Cluster analysis of eating-disorder-related content with hashtag #edrec0very on TikTok

Amar Abboodi, Havannah Tung

Motivation: Background

International Journal of
EATING DISORDERS

ORIGINAL ARTICLE

"This is just how I cope": An inductive thematic analysis of eating disorder recovery content created and shared on TikTok using #EDrecovery
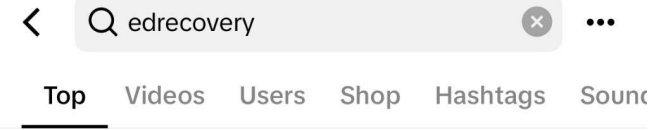
Shannon S. C. Herrick MA ✉, Laura Hallward MA, Lindsay R. Duncan PhD

Motivation: Background (cont'd)

- Eating disorders (ED)
- #edrecovery vs. #edrec0very ('zero')

Motivation: Goals and Objectives

- Thematic analysis of TikTok with #edrec0very
- Hashtag analysis to find other popular "turn-around" hashtags used for ED-recovery-related content

Method: Datasets

- Apify
  (https://apify.com/)
- Scrape all TikTok
  videos showing up as
  #edrec0very search
  results
  - Popularity features (e.g.
    likes, shares,
    comments, fans,
    followings)
  - Hashtags

|	authorMeta/avatar	authorMeta/bioLink	authorMeta/digg	authorMeta/fans	authorMeta/following	authorMeta/hea
1	https://p16-sign-useast2a.tiktokcdn.com/tos-useas...	NA	23800	1765	344	18
2	https://p16-common-sign-va.tiktokcdn-us.com/tos-...	NA	136	545	17	6
3	https://p16-common-sign-va.tiktokcdn-us.com/tos-...	NA	8693	38100	937	690
4	https://p16-common-sign-va.tiktokcdn-us.com/tos-...	NA	8633	24400	457	610
5	https://p19-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	2168	2676	494	15
6	https://p19-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	1034	3440	138	12
7	https://p16-common-sign-sg.tiktokcdn-us.com/tos-...	NA	10300	758	98	8
8	https://p16-common-sign-va.tiktokcdn-us.com/tos-...	NA	1083	618	522	
9	https://p19-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	126	236	41	3
10	https://p16-sign-useast2a.tiktokcdn.com/tos-useas...	NA	1036	3508	13	76
11	https://p19-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	261	207	44	
12	https://p16-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	25	57	7	
13	https://p16-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	657	420	150	2
14	https://p16-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	96900	1330	143	94
15	https://p19-pu-sign-useast8.tiktokcdn-us.com/tos-...	NA	7487	432	138	2
16	https://p16-sign.tiktokcdn-us.com/tos-useast5-avt-...	NA	6162	2527	63	68
17	https://p19-common-sign-sg.tiktokcdn-us.com/tos-...	NA	1866	17600	224	93
18	https://p16-sign-va.tiktokcdn.com/tos-maliva-avt-...	NA	12200	10500	14	50

Methods: Datasets (con'd)

- Scrape hashtags of videos from content creators of which the videos showed up in search result

