Twitch Streams: Detecting Hate

By: Kelsey Lane

Agenda

01 Business Understanding

02 Data Overview

03 Modeling

04 Suggestions

05 Future Research

01

Business Understanding

1 in 10 users reported they've been bullied or felt unsafe while streaming

Business

StreamElements (2018)

3 out of 4 users think Twitch is not effective in addressing bullying

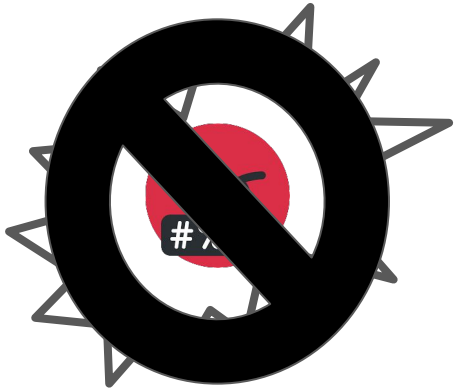
Business

StreamElements (2018)

So What Can We Do?

Business

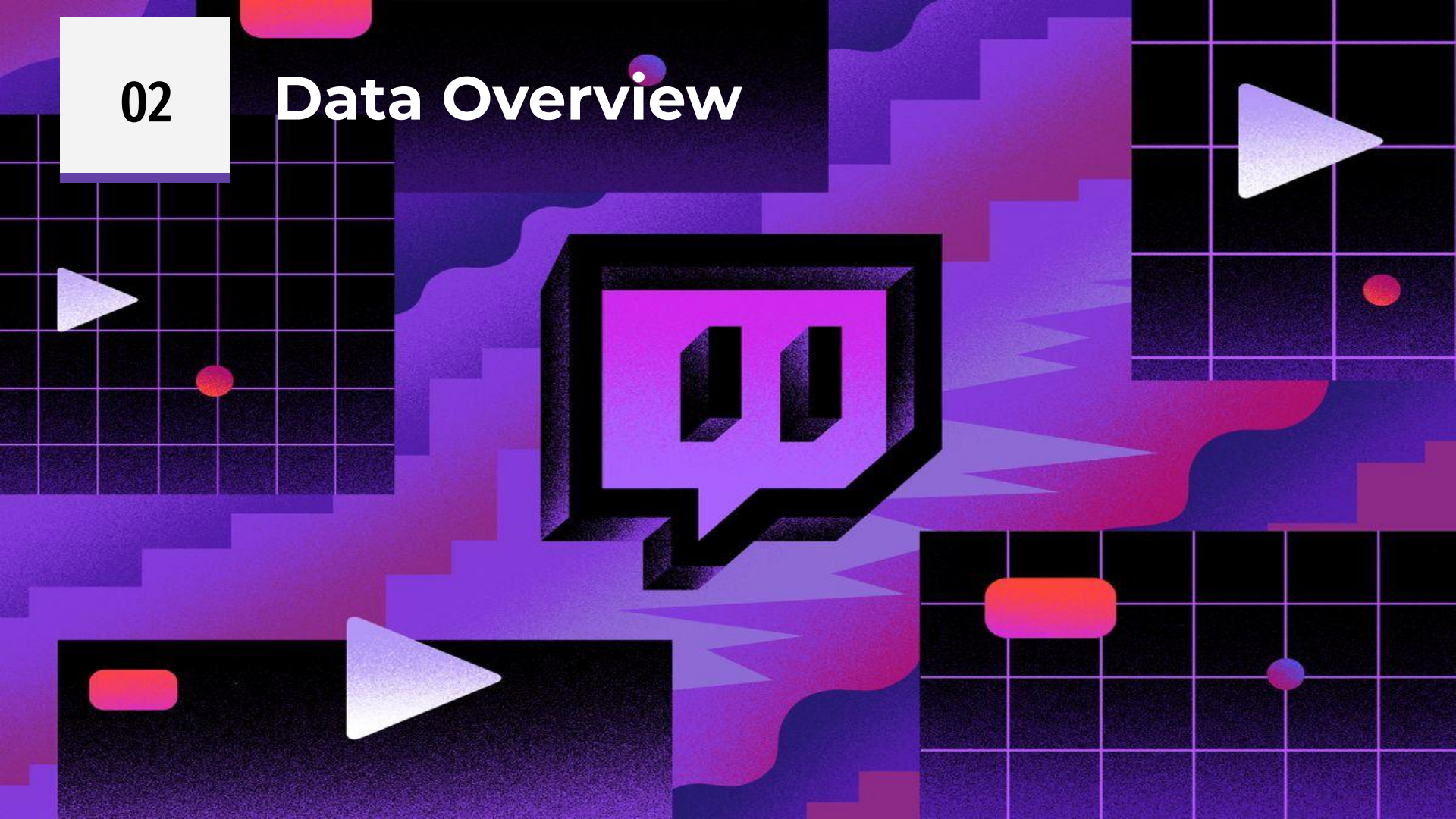
So What Can We Do?



