack and Miri Make a Porno  ...    0        0      0       0
0
Zanna, Don't!               ...    0        0      0       0
0
Zapped                      ...    0        0      0       0
0
Zatch Bell                  ...    0        1      0       0
0
Zathura                     ...    0        0      0       0
0
Zeke and Luther             ...    0        0      0       0
0
Zeke's Pad                  ...    0        0      0       0
0

                          zzzzzzzzzzzzz  zzzzzzzzzzzzzz  \
category
Z Nation                              0               0
Z for Zachariah                       0               0
Z.H.P.                                0               0
Zack and Miri Make a Porno            0               0
Zanna, Don't!                         0               0
Zapped                                0               0
Zatch Bell                            0               0
Zathura                               0               0
Zeke and Luther                       0               0
Zeke's Pad                            0               0

                          zzzzzzzzzzzzzzzz  zzzzzzzzzzzzzzzzz  \
category
Z Nation                                 0                   0
Z for Zachariah                          0                   0
Z.H.P.                                   0                   0
Zack and Miri Make a Porno               0                   0
Zanna, Don't!                            0                   0
Zapped                                   0                   0
Zatch Bell                               0                   0
Zathura                                  0                   0
Zeke and Luther                          0                   0
Zeke's Pad                               0                   0


zzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzzz
category
```

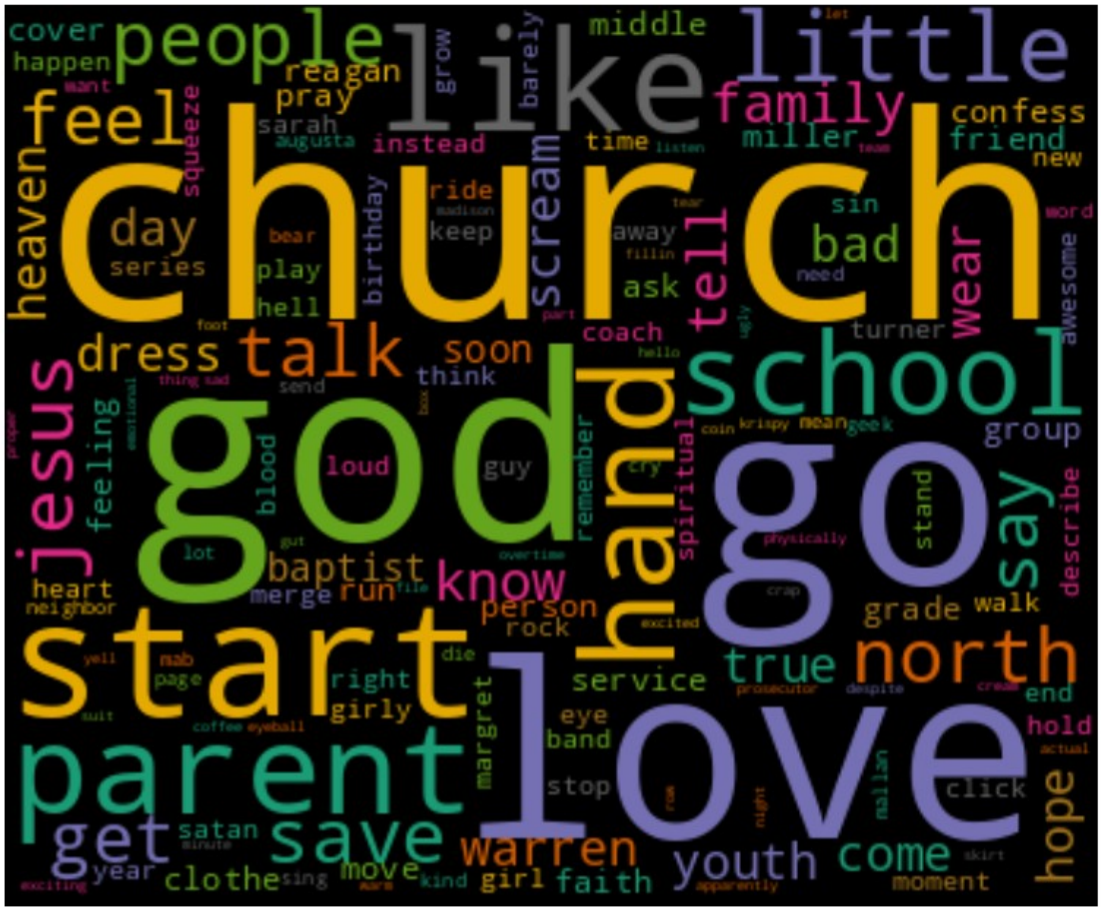
```
Z Nation
0
Z for Zachariah
0
Z.H.P.
0
Zack and Miri Make a Porno
0
Zanna, Don't!
0
Zapped
0
Zatch Bell
0
Zathura
0
Zeke and Luther
0
Zeke's Pad
0

[10 rows x 34485 columns]
```

```
# Importing wordcloud for plotting word clouds and textwrap for
wrapping longer text
from wordcloud import WordCloud
from textwrap import wrap

# Main method for generating word clouds (freq of different words in a
doc)
def generate_wordcloud(data,title):
  wc = WordCloud(width=400, height=330,
max_words=150,colormap="Dark2").generate_from_frequencies(data)
  plt.figure(figsize=(10,8))
  plt.imshow(wc, interpolation='bilinear')
  plt.axis("off")
  plt.title('\n'.join(wrap(title,60)),fontsize=13)
  plt.show()

# Transposing document term matrix
df_dtm=df_dtm.transpose()
# Plotting word cloud for each product
for index,product in enumerate(df_dtm.columns):

generate_wordcloud(df_dtm[product].sort_values(ascending=False),produc
t)
```

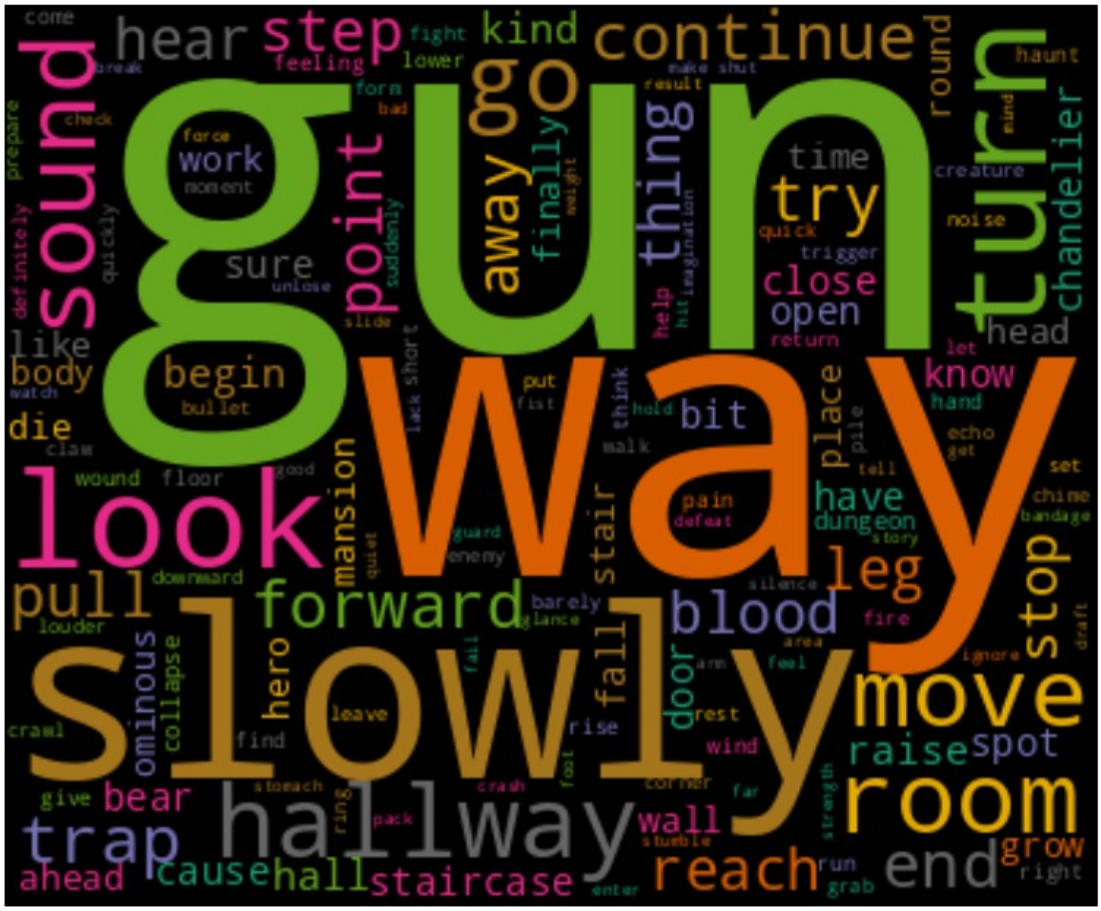
Z Nation


know
juliet
romeo
try
life
way
maybe
like
write
remember
thing
time
go
give
stop
heart
away
die
mack
version
forget
hand
start
poison
mercy
blood
sort
help
find
fall
kind
take
play
matter
talk
lose
apart
end
california
tooth
direction
wildflower
wake
drop
delight
believe
hollowness
man
meet
death
wait
middle
day
work
sure
dead
right
little
floor
pick
grab
clasp
jagged
sight
dagger
baby
clung
mouth
silent
copy
final
awakens
bullet
desire
leave
word
run
course
mean
pass
different
rib
cough
survivor
zombie
feel
simply
stay
date
drive
ensure
gentle
new
arc
hour
lv
get
cold
live
page
single
fight
reach
destroy
drag
lip
suppose
laugh
think
decide
finger
waitress
quiet
change
lullaby
open
turn
proposal
significance
say
eye
tombstone
shakespeare
violent
distant
intend
notice
place
walk
survive
mind
apocalypse
tragedy
child
half
flashback
destroyer
clothe
mother
accept
lullaby
tray
memory
teach
let
later
tragedy

Z for Zachariah



Z.H.P.


gun
way
slowly
look
sound
turn
hear
step
continue
kind
go
thing
try
time
work
point
away
finally
sure
close
open
head
chandelier
round
begin
bullet
like
body
die
know
bit
place
pull
forward
mansion
stair
have
dungeon
leg
blood
stop
move
room
hallway
trap
reach
end
staircase
hall
cause
ahead
bear
wall
fall
door
hero
ominous
collapse
leave
find
raise
spot
grow
wound
floor
pain
defeat
fire
silence
barely
glance
louder
downward
crawl
give
stomach
crash
corner
rest
wind
rise
foot
fall
arm
feel
area
ignore
draft
set
chime
bandage
haunt
creature
noise
quick
trigger
return
let
hand
echo
get
tell
pile
hold
think
walk
put
fist
lack
short
slide
suddenly
unlose
moment
force
check
prepare
definitely
quickly
break
feeling
fight
lower
form
bad
result
make
shut
mind
enemy
guard
quiet
good
watch
claw
come

Zack and Miri Make a Porno

Zanna, Don't!

