



# Metodología de estimación de población y vivienda 2022

San José, Costa Rica  
Julio 2023

## Presentación

El Instituto Nacional de Estadística y Censos (INEC) presenta la metodología para la estimación de población y vivienda 2022 con el objetivo de transparentar los procesos implementados y de permitir una mayor comprensión de los resultados obtenidos.

Las estimaciones de población y vivienda 2022 son el resultado de un complejo proceso técnico donde se implementaron tres metodologías distintas, las cuales se validaron y analizaron para determinar cuáles serían los datos definitivos que se divulgarían.

El INEC agradece profundamente al Centro Latinoamericano y Caribeño de Demografía (CELADE), a la División de Estadística de la Comisión Económica para América Latina (CEPAL), al Fondo de Población de Naciones Unidas (UNFPA) y al Centro Centroamericano de Población (CCP), por su asistencia técnica y acompañamiento en este proceso; así como a los diferentes equipos de la institución que participaron en este trabajo para brindar al país los principales datos demográficos del 2022.

## Tabla de contenido

Presentación .....	2
Símbolos y siglas.....	5
1. Introducción .....	6
2. Contexto en el que se realizó el Censo 2022.....	7
2.1 Planificación del Censo 2022.....	7
2.2 Pandemia por COVID-19.....	8
2.3 Dificultades presupuestarias para el uso de recursos .....	9
2.4 Faltante de personal censista.....	10
2.5 Cambios tecnológicos .....	13
2.6 Proceso de capacitación .....	15
2.7 No respuesta y rechazo .....	16
3. Estimación de la cobertura del Censo 2022 .....	17
4. Metodologías para la estimación de población y viviendas.....	19
4.1 Estimación de población basada en la información censal .....	20
4.2 Estimación demográfica de población basada en registros administrativo .....	28
4.3 Modelo bayesiano para estimación de población .....	33
4.4 Validación de resultados y definición de estimación de población final.....	37
4.5 Estimación de viviendas.....	38
5. Bibliografía .....	40

## Índice de cuadros

Cuadro 3. 1. Distribución de las viviendas según estado de cobertura, 2022.....	19
Cuadro 4. 1. Distribución de las viviendas desocupadas por tipo, 2000 y 2011 .....	25
Cuadro 4. 2. Distribución de las viviendas desocupadas reclasificadas, 2022 .....	27

## Índice de figuras

Figura 4. 1. Distribución de puntos de viviendas por resultado de entrevista. Distrito San Sebastián, 2022. ....	26
Figura 4. 2. Estimación del número de viviendas desocupadas y del número de habitantes en Costa Rica.....	36
Figura 4. 3. Estimación de la estructura poblacional (hombres y mujeres en edades quinquenales) en Costa Rica .....	37

## Índice de gráficos

Gráfico 4. 1. Promedio de ocupantes por vivienda según operación estadística, 2000-2022 .....	21
Gráfico 4. 2. Promedio de ocupantes por vivienda según Región de Planificación, 2000-2022.....	22

## Símbolos y siglas

### Símbolos

- (-) Presente en un cuadro indica que la cantidad es nula o cero.
- (.) Se utiliza para señalar decimales.
- () Espacio en blanco para los miles.

### Siglas

CCP	Centro Centroamericano de Población
CCSS	Caja Costarricense de Seguro Social
CELADE	Centro Latinoamericano de Demografía
CEPAL	Comisión Económica para América Latina y el Caribe
CONADECO	Confederación de Asociaciones de Desarrollo Comunal
COVID-19	Enfermedad causada por el virus del SARS-CoV-2
DGME	Dirección General de Migración y Extranjería
INEC	Instituto Nacional de Estadística y Censos
MEP	Ministerio de Educación Pública
OMS	Organización Mundial de la Salud
ONU	Organización de Naciones Unidas
RUP	Registro Único de Población
TEC	Instituto Tecnológico de Costa Rica
UNFPA	Fondo de Población de Naciones Unidas

## 1. Introducción

El XI Censo Nacional de Población y VII de Vivienda se llevó a cabo entre junio y setiembre de 2022, no obstante, debido a diferentes situaciones que se profundizan en este documento, la cobertura final obtenida fue parcial.

