CIFF
Centro Internacior
de Formación Fina

CIFF Trustees:

Santander

Universidad de Alcalá

Social Network Analysis

Lecture 1: Basic Concepts

David Martín-Corral Calvo
dmartincc@smartvel.com

Enero / Marzo

MASTER EN BA & BD

Lecture 1: Basic concepts
- Activity I (Gephi) + Assignment I

Lecture 2: Centrality in networks
- Activity II (Gephi) + 3*Presentations

Lecture 3: Communities in networks
- Activity III (R) + Assignment II + 3*Presentations

Lecture 4: Advanced concepts
- Activity IV (python) + 3*Presentations

Lecture 5: Neo4j
- Activity V (py2neo + Gephi) + 3*Presentations

Lecture 6: Neo4j Applications
- Work in Assignment II

Lecture 7: Social network applications
- Presentations Assignment II

Outline:

- Why do we study systems as networks?
- What are networks?
- Concepts: Nodes, Edges, Degree, Adjacency matrix, Edgelist, Adjacency list, Connected Components, Directed and undirected graphs.

Why do we study systems as networks?

*Over the past decade there has been a growing public fascination with the **complex "connectedness" of modern society**. At the heart of this fascination is the idea of a network. **"a pattern of interconnections among a set of things"***

Networks, Crowds, and Markets By David Easley

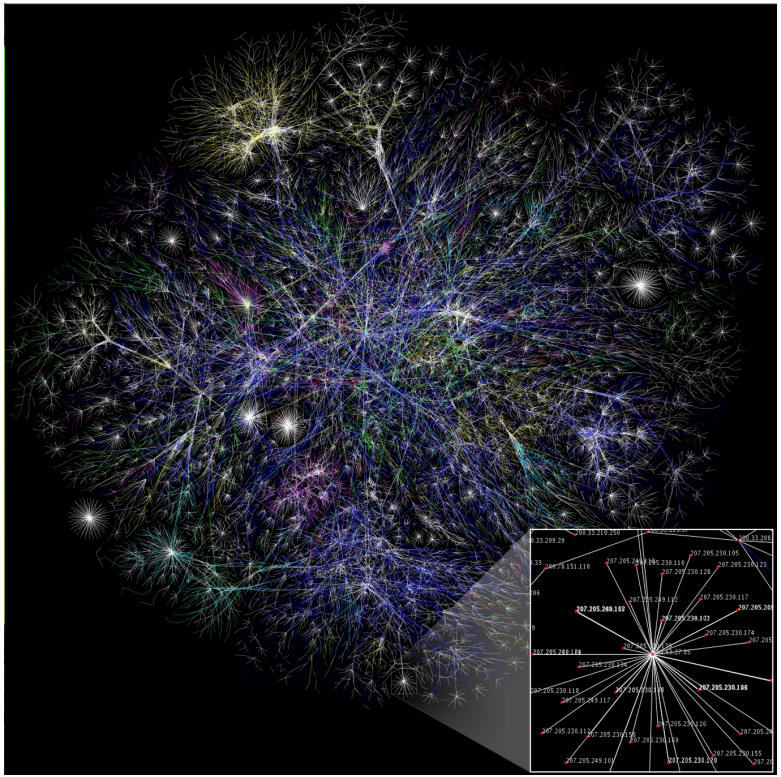
Why do we study systems as networks?

