Social network analysis of discussion forum data

WANG Hao

- **Forum Data**
- Clustering Structure
- Topic Modeling
- Conclusion

Forum data

course	#users	#posts	#threads	#forum	#comments
stschina1	835	1718	268	9	1033
newchinahistory1	1138	2641	434	12	1963
gastronomy	3725	9105	1474	14	6088

Tree view

General Discussion

An Introduction to Science and Technology Studies, by Sergio Sismondo

Wondering if there's an accessible copy of Sismondo within Coursera ?
Don't particularly want to buy the whole book just in order to read p.1-11 next week.
Did a google search, found several places that would let you read the latter part for free but
version ?

I agree. We may need some people with Academic reach to these documents.

> Yeah, other courses will provide **open online resources** even it's optional readings. Unti
> instructors will provide it until the formal opening of the course. I think **course staffs**

It is not usual practice on these courses to require reading of text which are not publicly, f
Even learned papers are usually to be found behind a paywall and are out of reach of the bulk

If the text is not accessible, the bulk of students will be unable to complete the reading ass

I wan't be able to buy this book before the course starts. Could someone scan those 11 pages a

Hello !
What about that?

Tree view
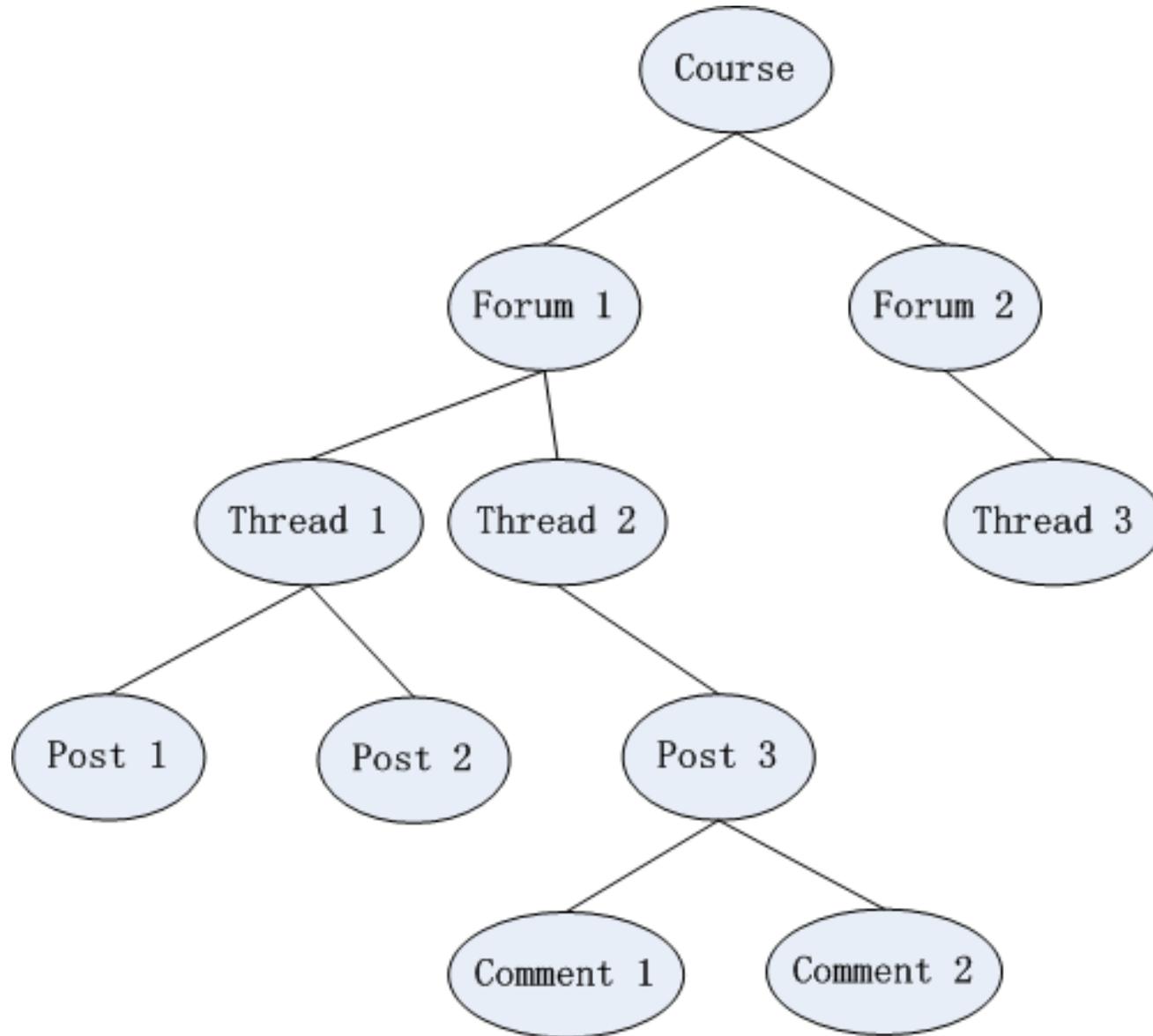
- **Forum Data**
- **Clustering Structure**
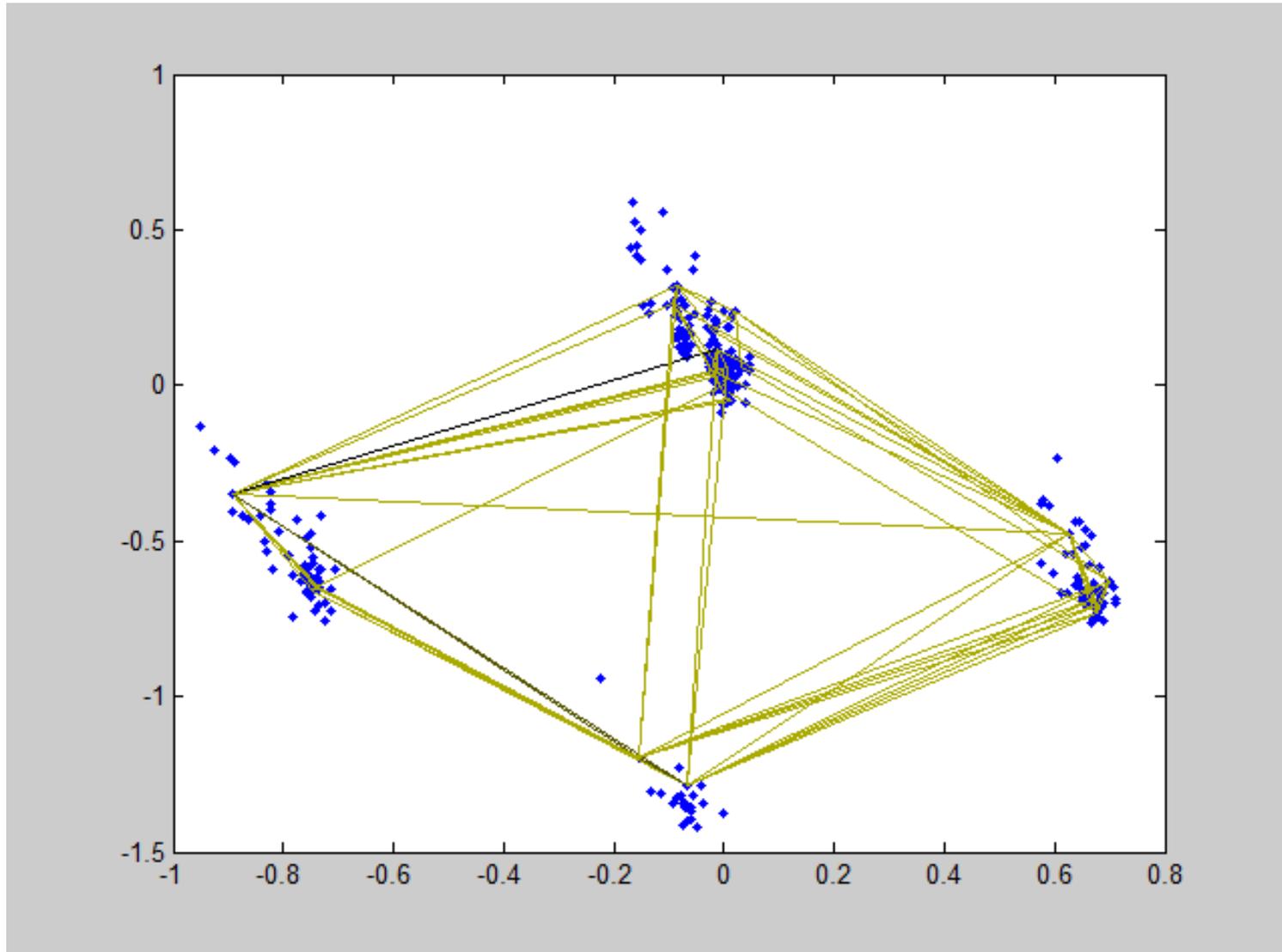- **Topic Modeling**
- **Conclusion**

Clustering using PCA

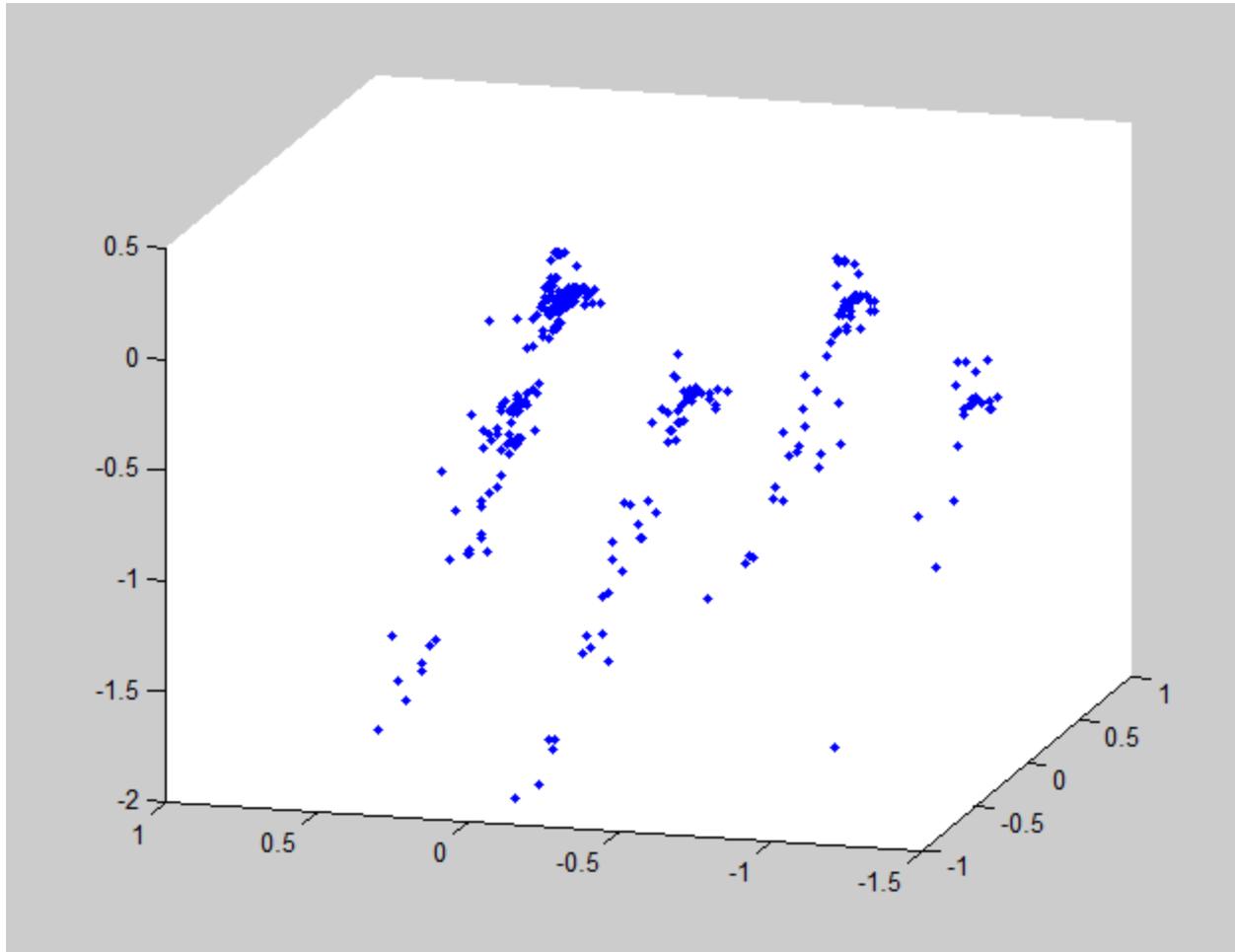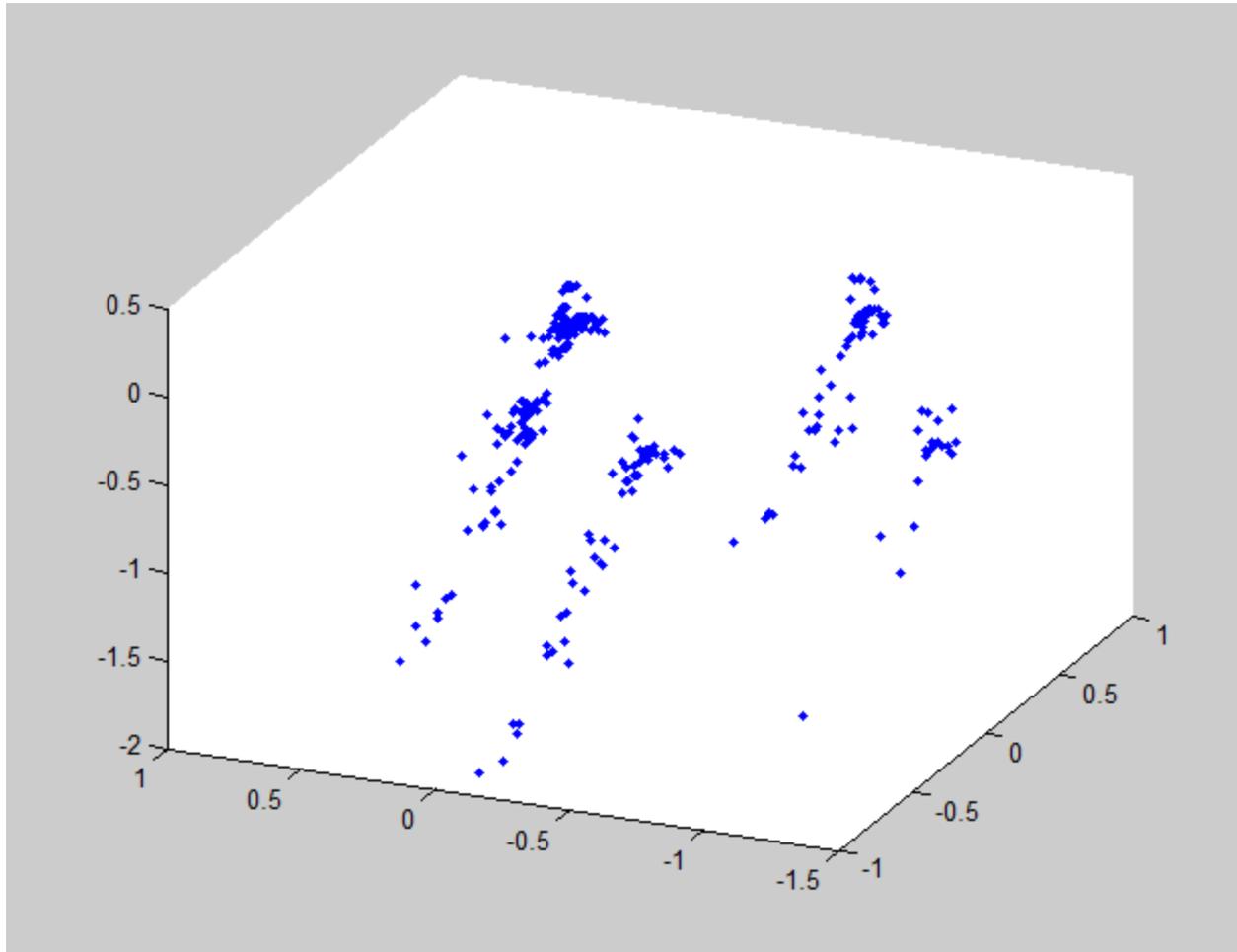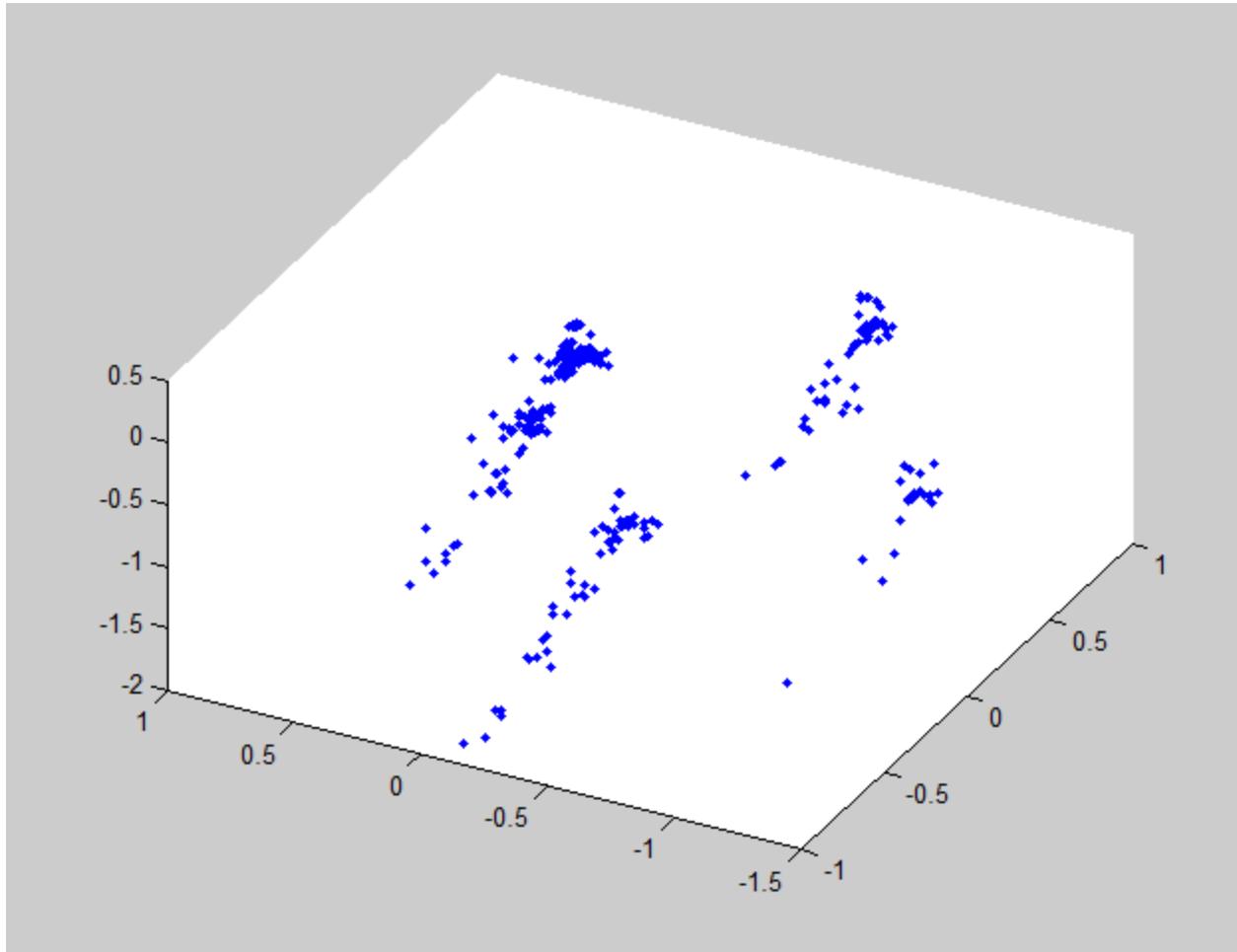| Thread / User	1	2	3	4	5	6
1		1				1
2						
3					1	
4	1		1			
5			1		1	

