

Aprende  
Teoría de Grafos  
con  
las tramas de corrupción  
en España

@mariamedp



# HELLO! My name is María



**MARÍA MEDINA**

 **@mariamedp**

Data Scientist en Microsoft

RLady + PyLady

## casos-aislados.com



## casos-aislados.com

<b>Caso 3 per cent</b> <ul style="list-style-type: none"><li>Partido/Org: <b>CDC</b></li><li>Lugar: <b>Cataluña</b></li><li>Coste: <b>1.800.000.000</b></li></ul> 	<b>Caso Abogados Melilla</b> <ul style="list-style-type: none"><li>Partido/Org: <b>PP</b></li><li>Lugar: <b>Melilla</b></li><li>Coste: <b>1.114.510</b></li></ul> 	<b>Caso Aceinsa</b> <ul style="list-style-type: none"><li>Partido/Org: <b>PP</b></li><li>Lugar: <b>Castilla y León</b></li><li>Coste: <b>N/A</b></li></ul> 	<b>Caso ACM</b> <ul style="list-style-type: none"><li>Partido/Org: <b>PSOE</b></li><li>Lugar: <b>Andalucía</b></li><li>Coste: <b>500.000</b></li></ul> 
<b>Caso Acuamed</b> <ul style="list-style-type: none"><li>Partido/Org: <b>PP</b></li><li>Lugar: <b>Comunidad de Madrid</b></li><li>Coste: <b>20.000.000</b></li></ul> 	<b>Caso ADIF/AVE La Sagrera</b> <ul style="list-style-type: none"><li>Partido/Org: <b>Empresas</b></li><li>Lugar: <b>Nacional</b></li><li>Coste: <b>133.000.000</b></li></ul> 	<b>Caso ADIF/AVE Murcia</b> <ul style="list-style-type: none"><li>Partido/Org: <b>Empresas</b></li><li>Lugar: <b>Región de Murcia</b></li><li>Coste: <b>25.000.000</b></li></ul> 	<b>Caso Adigsa</b> <ul style="list-style-type: none"><li>Partido/Org: <b>CDC</b></li><li>Lugar: <b>Cataluña</b></li><li>Coste: <b>120.031</b></li></ul> 

## casos-aislados.com

### Resumen

- Partido/Org: **CDC**
- Lugar: **Cataluña**
- Coste aprox.: **1.800.000.000 €**
- Total implicados: **10**
- Año de los hechos: **2008**
- Estado judicial: **Abierto**



Actualizado el: 28 de Enero del 2019

### Anuncios para la Caja A de CA

### Implicados



**SIXTE CAMBRA I SÁNCHEZ**

- Situación Judicial:** Imputado
- Cargo:** Ex-presidente del Puerto de Barcelona

---



**PARTIDO DEMÓCRATA EUROPEO CATALÁN**

- Situación Judicial:** Imputado
- Delito:** tráfico de influencias, cohecho y blanqueo de capitales
- Cargo:** Partido Político

Imputado por el juez De La Mata el 27/07/2018 en condición de personas jurídicas.

## Grafo de casos



Casos



Personas implicadas en ambos



Nº personas en común

## Grafo de implicados



Personas



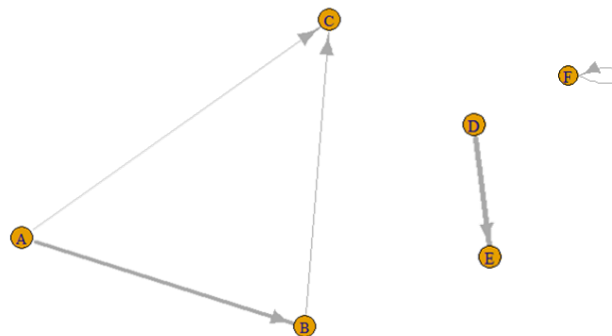
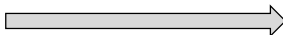
Implicadas en el mismo caso



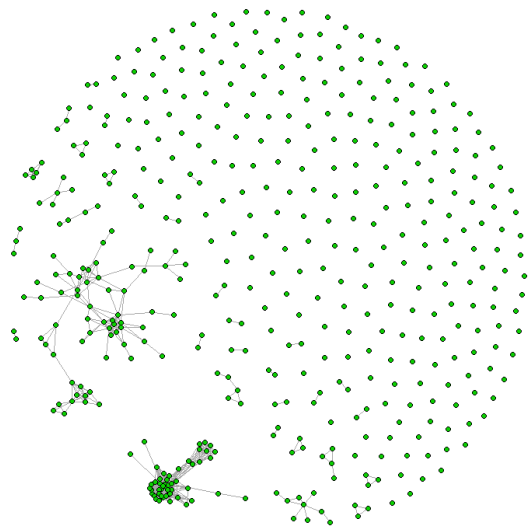
Nº casos en común

```
library(igraph)
g <- graph_from_data_frame(df_links)
```

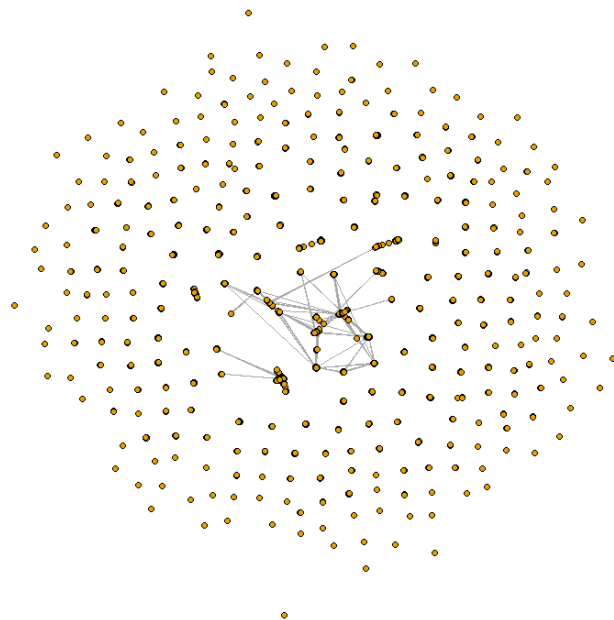
node1	node2	weight
A	B	3.5
A	C	1.5
B	C	2
D	E	5
F	F	1



