



J | JIPS
E | Essential
T | Toolkit

Guía de Muestreo para Situaciones de Desplazamiento & Ejemplos Prácticos

Junio, 2020



¡ESTAMOS AQUÍ!


Iniciar el ejercicio de caracterización


Establecer el proceso colaborativo


Diseñar la metodología


Implementar la recolección de datos


Procesar y analizar los datos


Validación, elaboración de informe y diseminación

Guía de Muestreo JET Partes I&II

Con el *JIPS Essential Toolkit* (JET, el kit de herramientas de caracterización de JIPS), emprendimos una misión con el fin de crear y disseminar herramientas y guías metodológicas accesibles que empoderen a nuestros colegas y socios de todo el mundo para construir una comprensión común sobre las situaciones de desplazamiento. Un ejercicio de caracterización, al producir una base de evidencia acordada capaz de informar el diseño de programas y políticas, puede contribuir al alcance de soluciones duraderas para las poblaciones desplazadas. A fin de compartir nuestras experiencias colectivas y asegurar que nadie se vea obligado a empezar desde cero, dividimos el proceso de caracterización en seis etapas, con un enfoque paso a paso y accesible. El JET es una colección de modelos, guías y *checklists* diseñadas para orientar a las personas durante la realización de un ejercicio de caracterización.

La Guía de Muestreo para Situaciones de Desplazamiento es parte de la etapa 3, que explica en detalle a las y los profesionales el diseño metodológico de un ejercicio de caracterización, y representa un elemento crítico de este proceso. El objetivo de diseñar una metodología es garantizar que el método aplicado y el análisis resultante realmente sean capaces de satisfacer las necesidades de información del contexto y que los resultados del análisis sean compatibles con los objetivos del ejercicio de caracterización. Una metodología de muestreo diseñada de manera adecuada asegura que las características de la muestra sean representativas —o al menos lo más representativas posible— de la población general, utilizando un enfoque comparativo. Por lo tanto, seleccionar y planificar correctamente el método de muestreo es la piedra angular para lograr la recolección de datos sólidos y confiables.

Dicho lo anterior, la Guía de Muestreo para Situaciones de Desplazamiento busca introducir brevemente algunos conceptos clave relacionados con el muestreo (por ejemplo, población objetivo, marco de muestreo y sesgo) y se enfoca en servir de referencia a las y los profesionales durante el proceso de selección de una muestra en situaciones complejas, como las de desplazamiento prolongado. Esta guía pretende ser lo más funcional posible: ofrece ejemplos concretos de metodologías de muestreo, tomados directamente de la experiencia de JIPS en el apoyo a ejercicios de caracterización a lo largo de los años, e incluye algunos recursos prácticos y materiales de referencia útiles.

La **Parte I** de la guía se centra en el muestreo como metodología y responde a preguntas clave sobre la selección del muestreo, la terminología, las opciones disponibles y la documentación. Antes de ahondar en los aspectos técnicos más detallados del proceso, esta sección ofrece definiciones y datos básicos relacionados con el muestreo. El contenido es presentado paso a paso y se revisan todas las opciones de muestreo disponibles, incluyendo una descripción detallada de cómo deben ser utilizadas, cómo decidir cuál opción elegir y qué errores se pueden esperar. De este modo, la Parte I es un documento de orientación que puede ser utilizado como referencia para comprender los aspectos teóricos y técnicos del muestreo. Considerando que existe una extensa literatura al respecto, este documento no espera cubrir a profundidad todos los aspectos posibles, sino que servir como un repaso y una introducción a conceptos clave.

La **Parte II** complementa a la Parte I de una manera muy práctica, presentando una selección de ejercicios de caracterización apoyados por JIPS. En esta sección, las personas usuarias de la guía encontrarán una descripción detallada de los métodos seleccionados, una explicación de por qué estos métodos fueron escogidos y cuáles fueron sus limitaciones en seis contextos de desplazamiento distintos (El Salvador, Grecia, Iraq, Kosovo, Somalia y Sudán). De este modo, las y los investigadores que estén interesados en llevar a cabo ejercicios de caracterización podrán vislumbrar su realidad operacional y, a partir de esto, formular hipótesis para contextos similares. Al compartir nuestra experiencia en el terreno, esperamos alentar e inspirar una comunidad de práctica que lidere el camino y establezca altos estándares en materia de manejo responsable de datos.

¿Quién es JIPS?

Esta guía fue elaborada por el equipo de JIPS, un servicio interagencial establecido en 2009 para apoyar a gobiernos y organizaciones humanitarias y de desarrollo en el diseño e implementación de ejercicios de caracterización de situaciones de desplazamiento. Nuestra misión principal es ofrecer apoyo —tanto en el terreno como de forma remota— mediante asistencia técnica, fortalecimiento de capacidades, formación, herramientas y orientación.

Reconocimientos

Esta Guía de Muestreo fue escrita como parte de la Fase 3 —Diseñar la Metodología— del JET. El proyecto es dirigido por JIPS con el fin de fortalecer capacidades relacionadas con la realización de ejercicios de caracterización. Se trata de un esfuerzo colectivo para compartir la experiencia, los conocimientos y los aprendizajes que hemos adquirido en el terreno y garantizar que nadie tenga que empezar desde cero al planear e implementar un ejercicio de caracterización.

Autoría: Marte Claussen (JIPS), Margharita Lundkvist-Houndoumadi (JIPS), Wilhelmina Welsch (JIPS), Dag Roll-Hansen (JIPS)

La elaboración de esta guía se benefició en gran medida de las contribuciones y del apoyo de nuestros socios de caracterización, incluyendo a Statistics Norway que, por medio de la colaboración que mantenemos, facilitó la participación de personas asesoras expertas.

Apreciamos y agradecemos a todos los miembros del equipo de JIPS por sus revisiones y aportaciones a lo largo del proceso, incluyendo un agradecimiento especial a Corina Demottaz por sus importantes comentarios y contribuciones.

Diagramación y diseño: Shorouq Manaseer y Mohammed Belo (VisuaLab)

Índice

<i>Guía de Muestreo JET Partes I&II</i>	2
Parte I: Sampling Guide for Displacement Situations	5
<i>Introducción</i>	6
<i>¿Por qué es fundamental elegir la metodología de muestreo adecuada?</i>	7
<i>¿Cuáles son los términos técnicos más importantes?</i>	7
<i>El proceso de selección de una metodología de muestreo</i>	9
<i>¿Cuáles son las distintas opciones de metodología?</i>	13
<i>La importancia de documentar el proceso</i>	24
<i>¿Dónde se puede aprender más?</i>	25
Parte II: Examples of sampling approaches in displacement situations	26
<i>Introducción y observaciones generales</i>	27
<i>Cuadro de resumen</i>	30
<i>Ejercicios de caracterización y metodologías de muestreo</i>	32
<i>Ejercicio de caracterización en El Salvador, 2016-2018</i>	32
<i>Ejercicio de caracterización en Tesalónica, Grecia, 2017-2019</i>	37
<i>Ejercicio de caracterización en Erbil, Región del Kurdistán de Iraq, 2015-2016</i>	43
<i>Ejercicio de caracterización en Kosovo, 2015-2016</i>	46
<i>Ejercicio de caracterización en Mogadiscio, Somalia, 2015-2016</i>	52
<i>Ejercicio de caracterización en Sudán, 2017-2019</i>	57

Parte I

Guía de Muestreo para Situaciones de Desplazamiento



Introducción

Cuando queremos recolectar información sobre un grupo de personas, tenemos que escoger entre recolectar información de todos los miembros del grupo o únicamente de una muestra de ellos. La primera alternativa es lo que se denomina “censo”. En teoría, un censo ofrece los datos más precisos. Sin embargo, en los contextos de desplazamiento, la realización de un censo muchas veces puede estar particularmente sujeta a sesgos, debido a una serie de motivos relacionados con aspectos específicos del contexto, así como con el tamaño y la complejidad operacional del ejercicio. Además, realizar un censo suele ser una tarea costosa y frecuentemente requiere años de planeación y preparación.

La segunda opción es lo que llamamos “encuesta por muestreo”. Para realizar una encuesta, seleccionamos una muestra (una porción de la población) que represente de la forma más precisa posible el total del grupo de interés del estudio. Con el diseño adecuado, una encuesta por muestreo es capaz de ofrecernos información con precisión suficiente para producir resultados confiables. Las encuestas, en muchos de los contextos de desplazamiento, pueden estar menos sujetas a determinados sesgos¹ y en la mayoría de las veces son más eficaces en materia de costos que un censo. Por lo tanto, las encuestas por muestreo suelen ser el método de recolección de información preferido en las situaciones de desplazamiento o en situaciones con un enfoque temático específico.

El objetivo de esta guía es presentar conceptos clave relativos al muestreo y servir como una referencia práctica para la selección de una muestra capaz de representar a la población de interés en un contexto de desplazamiento, así como ofrecer ejemplos pertinentes de métodos de muestreo que fueron utilizados en ejercicios de caracterización apoyados por JIPS². Aunque esta guía no pretende ofrecer una revisión exhaustiva de la teoría detrás de los conceptos estadísticos relativos a la selección de muestras, cuando sea posible se hará referencia a recursos que podrían ser útiles en este sentido.

Caracterización

Un ejercicio colaborativo que busca **establecer una comprensión común de las situaciones de desplazamiento** y de las circunstancias y características de las personas afectadas. Utiliza **enfoques de métodos mixtos, lo que suele incluir la aplicación de encuestas por muestreo** para recolectar y analizar datos de las poblaciones desplazadas, las comunidades que las acogen y otros grupos, y contextualiza la información obtenida tomando en cuenta el escenario económico, político y social más amplio. El objetivo general es crear **una base de evidencia integral y consensuada para informar intervenciones humanitarias y de desarrollo más eficaces, esfuerzos de incidencia y el desarrollo de políticas nacionales** para apoyar el alcance de soluciones duraderas para las poblaciones desplazadas.

Empezaremos con el análisis de por qué elegir el método de muestreo adecuado para una encuesta es fundamental para obtener resultados confiables. Luego, examinaremos algunos términos esenciales que serán de utilidad al leer el resto de la guía, así como al discutir sobre metodologías de muestreo con posibles socios o colegas. Finalmente, nos sumergiremos en el proceso de selección de un método de muestreo.

¹ Tales como los sesgos de inclusión y exclusión, los cuales serán explicados más adelante en este documento.

² Para conocer ejemplos prácticos de métodos de muestreo utilizados en ejercicios de caracterización apoyados por JIPS, ver la Parte II: Ejemplos de metodologías de muestreo en situaciones de desplazamiento.

¿Por qué es fundamental elegir la metodología de muestreo adecuada?

Elegir el método adecuado para seleccionar la muestra de una encuesta por muestreo es una parte crítica del proceso de diseño de la metodología general de un ejercicio de caracterización. Un método de muestreo diseñado adecuadamente nos permite suponer, con cierto grado de certeza, que las características de la muestra son válidas para la población más amplia a partir de la cual fue seleccionada (es decir, generalizar) y, por lo tanto, representa un prerequisite fundamental para la realización de un análisis sólido y riguroso.

Por otro lado, un método de muestreo diseñado de manera deficiente puede limitar drásticamente la utilidad de los datos recolectados para responder las preguntas de investigación. En resumen, el esfuerzo invertido en elegir y planificar un método de muestreo adecuado será recompensado futuramente durante el ejercicio de caracterización a través de la reducción del riesgo de errores y de potenciales gastos adicionales.

¿Cuáles son los términos técnicos más importantes?



Población objetivo

Es el grupo de población que los resultados de la encuesta por muestreo deben representar.

Ejemplo

"Personas refugiadas y solicitantes de asilo que llegaron a Tesalónica después de enero de 2015"



Marco de muestreo

Una lista de los individuos u hogares de la población objetivo, a partir de la cual se seleccionará la muestra. Podría ser, por ejemplo, una lista de direcciones. En un escenario ideal, se tendría acceso a una lista completa de toda la población objetivo. Sin embargo, este no suele ser el caso en los contextos de desplazamiento, por lo que se hace necesario encontrar otras formas de seleccionar una muestra representativa.



Muestra representativa

Al decidir cuál método de muestreo utilizar, la meta principal siempre debe ser que la muestra sea representativa de la población objetivo. Esto quiere decir que se debe seleccionar una muestra que se asemeje mucho a las características de la población objetivo y así refleje a dicha población de la manera más fiel posible. Por ejemplo, si el 50% de la población objetivo es del género femenino, idealmente la muestra también debe estar compuesta por un 50% de mujeres. Por lo general, seleccionar la muestra de manera aleatoria producirá una muestra representativa, con un nivel conocido de certeza. También se pueden tomar medidas adicionales para reforzar las posibilidades de obtener una muestra que sea representativa de la población objetivo por medio de la incorporación de estratos. El concepto de estratificación será explicado más adelante en este documento.

Nivel de confianza y margen de error

Estas dos variables describen el nivel de precisión de los resultados de una encuesta por muestreo: si se tiene un nivel de confianza de 95% con un margen de error de 5%, esto quiere decir que, si la encuesta es realizada en repetidas ocasiones (seleccionando una nueva muestra a partir del mismo marco de muestreo cada vez que la encuesta es repetida), en el 95% de las veces los resultados estarán a una distancia máxima de 5 puntos porcentuales de los valores reales de la población.

Tamaño de la muestra

Se refiere al número de individuos u hogares en la muestra. El tamaño de la muestra debe ser determinado con base en ciertas variables, como el nivel de confianza y el margen de error que se desea tener y la tasa de respuesta esperada. Muchas veces, la cantidad de fondos disponibles influye en la determinación de qué tamaños de muestra son factibles.

Tasa de respuesta

Es el porcentaje de individuos/hogares de la muestra que completan la encuesta. Cuanto más alta es la tasa de respuesta, más precisos serán los resultados. Por lo tanto, para obtener cifras correctas es fundamental tener una tasa de respuesta alta. Si la tasa de respuesta es baja, se puede intentar compensar esto con el uso de factores de ponderación. Es importante señalar que la "no respuesta" (esto es, la falta de respuesta) no puede ser solucionada mediante la inclusión en la muestra de nuevas personas u hogares que no habían sido incluidos inicialmente, dado que esto puede generar sesgos en los resultados de la encuesta y así ya no será posible afirmar que estos son precisos con determinado nivel de confianza y margen de error.

Sesgo

Se refiere a una situación en la que los resultados de una encuesta por muestreo sobrestiman o subestiman un parámetro poblacional (un valor estadístico que describe a la población objetivo) de forma sistemática. Los sesgos pueden ocurrir debido a distintos factores. Más adelante, esta guía describirá diferentes causas de sesgo.

Factor de ponderación

Un factor de ponderación es una variable utilizada en algunas ocasiones para atribuir un valor a cada observación de un conjunto de datos. Dicho factor atribuye un peso a determinada variable estadística para aumentar o disminuir su importancia con relación a las demás variables del conjunto de datos. En el contexto de las encuestas por muestreo, a veces se utilizan factores de ponderación para que las muestras se ajusten mejor a la población objetivo, tanto para el caso de muestras estratificadas no proporcionales como cuando se registran altos niveles de no respuesta, temas que serán abordados más adelante en esta guía.

El proceso de selección de una metodología de muestreo

¿Qué necesitamos saber antes de elegir la metodología de muestreo adecuada?

Existen varios aspectos clave y factores contextuales que deben ser definidos y tomados en cuenta al decidir qué método de muestreo utilizar en determinada situación de desplazamiento. A continuación, presentamos una lista de los más importantes:

-  *Determinar qué tipo de análisis es necesario para atender los objetivos*
-  *Determinar la población objetivo*
-  *Establecer el marco de muestreo*
-  *Tener en cuenta la disponibilidad de recursos*

Ejemplo

En una encuesta por muestreo realizada como parte de un *ejercicio de caracterización en Sudán en 2018*, una de las dos poblaciones objetivo eran hogares de personas desplazadas internamente (PDI) residentes en dos campamentos distintos. Al tomar la muestra para la encuesta, los hogares de PDI de algunas áreas de los campamentos fueron sobremuestreados o, en otras palabras, tuvieron una mayor probabilidad de ser incluidos en la muestra. El motivo del sobremuestreo de ciertas áreas estaba relacionado con medidas prácticas que tuvieron que ser tomadas en el terreno. De este modo, se usaron factores de ponderación para compensar el hecho de que algunos hogares habían tenido una mayor probabilidad de ser incluidos en la muestra. Los factores de ponderación redujeron la importancia de las respuestas de los hogares que habían sido sobremuestreados y aumentaron la importancia de las respuestas de los hogares que no habían sido sobremuestreados.

Determinar los resultados que se desea obtener del análisis

Al planear una encuesta por muestreo para un ejercicio de caracterización, uno de los primeros pasos debe ser determinar y definir claramente qué tipo de análisis las y los investigadores buscan realizar a partir de la encuesta para atender los objetivos del estudio. Antes de elegir el método de muestreo, será necesario discutir cuestiones como: ¿qué preguntas quiero que la encuesta sea capaz de contestar? ¿sobre quiénes quiero que hablen los resultados de la encuesta?

El tipo de análisis que se desea realizar debe ser discutido en detalle, dado que será determinante para seleccionar el tipo de método de muestreo más apropiado. Además de definir claramente la o las poblaciones objetivo del estudio (tema que será discutido de forma detallada a continuación), será necesario decidir si se desea poder desagregar los resultados de acuerdo con distintas características de la población objetivo y con cuáles criterios de desagregación se desea contar, lo que implica garantizar resultados representativos para diferentes subgrupos de la población objetivo. Por ejemplo, ¿se quiere examinar más de cerca los resultados de determinado subgrupo definido por ubicación o nacionalidad? Esto podría indicar que será necesario crear una muestra estratificada para asegurar resultados representativos para los subgrupos que se desea estudiar, tema que será discutido con mayor detalle a continuación.