Connectedness

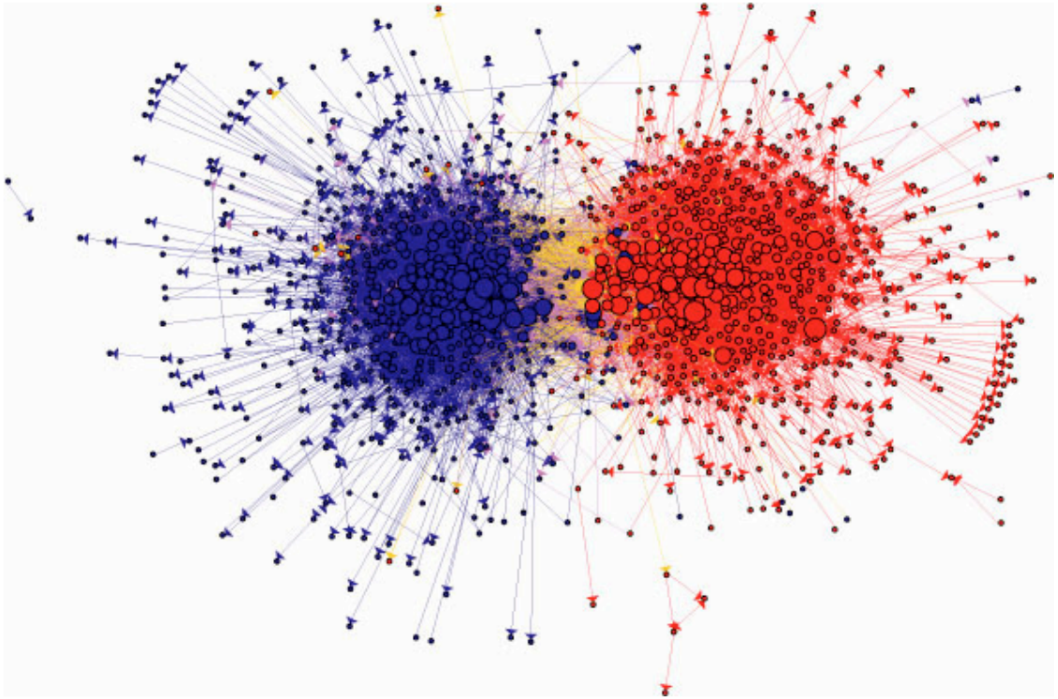
\+ Behaviour and Dynamics

Cause-Effect

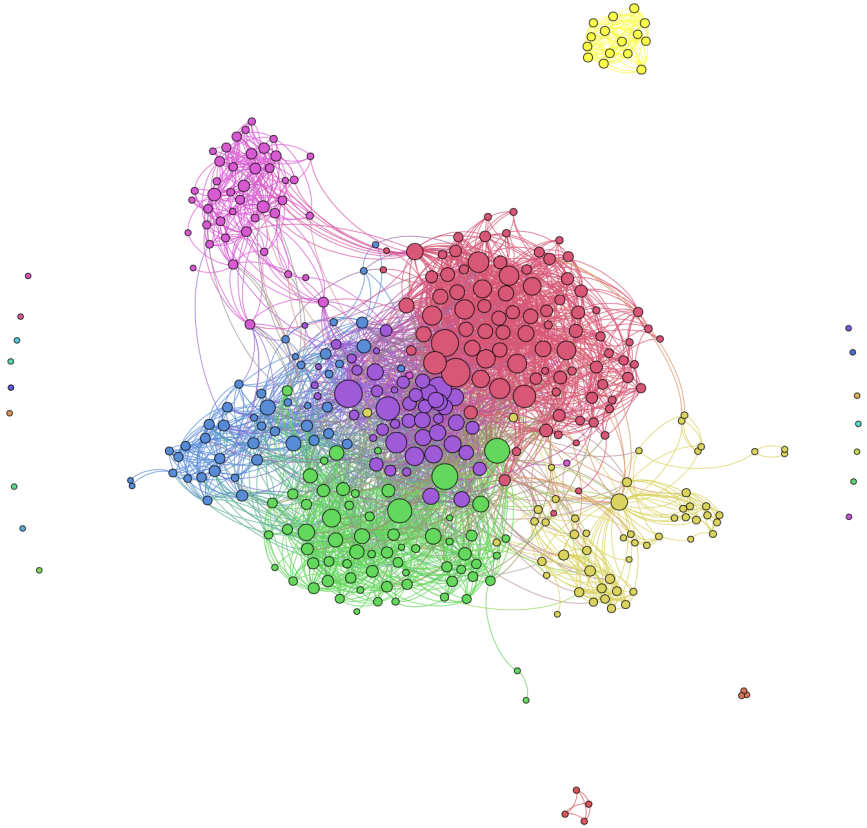
Internet