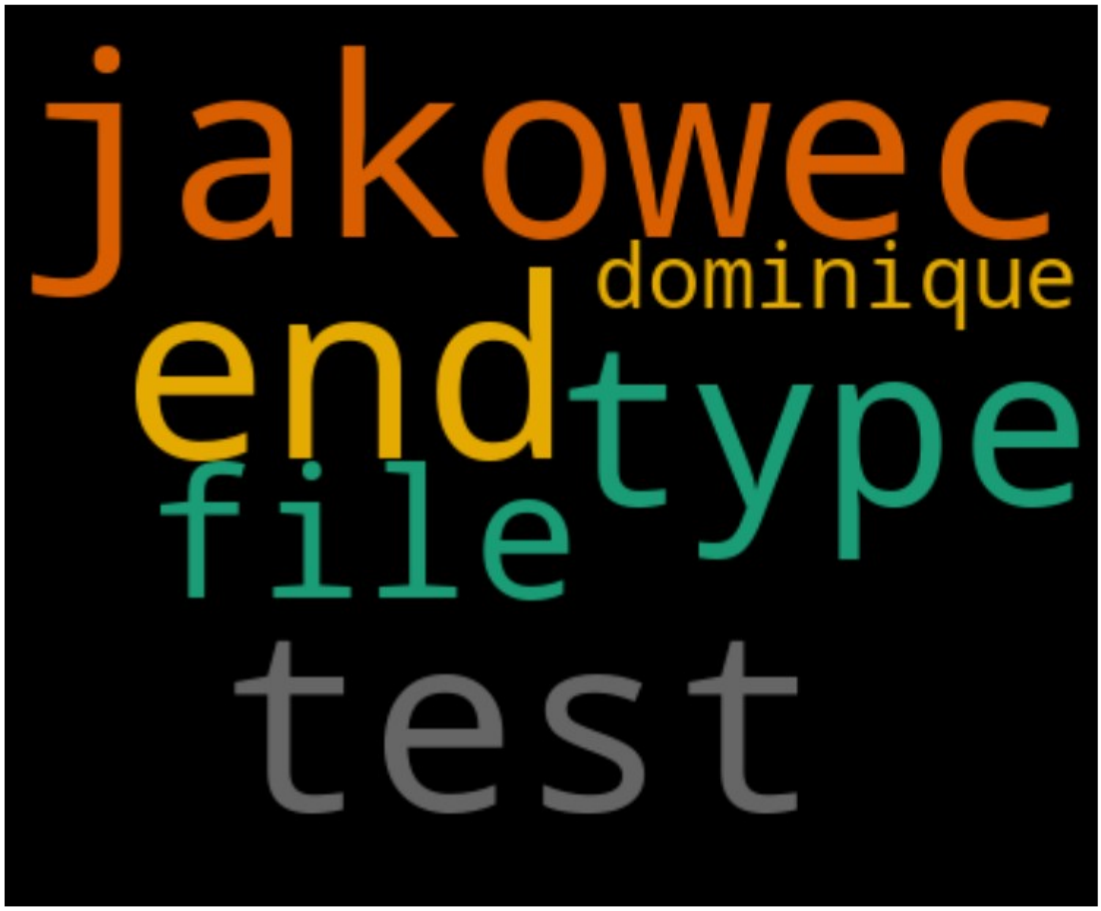
Zapped


jackson
zoey
look
know
want
reply
adam
walk
nod
think
kiss
like
chuckle
smile
talk
blush
away
taylor
love
date
sorry
shocked
guy
laugh
phone
shoulder
review
shake
mean
leave
well
see
family
add
say
time
turn
ask
bit
eye
movie
pack
thing
boy
hand
girl
hope
head
happen
get
app
go
take
start
tell
okay
wait
good
end
honest
oh
stop
not
state
fix
raise
scene
walk
hey
let
shot
horrible
brother
enjoy
happy
kinda
college
grin
promise
try
trust
cry
actually
file
help
hung
work
course
yes
neck
deep
disbelief
pov
break
ben
glad
great
mistake
realize
competition
moment
house
fun
begin
manipulate
trouble
funny
boyfriend
finish
pant
drabble
jackett
thank
believe
finally
roll
hurt
weird
zoeyjackson
confused
exclaim
anymore
lesson
real
feel
forget
match
use
zap
stick
share
summary
rest
dog
post
act
eyebrow
loose
stand
skill
section
sound
god
story
command
wide
place
sigh
learn
right
state
file
actually
hung
work
course
yes
neck
deep
disbelief

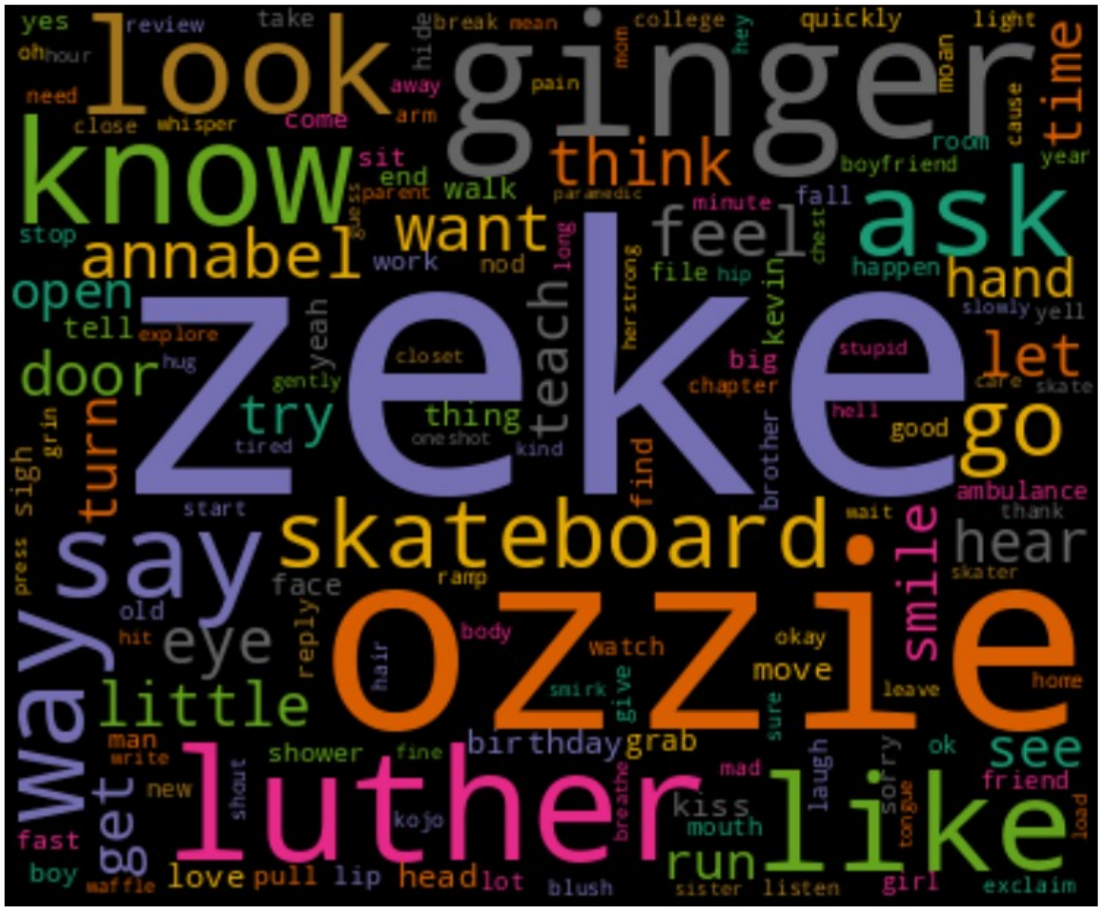
Zatch Bell


zatch
look
say
ask
kiyo
tia
megumi
know
think
eye
hand
destroy
close
start
turn
dinemo
spell
okay
tell
friend
away
let
willy
happy
hori
call
shout
fight
grab
wait
sit
face
help
open
give
love
battle
happen
yell
arm
big
hard
fire
folgore
giant
laugh
kiss
good
stand
lee
reply
evil
monrei
lose
light
thank
girl
get
feel
try
come
day
run
see
hold
new
begin
end
mamodo
sherry
speak
hair
kanchome
finally
human
child
red
body
pull
long
time
hear
thing
dark
brother
bed
little
room
door
reece
zeno
stop
take
wear
place
rest
walk
kolulu
chocolate
continue
ground
chapter
book
leave
kiko
oh
world
way
brago
koko
king
plush
want
like
go
fall
graviray
find
zofis
stare
smile
head
wonka
right
mean
not
mouth
yes
sleep
sure
shoot
soon
blaze
sorry
answer
make
well
hey
talk
voice
guy
heart
black
match
need
respond

Zathura



Zeke and Luther



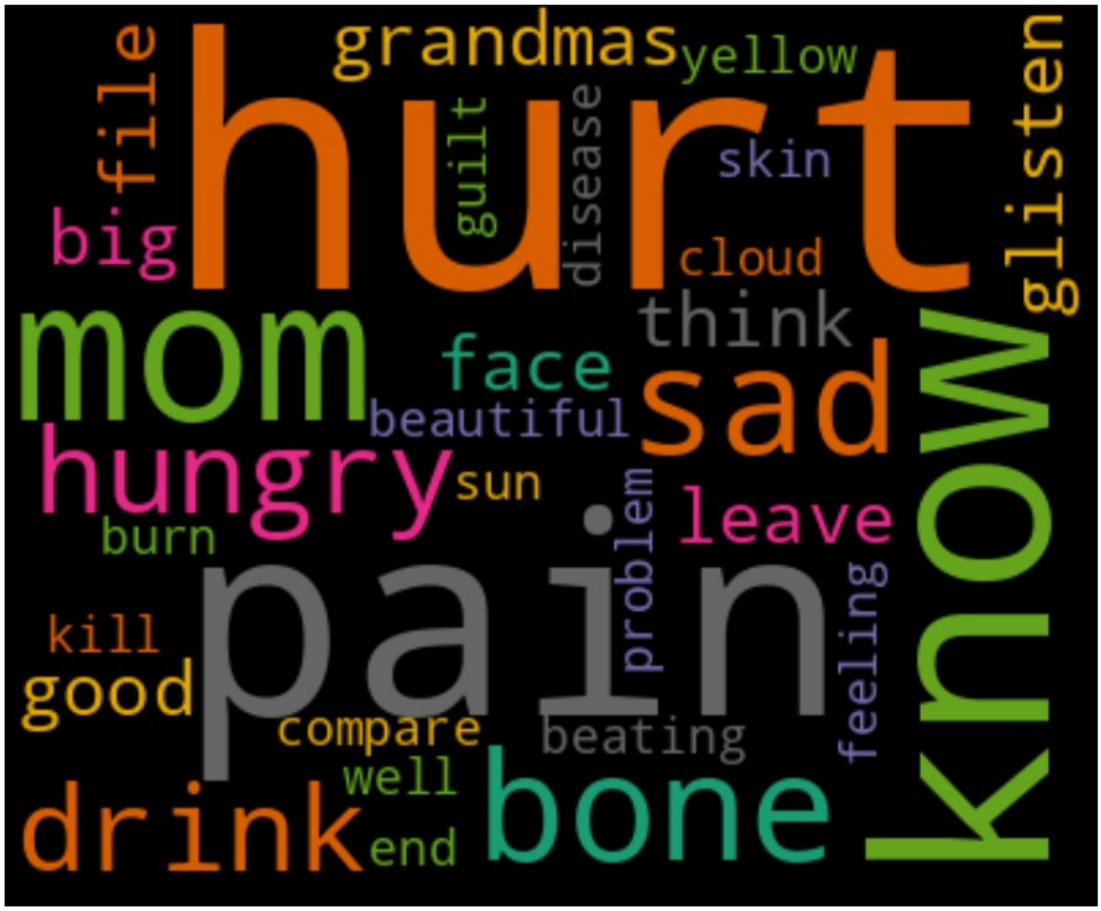
Zeke's Pad


jay
zeke
life
friendship
phone
way
like
want
pull
night
actually
simply
intuitively
rush
fix
sure
lately
mistake
sleep
far
enjoy
feel
live
possible
boredom
remember
section
need
matter
see
rely
type
friendless
away
month
file
write
pick
end
read
anymore
blast
meet
time
ear
suddenly
pillow
ring
tell
book
meaning
upload
problem
unlikely
help
have
lonelyness
collide
friend
careless
ago
get
somebody
finally
middle
difficulty
sleeping

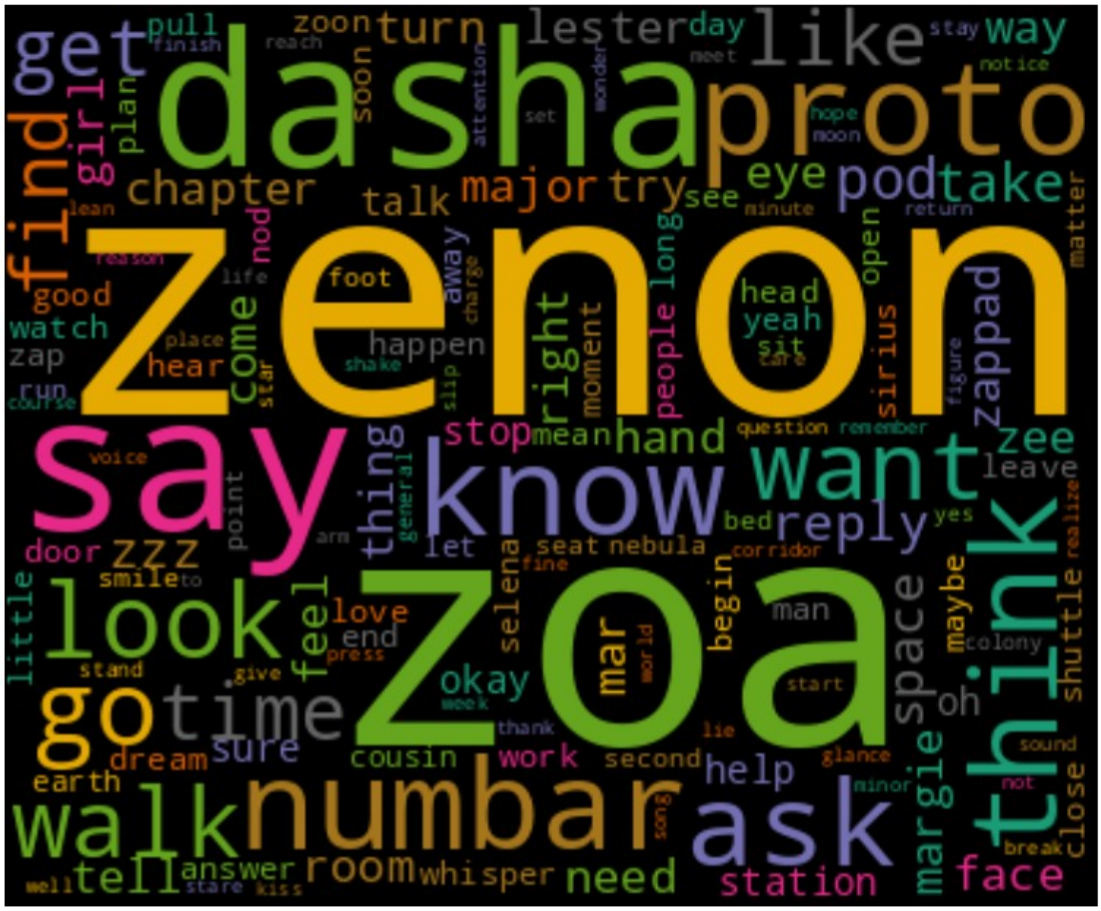
Zen


know
say
look
think
zen
go
tell
ruspanti
yes
hand
smile
want
love
little
aurelio
cigarette
ask
night
leave
turn
like
arianna
take
feel
work
coffee
stop
watch
lupi
pull
good
light
eye
hair
door
see
open
wait
time
money
get
come
man
meet
hold
bed
mouth
hear
need
face
away
long
car
room
talk
give
place
home
office
case
laugh
word
mind
lover
nice
sex
end
ring