	username	video_id	video_url	hashtags
1	abbonk0	7.485527e+18	*NA*	['#binge', '#ed', '#food', '#gym', '#gymtok', '#gymgirl', ...
2	abbonk0	7.485923e+18	*NA*	['#binge', '#bingeed', '#food', '#foodnoise', '#advice', '...
3	abbonk0	7.491109e+18	*NA*	['#gym', '#motivation', '#spreak', '#springbreak', '#hea...
4	abbonk0	7.490738e+18	*NA*	['#fyp', '#foryou', '#food', '#binge', '#relatable', '#gymt...
5	abbonk0	7.490345e+18	*NA*	['#recovery', '#ed', '#foryou', '#fyp', '#edrec0very', '#ha...
6	abbonk0	7.490043e+18	*NA*	['#food', '#foodnoise', '#fyp', '#foryou', '#gym', '#gymt...
7	abbonk0	7.489626e+18	*NA*	['#binge', '#bingefree', '#brainoverbinge', '#food', '#fo...
8	abbonk0	7.488905e+18	*NA*	['#nutrition', '#nutritiontips', '#binge', '#food', '#fyp', '...
9	abbonk0	7.488548e+18	*NA*	['#relatable', '#bigback', '#fat', '#eat', '#foryou', '#fyp', ...
10	abbonk0	7.488532e+18	*NA*	[]
11	abbonk0	7.488381e+18	*NA*	['#relatable', '#foryou', '#binge', '#calories', '#caloried...
12	abbonk0	7.488200e+18	*NA*	['#relatable', '#foryou', '#fyp', '#binge', '#bingefree', '#...
13	abbonk0	7.488166e+18	*NA*	[]
14	abbonk0	7.488163e+18	*NA*	['#bingefree', '#advice', '#foryou', '#fyp', '#gym', '#gy...
15	abbonk0	7.487937e+18	*NA*	['#oats', '#highprotein', '#protein', '#breakfast', '#recip...
16	abbonk0	7.487806e+18	*NA*	['#fy', '#fyp', '#relatable', '#food', '#foodnoise', '#binge...
17	abbonk0	7.487801e+18	*NA*	['#bodytransformation', '#food', '#gym', '#gymtok', '#r...
18	abbonk0	7.487688e+18	*NA*	['#energydrink', '#caffeine', '#caffeineaddiction', '#foo...
19	abbonk0	7.487416e+18	*NA*	['#fyp', '#relatable', '#food', '#foodnoise', '#gym', '#for...

Method: Hashtag Pre-Analysis

1. Load the TikTok Dataset
2. Preprocess Hashtags
3. Filter for Posts Containing #edrec0very
4. Count Co-occurring Hashtags
5. Plot the Results

Load the TikTok Dataset

↓

Preprocess Hashtags

↓

Filter for Posts Containing *#edrecovery*

↓

Count Co-occurring Hashtags

↓

Plot the Results
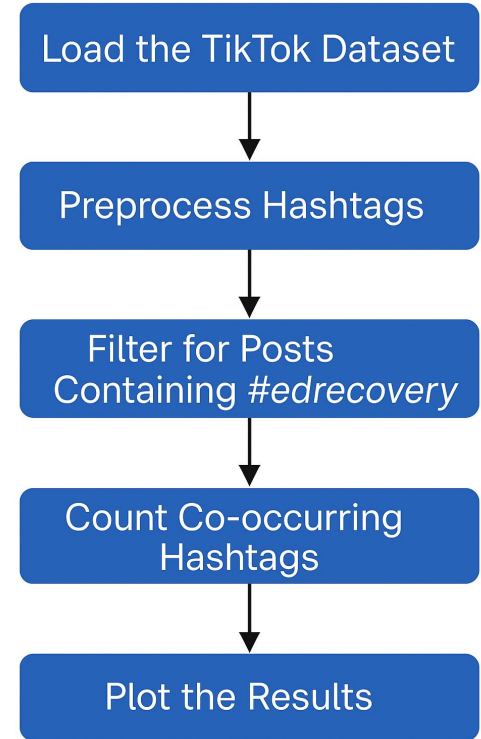
Method: clustering

1. Prepare Text Data
2. Create Document-Term Matrix (DTM)
3. Apply TF-IDF Weighting
4. Clean and Normalize
5. Clustering
6. Dimensionality Reduction
7. Visualization