Dicha situación generó la necesidad de trabajar en una estimación de población y de viviendas con el objetivo de brindar al país los principales indicadores demográficos para la toma de decisiones a nivel nacional y cantonal.

Para ello, el INEC trabajó con la asesoría de la División de Estadística de la CEPAL, con CELADE, con UNFPA y con el CCP para implementar diferentes metodologías estadísticas para la estimación de la población.

Se desarrollaron tres metodologías de estimación, a saber: una basada en los datos censales; otra que se fundamenta en registros administrativos de carácter demográfico y otra que se basa en un modelo bayesiano para estimación de población. Finalmente se realizó un proceso de validación y comparación de las tres metodologías con diferentes fuentes y los mismos datos censales para determinar la estimación definitiva.

En ese contexto y en aras de la transparencia y rigurosidad técnica que siempre ha caracterizado al INEC, en este documento se presentan, los factores que influyeron en la cobertura parcial del Censo 2022, la estimación de cobertura final y la metodología detallada de las tres estimaciones de población realizadas, así como de la estimación final de viviendas.

## 2. Contexto en el que se realizó el Censo 2022

### 2.1 Planificación del Censo 2022

Desde el año 2018 el INEC inició el trabajo de planificación del XI Censo Nacional de Población y VII de Vivienda, que es sin duda la operación estadística más compleja que desarrolla el país pues implica recolectar información de todas las viviendas y personas residentes habituales de Costa Rica.

Por tanto, se inició con una etapa de preparación que incluyó la conformación de un equipo de trabajo, la elaboración de un presupuesto y el diseño del perfil del proyecto. En esta primera etapa también se realizó la inscripción del proyecto en el Banco de Proyectos de Inversión del Ministerio de Planificación Nacional y Política Económica (MIDEPLAN).

El desarrollo de un proyecto de tal envergadura como un censo nacional requiere de un equipo interdisciplinario que sea responsable de los diferentes procesos implicados, a saber, cartografía, diseño metodológico, recolección de la información, procesamiento y análisis de datos, tecnologías de información, comunicación y por supuesto, administración y planificación.

Entre 2018 y 2021 se diseñaron e implementaron los diferentes procesos necesarios, los cuales se trabajaron en torno a un Plan Anual de Trabajo del proyecto, donde se establecieron metas trimestrales y, además, se elaboraron matrices de riesgos que eran revisadas periódicamente. En el periodo de planificación del censo nacional se llevaron a cabo las siguientes actividades:

- Actualización cartográfica de campo, donde se georreferenciaron todas las viviendas del país.
- Diseño del cuestionario censal, que incluyó una consulta a personas usuarias y expertas con la participación de más de 400 personas.
- Diseño del operativo de recolección de datos donde 34 equipos denominados Agencias Censales se encargaron de gestionar los recursos y realizar las coordinaciones necesarias en todo el país.
- Diseño del proceso de capacitación, el cual por primera vez en la historia de los censos se realizó de forma bimodal y, por tanto, requirió del diseño de un curso virtual y diferentes recursos didácticos.

- Diseño de los diferentes sistemas informáticos, tanto los que se desarrollaron en el INEC como aquellos que se contrataron. Entre ellos se encuentran el App Censista, el Autocenso, el Sistema Gestor de Operativo de Campo, Sistema de Procesamiento y Sistema Informatizado de Codificación.
- Diseño de una campaña de comunicación integral y de todos los productos comunicacionales relacionados.
- Coordinación de convenios con diferentes instituciones nacionales y locales para el apoyo durante la capacitación y la recolección de los datos.
- Gestión de la logística de distribución y resguardo de material censal por todo el país.
- Gestión de compra de los dispositivos móviles de captura y los demás insumos y servicios requeridos para la implementación del operativo de campo.
- Implementación de siete pruebas de campo, pruebas cognitivas y un Censo Piloto donde se censó todo el cantón de El Guarco de Cartago.

Todos los procesos contaron con productos específicos y fueron evaluados por el sistema de control interno de la institución, no obstante, a pesar de la amplia y detallada planificación, hubo factores que afectaron especialmente la recolección de los datos, los cuales se detallan a continuación.

## **2.2 Pandemia asociada a COVID-19**

En marzo de 2020 la Organización Mundial de la Salud (OMS) declaró el COVID-19 como pandemia. El día 16 de ese mismo mes se declara en Costa Rica el estado de emergencia nacional debido a la pandemia, la cual impuso una serie de restricciones a las instituciones, empresas y ciudadanía en general, con el objetivo de resguardar la salud de la población.

