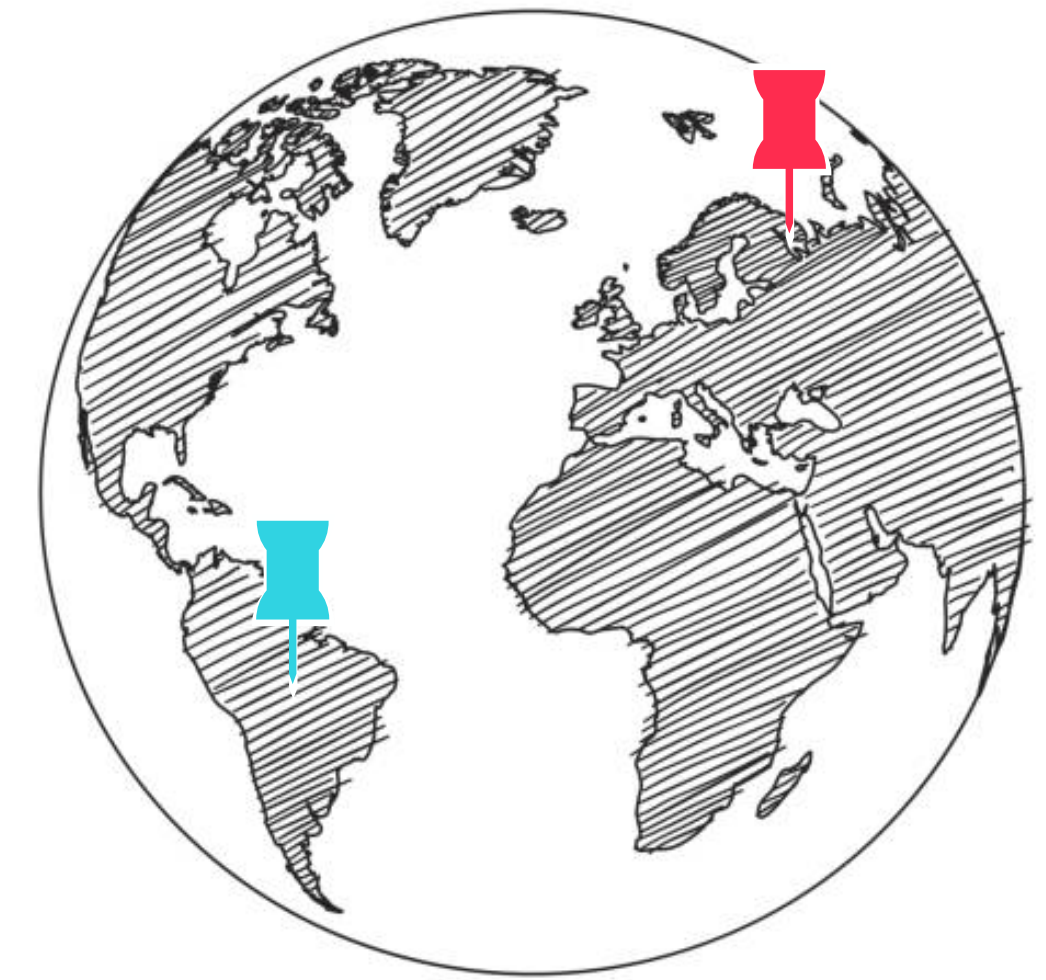




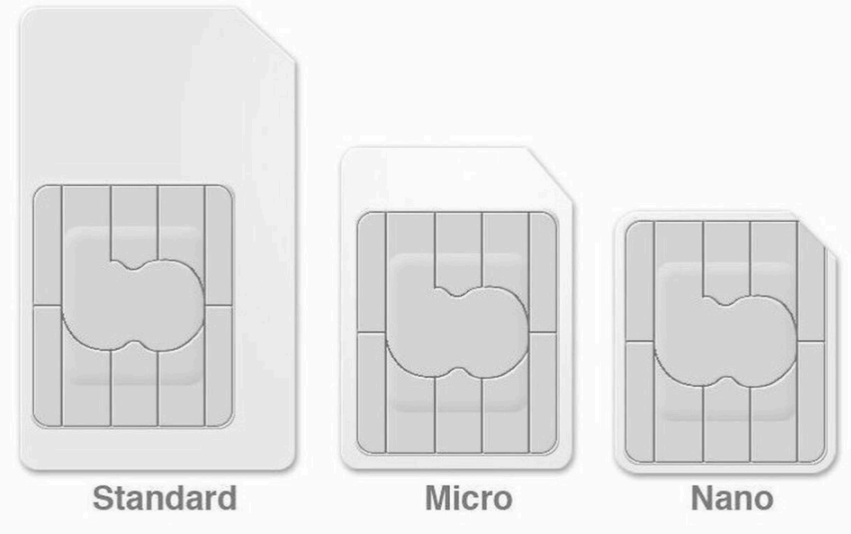
