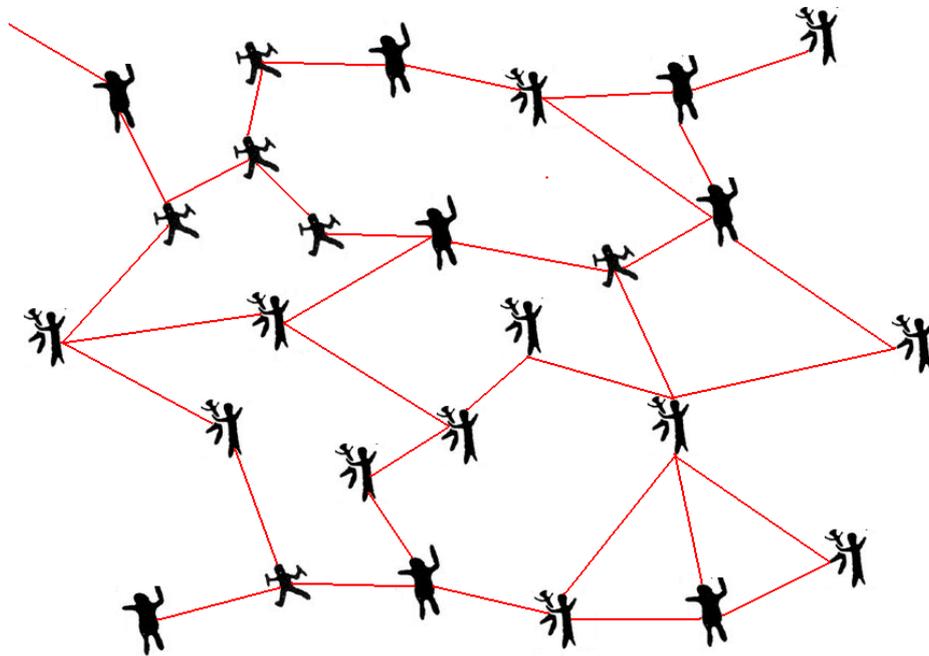


ANTROPOLOGÍA Y REDES SOCIALES

UNA INTRODUCCIÓN A UCINET6-NETDRAW, EGONET Y EL
ANÁLISIS COMPARADO CON SPSS



José Luis Molina y Javier Ávila (eds.)

Universidad Autónoma de Barcelona (UAB)

Universidad Nacional Federico Villareal (UNFV)

Agencia española de Cooperación Internacional (AECID)

ÍNDICE

INTRODUCCIÓN	3
OPERACIONES BÁSICAS CON UCINET 6	7
CÓMO INTRODUCIR DATOS DE RELACIONES (RETICULARES)	7
<i>HOJA DE DATOS</i>	7
<i>FICHERO DE TEXTO</i>	10
CÓMO CONVERTIR MATRICES DE MODO 2 EN MATRICES DE MODO 1	11
<i>COMPLETAR FILAS O COLUMNAS</i>	11
<i>AFILIACIONES</i>	15
CÓMO RECODIFICAR MATRICES	19
EL ANÁLISIS DE REDES SOCIOCÉNTRICAS CON UCINET 6 Y NETDRAW	22
INTRODUCCIÓN	22
DESCRIPCIÓN DEL EJERCICIO PRÁCTICO	24
CREAR UNA MATRIZ CON UCINET 6	25
GRAFICAR UNA RED CON NETDRAW	30
MEDIDAS DE CENTRALIDAD	36
<i>GRADO NODAL (DEGREE)</i>	37
<i>GRADO DE INTERMEDIACIÓN (BETWEENNESS)</i>	40
<i>CERCANÍA (CLOSENNESS)</i>	42
<i>REDES DENTRO DE REDES: CLIQUES</i>	44
<i>CLIQUES</i>	45
<i>N-CLIQUE, N-CLAN Y K-PLEX</i>	51
EL ANÁLISIS DE REDES PERSONALES CON EGONET	64
NOTAS SOBRE LA INSTALACIÓN DE EGONET	66
MÓDULO ADMINISTRADOR (EGONET AUTHOR)	67
<i>EGO</i>	68
<i>ALTER PROMPT</i>	70
<i>ALTER</i>	71
<i>ALTER PAIR</i>	72
MÓDULO CLIENTE (EGONET CLIENT)	74
LA ENTREVISTA CUALITATIVA	80
<i>VISUALIZACIÓN</i>	82
BIBLIOGRAFÍA	86
LA COMPARACIÓN CUANTITATIVA DE GRUPOS CON SPSS	87
A. ESTADÍSTICA DESCRIPTIVA Y GRÁFICOS	88
B. LA PRUEBA DE CHI-CUADRADO	91
C. LA PRUEBA DE LA -T- PARA MUESTRAS INDEPENDIENTES	98
<i>A. ESTADÍSTICA DESCRIPTIVA</i>	102
<i>B. LA PRUEBA DE LA CHI-CUADRADO</i>	106
<i>C. LA PRUEBA DE LA -T- PARA MUESTRAS INDEPENDIENTES</i>	107
APÉNDICE. LA DESVIACIÓN TÍPICA / LA DESVIACIÓN ESTÁNDAR	111

INTRODUCCIÓN

¿Qué son las *redes sociales*? Es necesario aclarar en primer lugar que se trata de un concepto polisémico. Un primer sentido se refiere a las redes sociales como “networking” entre personas mediante plataformas virtuales. Este sentido ha desplazado por su reciente popularidad a otros con mucha más tradición en las ciencias sociales. Un segundo sentido se refiere a las redes sociales como metáforas sobre la complejidad del mundo social. Un tercer sentido se refiere a las redes como estrategia de “empoderamiento” de actores locales (redes de organizaciones normalmente) para la mejora de una situación o la puesta en marcha de una iniciativa. Finalmente, un cuarto sentido se refiere a las redes sociales como herramienta para el análisis de las relaciones empíricas entre personas, organizaciones, países u otras unidades de análisis. Este último sentido es el que nos interesa desarrollar aquí. Para nosotros las redes sociales son medidas (imperfectas, como todas las medidas) de un mundo social y cultural en movimiento. Si somos capaces de interpretar adecuadamente esas medidas podremos ver simultáneamente las huellas de la historia, de las instituciones culturales y de la agencia de los individuos impresas en ellas. Y es aquí donde pensamos que reside su mayor interés.

Es importante destacar que desde la perspectiva del Análisis de Redes, la sociedad no se entiende sólo como agregado de individuos a nivel “micro” o de estructuras a nivel “macro”, sino como una “red de redes”, con adaptaciones mutuas.

Esta “red de redes” tiene estructuras, patrones o regularidades en las formas de vinculación de esos conjuntos relacionales. A primera vista las redes aparecen como un complejo difícil de desentrañar. Sin embargo, detrás de esa maraña existen patrones de interacción, que influyen en el comportamiento de las personas u organizaciones implicadas, aunque no sean conscientes de esta influencia. Desde este enfoque, la sociedad no la conforman los individuos, sino el agregado de individuos y patrones de relaciones, simultáneamente.

Desde esta perspectiva, el concepto de “red social” implica dos elementos: a) nodos y b) vínculos. En términos generales, los nodos pueden representar diferentes tipos de unidades de análisis (personas, hogares, comunidades,

grupos étnicos, organizaciones, instituciones, entre otros). Por su parte, los vínculos pueden representar diferentes tipos de relaciones entre nodos: a) semejanza, b) relaciones sociales, c) interacciones y d) flujos¹.

El Análisis de Redes estudia pues los vínculos entre personas (u otra unidad de análisis) y la influencia que ejercen estos vínculos en el comportamiento o la cognición. Sus antecedentes intelectuales se pueden encontrar en la sociología de Simmel, la sociometría de Moreno y la antropología social de la Escuela de Manchester.

La expresión “red social” fue acuñada por el antropólogo John Barnes (1954) en su trabajo sobre una aldea de pescadores noruegos. Igualmente es destacable el uso de esta expresión por Elisabeth Bott (1955) en su trabajo con familias de clase media inglesas. Durante los años 60 los estudios de redes sociales fueron populares en la antropología británica fue enorme, como se evidencia en el trabajo de Clyde Mitchell (1969), que recoge y sistematiza diversos estudios de antropología urbana en África. Estos primeros trabajos se realizaron con términos y conceptos desarrollados por los mismos autores, cuyo objetivo era capturar las transformaciones que la modernización y urbanización generaba en aldeas y comunidades tribales, inmersas en nuevos procesos de cambio cultural y diferenciación social.

Aunque los principios elementales de la aproximación ya estaban disponibles en esos años, las limitaciones técnicas fueron imposibles de superar. En efecto, si recoger datos sobre relaciones sociales es ya una tarea difícil, analizarlos sin ayuda de computadoras lo es más aún. Éste fue el caso de los primeros estudios de redes sociales, que se realizaron de manera artesanal, tan sólo con papel y lápiz, pero con mucha creatividad e imaginación. No cabe duda que el gran mérito de estos antropólogos fue poner en evidencia el complejo mundo de las redes sociales e iniciar su estudio sistemático, a pesar de las limitaciones técnicas.

¹ Los vínculos por semejanza son aquellos desarrollados entre nodos que tienen un mismo tipo de atributos (edad, sexo, estudios, etnia, clase, etc.). Los vínculos por relaciones sociales son aquellos que siguen roles específicos, como parentesco o amistad. Los vínculos por interacciones son aquellos mediados por pautas de comportamiento, como un saludo, una conversación o una invitación. Finalmente, los vínculos por flujos son aquellos donde hay intercambio de recursos, como dinero, información o favores.

Hoy en día el panorama es diferente. Existen programas para Análisis de Redes que facilitan el trabajo del investigador. A diferencia de hace medio siglo, el antropólogo/a puede realizar trabajo de campo con una portátil y recoger información sobre redes de manera directa, utilizando UCInet y EgoNet, entre muchos otros. Estos programas permiten recoger y analizar volúmenes complejos de información sobre redes y visualizar sus principales propiedades de estructura y composición.

En la actualidad, los escenarios de investigación antropológica “tradicionales” se vuelven más complejos. Categorías como “multiculturalismo” e “interculturalidad” describen la nueva heterogeneidad de estos escenarios locales y el incremento de sus conexiones no-locales. Hoy en día los “nativos” socializan entre matrices culturales “desterritorializadas”, más híbridas, creolizadas y translocales.

En un mundo cada vez más globalizado el “espacio etnográfico” no se puede delimitar sólo siguiendo criterios de geografía local (una isla, tribu o comunidad). Mas aún, cuando el “encaje” entre geografía, sociedad y cultura es menos nítido y las personas desarrollan sus vínculos en ámbitos locales y no-locales, cuya complejidad la etnografía clásica no ha podido recoger. En ese sentido, el Análisis de Redes se puede constituir en una herramienta estratégica para el desarrollo de nuevas etnografías multisituadas (*Multi-Sited Ethnography*), que permitan identificar y analizar las nuevas cartografías de sociabilidad y cultura de las personas.

Este manual es una introducción al uso del UCInet y Egonet. Ofrece una explicación de operaciones básicas para el estudiante. Los materiales de formación que se presentan aquí son en parte una reedición de materiales publicados con anterioridad en la colección *Quaderns* de la UAB (UCInet 6) y en parte escritos para esta edición (EgoNet) o publicados por primera vez después que hubiesen sido preparados para un curso de comparación transcultural de redes personales, impartido en el año 2006 por Chris McCarty y Miranda J. Lubbers en el programa de *Investigación etnográfica, teoría antropológica y comparación transcultural* en la UAB.

Esta nueva edición está financiada por el proyecto *C/017589/08 – Remesas y redes transnacionales: implicaciones para el desarrollo económico peruano* de la AECI y tiene el objetivo de servir de soporte a los talleres de formación sobre redes sociales incluidos en el proyecto.

Como en el caso de la edición anterior, este manual intenta llenar un espacio hasta ahora cubierto de forma incompleta: el de la autoformación en las diferentes aplicaciones del Análisis de redes sociales. Los textos teóricos disponibles en castellano empiezan a ser abundantes, sobre todo gracias a la Revista REDES (<http://revista-redes.rediris.es>). Sin embargo, los talleres y las matrices de datos, tanto reticulares como atributivos, están limitados prácticamente al web REDES (www.redes-sociales.net).

Cada capítulo tiene la voluntad de ser un taller autosuficiente. Es por ello que en ocasiones se pueden encontrar informaciones redundantes, pero siempre necesarias para la realización de la práctica propuesta.

Evidentemente este manual está pensado para Windows XP, excepto en el caso de EgoNet, escrito en Java y por tanto multiplataforma.

OPERACIONES BÁSICAS CON UCINET 6 ²

La experiencia nos enseña que las dificultades para el análisis de redes sociales residen, más que en cuestiones conceptuales o teóricas, en el aprendizaje de la mecánica de la transformación de datos. Este capítulo está dedicado a la realización de unas pocas prácticas iniciales que permitirán con toda seguridad moverse ágilmente por la variedad de posibilidades analíticas del programa.

CÓMO INTRODUCIR DATOS DE RELACIONES (RETICULARES)

Si bien a lo largo de este manual se encontrarán referencias a la manera de introducir datos y construir matrices, presentamos a continuación una breve síntesis de las dos formas de introducir datos que pueden utilizarse en Ucinet 6. La primera es directamente en una hoja de datos dentro del mismo programa; la segunda es mediante un fichero de texto que creamos en otra aplicación y que luego importamos. Veamos cada una de ellas.

HOJA DE DATOS

En la Ilustración 1 se presenta la pantalla inicial de Ucinet, en donde hemos resalto el botón que permite activar la matriz de datos de Ucinet 6.

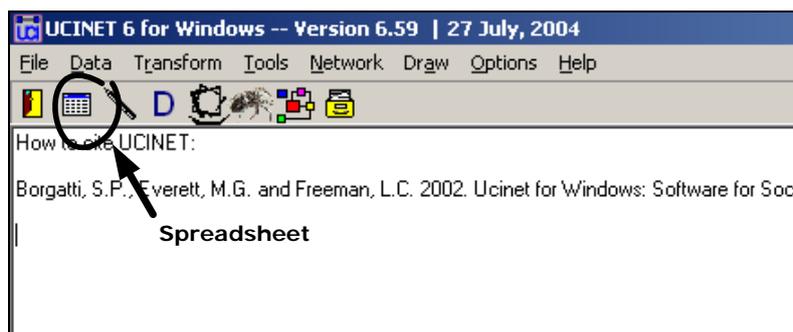


Ilustración 1. La hoja de datos (*spreadsheet*) de Ucinet 6

Esta matriz es útil para introducir pocos nodos de una forma rápida. Una vez clicado el botón indicado, se desplegará otra pantalla (ver ilustración 2). Haremos una pequeña práctica para introducir una red de solamente 5 nodos.

² Autor: José Luis Molina (Departament d'antropologia social i cultural, UAB). Ucinet6-Netdraw se puede descargar de <http://www.analytictech.com>.

La relación que mediremos es si las personas de la red han sido o no pareja sentimental estable. Dado que la relación “ser pareja de” es simétrica (si A es pareja de B, B es pareja de A) lo primero que haremos será seleccionar la opción de datos simétricos para introducir fácilmente las etiquetas de los nodos, antes de introducir los nombres de los integrantes de esta red. También indicaremos las dimensiones de la matriz: 5 filas y 5 columnas (ver ilustración 2)

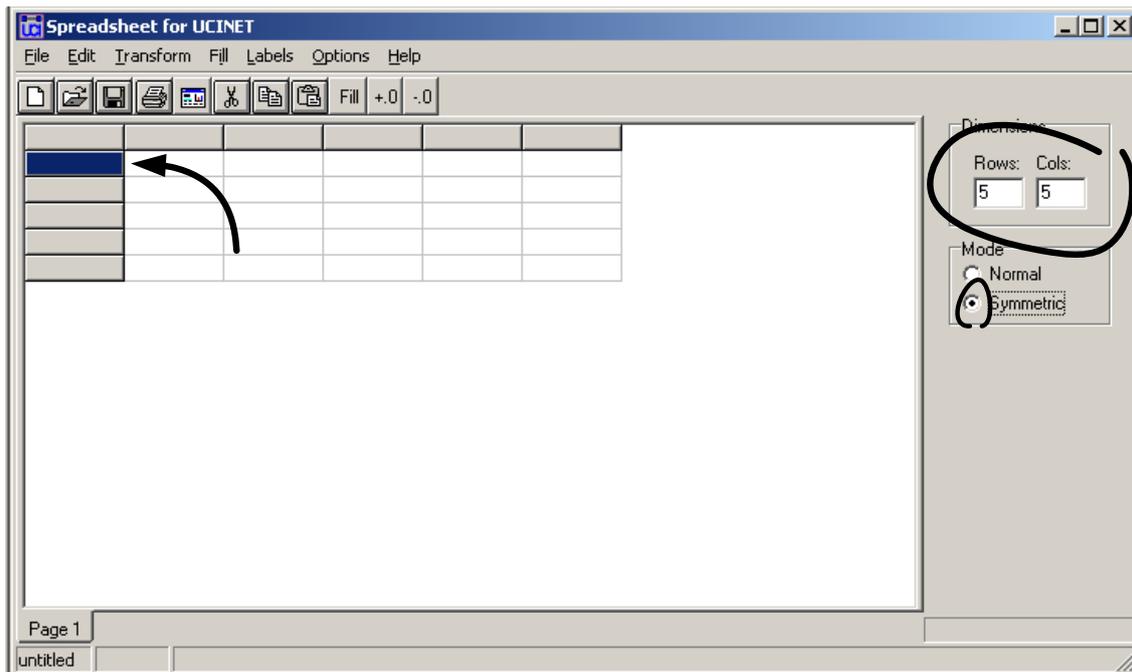


Ilustración 2. Hoja de datos de Ucinet 6.

A continuación clicaremos en la celda seleccionada y comenzaremos a escribir los nombres de los integrantes de la red: Penélope, Tom, Nikole, Melanie, Antonio. Una vez escritos copiaremos estas etiquetas de las filas a las columnas con el comando:

Labels>Copy rows to columns

A continuación indicaremos con un “1” si son o han sido (o han dicho que han sido) pareja estable. Como la matriz es simétrica sólo es necesario que indiquemos que Penélope tiene un “1” con Tom, Tom un “1” con Nikole y Melanie un “1” con Antonio. Con el comando

Fill>Missings with simmetric counterparts

rellenaremos los “1” que faltaban y con el comando

Fill>Blanks w/ 0s

rellenaremos con “0” las celdas en blanco. Por último, solamente nos queda guardar la matriz en un fichero:

File>Save As>prueba

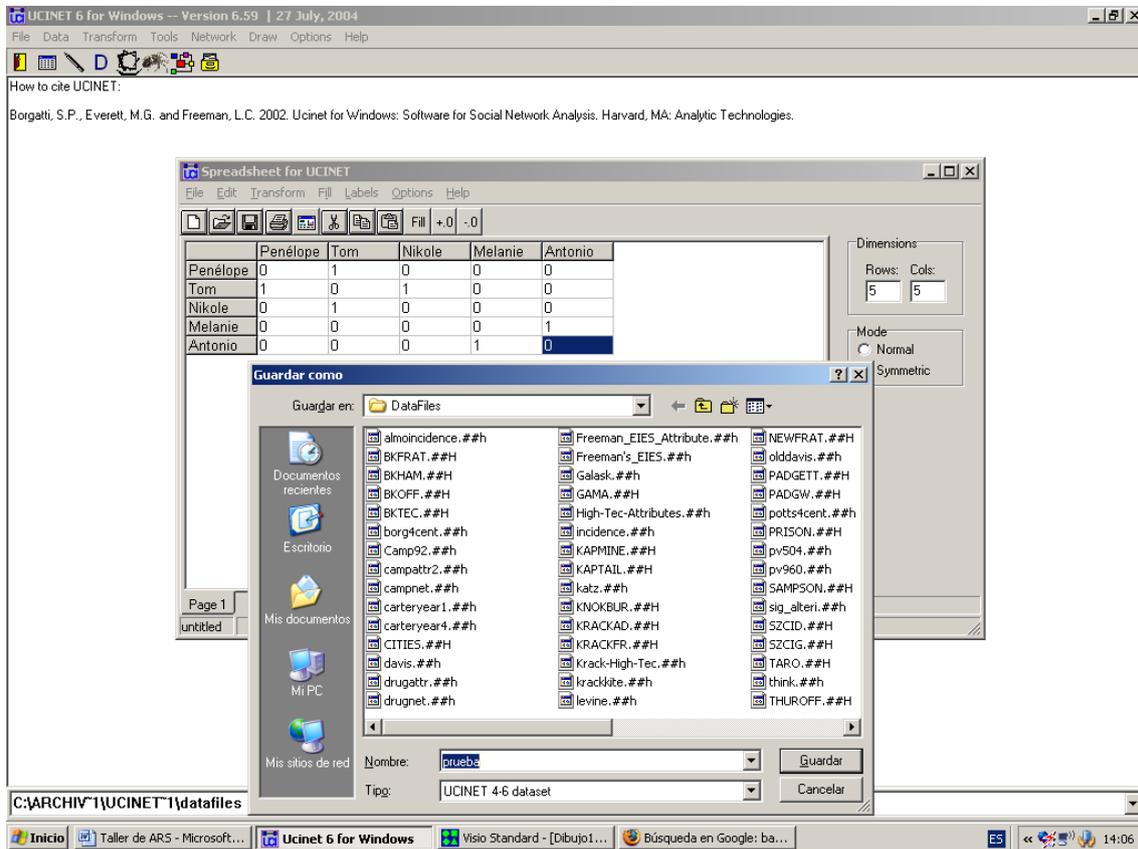


Ilustración 3. Guardar la matriz generada en la hoja de datos

Con esto ya tenemos generada una matriz de datos lista para ser analizada. El nombre asignado por Ucinet 6 será

prueba.##h

La hoja de datos de Ucinet 6 es compatible con la opción de copiar y pegar de Excel, por lo que en ocasiones puede ser útil trabajar en esa hoja de cálculo, abrir la hoja de datos de Ucinet y pegar la matriz seleccionada previamente. En este caso tenemos que tener cuidado que no queden filas ni columnas en blanco, indicando en el apartado correspondiente el número exacto de ambas.

FICHERO DE TEXTO

Cuando en el transcurso de una investigación se recogen datos relacionales, la lista de nodos suele ser de decenas o centenas. Por tanto, es mucho más práctico recoger los datos en otro formato, como por ejemplo

```
Penélope Tom 1
Tom Nikole 1
etc.
```

En lugar de introducir los datos directamente en la matriz, crearemos un fichero de texto, que luego importaremos a Ucinet 6.

Ucinet 6 dispone de muchos tipos de formato para los ficheros de texto, pero el más útil es el formato dl. Si quisiéramos incorporar los datos de la red que hemos realizado anteriormente en este formato deberíamos utilizar algún procesador de texto, y escribir lo siguiente:

```
dl n = 5 format = el1
labels embedded
data:
Penélope Tom 1
Tom Penélope 1
Tom Nikole 1
Nikole Tom 1
Melanie Antonio 1
Antonio Melanie 1
```

De esta forma se indica a Ucinet 6 cada una de las relaciones. Si copiamos estas instrucciones en un fichero de texto y lo guardamos como prueba2.txt podremos importarlo a Ucinet 6 de la siguiente forma:

```
Data>Import>DL>prueba2.txt
```

El resultado es el mismo que hemos obtenido con la hoja de datos. Ucinet crea un fichero llamado

```
Prueba2.##h
```

CÓMO CONVERTIR MATRICES DE MODO 2 EN MATRICES DE MODO 1

COMPLETAR FILAS O COLUMNAS

A menudo nos encontramos con matrices de modo 2 que tienen que ser convertidas en matrices de modo 1 para permitir algunos análisis. Las matrices de modo 2 resultan, por ejemplo, cuando los respondientes citan nombres libremente, sin escogerlos de una lista. Así tenemos en las filas los respondientes y en las columnas los nominados, que no suelen coincidir.³

Vamos a suponer que tenemos una encuesta por barrios a jóvenes barceloneses sobre dónde van a divertirse por la noche. Tendríamos una matriz del tipo los del barrio A van al barrio B. No todos los barrios reciben jóvenes, por lo que la lista de barrios de las filas no coincide con la lista de barrios de las columnas. También puede ser que se queden en el barrio, por lo que nos interesan las relaciones reflexivas.

Así tenemos un matriz de 10 barrios en las filas (los barrios de origen de los jóvenes) y de 24 barrios y otras poblaciones (los destinos) en las columnas. Los números indican el número de jóvenes que se desplazan de un barrio a otro.

³ En las matrices de modo 1, contamos con los mismos nodos en las columnas y en las filas, mientras que en las de modo 2 los actores de filas y columnas no son exactamente los mismos.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W
A	24	4	0	0	1	4	1	0	0	1	3	0	1	0	0	0	4	0	0	0	1	2	0	0
B	44	29	1	0	5	20	0	0	0	2	11	1	2	3	0	0	1	0	0	0	2	1	1	0
C	28	5	10	1	0	5	1	0	0	1	6	4	4	1	1	1	4	0	0	0	10	2	0	1
D	8	4	3	19	4	8	0	0	0	0	1	3	1	0	0	0	0	0	0	0	1	1	0	0
E	14	6	0	0	18	17	1	0	0	0	2	3	2	0	0	0	0	0	0	0	1	0	0	2
F	6	8	0	0	1	26	0	0	0	1	4	2	0	0	1	0	0	0	0	0	1	2	1	1
G	23	8	2	1	2	16	5	2	0	0	6	3	2	3	8	1	1	0	0	0	1	2	0	0
H	18	8	0	1	1	3	2	24	2	5	9	0	3	0	3	3	0	4	1	0	7	6	1	1
I	16	4	0	1	2	6	0	1	9	1	9	1	3	1	0	2	0	0	0	0	2	3	5	2
J	26	11	0	0	4	10	0	1	4	5	22	1	4	2	1	1	3	1	1	1	2	6	0	2

Tabla 1. Número de jóvenes que se trasladan para divertirse entre barrios de una ciudad y poblaciones anejas

El fichero de datos es el siguiente (página siguiente):

DL

NR=10, NC=24

FORMAT = FULLMATRIX DIAGONAL PRESENT

ROW LABELS:

"A"
"B"
"C"
"D"
"E"
"F"
"G"
"H"
"I"
"J"

ROW LABELS EMBEDDED

COLUMN LABELS:

"A"
"B"
"C"
"D"
"E"
"F"
"G"
"H"
"I"
"J"
"K"
"L"
"M"
"N"
"Ñ"
"O"
"P"
"Q"
"R"
"S"
"T"
"U"
"V"
"W"

COLUMN LABELS EMBEDDED

LEVEL LABELS:

"Page 1"

DATA:

	"A"	"B"	"C"	"D"	"E"	"F"	"G"	"H"	"I"	"J"	"K"	"L"	"M"	"N"	"Ñ"	"O"	"P"	"Q"	"R"	"S"	"T"			
"U"																								
"V"																								
"W"																								
"A"	24	4	0	0	1	4	1	0	0	1	3	0	1	0	0	0	4	0	0	0	1	2	0	0
"B"	44	29	1	0	5	20	0	0	0	2	11	1	2	3	0	0	1	0	0	0	2	1	1	0
"C"	28	5	10	1	0	5	1	0	0	1	6	4	4	1	1	1	4	0	0	0	10	2	0	1
"D"	8	4	3	19	4	8	0	0	0	0	1	3	1	0	0	0	0	0	0	0	1	1	0	0
"E"	14	6	0	0	18	17	1	0	0	0	2	3	2	0	0	0	0	0	0	0	1	0	0	2
"F"	6	8	0	0	1	26	0	0	0	1	4	2	0	0	1	0	0	0	0	0	1	2	1	1
"G"	23	8	2	1	2	16	5	2	0	0	6	3	2	3	8	1	1	0	0	0	1	2	0	0
"H"	18	8	0	1	1	3	2	24	2	5	9	0	3	0	3	3	0	4	1	0	7	6	1	1
"I"	16	4	0	1	2	6	0	1	9	1	9	1	3	1	0	2	0	0	0	0	2	3	5	2
"J"	26	11	0	0	4	10	0	1	4	5	22	1	4	2	1	1	3	1	1	1	2	6	0	2

Seleccionemos esos datos y copiémoslos en un fichero de texto con nombre “jovenes.txt”. Desde Ucinet 6 realizaremos la importación:

Data>Import>DL

Hay que tener cuidado que el nombre del fichero aparezca como el producto de la exportación en el subdirectorio adecuado, sino no podremos localizarlo posteriormente.

Abriremos la hoja de datos y cargaremos el fichero importado. A continuación cambiaremos el número de filas de la matriz a 24 y copiaremos las etiquetas de las columnas en las filas. Por último, rellenaremos con ceros las celdas en blanco y guardaremos una matriz de modo 1 con un nombre nuevo. El resultado debería ser el siguiente:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	24	4	0	0	1	4	1	0	0	1	3	0	1	0
B	44	29	1	0	5	20	0	0	0	2	11	1	2	3
C	28	5	10	1	0	5	1	0	0	1	6	4	4	1
D	8	4	3	19	4	8	0	0	0	0	1	3	1	0
E	14	6	0	0	18	17	1	0	0	0	2	3	2	0
F	6	8	0	0	1	26	0	0	0	1	4	2	0	0
G	23	8	2	1	2	16	5	2	0	0	6	3	2	3
H	18	8	0	1	1	3	2	24	2	5	9	0	3	0
I	16	4	0	1	2	6	0	1	9	1	9	1	3	1
J	26	11	0	0	4	10	0	1	4	5	22	1	4	2
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Ilustración 4. Matriz de modo 1 de 24 x 24 barrios o localidades

En este momento podemos hacer un primer análisis, por ejemplo con Netdraw y obtener una imagen de los desplazamientos de los jóvenes.

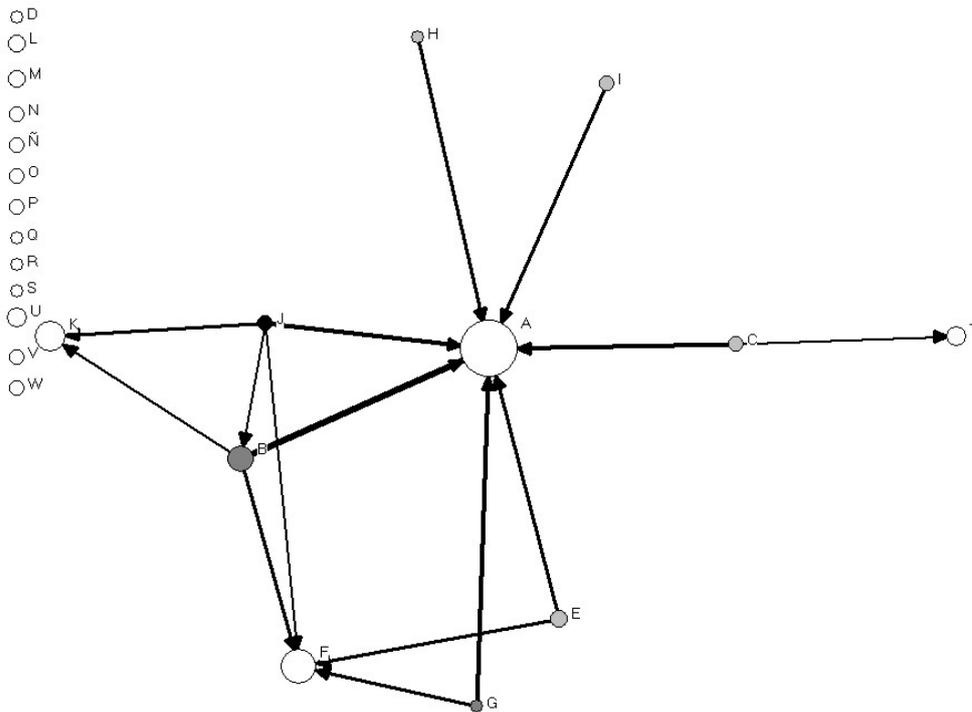


Ilustración 5. Desplazamientos nocturnos de jóvenes.

La escala de grises indica la proporción de jóvenes que salen de un barrio. El tamaño del nodo indica la proporción de jóvenes que visitan un barrio y el grosor de la línea la proporción de desplazamientos. Los nodos aislados indican que tienen muy pocos jóvenes con esta conducta. Parece claro que el barrio A recibe la mayor parte de las visitas nocturnas...

Para usar las potencialidades de Netdraw consultar el capítulo correspondiente.

AFILIACIONES

En ocasiones nos interesa conservar la matriz de modo 2 y hacer un análisis de filas por filas o de columnas por columnas. El caso típico consiste en el análisis hecho por Davis *et. al.* (1941) de un grupo de mujeres y los eventos sociales a los que asistieron. La matriz de modo 2 del ejemplo es la siguiente:

DL
 NR=18, NC=14
 FORMAT = FULLMATRIX DIAGONAL PRESENT
 ROW LABELS:

"EVELYN"
 "LAURA"
 "THERESA"
 "BRENDA"
 "CHARLOTTE"
 "FRANCES"
 "ELEANOR"
 "PEARL"
 "RUTH"
 "VERNE"
 "MYRNA"
 "KATHERINE"
 "SYLVIA"
 "NORA"
 "HELEN"
 "DOROTHY"
 "OLIVIA"
 "FLORA"

ROW LABELS EMBEDDED

COLUMN LABELS:

"E1"
 "E2"
 "E3"
 "E4"
 "E5"
 "E6"
 "E7"
 "E8"
 "E9"
 "E10"
 "E11"
 "E12"
 "E13"
 "E14"

COLUMN LABELS EMBEDDED

DATA:

	"E1"	"E2"	"E3"	"E4"	"E5"	"E6"	"E7"	"E8"	"E9"	"E10"	"E11"	"E12"	"E13"	"E14"
"EVELYN"	1	1	1	1	1	0	1	1	0	0	0	0	0	0
"LAURA"	1	1	1	0	1	1	1	1	0	0	0	0	0	0
"THERESA"	0	1	1	1	1	1	1	1	0	0	0	0	0	0
"BRENDA"	1	0	1	1	1	1	1	1	0	0	0	0	0	0
"CHARLOTTE"	0	0	1	1	1	0	1	0	0	0	0	0	0	0
"FRANCES"	0	0	1	0	1	1	0	1	0	0	0	0	0	0
"ELEANOR"	0	0	0	0	1	1	1	1	0	0	0	0	0	0
"PEARL"	0	0	0	0	0	1	0	1	1	0	0	0	0	0
"RUTH"	0	0	0	0	1	0	1	1	1	0	0	0	0	0
"VERNE"	0	0	0	0	0	0	1	1	1	0	0	1	0	0
"MYRNA"	0	0	0	0	0	0	0	1	1	1	0	1	0	0
"KATHERINE"	0	0	0	0	0	0	0	1	1	1	0	1	1	1
"SYLVIA"	0	0	0	0	0	0	1	1	1	1	0	1	1	1
"NORA"	0	0	0	0	0	1	1	0	1	1	1	1	1	1
"HELEN"	0	0	0	0	0	0	1	1	0	1	1	1	0	0
"DOROTHY"	0	0	0	0	0	0	0	1	1	0	0	0	0	0
"OLIVIA"	0	0	0	0	0	0	0	0	1	0	1	0	0	0
"FLORA"	0	0	0	0	0	0	0	1	0	1	0	0	0	0

De nuevo seleccionamos los datos anteriores y los copiamos en un fichero de texto. Este fichero lo importamos desde Ucinet, con el procedimiento que ya hemos mencionado. Esta matriz de 18 mujeres en las filas y 14 eventos sociales en las columnas puede ser analizada de forma que obtengamos una matriz de mujeres por mujeres (18 x 18) y una matriz de eventos por eventos (14 x 14). En cada caso el análisis nos permitirá identificar grupos de mujeres o de eventos, mujeres o eventos centrales o periféricos, en fin, obtener medidas estructurales sobre las relaciones.

