



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

---

FACULTAD DE CIENCIAS

La Aplicación de Métodos Estadísticos  
Multivariados en la Investigación de Mercados  
Un enfoque dirigido al estudio del  
comportamiento del consumidor

T E S I S

QUE PARA OBTENER EL TÍTULO DE:  
ACTUARIO

PRESENTA:  
MARÍA LUISA MONDRAGÓN MARTÍNEZ

DIRECTOR DE TESIS:  
DR. FERNANDO BRAMBILA PAZ



2008



FACULTAD DE CIENCIAS  
 Secretaría General  
 División de Estudios Profesionales  
 Votos Aprobatorios

**Act. Mauricio Aguilar González**  
 Jefe de la División de Estudios Profesionales  
 Facultad de Ciencias  
**P r e s e n t e**

Por este medio hacemos de su conocimiento que hemos revisado el trabajo escrito titulado:

**Aplicación de Métodos Estadísticos Multivariados en la Investigación de Mercados. Un enfoque dirigido al estudio del comportamiento del consumidor**

realizado por **Mondragón Martínez María Luisa** con número de cuenta **0-9650321-0** quien ha decidido titularse mediante la opción de **tesis** en la licenciatura en **Actuaría**. Dicho trabajo cuenta con nuestro voto aprobatorio.

Propietario Tutor Dr. Fernando Brambila Paz

Propietario Act. Jaime Vázquez Alamilla

Propietario Act. Francisco Sánchez Villarreal

Suplente Mat. José Benjamín Medina Vera

Suplente M. en C. Elena de Oteyza de Oteyza

Atentamente,

“POR MI RAZA HABLARÁ EL ESPÍRITU”

Ciudad Universitaria, D. F., a 14 de abril de 2008

EL COORDINADOR DEL COMITÉ ACADÉMICO DE LA LICENCIATURA EN ACTUARÍA

ACULTAD DE CIENCIAS  
 CONSEJO DEPARTAMENTAL  
 MATEMÁTICA  
 ACT. ROBERTO CÁNOVAS THERIOT

Señor sinodal: antes de firmar este documento, solicite al estudiante que le muestre la versión digital de su trabajo y verifique que la misma incluya todas las observaciones y correcciones que usted hizo sobre el mismo.

## AGRADECIMIENTOS

A Dios por permitirme llegar hasta este momento tan importante de mi vida y lograr otra meta más en mi carrera.

A mis mamis porque a pesar de la tormenta siempre están de pie, por su fortaleza que me ha acompañado día con día, por todo su apoyo y confianza incondicional, por todo su amor que me ha permitido ser quien soy ahora, no saben cuanto las quiero y como doy gracias a Dios por tenerlas aún conmigo. Esta tesis es para ustedes.

A mi papi por ser una parte fundamental y sumamente importante en mi vida, porque a pesar de que ya no estás aquí siempre estás conmigo, gracias por guiarme sobre el camino de la educación, con el tiempo espero llenarte de orgullo, te quiero mucho pa', esta tesis también es tuya.

A Coco por tu apoyo, comprensión y amor que me permite sentir poder lograr lo que me proponga, por escuchar, por estar y formar parte de mi vida.

A toda mi familia por todos los momentos que a lo largo de este andar vivimos juntos, por soportar mi carácter, mis tristezas y todas mis locuras, los quiero.

Al Dr. Fernando Brambila por asesorarme a lo largo de la tesis y acompañarme en este camino que hoy culmina en el presente proyecto, por compartir su conocimiento conmigo e inspirar en mi mucha admiración.

A mis profesores, que participaron en mi desarrollo profesional durante mi carrera, sin su ayuda y conocimientos no estaría en donde me encuentro ahora.

A Anita por tu amistad, tus conocimientos y apoyo que me brindaste, estoy segura que sin tu ayuda hubiera tardado mucho más en presentar este trabajo.

A todos mis amigos que estuvieron conmigo y compartimos tantas aventuras, experiencias, desveladas, llanto, caídas y triunfos.

## ÍNDICE

<b>CAPÍTULO 1</b>		
<b>PLANTEAMIENTO</b>	<b>1</b>	
1.1	FORMULACIÓN DEL PROBLEMA	1
1.2	OBJETIVOS	1
1.2.1	Objetivos generales	1
1.2.2	Objetivos específicos	1
<b>MARCO TEÓRICO</b>		<b>2</b>
1.3	ANTECEDENTES	2
1.4	PLANTEAMIENTO TEÓRICO	2
1.4.1	La mercadotecnia	2
1.4.1.1	El marketing	2
1.4.1.2	Valor, satisfacción y calidad	2
1.4.1.3	Mercados	3
1.4.2	Dirección de marketing	3
1.4.3	Concepto de marketing	3
1.4.3.1	Globalización	3
1.4.3.2	Segmentación de mercado	3
1.4.4	El mercado consumidor	4
1.4.4.1	Estrategias de segmentos y segmentación	4
1.4.5	Estrategia de mercado meta	5
1.4.5.1	Producto	5
1.4.5.2	Posicionamiento de producto	5
1.4.6	La mezcla de mercadotecnia (4P's)	5
1.4.6.1	Estrategia de producto	5
1.4.6.2	Estrategia de distribución (Plaza)	6
1.4.6.3	Estrategia de promoción	6
1.4.6.4	Mezcla de promoción	6
1.4.6.5	Estrategia de precios	6
1.4.7	Investigación de mercados	7
1.4.7.1	Carácter interdisciplinario de la investigación de mercados	7
1.4.7.2	Contribución de la investigación de mercados	7
1.4.7.3	Adaptación metodológica de un proyecto de investigación de mercados	8
1.4.7.4	El desarrollo de formatos para recopilar información	8
<b>CAPÍTULO 2</b>		
<b>ANÁLISIS MULTIVARIADO</b>		<b>11</b>
2.1	INTRODUCCIÓN A LOS MÉTODOS MULTIVARIADOS	11
2.1.1	Objetivos de los Métodos Multivariados	11
2.1.2	Clasificación de los Métodos Multivariados	11
<b>TÉCNICAS SELECCIONADAS DE ANÁLISIS MULTIVARIADO</b>		
2.2	ANÁLISIS DE COMPONENTES PRINCIPALES (ACP)	12
2.2.1	Obtención de las componentes principales	13
2.2.2	Varianzas de las componentes	16
2.2.3	Aplicación e interpretación del análisis de componentes principales	17
2.3	ANÁLISIS CLUSTER	18

2.3.1	Pasos del análisis de conglomerados	18
2.3.2	Distancias y Similitudes	19
2.3.3	Clusters no jerárquico	19
2.3.4	Clusters jerárquicos: Dendograma	21

## APÉNDICE

	El concepto de Distancia y su aplicación en Estadística Multivariada	24
a.	Supuestos de la Distancia Euclidiana	25
b.	La Distancia en aplicaciones estadísticas	26

## CAPÍTULO 3

### METODOLOGÍA Y DESARROLLO 30

3.1	NIVEL Y TIPO DE INVESTIGACIÓN	30
3.1.1	Tipo de investigación	30
3.1.2	Nivel de la investigación	30
3.2	DISEÑO DE LA INVESTIGACIÓN	30
3.3	MÉTODO DE MUESTREO	30
3.3.1	Muestreo probabilístico	31
3.4	CALCULO DEL TAMAÑO MUESTRAL	31
3.5	TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS	33
3.6	ANÁLISIS DE LA INFORMACIÓN	33
3.6.1	Objetivo de la investigación	33
3.6.2	Población objetivo estudiada	33
3.6.3	Variables sociodemográficas	34
3.6.3.1	Total de la muestra por sexo	34
3.6.3.2	Total de la muestra por sexo y nivel socioeconómico	34
3.6.3.3	Distribución de estado civil en el total de la muestra por grupos de edad en mujeres	35
3.6.3.4	Distribución de estado civil en el total de la muestra por grupos de edad en hombres	35
3.6.3.5	Total de la muestra por niveles socioeconómicos y delegaciones	37
3.6.4	Ruido publicitario	38
3.6.4.1	Top of mind de producto en mujeres	38
3.6.4.2	Top of mind de producto en hombres	39
3.6.4.3	Top of mind de marca en mujeres	41
3.6.4.4	Top of mind de marca en hombres	43
3.6.5	Nivel de conocimiento de las marcas de leche	44
3.6.5.1	Conocimiento espontáneo de marca de leches	45
3.6.5.2	Conocimiento ayudado de marca de leches	52
3.6.5.3	Awareness de marca de leches	53
3.6.6	Nivel de recordación publicitaria de marcas de leche	55
3.6.6.1	Recordación espontánea de marcas de leche publicitadas	55
3.6.6.2	Recordación ayudada de marcas de leche publicitadas	61
3.6.6.3	Awareness publicitario de marca de leches	62
3.6.7	Indicadores de preferencia y fidelidad hacia las marcas de leche	64
3.6.7.1	Marcas de leches que ha consumido en los últimos 3 meses	64
3.6.7.2	Marca de leches que ha consumido en el último mes	65
3.6.7.3	Marcas de leches que ha consumido en la última semana	65
3.6.7.4	Marca de leche favorita	69

3.6.7.5	Marca comprada habitualmente	74
3.6.7.6	Tipo de leche de la marca comprada habitualmente	78
3.6.7.7	Marca sustituta	81
3.6.7.8	Marca que se comprará la próxima vez	85
3.6.7.9	Marca que nunca se compraría	88
3.6.8	Indicadores de comportamiento de compra	91
3.6.8.1	Persona que decide la marca a comprar	91
3.6.8.2	Marca comprada en la última semana	96
3.6.8.3	Key drivers en la decisión de compra	99
3.6.8.4	Frecuencia de consumo del producto	102
3.6.8.5	Lugar de compra	103
3.6.8.6	Frecuencia de compra	104
3.6.8.7	Frasas asociadas a las marcas	107
3.6.9	Sección Alpura	109
3.6.9.1	Publicidad de leches especializadas Alpura	109
3.6.9.2	Recordación slogan o frase publicitaria Alpura	113
3.6.9.3	Comerciales de televisión	114
3.6.9.4	Opinión de las personas que actualmente no consumen leche Alpura	125
3.6.9.5	Indicadores de compra de la marca Alpura	129

#### **CAPÍTULO 4**

4.1	CONCLUSIONES	142
4.2	RECOMENDACIONES	202

#### **ANEXO 1**

ENCUESTA	151
----------	-----

#### **ANEXO 2**

##### **ALGUNOS ELEMENTOS PARA LA ENSEÑANZA DE LA ESTADÍSTICA**

1.1	DEFINICIÓN Y UTILIDAD DE LA ESTADÍSTICA	165
1.2	CLASIFICACIÓN DE LA ESTADÍSTICA	165
1.2.1	Según la etapa	165
1.2.2	Según la cantidad de variables estudiada	166
1.3	POBLACIÓN Y MUESTRA	166
1.4	ESTRUCTURA DEL DATO	167
1.5	LA MEDICIÓN	167
1.6	ESTADÍSTICA DESCRIPTIVA	168
1.6.1	Medidas de tendencia central	169
1.6.1.1	Media aritmética el valor medio	169
1.6.1.2	Mediana	170
1.6.1.3	Moda	171
1.6.1.4	Media ponderada	171
1.6.2	Medidas de dispersión	172
1.6.2.1	Rango	172
1.6.2.2	Desviación estándar	172
1.6.2.3	Varianza	172
1.6.2.4	Desviación media	173
1.6.2.5	Coefficiente de variación	174
1.7	CORRELACIÓN Y REGRESIÓN	175
1.7.1	Análisis de correlación	175
1.7.1.1	Cálculo gráfico de la correlación	176
1.7.1.2	Cálculo analítico de la correlación	177
1.7.2	Análisis de regresión	177
1.8	ESTADÍSTICA INFERENCIAL	178

1.8.1	La estimación de parámetros	178
1.8.2	Pruebas de hipótesis	179
1.8.2.1	Formulación de la hipótesis de investigación y obtención de datos	179
1.8.2.2	Formulación de la hipótesis alternativa y la hipótesis nula	179
1.8.2.3	Selección de la prueba de hipótesis más adecuada	180
1.8.2.4	Determinación del nivel de significación	180
1.8.2.5	Determinación del tamaño de la muestra	180
1.9	GLOSARIO ESTADÍSTIVO	182
	<b>BIBLIOGRAFÍA</b>	<b>184</b>

## **“La Aplicación de Métodos Estadísticos Multivariados en la Investigación de Mercados”**

### **Un enfoque dirigido al estudio del comportamiento del consumidor**

#### **RESUMEN**

La investigación de mercado es la recopilación, registro y análisis sistemático de datos relacionados con problemas del mercado de bienes y servicios; es un enfoque sistemático y objetivo hacia el desarrollo y provisión de información aplicables al proceso de toma de decisiones. Lo sistemático se refiere a la necesidad de que el proyecto de investigación esté bien organizado y planeado, la objetividad implica que la investigación de mercado se esfuerza por ser imparcial e insensible en la realización de sus responsabilidades.

Mejorar la calidad de los productos y servicios es una de las principales tareas que se debe llevar a cabo para mantener e incrementar la cuota de mercado, dada la alta competitividad, por lo que se requieren estudios periódicos que permitan conocer el nivel de satisfacción general, preferencias y expectativas de los actuales clientes, así como sus principales mercados para enfocar hacia ellos todo el esfuerzo.

El aporte fundamental de este trabajo consiste en la obtención de un procedimiento metodológico que sobre la base de los ya existentes permite realizar investigaciones de mercado, con un enfoque hacia el estudio del comportamiento del consumidor, las ventas y el estado de la competencia enfatizando en el desarrollo de formatos válidos y confiables para recopilar información y en el procesamiento y análisis de la misma.

#### **INTRODUCCIÓN**

La investigación de mercados abarca desde la encuesta y el estudio pormenorizado de la misma hasta la realización de estudios estadísticos que permitan analizar las tendencias en el consumo, y poder prever así la cantidad de productos y la localización de los mercados más rentables para un determinado tipo de bien o servicio. Los métodos estadísticos constituyen hoy una herramienta fuerte y necesaria para tomar decisiones efectivas, atendiendo a las crecientes exigencias que suponen los mercados actuales y la necesidad de incluir numerosas variables que caractericen el problema.

La investigación de mercado se ha basado tradicionalmente en soportes cuantitativos independientemente del medio utilizado (personal, telefónico, etc.). La encuesta estadística estructurada es su máximo exponente y su objetivo es reflejar a través de una muestra estadísticamente representativa la realidad social o económica que sustenta a un mercado concreto.

La investigación cuantitativa, determina una visión estadística, de datos concretos, generalmente numéricos y siempre cuantificables. Los estudios de tipo cuantitativo nos permiten responder a las siguientes preguntas: qué, dónde, cuándo y cómo compra la gente.

A efectos metodológicos, las técnicas cuantitativas se basan en agrupar y medir a los individuos muestrales en categorías, en función de variables preestablecidas, tales como pautas de consumo, rasgos sociodemográficos, ejes lógicos de segmentación, etc.

Por tanto, los resultados obtenidos a partir de técnicas cuantitativas se validan exclusivamente con criterios estadísticos, por lo que no son adecuadas para reconstruir las relaciones sociales que subyacen a la medición concreta. En el mejor de los casos, se limitan a describirla. Es decir, no descubren procesos, simplemente los cuantifican.



La calidad de los productos y servicios es determinante para la retención de los actuales clientes y la captación de otros, lo cual permite lograr aumentos en la cuota de mercados.

Por ello se sugiere la realización de un estudio de mercado que permita lograr un aumento en los niveles de satisfacción de los clientes actuales y el descubrimiento de nuevos mercados.

La tesis viene desarrollada por capítulos, siguiendo los criterios del método científico, los cuales son descritos a continuación:

**Capítulo 1.** Planteamiento, se identifica el problema que se convierte en el objeto de reflexión sobre el cual se percibe la necesidad de investigar y se plantean los objetivos respectivos; Marco teórico, se embarca la investigación a conocimientos existentes y se asume una posición frente a ello.

**Capítulo 2.** Análisis Multivariado; objetivos y clasificación de los métodos multivariados, algunas técnicas seleccionadas del análisis multivariado como es el caso del análisis por componentes principales y el cluster, el concepto de distancia y su aplicación en la estadística multivariada.

**Capítulo 3.** Metodología y Desarrollo (Caso Práctico); nivel y tipo de la investigación, diseño de la investigación, método de muestreo, cálculo del tamaño muestral, técnicas e instrumentos de recolección de datos y análisis de la información. Se muestran los resultados más relevantes de la investigación, analizando los resultados de las encuestas realizadas.

**Capítulo 4.** Conclusiones, recomendaciones; se realiza un compendio de las conclusiones y recomendaciones de la investigación

Se complementa con 2 anexos:

**Anexo 1.** Encuesta, por medio de la cual se obtuvo la información necesaria para validar el trabajo de investigación.

**Anexo 2.** Introducción a la Estadística, definición y utilidad de la estadística, clasificación, población y muestra, estructura del dato, la medición, estadística descriptiva, correlación y regresión, estadística inferencial y un glosario estadístico. Se presenta como un manual de consulta de quienes cursan asignaturas donde se enseña la estadística como una herramienta de la metodología de la investigación científica.

## CAPITULO 1

### PLANTEAMIENTO

#### 1.1 FORMULACIÓN DEL PROBLEMA

Tener un histórico de la recordación, preferencia y actitudes de marca. Midiendo también recordación publicitaria, indicadores de preferencia y compra de marcas propias y de la competencia.

#### 1.2 OBJETIVOS

##### 1.2.1 Objetivos generales

Obtener una medición bimestral de los principales indicadores del desempeño de las marcas, para el establecimiento y evaluación de estrategias de mercado, en la categoría de leches.

Conocer la tendencia de las marcas a lo largo del tiempo para establecer relación con la inversión y el trabajo que se realiza con la marca (producto propio y competencia).

##### 1.2.2 Objetivos específicos

Determinar el nivel de conocimiento de las marcas.

- Top of Mind
- Conocimiento espontáneo
- Conocimiento ayudado
- Awareness de marcas

Establecer el nivel de recordación publicitaria.

- Top of mind de marcas publicitadas
- Recordación espontánea de marcas publicitadas
- Recordación ayudada de marcas publicitadas
- Awareness publicitario

Determinar los indicadores de preferencia y fidelidad hacia las marcas.

- Marcas consumidas últimamente
- Marca favorita
- Marca comprada habitualmente
- Marca sustituta
- Marca que se comprará la próxima vez
- Marca que nunca se compraría

Indicadores de comportamiento de compra.

- Persona que decide la marca a comprar
- Marca comprada en la última semana
- Key drivers en la decisión de compra
- Frecuencia de consumo del producto
- Lugar de compra
- Frecuencia de compra

Permitirá esta información identificar la evolución y modificaciones en el hábito de compra y frecuencia de consumo del producto.

## **MARCO TEÓRICO**

### **1.3 ANTECEDENTES**

Las estrategias de mercado y publicitarias se solidifican al contar con información continua que mida los efectos de las acciones.

No sólo se deben hacer esfuerzos para construir una marca, una vez ésta en el mercado se debe realizar un seguimiento sobre su desempeño.

### **1.4 PLANTEAMIENTO TEÓRICO**

Es muy importante crear valor y satisfacción en los clientes, como también lograr que la gente conozca la marca y el producto.

Cuando la marca logra un prestigio es mucho más sencillo penetrar en nuevos mercados y diferentes mercados en todo el mundo.

Es importante que la empresa este consciente de la relación que tiene el cliente con la marca para de esta forma poder promover mejor el producto. Para esto es necesario comprender la mercadotecnia y todas sus estrategias de mercado.

#### **1.4.1 La mercadotecnia**

Para el marketing lo más importante son los clientes. Como objetivo principal es poder entregar satisfacción a los clientes actuales y con ello obtener su aceptación y recomendación del producto, logrando atraer nuevos clientes y la completa satisfacción de los mismos.

##### **1.4.1.1 El marketing**

Se puede definir el marketing como: El proceso social y gerencial por el que los individuos y grupos obtienen lo que necesitan y desean, creando e intercambiando productos y valor con otros. Por esto si el encargado de marketing entiende bien las necesidades de los consumidores, desarrolla productos que ofrecen bien las necesidades de los consumidores, desarrolla productos que ofrecen mayor valor, les asigna precio apropiado, y los distribuye y promueve de manera eficaz, esos productos se venderán muy fácilmente<sup>1</sup>.

Todas las personas por cultura y personalidad individual tienen deseos los cuales se convierten en necesidades, dependiendo del origen de la persona serán sus necesidades. Las personas son capaces de pagar más dinero por obtener los productos debido a que les otorgan un gran valor y por ello es necesario hacer todo lo posible por conocer las necesidades, deseos y demandas de los consumidores.

##### **1.4.1.2 Valor, satisfacción y calidad**

El valor para el cliente es la diferencia entre los valores que el cliente obtiene al poseer y usar un productos y los costos de obtener el producto. Muchos clientes utilizan algunos productos

---

<sup>1</sup> Kotler y Armstrong. 1999. P3

como demostración de una imagen o estatus. No siempre los clientes tienen atención con los valores y costos con una gran exactitud, solo actúan según el valor percibido<sup>2</sup>.

La satisfacción de los clientes está íntimamente ligada con la calidad. En años recientes, muchas empresas han adoptado programas de administración de calidad total, diseñados para mejorar constantemente la calidad de sus bienes, servicios y procesos de marketing. La calidad afecta directamente el desempeño de los productos y por tanto, la satisfacción de los clientes<sup>3</sup>.

#### **1.4.1.3 Mercados**

El intercambio es sólo una de las muchas formas que tiene la gente de obtener un objeto deseado. Mientras que el intercambio es el concepto central de marketing una transacción es una unidad de medida de marketing. Más allá de crear transacciones a corto plazo, quienes hacen marketing necesitan forjar relaciones a largo plazo con clientes, distribuidores, concesionarios y proveedores valiosos<sup>4</sup>.

#### **1.4.2 Dirección de marketing**

La dirección de marketing es el análisis, planeación, implementación y control de programas diseñados para crear, formar y mantener intercambios provechosos con los compradores meta y así alcanzar las metas de la organización. La dirección de marketing administra la demanda, lo que a su vez supone administrar las relaciones con los clientes<sup>5</sup>.

#### **1.4.3 Concepto de marketing**

El marketing tiene grandes retos, opera dentro de un entorno global dinámico, y obliga a repensar objetivos y prácticas de marketing. La rapidez de los cambios puede hacer que en un poco de tiempo las estrategias ganadoras de ayer sean anticuadas.

El auge tecnológico ha creado nuevas e interesantes formas de estudiar y seguir a los clientes, crear productos adaptados para satisfacer necesidades, distribuir productos de forma más eficiente y eficaz, y comunicarse con los clientes en grupos grandes o individuales<sup>6</sup>.

##### **1.4.3.1 Globalización**

Según Kotler y Armstrong (1999), la economía ha sufrido cambios radicales en las últimas dos décadas. Las distancias geográficas y culturales se han encogido gracias a los adelantos tecnológicos, esto ha permitido a las empresas expandir considerablemente su cobertura geográfica de mercados, compras y fabricación.

Las empresas de América Latina han enfrentado en sus propios países, además de la competencia estadounidense el reto del hábil marketing de empresas multinacionales europeas y asiáticas<sup>7</sup>.

##### **1.4.3.2 Segmentación de mercado**

Es un proceso de división del mercado en subgrupos homogéneos de clientes con el fin de realizar una estrategia comercial diferenciada para cada uno de ellos. La segmentación se puede realizar en virtud de atributos objetivos y subjetivos. La tabla muestra las formas:

---

<sup>2</sup> Ibídem. P6

<sup>3</sup> Kotler y Armstrong. 1999. P8

<sup>4</sup> Ibídem. P 9

<sup>5</sup> Ibídem. P 10

<sup>6</sup> Ibídem. P 19

<sup>7</sup> Ibídem. P 21

	Generales	Específicos
<b>Objetivos</b>	Demográficos Socioeconómicos Geográficos	Estructura de consumo Uso del producto Fidelidad a la marca/empresa Tipo de compra: Primera, renovación Situaciones de compra Lugar de compra
<b>Subjetivos</b>	Personalidad Estilo de vida	Beneficio buscado Actitudes Percepciones Preferencias

En el mercado, la gente con requerimientos manifiestos o latentes constituye mercados, y la gente con sólo requerimientos crea oportunidades mercadotecnicas. La gente con requerimientos y poder de compra vuelve eficiente a los mercados. Un mercado con gente implica la existencia de público con poder de compra, por lo tanto, un mercado está compuesto de gente con dinero que requiere mercancía y servicios, y la oportunidad básica en mercadotecnia es proporcionar a esa gente mercancía y servicios satisfactorios de sus requerimientos<sup>8</sup>.

Por mercados del consumidor se abarca a aquellos individuos y hogares que compran y consumen mercancía y servicios para satisfacer sus requerimientos personales.

#### 1.4.4 El mercado del consumidor

La distribución por edades de la población ayuda a comprender las clases de mercancía y servicio que solicitan los consumidores, así como por los diferentes tipos de comportamientos exhibidos por el consumidor<sup>9</sup>.

##### 1.4.4.1 Estrategias de segmentos y segmentación

En las estrategias de segmentos y segmentación se puede notar que se ha presentado el concepto de mercado, y se han descrito varios mercados de importancia. Estos mercados, ya sea para el consumidor o para la industria, se estudiaron como conjuntos que abrazaban muchas clases distintas de clientes individuales mediante el empleo de variantes como la edad y el ingreso para describir a todos los clientes. El mercado del consumidor es un conjunto formado por individuos quizá no idénticos pero suficientemente semejantes para presentar apariencia de homogeneidad.

Se define así el concepto de segmento de mercado, como un grupo homogéneo de clientes con atributos similares y para el cual se cumplen las siguientes características: es identificable y con potencial mensurable, es accesible, es sustancial (alcanza una magnitud suficiente para ser rentable), los segmentos de un mercado deben ser diferentes, es posible llegar a ellos con los recursos de la empresa y, son defendibles ante la intrusión de la competencia.

La segmentación de mercados comprende el desarrollo de programas mercadotécnicos, diseñados por separado para alcanzar las necesidades de uno o más segmentos particulares del mercado. El proceso consta de dos partes: (1) Identificación de los segmentos viables, (2) Formulación de programas que incluyen estas partes como meta.

<sup>8</sup> Bell. 1979. P 124

<sup>9</sup> *Ibidem*. P 125

Una estrategia de segmentación deberá involucrar la evolución de un programa diferente de mercadotecnia para cada consumidor.

Existen varios tipos de mercado, el más importante, desde muchos puntos de vista es el mercado del consumidor. Este mercado compuesto por individuos y casas, puede analizarse de acuerdo a varias características, misma que incluyen factores demográficos y socioeconómicos, consideraciones especiales y aspectos familiares. Es importante concienciar acerca de cada una de las áreas, así como de los cambios que sucede en ellas, para entender el mercado del consumidor<sup>10</sup>.

#### **1.4.5 Estrategia del mercado meta**

La estrategia de mercadotecnia se refiere a las actividades de seleccionar y describir uno o más mercados meta a desarrollar y mantener una mezcla de mercadotecnia que produzca intercambios mutuamente satisfactorios con el mercado objetivo<sup>11</sup>.

##### **1.4.5.1 Producto**

El conocimiento de que un producto proporciona beneficios significativos y únicos a un segmento considerable de consumidores no siempre asegura su éxito. Las acciones de los competidores, de los consumidores y de los intermediarios en el proceso de distribución, pueden impedir el logro de un volumen y ganancias satisfactorias. Por ello es importante que la definición de oportunidad del mercado se base no solo en los deseos de los consumidores o usuarios, sino también en la consideración cuidadosa de los deseos de las empresas distribuidoras y del comportamiento de los consumidores<sup>12</sup>.

##### **1.4.5.2 Posicionamiento del producto**

Descripción objetiva, breve; es decir, cómo se quiere que el consumidor lo vea. La posición del producto es la forma en la cual los consumidores definen el producto en lo que concierne a sus atributos importantes, es el lugar que ocupa el producto en la mente de los consumidores en relación con los productos de la competencia<sup>13</sup>.

#### **1.4.6 La mezcla de mercadotecnia (4P's)**

El término mezcla de mercadotecnia se refiere a una mezcla distinta de estrategia de producto, distribución, promoción y precios, diseñada para producir intercambios mutuamente satisfactorios con un mercado objetivo. La distribución se conoce algunas veces como lugar o plaza, lo que nos da las "4P's" de la mezcla de mercadotecnia: producto, precio, plaza y promoción<sup>14</sup>.

##### **1.4.6.1 Estrategia de producto**

Para que una estrategia de producto o una política de la línea de producto sea efectiva, debe surgir de una evaluación completa y objetiva de la situación de la empresa y de su sensibilidad

---

<sup>10</sup> *Ibidem*. P 153

<sup>11</sup> Lamb, Hair y McDaniel. 1998. P 38

<sup>12</sup> Huges. 1978. P 320

<sup>13</sup> Kotler. 1998. P 223 - 224

<sup>14</sup> Lamb, Hair y McDaniel. 1998. P 24

a las necesidades del mercado. Debe derivarse de un reconocimiento completo sobre las ventajas y desventajas de las empresas<sup>15</sup>.

El punto de inicio, es la oferta y la estrategia de producto. El producto incluye no solo la unidad física, sino también su empaque, garantía, servicio posterior a la venta, marca, imagen de la empresa, valor y muchos otros factores<sup>16</sup>.

#### **1.4.6.2 Estrategia de distribución (Plaza)**

Se refiere a todas las actividades de negocio relacionadas con el almacenamiento y transporte de materias primas o productos terminados. La meta de la distribución es tener la certeza de que los productos llegan en condiciones de uso a los lugares designados, cuando se necesitan<sup>17</sup>.

#### **1.4.6.3 Estrategia de promoción**

Se tiene que tomar en cuenta, los efectos que se buscan obtener de los clientes actuales y potenciales para crear una personalidad distinta.

El papel de la promoción en la mezcla de mercadotecnia consiste en fomentar intercambios mutuamente satisfactorios con los mercados meta mediante la información, educación, persuasión y recuerdo de los beneficios de una empresa o producto<sup>18</sup>.

#### **1.4.6.4 Mezcla de promoción**

La conjugación eficaz de la publicidad, ventas, promoción y relaciones públicas, obtienen una adecuada mezcla de promoción. Consiste en la combinación específica de herramientas de publicidad, ventas personales, promoción de ventas y relaciones públicas, que la empresa usa para alcanzar sus objetivos de publicidad y marketing<sup>19</sup>.

- Publicidad. Cualquier forma pagada de presentación y promoción no personal de ideas, bienes o servicios, por un patrocinador identificado
- Ventas personales. Presentación personal que hace la fuerza de ventas de la empresa, con el fin de vender y forjar relaciones con el cliente
- Promoción de ventas. Incentivos a corto plazo que fomentan la compra o venta de un producto
- Relaciones públicas. Forjar buenas relaciones con los diversos públicos de la empresa, mediante la obtención de publicidad favorable, la creación de una buena imagen corporativa y el manejo o bloqueo de los rumores o sucesos desfavorables
- Marketing directo. Comunicación directa con consumidores seleccionados cuidadosamente, con el fin de obtener una respuesta inmediata (el uso de correo, teléfono, correo electrónico y otras herramientas no personales para comunicarse directamente con consumidores específicos o solicitar una respuesta directa

#### **1.4.6.5 Estrategia de precios**

El precio es lo que el comprador da a cambio para obtener un producto. Suele ser flexible de los cuatro elementos de la mezcla de mercadotecnia.

El precio representa una importante arma competitiva y resulta fundamental para la empresa como un todo, porque el precio multiplicado por el número de unidades vendidas es igual al ingreso total de la empresa.

---

<sup>15</sup> Hughes. 1978. P 319

<sup>16</sup> Lamb, Hair y McDaniel. 1998. P 40

<sup>17</sup> Lamb, Hair y McDaniel. 1998. P 40

<sup>18</sup> Ibídem. P 40

<sup>19</sup> Lamb, Hair y McDaniel. 1998. P 461

### **1.4.7 Investigación de mercados**

Se puede definir como la recopilación y análisis de información, en lo que respecta al mundo de la empresa y del mercado, realizado de forma sistemática o expresa, para poder tomar decisiones dentro del campo del marketing.

Se trata en definitiva, de una potente herramienta que debe permitir a la empresa obtener la información necesaria para establecer las diferentes políticas, objetivos, planes y estrategias más adecuadas a sus intereses.

#### **1.4.7.1 Carácter interdisciplinario de la investigación de mercados**

Para poder llevar a un buen término un análisis de mercado, es necesario aplicar diversos conocimientos adquiridos a través de las siguientes materias:

- La economía aplicada, psicología y la sociología. En la medida en que el funcionamiento del sistema económico se apoya en decisiones de mercado, el análisis del comportamiento del consumidor necesita conocimientos de psicología; la sociología se hace necesaria para el estudio de los grupos e instituciones de mercado
- La filosofía por la destacada importancia que tiene la lógica en la investigación aplicada
- La estadística y las matemáticas por su aporte fundamental en la cuantificación de los hechos detectado en la investigación
- La comunicación, por el diálogo que se produce de forma permanente en el trabajo de campo
- La dirección empresarial, ya que los objetivos que se persiguen con la investigación están estrechamente ligados al diseño de una estrategia y al cumplimiento de unas metas de venta, precios, productos y distribución
- La capacidad innovadora, aplicada al desarrollo de nuevos métodos eficaces y diferenciados en el diseño de soluciones rentables

#### **1.4.7.2 Contribución de la investigación de mercados**

##### ***En la toma de decisiones***

La investigación de mercados proporciona la información necesaria para la maduración de decisiones básicas y de largo alcance de la empresa que requieren un análisis cuidadoso de los hechos. Cuando las soluciones alternativas de los problemas son complejas, la toma de decisiones sin auxilio es peligrosa.

##### ***En la tarea directiva***

La investigación de mercados proporciona al directivo conocimientos válidos sobre cómo tener los productos en el lugar, momento y precio adecuados. No garantiza soluciones correctas pero reduce considerablemente los márgenes de error en la toma de decisiones.

##### ***En la rentabilidad de la empresa***

Básicamente contribuye al aumento del beneficio empresarial, ya que:

- Permite adaptar mejor los productos a las condiciones de la demanda
- Perfecciona los métodos de promoción
- Hace por una parte más eficaz el sistema de ventas y el rendimiento de los vendedores, y por otra reduce el costo de ventas



- Impulsa a los directivos a la reevaluación de los objetivos previstos
- Estimula al personal al saber que su empresa tiene conocimiento completo de su situación en el mercado y que se dirige hacia unos objetivos bien seleccionados

#### **1.4.7.3 Adaptación metodológica de un proyecto de investigación de mercados**

La Investigación de Mercados proporciona información sobre varios aspectos del sistema de mercadeo fundamentalmente para la toma de decisiones.

Numerosos autores han desarrollado el concepto de Investigación de Mercados sobre la base de diferentes enfoques, que pueden resumirse en la siguiente definición:

“Es el diseño y recopilación de información confiable y con calidad para identificar y definir problemas y oportunidades de marketing con que se enfrenta la empresa, utilizando métodos requeridos para tomar adecuadas decisiones”

En cualquier investigación de mercados se desarrollan las siguientes fases:

- Establecer la necesidad de información
- Especificar los objetivos de la investigación
- Formular el diseño de la investigación y determinar las fuentes de información
- Desarrollar formatos para recopilar información
- Diseñar la muestra
- Procesamiento y análisis estadístico de la información con técnicas avanzadas
- Presentación de los resultados

Las necesidades de información constituyen el punto de partida del proceso de cualquier investigación. Para no correr el riesgo de recopilar información innecesaria debe existir una estrecha relación entre las preguntas del cuestionario, las necesidades de información y los objetivos de la investigación.

La formulación del diseño de la investigación está determinada por la naturaleza de los objetivos o los tipos de investigación.

#### **1.4.7.4 El desarrollo de formatos para recopilar información**

En la recolección de información primaria, el investigador juega un papel fundamental necesitando de su creatividad y eficiencia para preparar formatos que establezcan un enlace entre las necesidades de información y las preguntas a realizar.

##### ***Diseño y organización del cuestionario***

El diseño y la organización del cuestionario tiene gran importancia, la utilización de encuestas es por lo general menos costosa y rechazada que otras variantes. Deben tenerse en cuenta las siguientes fases:

- Diseño
- Organización
- Selección de la escala de medición
- Fraseología

Existen varios criterios generales a tener en cuenta para diseñar y organizar el cuestionario, entre los que se encuentran la brevedad, la claridad en el objetivo propuesto, el lenguaje y en el formato de respuestas, lo lógico e interesante.

### ***Selección de la escala de medición***

La selección del formato (escala) de respuesta es un aspecto fundamental que se necesita tener en cuenta para diseñar un cuestionario. Su extrema importancia radica también en que determina cómo puede utilizarse y procesarse estadísticamente la información procedente del mismo. Existen diferentes escalas de respuestas posibles, entre las que se encuentran las de final cerrado y las de final abierto.

También existen escalas gráficas de clasificación que son especialmente útiles cuando los encuestados hablan diferentes idiomas.

### ***Fraseología empleada en el cuestionario***

A pesar de no existir una confirmación absoluta de su inadecuación, la mayoría de los autores no recomiendan las preguntas en forma negativa, pues tienden a confundir al respondiente y se observa mayor estabilidad en las respuestas enunciadas en forma positiva.

La fraseología debe ser clara y accesible para todo tipo de encuestado, cuidando también de no influenciar o conducir respuestas deseadas o indeseadas.

### ***Implementación del cuestionario***

Luego de diseñado el cuestionario, la próxima etapa es su implementación.

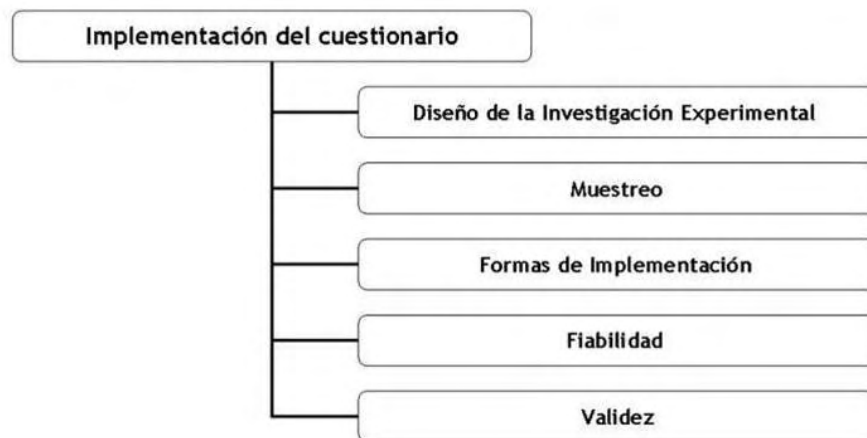


Figura 1.1 Implementación del cuestionario

### ***Diseño de la investigación experimental***

Existen diferentes diseños que se utilizan en la investigación de mercados, entre los que se encuentran, el diseño transversal y el longitudinal. Existen otros pero menos frecuentes como la comparación de grupos estáticos.

### ***Muestreo***

La posibilidad de que cualquier muestra represente a la población de donde fue extraída descansa en un grupo de factores.



Figura 1.2 Proceso de muestreo

Para establecer una expresión matemática que permita determinar la cantidad de unidades a muestrear es necesario tener en cuenta si el tamaño y la varianza de la población son conocidos o no. Una vez calculado el tamaño de la muestra, debe hacerse una estimación para la no respuesta, porque puede convertirse en una fuente de sesgo dentro de la muestra. (Lorente, 1999).

### **Formas de implementación**

La forma de implementación más utilizada es la aplicación directa mediante cuestionario, para la cual se requiere escoger los lugares y horarios más convenientes. Entre las más frecuentes se encuentran la entrevista personal, telefónica y por correo. La selección de una u otra forma dependerá de la complejidad y versatilidad, la cantidad y calidad de la información y el porcentaje de respuestas.

### **Fiabilidad y validez**

La fiabilidad se refiere al grado en que las puntuaciones percibidas en los diferentes ítems del cuestionario, están altamente interrelacionadas. Si un cuestionario es fiable, debe arrojar resultados estables, cuando lo aplican diferentes personas, bajo diferentes circunstancias.

Un método para estimar la confiabilidad de una escala o cuestionario es la fórmula para la estimación de la fiabilidad  $\alpha$  de Cronbach. Otro método es el conocido como estimación media de la fiabilidad, (Visauta, 1999).

No es suficiente que el cuestionario sea fiable sino que también debe ser válido. La validez expresa el grado en que un cuestionario mide lo que estaba diseñado para medir. Para medir la validez del cuestionario pueden utilizarse varios criterios correlacionales como la validez convergente, discriminante y de contenido, (Kinnear y Taylor, 1991).

## CAPÍTULO 2

### ANÁLISIS MULTIVARIADO

#### 2.1 INTRODUCCIÓN A LOS MÉTODOS MULTIVARIADOS

El análisis multivariado, definido en términos generales, son todos aquellos métodos estadísticos que están relacionados con el análisis de observaciones que contienen mediciones sobre más de una variable. Estos métodos permiten una comprensión global del fenómeno bajo estudio ya que indagan las relaciones que existen tanto entre variables como entre individuos.

Las aplicaciones de la estadística multivariada están presentes en distintas áreas o ramas de la ciencia, en investigación de mercado es útil para identificar características de los individuos con el propósito de determinar qué tipo de personas compran un determinado producto. Los datos multivariados surgen cuando a un mismo individuo se le mide más de una característica de interés. Un individuo puede ser un objeto o concepto que se puede medir.

Existen varios métodos de análisis multivariado, se presentarán sólo dos técnicas en este capítulo, para aquellos lectores que deseen profundizar en éstos, se recomienda consultar los siguientes autores: Hair, Anderson y Tatham (1995), Anderson T.W (1985), Grimm, L.G. and Yarnold P.R. Editors (1998).

##### 2.1.1 Objetivos de los Métodos Multivariados

- Simplificación. Los métodos multivariados son un conjunto de técnicas que permiten interpretar y visualizar conjuntos grandes de datos (tanto en individuos como en variables), a partir de su simplificación o reducción
- Relación. Encontrar relaciones entre variables, entre individuos y entre ambos
  - o Relación entre variables. Existe relación entre variables cuando las mismas miden características comunes
  - o Relación entre individuos. Existe relación entre individuos si alguno de ellos son semejantes entre sí

En los métodos multivariados, se supone que las variables están correlacionadas, pero las observaciones sobre los individuos son independientes. Generalmente se supone también que el conjunto de variables que intervienen en el análisis poseen una distribución normal multivariada. Esta suposición permite que el análisis multivariado se desarrolle paralelamente al correspondiente análisis univariado basado en una distribución normal.

##### 2.1.2 Clasificación de los Métodos Multivariados

El análisis multivariado en general se clasifica en dos grandes grupos\_

- Métodos explicativos. Como regresión lineal, análisis discriminante, regresión logística, modelos de respuesta probit, logic, modelos loglineales, entre otros
- Métodos descriptivos. Como análisis de conglomerados, análisis factorial, análisis de componentes principales, análisis de correspondencias simples y múltiples, etc.

## TÉCNICAS SELECCIONADAS DEL ANÁLISIS MULTIVARIADO

### 2.2 ANÁLISIS DE COMPONENTES PRINCIPALES (ACP)

El análisis de componentes principales es una técnica de análisis estadístico multivariante que se clasifica entre los métodos de simplificación o reducción de la dimensión y que se aplica cuando se dispone de un conjunto elevado de variables con datos cuantitativos persiguiendo obtener un menor número de variables, combinación lineal de las primitivas, que se denominan componentes principales o factores, cuya posterior interpretación permitirá un análisis simple del problema estudiado. Su aplicación es directa sobre cualquier conjunto de variables, a las que considera en bloque (a lo que podría llamarse técnica multivariada descriptiva), sin que se haya establecido jerarquías entre ellas, ni se necesite comprobar la normalidad de su distribución. Se trata por tanto de una técnica para el análisis de la interdependencia (ninguna variable destaca como dependiente principal).

El ACP permite describir, de un modo sintético, la estructura y las interrelaciones de las variables originales en el fenómeno que se estudia a partir de las componentes obtenidas, que, naturalmente, habrá que interpretar y nombrar. El mayor número posible de componentes coincide con el número total de variables. Quedarse con todas ellas no simplificaría el problema, por lo que se deberá seleccionar entre distintas alternativas aquéllas que, siendo pocas e interpretables, expliquen una proporción aceptable de la varianza global o inercia de la nube de puntos que suponga una razonable pérdida de información. Esta reducción de muchas variables a pocas componentes puede simplificar la aplicación sobre estas últimas de otras técnicas multivariantes, tales como clusters, regresión, entre otras.

El ACP permite reducir la dimensionalidad de los datos, transformando el conjunto de  $p$  variables originales en otro conjunto de variables  $q$  no correlacionadas, llamadas componentes principales. Las  $p$  variables son medidas sobre cada uno de los  $n$  individuos, obteniéndose una matriz de datos de orden  $np$  ( $p < n$ ). En el ACP existe la opción de usar la matriz de correlaciones o bien, la matriz de covarianzas. En la primera opción se le está dando la misma importancia a todas y a cada una de las variables; esto puede ser conveniente cuando se considera que todas las variables son igualmente relevantes. La segunda opción se puede utilizar cuando todas las variables tengan las mismas unidades de medida y además, cuando se juzga conveniente destacar cada una de las variables en función de su grado de variabilidad.

Las  $q$  nuevas variables (componentes principales) son obtenidas como combinaciones lineales de las variables originales. Los componentes se ordenan en función del porcentaje de varianza explicada. En este sentido, el primer componente será el más importante por ser el que explica mayor porcentaje de la varianza de los datos, quedando a criterio de cada analista decidir cuántos componentes se elegirán en el estudio.

Entre los usos más frecuentes del ACP están:

- Como técnica de análisis exploratorio que permite descubrir interrelaciones entre los datos y de acuerdo con los resultados, proponer los análisis estadísticos más apropiados
- Reducir la dimensionalidad de la matriz de datos con el fin de evitar redundancias y destacar relaciones. En la mayoría de los casos, tomando sólo los primeros componentes, se puede explicar la mayor parte de la variación total contenida en los datos originales
- Construir variables no observables (componentes) a partir de variables observables
- Bajo ciertas circunstancias, es de gran utilidad usar estos componentes incorrelacionados, como datos de entrada para otros análisis. Por ejemplo, en el caso de la regresión múltiple cuando las variables independientes presentan alta colinealidad es preferible haber la regresión sobre los componentes principales en lugar de usar las variables originales

### 2.2.1 Obtención de las componentes principales

En el análisis en componentes principales se dispone de una muestra de tamaño  $n$  acerca de  $p$  variables  $X_1, X_2, \dots, X_p$  (tipificadas o expresadas en desviaciones respecto a su media) inicialmente correlacionadas, para posteriormente obtener a partir de ellas un número  $k \leq p$  de variables incorrelacionadas,  $Z_1, Z_2, \dots, Z_p$  que sean combinación lineal de las variables iniciales y que expliquen la mayor parte de su variabilidad.

**La primera componente principal**, al igual que las restantes, se expresa como combinación lineal de las variables originales:

$$Z_{1i} = u_{11}X_{1i} + u_{12}X_{2i} + \dots + u_{1p}X_{pi}$$

Para el conjunto de las  $n$  observaciones muestrales esta ecuación puede expresarse matricialmente como:

$$\begin{bmatrix} Z_{11} \\ Z_{12} \\ \cdot \\ Z_{1n} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{p1} \\ X_{12} & X_{22} & \dots & X_{p2} \\ \cdot & \cdot & & \cdot \\ X_{1n} & X_{2n} & \dots & X_{pn} \end{bmatrix} \cdot \begin{bmatrix} u_{11} \\ u_{12} \\ \cdot \\ u_{1n} \end{bmatrix}$$

En notación abreviada se tendrá:  $Z_1 = Xu_1$

Tanto si las  $X_j$  están tipificadas, como si están expresadas en desviaciones respecto de su media muestral, la media de  $Z_1$  es cero, esto es,  $E(Z_1) = E(Xu_1) = E(Z)u_1 = 0$

La varianza de  $Z_1$  será:

$$V(Z_1) = \frac{\sum_{i=1}^n Z_{1i}^2}{n} = \frac{1}{n} Z_1' Z_1 = \frac{1}{n} u_1' X' X u_1 = u_1' \left[ \frac{1}{n} X' X \right] u_1 = u_1' V u_1$$

Si las variables están expresadas en desviaciones respecto a la media, la expresión  $\frac{1}{n} X' X$  (matriz de inercia) es la matriz de covarianzas muestral a la que se denomina  $V$  (caso más general) y para variables tipificadas  $\frac{1}{n} X' X$  es la matriz de correlaciones  $R$ .

La primera componente  $Z_1$  se obtiene de forma que su varianza sea máxima y sujeta a la restricción de que la suma de los pesos  $u_{1j}$  al cuadrado sea igual a la unidad, es decir, la variable de los pesos o ponderaciones  $(u_{11}, u_{12}, \dots, u_{1p})$  se toma normalizada.

Se trata entonces de hallar  $Z_1$  maximizando  $V(Z_1) = u_1' V u_1$  sujeta a la restricción  $\sum_{j=1}^p u_{1j}^2 = u_1' u_1 = 1$

Para resolver este problema de optimización con restricciones se aplica el método de los multiplicadores de Lagrange considerando la función lagrangiana:

$$L = u_1' V u_1 - \lambda (u_1' u_1 - 1)$$

Derivando respecto de  $u_1$  e igualando a cero, se tiene:

$$\frac{\partial L}{\partial u_1} = 2V u_1 - 2\lambda u_1 = 0 \Rightarrow (V - \lambda I) u_1 = 0$$

Se trata de un sistema homogéneo en  $u_1$ , que solo tiene solución si el determinante de la matriz de los coeficientes es nulo, es decir  $|V - \lambda I| = 0$ . Pero la expresión  $|V - \lambda I| = 0$  es equivalente a decir que  $\lambda$  es un valor propio de la matriz  $V$ .

En general la ecuación  $|V - \lambda I| = 0$  tiene  $n$  raíces  $\lambda_1, \lambda_2, \dots, \lambda_n$  que se pueden ordenar de mayor a menor  $\lambda_1 > \lambda_2 > \dots > \lambda_n$ .

En la ecuación  $|V - \lambda I| u_1 = 0$  se puede multiplicar por  $u_1'$  a la derecha, con lo que se tiene  $u_1' |V - \lambda I| u_1 = 0 \Rightarrow u_1' V u_1 = \lambda \Rightarrow V(Z_1) = \lambda$ . Por lo tanto para maximizar  $V(Z_1)$  se toma el mayor valor propio  $\lambda$  de la matriz  $V$ .

Tomando  $\lambda_1$  como el mayor valor propio de  $V$  y tomando  $u_1$  como su vector propio asociado normalizado ( $u_1' u_1 = 1$ ), ya se tiene definido el vector de ponderaciones que se aplica a las variables iniciales para obtener la primera componente principal, componente que vendrá definida como:

$$Z_1 = X u_1$$

**La segunda componente principal**, al igual que las restantes, se expresa como combinación lineal de las variables originales:

$$Z_{2i} = u_{21} X_{1i} + u_{22} X_{2i} + \dots + u_{2p} X_{pi}$$

Para el conjunto de las  $n$  observaciones muestrales esta ecuación puede expresarse matricialmente como:

$$\begin{bmatrix} Z_{21} \\ Z_{22} \\ \cdot \\ Z_{2n} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{p1} \\ X_{12} & X_{22} & \dots & X_{p2} \\ \cdot & \cdot & & \cdot \\ X_{1n} & X_{2n} & \dots & X_{pn} \end{bmatrix} \cdot \begin{bmatrix} u_{21} \\ u_{22} \\ \cdot \\ u_{2n} \end{bmatrix}$$

En notación abreviada se tendrá:  $Z_2 = X u_2$

Tanto si las  $X_j$  están tipificadas, como si están expresadas en desviaciones respecto de su media muestral, la media de  $Z_2$  es cero, esto es,  $E(Z_2) = E(X u_2) = E(Z) u_2 = 0$

La varianza de  $Z_2$  será:

$$V(Z_2) = \frac{\sum_{i=1}^n Z_{2i}^2}{n} = \frac{1}{n} Z_2' Z_2 = \frac{1}{n} u_2' X' X u_2 = u_2' \left[ \frac{1}{n} X' X \right] u_2 = u_2' V u_2$$

La segunda componente  $Z_2$  se obtiene de forma que su varianza sea máxima y sujeta a la restricción de que la suma de los pesos  $u_{2j}$  al cuadrado sea igual a la unidad, es decir, la variable de los pesos o ponderaciones  $(u_{21}, u_{22}, \dots, u_{2p})$  se toma normalizada.

Por otra parte como  $Z_1$  y  $Z_2$  han de estar incorrelacionados se tiene que:  $0 = E(Z_2' Z_1) = E(u_2' X' X u_1) = u_2' E(X' X) u_1 = u_2' V u_1$ , también se sabe que  $V u_1 = \lambda_1 u_1$  (ya que  $u_1$  es el vector propio de  $V$  asociado a su mayor valor propio  $\lambda_1$ ). Si se multiplica por  $u_1$  a la derecha se tiene:  $0 = u_2' V u_1 = \lambda_1 u_2' u_1 \Rightarrow u_2' u_1 = 0$  con lo que  $u_2$  y  $u_1$  son ortogonales.

Se trata entonces de hallar  $Z_2$  maximizando  $V(Z_2) = u_2' V u_2$  sujeta a las restricciones  $u_2' u_2 = 1$  y  $u_2' V u_1 = 0$

Para resolver este problema de optimización con dos restricciones se aplica el método de los multiplicadores de Lagrange considerando la función lagrangiana:

$$L = u_2' V u_2 - 2\mu(u_2' V u_1) - \lambda(u_2' u_2 - 1)$$

Derivando respecto de  $u_2$  e igualando a cero, se tiene:

$$\frac{\partial L}{\partial u_2} = 2V u_2 - 2\mu V u_1 - 2\lambda u_2 = 0$$

dividiendo por 2 y multiplicando por  $u_1'$  se tiene:

$$u_1' V u_2 - \mu u_1' V u_1 - \lambda u_1' u_2 = 0$$

Y como  $V u_1 = \lambda_1 u_1$  (ya que  $u_1$  es el vector propio de  $V$  asociado a su mayor valor propio  $\lambda_1$ , entonces  $u_1' V = \lambda_1 u_1'$  y se puede escribir la igualdad anterior como:

$$\lambda_1 u_1' u_2 - \mu V(Z_1) - \lambda u_1' u_2 = 0$$

Pero:  $u_1' u_2 = 0 \Rightarrow \mu V(Z_1) = 0 \Rightarrow \mu = 0$  de donde:  $\frac{\partial L}{\partial u_2} = 2V u_2 - 2\lambda u_2 = 0 \Rightarrow (V - \lambda I) u_2 = 0$

Se trata de un sistema homogéneo en  $u_2$ , que sólo tiene solución si el determinante de la matriz de los coeficientes es nulo, es decir,  $|V - \lambda I| = 0$ . Pero la expresión  $|V - \lambda I| = 0$  es equivalente a decir que  $\lambda$  es un valor propio de la matriz  $V$ .



En general la ecuación  $|V - \lambda I| = 0$  tiene  $n$  raíces  $\lambda_1, \lambda_2, \dots, \lambda_n$  que se pueden ordenar de mayor a menor  $\lambda_1 > \lambda_2 > \dots > \lambda_n$ .

