

SDAIA Academy T5C04 Bootcamps: Data Science
Natural Language Processing and Unsupervised Learning Project Presentation



What's trending on App Store?

Presented by: Nada Rambu

01

Problem and Solution

Project background and objectives

Objective

Learn more about the most popular app genre with high ratings using topics modeling techniques



02

Data Preprocessing

Bringing the data from the its source to be processed

Dataset

- 1."id" : App ID
- 2."track_name": App Name
- 3."size_bytes": Size (in Bytes)
- 4."currency": Currency Type
- 5."price": Price amount
- 6."ratingcounttot": User Rating counts (for all version)
- 7."ratingcountver": User Rating counts (for current version)
- 8."user_rating" : Average User Rating value (for all version)
- 9."userratingver": Average User Rating value (for current version)
- 10."ver" : Latest version code
- 11."cont_rating": Content Rating
- 12."prime_genre": Primary Genre
- 13."sup_devices.num": Number of supporting devices
- 14."ipadSc_urls.num": Number of screenshots showed for display
- 15."lang.num": Number of supported languages
- 16."vpp_lic": Vpp Device Based Licensing Enabled
- 17.app_desc: Application description

Dataset

- 1."id" : App ID
- 2."track_name": App Name
- 3."size_bytes": Size (in Bytes)
- 4."currency": Currency Type
- 5."price": Price amount
- 6."ratingcounttot": User Rating counts (for all version)
- 7."ratingcountver": User Rating counts (for current version)
- 8."user_rating" : Average User Rating value (for all version)**
- 9."userratingver": Average User Rating value (for current version)
- 10."ver" : Latest version code
- 11."cont_rating": Content Rating
- 12."prime_genre": Primary Genre
- 13."sup_devices.num": Number of supporting devices
- 14."ipadSc_urls.num": Number of screenshots showed for display
- 15."lang.num": Number of supported languages
- 16."vpp_lic": Vpp Device Based Licensing Enabled
- 17.app_desc: Application description**

Preprocessing

NLP Preprocessing Pipeline

1. Remove URLs
2. Remove punctuations
3. Remove non-alphabetic characters
4. Lower casing
5. Convert text to tokens
6. Remove stopwords
7. Apply lemmatization

Features



Photo & Video



Book



Travel



Productivity



Music



Weather



Education



Shopping



Food & Drink



Health & Fitness



Social Networking



Entertainment

Features



Photo & Video



Book



Travel



Productivity



Music



Weather



Education



Shopping



Food & Drink



Health & Fitness



Social Networking



~~Entertainment~~

Features



Features

855 Documents

11 Topics

1894 Unique words

03

Model Development

Train and test the model to estimate results

Models

LSA

`sklearn.decomposition.TruncatedSVD`

LDA

`sklearn.decomposition.LatentDirichletAllocation`

NMF

`sklearn.decomposition.NMF`

Models

LSA
(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

Social Networking?

Models

LSA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

Photo & Video

Models

LSA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

Weather

Models

LSA
(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

Weather?

Models

LSA
(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

Health & Fitness

Models

LSA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

Productivity

Models

LSA
(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

???

Models

LSA
(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

???

Models

LSA

(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	video	share	create	free	like	add	filter	effect	text
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card

???

Models

LSA
(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	
Topic 0	photo	video	share	create	free	like	add	filter	effect	text	
Topic 1	photo	video	collage	effect	filter	camera	frame	instagram	editing	sticker	
Topic 2	weather	photo	map	forecast	location	alert	radar	storm	data	day	
Topic 3	video	weather	forecast	radar	map	alert	current	storm	speed	location	
Topic 4	workout	exercise	photo	fitness	weight	account	training	period	plan	food	
Topic 5	file	pdf	document	text	support	note	cloud	device	image	scan	
Topic 6	workout	sleep	sound	exercise	fitness	meditation	effect	mode	weather	image	
Topic 7	sleep	music	sound	meditation	sky	star	song	alarm	night	period	
Topic 8	sky	star	moon	event	chart	night	calendar	device	planet	constellation	
Topic 9	food	life	sleep	toca	recipe	product	day	share	friend	card	???

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Weather

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Shopping

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Educational - Book

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Photo & Video

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Productivity

Models

LDA

(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Educational

Models

LDA

(with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

???