Para ello seleccionamos en Ucinet 6

Data>Affiliations

En la ventana de diálogo que aparece a continuación (Ilustración 2), seleccionamos el fichero Ucinet importado y elegimos si queremos obtener la nueva matriz a partir de las filas ("Row") o partir de las columnas ("Column", en el desplegable).

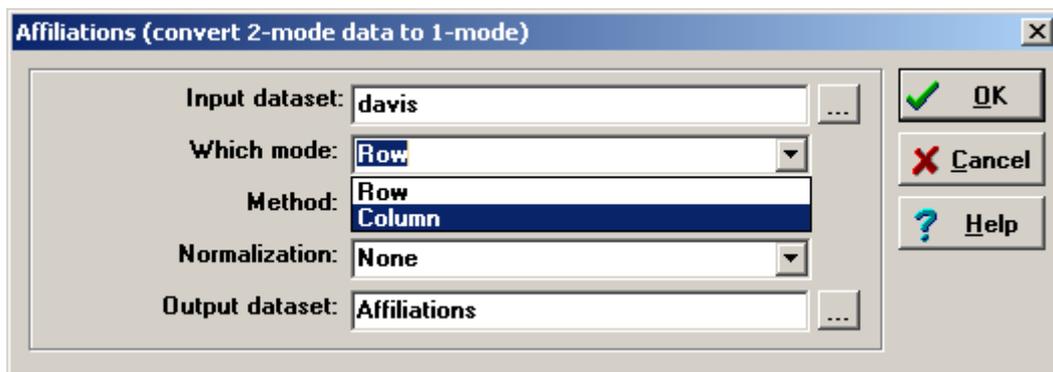


Ilustración 6. Ventana de diálogo de comando "Afilaciones"

Si elegimos "Row" obtendremos un fichero llamado Affiliations.##h de 18 x 18 listo con la siguiente matriz:

		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
	E	L	T	B	C	F	E	P	R	V	M	K	S	N	H	D	O	F	
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	EVELYN	8	6	7	6	3	4	3	3	3	2	2	2	2	2	1	2	1	1
2	LAURA	6	7	6	6	3	4	4	2	3	2	1	1	2	2	2	1	0	0
3	THERESA	7	6	8	6	4	4	4	3	4	3	2	2	3	3	2	2	1	1
4	BRENDA	6	6	6	7	4	4	4	2	3	2	1	1	2	2	2	1	0	0
5	CHARLOTTE	3	3	4	4	4	2	2	0	2	1	0	0	1	1	1	0	0	0
6	FRANCES	4	4	4	4	2	4	3	2	2	1	1	1	1	1	1	1	0	0
7	ELEANOR	3	4	4	4	2	3	4	2	3	2	1	1	2	2	2	1	0	0
8	PEARL	3	2	3	2	0	2	2	3	2	2	2	2	2	2	1	2	1	1
9	RUTH	3	3	4	3	2	2	3	2	4	3	2	2	3	2	2	2	1	1
10	VERNE	2	2	3	2	1	1	2	2	3	4	3	3	4	3	3	2	1	1
11	MYRNA	2	1	2	1	0	1	1	2	2	3	4	4	4	3	3	2	1	1
12	KATHERINE	2	1	2	1	0	1	1	2	2	3	4	6	6	5	3	2	1	1
13	SYLVIA	2	2	3	2	1	1	2	2	3	4	4	6	7	6	4	2	1	1
14	NORA	2	2	3	2	1	1	2	2	2	3	3	5	6	8	4	1	2	2
15	HELEN	1	2	2	2	1	1	2	1	2	3	3	3	4	4	5	1	1	1
16	DOROTHY	2	1	2	1	0	1	1	2	2	2	2	2	2	1	1	2	1	1
17	OLIVIA	1	0	1	0	0	0	0	1	1	1	1	1	1	2	1	1	2	2
18	FLORA	1	0	1	0	0	0	0	1	1	1	1	1	1	2	1	1	2	2

Un simple análisis de f-Groups (una técnica que busca grupos internamente densos y débilmente conectados que está disponible en Network>Subgroups>f-Groups) de esta matriz nos muestra la existencia de dos grupos de mujeres bien definidos:

Group 1:

EVELYN LAURA THERESA BRENDA CHARLOTTE FRANCES ELEANOR RUTH

Group 3:

VERNE MYRNA KATHERINE SYLVIA NORA HELEN

Levels of Ties Among Actors

		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	
		E	L	T	B	S	F	E	P	R	V	M	K	S	N	H	D	O	F	
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
1	EVELYN	2	2	2	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	
2	LAURA	2	2	2	2	1	2	2	1	1	1	1	1	1	1	1	1	1	0	0
3	THERESA	2	2	2	2	2	2	2	1	2	1	1	1	1	1	1	1	1	1	1
4	BRENDA	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	0	0
5	CHARLOTTE	1	1	2	2	2	1	1	0	1	1	0	0	1	1	1	0	0	0	0
6	FRANCES	2	2	2	2	1	2	1	1	1	1	1	1	1	1	1	1	0	0	0
7	ELEANOR	1	2	2	2	1	1	2	1	1	1	1	1	1	1	1	1	0	0	0
8	PEARL	1	1	1	1	0	1	1	2	1	1	1	1	1	1	1	1	1	1	1
9	RUTH	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	VERNE	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1
11	MYRNA	1	1	1	1	0	1	1	1	1	1	2	2	2	1	1	1	1	1	1
12	KATHERINE	1	1	1	1	0	1	1	1	1	1	2	2	2	2	1	1	1	1	1
13	SYLVIA	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1
14	NORA	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	1	1	1	1
15	HELEN	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1	1	1	1	1
16	DOROTHY	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	2	1	1	1
17	OLIVIA	1	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1	2	1	1
18	FLORA	1	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	2

Ilustración 7. Análisis de F-Groups de Ucinet 6

Por otra parte, un análisis de las columnas también puede ser de utilidad. Procederíamos de la misma forma pero seleccionando “Column” en lugar de “Row”.

CÓMO RECODIFICAR MATRICES

Ucinet 6 dispone de un amplio conjunto de procedimientos para manipular matrices. De todos ellos destacaremos la opción

Transform>Recode

Este procedimiento consta de dos partes (pestaña “Files” y pestaña “Recode”). En la pestaña **Files**, podemos seleccionar el fichero con los datos, las filas y las columnas que queremos recodificar. El botón **L** nos permite seleccionar diferentes subconjuntos de filas o columnas, presionando la tecla Ctrl y clicando las etiquetas de fila o columna correspondientes. Si no especificamos nada, la remodificación afectará a toda la matriz.

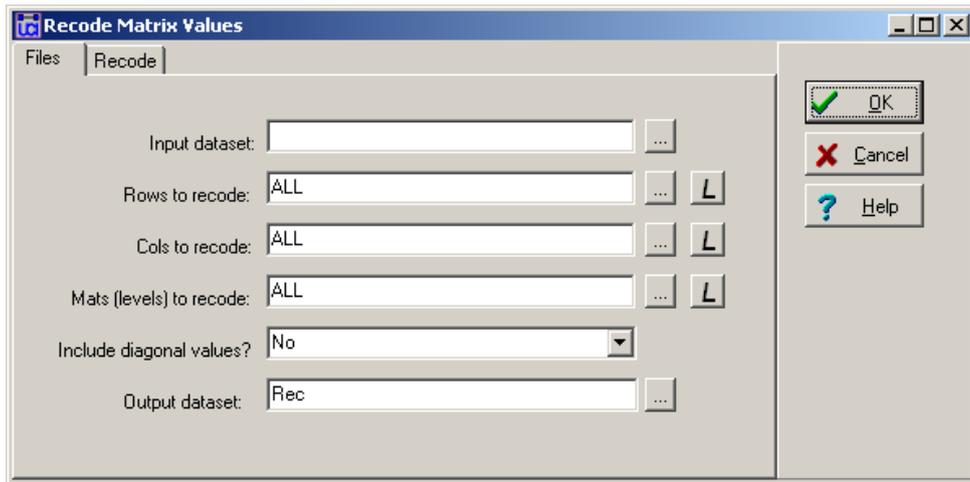


Ilustración 8. Opción "Recode" de Ucinet 6

Por último, en la pestaña **Recode** podemos especificar un conjunto de reglas de recodificación. Siguiendo con el ejemplo de los jóvenes que se mueven en la noche barcelonesa, podemos introducir el criterio de convertir en “0” todos los desplazamientos de menos de 15 jóvenes y en “1” todos los desplazamientos de 15 o superior. Así:

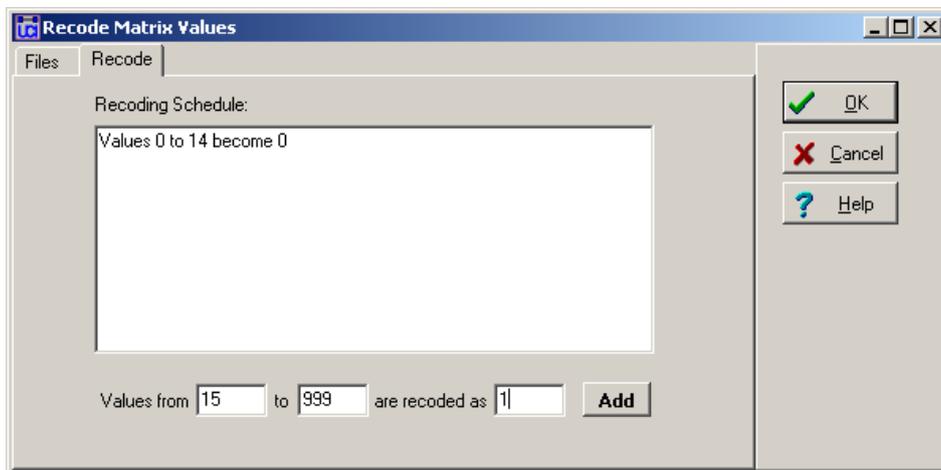


Ilustración 9. Recodificación de los valores de una matriz

Si clicamos la tecla **Add** dispondremos de las dos reglas definidas. El resultado de la recodificación (matriz Recjovenes) es el siguiente:

		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	
		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	A	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	B	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	C	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	D	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	E	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	F	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	G	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	H	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	I	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	J	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Si representamos gráficamente la matriz con NetDraw obtendremos una imagen bastante parecida a la que obtuvimos en la Ilustración 6.

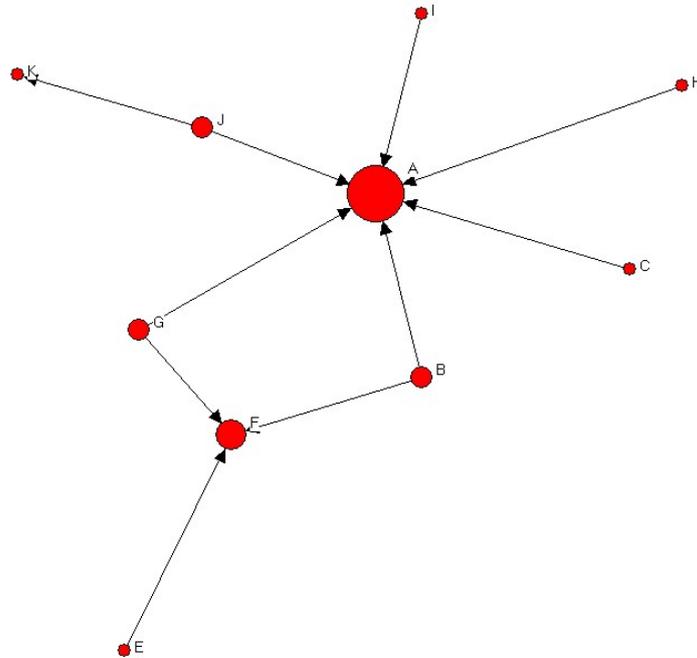


Ilustración 10. Grafo de la matriz recodificada de traslado nocturno de jóvenes en Barcelona

EL ANÁLISIS DE REDES SOCIOCÉNTRICAS CON UCINET 6 Y NETDRAW⁴

INTRODUCCIÓN

Este taller tiene como objetivo introducir conceptos básicos del análisis de datos reticulares (o análisis de redes sociales), utilizando UCINET 6 y NetDraw⁵. Los contenidos de este manual surgen de tres fuentes principales:

- Hanneman, Robert A., *Introducción a los métodos de análisis de redes sociales*. Departamento de Sociología de la Universidad de California Riverside. <http://wizard.ucr.edu/~rhannema/netwprks/text/textindex.html>
Versión en castellano en <http://www.redes-sociales.net/materiales>.
- Molina, José Luis (2001), *El análisis de redes sociales. Una introducción*. Barcelona: Edicions Bellaterra.
- Wasserman, Stanley and Faust, Katherine (1994), *Social network Analysis. Methods and Applications*. Cambridge MA: Cambridge University Press.

Pero ¿qué es el análisis de redes sociales? Veamos una definición:

El análisis de redes sociales se ocupa del estudio de las relaciones entre una serie definida de elementos (personas, grupos, organizaciones, países e incluso acontecimientos). A diferencia de los análisis tradicionales que explican, por ejemplo, la conducta en función de la clase social y la profesión, el análisis de redes sociales incorpora las relaciones y no sólo los atributos de los elementos. (Molina, *op.cit.* p. 13).

La particularidad del análisis de redes sociales radica entonces en el énfasis en las relaciones entre los elementos estudiados, entre sus propiedades relacionales, y no exclusivamente en las características monádicas (atributos individuales) de cada elemento.

⁴ Autora: Águeda Quiroga (UPF).

⁵ Borgatti, S.P. 2002. *NetDraw: Graph Visualization Software*. Harvard: Analytic Technologies <http://www.analytictech.com>.

Como señala Hanneman (*op.cit*), los datos sociológicos “tradicionales” se representan en una matriz rectangular en donde se inscriben las mediciones. En las filas se ubican los casos u observaciones (personas, organizaciones, países, etc.), y en las columnas, las variables seleccionadas (edad, tipo de organización, población, etc.). En cambio, en el análisis de redes sociales

(...) los datos de la ‘red’ (en su forma más pura), constituyen una matriz cuadrada de mediciones. Las filas de la matriz son los casos, sujetos u observaciones. Las columnas son el mismo conjunto de casos, sujetos y observaciones –allí está la diferencia clave con los datos convencionales. En cada celda de la matriz se describe una relación entre los actores. (p. 5)

Informante	Edad	Sexo	Ocupación
FFG	32	Mujer	Secretaria administrativa
BNM	20	Hombre	Estudiante
DFP	45	Hombre	Funcionario público
CCD	37	Mujer	Abogada
STR	26	Mujer	Docente

Tabla 2. Matriz con datos sociológicos “convencionales”

¿Quién es amigo de quién?					
	FFG	BNM	DFP	CCD	STR
FFG	-	1	0	0	1
BNM	1	-	1	0	0
DFP	0	1	-	1	1
CCD	0	0	1	-	0
STR	1	0	1	0	-

Tabla 3. Matriz con datos relacionales

En las Tablas 2 y 3 se presenta un ejemplo de los dos tipos de matrices señaladas por Hanneman. La primera se fija en las características de los individuos. La segunda, en las relaciones que mantienen dichos individuos entre sí. Así, se ha marcado con 1 la presencia de la relación de amistad y con 0 su ausencia.

Los elementos básicos del análisis de redes sociales son, entonces, los nodos (los puntos que representan a personas, grupos, países, etc.) y la relación o vínculo que nos interesa analizar (*amistad, enemistad, parentesco, comercio, etc.*) y que se establece entre tales nodos.

DESCRIPCIÓN DEL EJERCICIO PRÁCTICO

Comenzaremos trabajando con un ejemplo sencillo, a fin de introducir los conceptos básicos. Analizaremos las relaciones de conocimiento mutuo de un curso de estudiantes universitarios. Contamos con el listado de estudiantes (nombre y sexo) y la carrera a la que pertenecen. Además, cada uno de ellos ha respondido un cuestionario en el que se les preguntaba a quiénes de los demás integrantes de la clase conocían personalmente antes de iniciar el curso.

Así, los nodos serán en este caso los estudiantes, y el vínculo a analizar será la relación de conocimiento previo. Conviene precisar bien a qué propiedad relacional nos referimos. No se trataba en este caso de preguntar quiénes eran amigos de quiénes, ni de si algún estudiante tenía conocimiento indirecto de otro (ambos casos resultan problemáticos por diversas razones). Se trata de la propiedad “conocer a”, que en nuestro estudio (y en el cuestionario) se entendía como tener un mutuo conocimiento personal (que no tiene por qué implicar relación continuada, pero sí alguna interacción pasada). Con los datos de las tablas 4 y 5 elaboraremos las matrices que utilizaremos para el análisis

Nombre	Sexo (M/F)	Carrera
Andrés	M	Antropología
Carlos	M	Arqueología
Carme	F	Arqueología
Carmen	F	Antropología
Carol	F	Arqueología
Dolors	F	Antropología
Hugo	M	Sociología
Joan	M	Antropología
José	M	Psicología Social
Julio	M	Antropología
Liliana	F	Sociología
María	F	Psicología Social
Mariano	M	Arqueología
Martha	F	Arqueología
Nuria	F	Psicología Social
Pau	M	Antropología
Xavi	M	Psicología Social

Tabla 4. Listado de estudiantes

Nombre	Conoce a
Andrés	Carlos, Carmen, Dolors, Joan, Julio, Pau

Carlos	Carme, Carol, Andrés
Carme	Carlos, Carol
Carmen	Andrés, Dolors, Pau
Carol	Carlos, Carme
Dolors	Andrés, Carmen, José
Hugo	Liliana
Joan	Andrés, Julio, Pau, José
José	Dolors, Joan, María, Nuria, Xavi
Julio	Andrés, Joan
Liliana	Hugo
María	José, Nuria, Xavi
Mariano	Martha, Pau
Martha	Mariano, Pau
Nuria	José, María
Pau	Andrés, Carmen, Joan, Martha, Mariano
Xavi	José, María

Tabla 5. Resultado del cuestionario

CREAR UNA MATRIZ CON UCINET 6

El primer paso en nuestro análisis de redes del ejemplo planteado consiste en construir una matriz con la información de que disponemos, a fin de poder analizarla y graficarla. Para eso, utilizaremos el programa Ucinet 6.

Al iniciar dicho programa Ucinet 6 nos encontramos con la siguiente pantalla:

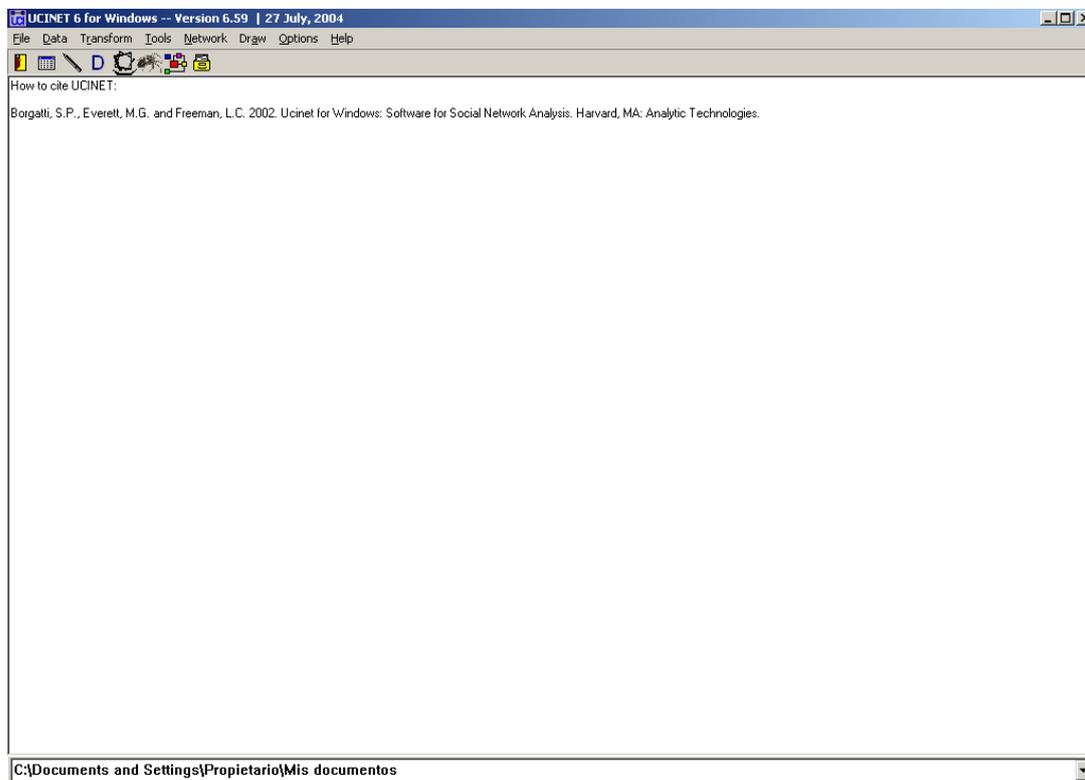


Ilustración 11. Ucinet 6

Al operar en un entorno Windows, el programa presenta características comunes a los programas elaborados para este sistema operativo. Cuenta con una barra de menús desplegables (*File, Data, Transform, Tools, Network, Draw, Options, Help*) más una serie de iconos de acceso directo. Si nos fijamos en la figura 13, veremos que en la parte inferior hay un rectángulo blanco, donde se lee “C:\Documents and Settings\Propietario\Mis documentos”. Esta dirección nos indica el directorio en el que estamos trabajando, y en el que por defecto guardará los archivos. Para definir otra carpeta o directorio se debe hacer seleccionar

File> Change Default Folder

Esta operación es importante hacerla en primer lugar.

A continuación debemos introducir nuestros datos. Para ello activaremos la función “Spreadsheet” (el segundo icono desde la izquierda, ver la Ilustración 2 y siguientes).

Antes de continuar con la introducción de datos en la matriz es conveniente realizar una aclaración sobre los tipos de vínculos. Los lazos o vínculos pueden ser orientados o no orientados (por ejemplo, *conocimiento previo* como una relación no orientada, *prestar dinero a alguien* como una relación orientada). Pero, por otra parte, también pueden ser o binarios o valorados. Son binarios cuando medimos la ausencia o presencia de la relación. En el caso que estamos analizando, si dos personas se conocen previamente le damos el valor 1 y si no, 0. Pero también podríamos querer valorar el tipo de relación que tienen dos personas, ponderando el vínculo, por ejemplo, entre 0 si no hay conocimiento, 1 si son apenas conocidos, 2 si son conocidos con una relación continuada, 3 si son amigos y 4 si son amigos muy cercanos. A lo largo de este ejemplo sólo trabajaremos con vínculos no orientados y binarios, pero es importante recordar que se pueden utilizar también vínculos orientados y

vínculos valorados, y que dentro de estos existe una gran pluralidad de variantes⁶.

Dado que, como hemos dicho, en nuestro ejemplo los vínculos no son orientados, cambiaremos la función a “Symmetric” (Ilustración 3) e introduciremos los nombres de los estudiantes en las filas y luego los copiaremos en las columnas

Labels> Copy rows to columns

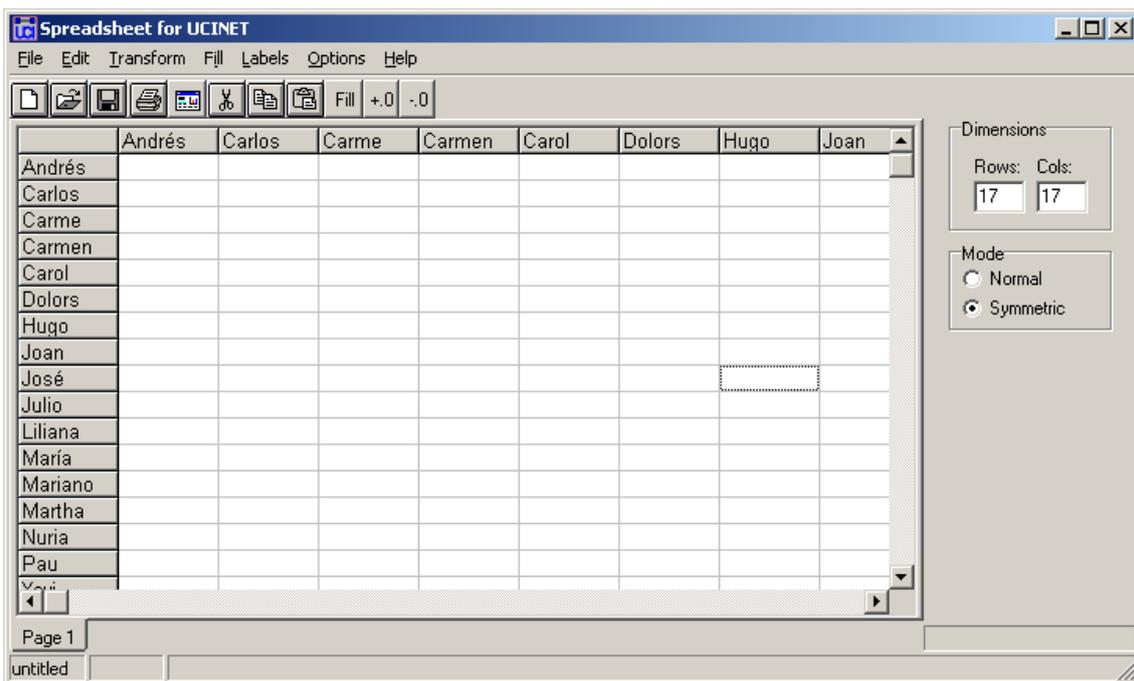


Ilustración 12. Hoja de datos con los estudiantes

Hemos introducido ya los nodos de la red que queremos analizar. A continuación introduciremos los valores de la relación. Así, siguiendo la información de la tabla 5, introduciremos manualmente la información correspondiente a cada par de personas. Introduciremos un 1 si hay relación y dejaremos la celda en blanco si no la hay:

⁶ Para más información sobre los tipos de vínculos y las escalas de medida, véase Hanneman, *op.cit*, cap 1:17-22.

	Dolores	Hugo	Joan	José	Julio	Lilianna	María	Mariano	Martha	Nuria	Pau	Xavi
Andrés	1											
Carlos		1										
Carme												
Carmen											1	
Carol												
Dolores				1								
Hugo						1						
Joan					1						1	
José			1				1					
Julio				1						1		1
Lilianna												
María										1		1
Mariano									1		1	
Martha								1			1	
Nuria							1					
Pau												1
Xavi												

Ilustración 13. Relación existente entre los estudiantes (conocimiento previo)

Una vez completa nos ubicaremos en la última celda (en este caso, en la intersección de la columna 17 con la fila 17) y desde allí haremos

Fill> Blanks w/0s

De esta manera, se completarán con ceros el resto de las celdas de la matriz:

Hemos completado así la matriz con las relaciones. La guardaremos con el nombre “estudiantes” (a partir del menú File> Save as).

A continuación crearemos dos nuevas matrices, en las que almacenaremos los datos atributivos de la Tabla 4 (*sexo* y *carrera* a la que pertenecen los estudiantes). Recordemos que los datos atributivos son propiedades individuales, y por lo tanto no relacionales, de los nodos que son objeto de nuestro análisis. Para ellos, abriremos la matriz que acabamos de crear (File>Open>estudiantes.###) y realizaremos nuevamente la operación “Save as”. Llamaremos “carrera” a este nuevo fichero (File>Save as>carrera).

Introduciremos en esta matriz (carrera.###) los datos correspondientes a la carrera universitaria que estudian los alumnos de la clase. Para aprovechar el listado de los nombres, eliminaremos la información de las columnas. Seleccionamos con el ratón las 17 columnas y apretamos la tecla “Suprimir” o “Delete”:

Una vez borrada la información de las columnas, cambiaremos el modo de la matriz a “Normal” (esta matriz ya no será cuadrada como la anterior, sino que tendrá 17 filas y 1 columna). Escribimos en el encabezado de la columna 1 (en la zona coloreada debajo del número 1) “carrera” (sin las comillas). Introduciremos en esa columna la información codificada de la siguiente manera: 1 para los estudiantes de antropología, 2 para los de arqueología, 3 para los de sociología y 4 para los de Psicología social:

	carrera
Andrés	1
Carlos	2
Carme	2
Carmen	1
Carol	2
Dolors	1
Hugo	3
Joan	1
José	4
Julio	1
Liliana	3
María	4
Mariano	2
Martha	2
Nuria	4
Pau	1
Xavi	4

Ilustración 14. Datos atributivos de los estudiantes (carrera)

Guardamos la matriz y elegimos “yes” en la opción guardar cambios.

A continuación, abrimos la matriz que acabamos de crear (File>Open>carrera.###h) y la guardamos con el nombre “sexo” (File>Save as>sexo). Realizamos sobre esta matriz la misma operación que en el caso anterior (borramos la columna 1) y escribimos como encabezado de la misma “sexo” (sin comillas) e introducimos la información codificada de la siguiente manera: 1 para varones y 2 para mujeres.

Cerramos la matriz y guardamos los cambios. Hemos introducido ya toda la información que utilizaremos en este taller.

GRAFICAR UNA RED CON NETDRAW

A continuación, veremos la representación gráfica de la primera matriz que hemos creado, la que resume la información sobre las relaciones existentes entre los estudiantes.

Una vez cerrada la función “Spreadsheet”, regresamos a la pantalla inicial de Ucinet 6. Allí hacemos un click sobre el icono que abre el programa NetDraw (el séptimo icono contando desde la izquierda). Se abre así una ventana con el programa NetDraw. Este programa se utiliza para visualizar gráficamente redes sociales. Veremos representada en grafos la matriz que hemos introducido. Para ello, abriremos la matriz “estudiantes.###” que hemos creado anteriormente desde:

File>Open>Ucinet dataset>Network

Al abrirla obtenemos el siguiente gráfico:

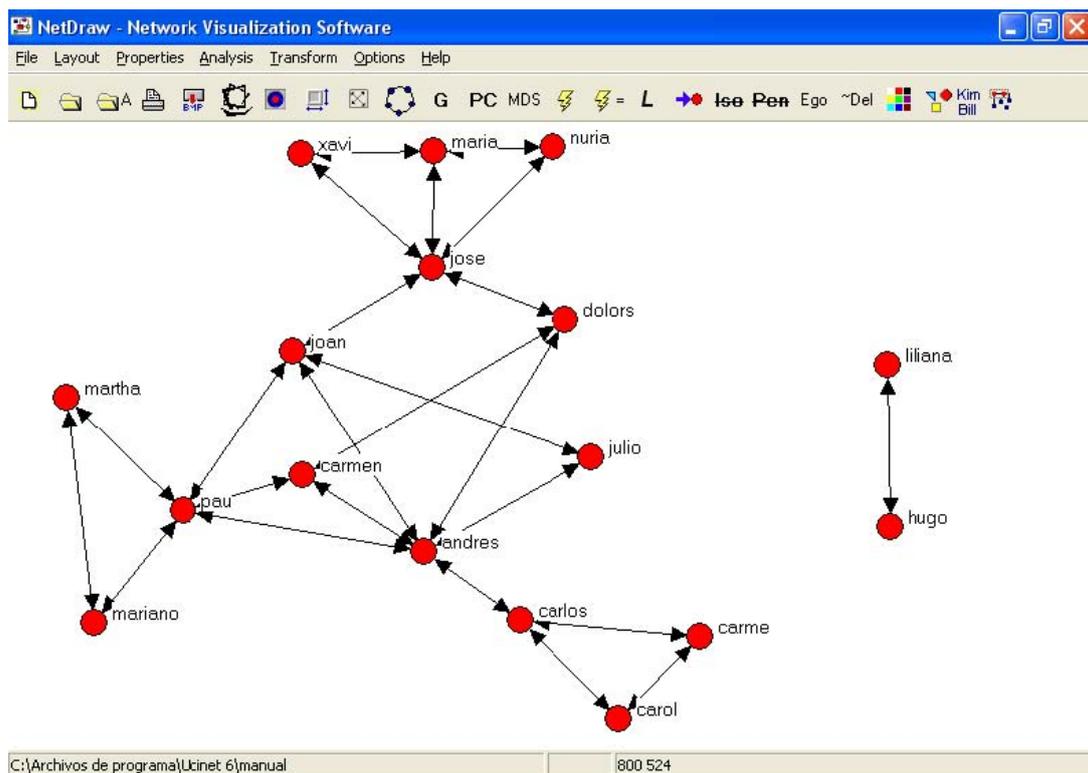


Ilustración 15. Red de los estudiantes con NetDraw

Vemos representadas gráficamente las relaciones que habíamos introducido en la matriz. Así, por cada par de personas que habíamos conectado en la matriz (mediante el 1 que señalaba la relación) observamos un vínculo que une los dos nodos. Como se trata de una relación no orientada, cada grafo tiene una

doble flecha (Hugo conoce a Liliana y Liliana conoce a Hugo). Un grafo no orientado también se representa como un segmento que une dos nodos, sin ninguna flecha que indique dirección.

Esta red sociométrica presenta una fotografía de las relaciones que existen al interior de un grupo (la clase, en este caso). A simple vista podemos observar que existen individuos más conectados que otros, que existen sub-grupos dentro de la red, etc. Podríamos entonces intentar establecer hipótesis acerca de las relaciones existentes. Por ejemplo, podríamos suponer que las personas que cursan las mismas carreras se relacionan más entre sí que las de carreras diferentes. Para ello, necesitamos complementar la información con los datos atributivos que hemos recopilado.

Para ello, añadiremos la matriz que contiene la información sobre las carreras a la matriz que representaba las relaciones de conocimiento y que ya hemos realizado. La operación es muy sencilla. Sin salir de la red de estudiantes (Ilustración 17) realizamos:

File>Open>Ucinet dataset>Attribute Data

Y seleccionamos el fichero “carrera.##h”.