En la ecuación  $(V - \lambda I)u_2 = 0$  se puede multiplicar por  $u_2'$  a la derecha, con lo que se tiene  $u_2'(V - \lambda I)u_2 = 0 \Rightarrow u_2'Vu_2 = \lambda \Rightarrow V(Z_2) = \lambda$ . Por lo tanto para maximizar  $V(Z_2)$  se toma el segundo mayor valor propio  $\lambda$  de la matriz  $V$  (el mayor ya se había tomado al obtener la primera componente principal)

Tomando  $\lambda_2$  como el mayor valor propio de  $V$  y tomando  $u_2$  como su vector propio asociado normalizado ( $u_2'u_2 = 1$ ), ya se tiene definido el vector de ponderaciones que se aplica a las variables iniciales para obtener la segunda componente principal, componente que vendrá definida como:

$$Z_2 = Xu_2$$

De forma similar, la componente  $h$ -ésima se define como  $Z_h = Xu_h$  donde  $u_h$  es el vector propio de  $V$  asociado a su  $h$ -ésimo mayor valor propio. Suele denominarse también a  $u_h$  eje factorial  $h$ -ésimo.

### 2.2.2 Varianzas de las componentes

En el proceso de obtención de las componentes principales presentado, la varianza de la componente  $h$ -ésima es:

$$V(Z_h) = u_h'Vu_h = \lambda_h$$

Es decir, la varianza de cada componente es igual al valor propio de la matriz  $V$  al que va asociada. Si la medida de la variabilidad de las variables originales es la suma de sus varianzas, dicha variabilidad será:

$$\sum_{h=1}^p V(X_h) = \text{traza}(V)$$

Ya que las varianzas de las variables son los términos que aparecen en la diagonal de la matriz de varianzas - covarianzas  $V$ .

Como  $V$  es una matriz real simétrica, por la teoría de la diagonalización de matrices, existe una matriz ortogonal  $P(P^{-1} = P')$  tal que  $PVP = D$  siendo  $D$  diagonal con los valores propios de  $V$  ordenados de mayor a menor en la diagonal principal. Por lo tanto:

$$\text{traza}(PVP) = \text{traza}(D) = \sum_{h=1}^p \lambda_h$$

Pero:  $\text{traza}(PVP) = \text{traza}(VPP') = \text{traza}(VI) = \text{traza}(V)$  con lo que ya se puede escribir:

$$\sum_{h=1}^p V(X_h) = \text{traza}(V) = \text{traza}(PVP) = \text{traza}(D) = \sum_{h=1}^p \lambda_h = \sum_{h=1}^p V(Z_h)$$

Se ha probado que la suma de las varianzas de las variables (inercia total de la nube de puntos) es igual a la suma de las varianzas de las componentes principales e igual a la suma de los valores propios de la matriz de varianzas-covarianzas muestral  $V$ .

### 2.2.3 Aplicación e interpretación del análisis de componentes principales

En la aplicación de un ACP, el primer paso consiste en la tipificación de las variables, ya que en la mayoría de los casos serán de magnitudes diferentes. La tipificación asegura que las diferentes magnitudes de las medidas obtenidas experimentalmente no influyan en el peso que posteriormente tendrán en el análisis. Sino se tipificaran, las variables con magnitudes elevadas dominarían la tendencia del ACP.

Antes de llevar a cabo el ACP propiamente dicho es recomendable la exploración de diferentes variables y el cálculo de la matriz de correlación, con el objeto de conocer las interrelaciones existentes entre las variables.

A continuación, se está en condiciones de aplicar el ACP. La primera información que se genera es la varianza explicada por cada uno de los componentes generados, que indica la varianza contenida en los datos originales que es explicada por cada uno de ellos. En general, la varianza explicada por los primeros componentes debe ser elevada para que se pueda obtener información válida a partir del estudio. Sin embargo en algunos estudios, la varianza explicada nunca es elevada, un ejemplo típico lo constituyen los estudios de tipo medioambiental, donde influyen multitud de variables que no son tenidas en cuenta, pues resulta imposible cuantificar la totalidad de variables que intervienen en un proceso que tiene lugar en la naturaleza. Los dos primeros componentes no suelen explicar más del 55 - 60% de la varianza, lo que se puede considerar suficiente para obtener información válida.

Los objetos de la matriz de datos ( $n$ ) se pueden representar en un espacio  $R^n$  según el valor calculado para los objetos respecto a cada componente, que se utilizan como coordenadas. Estos valores o pesos de los objetos o individuos para cada componente se calculan a partir de los valores propios y  $(u_{11}, u_{12}, \dots, u_{1p})$  que dan los coeficientes de cada variable en la combinación lineal generada, calculados a partir de la matriz de correlaciones:

El valor del primer objeto respecto al primer componente ( $Z_1$ ) sería:

$$Z_{1i} = u_{11}X_{1i} + u_{12}X_{2i} + \dots + u_{1p}X_{pi}$$

donde los coeficientes  $u$  se corresponden con  $(u_{11}, u_{12}, \dots, u_{1p})$  y  $X_1, X_2, \dots, X_p$  son las variables tipificadas para el primer objeto.

El cálculo de los valores  $Z_1, Z_2, \dots, Z_p$  permite representar los objetos en un espacio  $R^n$  definido por los  $n$  primeros componentes. Esto permite visualizar los objetos respecto a los primeros componentes. Puesto que la mayor parte de la varianza se debe explicar en los primeros componentes, las gráficas que se obtienen suelen estar referidas como máximo a los tres primeros componentes ( $R^3$ ), y en la mayoría de los casos sólo a los dos primeros. De igual manera, se pueden representar variables.

La interpretación del análisis consiste en encontrar las correlaciones positivas o negativas entre objetos o grupos de objetos, según su situación respecto a uno o varios componentes. Estas correlaciones están representadas por las distancias euclídeas en el espacio que se represente. Generalmente se obtiene suficiente información como para poder interpretar el significado de los componentes, relacionándolos con procesos físicos, químicos o biológicos dependiendo de la naturaleza del problema que se ha convertido en matriz.

## 2.3 ANÁLISIS CLUSTER

Es un conjunto de técnicas utilizadas para clasificar los objetos o casos en grupos homogéneos llamados conglomerados (clusters) con respecto a algún criterio de selección predeterminado. Los objetos dentro de cada grupo (conglomerado), son similares entre sí (alta homogeneidad interna) y diferentes a los objetos de los otros conglomerados o clusters (alta heterogeneidad externa). Es decir, que si la clasificación hecha es óptima, los objetos dentro de cada cluster estarán cercanos unos de otros y los cluster diferentes estarán muy apartados. Por ello, es también conocido como análisis de clasificación o taxonomía numérica.

El análisis cluster es un método estadístico multivariante de clasificación automática de datos. A partir de una tabla de casos - variables, trata de situar los casos (individuos) en grupos homogéneos, conglomerados o clusters, no conocidos de antemano pero sugeridos por la propia esencia de los datos, de manera que individuos que puedan ser considerados similares sean asignados a un mismo cluster, mientras que individuos diferentes (disimilares) se localicen en clusters distintos.

El enorme campo de aplicación en numerosas disciplinas, que se inició con la clasificación de las especies biológicas, ha propiciado la diversificación de este análisis, con denominaciones específicas tales como taxonomía numérica, taximetría, nosología, nosografía, etc. En la investigación de mercados es usado para la segmentación de mercados; comprensión del comportamiento del consumidor (identificación de grupos de consumidores homogéneos para analizar el comportamiento de cada grupo por separado); identificar oportunidades para nuevos productos, seleccionar mercados de prueba, reducción de datos con el fin de facilitar el manejo de la información.

La creación de grupos basados en similitud de casos exige una definición de este concepto, o de su complementario “distancia” entre individuos. La variedad de formas de medir diferencias multivariantes o distancias entre casos proporciona diversas posibilidades de análisis. El empleo de ellas, y el de las que continuamente siguen apareciendo, así como de los algoritmos de clasificación, o diferentes reglas matemáticas para asignar los individuos a distintos grupos, depende del fenómeno estudiado y del conocimiento previo de posible agrupamiento que de él se tenga.

### 2.3.1 Pasos del Análisis de Conglomerados

#### a. Formulación del problema

Lo más importante es la selección de las variables en las que se basará la agrupación. El conjunto de variables seleccionado debe describir la similitud entre los objetos en términos relevantes para el problema de investigación de mercados. Estas variables se seleccionan en base a investigaciones anteriores, la teoría o una consideración de las hipótesis que se prueban

#### b. Selección de una medida de similitud

Como el conglomerado agrupa objetos similares, se necesita una medida para evaluar las diferencias y similitudes entre objetos. La similitud (similitud) es una medida de correspondencia o semejanza entre los objetos que van a ser agrupados. Lo más común es medir la equivalencia en términos de la distancia entre los pares de objetos. Así, los objetos con distancias reducidas entre ellos son más parecidos entre sí que aquellos con distancias mayores y se agruparán por lo tanto, dentro del mismo cluster. Los tres métodos usados en la medición de la similitud son: las medidas de correlación y las medidas de distancia (usadas cuando se tienen variables métricas) y las medidas de asociación (usadas para variables categóricas)

### c. Estandarización de datos

Como las medidas de distancia son sensibles a la diferencia de escalas o de magnitudes hechas entre variables es necesaria la estandarización de datos para evitar que las variables con una gran dispersión tengan un mayor efecto en la similaridad. La forma de estandarización más común es restarle a cada observación la media de la variables y este resultado dividirlo entre su desviación estándar. Lo que se consigue con ello es eliminar las diferencias introducidas por las diferencias de escalas de las distintas variables (atributos) usados en el análisis.

Luego de seleccionar las variables y calcular las similaridades, se empieza con el proceso de agrupación, lo primero es seleccionar el algoritmo de agrupación para formar los grupos (clusters) y luego determinar el número de grupos que se van a formar. Estos dos procedimientos dependerán de los resultados que se obtengan y la interpretación derivada de ellos. Tipos de procedimientos de agrupación:

- Jerárquico. Se caracteriza por el desarrollo de una jerarquía o estructura de árbol (dendograma). De este modo, los clusters están formados solamente por la unión de los grupos existentes, así cualquier miembro de un cluster puede trazar su relación en un irrompible sendero que comenzaría con una simple relación
- No jerárquico. Asignan los casos a grupos diferenciados que el propio análisis configura, sin que unos dependan de otros

### 2.3.2 Distancias y Similitudes

La proximidad expresa la semejanza que existe entre individuos o variables. Es decir, es el grado de asociación que existe entre ellos. Las proximidades pueden medir la distancia o la similitud (similaridad) entre individuos o variables. El valor que se obtiene es una medida de distancia es tanto mayor cuanto más alejados están los individuos o puntos entre los que se mide. En las similitudes, al contrario de las distancias, el valor que se obtiene es tanto mayor cuanto más próximos están los elementos considerados. La correlación de Pearson y los coeficientes de Spearman y de Kendall son índices de similitud.

Más adelante en el apéndice se presentará y discutirá el concepto de distancia y posteriormente se examina su aplicación en Estadística Multivariada.

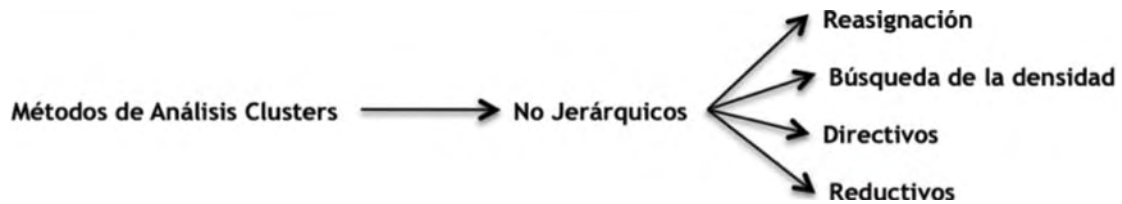
### 2.3.3 Clusters no jerárquicos

La clasificación de todos los casos de una tabla de datos en grupos separados que configura el propio análisis proporciona clusters no jerárquicos. Esta denominación alude a la no existencia de una estructura vertical de dependencia entre los grupos formados y, por consiguiente, éstos no se presentan en distintos niveles de jerarquía. El análisis precisa que se fije de antemano el número de clusters en que se quiere agrupar los datos.

Como puede no existir un número definido de grupos o, si existe, generalmente no se conoce, la prueba debe ser repetida con diferente número a fin de revisar la clasificación que mejor se ajuste al objetivo del problema, o la de más clara interpretación.

Los métodos no jerárquicos, también se conocen como métodos partitivos o de optimización, dado que tienen por objetivo realizar una sólo participación de los individuos en  $k$  grupos. Esto implica que se debe especificar los grupos que deben ser formados. Ésta es, posiblemente, la principal diferencia respecto de los métodos jerárquicos. La asignación de individuos a los grupos se hace mediante algún proceso que optimice el criterio de selección. Otra diferencia está en que estos

métodos trabajan con la matriz de datos original y no requieren conversión en una matriz de proximidades. Se agrupan los métodos no jerárquicos en las cuatro familias siguientes:



**Métodos de reasignación.** Permiten que un individuo asignado a un grupo en un determinado paso del proceso sea reasignado a otro grupo en un paso posterior si esto optimiza el criterio de selección. El proceso termina cuando no quedan individuos cuya reasignación permita optimizar el resultado que se ha conseguido.

Algunos de los algoritmos más conocidos dentro de estos métodos son el método de K-means (K-medias) de MeQuenn (1967), el Quick Cluster Analysis y el método de Forgy, los cuales se suelen agrupar bajo el nombre de métodos centroides o centros de gravedad. Por otra parte está el método de las nubes dinámicas, debido a Diday.

**Métodos de búsqueda de densidad.** Presentan una aproximación tipológica y una aproximación probabilística. En la primera aproximación, los grupos se forman buscando las zonas en las cuales se da una mayor concentración de individuos. Entre los algoritmos más conocidos dentro de estos métodos están el análisis modal de Wishart, el método de Taxmap de Carmichael y Sneath, y el método de Fortin. En la segunda aproximación, se parte del postulado de que las variables siguen una ley de probabilidad según la cual los parámetros varían de un grupo a otro. Se trata de encontrar los individuos que pertenecen a la misma distribución. Destaca en esta aproximación el método de las combinaciones de Wolf.

**Métodos directos.** Permiten clasificar simultáneamente a los individuos y a las variables. Las entidades agrupadas, ya no son los individuos o las variables, sino que son las observaciones, es decir, los cruces que configuran la matriz de datos.

**Métodos de reducción de dimensiones.** Este método consiste en buscar factores en el espacio de los individuos, correspondiendo cada factor a un grupo. La interpretación de los grupos puede ser compleja dado que cada individuo puede corresponder a varios factores diferentes.

Resulta muy intuitivo suponer que una clasificación correcta debe ser aquella en que la dispersión dentro de cada grupo formado sea la menor posible. Esta condición se denomina criterio de varianza, y lleva a seleccionar una configuración cuando la suma de las varianzas dentro de cada grupo (varianza residual) sea mínima.

Se han propuesto diversos algoritmos de clasificación no jerárquica, basados en minimizar progresivamente esta varianza, que difieren en la elección de los clusters provisionales que necesita el arranque del proceso y el método de asignación de individuos a los grupos. Aquí se describen los dos más utilizados:

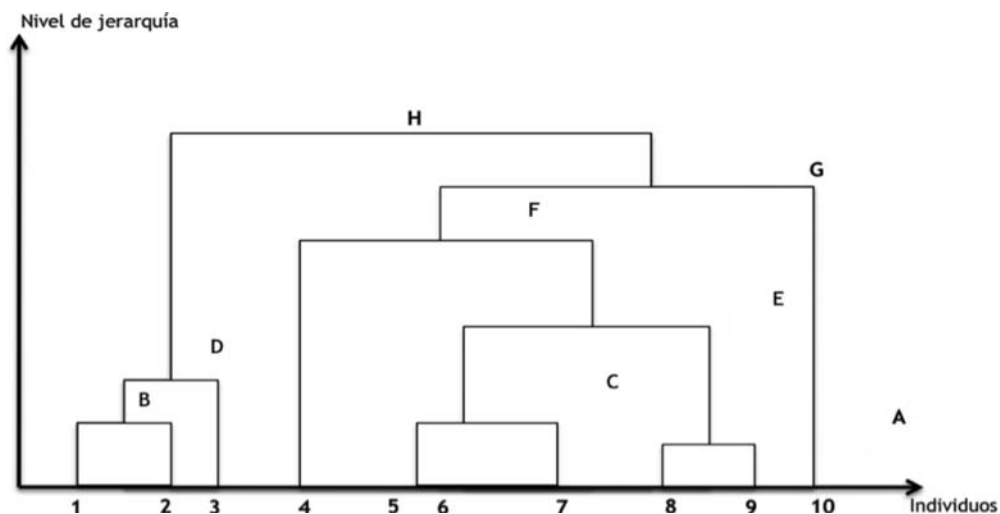
- a) **Algoritmo de las H-medias.** Parte de una primera configuración arbitraria de grupos con su correspondiente media, eligiendo un primer individuo de arranque de cada grupo y asignando posteriormente cada caso al grupo cuya media es más cercana. Una vez que todos los casos han sido ubicados, calcula de nuevo las medias o centroides y las toma en lugar de los primeros individuos como una mejor aproximación de los mismos, repitiendo el proceso mientras la varianza residual vaya disminuyendo. La partición de arranque define el número de clusters que, lógicamente, pueda disminuir si ningún caso es asignado a alguno de ellos

b) **Algoritmo de las K-medias.** El más importante desde los puntos de vista conceptual y práctico, parte también de unas medias arbitrarias y, mediante pruebas sucesivas, contrasta el efecto que sobre la varianza residual tiene la asignación de cada uno de los casos a cada uno de los grupos. El valor mínimo de varianza determina una configuración de nuevos grupos con sus respectivas medias. Se asignan otra vez todos los casos a estos nuevos centroides en un proceso que se repite hasta que ninguna transferencia puede ya disminuir la varianza residual; o se alcance el criterio de parada: un número limitado de pasos de iteración o simplemente, que la diferencia obtenida entre los centroides de dos pasos consecutivos sea menor que un valor prefijado. El procedimiento configura los grupos maximizando, a su vez, la distancia entre sus centros de gravedad. Como varianza total es fija, minimizar la residual hace máxima la factorial o intergrupos. Y puesto que minimizar la varianza residual es equivalente a conseguir que sea mínima la suma de distancias al cuadrado desde los casos a la media del cluster al que van a ser asignados, es esta distancia euclídea al cuadrado la utilizada por el método

Matemáticamente, un método de clasificación no jerarquizado consiste en formar un número prefijado de  $k$  de clases homogéneas excluyentes, pero con máxima divergencia entre las clases. Las  $k$  clases o clusters forman una única partición (clustering) y no están organizadas jerárquicamente ni relacionadas entre sí. La clasificación no jerárquica o de reagrupamiento tiene una estructura matemática menos precisa que la clasificación jerárquica.

### 2.3.4 Clusters jerárquicos: Dendograma

Partiendo de tantos grupos iniciales como individuos se estudian, se trata de conseguir agrupaciones sucesivas entre ellos de forma que progresivamente se vayan integrando en clusters los cuales, a su vez, se unirán entre sí en un nivel superior formando grupos mayores que más tarde se juntarán hasta llegar al cluster final que contiene todos los casos analizados. La representación gráfica de estas etapas de formación de grupos, a modo de árbol invertido, se denomina dendograma.



La figura que corresponde a un estudio de los individuos, muestra cómo el 8 y el 9 se agrupan en un primer cluster (A). En un nivel inmediatamente superior, se unen los individuos 1 y 2 (cluster B); y enseguida los 5, 6 y 7 (C). Un paso siguiente engloba el cluster B con el individuo 3 (D); y así sucesivamente hasta que todos ellos quedan estructurados al conseguir, en el nivel más alto, el cluster total (H) que reúne los 10 casos.

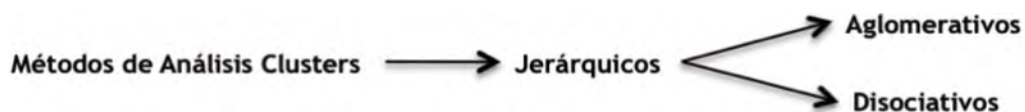
Evidentemente, la decisión de todas estas agrupaciones ha de tomarse en función de la similaridad multivariante (o de su contrario distancia) proporcionada por el conjunto de variables estudiadas, ya que en cada nivel de jerarquía se unen los dos clusters más cercanos. Es importante como paso previo a un análisis de clusters jerárquicos, la elección de una adecuada métrica de similaridad o disimilaridad. Se sabe que a partir de la tabla inicial de datos ( $n \times p$ ) es preciso calcular la matriz de distancias entre individuos ( $n \times n$ ). Se han descrito numerosas formas de medir distancias multivariantes. La conocida distancia euclídea es la más sencilla y utilizada: Se usa también en el análisis de componentes principales cuyos factores son, como se sabe, muchas veces datos previos para entrar en un análisis de clusters.

La segunda decisión que se debe tomar, es el algoritmo a emplear para la formación de grupos, definiendo a qué va a llamar “distancia entre clusters” para luego poder unir, a otro nivel jerárquico, aquellos clusters más próximos. Este concepto no existía en el análisis no jerárquico, puesto que allí no se unían los grupos. Es también muy variada y continuamente ampliada, la oferta de procedimientos de agrupación.

La aglomeración comienza con tantos grupos como individuos; cada uno de éstos constituye un cluster inicial. A medida que transcurren las etapas del proceso se van formando nuevos clusters por unión de dos individuos, de un individuo con un grupo previo, o de dos grupos anteriores entre los que exista la menor distancia.

El proceso finaliza con un único grupo (todos los individuos), pero constituido por aglomeraciones sucesivas en distintos niveles. Este es el fundamento de la agregación (ascendente); en contraposición con el proceso de disgregación (descendente), que opera de forma inversa: Parte del grupo total de individuos para llegar, tras varias etapas de partición, hasta tantos clusters como individuos. Característica importante de los métodos jerárquicos es el no permitir reasignaciones de grupos, es decir, que dos clusters (o dos individuos) que han sido unidos en un paso del proceso no pueden ya separarse en etapas sucesivas; lo que sí es posible en los métodos no jerárquicos (si bien en éstos es necesario fijar previamente el número de clusters deseado).

Se presenta a continuación una relación de los principales métodos de unión de grupos o algoritmos de clasificación jerárquica en la que, junto a su descripción, se comentan sus ventajas e inconvenientes como ayuda en la decisión del método apropiado.



Entre los *métodos aglomerativos* se tienen los siguientes:

**Método de las distancias mínimas o Enlace simple (single linkage).** Considera como distancia entre dos grupos la que responde al concepto de “vecinos más cercanos” (nearest neighbor), es decir, la separación que existe entre los individuos más próximos de uno y otro grupo. Aunque presenta buenas propiedades teóricas, su eficacia no ha sido la esperada en estudios de validación comparativa de métodos mediante ficheros simulados (Monte Carlo), por su tendencia al “encadenamiento” (un cluster realmente diferentes, si están próximos), y porque sólo considera la información de individuos extremos (los valores atípicos, outliers, pueden distorsionar la agrupación). Si bien no es adecuado para la obtención de grupos compactos, resulta de utilidad para clusters irregulares o alargados. El método consiste en ir agrupando los individuos que tienen menor distancia o mayor similitud, considerando como distancia entre dos clusters la distancia entre sus dos puntos más próximos.

**Método de las distancias máximas o Enlace completo (complete linkage).** Considera como distancia entre dos grupos la existente entre “vecinos más lejanos” (furthest neighbor), es decir, entre los individuos más separados de ambos grupos (máxima distancia que es posible encontrar entre un caso de un cluster y un caso de otro). Presenta una excesiva tendencia a producir grupos de igual diámetro, y se ve muy distorsionado ante valores atípicos moderados.

**Método del promedio entre grupos o Enlace promedio (average linkage).** Considera como distancia entre dos clusters, no la de los individuos más próximos ni más lejanos de ambos grupos, sino la distancia media entre todos los pares posibles de casos (uno de cada cluster). Tiende a producir clusters compactos, por lo que es utilizado y suele ser el método por defecto en los paquetes de software. Una variante de este método es el **método de la media ponderada (average linkage mitin groups)**, en el cual se combinan los grupos de tal forma que la distancia promedio entre todos los casos en el cluster resultante sea lo más pequeña posible.

**Método del centroide o Enlace centroide (centroid method).** Considera como distancia entre dos grupos la existente entre sus centros de gravedad, definidos por las medias aritméticas de las variables de los individuos que componen los clusters. Es el más robusto de los métodos jerárquicos ante la presencia de casos atípicos.

**Método de la mediana (median method).** Considera como distancia entre dos grupos la existente entre las medianas de las variables de los individuos que componen los clusters. De este modo, los dos clusters que se combinan se ponderan de forma equivalente al método centroide, pero independientemente del número de individuos que haya en cada grupo.

**Método de Ward o Enlace por mínima varianza (momento central de orden dos o pérdida de inercia mínima).** En este método se calcula la media de todas las variables de cada cluster, luego se calcula la distancia euclídea al cuadrado entre cada individuo y la media de su grupo y después se suman las distancias de todos los casos. En cada paso, los clusters que se forman son aquéllos que resultan con el menor incremento en la suma total de las distancias al cuadrado intracluster. Tiende a formar clusters esféricos o compactos, y del mismo tamaño.

La métrica normalmente considerada en los métodos hasta aquí descritos es la euclídea o la euclídea al cuadrado. Esta última se suele usar por omisión en programas estadísticos.

**Método de Enlace por densidad.** Utiliza un nuevo concepto de distancia entre dos individuos (A y B): Puesto que la nube de puntos no tendrá una densidad homogénea, existirán zonas de acúmulo de casos separadas por otras menos densas, lo que puede sugerir agrupamientos naturales entre los individuos. De esta forma, para unir dos de ellos, se les puede considerar como centros de dos esferas (multidimensionales) capaces de englobar un número prefijado de casos de la nube. El enlace por densidad es capaz de encontrar clusters irregulares y elongados, aunque ha demostrado menos eficacia en la configuración de clusters compactos.

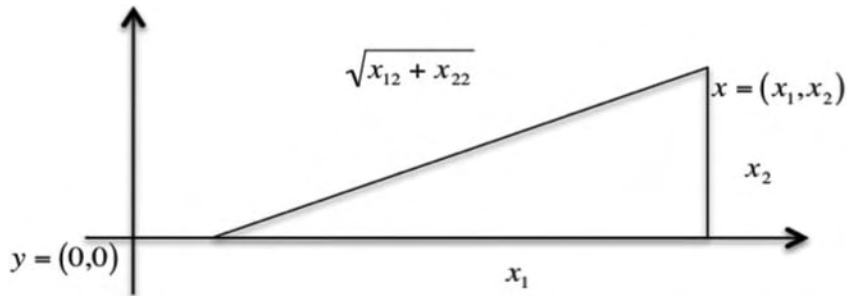
Los siete criterios anteriores para la clasificación jerárquica aglomerativa también son utilizables en el caso de los **métodos de clasificación disociativos**, si bien en la práctica los que más suelen usarse son el del promedio entre grupos y el de Ward. Además de ellos dentro del proceso disociativo destacan los que se presentan brevemente a continuación.

En los métodos disociativos, cuando el criterio de división toma en consideración cada variable observada una a una, el método recibirá el nombre de **monotético**. Por el contrario, cuando se toman en cuenta todas las variables simultáneamente el método se llamará **politético**.

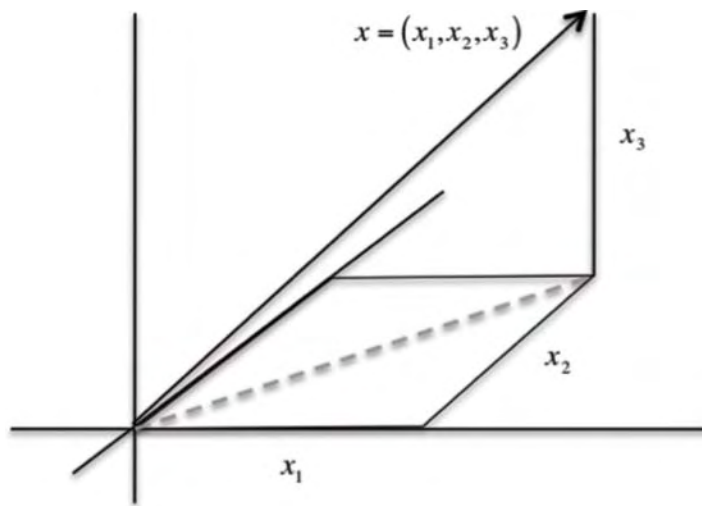
Otro método disociativo importante es el **detector automático de interacción**, conocido como **AID (Automatic Interaction Detector Method)**, que no nació como una técnica de clasificación y sin embargo, ha sido aceptada como tal. Los elementos básicos del AID son análogos a los de la regresión, con una variable dependiente y varias independientes. El objetivo de la AID consiste en







Cuando  $n = 3$  la Distancia Euclidiana tiene su fundamento en una generalización del Teorema de Pitágoras.



Al utilizar notación matricial y al tratar a los vectores de  $n$  componentes como matrices de  $n$  renglones y una columna podemos escribir:

$$d(x,y) = \sqrt{(x-y)'(x-y)} \text{ o equivalentemente, } d^2(x,y) = (x-y)'(x-y) \quad (2)$$

Donde  $(x-y)'$  es el transpuesto de  $(x-y)$ . Esta notación permitirá visualizar mejor las adaptaciones que se han hecho a la Distancia Euclidiana.

#### a. Supuestos de la Distancia Euclidiana

Con el fin de descubrir los supuestos inherentes a la Distancia Euclidiana se referirá al caso de dos dimensiones, esto es,  $n = 2$  y se considera un triángulo rectángulo. El triángulo rectángulo queda definido por los puntos  $(0,0)$ ,  $(x_1,0)$  y  $(x_1,x_2)$ . Los catetos son  $x_1$  y  $x_2$  y la hipotenusa es la raíz

cuadrada de  $x_1^2 + x_2^2$ . Como los catetos son longitudes, tanto  $x_1$  como  $x_2$  deben ser números reales positivos medidos en ciertas unidades de longitud (centímetros, decímetros, metros, etc.). La restricción de que  $x_1$  y  $x_2$  sean positivos no es necesaria en general para aplicar la fórmula (1).

El Teorema de Pitágoras supone que tanto  $x_1$  como  $x_2$  están medidos en las mismas unidades. Si son  $x_1$  y  $x_2$  están medidos en unidades distintas, el Teorema de Pitágoras produce una respuesta incorrecta. Por ejemplo, si son  $x_1 = 3$  metros y  $x_2 = 4$  pies y utilizamos el Teorema de Pitágoras para calcular la hipotenusa, es claro que la hipotenusa no mide ni 5 metros ni 5 pies. En cambio si  $x_1 = 3$  metros y  $x_2 = 4$  metros, podemos verificar, al medir la hipotenusa con una cinta métrica, que mide 5 metros.

Al considerar un triángulo rectángulo y graficarlo en un sistema de dos ejes, se está haciendo el supuesto de que los ejes son perpendiculares (suponiendo un sistema cartesiano). En un sistema donde los ejes no son perpendiculares el Teorema de Pitágoras no aplica.

Existen otros supuestos que pueden motivarse a partir de las aplicaciones estadísticas, las observaciones hechas aplican a cualquier número finito de dimensiones.

### b. La Distancia en aplicaciones estadísticas

En aplicaciones de a Estadística Multivariada, se caracterizan a personas, marcas, empresas, etc., por medio de vectores de mediciones. Así, una persona puede estar caracterizada por: estatura ( $= x_1$ ), edad ( $= x_2$ ), sexo ( $= x_3$ ), nivel socioeconómico ( $= x_4$ ), peso ( $= x_5$ ), etc. De tal manera que la persona  $k$  en una muestra es representada por un vector  $x_k$ . Las variables pueden ser medidas en escalas diferentes (metros, años, escalas nominales, etc.).

Diversas técnicas estadísticas, como el Análisis de Conglomerados (Cluster Analysis), Análisis de Componentes Principales (ACP) y el Análisis de Factores (Factor Analysis), están basadas en la distancia que existe entre personas o la distancia entre personas a una recta en el espacio.

En la siguiente tabla se muestra el peso y la estatura de 17 personas. Se caracteriza a una persona por medio de un vector de dos componentes ( $x_1 = peso, x_2 = estatura$ ) donde  $x_1$  está medida en kilogramos y  $x_2$  en metros.

Persona	$x_1 = \text{Peso}$	$x_2 = \text{Estatura}$
1	45	1.52
2	50	1.57
3	54	1.63
4	58	1.60
5	60	1.75
6	62	1.75
7	62	1.53
8	63	1.65
9	65	1.67
10	67	1.74
11	70	1.79
12	72	1.69
13	72	1.65
14	74	1.73
15	76	1.72
16	82	1.87
17	85	1.85
Media	65.7	1.689
Varianza	115.7	0.010
Desv. Est.	10.8	0.101
Correlación	0.794	

Si calculamos el cuadrado de la Distancia Euclidiana entre las personas 5 y 9 tenemos que:

$$d^2(x_5, x_9) = (60 - 65)^2 + (1.75 - 1.67)^2 = 25 + 0.0064 = 25.0064$$

Se observa que la diferencia en el peso es la determinante en el cálculo de la distancia. Lo mismo sucede con otros pares de personas; el peso es el factor crucial mientras que la estatura casi no afecta. Pero este hecho es derivado de la escala de medición; una unidad de medición en el peso es un kilo mientras que una unidad de medición en la estatura es un metro. Para que una persona  $(x_1, x_2)$  del mismo peso que la persona 5 ( $x_1 = 60$ ) esté a la misma distancia que la persona 9 (5) se necesita que la persona tuviera una estatura de 6.75 metros aproximadamente, que se obtiene al

resolver  $25.0064 = (60 - 60)^2 + (x_2 - 1.75)^2$ . Notemos que la variable peso tiene una varianza mayor que la variable estatura. La variable de mayor varianza dominará en el cálculo de las distancias<sup>2</sup>. Si se expresa la estatura en centímetros (si se hace un cambio de escala) entonces:

$$d^2(x_5, x_9) = (60 - 65)^2 + (175 - 167)^2 = 25 + 64 = 89$$

Y resulta que la estatura, para este par de personas, es la determinante en el cálculo de la distancia. Observar lo que produjo el cambio de escala; una unidad en la estatura es ahora un centímetro; antes una unidad era un metro.

En consecuencia, la escala de medición tiene impacto directo sobre la distancia. De esta manera, al cambiar la escala de una variable se puede hacer que pese más o menos en el cálculo de la distancia; al utilizar una escala más fina se incrementa su peso y al utilizar una escala más gruesa se resta peso.

Al expresar una variable en una escala más fina se hace que la desviación estándar (o equivalentemente la varianza) sea mayor. En consecuencia, el problema de las escalas distintas está relacionado con la situación en la que las distintas variables tienen varianzas distintas. Para resolver el problema de las distintas escalas de medición se han hecho varias propuestas que pueden resumirse como: escalar los datos respecto a una estadística (usando índices) y luego aplicar la distancia Euclidiana.

Por ejemplo, si la estadística es la media aritmética se transformaría los datos a índices respecto de a las distintas medias:

$$x = (x_1, x_2, \dots, x_n)' = x \left( \begin{matrix} x_1/m_1 & x_2/m_2 & \dots & x_n/m_n \end{matrix} \right)'$$

Donde las  $m$ 's son las medias aritméticas. El cuadrado de la distancia entre dos vectores  $x$  y  $y$  estaría dado por la fórmula:

.....

\* Suponer que  $X$  es un vector aleatorio de  $n$  componentes que tiene media  $\mu (= E(X))$  y matriz de varianzas y covarianzas  $\Sigma (= E(X - \mu)(X - \mu)')$ . Al calcular el promedio del cuadrado de la distancia de  $X$  a  $\mu$  se tiene que:

$$E[d^2(X - \mu)] = E[(X - \mu)'(X - \mu)] = E[tr(X - \mu)(X - \mu)'] = tr(\Sigma) = \sum_{k=1}^n Var(X_k)$$

Donde  $tr(\bullet)$  denota la traza de una matriz. La fórmula permite ver que la variable con la mayor varianza será la que domine la suma (que es el cuadrado de la distancia promedio a la media).

$$\frac{(x_1 - y_1)^2}{m_1^2} + \frac{(x_2 - y_2)^2}{m_2^2} + \dots + \frac{(x_n - y_n)^2}{m_n^2}$$

Se observa que se pueden utilizar índices respecto a mínimos, máximos, medianas, etc. Existe una propuesta que sugiere escalar los datos por las desviaciones estándar. Esta propuesta logra dos objetivos; por una parte elimina la escala y por otro logra que todas las variables tengan la misma varianza (igual a uno). Así, el cuadrado de la distancia entre dos vectores  $x$  y  $y$  estaría dado por la fórmula:

$$\frac{(x_1 - y_1)^2}{S_1^2} + \frac{(x_2 - y_2)^2}{S_2^2} + \dots + \frac{(x_n - y_n)^2}{S_n^2}$$

Donde las  $S$ 's son las desviaciones estándar. Al escalar por las desviaciones estándar las variables quedan expresadas en unidades que no necesariamente son fáciles de interpretar. Los escalamientos planteados llevan a la siguiente variación en la fórmula (1):

$$d_1(x,y) = \sqrt{w_1(x_1, y_1)^2 + w_2(x_2, y_2)^2 + \dots + w_n(x_n, y_n)^2}$$

Donde  $w_1, w_2, \dots, w_n$  son los factores de escalamiento.

Por otra parte las  $w$ 's pueden ser interpretadas como pesos o importancias de cada variable o dimensión y son número reales no negativo. Mientras una  $w$  sea mayor, mayor será la influencia de esa variable en el cálculo de la distancia. Si todos los pesos son iguales, esto es,  $w_1 = w_2 = \dots = w_n$ , se estará aplicando esencialmente la Distancia Euclidiana (en consecuencia, la Distancia Euclidiana da el mismo peso a todas las variables).

Si se reconsidera el ejemplo presentado (peso en kilos y estatura en centímetros), pero se asignan pesos distintos a las variables, por ejemplo:  $w_1 = 4$  y  $w_2 = 1$ , se tendrá:

$$d_1^2(x_5, x_9) = 4(60 - 65)^2 + (175 - 167)^2 = 100 + 64 = 164$$

Y el peso vuelve a ser el factor determinante en el cálculo de la distancia. Surge la pregunta de bajo qué criterio de optimalidad se puede seleccionar a las  $w$ 's, si definimos la matriz  $W$  como aquella matriz que tiene en la diagonal principal a las  $w$ 's y fuera de la diagonal ceros, esto es,

$W = \text{diag}(w_1, w_2, \dots, w_n)$ , se puede escribir:  $d_1^2(x,y) = (x,y)'W(x,y)$ . Es posible demostrar que  $d_1$  es una distancia y satisface las tres condiciones:

- $d(x,y) = d(y,x)$
- $d(x,y) \geq 0$  y es igual a cero si y sólo si  $x = y$
- $d(x,y) \leq d(x,z) + d(z,y)$  par

uno de los análisis multivariados que permite motivar una nueva modificación a la fórmula de la Distancia Euclidiana es la detección de observaciones aberrantes (outliers). Una observación aberrante es un vector de datos que es muy diferente del resto. Estas observaciones pueden distorsionar seriamente el análisis y deben ser identificadas lo antes posible. La decisión de incluirlas o excluirlas del análisis depende de cada caso particular.

Bajo la Distancia Euclidiana los puntos equidistantes del vector de medias son una hiperesfera (generalización de círculo y esfera) reflejando el hecho de iguales escalas, varianzas e importancias. Bajo la variación presentada de la Distancia Euclidiana, los puntos equidistantes del vector de medias son un hiperelipsoide paralelo a los ejes (generalización de elipse y elipsoide) reflejando la existencia de escalas, varianzas e importancias distintas. Ninguna de las dos distancias toma en cuenta la correlación existente entre las variables.

La Distancia de Mahalanobis toma en cuenta la correlación existente entre las variables estudiadas para determinar la distancia. Si la matriz de varianzas y covarianzas de las variables estudiadas es S, el cuadrado de la distancia de Mahalanobis entre los vectores  $x$  y  $y$  está definida como\_

$$d_M^2(x,y) = (x,y)' S^{-1}(x,y)$$

Donde  $S^{-1}$  es la matriz inversa de S. Con el fin de ver cómo interviene la correlación entre variables se revisará el caso particular cuando  $n = 2$ . Cuando dos variables intervienen en el análisis, la matriz S tiene en la diagonal las varianzas de las variables y fuera de la diagonal tiene la covarianza, de manera explícita:

$$d_M^2(x,y) = \left[ \frac{(x_1 - y_1)^2}{S_1^2} + \frac{(x_2 - y_2)^2}{S_2^2} - \frac{2r(x_1 - y_1)(x_2 - y_2)}{(S_1 - S_2)} \right] \div (1 - r^2)$$

Donde r es el coeficiente de correlación entre  $x_1$  y  $y_1$ , definido como la covarianza dividida entre el producto de las desviaciones estándar, esto es:

$$r = \frac{S_{12}}{(S_1 S_2)}$$

Se observa primero que si  $r = 0$ , los dos primeros términos corresponden a la Distancia Euclidiana con las variables escaladas por la desviación estándar, pero afectadas por el término  $\frac{1}{(1 - r^2)}$ . El tercer término es la novedad.

## CAPÍTULO 3

### METODOLOGÍA Y DESARROLLO

#### 3.1 NIVEL Y TIPO DE INVESTIGACIÓN

##### 3.1.1 Tipo de investigación

El tipo de investigación es cuantitativa, concluyente, la cual se recomienda cuando se requiere tomar decisiones con el respaldo de información con validez estadística; obteniendo resultados con base en una muestra, la cual refleja la opinión de una muestra representativa de la población.

Este tipo de investigación, contempla desde el diseño de un estudio hasta la entrega de un análisis y recomendaciones sobre los pasos a seguir desde un punto de vista de marketing. Proporciona resultados estadísticos sobre el comportamiento y las preferencias del consumidor. Descubre el quién, el qué, el cuándo, el dónde y el cómo de las preferencias, las tendencias y los hábitos de los consumidores.

##### 3.1.2 Nivel de la investigación

La investigación exploratoria es la etapa inicial o preliminar del proceso de investigación. En la investigación exploratoria, la información se recolecta de fuentes primarias y secundarias con el fin de suministrar información sobre el problema e identificar cursos de acción.

Se inició con una investigación exploratoria, la cuál se hizo de manera directa mediante de una entrevista cara a cara, logrando de esta manera información valiosa para poder usarla posteriormente al elaborar el cuestionario.

#### 3.2 DISEÑO DE LA INVESTIGACIÓN

El diseño de investigación es el plan básico que guía las fases de recolección de datos del proyecto de investigación. Especifica el tipo de información que se recolectará, las fuentes de datos y el procedimiento de datos.

El diseño de investigación será de tipo no experimental - transversal ya que se realizará una sola vez dentro de un tiempo determinado. Así mismo, se desarrolló un cuestionario estructurado que sirvió como instrumento para la recolección de datos mediante una entrevista cara a cara, y con esto es una investigación probabilística ya que todos los que fueron entrevistados tenían el mismo porcentaje de libertad para contestarlo

#### 3.3 MÉTODO DE MUESTREO

El muestreo es una herramienta de la investigación científica. Su función básica es determinar que parte de una realidad en estudio (población o universo) debe examinarse con la finalidad de hacer inferencias sobre dicha población. El error que se comete debido al hecho de que se obtienen conclusiones sobre cierta realidad a partir de la observación de sólo una parte de ella, se denomina error de muestreo. Obtener una muestra adecuada significa lograr una versión simplificada de la población, que reproduzca de algún modo sus rasgos básicos.

##### Terminología

- **Población objeto.** Conjunto de individuos de los que se quiere obtener información

- **Unidades de muestreo.** Número de elementos de la población, no solapados, que se van a estudiar. Todo miembro de la población pertenecerá a una y sólo una unidad de muestreo
- **Unidades de análisis.** Objeto o individuo del que hay que obtener información
- **Marco muestral.** Lista de unidades o elementos de muestreo
- **Muestra.** Conjunto de unidades o elementos de análisis sacados del marco

### 3.3.1 Muestreo probabilístico

El método otorga una probabilidad conocida de integrar la muestra a cada elemento de la población, y dicha probabilidad no es nula para ningún elemento.

Los métodos de muestreo no probabilísticos no garantizan la representatividad de la muestra y por lo tanto no permiten realizar estimaciones inferenciales sobre la población.

*(En algunas circunstancias los métodos estadísticos y epidemiológicos permiten resolver los problemas de representatividad aún en situaciones de muestreo no probabilístico, por ejemplo los estudios de caso-control, donde los casos no son seleccionados aleatoriamente de la población.)*

Entre los métodos de muestreo probabilísticos más utilizados en investigación se encuentran:

- Muestreo aleatorio simple
- Muestreo estratificado
- Muestreo sistemático
- Muestreo polietápico o por conglomerados

Para esta investigación se utiliza un muestreo probabilístico polietápico o por conglomerados, en el cual se realizan varias fases de muestreo sucesivas (polietápico). La necesidad de listados de las unidades de una etapa se limita a aquellas unidades de muestreo seleccionadas en la etapa anterior.

Entre las ventajas de este muestreo se encuentran:

- Es muy eficiente cuando la población es muy grande y dispersa
- No es preciso tener un listado de toda la población, sólo de las unidades primarias de muestreo

## 3.4 CÁLCULO DEL TAMAÑO MUESTRAL

Cada estudio tiene un tamaño muestral idóneo que permite comprobar lo que se pretende con la seguridad y precisión fijadas.

**¿De qué depende el tamaño muestral?**

- **Variabilidad del parámetro a estimar.** Datos previos, estudios piloto o usar 50% como peor estimación
- **Precisión.** Amplitud del intervalo de confianza. Si se estima prevalencia su formato será %
- **Nivel de confianza**  $(1 - \alpha)$ . Habitualmente 95% o 99%. Probabilidad complementaria al error admitido  $\alpha$

Si se aumenta el tamaño muestral  $n$ , se podrá mejorar la calidad de la estimación bien aumentando la precisión (disminuye amplitud del intervalo) o bien aumentando la seguridad (disminuye el error admitido).



Se puede aplicar las siguientes fórmulas para el cálculo del tamaño muestral (si el muestreo es aleatorio). Si se conoce el tamaño de la población se usa el método para poblaciones finitas. Si por el contrario el tamaño de la población es desconocido o infinito, se usa la otra alternativa.

Hay que tener en cuenta que una población infinita puede corresponder a una finita (conocida) en la que se ha definido un muestro con reemplazamiento (el mismo individuo puede salir muestreado varias veces).

Tamaño de la población infinito o desconocido:  $n = Z_{\alpha}^2 \frac{p \cdot q}{i^2}$

Tamaño de la población finito:  $n = Z_{\alpha}^2 \frac{N \cdot p \cdot q}{i^2(N-1) + Z_{\alpha}^2 \cdot p \cdot q}$

- $n$  Tamaño muestral
- $N$  Tamaño de la población, número total de historias
- $Z$  Valor correspondiente a la distribución de Gauss 1.96 para  $\alpha = 0.05$  y 2.58 para 1  $\alpha = 0.01$
- $p$  Prevalencia esperada del parámetro a evaluar. En caso de desconocerse, aplicar la opción más desfavorable ( $p = 0.5$ ) que hace mayor el tamaño muestral
- $q$   $1 - p$  (Si  $p = 30\%$ ,  $q = 70\%$ )
- $i$  Error que se prevé cometer. Por ejemplo, para un error del 10% se introduce en la fórmula el valor 0.1. Así, con un error de 10%, si el parámetro estimado resulta del 80% se tendría una seguridad del 95% (para  $\alpha = 0.05$ ) de que el parámetro real se sitúa entre el 70% y el 90%. La amplitud total del intervalo es el doble del error que se introduce en la fórmula

Utilizando los valores estándar indicados se efectúa el cálculo siguiente:

$$n = Z_{\alpha}^2 \frac{p \cdot q}{i^2} = \frac{(1.96)^2 \times (0.5(1-0.5))}{(0.05)^2} = \frac{3.8416 \times 0.25}{0.0025} = \frac{0.9604}{0.0025}$$

$$n = 384.16 \approx 384$$

Para esta investigación se utiliza un muestreo probabilístico polietápico o por conglomerados, en el cual se realizan varias fases de muestreo sucesivas (6 en el año). A fin de corregir la diferencia en el diseño, el tamaño de la muestra se multiplica por el efecto de diseño (D). Por lo general se presupone un efecto de diseño igual a 2 para las encuestas que utilizan una metodología de muestreo por conglomerados.

$$n \times D = 384 \times 2 = 768$$

El tamaño de la muestra se aumenta en un 5% para hacer frente a imprevistos como ausencia de respuesta o errores de registro.

$$n + 5\% = 768 \times 1.05 = 806.4 \approx 806$$

Se divide el tamaño de la muestra (N) por el número de conglomerados (6) a fin de determinar el número de observaciones por conglomerado.

$$N \div \text{No.conglomerados} = 806 \div 6 = 1.34.4 \approx 134$$

Se tomará un total de 100 entrevistas mensuales (200 bimestrales), las cuales para resultados totales, tienen un error de muestreo de 5%, trabajando con una confiabilidad del 95%.

***La muestra se distribuirá***

<b>NSE</b>	<b>C</b>	<b>C+</b>
TOTALES	100	100
MUJERES	50	50
HOMBRES	50	50

NSE Nivel Socioeconómico

C+ En este segmento se consideran a las personas con ingresos o nivel de vida ligeramente superior al medio

C En este segmento se considera a las personas con ingresos o nivel de vida medio

Se hizo un sorteo aleatorio de acuerdo a la proporción de niveles socioeconómicos para garantizar resultados válidos en todos los estratos a investigar. No se realizaron más de 35 entrevistas por punto muestral.

### **3.5 TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS**

Para la investigación se utilizaron las siguientes técnicas:

**La encuesta.** Técnica que permitió elaborar en función del problema planteado, la hipótesis y las variables identificadas, la información necesaria para validar el presente trabajo de investigación.

#### **ANEXO 1**

### **3.6 ANÁLISIS DE LA INFORMACIÓN**

Para el análisis se ocupó SPSS (Statistical Product for Service Solutions) y Statistica. Se levantaron 200 entrevistas efectivas. El levantamiento de campo se realizó del 1ro. de julio al 31 de agosto de 2006 Cuarto bimestre 2006

#### **3.6.1 Objetivo de la investigación**

Tener un histórico de la recordación, preferencia y actitudes de marca. Midiendo también recordación publicitaria, indicadores de preferencia de compra de nuestras marcas y de la competencia.

#### **3.6.2 Población objetivo estudiada**

Considerando como público objetivo a aquellas personas que consumen y/o compran leches.

Sexo: Hombres y Mujeres  
Edad: Mayores de 18 años  
NSE: C+, C  
Ubicación: Residentes de la Ciudad de México

### 3.6.3 Variables sociodemográficas

Para iniciar el análisis se hace una recodificación de las edades en 4 grupos:

- Jóvenes. De 18 a 28 años
- Adultos jóvenes. De 29 a 39 años
- Adultos maduros. De 40 a 50 años
- Adultos mayores. De 51 a 70 años

#### 3.6.3.1 Total de la muestra por sexo

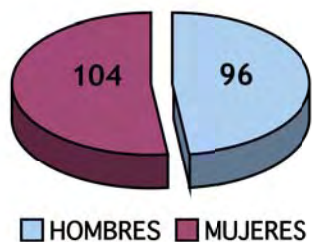
		SEXO			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	MUJER	104	52.0	52.0	52.0
	HOMBRE	96	48.0	48.0	100.0
	Total	200	100.0	100.0	

#### 3.6.3.2 Total de la muestra por sexo y nivel socioeconómico

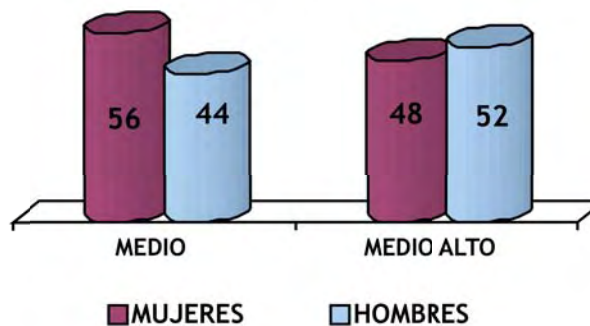
**SEXO \* NIVEL SOCIOECONOMICO**  
Crosstabulation  
Count

		NIVEL SOCIOECONOMICO		Total
		MEDIO	MEDIO ALTO	
SEXO	MUJER	56	48	104
	HOMBRE	44	52	96
Total		100	100	200

TOTAL DE LA MUESTRA  
POR SEXO



TOTAL DE LA MUESTRA POR NIVELES  
SOCIOECONÓMICOS



### 3.6.3.3 Distribución de estado civil en el total de la muestra por grupos de edad en mujeres

**ESTADO CIVIL \* GRUPOS DE EDADES Crosstabulation**  
Count

MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
ESTADO CIVIL	CASADO	4	11	21	6	42
	SOLTERO	20	13	9	3	45
	UNION LIBRE	4	3	4	0	11
	OTRO	0	0	2	4	6
Total		28	27	36	13	104

### 3.6.3.4 Distribución de estado civil en el total de la muestra por grupos de edad en hombres

**ESTADO CIVIL \* GRUPOS DE EDADES**  
Crosstabulation  
Count

HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
ESTADO CIVIL	CASADO	6	7	29	12	54
	SOLTERO	9	16	3	2	30
	UNION LIBRE	2	4	3	0	9
	OTRO	1	1	0	1	3
Total		18	28	35	15	96

#### Symmetric Measures

		Value	Asymp. Std. Error(a)	Approx. T(b)	Approx. Sig.
Interval by Interval	Pearson's R	-.341	.103	-3.521	.001(c)
Ordinal by Ordinal	Spearman Correlation	-.430	.092	-4.619	.000(c)
N of Valid Cases		96			

- a Not assuming the null hypothesis.
- b Using the asymptotic standard error assuming the null hypothesis.
- c Based on normal approximation.

#### Correlations

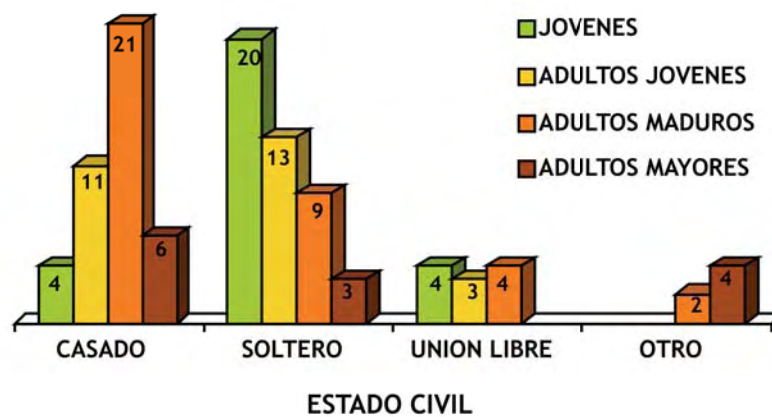
		ESTADO CIVIL	GRUPOS DE EDADES
ESTADO CIVIL	Pearson Correlation	1	-.341(**)
	Sig. (2-tailed)	.	.001
	N	96	96
GRUPOS DE EDADES	Pearson Correlation	-.341(**)	1
	Sig. (2-tailed)	.001	.
	N	96	96

\*\* Correlation is significant at the 0.01 level (2-tailed).