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Music?

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Social Networking?

Models

LDA (with CountVectorizer)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	weather	photo	map	tool	forecast	location	color	storm	alert	data
Topic 1	video	store	best	state	deal	save	location	product	free	shop
Topic 2	kid	fun	world	toca	life	friend	like	meditation	story	explore
Topic 3	photo	video	effect	filter	camera	calculator	mode	collage	frame	calculation
Topic 4	file	device	support	note	event	email	search	document	photo	view
Topic 5	child	word	fun	number	learning	learn	math	kid	award	book
Topic 6	account	period	current	end	free	purchase	hour	charged	user	itunes
Topic 7	sound	sleep	star	sky	audio	moon	support	mode	map	music
Topic 8	photo	create	image	text	share	add	tool	instagram	color	design
Topic 9	workout	exercise	fitness	apple	track	food	health	weight	training	day

Health & Fitness

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Photo & Video

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Educational

Models

NMF (with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Health & Fitness

Models

NMF (with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Weather

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Shopping

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Productivity

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Social Networking

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Food & Drink

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

Music

Models

NMF
(with tf-idf)

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Topic 0	photo	filter	collage	effect	tool	image	color	camera	frame	sticker
Topic 1	kid	child	fun	word	panda	learn	parent	activity	award	learning
Topic 2	workout	exercise	fitness	training	weight	plan	body	minute	goal	runtastic
Topic 3	weather	map	forecast	radar	location	alert	storm	current	data	snow
Topic 4	period	account	current	charged	end	month	hour	purchase	turned	renewal
Topic 5	file	document	pdf	text	note	drive	support	image	word	cloud
Topic 6	video	instagram	effect	clip	share	speed	movie	camera	photo	live
Topic 7	food	deal	item	store	shop	recipe	shopping	product	order	restaurant
Topic 8	music	sound	song	sleep	playlist	audio	meditation	artist	listen	midi
Topic 9	calculator	calculation	button	scientific	mode	basic	equation	number	function	history

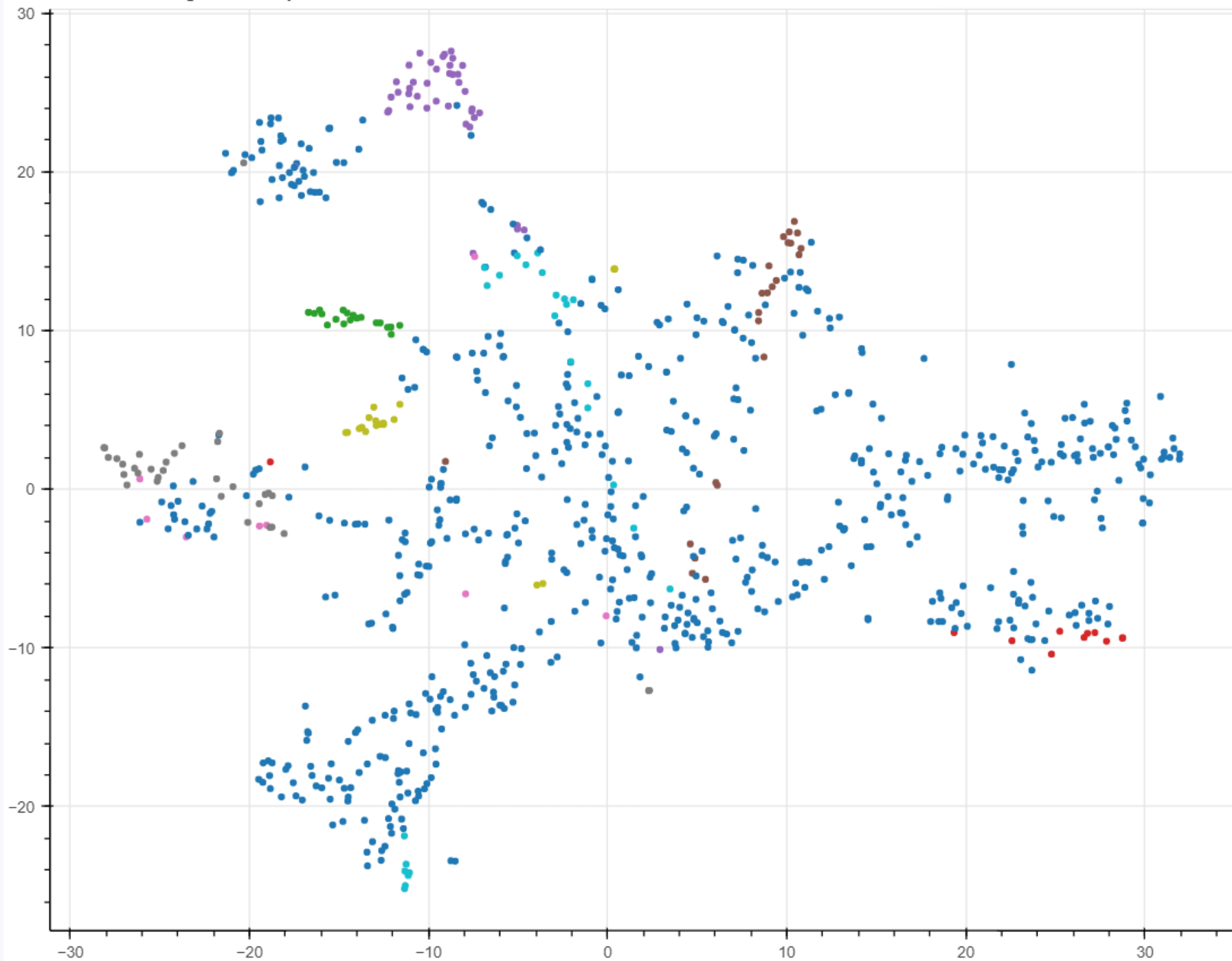
???

