Privacy and Security in Online Social Networks Department of Computer Science and Engineering Indian Institute of Technology, Madras

Lecture – 28 Gephi Network Visualization

Gephi is an open source network analysis and visualization software. Gephi is widely used in a number of research projects in academia, journalism, digital humanities etcetera. Gephi can also input data of social networks like Facebook and Twitter and generate graphs and clusters.

(Refer Slide Time: 00:57)



In this tutorial, we will learn in detail how to use gephi. Let us begin by understanding how to load data in gephi. Gephi can read many file **formats** including gml, graphml, pajek net file and uci net, dl files. You can open any of those by simply going to file and then clicking on open.

(Refer Slide Time: 01:04)

Overview	Data Laboratory	Preview		
Preview Settings (3)		0	60 Preview ()	010
		1		
		0)		
	Manage rende		Open	
		twecol	II-master	
	cho Properties-	arupama, Jaga dat arupama, Jaga ng arupama, Jaga ng arupama, Jaga ng arupama, Jaga ng arupama arupama, Jaga ng arupama arupa arupa arupama arupa aru	Sunday, 12 June, 2016 8:11 PM Monday, 13 June, 2016 1:20 2MM Monday, 13 June, 2016 1:20 2M Monday, 13 June, 2016 1:20 2M Sunday, 13 June, 2016 1:21 PM Sunday, 13 April, 2016 1:25 PM Sunday, 12 June, 2016 1:25 PM	
Preview ratio: 0 bpon: SvG/PDF/PNG		File Format	All Files CSV Files (* Cxv) Edge List (* Cdges) DL Files (DKET) (* d) Cophily Files (* Cdges) CDF Files (CdES) (* cget) CDF Files (* cget) *	

You can also look at the complete list of supported file **formats** by clicking on the drop down menu. Gephi can also read excel and csv files.

(Refer Slide Time: 01:28)



	edges.csv	UNREGISTERE
nodes.csv x edges.csv x		
1 Source, Target, Type, Id, Weight, Average Degree		
2 1,3,Undirected,1,1.0,1.0		85
3 2,10,Undirected,8,1.0,1.0		
4 3,4,Undirected,2,1.0,1.0		
5 4,5,Undirected,3,1.0,1.0		
6 5,10,Undirected,4,1.0,1.0		
7 8,2,Undirected,7,1.0,1.0		
8 8,9,Undirected,11,1.0,1.0		
9 10,1,Undirected,5,1.0,1.0		
10 10,6,Undirected, 10, 1.0, 1.0		
11 10,7,Undirected,9,1.0,1.0		
12 10,8,Undirected,6,1.0,1.0		
13 10,11,Undirected,12,1.0,1.0		
14 10,13,Undirected,14,1.0,1.0		
15 12,9,Undirected,13,1.0,1.0		
16		

To import data from csv, you will usually need to prepare two files, one containing nodes and their attributes and the other containing an edge list and edge attributes. The csv file containing nodes needs to include a column named id containing unique node ids. The edge list csv should have columns titled source and target containing node ids of the start and end node for each edge. In addition, you can also include a column called type indicating the type of each edge that is directed or undirected.

(Refer Slide Time: 02:22)



We already learnt about graphml format in the previous tutorial on social network analysis. Let us open an already existing gml file for this demo. If the data is in correct format, gephi will give you a summary of the data, which is the number of nodes, the number of edges and other information about the edge. Click on ok to successfully import the file.

(Refer Slide Time: 02:36)

Overview	Data Laboratory Preview			\mathcal{X}
Workspace 1 🔘				
Preview Settings 😳		💿 60 Preview 🕄		
Presets		8		
Default				
	Settings Manage renderers			
v Nodes				
Border Width	1.0			
Border Color	custom (0.0.0)			
opacity	100.0			
v Node Labels				
Show Labels				
Font	Arial 12 Plain			
Proportional size	v			
Color	custom (0,0,0)			
Shorten label				
Max characters	30			
Dutline size	0.0			
Dutline color	custom [255,255,255]	-		
Dutline opacity	80.0			
Вох				
Box color	parent	-		
Box opacity	100.0			
v Edges			h	
	-			
Preview ratio: 100%				
		Referch		
		W MOTON		
txport: SVG/PDF/PNG		Background	Reset zoom - +	

(Refer Slide Time: 02:44)

					Gephi 0	1.9.1 - Proje	ct 1							
Overview Data I	aborator	y 💌	Preview	_									X	11
Workspace 1 (C)													110	14.5.8
witter Streaming Importer 🔘 🔅	🗈 Data	Table ()												19.10
- Set Credentials :	Nodes	Edges 0 0	onfiguration	Q A	dd node 🕀 Ada	i edge 🛗 S	earch/Replac	e 💽 Import	Spreadsheet	🐮 Export	table	More actions	~ Filter:	Id
Define the surgery	ld *	Label	Interval	friends	user_id	žte –	followers	ffr	listed	shape	statuses	type	file	image
- Dennie the query .	72.0	virgilical		186.0	24543879	0.3929	840.0	4.5161	33.0	triangle-up	\$57.0	friends	anupama	/Users/A.
Load Query File Save Query	194.0	mstrohm		463.0	14065835	1.0	1560.0	3.3693	156.0	triangle-up	2800.0	friends	anupama	/Users/A_
	148.0	zelandiya		\$30.0	35714839	1.0311	2541.0	4.7943	262.0	triangle-up	3339.0	friends	anupama	/Users/A.
Monds to follow Us	8.0	dipestitech		451.0	84812171	0.719	153.0	0.355	11.0	triangle	142.0	triends	anupama	/Users/A.
	155.0	Jawbone		300.0	26276396	0.2561	63444.0	211.48	1498.0	triangle-up	19979.0	friends	anupama	/Users/A.
Add	140.0	CERCABILIT		0.0	2229460	0.2174	92.0	0.0	2.0	triangle	103.0	triends	anupama	/Users/A_
Biliti	7.0	hemankla		286.0	1185857_	0.289	546.0	1.2098	10.0	triangle-up	8/4.0	mends	anupama	/Users/A.
Delete	95.0	saravade		0.0	3/903380	0.2529	1403.0	295.8555	37.0	triangle-up	1980.0	menos	anupama	/Users/A.
	122.0	pycoders		2.0	9697370	0.1695	1255.0	11951.0	33.0	triangle-up	3293.0	friends	anupama	/Users/A.
	23.0	namanpu		105.0	40130310	0.1003	1303.0	0.1332	23.0	triangle-up	4373.0	friends	anupama	/USEIS/A
	33.0	meana_g		215.0	4313137	0.00	216003.0	1006 2328	6542.0	triangle	130.0	friends	anupama	/Users/A.
	178.0	coursera		513.0	5320332	0.6517	1225.0	2 6 2 8	97.0	triangle-up	7570.0	friends	anupama.	/Users/A
	192.0	ibrin		202.0	3143850	0.409	2005.0	5 1018	82.0	triangle-up	5876.0	friends	anupama	/lisers/A
	12.0	achich cu		47.0	38433184	0.2222	450.0	9 5745	10.0	triangle-up	1136.0	friends	aminama	/licers/A
	34.0	Enternant -		75.0	3314738	0.000	101.0	1 3467	10.0	triangle-up	56.0	friends	anupama	/lisars/A
	56.0	chravanh		575.0	55520719	0.0234	4344046.0	7554 8626	10149.0	triangle-up	18963.0	friends	aninama	/licers/A
	173.0	redditindia		682.0	3036331	0 1019	24045.0	35,2566	245.0	triangle-up	39161.0	friends	anunama	/lisers/A
	2.0	Rahy		1531.0	15234886	0 382	68280.0	44 5983	2608.0	triancle-un	16079.0	friends	aminama	/licers/A
8 - Choose Network Logic to apply :	6.0	twitter		144.0	783214	0.0163	\$\$20112	383341	90100.0	trianole-up	2881.0	friends	anunama	/lisers/A
	135.0	kkmookhev		401.0	3322853	0.3143	891.0	2.2219	28.0	triangle-up	681.0	friends	anupama	/Users/A
	171.0	WSDMSocial		113.0	3107101	0.5019	538.0	4.7611	27.0	triangle-up	54.0	friends	anupama	/Users/A.
	96.0	mrowebot		869.0	13092722	0.9338	1103.0	1.2693	103.0	triangle-up	5650.0	friends	anupama	/Users/A.
	***	n				A A334	****	1.0000	***		122.4		100000	41. 14
			1											
Connect														
connect			Add		Merge	Delete	Clear	Cop	y data to	Fill col	umn alue v	Duplicate		
			Contenting	-	Committee ()	CO-GIIIII -	corunn	ourer	summittee .	mold d to		Committee .		_

Now, we can view the data by going to the data laboratory tab in gephi. On right, you can see that there is a data table, which has two tabs - nodes and edges. The nodes tab contains the data table for nodes and node attributes.

