



Universidad de Chile
Facultad de Ciencias Económicas y Administrativas
Carrera de Ingeniería Comercial

SEGMENTACIÓN DE LA POBLACIÓN CHILENA

Seminario para optar al Título de Ingeniero Comercial

ALUMNO : FRANCISCO
ANTONIO GIADALAH ASTUDILLO

PROFESORES : ENRIQUE
FRANCISCO MANZUR MOBAREC
MIGUEL
RAFAEL MENDOZA HENRIQUEZ

Santiago de Chile
- 2003 -

**A la piedra viva, desechada por muchos, pero que vendría a ser la principal piedra del ángulo, escogida y preciosa.
Al que es Alfa y Omega, principio y fin. Al que es y que era y que ha de venir... el Todopoderoso.**

Agradecimientos

Primeramente le doy las gracias a Dios por su Amor, por ser mi cimiento y por poder vivir este momento, por que solo de Él es el dar y el quitar, el poner y el sacar, el abrir o el cerrar.

Doy gracias también a Antonio e Ingrid, mis padres, por todo el amor, los cuidados y el apoyo que me han brindado durante toda mi vida. Y en general por ser excelentes padres.

Agradezco a Denisse por regalarme su amor y por todo el apoyo que me ha dado.

Agradezco a los profesores Miguel Mendoza y Enrique Manssur, por su guía y por su gran disponibilidad para acercarme a ellos cada vez que lo necesitara.

Doy gracias a toda mi familia, en especial a mis abuelos, por su amor y su apoyo.

Doy gracias también al ministerio de Vidavision, en especial al pastor Alejandro Martínez, por todo lo que me han entregado.

Y finalmente doy gracias a todos los que de alguna forma me han apoyado, para llegar a este momento. A todos los que me han brindado su amor, cariño o amistad.

Tabla de contenidos

Introducción	1
Capitulo I. Métodos de Segmentación.	4
Capitulo II. Metodología	10
Capitulo III. Procedimientos	18
Capitulo IV. Resultados.	24
Capitulo V. Conclusiones.	58
Bibliografía	62
Anexo 1. Descriptores genéricos de segmentación	

- Anexo 2. Descriptores específicos de segmentación**
- Anexo 3. Cuestionario encuesta CASEN 2000**
- Anexo 4. Metodología de análisis de cluster**
- Anexo 5. Set de variables original**
- Anexo 6. Output Adjusted Rand Index**
- Anexo 7. Variables seleccionadas estructura 1**
- Anexo 8. Variables seleccionadas estructura 2**
- Anexo 9. Grafico Gap Statistic**
- Anexo 10. Lista de posesiones**

Introducción

La segmentación se refiere al proceso de homogeneización de un conjunto de entidades o elementos que por naturaleza son heterogéneos en varios subconjuntos, los cuales están en función de características similares.

El objetivo del presente trabajo es lograr la consecución de una segmentación de la población chilena. Una buena segmentación es la que cumple el objetivo de agrupar un conjunto de elementos, en este caso hogares, en grupos o cluster que están internamente conformados por hogares que tienen características similares entre sí, y que además difieren de las características de los hogares que fueron clasificados en otro grupo, debiendo además cumplirse que las variables con que se mide su similitud sean relevantes.

La motivación para tratar el tema anteriormente expuesto radica en la importancia que tiene para nuestra sociedad una buena asignación de recursos, tanto en el ámbito público, como en el sector empresarial, esta segmentación permite identificar deseos y necesidades de grupos específicos, entonces por ejemplo, en lo referente a políticas públicas se podría planificar de mejor manera a qué sectores se dirigen los subsidios, en qué se debe invertir y en definitiva hacer un mejor uso de los recursos. En lo referente al ámbito privado las empresas podrían desarrollar una estrategia comercial más eficiente, sería posible, a través esto, seleccionar los mercados objetivos permitiendo definir una mezcla comercial óptima para dirigirse a un grupo específico.

La investigación además debe considerarse como el primer paso para la creación de una segmentación geodemográfica, método que es ampliamente usado principalmente en Estados Unidos y Europa, en donde puede encontrarse disponible comercialmente.

La otra motivación que guía esta investigación es que pese a la importancia de una segmentación de este tipo, en la actualidad en el país no existe ninguna segmentación disponible alcanzada a través de criterios estadísticos rigurosos, lo

que hace cuestionable su calidad, en particular se observa que hoy en día en el país existen dos métodos en uso: el G.S.E o segmentación por grupos socioeconómicos que consiste en un conjunto de parámetros demográficos que describen a individuos que se pretende respondan homogéneamente a ciertos estímulos del marketing.

El G.S.E. adolece de algunas debilidades: éste se trata de categorías construidas en base a criterios y experiencia, no en base a una metodología estadística de clasificación, que definen, hasta ahora, las estrategias de mercado, la segmentación por G.S.E. además entrega perfiles elaborados para el gran Santiago, los cuales se pretenden proyectar a todo el país. Además, está el hecho de incluir la comuna de residencia en el G.S.E., como una variable explicativa, lo que no considera la coexistencia de diversos segmentos en ella. Por lo mismo, la ubicación geográfica de los segmentos es genérica y no específica. Por otro lado, el número de segmentos trabajados bajo este sistema, que en total suman 7, no necesariamente representan el número real de segmentos en los que se puede particionar nuestra población, cabe por tanto la posibilidad de la existencia de otros subgrupos, contenidos en los segmentos, a los cuales se podría acceder con un sistema de segmentación más fino.

La otra propuesta actual en nuestro país, es la Matriz de Clasificación Social de Adimark y J. Walter Thompson, ésta clasifica a la población sobre la base de dos variables: educación y ocupación del sostenedor del hogar, ésta se basa sólo en una pequeña muestra recogida sólo en algunas de las ciudades del país. Dado lo anterior este método presenta problemas similares al G.S.E.

En la tarea a realizar se hará uso de la Encuesta de caracterización Socioeconómica Nacional CASEN del año 2000, esto como respuesta a la necesidad de una segmentación real de nuestra población mucho más precisa y práctica. Para realizar lo anterior se deberá buscar una nueva metodología, que permita superar el problema impuesto por la inoperancia de muchos métodos, cuando existe una base de datos demasiado numerosa, como es el caso de esta investigación, en el que la encuesta CASEN posee más de sesenta y cinco mil observaciones. De esta manera como ya se mencionó se podrán obtener distintos y más numerosos segmentos en función de descriptores demográficos, todos ellos representativos a nivel nacional de nuestra realidad socioeconómica.

El capítulo I es un acercamiento a la teoría que está detrás de la segmentación. El capítulo II expone la nueva metodología que se propone para la realización de esta investigación, se explica cada proceso involucrado, estipulándose además la hipótesis de la investigación y la fuente de los datos.

El capítulo III explica los procedimientos involucrados en los análisis realizados.

En el capítulo IV se abordan los resultados del trabajo con sus análisis.

Finalmente el capítulo V contiene las conclusiones de la presente Investigación.

Capítulo I

Métodos de Segmentación

1.1 Segmentación de Mercados.

1.1.1 Definición

“La segmentación de mercado es el arte y la ciencia de separar a los clientes o empresas en distintos grupos, en función de determinadas características. Cada grupo no sólo será distinto, sino que los grupos serán idealmente excluyentes” (Paul E. Green, Abba M. Krieger; 1999).

Los elementos dentro de cada grupo serán más o menos similares y, al revés, los elementos a través de los grupos serán diferentes.

1.1.2 Beneficios

El desafío para los negocios es segmentar en la forma más útil. No es suficiente agrupar a las personas o cosas en grupos distintos basados solamente en la similitud de sus miembros. Para ser útiles, los grupos deben responder de forma diferente a las distintas acciones desplegadas por las empresas tanto en diseño, precio, promoción y distribución.

Si bien toda segmentación para ser efectiva, deberá poseer las características que a continuación se enuncian. (William J. Stanton, Michael J. Etzel, Bruce J. Walker; 1996)

- Tener segmentos identificables y medibles.
Las variables para identificar y perfilar los segmentos deben ser posibles de medir.
- Tener segmentos accesibles.

Los segmentos debieran ser tales que permitan llegar a ellos a través de medios que faculten tanto obtener información como dirigir una campaña de marketing eficiente.

➤ Tener segmentos rentables.

Que los segmentos sean lo suficientemente grandes y permanentes como para poder ser capaces de generar utilidades en el tiempo.

➤ Tener segmentos reactivos.

Que los segmentos respondan favorablemente ante los programas de marketing de las empresas. Esto permitirá obtener una respuesta mucho más eficiente de los mercados objetivos ante los estímulos de una determinada mezcla de marketing.

Cabe hacer notar el hecho que los análisis que se realizan, no se hacen para mercados estáticos. Los mercados cambian, son dinámicos, evolucionan. Las respuestas que se obtienen de los segmentos representan opiniones sostenidas en un punto específico en el tiempo y hoy en día esas opiniones están destinadas a cambiar.

1.2 Tipos de Segmentación.

Dadas las condiciones sociales y económicas, durante los primeros 50 años de este siglo, la mayoría de las empresas, persiguiendo una estrategia de mercado agregado, trataron de maximizar la eficiencia de producción y esfuerzos de marketing para fabricar uno o unos pocos productos estandarizados y venderlos a un gran mercado concentrado. (Boyd, Walker; 1993). Su principal estrategia de marketing era apelar a tantos clientes potenciales como fuera posible. Para ello, contaron sólo con su eficiencia de costos para ofrecer los precios más bajos.

En la actualidad sólo pocos mercados de productos se desarrollan bajo estas condiciones, esto, pues un solo producto estandarizado no satisface a la mayoría que lo necesita, ya que muchos clientes prefieren algo diferente. Así, la segmentación de mercado, ha llegado a ser una importante herramienta en el desarrollo de negocios y estrategias de marketing. Algunas razones para ello son:

1. Los mercados han disminuido su crecimiento producto que la población también ha retardado su crecimiento. Esto ha hecho que la competencia entre las empresas sea más agresiva, colocando el acento en las necesidades y deseos de sus clientes.
- 2 Ha aumentado el nivel de ingreso disponible, los niveles de educación, el crecimiento de los medios masivos de comunicación. Se han producido cambios sociológicos de vital importancia como el cambio de los antiguos roles al interior de la familia. Lo anterior, aunado a la irrupción de la información ha

producido clientes con más y variadas necesidades, gustos sofisticados y estilos de vida diferentes.

- 3 Se ha impuesto definitivamente el libre mercado como estructura económica, gatillando un agresivo proceso de competencia en los mercados.

No existe un solo tipo de segmentación de mercados, puesto que éstos estarán en función de las variables que se elijan como descriptores de los segmentos. Los métodos se pueden clasificar en dos grandes grupos: los descriptores de segmentación genéricos y los descriptores de segmentación específicos. (Boyd, Walker; 1993)

1.2.1 Descriptores de segmentación genéricos.

Corresponderían a las Necesidades del Consumidor, Comportamiento relacionado con el producto, Comportamiento relacionado con la persona, y Geográficos (Boyd, Walker 1993) . (Referirse a Anexo 1)

a) Necesidades del Consumidor.

Los beneficios deseados son los buscados por los consumidores de un producto o servicio. El consumidor finalmente comprará el producto que le provea el mejor conjunto de beneficios. Es preciso identificar estos beneficios específicos que buscan los consumidores ante un producto. Una vez conocidos los beneficios individuales, una segunda etapa consiste en describir las características demográficas y psicográficas.

b) Comportamiento relacionado con el producto.

La Tasa de Consumo del producto, se refiere a la tasa con que el público usa o consume un producto. Una clasificación usada para este indicador es la siguiente: no usuarios, usuarios mediano y grandes usuarios. Normalmente a una empresa le interesa sobre todo la última categoría. Las empresas centrarán sus esfuerzos de marketing en conservar a los que constituyen la mitad consumidora pesada de su marca y alientan al mismo tipo de usuario de la competencia para que cambie de marca.

Otro aspecto, la Lealtad, es medido por la frecuencia de compra del producto o servicio. Los consumidores de un producto difieren en su lealtad para una marca. El desafío para las empresas es mantener o proteger sus clientes leales y convertir a los nuevos.

La predisposición de compra es otro punto importante, los clientes potenciales pueden no estar usando un producto por diversas razones, dichas razones reflejan diferentes predisposiciones hacia la compra de un producto. Una segmentación de mercado basada en el conocimiento del producto o predisposición a comprar puede identificar a los no usuarios que sean los más probables en llegar a ser futuros compradores.

La Influencia de Compra es otro de los aspectos, a modo de ejemplo, muchos productos que son usados por la mayoría de los miembros de la familia son comprados por la dueña de casa, en otros las decisiones son tomadas en conjunto (marido y mujer). Por otro lado, existe otra gama donde la decisión es individual, etcétera.

La ocasión de Consumo se refiere a la premisa que, diferentes situaciones de uso dictan diferentes criterios de elección de marca. Así, las ocasiones de uso se pueden organizar en segmentos que tengan similares criterios de selección de marca. Si diferentes criterios de selección de marca operan en varias situaciones de uso, las estrategias de marketing se pueden desarrollar para reforzar la imagen de un producto en términos de lo apropiado que será para una ocasión de uso particular o un set de ocasiones de uso particular.

c) Comportamiento relacionado con la persona.

En la Innovación los individuos varían en su actitud, esto en función de su edad, ingresos, educación, ocupación, exposición a los medios y contactos sociales. Para la estrategia de marketing, el identificar y comprender las distintas características de innovación por parte de los clientes puede ser vital para implementar medidas en las distintas etapas del ciclo de vida de un producto.

1.2.2 Descriptores de segmentación específicos.

Corresponden a los Descriptores de comportamiento y Descriptores demográficos.(Boyd, Walker 1993) . (Referirse a Anexo 2)

a) Descriptores de comportamiento.

Con respecto al estilo de vida, se busca agrupar a los consumidores sobre la base de sus actividades, intereses y opiniones. El estilo de vida provee una de las descripciones más informativa de la conducta de los consumidores.

La Segmentación Psicográfica, se refiere a la segmentación basada en valores y estilos de vida. (Melvin Crask, Richard J. Fox, Roy G. Stout.; 1995). El sistema de segmentación Psicográfica más destacado es el VALS (Valores y

Estilo de Vida) ofrecido por Stanford Research Institute. El sistema VALS se basa en la premisa que el estilo de vida de una persona está dictada por su auto-orientación y los recursos (psicológicos, físicos, demográficos, medios materiales y capacidades) con que dispone el individuo. El sistema de segmentación VALS ha sido ligado a la mayoría de los sistemas de segmentación geodemográfica.

b) Descriptores Demográficos.

Estas variables, tales como edad, ingreso, tamaño del hogar o sexo, dan otra base para particionar a los consumidores en grupos. Fuentes sindicadas tales como la encuesta de Caracterización Socioeconómica Nacional es una fuente secundaria que proporciona información útil sobre el tamaño de tales segmentos.

La Segmentación por Grupo Socioeconómico es una medida integrada por un conjunto de características demográficas, elegidas y definidas sobre la base de criterios y experiencia. El indicador más común de la clase social incluye el nivel de ingreso, el nivel de escolaridad, tipo de ocupación y tipo de vecindario en que vive una persona, es por tanto, una medida demográfica compuesta.

La variable edad puede utilizarse como herramienta predictiva, por ejemplo, si se conoce la distribución de edad de la población en una fecha determinada, y las estimaciones de nacimientos y defunciones, podría predecirse el tamaño de varios grupos de edad en una fecha futura.