Las principales restricciones aplicadas fueron la implementación de teletrabajo en las instituciones públicas y las empresas privadas, los cierres de negocios con permiso sanitario, así como la limitación de circulación vehicular. Durante todo el 2020 y 2021 las restricciones sanitarias se relajaron o intensificaron dependiendo de las denominadas “olas” del virus, es decir, momentos con aumentos o disminuciones importantes de casos.



Para ese momento el INEC tenía dos años de trabajar en la planificación del censo nacional que sería originalmente en 2020. Sin embargo, la pandemia fue uno (no el único) de los factores para decidir posponerlo, primero a 2021 y luego a 2022.

Los cambios que se implementaron correspondieron a la reorganización de las tareas de las Agencias Censales para priorizar aquellas que podían realizarse en formato de teletrabajo; aplicar una metodología de actualización cartográfica que no requiriera las visitas en campo, para finalizar los distritos pendientes; se decidió cambiar la metodología para la modalidad de recolección de datos por autocenso durante el operativo censal, la cual estaba originalmente pensada para la recuperación de viviendas pendientes en un periodo posterior; la adquisición de equipo de resguardo sanitario (mascarillas y alcohol en gel) para cada persona que participaría en el operativo.

El Censo Piloto se pudo realizar en el cantón de El Guarco en la provincia de Cartago en noviembre de 2020, no obstante, por las restricciones no fue posible contratar personal exclusivo para eso, por lo que se tuvo que realizar todo el operativo con personal de las mismas agencias censales. Además, la capacitación, originalmente planteada en una estrategia bimodal (virtual/presencial), tuvo que hacerse 100% virtual.

Si bien el Censo Piloto cumplió con algunos de sus objetivos, como probar el cuestionario y los sistemas de captura de datos; los ajustes realizados debido a la pandemia tuvieron impacto pues no permitieron probar totalmente las condiciones del censo nacional, como el proceso de reclutamiento o la carga masiva del aplicativo de captura.

### **2.3 Dificultades presupuestarias para el uso de recursos**

La Ley 9694 del Sistema de Estadística Nacional, de junio 2019, brinda al INEC solvencia financiera para realizar los censos nacionales, por tanto, se contaba con los recursos necesarios para la implementación del proyecto censal, sin embargo, en el 2020 y 2021 enfrentó a las restricciones establecidas por la Ley 9635 denominada “Fortalecimiento de las finanzas públicas”, la cual establece una regla fiscal que define un límite de crecimiento en el gasto corriente de las instituciones públicas.

Debido a la incertidumbre sobre la utilización de los recursos disponibles para continuar con el proyecto en 2020 y 2021, se tomó la decisión de aplazar el censo nacional para el 2022

Sin embargo, fue necesario abogar por una modificación a la Ley 9694 para que fuera posible presupuestar los recursos requeridos.

La modificación a la ley se logró en noviembre de 2021, casi al cierre del año, lo cual mantuvo en vilo muchos aspectos, y dejó plazos muy cortos para el inicio del reclutamiento de personal censista, la compra de insumos como los dispositivos móviles, y la contratación de servicios como la infraestructura tecnológica que permitía la implementación de los sistemas desarrollados, así como la firma de convenios muy importantes como el que se necesitaba establecer con el Ministerio de Educación Pública (MEP) para el préstamo de instalaciones y de personal para facilitar la capacitación.

## **2.4 Faltante de personal censista**

El Censo 2022 fue el primer censo en Costa Rica que se realizó utilizando dispositivos móviles de captura y ese cambio tan importante generó modificaciones en casi todos los procesos del proyecto, uno de ellos, el personal para la recolección de datos.

Tradicionalmente el personal que recolectaba la información de los censos eran los y las docentes del MEP. Esto era posible gracias a que el periodo de recolección era de cinco días y, por tanto, a través de un convenio, se brindaba autorización para que durante esos días no se brindaran lecciones y el personal del MEP fuera a censar, de manera que recibían su salario y, además, el INEC les brindaba un incentivo económico adicional.