		inger tretter	_					11
Workspace 1 (C)								1000
witter Streaming Importer 🛈 🔅	🗅 Data Table 🔘	(i						1010
I - Set Credentials :	Nodes Edges	© Configuration	🚯 Add node 🕀 Add edg	e 🖁 Search/Replace	mport Spreadsheet	🐑 Export table 🛔	More actions Y Filter.	Sourc
Defendence.	Source	Target	Type	id	Label	Interval	Weight	
- Define the query :	72.0	181.0	Directed	0			2.681358	
Load Query File Save Query I	72.0	68.0	Directed	1			0.245086	
	72.0	75.0	Directed	2			2.148089	
Words to follow Us	72.0	177.0	Directed	3			1.709412	
	72.0	194.0	Directed	4			0.92562	
Add	72.0	192.0	Directed	5			0.998979	
	72.0	6.0	Directed	6			0.847353	
Delete	72.0	127.0	Directed	7			3.181063	
	72.0	36.0	Directed	8			1.953829	
	72.0	69.0	Directed	9			0.231794	
	72.0	85.0	Directed	10			0.145911	
	72.0	27.0	Directed	11			2.289996	
	72.0	58.0	Directed	12			0.157128	
	72.0	98.0	Directed	13			0.0918	
	72.0	82.0	Directed	14			0.037357	
	194.0	172.0	Directed	15			1.033954	
	194.0	170.0	Directed	16			0.650754	
	194.0	150.0	Directed	17			0.15679	
	194.0	93.0	Directed	18			0.973746	
- Choose Network Logic to apply :	194.0	112.0	Directed	19			3.643472	
	194.0	181.0	Directed	20			2.681358	
	194.0	85.0	Directed	21			0.145911	
	194.0	37.0	Directed	22			0.209978	
			~ ~ 1					
						144		
			•	1 1	1	1	II.	
Connect		Add	Merge	Delete Clear	Copy data to	Fill column	Duplicate	

(Refer Slide Time: 02:59)

And the edges tab contains the data table for edges and edge attributes. You will also notice an import spreadsheet button which is used to get nodes and edges data directly from excel or csv files. The panel at the bottom can help you manipulate data columns and modify the data. You can also change individual values directly by clicking on them and typing their new values.

(Refer Slide Time: 03:28)

					Gephi 0	.9.1 - Proje	ct 1						G. 1000	
😧 Overview 📋 Data L	aboratory		Preview		1								Y	í
Workspace 1 0						_	_	_	_	_	-	_	~	1
witter Streaming Importer ()	📃 Data T	able O												
1 - Set Credentials :	Nodes 1	Edges 🛛 Co	ofiguration	Q As	id node 💽 Add	edge 🛔 S	earch/Replac	e 🛃 Import	Spreadsheet	t 📳 Export t	able	More action	s ~ Filter:	Id
	Id	Label	Interval	friends	user_id	Hr.	followers	ffr	listed	shape	statuses	type	file	image
2 - Define the query :	72.0	virgiloal		188.0	24543879	0.3929	\$40.0	4.5161	33.0	triangle-up	557.0	friends	anupama	/Users/A
Load Ouery File Save Ouery	194.0	mstrohm		463.0	14065835	1.0	1560.0	3.3693	156.0	triangle-up	2800.0	friends	anupama	/Users/A
	148.0	zelandiya		\$30.0	35714839	1.0311	2541.0	4.7943	262.0	triangle-up	3339.0	friends	anupama	/Users/A
Words to follow Us	8.0	dipeshtech		431.0	84812171	0.719	153.0	0.355	11.0	triangle-d	142.0	friends	anupama	/Users/A
	155.0	Jawbone		300.0	26276396	0.2361	63444.0	211.48	1498.0	triangle-up	19979.0	friends	anupama	/Users/A
Add	140.0	CERCatIITD		0.0	2229460	0.2174	92.0	0.0	2.0	triangle-d	103.0	friends	anupama	/Users/A.
	7.0	hemankla		286.0	1183857	0.289	346.0	1.2098	10.0	triangle-up	874.0	friends	anupama	/Users/A
Delete	95.0	saravade		6.0	37903380	0.2529	1463.0	243.8333	37.0	triangle-up	1980.0	friends	anupama	/Users/A.
	61.0	pycoders		2.0	4847570	0.3744	23902.0	11951.0	895.0	triangle-up	\$293.0	friends	anupama	/Users/A.
	133.0	namanpu		245.0	96136510	0.1685	1365.0	5.5714	23.0	triangle-up	4573.0	friends	anupama	/Users/A.
	33.0	medha_g		105.0	4515157	0.0	14.0	0.1333	0.0	triangle-d	138.0	friends	anupama	/Users/A
	18.0	coursera		315.0	3520532	0.2064	316992.0	1006.3238	6543.0	triangle-up	2228.0	friends	anupama	/Users/A.
	178.0	gianiuca		508.0	\$1169895	0.6517	1335.0	2.628	87.0	triangle-up	7579.0	friends	anupama	/Users/A.
	193.0	ilhsin		393.0	2342859_	0.409	2005.0	5.1018	82.0	triangle-up	5836.0	friends	anupama	/Users/A.
	12.0	ashish_su		47.0	38433184	0.2222	450.0	9.5745	10.0	triangle-up	1136.0	friends	anupama	/Users/A.
	34.0	soumyav		75.0	3314738	0.099	101.0	1.3467	1.0	triangle-up	56.0	friends	anupama	/Users/A.
	56.0	shreyagh		575.0	55520719	0.0234	4344046.0	7554.8626	10149.0	triangle-up	18963.0	friends	anupama	/Users/A.
	173.0	redditindia		682.0	3036331	0.1019	24045.0	35.2566	245.0	triangle-up	39161.0	friends	anupama	/Users/A.
	2.0	Bitly		1581.0	15234886	0.382	68280.0	44.5983	2608.0	triangle-up	16079.0	friends	anupama	/Users/A.
- Choose Network Logic to apply :	6.0	twitter		144.0	783214	0.0163	5520112	383341	90100.0	triangle-up	2881.0	friends	anupama	/Users/A.
	135.0	kkmookhev		401.0	3322853	0.3143	891.0	2.2219	28.0	triangle-up	681.0	friends	anupama	/Users/A.
	171.0	WSDMSocial		113.0	3107101	0.5019	538.0	4.7611	27.0	triangle-up	\$4.0	friends	anupama	/Users/A.
	96.0	mrowebot		869.0	13092722	0.9338	1103.0	1.2693	103.0	triangle-up	\$650.0	friends	anupama_	/Users/A.
	85.0	DanielMR		325.0	14817871	0.8238	522.0	1.6062	43.0	triangle-up	430.0	friends	anupama	/Users/A.
	182.0	jvitak		440.0	94384796	0.7991	1877.0	4.2659	150.0	triangle-up	7871.0	friends	anupama	/Users/A.
		- arar		****			*****	** ****		and the second				
			I		0	i	п		lt			ii.		
Connect			Add		Merge	Delete	Clear	Cop	y data to	Fill colu	ma	Duplicate		

For instance, let us go to the nodes table, and change the number of friends of the first row from 186 to 188.

(Refer Slide Time: 03:45)

		Gephi 0.9.1 - Project	1	VY
Overviers	boratory Preview			
Appearance ()		Graph (2)		
Nodes Edges	€ ⊜ A T	Mouse selection Configurei		
Unique Attribute		4		
Choose an attribute	0			
	1	0		1
		1		
		2		
	a Antiv	10	N. Anolis	
Layout O		1	A state in	
Choose a layout	P	1	Crantale .	
		×	The second second	
0) for	0		
		2	ALC: NOT THE	
			N. C. A. S.	
<no properti<="" td=""><td>E2</td><td></td><td></td><td></td></no>	E2			
		P		
	0			
		A		
P Presets., Reset		TE T X Z	-()	- 1 2 4

Now, let us look at the over view tab to calculate the network measures and set up network visualization. Once you click the over view tab, you will notice the visualization

of your data in the blank space in the middle. Over view tab has many options like appearance, layout and a several panels to change the color, size and attributes of the visualization. Let us look at each one of them one by one. The appearance tab on the top left lets you change the node edge color and size based on categorical and continuous attributes. You can select various visual properties by going through the options.

The layout tab can help you select and customize one of the available layout algorithms. Various options on the left of the visualization help you with interactive selection of nodes and edges. And also lets you change the size and color of the nodes and edges manually. The edge pencil and the node pencil **can** also help you add nodes and edges manually. The options towards the bottom let you resize the color and size of the characteristics of nodes and edges.

(Refer Slide Time: 05:04)



The panel **at** the bottom can be expanded by clicking on **the** small up arrow on bottom right. This panel can help you change the color, size and other characteristics applying to all nodes edges and labels.