El sexo es un descriptor que tiene implicaciones no solamente por razones biológicas sino también por la asignación de roles de hombres y mujeres en nuestra sociedad.

El concepto de ciclo de vida familiar está dirigido a predecir el comportamiento de compra por una combinación de edad, estado civil y la presencia de niños.

En lo que se refiere al Ingreso, la disponibilidad de dinero para afrontar las necesidades es, por supuesto, un pre-requisito de la demanda. Este descriptor se define tanto como el ingreso total familiar o bien como ingreso per capita del hogar.

La ocupación puede relacionarse directamente con la necesidad de productos específicos

La educación está relacionada cercanamente con el ingreso, como ya se dijo, el ingreso por si solo no explica el comportamiento del gasto. El nivel educacional de los consumidores tiene un alto impacto en productos tales como viajes, libros, revistas, seguros y equipos fotográficos. Algunas empresas han encontrado muy rentable dirigirse a ciertos grupos de estudiantes ya sea universitarios o escolares.

Capítulo II

Metodología

2.1 Hipótesis

Se postula la existencia en la población chilena de ciertas características demográficas, que harían posible su clasificación en grupos colectivamente exhaustivos y mutuamente excluyentes, apropiados para la implementación de políticas públicas por organismos estatales y para políticas comerciales dentro de las empresas. Lo anterior está sujeto, a la posibilidad de hallar e implementar una metodología capaz de abordar este desafío.

2.2 Fuentes de datos.

La fuente de datos a utilizar en la investigación es La Encuesta de Caracterización Socioeconómica Nacional 2000 o CASEN 2000. Esta encuesta se ha llevado a cabo desde 1985 con una periodicidad de dos años.

El marco muestral de la encuesta CASEN se basa en los antecedentes del último Censo de Población y Vivienda de 1992, en cuanto a material cartográfico, como también a la información de población y vivienda. Este marco se complementó, además, con antecedentes de las Municipalidades sobre las nuevas construcciones surgidas a partir de 1992, especialmente en aquellas comunas con crecimiento post-censal significativo.

El tipo de muestreo utilizado es Estratificado por conglomerado probabilístico, La estratificación utilizada es de tipo geográfica. El país se dividió en estratos, se entendió como tal a la conjunción de división político-administrativa (comuna o agrupación de comunas) y área geográfica (urbana o rural). En el área urbana están constituidos por los sectores de empadronamiento censal y en el área rural como un conjunto de viviendas próximas. En cada conglomerado se seleccionan sectores censales con probabilidad proporcional al tamaño del éste medido por el

número de viviendas, dentro de cada sector seleccionado se eligen las viviendas a encuestar.

El universo de la encuesta representa a la población que habita en hogares particulares del país, se excluye las zonas de difícil acceso. El tamaño de la muestra alcanza los 65.036 hogares en la versión 2000.

La clasificación de Urbano - Rural, es la misma que el INE utilizó en el Censo de 1992; es decir:

Área Urbana: Conjunto de viviendas concentradas con más de 2.000 habitantes, o entre 1.001 y 2.000 habitantes, con el 50 % de su población económicamente activa, dedicada a actividades secundarias y/o terciarias.

Área Rural: Conjunto de viviendas concentradas o dispersas con 1.000 habitantes o menos o entre 1.001 y 2.000 habitantes, con menos del 50 % de su población económicamente activa, dedicada a actividades secundarias y/o terciarias.

La Encuesta CASEN, tiene en la actualidad representatividad nacional, regional y por zonas (Urbano y Rural).

En el anexo 3 se puede encontrar el cuestionario utilizado para levantar la encuesta CASEN 2000.

2.3 Métodos Estadísticos.

Al comenzar este trabajo, el objetivo que se quería alcanzar estaba claramente establecido. Sin embargo, lo que no parecía para nada claro era la forma de llegar a él. Al momento de empezar no había ninguna metodología conocida y que fuera implementable, que nos permitiera alcanzar dicho objetivo. La principal dificultad que se enfrentaba era la gran cantidad de observaciones disponibles, que se deberían someter a la segmentación. Lo que constituía una limitante por el hecho que esta cantidad de datos hacía imposible aplicar las aproximaciones que comúnmente se usan para segmentar, dado que la capacidades de los programas computacionales utilizados a este respecto, se veían complemente sobrepasadas. En particular nos estamos refiriendo al análisis de cluster, que es la denominación de un grupo de técnicas multivariantes cuyo principal propósito es agrupar objetos (en nuestro caso hogares), de tal forma que cada objeto es muy parecido a los que hay en el conglomerado con respecto a algún criterio de selección predeterminado. Dentro de esta familia de métodos existen dos grandes subgrupos, por un lado están los métodos jerárquicos que se caracterizan por el desarrollo de una jerarquía o estructura en forma de árbol, y por el otro están los métodos no jerárquicos o también conocidos como de K medias, cuya idea básica es asignar los objetos a K conglomerados de manera que la suma de cuadrados dentro estos sea minimizada. (Para una mayor explicación del análisis de cluster referirse al anexo 4).

Los métodos jerárquicos a pesar de que tienen varias características deseables están sujetos, a la limitante antes planteada, de incapacidad de abordar un problema con gran cantidad de datos, como es el caso. De hecho los programas computacionales que teníamos disponibles solo soportaban una muestra con un máximo con alrededor de 300 casos, lo que hacía imposible siquiera tomar un

subset de datos de la muestra original, ya que esta sería de menos del 1% de los datos, lo que de ninguna manera resultaría representativo. La alternativa natural ante esto eran los métodos de K medias, los cuales si tienen la ventaja de poder ser utilizados con grandes muestras. Sin embargo esta alternativa tampoco adolecía de inconvenientes. Siendo el principal la determinación del apropiado grupo de segmentos, asunto sobre el cual la alternativa de K medias no arroja ninguna luz. Otro problema del método es que la selección de los centros iniciales es arbitraria, y puede conducir a máximos locales. La solución más común para estos dos problemas planteada en la literatura, es el realizar previamente un procedimiento jerárquico, que de indicios del número adecuado de conglomerados, y que a la vez permite usar los centros finales de los grupos como iniciales para el procedimiento de K medias. Sin embargo como queda claro por lo anteriormente expuesto, esta opción no estaba disponible. Además se debió enfrentar otra dificultad no inherente al método, como lo fue el de seleccionar un set adecuado de variables que hiciera posible lograr una buena clasificación, lo que resulta de gran importancia dado el amplio espectro de variables candidatas a ser incluidas y dado que como se ha reflejado en la literatura sobre el tema, sólo un subgrupo de variables es realmente valioso para definir una estructura verdadera. (DeSarbo, Carroll, Clark, & Green, 1984; De Soete, DeSarbo, & Carroll, 1985; Gnanadesikan, Kettenring & Tsao, 1995; Milligan, 1989), y aún más la incorporación de variables que no contribuyen a definir esta estructura verdadera pueden efectivamente complicar u oscurecer la recuperación de ésta en un procedimiento de Análisis de Cluster ya sea jerárquico o no jerárquico. (Milligan, 1980; Milligan, 1989). En este tipo de trabajos, dicho problema usualmente ha sido enfrentado a través del uso de un análisis de componentes principales previo al análisis de cluster mismo, lo que permite reducir un gran set de variables a una cantidad manejable de componentes. Sin embargo, este proceso de dos etapas ha sido objeto de recientes críticas, por no hacer uso de toda la información disponible. (Arabie and Hubert, 1994).

Debido a todo lo expuesto anteriormente, se inició la búsqueda de nueva literatura sobre el tema, la que permitiera hacer frente a estos obstáculos encontrándose algunas posibles soluciones, varias de ellas descartadas, ya sea por la misma limitante de no poder trabajar con un gran set de datos o por la inexistencia de un programa computacional disponible que permitiese implementar dicha metodología.

Finalmente se encontraron metodologías, las que al menos de manera teórica, parecían posibles de implementar. La primera de ellas es la de "A Variable-Selection Heuristic for K-means clustering" (Michael J. Brusco and J. Dennis Cradit; 2001) que permitió abordar el problema de selección de un set de variables adecuado, la otra metodología es el "Gap Statistic" (Robert Tibshirani, Guenther Walter and Trevor Hastie; 2001) que hizo posible determinar el número correcto de cluster.

2.3.1 A Variable-Selection Heuristic for K-means clustering

Este método está basado en el "Adjusted Rand index" (Arabie and Hubert, 1985), índice que es utilizado para calcular la concordancia entre dos particiones específicas. El método incluye variables en una forma iterativa y además hace uso de la información de la suma de cuadrados entre cluster y de la suma total de cuadrados. Un adjusted Rand Index grande entre las variables iniciales y una variable j , sugeriría que al añadir dicha variable j , no se enmascararía la actual estructura.

Para una descripción detallada del algoritmo, usaremos la siguiente notación:

S = set de variables seleccionadas para la inclusión en el análisis de cluster.

U = set de variables descartadas, $S \cup U = [1, 2, \dots, D]$ y $S \cap U = [\Phi]$

W_{jk} = un vector fila ($1 \times M$) que define una partición desarrollada usando las variables j y k con $j = 1, \dots, D-1$ y $k = j+1, \dots, D$.

Los elementos del vector (w_{jki}) designan a qué cluster el objeto i es asignado.

Q = Una matriz simétrica ($D \times D$), con ceros en la diagonal principal y donde los restantes elementos $q_{jk} = q_{kj}$, definen el ratio entre la suma de cuadrados Inter Cluster y la suma total de cuadrados para la partición w_{jk} con $j = 1, \dots, D-1$ y $k = j+1, \dots, D$.

T = Es el valor umbral del adjusted Rand Index al seleccionar el primer par de variables.

y = Es un vector fila ($1 \times M$), que define una partición desarrollada usando las variables j perteneciente a S . Los elementos del vector y_i , designan el cluster al que el objeto i es asignado.

G_j = Es el adjusted Rand Index asociado con la partición p_j e y , para $j = 1, \dots, D$.

G_{min} = Es el mínimo valor permitido de G_j tal que la variable j puede ser seleccionada para la inclusión en el análisis de cluster.

G_{fac} = Es un factor que al ser multiplicado por G_j de la variable seleccionada más recientemente, sirve como un valor umbral para la inclusión de la siguiente variable.

Pasos del algoritmo

Paso inicial

Se comienza con $y = 0$; $p_j = 0$ para todo $j = 1, \dots, D$; $W_{jk} = 0$, para todo $j=1, \dots, D-1$ y $k= j+1, \dots, D$; $S = [\Phi]$, $U = [1, 2, 3, \dots, D]$.

Paso 1

Desarrollar una partición p_j de C cluster usando sólo la variable j , para $j = 1, \dots, D$.

Paso 2

Calcular el adjusted Rand Index para todas las $D(D-1)/2$ par de particiones p_j y p_k . Para $j = 1, \dots, D-1$ y $k = j+1, \dots, D$.

Paso 3

Desarrollar una partición W_{jk} de C cluster usando las variables j y k y calcular q_{jk} , para $j = 1, \dots, D-1$ y $k = j+1, \dots, D$.

Paso 4

Si el $\text{Max}_{j,k} (r_{jk}) \geq T$, luego $\varepsilon = \text{Max}_{j,k} (q_{jk} / r_{jk} \geq T)$ también $\varepsilon = \text{Max}_{j,k} (q_{jk})$. Luego j^* y k^* son variables tales que $q_{j^*k^*} = \varepsilon$, y $\eta = r_{j^*k^*}$. Entonces $S = S \cup (j^*, k^*)$ y $U = U - (j^*, k^*)$.

Paso 5

Desarrollar una partición y de C cluster usando todas las variables seleccionadas j pertenecientes a S .

Paso 6

Para cada variable no seleccionada j perteneciente a U , calcular el adjusted Rand Index G_j entre las particiones p_j e y .

Paso 7

Si $\lambda = \text{Max}_{j \in U} (G_j)$. Si $\lambda < G_{\min}$, o $\lambda < \eta \times G_{\text{fac}}$, entonces ir al paso 8. Si no j^* denotará la variable para la cual $G_{j^*} = \lambda$, tomando $\eta = \lambda$, y $S = S \cup (j^*)$ y $U = U - (j^*)$.

Paso 8

Las variables en el subconjunto S son seleccionadas para su inclusión y las variables en U son descartadas. El análisis de K medias es realizado usando sólo las variables en S .

En el paso 1 se desarrollan particiones individuales para cada una de las variables candidatas a ser incluidas a través del método de K medias. En el paso 2 se calcula el adjusted Rand Index para cada par de particiones individuales. En el paso 3 se realizan particiones para cada posible par de variables, además se calcula el ratio entre suma de cuadrados Inter cluster y la suma total de cuadrados. Al llegar al paso 4 se dispone de dos importante piezas de información para cada par de variables: (a) el adjusted Rand Index para cada par de variables y (b) una estimación de la homogeneidad de los cluster, dado por el ratio de las sumas de cuadrados. Entonces es escogido el par de variables con el mayor ratio para su inclusión, siempre que el adjusted Rand Index para este par de variables supere el valor umbral establecido (T). El objetivo es seleccionar un par de variables con un nivel de concordancia razonable y que al mismo tiempo produzca cluster homogéneos. Si ningún par de variables alcanza el valor umbral, entonces el que tenga mayor ratio es escogido. En el paso 5 todas las variables seleccionadas son usadas para generar una partición (y) y en el paso 6 se calcula el adjusted Rand Index entre esta última partición y cada una de las variables no seleccionadas hasta el momento. En tanto, en el paso 7, se selecciona la variable con el mayor adjusted Rand Index calculado en el paso 6, siempre que este sea mayor a los umbrales G_{\min} y G_{fac} . Si esto ocurre, se vuelve al paso 5, incluyendo

además la variable recién seleccionada. Esto continúa hasta que ninguna variable supere los umbrales establecidos, entonces se prosigue al paso número 8, que no es más que realizar el análisis de cluster con las variables que fueron seleccionadas.

2.3.2 Gap Statistic

Esta técnica que permite estimar el número de cluster en un set de datos es aplicable a prácticamente cualquier método de cluster. La idea básica detrás de él es la comparación del cambio de la dispersión dentro del cluster con la esperanza bajo una apropiada distribución de referencia nula.

Los pasos para la implementación del método son los siguientes:

Paso 1

Se Realiza el análisis de cluster con el set de variables originales a través de varias repeticiones cambiando el numero de cluster requeridos (K). Generando medidas de dispersión intra cluster W_k , donde $k= 1, 2, \dots, K$.

Paso 2

En este paso se deben generar B bases de datos de referencia, para lo cual existen dos alternativas:

- a) Cada variable de referencia es generada a través de una distribución uniforme sobre el rango de los valores observados en variable original correspondiente.
- b) Generar las variables de referencia a partir de una distribución uniforme sobre una matriz alineada con los componentes principales de los datos. En detalle, si X es nuestra matriz de datos de $n \times p$, se asume que las columnas tienen media 0 y se calcula el valor de descomposición singular $X = UDV^*$, donde $(*)$ =traspuesto. Luego se transforma a través de $X' = XV$ y se generan las variables de referencia Z' uniformemente, sobre los rangos de las columnas de X'. Finalmente se transforma a través de $Z = Z'V^*$, con lo que se obtiene los datos de referencia Z.

El primer método tiene la ventaja de la simplicidad, pero el segundo toma en cuenta la pendiente de la distribución de los datos.

Una vez que se tienen los datos de referencia, cada una de los B set de datos es sometido al mismo proceso que los datos originales (paso 1), obteniéndose medidas de dispersión intra cluster W_{kb}^* , donde $k= 1, 2, \dots, K$ y $b= 1, 2, \dots, B$.