Business

02

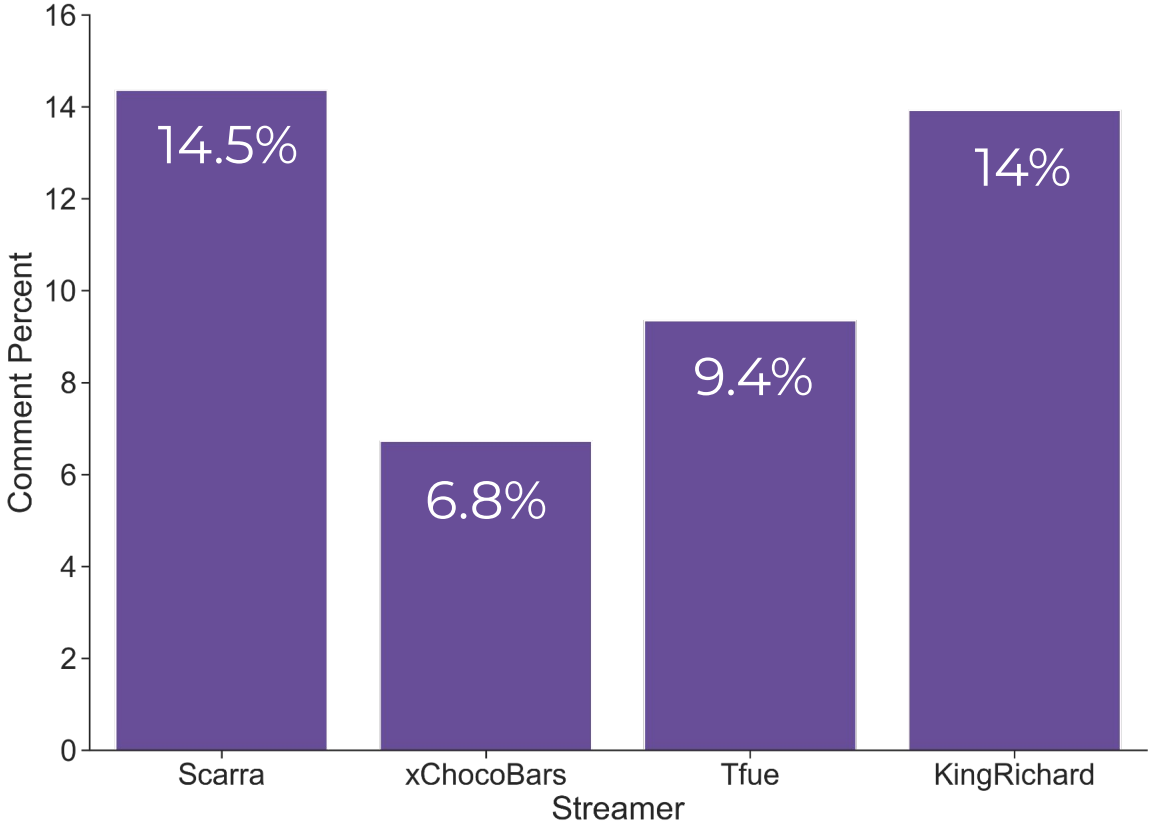
Data Overview



Data

- Twitch livestream chats
  - May - June 2018
  - 4 streamers
  - 9 total streams
- 102,700 total