(Refer Slide Time: 05:33)



Now, let us look at the filters and statistics tab on the right. You can apply filters to select specific nodes and edges from your network. The filters are applied to dragging and dropping the specific filter in the query space. There are several filters based on attributes like, equal can help you select elements with particular attribute values.

(Refer Slide Time: 06:02)



The partition filter can help you select the different levels of categorical attributes.

XI Data Labo atory Preview € 6 д т Q 1 F Drag I \$ to Arial-BoldWT, 32 H: T 1.7

ohi 0.9.1 - Pro

(Refer Slide Time: 06:14)

With the range filter, you can select nodes and edges with attribute values in a particular range. For instance, followers between 100 to 1000. Filters based on edges allow you to select edges with different properties like edge weight, edge type and self loops.

(Refer Slide Time: 06:32)

		Gephi 0.9.1 - Project 1	
Overview	Data Laboratory Preview		Yán (
Workspace 1 ()			
Appearance ()		Graph (C)	Filters O Statistics
Nodes Edges	● ⑤ A T	Gp Mouse selection (Configure)	Reset B B An
Unique Attribute		4	T Mutual Edge
Choose an attribute	0		> Coperator
	(-	0	T Degree Range
		1	T Ego Network
			T Has Self-loop
-		a second of	T In Degree Range
	a 👌 Apply		T Mutual Degree Range
Layout ()			T Out Degree Range
Choose a layout	2	Contraction of the second seco	Saved queries
	1.14	- 5.8 BAR	Queries Queries Queries Queries Queries Drag filter here
, i i i i i i i i i i i i i i i i i i i		·	
<no f<="" td=""><td>Properties></td><td></td><td></td></no>	Properties>		
		0	
		A A	
			b The
* / Fresets Reset			And and a second s

The topology filter allows **for** selection based on the network structure like components, k cores, degree, ranges etcetera.

(Refer Slide Time: 06:43)

		Gephi 0.9.1 - Project 1	
Overview Data Laboratory	Preview		YÃO.
Workspace 1 0			
Appearance ()	6	Graph (C	Filters 🗘 Statistics
Nodes Edges	● ① A TT	Gp Mouse selection Configure)	Reset
Unique Attribute		4	Mutual Edge Self-Loon
Choose an attribute	8		🙀 🤷 Operator
	6	0	INTERSECTION MASK (Edges)
		1	T NOT (Edges)
			T UNION
		a manth of	Topology
	= Apply	2º	T Ego Network
Layout ()	9		Giant Component Has Self-Joon
Choose a layout	0		🕆 In Degree Range
		× Line Li	Queries Oran fiber have
0) tur	0	Condy mer mire
		2	
		. the same the	
«No Properties»			
		<u>A</u>	
		<u>A</u>	
P Instein Inst		👔 🖬 🕆 🍸 🔍 🚡 🕘 🗛 🗛 Arial-BoldMT, 32 🕥 🌢) Thur

The operator allows you to combine other filters in different ways. For instance, find your intersection, union, compliment etcetera.

(Refer Slide Time: 06:57)

		Gephi 0.9.1 - Project 1		
Overview Data	Laboratory Preview		Y	ĬΛ
Workspace 1 0			7,6	4 4 4
Appearance ()	•	Graph ()	A Pitters Statistics ()	6
Nodes Edges		G Mouse selection (Configure)	Settings	
Unique Attribute		8	C Network Overview	
Character and the se		n	Average Degree	Run 0
Choose an attribute			Avg. Weighted Degree	Run 0
		M	Network Diameter	Run 0
		1	Graph Density	Run 0
		2	Modularity	Run 0
	a Atab	10 Mit Adams	PageRank	Run 0
		/ Nation	Connected Components	Run 0
rayour ()			E Node Overview	
Choose a layout	9		Avg. Clustering Coefficient	Run 0
		SALA	Eigenvector Centrality	Run 0
0	P Rett	0	Edge Overview	
		2	Avg. Path Length	Run 0
		a la factoria de la f	B Dynamic	
			# Nodes	Run 0
<no prope<="" td=""><td>rties></td><td></td><td># Edges</td><td>Run 0</td></no>	rties>		# Edges	Run 0
			Degree	Run 0
		P	Clustering Coefficient	Run 0
			25	
		A		
P. Presets Reset		TET A	A- Arial-BoldMT, 32 🌕 🏟	

Now let us look at the statistics tab, which is used to calculate the network node edge statistics. Once you click on this statistics tab, you will be able to see the different measures whose value can be calculated by clicking on the individual run button. For our data, let us try calculating the value of these measures.



(Refer Slide Time: 07:24)

Let us begin by computing the average degree. Once you click on run, there will be a degree report generated, which will tell you the average degree. In our case, it is 17.577 and also give you a degree distribution graph where the x-axis is the value of the degree distribution, and the y value is the number of nodes with that particular degree distribution. It will also give specific reports of the in-degree and out-degree values and corresponding distributions. Next, let us run the average weighted degree. In our data, the weight of the edges is one, therefore, the graph is a straight line.

(Refer Slide Time: 08:11)

Workspace 1 0					4.12
Appearance ()	6	Graph ()	101131 0	Filters Statistics ()	
Nodes Edges	 О А т 	G Mouse selection (Configure)		Settings	
Unique Attribute		4		Network Overview	
Phone as and as		n		Average Degree	17.577 Run
Undose an attribute		Graph Distance settings		· · · · Weighted Degree	13.335 Run
	-			work Diameter	Rum
	The average graph-distance	e between all pairs of nodes. Connected n	odes have graph distance 1. The	th Density	Run
	diameter is the longest gra	ph distance between any two nodes in the	network. (i.e. How far apart are	Iularity	Run
	the two most distant hode:	2/1		Rank	Run
	Directed	Normalize Cent	ralities in (1) 11	narted Components	Run
ayout O	Directed	- Homaize Cox	idnines in [0,1]	ede Oranjan	640
Choose a layout	O Undirected			Clustering Coefficient	Run
	Betweenness Centrality: Mei	asures how often a node appears on short	est paths between nodes in the netw	ork, master Centrality	Run
0	Closeness Centrality: The	e average distance from a given starting n	Invector Centrality	Nati	
	Eccentricity: The	distance from a given starting node to th	e farthest node from it in the netwo	k. Path Length	Run
				vnamir	
			Cancel OK	ades	Run
«No Propertie				amax/085	Run
				Degree	Run
				Chanterine Carlledon	
				clustering coefficient	Nan
		<u>A</u>			
		A			

Next, let us look at the network diameter. Network diameter is the average distance between all pair of nodes. We will select the directed option since we have a directed network, and click on OK.

(Refer Slide Time: 08:26)

Workspace 1 O	0 0 0 HTML Report			
Appearance () Nodes Edges	Graph Distance Report		tings	1
Unique Antibute	Parameters: Network Interpretation: directed Results: Diameter: 9 Facials: 0 Average Pail length 3.0906221625529464 Betweenness Centrality Distribution		Network Overview erage Degree 17.57 9. Weighted Degree 13.33 twork Diameter 9 9h Density dularity peRank nected Components	7 Run 6 5 Run 6 9 Run 6 Run 6 Run 6 Run 6 Run 6
Choose a layout	50 50 50 50 50 50 50 50 50 50 50 50 50 5		Node Overview 9. Clustering Coefficient envector Centrality Edge Overview 9. Path Length 3.09 Dynamic Lodes dges gree sterring Coefficient	Run (Run (Run (Run (Run (Run (Run (
	An Print La Copy Save	Close		

This **will** generate graph distance report which says that the average network diameter is 9.

(Refer Slide Time: 08:41)

Workspace 1 O				N 89 1	1.10
Appearance ()	0	Graph (C)	ICEICI IN	Filters Statistics ()	
Nodes Edges	T A G .	G Mouse selection Configure)		Settings	
Unique Attribute		4		Network Overview	
Choose an attribute	A			Average Degree 17.5	77 Run
Choose an annouce	•	8		Avg. Weighted Degree 13.3	35 Run
	0.00	Density settings		Network Diameter	9 Run
	Density			Graph Density	Right
	Measure graph h	s how close the network is to complete. A complete as all possible edges and density equal to 1.		Modularity	Run
				PageRank	Run
	O Directe	d		Connected Components	Run
Layout 0		cted		Node Overview	
Choose a layout				Avg. Clustering Coefficient	Run
				Eigenvector Centrality	Run
0, 				Edge Overview	
				Avg. Path Length 3.0	91 Run
		Cancel	OK	🗉 Dynamic	
All Deservations				# Nodes	Run
<ree propertiess<="" td=""><td></td><td>-</td><td></td><td># Edges</td><td>Run</td></ree>		-		# Edges	Run
				Degree	Run
		9		Clustering Coefficient	Run
		÷			

Now let us **compute** the graph density. The graph density measures how close the network is to complete. A complete network means that all the nodes in the graph are connected to each other. Click on OK.

