

Uso de herramientas tecnológicas para la Gestión de Conocimiento.

El caso RSE en San Lorenzo

Alicia Marchese¹, Ramiro Ingrassia²

¹ Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario, Rosario, Argentina, amarches@fcecon.unr.edu.ar

² Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario, Rosario, Argentina, ringrassia@fcecon.unr.edu.ar

Abstract: El presente trabajo versa sobre el relevamiento realizado en la ciudad de San Lorenzo relativo al conocimiento de la población en lo que hace a la Responsabilidad Social Empresaria y el uso de Tecnologías de Conocimiento. Se aplican diversas herramientas de análisis de datos y generación de conocimiento a partir de la experiencia de campo. Se concluye acerca de las ventajas y desventajas de cada una, proponiendo cursos de acción futura para culminar con la tarea planteada.

Palabras Claves: Gestión de Conocimiento - Cultura Urbana - Responsabilidad Social Empresaria - Clustering

1 Introducción

El presente aporte surge de las experiencias de investigación en Gestión de Conocimiento, utilizando herramientas tecnológicas. En ese recorrido se han incorporado conceptos de Administración de los Recursos Humanos y se han estudiado diferentes medios de procesamiento de la información y de generación de conocimiento. Sabemos que éste se convierte en un acervo común, social, que surge del aporte de las individualidades. Por otra parte, el objeto de nuestras investigaciones consiste en la reproducción del conocimiento y su aplicación a actividades innovadoras, y existe una tendencia creciente entre el público y los académicos a discriminar entre empresas que crean valor con mayor o menor grado de responsabilidad social.

La responsabilidad social empresarial presenta diferentes aspectos que incumben a distintas disciplinas. Desde una perspectiva sistémica, todas éstas pueden realizar su aporte y mejorar el rendimiento común. No obstante, el protagonismo de la Administración de Recursos Humanos y de las Ciencias Informáticas, denota una mayor dinámica que las lleva a cuestionarse y dar respuestas de manera más expeditiva que otras.

Más adelante, al abordaje se incorporan otras perspectivas disciplinarias, como la Administración Pública y la Economía. Esta “mirada” surge a partir de la identificación de los actores sociales involucrados: empresas – sociedad – estado. Para esta investigación específica se identificaron tres instituciones: Universidad – Estado – Empresas.

Y en este punto, debatido en la Economía, sobre quién debe “pagar” los costos de las externalidades, se asienta una propuesta para que tanto el Estado como el sector privado puedan solventar tales costos. Sin embargo, este punto queda abierto a distintas opiniones.

A modo de cierre, deseamos concluir de las experiencias enunciadas que el proceso de Generación de Conocimiento puede y debe ser profundizado, capitalizando los diversos intentos que desde las distintas disciplinas se aportan, buscando la participación de los usuarios y aportando los diferentes puntos de vista.

Durante el año pasado, hemos abordado el relevamiento preliminar y la construcción de indicadores que den cuenta y permitan realizar seguimiento y diagnóstico de este aspecto de la vida económica.

De esta manera hemos contrastado los indicadores presentados inicialmente por las empresas, con la información presentada por los factores intervinientes, sobre todo, los habitantes de las zonas involucradas.

En el caso estudiado, se establecieron las pautas para diseñar una estructura que permita implementar minería de datos, para brindar al Estado la información necesaria para el diseño de sus políticas públicas.

2 Objetivos de este trabajo

En esta etapa, habiendo concluido la primera etapa del relevamiento, nos proponemos:

- Evaluar los resultados de las tareas realizadas.
- Analizar las herramientas disponibles, para elegir las más eficientes para la elaboración y publicación de resultados.
- Proponer los indicadores - objetivos para la etapa siguiente de trabajo.

3 Presentación del Caso

La ciudad de San Lorenzo está situada en el centro de la Provincia de Santa Fe, con una superficie de 32 Km². y según datos del Censo Nacional de Población y Vivienda al año 2001, posee una población de 43.520 habitantes.

La importancia a nivel agrícola industrial de la ciudad comienza con el nacimiento del cordón industrial que luego de atravesar la crisis de los '90 vuelve a tomar preponderancia en la región, y por los guarismos económicos que maneja, en el país mismo.

La composición de su estructura económica presenta marcada diferencia entre el desarrollo de los diferentes sectores y en el tamaño de las empresas que lo componen, tal como fue reflejado en el Documento Base Municipalidad de San Lorenzo "Programa Más y Mejor Trabajo" de la Universidad Nacional de Rosario por la Facultad de Ciencias Económicas y Estadística y la Facultad de Ciencia Política y Relaciones Internacionales.

4 Metodología

En esta etapa de la investigación se relevó la información concerniente a la opinión de la comunidad sobre la actividad comercial de las empresas. La articulación del relevamiento se realizó con el aval de la Municipalidad de San Lorenzo y la participación de las Asociaciones Civiles Vecinales.

Para relevar la información de la comunidad se realizaron dos encuestas: una dirigida al presidente de la asociación vecinal y otra, de carácter domiciliaria, al Jefe de hogar. La primera fue bajo la modalidad censal, mientras que para la segunda se procedió a tomar una muestra. El tamaño muestral de los domicilios fue estimado en 1.527, el 10% de los domicilios existentes en la ciudad de San Lorenzo en función de las proyecciones del Censo Nacional de Vivienda y Población 2001.

4.1 Diseño de los Formularios

El objetivo de fue obtener una contrastación entre la opinión de los ciudadanos y los jefes vecinales con la información social de las empresas. Actualmente, en la ciudad de San Lorenzo, las empresas no están obligadas a presentar Balances Sociales. Excepto el capítulo referido a la situación laboral que exige la ley nacional n° 25.877. Los resultados del relevamiento realizado a las publicaciones en la Web se encuentran (Marchese, et. al., 2010)

En consecuencia, para la formulación de las preguntas de la encuesta se utilizó una normativa estandarizada representativa de todos los aspectos de la Responsabilidad Social Empresaria (RSE). Por tal motivo, se estimó que el estándar más abarcativo sobre esta temática son los de indicadores del Global Reporting Initiative (GRI, 2002-2006). Los protocolos (o capítulos) de esta normativa describen minuciosamente los indicadores a emplearse. Mientras que en otras normativas, se describe la categoría de indicadores pero no se especifica ninguno de estos. Tampoco se los relaciona con alguna norma contable, en el caso que fuera necesario. *El supuesto subyacente en este trabajo es que el Balance Social es emitido según los protocolos de la GRI.* Se tomaron los protocolos (capítulos) de indicadores concernientes a las siguientes áreas:

- Económicos,
- Medioambientales,
- Derechos humanos,
- Laborales,
- Responsabilidad sobre productos,
- Sociedad.