Method: Identify the Clusters

1. Extract valid documents from the TF-IDF matrix (m) — i.e., those with at least one non-zero term.
2. Filter out empty rows (if any) to focus only on meaningful content.
3. Cluster the documents using K-means based on their TF-IDF vectors.
4. Assign each video a cluster label by mapping doc_id to its cluster.
5. Join this cluster info back to the original dataset for further analysis.
6. Find the top 5 tags per cluster by unnesting the tags and counting their frequency within each cluster.
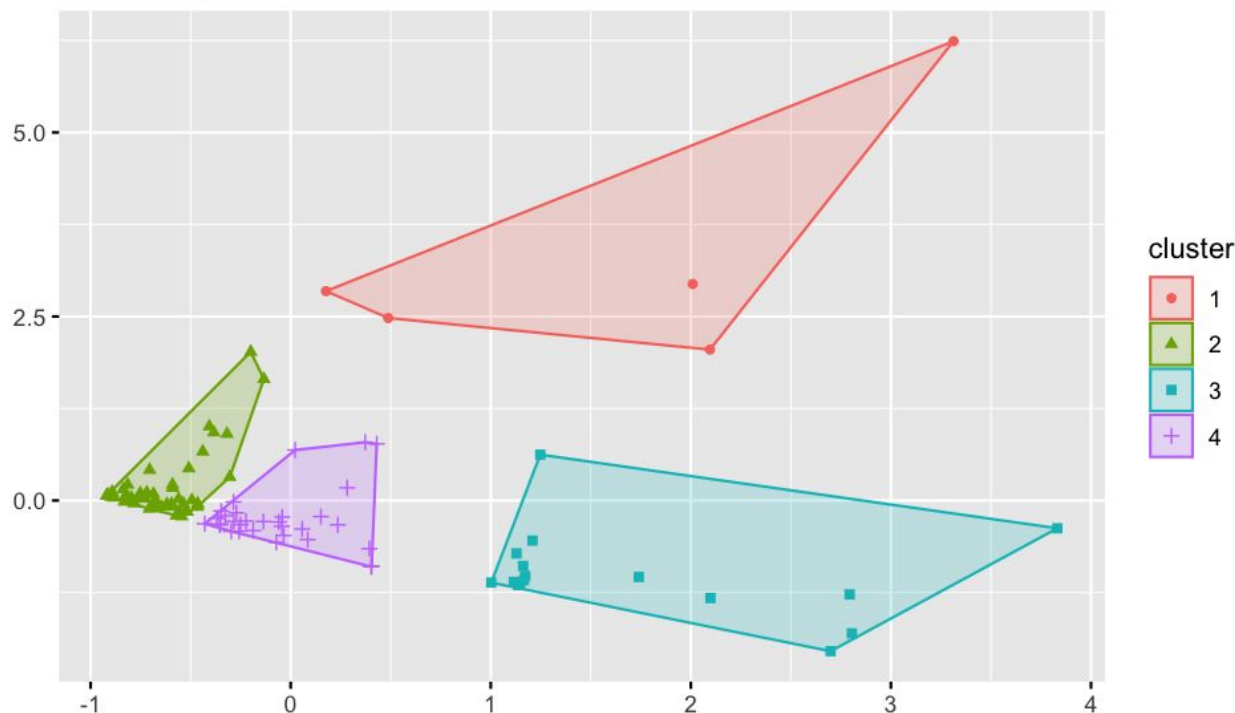
Results from Search Result Dataset