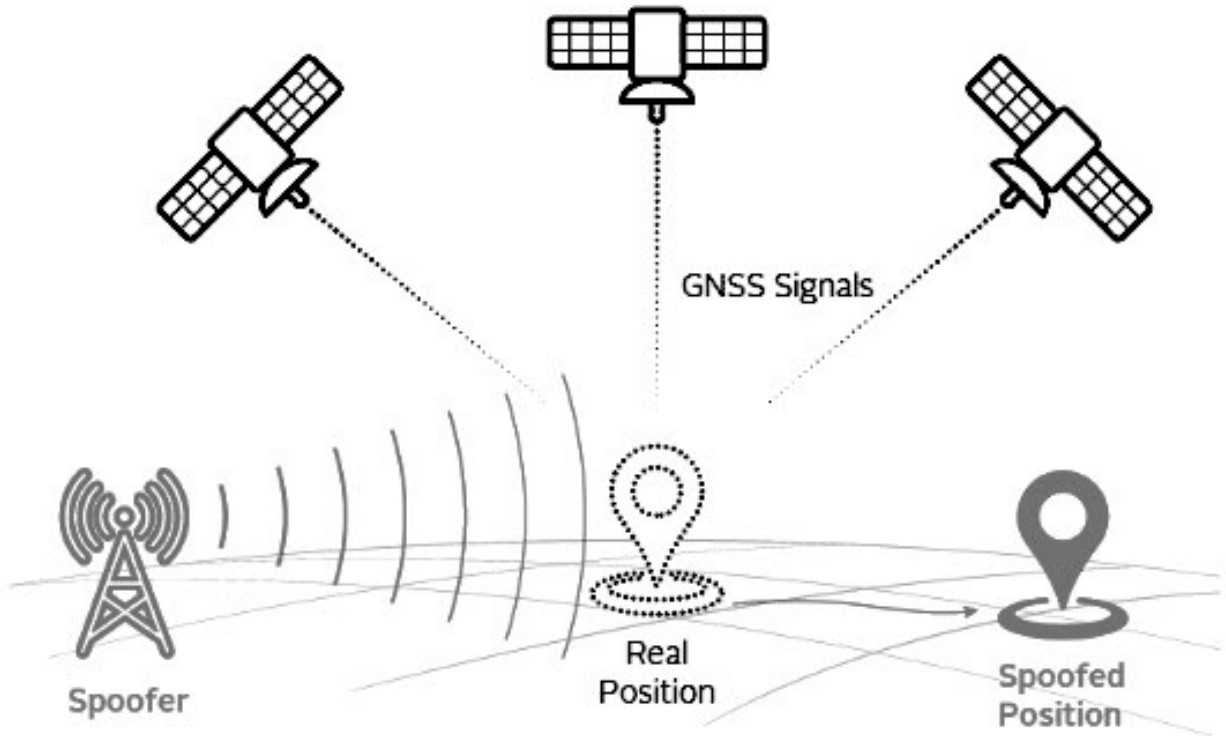
TIKTOK TREND ANALYSIS

