

CARING

Redes interactivas para la visualización de análisis sociopolíticos

Cristina Calvo (cristinacalvolopez@usal.es) & Modesto Escobar (modesto@usal.es)

Universidad de Salamanca

AECPA

Congreso Español de Ciencia Política y de la Administración
Lecciones y retos políticos de la pandemia

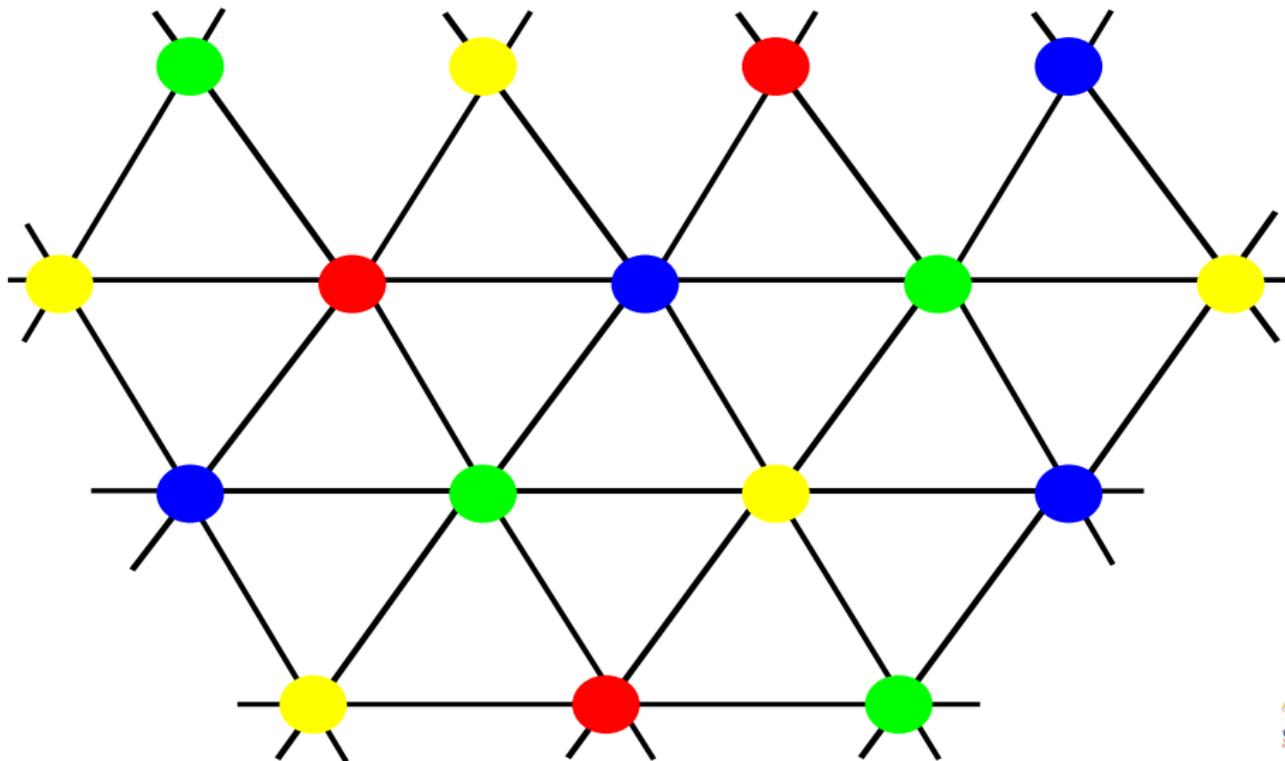
Online, 7 de julio de 2021

Contenido

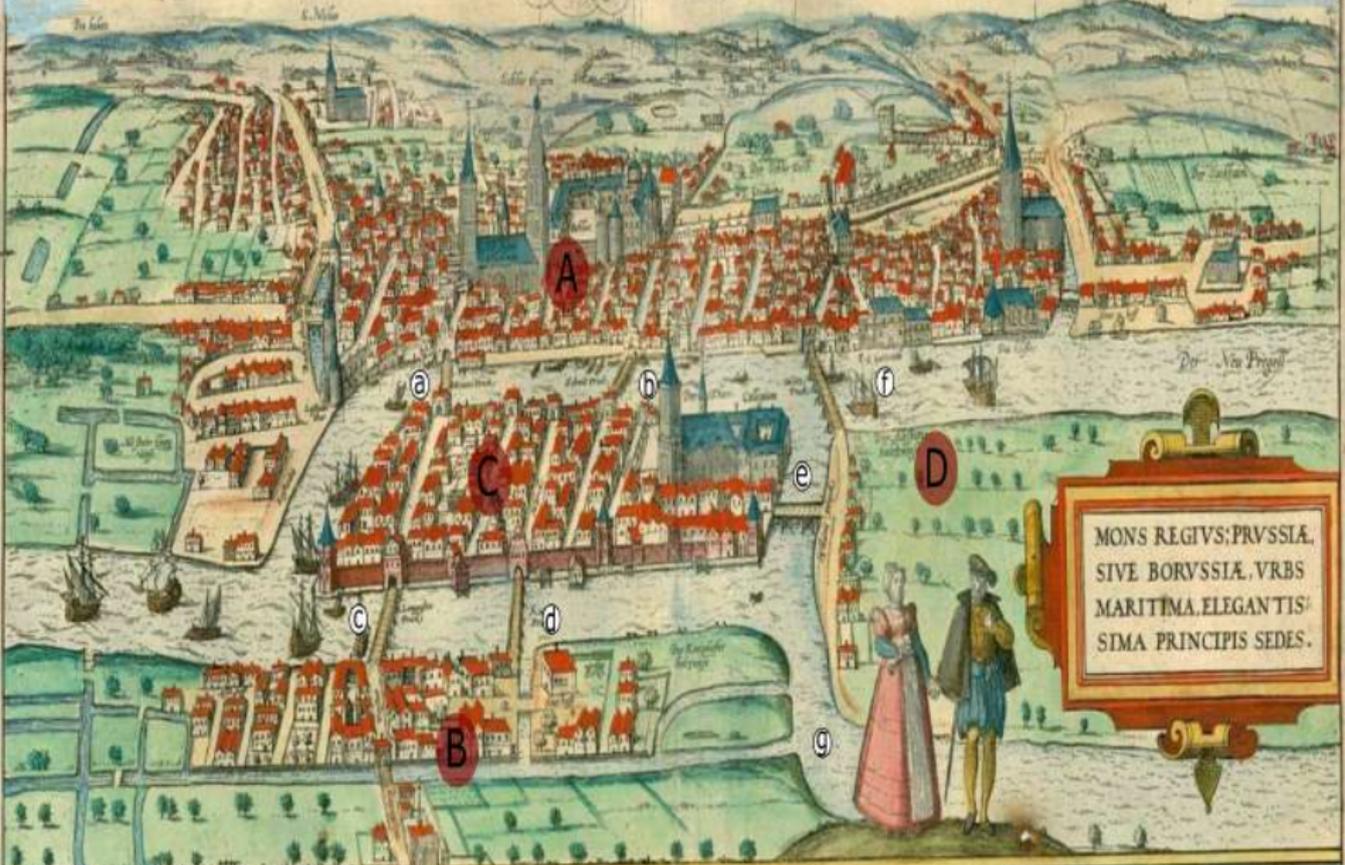
- 1 **Introducción**
 - Redes
 - Los grafos en la investigación social
- 2 CARING
 - Análisis de coincidencias
 - Datos
- 3 Software
 - El paquete netCoin de R
 - Construcción de la red
 - Modificación de parámetros
- 4 Más allá de CARING
- 5 Propuestas

¿Qué es una red?