Para luego calcular:

$$Gap(\hat{K}) = \frac{1}{B} \sum_b \log(W_{kb}^{\#}) - \log(W_k)$$

Paso 3

Llamemos $\Psi = \frac{1}{B} \sum_b \log(W_{kb}^{\#})$. Luego se calcula la desviación estándar:

$$sd_k = \left[\frac{1}{B} \sum_b \log \left(\frac{\Psi_{kb}^{\#}}{\Psi} \right) \right]^{1/2}.$$

Y luego se define:

$$s_k = sd_k \sqrt{1/B}.$$

Finalmente se escoge el numero de cluster según la siguiente regla:

$K_{\text{optimó}} = \text{el mas pequeño } K \text{ tal que } \text{Gap}(K) \geq \text{Gap}(K+1) s_{k+1}.$

Sin embargo se debe analizar toda la curva ya que puede haber mas de un máximo.

Para una idea mas intuitiva de lo que hace el método considérese lo siguiente: Si en realidad los datos tuvieran K bien separados cluster, se esperaría que si $k \leq K$, entonces $\log \left(\frac{\Psi_{kb}^{\#}}{\Psi} \right)$, es decir la dispersión intra cluster, decrecería mas rápidamente que su contraparte calculada a partir de una base de datos generada aleatoriamente. Y en cambio, si $k > K$ entonces esencialmente se estarían añadiendo centros de cluster innecesarios. Luego el álgebra nos muestra que $\log \left(\frac{\Psi_{kb}^{\#}}{\Psi} \right)$ debería decrecer mas lentamente que su tasa esperada. Con lo que el Gap statistic debe alcanzar su máximo cuando $k=K$.

Capítulo III

Procedimientos

Al comenzar el trabajo, se debe constatar que ya existía un trabajo previo con las variables a utilizar, éste consistió en lo siguiente: agrupación de los datos que estaban a nivel de individuos en forma de hogares, selección dentro de todas las preguntas que incluía la encuesta Casen 2000, de aquellas que tuvieran un mayor valor a la hora de segmentar la población chilena. En este proceso, algunas de las variables fueron seleccionadas directamente de las respuestas originales, mientras que en otros casos fueron creadas variables a partir de las ya existentes. El trabajo se inició con este set de variables más un par de ellas que fueron creadas y añadidas a continuación. El set completo de variables con que se comenzó a trabajar se puede observar en el anexo 5.

Una vez que se obtuvo el set de variables definitivo con el que se comenzaría a trabajar, la primera decisión que se debió enfrentar fue la de si se debían estandarizar o no los datos. Esto, dado que como ya se señaló anteriormente, se optó por no utilizar la metodología que consistía en realizar un análisis de componentes principales previo al análisis de cluster, dicha aproximación implícitamente produce la estandarización de las variables, llevándolas a una media de cero y a una desviación estándar de uno, lo que se conoce como Z-score estandarización. Pese a lo anterior, se sabía que la estandarización de las variables es común aún cuando no se realiza un análisis factorial previo y además el hacer esto, tiene mucho sentido cuando las variables originales están expresadas en distintas unidades (P. Green, C. Schaffer, 1996), sin duda, éste era el caso al cual se debió enfrentar, donde se tenían variables medidas en unidades muy disímiles, que se movían en rangos diametralmente distintos, lo que se comprobó a través de pruebas iniciales que afectaba ostensiblemente la consecución de un buen resultado a través del uso del método de k medias,

observando que las variables con mayores desviaciones estándar eran las que tenían el mayor impacto en el resultado.

Sin embargo la estandarización de las variables no está exenta de costos. Si existe una relación natural reflejada en la escala de las variables, entonces la estandarización puede no ser apropiada. Hay que tener en cuenta que ésta tiene impactos tanto conceptuales como empíricos.

Luego de evaluar todo lo anterior, se decidió finalmente el llevar a cabo la estandarización de las variables, pero con esto todavía no estaba todo dicho ya que había más de una opción para realizar esto. Finalmente se optó por la opción de dividir cada variable por su propio rango, alternativa que según Arabie and Hubert (1994), puede ser superior a estandarizar por Z-scores, con respecto a esto, anteriormente, también Milligan and Cooper (1988) encontraron que ofrecía los mejores resultados a lo largo de diferentes condiciones de error, medidas de similitud y algoritmos de cluster.

Ya con el set de datos estandarizados, se debió tomar la decisión sobre cual de las dos metodologías complementarias a la de K medias, que se habían escogido, debía ser aplicada primero, ya que de cierta forma había una interdependencia entre ellas. Dado que como input para el algoritmo de selección de variables era necesario establecer un número determinado de cluster o grupos, y por otro lado la selección de un número adecuado de cluster, a través del gap statistic dependía del set de variables ocupado para realizar el análisis. Pese a que se tenía una noción teórica de lo que sería más adecuado se decidió probarlo directamente.

Para esto, se implementó en primer lugar el gap statistic a partir de todo el set original de variables.

Los resultados del análisis fueron contundentes, éstos mostraron que el conjunto de datos no debían ser sometidos a ningún tipo de segmentación ya que no parecían ninguna división clara en ellos.

Estaba claro entonces el efecto de la presencia de variables enmascarantes que en lugar de ayudar a develar la existencia de la estructura verdadera subyacente en los datos, producían el oscurecimiento de ésta. Esto demostró la inutilidad de someter los datos al método del Gap Statistic sin previamente realizar una preselección de las variables que verdaderamente contribuyen a la consecución de una buena segmentación.

Una vez realizado lo anterior, se decidió entonces tomar como primer paso el algoritmo de selección de variables, lo que teóricamente era preferible desde el comienzo, por el hecho de que en el estudio en donde se presenta el método, éste muestra una gran estabilidad respecto a las variables seleccionadas, aún cuando se varíe el número de cluster establecido como input.

La implementación de éste, sin embargo, no estuvo exenta de dificultades. En primer lugar, no se contaba con algún paquete estadístico capaz de calcular directamente el adjusted rand index, paso fundamental para llevar a cabo el algoritmo. Para enfrentar esta dificultad, se debió estudiar el uso del módulo de sintaxis, incluido en el programa estadístico SPSS, a través del cual se encontró finalmente la forma de calcular este índice, aunque de una manera bastante engorrosa (Para ver un output del adjusted rand index referirse al anexo 6). El siguiente obstáculo a vencer tenía su origen más en los datos que en el algoritmo de selección de variables mismo. El problema se producía por la abundante

presencia de variables categóricas u otras que pese a ser de escala presentaban una pobre variabilidad. Esto sumado al hecho que el adjusted rand index es capaz de comparar sólo particiones de un igual número de grupos, producía una limitación en el número de cluster seleccionados como input. Esto debido que al realizar las particiones individuales para cada variable, como está establecido en uno de los pasos del algoritmo, estas variables con escasa variabilidad, no eran capaces de discriminar en un número mayor de grupos de lo que estaba dado por las categorías que ellas representaban o en el caso de las de escala por el pequeño rango en que se movían, por lo cual, se debió optar por establecer como input un número pequeño de cluster, específicamente cinco, aún con lo cual algunas de las variables debieron ser eliminadas porque su capacidad de discriminación era menor a cinco, éstas eran en su mayoría variables dicotómicas. Por otro lado, la presencia de esas variables provocaba un segundo efecto que alteraba la selección del par inicial de variables seleccionadas por el algoritmo. Esto debido a que la elección de éstas está basada en una medida de homogeneidad intracluster, seleccionándose el par de variables que arrojara un mayor valor en este estadístico, pero muchas de las variables categóricas por su naturaleza misma presentaban el valor máximo en el estadístico ya que al conformar los cluster, en cada uno de ellos se agrupaban sólo los casos exactamente iguales. Ante esto, se debió optar por obviar este primer paso en la selección de variables, lo que se reemplazó por la elección de una variable semilla escogida a partir de su importancia teórica, lo que se facilitó, por el conocimiento de la importancia que tienen las variables incluidas en el análisis.

Finalmente, antes de iniciar el análisis se optó por considerar algunas variables como grupos ya que por sus especiales características lo que importaba era la estructura que podían definir en su conjunto y no así individualmente.

De esta manera, se llevó a cabo el algoritmo de selección de variables, primeramente estableciendo como variable semilla el total de ingresos de cada hogar. De esta forma se llegó a seleccionar un set de veintiún variables, las que se pueden observar en el anexo 7. Sin embargo, luego de analizarlas, se observó que todas ellas apuntaban a un tipo determinado de características de los hogares. Esto hizo pensar en la posibilidad que era contemplada por Michael Brusco y Dennis Cradit. (2001), de que pudiera existir más de una sólo estructura verdadera en un set de datos. Para probar esto se eligió una segunda variable semilla entre las que todavía no habían sido seleccionadas y se hecho a correr el algoritmo, encontrándose efectivamente otro set de veintiséis variables complementario al anterior que apuntaba a otras características de los hogares, que se puede observar en el anexo 8. Llegado este punto la existencia de dos set de variables planteaba una nueva interrogante, las opciones eran las siguientes: trabajar con sólo uno de los dos set de variables, trabajar con ellos paralelamente o combinarlos haciendo un uso conjunto de ellos, esta última fue la escogida ya que sólo a través de ella se podría lograr una segmentación a través de todas las características de los hogares que estimaban importantes.

Ya conocidas las variables finales a ser incluidas en el análisis se procedió ahora sí a realizar la técnica del Gap statistic.

Los resultados arrojados por ésta fueron alentadores, el método claramente estaba funcionando, lo que además constituía una validación a la selección de

variables realizada. En lo específico, luego del análisis de la curva, se observó que había más de un máximo y por tanto había más de un número específico de cluster que podía ser seleccionado. Por lo que se procedió a una comparación de éstos. Se observó además que en varias de las particiones habían cluster nulos, refiriéndose con esto a que habían en varias de las particiones cluster conformados por un solo hogar, mientras que todo el resto estaban formados por un número mucho mayor. De aquí, que la solución escogida que nominalmente era de treinta y ocho cluster fuera en realidad de treinta y tres (Para ver el gráfico del Gap statistic referirse al anexo 9).

Pasada esta etapa, ya se tenía todo lo necesario para realizar el análisis de K medias, a través del set de variables seleccionadas y solicitando una solución con treinta y tres cluster.

Los resultados del análisis de cluster se mostraban contradictorios, por un lado los centros de los cluster tenían muchísimo sentido desde una perspectiva conceptual, pero por otro lado la variabilidad de los cluster en las variables más importantes era exacerbadamente grande, es decir, aunque las medias parecían correctas los rangos en que se movían las variables dentro del cluster eran altísimos. Esta situación desembocó en un profundo análisis de los posibles orígenes del problema. Se llegó a identificar dos posibles causas: Primero estaba la arbitrariedad de los puntos de los centros iniciales de los cluster, los que podían tener una incidencia en el resultado final. Segundo y probablemente más importante era la estandarización a la cual habían sido sometidas las variables, ésta por su naturaleza llevó a las variables conceptualmente más importantes y las que a la vez tenían una mayor variabilidad a tener un menor peso a la hora de realizar el análisis de cluster, que sus contrapartes que tenían una escasa variabilidad. Sin embargo, por otro lado, también se había comprobado que la abismante diferencia en las escalas en las que se movían las variables originales, hacían imposible llegar a un buen resultado ocupando éstas directamente. Entonces se estaba en un dilema, por un lado al estandarizar, las variables más importantes perdían peso impidiendo llegar a una buena solución, pero por otro lado tampoco se podía trabajar directamente sobre los datos originales. Para solucionar esto, la mirada se enfocó nuevamente en la literatura sobre el tema, en la búsqueda de una forma de estandarizar o transformar las variables que constituyera un término medio entre estos dos polos. Lamentablemente, en la literatura no se encontró nada parecido a lo que se buscaba, entonces, la única opción posible era que el mismo investigador encontrara alguna transformación matemática que cumpliera con los objetivos deseados los que consistían en: llevar a las escalas de las variables a un nivel razonablemente similar y que a la vez la relación natural de importancia de las variables siguiera viéndose reflejada en éstas.

Finalmente se encontró una fórmula que cumplía con estos objetivos, la que se presenta a continuación:

$$\frac{x_i}{Rango_{0,5\%}^\alpha}$$

Donde:

x_i son las observaciones en la variable X ;

$Rango_{0,5\%}$ es el rango en la variable X quitando el 0.5% de los datos en cada cola de la distribución;

α es un número que se mueve entre cero y uno.

El rango se calcula a partir de los datos, quitándosele 0.5% de cada cola, con el fin de eliminar los valores extremos, usualmente muy distintos del resto de las variables. Por otro lado el elevar esta suma a α produce el efecto requerido de dejar las variables en escalas razonablemente parecidas sin eliminar el efecto de la importancia relativa de cada una de estas. Este alfa debe ser escogido a manera de ensayo y error buscando el equilibrio que se necesita. En este caso, el alfa que se escogió fue de 0.76.

Con esto quedaba solucionado el problema de la estandarización, retomándose de esta manera el análisis de cluster. Antes de hacerlo, se haría un último ajuste, el que consistió en que en lugar de usar centros iniciales arbitrarios se usaron los centros de los cluster finales que se habían alcanzado en la solución anterior, con lo que quedaba resuelto además el otro problema que afectaba la consecución de un buen resultado.

Finalmente, haciendo uso de todo lo anterior, se echó a correr el programa de k medias, encontrándose resultados totalmente satisfactorios, lo que se determinó a través un exhaustivo análisis.

Llegado este punto sólo faltaba por realizar el perfilamiento de los cluster, en este momento se decidió fusionar tres de los cluster, los cuales tenían pocas observaciones y coincidía que eran los tres grupos de más altos ingresos, siendo muy similares en casi todas sus características. Con lo anterior, finalmente se llega a una solución de treinta y un cluster, cuyas descripciones se pueden encontrar en el capítulo de resultados.

Capitulo IV

Resultados de la Investigación

En este capítulo se presentan los resultados del trabajo efectuado, el cual consiste en treinta y un perfiles correspondientes a los grupos en que se dividió la población. Los resultados están ordenados según el promedio del ingreso percibido por el grupo, en forma creciente. A continuación se presenta una lista con los nombres con que fueron bautizados los grupos.

1	Soledad y Pobreza
2	Solitarios, Separados y Madres Solteras
3	Formales
4	Madrugadores
5	Solitarios acompañados
6	Jóvenes Convivientes
7	Arremangados
8	Montepiados
9	Familia y Esfuerzo
10	Formales emergentes
11	Años Plateados
12	Maduros solitarios
13	Papas por siempre
14	Convivientes
15	Prolíficos
16	Madurez en familia

17	Jóvenes, Formales y emergentes
18	Viudez temprana
19	Años Dorados
20	Autónomos
21	Achoclonados
22	Pujantes
23	Robles Otoñales
24	Colmenas
25	Madurez Tranquila
26	Cuello y Corbata
27	Nanas y Van
28	Profesionales exitosos
29	Emprendedores exitosos
30	Lujos y Viajes
31	Crema y Nata

Soledad y Pobreza (9,58% de la población)

Este grupo se caracteriza por tener hogares pequeños, con una media de 1,7 personas y con un jefe de hogar que sólo en un 23% de los casos tiene pareja. La mediana de la edad del jefe de hogar es de sesenta y cinco años. Con respecto a su escolaridad, la mediana es de tres años, con un porcentaje de analfabetismo del 38%. También es relevante el bajo ingreso que perciben, siendo en promedio de sólo \$ 68.000, y contribuyendo a él en promedio sólo 1,2 personas, encontrándose en el grupo un 62% de hogares con un jefe inactivo. Respecto a las personas que trabajan actualmente la mayoría se desempeña como agricultores o como trabajadores no calificados.

