CAN I MAKE MY OWN SOCIAL THREAT SCORE?

MASTERCHEN

@CHENB0X

WHAT BROUGHT ME HERE?

Psychology

Writer

Con Speaker

Hacker

2600

DEF CON

Telephreak

Teacher/Mentor

Martial Artist

Recon Village

BSidesLV

DC Skytalks

Self Defense

Stalking/Anti-Stalking

Amateur Data Scientist
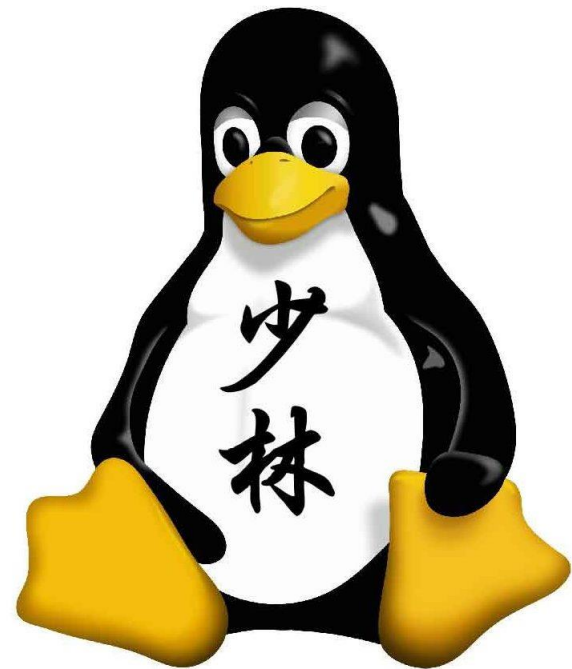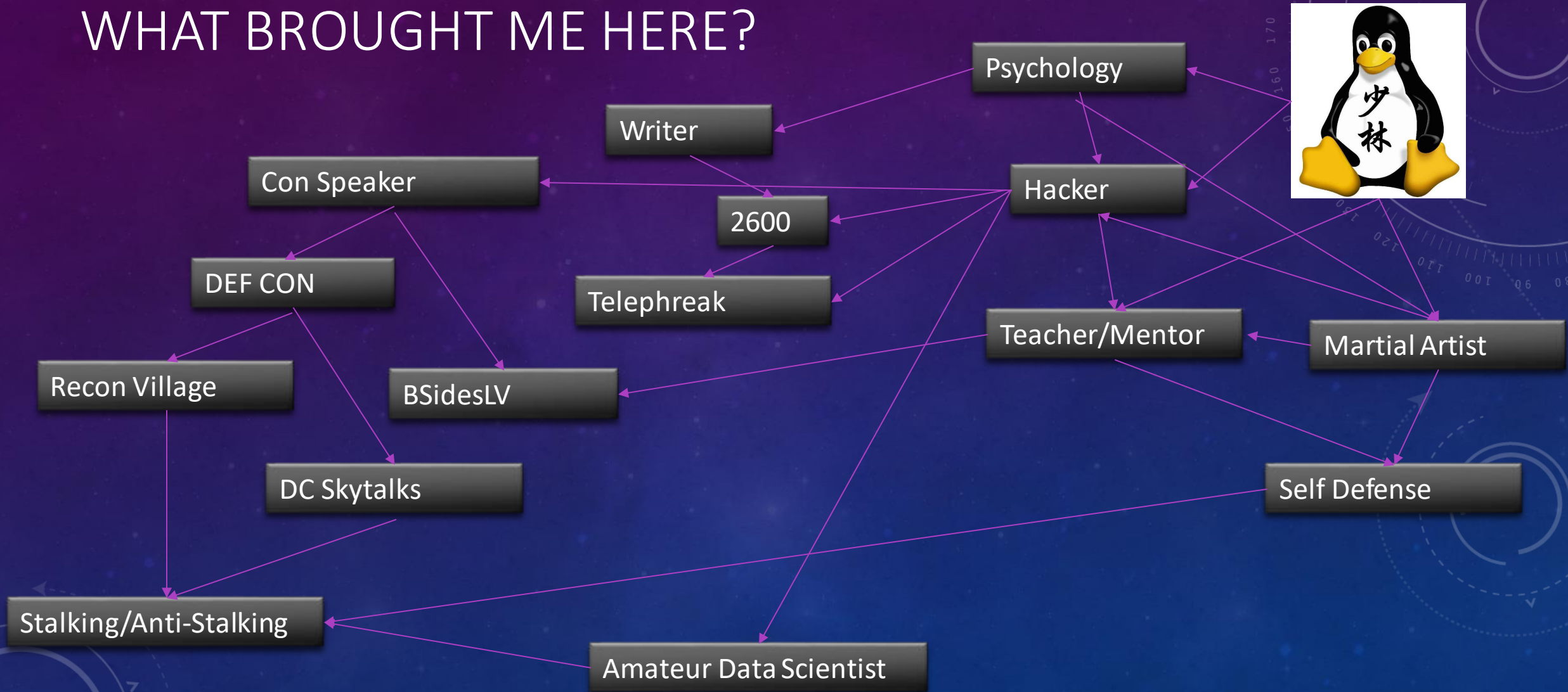
STANDARD DISCLAIMERS