Un conjunto de líneas entrelazadas

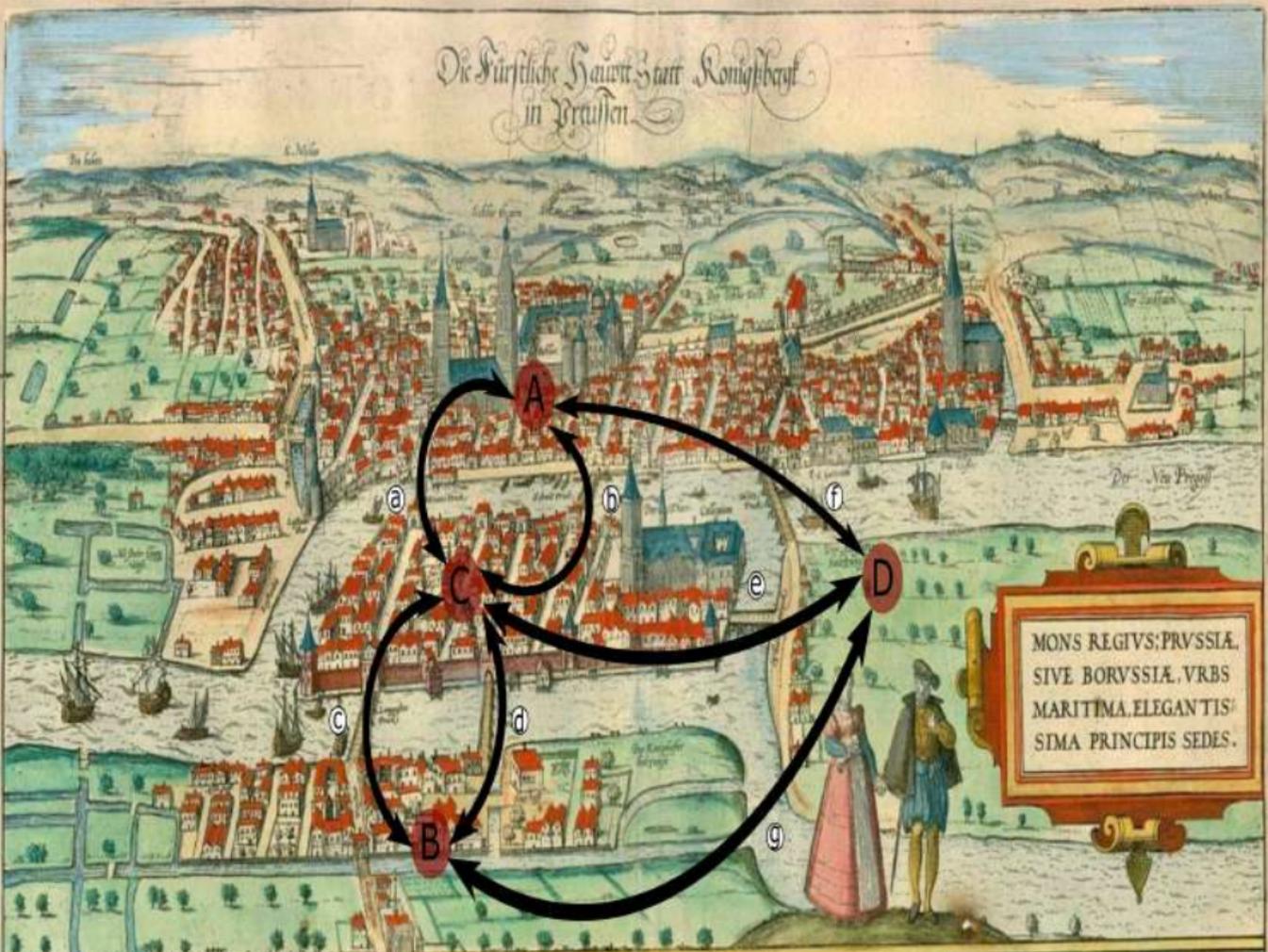


Die Fürstliche Haupt Stadt Königsberg
in Preussen

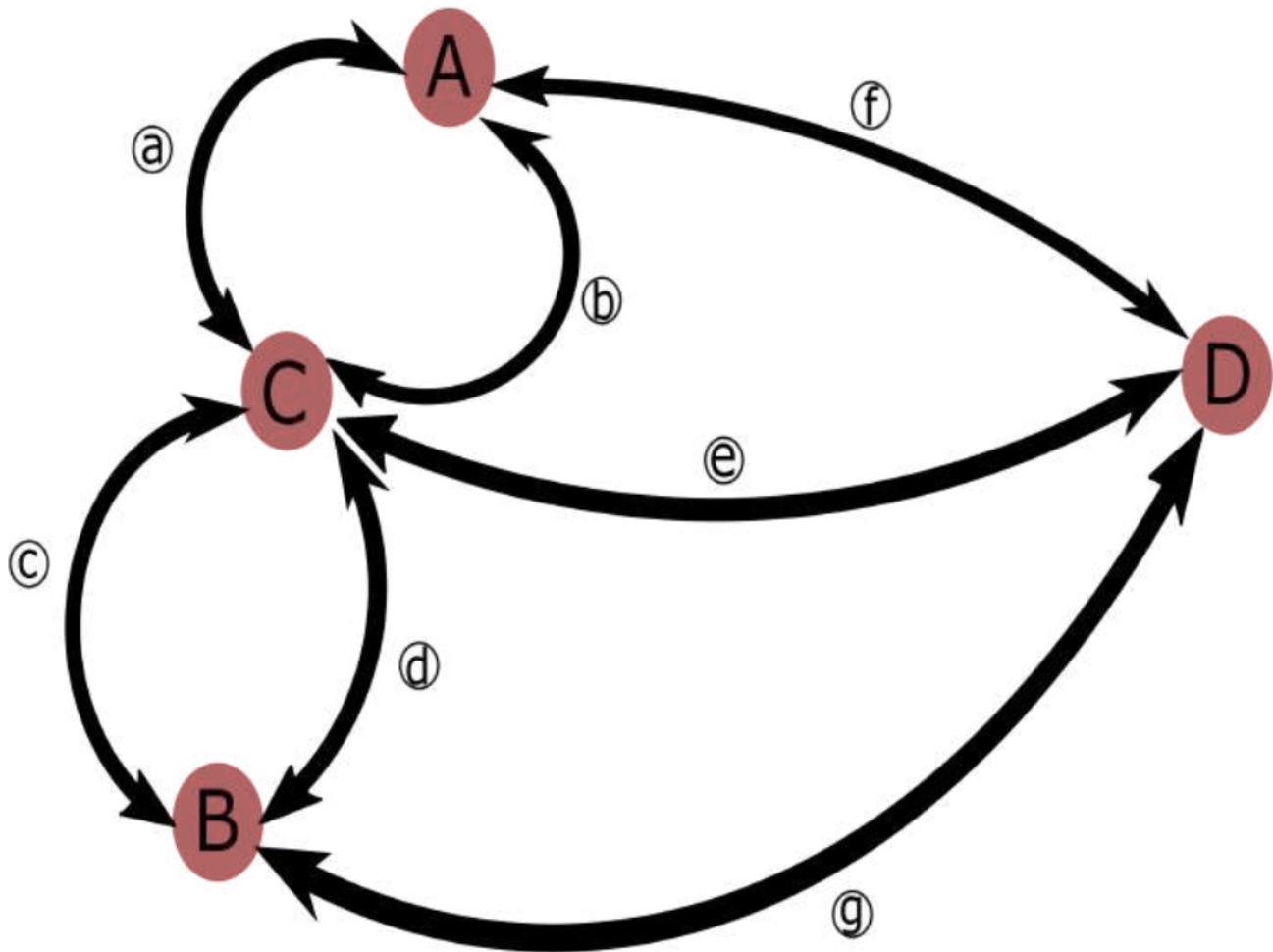


MONS REGIVS; PRVSSIA.
SIVE BORVSSIA. VRBS
MARITIMA. ELEGANTIS-
SIMA PRINCIPIS SEDES.

Die Fürstliche Haupt Stadt Königsberg
in Preussen



MONS REGIVS; PRVSSIA.
SIVE BORVSSIA. VRBS
MARITIMA. ELEGANTIS-
SIMA PRINCIPIS SEDES.