El grupo es predominantemente rural, alcanzando un 67%.

Sólo en un 18% de los hogares existen parejas casadas y en un 5% encontramos convivientes. Es importante la proporción de hogares con personas viudas (35%), a lo que se suma un 14% de hogares con personas separadas o anuladas.

La presencia de hijos es escasa, encontrándose éstos sólo en un 22% de los hogares y con una media de sólo 0,2 hijos por hogar.

En cuanto a la composición etarea el único rango que destaca, es el de sesenta y un años o más, que se encuentra presente en el 72% de los hogares

En cuanto a la educación, como es esperable, es muy baja la proporción de hogares en los que hay personas educándose en la actualidad, ya sea en cualquiera de los niveles educacionales, también es destacable la prácticamente nula existencia de personas con educación superior completa en el grupo

En cuanto a la vivienda en un 80% de los casos ésta pertenece a alguno de los miembros del hogar. Hay un importante porcentaje de casos en que la vivienda es una mediagua o mejora (15%). Igualmente importante son los casos en que la vivienda es un rancho, ruca o choza. (12%). Relacionado a lo anterior, se

encuentra que en el 26% de casos, la conservación de la vivienda es mala. También es relevante el alto porcentaje de hogares con urbanización deficiente. En cuanto a las posesiones, en un 70% de los casos, no se posee ninguno de los artículos mencionados en el anexo 10 y en promedio poseen sólo 0,3 de éstos.

Solitarios, Separados y Madres Solteras (4,6% de la población)

Este grupo se caracteriza por el hecho de que sólo en un 5% de los casos el jefe de hogar tiene pareja, esto a pesar de la relativa juventud de los jefes de hogar, que tienen una mediana de treinta y nueve años, formando hogares pequeños, con una mediana de dos personas. Esto en parte se explica por el 39% de personas separadas y por un 10% de personas viudas. También es interesante que el 62% de los casos el jefe de hogar es de sexo femenino. La mediana de la escolaridad del jefe de hogar es de ocho años, en promedio los hogares perciben un ingreso de sólo \$ 88.000, al que aportan en promedio menos de dos personas. Dicho ingreso se obtiene desempeñándose principalmente como trabajadores no calificados, en servicio doméstico, vendedores y agricultores.

En este grupo predomina la ubicación de los hogares en sectores urbanos con un 64%.

Respecto a la presencia de hijos, se encuentran en el 57% de los hogares, siendo la media de menos de uno por hogar.

En cuanto a la composición etarea, destaca el 70% de hogares con personas en el rango de edad entre treinta y uno y cincuenta años, rango en el que se ubicaría gran parte de los jefes de hogar del grupo, aunque se debe mencionar que todos los otros rangos entre cero y sesenta años tienen porcentajes significativos.

En cuanto a la educación, existe un 36% hogares con personas que actualmente se encuentran en educación básica y en un 10% hay personas en educación media. Respecto a las personas que no están actualmente educándose, es importante señalar que el porcentaje que tiene educación superior completa es irrelevante.

Concerniente a la vivienda, sólo en el 59% de los casos, ésta pertenece a alguno de los miembros del hogar, lo que se explica por lo bajo de los ingresos y por la

juventud de los jefes de hogar. En la mayoría de los casos la vivienda es una casa o casa en cité, sin embargo existe un significativo 11% de casos en que ésta es una mediagua o mejora y en un 4% es un rancho, ruca o choza. Relacionado a lo anterior se observa que hay un 13% de los casos en que la conservación de la vivienda es mala.

En lo que respecta a las posesiones en el 50% de los casos no se posee nada de lo estipulado en el anexo 10 y en promedio poseen menos de una.

Formales (9,6% de la población)

Este grupo se caracteriza porque el 100% de los jefes de hogar, tiene pareja y todos son bien constituidas. El jefe de hogar promedio es joven, con una mediana de treinta y seis años, formando hogares de tamaño típico con una media de cuatro personas. Relacionado a lo anterior también se observa que en el 95% de los hogares se encuentran hijos, siendo la media de 1,7, y que por otro lado los porcentajes de hogares con nietos o padres o suegros del jefe de hogar son irrelevantes. Todo lo anterior sumado, nos muestra que estos hogares tienen la conformación típica, es decir el matrimonio con sus hijos.

El jefe de hogar tiene una escolaridad media de ocho años y el ingreso del hogar alcanza los \$107.000, al que aportan en promedio poco menos de dos personas, cuyas ocupaciones principales son las de trabajadores no calificados, agricultores, y artesanos u operarios.

Respecto a la razón urbano rural, se observa que el grupo es balanceado, con un leve predominio de los hogares rurales (52%).

En lo que se refiere a la composición etarea, es posible observar que en un 90% de los hogares existen personas con edad entre cero y catorce años, lo que se explica con los hijos del matrimonio. En tanto que los cónyuges de este grupo se distribuyen en el rango de edad de treinta y uno a cincuenta años (78% de los hogares), y en el rango de veintiuno y treinta años (48%)

En cuanto a las personas que actualmente se encuentran estudiando, destaca el 65% que está cursando educación básica actualmente y respecto a los que ya no lo están haciendo se observa que existe un importante porcentaje con educación básica incompleta (49%) y un también considerable 41% con educación media incompleta. En este grupo es irrelevante el porcentaje de personas con educación superior completa.

Concerniente a la vivienda sólo un 60% de los casos , ésta pertenece a algún miembro del hogar, lo que se explicaría tanto por la juventud de los jefes de grupo, como por su bajo ingreso. La vivienda es en la mayoría de los casos una casa o casa en cité, sin embargo es relevante el porcentaje en que ésta es una mediagua o mejora (8%), sumado al 2,5% de los casos en que es un rancho, ruca o choza. Relacionado a lo anterior se observa que hay cerca de un 10% de vivienda en mal estado de conservación, mientras que la urbanización es en promedio regular. En lo referente a las posesiones, en promedio sólo poseen uno de los artículos mencionados en el anexo 10 y además existe un significativo 35% que no posee ninguno de éstos.

Madrugadores (2,67% de la poblacio)

Este grupo se caracteriza por estar formado por hogares numerosos, con una mediana de 6 personas. Y porque sus jefes de edad son relativamente jóvenes, siendo la mediana de la edad de 40 años. Lo numeroso del hogar se explicaría por la abundante presencia de hijos, pudiendo encontrárseles prácticamente en todos los hogares del grupo y en gran cantidad, siendo la media superior a 4 hijos por hogar. No siendo relevante la presencia de padres o suegros del jefe de hogar, y con nietos presentes en el 14% de los hogares. Lo que indica la predominancia de un núcleo familiar acotado, es decir en la mayoría de los casos tiene la forma de la familia típica, lo que es respaldado también por el hecho de que la mayoría de los jefes de hogar tiene una pareja (90%). Sumado a lo anterior se observa que en el 71% de los hogares hay personas casadas, y en 19% hay parejas de convivientes.

La escolaridad de los jefes de hogar del grupo es baja, con una mediana de 6 años. Relacionado a lo anterior esta los bajos ingresos en los hogares del grupo, con una media de \$107.000. Llama la atención además que para juntar ese ingreso contribuyen en promedio mas de 3 personas. Indicando que en la mayoría de los casos deben contribuir al ingreso ambos cónyuges, además de hijos de estos. Los que como es de esperar se desempeñan en trabajos de poca calificación y mal remunerados. Entre estos destacan los porcentajes de hogares con personas que se desempeñan como trabajadores no calificados, y los que lo hacen como agricultores.

Lo anterior también guarda relación por el nivel educacional del hogar, de lo que destaca primeramente el 83% de hogares con personas con educación básica incompleta, además de la prácticamente nula presencia de personas con educación superior completa en el grupo. En cuanto a las personas que se están

educando en la actualidad, se observa que en casi la totalidad de los hogares hay personas cursando educación básica, y prácticamente todos es mas de uno. Esta cifra cae a solo un 16% de hogares con personas en educación media, explicable por la menor presencia de personas en la edad correspondiente, pero seguramente también a la deserción escolar.

Respecto al la estructura etarea destaca la casi totalidad de hogares con personas con edad entre cero y catorce años, encontrándose en la mayoría de los casos entre dos y cinco personas con esta edad. Por otro lado en el 95% de los hogares hay personas con edad entre 31 y 50, siendo la moda de dos. Esto se explicaría porque la mayoría de los jefes de hogar y sus parejas caerían en este rango de edad.

En cuanto a la vivienda esta solo en un 70% de los casos pertenece a alguno de los miembros del grupo familiar. Lo que sin duda tiene relación con la juventud y con los bajos ingresos. Además es destacable que en cerca de un 12% de los casos la vivienda es una mediagua o mejora, y en un 6% de los casos es un rancho, ruca o choza. Relacionado a lo anterior también esta el cerca de 23% de viviendas en mal estado de conservación que hay en el grupo. A lo que se suma un gran porcentaje de hogares con una urbanización entre mala y regular. Lo que seguramente también guarda relación con la predominancia rural que hay en el grupo, estando cerca del 69% de estos ubicados en zonas rurales.

Respecto a las posesiones, los hogares del grupo están pobremente dotados, de hecho hay cerca de un 62% de hogares en que no poseen ninguno de los artículos estipulados en el anexo 10 , y en promedio solo poseen 0,5 de estos.

Solitarios acompañados (3,95% de la población)

Este grupo se caracteriza por tener jefes de hogar prácticamente en la totalidad de los casos sin una pareja y sin embargo ser un hogar numeroso, con una mediana de cinco personas por hogar. Esto sumado a que en un 75% de los casos hay parejas casadas y en un 6% convivientes. Esto indica un hogares con su núcleo extendido, es decir un gran porcentaje de hijos del jefe de hogar, viven con su pareja, siendo en su mayoría bien constituidos. Otro dato importante es el 97% de hogares con hijos, siendo la media de 2,4 y el 40% de hogares con nietos. Llama la atención también, que en el 14% de los hogares se hallen personas viudas y el 12% de hogares con personas separadas o anuladas, sólo cubre una pequeña parte del porcentaje de jefes de hogar sin pareja en el grupo, lo que podría indicar que muchos de éstos fueron padres solteros.

En el grupo predominan los hogares en las zonas rurales, con 66% de los casos.

La escolaridad del jefe de hogar es baja, con una mediana de tan solo tres años, con una tasa de analfabetismo que alcanza el 26%. El ingreso del hogar alcanza en promedio los \$114.000, al que contribuyen en promedio más de dos personas, ocupándose principalmente como trabajadores no calificados y agricultores.

Se observa una variada composición etarea, lo que puede ser explicado por el núcleo extendido de los hogares.

En cuanto a la educación en el 66% de los hogares hay personas en educación básica en un 29% en educación media, lo que podría explicarse tanto por la presencia de hijos y de nietos. La baja educación del grupo se ve reflejada además por el 73% de hogares con personas con educación básica incompleta y el 49% con educación media incompleta.

Concerniente a la vivienda, en el 82% de los casos ésta pertenece a alguno de los miembros del hogar, cifra alta si se toma en cuenta el bajo ingreso del grupo, lo

que se podría explicar por la mayor edad de los jefes de hogar. En la mayoría de los casos ésta es una casa o casa en cité, pero existe un 5% en que ésta es una mediagua o mejora y un 4% en que ésta es un rancho, ruca o choza.

Relacionado a lo anterior, se observa que hay un 17% de hogares en mal estado de conservación.

Con respecto a las posesiones, en el 52% de los hogares no se posee ninguno de los artículos mencionados en el anexo 10 y en promedio se posee menos de uno de éstos.

Jóvenes Convivientes (5,1% de la Población)

Este grupo se caracteriza por tener jefes de hogar jóvenes, con una mediana de treinta y seis años y que en casi la totalidad de los casos tienen parejas, pero la totalidad de ellos se refiere a convivientes, formando hogares con una mediana de cuatro personas, encontrándose hijos en el 88% de los casos, con una media de 1,5 hijos. La mediana de la escolaridad del jefe de hogar es de ocho años y el ingreso del hogar alcanza los \$115.000, al que aportan en promedio dos personas, desempeñándose principalmente como trabajadores no calificados, agricultores y artesanos u operarios.

La razón urbano rural está balanceada en lo concerniente a la ubicación de los hogares, con un 52% presentes en zona urbanas.

En cuanto a la composición etarea destaca el 85% de hogares con personas en edad entre cero y catorce, lo que se explica por la presencia de hijos y por otra parte también son significativos el 72% correspondiente al rango de edad de treinta y uno a cincuenta y el 54% correspondiente al rango de veintiuno a treinta años, éstos dos últimos en donde se ubicarían el jefe de hogar y su pareja.

Concerniente a la educación hay un 50% de hogares con personas en educación básica, también es importante el 56% de hogares con personas con educación básica incompleta y el 41% con educación media incompleta.

En lo se refiere a la vivienda, sólo en un 52% de los casos ésta pertenece a algún miembro del hogar, lo que se explica por la juventud de los jefes de hogar y sus bajos ingresos. También es destacable que aunque en la mayoría de los casos la vivienda es una casa o casa en cité, en un significativo 14% ésta es una mediagua o mejora y en el 4% de los casos es un rancho, ruca o choza. Relacionado con lo anterior se observa que en el grupo hay un 14% de viviendas en un mal estado de conservación.

En cuanto a las posesiones, en el 50% de los casos, no posee ninguno de los artículos mencionados en el anexo 10 y en promedio poseen menos de uno de estos.

Arremangados (4,63% de la población)

Este grupo se caracteriza por estar formados por hogares pequeños, con una media de tres personas por hogar y con un jefe de hogar, que se encuentra ya en una etapa de edad madura siendo la mediana de cincuenta y seis años. En la mayoría de los casos, éste tiene una pareja (86%), en la gran mayoría bien constituida, además se observa que existe un 5% de viudos en el hogar y un poco menos de un 8% de separados o anulados. Respecto a hijos del jefe de hogar, su presencia es relativamente baja, sólo en el 40% de los hogares, esto puede explicarse por el ciclo de vida de la familia, deduciéndose que una gran parte de los hijos ya han dejado el hogar, también está la presencia de nietos que alcanza el 21%.

La proporción urbano rural está bastante balanceada, mostrándose un 53% de las viviendas ubicadas en zonas urbanas.

Los jefes de hogar del grupo, en promedio tienen una escolaridad de cinco años. Cabe señalar que existe un 11% de analfabetismo entre éstos y el ingreso del hogar alcanza en promedio los \$150.000, al cual aportan poco menos de dos personas, los cuales se desempeñan principalmente como trabajadores no calificados, agricultores y artesanos u operarios.

En cuanto a la composición etarea en prácticamente todos los rangos de edad se encuentran porcentajes significativos, pese a lo cual destaca fuertemente el 97% de hogares con personas entre cincuenta y uno y sesenta años, donde se ubica una gran parte de los cónyuges de este grupo.

Concerniente a los niveles educacionales del grupo se observa que hay un 27% de hogares con personas con educación básica incompleta y un 36% de hogares con personas con educación media incompleta, y sólo un 3% tiene enseñanza superior completa. Actualmente se observa en este grupo un 23% de hogares con

personas en educación básica, y sólo un 9% en educación media y no despreciable 5% en educación superior.

En lo que se refiere a las viviendas en este grupo se observa que en el 85% de los casos ésta pertenece a alguno de los miembros del hogar, lo que posiblemente tenga relación con la edad de los jefes de hogar. Se observa que en la gran mayoría de los casos la vivienda es una casa o casa en cité, en un estado de conservación bueno o por lo menos aceptable en la mayoría de los casos, a pesar de lo cual existe un no despreciable 7% en que ésta es mala.