ROWAN SCHAEFER, 2022

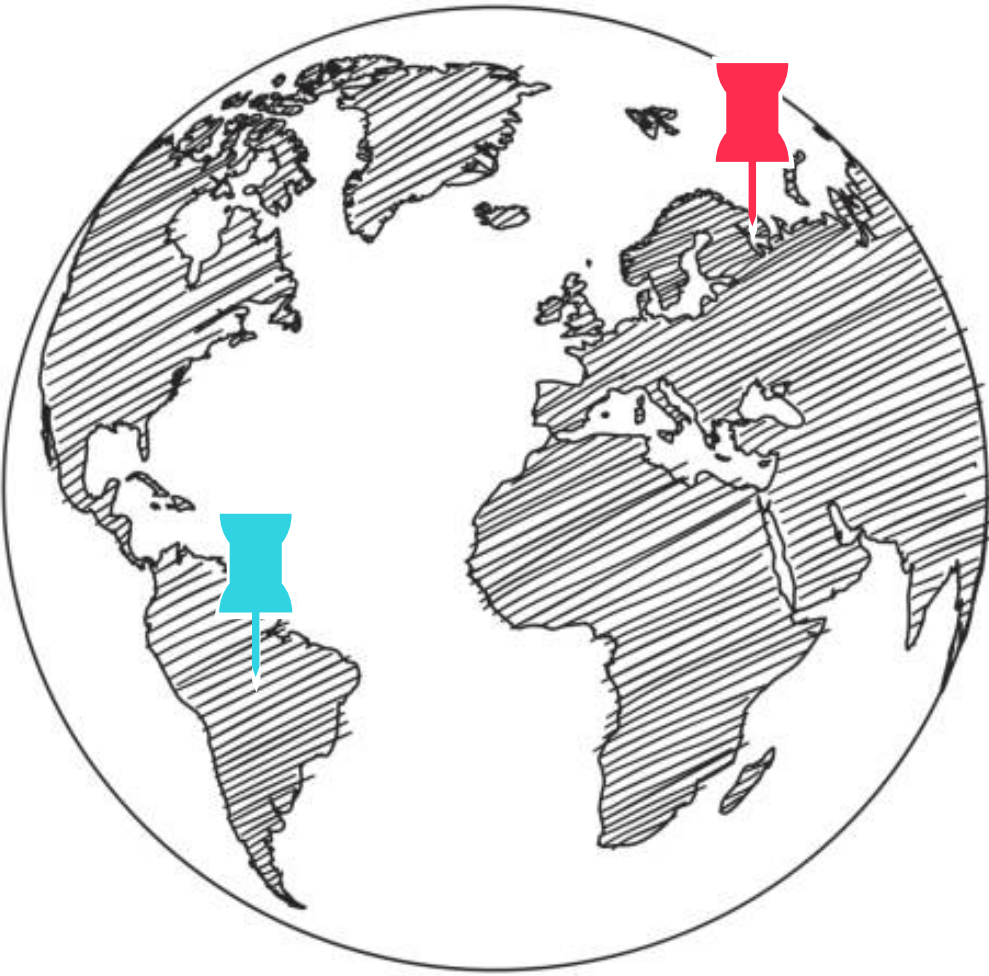


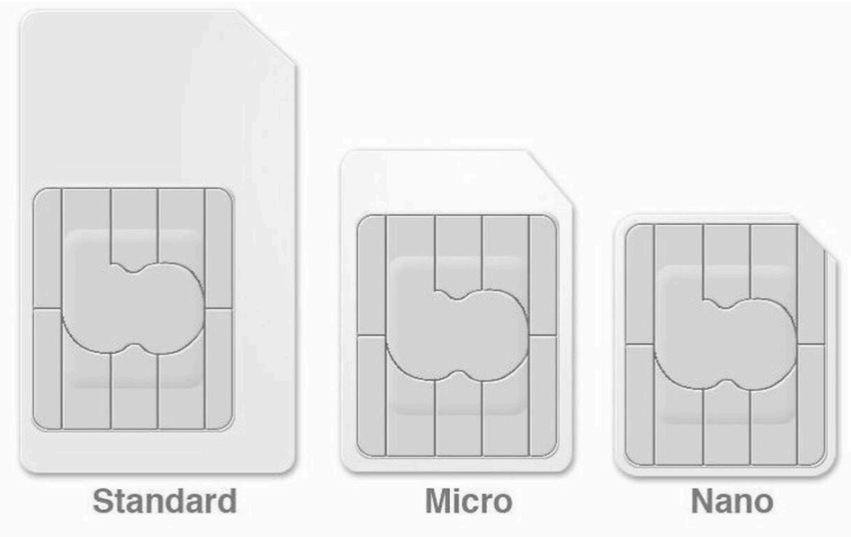
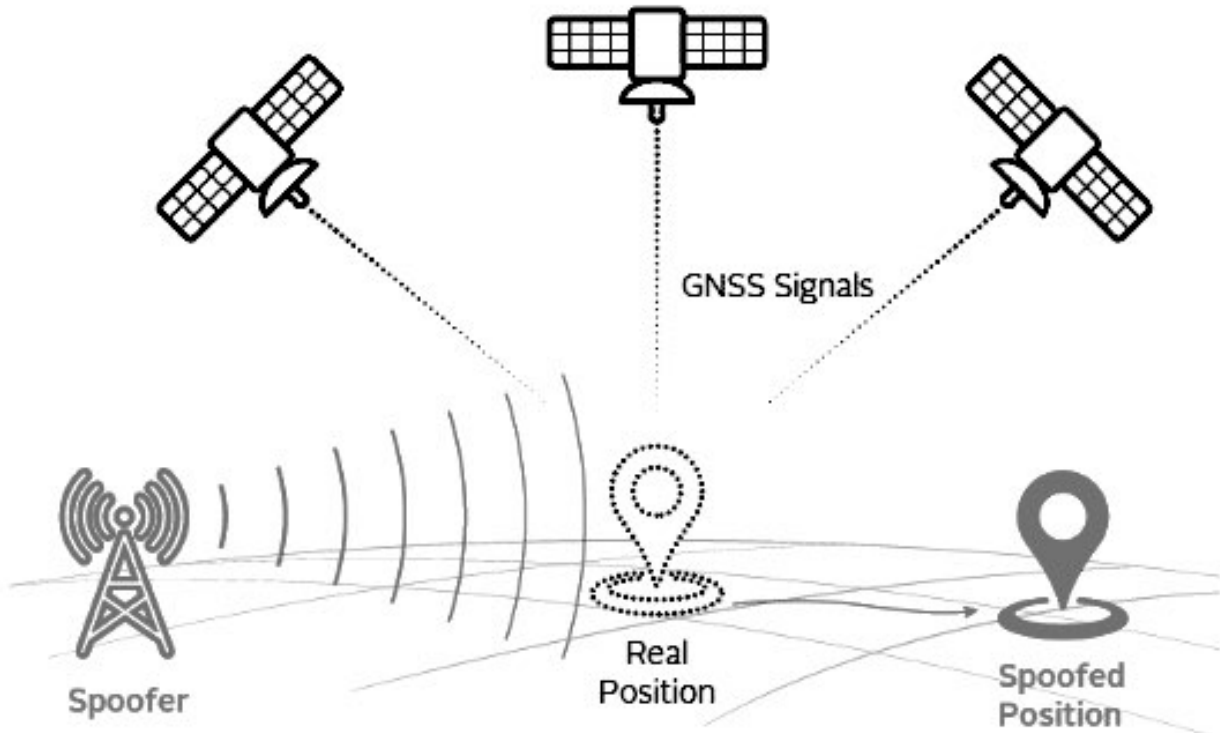
TIKTOK TREND ANALYSIS