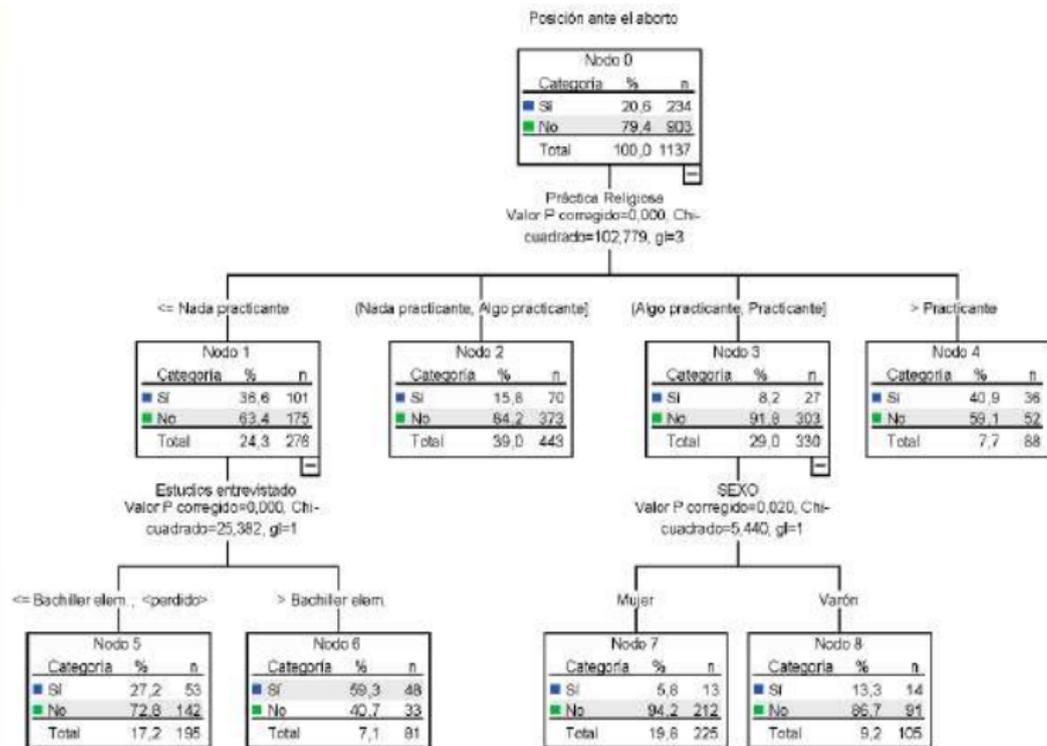
Grafos en la investigación social

Clasificación

- Grafos (Análisis de redes sociales)
- Árboles (Árboles de clasificación y regresión)
- Senderos (Análisis de senderos)
- Modelos causales y de medida (Modelos de ecuaciones estructurales)
- Redes en el análisis cualitativo (CAQDAS).
- Grafos reticulares interactivos de clasificación y regresión. (CARING)

