

Tweet Emotion Dynamics

Emotion Word Usage in Tweets from US and Canada

Krishnapriya (KP) Vishnubhotla

vkpriya@cs.toronto.edu



Computer Science
UNIVERSITY OF TORONTO



Saif M. Mohammad

saif.mohammad@nrc-cnrc.gc.ca



National Research
Council Canada

Conseil national de
recherches Canada

Overview: We Introduce

- **Tweet Emotion Dynamics (TED):** A computational framework to quantify changes in emotion word usage in tweets over time
 - Generalizable to any form of text utterances
- **Tweets from US and Canada (TUSC):** A dataset of geo-located tweets from the USA and Canada, 2015–2021
 - TUSC-City (2020–2021): collected from 46 cities
 - TUSC-Country (2015–2021): country-wide tweets

Overview: Questions of Interest

- What are the yearly trends in emotion word usage from 2015 to 2021?
 - How has the COVID-19 pandemic impacted these trends? (2019–2020–2021)
- How do Canada and the US compare in terms of emotion word usage?
 - How do these change over time?
 - Did the pandemic affect the two countries differently?
 - How much variation do we find at the city-level for both countries?
- What are the patterns of emotion word usage across time for individual users?
 - When aggregated at city/country–level, what trends do we see?

Dataset: Tweets From US and Canada (TUSC)

Data Collection

Twitter APIs

- Free API
 - Allows access to random tweets from the past week
 - Used to collect TUSC-City data
 - April 1 2020 – Dec 2021
- Academic API
 - Allows access to historical tweets
 - Used to collect TUSC-Country data
 - Jan 2015 – Dec 2021
 - More data from 2019–2021

Data Curation

- Tokenize the tweet text (Python package: Ttokenize)
- Keep one tweet per user per day
- English-language only
- Remove retweets
- Remove tweets with URLs and media links
- Remove tweets with less than three tokens
- Organize by:
 - Location: Country, City
 - Date of Posting: Year, Month

TUSC: Statistics

Dataset	Canada			USA		
	#tweets	# tweeters	Av.TpT	#tweets	# tweeters	Av.TpT
TUSC-Country						
2015	89,566	40,290	15.729	131,330	104,670	13.805
2016	93,280	40,994	16.164	133,413	109,110	14.305
2017	94,364	39,258	18.067	133,854	107,080	16.015
2018	95,403	38,866	21.763	133,066	105,227	19.394
2019	330,361	70,122	22.040	339,186	204,311	19.341
2015–2019	702,974	159,284	18.753	870,849	516,885	16.572
2020	321,176	57,465	22.123	503,976	250,080	19.698
2021	304,106	49,128	22.192	478,798	214,653	19.566
2015–2021	1,328,256	206,691	19.73	1,853,623	802,369	17.45
TUSC-City						
2020 (Apr–Dec)	15,039,503	716,063	19.275	23,470,855	2,669,081	17.556
2021	22,371,990	798,602	19.367	43,693,643	3,247,124	17.306
2020–2021	37,411,493	1,049,774	19.327	67,164,498	4,274,374	17.413

TUSC: Statistics

Canadians use ~2 more tokens per tweet!

Dataset	Canada			USA		
	#tweets	# tweeters	Av.TpT	#tweets	# tweeters	Av.TpT
TUSC-Country						
2015	89,566	40,290	15.729	131,330	104,670	13.805
2016	93,280	40,994	16.164	133,413	109,110	14.305
2017	94,364	39,258	18.067	133,854	107,080	16.015
2018	95,403	38,866	21.763	133,066	105,227	19.394
2019	330,361	70,122	22.040	339,186	204,311	19.341
2015–2019	702,974	159,284	18.753	870,849	516,885	16.572
2020	321,176	57,465	22.123	503,976	250,080	19.698
2021	304,106	49,128	22.192	478,798	214,653	19.566
2015–2021	1,328,256	206,691	19.73	1,853,623	802,369	17.45
TUSC-City						
2020 (Apr–Dec)	15,039,503	716,063	19.275	23,470,855	2,669,081	17.556
2021	22,371,990	798,602	19.367	43,693,643	3,247,124	17.306
2020–2021	37,411,493	1,049,774	19.327	67,164,498	4,274,374	17.413

Pre-pandemic +
pandemic years



Experiments

What are the trends in emotion word usage across years, and across countries?