Con respecto a las posesiones, se observa que en un 23% de los hogares del grupo no se posee ninguno de los artículos mencionados en el anexo 10 y en promedio poseen 1,4 de éstos.

Montepiados (6,97% de la población)

Este grupo se caracteriza por estar formado por hogares muy pequeños, con una mediana de dos personas y porque el jefe de hogar es de avanzada edad, siendo la media de setenta y tres años. Además sólo en el 42% de los casos el jefe de hogar tiene pareja, en la mayoría bien constituida, esto se explica fundamentalmente por el 45% de hogares con personas viudas y el 10% con personas anuladas o separadas. No son muchos los hogares en que se encuentran hijos, esta cifra sólo alcanza el 35%, con una media de menos de un hijo por hogar. También cabe señalar, que en un 21% de los hogares existe la presencia de nietos. De lo anterior se infiere que la mayor parte de los hogares del grupo, están en la última etapa del ciclo de vida, con un gran porcentaje de nidos vacíos.

En cuanto a la composición urbano rural, se observa el predominio de la primera, con un 65% de los casos.

En referencia a la escolaridad de los jefes de hogar del grupo, la mediana alcanza sólo los cuatro años, existiendo un índice de analfabetismo del 19% , el ingreso del grupo familiar es en promedio de \$158.000 al que aportan en promedio menos de dos personas, como es de esperar hay un alto porcentaje de hogares con personas jubiladas (62%), mientras que el porcentaje de jefes de hogar inactivo, sube hasta un 82%. De las personas laboralmente activas, éstas se desempeñan principalmente como trabajadores no calificados, agricultores, artesanos u operarios y vendedores.

En cuanto a la composición etarea, como es de esperar, el rango de edad que por lejos predomina es el de sesenta y uno o más años, presentes en prácticamente el 100% de los hogares del grupo.

En lo que respecta a la educación en los hogares del grupo hay escasa de personas estudiando en la actualidad y prácticamente nula la presencia de educación superior completa.

En lo respectivo a la vivienda en el 89% de los casos, ésta pertenece a alguno de los miembros del grupo familiar y en su gran mayoría son casas o casas en cité. Respecto a la conservación de las mismas hay un 6% de casos en que ésta es mala.

En cuanto a las posesiones existe un 22% de hogares en que no hay ninguno de los artículos listados en el anexo 10 y en promedio poseen 1,4 de éstos

Familia y Esfuerzo (5% de la población)

Este grupo se caracteriza porque en prácticamente en la totalidad de los casos, el jefe de hogar tiene pareja y se encuentra bien constituida. Siendo los hogares numerosos, con una mediana de cinco personas. Éstos son hogares en los que sus jefes son relativamente jóvenes, con una mediana de cuarenta y dos años, con una abundante presencia de hijos, encontrándose aquellos en prácticamente la totalidad de los hogares y con una media de tres hijos por hogar.

La proporción urbana rural está bastante balanceada en este grupo, con un 55% de los hogares ubicados en zonas urbanas.

La mediana de la escolaridad del jefe de hogar es de ocho años y el ingreso del hogar es en promedio de \$220.000, al que contribuyen en promedio dos personas, empleándose principalmente en trabajos de poca calificación

En cuanto a la composición etarea , vemos que hay dos rangos de edad que destacan. En primero en el 94% de los casos hay personas cuya edad fluctúa entre cero y catorce años, lo que se explica por la alto presencia de hijos y por el tampoco despreciable 11% con nietos. Además en el 98% hay personas en el rango de edad de treinta y uno y cincuenta años, en la mayoría dos, esto se explica por ser éste el rango de edad en que se encuentran la mayoría de los matrimonios del grupo.

En cuanto a la educación, existe un importante porcentaje de hogares con personas en educación básica (86%), explicable por la gran presencia de menores en este grupo. En cuanto a la personas de este grupo que no están educándose en la actualidad hay porcentajes significativos, tanto con educación básica incompleta, como con educación media incompleta. También es destacable el 5% de personas actualmente en educación superior.

En lo referente a la vivienda, en casi un 72%, ésta pertenece a algún miembro del grupo familiar y en la mayoría de los casos, se trata de una casa o casa en cité. Respecto a la conservación ésta es buena o aceptable, pero en poco más del 6%, ésta es mala.

Concerniente a las posesiones existe un 20% de los hogares en los que no se encuentra ninguno de los artículos mencionados en el anexo 10, y en promedio se posee un 1,4 de éstos.

Formales emergentes (5,8% de la población)

Este grupo destaca porque todos sus jefes de hogar tienen pareja, y todas son bien constituidas. Sobresale también la juventud de los jefe de hogar, siendo la mediana de treinta y cinco años, conformando hogares con una mediana de cuatro personas.

El 80% de los hogares está ubicado en zonas urbanas.

La mediana de la escolaridad de los jefes de hogar es de doce años, y el ingreso promedio del grupo es de \$232.000, al que contribuyen en promedio 1,5 personas. Las personas principalmente se desempeñan como artesanos, operarios u oficiales; operadores o montadores; trabajadores no calificados y vendedores.

Respecto a la presencia de hijos se les puede encontrar en la mayoría de los hogares (93%), con promedio de 2,4 hijos.

En cuanto a la composición etarea destaca el 85% de hogares con personas en el rango de edad de cero a catorce, dado principalmente por la gran presencia de hijos. Se destaca además el 51% de hogares con personas con edad de veintiuno a treinta años y el 75% en el rango de treinta y uno a cincuenta años, en estos dos últimos rangos se debieran ubicar las parejas que conforman los hogares.

Concerniente a las personas que se encuentran actualmente educándose, destaca el porcentaje de hogares con miembros en educación básica (56%) y respecto a los que no están educándose en la actualidad destaca el bajo porcentaje de hogares con personas con educación básica incompleta y por otro lado el 12% de hogares con personas con educación superior completa.

En cuanto a la vivienda sólo en el 60% de los casos ésta pertenece a algún miembro del hogar, esto puede deberse a la mayor juventud de los jefes de hogar

del grupo. En la mayoría de los casos, la vivienda es una casa o casa en cité y en un 5% se trata de un departamento.

La urbanización es en la mayoría de los casos buena o más que regular, lo que debe estar relacionado a la predominancia de hogares en zonas urbanas.

En cuanto a la conservación de la vivienda, ésta es en casi de la totalidad de los casos buena o aceptable.

Respecto a las posesiones, en promedio se poseen 2,2 de los artículos listados en el anexo 10 y sólo en un 7,5% no hay ninguno de ellos.

Años Plateados (3,57% de la población)

Este grupo se caracteriza por tener hogares relativamente pequeños, siendo la mediana de tres personas. También es característico que sólo en el 50% de los casos el jefe de hogar tiene una pareja, lo que seguramente esta relacionado, con su edad que en promedio es de sesenta y cinco años. Relacionado con lo anterior, está el hecho que en el 30% de los hogares se encuentran personas viudas, mientras que en el 10% de éstos hay personas separadas o anuladas.

Otra característica distintiva es la escasa escolaridad de los jefes de hogar, siendo la mediana de sólo dos años y con prácticamente la totalidad de los casos con seis o menos años de escolaridad. El índice de analfabetismo es también relevante alcanzando el 43%. Pese a lo anterior, el ingreso medio es de \$240.000, siendo en promedio dos las personas que aportan a éste. En este punto también es destacable indicar que en el 50% de los casos el jefe de hogar está inactivo y de las personas que trabajan destacan los porcentajes de años que desempeñan como agricultores y como trabajadores no calificados.

En cuanto a los hijos, éstos se pueden encontrar sólo en un 55% de los casos, siendo la media de menos de un hijo por hogar.

En cuanto a la composición etarea lo más destacable es que en un 80% de los hogares se encuentran personas de sesenta y uno o más años, siendo común encontrar entre uno o dos.

En cuanto a la educación, ésta es en la mayoría de los casos baja, un indicativo de esto es la prácticamente nula existencia de personas que hayan completado o estén cursando actualmente educación superior.

En cuanto a la vivienda, alrededor del 85%, ésta pertenece a algún miembro del hogar. Hay un porcentaje importante de casos en que la vivienda es una mediagua

o mejora (10%) y sólo un poco menor es el porcentaje en que se trata de un rancho, ruca, o choza.

La conservación de la vivienda es en promedio aceptable, pero en un 20% ésta es mala. También es relevante los porcentajes de hogares con urbanización deficiente.

En cuanto a las posesiones, los hogares del grupo están pobremente dotados, observándose que en más del 60% de los casos no se cuenta con ninguno de los artículos estipulados en el anexo 10.

Maduros solitarios (3,1% de la población)

Este grupo se caracteriza porque casi la totalidad de los jefes de hogar carece de una pareja, a pesar de su relativa juventud, siendo la mediana de la edad de 43 años. Conformando hogares pequeños con una mediana de dos personas. Entre las razones de la soledad del jefe de hogar podemos constatar, el 39% de hogares con personas separadas o anuladas en el grupo y el 25% de hogares con viudos, y seguramente un porcentaje que simplemente no se ha casado. Sumado a esto es también interesante señalar la predominancia de jefes de hogar de sexo femenino. Por otra parte también destaca el 59% de hogares con hijos en el hogar, el 10% de nietos, y el igual porcentaje de padres o suegros del jefe de hogar.

En cuanto a la escolaridad del jefe de hogar la mediana es de 11 años, y el ingreso promedio alcanza los \$246.000, al que contribuyen en promedio menos de dos personas, que principalmente se desempeñan como trabajadores no calificados, vendedores, oficinistas y artesanos u operarios.

En el grupo es predominantemente urbano, con un 80% de los casos en zona urbana.

En lo referido a composición etarea del hogar, hay que señalar que las edades son variadas, encontrándose que en todos los rangos de edad hay porcentajes significativos de hogares con personas en que pertenecen a él. Sobresaliendo algo sobre el resto el rango de edad de treinta y uno a cincuenta (65% de los hogares), lo que se explica por que en él, se ubican la mayoría de jefes de hogar del grupo.

En lo que respecta a la educación lo mas destacable son el 15% de hogares con personas con educación superior completa, y el 10% con personas que están cursando dicho nivel educacional.

Respecto a la vivienda esta solo en un 65% de los casos pertenece a alguno de los miembros del hogar. En la mayoría de los casos esta es una casa o casa en cite, aunque en un 7% de los casos esta es un departamento. La conservación de esta es en la mayoría buena o aceptable, mientras que la urbanización en la fran mayoría de los casos también esta en un buen nivel.

En cuanto a las posesiones los hogares del grupo en promedio poseen 2 de los artículos estipulados en el anexo 10 , y por otra parte hay no despreciable 15% de hogares que no tiene ninguno.

Papas por siempre (2,96% de la población)

Este grupo se caracteriza por estar conformado por hogares numerosos, con una mediana de seis personas, pero con un jefe de hogar de avanzada edad, que en promedio es de setenta años y que sólo en el 50% de los casos tiene pareja. Además puede observarse que en un 65% de los hogares hay parejas casadas y en un 10% de los casos hay dos, lo que sumado al 10% de hogares con convivientes nos da pie para inferir que en un porcentaje significativo los hogares del grupo tienen un grupo extendido, lo que también es corroborado por la alta presencia de nietos, presentes en el 87% de los hogares, mientras que en un 88% de los casos se encuentra la presencia de hijos, siendo la media de 1,7 de éstos por hogar. El alto porcentaje de jefes de hogar solitarios, se explicaría por el 44% de hogares con personas viudas, sumado al 23% con personas separadas o anuladas.

En cuanto a la proporción urbano rural de los hogares, es balanceada en forma perfecta .

El jefe de hogar presenta una escolaridad bastante baja, con una mediana de sólo tres años y con un índice de analfabetismo que alcanza el 31% entre éstos. El ingreso del hogar es en promedio de \$274.000 al que aportan en promedio poco más de tres personas. Es destacable que hay 48% de hogares con jubilados en el grupo y 74% de jefes de hogar inactivos. En cuanto a las personas laboralmente activas, éstas se desempeñan principalmente como trabajadores no calificados, agricultores, artesanos u operarios, vendedores, operadores y montadores y oficinistas.

En la composición etarea destaca el 94% de hogares con personas de sesenta y uno o más años, rango de edad donde se ubican casi la totalidad de los jefes de

hogar del grupo. También destaca el 82% de hogares con personas de cero a catorce años y el igual porcentaje en el rango de edad de treinta y uno a cincuenta años, lo que se explicaría por la presencia de nietos e hijos respectivamente.

En lo referente a la educación se observa un 64% de hogares con personas que actualmente están en educación básica, 17% en educación media y sólo un 4% en educación superior. Respecto a las personas que no están actualmente educándose, se observa que en un 49% de los hogares hay miembros con educación básica incompleta y en igual porcentaje con educación media incompleta.

Concerniente a la vivienda en un significativo 93% de los casos, ésta pertenece a alguno de los miembros del hogar, correspondiendo casi en la totalidad a casa o casa en cité y respecto a su conservación es destacable que existe casi cerca de un 11% en que la vivienda se encuentra en mal estado.

Es considerable destacar en cuanto a las posesiones, que en 28% no se posee ninguno de los artículos señalados en el anexo 10 y en promedio se poseen 1,2 de éstos.

Convivientes (2,74% de la población)

Este grupo presenta hogares de tamaño típico, en promedio de cuatro personas, con jefes de hogar relativamente jóvenes siendo la mediana de treinta y nueve años. Es característico del grupo que en la gran mayoría de los casos el jefe de hogar tiene pareja (97%) , pero en casi la totalidad de ellas no bien conformadas. En un alto porcentaje de los hogares se encuentran hijos (86%), pero la media sólo es de 1,3 hijos en el hogar. Cabe señalar también que en un 12% de los hogares existe la existencia de nietos.

En cuanto a la composición urbano rural predomina la primera con un 71% de los casos.

Respecto al nivel de escolaridad del jefe de hogar, la mediana del grupo es de nueve años y el ingreso en promedio alcanza los \$281.000 al que contribuyen en promedio poco más de dos personas, que se desempeñan principalmente como trabajadores no calificados, artesanos u operarios, vendedores, operadores o montadores, agricultores y en servicio doméstico.

En cuanto a la composición etarea, destaca el 78% de hogares con personas entre cero y catorce años atribuible a hijos y nietos del jefe de hogar y el 78% de hogares con personas entre treinta y uno y cincuenta años.

Respecto al nivel educacional hay un 50% de hogares con personas actualmente en educación básica y un 11% en educación media. En referencia a los hogares con personas que actualmente no están educándose hay un 34% con educación básica incompleta, y un 52% con educación media incompleta, mientras que sólo un 6% tiene educación superior completa.

Concerniente a la vivienda, en el 57% de los casos ésta pertenece a alguno de los miembros del hogar, esto puede deberse a la juventud del jefe de hogar y al bajo

ingreso que se percibe. La mayoría de las viviendas son casas o casas en cité, sin embargo en alrededor del 5% ésta es una mediagua o mejora y en un 4% de los casos es in departamento. En cuanto a la conservación de éstas en la gran mayoría es buena o aceptable, sin embargo hay un 6% en que es mala.