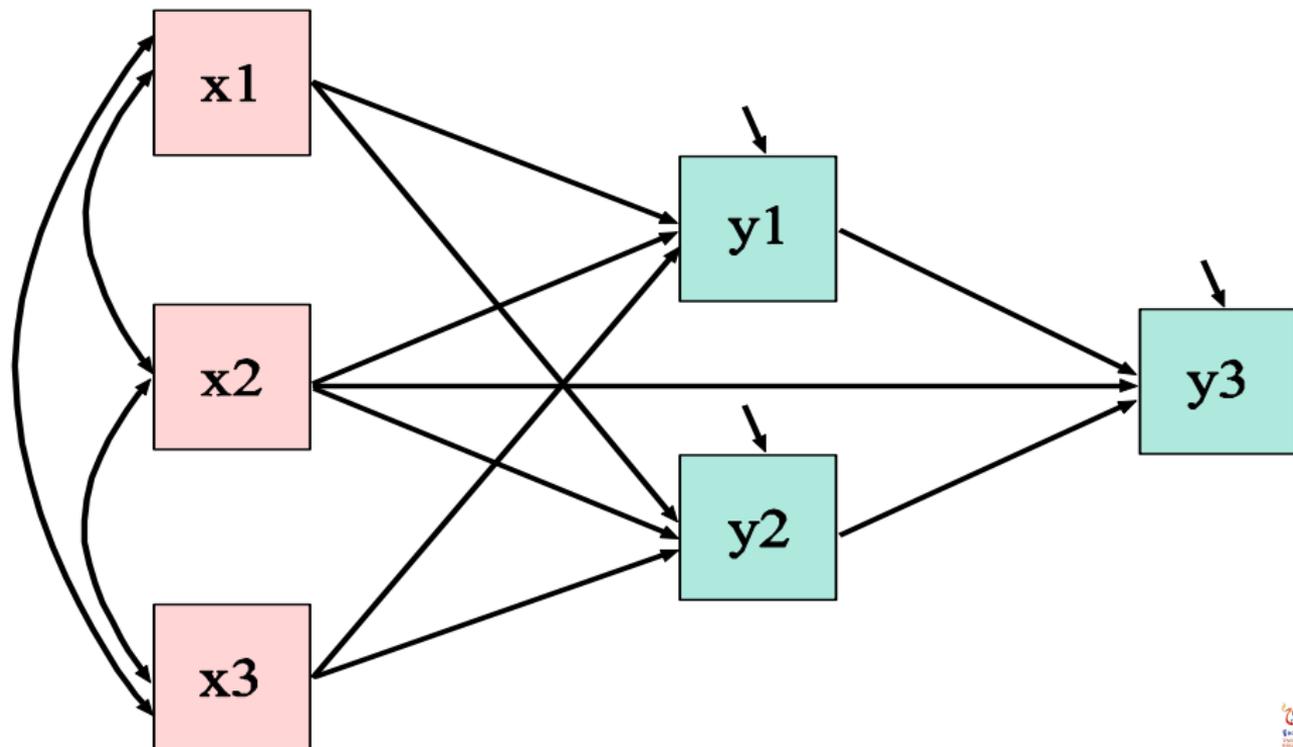
CART

Necesidad percibida de inmigrantes (Fuente: CIS-2511)



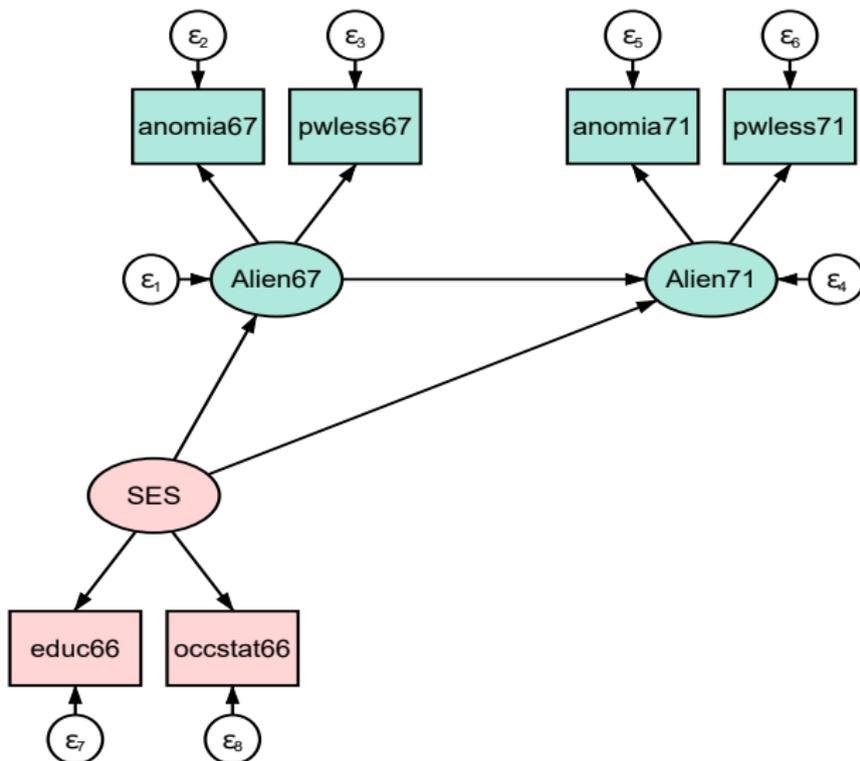
Análisis de senderos

Análisis de senderos con variables continuas (Muthén & Muthén 2017,32)



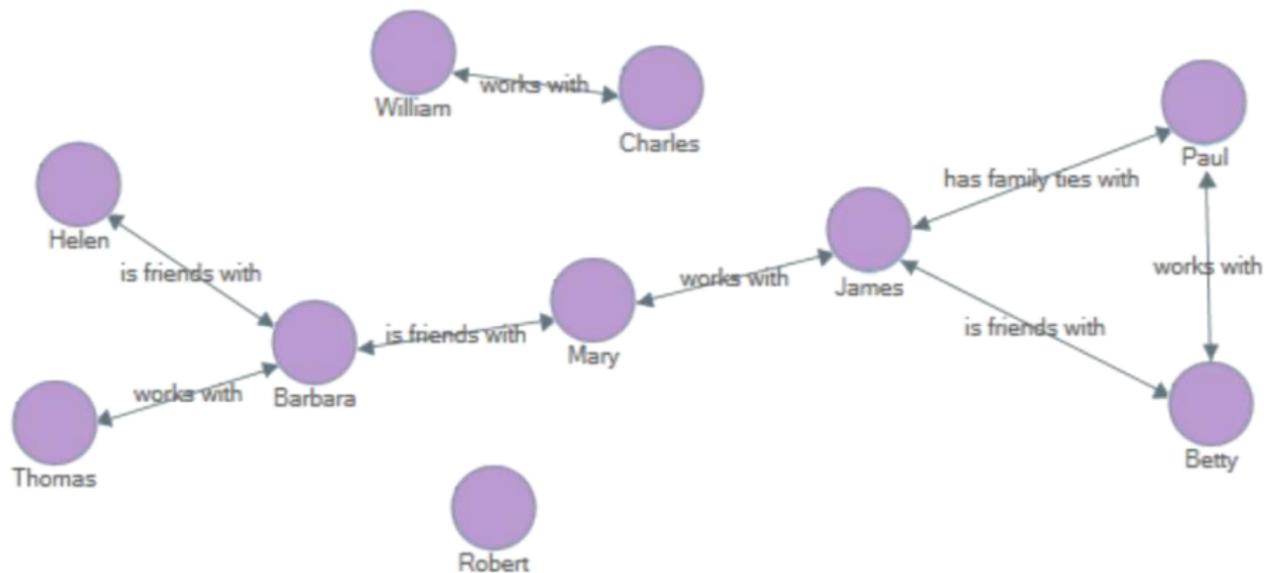
SEM

Ejemplo clásico de la alienación (Wheaton et al. 1977,84-136)