## Grafo de casos

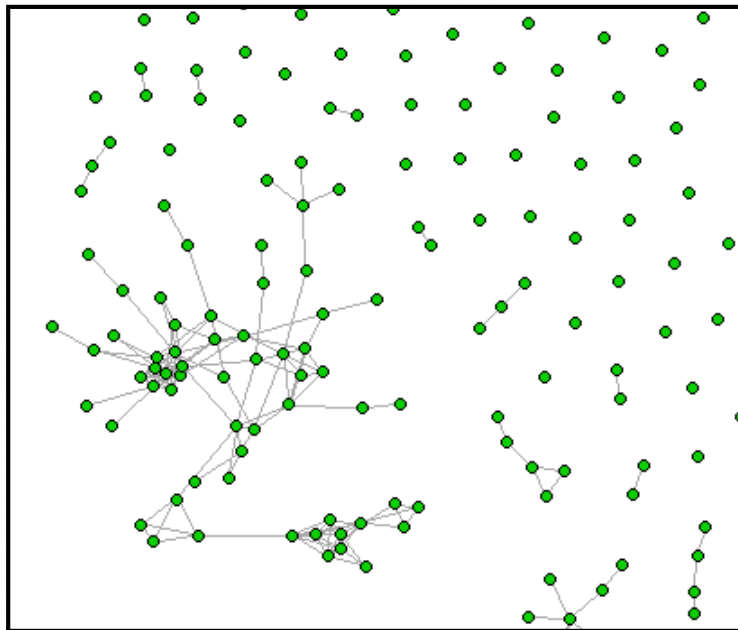


## Grafo de implicados

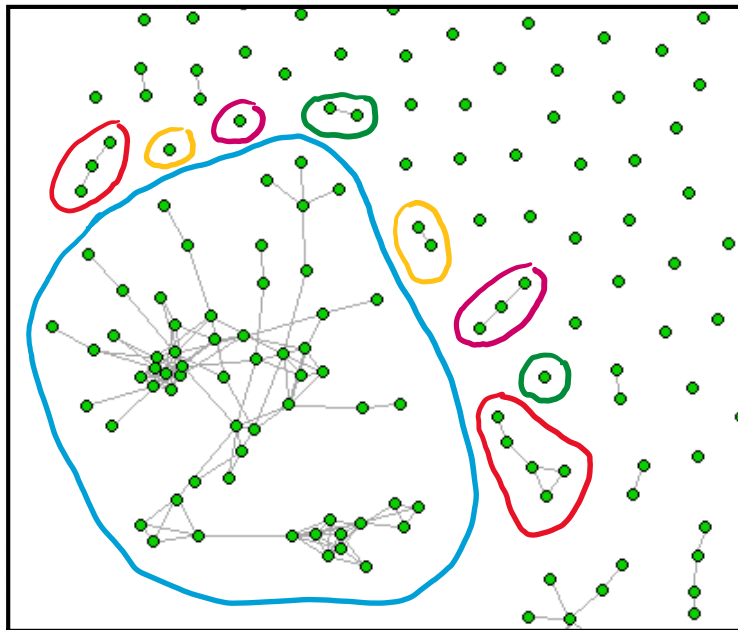




# Componentes conexos

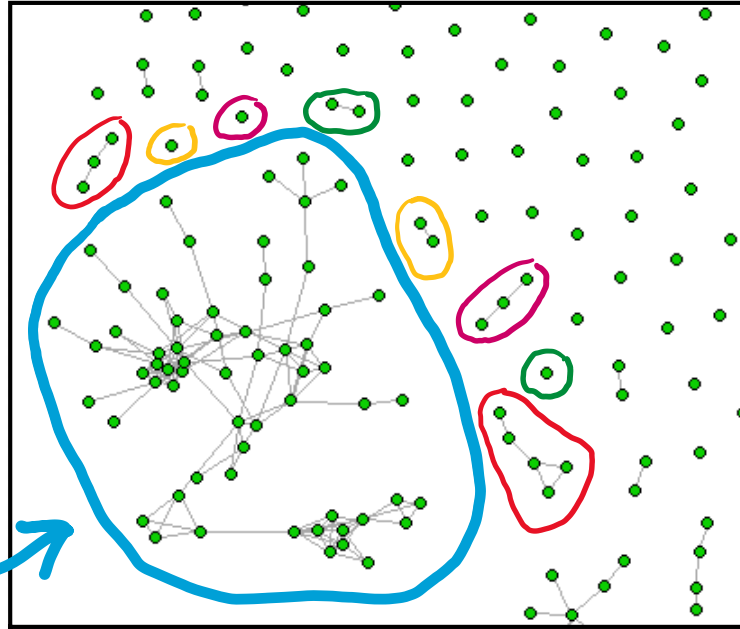


# Componentes conexas



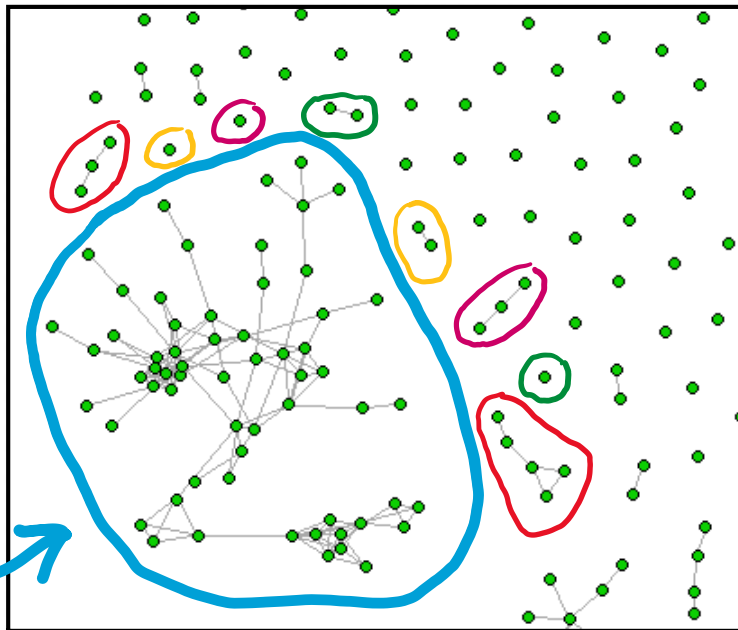
# Componentes conexas

Componente conexas  
más grande  
(***giant component***)



# Componentes conexas

Componente conexa  
más grande  
(***giant component***)



Componentes  
=  
tramas

```
g_components <- components(g)
```

```
> str(g_components)
```

```
List of 3
```

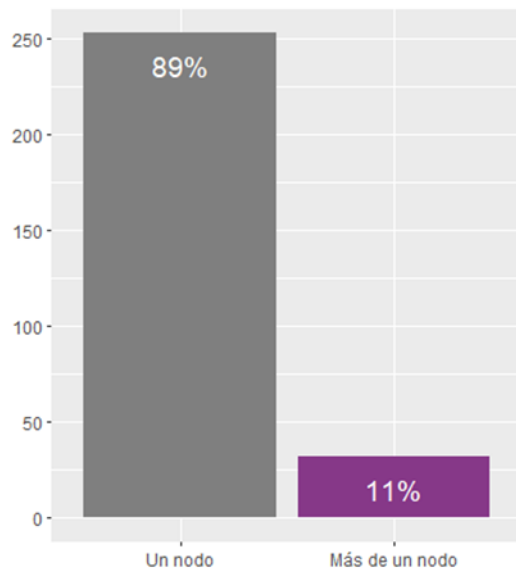
```
$ membership: Named num [1:6] 1 1 2 1 2 3
```

```
..- attr(*, "names")= chr [1:6] "A" "B" "D" "C" "E" "F"
```

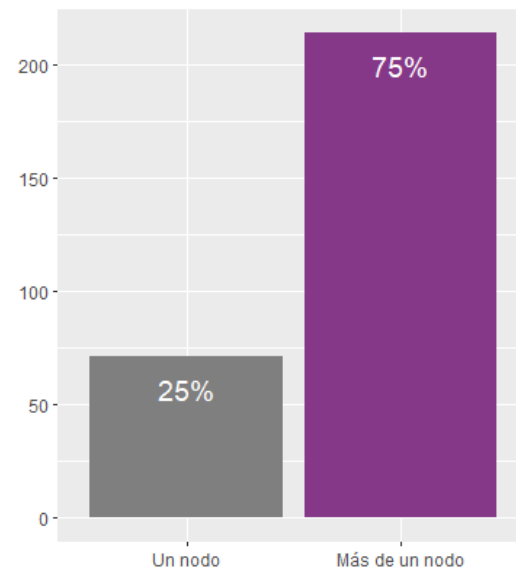
```
$ csize      : num [1:3] 3 2 1
```

```
$ no         : int 3
```

## Grafo de casos



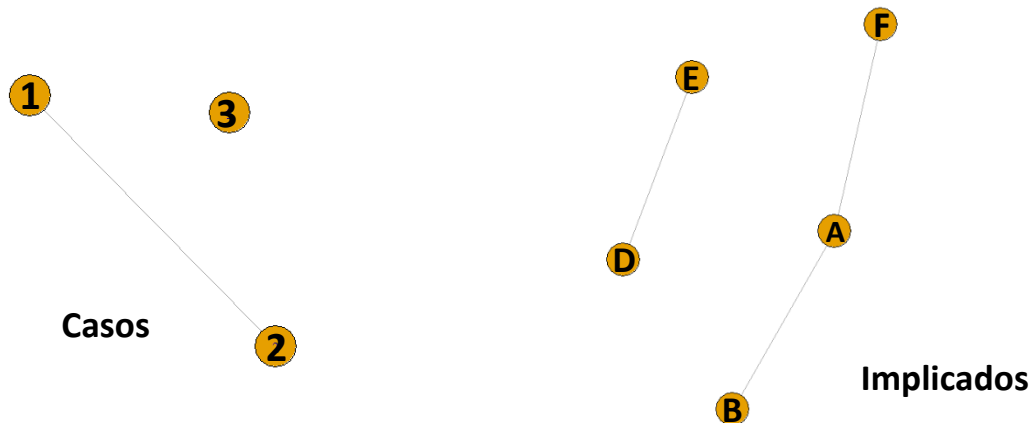
## Grafo de implicados



¿Cómo es el grafo de implicados de casos en tramas complejas?

¿Cómo es el grafo de implicados de casos en tramas complejas?