(Refer Slide Time: 09:02)

litterbrane 1.0	Constantion and a second secon	M
Appearance O	O HTML Report	iters Statistics 🔿
Nodes Edges	Graph Density Report	tings
Unique Attribute		Network Overview
and a second		erage Degree 17.577 Run
Choose an attribute	Parameters:	g. Weighted Degree 13.335 Run
	Network Interpretation: directed	twork Diameter 9 Run
	Results	uph Density 0.046 Run
	Daariby 0.046	dularity Run 1
	Density. 0.040	geRank Run
		nnected Components Run
ayout 🖸		Node Overview
Choose a layout		g. Clustering Coefficient Run
		envector Centrality Run
2		Edge Overview
		g. Path Length 3.091 Run
		Dynamic
		4odes Run
<no properties<="" td=""><td>Properties></td><td>idges Run</td></no>	Properties>	idges Run
		gree Run
		astering Coefficient Run
	Arint Copy y Save	Close

And the graph density in our case is 0.046.

(Refer Slide Time: 09:07)

Overview Data Labo	ratory Preview				XĂ	Ń
Workspace 1 ()						1
Appearance ()	e	Graph 🖸	(0110) E	Fitters Statistics ()		10
Nodes Edges	• <u>0</u> <u>A</u> 1	G Mouse selection Configure)		Settings		
Unique Attribute		14		Network Overview		
Choose an attribute				Average Degree	17.577 Ru	0 10
Choose an attribute	*	10		Avg. Weighted Degree	13.335 Rui	in O
		Modularity settings		Network Diameter	9 Ru	m O
	Modularity			Graph Density	0.046 Ru	m 🕐
	Community	detection algorithm.		Modularity	Rat	0 10
				PageRank	Ru	0 11
	Random	Produce a better decomposition but increases computation tim	2	Connected Components	Ru	0 10
Layout O	🛃 Use weig	ghts Use edge weight		Node Overview		
Choose a layout				Avg. Clustering Coefficient	Rut	0 10
	Resolution:	Lower to get more communities (smaller ones) and righer than 1.0 to get less communities (bigger ones).		Eigenvector Centrality	Rul	n 0
0	1.0			Edge Overview		
				Avg. Path Length	3.091 Rui	in O
		Cancel		🗉 Dynamic		
0.00				# Nodes	Rui	0 11
<no properties<="" td=""><td>».</td><td></td><td></td><td># Edges</td><td>Ru</td><td>n 0</td></no>	».			# Edges	Ru	n 0
				Degree	Ru	n 0
		P		Clustering Coefficient	Rat	n 0
				· · · · · · · · · · · · · · · · · · ·		
P Presett Reset		YET T A TO A Arial-BoldM	T, 32 🕘 🎄			

Now, let us compute modularity. Modularity is the measure for communities in a network which we have already covered in the previous tutorial. For our data, we will

unselect the use weight option, because we do not have a weighted graph and click on ok.

(Refer Slide Time: 09:24)

Workspace 1 ()	0.0.0	HTML Report		4123
Appearance () Nodes Edges	Modularity Report		tings	
Unique AttributeChoose an attribute	Parameters: Randomize: On Use edge weights: Off Resolution: 1.0 Resolution: 0.374 Modularity with resolution: 0.374 Number Of Communities: 1.8		Network Overview erape Degree 12 g. Weighted Degree 12 swork Diameter 12 up Density 02 dularity 02 geRnik nnected Components	1.577 Run 0 3.335 Run 0 9 Run 0 9 Run 0 1.046 Run 0 1.374 Run 0 Run 0 Run 0
Californi ()	Siz	e Distribution	Node Overview	
Choose a layout	55 (*	g. Clustering Coefficient envector Centrality	Run 0
0	50		Edge Overview	
	45		g. Path Length	3.091 Run 🙂
	(s) 40		Dynamic	
	2 35		 łodes	Run 0
<no pro<="" td=""><td>pertiesa 5 30-</td><td></td><td> idges</td><td>Run 0</td></no>	pertiesa 5 30-		 idges	Run 0
	đ 25 ·		gree	Run 0
	20		istering Coefficient	Run 0
	Print Come State		loca	

In our case, we can see that the numbers of communities are 18 and the modularity score is 0.374.

(Refer Slide Time: 09:41)

Workspace 1 Q		
Appearance 🔘	Graph O	Filters Statistics ()
Nodes Edges	D A TT G Mouse selection Configurei	Settings
Unique Attribute		Network Overview
Choose an attribute	0	Average Degree 17.577 Run
Character an account	Page Rank settings	Avg. Weighted Degree 13.335 Run
	PageRank	Network Diameter 9 Run
	Ranks nodes "pages" according to how often a user following links will non	n-randomly Graph Density 0.046 Run
	reach the node "page".	Modularity 0.374 Run
	and the second s	PageRank Run
	ODirected Probability (p): 0.83	Connected Components Bun
ayout O	Used to simulate the user randomly restarting the web-surn	ing.
Choose a layout	Epsilon: 0.001	Avg. Clustering Coefficient Run
	Stopping criterion, the smaller this value, the longer converg	gence will take. Economic Controlity Run
0		e Edge Chamien
	Use edge weight	Avo, Path Lenoth 3.091 Run
		P Dynamic
	Canal	# Nodes Run
«No Properties»	Lexo	# Edges Run
		Dentes
	6	Character Codificient Day
		Clustering Coefficient

Now let us compute the page rank. You can change the p and epsilon values as per the page rank algorithm or keep it as default and click on OK.

Workspace 1 ()	ITML Report			4 1 (27)
Appearance ()	PageRank Report		iters Statistics ()	1
Nodes Edges	rugenank neport		Retword Overslaw	
Unique Attribute			erace Decree 17.	577 Run @
Choose an attribute	Parameters:		n Walahtad Danna 17	225 840 0
	Epsilon = 0.001		g. weighten Degree 15.	0 2-0 0
	Probability = 0.85		cwork plameter	9 (Kan 0
	Results:		iph Density 0.	J46 Run 0
	PageRank Distribution		dularity 0.	374 Run 🙂
			geRank	Run (0
Layout O	25.0		nnected Components	Run 0
Channelman	22.5		Node Overview	Party of
Choose a layout	20.0		g. Clustering Coefficient	Run 0
0	17.5		envector Centrality	Run 0
	15.0		Edge Overview	
	5 12.5		g. Path Length 3.	J91 Run 🔍
	0 10 0		Dynamic	Run (B
chio Pr	contract of the second second		eoues	Nati 0
	75		loges	Kun 0
	5.0		gree	Run 0
	25		astering Coefficient	Run 0
	Print De Conv D Save	Close		

(Refer Slide Time: 09:50)

This will compute the values of the page rank algorithm for your graphs.

(Refer Slide Time: 10:04)

000	Gephi 0.9.1 - Project 1	
😝 Overview 📄 Data Laboratory	Preview	
Workspace 1 ()		
Appearance ()	Graph C	Filters Statistics
Nodes Edges	Φ Δ π K Mouse selection (Configure)	Settings
Unique Attribute	R.	Network Overview
Chesro an attribute	Consultad Community utilizer	Average Degree 17.577 Run ()
Choose an attribute	Connected Components settings	Avg. Weighted Degree 13.335 Run (9
	Connected Components	Network Diameter 9 Run ()
	Determines the number of connected components in the network.	Graph Density 0.046 Run (9
		Modularity 0.374 Run (9
	O Directed Detects strongly & weakly connected components	PageRank Run ()
	Indirected Detects only weakly connected components	Connected Components Run 0
Layout ()		Node Overview
Choose a layout		Avg. Clustering Coefficient Run 0
		Eigenvector Centrality Run 0
•		Edge Overview
		Avg. Path Length 3.091 Run 0
		🖻 Dynamic
		# Nodes Run @
<no properties=""></no>	Cancel	# Edges Run 0
		Degree Run 0
	P	Clustering Coefficient Run 0
Presets Reset	PET T S C A. A. Arial-BoldMT, 32	▲

Let us see how many connected components are there in our data. Click on OK.

(Refer Slide Time: 10:08)

Workspace 1 ()	0.0	HTML Report			10 00	a tác a
Appearance () Nodes Edges	Connected Cor	nponents Report		tings		(8
Unique Attribute				Network Overview		-
Choose an attribute	Parameters:			erage Degree	17.577	Run O
	Natural Interpretation: direct	ted.		g. Weighted Degree	13.335	Run 📀
	interview interpretation, uner			twork Diameter	9	Run 0
	Results:			uph Density	0.046	Run 🕐
	Number of Weakly Connecte	d Components: 14		dularity	0.374	Run 📀
	Number of Stronigy Connect	ed Components: 65		geRank		Run 😃
	_			nnected Components	14	Run (0)
Layout ()		Size Distribution		Node Overview		
Choose a layout	13 +			g. Clustering Coefficient		Run 0
	12			envector Centrality		Run 0
	11			Edge Overview		
	10	•		g, Path Length	3.091	Run 🙂
	9			Dynamic		
	¥ 7			lodes		Run 0
crio Prope	Desa 5 6			idges		Run 0
	5			gree		Run 0
	4			istering Coefficient		Run 0
	Print Copy Save		Close			

We can see that we have 65 strongly connected components; and 14 weakly connected components in a data.