No obstante, es importante señalar desde ya que incorporar muchos estratos puede aumentar drásticamente el tamaño de muestra necesario para producir resultados representativos para cada estrato, lo que también hará con que se necesite una mayor cantidad de recursos para realizar la encuesta. Por este motivo, se debe considerar cuidadosamente cuáles son los tipos de niveles de análisis necesarios para atender los objetivos de la encuesta.

Determinar la población objetivo

La población objetivo es el grupo de población que se quiere representar con los resultados de la encuesta por muestreo. Es fundamental que la población objetivo esté claramente definida, tanto en el espacio como en el tiempo. Un ejemplo de una población objetivo bien definida puede ser encontrado en el ejercicio de caracterización apoyado por JIPS en Tesalónica, Grecia, en 2017:

Ejemplo

“Personas refugiadas y solicitantes de asilo que llegaron a Tesalónica después de enero de 2015”

Esta población objetivo ha sido definida en el espacio —Tesalónica— y en el tiempo —de enero de 2015 hasta el momento de la encuesta—. En este ejemplo, también es importante aclarar quiénes son consideradas personas refugiadas o solicitantes de asilo. En este caso, se definió a personas refugiadas y solicitantes de asilo como aquellas que poseían documentos que demostraran que contaban con un prerregistro de solicitante de asilo, un registro de solicitante de asilo completo, una resolución de reconocimiento de la condición de persona refugiada, un permiso de residencia o una solicitud de asilo pendiente de recurso administrativo.

Uno de los componentes fundamentales de una caracterización es el enfoque comparativo, esto es, comparar la situación de diferentes grupos de población, por ejemplo: personas desplazadas internamente (PDI), personas refugiadas, personas retornadas, migrantes económicos y poblaciones de acogida. Incluso si el objetivo principal del ejercicio de caracterización es evaluar la situación de solo uno de estos grupos de población, muchas veces es necesario analizar la situación de los otros grupos de la zona (por ejemplo, del mismo municipio, región o país) para comprender si la situación y las características son específicas de dicho grupo de población o son compartidas por todos.

Si se decide realizar una encuesta por muestreo para más de un grupo poblacional, será necesario definir varias poblaciones objetivo y todas deben contar con definiciones claras, como se describió anteriormente. Dicho esto, definir poblaciones objetivo en un contexto de desplazamiento no siempre es una tarea fácil; por ejemplo, llegar a un acuerdo sobre las definiciones de las poblaciones objetivo puede representar un desafío para los socios. En algunos casos, existen definiciones reconocidas internacionalmente que pueden servir de referencia (por ejemplo, personas refugiadas³ y PDI⁴). En otros, sin embargo, no existen directrices sobre cómo definir determinadas poblaciones objetivo en contextos de desplazamiento (por ejemplo, cómo definir una “población de acogida”), y se hace necesario construir y llegar a un acuerdo sobre una definición técnica con los socios interesados de cada estudio específico.

³ Grupo de Expertos sobre Estadísticas de Refugiados y Desplazados Internos (2018). Recomendaciones Internacionales sobre Estadísticas de Refugiados. <https://www.jips.org/jips-publication/international-recommendations-on-refugee-statistics/>

⁴ Grupo de Expertos sobre Estadísticas de Refugiados y Desplazados Internos (2018). Recomendaciones Internacionales sobre Estadísticas de PDI. <https://unstats.un.org/unsd/statcom/51st-session/documents/BG-item-3n-international-recommendations-on-IDP-statistics-E.pdf>

Si bien la tarea de definir claramente las poblaciones objetivo puede parecer un desafío en algunas ocasiones, queremos destacar la **importancia de trabajar hacia el establecimiento de definiciones claras de las poblaciones objetivo en cuestión. Este es un prerrequisito fundamental para la realización de una encuesta por muestreo** y facilitará el trabajo a la hora de seleccionar la muestra, aplicar la encuesta y analizar los resultados.

Establecer el marco de muestreo

El siguiente paso en el proceso de selección de una muestra es establecer un marco de muestreo, esto es, una lista de individuos u hogares de la cual se puede extraer una muestra. Idealmente, el marco de muestreo debe **incluir a todas las personas u hogares de la población objetivo**. La posibilidad de crear o tener acceso a una lista exhaustiva y actualizada de la población objetivo es el escenario ideal para cualquier encuesta por muestreo y el mejor punto de partida para diseñar un método de muestreo que permita la selección de una muestra representativa.

Para averiguar si existe una lista de este tipo, o si es posible crearla a partir de la compilación de listas provenientes de diversas fuentes, es necesario analizar datos secundarios. Algunas fuentes relevantes para esto pueden ser, por ejemplo, registros de población que identifiquen a personas u hogares desplazados, datos censales recientes, listas de personas residentes en campamentos o listas de personas beneficiarias de programas en materia de desplazamiento.

No obstante, la disponibilidad de listas que identifiquen a personas u hogares desplazados varía de un contexto de desplazamiento a otro y a veces no es posible crear un marco de muestreo confiable a partir de datos secundarios. Esto puede ocurrir, por ejemplo, por tratarse de datos sensibles de compartir. Aun si se tiene acceso a una lista de la población objetivo, puede ser que no se reciba información sobre cómo la lista fue compilada y, por lo tanto, no siempre es posible saber si las listas tienen calidad suficiente.

A falta de una lista preexistente de la población objetivo, será necesario crear una. Sin embargo, en algunos contextos de desplazamiento puede no ser posible crear una lista completa y precisa de toda la población objetivo. Esto puede ocurrir por diversos motivos, por ejemplo, porque se tendría que buscar establecer contacto con poblaciones difíciles de alcanzar, como las que viven en lugares con acceso limitado, las que tienen un alto nivel de movilidad, las que se encuentran dispersas a lo largo de áreas extensas y/o las que desean permanecer "invisibles", o a veces simplemente porque no hay fondos suficientes. Crear una lista de este tipo a partir de la enumeración de la población objetivo requiere una gran cantidad de recursos, de modo que los fondos suelen ser un factor limitante.

El hecho de contar o no con la posibilidad de crear o acceder a un marco de muestreo preciso y completo tiene implicaciones para el diseño del método de selección de la muestra, así como para la interpretación de los hallazgos, es decir, **¿sobre qué y quiénes realmente podemos afirmar algo con base en los hallazgos de la encuesta? Estas implicaciones serán abordadas con mayor profundidad más adelante en la guía.**

En los siguientes cuadros, describimos con ejemplos algunos escenarios que podrían presentarse al momento de establecer un marco de muestreo para un ejercicio de caracterización:

Escenario ideal

Un registro exhaustivo y actualizado está disponible

En un ejercicio de caracterización realizado en Kosovo en 2018, una de las varias poblaciones objetivo eran PDI serbias residentes en áreas específicas de Kosovo. Los socios de la caracterización lograron acceder a una lista actualizada de todos los hogares de PDI serbias a partir de una base de datos gestionada por un organismo público en Serbia.

Es posible crear una lista exhaustiva y actualizada combinando distintas fuentes

En un ejercicio de caracterización realizado en Tesalónica en 2017, una de las poblaciones objetivo eran personas refugiadas que habían llegado a Tesalónica después de enero de 2015. Para construir el marco de muestreo de esta población objetivo, el equipo encargado de la caracterización combinó la información de una base de datos de las Naciones Unidas sobre personas refugiadas (la base de datos de registro ProGres, de ACNUR) con la lista de residentes de un campamento para personas refugiadas. Sin embargo, se constató que la información de la base de datos estaba desactualizada, lo que nos lleva al siguiente escenario.

Escenario menos ideal, pero más común

Los registros tienen sesgos y/o no están actualizados

En los contextos de desplazamiento, es poco común encontrar listas de registro que sean totalmente precisas. Muchas veces, esto ocurre porque la población en cuestión es muy móvil, lo que a su vez hace que sea difícil mantener una lista actualizada con direcciones u otra información de contacto. En el ejercicio de caracterización en Tesalónica, se hizo evidente que el equipo tendría que encontrar una forma de actualizar el marco de muestreo desfasado. Los métodos más comunes para actualizar un marco de muestreo pueden implicar la enumeración o la compilación de fuentes secundarias, tales como listas de las personas beneficiarias de determinados programas o listas gestionadas por autoridades locales.

Se necesita construir un marco de muestreo

Un ejemplo de un escenario donde no estaba disponible una lista de la población objetivo, y tampoco era posible construir una a partir de la combinación de fuentes preexistentes, es el ejercicio de caracterización realizado en El Salvador en 2016, cuya población objetivo eran hogares desplazados por la violencia entre 2006 y 2016. De este modo, el equipo tuvo que encontrar una manera alternativa de crear un marco de muestreo y seleccionar una muestra representativa. Se escogió un método de muestreo por conglomerados estratificados que incluía una enumeración. Este método será descrito más adelante.

Tener en cuenta la disponibilidad de recursos

Finalmente, otro factor importante que se debe considerar al elegir una metodología de muestreo es la cantidad de tiempo y fondos disponibles para la aplicación de la encuesta. Por ejemplo, hacer una enumeración completa en un área extensa puede ser una tarea costosa debido a los costos de contratar, capacitar y pagar personas para llevarla a cabo. De manera similar, trasladarse a lo largo de áreas extensas para acceder a la población objetivo que ha sido seleccionada para la muestra de una encuesta puede generar gastos de viaje altos, y muchas veces la cantidad de fondos disponibles también influirá en el tamaño que puede tener la muestra. Las limitaciones de tiempo también pueden afectar la viabilidad de los métodos de muestreo.

De este modo, la disponibilidad de tiempo y de fondos debe ser tomada en cuenta durante todo el proceso de selección del método de muestreo. La cantidad de recursos disponibles puede afectar particularmente el tipo de análisis que puede ser realizado (los niveles de desagregación) y a veces incluso el número de poblaciones objetivo o su definición.

¿Cuáles son las distintas opciones de metodología?

En esta sección describimos diferentes métodos que pueden ser usados para la selección de muestras, los aspectos que deben ser considerados al decidir cuál método utilizar, las limitaciones de cada método, así como algunos ejemplos sobre cómo pueden ser utilizados dependiendo del contexto. Analizaremos cada método por separado; sin embargo, en muchos casos es posible utilizar una combinación de varios métodos, como demuestran los ejemplos presentados en la Parte II.

Los métodos de muestreo han sido divididos en dos categorías: métodos probabilísticos y métodos no probabilísticos. En las *muestras probabilísticas*, cada miembro de la población objetivo tiene una probabilidad conocida de ser seleccionado. Esto quiere decir que es posible calcular qué tan probable es que un miembro sea incluido en la muestra final y **qué tan bien la información resultante de la muestra representa a la población objetivo** (ver la definición de intervalo de confianza y margen de error, presentada anteriormente).

En los *muestreos no probabilísticos*, se seleccionan miembros de la población objetivo de una manera no aleatoria. Lamentablemente, **el muestreo no probabilístico no permite saber con certeza en qué medida la información de la muestra difiere de la población objetivo; esto es algo que simplemente se desconoce**. Sin embargo, existen algunas formas de superar parte de la incertidumbre, las cuales serán abordadas más adelante.

Optar por un método de muestreo probabilístico siempre debe ser la opción preferida, aunque en los contextos de desplazamiento esto no siempre es viable. En caso de que no haya otra opción que recurrir a un muestreo no probabilístico, se debe buscar minimizar los sesgos que este tipo de método ocasiona. Describiremos una estrategia para esto más adelante.

Muestreo probabilístico

Cuando cada miembro de la población objetivo tiene una probabilidad conocida de ser incluido en la muestra, también es posible conocer la incertidumbre (es decir, el nivel de confianza y el margen de error) de los resultados, lo que permite evaluar en qué medida los hallazgos de la encuesta pueden ser generalizados para el total de la población objetivo. Esta información también puede ser compartida con el público o con las y los usuarios potenciales de los resultados de la encuesta, lo que aumenta la confianza y la credibilidad de los hallazgos. **Para utilizar un método de muestreo probabilístico, es necesario contar con un marco de muestreo.**

Cuando se tiene acceso a un marco de muestreo preexistente

Todos los métodos presentados a continuación requieren contar con acceso a un marco de muestreo actualizado, preciso y con buena cobertura de la población objetivo.

Muestra aleatoria simple



En un muestreo aleatorio simple, personas u hogares son seleccionados aleatoriamente a partir del marco de muestreo completo, y cada persona en el marco tiene la misma probabilidad de ser seleccionada. Este método es considerado la forma más básica de muestreo aleatorio y suele ser utilizado en combinación con otro método como, por ejemplo, el muestreo estratificado. La selección de una muestra aleatoria simple puede ser realizada a través de Excel o de algún otro programa informático que permita escoger una selección aleatoria de personas u hogares para la muestra.

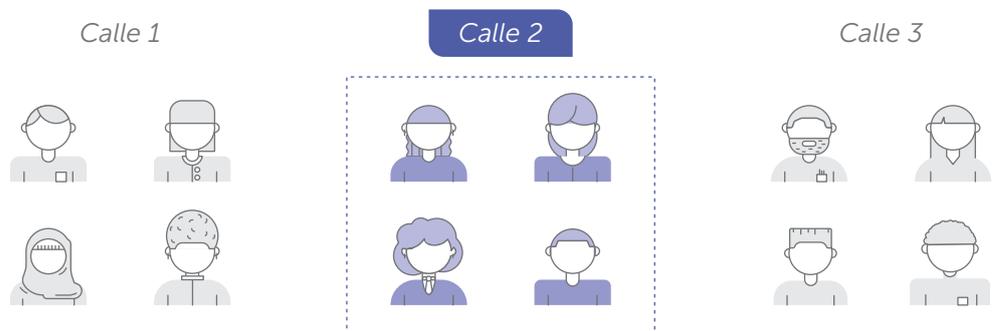
Muestra sistemática



Este método de muestreo es similar al muestreo aleatorio simple, en el sentido de que todas las personas de la población objetivo tienen la misma probabilidad de ser seleccionadas. Sin embargo, en el muestreo sistemático la muestra es seleccionada de una lista utilizando un intervalo fijo. Imaginemos que existe una lista en papel con las direcciones de todos los hogares de la población objetivo y se selecciona cada 12.^a dirección de la lista para conformar la muestra. Este método es una manera fácil de seleccionar la muestra, pero debe ser utilizado con cautela en caso de que exista algún patrón cíclico en la lista de hogares. Supongamos, por ejemplo, que la lista de direcciones está ordenada por calle. Entonces, cada 12.^o hogar seleccionado puede tener algo en común, como estar siempre situado en la planta baja o ser una casa de esquina. Estas características, a su vez, podrían afectar las respuestas de las personas encuestadas que conforman la muestra. Por lo tanto, si la lista está ordenada de alguna manera sistemática, es preferible que se intente reordenarla antes de tomar la muestra.

El método de muestreo sistemático es una técnica ampliamente utilizada, pero debe ser empleada con cautela para asegurar que ofrezca resultados tan buenos como los del muestreo aleatorio simple. El muestreo sistemático también es frecuentemente usado en combinación con algún otro método, como el muestreo estratificado.

Muestra por conglomerados



Agrupar de acuerdo con la ubicación y elegir un grupo

En una muestra por conglomerados, se seleccionan grupos de personas en lugar de individuos, tales como asentamientos, escuelas, distritos, entre otros. La ventaja de escoger un método de muestreo por conglomerados es que estos suelen ser menos costosos y más fáciles de implementar desde el punto de vista administrativo. Por ejemplo, si se desea aplicar una encuesta por muestreo a las y los estudiantes de un país, sería más rápido y menos costoso seleccionar aleatoriamente clases escolares y entrevistar a todas las personas de una clase que seleccionar aleatoriamente a estudiantes que podrían estar dispersos en un gran número de escuelas y clases a lo largo del país.

Una limitación del método de muestreo por conglomerados es que las personas que conforman el grupo seleccionado pueden ser más parecidas entre sí. Esto puede reducir la representatividad de la muestra y, consecuentemente, aumentar la incertidumbre de los resultados en comparación con un muestreo aleatorio simple o con un muestreo sistemático.

Muestra estratificada



En una muestra estratificada, se divide a la población objetivo en subgrupos que no se superponen, llamados estratos, para un mejor control de la muestra. La división en estratos es realizada con base en información auxiliar sobre alguna característica de la población objetivo, como el lugar de residencia (por ejemplo, áreas urbanas/rurales, regiones o municipios), la posición socioeconómica o variables demográficas. En el muestreo estratificado, cada subgrupo es tratado como una población objetivo independiente, y se selecciona una muestra separada para cada estrato. Dividir a la población objetivo en distintos estratos suele tener diversas ventajas.

En primer lugar, un muestreo estratificado puede aumentar la eficiencia y la representatividad de la muestra. Con una muestra aleatoria simple, en promedio se obtendrá una muestra representativa, pero existe un mayor riesgo de tener subgrupos sobre o subrepresentados en la muestra. Estratificar la muestra es una forma de intentar evitar sesgos y obtener estimaciones más precisas de la población objetivo. Además, si todos los subgrupos son homogéneos en lo que se refiere a las variables que se pretende estudiar, será posible obtener resultados con la calidad deseada a partir de una muestra más reducida en comparación con una muestra no estratificada.

En segundo lugar, el método de muestreo estratificado puede garantizar la representación de subgrupos pequeños de la población objetivo. Al tratar un subgrupo pequeño como una población objetivo independiente, es posible incluir una muestra representativa de este subgrupo en el estudio.