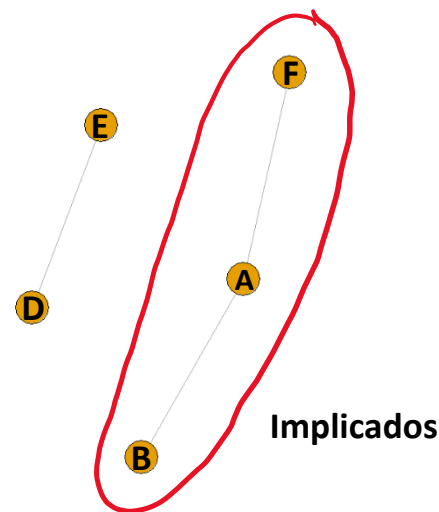
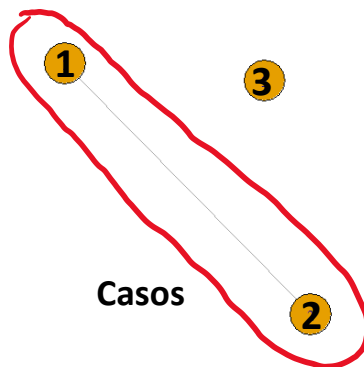
caso	implicado
1	A
1	B
2	A
2	C
3	D
3	E



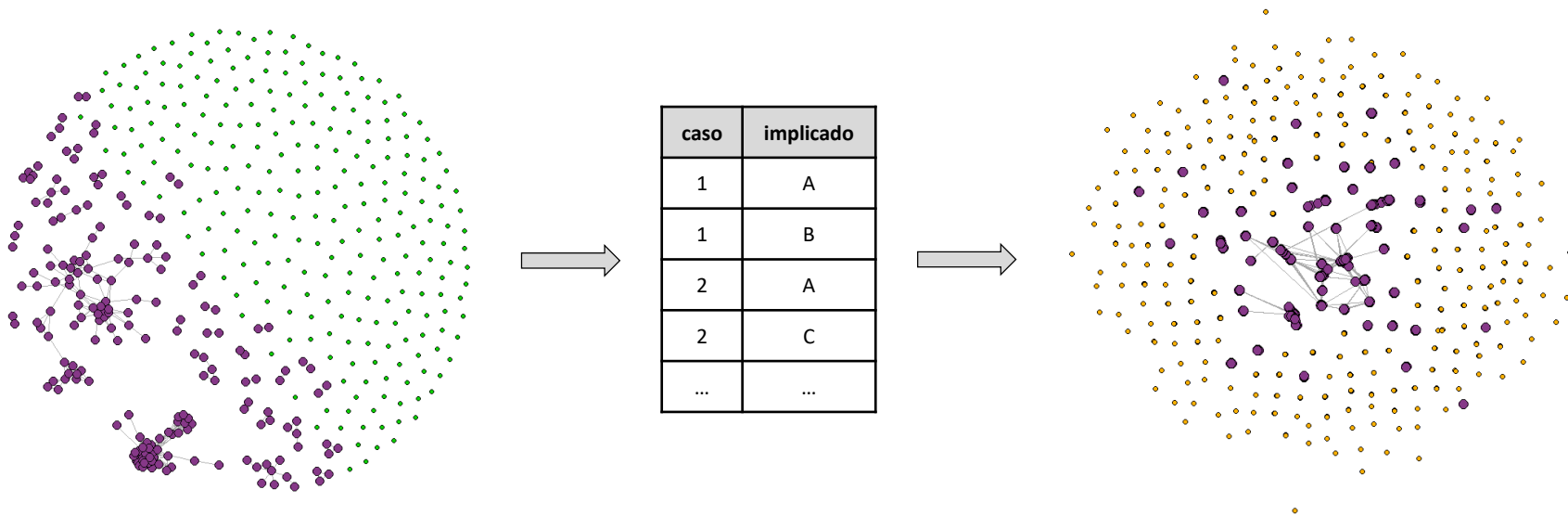


¿Cómo es el grafo de implicados de casos en tramas complejas?

caso	implicado
1	A
1	B
2	A
2	C
3	D
3	E



¿Cómo es el grafo de implicados de casos en tramas complejas?



¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$ccsize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- subset(implicados_casos, caso %in% casos_tramas)

ind_impl_trama <- match(implicados_tramas$nombre_implicado, v(g_impl)$name)
g_impl_trama <- induced_subgraph(g_impl, vids=ind_impl_trama)
```

¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$csize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- > str(casos_comp)
ind_impl_trama <- ma
g_impl_trama <- indu
```

```
List of 3
 $ membership: Named num [1:3] 1 1 2
  ..- attr(*, "names")= chr [1:3] "1" "2" "3"
 $ csize      : num [1:2] 2 1
 $ no        : int 2
```

¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$csize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- > str(casos_comp)

ind_impl_trama <- ma
g_impl_trama <- indu
```

```
List of 3
 $ membership: Named num [1:3] 1 1 2
  ..- attr(*, "names")= chr [1:3] "1" "2" "3"
 $ csize      : num [1:2] 2 1
 $ no        : int 2
```

¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$csizes > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- > str(casos_comp)

ind_impl_trama <- ma
g_impl_trama <- indu
```

```
List of 3
 $ membership: Named num [1:3] 1 1 2
  ..- attr(*, "names")= chr [1:3] "1" "2" "3"
 $ csizes      : num [1:2] 2 1
 $ no          : int 2
```

¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$csize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- > str(casos_comp)

ind_impl_trama <- ma
g_impl_trama <- indu
```

```
List of 3
 $ membership: Named num [1:3] 1 1 2
  ..- attr(*, "names")= chr [1:3] "1" "2" "3"
 $ csize      : num [1:2] 2 1
 $ no        : int 2
```

¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$ccsize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- subset(implicados_casos, caso %in% casos_tramas)
```

```
ind_impl_trama <- match(implicados_tramas$nombre_implicado, nombres_implicados)
g_impl_trama <- induced_subgraph(g_impl, vids=ind_impl_trama)
```

```
> implicados_tramas
```

	caso	nombre_implicado
1	1	A
2	1	B
3	2	A
4	2	C



¿Cómo es el grafo de implicados de casos en tramas complejas?

```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$ccsize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- subset(implicados_casos, caso %in% casos_tramas)

ind_impl_trama <- match(implicados_tramas$nombre_implicado, v(g_impl)$name)
g_impl_trama <- induced_subgraph(g_impl, vids=ind_impl_trama)
```

```
> implicados_tramas
```

	caso	nombre_implicado
1	1	A
2	1	B
3	2	A
4	2	C

¿Cómo es el grafo de implicados de casos en tramas complejas?