Social networks and clustering-stschina1
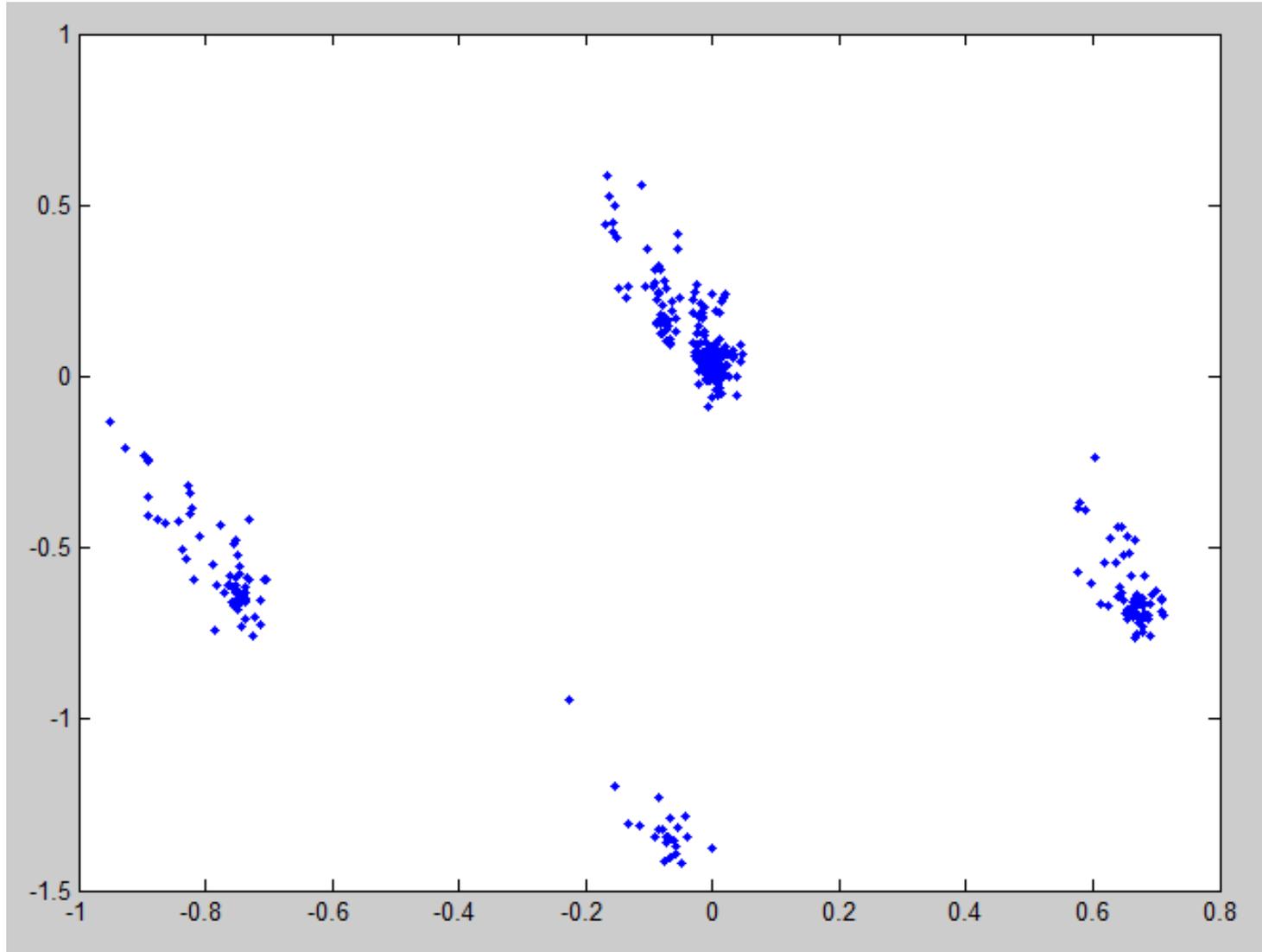
Social networks and clustering-stschina1
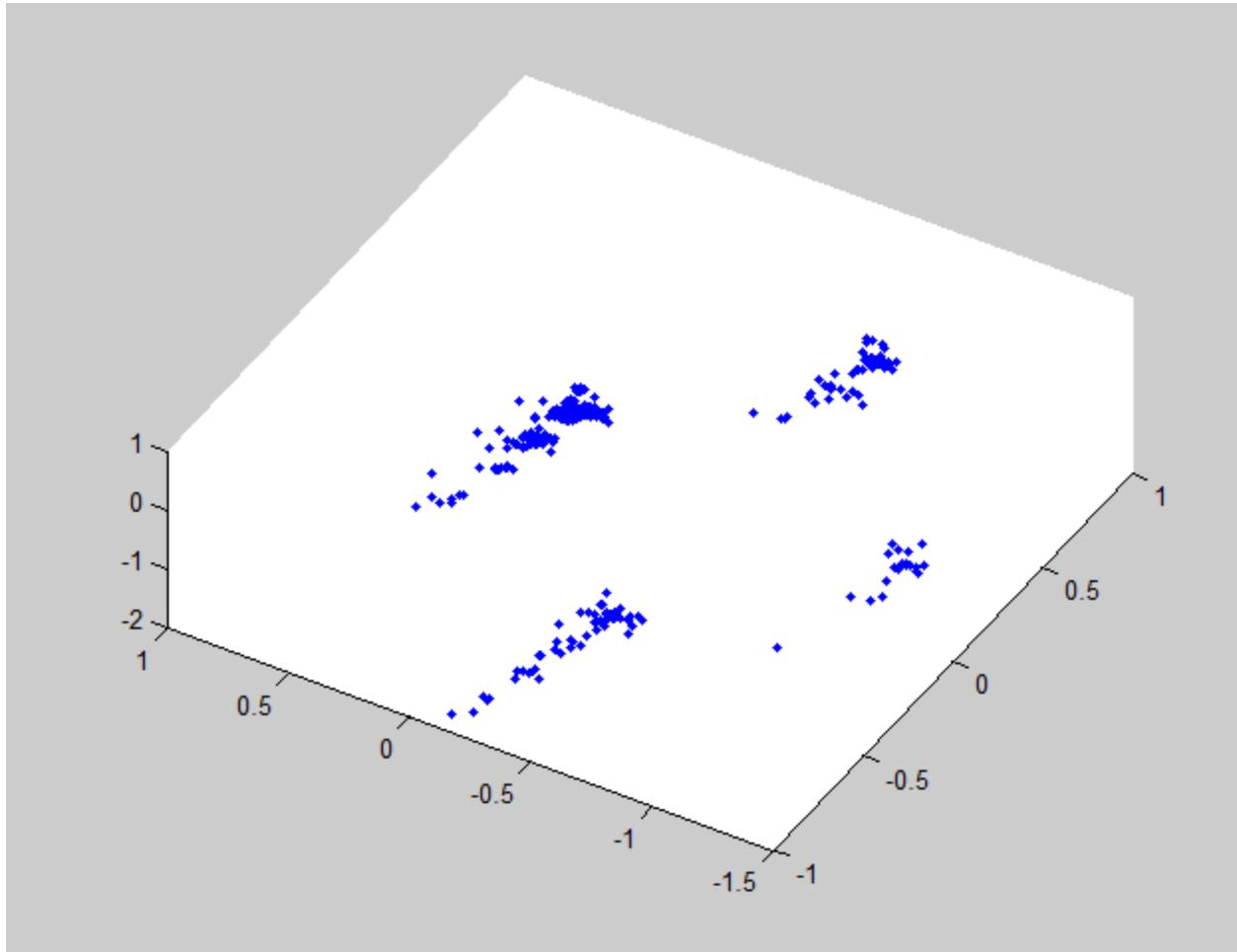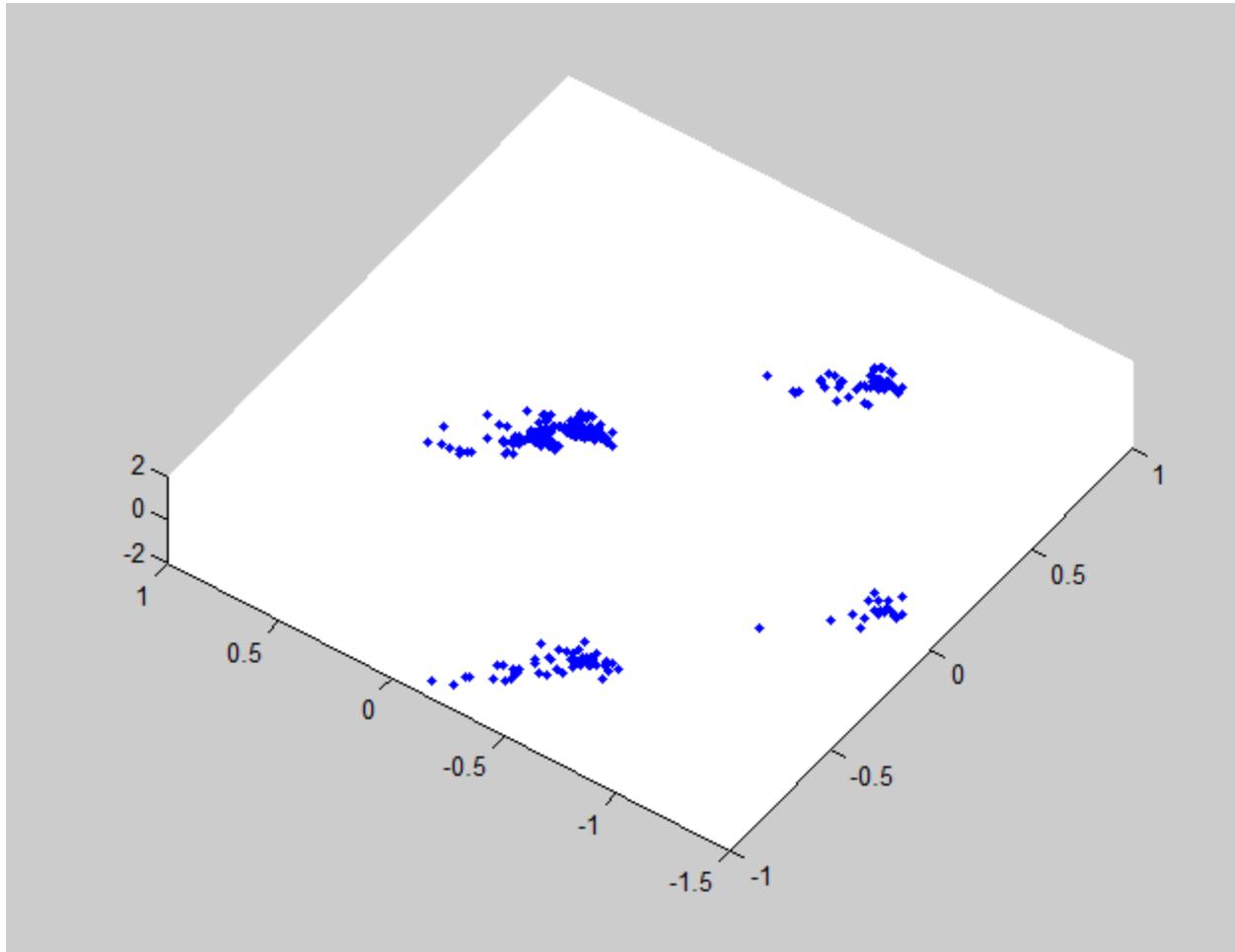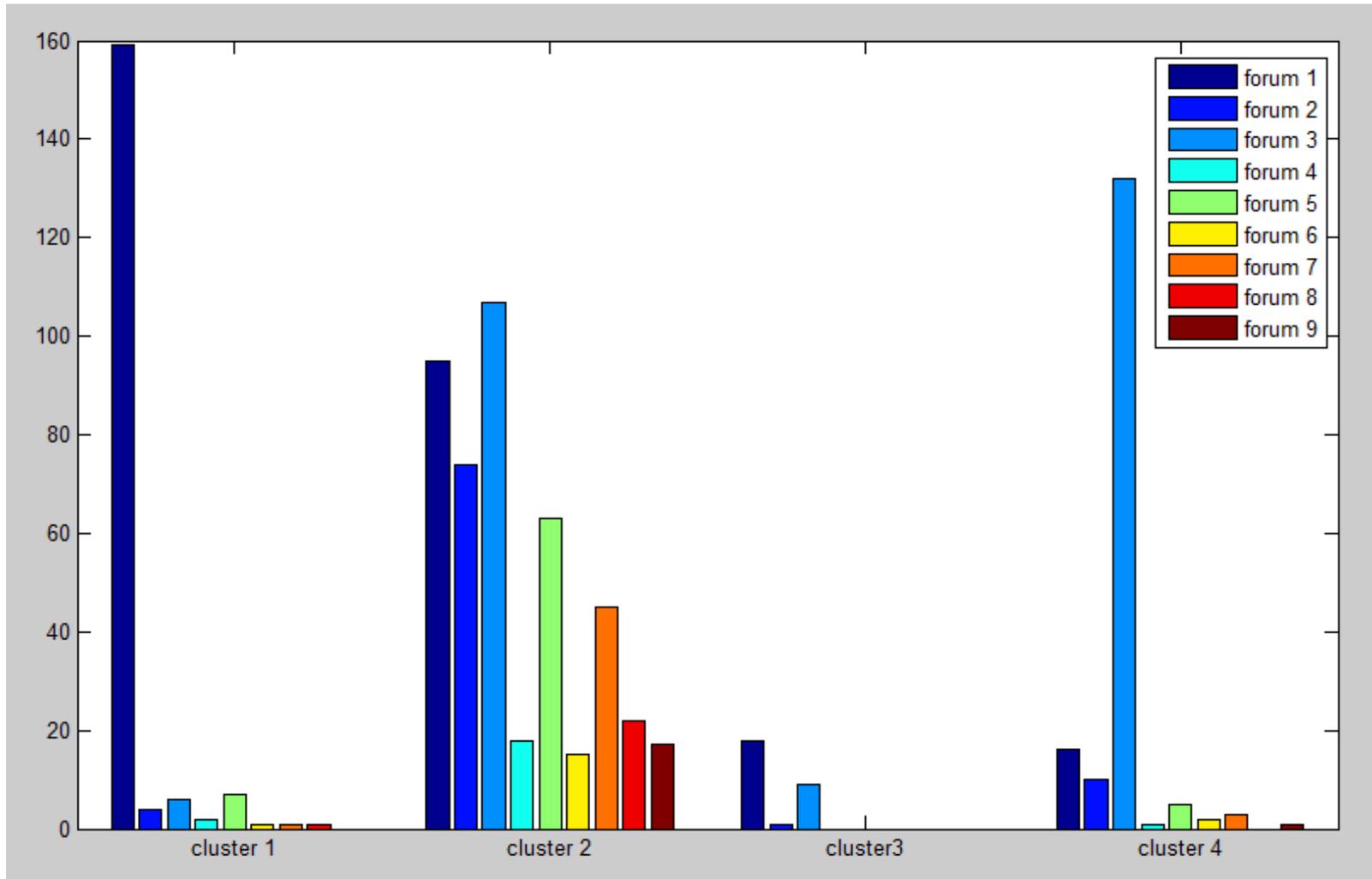
Social networks and clustering-stschina1