IANAL

IANAS

INADS

My Twitter account and any other social media is subject to usage as a test bed for theory. You are all my guinnea pigs. Ye be warned, but also, THANK YOU!

CREDIT SCORES – BOTH FINANCIAL AND SOCIAL

- Risk scores based on accumulated metrics

- Used to influence behavior of the subject

- Provides "at-a-glance" summary of the subject within the context of the score

THREAT SCORES

- Uncover unknown imminent threats

- Tool for the person/team/entity in a position of defense

- Risk score in the opposite direction of "conventional" credit scores

- Pays attention to what matters*

Social Threat Score		
	What?	Score to determine the probability and risk that a profile/account/entity may pose
	Who?	Stalkers / Bullies
	Why?	Self Defense / Advanced knowledge / Psychological interest
	How?	Descriptive Statistics / Meaningful Metrics
	Where?	Twitter
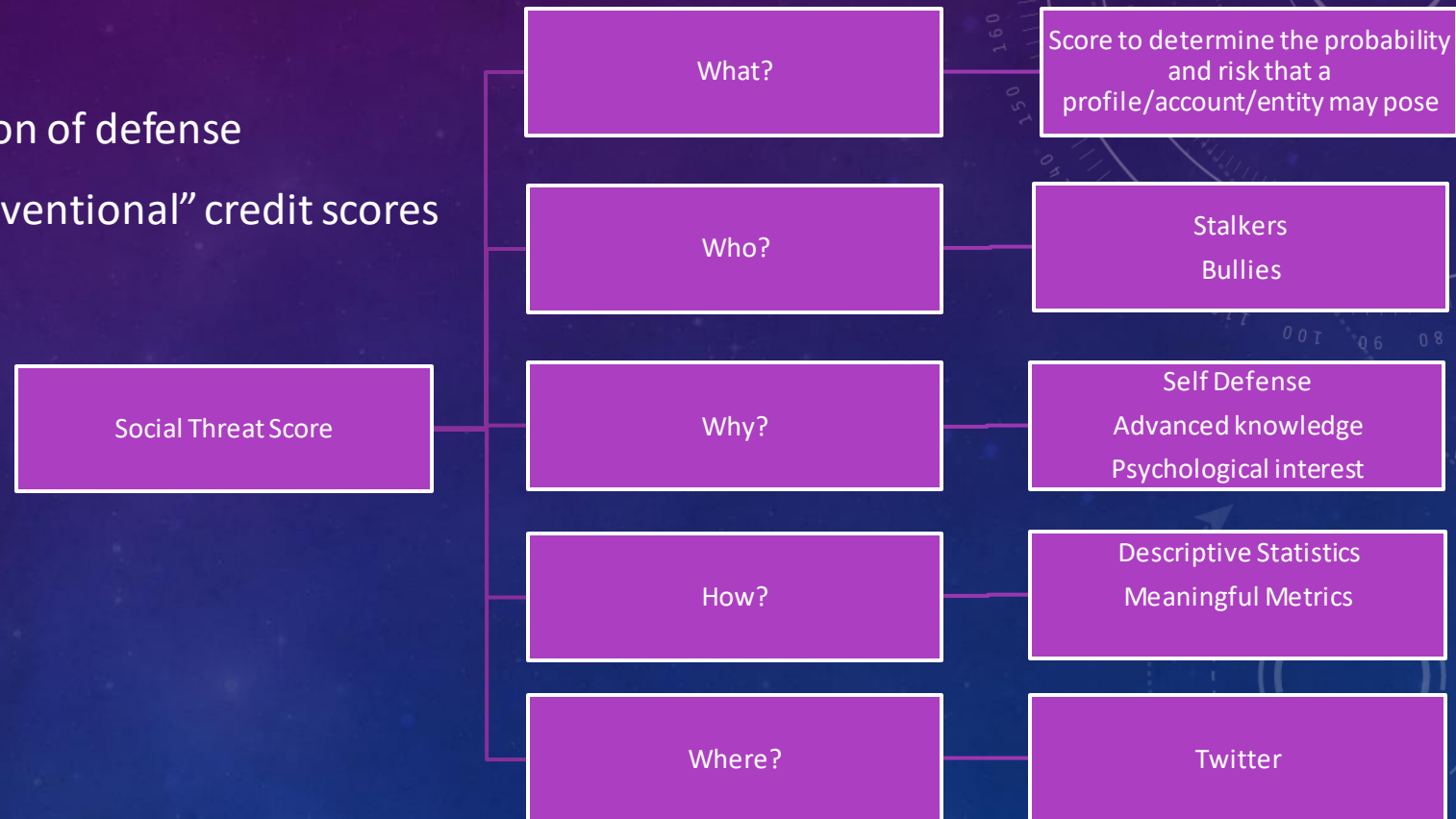
INITIAL THOUGHTS AND HYPOTHESES