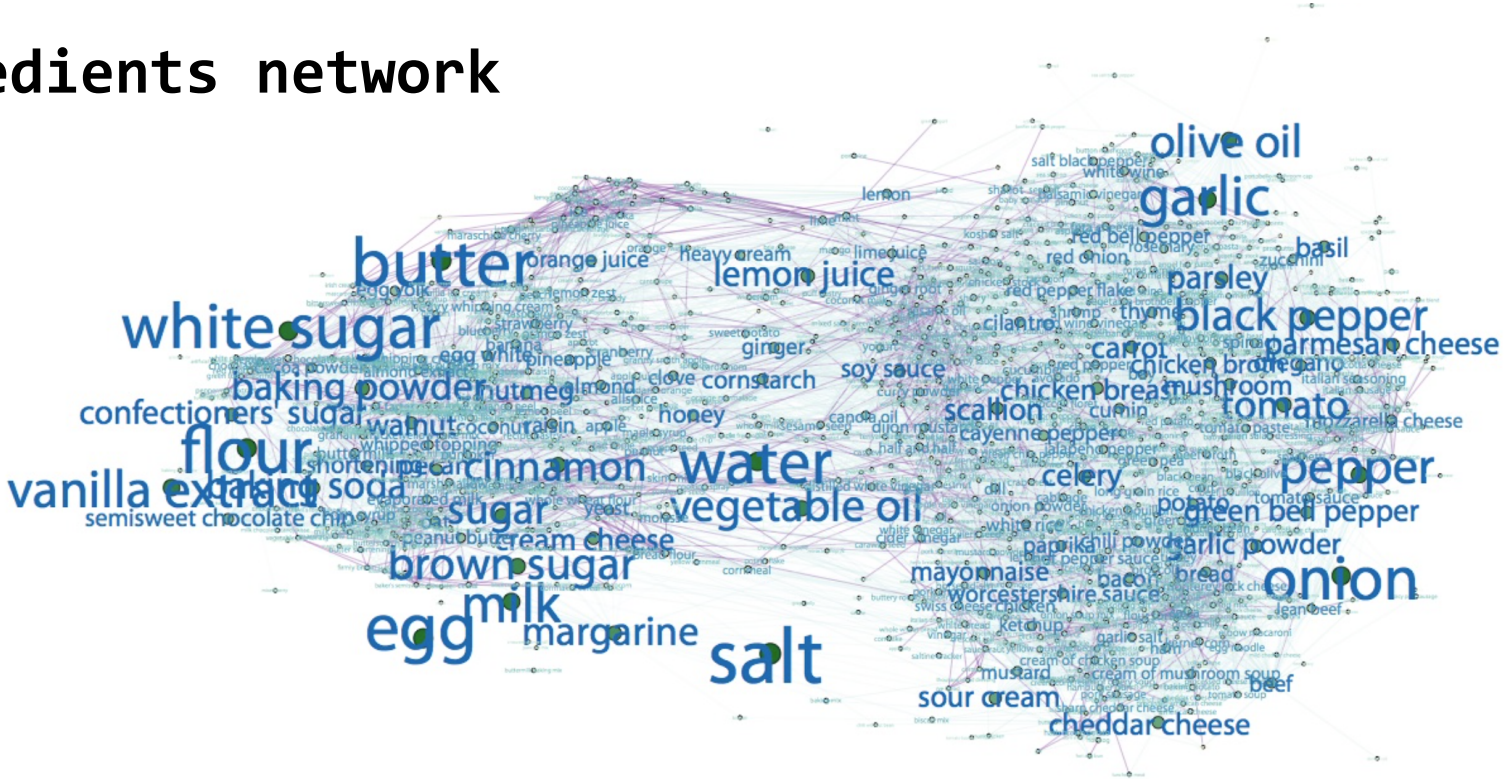
Political blog



Facebook

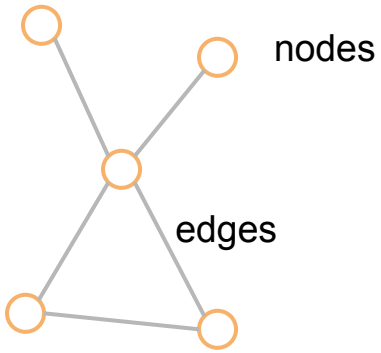


Ingredients network



What are networks?

