



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ciencias Naturales y Matemáticas

Ingeniería en Estadística Informática

**“DESARROLLO DE UN MODELO ESTADÍSTICO PARA
MEDIR LAS PREFERENCIAS DE CONSUMO DE LA
POBLACIÓN ECUATORIANA”**

PROYECTO INTEGRADOR

Previo a la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

Presentado por:

Mario Nicolás Plaza Ruíz

Guayaquil – Ecuador

2015

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL FACULTAD DE CIENCIAS
NATURALES Y MATEMÁTICAS**

INGENIERÍA EN ESTADÍSTICA E INFORMÁTICA

INFORME DEL PROFESOR DE LA MATERIA INTEGRADORA

Habiendo sido nombrado PROFESOR DE LA MATERIA INTEGRADORA del
señor,

MARIO NICOLÁS

PLAZA RUÍZ

Con el tema “DESARROLLO DE UN MODELO ESTADÍSTICO PARA MEDIR
LAS PREFERENCIAS DE CONSUMO DE LA POBLACIÓN ECUATORIANA”,
previa a la obtención del título de **INGENIERO EN ESTADÍSTICA
INFORMÁTICA**, me permito informar que he leído el contenido del proyecto
integrador, luego de lo cual indico que estoy de acuerdo en que el mismo se lo
ha desarrollado conforme a los lineamientos de la Unidad de Titulación Especial
de la ESPOL.

ING. SONNIA PAOLA REYES RAMOS

PROFESOR DE LA MATERIA INTEGRADORA

Guayaquil, 19 de Febrero del 2016

AGRADECIMIENTO

A Dios, por darme la fortaleza para llegar al
cúlmino de esta etapa;

A mis madres que siempre me han apoyado en
cada paso que he dado y en cada decisión que he
tomado.

DEDICATORIA

Este trabajo lo dedico a Dios, quien me ha acompañado desde el inicio hasta
el final de mi vida universitaria.

A mi familia, quienes me han enseñado lo que es el arduo trabajo y las
recompensas del mismo.

A mis profesores quienes a través de sus conocimientos me han enseñado
caminos nuevos y metas nuevas en cada materia de la cual he sido parte.

DECLARACIÓN EXPRESA

"La responsabilidad del contenido de este Trabajo Final de Graduación, nos corresponde exclusivamente; y el patrimonio intelectual del mismo a la Escuela Superior Politécnica del Litoral".

(Reglamento de Graduación de la ESPOL)

Mario Nicolás Plaza Ruíz

RESUMEN

El presente trabajo nos ayuda a conocer modelos de preferencia de consumo con respecto a diferentes categorías. La finalidad del mismo es poder determinar el comportamiento de la variable de interés, a través de un conjunto de variables que van brindarnos luces sobre la preferencia de determinadas marcas por sobre otras.

El primer capítulo consta de antecedentes de la empresa que colaboro con el estudio, la definición del problema que se va a abordar en el proyecto, su respectiva justificación y los objetivos que nos motivan a poder llevar adelante el proyecto.

En el segundo capítulo se detalla la metodología que va a ser utilizada en el proyecto. Especificando los tipos de gráficos a ser utilizados en el análisis univariado, así como también una explicación de las diferentes variables presentes en el estudio. Procedemos a continuación la explicación de las variables presentes en el estudio, para culminar con la metodología bivariada y multivariada.

En el tercer capítulo procedemos a la analizar las variables en base a la metodología expuesta en el capítulo anterior. En la sección univariada se muestran los gráficos de interés para mostrar un preámbulo de la situación global de la población objetivo, y poder conocer con el porcentaje de consumo dentro de cada categoría de interés. En la sección inferencial

procedemos a analizar las variables que van a formar parte de cada uno de los modelos de regresión logística multinomial expuestos en cada categoría del estudio, para poder determinar si existe evidencia estadística para determinar si dichas variables pasen a formar parte del modelo o no. Y en la sección multivariada procedemos a utilizar la regresión logística multinomial para cada una de las categorías presentadas en el proyecto.

El informe culmina con las conclusiones y recomendaciones que se consideran de mayor interés en base a los resultados obtenidos.

Contenido

AGRADECIMIENTO	iii
DEDICATORIA	iv
RESUMEN	vi
Introducción.....	1
1.1. Antecedentes.....	2
1.1.1. Antecedentes de la empresa.....	2
1.1.2. Descripción de los Servicios de la empresa	3
1.1.3. MISIÓN	7
1.1.4. VISIÓN.....	7
1.1.5. POLITICAS.....	7
1.1.6. OBJETIVOS.....	9
1.2. Definición del problema.....	9
1.3. Justificación.....	10
1.4. Objetivos.....	11
1.4.1. Objetivo General.....	11
1.4.2. Objetivos específicos.....	11
2.1. Metodología descriptiva	13
2.1.1. Histograma	13
2.1.2. Diagrama de pastel.....	14
2.2. Variables de estudio	14
2.2.1. Descripción y codificación de las variables.....	14
2.2.2. Variables de la Sección I – Identificación Personal.	15
2.2.3. Variables de la Sección II – Posicionamiento (sin categoría).	17
2.2.4. Variables de la Sección III – Posicionamiento por Categoría.	18
2.2.5. Variables de la Sección IV – Orden de importancia.	21
2.3. Metodología Inferencial.....	22
2.3.1. Tablas de Contingencia.	23
2.3.2. Regresión logística	25

2.3.2.1.	Factores de confusión	27
2.3.2.2.	Concepto de interacción	27
2.3.2.3.	Regresión logística binaria	28
2.3.2.4.	Estimación de los coeficientes del modelo	33
2.3.2.5.	El estadístico de Wald	33
2.3.2.6.	El estadístico G de razón de verosimilitud	34
2.3.2.7.	La prueba de Score.....	34
2.3.2.8.	Tipos de modelos de regresión logística.....	35
2.3.2.9.	Medidas de confiabilidad del modelo.....	36
2.3.2.9.1.	Devianza	36
2.3.2.9.2.	Criterio AIC de Akaike	37
2.3.2.9.3.	Prueba de bondad de ajuste de Hosmer-Lemeshov.....	37
2.3.2.10.	Regresión logística multinomial.....	37
2.3.2.11.	Modelos Logit para respuestas nominales	38
2.3.2.12.	Modelos Acumulados para datos ordinales.....	40
3.1.	Análisis univariado	44
3.1.1.	Identificación personal del entrevistado.	45
3.1.2.	Posicionamiento por categoría.....	57
3.1.3.	Variables de la Sección IV – Orden de importancia.	91
3.2.	Análisis inferencial	95
3.3.	Regresión Logística Multinomial	147
	Conclusiones	190
	Recomendaciones.....	196
4.1.	Anexos de Tablas de Contingencia de Telefonía Celular.....	197
4.2.	Anexos de Tablas de Contingencia de Bebidas Gaseosas.	201
4.3.	Anexos de Tablas de Contingencia de Bebidas Alcohólicas.	204
4.4.	Anexos de Tablas de Contingencia de Atún.....	206
4.5.	Anexos de Tablas de Contingencia de Chocolate	207
4.6.	Anexos de Tablas de Contingencia de Café.	209
5.1.	Anexos de programación en R de Telefonía Celular.....	210

5.2.	Anexos de programación en R de Bebidas Gaseosas.	211
5.3.	Anexos de programación en R de Chocolate en Barra.....	213
5.4.	Anexos de programación en R de Café.	214
5.5.	Anexos de programación en R de Atún.....	215
5.6.	Anexos de programación en R de Atún.....	217
	Bibliografía	219

Introducción

Vivimos en un mundo en donde las guerras de mercado se ganan a través de las decisiones de compra de los clientes. Se debe tener en cuenta que nuevas marcas siempre se encuentran apareciendo en el mercado, por lo que es de vital importancia fortalecer las marcas y conseguir permanecer en las mentes de los clientes actuales con miras hacia clientes potenciales.

Debido a los cambios vertiginosos del mundo actual es indispensable que las empresas realicen investigaciones de mercado, para obtener información que les permita pensar y ejecutar estrategias adecuadas para comercializar sus diferentes productos.

Por estas razones este trabajo pretende dilucidar la marca de preferencia en la mente de los ecuatorianos de forma general, y dentro de distintas categorías de productos. Obteniendo información sobre la predilección de las marcas según edad, género, ciudad y ocupación; para así poder determinar qué tipo de publicidad sería la más conveniente en base al nicho de mercado que se encuentra prefiriendo la marca. Además se tiene en consideración la importancia de atributos presentes al momento de elegir una marca.

CAPITULO I

1. Descripción de la empresa

1.1. Antecedentes

1.1.1. Antecedentes de la empresa

La empresa XYZ es una empresa especializada en estudios de mercado de tipo cuantitativo, sistema de investigación que permite sustentar los planes de marketing, producción, tesorería, financiación e inversión de una empresa, el mencionado sistema ofrece muestreos estadísticos con validez del marco muestral.

La empresa es contratada por empresas de diversa índole en todo el país, pues aun cuando sus oficinas se encuentra en Guayaquil, es contratada por empresas que se encuentran localizadas en otras ciudades, es por ello que entre las metas de esta empresa en el futuro es abrir nuevas oficinas en otros ciudades e incluso establecer alianzas estratégicas con empresas extranjeras.

Enfocándose a las necesidades y exigencias de cada cliente, el servicio que ofrece es personalizado, de ésta manera han logrado cumplir con las más altas exigencias de calidad. Cuenta con unos 13 años en el mercado nacional, y realiza proyectos a nivel de todo el país.

1.1.2. Descripción de los Servicios de la empresa

Entre los servicios que oferta la empresa XYZ tenemos los siguientes:

1. Investigaciones de mercado

1.1. Asesoría en estudios de mercado. Tipo de estudio

1.2. Diseño de encuestas y/o formularios.

1.3. Aplicaciones de campo.

1.4. Análisis estadístico.

1.5. Informe y análisis de resultados.

1.6. Estudios de campo

1.6.1. Pantry Check

1.6.2. Distribution Check

1.6.3. Home Panel

1.6.4. Store Audit

1.6.5. Perfil de clientes

1.6.6. Estilos de vida

1.6.7. Investigaciones de producto

1.6.8. Prueba de nuevos productos

1.6.9. Pruebas de concepto

1.6.10. Potencial de mercado

1.6.11. Investigaciones de efectividad publicitaria

1.6.12. Investigaciones de medios

1.7. Sondeos de opinión

1.7.1. Economía

1.7.2. Política

1.7.3. Ambiente financiero

1.7.4. Estudios de demanda especializados

1.8. Área de estudios urbanos

1.8.1. Estudios de satisfacción con los servicios públicos.

1.8.1.1. Estudios de actitud y opinión de los ciudadanos ante diferentes temas.

1.8.1.2. Estudios de demanda de los ciudadanos.

1.8.2. Estudios a grupos sociales

1.8.2.1. Opiniones y actitudes colectivos sociales

1.8.2.2. Estudios de preferencia.

1.8.3. Estudios de opinión pre – electoral.

1.8.3.1. Predicción de voto.

1.8.3.2. Satisfacción del electorado.

1.8.3.3. Imagen de líderes políticos.

1.8.3.4. Barómetro de clima social.

1.8.4. Mapas perceptuales

1.8.4.1. Variables que definen un perfil

1.8.4.2. Comparativo de perfiles

1.8.5. Estudio del nivel de satisfacción de un producto y/o servicio.

1.8.6. Censos

1.8.6.1. Poblacionales

1.8.6.2. De la Microempresa

1.8.6.3. Censos de impacto y reacción del ciudadano.

2. Desarrollo de aplicaciones

2.1. Para Base de Datos

2.2. Actualizaciones

3. Proyectos

- 3.1. Diseño conceptual de la investigación
 - 3.2. Diseño de cuestionarios
 - 3.3. Aplicaciones de prueba
 - 3.4. Desarrollo aplicación de proyecto
 - 3.5. Herramientas del marketing basados en la estadística.
 - 3.6. Generación de resultados de información.
 - 3.7. Definición socio – económica de sus clientes.
 - 3.8. Estudios de perfiles de mercado
 - 3.9. Test de marca
4. Consultores de Apoyo
 - 4.1. La experiencia del componente humano y técnico que conforma Tendencia, permite asesorar a los clientes cuando éstos necesiten convertir a la investigación en herramienta de toma de decisiones en el Mercado.
5. Consultoría en Marketing
 - 5.1. Estrategias de marketing.
 - 5.2. Herramientas para facilitar el acceso al mercado de manera práctica y efectiva.
 - 5.3. Implementación de base de datos.

5.4. Análisis de clientes actuales y potenciales.

5.5. Aplicación de estrategias de mercadeo tradicional.

6. Digitación

6.1. Minería de datos, Tabulación y Depuración para recolección de información correcta para toma de decisiones.

1.1.3. MISIÓN

Proporcionar al cliente alternativas para mejorar su negocio, explotando la investigación del mercado y lograr la satisfacción del cliente para ambos llegar a la excelencia.

1.1.4. VISIÓN

Logar alianzas estratégicas con empresas nacionales y extranjeras para estar a la vanguardia en el mercado nacional e internacional.

1.1.5. POLITICAS

- Proporcionar al cliente información que satisfaga su creciente necesidad de información acerca del mercado en nuestro país.
- Establecer y mantener vínculos de comunicación con los clientes.

- Para la elaboración de un estudio que se realice en varias ciudades del país, la información debe ser recopilada al mismo tiempo.
- El personal que requiera ausentarse del lugar de trabajo deberá solicitar permiso a una de las siguientes autoridades: Gerente General, Gerente de Proyecto, Gerente Comercial.
- Ejecutar sanciones sobre el personal que tarde en incorporarse a sus actividades.
- Establecer incentivos para el personal de mayor rendimiento en el área de producción.
- Implantar incentivos para el personal administrativo que firme contrato con la empresa por el período mínimo de 1 año.
- No se permite el ingreso del personal ajeno a la empresa, durante la jornada laborable.
- Todo el personal está en la obligación de registrar su jornada laboral a través de un mecanismo de control.
- Desarrollar programas de capacitación al personal subcontratado para concientizar al mismo la importancia que conlleva la recolección de los datos.

1.1.6. OBJETIVOS

- Entregar al cliente los estudios solicitados en un lapso no mayor a un mes y medio.
- Incrementar cada seis meses 2% de la participación en el mercado.
- Reducir en 40% los problemas de errores y no respuestas dentro de los cuestionarios entregados por los entrevistadores en cada estudio.
- Disminuir en un plazo de dos meses el 15% del número de atrasos registrados.
- Disminuir en el lapso de un mes el 20% del número de ausencias registradas.
- Destinar el 2.5% sobre los proyectos adquiridos en un mes para el pago de incentivos del personal subcontratado que se destaque en cada proyecto.
- Aumentar el 10% del desempeño laboral en el personal administrativo mediante indicadores de control.

1.2. Definición del problema

El proyecto se lo realiza para determinar las preferencias de productos de consumo masivo en algunas ciudades del Ecuador, a través de la medición del conocimiento de la población objetivo. Primero un análisis de

forma general sobre la primera marca comercial que se le viene a la mente y las razones de la misma; para posteriormente realizar una segmentación según distintas categorías, las cuales son:

- Empresas de telefonía celular.
- Bebidas Alcohólicas (en general).
- Bebidas gaseosas.
- Atún.
- Chocolate en barra.
- Café (en general).

Dentro de cada categoría se desea conocer si la utiliza o consume, según sea el caso; la primera que se le viene a la mente, la de preferencia y las razones de preferencia. Luego se analizan atributos los cuales se solicita se coloquen según el orden de importancia que le otorga cada ente de investigación.

1.3. Justificación

En base al proyecto cada empresa puede determinar si la publicidad que está utilizando realmente llega al nicho de mercado objetivo, en base al marketing efectuado por la misma. Si dicho objetivo interno de la empresa se está consiguiendo, se puede obtener información adicional, sobre el nicho de mercado, tal como su género, edad, profesión y ciudad. Además también se

determina si las preferencias de la marca se basan en la publicidad que se está trabajando o en características externas a la misma. Estos resultados podrían contrastar a las hipótesis internas de la empresa sobre las creencias de la efectividad de su departamento de marketing. Ya que se han dado casos conocidos en los cuales se apunta a determinado nicho de mercado, pero se termina trabajando con otro completamente diferente.

1.4. Objetivos

1.4.1. Objetivo General

El objetivo general del proyecto es poder determinar la probabilidad de preferencia de una marca, y el Top of Mind en base a cuatro variables intrínsecas de los individuos de la población:

1. Género.
2. Edad.
3. Ciudad.
4. Ocupación.

1.4.2. Objetivos específicos

1. Determinar la probabilidad de preferencia de una marca de telefonía y el top of Mind, según género, edad, ciudad y ocupación.

2. Determinar la probabilidad de preferencia de una marca de bebidas alcohólicas y el top of Mind, según género, edad, ciudad y ocupación.
3. Determinar la probabilidad de preferencia de una marca de atún y el top of Mind, según género, edad, ciudad y ocupación.
4. Determinar la probabilidad de preferencia de una marca de chocolate en barra y el top of Mind, según género, edad, ciudad y ocupación.
5. Determinar la probabilidad de preferencia de una marca de café y el top of Mind, según género, edad, ciudad y ocupación.

CAPITULO II

2. Metodología

2.1. Metodología descriptiva

Dentro del análisis univariado se ha escogido los histogramas y pies debido a que son datos categóricos y esta es la mejor forma de poder apreciarlos.

2.1.1. Histograma

Un histograma no es más que una representación gráfica de una variable en forma de barras, en la cual la superficie de cada barra es proporcional a la frecuencia de los valores representados. Usualmente en el eje vertical se encuentran los valores de frecuencia, y en el eje horizontal los valores que toma cada variable; siendo este valor normalmente la marca de clase, que no es más que la mitad del intervalo en que se encuentran los datos. Los histogramas son mayormente utilizados para variables continuas, como por ejemplo la edad de los entrevistados, siendo agrupados los datos en clases o valores continuos. Las variables que van a ser explicadas a través de histogramas o de gráficos de barras son: Edad, Actividad a la que se dedica, Primera marca que se le viene a la mente, Razones de por qué recuerda esa marca, TOM (Top of Mind), Preferencia, Razones de

preferencia y el Orden de importancia de los atributos al momento de elegir una marca. Para esto procedemos a obtener la frecuencia de datos dentro de cada variable que se va a realizar el análisis, para posteriormente proceder a elaborar la gráfica.

2.1.2. Diagrama de pastel

Un gráfico de pie o de pastel es un recurso estadístico utilizado para poder representar porcentajes referentes a características de interés, o su respectiva proporción. Este tipo de gráficos es utilizado cuando al investigador le interesa mostrar la proporción de una determinada característica con respecto al total de la misma. O en aquellas ocasiones en las cuales no interesa tan solo mostrar el número de veces en las que se da una determinada característica, sino que se pueda visualizar la respectiva proporcionalidad. El gráfico de pastel va a ser utilizado con las variables Género y Utilización o consumo dentro de las distintas categorías.

2.2. Variables de estudio

2.2.1. Descripción y codificación de las variables.

El cuestionario está constituido por cuatro secciones, la primera sección que es información personal del entrevistado. La segunda sección trata de la primera marca que se le viene a la mente. La tercera sección consta de las marcas según categorías; siendo ya especificadas en

una categoría determinada, según sean estas de: Empresas de Telefonía, Bebidas gaseosas-Colas, Bebidas alcohólicas, Atún, Chocolate en barra, Café; en donde se especifica la primera que se le viene a la mente, otras marcas, preferencias y razones de preferencias. La cuarta sección es el orden de importancia de determinados atributos a la hora de elegir una marca.

Este estudio consta de un total de 40 variables y mediante la codificación de cada una se podrá desarrollar e interpretar de manera más sencilla y comprensiva, los análisis estadísticos respectivos.

De las 39 variables de este estudio, solo una de ellas es cuantitativa y esa es la variable Edad (Variable X_2), y el resto son todas variables cualitativas.

2.2.2. Variables de la Sección I – Identificación Personal.

Estas variables incluyen datos como nombre, edad, género, ciudad y actividad a la que se dedica cada individuo.

Variable X_1 : Nombre

Esta es una variable cualitativa y nos indica el nombre del entrevistado para poder tener un registro de los entes investigados.

Esta variable no se codifica.

Variable X_2 : Edad

Esta es una variable cuantitativa y numérica ordinal, a través de la cual se registra la edad del entrevistado. Cabe destacar que para efecto de esta

investigación solo se consideraron personas mayores de edad, es decir, mayores a los 18 años.

Esta variable no se codifica.

Variable X₃: Género

Esta variable es cualitativa binomial de carácter nominal y permite identificar el género al cual pertenece el entrevistado.

Codificación:

Hombre	1
Mujer	2

Variable X₄: Ciudad

Ciudad es la variable correspondiente a la ciudad en la cual se realizó el estudio, las cuales fueron:

1. Guayaquil
2. Quito
3. Cuenca
4. Machala
5. Manta

Variable X₅: Actividad

Actividad es una variable cualitativa nominal, permite conocer la actividad profesional o no profesional del entrevistado, para efecto del estudio procedemos a cuantificarla.

2.2.3. Variables de la Sección II – Posicionamiento (sin categoría).

Las variables pertenecientes a esta sección nos indica la primera marca que el entrevistado se le viene a la mente sin necesidad de encasillarla en algún tipo de clasificación, mostrándonos la primera marca que el entrevistado tiene presente.

PREGUNTA 1: De las marcas en general. ¿Cuál es la primera marca comercial que se le viene a la mente en este momento?

Variable X₆: Marca

Mediante esta variable cualitativa de carácter nominal, se establece la primera marca que el entrevistado tiene presente en la mente. Para efectos del estudio procedemos a cuantificarla para su posterior análisis.

PREGUNTA 2: ¿Por qué recuerda esta marca?

Variable X₇: Porque

La variable nominal cualitativa “porque”, establece las razones por las cuales el entrevistado recuerda la marca anteriormente mencionada. Para efectos de este estudio esta variable será cuantificada.

2.2.4. Variables de la Sección III – Posicionamiento por Categoría.

Las variables pertenecientes a esta sección nos indica la primera marca que tiene presente el entrevistado pero en las distintas categorías que han sido seleccionadas. Las categorías presentes en esta sección son 6:

1. Empresas de telefonía celular
2. Bebidas gaseosas – Colas
3. Bebidas alcohólicas (en general)
4. Atún
5. Chocolate en barra
6. Café (en general)

PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

Variable X₈: Utiliza/Consume

Variable X₁₂: Utiliza/Consume

Variable X₁₆: Utiliza/Consume

Variable X₂₀: Utiliza/Consume

Variable X₂₄: Utiliza/Consume

Variable X₂₈: Utiliza/Consume

Por medio de esta variable cualitativa binomial se conoce si los entrevistados utilizan/consumen o no cada uno de los productos o servicios que se

encuentran en cada una de las seis categorías. Si el entrevistado utiliza/consume de los productos o servicios entonces se procede a realizar las siguientes preguntas.

Codificación:

No 1

Si 2

PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

Variable X₉: TOP

Variable X₁₃: TOP

Variable X₁₇: TOP

Variable X₂₁: TOP

Variable X₂₅: TOP

Variable X₂₉: TOP

Mediante esta variable cualitativa de carácter nominal, se establece la primera marca que el entrevistado tiene presente en la mente, para cada una de las diferentes categorías. Para efectos del estudio procedemos a cuantificarla para su posterior análisis.

PREGUNTA 5: ¿Cuál es la marca que usted prefiere?

Variable X₁₀: Prefiere

Variable X₁₄: Prefiere

Variable X₁₈: Prefiere

Variable X₂₂: Prefiere

Variable X₂₆: Prefiere

Variable X₃₀: Prefiere

Estas variables son cualitativas de carácter nominal, e indica la preferencia de alguna marca que tiene el entrevistado en las distintas categorías. Para efectos del presente estudio se procedió a cuantificar las variables.

PREGUNTA 6: ¿Por qué prefiere la marca?

Variable X₁₁: Razones de preferencia

Variable X₁₅: Razones de preferencia

Variable X₁₉: Razones de preferencia

Variable X₂₃: Razones de preferencia

Variable X₂₇: Razones de preferencia

Variable X₃₁: Razones de preferencia

Estas variables son cualitativas de carácter nominal, e indica las razones de preferencia de alguna marca que tiene el entrevistado en las distintas categorías. Para efectos del presente estudio se procedió a cuantificar las variables.

2.2.5. Variables de la Sección IV – Orden de importancia.

Las variables que se encuentran presentes en esta sección nos indican el orden de importancia que tiene para cada entrevistado los atributos que ellos consideran importantes al momento de elegir una marca. Para esto se dieron a elegir a los entrevistados una lista de 9 atributos para que los colocaran en el orden de importancia que ellos le daban a cada uno. La lista de atributos es la siguiente:

1. Color
2. Diseño
3. Tradición
4. Calidad
5. Marca Confiable (Confianza)
6. Envase
7. Presentación
8. Precio
9. Logotipo

PREGUNTA 7: En orden de importancia ¿Cuáles son los atributos que usted considera importantes al momento de elegir una marca?

Variable X₃₂: Primer lugar

Variable X₃₃: Segundo lugar

Variable X₃₄: Tercer lugar

Variable X₃₅: Cuarto lugar

Variable X₃₆: Quinto lugar

Variable X₃₇: Sexto lugar

Variable X₃₈: Séptimo lugar

Variable X₃₉: Octavo lugar

Variable X₄₀: Noveno lugar

Estas variables son cualitativas ordinales que permiten ordenar a los atributos antes mencionados según el orden especificado de cada entrevistado, dándonos a conocer cuáles son las características que cada uno considera importante a la hora de elegir una determinada marca. Para efectos del estudio se procedió a cuantificar las opciones para posteriormente ser colocadas en el respectivo orden asignado.

2.3. Metodología Inferencial.

Se va a proceder a realizar el análisis bivariado con respecto a las variables de interés, la cual va a ser analizada a través de las Tablas de Contingencia, a través de la cual se analiza la independencia de la variable dependiente con respecto a las variables dependientes.

2.3.1. Tablas de Contingencia.

Para poder analizar la relación de dependencia o independencia entre variables cualitativas nominales se debe estudiar su distribución conjunta o tabla de contingencia.

La tabla de contingencia es una tabla de doble entrada, en la cual en cada casilla se va a colocar el número de casos que poseen un nivel de los factores analizados, tanto en las filas como en las columnas.

Tabla I				
Ejemplo de Tabla de Contingencia				
		SEXO		
		HOMBRE	MUJER	MARGINAL
FUMA	SI	n_{11}	n_{12}	$n_{1.}$
	NO	n_{21}	n_{22}	$n_{2.}$
	MARGINAL	$n_{.1}$	$n_{.2}$	N

Fuente: Análisis de datos cualitativos

Elaboración: José Vicéns Otero y Eva Medina Moral

Donde

n_{ij} = número de observaciones que tienen el atributo i y j

$n_{i.}$ = número de individuos que tienen el atributo i (marginal i)

$n_{.j}$ = número de individuos que tienen el atributo j (marginal j)

La tabla de contingencia está definida por la cantidad de atributos que se analizan de forma conjunta y el número de modalidades o niveles de los mismos. En el ejemplo propuesto es una tabla de contingencia de 2x2 (FUMA y SEXO) y cada uno tiene dos niveles.

Normalmente las tablas de contingencia tienen por finalidad lo siguiente:

1. Organizar la información presente en un experimento si es binomial, es decir, dos factores (variables cualitativas).
2. Con la tabla de contingencia se puede analizar si existe dependencia o independencia entre los niveles de las variables. Si dos variables son independientes entonces los valores que tome una de ellas no va a estar influenciado por la modalidad que tome la otra.

Para determinar la dependencia de las variables se utiliza un contraste estadístico basado en el estadístico χ^2 , la cual nos va a permitir afirmar con un nivel de confianza estadístico si los niveles de determinada variable cualitativa influyen en los niveles de la otra variable.

Para poder analizar las variables vamos a hacer uso del nivel de significancia (p), si dicho valor es muy pequeño ($p < 0,05$) entonces vamos a rechazar la

hipótesis nula, aceptando que los atributos son dependientes. Sin embargo, si el nivel de significancia fuera mayor a 0,05, entonces se acepta la hipótesis nula de independencia. Siendo entonces las hipótesis de la siguiente forma:

$H_0 =$ La variable Y es independiente de la variable X

Vs

$H_1 =$ La variable Y no es independiente de la variable X

2.3.2. Regresión logística

La regresión logística es un conjunto de métodos estadísticos que desea conocer cuál es la relación existente entre una variable dependiente cualitativa dicotómica o con más valores con otra u otras variables o covariables explicativas independientes, las cuales pueden ser cualitativas o cuantitativas.

Si las covariables fueran cualitativas entonces deben ser dicotómicas, es decir, tomar valores de 0 para su ausencia y 1 para su presencia. Este tipo de codificación es de suma importancia a la hora de interpretar el modelo. En cambio si la covariable cualitativa tuviera más de dos categorías, pues se debe de transformar en varias covariables cualitativas dicotómicas ficticias o comúnmente llamadas variables dummy, tomando una de las categorías como de referencia, y por lo tanto, cada categoría ingresa de forma individual

en el modelo. Se puede decir que de forma general si la covariable cualitativa posee n categorías, pues se deberán realizar $n-1$ covariables dummies.

Todo modelo de regresión logística nos permite cuantificar la relación que existe entre cada una de las covariables y la variable dependiente, clarificando la existencia de interacción y confusión de las covariables con la variable dependiente, en pocas palabras conocer la *posibilidad* para cada covariable. Y también clasificar individuos dentro de las categorías presente y ausente de la variable dependiente, según la probabilidad de que pertenezca a una de ellas según la presencia de determinadas covariables.

El propósito del análisis es predecir cuál es la probabilidad de que a alguien le ocurra cierto evento, determinando qué variables pesan más para aumentar o disminuir la probabilidad de que a alguien le suceda dicho evento.

Entre los indicadores de importancia para poder apreciar el ajuste del modelo se encuentra el doble logaritmo del estadístico de verosimilitud (likelihood), el cual tiene una distribución similar a χ^2 y compara los valores predichos con los observados en dos momentos, en el modelo sin variables independientes,

es decir solo con la constante, y con las variables predictoras. En conclusión el valor de verosimilitud debería disminuir en ambas instancias, y por ende tender a cero cuando el modelo predice bien.

2.3.2.1. Factores de confusión

En el proceso de selección del modelo más adecuado, se debe de considerar un aspecto adicional, especialmente si el método a usar es el método manual, el cual obliga a todas las variables a entrar en el modelo, siendo el investigador quien paso a paso va construyendo el modelo de regresión más conveniente.

En el proceso de incorporar y eliminar variables, si se elimina una variable de uno de los modelos estimados, se debe observar si el modelo que resulta de excluir dicha variable varía significativamente con el modelo que contenía a la misma. Si sucede esto, entonces la variable que se ha eliminado se considera como un *factor de confusión*, la cual no muestra relación con la variable que se está estudiando pero sí con las otras variables que pueden sí tener significancia con la variable de estudio. Si esto ocurriese lo más conveniente es incluir a dicha variable en el modelo final.

2.3.2.2. Concepto de interacción

A veces se suele introducir términos independientes que son únicos, y también las interacciones entre variables de cualquier

orden, siempre y cuando se considere que éstas sean de interés y que afecten el resultado.

Si se van a introducir este tipo de variables se debe de considerar un orden jerárquico, es decir que si se va a introducir un término de interacción de orden superior ($x \cdot y \cdot z$), también debe introducirse los términos de orden inferior ($x \cdot y, x \cdot z, y \cdot z$), y obviamente los términos independientes de las variables de interacción (x, y, z).

2.3.2.3. Regresión logística binaria

Este tipo de modelo son los de mayor interés por su utilidad en diferentes ciencias. Sigue manteniendo las mismas consideraciones con respecto a las variables que el modelo de regresión logística, la cual es de suma importancia al momento de usar paquetes estadísticos que no recodifiquen las variables de forma automática.

La ecuación fundamental en los modelos de regresión logística es:

$$Pr(y = 1|x) = \frac{\exp(b_0 + \sum_{i=1}^n b_i x_i)}{1 + \exp(b_0 + \sum_{i=1}^n b_i x_i)}$$

En donde:

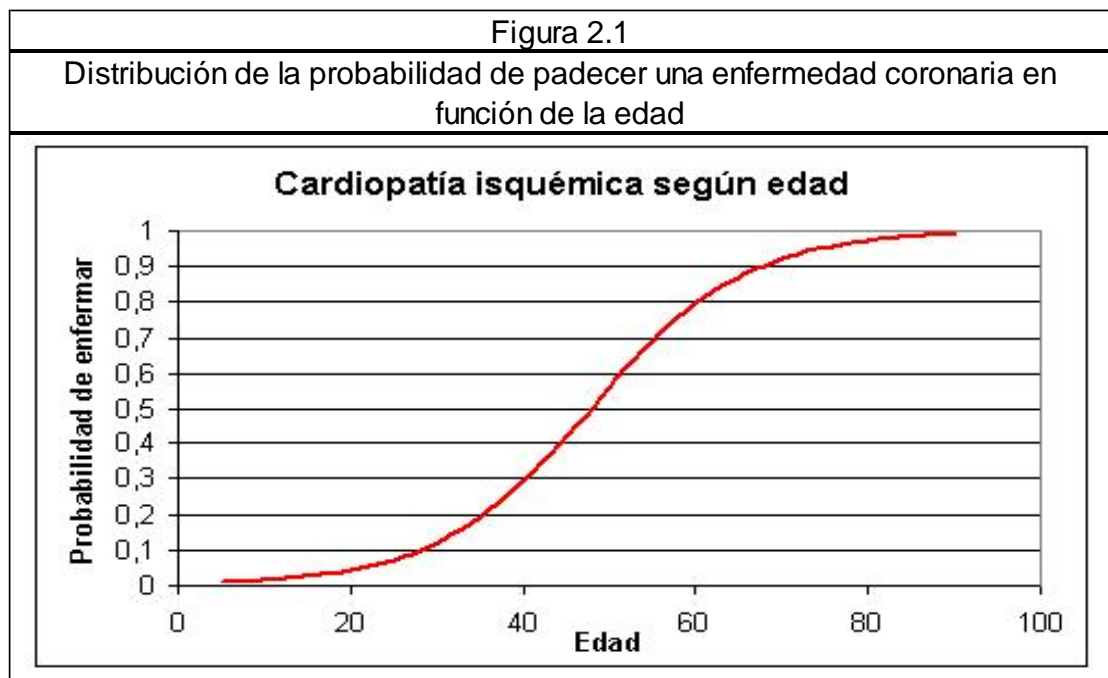
$Pr(y = 1|x)$ es la probabilidad de que y tome el valor de 1 que representa la presencia de la característica estudiada, en presencia de las covariables x ;

x es un conjunto de n covariables $\{x_1, x_2, \dots, x_n\}$ que forman parte del modelo;

b_0 es la constante del modelo o término independiente;

b_i los coeficientes de las covariables.

Un ejemplo gráfico es el siguiente, en el cual se puede apreciar la relación que existe entre la variable dependiente y la covariable edad, que no es una recta sino más bien de forma sigmoidea.



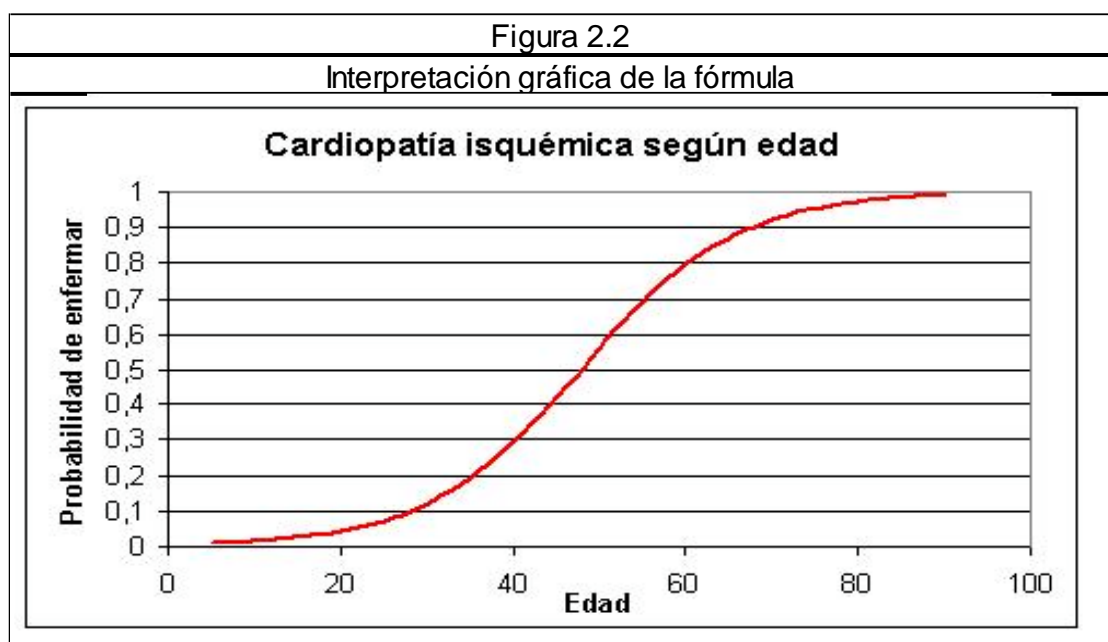
Fuente: Regresión logística

Elaboración: SEQC

Al dividir la expresión anterior por su complementario, es decir, si construimos su *posibilidad* se obtiene la siguiente expresión:

$$\frac{Pr(y = 1|x)}{1 - Pr(y = 1|x)} = \exp\left(b_0 + \sum_{i=1}^n b_i x_i\right)$$

Sin embargo aún esta expresión es compleja de interpretar, por lo cual vamos a usar la representación gráfica que podemos apreciar en la Figura 2.2:



Fuente: Regresión logística

Elaboración: SEQC

Si aplicamos ahora el logaritmo natural, se va a obtener una ecuación lineal de manejo más sencillo y de mejor comprensión:

$$\log\left(\frac{\Pr(y = 1|x)}{1 - \Pr(y = 1|x)}\right) = b_0 + \sum_{i=1}^n b_i x_i$$

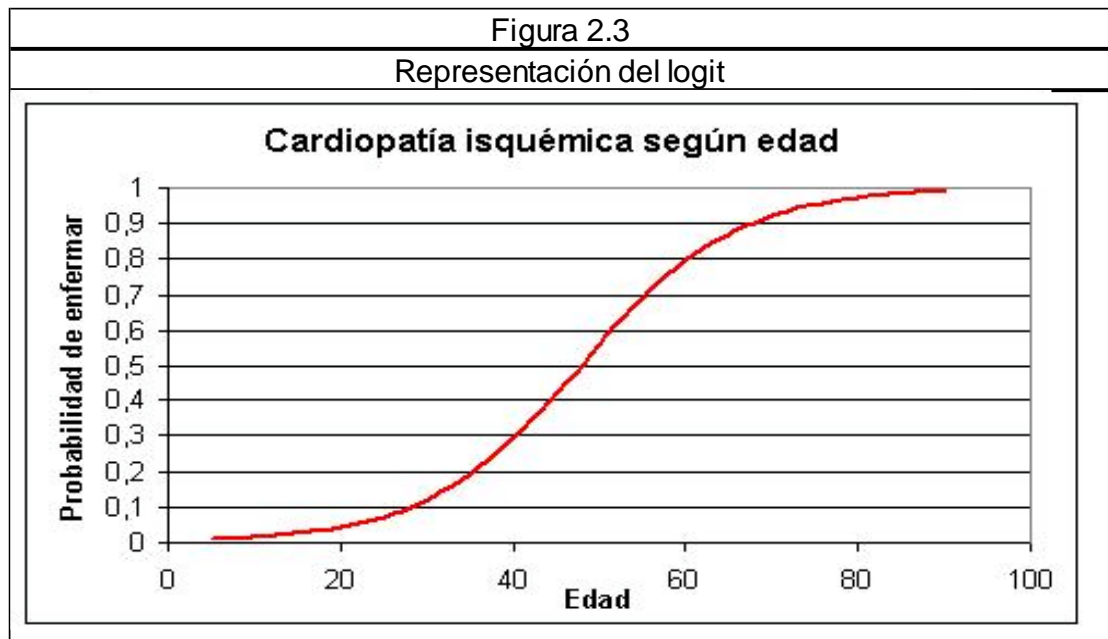
Simplificando tenemos:

$$\log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_i x_i$$

En esta expresión tenemos a la izquierda el llamado *logit*, que no es más que el logaritmo natural de la *posibilidad* de la variable dependiente. Y el término a la derecha de la igualdad es la expresión de una recta, que es idéntica a la del modelo de regresión lineal:

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

Con estas últimas expresiones se puede representar el *logit* de la siguiente forma:



Fuente: Regresión logística

Elaboración: SEQC

Sin embargo existe una gran diferencia entre el modelo de regresión lineal y el de regresión logística, la cual radica en que la regresión lineal asume que los errores estándar de cada coeficiente siguen una distribución normal con media 0 y varianza constante, y en la regresión logística no se puede realizar estas presunciones puesto que la variable no es continua. Por lo que se procede a llamar ε al posible error de predicción para cada covariable x_i el cual va a depender del valor que la variable dependiente llegue a tomar, tal como lo podemos apreciar a continuación:

$$y = Pr(x) + \varepsilon \quad \begin{cases} y = 1 \Rightarrow \varepsilon = 1 - Pr(x) \\ y = 0 \Rightarrow \varepsilon = -Pr(x) \end{cases}$$

Lo que significa que ε sigue una distribución binomial, con media y varianza proporcionales al tamaño muestral y a $Pr(y = 1|x_1)$ que no es más que la probabilidad de que $y = 1$ dada la presencia de x_1 .