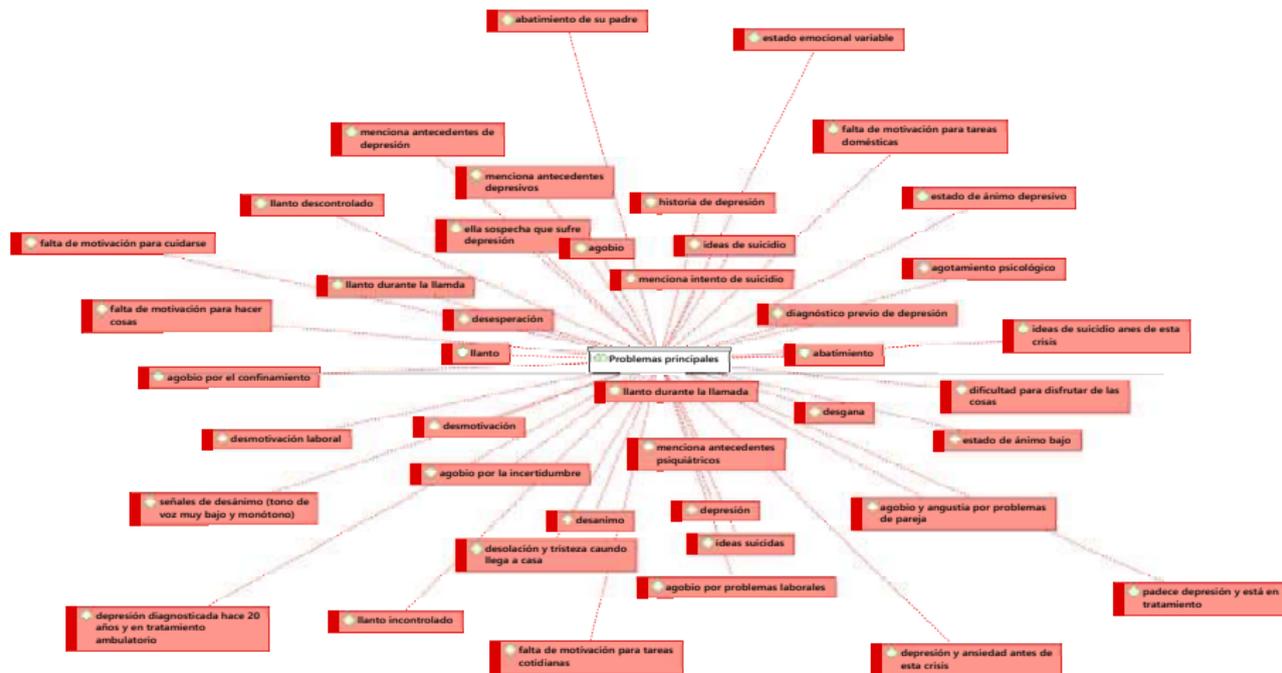
CAQDAS: I. NVivo

Compute-Assisted Qualitative Data Analysis (<http://help-nv11.qsrinternational.com/>)



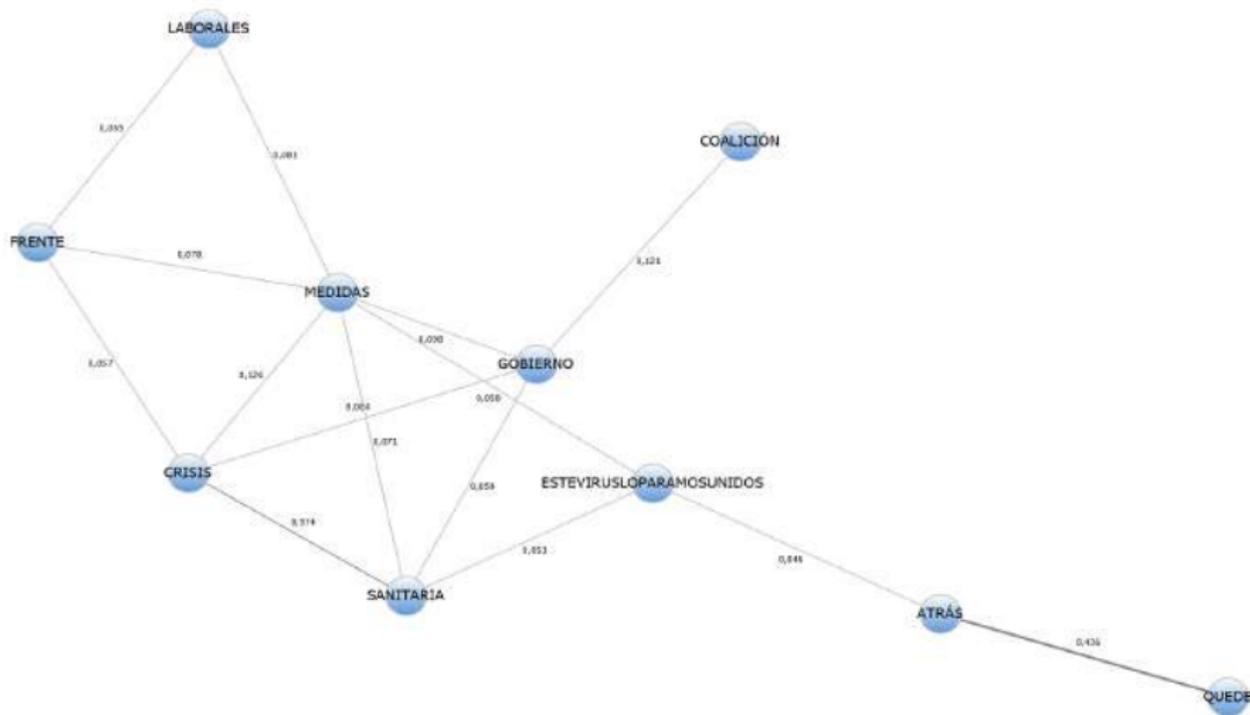
CAQDAS: II. Atlas.ti

<https://alumni.usal.es/servicio-atencion-psicologica-habilitado-la-facultad-psicologia-la-usal/>



CAQDAS: III. WordStat

Análisis de los tuits emitidos por ministros de UP (Abril-Junio 2020)



Contenido

- 1 Introducción
 - Redes
 - Los grafos en la investigación social
- 2 **CARING**
 - Análisis de coincidencias
 - Datos
- 3 Software
 - El paquete netCoin de R
 - Construcción de la red
 - Modificación de parámetros
- 4 Más allá de CARING
- 5 Propuestas

Análisis de coincidencias

Definición

- netCoin: Marco analítico visual que combina modelos estadísticos y gráficos para representar relaciones entre variables.
- El análisis de coincidencia persigue detectar qué sujetos, objetos, atributos o sucesos tienden a aparecer simultáneamente.
- Con este marco se pueden realizar tanto...
 - Modelos bivariados (análisis de coincidencias) y correlaciones.
 - Modelos multivariados como regresiones lineales, logísticas y modelos de ecuaciones estructurales (SEM).

Datos

Fuente: CIS-3.318 (Barómetro de abril de 2021)

- Vamos a ver los grafos reticulares de coincidencias con datos de un estudio sobre COVID realizado por el CIS.
- En este ejemplo se emplearán 6 variables: test y síntomas de COVID, género, edad, ccaa y tamaño de municipio
- Los síntomas y el test se considerarán variables endógenas (dependientes), mientras el resto serán tratadas como exógenas (independientes).