Networks are sets of nodes connected by edges.



Networks = Graphs

points	lines	
vertices	edges,arcs	Math
nodes	links	Computer science
sites	bonds	Physics
actors	ties, relations	Sociology

Network components: Edges

Directed

A->B

A likes B, A is following B, A is B's father

Undirected

A <-> B or A-B

A and B like each other

A and B are couple

A and B work together

Network components: Edges attributes

Weight (e.g. number of communication)

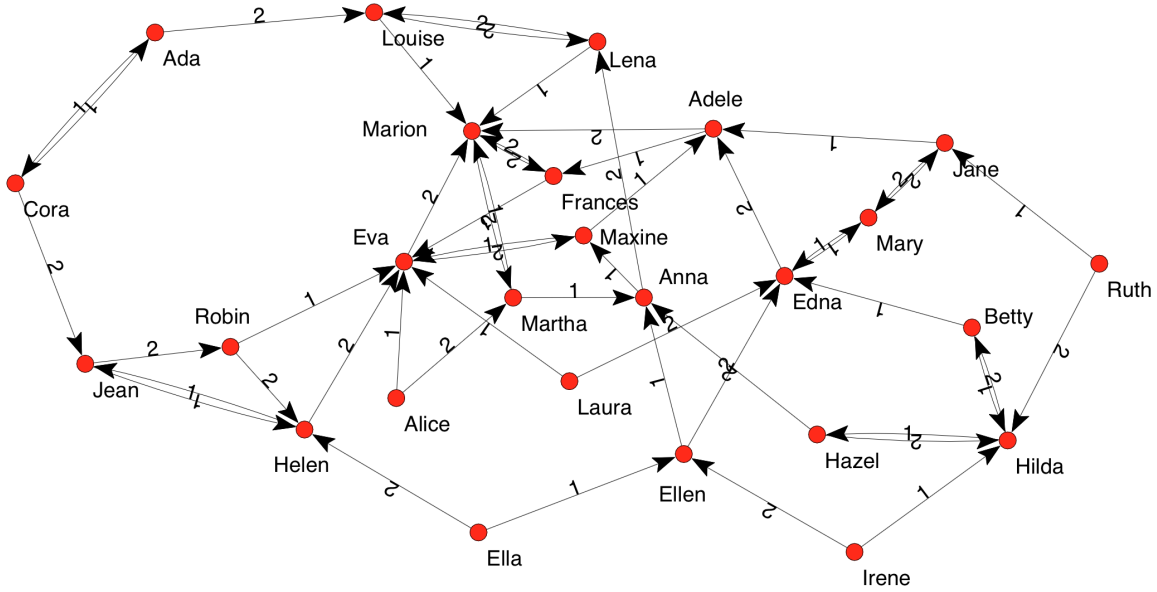
Ranking (best friend, second best friend,...)

Type (friend, relative, classmate,...)

Properties depending on the structure of the rest of the graph

Network components: Directed networks

Girls' school dormitory dining-table partners, 1st and 2nd choices (Moreno, The sociometry render, 1960)



Data representations

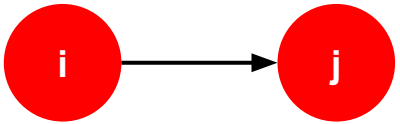
Adjacency Matrix

Edgelist

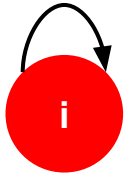
Adjacency List

Data representations: Adjacency Matrices

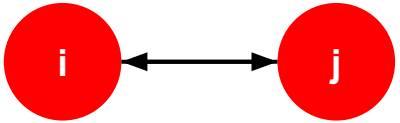
Represents edges (who is adjacent to whom) as matrix