2.3.2.4. Estimación de los coeficientes del modelo

Para estimar los coeficientes y sus errores estándar se lo realiza a través de la estimación de máxima verosimilitud, es decir, estimaciones que maximicen la probabilidad de obtener los valores de la variable dependiente Y proporcionados por los datos de la muestra. Dichas estimaciones no se calculan de forma directa sino más bien de forma iterativa como el método de Newton-Raphson. De todos estos métodos se obtienen los coeficientes de regresión, los errores estándar y las covarianzas entre las covariables del modelo.

Para poder comprobar la significación de los coeficientes de regresión se utilizan tres métodos estadísticos: el estadístico de Wald, el estadístico G de razón de verosimilitud y la prueba de Score.

2.3.2.5. El estadístico de Wald

Este estadístico se encarga de contrastar la hipótesis de que un coeficiente aislado sea diferente de 0, siguiendo una distribución normal con media 0 y varianza 1. Su valor viene dado por el cociente entre el

valor del coeficiente y su respectivo error estándar. Lo que nos indica es que si dicho coeficiente es diferente de 0 merece la pena su conservación en el modelo. En cambio si el modelo es de error estándar grande, el estadístico de Wald puede dar error tipo II, por lo que tampoco se recomienda su uso si se emplean variables de diseño.

2.3.2.6. El estadístico G de razón de verosimilitud

Este estadístico contrasta cada modelo que surge de eliminar cada covariable frente al modelo completo. El estadístico G sigue una χ^2 con un grado de libertad, por lo que no se asume normalidad. La ausencia de significación implica que da igual si la variable se encuentra presente o no, por lo que según el principio de parsimonia dicha variable debe ser eliminada puesto que no aporta en nada. Esta prueba no asume ninguna distribución, por lo que se vuelve la más recomendada para estudiar la significación de los coeficientes.

2.3.2.7. La prueba de Score

El cálculo para este estadístico para el caso de una única variable viene dada por:

$$S = \frac{\sum_{i=1}^n x_i(y_i - \bar{y})}{\sqrt{\bar{y}(1 - \bar{y}) \sum_{i=1}^n (x_i - \bar{x})^2}}$$

Si el caso fuera que se trataran de múltiples covariables se debe utilizar entonces cálculo matricial, que a pesar de que no se requiera un cálculo iterativo, se tienen dos puntos en contra:

1. Este estadístico se va incrementando según aumenta el número de covariables.
2. También asume normalidad con media 0 y varianza 1.

De la misma forma que los estadísticos anteriores, si alcanza la significación nos indica que dicha covariable debe permanecer en el modelo. Cuando la covariable es cualitativa con n categorías, con $n > 2$, se va a analizar el modelo la significación de cada una de sus $n - 1$ variables ficticias, así como la significación global de la covariable comparando la presencia de la misma en bloque frente a la ausencia en bloque de sus $n - 1$ variables ficticias.

2.3.2.8. Tipos de modelos de regresión logística

Modelo logístico univariante simple

$$y = \frac{\exp(a + bx)}{1 + \exp(a + bx)}$$

Modelo logístico multivariante simple

$$y = \frac{\exp(\beta_{i,0} + \beta_1 x_1 + \beta_{i,2} x_2 + \dots + \beta_{i,k} x_k)}{1 + \exp(\beta_{i,0} + \beta_1 x_1 + \beta_{i,2} x_2 + \dots + \beta_{i,k} x_k)}$$

Modelo logístico multivariante simple

$$y_i = \frac{\exp(a + bx)}{1 + \exp(a + bx)}$$

Modelo logístico multivariante múltiple

$$y_i = \frac{\exp(\beta_{i,0} + \beta_1 x_1 + \beta_{i,2} x_2 + \dots + \beta_{i,k} x_k)}{1 + \exp(\beta_{i,0} + \beta_1 x_1 + \beta_{i,2} x_2 + \dots + \beta_{i,k} x_k)}$$

2.3.2.9. Medidas de confiabilidad del modelo

2.3.2.9.1. Devianza

Es parecida a la suma de cuadrados del error de la regresión lineal y se define de la siguiente forma:

$$D = -2 \sum_{i=1}^n \left(y_i \log\left(\frac{\hat{p}}{y_i}\right) + (1 - y_i) \log\left(\frac{1 - \hat{p}}{1 - y_i}\right) \right)$$

Si D es mayor que una χ^2 con $n - p$ grados de libertad para un nivel de significación dado entonces el modelo logístico es confiable.

2.3.2.9.2. Criterio AIC de Akaike

Se define como

$$AIC = D + 2(p + 1)$$

Donde p es el número de variables predictoras.

2.3.2.9.3. Prueba de bondad de ajuste de Hosmer-Lemeshov

Se define como

$$c = \sum_{i=1}^g \frac{(O_i - n'_i \bar{p}_i)^2}{n'_i \bar{p}_i (1 - \bar{p}_i)}$$

En donde

g es el número de grupos;

n'_i es el número de observaciones en el i -ésimo grupo;

O_i es la suma de las Y en el i -ésimo grupo; y

\bar{p}_i es el promedio de las p_i en el i -ésimo grupo.

2.3.2.10. Regresión logística multinomial

Consideremos a un modelo de regresión logística como aquel cuya variable dependiente de tipo nominal tiene más de dos categorías, siendo una extensión de la regresión logística binaria clásica. Las

variables de respuesta pueden ser continuas (regresores) como también categóricas (factores).

2.3.2.11. Modelos Logit para respuestas nominales

Para este tipo de modelos vamos a nombrar a J como la cantidad de categorías existentes en la variable Y y $\{\pi_1, \dots, \pi_J\}$ serán las probabilidades de las diferentes respuestas que van a satisfacer $\sum_j \pi_j = 1$

Si se sabe que contamos con n observaciones diferentes localizadas en las J categorías, comprenderemos que la distribución de probabilidad de las observaciones presentes en las J categorías va a seguir una distribución multinomial. La cual modeliza la probabilidad de cada una de las posibles formas en que n observaciones puedan repartirse entre las J categorías.

Como se sabe que la escala es nominal, entonces el orden entre las distintas categorías es irrelevante. Se procede a tomar una categoría como base, en este caso tomaremos la última que es J , y vamos a definir un modelo logit con respecto a ella:

$$\log\left(\frac{\pi_j}{\pi_J}\right) = \alpha_j + \beta_j x$$

Donde $j = 1, \dots, J - 1$

De esta fórmula podemos concluir que el modelo cuenta con $J - 1$ ecuaciones, cada una con su propio parámetro, y los efectos van a variar dependiendo de la categoría que se haya considerado como base.

Cuando $J = 2$, entonces el modelo equivale a una única ecuación $\log(\pi_1/\pi_2) = \text{logit}(\pi_1)$ obteniéndose de esta forma el modelo de regresión logística estándar.

Los logits de cualquier pareja de categorías son determinados a través de la ecuación general logit con respecto a la categoría base J . Considerando entonces dos categorías a y b tenemos que:

$$\begin{aligned} \log\left(\frac{\pi_a}{\pi_b}\right) &= \log\left(\frac{\pi_a/\pi_J}{\pi_b/\pi_J}\right) = \log\left(\frac{\pi_a}{\pi_J}\right) - \log\left(\frac{\pi_b}{\pi_J}\right) = (\alpha_a + \beta_a x) - (\alpha_b + \beta_b x) \\ &= (\alpha_a - \alpha_b) + (\beta_a - \beta_b x) \end{aligned}$$

Haciéndolo así, entonces la ecuación que se va a usar para las categorías a y b van a tener la forma $\alpha + \beta x$ en donde $\alpha = (\alpha_a - \alpha_b)$ y $\beta = (\beta_a - \beta_b)$.

2.3.2.12. Modelos Acumulados para datos ordinales

Al tener las respuestas de la variable categórica de forma ordinal entonces se puede utilizar el modelo logit acumulados.

La probabilidad acumulada de una variable Y no es más que la probabilidad de que Y sea menor o igual que un determinado valor j . En consecuencia, para una categoría j podemos definir a la probabilidad acumulada como

$$P(Y \leq j) = \pi_1 + \dots + \pi_j$$

Para $j = 1, \dots, J$

Las probabilidades acumuladas reflejan el orden entre las diferentes categorías:

$$P(Y \leq 1) \leq P(Y \leq 2) \leq \dots \leq P(Y \leq J) = 1$$

Tenemos entonces que los logits de las probabilidades acumuladas son:

$$\text{logit}[P(Y \leq j)] = \log \left[\frac{P(Y \leq j)}{1 - P(Y \leq j)} \right] = \log \left[\frac{\pi_1 + \dots + \pi_j}{\pi_{j+1} + \dots + \pi_J} \right] = \log \left[\frac{\sum_{i=1}^j \pi_i}{\sum_{i=j+1}^J \pi_i} \right]$$

Para $j = 1, \dots, J - 1$

El modelo logit de posibilidades proporcionales, al considerar d variables explicativas $x = (x_1, \dots, x_d)$, podemos expresarlo como:

$$\text{logit}[P(Y \leq j|x)] = \alpha_j + \sum_{i=1}^d \beta_i x_i$$

Para $j = 1, \dots, J - 1$

Debemos de saber que cada variable independiente solo tiene un coeficiente el cual no depende del valor j . Conocemos que la dependencia con respecto al valor j aparece únicamente en el coeficiente α_j .

Si consideramos que $j' < j$ entonces

$$\text{logit}[P(Y \leq j'|x)] \leq \text{logit}[P(Y \leq j|x)]$$

Tenemos que verificar que $\alpha_{j'} \leq \alpha_j$, en conclusión los valores de α_j aumentan paralelamente con los valores de j .

Vamos a considerar dos vectores de variables independientes

$$x_1 = (x_{11}, \dots, x_{1d})'$$

$$x_2 = (x_{21}, \dots, x_{2d})'$$

Entonces

$$\text{logit}[P(Y \leq j|x_1)] - \text{logit}[P(Y \leq j|x_2)] = \sum_{i=1}^d \beta_i (x_{1i} - x_{2i}).$$

Sin embargo, también tenemos

$$\text{logit}[P(Y \leq j|x_1)] - \text{logit}[P(Y \leq j|x_2)] = \log \left[\frac{\frac{P(Y \leq j|x_1)}{P(Y > j|x_1)}}{\frac{P(Y \leq j|x_2)}{P(Y > j|x_2)}} \right]$$

Lo cual lo podemos reescribir de la siguiente forma

$$\log \left[\frac{\frac{P(Y \leq j|x_1)}{P(Y > j|x_1)}}{\frac{P(Y \leq j|x_2)}{P(Y > j|x_2)}} \right] = \sum_{i=1}^d \beta_i (x_{1i} - x_{2i})$$

La parte izquierda de la ecuación se la conoce como la *posibilidad acumulada*.

De esta manera, las posibilidades para un valor de respuesta menor o igual que j para $x = x_1$ es igual a $\exp(\sum_{i=1}^d \beta_i (x_{1i} - x_{2i}))$ veces las posibilidades para $x = x_2$ y dicha proporcionalidad no va a depender del valor j . Por esta razón se denomina al modelo de *posibilidades proporcionales*.

Vamos a suponer que los vectores x_1 y x_2 son iguales a excepción de la i -ésima componente. En otras palabras, ambos entes de investigación coinciden en casi todos los valores, salvo en dicha componente. Aparte de esto, vamos a suponer que difieren en una unidad en la i -ésima componente. Bajo dichas hipótesis tenemos la siguiente ecuación:

$$\frac{\frac{P(Y \leq j|x_1)}{P(Y > j|x_1)}}{\frac{P(Y \leq j|x_2)}{P(Y > j|x_2)}} = e^{\beta_i(x_{1i}-x_{2i})} = e^{\beta_i}$$

De esto podemos deducir que si mantenemos a las demás variables constantes, el cambio de las posibilidades de la función de distribución de Y condicionada a x_1 y a x_2 va a ser igual a e^{β_i} .

CAPITULO III

3. Análisis de datos

3.1. Análisis univariado

Este capítulo denominado Análisis Estadístico, tiene como objetivo fundamental el de responder a varias inquietudes que se tienen con respecto a las principales marcas que se encuentran presentes en la mente de los consumidores y cuáles de éstas son las de su preferencia.

El análisis estadístico univariado fue desarrollado para las 40 variables presentes en esta investigación, con la finalidad de conocer el comportamiento de cada una de ellas, entre los análisis efectuados se encuentran los de medidas de tendencia central, las medidas de dispersión o a través de tablas de frecuencia relativa y absoluta. Cabe mencionar que también se realizaran gráficos que nos ayuden a visualizar y comprender mejor los resultados obtenidos de dicha investigación y respectivo análisis.

El análisis univariado presente en este capítulo estará dividido en cuatro partes: La primera sección que es información personal del entrevistado. La

segunda sección trata de la primera marca que se le viene a la mente. La tercera sección consta de las marcas según categorías; siendo ya especificadas en una categoría determinada, según sean estas de: Empresas de Telefonía, Bebidas gaseosas-Colas, Bebidas alcohólicas, Atún, Chocolate en barra, Café; en donde se especifica la primera que se le viene a la mente, otras marcas, preferencias y razones de preferencias. La cuarta sección es el orden de importancia de determinados atributos a la hora de elegir una marca.

3.1.1. Identificación personal del entrevistado.

Aquí se va a presentar el análisis estadístico univariado de las variables de la Sección I, Identificación personal del entrevistado, las cuales son:

- ❖ Edad
- ❖ Género
- ❖ Ciudad
- ❖ Actividad a la que se dedica

De éstas variables, solo la variable edad es cuantitativa, por lo que, procederemos a realizar un análisis de su estadística descriptiva, distribución

de frecuencias y además, se le aplicará la Prueba de Bondad de Ajuste donde se planteará como Hipótesis Nula que la variable Edad sigue una distribución Normal.

Variable X₂: EDAD

La variable EDAD es de tipo cuantitativa, por lo tanto se realizó el análisis descriptivo respectivo de la misma (véase Tabla II); en lo que respecta a las medidas de tendencia central, la edad promedio de los entrevistados es 35,18 años de edad, la mediana indica que el 50% de los entrevistados tienen una edad menor o igual a 33 años, también se puede observar que la edad de mayor frecuencia (moda) es 20 años. Además se puede establecer que existe al menos un entrevistado con una edad mínima de 18 años y alguien con una edad máxima de 70 años de edad; en cuanto a las medidas de dispersión 12,67 años es la desviación estándar de los datos con respecto a la media.

TABLA II
Estadística Descriptiva de la Edad

Media		35,18
Mediana		33
Moda		20
Desviación Standard		12,67
Varianza		160,49
Sesgo		0,68
Kurtosis		-0,19
Edad Mínima		18
Edad Máxima		70
Primer Cuartil	25%	25
Segundo Cuartil	50%	33
Tercer Cuartil	75%	43

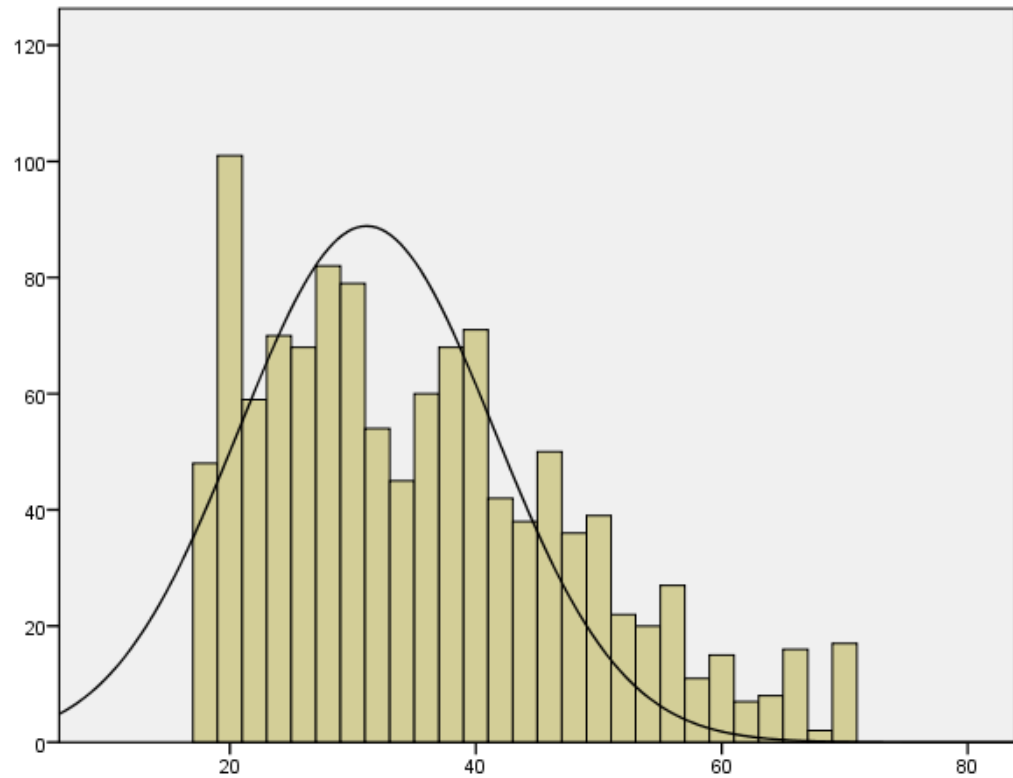
Fuente y Elaboración: Mario Plaza

La distribución de esta variable es asimétrica hacia la derecha, pues tiene como resultado un sesgo con valor positivo el cual es 0,68 lo cual permite establecer que la mayor concentración de los datos se encuentra a la izquierda de la media (Véase Gráfico 3.1).

El primer cuartil indica que el 25% de los entrevistados tiene una edad inferior a los 25 años, el segundo cuartil que el 50% tiene menos de 33 años y el tercer cuartil que el 75% de los entrevistados son menores de 43 años.

Gráfico 3.1

Histograma de frecuencias de la edad de los entrevistados



Fuente y Elaboración: Mario Plaza

Variable X₃: GÉNERO

En el Gráfico 3.2 podemos apreciar que el 59% de la población objetivo son mujeres y que el 41% restante son hombres. Es lógico presenciar que el porcentaje de mujeres sea más alto que el de los hombres, debido a que la encuesta fue efectuada en horario de trabajo, por lo que en su gran mayoría se encontraban amas de casa. Sin embargo el proyecto también fue

realizado en días de fin de semana por lo que explica el porcentaje de hombres.

Tabla III		
Distribución de Frecuencia del Género		
Género	Número de entrevistados	Frecuencia Relativa
Hombre	477	0,41
Mujer	678	0,59
Total	1155	1,00

Fuente y Elaboración: Mario Plaza



Variable X₄: CIUDAD

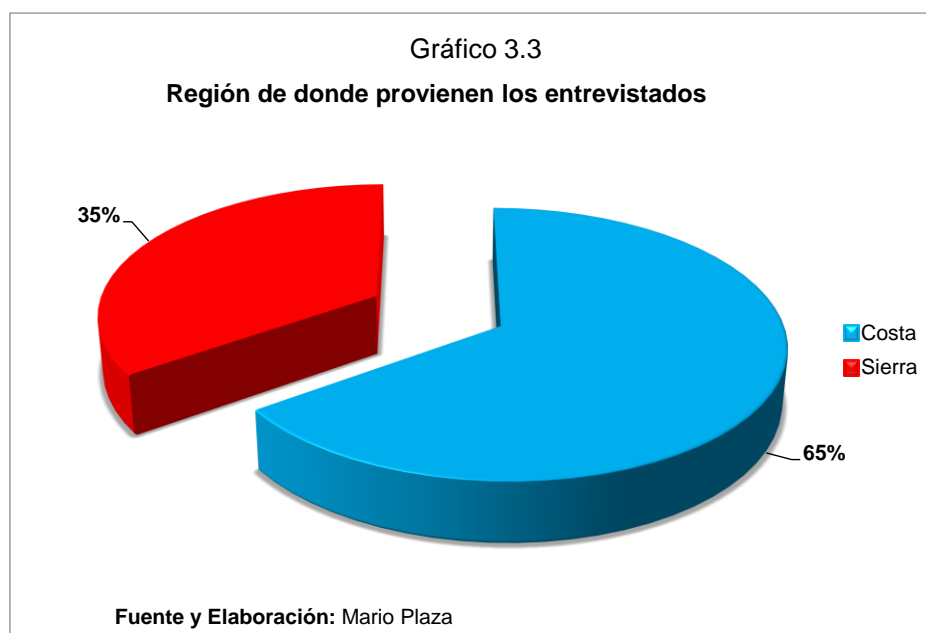
El estudio fue realizado en cinco ciudades: Guayaquil, Quito, Cuenca, Machala, Manta. Por cuestiones de análisis se van a recodificar las variables

para presentar a las ciudades en base a la región a la cual pertenecen: Costa y Sierra.

En el Gráfico 3.3 podemos apreciar que el 65% de los entes investigados pertenecen a la Costa y el 35% restante pertenece a la región Sierra.

Tabla IV		
Distribución de Frecuencia de la Región		
Región	Número de entrevistados	Frecuencia Relativa
Costa	753	0,65
Sierra	402	0,35
Total	1155	1

Fuente y Elaboración: Mario Plaza



Variable X₅: ACTIVIDAD

Debido a la cantidad de Actividades ejecutadas por los entes de investigación se determinó separar en dos grupo a la misma. En el primer grupo las Actividades con mayor peso porcentual, y en el segundo grupo las de menor peso. Esto lo podemos apreciar en las siguientes tablas y gráficos: Tabla V, Tabla VI, Gráfico 3.4 y Gráfico 3.5.

Tabla V		
Distribución de Frecuencia de la Actividad		
Actividad	Número de entrevistados	Frecuencia Relativa
Ama de casa	259	0,33
Estudiante	190	0,24
Comerciante	189	0,24
Empleado Público	147	0,19
Total	785	1,00

Fuente y Elaboración: Mario Plaza

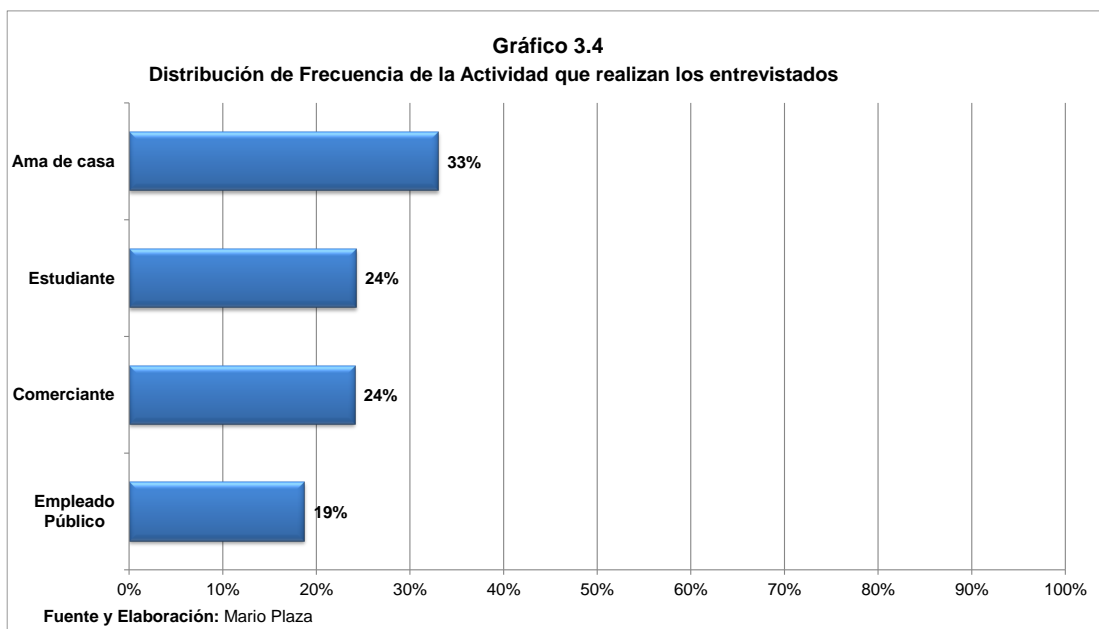
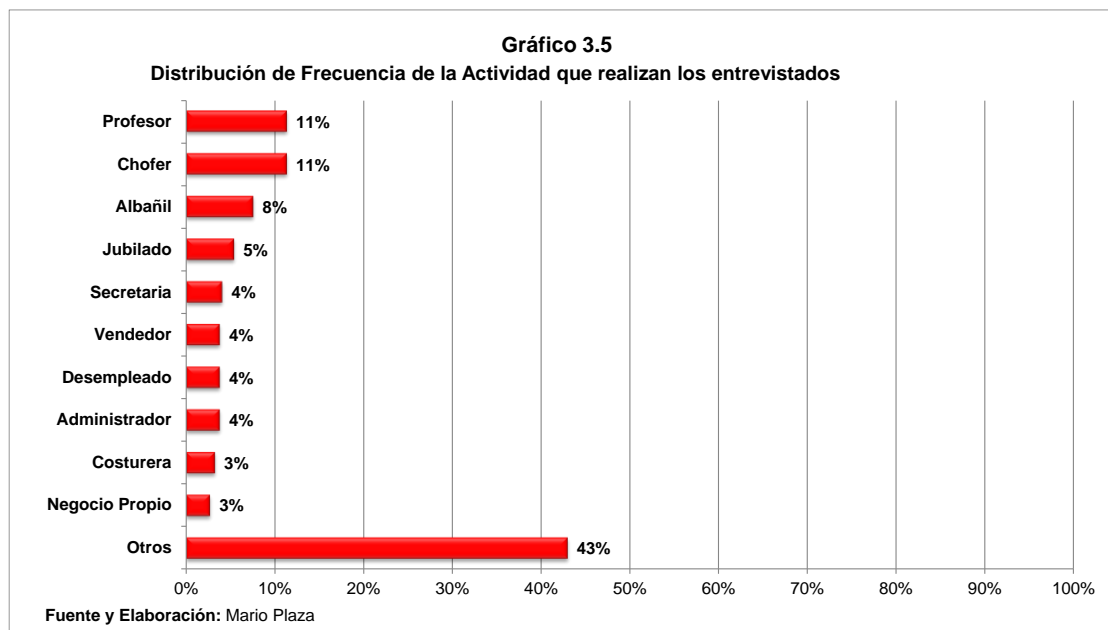


Tabla VI		
Distribución de Frecuencia de la Actividad		
Actividad	Número de entrevistados	Frecuencia Relativa
Chofer	42	0,11
Profesor	42	0,11
Albañil	28	0,08
Jubilado	20	0,05
Secretaria	15	0,04
Administrador	14	0,04
Desempleado	14	0,04
Vendedor	14	0,04
Costurera	12	0,03
Negocio Propio	10	0,03
Otros	159	0,43
Total	370	1,00

Fuente y Elaboración: Mario Plaza



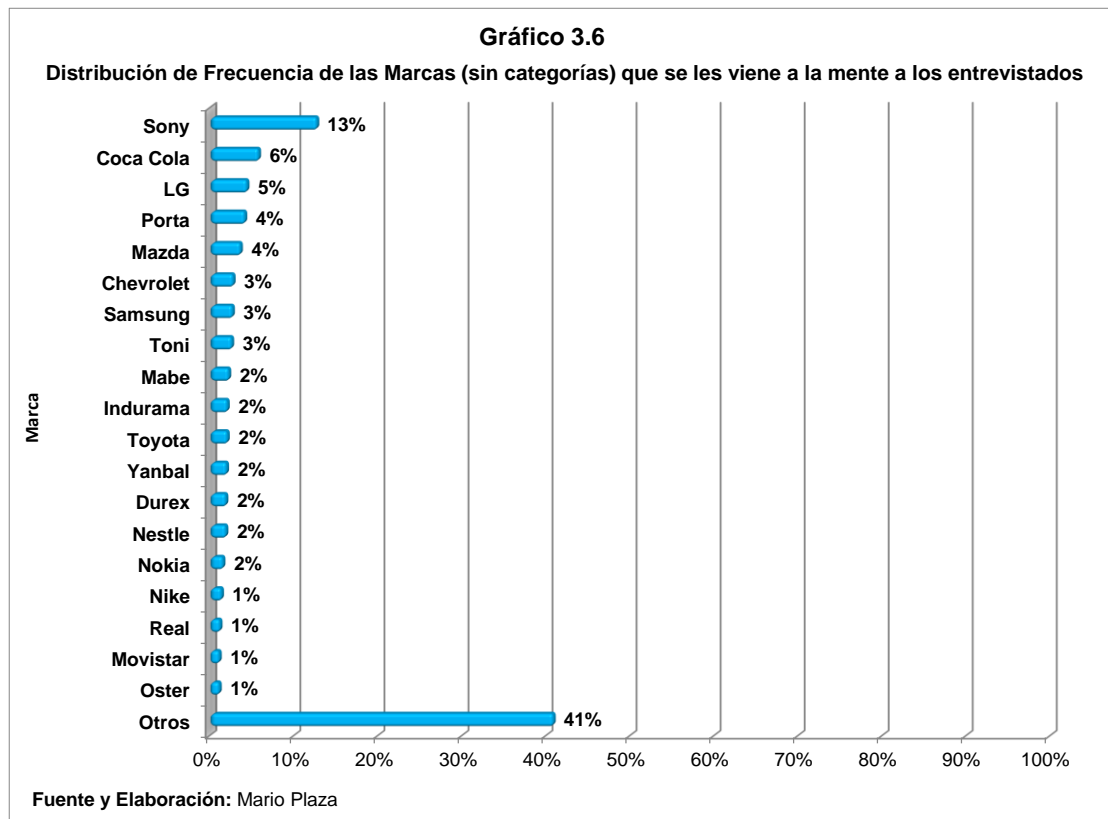
Variable X₆: Marca

Esta variable permitirá conocer cuál es la marca que se encuentra presente en la mente de los entrevistados sin que exista ningún tipo de direccionamiento.

Los resultados de esta variable lo podemos apreciar en la Tabla VII y en el Gráfico 3.6. Como podemos apreciar, la marca que más es recordada en la mente de los entrevistados es Sony con un total de 148 entrevistados que representa el 13%, seguido de Coca Cola con 68 entrevistados representado un 6%, en tercer lugar tenemos a LG con 52 entes de investigación y tiene un 5%, la categoría de Otros conlleva un gran porcentaje debido a que se han agrupado una gran cantidad de marcas cuyos pesos de forma unitaria no aporta en nada con información.

Tabla VII		
Distribución de Frecuencia de la Marca (sin categoría)		
Marca	Número de entrevistados	Frecuencia Relativa
Sony	148	0,128
Coca Cola	68	0,059
LG	52	0,045
Porta	49	0,042
Mazda	43	0,037
Chevrolet	33	0,029
Samsung	32	0,028
Toni	31	0,027
Mabe	27	0,023
Toyota	25	0,022
Indurama	25	0,022
Yanbal	24	0,021
Nestle	23	0,020
Durex	23	0,020
Nokia	19	0,016
Nike	17	0,015
Real	15	0,013
Oster	14	0,012
Movistar	14	0,012
Otros	473	0,410
Total	1155	1

Fuente y Elaboración: Mario Plaza



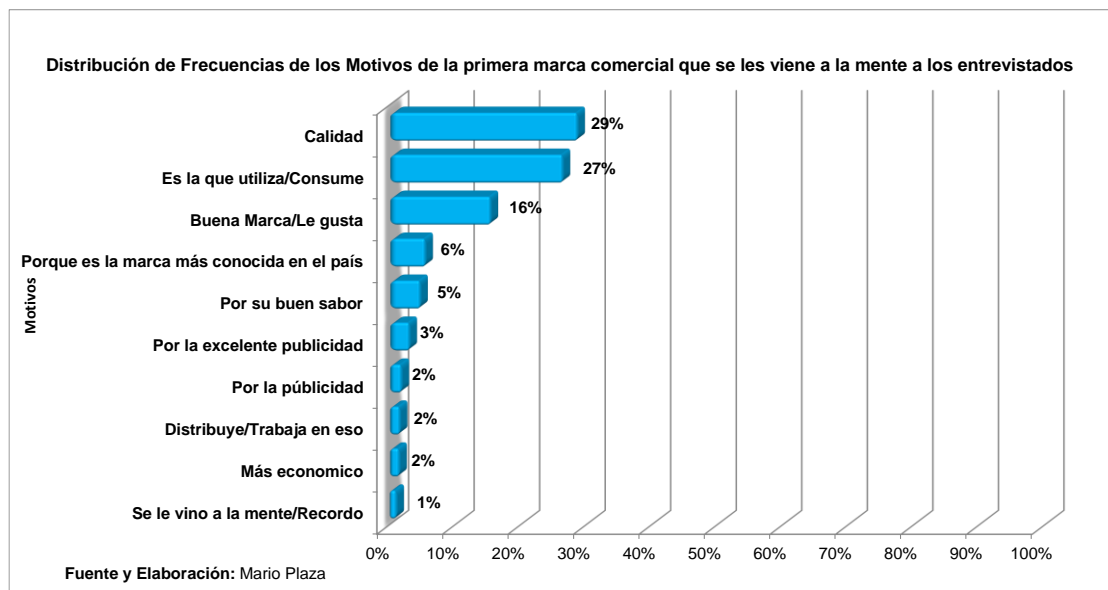
Variable X7: MOTIVO

Los motivos por los cuales los entrevistados recordaron las marcas en la pregunta anterior lo tenemos detallado en la Tabla VIII y el Gráfico 3.7. Se pueden apreciar tres motivos que sobresalen entre los demás, indicando que son las razones principales por las cuales recuerdan a determinadas marcas, la primera marca es Calidad con 333 entrevistados que representa un 29%, la segunda porque Es la que utiliza/Consume con 307 entrevistados, demostrando el hecho de que alguien utilice determinada marca ayuda a que ésta se encuentre presente en la mente del usuario; y en tercer lugar que Le

gusta, independientemente de si la usa o no, el hecho de que sienta un gusto por la marca ayuda a que la tenga en su memoria.

Distribución de Frecuencia de los Motivos de la primera marca comercial que se le viene a la mente		
Motivos	Número de entrevistados	Frecuencia Relativa
Calidad	333	0,2883
Es la que utiliza/Consume	307	0,2658
Buena Marca/Le gusta	180	0,1558
Porque es la marca más conocida en el país	65	0,0563
Por su buen sabor	56	0,0485
Por la excelente publicidad	37	0,0320
Por la publicidad	23	0,0199
Distribuye/Trabaja en eso	20	0,0173
Más economico	19	0,0165
Se le vino a la mente/Recordo	14	0,0121
Por sus modelos	13	0,0113
Por su tecnología	10	0,0087
Otros	78	0,0675
Total	1155	1

Fuente y Elaboración: Mario Plaza



3.1.2. Posicionamiento por categoría.

Posicionamiento por categoría forma parte de la Sección III del formulario. En esta sección deseamos conocer: el uso o consumo de ciertos productos referentes a 6 categorías, la primera marca que se encuentra presente en la mente de los entrevistados, la marca de preferencia y las razones de la misma. Las 6 categorías son:

1. Empresas de Telefonía Celular
2. Bebidas gaseosas – Colas
3. Bebidas alcohólicas (en general)
4. Atún
5. Chocolate en barra
6. Café (en general)

Para esto se realizan cuatro preguntas para cada una de las categorías mencionadas anteriormente, las preguntas fueron realizadas de forma general, de tal manera que va adaptándose dependiendo de la categoría a la cual se refiere. En total esta sección está conformada por 24 variables que son de carácter cualitativas y que procedemos a cuantificarlas por objeto de estudio, y por lo tanto se les realizará un análisis de frecuencia mediante tablas de distribución y gráficos estadísticos.

Vamos a proceder el análisis de cada una de las categorías, empezando por Empresas de Telefonía Celular.

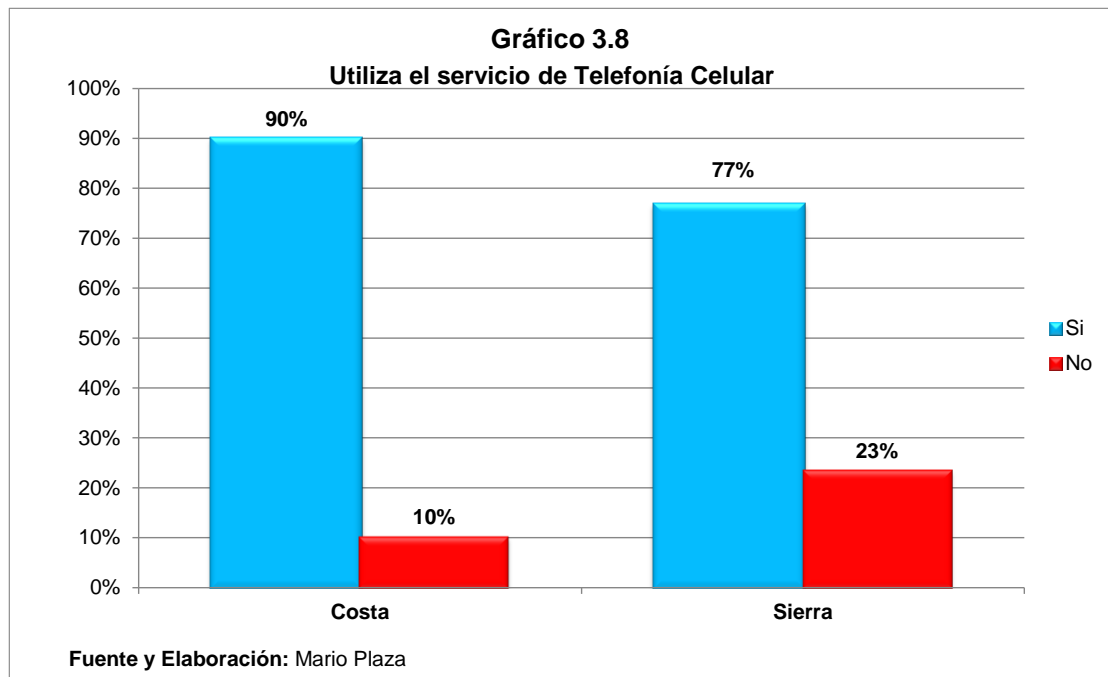
PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

Variable X₈: Utiliza/Consume

La variable X₈ trata sobre la utilización de telefonía celular. En la Tabla IX y el Gráfico 3.8 podemos apreciar que del total de entrevistados en la Costa existen 677 que si utilizan el servicio de alguna empresa de telefonía celular y en la Sierra 308; en cambio 76 individuos en la Costa no usan celular y en la Sierra 94. De la base que se forman con los que Si utilizan procedemos con el análisis.

Tabla IX				
Distribución de Frecuencia del Uso de Telefonía Celular				
Uso de Telefonía Celular	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Si	677	308	0,90	0,77
No	76	94	0,10	0,23
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



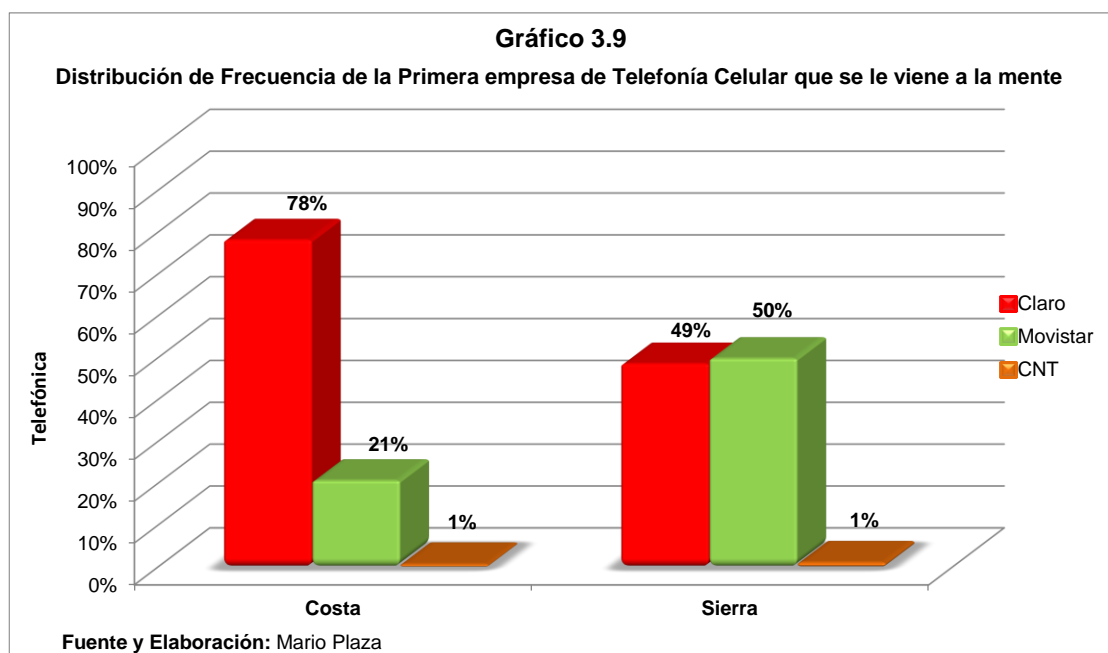
PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

Variable X₉: TOP

La variable X₉ nos da a conocer la primera compañía de Telefonía celular que se encuentra en la mente del entrevistado. Como podemos apreciar en la Tabla X y el Gráfico 3.9, la primera empresa que se le viene a la mente a los entrevistados es Claro en la Costa con un 78%, en la Sierra con un 49%, seguida de Movistar en la Costa con un 21% y en la Sierra un 50%; y por último lugar CNT con tan solo un 1% en ambas regiones.