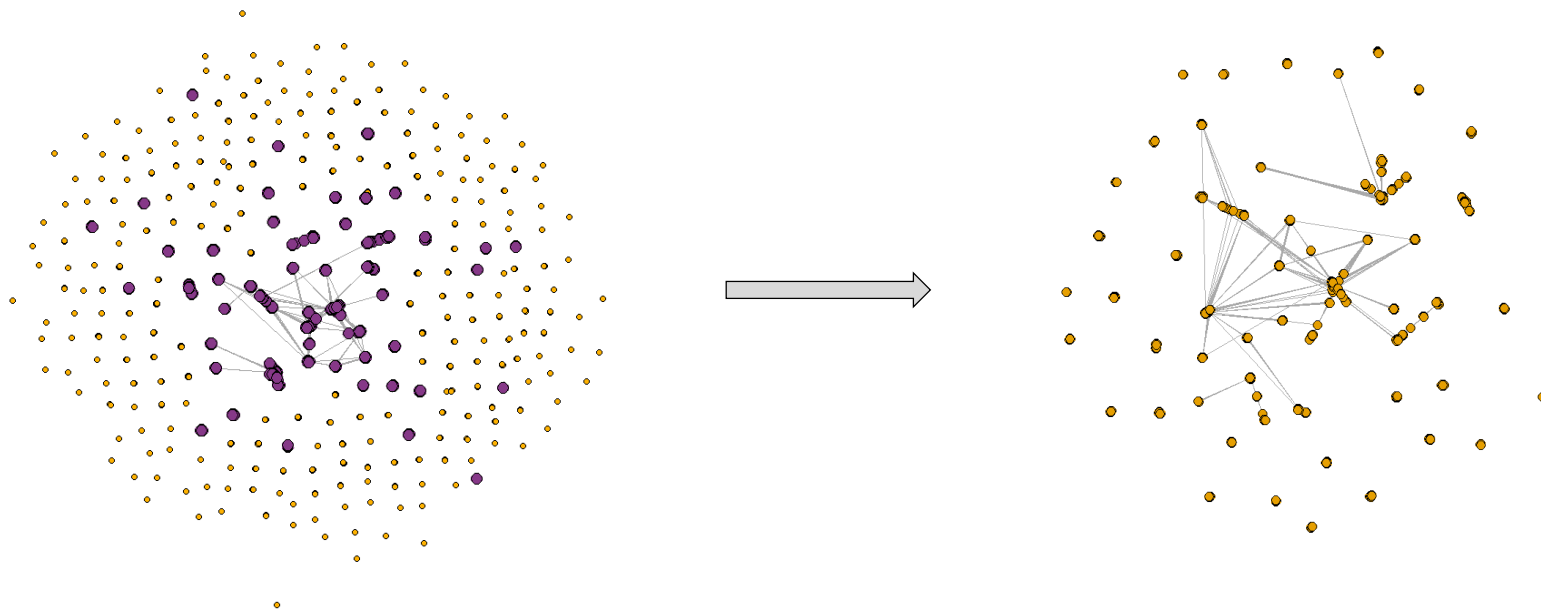
```
casos_comp <- components(g_casos)

ind_comp_grandes <- which(casos_comp$ccsize > 1)
ind_casos_tramas <- casos_comp$membership %in% ind_comp_grandes
casos_tramas <- names(casos_comp$membership)[ind_casos_tramas]

implicados_tramas <- subset(implicados_casos, caso %in% casos_tramas)

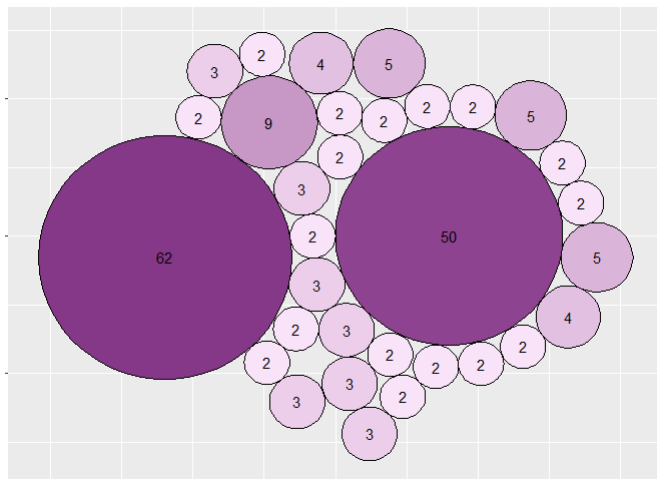
ind_impl_trama <- match(implicados_tramas$nombre_implicado, v(g_impl)$name)
g_impl_trama <- induced_subgraph(g_impl, vids=ind_impl_trama)
```

¿Cómo es el grafo de implicados de casos en tramas complejas?



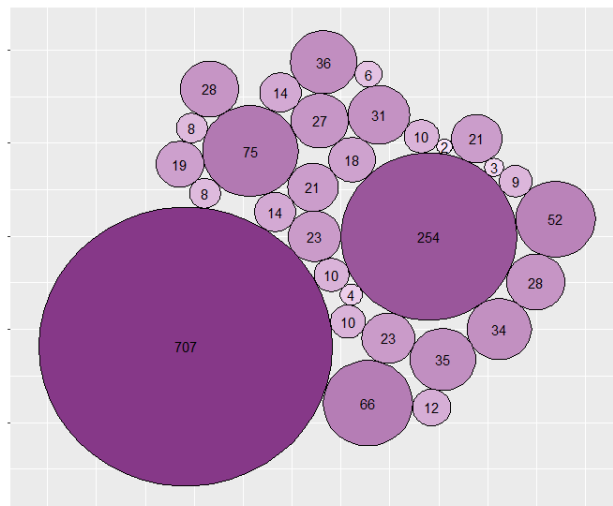
# Grafos inducidos – componentes conexas

## Grafo de casos



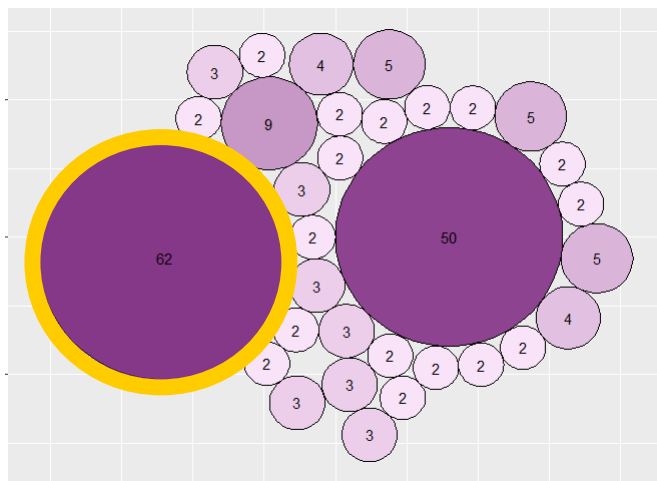
Tamaño de componentes conexas

## Grafo de implicados



Tamaño de componentes conexas

## Grafo de casos



## Tamaño de componentes conexas

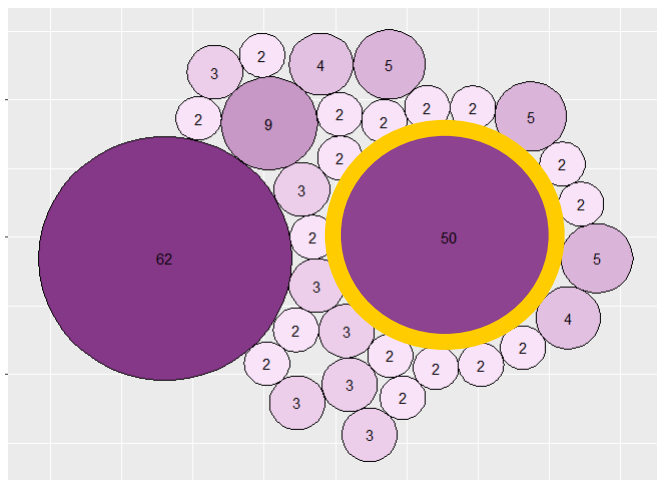
## Madrid/Valencia/Baleares

PP/UM/...

845 implicados

- Casos Gürtel
- Casos Púnica
- Casos Lezo
- Caso Nóos
- Caso Trajes
- Caso Tarjetas Black
- Caso Bárcenas
- Caso Palma Arena
- Caso Taula
- ...

## Grafo de casos



Tamaño de componentes conexas

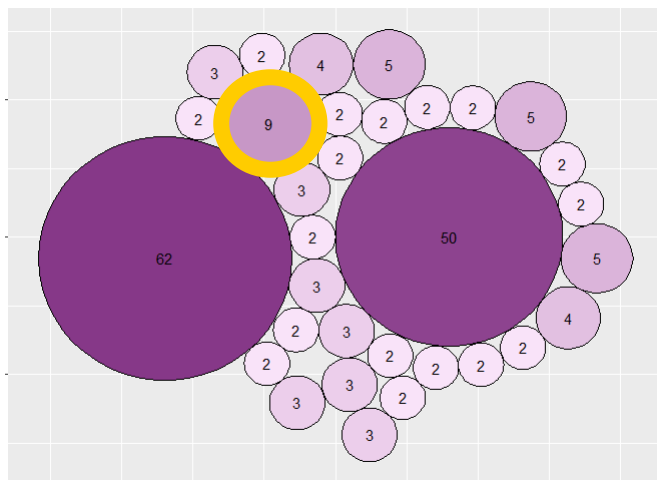
Andalucía

PSOE/UGT

420 implicados

- Casos ERE
- Caso Facturas Falsas
- Caso de los Avales
- Caso Cursos Formación
- ...