Social networks and clustering-stschina1

Social networks and clustering-stschina1

Social networks and clustering-stschina1

Social networks and clustering-stschina1

What clusters might mean

Forum distribution

What clusters might mean

User type – cluster 1

Passionate, outgoing, helpful <u>sometimes</u> talking about serious questions

User **80e5db41946a6463192efc416acddb63bc450480**

Post 516: As a scientist, don't you think ……

Post 402: I am taking this course because ……

Post 404: Jos Muyres found a download on the Wiley site. The address is ……

Post 407: Hi Everyone. My name is Barbara ……

What clusters might mean

User type – cluster 2

Less outgoing,
<u>Almost always</u> talking about serious questions

User **7643d42794487a9214c9a1e206ea7076bad0ab31**

Post 1465: Innovation being the practice of ……
Post 1688: Many/most modern products require ……
Thread 77: The idea that the world was ……
Thread 80: I have a major concern with ……

What clusters might mean

User type – cluster 3

Passionate, not very helpful, <u>hardly</u> talking about something productive

User **95424a76f82f273fc03f5fb587ebc4de3f1357c8**

Post 34: Hello! I am Chamath. I am in St. Paul ……
Post 47: Hi Andy. I'm Chamath in St. Paul ……
Post 822: Thanks Linda, for the instructions and the ……
Post 1222: How do I join the group ……

What clusters might mean

User type – cluster 4

Passionate, not very helpful,
<u>hardly</u> talking about serious stuff
<u>sometimes</u> talking about things that are
non-lecture-related but interesting

User **3c224926ee2ec05205f3c21ba2914f787e5bfa03**

Post 44: I am Julia from near Leeds in the UK ……

User **3c224926ee2ec05205f3c21ba2914f787e5bfa03**

Post 107: what does the tag say in countries where
English is not spoken?

Post 121: Good to learn from a professor who works there.

Other datasets – newchinahistory1
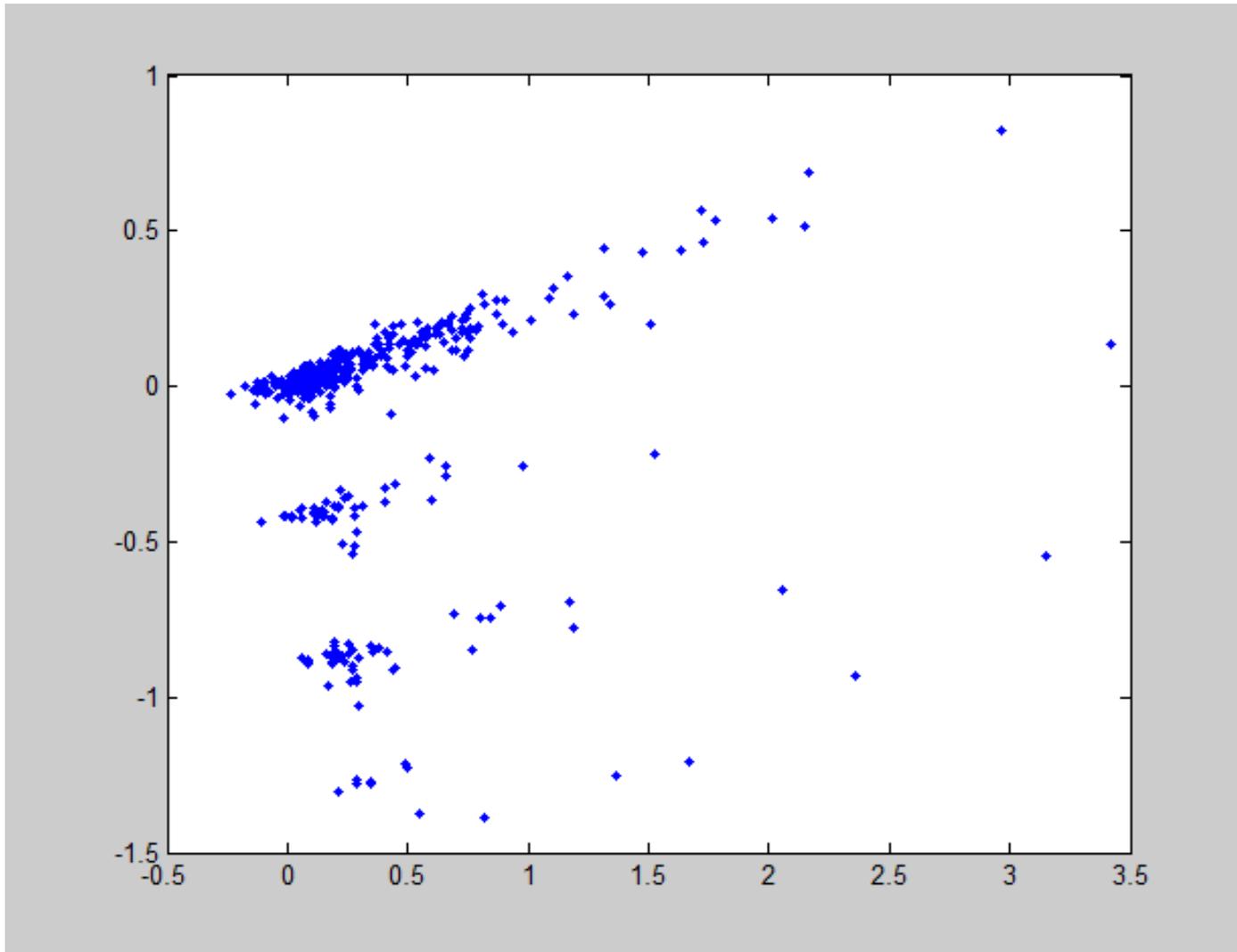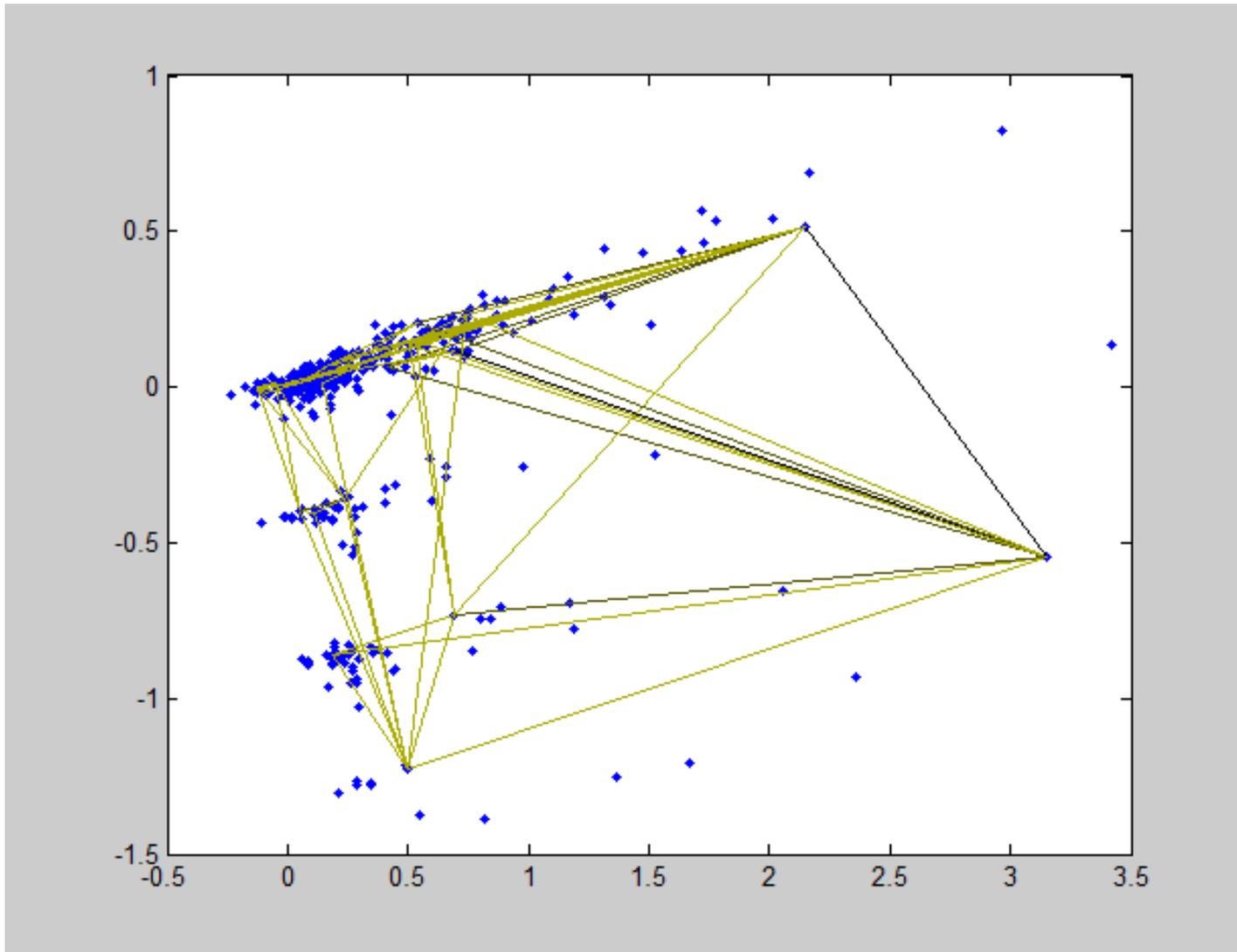
Other datasets – newchinahistory1