Preguntas endógenas del cuestionario

Síntomas y test del COVID

**P.4A2 ¿Qué síntomas principales experimentó Ud.?
(RESPUESTA MÚLTIPLE). (ENTREVISTADOR/A: MARCAR
TODOS LOS QUE DIGA).**

[SINTOMAS/ENTRE]

<i>Fiebre</i>	1
<i>Tos seca</i>	2
<i>Cansancio</i>	3
<i>Molestias y dolores musculares</i>	4
<i>Dolor de garganta</i>	5
<i>Diarrea</i>	6
<i>Conjuntivitis</i>	7
<i>Dolor de cabeza</i>	8
<i>Pérdida del sentido del olfato y/o gusto</i>	9
<i>Erupciones cutáneas</i>	10
<i>Inflamación o dermatitis de la lengua</i>	11
<i>Dermatitis en manos y pies</i>	12
<i>Dificultad para respirar o sensación de falta de aire</i>	13
<i>Dolor o presión en el pecho</i>	14
<i>Incapacidad para hablar o moverse</i>	15
<i>Ninguno</i>	97
<i>Otro/s ¿Cuál/es?</i>	96
<i>N.S./No recuerda</i>	98
<i>N.C.</i>	99

P.4A5 ¿Le hicieron a Ud. la prueba del coronavirus?

[PRUEBA/ENTRE]

<i>Sí</i>	1
<i>No</i>	2
<i>N.S.</i>	8
<i>N.C.</i>	9

P.4A6 ¿Y le diagnosticaron a Ud. finalmente infección por coronavirus?

[CORONA/ENTRE]

<i>Sí</i>	1
<i>No</i>	2
<i>N.C.</i>	9

Contenido

- 1 Introducción
 - Redes
 - Los grafos en la investigación social
- 2 CARING
 - Análisis de coincidencias
 - Datos
- 3 **Software**
 - El paquete netCoin de R
 - Construcción de la red
 - Modificación de parámetros
- 4 Más allá de CARING
- 5 Propuestas

netCoin

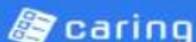
Vínculos en la red

- netCoin R package
- netCoin vignettes (R)
- netCoin paper JSS
- CARING: GUI para el uso de netCoin (funciones: surCoin y surScat)



Primera pantalla para cargar el fichero

<http://caring.usal.es>

[NUESTRO EQUIPO](#)[CONTACTO](#)

Coincidencias en Encuestas

El objetivo de este proyecto es integrar las técnicas estadísticas tradicionales con herramientas de aprendizaje automático y de análisis de redes sociales con el fin de obtener pantallas visuales e interactivas de grandes datos.

Crema tu red en solo 3 pasos



Empieza ahora



Subir archivo

Acepta archivos csv, tsv, xlsx, sav y dta

Cargar datos de demostración



Pantalla del fichero subido

<http://caring.usal.es>



¡Datos cargados!
¿Qué quieres hacer ahora?

[Añadir parámetros](#) [Crear red](#)

[Subir archivo](#)

Acepta archivos csv, tsv, xlsx, sav y dta

[Cargar datos de demostración](#)

Archivo a convertir: DA3318.dta

[Seleccionar todo](#)

[Deseleccionar todo](#)

Show 10 entries

Search:

<input checked="" type="checkbox"/> Fiebre	<input checked="" type="checkbox"/> Tos	<input checked="" type="checkbox"/> Casancio	<input checked="" type="checkbox"/> Dolores_musculares	<input checked="" type="checkbox"/> Dolor_garganta	<input checked="" type="checkbox"/> Diarrea	<input checked="" type="checkbox"/> Dolor_cabeza	<input checked="" type="checkbox"/> Anosmia	<input checked="" type="checkbox"/> Problemas
No	No	No	No	No	No	No	No	No
No	No	Sí	No	No	No	No	No	No
Sí	No	No	Sí	No	Sí	Sí	Sí	No
No	No	No	No	No	No	No	No	No

Controles específicos de CARING (en la ventana red)

<http://caring.usal.es>

Cargar datos

Añadir parámetros

Red

[Abrir en nueva pestaña](#)

[Modificar parámetros](#)

Los parámetros del gráfico

<http://caring.usal.es>

Los parámetros

Los parámetros son información dada al programa para que se comporte de uno u otro modo.

Se han clasificado los parámetros en cuatro clases:

- De variables
- De nodos
- De enlaces
- De red

Parámetros de variables

<http://caring.usal.es>

Parámetros de variables

Con ellos podemos determinar las características de nuestras variables

- Dicotómicas (solo representa una categoría de la variable)
- Exógenas o independientes (no interesa la relación entre sus categorías)
- Peso (variable cuantitativa que sirve para ponderar la muestra)

Añadir parámetros

variables	religion	ideologia	voto	PESO
variables	Dicotomías			
nodos	<input type="text" value="No practicante"/> <input type="text" value="No creyente"/> <input type="text" value="Practicante"/> <input type="text" value="NA"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
links	Exógenas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
red	Peso	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Parámetros de nodos (categorías)

<http://caring.usal.es>

Parámetros de nodos

Con ellos podemos determinar el comportamiento de nuestras categorías

- Pareados (omisión de casos perdidos por pares de categorías)
- Mínima/Máxima frecuencia para representar las categorías
- Filtro de grado (escoge solo categorías con x relaciones con otras)

Añadir parámetros

variables >	Submuestra ⓘ FALSO ▾	Pareado ⓘ VERDADER ▾	Mínimo ⓘ 200	Máximo ⓘ mayor que 0	Filtro de grado ⓘ 1
nodos >	Bipolaridad de Nodos ⓘ FALSO ▾	Color por Defecto ⓘ <input type="text"/>			
links >					
red >					

Parámetros de enlaces (cor[relaciones])

<http://caring.usal.es>

Parámetros de enlaces

Con ellos podemos determinar qué categorías se van a enlazar

- Procedimientos (cálculo de 25 posibles estadísticos)
- Criterio (estadístico para determinar el vínculo entre categorías)
- Corrección del criterio probabilístico (Bonferroni)
- Bipolaridad (para estadísticos con signo)

Añadir parámetros

variables >	Procedimientos ⓘ	Criterio ⓘ	Bonferroni ⓘ	Apoyo ⓘ	Min. vínculos ⓘ
nodos >	Gower	Z	FALSO	mayor que 0	mayor que 0
links >	Haberman				
red >	Hamann				
	Jaccard				
	Max. vínculos ⓘ	Bipolaridad de Vínculos ⓘ			
	mayor que 0	FALSO			

Estadísticos disponibles

Coefficientes de correlación, de semejanza y estadísticos de significación

Tipo	Medidas (con abreviación)
Frecuencias	Frecuencias (f), frecuencias relativas (x), frecuencias condicionales (i, ii)*
Niveles	Nivel de coincidencia (cc), Nivel probable (cp)
F. esperadas	Esperadas (e), Intervalo de confianza (con)
Matching	<i>Matching</i> (m), <i>Rogers</i> (t), <i>Gower</i> (g), <i>Sneath</i> (s), <i>Anderberg</i> (and)
Jaccard	<i>Jaccard</i> (j), <i>Dice</i> (d), <i>antiDice</i> (a), <i>Ochiai</i> (o), <i>Kulczynski</i> (k)
Rusell	<i>Rusell</i> (r)
Pearson	<i>Pearson</i> (p), <i>Haberman</i> (h), <i>Yule</i> (y), <i>Hamann</i> (ham), <i>odds ratio</i> (od)
Probabilísticas	<i>p</i> valor de <i>Haberman</i> (z), significación hipergeométrica (hyp)

* i: condicionado por fuente; ii: condicionado por destino.