## Grafo de casos



Tamaño de componentes conexas

Galicia

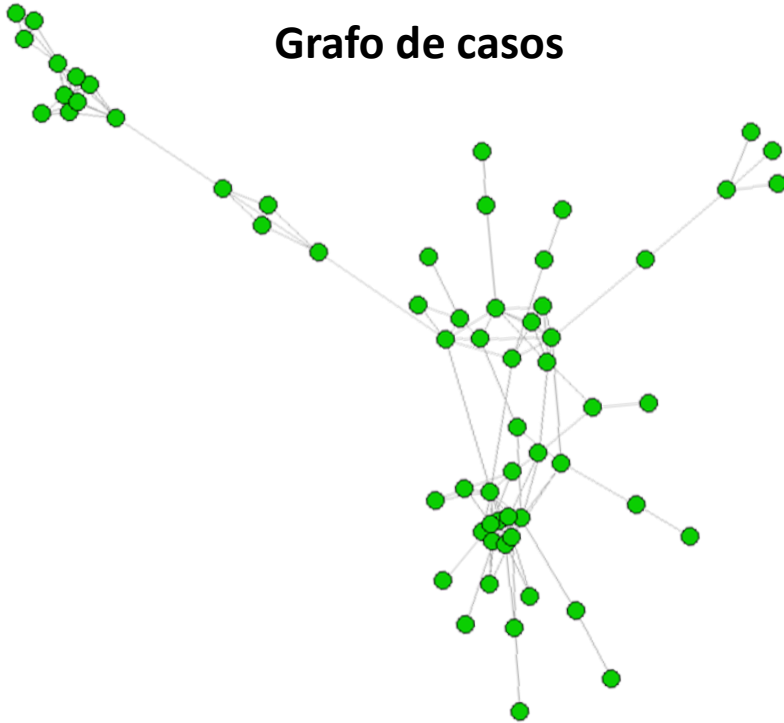
PP/PSOE

87 implicados

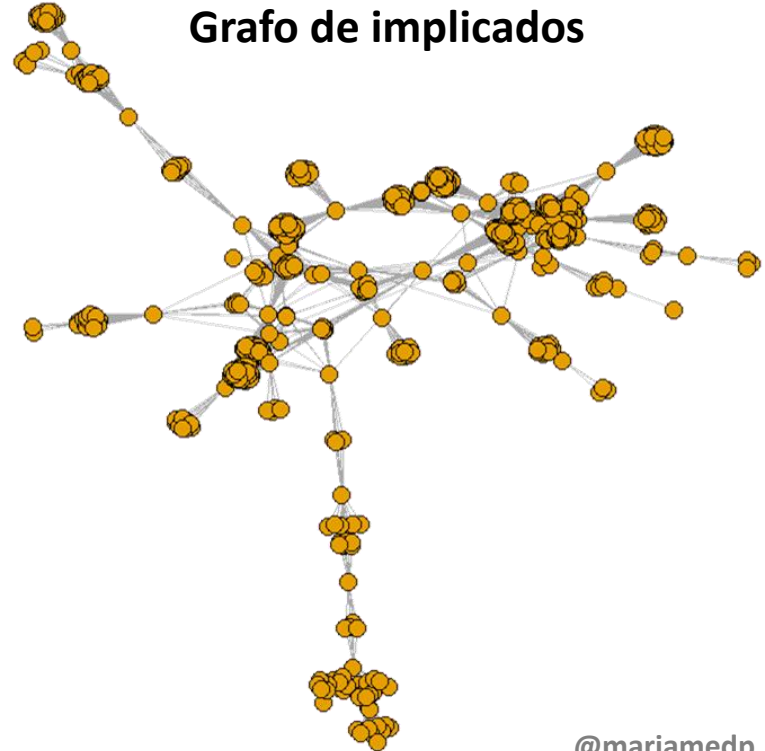
- Casos Pokémon
- Caso Campeón
- Caso Baltar
- ...

# Giant component (Trama 1)

**Grafo de casos**



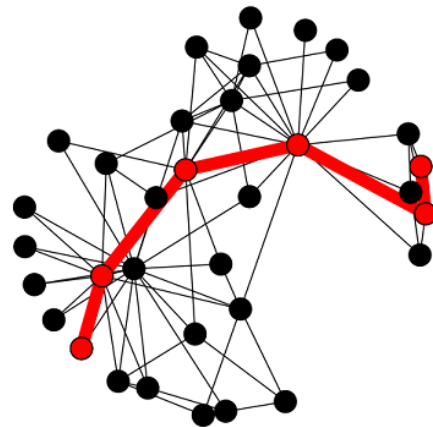
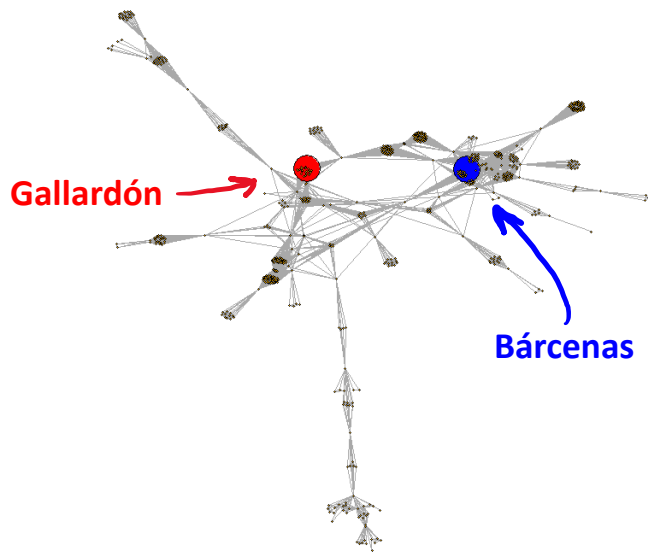
**Grafo de implicados**





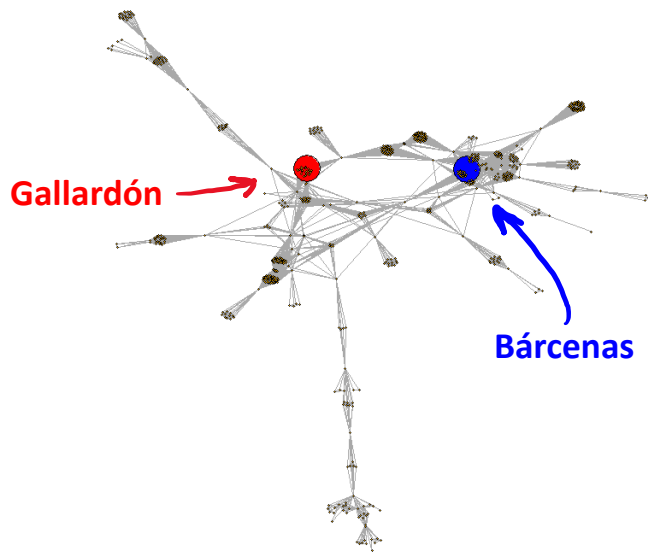
# Caminos más cortos (shortest path)

¿Cuántos pasos hay entre Gallardón y Bárcenas?



# Caminos más cortos (shortest path)

¿Cuántos pasos hay entre Gallardón y Bárcenas?

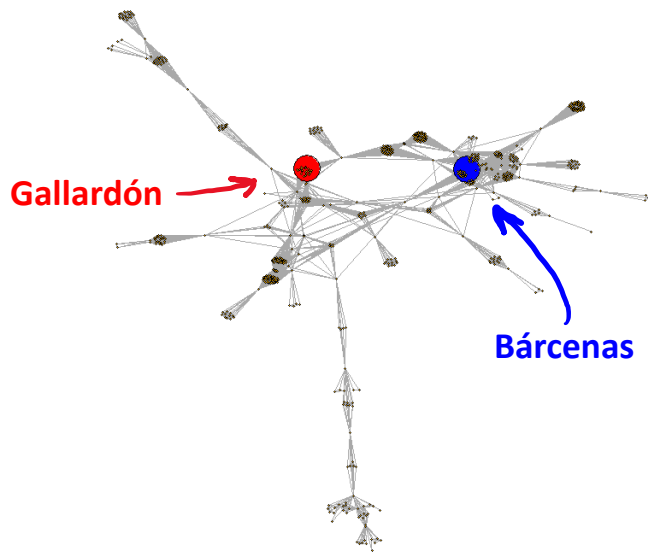


```
paths <- shortest_paths(g,  
  from="Gallardón", to="Bárcenas")
```

```
> paths$vp[1]  
+ 5/707 vertices, named, from 034492f:  
[1] Alberto Ruiz-Gallardón Jiménez  
[2] Ignacio González González  
[3] Francisco Javier López Madrid (Compiyogui)  
[4] Alberto López Viejo  
[5] Luis Francisco Bárcenas Gutiérrez
```

# Caminos más cortos (shortest path)

¿Cuántos pasos hay entre Gallardón y Bárcenas?



```
path_length <- distances(g,  
  v="Gallardón", to="Bárcenas")
```