La distribución del tamaño de la muestra entre los estratos puede ser proporcional o no proporcional al tamaño de la población objetivo en cada estrato. En este último caso, la probabilidad de ser incluido en la muestra varía entre los distintos estratos. Al estimar una media poblacional con base en los datos de la encuesta, esta representación desigual debe ser compensada a través del uso de factores de ponderación. **El factor de ponderación corresponde al número de personas u hogares representados por cada persona u hogar de la muestra. No se utilizan factores de ponderación al realizar análisis que comparan unos estratos con otros; dichos factores solo son aplicados al calcular cifras que cubren toda la muestra.**

Ejemplo

En un ejercicio de caracterización realizado en Kosovo en 2015, una de las poblaciones objetivo eran hogares de PDI serbias que habían sido desplazadas entre enero de 1998 y marzo de 2004. Una lista completa de la población objetivo (un marco de muestreo) estaba disponible para el equipo de la caracterización. La mayor parte de la población objetivo vivía en municipios del norte de Kosovo y una proporción menor vivía en los municipios del sur.

Los socios del ejercicio de caracterización tenían información que indicaba que la situación de las PDI que vivían en el sur era distinta a la de las que vivían en los municipios del norte. Por este motivo, era recomendable contar con la posibilidad de desagregar los resultados por ubicación (norte y sur) para que fuera posible comparar la situación de la población objetivo en los dos lugares. Por lo tanto, se decidió estratificar a la población objetivo entre municipios del norte y municipios del sur para asegurar que la muestra general de PDI fuera representativa de la población objetivo de ambas regiones. El tamaño de la muestra fue distribuido igualmente entre los municipios del sur y del norte, es decir, de manera no proporcional al tamaño de la población objetivo en cada estrato, por lo que para el análisis de los hallazgos se requirió el uso de factores de ponderación cada vez que los resultados eran calculados a nivel nacional.

El equipo de la caracterización también quería poder comparar las situaciones de las PDI en las zonas urbanas y rurales. De este modo, se estratificó la muestra nuevamente con base en este criterio y de manera proporcional al tamaño de la población objetivo en áreas urbanas y rurales. Por ende, no fue necesario utilizar factores de ponderación para compilar los hallazgos de estos estratos.

Se utilizó el método de muestreo aleatorio simple para seleccionar la muestra de cada uno de los estratos.

Cuando no se tiene acceso a un marco de muestreo preexistente

En los contextos de desplazamiento, es común no tener acceso a un marco de muestreo completo y actualizado. Esto quiere decir que muchas veces es necesario construir uno. Sin embargo, identificar y enumerar a todas las personas de la población objetivo de determinado estudio suele ser una tarea demasiado ardua y costosa. Por lo tanto, no podemos depender exclusivamente de los métodos descritos anteriormente. Una estrategia usual para hacer frente a este desafío es elegir un método de muestreo en dos etapas.

Muestra aleatoria bietápica



Tal como lo sugiere su nombre, este método se realiza en dos etapas.

En la primera etapa, se divide el área donde reside la población en áreas geográficas más pequeñas llamadas Unidades Primarias de Muestreo (UPM), las cuales pueden ser municipios, colonias, manzanas o segmentos censales⁵. Luego, se hace una lista de las UPM y se extrae una muestra de estas, ya sea de manera aleatoria o sistemática. El número de UPM que deben ser seleccionadas depende del número total de hogares que se desea incluir en la muestra y del tamaño de las UPM.

Si se sabe cuántas personas de la población objetivo viven en cada unidad, la muestra debe ser tomada con una probabilidad proporcional al número de personas de la población objetivo residentes en cada área. De este modo, las áreas con muchos miembros de la población objetivo tendrán una mayor probabilidad de ser seleccionadas: si en el área A vive el doble de población objetivo que en el área B, la probabilidad de que un individuo u hogar del área A sea seleccionado es dos veces más alta que la de aquellos en el área B.

En la segunda etapa, se prepara una lista (un marco de muestreo) de la población objetivo (por ejemplo, hogares de PDI) que vive en cada una de las UPM seleccionadas. Esto generalmente se hace a partir de la enumeración de todos los hogares de las unidades seleccionadas o a través de consultas a la administración local o a organizaciones locales. Luego, con base en estas listas, se extrae una muestra aleatoria de cada UPM. Los hogares son las llamadas Unidades Secundarias de Muestreo (USM). Si las Unidades Primarias de Muestreo tienen tamaños similares, se debe seleccionar el mismo número de hogares de cada una de ellas. En cambio, si las Unidades Primarias de Muestreo tienen tamaños distintos, el número de Unidades Secundarias de Muestreo seleccionadas en cada UPM debe ser proporcional al número de USM en cada UPM.

Uno de los beneficios de realizar el muestreo en dos etapas es tener una muestra de hogares que están ubicados en la misma área, de modo que los costos de traslado de las y los enumeradores disminuyen. Sin embargo, seleccionar la muestra en dos etapas aumenta la incertidumbre de los resultados y genera algo llamado "efecto de diseño". Esto implica que se tendrá que aumentar el tamaño de la muestra en comparación con el tamaño que sería necesario si se hubiera utilizado un muestreo aleatorio simple.

⁵ Área geográfica establecida por la oficina nacional de estadísticas de un país con el fin de realizar un censo.

Ejemplo

En un ejercicio de caracterización realizado en Mogadiscio, Somalia, en 2015, las poblaciones objetivo eran PDI, migrantes económicos y comunidades de acogida residentes en asentamientos informales de tres distritos de la ciudad de Mogadiscio. Se hizo un mapeo de todos los asentamientos existentes en estos distritos, seguido de una enumeración de los hogares de todos los asentamientos identificados con el fin de agruparlos en las distintas poblaciones objetivo.

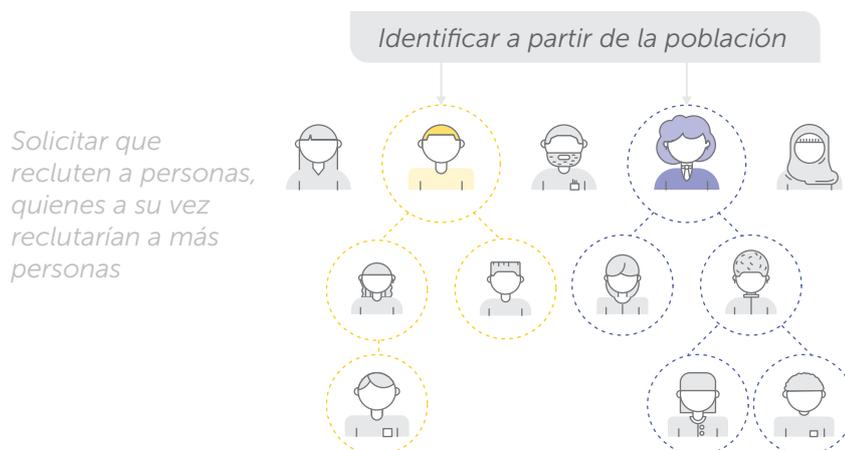
Luego, se utilizó un método de muestreo en dos etapas para la población objetivo de PDI. Los asentamientos fueron estratificados entre los tres distritos. En la primera etapa, se seleccionaron aleatoriamente asentamientos (unidades primarias de muestreo) de cada uno de los distritos, con base en una probabilidad proporcional al tamaño de la población desplazada internamente en cada asentamiento. En la segunda etapa, se tomó una muestra aleatoria de hogares (unidades secundarias de muestreo) de cada uno de los asentamientos seleccionados, con base en las listas elaboradas durante la enumeración. Los hogares seleccionados fueron entrevistados para la encuesta por muestreo.

Muestreo no probabilístico

En los contextos donde un marco de muestreo preexistente no está disponible, se debe preferir el uso de un método bietápico. Sin embargo, si a pesar de que se ha mapeado a las poblaciones resulta particularmente difícil encontrar a los hogares desplazados, o si no existen los recursos para llevar a cabo una enumeración completa de las unidades primarias de muestreo, puede ser necesario recurrir al muestreo no probabilístico. Utilizar técnicas no probabilísticas para la selección de una muestra probablemente generará resultados más sesgados en comparación con los que se obtienen a partir de las técnicas probabilísticas, ya que muchas veces algunos grupos no quedan representados en la muestra. Además, es difícil evaluar si una muestra no probabilística es representativa. En otras palabras, es difícil —si no es que imposible, considerando las herramientas que estén a nuestro alcance— calcular la incertidumbre de cualquier estimación que se realice. Si se decide utilizar un método de muestreo no probabilístico, se debe asegurar que siempre se evidencien estas limitaciones cuando se comuniquen los resultados del estudio.

No obstante, se ha comprobado que, cuando utilizada con cautela y con medidas adicionales para aumentar la representatividad, esta técnica ofrece resultados razonablemente sólidos, a pesar de que los resultados son más inciertos y que se necesita un tamaño de muestra más grande en comparación con el muestreo sistemático o el aleatorio.

Muestra “bola de nieve”



Una de las formas más básicas y conocidas del muestreo no probabilístico es el método llamado “bola de nieve”. El muestreo bola de nieve se basa en las redes de contactos de la población objetivo. Este método de muestreo implica identificar a algunas personas que pertenecen a la población objetivo y preguntarles si podrían reclutar o referir a otras personas que también pertenecen a la misma población objetivo. A las personas reclutadas por los sujetos iniciales, a su vez, se les pide que recluten o refieran a otras personas, y así el proceso continúa y la muestra crece. Se debe continuar este proceso hasta que se alcance el tamaño de muestra necesario (el cual debería ser mayor que en un muestreo aleatorio simple) o que se llegue al punto de saturación. Llegar al punto de saturación quiere decir alcanzar un punto donde se está seguro de que todas las personas de la población objetivo han sido identificadas y no se reciben nuevas referencias. Esta es la única forma de asegurar que se ha logrado una muestra representativa (en la práctica, un conteo completo), ya que toda la población objetivo ha sido incluida. En estos casos, el muestreo bola de nieve también puede ser usado para crear un marco de muestreo, a partir del cual se puede seleccionar una muestra aleatoria simple. Sin embargo, puede ser difícil saber si realmente se ha alcanzado el punto de saturación, aun si no se reciben nuevas referencias. Esto podría ocurrir porque en los contextos de desplazamiento las personas a veces quieren permanecer invisibles o porque no ha sido posible identificar a personas con redes de contactos pequeñas o inexistentes. Alcanzar el punto de saturación puede ser posible en los casos donde las poblaciones objetivo son relativamente pequeñas, pero es bastante raro con una población objetivo grande.

Una limitación importante del muestreo bola de nieve es el hecho de que las personas encuestadas probablemente reclutarán a personas con características similares a las suyas, lo que disminuye la representatividad. Además, personas con redes de contactos amplias tienen mayor probabilidad de ser incluidas en la muestra y, a su vez, referirán a relativamente más personas en comparación con quienes poseen redes de contactos más reducidas. Considerando lo anterior, una forma de aumentar la probabilidad de obtener una muestra representativa es garantizar que el conjunto inicial de personas encuestadas difiera en características importantes, como determinadas variables demográficas y socioeconómicas.

Otra limitación del método bola de nieve es que en algunos contextos las personas encuestadas pueden no querer referir a otras personas de la población objetivo con el fin de protegerlas, ya que ser identificadas como parte de la población objetivo podría representar una amenaza. Esto es conocido como “enmascaramiento”. Una forma de evitar el enmascaramiento es solicitar a las personas encuestadas que sean ellas mismas quienes recluten a las otras personas, en lugar de que solo proporcionen nombres e información de contacto al equipo encargado de la caracterización. De este modo, las personas tienen la posibilidad de escoger si desean ser encuestadas o no sin haber sido previamente identificadas por el equipo de la caracterización.

Muestra dirigida por las personas encuestadas

El muestreo dirigido por las personas encuestadas es similar al muestreo bola de nieve, pero ha sido desarrollado como un intento de compensar el hecho de que la muestra ha sido seleccionada de una manera no aleatoria. Este método combina el muestreo bola de nieve con un modelo matemático y busca generar muestras que sean independientes de las personas con las cuales el proceso de muestreo empezó, además de corregir las diferencias entre los tamaños de las redes de contactos de las personas⁶.

En primer lugar, este método trata de limitar el sesgo relacionado con las diferencias entre los tamaños de las redes de contactos de las personas a través del establecimiento de un límite al número de referencias que pueden ser hechas por cada persona (generalmente este límite es de tres referencias). Para implementar esta restricción, se entrega un número limitado de cupones a cada persona encuestada para que los transfiera a otros miembros de la población objetivo. El cupón de cada persona encuestada tiene un número de identificación único, y también contiene el número de identificación del cupón de la persona que la ha reclutado. De esta manera es posible controlar quién ha sido referido por quién, lo que también podría ayudar a evitar que las personas sean referidas más de una vez. En segundo lugar, el método intenta corregir las diferencias entre los tamaños de las redes de contactos a partir de la aplicación de factores de ponderación a los datos de las personas encuestadas, de acuerdo con el tamaño de su red de contactos. Los datos proporcionados por personas con redes de contactos más reducidas adquieren mayor peso que los datos proporcionados por personas con redes de contactos más amplias. Para lograr esto, durante el proceso de recolección de información es necesario obtener información sobre el tamaño de la red de contactos de cada persona encuestada (por ejemplo, incluyendo una pregunta sobre cuántas personas de la población objetivo cada persona encuestada conoce).

Este método también sugiere que se ofrezcan incentivos (por ejemplo, pequeñas cantidades de dinero) a las personas encuestadas, tanto por participar como por reclutar a más personas. Ofrecer una pequeña remuneración para cada nuevo reclutamiento puede generar una mayor disposición de participar en la encuesta y, consecuentemente, un aumento de la muestra.

El reclutamiento de nuevas personas para participar en la encuesta debe continuar hasta que se obtenga el tamaño de muestra requerido, en el caso de que se conozca el tamaño de la población objetivo, o hasta que se haya alcanzado el llamado equilibrio. Se considera que se ha alcanzado un equilibrio cuando las características de las personas encuestadas al inicio ya no tienen una influencia sobre la muestra, como cuando, por ejemplo, la proporción de hombres y mujeres permanece estable en los reclutamientos.

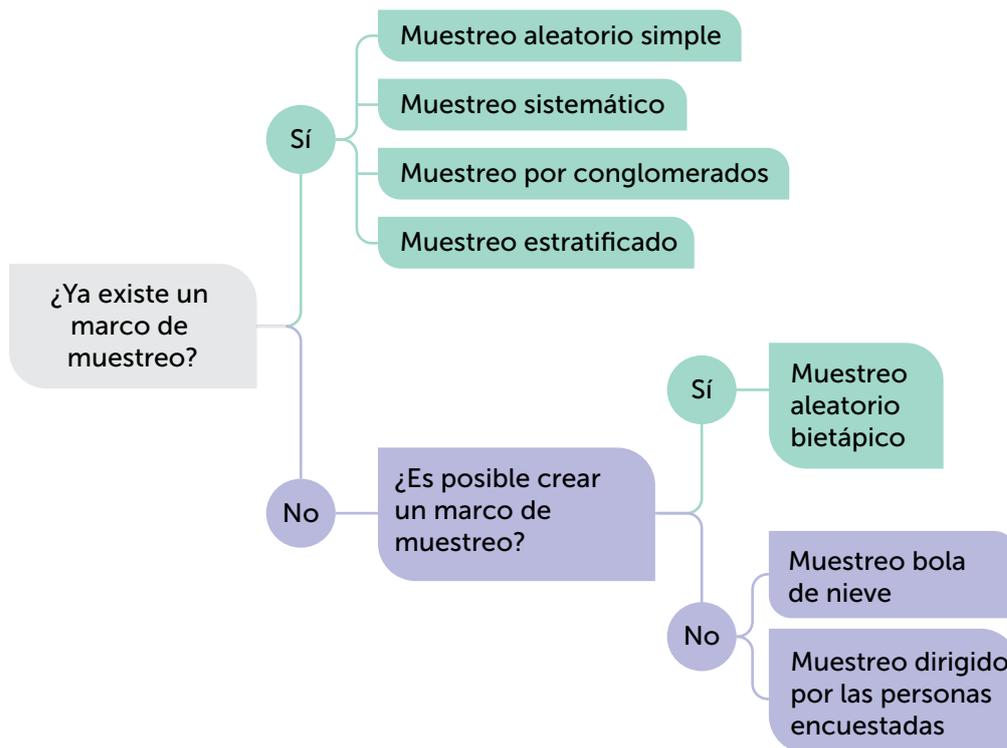
Una de las limitaciones del método de muestreo dirigido por las personas encuestadas es el hecho de que está basado en algunas suposiciones que deben ser verdaderas para que el método pueda generar muestras representativas. Por ejemplo, la suposición de que las personas con redes de contactos más amplias tienden a ser sobremuestreadas. Además, es necesario utilizar herramientas de análisis de datos especializadas para analizar los resultados. Sin embargo, existen herramientas gratuitas disponibles en línea como, por ejemplo, la RDSAT⁷.

⁶ Heckathorn, Douglas D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. <http://www.respondentdrivensampling.org/>

⁷ <http://www.respondentdrivensampling.org/>

¿Cómo decidir cuál es el mejor método de muestreo?

A fin de sintetizar la discusión anterior sobre los distintos enfoques metodológicos que pueden ser utilizados para la selección de muestras, hemos elaborado este sencillo "árbol de decisiones" para orientar a las personas usuarias en su decisión sobre cuál método de muestreo utilizar.



Determinando el tamaño de la muestra: ¿a cuántas personas deberíamos encuestar?

Para determinar el tamaño de la muestra de una encuesta puramente aleatoria, se debe definir, con base en el contexto, el nivel de incertidumbre que se está dispuesto a aceptar. Generalmente, la decisión sobre el nivel aceptable de incertidumbre es influenciada por las implicaciones financieras y por cómo se planea utilizar los resultados. Para un ejercicio de caracterización, un nivel de confianza de 95% con un margen de error de +/- 5% es suficiente para analizar las tendencias generales dentro de una población. El tamaño de la población general también afecta el tamaño de muestra necesario, pero por lo general esto deja de variar significativamente si la población es de 20,000 personas o más.