Aij = 1 if node j has an edge to node i
= 0 if node j does not have an edge to i

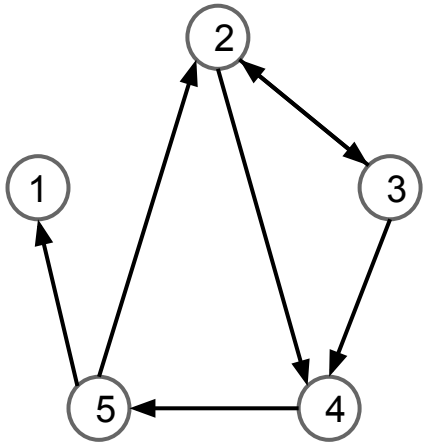


Aii = 0 unless the network has self-loops



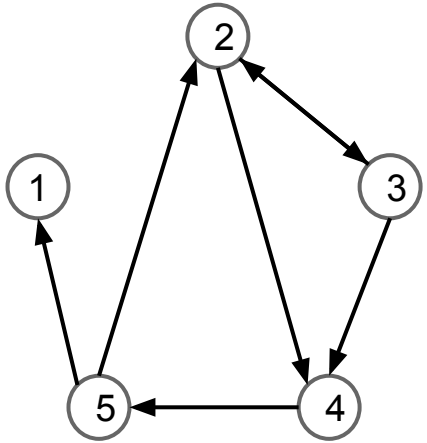
Aij = Aji if the network is undirected, or if i and j share a reciprocated edge

Data representations: Adjacency Matrix



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Data representations: Edge list

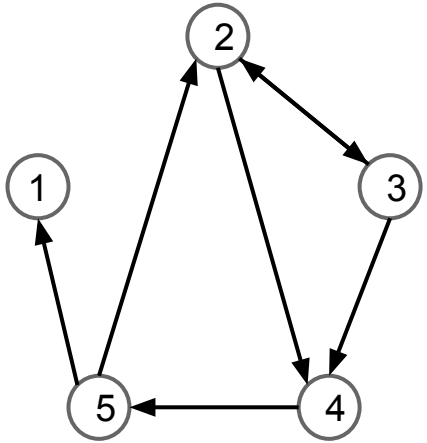


Edgelist =

2, 3
2, 4
3, 2
3, 4
4, 5
5, 2
5, 1

Data representations: Adjacency list

is easier to work with if network is large and sparse, retrieves all neighbours for a node.



1:
2: 3, 4
3: 2, 4
4: 5
5: 1, 2

Metrics

Degree

Connected Components

Metrics: Degree

Which node has the most edges?

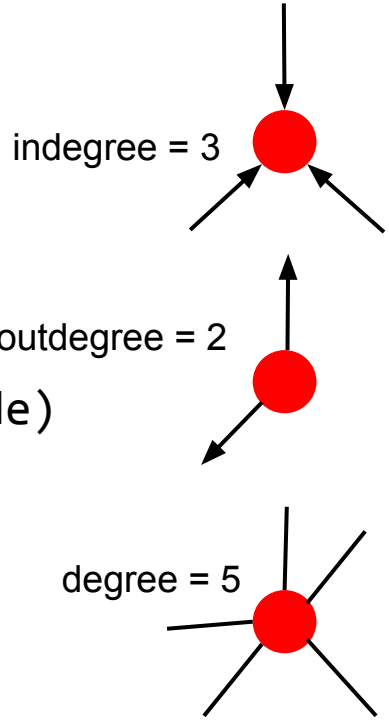
Node network properties

from immediate connections

indegree (how many edges arrives to a node)
outdegree (how many edges come out from a node)
degree (number of edges in a node)

from entire graph

centrality



Metrics: Degree from matrix values

Indegree = $\sum_{j=1} A_{ij}$ $\sum_{j=1} A_{3j} = 2$
(rows)

Outdegree = $\sum_{i=1} A_{ij}$ $\sum_{i=1} A_{i3} = 1$
(columns)

Degree = Indegree + Outdegree

A =

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Metrics: Connected components

Strongly connected components: Each node within the component can be reached from every other node in the component by following directed links.