Percent Negative Comments By Streamer

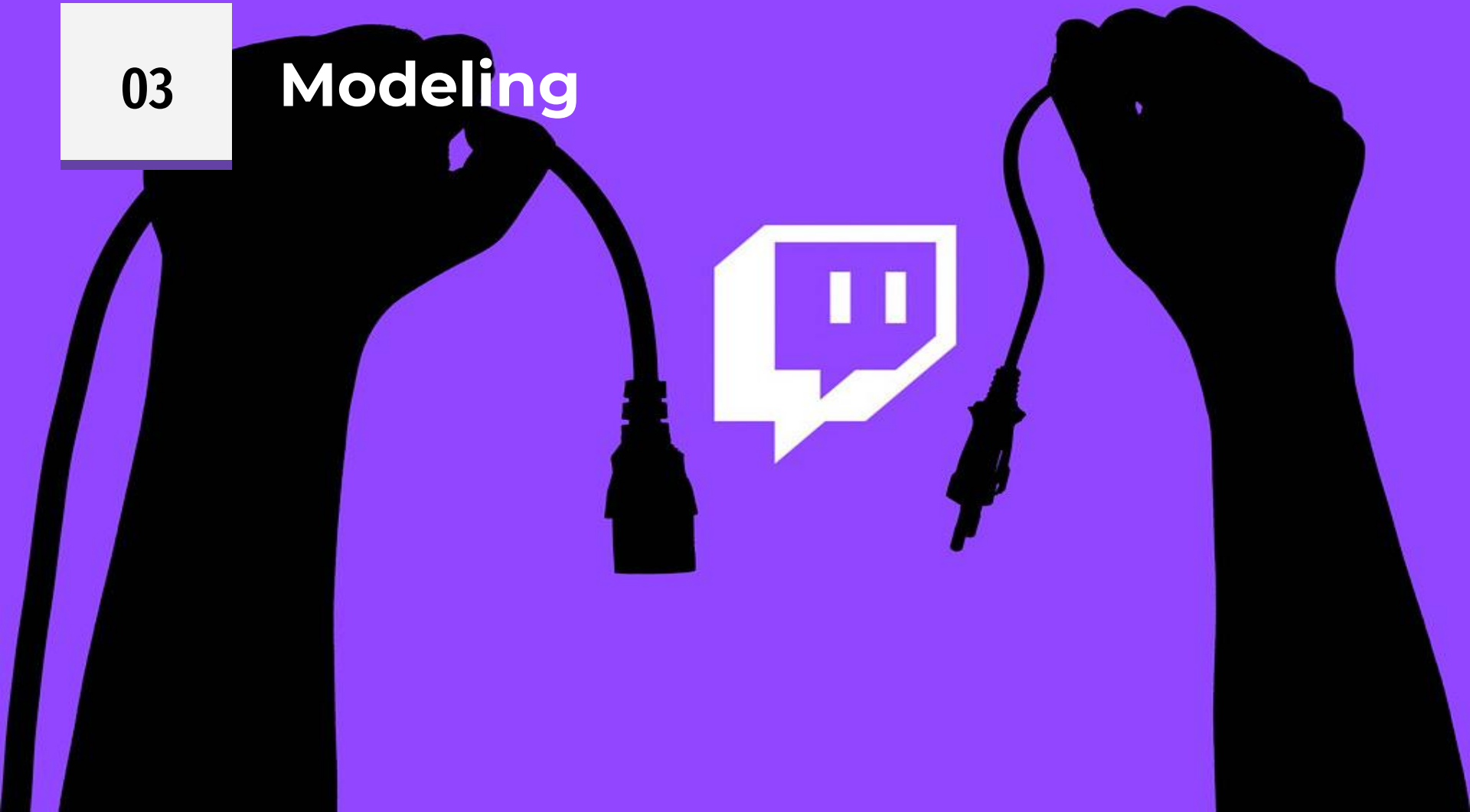


Data

think
got
u
lol
hate
damn
hard
dont
bad
banned
ass
one
people
now
damage
chat.
rip
go
stream
make
game
ban
time
wtf
guy
shit
stop
fuck
trash