Sin embargo, al tomar la decisión de incorporar los dispositivos móviles en la captura de la información, también se debió ampliar el periodo de recolección, por varios motivos: por la cantidad de dispositivos que se podía adquirir, por la logística de instalación del aplicativo y por la necesidad de tener más espacio de reacción ante cualquier eventualidad. Ante eso, se amplió el periodo a 18 días para lo que se requería un total de, aproximadamente, 14 mil personas distribuidas entre las coordinaciones de centros operativos, las supervisiones de sección y las personas censistas.

Ante esos cambios y, además, ante la crisis educativa que vivió el país debido a todos los ajustes que debió hacer el MEP ante la pandemia de COVID-19 se determinó que para el Censo 2022 no sería posible trabajar con los y las docentes, al menos no para la recolección

de datos, aunque sí se tenía previsto que apoyaran como personas facilitadoras en el proceso de capacitación.

Esa decisión implicó que el INEC debía buscar nuevas alternativas para contratar al personal de campo, lo cual no era tarea sencilla pues hay pocas experiencias nacionales de una contratación tan masiva para tan poco tiempo (aproximadamente un mes de contratación). Por ello, en primera instancia se realizó una investigación de las diferentes opciones.

Surgieron varias opciones, todas con ventajas y limitaciones. Voluntariado, contratación por servicios profesionales, contratación terciarizada o que el mismo INEC contratara utilizando plazas temporales. Según todos los análisis, la opción con menos riesgos era la última, es decir, que se crearan plazas temporales y que fuera el mismo INEC, quien ya ha tenido la experiencia de manejar planillas muy grandes de forma temporal, hiciera el proceso de reclutamiento y selección.

Para que eso fuera realidad era necesario acudir al Ministerio de Hacienda y a la Secretaría Técnica de la Autoridad Presupuestaria para que se autorizara al INEC la creación de las 14 mil plazas temporales. No obstante, no hubo respuesta positiva ante esa solicitud y por tanto hubo que optar por otra opción más riesgosa.

La otra alternativa era buscar una entidad que pudiese gestionar la planilla, realizar todo el proceso de reclutamiento, selección y administración del personal para todo el país. Inicialmente se realizó un estudio de mercado para determinar posibles empresas privadas que pudieran ofrecer el servicio, sin embargo, no se ubicaron empresas dispuestas a brindarlo por su complejidad.

Ante ese panorama, surgió la posibilidad de trabajar con la Confederación de Asociaciones de Desarrollo Comunal (CONADECO) que es la organización que aglomera a las diferentes asociaciones de desarrollo comunal del país, a través de Federaciones y que tiene presencia en prácticamente todo el territorio nacional.

La ventaja de trabajar con CONADECO era justamente su amplio conocimiento de los territorios, su experiencia con el trabajo de base y su capacidad de movilizar a las comunidades por un objetivo. Sin embargo, nunca habían tenido la experiencia de manejar tanto personal de forma remunerada.

La metodología utilizada consistió en que el INEC contrató a 14 Federaciones, las cuales a su vez conformaban CONADECO. Cada federación realizó la gestión de una planilla. Además, también se contrató personal que brindarían apoyo administrativo y logístico, como una

coordinación administrativa, una coordinación financiera, profesionales en recursos humanos y asistentes de planillas para cada federación.

Sin embargo, el tiempo fue uno de los principales problemas, ya que, debido a que al final de 2021, por la incertidumbre de la continuidad o no del proyecto, se paralizaron todas las negociaciones, no fue sino hasta principios de 2022 que se logró comenzar a trabajar con CONADECO.

Si bien hubo mucha voluntad de trabajar coordinadamente, algunas federaciones debían realizar trámites previos para poder ser contratadas, lo cual fue atrasando algunas de las actividades planificadas y algunos contratos se firmaron muy cerca de la fecha de inicio del operativo de campo.

El reclutamiento del personal inició con listas previas que cada federación realizó, donde las personas interesadas en participar se anotaban. En ese proceso, denominado pre-reclutamiento, se llegó a tener enlistadas a más de 20 mil personas, de manera que para ese momento se estaban cumpliendo los objetivos establecidos.