sit
finger
oh
inspector
voice
help
story
play
bridge
try
push
cabal
apartment
perotto
outside
shake
start
early
step
table
husband
sure
business
stand
5
drop
day
rest
stay
maybe
evening
drive
murder
dark
not
live
big
slowly
phone
miss
mean
dress
afm
find
breast
walk
high
lip
kiss
catch
head
ludo
floor
god
thing
slide
way
pick
luca
flat
hueber
key
box
lean
till
minister
beautiful
tania
friend
lie
body
touch
green
inside
well
cardinal
client
morning
move

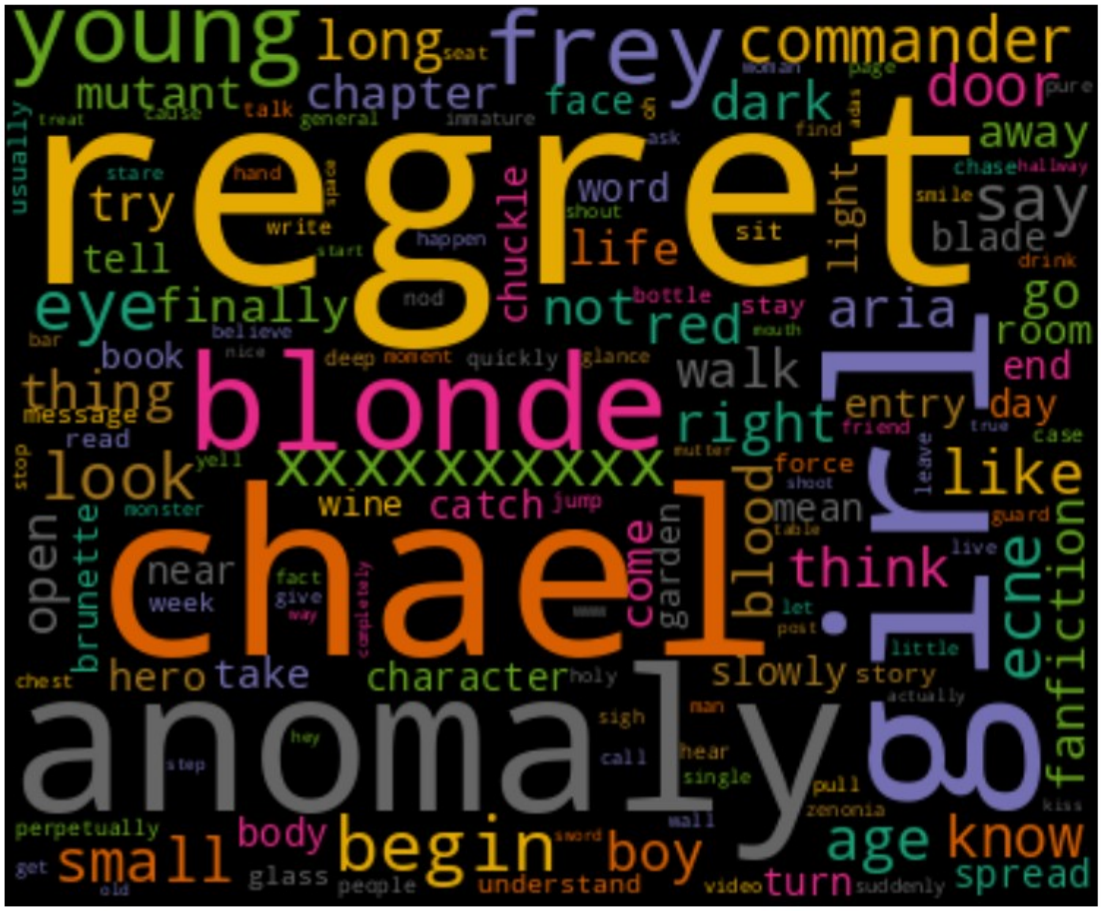
Zenda


hurt
grandmas
yellow
file
guilt
disease
skin
big
cloud
glisten
mom
think
face
sad
know
beautiful
hungry
sun
problem
leave
burn
pain
feeling
kill
good
compare
beating
well
drink
bone
end

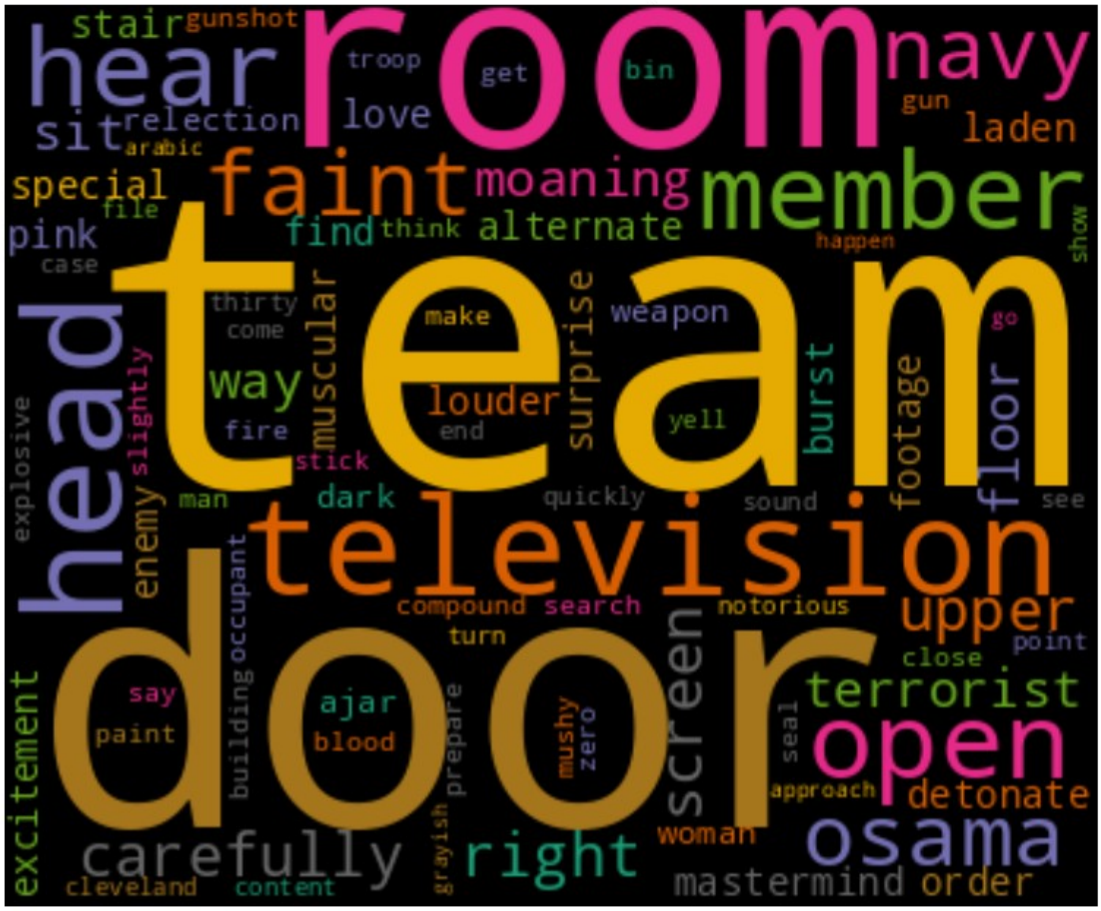
Zenon


zenon
dasha
proto
zoa
say
know
want
think
look
go
time
walk
numbar
ask
find
get
like
turn
lester
way
pull
zoon
girl
plan
soon
chapter
talk
major
try
eye
pod
take
good
watch
come
hear
happen
right
people
long
head
yeah
sirius
open
zappad
zee
stop
mean
hand
reply
leave
thing
door
zzz
bed
space
maybe
shuttle
close
face
station
margie
help
work
second
need
room
whisper
tell
answer
sure
dream
cousin
earth
okay
feel
love
end
mar
begin
man
oh
little
smile
selena
seat
nebula
let
stand
give
press
thank
lie
start
glance
minor
not
sound
break
realize
matter
foot
away
nod
life
reason
lean
minute
return
hope
moon
stay
notice
meet
wonder
attention
set
reach
finish
yes
corridor
colony
figure
remember
question
sit
care
shake
slip
star
course
run
zap
place
voice
point
general
arm
fine
song
world
meek
thank
well
stare
kiss

Zenonia


young
long
frey
commander
mutant
chapter
face
dark
door
away
regret
try
word
say
tell
life
light
blade
eye
finally
not
red
aria
go
book
walk
room
thing
blonde
girl
end
entry
day
look
right
like
xxxxxxxxxx
wine
catch
blood
mean
open
chael
come
garden
near
think
ecne
fanfiction
brunette
hero
take
character
slowly
story
anomaly
boy
age
know
small
body
begin
boy
spread
glass
people
understand
turn
suddenly
video
perpetually

Zero Dark Thirty



Zero Day

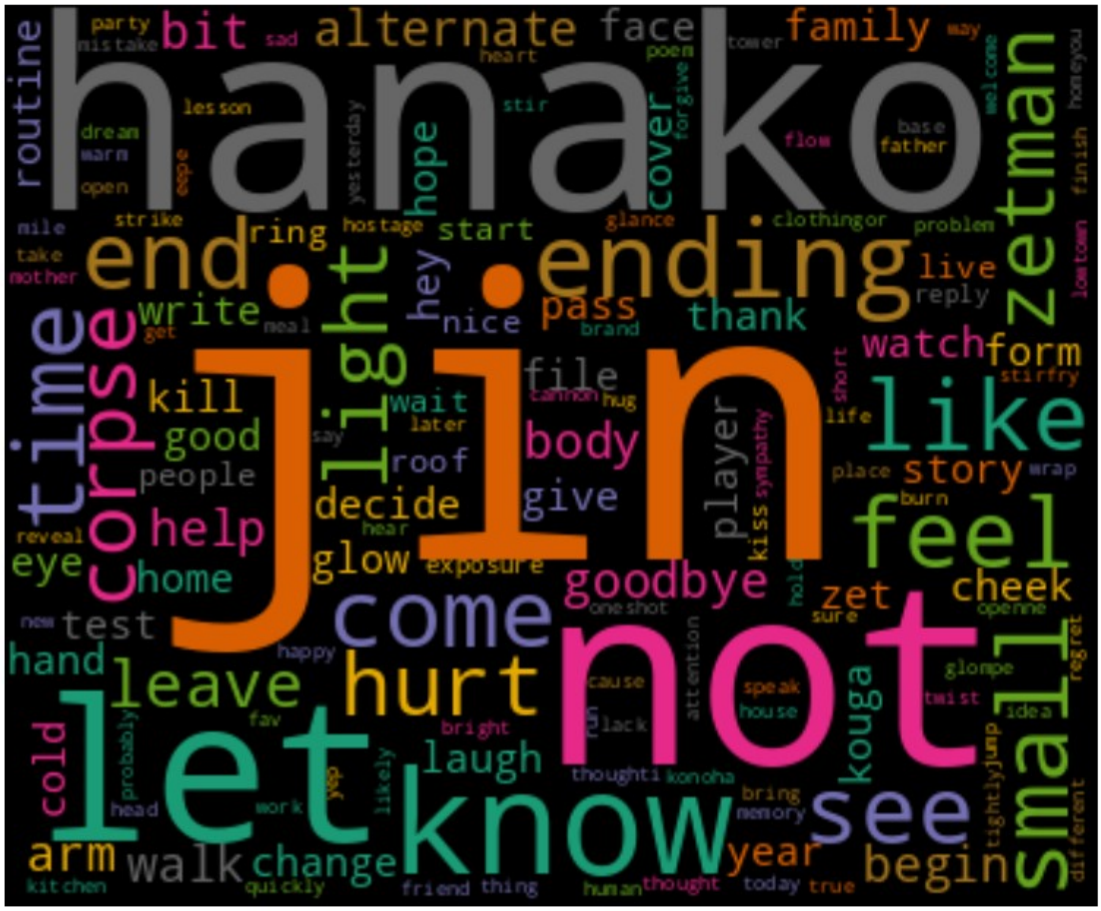

look
walk
call
kill
say
gun
girl
boy
voice
time
floor
table
leave
pick
blood
scream
run
bullet
know
try
stand
pull
hear
shoot
sound
start
little
phone
chair
ask
eye
room
way
thing
face
turn
yell
stop
sense
die
desk
sob
shut
kill
go
find
make
pick
hold
come
end
hand
past
get
shriek
place
damn
echo
laugh
victim
away
point
rifle
answer
couch
second
wait
stupid
smile
let
friend
quiet
collapse
magazine
man
study
grin
people
push
power
shake
oh
andre
add
want
long
see
pack
grow
feel
fucking
fall
anger
jump
dead
right
far
world
woman
death
cop
swing
finally
scream
whimper
life
alive
drop
word
idea
greg
help
school
hope
repeat
hey
tell
happen
glass
hide
sure
fire
red
good
kick
set
laughter
trigger
outside