Tabla X				
Tabla de Frecuencia de la Primera empresa de Telefonía Celular que se le viene a la mente				
Empresas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Claro	530	150	0,78	0,49
Movistar	141	154	0,21	0,50
CNT	6	4	0,01	0,01
Total	677	308	1	1

Fuente y Elaboración: Mario Plaza



PREGUNTA 5: ¿Cuál es la marca que usted prefiere?

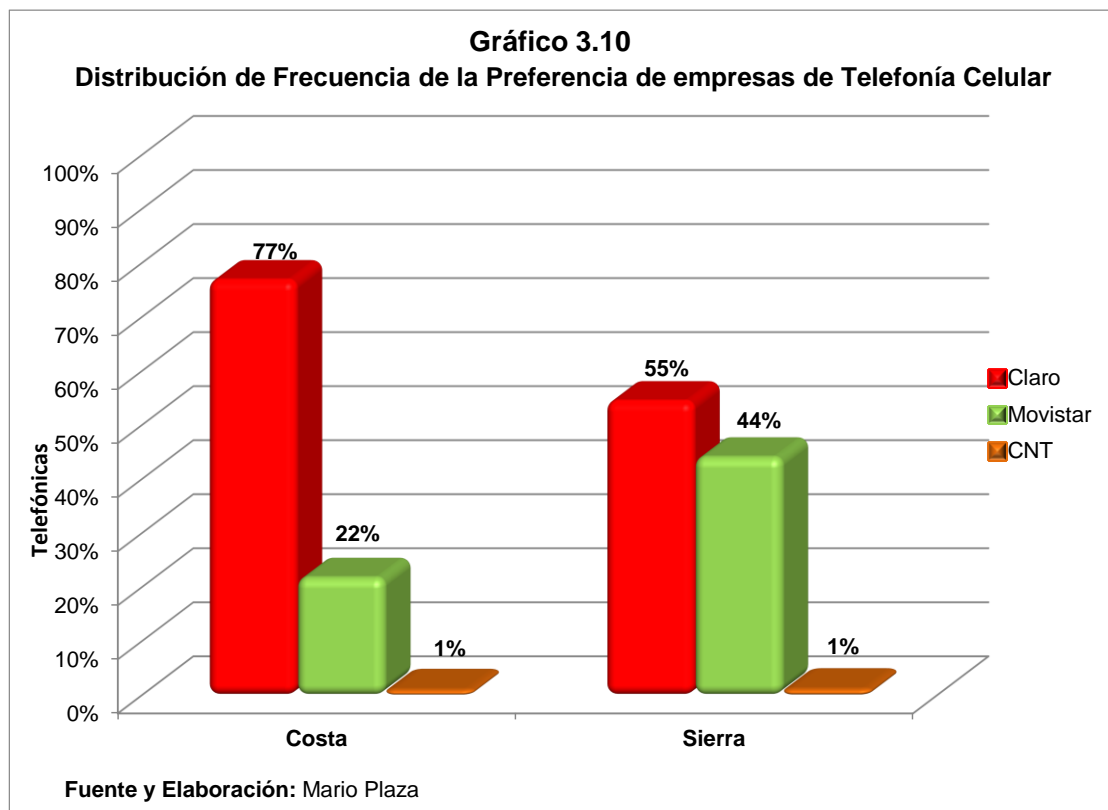
Variable X_{10} : Prefiere

La variable X_{10} tiene que ver con la empresa de telefonía celular que el entrevistado prefiere. Dicha respuesta la podemos apreciar en la Tabla XI y en el Gráfico 3.10, en el cual la empresa Claro cuenta en la Costa con el 77% de preferencia y en la Sierra 55%, seguida de la empresa Movistar con un

22% en la Costa y 44% en la Sierra, y al final CNT con un 1% en ambas regiones. No resulta sorpresa que el TOP y Preferencia mantengan relación a tal punto que en sus porcentajes lleven una alta similaridad. Y vamos a ver las razones en las cuales se sustenta esta hipótesis con la siguiente variable.

Tabla XI				
Tabla de Frecuencia de la empresa de Telefonía Celular que Prefiere				
Empresas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Claro	521	168	0,77	0,55
Movistar	148	136	0,22	0,44
CNT	8	4	0,01	0,01
Total	677	308	1	1

Fuente y Elaboración: Mario Plaza

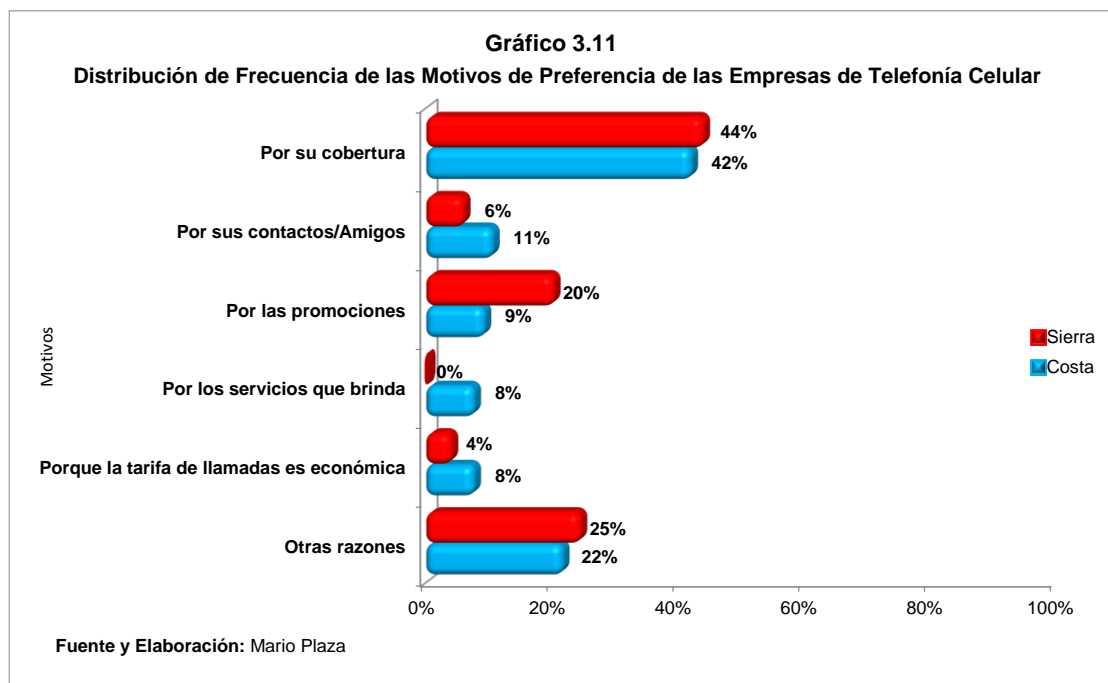


PREGUNTA 6: ¿Por qué prefiere la marca?**Variable X₁₁: Razones de preferencia**

La variable X₁₁ nos permite conocer todas las razones por las cuales el entrevistado prefiere tal o cual marca. Para este propósito tenemos la Tabla XII y el Gráfico 3.11, en las cuales podemos apreciar que una de las razones con más peso porcentual para los entrevistados es Por su cobertura con 42% en la Costa y un 44% en la Sierra, seguida de Por sus contactos/Amigos con un 11% en la Costa y un 6% en la Sierra, después Por las promociones con un 9% en la Costa y un 20% en la Sierra. De esta razón de preferencia podemos indicar que los entrevistados consideran que usan determinada operadora debido a que las llamadas con sus contactos son más económicas a que si fueran de otra. Estas tres razones nos entregan un total de 62% en la Costa y un 70% en la Sierra, la cual en si tiene que ver con el uso de la misma; es decir, que la mayoría de entrevistados nombran y prefieren a una determinada empresa porque la usan. También podemos apreciar que existen razones que son semejantes en ambas regiones y otras en las cuales no es así.

Motivos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Por su cobertura	286	136	0,42	0,44
Por sus contactos/Amigos	73	19	0,11	0,06
Por las promociones	64	63	0,09	0,20
Por los servicios que brinda	53	1	0,08	0,00
Porque la tarifa de llamadas es económica	53	13	0,08	0,04
Otras razones	148	76	0,22	0,25
Total	677	308	1	1

Fuente y Elaboración: Mario Plaza



PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

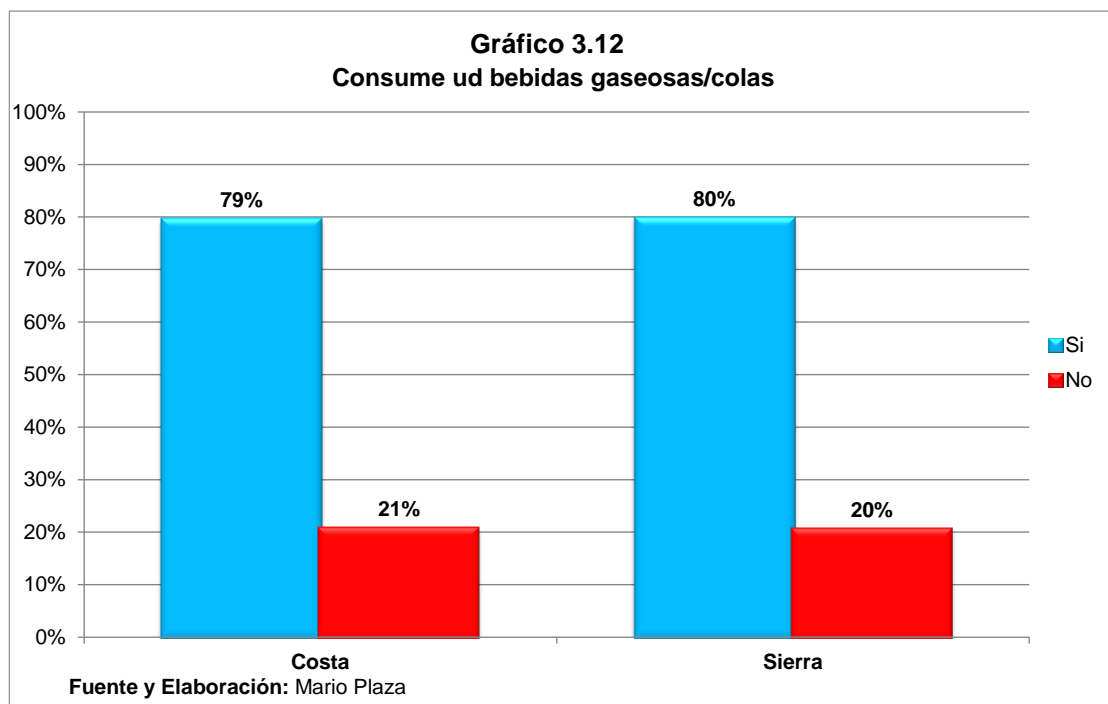
Variable X₁₂: Utiliza/Consume

La variable X₁₂ nos permite conocer el porcentaje de consumo de Bebidas gaseosas entre los entrevistados. Estos porcentajes lo podemos apreciar en

la Tabla XIII y el Gráfico 3.12, los cuales nos indican que el 79% de la población objetivo en la Costa consume algún tipo de gaseosa, mientras que en la Sierra es un 80%. Es con estos porcentajes con los cuales se procede a realizar el análisis de ésta categoría.

Tabla XIII				
Distribución de Frecuencia del Consumo de Bebidas Gaseosas-Colas				
Consumo de Bebidas Gaseosas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Si	598	320	0,79	0,80
No	155	82	0,21	0,20
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



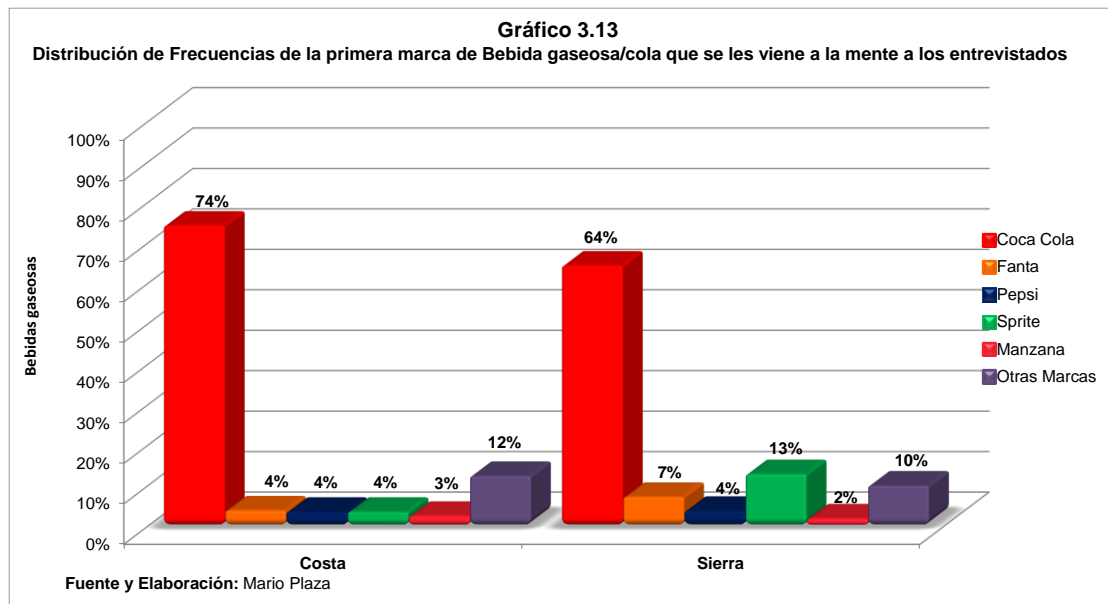
PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

Variable X₁₃: TOP

La variable X₁₃ nos da a conocer la primera marca de Bebida gaseosa que el encuestado tiene en mente. Lo cual podemos observarlo en la Tabla XIV y el Gráfico 3.13, que nos indican que la marca más recordada es Coca Cola en ambas regiones del país, en la Costa con un 74% y en la Sierra con un 64%, seguida muy de lejos por Fanta con un 4% en la Costa y un 7% en la Sierra, y después Pepsi con un 4% en la Costa y el mismo porcentaje en la Sierra.

Tabla XIV				
Tabla de Frecuencia de la Primera marca de Bebida gaseosa/Cola que se le viene a la mente				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Coca Cola	444	206	0,74	0,64
Fanta	23	23	0,04	0,07
Pepsi	21	12	0,04	0,04
Sprite	21	41	0,04	0,13
Manzana	15	6	0,03	0,02
Otras Marcas	74	32	0,12	0,10
Total	598	320	1	1

Fuente y Elaboración: Mario Plaza



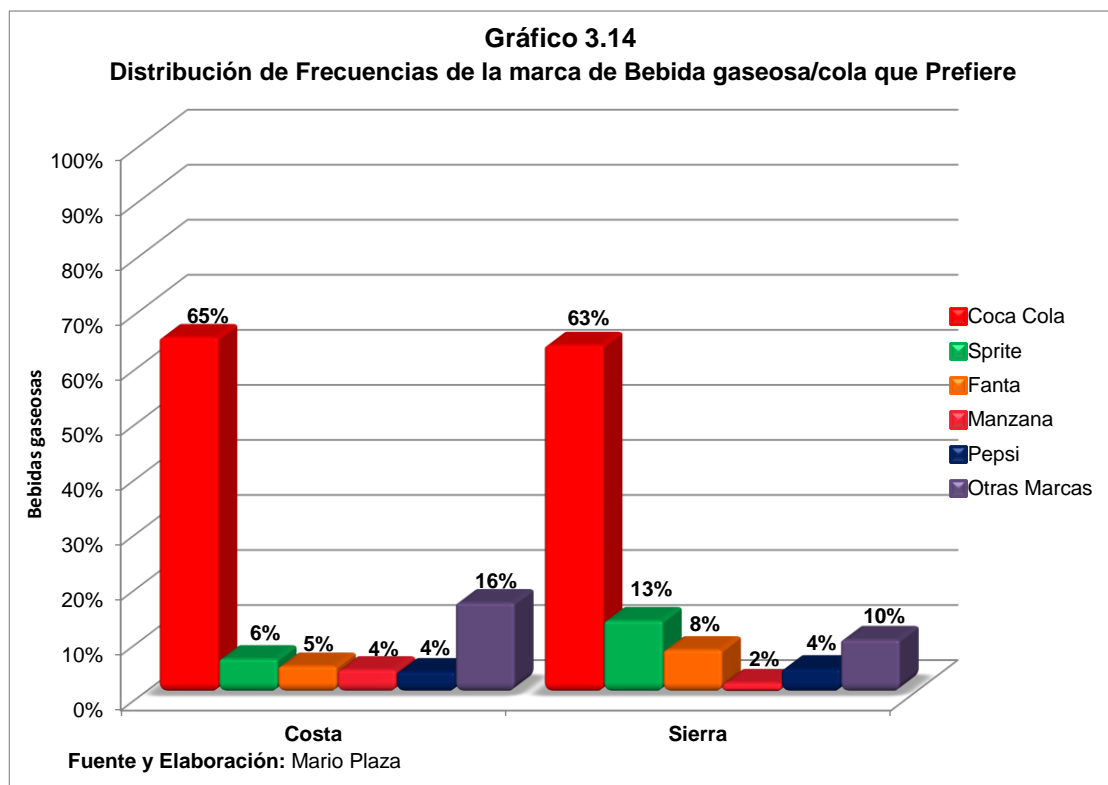
PREGUNTA 5: ¿Cuál es la marca que usted prefiere?

Variable X_{14} : Prefiere

La variable X_{14} tiene que ver con la marca de bebida gaseosa que prefiere el entrevistado. Esta información podemos apreciarla en la Tabla XV y el Gráfico 3.14, que indican que similar al TOM la marca preferida en ambas regiones es Coca Cola con un 65% en la Costa y un 63% en la Sierra, seguida muy de lejos por Sprite con un 6% de preferencia en la Costa y un 13% en la Sierra; en tercer lugar la marca Fanta con un 5% en la Costa y un 8% en la Sierra.

Tabla XV				
Tabla de Frecuencia de la Marca de Bebida gaseosa/cola que Prefiere				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Coca Cola	386	202	0,65	0,63
Sprite	37	42	0,06	0,13
Fanta	29	25	0,05	0,08
Manzana	25	6	0,04	0,02
Pepsi	23	14	0,04	0,04
Otras Marcas	98	31	0,16	0,10
Total	598	320	1	1

Fuente y Elaboración: Mario Plaza

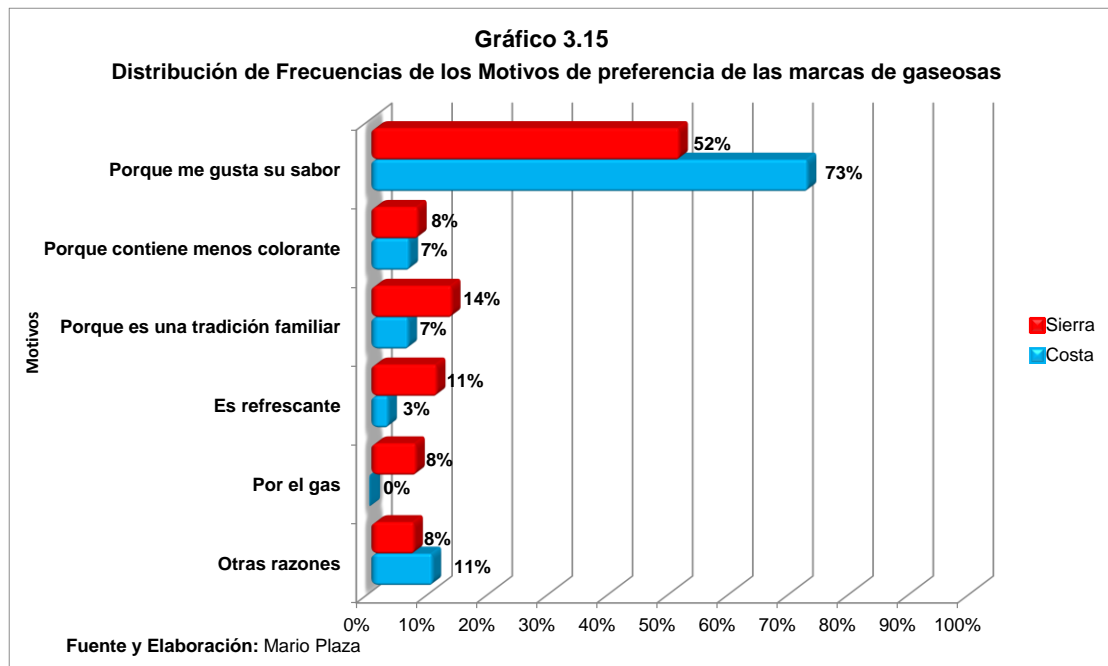


PREGUNTA 6: ¿Por qué prefiere la marca?**Variable X₁₅: Razones de preferencia**

La variable X₁₅ nos permite conocer todas las razones por las cuales el encuestado prefiere determinada marca de bebida gaseosa. Podemos apreciar dicha información en la Tabla XVI y en el Gráfico 3.15. La razón con mayor porcentaje es Porque me gusta su sabor con un 73% en la Costa y 52% en la Sierra, seguida de la razón Porque contiene menos colorante con un 7% en la Costa y un 8% en la Sierra, en tercer lugar tenemos la razón Porque es una tradición familiar con un 7% en la Costa; sin embargo, esta razón cuenta con un 14% en la Sierra.

Tabla XVI				
Distribución de Frecuencia de los Motivos de preferencia de las marcas de gaseosas				
Motivos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Porque me gusta su sabor	436	165	0,73	0,52
Porque contiene menos colorante	40	26	0,07	0,08
Porque es una tradición familiar	39	44	0,07	0,14
Es refrescante	19	36	0,03	0,11
Por el gas	1	25	0,00	0,08
Otras razones	63	24	0,11	0,08
Total	598	320	1	1

Fuente y Elaboración: Mario Plaza



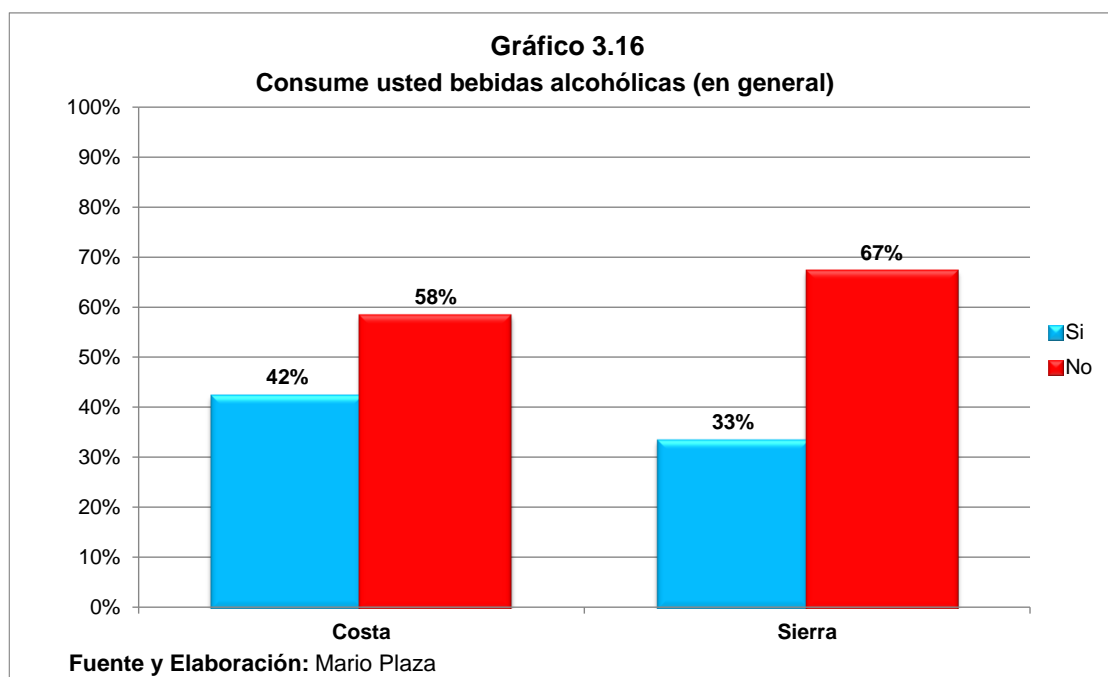
PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

Variable X₁₆: Utiliza/Consume

La variable X₁₆ nos permite conocer el porcentaje de consumo de Bebidas Alcohólicas (en general) entre los entrevistados. Estos datos podemos apreciarlos en la Tabla XVII y el Gráfico 3.16, que nos indican que un 42% de la población en la Costa consume algún tipo de bebida alcohólica, y en la Sierra un 33%. Estos valores son bajos debido a que debemos recordar que en el estudio hay un 59% de Mujeres y solo un 41% de Hombres.

Tabla XVII				
Distribución de Frecuencia del Consumo de Bebidas Alcohólicas (en general)				
Consumo de Bebidas Alcohólicas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Si	316	133	0,42	0,33
No	437	269	0,58	0,67
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

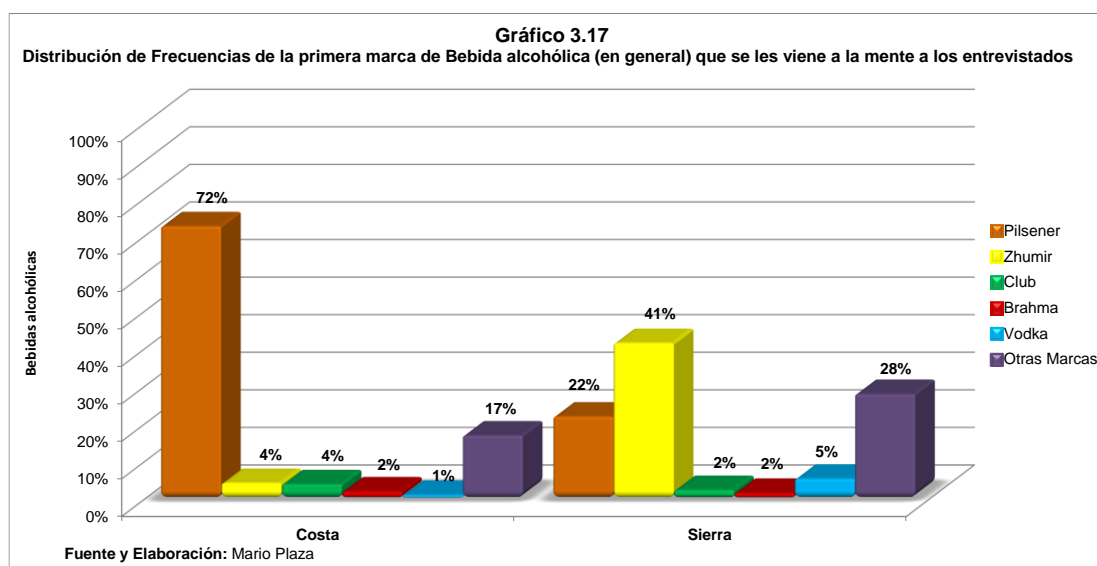
Variable X₁₇: TOP

La variable X₁₇ nos permite conocer el porcentaje de encuestados que tienen determinada marca en su mente en primer lugar. En la Tabla XVIII y en el Gráfico 3.17 podemos apreciar que la marca con mayor nivel de recordación

es Pilsener con un 72% en la Costa y tan solo un 22% en la Sierra, en segundo lugar para la Costa viene la marca Zhumir con un 4%, sin embargo esta marca viene a ser la primera en la Sierra con un 41%. En tercer lugar para la Costa tenemos la marca Club con un 4%; en cambio, en la Sierra en tercer lugar se encuentra la marca Vodka con un 5%.

Tabla XVIII				
Tabla de Frecuencia de la Primera marca de Bebida alcohólica (en general) que se le viene a la mente				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Pilsener	229	29	0,72	0,22
Zhumir	13	55	0,04	0,41
Club	12	3	0,04	0,02
Brahma	6	2	0,02	0,02
Vodka	3	7	0,01	0,05
Otras Marcas	53	37	0,17	0,28
Total	316	133	1	1

Fuente y Elaboración: Mario Plaza

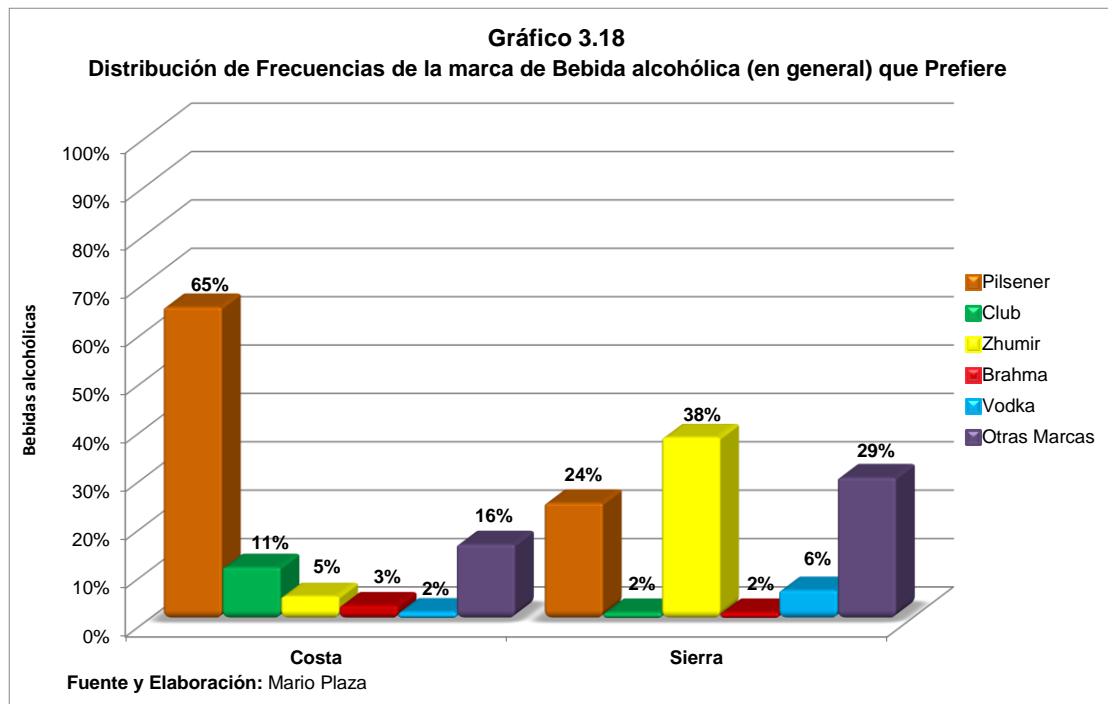


PREGUNTA 5: ¿Cuál es la marca que usted prefiere?**Variable X₁₈: Prefiere**

La variable X₁₈ nos entrega información referente al porcentaje de encuestados que prefieren determinada marca de Bebida Alcohólica. En la Costa en primer lugar tenemos a Pilsener con un 65%, en segundo lugar a Club con un 11% y en tercer lugar a Zhumir con un 5%. En la Sierra en cambio se tiene a Zhumir en primer lugar con un 38% de preferencia, en segundo lugar a Pilsener con un 24% y en tercer lugar a Vodka con un 6%. Esta información la podemos apreciar en la Tabla XIX y el Gráfico 3.18.

Tabla XIX				
Tabla de Frecuencia de la Marca de Bebida alcohólica (en general) que Prefiere				
Marcas	Número de		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Pilsener	204	32	0,65	0,24
Club	34	2	0,11	0,02
Zhumir	15	50	0,05	0,38
Brahma	9	2	0,03	0,02
Vodka	5	8	0,02	0,06
Otras Marcas	49	39	0,16	0,29
Total	316	133	1	1

Fuente y Elaboración: Mario Plaza



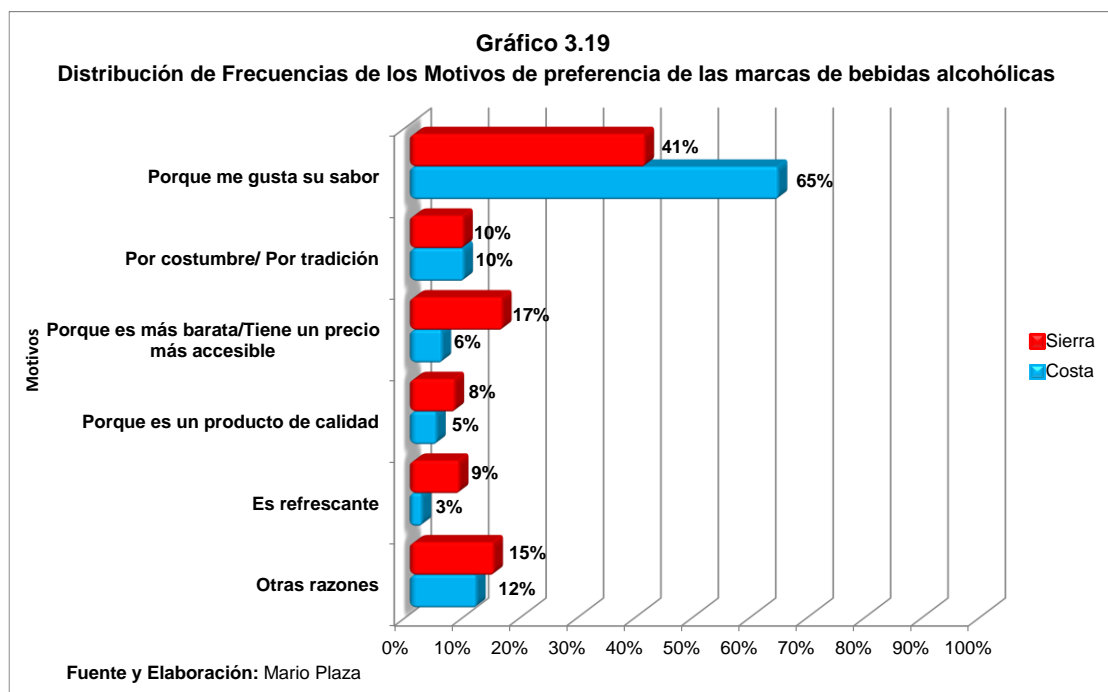
PREGUNTA 6: ¿Por qué prefiere la marca?

Variable X_{19} : Razones de preferencia

En la variable X_{19} conocemos las razones por las cuales los entrevistados prefieren determinada marca. En primer lugar tenemos la razón Porque me gusta su sabor con un 65% en la Costa y un 41% en la Sierra. En segundo lugar tenemos la razón Por costumbre/Por tradición con un 10 en ambas regiones. Y en tercer lugar tenemos Porque es más barata/Tiene un precio más accesible con un 6% en la Costa y un 17% en la Sierra. Podemos apreciar esta información en la Tabla XX y en el Gráfico 3.19.

Tabla XX				
Distribución de Frecuencia de los Motivos de preferencia de las marcas de bebidas alcohólicas (en general)				
Motivos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Porque me gusta su sabor	204	55	0,65	0,41
Por costumbre/ Por tradición	31	13	0,10	0,10
Porque es más barata/Tiene un precio más accesible	19	22	0,06	0,17
Porque es un producto de calidad	16	11	0,05	0,08
Es refrescante	8	12	0,03	0,09
Otras razones	38	20	0,12	0,15
Total	316	133	1	1

Fuente y Elaboración: Mario Plaza



PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

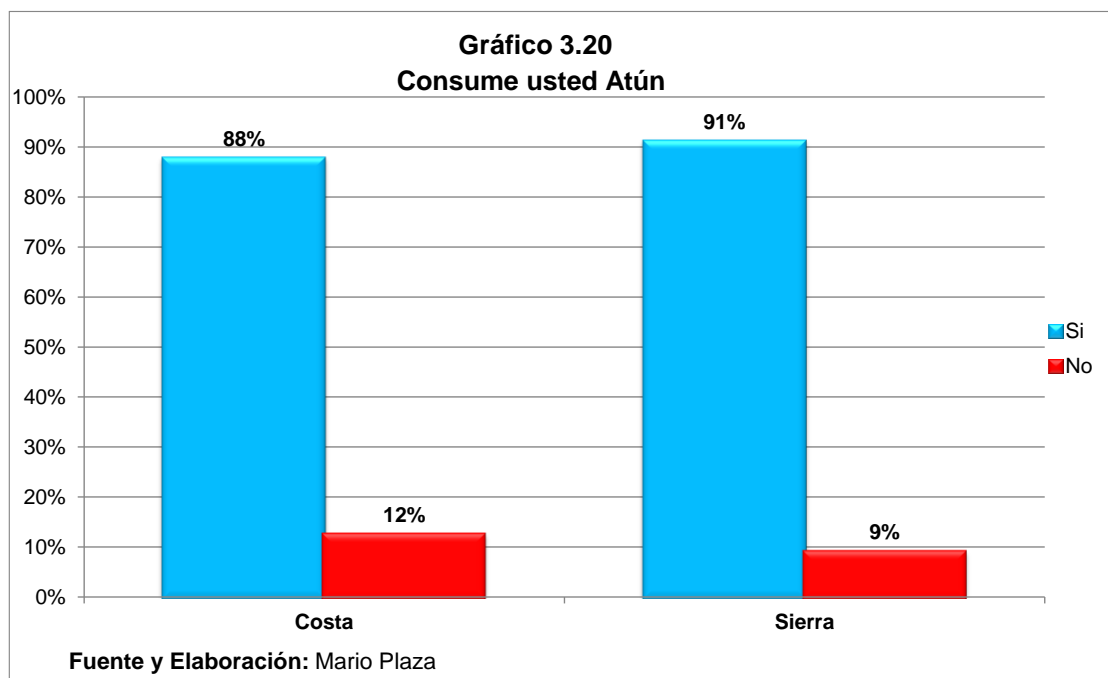
Variable X_{20} : Utiliza/Consume

La variable X_{20} nos permite conocer el porcentaje de entes de investigación que consumen Atún. Dichos porcentajes podemos apreciarlos en la Tabla

XXI y el Gráfico 3.20, los cuales nos indican que un 88% en la Costa consume atún y un 91% en la Sierra lo hace. Con estos porcentajes procedemos a realizar los diferentes análisis.

Tabla XXI				
Distribución de Frecuencia del Consumo de Atún				
Consumo de Atún	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Si	660	366	0,88	0,91
No	93	36	0,12	0,09
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



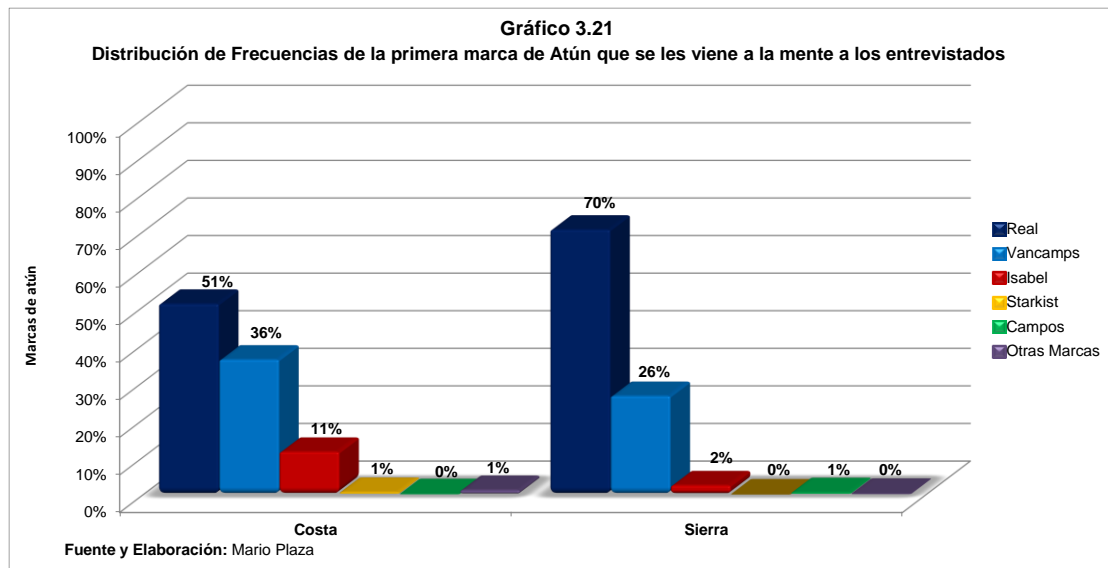
PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

Variable X₂₁: TOP

La variable X₂₁ nos permite conocer las marcas que los entrevistados tienen en sus mentes al momento de que se le pregunte por una marca de atún. Esta información la podemos apreciar en la Tabla XXII y el Gráfico 3.21, en donde se observa que la marca de atún Real en la Costa tiene un 51% y en la Sierra cuenta con un 70%, Van Camps en la Costa cuenta con un 36% de recordación y en la Sierra un 26%, seguidos muy de lejos por Isabel con un 11% en la Costa y un 2% en la Sierra.