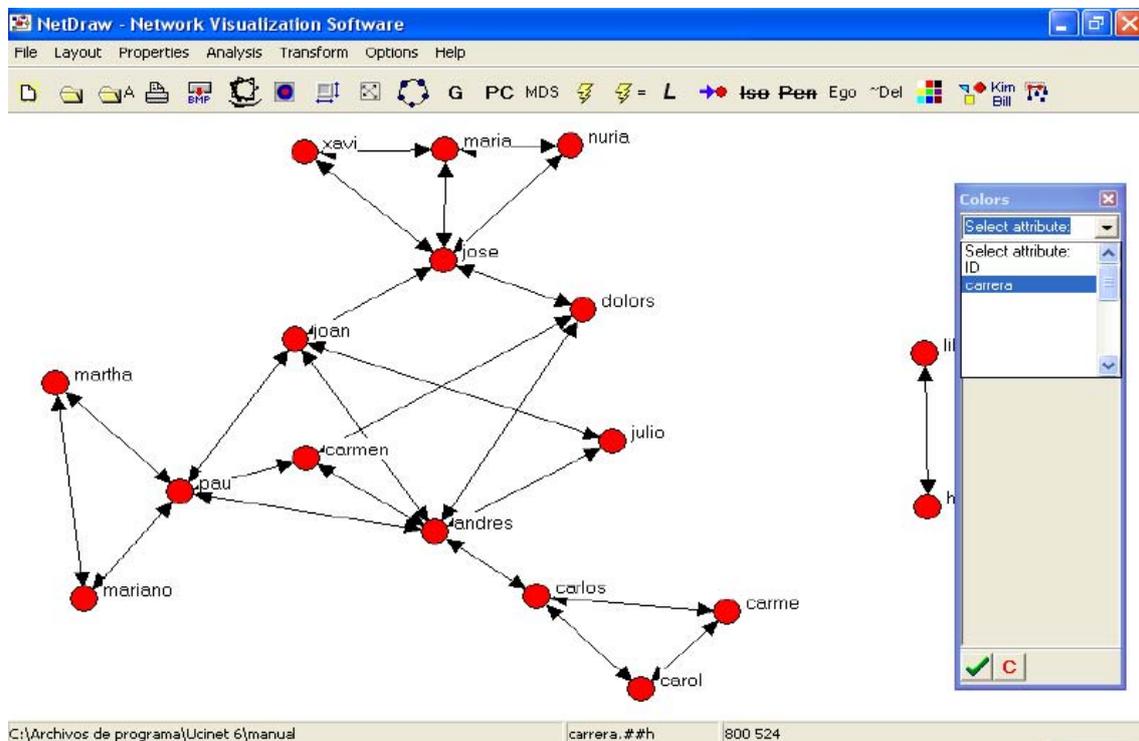


Ilustración 16. Red de estudiantes y la información sobre sus estudios

Veremos los cuatro códigos que hemos introducido para las carreras, cada uno de ellos asociado a un color diferente. Si quisiéramos cambiar el color adjudicado a alguno, sólo hay que hacer click sobre el rectángulo de color que se encuentra a la izquierda del código y seleccionar el color que deseamos en la paleta que se despliega. En este caso, cambiaremos el color correspondiente al código 2, de negro a verde:

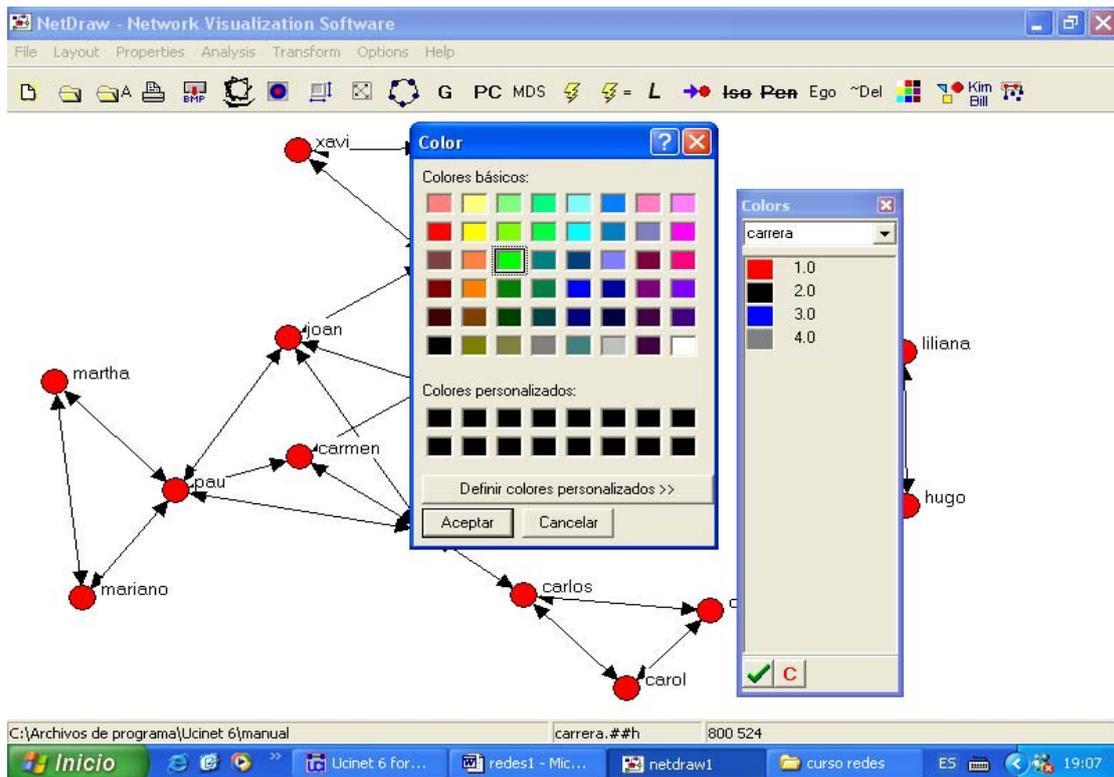


Ilustración 17. Cambiando los colores de las categorías de la variable atributiva "carrera"

Una vez efectuado el cambio, hacemos click sobre el botón "Aceptar". Vemos entonces que a cada persona le corresponde el color de la carrera que está estudiando: rojo para Antropología, verde para Arqueología, azul para Sociología y gris para Psicología Social.

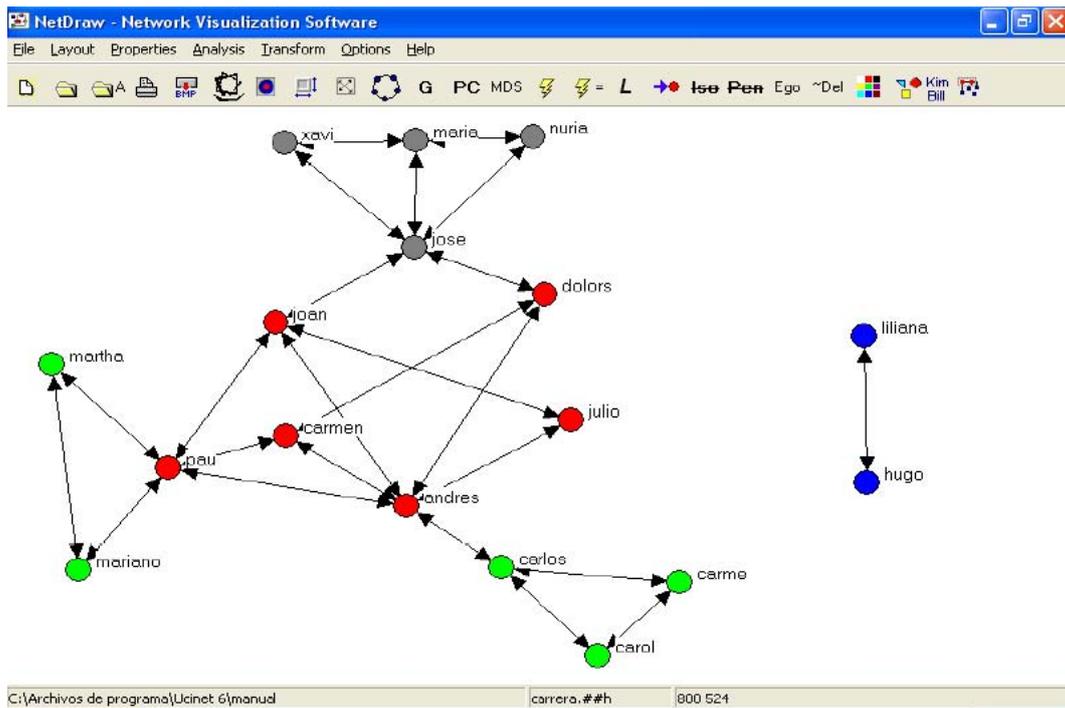


Ilustración 18. Representación de la red con el atributo correspondiente

A continuación, introduciremos los datos correspondientes al sexo, utilizando formas diferentes para los nodos, según representen hombres o mujeres. Introducimos entonces la matriz correspondiente al sexo (sexo.##h). Y seleccionamos la función “Choose node shapes according to attribute values” (tercer icono contando desde la derecha). Se abrirá una ventana similar a la que utilizamos para definir los colores. Seleccionamos el atributo “sexo”. Un consejo: si en el menú desplegable (“select attribute”) no vemos el atributo que queremos introducir, debemos descender hacia el final de la ventana, hasta encontrarlo.

A continuación seleccionaremos un triángulo para identificar a los hombres (codificados con el valor “1”) y un círculo para identificar a las mujeres (codificadas con el valor “2”):

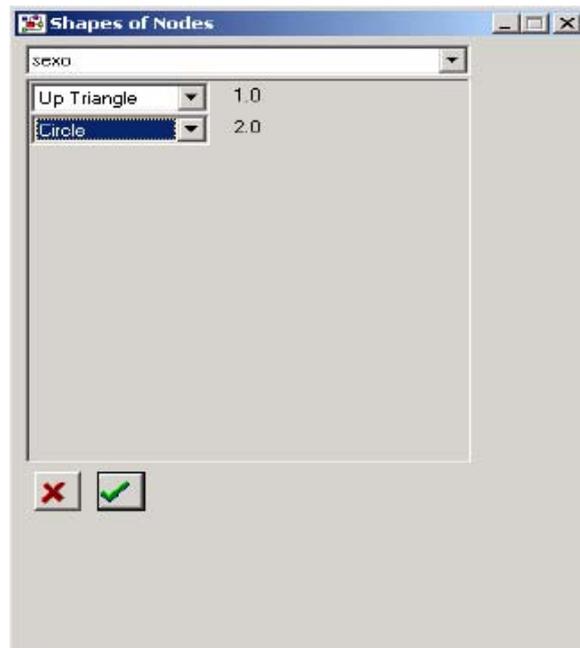


Ilustración 19. Variable atributiva "sexo"

De esta manera hemos introducido los dos atributos en la red. Y obtenemos el siguiente resultado:

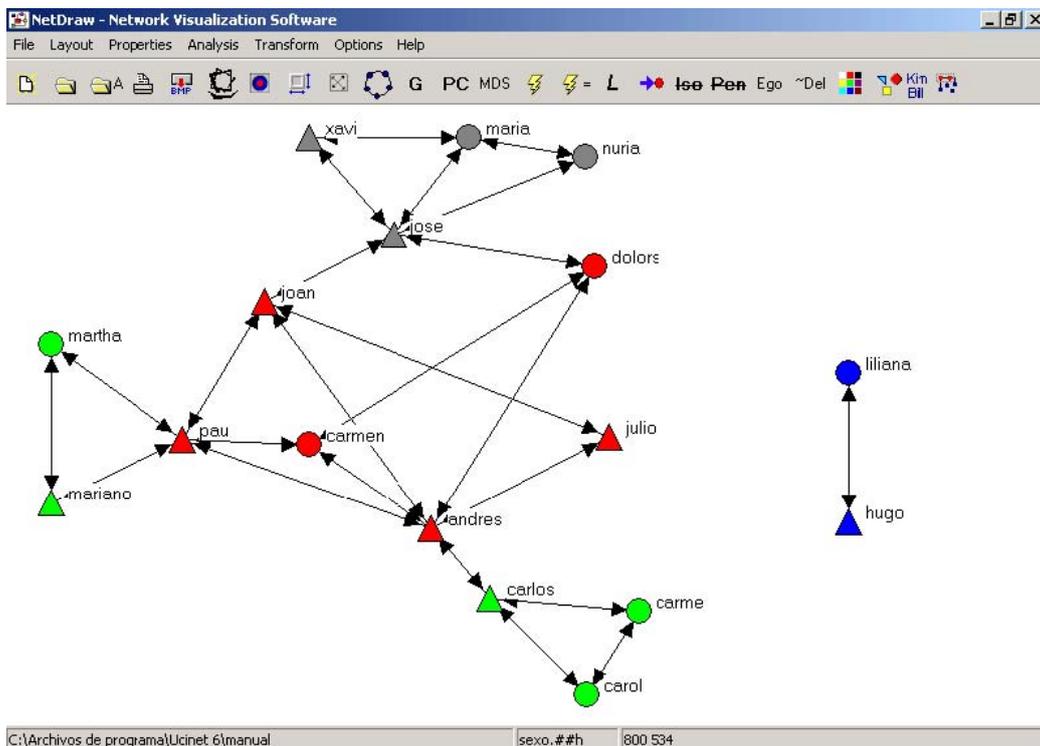


Ilustración 20. Red social con indicación del sexo (forma) y los estudios (color)

Hemos conseguido una descripción más detallada de la red de interacciones. Como podemos observar en la Ilustración 22, existen subgrupos en el interior

de nuestra red así como personas con más conexiones que otras. Así, los que cursan la carrera de Antropología (marcados en rojo) se conocían todos entre sí antes de iniciar el curso. Lo mismo pasa con los estudiantes de Psicología Social (nodos grises) y los de Sociología (nodos azules). Sin embargo, no sucede lo mismo con los estudiantes de Arqueología (verde), entre los que podemos observar dos grupos diferenciados. Por el contrario, no parece haber ningún patrón diferenciado de relación entre hombres y mujeres.

En lo que hace a las personas, algunos estudiantes sólo conocen a una o dos personas del grupo, mientras que otras (como Andrés o Joan) conocen a mucha gente. Pero para saber con certeza cuáles son las características de esta red, recurriremos al cálculo de las medidas de centralidad, a fin de realizar una primera aproximación al análisis de esta red social. Antes de pasar a esta sección, exportaremos la imagen de la red que hemos obtenido, para poder utilizarla posteriormente. Para ello, desde la función “File” de la barra de menús seleccionamos la función “Save diagram as” y luego el formato que elijamos (en este caso, “Bitmap”)

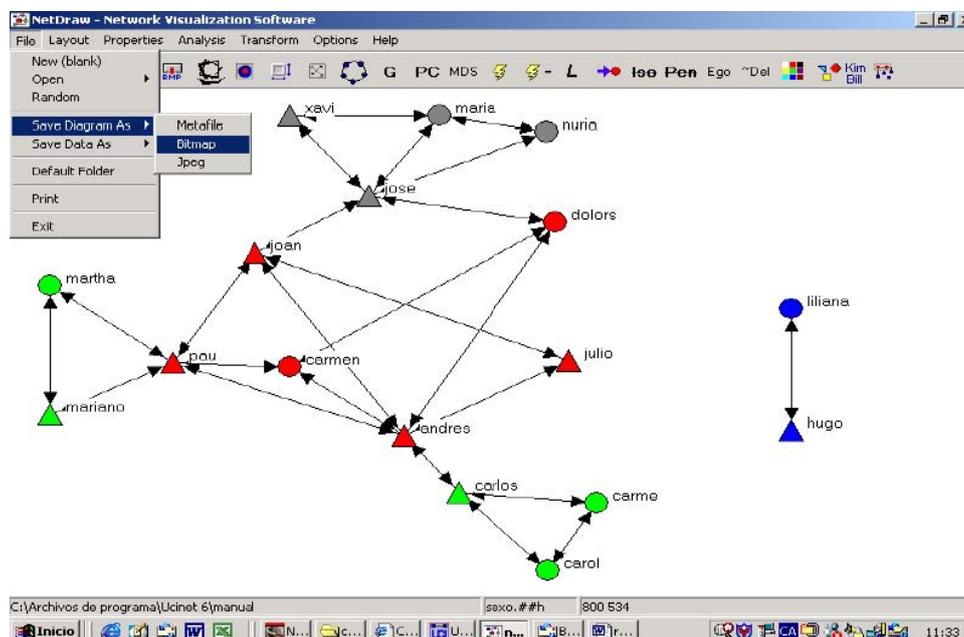


Ilustración 21. Guardando la red como imagen

La guardaremos con el nombre “redestudiantes”. Para concluir la primera parte de la práctica, crearemos un documento en Word (con el nombre “informe práctica.doc”) e insertaremos la imagen que hemos creado (Insertar>Imagen>Desde archivo):

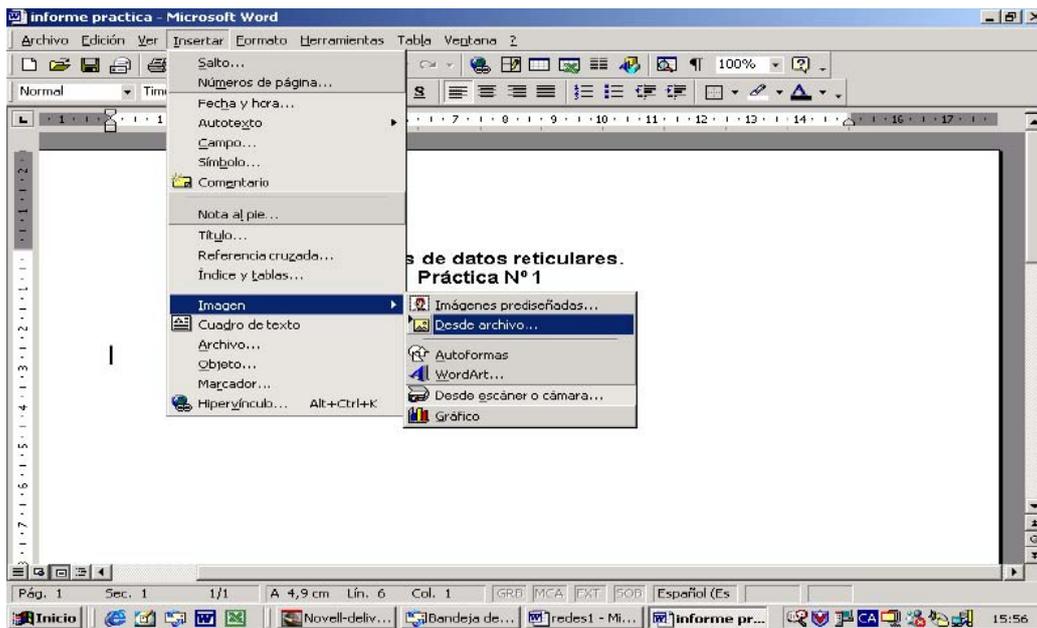


Ilustración 22. Insertando una red en Word

Seleccionamos la imagen a insertar (“redestudiantes”), y guardamos los cambios realizados en el informe.

MEDIDAS DE CENTRALIDAD

Como señala Hanneman (*op.cit.*),

La perspectiva de redes implica tener en cuenta múltiples niveles de análisis. Las diferencias entre los actores son interpretadas en base a las limitaciones y oportunidades que surgen de la forma en que éstos están inmersos en las redes; la estructura y el comportamiento de las redes está basado en y activado por las interacciones locales entre los actores. (...) Las diferencias en cómo los individuos están conectados puede ser extremadamente útil para entender sus atributos y comportamiento. Muchas conexiones significan a menudo que los individuos se exponen todavía a más y más diversa información. (cap V: p. 3)

Podemos realizar una primera aproximación a la estructura de una red social mediante el análisis de tres medidas de centralidad: rango (*degree*), grado de intermediación (*betweenness*) y cercanía (*closeness*). Por medida de centralidad se entiende un conjunto de algoritmos calculado sobre cada red que nos permite conocer la posición de los nodos en el interior de la red y la

estructura de la propia red. Veamos qué significa cada una de estas medidas, cómo se calculan y cómo se interpretan.

GRADO NODAL (DEGREE)

El grado nodal o rango es el número de lazos directos de un actor (o nodo), es decir con cuántos otros nodos se encuentra directamente conectado. Si volvemos a la figura 23, por ejemplo, podemos observar que Martha está conectada a dos personas: Mariano y Pau. Su rango, entonces es 2. Andrés, en cambio, conoce a seis personas: su rango es 6. En este caso es sencillo calcularlo a simple vista. Pero para trabajar con mayor precisión, utilizaremos Ucinet 6. Desde el menú Network seleccionaremos la función Centrality y luego Degree:

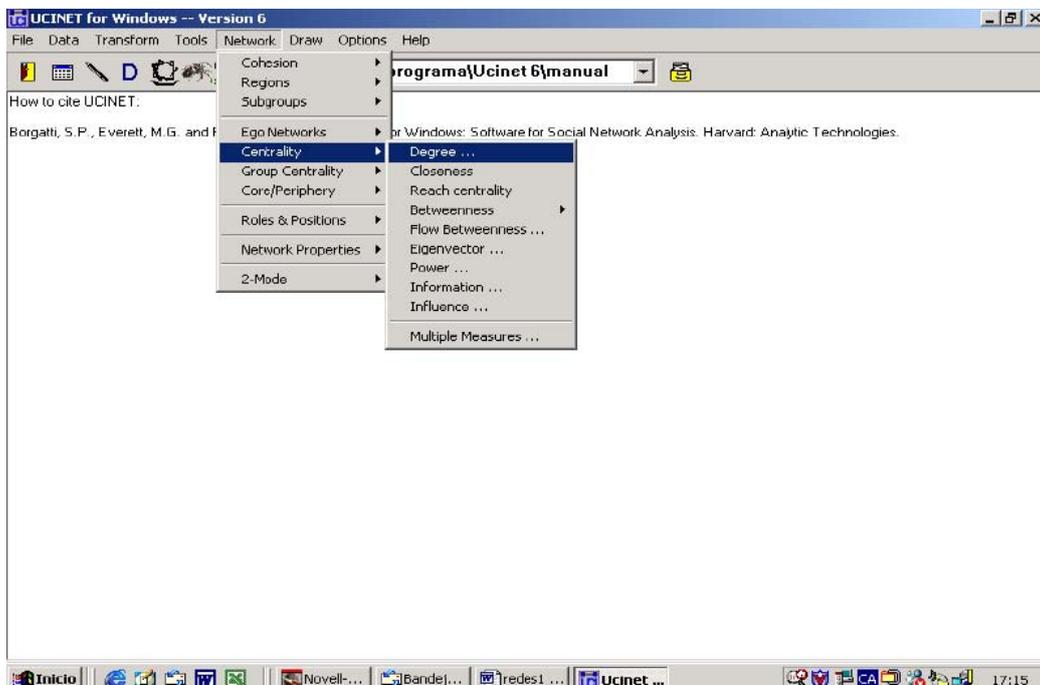


Ilustración 23. Calculando el grado nodal (Degree) con Ucinet 6

En la pantalla que se despliega seleccionaremos la matriz que contiene los datos que queremos analizar. En este caso, analizaremos la matriz “estudiantes.##h”, la que contiene la información sobre las relaciones de conocimiento entre los estudiantes de la clase. Para ello, haremos click con el ratón en el botón de la fila “input dataset” (marcado con tres puntos):



Ilustración 24. Seleccionando el fichero "estudiantes"

La primera fila ("Input dataset") indica el fichero que utilizará para calcular el rango. La segunda fila ("Treat data as symmetric") pregunta si se trata de una matriz simétrica (la relación de A-B vale para B-A) o no. En este caso, la respuesta es "sí". La tercera fila (Include diagonal values?) pregunta si debe tener en cuenta los valores de la diagonal (la relación entre A y A, B y B, C y C, etc.), y responderemos "no". La cuarta fila informa con qué nombre se guardarán los resultados, en el directorio que tengamos predeterminado. Haciendo click en el icono de los tres puntos, se puede definir otro directorio y modificar el nombre del archivo con los datos de salida si así lo deseamos. Una vez realizados los cambios que queremos hacer, hacemos click en "aceptar". Obtenemos la siguiente pantalla:

FREEMAN'S DEGREE CENTRALITY MEASURES:

Diagonal valid? NO
 Model: SYMMETRIC
 Input dataset: C:\Archivos de programa\Ucinet 6>manual\estudiantes

	1	2	3
	Degree	NrmDegree	Share
1 andres	6.000	37.500	0.125
16 pau	5.000	31.250	0.104
9 jose	5.000	31.250	0.104
8 joan	4.000	25.000	0.083
2 carlos	3.000	18.750	0.063
6 dolors	3.000	18.750	0.063
4 carmen	3.000	18.750	0.063
12 maria	3.000	18.750	0.063
3 carme	2.000	12.500	0.042
15 nuria	2.000	12.500	0.042
5 carol	2.000	12.500	0.042
13 mariano	2.000	12.500	0.042
17 xavi	2.000	12.500	0.042
10 julio	2.000	12.500	0.042
14 martha	2.000	12.500	0.042
11 lilliana	1.000	6.250	0.021
7 hugo	1.000	6.250	0.021

DESCRIPTIVE STATISTICS

	1	2	3
	Degree	NrmDegree	Share
1 Mean	2.824	17.647	0.088

Ilustración 25. Resultados del cálculo del grado nodal

Vemos la lista de todos los nodos de la red, ordenados de mayor rango (más número de conexiones) a menor rango (menor número de conexiones). Se reproduce a continuación la información obtenida en las tres primeras columnas:

	Degree	NrmDegree
Andrés	6.0	37.500
Pau	5.0	31.250
José	5.0	31.250
Joan	4.0	25.000
Carlos	3.0	18.750
Dolors	3.0	18.750
Carmen	3.0	18.750
María	3.0	18.750
Carme	2.0	12.500
Nuria	2.0	12.500
Carol	2.0	12.500
Mariano	2.0	12.500
Xavi	2.0	12.500
Julio	2.0	12.500
Martha	2.0	12.500
Liliana	1.0	6.250
Hugo	1.0	6.250

La columna “Nrmdegree” indica el rango normalizado, es decir, el porcentaje de conexiones que tiene un nodo sobre el total de la red.

El análisis del rango o grado nodal nos indica que la persona más conectada en este grupo, la de mayor centralidad, es Andrés, y que Liliana y Hugo son los menos centrales, con un rango de 1 cada uno. Podríamos suponer que Andrés es la persona que controla mayor cantidad de información. El grado nodal puede ser considerado una medida que permite acceder al índice de accesibilidad a la información que circula por la red. Si, por ejemplo, en el interior del grupo circularan ciertos rumores, los actores con un grado nodal más alto tendrán mayores probabilidades de escucharlos y difundirlos. El rango también puede ser interpretado como el grado de oportunidad de influir o ser influido por otras personas en la red. Pongamos por ejemplo el caso de la red que estamos analizando. Supongamos que nos interesa realizar algunas modificaciones en el dictado de la asignatura o que queremos conocer la valoración de los estudiantes sobre su desarrollo. En este caso Andrés (con un

grado nodal de 6, el más alto de la red) parece ser un buen informante, alguien que tiene acceso a bastante información de la que circula por el grupo. Podemos utilizar el rango también, por ejemplo, como un método de selección de personas para entrevistas o negociaciones.

Pero debemos complementar este análisis con las otras medidas de centralidad, para obtener un panorama más completo. Antes de hacerlo, nos referiremos a los estadísticos descriptivos que acompañan la información:

Output Log #1

Log File Number 1

Node	Degree	NrmDegree	Share
17 xavi	2.000	12.500	0.042
10 julio	2.000	12.500	0.042
14 martha	2.000	12.500	0.042
11 liliana	1.000	6.250	0.021
7 hugo	1.000	6.250	0.021

DESCRIPTIVE STATISTICS

	1 Degree	2 NrmDegree	3 Share
1 Mean	2.824	17.647	0.000
2 Std Dev	1.382	8.638	0.000
3 Sum	48.000	300.000	0.000
4 Variance	1.910	74.611	0.000
5 SSQ	168.000	6562.500	0.000
6 MCSSQ	32.471	1268.382	0.000
7 Euc Ncrm	12.961	81.009	0.000
8 Minimum	1.000	6.250	0.000
9 Maximum	6.000	37.500	0.000

Network Centralization = 22.50%
Homogeneity = 7.29%

NOTE: For valued data, both the normalized centrality and the centralization index may be larger

Actor-by-centrality matrix saved as dataset FreemanDegree

Running time: 00:00:01
Output generated: 23 jun 03 17:26:54
Copyright (c) 1999-2000 Analytic Technologies

Ilustración 26. Estadísticos descriptivos del cálculo del grado nodal

Los estadísticos descriptivos brindan información sobre los valores que tomó el rango en el conjunto de la red. Vemos que el rango promedio de la red (“Mean”) es 2.824, y que los valores oscilan entre 1 y 6 (la menor y mayor cantidad de lazos).

Incorporaremos esta información al informe de la práctica, de manera muy sencilla: seleccionamos con el ratón el texto y seleccionamos la función “copiar” (Edit>Copy). Pegamos la información en el informe en Word que estamos elaborando (Edición>Pegar).

GRADO DE INTERMEDIACIÓN (BETWEENNESS)

El grado de intermediación indica la frecuencia con que aparece un nodo en el tramo más corto (o geodésico) que conecta a otros dos. Es decir, muestra

cuando una persona es intermediaria entre otras dos personas del mismo grupo que no se conocen entre sí (lo que podríamos denominar “persona puente”). Veamos cuáles son los valores para el grado de intermediación en nuestra red para localizar un ejemplo.

Para calcular el grado de intermediación seleccionamos el menú Network>Centrality>Betweenness>Nodes:

En la pantalla que se despliega elegimos la matriz que contiene los datos que vamos a analizar (“estudiantes.###”) y hacemos click en el botón “ok”. Obtenemos como resultado el valor del grado de intermediación de cada nodo, ordenados de mayor a menor:

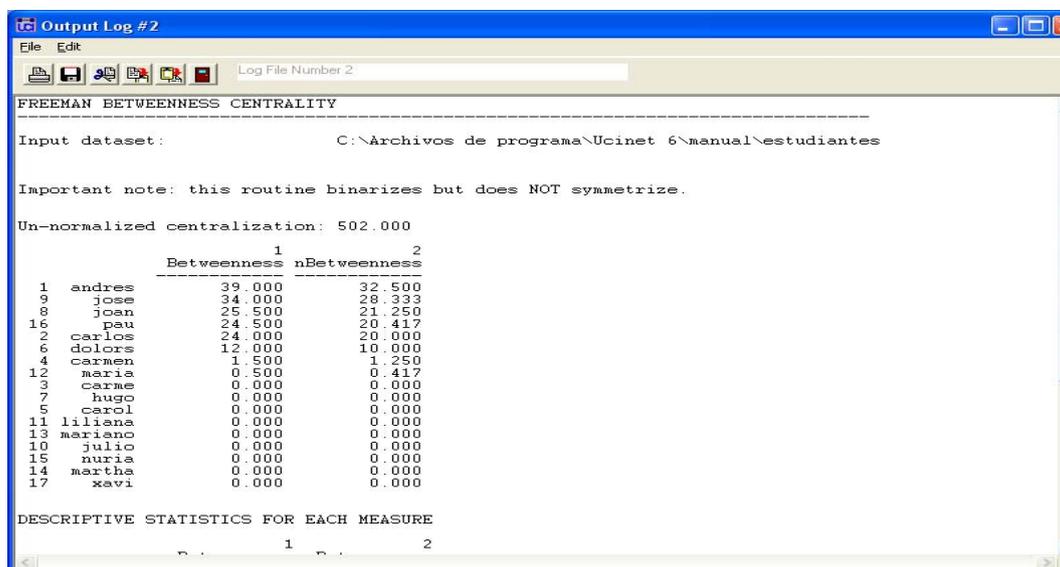


Ilustración 27. Resultados del cálculo del grado de intermediación

Se reproducen a continuación los resultados obtenidos:

	Betweenness	nBetweenness
Andrés	39.0	32.5
José	34.0	28.3
Joan	25.5	21.2
Pau	24.5	20.4
Carlos	24.0	20.0
Dolors	12.0	20.0
Carmen	1.5	10.0
María	0.5	1.2
Carme	0.0	0.4
Hugo	0.0	0.0
Carol	0.0	0.0
Liliana	0.0	0.0
Mariano	0.0	0.0

Julio	0.0	0.0
Nuria	0.0	0.0
Martha	0.0	0.0
Xavi	0.0	0.0

Andrés y José son las personas con un mayor grado de intermediación. Si observamos la figura 23 vemos que Andrés es la persona que conecta al grupo de los antropólogos con uno de los grupos de los arqueólogos (el formado por Carlos, Carme y Carol) y José conecta al grupo de los estudiantes de psicología social con los de antropología. Pero Joan, Pau y Carlos también tienen valores altos de grado de intermediación. La relación entre Joan y Pau permite conectar tres grupos: arqueólogos, antropólogos y psicólogos sociales. En el caso de Carlos, su relación con Andrés vincula a Carol y Carme con un grupo mayor. Si esa relación se rompiera, el grupo quedaría aislado.

Realizaremos la misma operación que en el caso del rango: seleccionaremos los resultados obtenidos y con la función “copiar” incorporaremos la información en el documento del informe.

CERCANÍA (CLOSENNESS)

El grado de cercanía indica la cercanía de un nodo respecto del resto de la red. Representa la capacidad que tiene un nodo de alcanzar a los demás. Veremos su interpretación a partir de un ejemplo.