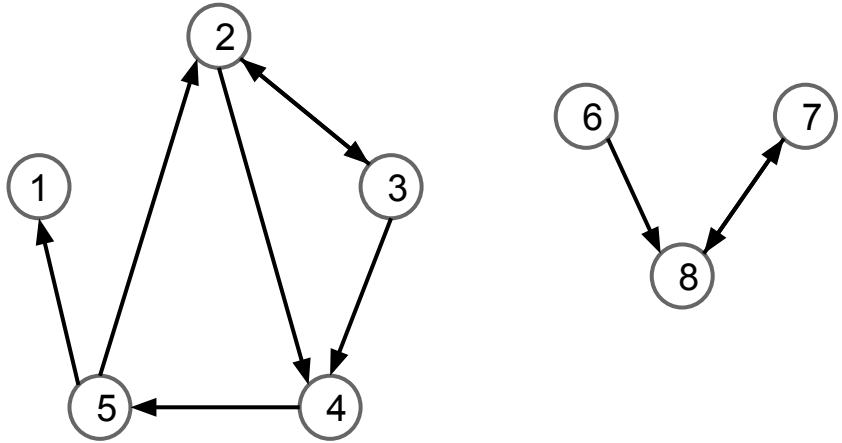
components:

5, 4, 3, 2

1

7, 8

6



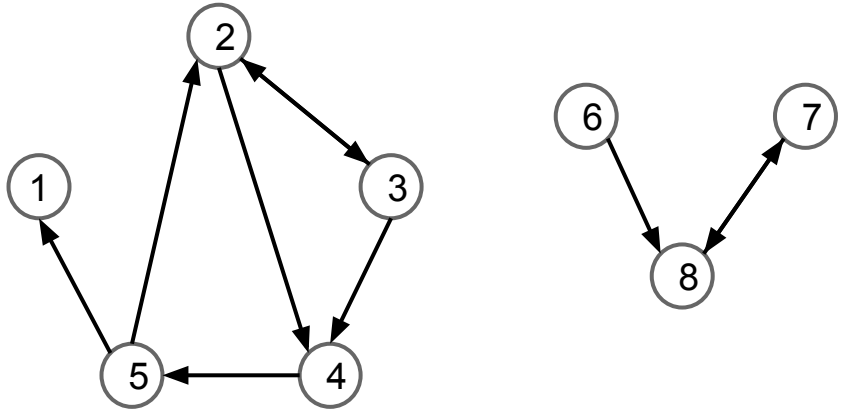
Metrics: Connected components

Weakly connected components: every node can be reached from every other node by following links in either direction.

components:

1, 2, 3, 4, 5

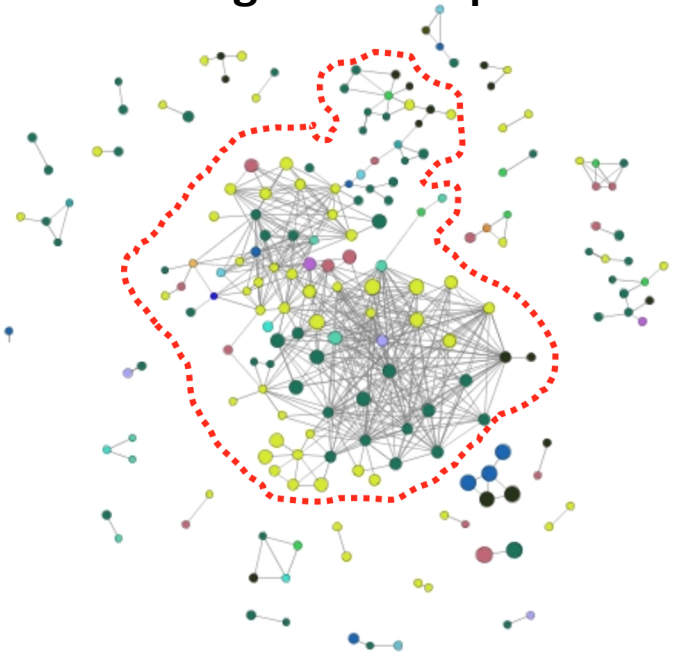
6, 7, 8



In undirected components on talks simply about connected components

Metrics: Connected components

If the largest component encompasses a significant fraction of the graph, it is called **the giant component**



Recap:

- Networks can be represented as matrices
- Useful metrics:
  - Degree
  - Connected components

Activity: Facebook network analysis

We are going to analyse facebook's user networks (with Gephi).