Tabla XXII				
Tabla de Frecuencia de la Primera marca de Atún que se le viene a la mente				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Real	335	258	0,51	0,70
Vancamps	237	96	0,36	0,26
Isabel	75	9	0,11	0,02
Starkist	4	0	0,01	0,00
Campos	1	2	0,00	0,01
Otras Marcas	8	1	0,01	0,00
Total	660	366	1	1

Fuente y Elaboración: Mario Plaza



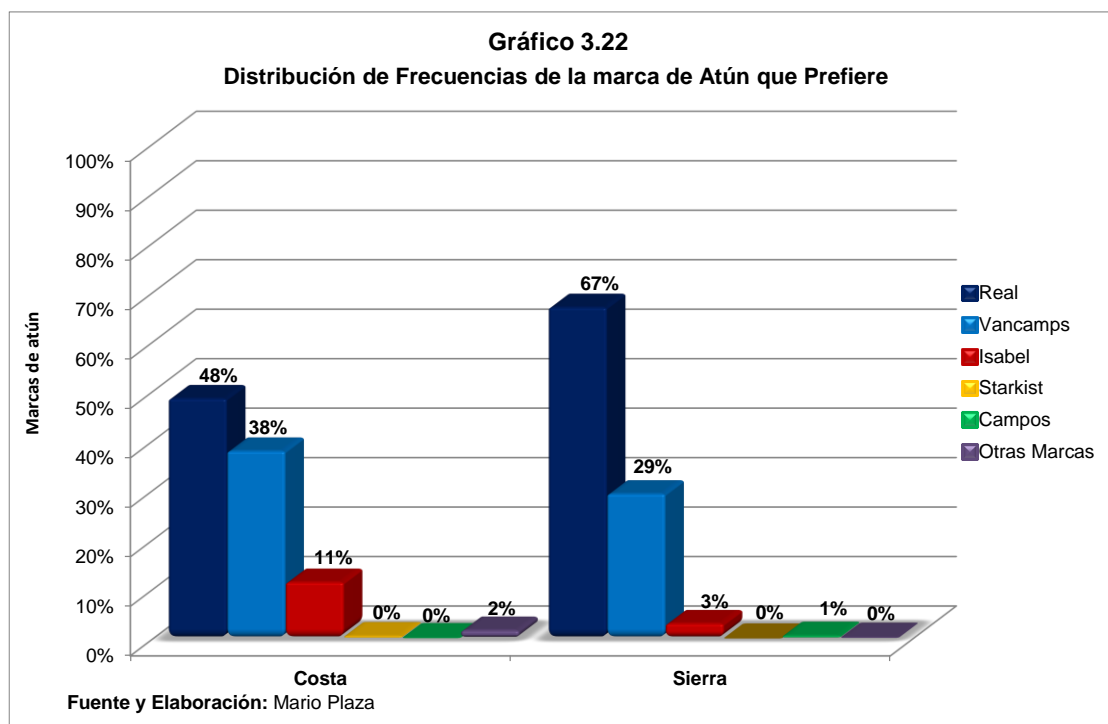
PREGUNTA 5: ¿Cuál es la marca que usted prefiere?

Variable X₂₂: Prefiere

En la variable X₂₂ podemos apreciar el porcentaje de entrevistados que prefieren determinada marca de Atún. En la Tabla XXIII y en el Gráfico 3.22 encontramos estos datos. Muy similar al TOM tenemos que atún Real tiene un 48% de preferencia y en la Sierra un 67%. Atún Van Camps cuenta con un 38% de preferencia en la Costa y un 29% en la Sierra. Seguidas muy de lejos por Isabel con un 11% en la Costa y un 3% en la Sierra.

Tabla XXIII				
Tabla de Frecuencia de la Marca de Atún que Prefiere				
Marcas	Número de		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Real	320	245	0,48	0,67
Vancamps	250	107	0,38	0,29
Isabel	75	11	0,11	0,03
Starkist	3	0	0,00	0,00
Campos	1	2	0,00	0,01
Otras Marcas	11	1	0,02	0,00
Total	660	366	1	1

Fuente y Elaboración: Mario Plaza

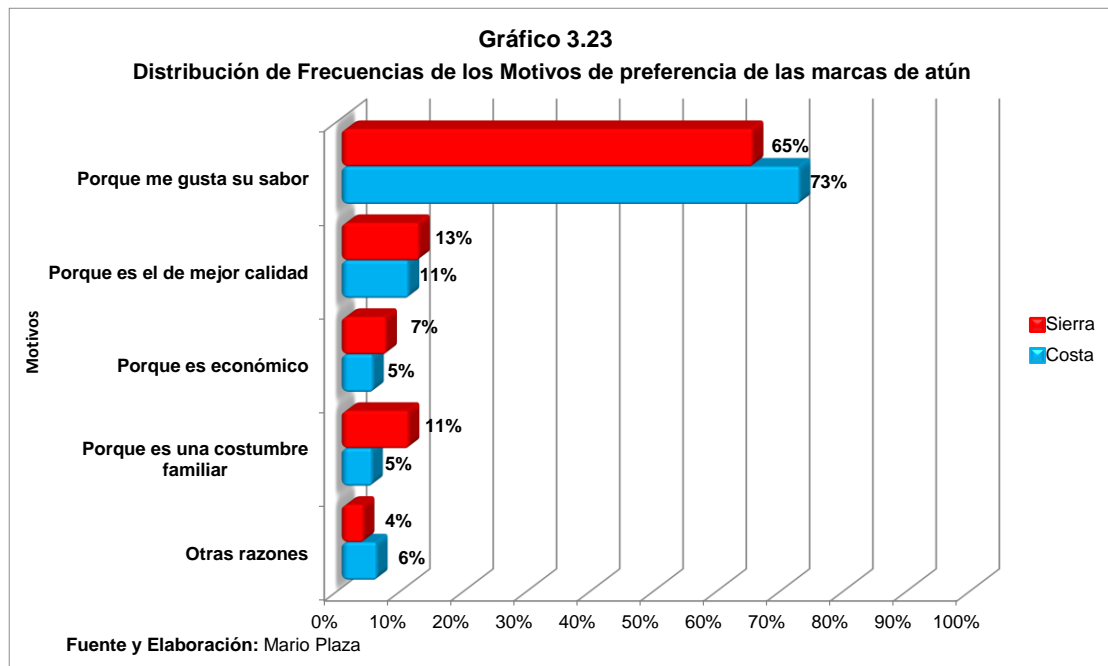


PREGUNTA 6: ¿Por qué prefiere la marca?**Variable X₂₃: Razones de preferencia**

La variable X₂₃ indica las razones de preferencia de cada marca. La preferencia por El sabor cuenta con un 73% en la Costa y un 65% en la Sierra. La Calidad es otra razón con un 11% en la Costa y un 13% en la Sierra. Y la Economía con un 5% en la Costa y un 7% en la Sierra.

Tabla XXIV				
Distribución de Frecuencia de los Motivos de preferencia de las marcas de atún				
Motivos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Porque me gusta su sabor	480	239	0,73	0,65
Porque es el de mejor calidad	72	46	0,11	0,13
Porque es económico	35	27	0,05	0,07
Porque es una costumbre familiar	34	40	0,05	0,11
Otras razones	39	14	0,06	0,04
Total	660	366	1	1

Fuente y Elaboración: Mario Plaza



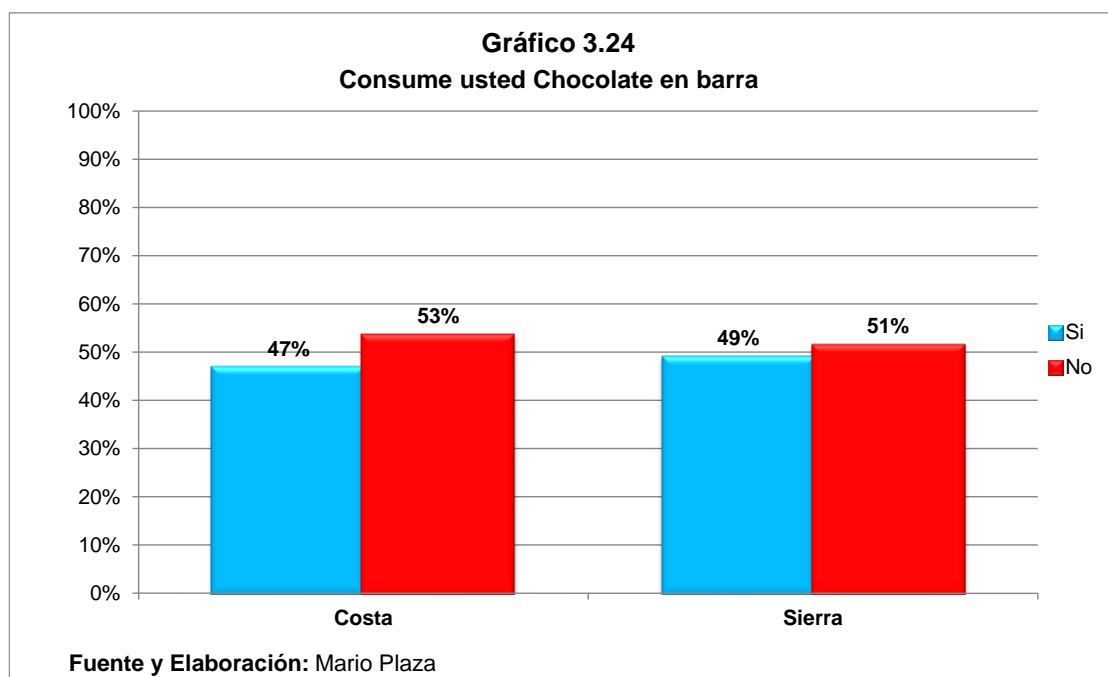
PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

Variable X₂₄: Utiliza/Consume

La variable X₂₄ nos indica el porcentaje de consumo de Chocolate en barra en la Costa y Sierra. Con un porcentaje del 47% de consumo en la Costa y uno de 49% en la Sierra procedemos a analizar la información con estos datos. Podemos apreciar de mejor manera la información previamente mencionada en la Tabla XXV y el Gráfico 3.24.

Tabla XXV				
Distribución de Frecuencia del Consumo de Chocolate en barra				
Consumo de Chocolate	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Si	351	196	0,47	0,49
No	402	206	0,53	0,51
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

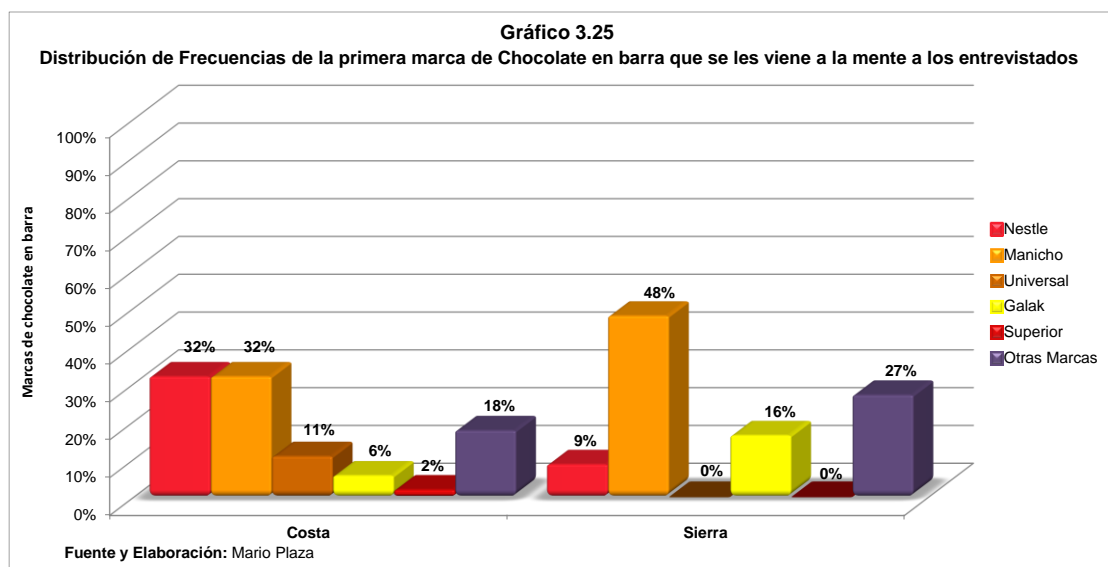
Variable X₂₅: TOP

La variable X₂₅ nos permite conocer el TOM con respecto a las marcas de Chocolate en barra en las regiones Costa y Sierra. En la Costa se tiene un empate con la marca Nestlé y Manicho con un 32%, a diferencia de la Sierra

en la que la primera solo cuenta con un 9% y la segunda con un 48%. Seguido muy de lejos con la Universal con solo un 11% en la Costa. La Tabla XXVI y el Gráfico 3.25 nos permiten conocer con mejor detalle esta información y la de otras marcas.

Tabla de Frecuencia de la Primera marca de Chocolate en barra que se le viene a la mente				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Nestle	112	17	0,32	0,09
Manicho	112	94	0,32	0,48
Universal	38	0	0,11	0,00
Galak	20	32	0,06	0,16
Superior	7	0	0,02	0,00
Otras Marcas	62	53	0,18	0,27
Total	351	196	1	1

Fuente y Elaboración: Mario Plaza

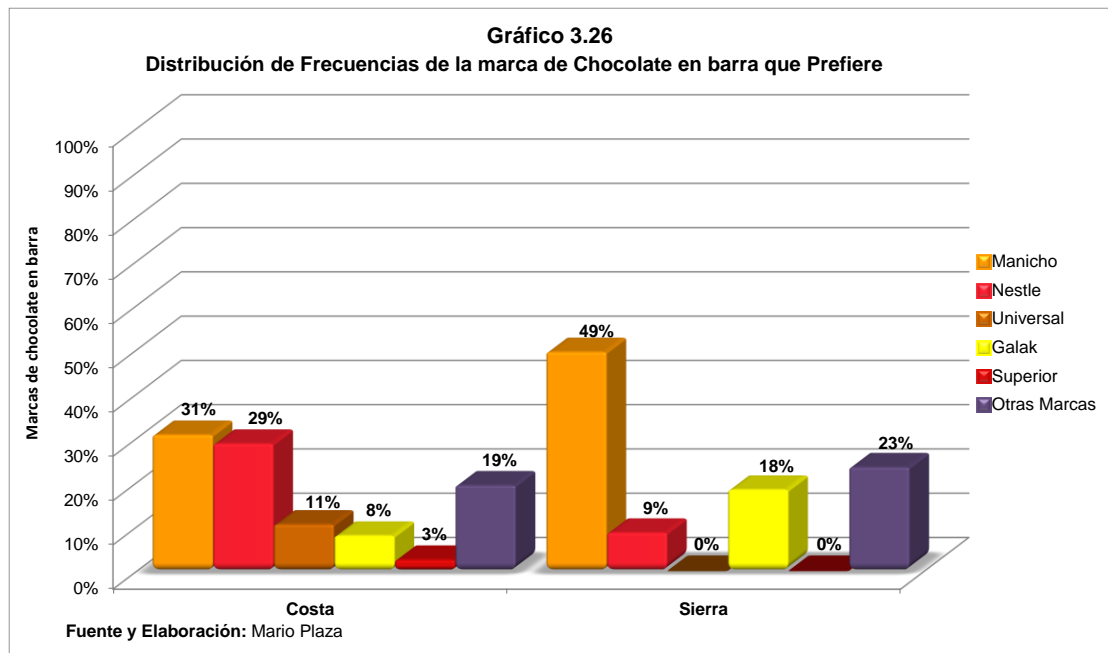


PREGUNTA 5: ¿Cuál es la marca que usted prefiere?**Variable X₂₆: Prefiere**

La variable X₂₆ nos indica las marcas de preferencia de Chocolate en barra en las regiones de la Costa y Sierra ecuatoriana. Podemos apreciar estos datos en la Tabla XXVII y el Gráfico 3.26, los cuales nos muestran que la marca que prefieren en ambas regiones es Manicho con un 31% en la Costa y 49% en la Sierra, con una reñida posición viene Nestlé con un 29% en la Costa y un 9% en la Sierra, muy de lejos continúa Universal con un 11% de preferencia solo en la Costa.

Tabla XXVII				
Tabla de Frecuencia de la Marca de Chocolate en barra que Prefiere				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Manicho	108	97	0,31	0,49
Nestle	101	17	0,29	0,09
Universal	37	0	0,11	0,00
Galak	28	36	0,08	0,18
Superior	9	0	0,03	0,00
Otras Marcas	68	46	0,19	0,23
Total	351	196	1	1

Fuente y Elaboración: Mario Plaza



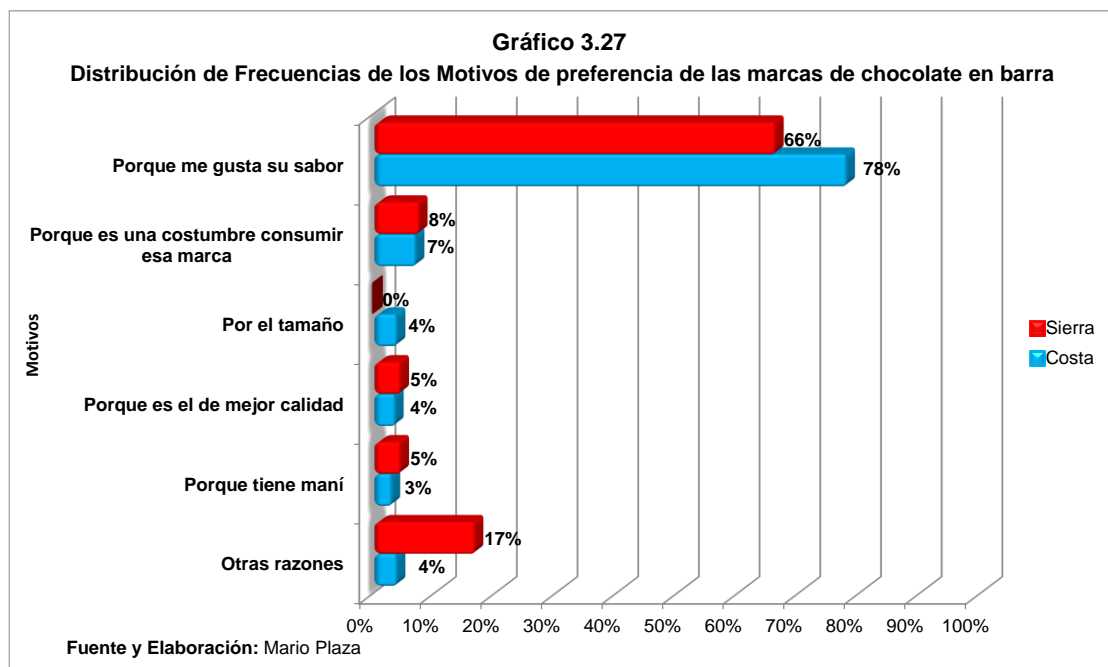
PREGUNTA 6: ¿Por qué prefiere la marca?

Variable X₂₇: Razones de preferencia

La variable **X₂₇** indica las razones de preferencia de las marcas de Chocolate en barra en las regiones de la Costa y Sierra. En primer lugar tenemos de forma indiscutible en ambas regiones al Sabor con un 78% en la Costa y un 66% en la Sierra, seguido muy lejos por la Costumbre con un 7% en la Costa y un 8% en la Sierra.

Distribución de Frecuencia de los Motivos de preferencia de las marcas de Chocolate en barra				
Motivos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Porque me gusta su sabor	274	130	0,78	0,66
Porque es una costumbre consumir esa marca	25	15	0,07	0,08
Por el tamaño	14	0	0,04	0,00
Porque es el de mejor calidad	13	9	0,04	0,05
Porque tiene maní	11	9	0,03	0,05
Otras razones	14	33	0,04	0,17
Total	351	196	1	1

Fuente y Elaboración: Mario Plaza



PREGUNTA 3: ¿Usted utiliza o consume de los siguientes tipos de productos?

Variable X₂₈: Utiliza/Consume

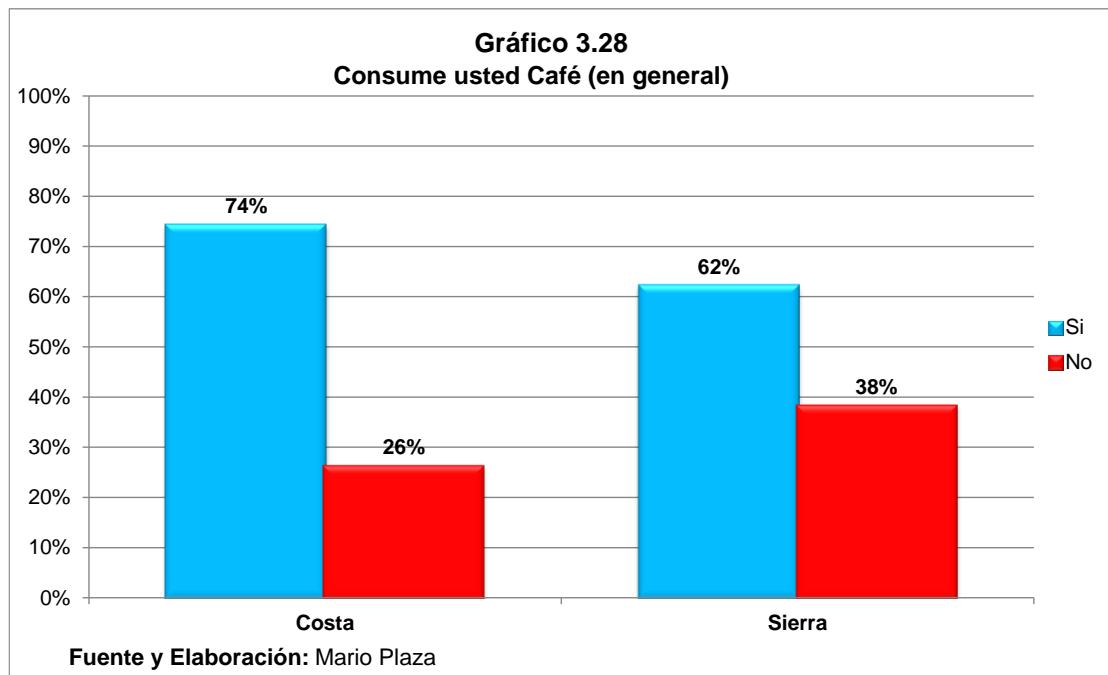
La variable X₂₈ indica el porcentaje de consumo de Café en la Costa y Sierra.

Los datos podemos apreciarlos en la Tabla XXIX y el Gráfico 3.28, que nos

muestran que el consumo de café en la Costa es del 74% y en la Sierra del 62%.

Tabla XXIX				
Distribución de Frecuencia del Consumo de Café				
Consumo de Café	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Si	557	249	0,74	0,62
No	196	153	0,26	0,38
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



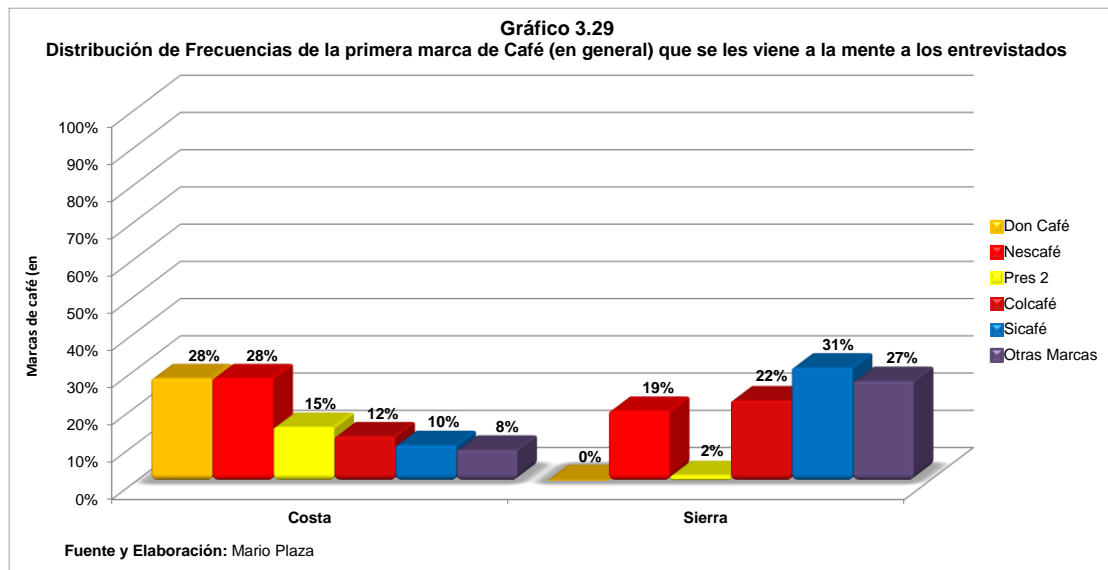
PREGUNTA 4: Según la categoría o Producto ¿Cuál es la marca que usted recuerda en estos momentos?

Variable X₂₉: TOP

La variable X₂₉ muestra a través de la Tabla XXX y del Gráfico 3.29 el TOM entre las diferentes marcas de café, entre las cuales tenemos a Don Café y Nescafé con el 28% en la Costa, y solo se cuenta con valor del TOM de Nescafé en la Sierra con un 19%.

Tabla XXX				
Tabla de Frecuencia de la Primera marca de Café (en general) que se le viene a la mente				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Don Café	154	1	0,28	0,00
Nescafé	154	47	0,28	0,19
Pres 2	81	4	0,15	0,02
Colcafé	67	54	0,12	0,22
Sicafé	54	76	0,10	0,31
Otras Marcas	47	67	0,08	0,27
Total	557	249	1	1

Fuente y Elaboración: Mario Plaza



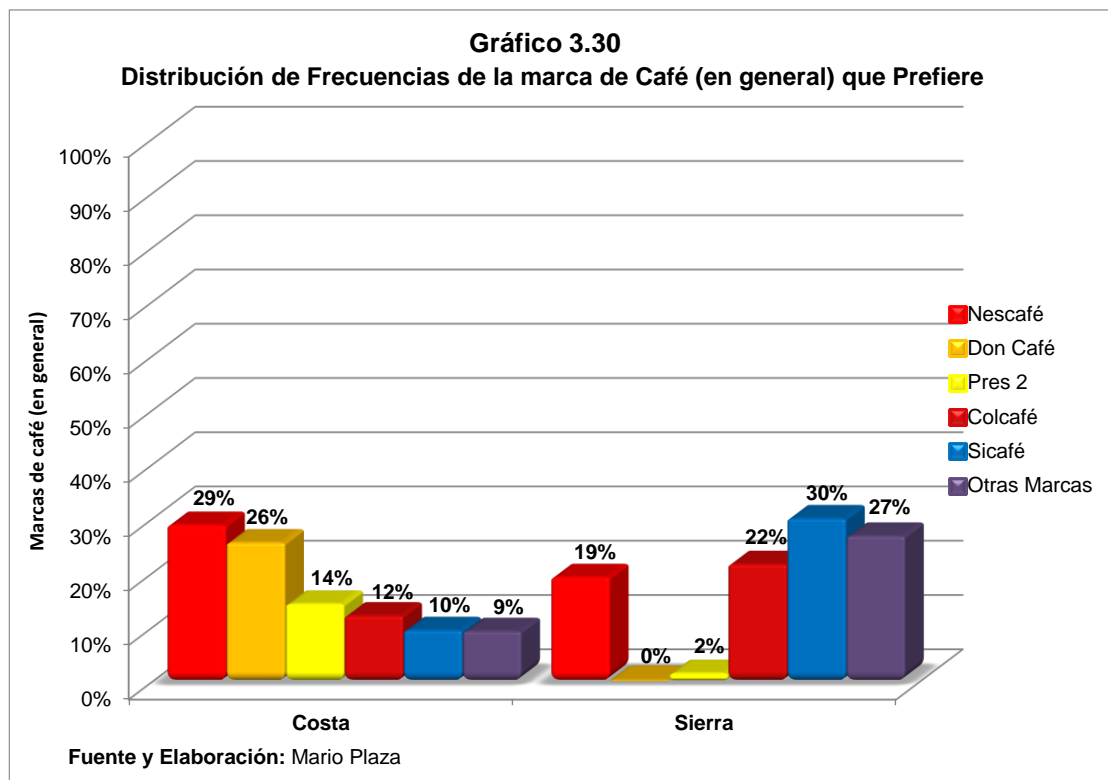
PREGUNTA 5: ¿Cuál es la marca que usted prefiere?

Variable X_{30} : Prefiere

La variable X_{30} nos indica las marcas de café que los entrevistados prefieren. Esta información la podemos apreciar en la Tabla XXXI y en el Gráfico 3.29, en las cuales tenemos a Nescafé con un 29% en la Costa y un 19% en la Sierra. Don Café solo tiene preferencia en la Costa con un 26%. Y Pres 2 con un 14% en la Costa y 2% en la Sierra.

Tabla XXXI				
Tabla de Frecuencia de la Marca de Café (en general) que Prefiere				
Marcas	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Nescafé	161	48	0,29	0,19
Don Café	143	1	0,26	0,00
Pres 2	80	4	0,14	0,02
Colcafé	68	54	0,12	0,22
Sicafé	53	75	0,10	0,30
Otras Marcas	52	67	0,09	0,27
Total	557	249	1	1

Fuente y Elaboración: Mario Plaza



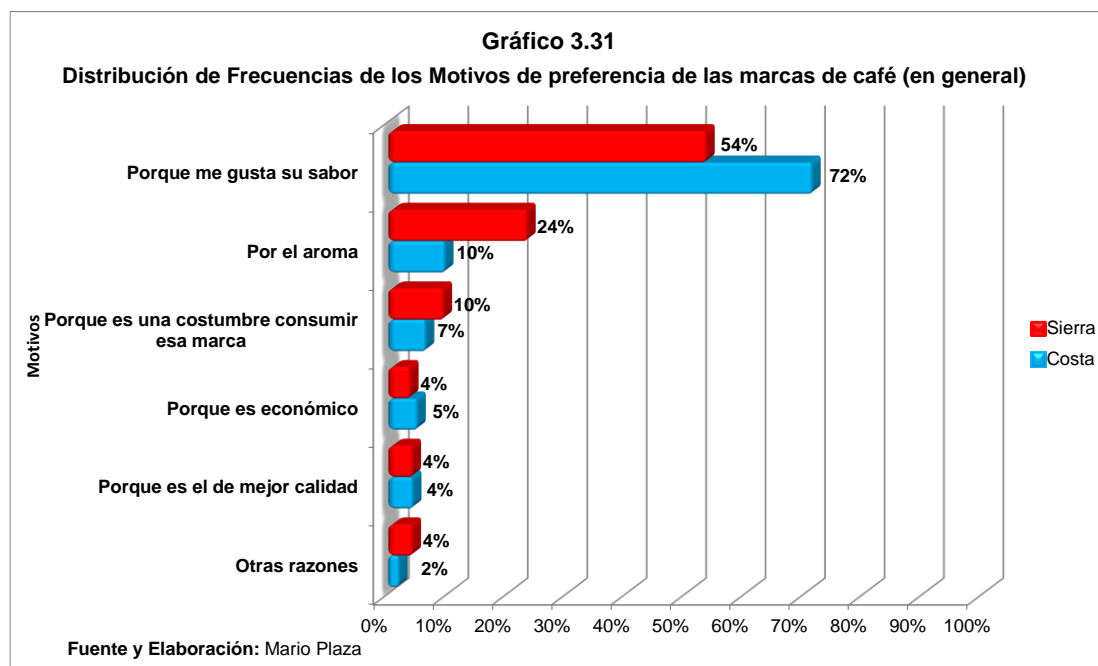
PREGUNTA 6: ¿Por qué prefiere la marca?

Variable X₃₁: Razones de preferencia

La variable X₃₁ nos indica las razones de preferencia de las marcas de café, que podemos apreciar en la Tabla XXXII y en el Gráfico 3.31, siendo la razón Porque me gusta su sabor la primera con un 72% en la Costa y un 54% en la Sierra, después El aroma con un 10% en la Costa y un 24% en la Sierra.

Distribución de Frecuencia de los Motivos de preferencia de las marcas de café (en general)				
Motivos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Porque me gusta su sabor	399	134	0,72	0,54
Por el aroma	55	59	0,10	0,24
Porque es una costumbre consumir esa marca	37	24	0,07	0,10
Porque es económico	29	10	0,05	0,04
Porque es el de mejor calidad	25	11	0,04	0,04
Otras razones	12	11	0,02	0,04
Total	557	249	1	1

Fuente y Elaboración: Mario Plaza



3.1.3. Variables de la Sección IV – Orden de importancia.

En esta sección se han de analizar los atributos entregados al encuestado para que el mismo los coloque en orden de importancia según le parezca. La lista de atributos es la siguiente:

1. Color
2. Diseño
3. Tradición
4. Calidad
5. Marca Confiable (Confianza)
6. Envase
7. Presentación
8. Precio
9. Logotipo

PREGUNTA 7: En orden de importancia ¿Cuáles son los atributos que usted considera importantes al momento de elegir una marca?

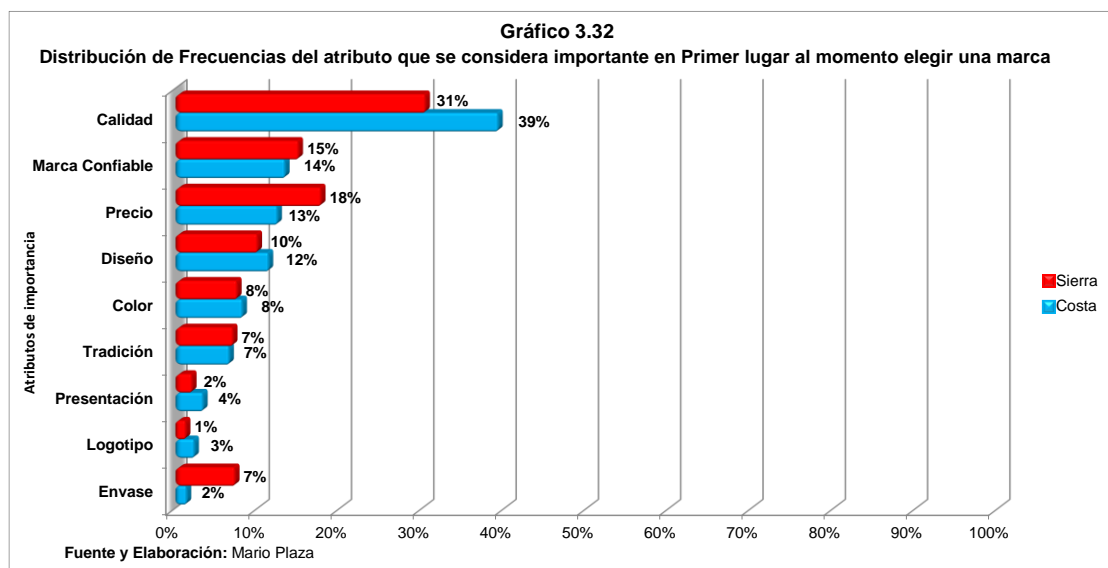
Variable X₃₂: Primer lugar

La variable X₃₂ nos indica el orden de importancia entregado por los entes de investigación según determinados atributos al momento de elegir una marca, en primer lugar. Tenemos a Calidad con un 39% en Costa y un 31% en la Sierra, Marca Confiable con un 14% en la Costa y un 15% en la Sierra y Precio con un 13% en la Costa y un 18% en la Sierra. Podemos apreciar que

estos atributos tienen diferentes ponderaciones en las diferentes regiones, podemos apreciarlo en la Tabla XXXIII y en el Gráfico 3.32.

Tabla XXXIII				
Tabla de Frecuencia del atributo que se considera importante en Primer lugar al momento de elegir una marca				
Atributos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Calidad	297	123	0,39	0,31
Marca Confiable	102	61	0,14	0,15
Precio	95	72	0,13	0,18
Diseño	87	41	0,12	0,10
Color	63	31	0,08	0,08
Tradición	51	29	0,07	0,07
Presentación	27	9	0,04	0,02
Logotipo	19	6	0,03	0,01
Envase	12	30	0,02	0,07
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



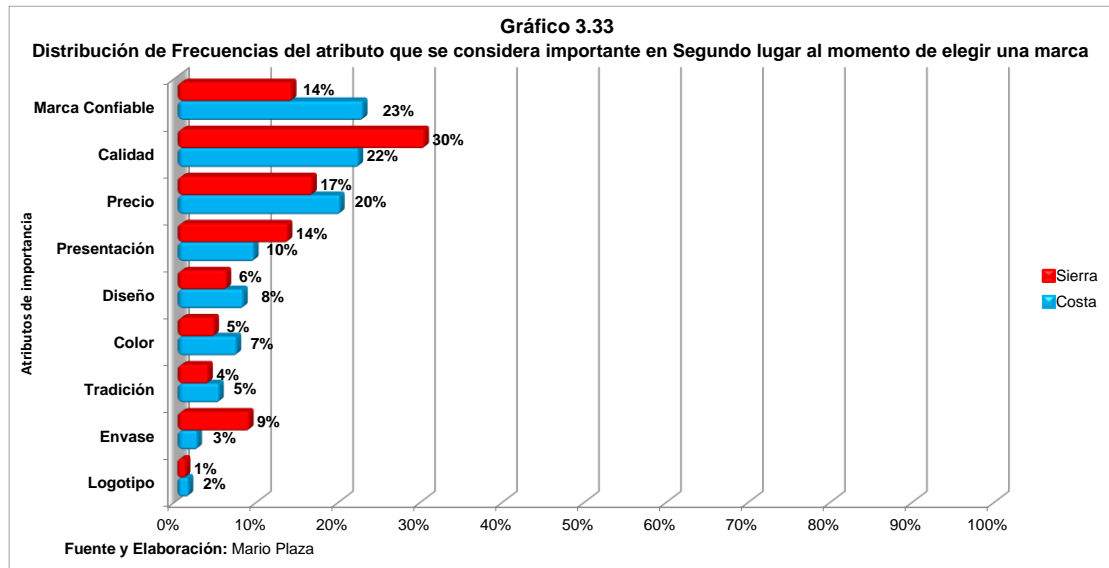
PREGUNTA 7: En orden de importancia ¿Cuáles son los atributos que usted considera importantes al momento de elegir una marca?

Variable X₃₃: Segundo lugar

La variable X₃₃ indica el segundo lugar de importancia de un atributo a la hora de escoger una marca. Según la Tabla XXXIV y el Gráfico 3.33 los atributos que ocupan el segundo lugar de importancia son Marca confiable con 23% en la Costa y un 14% en la Sierra, Calidad con un 22% en la Costa y un 30% en la Sierra y Precio con 20% en la Costa y 17% en la Sierra.

Tabla XXXIV				
Tabla de Frecuencia del atributo que se considera importante en Segundo lugar al momento de elegir una marca				
Atributos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Marca Confiable	172	57	0,23	0,14
Calidad	168	122	0,22	0,30
Precio	151	67	0,20	0,17
Presentación	72	55	0,10	0,14
Diseño	62	25	0,08	0,06
Color	56	19	0,07	0,05
Tradición	40	16	0,05	0,04
Envase	20	36	0,03	0,09
Logotipo	12	5	0,02	0,01
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



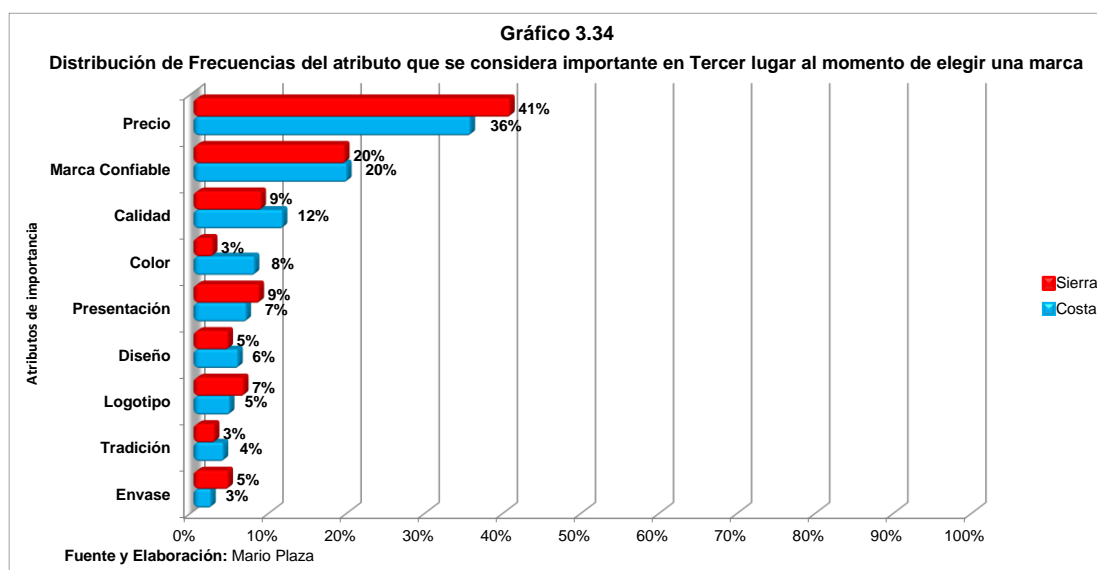
PREGUNTA 7: En orden de importancia ¿Cuáles son los atributos que usted considera importantes al momento de elegir una marca?

Variable X_{34} : Tercer lugar

La variable X_{34} nos indica el atributo que es de importancia en tercer lugar, y lo podemos apreciar en la Tabla XXXV y el Gráfico 3.34, en donde Precio cuenta con un 36% en la Costa y 41% en la Sierra, Marca Confiable un 20% en la Costa y Sierra, y Calidad con un 12% en la Costa y un 9% en la Sierra.