(Refer Slide Time: 10:27)

😝 Overview 📄 Data Laborato	y Preview		Ľ٨
Workspace 1 (3)			
Appearance ()	Graph C	Filters Statistics ()	
Nodes Edges	P D A T G Mouse selection (Configure)	Settings	
Unique Attribute	R	Network Overview	
Channel and the		Average Degree	17.577 Run (
Unoose an attribute	Charterine Confident antilians	Avg. Weighted Degree	13.335 Run
	Clustering Coemcient settings	Network Diameter	9 Run
	Clustering Coefficient The clustering coefficient along with the many chortest outh can indicate a "small-world"	Graph Density	0.046 Run
	effect. It indicates how nodes are embedded in their neighborhood. The average give an	Modularity	0.374 Run
	overall indication of the clustering in the network.	DearDealt	0.374 1141
		Payenank	Rati
ayout ()	O Directed	Connected Components	14 Run
Chanza a lausuit	OUndirected	Node Overview	
Choose a layout		Avg. Clustering Coefficient	Ş n
P		Eigenvector Centrality	Run
		Edge Overview	
		Avg. Path Length	3.091 Run
		Dynamic	
Alla Desperations	Cancel OK	# Nodes	Run
CN0 PT000000		# Edges	Run
		Degree	Run
	P	Clustering Coefficient	Run
	? A		
	<u>a</u>		
Province Provent	Y HI Y T X Y O A. A. Arial-BoldMT, 32	۵.	

Now let us also look at the node specific attributes which is the average clustering coefficient. Let us click on run.

(Refer Slide Time: 10:33)

Overview Data Laboratory	Gephi 0.9.1 - Project 1		¥ ž'n
Workspire 1 O			\mathbf{M}
Appearance ()	Graph O	Fitters Statistics	
Nodes Edges	A T / Mouse selection (Configure)	Settings	
Unince Attribute		Network Overview	
Unique Allinosie		Average Degree	17.577 Run 0
Choose an attribute	Cinemanter Centrality entitiese	Avg. Weighted Degree	13.335 Run (
	Bigenvector Centrality securitys	Network Diameter	9 Run (3
	Eigenvector Centrality	Graph Density	0.046 Run 3
	A measure of node importance in a network based on a node's connections.	Modularity	0 374 Run 0
		ProsPank	Run (
	O Directed	rayenalik	Id Pue 0
Layout 💿	Oliverand N	Connected Components	14 Kun 0
Choose a layout	Undurected	Avn. Clustering Coefficient	0.256 8up 0
	Number of iterations: 100	Figure and a control in the	0.230 844 0
0		Eigenvector Centrality	Kin 0
		Avg. Path Length	3 091 Run @
		R Denamic	3.031
	Carrel	# Nodes	Run 0
<no properties=""></no>	Canto On	# Edges	Run 0
		Degree	Run (I
	8	Chatterine Conflictent	Aug d
	P	Clustering Coefficient	KUN U
	2 4		
	A		
Presets Reset	PET X 2 C A. A. Arial-BoldWT, 32] ● ▲	

And let us also compute the eigen vector centrality of the nodes.

(Refer Slide Time: 10:38)

Workspace 1 💿	HTML Report		1.0
ppearance ()	Parameters	iters Statistics 🛇	
iodes Edges	raidineters.	tings	
inique Attribute	Network Interpretation: directed Number of iterations: 100	Network Overview	
Change on ottaliute	Sum change: 0.010901034387345644	erage Degree	17.577 Run
Choose an attribute	Results:	g. Weighted Degree	13.335 Run
	Figure star Cantrality Distribution	twork Diameter	9 Run
	Eigenvector Centrality Distribution	aph Density	0.046 Run
	25.0	dularity	0.374 Run
	22.5	neRank	Run
	20.0	annuted Commenter	14 10
O tuo	17.	Nede Oversiens	14
Choose a layout	1/3	n Clustering Coefficient	0.256 Run
	E ^{15.0}	gr entreling exercisity	0.230
	8 12.5	envector Centrality	Nun
	10.0	a Path Leasth	2.001 8.0
	7.5	Braunic Congen	3.031
	50	lodes	Run
<no propertie:<="" td=""><td></td><td>Here</td><td>Pue</td></no>		Here	Pue
		, uges	Num
	0.0	gree	Run
	0 I	astering Coefficient	Run

Once these measures are calculated, many of these will be available in the data laboratory and can also be used for the visualization. For example, after computing the eigen vector centrality, it will allow you to resize nodes based on that particular attribute.

(Refer Slide Time: 11:14)

		Gephi 0.9.1 - Project 1	
😝 Overview 📋 Data Labo	ratory Preview		
Workspace 1 0			
Appearance 🔘	9	Graph ()	A A Filters Statistics ()
Nodes Edges		Gp Mouse selection Configure	Settings
Unique Attribute		24	Network Overview
Choose an attribute	A		Average Degree 17.577 Run
Strongly-Connected ID	-	0	Avg. Weighted Degree 13.335 Run
type			Network Diameter 9 Run
Closeness Centrality friends		1	Graph Density 0.046 Run
Eigenvector Centrality	1	*	Modularity 0.374 Run
shape Clustering Coefficient		D Martine 1	PageRank Run
Degree	-	/ North	Connected Components 14 Run
Layout or		1 2 2 2 4 1	E Node Overview
Choose a layout	2		Avg. Clustering Coefficient 0.256 Run
		1 Sales	Eigenvector Centrality Run
	P for	•	Edge Overview
		2	Avg, Path Length 3.091 Run
		A Party and	B Dynamic
No Proveding			# Nodes Run
<no properties<="" td=""><td>8.</td><td></td><td># Edges Run</td></no>	8.		# Edges Run
			Degree Run
		P	Clustering Coefficient Run
		A A	
		Internet and the local program	
P. Presett Reset		T A C A A	Arial-BoldMT, 32 a

We can see **it** there in the left panel. Let us go to nodes, attributes, and you will be able to see all the measures which we have computed just now, like the strongly connected components, eigenvector centrality, **clustering co-efficient** etcetera. All of these options were not present earlier.

(Refer Slide Time: 11:50)



Now, let us collapse the filters and statistic tab and play around with the visualization of the data a bit. Currently, because the data is very dense, the nodes and edges are all mixed up. Therefore, let us choose a layout so that we can see the data more efficiently, click on choose a layout drop down menu and select Fruchterman Reingold algorithm. And then click on run.

(Refer Slide Time: 12:02)

	Gephi 0.9.1 - Project 1	
😧 Overview 🚺 Data Labo	atory Preview	
Workspace 1 ()		
Appearance ()	Graph Q	101101 1016
Nodes Edges	D A T Mouse selection (Configure)	Ĩ
Unique Attribute	2	
Choose an attribute		·
	· · · · ·	
	1	
	· · · · · · · · · · · · · · · ·	
	= D Apply	
Layout O		
Fruchterman Reingold		A Carlos And
0		
•		
v Fruchterman Reingold Area 10000		
Gravity 10.0		
Speed 1.0		
	e Alto	
		V VVA-
Fruchterman Reingold		
Presets Reset	TH. T X 2 A	▲ Arial-BoldMT, 32

Let the algorithm run for a while. And after it starts to look stabilized, click on stop.

(Refer Slide Time: 12:31)



Now let us adjust the size of the nodes and edges based on the various attributes, which we just calculated. Let us first change the size of the nodes. In the appearance tab, on the top left, click on nodes, then select the size from the right side options, click on attribute and let us choose an attribute. Let us say we want to size the nodes according to how many followers that specific node has. Click on followers and then apply.

(Refer Slide Time: 12:50)



You will notice that some nodes now are bigger. These are probably the nodes which have large number of followers and which has celebrity accounts in the twitter network.