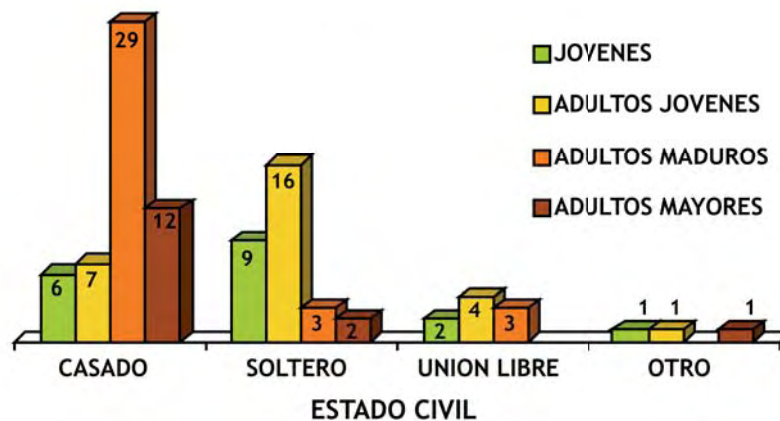
Si hay correlación inversamente proporcional aunque es baja ( $r = -0.341$  con un nivel de significancia de  $p = .001$  a dos colas con una  $\alpha = 0.01$ ) con respecto a la edad recodificada y el estado civil para el caso de los hombres. Esto es a medida que los hombres entran en el siguiente grupo de edad, van dejando su condición de casados para entrar en otras categorías de estado civil como unión libre.

### DISTRIBUCIÓN DE ESTADO CIVIL EN EL TOTAL DE LA MUESTRA POR GRUPOS DE EDAD

#### MUJERES



#### HOMBRES

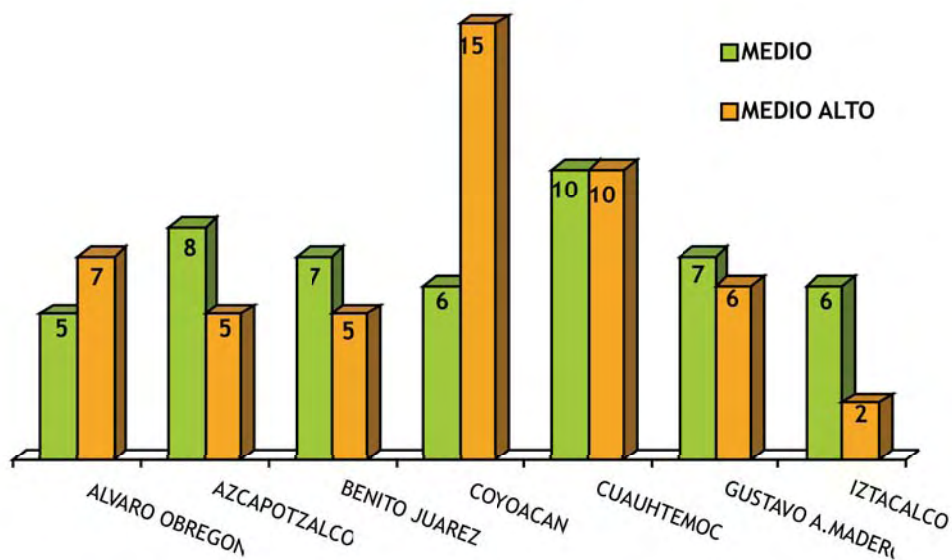


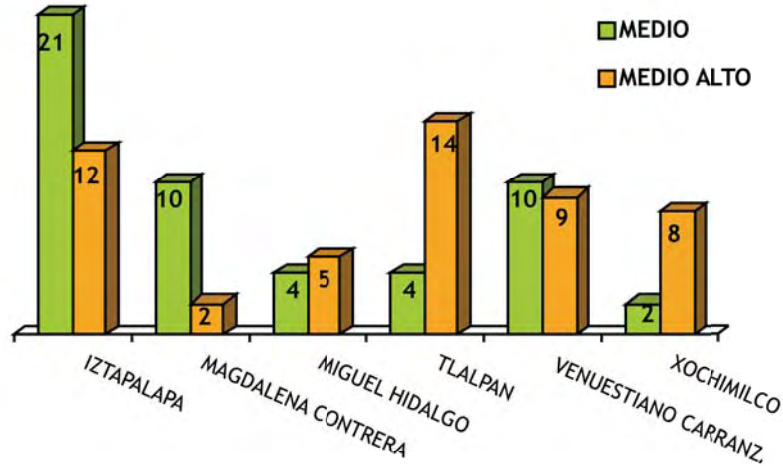
### 3.6.3.5 Total de la muestra por niveles socioeconómicos y delegaciones

**DELEGACION \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**  
 Count

		NIVEL SOCIOECONOMICO		Total
		MEDIO	MEDIO ALTO	
DELEGACION	ALVARO OBREGON	5	7	12
	AZCAPOTZALCO	8	5	13
	BENITO JUAREZ	7	5	12
	COYOACAN	6	15	21
	CUAUHTEMOC	10	10	20
	GUSTAVO A MADERO	7	6	13
	IZTACALCO	6	2	8
	IZTAPALAPA	21	12	33
	MAGDALENA CONTRERAS	10	2	12
	MIGUEL HIDALGO	4	5	9
	TLALPAN	4	14	18
	VENUSTIANO CARRANZA	10	9	19
	XOCHIMILCO	2	8	10
	Total	100	100	200

**TOTAL DE LA MUESTRA**  
**POR NIVELES SOCIOECONÓMICOS Y DELEGACIONES**





### 3.6.4 Ruido publicitario

#### 3.6.4.1 Top of mind de producto en mujeres

**Case Processing Summary**

MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
1RA PUBLICIDAD DE PRODUCTOS EN TODOS LOS MEDIOS * GRUPOS DE EDADES	97	100.0%	0	.0%	97	100.0%

**1RA PUBLICIDAD DE PRODUCTOS EN TODOS LOS MEDIOS \* GRUPOS DE EDADES**  
Crosstabulation  
Count

TOM DE PRODUCTOS MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
1RA PUBLICIDAD DE PRODUCTOS EN TODOS LOS MEDIOS	REFRESCOS	9	5	6	2	22
	TELEFONIA / CELULARES	0	0	2	0	2
	CERVEZA	1	1	0	1	3
	POLITICA	1	0	1	0	2
	CIGARROS	1	0	0	0	1
	SHAMPOO	1	2	3	1	7
	LECHES	1	2	3	2	8
	TIENDAS DEPARTAMENTALES	1	2	4	1	8
SUPERMERCADOS	2	2	2	0	6	

ROPA	0	1	0	0	1
CEREAL	1	0	2	1	4
AUTOS	0	0	1	0	1
PRODUCTOS PANIFICADOS/GALLLETAS	0	1	2	0	3
ELECTRONICOS	1	0	0	0	1
PRODUCTOS LACTEOS	0	4	0	2	6
AGUA EMBOTELLADA	1	0	1	0	2
BELLEZA / COSMETICOS	1	2	1	1	5
FRITURAS / PAPAS	2	1	0	0	3
JOYERIA / RELOJES	0	0	1	0	1
ZAPATOS / TENIS	1	1	0	0	2
DENTRIFICOS	1	0	1	0	2
PAÑALES	0	1	0	0	1
TOALLAS SANITARIAS	0	0	2	1	3
RESTAURANTES COMIDA RAPIDA	0	1	1	0	2
PAPEL SANITARIO	1	0	0	0	1
<b>Total</b>	<b>26</b>	<b>26</b>	<b>33</b>	<b>12</b>	<b>97</b>

Para las mujeres el Top of Mind son los refrescos (22 veces) y de estos la mayoría ocurre en el grupo de Jóvenes (9), seguido por leches (8) y tiendas departamentales (8) y en tercer lugar shampoo (7), en el caso de leches (3) ocurrieron en el grupo de adultos maduros caso similar al de las tiendas con (4) para este grupo.

### 3.6.4.2 Top of mind de producto en hombres

#### Case Processing Summary

HOMBRES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
1RA PUBLICIDAD DE PRODUCTOS EN TODOS LOS MEDIOS * GRUPOS DE EDADES	84	100.0%	0	.0%	84	100.0%

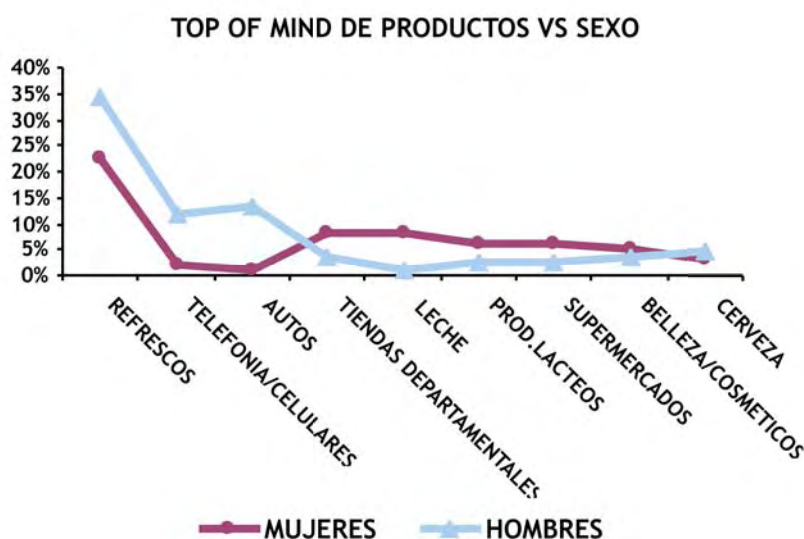


**1RA PUBLICIDAD DE PRODUCTOS EN TODOS LOS MEDIOS \* GRUPOS DE EDADES**  
**Crosstabulation**  
**Count**

TOM DE PRODUCTOS HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
1RA PUBLICIDAD DE PRODUCTOS EN TODOS LOS MEDIOS	REFRESCOS	5	7	16	1	29
	TELEFONIA / CELULARES	1	2	4	3	10
	CERVEZA	0	0	4	0	4
	POLITICA	0	1	0	1	2
	CIGARROS	2	0	0	1	3
	LECHES	0	1	0	0	1
	TIENDAS DEPARTAMENTALES	0	3	0	0	3
	SUPERMERCADOS	0	2	0	0	2
	ROPA	1	0	0	1	2
	CEREAL	0	1	0	0	1
	BANCOS	0	0	0	1	1
	AUTOS	2	3	5	1	11
	PRODUCTOS LACTEOS	1	1	0	0	2
	AGUA EMBOTELLADA	1	1	0	0	2
	BELLEZA / COSMETICOS	0	1	1	1	3
	FRITURAS / PAPAS	0	0	0	1	1
	DENTRIFICOS	2	0	0	0	2
	AEROLINEAS	0	0	1	0	1
	DESODORANTES	2	0	1	0	3
RESTAURANTES COMIDA RAPIDA	0	0	1	0	1	
<b>Total</b>	<b>17</b>	<b>23</b>	<b>33</b>	<b>11</b>	<b>84</b>	

Los productos mas frecuentes son refrescos (29), autos (11), telefonía / celulares (10) y cerveza (4). De éstos el refresco tiene su mayor frecuencia en los adultos maduros (16), seguido por adultos jóvenes (7), para el caso de autos sigue el mismo comportamiento con 5 y 3 veces; para el caso de telefonía / celulares (4) adultos maduros y 2 adultos jóvenes, las leches sólo se presentan una vez en el grupo de adultos jóvenes.

Son ellas quienes tienen más "ruido" de productos en la cabeza en comparación con los hombres.



Es cierto que hay Top of Mind en el que coinciden ambos sexos (caso de los refrescos 22 mujeres y 29 para hombres), pero algunos de los productos solo se presentan en un sexo y en el otro no, caso autos 11 menciones en hombres vs. 1 mención en mujeres o por ejemplo en el caso telefonía / celulares con 10 para hombres y 2 para mujeres. Tal vez esto puede estar ligado a la parte laboral en cada sexo.

### 3.6.4.3 Top of mind de marca en mujeres

#### Case Processing Summary

MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
1RA PUBLICIDAD DE MARCA EN TODOS LOS MEDIOS * GRUPOS DE EDADES	97	100.0%	0	.0%	97	100.0%

#### 1RA PUBLICIDAD DE MARCA EN TODOS LOS MEDIOS \* GRUPOS DE EDADES

##### Crosstabulation

Count

TOM MARCA MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
1RA PUBLICIDAD DE MARCA EN TODOS LOS MEDIOS	COCA COLA	8	5	6	2	21
	PEPSI COLA	1	0	0	0	1
	TELCEL	0	0	2	0	2
	CORONA	1	0	0	1	2
	PRD	1	0	0	0	1

MARLBORO	0	1	0	0	1
CAMEL	1	0	0	0	1
SEDAL	1	2	3	1	7
ALPURA	0	3	1	4	8
LALA	1	0	2	0	3
LIVERPOOL	0	1	2	1	4
LOREAL	1	1	0	0	2
MARTI	0	1	0	0	1
IFE	0	0	1	0	1
KELLOGS	1	0	2	1	4
VW	0	0	1	0	1
SONY	1	0	0	0	1
DANONE	0	2	0	0	2
BIMBO	0	0	2	0	2
PALACIO DE HIERRO	0	1	2	0	3
WALMART	2	0	1	0	3
COMERCIAL MEXICANA	0	2	1	0	3
BONAFONT	1	0	1	0	2
OREO	0	1	0	0	1
SABRITAS	2	1	0	0	3
CITIZEN	0	0	1	0	1
CHANNEL	0	0	1	0	1
ANDREA	1	1	0	0	2
COLGATE	1	0	1	0	2
HUGGIES	0	1	0	0	1
HINDS	0	0	0	1	1
PRIMAVERA	0	1	0	0	1
OFFICE DEPOT	1	0	0	0	1
SABA	0	0	1	0	1
KOTEX	0	0	1	1	2
BURGER KING	0	1	1	0	2
REVLON	0	1	0	0	1
PETALO	1	0	0	0	1
Total	26	26	33	12	97

La Top of Mind de Marca es Coca Cola (21 de 97) de los cuales 8 son del grupo de jóvenes , segundo lugar Alpura (8) en el grupo de adultos jóvenes (3) y adultos mayores (4) y en tercer lugar Sedal con (7), la competencia Lala (3).

### 3.6.4.4 Top of mind de marca en hombres

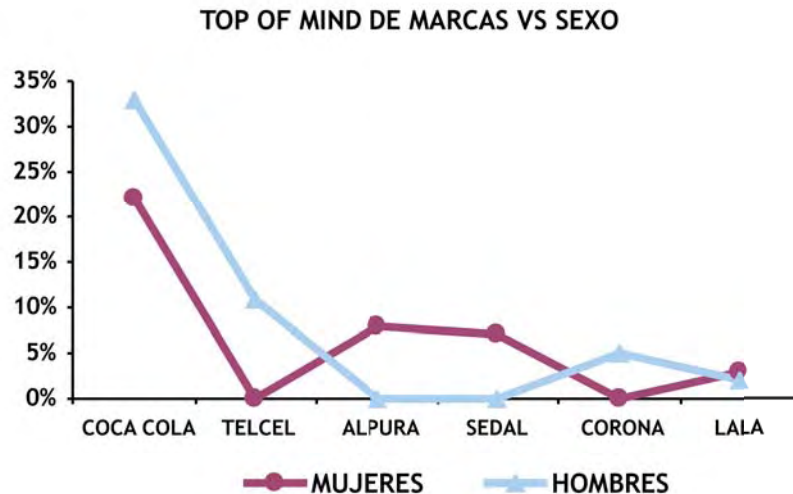
#### Case Processing Summary

HOMBRES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
1RA PUBLICIDAD DE MARCA EN TODOS LOS MEDIOS * GRUPOS DE EDADES	84	100.0%	0	.0%	84	100.0%

#### 1RA PUBLICIDAD DE MARCA EN TODOS LOS MEDIOS \* GRUPOS DE EDADES Crosstabulation Count

TOM MARCA HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
1RA PUBLICIDAD DE MARCA EN TODOS LOS MEDIOS	COCA COLA	5	7	15	1	28
	TELCEL	1	2	4	2	9
	CORONA	0	0	4	0	4
	PRD	0	1	0	0	1
	PAN	0	0	0	1	1
	MARLBORO	2	0	0	1	3
	LALA	1	1	0	0	2
	LIVERPOOL	0	2	0	0	2
	GIGANTE	0	1	0	0	1
	KELLOGS	0	1	0	0	1
	BANCOMER	0	0	0	1	1
	FORD	1	0	2	0	3
	VW	1	3	2	1	7
	DANONE	0	1	0	0	1
	CHRYSLER	0	0	1	0	1
	WALMART	0	1	0	0	1
	BONAFONT	1	1	0	0	2
	SABRITAS	0	0	0	1	1
	CHANNEL	0	0	0	1	1
	PEÑAFIEL	0	0	1	0	1
	COLGATE	2	0	0	0	2
	AEROMEXICO	0	0	1	0	1
	AXE	2	0	1	0	3
	NAUTICA	0	0	0	1	1
	TELMEX	0	0	0	1	1
	HINDS	0	1	0	0	1
	OFFICE DEPOT	0	1	0	0	1
	KFC	0	0	1	0	1
REVLON	0	0	1	0	1	
LEVI'S	1	0	0	0	1	
<b>Total</b>		17	23	33	11	84

La marca Top es Coca Cola, al igual que en el caso de las mujeres, para los hombres tiene frecuencia de 28 de los cuales la mayor parte (15) queda en el grupo de los adultos maduros, seguido por el grupo de los adultos jóvenes con (7) y el grupo de jóvenes con 5 .  
 La marca 2 es Telcel con 9 menciones de las cuales 4 se encuentran en el grupo de los adultos maduros, 2 para los adultos jóvenes y 2 para los adultos mayores.  
 La marca 3 es corona con solo 4 menciones en el grupo de adultos maduros, en la categoría de Leches Lala aparece con 2 menciones, una en jóvenes y la otra en adultos jóvenes y Alpura no aparece como marca conocida de los hombres.



Es cierto que hay top of mind en el que coinciden ambos sexos como es el caso de Coca Cola con 21 menciones en Mujeres y 28 en Hombres, podemos observar en lo que se refiere a Telefonía / Celulares, los Hombres recuerdan más a Telcel con 9 menciones donde 4 de las cuales se encuentran en los Adultos Maduros, esto puede ser debido a que a los hombres les interesa más todo lo relacionado con la tecnología, caso contrario en las mujeres; dado que podemos observar en el caso de Alpura con 8 menciones y Sedal con 7 menciones, marcas relacionadas con productos utilizados en el hogar. En la marca Corona vemos claramente que los Hombres la tienen más en la mente con 4 menciones y en Mujeres no hay mención de la marca y por último observamos a Lala también muy pareja en el número de menciones con 3 en Mujeres y 2 en Hombres.

### 3.6.5 Nivel de conocimiento de las marcas de leche

“Top of mind” es la primera marca que mencionan los consumidores al consultársele sobre un producto.

### 3.6.5.1 Conocimiento espontáneo de marca de leches

#### Top of mind por grupo de edad en mujeres

#### Case Processing Summary

TOM MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES * GRUPOS DE EDADES	104	100.0%	0	.0%	104	100.0%

#### TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES \* GRUPOS DE EDADES

#### Crosstabulation

Count

TOM MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
TOP OF MIND DE MENCION ESPONTANEA MARCA LECHES	AL DIA	0	1	2	1	4
	ALPURA	9	7	17	5	38
	ALPURA CLASICA	0	1	0	0	1
	ALPURA LIGHT	0	1	2	0	3
	ALPURA 2000	1	1	0	1	3
	ALPURA 2000 LIGHT	0	0	2	0	2
	ALPURA 40 Y TANTOS	0	1	1	0	2
	ALPURA DESLACTOSADA	0	0	1	0	1
	FORTILECHE	1	0	0	0	1
	LALA	11	9	4	3	27
	LALA LIGHT	2	1	0	0	3
	LALA PREMIUM ENTERA	0	0	1	0	1
	MILECHE	0	1	0	0	1
	NUTRILECHE	1	1	2	1	5
	PARMALAT	1	0	1	1	3
	SAN MARCOS	1	0	0	0	1
SVELTY	0	0	2	0	2	
OTRA	1	3	1	1	6	
Total		28	27	36	13	104

El top of mind de marcas de leches en mujeres es Alpura con 38 menciones de un total de 104, de las cuales 17 fueron en Adultos Maduros.

## Top of mind por grupo de edad en hombres

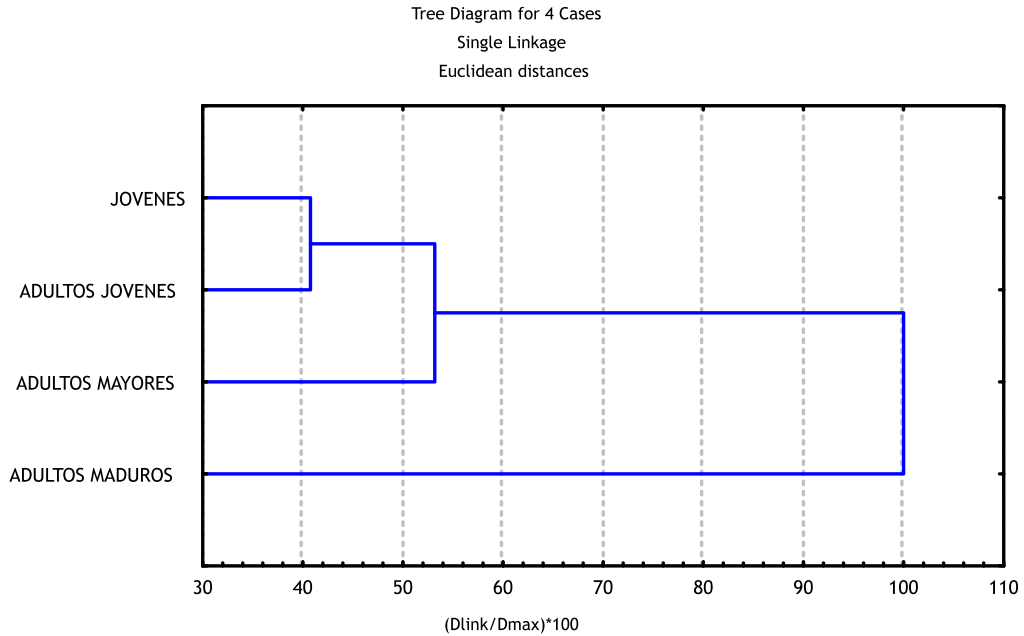
### Case Processing Summary

TOM HOMBRES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES * GRUPOS DE EDADES	96	100.0%	0	.0%	96	100.0%

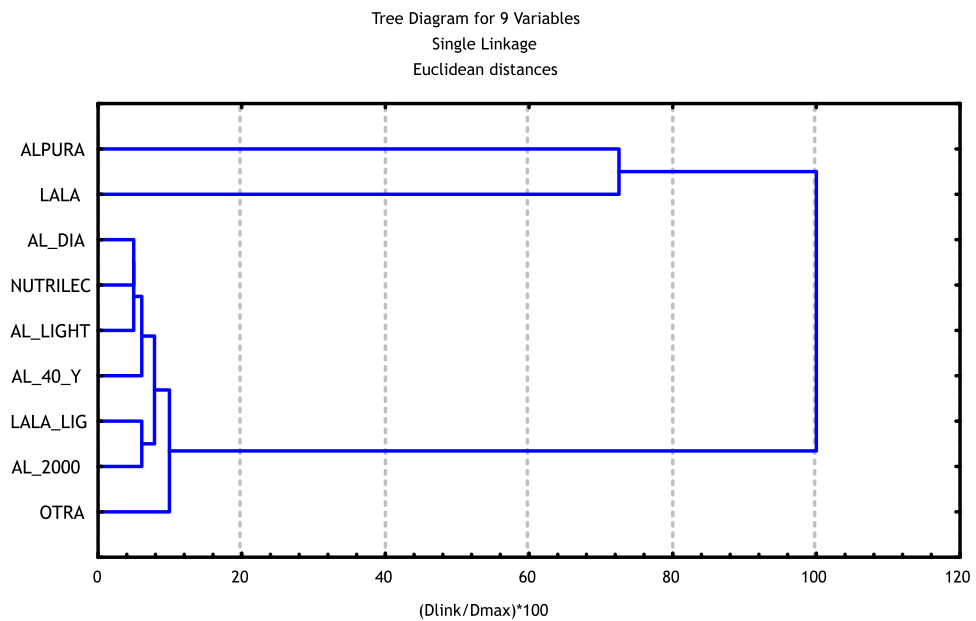
### TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES \* GRUPOS DE EDADES Crosstabulation Count

TOM HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
TOP OF MIND DE MENCION ESPONTANEA MARCA LECHES	AL DIA	2	1	2	1	6
	ALPURA	4	10	14	6	34
	ALPURA LIGHT	0	0	2	1	3
	ALPURA 2000	1	0	1	0	2
	ALPURA 2000 LIGHT	0	0	0	1	1
	ALPURA 40 Y TANTOS	1	0	2	0	3
	BOREAL	1	0	0	0	1
	CONASUPO/LICONSA	1	0	1	0	2
	LALA	7	14	8	6	35
	LALA LIGHT	1	0	2	0	3
	NUTRILECHE	0	1	2	0	3
	SAN MARCOS	0	1	0	0	1
	OTRA	0	1	1	0	2
Total		18	28	35	15	96

El Top Of Mind de marcas de leches en Hombres es Lala con 35 menciones de un total de 96, de las cuales 14 fueron en Adultos Jóvenes. Solo se mencionaron 13 marcas.



Se observa que los jóvenes y adultos jóvenes son los que más conocimiento tienen sobre las marcas de leche, ello puede ser debido a que están más expuestos a la publicidad, después están los adultos mayores y en menor medida por los adultos maduros.



Para ambos sexos, Alpura y Lala son las marcas con puntaje más alto por lo cual hacen un gran grupo, separado de las otras marcas de leche y leche especializada.



**Top of mind por nivel socioeconómico en mujeres**

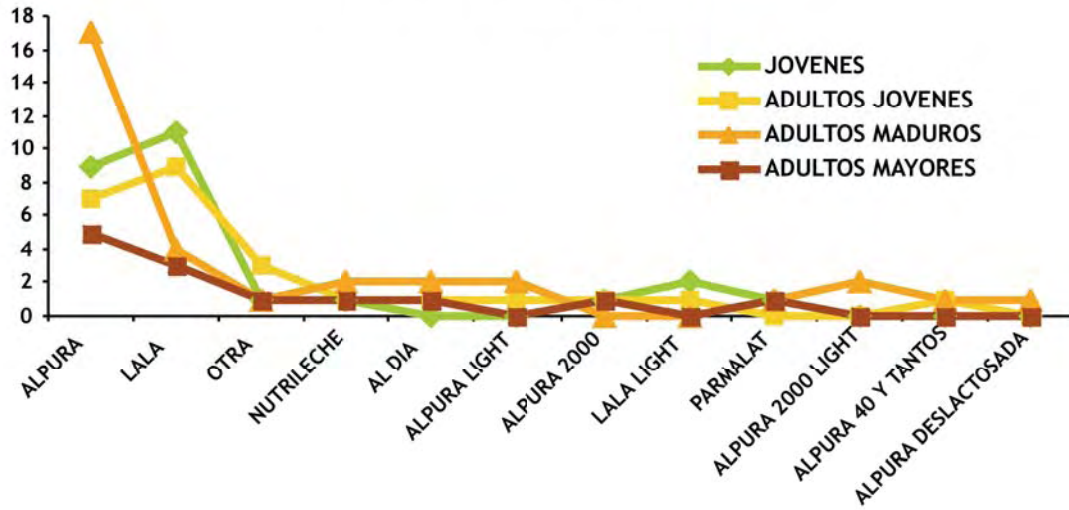
**Case Processing Summary**

MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES * NIVEL SOCIOECONOMICO	104	100.0%	0	.0%	104	100.0%

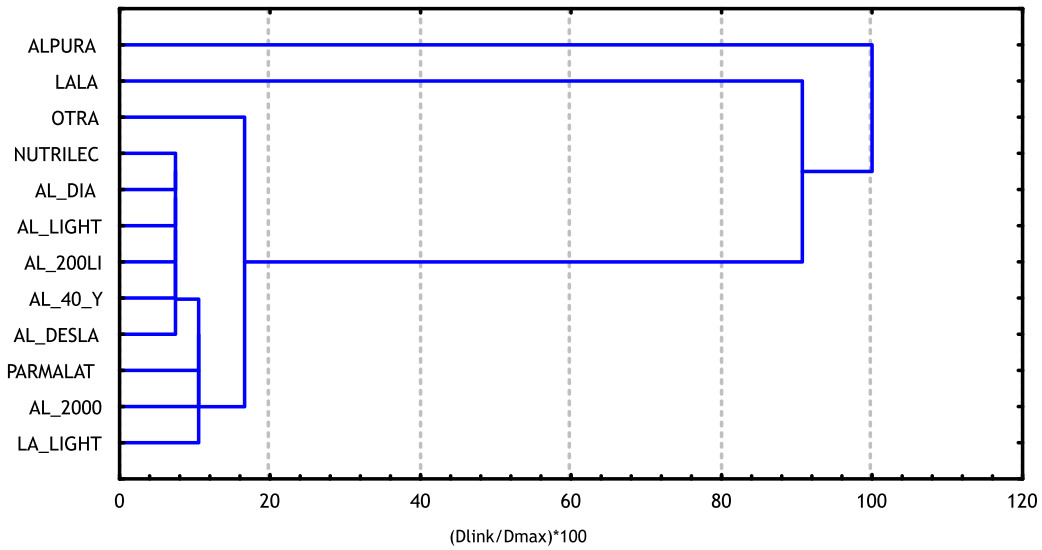
**TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**  
**Count**

		NIVEL SOCIOECONOMICO		Total
		MEDIO	MEDIO ALTO	
TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES	AL DIA	1	3	4
	ALPURA	23	15	38
	ALPURA CLASICA	0	1	1
	ALPURA LIGHT	1	2	3
	ALPURA 2000	2	1	3
	ALPURA 2000 LIGHT	0	2	2
	ALPURA 40 Y TANTOS	2	0	2
	ALPURA DESLACTOSADA	0	1	1
	FORTILECHE	1	0	1
	LALA	12	15	27
	LALA LIGHT	2	1	3
	LALA PREMIUM ENTERA	0	1	1
	MILECHE	0	1	1
	NUTRILECHE	3	2	5
	PARMALAT	3	0	3
	SAN MARCOS	0	1	1
	SVELTY	2	0	2
OTRA	4	2	6	
Total		56	48	104

TOP OF MIND DE MARCA DE LECHES  
POR GRUPO DE EDAD EN MUJERES



Tree Diagram for 12 Variables  
Single Linkage  
Euclidean distances



La tendencia es mayor para Alpura y Lala, sin embargo para las otras leches no es tan alto por ello forman un grupo poco homogéneo.

Top of mind por nivel socioeconómico en hombres

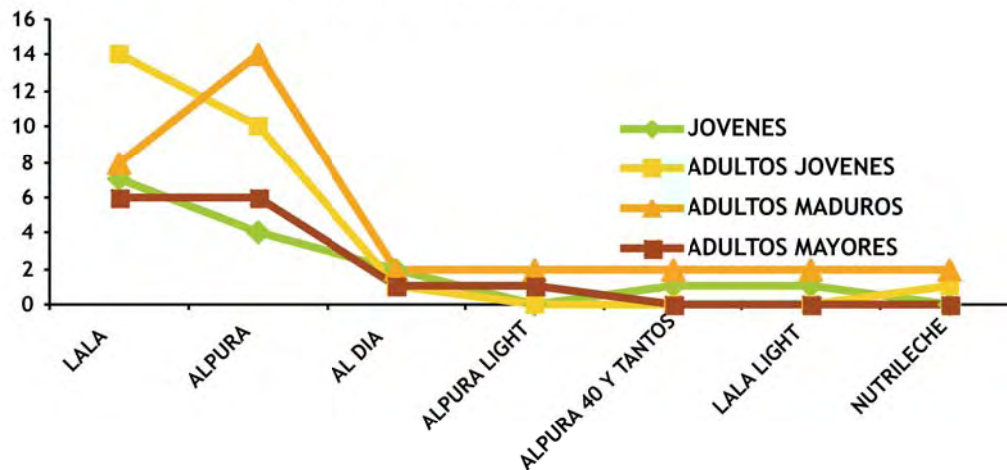
### Case Processing Summary

HOMBRES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES * NIVEL SOCIOECONOMICO	96	100.0%	0	.0%	96	100.0%

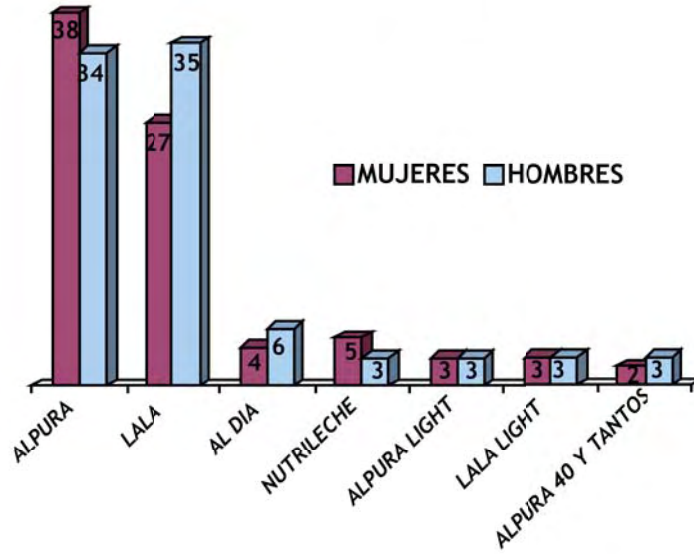
### TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES \* NIVEL SOCIOECONOMICO Crosstabulation Count

HOMBRES		NIVEL SOCIOECONOMICO		Total
		MEDIO	MEDIO ALTO	
TOP OF MIND MENCION ESPONTANEA MARCA DE LECHES	AL DIA	3	3	6
	ALPURA	14	20	34
	ALPURA LIGHT	1	2	3
	ALPURA 2000	1	1	2
	ALPURA 2000 LIGHT	0	1	1
	ALPURA 40 Y TANTOS	2	1	3
	BOREAL	1	0	1
	CONASUPO/LICONSA	2	0	2
	LALA	18	17	35
	LALA LIGHT	0	3	3
	NUTRILECHE	1	2	3
	SAN MARCOS	0	1	1
	OTRA	1	1	2
	Total	44	52	96

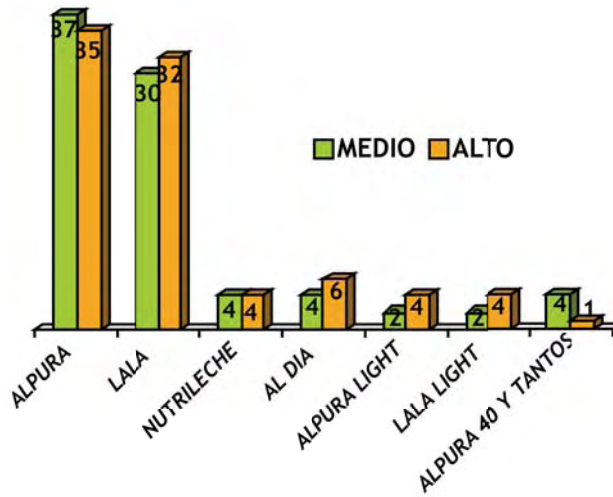
### TOP OF MIND DE MARCA DE LECHES POR GRUPO DE EDAD EN HOMBRES



TOP OF MIND DE MARCA DE LECHE  
MUJERES VS HOMBRES



TOP OF MIND DE MARCA DE LECHE  
POR NIVELES SOCIOECONÓMICOS

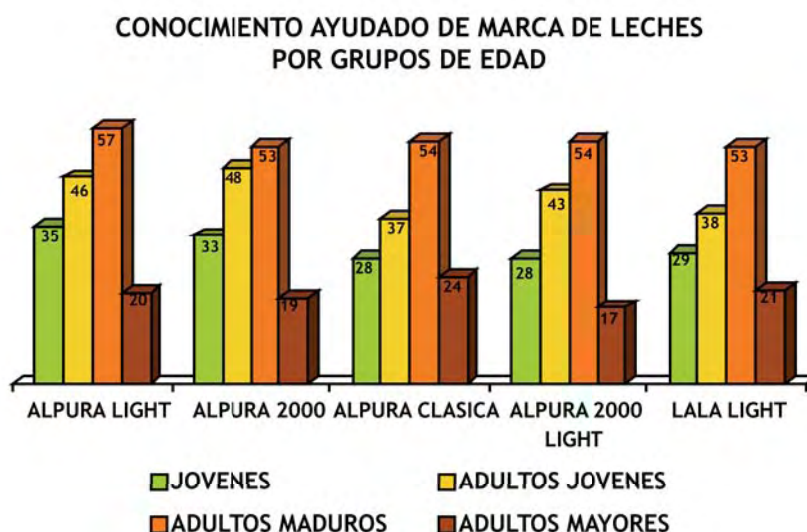


### 3.6.5.2 Conocimiento ayudado de marca de leches

Para ver la distribución de frecuencias del conocimiento ayudado de marcas de leche con respecto a los grupos de edad, pedimos al programa una tabla de frecuencias.

TOTAL DE RECORDACION AYUDADA DE MARCAS DE TODOS LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MARCAS
111	29	30	40	12	AL DIA
36	4	15	12	5	ALPURA
143	28	37	54	24	ALPURA CLASICA
133	24	39	55	15	ALPURA SEMIDESCREMADA
158	35	46	57	20	ALPURA LIGHT
153	33	48	53	19	ALPURA 2000
120	17	30	52	21	ALPURA 2000 SEMIDESCREMADA
142	28	43	54	17	ALPURA 2000 LIGHT
104	21	28	46	9	ALPURA 40 Y TANTOS
100	24	35	31	10	ALPURA DESLACTOSADA
45	13	10	15	7	ALPURA LIGHT DESLACTOSADA
31	7	9	9	6	ALPURA LIGHT EXTRA
51	16	17	12	6	ALPURA SIN COLESTEROL
110	27	25	48	10	ALPURA FRUTAL
24	8	7	6	3	ALPURA FRUTAL ULTRAPASTEURIZADA
30	11	10	8	1	ALPURA FRUTAL PARA LLEVAR
45	9	16	10	10	ALPURA EN POLVO
59	19	20	14	6	ALPURA KIDS
57	6	23	24	4	BADEN
110	22	30	42	16	BOREAL
132	30	38	52	12	CONASUPO/LICONSA
38	5	12	17	4	DESLY
113	27	33	41	12	FORTILECHE
29	7	8	10	4	LALA
35	12	11	5	7	LALA DESARROLLO
86	17	24	34	11	LALA DESLACTOSADA
141	29	38	53	21	LALA LIGHT
126	27	34	48	17	LALA PREMIUM ENTERA
58	11	20	21	6	LALA SEMI (SEMIDESCREMADA)
46	12	15	11	8	LALA SILUETTE PLUS LIGHT
27	10	3	8	6	LALA SUPREMA
14	4	2	5	3	LALA SUPREMA LIGHT
130	30	32	52	16	MILECHE
125	26	33	47	19	NUTRILECHE
126	28	39	47	12	PARMALAT
80	14	26	33	7	SAN MARCOS
98	16	30	39	13	SVELTY
7	1	3	3	0	OTRA

En general la recordación ayudada de marca es más alta en grupos de Adultos Maduros, seguida por el grupo de Adultos Jóvenes, Jóvenes y Adultos Mayores.



### 3.6.5.3 Awareness de marca de leches

El Awareness es simplemente el conocimiento de que la marca existe, es a través del Awareness que un consumidor pronto puede considerar la posibilidad de comprar un producto.

#### Awareness de marca de leche en grupos de edad de mujeres

Se presenta a continuación las 10 marcas que se mencionaron con mayor frecuencia.

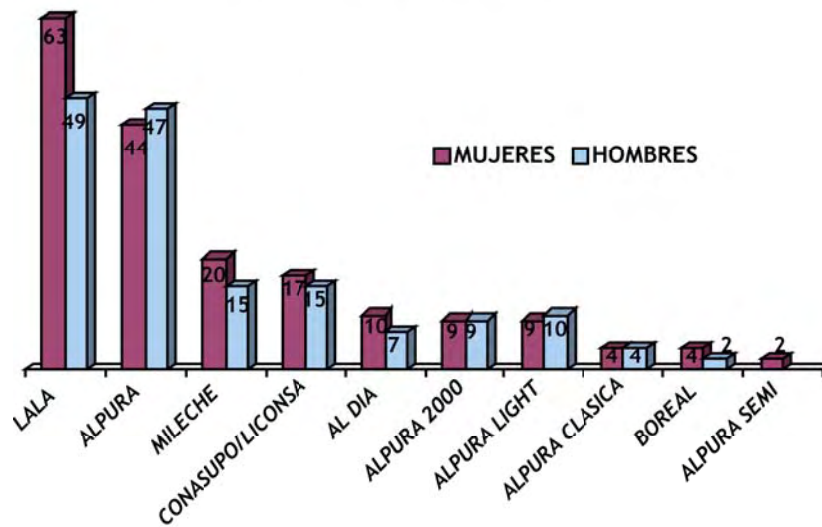
MUJERES	JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	TOTAL
LALA	14	14	25	10	63
ALPURA	16	10	13	5	44
MILECHE	3	10	6	1	20
CONASUPO/LICONSA	1	3	7	6	17
AL DIA	3	2	3	2	10
ALPURA 2000	1	1	5	2	9
ALPURA LIGHT	4	1	3	1	9
ALPURA CLASICA	1	1	1	1	4
BOREAL	1	3	0	0	4
ALPURA SEMI	1	0	1	0	2

## Awareness de marca me leche en grupos de edad de hombres

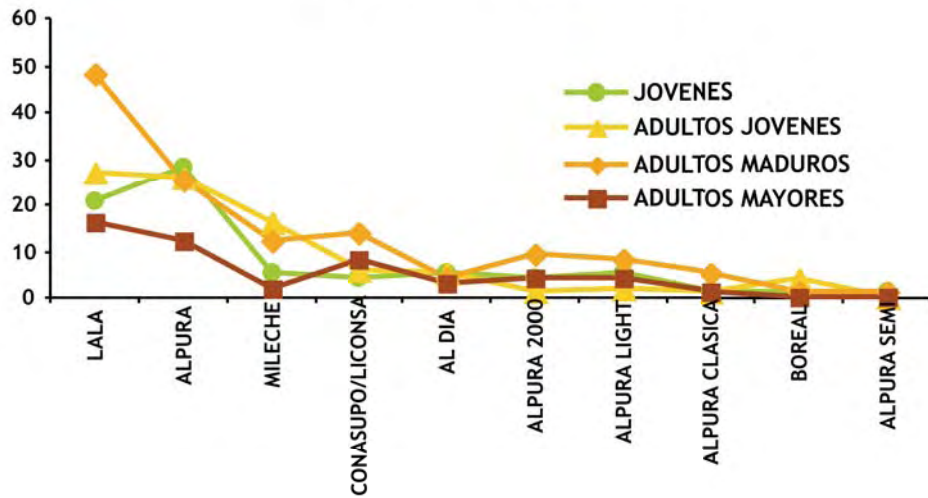
Se presenta a continuación las 10 marcas que se mencionaron con mayor frecuencia.

HOMBRES	JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	TOTAL
LALA	7	13	23	6	49
ALPURA	12	16	12	7	47
MILECHE	2	6	6	1	15
CONASUPO/LICONSA	3	3	7	2	15
AL DIA	2	3	1	1	7
ALPURA 2000	3	0	4	2	9
ALPURA LIGHT	1	1	5	3	10
ALPURA CLASICA	0	0	4	0	4
BOREAL	0	1	1	0	2
ALPURA SEMI	0	0	0	0	0

## AWARENESS DE MARCAS DE LECHE POR SEXO



### AWARENESS DE MARCAS DE LECHE



### 3.6.6 Nivel de recordación publicitaria de marcas de leche

#### 3.6.6.1 Recordación espontánea de marcas de leche publicitadas

##### Top of mind por grupo de edad en mujeres

#### Case Processing Summary

MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHE EN EL ULTIMO MES * GRUPOS DE EDADES	103	99.0%	1	1.0%	104	100.0%



**TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHES EN EL ULTIMO MES \***  
**GRUPOS DE EDADES Crosstabulation**  
 Count

MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
TOP OF MIND DE MENCION ESPONTANEA PUBLICIDAD DE MARCA LECHES EN EL ULTIMO MES	ALPURA	15	8	21	6	50
	ALPURA CLASICA	0	1	0	0	1
	ALPURA LIGHT	0	2	2	0	4
	ALPURA 2000	0	0	0	1	1
	ALPURA 2000 LIGHT	1	0	2	1	4
	ALPURA 40 Y TANTOS	0	1	1	0	2
	ALPURA DESLACTOSADA	1	0	1	0	2
	ALPURA FRUTAL	0	0	0	1	1
	LALA	8	11	6	3	28
	LALA LIGHT	0	1	0	1	2
	LALA PREMIUM ENTERA	0	0	1	0	1
	PARMALAT	1	0	0	0	1
	SVELTY	1	0	0	0	1
	OTRA	1	2	2	0	5
Total		28	26	36	13	103

El Top Of Mind de marcas publicitadas en el grupo de las Mujeres fue Alpura con 50 de 103 menciones, de las cuales 21 corresponden al grupo de edad de Adultos Maduros y 15 en Jóvenes.

**Top of mind Por nivel socioeconómico en mujeres**

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHES EN EL ULTIMO MES * NIVEL SOCIOECONOMICO	103	99.0%	1	1.0%	104	100.0%

**TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHE EN EL ULTIMO MES \*  
NIVEL SOCIOECONOMICO Crosstabulation  
Count**

		NIVEL SOCIOECONOMICO		Total
		MEDIO	MEDIO ALTO	
TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHE EN EL ULTIMO MES	ALPURA	28	22	50
	ALPURA CLASICA	0	1	1
	ALPURA LIGHT	1	3	4
	ALPURA 2000	1	0	1
	ALPURA 2000 LIGHT	1	3	4
	ALPURA 40 Y TANTOS	2	0	2
	ALPURA DESLACTOSADA	1	1	2
	ALPURA FRUTAL	1	0	1
	LALA	14	14	28
	LALA LIGHT	1	1	2
	LALA PREMIUM ENTERA	0	1	1
	PARMALAT	1	0	1
	SVELTY	1	0	1
	OTRA	4	1	5
Total		56	47	103

**Top of mind por grupo de edad en hombres**

**Case Processing Summary**

HOMBRES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHE EN EL ULTIMO MES * GRUPOS DE EJES	93	96.9%	3	3.1%	96	100.0%

**TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHES EN EL ULTIMO MES \***  
**GRUPOS DE EDADES**  
**Crosstabulation**  
**Count**

HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
TOP OF MIND MARCA MENCION ESPONTANEA PUBLICIDAD DE LECHES EN EL ULTIMO MES	AL DIA	1	0	0	1	2
	ALPURA	10	14	17	6	47
	ALPURA CLASICA	0	0	1	0	1
	ALPURA LIGHT	0	0	2	1	3
	ALPURA 2000	0	0	0	1	1
	ALPURA 2000 LIGHT	0	0	0	1	1
	ALPURA 40 Y TANTOS	0	0	0	1	1
	ALPURA DESLACTOSADA	0	0	2	0	2
	FORTILECHE	0	1	0	0	1
	LALA	3	9	7	3	22
	LALA DESLACTOSADA	0	1	0	0	1
	LALA LIGHT	1	0	2	0	3
	NUTRILECHE	0	0	1	0	1
	OTRA	2	1	3	1	7
Total		17	26	35	15	93

El Top Of Mind de marcas publicitadas en el grupo de Hombres es también Alpura con 47 de 93 menciones, de las cuales 17 corresponden al grupo de edad de Adultos Maduros, 14 en Adultos Jóvenes.

**Top of mind por nivel socioeconómico en hombres**

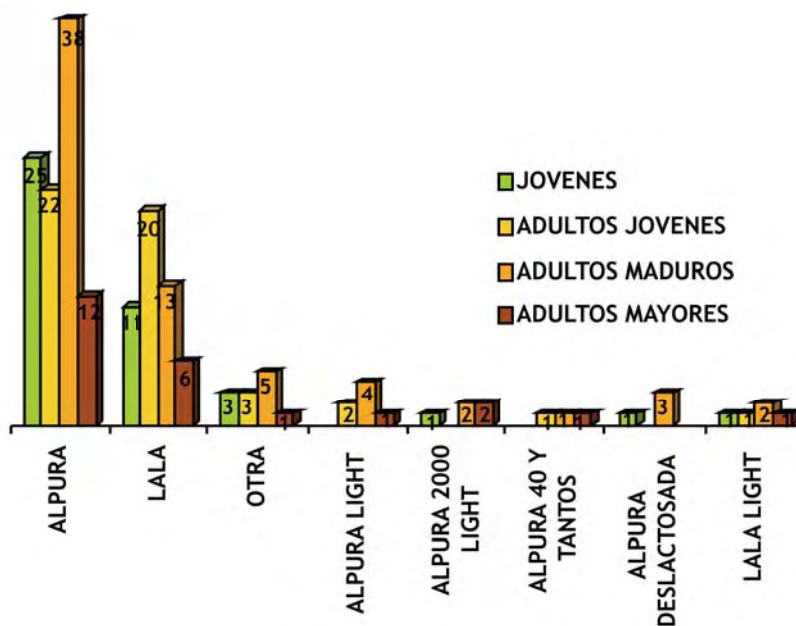
**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHES EN EL ULTIMO MES * NIVEL SOCIOECONOMICO	93	96.9%	3	3.1%	96	100.0%

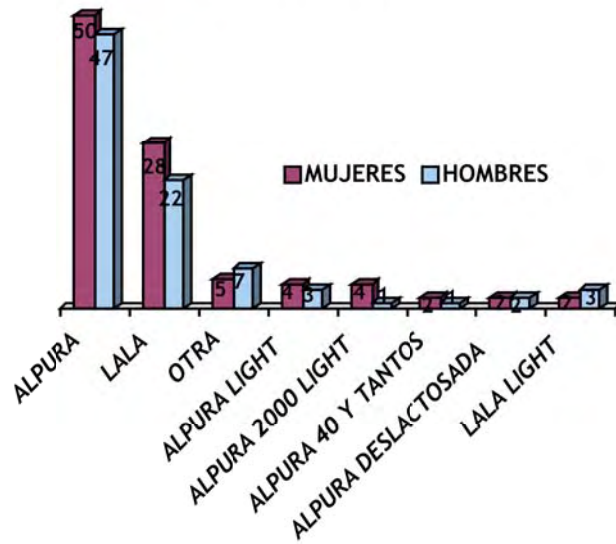
**TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHES EN EL ULTIMO MES \*  
NIVEL SOCIOECONOMICO Crosstabulation  
Count**

		NIVEL SOCIOECONOMICO		
		MEDIO	MEDIO ALTO	Total
TOP OF MIND MENCION ESPONTANEA PUBLICIDAD DE MARCA DE LECHES EN EL ULTIMO MES	AL DIA	0	2	2
	ALPURA	19	28	47
	ALPURA CLASICA	1	0	1
	ALPURA LIGHT	1	2	3
	ALPURA 2000	1	0	1
	ALPURA 2000 LIGHT	0	1	1
	ALPURA 40 Y TANTOS	0	1	1
	ALPURA DESLACTOSADA	2	0	2
	FORTILECHE	1	0	1
	LALA	13	9	22
	LALA DESLACTOSADA	1	0	1
	LALA LIGHT	0	3	3
	NUTRILECHE	0	1	1
	OTRA	4	3	7
<b>Total</b>		<b>43</b>	<b>50</b>	<b>93</b>

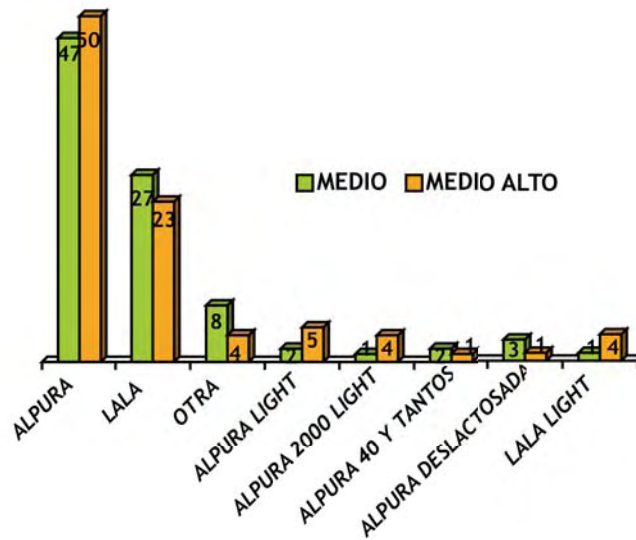
**TOP OF MIND DE MARCAS DE LECHES PUBLICITADAS  
POR GRUPO DE EDAD**



TOP OF MIND DE MARCA DE LECHE PUBLICITADAS  
MUJERES VS HOMBRES



TOP OF MIND DE LECHE PUBLICITADAS  
POR NIVEL SOCIOECONÓMICO



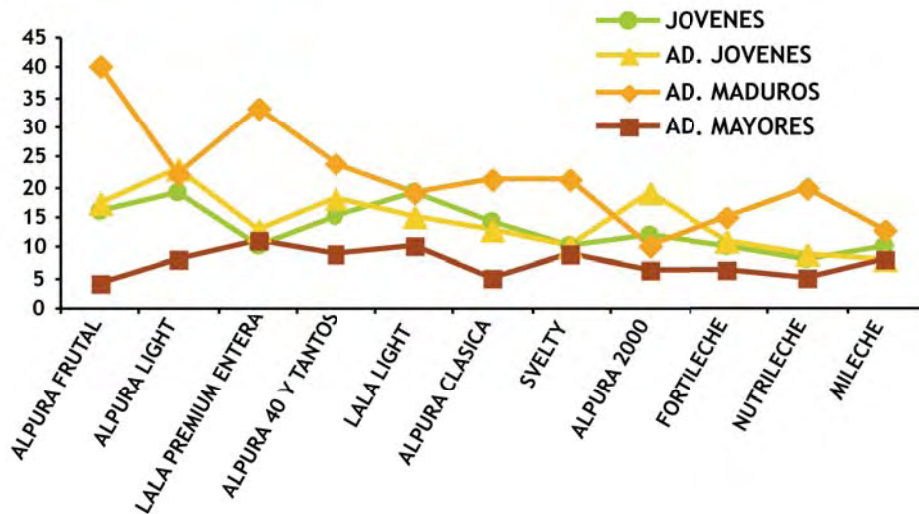
### 3.6.6.2 Recordación ayudada de marca de leches publicitadas

Para ver la distribución de frecuencias de la recordación ayudada de marcas de leche publicitadas con respecto a los grupos de edad, se hace mediante una tabla de frecuencias.

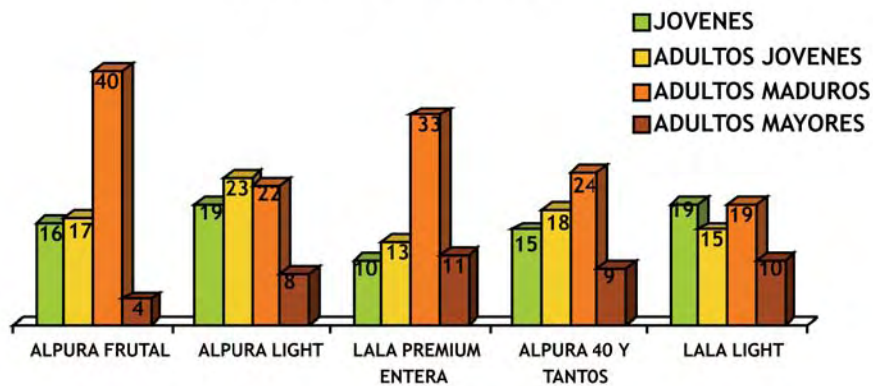
Se consideró todas las menciones de recordación ayudada de marca de leches publicitadas. Estos resultados se pasaron a Excel para sumar la frecuencia en los grupos de edad y dar la posición de la recordación ayudada de marcas publicitadas. En general la recordación ayudada de marca es más alta en grupos de Adultos Maduros, seguida por el grupo de Adultos Jóvenes, Jóvenes y Adultos Mayores.