Other datasets – newchinahistory1
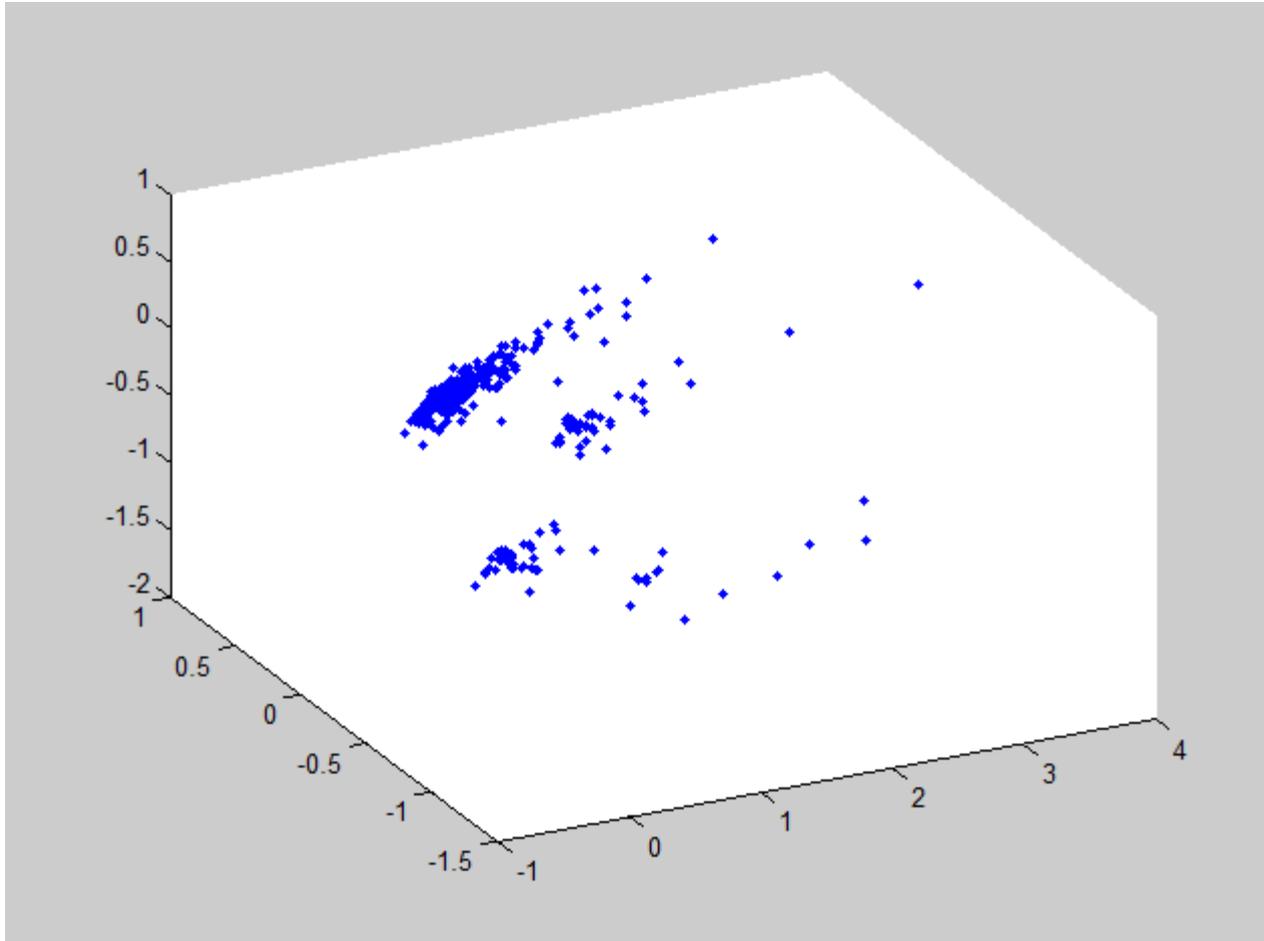
Other datasets – newchinahistory1
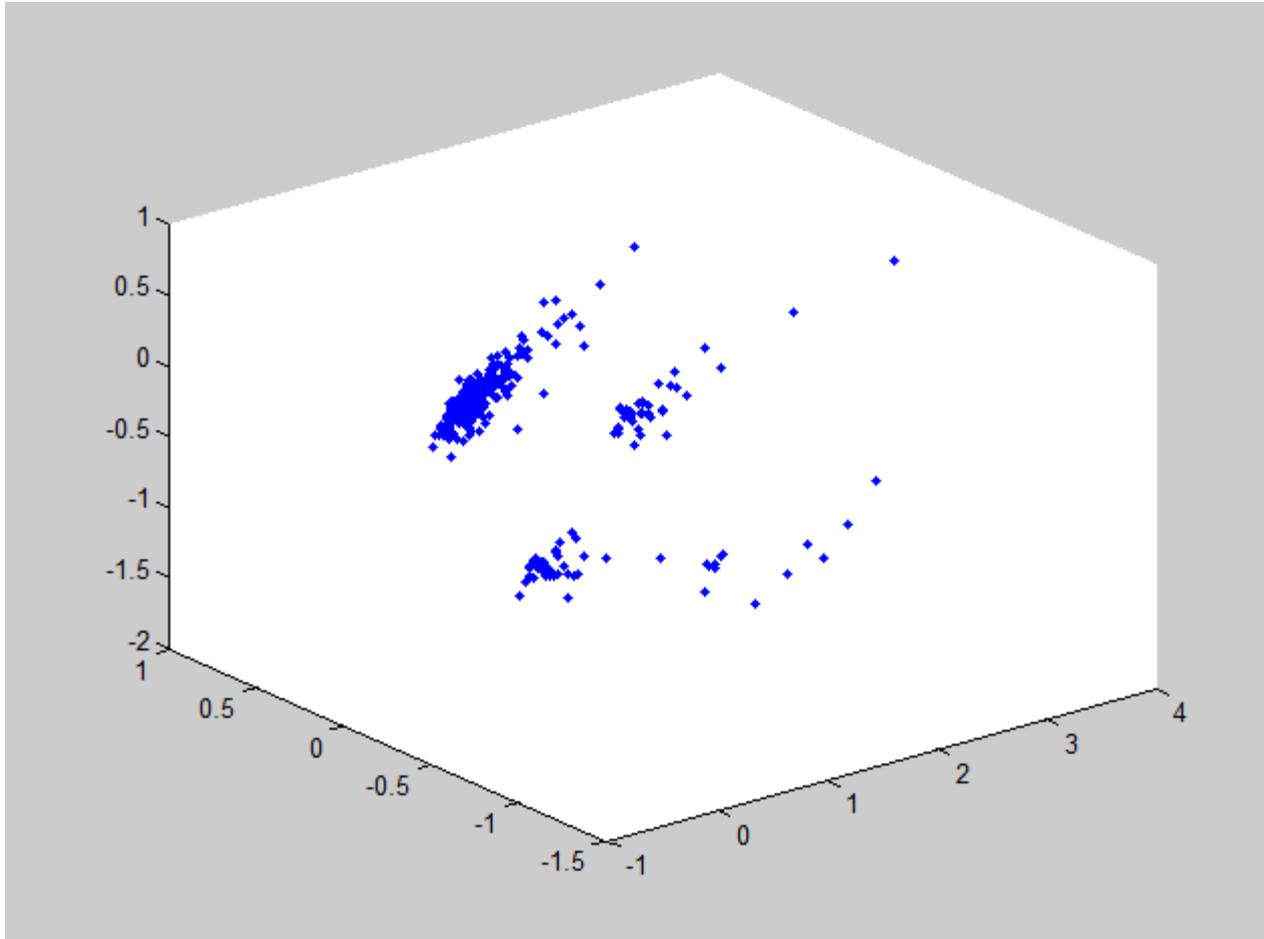
Other datasets – newchinahistory1
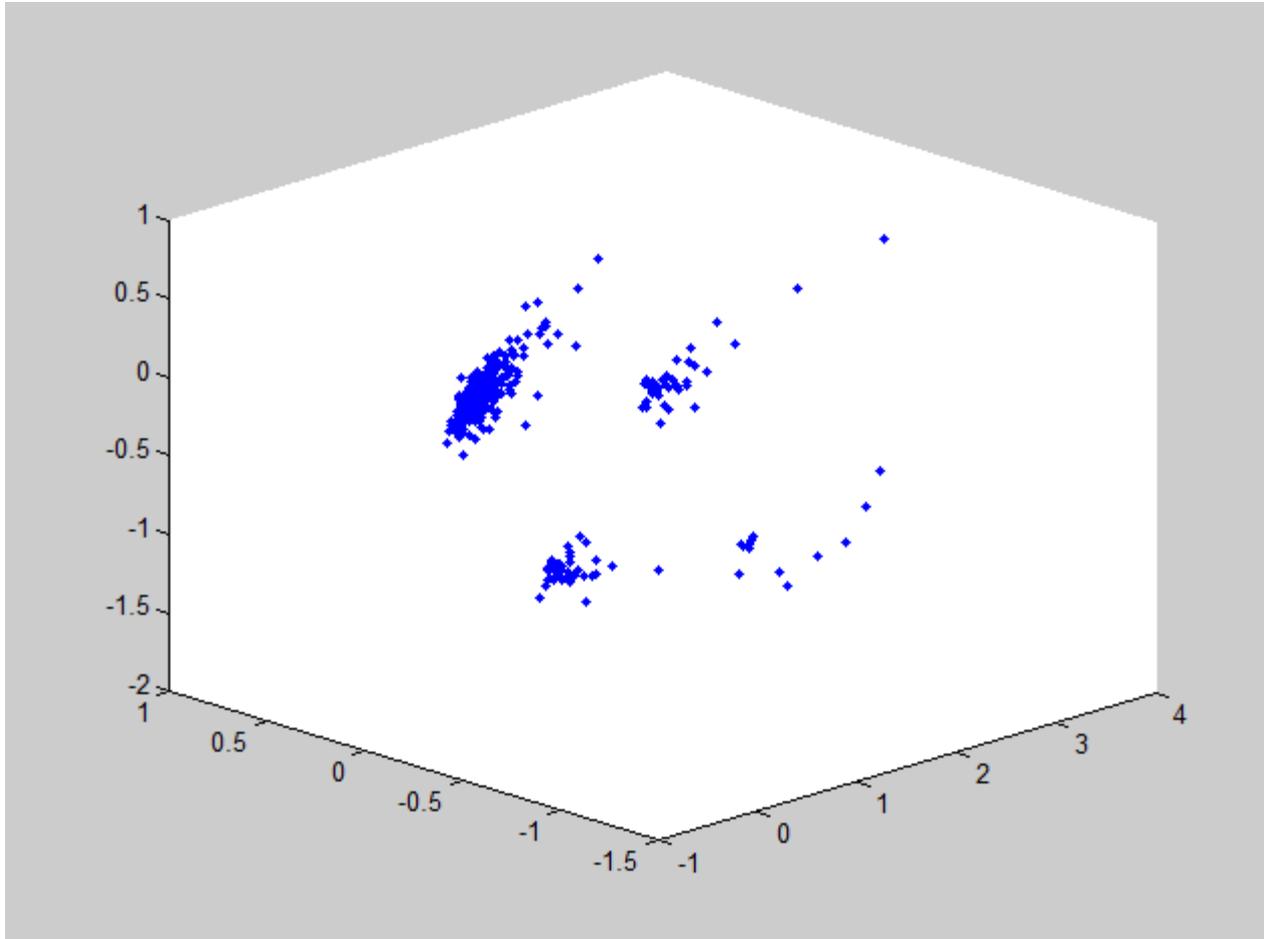
Other datasets – newchinahistory1
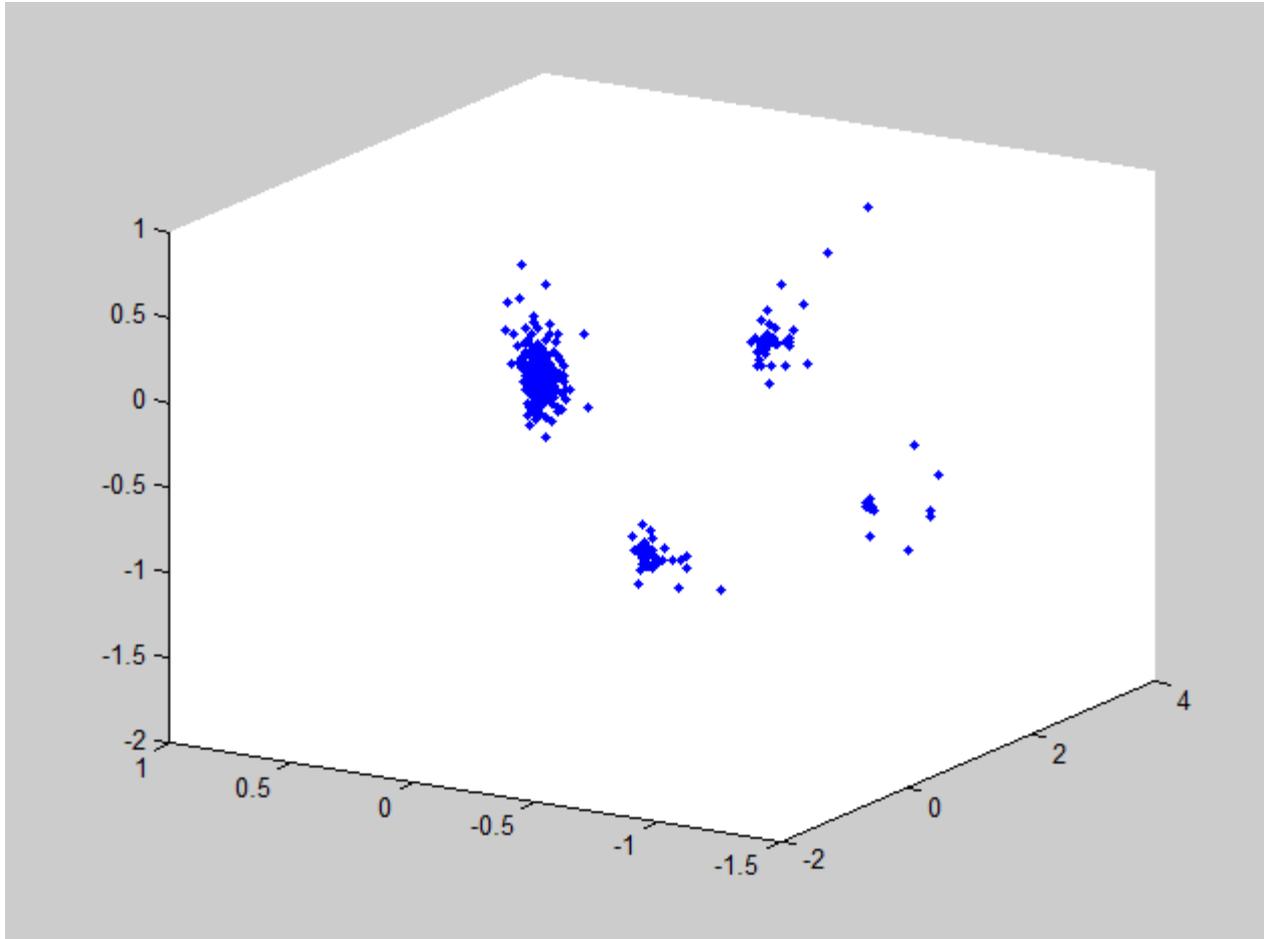