En lo que respecta a las posesiones en un 19% de los casos, no se posee ninguno de los artículos mencionados en el anexo 10 y en promedio se posee sólo un 1,6 de éstos

Prolíficos (1,68% de la población)

Este grupo se caracteriza por tener hogares muy numerosos, con una mediana de ocho personas. El jefe de hogar tiene una edad promedio de cincuenta y dos años. En el 82% de los casos el jefe de hogar tiene pareja, se observa que en el 75% de los hogares hay parejas casadas e incluso en el 15% hay dos en el mismo hogar, además 23% de los hogares tiene parejas de convivientes. Considerando lo anterior se puede concluir que hay un porcentaje significativo de hogares, con su núcleo extendido, lo que también se ve corroborado por el significativo 66% de hogares con nietos. La gran cantidad de personas en el hogar también se explica por la abundante presencia de hijos, a los cuales se les puede encontrar en prácticamente la totalidad de los hogares y con una media altísima de 4,1 en el hogar.

Tampoco puede dejar de mencionarse el 16% de hogares con personas separadas o anuladas, y el 12% de hogares con personas viudas, factores que explicarían el porcentaje de casos en que el jefe de hogar no tiene pareja.

La mediana de la escolaridad del jefe de hogar es de sólo cuatro años, puede observarse un significativo 19% de analfabetismo. El ingreso del grupo familiar alcanza los \$316.000, lo que se explicaría por las más de cuatro personas que en promedio aportan a él. Las personas del grupo se desempeñan principalmente como trabajadores no calificados, agricultores, artesanos u operarios y en servicio doméstico.

Predomina en este grupo la ubicación de las viviendas en zonas rurales alcanzando el 61%.

En cuanto a la composición etarea del grupo destaca el hecho de que en casi la totalidad de los hogares hay personas en el rango de edad de cero a catorce años,

lo que no es de extrañarse tomando en cuenta la abundante presencia de hijos y nietos, de hecho este mismo factor explica el 79% de hogares con personas entre quince y veinte años, y el 82% de hogares con personas de veintiuno a treinta años, en tanto que el 88% de personas en el rango entre treinta y uno y cincuenta y el 51% en el rango de cincuenta y uno a sesenta se explicaría principalmente por los jefes de hogar y sus parejas.

Con respecto a la educación hay un 77% de hogares con personas en educación básica y sólo un 29% en educación media, lo que podría ser un signo de alta deserción escolar y en sólo en un 4% de los hogares existen personas en educación superior. En cuanto a las personas que no se están educando en la actualidad, se observa un significativo 80% de hogares con personas con educación básica incompleta y un 64% con educación media incompleta.

En el 80% de los casos la vivienda pertenece a alguno de los miembros del grupo familiar, tratándose en la mayoría de los casos de una casa o casa en cité, y en un 3% es una mediagua o mejora. En cuanto a la conservación hay un significativo 15% en que ésta es mala.

Concerniente a las posesiones, se observa que en el 38% de los casos no se posee ninguno de los artículos mencionados en el anexo 10 y en promedio se posee uno de éstos.

Madurez en familia (3,58% de la población)

Este grupo se caracteriza por estar conformado con hogares bastantes típicos. La mediana del número de personas en el hogar es de cuatro y la edad promedio del jefe de hogar es de cincuenta y ocho años, los cuales, en la mayoría de los casos tienen pareja (95%), casi en su totalidad bien constituidas. Cabe mencionar la existencia de un 7% de hogares con personas separadas y un 3% con personas viudas. En cuanto a la presencia de hijos, se les puede encontrar en el 77% de los hogares, siendo la media de 1,2 hijos por hogar, es también destacable el 25% de hogares con nietos.

En este grupo predominan los hogares ubicados en zonas urbanas (63%).

La mediana de la escolaridad del jefe de hogar es de sólo seis años. El ingreso promedio en el hogar de este grupo alcanza los \$317.000 al que aportan en promedio poco más de dos personas, las que se desempeñan principalmente como trabajadores no calificados, artesanos u operarios, agricultores, vendedores y operadores o montadores. Cabe mencionar que en este grupo hay un 26% de hogares con personas jubiladas.

En cuanto a la composición etarea, hay porcentajes significativos en prácticamente todos los rangos de edad, sin embargo en este grupo, destaca el porcentaje de hogares con personas entre cincuenta y uno y sesenta años (83%), lo que refleja la mayor edad de los jefes de hogar de este grupo.

En lo que a educación se refiere vemos que existe un 25% con personas con educación básica incompleta y un 49% con educación media incompleta. Se destaca además la existencia de un 8% de hogares con personas con educación superior completa. Con respecto a las personas que actualmente se encuentran

estudiando se observan los siguientes porcentajes: un 22% en educación básica, un 12% en educación media, y un significativo 10% en educación superior. Concerniente a la vivienda de este grupo, en un 87% de los casos ésta pertenece a alguno de los miembros del grupo familiar e incluso existe un 10% en que se poseen dos. En la gran mayoría se trata de una casa o casa en cité. La conservación de la vivienda es en su gran mayoría óptima o aceptable. Referente a las posesiones del hogar en este grupo en promedio se observan 2,1 de los artículos mencionados en el anexo 10, pero existe un considerable 10% que no posee ninguno.

Jóvenes, Formales y Emergentes (4,38% de la población)

Este grupo se caracteriza por tener jefes de hogar relativamente jóvenes, con una mediana de cuarenta años y además porque prácticamente el 100% de ellos tiene pareja, siendo la gran mayoría bien constituida, formando hogares de tamaño normal, con una mediana de cuatro personas. Relacionado a lo anterior, se observa que en el 95% de los hogares, se encuentran hijos, siendo la media de 2. El jefe de hogar en este grupo tiene un nivel educacional aceptable, siendo la mediana de la escolaridad de doce años, obteniéndose un ingreso promedio del grupo familiar de \$ 397.000 , al que aportan en promedio poco menos de dos personas, cuyas principales ocupaciones son las de artesanos u operarios; vendedores; trabajadores no calificados; operadores o montadores; oficinistas y técnicos.

El grupo es predominantemente urbano, observándose un 80% de hogares ubicados en tales zonas.

En cuanto a la composición etarea, destaca el 82% de hogares con personas en el rango de edad de cero a catorce años, lo que se explica por la presencia de hijos en el hogar. El otro rango destacable es el que va desde treinta y uno a cincuenta años (88% de los hogares), lo que se explica por el hecho de que en este rango de edad, se ubica la mayoría de cónyuges del grupo.

Con respecto a la educación, destaca el 23% de hogares con personas con educación superior completa y el 11% de hogares con personas que actualmente cursan este nivel educacional.

En lo que respecta a la vivienda, sólo en un 68% de los casos ésta pertenece a algún miembro del hogar, lo que se explicaría por la relativa juventud de los jefes

de hogar, ésta es en la mayoría una casa o casa en cité, aunque en un 5% de los casos se trata de un departamento, su conservación es en prácticamente la totalidad de los casos, buena o aceptable, mientras que la urbanización, en la gran mayoría de los casos está en un buen nivel.

Relativo a las posesiones, vemos que este grupo está medianamente bien dotado, siendo la media de 2,9 de los artículos mencionados en el anexo 10.

Viudez temprana (0,87% de la Población)

El grupo está formado por hogares con una mediana de cinco personas. Y con un jefe de hogar promedio con 44 años de edad. En la mayoría de los casos este tiene una pareja (95%). Relacionado con lo anterior Esta el 80% de hogares con personas casadas en el grupo y además existe un 8% de estos en que hay dos parejas casadas en el mismo hogar. Además de un 10% de convivientes. También hay que señalar que en el 13% de los hogares hay personas separadas o anuladas y en un 83% de los casos se encuentran personas viudas. Última cifra que sin duda guarda relación con una de las características más distintivas del grupo que es la alta presencia de padres o suegros del jefe de hogar, encontrándoseles en el 97% de los hogares del grupo, pero en un 12% de los casos estos se encuentran también con su pareja. Por otra parte tampoco deja de ser relevante el 12,6% de hogares con nietos. Sumando todo lo anterior se puede concluir que este grupo se caracteriza por tener hogares con núcleos extendidos. En el grupo hay una cierta predominancia de hogares ubicados en zonas urbanas con un 65% de los casos.

Relativo a la escolaridad del jefe de hogar la mediana del grupo es de nueve años y el ingreso del grupo familiar es en promedio de \$397.000, ingreso al que aportan en promedio poco más de tres personas, los que se desempeñan principalmente como trabajadores no calificados, artesanos u operarios, agricultores, vendedores, operadores o montadores y oficinistas.

En cuanto a la composición etaria aunque en todos los rangos de edad hay porcentajes significativos, existen tres que destacan sobre los otros: en primer lugar el de cero a catorce años (72% de los casos), donde se ubicarían nietos y

buena parte de los hijos del jefe de hogar; luego destaca el 84% de hogares con personas con edad entre treinta y uno y cincuenta años, donde se ubicarían primordialmente los jefes de hogar del grupo con sus parejas y por último destaca el 93% de hogares con personas con sesenta y uno o más años, lo que se explicaría por la alta presencia de padres o suegros del jefe de hogar.

En lo que se refiere a educación un 58% de los hogares tiene personas cursando actualmente educación básica, un 18% educación media y un relevante 8% cursando educación superior. En cuanto a las personas que no están educándose actualmente hay un 35% de hogares con personas con educación básica incompleta, lo que aumenta al 58% al tratarse de educación media incompleta, también destaca el 9% de personas con educación superior completa en el grupo.

Respecto a la vivienda en el 86% de los casos, ésta pertenece a alguno de los miembros del grupo familiar, siendo prácticamente en la totalidad de los casos, una casa o casa en cité. En cuanto a su conservación cabe mencionar que en un no despreciable 5%, ésta es mala.

En lo que se refiere a posesiones hay un 13% de hogares en que no se posee ninguno de los artículos nombrados en el anexo 10 y en promedio poseen 2,1 de éstos.

Años Dorados (3,35% de la Población)

Este grupo se caracteriza por tener hogares relativamente pequeños, con una mediana de tres personas y con jefes de hogar de avanzada edad, siendo la media de setenta y dos años, teniendo éste pareja en el 65% de los casos, siendo en la mayoría bien constituidos. Es también relevante mencionar el porcentaje de personas viudas que alcanza el 29%, y el porcentaje de separados o anulados que es de un 15%. En cuanto a la escolaridad del jefe de hogar, la mediana es de cinco años, con un índice del 13% de analfabetismo. En cuanto al ingreso, la media es de \$415.000, al que aportan en promedio alrededor de dos personas. Hay que destacar que en un 63% de los hogares existen personas jubiladas y quienes actualmente trabajan se desempeñan principalmente como trabajadores no calificados, agricultores, vendedores y otras actividades de baja calificación. En cuanto a la proporción urbano rural, domina la primera, con un 63% de los hogares en zonas urbanas.

En cuanto a la presencia de hijos se les puede encontrar en el 57% de los hogares, en la mayoría de éstos entre uno y dos. También es significativa la proporción de nietos en el hogar, presentes en el 27% de los hogares.

En cuanto a la composición etarea, ésta es variada destacándose la proporción de personas con sesenta y uno o más años, presentes prácticamente en la totalidad de los hogares, siendo la moda de dos personas por cada hogar en ese rango.

En lo referente a la vivienda, en un 94% ésta pertenece a alguno de los miembros del hogar, tratándose en la mayoría de los casos de una casa o casa en cité. La conservación de éstas es, en la mayoría de los casos, buena o aceptable, aunque hay un 5% en que es mala.

Concerniente a las posesiones, en promedio se poseen dos de los artículos presentados en el anexo 10. En un 16% de los casos, no se posee ninguno de éstos.

Autónomos (16,5% de la población)

Este grupo se caracteriza por presentar hogares pequeños, con una mediana de dos personas y con un jefe de hogar que sólo en el 15% de los casos tiene pareja, esto sumado a que en la mitad de los casos el jefe de hogar es de sexo femenino. Como posibles explicaciones al bajo porcentaje de jefes de hogar con pareja, se puede constatar el 33% de hogares en que hay personas separadas o anuladas y el 19% con personas viudas, además de la posibilidad de personas solteras. Es interesante además mencionar el 50% de los hogares con hijos, con media de menos de un hijo por hogar. También cabe señalar que hay un no despreciable 7% de hogares con nietos y un 6% con padres o suegros del jefe de hogar. En cuanto a la escolaridad del jefe de hogar, la mediana es de doce años y el ingreso del hogar alcanza los \$ 455.000, monto al que aportan en promedio menos de dos personas. Las personas en el hogar se desempeñan principalmente como profesionales, vendedores, artesanos u operarios, oficinistas, trabajadores no calificados, técnicos y personas en el sector público. También debe señalarse que existe un 8% de hogares con jubilados.

En el grupo predomina los hogares ubicados en zonas urbanas, con 74%.

En cuanto a la composición etarea se puede señalar que las edades son variados, encontrándose porcentajes relevantes en todos rangos de edad, siendo el único que sobresale el de treinta y uno a cincuenta años (61%) de los hogares.

Concerniente a la educación, lo más destacable es el 32% de hogares con personas con educación superior completa y el 14% que cursa actualmente ese nivel educacional.

Con respecto a la vivienda, sólo en el 66% ésta pertenece a algún miembro del hogar. En la mayoría en los casos la vivienda es una casa o casa en cité, pero en un 9% de los casos corresponde a un departamento. Con respecto también a la vivienda se observa que en casi la totalidad de los casos la conservación de ésta es buena o aceptable.

Respecto a las posesiones se observa que el 16% de los hogares de este grupo no posee ninguno de los artículos mencionados en el anexo 10, pero en promedio, se poseen 2,5 de éstos, siendo variado el grupo en este sentido.

Achoclonados (2,54% de la población)

Este grupo se caracteriza principalmente por estar formado por hogares numerosos, con una mediana de seis personas, esto pese a que sólo el 77% de los jefes de hogar tiene pareja. Sin embargo si observamos el porcentaje de casados, vemos que éste alcanza a un 72%. Esto sumado a un 12% de hogares en que se encuentran dos parejas casadas, además del 22% de hogares con convivientes, lo anterior nos da claras señales de un porcentaje importante de hogares con su núcleo extendido, es decir, con alguno de los hijos del jefe de hogar emparejado. Otro factor que explicaría lo numeroso del hogar es la alta presencia de hijos, encontrándose casi en la totalidad de los hogares, siendo la media de 2,7 hijos en el hogar y así también la presencia de nietos, con un importante 55% de los hogares.

La proporción urbano rural está bastante balanceada en el grupo, encontrándose un 55% de los casos en zonas urbanas.

La mediana de la edad del jefe de hogar es de cincuenta y cinco años, en cuanto a la escolaridad de éste la mediana es de seis años, con una tasa de analfabetismo del 13%. A pesar de lo anterior, el promedio del ingreso del hogar alcanza los \$488.000, lo que se explicaría por el número de personas que contribuyen a éste, que en promedio es de 3,5. Es importante señalar que hay un 21% de jubilados en el grupo. En cuanto a las personas laboralmente activas, la mayoría se desempeña en trabajos de poca calificación.

Respecto a la composición etarea las edades son variadas, con porcentajes significativos de personas en prácticamente todos los rangos de edad.

Concerniente a la educación, se observa que hay un 58% de hogares con personas actualmente en educación básica, también se puede notar que existe un alto porcentaje de personas con educación media incompleta (62%) y también un alto porcentaje de hogares con personas con educación básica incompleta (53%). Respecto a la educación superior, en un 7% de los hogares existen personas con este nivel alcanzado, mientras que en alrededor del 9% hay personas cursándola actualmente.

En cuanto a la vivienda en un 86% pertenece a algún miembro del hogar y en un 11% poseen una segunda vivienda y en la mayoría de los casos la vivienda es una casa o casa en cité. La conservación de ésta es en su mayoría buena o aceptable, aunque en un no menor 5% de los casos ésta es mala.