Top 20 hashtags used in #edrec0very search results
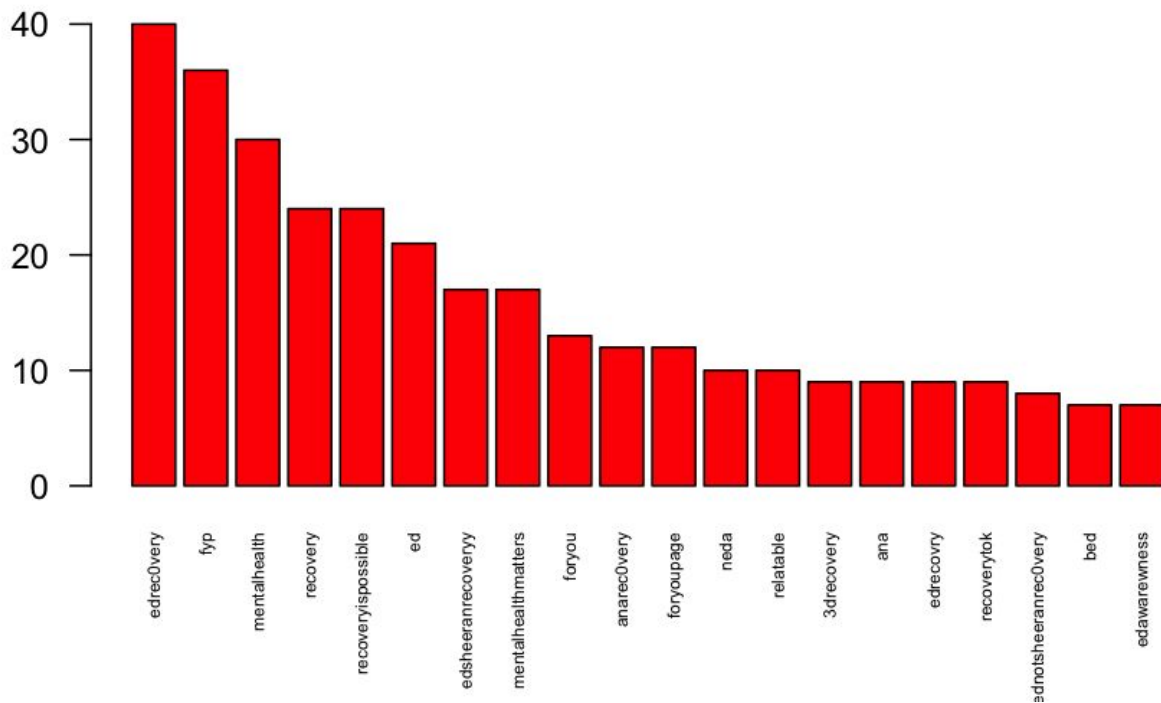
Results from Search Result Dataset (cont'd)

Popularity cluster result

Cluster 1: #foryou, #ana, #bed, #diet, #dietitian #ed #edd #eddtt #edsheeran

Cluster 2: #fyp, #mentalhealth, #recovery, #ed, #recoveryispossible, #edsheeranrecoveryy, #relatable

Cluster 3: #milosaysyoumatter, #miloisproudofyou, #edrecovery, #recovery, #recoverytok, #edsheeranrecoveryy, #fyp, #mentalhealth

Cluster 4: #mentalhealth, #recoveryispossible, #fyp, #mentalhealthmatters, #anarec0very, #ed, #3drecovery

Results from Search Result Dataset (cont'd)