Parámetros de redes

<http://caring.usal.es>

Parámetros de redes

Con ellos podemos manejar el grafo en su conjunto

- Título, nota, ayuda, lenguaje, fuente, fondo
- Dirigido, diagonal
- Comunidad, disposición y coordenadas

Añadir parámetros

variables >	Título ⓘ	Nota ⓘ	Ayuda ⓘ	Ayuda mostrada ⓘ	
nodos >	<input type="text"/>	<input type="text"/>	<input type="text"/>	FALSO v	
links >	Lenguaje ⓘ	Escalado de Fuente ⓘ	Controles ⓘ	Fondo ⓘ	
red >	es v	mayor que 0	Botones de exportación Tablas Botones de selección Barra lateral	<input type="text"/>	
	Comunidad ⓘ	Disposición ⓘ	Direccional ⓘ	Diagonal ⓘ	Mostrar Coordenadas ⓘ
	-- nada -- v	Force-Atlas v	FALSO v	FALSO v	FALSO v

Comunidades

Comunidades obtenibles mediante CARING

Comunidad	Argumento	Abreviatura
Edge-betweenness	cluster_edge-betweenness	ed
Fast-greedy	cluster_fast_greedy	fa
Label propagation	cluster_label_prop	la
Leading eigenvector	cluster_leading_eigen	le
Louvain	cluster_louvain	lo
Optimal modularity	cluster_optimal	op
Sping glass	cluster_spinglass	sp
Walktrap	cluster_walktrap	wa

Disposiciones de los grafos

Tipos de gráficos

- Grafos
 - Geométricos
 - Circular
 - Estrella
 - Rejilla
 - Espiral
 - Físicos
 - Fruchterman-Reingold
 - Kamada-Kawai
 - Atlas-force2
 - Estadísticos
 - mds (escalas multidimensionales)
 - pca (análisis de componentes principales)
 - mca (análisis de correspondencias múltiples)
 - biplot

Disposiciones posibles en CARING

Modalidades de disposición

Disposición	igraph	Abrev.
Random disposition of vertices	layout.random	ra
Rectangular grid disposition	layout.grid	gr
Circle distributed vertexes	layout.circle	ci
Star disposition of vertices	layout.star	st
Forced directed layout	layout.fruchterman.reingold	fr
Forced directed layout	layout.kamada.kawai	ka
Forced directed layout (GEM)	layout.gem	ge
Atlas Force-2		fo
Simulated annealing algorithm	layout.davidson.harel	da
Multidimensional scaling coordinates	layout.mds	md
Tidy arrangement of vertices	layout.reingold.tilford	re
Layered directed acyclic graphs	layout.sugiyama	su
Large scale graphs	layout.drl	dr
Large graph layout	layout.lgl	lg
Principal components dimensions		pca
Multiple correspondence analysis		mca

Ilustración de la representación de relaciones entre categorías

Red de síntomas, test y categorías sociodemográficas

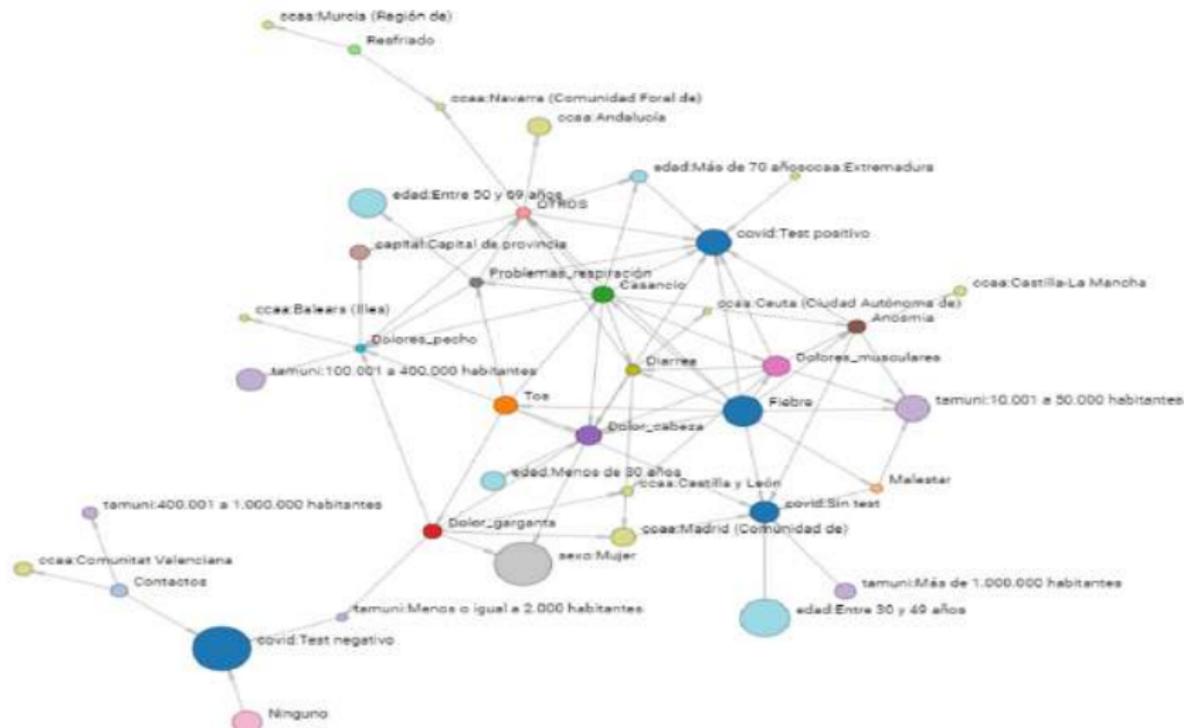
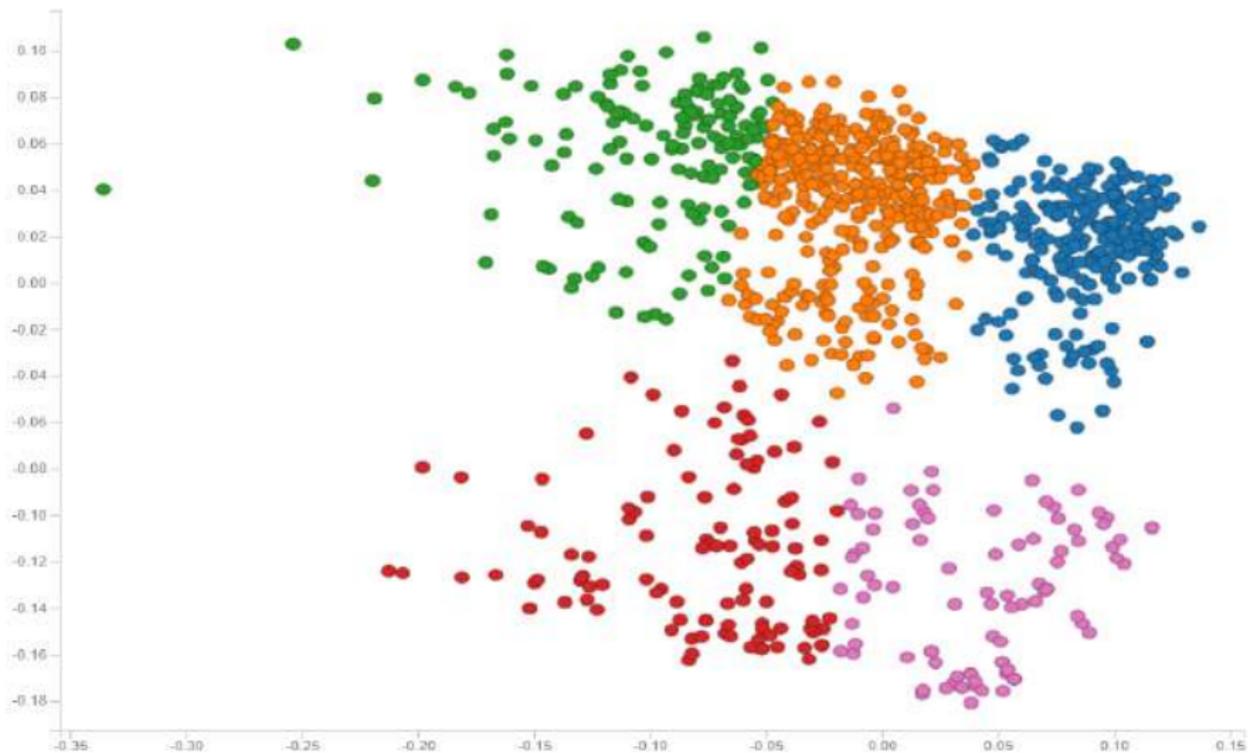