De esta manera, la formulación de las preguntas siguió el siguiente razonamiento, como por ejemplo:

Tomando al indicador SO81: *Valor monetario de sanciones y multas significativas y número total de sanciones no monetarias derivadas del incumplimiento de las leyes y regulaciones*, se consulta:

17. ¿Ha denunciado o participado en denuncias a empresas por privarle de alguno de sus derechos civiles?

SI NO NS/NC

En caso afirmativo, nombre las empresas

En la formulación del cuestionario participaron profesionales de distintas disciplinas como: arquitectos, médicos y profesionales de las ciencias económicas. Se concluyó en la confección de dos formularios, uno para vecinalistas y otro para hogares.

De esta manera, la encuesta dirigida a los presidentes de asociaciones vecinalistas cuenta con 14 preguntas divididas en las siguientes categorías:

- Niveles de Infraestructura del barrio
- Niveles de espacios públicos – espacios verdes – arbolado
- Tratamiento de desechos y basura doméstica
- Transporte

¹ SO: Protocolo de indicadores de Sociedad.

Vivienda Pública

Patrimonio histórico

Participación social barrial

Por su parte, la encuesta de opinión sobre responsabilidad social dirigida a los jefes de hogar de los domicilios de San Lorenzo, consta de 38 preguntas agrupadas en dos partes y cada una de ellas dividida en los siguientes bloques:

Parte I: Grupo Familiar

Se describe a la familia que habita el domicilio.

- ✓ Bloque A: Descripción de la unidad familiar

Se describe a la familia en cuanto a cantidad de miembros por edad y sexo, enfermedades que padecen y nivel socio-económico.

- ✓ Bloque B: La unidad familiar y su entorno urbano

Describe la relación de la familia con las empresas de la zona.

Parte II: Jefe del Hogar

- ✓ Bloque C: Descripción socio-laboral

Preguntas relacionadas al trabajo que desempeña el jefe del hogar.

- ✓ Bloque D: Interacción con los vecinos del barrio

Preguntas dirigidas a evaluar la relaciones entre los vecinos y su participación en las asociaciones vecinales de su barrio.

- ✓ Bloque E: Aspectos como consumidor

Preguntas referidas al jefe del hogar como consumidor, la finalidad es saber si los ciudadanos están conscientes de la cohesión que pueden lograr y la visión que tienen de las empresas productoras de alimentos.

- ✓ Bloque F: Niveles de infraestructura del barrio

Preguntas dirigidas a al jefe de hogar para evaluar las diferencias urbanísticas que se tienen en un mismo barrio o vecinal.

5 Resultados

5.1 Población en estudio

De acuerdo a los datos del Censo Nacional de Población disponible, se convocó a un asesor estadístico para el diseño muestral. Posteriormente se procedió a la definición de los domicilios y a corrección de la cartografía.

Contrariamente a lo planificado, los miembros de la comunidad respondieron con dispar motivación a la convocatoria. Luego de numerosas reuniones, se logró un conjunto de respuestas válidas en número inferior al previsto, si bien representativo de las áreas geográficas más interesantes. De igual forma, se analizaron las causas de los rechazos, por resultar de interés para este tipo de análisis con alto componente político – social y cualitativo.

Las vecinales ausentes, pueden clasificarse en:

No aceptan colaborar en la repartición de la encuesta debido a que:

- ✓ Los objetivos planteados por el proyecto no son los que persigue la vecinal

No pueden realizar el reparto de la encuesta:

- ✓ Aquellas vecinales en donde el presidente es el promotor principal (unipersonales) y este tiene problemas que paralizan la mayor cantidad de actividades.

No tienen interés en la encuesta:

- ✓ Se trata de vecinales a las que se convocó varias veces (más de 4), mediante diálogo con el presidente o vicepresidente. Las reuniones pactadas fueron suspendidas o postergadas sucesivamente.

No funciona la vecinal:

- ✓ La vecinal tiene problemas de funcionamiento o está acéfala

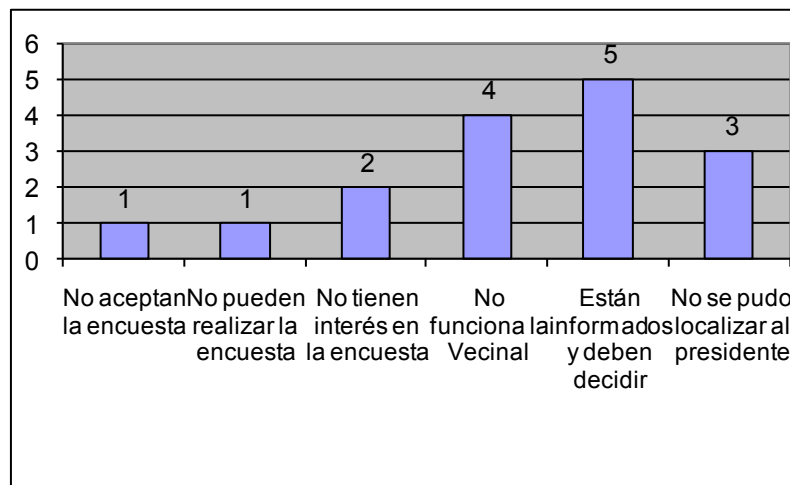
Están informados de la encuesta y la comisión está evaluando su colaboración:

- ✓ Se buscó contactar al presidente de la vecinal o algún miembro de la comisión y no han dado una respuesta. Una de las razones es que la comisión no está reuniéndose semanalmente.

No pudo localizarse al presidente de la vecinal:

- ✓ Las vecinales funcionan pero no se pudo comunicar con el presidente de la vecinal.

Gráfico 1: Motivos de la no participación en la encuesta



Fuente: Elaboración propia

Se culminó esta etapa exploratoria, con los siguientes resultados parciales:

Tabla 1: Totales de encuestas por barrio

Vecinal	Tamaño Muestral			Encuestas Procesadas	
	Estimado	Real	Diferencia	Efectivas	% Efect./Real
1° de Julio	34	42	8		0,00
2 de Abril	31	22	-9	14	63,64
Bouchard	74	94	20		0,00
Del Combate	94	116	22	35	30,17
Díaz Vélez	137	106	-31	12	11,32
El Pino	55	66	11		0,00
José Hernández	102	120	18		0,00
Mitre	98	110	12	38	34,55

Tabla 1: Totales de encuestas por barrio (cont.)

Vecinal	Tamaño Muestral			Encuestas Procesadas	
	Estimado	Real	Diferencia	Efectivas	% Efect./Real
Morando	20	3	-17	3	100,00
Rivadavia	34	54	20	47	87,04
San Eduardo	11	3	-8		0,00
San Martín	57	50	-7	19	38,00
Villa Felisa	46	55	9	42	76,36
Totales	793	841	48	210	24,97

Fuente: Elaboración propia

No obstante, los resultados serán un valioso punto de partida para la profundización del relevamiento, ya que las vecinales interesadas, se encuentran en torno a las industrias más determinantes.