Para calcularlo, procederemos de la misma manera que en las dos medidas de centralidad anteriores, es decir, a partir del menú “Network” (Network>Centrality>Closeness). Seleccionamos el archivo a analizar y damos click en el botón “ok”. El resultado de la operación, al igual que en los casos anteriores, es la lista ordenada de los valores de cercanía, y también el de su opuesto, lejanía (*farness*):

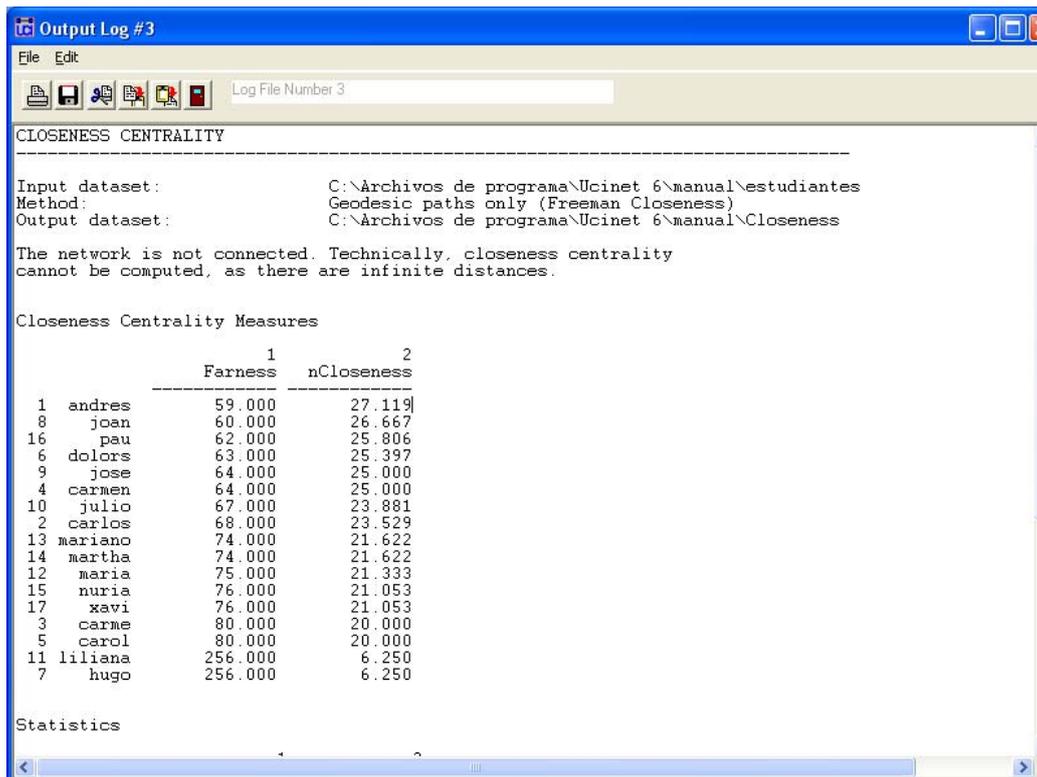


Ilustración 28. Cálculo de la cercanía

Se reproducen a continuación los resultados:

	Farness	Closeness
Andrés	59.0	27.11
Joan	60.0	26.66
Pau	62.0	25.80
Dolors	63.0	25.39
José	64.0	25.00
Carmen	64.0	25.00
Julio	67.0	23.88
Carlos	68.0	23.52
Mariano	74.0	21.62
Martha	74.0	21.62
Maria	75.0	21.33
Nuria	76.0	21.05
Xavi	76.0	21.05
Carme	80.0	20.0
Carol	80.0	20.0
Liliana	256.0	6.25
Hugo	256.0	6.25

Al igual que en las dos medidas de centralidad anteriormente analizadas, Andrés es la persona de la red que cuenta con un grado de cercanía más grande. Tiene una mayor capacidad para acceder al resto de los nodos de la

red. Sin embargo, los resultados para el resto de las personas no son iguales a los obtenidos en las mediciones anteriores. Si volvemos a los resultados del grado nodal vemos que, por ejemplo, Dolors tenía el mismo rango que Carlos, Carmen y María (conocían a tres personas cada uno). Sin embargo, el grado de cercanía de Dolors es mayor que el de ellos. No sólo es importante el número de personas que conoces, sino quiénes son esas personas, cuál es su grado de conexiones. Si observamos la red vemos que Dolors conoce a Andrés y a José, las dos personas con mayor grado de cercanía de toda la red. De esta manera, su probabilidad de acceder al resto de los nodos es más alta. “Una persona poco conectada con el resto (baja centralidad, bajo grado de intermediación) por el solo hecho de estar conectada con una persona ‘importante’ puede tener una alta cercanía” (Molina, *op.cit.* p. 79).

Finalizaremos ahora el ejercicio, copiando los resultados y llevándolos al informe que hemos elaborado.

REDES DENTRO DE REDES: CLIQUES

Regresemos ahora al gráfico de nuestra red. Una característica que podemos observar a simple vista es que dentro de ella es posible identificar grupos o sub-grafos. Por ejemplo, Carlos, Carme y Carol constituyen un grupo: están relacionados entre ellos (de hecho, tienen se encuentran más vinculados entre ellos que con el resto de la red). Este subgrupo recibe el nombre técnico de *clique* (si bien, como veremos posteriormente, podríamos hablar de dos acepciones de clique: en sentido “duro” y en sentido “blando”). Un clique (en sentido “blando”) es un grupo dentro de una red.

El análisis de sub-grafos, o de cliques, es un tipo de aproximación a la estructura de la red, una aproximación de “abajo hacia arriba”, como señala Hanneman (*op. cit.*):

La noción [de clique] parte de los vínculos simples para “construir” la red. Un mapa de toda la red puede ser construido examinando los tamaños de los distintos cliques y agrupaciones de tipo de clique, notando sus tamaños y yuxtaposiciones. Este tipo de aproximación, acerca de las subestructuras de las redes, tienden a enfatizar cómo lo macro puede surgir de lo micro. Tienden a enfocar nuestra atención primero en los individuos y en entender cómo están

inmersos en la estructura mayor de la red a partir de los grupos yuxtapuestos. Esta idea aparentemente obvia se tiene que destacar porque también es posible aproximarse a la cuestión de la subestructuras de las redes desde una perspectiva de arriba hacia abajo. Ambos aspectos son valiosos y complementarios. (cap V: p. 6)

Ucinet permite analizar las subestructuras de la red, a partir de una serie de medidas de agrupamiento, que revisaremos a continuación

CLIQUEES

Hemos utilizado anteriormente la palabra clique como sinónimo de subgrupo dentro de una red. Pero en sentido estricto se denomina clique a un conjunto de nodos o actores que tienen todos los vínculos posibles entre ellos. Los actores que conforman un clique deben ser más de dos, por lo general se trabajan cliques de tres y más integrantes. Un grupo de nodos que tienen entre ellos todos los vínculos posibles se denomina “subgrafo máximo completo”.

Pediremos a Ucinet que identifique todos los cliques de más de tres integrantes que se encuentran en nuestra red. Para ello, desde la pantalla inicial de Ucinet seleccionaremos el menú “Network” y luego Subgroups>Cliques:

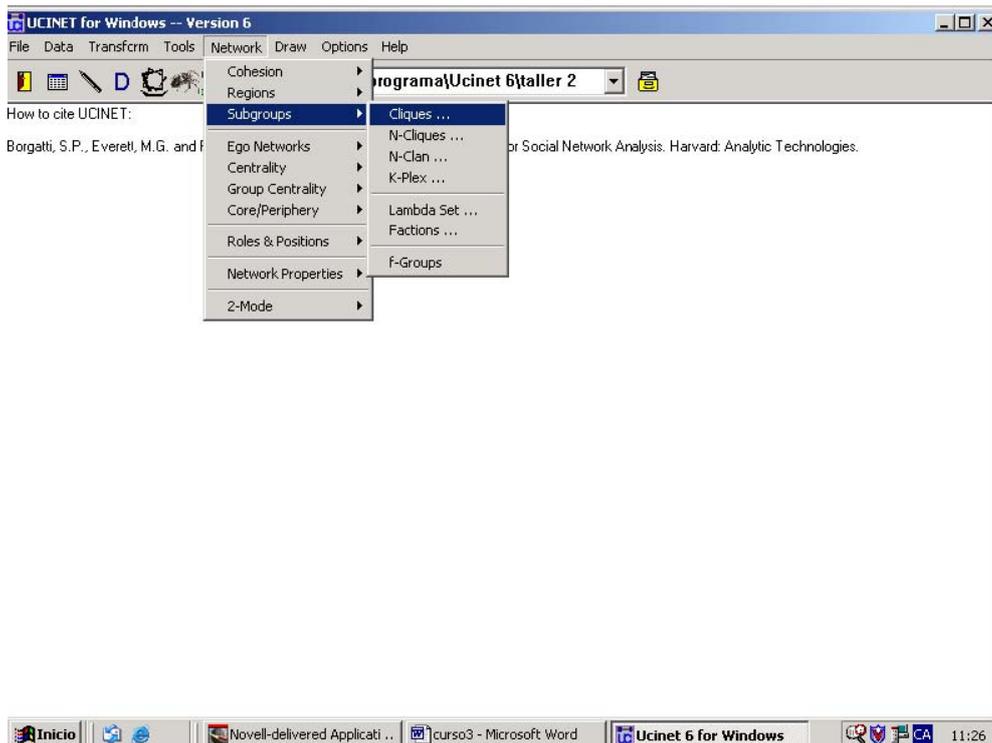


Ilustración 29. Localizando cliques en la red

En el cuadro de diálogo que se despliega (que es muy similar a los que utilizamos con las medidas de centralidad) seleccionaremos en primer lugar la matriz en donde se encuentran los datos que queremos analizar (Input dataset). En este caso es la que hemos estado utilizando hasta ahora (estudiantes.###h). La segunda opción es el número mínimo de integrantes del clique. La opción por defecto es 3 y en este caso la mantendremos. A continuación se nos pregunta si queremos que se analice el patrón de superposiciones (*Analyze pattern of overlaps*). Esto significa que podemos saber cuándo un integrante de un clique pertenece a otros, y con cuales miembros de la red comparte cliques. Es una información muy interesante, que analizaremos con más detalle cuando veamos los resultados. Ignoraremos por ahora el resto de las opciones, ya que es necesario que en primer lugar aprendamos las cuestiones básicas de esta medida.

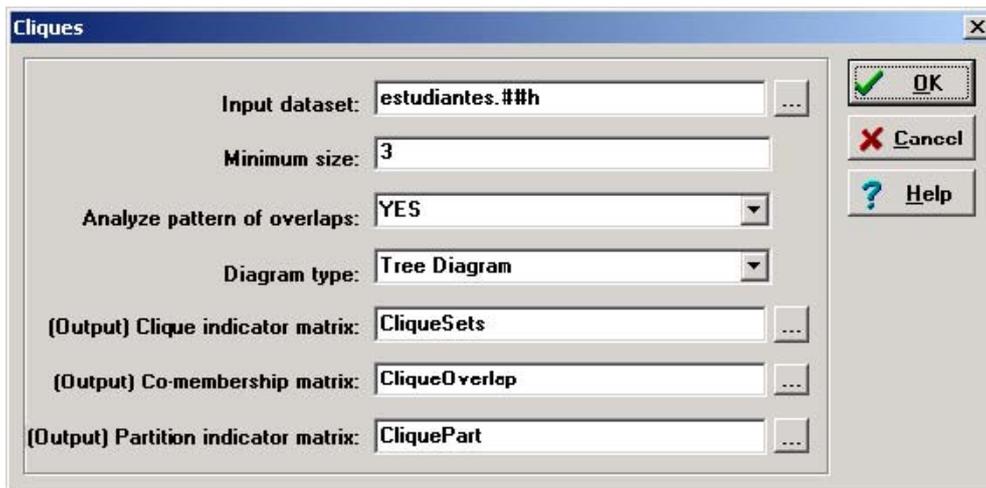


Ilustración 30. Ventana de diálogo de la opción Cliques

Al hacer click en “OK” obtenemos las siguiente pantalla:

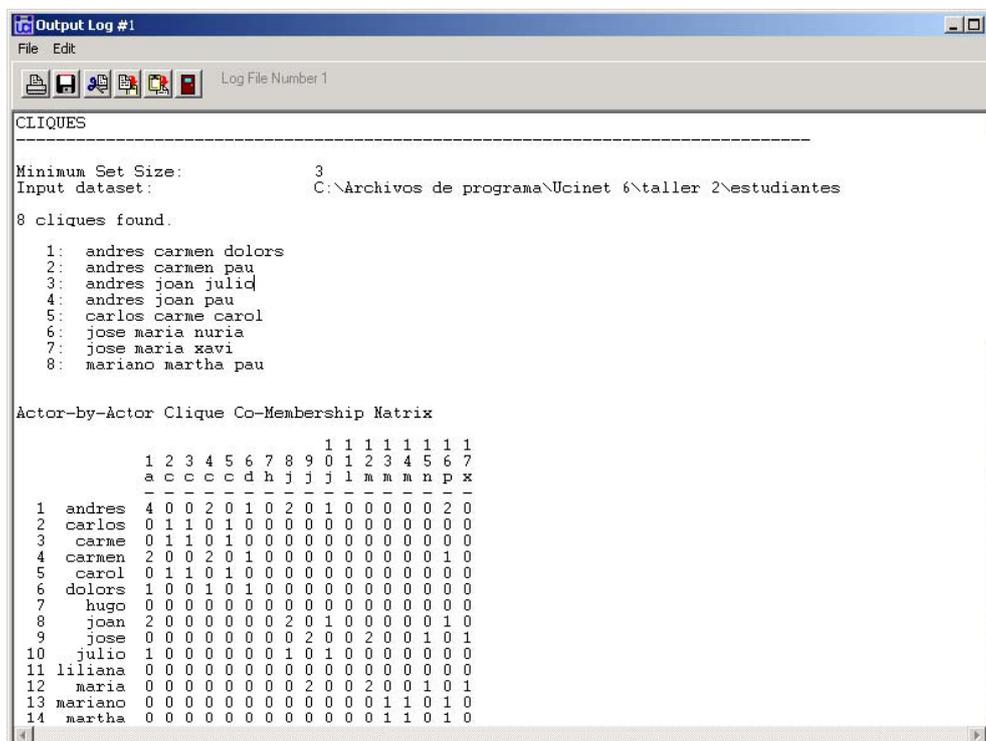


Ilustración 31. Ventana de datos de los resultados del cálculo de los Cliques

El primer dato que obtenemos es que se han encontrado 8 cliques en nuestra red. Transcribo a continuación los resultados:

8 cliques found	
1	Andrés- Carmen- Dolors
2	Andrés- Carmen- Pau
3	Andrés- Joan- Julio
4	Andrés- Joan- Pau
5	Carlos- Carme- Carol

		1 1 1 1 1 1 1 1																
		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7
		a	c	c	c	c	d	h	j	j	j	l	m	m	m	n	p	x
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	andres	4	0	0	2	0	1	0	2	0	1	0	0	0	0	0	2	0
2	carlos	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
3	carme	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	carmen	2	0	0	2	0	1	0	0	0	0	0	0	0	0	0	1	0
5	carol	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
6	dolors	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
7	hugo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	joan	2	0	0	0	0	0	0	2	0	1	0	0	0	0	0	1	0
9	jose	0	0	0	0	0	0	0	0	2	0	0	2	0	0	1	0	1
10	julio	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
11	liliana	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	maria	0	0	0	0	0	0	0	2	0	0	2	0	0	1	0	1	0
13	mariano	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0
14	martha	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	0
15	nuria	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
16	pau	2	0	0	1	0	0	0	1	0	0	0	1	1	0	3	0	0
17	xavi	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0

Como habíamos visto al analizar los cliques, Andrés pertenece a cuatro cliques diferentes, seguido por Pau, que pertenece a tres. La mayoría de los actores sólo pertenece a uno, mientras que tenemos dos actores aislados, Liliana y Hugo, que no pertenecen a ninguno, ya que, como vimos en el análisis de la graficación de la red y en las medidas de centralidad, constituyen una díada aislada del resto de la red.

Para obtener los otros dos datos que presenta esta matriz (con quién se comparte clique y qué número de cliques se comparte con cada actor), podemos mirar la línea horizontal o la vertical de cada actor:

		1 1 1 1 1 1 1 1																
		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7
		a	c	c	c	c	d	h	j	j	j	l	m	m	m	n	p	x
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	andres	4	0	0	2	0	1	0	2	0	1	0	0	0	0	0	2	0
2	carlos	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
3	carme	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	carmen	2	0	0	2	0	1	0	0	0	0	0	0	0	0	0	1	0
5	carol	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
6	dolors	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
7	hugo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	joan	2	0	0	0	0	0	2	0	1	0	0	0	0	0	0	1	0
9	jose	0	0	0	0	0	0	0	2	0	0	2	0	0	1	0	1	0
10	julio	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0

```

11 liliانا 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
12 maria 0 0 0 0 0 0 0 0 2 0 0 2 0 0 1 0 1
13 mariano 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0
14 martha 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0
15 nuria 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0
16 pau 2 0 0 1 0 0 0 1 0 0 0 0 0 1 1 0 3 0
17 xavi 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1

```

Podemos observar, por ejemplo, que Andrés no comparte ningún clique con Carlos o Carme, pero comparte dos cliques con Carmen, Joan y Pau, y uno con Dolors y con Julio. O que Dolors comparte un clique con Andrés y otro con Carmen, y ninguno más con el resto de los actores de la red, y constatar el aislamiento de Liliانا y Hugo. Pero también podemos analizar qué cliques comparten miembros entre sí. En este caso, nuestra unidad ya no son los actores, sino los cliques.

Clique-by-Clique Co-membership matrix

```

  1 2 3 4 5 6 7 8
-----
1 6 2 1 1 0 0 0
2 2 6 1 2 0 0 1
3 1 1 6 2 0 0 0
4 1 2 2 6 0 0 1
5 0 0 0 0 6 0 0
6 0 0 0 0 0 6 2 0
7 0 0 0 0 0 2 6 0
8 0 1 0 1 0 0 0 6

```

Esta matriz presenta en las filas y las columnas los cliques y leeremos la línea vertical o la horizontal, ignorando la diagonal:

```

  1 2 3 4 5 6 7 8
-----
1 6 2 1 1 0 0 0
2 2 6 1 2 0 0 1
3 1 1 6 2 0 0 0
4 1 2 2 6 0 0 1
5 0 0 0 0 6 0 0
6 0 0 0 0 0 6 2 0
7 0 0 0 0 0 2 6 0
8 0 1 0 1 0 0 0 6

```

Vemos que el clique N° 1 (integrado por Andrés, Carmen y Dolors) comparte dos integrantes con el clique N° 2 (integrado por Andrés, Carmen y Pau). Andrés y Carmen son los dos integrantes que se repiten en ambos cliques. El clique N° 1 también comparte un integrante con el clique N° 3 y con el N° 4 (Andrés es el integrante que se repite en ambos casos).

A continuación, copiaremos los resultados obtenidos en un documento de word, que guardaremos con el nombre de “informe practica 2”):

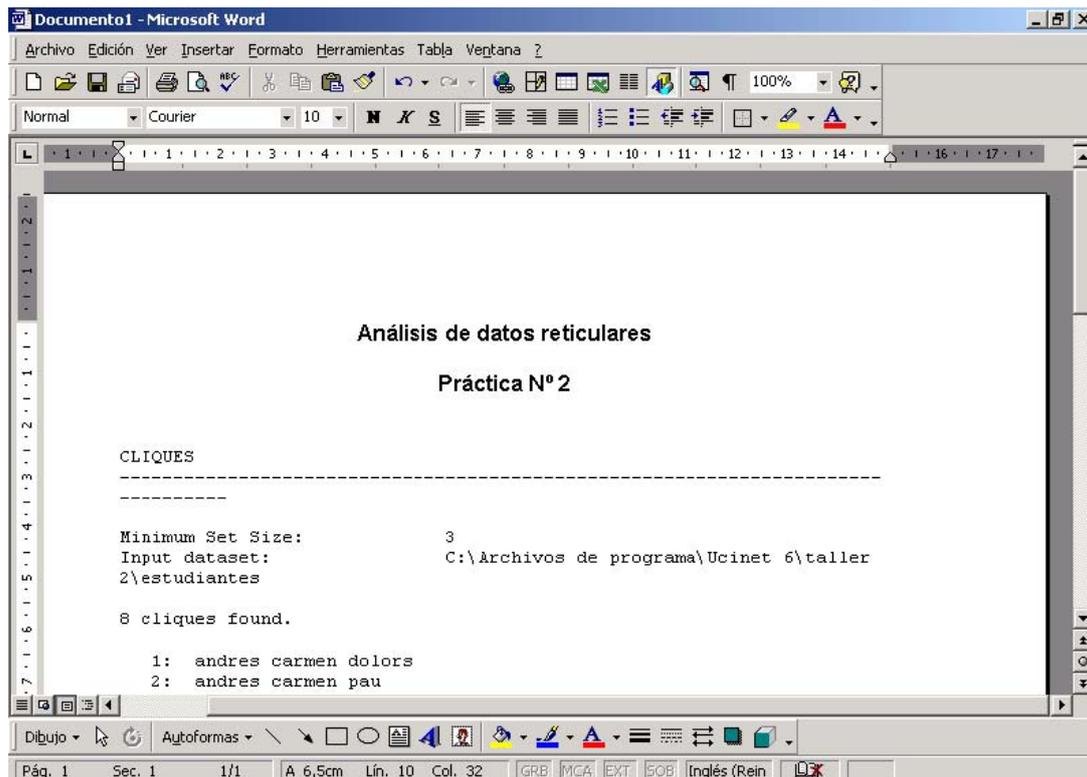


Ilustración 32. Guardando los resultados en Word

N-CLIQUE, N-CLAN Y K-PLEX

Como señala Hanneman (*op. cit*) en algunas ocasiones, la definición “dura” de clique (sub-grafo máximo completamente conectado) es demasiado estricta para nuestros fines. Puede interesarnos conocer cliques en los que algunos de sus miembros no estén conectados con todos los demás integrantes, por ejemplo. Existen una serie de medidas de agrupación que relajan la condición de subgrafo máximo completo: N-clique, N-clan y k-plex, que revisaremos a continuación.

Según la primera medida, N-clique, un actor es miembro de un clique si está conectado con todos los miembros del grupo a una distancia mayor que uno. Generalmente se utiliza una distancia de dos (lo que equivale a ser amigo de un amigo).

Aplicaremos N-clique a nuestros datos. Para ellos, seleccionaremos el menú Network y a continuación Subgroups>N-cliques.

En el cuadro de diálogo que se despliega, seleccionaremos en primer lugar la matriz que contiene los datos a analizar (estudiantes.###). La segunda opción nos indica el tamaño de la n (la distancia que estableceremos). Mantendremos la opción por defecto, que es 2. La tercera opción es el tamaño mínimo de los cliques que identificaremos (mantendremos 3) y la cuarta nos pregunta si queremos obtener la matriz de superposiciones o co-membrecía (seleccionaremos “yes”). Como en el caso del análisis de cliques, ignoraremos ahora el resto de las posibilidades.

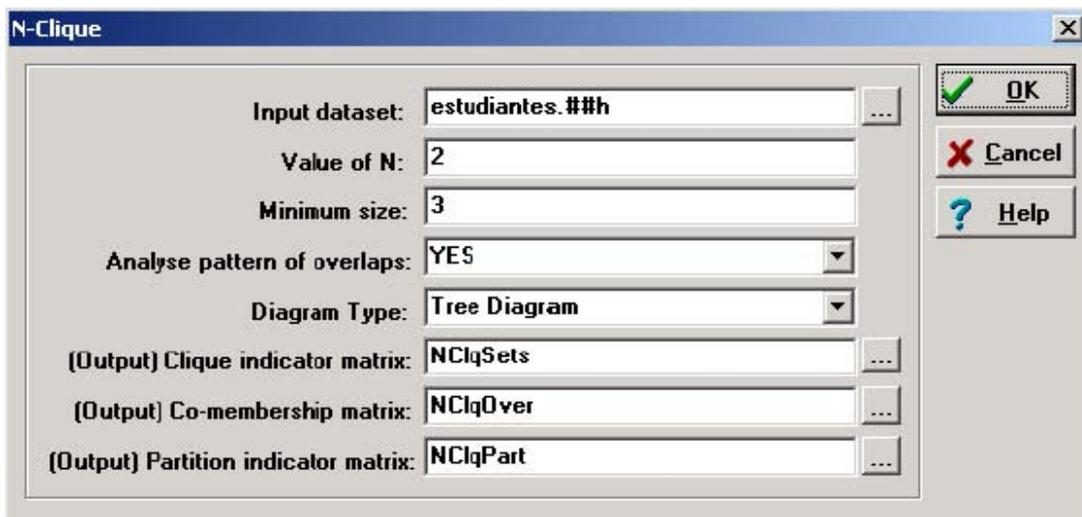


Ilustración 33. Ventana de diálogo de N-Cliques

Obtenemos el siguiente resultado:

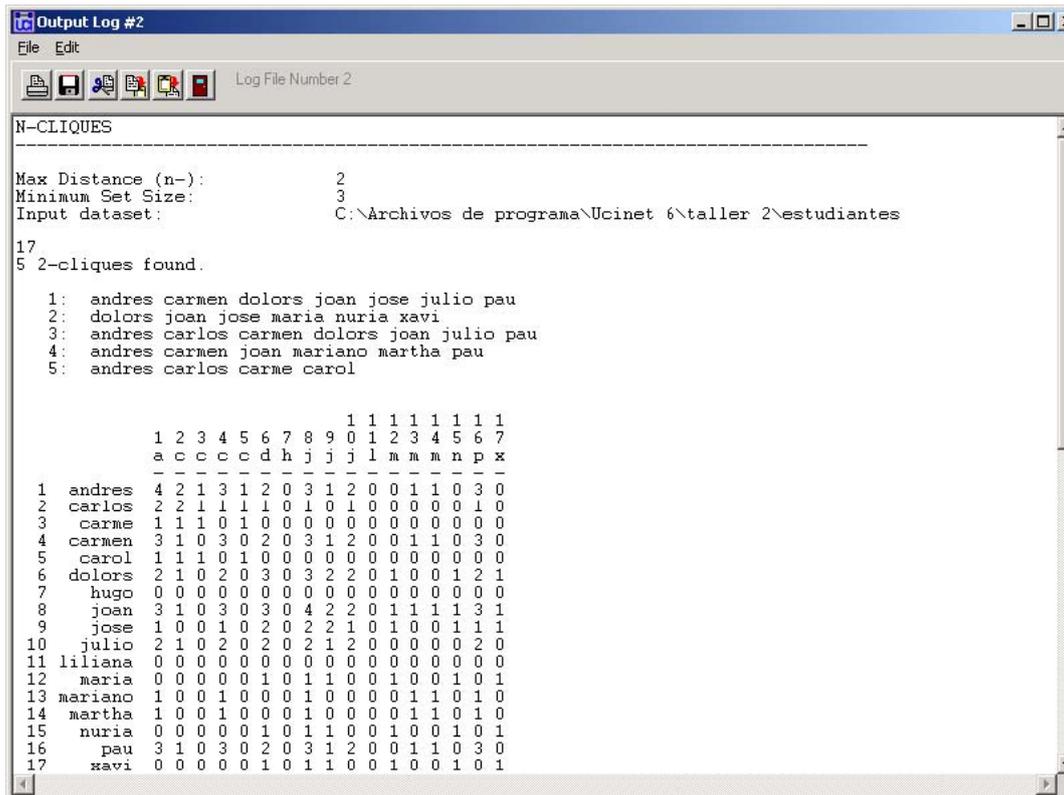


Ilustración 34. Resultados de N-Cliques

La primera diferencia en relación con el análisis anterior es que hemos obtenido tres cliques menos (ahora son cinco) pero de mayor tamaño. En efecto, la condición “conozco a alguien que lo conoce” nos devuelve grupos más grandes. Reproduzco a continuación los resultados de la primera matriz:

5 2-cliques found.	
1	andres carmen dolors joan jose julio pau
2	dolors joan jose maria nuria xavi
3	andres carlos carmen dolors joan julio pau
4	andres carmen joan mariano martha pau
5	andres carlos carme carol

Se han identificado dos cliques con siete miembros, dos con seis y uno con cuatro. N-clique es una medida mucho más inclusiva que clique. Por esta razón es importante ser cuidadosos a la hora de utilizarla, ya que podrían incluirse dentro de un mismo grupo actores que realmente no tienen una pertenencia clara.

En lo que hace a los actores que integran cada grupo, vemos que Andrés continua teniendo un protagonismo importante (integra cuatro de los cinco

cliques). Pero aparece una modificación: ya no es el único actor destacado, sino que tenemos a otro actor (Joan) integrando también cuatro cliques. Joan ha doblado su participación respecto de la medición anterior. Y este caso es un buen ejemplo para graficar la diferencia entre clique y N-clique. De acuerdo con las medidas de centralidad que hemos calculado anteriormente, Joan tiene un rango de 4, es decir, conoce a cuatro personas en la red: Andrés, Julio, Pau y José. Cuando calculamos los cliques, vimos que, de acuerdo con los datos de la tabla 5, Joan forma un clique junto con Andrés y Julio y otro con Andrés y Pau. No integra ningún clique con José, que es la cuarta persona que conoce en la red, ya que no tiene relación con el resto de los integrantes de estos cliques (José forma un clique con María y Nuria y otro con María y Xavi). Pero al aligerar la condición de la relación, a través de $N=2$, tenemos que pertenece al clique N° 1 porque conoce directamente a Andrés, Julio, José y Pau, e indirectamente a Carmen (a través de Andrés o de Pau) y a Dolors (a través de José o de Andrés). En el clique N° 2 sólo conoce directamente a José, pero “hereda” sus relaciones con María, Nuria y Xavi, y llega nuevamente a Dolors a través de Andrés o de Pau (aunque ellos no integren el clique. Retomaremos este punto en el análisis de N-clan). El mismo proceso sucede en los cliques N° 3 y 4.

Veamos ahora la matriz de solapamientos o co-membrecía. Leeremos la diagonal en primer lugar:

		1 1 1 1 1 1 1 1																
		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7
		a	c	c	c	d	h	j	j	j	l	m	m	m	n	p	x	
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	andres	4	2	1	3	1	2	0	3	1	2	0	0	1	1	0	3	0
2	carlos	2	2	1	1	1	1	0	1	0	1	0	0	0	0	0	1	0
3	carme	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	carmen	3	1	0	3	0	2	0	3	1	2	0	0	1	1	0	3	0
5	carol	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
6	dolors	2	1	0	2	0	3	0	3	2	2	0	1	0	0	1	2	1
7	hugo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	joan	3	1	0	3	0	3	0	4	2	2	0	1	1	1	1	3	1
9	jose	1	0	0	1	0	2	0	2	2	1	0	1	0	0	1	1	1
10	julio	2	1	0	2	0	2	0	2	1	2	0	0	0	0	0	2	0
11	liliana	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	maria	0	0	0	0	0	1	0	1	1	0	0	1	0	0	1	0	1
13	mariano	1	0	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0
14	martha	1	0	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0
15	nuria	0	0	0	0	0	1	0	1	1	0	0	1	0	0	1	0	1
16	pau	3	1	0	3	0	2	0	3	1	2	0	0	1	1	0	3	0
17	xavi	0	0	0	0	0	1	0	1	1	0	0	1	0	0	1	0	1

Ha aumentado el número de actores con los que se comparte clique. Esto se debe a la mayor inclusividad que tiene N-clique en relación con clique. Si vemos los resultados obtenidos ahora por los dos actores que tomamos como ejemplo a la hora de analizar la co-membrecía en la tabla 9, Andrés y Dolors, vemos que en ambos casos ha aumentado el número de personas con los que comparten clique: Andrés pasa de 5 a 11 y Dolors de 2 a 11. Y además aumenta el número de participaciones compartidas con cada actor.

Una manera de restringir N-clique es utilizando N-clan. Como hemos mencionado, en ocasiones al aplicar N-clique se muestra conectividad a través de actores que no son miembros del clique (el caso de la relación entre Dolors y Joan, en el clique 2, mediada por actor que no formaba parte de ese clique). Si en cambio utilizamos N-clan, introducimos una cláusula en el análisis: la relación continúa siendo mediada (accedo a alguien porque es conoce a una persona que yo conozco) pero todos los vínculos deben alcanzarse mediante otros miembros del clique.

Veamos lo que sucede con nuestros datos cuando aplicamos N-clan. El procedimiento es similar al resto de las medidas de agrupamiento: (Menú Network>Subgroups>N-clan).

El cuadro de diálogo es como los que habíamos utilizado anteriormente: seleccionamos la matriz que contiene los datos; determinamos el valor de N (que este caso será 2, al igual que en N-clique); pedimos cliques de tres o más integrantes y analizaremos el patrón de superposiciones.

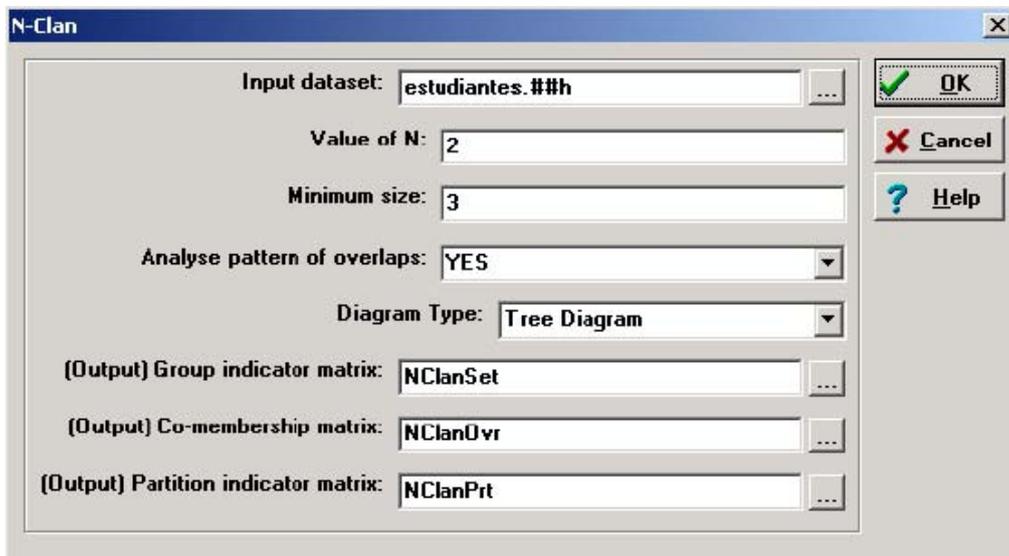


Ilustración 35. Ventana de diálogo de N-Clan

Obtenemos la siguiente pantalla de resultados:

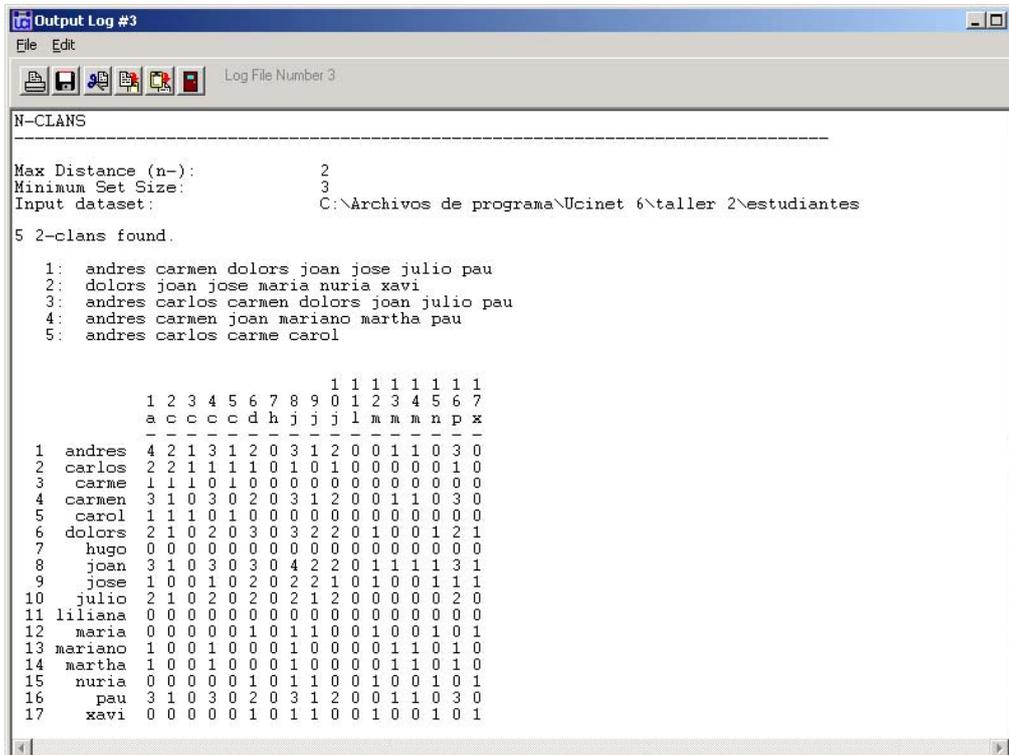


Ilustración 36. Resultados N-Cliques

Los resultados son muy similares a los obtenidos mediante N-clique: cinco cliques. Veamos su composición:

5 2-clans found.	
1	andres carmen dolors joan jose julio pau
2	dolors joan jose maria nuria xavi
3	andres carlos carmen dolors joan julio pau
4	andres carmen joan mariano martha pau
5	andres carlos carme carol

Aunque la composición es idéntica al caso de N-clique (ya que la definición de N-clique se encuentra incluida en la de N-clan), el procedimiento es diferente. Veamos el ejemplo de Andrés. Pertenece al clan N° 5 por su relación con Carlos, ya que no conoce directamente a Carme o a Carol. Pero no podría invocar a ningún otro actor para pertenecer al clan si este no perteneciera también. N-clan es una modificación menor de N-clique, y como hemos observado en el ejemplo, en muchos casos se obtienen idénticos resultados, pero ambas medidas se complementan, ya que permiten asegurarnos de que los actores que incluimos en un clique realmente pertenecen a él.