draw
fool
book
silent
hunt
maybe
well
belt
probably
god
take
understand
continue
library
strength
hall
air
agree
cold

Zero Wing

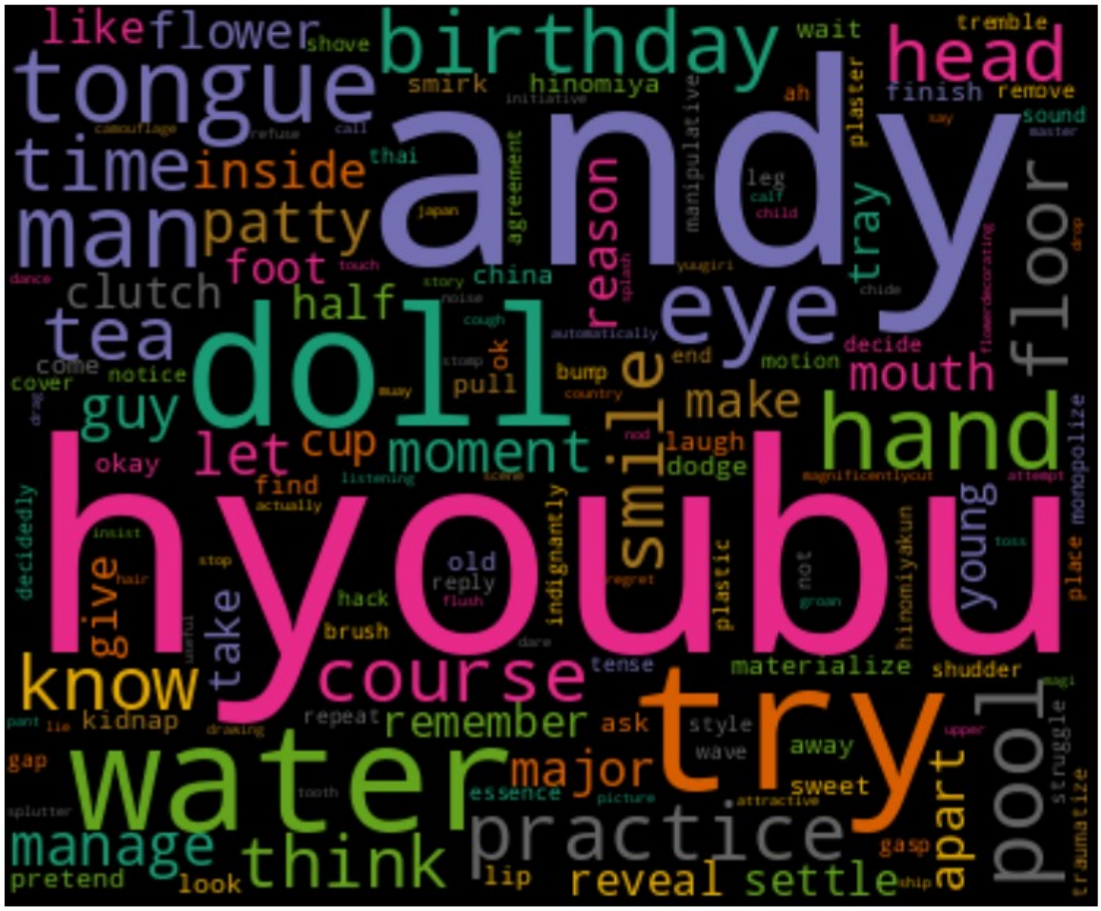

dog
ship
captain
ha
zag
game
base
cat
end
stage
like
face
start
time
crew
wing
beat
enemy
soul
life
zero
screen
final
great
zig
belong
know
die
defeat
strike
fighter
story
unfairness
incorrect
file
undo
piece
justice
bomb
remain
basis
chance
fill
anguish
word
good
go
chaps
embrace
survive
silent
force
deep
swift
death
tell
happen
star
strong
route
voice
vessel
victory
try
cry
make
gentleman
defiance
blaze
human
fall
less
calm
friend
rule
set
figure
operator
freeze
come
mortal
eye
war
napalm
evil
watch
reach
emotion
fear
dark
take
courage
mind
play
attack
main
destroy
space
black
mean
god
hear
bring
way
place
begin
sense
cold
destruction
disclaimer
bad
laughter
weekly
grip
bright
decide
stomach
ad
distraught
whisper
frustration
mercy
turn
russet
turnon
man
aybasbtu
disbelief
brief
silence
run
existence
correct
tear
beautiful
instant
natolsii
breathe
medium
close
narration
exclamation
mechanic
form
fuel
distant
team

Zeta Project

Zetman


hanako
jin
ending
let
not
know
small
like
feel
zetman
time
corpse
light
end
come
hurt
leave
see
begin
kouga
zet
goodbye
body
player
give
decide
help
home
eye
glow
kill
good
people
write
pass
thank
watch
form
story
cheek
hand
arm
walk
change
laugh
cold
year
test
bit
alternate
face
family
routine
hope
cover
start
ring
hey
nice
file
wait
roof
exposure
oneshot
sympathy
kiss
hug
cannon
later
say
short
life
place
burn
wrap
stirfry
reply
live
lowtown
finish
honeyou
welcome
way
tower
poem
heart
sad
party
mistake
lesson
stir
forgive
yesterday
flow
base
father
dream
warm
open
strike
hostage
glance
clothingor
problem
mile
take
mother
get
meal
brand
openne
regret
glompe
twist
idea
different
tightlyjump
human
thought
today
true
thoughti
konoha
bring
memory
friend
thing
quickly
kitchen
work
head
probably
likely
yep
fav
bright
black
cause
speak
house
attention
hold
sure
now
happy
reveal
hear

Zettai Karen Children

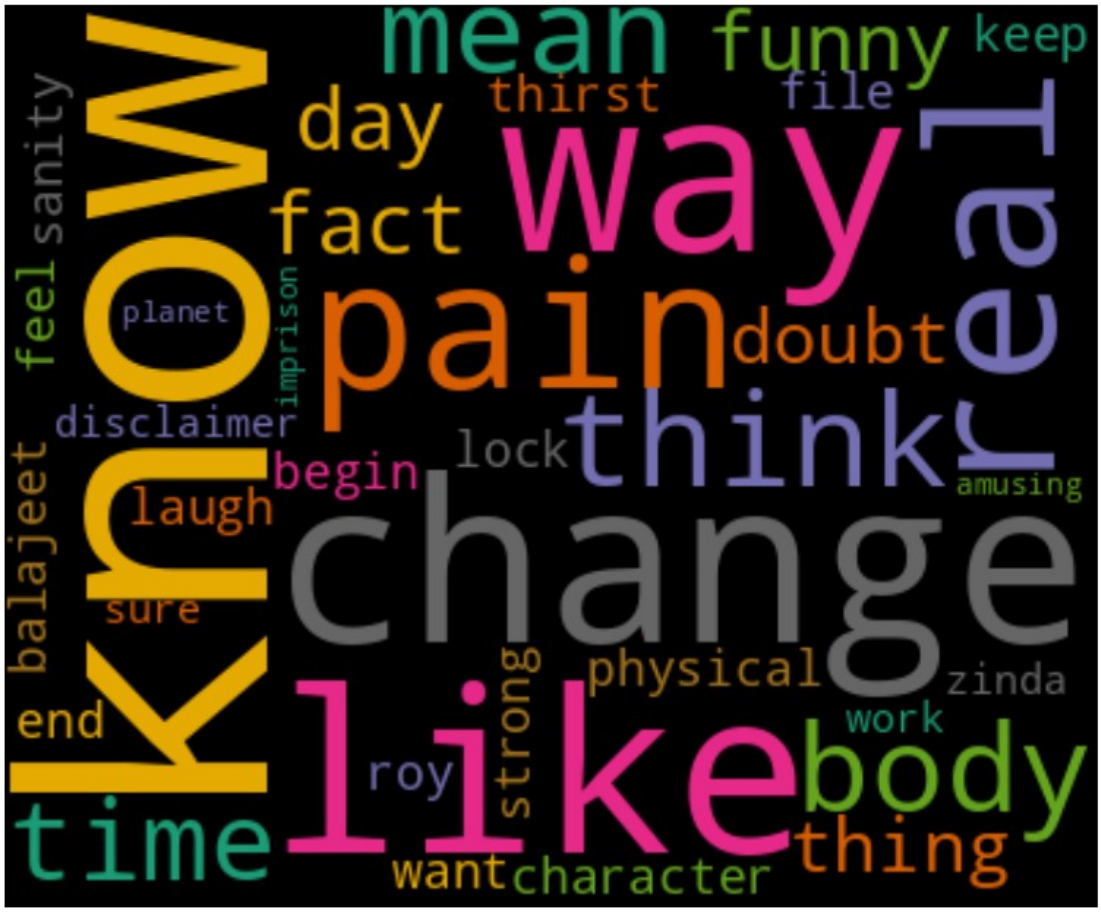


Zevo-3


```
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2332 (\N{DEVANAGARI LETTER JA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Matplotlib currently does not support Devanagari
natively.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2364 (\N{DEVANAGARI SIGN NUKTA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2306 (\N{DEVANAGARI SIGN ANUSVARA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
UserWarning: Glyph 2342 (\N{DEVANAGARI LETTER DA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
/usr/local/lib/python3.9/dist-packages/IPython/core/pylabtools.py:151:
```

```
UserWarning: Glyph 2366 (\N{DEVANAGARI VOWEL SIGN AA}) missing from
current font.