ROWAN SCHAEFER, 2022

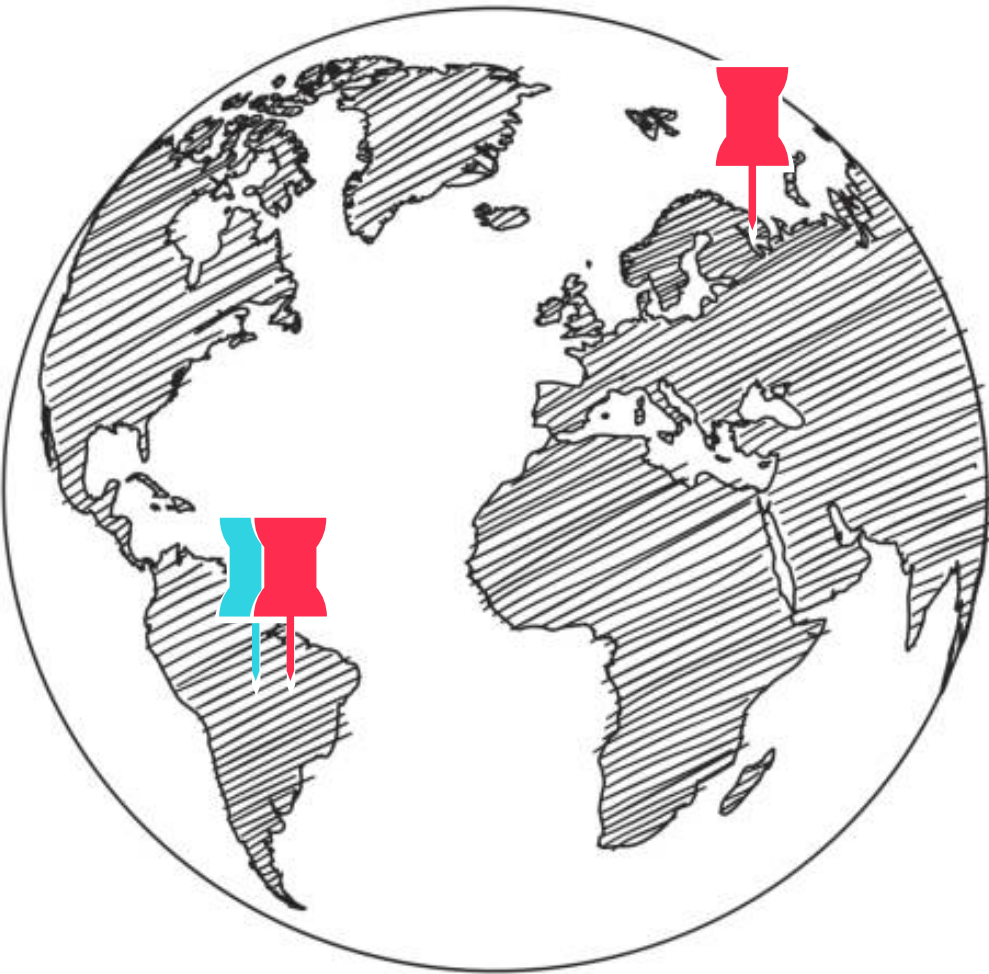


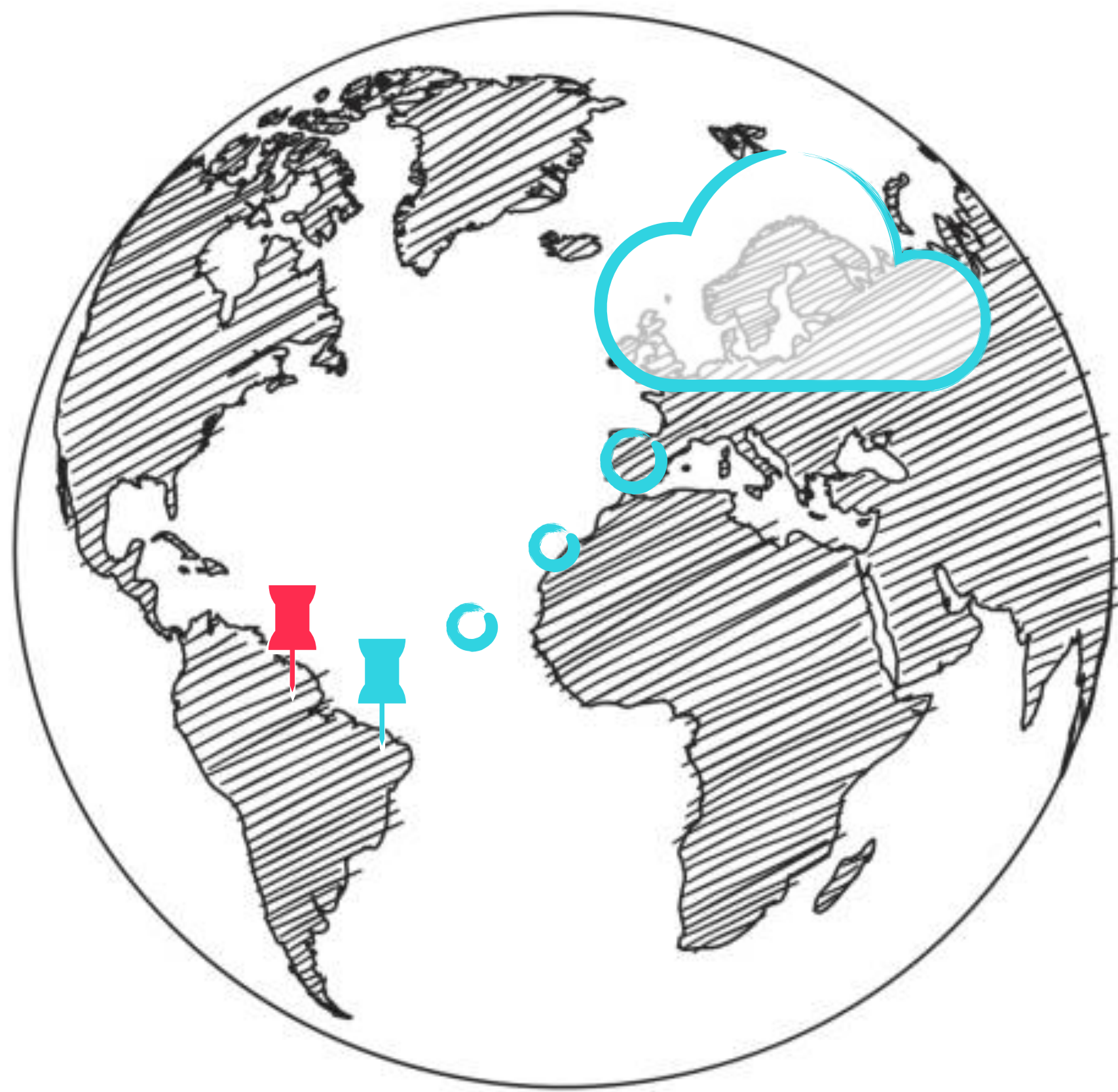
Status	
Status	<div><div></div> Connected</div>
DHCP Server	10.10.10.1
IP Address (vlan1)	10.10.10.16/24,fe80:0:....
Gateway	10.10.10.1
Primary DNS	8.8.8.8
Secondary DNS	
External IP Address	122.172.108.135
No. Clients	0
Uptime	3h 5m