Tabla 2: Distribución de empresas por rubro y barrio

Barrio	Rubro	Cant_Empresas
2 de Abril	Industria	2
2 de Abril	servicios	1
Total 2 de Abril		3
Díaz Velez	Comercio	2
Díaz Velez	Industria	4
Díaz Velez	servicios	10
Total Díaz Velez		16
El Combate	Comercio	1
El Combate	Industria	1
El Combate	servicios	19
Total El Combate		21
Mitre	servicios	3
Total Mitre		3
Rivadavia	servicios	2
Total Rivadavia		2
San Martín	Industria	1
San Martín	Servicios	3
Total San Martín		4
Villa Felisa	servicios	2
Total Villa Felisa		2
Total general		51

Fuente: Elaboración propia

5.2 Análisis de Conglomerados (cluster)

La opinión del jefe de hogar sobre la responsabilidad social de las empresas está plasmada en los bloques D, E y F de la encuesta. De esta manera, se identifica como variables dependientes a aquellas que se encuentran en dichos apartados. Mientras

que las variables independientes son aquellas que describen el entorno familiar-económico-social del jefe de hogar. Estas últimas se encuentran reflejadas en los bloques A, B y C de la encuesta.

En esta primera instancia, se desconoce las características o atributos que influyen de manera decisiva en la opinión de los jefes de hogar. El objetivo que se plantea es realizar una investigación exploratoria dentro de las variables independientes. Se intentará encontrar distintos grupos familiares que en ellos se unifiquen características comunes basándose en las observaciones realizadas. Por tal motivo, se decide realizar un análisis de conglomerados para que cada grupo (conglomerado o cluster) sea homogéneo respecto a las variables utilizadas para caracterizarlo y tales grupos sean lo más distintos posibles uno de los otros.

Como no se tiene una presunción de cuál puede ser el número de grupos naturales se decide realizar una experiencia combinada entre los métodos jerárquicos y no jerárquicos. Con el método jerárquico se espera obtener una aproximación de la cantidad de grupos, para luego, optimizar la división con el método no jerárquico.

5.2.1 Método Jerárquico

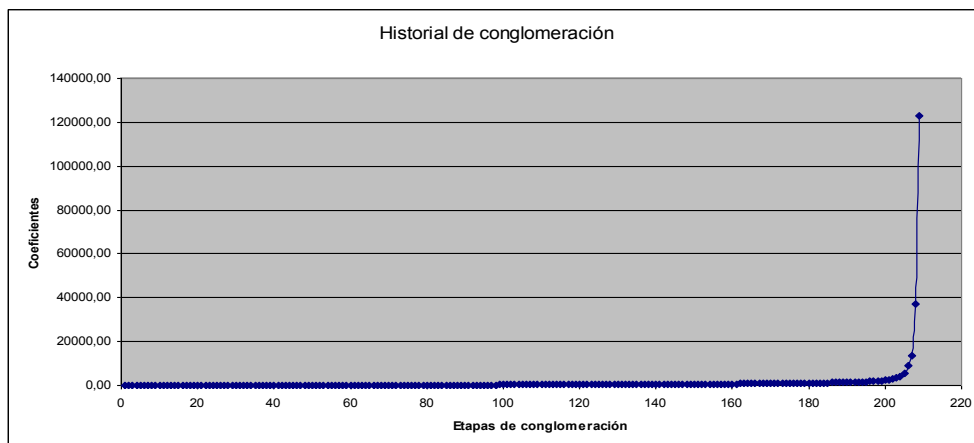
En el análisis por el método jerárquico se utilizaron las observaciones de los siguientes atributos:

- ✓ Fem_0a15
- ✓ Fem_16a30
- ✓ Fem_31a45
- ✓ Fem_46a60
- ✓ Fem_61yMas
- ✓ Masc_0a15
- ✓ Masc_16a30
- ✓ Masc_31a45
- ✓ Masc_46a60
- ✓ Masc_61yMas
- ✓ Trabajan
- ✓ Quieren_Trabajar
- ✓ Subsidios
- ✓ Alergia
- ✓ Cancer
- ✓ Broncoespasmos
- ✓ Jefe_Edad

El método de agrupamiento utilizado es el de Ward pues en comparación al resto de los métodos (centroide, vecino más cercano, vecino más lejano y vinculación promedio) agrupó a las observaciones en un número razonable de grupos. Además, porque tiende a maximizar la homogeneidad dentro de cada grupo.

Como resultado del análisis se obtiene el dendrograma que permite ver los distintos conglomerados en las sucesivas etapas de agrupamiento. Dado el tamaño del dendrograma y del historial de conglomeración, pues existen 210 observaciones, se expone el gráfico n° 1. En este, se intenta resumir la información de los anteriores. En el eje de las x, se muestran las sucesivas etapas en las cuales se fusionan los grupos. En el eje de las y, se muestran los coeficientes de aglomeración que indican la homogeneidad que hay entre cada grupo identificado.

Gráfico n° 1: Coeficientes según las etapas de conglomeración

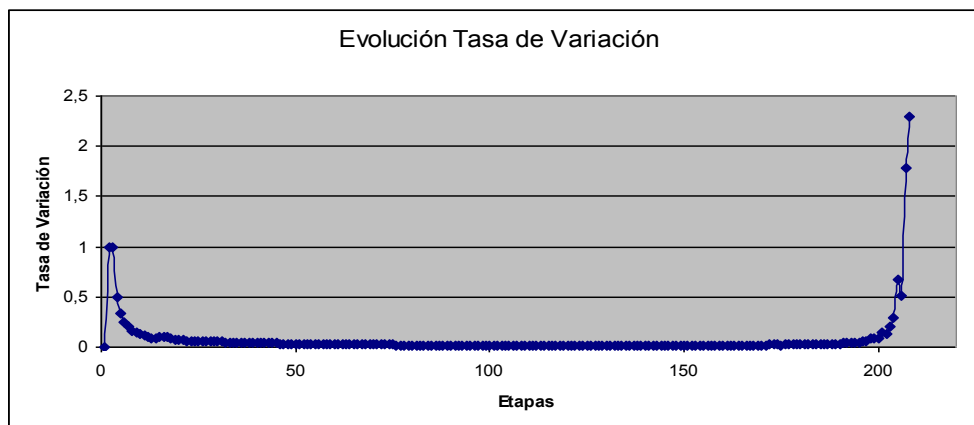


A medida que aumentan las etapas, mayores grupos son identificados. El coeficiente en cada etapa no es muy diferente al de la etapa anterior. Por tal motivo, el comportamiento de este último es lineal y se interpreta que los grupos que se van identificando son similares entre sí. Debe aclararse que la recta no es paralela al eje de las x, esto se debe a la cantidad de observaciones. En realidad, la recta es ascendente pues el coeficiente va acumulando las “similitudes” de cada grupo encontrado. Sin embargo, cuando se produce un “salto” significativo se interpreta que la fusión en dicha etapa es forzada pues se encontraron grupos disímiles entre sí. Por tal motivo, a partir de la etapa 200 comienza a ser ascendente perpendicular al eje de las x. En esas etapas es donde la fusión de grupos debe terminar pues estos son muy distintos entre sí.