- Twitter, as a *publicly available database of sentiment*, provides enough key indicators to determine if an entity has the *influence and motivation to promote adverse activity* against another entity; namely our own.

- Do **NOT** judge a book(person) by the cover, **BUT** get your computer to read the book and give you the cliff notes with a reasonable expectation of accuracy :-D

Word Frequency

Tweet Frequency

Active Timeline: 4160 days, 0:00:00
Average Tweets Per Day: 2.307932692307692

Verified Percentage

Sentiment Analysis

Verified: 27
Unverified: 968
Verified Percentage: 0.0054

Positive Sentiment Percentage: 0.2830601092896175
Negative Sentiment Percentage: 0.12415300546448088
Neutral Sentiment Percentage: 0.5927868852459016

Follow Ratio: 0.0003539509712467888

Follow Ratio

Meaningful Metrics
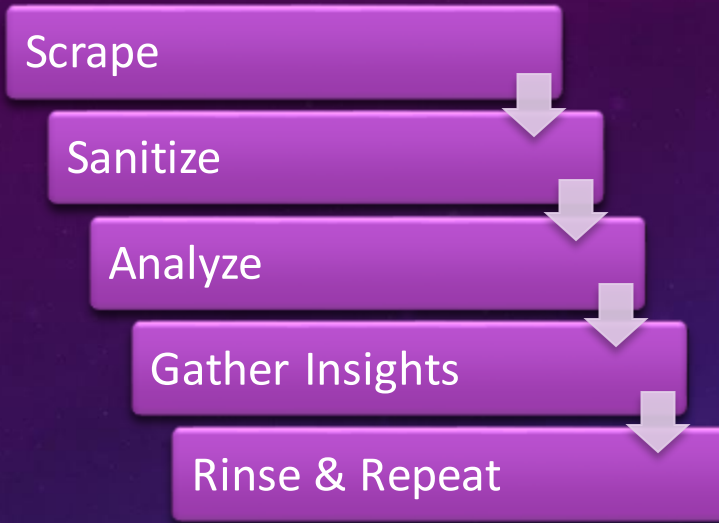
Hashtag Usage

---Top hashtags---
womenshistorymonth 24
debates2020 18
poetryin13 17
hangtenstories 15
writingcommunity 13
sixwordstory 9
covid19 6
poetry 5
prompt 5
christmascountdown 5
election2020 5
vss365 4
scifaikusaturday 3
films 3
filmtwitter 3
mtndewmajormelon 3
mypromotionentry 3
31daysofhorrormovies 3
vpdebate 3
wearamask 3