La última medida que analizaremos es K-plex. Su definición es muy sencilla: un actor puede ser miembro de un clique siempre y cuando tenga vínculos directos con todos excepto K miembros del grupo.

Hanneman dice que, si bien

(...) el enfoque K-plex parece tener bastante en común con el N-clique, el análisis basado en K-plex a menudo arroja un cuadro bastante diferente de las sub-estructuras del grafo. En vez de las agrupaciones grandes y concatenadas que a veces produce el análisis N-clique, el análisis K-plex tiende a encontrar número relativamente grandes de pequeñas agrupaciones. Esto tiende a llamar la atención sobre yuxtaposiciones y co-presencia (centralización) más que en solidaridad y accesibilidad. (cap- V: p. 11)

K-plex, entonces, muestra información sobre círculos sociales superpuestos. Veamos qué sucede cuando aplicamos K-plex a nuestros datos. Daremos a K el valor de 2 (un actor será miembro de un clique si conoce a todos menos dos actores del mismo). Seleccionamos el menú Network>Subgroups>K-plex. El

cuadro de diálogo es similar a todos los anteriores. Seleccionamos el documento a analizar, determinamos el valor de K (2), declaramos que los grupos deben tener 3 o más integrantes y pedimos la matriz de superposiciones:

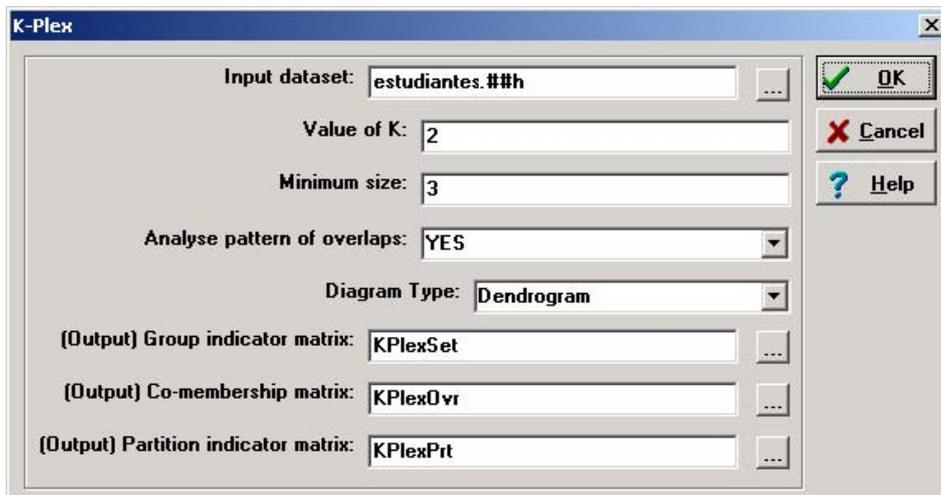


Ilustración 37. Ventana de diálogo de K-plex

Obtenemos el siguiente resultado:

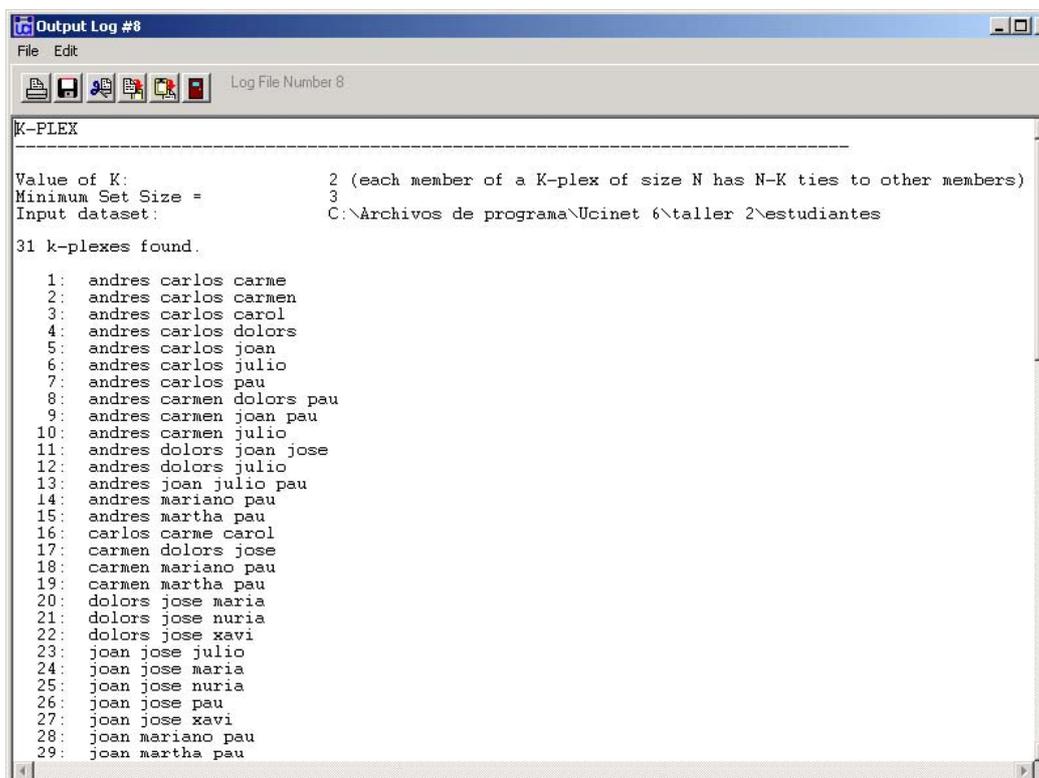


Ilustración 38. Resultados de K-Plex

Nuestra red consta de 31 K-plex, es decir 31 subgrupos en donde un actor conoce a todos los integrantes menos a dos. Transcribo los resultados de la primera tabla:

31 k-plexes found.	
1	andres carlos carne
2	andres carlos carmen
3	andres carlos carol
4	andres carlos dolors
5	andres carlos joan
6	andres carlos julio
7	andres carlos pau
8	andres carmen dolors pau
9	andres carmen joan pau
10	andres carmen julio
11	andres dolors joan jose
12	andres dolors julio
13	andres joan julio pau
14	andres mariano pau
15	andres martha pau
16	carlos carne carol
17	carmen dolors jose
18	carmen mariano pau
19	carmen martha pau
20	dolors jose maria
21	dolors jose nuria
22	dolors jose xavi
23	joan jose julio
24	joan jose maria
25	joan jose nuria
26	joan jose pau
27	joan jose xavi
28	joan mariano pau
29	joan martha pau
30	jose maria nuria xavi
31	mariano martha pau

Vemos una gran mayoría de K-plexes de tres integrantes (26) y sólo cinco con más de tres. Andrés es el actor que se repite mayor cantidad de veces, como podemos comprobar al mirar la matriz de superposiciones: está presente en 15 grupos. Pero también hay otros actores con participaciones importantes. Al utilizar esta medida, todos los actores aumentan el número de cliques a los que perteneces.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
		an	ca	ca	ca	ca	do	hu	jo	jo	ju	li	ma	ma	ma	nu	pa	xa
1	andres	15	7	1	4	1	4	0	4	1	4	0	0	1	1	0	6	0
2	carlos	7	8	2	1	2	1	0	1	0	1	0	0	0	0	0	1	0
3	carme	1	2	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	carmen	4	1	0	7	0	2	0	1	1	1	0	0	1	1	0	4	0
5	carol	1	2	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0
6	dolors	4	1	0	2	0	8	0	1	5	1	0	1	0	0	1	1	1
7	hugo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	joan	4	1	0	1	0	1	0	11	6	2	0	1	1	1	1	5	1
9	jose	1	0	0	1	0	5	0	6	11	1	0	3	0	0	3	1	3
10	julio	4	1	0	1	0	1	0	2	1	5	0	0	0	0	0	1	0
11	liliana	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	maria	0	0	0	0	0	1	0	1	3	0	0	3	0	0	1	0	1
13	mariano	1	0	0	1	0	0	0	1	0	0	0	0	4	1	0	4	0
14	martha	1	0	0	1	0	0	0	1	0	0	0	0	1	4	0	4	0
15	nuria	0	0	0	0	0	1	0	1	3	0	0	1	0	0	3	0	1
16	pau	6	1	0	4	0	1	0	5	1	1	0	0	4	4	0	12	0
17	xavi	0	0	0	0	0	1	0	1	3	0	0	1	0	0	1	0	3

Algo similar observamos al analizar la línea horizontal (o vertical):

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
		an	ca	ca	ca	ca	do	hu	jo	jo	ju	li	ma	ma	ma	nu	pa	xa
1	andres	15	7	1	4	1	4	0	4	1	4	0	0	1	1	0	6	0
2	carlos	7	8	2	1	2	1	0	1	0	1	0	0	0	0	0	1	0
3	carme	1	2	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	carmen	4	1	0	7	0	2	0	1	1	1	0	0	1	1	0	4	0
5	carol	1	2	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0
6	dolors	4	1	0	2	0	8	0	1	5	1	0	1	0	0	1	1	1
7	hugo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	joan	4	1	0	1	0	1	0	11	6	2	0	1	1	1	1	5	1
9	jose	1	0	0	1	0	5	0	6	11	1	0	3	0	0	3	1	3
10	julio	4	1	0	1	0	1	0	2	1	5	0	0	0	0	0	1	0
11	liliana	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	maria	0	0	0	0	0	1	0	1	3	0	0	3	0	0	1	0	1
13	mariano	1	0	0	1	0	0	0	1	0	0	0	0	4	1	0	4	0
14	martha	1	0	0	1	0	0	0	1	0	0	0	0	1	4	0	4	0
15	nuria	0	0	0	0	0	1	0	1	3	0	0	1	0	0	3	0	1
16	pau	6	1	0	4	0	1	0	5	1	1	0	0	4	4	0	12	0

Aumenta el número de cliques que se comparten con el resto de integrantes de la red. Pero, siguiendo a Hanneman, eliminaremos los cliques de sólo tres integrantes, ya que nos dan poca información adicional (sólo hace falta conocer a un miembro para formar parte del grupo). Repetiremos el procedimiento, pero esta vez declararemos que el tamaño mínimo del grupo deberá ser de 4 nodos. Transcribo a continuación los resultados obtenidos:

1	andres carmen dolors pau
2	andres carmen joan pau
3	andres dolors joan jose
4	andres joan julio pau
5	jose maria nuria xavi

Sólo existen cinco sub.grupos que cumplan con las características solicitadas. Y una vez más Andrés es el actor más destacado, con cuatro participaciones. Pero también lo son Joan y Pau, con tres participaciones cada uno.

		1 1 1 1 1 1 1 1																
		1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7
		a	c	c	c	d	h	j	j	j	l	m	m	m	n	p	x	
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	andres	4	0	0	2	0	2	0	3	1	1	0	0	0	0	0	3	0
2	carlos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	carme	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	carmen	2	0	0	2	0	1	0	1	0	0	0	0	0	0	0	2	0
5	carol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	dolors	2	0	0	1	0	2	0	1	1	0	0	0	0	0	0	1	0
7	hugo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	joan	3	0	0	1	0	1	0	3	1	1	0	0	0	0	0	2	0
9	jose	1	0	0	0	0	1	0	1	2	0	0	1	0	0	1	0	1
10	julio	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0
11	liliana	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	maria	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	1
13	mariano	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	martha	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	nuria	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	1	0
16	pau	3	0	0	2	0	1	0	2	0	1	0	0	0	0	0	3	0
17	xavi	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	1	0

K-plex presenta grupos más “legítimos” en cierto sentido, ya que elimina la presencia de los intermediarios (tengo que conocer directamente a todos los miembros menos a dos, en este caso).

Al igual que en el caso de las medidas de centralidad, las medidas de agrupamiento son complementarias, y la combinación de ellas nos devuelve una imagen más ajustada de la estructura de la red.

Para terminar con este ejercicio, pegaremos los resultados obtenidos en el documento de Word.

EL ANÁLISIS DE REDES PERSONALES CON EGONET⁷

Existen dos perspectivas diferentes para el estudio de las redes sociales: la sociocéntrica y la egocéntrica. La perspectiva sociocéntrica suele identificar para cada contexto institucional una red diferente. Así, si utilizamos como generador de nombres, “señala por favor a quiénes elegirías como representantes de esta lista”, o “dime a quién pedirías prestado dinero”, la red de respondientes y personas nominadas será diferente en cada caso. Naturalmente, podemos recoger diferentes contextos y combinarlos, pero la perspectiva sigue siendo la misma: un conjunto definido de personas o actores que tienen relaciones entre sí por algún tipo de relación (o contexto institucional) previamente definido.

La perspectiva egocéntrica, en cambio, parte de un ego o punto focal y se interroga por *todas* las relaciones existentes. Naturalmente, de esta forma se recogen (o se deberían recoger) *todos los contextos institucionales* en los que se mueve ego, es decir, familia, vecinos, compañeros de estudio o trabajo, grupo religioso o cualquier institución relevante para su vida social. Estas instituciones varían según la cultura, naturalmente.

Ahora bien, ¿cómo recoger o aproximarse la mirada de interacciones que se producen alrededor de una persona? La respuesta a esta pregunta depende también de la perspectiva adoptada. En general podemos afirmar que las redes personales están constituidas por un centro de lazos fuertes que cambia lentamente a lo largo de la vida y una periferia más amplia y dinámica. El tamaño de esta red varía enormemente en la escala social, de forma que las personas con cargos políticos o administrativos de importancia suelen tener unas redes muy amplias comparadas con las redes de personas marginales en una sociedad. Si atendemos a las relaciones acumuladas a lo largo de una vida (muchas de ellas se olvidan) el resultado es de 3500 personas (*Cf.* Molina, 2005 para una explicación detallada de los métodos). Si atendemos a las relaciones *estables*, podemos identificar una media de unas 300 personas, con

⁷ Autor: José Luis Molina, Sarah Hoeksma y Sílvia Gómez i Mestres (Departament d'Antropologia social i cultural, UAB).

una desviación típica importante en función de la clase social y el sexo. Si atendemos a las relaciones fuertes e íntimas, el número se puede reducir a menos de una veintena. En el siguiente gráfico podemos apreciar estas diferencias.

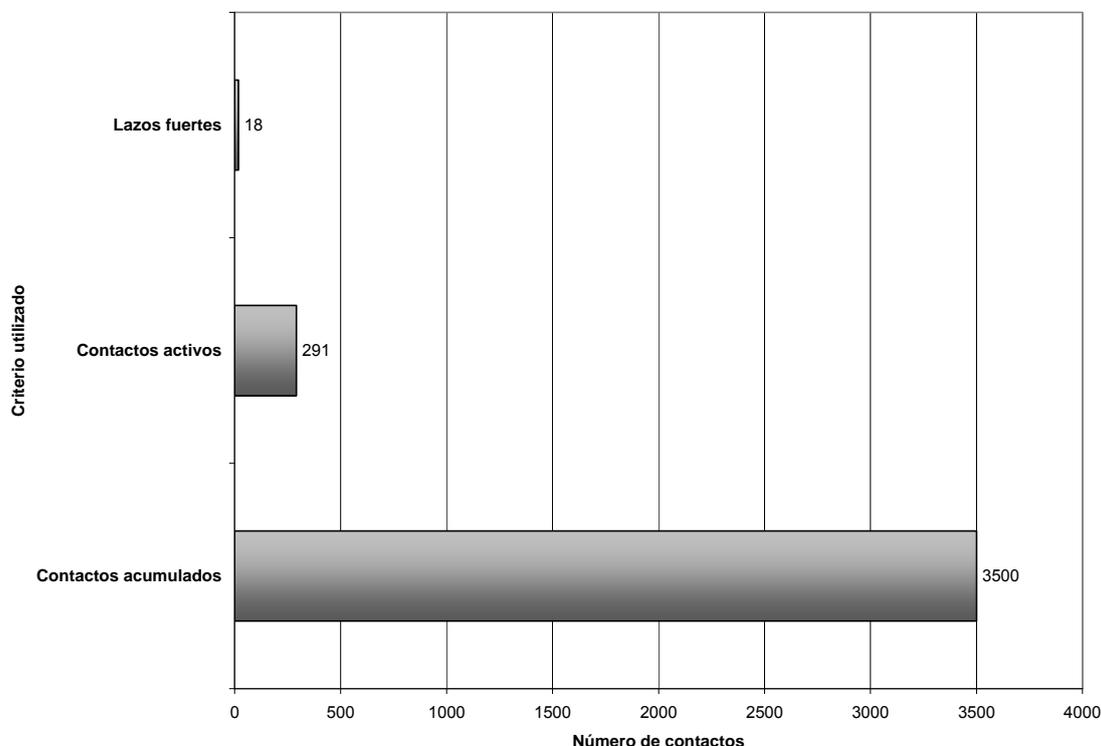


Ilustración 39. Las diferentes estimaciones del tamaño de la red personal

Christopher McCarty (2002) sugiere un cambio en el enfoque del estudio de redes personales, pasando de la estimación del tamaño a la estimación de la *estructura* de las redes personales. La pregunta es: *¿cuál es el mínimo de número de alteri sobre los que hay que indagar para disponer de una idea aproximada de la estructura de la red personal?* McCarty propone que un número entre 30 y 45 personas tiene que ser suficiente si el generador de nombres utilizado permite obtener una lista libre de personas de todas las categorías sociales. Por ejemplo:

Por favor, díganos una lista de 40 personas que Ud. conozca. Intente que todas las categorías importantes de personas (parientes, amigos, compañeros, vecinos...) estén representados en la lista.

El generador utilizado tiende a producir lazos íntimos al principio, pero la lista es lo suficientemente larga como para recoger información de diferentes áreas

de la estructura de las redes personales, incluidos los lazos débiles. A continuación es necesario indagar sobre quién conoce a quién, aparte de recoger datos atributivos sobre ego y los *alteri*. Esta parte es la más larga pues una lista de 40 personas implica 780 preguntas de si dos personas se conocen o tienen algún tipo de relación que interese recoger. Para obtener esta información es necesaria la ayuda del programa Egonet. Por último, se presenta un informe a ego sobre su propia red personal. En todos los casos, los informantes se han mostrado sorprendidos por la calidad de la información mostrada sobre sus propias redes.

El estudio de las redes personales tiene la inmensa ventaja que permite aplicar los métodos de muestreo de poblaciones (a diferencia del enfoque sociocéntrico) y que los resultados acerca de la composición y la estructura de cada red personal pueden ser convertidos en variables atributivas de los casos y formar parte de los modelos explicativos sobre una variedad de temas (salud, discursos, rendimiento académico, aculturación ...).

En este ejercicio realizaremos tres actividades. La primera, con el módulo de administración de Egonet, diseñaremos un cuestionario que nos permita aproximarnos a nuestra propia red personal. La segunda actividad consistirá en recoger datos de nuestra red personal (y, en su caso, la de otras personas) y visualizarla y analizarla con Visualizer, el módulo de visualización incorporado en Egonet. La tercera actividad consistirá en exportar los datos recogidos a SPSS.

NOTAS SOBRE LA INSTALACIÓN DE EGONET

Puede descargarse Egonet de la página <http://sourceforge.net/projects/egonet/>. Es posible descargar un fichero comprimido con la versión java multiplataforma (*.jar) o una versión para Windows XP (*.exe). Cada fichero comprimido contiene los dos módulos de **EgoNet**, **EgoNet Author** y **EgoNet Client**. Una vez descomprimida una de las dos versiones se disponen de dos programas ejecutables en un subdirectorío. Los diferentes estudios, las diferentes entrevistas y estadísticas se crearán en subdirectoríos automáticamente creados dentro de ese subdirectorío.

Es importante recordar que solamente es posible visualizar o calcular estadísticas de las entrevistas creadas con un estudio determinado. Si éste se modifica no se podrá acceder a las entrevistas realizadas con la versión anterior.

MÓDULO ADMINISTRADOR (EGONET AUTHOR)

El módulo Administrador de Egonet se llama **Egonet Author** (administrador de cuestionarios). Una vez clicamos el icono aparece la siguiente pantalla:

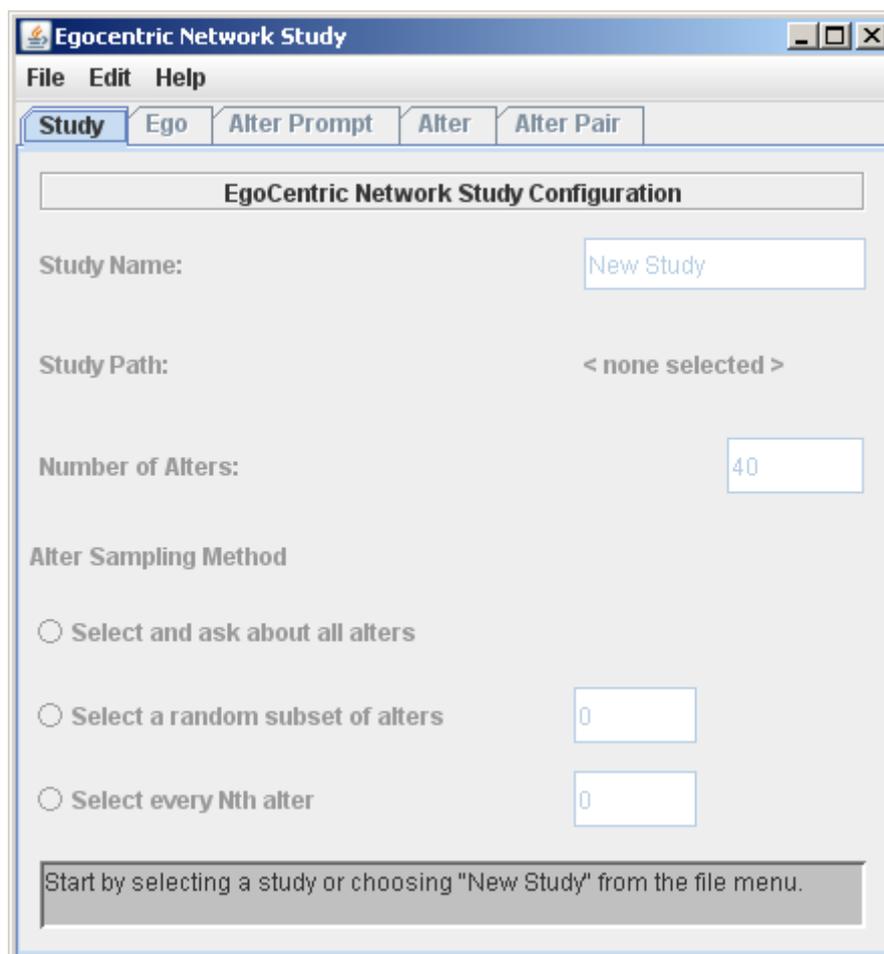


Ilustración 40. Pantalla de identificación de Egonet Author

A continuación crearemos un nuevo estudio para estudiar, por ejemplo, redes transnacionales.

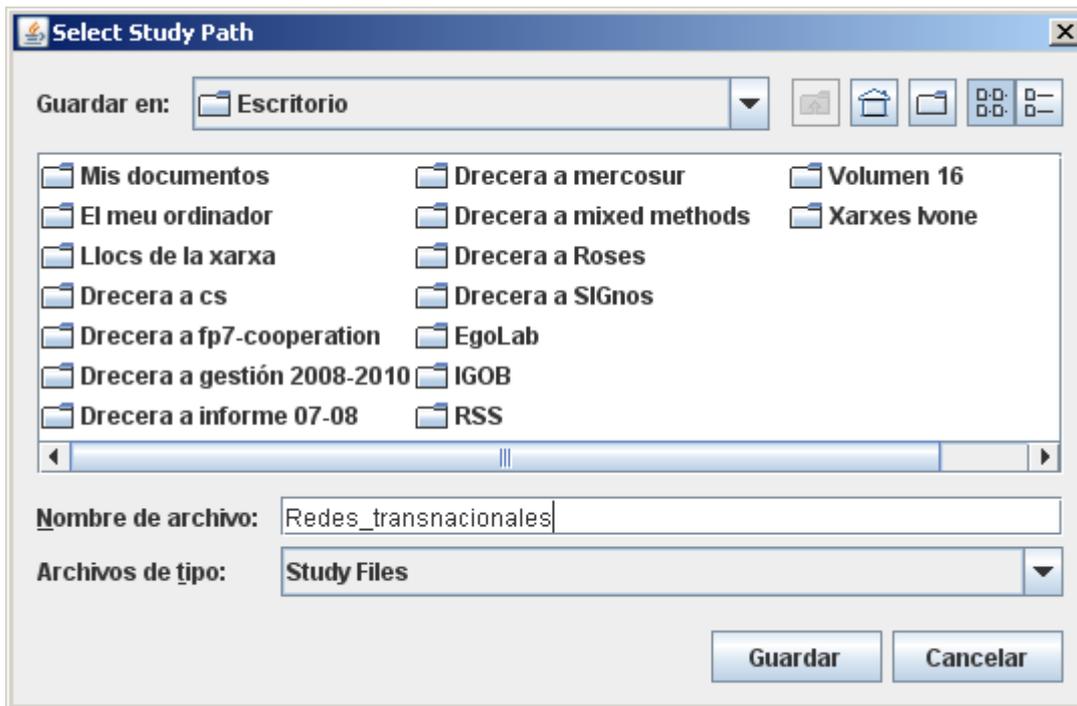


Ilustración 41. Creación de un nuevo estudio

A continuación pondremos 30 en el número de alteri que pediremos a cada ego (Number of Alters). Evidentemente, el número de alteri puede ser cualquier otro.

Los módulos que componen **EgoNet Author** son “Ego”, las variables identificativas de Ego, “Alter Prompt”, la pregunta o preguntas que utilizaremos para generar nombres o nominaciones de personas, “Alters”, las variables que pediremos para cada una de las personas nominadas y “Alter Pair”, la pregunta que utilizaremos para saber si dos personas nominadas tienen algún tipo de relación entre sí. Veamos cada uno de los módulos.

Ego

Ya estamos en condiciones de crear preguntas para cada Ego. La información Nombre y Apellido aparece automáticamente cuando se ejecuta **EgoNet**, por lo que no es necesario introducir estas preguntas. Clicaremos el botón **New** e introduciremos la primera pregunta, la edad de Ego.

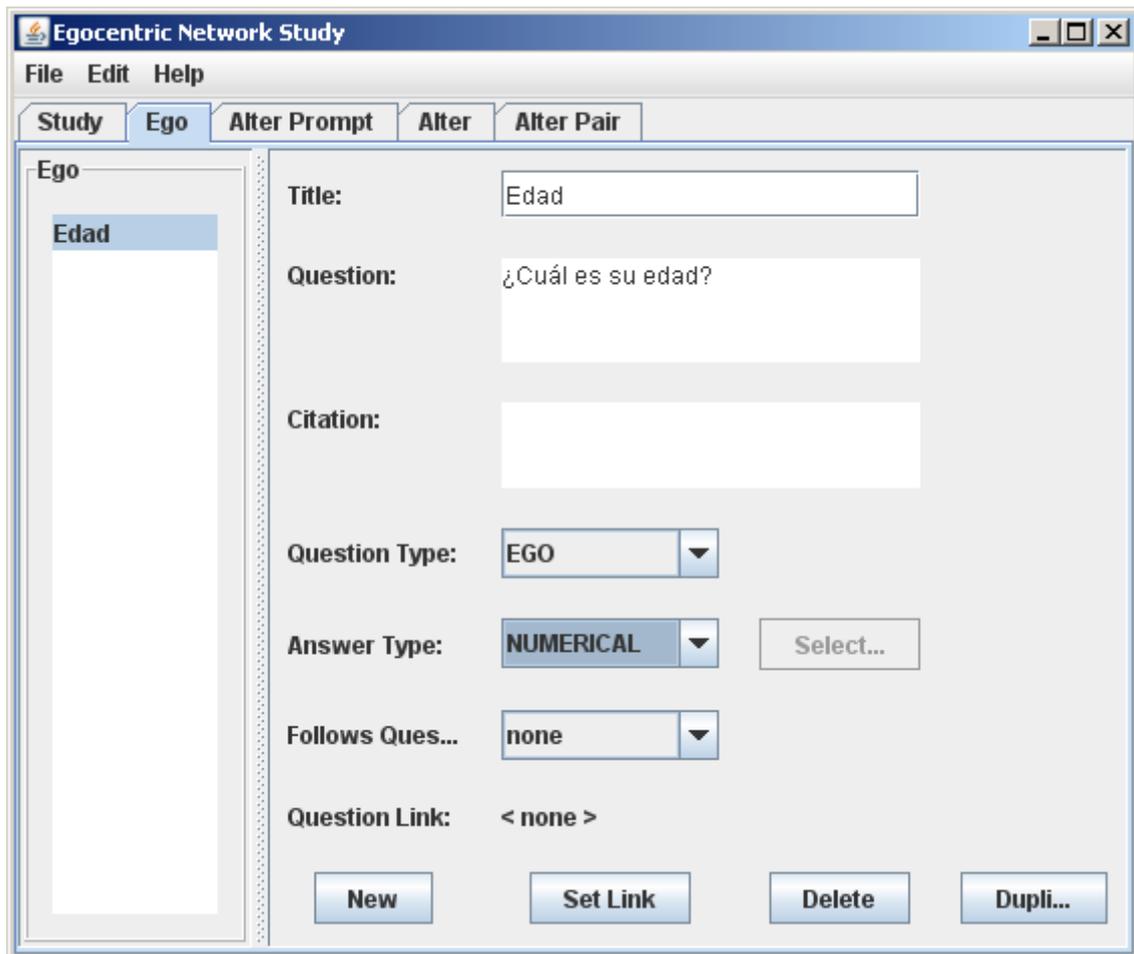


Ilustración 3. Variable atributiva de Ego

Observamos que en el campo Question se introduce el literal de la pregunta, en el campo Title, una cadena de caracteres y en Answer Type hemos seleccionado “Numerical” para poder introducir directamente los años como un número entero. Siempre estaremos a tiempo después de convertir la edad en categorías en SPSS o R.

Si clicamos el botón **New** de nuevo, podremos introducir una nueva variable, el sexo en este caso. Para ello introducimos un literal (“¿Es Ud. hombre o mujer?”), el nombre de la variable (“sexo”) y seleccionamos el botón **Selections** (si en Answer type figura la opción “Categorical”). Al hacerlo se nos abre una ventana de gestión de categorías.

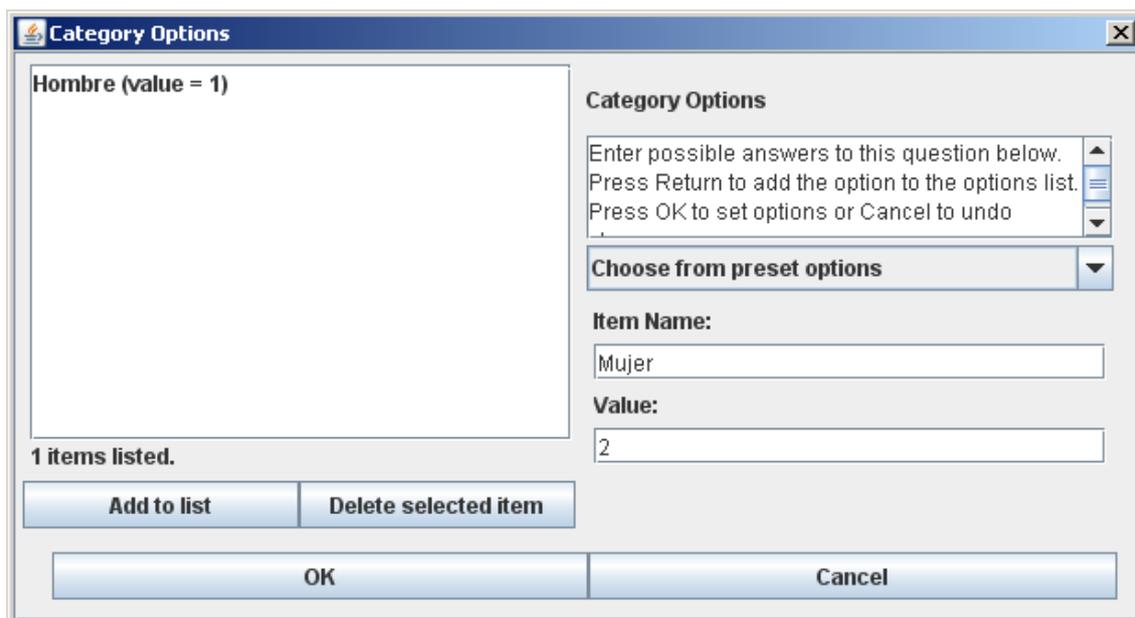


Ilustración 4. Gestor de categorías

Para crear una nueva categoría ponemos en Item Name el literal y en Value el número correspondiente y clicamos **Add to list** (botón izquierdo de la pantalla). Clicamos OK y ya tenemos añadida una nueva variable categórica para Ego, volviendo a la pantalla anterior.