A partir de los coeficientes de conglomeración, se puede calcular la tasa de variación entre coeficientes de cada uno de las etapas. Cuando dicha tasa sea drásticamente superior a la anterior, será el momento de detener las fusiones. De esta manera se “refuerza” la idea anteriormente expuesta.

En gráfico n° 2 la evolución de la tasa de variación durante las sucesivas etapas de agrupación. En el gráfico anterior puede visualizarse que desde un comienzo hay un salto repentino para después descender y estabilizarse alrededor de 0. Al final, surge otro salto brusco cercano entre 1,5 y 2, y continúa ascendiendo. En el primer salto, se ubica en la etapa de fusión 8, en donde se identificaron 200 grupos.

Gráfico n° 2: Tasa de Variación de los Coeficientes de Conglomeración



Mientras que en el segundo salto, se da en la etapa 207, en donde existen 3 conglomerados. Este último “salto” coincide con el que sufre la recta de los coeficientes. De esta manera, el número de conglomerados se podría establecer en 3. Sin embargo, para una mejor conclusión se pueden calcular los estadísticos descriptivos para evaluar la consistencia del agrupamiento en 3 conglomerados.

En las siguientes tablas se exponen las estadísticas descriptivas para cada uno de los atributos en los conglomerados identificados. De esta manera, se podrían contrastar con algunas observaciones realizadas durante el relevamiento.

Conglomerado 1:

Cantidad de observaciones: 98

Tabla 3: Estadísticas descriptivas para el Conglomerado 1

	Mínimo	Máximo	Media	Desv. típ.
Fem_0a15	0	5	,79	,987
Fem_16a30	0	4	,71	,732
Fem_31a45	0	1	,60	,492
Fem_46a60	0	1	,11	,317
Fem_61yMas	0	2	,04	,245
Masc_0a15	0	4	,79	,876
Masc_16a30	0	3	,60	,714
Masc_31a45	0	1	,52	,502
Masc_46a60	0	1	,13	,341
Masc_61yMas	0	1	,05	,221
Trabajan	0	4	1,52	,828
Quieren_Trabajar	0	3	,50	,763
Subsidios	0	2	,17	,407
Alergia	0	5	,83	1,227
Cancer	0	1	,06	,241
Broncoespasmos	0	5	,56	1,026
Jefe_Edad	19	50	37,47	7,297

En este conglomerado, las familias estarían formadas por cónyuges adultos-jóvenes con hijos pequeños. El alto valor en promedio que tienen las personas que trabajan presupone que ambos cónyuges trabajan. Además, la edad media es de 37 para el Jefe de familia, lo cual podría ratificar lo anterior. También, podría estimarse que la presencia de jóvenes aumenta considerablemente los valores de las enfermedades (Alergia y Broncoespasmos).

Conglomerado 2:

Cantidad de observaciones: 51

Al analizar el siguiente conglomerado se visualiza que en estas familias preponderan notoriamente las mujeres menores de 16 años. La unidad familiar

estaría conformada por cónyuges jóvenes con hijos pequeños. También un alto número de personas que quieren trabajar, siendo el mayor de todos los conglomerados. En consecuencia, podría ser el caso de cónyuges jóvenes uno de los cuales trabaja. Como en el conglomerado anterior, podría estimarse que la presencia de jóvenes aumenta considerablemente los valores de las enfermedades (Alergia y Brocoespasmos).

Tabla 4: Estadísticas descriptivas para el conglomerado 2

	Mínimo	Máximo	Media	Desv. típ.
Fem_0a15	0	21	1,12	3,004
Fem_16a30	0	4	,59	,853
Fem_31a45	0	2	,43	,608
Fem_46a60	0	2	,25	,483
Fem_61yMas	0	1	,16	,367
Masc_0a15	0	4	,61	1,115
Masc_16a30	0	3	,51	,731
Masc_31a45	0	2	,37	,528
Masc_46a60	0	1	,22	,415
Masc_61yMas	0	1	,24	,428
Trabajan	0	3	1,12	,791
Quieren_Trabajar	0	3	,69	,927
Subsidios	0	2	,29	,540
Alergia	0	4	,94	1,047
Cancer	0	2	,14	,401
Broncoespasmos	0	4	,88	,973
Jefe_Edad	0	0	,00	,000

Conglomerado 3:

Cantidad de observaciones: 61

En este conglomerado, las familias parecen estar mayormente compuestas por mujeres jubiladas y muy pocos jóvenes y adultos. Otro atributo que se destaca dentro de este conglomerado es el de personas que trabajan, el cual puede corresponder a los restantes miembros de la familia. Igualmente, este atributo tiene el valor más bajo del resto de los conglomerados.

Tabla 5: Estadísticas descriptivas para el Conglomerado 3

	Mínimo	Máximo	Media	Desv. típ.
Fem_0a15	0	2	,23	,589
Fem_16a30	0	4	,43	,718
Fem_31a45	0	2	,21	,487
Fem_46a60	0	1	,59	,496
Fem_61yMas	0	2	,39	,525
Masc_0a15	0	3	,28	,636
Masc_16a30	0	3	,38	,637
Masc_31a45	0	1	,13	,340
Masc_46a60	0	2	,41	,528
Masc_61yMas	0	1	,26	,444
Trabajan	0	5	1,30	1,256
Quieren_Trabajar	0	3	,39	,781
Subsidios	0	1	,13	,340
Alergia	0	3	,79	,915
Cancer	0	2	,13	,386
Broncoespasmos	0	4	,54	,808
Jefe_Edad	50	86	62,66	10,211

Las características que se resaltan en los conglomerados 1 y 3 coinciden con lo percibido durante el relevamiento a las familias. Además, en los conglomerados 1 y 2 parece respetarse una observación realizada por distintos vecinalistas: la presencia de jóvenes y la cantidad de enfermos (en particular la alergia y el broncoespasmos). La agrupación en menos o más conglomerados distorsionaría estas últimas observaciones.

En consecuencia, se establece como 3 al número de conglomerados por dos razones:

- ✓ Desde un punto de vista “cuantitativo”, la tasa de variación de los coeficientes de los conglomerados tiene un “gran salto” en la etapa 207;
- ✓ Desde un punto de vista “cualitativo”, las agrupaciones coincide con las percepciones de los integrantes del grupo de investigación y los presidentes vecinales.

