



TARGET GROUP INDEX

Manual Metodológico de Clientes - 2023

1. Objetivo del Estudio

Target Group Index proporciona un conocimiento exhaustivo de los comportamientos de compra, la exposición y el consumo de medios de comunicación (incluidas las actividades en línea), los productos, las marcas, las opiniones y las actitudes.

Target Group Index ofrece numerosos cruces y combinaciones de información para conectar a los consumidores, las marcas y los vehículos. Con esta inteligencia conectada, aportamos información sobre las personas, sus comportamientos, los medios de comunicación, los productos y las marcas que consumen y cómo sus actitudes dan forma a sus pensamientos.

2. Cobertura del Estudio

El área de investigación y la entrega de datos tiene representatividad nacional.

3. Características Generales

El Estudio TGI presenta las siguientes características:

CARACTERÍSTICAS GENERALES	
Año	2023
Nombre de la base de datos	TGI PE 2023 R1
Mes de entrega de la base de datos	May/2023
Período de campo	15/05/2022 – 21/06/2022 y 03/01/2023 – 21/02/2023
Tamaño de muestra TOTAL	10.183
Tamaño de muestra en la OLA	5.098
Tamaño de población individual	18.888.242
Método de recopilación de datos	60% online 40% Face to Face
Frecuencia entrega base	2 veces al año

3.1. Muestra

El Estudio TGI corresponde a una entrevista aplicada a personas de 12 años a más que residen en el universo de la investigación. La muestra es controlada con las siguientes variables marginales:

- Región
- Nivel Socioeconómico (NSE)
- Género
- Edad

3.2. Universo

El Estudio TGI está respaldado por fuentes oficiales de información provenientes del Instituto Nacional de Estadística e Informática (INEI), como: Sistema de Consulta de Base de Datos REDATAM (Censos Nacionales 2017: XII de Población, VII de Vivienda y III de Comunidades Indígenas), la Cartografía asociada y los Niveles Socioeconómicos de APEIM.

Región	REDATAM, 12+ (000)	REDATAM, 12+ (%)	Sample Size 12+	Sample Size 12+ (%)	Weighted 12+ (000)	Weighted 12+ (%)
LIMA	9128	48%	4697	46%	9128	48%
Trujillo	771	4%	760	7%	771	4%
Chiclayo	490	3%	755	7%	490	3%
Piura	459	2%	754	7%	459	2%
Arequipa	891	5%	782	8%	891	5%
Cusco	407	2%	772	8%	407	2%
Huancayo	327	2%	754	7%	327	2%
Resto del País	6415	34%	909	9%	6415	34%
Total	18888	100%	10183	100%	18888	100%

4. Cuestionario

El cuestionario aplicado para el estudio TGI colecta la siguiente información dividido en las partes Core y Versión.

CONTENIDO DEL CUESTIONARIO	
CORE 1	<ul style="list-style-type: none"> • Demográfico Individual / Hogar / Jefe de Familia / Comprador principal • Bienes del Hogar/ Eventos de la vida • Consumo de productos del hogar (limpieza, bebidas, alimentos y bebés) • Compras del mercado • Mascotas • Finanzas • Auto • Preguntas Sensibles • Opiniones & Actitudes
CORE 2	<ul style="list-style-type: none"> • Tecnología • Televisión / Video on demand • Opiniones & Actitudes
CORE 3	<ul style="list-style-type: none"> • Internet • Redes Sociales • Videos Online • E-commerce • Opiniones & Actitudes
VERSIÓN A	<ul style="list-style-type: none"> • Deportes y Ocio • Cine • Apuestas Online • Vacaciones y Viajes • Ropa, Accesorios y Artículos • Centros Comerciales / Tiendas de Departamento • Movilidad (OOH) • Opiniones & Actitudes
VERSIÓN B	<ul style="list-style-type: none"> • Alimentos [consumo individual] • Restaurante/Cafetería • Periódicos • Revistas • Radio • Streaming de Música/Podcast • Diario de Actividades • Opiniones & Actitudes

VERSIÓN C

- Bebidas [consumo individual]
- Bebidas alcohólicas
- Tabaco
- Higiene Personal
- Maquillaje
- Marketing Boca a Boca
- Remedios y Salud
- Juegos
- Opiniones & Actitudes

5. Ponderación

La técnica de ponderación en el Estudio TGI es el RIM Weighting, esta técnica permite calcular el peso (proyección) para cada entrevistado a través de algoritmos e interacciones, luego al sumar los pesos de las personas estás convergen en los universos controlados.

Las variables utilizadas en el proceso de ponderación son:

- Ciudad
- Nivel Socioeconómico
- Género
- Edad
- Ola

6. Metodología

Para el estudio TGI, la metodología utilizada es 60% online y 40% Face to Face. La información es colectada por medio del cuestionario versionado de la siguiente manera:

- Una parte denominada CORE, la cual es aplicada al 100% de la muestra y que tiene un tiempo total de aplicación de 63 minutos.
- La parte VERSIONADA tiene una aplicación de 12 minutos, y es aplicada cada versión al 33% de la muestra (Versión A, B y C); considerando muestras con características espejos, es decir con las mismas informaciones demográficas. La aplicación de las versiones es aleatoria.