Para facilitar el uso de categorías ya creadas, EgoNet permite duplicar una pregunta clicando en el botón **Duplicate**, con el objeto de que se pueda editar a continuación, haciendo los cambios necesarios.

También es posible establecer un orden de flujo de las preguntas señalando en cada una de ellas la pregunta inmediatamente anterior con el botón **Set Link**.

Para *borrar* una variable, simplemente nos colocamos en el panel izquierdo, seleccionamos la variable con el cursor y clicamos **Delete** en la parte derecha de la pantalla.

ALTER PROMPT

En el siguiente módulo clicaremos el botón **New** e introduciremos “Generador de nombres” en el campo Title y la expresión “Por favor, escriba una lista de 30 personas que conozca por su nombre de pila y viceversa y con las que haya tenido al menos algún contacto personal en los dos últimos años y con la que sea posible contactar de nuevo si fuese necesario. Intente que todas las categorías de contactos estén representados”.

Ya disponemos de una pregunta generadora de nombres. Podemos añadir otras clicando de nuevo el botón **New**.

Siempre es útil trabajar con acrónimos, por ejemplo MigCer para Miguel de Cervantes. De esta forma las personas que responden el cuestionario pueden recordar fácilmente a quién corresponde el acrónimo y se establece una primera medida de confidencialidad.

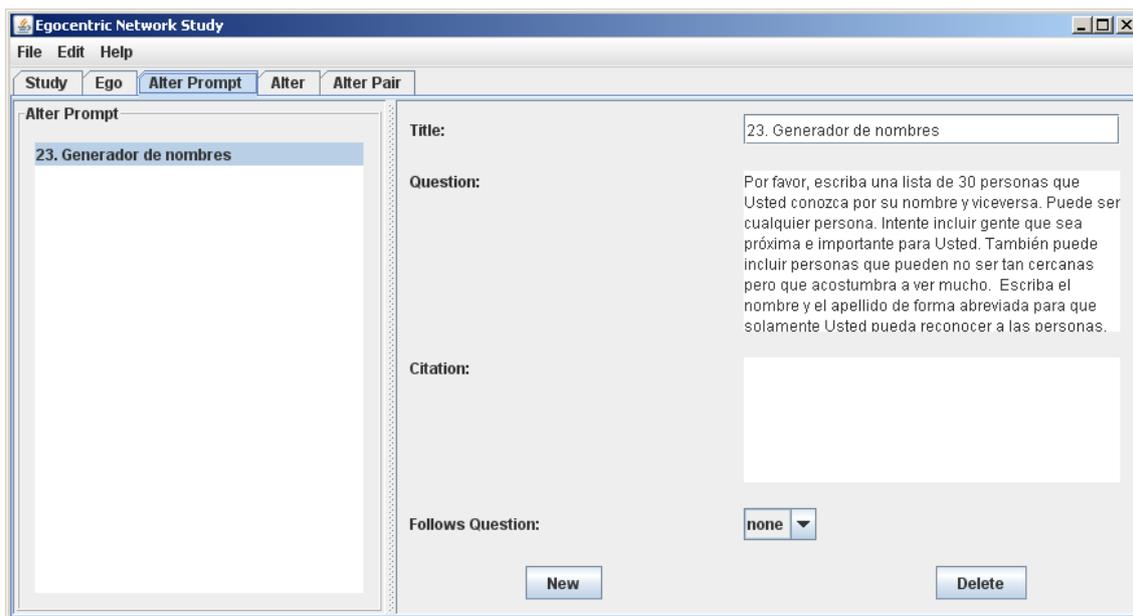


Ilustración 5. Alter Prompt. Este módulo permite introducir generadores de nombres.

ALTER

En este módulo introduciremos variables atributivas del mismo modo que el módulo para Ego. En este caso introduciremos también la *edad*, el *sexo* y la *categoría ocupacional*, por ejemplo. Otras variables, como la *nacionalidad* o la *ciudad de residencia habitual*, exigirían bien un listado de países adaptado a los respondientes, bien un espacio en blanco para la introducción de un literal. Es importante limitar el número de variables de este módulo, pues cada variable tendrá que ser contestada 30 veces (una vez por cada alter o persona nominada).

Una vez creadas las preguntas el aspecto debería ser éste:

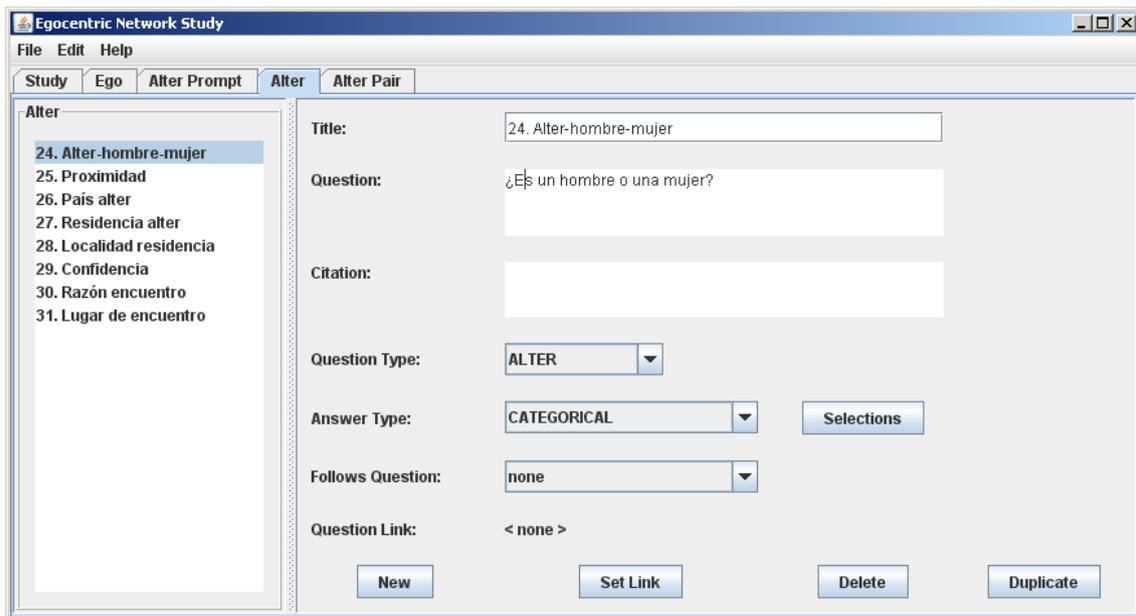


Ilustración 6. Preguntas del módulo Alter

Como en el módulo “Ego”, disponemos de las opciones **Set Link Delete** y **Duplicate**.

Y ahora nos resta describir el módulo final.

ALTER PAIR

En este módulo introduciremos una sola pregunta para relacionar los alteri entre sí. Esta pregunta es la siguiente:

- “¿Estas personas se relacionarían entre sí Ud no estuviese?” (categórica, “Muy probablemente (2), probablemente (1), no es probable (0)”).

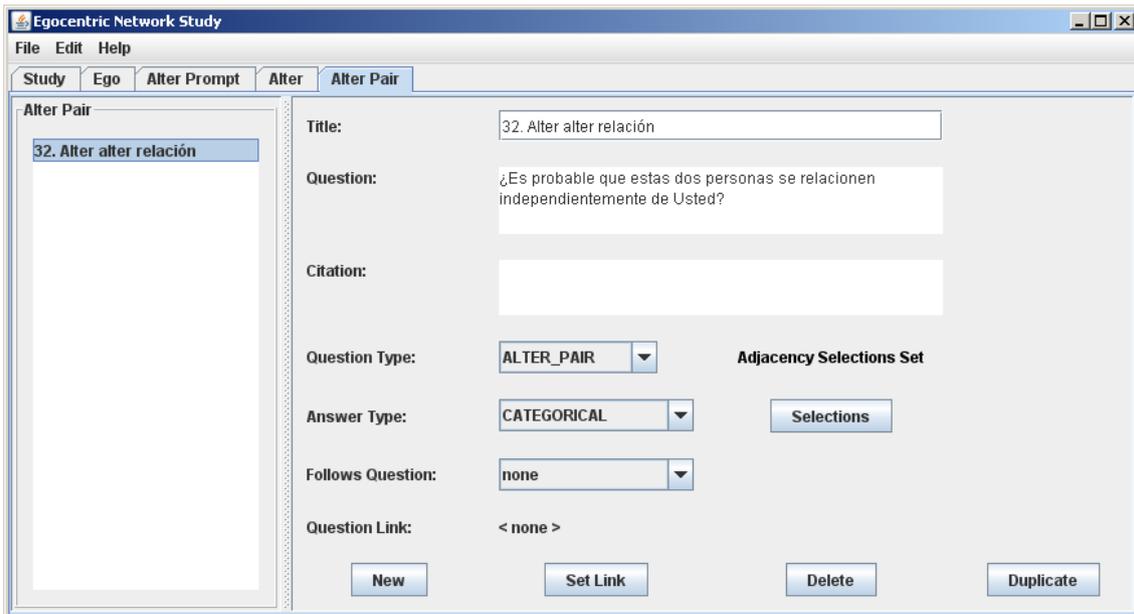


Ilustración 7. Preguntas del módulo Alter

Ahora tenemos que indicar cuáles de las respuestas previstas a la pregunta crea una relación entre dos Alteri. Esto lo indicaremos posicionándonos en las respuestas que crean la relación y clicando el botón “Mark selected item adjacent”).

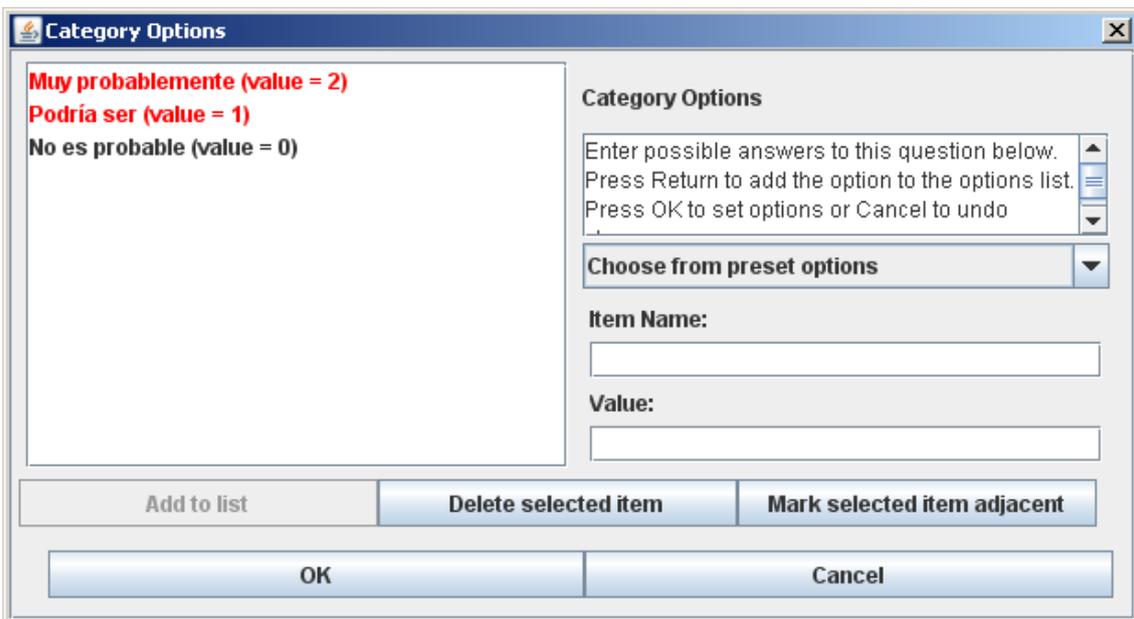


Ilustración 8. Preguntas del módulo Alter

De esta forma, siempre que se seleccionen una de estas dos respuestas se establecerá un vínculo entre el par de Alteri seleccionados.

Clicando OK volvemos a la pantalla anterior.

Una vez completado el cuestionario guardaremos el estudio (File>Save Study) con el nombre remesas, por ejemplo. Con la opción File>Save Study as podemos cambiar el nombre. Los estudios diseñados con **EgoNet Author** tienen la extensión *.ego.

Con File>Exit saldremos del módulo de administración de Egonet con nuestro cuestionario listo para ser utilizado.

MÓDULO CLIENTE (EGONET CLIENT)

La pantalla que obtenemos cuando ejecutamos **Egonet Client** es la siguiente:

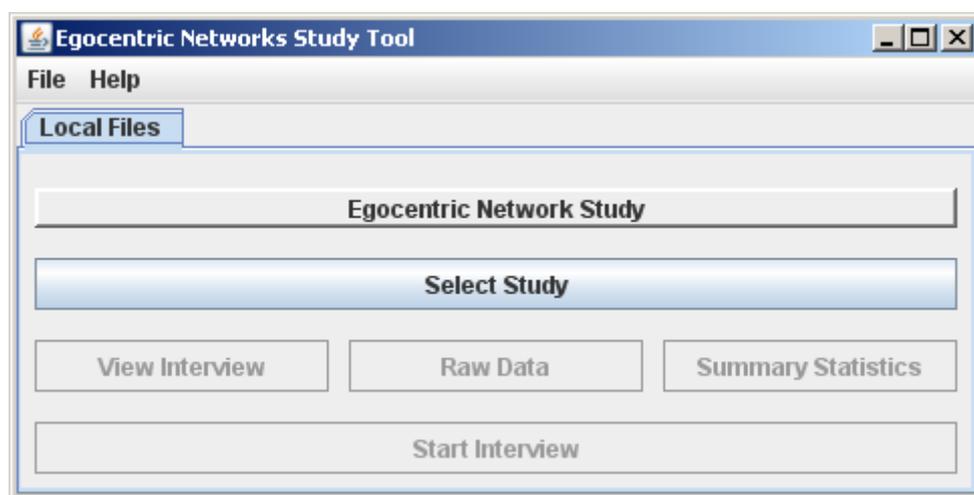


Ilustración 48. Pantalla principal de EgoNet cliente

En esta pantalla podemos ver diferentes botones.

Select Study nos permite seleccionar el estudio que queremos aplicar. Una vez seleccionado un estudio se activan en resto de botones.

Con **Start Interview** podemos iniciar una entrevista con el estudio seleccionado. Una vez seleccionada esta opción aparece la pantalla siguiente:

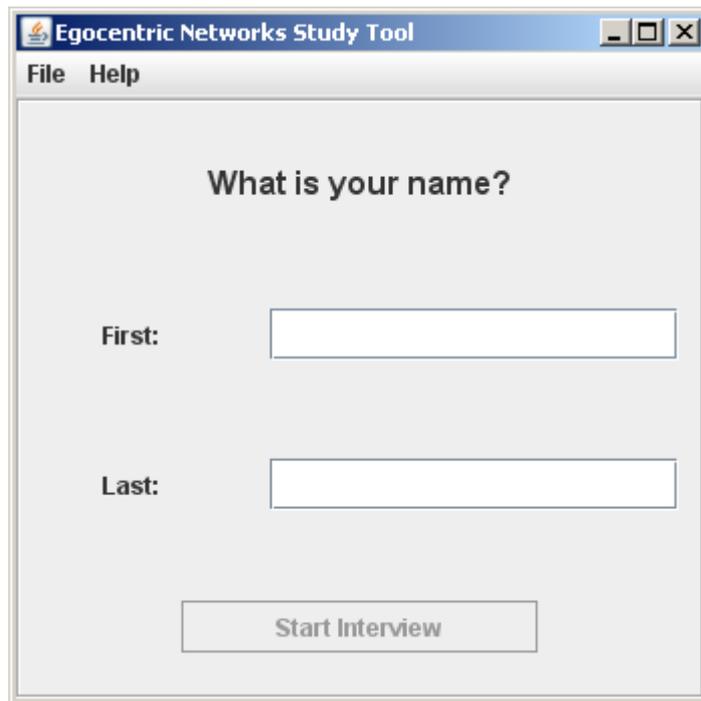


Ilustración 49. Pantalla de identificación de Egonet (módulo cliente)

El nombre y el apellido pueden ser auténticos, pero siempre es recomendable la utilización de acrónimos. Introducimos, por ejemplo, “Juan” y “Ruiz” en cada campo (**First** y **Last**). Una vez introducido el Nombre y el Apellido o uno de los dos, se activa el botón **Start Interview** y aparece la primera pregunta que habíamos previsto para Ego.

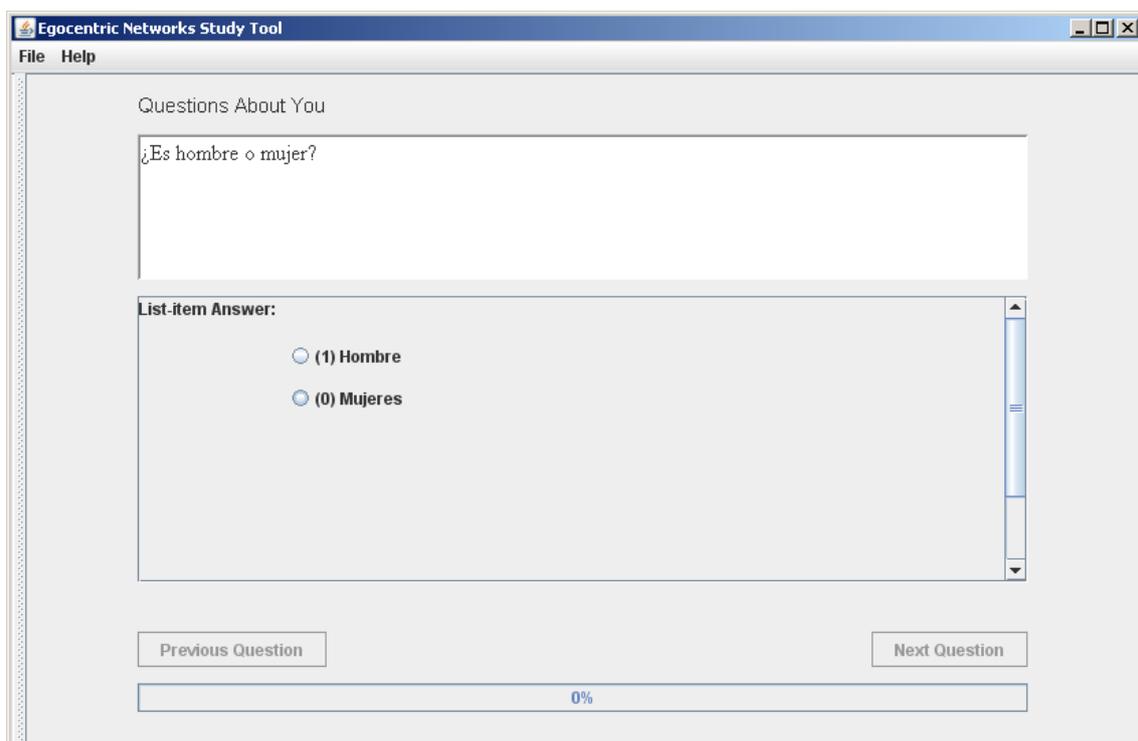


Ilustración 50. Módulo con las preguntas para Ego.

Si volvemos a clicar el botón **Next Question** obtenemos la siguiente pantalla con las preguntas previstas para Ego y así sucesivamente hasta que entramos en el siguiente módulo.

En el módulo **Alter Prompt** podemos ir introduciendo la lista de personas que sugiere el generador de nombres que hemos introducido entrando First Name y/o Last Name y la tecla **Return** o bien **Add to list**. Una vez la lista esté completa será posible continuar con el siguiente módulo.

Whom do you know?

Por favor, escriba una lista de 30 personas que Usted conozca por su nombre y viceversa. Puede ser cualquier persona. Intente incluir gente que sea próxima e importante para Usted. También puede incluir personas que pueden no ser tan cercanas pero que acostumbra a ver mucho. Escriba el nombre y el apellido de forma abreviada para que solamente Usted pueda reconocer a las personas. Es importante que no abrevie demasiado para poder reconocerlas más tarde. Por ejemplo: Mig Cervan por "Miguel de Cervantes"

Juan Salvador (value = -1)
Ramón Valdés (value = -1)

Your Acquaintances

Enter the names of 30 people. After entering 30 names you can continue.

First Name:
Laura

Last Name:
Rota

2 items listed.

Add to list Delete selected item

Previous Question Next Question

3%

Ilustración 51. Módulo con las preguntas generadoras de nombres.

Ahora viene la parte más larga de la entrevista, la atribución de variables a los Alteri. Cada una de las preguntas que hemos previsto para los Alteri será realizada para cada persona nominada.

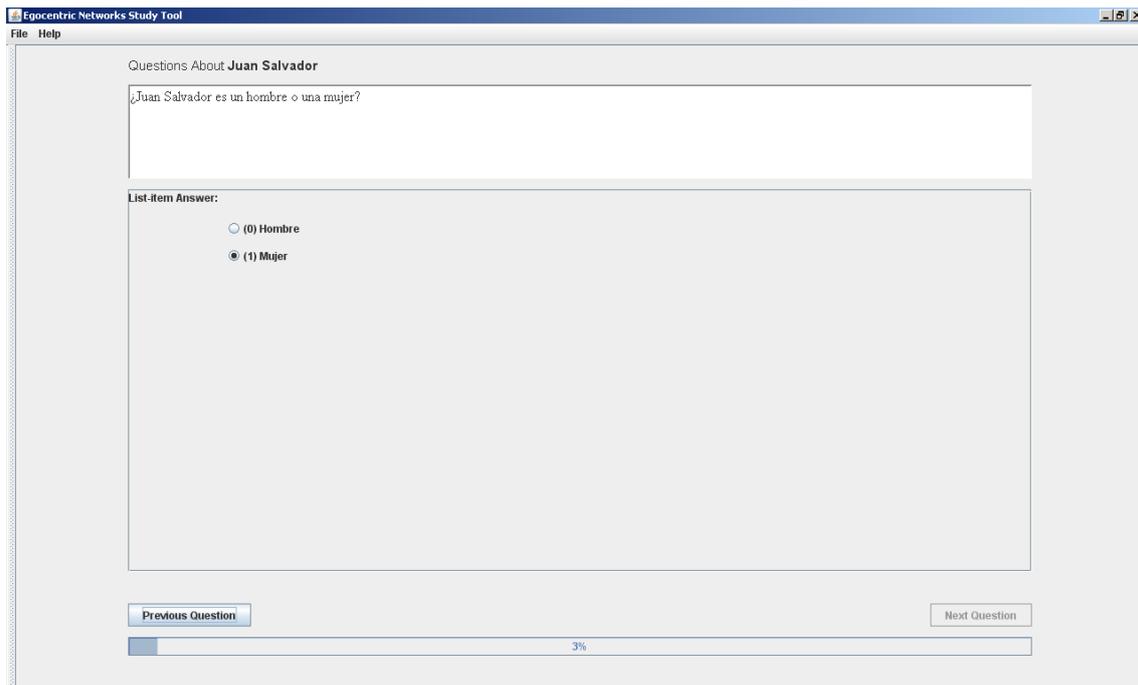


Ilustración 52. Módulo con las preguntas sobre los Alteri.

Una vez cumplimentada esta parte nos aparece el último módulo de relación, indicando cada pareja de nominados:

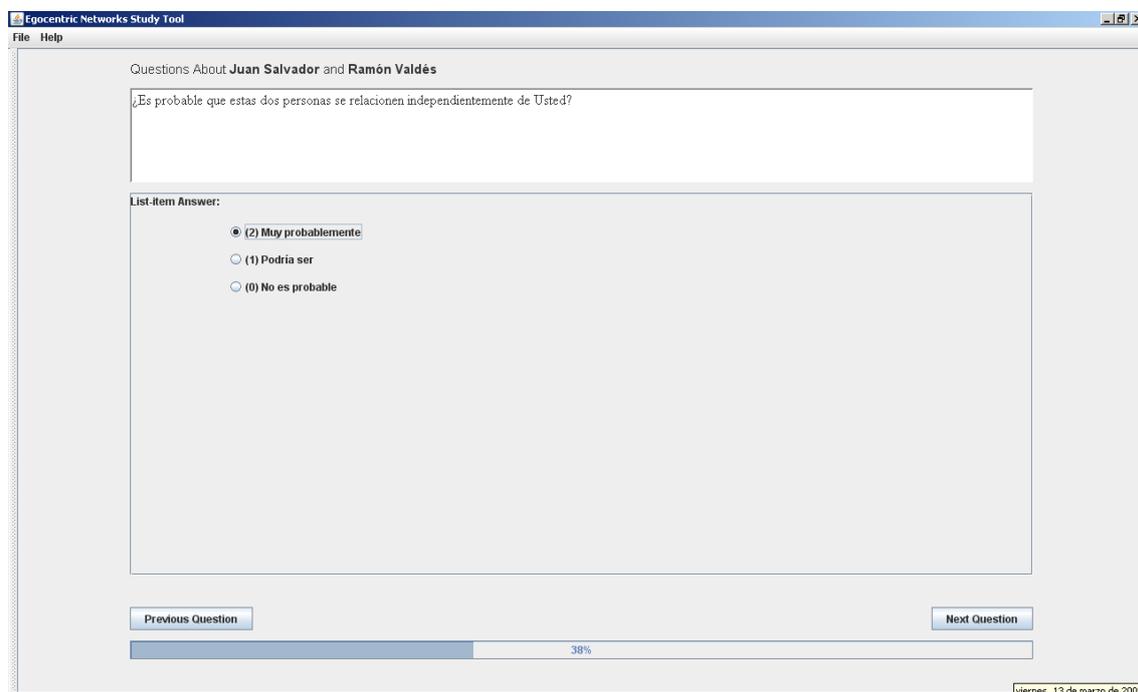


Ilustración 53. Módulo cuarto para la relación entre las personas nominadas

Se observará que es posible clicar la opción adecuada o bien teclear el número, lo cual suele ser más rápido.

Una vez completado este módulo, aparece el mensaje Completed Interview y volvemos a la pantalla principal de **EgoNet Client**. Clicando la opción **View Interview** podemos seleccionar en el subdirectorio creado por **Egonet Client** llamado Interviews, el fichero correspondiente a la persona deseada. Los ficheros con las entrevistas tienen la extensión *.int.

Después de un mensaje de aviso el ordenador calcula el gráfico resultante y nos muestra la pestaña Interview en la que es posible corregir algún valor equivocado, , que nos presenta un análisis de las características de Ego y Graph un gráfico con la red personal de la persona entrevistada.

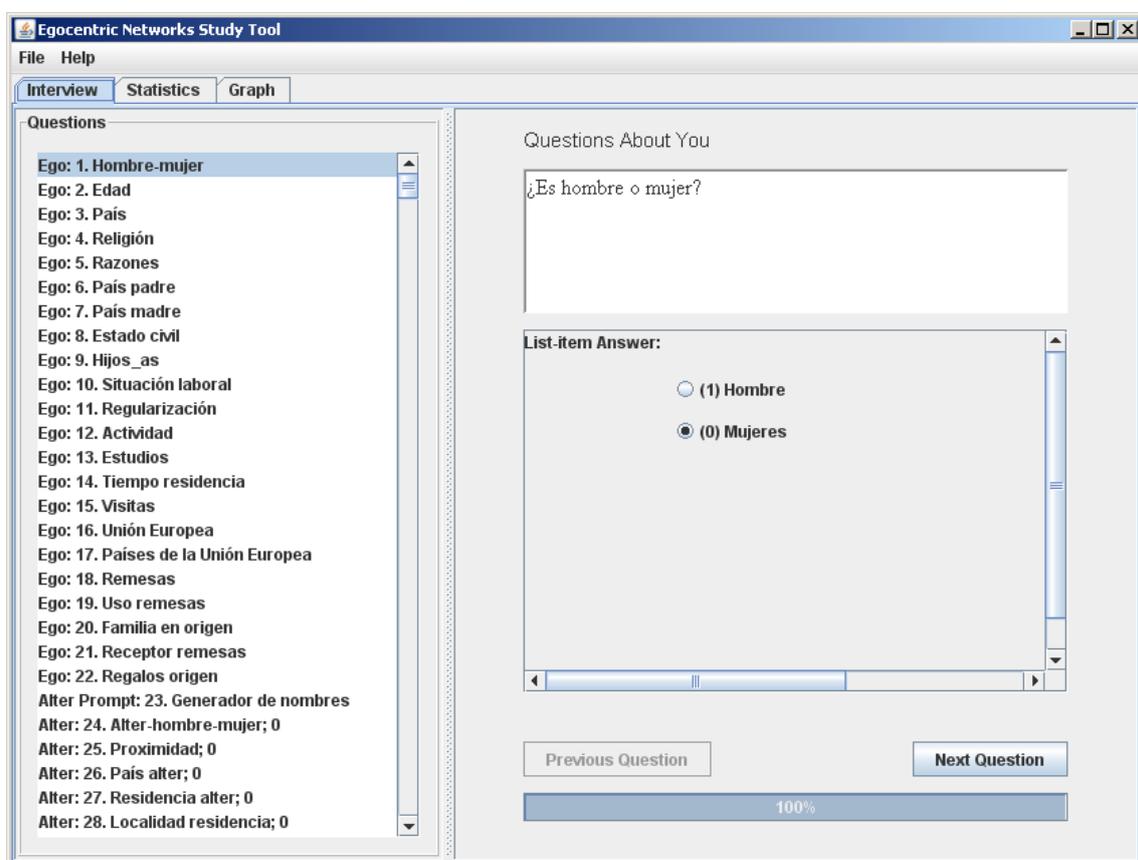


Ilustración 54. Visualización de una entrevista.

En la pestaña Statistics disponemos de un conjunto de informaciones sobre la red personal de la persona entrevistada, como el grado de intermediación de los Alteri de Ego o el número de cliques (o grupos en el que todos los Alteri tienen relaciones entre sí) o componentes (subgrafos completamente conectados de la red personal, es decir, que es posible encontrar un camino que una a cualesquiera par de nodos del subgrafo que forma el componente).

The screenshot shows the 'Egocentric Networks Study Tool' window with the 'Statistics' tab selected. The interface includes a menu bar (File, Help) and sub-tabs for 'Interview', 'Statistics', and 'Graph'. Under 'Statistics', there are sub-tabs for 'Degree Centrality', 'Closeness Centrality', 'Betweenness Centrality', 'Cliques', and 'Components'. The main area is divided into 'Structural Measures' and 'Compositional Summary'. The data table below shows the following values:

Structural Measures		Compositional Summary	
Degree Centrality Maximum	Thelma madre		20
Closeness Centrality Minimum	Thelma madre		23.577
Betweenness Centrality Maximum	Thelma madre		120.2
Number of Cliques	6		
Number of Components	2		

Ilustración 55. Estadísticas de una entrevista.

Por último, la pestaña Graph nos permite visualizar la red personal de la persona entrevistada y asignar variables visuales a partir de las variables recogidas de los Alteri, de forma que es posible explorar las diferentes combinaciones de variables de composición (características de los Alteri) y estructura (estructura de relaciones de los Alteri).

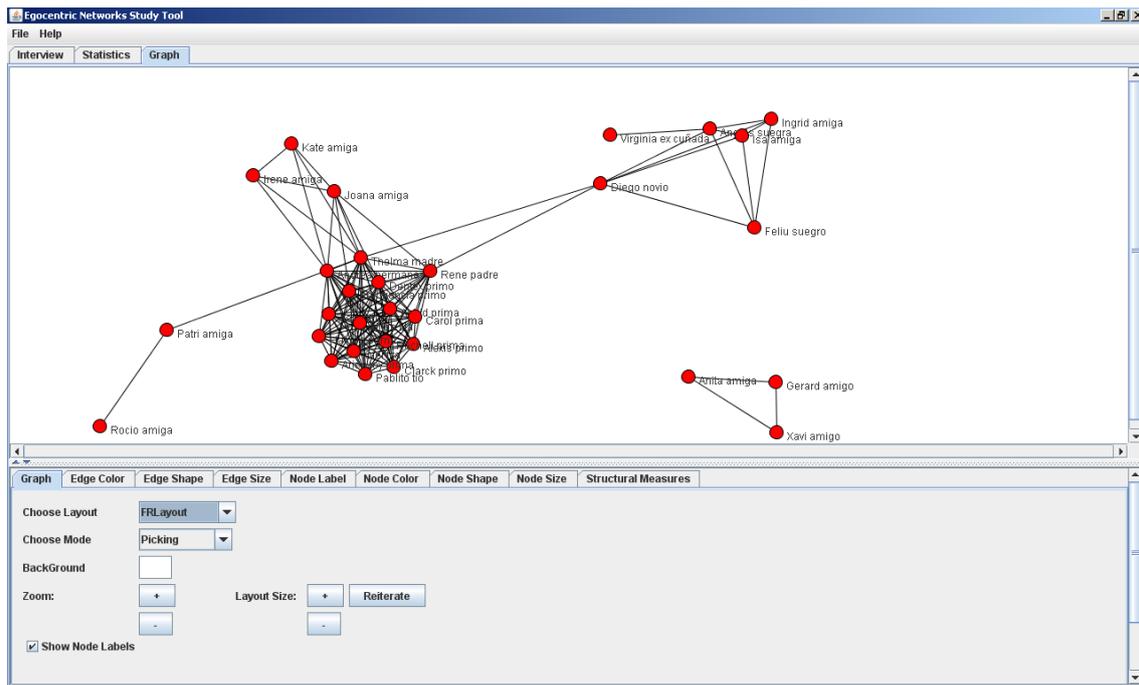


Ilustración 56. Grafo de una red personal con dos componentes.

Esta visualización es utilizada para realizar una entrevista con Ego y explorar conjuntamente su mundo social. Veámoslo con detalle en el siguiente apartado.

LA ENTREVISTA CUALITATIVA

La visualización de la red personal es un recurso extraordinario para indagar conjuntamente con el informante acerca del significado de los diferentes grupos, posición de personas o distribución de variables de su red personal. En este apartado mostramos un ejemplo tomado de un proyecto de investigación que ilustra las posibilidades que ofrece el módulo de visualización de Egonet. Estas entrevistas pueden grabarse utilizando el mismo ordenador con *Windows Media Encoder*, por ejemplo.

Para realizar esta explicación se ha diseñado una red personal con pocos Alteri. Es importante recordar que situando el ratón sobre cada nodo se puede acceder a toda la información introducida sobre el nodo.

Se puede seleccionar un nodo (o varios) y desplazarlo hacia donde queramos. Para hacerlo hay que escoger en Choose Mode> Picking y seleccionar a continuación los nodos que nos interese mover.