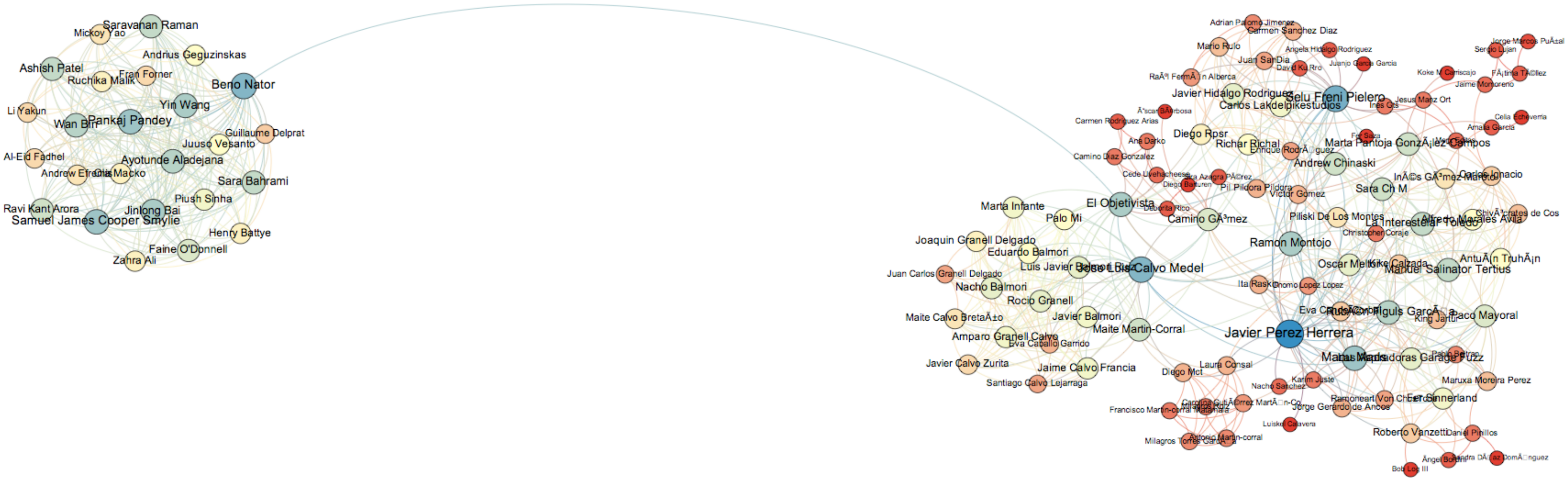
1. Go to Facebook App GetNET
2. Choose which user data you'd like to include
3. Be patient if you have many friends
4. Open gephi and load your file
5. Layout Force Atlas 2, set gravity under 50
6. Select Giant Component
7. Size nodes by degree
8. Color nodes by degree