Other datasets – newchinahistory1
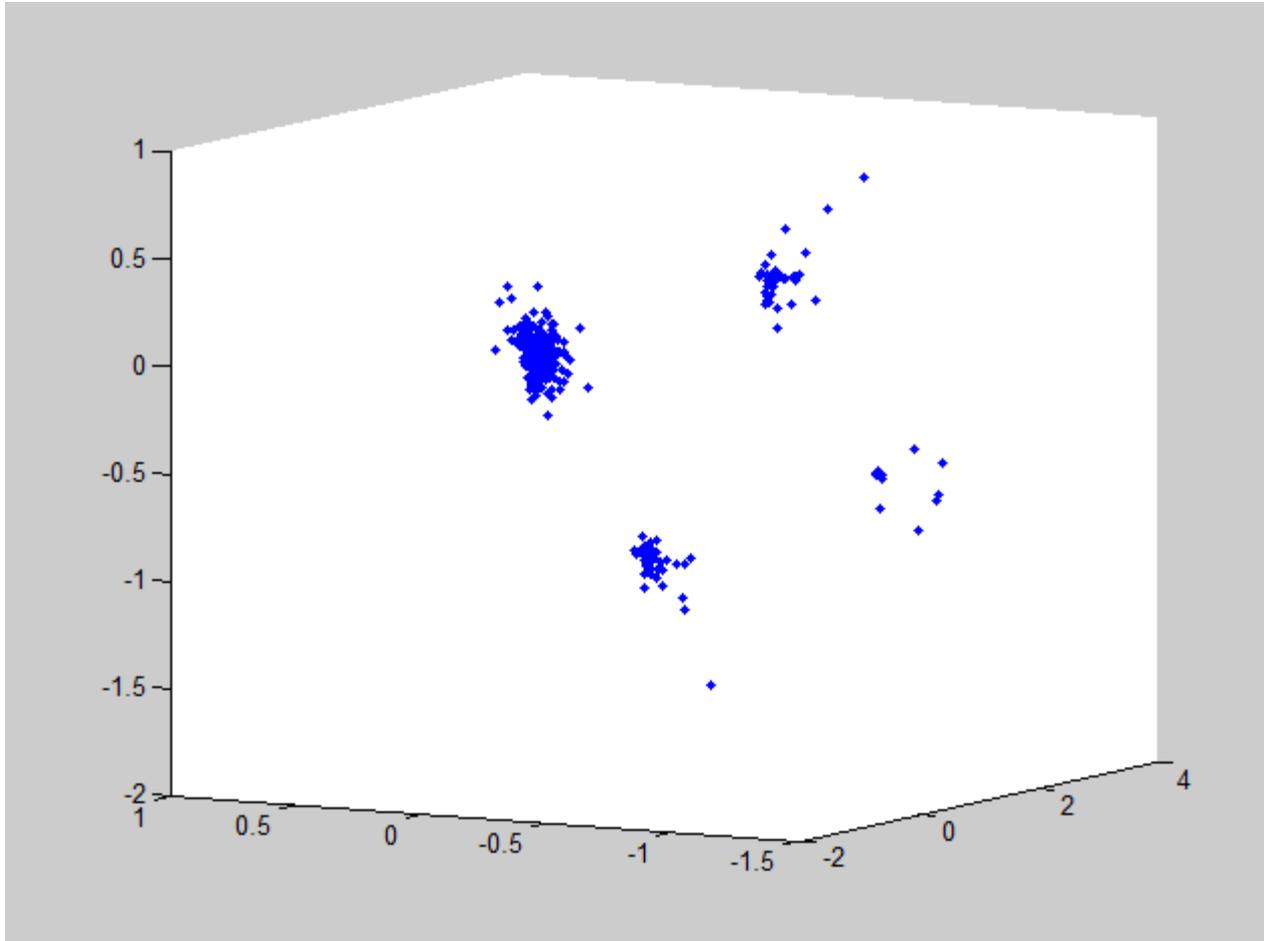
Other datasets – newchinahistory1
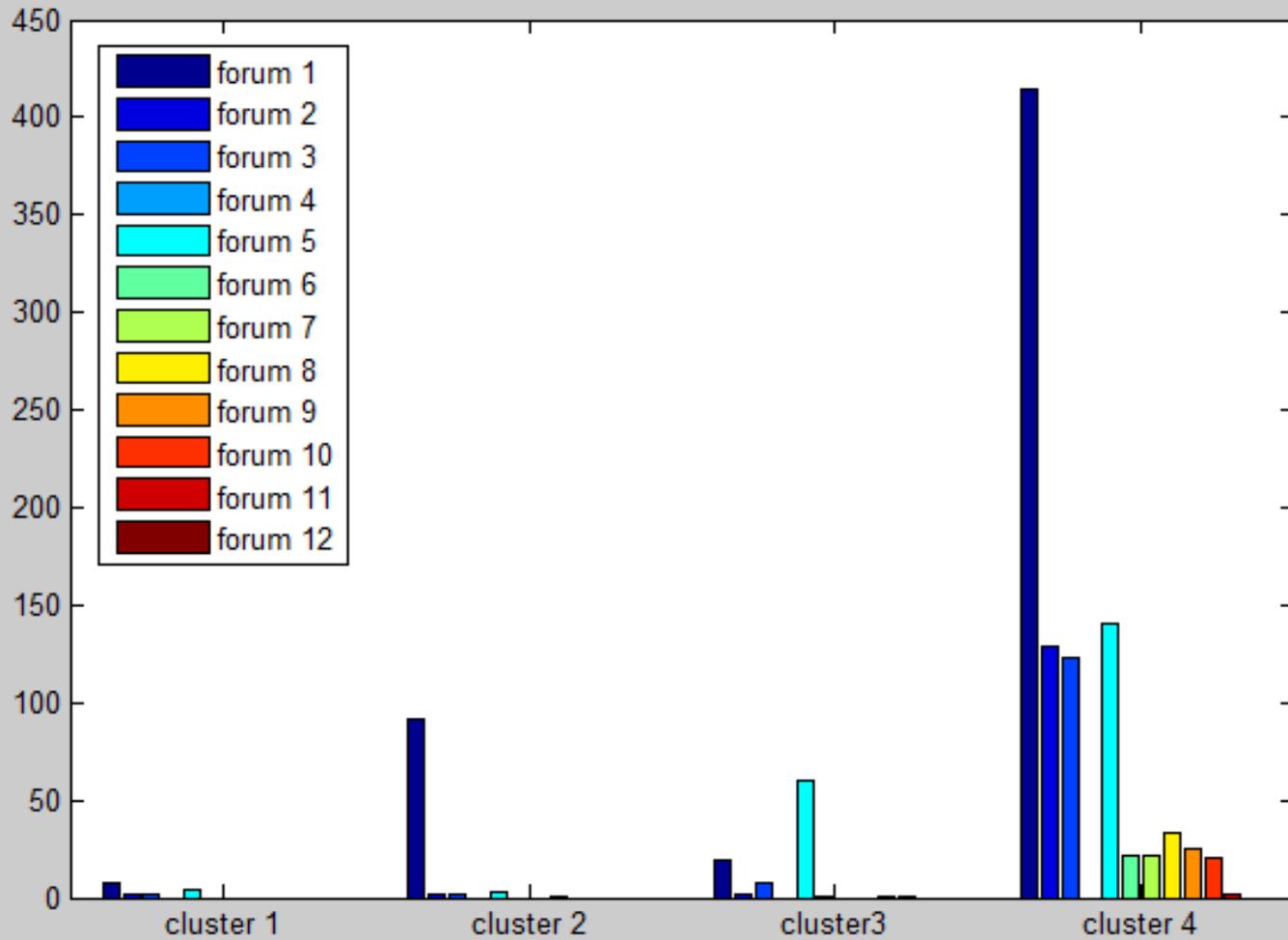
Other datasets – gastronomy
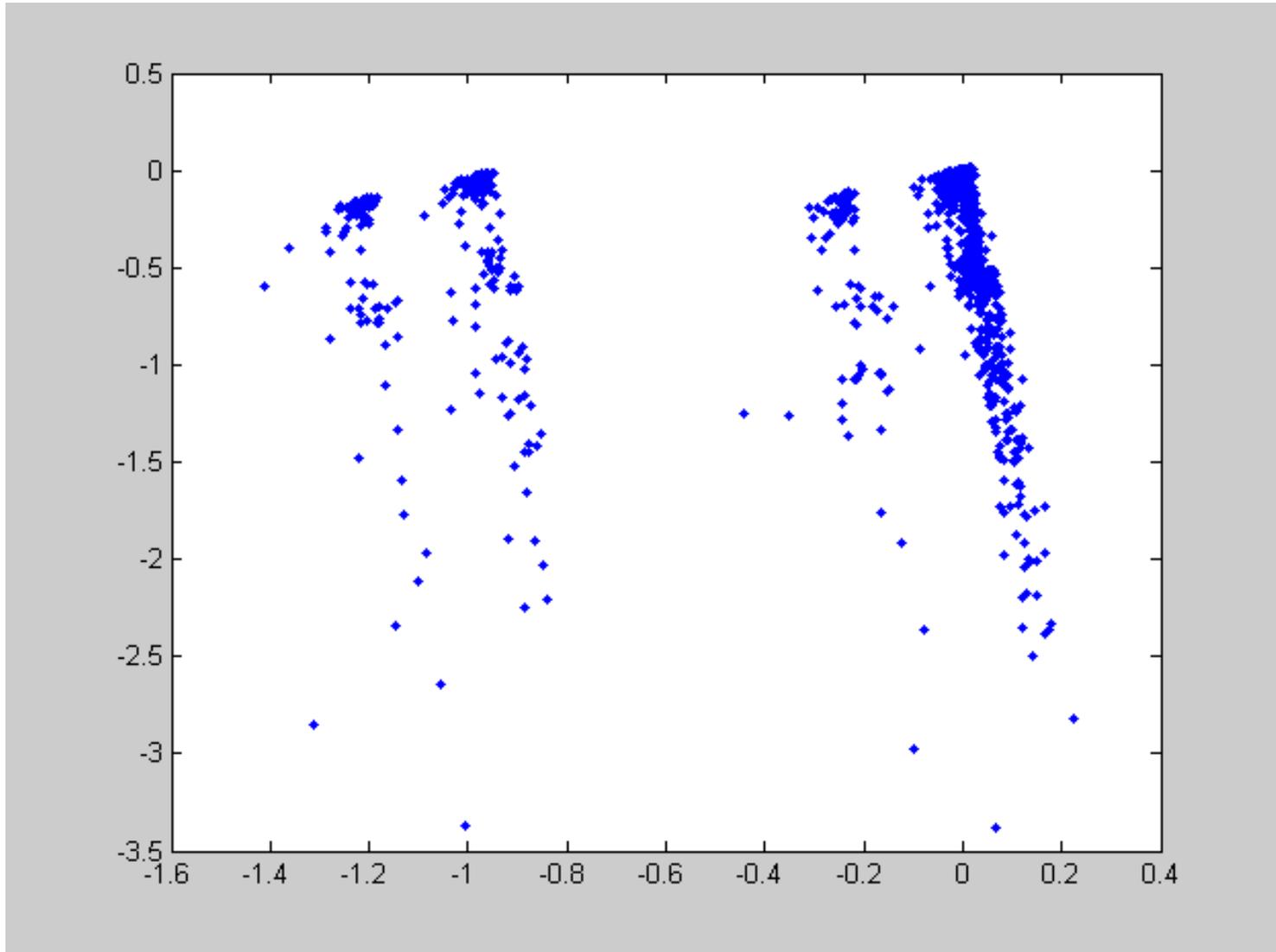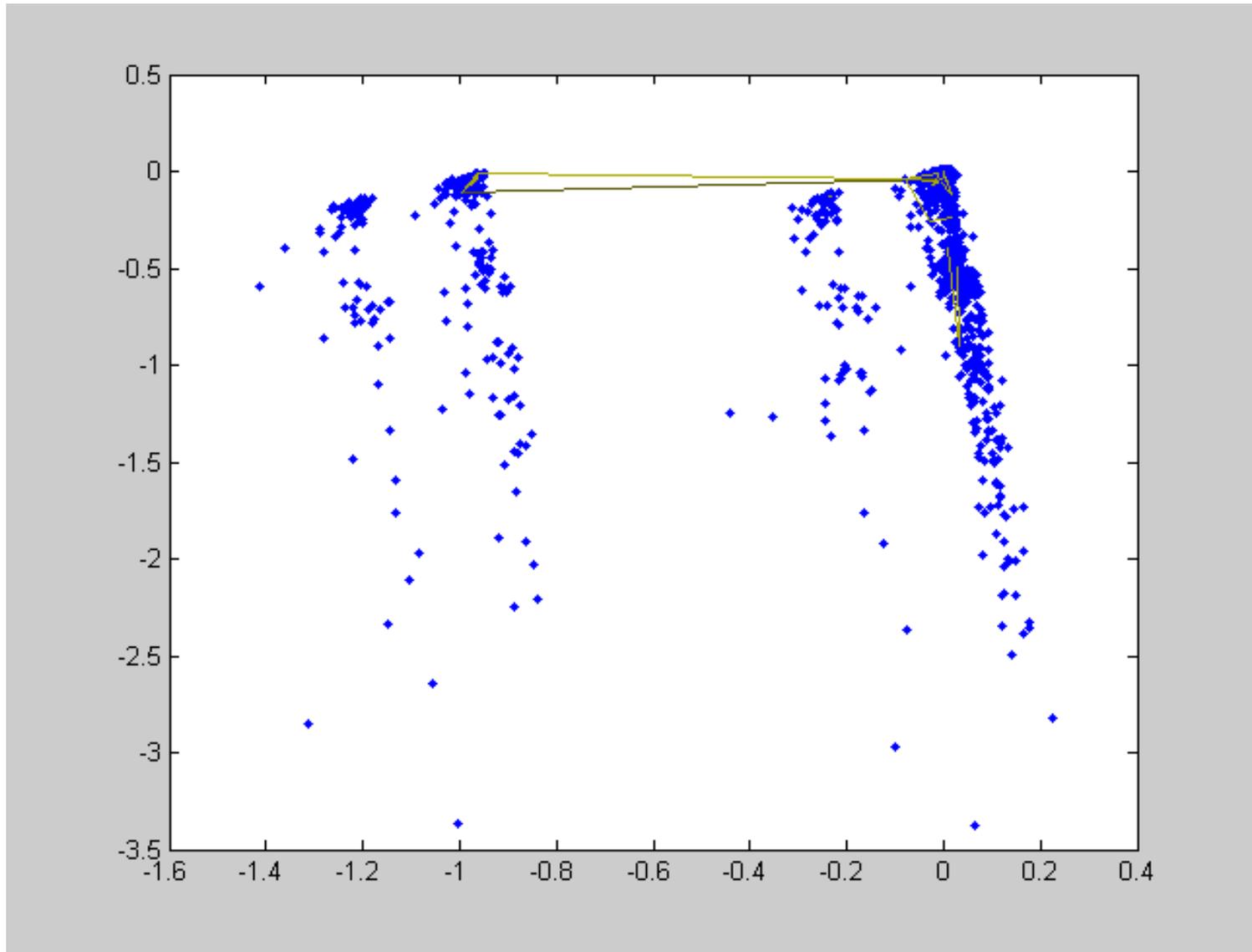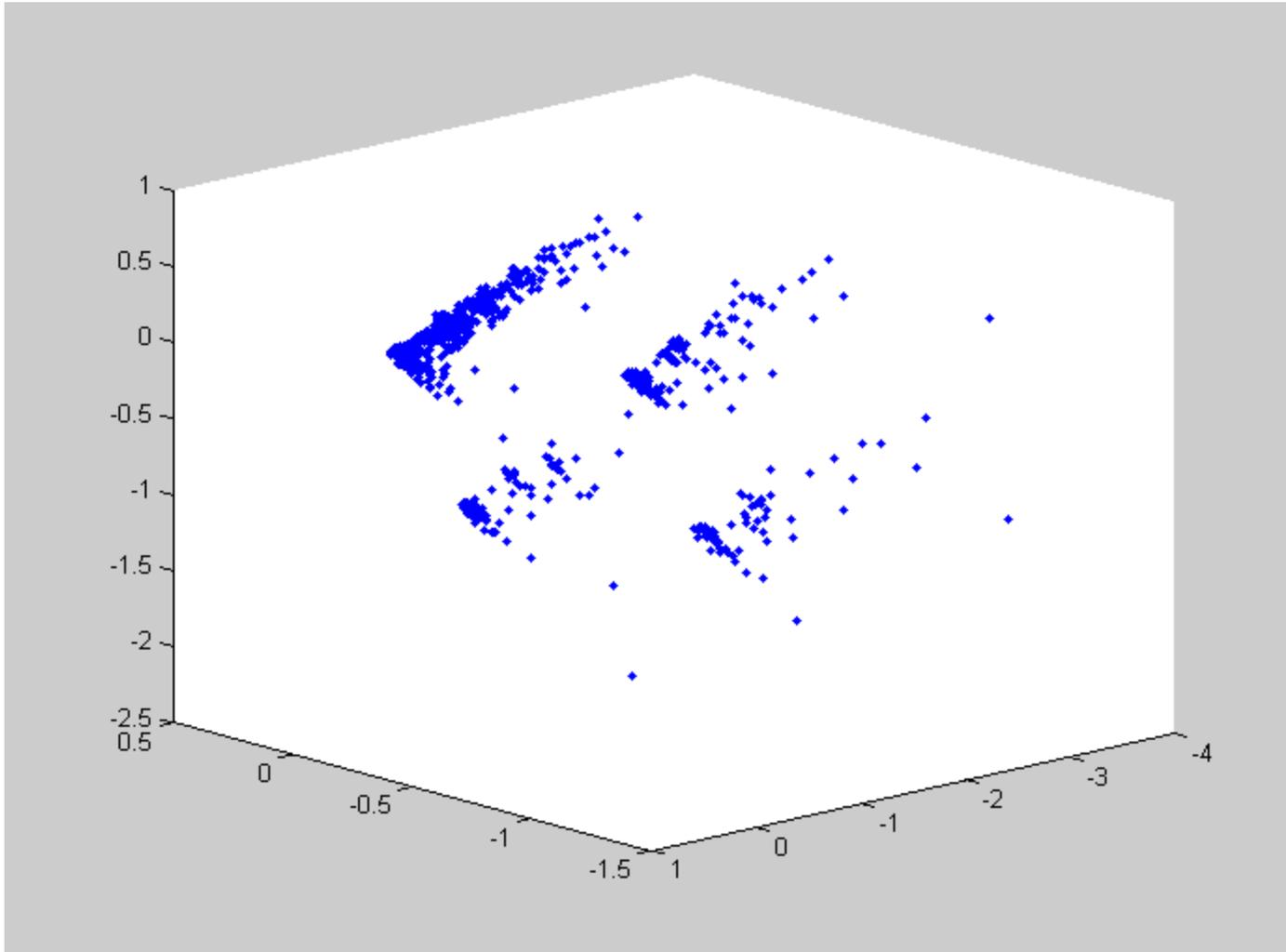
Other datasets – gastronomy