El proceso formal de reclutamiento debía hacerse en un formulario web que a su vez era el sistema que daba seguimiento a cada persona durante la capacitación, la selección y la asignación de su carga de trabajo. Si bien era un sistema simple, para muchas personas fue un primer obstáculo, ya que requerían un correo electrónico y también tener conocimientos básicos del uso de una computadora. A pesar de que se habían establecido perfiles para las personas a contratar, donde se incluía como requisito tener conocimientos mínimos del uso de computadora o dispositivos móviles, fue evidente que algunas personas de las que estaban pre-reclutadas no cumplían ese requisito y ahí comenzó a disminuir la lista.

Otro de los factores que disminuyó la lista fue que algunas personas se anotaron para los tres puestos disponibles (coordinación de centro operativo, supervisión de sección y censistas) sin embargo posteriormente al determinar que no cumplían alguno de los requisitos o al enterarse que se debía cumplir con un proceso de capacitación, preferían salir de la lista y no continuar con el reclutamiento. Además, también hubo lugares del país donde no había personas interesadas en el trabajo, esto sucedió sobre todo en zonas turísticas.

Finalmente, al cierre del periodo establecido para el reclutamiento del personal no se logró tener la cantidad total, particularmente para el puesto de censistas hubo un déficit de

aproximadamente 2000 personas. Eso generó que se ampliara el reclutamiento incluso posterior al inicio del operativo, no obstante, eso implicó ajustes en los demás procesos asociados, como la capacitación. Por tanto, al finalizar la primera semana del operativo, se decidió no seguir contratando y continuar con las personas que ya estaban contratadas.

Posterior a la finalización del periodo original de recolección (28 de junio), se amplió el contrato a unas 1500 personas quienes continuaron recolectando información durante tres semanas más en las áreas que no habían podido ser cubiertas.

## 2.5 Cambios tecnológicos

Desde el inicio del proyecto se había decidido que el censo nacional se realizaría potenciando el uso de la tecnología en todos los procesos, desde la actualización cartográfica, el diseño del operativo de campo, la recolección de los datos, la capacitación, la gestión del personal y por supuesto, el procesamiento y análisis de la información.

Ante eso, el INEC buscó contratar a personas expertas en la materia para que asesoraran en ese camino y que desarrollaran la aplicación para recolección de datos, tanto en campo como vía web y también, el sistema para la gestión del personal y el seguimiento de la cobertura, lo cual suele ser de lo más complejo en una operación estadística tan grande. Después de un proceso de análisis de mercado, se contrató a Go Touch, una iniciativa de investigación en informática del Instituto Tecnológico de Costa Rica (TEC), quienes se han especializado en el desarrollo de aplicaciones móviles.

El objetivo era tener una serie de sistemas integrados que gestionaran desde el reclutamiento y la selección del personal, hasta la asignación de cargas de trabajo y posteriormente el seguimiento de la cobertura durante el operativo. De esa manera, el mismo usuario que las personas creaban para reclutarse sería el que utilizarían para ingresar a la aplicación de recolección de datos (denominada App Censista) y para el módulo de supervisión.

Esa integración si bien buscaba facilitar la trazabilidad de la información y garantizar el control, generó cuellos de botella pues algunas personas no recordaban el usuario que habían creado, habían utilizado correos mal escritos, correos saturados o incluso, algunos utilizaron correos de otras personas, de manera que no podían tener acceso al resto de los procesos, incluyendo su carga de trabajo.

Como complemento a los sistemas desarrollados también fue necesario obtener un sistema de gestión para los dispositivos móviles, con el objetivo de que el INEC pudiera tener el control de ellos, limitara el acceso a aplicaciones no requeridas para el trabajo y también para facilitar la instalación de la App Censista. El sistema utilizado fue Workspace de Google el cual permitía todas esas funcionalidades, pero significaba que cada persona debía instalar una cuenta en el dispositivo que se le asignara.

Si bien era un proceso sencillo, en un operativo tan masivo, la mínima complejidad se vuelve grande y, por tanto, esa instalación también fue un cuello de botella, ya que, a pesar de la capacitación y los recursos explicativos, no todas las personas lograron instalarlo en el momento que se requería.

Esos cuellos de botella tuvieron como consecuencia que el primer día del operativo no todas las personas habían logrado acceder a la aplicación y por tanto a su carga de trabajo. Eso fue disminuyendo durante los primeros días, sin embargo, hubo personas que durante todo el operativo no lograron ingresar y por tanto debieron trabajar con cuestionarios en papel.