Las calculadoras de tamaños de muestra en línea (por ejemplo, Raosoft⁸) pueden ayudar a determinar el tamaño de muestra necesario para una encuesta. A continuación, se presenta un ejemplo de los resultados obtenidos para distintos tamaños poblacionales, niveles de confianza y márgenes de error.

Ejemplo

	Margen de error de 5%	Margen de error de 2%	Margen de error de 1%
Para una población de 20,000			
Nivel de confianza de 90%	270	1,560	5,060
Nivel de confianza de 95%	380	2,140	6,490
Nivel de confianza de 99%	640	3,440	9,070
	Margen de error de 5%	Margen de error de 2%	Margen de error de 1%
Para una población de 150,000			
Nivel de confianza de 90%	270	1,670	6,470
Nivel de confianza de 95%	380	2,360	9,030
Nivel de confianza de 99%	660	4,040	14,940

El tamaño de la muestra también debe tomar en cuenta la complejidad del diseño de la metodología de muestreo, descrita por el efecto de diseño. En resumen, si se va a aplicar un método de muestreo bietápico, será necesario aumentar el tamaño de la muestra debido a la complejidad del método. En este caso, se recomienda utilizar un efecto de diseño de 1,8.

La tasa de respuesta también es relevante para el tamaño de la muestra. Se recomienda estimar la tasa de respuesta con base en encuestas previas realizadas en el mismo contexto o en otro tipo de información que permita hacer una suposición fundamentada respecto a la tasa de respuesta esperada. Si no se cuenta con información adicional, normalmente se considera una tasa de respuesta de 50% al calcular el tamaño de la muestra.

Es importante señalar que generalmente las calculadoras de tamaños de muestras se basan en el muestreo aleatorio o sistemático; si se utiliza el muestreo dirigido por las personas encuestadas, el tamaño de la muestra debe ser aumentado aún más de lo que recomienda la calculadora. Calcular qué tan grande debe ser este aumento es difícil y depende mucho de la cantidad de recursos disponibles.

8 <http://www.raosoft.com/samplesize.html>

¿Cuáles son los errores que podrían ocurrir?

Los hallazgos de una encuesta por muestreo nunca serán capaces de reflejar completamente las características reales de una población objetivo: siempre habrá algún nivel de error. Sin embargo, como hemos visto anteriormente, algunos métodos de muestreo suelen generar menos errores que otros. Los errores que se presentan en las encuestas por muestreo pueden ser divididos entre *errores de muestreo* y *errores no muestrales*⁹.

Los *errores de muestreo* son errores que resultan del diseño del proceso de muestreo y ocurren precisamente porque los hallazgos están basados en una muestra, y no en la población objetivo completa. El *error de estimación* es la diferencia entre los valores encontrados en la muestra y los valores reales de la población objetivo (ver la definición de *margen de error* en la sección de términos técnicos). Otro tipo de error de muestreo es el llamado *error de especificación*, que ocurre cuando la probabilidad real de ser incluido en la muestra difiere de la probabilidad que se pretendía tener de acuerdo con el diseño del proceso de muestreo. Nos enfrentaríamos a un error de especificación si, por ejemplo, una misma persona o un mismo hogar hubieran sido enlistados dos o más veces en el marco de muestreo.

Los *errores no muestrales* son errores que ocurren al momento de obtener las respuestas a las preguntas del cuestionario. Los errores no muestrales incluyen situaciones en las que personas que no cumplen con los criterios para ser parte de la población objetivo son incluidas (*error de inclusión*) y situaciones en las que personas que sí cumplen con los criterios para ser parte de la población objetivo quedan fuera (*error de exclusión*). Los equipos de caracterización de situaciones de desplazamiento están constantemente expuestos al riesgo de enfrentarse a errores de inclusión y exclusión, por lo que deben estar especialmente atentos a este riesgo. Otro tipo de error no muestral es el *error de medición*, el cual ocurre cuando la persona encuestada no entiende bien la pregunta o no quiere proporcionar la respuesta correcta, o cuando la persona que realiza la entrevista se equivoca al registrar la respuesta. También se puede generar un error de medición si las personas que realizan la entrevista hacen preguntas tendenciosas o si la persona encuestada tiene dificultad en recordar la respuesta correcta como, por ejemplo, cuando se hacen preguntas sobre periodos anteriores. Un último tipo de error no muestral es la *no respuesta*: cuando personas que fueron seleccionadas para la muestra se rehúsan a contestar el cuestionario por completo o se abstienen de contestar a algunas de sus preguntas.

⁹ Bethlehem, J. (2010). Selection bias in web surveys. https://pure.uva.nl/ws/files/1040458/92124_330481.pdf

La importancia de documentar el proceso

Cómo los resultados de una encuesta por muestreo deberían ser utilizados e interpretados depende mucho del diseño del método de muestreo, como hemos visto anteriormente. Por este motivo, es muy importante ser transparentes con los socios y con el público en general sobre cómo el método de muestreo fue diseñado y cómo la muestra terminó siendo seleccionada al final, así como sobre las limitaciones del método aplicado. La realización de ejercicios de caracterización en situaciones de desplazamiento suele conllevar una serie de acontecimientos inesperados, incluso en lo que se refiere a la selección de las muestras. Esto implica que no siempre será posible aplicar el método de muestreo previsto inicialmente. En estos casos, es importante recordar que lo que se comunique al público debe reflejar el método que realmente fue aplicado a la hora de seleccionar la muestra. Por ejemplo, si se hubiera planeado un método de muestreo bietápico que contemplara la realización de enumeraciones y la selección aleatoria de unidades de muestreo, pero debido a algún evento inesperado se terminó teniendo que implementar el muestreo bola de nieve, el método bola de nieve es el que terminó siendo aplicado y el que debe ser comunicado al público.

Además de la transparencia hacia los socios y el público en general, es altamente recomendable documentar minuciosamente la metodología de muestreo que ha sido diseñada y todos los pasos que se han tomado para obtener la muestra de la encuesta, con el fin de contar con un registro para futuras referencias internas. Esto incluye describir y explicar los desafíos encontrados a lo largo del proceso y cómo estos fueron resueltos. Documentar el proceso de selección de la muestra puede servir tanto para quien lo haga como para sus colegas a la hora de desarrollar una metodología de muestreo para futuros ejercicios de recolección de datos.

¿Dónde se puede aprender más?

El Departamento de Asuntos Económicos y Sociales de la División de Estadística de las Naciones Unidas ha elaborado un manual sobre la selección de muestras para encuestas en hogares llamado "Diseño de Muestras para Encuestas de Hogares: Directrices Prácticas". Este manual está disponible en ruso, francés, árabe, inglés y español:

https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf

El manual sobre "Encuestas de Hogares en los Países en Desarrollo y en Transición" también ofrece una gran cantidad de información útil. Este manual presenta mejores prácticas sobre varios aspectos relacionados con la realización de encuestas de hogares en países en desarrollo y en transición, incluyendo el diseño de la muestra, la implementación de la encuesta, los errores no muestrales, los costos de realización de la encuesta y el análisis de los datos resultantes. La publicación está disponible en árabe, chino, inglés, francés, ruso y español:

https://unstats.un.org/unsd/publication/seriesf/seriesf_96s.pdf

El "Integrated Framework for Household Survey", elaborado por el ACNUR, es un kit de herramientas para facilitar procesos de diseño, recolección y análisis, entre otros aspectos relacionados con el muestreo:

<https://unhcr.github.io/Integrated-framework-household-survey/Sampling.html>

El ACNUR también ofrece orientación a través de la "Standardised Expanded Nutrition Survey (SENS)", la cual puede ser útil para la implementación de encuestas de hogares para PDI y personas refugiadas:

<http://sens.unhcr.org/>

El UNICEF proporciona diversas herramientas con sus "Encuestas de Indicadores Múltiples por Conglomerados (MICS)", incluyendo una herramienta para el mapeo y listado de hogares, información que puede ser útil para quienes consideren llevar a cabo un muestreo bietápico con enumeración:

<http://mics.unicef.org/tools?round=mics5>

Ver también los manuales y materiales de capacitación del programa de "Encuestas de Demografía y Salud (DHS)" de la Agencia de Estados Unidos para el Desarrollo Internacional (USAID):

<http://www.dhsprogram.com/publications/publication-dhsm4-dhs-questionnaires-and-manuals.cfm>

<http://www.dhsprogram.com/What-We-Do/Student-Resources.cfm>

La publicación "Survey Sampling Reference Guidelines" de Eurostat ofrece un buen panorama de los métodos de muestreo probabilístico:

<https://ec.europa.eu/eurostat/documents/3859598/5901961/KS-RA-08-003-EN.PDF/833f7740-0589-47e1-99a5-c14878a2c1a8>

Ver también este sitio web sobre el muestreo dirigido por las personas encuestadas:

<http://www.respondentdrivensampling.org/>



J | JIPS
E | Essential
T | Toolkit

Parte II

Ejemplos de metodologías de muestreo en situaciones de desplazamiento



Introducción y observaciones generales

El objetivo de esta sección es repasar y consolidar distintas metodologías de muestreo utilizadas en algunos de los ejercicios de caracterización apoyados por JIPS en años recientes. El documento ofrece una descripción detallada de los diferentes métodos aplicados, explica por qué estos métodos fueron escogidos e identifica sus posibles limitaciones.

La realidad operacional de los contextos de desplazamiento presenta una serie de limitaciones que afectan la elección de los métodos de muestreo. A veces, esto nos puede obligar a escoger la “segunda mejor opción” entre los métodos de muestreo probabilístico, o incluso tener que optar por un método de muestreo no probabilístico. Sin embargo, si se garantiza la transparencia sobre cómo la muestra fue seleccionada y si se documenta el proceso de manera explícita y minuciosa en la sección metodológica del informe final de la caracterización, estos métodos todavía pueden ofrecer encuestas con resultados útiles.

Este documento busca ofrecer ejemplos de cómo se llevó a cabo el muestreo en distintos ejercicios de caracterización para garantizar que el método empleado fuera “suficientemente bueno” a pesar de las limitaciones impuestas por cada contexto. Específicamente, el documento examina la metodología de muestreo de seis caracterizaciones, realizadas en El Salvador, Grecia, Iraq, Kosovo, Somalia y Sudán. Queremos que este sea un documento vivo y que se puedan incluir más ejemplos a medida que se completen nuevos ejercicios de caracterización. De este modo, esperamos inspirar y promover una comunidad de práctica que valore la transparencia y lidere el camino en materia de manejo responsable de datos.

Entre las principales conclusiones resultantes del análisis de los ejercicios de caracterización mencionados, se encuentran las siguientes:

Es importante documentar la metodología de muestreo y sus limitaciones.

La falta generalizada de documentación de la metodología de muestreo hace que **la utilidad de los resultados de la encuesta sea limitada** para todas aquellas personas que no sean los propios socios de la caracterización. Cuando la metodología y sus limitaciones son compartidas públicamente, no solo es más fácil comprender y utilizar los datos generados, sino que también compartirlos, lo que a su vez permite asegurar un manejo responsable de los datos.

La carencia de documentación y transparencia en torno a las limitaciones de los métodos de muestreo puede resultar en **una interpretación errónea de los resultados** (esto es, sobre en qué medida estos pueden ser generalizados), lo que contribuye a la idea común de que las encuestas por muestreo puramente probabilísticas son la norma, incluso en los contextos de desplazamiento. Sin embargo, todos los estudios empíricos demuestran que esta es más bien la excepción.

Las limitaciones del método de muestreo y los objetivos del ejercicio deben ser considerados en conjunto y su interdependencia debe ser comunicada de forma clara a fin de manejar las expectativas sobre los análisis que podrán ser realizados.

Los objetivos específicos del ejercicio y el análisis que se espera realizar determinan los requisitos que debe cumplir el método de muestreo. Por este motivo, es necesario considerar cuidadosamente cómo formular los objetivos del ejercicio de la forma más concreta posible. En algunas ocasiones, el método de muestreo será limitado por la realidad operacional y lo que es viable en el contexto. En estos casos, es fundamental revisar los objetivos y asegurar que estos sean coherentes con el análisis factible.

Diseñar el método de muestreo requiere una revisión exhaustiva de los datos poblacionales disponibles, definiciones de la población y una consideración minuciosa del tipo de análisis que podrá responder mejor a los objetivos.

Para crear un marco de muestreo, se requiere contar con información sobre el tamaño de la población objetivo y su distribución geográfica. Sin embargo, por distintas razones, en la mayoría de los casos no están disponibles registros actualizados. En la práctica, existen otras fuentes que pueden ser suficientes para la creación de un marco de muestreo. En estos casos, se requieren varios pasos preparatorios antes de la realización de la encuesta (por ejemplo, un muestreo multietápico). Muchas veces será necesario llevar a cabo un ejercicio de enumeración.

Usualmente, para la caracterización de situaciones de desplazamiento se seleccionan muestras de poblaciones desplazadas y no desplazadas para permitir un análisis comparativo. Será necesario establecer definiciones claras, operativas y acordadas para estas dos poblaciones objetivo. Además, el análisis también puede buscar una comparación entre subgrupos, con base en aspectos como la ubicación geográfica, características de alojamiento, etc. Esto requerirá marcos y métodos de muestreo adicionales y específicos. La identificación de los subgrupos requiere una muy buena comprensión del contexto de desplazamiento en cuestión y de los programas o políticas que serán informados a partir de los resultados del ejercicio.

En muchos casos, será necesario utilizar métodos de muestreo diferentes para cada grupo objetivo, dado que la disponibilidad de marcos de muestreo puede variar entre los distintos grupos. Esto puede limitar la posibilidad de hacer comparaciones entre distintos grupos objetivo a partir de los resultados de una encuesta por muestreo.

La realidad operacional de la recolección de datos puede influenciar el diseño final del muestreo; es fundamental que los cambios que ocurran estén reflejados de forma transparente en la descripción final de la metodología de muestreo.

Un desafío común durante la etapa de recolección de datos es la existencia de información personal desactualizada, como direcciones, números de teléfono, etc. Esto afecta la validez del método de muestreo escogido y debe ser tomado en cuenta para determinar el tamaño de muestra necesario. Algunos factores dinámicos de la realidad operacional también pueden influenciar el método de muestreo y el ejercicio en general. Estos factores incluyen la inestabilidad de la situación de seguridad, el acceso a las poblaciones y áreas que serán analizadas, traslados de la población objetivo, grupos de población que no quieren ser identificados, etc. Consecuentemente, muchas veces es necesario ajustar o incluso cambiar completamente el diseño del muestreo. Es fundamental que estos cambios estén reflejados en la descripción original de la metodología de muestreo.

Antes de elegir un método de muestreo, (casi¹) siempre se debe contar con al menos tres elementos: **(1)** los objetivos de la investigación, **(2)** la o las poblaciones objetivo que se desean estudiar y **(3)** un marco de muestreo ya establecido o al menos un plan para obtener uno. Por lo tanto, estos tres elementos aparecen en la descripción de cada uno de los seis ejercicios presentados, antes de entrar al análisis detallado de los métodos de muestreo. Al final de cada ejemplo, también se presentan los principales desafíos y limitaciones del método seleccionado.

En la siguiente sección, presentamos un cuadro de resumen con algunos aspectos clave de la metodología de muestreo utilizada en cada uno de los ejercicios de caracterización. Este cuadro puede servir de referencia para buscar ejemplos con determinadas características. El resto del documento presenta un panorama de la metodología de muestreo de los distintos ejercicios. Para una explicación de los diferentes términos técnicos y métodos de muestreo, consultar la Guía de Muestreo para Situaciones de Desplazamiento².

Caracterización

Un enfoque/sistema colaborativo de recolección de datos que busca **establecer una comprensión común de las situaciones de desplazamiento** y de las circunstancias y características de las personas afectadas. Utiliza **enfoques de métodos mixtos, lo que suele incluir la aplicación de encuestas por muestreo** para recolectar y analizar datos de las poblaciones desplazadas, las comunidades que las acogen y otros grupos, y contextualiza la información obtenida tomando en cuenta el escenario económico, político y social más amplio. El objetivo general es crear **una base de evidencia integral y consensuada para informar intervenciones humanitarias y de desarrollo más eficaces, esfuerzos de incidencia y el desarrollo de políticas nacionales** para apoyar el alcance de soluciones duraderas para las poblaciones desplazadas.

¹ La excepción a esta regla se refiere al uso del muestreo bola de nieve o del muestreo dirigido por las personas encuestadas cuando no existe un marco de muestreo confiable y, por lo tanto, no se cumple con el punto (3).

² Parte I de esta Guía de Muestreo.