Other datasets – gastronomy

Other datasets – gastronomy
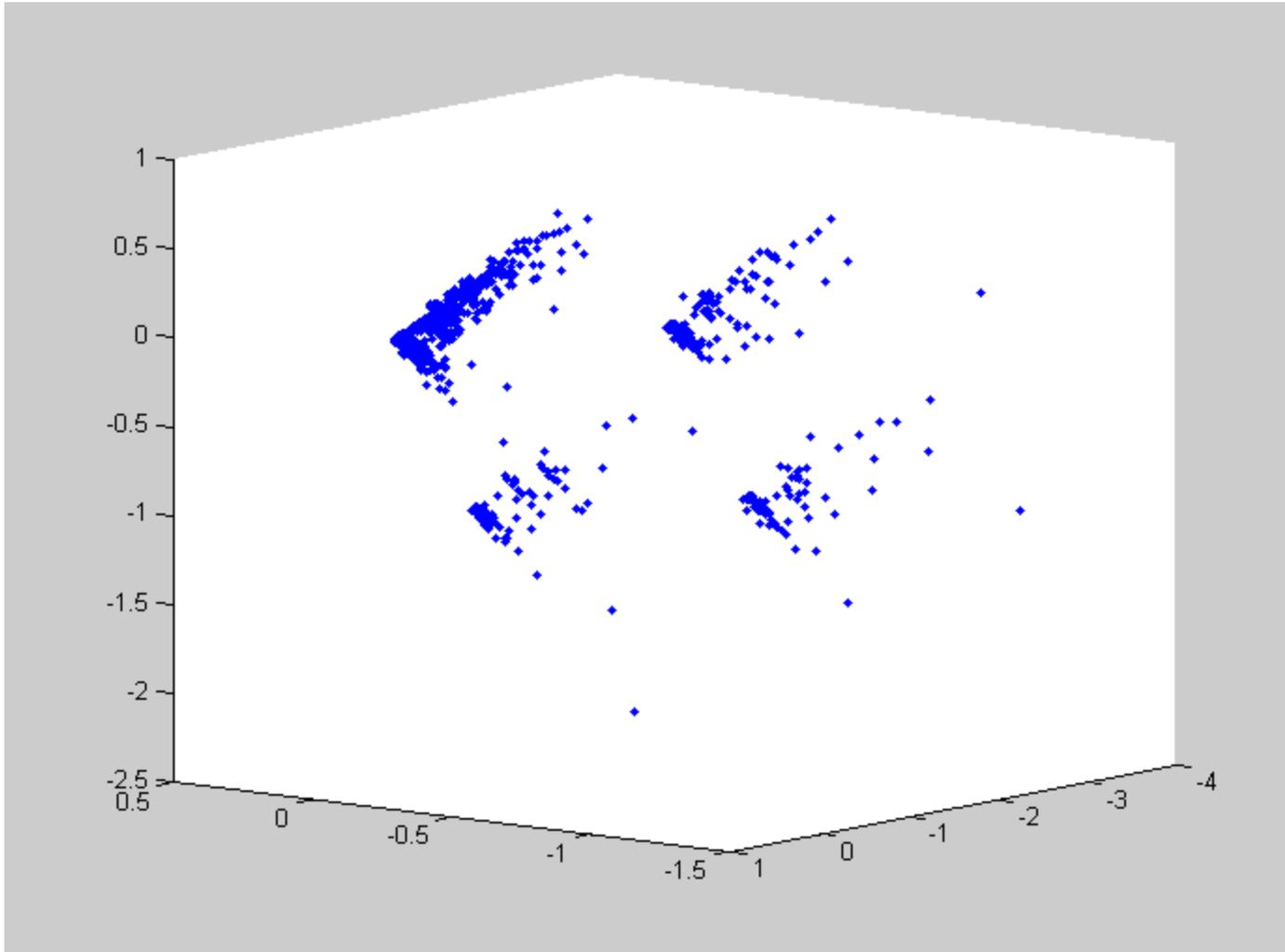
Other datasets – gastronomy
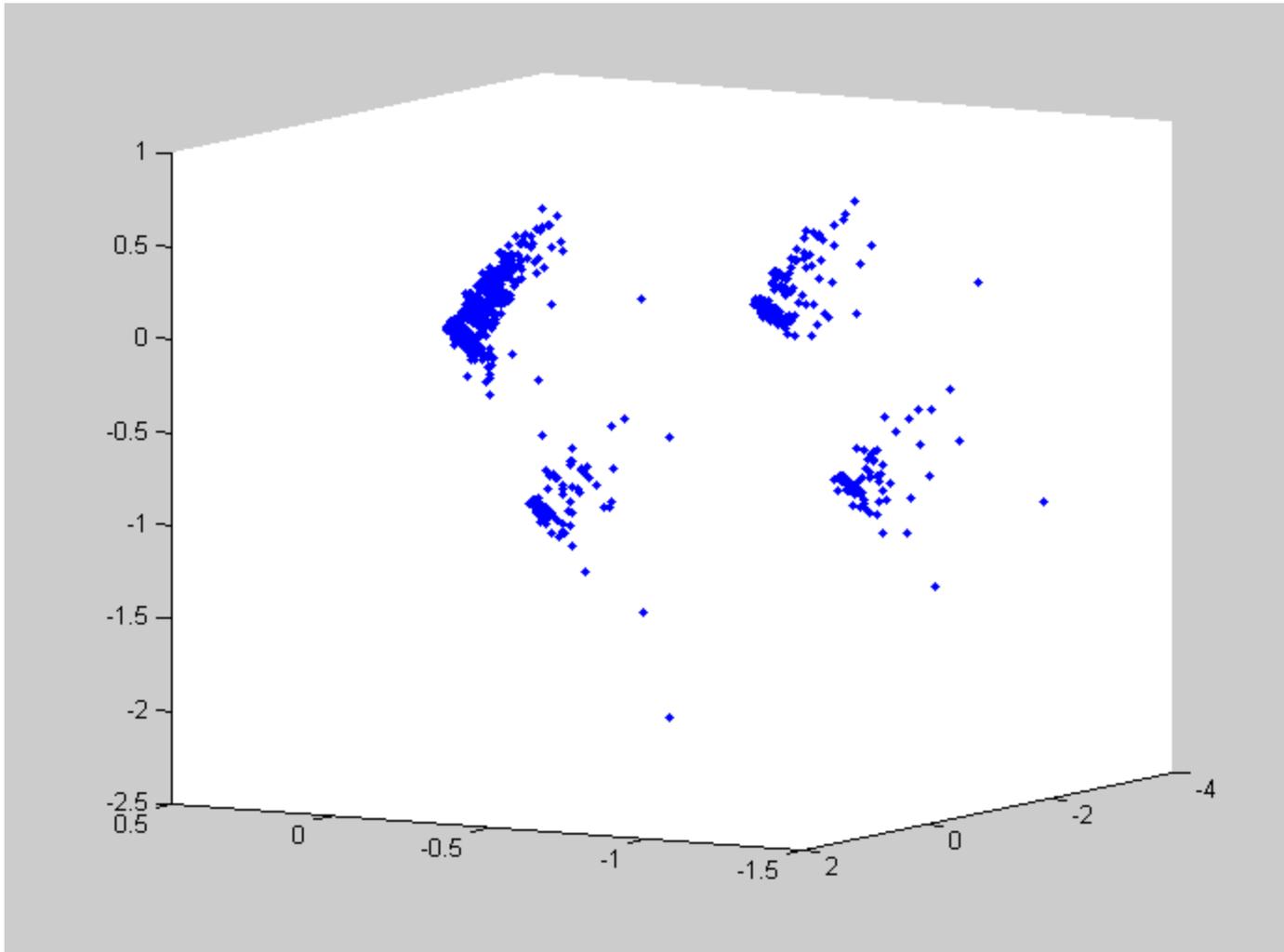
Other datasets – gastronomy
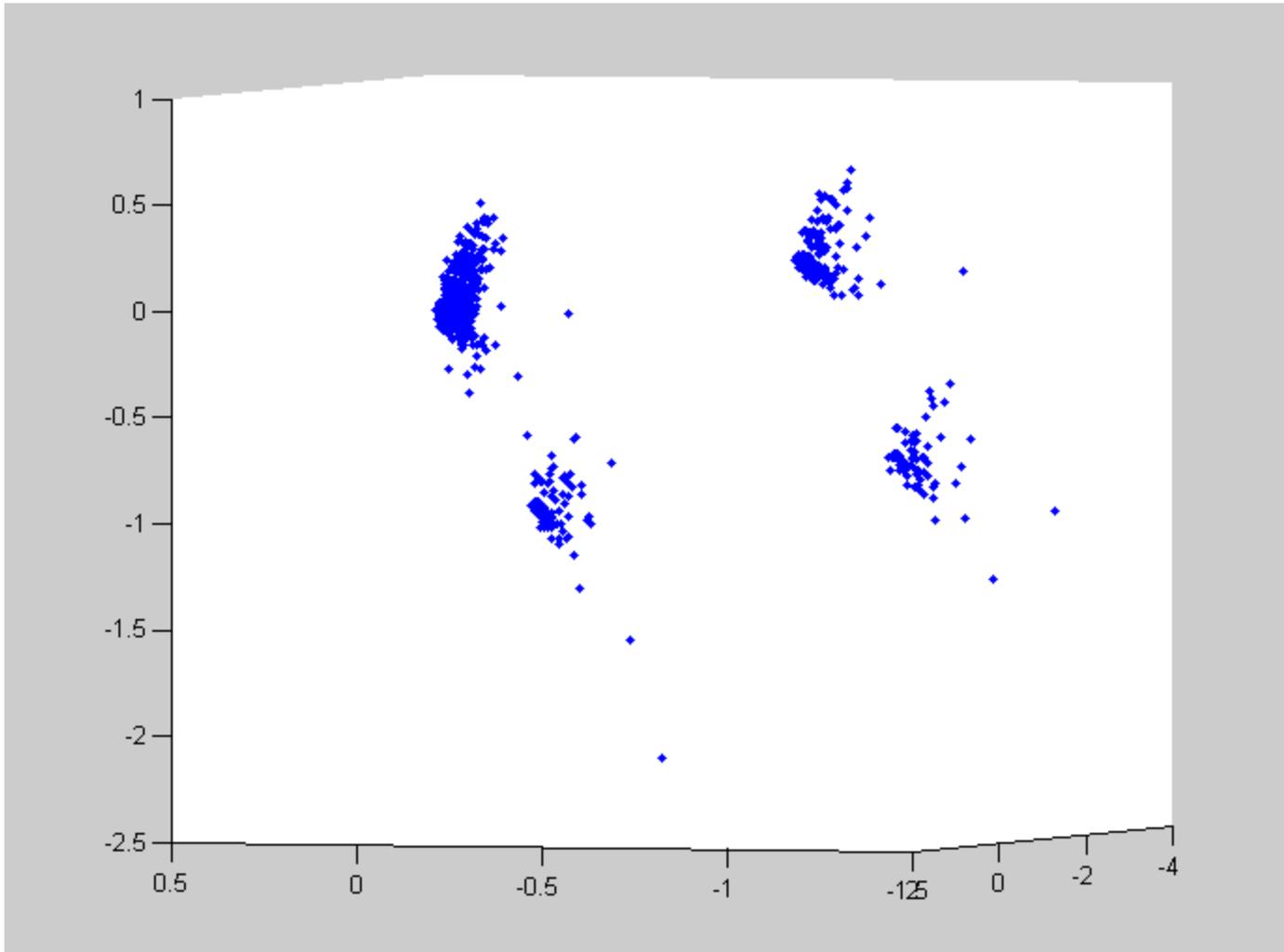
Other datasets – gastronomy
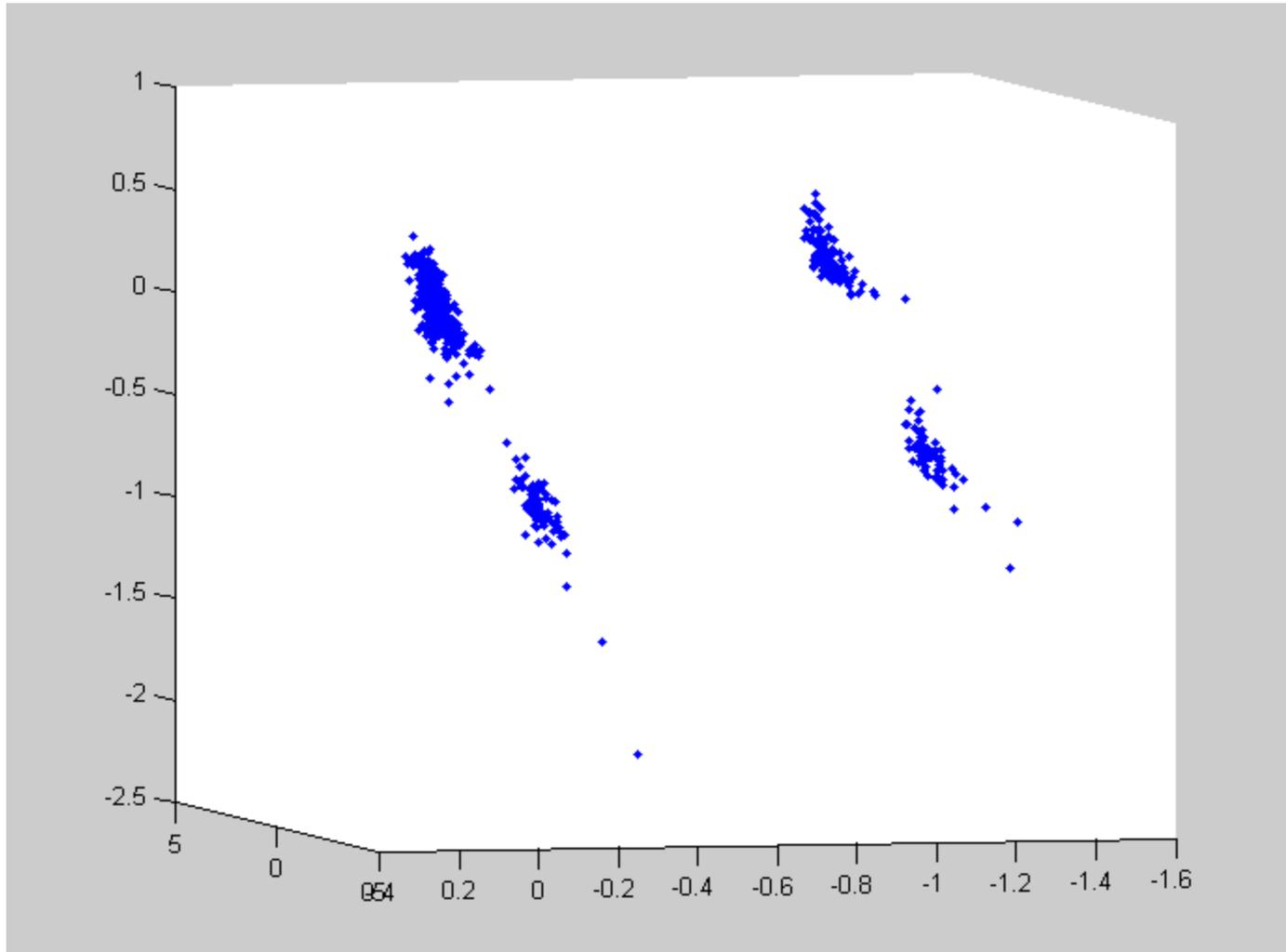
Other datasets – gastronomy