  fig.canvas.print_figure(bytes_io, **kw)
```

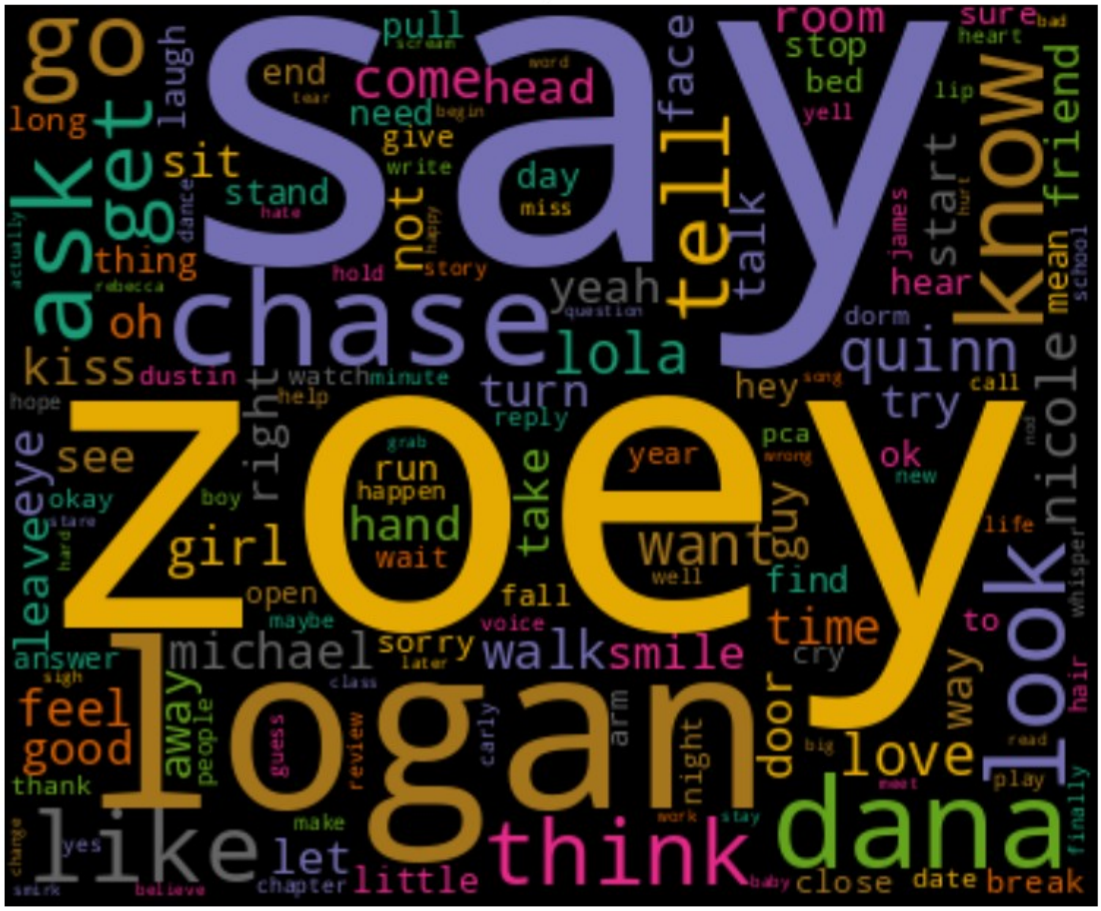
Zinda/□□□□□



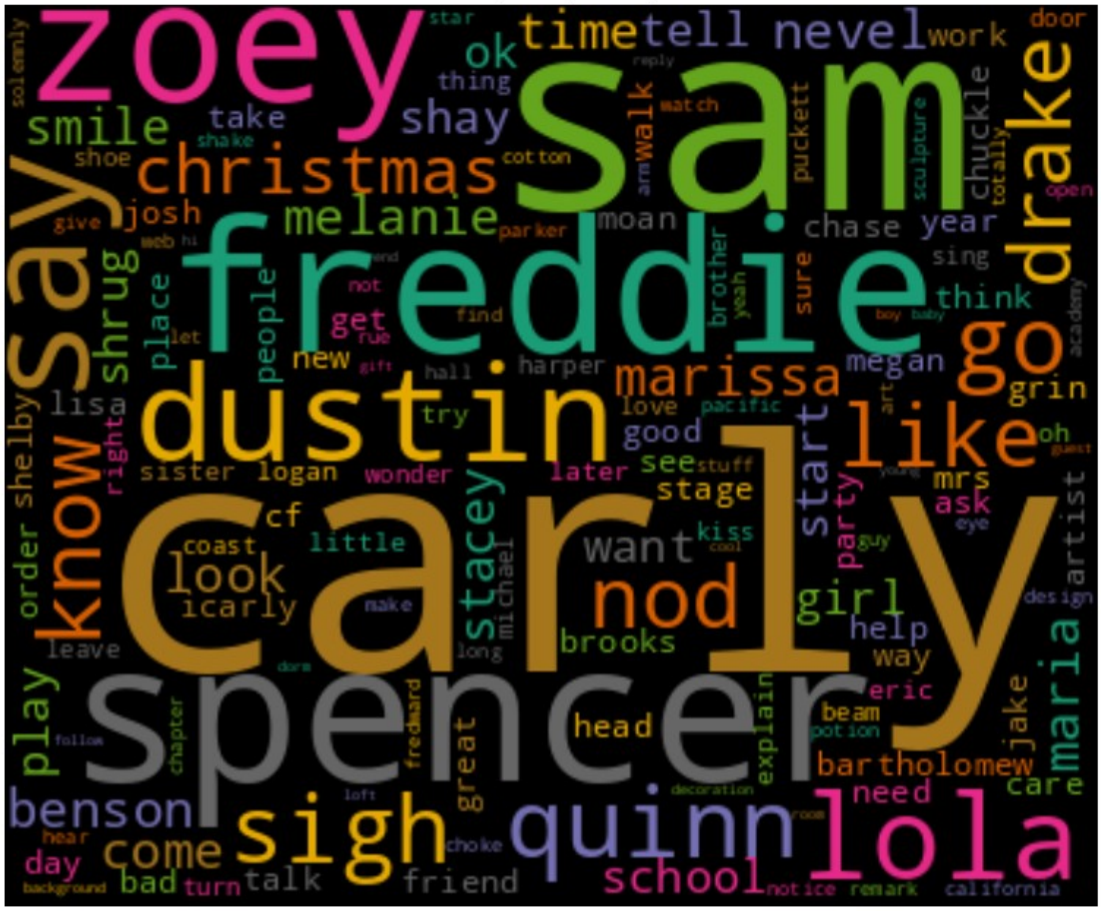
Zodiac P.I.


look
shoe
turn
run
hiro
kid
michelle
pizza
christmas
sarah
police
bloody
boy
like
hand
face
money
momma
end
walk
home
say
coin
wave
man
ask
quickly
finally
couple
little
beautiful
hurry
try
cashier
pocket
door
need
open
people
see
want
wrap
smile
stare
file
go
sir
get
stand
give
pull
mall
buy
shoot
dumb
scream
shopping
place
found
time
mama
blood
sick
spell
sight
close
clothe
review
soz
gladly
explain
fidget
pay
bloddy
gentle
grin
newsong
kitchen
lily
size
follow
watch
great
count
bath
hoodthen
opend
thn
pipethey
shoppingnichelle
raneveryone
hum
feeling
spare
dress
sure
stop
hot
bill
fun
busy
day
minute
grab
pizze
wallet
touvhe
hastily
practically
start
stude
use
take
coat
pizzathe
daddy
hood
person
feel
snap
jesus
hear
tonight
later
exit
rush
forst
lately
disappear
inside
maybe
pair
believe
break
cafewhen
register
counter
disbelief
spend
excited
wait
ps
eye
attention
pi
and
sister

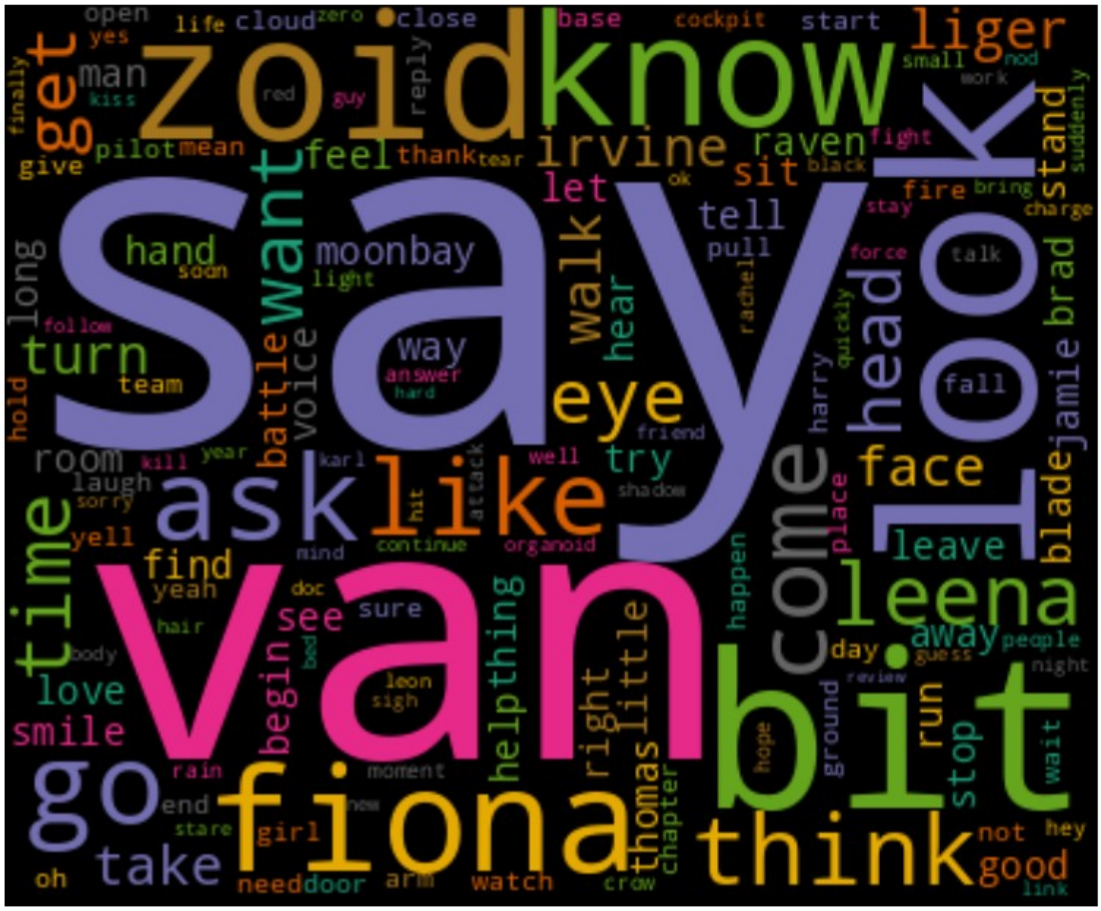
Zoey 101


say
zoey
logan
chase
dana
go
get
ask
know
tell
look
like
think
love
room
pull
come
head
face
stop
bed
sure
heart
long
laugh
end
need
give
write
sit
stand
day
miss
not
talk
start
hear
friend
mean
thing
oh
yeah
lola
quinn
nicole
kiss
dustin
watch
minute
turn
hey
try
call
dorm
hope
help
reply
see
eye
leave
right
run
happen
hand
wait
take
year
pca
ok
girl
want
guy
find
okay
boy
open
fall
maybe
voice
time
michael
walk
smile
sorry
later
cry
to
way
answer
sigh
feel
good
away
people
thank
guess
review
carly
arm
night
door
hair
make
yes
let
little
chapter
close
date
break
finally
play
meet
work
stay
baby
believe
smirk
change
life
whisper
read
mad
wrong
new
song
well
class

Zoey 101, iCarly



Zoids

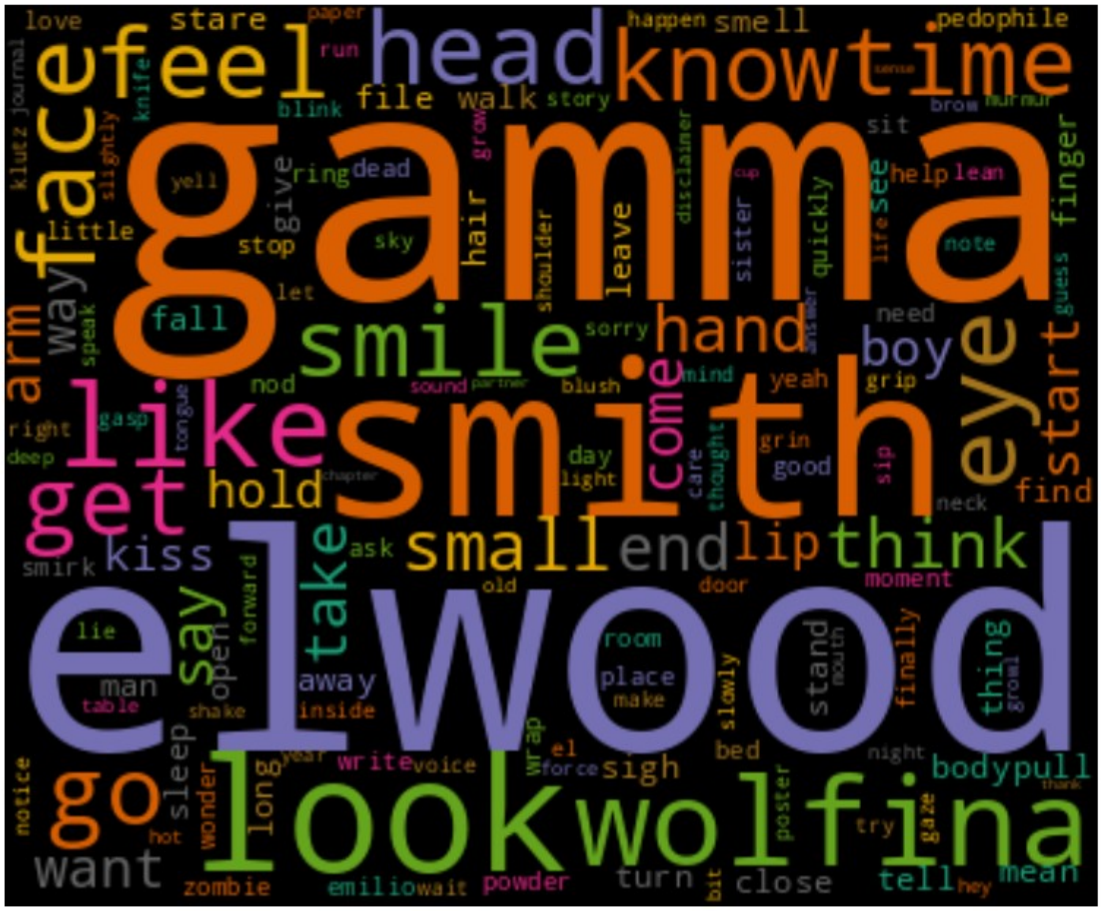

say