Hashtag cluster result
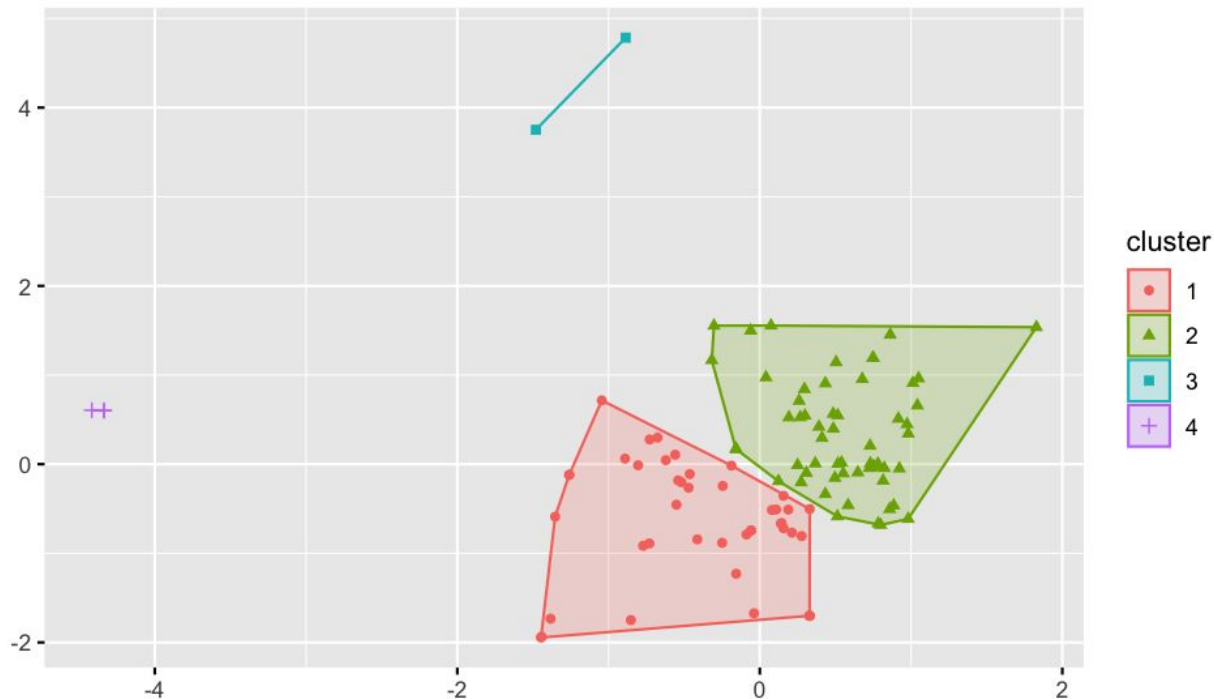
Cluster1: #mentalhealth, #fyp, #recoveryispossible, #mentalhealthmatters, #edsheeranrecoveryy, #anarec0very
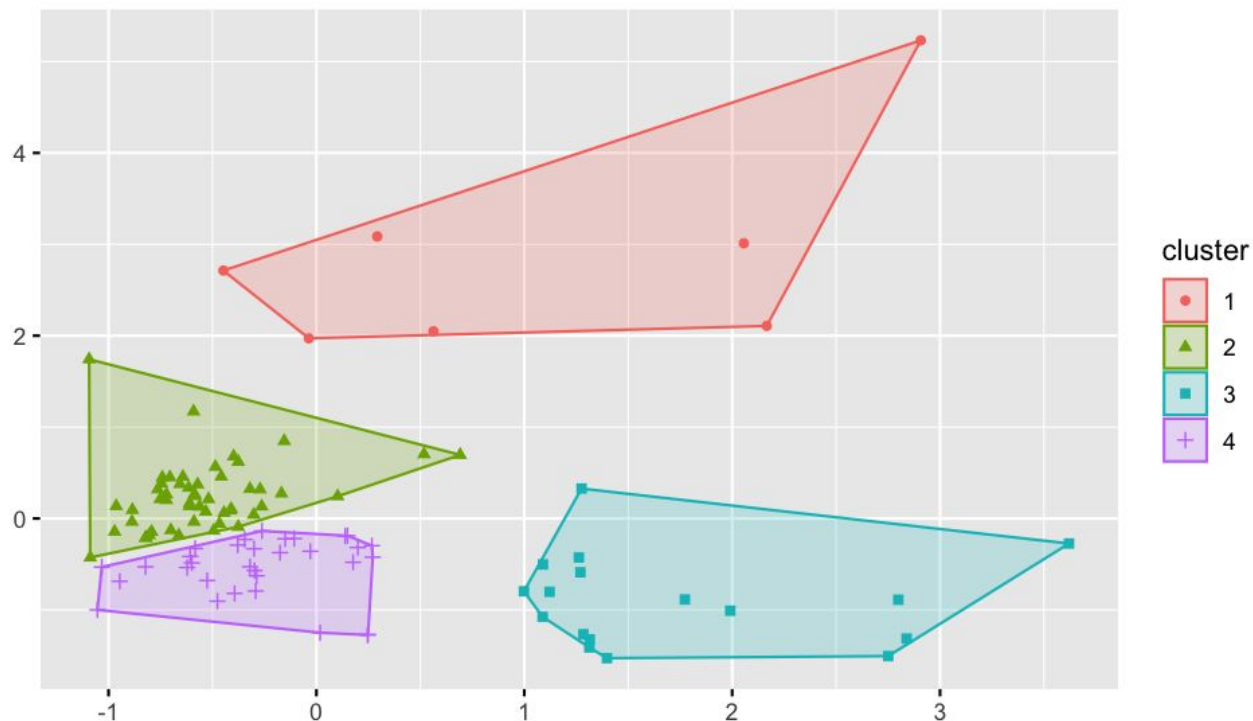
Cluster2: #fyp, #ed, #recovery, #mentalhealth, #recoveryispossible, #ana, #foryou, #recoverytok