THE PROCESS

Scrape

Sanitize

Analyze

Gather Insights

Rinse & Repeat

```
Verified: 27
Unverified: 968
Verified Percentage: 0.0054
```

```
Means: Following 5815.747600767754, Followers 10302.415547024952 Tweets 6934.265834932821
Medians: Following 801.0, Followers 245.0 Tweets 694.5
Modes: Following 0     69
dtype: int64, Followers 0     10
dtype: int64 Tweets 0     0
dtype: int64
Mins: Following 2, Followers 1 Tweets 0
Maxes: Following 1469793, Followers 2498255 Tweets 741584
S. Deviations: Following 51837.986230493865, Followers 90581.36623660615 Tweets 30219.037135394614
```

1. Scraped and analyzed my own followers first for "immeforte" threats
   - Descriptive statistics
     - Mean, Median, Mode, S. Deviation
   - Find the "influencers"
     - High follower counts, follow ratios, verified
     - Remove corporate accounts, bots, etc
   - Grab timeline of "influencers"
     - Normalize (remove punctuation and lower case)
     - Check for various frequencies and compare
   - Spot check sample size of the follower population
     - Bullying behavior
2. Search for notable profiles to see comparison
   - Journalists
   - Celebrities
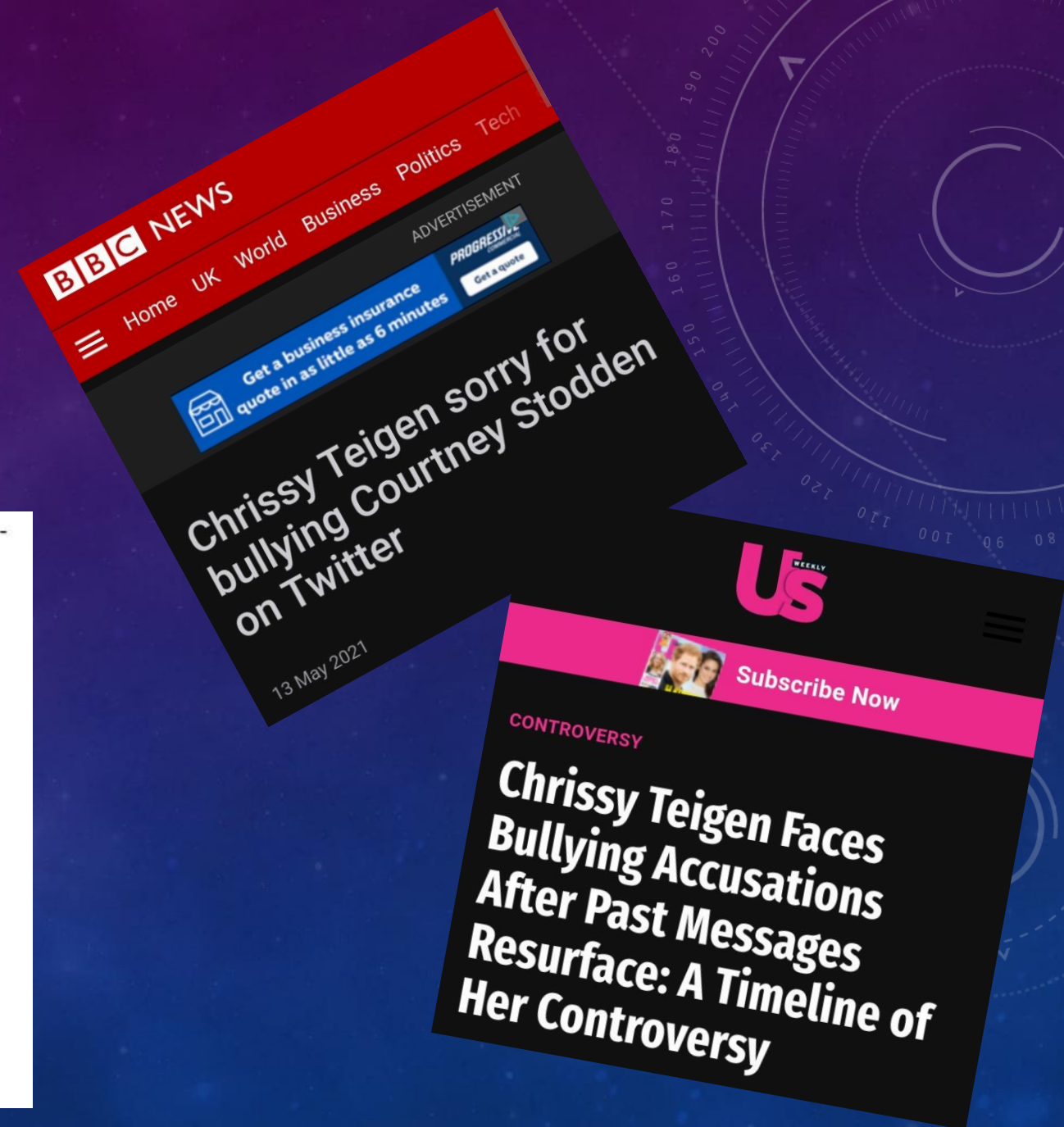   - Politicians
   - Tech moguls