Status	
Status	<div><div></div> Connected</div>
DHCP Server	10.10.10.1
IP Address (vlan1)	10.10.10.16/24,fe80:0:...
Gateway	10.10.10.1
Primary DNS	8.8.8.8
Secondary DNS	
External IP Address	122.172.108.135
No. Clients	0
Uptime	3h 5m







P R O B L E M S T A T E M E N T

For TikTok creators who don't have the time to manage multiple accounts, devices, or SIM cards, a solution is needed to better analyze trends in non-US countries. Currently there are no web apps that allow you to cross check trending content in different countries, and tricking TikTok into thinking you live somewhere else is difficult.



G O A L

My primary goal was to build an app that could provide statistics on ALL countries rather than just the USA or select larger countries.

C A V E A T

There were some reasons why this hadn't been done by anyone else yet

這, 個



Source: derry

R E S E A R C H Q U E S T I O N S



What are the top trending sounds/hashtags in [country 1] ?

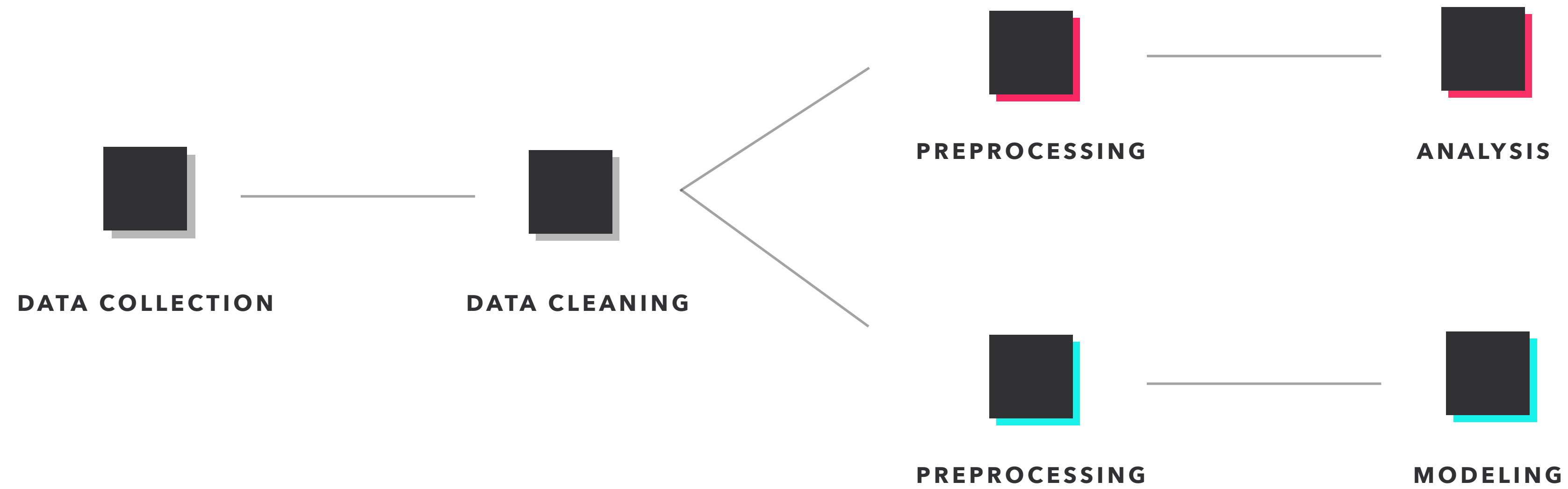


What is the intersection of top trending sounds/hashtags in [country 1] , [country 2] , and [country 3] ?



Bonus: A recommender model for FYP videos.

P R O C E S S



D A T A C O L L E C T I O N



Tiktok-All-In-One from FastApi, ^{*} ~10 rows per minute



74,197 rows total for 119 countries (A-K, with some extras.)