En promedio estos hogares poseen 2,9 de los artículos citados en el anexo10.

Es importante señalar que en el 18% de los casos no se posee ninguno.

Pujantes (3,3% de la población)

Este grupo se caracteriza por tener jefes de hogar relativamente jóvenes, con una mediana de 41 años. Y por que tiene pareja en prácticamente la totalidad de los casos, siendo en la mayoría bien constituidos, formando hogares de tamaño típico, con una mediana de 4 personas. Además los jefes de hogar tienen un nivel educacional aceptable, con una mediana de 12 años de escolaridad. Obteniendo un ingreso en promedio de \$584.000. Al que contribuyen en promedio poco menos de 2 personas, Y con una moda justamente de dos. Indicando que en un gran porcentaje de casos trabajarían ambos cónyuges. En cuanto a las ocupaciones , vemos que las más comunes son la de oficinista; vendedor; como profesionales; artesanos, operarios y técnicos.

En el grupo el 85% de los hogares se ubican en zonas urbanas.

En cuanto a hijos, se les puede encontrar la gran mayoría de los hogares(93%), siendo de poco menos de dos hijos por hogar. Hay que señalar además que los porcentaje de hogares con nietos o padres u suegros del jefe de hogar son bajos. Indicando que en la mayoría de los casos el hogar tiene la constitución típica.

Respecto a la composición etarea destaca el 76% de hogares con edad entre cero y catorce explicado por la presencia de hijos en el hogar. Los que también explican el 38% de hogares con personas entre quince y veinte. En tanto la mayoría de los cónyuges se concentra en el rango de edad de treinta y uno a cincuenta, donde hay personas en el 85% de los hogares. Pero a la vez hay un no despreciable 44% de hogares con personas entre veintiuno y treinta que sería

explicado tanto por los cónyuges, como por hijos de los matrimonios de más edad en el grupo.

En cuanto a las personas actualmente educándose, hay un 55% de hogares con personas actualmente cursando educación básica y un 23% que lo está haciendo en educación media. Pero sin duda lo más destacable es el casi 19% de hogares con personas cursando educación superior. Y concerniente a las que no están estudiando en la actualidad, también destaca el 13% de hogares con personas con educación superior completa.

Referido a la vivienda se observa que en un 71% de los casos esta pertenece a alguno de los miembros del grupo familiar, e incluso hay un 10% de hogares en que poseen una segunda vivienda. Esta en la mayoría de los casos es una casa o casa en cité, mientras que en cerca del 6% esta es un departamento. La conservación de esta es en casi la totalidad de los casos buena o aceptable. Mientras que la urbanización también está en un buen nivel en la gran mayoría de los casos.

En cuanto a posesiones estos hogares están relativamente bien equipados poseyendo en promedio 3,6 de los artículos mencionados en el anexo 10.

Robles Otoñales (2,35% de la Población)

Este grupo se caracteriza por estar formado por hogares pequeños, con una mediana de tres personas y con jefes de hogar mayores, cuya mediana es de sesenta y dos años. Además se observa que sólo el 54% de los jefes de hogar tiene pareja, de los cuales la mayoría sería bien constituida. El amplio porcentaje de jefes de hogar solitarios se vería explicado por el 21% de hogares con personas separadas o anuladas y el 30% de hogares con personas viudas. Aunque a esta última cifra también deben contribuir seguramente el 8% de padres y suegros del jefe de hogar encontrados en el grupo. En cuanto a la presencia de hijos, se observan sólo en el 64% de los hogares, con una media de 1,7 hijos por hogar. En tanto que en un 22% de los hogares, se observa la presencia de nietos. Concerniente a la ubicación urbano rural de la vivienda, predomina la primera con 70% de los casos.

El jefe de hogar, tiene en promedio ocho años de educación y el ingreso del hogar alcanza en promedio los \$648.000, al cual contribuyen poco más de dos personas, desempeñándose en rubros diversos, entre los principales se encuentran los que están en el sector público, trabajadores no calificados, vendedores, oficinistas, agricultores, profesionales, técnicos, artesanos u operarios, operadores o montadores y como patrones. También es de gran relevancia en este grupo el porcentaje de hogares con jubilados, que alcanza el 45%.

Respecto a la composición etaria aunque en todos los niveles de edad hay porcentajes significativos, los más pronunciados se encuentran sobre los treinta

años, encontrándose la mayor presencia en el rango de sesenta y uno o más años (61% de los casos), lo que se explicaría por la edad de los jefes de hogar y de sus parejas en caso de tenerlas.

En lo que se refiere a la educación lo que más destaca es el 23% de hogares con personas con educación superior completa y el 11% de hogares con personas cursando actualmente ese nivel educacional.

En lo que se refiere a la vivienda en un 87% de los casos, ésta pertenece a alguno de los miembros del grupo familiar e incluso en un significativo 20% se posee una segunda vivienda, ésta en la gran mayoría de los casos es una casa o casa en cité aunque existe un 5% en que se trata de un departamento. La conservación de las viviendas es en su gran mayoría buena o al menos aceptable. En cuanto a las posesiones en los hogares del grupo, en promedio se posee 2,7 de los artículos indicados en el anexo 10, y además se observa que en un 9% no se posee ninguno.

Colmenas (0,11% de la población)

Este grupo se caracteriza por poseer hogares numerosos, con una mediana de siete personas. Prácticamente en casi la totalidad de los hogares se encuentran hijos del jefe de hogar, siendo la media alrededor de tres hijos, además en muchos de los casos parece existir un núcleo familiar extendido, es decir, que en una proporción importante existen hijos del jefe de hogar viviendo con una pareja. También es importante la existencia de nietos en el hogar, estando presentes en el 70% de los casos.

La proporción urbano rural de los hogares es de alrededor de sesenta a cuarenta. La edad del jefe de hogar es variada, la media es de cincuenta y nueve años. En cuanto a su educación, ésta es baja, con una mediana de seis años; sin embargo, tomando en cuenta todo el hogar existe una proporción interesante de hogares con personas con educación superior completa (12%) y una proporción similar con personas actualmente en educación superior.

A pesar de que el jefe de hogar tiene una educación bastante limitada, el ingreso total del hogar es bastante interesante, con una media de \$719.000, esto podría explicarse por el alto número de personas que aportan al ingreso del hogar, siendo la media de alrededor de cuatro personas, las que se desempeñan principalmente en ocupaciones de baja calificación aunque hay un no despreciable 10% de hogares con personas que se desempeñan como técnicos y alrededor de un 6% de los hogares con profesionales.

En cuanto a la composición etarea del hogar, cabe señalar que las edades son variadas.

En lo que se refiere a vivienda casi en el 90% de los casos ésta pertenece a alguno de los miembros del grupo familiar y en casi la totalidad de los casos se trata de una casa o casa en cité, cuya conservación es en la mayoría de los casos buena o aceptable, pero en un no menos importante 5%, ésta es mala.

En cuanto a las posesiones, los hogares están solo regularmente dotados, teniendo promedio solo dos de los artículos mencionados en el anexo 10, e incluso hay alrededor de un 15% de hogares que no tiene ninguno de estos, que incluyen algo tan básico como el refrigerador.

Madurez Tranquila (2,86% de la Población)

El grupo posee hogares con una media de cuatro personas, con un jefe de hogar que posee una edad promedio de cuarenta y siete años. En el 84% de los casos el jefe de hogar tiene pareja, relacionado a lo anterior se observa que hay un 77% de hogares con personas casadas y un 4% en que hay dos parejas. Además de un 10% de convivientes. El porcentaje de casos en que el jefe de hogar no tiene pareja, puede explicarse por el 10% de hogares con personas anuladas y el 9% en que hay personas viudas, también es importante el porcentaje de hogares con hijos (81%), siendo la media de 1,6 hijos, siendo relevante también el porcentaje de nietos en el hogar (12%) y el de padres o suegros (7%).

El grupo es predominantemente urbano, con un 80% de los hogares en estas zonas.

Respecto a la escolaridad del jefe de hogar la media es de doce años. El ingreso promedio del grupo es de \$838.000, al que aportan en promedio poco más de dos personas, las que se desempeñan principalmente como profesionales, en el sector público, como técnicos, vendedores y oficinistas.

En cuanto a la composición etarea destaca el 54% de hogares con personas en el rango de edad entre treinta y uno y cincuenta años y el 58% de hogares con personas entre cero y catorce años

En cuanto a la educación se observa que hay un 40% de hogares con personas en educación básica y un 19% en educación media. Pero sin duda lo más

relevante en el grupo es el 19% de personas actualmente en educación superior y en cuanto a los que no están educándose actualmente la presencia del 47% de hogares con personas en educación superior, existiendo incluso un 20% de hogares que tienen dos personas con este nivel educacional.

En lo que se refiere a la vivienda en un 78% de los casos, ésta pertenece a algún miembro del hogar e incluso hay un 17% de casos en que se posee una segunda vivienda. En la mayoría de los casos la vivienda es una casa o casa en cité, mientras que en un 7% es un departamento. La conservación de la vivienda es en prácticamente la totalidad de los casos buena o aceptable, sucediendo algo similar con la urbanización.

En lo que concierne a las posesiones, en promedio poseen 3,9 de los artículos mencionados en el anexo 10.

Cuello y Corbata (2,3% de la población)

Este grupo está formado por hogares de tamaño típico, con una mediana de cuatro personas. La mediana de edad del jefe de hogar es de cincuenta años, siendo sin embargo heterogéneo el grupo a este respecto. En la mayoría de los casos el jefe de hogar tiene pareja (81%). La mediana de la escolaridad de los jefes de hogar es de doce años. Estos hogares suman en promedio un ingreso de \$1.100.000, lo que se explicaría por el significativo porcentaje de hogares con personas con educación superior completa (50%), a lo cual que en el 22% de los hogares se encuentran dos personas con este nivel educacional. Además se debe señalar que en promedio para llegar a ese nivel de ingreso, son pocos más de dos las personas que aportan. Las principales ocupaciones en el grupo son las de desempeñarse como, profesionales, en el sector público, como técnico, vendedores y como patrón.

También se debe agregar que en este grupo existe un 26% de jubilados.

En cuanto a la ubicación de los hogares éstos se encuentran predominantemente en zonas urbanas (79%).

Respecto a la composición del hogar, se observa que en el 75% de los hogares hay parejas casadas y aún en un 6% hay dos parejas en el mismo hogar. Además de presentarse un 10% en que hay parejas de convivientes. Podemos encontrar que en un 12% de los casos hay personas separadas o anuladas y en un 15%

encontramos personas viudas. Respecto a la presencia de hijos, se les puede encontrar en la mayoría de los hogares (82%), siendo la media de 1,7, observándose además un no menor 10% de hogares con nietos y un 8% de hogares con padres o suegros.

En lo concerniente a la composición etarea, las edades son diversas, con porcentajes significativos en todos los rangos de edad.

En cuanto a las personas que se educan actualmente, destaca el 22% de hogares con personas en educación superior, mientras que en el 39% hay personas en educación básica.

Respecto a la vivienda en un 82% ésta pertenece a algún miembro del hogar, siendo destacable además que en un 20% de los casos se posee una segunda vivienda. En la mayoría de los casos la vivienda corresponde a una casa o casa en cité. Aunque existe un 6% en que ésta es un departamento. La conservación de ésta, es prácticamente la totalidad de los casos buena o aceptable, mientras que la urbanización en la gran mayoría de los casos está en un buen nivel.

En lo que se refiere a posesiones en promedio los hogares del grupo poseen 4,1 de los artículos que se enuncian en el anexo 10.

Nanas y Van (1,4% de la Población)

Este grupo está conformado por hogares con tamaño normal, con una mediana de cuatro personas. La media de la edad del jefe de hogar es de cincuenta años, siendo heterogéneo el grupo en ese sentido. En el 85% de los casos el jefe de hogar tiene pareja, la mayoría bien constituidas, destacando además un 12% de los hogares con personas viudas y el similar porcentaje con personas separadas o anuladas.

En este grupo el 78% de los hogares están ubicados en zonas urbanas.

Con respecto a la escolaridad del jefe de hogar la mediana es de trece años, teniendo un ingreso promedio cercano a \$1.500.000 , al cual contribuyen en promedio alrededor de aproximadamente dos personas. En el grupo hay un 11% de hogares con el jefe de hogar inactivo, mientras que en el 22% de los hogares encontramos personas jubiladas. Respecto a las ocupaciones de quienes se encuentran actualmente trabajando, destacan los que se desempeñan como profesionales, en el sector público, como patrones y los técnicos.

En cuanto a la presencia de hijos en los hogares, se les puede encontrar en el 80% de éstos, con una media de 1,7 hijos por hogar. También es importante el 14% de hogares con nietos y el 7% de hogares con padres o suegros del jefe de hogar.

En cuanto a la composición etarea, ésta es diversa, destacando el 75% de hogares con personas entre treinta y uno y cincuenta años.

Con respecto a la educación, lo más significativo es el alto porcentaje de hogares con personas con educación superior completa, que alcanza el 59%, encontrándose incluso un 31% de hogares con dos personas con este nivel educacional. En cuanto a las personas que se encuentran actualmente educándose, lo más destacable, son las que están en educación superior, presentes en el 23% de los casos.

Concerniente a la vivienda, en un 80% de los casos pertenece a alguno de los miembros del hogar, siendo destacable además el 26% de casos en que se posee una segunda vivienda. En la mayoría de los casos, ésta es una casa o casa en cité, aunque existe un no menor 8% de casos en que se trata de un departamento. La conservación de la vivienda en la mayoría de los casos es óptima.

En cuanto a las posesiones, se observa que estos son hogares bien dotados, que en promedio poseen 4.6 de los artículos mencionados en el anexo 10.

Profesionales exitosos (0,87% de la población)

Este grupo se caracteriza por tener jefes de hogar altamente educados, con una mediana de quince años de escolaridad. Asociado también a esto, se encuentra el alto ingreso en los hogares del grupo, siendo el promedio, de \$ 2.000.000 , dicho ingreso se alcanza con la contribución en promedio de alrededor de dos personas.

Este grupo se encuentra predominantemente ubicado en zonas urbanas (80%).

Son hogares de tamaño normal con una mediana de cuatro personas, en el 80% de los casos, el jefe de hogar tiene una pareja y en la mayoría son bien constituidos. La edad del jefe de hogar es variada, con una media de cincuenta años, existe un porcentaje no despreciable de hogares con personas viudas (12%) y en un porcentaje poco menor encontramos personas separadas o anuladas en el grupo.

Prácticamente en el 80% de los hogares se encuentra la presencia de hijos, en la mayoría de los casos entre uno y tres. Además en un 13% de los casos se encuentran nietos en el hogar.

En cuanto a la educación destaca que en alrededor del 70% de los casos hay personas con educación superior completa, siendo incluso significativo el porcentaje de hogares con dos personas con este nivel educacional. También es relevante el 22% de casos que actualmente están en educación superior. Otra

cosa digna de destacar es que en este grupo es casi insignificante el porcentaje de personas con educación básica incompleta.

En cuanto a las ocupaciones, las más relevantes son las personas que se desempeñan como profesionales, en el sector público, como técnicos o como patronos.

En lo que se refiere a vivienda, en el 85% de los casos, ésta es propiedad de alguno de los miembros del grupo familiar. Es destacable que en un 25% de los casos, se posee una segunda vivienda. En la mayoría de los casos ésta se trata de una casa o casa en cité, pero también hay un porcentaje relevante en que la vivienda es un departamento. En casi la totalidad de los casos, la vivienda se encuentra en un buen estado de conservación.