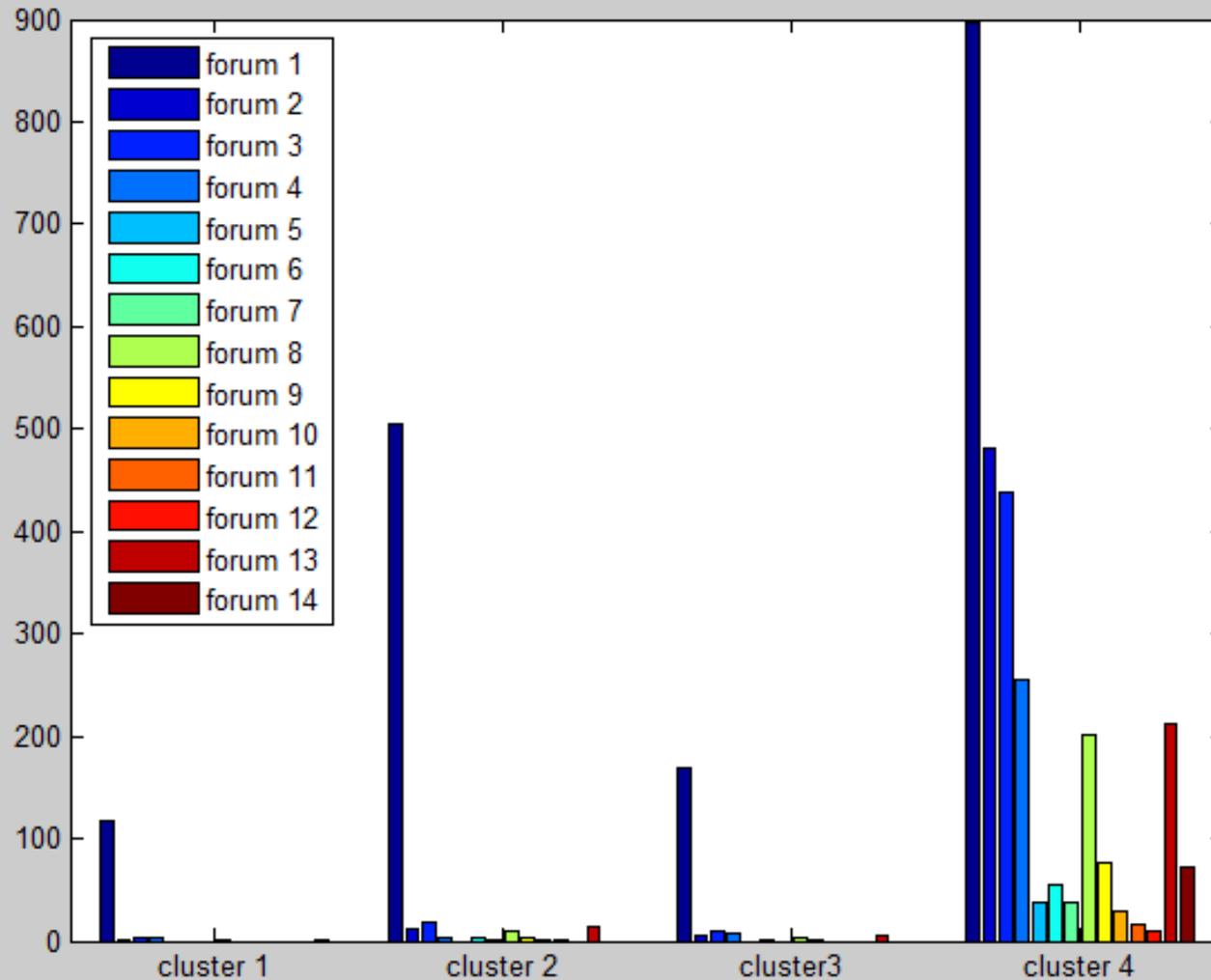
- **Forum Data**
- **Clustering Structure**
- **Topic Modeling**
- **Conclusion**

Topic modeling

Latent Dirichlet Allocation

VS

Bag of Words

LDA

LDA for gastsci

bar graph per week
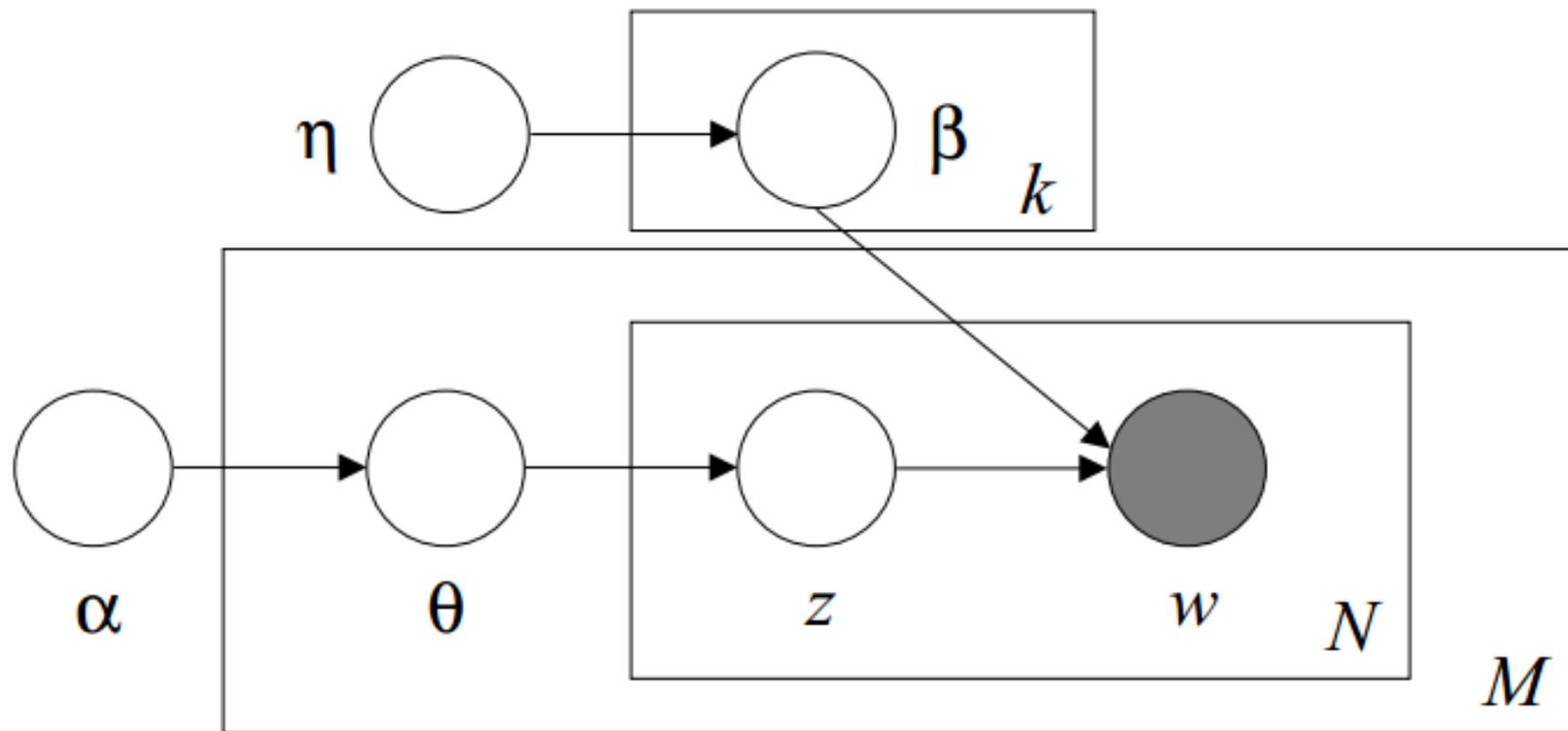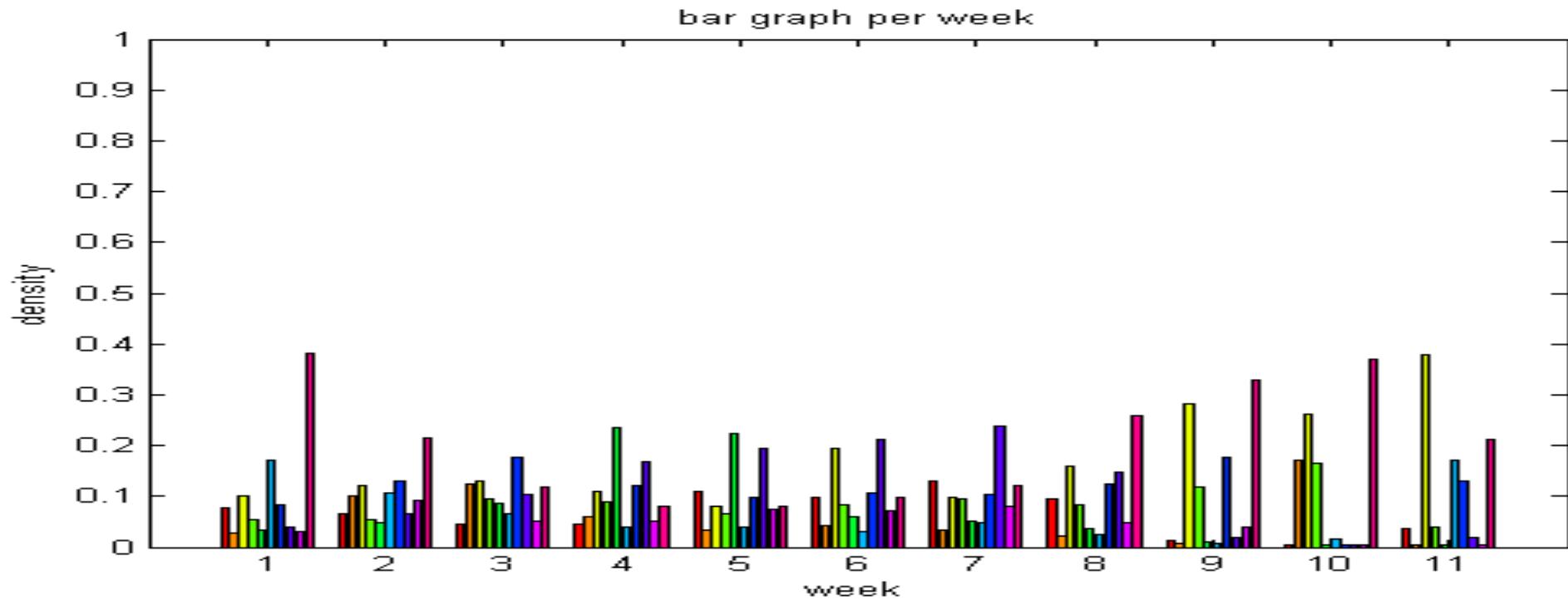
LDA for gastsci