Tabla XXXV				
Tabla de Frecuencia del atributo que se considera importante en Tercer lugar al momento de elegir una marca				
Atributos	Número de entrevistados		Frecuencia Relativa	
	Costa	Sierra	Costa	Sierra
Precio	269	164	0,36	0,41
Marca Confiable	150	79	0,20	0,20
Calidad	88	36	0,12	0,09
Color	61	11	0,08	0,03
Presentación	53	35	0,07	0,09
Diseño	45	19	0,06	0,05
Logotipo	37	27	0,05	0,07
Tradición	31	12	0,04	0,03
Envase	19	19	0,03	0,05
Total	753	402	1	1

Fuente y Elaboración: Mario Plaza



3.2. Análisis inferencial

En la sección del Análisis inferencial vamos a proceder con el análisis de las tablas de contingencia, las cuales nos ayudan a establecer si existe o no independencia entre la variable explicada y las variables

explicativas a través del valor p obtenido en la prueba de hipótesis. En esta sección hemos de colocar las tablas referentes a las pruebas de hipótesis y en los Anexos se han de colocar las tablas de contingencia.

Para cada categoría se ha de verificar a través del valor p si cada variable puede formar parte o no del modelo estadístico.

Categoría: Empresas de Telefonía Celular

La variable de interés en esta categoría es Preferencia Celular, la cual se va a analizar con las diferentes posibles variables que pueden conformar el modelo. Para todas y cada una de las posibles variables que conformen el modelo se realiza la siguiente prueba de hipótesis:

H_0 : Las variables son independientes

Vs.

H_1 : Rechazamos H_0

Variable: Ciudad

En la Tabla XXXVI se aprecia el valor p de la tabla de contingencia Preferencia Celular VS Ciudad, cuyo valor es menor a 0,05, por lo que cae en la región de rechazo, y podemos concluir que existe evidencia estadística para rechazar la hipótesis nula y por lo tanto las variables no son

independientes, por ende la variable Ciudad va a formar parte del modelo dentro de ésta categoría.

Tabla XXXVI		
Valor p de la Tabla de contingencia de Preferencia Celular VS Ciudad		
Chi Cuadrado	Grados de libertad	Valor p
86.427	8	.000

Fuente y Elaboración: Mario Plaza

Variable: Región

En la Tabla XXXVII se muestra el valor p de la variable de interés con la variable Región, el cual es menor a 0,05 por lo que se concluye que existe evidencia para rechazar la hipótesis nula, mostrando que ambas variables no son independientes y por lo tanto la variable Región pasa a formar parte del modelo.

Tabla XXXVII		
Valor p de la Tabla de contingencia de Preferencia Celular VS Región		
Chi Cuadrado	Grados de libertad	Valor p
51.719	2	.000

Fuente y Elaboración: Mario Plaza

Variable: Edad

En la Tabla XXXVIII se puede apreciar que el valor p correspondiente a la tabla de contingencia de la variable de interés vs Edad es mayor a 0,05, por lo que concluimos que no existe suficiente evidencia estadística como para

rechazar la hipótesis nula. Por lo tanto la variable Edad no va a formar parte de modelo dentro de la categoría Telefonía celular.

Tabla XXXVIII		
Valor p de la Tabla de contingencia de Preferencia Celular VS Edad		
Chi Cuadrado	Grados de libertad	Valor p
97.405	100	.555

Fuente y Elaboración: Mario Plaza

Variable: Género

En la Tabla XXXIX podemos apreciar que el valor p no es menor a 0,05, por lo tanto se concluye que no existe evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Género no va a formar parte del modelo de la categoría Telefonía Celular.

Tabla XXXIX		
Valor p de la Tabla de contingencia de Preferencia Celular VS Género		
Chi Cuadrado	Grados de libertad	Valor p
1.863	2	.394

Fuente y Elaboración: Mario Plaza

Variable: Actividad

En la Tabla XL podemos apreciar que el valor p es mayor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Actividad no va a formar parte del modelo de la categoría Telefonía Celular.

Tabla XL		
Valor p de la Tabla de contingencia de Preferencia Celular VS Actividad		
Chi Cuadrado	Grados de libertad	Valor p
106.315	132	.951

Fuente y Elaboración: Mario Plaza

Variable: Top Celular

En la Tabla XLI podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula y determinar que las variables no son independientes. Por lo tanto la variable Top Celular va a formar parte del modelo dentro de la categoría Telefonía Celular.

Tabla XLI		
Valor p de la Tabla de contingencia de Preferencia Celular VS Top Celular		
Chi Cuadrado	Grados de libertad	Valor p
1.342.455	4	.000

Fuente y Elaboración: Mario Plaza

Variable: Razones de Preferencia

En la Tabla XLII podemos ver que el valor p es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula de independencia. Por lo que la variable razones de Preferencia pasa a formar parte del modelo dentro de la categoría Telefonía Celular.

Tabla XLII		
Valor p de la Tabla de contingencia de Preferencia Celular VS Razones de Preferencia		
Chi Cuadrado	Grados de libertad	Valor p
228.140	10	.000

Fuente y Elaboración: Mario Plaza

Variable: 1er Atributo

En la Tabla XLIII podemos apreciar que el valor p es inferior a 0,05 por lo que se rechaza la hipótesis nula que indica que las variables son independientes. En conclusión la variable 1er Atributo pasa a formar parte del modelo dentro de la categoría Telefonía Celular.

Tabla XLIII		
Valor p de la Tabla de contingencia de Preferencia Celular VS 1er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
28.245	16	.030

Fuente y Elaboración: Mario Plaza

Variable: 2do Atributo

En la Tabla XLIV podemos ver que el valor p no es mayor que 0,05 por lo que no existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto la variable 2do Atributo no va a formar parte del modelo dentro de la categoría Telefonía Celular.

Tabla XLIV		
Valor p de la Tabla de contingencia de Preferencia Celular VS 2do Atributo		
Chi Cuadrado	Grados de libertad	Valor p
24.821	16	.073

Fuente y Elaboración: Mario Plaza

Variable: 3er Atributo

En la Tabla XLV podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 3er Atributo no va a formar parte del modelo de la categoría Telefonía Celular.

Tabla XLV		
Valor p de la Tabla de contingencia de Preferencia Celular VS 3er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
21.906	16	.146

Fuente y Elaboración: Mario Plaza

Variable: 4to Atributo

En la Tabla XLVI podemos apreciar que el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 4to Atributo pasa a formar parte del modelo de la categoría Telefonía Celular.

Tabla XLVI		
Valor p de la Tabla de contingencia de Preferencia Celular VS 4to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
36.408	16	.003

Fuente y Elaboración: Mario Plaza

Variable: 5to Atributo

En la Tabla XLVII se puede apreciar que el valor p es menor a 0,05 por lo que podemos concluir que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 5to Atributo pasa a formar parte del modelo de la categoría Telefonía Celular.

Tabla XLVII		
Valor p de la Tabla de contingencia de Preferencia Celular VS 5to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
44.834	16	.000

Fuente y Elaboración: Mario Plaza

Variable: 6to Atributo

En la Tabla XLVIII podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 6to Atributo no va a formar parte del modelo de la categoría Telefonía Celular.

Tabla XLVIII		
Valor p de la Tabla de contingencia de Preferencia Celular VS 6to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
24.001	16	.089

Fuente y Elaboración: Mario Plaza

Variable: 7mo Atributo

Como podemos apreciar en la tabla XLIX el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 7mo Atributo pasa a formar parte del modelo de la categoría Telefonía Celular.

Tabla XLIX		
Valor p de la Tabla de contingencia de Preferencia Celular VS 7mo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
36.362	16	.003

Fuente y Elaboración: Mario Plaza

Variable: 8avo Atributo

Como se puede ver en la Tabla L el valor p es mayor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 8avo Atributo no va a formar parte del modelo dentro de la categoría Telefonía Celular.

Tabla L		
Valor p de la Tabla de contingencia de Preferencia Celular VS 8avo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
21.469	16	.161

Fuente y Elaboración: Mario Plaza

Variable: 9no Atributo

La Tabla LI muestra el valor p que es inferior a 0,05 por lo que se concluye que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 9no atributo pasa a formar parte del modelo de la categoría de Telefonía Celular.

Tabla LI		
Valor p de la Tabla de contingencia de Preferencia Celular VS 9no Atributo		
Chi Cuadrado	Grados de libertad	Valor p
35.181	16	.004

Fuente y Elaboración: Mario Plaza

Categoría: Bebidas Gaseosas-Colas

La variable de interés en esta categoría es Preferencia Gaseosa, la cual se va a analizar con las diferentes posibles variables que pueden conformar el modelo. Para todas y cada una de las posibles variables que conformen el modelo se realiza la siguiente prueba de hipótesis:

H_0 : Las variables son independientes

Vs.

H_1 : Rechazamos H_0

Variable: Ciudad

En la Tabla LII se aprecia el valor p de la tabla de contingencia Preferencia Gaseosa VS Ciudad, cuyo valor es menor a 0,05, por lo que cae en la región de rechazo, y podemos concluir que existe evidencia estadística para rechazar la hipótesis nula y por lo tanto las variables no son independientes, por ende la variable Ciudad va a formar parte del modelo dentro de ésta categoría.

Tabla LII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Ciudad		
Chi Cuadrado	Grados de libertad	Valor p
60.863	20	.000

Fuente y Elaboración: Mario Plaza

Variable: Región

En la Tabla LIII se muestra el valor p de la variable de interés con la variable Región, el cual es menor a 0,05 por lo que se concluye que existe evidencia para rechazar la hipótesis nula, mostrando que ambas variables no son independientes y por lo tanto la variable Región pasa a formar parte del modelo.

Tabla LIII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Región		
Chi Cuadrado	Grados de libertad	Valor p
24.922	5	.000

Fuente y Elaboración: Mario Plaza

Variable: Edad

En la Tabla LIV se puede apreciar que el valor p correspondiente a la tabla de contingencia de la variable de interés vs Edad es mayor a 0,05, por lo que concluimos que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Edad no va a formar parte de modelo dentro de la categoría Bebidas Gaseosas-Colas.

Tabla LIV		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Edad		
Chi Cuadrado	Grados de libertad	Valor p
230.160	255	.866

Fuente y Elaboración: Mario Plaza

Variable: Género

En la Tabla LV podemos apreciar que el valor p no es menor a 0,05, por lo tanto se concluye que no existe evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Género no va a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LV		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Género		
Chi Cuadrado	Grados de libertad	Valor p
8.754	5	.119

Fuente y Elaboración: Mario Plaza

Variable: Actividad

En la Tabla LVI podemos apreciar que el valor p es mayor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Actividad no va a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LVI		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Actividad		
Chi Cuadrado	Grados de libertad	Valor p
331.251	320	.321

Fuente y Elaboración: Mario Plaza

Variable: Top Gaseosa

En la Tabla LVII podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula y determinar que las variables no son independientes. Por lo tanto la variable Top Gaseosa va a formar parte del modelo dentro de la categoría Bebidas Gaseosas-Colas.

Tabla LVII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Top Gaseosa		
Chi Cuadrado	Grados de libertad	Valor p
2.677.684	25	.000

Fuente y Elaboración: Mario Plaza

Variable: Razones de Preferencia

En la Tabla LVIII podemos ver que el valor p es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula de independencia. Por lo que la variable razones de Preferencia pasa a formar parte del modelo dentro de la categoría Bebidas Gaseosas-Colas.

Tabla LVIII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS Razones de Preferencia		
Chi Cuadrado	Grados de libertad	Valor p
272.518	25	.000

Fuente y Elaboración: Mario Plaza

Variable: 1er Atributo

En la Tabla LIX podemos apreciar que el valor p no es inferior a 0,05 por lo que no existe suficiente evidencia estadística como para rechazar la hipótesis nula que indica que las variables son independientes. En conclusión la variable 1er Atributo no pasa a formar parte del modelo dentro de la categoría Bebidas Gaseosas-Colas.

Tabla LIX		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 1er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
54.659	40	.061

Fuente y Elaboración: Mario Plaza

Variable: 2do Atributo

En la Tabla LX podemos ver que el valor p es mayor a 0,05 por lo que existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto la variable 2do Atributo va a formar parte del modelo dentro de la categoría Bebidas Gaseosas-Colas.

Tabla LX		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 2do Atributo		
Chi Cuadrado	Grados de libertad	Valor p
34.588	40	.712

Fuente y Elaboración: Mario Plaza

Variable: 3er Atributo

En la Tabla LXI podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 3er Atributo no va a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LXI		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 3er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
38.620	40	.532

Fuente y Elaboración: Mario Plaza

Variable: 4to Atributo

En la Tabla LXII podemos apreciar que el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 4to Atributo pasa a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LXII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 4to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
68.488	40	.003

Fuente y Elaboración: Mario Plaza

Variable: 5to Atributo

En la Tabla LXIII se puede apreciar que el valor p no es menor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 5to Atributo no pasa a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LXIII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 5to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
50.324	40	.127

Fuente y Elaboración: Mario Plaza

Variable: 6to Atributo

En la Tabla LXIV podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 6to Atributo no va a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LXIV		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 6to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
52.695	40	.086

Fuente y Elaboración: Mario Plaza

Variable: 7mo Atributo

Como podemos apreciar en la Tabla LXV el valor p no es menor a 0,05 por lo que no existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 7mo Atributo no pasa a formar parte del modelo de la categoría Bebidas Gaseosas-Colas.

Tabla LXV		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 7mo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
49.668	40	.141

Fuente y Elaboración: Mario Plaza

Variable: 8avo Atributo

Como se puede ver en la Tabla LXVI el valor p no es mayor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 8avo Atributo pasa a formar parte del modelo dentro de la categoría Bebidas Gaseosas-Colas.

Tabla LXVI		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 8avo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
64.153	40	.009

Fuente y Elaboración: Mario Plaza

Variable: 9no Atributo

La Tabla LXVII muestra el valor p que es inferior a 0,05 por lo que se concluye que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 9no atributo pasa a formar parte del modelo de la categoría de Bebidas Gaseosas-Colas.

Tabla LXVII		
Valor p de la Tabla de contingencia de Preferencia Gaseosa VS 9no Atributo		
Chi Cuadrado	Grados de libertad	Valor p
66.660	40	.005

Fuente y Elaboración: Mario Plaza

Categoría: Bebidas Alcohólicas (en general)

La variable de interés en esta categoría es Preferencia Alcohol, la cual se va a analizar con las diferentes posibles variables que pueden conformar el modelo. Para todas y cada una de las posibles variables que conformen el modelo se realiza la siguiente prueba de hipótesis:

H_0 : Las variables son independientes

Vs.

H_1 : Rechazamos H_0

Variable: Ciudad

En la Tabla LXVIII se aprecia el valor p de la tabla de contingencia Preferencia Alcohol VS Ciudad, cuyo valor es menor a 0,05, por lo que cae en la región de rechazo, y podemos concluir que existe evidencia estadística para rechazar la hipótesis nula y por lo tanto las variables no son independientes, por ende la variable Ciudad va a formar parte del modelo dentro de ésta categoría.

Tabla LXVIII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Ciudad		
Chi Cuadrado	Grados de libertad	Valor p
225.785	20	.000

Fuente y Elaboración: Mario Plaza

Variable: Región

En la Tabla LXIX se muestra el valor p de la variable de interés con la variable Región, el cual es menor a 0,05 por lo que se concluye que existe evidencia para rechazar la hipótesis nula, mostrando que ambas variables no son independientes y por lo tanto la variable Región pasa a formar parte del modelo.

Tabla LXIX		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Región		
Chi Cuadrado	Grados de libertad	Valor p
125.130	5	.000

Fuente y Elaboración: Mario Plaza

Variable: Edad

En la Tabla LXX se puede apreciar que el valor p correspondiente a la tabla de contingencia de la variable de interés vs Edad es mayor a 0,05, por lo que concluimos que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Edad no va a formar parte de modelo dentro de la categoría Bebidas Alcohólicas.

Tabla LXX		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Edad		
Chi Cuadrado	Grados de libertad	Valor p
221.619	240	.797

Fuente y Elaboración: Mario Plaza

Variable: Género

En la Tabla LXXI podemos apreciar que el valor p no es menor a 0,05, por lo tanto se concluye que no existe evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Género no va a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXI		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Género		
Chi Cuadrado	Grados de libertad	Valor p
5.187	5	.393

Fuente y Elaboración: Mario Plaza

Variable: Actividad

En la Tabla LXXII podemos apreciar que el valor p no es mayor a 0,05 por lo que podemos concluir que existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Actividad pasa a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Actividad		
Chi Cuadrado	Grados de libertad	Valor p
337.406	270	.003

Fuente y Elaboración: Mario Plaza

Variable: Top Alcohol

En la Tabla LXXIII podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula y determinar que las variables no son independientes. Por lo tanto la variable Top Alcohol pasa a formar parte del modelo dentro de la categoría Bebidas Alcohólicas.

Tabla LXXIII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Top Alcohol		
Chi Cuadrado	Grados de libertad	Valor p
1.189.188	25	.000

Fuente y Elaboración: Mario Plaza

Variable: Razones de Preferencia

En la Tabla LXXIV podemos ver que el valor p es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula de independencia. Por lo que la variable razones de Preferencia pasa a formar parte del modelo dentro de la categoría Bebidas Alcohólicas.

Tabla LXXIV		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS Razones de Preferencia		
Chi Cuadrado	Grados de libertad	Valor p
130.859	25	.000

Fuente y Elaboración: Mario Plaza

Variable: 1er Atributo

En la Tabla LXXV podemos apreciar que el valor p no es inferior a 0,05 por lo que no existe suficiente evidencia estadística como para rechazar la hipótesis nula que indica que las variables son independientes. En conclusión la variable 1er Atributo no pasa a formar parte del modelo dentro de la categoría Bebidas Alcohólicas.

Tabla LXXV		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 1er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
40.670	40	.441

Fuente y Elaboración: Mario Plaza

Variable: 2do Atributo

En la Tabla LXXVI podemos ver que el valor p es mayor a 0,05 por lo que no existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto la variable 2do Atributo no va a formar parte del modelo dentro de la categoría Bebidas Alcohólicas.

Tabla LXXVI		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 2do Atributo		
Chi Cuadrado	Grados de libertad	Valor p
43.199	40	.336

Fuente y Elaboración: Mario Plaza

Variable: 3er Atributo

En la Tabla LXXVII podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 3er Atributo no va a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXVII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 3er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
36.079	40	.647

Fuente y Elaboración: Mario Plaza

Variable: 4to Atributo

En la Tabla LXXVIII podemos apreciar que el valor p no es menor a 0,05 por lo que no existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 4to Atributo no pasa a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXVIII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 4to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
55.246	40	.055

Fuente y Elaboración: Mario Plaza

Variable: 5to Atributo

En la Tabla LXXIX se puede apreciar que el valor p no es menor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 5to Atributo no pasa a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXIX		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 5to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
52.612	40	.087

Fuente y Elaboración: Mario Plaza

Variable: 6to Atributo

En la Tabla LXXX podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 6to Atributo no va a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXX		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 6to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
42.189	40	.377

Fuente y Elaboración: Mario Plaza

Variable: 7mo Atributo

Como podemos apreciar en la Tabla LXXXI el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 7mo Atributo pasa a formar parte del modelo de la categoría Bebidas Alcohólicas.

Tabla LXXXI		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 7mo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
84.270	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 8avo Atributo

Como se puede ver en la Tabla LXXXII el valor p no es mayor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 8avo Atributo pasa a formar parte del modelo dentro de la categoría Bebidas Alcohólicas.

Tabla LXXXII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 8avo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
83.186	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 9no Atributo

La Tabla LXXXIII muestra el valor p que es inferior a 0,05 por lo que se concluye que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 9no atributo pasa a formar parte del modelo de la categoría de Bebidas Alcohólicas.

Tabla LXXXIII		
Valor p de la Tabla de contingencia de Preferencia Alcohol VS 9no Atributo		
Chi Cuadrado	Grados de libertad	Valor p
86.021	40	.000

Fuente y Elaboración: Mario Plaza

Categoría: Atún

La variable de interés en esta categoría es Preferencia Atún, la cual se va a analizar con las diferentes posibles variables que pueden conformar el modelo. Para todas y cada una de las posibles variables que conformen el modelo se realiza la siguiente prueba de hipótesis:

H_0 : Las variables son independientes

Vs.

H_1 : Rechazamos H_0

Variable: Ciudad

En la Tabla LXXXIV se aprecia el valor p de la tabla de contingencia Preferencia Atún VS Ciudad, cuyo valor es menor a 0,05, por lo que cae en la región de rechazo, y podemos concluir que existe evidencia estadística para rechazar la hipótesis nula y por lo tanto las variables no son independientes, por ende la variable Ciudad va a formar parte del modelo dentro de ésta categoría.

Tabla LXXXIV		
Valor p de la Tabla de contingencia de Preferencia Atún VS Ciudad		
Chi Cuadrado	Grados de libertad	Valor p
509.536	20	.000

Fuente y Elaboración: Mario Plaza

Variable: Región

En la Tabla LXXXV se muestra el valor p de la variable de interés con la variable Región, el cual es menor a 0,05 por lo que se concluye que existe evidencia para rechazar la hipótesis nula, mostrando que ambas variables no son independientes y por lo tanto la variable Región pasa a formar parte del modelo.

Tabla LXXXV		
Valor p de la Tabla de contingencia de Preferencia Atún VS Región		
Chi Cuadrado	Grados de libertad	Valor p
46.067	5	.000

Fuente y Elaboración: Mario Plaza

Variable: Edad

En la Tabla LXXXVI se puede apreciar que el valor p correspondiente a la tabla de contingencia de la variable de interés vs Edad es mayor a 0,05, por lo que concluimos que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Edad no va a formar parte de modelo dentro de la categoría Atún.

Tabla LXXXVI		
Valor p de la Tabla de contingencia de Preferencia Atún VS Edad		
Chi Cuadrado	Grados de libertad	Valor p
226.992	255	.896

Fuente y Elaboración: Mario Plaza

Variable: Género

En la Tabla LXXXVII podemos apreciar que el valor p no es menor a 0,05, por lo tanto se concluye que no existe evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Género no va a formar parte del modelo de la categoría Atún.

Tabla LXXXVII		
Valor p de la Tabla de contingencia de Preferencia Atún VS Género		
Chi Cuadrado	Grados de libertad	Valor p
6.154	5	.292

Fuente y Elaboración: Mario Plaza

Variable: Actividad

En la Tabla LXXXVIII podemos apreciar que el valor p es mayor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Actividad no pasa a formar parte del modelo de la categoría Atún.

Tabla LXXXVIII		
Valor p de la Tabla de contingencia de Preferencia Atún VS Actividad		
Chi Cuadrado	Grados de libertad	Valor p
294.164	330	.923

Fuente y Elaboración: Mario Plaza

Variable: Top Atún

En la Tabla LXXXIX podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula y determinar que las variables no son independientes. Por lo tanto la variable Top Atún pasa a formar parte del modelo dentro de la categoría Atún.

Tabla LXXXIX		
Valor p de la Tabla de contingencia de Preferencia Atún VS Top Atún		
Chi Cuadrado	Grados de libertad	Valor p
2.827.718	25	.000

Fuente y Elaboración: Mario Plaza

Variable: Razones de Preferencia

En la Tabla XC podemos ver que el valor p es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula de independencia. Por lo que la variable razones de Preferencia pasa a formar parte del modelo dentro de la categoría Atún.

Tabla XC		
Valor p de la Tabla de contingencia de Preferencia Atún VS Razones de Preferencia		
Chi Cuadrado	Grados de libertad	Valor p
72.944	20	.000

Fuente y Elaboración: Mario Plaza

Variable: 1er Atributo

En la Tabla XCI podemos apreciar que el valor p no es inferior a 0,05 por lo que no existe suficiente evidencia estadística como para rechazar la hipótesis nula que indica que las variables son independientes. En conclusión la variable 1er Atributo no pasa a formar parte del modelo dentro de la categoría Atún.

Tabla XCI		
Valor p de la Tabla de contingencia de Preferencia Atún VS 1er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
37.400	40	.588

Fuente y Elaboración: Mario Plaza

Variable: 2do Atributo

En la Tabla XCII podemos ver que el valor p es mayor a 0,05 por lo que no existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto la variable 2do Atributo no va a formar parte del modelo dentro de la categoría Atún.

Tabla XCII		
Valor p de la Tabla de contingencia de Preferencia Atún VS 2do Atributo		
Chi Cuadrado	Grados de libertad	Valor p
53.996	40	.069

Fuente y Elaboración: Mario Plaza

Variable: 3er Atributo

En la Tabla XCIII podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 3er Atributo no va a formar parte del modelo de la categoría Atún.

Tabla XCIII		
Valor p de la Tabla de contingencia de Preferencia Atún VS 3er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
47.728	40	.187

Fuente y Elaboración: Mario Plaza

Variable: 4to Atributo

En la Tabla XCIV podemos apreciar que el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 4to Atributo pasa a formar parte del modelo de la categoría Atún.

Tabla XCIV		
Valor p de la Tabla de contingencia de Preferencia Atún VS 4to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
94.826	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 5to Atributo

En la Tabla XCV se puede apreciar que el valor p es menor a 0,05 por lo que podemos concluir que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 5to Atributo pasa a formar parte del modelo de la categoría Atún.

Tabla XCV		
Valor p de la Tabla de contingencia de Preferencia Atún VS 5to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
59.791	40	.023

Fuente y Elaboración: Mario Plaza

Variable: 6to Atributo

En la Tabla XCVI podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 6to Atributo pasa a formar parte del modelo de la categoría Atún.

Tabla XCVI		
Valor p de la Tabla de contingencia de Preferencia Atún VS 6to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
67.929	40	.004

Fuente y Elaboración: Mario Plaza

Variable: 7mo Atributo

Como podemos apreciar en la Tabla XCVII el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 7mo Atributo pasa a formar parte del modelo de la categoría Atún.

Tabla XCVII		
Valor p de la Tabla de contingencia de Preferencia Atún VS 7mo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
107.208	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 8avo Atributo

Como se puede ver en la Tabla XCVIII el valor p no es mayor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 8avo Atributo pasa a formar parte del modelo dentro de la categoría Atún.

Tabla XCVIII		
Valor p de la Tabla de contingencia de Preferencia Atún VS 8avo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
116.699	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 9no Atributo

La Tabla XCIX muestra el valor p que es inferior a 0,05 por lo que se concluye que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 9no atributo pasa a formar parte del modelo de la categoría de Atún.

Tabla XCIX		
Valor p de la Tabla de contingencia de Preferencia Atún VS 9no Atributo		
Chi Cuadrado	Grados de libertad	Valor p
114.207	40	.000

Fuente y Elaboración: Mario Plaza

Categoría: Chocolate en barra

La variable de interés en esta categoría es Preferencia Chocolate, la cual se va a analizar con las diferentes posibles variables que pueden conformar el modelo. Para todas y cada una de las posibles variables que conformen el modelo se realiza la siguiente prueba de hipótesis:

H_0 : Las variables son independientes

Vs.

H_1 : Rechazamos H_0

Variable: Ciudad

En la Tabla C se aprecia el valor p de la tabla de contingencia Preferencia Chocolate VS Ciudad, cuyo valor es menor a 0,05, por lo que cae en la región de rechazo, y podemos concluir que existe evidencia estadística para rechazar la hipótesis nula y por lo tanto las variables no son independientes, por ende la variable Ciudad va a formar parte del modelo dentro de ésta categoría.

Tabla C		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Ciudad		
Chi Cuadrado	Grados de libertad	Valor p
127.965	20	.000

Fuente y Elaboración: Mario Plaza

Variable: Región

En la Tabla CI se muestra el valor p de la variable de interés con la variable Región, el cual es menor a 0,05 por lo que se concluye que existe evidencia para rechazar la hipótesis nula, mostrando que ambas variables no son independientes y por lo tanto la variable Región pasa a formar parte del modelo.

Tabla CI		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Región		
Chi Cuadrado	Grados de libertad	Valor p
73.623	5	.000

Fuente y Elaboración: Mario Plaza

Variable: Edad

En la Tabla CII se puede apreciar que el valor p correspondiente a la tabla de contingencia de la variable de interés vs Edad es menor a 0,05, por lo que concluimos que existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Edad va a formar parte de modelo dentro de la categoría Chocolate en barra.

Tabla CII		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Edad		
Chi Cuadrado	Grados de libertad	Valor p
304.282	255	.019

Fuente y Elaboración: Mario Plaza

Variable: Género

En la Tabla CIII podemos apreciar que el valor p no es menor a 0,05, por lo tanto se concluye que no existe evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Género no va a formar parte del modelo de la categoría Chocolate en barra.

Tabla CIII		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Género		
Chi Cuadrado	Grados de libertad	Valor p
9.260	5	.099

Fuente y Elaboración: Mario Plaza

Variable: Actividad

En la Tabla CIV podemos apreciar que el valor p es mayor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Actividad no pasa a formar parte del modelo de la categoría Chocolate en barra.

Tabla CIV		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Actividad		
Chi Cuadrado	Grados de libertad	Valor p
244.023	280	.941

Fuente y Elaboración: Mario Plaza

Variable: Top Chocolate

En la Tabla CV podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula y determinar que las variables no son independientes. Por lo tanto la variable Top Chocolate pasa a formar parte del modelo dentro de la categoría Chocolate en barra.

Tabla CV		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Top Chocolate		
Chi Cuadrado	Grados de libertad	Valor p
1.742.926	25	.000

Fuente y Elaboración: Mario Plaza

Variable: Razones de Preferencia

En la Tabla CVI podemos ver que el valor p es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula de independencia. Por lo que la variable razones de Preferencia pasa a formar parte del modelo dentro de la categoría Chocolate en barra.

Tabla CVI		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS Razones de Preferencia		
Chi Cuadrado	Grados de libertad	Valor p
73.519	25	.000

Fuente y Elaboración: Mario Plaza

Variable: 1er Atributo

En la Tabla CVII podemos apreciar que el valor p es inferior a 0,05 por lo que existe suficiente evidencia estadística como para rechazar la hipótesis nula que indica que las variables son independientes. En conclusión la variable 1er Atributo pasa a formar parte del modelo dentro de la categoría Chocolate en barra.

Tabla CVII		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 1er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
59.449	40	.024

Fuente y Elaboración: Mario Plaza

Variable: 2do Atributo

En la Tabla CVIII podemos ver que el valor p es mayor a 0,05 por lo que no existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto la variable 2do Atributo no va a formar parte del modelo dentro de la categoría Chocolate en barra.

Tabla CVIII		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 2do Atributo		
Chi Cuadrado	Grados de libertad	Valor p
43.929	40	.309

Fuente y Elaboración: Mario Plaza

Variable: 3er Atributo

En la Tabla CIX podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 3er Atributo no va a formar parte del modelo de la categoría Chocolate en barra.

Tabla CIX		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 3er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
39.629	40	.487

Fuente y Elaboración: Mario Plaza

Variable: 4to Atributo

En la Tabla CX podemos apreciar que el valor p no es menor a 0,05 por lo que no existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 4to Atributo no pasa a formar parte del modelo de la categoría Chocolate en barra.

Tabla CX		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 4to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
51.537	40	.105

Fuente y Elaboración: Mario Plaza

Variable: 5to Atributo

En la Tabla CXI se puede apreciar que el valor p no es menor a 0,05 por lo que podemos concluir que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 5to Atributo no pasa a formar parte del modelo de la categoría Chocolate en barra.

Tabla CXI		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 5to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
40.785	40	.436

Fuente y Elaboración: Mario Plaza

Variable: 6to Atributo

En la Tabla CXII podemos apreciar que el valor p no es menor a 0,05 por lo que se concluye que no existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 6to Atributo no pasa a formar parte del modelo de la categoría Chocolate en barra.

Tabla CXII		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 6to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
36.105	40	.646

Fuente y Elaboración: Mario Plaza

Variable: 7mo Atributo

Como podemos apreciar en la Tabla CXIII el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 7mo Atributo pasa a formar parte del modelo de la categoría Chocolate en barra.

Tabla CXIII		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 7mo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
56.636	40	.042

Fuente y Elaboración: Mario Plaza

Variable: 8avo Atributo

Como se puede ver en la Tabla CXIV el valor p no es mayor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 8avo Atributo pasa a formar parte del modelo dentro de la categoría Chocolate en barra.

Tabla CXIV		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 8avo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
90.391	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 9no Atributo

La Tabla CXV muestra el valor p que es inferior a 0,05 por lo que se concluye que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 9no atributo pasa a formar parte del modelo de la categoría de Chocolate en barra.

Tabla CXV		
Valor p de la Tabla de contingencia de Preferencia Chocolate VS 9no Atributo		
Chi Cuadrado	Grados de libertad	Valor p
81.095	40	.000

Fuente y Elaboración: Mario Plaza

Categoría: Café (en general)

La variable de interés en esta categoría es Preferencia Café, la cual se va a analizar con las diferentes posibles variables que pueden conformar el modelo. Para todas y cada una de las posibles variables que conformen el modelo se realiza la siguiente prueba de hipótesis:

H_0 : Las variables son independientes

Vs.

H₁: Rechazamos H₀

Variable: Ciudad

En la Tabla CXVI se aprecia el valor p de la tabla de contingencia Preferencia Café VS Ciudad, cuyo valor es menor a 0,05, por lo que cae en la región de rechazo, y podemos concluir que existe evidencia estadística para rechazar la hipótesis nula y por lo tanto las variables no son independientes, por ende la variable Ciudad va a formar parte del modelo dentro de ésta categoría.

Tabla CXVI		
Valor p de la Tabla de contingencia de Preferencia Café VS Ciudad		
Chi Cuadrado	Grados de libertad	Valor p
492.628	20	.000

Fuente y Elaboración: Mario Plaza

Variable: Región

En la Tabla CXVII se muestra el valor p de la variable de interés con la variable Región, el cual es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula, mostrando que ambas variables no son independientes y por lo tanto la variable Región pasa a formar parte del modelo.

Tabla CXVII		
Valor p de la Tabla de contingencia de Preferencia Café VS Región		
Chi Cuadrado	Grados de libertad	Valor p
186.735	5	.000

Fuente y Elaboración: Mario Plaza

Variable: Edad

En la Tabla CXVIII se puede apreciar que el valor p correspondiente a la tabla de contingencia de la variable de interés vs Edad no es menor a 0,05, por lo que concluimos que no existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Edad no pasa a formar parte de modelo dentro de la categoría Café.

Tabla CXVIII		
Valor p de la Tabla de contingencia de Preferencia Café VS Edad		
Chi Cuadrado	Grados de libertad	Valor p
278.124	250	.107

Fuente y Elaboración: Mario Plaza

Variable: Género

En la Tabla CXIX podemos apreciar que el valor p es menor a 0,05, por lo tanto se concluye que existe evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Género pasa a formar parte del modelo de la categoría Café.

Tabla CXIX		
Valor p de la Tabla de contingencia de Preferencia Café VS Género		
Chi Cuadrado	Grados de libertad	Valor p
17.444	5	.004

Fuente y Elaboración: Mario Plaza

Variable: Actividad

En la Tabla CXX podemos apreciar que el valor p no es mayor a 0,05 por lo que podemos concluir que existe suficiente evidencia estadística como para rechazar la hipótesis nula. Por lo tanto la variable Actividad pasa a formar parte del modelo de la categoría Café.

Tabla CXX		
Valor p de la Tabla de contingencia de Preferencia Café VS Actividad		
Chi Cuadrado	Grados de libertad	Valor p
368.609	310	.012

Fuente y Elaboración: Mario Plaza

Variable: Top Café

En la Tabla CXXI podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula y determinar que las variables no son independientes. Por lo tanto la variable Top Café pasa a formar parte del modelo dentro de la categoría Café.

Tabla CXXI		
Valor p de la Tabla de contingencia de Preferencia Café VS Top Café		
Chi Cuadrado	Grados de libertad	Valor p
3,284791	25	.000

Fuente y Elaboración: Mario Plaza

Variable: Razones de Preferencia

En la Tabla CXXII podemos ver que el valor p es menor a 0,05 por lo que se concluye que existe evidencia estadística para rechazar la hipótesis nula de independencia. Por lo que la variable razones de Preferencia pasa a formar parte del modelo dentro de la categoría Café.

Tabla CXXII		
Valor p de la Tabla de contingencia de Preferencia Café VS Razones de Preferencia		
Chi Cuadrado	Grados de libertad	Valor p
82.722	25	.000

Fuente y Elaboración: Mario Plaza

Variable: 1er Atributo

En la Tabla CXXIII podemos apreciar que el valor p es inferior a 0,05 por lo que existe suficiente evidencia estadística como para rechazar la hipótesis nula que indica que las variables son independientes. En conclusión la variable 1er Atributo pasa a formar parte del modelo dentro de la categoría Café.

Tabla CXXIII		
Valor p de la Tabla de contingencia de Preferencia Café VS 1er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
81.384	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 2do Atributo

En la Tabla CXXIV podemos ver que el valor p no es mayor a 0,05 por lo que existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto la variable 2do Atributo pasa a formar parte del modelo dentro de la categoría Café.

Tabla CXXIV		
Valor p de la Tabla de contingencia de Preferencia Café VS 2do Atributo		
Chi Cuadrado	Grados de libertad	Valor p
87.258	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 3er Atributo

En la Tabla CXXV podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 3er Atributo pasa a formar parte del modelo de la categoría Café.

Tabla CXXV		
Valor p de la Tabla de contingencia de Preferencia Café VS 3er Atributo		
Chi Cuadrado	Grados de libertad	Valor p
72.776	40	.001

Fuente y Elaboración: Mario Plaza

Variable: 4to Atributo

En la Tabla CXXVI podemos apreciar que el valor p es menor a 0,05 por lo que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 4to Atributo pasa a formar parte del modelo de la categoría Café.

Tabla CXXVI		
Valor p de la Tabla de contingencia de Preferencia Café VS 4to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
102.305	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 5to Atributo

En la Tabla CXXVII se puede apreciar que el valor p es menor a 0,05 por lo que podemos concluir que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 5to Atributo pasa a formar parte del modelo de la categoría Café.

Tabla CXXVII		
Valor p de la Tabla de contingencia de Preferencia Café VS 5to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
115.259	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 6to Atributo

En la Tabla CXXVIII podemos apreciar que el valor p es menor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 6to Atributo pasa a formar parte del modelo de la categoría Café.

Tabla CXXVIII		
Valor p de la Tabla de contingencia de Preferencia Café VS 6to Atributo		
Chi Cuadrado	Grados de libertad	Valor p
98.218	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 7mo Atributo

Como podemos apreciar en la Tabla CXXIX el valor p es menor a 0,05 por lo que existe evidencia estadística para rechazar la hipótesis nula de independencia de las variables. Por lo tanto la variable 7mo Atributo pasa a formar parte del modelo de la categoría Café.

Tabla CXXIX		
Valor p de la Tabla de contingencia de Preferencia Café VS 7mo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
124.579	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 8avo Atributo

Como se puede ver en la Tabla CXXX el valor p no es mayor a 0,05 por lo que se concluye que existe suficiente evidencia estadística como para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 8avo Atributo pasa a formar parte del modelo dentro de la categoría Café.

Tabla CXXX		
Valor p de la Tabla de contingencia de Preferencia Café VS 8avo Atributo		
Chi Cuadrado	Grados de libertad	Valor p
123.557	40	.000

Fuente y Elaboración: Mario Plaza

Variable: 9no Atributo

La Tabla CXXXI muestra el valor p que es inferior a 0,05 por lo que se concluye que existe suficiente evidencia estadística para rechazar la hipótesis nula de independencia de variables. Por lo tanto la variable 9no atributo pasa a formar parte del modelo de la categoría de Café.

Tabla CXXXI		
Valor p de la Tabla de contingencia de Preferencia Café VS 9no Atributo		
Chi Cuadrado	Grados de libertad	Valor p
125.266	40	.000

Fuente y Elaboración: Mario Plaza

3.3. Regresión Logística Multinomial

Después de haber realizado las tablas de contingencia para determinar independencia entre las variables procedemos a verificar cómo queda el modelo previamente para cada una de las categorías, para que el siguiente paso sea determinar en base a técnicas de regresión logística multinomial si las variables realmente van a formar parte del modelo.

El primer paso es determinar los respectivos valores del estadístico AIC para cada modelo de cada categoría. Y posteriormente confirmar con el valor p si la variable es o no un factor significativo para el modelo.

Categoría: Telefonía Celular

En base a los resultados obtenidos a través de las Tablas de contingencia, el modelo dentro de ésta categoría es el siguiente:

Preferencia Celular ~ Ciudad + Región + Top Celular + Razones de Preferencia + 1er Atributo + 4to Atributo + 5to Atributo + 7mo Atributo + 9no Atributo

Cuyo valor AIC es 599,96.

En consecuencia consideramos un segundo modelo cuyo valor AIC sea inferior para poder predecir de forma más efectiva la variable dependiente, quedando de la siguiente forma:

Preferencia Celular ~ Ciudad + Región + Top Celular + Razones de Preferencia

Cuyo valor AIC es 489,72.

En base a este modelo procedemos a obtener los coeficientes a través del software estadístico R-Studio. Debido a que el método de regresión logística multinomial toma una categoría como base vamos en este caso a tener dos modelos. Sabemos que dentro de Telefonía celular contamos con las empresas: Claro, Movistar y CNT. En este caso la categoría que el software consideró como categoría con la cual comparar es Claro, por lo tanto tenemos dos modelos en base a las otras dos categorías, las cuales miden la posibilidad de preferir la empresa X por encima de la empresa Y. Los coeficientes de ambos modelos podemos apreciarlos en las Tablas CXXXII y CXXXIII.