Ilustración de coincidencias

Gráfico de dispersión con factorial de correspondencias y conglomerados



Contenido

- 1 Introducción
 - Redes
 - Los grafos en la investigación social
- 2 CARING
 - Análisis de coincidencias
 - Datos
- 3 Software
 - El paquete netCoin de R
 - Construcción de la red
 - Modificación de parámetros
- 4 Más allá de CARING
- 5 Propuestas

¿Cómo hacer estos gráficos con más opciones?

netCoin

- Todo esto y más se puede hacer con netCoin, un paquete gratuito de R.
 - netCoin calcula
 - Una matriz de coincidencias a partir de una matriz de incidencias (coin).
 - Matrices de semejanzas (distancias) a partir de una matriz de coincidencias (sim).
 - netCoin representa gráficamente
 - Desde una lista de vínculos (netCoin)
 - Desde un conjunto de variables cuantitativas (netCorr)
 - Desde un conjunto de variables dicotómicas (allNet)
 - Desde un conjunto de datos en SPSS o Stata (surCoin and glmCoin)
 - Desde un objeto lavaan (pathCoin)
 - netCoin convierte
 - Una matriz de incidencias/coincidencias/semejanzas/distancias en una lista de vínculos.
 - Un gráfico netCoin en otro igrph o Pajek.

netCoin

Otros ejemplos interactivos

- Bases de datos
 - M. de Unamuno's album
 - British Library
- Encuestas
 - CIS
 - Series de TV
 - Otras iniciativas
- Coincidencias históricas
 - Historia de la ciencia
 - History de la cultura
- Análisis de contenido
 - TST
 - Blogs de mujeres
- Modelos de ecuaciones estructurales
 - Innovacion empresarial (CIE)

Contenido

- 1 Introducción
 - Redes
 - Los grafos en la investigación social
- 2 CARING
 - Análisis de coincidencias
 - Datos
- 3 Software
 - El paquete netCoin de R
 - Construcción de la red
 - Modificación de parámetros
- 4 Más allá de CARING
- 5 Propuestas

Algunas propuestas

- En las ciencias sociales deberíamos asumir un papel importante en el mundo emergente del big data.
- Hemos de ser conscientes del enorme potencial de los grandes conjuntos de datos construidos mediante las interacciones sociales (datos transaccionales, tuits, ...)
- Podemos contribuir aportando métodos robustos y visuales en al análisis de datos e interpretaciones ilustrativas.

Algunas propuestas

- Proponemos gráficos **reticulares** de coincidencias (y regresión) como un marco analítico visual **interactivo**.
 - Representar las variables o categorías como nodos.
 - Dimensionar los nodos por su importancia (frecuencia) en lugar de por su grado.
 - Dotar a los nodos de significado mediante colores, formas o imágenes además de texto.
 - Representar como enlaces las relaciones o asociaciones (positivas) entre variables o categorías.
 - Dimensionar los enlaces en función de su significación (valores z o t).
- Todo esto (y mucho más) se puede hacer con el paquete netCoin, gratuito en el CRAN de R.
- Sin embargo, para divulgar este tipo de gráficos interactivos se ha diseñado una página web (CARING), donde se suben los datos para obtener la red automáticamente.
- Aunque sea automático el proceso, hay muchos parámetros que pueden ser cambiados.

Última diapositiva

Agradecimiento

Muchas gracias.
modesto@usal.es

Bibliografía

Análisis reticular de coincidencias

Bibliografía sobre análisis reticular de coincidencias

- Escobar, M. (2009). Redes semánticas en textos periodísticos: propuestas técnicas para su representación. *Empiria*, 17, 13-39.
- Escobar, M., y Gómez Isla, J. (2015). "La expresión de la identidad a través de la imagen: los archivos fotográficos de Miguel de Unamuno y Joaquín Turina". *Revista Española de Investigaciones Sociológicas*, 152, 23-46.
- Escobar, M. (2015). "Studying Coincidences with Network Analysis and Other Multivariate Tools". *The Stata Journal*, 15(4), 1118-1156.
- Escobar, M. (2016). "Ensayo sobre las coincidencias". En A. Almarcha, P. González, y L. Román (Eds.), *Donde la Sociología te lleve*. A Coruña: Universidad de A Coruña.
- Escobar, M., y C. Tejero (2018). "El análisis reticular de coincidencias". *Empiria*, 39, 129-148.
- Escobar, M. y L. Martínez (2020) "Network Coincidence Analysis: the netCoin R Package". *International Journal of Statistics*, 93(11), 1-32.