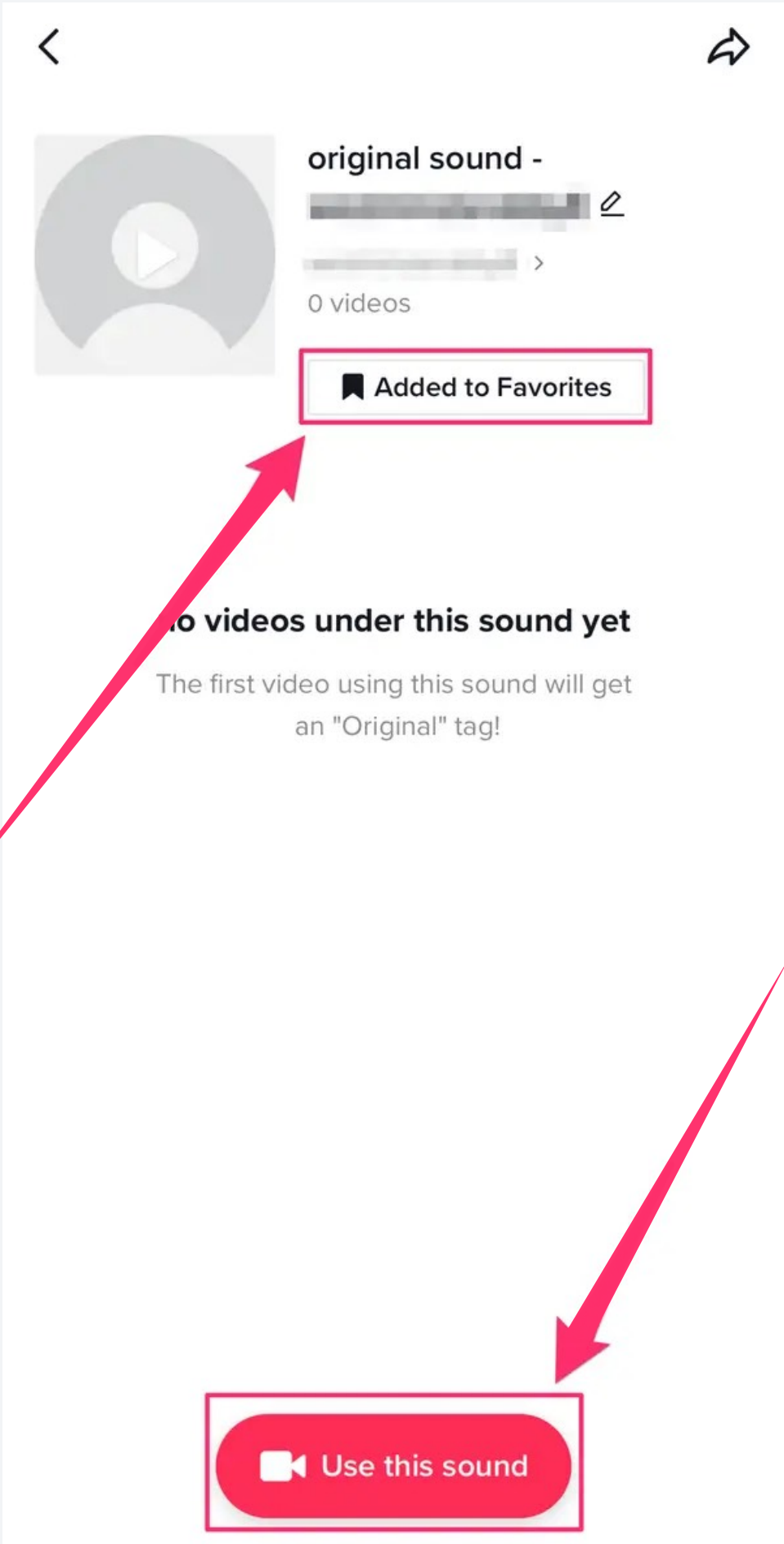


Future data collection will need to be done with a virtual machine.

CLEANING & PREPROCESSING

11

Want to go viral? Sounds are more important than songs!



Combined matched_pgc_sound authors and artists with sound authors and artists

CLEANING & PREPROCESSING

1 2

Preserved all hashtags,
didn't lemmatise

	hashtag	count
0	#fyp	18295.0
1	#foryou	9612.0
2	#viral	8365.0
3	#fypシ	6375.0
4	#foryoupage	5542.0