Tabla CXXXII							
Valores de los Coeficientes del modelo de Telefonía Celular							
Categorías	(Intercept)	Region Sierra	Razones_Pref Por las promociones	Razones_PrefPor los servicios que brinda	Razones_PrefPor su cobertura	Razones_PrefPor sus contactos/Amigos	Razones_PrefPorque la tarifa de llamadas es económica
CNT	-8,155333	4,2846852	0,04056133	-26,64832	-15,53149	-20,9084699	-27,542612
Movistar	-2,314299	-0,436266	1,562858	-1,09929	-1,085229	-0,6521778	1,016074

Fuente y Elaboración: Mario Plaza

Tabla CXXXIII						
Valores de los Coeficientes del modelo de Telefonía Celular						
Categorías	CiudadGuayaquil	Ciudad Machala	CiudadManta	CiudadQuito	TopCnt	TopMovistar
CNT	5,1118476	-7,036155	-10,515711	-17,581403	32,14245	3,018871
Movistar	-0,3873507	-0,427181	-1,063501	-0,4885015	-14,09939	5,004178

Fuente y Elaboración: Mario Plaza

En base a los valores de los coeficientes de las tablas anteriores, los modelos son los siguientes:

$$\begin{aligned} \eta_{Movistar} = & -2,314299 - 0,4362663 \text{Región (Sierra)} \\ & + 1,562858 \text{ Razon de preferencia (Por las promociones)} \\ & - 1,09929 \text{ Razon de preferencia (Por los servicios que brinda)} \\ & - 1,085229 \text{ Razon de preferencia (Por su cobertura)} \\ & - 0,6521778 \text{ Razon de preferencia (Por sus contactos)} \\ & + 1,016074 \text{ Razon de preferencia (Por tarifa económica)} \\ & - 0,3873507 \text{ Ciudad (Guayaquil)} - 0,4271808 \text{ Ciudad (Machala)} \\ & - 1,063501 \text{ Ciudad (Manta)} - 0,488502 \text{ Ciudad (Quito)} \\ & - 14,09939 \text{ TOP(CNT)} + 5,004178 \text{ TOP(Movistar)} \end{aligned}$$

El modelo nos indica que aquellos cuya Razón de preferencia sean Las promociones y Las Tarifas económicas y que en la pregunta del TOM hayan mencionado a la empresa Movistar aumentan la posibilidad de que prefieran a Movistar por encima de Claro.

$$\begin{aligned} \eta_{CNT} = & -8,155333 + 4,2846852 \text{Región (Sierra)} \\ & + 0,04056133 \text{ Razon de preferencia (Por las promociones)} \\ & - 26,64832 \text{ Razon de preferencia (Por los servicios que brinda)} \\ & - 15,531489 \text{ Razon de preferencia (Por su cobertura)} \\ & - 20,9084599 \text{ Razon de preferencia (Por sus contactos)} \\ & - 27,542612 \text{ Razon de preferencia (Por tarifa económica)} \\ & + 5,1118476 \text{ Ciudad (Guayaquil)} - 7,0361546 \text{ Ciudad (Machala)} \\ & - 10,515711 \text{ Ciudad (Manta)} - 17,5814 \text{ Ciudad (Quito)} \\ & + 32,14245 \text{ TOP(CNT)} + 3,018871 \text{ TOP(Movistar)} \end{aligned}$$

El modelo nos indica que aquellos que sean de la Región Sierra, ciudad Cuenca, de la ciudad de Guayaquil, que su razón de preferencia sea Por las promociones y que en la pregunta del TOM hayan mencionado Movistar o CNT, aumentan la posibilidad de escoger a la empresa CNT por encima de Claro.

Los valores de las posibilidades más altas que determinan que prefieren a Movistar por encima de Claro, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Claro por encima de Movistar se encuentran en la Tabla CXXXIV.

Tabla CXXXIV						
Valores de las posibilidades del modelo de Telefonía Celular categoría Movistar						
(Intercept)	RegionCosta	TopMovistar	CiudadCuenca	Razones_PrefPor las promociones	Valor de la posibilidad	70,29755369
(Intercept)	RegionCosta	TopMovistar	CiudadGuayaquil	Razones_PrefPor las promociones	Valor de la posibilidad	47,72170377
(Intercept)	RegionSierra	TopMovistar	CiudadManta	Razones_PrefPor su cobertura	Valor de la posibilidad	1,110580332
(Intercept)	RegionSierra	TopMovistar	CiudadManta	Razones_PrefPor los servicios que brinda	Valor de la posibilidad	1,095073736
(Intercept)	RegionSierra	TopCnt	CiudadManta	Razones_PrefPor su cobertura	Valor de la posibilidad	5,61017E-09
(Intercept)	RegionSierra	TopCnt	CiudadManta	Razones_PrefPor los servicios que brinda	Valor de la posibilidad	5,53183E-09

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a CNT por encima de Claro, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Claro por encima de CNT se encuentran en la Tabla CXXXV.

Tabla CXXXV						
Valores de las posibilidades del modelo de Telefonía Celular categoría CNT						
(Intercept)	RegionSierra	TopCnt	CiudadGuayaquil	Razones_PrefPor las promociones	Valor de la posibilidad	3,28058E+14
(Intercept)	RegionSierra	TopCnt	CiudadGuayaquil	Razones_PrefOtras razones	Valor de la posibilidad	3,15018E+14
(Intercept)	RegionSierra	TopCnt	CiudadMachala	Razones_PrefPor sus contactos/Amigos	Valor de la posibilidad	1,387047934
(Intercept)	RegionCosta	TopMovistar	CiudadGuayaquil	Razones_PrefPor las promociones	Valor de la posibilidad	1,016074761
(Intercept)	RegionSierra	TopClaro	CiudadQuito	Razones_PrefPor los servicios que brinda	Valor de la posibilidad	1,28906E-21
(Intercept)	RegionSierra	TopClaro	CiudadQuito	Razones_PrefPorque la tarifa de llamadas es económica	Valor de la posibilidad	5,27094E-22

Fuente y Elaboración: Mario Plaza

Los demás valores de posibilidades se encuentran en la sección Anexos.

En la tabla de clasificación que es la Tabla CXXXVI nos permite observar que de los 689 entrevistados que prefirieron Claro, 649 fueron pronosticados como tal, es decir un 94%. En conclusión el modelo nos proporciona un 92,6% de correcta clasificación en forma general.

Tabla CXXXVI				
Tabla de Clasificación				
Observado	Pronosticado			
	Claro	Movistar	CNT	Porcentaje Correcto
Claro	649	39	1	94.2%
Movistar	29	255	0	89.8%
CNT	2	2	8	66.7%
Porcentaje global	69.0%	30.1%	.9%	92.6%

Fuente y Elaboración: Mario Plaza

Categoría: Bebidas Gaseosas-Colas

En base a los resultados obtenidos a través de las Tablas de contingencia, el modelo dentro de ésta categoría es el siguiente:

Preferencia Gaseosa ~ Ciudad + Región + Top Gaseosa + Razones de Preferencia + 4to Atributo + 8avo Atributo + 9no Atributo

Cuyo valor AIC es 1116,31.

En consecuencia consideramos un segundo modelo cuyo valor AIC sea inferior para poder predecir de forma más efectiva la variable dependiente, quedando de la siguiente forma:

Preferencia Gaseosa ~ Ciudad + Región + Top Gaseosa + Razones de Preferencia

Cuyo valor AIC es 998,6.

En base a este modelo procedemos a obtener los coeficientes a través del software estadístico R-Studio. Debido a que el método de regresión logística multinomial toma una categoría como base vamos en este caso a tener cinco modelos. Sabemos que dentro de Bebidas Gaseosas-Colas contamos con las marcas: Coca Cola, Fanta, Manzana, Sprite, Pepsi y Otras Marcas. En este caso la categoría que el software consideró como categoría con la cual comparar es Coca Cola, por lo tanto tenemos cinco modelos en base a las otras categorías, las cuales miden la posibilidad de preferir la empresa X por encima de la empresa Y. Los coeficientes de los modelos podemos apreciarlos en las Tablas CXXXVII, CXXXVIII y CXXXIX.

Tabla CXXXVII						
Valores de los Coeficientes del modelo de Bebidas Gaseosas-Colas						
Categorías	(Intercept)	Region Sierra	Razones_Pref Otras razones	Razones_PrefPor el gas	Razones_PrefPor que contiene menos colorante	Razones_Pref Porque es una tradición familiar
Sprite	-2,648208	-0,8564556	-0,8250828	0,07312478	2,9218077	-121,657514
Fanta	-3,781243	-1,1315506	-0,3447377	1,98665923	3,7431716	-1,66388
Manzana	-4,210045	-1,5631235	-36,2942256	-42,44604237	3,4625571	-32,972789
Pepsi	-3,557255	-1,49189	-2,82294669	-1,49203568	-0,6316613	-40,57094
Otras Marcas	-2,874925	-1,2654912	1,0886947	0,57948351	2,7879397	-3,228725

Fuente y Elaboración: Mario Plaza

Tabla CXXXVIII					
Valores de los Coeficientes del modelo de Bebidas Gaseosas-Colas					
Categorías	Razones_PrefP orque me gusta su sabor	Ciudad Guayaquil	Ciudad Machala	Ciudad Manta	Ciudad Quito
Sprite	-1,2540074	0,1080157	0,4759221	-2,3756906	-0,5231451
Fanta	0,9217035	-1,0608411	-0,7415878	-0,8472632	-0,2906761
Manzana	0,6174394	-0,1927925	0,3583294	-2,8124588	-2,2037379
Pepsi	-0,6011471	-0,1434643	0,324169	-2,2460702	1,6543147
Otras Marcas	0,4843526	-0,1423481	-0,2332967	-1,2337895	-0,1371354

Fuente y Elaboración: Mario Plaza

Tabla CXXXIX					
Valores de los Coeficientes del modelo de Bebidas Gaseosas-Colas					
Categorías	TopFanta	Top Manzana	TopOtras Marcas	TopPepsi	TopSprite
Sprite	5,579648	-43,81757	2,70279	45,56994	7,571803
Fanta	8,000871	-31,530007	3,475248	10,1272	3,765659
Manzana	-5,938647	7,740326	3,693479	191,70086	4,845764
Pepsi	-8,714219	-13,31925	-6,453623	194,99206	-7,376956
Otras Marcas	4,485634	3,466636	5,892949	189,82467	2,716698

Fuente y Elaboración: Mario Plaza

En base a los valores de los coeficientes de las tablas anteriores, los modelos son los siguientes:

$$\begin{aligned}
 \eta_{Pepsi} = & -3,557255 - 1,49189 \text{ Región (Sierra)} \\
 & + 2,822946 \text{ Razon de preferencia (Otras razones)} \\
 & - 1,49203 \text{ Razon de preferencia (Por el gas)} \\
 & - 0,631666 \text{ Razon de preferencia (Contiene menos colorante)} \\
 & - 40,57094 \text{ Razon de preferencia (Tradición)} \\
 & - 0,6011471 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & - 0,1434643 \text{ Ciudad (Guayaquil)} + 0,324169 \text{ Ciudad (Machala)} \\
 & - 2,2460702 \text{ Ciudad (Manta)} + 1,6543147 \text{ Ciudad (Quito)} \\
 & - 8,714219 \text{ TOP(Fanta)} - 13,31925 \text{ TOP(Manzana)} \\
 & - 6,453623 \text{ TOP(Otras marcas)} + 194,99206 \text{ TOP(Pepsi)} \\
 & - 7,376698 \text{ TOP(Sprite)}
 \end{aligned}$$

El modelo nos indica que aquellos cuya Razón de preferencia sean Otras razones, de las ciudades de Machala y Quito y que en la pregunta del TOM

hayan mencionado a la marca Pepsi aumentan la posibilidad de que prefieran a Pepsi por encima de Coca Cola.

$$\begin{aligned} \eta_{\text{Sprite}} = & -2,648208 - 0,8564556 \text{ Región (Sierra)} \\ & - 0,8250828 \text{ Razon de preferencia (Otras razones)} \\ & + 0,07312478 \text{ Razon de preferencia (Por el gas)} \\ & + 2,9218077 \text{ Razon de preferencia (Contiene menos colorante)} \\ & - 121,657514 \text{ Razon de preferencia (Tradición)} \\ & - 121,2540074 \text{ Razon de preferencia (Me gusta su sabor)} \\ & + 0,1080157 \text{ Ciudad (Guayaquil)} + 0,4759221 \text{ Ciudad (Machala)} \\ & - 2,3756906 \text{ Ciudad (Manta)} - 0,5231451 \text{ Ciudad (Quito)} \\ & + 5,579648 \text{ TOP(Fanta)} - 43,81757 \text{ TOP(Manzana)} \\ & + 2,70279 \text{ TOP(Otras marcas)} + 45,56994 \text{ TOP(Pepsi)} \\ & + 7,571803 \text{ TOP(Sprite)} \end{aligned}$$

El modelo nos indica que aquellos que sus razones de preferencia sean Por el gas y Contiene menos Colorante, procedentes de las ciudades de Guayaquil y Machala, que en la pregunta del TOM hayan mencionado Fanta, Otras marcas, Pepsi o Sprite, aumentan la posibilidad de escoger a la Sprite por encima de Coca Cola.

$$\begin{aligned} \eta_{\text{fanta}} = & -3,781243 - 1,1315506 \text{ Región (Sierra)} \\ & - 0,3447377 \text{ Razon de preferencia (Otras razones)} \\ & + 1,98665923 \text{ Razon de preferencia (Por el gas)} \\ & + 3,7431716 \text{ Razon de preferencia (Contiene menos colorante)} \\ & - 1,66388 \text{ Razon de preferencia (Tradición)} \\ & + 0,9217035 \text{ Razon de preferencia (Me gusta su sabor)} \\ & - 1,0608411 \text{ Ciudad (Guayaquil)} - 0,7415878 \text{ Ciudad (Machala)} \\ & - 0,8472632 \text{ Ciudad (Manta)} - 0,2906761 \text{ Ciudad (Quito)} \\ & + 8,000871 \text{ TOP(Fanta)} - 31,53007 \text{ TOP(Manzana)} \\ & + 3,475248 \text{ TOP(Otras marcas)} + 10,1272 \text{ TOP(Pepsi)} \\ & + 3,765659 \text{ TOP(Sprite)} \end{aligned}$$

El modelo nos indica que aquellos que sus razones de preferencia sean Por el gas, Contiene menos Colorante y Me gusta su sabor, que en la pregunta del TOM hayan mencionado Fanta, Otras marcas, Pepsi o Sprite, aumentan la posibilidad de escoger a la Fanta por encima de Coca Cola.

$$\begin{aligned} \eta_{manzana} = & -4,210045 - 1,5631235 \text{ Región (Sierra)} \\ & - 36,2942256 \text{ Razon de preferencia (Otras razones)} \\ & - 42,4460424 \text{ Razon de preferencia (Por el gas)} \\ & + 3,4625571 \text{ Razon de preferencia (Contiene menos colorante)} \\ & - 32,972789 \text{ Razon de preferencia (Tradición)} \\ & + 0,6174394 \text{ Razon de preferencia (Me gusta su sabor)} \\ & - 0,1927925 \text{ Ciudad (Guayaquil)} + 0,3583294 \text{ Ciudad (Machala)} \\ & - 2,8124588 \text{ Ciudad (Manta)} - 2,2037379 \text{ Ciudad (Quito)} \\ & - 5,938647 \text{ TOP(Fanta)} + 7,740326 \text{ TOP(Manzana)} \\ & + 3,693479 \text{ TOP(Otras marcas)} + 191,70086 \text{ TOP(Pepsi)} \\ & + 4,845764 \text{ TOP(Sprite)} \end{aligned}$$

El modelo nos indica que aquellos que sus razones de preferencia sean Contiene menos Colorante y Me gusta su sabor, procedentes de la ciudad de Machala y que en la pregunta del TOM hayan mencionado Manzana, Otras marcas, Pepsi o Sprite, aumentan la posibilidad de escoger a la Manzana por encima de Coca Cola.

$$\begin{aligned}
 \eta_{\text{Otras marcas}} = & -2,874925 - 1,2654912 \text{ Región (Sierra)} \\
 & + 1,0886947 \text{ Razon de preferencia(Otras razones)} \\
 & + 0,57948351 \text{ Razon de preferencia (Por el gas)} \\
 & + 2,7879397 \text{ Razon de preferencia (Contiene menos colorante)} \\
 & - 3,228725 \text{ Razon de preferencia (Tradición)} \\
 & + 0,4843526 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & - 0,1423481 \text{ Ciudad (Guayaquil)} - 0,2332967 \text{ Ciudad (Machala)} \\
 & - 1,2337895 \text{ Ciudad (Manta)} - 0,1371354 \text{ Ciudad (Quito)} \\
 & + 4,485634 \text{ TOP(Fanta)} + 3,466636 \text{ TOP(Manzana)} \\
 & + 5,892949 \text{ TOP(Otras marcas)} + 189,82467 \text{ TOP(Pepsi)} \\
 & + 2,716698 \text{ TOP(Sprite)}
 \end{aligned}$$

El modelo nos indica que aquellos que sus razones de preferencia sean Otras razones, Por el gas, Contiene menos Colorante y Me gusta su sabor y que en la pregunta del TOM hayan mencionado Fanta, Manzana, Otras marcas, Pepsi o Sprite, aumentan la posibilidad de escoger Otras marcas por encima de Coca Cola.

Los valores de las posibilidades más altas que determinan que prefieren a Sprite por encima de Coca Cola, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Coca Cola por encima de Sprite se encuentran en la Tabla CXL.

Tabla CXL						
Valores de las posibilidades del modelo de Telefonía Celular categoría Sprite						
(Intercept)	RegionCosta	TopPepsi	CiudadMachala	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	1,30703E+20
(Intercept)	RegionCosta	TopPepsi	CiudadGuayaquil	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	9,04705E+19
(Intercept)	RegionCosta	TopOtras Marcas	CiudadCuenca	Razones_PrefPor el gas	Valor de la posibilidad	1,136219791
(Intercept)	RegionCosta	TopOtras Marcas	CiudadCuenca	Razones_PrefPorque es refrescante	Valor de la posibilidad	1,056099073
(Intercept)	RegionCosta	TopManzana	CiudadManta	Razones_PrefPorque es una tradición familiar	Valor de la posibilidad	8,9791E-75
(Intercept)	RegionSierra	TopManzana	CiudadManta	Razones_PrefPorque es una tradición familiar	Valor de la posibilidad	3,8131E-75

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Fanta por encima de Coca Cola, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Coca Cola por encima de Fanta se encuentran en la Tabla CXXI.

Tabla CXXI						
Valores de las posibilidades del modelo de Bebidas Gaseosas categoría Fanta						
(Intercept)	RegionSierra	TopFanta	CiudadGuayaquil	Razones_PrefOtras razones	Valor de la posibilidad	5,378979112
(Intercept)	RegionCosta	TopPepsi	CiudadCuenca	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	5,9462E+19
(Intercept)	RegionSierra	TopSprite	CiudadMachala	Razones_PrefPor el gas	Valor de la posibilidad	1,102893113
(Intercept)	RegionCosta	TopOtras Marcas	CiudadManta	Razones_PrefPor el gas	Valor de la posibilidad	1,062838363
(Intercept)	RegionSierra	TopSprite	CiudadManta	Razones_PrefPor el gas	Valor de la posibilidad	0,992291296
(Intercept)	RegionCosta	TopCoca Cola	CiudadCuenca	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	0,962644206

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a la marca Manzana por encima de Coca Cola, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Coca Cola por encima de Manzana se encuentran en la Tabla CXXII.

Tabla CXXII						
Valores de las posibilidades del modelo de Bebidas Gaseosas categoría Manzana						
(Intercept)	RegionCosta	TopPepsi	CiudadMachala	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	1,21791E+83
(Intercept)	RegionCosta	TopPepsi	CiudadCuenca	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	8,5113E+82
(Intercept)	RegionCosta	TopOtras Marcas	CiudadCuenca	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,106136596
(Intercept)	RegionSierra	TopSprite	CiudadMachala	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,049552938
(Intercept)	RegionSierra	TopFanta	CiudadQuito	Razones_PrefPor el gas	Valor de la posibilidad	3,3302E-25
(Intercept)	RegionSierra	TopFanta	CiudadManta	Razones_PrefPor el gas	Valor de la posibilidad	1,81178E-25

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a la marca Pepsi por encima de Coca Cola, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que

determinan que prefieren Coca Cola por encima de Pepsi se encuentran en la Tabla CXLIII.

Tabla CXLIII						
Valores de las posibilidades del modelo de Bebidas Gaseosas categoría Pepsi						
(Intercept)	RegionCosta	TopPepsi	CiudadQuito	Razones_PrefPorque es refrescante	Valor de la posibilidad	7,20343E+83
(Intercept)	RegionCosta	TopPepsi	CiudadQuito	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	3,94879E+83
(Intercept)	RegionCosta	TopCoca Cola	CiudadQuito	Razones_PrefPorque es refrescante	Valor de la posibilidad	0,149129489
(Intercept)	RegionCosta	TopCoca Cola	CiudadQuito	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	0,081750169
(Intercept)	RegionCosta	TopManzana	CiudadManta	Razones_PrefPorque es una tradición familiar	Valor de la posibilidad	1,1897E-26
(Intercept)	RegionSierra	TopManzana	CiudadManta	Razones_PrefPorque es una tradición familiar	Valor de la posibilidad	2,67619E-27

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a la agrupación Otras Marcas por encima de Coca Cola, los valores que determinan que las posibilidades de escoger ambas empresas sean cercanas y las que determinan que prefieren Coca Cola por encima de la agrupación Otras Marcas se encuentran en la Tabla CXLIV.

Tabla CXLIV						
Valores de las posibilidades del modelo de Bebidas Gaseosas categoría Otras Marcas						
(Intercept)	RegionCosta	TopPepsi	CiudadCuenca	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	2,52365E+82
(Intercept)	RegionCosta	TopPepsi	CiudadQuito	Razones_PrefPorque contiene menos colorante	Valor de la posibilidad	2,20025E+82
(Intercept)	RegionSierra	TopFanta	CiudadMachala	Razones_PrefPorque es refrescante	Valor de la posibilidad	1,118424613
(Intercept)	RegionCosta	TopSprite	CiudadMachala	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,097273976
(Intercept)	RegionSierra	TopCoca Cola	CiudadMachala	Razones_PrefPorque es una tradición familiar	Valor de la posibilidad	0,000499233
(Intercept)	RegionSierra	TopCoca Cola	CiudadManta	Razones_PrefPorque es una tradición familiar	Valor de la posibilidad	0,000183567

Fuente y Elaboración: Mario Plaza

Los demás valores de posibilidades se encuentran en la sección Anexos.

En la tabla de clasificación que es la Tabla CXLV nos permite observar que de los 588 entrevistados que prefirieron Coca Cola, 585 fueron pronosticados como tal, es decir un 99,5%. En conclusión el modelo nos proporciona un 70,8% de correcta clasificación en forma general.

Tabla CXLV							
Tabla de Clasificación del modelo de Bebidas Gaseosas-Colas							
Observado	Pronosticado						
	Otras Marcas	Coca Cola	Fanta	Pepsi	Sprite	Manzana	Porcentaje Correcto
Otras Marcas	0	124	0	0	3	2	0.0%
Coca Cola	0	585	0	0	2	1	99.5%
Fanta	0	42	0	0	11	1	0.0%
Pepsi	0	7	0	0	30	0	0.0%
Sprite	0	20	0	0	53	6	67.1%
Manzana	0	11	0	0	8	12	38.7%
Porcentaje global	0.0%	85.9%	0.0%	0.0%	11.7%	2.4%	70.8%

Fuente y Elaboración: Mario Plaza

Categoría: Bebidas Alcohólicas (en general)

En base a los resultados obtenidos a través de las Tablas de contingencia, el modelo dentro de ésta categoría es el siguiente:

Preferencia Alcohol ~ Ciudad + Región + Actividad + Top Alcohol + Razones de Preferencia + 7mo Atributo + 8avo Atributo + 9no Atributo

Cuyo valor AIC es 1119,31.

En consecuencia consideramos un segundo modelo cuyo valor AIC sea inferior para poder predecir de forma más efectiva la variable dependiente, quedando de la siguiente forma:

Preferencia Alcohol ~ Ciudad + Región + Top Alcohol + Razones de Preferencia

Cuyo valor AIC es 530,09.

En base a este modelo procedemos a obtener los coeficientes a través del software estadístico R-Studio. Debido a que el método de regresión logística multinomial toma una categoría como base vamos en este caso a tener cinco

modelos. Sabemos que dentro de Bebidas Gaseosas-Colas contamos con las marcas: Pilsener, Zhumir, Club, Brahma, Vodka y Otras Marcas. En este caso la categoría que el software consideró como categoría con la cual comparar es Brahma, por lo tanto tenemos cinco modelos en base a las otras categorías, las cuales miden la posibilidad de preferir la empresa X por encima de la empresa Y. Los coeficientes de los modelos podemos apreciarlos en las Tablas CXLVI, CXLVII y CXLVIII.

Tabla CXLVI						
Valores de los Coeficientes del modelo de Bebidas Alcohólicas						
Categorías	(Intercept)	Region Sierra	Razones_Pref Otras razones	Razones_Pref Por costumbre/ Por tradición	Razones_Pref Porque es más barata/Tiene un precio más accesible	Razones_Pref Porque es un producto de calidad
Pilsener	48,16254	-29,6671	-31,80893	-38,84961	-315,853711	-53,00115
Club	80,54315	-894,9355	-740,91987	-792,00559	-1030,83012	-400,1661
Otras marcas	-442,88159	-152,1732	287,10411	279,05272	3,135241	267,78970
Vodka	-428,96411	-263,6852	39,47387	89,52026	-25,363526	235,56954
Zhumir	-191,73504	-18,087	490,62837	484,20408	208,956124	469,85615

Fuente y Elaboración: Mario Plaza

Tabla CXLVII					
Valores de los Coeficientes del modelo de Bebidas Alcohólicas					
Categorías	Razones_Pref Porque me gusta su sabor	Ciudad Guayaquil	Ciudad Machala	Ciudad Manta	Ciudad Quito
Pilsener	-315,605774	-29,24092	-27,70228	134,77284	0,8202115
Club	-664,233997	502,71598	154,79872	317,96396	1050,00100299
Otras Marcas	3,467411	-152,13044	-150,73306	12,15507	1,1171626
Vodka	-27,79148	57,15862	57,9262	-280,36376	321,8041842
Zhumir	208,320707	-18,93624	-17,34281	-137,36898	-0,8756344

Fuente y Elaboración: Mario Plaza

Tabla CXLVIII					
Valores de los Coeficientes del modelo de Bebidas Alcohólicas					
Categorías	TopClub	TopOtras Marcas	TopPilsener	TopVodka	TopZhumir
Pilsener	-221,67157	771,3269	299,770835	-128,10293	739,5474
Club	650,64532	542,0033	82,698876	-36,97572	364,1328
Otras Marcas	810,67832	1068,3078	590,89797	934,75569	1035,2261
Vodka	58,18910	874,3289	240,5143	744,59022	842,3237
Zhumir	-37,95536	475,6297	1,091685	346,36324	449,1645

Fuente y Elaboración: Mario Plaza

En base a los valores de los coeficientes de las tablas anteriores, los modelos son los siguientes:

$$\begin{aligned} \eta_{Pilsener} = & 48,16254 - 29,6671 \text{ Región (Sierra)} \\ & - 31,80893 \text{ Razon de preferencia (Otras razones)} \\ & - 38,84961 \text{ Razon de preferencia (Por tradición)} \\ & - 315,853711 \text{ Razon de preferencia (Es más barata)} \\ & - 53,00115 \text{ Razon de preferencia (Calidad)} \\ & - 315,605774 \text{ Razon de preferencia (Me gusta su sabor)} \\ & - 29.24092 \text{ Ciudad (Guayaquil)} - 27,70228 \text{ Ciudad (Machala)} \\ & + 134,77284 \text{ Ciudad (Manta)} + 0,8202115 \text{ Ciudad (Quito)} \\ & - 221,67157 \text{ TOP(Club)} + 771,3269 \text{ TOP(Otras Marcas)} \\ & + 299,770835 \text{ TOP(Pilsener)} - 128,10293 \text{ TOP(Vodka)} \\ & + 739,5474 \text{ TOP(Zhumir)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que provengan de las ciudades de Manta o Quito y que en la pregunta del TOM hayan mencionado entre las opciones Otras Marcas, Pilsener, Vodka o Zhumir, aumentan la posibilidad de que prefieran a Pilsener por encima de Brahma.

$$\begin{aligned}
 \eta_{Club} = & 80,54315 - 894,9355 \text{ Región (Sierra)} \\
 & - 740,91987 \text{ Razon de preferencia (Otras razones)} \\
 & - 792,00559 \text{ Razon de preferencia (Por tradición)} \\
 & - 1030,83012 \text{ Razon de preferencia (Es más barata)} \\
 & - 400,1661 \text{ Razon de preferencia (Calidad)} \\
 & - 664,233997 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & + 502,71598 \text{ Ciudad (Guayaquil)} + 154,79872 \text{ Ciudad (Machala)} \\
 & + 317,96396 \text{ Ciudad (Manta)} + 1050,00100299 \text{ Ciudad (Quito)} \\
 & + 650,64532 \text{ TOP(Club)} + 542,0033 \text{ TOP(Otras Marcas)} \\
 & + 82,698876 \text{ TOP(Pilsener)} - 36,97572 \text{ TOP(Vodka)} \\
 & + 364,1328 \text{ TOP(Zhumir)}
 \end{aligned}$$

El modelo nos indica que aquellos encuestados que provengan de las ciudades de Guayaquil, Machala, Manta o Quito y que en la pregunta del TOM hayan mencionado entre las opciones Club, Otras Marcas, Pilsener, Vodka o Zhumir, aumentan la posibilidad de que prefieran a Club por encima de Brahma.

$$\begin{aligned}
 \eta_{Otras Marcas} = & -442,88159 - 152,1732 \text{ Región (Sierra)} \\
 & + 287,10411 \text{ Razon de preferencia (Otras razones)} \\
 & + 279,05272 \text{ Razon de preferencia (Por tradición)} \\
 & + 3,135241 \text{ Razon de preferencia (Es más barata)} \\
 & + 267,78970 \text{ Razon de preferencia (Calidad)} \\
 & - 664,233997 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & - 152,13044 \text{ Ciudad (Guayaquil)} - 150,73306 \text{ Ciudad (Machala)} \\
 & + 12,15507 \text{ Ciudad (Manta)} + 1,1171626 \text{ Ciudad (Quito)} \\
 & + 810,67832 \text{ TOP(Club)} + 1068,3078 \text{ TOP(Otras Marcas)} \\
 & + 590,89797 \text{ TOP(Pilsener)} + 934,75569 \text{ TOP(Vodka)} \\
 & + 1035,2261 \text{ TOP(Zhumir)}
 \end{aligned}$$

El modelo nos indica que aquellos encuestados que hayan mencionado cualquier razón de preferencia menos Me gusta su sabor, provengan de las

ciudades de Manta o Quito y que en la pregunta del TOM hayan mencionado cualquier marca de bebida alcohólica, aumentan la posibilidad de que prefieran a Otras Marcas por encima de Brahma.

$$\begin{aligned} \eta_{Vodka} = & -428,96411 - 263,6852 \text{ Región (Sierra)} \\ & + 39,47387 \text{ Razon de preferencia (Otras razones)} \\ & + 89,52026 \text{ Razon de preferencia (Por tradición)} \\ & - 25,363526 \text{ Razon de preferencia (Es más barata)} \\ & + 235,56954 \text{ Razon de preferencia (Calidad)} \\ & - 27,79148 \text{ Razon de preferencia (Me gusta su sabor)} \\ & + 57,15862 \text{ Ciudad (Guayaquil)} - 57,9262 \text{ Ciudad (Machala)} \\ & - 280,36376 \text{ Ciudad (Manta)} + 321,8041842 \text{ Ciudad (Quito)} \\ & + 58,18910 \text{ TOP(Club)} + 874,3289 \text{ TOP(Otras Marcas)} \\ & + 240,5143 \text{ TOP(Pilsener)} + 744,59022 \text{ TOP(Vodka)} \\ & + 842,3237 \text{ TOP(Zhumir)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que hayan mencionado cualquier razón de preferencia a excepción de Es más barata o Me gusta su sabor, que provengan de las ciudades de Guayaquil o Quito y que en la pregunta del TOM hayan mencionado cualquier marca de bebida alcohólica, aumentan la posibilidad de que prefieran Vodka por encima de Brahma.

$$\begin{aligned}
 \eta_{Zhumir} = & -191,73504 - 18,087 \text{ Región (Sierra)} \\
 & + 490,62837 \text{ Razon de preferencia (Otras razones)} \\
 & + 484,20408 \text{ Razon de preferencia (Por tradición)} \\
 & + 208,956124 \text{ Razon de preferencia (Es más barata)} \\
 & + 469,85615 \text{ Razon de preferencia (Calidad)} \\
 & + 208,320707 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & - 18,93624 \text{ Ciudad (Guayaquil)} - 17,34281 \text{ Ciudad (Machala)} \\
 & - 137,36898 \text{ Ciudad (Manta)} - 0,8756344 \text{ Ciudad (Quito)} \\
 & - 37,95536 \text{ TOP(Club)} + 475,6297 \text{ TOP(Otras Marcas)} \\
 & + 1,091685 \text{ TOP(Pilsener)} + 346,36324 \text{ TOP(Vodka)} \\
 & + 449,1645 \text{ TOP(Zhumir)}
 \end{aligned}$$

El modelo nos indica que aquellos encuestados que hayan mencionado cualquier razón de preferencia y que en la pregunta del TOM hayan mencionado cualquier marca de bebida alcohólica menos Club, aumentan la posibilidad de que prefieran Zhumir por encima de Brahma.

Los valores de las posibilidades más altas que determinan que prefieren a Pilsener por encima de Brahma, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Brahma por encima de Pilsener se encuentran en la Tabla CXLIX.

Tabla CXLIX						
Valores de las posibilidades del modelo de Bebidas Alcohólicas categoría Pilsener						
(Intercept)	RegionSierra	TopOtras Marcas	CiudadMachala	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	9,2561E+307
(Intercept)	RegionSierra	TopOtras Marcas	CiudadGuayaquil	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	1,987E+307
(Intercept)	RegionSierra	TopPilsener	CiudadCuenca	Razones_PrefPorque es más barata/Tiene un precio más accesible	Valor de la posibilidad	11,16254525
(Intercept)	RegionCosta	TopVodka	CiudadManta	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	6,241995981
(Intercept)	RegionSierra	TopClub	CiudadGuayaquil	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	9,9246E-239
(Intercept)	RegionSierra	TopClub	CiudadGuayaquil	Razones_PrefPorque es más barata/Tiene un precio más accesible	Valor de la posibilidad	7,7452E-239

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Club por encima de Brahma, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Brahma por encima de Club se encuentran en la Tabla CL.

Tabla CL						
Valores de las posibilidades del modelo de Bebidas Alcohólicas categoría Club						
(Intercept)	RegionCosta	TopZhumir	CiudadQuito	Razones_PrefPor costumbre/ Por tradición	Valor de la posibilidad	1,4665E+305
(Intercept)	RegionCosta	TopVodka	CiudadQuito	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	1,383E+301
(Intercept)	RegionCosta	TopPilsener	CiudadGuayaquil	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	5,606961776
(Intercept)	RegionCosta	TopBrahma	CiudadManta	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	0,190331117
(Intercept)	RegionSierra	TopVodka	CiudadMachala	Razones_PrefEs refrescante	Valor de la posibilidad	3,0463E-303
(Intercept)	RegionCosta	TopVodka	CiudadCuenca	Razones_PrefOtras razones	Valor de la posibilidad	1,3921E-303

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Otras Marcas por encima de Brahma, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Brahma por encima de Otras Marcas se encuentran en la Tabla CLI.

Tabla CLI						
Valores de las posibilidades del modelo de Bebidas Alcohólicas categoría Otras Marcas						
(Intercept)	RegionCosta	TopZhumir	CiudadMachala	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	1,2275E+308
(Intercept)	RegionSierra	TopZhumir	CiudadQuito	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	8,8866E+307
(Intercept)	RegionCosta	TopPilsener	CiudadMachala	Razones_PrefPorque es más barata/Tiene un precio más accesible	Valor de la posibilidad	1,519773028
(Intercept)	RegionSierra	TopPilsener	CiudadQuito	Razones_PrefPorque es más barata/Tiene un precio más accesible	Valor de la posibilidad	1,100300803
(Intercept)	RegionCosta	TopBrahma	CiudadGuayaquil	Razones_PrefEs refrescante	Valor de la posibilidad	3,8865E-259
(Intercept)	RegionSierra	TopBrahma	CiudadCuenca	Razones_PrefEs refrescante	Valor de la posibilidad	3,7238E-259

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Vodka por encima de Brahma, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que

determinan que prefieren Brahma por encima de Vodka se encuentran en la Tabla CLII.

Tabla CLII						
Valores de las posibilidades del modelo de Bebidas Alcohólicas categoría Vodka						
(Intercept)	RegionCosta	TopZhumir	CiudadQuito	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,6139E+307
(Intercept)	RegionSierra	TopZhumir	CiudadQuito	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	1,1671E+307
(Intercept)	RegionSierra	TopVodka	CiudadManta	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	1269,895636
(Intercept)	RegionSierra	TopOtras Marcas	CiudadManta	Razones_PrefPor costumbre/ Por tradición	Valor de la posibilidad	0,000104753
(Intercept)	RegionSierra	TopBrahma	CiudadCuenca	Razones_PrefEs refrescante	Valor de la posibilidad	1,5354E+301
(Intercept)	RegionSierra	TopPilsener	CiudadManta	Razones_PrefOtras razones	Valor de la posibilidad	1,0547E+301

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Zhumir por encima de Brahma, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Brahma por encima de Zhumir se encuentran en la Tabla CLIII.

Tabla CLIII						
Valores de las posibilidades del modelo de Bebidas Alcohólicas categoría Zhumir						
(Intercept)	RegionSierra	TopZhumir	CiudadCuenca	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	1,0024E+308
(Intercept)	RegionCosta	TopZhumir	CiudadGuayaquil	Razones_PrefPorque es un producto de calidad	Valor de la posibilidad	4,2877E+307
(Intercept)	RegionCosta	TopPilsener	CiudadMachala	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,397300275
(Intercept)	RegionSierra	TopPilsener	CiudadCuenca	Razones_PrefPorque es más barata/Tiene un precio más accesible	Valor de la posibilidad	1,253286123
(Intercept)	RegionCosta	TopClub	CiudadManta	Razones_PrefEs refrescante	Valor de la posibilidad	3,8738E+160
(Intercept)	RegionSierra	TopClub	CiudadManta	Razones_PrefEs refrescante	Valor de la posibilidad	5,4082E+168

Fuente y Elaboración: Mario Plaza

Los demás valores de posibilidades se encuentran en la sección Anexos.

En la tabla de clasificación que es la Tabla CLIV nos permite observar que de los 236 entrevistados que prefirieron Pilsener, 217 fueron pronosticados como tal, es decir un 91,9%. En conclusión el modelo nos proporciona un 71,3% de correcta clasificación en forma general.

Tabla CLIV							
Tabla de Clasificación del modelo de Bebidas Alcohólicas							
Observado	Pronosticado						
	Otras Marcas	Pilsener	Brahma	Club	Zhumir	Vodka	Porcentaje Correcto
Otras Marcas	31	52	0	0	3	2	35.2%
Pilsener	18	217	0	0	1	0	91.9%
Brahma	0	10	0	0	1	0	0.0%
Club	0	25	0	9	2	0	25.0%
Zhumir	0	5	0	2	57	1	87.7%
Vodka	3	2	0	0	2	6	46.2%
Porcentaje global	11.6%	69.3%	0.0%	2.4%	14.7%	2.0%	71.3%

Fuente y Elaboración: Mario Plaza

Categoría: Atún

En base a los resultados obtenidos a través de las Tablas de contingencia, el modelo dentro de ésta categoría es el siguiente:

Preferencia Atún ~ Ciudad + Región + Top Atún + Razones de Preferencia + 4to Atributo + 5to Atributo + 6to Atributo + 7mo Atributo + 8avo Atributo + 9no Atributo

Cuyo valor AIC es 1115,48.

En consecuencia consideramos un segundo modelo cuyo valor AIC sea inferior para poder predecir de forma más efectiva la variable dependiente, quedando el siguiente modelo:

Preferencia Atún ~ Ciudad + Región + Top Atún + Razones de Preferencia

Cuyo valor AIC es 814,07.

Sin embargo a través del valor p , se determina que se proceda a retirar una variable más, quedando el modelo de la siguiente forma:

Preferencia Atún ~ Región + Top Atún + Razones de Preferencia

Cuyo valor AIC es 814,03.

En base a este modelo procedemos a obtener los coeficientes a través del software estadístico R-Studio. Debido a que el método de regresión logística multinomial toma una categoría como base vamos en este caso a tener cinco modelos. Sabemos que dentro de Atún contamos con las marcas: Real, Van Camps, Starkist, Isabel, Campos y Otras Marcas. En este caso la categoría que el software consideró como categoría con la cual comparar es Campos, por lo tanto tenemos cinco modelos en base a las otras categorías, las cuales miden la posibilidad de preferir la empresa X por encima de la empresa Y. Los coeficientes de los modelos podemos apreciarlos en las Tablas CLV y CLVI.