5.2.2 Método no Jerárquico

Posteriormente, se utiliza un método no jerárquico para maximizar la homogeneidad de las observaciones asignados a un mismo cluster y la heterogeneidad entre los distintos conglomerados.

Para realizar este procedimiento, se contrasta la salida de Weka y del módulo de minería de datos del SQL Server 2000.

a) Weka

Para realizar el agrupamiento no jerárquico se utilizó el algoritmo simple K-medias. Se indica que la cantidad de conglomerados es 3 y se deja al algoritmo el cálculo de los centroides iniciales dado que no tiene una opción para ingresarlos. En esta

Tabla 6: Centroides finales de los conglomerados

Cluster centroids:

Attribute	Full Data (210)	Cluster#		
		0 (65)	1 (112)	2 (33)
Fem_0a15	0.7048	0.2769	1.1071	0.1818
Fem_16a30	0.6	0.5846	0.6696	0.3939
Fem_31a45	0.4476	0.1538	0.6786	0.2424
Fem_46a60	0.2857	0.7231	0.0179	0.3333
Fem_61yMas	0.1714	0.2462	0.0446	0.4545
Masc_0a15	0.5952	0.2923	0.8571	0.303
Masc_16a30	0.5143	0.4923	0.5714	0.3636
Masc_31a45	0.3714	0.0923	0.5625	0.2727
Masc_46a60	0.2333	0.6615	0.0446	0.0303
Masc_61yMas	0.1571	0	0	1
Trabajan	1.3571	1.5231	1.3661	1
Quieren_Trabajar	0.5143	0.5385	0.5804	0.2424
Subsidios	0.1905	0.1538	0.1518	0.3939
Alergia	0.8429	0.9538	0.875	0.5152
Cancer	0.1	0.1077	0.0804	0.1515
Broncoespasmos	0.6333	0.6154	0.6964	0.4545
Jefe_Edad	35.6857	49.0154	26.7143	39.8788

agrupación de las observaciones, la distribución de estas fue:

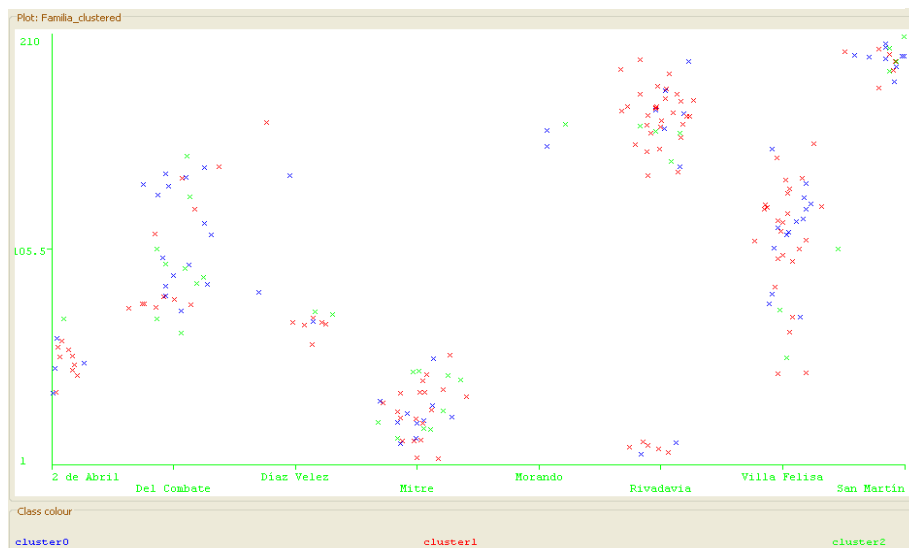
En el conglomerado 0 (30% de los casos), el grupo familiar estaría conformado por adultos (46 a 60 años) con una presencia de jóvenes mujeres entre 16 y 30 años. La edad media del jefe de hogar oscila en los 49 años, lo cual puede ratificar la suposición anterior. La cantidad de personas que trabajan es importante, lo cual indicaría que trabajan ambos cónyuges. La presencia de casos de alergia y broncoespasmos es alta, especialmente los primeros.

En el conglomerado 1 (53% de los casos), el grupo familiar lo componen cónyuges jóvenes con hijos adolescentes. La media de la edad del jefe de hogar (27 años aprox.) ratificaría esta idea. El valor representativo de los miembros que trabajan señalaría que uno de los cónyuges trabaja. Aunque, las expectativas de trabajar son altas.

En el conglomerado 2 (16% de los casos) son familias con miembros adultos y mayores, principalmente varones de más de 61 años. Además, hay una presencia mayor de jóvenes entre 16 y 30 años que adultos entre 31 y 45 años. Si bien hay una presencia significativa de las personas que trabajan, también lo hay de aquellas que perciben subsidios.

A continuación, a través de un gráfico de puntos, se expone la distribución de los clusters en los barrios relevados. En el presente gráfico, en el eje de las x están los barrios relevados y en el eje de las y, las observaciones. En esta primera instancia podría estimarse que:

Gráfico 3: Dispersión de los conglomerados por barrio



El grupo familiar conformado por personas mayores (cluster 2) se encuentra mayormente presente en los barrios: Del Combate, Mitre. Y en menor medida, en los barrios: 2 de Abril, Díaz Vélez, Rivadavia y Villa Felisa.

La familia conformada por personas jóvenes y adultos (cluster 0) está distribuida mayormente por los barrios: Del Combate, Mitre, Villa Felisa y San Martín.

La unidad familiar de cónyuges jóvenes e hijos pequeños (cluster 3), están distribuidas por todos los barrios pero con menor presencia en: San Martín y Del Combate.

b) SQL Server 2000

Analizar los datos en el módulo de minería de datos del SQL Server 2000 tiene una ventaja respecto de los otros software's. El algoritmo de esta herramienta permite trabajar con datos paramétricos como no paramétricos al mismo tiempo. El analista se desprende de este problema y puede concentrarse en los atributos a seleccionar. De esta manera, para realizar el análisis de conglomerados, se seleccionaron todos los atributos del bloque A y B. Además, en la interfaz del usuario se unifica las tablas de datos con gráficos de clusters.

Como en el caso anterior se le señaló el número de cluster a obtener y los resultados se exponen a continuación.

a) Conglomerado 1: 86 casos

Tabla 7: Frecuencia que adoptan los distintos atributos del cluster 1

Atributos	Valores
Fem 16a30	3
Alergia	5
Alergia	4
Masc 31a45	2
Masc 0a15	4
Fem 0a15	3
Fem 31a45	1
Fem 0a15	2
Fem 0a15	1
Fem 46a60	0
Fem 16a30	2
Subsidios	1
Masc 46a60	0
Masc 31a45	1
Trabajan	1
Propietarios	no son propietarios
Proximidad	A menos de 100m
Conoce Incidencia	Desconoce Incidencia
Fem 61y Mas	0
Barrio	Rivadavia

En este conglomerado, la unidad familiar estaría compuesta por cónyuges jóvenes-adultos (entre 16 y 30 años) y por jóvenes menores a 15 años. La edad promedio del jefe de hogar está alrededor de los 22 años, lo cual permitiría confirmar lo anterior. En cuanto a la cantidad de personas que trabajan sólo llega a 1 en el 53% de los casos. Esto presupone que sólo trabaja un cónyuge.