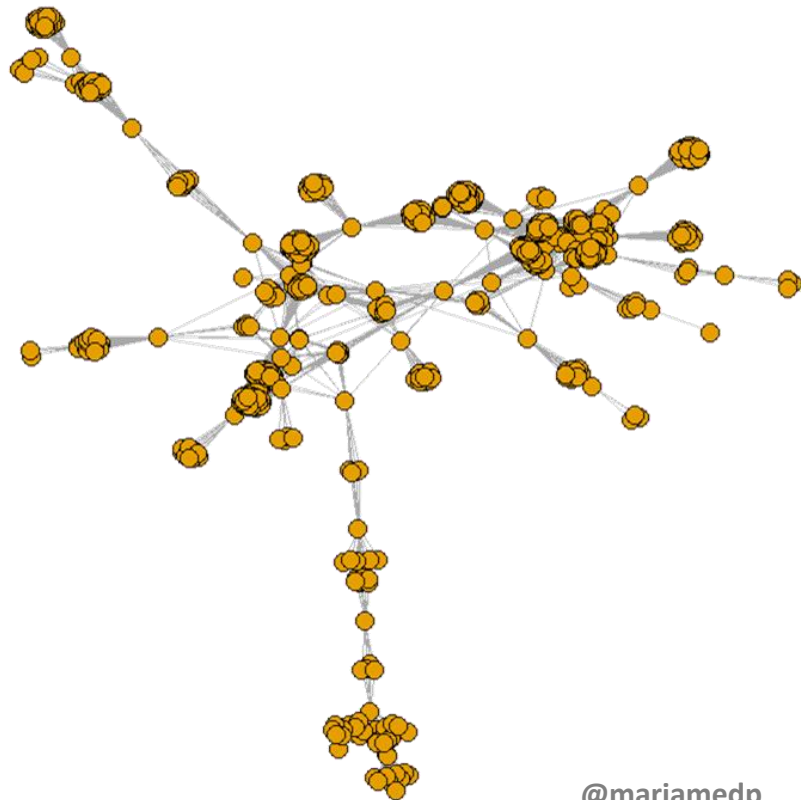
```
> path_length  
      Bárcenas  
Gallardón      4
```

# Small World Network

```
paths_lengths <- distances(g)
```

```
> dim(paths_lengths)
```

```
[1] 707 707
```



# Small World Network

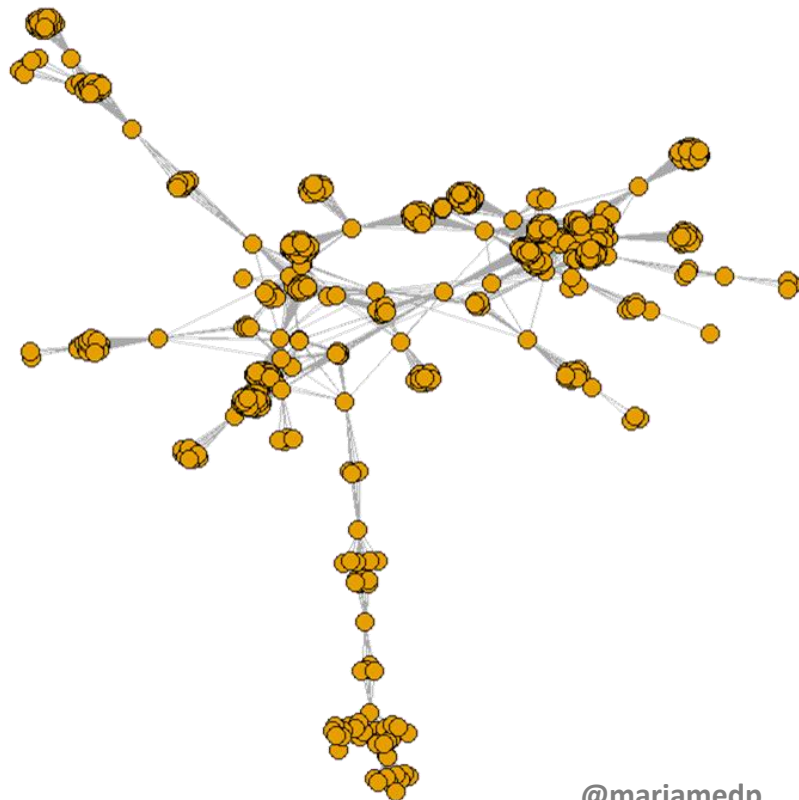
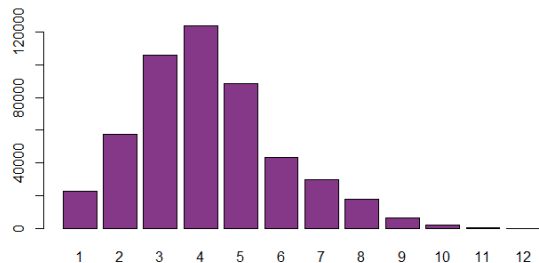
```
paths_lengths <- distances(g)
```

```
> dim(paths_lengths)
```

```
[1] 707 707
```

```
> mean(paths_lengths)
```

```
[1] 4.194039
```



## Centralidad de grado *¿Cuántos amigos tengo?*

↳ Grado del nodo

## Betweenness *¿Cómo de imprescindible soy?*

↳ Nº de shortest paths que pasan por el nodo

## Closeness *¿Cómo de cerca estoy de todo el mundo?*

↳ Longitud de shortest paths a todos los nodos de la red

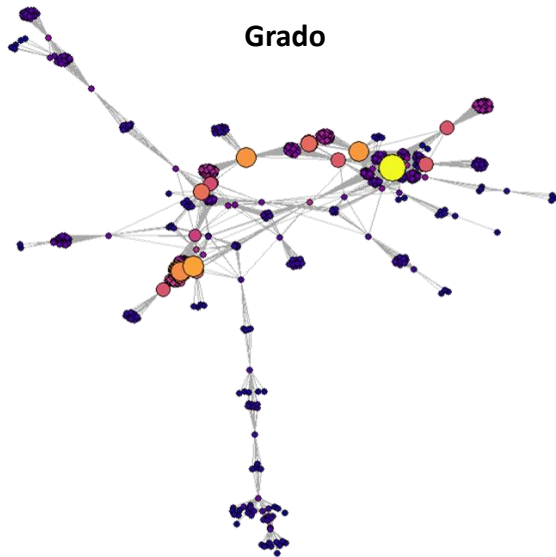
```
nodes_degree <- degree(g)
```

```
nodes_betweenness <- betweenness(g)
```