PUBLICIDAD AYUDADA DE TODOS LOS GRUPOS DE EDADES Y AMBOS	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MARCAS
38	11	7	16	4	AL DIA
28	3	9	10	6	ALPURA
53	14	13	21	5	ALPURA CLASICA
17	4	5	3	5	ALPURA SEMIDESCREMADA
72	19	23	22	8	ALPURA LIGHT
47	12	19	10	6	ALPURA 2000
19	1	7	7	4	ALPURA 2000 SEMIDESCREMADA
30	9	10	6	5	ALPURA 2000 LIGHT
66	15	18	24	9	ALPURA 40 Y TANTOS
22	6	7	7	2	ALPURA DESLACTOSADA
14	4	4	4	2	ALPURA LIGHT DESLACTOSADA
9	2	3	2	2	ALPURA LIGHT EXTRA
18	7	5	5	1	ALPURA SIN COLESTEROL
77	16	17	40	4	ALPURA FRUTAL
5	2	1	2	0	ALPURA FRUTAL ULTRAPASTEURIZADA
11	3	3	4	1	ALPURA FRUTAL PARA LLEVAR
5	1	2	2	0	ALPURA EN POLVO
20	6	9	2	3	ALPURA KIDS
2	0	0	2	0	BADEN
2	1	0	0	1	BOREAL
5	0	1	3	1	CONASUPO/LICONSA
1	0	0	0	1	DESLY
42	10	11	15	6	FORTILECHE
23	4	8	8	3	LALA
6	1	2	3	0	LALA DESARROLLO
15	6	5	4	0	LALA DESLACTOSADA
63	19	15	19	10	LALA LIGHT
67	10	13	33	11	LALA PREMIUM ENTERA
5	1	1	1	2	LALA SEMI (SEMIDESCREMADA)
13	3	3	4	3	LALA SILUETTE PLUS LIGHT
1	0	0	1	0	LALA SUPREMA LIGHT
39	10	8	13	8	MILECHE
42	8	9	20	5	NUTRILECHE
8	2	3	2	1	PARMALAT
3	2	0	0	1	SAN MARCOS
50	10	10	21	9	SVELTY
10	4	1	5	0	OTRA

### RECORDACIÓN AYUDADA DE MARCA DE LECHE PUBLICADAS POR GRUPOS DE EDADES



### RECORDACIÓN AYUDADA DE MARCA DE LECHE PUBLICADAS PRIMERAS 5 POSICIONES



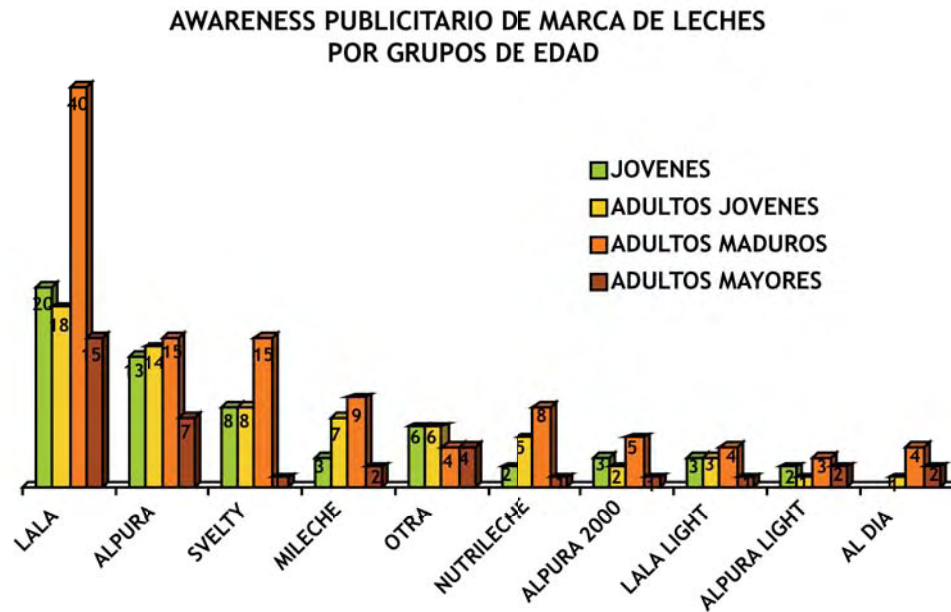
#### 3.6.6.3 Awareness publicitario de marca de leches

El awareness es simplemente el conocimiento de que la marca existe, es a través del awareness que un consumidor pronto puede considerar la posibilidad de comprar un producto.

Se consideró de la segunda a la quinta recordación ayudada de marca de leches publicadas. Estos resultados se pasaron a Excel para sumar la frecuencia en los grupos de edad y dar la posición de la recordación ayudada.

MARCAS	JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	SUMA
AL DIA	0	1	4	2	7
ALPURA	13	14	15	7	49
ALPURA CLASICA	1	1	2	0	4
ALPURA SEMIDESCREMADA	1	0	0	0	1
ALPURA LIGHT	2	1	3	2	8
ALPURA 2000	3	2	5	1	11
ALPURA 40 Y TANTOS	2	0	2	2	6
ALPURA DESLACTOSADA	0	0	1	2	3
ALPURA LIGHT DESLACTOSADA	0	0	1	0	1
ALPURA LIGHT EXTRA	1	0	0	0	1
ALPURA SIN COLESTEROL	0	0	0	1	1
ALPURA FRUTAL	0	0	3	1	4
CONASUPO/LICONSA	1	0	0	1	2
FORTILECHE	1	2	1	0	4
LALA	20	18	40	15	93
LALA LIGHT	3	3	4	1	11
LALA PREMIUM ENTERA	0	0	4	2	6
MILECHE	3	7	9	2	21
NUTRILECHE	2	5	8	1	16
SVELTY	8	8	15	1	32
OTRA	6	6	4	4	20

Lala con 93 menciones, Alpura con 49, Svelty con 39, Mileche 21, Otra 20, Nutrileche, Alpura 2000 y Lala Light con 11, Alpura Light 8 y Al Dia 7.





### 3.6.7 Indicadores de preferencia y fidelidad hacia las marcas de leche

#### 3.6.7.1 Marcas de leches que ha consumido en los últimos 3 meses

Se observará la frecuencia de las marcas de leches que se han consumido en los últimos 3 meses con respecto a los grupos de edad.

Los datos se pasaron a Excel para sumar las frecuencias, quedando así las primeras 10 posiciones:

- |                                |  |
|--------------------------------|--|
| 1. Lala (50 menciones)         | 7. Otra (14 menciones)                 |
| 2. Alpura (45 menciones)       | 8. Alpura Clásica (12 menciones)       |
| 3. Alpura 2000 (32 menciones)  | 9. Alpura 2000 Light (11 menciones)    |
| 4. Al día (21 menciones)       | 10. Conasupo / Liconsa (11 menciones)  |
| 5. Nutrileche (19 menciones)   | 11. Lala Premium Entera (11 menciones) |
| 6. Alpura Light (16 menciones) | 12. Lala Light (10 menciones)          |

TOTAL DE TODOS LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MARCAS CONSUMIDAS EN LOS ULTIMOS 3 MESES
21	4	6	10	1	AL DIA
45	12	13	14	6	ALPURA
12	1	4	6	1	ALPURA CLASICA
6	1	2	2	1	ALPURA SEMIDESCREMADA
16	1	4	7	4	ALPURA LIGHT
32	3	9	14	6	ALPURA 2000
1	0	1	0	0	ALPURA 2000 SEMIDESCREMADA
11	0	3	6	2	ALPURA 2000 LIGHT
2	0	1	1	0	ALPURA 40 Y TANTOS
1	1	0	0	0	ALPURA DESLACTOSADA
2	0	1	1	0	ALPURA LIGHT DESLACTOSADA
1	0	1	0	0	ALPURA SIN COLESTEROL
2	0	1	1	0	ALPURA FRUTAL
1	0	0	0	1	ALPURA EN POLVO
1	0	1	0	0	ALPURA KIDS
11	1	6	3	1	CONASUPO/LICONSA
50	19	14	12	5	LALA
1	0	0	0	1	LALA DESLACTOSADA
10	2	2	4	2	LALA LIGHT
11	1	2	6	2	LALA PREMIUM ENTERA
1	0	0	0	1	LALA SILUETTE PLUS LIGHT
2	1	1	0	0	MILECHE
19	7	3	8	1	NUTRILECHE
4	0	2	0	2	PARMALAT
5	1	0	3	1	SAN MARCOS
9	3	1	4	1	SVELTY
14	6	5	2	1	OTRA

### 3.6.7.2 Marca de leches que ha consumido en el último mes

Se observa la frecuencia de las marcas de leches que se han consumido en el último mes con respecto a los grupos de edad.

Los datos se pasaron a Excel para sumar las frecuencias, quedando así las primeras 10 posiciones:

- |                                |  |
|--------------------------------|--|
| 1. Lala (53 menciones)         | 7. Otra (13 menciones)                 |
| 2. Alpura (40 menciones)       | 8. Alpura Clásica (12 menciones)       |
| 3. Alpura 2000 (32 menciones)  | 9. Alpura 2000 Light (11 menciones)    |
| 4. Nutrileche (19 menciones)   | 10. Conasupo / Liconsa (10 menciones)  |
| 5. Al día (16 menciones)       | 11. Lala Premium Entera (10 menciones) |
| 6. Alpura Light (16 menciones) | 12. Svelty (10 menciones)              |

TOTAL DE TODOS LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MARCAS CONSUMIDAS EN EL ULTIMO MES
16	2	4	9	1	AL DIA
40	10	12	13	5	ALPURA
12	1	4	6	1	ALPURA CLASICA
5	0	2	2	1	ALPURA SEMIDESCREMADA
16	1	4	7	4	ALPURA LIGHT
32	3	9	13	7	ALPURA 2000
11	0	3	6	2	ALPURA 2000 LIGHT
1	0	1	0	0	ALPURA 40 Y TANTOS
1	1	0	0	0	ALPURA DESLACTOSADA
2	0	1	1	0	ALPURA LIGHT DESLACTOSADA
1	0	1	0	0	ALPURA SIN COLESTEROL
1	0	0	0	1	ALPURA EN POLVO
1	0	1	0	0	ALPURA KIDS
10	1	5	3	1	CONASUPO/LICONSA
53	20	17	11	5	LALA
1	0	0	0	1	LALA DESLACTOSADA
7	1	2	4	0	LALA LIGHT
10	1	1	6	2	LALA PREMIUM ENTERA
1	0	0	0	1	LALA SILUETTE PLUS LIGHT
2	1	1	0	0	MILECHE
19	6	4	8	1	NUTRILECHE
3	0	1	0	2	PARMALAT
5	1	0	3	1	SAN MARCOS
10	3	2	4	1	SVELTY
13	6	4	2	1	OTRA

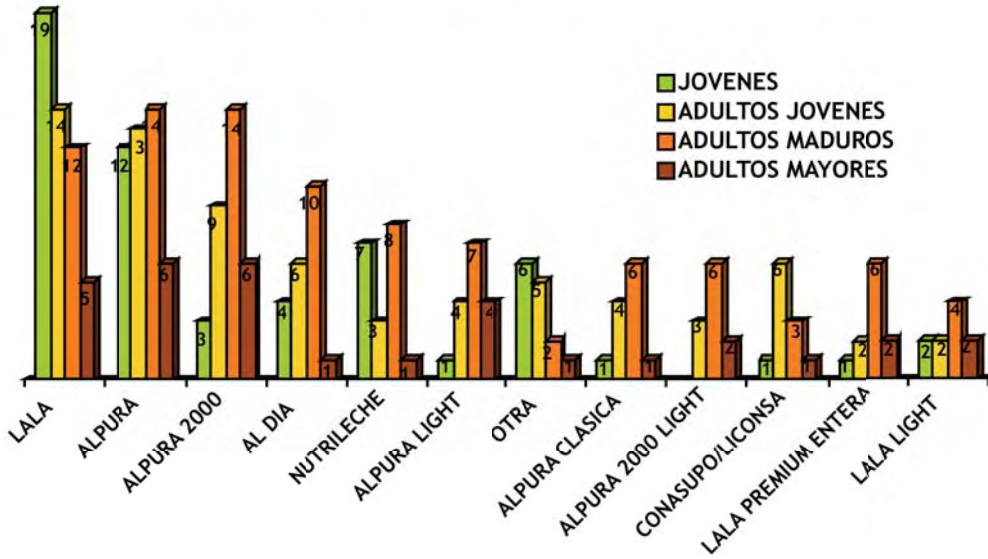
### 3.6.7.3 Marcas de leches que ha consumido en la última semana

Se observa la frecuencia de las marcas de leche que se han consumido en la última semana con respecto a los grupos de edad. Los datos se pasaron a Excel para sumar las frecuencias, quedando así las 10 primeras posiciones:

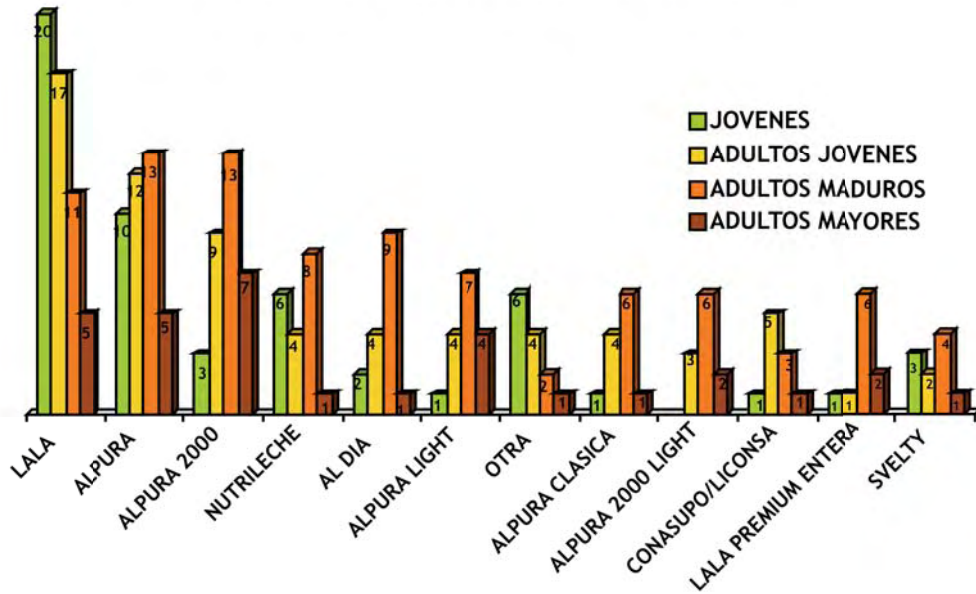
- |                                |                                       |
|--------------------------------|---------------------------------------|
| 1. Lala (50 menciones)         | 7. Alpura Clásica (12 menciones)      |
| 2. Alpura (36 menciones)       | 8. Otra (10 menciones)                |
| 3. Alpura 2000 (31 menciones)  | 9. Alpura 2000 Light (9 menciones)    |
| 4. Al día (17 menciones)       | 10. Conasupo / Liconsa (9 menciones)  |
| 5. Nutrileche (17 menciones)   | 11. Lala Premium Entera (8 menciones) |
| 6. Alpura Light (15 menciones) | 12. Svelty (7 menciones)              |

TOTAL DE TODOS LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MARCAS CONSUMIDAS EN LA ULTIMA SEMANA
17	3	4	9	1	AL DIA
36	8	10	13	5	ALPURA
12	1	4	6	1	ALPURA CLASICA
4	0	2	2	0	ALPURA SEMIDESCREMADA
15	1	4	7	3	ALPURA LIGHT
31	3	8	13	7	ALPURA 2000
9	0	2	5	2	ALPURA 2000 LIGHT
1	0	1	0	0	ALPURA 40 Y TANTOS
1	1	0	0	0	ALPURA DESLACTOSADA
1	0	0	1	0	ALPURA LIGHT DESLACTOSADA
1	0	1	0	0	ALPURA KIDS
9	1	5	2	1	CONASUPO/LICONSA
50	19	16	10	5	LALA
1	0	1	0	0	LALA DESLACTOSADA
6	1	2	3	0	LALA LIGHT
8	1	0	5	2	LALA PREMIUM ENTERA
1	0	0	0	1	LALA SILUETTE PLUS LIGHT
2	1	1	0	0	MILECHE
17	5	3	7	2	NUTRILECHE
3	0	1	0	2	PARMALAT
5	1	0	3	1	SAN MARCOS
7	3	0	3	1	SVELTY
10	6	4	0	0	OTRA

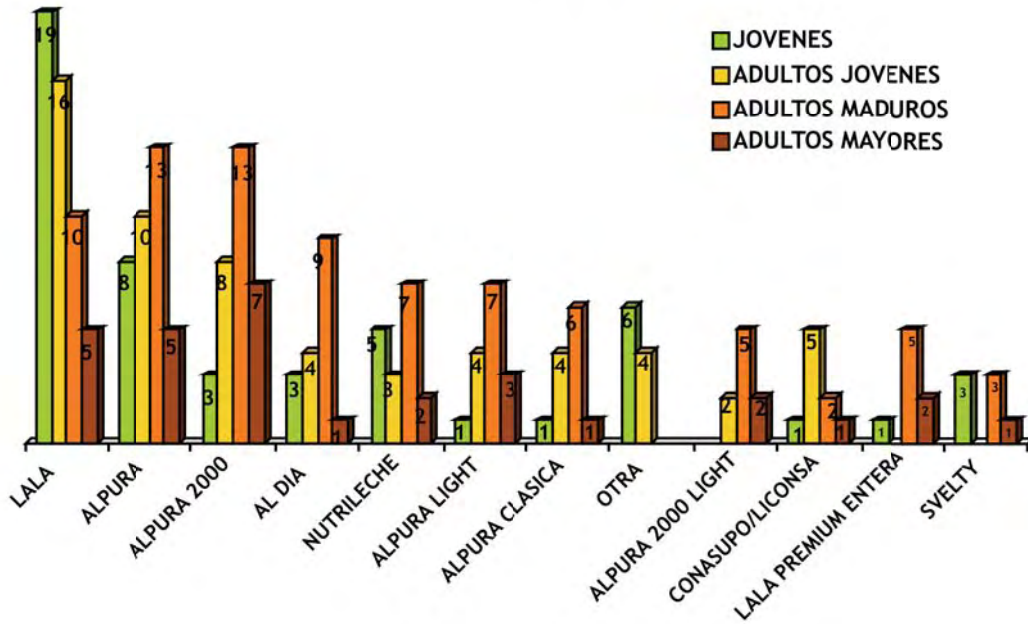
### MARCAS DE LECHES CONSUMIDAS EN LOS ÚLTIMOS 3 MESES



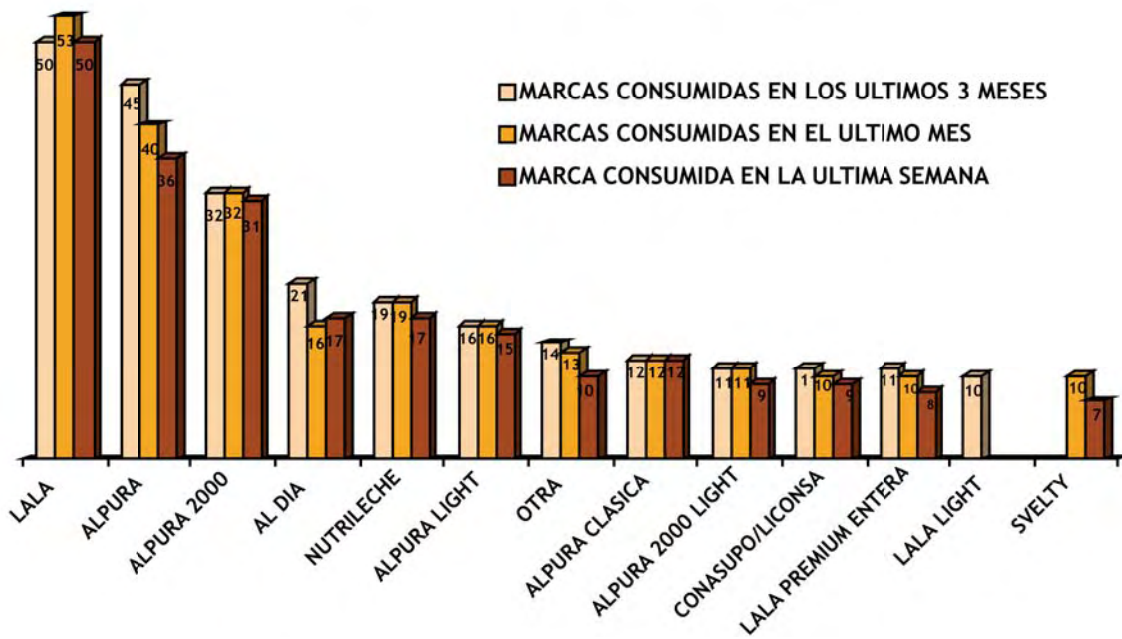
### MARCAS DE LECHES CONSUMIDAS EN EL ÚLTIMO MES



MARCAS DE LECHES CONSUMIDAS EN LA ÚLTIMA SEMANA  
POR GRUPOS DE EDAD



MARCAS CONSUMIDAS ÚLTIMAMENTE



### 3.6.7.4 Marca de leche favorita

Se observa cuál es la marca favorita con respecto a los grupos de edad.

	GRUPOS DE EDADES			
	JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES
	MARCA DE LECHE FAVORITA	MARCA DE LECHE FAVORITA	MARCA DE LECHE FAVORITA	MARCA DE LECHE FAVORITA
AL DIA	2	4	6	1
ALPURA	9	8	12	4
ALPURA CLASICA	1	2	6	1
ALPURA SEMIDESCREMADA	1	1		
ALPURA LIGHT		3	5	3
ALPURA 2000	2	7	10	5
ALPURA 2000 LIGHT		2	5	2
ALPURA 40 Y TANTOS		1		
ALPURA DESLACTOSADA	1			
ALPURA LIGHT DESLACTOSADA			1	
ALPURA KIDS		1		
CONASUPO/LICONSA		2	1	
LALA	16	13	10	4
LALA LIGHT	1	2	2	
LALA PREMIUM ENTERA	1		4	2
LALA SILUETTE PLUS LIGHT				1
MILECHE		2		
NUTRILECHE	4	4	4	1
PARMALAT				1
SAN MARCOS	1		2	1
SVELTY	3	1	3	1
OTRA	4	2		1

Se observa cuál es la marca favorita con respecto al sexo y al nivel socioeconómico.

#### Nivel Medio

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA DE LECHE FAVORITA * SEXO * NIVEL SOCIOECONOMICO	100	100.0%	0	.0%	100	100.0%

**MARCA DE LECHE FAVORITA \* SEXO \* NIVEL SOCIOECONOMICO**  
Crosstabulation

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO	MARCA DE LECHE FAVORITA	AL DIA	3	3	6
		ALPURA	7	6	13
		ALPURA CLASICA	3	3	6
		ALPURA SEMIDESCREMADA	1	1	2
		ALPURA LIGHT	3	0	3
		ALPURA 2000	10	6	16
		ALPURA 2000 LIGHT	2	1	3
		CONASUPO/LICONSA	1	0	1
		LALA	14	13	27
		LALA PREMIUM ENTERA	3	1	4
		LALA SILUETTE PLUS LIGHT	1	0	1
		NUTRILECHE	4	5	9
		PARMALAT	1	0	1
		SAN MARCOS	0	1	1
		SVELTY	2	1	3
OTRA	1	3	4		
Total		56	44	100	

**Nivel Medio Alto**

**Case Processing Summary**

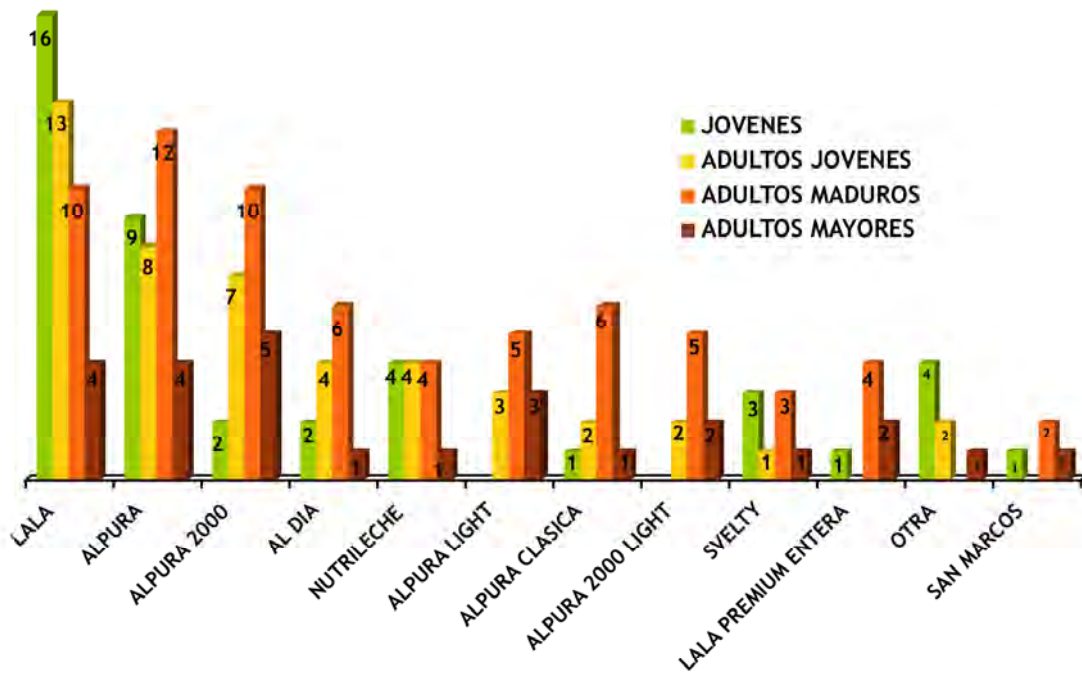
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA DE LECHE FAVORITA * SEXO * NIVEL SOCIOECONOMICO	100	100.0%	0	.0%	100	100.0%

**MARCA DE LECHE FAVORITA \* SEXO \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**  
**Count**

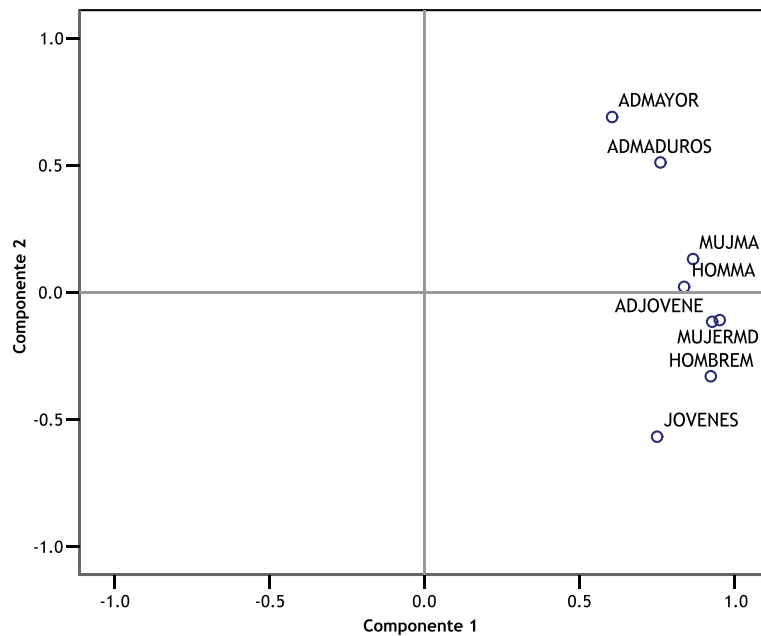
NIVEL SOCIOECONOMICO		SEXO		Total	
		MUJER	HOMBRE		
MEDIO ALTO	MARCA DE LECHE FAVORITA	AL DIA	4	3	7
		ALPURA	6	14	20
		ALPURA CLASICA	2	2	4
		ALPURA LIGHT	4	4	8
		ALPURA 2000	4	4	8
		ALPURA 2000 LIGHT	4	2	6
		ALPURA 40 Y TANTOS	1	0	1
		ALPURA DESLACTOSADA	0	1	1
		ALPURA LIGHT DESLACTOSADA	1	0	1
		ALPURA KIDS	1	0	1
		CONASUPO/LICONSA	1	1	2
		LALA	7	9	16
		LALA LIGHT	1	4	5
		LALA PREMIUM ENTERA	2	1	3
		MILECHE	1	1	2
		NUTRILECHE	2	2	4
		SAN MARCOS	1	2	3
		SVELTY	5	0	5
		OTRA	1	2	3
Total		48	52	100	



### MARCA DE LECHE FAVORITA POR GRUPOS DE EDAD

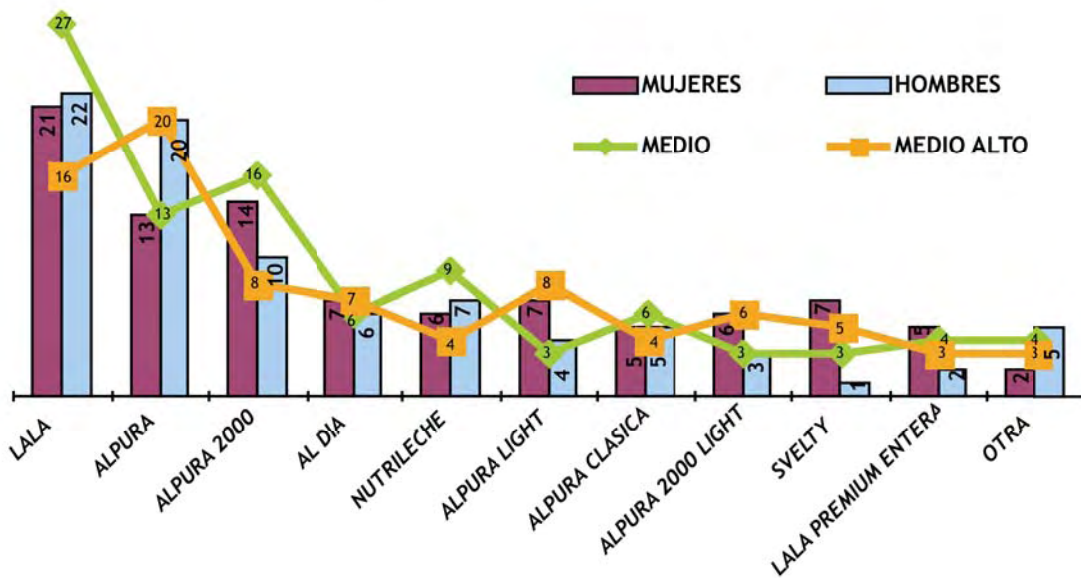


### ANÁLISIS POR COMPONENTES PRINCIPALES LECHE FAVORITA POR GRUPOS DE EDAD Y NSE

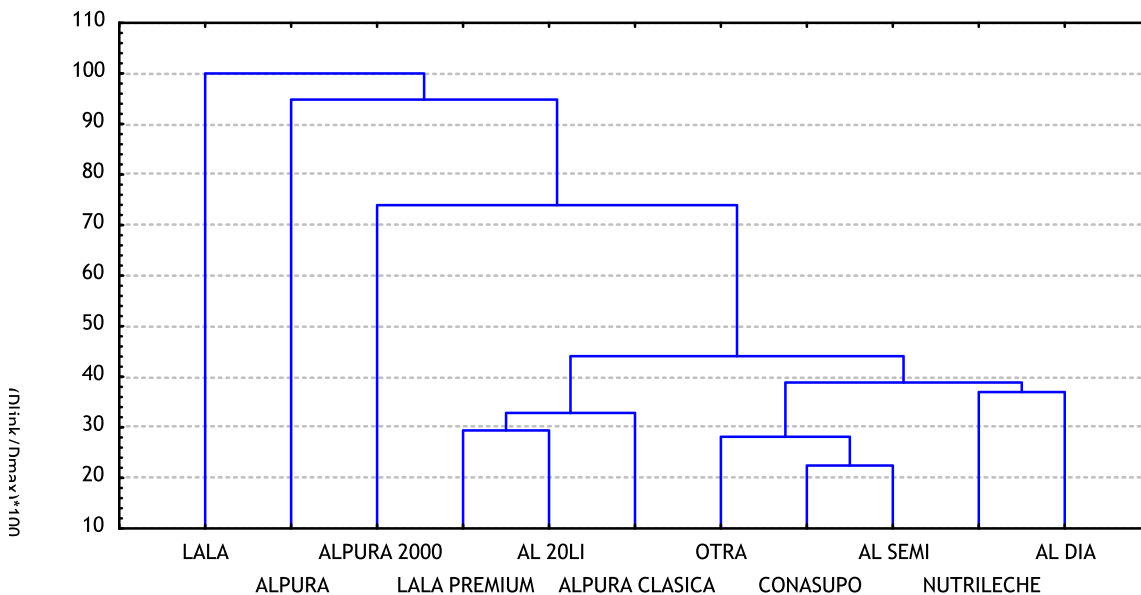


Con respecto a las leches mencionadas, se observa que los Adultos Mayores si bien tienen mayor TOM, son ellos quienes menos consumen, mientras que el grupo de los Adultos Maduros sobre todo en mujeres del NSE Medio Alto son quienes más consumen leche. Los Adultos Jóvenes se encuentran en segundo lugar con respecto al puntaje y de éstos, las mujeres de NSE Medio son quienes más consumen. Los Jóvenes tienen un bajo consumo de leche aunque tienen preferencia por Lala.

MARCA DE LECHE FAVORITA POR SEXO Y NSE



Tree Diagram for 11 Cases  
Single Linkage  
Euclidean distances



Se observa que Lala continúa siendo la preferida en el NSE medio y de los jóvenes, en segundo lugar tenemos a Alpura que es la preferida de los hombres de NSE medio alto y de las mujeres de NSE medio, en tercer lugar está Alpura 2000 con una preferencia homogénea según sexo y NSE.

Nutrileche y Al día son preferidas de igual manera en el NSE medio y medio alto, así como en los dos sexos, después aparece Lala Premium y Alpura 2000 Light con una preferencia muy similar en NSE y sexo, el siguiente es Alpura Semi y Conasupo.

### 3.6.7.5 Marca comprada habitualmente

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA * GRUPOS DE EDADES	113	56.5%	87	43.5%	200	100.0%

#### MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA \* GRUPOS DE EDADES

##### Crosstabulation

Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA	AL DIA	1	1	3	0	5
	ALPURA	2	4	5	4	15
	ALPURA CLASICA	1	1	5	1	8
	ALPURA SEMIDESCREMADA	0	0	1	0	1
	ALPURA LIGHT	0	2	3	2	7
	ALPURA 2000	1	5	8	3	17
	ALPURA 2000 LIGHT	0	2	5	2	9
	ALPURA DESLACTOSADA	1	0	0	0	1
	ALPURA LIGHT DESLACTOSADA	0	0	1	0	1
	ALPURA KIDS	0	1	0	0	1
	CONASUPO/LICONSA	1	1	0	1	3
	LALA	5	7	4	1	17
	LALA LIGHT	0	1	1	0	2
	LALA PREMIUM ENTERA	1	0	5	1	7
	LALA SILUETTE PLUS LIGHT	0	0	0	1	1
	MILECHE	0	1	0	0	1
	NUTRILECHE	3	2	2	0	7
	PARMALAT	0	1	0	1	2
SAN MARCOS	1	0	0	0	1	
SVELTY	0	0	2	1	3	
OTRA	2	1	0	1	4	
Total		19	30	45	19	113

### Symmetric Measures

		Value	Asymp. Std. Error(a)	Approx. T(b)	Approx. Sig.
Interval by Interval	Pearson's R	-.207	.097	-2.225	.028(c)
Ordinal by Ordinal	Spearman Correlation	-.161	.098	-1.722	.088(c)
N of Valid Cases		113			

- a Not assuming the null hypothesis.  
 b Using the asymptotic standard error assuming the null hypothesis.  
 c Based on normal approximation.

Posiciones:

1. Alpura 2000 y Lala con 17 menciones
2. Alpura (15 menciones)
3. Alpura 2000 Light (9 menciones)
4. Alpura Clásica (8 menciones)
5. Alpura Light, Lala Premium Entera y Nutrileche (7 menciones)

Si hay correlación baja ( $r = -0.207, p = 0.028$ ) e inversamente proporcional con respecto al grupo de edad y la marca que compran con mayor frecuencia.

Se observa cuál es la marca comprada con mayor frecuencia con respecto al sexo y al nivel socioeconómico.

### Nivel Medio

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA * SEXO * NIVEL SOCIOECONOMICO	56	56.0%	44	44.0%	100	100.0%

#### MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA \* SEXO \* NIVEL SOCIOECONOMICO

##### Crosstabulation

Count

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO	MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA	ALPURA	5	2	7
		ALPURA CLASICA	1	2	3
		ALPURA LIGHT	3	0	3
		ALPURA 2000	6	5	11
		ALPURA 2000 LIGHT	2	2	4
		CONASUPO/LICONSA	2	1	3
		LALA	6	2	8
		LALA PREMIUM ENTERA	3	0	3
		LALA SILUETTE PLUS LIGHT	1	0	1
		NUTRILECHE	3	2	5

	PARMALAT	2	0	2
	SVELTY	2	1	3
	OTRA	1	2	3
	Total	37	19	56

**Nivel Medio Alto**

**Case Processing Summary**

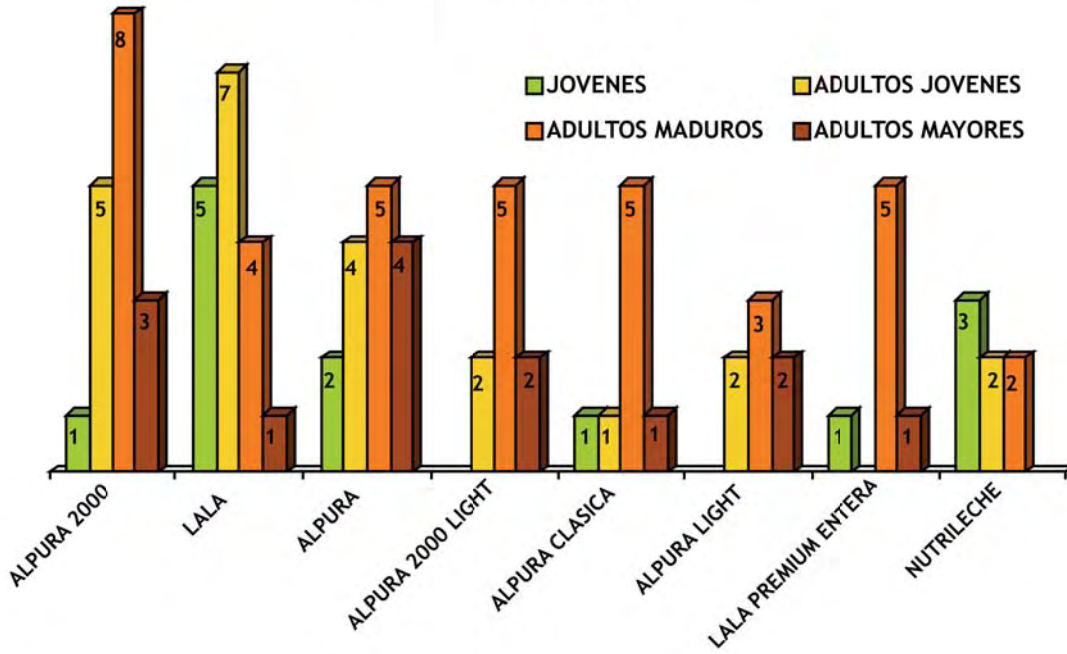
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA * SEXO * NIVEL SOCIOECONOMICO	57	57.0%	43	43.0%	100	100.0%

**MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA \* SEXO \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**

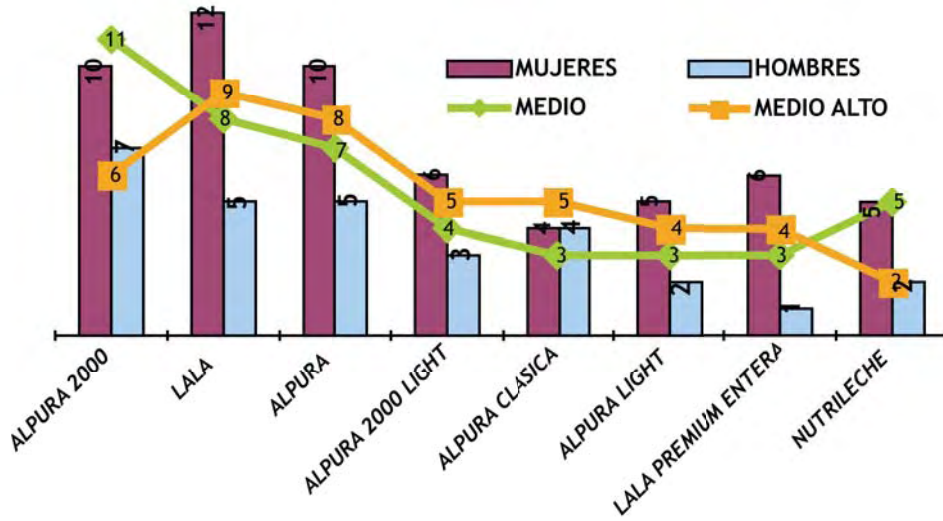
Count

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO ALTO	MARCA DE LECHE QUE COMPRA CON MAYOR FRECUENCIA	AL DIA	4	1	5
		ALPURA	5	3	8
		ALPURA CLASICA	3	2	5
		ALPURA SEMIDESCREMADA	0	1	1
		ALPURA LIGHT	2	2	4
		ALPURA 2000	4	2	6
		ALPURA 2000 LIGHT	4	1	5
		ALPURA DESLACTOSADA	0	1	1
		ALPURA LIGHT DESLACTOSADA	1	0	1
		ALPURA KIDS	1	0	1
		LALA	6	3	9
		LALA LIGHT	0	2	2
		LALA PREMIUM ENTERA	3	1	4
		MILECHE	0	1	1
		NUTRILECHE	2	0	2
		SAN MARCOS	1	0	1
		OTRA	1	0	1
Total			37	20	57

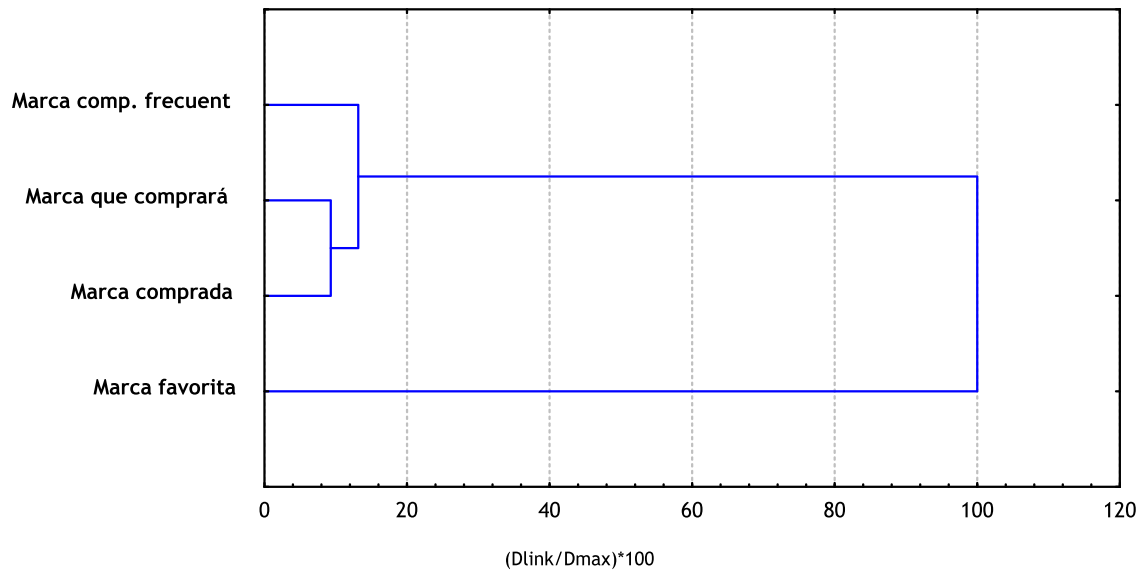
MARCA DE LECHE COMPRADA CON MAYOR FRECUENCIA  
POR GRUPOS DE EDAD



MARCA DE LECHE COMPRADA CON MAYOR FRECUENCIA  
POR SEXO Y NSE



Tree Diagram for 4 Variables  
 Single Linkage  
 Euclidean distances



La marca más frecuentemente comprada y la marca comprada son similares, mientras que difieren un poco la marca favorita y la que comprarán la próxima vez. Se nota que la marca favorita no es la que siempre se compra.

### 3.6.7.6 Tipo de leche de la marca comprada habitualmente

Se observa de qué tipo de leche es la marca comprada con mayor frecuencia con respecto a los grupos de edad.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA * GRUPOS DE EDADES	113	56.5%	87	43.5%	200	100.0%

**DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA \* GRUPOS DE EDADES**  
**Crosstabulation**  
**Count**

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA	LECHE PASTEURIZADA	14	20	16	8	58
	LECHE ULTRA PASTEURIZADA	1	7	24	8	40
	AMBAS	1	0	4	0	5
	NO SABE	3	3	1	3	10
Total		19	30	45	19	113

Ahora se observa de qué tipo es la marca comprada con mayor frecuencia con respecto al sexo y al nivel socioeconómico.

**Nivel Medio**

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA * SEXO * NIVEL SOCIOECONOMICO	56	56.0%	44	44.0%	100	100.0%

**DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA \* SEXO \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**  
**Count**

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO	DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA	LECHE PASTEURIZADA	19	12	31
		LECHE ULTRA PASTEURIZADA	11	6	17
		AMBAS	3	0	3
		NO SABE	4	1	5
Total			37	19	56



Nivel Medio Alto

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA * SEXO * NIVEL SOCIOECONOMICO	57	57.0%	43	43.0%	100	100.0%

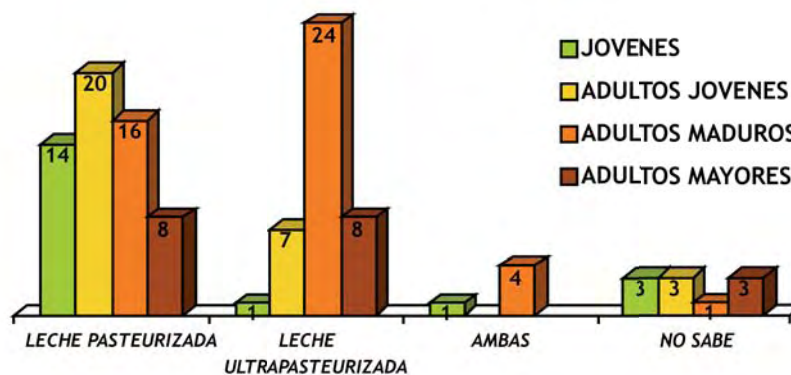
DE QUE TIPO ES LA MARCA DE LECHE QUE SE COMPRA CON MAYOR FRECUENCIA \* SEXO \* NIVEL SOCIOECONOMICO

Crosstabulation

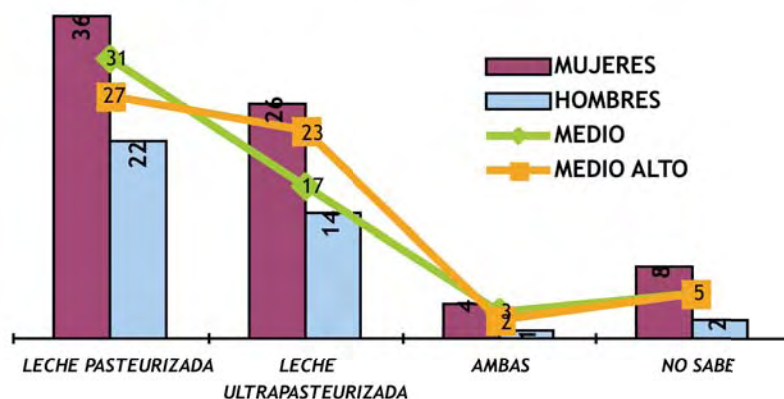
Count

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO ALTO	DE QUE TIPO ES LA MARCA CON MAYOR DE LECHE QUE SE COMPRA FRECUENCIA	LECHE PASTEURIZADA	17	10	27
		LECHE ULTRA PASSTEURIZADA	15	8	23
		AMBAS	1	1	2
		NO SABE	4	1	5
	Total		37	20	57

TIPO DE LECHE DE LA MARCA COMPRADA CON MAYOR FRECUENCIA POR GRUPOS DE EDAD



**TIPO DE LECHE DE LA MARCA COMPRADA CON MAYOR FRECUENCIA POR SEXO Y NSE**



### 3.6.7.7 Marca sustituta

Se observa qué harían las personas si donde acostumbran a realizar sus compras de leche líquida no encontrarán la marca que consumen en su hogar con mayor frecuencia, con respecto a los grupos de edad.

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE HARIA SI DONDE ACOSTUMBRA A COMPRAR LECHE NO ENCONTRARA SU MARCA * GRUPOS DE EDADES	113	56.5%	87	43.5%	200	100.0%

#### QUE HARIA SI DONDE ACOSTUMBRA A COMPRAR LECHE NO ENCONTRARA SU MARCA \* GRUPOS DE EDADES Crosstabulation

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE HARIA SI DONDE ACOSTUMBRA A COMPRAR LECHE NO ENCONTRARA SU MARCA	IRIA A OTRA TIENDA A BUSCAR MI MARCA	14	18	26	7	65
	COMPRARIA LA MISMA MARCA PERO EN ALGUNA OTRA VARIEDAD O TIPO	0	1	2	4	7

	COMPRARIA OTRA MARCA	5	9	15	6	35
	NO COMPRARIA NINGUNA MARCA/REGRESARI A MAS TARDE	0	2	1	2	5
	COMPRARIA OTRO PRODUCTO QUE NO FUESE LECHE	0	0	1	0	1
Total		19	30	45	19	113

La mayoría respondió que iría a otra tienda a buscar su marca (65 menciones de 113) de los cuales 26 menciones fueron en Adultos maduros y 18 en Adultos Jóvenes. Los que respondieron que comprarían otra marca fueron 35 de 113 menciones, de los cuales 15 fueron de Adultos Maduros, 9 Adultos Jóvenes, 6 Adultos Mayores y 5 Jóvenes. Ahora se observa cuál es la marca sustituta que comprarían aquellos que contestaron que comprarían otra marca en caso de no encontrar su marca donde acostumbran a comprar leche líquida.

Se observa qué marca sustituta es la que eligen las mujeres con respecto al grupo de edad.

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA SUSTITUTA * GRUPOS DE EDADES	19	18.3%	85	81.7%	104	100.0%

#### MARCA SUSTITUTA \* GRUPOS DE EDADES Crosstabulation Count

MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
MARCA SUSTITUTA	AL DIA	1	0	0	1	2
	ALPURA	1	2	3	0	6
	ALPURA LIGHT	0	0	1	0	1
	FORTILECHE	1	0	0	0	1
	LALA	1	1	2	0	4
	LALA LIGHT	0	0	1	0	1
	LALA PREMIUM ENTERA	0	1	0	0	1
	NUTRILECHE	0	0	1	0	1
	SVELTY	0	0	1	0	1
	OTRA	0	1	0	0	1
Total		4	5	9	1	19

Se observa qué marca sustituta es la que eligen los hombres con respecto al grupo de edad.

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA SUSTITUTA * GRUPOS DE EDADES	16	16.7%	80	83.3%	96	100.0%

### MARCA SUSTITUTA \* GRUPOS DE EDADES Crosstabulation

Count

HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
MARCA SUSTITUTA	AL DIA	0	1	3	0	4
	ALPURA	0	0	1	1	2
	APLURA CLASICA	0	1	0	0	1
	ALPURA SEMIDESCREMADA	0	0	0	1	1
	ALPURA 2000	0	1	0	1	2
	LALA	1	0	0	2	3
	LALA LIGHT	0	1	0	0	1
	LALA PREMIUM ENTERA	0	0	1	0	1
	MILECHE	0	0	1	0	1
Total		1	4	6	5	16

Ahora se observará cuál es la marca sustituta con respecto al sexo y al nivel socioeconómico.

### Nivel Medio

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA SUSTITUTA * SEXO * NIVEL SOCIOECONOMICO	21	21.0%	79	79.0%	100	100.0%

**MARCA SUSTITUTA \* SEXO \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**  
 Count

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO	MARCA SUSTITUTA	AL DIA	2	4	6
		ALPURA	3	1	4
		APLURA CLASICA	0	1	1
		ALPURA SEMIDESCREMADA	0	1	1
		ALPURA 2000	0	1	1
		FORTILECHE	1	0	1
		LALA	2	0	2
		LALA LIGHT	1	1	2
		LALA PREMIUM ENTERA	0	1	1
		MILECHE	0	1	1
		NUTRILECHE	1	0	1
Total			10	11	21

**Nivel Medio Alto**

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA SUSTITUTA * SEXO * NIVEL SOCIOECONOMICO	14	14.0%	86	86.0%	100	100.0%

**MARCA SUSTITUTA \* SEXO \* NIVEL SOCIOECONOMICO**  
**Crosstabulation**

Count

NIVEL SOCIOECONOMICO			SEXO		Total
			MUJER	HOMBRE	
MEDIO ALTO	MARCA SUSTITUTA	ALPURA	3	1	4
		ALPURA LIGHT	1	0	1
		ALPURA 2000	0	1	1
		LALA	2	3	5
		LALA PREMIUM ENTERA	1	0	1
		SVELTY	1	0	1
		OTRA	1	0	1
		Total			9

### 3.6.7.8 Marca que se comparará la próxima vez

Ahora se observa qué marca se comparará la próxima vez que se compre leche líquida con respecto a los grupos de edad.