THE TEIGEN BASELINE

- In the news recently for cyber bullying

- Highly influential; "celebrity" status

- Makes for a good case study and baseline

```
ID: 39364684
Username: chrissyteigen
User Bio: de-motivational speaker
Account birthdate: 2009-05-11 22:50:04
Followers Count: 13524472
Friends Count: 4787
Tweets: 14383
Follow Ratio: 0.0003539509712467888
```

```
---chrissyteigen---
Positive Sentiment Percentage: 0.4161280393967375
Negative Sentiment Percentage: 0.17543859649122806
Neutral Sentiment Percentage: 0.40843336411203446
```
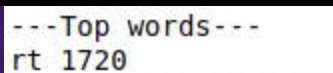
```
---Top words---
lol 246
it's 205
like 192
i'm 187
rt 179
oh 139
people 135
know 134
 130
one 128
don't 120
love 113
get 109
yashar 105
good 99
god 93
u 92
really 92
would 87
much 81
```


BBC NEWS
Chrissy Teigen sorry for bullying Courtney Stodden on Twitter

13 May 2021

CONTROVERSY
Chrissy Teigen Faces Bullying Accusations After Past Messages Resurface: A Timeline of Her Controversy

CASE SAMPLES – NEGATIVE SENTIMENT


```
---Top words---
rt 829
 151
like 133
people 129
it's 125
i'm 93
don't 93

one 84
thank 83
new 78
know 77
infosec 73
work 73
think 71
thanks 71
us 67
get 67
time 64
would 62
```

Positive Sentiment Percentage: 0.2830601092896175
Negative Sentiment Percentage: 0.12415300546448088
Neutral Sentiment Percentage: 0.5927868852459016


```
---Top words---
rt 1720

biden 279
trump 255
like 166
it's 164
media 163
people 161
 155
one 142
arthurschwartz 128
amp 124
would 122
democrats 114
us 110
joe 106
get 101
new 98
want 95
project 94
```

Positive Sentiment Percentage: 0.2830349531116795
Negative Sentiment Percentage: 0.15004262574595056
Neutral Sentiment Percentage: 0.56692242114237


```
---Top words---
rt 446



trump 236
people 197
like 171
im 153
one 145
youre 128
dont 121
 117
know 105
thats 98
right 83
first 82
said 81
get 80
time 76
didnt 75
```

Positive Sentiment Percentage: 0.23374091795144428
Negative Sentiment Percentage: 0.1213893319156477
Neutral Sentiment Percentage: 0.644869750132908