En cuanto a las posesiones, los hogares de este grupo están bien dotados, teniendo en promedio cinco de los artículos mencionados en el anexo 10.

Emprendedores exitosos (0,53% de la Población)

Este grupo está formado por hogares de tamaño típico, con una mediana de 4 personas. Los jefes de hogar tienen en promedio 50 años, aunque el grupo es variado en este aspecto. Estos se caracterizan principalmente por su alto nivel educacional, siendo la media de 14 años de escolaridad y la mediana de 16. La gran mayoría de los jefes de hogar tiene pareja (88%), siendo en la mayoría bien constituidos. El porcentaje de jefes de hogar sin pareja se explicaría por lo menos en parte por el 9% de hogares con personas viudas y el 8% con personas separadas o anuladas. Por otro lado, en el 80% de los hogares encontramos hijos, siendo la media de 1,7. También se debe mencionar el 8% de hogares con nietos y el 4% de hogares con padres o suegros.

En el grupo predominan los hogares en zonas urbanas (79%).

En cuanto al ingreso éste es en promedio de \$3.000.000 al que contribuyen en promedio poco más de dos personas, las que se desempeñan principalmente como patrón, en el sector público y como profesionales.

También relacionado a lo anterior está el 70% de hogares con personas con educación superior completa encontrándose incluso el 44% de los casos, dos personas con este nivel educacional en el mismo hogar. En cuanto a las personas que están actualmente estudiando hay un 38% con personas en educación básica, un 19% en educación media y un significativo 20% en educación superior.

En lo que se refiere a la composición etarea destaca el 76% de hogares con personas en el rango de edad de treinta y uno a cincuenta años, aunque hay que señalar que todos los demás rangos de edad tienen porcentajes significativos.

Concerniente a la vivienda, en un 87% de los casos ésta pertenece a alguno de los miembros del hogar, y en un significativo 38% se posee además una segunda vivienda. En la mayoría de los casos la vivienda es una casa o casa en cité, aunque en este segmento también es significativo el porcentaje en que son departamentos (10%) y el porcentaje en que se trata de casas en condominio (5%). La conservación de éstas es en su gran mayoría buena al igual que la urbanización.

Referente a las posesiones, los hogares del grupo poseen en promedio 5,2 de los artículos mencionados en el anexo 10.

Lujos y Viajes (0,25% de la población)

Los hogares de este grupo tienen un tamaño medio de cuatro personas. La edad del jefe de hogar es en promedio de cincuenta años. La gran mayoría de éstos tienen parejas (93%) y en su mayoría son bien constituidas. La diferencia que queda se explicaría por el 8% de hogares con personas separadas y el 7% de hogares con personas viudas. Respecto a la presencia de hijos éstos se pueden encontrar en el 80% de los hogares, siendo la media de 1,9 hijos, además se debe señalar que hay un 8% de hogares con nietos y un 4% de hogares con padres o suegros del jefe de hogar.

En el grupo domina la presencia de hogares ubicados en zonas urbanas (alcanza un 70%)

Este grupo se caracteriza por la alta escolaridad de los jefes de hogar, siendo la mediana de dieciséis años y también por el alto ingreso del hogar, el que en promedio es de \$4.480.000, al cual aportan poco más de dos personas, las cuales se desempeñan principalmente como patrón, en el sector público y como profesionales. Relacionado con lo anterior, se observa que en el 63% de los hogares hay personas con educación superior completa y en el 42% se encuentran dos personas con este nivel educacional. En cuanto a las personas que se encuentran actualmente educándose, vemos que hay un 37% de hogares con personas actualmente en educación básica, y un 23% en educación media,

además es muy significativo el porcentaje de hogares con personas en educación superior que es de un 22%.

La composición etarea en este grupo es variada, con porcentajes relevantes en todos los rangos de edad, de entre los cuales sobresale el que desde treinta y uno a cincuenta años, presente en el 80% de los casos.

Respecto a la vivienda, en un 87% de los casos ésta pertenece a alguno de los miembros del hogar y en un significativo 41% se posee además una segunda vivienda. En la mayoría de los casos, la vivienda es una casa o casa en cité, pero sin embargo es significativo el porcentaje en que ésta es un departamento (10%) y también los casos en que se trata de una casa en condominio (5%). La conservación de la vivienda es en la gran mayoría de los casos buena, al igual que la urbanización.

En cuanto a las posesiones, en promedio los hogares del grupo poseen 5,1 de los artículos mencionados en el anexo 10.

Crema y Nata (0,19% de la población)

Este grupo está formado por hogares con tamaño típico, con una mediana de 4 personas, los jefes de hogar tienen una edad promedio de 52 años, pero el grupo es variado respecto a esto. En la mayoría de los casos éstos tienen pareja (94%), la mayoría bien constituida. Además en alrededor de un 10% de los casos, se encuentran personas viudas en el hogar y en un 8% personas anuladas o separadas. En cuanto a la presencia de hijos, se les encuentra en el 80% de los hogares del grupo, siendo la media de dos. Además hay poco menos de un 6% de hogares con nietos.

En el grupo predominan las viviendas ubicadas en zonas urbanas, que alcanza el 70% de los casos.

El grupo se caracteriza marcadamente por el alto nivel educacional de los jefes de hogar, que tienen una media de catorce años de escolaridad y una mediana de diecisiete años y además por su alto nivel de ingreso cuya mediana es de \$8.200.000, ingreso al que aportan dos personas en promedio, desempeñándose principalmente como patrones en el sector público y como profesionales.

Relacionado también a lo anterior está el 73% de hogares en que hay personas con educación superior completa y en un 43% de los casos hay dos personas con este nivel educacional. De los que estudian actualmente se encuentran en los

siguientes porcentajes en los hogares: 37% en educación básica, 28% en educación media y un significativo 27% en educación superior.

En lo que concierne a la composición etarea ésta es variada, existen porcentajes significativos en todos los rangos de edad.

En lo que se refiere a la vivienda, en un 92% de los casos, ésta pertenece a algún miembro del hogar y en un 40% se posee una segunda vivienda. La mayoría de las viviendas corresponde a casa o casa en cité, pero en este grupo hay un significativo 12% en que ésta es una casa en condominio. La conservación de la vivienda es prácticamente en la totalidad de los casos buena.

Respecto a las posesiones, los hogares del grupo poseen en promedio 5,2 de los artículos mencionados en el anexo 10.

Capitulo V

Conclusiones

Con la investigación ya finalizada, se puede afirmar respecto a la hipótesis que se planteó, que ésta es validada por los resultados arrojados en el trabajo. En particular, respecto a la búsqueda e implementación de una metodología capaz de enfrentar el desafío planteado, puede afirmarse que ésta fue exitosa., constituyéndose en una herramienta valiosa a la hora de abordar problemas similares. Entre sus características está la gran ventaja de poder ser usada en situaciones en que se dispone de una base de datos particularmente grande, problema que impide el uso de una gran cantidad de metodologías destinadas a segmentar, como lo son por ejemplo los métodos jerárquicos de análisis de cluster. Para enfrentar dicho problema se hizo uso del método del algoritmo de k medias, aprovechando sus virtudes, pero a la vez complementándolo para así superar sus falencias.

Por otro lado, respecto a la afirmación propuesta en la hipótesis de que en la población chilena existen ciertas características demográficas, que hacen posible su agrupación, en grupos colectivamente exhaustivos y mutuamente excluyentes, ésta es confirmada por el éxito del procedimiento, donde se llega a una taxonomía de la población, dividiéndola en 31 grupos o conglomerados, en donde los grupos muestran ser internamente homogéneos, es decir, cada grupo está conformado por hogares entre sí similares, pero diferentes a los hogares de los otros grupos.

Dichas diferencias están dadas precisamente por diferentes características en variables demográficas, tales como ingreso, educación, edades, tamaño del hogar, etcétera.

En lo que respecta a la metodología es importante señalar algunas observaciones, ya que son importantes a la hora de realizar una buena segmentación.

En primer lugar, en la realización del trabajo, resultó evidente la importancia de poseer un set de variables adecuado, o como corolario, que es de vital importancia, eliminar aquellas variables que no contribuyen al análisis, quedando claro que éstas, no sólo no aportan sino que en realidad entorpecen en gran manera el análisis, dificultando la recuperación de una estructura verdadera dentro de los datos. Al respecto, el método de selección de variables empleado, demostró una gran efectividad, lo que quedó demostrado por el cambio en los resultados en el Gap Statistic, cuando se pasó de la situación en que se tomaron en cuenta todas las variables, a la que sólo consideró aquellas que habían sido previamente seleccionadas. En el primer caso, el resultado mostraba la ausencia de cualquier estructura en los datos, sugiriendo que éstos no debían ser sometidos a un análisis de conglomerado, dicha situación se vio claramente revertida en el segundo caso, mostrando la validez de la afirmación de que en la mayoría de las veces solo un subset de variables tiene en realidad valor, a la hora de definir una estructura de cluster.

En segundo lugar, un punto que demostró ser especialmente sensible a la hora de llevar adelante una segmentación a través del método de K medias, fue la de la estandarización de las variables, mostrándose durante el trabajo que los resultados pueden variar en gran manera dependiendo del proceso al que sean sometidos los datos. Sin embargo, por otro lado, también se hizo evidente durante el proceso, que cuando existen variables medidas en escalas tan disímiles, como el caso en cuestión y que se mueven en rangos tan diferentes, no resulta efectivo la aplicación directa del algoritmo de K medias, sin que los datos hayan sido sometidos a un proceso previo, esto se debe a que por la forma en que funciona el algoritmo, las variables que se mueven en rangos mayores, tienen más peso en el análisis, produciéndose así, un proceso de aglomeración poco balanceado, en el que pueden tener ingerencia sólo aquellas variables con más peso, relegando a las demás, pero como ya se mencionó la estandarización de las variables también tiene sus costos. Si existe una relación natural reflejada en la escala de las variables, como es el caso, la estandarización produce que ésta se pierda restando importancia a las variables conceptualmente más importantes, lo que luego lleva a la conformación de cluster mucho más homogéneos respecto a éstas, pudiendo afectar el fin último de la segmentación, consistente en que los grupos son internamente homogéneos y externamente heterogéneos. Tomando en cuenta todo lo anterior, una transformación del tipo que se usó en la investigación, puede ser un gran aporte, ya que constituye un equilibrio entre dos extremos, es decir, lleva a las variables a escalas razonablemente parecidas y a la vez no se pierde la relación natural de importancia que hay entre ellas.

Otro punto, que debe ser objeto de mayores consideraciones, a la hora de efectuarse un análisis de K medias, es el que concierne a los centros iniciales especificados al comenzar el algoritmo. La determinación de éstos puede ser un factor importante a la hora de buscar obtener un buen resultado, se debe señalar que no queda clara la efectividad del algoritmo al hacer uso de centros iniciales determinados automáticamente por el paquete estadístico, parece por lo tanto recomendable, realizar un análisis previo de éstos, para constatar que teóricamente tengan sentido.

Otra característica digna de destacar, respecto a la metodología utilizada es que ésta permite obviar la realización de un análisis factorial previo al análisis de cluster, proceso de dos etapas, que pese a que es ampliamente utilizado ha sido objeto de recientes críticas, por no hacer uso de la totalidad de la información disponible en las variables.

Con Respecto a los resultados mismos de la investigación, se puede concluir que se ha logrado llegar a una tipología rica y variada de la población chilena, dividiendo esta en treinta y un segmentos representativos de la realidad nacional. Los que se caracterizan y se diferencian entre si por una serie de variables demográficas, conformando grupos homogéneos de hogares.

Esta segmentación sin duda se constituye en una valiosa herramienta para al toma de decisiones y la asignación de recursos tanto en el ámbito publico como en el privado.

Además hay que señalar que el resultado obtenido, deja libre el camino para la creación de una segmentación geodemografica del país, derribándose así el principal obstáculo para llegar a ella. Lo que constituía el objetivo principal de la investigación.

Final mente en un ámbito personal, quiero manifestar lo estimulante que fue esta investigación, fueron meses llenos de desafíos, de intenso trabajo, y de mucha reflexión, que sin duda llevaron a un enriquecimiento personal. Y ahora que el trabajo concluye guardo una satisfacción por el buen trabajo realizado, que sin duda considero como completamente exitoso.

Bibliografía

- Arabie, P. y Hubert L. (1994) "Cluster Analysis in Marketing Research", In R. P. Bagozzi (Ed), *Advanced Methods in Marketing Research*, 160-189.
- Bacher, Johann (2002) "Cluster Analysis", University Erlangen-Nuremberg.
- Bello Cristian, Alejandro González y Cristian Chamorro (2000), "Geodemografía: Una Nueva Segmentación para Chile", Seminario para optar al título de Ingeniero Comercial, Universidad de Chile.
- Boyd and Walker (1993) "Customized Publication for Marketing Management", Irwin Professional Publishing.
- Breckenridge, James N. (2000) "Validating Cluster Analysis: Consistent Replication and Symmetry", *Multivariate Behavioral Research*, 35(2), 261-285.
- Brusco, Michael J. y J. Dennis Cradit (2001) "A Variable Selection Heuristic For K-Means Clustering", *Psychometrika*, vol. 66, No. 2, 249-270.
- Calinski, R.B. y J. Harabasz (1974) "A dendrite method for cluster analysis", *Communs Statist.*, 3, 1-27.

- Carmone, Frank J. , Ali Kara and Sarah Maxwell (1999) “HINoV: A New Model to Improve Market Segment Definition by Identifying Noisy Variables”, *Journal of Marketing Research*, Vol. XXXVI, 501-509.
- Everitt, B. S. (1992) “Cluster Analysis”, London: Edward Arnold.
- Fielder, John , John J. McDonald (1993) “Market Figmentation: Clustering on Factor Scores versus Individual Variables”, Paper presented to the AMA Advanced Research Techniques Forum.
- Fowlkes, E.B. y C.L. Mallows (1983) “A method for comparing two hierarchical clusterings (with comments and rejoinder)”, *Journal of the American Statistical Association*, 78, 553-584.
- Gnanadesikan, R. , J. R. Kettenring y J. L. Tsao (1995) “Weighting and Selection of Variables for Cluster Analysis”, *Journal of Classification*, 12, 113-136.
- Gordon, A. (1996) “Null models in cluster validation”, eds W. Gaul and D. Pfeifer, 32-44, New York: Springer.
- Green, Paul E. y Catherine M. Schaffer (1996) “An Empirical Comparison of Standardization Methods in Cluster Analysis”, *Multivariate Behavioral Research*, 31 (2), 149-167.
- Kaufman, L. y P. Rosseeuw (1990) “Finding Groups in Data: an Introduction to Cluster Analysis”, New York: Wiley.
- Malhotra Naresh K. “Investigación de Mercados: Un Enfoque Práctico”, segunda edición, Prentice Hall, Inc.
- Milligan, G. W. y M. C. Cooper (1985) “An Examination of Procedures for Determining the Number of Cluster in a Dataset”, *Psychometrika*, 50, 159-179.
- Milligan, G. W. y M. C. Cooper (1988) “ A Study of Variable Standardization”, *Journal of Classification*, 5, 181-204.
- Milligan, G. W. (1995) “Issues in Applied Classification”, *CSNA Newsletter*, 36-38.
- Tibshirani, Robert , Guenther Walther y Trevor Hastie (2001) “Estimating the number of cluster in a data set via the gap statistic”, *Journal of Royal Statistical Society*, 63, Part 2, 411-423.
- Wedel, Michel y Wagner Kamakura “Market Segmentation, Conceptual and Methodological Foundations”, Second Edition, International Series In Quantitative Marketing. 2000.