Cuadro de resumen

	Población o poblaciones objetivo	Marco de muestreo	Método de muestreo
El Salvador, 2016-2018	Población desplazada internamente en El Salvador entre 2006 y 2016 por motivo de violencia	No estaba disponible una lista completa de la población de PDI. El total de la población de El Salvador sirvió como marco de muestreo	Muestra aleatoria estratificada
	Población de El Salvador que no fue desplazada entre 2006 y 2016 por motivo de violencia	El total de la población de El Salvador sirvió como marco de muestreo	Muestra estratificada bietápica
Tesalónica, Grecia, 2017-2019	Personas refugiadas y solicitantes de asilo que llegaron a Tesalónica después de enero de 2015	Construido a partir de la combinación de dos fuentes distintas que estaban desactualizadas	Censo y muestra bola de nieve (Intento de muestra aleatoria estratificada)
	Personas que no tenían documentos de los servicios de asilo	Lista indisponible, estimación del número total de hogares disponible	Muestra bola de nieve
Erbil, Región del Kurdistán de Iraq, 2015-2016	Personas refugiadas sirias residentes en áreas específicas de la provincia de Erbil	Lista disponible, pero se desconocía si estaba completa	Muestra aleatoria simple
	Población desplazada internamente desde diciembre de 2013 residente en áreas específicas de la provincia de Erbil	Lista disponible, pero se desconocía si estaba completa	Muestra estratificada bietápica
	Comunidad de acogida residente en áreas específicas de la provincia de Erbil	Lista disponible, pero se desconocía si estaba completa	Muestra estratificada bietápica

	Población o poblaciones objetivo	Marco de muestreo	Método de muestreo
Kosovo, 2015-2016	Personas albanesas desplazadas internamente entre enero de 1998 y el final de marzo de 2004, residentes en municipios específicos	Construido a partir de la combinación de fuentes, pero desactualizado	Muestra bola de nieve (Intento de muestra aleatoria estratificada)
	Personas serbias desplazadas internamente entre enero de 1988 y el final de marzo de 2004, residentes en alojamientos privados en municipios específicos	Lista disponible	Muestra aleatoria estratificada
	Personas serbias desplazadas internamente entre enero de 1988 y el final de marzo de 2004, residentes en Centros Colectivos en municipios específicos	Lista disponible	Censo
	Personas Roms/Ashkalíes/Egipcias (RAE) desplazadas internamente entre enero de 1998 y el final de marzo de 2004, residentes en municipios específicos	Compilado con base en dos fuentes distintas, las cuales estaban desactualizadas e incompletas	Censo y muestra bola de nieve
Mogadiscio, Somalia, 2015-2016	PDI residentes en asentamientos informales en Mogadiscio	Lista indisponible, marco creado a través de mapeo y enumeración	Muestra estratificada bietápica
	Migrantes económicos residentes en asentamientos informales en Mogadiscio	Lista indisponible, marco creado a través de mapeo y enumeración	Muestra aleatoria simple
	Comunidades de acogida residentes en asentamientos informales en Mogadiscio	Lista indisponible, marco creado a través de mapeo y enumeración	Muestra aleatoria simple
Sudán, 2017-2019	PDI residentes en dos campamentos: Abu Shouk y El Fasher	Lista indisponible, marco creado a través de mapeo y enumeración	Muestra por conglomerados bietápica
	Población no desplazada residente en la periferia y en la ciudad de El Fasher	Lista indisponible, marco creado a través de mapeo y enumeración	Muestra por conglomerados bietápica

Ejercicios de caracterización y metodologías de muestreo

Ejercicio de caracterización en El Salvador, 2016-2018



Objetivos

Los objetivos de este ejercicio eran identificar el número de familias desplazadas por la violencia en El Salvador en años recientes y comprender mejor las condiciones de vida y necesidades específicas de las víctimas de desplazamiento³. Específicamente, el estudio buscaba

- (1) Obtener estimaciones estadísticas del **número de familias e individuos** que se han tenido que movilizar internamente a causa de la violencia criminal, con estimaciones representativas a nivel nacional;
- (2) Identificar **las causas e impactos específicos** de la movilidad interna causada por la violencia criminal;
- (3) Caracterizar **el perfil y la situación** de la población que se ha tenido que movilizar internamente a causa de la violencia criminal.

Para obtener la información requerida, el estudio adoptó un enfoque de métodos mixtos, utilizando técnicas de recolección de datos cualitativas y cuantitativas, incluyendo una encuesta por muestreo que será el foco del presente análisis.

Población objetivo

Este ejercicio de caracterización tuvo dos poblaciones objetivo:

- (1) La principal población objetivo del estudio fueron **los miembros de los hogares⁴ en donde al menos una de las personas se había desplazado internamente** entre 2006 y 2016 debido a la violencia en El Salvador.
- (2) Además, también se definió otra población objetivo para comparar la situación de los hogares de PDI a la de los hogares de personas no desplazadas. Se determinó que la población objetivo de comparación serían los miembros de los hogares donde ninguna persona se había desplazado internamente entre 2006 y 2016 debido a la violencia en El Salvador.

³ Caracterización de la movilidad interna a causa de la violencia en El Salvador (2018). <https://www.jips.org/uploads/2018/10/Profiling-Report-ElSalvador-ES.pdf>

⁴ Definido en el estudio como una o varias personas que viven juntas en una misma vivienda y comparten gastos para proveer y satisfacer sus necesidades alimenticias.

Marco de muestreo

No había un registro de los hogares o personas desplazadas internamente disponible para los socios de la caracterización. La falta de un registro también fue parte de las razones para la realización de este ejercicio de caracterización, ya que uno de los objetivos del estudio era ofrecer una estimación del número de familias desplazadas por la violencia. Sin embargo, debido a restricciones de tiempo y recursos, no era posible construir una lista completa de todas las familias desplazadas por la violencia. De este modo, fue necesario explorar formas alternativas de crear un marco de muestreo que permitiera seleccionar una muestra representativa y estimar el número de personas desplazadas.

Unidades de muestreo y de análisis



Método de muestreo

En este ejercicio de caracterización, los objetivos consistían tanto en poder ofrecer una estimación del número de hogares desplazados debido a la violencia a nivel nacional (objetivo **(1)**) como en proporcionar información sobre las características y condiciones de las personas que conformaban este grupo (objetivos **(2)** y **(3)**). Debido a que no existían listas de la principal población objetivo y no era posible hacer una enumeración completa de todo el país, los socios de la caracterización tuvieron que diseñar un método de muestreo que les permitiera estimar el total de la población de PDI a nivel nacional a partir de una muestra. Además, para cumplir con los objetivos relacionados con las características y condiciones de los hogares desplazados, los socios de la caracterización necesitaban asegurar que la muestra fuera representativa de la población objetivo.

Para la población objetivo de personas desplazadas, se eligió un método de muestreo aleatorio estratificado y por conglomerados. Para la población objetivo no desplazada, se eligió un método de muestreo bietápico. El proceso de selección de las dos muestras puede ser resumido de la siguiente manera:



5 Área geográficamente definida y establecida con el fin de realizar un censo. Tales áreas pueden ser maneras útiles de dividir geográficamente áreas grandes (por ejemplo, municipios) en áreas más pequeñas para seleccionar una muestra para un ejercicio de caracterización.

6 Una o varias personas que viven juntas en una misma vivienda y comparten gastos para proveer y satisfacer sus necesidades alimenticias.

En el muestreo estratificado, se divide a la población objetivo en diferentes estratos (grupos) y se selecciona una muestra a partir de cada estrato. Un muestreo estratificado no proporcional quiere decir que el tamaño de la muestra tomada de determinado estrato no es proporcional a su tamaño relativo, sino que es determinado con base en otra característica del estrato. En un muestreo por conglomerados, se seleccionan grupos de personas en lugar de personas. En los casos donde no existe un marco de muestreo, un método bastante utilizado para identificar a los miembros de la población objetivo consiste en seleccionar una muestra de conglomerados y hacer una enumeración. A continuación, se presenta una descripción detallada de cómo se llevó a cabo el muestreo.

Paso 1 : Estratificación

Una de las opciones hubiera sido seleccionar aleatoriamente determinado número de municipios de El Salvador para realizar una enumeración. Sin embargo, aunque este método hubiera podido ofrecer información suficiente para estimar el número de hogares de PDI en el país, la muestra de hogares desplazados sería demasiado pequeña para cumplir con los demás objetivos mencionados anteriormente, es decir, comprender mejor la situación y las necesidades específicas de los hogares desplazados. Para compensar esto, se introdujeron estratos para que fuera posible sobremuestrear en las áreas donde las fuentes indicaban que probablemente habría mayor proporción de hogares desplazados.

Al principio, los socios de la caracterización reunieron información sobre la distribución de los hogares de PDI a lo largo del país, con el fin de informar el proceso de estratificación. A través de una revisión de las fuentes existentes, se constató que había 20 municipios de El Salvador donde se esperaba encontrar un mayor número de hogares desplazados por la violencia que en otros municipios. Las fuentes analizadas consistían en estadísticas demográficas y migratorias, así como estadísticas sobre la incidencia de violencia a nivel municipal.

A partir de una lista de los segmentos censales⁷ de cada municipio del país, compartida con los socios de la caracterización por la Dirección General de Estadísticas y Censos de El Salvador, fue posible realizar una subdivisión adicional de los municipios. Los segmentos censales funcionaron como conglomerados y unidades primarias de muestreo. Se llevaron a cabo entrevistas y talleres participativos con informantes clave en los 20 municipios priorizados, a partir de los cuales se pudo generar una lista de 648 segmentos censales —de un total de 4,427— donde se esperaba encontrar una mayor concentración de hogares de PDI.

Con base en esta información, los segmentos censales del país fueron divididos en tres estratos:

- (1) 648 segmentos censales prioritarios en los 20 municipios priorizados
- (2) 3,779 segmentos censales restantes en los 20 municipios priorizados
- (3) 7,996 segmentos censales restantes en el país

Paso 2 : Determinación del tamaño de la muestra

Se llevó a cabo una selección aleatoria de segmentos censales en cada uno de los tres estratos. El tamaño de la muestra de segmentos censales fue determinado con base en criterios técnicos de precisión (95% de confianza) para estimaciones de baja

⁷ Áreas geográficas más pequeñas que los municipios, establecidas por la Dirección General de Estadísticas y Censos de El Salvador.

incidencia⁸, suponiendo una tasa de no respuesta elevada⁹ y tomando en cuenta el efecto de diseño significativo¹⁰ causado por la estratificación. Se determinó que el tamaño de la muestra sería de 542 segmentos censales, lo equivalente a una meta estimada de 40,650 hogares.

Paso 3 : Selección de una muestra de conglomerados (segmentos censales)

La muestra fue distribuida de manera no proporcional entre los estratos para garantizar que habría una cantidad mínima de población objetivo en la muestra. A los segmentos censales de los dos estratos prioritarios se les atribuyó una mayor probabilidad de ser seleccionados para la muestra que a los segmentos censales del estrato restante. Luego, se seleccionaron aleatoriamente segmentos de cada estrato. La siguiente tabla presenta la distribución de los segmentos censales seleccionados en los tres estratos.

Estrato	# de segmentos en El Salvador	Distribución de segmentos en El Salvador	# de segmentos en la muestra	Distribución de segmentos en la muestra
1	648	5%	184	34%
2	3,779	30%	125	23%
3	7,996	64%	233	43%
Total	12,423	100%	542	100%

Debido a la distribución no proporcional de los segmentos censales entre los estratos, durante el análisis de los resultados, cuando los hallazgos eran considerados a nivel nacional, era necesario introducir factores de ponderación para compensar la representación desigual de los segmentos en la muestra. Estos factores de ponderación se denominan ponderaciones de muestreo o, alternativamente, ponderaciones de probabilidad inversa, y generalmente se calculan considerando la cantidad de hogares que cada hogar de la muestra representa (ponderación de muestreo = número total de hogares en un estrato/número de hogares en la muestra del estrato). Para obtener el número de hogares de cada estrato, se puede dividir la población total del estrato por el tamaño promedio de los hogares. Aunque en este ejercicio de caracterización los socios no documentaron cómo los factores de ponderación fueron calculados, se puede suponer que se aplicó este método bastante común de calcular ponderaciones de muestreo.

Paso 4 : Enumeración de los conglomerados (segmentos censales) seleccionados

Después de seleccionar la muestra de 542 segmentos censales, se llevó a cabo una enumeración de todos los hogares de 501 de los 542 segmentos seleccionados. Las y los enumeradores no pudieron acceder a los 41 segmentos censales restantes debido a las condiciones de seguridad. También se registraron casos de no respuesta, aunque la tasa fue menor a la esperada (37%). Los hogares enumerados eran clasificados como desplazados o no desplazados. A partir de la cantidad de hogares desplazados

⁸ En un estudio similar realizado en Honduras, la proporción promedio de hogares de PDI fue de 4%.

⁹ La tasa de no respuesta en el estudio de Honduras ascendió a cerca del 50%.

¹⁰ Implica un aumento de la incertidumbre con relación a los hallazgos debido a la complejidad del método de muestreo.

identificados en cada segmento, fue posible extrapolar los números a nivel nacional¹¹ para cumplir con el objetivo (1). Los resultados de la enumeración también permitieron conformar una muestra de hogares de PDI, así como un marco de muestreo para la selección de hogares de personas no desplazadas para la encuesta por muestreo que abordaría los dos objetivos restantes.

Paso 5 : Muestra de hogares de PDI resultante

La enumeración resultó en la identificación de 466 hogares desplazados por la violencia, los cuales estaban conformados por 1,811 personas. Todos estos hogares fueron incluidos en la muestra y se les solicitó que participaran en la encuesta por muestreo. 431 de estos hogares, lo equivalente a 1,668 personas, aceptaron participar y completaron la encuesta.

Paso 6 : Selección de la muestra de hogares no desplazados

A partir de los segmentos censales seleccionados, también se tomó una muestra de la población objetivo de comparación, a través de un muestreo aleatorio simple. Para cada tres encuestas aplicadas a un hogar desplazado, se aplicaba una encuesta a un hogar no desplazado seleccionado aleatoriamente en el mismo segmento censal, con base en la lista resultante de la enumeración. Debido a que la selección de la muestra de los hogares no desplazados fue realizada en dos etapas (es decir, primero se tomó una muestra de segmentos censales y luego se tomó una muestra de hogares), este método puede ser denominado muestreo “bietápico”.

Desafíos y limitaciones

El ejercicio de caracterización realizado en El Salvador ilustra un escenario que muchas veces predomina en los contextos de desplazamiento: la falta de un registro de la población objetivo que pueda servir como marco de muestreo. La falta de un marco tiene implicaciones sobre el diseño del método de muestreo y suele generar mayor complejidad, lo que a su vez puede aumentar la incertidumbre con relación a los hallazgos. Esto pasó en el ejercicio realizado en El Salvador, donde los socios de la caracterización decidieron usar un muestreo por conglomerados estratificados y un muestreo bietápico.

Otro desafío común al realizar encuestas por muestreo es la **no respuesta**, esto es, cuando no se logra establecer contacto con personas que han sido seleccionadas para ser encuestadas o cuando algunas de las personas seleccionadas deciden no participar en la encuesta. Como se mencionó anteriormente, este ejercicio registró una tasa de no respuesta considerable durante el proceso de enumeración, aunque esta fue inferior a la esperada. Asimismo, los socios de la caracterización se enfrentaron a restricciones de seguridad que no les permitieron enumerar todos los segmentos censales de la muestra. Estos dos factores pueden haber generado sesgos en los resultados.

Finalmente, la **distribución de género** no fue considerada durante la aplicación de la encuesta en los hogares desplazados y no desplazados seleccionados. Esto también puede haber causado sesgos en los resultados.

¹¹ Para una explicación de cómo se llevó a cabo la extrapolación, ver el Anexo 1 del informe de la caracterización: <https://www.jips.org/uploads/2018/10/Profiling-Report-ElSalvador-ES.pdf>

Ejercicio de caracterización en Tesalónica, Grecia, 2017-2019



Objetivos

Este ejercicio de caracterización tenía la finalidad de ofrecer al municipio de Tesalónica y a los demás actores pertinentes una base de evidencia capaz de informar el diseño de iniciativas de incidencia, políticas, sistemas de prestación de servicios y programas de integración que se adaptaran mejor a las necesidades de las personas afectadas por el desplazamiento en el área metropolitana de Tesalónica.

Los objetivos específicos que se acordaron para este ejercicio fueron los siguientes:

- (1) Elaborar un **perfil demográfico** de las personas refugiadas, solicitantes de asilo y llegadas espontáneas, desagregado por sexo, edad y otros criterios de diversidad pertinentes;
- (2) **Analizar las capacidades, vulnerabilidades y mecanismos de supervivencia** de las poblaciones objetivo;
- (3) **Evaluar el nivel de integración** de las poblaciones objetivo, con foco en el acceso a servicios;
- (4) Producir un conjunto de indicadores para medir el nivel de integración.

Población objetivo

Este ejercicio de caracterización tenía dos poblaciones objetivo principales:

- (1) **Personas refugiadas y solicitantes de asilo que llegaron a Tesalónica después de enero de 2015** residentes en viviendas proporcionadas por el programa de alojamiento ESTIA¹², en el centro de recepción de Diavata¹³ o en alojamiento independiente. Esta población objetivo incluye a las personas que poseían documentos que demostraran que contaban con un prerregistro de solicitante de asilo, un registro de solicitante de asilo completo, una resolución de reconocimiento de la condición de persona refugiada, un permiso de residencia o una solicitud de asilo pendiente de recurso administrativo.

¹² El programa de Ayuda de Emergencia para la Integración y el Alojamiento (ESTIA) proporciona alojamiento para personas refugiadas y solicitantes de asilo en apartamentos en áreas urbanas. Este programa es gestionado por el ACNUR en coordinación con 12 organizaciones socias en el norte y centro de Grecia.

¹³ Instalaciones de alojamiento provisional creadas para acoger a personas refugiadas y solicitantes de asilo que quedaron atrapadas en el lado griego de la frontera entre Grecia y Macedonia del Norte tras el cierre de dicha frontera a principios de 2016. Estas instalaciones no eran adecuadas para alojamiento a largo plazo.

(2) Nacionales de terceros países sin registro del Servicio de Asilo Griego. Esto incluye a las personas que contaban con notas de la policía, personas cuya documentación que recibieron del Estado griego había expirado y personas que no habían recibido documentación del Estado griego. El estudio incluyó a las personas que residían en el centro de recepción de Diavata de manera no oficial, personas que habían sido acogidas por individuos en Tesalónica, así como personas en situación de calle.

Específicamente, el alcance geográfico de la caracterización incluía los seis distritos del municipio de Tesalónica y los municipios vecinos de Kalamaria, Neapolis-Sikeon, Pavlou Mela, Kordeliu-Evosmos, Ampelokipon-Menemenis, Delta, Thermis y Oreokastro. Las personas residentes en el centro de recepción de Diavata también fueron incluidas en el ejercicio.

Las personas refugiadas y solicitantes de asilo que no contaban con una dirección de domicilio válida no fueron incluidas en la población objetivo del estudio debido a la falta de información disponible sobre este grupo. Asimismo, las personas refugiadas y solicitantes de asilo de nacionalidad turca fueron excluidas de la encuesta por muestreo porque no se pudo acceder a información exacta sobre sus direcciones. Sin embargo, se realizaron discusiones de grupos focales con la población turca.