'cook'   'cours'   'scienc'   'food'   'learn'   'love'   'new'   'recip'   'interest'
'cook'   'cours'   'scienc'   'food'   'learn'   'love'   'new'   'recip'   'interest'
'chocol'   'assign'   'would'   'time'   'wine'   'tongu'   'meal'   'question'   'experi'
'food'   'color'   'like'   'think'   'green'   'sweet'   'red'   'fruit'   'sour'
'food'   'color'   'like'   'think'   'green'   'sweet'   'red'   'fruit'   'sour'
'dish'   'add'   'meat'   'use'   'salt'   'cook'   'oil'   'sauc'   'fri'
'dish'   'add'   'meat'   'use'   'salt'   'cook'   'oil'   'sauc'   'fri'
'cook'   'cours'   'scienc'   'food'   'learn'   'love'   'new'   'recip'   'interest'
'cook'   'cours'   'scienc'   'food'   'learn'   'love'   'new'   'recip'   'interest'
'cook'   'cours'   'scienc'   'food'   'learn'   'love'   'new'   'recip'   'interest'
'thread'   'video'   'quot'   'post'   'forum'   'cours'   'lectur'   'week'   'scigast'

LDA for stschina1

bar graph per week

LDA for stschina1

'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'im'   'studi'   'cours'   'name'   'everyon'   'meet'   'chines'   'take'   'hope'
'im'   'studi'   'cours'   'name'   'everyon'   'meet'   'chines'   'take'   'hope'
'video'   'lectur'   'assign'   'week'   'cours'   'submit'   'time'   'download'   'late'
'video'   'lectur'   'assign'   'week'   'cours'   'submit'   'time'   'download'   'late'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'assign'   'grade'   'peer'   'cours'   'essay'   'request'   'regrad'   'evalu'   'word'
'im'   'studi'   'cours'   'name'   'everyon'   'meet'   'chines'   'take'   'hope'
'innov'   'system'   'altern'   'model'   'approach'   'one'   'linear'   'technolog'
'advantag'

Componential Counting Grids

CCG for gastsci: Week 1

make
like
tend

sugar
enhanc
asparagu

sugar
fruit
syrup

bottl
sheep
coke

colour	green	yellow	sour
thought	blue	green	sweet
gener	yellow	brown	brown
orang	black	lemon	exampl
purpl	white	juic	usual
often	fruit	appl	depend
prefer	alway	one	hot
notic	banana	seem	mean
yogurt	mango	although	prevent
cooki	chip	seem	iron
nois	bread	better	metal
popcorn	chose	second	cast

CCG for gastsci : Week 2

soy al 33	quantiti part humbl	problembulli abusefor id	chip nose -
mushroom grow bread	yesterday made sandwich	dish flavor onion	ingredi potato husband
leav mix pie	made spice slice	pepper fresh cucumb	chees combin corn
onion garlic oliv	salad rice ad	tomato sauc green	spici delici chile

con
que
por

burger
indian
rhubarb

seed
potato
crust

add
bean
onion

CCG for gastsci : Week 3

new	scienc	cours	school
curiou	forward	forward	practic
favorit	underr	canada	taught
chemic	scienc	learn	teach
involv	glad	educ	new
creat	chemistri	culinari	excit
happi	famili	cours	year
comfort	behind	learn	new
re	prepar	motiv	wish
know	start	go	hope
stuff	still	sinc	year
haven	next	abl	ago

less	eat	peopl	know
chocol	peopl	know	stuff
gummi	ever	tri	haven
cracker	good	meal	one
bear	piec	mayb	end
satieti	real	doesn	particip
sens	feel	ask	answer
mouth	good	full	question
affect	four	10	get
like	would	test	question
matter	guess	result	sure
abl	surpris	wasn	suppos

CCG for gastsci : Week 4

CCG for gastsci : Week 5

make
like
tend

sugar
enhanc
asparagu

sugar
fruit
syrup

bottl
sheep
coke

worm
said
milkwe

CCG for gastsci : Week 6

feel	ask	answer	got
good	full	question	get
four	10	get	complet

would	test	question	got
guess	result	sure	right
surpris	wasn	suppos	order

flavour	respons	would	red
sensit	option	think	think
tongu	astring	pair	mistak

tongu	certain	colour	green
map	come	thought	blue
area	cocoa	gener	yellow

anyth	english	orang	black
coffe	away	purpl	white
particular	afraid	often	fruit

CCG for gastsci : Week 7

cell	method	cook	home	basic
form	slow	devic	kitchen	knowledg
make	also	insteadi	mani	inform
break	receptor	part	mani	cultur
caus	requir	appli	relat	explain
ph	odor	vide	keep	consid
well	work	mean	far	anyon
process	well	import	ve	simpl
brine	level	creat	area	easi
case	avail	egg	far	ingredi
rather	isn	possibl	long	keep
dessert	need	happen	term	bird

CCG for gastsci : Week 8

food	happi	famili	cours	year
remind	comfort	behind	learn	new
consid	re	prepar	motiv	wish
peopl	know	start	go	hope
know	stuff	still	sinc	year
tri	haven	next	abl	ago
meal	one	week	time	love
mayb	end	first	late	zone
doesn	particip	two	week	two
ask	answer	got	assign	deadlin
full	question	get	grade	date
10	get	complet	submit	view

CCG for gastsci : Week 9

chocol	less	eat	peopl	know
rate	chocol	peopl	know	stuff
candi	gummi	ever	tri	haven
chocol	cracker	good	meal	one
specif	bear	piec	mayb	end
dark	satieti	real	doesn	particip
sensat	sens	feel	ask	answer
mouth	mouth	good	full	question
intens	affect	four	10	get
textur	like	would	test	question
even	matter	guess	result	sure
tast	abl	surpris	wasn	suppos

CCG for gastsci : Week 10

well	work	mean	far
process	well	import	ve
brine	level	creat	area
case	avail	egg	far
rather	isn	possibl	long
dessert	need	happen	term
rather	need	pleas	lectur
qualiti	re	list	pleas
asian	option	help	slide
top	chang	video	back
amaz	problem	help	note
freez	flash	problem	watch

sugar
enhanc
asparagu

sugar
fruit
syrup

bottl
sheep
coke

worm
said
milkwe

CCG for gastsci : all

CCG for stschina1 : Week 1

wrong	use	digit	present
context	histori	en	abil
formul	peopl	later	real
paradigm	thought	shall	gdp
find	opinion	show	term
kuhn	unless	dead	academ
reason	religi	even	would
leav	mani	cant	support
seem	reason	money	countri
understand	peopl	influenc	world
idea	talk	think	nation
believ	live	econom	long
better	like	polit	polici
us	import	bad	other
present	profession	power	industri

CCG for stschina1 : Week 2

east	chemic	interest	anyon	free	go
west	might	regard	chose	journal	wonder
abstract	emerg	caredeba	eastern	almost	guy
		njan			
case	engin	us	curiou	natur	avail
refer	technolog	west	natur	everi	consid
word	manag	get	elsewher	saw	similar
india	usa	world	class	forum	text
oil	industri	area	interest	recommend	coursera
industri	compani	avoid	subject	display	cost
countri	spend	futur	class	tag	see
money	alreadi	exampl	gain	materi	ii
educ	manufactur	suppos	tri	unless	ago

CCG for stschina1 : Week 3

quit	technolog	need	world
take	interest	anybodi	tradit
impact	need	dont	technolog
china	mani	come	probabl
enough	coupl	compani	foreign
go	often	mind	though
mandarin	languag	notic	reason
area	one	also	around
edit	found	understand	thing
learn	learn	sourc	subject
form	hope	happen	project
done	googl	greater	like

curiou	natur	avail	mean
natur	everi	consid	altern
elsewher	saw	similar	research
class	forum	text	public
interest	recommend	coursera	manufactur
subject	display	cost	right
class	tag	see	top
gain	materi	ii	data
tri	unless	ago	activ
idea	innov	www	critiqu
live	link	core	written
peopl	url	andi	emili

CCG for stschina1: Week 4

limit	somewhat	often	western
offer	give	encourag	japanes
idea	viewpoint	wonder	gave

state	blogspot	email	read
two	phone	address	formal
html	au	banker	consid

professor	en	intern	asian
http	lo	via	idea
brand	hao	upcom	climat

student	el	alguien	beij
jose	lo	eso	toni
marathon	de	es	anthoni

experi
poor
individu

scientist
must
valid

scientif
scienc
definit

CCG for stschina1: Week 5

histori	realli	point	say
realli	educ	onlin	like
respect	rule	fun	point
would	keep	ever	us
like	idea	human	american
industri	thing	friend	peopl
alway	well	name	chines
find	china	chines	china
compani	life	whole	french
write	china	chines	china
western	want	china	chines
enough	compani	sever	year

exampl	use	wonder	transfer
lockin	made	excit	dont
assign	thought	learn	regard
limit	somewhat	often	western
offer	give	encourag	japanes
idea	viewpoint	wonder	gave
state	blogspot	email	read
two	phone	address	formal
html	au	banker	consid
professor	en	intern	asian
http	lo	via	idea
brand	hao	upcom	climat

CCG for stschina1 : Week 6

exampl	discuss	might	mean
challeng	like	done	word
first	could	set	answer
someon	class	found	dog
also	differ	fantast	main
two	curios	pure	littl
histori	realli	point	say
realli	educ	onlin	like
respect	rule	fun	point
would	keep	ever	us
like	idea	human	american
industri	thing	friend	peopl

exampl	use	wonder	transfer
lockin	made	excit	dont
assign	thought	learn	regard
limit	somewhat	often	western
offer	give	encourag	japanes
idea	viewpoint	wonder	gave
state	blogspot	email	read
two	phone	address	formal
html	au	banker	consid
professor	en	intern	asian
http	lo	via	idea
brand	hao	upcom	climat

CCG for stschina1 :
Week 7

great	human	way	time
problem	exist	theori	mathemat
concern	social	differ	part

unfortun	polit	sound	physic
hand	possibl	import	etc
staff	mind	also	studi

realiti	human	human	sociolog
perform	activ	use	human
insight	give	hear	biolog

finland	howev	later	develop
complex	extend	propos	philosophi
decid	develop	classifi	dr

much	big	plan	idea
could	interest	done	experi
chines	local	even	would

CCG for stschina1 : Week 8

social	realiti	human	human
truth	perform	activ	use
make	insight	give	hear

societi	finland	howev	later
chang	complex	extend	propos
construct	decid	develop	classifi

societi	give	phd	field
slide	liber	commun	book
continu	idea	worker	peopl

softwar	base	time	work
attract	mistak	100	year
suggest	make	fix	tech

may	depart	com	professor
agenc	manag	www	http
region	particularli	public	brand

product	get	earthquak	student
guess	american	20	jose
bank	week	player	marathon

long	question	post	everi
smart	qualiti	doubt	philosophi
pr	increas	work	low

CCG for stschina1 : Week 9

anyon	free	go	busi
chose	journal	wonder	research
eastern	almost	guy	identifi
curiou	natur	avail	mean
natur	everi	consid	altern
elsewher	saw	similar	research
class	forum	text	public
interest	recommend	coursera	manufactur
subject	display	cost	right
class	tag	see	top
gain	materi	ii	data
tri	unless	ago	activ

somewhat	often	western	perhap
give	encourag	japanes	spain
viewpoint	wonder	gave	industri
blogspot	email	read	professor
phone	address	formal	commun
au	banker	consid	far
en	intern	asian	nick
lo	via	idea	set
hao	upcom	climat	perhap
el	alguien	beij	school
lo	eso	toni	thoma
de	es	anthoni	condit

CCG for stschina1: all

- **Forum Data**
- **Clustering Structure**
- **Topic Modeling**
- **Conclusion**

Possible future work

- post recommendation
  - click trajectory
  - social network
- peer grading
  - click stream and forum
- assigning reviewers
  - implicit social networks