Data

03 Modeling



Sentiment Analysis

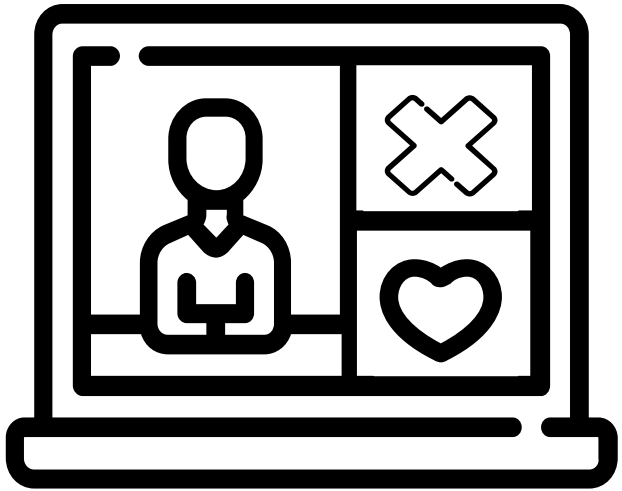
- Identify negative chats
  - Find frequent words
- 76% Recall

Model

04 Suggestions

Suggestions

- Flag hate chats
- Mark frequent offenders
- Build stronger filter
    - Focus on misspellings



Suggest

05

Future Research

Future Research

- Model specific categories
- Other languages
- Better labels

Future

kelsklane

kelsklane@gmail.com

Thank you!

Questions?

Appendix

Methods

- Keep uppercase + stopwords
- Tweet Tokenizer + Lemmatize
- Count Vectorizer

Engineered Features

- Time of day
- Ratio capital letters
- Offset
- Total comments per user
- Any emotes

Chat Examples

Positive/Neutral:

- Congrats king! Big fan!

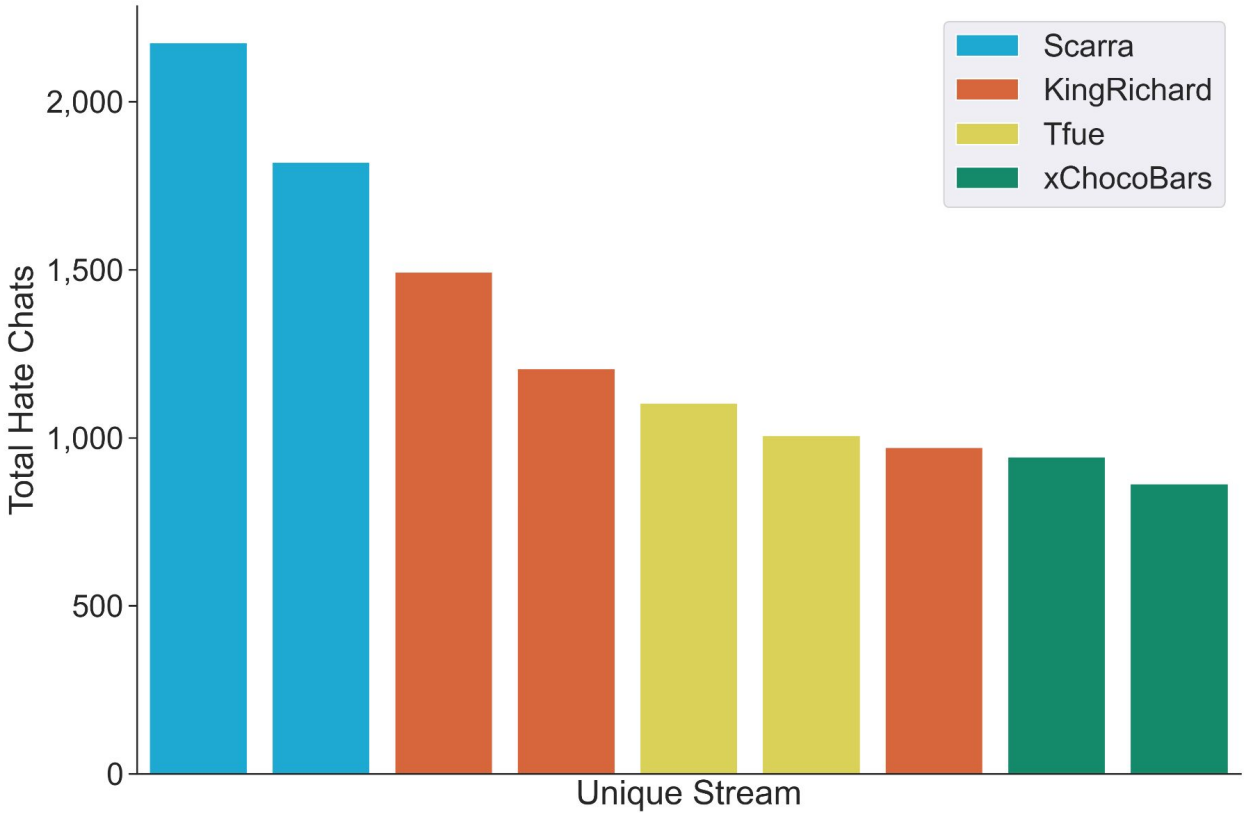
Hate:

- i'm better than this trash man

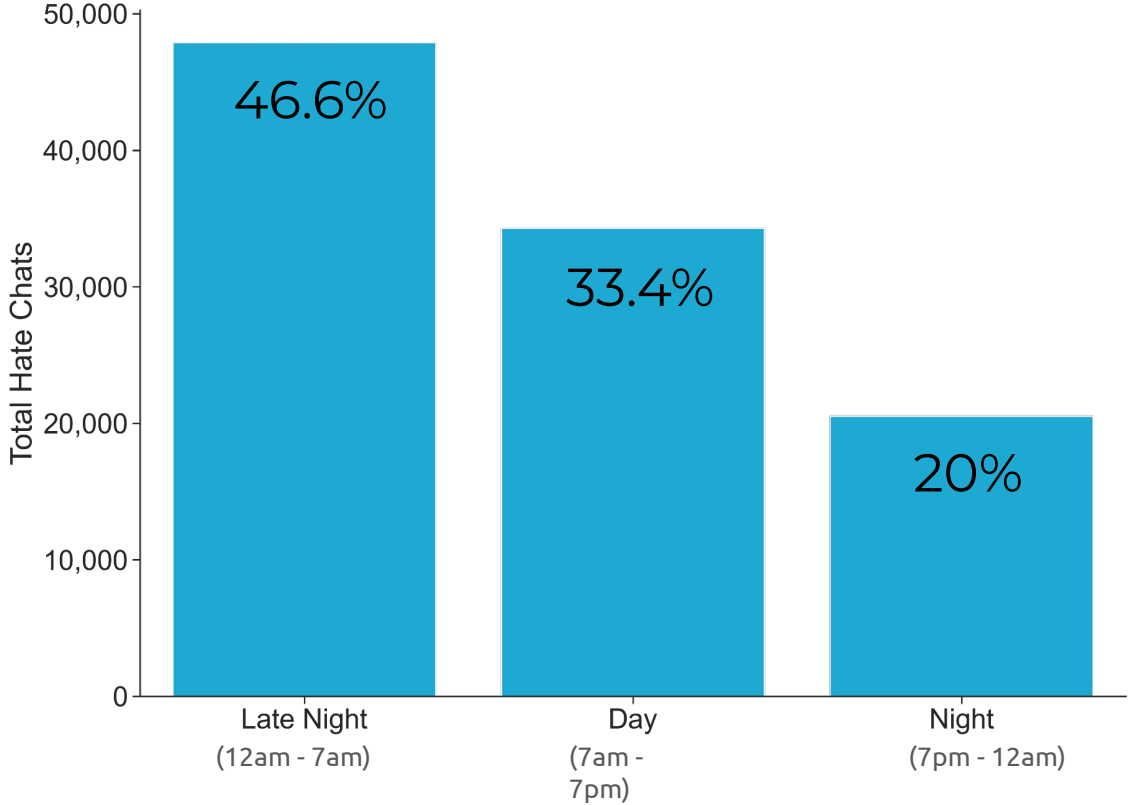
Negative, non-hate:

- i wanna sub but i’m broke FailFish

Number of Hate Comments For Each Stream

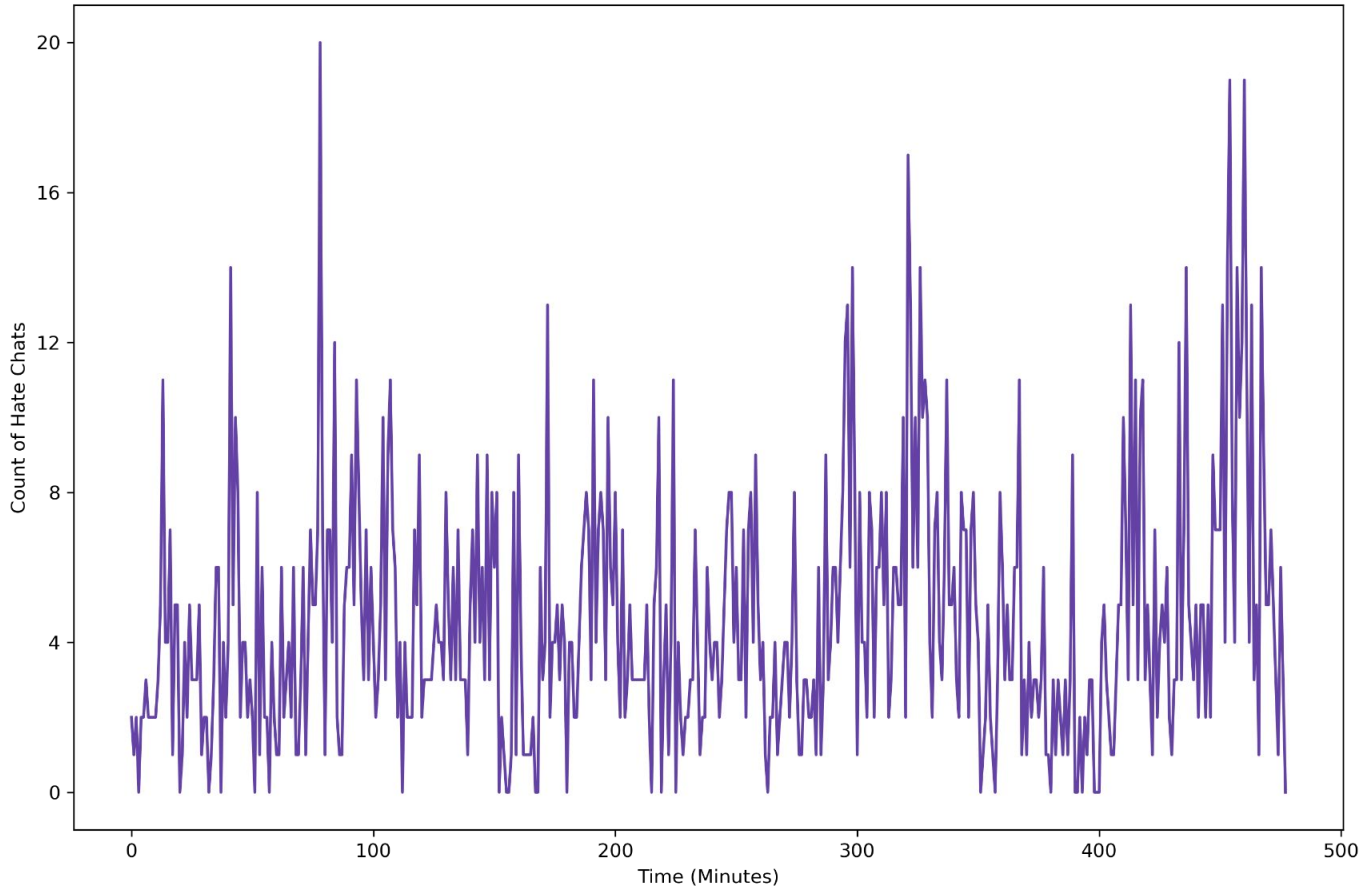


Number of Hate Comments By Time of Day

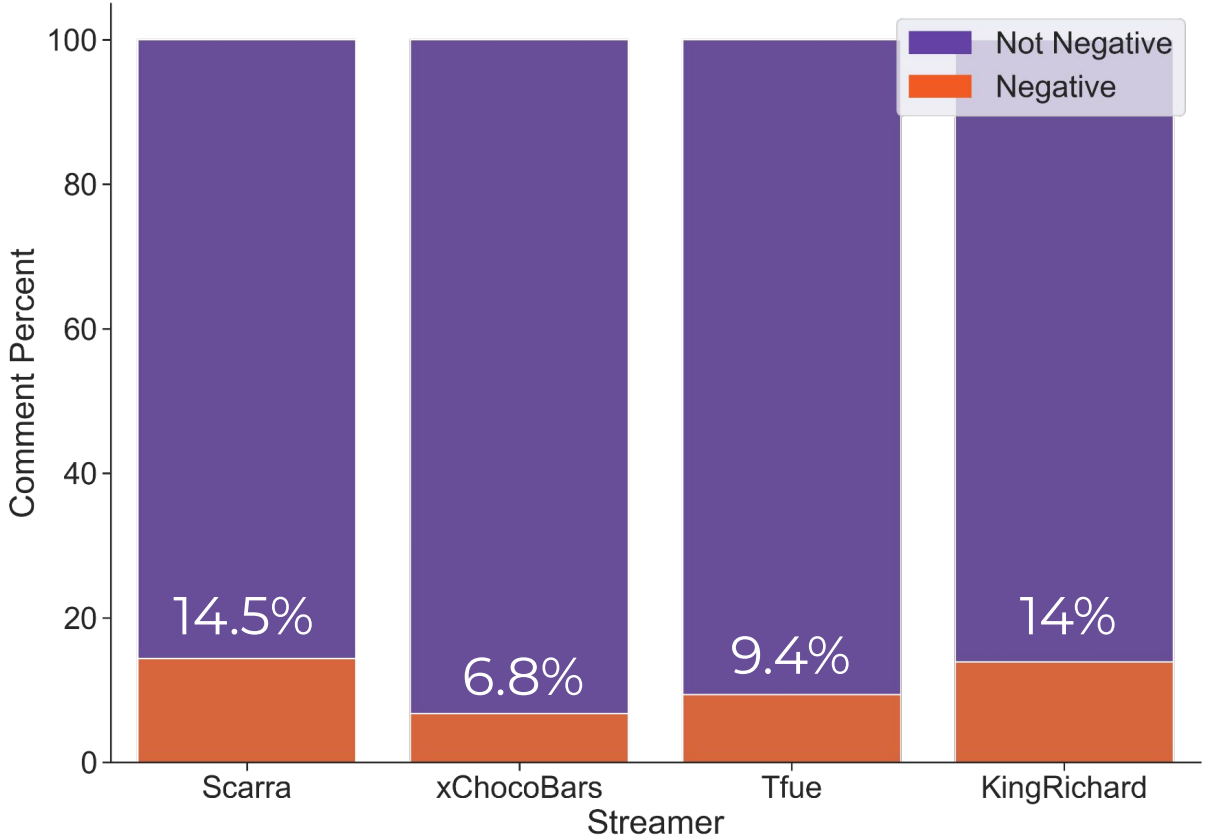


Data

Number of Hate Comments Over Time For One Stream



Percent Negative Comments By Streamer



Data

Credits:

Image:
https://www.polygon.com/22700564/twitch-hate-raids-safety-feature-streaming

Survey:
https://blog.streamelements.com/streamelements-analysis-on-twitch-bullying-c3f2b2240318