También hubo algunas zonas fronterizas donde los mapas digitales que debían ir incorporados en la App Censista tuvieron problemas con las coordenadas, por tanto, esas zonas también se tuvieron que censar con cuestionarios en papel, pues las personas censistas asignadas no iban a poder ingresar a esos mapas.

Además de las complicaciones asociadas al uso de los sistemas, otro factor que complejizó más la situación fue que el gobierno de Costa Rica sufrió un ciberataque aproximadamente dos meses antes del inicio del censo, el cual afectó a varias instituciones cuyos sistemas dejaron de operar por varias semanas. Si bien el INEC no fue víctima del ataque, se tuvieron que tomar nuevas medidas de ciberseguridad aún más estrictas lo que, a su vez, complicó más el ingreso a los sistemas por parte de las personas usuarias.

## 2.6 Proceso de capacitación

Por primera vez en un censo nacional se implementó una capacitación híbrida o bimodal, donde hubo un curso virtual de autoaprendizaje para cada rol dentro del operativo de campo y, además, dos días de capacitación presencial donde se realizaron prácticas y se hizo la instalación de la App Censista.

Este proceso se vio afectado directamente por los atrasos y las dificultades surgidas durante el reclutamiento del personal pues al tener que ampliar los plazos, también hubo que modificar las fechas de cierre de los cursos e incluso, realizar un curso resumido para aquellas personas que fueron contratadas cuando el operativo ya había iniciado.

En el caso de la modalidad presencial, originalmente se había planteado bajo la metodología de cascada, donde un primer grupo de personas funcionarias del INEC capacitarían a personas docentes del MEP quienes facilitarían los talleres presenciales para las personas supervisoras y censistas.

Sin embargo, debido a la incertidumbre sobre la continuidad del proyecto, no se logró firmar el convenio con el MEP con suficiente antelación. A pesar de todas las gestiones, no fue posible que el jerarca del gobierno saliente firmara el convenio y fue necesario esperar al cambio de poderes para que la nueva ministra lo firmara. Eso sucedió tan solo dos semanas antes del inicio del proceso de capacitación presencial.

Con tan poco tiempo de reacción y ante una estructura muy jerárquica establecida en el MEP, la información no se compartió con oportunidad entre los centros educativos y, por tanto, solo el 50% de las personas MEP que se requerían se presentaron a recibir su capacitación. Esto implicó que más de cien personas funcionarias del INEC tuvieran que ir a impartir capacitación al último nivel y, además, que las personas coordinadoras de centro operativo tuvieran que asumir esa función en algunos lugares. También se realizaron capacitaciones de contingencia que resumían los principales contenidos y que se enfocaban en la instalación de la App Censista y la práctica.

## 2.7 No respuesta y rechazo

A pesar de todos los esfuerzos que el INEC realizó para superar todos los retos ya reseñados y lograr visitar más del 90% de las áreas de empadronamiento establecidas, el obstáculo que no fue posible superar fue el de la apatía de la población, que generó porcentajes de rechazo y no respuesta nunca antes vistos en un censo nacional.

Esa actitud se enmarca en un contexto mucho más amplio que el mismo censo, es el resultado de un proceso social que está viviendo el país y que se ha visibilizado en otros momentos relevantes como, por ejemplo, el abstencionismo en los procesos electorales.

La apatía política usualmente no se asocia con el desinterés en participar en una operación estadística, sin embargo, cuando se trata de un censo nacional como proceso para generar datos que tienen relevancia política, claro que se relaciona, especialmente con el componente de desafección, es decir, con la indiferencia y el desinterés de los asuntos que de una u otra forma requieren participación o acción de la ciudadanía (Cebrián, 2011) (Romano, 2015).

Los factores causantes de esa apatía pueden ser diversos, entre ellos desconfianza, insatisfacción, desmotivación, exclusión, desigualdad, entre otros, que pueden gestarse en la población a través de largos procesos o por situaciones particulares de alguna coyuntura. En el caso costarricense, es sin duda necesario un análisis a profundidad de esa apatía, no obstante, es evidente que, para el Censo 2022, algunos hechos particulares y casos de corrupción agravaron esas actitudes.