El cuestionario se presenta por medio de un link para ambos tipos de metodología, para la parte Online el proveedor colecta la información de manera independiente, tomando en cuenta las cuotas planeadas; y para la parte F2F la colecta se realiza mediante el uso de Tablet, en el cual se encuentra el desarrollo del link con la información del cuestionario.

7. Fusión de datos

Para hacer la composición de la base de datos con respuestas para el 100% de la muestra, se utiliza un modelaje de fusión de datos que busca a través de hooks informaciones similares entre las partes. Los hooks son las variables a tener en cuenta para identificar cuál encuestado va donar sus respuestas a otros. Dentro de estas variables utilizadas se encuentran las siguientes:

- Regiones
- Actividades realizadas en internet mientras ve videos
- Consumo de Videos Online, Sitios Web y Mensajes Instantáneos
- Acceso a Aplicaciones y Redes Sociales
- Dispositivo que accede a Internet
- Clasificación Social
- Publicidad en el Teléfono Móvil
- Consumo de Internet
- Rango de edad
- Consumo de Televisión
- Empleo
- Ama de Casa

8. DataScribe

Para nuevas marcas, categorías y variables en la base de datos, se utilizará un modelado en los datos que aplicará una probabilidad de todos los elementos nuevos en la base de datos combinada para el período anterior.

La principal ventaja de este proceso es obtener datos completos [12 meses] para cada novedad dentro de la Ola y ya no existe la necesidad de crear bases únicas.

Los indicadores de cambios [flags] continúan mostrándose en la herramienta de análisis Choices, lo que permite a los usuarios identificar qué datos se modelan para que se pueda tomar la precaución adecuada durante el análisis.

Para cada variable de período parcial individual, el algoritmo selecciona las mejores variables de predicción y las utiliza para generar un modelo de probabilidad que luego se utiliza para imputar datos para cada encuestado.

9. Control de Calidad Pre Producción

Para el estudio TGI, al culminar el campo, se realizan los siguientes procesos de control de calidad de las bases de datos de acuerdo con el tipo de metodología

9.1. Metodología Online

Para la parte online se identifican las encuestas consideradas válidas, sobre las cuales se obtiene el nivel socioeconómico y se genera el listado de las encuestas a considerar para el procesamiento de la información.

9.2. Metodología F2F

Para la parte F2F, al culminar el campo se realizan los siguientes procedimientos:

9.2.1. Supervisión

La supervisión de encuestas se realiza al 20% de la base total CAPI, la supervisión puede ser de tipo presencial o telefónica. Este procedimiento conlleva la revisión y eventual anulación de encuestas que no superen el proceso de supervisión. Este procedimiento se realiza por medio de la ficha de supervisión y de acuerdo con los lineamientos establecidos para este proceso.

9.2.2. Revisión de cuotas

Se verifica que se cumplan con las cuotas establecidas de acuerdo con la muestra contemplada para el estudio. Finalmente se identifican las encuestas finales, sobre las cuales se obtiene el nivel socioeconómico y se genera el listado de las encuestas a considerar para el procesamiento de la información.

10. Control de Calidad Post Producción

El objetivo principal del Control de Calidad Postproducción es validar que la información mostrada en el Software Choices sea la colectada en campo y que haya superado los respectivos controles. Dentro del proceso de Control de Calidad posterior al procesamiento se encuentran las siguientes actividades:

- Comprobación de las etiquetas mostradas, de acuerdo con la información colectada y validar el texto armonizado para un mejor entendimiento de las bases.
- Verificación del ranking de marcas de productos, validando que la tendencia se mantenga acorde al mercado y a lo colectado en el levantamiento de encuestas.
- Se realiza un análisis de tendencias, para revisar la evolución de la información respecto a las bases anteriores, verificando que se mantenga estable o que la variación tenga coherencia de acuerdo con comportamiento del consumidor y características generales.

Las observaciones que se encuentren en el control de calidad post producción son documentadas y compartidas con el equipo de procesamiento, con el fin de subsanarlas y garantizar la calidad de los datos en la entrega final.

11. Datos de Encuestados

La colecta de información y retención de datos se da de acuerdo con la metodología de la siguiente manera:

11.1. Metodología Online

Para la parte online, el proveedor del panel basa su metodología en las siguientes características:

- Validación de Identidad: Se da al momento del reclutamiento a través de comparación de informaciones con una base de datos de terceros, garantizando el nivel de validez.
- Validación de la Dirección IP: Se captura la dirección IP para validar el país y la región, detectando servidores proxy y contra una lista de servidores fraudulento garantizando que las respuestas provengan de la ubicación esperada.
- Detector de honestidad: Este enfoque analiza las respuestas para actividades de baja incidencia y preguntas comparativas, identificando combinaciones improbables. Se realiza con el objetivo de calificar la encuesta e identificar a los encuestados como “Honestos”; esta identificación se realiza cada 2 meses, bloqueándose a los encuestados considerados “deshonestos”.