Tabla CLV						
Valores de los Coeficientes del modelo de Atún						
Categorías	(Intercept)	RegionSierra	Razones_Pref Porque es económico	Razones_Pref Porque es el de mejor calidad	Razones_Pref Porque es una costumbre	Razones_Pref Porque me gusta su sabor
Isabel	-22,97477	-37,18203	-16,66923	-37,691535	-18,9988	-71,74802
Otras Marcas	-32,03812	-38,57858	34,94448	-6,909111	33,24654	-20,18157
Real	-47,62693	-36,50113	-17,93732	-38,812021	-20,44224	-72,26667
Starkist	-87,20396	-49,76743	25,19155	-12,569007	31,47906	48,61498
Van Camps	-28,8326	-36,70551	-19,83774	-37,567681	-19,14806	-71,66166

Fuente y Elaboración: Mario Plaza

Tabla CLVI					
Valores de los Coeficientes del modelo de Atún					
Categorías	TopIsabel	TopOtras Marcas	TopReal	TopStarkist	TopVancamps
Isabel	154,42742	204,417	158,37931	-23,291529	143,37692
Otras Marcas	22,62465	164,1073	115,99829	-8,111549	86,48378
Real	177,55983	230,7505	187,56286	125,138354	169,86701
Starkist	94,63725	-48,8795	56,04019	140,245645	86,60399
Van Camps	156,75516	23,1805	165,60113	106,686903	153,62137

Fuente y Elaboración: Mario Plaza

En base a los valores de los coeficientes de las tablas anteriores, los modelos son los siguientes:

$$\begin{aligned}
 \eta_{Isabel} = & -22,97477 - 37,18203 \text{ Región (Sierra)} \\
 & - 16,66923 \text{ Razon de preferencia (Es más barato)} \\
 & - 37,691535 \text{ Razon de preferencia (Calidad)} \\
 & - 18,9988 \text{ Razon de preferencia (Tradicion)} \\
 & - 71,74802 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & - 154,42742 \text{ TOP(Isabel)} - 204,417 \text{ TOP(Otras Marcas)} \\
 & + 158,37931 \text{ TOP(Real)} - 23,291529 \text{ TOP(Starkist)} \\
 & + 143,37692 \text{ TOP(Vancamps)}
 \end{aligned}$$

El modelo nos indica que aquellos encuestados que en la pregunta del TOM hayan mencionado las marcas Real o Van Camps, aumentan la posibilidad de que prefieran a Isabel por encima de Campos.

$$\begin{aligned}
\eta_{Otras\ Marcas} = & -32,03812 - 38,57858 \text{ Región (Sierra)} \\
& + 34,94448 \text{ Razon de preferencia (Es más barato)} \\
& - 6,909111 \text{ Razon de preferencia (Calidad)} \\
& + 33,24654 \text{ Razon de preferencia (Tradicion)} \\
& - 20,18157 \text{ Razon de preferencia (Me gusta su sabor)} \\
& + 22,62465 \text{ TOP(Isabel)} + 164,1073 \text{ TOP(Otras Marcas)} \\
& + 115,99829 \text{ TOP(Real)} - 8,111549 \text{ (Starkist)} \\
& + 86,48378 \text{ TOP(Vancamps)}
\end{aligned}$$

El modelo nos indica que aquellos encuestados que en las razones de preferencia hayan mencionado Es más barato o Tradición, que en la pregunta del TOM hayan mencionado las marcas Isabel, Real, Van Camps u Otras Marcas aumentan la posibilidad de que prefieran Otras Marcas en vez de Campos.

$$\begin{aligned}
\eta_{Real} = & -47,62693 - 36,50113 \text{ Región (Sierra)} \\
& - 17,93732 \text{ Razon de preferencia (Es más barato)} \\
& - 38,812021 \text{ Razon de preferencia (Calidad)} \\
& - 20,44224 \text{ Razon de preferencia (Tradicion)} \\
& - 72,26667 \text{ Razon de preferencia (Me gusta su sabor)} \\
& + 177,55983 \text{ TOP(Isabel)} + 203,7505 \text{ TOP(Otras Marcas)} \\
& + 187,56286 \text{ TOP(Real)} + 125,138354 \text{ (Starkist)} \\
& + 169,86701 \text{ TOP(Vancamps)}
\end{aligned}$$

El modelo nos indica que aquellos encuestados que en la pregunta del TOM hayan mencionado cualquier marca de atún aumentan la posibilidad de que prefieran Real en vez de Campos.

$$\begin{aligned}
 \eta_{Starkist} = & -87,20396 - 49,76743 \text{ Región (Sierra)} \\
 & + 25,19155 \text{ Razon de preferencia (Es más barato)} \\
 & - 12,569007 \text{ Razon de preferencia (Calidad)} \\
 & + 31,47906 \text{ Razon de preferencia (Tradicion)} \\
 & + 48,61498 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & + 94,63725 \text{ TOP(Isabel)} - 48,8795 \text{ TOP(Otras Marcas)} \\
 & + 56,04019 \text{ TOP(Real)} + 140,245645 \text{ (Starkist)} \\
 & + 86,60399 \text{ TOP(Vancamps)}
 \end{aligned}$$

El modelo nos indica que aquellos encuestados que en las razones de preferencia hayan mencionado cualquiera a excepción de Calidad, que en la pregunta del TOM hayan mencionado cualquier marca de atún aumentan la posibilidad de que prefieran Starkist en vez de Campos.

$$\begin{aligned}
 \eta_{VanCamps} = & -28,8326 - 36,70551 \text{ Región (Sierra)} \\
 & - 19,83774 \text{ Razon de preferencia (Es más barato)} \\
 & - 37,567681 \text{ Razon de preferencia (Calidad)} \\
 & - 19,14806 \text{ Razon de preferencia (Tradicion)} \\
 & - 71,66166 \text{ Razon de preferencia (Me gusta su sabor)} \\
 & + 156,75516 \text{ TOP(Isabel)} + 231805 \text{ TOP(Otras Marcas)} \\
 & + 165,60113 \text{ TOP(Real)} + 106,686903 \text{ (Starkist)} \\
 & + 153,62137 \text{ TOP(Vancamps)}
 \end{aligned}$$

El modelo nos indica que aquellos encuestados que en la pregunta del TOM hayan mencionado cualquier marca de atún aumentan la posibilidad de que prefieran Van Camps en vez de Campos.

Los valores de las posibilidades más altas que determinan que prefieren a Van Camps por encima de Campos, los valores que determinan que las

posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Campos por encima de Van Camps se encuentran en la Tabla CLVII.

Tabla CLVII					
Valores de las posibilidades del modelo de Atún categoría Van Camps					
(Intercept)	RegionCosta	TopReal	Razones_PrefOtras Razones	Valor de la posibilidad	2,4993E+59
(Intercept)	RegionCosta	TopIsabel	Razones_PrefOtras Razones	Valor de la posibilidad	3,59801E+55
(Intercept)	RegionSierra	TopStarkist	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	35,91345441
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefOtras Razones	Valor de la posibilidad	0,003510138
(Intercept)	RegionSierra	TopOtras Marcas	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	3,03431E-50
(Intercept)	RegionSierra	TopCampos	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	2,59954E-60

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Isabel por encima de Campos, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Campos por encima de Isabel se encuentran en la Tabla CLVIII.

Tabla CLVIII					
Valores de las posibilidades del modelo de Atún categoría Isabel					
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefOtras Razones	Valor de la posibilidad	6,30027E+78
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefPorque es económico	Valor de la posibilidad	3,63082E+71
(Intercept)	RegionSierra	TopIsabel	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	6045620726
(Intercept)	RegionSierra	TopVan Camps	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	95999,66684
(Intercept)	RegionSierra	TopCampos	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	5,18161E-58
(Intercept)	RegionSierra	TopStarkist	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	3,97267E-68

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Otras Marcas por encima de Campos, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que

determinan que prefieren Campos por encima de Otras Marcas se encuentran en la Tabla CLVIX.

Tabla CLIX					
Valores de las posibilidades del modelo de Atún categoría Otras Marcas					
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefPorque es económico	Valor de la posibilidad	3,4128E+72
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefPorque es una costumbre familiar	Valor de la posibilidad	6,24748E+71
(Intercept)	RegionCosta	TopCampos	Razones_PrefPorque es económico	Valor de la posibilidad	18,29010128
(Intercept)	RegionCosta	TopCampos	Razones_PrefPorque es una costumbre familiar	Valor de la posibilidad	3,34819033
(Intercept)	RegionSierra	TopCampos	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	3,68818E-40
(Intercept)	RegionSierra	TopStarkist	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,10665E-43

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Real por encima de Campos, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Campos por encima de Real se encuentran en la Tabla CLX.

Tabla CLX					
Valores de las posibilidades del modelo de Atún categoría Real					
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefOtras Razones	Valor de la posibilidad	3,38498E+79
(Intercept)	RegionCosta	TopOtras Marcas	Razones_PrefPorque es económico	Valor de la posibilidad	5,48879E+71
(Intercept)	RegionCosta	TopStarkist	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	189,5691755
(Intercept)	RegionSierra	TopStarkist	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	9,009440752
(Intercept)	RegionSierra	TopCampos	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	4,05323E-54
(Intercept)	RegionSierra	TopCampos	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,19848E-68

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Starkist por encima de Campos, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Campos por encima de Starkist se encuentran en la Tabla CLXI.

Tabla CLXI					
Valores de las posibilidades del modelo de Atún categoría Starkist					
(Intercept)	RegionCosta	TopStarkist	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,40906E+44
(Intercept)	RegionCosta	TopStarkist	Razones_PrefPorque es una costumbre familiar	Valor de la posibilidad	5,09206E+36
(Intercept)	RegionCosta	TopReal	Razones_PrefPorque es una costumbre familiar	Valor de la posibilidad	1,370656744
(Intercept)	RegionCosta	TopVan Camps	Razones_PrefOtras Razones	Valor de la posibilidad	0,548828101
(Intercept)	RegionSierra	TopOtras Marcas	Razones_PrefOtras Razones	Valor de la posibilidad	1,9319E-81
(Intercept)	RegionSierra	TopOtras Marcas	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	6,71945E-87

Fuente y Elaboración: Mario Plaza

Los demás valores de posibilidades se encuentran en la sección Anexos.

En la tabla de clasificación que es la Tabla CLXII nos permite observar que de los 357 entrevistados que prefirieron Van Camps, 314 fueron pronosticados como tal, es decir un 88%; y de los 565 encuestados que prefirieron Real, 538 fueron pronosticados como tal, es decir un 95,2%. En conclusión el modelo nos proporciona un 90,4% de correcta clasificación en forma general.

Tabla CLXII							
Tabla de Clasificación del modelo de Atún							
Observado	Pronosticado						Porcentaje Correcto
	Otras Marcas	Vancamps	Real	Isabel	Starkist	Campos	
Otras Marcas	0	5	7	0	0	0	0.0%
Vancamps	0	314	39	4	0	0	88.0%
Real	0	17	538	10	0	0	95.2%
Isabel	0	5	9	72	0	0	83.7%
Starkist	0	1	0	2	0	0	0.0%
Campos	0	0	0	0	0	3	100.0%
Porcentaje global	0.0%	33.3%	57.8%	8.6%	0.0%	.3%	90.4%

Fuente y Elaboración: Mario Plaza

Categoría: Chocolate

En base a los resultados obtenidos a través de las Tablas de contingencia, el modelo dentro de ésta categoría es el siguiente:

Preferencia Chocolate ~ Ciudad + Región + Edad + Top Chocolate + Razones de Preferencia + 1er Atributo + 7mo Atributo + 8avo Atributo + 9no Atributo

Cuyo valor AIC es 794,12.

En consecuencia consideramos un segundo modelo cuyo valor AIC sea inferior para poder predecir de forma más efectiva la variable dependiente, quedando el siguiente modelo:

Preferencia Chocolate ~ Región + Edad + Top Chocolate

Cuyo valor AIC es 639,73.

En base a este modelo procedemos a obtener los coeficientes a través del software estadístico R-Studio. Debido a que el método de regresión logística multinomial toma una categoría como base vamos en este caso a tener cinco modelos. Sabemos que dentro de Chocolate contamos con las marcas: Nestlé, Manicho, Universal, Galak, Superior y Otras Marcas. En este caso la categoría que el software consideró como categoría con la cual comparar es Galak, por lo tanto tenemos cinco modelos en base a las otras categorías, las cuales miden la posibilidad de preferir la empresa X por encima de la

empresa Y. Los coeficientes de los modelos podemos apreciarlos en la Tabla CLXIII.

Tabla CLXIII								
Valores de los Coeficientes del modelo de Chocolate								
Categorías	(Intercept)	RegionSierra	Edad	TopManicho	TopNestle	TopOtras Marcas	TopSuperior	TopUniversal
Manicho	-2,666319	0,3514999	-0,023778779	6,977224	2,404138	3,226986	-7,513712	2,707786
Nestlé	-4,113816	-0,7030516	0,021919744	2,964742	6,568966	3,131867	-3,10448	3,361301
Otras Marcas	-2,594144	-0,5036932	-0,007402006	2,791897	3,26194	5,72443	-3,681777	3,523933
Superior	-5,528732	-15,1263299	0,076029164	-19,864757	1,988801	-16,170772	51,421641	-10,480146
Universal	-3,747069	-16,7049808	0,028609656	1,454083	3,032887	-15,261569	49,450233	5,422767

Fuente y Elaboración: Mario Plaza

En base a los valores de los coeficientes de la tabla anterior, los modelos son los siguientes:

$$\eta_{Manicho} = -2,666319 + 0,3514999 \text{ Región (Sierra)} - 0,0237778779 \text{ Edad} \\ + 6,977224 \text{ TOP(Manicho)} + 2,404138 \text{ TOP(Nestle)} \\ + 3,226986 \text{ TOP(Otras marcas)} - 7,513712 \text{ TOP(Superior)} \\ + 2,707786 \text{ TOP(Universal)}$$

El modelo nos indica que aquellos encuestados provenientes de la Sierra, que en la pregunta del TOM hayan mencionado cualquier marca de chocolate, aumentan la posibilidad de que prefieran a Manicho por encima de Galak.

$$\begin{aligned} \eta_{Nestle} = & -4,113816 - 0,7030516 \text{ Región (Sierra)} + 0,021919744 \text{ Edad} \\ & + 2,964742 \text{ TOP(Manicho)} + 6,568966 \text{ TOP(Nestle)} \\ & + 3,131867 \text{ TOP(Otras marcas)} - 3,10448 \text{ TOP(Superior)} \\ & + 3,361301 \text{ TOP(Universal)} \end{aligned}$$

El modelo nos indica que aquellos encuestados mayores de edad, que en la pregunta del TOM hayan mencionado cualquier marca de chocolate, aumentan la posibilidad de que prefieran a Nestlé por encima de Galak.

$$\begin{aligned} \eta_{Otras Marcas} = & -2,594144 - 0,5036932 \text{ Región (Sierra)} - 0,007402006 \text{ Edad} \\ & + 2,791897 \text{ TOP(Manicho)} + 3,26194 \text{ TOP(Nestle)} \\ & + 5,72443 \text{ TOP(Otras marcas)} - 3,681777 \text{ TOP(Superior)} \\ & + 3,523933 \text{ TOP(Universal)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que en la pregunta del TOM hayan mencionado cualquier marca de chocolate a excepción de Superior, aumentan la posibilidad de que prefieran Otras Marcas por encima de Galak.

$$\begin{aligned} \eta_{Superior} = & -5,528732 - 15,1263299 \text{ Región (Sierra)} + 0,076029164 \text{ Edad} \\ & - 19,864757 \text{ TOP(Manicho)} + 1,988801 \text{ TOP(Nestle)} \\ & - 16,170772 \text{ TOP(Otras marcas)} + 51,421641 \text{ TOP(Superior)} \\ & - 10,480146 \text{ TOP(Universal)} \end{aligned}$$

El modelo nos indica que aquellos encuestados mayores de edad, que en la pregunta del TOM hayan mencionado las marcas Nestlé o Superior, aumentan la posibilidad de que prefieran a Superior por encima de Galak.

$$\begin{aligned} \eta_{La\ Universal} = & -3,747069 - 16,7049808\ Region\ (Sierra) + 0,028609656\ Edad \\ & + 1,454083\ TOP(Manicho) + 3,032887\ TOP(Nestle) \\ & - 15,261569\ TOP(Otras\ marcas) + 49,450233\ TOP(Superior) \\ & + 5,422767\ TOP(Universal) \end{aligned}$$

El modelo nos indica que aquellos encuestados mayores de edad, que en la pregunta del TOM hayan mencionado las marcas Manicho, Nestlé, Superior o La Universal, aumentan la posibilidad de que prefieran a La Universal por encima de Galak.

Los valores de las posibilidades más altas que determinan que prefieren a Manicho por encima de Galak, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Galak por encima de Manicho se encuentran en la Tabla CLXIV.

Tabla CLXIV					
Valores de las posibilidades del modelo de Chocolate categoría Manicho					
(Intercept)	Edad	RegionSierra	TopManicho	Valor de la posibilidad	103,4021879
(Intercept)	Edad	RegionCosta	TopManicho	Valor de la posibilidad	72,75708001
(Intercept)	Edad	RegionSierra	TopNestle	Valor de la posibilidad	1,067735575
(Intercept)	Edad	RegionCosta	TopUniversal	Valor de la posibilidad	1,017845584
(Intercept)	Edad	RegionSierra	TopSuperior	Valor de la posibilidad	5,26255E-05
(Intercept)	Edad	RegionCosta	TopSuperior	Valor de la posibilidad	3,7029E-05

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Nestlé por encima de Galak, los valores que determinan que las posibilidades

de escoger ambas empresas son cercanas y las que determinan que prefieren Galak por encima de Nestlé se encuentran en la Tabla CLXV.

Tabla CLXV					
Valores de las posibilidades del modelo de Chocolate categoría Nestle					
(Intercept)	Edad	RegionCosta	TopNestle	Valor de la posibilidad	11,90632467
(Intercept)	Edad	RegionSierra	TopNestle	Valor de la posibilidad	5,894490753
(Intercept)	Edad	RegionCosta	TopUniversal	Valor de la posibilidad	0,481622216
(Intercept)	Edad	RegionCosta	TopOtras Marcas	Valor de la posibilidad	0,382881684
(Intercept)	Edad	RegionCosta	TopSuperior	Valor de la posibilidad	0,000749296
(Intercept)	Edad	RegionSierra	TopSuperior	Valor de la posibilidad	0,000370956

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren Otras Marcas por encima de Galak, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Galak por encima de Otras Marcas se encuentran en la Tabla CLXVI.

Tabla CLXVI					
Valores de las posibilidades del modelo de Chocolate categoría Otras Marcas					
(Intercept)	Edad	RegionCosta	TopOtras Marcas	Valor de la posibilidad	22,71178594
(Intercept)	Edad	RegionSierra	TopOtras Marcas	Valor de la posibilidad	13,72461305
(Intercept)	Edad	RegionCosta	TopManicho	Valor de la posibilidad	1,209674112
(Intercept)	Edad	RegionSierra	TopNestle	Valor de la posibilidad	1,169645597
(Intercept)	Edad	RegionCosta	TopSuperior	Valor de la posibilidad	0,001867186
(Intercept)	Edad	RegionSierra	TopSuperior	Valor de la posibilidad	0,00112833

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren Superior por encima de Galak, los valores que determinan que las

posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Galak por encima de Superior se encuentran en la Tabla CLXVII.

Tabla CLXVII					
Valores de las posibilidades del modelo de Chocolate categoría Superior					
(Intercept)	Edad	RegionCosta	TopSuperior	Valor de la posibilidad	9,20569E+19
(Intercept)	Edad	RegionSierra	TopSuperior	Valor de la posibilidad	2,48184E+13
(Intercept)	Edad	RegionCosta	TopNestle	Valor de la posibilidad	0,031307367
(Intercept)	Edad	RegionCosta	TopGalak	Valor de la posibilidad	0,004284708
(Intercept)	Edad	RegionSierra	TopOtras Marcas	Valor de la posibilidad	1,09588E-16
(Intercept)	Edad	RegionSierra	TopManicho	Valor de la posibilidad	2,72575E-18

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren La Universal por encima de Galak, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Galak por encima de La Universal se encuentran en la Tabla CLXVIII.

Tabla CLXVIII					
Valores de las posibilidades del modelo de Chocolate categoría Universal					
(Intercept)	Edad	RegionCosta	TopSuperior	Valor de la posibilidad	7,26201E+19
(Intercept)	Edad	RegionSierra	TopSuperior	Valor de la posibilidad	4,03809E+12
(Intercept)	Edad	RegionCosta	TopUniversal	Valor de la posibilidad	5,497578134
(Intercept)	Edad	RegionCosta	TopNestle	Valor de la posibilidad	0,503801799
(Intercept)	Edad	RegionSierra	TopGalak	Valor de la posibilidad	1,34962E-09
(Intercept)	Edad	RegionSierra	TopOtras Marcas	Valor de la posibilidad	3,17832E-16

Fuente y Elaboración: Mario Plaza

Los demás valores de posibilidades se encuentran en la sección Anexos.

En la tabla de clasificación que es la Tabla CLXIX nos permite observar que de los 118 entrevistados que prefirieron Nestlé, 94 fueron pronosticados como tal, es decir un 79,7%; y de los 205 encuestados que prefirieron Manicho, 205 fueron pronosticados como tal, es decir un 95,1%. En conclusión el modelo nos proporciona un 69,8% de correcta clasificación en forma general.

Tabla CLXIX							
Tabla de Clasificación del modelo de Chocolate							
Observado	Pronosticado						
	Otras Marcas	Nestle	Manicho	Galak	Superior	Universal	Porcentaje Correcto
Otras Marcas	64	40	6	0	0	4	56.1%
Nestle	18	94	4	0	0	2	79.7%
Manicho	6	3	195	0	0	1	95.1%
Galak	6	6	50	0	0	2	0.0%
Superior	0	2	7	0	0	0	0.0%
Universal	0	5	3	0	0	29	78.4%
Porcentaje global	17.2%	27.4%	48.4%	0.0%	0.0%	6.9%	69.8%

Fuente y Elaboración: Mario Plaza

Categoría: Café

En base a los resultados obtenidos a través de las Tablas de contingencia, el modelo dentro de ésta categoría es el siguiente:

Preferencia Café ~ Ciudad + Región + Género + Actividad + Top Café + Razones de Preferencia + 1er Atributo + 2do Atributo + 3er Atributo + 4to Atributo + 5to Atributo + 6to Atributo + 7mo Atributo + 8avo Atributo + 9no Atributo

Cuyo valor AIC es 1425,55.

En consecuencia consideramos un segundo modelo cuyo valor AIC sea inferior para poder predecir de forma más efectiva la variable dependiente, quedando de la siguiente forma:

Preferencia Café ~ Ciudad + Región + Género + Top Café + Razones de Preferencia

Cuyo valor AIC es 723,53.

Sin embargo en base al valor p se procede a eliminar una variable, la cual es Género, por lo tanto el modelo queda de la siguiente manera:

Preferencia Café ~ Ciudad + Región + Top Café + Razones de Preferencia

Cuyo valor AIC es 720,15.

En base a este modelo procedemos a obtener los coeficientes a través del software estadístico R-Studio. Debido a que el método de regresión logística multinomial toma una categoría como base vamos en este caso a tener cinco modelos. Sabemos que dentro de Café contamos con las marcas: Don café, Nescafé, Pres2, Colcafé, Sicafé y Otras Marcas. En este caso la categoría que el software consideró como categoría con la cual comparar es Colcafé, por lo tanto tenemos cinco modelos en base a las otras categorías, las cuales miden la posibilidad de preferir la empresa X por encima de la empresa Y.

Los coeficientes de los modelos podemos apreciarlos en las Tablas CLXX y CLXXI.

Tabla CLXX								
Valores de los Coeficientes del modelo de Café								
Categorías	(Intercept)	RegionSierra	Razones_Pref Por el aroma	Razones_Pref Porque es económico	Razones_Pref Porque es el de mejor	Razones_Pref Porque es una costumbre	Razones_Pref Porque me gusta su	Ciudad Guayaquil
Don Café	-3,433319	-36,0994464	-12,417786	-10,802819	-10,179126	-9,222737	-10,931029	10,7185947
Nescafé	5,617776	1,3455814	-11,546308	-13,200952	-9,79862	-9,785852	-10,269984	1,8627929
Otras Marcas	-38,419788	-8,5381463	-11,389795	-13,463196	-8,778828	-9,763379	-9,552293	-9,9069399
Pres 2	2,568611	-0,2502648	-7,991249	-8,545608	-7,768692	-6,627617	-7,712388	0,7632926
Sí Café	6,05466	2,8497309	-13,33389	-9,27716	-12,268375	-10,76916	-11,62991	1,0852419

Fuente y Elaboración: Mario Plaza

Tabla CLXXI								
Valores de los Coeficientes del modelo de Café								
Categorías	Ciudad Machala	Ciudad Manta	Ciudad Quito	TopDon Café	TopNescafé	TopOtras Marcas	TopPres 2	TopSícafé
Don Café	10,8875516	11,059981	44,5721044	7,87096	3,811927	16,046878	4,069754	5,402907
Nescafé	1,3896254	1,019776	-0,0939578	5,177279	7,006262	62,162551	4,488775	4,942263
Otras Marcas	-8,7835026	-11,191199	-1,7932605	58,300461	57,160496	120,696696	58,655881	19,570176
Pres 2	0,7441891	1,311394	-0,3758477	4,420609	4,339972	11,57746	8,737333	-48,095804
Sí Café	1,5281936	0,591494	-1,6629775	3,382566	2,759359	2,702372	-45,794129	9,050236

Fuente y Elaboración: Mario Plaza

En base a los valores de los coeficientes de las tablas anteriores, los modelos son los siguientes:

$$\eta_{Don\ Café} = -3,433319 - 36,0994464 \text{ Región (Sierra)}$$

$$- 12,417786 \text{ Razon de preferencia (Por el aroma)}$$

$$- 10,802819 \text{ Razon de preferencia (Es más barato)}$$

$$- 10,179126 \text{ Razon de preferencia (Calidad)}$$

$$- 9,222737 \text{ Razon de preferencia (Tradición)}$$

$$- 10,931029 \text{ Razon de preferencia (Me gusta su sabor)}$$

$$+ 10,7185947 \text{ Ciudad (Guayaquil)}$$

$$+ 10,8875516 \text{ Ciudad (Machala)} + 11,059981 \text{ Ciudad (Manta)}$$

$$+ 44,5721044 \text{ Ciudad (Quito)} + 7,87096 \text{ TOP(Don Café)}$$

$$+ 3,811927 \text{ TOP(Nescafé)} + 16,046878 \text{ TOP(Otras marcas)}$$

$$+ 4,069754 \text{ TOP(Pres 2)} + 5,402907 \text{ TOP(Si Café)}$$

El modelo nos indica que aquellos encuestados que provengan de cualquier ciudad y que en la pregunta del TOM hayan mencionado cualquier marca, aumentan la posibilidad de que prefieran a Don Café por encima de Colcafé.

$$\begin{aligned} \eta_{Nescafé} = & 5,617776 + 1,3455814 \text{ Región (Sierra)} \\ & - 11,546308 \text{ Razon de preferencia (Por el aroma)} \\ & - 13,200952 \text{ Razon de preferencia (Es más barato)} \\ & - 9,785852 \text{ Razon de preferencia (Calidad)} \\ & - 10,269984 \text{ Razon de preferencia (Tradición)} \\ & - 0,6011471 \text{ Razon de preferencia (Me gusta su sabor)} \\ & + 1,8627929 \text{ Ciudad (Guayaquil)} + 1,3896254 \text{ Ciudad (Machala)} \\ & + 1,019776 \text{ Ciudad (Manta)} - 0,09395778 \text{ Ciudad} \\ & + 5,177279 \text{ TOP(Don Café)} + 7,006262 \text{ TOP(Nescafé)} \\ & + 62,162551 \text{ TOP(Otras marcas)} + 4,488775 \text{ TOP(Pres 2)} \\ & + 4,942263 \text{ TOP(Si Café)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que provengan de cualquier ciudad y que en la pregunta del TOM hayan mencionado cualquier marca, aumentan la posibilidad de que prefieran a Nescafé por encima de Colcafé.

$$\begin{aligned} \eta_{Otras Marcas} = & -38,419788 - 8,5381463 \text{ Región (Sierra)} \\ & - 11,389795 \text{ Razon de preferencia (Por el aroma)} \\ & - 13,463196 \text{ Razon de preferencia (Es más barato)} \\ & - 8,778828 \text{ Razon de preferencia (Calidad)} \\ & - 9,763379 \text{ Razon de preferencia (Tradición)} \\ & - 9,552293 \text{ Razon de preferencia (Me gusta su sabor)} \\ & - 9,9069399 \text{ Ciudad (Guayaquil)} - 8,785026 \text{ Ciudad (Machala)} \\ & - 11,191199 \text{ Ciudad (Manta)} - 1,79326045 \text{ Ciudad (Quito)} \\ & + 58,300461 \text{ TOP(Don Café)} + 57,160496 \text{ TOP(Nescafé)} \\ & + 120,696696 \text{ TOP(Otras marcas)} + 58,655881 \text{ TOP(Pres 2)} \\ & + 19,570176 \text{ TOP(Si Café)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que en la pregunta del TOM hayan mencionado cualquier marca, aumentan la posibilidad de que prefieran Otras Marcas por encima de Colcafé.

$$\begin{aligned} \eta_{Pres2} = & 48,16254 - 894,9355 \text{ Región (Sierra)} \\ & + 2,822946 \text{ Razon de preferencia (Por el aroma)} \\ & - 1,49203 \text{ Razon de preferencia (Es más barato)} \\ & - 0,631666 \text{ Razon de preferencia (Calidad)} \\ & - 40,57094 \text{ Razon de preferencia (Tradición)} \\ & - 0,6011471 \text{ Razon de preferencia (Me gusta su sabor)} \\ & - 0,1434643 \text{ Ciudad (Guayaquil)} + 0,324169 \text{ Ciudad (Machala)} \\ & - 2,2460702 \text{ Ciudad (Manta)} + 1,6543147 \text{ Ciudad (Quito)} \\ & - 8,714219 \text{ TOP(Don Café)} - 13,31925 \text{ TOP(Nescafé)} \\ & - 6,453623 \text{ TOP(Otras marcas)} + 194,99206 \text{ TOP(Pres 2)} \\ & - 7,376698 \text{ TOP(Si Café)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que hayan mencionado como razón de preferencia El Aroma y que en la pregunta del TOM hayan mencionado Pres2, aumentan la posibilidad de que prefieran a Pres2 por encima de Colcafé.

$$\begin{aligned} \eta_{Si\ Café} = & 48,16254 - 894,9355 \text{ Región (Sierra)} \\ & + 2,822946 \text{ Razon de preferencia (Por el aroma)} \\ & - 1,49203 \text{ Razon de preferencia (Es más barato)} \\ & - 0,631666 \text{ Razon de preferencia (Calidad)} \\ & - 40,57094 \text{ Razon de preferencia (Tradición)} \\ & - 0,6011471 \text{ Razon de preferencia (Me gusta su sabor)} \\ & - 0,1434643 \text{ Ciudad (Guayaquil)} + 0,324169 \text{ Ciudad (Machala)} \\ & - 2,2460702 \text{ Ciudad (Manta)} + 1,6543147 \text{ Ciudad (Quito)} \\ & - 8,714219 \text{ TOP(Don Café)} - 13,31925 \text{ TOP(Nescafé)} \\ & - 6,453623 \text{ TOP(Otras marcas)} + 194,99206 \text{ TOP(Pres 2)} \\ & - 7,376698 \text{ TOP(Si Café)} \end{aligned}$$

El modelo nos indica que aquellos encuestados que hayan mencionado como razón de preferencia El Aroma, que provengan de las ciudades Machala o Quito y que en la pregunta del TOM hayan mencionado Pres2, aumentan la posibilidad de que prefieran a Si Café por encima de Colcafé.

Los valores de las posibilidades más altas que determinan que prefieren a Don Café por encima de Colcafé, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Colcafé por encima de Don Café se encuentran en la Tabla CLXXII.

Tabla CLXXII						
Valores de las posibilidades del modelo de Café categoría Don Café						
(Intercept)	RegionCosta	TopOtras Marcas	CiudadQuito	Razones_PrefOtras Razones	Valor de la posibilidad	6,8457E+24
(Intercept)	RegionCosta	TopDon café	CiudadQuito	Razones_PrefOtras Razones	Valor de la posibilidad	1,92603E+21
(Intercept)	RegionCosta	TopOtras Marcas	CiudadCuenca	Razones_PrefPor el aroma	Valor de la posibilidad	1,216250785
(Intercept)	RegionCosta	TopNescafe	CiudadGuayaquil	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,180778185
(Intercept)	RegionSierra	TopColcafe	CiudadCuenca	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,21298E-22
(Intercept)	RegionSierra	TopColcafe	CiudadCuenca	Razones_PrefPor el aroma	Valor de la posibilidad	2,7426E-23

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren a Nescafé por encima de Colcafé, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Colcafé por encima de Nescafé se encuentran en la Tabla CLXXIII.

Tabla CLXXIII						
Valores de las posibilidades del modelo de Café categoría Nescafé						
(Intercept)	RegionSierra	TopOtras Marcas	CiudadGuayaquil	Razones_PrefOtras Razones	Valor de la posibilidad	6,76085E+30
(Intercept)	RegionSierra	TopOtras Marcas	CiudadMachala	Razones_PrefOtras Razones	Valor de la posibilidad	4,21218E+30
(Intercept)	RegionSierra	TopSi café	CiudadMachala	Razones_PrefPorque es económico	Valor de la posibilidad	1,09888255
(Intercept)	RegionCosta	TopSi café	CiudadManta	Razones_PrefPor el aroma	Valor de la posibilidad	1,034074682
(Intercept)	RegionCosta	TopColcafe	CiudadCuenca	Razones_PrefPorque es económico	Valor de la posibilidad	0,000508942
(Intercept)	RegionCosta	TopColcafe	CiudadQuito	Razones_PrefPorque es económico	Valor de la posibilidad	0,000463301

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren Otras Marcas por encima de Colcafé, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Colcafé por encima de Otras Marcas se encuentran en la Tabla CLXXIV.

Tabla CLXXIV						
Valores de las posibilidades del modelo de Café categoría Otras Marcas						
(Intercept)	RegionCosta	TopOtras Marcas	CiudadCuenca	Razones_PrefOtras Razones	Valor de la posibilidad	5,40017E+35
(Intercept)	RegionCosta	TopOtras Marcas	CiudadQuito	Razones_PrefOtras Razones	Valor de la posibilidad	8,98678E+34
(Intercept)	RegionCosta	TopPres8	CiudadMachala	Razones_PrefPor el aroma	Valor de la posibilidad	1,064808957
(Intercept)	RegionCosta	TopNescafe	CiudadGuayaquil	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	1,05647733
(Intercept)	RegionSierra	TopColcafe	CiudadGuayaquil	Razones_PrefPorque es económico	Valor de la posibilidad	2,86357E-31
(Intercept)	RegionSierra	TopColcafe	CiudadManta	Razones_PrefPorque es económico	Valor de la posibilidad	7,92794E-32

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren Pres2 por encima de Colcafé, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que

determinan que prefieren Colcafé por encima de Pres2 se encuentran en la Tabla CLXXV.

Valores de las posibilidades del modelo de Café categoría Pres 2						
(Intercept)	RegionCosta	TopOtras Marcas	CiudadManta	Razones_PrefOtras Razones	Valor de la posibilidad	5165254,845
(Intercept)	RegionSierra	TopOtras Marcas	CiudadManta	Razones_PrefOtras Razones	Valor de la posibilidad	4021639,447
(Intercept)	RegionSierra	TopNescafe	CiudadCuenca	Razones_PrefPorque es una costumbre consumir esa marca	Valor de la posibilidad	1,031177342
(Intercept)	RegionCosta	TopDon café	CiudadMachala	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,0212436
(Intercept)	RegionCosta	TopSi café	CiudadQuito	Razones_PrefPorque es económico	Valor de la posibilidad	2,25555E-24
(Intercept)	RegionSierra	TopSi café	CiudadQuito	Razones_PrefPorque es económico	Valor de la posibilidad	1,75616E-24

Fuente y Elaboración: Mario Plaza

Los valores de las posibilidades más altas que determinan que prefieren Sícafé por encima de Colcafé, los valores que determinan que las posibilidades de escoger ambas empresas son cercanas y las que determinan que prefieren Colcafé por encima de Sícafé se encuentran en la Tabla CLXXVI.

Valores de las posibilidades del modelo de Café categoría Sí Café						
(Intercept)	RegionSierra	TopSi café	CiudadMachala	Razones_PrefOtras Razones	Valor de la posibilidad	289255373,1
(Intercept)	RegionSierra	TopSi café	CiudadGuayaquil	Razones_PrefOtras Razones	Valor de la posibilidad	185741930,9
(Intercept)	RegionSierra	TopNescafe	CiudadCuenca	Razones_PrefPorque me gusta su sabor	Valor de la posibilidad	1,034418983
(Intercept)	RegionSierra	TopDon café	CiudadCuenca	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	1,018755618
(Intercept)	RegionCosta	TopPres22	CiudadQuito	Razones_PrefPorque es el de mejor calidad	Valor de la posibilidad	4,90974E-24
(Intercept)	RegionCosta	TopPres20	CiudadQuito	Razones_PrefPor el aroma	Valor de la posibilidad	1,69165E-24

Fuente y Elaboración: Mario Plaza

Los demás valores de posibilidades se encuentran en la sección Anexos.

En la tabla de clasificación que es la Tabla CLXXVII nos permite observar que de los 144 entrevistados que prefirieron Don Café, 131 fueron pronosticados como tal, es decir un 91%, y de los 128 encuestados que prefirieron SiCafé, 114 fueron pronosticados como tal, es decir un 89,1%. En

conclusión el modelo nos proporciona un 87,2% de correcta clasificación en forma general.

Tabla CLXXVII							
Tabla de Clasificación del modelo de Café							
Observado	Pronosticado						Porcentaje Correcto
	Otras Marcas	Don Café	Nescafé	Colcafé	Sicafé	Pres 2	
Otras Marcas	112	3	2	0	0	2	94.1%
Don Café	0	131	7	1	3	2	91.0%
Nescafé	2	15	182	3	4	3	87.1%
Colcafé	0	2	27	91	1	1	74.6%
Sicafé	0	1	3	2	114	8	89.1%
Pres 2	0	2	4	1	4	73	86.9%
Porcentaje global	14.1%	19.1%	27.9%	12.2%	15.6%	11.0%	87.2%

Fuente y Elaboración: Mario Plaza

Conclusiones

1. El 59% de la población objetivo del estudio es femenina, lo cual tiene consistencia debido a que la encuesta fue realizada en horarios en que las amas de casa se encuentran en el hogar.
2. El 65% de los entrevistados pertenecen a la Costa y el 35% a la Sierra.
3. La variable Actividad fue dividida en dos partes. La primera cuenta con 785 entrevistados, de los cuales el 33% es ama de casa y el 24% son estudiantes.
4. La marca que fue más recordada sin ningún tipo de direccionamiento fue Sony con un 12,8%.
5. La principal razón por la que recuerdan estas marcas es por Calidad con un 29% y por Consumo un 27%.
6. Con respecto a las categorías de consumo fueron analizadas a través de la variable región. El uso de telefonía celular en la Costa es del 90% y en la Sierra es del 77%. El TOM más alta en la Costa pertenece indiscutiblemente a Claro con un 78%; sin embargo en la Sierra es más disputado el puesto, ya que Movistar cuenta con un 50% y Claro con un 49%. Lo mismo ocurre con la preferencia en la Costa Claro cuenta con un 77% y movistar solo con un 22%; en cambio en la Sierra Claro cuenta con un 55% y Movistar con un 44%. Entre las razones por las que se prefiere a determinada marca están: Cobertura

en la Costa con un 44% y en la Sierra con un 42%; Por sus contactos con un 11% en la Costa y 6% en la Sierra; y Por las promociones con 9% en la Costa y 20% en la Sierra.

7. El consumo de Bebidas Gaseosas en la Costa es del 79% y en la Sierra del 80%. En ambas regiones la marca con mayor TOM es Coca Cola con un 74% en la Costa y un 64% en la Sierra. En cuestión de preferencias la marca con mayor porcentaje de preferencia es Coca Cola con un 65% en la Costa y un 63% en la Sierra. Entre las razones de preferencia están: Porque me gusta su sabor con un 73% en la Costa y un 52% en la Sierra; Porque contiene menos colorante con un 7% en la Costa y un 8% en la Sierra; y Porque es una tradición familiar con un 7% en la Costa y un 14% en la Sierra. Estas personas que consideran la preferencia de determinada marca como preferencia son aquellas que son **fieles** a la marca.
8. El porcentaje de personas que consumen bebidas alcohólicas es del 42% en la Costa y del 33% en la Sierra, es considerablemente bajo debido a que los formularios fueron realizados en horas en que estaban las amas de casa y en su gran mayoría de la población son mujeres. De estos porcentajes un 72% respondieron como primera marca a Pilsener en la Costa que se les viene a la mente en primer lugar, mientras que en la Sierra la primera marca es Zhumir con un 41%, seguida de Pilsener con un 22%. Con respecto a la preferencia

no existe gran variante, en la Costa Pilsener con 65% seguida de Club con un 11%; y en la Sierra Zhumir cuenta con 38%, seguida de Pilsener con 24% y Vodka con 6%. Entre las razones de preferencia tenemos: Me gusta su sabor con un 65% en la Costa y 41% en la Sierra, Tradición con 10% en ambas regiones, y Porque es más barata con un 6% en la Costa y un 17% en la Sierra. Por lo que se puede apreciar el costo no es un gran problema para ninguna de las dos regiones.