van
bit
zoid
know
look
liger
irvine
raven
want
feel
walk
eye
head
come
face
leena
ask
like
time
go
fiona
think
take
turn
hand
moonbay
way
tell
let
sit
stand
blade
jamie
brad
leave
away
find
see
love
smile
begin
helpthing
right
little
thomas
stop
run
good
not
get
give
pilot
mean
thank
tear
fight
fire
bring
charge
stay
force
talk
fall
pull
rachel
harry
quickly
suddenly
long
follow
hold
team
battle
voice
answer
hear
try
well
friend
shadow
room
laugh
sorry
yell
kill
year
karl
attack
hit
mind
continue
organoid
place
happen
yeah
doc
hair
body
bed
leon
sigh
sure
day
people
guess
night
review
ground
hope
wait
hey
rain
moment
end
stare
girl
chapter
oh
need
door
arm
watch
crow
link
open
yes
life
cloud
zero
close
base
cockpit
start
small
nod
work
kiss
red
guy
reply
finally
ok
black
soon
light
hard
now

Zom-B


book
love
file
end
fanfiction
thank
series
question
guy
plz
add
think
actual
start
read
review
people
leave
fanfic

Zombie Fallout

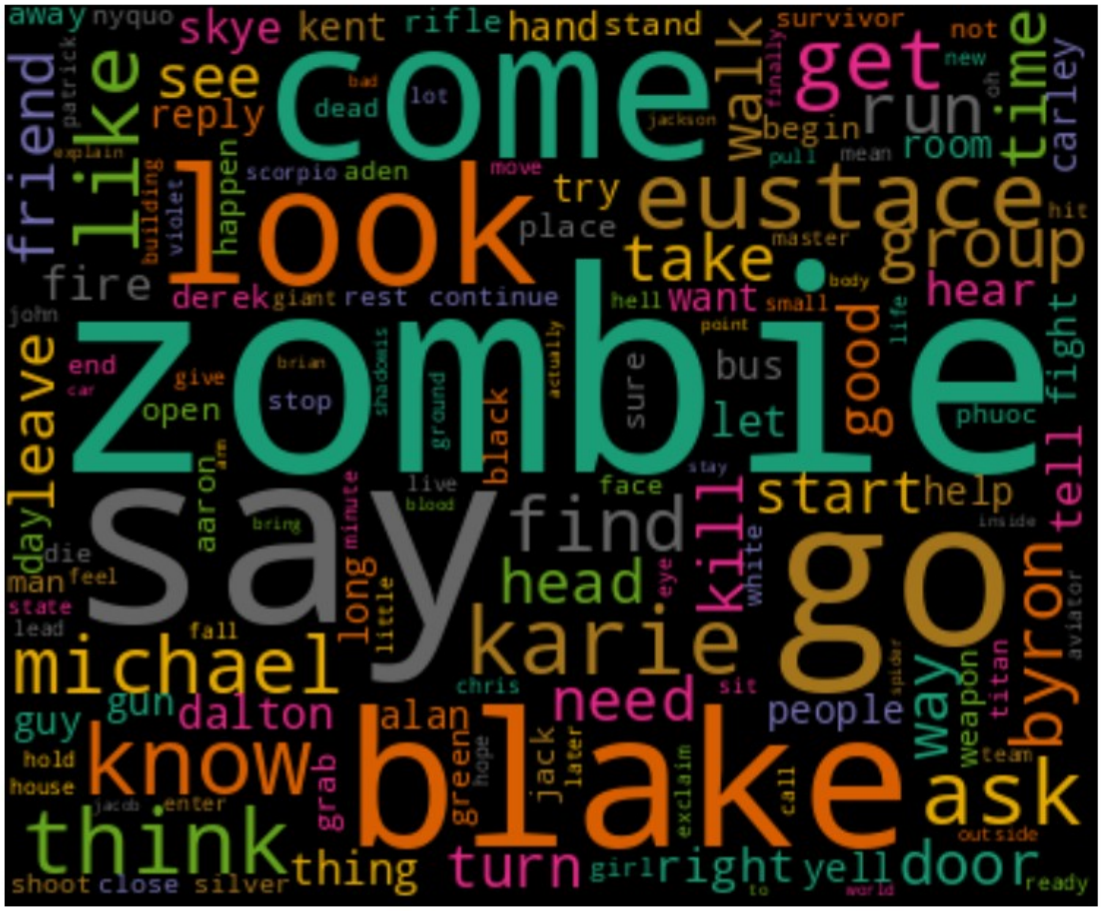
Zombie Powder



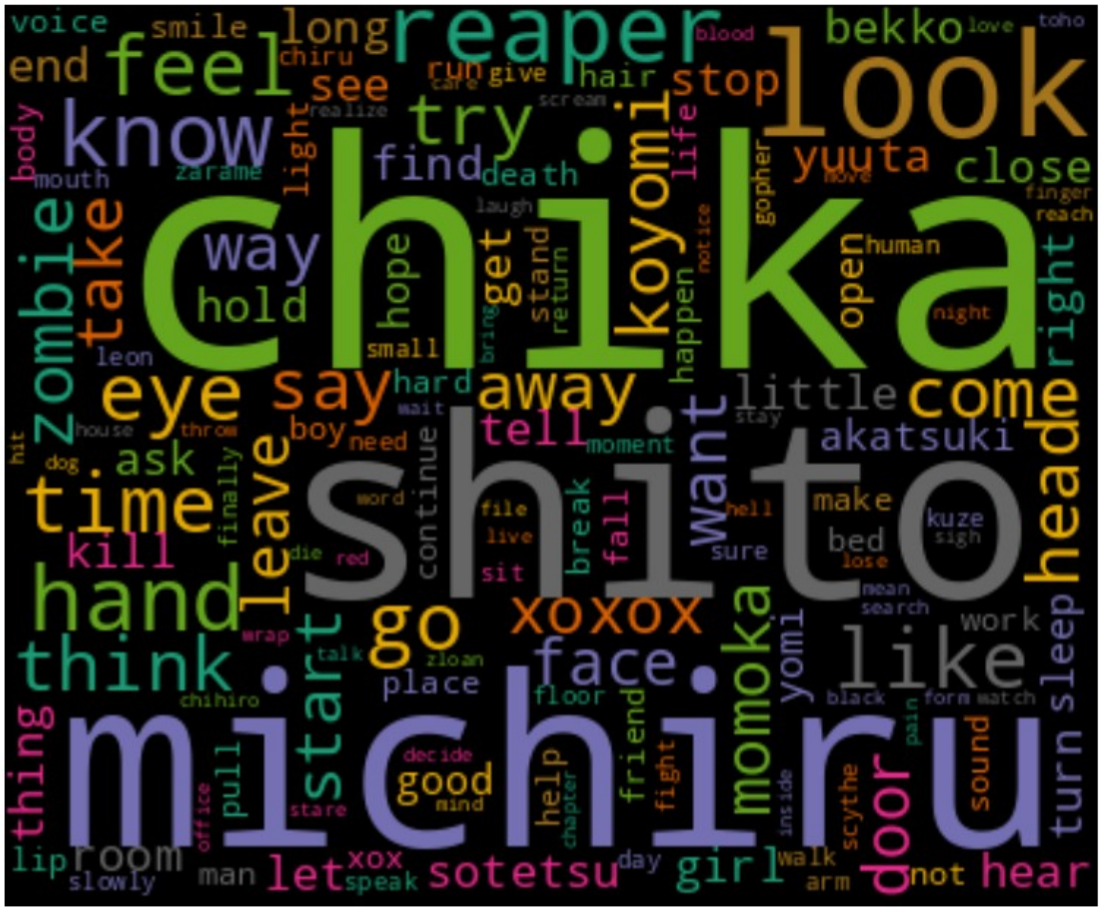
Zombie Prom


zelogo
scream
aaaaah
know
way
miss
knife
want
like
look
tell
pain
feel
need
end
leave
surgery
cause
say
ghost
come
hold
love
ass
blood
hair
hurt
room
schedule
vintage
get
promise
girl
fuck
stick
mind
open
talk
lay
file
wake
walk
start
stand
reply
kill
stay
mean
daehyun
junhong
sob
tell
deep
fast
wear
white
bleed
black
slowly
tonight
wait
hospital
thisi
happend
shame
follow
deserve
camera
cry
brownhaire
waitingthe
protect
rest
thing
shock
silent
smile
girlfriend
tomorrow
happen
upi
herei
exactly
honest
yonggguk
soyou
kid
amake
minutei
untill
downstais
later
night
morei
againi
practice
happendi
accidentthey
together
hearthyungi
chaw
surgeon
tonights
dae
zelohishe
thatthe
sure
jagiyayou
waitingthe
dose
suddenly
create
let
space
ll
music
unconscious
condition
wrist
jean
life
mix
voice
happily