CASE SAMPLES – POSITIVE SENTIMENT
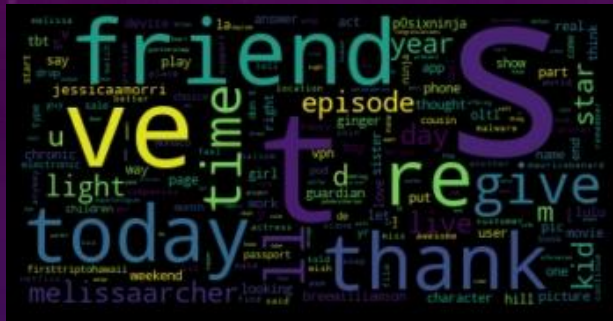

```
---Top words---
rt 794
thank 358

 168
defcon 143
thanks 139
happy 121
amp 106
one 106
new 98
like 97

love 90
it's 87
i'm 87
nitotv 87
birthday 76
guardianiosapp 71
get 69
see 68
```

```
Positive Sentiment Percentage: 0.3956043956043956
Negative Sentiment Percentage: 0.07171775592828225
Neutral Sentiment Percentage: 0.5326778484673221
```

*


```
---Top words---
astrology 1860
new 933
post 927
 305
guru 55
life 51
quote 51
practice 44
rt 43
suffering 43
beings 39
love 39
jetsunma 38
compassion 37
rinpoche 37
♥ 36
hung 35
amp 34
mind 34
us 33
```
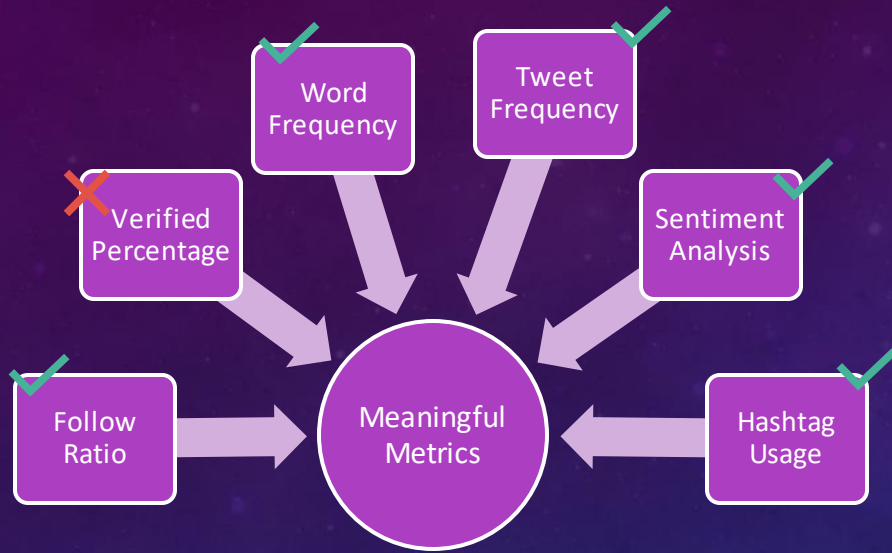
```
Positive Sentiment Percentage: 0.36031892057651027
Negative Sentiment Percentage: 0.03189205765102729
Neutral Sentiment Percentage: 0.6077890217724624
```

THERE WAS SOMETHING ABOUT A SCORE?

Word Frequency ✓

Tweet Frequency ✓

Verified Percentage ✗

Sentiment Analysis ✓

Follow Ratio ✓

Meaningful Metrics

Hashtag Usage ✓

| Low |
| Medium |
| High |
| Severe |

- Scoring is...hard
  - Delta with combination of factors
  - Weighting
  - Percentage breakdown

INSIGHTS & IMPROVEMENTS

Insights

- Word frequency itself doesn't seem to be any indicator
- Negative sentiment may also be self deprecating…sadly, it's the highest negative percentage seen in the scrape (aside from a DJT parody account
- The "You" tendency

Improvements

- Profanity tracking
- Word/phrase searching
- Graphing the data
- Fine-tuning sentiment analysis

RESOURCES

- https://merics.org/en/report/chinas-social-credit-system-2021-fragmentation-towards-integration

- https://www.myfico.com/credit-education/whats-in-your-credit-score

- https://us-cert.cisa.gov/CISA-National-Cyber-Incident-Scoring-System

- https://GitHub.com/MasterChenb0x/SocialThreatScore