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA QUE COMPRARA LA PROXIMA VEZ * GRUPOS DE EDADES	113	56.5%	87	43.5%	200	100.0%

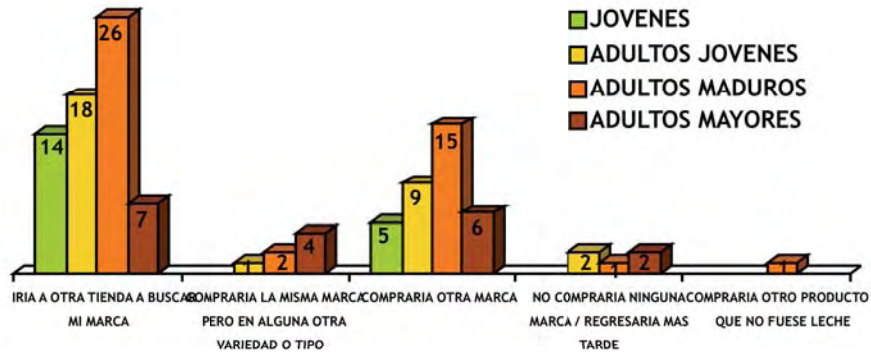
Las principales posiciones quedan de la siguiente manera:

1. Lala 19 menciones de 113
2. Alpura 2000 (15)
3. Alpura (14)
4. Alpura Clásica (8) y Alpura 2000 Light (8)

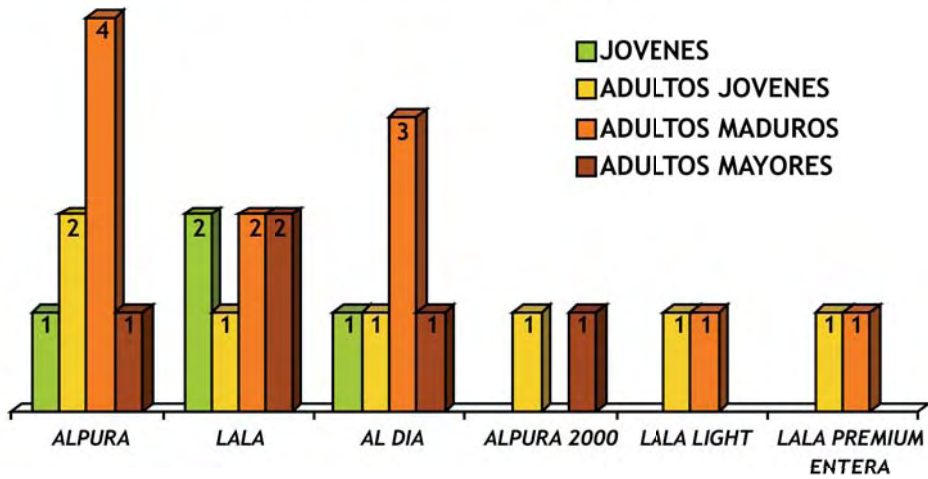
#### MARCA QUE COMPRARA LA PROXIMA VEZ \* GRUPOS DE EDADES Crosstabulation Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
MARCA QUE COMPRARA LA PROXIMA VEZ	AL DIA	1	3	3	0	7
	ALPURA	1	4	5	4	14
	ALPURA CLASICA	1	1	5	1	8
	ALPURA SEMIDESCREMADA	0	0	1	0	1
	ALPURA LIGHT	0	2	3	2	7
	ALPURA 2000	1	4	7	3	15
	ALPURA 2000 LIGHT	0	2	4	2	8
	ALPURA DESLACTOSADA	1	0	0	0	1
	ALPURA LIGHT DESLACTOSADA	0	0	1	0	1
	ALPURA KIDS	0	1	0	0	1
	LALA	6	7	5	1	19
	LALA LIGHT	0	1	1	0	2
	LALA PREMIUM ENTERA	1	0	4	2	7
	LALA SILUETTE PLUS LIGHT	0	0	0	1	1
	MILECHE	0	1	0	0	1
	NUTRILECHE	3	2	2	0	7
	PARMALAT	0	1	0	1	2
	SAN MARCOS	1	0	1	0	2
	SVELTY	1	0	3	1	5
	OTRA	2	1	0	1	4
Total		19	30	45	19	113

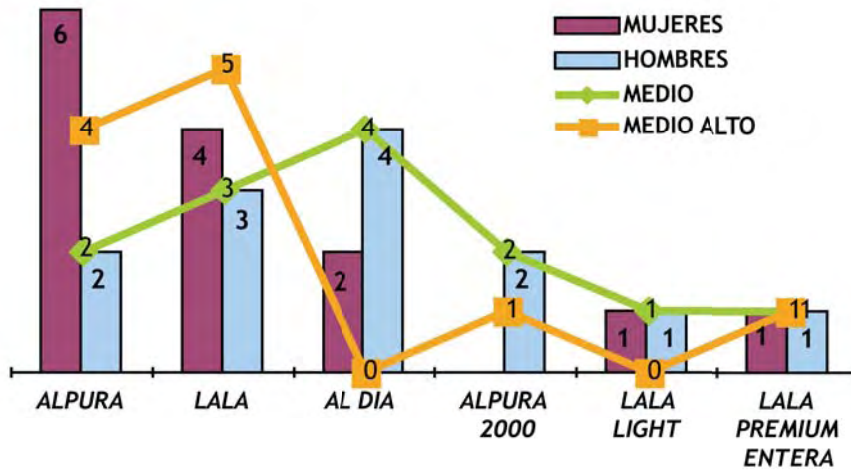
¿QUÉ HARÍA SI DONDE ACOSTUMBRA A COMPRAR LECHE NO ENCONTRARA SU MARCA?



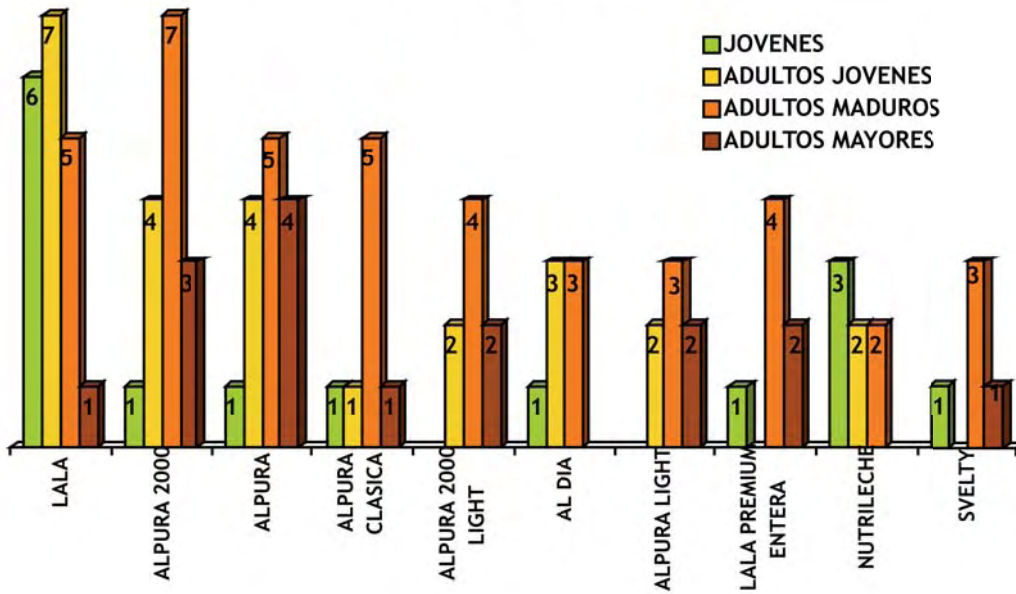
MARCA SUSTITUTA POR GRUPOS DE EDAD



MARCA SUSTITUTA POR SEXO Y NSE



MARCA QUE SE COMPRARÁ LA PRÓXIMA VEZ POR GRUPOS DE EDAD





### 3.6.7.9 Marca que nunca se compraría

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA QUE NUNCA COMPRARIA * GRUPOS DE EDADES	107	53.5%	93	46.5%	200	100.0%

#### MARCA QUE NUNCA COMPRARIA \* GRUPOS DE EDADES Crosstabulation Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
MARCA QUE NUNCA COMPRARIA	AL DIA	0	1	0	1	2
	ALPURA	1	0	0	0	1
	ALPURA LIGHT	1	1	1	1	4
	ALPURA 2000 LIGHT	0	0	0	2	2
	ALPURA 40 Y TANTOS	0	1	1	0	2
	ALPURA DESLACTOSADA	0	1	1	0	2
	ALPURA FRUTAL	0	0	1	0	1
	ALPURA FRUTAL PARA LLEVAR	0	0	1	0	1
	ALPURA EN POLVO	0	0	0	1	1
	BADEN	1	3	4	2	10
	BOREAL	1	3	4	2	10
	CONASUPO/LICONS A	3	6	15	4	28
	DESLY	1	2	9	0	12
	FORTILECHE	2	2	2	1	7
	LALA	2	1	0	0	3
	LALA LIGHT	0	0	2	1	3
	LALA SUPREMA LIGHT	0	1	0	0	1
	MILECHE	3	0	0	1	4
	NUTRILECHE	1	3	2	1	7
	PARMALAT	1	1	1	0	3
SAN MARCOS	0	2	0	0	2	
OTRA	0	1	0	0	1	
Total		17	29	44	17	107

Las principales posiciones son las siguientes:

1. Conasupo/ Liconsa (28)
2. Desly (12)
3. Boreal, Baden (10)
4. Fortileche, Nutrileche (7)

No hay correlación ( $r = -0.180, p = 0.064$ ) entre la marca de leche que nunca compraría con respecto a los grupos de edad de ambos sexos.

### ¿Por Qué Nunca Compraría Esa Marca?

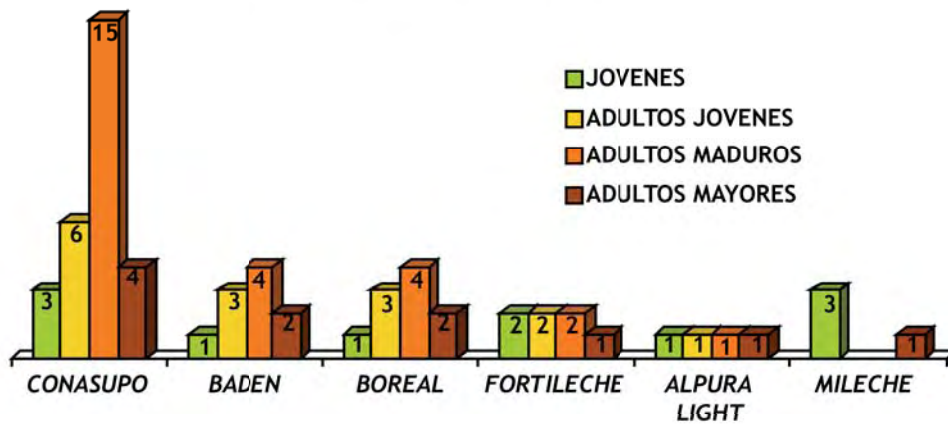
#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PORQUE NUNCA LA COMPRARÍA 1 * GRUPOS DE EDADES	107	53.5%	93	46.5%	200	100.0%
PORQUE NUNCA LA COMPRARÍA 2 * GRUPOS DE EDADES	21	10.5%	179	89.5%	200	100.0%
PORQUE NUNCA LA COMPRARÍA 3 * GRUPOS DE EDADES	1	.5%	199	99.5%	200	100.0%

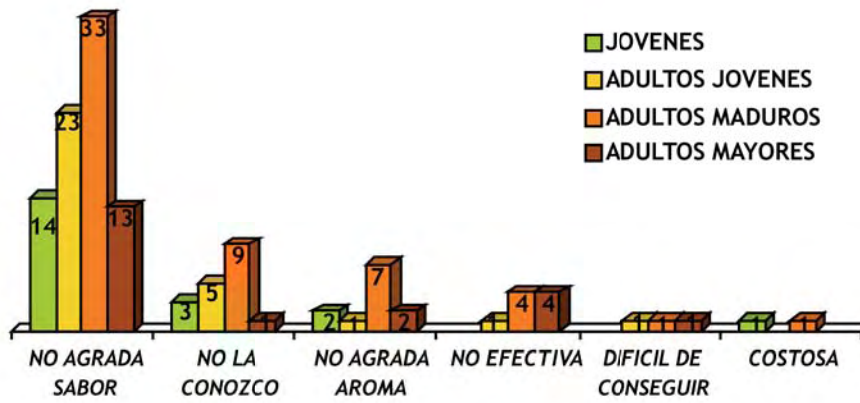
#### Crosstab Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
PORQUE NUNCA LA COMPRARÍA 1	DIFICIL DE CONSEGUIR	0	1	1	1	3
	ENGORDA	0	0	0	1	1
	NO AGRADA EL AROMA	2	1	7	2	12
	NO AGRADA EL SABOR	12	22	26	10	70
	COSTOSA	1	0	0	0	1
	NO EFECTIVA	0	0	1	2	3
	PRECIO	0	0	1	0	1
	NO LA CONOZCO	2	5	8	1	16
Total		17	29	44	17	107

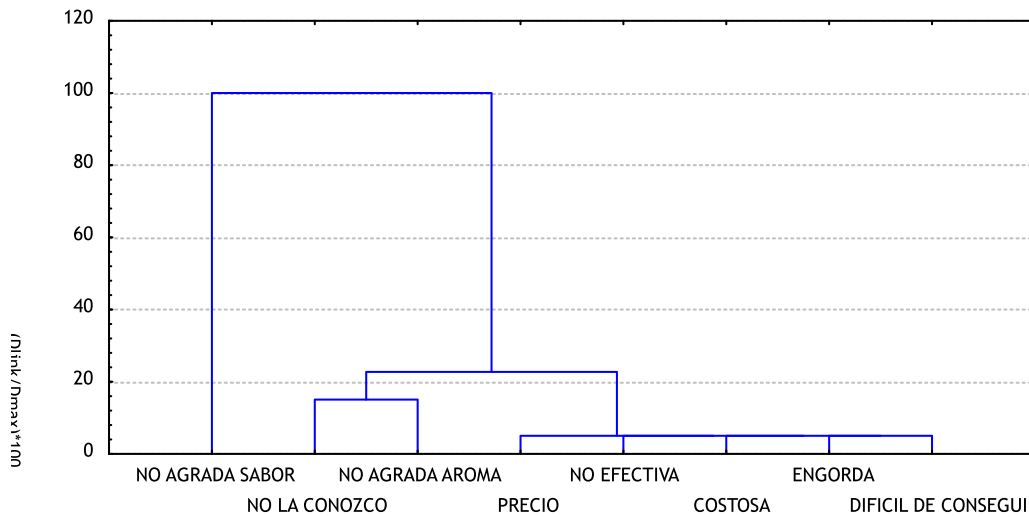
### MARCA QUE NUNCA SE COMPRARÍA POR GRUPOS DE EDAD



### ¿POR QUÉ NUNCA LA COMPRARÍA?



Tree Diagram for 8 Cases  
Single Linkage  
Euclidean distances



La causa principal por la cual nunca comprarían cierto marca de leche es porque no agrada el sabor, después porque no la conocen y en tercer lugar porque no agrada el aroma, los otros causas no son tan determinantes en la decisión de compra.

### 3.6.8 Indicadores de comportamiento de compra

#### 3.6.8.1 Persona que decide la marca a comprar

Se observa quién decide la marca a comprar, se verá primero el caso para las mujeres con respecto a los grupos de edad.

#### Case Processing Summary

MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR * GRUPOS DE EDADES	104	100.0%	0	.0%	104	100.0%

#### PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR \* GRUPOS DE EDADES Crosstabulation Count

MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR	YO MISMO	14	17	32	10	73
	OTRA PERSONA	14	10	4	3	31
Total		28	27	36	13	104

#### Symmetric Measures

		Value	Asymp. Std. Error(a)	Approx. T(b)	Approx. Sig.
Interval by Interval	Pearson's R	-.296	.095	-3.128	.002(c)
Ordinal by Ordinal	Spearman Correlation	-.304	.094	-3.224	.002(c)
N of Valid Cases		104			

- a Not assuming the null hypothesis.
- b Using the asymptotic standard error assuming the null hypothesis.
- c Based on normal approximation.

En general para el caso de las mujeres, ellas son las que deciden que marca de leche comprar (73 menciones de 104). Solo para 31 es otra persona quien toma la decisión de que marca comprar.

Hay una correlación baja ( $r = -0.296, p = 0.002$ ) e inversamente proporcional con respecto a quien toma la decisión de que marca comprar, esto es a medida que pasan al siguiente grupo de edad son ellas quienes deciden que marca de leche comprar.

Ahora se observa si es que hay alguna correlación entre el estado civil con respecto a los grupos de edad para el caso de la mujeres.

### Case Processing Summary

MUJERES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR * GRUPOS DE EDADES * ESTADO CIVIL	104	100.0%	0	.0%	104	100.0%

### PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR \* GRUPOS DE EDADES \* ESTADO CIVIL Crosstabulation Count

ESTADO CIVIL			GRUPOS DE EDADES				Total
			JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
CASADO	PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR	YO MISMO	4	9	19	4	36
		OTRA PERSONA	0	2	2	2	6
	Total		4	11	21	6	42
SOLTERO	PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR	YO MISMO	6	6	8	3	23
		OTRA PERSONA	14	7	1	0	22
	Total		20	13	9	3	45
UNION LIBRE	PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR	YO MISMO	4	2	3		9
		OTRA PERSONA	0	1	1		2
	Total		4	3	4		11
OTRO	PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR	YO MISMO			2	3	5
		OTRA PERSONA			0	1	1
	Total				2	4	6

Como se puede observar las mujeres casadas por lo general son quienes deciden qué marca de leche comprar, en las solteras en el grupo de Jóvenes (14) es otra persona quien decide que marca de leche comprar, esto puede ser porque no son jefe de familia y es su madre quien toma la decisión, en unión libre pasa algo similar al de las casadas ya que son ellas quien deciden que marca de leche comprar.

Ahora se observa el caso de los hombres con respecto al grupo de edad.

### Case Processing Summary

HOMBRES	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR * GRUPOS DE EDADES	96	100.0%	0	.0%	96	100.0%

### PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR \* GRUPOS DE EDADES Crosstabulation

Count

HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR	YO MISMO	5	12	13	9	39
	OTRA PERSONA	13	16	22	6	57
Total		18	28	35	15	96

Aquí ocurre el caso contrario al de las mujeres, por lo general son otras personas quien deciden qué marca comprar.

Ahora se observa si es que hay alguna correlación entre el estado civil con respecto a los grupos de edad para el caso de los hombres.

### Case Processing Summary

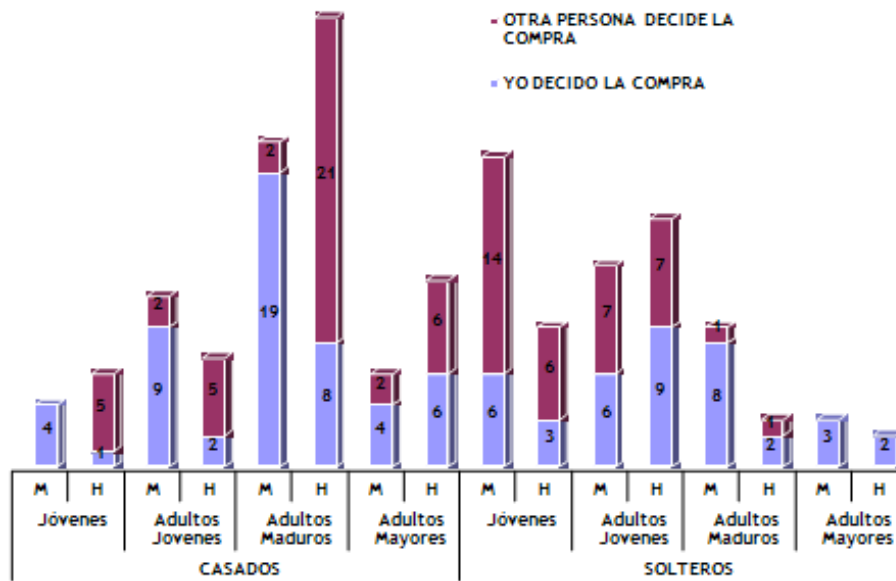
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR * GRUPOS DE EDADES * ESTADO CIVIL	96	100.0%	0	.0%	96	100.0%

**PERSONA QUE DECIDE QUE MARCAS DE LECHE COMPRAR \* GRUPOS DE EDADES \* ESTADO CIVIL**  
**Crosstabulation**  
**Count**

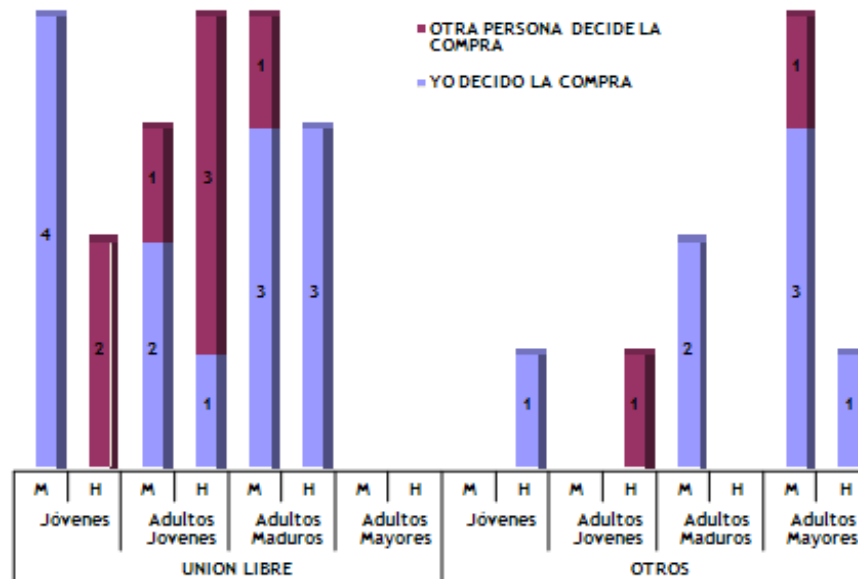
ESTADO CIVIL			GRUPOS DE EDADES				Total
			JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
CASADO	PERSONA QUE DECIDE QUE MARCAS	YO MISMO	1	2	8	6	17
	DE LECHE COMPRAR	OTRA PERSONA	5	5	21	6	37
	Total		6	7	29	12	54
SOLTERO	PERSONA QUE DECIDE QUE MARCAS	YO MISMO	3	9	2	2	16
	DE LECHE COMPRAR	OTRA PERSONA	6	7	1	0	14
	Total		9	16	3	2	30
UNION LIBRE	PERSONA QUE DECIDE QUE MARCAS	YO MISMO		1	3		4
	DE LECHE COMPRAR	OTRA PERSONA	2	3			5
	Total		2	4	3		9
OTRO	PERSONA QUE DECIDE QUE MARCAS	YO MISMO	1			1	2
	DE LECHE COMPRAR	OTRA PERSONA		1			1
	Total		1	1		1	3

Como se puede observar en los hombres casados 37 no deciden qué marca de leche comprar, 21 de los cuales son Adultos Maduros, son sus esposas quienes toman la decisión; en los solteros solo 16 de ellos toman la decisión y 14 lo decide otra persona, el grueso de ellos se encuentra en los Jóvenes y Adultos Jóvenes, quienes aun no son jefe de familia, es por ello que otra persona decide la compra.

### PERSONA QUE DECIDE LA MARCA A COMPRAR



### PERSONA QUE DECIDE LA MARCA A COMPRAR





### 3.6.8.2 Marca comprada en la última semana

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
COMPRO LECHES EN LA ULTIMA SEMANA * GRUPOS DE EDADES	113	56.5%	87	43.5%	200	100.0%

#### COMPRO LECHES EN LA ULTIMA SEMANA \* GRUPOS DE EDADES Crosstabulation Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
COMPRO LECHES EN LA ULTIMA SEMANA	SI	18	29	45	19	111
	NO	1	1	0	0	2
Total		19	30	45	19	113

La mayoría (111 de 113 menciones) compró leche en la ultima semana, el grupo de edad mejor representado fue Adultos Maduros, seguido por Adultos Jóvenes, Jóvenes y Adultos Mayores.

#### ¿Qué Marca Compró?

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARCA QUE COMPRÓ * GRUPOS DE EDADES	111	55.5%	89	44.5%	200	100.0%

**MARCA QUE COMPRÓ \* GRUPOS DE EDADES**  
**Crosstabulation**  
**Count**

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
MARCA QUE COMPRÓ	AL DIA	0	3	2	0	5
	ALPURA	2	4	4	4	14
	ALPURA CLASICA	1	1	5	1	8
	ALPURA SEMIDESCREMADA	0	0	1	0	1
	ALPURA LIGHT	0	2	3	2	7
	ALPURA 2000	1	4	7	3	15
	ALPURA 2000 LIGHT	0	2	5	2	9
	ALPURA DESLACTOSADA	1	0	0	0	1
	ALPURA LIGHT DESLACTOSADA	0	0	1	0	1
	ALPURA KIDS	0	1	0	0	1
	CONASUPO/LICONSA	0	0	0	1	1
	LALA	5	7	6	1	19
	LALA LIGHT	0	1	1	0	2
	LALA PREMIUM ENTERA	1	0	4	1	6
	LALA SILUETTE PLUS LIGHT	0	0	0	1	1
	MILECHE	0	1	1	0	2
	NUTRILECHE	3	2	2	0	7
	PARMALAT	0	0	0	1	1
	SAN MARCOS	1	0	0	0	1
	SVELTY	1	0	3	1	5
OTRA	2	1	0	1	4	
Total		18	29	45	19	111

**Symmetric Measures**

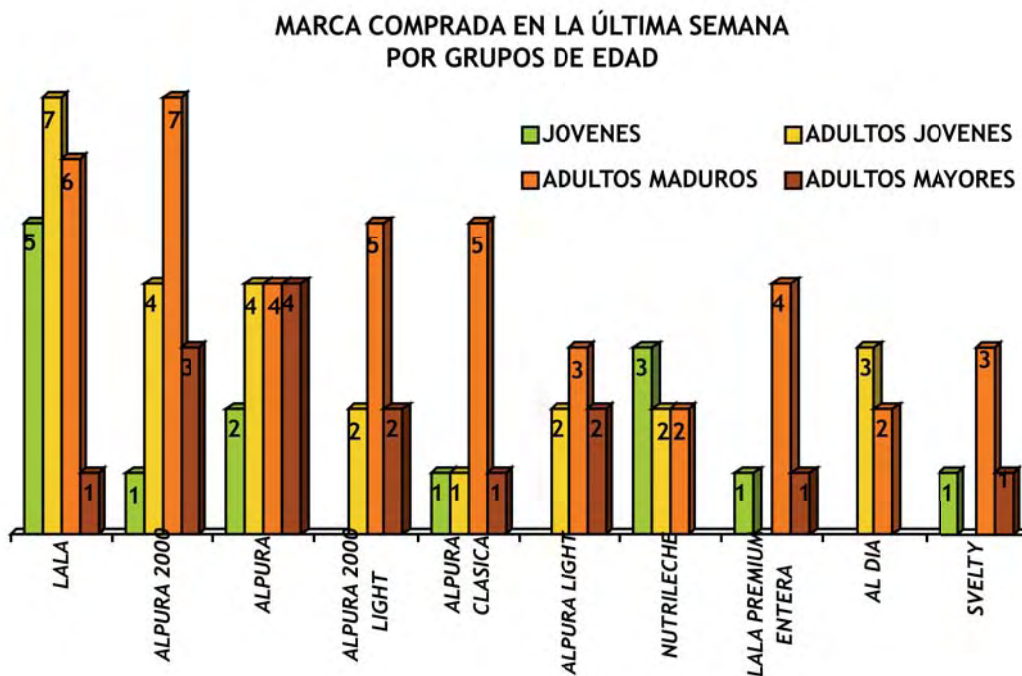
		Value	Asymp. Std. Error(a)	Approx. T(b)	Approx. Sig.
Interval by Interval	Pearson's R	-.198	.097	-2.111	.037(c)
Ordinal by Ordinal	Spearman Correlation	-.144	.099	-1.514	.133(c)
N of Valid Cases		111			

- a Not assuming the null hypothesis.
- b Using the asymptotic standard error assuming the null hypothesis.
- c Based on normal approximation.

Las principales posiciones:

1. Lala con (19 menciones)
2. Alpura 2000 (15 menciones)
3. Alpura (14 menciones)
4. Alpura 2000 Light (9 menciones)
5. Alpura Clásica (8 menciones)
6. Alpura Light y Nutrileche (7 menciones)
7. Lala Premium Entera (6 menciones)
8. Al Día y Svelty (5 menciones)

Si hay una correlación baja ( $r = -0.198, p = 0.037$ ) e inversamente proporcional entre el grupo de edad y la marca de compra, esto es a medida que pasan al siguiente grupo, disminuye la compra de marcas de leche.



### 3.6.8.3 Key drivers en la decisión de compra

Ahora se observa el motivo de compra de las 5 primeras posiciones de las marcas compradas en la última semana. Se observará primero el caso para Lala que tuvo 19 menciones de 111.

TOTAL DE LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	LALA
1	0	1	0	0	NO TIENE CONSERVADORES
3	0	1	1	1	ES NUTRITIVA
4	0	2	2	0	TIENE LECHE DE LA MAS ALTA CALIDAD
13	4	5	4	0	POR TRADICION
11	2	4	5	0	TIENE LA LECHE CON EL MEJOR SABOR
2	0	1	1	0	ES LA MARCA PREFERIDA DE LOS NIÑOS
4	1	1	1	1	ES CREMOSA
3	0	2	1	0	ES MAS NATURAL
5	2	2	1	0	SE CONSIGUE FACIL

Alpura 2000 con 15 menciones de 111:

TOTAL DE LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	ALPURA 2000
2	0	0	1	1	ES NUTRITIVA
7	1	0	4	2	TIENE LECHE DE LA MAS ALTA CALIDAD
7	0	1	4	2	POR TRADICION
1	0	1	0	0	LA RECETO EL MEDICO
1	0	0	1	0	LE DA MAS VALOR POR SU DINERO
3	0	2	1	0	TIENE LA LECHE CON EL MEJOR SABOR
2	1	0	1	0	ES LA MARCA PREFERIDA DE LOS NIÑOS
2	0	1	1	0	ES MAS NATURAL
1	0	0	1	0	TIENE EMPAQUES MAS ATRACTIVOS
1	0	1	0	0	SE CONSIGUE FACIL

Alpura con 14 menciones de 111.

TOTAL DE LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	ALPURA
1	0	0	0	1	ES NUTRITIVA
4	0	2	1	1	TIENE LECHE DE LA MAS ALTA CALIDAD
9	1	2	3	3	POR TRADICION
6	1	1	1	3	TIENE LA LECHE CON EL MEJOR SABOR
4	1	0	1	2	ES CREMOSA
2	0	2	0	0	SE CONSIGUE FACIL

Alpura 2000 Light con 9 menciones de 111:

TOTAL DE LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	ALPURA 2000 LIGHT
1	0	0	0	1	ES NUTRITIVA
4	0	0	3	1	TIENE LECHE DE LA MAS ALTA CALIDAD
5	0	1	3	1	POR TRADICION
4	0	0	4	0	LA RECETO EL MEDICO
3	0	1	0	2	TIENE LA LECHE CON EL MEJOR SABOR
2	0	0	1	1	TIENE LA CONSISTENCIA MAS ADECUADA
2	0	0	1	1	SE CONSIGUE FACIL

Alpura Clásica con 8 menciones de 111:

TOTAL DE LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	ALPURA CLÁSICA
1	0	0	1	0	ES NUTRITIVA
4	0	1	2	1	TIENE LECHE DE LA MAS ALTA CALIDAD
6	1	1	3	1	POR TRADICION
6	1	0	4	1	TIENE LA LECHE CON EL MEJOR SABOR
3	0	0	2	1	ESTA DISPONIBLE EN VARIAS PRESENTACIONES Y TAMAÑOS
1	0	0	0	1	ES CREMOSA
3	0	1	2	0	TIENE LA CONSISTENCIA MAS ADECUADA
2	0	0	2	0	TIENE EMPAQUES MAS ATRACTIVOS

¿Por qué comenzó a utilizar esa marca?

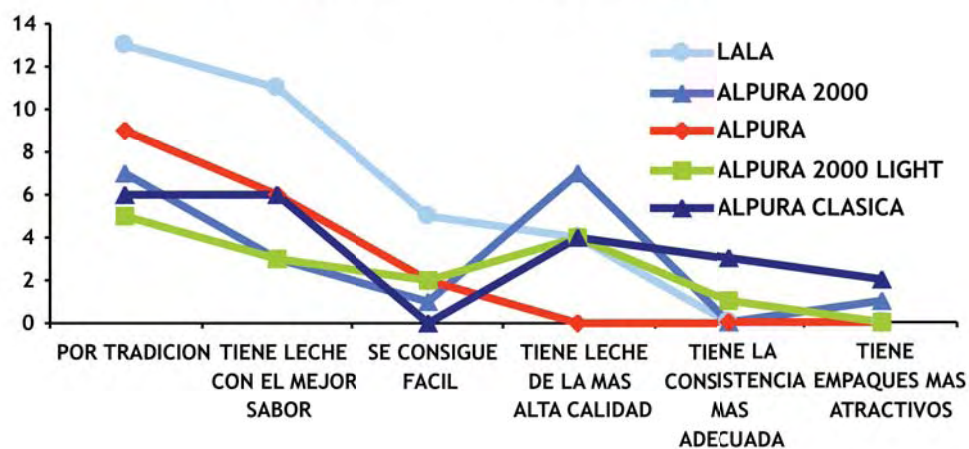
#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PORQUE EMPEZO A UTILIZAR ESTA MARCA DE LECHE 1 * GRUPOS DE EDADES	111	55.5%	89	44.5%	200	100.0%
PORQUE EMPEZO A UTILIZAR ESTA MARCA DE LECHE 2 * GRUPOS DE EDADES	49	24.5%	151	75.5%	200	100.0%
PORQUE EMPEZO A UTILIZAR ESTA MARCA DE LECHE 3 * GRUPOS DE EDADES	3	1.5%	197	98.5%	200	100.0%

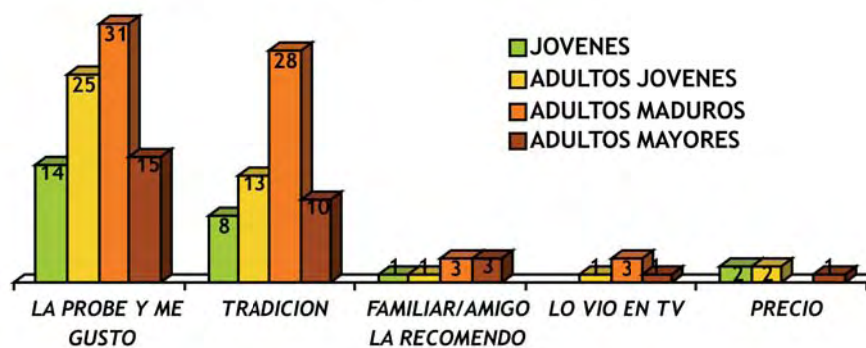
**PORQUE EMPEZO A UTILIZAR ESTA MARCA DE LECHE 1 \* GRUPOS DE EDADES Crosstabulation Count**

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
PORQUE EMPEZO A UTILIZAR ESTA MARCA DE LECHE 1	LA PROBE Y ME GUSTO	14	25	31	15	85
	UN FAMILIAR/AMIGO LA RECOMENDO	0	0	2	2	4
	LO VIO EN TV	0	0	2	0	2
	TRADICION	4	4	10	2	20
Total		18	29	45	19	111

**¿POR QUÉ COMPRO ESA MARCA DE LECHE?**



**¿POR QUÉ COMENZÓ A UTILIZAR ESA MARCA DE LECHE?**



### 3.6.8.4 Frecuencia de consumo del producto

TOTAL DE TODOS LOS GRUPOS DE EDADES Y AMBOS SEXOS.	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MOMENTO EN QUE SE CONSUME LECHE
175	43	46	64	22	DESAYUNO
6	0	4	1	1	MEDIA MAÑANA
5	0	3	1	1	CON LA COMIDA
10	2	5	2	1	EN LA TARDE
51	13	16	10	12	CON LA CENA
122	17	35	52	18	POR LA NOCHE

#### ¿EN QUÉ MOMENTO CONSUMEN LECHE?



### 3.6.8.5 Lugar de compra

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
LUGAR DONDE ACOSTUMBRA COMPRAR MAS FRECUENTEMENTE LECHE 1 * GRUPOS DE EDADES	111	55.5%	89	44.5%	200	100.0%
LUGAR DONDE ACOSTUMBRA COMPRAR MAS FRECUENTEMENTE LECHE 2 * GRUPOS DE EDADES	26	13.0%	174	87.0%	200	100.0%
LUGAR DONDE ACOSTUMBRA COMPRAR MAS FRECUENTEMENTE LECHE 3 * GRUPOS DE EDADES	1	.5%	199	99.5%	200	100.0%
LUGAR DONDE ACOSTUMBRA COMPRAR MAS FRECUENTEMENTE LECHE 4 * GRUPOS DE EDADES	1	.5%	199	99.5%	200	100.0%

#### LUGAR DONDE ACOSTUMBRA COMPRAR MAS FRECUENTEMENTE LECHE 1 \* GRUPOS DE EDADES

#### Crosstabulation Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
LUGAR DONDE ACOSTUMBRA COMPRAR MAS FRECUENTEMENTE LECHE 1	TIENDAS DE MEMBRESIA	2	4	20	1	27
	SUPERMERCADOS	9	20	22	11	62
	TIENDA DE LA ESQUINA	7	5	3	7	22
Total		18	29	45	19	111



### 3.6.8.6 Frecuencia de compra

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
FRECUENCIA CON QUE COMPRA LECHE * GRUPOS DE EDADES	110	55.0%	90	45.0%	200	100.0%

#### FRECUENCIA CON QUE COMPRA LECHE \* GRUPOS DE EDADES Crosstabulation Count

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
FRECUENCIA CON QUE COMPRA LECHE	DIARIO	2	2	3	3	10
	2 O 3 VECES A LA SEMANA	4	7	3	4	18
	SEMANALMENTE	5	5	10	1	21
	CADA 10 DIAS	0	1	1	0	2
	CADA 15 DIAS	5	10	21	4	40
	CADA 20 DIAS	0	3	2	4	9
	MENSUALMENTE	2	1	4	3	10
Total		18	29	44	19	110

Ahora se observa la facilidad o dificultad para encontrar mi marca en tiendas de membresía, supermercados, supercito y tienda de la esquina.

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN TIENDAS DE MEMBRESIA * GRUPOS DE EDADES	28	14.0%	172	86.0%	200	100.0%
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN SUPERMERCADO * GRUPOS DE EDADES	83	41.5%	117	58.5%	200	100.0%

FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN SUPERCITO * GRUPOS DE EDADES	1	.5%	199	99.5%	200	100.0%
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN TIENDAS DE LA ESQUINA * GRUPOS DE EDADES	27	13.5%	173	86.5%	200	100.0%

**FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN TIENDAS DE MEMBRESIA \* GRUPOS DE EDADES**  
**Crosstabulation**  
**Count**

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN TIENDAS DE MEMBRESIA	SE ENCUENTRA EN TODOS LADOS	2	4	16	2	24
	SE ENCUENTRA EN LA MAYORIA DE LOS LUGARES	0	0	4	0	4
Total		2	4	20	2	28

**FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN SUPERMERCADO \* GRUPOS DE EDADES**  
**Crosstabulation**  
**Count**

		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN SUPERMERCADO	SE ENCUENTRA EN TODOS LADOS	7	14	37	10	68
	SE ENCUENTRA EN LA MAYORIA DE LOS LUGARES	3	7	2	2	14
	SE ENCUENTRA EN POCOS LUGARES	0	1	0	0	1
Total		10	22	39	12	83

### FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN SUPERCITO \* GRUPOS DE EDADES

Crosstabulation  
Count

		GRUPOS DE EDADES	
		ADULTOS MAYORES	Total
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN SUPERCITO	SE ENCUENTRA EN POCOS LUGARES	1	1
Total		1	1

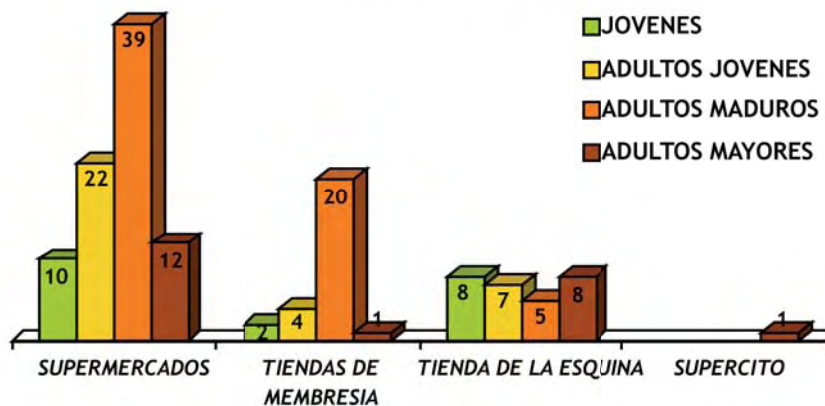
### FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN TIENDAS DE LA ESQUINA \* GRUPOS DE EDADES

Crosstabulation  
Count

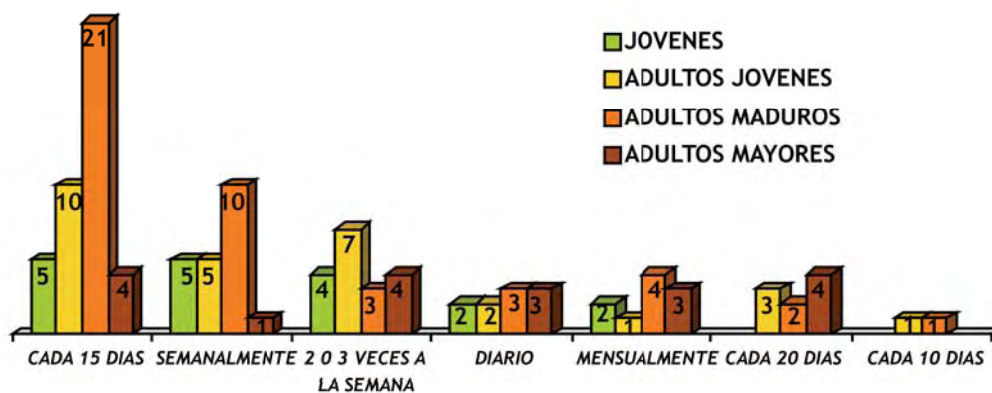
		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
FACILIDAD O DIFICULTAD PARA ENCONTRAR LECHE EN TIENDAS DE LA ESQUINA	SE ENCUENTRA EN TODOS LADOS	7	5	5	6	23
	SE ENCUENTRA EN LA MAYORIA DE LOS LUGARES	1	2	0	0	3
	SE ENCUENTRA EN POCOS LUGARES	0	0	0	1	1
Total		8	7	5	7	27

Las marcas de leche que prefieren las personas se encuentran principalmente en supermercados (68 de 83 menciones), en tiendas de membresía (24 de 28 menciones), incluso en la tiendita de la esquina con 23 de 27 menciones.

### LUGAR DE COMPRA



### FRECUENCIA DE COMPRA

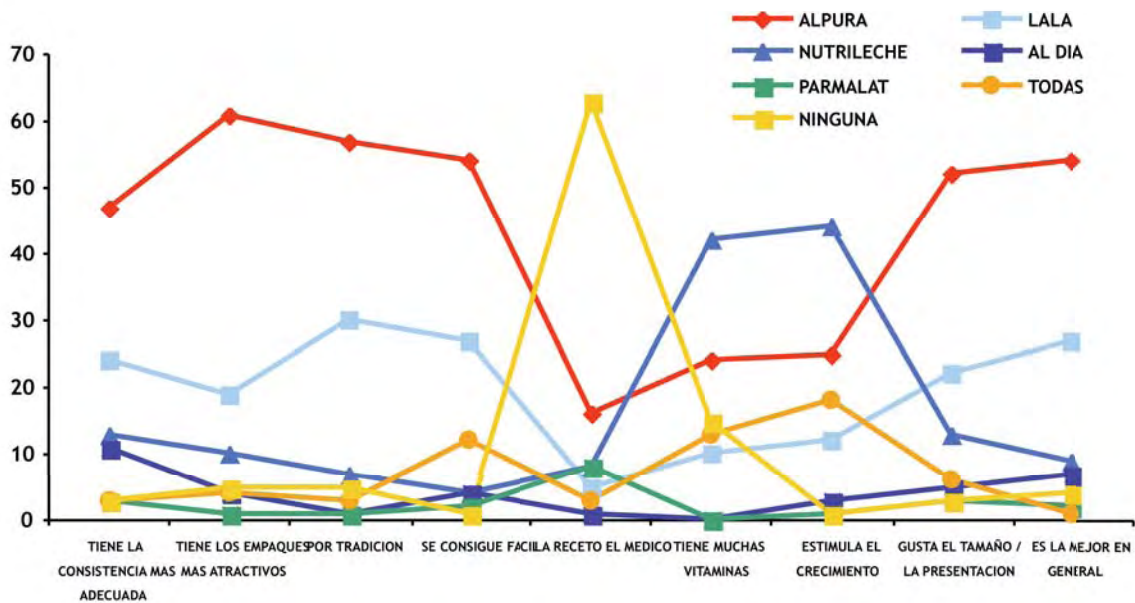
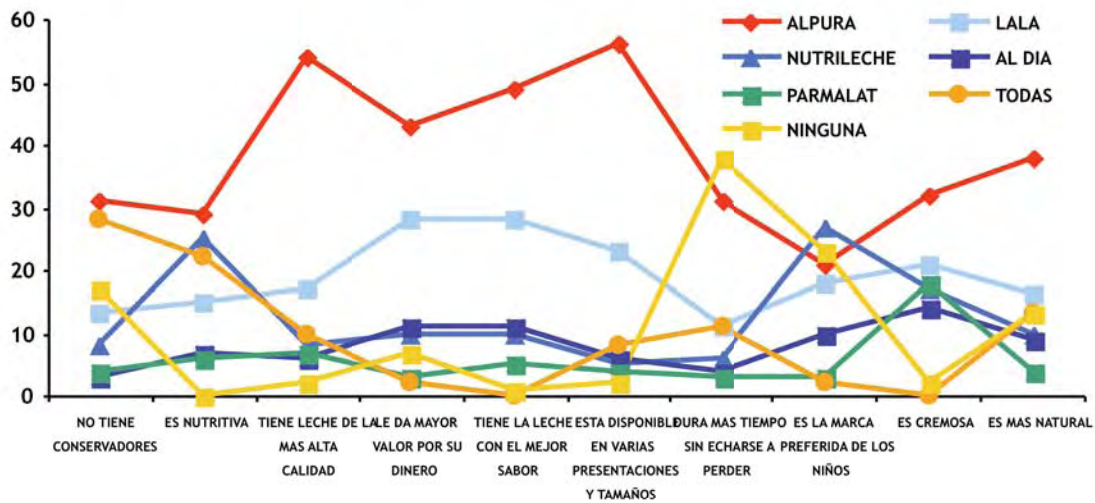


### 3.6.8.7 Frases asociadas a las marcas

FRASES ASOCIADAS	MARCAS						
	ALPURA	LALA	NUTRILECHE	AL DIA	PARMALAT	TODAS	NINGUNA
NO TIENE CONSERVADORES	31	13	8	3	4	28	17
ES NUTRITIVA	29	15	25	7	6	22	0
TIENE LECHE DE LA MAS ALTA CALIDAD	54	17	8	6	7	10	2
LE DA MAYOR VALOR POR SU DINERO	43	28	10	11	3	2	7
TIENE LA LECHE CON EL MEJOR SABOR	49	28	10	11	5	0	1
ESTA DISPONIBLE EN VARIAS PRESENTACIONES Y TAMAÑOS	56	23	5	6	4	8	2
DURA MAS TIEMPO SIN ECHARSE A PERDER	31	11	6	4	3	11	38
ES LA MARCA PREFERIDA DE LOS NIÑOS	21	18	27	10	3	2	23
ES CREMOSA	32	21	17	14	18	0	2
ES MAS NATURAL	38	16	10	9	4	13	13
TIENE LA CONSISTENCIA MAS ADECUADA	47	24	13	11	3	3	3
TIENE LOS EMPAQUES MAS ATRACTIVOS	61	19	10	4	1	4	5
POR TRADICION	57	30	7	1	1	3	5

SE CONSIGUE FACIL	54	27	4	4	2	12	1
LA RECETO EL MEDICO	16	5	8	1	8	3	63
TIENE MUCHAS VITAMINAS	24	10	42	0	0	13	15
ESTIMULA EL CRECIMIENTO	25	12	44	3	1	18	1
GUSTA EL TAMAÑO / LA PRESENTACION	52	22	13	5	3	6	3
ES LA MEJOR EN GENERAL	54	27	9	7	2	1	4

### FRASES ASOCIADAS A LAS MARCAS



### 3.6.9 Sección Alpura

#### 3.6.9.1 Publicidad de leches especializadas Alpura

Se observa en donde se vio o escuchó publicidad de la marca Alpura para las leches especializadas (Alpura 40 y tantos, Alpura 2000 sin colesterol, Alpura 2000 Light Extra, Alpura deslactosada), en mujeres con respecto a los grupos de edad.

TOTAL DE TODOS LOS GRUPOS DE EDADES	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MUJERES
62	20	15	22	5	TELEVISION ABIERTA
2	2	0	0	0	RADIO
8	1	3	4	0	REVISTAS
11	5	3	2	1	TIENDA
2	1	0	0	1	PARABUSES
15	5	4	4	2	MEDIOS DE TRANSPORTE/CAMIONES REPARTIDORES
18	6	3	8	1	ESPECTACULARES
1	0	0	1	0	EVENTOS
2	0	0	1	1	FOLLETO
2	0	1	1	0	TELEVISION DE PAGA
19	4	3	10	2	ENVASES

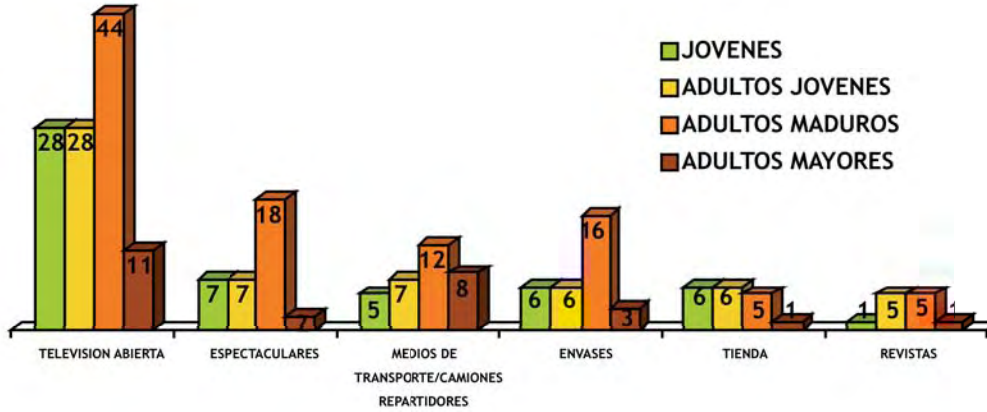
Posiciones: Televisión Abierta (62 menciones), Envases (19 menciones), Espectaculares (18 menciones), Medios de Transporte /Camiones Repartidores (15 menciones), Tienda (11 menciones).

Ahora se observa en donde se vio o escuchó publicidad de la marca Alpura para las leches especializadas (Alpura 40 y tantos, Alpura 2000 sin colesterol, Alpura 2000 Light Extra, Alpura deslactosada), en hombres con respecto a los grupos de edad.

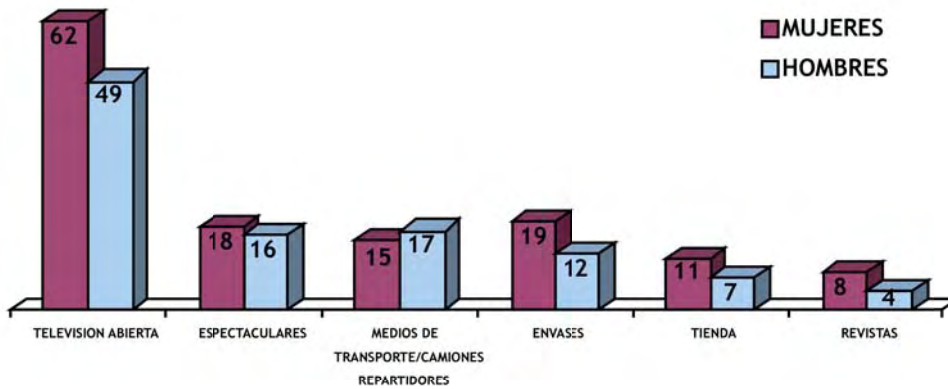
TOTAL DE TODOS LOS GRUPOS DE EDADES	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	HOMBRES
49	8	13	22	6	TELEVISION ABIERTA
2	0	1	1	0	RADIO
1	0	0	0	1	PRENSA
4	0	2	1	1	REVISTAS
7	1	3	0	3	TIENDA
1	0	1	0	0	VALLAS
2	0	1	1	0	PARABUSES
17	0	3	8	6	MEDIOS DE TRANSPORTE/CAMIONES REPARTIDORES
16	1	4	10	1	ESPECTACULARES
1	0	1	0	0	FOLLETO
2	0	1	1	0	TELEVISION DE PAGA
12	2	3	6	1	ENVASES

Posiciones: Televisión Abierta (49 menciones), Medios de Transporte /Camiones Repartidores (17 menciones), Espectaculares (16 menciones), Envases (12 menciones),

¿EN DONDE VIO O ESCUCHÓ PUBLICIDAD DE LAS LECHES ESPECIALIZADAS ALPURA?  
POR GRUPOS DE EDAD



¿EN DONDE VIO O ESCUCHÓ PUBLICIDAD DE LAS LECHES ESPECIALIZADAS ALPURA?  
MUJERES VS HOMBRES



¿Qué recuerda haber visto, leído o escuchado de la publicidad o artículos que hablen de las leches especializadas Alpura?

Se observa primero para el caso de las mujeres por medio de una tabla de frecuencias.

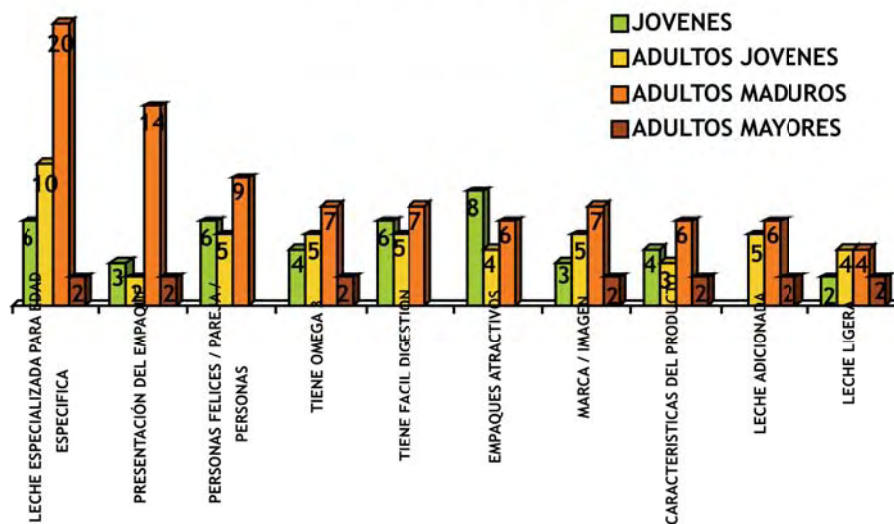
TOTAL DE TODOS LOS GRUPOS DE EDADES	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	MUJERES
11	3	3	4	1	TIENE OMEGA 3
1	1	0	0	0	ADICIONADA CON FIBRA Y ACIDO FOLICO
9	5	2	2	0	TIENE FACIL DIGESTION
10	1	0	8	1	PRESENTACIÓN DEL EMPAQUE
11	3	3	4	1	MARCA / IMAGEN
10	3	2	4	1	CARACTERISTICAS DEL PRODUCTO
11	4	2	5	0	PERSONAS FELICES / PAREJA / PERSONAS
16	4	4	7	1	LECHE ESPECIALIZADA PARA EDAD ESPECIFICA
2	0	1	0	1	LECHE DELICIOSA / RICA
9	0	4	4	1	LECHE ADICIONADA
15	7	2	6	0	EMPAQUES ATRACTIVOS
3	1	2	0	0	ADECUADA PARA PERSONAS INTOLERANTES A LA LACTOSA / DESLACT.
7	2	2	3	0	LECHE LIGERA
6	1	3	0	2	AYUDA AL CRECIMIENTO
2	2	0	0	0	CUIDA EL COLESTEROL / CUIDA EL CORAZON
1	0	1	0	0	CALIDAD ALPURA, CONFIANZA PURA

Ahora se observa para el caso de los hombres.