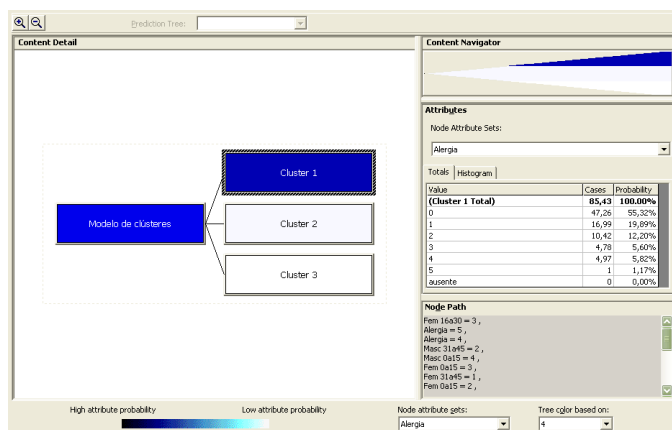
Este tipo de familia se encuentra mayormente en Barrio Rivadavia (aprox. el 32% de los casos), seguido por Barrio Mitre (24% de los casos) y Barrio Villa Felisa (15% de los casos). Las enfermedades tienen una alta presencia. ***Se mantiene la observación que a mayor presencia de jóvenes mayor aparición de casos de enfermedad: alergia y broncoespasmos.***

En comparación con el resto de los conglomerados, es muy frecuente que 4 integrantes padezcan de alergias. Mientras que en el caso de broncoespasmos, es frecuente que de 3 integrantes sufran esta enfermedad. Esto puede visualizarse en los siguientes gráficos:

En la ventana Content Detail se visualizan los tres conglomerados. Al seleccionar el atributo 'Alergias' (en el Node attribute sets), se indica que el color que deberán adoptar los diferentes cluster sea en base al número de integrantes con alergias (Tree color based on). Según el continuo 'High attribute probability' – 'Loss attribute probability', el color más oscuro indica que el valor seleccionado en 'Tree color based on' tiene la mayor probabilidad de ocurrencia. En este ejemplo, cuando el número de integrantes de la familia que padezcan alergias es 4, el cluster 1 es el

más oscuro. Es decir, en este cluster *se encuentra la mayor probabilidad que aparezcan 4 y 5 miembros de una familia con alergia*.

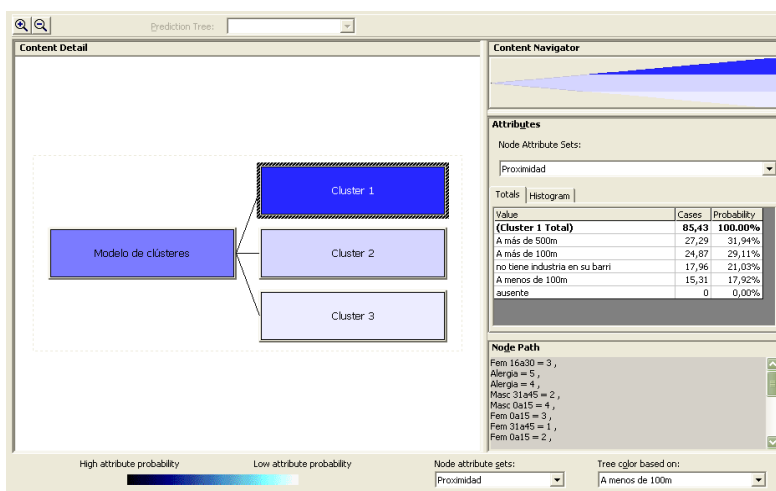
Ilustración 1: Cluster's en función de la cantidad de personas con alergias



Sin embargo, dentro de este cluster las familias que tienen 4y 5 miembros con alergia son el 5,82%y el 1,17% respectivamente. Estas dos observaciones anteriores permiten estimar que son muy puntuales las familias con enfermos de alergias y broncoespasmos.

En cuanto a la proximidad a empresas, es el conglomerado que *posee la mayor probabilidad que una familia esté "a menos de 100m" de distancia*. Como puede visualizarse en el gráfico siguiente.

Ilustración 2: Densidad en función de la proximidad a empresas



Dentro del conglomerado, el 32% (aprox.) de los hogares se encuentra a más de 500m, el 29% a más de 100m, el 18% (aprox.) a menos de 100m y el 21% no tiene industria en su barrio. Esto responde a la distribución que tienen los establecimientos en los tres barrios que "pertenecen" al cluster 1: Barrio Rivadavia, Barrio Mitre y Barrio Villa Felisa.

Sin embargo, en este conglomerado prevalece mayoritariamente los encuestados con un desconocimiento de casos de empresas que inciden en la salud de la

población. Esto puede sustentarse en que los casos con enfermedades están concentrados en pocas familias.

b) Conglomerado 2: 76 casos

Tabla 8: Frecuencia que adoptan los distintos atributos del cluster 2

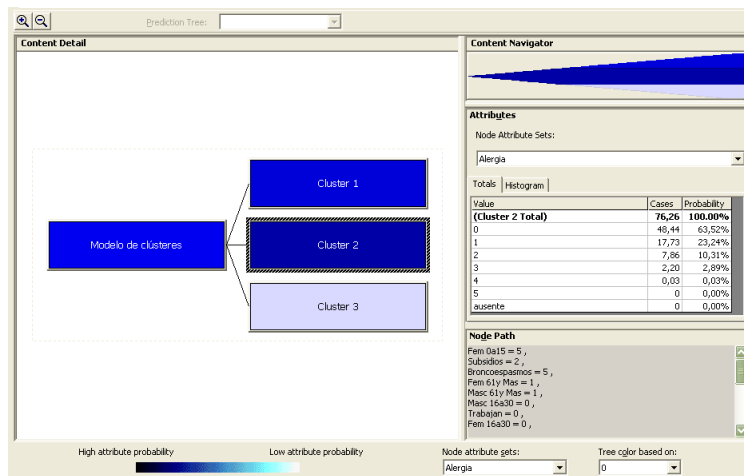
Atributos	Valores
Fem 0a15	5
Subsidios	2
Broncoespasmos	5
Fem 61y Mas	1
Masc 61y Mas	1
Masc 16a30	0
Trabajan	0
Fem 16a30	0
Barrio	Del Combate
Jefe Edad	> 86
Broncoespasmos	0
Masc 0a15	0
Fem 0a15	0
Quieren Trabajar	0

Atributos (cont.)	Valores (cont.)
Masc 46a60	0
Fem 31a45	0
Alergia	0
Barrio	San Martín
Fem 61y Mas	2
Barrio	Villa Felisa

En este conglomerado las familias están conformadas por cónyuges adultos-mayores (65%). La edad del jefe de hogar en promedio es alrededor de los 46 años. De todos los conglomerados es aquel que menos miembros de la familia trabaja. Sin embargo, dentro de éste trabajan al menos 1 integrante en el 42% y no trabaja ningún miembro cerca del 30% de los casos. Los barrios en los cuales se distribuye este conglomerado son: Barrio Del Combate (30%), Barrio Villa Felisa (25%), Barrio San Martín y Mitre (con el 15%).