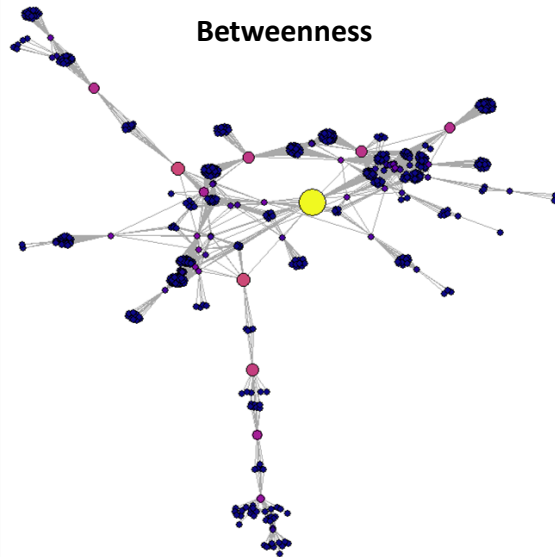
```
nodes_closeness <- closeness(g)
```

# Medidas de centralidad

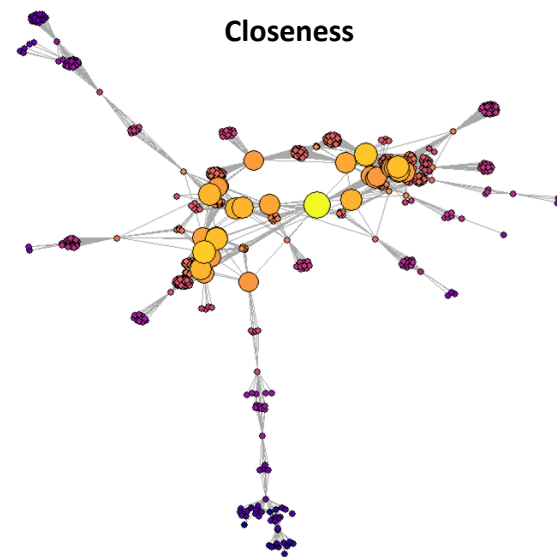
Grado



Betweenness



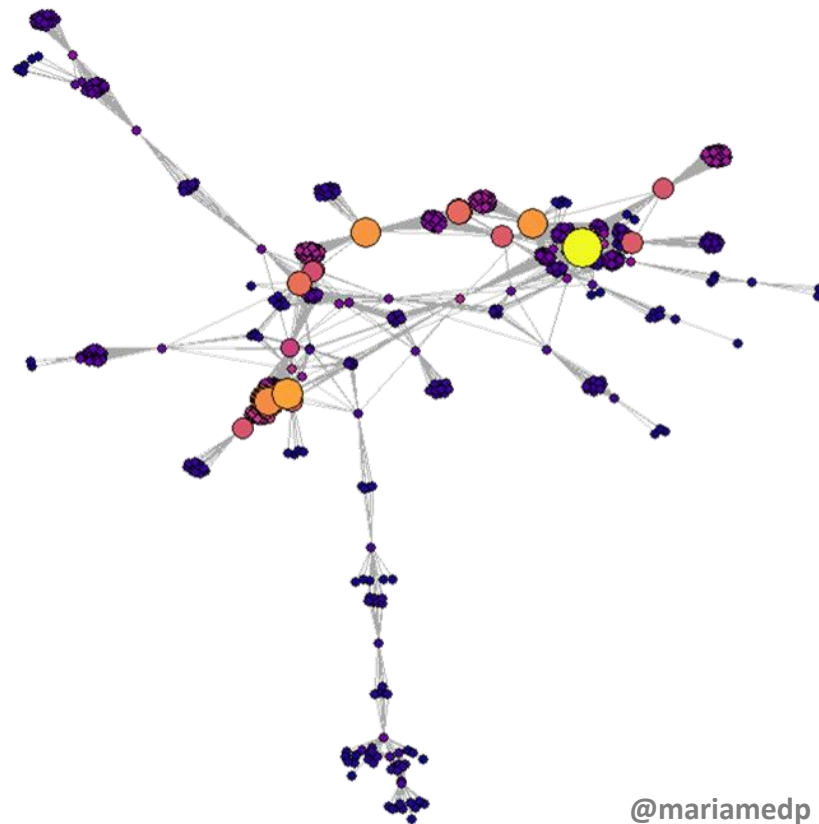
Closeness





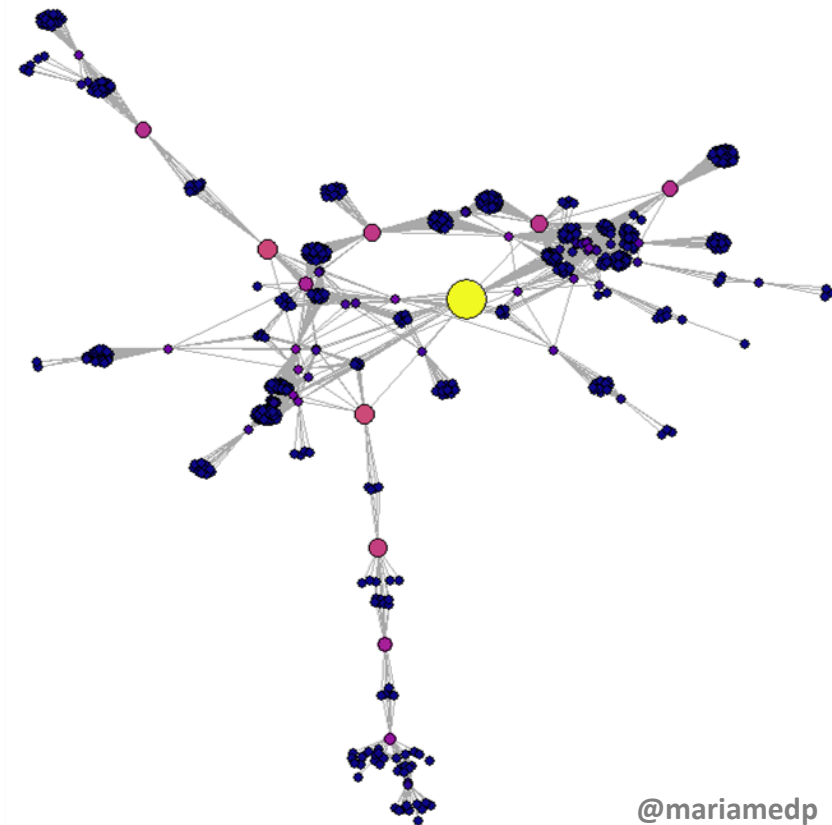
# Centralidad de grado

Implicado	Grado
Francisco Correa Sánchez	135
Pablo Crespo Sabarís	133
Arturo Fernández Álvarez	107
Partido Popular	103
Carlos Turró Homedes	102



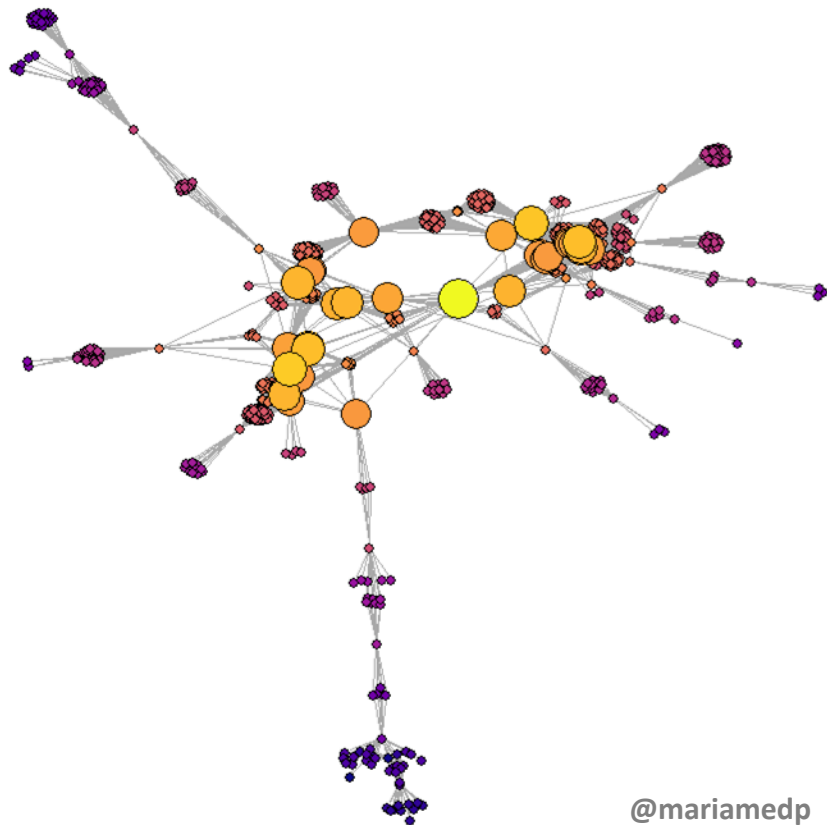
# Betweenness

Implicado	Betweenness (/1000)
Alberto López Viejo	7.62
Daniel Mercado Lozano	3.88
José Martínez Nicolás	3.82
Jaume Matas Palou	3.61
Carlos Turró Homedes	3.41

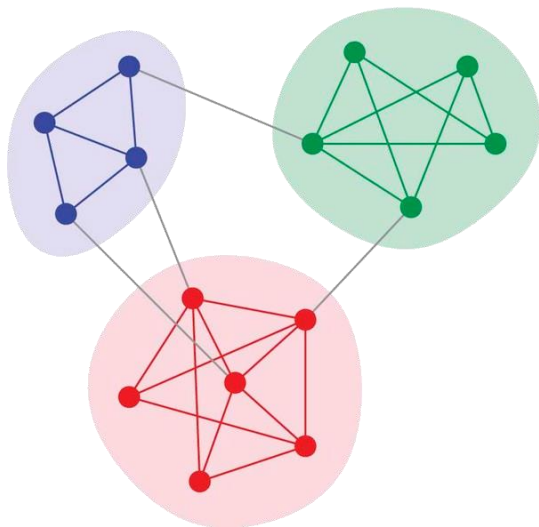


# Closeness

Implicado	Closeness (x1000)
Alberto López Viejo	0.520
Fco. Javier López Madrid (Compiyogui)	0.482
Partido Popular	0.477
Juan Miguel Villar Mir	0.471
J. Miguel Trabada Guijarro	0.471



Clustering de nodos en base a los enlaces y sus pesos.