(Refer Slide Time: 13:19)

	Gephi 0.9.1 - Project 1	
Overview Data Laboratory	Preview	YIN
Workspace 1 0		
Appearance ()	Craph C	ECHECH CO
Nodes Edges 🗣 句	A T Go Mouse selection Configurei	
Unique Attribute	Q.	
followers	E the advant	
Min size: 5 C Max size: 50 C	9	-
	1 townstation and and a subsections	•
🗐 Spline	Q	and the second sec
	Apply 10	Company and the second
Layout O		
Fruchterman Reingold	· · · · ·	
	X	and the second s
0	Run 0 seet Supers"	The second se
v Fruchterman Reingold	2	VARCE
Area 10000.0		
Speed 10		and a second
		a second s
	()	
Fruchterman Reingold		
V hours - Dece	A A Aria-BoldMT, 32	-0- 12 4
C Preses Reset		

We can see which users these are by enabling the node labels. To enable the node labels click on T in the bottom panel which is show node labels. Now, you will be able to see

which are these users which have a large number of followers. This particular node is the official account of twitter; and this one is YouTube.

000		Gephi 0.9.1 - Project 1
Overview Data I	Laboratory Preview	
Workspace 1 ()		
Appearance ()	•	Graph ()
Nodes Edges	€ <u>6</u> A T	Gp Mouse selection (Configure)
Unique Attribute		
In-Degree	0	Land the state of
Min size: 5 1 Max size: 50 1		
	= 🚺 Apply	D REFERENCE
Fruchterman Reingold	D Run	
v Enuchterman Reinspold		and a second sec
Area 10	000.0	hand the second se
Gravity 10	1.0	and the second
Speed 1.0	0	
Fruchterman Reingold	9	
Presets Reset		▼ T S -0 [A ⁺] A ⁺ [

(Refer Slide Time: 13:56)

Let us try **changing** the parameter based on which we want to size the nodes. Click on the drop down menu again, and let us select in-degree this time, click on apply. Since, it is a two **hop** network of my own twitter data, therefore the in-degree **would** signify that these are the accounts which are most followed in my two **hop** twitter network. We can adjust the minimum and the maximum size of the nodes by changing these values.

(Refer Slide Time: 14:28)



For instance, let us change the minimum size from 5 to 10 to make the even smallest node bigger and change the maximum size to let us say 70. Click on apply. And you will notice that the size of all the nodes is now bigger.

(Refer Slide Time: 15:06)

			Gephi 0.9.1 - Project	1	
0verview	Data Laboratory	Preview			Yin.
Workspace 1 ()	_				
Appearance 🔘		0	Graph ()		ICHICI IN C
Nodes Edges		TT A G •	G Mouse selection Configurei		ī
Unique Attribute			4		
Choose an attribute		0		Recolling adapted	
ffr Weighted In-Degree			0		num and a second
user_id Weighted Out-Degree Ifr			2		
listed Betweenness Centrality			D		and
Leyou to T			1		Section *
Fruchterman Reingold		8	/		
0		D Run	0		
v Fruchterman Reingold			1		and services and
Area	10000.0		6.9841 m		and the second second second
Gravity	10.0				And an and a second
Speed	1.0				
			P		and a second sec
Fruchterman Reingold		0	<u>A</u> A	A CONTRACT OF A	(Regel)
P Presets Reset			VE- T XX 2	- A· A· Arial-BoldMT, 32	

Now to make the graph more visible, let us try changing the color of the nodes. Now again, select nodes, and after that from the right, instead of size, select the color option, then go to attributes., and click on the drop down menu. Let us color the nodes based on the community, which they belong to. To do that, we will need to select the option modularity class.

(Refer Slide Time: 15:12)



As soon as you click the modularity class option, it will tell you how many nodes exist in each community, click on apply.

(Refer Slide Time: 15:26)



Now, the nodes are better visible you can zoom in by centering the graph and then scrolling down. Let us look at few more options. In the nodes, we can also change the label color by click on the A option. Let us color the labels based on the attributes available in the drop down menu. Let us choose out-degree and click on apply.

(Refer Slide Time: 15:52)



If you notice that the labels are not visible, you can change the color pallet. We can also change the size of the labels by clicking on the T option. Click on T, and either we can change the size of all the node labels.



(Refer Slide Time: 16:16)

Let us change from 1 to 5 and see what happens, the text becomes too big.

(Refer Slide Time: 16:40)



Let us change it back to 1, we can also resize the node labels by clicking on the attribute tab, and have a continuous function. Let us choose out-degree and click - apply. Now the nodes which have higher out-degree have a larger label which means that users who are following other users **a** lot will have a larger text. In this case, this particular account srishti underscore gupta 14 **is** the account which follows maximum number of users in **my** twitter **data**.

(Refer Slide Time: 17:08)



Let us see what happens if we choose in-degree. Select in-degree from the drop down menu and click on apply. Now, the size of the labels will be in line with the size of the nodes, because that is how we have sized our nodes.

(Refer Slide Time: 17:30)



Let us change the node labels back to the unique attribute by selecting one so that all the nodes are of the same size. Now we have all the node labels back to the same size which is equal to 1.

(Refer Slide Time: 18:15)



Let us explore more options to adjust the visualization of these graphs. We can also change the appearance of the edges; we can change the color, the label color and the label size of the edges. In our case, we do not have any labels for the edges. Therefore, we can play around with the color of the edges. Click on color, then attributes and then choose the available attributes. In our case, it is weight, click on apply.

(Refer Slide Time: 18:27)



If we want to give the same color to all the edges, click back on unique and select the edge color and then click on apply. This will make all the edges appear grey.

(Refer Slide Time: 18:48)



You can also try running the other available layouts in the layout tab. For instance, let us try running contraction; contraction will just zoom out the graphs in the current layout.

(Refer Slide Time: 19:24)



Let us try ForceAtlas. After the graph stabilizes, click on stop to analyze the graph further. Remember that if your graph goes off the screen, then find out this small

magnifying glass, which is used to center the graph in your blank space. And then you scroll down to zoom in the graph in the particular area which you want to see.



(Refer Slide Time: 19:48)

Let us again select the Fruchterman Reingold algorithm and click on run. When the graph seems stabilized, click on stop. Scroll up to zoom out of the magnified graph.

(Refer Slide Time: 20:22)



Now, let us look at more visualization options provided by Gephi. Let us look at the options on the left of the layout. The first option is the direct selection tool. It can help you select specific nodes and their neighbors. Click on the direct selection nodes and hover over a particular node. For instance, when we hover over this node it selects the particular node and it is neighbors. Now **it's** neighbors are automatically getting selected because of an option which has been opted n.

(Refer Slide Time: 20:53)



Let us look at that option. Notice the up arrow on the bottom panel, expand it, go to global, and you will see that the auto select neighbor option is ticked. Uncheck that and press the down arrow again. Now again hover over the previous node, this time the neighbors of this particular node will not be selected. If we want to turn back that option then there comes the up arrow again, and again select the auto select neighbor option. Now you will see the same happening.

(Refer Slide Time: 21:38)



Let us look at the next option. The next option is the rectangular selection. Click on the rectangular selection. This will enable you to select a part of the node. Drag and select **a** rectangular area to select all the nodes and their neighbors in that particular area.

(Refer Slide Time: 21:55)

		Gephi 0.9.1 - Project 1
Overview Data	Laboratory Preview	
Workspace 1 0		
Appearance ()	9	Craph (C
Nodes Edges	🕈 🕤 🛓 т	r 🕼 Rectangle selection
Unique Attribute		No. of the second secon
Size: 1 🗘		Non-Sector P Residence P Providence P Providence P Residence P Residence P Residence P Residence P Residence
Layout () Fruchterman Reingold	= Apply	
0	D Run	0
v Fruchterman Reingold		2
Area 1	0000.0	
Gravity 1	0.0	and the second se
Fruchterman Reingold	3	
P Present Record		T N = T N = - → A· A· A· A· a-ad-bolow.32 → → ■3 ▲

If you want to select only the specific nodes in the selective rectangular area, then we can again click on the arrow on the bottom panel and uncheck the auto select neighbor. Click the down arrow. And this time, when you select a particular area, only the nodes within that area will be selected.



(Refer Slide Time: 22:06)

In this graph, we can see that two of these particular nodes are very close to each other.

(Refer Slide Time: 22:34)

			Gephi 0.9.1 - Project 1		
Overview	Data Laboratory	Preview	1		Y In
Workspace 1 0					110
Appearance ()		Graph	0		
Nodes Edges		● D A T G	Dragging (Configure)		Ĩ
Unique Attribute		4		shundan fördeforte Garde 211 instante sytud	bitry Inter
Size: 1 0		2 2 2 2 2 2	Annahring Backborg Jar Nawas Backborg B	Anumbers of the Gray as the Cardeo and the second and the second and the second and the second and the second and the second and the second and the second a	Bandfindar Killer Gener ungest Veller Triglier ensahlget Mit Gener ensahlget Mit Gener ensahlget ensa
Layout O Fruchterman Reingold	3	P Apply /	antisatives antipos Restards antipos Constitutions and antipos attes vanishing an antipos		And an and and an
0		Run 0		and an and an and an and	And Ballins
v Fruchterman Reingold		1	CHILDRELING HUNDRELING		Dank Joseph Achter
Area	10000.0				Program Children Provinsi
Gravity	10.0		andaya unday	The Andrew Contraction of the Andrew Contrac	After Partial and death
Speed Fruchterman Reingold	1.0	() () () () () () () () () ()	tadara acadacias acada acadacias acadacias acadacias acadacias acadacias acadacias acadacias acada acadacias acadac		Landrey Branner Den Den Landrey References Landrey References
Presets Reset			B T S 2 2 -0-	A· A· Arial-BoldMT, 32	- -

Let us say we want to make a small adjustment and make these nodes **a** little far away. To do that, we can select the drag option **by** clicking on it. Then, going to the node which is **being overlapped** with the other and dragging **it** just a little further **off.** The drag **tool** is used for manual adjustment of the nodes.