En cuanto a las enfermedades es el conglomerado con menos presencia de miembros con alergia o broncoespasmos. Como puede visualizarse en los gráficos siguientes:

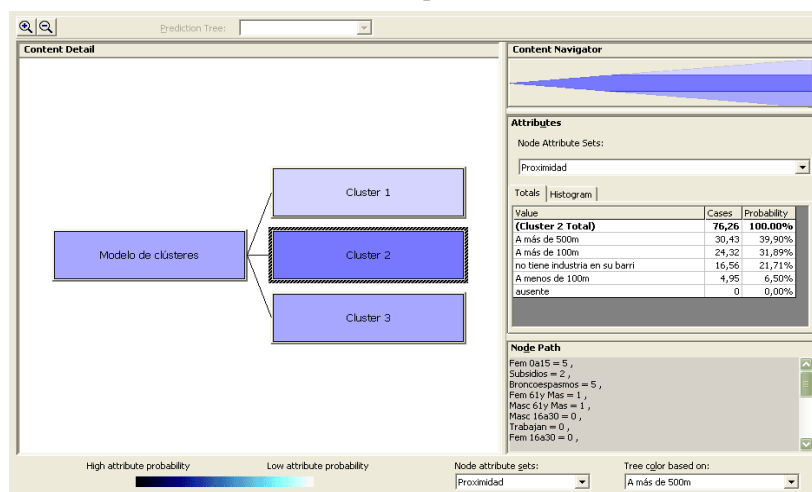
Ilustración 3: Densidad en función de 0 (cero) casos con enfermedades



Además, se puede visualizar que el cluster que sigue en orden a la cantidad de miembros con 0 alergias es el cluster 1. Se ratifica, entonces, la concentración en pocas familias de la presencia de 4 o 5 miembros con alergias.

En cuanto a la proximidad con establecimientos industriales, los grupos familiares están entre 100 m y más de 500m de distancia.

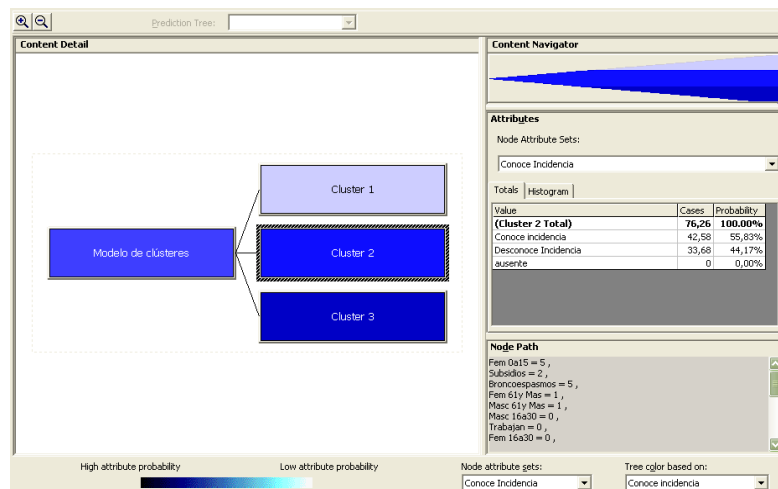
Ilustración 4: Densidad en función de la proximidad a más de 500m



Como puede visualizarse en el gráfico anterior, es el conglomerado con la probabilidad más grande que una familia esté a más de 500m de distancia de un establecimiento industrial.

A pesar de lo anterior, en este conglomerado el 56% (aprox.) de los encuestados conoce de casos de empresas regionales que inciden en la salud de la población. Y también, en orden a aquellos que conocen casos de empresas que inciden en la salud de la población, este está en el 2º lugar. Como se visualiza en el siguiente gráfico.

Ilustración 5: Densidad en función del atributo 'Conoce Incidencia'



Además, se comprueba que en el conglomerado 1 se encuentran los encuestados que mayor desconocimiento tienen de casos de empresas que inciden en la salud de la población.

c) Conglomerado 3: 48 casos

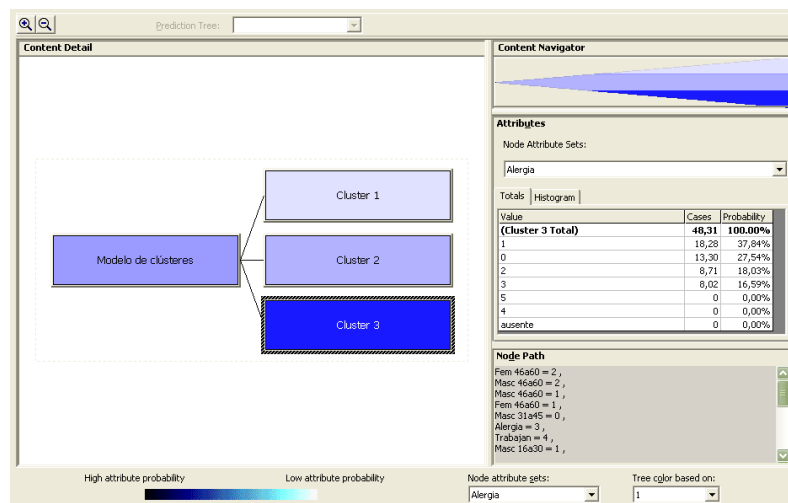
Tabla 9: Frecuencia que adoptan los distintos atributos del cluster 3

Atributos	Valores
Masc 46a60	2
Fem 46a60	1
Masc 31a45	0
Alergia	3
Trabajan	4
Masc 16a30	1
Fem 0a15	0
Reclamo Empresa	Sí reclamó
Trabajan	3
Proximidad	A más de 100m
Broncoespasmos	1
Alergia	1
Conoce Incidencia	Conoce incidencia
Masc 61y Mas	0
Barrio	San Martín
Fem 61y Mas	0
Fem 16a30	1

En este conglomerado, posiblemente, los grupos familiares están formados por jóvenes mujeres y cónyuges adultos. Alrededor del 60% trabajan dos de los miembros familiares. La edad del jefe de hogar es en promedio de 43 años. Los barrios en los cuales se distribuyen las familias pertenecientes a este conglomerado son: Barrio Villa Felisa (21%), Barrio Rivadavia (20%), Barrio San Martín y Mitre (16% y 13% respectivamente).