Cluster3: #gymtok, #binge, #bingefree, #coloriedeficit, #fitness

Cluster4: #3drecovery, #anarec0very, #anarec0vry, #eupd, #eupdrecovery

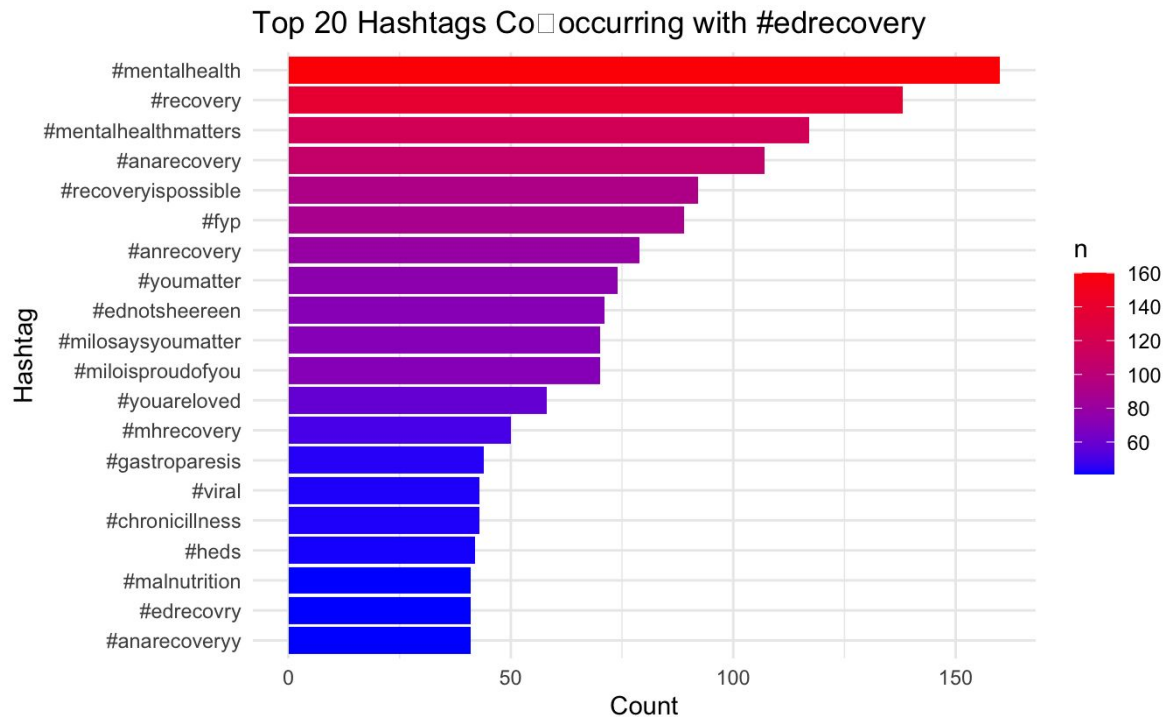# Results from Search Result Dataset (cont'd)

All feature cluster result

Cluster1: #foryou, #foryoupage, #fyp, #recoveryispossible, #weightloss, #3drecovery, #ana, #bed, #binge

Cluster2: #fyp, #ed, #mentalhealth, #recovery, #recoveryispossible, #relatable, #edsheeranrecoveryy

Cluster3: #milosaysyoumatter, #miloisproudofyou, #edrecovery, #recovery, #recoverytok, #edsheeranrecoveryy

Cluster4: #entalhealth, #mentalhealthmatters, #anarec0very, #recovery, #recoveryispossible, #edsheeanrecoveryy, #fyp

Results from User Profile Dataset

Top 20 Hashtags Co□occurring with #edrecovery

Results from User Profile Dataset (cont'd)