- Respondientes únicos de la encuesta: Se utilizan huellas dactilares digitales patentadas para identificar y eliminar duplicados del estudio. Funciona junto a la validación de la dirección IP evitando que los encuestados ingresen más de una vez.
- Compromiso: Verifica si el encuestado está atento y comprometido con la encuesta a través de controles de velocidad y calificaciones de satisfacción; asegurando la calidad en las respuestas.

La gestión de incentivos y confidencialidad de datos de todos los panelistas para la metodología Online, se encuentra a cargo del proveedor del Panel Online.

11.2. Metodología F2F

Para la parte F2F, la identificación de los encuestados se da a partir de la selección. La selección aleatoria permite identificar el cluster, y a través de la metodología de cuotas se identifica a la persona participante del estudio.

Dentro de la información colectada de los encuestados, se encuentran los datos demográficos y la información de contacto. Los cuales son tratados de acuerdo con la política de privacidad de datos de Kantar Ibope Media, y cuyo consentimiento para el uso de datos personales es dado por el encuestado al realizarse la encuesta.

La colecta F2F apoya la motivación del encuestado para participar del estudio, con la entrega de incentivos; los cuales son entregados de acuerdo con el nivel del avance del campo en el que se encuentre y que es entregado al finalizar la encuesta, y entre los tipos de incentivos se cuenta con los siguientes:

- Utensilios de cocina (set de cubiertos o similares)
- Accesorios electrónicos (USB o similares)
- Variedades (polos, libretas o similares)

12. Error Muestral

TGI es un estudio de metodología no probabilística, por ello no presenta errores muestrales, sin embargo, otorga la oportunidad de trabajar los datos que se obtienen de la base Choices en un intervalo de confianza, asumiendo un muestreo aleatorio simple y una eficiencia estadística del 80%.

Esto quiere decir que la proporción que se obtiene de los cruces de la base, va a tener una probabilidad del 95% de oscilar en el intervalo calculado para este dato.

Para esto usaremos la siguiente formula:

- Intervalo de confianza al 95%:

$$IC = p \pm 1.96 * SE(p)$$

Donde:

- Proporción en la población con la característica de interés:

$$p = \frac{r}{N}$$

r = Suma de pesos de los encuestados con la característica de interés

N = Suma de pesos de todos los encuestados

- Error estándar de la estimación:

$$(SE)(p) = \sqrt{\frac{p(1-p)}{n_e}}$$

$n_e = n * eficiencia (0.8)$

n = Cantidad de muestra total del target

Ejemplo:

Target: Mujeres consumidoras de gaseosas en los últimos 12 meses.

	elements	total	GASEOSAS: CONSUMIÓ: Consumidores- U12m
total	Sample	10183	6251
	(000)	18888	11129
	vert%	100%	100%
	horz%	100%	58.9%
DATOS BÁSICOS: GÉNERO: Entrevistado:Masculino	Sample	4967	3024
	(000)	9195	5463
	vert%	48.7%	49.1%
	horz%	100%	59.4%
DATOS BÁSICOS: GÉNERO: Entrevistado:Femenino	Sample	5216	3227
	(000)	9693	5666
	vert%	51.3%	50.9%
	horz%	100%	58.5%

Note: In the original image, the values 5216 and 58.5% are highlighted with red boxes, and callouts labeled 'n' and 'p' point to them.

1. Para obtener la proporción de la población dividimos la población del target entre la población total

$$p = \frac{r}{N} = \frac{5666}{9693} = 58.45\%$$

2. Para obtener la desviación estándar utilizamos la siguiente formula:

$$SE(p) = \sqrt{((p(1 - p))/n_e)}$$

Donde n_e es el tamaño de muestra efectivo del target, el cual se calcula multiplicando el valor por la eficiencia estadística que es 80%.

$$SE(p) = \sqrt{\left(\frac{58.45\%(1 - 58.45\%)}{5216 * 80\%}\right)}$$

$$SE(p) = 0.0076$$

3. Reemplazamos los valores obtenidos para generar los intervalos de confianza

$$p + 1.96 \times SE(p) = 58.45\% + 1.96 \times 0.0076 = 59.95\%$$

$$p - 1.96 \times SE(p) = 58.45\% - 1.96 \times 0.0076 = 56.96\%$$

El Intervalo de confianza para las mujeres consumidoras de Gaseosas en los últimos 12 meses se encuentra entre el 56.96% al 59.95%.

13. Sistema de Gestión de Calidad

Declaramos que Kantar Ibope Media cumple con los requisitos según nuestra declaración de aplicabilidad de la Norma ISO 20252.

Mantenemos información documentada para replicabilidad del proceso y esta información se encuentra disponible en el documento bajo petición.