Emotion Word Usage in TUSC

- What kinds of emotions?
 - Valence (V): positiveness – negativeness
 - Arousal (A): active – sluggish
 - Dominance (D): feeling in control – feeling things are out of control
- How do we measure the V/A/D score of a tweet?
 - Average of the V/A/D scores of constituent tokens
- How do we measure V/A/D score of a word?
 - Word–emotion score lexicons
 - NRC VAD Lexicon: 20K English Terms with real-valued VAD scores

Research Questions

- Are there notable trends across years in the valence, arousal, and dominance of tweets?
- Are we tweeting with more positive words, more negative words, more high arousal words, etc. than in past years?
- How are Canada and US different in terms of emotion word usage?
- How has COVID-19 impacted these metrics?

1. Average V/A/D of Tweets

- Valence scores of tweets average ~0.65

Canada – USA

- Canadian tweets are more positive, on average (higher valence)

Across Years

- Tweets from 2020 have the lowest valence
- Slight decrease from 2015–2021

Digging deeper

- June 2020 has the lowest valence values for both countries – first COVID-19 peak? Black Lives Matter movement?
- The final months of 2021 have the highest valence – heading back to normal?

Valence		Ave. 	
Dataset	Year	Canada	USA
TUSC-Country	2015	0.680	0.652
	2016	0.677	0.653
	2017	0.672	0.648
	2018	0.675	0.649
	2019	0.674	0.648
	2015--2019	0.676	0.650
	2020	0.669	0.644
TUSC-City	2020	0.659	0.644
	2021	0.665	0.653

1. Average V/A/D of Tweets

- Arousal scores of tweets average ~0.45

Canada - USA

- American tweets are more aroused/active

Across Years

- Peak arousal in 2019 for the USA
- Scores drop from 2017-2021 for Canada

Digging deeper

- March/April 2020: lowest arousal values for both countries

Arousal	Ave. 	Canada	USA
TUSC-Country	2015	0.436	0.439
	2016	0.440	0.444
	2017	0.442	0.447
	2018	0.438	0.446
	2019	0.436	0.449
	2015--2019	0.439	0.445
	2020	0.434	0.448
TUSC-City	2020	0.446	0.454
	2021	0.446	0.454

1. Average V/A/D of Tweets

- Dominance scores of tweets average ~0.55

Canada - USA

- Canadian tweets have higher dominance scores (more in-control)

Across Years

- Lowest in 2015, highest in 2021 (more volatile in the US)

Digging deeper

- April 2020: lowest dominance values for both countries

Dominance	Ave. 	Canada	USA
TUSC-Country	2015	0.563	0.535
	2016	0.568	0.545
	2017	0.569	0.549
	2018	0.578	0.556
	2019	0.578	0.555
	2015--2019	0.571	0.548
	2020	0.576	0.558
	2021	0.578	0.561
TUSC-City	2020	0.573	0.562
	2021	0.576	0.564

2. Proportion of Tweets with Emotional Terms

- Split the V-A-D lexicons into two parts
 - Low V-A-D: scores between 0 and 0.33
 - High V-A-D: scores between 0.67 and 1
- Find percentage of tweets containing at least one token belonging to each sub-lexicon
- Test if differences between places and years are statistically significant (paired t-Test, p -value 0.001)

- **Valence:** we post more tweets with positive words (High Valence) than with negative words (Low Valence) – about 100% more.
- **Arousal:** we post more tweets with low arousal words than high (~40% more).
- **Dominance:** we post more tweets with high dominance words than low (~35% more)

Canada – USA

- Canadian tweeters use:
 - more high-valence words AND fewer low-valence words
 - more high dominance words
 - slightly more low arousal words

2020 – 2021

- We tweeted more low valence (negative) words in 2020

Dataset	Year	Low Valence		High Valence	
		Canada	USA	Canada	USA
TUSC Country	2015	36.3	38.4	80.2	75.4
	2016	37.6	39.6	80.8	76.7
	2017	40.8	43.2	82.9	79.0
	2018	42.7	45.7	83.2	80.2
	2019	42.9	45.3	82.7	79.4
	2015--2019	40.0	42.4	82.0	78.1
	2020	43.8	46.1	82.3	79.2
	2021	42.6	44.5	82.5	79.7

Dataset	Year	Low Arousal		High Arousal	
		Canada	USA	Canada	USA
TUSC Country	2015	55.8	51.1	40.9	38.1
	2016	56.4	52.0	42.1	40.0
	2017	59.4	55.6	45.0	43.5
	2018	61.4	57.8	47.1	46.0
	2019	60.8	56.7	46.5	46.1
	2015--2019	58.8	54.6	44.3	42.7
	2020	60.8	57.0	45.7	45.9
	2021	61.1	57.2	45.6	46.1