PCA of Videos Colored by K☐means Cluster (k=5)
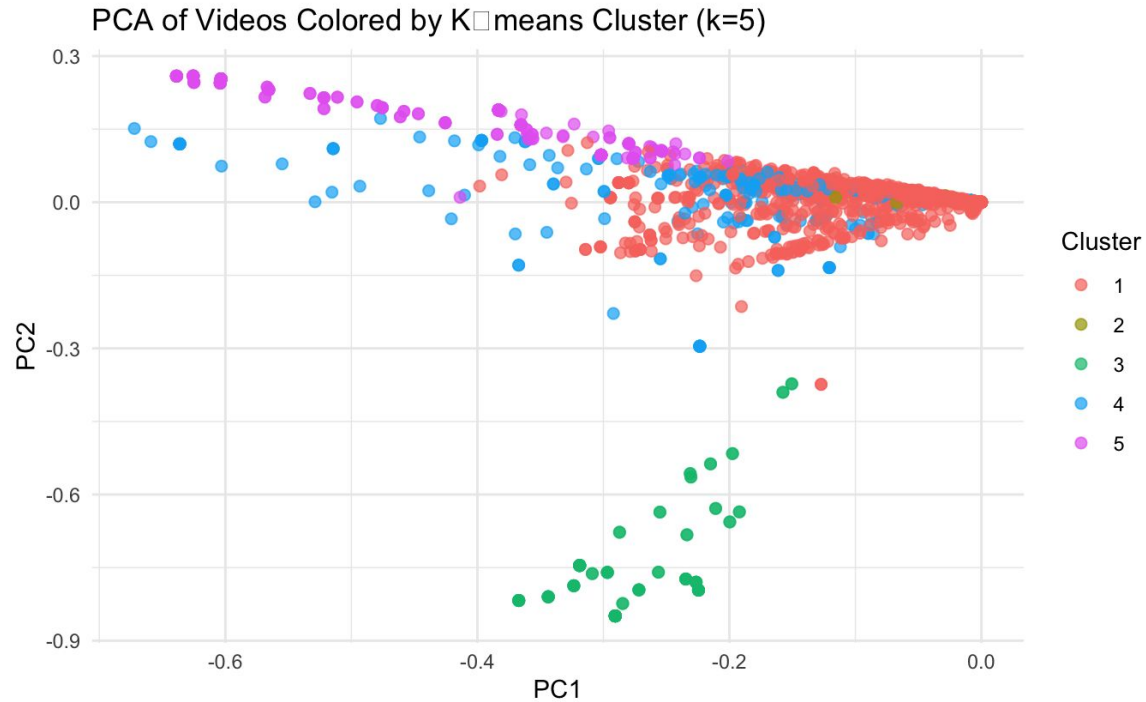
Results from User Profile Dataset (cont'd)

- Cluster 1: #vent — Emotional Expression & Personal Struggles
- Cluster 2: #real — Authenticity and Anti-Perfection
- Cluster 3: #london — Aesthetic & Cultural Identity
- Cluster 4: #fyp — Mainstream Recovery & Mental Health
- Cluster 5: #collegelife — Student & Nostalgia Niche

Discussion

- Popular "turn-around" hashtags for #edrecovery:
  - #3drecovery, #anarecovery, #anarec0very, #edsheeranrecovery
- What else could be done?
  - Mapping clustering results to user's metadata?

Limitations and future directions

- Smaller and biased datasets
  - More recent and popular posts
  - Future directions: research API

Limitations and future directions (cont'd)

- Lack of similar data mining literature on the topic to compare result
  - Future directions: content analysis by human researchers

Reference

- Herrick SSC, Hallward L, Duncan LR. "This is just how I cope": An inductive thematic analysis of eating disorder recovery content created and shared on TikTok using #EDrecovery. Int J Eat Disord. 2021 Apr;54(4):516-526. doi: 10.1002/eat.23463. Epub 2020 Dec 31. PMID: 33382136.
- Au ES, Cosh SM. Social media and eating disorder recovery: An exploration of Instagram recovery community users and their reasons for engagement. Eat Behav. 2022 Aug;46:101651. doi: 10.1016/j.eatbeh.2022.101651. Epub 2022 Jun 22. PMID: 35760017.