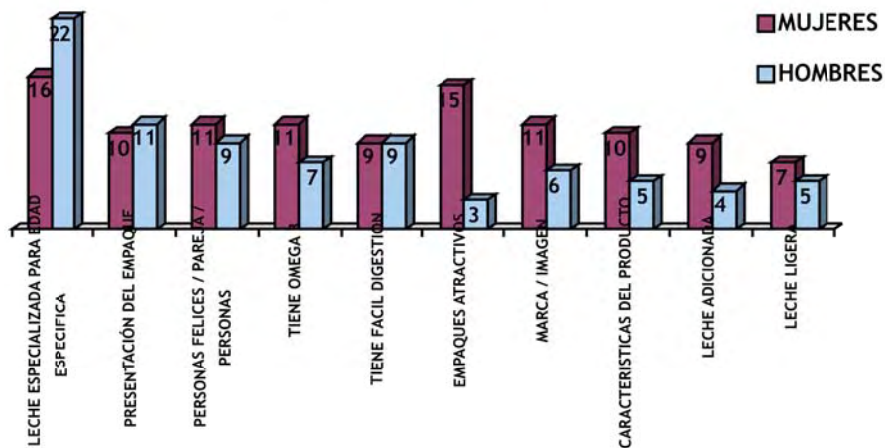
TOTAL DE TODOS LOS GRUPOS DE EDADES	SUMA DE JOVENES	SUMA DE ADULTOS JOVENES	SUMA DE ADULTOS MADUROS	SUMA DE ADULTOS MAYORES	HOMBRES
7	1	2	3	1	TIENE OMEGA 3
1	0	0	1	0	ADICIONADA CON FIBRA Y ACIDO FOLICO
9	1	3	5	0	TIENE FACIL DIGESTION
11	2	2	6	1	PRESENTACIÓN DEL EMPAQUE
6	0	2	3	1	MARCA / IMAGEN
5	1	1	2	1	CARACTERISTICAS DEL PRODUCTO
9	2	3	4	0	PERSONAS FELICES / PAREJA / PERSONAS
22	2	6	13	1	LECHE ESPECIALIZADA PARA EDAD ESPECIFICA
2	0	1	1	0	LECHE DELICIOSA / RICA
4	0	1	2	1	LECHE ADICIONADA
3	1	2	0	0	EMPAQUES ATRACTIVOS
2	0	1	0	1	ADECUADA PARA PERSONAS INTOLERANTES A LA LACTOSA / DESLACT.
5	0	2	1	2	LECHE LIGERA
3	0	1	0	2	AYUDA AL CRECIMIENTO
2	0	0	1	1	CUIDA EL COLESTEROL / CUIDA EL CORAZON
2	0	0	1	1	CALIDAD ALPURA, CONFIANZA PURA



¿QUÉ RECUERDA HABER VISTO, LEÍDO O ESCUCHADO DE LA PUBLICIDAD O ARTÍCULOS QUE HABLEN DE LAS LECHEES ESPECIALIZADAS ALPURA?  
POR GRUPOS DE EDAD



¿QUÉ RECUERDA HABER VISTO, LEÍDO O ESCUCHADO DE LA PUBLICIDAD O ARTÍCULOS QUE HABLEN DE LAS LECHEES ESPECIALIZADAS ALPURA?  
MUJERES VS HOMBRES



### 3.6.9.2 Recordación slogan o frase publicitaria Alpura

Se observa primero para el caso de las mujeres, si recuerdan o no el slogan y qué es lo que recuerdan del slogan.

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SLOGAN O FRASE PUBLICITARIA QUE RECUERDA * GRUPOS DE EDADES	24	96.0%	1	4.0%	25	100.0%

**SLOGAN O FRASE PUBLICITARIA QUE RECUERDA \* GRUPOS DE EDADES Crosstabulation Count**

MUJERES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
SLOGAN O FRASE PUBLICITARIA QUE RECUERDA	MUY RICA	0	0	8	3	11
	VACA / VAQUITA	5	2	5	1	13
Total		5	2	13	4	24

Quienes recuerdan el slogan: “Vaca/vaquita” (13 de 24 menciones) y “Muy Rica” (11).

Ahora se observa el caso de los hombres, si recuerdan o no el slogan y qué es lo que recuerdan del slogan.

**Case Processing Summary**

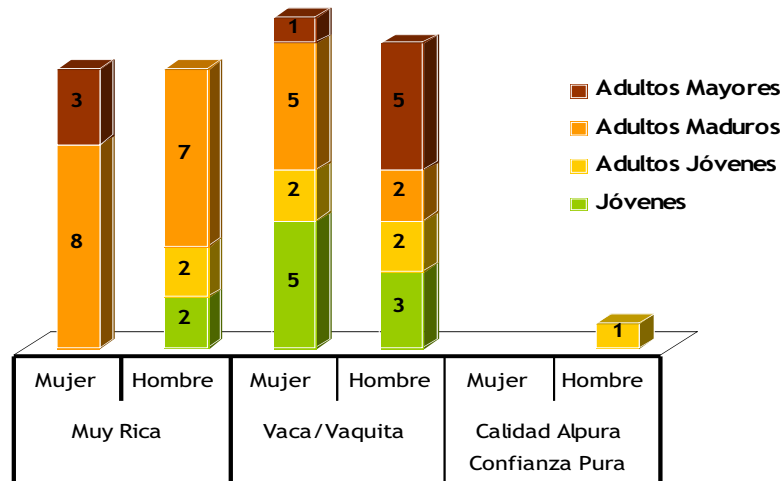
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SLOGAN O FRASE PUBLICITARIA QUE RECUERDA * GRUPOS DE EDADES	24	82.8%	5	17.2%	29	100.0%

**SLOGAN O FRASE PUBLICITARIA QUE RECUERDA \* GRUPOS DE EDADES Crosstabulation Count**

HOMBRES		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
SLOGAN O FRASE PUBLICITARIA QUE RECUERDA	MUY RICA	2	2	7	0	11
	VACA / VAQUITA	3	2	2	5	12
	CALIDAD ALPURA CONFIANZA PURA	0	1	0	0	1
Total		5	5	9	5	24

Quienes recuerdan el slogan: “Muy Rica” (12 de 24 menciones) y “Vaca/vaquita” (11), solo una persona recordó el slogan de Alpura “Calidad Alpura, confianza pura”

### RECORDACIÓN SLOGAN O FRASE PUBLICITARIA ALPURA



### 3.6.9.3 Comerciales de televisión

#### Comercial 1

A los encuestados se les leyó la descripción del comercial de Alpura 40 y tantos, observaremos de las personas que recordaron el comercial qué tanto les gustó, para ambos sexos por grupos de edad.

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TANTO LE GUSTA COMERCIAL 1 * GRUPOS DE EDADES	65	100.0%	0	.0%	65	100.0%

**QUE TANTO LE GUSTA COMERCIAL 1 \* GRUPOS DE EDADES**

**Crosstabulation  
Count**

COMERCIAL 1		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TANTO LE GUSTA COMERCIAL 1	NADA	0	3	0	0	3
	MUY POCO	2	1	2	1	6
	REGULAR	3	7	13	5	28
	MUCHO	5	6	12	3	26
	MUCHISIMO	1	0	1	0	2
Total		11	17	28	9	65

¿Qué cree que están tratando de comunicarle en este comercial?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TRATAN DE TRANSMITIR CON COMERCIAL 1A * GRUPOS DE EDADES	65	32.5%	135	67.5%	200	100.0%
QUE TRATAN DE TRANSMITIR CON COMERCIAL 1B * GRUPOS DE EDADES	34	17.0%	166	83.0%	200	100.0%

**Crosstab  
Count**

COMERCIAL 1		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TRATAN DE TRANSMITIR CON COMERCIAL 1A	AYUDA A MEJORAR LA CONDICION FISICA	2	1	7	0	10
	LECHE NUTRITIVA / BUENA	1	2	4	2	9
	CONTRIBUYE AL BIENESTAR Y A LA BUENA SALUD	3	6	5	4	18
	LECHE ESPECIALIZADA PARA EDAD ESPECIFICA	3	7	8	2	20
	ES BUENA PARA LOS NIÑOS	0	0	2	0	2
	LECHE ADICIONADA	1	0	1	0	2
	TOMAR LECHE NO IMPORTANDO LA EDAD	1	1	1	1	4
Total		11	17	28	9	65

**Crosstab  
Count**

COMERCIAL 1		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TRATAN DE TRANSMITIR CON COMERCIAL 1B	LECHE NUTRITIVA / BUENA	0	1	0	0	1
	CONTRIBUYE AL BIENESTAR Y A LA BUENA SALUD	2	1	11	2	16
	LECHE ESPECIALIZADA PARA EDAD ESPECIFICA	0	2	6	1	9
	LECHE ADICIONADA	0	2	3	0	5
	TOMAR LECHE NO IMPORTANDO LA EDAD	1	1	1	0	3
Total		3	7	21	3	34

**Comercial 2**

A los encuestados se les leyó la descripción del comercial de Alpura 2000 Light Extra y Alpura 2000 sin colesterol, observaremos de las personas que recordaron el comercial qué tanto les gusta, para ambos sexos por grupos de edad.

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TANTO LE GUSTA COMERCIAL 2 * GRUPOS DE EDADES	43	100.0%	0	.0%	43	100.0%

**QUE TANTO LE GUSTA COMERCIAL 2 \* GRUPOS DE EDADES**

**Crosstabulation  
Count**

COMERCIAL 2		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TANTO LE GUSTA COMERCIAL 2	MUY POCO	1	2	1	3	7
	REGULAR	8	6	12	2	28
	MUCHO	0	3	3	0	6
	MUCHISIMO	0	1	0	1	2
Total		9	12	16	6	43

¿Qué cree que están tratando de comunicarle en este comercial?

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TRATAN DE TRANSMITIR CON COMERCIAL 2A * GRUPOS DE EDADES	43	21.5%	157	78.5%	200	100.0%
QUE TRATAN DE TRANSMITIR CON COMERCIAL 2B * GRUPOS DE EDADES	19	9.5%	181	90.5%	200	100.0%

### Crosstab Count

COMERCIAL 2		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TRATAN DE TRANSMITIR CON COMERCIAL 2A	CUIDA TU FIGURA	0	1	2	1	4
	NUTRICION	1	1	1	1	4
	LECHE ESPECIALIZADA PARA UNA ETAPA DETERMINADA	2	3	0	0	5
	VALORES FAMILIARES	0	1	0	0	1
	LECHE ADICIONADA CON VITAMINAS / FIBRA	2	1	6	0	9
	LECHE NUTRITIVA / BUENA	3	1	4	1	9
	APORTE DE CALCIO PARA FORTALECER HUESOS	0	0	0	1	1
	CONTRIBUYE AL BIENESTAR Y BUENA SALUD	1	4	3	2	10
Total		9	12	16	6	43

### Crosstab Count

COMERCIAL 2		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TRATAN DE TRANSMITIR CON COMERCIAL 2B	CUIDA TU FIGURA	1	0	1	1	3
	NUTRICION	0	1	1	0	2
	LECHE ADICIONADA CON VITAMINAS / FIBRA	1	1	0	0	2

	LECHE NUTRITIVA / BUENA	0	1	1	0	2
	APORTE DE CALCIO PARA FORTALECER HUESOS	0	0	2	0	2
	CONTRIBUYE AL BIENESTAR Y BUENA SALUD	2	1	4	1	8
Total		4	4	9	2	19

### Comercial 3

A los encuestados se les leyó la descripción del comercial de Alpura Deslactosada, observaremos de las personas que recordaron el comercial qué tanto les gusta, para ambos sexos por grupos de edad.

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TANTO LE GUSTA COMERCIAL 3 * GRUPOS DE EDADES	54	96.4%	2	3.6%	56	100.0%

#### QUE TANTO LE GUSTA COMERCIAL 3 \* GRUPOS DE EDADES

##### Crosstabulation

Count

COMERCIAL 3		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TANTO LE GUSTA COMERCIAL 3	NADA	0	1	0	0	1
	MUY POCO	1	0	0	1	2
	REGULAR	6	8	13	1	28
	MUCHO	4	9	4	3	20
	MUCHISIMO	0	0	2	1	3
Total		11	18	19	6	54

¿Qué cree que están tratando de comunicarle en este comercial?

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TRATAN DE TRANSMITIR CON COMERCIAL 3A * GRUPOS DE EDADES	56	28.0%	144	72.0%	200	100.0%
QUE TRATAN DE TRANSMITIR CON COMERCIAL 3B * GRUPOS DE EDADES	13	6.5%	187	93.5%	200	100.0%

**Crosstab  
Count**

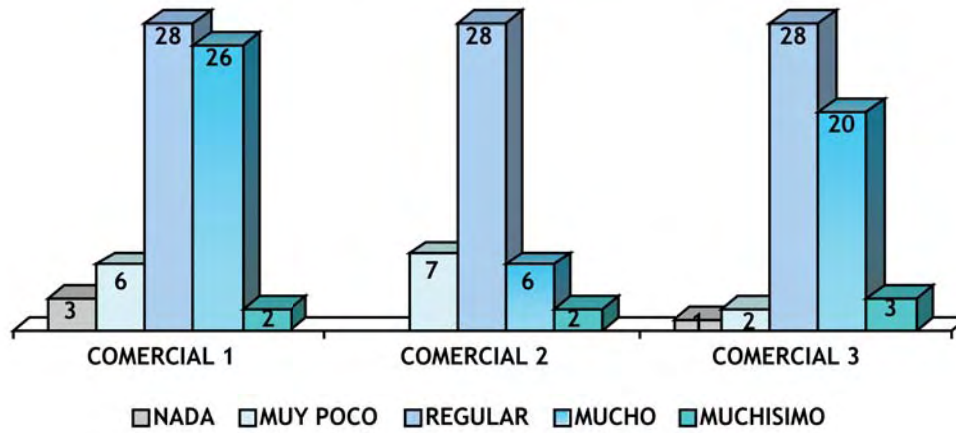
COMERCIAL 3		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TRATAN DE TRANSMITIR CON COMERCIAL 3A	AYUDA AL DESARROLLO Y CRECIMIENTO DE LOS NIÑOS	8	13	15	6	42
	PRODUCTO SANO / CONFIABLE	1	3	1	0	5
	VALORES FAMILIARES	0	3	0	0	3
	FORTALECE HUESOS Y DIENTES	0	0	1	0	1
	LECHE NUTRITIVA / BUENA	2	0	3	0	5
Total		11	19	20	6	56

**Crosstab  
Count**

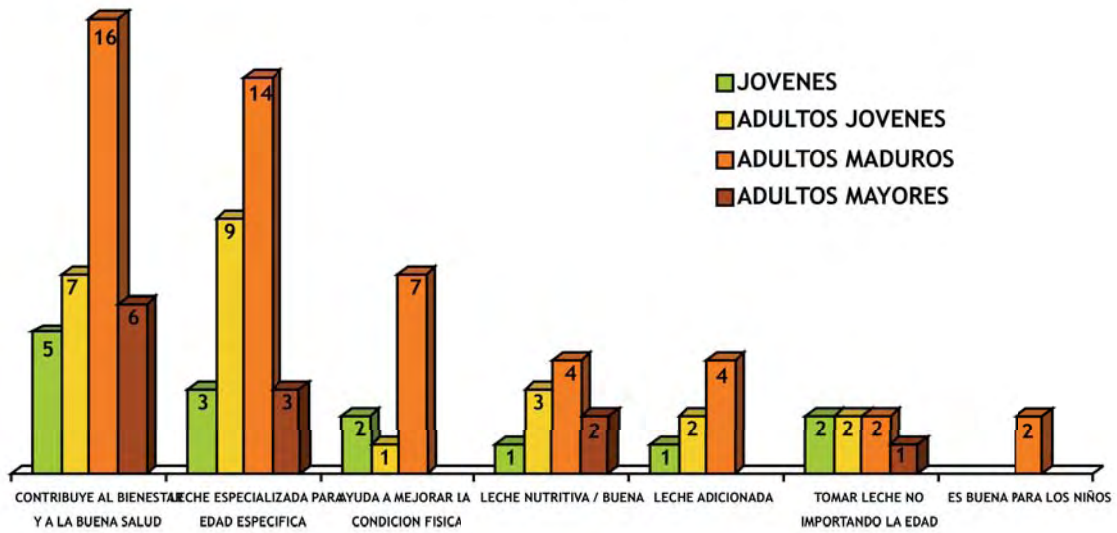
COMERCIAL 3		GRUPOS DE EDADES				Total
		JOVENES	ADULTOS JOVENES	ADULTOS MADUROS	ADULTOS MAYORES	
QUE TRATAN DE TRANSMITIR CON COMERCIAL 3B	AYUDA AL DESARROLLO Y CRECIMIENTO DE LOS NIÑOS	0	1	1	0	2
	PRODUCTO SANO / CONFIABLE	2	2	0	0	4
	VALORES FAMILIARES	0	1	1	0	2
	FORTALECE HUESOS Y DIENTES	0	0	2	1	3
	LECHE NUTRITIVA / BUENA	2	0	0	0	2
Total		4	4	4	1	13



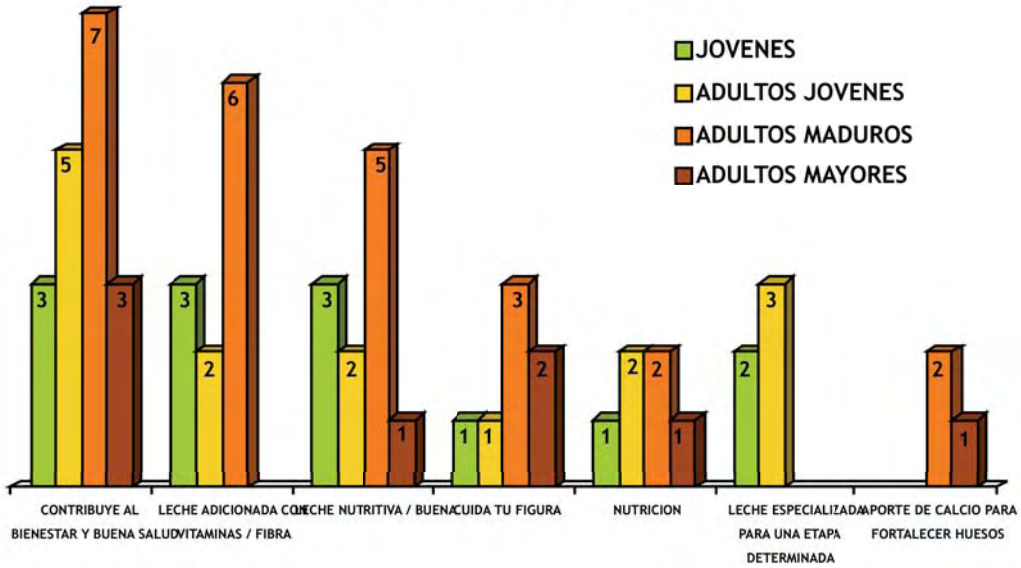
¿QUÉ TANTO LE GUSTARON LOS COMERCIALES?



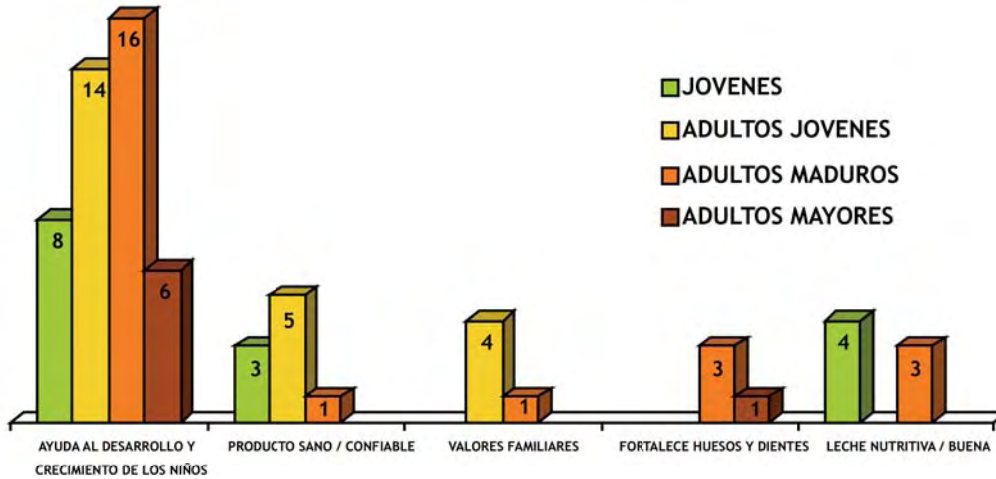
¿QUÉ TRATO DE COMUNICARLE EL COMERCIAL 1?  
POR GRUPOS DE EDAD



¿QUÉ TRATO DE COMUNICARLE EL COMERCIAL 2?  
POR GRUPOS DE EDAD



¿QUÉ TRATÓ DE COMUNICARLE EL COMERCIAL 3?  
POR GRUPOS DE EDAD



¿Cuál es el mensaje principal de esta publicidad?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MENSAJE PRINCIPAL DE ESTA PUBLICIDAD 1 * SEXO	110	55.0%	90	45.0%	200	100.0%
MENSAJE PRINCIPAL DE ESTA PUBLICIDAD 2 * SEXO	37	18.5%	163	81.5%	200	100.0%

**Crosstab Count**

		SEXO		Total
		MUJER	HOMBRE	
MENSAJE PRINCIPAL DE ESTA PUBLICIDAD 1	MARCA UNICA / INNOVADORA	3	4	7
	LECHE ESPECIALIZADA PARA UNA ETAPA DETERMINADA	3	3	6
	CONFIANZA Y EXPERIENCIA	3	1	4
	VALORES FAMILIARES	4	3	7
	CONTRIBUYE AL BIENESTAR Y A LA BUENA SALUD	35	24	59
	AYUDA AL CRECIMIENTO Y DESARROLLO DE LOS NIÑOS	6	2	8
	LECHE PARA TODAS LAS EDADES	6	4	10
	LECHE NUTRITIVA	4	5	9
Total		64	46	110

**Crosstab Count**

		SEXO		Total
		MUJER	HOMBRE	
MENSAJE PRINCIPAL DE ESTA PUBLICIDAD 2	LECHE ESPECIALIZADA PARA UNA ETAPA DETERMINADA	1	1	2
	CONFIANZA Y EXPERIENCIA	2	1	3
	VALORES FAMILIARES	0	1	1
	CONTRIBUYE AL BIENESTAR Y A LA BUENA SALUD	6	5	11

	AYUDA AL CRECIMIENTO Y DESARROLLO DE LOS NIÑOS	2	3	5
	LECHE PARA TODAS LAS EDADES	3	5	8
	LECHE NUTRITIVA	7	0	7
Total		21	16	37

¿Qué tanto le gusto en general la publicidad que vio o escuchó de las leches especializadas de Alpura?

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TANTO GUSTO EN GENERAL LA PUBLICIDAD * SEXO	110	55.0%	90	45.0%	200	100.0%

#### QUE TANTO GUSTO EN GENERAL LA PUBLICIDAD \* SEXO

#### Crosstabulation

Count

		SEXO		
		MUJER	HOMBRE	Total
QUE TANTO GUSTO EN GENERAL LA PUBLICIDAD	NADA	0	1	1
	MUY POCO	7	6	13
	REGULAR	26	17	43
	MUCHO	30	19	49
	MUCHISIMO	1	3	4
Total		64	46	110

¿Qué tanto diría usted en general de la publicidad que vio o escuchó de las leches especializadas de Alpura?

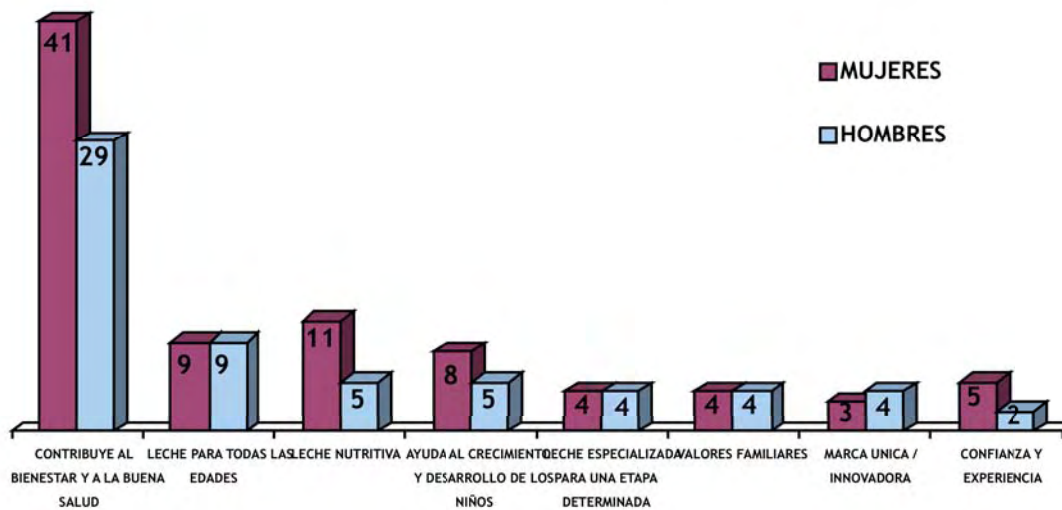
#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TANTO DIRIA DE LA PUBLICIDAD EN GENERAL * SEXO	110	55.0%	90	45.0%	200	100.0%

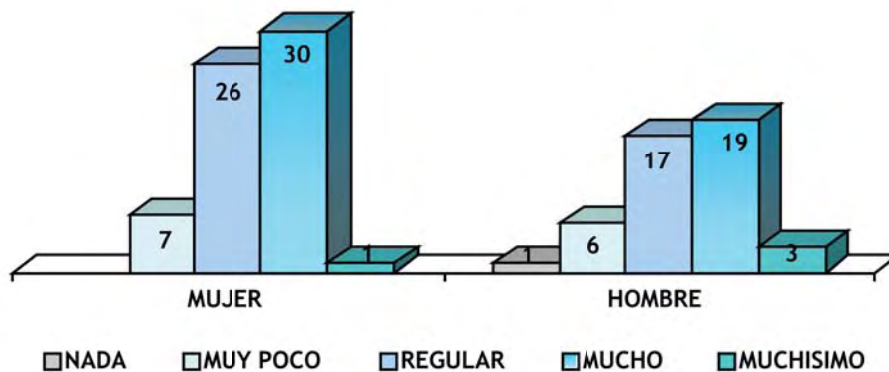
**QUE TANTO DIRIA DE LA PUBLICIDAD EN GENERAL \* SEXO**  
**Crosstabulation**  
 Count

		SEXO		Total
		MUJER	HOMBRE	
QUE TANTO DIRIA DE LA PUBLICIDAD EN GENERAL	NADA	2	2	4
	MUY POCO	3	10	13
	REGULAR	30	18	48
	MUCHO	24	16	40
	MUCHISIMO	5	0	5
Total		64	46	110

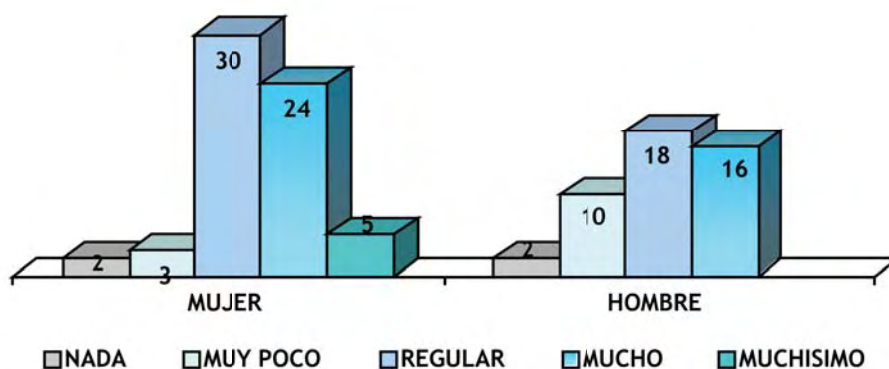
**MENSAJE PRINCIPAL DE LA PUBLICIDAD DE LECHEES ESPECIALIZADAS ALPURA**



**¿QUÉ TANTO LE GUSTÓ EN GENERAL LA PUBLICIDAD DE LECHEES ESPECIALIZADAS DE ALPURA?**



**¿QUÉ TANTO DIRÍA EN GENERAL LA PUBLICIDAD DE LECHEES ESPECIALIZADAS DE ALPURA?**



**3.6.9.4 Opinión de las personas que actualmente no consumen leche alpura**

De las personas que alguna vez consumieron leche líquida de la marca Alpura, se observa de qué tipo fue la leche que consumieron.

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
DE QUE TIPO * SEXO	79	100.0%	0	.0%	79	100.0%

**DE QUE TIPO \* SEXO**

**Crosstabulation**

Count

DE QUE TIPO		SEXO		Total
		MUJER	HOMBRE	
LECHE PASTEURIZADA		26	18	44
LECHE ULTRAPASTEURIZADA		10	3	13
AMBAS		3	5	8
NO SABE		2	12	14
Total		41	38	79

¿Cuándo fue la última vez que probó leche Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
CUANDO FUE LA ULTIMA VEZ QUE PROBO LECHE ALPURA * SEXO	79	39.5%	121	60.5%	200	100.0%

**CUANDO FUE LA ULTIMA VEZ QUE PROBO LECHE ALPURA \* SEXO Crosstabulation Count**

		SEXO		Total
		MUJER	HOMBRE	
CUANDO FUE LA ULTIMA VEZ QUE PROBO LECHE ALPURA	HACE UN MES O MENOS	5	5	10
	HACE 2 O 3 MESES	4	2	6
	HACE 4 O 6 MESES	6	3	9
	HACE UN AÑO	10	9	19
	HACE 2 O 3 AÑOS	11	7	18
	HACE 4 O 5 AÑOS	1	5	6
	HACE MAS DE 5 AÑOS	4	7	11
Total		41	38	79

¿Qué opina de la leche Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE OPINA USTED DE LA LECHE ALPURA 1 * SEXO	79	39.5%	121	60.5%	200	100.0%
QUE OPINA USTED DE LA LECHE ALPURA 2 * SEXO	14	7.0%	186	93.0%	200	100.0%

**Crosstab Count**

		SEXO		Total
		MUJER	HOMBRE	
QUE OPINA USTED DE LA LECHE ALPURA 1	TIENE UNA BUENA CALIDAD	11	13	24
	NO AGRADA SU SABOR	5	3	8
	AGRADA SABOR	7	3	10
	MARCA CON EXPERIENCIA EN EL MERCADO	0	1	1

	ES BUENA LECHE / LECHE FRESCA	16	14	30
	PREFERENCIA OTRA MARCA / CONSUMO DE OTRA MARCA	0	4	4
	NO AGRADA CONSISTENCIA	2	0	2
Total		41	38	79

**Crosstab  
Count**

		SEXO		Total
		MUJER	HOMBRE	
QUE OPINA USTED DE LA LECHE ALPURA 2	TIENE UNA BUENA CALIDAD	2	0	2
	NO AGRADA SU SABOR	2	1	3
	AGRADA SABOR	1	1	2
	MARCA CON EXPERIENCIA EN EL MERCADO	2	0	2
	ES BUENA LECHE / LECHE FRESCA	1	2	3
	PREFERENCIA OTRA MARCA / CONSUMO DE OTRA MARCA	2	0	2
Total		10	4	14

**¿Qué tanto diría de la publicidad de Alpura? ¿Le invita a comprar sus productos?**

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE TANTO DIRIA USTED DE LA PUBLICIDAD DE ALPURA INVITA A COMPRAR SUS PRODUCTOS * SEXO	80	40.0%	120	60.0%	200	100.0%

**QUE TANTO DIRIA USTED DE LA PUBLICIDAD DE ALPURA INVITA A COMPRAR SUS PRODUCTOS \* SEXO**

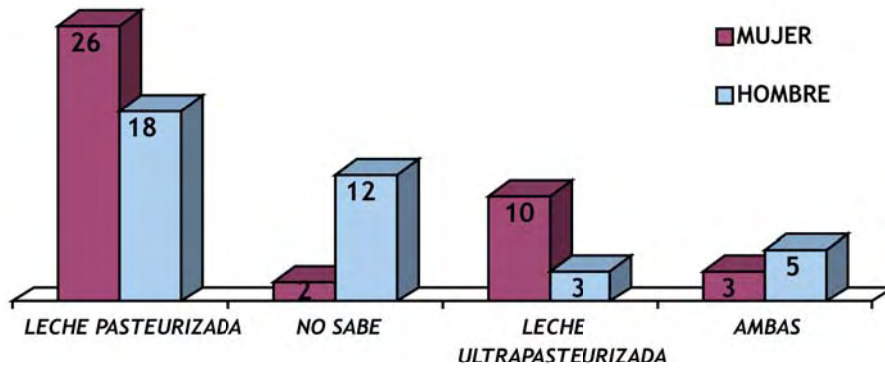
**Crosstabulation  
Count**

		SEXO		Total
		MUJER	HOMBRE	
QUE TANTO DIRIA USTED QUE LA PUBLICIDAD DE	NADA	4	2	6
	MUY POCO	3	6	9
	REGULAR	27	20	47

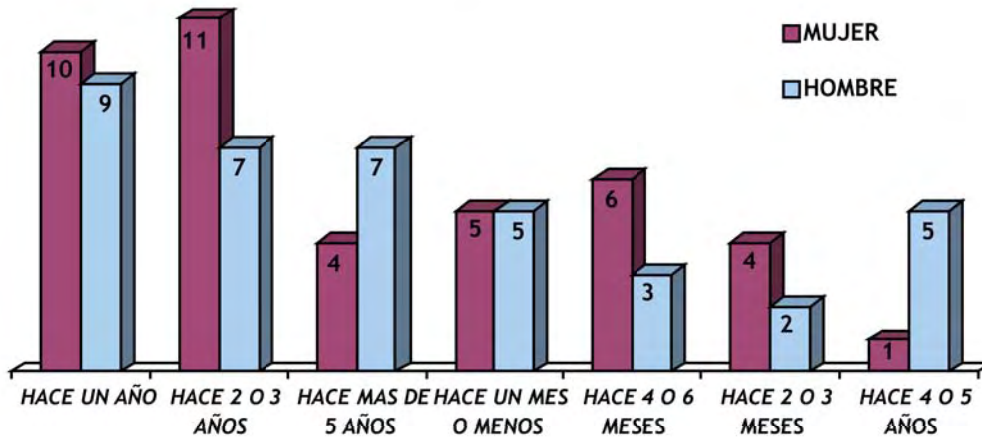


ALPURA INVITA A COMPRAR SUS PRODUCTOS	MUCHO	7	9	16
	MUCHISIMO	1	1	2
Total		42	38	80

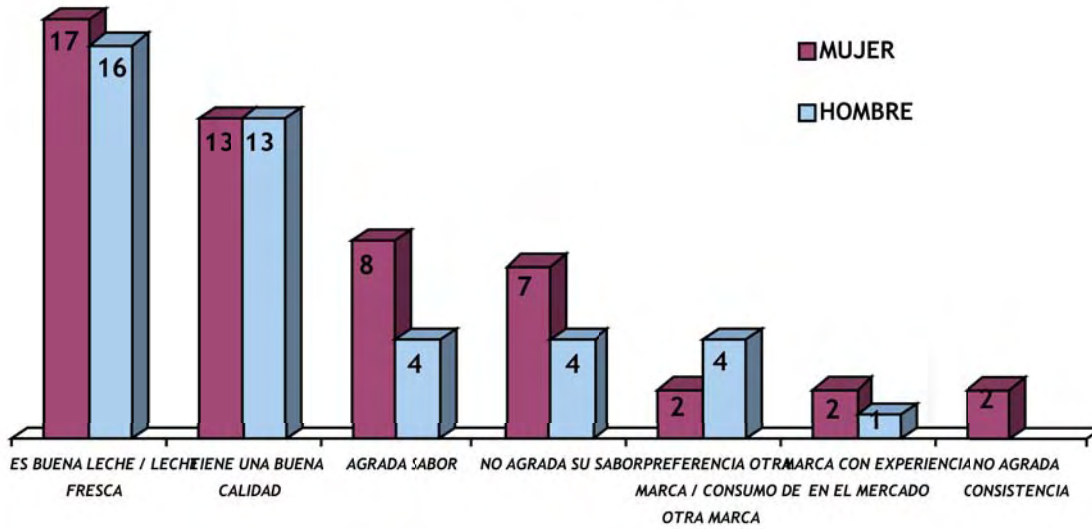
TIPO DE LECHE QUE SE CONSUMIÓ DE ALPURA POR SEXO



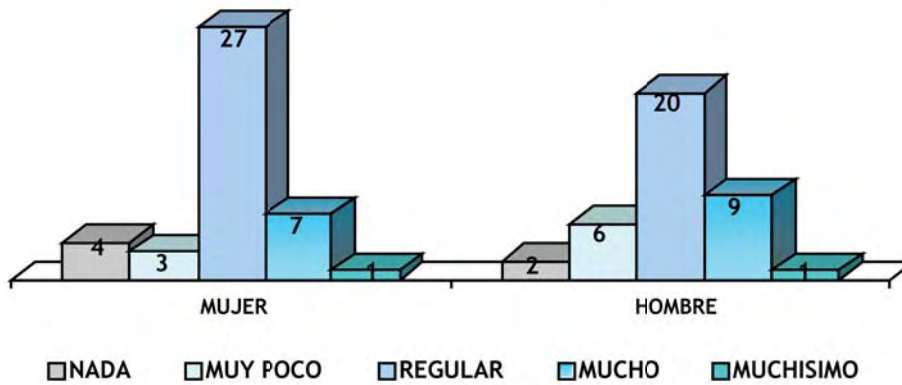
¿CUANDO FUÉ LA ÚLTIMA VEZ QUE PROBÓ ALPURA? POR SEXO



### OPINIÓN DE LA LECHE ALPURA POR SEXO



### LA PUBLICIDAD DE ALPURA ¿LE INVITA A COMPRAR SUS PRODUCTOS?



#### 3.6.9.5 Indicadores de compra de la marca Alpura

¿Qué tendría que hacer Alpura para que la consuma más?

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TOMARIA LECHE ALPURA SI 1 * SEXO	199	99.5%	1	.5%	200	100.0%
TOMARIA LECHE ALPURA SI 2 * SEXO	17	8.5%	183	91.5%	200	100.0%
TOMARIA LECHE ALPURA SI 3 * SEXO	1	.5%	199	99.5%	200	100.0%

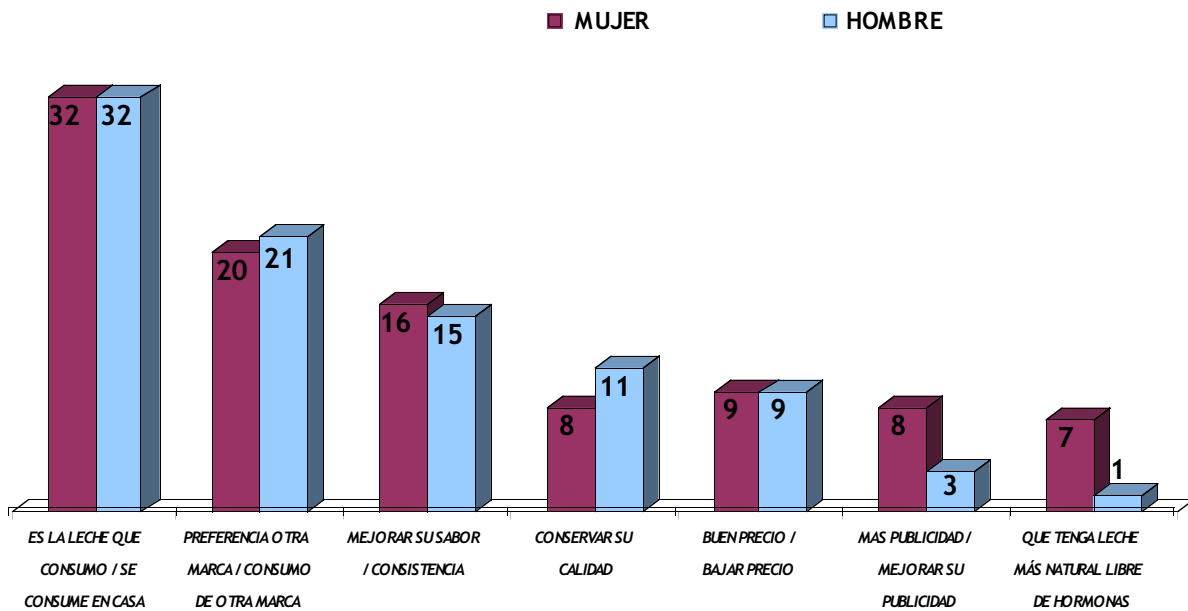
**Crosstab  
Count**

		SEXO		Total
		MUJER	HOMBRE	
TOMARIA LECHE ALPURA SI 1	CONSERVAR SU CALIDAD	8	10	18
	BUEN PRECIO / BAJAR PRECIO	8	6	14
	ES LA LECHE QUE CONSUMO / SE CONSUME EN CASA	32	32	64
	MEJORAR SU SABOR / CONSISTENCIA	15	15	30
	DISPONIBLE EN VARIAS PRESENTACIONES Y TAMAÑOS	1	0	1
	MUESTRAS EN CENTROS COMERCIALES	2	3	5
	MAS PUBLICIDAD / MEJORAR SU PUBLICIDAD	5	3	8
	QUE TENGA LECHE MÁS NATURAL LIBRE DE HORMONAS	7	1	8
	QUE TENGA MAYOR CONTENIDO VITAMINICO	2	1	3
	CAMBIAR EL SABOR DE ALPURA 40 Y TANTOS	2	2	4
	MEJORAR PRESENTACION / EMPAQUES MAS ATRACTIVOS	0	3	3
	PREFERENCIA OTRA MARCA / CONSUMO DE OTRA MARCA	20	19	39
	QUE SEA UNA LECHE MAS LIGERA	2	0	2
	<b>Total</b>	<b>104</b>	<b>95</b>	<b>199</b>

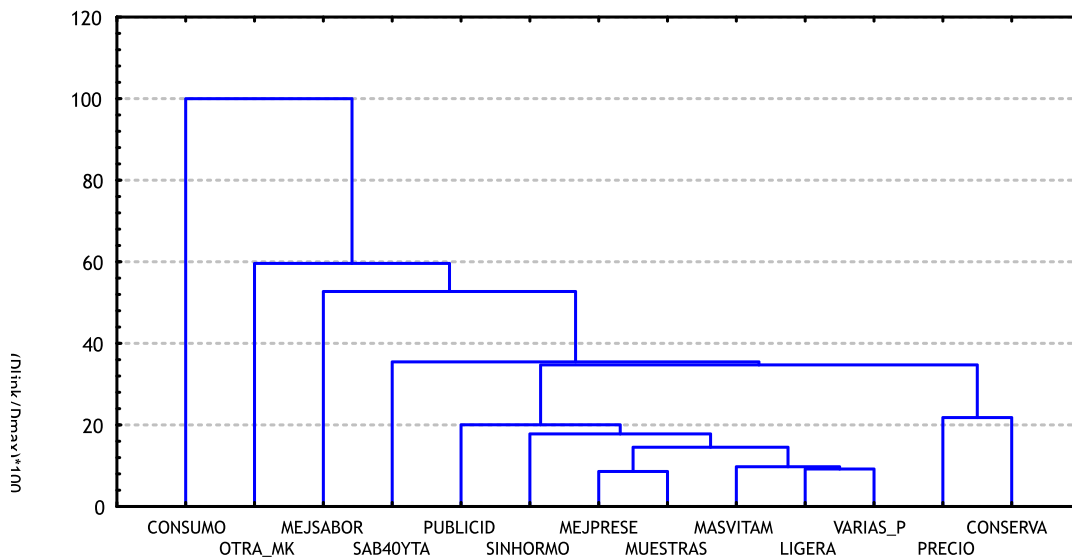
Posiciones: Es la leche que consumo/ se consume en casa (64), Preferencia otra marca/ consumo de otra marca (41), Mejorar su sabor / consistencia (31), Conservar su calidad (19), Buen precio/ bajar precio (18), Mas publicidad / mejorar su publicidad (11), Que tenga leche más natural libre de hormonas (8)

No hay diferencia significativa entre las peticiones de los sexos para aumentar su consumo.

### ¿QUÉ TENDRÍA QUE HACER ALPURA PARA QUE SE CONSUMIERA MÁS?



Tree Diagram for 13 Variables  
Single Linkage  
Euclidean distances



Al hacerles la pregunta ¿Qué tendría que hacer Alpura para que se consumiera más? La respuesta que destacó fue que era la marca que se consumía, seguida por consumo otra marca, en tercer lugar está la recomendación de mejorar su sabor y mejorar el sabor de Alpura 40 y tantos, conservar su calidad y precio, las demás respuestas hacen un grupo semejante.

¿Qué tan probable es que la próxima vez que compre leche líquida, compre de la marca Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
PROBABILIDAD QUE COMPRE LECHE ALPURA LA PROXIMA VEZ * SEXO	200	100.0%	0	.0%	200	100.0%

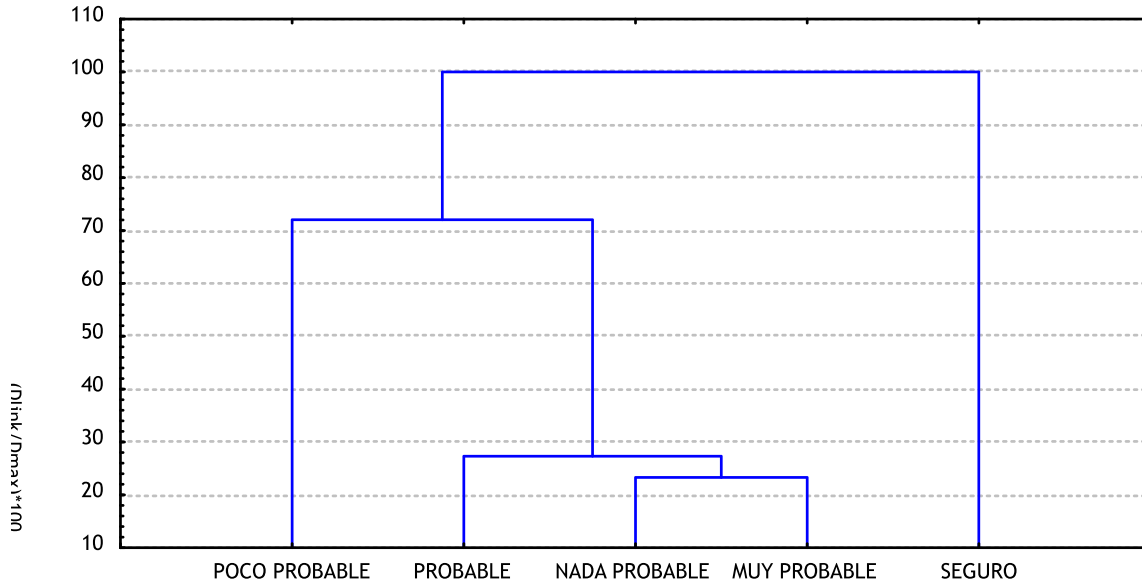
**PROBABILIDAD QUE COMPRE LECHE ALPURA LA PROXIMA VEZ \* SEXO Crosstabulation Count**

		SEXO		Total
		MUJER	HOMBRE	
PROBABILIDAD QUE COMPRE LECHE ALPURA LA PROXIMA VEZ	SEGURO	43	39	82
	MUY PROBABLE	12	7	19
	PROBABLE	11	15	26
	POCO PROBABLE	27	23	50
	NADA PROBABLE	11	12	23
Total		104	96	200

Tree Diagram for 5 Variables

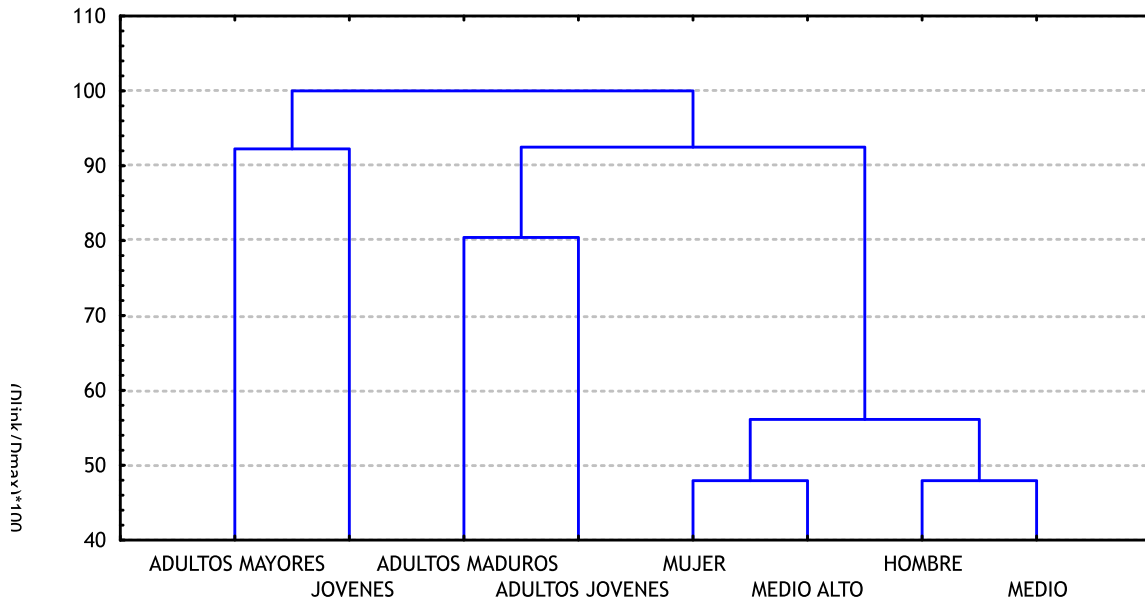
Single Linkage

Euclidean distances



Seguro es la respuesta con más menciones y el grupo de adultos maduros, seguido por el grupo de adultos jóvenes, en segundo lugar la respuesta es poco probable.

Tree Diagram for 8 Cases  
Single Linkage  
Euclidean distances



Se observa que adultos mayores y jóvenes son quienes menos menciones tienen para seguro, los adultos jóvenes y adultos maduros son quienes más menciones tienen para seguro y poco probable, y las mujeres con NSE medio alto son los que seguro comprarían Alpura la próxima vez que compren leche líquida.

¿Por qué es seguro que usted compre leche líquida de la marca Alpura?

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1 * SEXO	82	100.0%	0	.0%	82	100.0%
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 2 * SEXO	21	25.6%	61	74.4%	82	100.0%

Crosstab  
Count

		SEXO		Total
		MUJER	HOMBRE	
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA	AGRADA SU SABOR	3	4	7
	PRECIO COMPETITIVO	1	0	1

VEZ 1	MARCA PREFERIDA / GUSTA MARCA	28	27	55
	POR TRADICION	5	5	10
	CONFIANZA EN LA MARCA	1	1	2
	CONSUMO OTRA MARCA / PREFERENCIA DE OTRA MARCA	1	1	2
	MARCA PREFERIDA DE LOS NIÑOSO	2	1	3
	SE ENCUENTRA EN TODOS LADOS	1	0	1
	LECHE DE LA MAS ALTA CALIDAD	1	0	1
Total		43	39	82

**Crosstab  
Count**

		SEXO		Total
		MUJER	HOMBRE	
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 2	AGRADA SU SABOR	2	2	4
	PRECIO COMPETITIVO	0	1	1
	MARCA PREFERIDA / GUSTA MARCA	0	2	2
	POR TRADICION	2	7	9
	SE ENCUENTRA EN TODOS LADOS	1	1	2
	LECHE DE LA MAS ALTA CALIDAD	2	0	2
	LECHE DE TODO TIPO	1	0	1
Total		8	13	21

¿Por qué es probable que usted compre leche liquida de la marca Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1 * SEXO	26	100.0%	0	.0%	26	100.0%
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 2 * SEXO	1	3.8%	25	96.2%	26	100.0%

**Crosstab  
Count**

		SEXO		Total
		MUJER	HOMBRE	
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1	AGRADA SU SABOR	0	1	1
	MARCA PREFERIDA / GUSTA MARCA	2	2	4
	CONFIANZA EN LA MARCA	0	1	1
	CONSUMO OTRA MARCA / PREFERENCIA DE OTRA MARCA	8	8	16
	SE ENCUENTRA EN TODOS LADOS	0	3	3
	LECHE DE LA MAS ALTA CALIDAD	1	0	1
Total		11	15	26

¿Por qué es poco probable que usted compre leche liquida de la marca Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1 * SEXO	50	100.0%	0	.0%	50	100.0%
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 2 * SEXO	4	8.0%	46	92.0%	50	100.0%

**Crosstab  
Count**

		SEXO		Total
		MUJER	HOMBRE	
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1	AGRADA SU SABOR	1	0	1
	MARCA PREFERIDA / GUSTA MARCA	0	1	1
	CONSUMO OTRA MARCA / PREFERENCIA DE OTRA MARCA	24	22	46
	NO AGRADA SU SABOR / CONSISTENCIA	1	0	1
	LECHE DE TODO TIPO	1	0	1
Total		27	23	50



¿Por qué es nada probable que usted compre leche líquida de la marca Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1 * SEXO	23	100.0%	0	.0%	23	100.0%
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 2 * SEXO	6	26.1%	17	73.9%	23	100.0%

**Crosstab Count**

		SEXO		
		MUJER	HOMBRE	Total
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1	MARCA PREFERIDA / GUSTA MARCA	1	2	3
	CONSUMO OTRA MARCA / PREFERENCIA DE OTRA MARCA	9	10	19
	NO AGRADA SU SABOR / CONSISTENCIA	1	0	1
Total		11	12	23

¿Por qué es muy probable que usted compre leche líquida de la marca Alpura?