Marco de muestreo

Personas refugiadas y solicitantes de asilo: se construyó un marco de muestreo mediante la combinación de dos fuentes distintas, siendo estas la base de datos ProGres de ACNUR (una lista completa de las personas beneficiadas por el programa de alojamiento urbano ESTIA y de las personas que contaban con alojamiento independiente, incluyendo su información de contacto) y una lista de todas las personas residentes en el centro de recepción de Diavata.

Nacionales de terceros países sin registro del Servicio de Asilo Griego: ante la falta de registros o listas completas de la población de interés, los socios de la caracterización tuvieron que buscar otras maneras de diseñar un método de muestreo adecuado. Distintos proveedores de servicios locales contaban con información sobre esta población objetivo, pero el equipo de la caracterización no pudo obtener acceso a los datos debido a consideraciones en materia de protección. Sin embargo, a partir de la comparación de la información agregada de diversos proveedores de servicios, se pudo estimar que había alrededor de 200 hogares pertenecientes a esta población objetivo.

Unidades de muestreo y de análisis



Método de muestreo

Los objetivos del ejercicio –comprender mejor la situación y las características de las poblaciones objetivo– implicaban que los resultados de la encuesta por muestreo deberían, idealmente, representar a ambas poblaciones objetivo. Por lo tanto, un muestreo basado en probabilidad sería el método apropiado para garantizar una muestra representativa con un nivel de confianza y un margen de error conocidos. Sin embargo, esto no era posible dado que, como se mencionó anteriormente, los socios de la caracterización solo contaban con un marco de muestreo para una de las dos poblaciones objetivo.

La existencia de una lista completa de la población objetivo de personas refugiadas y solicitantes de asilo permitía, en teoría, extraer una muestra basada en probabilidad, mientras que la falta de un registro de las personas que no contaban con documentos de los servicios de asilo implicaba el uso de una estrategia de muestreo no probabilística que, en este caso, involucró un muestreo bola de nieve. Aunque este no era un escenario ideal, y significaba que los resultados de las encuestas por muestreo de los dos grupos de población no podrían ser comparados directamente, el equipo de la caracterización no tenía otras alternativas factibles. A pesar de que no se podría evaluar el nivel de representatividad y que los hallazgos no podrían ser generalizados para el total de la población objetivo con certeza¹⁴, este método todavía podría proporcionar datos útiles acerca de un grupo de población sobre el cual se contaba con muy poca información.

A continuación, se presentan descripciones más detalladas del método de muestreo elegido para cada una de las dos poblaciones objetivo.

Personas refugiadas y solicitantes de asilo

Inicialmente, la metodología elegida para esta población objetivo había sido el muestreo aleatorio estratificado. Sin embargo, durante la ejecución del ejercicio el método tuvo que ser modificado debido a motivos que serán explicados a fondo a continuación. El método que se había planeado utilizar inicialmente puede ser resumido de la siguiente manera:



Paso 1 : Estratificación

La población objetivo fue dividida en tres estratos con base en los distintos tipos de alojamiento. Los tres estratos fueron definidos de esta manera porque se esperaba que la situación de las personas en cada tipo de alojamiento presentara diferencias significativas debido a las características distintas de sus viviendas. De este modo, era recomendable comparar la situación de las personas de acuerdo con su tipo de alojamiento y, por lo tanto, era necesario asegurar que se tomara una muestra suficiente de cada uno de los tres tipos de alojamiento para que fuera posible compilar resultados representativos para cada grupo.

La población objetivo fue agrupada en los siguientes estratos:

- (1) Personas refugiadas y solicitantes de asilo que residían en apartamentos proporcionados por el programa de alojamiento urbano ESTIA (1,280 hogares).
- (2) Personas refugiadas y solicitantes de asilo residentes en alojamiento independiente en Tesalónica, esto es, personas que alquilaban apartamentos o habían sido acogidas por amigos, familiares o voluntarios (482 hogares).
- (3) Personas refugiadas y solicitantes de asilo registradas oficialmente como residentes del centro de recepción de Diavata (148 hogares).

Paso 2 : Determinación del tamaño de la muestra

Para el estrato (3), se decidió hacer un conteo completo de todos los hogares, debido al número relativamente reducido de hogares del centro de recepción. Para los estratos (1) y (2), el tamaño de la muestra fue calculado para un nivel de confianza de 95% y

¹⁴ Excepto si se alcanzara el punto de saturación, es decir, el punto donde todas las personas de la población objetivo han sido identificadas y ya no hay nuevas referencias.

un margen de error deseado de 5%. Asimismo, previendo una tasa de no respuesta de 20%, se ajustó el tamaño de la muestra de cada estrato para reflejar el tamaño reducido de la población objetivo. Se determinó un tamaño de muestra de 230 hogares para el estrato (1) y de 190 hogares para el estrato (2).

Paso 3 : Selección de la muestra

Como se mencionó anteriormente, todas las personas del estrato (3) fueron seleccionadas para la muestra. Se distribuyeron cuestionarios a todos los hogares en el centro de recepción y nadie se rehusó a contestar la encuesta, de modo que la tasa de respuesta de este estrato fue de 100%. Sin embargo, es importante señalar que la población de este estrato era relativamente pequeña en comparación con la de los demás estratos y que, para algunas variables, el número de observaciones fue muy bajo, lo que hizo que el análisis de estos resultados fuera más incierto.

Después de que se determinó el tamaño de las muestras de los estratos (1) y (2), se tomó una muestra aleatoria simple de cada uno de estos estratos. La probabilidad de inclusión en la encuesta fue distinta entre los estratos: en el estrato (1), 18% de la población objetivo fue incluida en la muestra; en el estrato (2), 39% de la población objetivo fue incluida y en el estrato (3) 100% de la población objetivo fue incluida. Al estimar una media poblacional para toda la muestra, esta representación desigual tendría que ser compensada a través de factores de ponderación. Estos factores, también denominados ponderaciones de probabilidad inversa, equivalen al número de hogares representados por cada hogar seleccionado para la muestra. Los factores de ponderación solo serían utilizados al calcular cifras referentes a los tres estratos (por ejemplo, al estimar el ingreso promedio de toda la población objetivo) y no eran necesarios al estimar un promedio referente a solo uno de los estratos (por ejemplo, el ingreso promedio de la población en el centro de recepción de Diavata).

Al intentar contactar por teléfono a las personas seleccionadas para la muestra en los estratos (1) y (2), pronto se hizo evidente que los números de teléfono de la mayoría de las personas de la lista estaban desactualizados a causa de cambios frecuentes de sus tarjetas SIM. Debido a que se contaba con una cantidad limitada de recursos, no era posible visitar a estos hogares sin una confirmación previa, de modo que estas personas no podrían ser contactadas. Esto significaba que el equipo de la caracterización tendría una tasa de no respuesta demasiado elevada para permitir que los resultados de la encuesta fueran representativos. Por lo tanto, se decidió ya no utilizar una muestra seleccionada aleatoriamente para estos dos estratos sino, en cambio, entrevistar a todas las personas de estos estratos con las que se lograra establecer contacto, esperando así obtener una tasa de respuesta aceptable.

Se realizó un censo en los dos estratos, combinado con un muestreo bola de nieve, donde las personas encuestadas referían a otros hogares pertenecientes a la población objetivo para que fueran incluidos en el ejercicio. La tasa de respuesta en el estrato (1) fue de 34,2% (de un total de 1,280 hogares) y en el estrato (2) fue de 41,2% (de un total de 482 hogares).

Paso 4 : Análisis de representatividad y ponderación

posestratificación

Las tasas de no respuesta relativamente altas registradas en los estratos (1) y (2) generaron un riesgo de sesgo en los resultados, ya que es posible que las personas que respondieron el cuestionario no reflejen adecuadamente las características de toda la población objetivo. Por este motivo, algunas características demográficas clave (división por edad, nacionalidad y género) de las personas encuestadas fueron analizadas y comparadas con la información del total de la población objetivo (que estaba disponible en la base de datos ProGres) para evaluar la representatividad de

las personas que respondieron el cuestionario. La distribución de las características de las personas encuestadas se asemejaba bastante a la de la población objetivo. Con base en esto, se concluyó que se podría considerar que los resultados de la encuesta eran representativos.

Realizar un censo suponía que no era necesario aplicar factores de ponderación. Sin embargo, para compensar las tasas de respuesta distintas de los tres estratos, se calcularon factores de ponderación posestratificación, los cuales eran aplicados cuando las personas encuestadas de los tres estratos eran analizadas en conjunto, es decir, cuando la variable de interés no era desglosada por estratos.

Personas sin documentación de los servicios de asilo

Para esta población objetivo, se realizó un muestreo no probabilístico a través del método bola de nieve. El muestreo bola de nieve consiste en seleccionar a algunas personas de la población objetivo para ser entrevistadas y solicitarles que refieran a otros miembros de dicha población objetivo para que participen en el estudio. Las nuevas referencias de posibles personas a ser encuestadas se reciben en "olas" hasta que se alcance el número deseado de personas para la muestra. El método utilizado puede ser resumido de la siguiente manera:



Paso 1 : Identificación de personas para ser entrevistadas

Algunos proveedores de servicios locales que ofrecían asistencia a personas que no poseían documentos de los servicios de asilo fueron involucrados en el ejercicio para contactar a posibles participantes y hacer las citas para las entrevistas. Además, se solicitaba a las personas contactadas que refirieran a otros miembros de la población objetivo. Se había estimado que el número aproximado de hogares en la población objetivo era de 200 y, por lo tanto, la meta era llegar a este número de hogares y así alcanzar el punto de saturación o, en otras palabras, lograr un conteo completo de la población objetivo.

Paso 2 : Realización de las entrevistas

El número de hogares entrevistados fue superior al que se había previsto originalmente. El equipo entrevistó a 451 personas que conformaban 227 hogares. La muestra contemplaba distintas situaciones de alojamiento: 222 personas no tenían un hogar (vivían en situación de calle, en edificios inacabados o abandonados, o en albergues para personas sin hogar), 161 personas vivían de manera no oficial en el centro de recepción de Diavata y 27 personas vivían en alojamientos alquilados en Tesalónica.

Algo que podría explicar el gran número de personas entrevistadas es el flujo de llegadas a Tesalónica que ocurrió en el mismo mes en que se realizaron las entrevistas. Los resultados de la encuesta corroboraron esta teoría, ya que al momento de la entrevista más de la mitad de las personas encuestadas habían estado en Tesalónica por menos de un mes. Este número elevado de llegadas recientes generó una mayor incertidumbre con relación a la estimación del tamaño total de la población objetivo. Además, muchas de las personas que fueron contactadas decidieron no conceder una entrevista. Consecuentemente, no fue posible evaluar en qué medida las personas entrevistadas eran representativas de la población objetivo.

Desafíos y limitaciones

El equipo de la caracterización se enfrentó a diversos desafíos que limitaron las posibilidades de aplicar el método de muestreo preferido para la realización de la encuesta, tanto en la etapa inicial de planificación como durante la ejecución del ejercicio.

Para la población objetivo de personas solicitantes de asilo y refugiadas, los obstáculos al contactar a los miembros de la muestra que se había tomado impidieron la realización de una encuesta basada en una muestra seleccionada aleatoriamente. La elevada **tasa de no respuesta** de la población objetivo que finalmente se logró incluir generó incertidumbre con relación a los resultados. Sin embargo, el análisis de la representatividad de la muestra demostró que las personas encuestadas parecían ser representativas de la población objetivo en algunas características clave, por lo que se presumió que los hallazgos eran representativos y podrían ser extrapolados para el total de la población objetivo. No obstante, se debe señalar que existen otras características que no fueron evaluadas y podrían representar diferencias entre las personas encuestadas y la población objetivo, lo que generaría más incertidumbre sobre la representatividad.

Para la población objetivo de nacionales de terceros países no registrados por el Servicio de Asilo, la **falta de un marco de muestreo** ocasionó la utilización de una muestra no probabilística. Durante la etapa de recolección de datos, **un gran número de miembros de la población objetivo decidió no participar** en la encuesta, lo que generó más complicaciones. Estos factores, en conjunto, significaban que los resultados eran inciertos y que no se podría afirmar que los hallazgos eran representativos de toda la población objetivo.

Ejercicio de caracterización en Erbil, Región del Kurdistán de Iraq, 2015-2016



Objetivos

La provincia de Erbil, donde vive un total de 2.01 millones de personas (cifras de 2014), alberga la capital de la Región del Kurdistán Iraquí. Al momento de la realización del ejercicio de caracterización, la provincia —como el resto de la región del Kurdistán— se encontraba profundamente afectada por oleadas de desplazamiento ocasionadas por los conflictos en Siria y en el resto de Iraq, y por una crisis financiera generalizada que afectaba a los sectores público y privado de la economía.

Si bien existía una cantidad significativa de información disponible sobre las PDI y las personas refugiadas que vivían en campamentos, se tenía menos información sobre aquellas que residían fuera de esos lugares. Además, hasta ese momento la mayoría de las estrategias para atenuar los efectos del desplazamiento se había enfocado en atender a las necesidades de las poblaciones desplazadas y refugiadas, mientras que las necesidades de las comunidades de acogida no recibían la misma atención.

Los objetivos específicos del estudio fueron los siguientes:

- (1) Elaborar **perfiles demográficos** desagregados por género, edad y condición de desplazamiento (es decir, personas refugiadas, PDI y comunidades de acogida) en las zonas de interés;
- (2) Elaborar perfiles en las zonas urbanas con alta concentración de poblaciones desplazadas fuera de campamentos;
- (3) **Analizar las capacidades, vulnerabilidades y mecanismos de supervivencia** de la población de estas zonas;
- (4) **Analizar las relaciones** entre las poblaciones desplazadas y las poblaciones que las acogen;
- (5) Analizar la resiliencia en las zonas urbanas en relación con la disponibilidad y las limitaciones de los servicios;
- (6) Generar un conjunto de datos disponible para el Gobierno Regional del Kurdistán y para la comunidad humanitaria/de desarrollo.

Los objetivos (1), (3) y (4) serían alcanzados a través de la aplicación de una encuesta por muestreo en hogares.

Población objetivo

Se definieron tres poblaciones objetivo:

- (1) Personas refugiadas sirias;

- (2) Población desplazada internamente desde diciembre de 2013;
- (3) Comunidad de acogida (es decir, personas que no pertenecen a las categorías mencionadas anteriormente, incluyendo a migrantes económicos).

Todas residentes en las siguientes áreas geográficas:

- Centro del distrito de Erbil, conformado por las ciudades de Hawler (Nawandy Hawler) y Ainkawa;
- Periferia del distrito de Erbil, conformada por los núcleos urbanos ubicados en el entorno inmediato del centro del distrito de Erbil y directamente conectados a él: Baharka, Bnaslawa (Dashty Hawler Centre), Daratu, Kasnazan, Khabat Centre (Nawandy Khabat) y Rizgari;
- Centros urbanos al interior de la provincia: Harir, Koya Centre, Shaqlawa Centre y Soran Centre (Nawandy Soran), en conjunto con Diana.

Marco de muestreo

Se establecieron tres marcos de muestreo distintos a partir de diferentes fuentes secundarias, uno para cada población objetivo:

- (1) Personas refugiadas sirias: se utilizaron datos de ProGres, la base de datos de ACNUR para el registro de personas refugiadas;
- (2) Población desplazada internamente desde diciembre de 2013: se utilizaron datos de la primera etapa del Registro Integral de Personas Desplazadas (CRDP), realizado por la Oficina de Estadística de la Región del Kurdistán (KRSO) en junio de 2015;
- (3) Comunidad de acogida: se utilizaron datos de 2009 que habían sido recolectados durante el proceso de preparación para un censo que fue planeado pero no se llevó a cabo.

Unidades de muestreo y de análisis



Método de muestreo

Los objetivos del ejercicio implicaban describir las características y la situación de las poblaciones objetivo, así como analizar las relaciones entre los distintos grupos. Para atender estos objetivos sería necesario seleccionar muestras que fueran representativas de las poblaciones objetivo y, por lo tanto, un método de muestreo probabilístico era la opción preferida. Los socios de la caracterización también querían poder comparar los resultados entre las áreas geográficas, así como entre las distintas poblaciones objetivo de cada área. Para esto, el equipo de la caracterización tendría que diseñar un método que permitiera extraer muestras suficientemente grandes de las tres poblaciones objetivo en cada área geográfica. Se diseñó un proceso de muestreo específico para cada una de las tres poblaciones objetivo. La Oficina de Estadística de la Región del Kurdistán diseñó la metodología de muestreo (en parte en colaboración con JIPS), seleccionó las muestras y realizó la recolección de los datos. A continuación, se presenta una

breve descripción de los métodos utilizados para cada población objetivo.

Personas refugiadas sirias

Se seleccionó una muestra aleatoria simple a partir del marco de muestreo. En cierta medida, el marco de muestreo estaba desactualizado y presentaba deficiencias, pero no había posibilidad de actualizarlo. No obstante, sí fue posible realizar entrevistas con los hogares seleccionados para la muestra. No se cuenta con información sobre la tasa de respuesta. Si la tasa de no respuesta fuera elevada, podría existir sesgos en los resultados, lo que disminuiría la representatividad de la muestra.

PDI

Se utilizó un método de muestreo estratificado bietápico para seleccionar la muestra de la población desplazada internamente.

Etapa 1: Se seleccionaron los distritos con la mayor concentración de PDI de los tres estratos. La Oficina de Estadística de la Región del Kurdistan contaba con información sobre qué distritos tenían las concentraciones más altas de PDI debido a la primera fase del Registro Integral de Personas Desplazadas. La distribución de la muestra entre los diferentes distritos debería ser proporcional al tamaño de la población de PDI en cada distrito. Se escogió y enumeró un número determinado de barrios en cada uno de los distritos. JIPS desconoce si los barrios fueron seleccionados aleatoriamente o con base en la concentración de PDI.