Assignment: Facebook network analysis

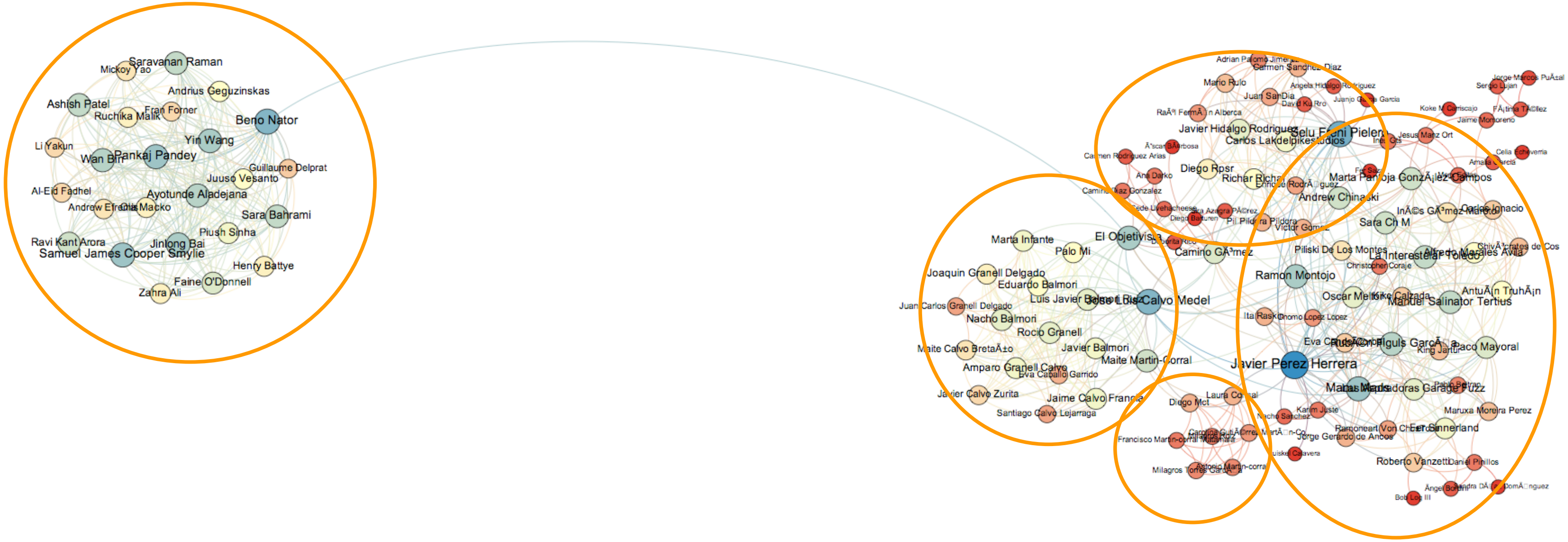
Presentation with your Facebook network analysis, explaining why your network looks the way it looks, answering the following questions:

brief explanation of your network (visual graph)
number of nodes:
number of edges:
number of connected components:
size of largest connected component:
maximum degree:

3 slides


Saravanan Raman
Mickoy Yao
Andrius Geguzinskas
Ashish Patel
Ruchika Malik
Fran Forner
Beno Nator
Yin Wang
Li Yakun
Pankaj Pandey
Guillaume Delprat
Juuso Vesanto
Al-Eid Fadhel
Ayotunde Aladejana
Sara Bahrami
Piush Sinha
Ravi Kant Arora
Jinlong Bai
Samuel James Cooper Smylie
Henry Battye
Faine O'Donnell
Zahra Ali
Adrian Palomo Jimenez
Carmen Sanchez Diaz
Mario Rulo
Juan SanDia
Javier Hidalgo Rodriguez
Selu Freni Pielero
Carlos Lakdelpikestudios
Diego Rpsr
Richar Richal
Marta Pantoja
Andrew Chinaski
Sara Ch M
El Objetivista
Marta Infante
Palo Mi
Camino GÃ³mez
Ramon Montojo
Joaquin Granell Delgado
Eduardo Balmori
Juan Carlos Granell Delgado
Nacho Balmori
Rocio Granell
Maite Calvo BretaÃ±o
Javier Balmori
Maite Martin-Corral
Amparo Granell Calvo
Eva Caballo Garrido
Javier Calvo Zurita
Jaime Calvo Francia
Santiago Calvo Lejarraga
Javier Perez Herrera
Oscar Melfon
Manuel Salinator Tertius
Paco Mayoral
Laura Consal
Diego Mct
Maruxa Moreira Perez
Roberto Vanzetti
Francisco Martin-corral
Milagros Torres

CIFF
Centro Internacional
de Formación Financiera



Important Links:

Networks, Crowds, and Markets (Chapter 1 and 2)

Suggested readings:

Learn how to use Gephi

Gephi tutorial on how to analyze facebook networks

The science of friendship

CIFF

Centro Internacional
de Formación Financiera

CIFF Trustees:

Santander

Universidad de Alcalá