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1 * SEXO	19	100.0%	0	.0%	19	100.0%
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 2 * SEXO	2	10.5%	17	89.5%	19	100.0%

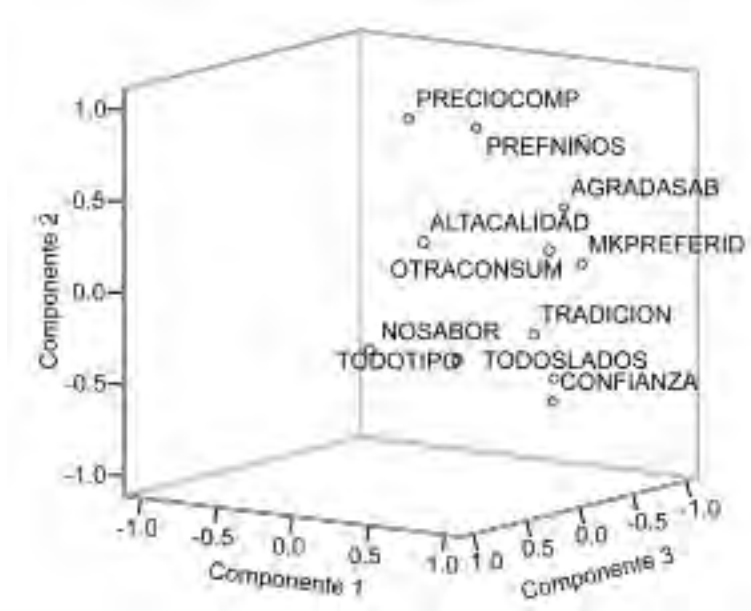
Crosstab  
Count

		SEXO		Total
		MUJER	HOMBRE	
RAZON POR LA QUE COMPRARA O NO ALPURA LA PROXIMA VEZ 1	AGRADA SU SABOR	0	1	1
	PRECIO COMPETITIVO	1	0	1
	MARCA PREFERIDA / GUSTA MARCA	4	3	7
	POR TRADICION	3	1	4
	CONFIANZA EN LA MARCA	0	1	1
	CONSUMO OTRA MARCA / PREFERENCIA DE OTRA MARCA	1	0	1
	MARCA PREFERIDA DE LOS NIÑOS	1	0	1
	SE ENCUENTRA EN TODOS LADOS	2	1	3
Total		12	7	19

RAZONES POR LAS QUE PUEDE COMPRAR O NO LECHE ALPURA



### Gráfico de Componentes

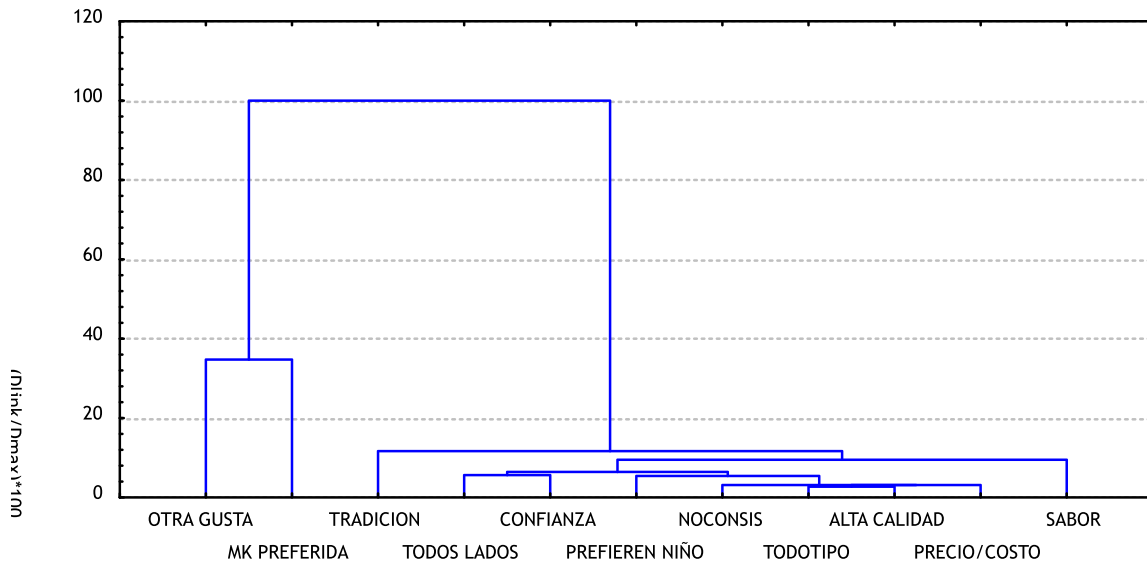


Precio competitivo y preferida por los niños son razones con poco menciones y quien hace estas menciones es el NSE medio alto. La principal razón para consumir o no la marca Alpura es que es la leche que agrada o que gusta otra marca, más mención en el grupo de adultos maduros y adultos jóvenes, los jóvenes mencionan que prefieren otra marca (en este caso Lala es la marca preferida por los jóvenes)

Tree Diagram for 11 Variables

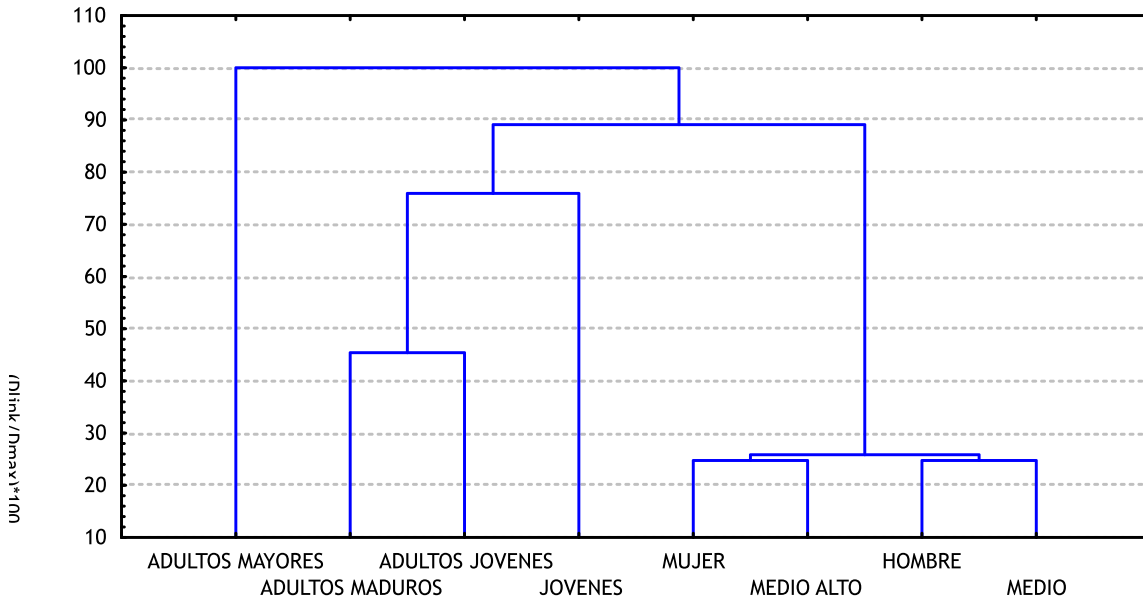
Single Linkage

Euclidean distances



Las razones por las que se consumiría o no la marca Alpura son: agrada o gusta otra marca, después tradición, sabor y al final el conjunto de las otras razones. El orden es más claro que en el Análisis de Componentes Principales

Tree Diagram for 8 Cases  
Single Linkage  
Euclidean distances



Se puede observar que los que menos razones reportan son los adultos mayores seguidos de los jóvenes y adultos maduros, los jóvenes son quienes más razones reportan.

¿Qué es lo que más le gusta de la marca de leche líquida Alpura?

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE ES LO QUE MAS GUSTA DE LA MARCA DE LECHE ALPURA 1 * SEXO	187	93.5%	13	6.5%	200	100.0%
QUE ES LO QUE MAS GUSTA DE LA MARCA DE LECHE ALPURA 2 * SEXO	33	16.5%	167	83.5%	200	100.0%

Crosstab  
Count

		SEXO		Total
		MUJER	HOMBRE	
QUE ES LO QUE MAS GUSTA DE LA MARCA DE LECHE ALPURA 1	AGRADA SU SABOR	32	28	60
	TIENE LA CONSISTENCIA MAS ADECUADA	8	3	11
	SE ENCUENTRA EN TODOS LADOS	12	12	24

	ES SEGURA / CONFIABLE	9	7	16
	GUSTA EL TAMAÑO / LA PRESENTACION	10	6	16
	LECHE ADICIONADA	1	0	1
	LECHE DE LA MAS ALTA CALIDAD	12	17	29
	LECHE DE TODO TIPO	7	6	13
	GUSTA SU PUBLICIDAD / PUBLICIDAD ATRACTIVA	5	8	13
	MARCA CON EXPERIENCIA EN EL MERCADO	3	1	4
Total		99	88	187

¿Qué es lo que más le disgusta de la marca de leche líquida Alpura?

#### Case Processing Summary

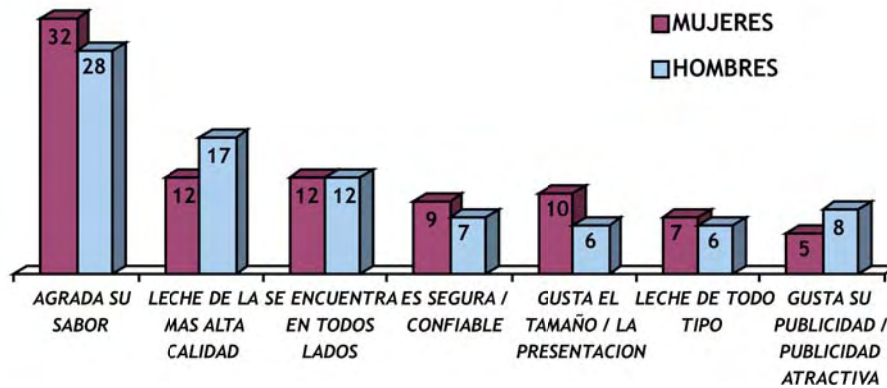
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
QUE ES LO QUE MAS LE DISGUSTA DE LA MARCA DE LECHE ALPURA 1 * SEXO	163	81.5%	37	18.5%	200	100.0%
QUE ES LO QUE MAS LE DISGUSTA DE LA MARCA DE LECHE ALPURA 2 * SEXO	21	10.5%	179	89.5%	200	100.0%

#### Crosstab Count

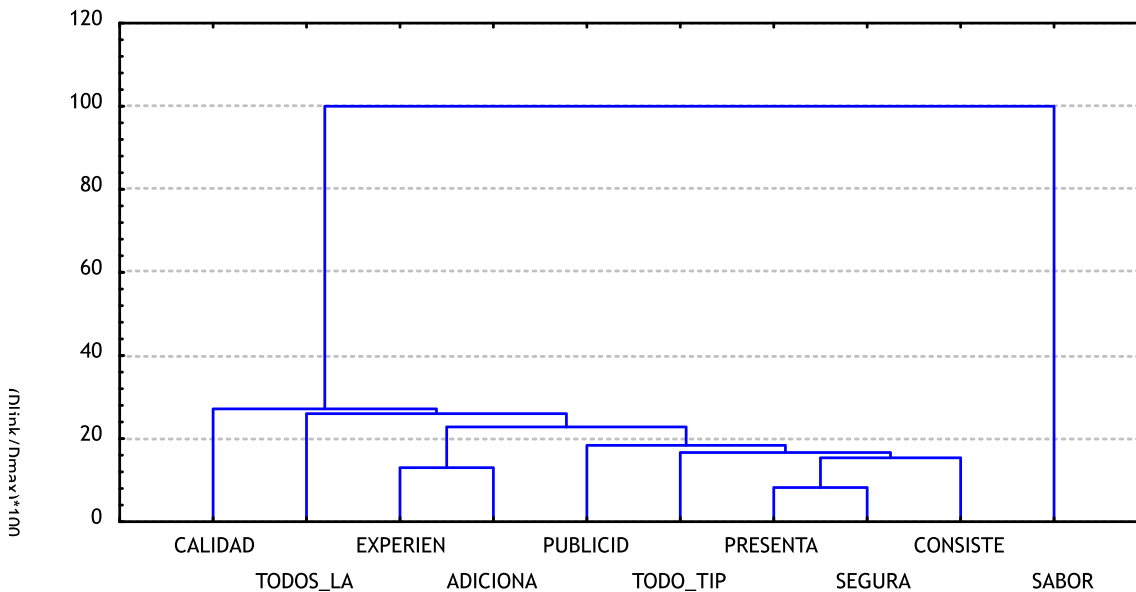
		SEXO		Total
		MUJER	HOMBRE	
QUE ES LO QUE MAS LE DISGUSTA DE LA MARCA DE LECHE ALPURA 1	PRECIO ALTO	7	3	10
	NO RENEVAN SU IMAGEN	0	1	1
	TODO ME AGRADA	33	33	66
	NO AGRADA SABOR	20	20	40
	NO TIENE LECHE LIBRE DE HORMONAS	4	0	4
	NO AGRADA PRESENTACION / EMPAQUES	0	5	5
	ENVASE DIFICIL DE ABRIR	4	9	13
	NO AGRADA EL SABOR DE ALPURA 40 YTANTOS	5	5	10

	NO AGRADA EL SABOR DE ALPURA DESLACTOSADA	5	0	5
	NO AGRADA CONSISTENCIA	6	3	9
Total		84	79	163

**¿QUÉ ES LO QUE MÁS LE GUSTA DE LA MARCA DE LECHE LÍQUIDA ALPURA?**



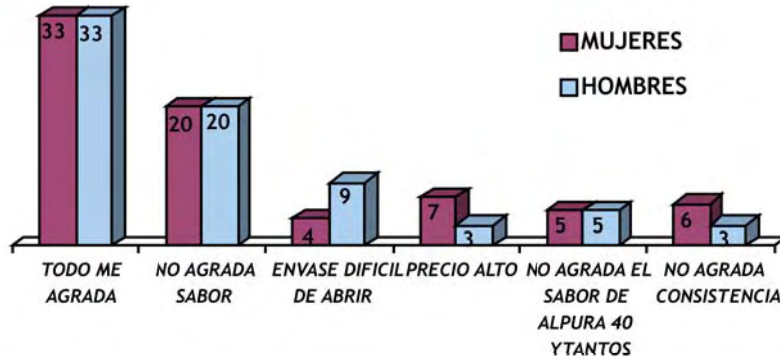
Tree Diagram for 10 Variables  
Single Linkage  
Euclidean distances



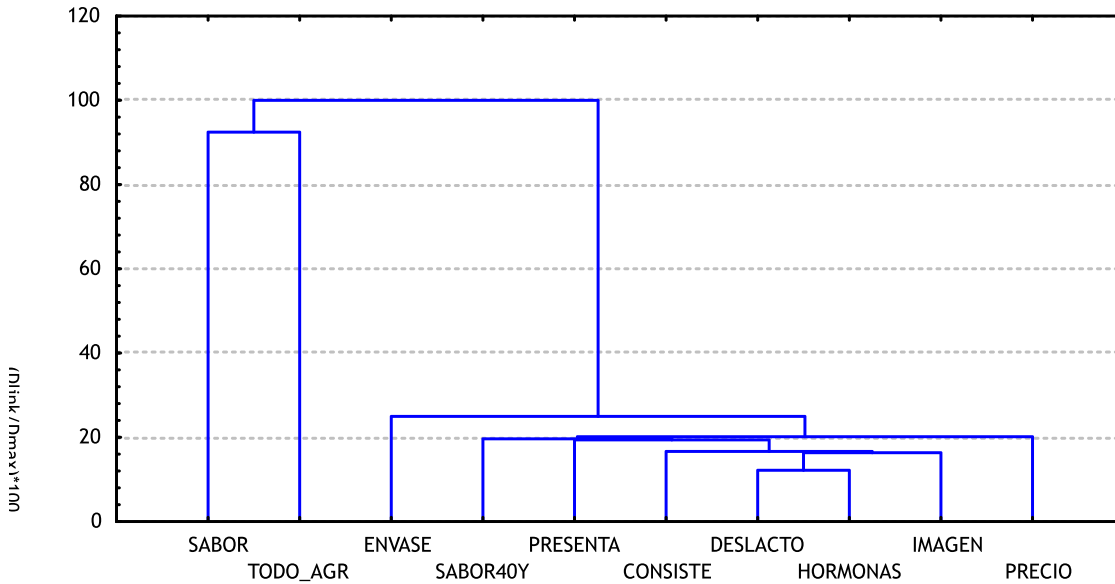
El sabor es muy importante en el gusto de las personas sin importar el NSE, el grupo de edad, ni el sexo. En segundo lugar es la Calidad (Leche de la más alta calidad), sobre todo para los hombres adultos maduros de NSE medio alto. Después está la publicidad junto con leche de todo tipo, la publicidad es preferida por los hombres adultos maduros y jóvenes en el NSE

medio alto, mientras que la leche de todo tipo es algo que mencionan más las mujeres correspondientes al grupo de edad de adultos maduros en NSE medio. Se observa el gusto por la consistencia por parte de los adultos jóvenes y adultos maduros.

### ¿QUÉ ES LO QUE MÁS LE DISGUSTA DE LA MARCA DE LECHE LÍQUIDA ALPURA?



Tree Diagram for 10 Variables  
Single Linkage  
Euclidean distances

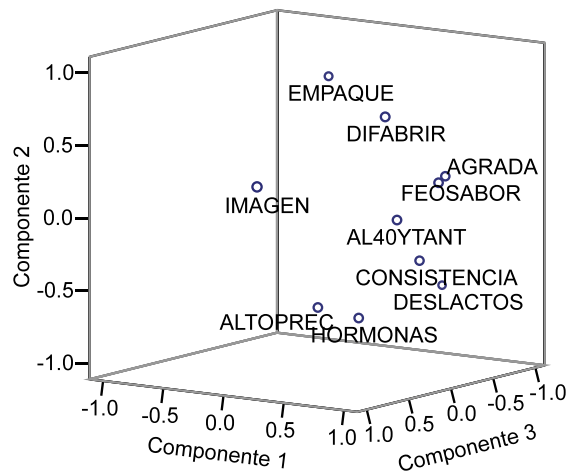


Aunque la mayoría dijo que todo le agradaba, la característica con más menciones es que les disgusta el sabor, sobre todo a los adultos maduros con NSE medio alto. Los adultos jóvenes y adultos maduros mencionaron que su envase era difícil de abrir. Es interesante observar que Alpura 40 y tantos no agrada el sabor al grupo al que va dirigido (adultos maduros).

En el ACP se observa que lo que más disgusta de la marca es el sabor, es la que más menciones presenta en los diferentes grupos de edad, sexo y NSE. La imagen no es una característica que desagrada (sólo tiene una mención de un hombre de NSE medio) por ello se observa en el ACP alejada de las características principales. De manera similar el que sea una leche con hormonas, no agrada a las mujeres jóvenes de NSE medio.

La consistencia, el desagrado de la leche deslactosada y de alpura 40 y tantos, son un grupo de cualidades que desagradan de Alpura.

### Análisis de Componentes





## CAPÍTULO 4

### 4.1 CONCLUSIONES

Es evidente que la evolución de la sociedad igualmente modifica los hábitos y razones de compra en el hogar.

En el país se observa un progreso en el nivel educativo del ama de casa. Si bien se aprecia un comportamiento escalonado en el grado escolar alcanzado en cada nivel socioeconómico, es evidente el avance del ama de casa en este terreno. En consecuencia a mayor nivel educativo, mayor análisis y conciencia en la decisión de compra.

Factores como la edad, actitudes y preferencias de lugar de compra, constituyen un perfil del consumidor, el cual resulta útil para entender su comportamiento y definir cómo acercarse a él de manera eficiente para evitar el desperdicio de recursos.

Para analizarlos se crearon 4 grupos de edad:

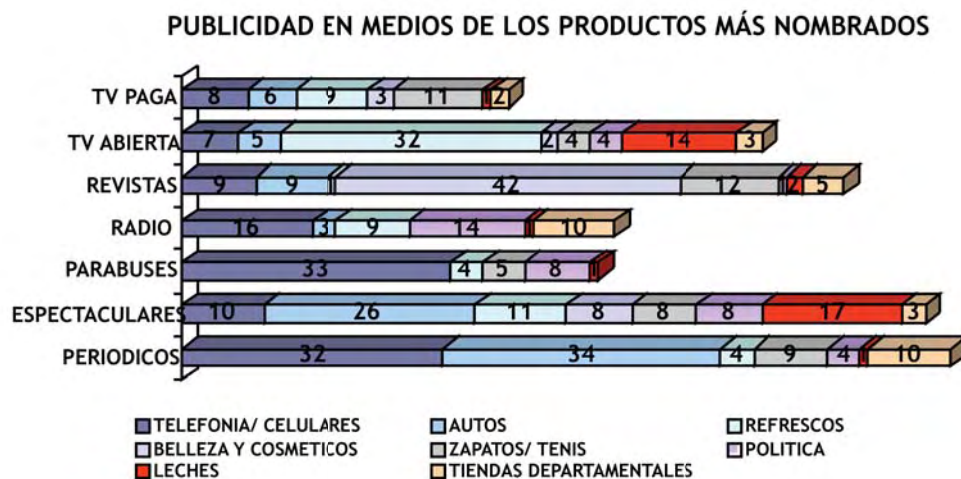
- Jóvenes. De 18 a 28 años
- Adultos jóvenes. De 29 a 39 años
- Adultos maduros. De 40 a 50 años
- Adultos mayores. De 51 a 70 años

Es evidente que las mujeres tienen más “ruido publicitario” en la cabeza en comparación con los hombres. Sin embargo para ciertas categorías de productos hay mayor mención en los hombres, como es el caso de telefonía/celulares, autos y cerveza; caso contrario en ellas que están por arriba en tiendas departamentales, leches, productos lácteos, supermercados y belleza/cosméticos.

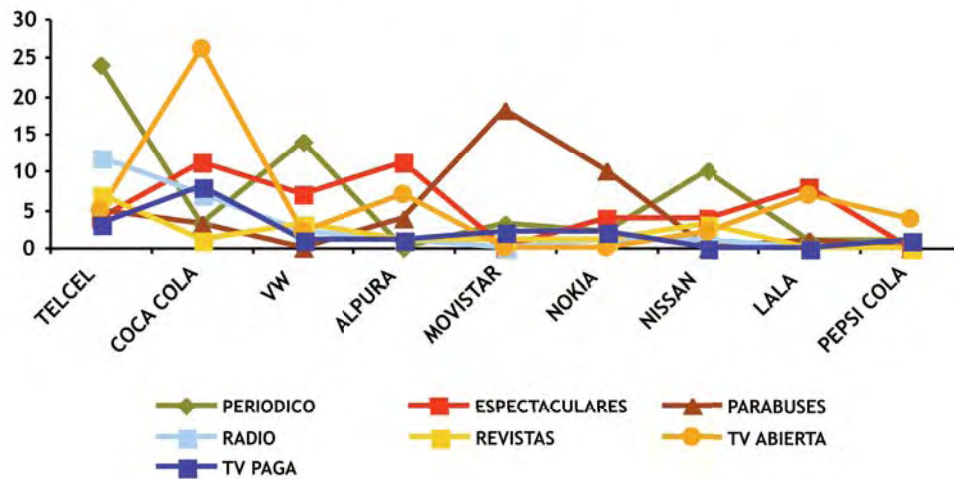
El producto más mencionado en ambos sexos fue el refresco y como marca Coca Cola; en la categoría de leches, se mencionaron Alpura con mayor mención que Lala, pero solo en Mujeres.

En espectaculares es donde tiene más mención la categoría de leches (17 menciones), Alpura se encuentra por arriba de Lala, 11 menciones contra 8 respectivamente.

En televisión abierta se tuvieron 14 menciones de la categoría y en este medio las menciones fueron iguales para las dos marcas (Alpura y Lala).



### PUBLICIDAD EN MEDIOS DE LAS MARCAS MÁS NOMBRADAS



Si bien es cierto que Alpura en el ruido publicitario está por arriba de Lala, es la segunda marca más consumida, mientras que la primera es Lala.

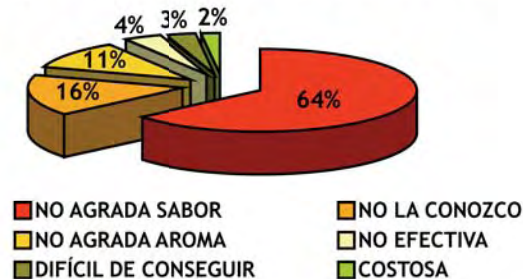
También se observa que Alpura tiene las mejores frases asociadas, pero no tiene el primer lugar en consumo; aunque lala le sigue en frecuencia de frases asociadas, es más consumida y ello porque tiene mejor sabor que Alpura.

Las personas se ven motivadas al consumo principalmente por tradición, leche con mejor sabor y leche de la más alta calidad. El que la leche tenga un sabor agradable es determinante para que ésta sea consumida, recordar que el 53% empezó a consumir la marca porque la probaron y les gustó, mientras que las marcas de leche que no se consumirían muestran como primer motivo "No agrada sabor" con un 64%.

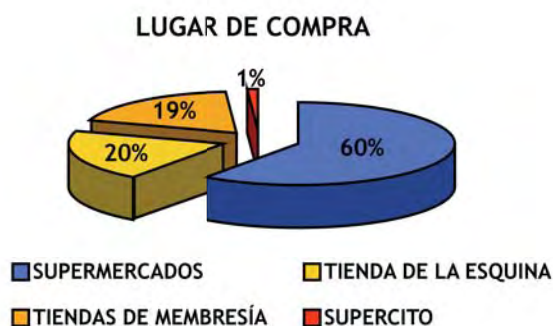
### ¿POR QUÉ COMENZÓ A UTILIZAR ESA MARCA?



### ¿POR QUÉ COMENZÓ A UTILIZAR ESA MARCA?



Desde la perspectiva de los canales de distribución, la expansión del autoservicio continúa avanzando a costa de formatos tradicionales. Se puede ver que el 60% compran la leche en los supermercados.



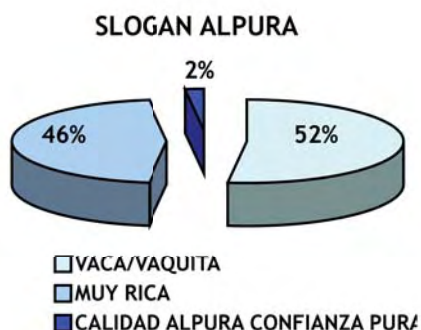
En consecuencia el comercio organizado con su oferta de eficiencia y servicio seguramente continuará en esta línea por algunos años más, facilitando el acceso al consumidor y acentuando la guerra entre marcas por conquistar su preferencia, o mejor dicho su lealtad.

El consumidor está cada vez más expuesto a una creciente gama de opciones sobre la cuál tomará una decisión de compra, y no sólo eso, adoptará nuevas categorías, intentará nuevas marcas y descartará otras.

La ejecución en el punto de venta es fundamental para lograr que una marca sea considerada, la tendencia es a sustituir productos, no necesariamente a incrementarlos.

### Slogan Alpura

Solo una persona recordó el slogan actual de Alpura “Calidad Alpura, confianza pura”. El 46% menciona “Muy rica” que era el slogan anterior.



Es curioso que el 52% al preguntarle el slogan mencione la “vaca/vaquita”. Esto se puede traducir en que el consumidor tiene muy presente el logo de la marca.



### OPINIÓN DE ALPURA DE QUIENES NO LA CONSUMEN

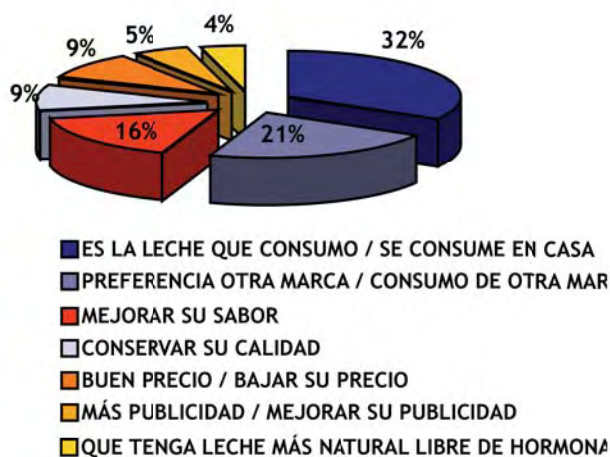


El 34% opina que tiene una buena calidad, el 22% que es buena leche/leche fresca y el 16% que agrada el sabor, sin embargo son personas que actualmente consumen otra marca y que en algún momento consumieron Alpura.

### ¿Qué tendría que hacer Alpura para que se consumiera más?

El 16% mencionaron “mejorar su sabor” y 9% “conservar su calidad”, esto apoya a lo mencionado anteriormente, sobre la importancia del buen sabor de la leche como factor determinante para el consumo. (Ver Cluster P.131)

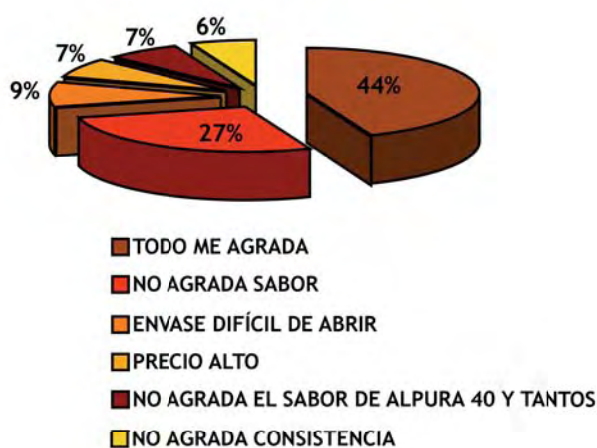
### ¿QUÉ TENDRÍA QUE HACER ALPURA PARA QUE SE CONSUMIERA MÁS?



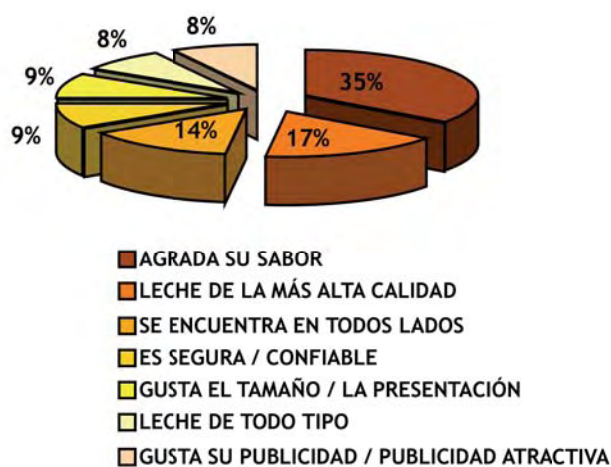
### ¿Qué es lo que más le disgusta de la marca de leche líquida Alpura?

Como se puede observar en la siguiente gráfica, el factor del sabor sigue siendo muy importante para que la gente consuma la leche Alpura (27%), además que se puede notar que al 7% no agrada el sabor de Alpura 40 y tantos (leche especializada).

#### ¿QUÉ ES LO QUE MÁS LE DISGUSTA DE LA MARCA DE LECHE LÍQUIDA ALPURA?



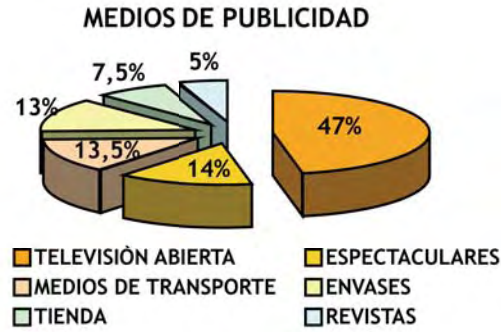
#### ¿QUÉ ES LO QUE MÁS LE GUSTA DE LA MARCA DE LECHE LÍQUIDA ALPURA?



En general tiene puntos positivos la marca, de las 187 personas que contestaron la pregunta el 35% dice que tiene buen sabor (personas que actualmente consumen la marca) el 17% leche de la más alta calidad y 14% se encuentra en todos lados.

### Leches especializadas Alpura

El 47% mencionó que vio publicidad de las leches especializadas de Alpura en televisión abierta, y 14% en espectaculares.

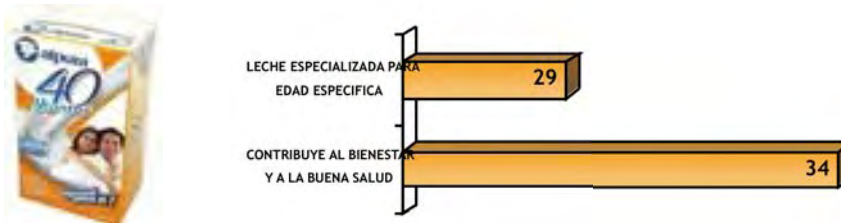


Y de esta publicidad principalmente recuerdan que es una leche especializada para edad específica y la presentación de sus empaque, describiendo las imágenes (personas) y características del producto.



### Comerciales de televisión

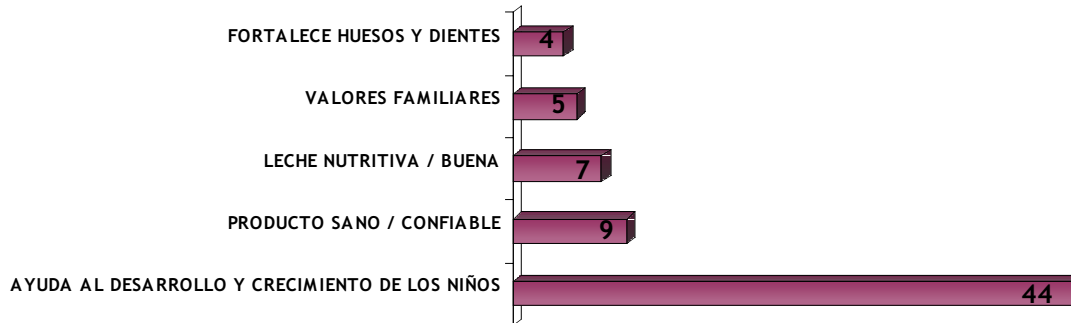
Se les leyó sólo la descripción de los comerciales, el primer comercial fue el más recordado (65 menciones) y el más identificado como publicidad de “Alpura 40 y tantos”.



El comercial 2 fue recordado sólo por 43 personas, sin embargo no fue identificado como publicidad de “Alpura Light Extra y Alpura sin colesterol”



Al leerles la descripción del comercial 3, 54 personas mencionaron recordarlo, sin embargo en todas ellas hay una confusión entre la parte visual del comercial y lo que se dice en él, ya que es publicidad de “Alpura Deslactosada” y fue asociado sólo como un producto para niños.



El consumidor mexicano continuará expuesto a más opciones de producto y canales.

El éxito de un producto depende de su alineación a segmentos específicos. Comprender estos segmentos y accionar la mezcla de mercadotecnia sobre ellos mejorará el desempeño competitivo.

Tomando en cuenta el grupo de edad al que pertenecen, su actitud, medios a los que están expuestos y lugares donde realizan sus compras, es posible alinear las estrategias de comunicación y canales de comercialización.

Teniendo en cuenta esto, se dará el primer paso para alcanzar la lealtad de los consumidores

## 4.2 RECOMENDACIONES

Mucho se ha hablado sobre la influencia que tienen las mujeres en el ámbito familiar y en el social, principalmente cuando el campo laboral se expande a otros ámbitos para la mujer, ya que aumenta su poder adquisitivo y, por lo tanto, de decisión sobre lo que adquiere, modificando sus modos de consumo dependiendo de sus necesidades individuales y familiares, básicas y adquiridas, sobre todo por las exigencias del entorno laboral.

Este fenómeno ha tenido repercusiones muy fuertes en el ámbito económico, principalmente en los sectores industrial y de servicios, sobre todo en el comercio, ya que estos son los que proveen los bienes de consumo a la población.

Las mujeres en general son las que deciden el destino que se le da al ingreso familiar, pues fungen como las administradoras de él, y las que comúnmente detectan las necesidades de todos los miembros como individuos y como grupo.

Es por ello que se debe dirigir más publicidad hacia ellas, principalmente en los grupos de jóvenes y adultos jóvenes, ya que cuando ellas se vuelven jefes de familia (casadas o en unión libre) son quienes toman la decisión de la marca a comprar. Así pasando a los siguientes grupos de edad, consumirán la marca “por tradición”.

El individuo tiene intenciones de consumir productos alimenticios de acuerdo con su gusto por ciertos sabores y a las necesidades de algunos nutrientes. La preferencia de sabores, es en esencia aprendida y proviene de la interacción de fenómenos neurofisiológicos y psicológicos (especialmente afectivos). Sin embargo, el determinante más fuerte de la elección es la situación en la cual la persona escoge, es por ello que se debe mejorar el sabor de la leche ya que éste es un factor determinante para su consumo.

Con los procesos de vida tendientes a “mejorar” la salud y sobre todo la imagen, a través de evitar el sobre peso y obesidad, se popularizan los alimentos bajos en grasa, convirtiéndose en una opción de compra recurrente por parte de ambos sexos, sobre todo de la clase media a alta (dado su costo).

Las leches especializadas buscando posicionarlas como una leche saludable, buscan aminorar la preocupación de sus consumidores en aspectos de salud y bienestar, ayudándoles a estar bien hoy para mantenerse así en el futuro. Es por ello que se debe de poner mayor atención en la creación de los comerciales de éstas leches, ya que sólo Alpura 40 y tantos es la más identificada por los consumidores, por el impacto de su comercial, mientras que el comercial de Alpura Light Extra y Alpura sin Colesterol fue recordado pero no fue identificado como publicidad de estas marcas y en el comercial de Alpura Deslactosada, la gente pone más atención a la parte visual, más no al audio, ya que este comercial fue identificado como un producto sólo para niños.

El grupo de adultos maduros es el que más se debe cuidar - esta tendencia se puede ver que se conserva a través de todo el estudio - ya que son ellos los que consumen más, seguidos de adultos jóvenes, jóvenes y adultos mayores.

### Puntos relevantes

- Mejorar el sabor de la leche, siendo éste un factor determinante para su consumo
- Se debe dirigir más publicidad hacia las mujeres, principalmente para los grupos de jóvenes y adultos jóvenes
- La mayoría no dejaría de consumir su marca por probar otra y hay resistencia en probar otras marcas a medida que se pasa al siguiente grupo de edad
- Poner atención a los comerciales para que el consumidor pueda identificar perfectamente cada producto de la marca



# ANEXOS

ANEXO 1

Encuesta

No. Encuesta \_\_\_\_\_

Buenos días / Tardes / Noches, mi nombre es..... Estamos realizando un estudio sobre los hábitos de consumo de productos y la efectividad de la publicidad, para el cual agradecemos su colaboración respondiendo algunas preguntas

**Sí** Continúe  **No** Termine

P1. Ud. O alguna persona que viva en su hogar trabaja para algún periódico, revista, emisora radial, programadora de TV, empresa de investigación de mercado o agencia de publicidad?

**Sí** Agradezca y termine  **No** Continúe

P2. Cuántas MUJERES Y HOMBRES de 19 a 65 años viven en su hogar, inclúyase, pero no incluya el servicio domestico (ENC: si NO reúnen esta condición, terminar; si SI vaya a P3.) \_\_\_\_\_

P3. De estas personas, cuántas se encuentran entre: (pregunte según la cuota que deba realizar)

	HOMBRES	MUJERES
18 - 28		
29 - 39		
40 - 50		
51 - 61		
62 - +		

**Encuestador: Registre el número de grupos existentes y seleccione el grupo de edad que resulte de la intersección de la casilla del número del cuestionario a realizar.**

Podría ... (cite el nombre)... responderme la entrevista?

**Sí** Datos de encuesta  **No** Agradecer

Nombre del entrevistado:				
Dirección:				
Delegación:			Colonia	
# de Focos	Hasta 5 focos ( 1 )		De 6 a 15 focos ( 1 )	Más de 16 focos ( 1 )
# de Habitaciones	De 1 a 3 habitaciones ( 2 )		De 4 a 6 habitaciones ( 2 )	Más de 7 habitaciones ( 2 )
Nivel de educación del jefe de familia	Hasta Preparatoria ( 3 )		Licenciatura incompleta o más ( 3 )	Maestría y Doctorado ( 3 )
NSE:	1 MEDIO		2 MEDIO ALTO	
Sexo:	1 MUJER		2 HOMBRE	
Edad:				
Estado civil	1 CASADO	2 SOLTERO	3 UNION LIBRE	4 OTROS

**RUIDO PUBLICITARIO**

1. ¿En el último mes, de qué empresas, productos y marcas recuerda haber visto u oído publicidad? Por favor mencione todas las empresas o productos y marcas que recuerde. ¿Algún otro? *Encuestador: anote la categoría de producto o la empresa / institución en la columna de PRODUCTO y la/s marcas recordadas para cada tipo de producto escribalas en la columna contigua, en el renglón correspondiente. En todos los casos insista: ¿Alguna otra? Si sólo menciona marcas pedir que aclare de qué producto se trata.*
2. Específicamente en PERIÓDICOS, de ¿Cuáles empresas, productos y marcas recuerda haber visto alguna publicidad en el último mes?
3. ¿Y en ESPECTACULARES?
4. Más específicamente, en PARABUSES, ¿De cuáles empresas, productos y marcas recuerda haber visto publicidad en el último mes?
5. Y específicamente por RADIO, en el último mes, ¿De cuáles empresas, productos y marcas recuerda haber escuchado publicidad?
6. Y en REVISTAS ¿Cuáles recuerda haber visto en el último mes?
7. Pensando específicamente en la TELEVISIÓN ABIERTA ¿De cuáles empresas, productos y marcas recuerda haber visto u oído publicidad en el último mes?
8. ¿Y en la TELEVISIÓN DE PAGA?

<b>P1. TODOS LOS MEDIOS</b>		<b>P2. PERIÓDICOS</b>		<b>P3. ESPECTACULARES</b>	
-----------------------------	--	-----------------------	--	---------------------------	--

¿Recuerda?		¿Recuerda?		¿Recuerda?	
SI	NO	SI	NO	SI	NO
( 1 )	( 2 )	( 1 )	( 2 )	( 1 )	( 2 )

PRODUCTO	MARCA	PRODUCTO	MARCA	PRODUCTO	MARCA

<b>P4. PARABUSES</b>		<b>P5. RADIO</b>		<b>P6. REVISTAS</b>	
----------------------	--	------------------	--	---------------------	--

¿Recuerda?		¿Recuerda?		¿Recuerda?	
SI	NO	SI	NO	SI	NO
( 1 )	( 2 )	( 1 )	( 2 )	( 1 )	( 2 )

PRODUCTO	MARCA	PRODUCTO	MARCA	PRODUCTO	MARCA

<b>P7. TELEVISION ABIERTA</b>		<b>P8. TELEVISION DE PAGA</b>	
-------------------------------	--	-------------------------------	--

¿Recuerda?		¿Recuerda?	
SI	NO	SI	NO
( 1 )	( 2 )	( 1 )	( 2 )

PRODUCTO	MARCA	PRODUCTO	MARCA

1. ¿Qué marcas de LECHES ha oído nombrar aunque nunca las haya consumido? ¿Alguna otra? **(Registrar 1ª respuesta en la columna de 1ª mención y las restantes en Otras menciones)**
2. ¿Y para cuáles marcas de LECHES que me acaba de mencionar ha visto o ha escuchado publicidad en el último mes? ¿Alguna otra? **(Registrar 1ª respuesta en la columna de 1ª mención y las restantes en Otras menciones)**
3. ¿Conoce o ha oído nombrar la marca de LECHES...? **(Mencionar aquellas marcas que NO mencionó en Preg. 1)**
4. ¿Recuerda haber visto, leído o escuchado publicidad para las LECHES...? **(Mencionar aquellas marcas que NO mencionó en Preg. 2)**
5. ¿Qué marcas de LECHES ha consumido usted o alguno de los miembros de su hogar en los últimos tres meses? **(RM= Respuesta Múltiple)**
6. En el último mes ¿Qué marca de LECHES ha consumido? **(RM)**
7. ¿En la última semana, qué marca de LECHES ha consumido? **(RM)**
8. ¿Qué marca de LECHES es su favorita? **(RS= 1 respuesta)**
9. ¿En qué momento consumen LECHES? **(RM) MOSTRAR TARJETA 1**

Desayuno	Con la comida	Con la cena
1	3	5
Media mañana	En la tarde	Por la noche
2	4	6
Otro ¿Cuál? _____		

10. ¿Me podría decir quién decide qué marcas de LECHES comprar?
  1. Yo mismo
  2. Otra persona      Quién? \_\_\_\_\_ **(Vaya a preg. 23)**
11. ¿Qué marca de LECHES es la que compra con mayor frecuencia? **(RS)**
- 11ª. ¿De qué tipo es la marca de LECHE que compra con mayor frecuencia? **MOSTRAR TARJETA 2**

Leche Pasteurizada	1
Leche Ultrapasteurizada	2
Ambas	3
No sabe	4

12. Si Ud. Fuera al lugar donde acostumbre realizar sus compras de leche líquida y no encontrara la marca que consume en su hogar con mayor frecuencia **(LEER MARCA PREG. 11)** ¿Qué haría usted? **MOSTRAR TARJETA 3**

Iría a otra tienda a buscar mi marca	1	No compraría ninguna marca / regresaría más tarde	4
Compraría la misma marca pero en alguna otra variedad o tipo	2	Compraría otro producto que no fuese leche	5
Compraría otra marca	3	¿Qué otro producto? _____	
¿Qué marca? <b>MARQUE EN LA TABLA A CONTINUACIÓN COL. PREG.12 SUSTITUTA)</b>		¿De qué marca? _____	

13. ¿Qué marca compraría la próxima vez? **(RS)**
14. ¿Qué marca de LECHES nunca compraría? **(RS)**
15. ¿Por qué nunca la compraría? **(RM) MOSTRAR TARJETA 4**

Difícil de conseguir	No agrada el aroma	Costosa
1	3	5
Engorda	No agrada el sabor	No efectiva
2	4	6
Precio		
7		
Otro ¿Cuál? _____		

16. ¿Compró LECHES en la última semana?

SI	1	NO	2	¿Porqué? <b>(Especificar)</b> <b>(Vaya a Preg.23)</b>
----	---	----	---	---

17. ¿Qué marca compró? (RS)

	P1		P2		P3	P4	P5	P6	P7	P8	P11	P12	P13	P14	P17
	1ª Men.	Otras Men.	1ª Men.	Otras Men.	Ayuda Marca	Ayuda Public.	3 Meses	Mes	Semana	Fav.	+ FREC.	Sust.	Prox.	No comp.	Compro
	RS	RM	RS	RM	RM	RM	RM	RM	RM	RS	RS	RS	RS	RS	RS
NINGUNA	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
AL DIA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALPURA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ALPURA CLASICA	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
ALPURA SEMIDESCREMADA	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
ALPURA LIGHT	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
ALPURA 2000	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
ALPURA 2000 SEMIDESCREMADA	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
ALPURA 2000 LIGHT	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
ALPURA 40 Y TANTOS	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
ALPURA DESLACTOSADA	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
ALPURA LIGHT DESLACTOSADA	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
ALPURA LIGHT EXTRA	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
ALPURA SIN COLESTEROL	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
ALPURA FRUTAL	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
ALPURA FRUTAL ULTRAPASTEURIZADA	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
ALPURA FRUTAL PARA LLEVAR	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
ALPURA EN POLVO	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
ALPURA KIDS	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
BADEN	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
BOREAL	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
CONASUPO / LICONSA	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
DESLY	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
FORTILECHE	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23
LALA	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
LALA DESARROLLO	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
LALA DESLACTOSADA	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26
LALA LIGHT	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
LALA PREMIUM (ENTERA)	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
LALA SEMI (SEMIDESCREMADA)	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
LALA SILUETTE PLUS LIGHT	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
LALA SUPREMA	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31
LALA SUPREMA LIGHT	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
MILECHE	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
NUTRILECHE	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34
PARMALAT	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
SAN MARCOS	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36
SVELTY	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
Otra (Especificar)	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38

18. ¿Por qué compró LECHES marca...? (Vea Preg.17 y cite la marca comprada la ultima semana) (RM) TARJETA 5

No tiene conservadores	1	Le da mayor valor por su dinero	8	Es cremosa	14
Es nutritiva	2	Tiene la leche con el mejor sabor	9	Es más natural	15
Tiene leche de la más alta calidad	3	Esta disponible en varias presentaciones y tamaños	10	Tiene la consistencia más adecuada	16
Tiene el precio más adecuado	4	Dura más tiempo sin echarse a perder	11	Tiene empaques más atractivos	17
Por tradición	5	Es la marca preferida de los niños	12	Se consigue fácil	18
La recetó el médico	6	Tiene muchas vitaminas	13	Estimula el crecimiento	19
Gusta el tamaño/la presentación	7	Otra ¿Cuál?			

19. ¿Porqué empezó a utilizar esta marca de LECHES? (RM)MOSTRAR TARJETA 6

La probé y me gustó	1	Tradición	5	Catálogo	9
Un familiar / amigo recomendó	2	Promotora	6	Demostración	10
Lo vio en TV	3	Promoción	7		
Lo vio en una revista	4	Precio	8		

Otro ¿Cuál? \_\_\_\_\_

20. ¿A dónde acostumbra comprar más frecuentemente LECHES? (RM)MOSTRAR TARJETA 7

Tiendas de membresía	1	Tienda de la esquina	4
Supermercados	2		
Supercito	3		
Otro ¿Cuál?	_____		

21. ¿Con qué frecuencia compra LECHES? (RS)MOSTRAR TARJETA 8

Diario	1	Cada 15 días	5
2 o 3 veces a la semana	2	Cada 20 días	6
Semanalmente	3	Mensualmente	7
Cada 10 días	4	Más esporádicamente	8

22. Cuando va a comprar LECHES ¿Cuál frase demuestra mejor la facilidad o dificultad para encontrarlas? TARJETA 9

	TIENDAS DE MEMBRESIA	SUPERMERCADO	SUPERCITO	TIENDAS DE LA ESQUINA
1. Se encuentra en todos lados	1	1	1	1
2. Se encuentra en la mayoría de los lugares	2	2	2	2
3. Se encuentra en pocos lugares	3	3	3	3
4. No se encuentra en ningún lado	4	4	4	4

23. Le voy a leer una serie de frases y me gustaría que me dijera para qué marca o marcas de LECHES aplica mejor cada frase. No importa que nunca haya consumido, simplemente por lo que usted crea o haya escuchado sobre las marcas de LECHES. MOSTRAR TARJETA 10

	ALPURA	LALA	PARMALAT	AL DIA	NUTRILECHE	TODAS	NINGUNA
1. No tiene conservadores	1	2	3	4	5	6	7
2. Es nutritiva	1	2	3	4	5	6	7
3. Tiene leche de la más alta calidad	1	2	3	4	5	6	7
4. Le da mayor valor por su dinero	1	2	3	4	5	6	7
5. Tiene la leche con el mejor sabor	1	2	3	4	5	6	7
6. Está disponible en varias presentaciones y tamaños	1	2	3	4	5	6	7
7. Dura más tiempo sin echarse a perder	1	2	3	4	5	6	7
8. Es la marca preferida de los niños	1	2	3	4	5	6	7
9. Es cremosa	1	2	3	4	5	6	7
10. Es más natural	1	2	3	4	5	6	7
11. Tiene la consistencia más adecuada	1	2	3	4	5	6	7
12. Tiene los empaques más atractivos	1	2	3	4	5	6	7
13. Por tradición	1	2	3	4	5	6	7
14. Se consigue fácil	1	2	3	4	5	6	7
15. Las recetó el médico	1	2	3	4	5	6	7
16. Tiene muchas vitaminas	1	2	3	4	5	6	7
17. Estimula el crecimiento	1	2	3	4	5	6	7
18. Gusta el tamaño / la presentación	1	2	3	4	5	6	7
19. Es la mejor en general	1	2	3	4	5	6	7

**ATENCIÓN ENCUESTADOR:** Si el entrevistado mencionó haber visto publicidad de ALPURA 40 Y TANTOS, ALPURA 2000 SIN COLESTEROL, ALPURA 2000 LIGHT EXTRA, ALPURA DESLACTOSADA, vaya a Preg. 25. Si NO mencionó recordar publicidad de estas marcas, pase a Preg. 27.

24. ¿Ha visto u oído alguna vez la campaña publicitaria de ALPURA 40 Y TANTOS, ALPURA 2000 SIN COLESTEROL, ALPURA 2000 LIGHT EXTRA, ALPURA DESLACTOSADA? **MOSTRAR TARJETA 11**

SI  1 NO  2 (Vaya a Preg.37)

25. ¿Dónde vio o escucho publicidad de la marca ALPURA 40 Y TANTOS, ALPURA 2000 SIN COLESTEROL, ALPURA 2000 LIGHT EXTRA, ALPURA DESLACTOSADA? **MOSTRAR TARJETA 12**

Televisión abierta	1	Camiones/Medios de transportes/Camiones repartidores	8
Radio	2	Espectaculares	9
Prensa	3	Eventos	10
Revistas	4	Folleto	11
Tienda	5	Televisión de paga	12
Vallas	6	Envases	13
Parabuses	7		

26. Por favor dígame ¿Qué es lo que recuerda haber visto, leído o escuchado de la publicidad o artículos que hablen de la marca ALPURA 40 Y TANTOS, ALPURA 2000 SIN COLESTEROL, ALPURA 2000 LIGHT EXTRA, ALPURA DESLACTOSADA? ¿Qué más? ¿Algo más? (RM)


27. ¿Sabe cuál es o recuerda el slogan o frase publicitaria de ALPURA?

SI  1 ¿Cuál es? \_\_\_\_\_ NO  2

**SI NO MENCIONÓ LA FRASE “CALIDAD ALPURA, CONFIANZA PURA” CONTINUÉ, DE LO CONTRARIO PASE A P.29)**

28. La frase publicitaria o slogan es “CALIDAD ALPURA, CONFIANZA PURA” ¿La ha escuchado o la ha visto?

SI  1 NO  2

**COMERCIALES**

Ahora le voy a leer la descripción de unos comerciales

**COMERCIAL 1:** Es un comercial que describe que la vida comienza en la pancita de una mujer embarazada pero que la buena vida se comienza a vivir a partir de los 40 y describe a un hombre viviendo la buena vida en su oficina donde están celebrando su cumpleaños numero 40 con un pastel y es felicitado por sus compañeros de trabajo y a el se le ve feliz, en otra escena esta con su pareja en un muelle felices los dos, en la siguiente escena el se esta sirviendo un vaso de leche y lo toma, la siguiente escena lo muestra con un niño y un perro y en la escena final el esta nadando en una alberca

**COMERCIAL 2:** En la escena sale un hombre sirviéndose un vaso de leche y leyendo un diario, después la imagen de una mujer en un sofá, leyendo una revista y bebiendo un vaso de leche.

**COMERCIAL 3:** Esta una mamá sirviéndole un vaso con leche al niño, el niño bebe el vaso hasta el final y le pide más, se ve la imagen de ellos en una mesa en la cocina de su hogar al fondo.

29. ¿Recuerda haber visto algún comercial?

**SI NO RECUERDA NINGÚN COMERCIAL PASE A PREG.37, DE LO CONTRARIO CONTINÚE**

30. ¿Qué tanto le gusta este comercial? **MOSTRAR TARJETA 13**

COMERCIAL		PREG. 29		PREG. 30				
		RECUERDA	NO RECUERDA	QUE		TANTO	LE GUSTA	
		SI	NO	NADA	MUY POCO	REGULAR	MUCHO	MUCHISIMO
1	COMERCIAL 1	1	2	1	2	3	4	5
2	COMERCIAL 2	1	2	1	2	3	4	5
3	COMERCIAL 3	1	2	1	2	3	4	5

31. Además de que compre el producto ¿Qué cree usted que están tratando de comunicarle en este comercial?  
**MENCIONAR COMERCIAL 1**

_____		_____	
_____		_____	
_____		_____	
_____		_____	
_____		_____	
_____		_____	

32. Además de que compre el producto ¿Qué cree usted que están tratando de comunicarle en este comercial?  
**MENCIONAR COMERCIAL 2**

_____		_____	
_____		_____	
_____		_____	
_____		_____	
_____		_____	
_____		_____	

33. Además de que compre el producto ¿Qué cree usted que están tratando de comunicarle en este comercial?  
**MENCIONAR COMERCIAL 3**

_____		_____	
_____		_____	
_____		_____	
_____		_____	
_____		_____	
_____		_____	

34. Según usted ¿Cuál es el mensaje principal de esta publicidad? ¿Qué más? ¿Algo más?

---



---



---



---



---



35. En una escala de 1 al 5, donde 1 significa que no le gustó nada y 5 que le gustó muchísimo ¿Qué tanto le gustó en general la publicidad que vio o escuchó de ALPURA 40 Y TANTOS, ALPURA 2000 SIN COLESTEROL, ALPURA 2000 LIGHT EXTRA, ALPURA DESLACTOSADA? **MOSTRAR TARJETA 13**

NADA	MUY POCO	REGULAR	MUCHO	MUCHÍSIMO
1	2	3	4	5

36. En una escala de 1 a 5, donde 1 significa nada y 5 muchísimo ¿Qué tanto diría usted que la publicidad en general de ALPURA 40 Y TANTOS, ALPURA 2000 SIN COLESTEROL, ALPURA 2000 LIGHT EXTRA, ALPURA DESLACTOSADA; le invita a comprar sus productos? **MOSTRAR TARJETA 13**

NADA	MUY POCO	REGULAR	MUCHO	MUCHÍSIMO
1	2	3	4	5

**SECCIÓN ALPURA:**

**SI NO HA CONSUMIDO ALPURA DE LA PREG. 5 A LA 7 CONTINUE, DE LO CONTRARIO PASE A PREG. 41**

37. ¿Alguna vez ha consumido ALPURA?