Etapa 2: Se seleccionó una muestra de hogares de PDI en cada barrio, con base en la información resultante de la enumeración.

Comunidad de acogida

El mismo método de muestreo estratificado bietápico utilizado para las PDI fue aplicado para seleccionar la muestra de las comunidades de acogida, con la excepción de que se usó una unidad primaria de muestreo distinta en la primera etapa: manzanas, en lugar de distritos. JIPS no cuenta con información sobre por qué se decidió utilizar una unidad de muestreo diferente y cómo las manzanas fueron seleccionadas (de manera aleatoria o no).

Desafíos y limitaciones

La metodología de muestreo fue documentada de manera insuficiente en el informe de la caracterización. Esto limita la utilidad de los resultados de la encuesta para cualquiera que no haya formado parte del equipo que realizó el ejercicio.

La muestra de personas refugiadas sirias fue seleccionada a partir de un marco de muestreo desactualizado, de modo que es difícil determinar en qué medida la información obtenida a través de los hogares seleccionados es representativa del total de la población de cada estrato.

Asimismo, debido a que el criterio para seleccionar los distritos fue escoger a aquellos con mayor concentración de población desplazada internamente, no sería posible capturar información sobre las PDI y las comunidades de acogida de los distritos con menor concentración de población desplazada. En este sentido, cabe señalar que la situación de los hogares podría ser distinta entre los distritos con mayor y menor concentración de PDI. Además, si la población desplazada de los distritos con menor concentración de PDI representa una parte del total de la población desplazada del estrato que no puede dejar de ser considerada, la muestra seleccionada no sería representativa de las poblaciones de PDI y de las comunidades de acogida de esos distritos, sino que solo de aquellas de los distritos con alta concentración de PDI.

Ejercicio de caracterización en Kosovo, 2015-2016



Objetivos

La población forzada a desplazarse debido al conflicto de 1998/99 y a los disturbios de 2004 en Kosovo se ha enfrentado a una situación de desplazamiento prolongado. Al momento del ejercicio, las estimaciones indicaban que la gran mayoría de las personas desplazadas al interior de Kosovo residía en alojamientos privados, mientras que una pequeña parte residía en centros colectivos. El ejercicio de caracterización buscaba informar una política en materia de soluciones duraderas —incluyendo la estrategia de soluciones duraderas del Ministerio para las Comunidades y el Retorno de Kosovo— al ofrecer un análisis basado en evidencia de la situación de desplazamiento, de acuerdo con el Marco IASC de Soluciones Duraderas para las PDI.

Los objetivos específicos fueron los siguientes:

- (1) Elaborar un **perfil demográfico** de la población desplazada internamente en Kosovo, desagregado por edad, sexo, ubicación y diversidad;
- (2) Realizar un **análisis integral de la situación de desplazamiento** con relación a la integración social y económica de las personas desplazadas, los desafíos y vulnerabilidades vinculadas con el desplazamiento, el disfrute de derechos, los recursos y capacidades de las PDI, así como sus intenciones y planes futuros;
- (3) Mejorar **la capacidad de las instituciones de promover y diseñar una programación conjunta en favor de soluciones duraderas** para las PDI a través de la identificación de las prioridades específicas de estos grupos y de una mejor coordinación del análisis humanitario y de desarrollo.

Población objetivo

El estudio incluyó a cuatro poblaciones objetivo principales, las cuales habían sido obligadas a abandonar sus lugares de residencia permanente y desplazarse internamente en Kosovo entre enero de 1998 y el final de marzo de 2004, en particular como resultado o para evitar los efectos de un conflicto armado, de la situación de violencia generalizada y de violaciones de los derechos humanos. El alcance geográfico de las poblaciones objetivo del estudio abarcaba los municipios con la concentración más alta de personas desplazadas. Las personas de otros orígenes étnicos no fueron incluidas en el estudio.

- (1) **PDI albanesas** en Prishtinë/Pristina, en Mitrovica Sur y en Vushtrri/Vucitrn;
- (2) **PDI serbias en alojamientos privados** en Zveqan/Zvecan, Leposaviq/Leposavić, Mitrovica Norte, Zubin Potok, Graçanicë/Graçanica, Fushë Kosovë/Kosovo Polje, Lipjan/Lipljane, Shtërpçë/Strpce, Gjiilan/Gnjilane, Kamenicë/Kamenica, Viti/Vitina;

- (3) **PDI serbias en centros colectivos** (todos los centros colectivos de Kosovo) y
- (4) **PDI Roms/Ashkalíes/Egipcias (RAE)** en Fush K./K. Polje, Gracanice/Gracanica, Ferizaj/Urosevac, Pejë/Peć, Gjakovë/Dakova y en los demás municipios donde se pudo identificar a personas RAE desplazadas internamente.

También se estableció un grupo de comparación con el fin de comparar los hallazgos referentes a las PDI con la situación de las poblaciones no desplazadas de Kosovo. Se decidió que el grupo de comparación sería la población general de Kosovo, y que se utilizarían datos secundarios para obtener la información pertinente sobre esta población objetivo.

Marco de muestreo

La disponibilidad de una lista o registro único y actualizado variaba entre los cuatro estratos. En algunos casos, los marcos de muestreo fueron compilados a partir de diversas fuentes. A continuación, se presentan las distintas fuentes utilizadas para cada una de las poblaciones objetivo.

PDI albanesas: se generó una lista de 1,167 hogares y 5,879 individuos. La fuente fue un censo realizado en 2011 por la Agencia de Estadística de Kosovo (KAS). Con base en un conjunto de preguntas del censo que permitían determinar si la persona encuestada pertenecía a la población albanesa desplazada, se extrajo una lista de las PDI albanesas. Para verificar esta lista, se cruzó la información extraída del censo con fuentes complementarias: una lista de las PDI albanesas que habían presentado reclamaciones por pérdida de propiedad a la Agencia de Propiedad de Kosovo (la mayoría de estas relacionadas con desplazamientos causados por situaciones de conflicto), una lista de las PDI albanesas que habían contactado directamente al Consejo Danés para los Refugiados (DRC) y a la Oficina Administrativa de Mitrovica Norte (MNAO) en Mitrovicë, así como a través de líderes comunitarios de la población albanesa desplazada internamente que habían establecido contacto con el ACNUR y el DRC.

PDI serbias en alojamientos privados: una lista completa de 3,872 hogares y 16,383 individuos estaba disponible. La fuente fue la base de datos del Comisariado Serbio para los Refugiados y la Migración (SCRM), que contaba con datos actualizados sobre los hogares de PDI, incluyendo nombres, número de personas en el hogar, direcciones e información de contacto. El marco de muestreo proporcionado por el SCRM fue verificado a través de la comparación con información del ACNUR.

PDI serbias residentes en centros colectivos: una lista completa de 140 hogares estaba disponible. Las fuentes fueron la base de datos del Comisariado Serbio para los Refugiados y la Migración (SCRM) y una base de datos de ACNUR combinadas.

PDI Roms/Ashkalíes/Egipcias (RAE): se creó una lista de 137 hogares y 638 individuos a través de la combinación de una base de datos de ACNUR y las listas de otras organizaciones que ofrecían asistencia a este grupo. Sin embargo, la base de datos de ACNUR tenía varios años y no había sido actualizada. Debido a la vulnerabilidad económica de estos hogares, se esperaba que existiera cierto grado de movilidad, lo que probablemente disminuiría la precisión del marco de muestreo.

Unidades de muestreo y de análisis



Método de muestreo

Los objetivos del ejercicio eran generar una mejor comprensión de las características y situaciones de las poblaciones objetivo, lo que requería que los hallazgos de la encuesta por muestreo fueran representativos de cada una de estas poblaciones. Para esta finalidad, un método de muestreo probabilístico sería la opción preferida. Sin embargo, debido a la disponibilidad limitada de marcos de muestreo actualizados, fue necesario utilizar métodos no probabilísticos para algunas de las poblaciones objetivo.

Como resultado, los métodos para la selección de las muestras fueron adaptados para cada una de las cuatro poblaciones objetivo y diseñados de forma que los hallazgos fueran comparables entre los grupos de población, y no con el fin de ofrecer un panorama de la población desplazada como un todo. A continuación, se describen los diferentes métodos aplicados para cada una de las poblaciones objetivo.

PDI albanesas

Inicialmente, se había planeado utilizar un muestreo aleatorio estratificado para esta población objetivo. Sin embargo, durante el proceso el equipo de la caracterización se vio obligado a abandonar este método debido a desafíos que serán descritos más adelante. Como último recurso, decidieron utilizar el muestreo bola de nieve. El método de muestreo que se había planeado inicialmente puede ser resumido de la siguiente manera:



Paso 1 : Estratificación

Primero, la población objetivo fue dividida en tres estratos con base en áreas administrativas:

- (1) Municipio de Prishtinë/Pristina
- (2) Municipio de Mitrovica Sur
- (3) Municipio de Vushtrri/Vucitrn

Dentro de cada uno de los tres estratos, la población objetivo fue nuevamente estratificada en:

- (1) Áreas urbanas
- (2) Áreas rurales

Paso 2 : Determinación del tamaño de la muestra

Se calculó un tamaño de muestra de 510 hogares con base en un nivel de confianza deseado de 95%, un margen de error deseado de 5% y una tasa de no respuesta esperada de 15%.

Paso 3 : Selección de la muestra

El tamaño de la muestra fue distribuido proporcionalmente entre todos los estratos según el tamaño de la población de PDI albanesas en cada estrato, de acuerdo con la información del censo de 2011 de la Agencia de Estadística de Kosovo (KAS). Luego,

se tomó una muestra aleatoria de cada uno de los estratos con base en la lista extraída del censo.

Después de haber iniciado el trabajo de campo, al momento de contactar a los hogares seleccionados para la muestra con el fin de realizar las entrevistas, se constató que no fue posible encontrar a familias albanesas desplazadas internamente en casi ninguna de las direcciones que aparecían en la lista. En algunos de los casos, se descubrió que las familias que residían en estas direcciones cuando se realizó el censo se habían mudado a otro lugar. Sin embargo, en la mayoría de los casos, no se pudo conocer el motivo por el cual no fue posible encontrar a hogares de PDI en las direcciones en cuestión.

Por este motivo, el equipo de la caracterización abandonó el plan de muestreo original y decidió que sería necesario utilizar el método de muestreo bola de nieve, pero que todavía se buscaría alcanzar el mismo tamaño de muestra y mantener la distribución entre los estratos que se había planeado inicialmente. Los hogares de PDI albanesas que conformaban la muestra original y pudieron ser contactados en las direcciones de la lista del censo sirvieron como punto de referencia y se les solicitó apoyo para identificar a otros hogares de PDI albanesas.

El equipo de la caracterización logró acercarse al tamaño de muestra deseado, dado que en total fueron entrevistados 502 hogares. Sin embargo, la distribución final de los hogares entrevistados entre los municipios correspondientes a los tres estratos no ocurrió completamente conforme a lo planeado (como se demuestra en la tabla a continuación). Debido a la aplicación de un método no probabilístico, la representatividad de la muestra para el total de la población objetivo era incierta.

Municipio	Distribución planeada de los hogares	Distribución final de los hogares
Prishtinë/Pristina	114	45
Mitrovica Sur	315	435
Vushtrri/Vucitrn	81	12
Otro	0	10
Total	510	502

PDI serbias en alojamientos privados

Para seleccionar la muestra de esta población objetivo, se eligió el muestreo aleatorio estratificado. El método aplicado puede ser resumido de la siguiente manera:

Paso 1 : Estratificación

La población objetivo fue dividida en dos estratos con base en su ubicación geográfica:

- (1) Municipios del norte (donde se encontraba la mayor parte de la población de PDI serbias)
- (2) Municipios del sur (donde se encontraba la menor parte de la población de PDI serbias)



Dentro de estos estratos, la población objetivo fue dividida nuevamente en otros dos estratos:

- (1) Áreas urbanas
- (2) Áreas rurales

Paso 2 : Determinación del tamaño de la muestra

Se estimó un tamaño de muestra de 617 hogares con base en un nivel de confianza deseado de 95%, un margen de error deseado de 5% y una tasa esperada de no respuesta de 15%.

Paso 3 : Selección de la muestra

El tamaño de la muestra fue distribuido igualmente entre el estrato de municipios del norte y el de municipios del sur. Esta distribución generó la necesidad de utilizar factores de ponderación durante el análisis de los datos, para compensar el tamaño distinto de la población de PDI serbias en cada estrato. Los factores de ponderación correspondían al número de hogares que cada hogar seleccionado en estos dos estratos representaba. La ponderación sería utilizada siempre que los resultados fueran presentados a nivel global, es decir, cuando se agregaran los hallazgos de estos dos estratos. Al presentar los hallazgos de cada estrato de manera separada, los factores de ponderación no serían aplicados.

El tamaño de la muestra fue distribuido de manera proporcional entre los estratos urbanos y rurales, de acuerdo con el tamaño de la población de PDI serbias residentes en cada estrato de los municipios del norte y del sur. A partir del marco de muestreo, se tomó una muestra de cada uno de los estratos y se realizaron entrevistas con los hogares seleccionados. En total, 567 hogares contestaron la encuesta, lo que representa una tasa de respuesta alta, de 92%.

PDI serbias en los centros colectivos

Debido al tamaño reducido de esta población objetivo (140 hogares), se decidió incluir a todos los hogares en la encuesta, lo que significaba realizar un censo para este grupo. La lista de las personas de la población objetivo estaba actualizada y 135 hogares completaron la encuesta, lo que corresponde a una tasa de respuesta alta, de 96%.

PDI Roms, Ashkalíes y Egipcias (RAE)

Como en el caso de las PDI serbias residentes en los centros colectivos, el equipo de la caracterización decidió realizar un conteo completo de los hogares RAE identificados, debido a su número reducido. Se había identificado un total de 137 hogares, pero dado que la lista compilada de los miembros de la población objetivo estaba algo desactualizada se previó que no sería posible contactar a varios de los hogares. Ante esta situación, se utilizó un muestreo bola de nieve para complementar la lista de hogares de la base de datos de ACNUR.



Se contactó a los hogares de la lista para la realización de las entrevistas y se les solicitó que refirieran a otros miembros de la población objetivo que conocieran.

A partir de este método, se logró entrevistar a 123 hogares, o 491 individuos. Los hogares encuestados estaban ubicados principalmente en Mitrovicë/Mitrovica Sur (39 hogares) y Gračanica/Gračanicë (26 hogares) y, en menor cantidad, en Obiliq/Obilić (11 hogares) y Fushë Kosovë/Kosovo Polje (13 hogares). A través del método bola de nieve, las y los encuestadores pudieron identificar que estaban cerca de alcanzar el punto de saturación, es decir, de haber identificado y entrevistado a todas las personas RAE desplazadas internamente en los municipios mencionados. Esto les permitió concluir que los hallazgos podrían ser considerados representativos de toda la población objetivo.

Desafíos y limitaciones

La metodología de muestreo original tuvo que ser rediseñada, dado que la falta de precisión del marco de muestreo proporcionado por el censo generó la necesidad de aplicar un método distinto. El muestreo no probabilístico que finalmente se utilizó para seleccionar una muestra de PDI albanesas no alcanzó el punto de saturación (fueron entrevistados 502 hogares de un total estimado de 1,167), de modo que no era posible evaluar con certeza la representatividad de la muestra. Asimismo, la mayoría de los hogares de la muestra fueron identificados en el municipio de Mitrovicë/Mitrovica, con un menor número de hogares identificados en Prishtinë/Pristina. Debido a que se esperaba observar cierta movilidad de hogares de PDI que se habían mudado a Prishtinë/Pristina durante los últimos años, se concluyó que la muestra construida a partir del método bola de nieve solo podría ser considerada representativa para el municipio de Mitrovicë/Mitrovica.

La disponibilidad de información secundaria sobre la población general de Kosovo era limitada. Se utilizaron tres fuentes distintas: el Censo de Población y Vivienda de 2011, la Encuesta de Indicadores Múltiples por Conglomerados 2013-2014 del UNICEF y la encuesta Mosaic de 2015 del PNUD. Comparar los resultados de la encuesta realizada con personas desplazadas en 2016 con datos del censo de 2011 no era un escenario ideal. Considerando que los datos se referían a dos momentos distintos en el tiempo, comparar los resultados de las dos fuentes podría representar un problema, dado que las características relevantes de la población general de Kosovo podrían haber cambiado significativamente entre 2011 y 2016.

Ejercicio de caracterización en Mogadiscio, Somalia, 2015-2016



Objetivos

Dos décadas de conflictos armados, así como de sequías e inundaciones severas y recurrentes, han obligado a una parte considerable de la población somalí a abandonar sus hogares. Mogadiscio alberga la mayor población estimada de personas en situación de desplazamiento interno prolongado del país, la cual vive principalmente en asentamientos informales a lo largo de la ciudad y sus alrededores.

El objetivo general del ejercicio de caracterización era ofrecer una base de evidencia para informar la planificación de soluciones duraderas para las PDI a través de actividades de incidencia y de respuestas conjuntas en materia humanitaria y de desarrollo. De manera específica, el ejercicio buscaba:

- (1) Ofrecer una **estimación desagregada del número de PDI** residentes en asentamientos;
- (2) Identificar el desplazamiento interno a través del **análisis del historial de migración de los hogares**;
- (3) **Analizar la situación actual de distintas poblaciones objetivo**, incluyendo su situación socioeconómica, condiciones de vida, acceso a servicios básicos y preocupaciones en materia de protección;
- (4) **Comprender la resiliencia de las familias**, es decir, las habilidades, capacidades, necesidades especiales y mecanismos de supervivencia de la población objetivo que influyen en su toma de decisiones para el futuro.