Sin embargo, Costa Rica no es el único país que enfrenta esta apatía, otros países han evidenciado, incluso desde rondas pasadas que hay un incremento en las entrevistas no realizadas o rechazadas. De hecho, la Fundación Russell Sage publicó en 2006 el libro *The Hard Count. The Political and Social Challenges of Census Mobilization*, que tuvo como objetivo explorar la participación en el Censo 2000 de Estados Unidos y evaluar la campaña de comunicación de cara al Censo 2010 pero, además, revisa una serie de elementos contextuales, principalmente políticos, que intervienen en esa participación.

El libro concluye que uno de los factores asociados a la falta de cooperación en el censo se refiere a la desconfianza de la población sobre cómo el gobierno utilizará sus datos y con quién los compartirá, lo cual implica a su vez una desconfianza en la institución productora y en el manejo confidencial de los datos. Además, se encontró una relación entre el tipo de



hogar y la cooperación con el censo, donde se determinó que hay una mayor cooperación entre los hogares convencionales (parejas con o sin hijos), que los hogares no convencionales (hogares monoparentales, unipersonales o conformados por personas no familiares), no obstante, es difícil saber por qué eso es así, posiblemente se relaciona con estilos de vida y uso del tiempo (Hillygus, Nie, & al., 2006).

Si bien es necesario ahondar mucho más en las causas de la no respuesta y el rechazo a brindar información, lo cierto es que afectó profundamente los resultados del Censo 2022, pues si bien la cobertura implica visitar todas las áreas definidas, también es requisito fundamental, completar todas las entrevistas y eso no fue posible en un porcentaje importante de viviendas.

### **3. Estimación de la cobertura del Censo 2022**

Para el monitoreo de la cobertura durante la recolección de información del Censo 2022 se contó con un sistema denominado Sistema Gestor de Procesamiento y, además, con la transferencia continua de entrevistas, lo que permitía ir dando seguimiento al avance del operativo.

Por tanto, de forma temprana se identificaron problemas de cobertura, especialmente en zonas donde no fue posible contratar la totalidad del personal requerido, lo cual dejaba al descubierto las áreas de empadronamiento que no tenían una persona censista asignada.

En ese contexto, se toma la decisión de ampliar el periodo de recolección hasta el 30 de setiembre de 2022, es decir, prácticamente tres meses más de lo que estaba originalmente planificado (8 al 28 de junio). Para cubrir ese periodo adicional se extendió la contratación a aproximadamente 1 500 personas, por tres semanas más y posteriormente, se contó con el apoyo de personas voluntarias y del mismo INEC.

Además, se promovió el uso del autocenso a través de los medios de comunicación y se creó un visor en línea que permitía solicitar un código sin necesidad de tener que esperar que una persona censista se lo entregara. También se brindó la posibilidad de hacer la entrevista telefónica y se dio seguimiento para promover la participación.

Si bien se logró disminuir el porcentaje de viviendas no visitadas de 28% al final del periodo ordinario a 12% al final del periodo extraordinario, el porcentaje de viviendas visitadas que

rechazaron la entrevista o bien que nunca se logró contacto para la entrevista, no se logró disminuir.

Finalmente, debido a la complejidad administrativa de continuar con este proceso y, a que no se estaban logrando los resultados esperados, el 30 de setiembre se da por finalizado el operativo de recolección.

Posteriormente, se realizaron las siguientes acciones para determinar la cobertura final:

- **Integración de datos:** si bien más del 95% de los datos se recolectaron utilizando los dispositivos móviles de captura, se recibieron 32 463 entrevistas de autocensos y también se aplicaron aproximadamente 16 mil entrevistas en cuestionarios de papel, los cuales pasaron por un proceso de crítica y digitación. Además, también se obtuvo información de 140 viviendas colectivas por medio de sus registros administrativos, debido a la complejidad para realizar las entrevistas cara a cara, entre ellas los centros penitenciarios. Por tanto, fue necesario realizar un proceso para integrar esas diferentes fuentes y así contar en una sola base de datos con todas las viviendas.
- **Depuración de la base de viviendas:** durante la actualización cartográfica de campo se georreferenciaron todas las viviendas que formaron parte de los mapas que llevaron las personas censistas. La base inicial contenía 1 770 375 viviendas. Sin embargo, las personas censistas por su parte también realizaron una actualización de esas viviendas pues al visitarlas determinaron si se mantenían ahí, si ya no existían o bien, si había viviendas nuevas. Por tanto, al finalizar el operativo se determinó que se crearon 203 125 viviendas nuevas y se descartaron 144 605. Además, también se realizó una depuración de puntos duplicados y erróneos, con lo cual se obtuvo una base final con 1 769 081 viviendas. Esta base final es la que se tomó como referencia para la estimación de la cobertura de viviendas.