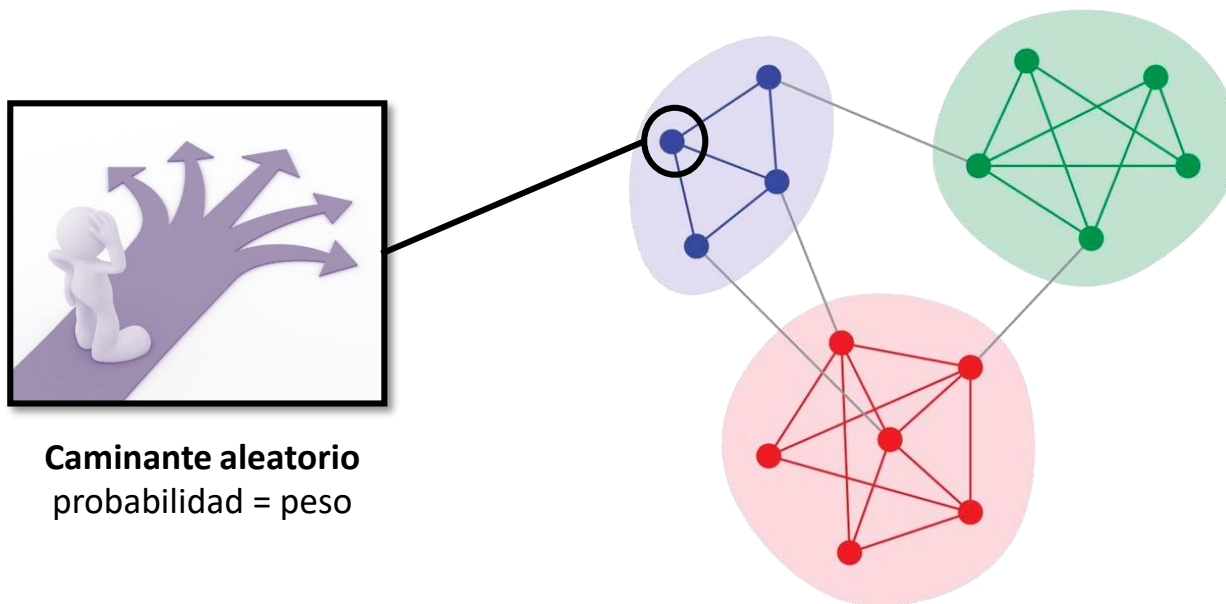


**Comunidad** = nodos densamente conectados

**Modularidad** = medida de lo bien separados que quedan los nodos

# Algoritmo Walktrap

Es un algoritmo basado en **caminos aleatorios** de  $N$  pasos.



**Caminante aleatorio**  
probabilidad = peso

# Algoritmo Walktrap

```
communities <- cluster_walktrap(g, steps=5)
```

```
> modularity(communities)
```

```
0.8023952
```

```
> length(communities)
```

```
17
```

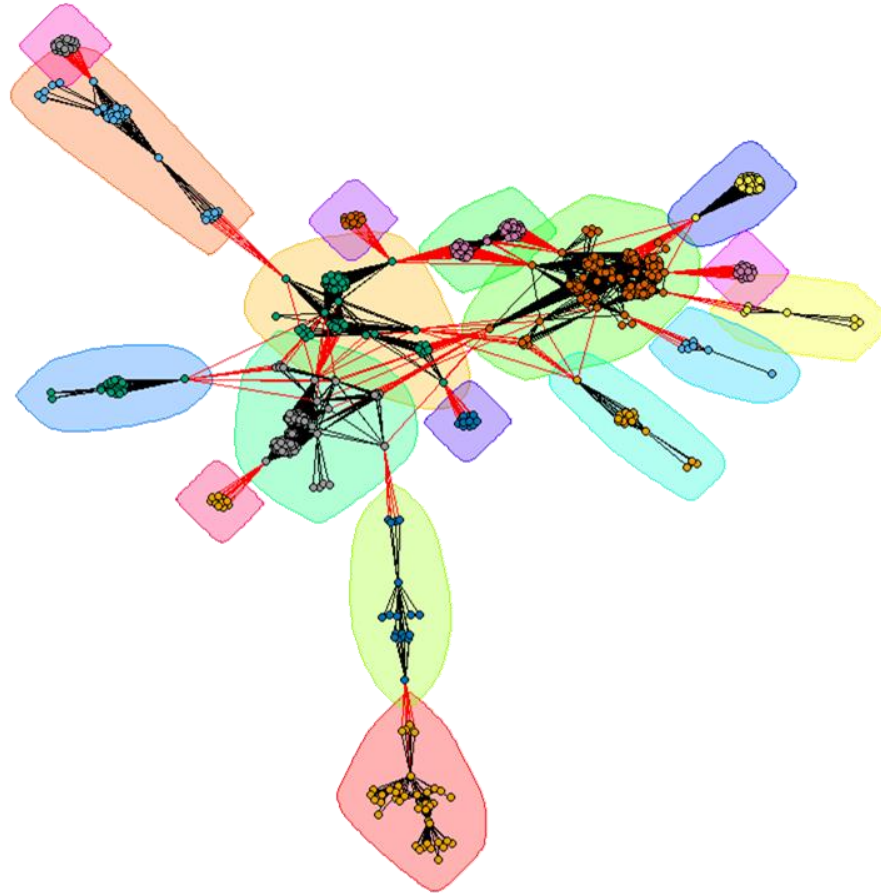
```
> sizes(communities)
```

```
Community sizes
```

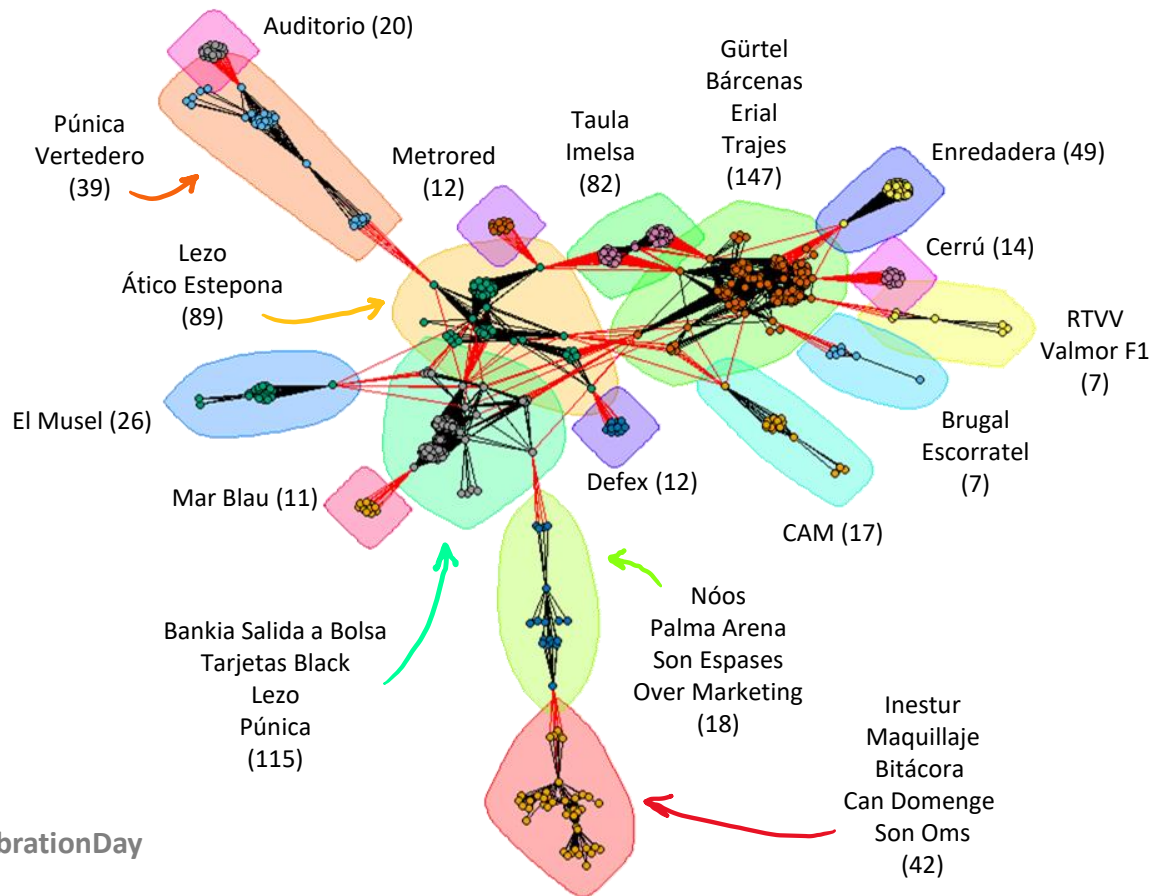
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
42	39	89	7	18	147	82	115	17	7	26	49	12	12	14	20	11

```
> plot(communities, g)
```

# Comunidades trama 1



# Comunidades trama 1







**LADA ADAMIC**

**ladamic.com**

Curso online (SI 508) en Open Michigan:

<https://open.umich.edu/find/open-educational-resources/information/si-508-networks-theory-application>



**¡ GRACIAS !**

@mariamedp