(Refer Slide Time: 23:07)



Let us look at the next option. The next option is the painter tool, which is used to color nodes by selecting the specific nodes. Select painter nodes and select the color, which you want to color your new node. Then go to your particular node and click on it. This will give it the color, which we have selected.



(Refer Slide Time: 24:25)

Now let us look at the option to add additional nodes using the node pencil. Select the node pencil and click somewhere in the empty space. This will create a node. You can add as many nodes as the number of new clicks, which you make. You can also add edges between the newly created nodes by clicking on the edge pencil and creating specific edges. You need to first select the source node and then the target node. Let us create some more edges.

(Refer Slide Time: 24:56)



We can also **edit** the node attributes by selecting the edit tool. Click on the edit tool and then click on a specific node. This will elaborate the properties of the nodes in the left panel. We can change the node properties, as you like. Let us change the size of the node from 10 to 25; you can see that the size of the node has increased. Similarly, you can **change** the other parameters.

(Refer Slide Time: 25:15)



You can also change the background color of the Gephi visualization by clicking on the bulb icon on the bottom. Clicking on the bulb icon will toggle between black and white backgrounds.

0.0.0			Gephi 0.	9.1 - Project 1		
Overview	Data Laboratory	Preview				YAA
Workspace 1 O						121
Appearance ()		00	raph O			ICHICI CO C
Nodes Edges		· OAT	Rectangle selection			1
Unique Attribute		4			1000	
Size: 1 C			Sa	we As	- Bogenett - marginett	
		Sa	ve As: screenshot,	014148.png	A	tatar Bina angkat Yadan
		Name aopfgifeimeiolaj Applications	eknfidlijpoebgc	 Date Modified Tuesday, 8 March, 2016 1:24 PM Friday, 22 July, 2016 1:35 PM 		alandager Manganadager
Layout ()	_	APWGphishnew.t	it.	Tuesday, 24 September, 2013 4:50	PM PM	Perf
Fruchterman Reingold		benhana.mov		Wednesday, 6 April, 2016 2 27 PM Thursday, 16 June, 2016 10:05 PM	Realized and	
0		comproacc		Thursday, 16 June, 2016 10:05 PM		
0		Compromised-ad	counts	Thursday, 28 July, 2016 1:57 PM Tuesday, 5 February 2013 11:43 Pf		Amalian
* Fruchterman Reingold		Despendra Rag	wamh(MTech, MT.	. Sunday, 6 March, 2016 12:39 AM		Ball Queen
Area	10000.0	Desktop	•	Monday, 12 September, 2016 12:23	PM	Martine State State State
Speed	1.0	Fi	le Format: PNG ('png)	Baller Branger	Public producest
		New Folder		Cancel		Sectores Generality 1
Fruchterman Reincold		0	I			utataya Ayutt
		A	8	Section 21	and the state of t	
Presets Reset			E T	X = A A	Arial-BoldMT, 32	- ∎3 ≜

(Refer Slide Time: 25:36)

Gephi also lets you take a screen shot of the generated visualization by clicking on the camera button. This will generate a screen shot in png format, and you can save it anywhere you want.

(Refer Slide Time: 26:14)



Now, let us look at the options in the bottom panel in more detail. Expand the bottom panel using the up arrow at the bottom, and click on global. From where you can customize the background color not just black and white, but also a customized color. Click on the background color and you can choose any color, which you want. You can change it back to the default color by again clicking on the background color option.

(Refer Slide Time: 26:38)



Let us click at the edges tab. The show option is clicked. If you uncheck the option, all the edges will be invisible.

(Refer Slide Time: 26:44)

		Gephi 0.9.1 - Project 1	
Overview	Data Laboratory Preview		Y in
Workspace 1 0			\sim
Appearance ()	÷	Graph O	ICHER OF
Nodes Edges		G Rectangle selection	1
Unique Attribute		1	
Size 1 0	= 🔰 Apply		
Layout () Fruchterman Reingold	•	/ / x	0
0	👂 Run		a man a man
v Fruchterman Reingold		A	and the second se
Area	10000.0	A	
Gravity	10.0		La particular conten
Speed	1.0		x Ariul-BoldMT, 32 ── ─ ■ (3) ♥
			Clobal Edges Labels
Fruchterman Reingold	0	Edge default color: Scale Se Use node color: Scale Se	election color tr Brac Az

Click it back to make the edges visible. From this panel, you can also decide whether you want to keep the color of the edges same as the nodes or not. If you choose the use node color option then all the edges will become of the color of these specific nodes, which they are attached to.

(Refer Slide Time: 27:13)



You can also change the font of the node labels by clicking on the font value at the bottom. Select the new font, which you want, and click on ok, this will change the font for all the node labels.

(Refer Slide Time: 28:08)



Now, let us look at the filters tab in Gephi on the right. Click on the filters tab. And let us select a filter to choose specific nodes and edges. Let us click on the attributes filters and click on range. Click on followers and define the range.



(Refer Slide Time: 28:35)

Let us say we want to keep the number of nodes, let us say we want to keep the nodes which have the followers between 8 to 4 million. Use this sliders accordingly and click on filter. This operation will remove the nodes which do not fall in to the particular range which we have defined.

(Refer Slide Time: 29:16)



We can also have additional filters. For instance, let us choose a filter based on the outdegree. Select it in the queries and drag the slider to change the values according to what you want. We can see that there are only few nodes, which have out-degree between 24 to 116. If click on stop then we will be able to see the entire network. Let us change the value of the range of the out-degree settings.

(Refer Slide Time: 30:45)



Remember that we also learned about the operator filter. We can also use multiple operations to combine the filters in the queries tab. Let us try using the intersection. Select the already existing filters, which is range based on out-degree and range based on followers as the sub filters in the intersection filter. This will give an overlap of the two filters which is the nodes which have an out-degree between 4 to 116, and also the nodes which have the number of followers between 8 to 4.1 millions. If you click on stop, you will be able to see the entire network.

(Refer Slide Time: 31:16)



You can use various other filters in similar fashion. You can also save the filtered graph by first applying the filter, and then clicking on the screenshot tool to save this specific graph.

(Refer Slide Time: 32:36)



Now, let us collapse the **filters tab** and look at the preview tab of Gephi. The preview tab will be on the top right. The preview tab provides advanced option to adjust the visualization of the generated graph. Once you click on refresh, you will be able to see the graph visualization on your right.

(Refer Slide Time: 32:43)



In this graph visualization, the edges are curved, because that is what is present in the default options; scroll down in the presets default in edges and unselect the curved option again click on refresh. Now the edges will be straight and not curved.

Gephi 0.9.1 - Project 1 ... XI .OH Data La ory Preview Presets Default age 0 Box color Box opacity T Edges Show Edges 1.0 hickness Rescale wei mixed 100.0 sacity inved 0.0 * Edge Arn 3.0 v Edge Labels Arel 10 Pe rten labe 30 0.0 Aax characte . Dutline size Dutline color view ratio: 100% 🖗 Refresh xt: SVC/PDF/PNC und Reset zoom

(Refer Slide Time: 33:10)

If you want all the edges to be of the same thickness, then click on rescale weight and click on refresh. This will cause all the edges to be of the same thickness.

(Refer Slide Time: 33:31)



Right now by default, we are not even able to see the node labels. To be able to see the node labels, click on the show labels option under the node labels, and again press on refresh, now you will be able to see the node labels and you can adjust your font size.

(Refer Slide Time: 33:50)



Let us change the font size to 9, and press refresh.

(Refer Slide Time: 34:08)



The preview tab has **it's** own export command at the bottom left. If you click on the export command, it will give you several options to store the generated file in either pdf format, png or svg.