En cuanto a los casos de alergia y broncoespasmos, **en un 70% de los casos al menos 1 integrante del grupo familiar padece alguna de estas**. A diferencia del conglomerado 1, los casos no están concentrados en pocas familias sino dispersos en todos los grupos familiares. Así, puede visualizarse en el siguiente gráfico:

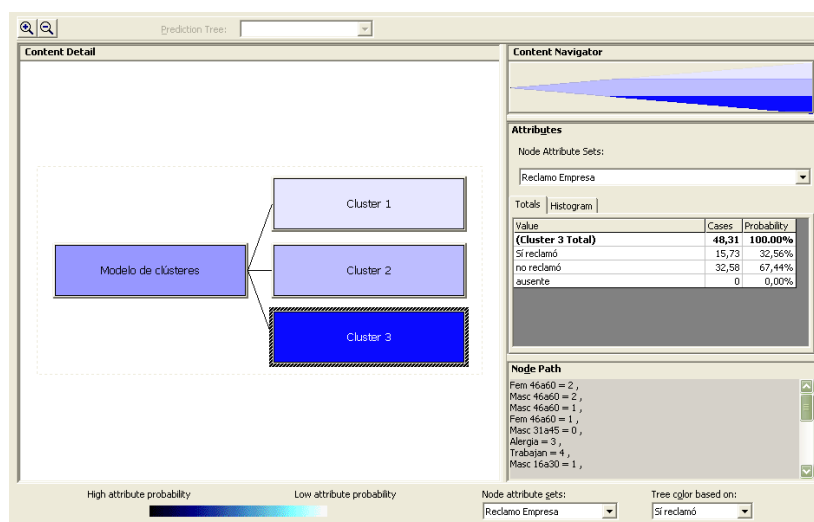
Ilustración 6: Densidad en función a casos de alergia en el Cluster 3



Es el conglomerado en el cual es más probable encontrar 1 miembro con alergia. En la tabla que se visualiza en la ventana 'Atributos', se exponen la frecuencia de apareciendo de los miembros con alergias. En el 38% (aprox.) de los casos de este cluster, se encuentra 1 miembro con alergia; con el 18% 2 miembros con alergia; y con el 17% 3 miembros con alergia. En el caso de broncoespasmos existe un comportamiento similar.

Con respecto a la proximidad a establecimientos industriales, el 51% está entre los 100m y 500m; y el 36% a más de 500m. En relación a este aspecto, se sitúa en primer lugar en función del conocimiento de empresas que inciden en la salud de la población. Alrededor del 66% de los encuestados conoce algún caso de incidencia. En consecuencia, **es el conglomerado que con mayor frecuencia se encuentran familias que reclamaron por la actividad empresarial.**

Ilustración 7: Densidad en función del reclamo a empresas



Dentro de este conglomerado, realizaron un reclamo formal cerca del 33% de los encuestados. Este bajo nivel de reclamo está asociado a un alto nivel de desconocimiento de dónde asesorarse para realizar el reclamo, el porcentaje es alrededor del 80%.

6 Conclusión y Propuestas

En esta primera etapa explorativa de las características de las familias que participaron en la encuesta de opinión se concluye que existen 3 grupos familiares.

El proceso de exploración se realizó aplicando técnicas del análisis multivariado y herramientas de minería de datos. En cuanto a las herramientas anteriores disponibles, se pudo contrastar la "maleabilidad" de ambos software para la realización de la investigación social.

En este primer informe, hemos encontrado ciertas características o aptitudes del grupo social en estudio.

Al proponernos relevar la situación socio-política y el conocimiento de la población sobre la problemática socio económico ambiental, hemos encontrado evidencias por momentos impensadas:

Aspectos técnicos:

- La cantidad de información obtenida requiere de una mayor dedicación en el futuro, ya que el análisis hasta el momento, de sólo dos bloques (20 de las 38 respuestas) generó una gran cantidad de conclusiones, algunas de las cuales se exponen más arriba.

- El procesamiento con SQL Server resulta más sencillo a priori, y como todo producto Microsoft, automatiza tanto la operatoria que no permite aplicar otros algoritmos diferentes a los previstos. Por esa razón se trató de encontrar otras herramientas que favorecieran el desarrollo de salidas más específicas.

- Tanto Weka como SPSS resultaron más adecuadas para ello, si bien requieren de un importante esfuerzo de capacitación.

Aspectos Sociales:

- La experiencia demostró que la población necesita un mayor grado de compromiso con la problemática en estudio, que sus representantes no se encuentran en condiciones de llevar adelante la tarea y que nuestro equipo no cuenta con los recursos suficientes para completar la misma.

Propuestas:

- Trabajar con los datos, obteniendo las primeras conclusiones relativas a características de la población, y su relación con las opiniones y la cultura expresadas.

- Aplicar la metodología aprendida a la muestra prevista originalmente y a otras poblaciones.

Referencia Bibliográfica

- Chang Castillo, H. (2010, Enero/Junio). El modelo de la triple hélice como un medio para la vinculación entre la universidad y empresa. *Revista Nacional de Administración*. 1. (1): 85-94.
- Colobrans, J. (2010, Enero/Febrero). De la triple hélice a la innovación social: ¿Qué está ocurriendo en el citilab de Cornellà? La factoria. Disponible en: <http://www.revistalafactoria.eu/articulo.php?id=520>
- Global Reporting Initiative. (2006). *Guía para la elaboración de memorias de Sostenibilidad*. (Versión 3.0). Disponible en: http://www.globalreporting.org/NR/rdonlyres/415F297B-5289-4160-8B6C-7CC034D5BE52/0/G3_GuidelinesESP.pdf
- Marchese, A., Gaibazzi, M., Ingrassia, R., Pérez Malo, N. (2009, Noviembre). *Sistema de conocimiento en organismos públicos*. Trabajo presentado en las XIV Jornadas “Investigaciones en la Facultad” de Ciencias Económicas y Estadística, Rosario, Argentina.
- Martín Blanco, J. (2007, Diciembre). La Universidad cubana del siglo XXI: la Sede Universitaria Municipal. *Revista Iberoamericana de Educación*. 44: 1-6.
- Sutz, J. (2002, Junio). Comentarios a las reflexiones de Ferry Shinn: La triple hélice y la nueva producción del conocimiento enfocadas como campos socio cognitivos. *Redes*. 18. (9): 213-223.