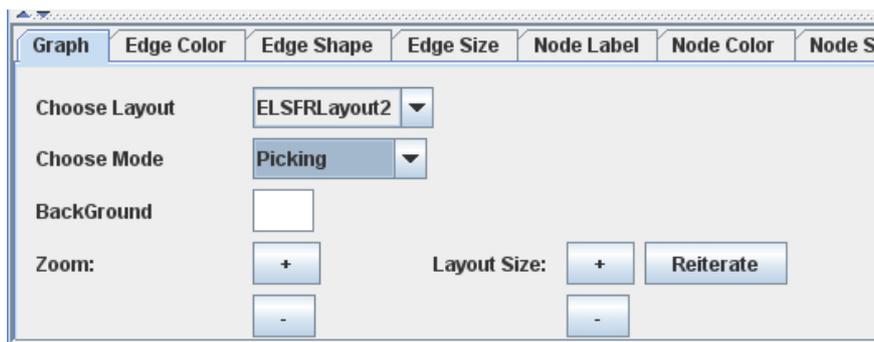


Ilustración 57. Desplazamiento de los nodos del grafo.

Una vez seleccionados, los nodos cambian presentan sus etiquetas de color azul. Para moverlos hay que clicar sobre uno de los nodos seleccionados y arrastrarlo. Si queremos volver a ver la primera visualización que nos presentó EgoNet seleccionaremos Choose Layout> ELSFRLayout2.

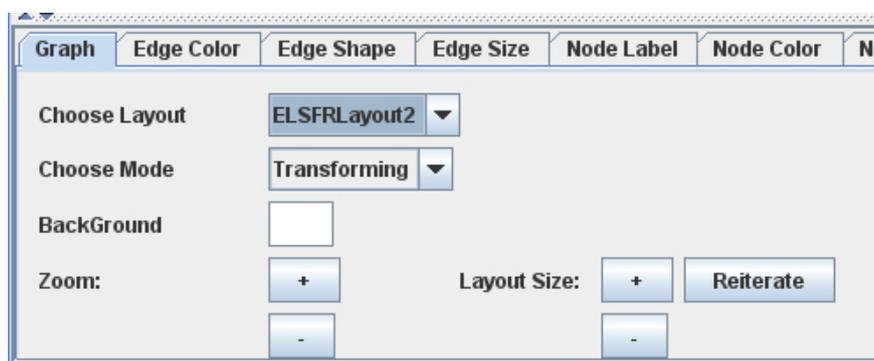


Ilustración 58. Visualización de una entrevista.

Aunque cada investigación tiene objetivos diferentes, exponemos a continuación un conjunto de visualizaciones que pueden ser utilizadas en diferentes casos. Cada visualización es una combinación de diferentes variables visuales por una parte (forma, color, etiqueta, tamaño) que presentan variables de composición (sexo, país de origen, cercanía percibida, por ejemplo) y estructura (betweenness, degree...). Las diferentes combinaciones las denominamos “Visualización 1”, “Visualización 2” y “Visualización 3”.

	Visualización 1	Visualización 2	Visualización 3
Forma	Sexo	Sexo	Sexo
Color	País Origen	Betweenness	Confidencia
Etiqueta	Alter	Alter	Razón Encuentro
Tamaño	Proximidad	Centralidad	Centralidad

VISUALIZACIÓN

Para definir la Forma realizaremos la siguiente secuencia de comandos: Node Shape> Choose question to shape: Alter-hombre-mujer> Hombre: triangle, mujer: circle> Apply Size.

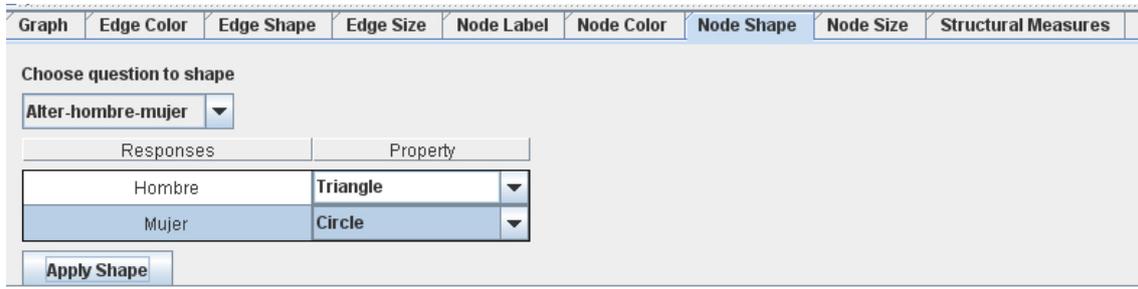


Ilustración 59. Aplicación de la forma.

Esta caracterización la mantendremos durante toda la entrevista.

Para definir el Color seleccionamos Node Color> Choose question to color: País Alter> Apply Color.

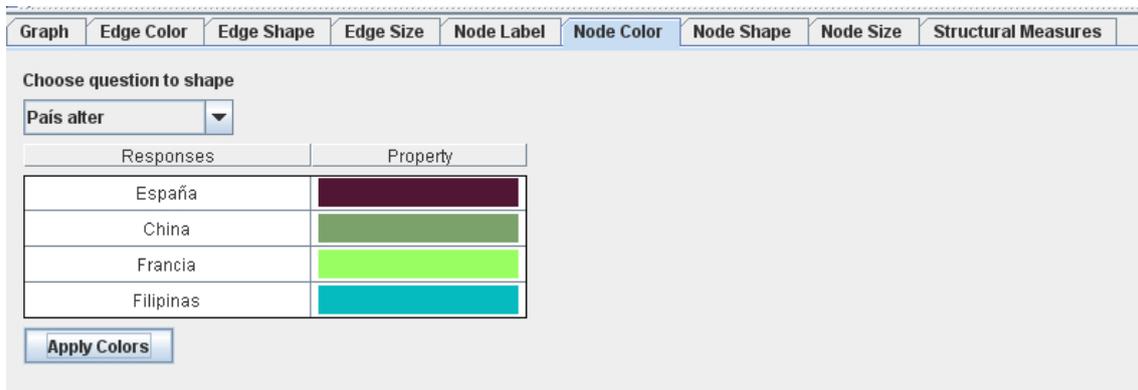


Ilustración 60. Aplicación del color.

Por ahora utilizaremos la Etiqueta que aparece por defecto..

Para aplicar el Tamaño seleccionamos Node Size> Choose question to size: Proximidad> Apply Size.

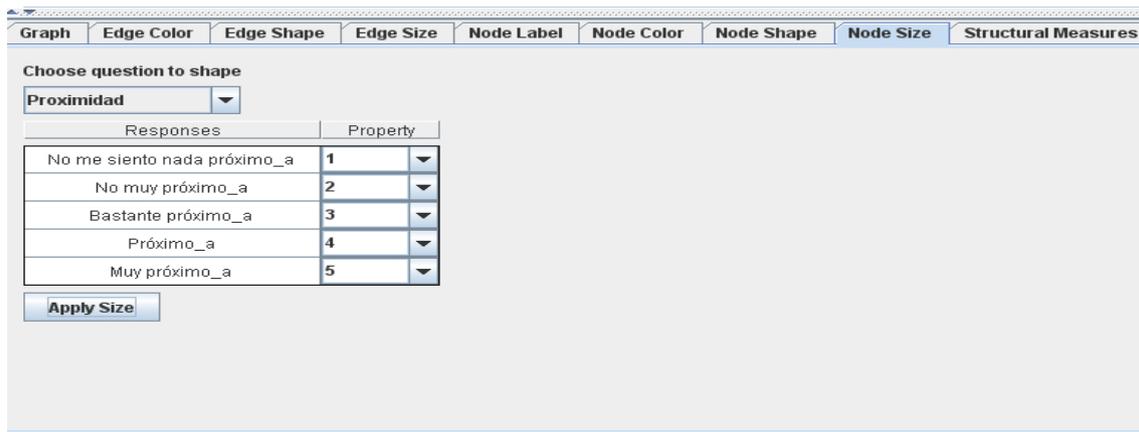


Ilustración 61. Aplicación del tamaño.

El resultado es el siguiente:

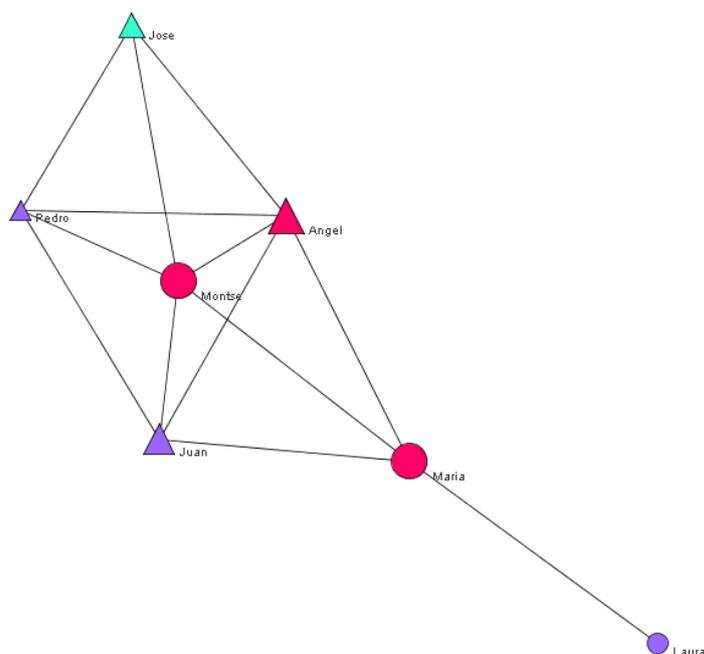


Ilustración 62. Visualización 1.

Esta visualización nos permite realizar interesantes preguntas a la persona entrevistada, como por ejemplo por qué Laura solamente conoce a Marta o quiénes son los que forman el grupo tan interrelacionado que aparece arriba a la izquierda.

Una vez obtenida la información es posible pasar a la segunda visualización.

En este caso, la **Forma** será la misma utilizada en la visualización anterior, al igual que la **Etiqueta** (el nombre de cada Alter)

Para el **Color** y **Tamaño** seleccionaremos Structural Measures> Size nodes > DegreeCentrality>Apply Size; Color nodes based on> BetweennessCentrality> Apply Color.

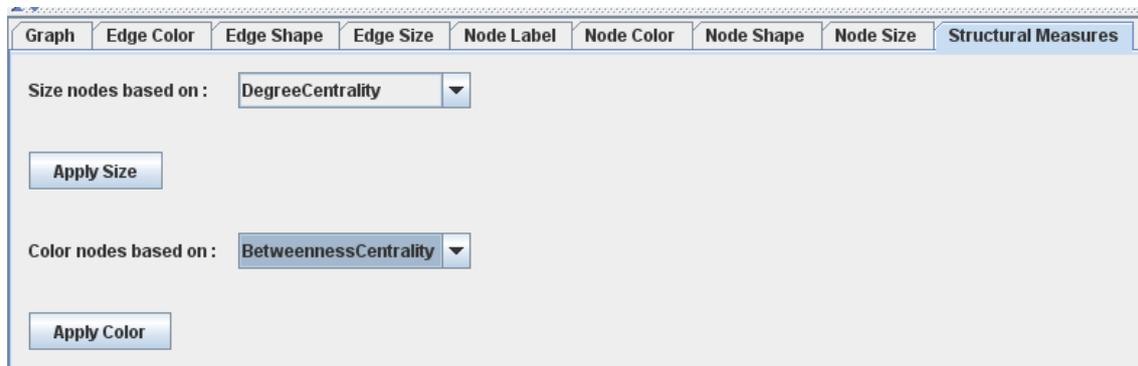


Ilustración 63. Modificando el Tamaño (Size) y el Color para la visualización 2.

El resultado es el siguiente:

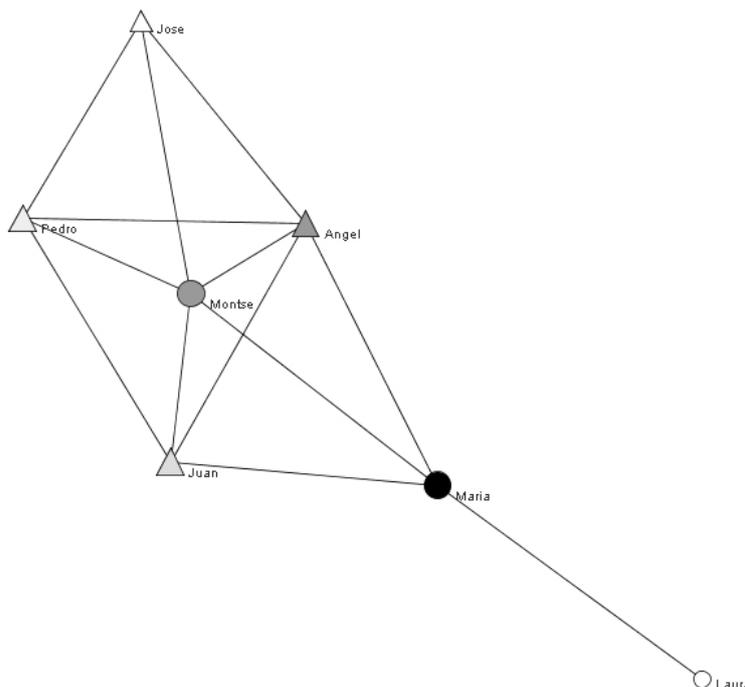


Ilustración 63. Visualización 2.

Para interpretarlo hay que tener en cuenta que el tamaño nos indica la centralidad de las personas en la red y el color las personas que unen diferentes grupos (cuanto más oscuras más grado de intermediación).

Por último, para la visualización 3 aplicaremos el siguiente esquema:

Mantendremos la **Forma** y el Tamaño y modificaremos el **Color** y la **Etiqueta**. Para modificar el **Color** seleccionaremos Node Color> Choose question to color: Confidencia> Apply Color.

Para modificar la Etiqueta seleccionaremos Node Label> Label by question: Razón encuentro> Apply Label.

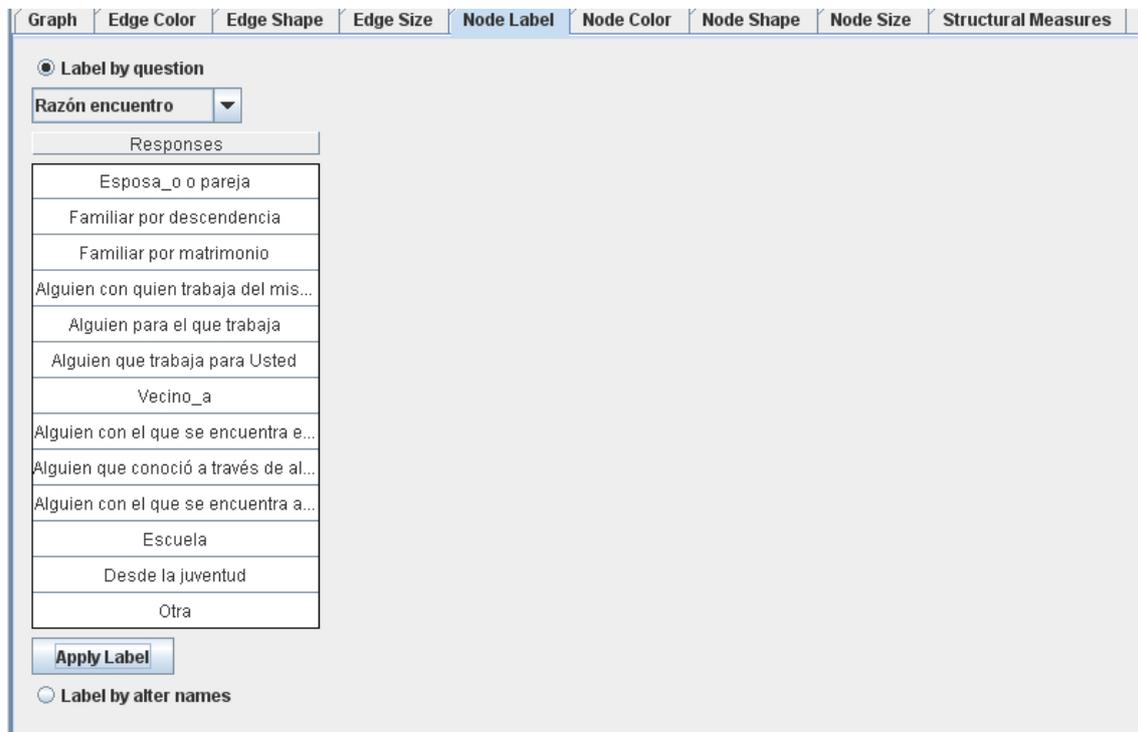


Ilustración 64. Modificación de la etiqueta para la visualización 3.

El resultado es el siguiente:

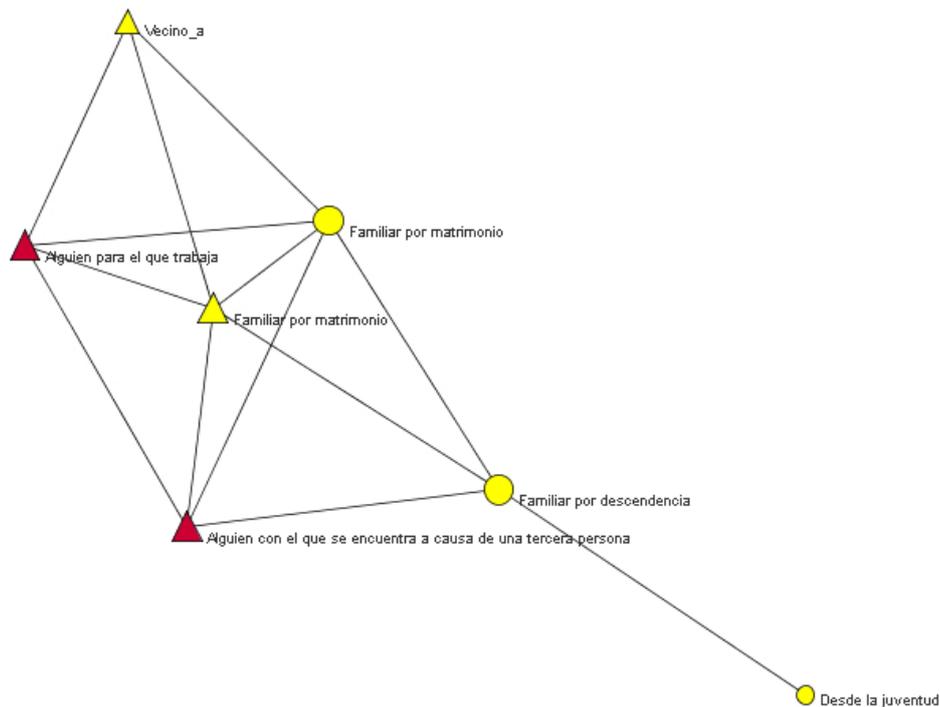


Ilustración 65. Modificación de la etiqueta para la visualización 3.

Los ejemplos pueden multiplicarse. El siguiente paso será exportar a SPSS o R todas las entrevistas realizadas con sus correspondientes variables e iniciar el análisis.

BIBLIOGRAFÍA

Davis, A.; Gardner, B.B.; Gardner, M.R. (1941). *Deep South. A Social Anthropological Study of Caste and Class*. Chicago: University of Chicago Press.

Molina, José Luis (2005). "El análisis de las redes personales", *Empiria*, 10 Julio-Diciembre (71-106).

LA COMPARACIÓN CUANTITATIVA DE GRUPOS CON SPSS⁸

Mientras los investigadores de *redes sociocéntricas* suelen recoger datos de una sola red, los investigadores de *redes personales* recogen datos de diferentes redes. Estas redes se pueden considerar también como muestras estadísticas representativas respecto a una población más amplia y ser analizadas con SPSS.

Por regla general, el estudio de redes sociocéntricas se realiza de manera individual. Sin embargo, en el caso de las redes personales, los investigadores pueden analizar un conjunto de redes de manera simultánea.

Un investigador puede recolectar redes personales en una localidad y observar diferencias según algún atributo de ego (edad, sexo, nacionalidad, ingresos, etc.). Por ejemplo, se podría investigar si las redes de hombres y mujeres difieren entre sí en la frecuencia media con que mantienen contacto con sus Alteri, o si el porcentaje de familiares en las redes covaria con la densidad.

Para esto tenemos que convertir las características de estructura y composición de las redes en variables atributivas (p.ej., la frecuencia media de contacto, el porcentaje de familiares en la red, la densidad). Luego de eso, podremos recurrir a procedimientos estadísticos para averiguar si las diferencias (o semejanzas) observadas en la muestra son lo suficientemente consistentes para generalizarlas hacia una población más amplia. Por supuesto, un requisito para tal generalización es que la muestra sea estadísticamente representativa de la población y que las personas en la muestra hayan sido seleccionadas de forma independiente (una muestra aleatoria).

Hay varios procedimientos estadísticos que se pueden aplicar. Este capítulo se enfoca en *la comparación de las características de las redes personales entre*

⁸ Autora: Miranda J. Lubbers (Departament d'Antropologia social de la UAB).

dos subgrupos de personas. Es necesario recordar algunos conceptos de estadística y entender como los atributos de las redes personales se pueden convertir en variables de análisis estadístico.

Este capítulo se divide en dos partes. La primera es un repaso de los conceptos estadísticos más importantes para el análisis de redes personales con SPSS. La segunda es el taller de autoaprendizaje de análisis de redes personales con SPSS. Para este taller está disponible una base de datos (fichero) con datos seleccionados del proyecto “Development of a Social Network Measure of Acculturation and its Application to Immigrant Populations in South Florida and Northeastern Spain” (National Science Foundation, BCS-0417429) dirigido por Chris McCarty y José Luis Molina. (www.egoredes.net).

Con estos datos, todos agregados al nivel de Ego, compararemos dos grupos diferenciados por su nacionalidad de origen con respecto a una característica (una *variable*). Por ejemplo, podremos investigar si los inmigrantes de origen senegalés residentes en Barcelona tienen mayor número de españoles en sus redes personales con relación a los dominicanos. En este ejemplo hay dos subgrupos (*casos*): a) los inmigrantes senegaleses y b) los inmigrantes dominicanos; y una *variable*, el número de españoles en su red personal.

A. ESTADÍSTICA DESCRIPTIVA Y GRÁFICOS

Es habitual empezar los análisis cuantitativos explorando y describiendo las variables una por una. Por ejemplo, para describir la muestra, solemos dar información sobre el número de entrevistados en cada subgrupo escogido, la edad media de la muestra, y el porcentaje de hombres o mujeres. Para eso, usamos *estadística descriptiva* (1). Esos mismos datos también se pueden describir usando gráficos. En ambos casos, podemos probar ciertas hipótesis sobre el estudio. Por ejemplo, si en ciertas variables hay diferencias entre grupos o si hay semejanzas entre variables. Esto se suele llamar *Inferencia estadística* (2). En ambos casos también decidimos sobre el tipo de análisis, tomando en cuenta la *escala de medición* de una variable.

En términos generales hay cuatro escalas de medición:

- Variables categóricas:

- **Escala nominal:** Los valores de una variable clasifican los casos en categorías sin un orden intrínseco, de modo que todos los que pertenezcan a la misma categoría sean equivalentes respecto del atributo en estudio.
 - P.ej.: *Religión: las categorías 'cristiano', 'musulmán', 'hindú', 'otra religión' y 'sin religión'. No hay un orden intrínseco en estas categorías, como de bajo a alto, de claro a oscuro, de poco a mucho, o de no religioso a muy religioso, etcétera.*
 - P.ej.: *Lugar de nacimiento.*
- **Escala ordinal:** Los valores representan categorías que tienen un cierto orden intrínseco, pero las distancias entre las categorías no tienen importancia.
 - P.ej.: *Nivel de estudios realizados: p.ej. las categorías 'sin estudios', 'estudios primarios', 'secundaria', 'barra profesional', 'universitarios'. Las categorías están ordenados de menor a mayor, pero las diferencias entre dos categorías no tienen sentido. La diferencia entre 'sin estudios' y 'estudios primarios' no es igual a la existente entre 'primaria' y 'secundaria'.*
 - P.ej.: *Proximidad: p.ej., las categorías 'no me siento nada próximo', 'no muy próximo', 'bastante próximo', 'muy próximo'.*
- Variables de escala:
 - **Escala de intervalo:** Los valores de dichas variables tienen distancias entre sí que son iguales, pero no tienen un punto cero propio como origen. Hay pocas variables que tienen verdaderamente esta escala de medida:
 - P.ej.: *Temperatura en grados Celsius: Hay un orden de frío a caliente, y la diferencia entre 5 y 10 grados es lo mismo que entre 10 y 15 grados. Pero, el punto cero es arbitrario, una ausencia de temperatura. Por ello, no se puede decir con una temperatura de 20 grados que hace dos veces más calor que con una temperatura de 10 grados.*
 - P.ej.: *El año en el calendario. Hay un orden de más antiguo a más reciente, y la diferencia entre 1897 y 1903 es la misma que entre 1903 y 1909. No hay un punto cero natural, el año 0 es arbitrario.*
 - **Escala de razón:** Las valores tienen un punto cero propio como origen, o sea es posible decir que un valor es el doble de otro valor.
 - P.ej.: *Los ingresos, medidos en euros. Alguien que gana 3000 euros al mes gana dos veces más que alguien que gana 1500 euros al mes.*
 - P.ej.: *Edad, medida en años.*

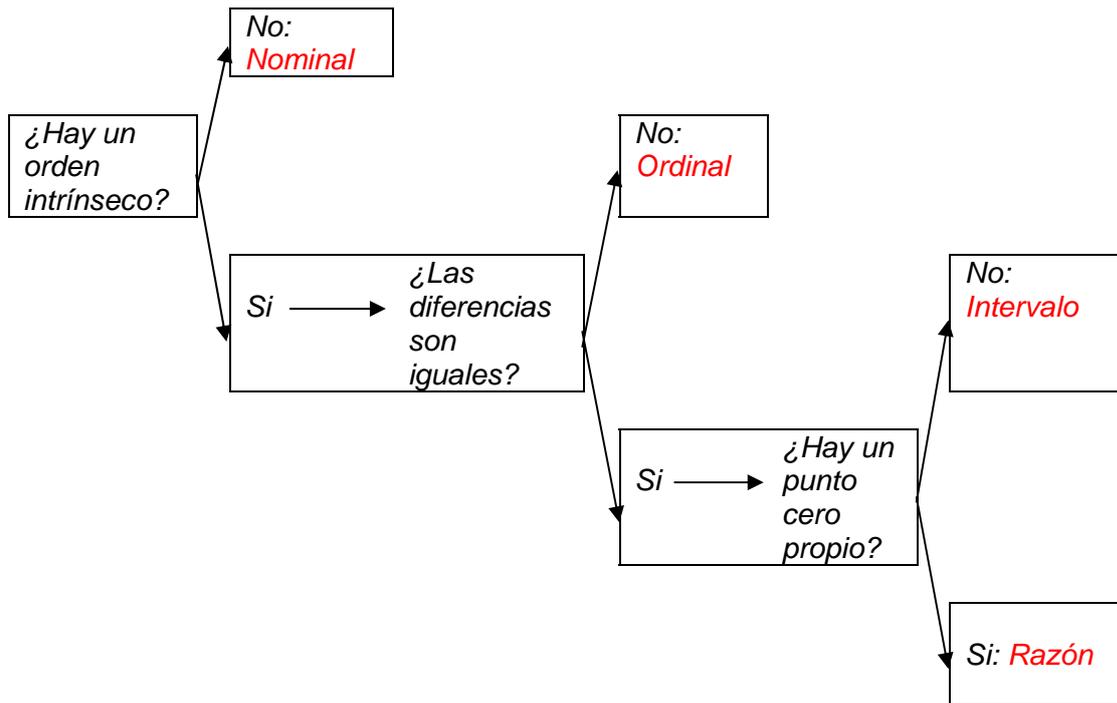


Ilustración 66. Árbol de decisión para determinar la escala de medición.

La escala de medición es importante porque determina las operaciones estadísticas que se permiten hacer: Para comparar casos en un variable nominal, sólo podemos decir si son iguales o diferentes. Con variables ordinales, además podemos comprobar si son mayores o menores. Variables de intervalos permiten la adición y la substracción y variables de razón además la multiplicación y la división. Las pruebas que podemos utilizar para comparar los valores a una variable entre dos grupos dependen por lo tanto de la escala de medición de la variable.

En esta práctica usamos los siguientes estadísticos, gráficos y pruebas (vease la Tabla 1). No obstante, hay muchas más opciones (por ejemplo, para comparar la diferencia entre dos grupos en una variable ordinal, se utiliza habitualmente la Prueba de Mann-Whitney).

Queremos ... Escala de medición	(1) Describir una variable <i>(Describir gráficamente)</i>	(2) Comparar la diferencia entre dos o más grupos en una variable estadísticamente
Categorica	Frecuencias <i>(Gráfico de Sectores)</i>	La prueba de la chi-cuadrado
Escala	Estadísticos descriptivos (promedio, desviación típica) <i>(Histograma)</i>	Con dos grupos: La prueba de la <i>t</i> - para muestras independientes (Con más grupos, o si hay 'covariables': el Análisis de Varianza)

Tabla 1. Resumen de los estadísticos, gráficos y pruebas en esta práctica.

Para calcular las frecuencias y estadísticas descriptivas se utiliza un programa estadístico como SPSS (*“Statistical Package for the Social Sciences”*). Sin embargo, para entender cómo aplicar adecuadamente estas pruebas e interpretar sus resultados, es necesario primero entender los conceptos teóricos detrás de estas pruebas. En las siguientes secciones explicaremos estos conceptos. Luego, al final, proponemos un taller de SPSS. Si ya conoce la teoría, puede ir directamente al taller de SPSS.

B. LA PRUEBA DE CHI-CUADRADO (χ^2)

Teoría

La pregunta que buscamos analizar con la prueba de chi-cuadrado es del siguiente tipo:

¿Hay una relación entre <dos variables categóricas>?

La hipótesis nula H0: Las dos variables son independientes
(No hay una relación).

La hipótesis alternativa H1: Las dos variables son interdependientes.

Para aplicar la prueba a las diferencias entre dos o más grupos, se puede traducir esta pregunta en la siguiente:

¿Difieren dos (o más) grupos de casos en la distribución de <una variable categórica>?

La hipótesis nula H0: No hay diferencia.

La hipótesis alternativa H1: Hay diferencia.

Por ejemplo: En muchos países en Europa, todavía hay una gran diferencia en la participación laboral de mujeres y hombres en las universidades. No sólo trabajan menos mujeres en las universidades, sino también el nivel en el que trabajan es diferente. Vamos a investigar hipotéticamente si el *segundo aspecto* es verdad en el caso de España. Para eso nos limitamos al profesorado contratado laboral (excluimos los profesores de los cuerpos docentes – es decir, profesores titulares de universidad, catedráticos de escuelas universitarias y titulares de escuelas universitarias). Distinguimos cuatro categorías en la variable ‘nivel del profesorado’: catedrático, profesor agregado, profesor lector, y asociado.

Las frecuencias observadas

Suponga que hemos tomado una muestra aleatoria y representativa de 1937 personas de todas las universidades de España⁹. En la muestra hay 1251 hombres y 686 mujeres. De esta muestra 337 personas son catedráticos, 35 profesores agregados, 168 profesores lectores, y 1397 asociados. Podemos rellenar una tabla con estas frecuencias, que se llaman ‘*los marginales*’ (ver Tabla 2).

	Mujeres	Hombres	Total
Catedrático			337
Profesor agregado			35
Profesor lector			168
Asociado (incluso médico)			1.397
Total	686	1251	1.937

Tabla 2. Los marginales.

De los hombres, 263 son catedráticos, 25 son profesores agregados, 89 son profesores lectores, y 874 son asociados. De las mujeres, 74 son catedráticas,

⁹ De hecho, no era una muestra. Los números en la Tabla 2 son los números totales observados en la UAB 2005-2006. Por supuesto, si queremos un estudio nacional, sería más fácil preguntar esa información a cada universidad al nivel agregado que tomar una muestra de los empleados. Pero para los fines de esta práctica trataremos estos datos como si fueran una muestra.

10 son profesoras agregadas, 79 son profesoras lectoras, y 523 son asociadas. Podemos añadir estas frecuencias en nuestra tabla (ver Tabla 3):

	Mujeres	Hombres	Total
Catedrático	74	263	337
Profesor agregado	10	25	35
Profesor lector	79	89	168
Asociado (incluso médico)	523	874	1.397
Total	686	1251	1.937

Tabla 3. Profesorado (hombres y mujeres) en las universidades españolas.

Las frecuencias en la tabla representan las frecuencias que hemos observado en la muestra, por eso las llamamos '*las frecuencias observadas*'.

Ahora, nos interesa saber lo siguiente: ¿La distribución entre los niveles es igual para hombres que para mujeres? Podemos ver que a cada nivel, hay más hombres que mujeres, pero eso sólo refleja que en general trabajan más hombres en la universidad que mujeres. No nos cuenta si las mujeres tienen posiciones iguales a hombres dentro de la universidad.

Las frecuencias esperadas

Si miramos cifras totales, podemos calcular que $337/1937 = 17\%$ es catedrático, $35 / 1937 = 2\%$ es profesor agregado, $168 / 1937 = 9\%$ es profesor lector, y $1397 / 1937 = 72\%$ es asociado.

Si hombres y mujeres no difieren entre sí, es de esperar que la distribución sea igual para mujeres y para hombres. Entonces, esperaríamos que el 17% de todas las mujeres (686) fuesen catedráticas, 2 % fuesen profesoras agregadas, 9% fuesen profesoras lectoras, y 72% fuesen asociadas. Y esperamos también que el 17% de todos los *hombres* (1251) fuesen catedráticos,.... etcétera.

	Mujeres	Hombres	Total
Catedrático	17% × 686	17% × 1251	337 (17%)
Profesor agregado	2% × 686	2% × 1251	35 (2%)
Profesor lector	9% × 686	9% × 1251	168 (9%)
Asociado (incluso médico)	72% × 686	72% × 1251	1397 (72%)
Total	686	1251	1937

Tabla 4. Frecuencias que esperamos si no hubiera diferencia (si la hipótesis nula fuera verdad)

Como ya hemos dicho, esperamos que 17% de las 686 mujeres fuesen catedráticas. Porque el 17% resulta de $337 / 1937$, podemos también concluir que esperamos que $686 \times 337/1937$ sean catedráticas (ver Tabla 5). *En general: para calcular las frecuencias esperadas tomamos el producto de los dos marginales correspondientes de una casilla y lo dividimos por el total.*

	Mujeres	Hombres	Total
Catedrático	$686 \times 337 / 1937$		337 (17%)
Profesor agregado			35 (2%)
Profesor lector			168 (9%)
Asociado (incluso médico)			1397 (72%)
Total	686	1251	1937

Tabla 4. Manera de calcular las frecuencias esperadas.