- **CountVectorizer**
- **Custom stopwords**

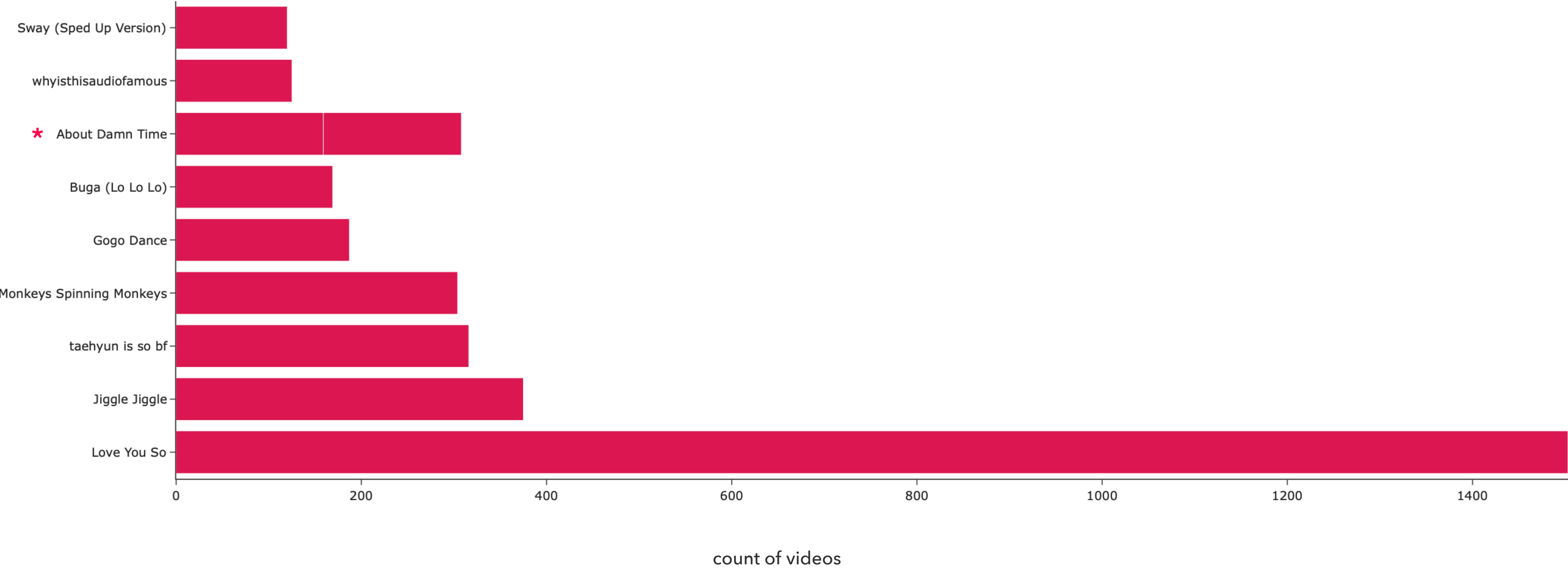
Custom stop words: common hashtags
(#viral, #fyp, #abcde, etc)

D A T A A N A L Y S I S

TOP 10 TRENDING HASHTAGS - ENGLISH SPEAKING COUNTRIES

	word	count
0	#funny	100
1	#trende	81
2	#johnnydepp	55
3	#comedy	49
4	#jamaicantiktok	45
5	#duet	43
6	#jamaicatiktok	41
7	#meme	40
8	#amberheard	37
9	#justiceforjohnnydepp	34

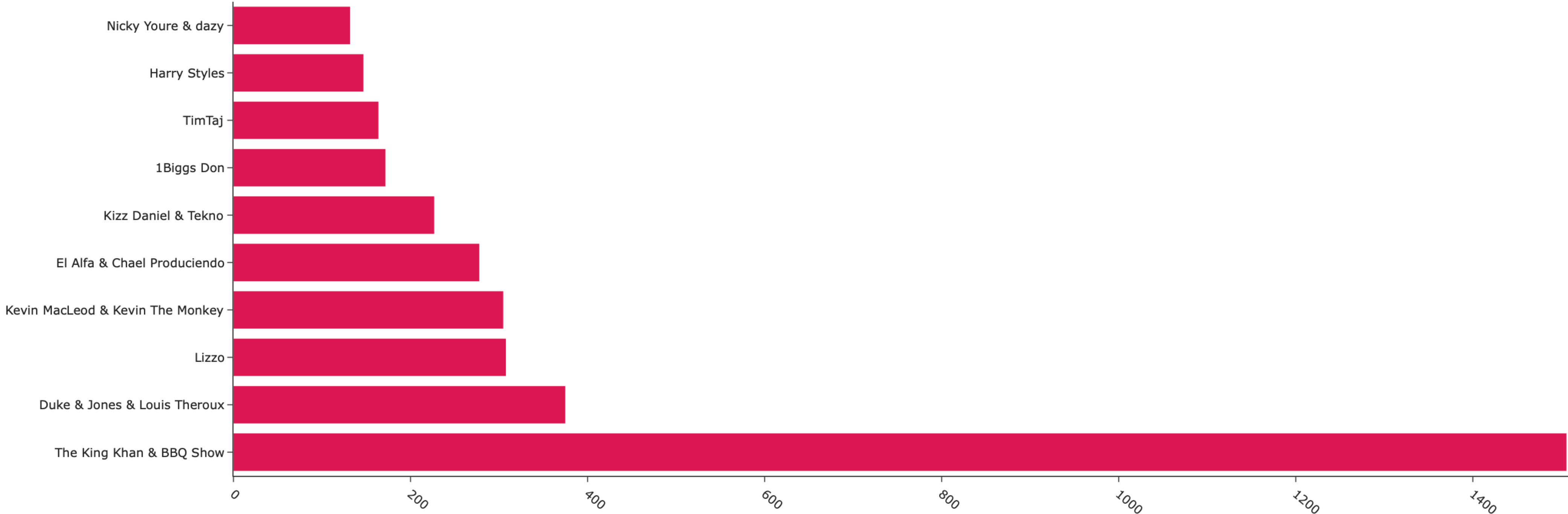
TOP 10 TRENDING SOUNDS - WORLDWIDE

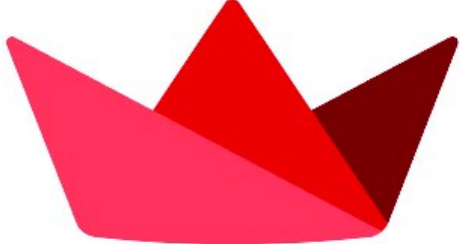


TOP 10 TRENDING ARTISTS - WORLDWIDE

1 5

trending artist





:)