04

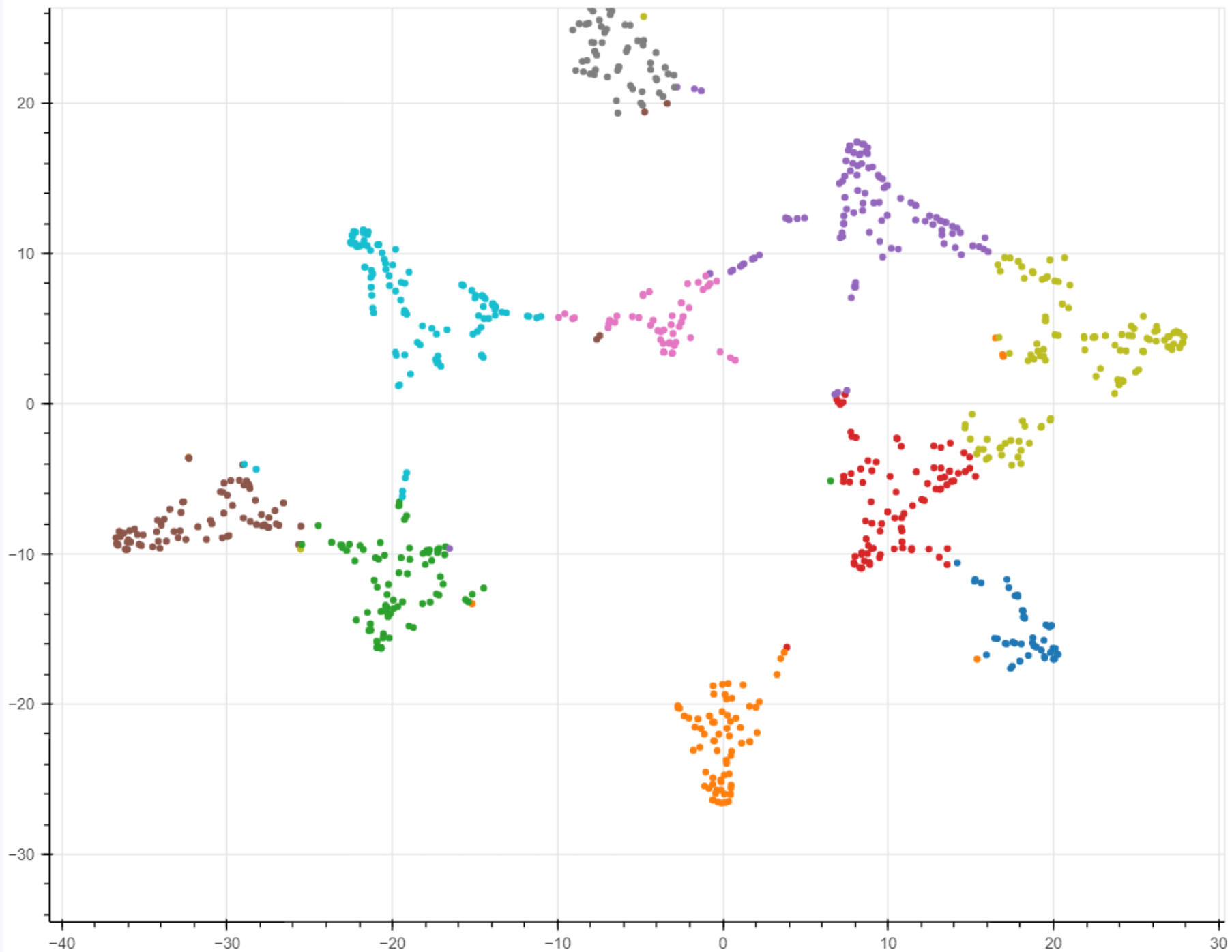
Data Visualization

Gaining insights from the data

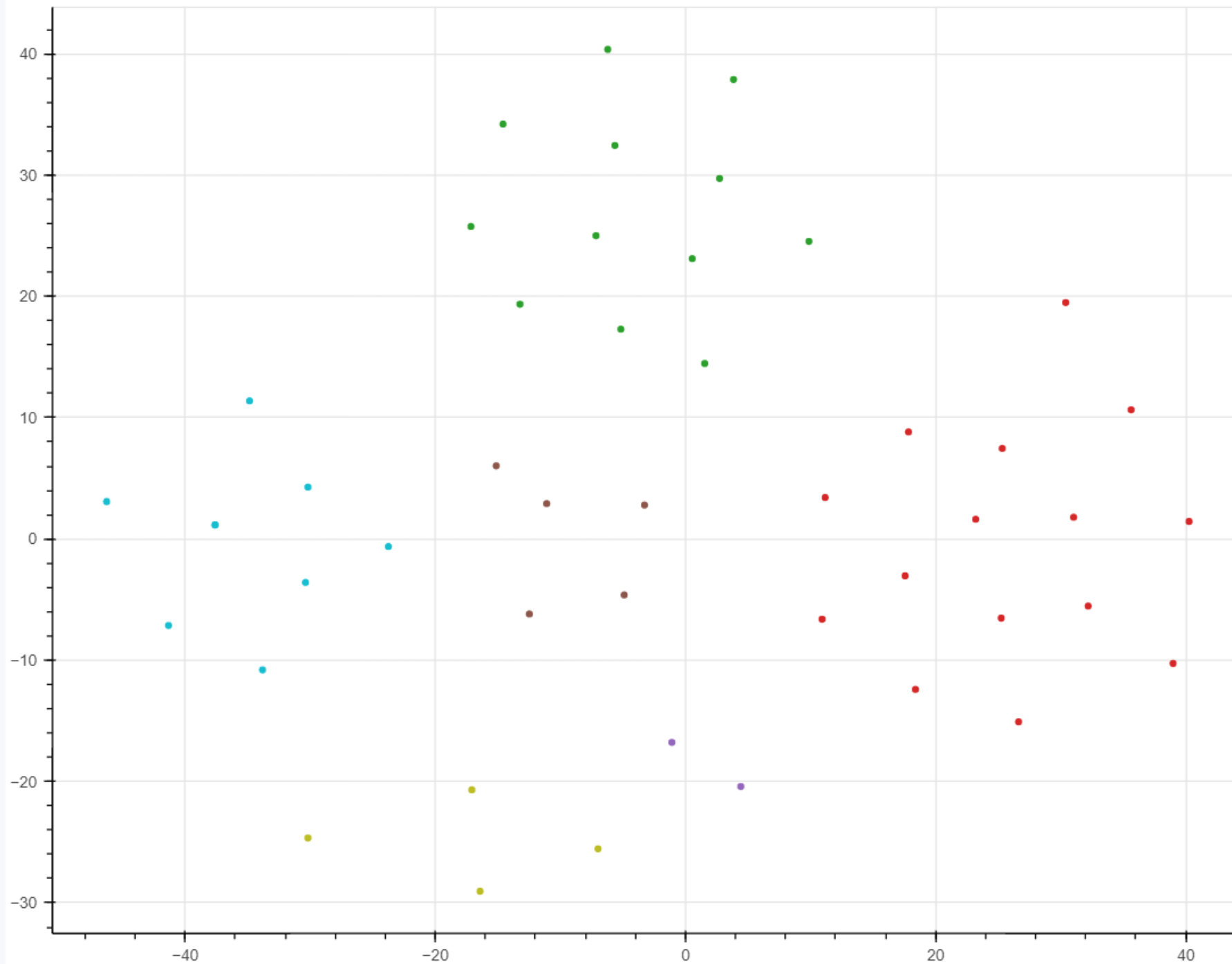
t-SNE Clustering of LSA Topics