Y lo haremos en cada casilla, por ejemplo la casilla (profesor lector, hombre).... (ver Tabla 5).

	Mujeres	Hombres	Total
Catedrático	$686 \times 337 / 1937$	$1251 \times 337 / 1937$	337 (17%)
Profesor agregado	$686 \times 35 / 1937$	$1251 \times 35 / 1937$	35 (2%)
Profesor lector	$686 \times 168 / 1937$	$1251 \times 168 / 1937$	168 (9%)
Asociado (incluso médico)	$686 \times 1397 / 1937$	$1251 \times 1397 / 1937$	1397 (72%)
Total	686	1251	1937

Tabla 5. Manera de calcular las frecuencias esperadas.

Después de realizar estas operaciones, las frecuencias esperadas son las que aparecen en la Tabla 6 ($686 \times 337 / 1937 = 119,35$, etcétera). Advierte que las casillas de cada fila y cada columna suman de nuevo a sus marginales.

	Mujeres	Hombres	Total
Catedrático	119,35	217,65	337
Profesor agregado	12,40	22,60	35
Profesor lector	59,50	108,50	168
Asociado (incluso médico)	494,76	902,24	1397
Total	686	1251	1937

Tabla 6. Frecuencias esperadas.

Ahora tenemos dos tablas, una con las *frecuencias observadas* (O) (Tabla 2) y una con las *frecuencias esperadas* (E) (Tabla 6). Cuando las comparamos, podemos ver algunas diferencias interesantes. Por ejemplo, parece que en la hipótesis de igualdad, hay 119,35 catedráticos femeninos, pero, en verdad, son sólo 74. Son muchas menos que en una situación de igualdad. Pero en lugar de examinar cada casilla, necesitamos también una medida para la diferencia total. Necesitamos una manera de decidir si podemos rechazar la hipótesis nula o no.

Por eso, en primer lugar vamos a medir el tamaño de cada diferencia (de cada casilla). Más específicamente, para cada casilla, calculamos

$$(O - E)^2 / E$$

Por ejemplo, para la casilla (mujer, catedrática), la frecuencia observada era 74, y la frecuencia esperada 119,35. Entonces, el tamaño de la diferencia es:

$$(74 - 119,35)^2 / 119,35 = 17,23.$$

Lo calculamos para cada casilla. En la tabla 7, ponemos todos los resultados.

	Mujeres	Hombres	Total
Catedrático	17,23	9,45	
Profesor agregado	0,46	0,25	
Profesor lector	6,39	3,50	
Asociado (incluso médico)	1,61	0,88	
Total			39,77

Tabla 7. Los resultados del cálculo $(O - E)^2 / E$.

La suma de estos resultados es el valor de la prueba 'Chi-cuadrado'.

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

En nuestro ejemplo:

$$\chi^2 = 17,23 + 0,46 + 6,39 + 1,61 + 9,45 + 0,25 + 3,50 + 0,88 = 39,77$$

Si no hay ninguna diferencia entre los dos grupos, el producto es cero (porque la diferencia sería cero, y si dividimos cero por cualquier número, sigue siendo cero). Si hay una diferencia (es decir, si el valor es más de cero), debemos saber si esta diferencia es lo suficientemente consistente para concluir que esta diferencia no es producto de la casualidad, sino que tenemos el nivel de confianza necesario para generalizarla a la población.

Como sabemos que muchas variables de 'razón' siguen una *distribución normal*, y que este conocimiento nos da una información importante, también se conoce que, bajo la hipótesis nula, el estadístico chi-cuadrado sigue una *distribución chi-cuadrado* con $(f-1)(c-1)$ grados de libertad.

- *¿Qué son los grados de libertad?* Se refiere al número de valores en el cálculo de un estadístico que es libre de variar. En general, una tabla tiene $(f - 1)(c - 1)$ grados de libertad, siendo f el número de filas, y c el número de columnas. En el ejemplo hay 4 filas y 2 columnas, entonces hay $gl = (4-1) \times (2-1) = 3$ grados de libertad. Los denominan 'grados de libertad', porque cuando has determinado los valores de tantas casillas (en el ejemplo 3 casillas), las otras casillas ya no son *libres de variar* (siempre cuando los marginales están conocidos). En el ejemplo, si conoces los valores de tres casillas, puedes calcular las otras (ver Tabla 8). P.ej., si hay 337 catedráticos y 74 de ellos son mujeres, sabes que $337 - 74$ son hombres. La casilla (catedrático, hombre) ya no es libre de variar.

	Mujeres	Hombres	Total
Catedrático	74		337
Profesor agregado	10		35
Profesor lector	79		168
Asociado (incluso médico)			1397
Total	686	1251	1937

Tabla 8. Tres grados de libertad.

Hay que buscar el valor chi-cuadrado con los grados de libertad en una distribución de chi-cuadrado, para ver si la diferencia es suficientemente grande. Antes, necesita decidir el *valor de significación*. Es decir, el riesgo asumido de cometer el error de rechazar la hipótesis nula cuando de hecho (en la población) esa hipótesis es verdad. Entonces se corre el riesgo de decir que no hay una diferencia cuando en realidad sí que lo hay.

Con muestras, nunca podemos estar 100% seguros de que nuestra conclusión es válida en la población. *Normalmente, aceptamos un riesgo de 5%* de que rechazamos la hipótesis nula cuando no hay diferencia. Ahora bien, si encontramos en la distribución de chi-cuadrado:

- un valor de $p \leq ,05$, concluimos (con 95% seguridad) que hay una diferencia entre los dos grupos. Tenemos un 95% en nuestro favor para generalizar sin temor, sin equivocarnos, la hipótesis alternativa (y un 5% en nuestra contra). Decimos que la diferencia es *significativa*.
- un valor de $p > ,05$, concluimos que la diferencia entre los dos grupos no es significativa, o sea que hay más probabilidad de no encontrar la diferencia en la población.

Cuando el tamaño de la muestra es muy grande, solemos tomar un riesgo más pequeño de equivocarnos (1%), porque en las muestras más grandes se pueden detectar pequeñas diferencias con más precisión.

La significación estadística NO quiere decir que las diferencias son importantes, ni nos ayuda con la interpretación de la relación (por eso, necesitamos investigar la tabla). Sólo puede decir sí podemos tener confianza de que la diferencia entre dos grupos (o la relación entre dos variables) que hemos encontrado en la muestra también existe en la población (y no podemos atribuirlo a errores experimentales).

Pre-condiciones:

Para usar la prueba de la chi-cuadrado necesitamos (además de una muestra aleatoria):

- Por lo menos 50 casos en total (aunque hay diferentes estándares según la fuente que se consulta, por ejemplo 20 y 100);
- Por lo menos una frecuencia *esperada* de
 - 5 casos en cada casilla para una tabla de 2 por 2,
 - 5 casos en por lo menos 80% de las casillas para tablas más grandes, **y** ninguna casilla con 0 casos.
- Cada caso sólo puede estar en *una* casilla de la tabla.

Cómo escribir los resultados de la prueba de la chi-cuadrado en un informe

Si, por ejemplo, el valor del χ^2 es 10,51, hay 90 casos, 4 grados de libertad y el valor del p es ,03, escribimos: $\chi^2 (4, N = 90) = 10,51, p < .05$

Si el valor del χ^2 es 4,51, hay 90 casos, 4 grados de libertad y el valor del p es ,34, escribimos ('ns' = no es significativa): $\chi^2 (4, N = 90) = 4,51, ns$

C. LA PRUEBA DE LA -T- PARA MUESTRAS INDEPENDIENTES

Teoría

La pregunta que probamos con la prueba de la -t- para muestras independientes es típicamente:

¿Difieren dos grupos de casos distintos en los dos promedios en <una variable de escala>?

La hipótesis nula H0: No hay diferencia (los promedios son iguales).

La hipótesis alternativa H1: Hay diferencia (los promedios no son iguales).

Ejemplo: Leung, Pe-Pua, & Karnilowicz (2006) investigaron las diferencias entre jóvenes de diferentes grupos étnicos según ocho variables psicológicas. Por ejemplo, investigaron si los anglo-australianos difieren de los filipino-australianos en síntomas psicológicos. Estudiaron los síntomas con una escala de 15 variables, cada una con cinco categorías ordinales, que miden depresión, ansiedad y síntomas psicósomáticos. Es una variable de escala, en el sentido que valores bajos indican pocos síntomas psicológicos, y valores altos indican muchos síntomas psicológicos. Los datos son los siguientes (ver Tabla 9). (*¡Si has olvidado lo que es la desviación típica, por favor echa un vistazo al apéndice!*)

	Filipino-Australianos	Anglo-Australianos
\bar{X} = el promedio de síntomas psicológicos	37,98	36,80
SD = la desviación típica de síntomas psicológicos	9,64	9,31
N = el número de entrevistados	106	157

Tabla 9. Las diferencias en los síntomas psicológicos de dos colectivos.

Fuente: Leung, C., Pe-Pua, R., & Karnilowicz, W. (2006). Psychological adaptation and autonomy among adolescents in Australia: A comparison of Anglo-Celtic and three Asian groups. *International Journal of Intercultural Relations*, 30, 99-118.

Si comparamos las dos medias, podemos ver que los síntomas psicológicos de los Anglo-Australianos son más bajos que los de los Filipino-Australianos. No obstante, la pregunta es: ¿Será la diferencia suficientemente grande para generalizar la diferencia a la población?

En general, si encontramos una diferencia entre dos grupos, es necesario saber cuál debe ser el tamaño de la diferencia para concluir que en la población a la que queremos generalizar haya una diferencia entre los dos grupos también.

Eso depende no sólo de la diferencia entre las dos medias, sino también de la variabilidad en los dos grupos y la dispersión de los datos. Si los casos difieren poco entre sí, una pequeña diferencia puede ser importante, pero si los casos difieren mucho, podemos encontrar una pequeña diferencia entre los dos grupos por una casualidad.

Para saber si la diferencia entre dos medias es diferente *estadísticamente*, debemos hacer una prueba: La prueba de la $-t-$ para muestras independientes (*muestras independientes* quiere decir que los grupos contienen casos distintos. También hay una prueba de la $-t-$ para muestras relacionadas, que se utiliza, por ejemplo, si las mismas personas tienen valores a una escala medida antes y después de la inmigración). La prueba de la $-t-$ considera la diferencia entre los promedios de dos grupos relativamente a sus dispersiones.

Para ello, la prueba divide la diferencia entre los dos promedios por una medida de la variabilidad.

$$t = \frac{\text{la diferencia entre promedios}}{\text{el error estándar de la diferencia}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\text{var}_1}{n_1} + \frac{\text{var}_2}{n_2}}}$$

En nuestro ejemplo (ver la tabla en la página anterior):

$$t = \frac{37,98 - 36,80}{\sqrt{\frac{9,64^2}{106} + \frac{9,31^2}{157}}} = \frac{1,18}{1,20} = 0,98$$

Bajo la hipótesis nula, el estadístico de la $-t-$ sigue una distribución de la $-t-$ con

$$gl = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

Se llaman grados de libertad (gl) por que, si sabemos los valores (p.ej. el estatus ocupacional) de todos los miembros de un grupo *menos uno*, y sabemos el promedio, podemos calcular el valor del último también. Y para el segundo grupo vale lo mismo.

En nuestro ejemplo:

$$gl = n_1 + n_2 - 2 = 106 + 157 - 2 = 261$$

Buscamos el valor de la $-t-$ con sus grados de libertad en una tabla de la distribución de la $-t-$, y parece que $p = ,32$, o sea $p > ,05$. Es decir, la diferencia entre los dos grupos NO desvia significativamente de cero (vease la prueba de la chi-cuadrado). Podemos concluir que los anglo-australianos no tienen más síntomas psicológicos que los filipino-australianos.

Pre-condiciones:

Para usar la prueba de la $-t-$ necesitamos (además de una muestra aleatoria):

- Por lo menos 10 casos por grupo;
- En cada grupo, la variable debe tener una distribución normal;
- Las varianzas de los dos grupos no tienen que diferir mucho entre sí.

Cómo escribir los resultados de la prueba de la $-t-$ en un informe o una tesina

Cuando el valor de la $-t-$ es 2,01, los grados de libertad 98 y el valor del p ,02, escribimos:

$t(98) = 2,01$, $p < ,05$. Cuando el valor de la $-t-$ es 0,50, los grados de libertad 98 y el valor del p es ,62, escribimos (ns = no es significativo): $t(98) = 0,50$, ns

Práctica con SPSS

Para ésta práctica vamos a utilizar un archivo de datos denominado “*inmigrantes_España.sav*”, que contiene una pequeña selección de datos del proyecto “*Development of a Social Network Measure of Acculturation and its Application to Immigrant Populations in South Florida and Northeastern Spain*” (National Science Foundation, BCS-0417429) dirigido por Chris McCarty y José Luis Molina. Este archivo contiene algunos datos de inmigrantes dominicanos y senegambianos en España. El archivo se puede bajar desde www.egoredes.net/

El programa SPSS Statistics¹⁰ implementa las tres fases para realizar un estudio estadístico:

- Fase 1. Implementar de forma organizada y ordenada la base de datos.
- Fase 2. Manipular los datos (p.ej. tomar una selección de los casos, recodificar o categorizar una variable).
- Fase 3. Analizar los datos.

Básicamente, hay tres ventanas diferentes en SPSS:

- **Editor de datos:** Esto es la ventana principal y la primera ventana que aparece nada más de acceder al programa. Hay dos pestañas:
 - Vista de datos.
 - Vista de variables.
- **Visor de resultados:** Los resultados generados por los procedimientos realizados se muestran en el visor. Esta ventana se abre de forma automática en el momento que se ejecuta el primer análisis.
- **Editor de sintaxis** (opcional): En lugar de trabajar con el menú, también podemos usar la sintaxis, o cuando usamos el menú, podemos pegar la sintaxis generada para guardarla. Por ahora, no lo usaremos, pero si va a trabajar más a menudo con SPSS, es aconsejable aprenderlo. Con la sintaxis, siempre puede recordar qué manipulaciones ha hecho con los datos y puede repetir los análisis fácilmente.

Vamos a conocer el programa SPSS respondiendo a las preguntas siguientes. (Si ya conoce a SPSS, puede dar sus respuestas trabajando con SPSS directamente. Si todavía no conoce a SPSS o si ha olvidado cómo funciona, siga los pasos debajo de las preguntas).

¹⁰ Utilizamos la versión 17.

A. ESTADÍSTICA DESCRIPTIVA

Preguntas para responder en la sección A

1. ¿El caso con ID 50, es hombre o mujer? ¿De qué país de origen es? ¿Qué edad tiene? ¿Cuántos hijos tiene? ¿Y el caso con ID 99?
2. ¿Cuáles son las categorías de la variable religión?
3. ¿Qué porcentaje de los inmigrantes han sufrido racismo en España?
4. Representa la distribución de experiencias de racismo de forma gráfica.
5. ¿Cuáles son los años medios que los inmigrantes residen en España?
6. Representa la distribución del número de hijos de forma gráfica. ¿Puedes decir que el número de hijos tiene una distribución normal?

Paso a paso

1. Abrir un archivo y orientación

- a. Abre el programa SPSS (Cancela el cuadro de diálogo que surge). Se abre de forma automática el **editor de datos**. Lo que vemos es un matriz vacío. Hay que introducir los datos manualmente (Etapa 1), o abrir una base de datos de SPSS previamente guardada (lo último es lo que vamos a hacer).
- b. Selecciona dentro del menú 'Archivo' la opción 'Abrir' y dentro de ella 'Datos'. Ahora, selecciona el archivo '*inmigrantes_españa.sav*' y pulse el botón 'Abrir'. Advierta que cambiando el tipo de archivo en la ventana de abrir datos, se pueden abrir también archivos de Excel (por ejemplo, los archivos previamente exportados desde Egonet).
- c. Ahora la matriz contiene datos. Cada fila tiene los datos de 1 entrevistado (de un 'caso') que estaban recogidos con Egonet (las 119 filas representan los 119 casos), y cada columna tiene los datos de 1 *variable*. Por ejemplo, la fila 1 tiene los datos del entrevistado con ID 1, su país de origen tiene el valor '2' y su sexo el valor '1'. Para ver el significado de los valores puedes pulsar el icono de la barra de herramientas que muestra una etiqueta (o elige dentro del menú 'Ver' la opción 'Etiquetas de valor'). Ahora puedes ver que la persona con ID 1 es un hombre dominicano.
- d. Ahora puedes responder a la primera pregunta.
- e. También puedes ver que a veces, los casos no tienen valores, como el caso 2 y el caso 18 a la variable '*racismo*'. Eso quiere decir que estas personas no tienen respuestas a la pregunta correspondiente ('¿Cuántas situaciones de racismo ha sufrido?'), por ejemplo porque no quisieron responder, o porque el entrevistador olvidó preguntarlo. Los llamamos *valores perdidos*.

- f. En la esquina inferior izquierda, puedes ver dos pestañas que permiten cambiar entre la '*Vista de datos*' (la que muestra los datos) y la '*Vista de variables*' (la que muestra una lista de variables). Ahora está en la vista de datos. Pulsa en la pestaña '*Vista de variables*'. Aquí vemos una lista de todas las variables en el archivo y sus definiciones. Si por ejemplo, pulsamos en la casilla 'valores' de '*país de origen*', podemos ver cuáles son los valores de 'país de origen' en esta muestra.
- g. Puede responder a la pregunta 2.
- h. En la *barra de menú* (arriba de la pantalla), aparecen diversas opciones. Algunas de ellas ('Archivo', 'Edición', 'Ver', 'Ventana', y 'Ayuda') son comunes a muchos programas basados en Windows, y otras son específicas de SPSS. Las últimas son:
 - A. Datos: Contiene opciones para hacer cambios que afectan a todo el archivo de datos (como unir archivos, seleccionar casos, transponer variables y casos) (Fase 2).
 - B. Transformar: Permite efectuar cambios sobre las variables seleccionadas y crear nuevas variables (Fase 2).
 - C. Analizar: Desde esta opción se ejecutan todos los procedimientos estadísticos (Fase 3).
 - D. Gráficos: Desde esta opción se ejecutan todos los procedimientos estadísticos de carácter gráfico (Fase 3).

Vamos a explorar la barra del menú.

2. Frecuencias

- a. Para calcular frecuencias (un procedimiento estadístico), hay que seleccionar el menú 'Analizar', dentro de éste elegimos la opción 'Estadísticos Descriptivos' y después 'Frecuencias'.
- b. En el cuadro de diálogo que se abre, selecciona la variable 'racismo' y clicla la flecha, para que el nombre de la variable aparezca en el recuadro 'Variables'. A la derecha de este recuadro se muestran algunas opciones extras, pero por ahora, lo dejamos. Pulsa el botón 'Aceptar'.
- c. Ahora se abre de forma automática una nueva ventana, **el Visor de resultados**, que muestra los resultados.
- d. ¿Ve que la tabla muestra tres columnas además de las frecuencias?: porcentaje, porcentaje válido y porcentaje acumulado.

<i>Pregunta:</i> ¿Cuáles son las diferencias entre esas columnas?

- e. Ahora puede responder a la pregunta 3.

3. Gráfico de sectores

- a. Para presentar las frecuencias de una variable categórica de forma más atractiva, puedes hacer un **gráfico de sectores**. Para ello, podemos seleccionar el menú 'Gráficos' dentro de ello la opción 'Generador de gráficos'. Ahora se abre una ventana interactiva. Pulse en la pestaña 'Galería' si no está seleccionada, y dentro de ella seleccione 'Sectores/Polar'. Arrastra el icono correspondiente al gráfico de sectores a la zona grande encima de la galería.
- b. Ahora hay que arrastrar la variable racismo desde la lista de variables a la zona de colocación denominado '¿Sectores por?' en la zona grande. Una vez que el generador muestra el gráfico con los datos deseados, pulse el botón 'Aceptar'.

Como hemos visto antes, también podemos hacer el gráfico dentro de la opción 'Frecuencias'. Para este alternativo, volvemos a la opción 'Frecuencias' (la variable racismo ya fue seleccionada la vez anterior que abrimos esta ventana). Pulsamos el botón 'Gráficos', y dentro de la ventana que se abre seleccionamos 'Gráficos de sectores' y después 'Continuar'. En la ventana a la que hemos vuelto pulsamos 'Aceptar'. El gráfico se muestra en el Visor de resultados.

- c. *(Opcional): Clique doble al gráfico en el visor de resultados. Ahora surge una nueva ventana: el **Editor de gráficos**. Con este editor, podemos obtener representaciones con efectos más llamativos. Por ejemplo, clique a uno de los sectores en el gráfico, elija el menú 'Elementos' y elija la opción 'Desgajar sector'. Mire el cambio. También puede, por ejemplo, cambiar colores de sectores clicando doble a un color en la leyenda. Ahora surge un nuevo cuadro de diálogo, dentro de ello, elija la pestaña 'relleno y borde' si no está seleccionado. Clique a la casilla 'relleno' y elige después un color. Pulse en el botón 'aplicar' (y cierre el cuadro de diálogo).*

Explore el editor de gráficos poniendo por ejemplo un título al gráfico y exportelo.

4. Promedio y desviación típica

- a. Para calcular el promedio y la desviación típica de una variable, hay que seleccionar el menú 'Analizar', dentro de éste la opción 'Estadísticos Descriptivos' y después 'Descriptivos'.
- b. En el cuadro de diálogo que se abre ahora, selecciona la variable 'años de residencia' y pulse la flecha, para que el nombre de la variable aparezca en el recuadro 'Variable'. Todavía no pulse 'Aceptar'.
- c. Dentro de éste cuadro de diálogo, clicas en el botón 'Opciones'.
- d. En el cuadro de diálogo que se abre ahora, se pueden elegir distintos estadísticas que se desean incluir en la tabla de

resultados. Elige por lo menos 'Media' y 'Desviación típica'. Pulse el botón 'Continuar' para regresar al cuadro de diálogo anterior.

- e. Esta vez no tiene sentido la solicitud de la tabla de frecuencias, porque esta variable tiene un gran número de valores diferentes, así que dicha tabla no es informativa. Por eso, optamos por suprimir la tabla. Para ello, pulsa en 'muestra frecuencias' para que ésta casilla esté vacía. Pulsa el botón 'aceptar'.
- f. Ahora se abre la ventana visor, que muestra los resultados.
- g. Responde a la pregunta 5.

5. Histograma

Para presentar la distribución de una variable de escala de forma más atractiva, puede hacer un histograma. Para responder a la pregunta 6:

- a. En el menú 'Gráficos', elige la opción 'Generador de gráficos'. En la galería, elija 'histograma'. Arrastra el primer icono a la zona grande encima de la galería.
- b. Arrastra la variable 'número de hijos' a la zona de colocación denominada '¿Eje X?'.
c. Pulsa el botón 'Aceptar' y los resultados aparecen en la ventana Visor. Si quieres, puedes editar el histograma clicando doble al gráfico en el visor de resultados.
- d. Ahora puedes responder a la pregunta 6.

Para guardar los resultados / copiar tablas y gráficos a Word:

Para guardar los resultados, selecciona el visor de resultados y elige dentro del menú 'archivo' la opción 'guardar como'. Escribe un nombre y pulsa 'aceptar'. Observa que por ahora sólo has guardado los resultados. Si quieres guardar los datos (por ejemplo cuando has manipulado los datos) o la sintaxis, necesitas hacerlo por separado (para guardar los datos, cambia al editor de datos y elige dentro del menú 'archivo' la opción 'guardar como'; para guardar el sintaxis, cambia al editor de sintaxis y elige dentro del menú 'archivo' la opción 'guardar como'). Por defecto, la extensión para el fichero de los resultados es .spo, para los datos .sav y para la sintaxis .sps.

Si más adelante, quieres abrir uno de los archivos dentro de SPSS, elige el menú 'archivo', dentro de esto 'abrir' y después el tipo de archivo que quiere leer: datos, sintaxis, o resultados. Después elige el archivo y ábrelo.

Para insertar una tabla o gráfico de SPSS en Word, se puede simplemente copiar una tabla de SPSS y pegar directamente en Word.

B. LA PRUEBA DE LA CHI-CUADRADO

Preguntas en la sección B:

7. ¿Los dominicanos y los senegambianos difieren entre sí en el tipo de la red? ¿Puedes describir la diferencia?
8. ¿Los dominicanos y los senegambianos difieren entre sí en sus experiencias de racismo? ¿Puede describir la diferencia?

Paso a paso

- a. Para hacer una tabla de contingencia, hay que seleccionar el menú 'Analizar', dentro de éste elegimos: 'Estadísticos Descriptivos' y dentro de ella: 'Tablas de contingencia'.
- b. En el cuadro de diálogo correspondiente, debemos definir qué variables categóricas definirán las filas y las columnas de la tabla. Hay que seleccionar primeramente la variable de agrupación, 'país_origen', y pulse la flecha más arriba, para que el nombre de la variable aparezca en el recuadro 'filas'.
- c. Selecciona después la variable de interés, 'tipo_red' y pulsa la segunda flecha para que el nombre de la variable aparezca en el recuadro 'columnas'.

Obsérvese que en el mismo cuadro de diálogo se pueden definir, mediante otras variables cualitativas, nuevas dimensiones a esta tabla, es decir, estratificar la muestra de acuerdo a criterios adicionales. Pero no queremos esto.

- d. En los resultados, no sólo queremos ver las frecuencias por casilla, sino también el porcentaje que cada país de origen tiene qué tipo de red. Por eso pulse el botón 'Casillas' en el mismo cuadro de diálogo. En el nuevo cuadro de diálogo que se abre, seleccione la casilla 'fila' (detrás de 'porcentajes'). Eso quiere decir que la ventana visor muestra los porcentajes sumando a 100% *por cada fila*. Si elige 'columna', cada columna suma a 100%. Si eliges 'total', todas las casillas juntas suman a 100%. Pulsa el botón 'continuar' para regresar al cuadro de diálogo anterior.
- e. También queremos ver el valor de la prueba de la chi-cuadrado. Para ello pulsa el botón 'Estadísticos'. Entre los estadísticos, selecciona el chi-cuadrado para el contraste de independencia. Pulsa 'continuar'. Pulsa 'aceptar' en la ventana de diálogo en la que has regresado.
- f. Ahora puedes responder a la pregunta 7, interpretando los resultados de la prueba y los porcentajes (En la tabla, 'Sig asintótica (bilateral)' refiere al valor del p).
- g. Repite los pasos con la variable 'racismo' para responder a la pregunta 8.

C. LA PRUEBA DE LA $-t-$ PARA MUESTRAS INDEPENDIENTES

Preguntas en la sección C:

9. ¿Conocen los dominicanos a más españoles que los senegambianos?
10. ¿Las redes personales de los dominicanos son igual de densas que las redes de los senegambianos?

Paso a paso

- a. Dentro del menú 'Analizar', hay que seleccionar la opción 'Comparar medias' y dentro de ella 'Prueba T para muestras independientes'.
- b. Ahora se abre un cuadro de diálogo en el que hay que especificar la variable que quiere contrastar y la variable de agrupación, que define los dos grupos.
- c. En la lista de variables, selecciona la variable de interés (número de españoles en la red) y pulsa la flecha más arriba para que aparezca en el recuadro 'Variables para contrastar'.
- d. Ahora selecciona la variable de agrupación en la lista de variables ('país de origen') y pulsa la segunda flecha para que aparezca en el recuadro 'Variable de agrupación'.
- e. En el recuadro de la variable de agrupación puedes ver dos signos de interrogación: (? ?). Eso significa que todavía hay que definir los valores de la variable de agrupación que indican los dos grupos que quiere comparar. Pulsa el botón 'Definir grupos' e introduce los dos valores que definen los dos grupos. Por ejemplo, si hay una variable 'sexo', en que 1 = mujer, 2 = hombre, escribe en el cuadro de diálogo detrás de 'usar valores especificados' por grupo 1 '1' y por grupo 2 '2'. Si la variable 'sexo' estaba definido así, que 0 = mujer, 1 = hombre, escriba en el cuadro de diálogo detrás de 'usar valores especificados' por grupo 1 '0' y por grupo 2 '1'. Ya hemos visto que la variable 'país_origen' tiene los valores 1 y 2 entonces...
- f. Después de entrar los valores, pulse el botón 'Continuar' para regresar al cuadro de diálogo anterior. Ahora puede ver que los signos de interrogación (?, ?) están sustituidos por (1, 2) (o los valores que haya definido).
- g. Pulse el botón 'Aceptar'. La ventana visor muestra los resultados.
- h. Ahora puedes responder a la pregunta 9. En promedio, ¿a cuántos españoles conocen los dominicanos? ¿Y a cuántos españoles conocen los senegambianos? ¿Qué puedes concluir de la prueba de la $-t-$ (consulta la página siguiente para una explicación más profunda de la tabla)?
- i. Repita los pasos con la variable 'densidad' para responder a la pregunta 10.

Para poder utilizar la prueba de la $-t-$, es necesario que la variable siga una distribución normal. En la selección de datos que utilizamos en el ejemplo, no es así, como se puede averiguar con estadística descriptiva y un histograma con distribución normal (si quieres, pruébalo). En general, es importante explorar la distribución de las variables de escala antes de hacer pruebas y si hace falta, hacer alguna transformación de los datos.

Interpretación del resultado de la prueba de la $-t-$

El resultado de esta prueba es más complicado que lo explicado anteriormente. ¿Por qué? Como hemos explicado anteriormente una suposición de la prueba de la $-t-$ es que las varianzas de los dos grupos son más o menos iguales. Si las varianzas no son iguales, podemos usar un procedimiento que es un poco más complicado, pero por suerte, está integrado en SPSS. La tabla de los resultados es un poco más complicado porque necesitamos decidir si tenemos varianzas iguales o no, para decidir si usamos los resultados de un procedimiento u otro.

La primera tabla muestra el número de casos, la desviación típica, y el error típico de la media para los dos grupos.

La segunda tabla es similar a la siguiente (ver Tabla 2; *en este caso, la pregunta era si los chicos y las chicas de una cierta edad difieren en masa corporal*):

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Índice de masa corporal	Se han asumido varianzas iguales	,550	,460	,502	98	,617	,4980	,9917	-1,4699	2,4659
	No se han asumido varianzas iguales			,505	97,054	,615	,4980	,9866	-1,4601	2,4561

Tabla 11. Resultados de la prueba de la $-t-$ en SPSS.

De hecho, existen dos pruebas en la segunda tabla. Primeramente, necesitamos saber si los dos grupos tienen varianzas más o menos iguales. [Recuerda que la varianza es el cuadrado de la desviación típica, así que se

puede tener una idea de esta igualdad mirando a la tabla anterior]. La primera parte de la segunda tabla muestra 'la prueba de Levene para la igualdad de varianzas' (la prueba de la $-F-$).

La hipótesis nula de dicha prueba es que las *varianzas* de los dos grupos son iguales (es decir, la diferencia entre las varianzas es cero). Si el valor de p (en SPSS 'Sig.') es menos de ,05, quiere decir que rechazamos la hipótesis nula (o sea, concluimos que hay diferencia entre las varianzas). Si el valor de p es más que ,05, quiere decir que no rechazamos la hipótesis nula (no hay diferencia entre las varianzas). Ahora,

1. si las varianzas no son iguales (la prueba de la $-F-$ tiene $p \leq ,05$), siga la segunda fila ('no se han asumido varianzas iguales').
2. si las varianzas son iguales (la prueba de la $-F-$ tiene $p > ,05$), siga la primera fila ('se han asumido varianzas iguales').

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Índice de masa corporal	Se han asumido varianzas iguales	,550	,460	,502	98	,617	,4980	,9317	-1,4699	2,4659
	No se han asumido varianzas iguales			,505	97,054	,615	,4980	,9366	-1,4601	2,4661

Tabla 12. Resultados de la prueba de la $-t-$ en SPSS.

En el ejemplo, $p = ,46$, así que seguimos la primera fila.

Si hemos decidido si las varianzas son iguales o no, vamos a la prueba de interés, la *prueba T para igualdad de medias*. Evaluamos el valor de la $-t-$ y el valor del p de la fila correspondiente.

La hipótesis nula de la prueba del $-t-$ es que los *promedios* de los dos grupos son iguales (la diferencia entre los promedios es cero). La conclusión es:

3. Si el valor del $p \leq ,05$, rechazamos la hipótesis nula (o sea, hay diferencia entre los promedios).
4. Si el valor del $p > ,05$, quiere decir que no rechazamos la hipótesis nula (no hay diferencia entre los promedios).

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Índice de masa corporal	Se han asumido varianzas iguales	,550	,400	,502	98	,017	,4860	,9917	-1,4099	2,4059
	No se han asumido varianzas iguales			,606	97,061	,616	,1080	,0866	1,1601	2,1661

Tabla 13. Resultados de la prueba de la t en SPSS.

En el ejemplo, $t = 0,50$ con 98 grados de libertad y $p = ,62$, así que concluimos que no hay diferencia en los promedios de los dos grupos. O sea, concluimos que los chicos y las chicas **NO** difieren entre sí en masa corporal.

APÉNDICE. LA DESVIACIÓN TÍPICA / LA DESVIACIÓN ESTÁNDAR

La desviación típica (o la desviación estándar) es una medida de la dispersión de los valores en una variable en una muestra, es decir, una indicación del alcance en que los valores oscilan del promedio. Cuando una variable tiene una distribución normal, la desviación típica nos informa que el 68% de los casos se encontrarán dentro de 1 desviación de la media, y que el 95% se encontrarán dentro de 2 desviaciones de la media.

Por ejemplo, la edad de tres profesores es 33, 40 y 45. Calculamos la desviación típica de la siguiente manera.

- En primer lugar calculamos que la media de edad, es $(33 + 40 + 45) / 3 = 39$.
- Después, vamos a restar la media a las edades (ver tercera columna en la tabla siguiente) para obtener las diferencias de la media
- Después, tomamos el cuadrado de estas restas (ver ultima columna de la tabla siguiente)
- Y por fin, sumamos los cuadrados. Como muestra la tabla, esta suma es 73.
- La varianza es igual a esta suma dividido por $(n - 1)$ (el número de casos menos 1), en nuestro caso 73 dividido por $(3-1) = 73 / 2 = 36.5$
- La desviación típica es la raíz cuadrada de la variable. En nuestro caso $\sqrt{36.5} = 6.04$.

Persona (i)	X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
1	33	-6	36
2	40	1	1
3	45	6	36
Total	\bar{X} (promedio) = 39		$\sum (X_i - \bar{X})^2 = 73$

Tabla 10. Desviación típica.

La fórmula de la desviación típica es:

$$SD = \sqrt{\frac{\sum (X_i - \bar{X})^2}{(n-1)}}$$

En nuestro caso $SD = \sqrt{\frac{73}{(3-1)}} = 6,04$

La desviación típica es la raíz cuadrada de la varianza, así que la varianza (var) es el cuadrado de la desviación típica.