Dataset	Year	Low Dominance		High Dominance	
		Canada	USA	Canada	USA
TUSC Country	2015	38.6	38.4	53.1	45.7
	2016	39.2	39.1	55.1	48.5
	2017	42.3	42.2	58.4	52.6
	2018	44.1	44.6	61.5	56.0
	2019	43.8	44.1	61.4	55.6
	2015--2019	41.6	41.7	57.9	51.7
	2020	44.4	44.3	61.0	56.1
	2021	43.8	43.4	61.5	56.5

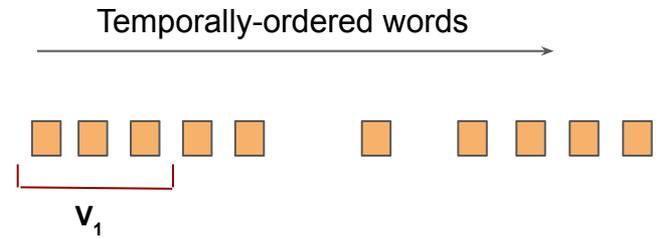
3. Tweet Emotion Dynamics

- Emotion Dynamics (ED): Psychology framework (Kuppens et al. 2010; Hollenstein 2015)
- Utterance Emotion Dynamics (UED): Computational framework to capture ED metrics using utterances (Hipson and Mohammad 2021)
- Tweet Emotion Dynamics (TED): UED applied to tweets
- Motivation: explore and benchmark *individual tweeter* behaviour and emotion word usage metrics across time
- We use TUSC-City, and consider users with at least 100 tweets in a year

TED Metrics

For each tweeter:

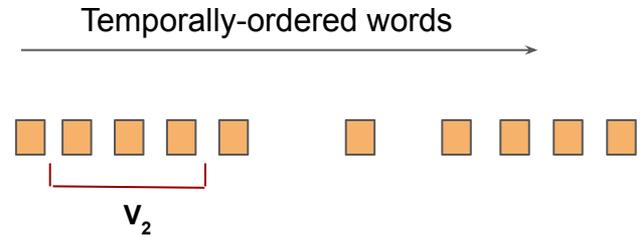
- Create a temporally-ordered sequence of words from their utterances
- Apply a sliding window of word-emotion score averages



TED Metrics

For each tweeter:

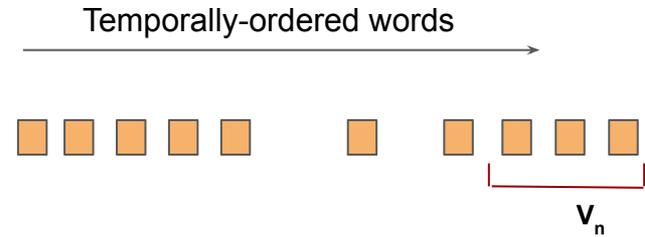
- Create a temporally-ordered sequence of words from their utterances
- Apply a sliding window of word-emotion score averages



TED Metrics

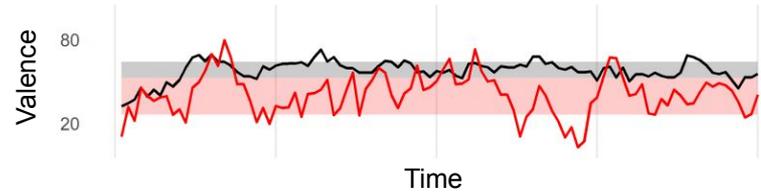
For each tweeter:

- Create a temporally-ordered sequence of words from their utterances
- Apply a sliding window of word-emotion score averages



TED Metrics

- Emotion word density
 - Average emotional state
 - Mean of V/A/D values of each window
- Variability
 - Standard deviation of window V/A/D values
- Home Base
 - Subspace of high-probability emotion states
 - Mean \pm Variability

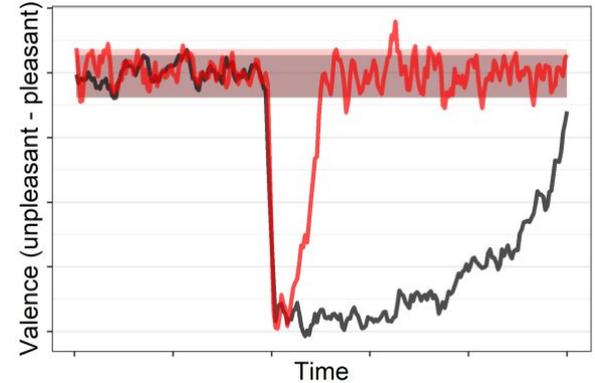


TED Metrics

- Displacement metrics:
 - Displacement count, displacement length, peak distance
 - Rise and recovery rates

We distinguish between:

- When the emotion state is lower than the home base range:
 - Home-Low rate (Hm-Lo)
 - Low-Home rate (Lo-Hm)
- When the emotion state is higher than the home base range:
 - Home-High rate (Hm-Hi)
 - High-Home rate (Hi-Hm)

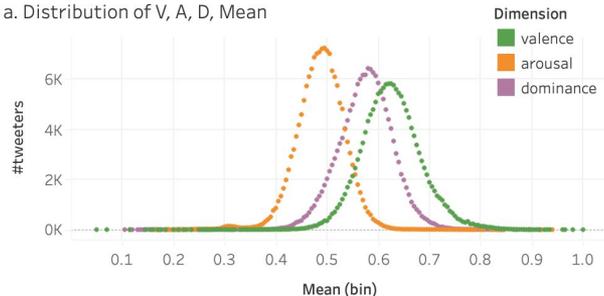


TED: Research Questions

- What is the usual range and distribution of various metrics of TED for American and Canadian tweeters?
- Are there notable differences in the distributions of TED metrics across American and Canadian tweeters?
- Are there notable differences in the distributions of TED metrics across 2020 and 2021?

V-A-D means follow a normal distribution

a. Distribution of V, A, D, Mean



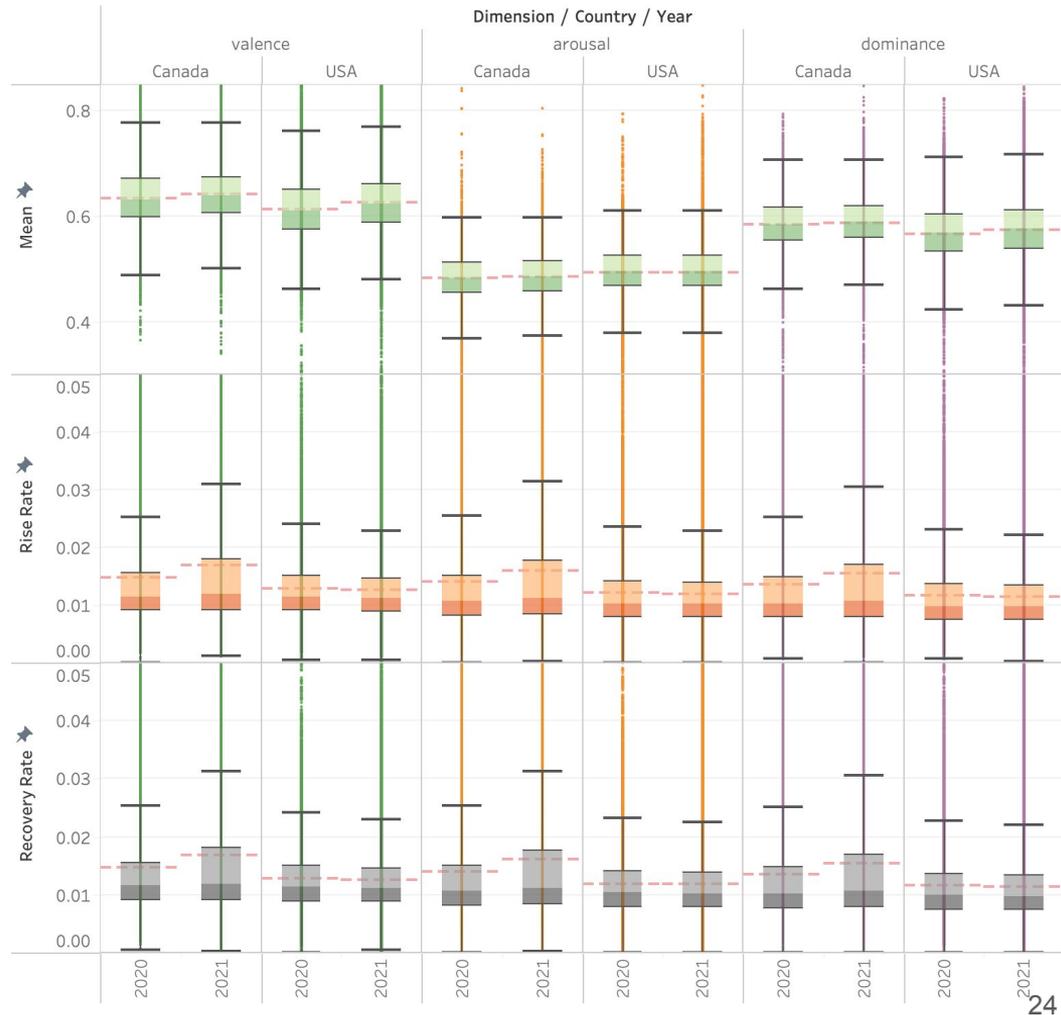
Box-and-whisker Plots

Mean:

- Mean valence is lower in 2020
- Canadian tweeters have:
 - Higher median valence and dominance
 - Lower median arousal

Rise and recovery rates:

- Medians are roughly equal
- Large third-quarter range for Canada 2021
 - Tweeters were quick to jump in and out of home base



Comparing Hm–Hi and Hm–Lo Rates

Arousal

<i>Hm-Hi Rate</i>	2020	0.0129	0.0130
	2021	0.0125	0.0129
<i>Hi-Hm Rate</i>	2020	0.0127	0.0130
	2021	0.0126	0.0128
<i>Hm-Lo Rate</i>	2020	0.0121	0.0127
	2021	0.0118	0.0123
<i>Lo-Hm Rate</i>	2020	0.0120	0.0125
	2021	0.0117	0.0121

Tweeters are slower to rise to and fall from negative arousal – more passive – states than from higher arousal states.

Valence

<i>Hm-Hi Rate</i>	2020	0.0118	0.0129
	2021	0.0113	0.0121
<i>Hi-Hm Rate</i>	2020	0.0118	0.0129
	2021	0.0115	0.0122
<i>Hm-Lo Rate</i>	2020	0.0143	0.0149
	2021	0.0140	0.0145
<i>Lo-Hm Rate</i>	2020	0.0141	0.0148
	2021	0.0139	0.0144

Tweeters are quicker to rise to and fall from negative valence states than from positive displacements.

Dominance

<i>Hm-Hi Rate</i>	2020	0.0114	0.0118
	2021	0.0111	0.0115
<i>Hi-Hm Rate</i>	2020	0.0114	0.0119
	2021	0.0111	0.0114
<i>Hm-Lo Rate</i>	2020	0.0127	0.0129
	2021	0.0124	0.0126
<i>Lo-Hm Rate</i>	2020	0.0126	0.0128
	2021	0.0123	0.0125

Tweeters are quicker to rise to and fall from negative dominance states than from higher dominance displacements.

TED: City-as-a-speaker

Consider all tweets from a particular city as being from a single “tweeter”.

Highlights:

- Canadian cities again have higher valence, lower arousal, higher dominance
- Valence values rise from 2020 to 2021
- Quebec and Windsor have the higher Canadian arousal rates
- Nashville, Seattle, San Jose, San Francisco have dominance and arousal rates comparable to Canadian cities

These metrics are potentially useful for city-level analyses of emotion word usage and change.

Country	City	valence	
		2020	2021
Canada	Brampton	0.634	0.642
	Calgary	0.636	0.647
	Edmonton	0.643	0.649
	Etobicoke	0.637	0.643
	Halifax	0.649	0.655
	Hamilton	0.644	0.651
	Laval	0.638	0.645
	London	0.649	0.656
USA	Austin	0.627	0.639
	Boston	0.625	0.641
	Charlotte	0.627	0.638
	Chicago	0.617	0.629
	Columbus	0.625	0.637
	Dallas	0.617	0.630
	Denver	0.624	0.636
	Detroit	0.614	0.626
	ElPaso	0.629	0.638

Summary

Our findings:

- Metrics to quantify emotion word usage in tweets
- Significant differences in emotion word usage between Canada and the USA
- Changes in emotion word usage in 2021 when compared to previous years

Future exploration:

- Word-level analysis with treemaps:
 - What words are driving these changes in emotion word usage?
 - Are they related to specific real-world events (COVID-19 pandemic, Black Lives Matter movement, US Presidential elections)?
- TED metrics of individual users
 - Collaboration with UNC Affective Science Lab: correlation with mental, emotional, physiological health of populations?

Ethical Considerations

- Automatic Emotion Recognition (AER)
 - emotions conveyed by people via the language they use
 - Not equivalent to people's "true" inner emotions
- Comparisons based on aggregate patterns
 - Not of specific individuals
- Effective use of emotion lexicons: Mohammad 2020 (<https://arxiv.org/abs/2011.03492>)
- Ethics sheet for AER: Mohammad 2022 (<https://arxiv.org/abs/2109.08256>)

Resources

- Paper: <https://arxiv.org/abs/2204.04862>
- Data and Code: <https://github.com/priya22/emotiondynamics>
 - Python scripts to calculate UED metrics for any temporally-ordered text data
- Poster presentation at LREC 2022: Thursday June 23 (9:30-10:50)
- Contact:
 - vkpriya@cs.toronto.edu
 - saif.mohammad@nrc-cnrc.gc.ca