Población objetivo

Tres poblaciones objetivo fueron incluidas en la encuesta por muestreo del ejercicio de caracterización:

- (1) **PDI**: personas de origen somalí que habían sido desplazadas de su lugar de origen hacia o al interior de Mogadiscio, sobre todo como resultado de conflictos, desastres, situaciones de inseguridad o violaciones de derechos.
- (2) **Migrantes económicos**: personas de origen somalí que habían abandonado su lugar de origen y llegado a Mogadiscio principalmente para buscar medios de subsistencia y por otras razones económicas.
- (3) **Comunidades de acogida**: personas originarias de Mogadiscio que nunca habían salido de su lugar de origen y vivían en asentamientos informales en esta ciudad.

Las siguientes poblaciones objetivo también fueron incluidas en el ejercicio de caracterización, pero no en la encuesta por muestreo. La información relativa a estos grupos fue obtenida a través de métodos cualitativos de recolección de datos:

- (4) **Personas retornadas:** personas de origen somalí que habían huido de Somalia hacia otros países principalmente a causa de conflictos o desastres y luego retornaron a Somalia.
- (5) **Personas refugiadas:** personas originarias de otros países que huyeron a Somalia (y particularmente a Mogadiscio) para buscar protección internacional.

Marco de muestreo

No se contaba con una lista completa que identificara a las PDI, migrantes económicos y comunidades de acogida residentes en los asentamientos informales de Mogadiscio, lo cual también fue parte de los antecedentes que motivaron este ejercicio. Se decidió que los marcos de muestreo tendrían que ser construidos a través de la realización de un mapeo de los asentamientos informales y de la enumeración de los hogares de estos asentamientos.

Unidades de muestreo y de análisis



Método de muestreo

De acuerdo con los objetivos del ejercicio, el equipo de la caracterización tendría que diseñar una metodología de muestreo capaz de ofrecer una estimación del número total de PDI residentes en asentamientos en Mogadiscio, así como información sobre la situación y las características de las PDI, migrantes económicos y comunidades de acogida, garantizando que los hallazgos pudieran ser extrapolados para el total de las tres poblaciones objetivo. Los socios de la caracterización también deseaban hacer comparaciones entre las tres poblaciones objetivo —PDI, migrantes económicos y comunidades de acogida— y, para el caso de la población objetivo de PDI, entre las distintas ubicaciones geográficas. A continuación, se describen los métodos utilizados para cada población objetivo.

PDI

Para esta población objetivo, se eligió un método de muestreo bietápico por conglomerados estratificados. Los asentamientos serían los conglomerados (es decir, unidades primarias de muestreo) y los hogares serían las unidades secundarias de muestreo. El método puede ser resumido de la siguiente manera:


Muestra probabilística

Paso 1 : Mapeo y enumeración

El equipo de la caracterización inició sus actividades con un mapeo de todos los asentamientos informales de Mogadiscio donde se esperaba encontrar PDI. Se buscó identificar dichos asentamientos a partir de la revisión de la información existente sobre sus nombres, coordenadas GPS y estimaciones de población. El mapeo reveló muchos vacíos de información, sobre todo en los dos distritos (Daynille y Kaxda) que eran conocidos por albergar la mayor cantidad de población desplazada internamente y que tenían el mayor número de asentamientos. Por este motivo, se decidió llevar a cabo un ejercicio de enumeración para identificar y contar a todas las personas que vivían en los asentamientos. Esto produciría estimaciones de la población de PDI residentes en los asentamientos (lo que cumpliría con el objetivo **(1)**) y también proporcionaría marcos de muestreo para las tres poblaciones objetivo que serían incluidas en la encuesta.

La enumeración buscaba realizar un censo completo de los hogares de todos los asentamientos identificados durante la fase de mapeo. Además, el equipo de la caracterización colaboró con comisionados de distrito para identificar asentamientos adicionales en cada distrito que sería enumerado. En los distritos de Daynille y Kaxda, también se utilizaron imágenes aéreas de Google Earth para dividir el área en cuadrantes y crear zonas de enumeración más pequeñas y manejables.

El ejercicio de enumeración también permitió identificar los asentamientos que estaban vacíos para que fueran descartados del proceso de muestreo. En cada hogar enumerado, la estructura exterior del alojamiento era pintada con un código único para evitar que un hogar fuera enumerado más de una vez y para facilitar la identificación de los hogares seleccionados para la encuesta por muestreo. Las preguntas que se hicieron durante la enumeración buscaban identificar a los miembros de las diferentes poblaciones objetivo. La enumeración identificó a 68,796 hogares de PDI, conformados por 399,292 individuos.

Paso 2 : Estratificación

Para cumplir con los objetivos de analizar la situación y las distintas características de los hogares de PDI, sería necesario seleccionar una muestra representativa de esta población objetivo. De este modo, los hallazgos referentes a la muestra podrían ser considerados válidos para el total de la población objetivo. Asimismo, las partes interesadas querían poder comparar los hallazgos referentes a las PDI de los distritos de Daynille y Kaxda con los de los demás distritos, debido a que los dos distritos mencionados albergaban a la mayor parte de la población desplazada internamente.

En vista de lo anterior, la población objetivo fue dividida en tres estratos, con base en la ubicación geográfica de la población de PDI:

- (1)** Distrito de Daynille (alberga a 35% de los hogares de PDI en Mogadiscio)

- (2) Distrito de Kaxda (alberga a 20% de los hogares de PDI en Mogadiscio)
- (3) Los demás distritos de Mogadiscio

Paso 3 : Determinación del tamaño de la muestra

El tamaño de la muestra debería ser representativo con un nivel de confianza de 95% y un margen de error de 5% para la muestra general, mientras que cuando esta estuviera dividida en estratos el margen de error sería de 10%. Se estableció un tamaño de muestra de 612 hogares.

Paso 4 : Selección de una muestra de asentamientos

Después del mapeo y de la enumeración de los asentamientos, así como de la estratificación por áreas geográficas, se hizo una selección aleatoria de asentamientos con una probabilidad proporcional al tamaño de la población de PDI en cada estrato.

Paso 5 : Selección de una muestra de hogares

Después de la selección de los asentamientos, se tomó una muestra aleatoria de hogares de cada asentamiento seleccionado para conformar la muestra final.

Migrantes económicos

La base de datos resultante de la enumeración de los asentamientos también pudo ser utilizada como marco de muestreo para esta población objetivo. A partir del conjunto de hogares que habían sido identificados como migrantes económicos durante el proceso de enumeración, se seleccionó una muestra aleatoria simple.



La muestra era representativa con un nivel de confianza de 95% y un margen de error de 5%. Se determinó un tamaño de muestra de 638 hogares.

Comunidad de acogida

Como en el caso de los migrantes económicos, los resultados del proceso de enumeración pudieron ser utilizados como marco de muestreo para la población objetivo de comunidad de acogida. A partir del conjunto de hogares que fueron identificados como parte de la comunidad de acogida, se seleccionó una muestra aleatoria simple.



La muestra era representativa con un nivel de confianza de 95% y un margen de error de 5%. Se determinó un tamaño de muestra de 638 hogares.

Desafíos y limitaciones

La situación inestable de Mogadiscio, en particular por el desalojo de hogares de los asentamientos, representó un desafío significativo para el equipo de la caracterización.

Una intensificación de los desalojos en marzo y abril de 2015 ocasionó la suspensión del ejercicio durante ese periodo. **Para mitigar los efectos de esta situación sujeta a cambios abruptos, se minimizó el intervalo de tiempo entre la enumeración y la aplicación de la encuesta en los hogares.** No obstante, los desalojos siguieron representando un desafío para el ejercicio de caracterización durante la etapa de recolección de datos, aunque en menor escala. Debido a que los desalojos resultaban principalmente en desplazamientos intramunicipales, es decir, de un asentamiento de PDI a otro, su impacto en las cifras generales y en los principales desafíos enfrentados por la población de Mogadiscio fue limitado. Consecuentemente, este ejercicio de caracterización todavía fue capaz de ofrecer un análisis completo de la magnitud y de las tendencias de la situación de desplazamiento, así como un análisis temático para apoyar la planificación de soluciones duraderas.

Durante la aplicación de la encuesta, dos distritos –Heliwa y Abdulaziz– tuvieron que ser excluidos por motivos de seguridad, de modo que los resultados de la caracterización no pueden ser generalizados para estas áreas. Sin embargo, ambos distritos fueron incluidos en el mapeo de instalaciones y en el área de enumeración.

Ejercicio de caracterización en Sudán, 2017-2019



Objetivos

En 2003, un conflicto violento estalló en la región de Darfur. Quince años después, la mayoría de las PDI de la región todavía vivía en campamentos que parecían asentamientos permanentes y seguía dependiendo de asistencia humanitaria. Ante la falta de resolución de la situación de desplazamiento prolongado, encontrar soluciones duraderas para las PDI y las personas retornadas se había vuelto aún más importante.

El ejercicio de caracterización fue realizado para atender a la necesidad de contar con datos exhaustivos que ofrecieran un panorama preciso de la situación enfrentada por las personas desplazadas residentes en los campamentos a las afueras de la ciudad de El Fasher, en Darfur del Norte. Los objetivos del ejercicio eran los siguientes:

- (1) Ofrecer un perfil integral de las PDI** residentes en los campamentos de Abu Shouk y El Salam, así como en áreas de la ciudad vecina de El Fasher. El perfil incluiría datos demográficos fundamentales, información socioeconómica y relativa a los medios de subsistencia, una perspectiva sobre las necesidades de las PDI, sus vulnerabilidades, mecanismos de supervivencia y capacidades. La caracterización también debería considerar las intenciones y planes de las PDI con relación al lugar de asentamiento, así como los factores de incentivo y disuasión que podrían promover un nuevo traslado, el retorno o la integración local.
- (2) Ofrecer un análisis comparativo** enfocado tanto en poblaciones desplazadas como no desplazadas para identificar las vulnerabilidades específicas de las PDI causadas por el desplazamiento. Este análisis buscaba promover una mejor comprensión de los obstáculos y oportunidades para el alcance de soluciones duraderas.
- (3) Ayudar a informar las respuestas programáticas** del Gobierno de Sudán y de los actores humanitarios y de desarrollo y fortalecer la resiliencia de las comunidades de acogida en conformidad con el Marco IASC de Soluciones Duraderas para las PDI.
- (4) Realizar un proyecto piloto de un ejercicio de caracterización del desplazamiento y de planificación conjunta de soluciones duraderas** con miras a replicar el modelo en otros contextos de desplazamiento en Sudán. Como tal, el piloto pretendía facilitar el desarrollo de herramientas de recolección de datos y metodologías acordadas que eventualmente pudieran ser reproducidas con las adaptaciones necesarias.

Población objetivo

Las poblaciones objetivo del ejercicio fueron las siguientes:

- (1) La población de PDI residente en los campamentos de Abu Shouk y de El Salam;
- (2) La población no desplazada residente en la periferia de El Fasher (un área urbana ubicada en las afueras de El Fasher, formando un radio de 3km alrededor de los dos campamentos) y en "El Fasher urbana" (los barrios urbanos de la ciudad de El Fasher).

Marco de muestreo

No había registros u otras listas de la población objetivo disponibles para el equipo de la caracterización. Se decidió que sería necesario crear un marco de muestreo como parte del proceso de selección de la muestra.

Unidades de muestreo y de análisis



Método de muestreo

Los objetivos del ejercicio estaban enfocados en ofrecer un análisis de las situaciones y características de las PDI en comparación con las de sus vecinos no desplazados. También era conveniente comparar a las PDI de los dos distintos campos, así como a las personas no desplazadas residentes en la periferia y en la ciudad, ya que se esperaba que estos grupos se enfrentarían a situaciones distintas según su ubicación. De este modo, el método de muestreo a ser aplicado tendría que producir una muestra que fuera representativa de todos estos subgrupos.

Ante la ausencia de un marco de muestreo preexistente, el equipo de la caracterización decidió llevar a cabo un muestreo bietápico por conglomerados estratificados. El método puede ser descrito de la siguiente manera:



Paso 1 : Estratificación

La población objetivo fue dividida en cuatro estratos, de acuerdo con los distintos grupos que serían comparados:

- PDI en el campamento de Abu Shouk
- PDI en el campamento de El Salam
- Población no desplazada en la periferia de El Fasher
- Población no desplazada en la ciudad de El Fasher (“El Fasher urbana”)

Luego, los cuatro estratos fueron divididos en conglomerados más pequeños que supuestamente tenían tamaños poblacionales similares, con base en cuadrantes creados a partir de un mapa de las áreas en cuestión.

Paso 2 : Determinación del tamaño de la muestra

En cada uno de los estratos, una muestra aleatoria de conglomerados sería seleccionada y posteriormente enumerada. Enseguida, a partir de los conglomerados enumerados, se seleccionarían hogares para la muestra final. Se determinó que el tamaño de la muestra sería de 3,000 hogares, lo que debería equivaler a 252 conglomerados. El tamaño previsto de las muestras de hogares y de conglomerados fue distribuido entre los estratos de la siguiente forma:

Tipo de población	Estrato	Número de hogares	Número de conglomerados
PDI	Campamento de Abu Shouk	996	84
	Campamento de El Salam	996	84
Comunidad de acogida	Periferia de El Fasher	504	42
	El Fasher urbana	504	42
Total		3,000	252

Paso 3 : Selección de una muestra de conglomerados

Se seleccionó una muestra aleatoria simple de conglomerados en cada uno de los estratos. No había cifras confiables de la población de los conglomerados, de modo que estos no fueron seleccionados con una probabilidad proporcional al tamaño poblacional. Por este motivo, al agregar los hallazgos de diferentes estratos para calcular una media poblacional, sería necesario aplicar factores de ponderación.

Paso 4 : Enumeración de los conglomerados seleccionados

El siguiente paso fue enumerar todos los hogares de cada uno de los conglomerados seleccionados. Este ejercicio resultó en **un número de conglomerados enumerados significativamente inferior al que se había planeado inicialmente. Por un lado, esto ocurrió porque se constató que algunos de los conglomerados seleccionados eran inaccesibles para el equipo, principalmente en el campamento de El Salam.** Además, los conglomerados que albergaban menos de 12 hogares fueron descartados de la muestra. Finalmente, algunos de los hogares enumerados habían sido incluidos dos veces en la lista o se encontraban fuera de los límites del conglomerado, y estos

también fueron descartados de la muestra.

Paso 5 : Selección de una muestra de hogares

La enumeración de los hogares de los conglomerados seleccionados creó un marco de muestreo a partir del cual, a través de un muestreo aleatorio simple, se podría tomar una muestra final de hogares para la realización de las entrevistas. Para la selección de esta muestra, los datos de la enumeración fueron restringidos a hogares de PDI en los dos estratos de campamentos y a hogares no desplazados en los dos estratos de El Fasher. Inicialmente, se había planeado que el número de hogares a ser seleccionados sería distribuido igualmente entre los conglomerados, seleccionándose 12 hogares de cada uno de ellos. Sin embargo, con un menor número de conglomerados enumerados, seleccionar 12 hogares por conglomerado para las entrevistas hubiera resultado en un tamaño de muestra significativamente inferior al que se había planeado, de 3,000 hogares. Por lo tanto, para compensar el número reducido de conglomerados, se hizo una selección aleatoria de algunos conglomerados de cada estrato para que fueran sobremuestreados. En dichos conglomerados, se realizarían 24 entrevistas. De este modo, solo los conglomerados que contaban con más de 23 hogares habían sido elegibles para este sobremuestreo.

Se calcularon ponderaciones de muestreo, que fueron aplicadas a los conglomerados de acuerdo con su tamaño y con el número de hogares seleccionados en cada uno de ellos (12 o 24). Los resultados ponderados no pudieron ser sometidos a prueba con base en datos de la población debido a la falta de información actualizada y confiable. Sin embargo, para los campamentos, los tamaños de población calculados a partir de las ponderaciones se acercaban a las estimaciones más recientes realizadas por la Organización Internacional para las Migraciones (OIM).

La siguiente tabla presenta la distribución final real de los conglomerados enumerados, hogares entrevistados e individuos comprendidos en la muestra de hogares entrevistados para cada uno de los cuatro estratos:

Tipo de población	Estrato	Número de conglomerados enumerados	Número de hogares entrevistados	Número de individuos en los hogares entrevistados
PDI	Campamento de Abu Shouk	82	996	5,849
	Campamento de El Salam	50	986	5,960
Comunidad de acogida	Periferia de El Fasher	40	509	3,303
	El Fasher urbana	41	511	3,376
Total		213	3,002	18,533

Desafíos y limitaciones

La muestra de los dos estratos de El Fasher (área urbana y periferia) incluía únicamente a hogares no desplazados. Por lo tanto, se debe tener en cuenta que los resultados referentes a estos estratos no representan a la situación de todos los residentes de la periferia y de los barrios urbanos, debido a la existencia de hogares desplazados en estos estratos que no fueron incluidos en la muestra. De manera específica, la enumeración señaló que alrededor de 28% de los hogares de la periferia y 21% de los hogares del área urbana de El Fasher eran de PDI. De este modo, el análisis comparativo se concentró en las diferencias entre los grupos de población de acuerdo con su condición de desplazamiento, y no puede ser utilizado para comparar las áreas como tal.

Los objetivos del ejercicio, en particular el de informar programas de integración local por zonas, se habrían beneficiado de una comparación adicional por área geográfica (centro de la ciudad vs. periferia vs. campamentos). Para esto, hubiera sido necesario utilizar una muestra más grande (para incluir a dos grupos objetivo en el estrato de la periferia y en el del área urbana, es decir, tanto población desplazada como no desplazada) o reducir cada muestra tomada (con el fin de no invertir más recursos). Esto a su vez produciría muestras con menor poder estadístico, pero podría haber añadido relevancia programática al análisis. Estas son consideraciones importantes que los socios deben tomar en cuenta al principio para llegar a un acuerdo sobre qué es lo más importante.