type
apart
explain
person
long
bringht
tear
high
huge
character
take
stop
time
hell
hour
think
sit
junhong

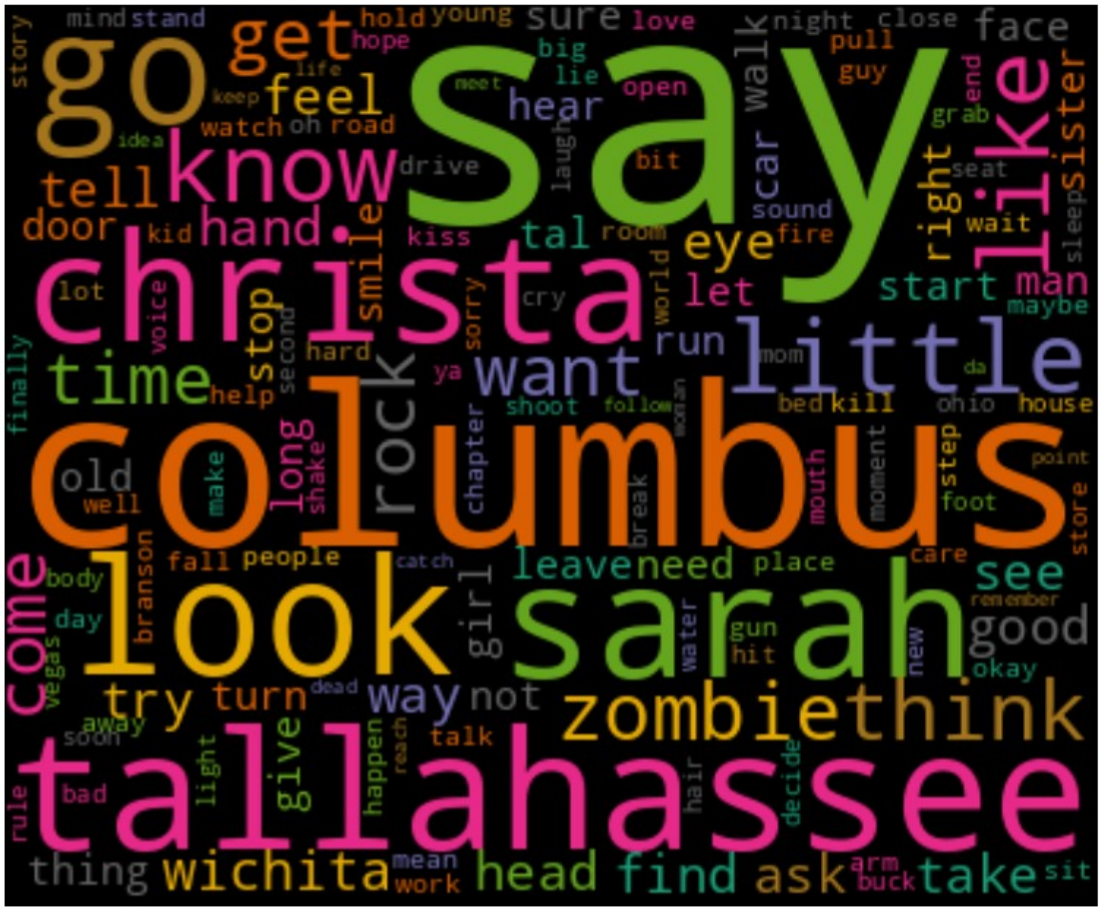
Zombie Survival Guide



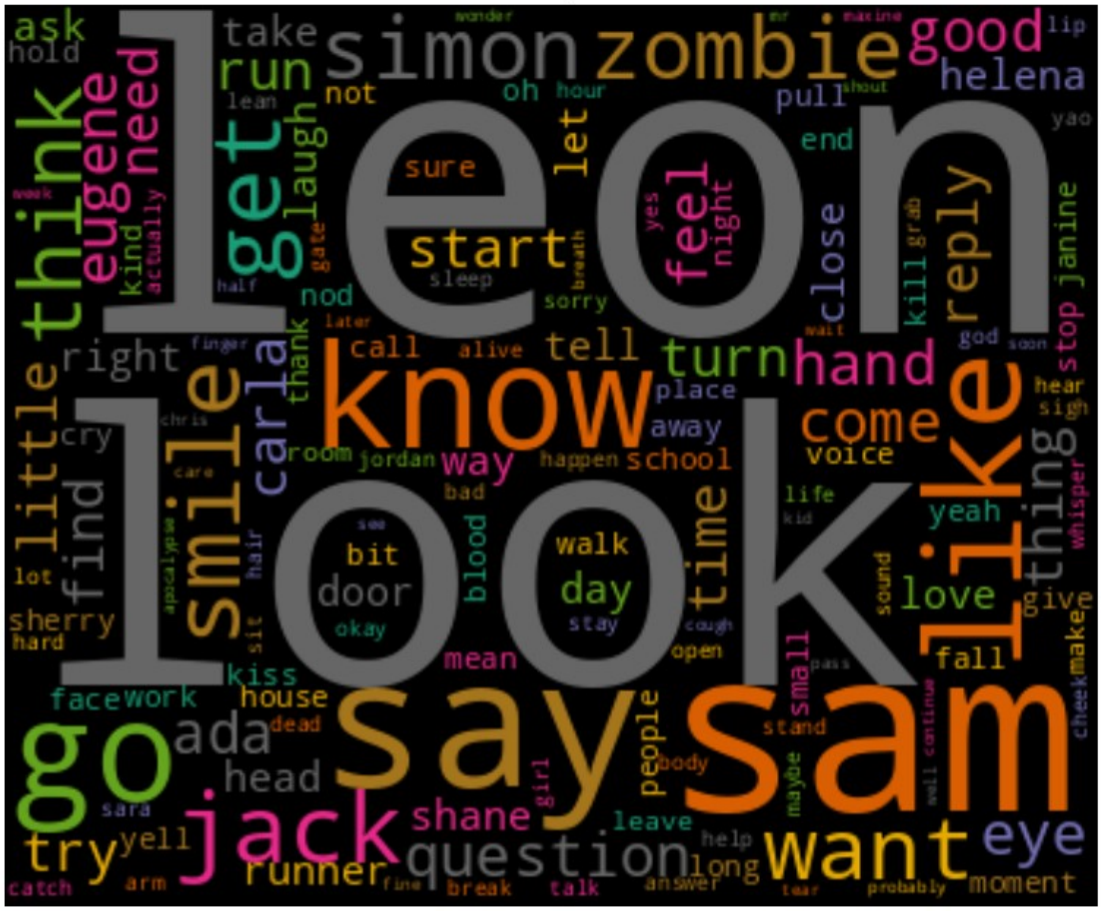
Zombie-Loan


chika
shito
michiru
look
reaper
feel
know
try
koyomi
yuuta
close
bekko
stop
see
long
end
voice
smile
zombie
take
way
hold
eye
say
away
little
come
head
right
time
ask
leave
tell
want
akatsuki
hand
kill
go
xoxox
face
momoka
like
sleep
work
think
start
place
yomi
thing
room
let
sotetsu
girl
door
turn
sound
hear
not
good
help
pull
lip
man
slowly
xox
speak
day
walk
arm
scythe
inside
fight
friend
floor
chihiro
black
form
watch
pain
sigh
kuze
lose
search
mean
bed
make
sure
hell
break
fall
continue
word
file
live
sit
die
red
wrap
talk
zloan
decide
mind
stare
office
chapter
hit
dog
house
throw
finally
boy
need
moment
wait
stay
hard
small
leon
bring
return
stand
hope
get
happen
notice
gopher
move
laugh
death
find
zaraME
mouth
body
light
realize
scream
run
give
care
chiru
blood
love
toho
finger
reach
human
night
open
life

Zombieland


say
columbus
tallahassee
look
sarah
christa
go
get
know
like
little
time
want
rock
zombie
think
wichita
head
find
ask
take
see
leaveneed
good
come
try
turn
way
not
give
tell
feel
hear
hand
door
eye
start
man
sister
face
sure
walk
car
right
smile
stop
long
girl
people
place
thing
mind
stand
hold
young
love
night
close
pull
guy
hope
big
lie
open
watch
road
drive
bit
seat
grab
end
wait
sleep
fire
room
tal
kiss
kid
lot
voice
cry
sorry
world
let
run
mom
maybe
da
ya
hard
second
shake
make
old
well
chapter
shoot
follow
bed
kill
ohio
house
break
mouth
moment
step
foot
point
store
care
remember
fall
catch
body
branson
day
vegas
water
gun
hit
new
okay
dead
away
soon
talk
reach
happen
hair
decide
bad
light
rule
mean
work
arm
buck
sit
laugh
sound
drive
idea
keep
life
meet

Zombies, Run!