t-SNE Clustering of LDA Topics



t-SNE Clustering of NMF Topics



Visualization



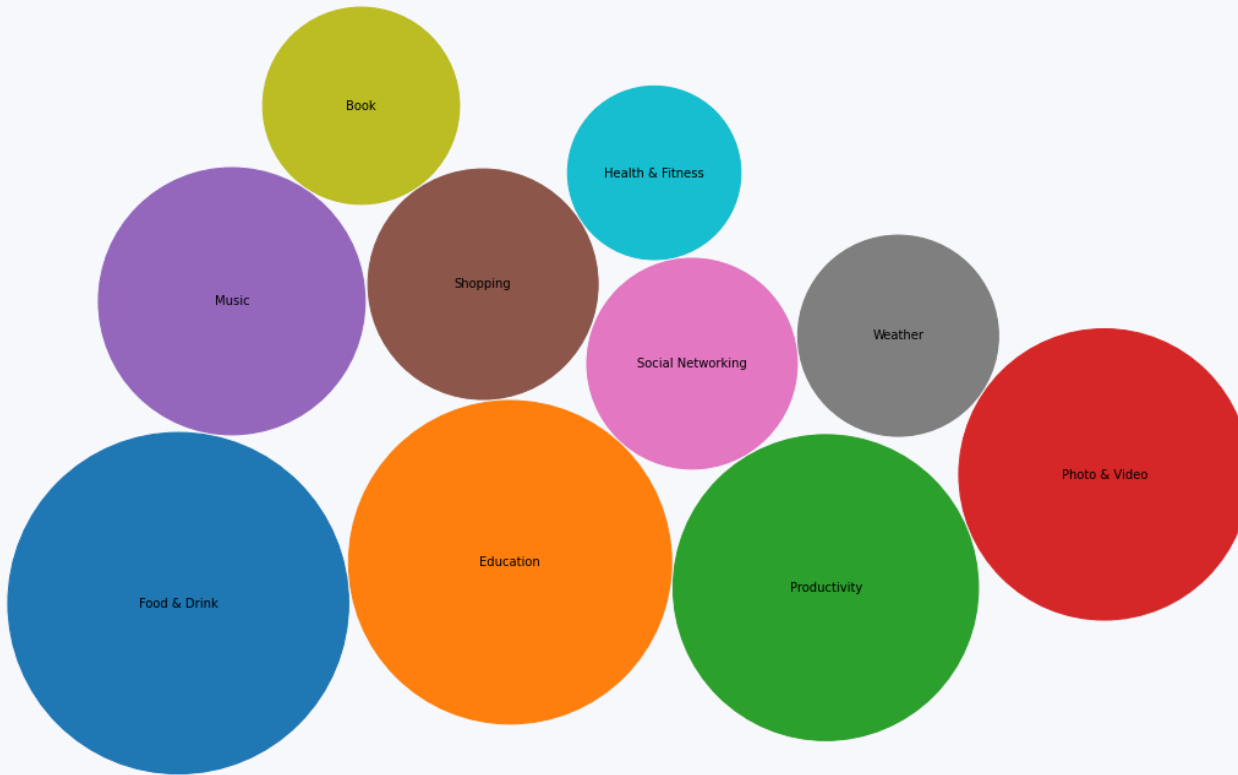
05

Conclusion

Some recommendation to solve the problem

Insights

Most Discussed Topic by NMF



According to the bubble chart, the top 5 most populous app genre are:

1. Food & Drink
2. Education
3. Productivity
4. Photo & Video
5. Music



Thank You!