P R E P R O C E S S I N G F O R M O D E L I N G

Excluded hashtags from lemmatization

Preserved # and @ sign

	hashtag	count
0	#fyp	18295.0
1	#foryou	9612.0
2	#viral	8365.0
3	#fypシ	6375.0
4	#foryoupage	5542.0

- **TfidfVectorizer**
- **SpaCy stopwords, ++**
- **English (ish) only!**

Custom stop words: boring words, gen z slang,
and common hashes (#viral, #fyp, #wyd, etc)

M O D E L I N G



Used SpaCy's en_core_web_sm model



TfidfVectorizer – min_df of **3**, max_df of **.30**



Cosine similarity using vectored video description text

	video id	similarity	desc
0	7094735704051305733	1.000000	Him naah leff mi out 🏆🙌❤️ #perryinchrist #stayinchrist #jamaica #christian #viral #caribbean #jamaicatiktok
1	7098054758610652421	0.547723	#jamaica #foryourpage #foryou #fyp #jamaicatiktok #jamaicatiktok #thevoice #tessannechin
2	7093585398084947206	0.516398	#yardiechallenge @Jazzy Vybz Muzik #djjazzy #jazzyvybzmuzik #dancehall 🇯🇲 @n_asti.a #riflewalk #leggodibird #weddi #jiggy #jamaicatiktok #fyp #jamaica #viral #caribeantiktok #dancehallvibes
3	7092501240898424110	0.516398	When Ella won't listen to me #jamaica #jamaicatiktok #fyp
4	7072816142208978222	0.516398	Beef Patties. #jamaica #jamaicatiktok #proudjamaican #jamaicancooking #jamaicanfood #cookingtiktok
5	7101806922424225029	0.516398	Biggsdon New Chune Preview 🔥🔥 #fy #fyp #foryoupage #fypシ #dancehall #reggae #biggsdon #biggsdonchallenge #jamaica #jamaicatiktok
6	7098095190094777606	0.507093	Dropsy run inna mi family bbz 😔 #fyp #foryoupage #viral #fypシ #trend #text #search #phone #relationship #caribbean #jamaica #swizzledioptimist
7	7091472593290562821	0.447214	Caption dis 😂 #fyp #viral #unknown__iix #derrickmorgan #foryou #jamaicatiktok
8	7081050152735165701	0.447214	Co mi Likkle rass #fyp #foryoupage 🇯🇲 🇧🇲 #tiiaababy #bigbattytiiaa
9	7090963964716125446	0.447214	Lol mi cyaa badda 🤔

	video id	similarity	desc
0	7102146577741073710	0.333333	Practice makes progress folks! #fyp #bridgerton #musiciansoftiktok #cello
1	7100050747319995650	0.333333	#fyp #viral #busker #melbourne #inoxia #americanboy #OscarsAtHome #artistsoftiktok
2	7101832423817628970	0.288675	Take me back #traveltiktok #travel #spain #foryou #fypシ #artistsoftiktok #artistatiktok #boricua
3	7103681011262754054	0.288675	coming out June rd 💖👯 #countrymusic #coverart #singersongwriter #newmusic #sayittomyself
4	7102166885067754795	0.258199	We got to celebrate #EverythingButOn going #... it was a great time being around everyone who works hard to keep this ship goin. Thank y'all for streaming 🤘 #countrymusic #countrysong #singersongwriter

	video id	similarity	desc
0	7101754758817516806	1.000000	Any girls ever have this issue at the gym?! 😂 @heyitzhailzz_ #gymtok #gymlife #fittok
1	7091345141008502021	0.670820	This trend was made for gym girls #fyp #gymtok
2	7096300319100898566	0.632456	just positive gym things 😊. #FitTok #gymculture #workouttips #gymlife
3	7102934095961001259	0.516398	Filming in the gym does NOT give you priority over others. #MYOB #gymtok
4	7103235615285693738	0.516398	Info in comments #proteinpancakes #pancakes #protein #bulking #FitTok #gymtok
5	7098811142595693825	0.447214	I got the girls with big girls ❤️
6	7097359634490510597	0.447214	That one girl on our FYP 😭😂🔥
7	7099535560254967046	0.447214	#fyp #foryou #kylecookefitness #viral #gymtok
8	7093953318384307461	0.447214	Best girl #foryou #fyp #besties @a.m.y.l.e.i.g.h
9	7094569957421223174	0.447214	Tik tok do your thing! Let's help a girl out #norway #cheater #girlssupportgirls

C O N C L U S I O N S



Analysis for this task was fairly straightforward, but pulling data was difficult - will probably be broken within a year.



NLP alone vs. strict location boundaries - more research needed on TikTok's current recommender algorithm.

NEXT STEPS



Build a more time effective scraping system with AWS, automate with Selenium for longevity.



Add additional fields - creator ID, song title, song artist, FYP location, and origin country for recommender.



Pull in related video suggestions from API to assess model performance



Adding another layer of meaning by using nlp on description text, then comparing word vectors to build recommender (?)



THANK YOU!

QUESTIONS?