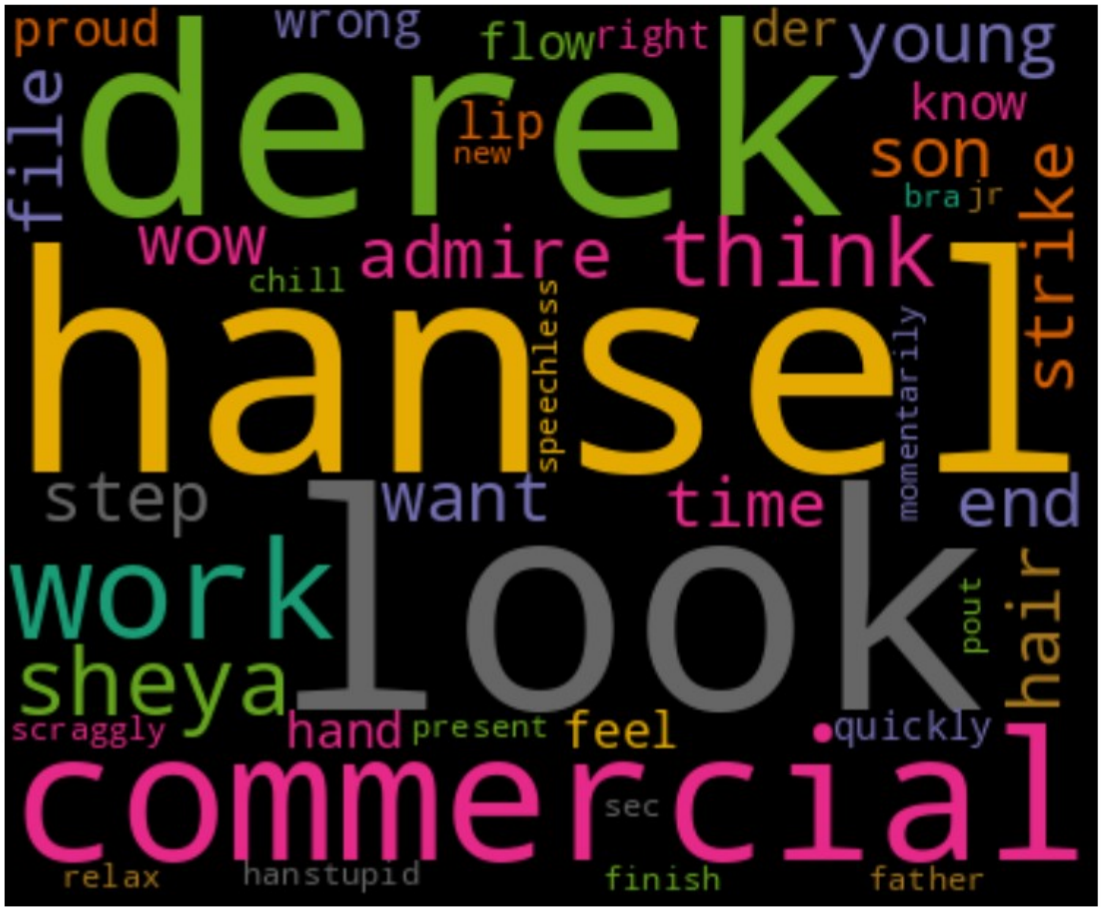

leon
look
like
know
say
sam
go
ada
jack
simon
zombie
good
helena
run
get
think
eugene
need
ask
take
hold
laugh
not
oh
hour
pull
let
end
sure
start
feel
night
close
reply
janine
stop
smile
carla
little
right
tell
turn
hand
call
alive
thank
nod
sorry
sleep
place
away
come
voice
school
way
room
jordan
happen
time
bad
blood
door
bit
walk
day
stay
yeah
love
thing
give
make
cheek
whisper
hear
sigh
find
sherry
lot
hard
okay
mean
open
kiss
fall
small
stand
people
body
face
work
house
dead
head
shane
girl
leave
try
yell
sara
question
long
want
eye
runner
fine
break
talk
answer
moment
probably
tear
catch
arm
maybe
wall
continue
sound
life
kid
cry
chris
care
hair
sit
apologize
see
lip
yao
kill
grab
god
soon
wait
shout
lean
gate
half
breath
yes
later
finger
wander
mr
maxine
pass
cough

Zone of The Enders

Zoo Tycoon


duke
heehee
animal
like
zoo
say
people
guest
bear
time
forest
exhibit
blue
stuff
life
course
enclosure
bath
house
blood
get
hippos
enjoy
serengeti
late
bird
river
know
start
shop
marathon
go
ape
god
cart
kid
drink
ice
water
new
little
stand
endanger
call
company
elephant
wild
thing
girl
kind
penguin
hear
mouth
heart
plush
lion
hey
look
cub
boy
croc
body
small
gorilla
area
voice
cold
stop
take
work
white
panda
big
different
african
come
red
hot
kiss
tiger
rock
dog
book
head
nail
way
good
sit
cage
middle
beast
picture
marvel
joseph
end
gift
perfect
yelled
dragon
toy
oh
ass
spot
mother
sell
crocodile
land
glover
think
capital
giant
slowly
polar
tycoon
komodo
savannah
gibbon
lemur
gitzlaff
bengal
warn
family
wide
man
eat
orangutan
hem
today
able
startled
fit
near
park
file
want
find
fill
home
light
usual
door
temperate
baby
give
variety
couple

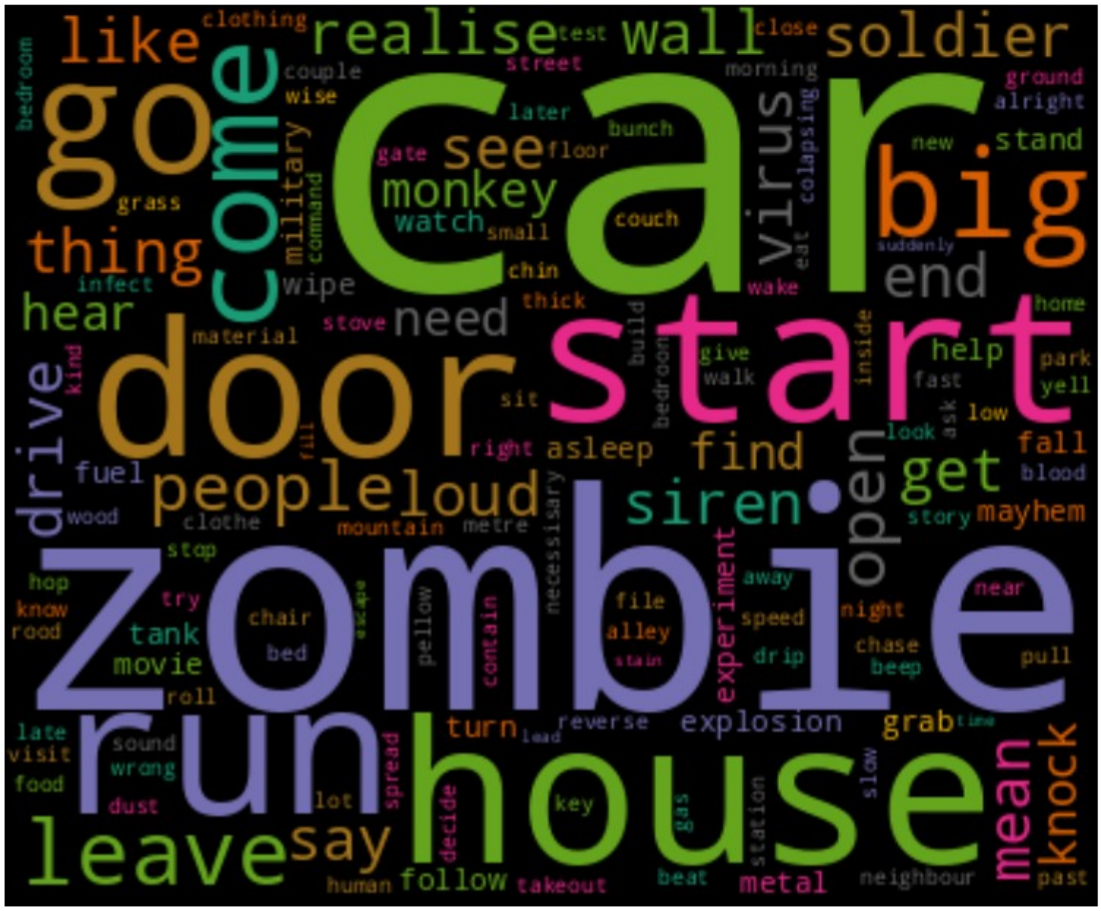
Zoolander


derek
hansel
look
commercial
work
sheya
wow
admire
think
young
know
son
strike
file
lip
new
proud
wrong
flow
right
der
bra
jr
chill
speechless
momentarily
step
want
time
end
hair
pout
scraggly
hand
present
feel
quickly
sec
relax
hanstupid
finish
father

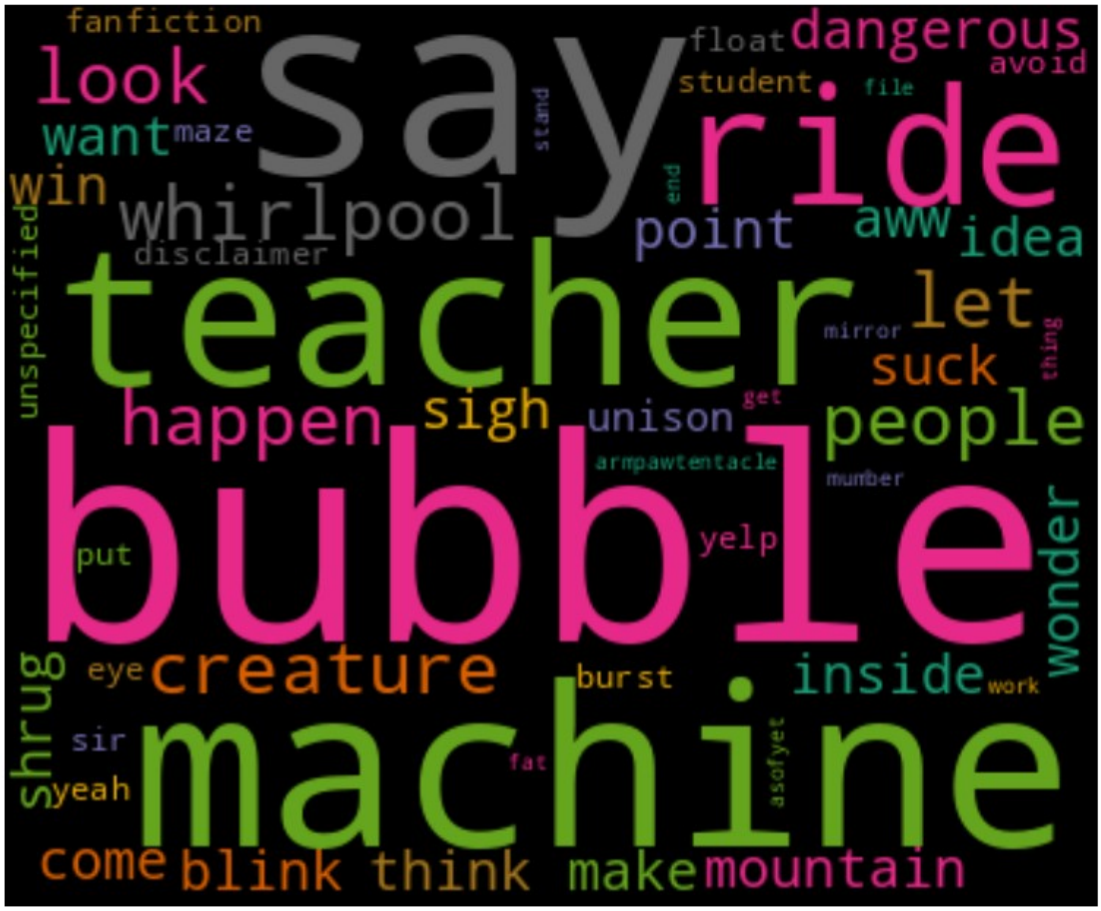
Zoom


summer
go
connor
brother
jackie
dylan
like
know
look
jack
chris
time
toby
try
eye
team
say
get
watch
run
room
think
year
feel
take
power
want
face
old
guy
long
close
happen
little
leave
thing
day
turn
hand
away
finally
tear
thirty
sure
sound
lavinia
new
kill
grin
pant
talk
right
start
lip
file
freak
tell
touch
arm
boy
girl
grant
come
good
bed
cry
end
home
give
hard
pull
parent
whisper
wonder
yeah
way
door
ask
close
nod
ready
believe
friend
vortex
mindsight
realize
cindy
open
small
sleep
deep
pray
send
notice
school
smile
stop
scream
anymore
young
lyl
soon
oh
keep
hold
okay
walk
sit
radiation
slowly
zenith
matter
book
find
not
forget
sigh
hair
marsha
moment
body
father
light
sense
thought
gamma
fall
accept
well
hear
listen
let
frown
work
job
sob
hope
rest
enzo
need
private
night
head
see
zoom
bad

Zoombie Blondes



Zoombinis


fanfiction
say
float
dangerous
avoid
look
student
file
ride
want
maze
stand
win
whirlpool
end
point
aww
idea
unspecified
disclaimer
teacher
let
mirror
thing
suck
happen
sigh
unison
get
people
armpawtentacle
number
bubble
wonder
yelp
put
shrug
eye
creature
burst
inside
work
sir
fat
asofyet
machine
yeah
come
blink
think
make
mountain

Zoop


anthony
look
blush
walk
kissian
hand
smile
room
bed
want
way
eye
whisper
shirt
stand
sit
dare
meet
pull
open
feel
lie
change
time
camera
bathroom
towel
push
find
end
yes
speed
hold
know
pass
lead
hair
touch
pant
leave
day
speak
worry
jack
round
happen
set
wait
slip
thank
drink
hour
lip
fall
take
think
beer
wanna
hear
move
clean
pepsi
file
sofa
shower
bench
break
nod
sound
guy
door
louder
turn
tip
swift
fanfic
say
start
fish
pretty
dream
ask
make
jacksepticeye
roommate
house
climax
cover
top
gently
rate
long
slump
pair
diaper
straight
like
redden
follow
arm
hip
sip
wink
yeah
dude
hood
fast
face
reach
explode
floor
shock
bother
upstairs
right
position
youtuber
stair
oh
red
wrap
youto
smosh
tongue
stupid
not
tomorrow
shake
andwho
downstair
ok
get
idol
show
explore
thirsty
fault
blue
god
go
whatoh
ianthony
anymore actually
embarrassment

Zorro


```
#Saving the vectorized data to a csv
dtm_file_path = 'dtm_data.csv'
cleaned_data_path = 'cleaned_data.csv'

# Save DataFrame to CSV
df_dtm.to_csv(dtm_file_path, index=False)
df.to_csv(cleaned_data_path, index=False)
```