9. El consumo de atún es de 88% en la Costa y del 91% en la Sierra. De entre los cuales en la pregunta del TOM en la Costa se tienen a Real con un 51%, seguido de Van Camps con un 36%; mientras que en la Sierra Real cuenta con un 70% y Van Camps con un 26%. La preferencia en la Costa la tiene Real con un 48% y Van Camps con un 38%, y en la Sierra Real cuenta con un 67% y Van Camps con un 29%. Entre las razones de preferencia tenemos: Me gusta su sabor con un 73% en la Costa y un 65% en la Sierra, Calidad con un 11% en la Costa y un 13% en la Sierra, y Economía con un 5% en la Costa y un 7% en la Sierra.
10. El consumo de Chocolate en barra es del 47% en la Costa y del 49% en la Sierra. En el TOM hay un empate con el 32% en la Costa entre Nestlé y Manicho, y en la Sierra 48% Manicho y un 9% Nestlé. La marca de preferencia cuenta Manicho en la Costa con un 31% y un

29% Nestlé, en cambio en la Sierra Manicho cuenta con un 49%, mientras que Nestlé solo con el 9%. Entre las razones de preferencia tenemos: Me gusta su sabor con un 78% en la Costa y un 66% en la Sierra y Costumbre un 7% en la Costa y 8% en la Sierra.

11. El consumo de Café en la Costa es del 74% y en la Sierra del 62%.

Con respecto al TOM en la Costa hay un empate entre Don Café y Nescafé con un 28%, seguido de Pres2 con un 15%, en cambio en la Sierra es Sicafé con un 31%, seguido de Colcafé con un 22% y Nescafé con un 19%. En cuestión de preferencias tenemos que en la Costa prefieren a Nescafé con un 29%, seguido de Don Café con un 26% y Pres2 con un 14%; en cambio en la Sierra prefieren a Sicafé con un 30%, seguido de Colcafé con un 22% y Nescafé con un 22%. Entre las razones de preferencia tenemos: Sabor con un 72% en la Costa y un 54% en la Sierra; Aroma con un 10% en la Costa y un 24% en la Sierra; y, Tradición con un 7% en la Costa y un 10% en la Sierra.

12. Con respecto a la importancia de determinados atributos a la hora de elegir una marca, el atributo que se escoge como primer lugar es Calidad con un 39% en la Costa y un 31% en la Sierra, Marca Confiable un 14% en la Costa y 15% en la Sierra, y Precio con un 13% en la Costa y un 18% en la Sierra.

13. Para elegir a Movistar por encima de Claro se debe haber mencionado en el TOM a Movistar, ser de Cuenca y entre las razones de

preferencia las Promociones. En cambio para preferir CNT ser de Guayaquil, haber mencionado en el TOM a CNT y la razón de preferencia las Promociones.

14. Para elegir a Pepsi en vez de Coca Cola, se debe ser de Quito, en el TOM mencionar Pepsi y la razón de preferencia Porque es refrescante. En cambio con Sprite ser de Guayaquil o Machala, en el TOM haber mencionado Pepsi y la Razón de preferencia Porque contiene menos colorante.
15. En la categoría de Bebidas Alcohólicas para preferir Pilsener en vez de Brahma, ser de Machala o Guayaquil, y la razón de preferencia la Calidad del producto. Con respecto a Club ser de Quito, en el TOM mencionar Zhumir o Vodka, y entre las razones de preferencia la Calidad o la Tradición.
16. En la categoría de atún, para que prefieran a Campos e vez de Van Camps el entrevistado debe ser de la Costa y haber mencionado Otras razones en las razones de preferencia. Con respecto a Campos Vs Real se debe ser de la Sierra, haber mencionado a Campos en el TOM y entre las razones de preferencia Calidad y Sabor.
17. En la categoría de Chocolate en barra, para preferir a Manicho en vez de Galak, es indistinta la región, que a menor edad mayor preferencia y que en el TOM hayan mencionado Manicho. En cambio con Nestlé

haber mencionado en el TOM la misma marca, indistinto la región y a mayor edad mayor preferencia.

18. En la categoría de Café, para preferir a Colcafé en vez de a Don Café el encuestado debe ser de Cuenca y preferir el Sabor o Aroma como razón de preferencia. Y con respecto a Colcafé vs Nescafé ser de Cuenca o Quito, haber mencionado en el TOM a Colcafé y la razón de preferencia Porque es económico.

Recomendaciones

1. Se recomienda que el estudio debe tener un porcentaje similar o cercano entre hombres y mujeres, específicamente para el consumo de determinados productos.
2. El porcentaje de la población objetivo en base a la región debería manejarse de forma equitativa también, que puede ser según la cantidad de ciudades y/o población objetivo.
3. El estudio puede ser considerado como prueba piloto si alguna empresa desea ahondar en cada categoría de producto para lo cual todas aquellas preguntas viertas se han de cerrar en base a las respuestas obtenidas en el mismo.
4. Se puede escoger determinada marca específica como marca base para compararla con las demás marcas de competencia lo cual va a depender de quien desea realizar dichas comparaciones.
5. Se recomienda que al realizar un estudio basado en el presente informe o uno similar, se debe colocar entre las variables de interés el Nivel Socioeconómico de los entrevistados.

4. Anexos de Tablas de Contingencia

4.1. Anexos de Tablas de Contingencia de Telefonía Celular.

Tabla A.1						
Tabla de Contingencia de Preferencia Celular vs Ciudad						
Marcas	Ciudad					
	Guayaquil	Quito	Machala	Cuenca	Manta	Total
Claro	307	103	121	65	93	689
Movistar	59	90	70	46	19	284
CNT	7	1	0	3	1	12
Total	373	194	191	114	113	985

Fuente y Elaboración: Mario Plaza

Tabla A.2			
Tabla de Contingencia de Preferencia Celular vs Región			
Marcas	Región		
	Costa	Sierra	Total
Claro	521	168	689
Movistar	148	136	284
CNT	8	4	12
Total	677	308	985

Fuente y Elaboración: Mario Plaza

Tabla A.3 a									
Tabla de Contingencia de Preferencia Celular vs Edad									
Marcas	Edad								
	18	19	20	21	22	23	24	25	26
Claro	30	29	32	20	19	20	25	27	15
Movistar	13	9	16	6	6	4	11	9	4
CNT	0	0	0	0	0	0	0	1	0
Total	43	38	48	26	25	24	36	37	19

Fuente y Elaboración: Mario Plaza

Tabla A.3 b									
Tabla de Contingencia de Preferencia Celular vs Edad									
Marcas	Edad								
	27	28	29	30	31	32	33	34	35
Claro	20	35	18	29	6	21	13	11	24
Movistar	5	10	8	17	7	9	7	5	13
CNT	1	0	0	0	0	0	0	2	1
Total	26	45	26	46	13	30	20	18	38

Fuente y Elaboración: Mario Plaza

Tabla A.3 c									
Tabla de Contingencia de Preferencia Celular vs Edad									
Marcas	Edad								
	36	37	38	39	40	41	42	43	44
Claro	11	20	26	11	32	7	19	7	12
Movistar	3	5	6	6	10	3	4	7	5
CNT	1	0	0	0	1	0	1	1	0
Total	15	25	32	17	43	10	24	15	17

Fuente y Elaboración: Mario Plaza

Tabla A.3 d									
Tabla de Contingencia de Preferencia Celular vs Edad									
Marcas	Edad								
	45	46	47	48	49	50	51	52	53
Claro	20	12	11	5	10	9	5	6	7
Movistar	10	3	7	4	6	7	4	6	3
CNT	1	0	0	0	2	0	0	0	0
Total	31	15	18	9	18	16	9	12	10

Fuente y Elaboración: Mario Plaza

Tabla A.3 e									
Tabla de Contingencia de Preferencia Celular vs Edad									
Marcas	Edad								
	54	55	56	57	58	59	60	61	62
Claro	5	13	8	4	2	1	5	3	4
Movistar	4	2	2	0	2	2	3	0	0
CNT	0	0	0	0	0	0	0	0	0
Total	9	15	10	4	4	3	8	3	4

Fuente y Elaboración: Mario Plaza

Tabla A.3 f							
Tabla de Contingencia de Preferencia Celular vs Edad							
Marcas	Edad						Total
	63	64	65	66	68	70	
Claro	2	2	6	3	1	6	689
Movistar	1	1	1	1	0	7	284
CNT	0	0	0	0	0	0	12
Total	3	3	7	4	1	13	985

Fuente y Elaboración: Mario Plaza

Tabla A.4			
Tabla de Contingencia de Preferencia Celular vs Género			
Marcas	Género		
	Hombre	Mujer	Total
Claro	295	394	689
Movistar	133	151	284
CNT	4	8	12
Total	432	553	985

Fuente y Elaboración: Mario Plaza

Tabla A.5 a									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Comerciante	Ama de casa	Transportista	Empleado Público	Niñero/a	Estilista	Chofer	Albañil	Jubilado
Claro	113	134	1	77	2	5	28	23	4
Movistar	42	56	0	57	1	1	10	3	5
CNT	3	5	0	3	0	0	0	0	0
Total	158	195	1	137	3	6	38	26	9

Fuente y Elaboración: Mario Plaza

Tabla A.5 b									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Abogado	Mécanico	Estudiante	Secretaria	Costurera	Marketing	Profesor	Pintor	Guardia
Claro	2	7	123	10	5	2	32	5	0
Movistar	5	1	42	4	4	0	6	1	1
CNT	0	0	0	0	0	0	0	0	0
Total	7	8	165	14	9	2	38	6	1

Fuente y Elaboración: Mario Plaza

Tabla A.5 c									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Negocio Propio	Ingeniero Comercia	Soldador	Ferretero	Radio Técnico	Ingeniero Agropecua	Asesor Comercial	Refrigeración	Artesania
Claro	6	5	3	1	4	3	5	1	4
Movistar	3	0	0	1	1	0	1	0	0
CNT	0	0	0	0	0	0	1	0	0
Total	9	5	3	2	5	3	7	1	4

Fuente y Elaboración: Mario Plaza

Tabla A.5 d									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Bailarín	Administrador	Modista	Enfermera	Policia	Cantante	Carpintero	Árbitro de Futbol	Evanista
Claro	0	7	1	1	1	1	2	1	5
Movistar	2	7	1	0	0	0	0	0	0
CNT	0	0	0	0	0	0	0	0	0
Total	2	14	2	1	1	1	2	1	5

Fuente y Elaboración: Mario Plaza

Tabla A.5 e									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Cocinero	Desempleado	Ejecutivo de Cobran	Asistente Administrat	Investigador	Servidor Público	Teleoperadora	Contador	Ganadero
Claro	3	6	1	6	0	1	0	3	1
Movistar	1	3	0	0	1	0	1	1	0
CNT	0	0	0	0	0	0	0	0	0
Total	4	9	1	6	1	1	1	4	1

Fuente y Elaboración: Mario Plaza

Tabla A.5 f									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Decorador	Ingeniero Civil	Médico	Marino Mercante	Agente de Ventas	Ingeniero en Sistem	Reciclador	Atención al Cliente	Electricista
Claro	2	1	4	1	2	2	1	5	1
Movistar	0	0	1	0	1	1	0	3	2
CNT	0	0	0	0	0	0	0	0	0
Total	2	1	5	1	3	3	1	8	3

Fuente y Elaboración: Mario Plaza

Tabla A.5 g									
Tabla de Contingencia de Preferencia Celular vs Actividad									
Marcas	Actividades								
	Entrenamiento Gene	Agricultor	Cajero	Economista	Diseñador Gráfico	Misionero	Farmacéutico	Pescador Artesana	de Recursos Hum
Claro	2	2	0	4	0	1	1	3	1
Movistar	1	0	1	2	1	0	2	2	0
CNT	0	0	0	0	0	0	0	0	0
Total	3	2	1	6	1	1	3	5	1

Fuente y Elaboración: Mario Plaza

Tabla A.5 h					
Tabla de Contingencia de Preferencia Celular vs Actividad					
Marcas	Actividades				
	Mensajero	Vendedor	Militar	Licenciado	Total
Claro	1	10	0	0	689
Movistar	1	2	1	1	284
CNT	0	0	0	0	12
Total	2	12	1	1	985

Fuente y Elaboración: Mario Plaza

Tabla A.6				
Tabla de Contingencia de Preferencia Celular vs Top Celular				
Marcas	Top Celular			
	Claro	Movistar	CNT	Total
Claro	649	39	1	689
Movistar	29	255	0	284
CNT	2	1	9	12
Total	680	295	10	985

Fuente y Elaboración: Mario Plaza

Tabla A.7							
Tabla de Contingencia de Preferencia Celular vs Razón preferencia							
Marcas	Razón de preferencia						
	Otras razones	Por los servicios que brinda	Porque la tarifa de llamadas es económica	Por su cobertura	Por las promociones	Por sus contactos/Amigos	Total
Claro	140	44	25	370	34	76	689
Movistar	80	10	38	52	89	15	284
CNT	4	0	3	0	4	1	12
Total	224	54	66	422	127	92	985

Fuente y Elaboración: Mario Plaza

Tabla A.8										
Tabla de Contingencia de Preferencia Celular vs 1er Atributo										
Marcas	1er Atributo									
	Color	Diseño	Tradición	Calidad	Marca Confiable	Envase	Presentación	Precio	Logotipo	Total
Claro	56	67	47	289	102	16	24	73	15	689
Movistar	20	38	18	87	42	9	8	54	8	284
CNT	0	0	1	6	1	0	0	3	1	12
Total	76	105	66	382	145	25	32	130	24	985

Fuente y Elaboración: Mario Plaza

Tabla A.9										
Tabla de Contingencia de Preferencia Celular vs 2do Atributo										
Marcas	2do Atributo									
	Color	Diseño	Tradición	Calidad	Marca Confiable	Envase	Presentación	Precio	Logotipo	Total
Claro	48	50	38	158	143	21	73	147	11	689
Movistar	19	26	9	77	56	16	26	49	6	284
CNT	0	0	1	2	7	0	2	0	0	12
Total	67	76	48	237	206	37	101	196	17	985

Fuente y Elaboración: Mario Plaza

Tabla A.10										
Tabla de Contingencia de Preferencia Celular vs 3er Atributo										
Marcas	3er Atributo									
	Color	Diseño	Tradición	Calidad	Marca Confiable	Envase	Presentación	Precio	Logotipo	Total
Claro	37	37	23	88	136	18	47	263	40	689
Movistar	24	18	17	23	59	6	25	94	18	284
CNT	0	1	0	2	0	1	2	6	0	12
Total	61	56	40	113	195	25	74	363	58	985

Fuente y Elaboración: Mario Plaza

4.2. Anexos de Tablas de Contingencia de Bebidas

Gaseosas.

Tabla B.1						
Tabla de Contingencia de Preferencia Gaseosa vs Ciudad						
Marcas	Ciudad					Total
	Guayaquil	Quito	Machala	Cuenca	Manta	
Otras Marcas	50	24	26	7	22	129
Coca Cola	217	132	113	70	56	588
Fanta	11	21	14	4	4	54
Pepsi	14	13	5	1	4	37
Sprite	19	22	13	20	5	79
Manzana	19	1	5	5	1	31
Total	330	213	176	107	92	918

Fuente y Elaboración: Mario Plaza

Tabla B.2			
Tabla de Contingencia de Preferencia Gaseosa vs Región			
Marcas	Región		
	Costa	Sierra	Total
Otras Marcas	98	31	129
Coca Cola	386	202	588
Fanta	29	25	54
Pepsi	23	14	37
Sprite	37	42	79
Manzana	25	6	31
Total	598	320	918

Fuente y Elaboración: Mario Plaza

Tabla B.3			
Tabla de Contingencia de Preferencia Gaseosa vs Género			
Marcas	Género		
	Hombre	Mujer	Total
Otras Marcas	45	84	129
Coca Cola	243	345	588
Fanta	30	24	54
Pepsi	13	24	37
Sprite	38	41	79
Manzana	13	18	31
Total	382	536	918

Fuente y Elaboración: Mario Plaza

Tabla B.4 a									
Tabla de Contingencia de Preferencia Gaseosa vs Actividad									
Marcas	Actividades								
	Comerciante	Ama de casa	Empleado Público	Niñero/a	Estilista	Chofer	Albañil	Jubilado	Abogado
Otras Marcas	13	35	11	1	1	6	3	1	1
Coca Cola	95	145	74	0	1	24	17	8	3
Fanta	10	7	13	0	1	3	1	0	0
Pepsi	8	10	4	0	0	3	1	1	0
Sprite	7	15	11	1	0	0	2	3	1
Manzana	5	4	3	0	1	0	0	0	0
Total	138	216	116	2	4	36	24	13	5

Fuente y Elaboración: Mario Plaza

Tabla B.5							
Tabla de Contingencia de Preferencia Gaseosa vs Top Gaseosa							
Marcas	Top Gaseosa						
	Otras Marcas	Coca Cola	Fanta	Pepsi	Sprite	Manzana	Total
Otras Marcas	89	32	3	2	1	2	129
Coca Cola	9	575	1	0	2	1	588
Fanta	4	11	37	0	2	0	54
Pepsi	0	7	0	30	0	0	37
Sprite	2	16	5	0	56	0	79
Manzana	2	9	0	1	1	18	31
Total	106	650	46	33	62	21	918

Fuente y Elaboración: Mario Plaza

Tabla B.6							
Tabla de Contingencia de Preferencia Gaseosa vs Razón preferencia							
Marcas	Razón de preferencia						
	Otras razones	Porque me gusta su sabor	Porque es una tradición familiar	Porque contiene menos colorante	Es refrescante	Por el gas	Total
Otras Marcas	22	81	1	11	9	5	129
Coca Cola	57	403	81	6	33	8	588
Fanta	2	35	1	8	4	4	54
Pepsi	2	28	0	1	3	3	37
Sprite	4	30	0	34	5	6	79
Manzana	0	24	0	6	1	0	31
Total	87	601	83	66	55	26	918

Fuente y Elaboración: Mario Plaza

Tabla B.7										
Tabla de Contingencia de Preferencia Gaseosa vs 1er Atributo										
Marcas	1er Atributo									
	Color	Diseño	Tradición	Calidad	Marca Confiable	Envase	Presentación	Precio	Logotipo	Total
Otras Marcas	9	13	6	49	19	5	6	21	1	129
Coca Cola	45	73	50	206	74	26	15	84	15	588
Fanta	8	8	6	12	11	1	2	6	0	54
Pepsi	0	3	0	15	5	1	1	10	2	37
Sprite	4	7	5	35	8	3	1	14	2	79
Manzana	1	2	1	11	12	1	1	2	0	31
Total	67	106	68	328	129	37	26	137	20	918

Fuente y Elaboración: Mario Plaza

Tabla B.8										
Tabla de Contingencia de Preferencia Gaseosa vs 2do Atributo										
Marcas	2do Atributo									
	Color	Diseño	Tradición	Calidad	Marca Confiable	Envase	Presentación	Precio	Logotipo	Total
Otras Marcas	3	9	8	32	26	9	15	23	4	129
Coca Cola	44	43	29	137	108	33	71	116	7	588
Fanta	5	7	0	16	10	1	2	11	2	54
Pepsi	4	2	3	9	7	1	5	6	0	37
Sprite	5	7	0	20	17	4	7	18	1	79
Manzana	1	1	3	10	7	1	2	5	1	31
Total	62	69	43	224	175	49	102	179	15	918

Fuente y Elaboración: Mario Plaza

Tabla B.9										
Tabla de Contingencia de Preferencia Gaseosa vs 3er Atributo										
Marcas	3er Atributo									
	Color	Diseño	Tradición	Calidad	Marca Confiable	Envase	Presentación	Precio	Logotipo	Total
Otras Marcas	10	8	3	17	22	4	13	48	4	129
Coca Cola	32	29	21	66	115	21	43	224	37	588
Fanta	4	7	3	5	9	2	3	13	8	54
Pepsi	1	3	2	3	10	1	2	13	2	37
Sprite	4	5	5	6	18	4	7	28	2	79
Manzana	3	0	0	3	5	2	3	15	0	31
Total	54	52	34	100	179	34	71	341	53	918

Fuente y Elaboración: Mario Plaza

4.3. Anexos de Tablas de Contingencia de Bebidas

Alcohólicas.

Tabla C.1						
Tabla de Contingencia de Preferencia Alcohol vs Ciudad						
Marcas	Ciudad					
	Guayaquil	Quito	Machala	Cuenca	Manta	Total
Otras Marcas	12	28	24	11	13	88
Pilsener	115	11	58	21	31	236
Brahma	7	1	1	1	1	11
Club	31	2	2	0	1	36
Zhumir	4	14	11	36	0	65
Vodka	3	8	2	0	0	13
Total	172	64	98	69	46	449

Fuente y Elaboración: Mario Plaza

Tabla C.2			
Tabla de Contingencia de Preferencia Alcohol vs Región			
Marcas	Región		
	Costa	Sierra	Total
Otras Marcas	49	39	88
Pilsener	204	32	236
Brahma	9	2	11
Club	34	2	36
Zhumir	15	50	65
Vodka	5	8	13
Total	316	133	449

Fuente y Elaboración: Mario Plaza

Tabla C.4			
Tabla de Contingencia de Preferencia Alcohol vs Género			
Marcas	Género		
	Hombre	Mujer	Total
Otras Marcas	53	35	88
Pilsener	128	108	236
Brahma	8	3	11
Club	25	11	36
Zhumir	34	31	65
Vodka	8	5	13
Total	256	193	449

Fuente y Elaboración: Mario Plaza

Tabla C.6							
Tabla de Contingencia de Preferencia Alcohol vs Top Alcohol							
Marcas	Top Alcohol						
	Otras Marcas	Pilsener	Brahma	Club	Zhumir	Vodka	Total
Otras Marcas	74	6	0	2	4	2	88
Pilsener	9	223	0	0	4	0	236
Brahma	0	4	7	0	0	0	11
Club	0	23	0	13	0	0	36
Zhumir	2	2	1	0	59	1	65
Vodka	5	0	0	0	1	7	13
Total	90	258	8	15	68	10	449

Fuente y Elaboración: Mario Plaza

Tabla C.7							
Tabla de Contingencia de Preferencia Alcohol vs Razón preferencia							
Marcas	Razón de preferencia						
	Otras razones	Porque me gusta su sabor	Porque es más barata/Tiene un precio más accesible	Por costumbre/ Por tradición	Porque es un producto de calidad	Es refrescante	Total
Otras Marcas	10	60	5	4	8	1	88
Pilsener	43	127	9	29	10	18	236
Brahma	0	9	2	0	0	0	11
Club	0	30	0	0	5	1	36
Zhumir	5	23	24	11	2	0	65
Vodka	0	10	1	0	2	0	13
Total	58	259	41	44	27	20	449

Fuente y Elaboración: Mario Plaza

4.4. Anexos de Tablas de Contingencia de Atún.

Tabla D.1						
Tabla de Contingencia de Preferencia Atún vs Ciudad						
Marcas	Ciudad					Total
	Guayaquil	Quito	Machala	Cuenca	Manta	
Otras Marcas	2	1	8	0	1	12
Vancamps	208	41	24	66	18	357
Real	120	198	173	47	27	565
Isabel	14	5	11	6	50	86
Starkist	1	0	0	0	2	3
Campos	1	2	0	0	0	3
Total	346	247	216	119	98	1026

Fuente y Elaboración: Mario Plaza

Tabla D.2			
Tabla de Contingencia de Preferencia Atún vs Región			
Marcas	Región		
	Costa	Sierra	Total
Otras Marcas	11	1	12
Vancamps	250	107	357
Real	320	245	565
Isabel	75	11	86
Starkist	3	0	3
Campos	1	2	3
Total	660	366	1026

Fuente y Elaboración: Mario Plaza

Tabla D.4			
Tabla de Contingencia de Preferencia Atún vs Género			
Marcas	Género		
	Hombre	Mujer	Total
Otras Marcas	5	7	12
Vancamps	162	195	357
Real	219	346	565
Isabel	36	50	86
Starkist	1	2	3
Campos	0	3	3
Total	423	603	1026

Fuente y Elaboración: Mario Plaza

Tabla D.6							
Tabla de Contingencia de Preferencia Atún vs Top Atún							
Marcas	Top Atún						Total
	Otras Marcas	Vancamps	Real	Isabel	Starkist	Campos	
Otras Marcas	5	0	7	0	0	0	12
Vancamps	0	314	39	2	2	0	357
Real	3	14	538	9	1	0	565
Isabel	1	4	9	72	0	0	86
Starkist	0	1	0	1	1	0	3
Campos	0	0	0	0	0	3	3
Total	9	333	593	84	4	3	1026

Fuente y Elaboración: Mario Plaza

4.5. Anexos de Tablas de Contingencia de Chocolate

Tabla E.1						
Tabla de Contingencia de Preferencia Chocolate vs Ciudad						
Marcas	Ciudad					Total
	Guayaquil	Quito	Machala	Cuenca	Manta	
Otras Marcas	32	23	27	23	9	114
Nestle	50	11	42	6	9	118
Manicho	79	40	21	57	8	205
Galak	15	27	6	9	7	64
Superior	8	0	1	0	0	9
Universal	28	0	6	0	3	37
Total	212	101	103	95	36	547

Fuente y Elaboración: Mario Plaza

Tabla E.2			
Tabla de Contingencia de Preferencia Chocolate vs Región			
Marcas	Región		
	Costa	Sierra	Total
Otras Marcas	68	46	114
Nestle	101	17	118
Manicho	108	97	205
Galak	28	36	64
Superior	9	0	9
Universal	37	0	37
Total	351	196	547

Fuente y Elaboración: Mario Plaza

Tabla E.4			
Tabla de Contingencia de Preferencia Chocolate vs Género			
Marcas	Género		
	Hombre	Mujer	Total
Otras Marcas	36	78	114
Nestle	33	85	118
Manicho	84	121	205
Galak	23	41	64
Superior	3	6	9
Universal	8	29	37
Total	187	360	547

Fuente y Elaboración: Mario Plaza

Tabla E.6							
Tabla de Contingencia de Preferencia Chocolate vs Top Chocolate							
Marcas	Top Chocolate						
	Otras Marcas	Nestle	Manicho	Galak	Superior	Universal	Total
Otras Marcas	97	7	4	2	0	4	114
Nestle	4	108	3	1	0	2	118
Manicho	7	2	193	2	0	1	205
Galak	7	5	5	45	0	2	64
Superior	0	2	0	1	6	0	9
Universal	0	5	1	1	1	29	37
Total	115	129	206	52	7	38	547

Fuente y Elaboración: Mario Plaza

4.6. Anexos de Tablas de Contingencia de Café.

Tabla F.1						
Tabla de Contingencia de Preferencia Café vs Ciudad						
Marcas	Ciudad					Total
	Guayaquil	Quito	Machala	Cuenca	Manta	
Otras Marcas	23	14	24	53	5	119
Don Café	94	1	32	0	17	144
Nescafé	123	39	25	9	13	209
Colcafé	23	49	33	5	12	122
Sicafé	14	42	38	33	1	128
Pres 2	24	3	15	1	41	84
Total	301	148	167	101	89	806

Fuente y Elaboración: Mario Plaza

Tabla F.2			
Tabla de Contingencia de Preferencia Café vs Región			
Marcas	Región		
	Costa	Sierra	Total
Otras Marcas	52	67	119
Don Café	143	1	144
Nescafé	161	48	209
Colcafé	68	54	122
Sicafé	53	75	128
Pres 2	80	4	84
Total	557	249	806

Fuente y Elaboración: Mario Plaza

Tabla F.4			
Tabla de Contingencia de Preferencia Café vs Género			
Marcas	Género		
	Hombre	Mujer	Total
Otras Marcas	47	72	119
Don Café	46	98	144
Nescafé	80	129	209
Colcafé	31	91	122
Sicafé	59	69	128
Pres 2	40	44	84
Total	303	503	806

Fuente y Elaboración: Mario Plaza

Tabla F.6							
Tabla de Contingencia de Preferencia Café vs Top Café							
Marcas	Top Café						
	Otras Marcas	Don Café	Nescafé	Colcafé	Sicafé	Pres 2	Total
Otras Marcas	112	3	2	0	0	2	119
Don Café	0	132	5	2	3	2	144
Nescafé	2	15	181	4	4	3	209
Colcafé	0	2	6	112	1	1	122
Sicafé	0	1	3	2	122	0	128
Pres 2	0	2	4	1	0	77	84
Total	114	155	201	121	130	85	806

Fuente y Elaboración: Mario Plaza

5. Anexos de Programación en R

5.1. Anexos de programación en R de Telefonía Celular.

```

setwd("C:/MARIO PLAZA/TESIS/bases/telefonía_celular")
library (plyr)
library(stats)
library(nnet)
BASE=data.frame(read.delim("base_telefonía_celular.txt",header=TRUE))
#categoria celulares
DATA=data.frame(Preferencia_Cel=BASE$pref_celular,
                Genero=BASE$genero,
                Edad=BASE$edad,
                Ciudad=BASE$ciudad,
                Actividad=BASE$actividad,
                Region=BASE$región,
                Importancia_Prim=BASE$imp_primera,
                Importancia_Seg=BASE$imp_segunda,
                Importancia_Terc=BASE$imp_tercera,
                Importancia_Cuarta=BASE$imp_cuarta,
                Importancia_Quinta=BASE$imp_quinta,
                Importancia_Sexta=BASE$imp_sexta,
                Importancia_Sept=BASE$imp_septima,
                Importancia_Octava=BASE$imp_octava,

```

```

    Importancia_Novena=BASE$imp_novena,
    Top=BASE$top_celular,
    Razones_Pref=BASE$raz_pref_celular)
DATA=DATA[complete.cases(DATA), ]

```

```

Modelo1=multinom(Preferencia_Cel~Region+Razones_Pref+Ciudad+Top
+Importancia_Prim+Importancia_Cuarta+Importancia_Quinta+Importancia
_Sept+Importancia_Novena,DATA,family = "ml")
step(Modelo1)
summary(Modelo1)
Modelo2=multinom(Preferencia_Cel~Region+Razones_Pref+Ciudad+Top
,DATA,family = "ml")
step(Modelo2)
summary(Modelo2)
Modelo3=multinom(Preferencia_Cel~Razones_Pref+Ciudad+Top,DATA,f
amily = "ml")
step(Modelo3)
summary(Modelo3)
#el valor de z es el estadístico de Wald
z=summary(Modelo2)$coefficients/summary(Modelo2)$standard.errors
z
#el valor p
p=(1-pnorm(abs(z), 0, 1))*2
p
#coeficientes del modelo
coef(Modelo2)
exp(coef(Modelo2))
exp(cbind(OR=coef(Modelo1),confint(Modelo2)))

```

5.2. Anexos de programación en R de Bebidas Gaseosas.

```

setwd("C:/MARIO PLAZA/TESIS/bases/gaseosas")
library (plyr)
library(stats)
library(nnet)
BASE=data.frame(read.delim("base_gaseosa.txt",header=TRUE))
#categoria gaseosas
DATA=data.frame(Preferencia_Gaseosa=BASE$pref_gaseosa,

```

```

Genero=BASE$genero,
Edad=BASE$edad,
Ciudad=BASE$ciudad,
Actividad=BASE$actividad,
Region=BASE$región,
Importancia_Prim=BASE$imp_primera,
Importancia_Seg=BASE$imp_segunda,
Importancia_Terc=BASE$imp_tercera,
Importancia_Cuarta=BASE$imp_cuarta,
Importancia_Quinta=BASE$imp_quinta,
Importancia_Sexta=BASE$imp_sexta,
Importancia_Sept=BASE$imp_septima,
Importancia_Octava=BASE$imp_octava,
Importancia_Novena=BASE$imp_novena,
Top=BASE$top_gaseosa,
Razones_Pref=BASE$raz_pref_gaseosa)
DATA=DATA[complete.cases(DATA), ]
Modelo1=multinom(Preferencia_Gaseosa~Region+Razones_Pref+Ciudad
+Top+Importancia_Cuarta+Importancia_Octava+Importancia_Novena,DA
TA,family = "ml")
step(Modelo1)
summary(Modelo1)
Modelo2=multinom(Preferencia_Gaseosa~Region+Razones_Pref+Ciudad
+Top,DATA,family = "ml")
step(Modelo2)
summary(Modelo2)
summary(DATA$Importancia_Sept)
z=summary(Modelo2)$coefficients/summary(Modelo2)$standard.errors
z
p=(1-pnorm(abs(z), 0, 1))*2
p
coef(Modelo2)
exp(coef(Modelo3))
exp(cbind(OR=coef(Modelo1),confint(Modelo3)))

```

5.3. Anexos de programación en R de Chocolate en Barra.

```

setwd("C:/MARIO PLAZA/TESIS/bases/chocolate")
library (plyr)
library(stats)
library(nnet)
BASE=data.frame(read.delim("base_chocolate.txt",header=TRUE))
#categoria chocolate
DATA=data.frame(Preferencia_Chocolate=BASE$pref_chocolate,
  Genero=BASE$genero,
  Edad=BASE$edad,
  Ciudad=BASE$ciudad,
  Actividad=BASE$actividad,
  Region=BASE$región,
  Importancia_Prim=BASE$imp_primera,
  Importancia_Seg=BASE$imp_segunda,
  Importancia_Terc=BASE$imp_tercera,
  Importancia_Cuarta=BASE$imp_cuarta,
  Importancia_Quinta=BASE$imp_quinta,
  Importancia_Sexta=BASE$imp_sexta,
  Importancia_Sept=BASE$imp_septima,
  Importancia_Octava=BASE$imp_octava,
  Importancia_Novena=BASE$imp_novena,
  Top=BASE$top_chocolate,
  Razones_Pref=BASE$raz_pref_chocolate)
DATA=DATA[complete.cases(DATA), ]
Modelo1=multinom(Preferencia_Chocolate~Region+Razones_Pref+Ciudad+Edad+Top+Importancia_Prim+Importancia_Sept+Importancia_Octava+Importancia_Novena,DATA,family = "ml")
step(Modelo1)
summary(Modelo1)
Modelo2=multinom(Preferencia_Chocolate~Region+Edad+Top,DATA,family = "ml")
step(Modelo2)
summary(Modelo2)
summary(DATA$Importancia_Sept)
z=summary(Modelo2)$coefficients/summary(Modelo2)$standard.errors
z

```

```
p=(1-pnorm(abs(z), 0, 1))*2
p
```

5.4. Anexos de programación en R de Café.

```
setwd("C:/MARIO PLAZA/TESIS/bases/cafe")
library (plyr)
library(stats)
library(nnet)
BASE=data.frame(read.delim("base_cafe.txt",header=TRUE))
DATA=data.frame(Preferencia_Cafe=BASE$pref_cafe,
  Genero=BASE$genero,
  Edad=BASE$edad,
  Ciudad=BASE$ciudad,
  Actividad=BASE$actividad,
  Region=BASE$región,
  Importancia_Prim=BASE$imp_primera,
  Importancia_Seg=BASE$imp_segunda,
  Importancia_Terc=BASE$imp_tercera,
  Importancia_Cuarta=BASE$imp_cuarta,
  Importancia_Quinta=BASE$imp_quinta,
  Importancia_Sexta=BASE$imp_sexta,
  Importancia_Sept=BASE$imp_septima,
  Importancia_Octava=BASE$imp_octava,
  Importancia_Novena=BASE$imp_novena,
  Top=BASE$top_cafe,
  Razones_Pref=BASE$raz_pref_cafe)
DATA=DATA[complete.cases(DATA), ]
Modelo1=multinom(Preferencia_Cafe~Genero+Region+Razones_Pref+Ciudad+Top+Actividad+Importancia_Prim+Importancia_Seg+Importancia_Terc+Importancia_Cuarta+Importancia_Quinta+Importancia_Sexta+Importancia_Sept+Importancia_Octava+Importancia_Novena,DATA,family = "ml")
step(Modelo1)
summary(Modelo1)
Modelo2=multinom(Preferencia_Cafe~Region+Razones_Pref+Top+Importancia_Prim+Importancia_Seg+Importancia_Terc+Importancia_Cuarta+Importancia_Quinta+Importancia_Sexta+Importancia_Sept+Importancia_Octava+Importancia_Novena,DATA,family = "ml")
```



```

step(Modelo2)
summary(Modelo2)
summary(DATA$Importancia_Sept)
Modelo3=multinom(Preferencia_Cafe~Genero+Region+Razones_Pref+Ciudad+Top+Actividad,DATA,family = "ml")
step(Modelo3)
summary(Modelo3)
Modelo4=multinom(Preferencia_Cafe~Genero+Region+Razones_Pref+Ciudad+Top,DATA,family = "ml")
step(Modelo4)
summary(Modelo4)
z=summary(Modelo4)$coefficients/summary(Modelo4)$standard.errors
z
p=(1-pnorm(abs(z), 0, 1))*2
p
Modelo5=multinom(Preferencia_Cafe~Region+Razones_Pref+Ciudad+Top,DATA,family = "ml")
step(Modelo5)
summary(Modelo5)

```

5.5. Anexos de programación en R de Atún.

```

setwd("C:/MARIO PLAZA/TESIS/bases/atun")
library (plyr)
library(stats)
library(nnet)
BASE=data.frame(read.delim("base_atun.txt",header=TRUE))
#categoria atun
DATA=data.frame(Preferencia_Atun=BASE$pref_atun,
                Genero=BASE$genero,
                Edad=BASE$edad,
                Ciudad=BASE$ciudad,
                Actividad=BASE$actividad,
                Region=BASE$región,
                Importancia_Prim=BASE$imp_primera,
                Importancia_Seg=BASE$imp_segunda,
                Importancia_Terc=BASE$imp_tercera,
                Importancia_Cuarta=BASE$imp_cuarta,

```

```

Importancia_Quinta=BASE$imp_quinta,
Importancia_Sexta=BASE$imp_sexta,
Importancia_Sept=BASE$imp_septima,
Importancia_Octava=BASE$imp_octava,
Importancia_Novena=BASE$imp_novena,
Top=BASE$top_atun,
Razones_Pref=BASE$raz_pref_atun)
DATA=DATA[complete.cases(DATA), ]
Modelo1=multinom(Preferencia_Atun~Region+Razones_Pref+Ciudad+Top+Importancia_Cuarta+Importancia_Quinta+Importancia_Sexta+Importancia_Sept+Importancia_Octava+Importancia_Novena,DATA,family = "ml")
step(Modelo1)
summary(Modelo1)
Modelo2=multinom(Preferencia_Atun~Region+Razones_Pref+Ciudad+Top,DATA,family = "ml")
step(Modelo2)
summary(Modelo2)
summary(DATA$Actividad)
z=summary(Modelo2)$coefficients/summary(Modelo2)$standard.errors
z
p=(1-pnorm(abs(z), 0, 1))*2
p
coef(Modelo2)
exp(coef(Modelo2))
exp(cbind(OR=coef(Modelo1),confint(Modelo2)))
Modelo3=multinom(Preferencia_Atun~Region+Razones_Pref+Top,DATA,family = "ml")
step(Modelo3)
summary(Modelo3)
z=summary(Modelo3)$coefficients/summary(Modelo3)$standard.errors
z
p=(1-pnorm(abs(z), 0, 1))*2
p
coef(Modelo2)
exp(coef(Modelo2))
exp(cbind(OR=coef(Modelo1),confint(Modelo2)))

```

5.6. Anexos de programación en R de Atún.

```

setwd("C:/MARIO PLAZA/TESIS/bases/bebidas alcoholicas")
library (plyr)
library(stats)
library(nnet)
BASE=data.frame(read.delim("base_alcohol.txt",header=TRUE))
DATA=data.frame(Preferencia_Alcohol=BASE$pref_alcohol,
  Genero=BASE$genero,
  Edad=BASE$edad,
  Ciudad=BASE$ciudad,
  Actividad=BASE$actividad,
  Region=BASE$región,
  Importancia_Prim=BASE$imp_primera,
  Importancia_Seg=BASE$imp_segunda,
  Importancia_Terc=BASE$imp_tercera,
  Importancia_Cuarta=BASE$imp_cuarta,
  Importancia_Quinta=BASE$imp_quinta,
  Importancia_Sexta=BASE$imp_sexta,
  Importancia_Sept=BASE$imp_septima,
  Importancia_Octava=BASE$imp_octava,
  Importancia_Novena=BASE$imp_novena,
  Top=BASE$top_alcohol,
  Razones_Pref=BASE$raz_pref_alcohol)
DATA=DATA[complete.cases(DATA), ]
Modelo1=multinom(Preferencia_Alcohol~Region+Razones_Pref+Ciudad+
Top+Actividad+Importancia_Sept+Importancia_Octava+Importancia_Nov
ena,DATA,family = "ml")
step(Modelo1)
summary(Modelo1)
Modelo2=multinom(Preferencia_Alcohol~Region+Razones_Pref+Ciudad+
Top,DATA,family = "ml")
step(Modelo2)
summary(Modelo2)
summary(DATA$Top)
z=summary(Modelo2)$coefficients/summary(Modelo2)$standard.errors
z
p=(1-pnorm(abs(z), 0, 1))*2

```

```
p
coef(Modelo2)
exp(coef(Modelo2))
exp(cbind(OR=coef(Modelo1),confint(Modelo2)))
```

Bibliografía

1. Agresti, A. (2002). *Categorical Data Analysis*. Florida: Wiley Interscience.
2. Fernández, S. d. (2011). *Regresión logística*. MADRID: UAM.
3. Fernández, V. P., & Fernández, R. S. (2004). *Regresión logística multinomial*. Madrid: Universidad de Valladolid.
4. SEQC. (2014). *Regresión logística*. Madrid.
5. Vicens Otero, J., & Medina Moral, E. (Enero de 2005). *Análisis de datos cualitativos*. Madrid, España.