SI  1

NO  2 **Vaya a Preg. 42**

38. ¿De qué tipo?

Leche Pasteurizada	1
Leche Ultrapasteurizada	2
Ambas	3
No sabe	4

39. ¿Cuándo fue la última vez que probó leche ALPURA? **ESPON TÁNEA (RS)**

Hace un mes o menos	1
Hace 2 - 3 meses	2
Hace 4 - 6 meses	3
Hace un año	4
Hace 2 - 3 años	5
Hace 4 - 5 años	6
Hace más de 5 años	7

40. ¿Qué opina usted de la leche ALPURA?

---



---



---

41. En una escala de 1 a 5 donde 1 significa nada y 5 muchísimo ¿Qué tanto diría usted que la publicidad en general de ALPURA, le invita a comprar sus productos? **MOSTRAR TARJETA 13**

NADA	MUY POCO	REGULAR	MUCHO	MUCHÍSIMO
1	2	3	4	5

42. ¿Qué tendría que hacer ALPURA para que la consuma más seguido? ¿Algo más?

---



---



---

43. Ahora vamos a hablar de la marca ALPURA. ¿Qué tan probable es que usted compre leche líquida de marca ALPURA la próxima vez que usted compre leche líquida? **(RS) MOSTRAR TARJETA 14**

Seguro	1
Muy probable	2
Probable	3
Poco probable	4
Nada probable	5

44. ¿Por qué razón o razones dice usted que es \_\_\_\_\_ **MENCIONE RESPUESTA ANTERIOR**, que usted compre leche líquida de la marca ALPURA la próxima vez?

---

---

---

45. Por lo que usted sabe o imagina ¿Qué es lo que más le gusta a usted de la marca de leche líquida ALPURA? ¿Qué más? ¿Algo más?

---

---

---

46. Y por lo que usted sabe o imagina ¿Qué es lo que más le disgusta a usted de la marca de leche líquida ALPURA? ¿Qué más? ¿Algo más?

---

---

---

**SI CONTESTO NADA EN P.45 Y P.46 INSISTA, HAY ALGO POR MÍNIMO QUE LE HAYA GUSTADO O DISGUSTADO**

## PERFIL DEL CONSUMIDOR Y DEL HOGAR

Para terminar, me podría dar algunos datos personales?

1. Nivel de Estudios:

a. Cuál es el nivel educativo más alto que usted haya cursado?

Sin estudio	1
Primaria incompleta	2
Primaria completa	3
Secundaria incompleta	4
Secundaria completa	5
Preparatoria incompleta	6
Preparatoria completa	7
Universitario incompleto	8
Universitario completo	9
Especialización, maestría, PHD	10

2. Tenencia de vivienda:

a. La casa en la que Ud. vive es?

Rentada	1
Propia	2

3. Tiene Ud. Hijos?

Sí	1	De 0 a 3 años	1
		De 3 a 4 años	2
		De 5 a 6 años	3
No	2	De 10 en adelante	4

4. Servicio domestico:

a. Qué cantidad de servicio domestico tiene en su casa? \_\_\_\_\_

5. Ocupación:

a. Cuál es su ocupación actual?

Trabaja de tiempo completo	1
Trabaja medio tiempo	2
Estudia ½ tiempo y trabaja ½ tiempo	3
Pensionado / jubilado	4
Desempleado	5
Estudiante y no trabaja	6
Se dedica al hogar	7
Otro ¿Cuál? _____	

6. Si trabaja, rol de la empresa:

a. En su trabajo Ud?

Trabajador en empresa familiar	1
Trabajador independiente/no profesional	2
Operario o empleado de base	3
Supervisor o mando medio	4
Trabajador independiente profesional	5
Directivo	6

7. Numero de empleados de la empresa:  
 a. La empresa en la que Ud. trabaja tiene:

1 empleado	1
De 2 a 10 empleados	2
De 11 a 50 empleados	3
De 51 a 200 empleados	4
De 201 a 500 empleados	5
Mas de 500 empleados	6
NS /NR	

8. Transporte:  
 a. Cómo se transporta generalmente?

Transporte Público	1
Motocicleta / bicicleta	2
Taxi	3
Automóvil particular	4

Si tiene automóvil:

De que año es su automóvil?	
El auto que tiene es propio o familiar	1
El auto es de la empresa o del negocio	0

9. Rol en el hogar:  
 a. En su hogar Ud. es?

Jefe del hogar (aporta para el mercado)	1
Ama de casa / pareja / cónyuge	2
Es hijo/a	3
Hermano/Cuñado	4
Nieto/a	5
Otra relación familiar	6
Relación no familiar	7

10. Tenencia de aparatos:

Horno de microondas	1
Sacadora de ropa	2
Lavaplatos eléctrico	3
PC de escritorio	4
Juegos de video	5
Aire acondicionado	6
Cámara de video	7
Video grabadora	8
DVD	9
Calentador de agua	10

11. ¿Qué religión profesa?

Católica	1
Cristiana	2
Evangélica	3
Mahometana	4
Protestante	5
Rosacruz	6
Testigo de Jehová	7
Otra ¿Cuál? _____	

12. ¿Cuántos baños con regadera tiene en su casa? \_\_\_\_\_
13. ¿Cuántos baños sin regadera tiene en su casa? \_\_\_\_\_
14. ¿Cuántos viajes internacionales de placer ha hecho en los últimos seis meses? \_\_\_\_\_
15. ¿Cuántos viajes nacionales de placer ha hecho en los últimos seis meses? \_\_\_\_\_

## ANEXO 2

### ALGUNOS ELEMENTOS PARA LA ENSEÑANZA DE LA ESTADÍSTICA

El presente texto se presenta como material de consulta para alumnos de diversas carreras universitarias de grado y posgrado que cursan asignaturas donde se enseña la estadística como una herramienta de la metodología de la investigación científica.

Se brinda un panorama general e introductorio de los principales temas de una disciplina que opera en dos grandes etapas: la estadística descriptiva y la estadística inferencial.

#### 1.1. DEFINICIÓN Y UTILIDAD DE LA ESTADÍSTICA

La estadística es una ciencia que estudia las características de un conjunto de casos para hallar en ellos regularidades en el comportamiento, que sirve para describir el conjunto y para efectuar predicciones.

La estadística puede dar respuesta a muchas de las necesidades que la sociedad actual plantea. Su tarea fundamental es la reducción de datos, con el objetivo de representar la realidad y transformarla, predecir su futuro o simplemente conocerla.



La estadística puede aplicarse a cualquier ámbito de la realidad, y por ello es utilizada en física, química, biología, medicina, astronomía, psicología, sociología, demografía, etc.

#### 1.2 CLASIFICACIONES DE LA ESTADÍSTICA

Existen varias formas de clasificar los estudios estadísticos:

##### 1.2.1 Según la etapa

Hay una estadística descriptiva y una estadística inferencial. La primera etapa se ocupa de describir la muestra, y la segunda etapa infiere conclusiones a partir de los datos que describen la muestra (por ejemplo, conclusiones con respecto a la población).

Tanto la estadística descriptiva como la estadística inferencial se ocupan de obtener datos nuevos. La diferencia radica en que la estadística descriptiva procede a resumir y organizar

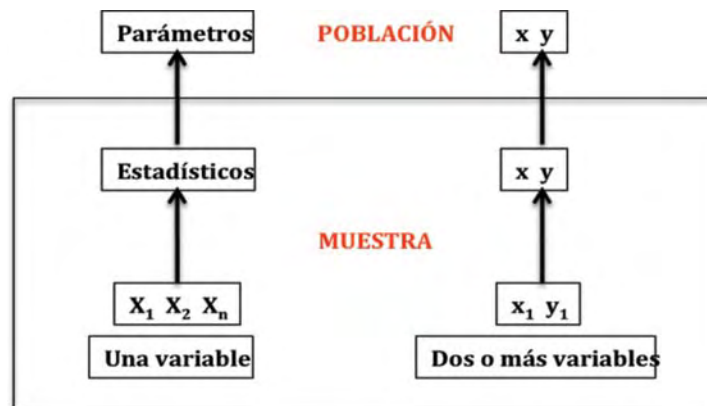
esos datos para facilitar su análisis e interpretación, y la estadística inferencial procede a formular estimaciones y probar hipótesis acerca de la población a partir de esos datos resumidos y obtenidos de la muestra. Puesto que estas últimas operaciones llevarán siempre a conclusiones que tienen algún grado de probabilidad, la teoría de la probabilidad constituye una de sus herramientas principales. Debe tenerse en cuenta que en sí misma la teoría de la probabilidad no forma parte de la estadística porque es otra rama diferente de la matemática, pero es utilizada por la estadística como instrumento para lograr sus propios objetivos.

La estadística descriptiva también incluye - explícita o implícitamente - consideraciones probabilísticas, aunque no resultan ser tan importantes como en la estadística inferencial.

### 1.2.2 Según la cantidad de variables estudiada

Desde este punto de vista hay una estadística univariada (estudia una sola variable), una estadística bivariada (estudia la relación entre dos variables), y una estadística multivariada (estudia tres o más variables)

El siguiente esquema ilustra la relación entre dos clasificaciones de la estadística: descriptiva / inferencial y univariada / bivariada.



### 1.3 POBLACIÓN Y MUESTRA

Puesto que la estadística se ocupa de una gran cantidad de datos, debe primeramente definir de cuáles datos se va a ocupar. El conjunto de datos de los cuales se ocupa un determinado estudio estadístico se llama **población**.

No debe confundirse la población en sentido demográfico y la población en sentido estadístico. La población en sentido demográfico es un conjunto de individuos, mientras que una población en sentido estadístico es un conjunto de datos referidos a determinada característica o atributo de los individuos.

Los datos de la totalidad de una población pueden obtenerse a través de un censo. Sin embargo, en la mayoría de los casos no es posible hacerlo por razones de esfuerzo, tiempo y recursos, razón por la cual se extrae, de la población, una muestra, mediante un procedimiento llamado muestreo. Se llama **muestra** a un subconjunto de unidades de análisis de una población dada, destinada a suministrar información sobre la población. Para que este subconjunto de unidades de análisis sea de utilidad estadística, deben reunirse ciertos requisitos en la selección de los elementos.

## 1.4 ESTRUCTURA DEL DATO

Los datos son la materia prima con que trabaja la estadística, puede definirse técnicamente dato como una categoría asignada a una variable de análisis. Todo dato tiene al menos tres componentes: una unidad de análisis, una variable y una categoría.

La **unidad de análisis** es el elemento del cual se predica una propiedad y característica. Puede ser una persona, una familia, un animal, una sustancia o un objeto.

La **variable** es la característica, propiedad o atributo que se predica de la unidad de análisis. Por ejemplo, edad, ingreso de un individuo, sexo, etc.

Pueden entonces también definirse población estadística (o simplemente población) como el conjunto de datos acerca de unidades de análisis (individuos, objetos) en relación a una misma característica, propiedad o atributo (variable).

La **categoría** es cada una de las posibles variaciones de una variable. Categorías de la variable de sexo son masculino y femenino, de la variable ocupación pueden ser arquitecto, médico, etc. Cuando la variable se mide cuantitativamente, es decir cuando se expresa numéricamente, a la categoría suele llamarse valor. En estos casos, el dato incluye también una unidad de medida, como por ejemplo años, cantidad de hijos, grados de temperatura, etc. El **valor** es entonces, cada una de las posibles variaciones de una variable cuantitativa.

Un **dato individual** es un dato de un solo individuo, mientras que un **dato estadístico** es un dato de una muestra o de una población en su conjunto.

Los datos estadísticos que describen una muestra suelen llamarse **estadísticos** (por ejemplo, el promedio de ingresos mensuales de las personas de una muestra), mientras que los datos estadísticos descriptores de una población suelen llamarse **parámetros** (por ejemplo, el promedio de ingresos mensuales de las personas de una población).

## 1.5 LA MEDICIÓN

Los datos se obtienen a través de un proceso llamado medición. Desde este punto de vista, puede definirse medición como el proceso por el cual se asigna una categoría (o valor) a una variable, para determinada unidad de análisis.

Se pueden hacer mediciones con mayor o menor grado de precisión. Cuanto más precisa sea la medición, más información nos suministra sobre la variable, y por tanto, sobre la unidad de análisis.

Los diferentes grados de precisión o de contenido informativo de una medición se suelen caracterizar como niveles de medición:

- Nominal. Permite clasificar a las unidades de análisis en categorías. Por ejemplo, sexo - varón y mujer.
- Ordinal. Además de clasificar a los elementos en distintas categorías, permite establecer una relación de orden de las mismas. Por ejemplo, clase social baja, media y alta.
- Intervalar. Permite clasificar, ordenar y medir la distancia entre las diferentes categorías. Por ejemplo, edad.
- Razón

**Nivel de medición nominal.** Se tienen dos o más categorías de variables, donde dichas categorías carecen de orden de jerarquía. Por ejemplo el número asignado a cada jugador de un equipo de fútbol.



**Nivel de medición ordinal.** Se tienen varias categorías, pero además mantienen un orden de mayor a menor. Por ejemplo, nombrar la posición de los equipos en un campeonato de fútbol.

**Nivel de medición por intervalos.** Además de haber orden y jerarquía entre categorías, se establecen intervalos iguales en la medición. Por ejemplo, nombrar la posición y el puntaje que obtuvo un equipo en un campeonato de fútbol.

**Nivel de medición de razón.** Además de tenerse todas las categorías del nivel de intervalos, implica parte de un cero real y absoluto donde no existe propiedad a medir. Por ejemplo, el equipo campeón de un torneo de fútbol obtuvo mayor puntaje, pero no fue el mejor equipo en la temporada.

**Escala nominal o de clasificación.** Clasifican objetos reales según ciertas características, tipologías o nombres, dándoles un nombre o símbolo. Los objetos se clasifican con relación a una igualdad o equivalencia de un aspecto o característica.

**Escala ordinal o jerárquica.** Se establecen posiciones relativas de objetos o individuos con relación a una características, sin que reflejen distancias entre ellos.

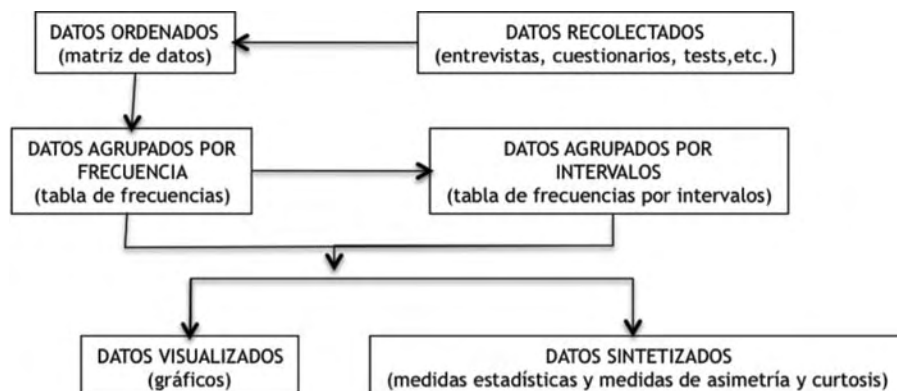
**Escala de intervalos o de distancias iguales.** Representan un nivel de medición más preciso que los anteriores, no solo establece un orden en las posiciones relativas de los objetos o individuos, sino que se mide también la distancia entre los intervalos.

## 1.6 ESTADÍSTICA DESCRIPTIVA

El propósito fundamental de la estadística descriptiva es resumir y organizar una gran cantidad de datos referentes a una muestra o a una población. Se supone que los datos resumidos y organizados permiten describir adecuadamente la muestra o la población a los efectos de conocerla y, eventualmente, utilizarlos en la estadística inferencial para obtener conclusiones a partir de ellos.

Para resumir y organizar los datos se utilizan diferentes procedimientos, llamados técnicas descriptivas: la matriz de datos permite ordenarlos, las tablas de frecuencias (o tablas de distribución de frecuencias) permiten agruparlos, los gráficos permiten visualizarlos, y las medidas estadísticas y las medidas de asimetría y curtosis permiten resumirlos reduciéndolos a un solo dato.

A medida que se van utilizando estos procedimientos, los datos van quedando cada vez más resumidos y organizados. El empleo de dichos procedimientos propios de la estadística descriptiva sigue un orden determinado, como puede apreciarse en el siguiente esquema:



Como puede verse:

- a) Los datos quedan recolectados mediante entrevistas, cuestionarios, tests, etc.,
- b) Los datos quedan ordenados mediante una matriz de datos (lo cual permite resumir la información en pocas páginas)
- c) Los datos quedan agrupados mediante tablas de frecuencias (lo cual permite resumir la información en una sola página)
- d) Los datos quedan visualizados mediante gráficos
- e) Los datos quedan sintetizados mediante las medidas estadísticas y otras (lo cual permite resumir la información en uno o dos renglones)

Puede entonces decirse que, mediante una matriz de datos, una tabla de frecuencias, un gráfico o con medidas estadísticas, etc., la muestra o la población (conjunto de datos) puede quedar adecuadamente descrita.

Estas sucesivas abstracciones estadísticas implican:

- a) La reducción del espacio físico donde queda guardada la nueva información y,
- b) La desaparición de considerable información irrelevante

### **Medidas estadísticas**

Los datos individuales pueden ser sintetizados mediante medidas de posición y medidas de dispersión.

Las medidas de posición son datos estadísticos que intentan representar un conjunto de datos individuales con respecto a una variable.

Esta definición se debe a tres cuestiones:

- Son medidas estadísticas, es decir, no son medidas individuales. Una medida de posición representa a todo un conjunto de datos, y no son los datos individuales. Por ejemplo, un promedio de edades representa a todas las edades de un grupo, y no es la edad individual de uno de sus miembros, aunque pueda coincidir numéricamente con ella
- Son medidas representativas, es decir, intentan representar y sintetizar a todas las medidas individuales. El conjunto de todas las medidas individuales puede recibir diversos nombres, tales como muestra y población, con lo cual tiene sentido afirmar las proposiciones tales como una medida de posición representa una muestra o población
- Son medidas que miden una variable, es decir, algún atributo o propiedad de los objetos.

Las medidas de posición pueden ser de tendencia central, las cuales se refieren a los valores de la variable que suelen estar en el centro de la distribución. Los ejercicios presentados en este capítulo fueron extraídos del curso de Estadística I impartido por el profesor Crisóforo García, ayudante del profesor Joaquín Curiel Cañedo.

#### **1.6.1 Medidas de tendencia central**

Una de las características más sobresalientes de la distribución de datos es su tendencia a acumularse hacia el centro de la misma. Esta característica se denomina tendencia central. Las medidas de tendencia central más usuales son:

##### **1.6.1.1 Media aritmética el valor medio**

La media aritmética de n valores, es igual a la suma de todos ellos dividida entre n. Tenemos:

$$\bar{x} = \frac{\sum fx}{n}$$

Si se cuenta con una distribución de datos entonces se aplica la fórmula:  $\bar{x} = \frac{\sum fx}{N}$

**Ejemplo 1.** Mediante los siguientes datos, hallar la media aritmética. 10, 8, 6, 5, 10, 7.

Solución 1:  $\bar{x} = \frac{10+8+6+5+10+7}{6} = 8$

**Ejemplo 2.** Mediante la siguiente distribución de frecuencias que nos muestra los espesores en pulgadas de recipientes de acero, hallar la media aritmética.

Espesores en pulg	f
0.307 - 0.310	3
0.311 - 0.314	5
0.315 - 0.318	5
0.319 - 0.322	22
0.323 - 0.326	14
0.327 - 0.330	1
	N= 50

Solución 2:

Espesores en pulg	f	X	fX
0.307 - 0.310	3	0.3085	0.9255
0.311 - 0.314	5	0.3125	1.5625
0.315 - 0.318	5	0.3165	1.5825
0.319 - 0.322	22	0.3205	7.0510
0.323 - 0.326	14	0.3245	4.5430
0.327 - 0.330	1	0.3285	0.3285
	N= 50		$\sum fX = 15.9930$

$$x = \frac{15.9930}{50} = 0.3199$$

### 1.6.1.2 Mediana

La mediana es el punto central de una serie de datos, para datos agrupados la mediana viene

dada por:  $Mediana = Li + \frac{(N/2 - \sum fi)c}{fm}$

Donde:

- Li Frontera inferior de la clase de la mediana
- N Número de datos (frecuencia total)
- $\sum fi$  Suma de las frecuencias de las clases inferiores a la de la mediana
- fm Frecuencia de la clase de la mediana
- c Anchura del intervalo de la clase de la mediana

**Ejemplo 1.** Hallar la mediana en los siguientes datos: 25, 30, 28, 26, 32.

Solución 1: Se ordenan en forma creciente o decreciente y se toma el valor central. 25, 26, 28, 30, 32. Mediana=28

**Ejemplo 2.** Hallar la mediana en los siguientes datos: 7, 10, 15, 13, 10, 12.

Solución 2: Al ordenar se tiene 7, 10, 10, 12, 13, 15, pero como el número de datos es par, se toma la media aritmética de los dos internos.

$$\text{Mediana} = \frac{10 + 12}{2} = 11$$

### 1.6.1.3 Moda

Es aquel valor de mayor frecuencia, la moda puede ser no única e inclusive no existir. Para

distribuciones de frecuencia, la moda viene dada por:  $\text{Moda} = Li + \left( \frac{d1}{d1 + d2} \right) c$

Donde:

$Li$  Frontera inferior de la clase modal

$d1$  Es la frecuencia de la clase modal menos la frecuencia de la anterior

$d2$  Es la frecuencia de la clase modal menos la frecuencia de la clase siguiente

$c$  Anchura del intervalo de clase modal

**Ejemplo 1.** Hallar la moda en los siguientes datos: 16, 18, 15, 20, 16.

Solución 1: Moda=16

**Ejemplo 2.** Hallar la moda en la siguiente distribución de frecuencias, la cual nos muestra los diámetros en pulgadas de 60 cojinetes de bolas fabricados por una compañía.

Clases	f	Solución
0.724 - 0.727	5	El intervalo 00.732 - 0.735 contiene la clase modal  $\text{Moda} = 0.7325 + \frac{11(.004)}{11 + 5}$  $\text{Moda} = 0.7353$
0.728 - 0.731	9	
0.732 - 0.735	20	
0.737 - 0.739	15	
0.740 - 0.743	8	
0.744 - 0.747	3	
	N=60	

### 1.6.1.4 Media ponderada

Existe otra medida de tendencia central, la media ponderada.

**Ejemplo 1.** Considerar los siguientes datos:

n	x
100	34
50	37
200	35

¿Cuál es la mejor evaluación de la media general=

Solución 1:  $\frac{34(100) + 37(50) + 35(200)}{100 + 50 + 200} = 35$

### 1.6.2 Medidas de dispersión

Las medidas de dispersión, llamadas también medidas de variabilidad o de variación, son datos estadísticos que informan acerca del grado de dispersión o variabilidad de los datos individuales de una muestra o una población, respecto de una variable. En otras palabras, indican el grado de homogeneidad o de heterogeneidad del conjunto de los datos.

Algunas de las medidas de dispersión más usuales son:

#### 1.6.2.1 Rango

Es la diferencia entre el dato mayor y el dato menor:  $R = X \text{ max.} - X \text{ min.}$

#### 1.6.2.2 Desviación estándar

La desviación estándar o desviación tipo se define como la raíz cuadrada de los cuadrados de las desviaciones de los valores de la variable respecto a su media.

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \text{ datos no agrupados}$$
$$S = \sqrt{\frac{\sum f(X_i - \bar{X})^2}{n}} \text{ datos agrupados}$$

Donde:

$S$	Es la desviación estándar de la muestra
$x_i$ ó $X_i$	Es cada uno de los datos (no agrupados) o la marca de clase (agrupados)
$\bar{x}$ ó $\bar{X}$	Es la media muestral
$n$	Es el número total de daos (tamaño de la muestra)

#### 1.6.2.3 Varianza

Es el cuadrado de la desviación estándar.

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n} \text{ datos no agrupados}$$
$$S^2 = \frac{\sum f(X_i - \bar{X})^2}{n} \text{ datos agrupados}$$

Donde:

$S^2$	Es la varianza de la muestra
$x_i$ ó $X_i$	Es cada uno de los datos (no agrupados) o la marca de clase (agrupados)
$\bar{x}$ ó $\bar{X}$	Es la media muestral
$n$	Es el número total de daos (tamaño de la muestra)

**Ejemplo 1.** Hallar la desviación estándar y la varianza de la siguiente serie de datos: 10, 18, 15, 12, 3, 6, 5, 7.

Solución 1:

$(x_i - \bar{x})^2$
$(10-9.5)^2 = 0.25$
$(18-9.5)^2 = 72.25$
$(15-9.5)^2 = 30.25$
$(12-9.5)^2 = 6.25$
$(3-9.5)^2 = 42.25$
$(6-9.5)^2 = 12.25$
$(5-9.5)^2 = 20.25$
$(7-9.5)^2 = 6.25$
$(x_i - \bar{x})^2 = 190$

$$\bar{x} = 9.5$$

$$S = \sqrt{\frac{190}{8}} = \sqrt{23.75} = 4.87$$

**Ejemplo 2.** Hallar la desviación estándar y la varianza para la siguiente distribución de frecuencias.

Clases	F	Solución	$f(X_i - \bar{X})^2$
10-15	2	$\bar{x} = 26.04$	$2(12.5-26.04) = 366.7$
16-21	8		$8(18.5-26.04) = 454.8$
22-27	13		$13(24.5-26.04) = 46.9$
28-33	10		$10(30.5-26.04) = 168.1$
34-39	6		$6(36.5-26.04) = 656.5$
	39		$\sum f(X_i - \bar{X})^2 = 1693$

$$S = \sqrt{\frac{1693}{39}} = \sqrt{43.4} = 6.6$$

$$S^2 = 43.4$$

#### 1.6.2.4 Desviación media

Se conoce también como promedio de desviación. Para una serie de N valores, puede calcular a través de la siguiente expresión:

$$D.M. = \frac{\sum |x - \bar{x}|}{N}$$

$|x - \bar{x}|$  = valor absoluto de las desviaciones de los x valores respecto a la media.

Y para datos agrupados se tiene:  $D.M. = \frac{\sum f|X - \bar{X}|}{N}$

**Ejemplo 1.** Hallar la desviación media de: 4, 6, 12, 16, 22.

Solución 1:  $\bar{x} = \frac{4 + 6 + 12 + 16 + 22}{5} = 12$

$ x - \bar{x} $
4-12 = 8
6-12 = 6
12-12 = 0
16-12 = 4
22-12 = 10
$\sum  x - \bar{x}  = 28$

$$D.M. = \frac{28}{5} = 5.6$$

**Ejemplo 2.** Hallar la desviación media en la siguiente distribución de frecuencias.

Solución 2:

Clases	$f$	$X$	$fX$	$f X - \bar{X} $
8-10	3	9	27	3(9-15.8)=20.4
11-13	6	12	72	6(12-15.8)=22.8
14-16	9	15	135	9(15-15.8)=7.2
17-19	11	18	198	11(18-15.8)=24.2
20-22	5	21	105	5(21-15.8)=26
	N=34			$\sum f X - \bar{X}  = 100.6$

$$\bar{X} = \frac{537}{34} = 15.8 \quad D.M. = \frac{100.6}{34} = 3$$

### 1.6.2.5 Coeficiente de variación

Es la relación que existe entre la  $S$  y la  $\bar{X}$ , expresada en términos de porcentaje y se expresa:

$$C.V. = \frac{S}{\bar{X}} (100)$$

Donde:

$C.V.$  Es el coeficiente de variación de la muestra  
 $S$  Es la desviación estándar de la muestra  
 $\bar{X}$  Es la media muestral

**Ejemplo 1.** Hallar el coeficiente de variación de una serie de datos cuya  $S=2$  y  $\bar{X} = 16$ .

Solución 1:

$$C.V. = \frac{2}{16} (100) = 12.5\%$$

## 1.7 CORRELACIÓN Y REGRESIÓN

El análisis de correlación permite averiguar el tipo y el grado de asociación estadística entre dos o más variables, mientras que el análisis de regresión permite hacer predicciones sobre la base de la correlación detectada.

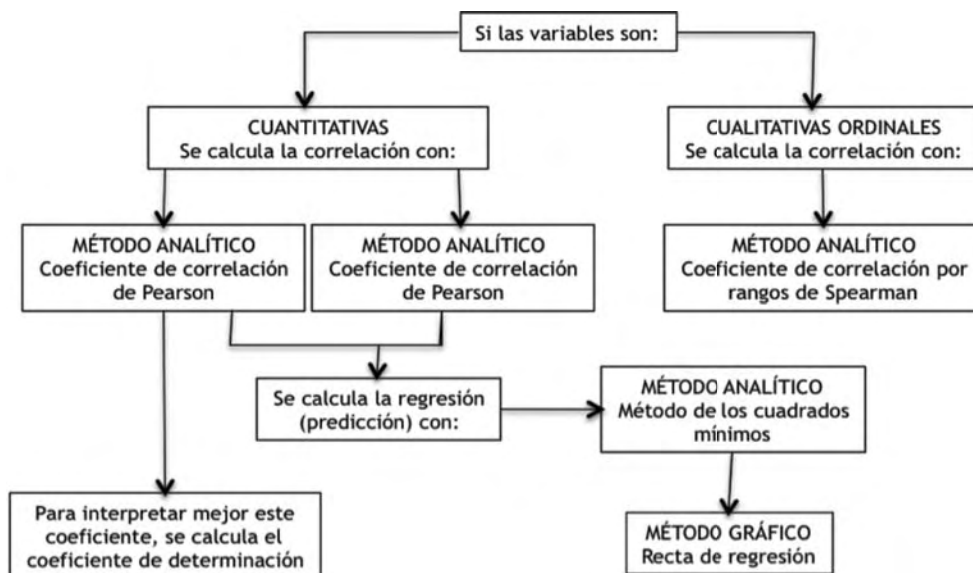
Una vez realizado el análisis de correlación, pueden obtenerse dos resultados: que haya correlación o que no la haya. Si hay correlación, entonces se emprende un análisis de regresión, el cual consiste en predecir cómo seguirán variando esas variables según nuevos valores.

En general el análisis de correlación se realiza conjuntamente con el análisis de regresión. Mientras el análisis de correlación busca asociaciones, el análisis de regresión busca predicciones, es decir, predecir el comportamiento de una variable a partir del comportamiento de la otra.

Vale la pena aclarar que en el contexto de un estudio científico, no basta con determinar el grado de correlación entre dos variables en una muestra. Es necesario además establecer, mediante una prueba de significación, si la correlación establecida en la muestra puede extenderse a toda la población con un determinado nivel de confianza. Esta tarea corresponderá a la estadística inferencial.

Correlación y causalidad. El hecho de que dos variables estén correlacionadas, no significa necesariamente que una sea la causa y la otra el efecto: la correlación no siempre significa causalidad.

El siguiente esquema permite visualizar algunos pasos posibles para llevar a cabo un análisis de correlación seguido de un análisis de regresión.



### 1.7.1 Análisis de correlación

El análisis de correlación busca establecer esencialmente tres cosas:



- Presencia o ausencia de correlación. Dadas dos o más variables, si existe o no correlación entre ellas
- Tipo de correlación. En caso de existir correlación, si esta correlación es directa o inversa. En la correlación directa, ambas variables aumentan (o disminuyen) concomitantemente, y en la correlación inversa ambas variables varían inversamente, o también pueden decirse en relación inversamente proporcional, lo que significa que cuando una aumenta la otra disminuye, o viceversa.
- Grado de correlación. El grado o intensidad de la correlación, es decir, cuánta correlación tienen en términos numéricos

Para hacer todas estas averiguaciones, se puede recurrir a tres procedimientos:

- El método tabular. Una correlación podría constatararse con la simple visualización de tablas de correlación, pero habitualmente las cosas no son tan fáciles, sobre todo porque hay bastante mayor cantidad de datos, y porque estos casi nunca registran los mismos incrementos para ambas variables. Por lo tanto, debe abandonarse la simple visualización de tablas y utilizar procedimientos más confiables, como los gráficos (diagramas de dispersión o dispersiogramas) y los analíticos (por ejemplo el coeficiente de Pearson)
- El método gráfico. Consiste en trazar un diagrama de dispersión
- El método analítico. Consiste en aplicar una fórmula que permita conocer no sólo el tipo de correlación (directa o inversa) sino también una medida cuantitativa precisa del grado de correlación. La fórmula del coeficiente de Pearson es un ejemplo típico para medir correlación entre variables cuantitativas

### 1.7.1.1 Cálculo gráfico de la correlación

Un gráfico es mucho mejor que una tabla para apreciar rápidamente si hay o no correlación entre variables.

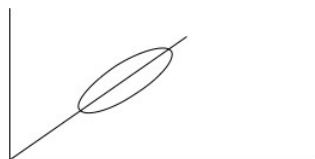
El diagrama de dispersión es básicamente una nube de puntos, donde cada punto corresponde al entrecruzamiento de cada par de valores de X e Y. Este diagrama puede realizarse independientemente del cálculo analítico de la correlación.

Requisitos:

- Que los puntajes sean lineales (se agrupan en una recta), ya sea en forma directamente proporcional o inversamente proporcional.



- Que haya homocedasticidad de varianza (se agrupan en una elipse), es decir que los datos se distribuyen alrededor de esa línea recta.



### 1.7.1.2 Cálculo analítico de la correlación

La correlación se calcula analíticamente mediante ciertos coeficientes, que serán distintos según se trate de correlacionar variables nominales, ordinales o cuantitativas, y según se trate de otras consideraciones.

Si bien existen muchos coeficientes de correlación, se explicará uno de los más utilizados: el coeficiente de correlación lineal de Pearson (para variables cuantitativas).

#### Coeficiente de correlación lineal de Pearson

Este coeficiente (que se designa con “r”), fue creado por Karl Pearson (1857-1936) para relacionar variables cuantitativas.

El coeficiente de Pearson es un número comprendido entre -1 y +1, y que posee un determinado signo (positivo o negativo). El valor numérico indica cuanta correlación hay, mientras que el signo indica qué tipo de correlación es (directa si el signo es positivo, inversa si es negativo)

Interpretación:

**Fuerza.** Esta indica el grado de relación que hay entre dos variables. De acuerdo al valor obtenido que solo puede ir de -1 a +1, pasando por valores positivos y negativos.

Valor	Interpretación de la correlación
-1	Correlación inversa máxima
-0.15	Baja correlación inversa
0	Correlación nula
+0.70	Alta correlación directa
+1	Correlación directa máxima

**Dirección.** Indica hacia que lado de la curva se encuentra la correlación. La dirección le da el signo, los signos se interpretan de la siguiente forma.

Signo	Interpretación
+ (positivo)	Directamente proporcional
- (negativo)	Inversamente proporcional

**Significancia.** Indica cuál es la probabilidad de aceptar o rechazar la hipótesis alternativa. Mayor o igual a 0.05 es significativo.

### 1.7.2 Análisis de regresión

El objetivo del análisis de regresión es establecer una predicción acerca del comportamiento de una variable Y conociendo el correspondiente valor de X (o viceversa) y el grado de correlación existente entre ambas variables.

Para ello es preciso conocer la llamada recta de regresión, que es la recta imaginaria que mejor representa el conjunto de pares de valores de las variables X e Y.

Se puede definir a la regresión como una correlación matemática basada en la ecuación de la recta modificada. Existen varios tipos de regresión y todos se basan en modificaciones de la fórmula de regresión lineal:

$$Y = a + b \cdot X \text{ (ecuación matemática de la recta)}$$

$Y$  es la variable dependiente (de estudio) y  $X$  la variable independiente.  
 $a$  y  $b$  son factores derivados de la ecuación matemática de la recta.

Básicamente por medio de la regresión se pretende predecir el valor de una variable llamada genéricamente “ $Y$ ”, a través de otra variable llamada “ $X$ ”.

La regresión se presenta mediante un coeficiente  $R$  que oscila entre  $-1$  y  $+1$ , cuando la variable dependiente  $Y$  aumenta ante incrementos de la variable independiente  $X$ , el  $R$  es positivo y oscila entre  $0$  y  $1$ . A su vez cuando  $Y$  disminuye ante incrementos de  $X$  el  $R$  es negativo, entre  $0$  y  $-1$ .

Los modelos de regresión no siempre son lineales y se basan en la ecuación pura de la recta. Existen también modificaciones de esta ecuación de tal manera que se pueden practicar análisis de regresión cuadrática, cúbica, logarítmica, logística, etc. Además la regresión puede ser simple o múltiple, constituyendo un tipo de análisis multivariado.

## 1.8 ESTADÍSTICA INFERENCIAL

A diferencia de la estadística descriptiva, la estadística inferencial va más allá de sólo la descripción de la muestra por cuanto se propone, a partir del examen de ésta última, inferir una conclusión acerca de la población, con un cierto nivel de confianza (o, complementariamente, con un cierto nivel de error).

Las muestras de las cuales se ocupa la estadística inferencial son muestras probabilísticas, es decir, aquellas en las cuales es posible calcular el error cometido al estimar una característica poblacional.

La inferencia estadística se ocupa de dos cuestiones: la estimación de parámetros y la prueba de hipótesis, aunque por lo general, la mayoría de las aplicaciones de la estadística inferencial pertenecen al área de la prueba de hipótesis.

En la estimación de parámetros se busca determinar una característica de la población a partir de los datos de la muestra, en cambio en la prueba de hipótesis, se reúne datos en una muestra para validar cierta hipótesis relativa a una población

### 1.8.1 La estimación de parámetros

Esta tarea consiste en, partiendo de ciertos valores de la muestra llamados estadísticos o estadígrafos (por ejemplo la media aritmética muestral), inferir ciertos otros valores de la población llamados parámetros (por ejemplo la media aritmética poblacional o esperanza).

Existen dos tipos de estimación de parámetros: la estimación puntual y la estimación intercalar.

- Estimación puntual. Consiste en inferir un determinado valor para el parámetro. Por ejemplo, inferir que la población debe tener puntualmente una media aritmética de  $80$
- Estimación intercalar. Consiste en inferir dentro de qué intervalo de valores estará el parámetro con un determinado nivel de confianza. Por ejemplo, inferir que la población debe tener una media aritmética entre  $75$  y  $83$ , con un nivel de confianza de  $0.95$  (esto es, hay un  $95\%$  de probabilidades de que el parámetro poblacional se encuentre entre  $75$  y  $83$ ) o, si se quiere, con un nivel de riesgo de  $0.05$  (esto es, hay un  $5\%$  de probabilidades de que el parámetro no se encuentra entre estos valores)

Hay diferentes procedimientos de estimación de parámetros, según se trate de estimar medias o proporciones, o según se trate de estimar parámetros de variables cualitativas (con dos categorías o con más de dos categorías), o cuantitativas.

### 1.8.2 Pruebas de hipótesis

Las pruebas de hipótesis se utilizan para probar alguna hipótesis de investigación científica. Se clasifican en nulas y alternativas, presentan las siguientes características:

- Implican la relación entre variables
- Se plantean sólo cuando se aplica estadística en la investigación

Existe una gran variedad de pruebas de hipótesis, pero todas ellas tienen en común una determinada secuencia de operaciones, que son las siguientes:

#### 1.8.2.1 Formulación de la hipótesis de investigación y obtención de los datos

Se comienza formulando la hipótesis que se pretende probar, luego se diseña un experimento y se ejecuta para obtener datos que permitan aceptar o no la hipótesis. Estrictamente este primer paso no forma parte de la prueba estadística de la hipótesis pero es una condición necesaria para realizarla, y ello por tres motivos:

- a) Si no hay datos empíricos, no puede realizarse ningún estudio estadístico
- b) Si los datos obtenidos en el experimento o en la observación son lo suficientemente convincentes como para aceptar o rechazar la hipótesis de investigación, no será necesario emplear una prueba estadística de hipótesis, con lo cual, este primer paso permite decidir si cabe o no aplicarla, aún cuando se sepa que en la inmensa mayoría de los casos sí cabe hacerlo
- c) Para obtener datos se utiliza un determinado diseño de investigación, y la elección de la prueba estadística de hipótesis más adecuada dependerá del tipo de diseño de investigación utilizado

#### 1.8.2.2 Formulación de la hipótesis alternativa y la hipótesis nula

Si la prueba estadística resulta necesaria, la hipótesis de investigación es reformulada en términos estadísticos, obteniéndose la hipótesis alternativa ( $H_a$ ). A continuación, se formula, en los mismos términos, la hipótesis nula ( $H_0$ ), que es la opuesta de la alternativa. Ambas reformulaciones incluyen consideraciones del tipo “hay o no hay una diferencia significativa entre...”

Si se supone una hipótesis dada (sobre una población) es cierta, pero los resultados hallados en una muestra aleatoria difieren mucho de los esperados bajo tal hipótesis, se dirá que las diferencias observadas son significativas y se tendría que rechazar la hipótesis.

Los procedimientos que permiten determinar si las muestras observadas difieren significativamente de los resultados esperados y que permiten decidir si se acepta o se rechaza la hipótesis se llaman contrastes o test de hipótesis.

Contrastar una hipótesis estadística consiste en aceptar o rechazar, con un cierto grado de confianza (medida numéricamente) y previa observación de muestras de una población, una hipótesis hecha sobre dicha población. La idea es contrastar una hipótesis, consiste en tomar una muestra de observaciones y calcular el estadístico correspondiente al parámetro que se quiere contrastar, y si el valor hallado para el estadístico es próximo al supuesto para el parámetro se acepta la hipótesis y, en caso contrario, se rechaza.

Una vez realizado el contraste de hipótesis, se habrá optado por una de las dos hipótesis (nula o alternativa), y la decisión tomada coincidirá o no con la que en realidad es cierta.

Al estimar parámetros o probar hipótesis pueden cometerse errores:

- Error de tipo I. Consiste en rechazar la hipótesis nula cuando en realidad es verdadera. O sea, creer que la muestra no es representativa de la población, cuando sí lo es
- Error de tipo II. Consiste en aceptar la hipótesis nula cuando en realidad es falsa. O sea, creer que la muestra si es representativa de la población cuando no lo es.

Se pueden dar los cuatro casos siguientes:

	$H_0$ es cierta	$H_a$ es cierta
Se escogió ( $H_0$ )	No hay error	Error de tipo II
Se escogió ( $H_a$ )	Error de tipo I	No hay error

Si la probabilidad de cometer error de tipo I está unívocamente determinada, su valor se suele denotar por la letra griega  $\alpha$ , y en las mismas condiciones, se denota por  $\beta$  la probabilidad de cometer el error de tipo II, esto es:

Cuando es necesario diseñar un contraste de hipótesis, sería deseable hacerlo de tal manera que las probabilidades de ambos tipos de error fueran tan pequeñas como fuera posible. Sin embargo, con una muestra de tamaño prefijado, disminuir la probabilidad del error tipo I,  $\alpha$ , conduce a incrementar la probabilidad del error tipo II,  $\beta$ .

Al contrastar una hipótesis, la máxima probabilidad con la se va a correr el riesgo de rechazar una hipótesis que debiera ser aceptada se llama nivel de significación del contraste.

### 1.8.2.3 Selección de la prueba de hipótesis más adecuada

Hay una gran cantidad de pruebas de hipótesis y su elección depende de la hipótesis alternativa que se formule, del número de casos examinados, del nivel de medición utilizado, etc.

Estas pruebas sirven para resolver los siguiente problemas:

- Reconocer si un grupo de muestras es homogéneo
- Determinar, en la variabilidad de una población de medidas, la parte que corresponde al azar y la que debe atribuirse a causa de variación sistemáticas, llamadas causas controladas o asignadas

### 1.8.2.4 Determinación del nivel de significación

El nivel de significación  $\alpha$  es la probabilidad de rechazar ( $H_0$ ) siendo esta verdadera (error tipo I). Se debe elegir el nivel de significación, es decir, la probabilidad de equivocarse en el sentido indicado.

El nivel de significación elegido dependerá de la importancia práctica de la investigación. Lo usual es especificar un nivel de significación  $\alpha$  (probabilidad de cometer el error tipo I) y no el nivel de significación  $\beta$  (probabilidad de cometer el error tipo II).

### 1.8.2.5 Determinación del tamaño de la muestra

En principio, el tamaño de la muestra  $n$  ya fue determinado en el momento de elegir y ejecutar el diseño de investigación: cuanto mayor haya sido el tamaño de la muestra elegido, menor será el error de  $\alpha$ . Sin embargo, también puede procederse al revés: si se elige un determinado nivel  $\alpha$ , puede determinarse por medios matemáticos el tamaño de la muestra  $n$

adecuado a ese nivel. Así por ejemplo, en general si se desea un menor margen de error, se deberá aumentar el tamaño de la muestra.

Además del tamaño de la muestra, deberán también determinarse la curva operativa característica y el poder de eficiencia de la prueba (o también potencia), definido este último como la probabilidad de rechazar la hipótesis nula cuando es realmente falsa. Por consiguiente, el poder de eficiencia se define como  $1 - \beta$ , es decir 1 menos la probabilidad del error II (no rechazar la hipótesis nula cuando es realmente falsa). Se debe tener presente:

$\alpha$	Nivel de error tipo I	Es la probabilidad de cometer el error tipo I. Probabilidad de rechazar la $H_0$ cuando es verdadera
$1 - \alpha$	Nivel de confianza	Es la probabilidad de no cometer el error tipo I. Probabilidad de no rechazar (aceptar) la $H_0$ cuando es verdadera
$\beta$	Nivel de error tipo II	Es la probabilidad de cometer el error tipo II. Probabilidad de no rechazar (aceptar) la $H_0$ cuando es falsa
$1 - \beta$	Nivel de confianza	Es la probabilidad de no cometer el error tipo II. Probabilidad de rechazar la $H_0$ cuando es falsa. Se llama poder de eficiencia o potencia de la prueba.

## 1.9 GLOSARIO ESTADÍSTICO

### Análisis Estadísticos

En el análisis divariado se investiga la influencia e una variable independiente por vez con respecto a la variable dependiente.

En el análisis multivariado se investiga la influencia de dos o más variables independientes, junto o no a una o más variables asociadas (covariables o cofactores) sobre una o más variables dependientes.

En el análisis univariado se describen las características de una variable por vez. También se le llama estadística descriptiva.

**Correlación.** El concepto de correlación y regresión se basa en el grado de relación que poseen dos variables numéricas entre sí.

El coeficiente de correlación permite observar, si entre dos variables existe o no una relación o dependencia. El coeficiente de regresión permite predecir matemáticamente el grado de relación matemática entre dos variables.

**Desvío.** Es la diferencia entre la  $m$  y cada uno de los valores de la muestra ( $x$ ).

**Desvío Estándar (SD).** Es la raíz cuadrada de la Varianza y corresponde al 68% del área de la curva alrededor de  $m$ .

**Distribución Asimétrica.** Aquella que no guarda la relación de Gauss en su distribución de frecuencia.

**Distribución Normal.** La distribución normal es en forma de campana, habitualmente llamada distribución de Gauss. Es simétrica en torno a su media ( $m$ ); la media, mediana y moda son iguales; el área total de la curva por encima del eje basal  $x$  es la unidad del área=1, por lo tanto cada sector de derecha e izquierda tiene un valor de 0.5. Si se trazan líneas perpendiculares a un desvío estándar ( $s$ ) de distancia de la media, se obtiene un 68% del área de la curva. Dos desvíos estándar encierran un 95%

**Error Estándar.** Es la división entre el SD y la raíz cuadrada de  $n$ .

**Frecuencia (f).** Número de veces en que se repite una característica o las categorías de una variable discreta. Generalmente se acompaña del porcentaje relacionado a esa frecuencia.

**Hipótesis alternativa  $H_a$ .** Es la HP que pretende comprobar el investigador en su muestra. Plantea la posibilidad de que exista una relación causa efecto en el evento de estudio.

**Hipótesis nula  $H_0$ .** Es lo contrario de la alternativa, o sea que no existe relación causa efecto y el evento de estudio se debe puramente al azar.

**Individuo.** Es la unidad mínima que se estudia. En medicina habitualmente es el paciente y en el caso de personas sanas se denomina sujeto o persona. También pueden estudiarse otros como: animales de experimentación, datos de laboratorio, exámenes, etc. (en estos casos se denomina observación).

Las poblaciones pueden ser clasificadas básicamente como sigue:

**Media.** Es la suma de todos los valores dividido por el número de casos  $n$ .

**Mediana.** Corresponde al valor central de la serie de datos observada.

**Moda.** Valor más frecuente de una serie de datos.

**Muestra.** Es el grupo de pacientes u observaciones que se estudiará, la cual debe haberse elegido al azar y ser representativa de un conjunto, población o universo, cuyas características debe reproducir en pequeño lo más exacto posible. A partir del análisis de la muestra, obtenida correctamente y al azar, se pueden hallar conclusiones que sean extrapolables a la población de origen. Para elegir la muestra debe apelarse.

**p.** Es la probabilidad de error al comparar dos o más muestras o grupos cuando aseguramos que ambos son diferentes. O sea que  $p$  es la probabilidad en el sentido de la significación estadística. Obtener una  $p < 0.05$  significa que tenemos un 5% de probabilidades de error en las conclusiones, por lo cual la probabilidad de equivocarnos es baja.

**Población.** Conjunto de individuos, sujetos u observaciones con alguna característica en común. Conjunto de elementos de la misma especie que pretende estudiar en una investigación científica y de la cual se obtiene una muestra.

**Variable Asociada.** Se denomina así a aquella variable independiente que no modifica por su sola presencia a la variable dependiente.

**Variable Cualitativa.** Son variables que representan cualidades de la muestra, como por ejemplo la evolución del paciente hacia la mejoría o la muerte, color de ojos de un grupo de personas, sexo, etc. Estas variables son llamadas Categóricas o Discretas.

**Variable Cuantitativa.** Es la que se puede medir. Habitualmente es llamada variable Numérica o Continua, o sea que posee una continuidad. Por ejemplo la edad, altura, peso, frecuencia cardiaca o respiratoria, dosis de un medicamento.

**Variable Dependiente.** Es la variable motivo de nuestro interés, cuyos valores dependen de otras variables que pueden influir en ella. También se llama variable de respuesta. Por ejemplo la sobrevida, respuesta al tratamiento, evolución, etc.

**Variable Independiente.** Es la que modifica de una u otra manera a la variable dependiente, llamándose también según el caso factor de riesgo, factor predictivo, etc.

**Variable.** Es una característica o propiedad determinada del individuo, sea medible o no. Esta propiedad hace que las personas de un grupo puedan difereir de las de otro grupo en la muestra o población.

**Variabes Categóricas Dicotómicas.** Son las que tienen dos valores fijos y excluyentes entre sí como la evolución, presencia o ausencia de una enfermedad o característica en la muestra.

**Variabes Categóricas Nominales.** Son variables cualitativas que no permiten establecer un orden, por ejemplo la raza, que puede ser blanca, negra, caucásica, etc., o los grupos sanguíneos A, B, AB. También son excluyentes entre sí, o sea que cada paciente pertenece a una u otra categoría pero no a dos al mismo tiempo.

**Variabes Categóricas Ordinales.** Estas sí permiten establecer un orden determinado. También son excluyentes entre sí.



## BIBLIOGRAFÍA

1. ABASCAL, E Y GRANDE I. (1989). Métodos Multivariantes para la Investigación Comercial. Editorial Ariel.
2. AMBROSIO (2000). Plan de Marketing. México: Prentice Hall
3. ANDERSON, T.W. (1985). An Introduction to Multivariate Statistical Analysis, Editora Revolución, La Habana.
4. ANSOFF, IGOR (1993). El Planteamiento Estratégico. Nueva tendencia de la Administración. / H Ansoff, R.P. Declerk y Hayes R. I. / Editorial Trillas, México
5. AZORÍN, F Y SÁNCHEZ-CRESPO, J. L. (1986). Métodos y Aplicaciones para Muestreo. Editorial Alianza. Madrid.
6. BELL, MARTIN, L. (1978). Mercadotecnia, Conceptos y Estrategias. Segunda Edición. México Continental.
7. CANAVOS, G. (1988). Probabilidad y Estadística. Aplicaciones y Métodos. España: McGraw Hill
8. CUADRAS, C.M. Y OTROS (1991). Métodos de Análisis Multivariante. EUNIBAR, Barcelona.
9. FERNANDEZ, VALINAS Y RICARDO (2002). Manual para elaborar un Plan de Mercadotecnia. México: Internacional Thomson Editores
10. FICHER, L. (2001). Mercadotecnia 4ª Edición. México. McGraw Hill. Interamericana de México S.A. de C.V.
11. FREUND, J. y SIMON, G. (1994). Estadística Elemental. México: Prentice Hall Hispanoamericana
12. GILBERT, Y CHURCHILL (2003). Investigación de Mercado. 4ª Edición. México: Internacional Thomson Editores S.A. de C.V.
13. GRIMM, L.G. AND YARNOLD, P.R EDITORS (1998). Reading and Understanding Multivariate Statics, American Psychological Association, Washington, D.C.
14. HAIR, J.F., ANDERSON, R.E., TATHAM, R.L. Y BLACK, W.C. (1995). Multivariate Data Analysis, Fourth Edition, Prentice Hall, Upper Saddle River, New Jersey.
15. HARNETT, D. Y MURPHY, J. (1987). Introducción al Análisis Estadístico. Addison-Wesley Iberoamerica.
16. HUGHES, DAVID, G. (1979). Mercadotecnia, Planeación Estratégica. Tercera Edición. USA: Addison - Wesley
17. HUGHES, DAVID, G. (1996). Mercadotecnia. 4ª Edición. México: Addison - Wesley Iberoamericana
18. KINNEAR, T.C., JAMES R.T. (1991). Investigación de mercados, un enfoque aplicado. Editorial Mc Graw Hill. Tercera Edición.

19. KINNEAR, C. THOMAS & TAYLOR, R. JAMES (1998). Investigación de Mercados. Quinta Edición. USA: McGraw - Hill Internacional, S.A.
20. KOLMOGOROV, A.N. Y FOMIN, S.V. (1970). Introductory Real Analysis, Prentice Hall, Inc., Englewood Cliffs, N.J.
21. KOTLER, PHILIPS (1998). Fundamentos de Mercadotecnia. 4ª Edición. USA: Prentice - Hall, Inc. Kotler, Philip & Armstrong, Gary (2001)
22. KOTLER, PHILIPS & AMSTRONG, GARY (2001). Marketing, 8ª Edición. México: Perason Education de México S.A. de C.V.
23. LAMB, HAIR Y MCDANIEL (1994). Marketing. México: Thomson Editores, S.A. de C.V.
24. LAMB, HAIR Y MCDANIEL (1998). Marketing. 4a Edición. México: Thomson Editores, S.A. de C.V.
25. MORRISON, D.F. (1976). Multivariate Statistical Methods, Second Edition, McGraw Hill Book Company, New York.
26. PEREZ LÓPEZ CÉSAR (2004). Técnicas de Análisis Multivariante de Datos. Editorial Prentice Hall
27. SIEGEL SIDNEY (1976). Estadística no Paramétrica. Editorial Trillas.
28. SHARON L. LOHR (2000). Muestreo. Diseño y Análisis. International Thomson.
29. VISAUTA VINACUA JOAN CARLES MARTORI I CAÑAS, BIENVENIDO (1999). Análisis estadístico con SPSS para Windows. Estadística Multivariante 2da. Edición. Mc Graw - Hill. Madrid.
30. WONNACOTT THOMAS H (1990). Introductory statistics. 5ta. Edición. Editorial Limusa