(Refer Slide Time: 34:55)

			Gep	hi 0.9.1 - Project 1				
Overview	Data Laboratory	Preview					Y	ĬΛ
Workspace 1 O							~~~	4 1 20 1 4
Appearance ()		e)	Graph 🛈					1 1 1 6
Nodes Edges			G Mouse selectio	n Configure)				ī
Unique Attribute			4					Filter
Size: 1		Apply	□					s 44 Statistics
Layout () Fruchterman Reingold		Run	/ / x					
* Fruchterman Reingold	10000		2		X Delete	^D		
Area	10000.0				Move to			
Speed	1.0				Copy to			
					Settle	^8 ^F		
Fruchterman Reingold		0			Selection data la	boratory ^L		
Presets Reset			₹ H· T	1820	A· A· BaskOld	Face, 32]	12	4

Now let us go back to the overview tab to view the visualization. In case, you want to know in detail about a particular node in the data laboratory, then you can first go to the select tool, **hover** over particular node, right click, and then click on the select in data laboratory option.

(Refer Slide Time: 35:25)

			Gep	hi 0.9.1 - Proje	ct 1							
Overview 🚺 Data L	aboratory	Preview									Y	ĭΛ
Workspace 1 ()											\sim	1.0
Twitter Streaming Importer ()	🚍 Data Table 🔘											
1 - Set Credentials :	Nodes Edges 0 C	Configuration	🚺 Add node 📀	Add edge	iearch/Replace	🖄 Import Spread	isheet 📃 Exp	ort table	🛔 More a	ictions .	Filter	Id
	ld Label Interva	friends user_i	d Ifr follow_	ffr listed	shape statu.	type file	image In-De	Out	Degree V	leighted	Weighted	Weighte
2 - Define the query :	47.0 googl	369.0 2270		178 1351.0	trian 3955.	0 friends anup	/Use 2	3	5 3	.94317	14.838535	18.78
Load Query File Save Query I	44.0 sema	903.0 2298	. 1.2579 946.0	1.0476 119.0	trian 3725.	0 friends anup	/Use 5	13	18 5	.533274	19.9282	25.46
	19.0 twitte	47.0 6253	0.022 5973	1270 1311	trian 3575.	0 friends anup	/Use 18	2	20 1	9.452	2.59179	22.04
Words to follow Us	58.0 neko	165:0 1474	0.7436 780.0	4.7273 58.0	trian 464.0	friends anup	/Use 17	17	34 2	.671182	23.381355	26.05
and the second se	40.0 Provide States	inter liter	In en12[1202.0]	10.1 064.6	tran 5328.	Offriends anup.	/Use 1	0	μμ	.933938	0.0	1.933
Add	13.0 kg < cult 1	luue			rian 524.0	friends anup	/Use 0	0	0 0	.0	0.0	0.0
Delete	Solec	t on Overview	÷		rian 1827.	o menos anup	/USE 9	0	4 3	.3855555	0.0	3.383
Delete	Selec	t neighbour r	nodes on table		Han. 9338.	friends anup	/USE 22	0	10 0	3,301	16 226014	15 60
	10 Selec	t related edg	es		rian 491.0	friends anup.	/Use 3 /Ilica 14	35	49 2	894316	21 371964	24.26
	175.0 M X Delet	e			rian. 60.0	friends anun	/lise_1	0	1 0	032068	0.0	0.032
	127.0 ra Move to				rian 9575.	0 friends anup	/Use 15	9	24 4	7.715	13.19404	60.90
	118.0 Pi Copyto.				rian 1236.	0 friends anup	/Use 4	5	9 3	.181467	4,480447	7.661
	60.0 h				rian 1561.	friends anup	/Use_ 44	13	57 2	01.47	23.560127	225.0
	9.0 di Clear				rian 35.0	friends anup	/Use 0	0	0 0	.0	0.0	0.0
	14.0 p	write data to	the other select	sd nodes	rian 36.0	friends anup	/Use 16	3	19 0	.297521	2.077792	2.375
	57.0 D a Settle				rian 2432.	friends anup	/Use 6	0	6 5	5.100	0.0	55.10
	69.0 Ju Erree				rian 748.0	friends anup	/Use 45	12	57 1	0.430	19.265581	29.69
2. Chance Network Logic to sender	62.0 A Set n	ode size			rian 5205.	0 friends anup	/Use 10	1	11 2	0.737	0.592308	21.32
3 - Choose Network Logic to apply :	22.0 vi				rian 319.0	friends anup	/Use 13	19	32 1	.589498	21.685135	23.27
	191.0 In Merg	e nodes			rian 9136.	_ friends anup	/Use 2	0	2 1	64.61	0.0	164.6
	172.0 m Link t	o nodes			rian 2497.	0 friends anup	/Use 12	0	12 1	2.407	0.0	12.40
	153.0 m Dupli	cate			nan 5290.	o menos anup	/USE Z	9	11 3	.371574	19.646572	25.01
	SZ.0 G Cell				nan 2026.	u menos anup	/USE 8	3	11 0	.804566	5.865047	12.66
	120.0 Auguant	04.0 0314		VIJ 110	-fill 94.0	triends anup	/USE					
		1	۲	1	11	LL.		11	ii			
Connect		Add column	Merge columns	Delete column	Clear	Copy data other colum	to Fill nn ~ with	column a value ~	Dupli	cate		

If you go to the data laboratory now, that particular row will be highlighted. And you can look at the particular node in more detail and the nodes tab. Let say, you want to select a particular node in the visualization from the data laboratory, then go to the particular row, right click and then select select on overview.

(Refer Slide Time: 35:31)



Now when you go to the overview, that particular node will be highlighted. This is an easy way to refer to nodes back and forth from the data laboratory to the visualization.

(Refer Slide Time: 36:14)

		Gephi	0.9.1 - Project 1
Overview	Duta Laboratory 📃 Preview		410
Workspace 1 O			\sim
Preview Settings ()		9	60 Provine ()
 Presets 		1	inde part inde part
Tag Cloud		6	ewy DataPortalindia
	Securitys Manage renderers		Prayteary Philippy anton Section Reamers Machines North/Relines CalvandFibble Main
v Nodes		1	internal k satyath **** shreyaghoshal war
Border Width	1.0		Fa Asset proders former progenitar geogenitar geogenitar
Border Color	custom [0,0,0]		Jackey By hada maskad maskad
opacity	0		ACC Phanteepects Discounting Discounting
v Node Labels			the second and achich sureka
Show Labels	8		Dropboxshithandgoyal
Font	Arial 12 Plain	13	wante vanie protect with twitter vani sivasankamayak chora
Proportional size	8		flowingdata categorian The Pring and a poor varang
Color	custom [0,0,0]	-	infochimps vinagaksysamanin_peini-
Shorten label			simpleved karpany gtc "Twittered bara jkroks ankar-
Max characters	30		nytgraphics mrogaticomparation
Outline size	0.0		man average IIIIaSOM walker non soom UneeraiSanghi
Outline color	custom [255,255,255]	100	DatahBeaufu niponabarita Warehenzokiaci Der WsiatD2015-
Outline opacity	80.0		bigdata ² 2000018 researcepterredova
Box	8		zelandya Kristinal Barnan Römert etweet and lanchi ya
Box color	parent	123	betetasibe
Box opacity	80.0		han habawayayamyardi
* Edges			waterman mstringing micer with antingo Bill
Show Edges			
Thickness	1.0		humanatia 60urtera anima
Pecrale weight			truthyatiosanakenutmekozzanga uring
Preview ratio: 100%			promyoliave/Ven/Ven/Ven/En/20-EDQuity www20/Nei[barted Preventance
Export SVG/PDF/PNG		👳 Rofresh	Background Reset zoom - +

Let us go back to the preview tab. Let us look at more options in the drop down menu; click on the tag cloud and press refresh. This will generate a tag cloud of the node labels.

(Refer Slide Time: 36:44)



We can also select a subset of the graph in a new work space while maintaining the current work space. To do that, let us click on the filters tab again, and apply a filter. Let us choose the previously existing filter which we just tried and click on filter.

(Refer Slide Time: 37:00)



Look at the second option at the top of the filters tab which says export filter graph to a new work space. Click on the option and it will generate a new work space which is work space two. Go to the work space two. And you will be able to see only the subset of the graph based on the filter which we choose earlier. Now, this does not have any existing filters applied to it, you can consider it as a fresh graph to work over and apply further filters, change in visualization etcetera over it.

(Refer Slide Time: 37:44)



As an example let us try changing the layout of this sub graph. Click on forceatlas 2 and apply the algorithm. Click on stop after the graph stabilizes. You can zoom in to see the resulting graph.

(Refer Slide Time: 37:53)



You can also click on **no** overlap, so that the nodes no longer **overlap** each other.



(Refer Slide Time: 38:37)

You can also recolor the nodes based on any new attributes, which you want. For instance, now for this sub graph, let us choose an attribute which is in-degree and apply.

This way you can create new work spaces by filtering portions of the original graph and have more detailed analysis on sub section of a graph.



(Refer Slide Time: 38:47)

To go back to the original graph, you can just click on work space one from the top left corner.

This completes our tutorial on Gephi. And you can generate more graphs based on any data, which has nodes and edges.