Con esa información y los datos depurados, se determinó que se logró visitar el 88% de las viviendas del país, es decir 1 550 168, de las cuales 1 074 258 fueron censadas (60,7%); 136 099 viviendas donde se rechazó la entrevista (7,7%) y 339 901 (19,2%) quedaron pendientes, es decir, se visitó la vivienda, pero no se logró obtener la entrevista. El total de viviendas que no fueron visitadas fue de 218 823 lo que representa un 12,4%.

Cuadro 3. 1. Distribución de las viviendas según estado de cobertura, 2022

### CUADRO 3.1

#### Distribución de las viviendas según estado de cobertura, 2022

Estado de las viviendas	Absoluto	Porcentaje
<b>Total</b>	<b>1 769 081</b>	<b>100,0</b>
<b>Visitadas</b>	<b>1 550 258</b>	<b>87,6</b>
Censadas	1 074 258	60,7
Rechazadas	136 099	7,7
Pendientes	339 901	19,2
<b>No visitadas</b>	<b>218 823</b>	<b>12,4</b>

Fuente: INEC-Costa Rica. Censo 2022.

Debido a la cobertura parcial fue necesario realizar una estimación de población para poder brindar un dato adecuado para su uso. Por tanto, a continuación, se detallan las metodologías utilizadas para ello.

## 4. Metodologías para la estimación de población y viviendas

Para la implementación de las metodologías de estimación de población se contó con el valioso apoyo técnico y asesoría de la División de Estadística de la CEPAL, del Centro Latinoamericano y del Caribe de Demografía (CELADE), del Fondo de Población de Naciones Unidas (UNFPA) y del Centro Centroamericano de Población (CCP). Gracias a esta colaboración fue posible contar con diferentes métodos de estimación basados en técnicas estadísticas que permitieron tener un dato final robusto y de calidad.

Se desarrollaron tres metodologías para la estimación de población, una basada en la misma información censal y los promedios de ocupantes por vivienda; otra que utilizó información demográfica de registros administrativos; y una tercera que implementó un modelo predictivo bayesiano para la estimación de población a partir de las viviendas y diferentes covariables. A continuación, se detalla cada metodología.

## 4.1 Estimación de población basada en la información censal

La primera metodología se basa en estimar la población, a nivel de distrito, a partir de los promedios de ocupantes con las viviendas que sí se lograron censar. Para ello, primero se debieron analizar y tomar decisiones sobre algunos factores que podían incidir en la estimación, ya que por la cobertura parcial es necesario tener la mayor certeza posible de que los datos utilizados son adecuados y están generando estimaciones coherentes. Los aspectos analizados fueron los siguientes:

- Consistencia del promedio de ocupantes por vivienda por distrito.
- Ajuste de clasificación en las viviendas desocupadas.
- Clasificación de viviendas no visitadas.

Los procesos se detallan a continuación:

### 4.1.1 Consistencia del promedio de ocupantes por distrito

Debido a la cobertura parcial del Censo 2022 fue necesario realizar una validación y ajuste de los promedios de ocupantes por vivienda, ya que ese dato sería el principal para esta metodología de estimación.

Para ello se implementó un modelo de regresión que, a través de una función lineal, busca modelar una variable Y en términos de un conjunto de variables predictoras X, mediante la siguiente fórmula:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Donde:

- $\beta_i$   $i=0,1,2,3$  son parámetros desconocidos; los cuales se estiman a partir de los datos utilizados para predecir. En este caso se utilizaron los promedios de personas observados para la Encuesta Nacional de Ingresos y Gastos de los Hogares 2014 y 2018, Encuesta Nacional de Hogares 2010 a 2022 y los Censos 2000, 2011 y 2022.
- $\varepsilon$  es el error entre la distancia de un valor de la respuesta y el promedio condicional dado un valor de  $X_i$ .

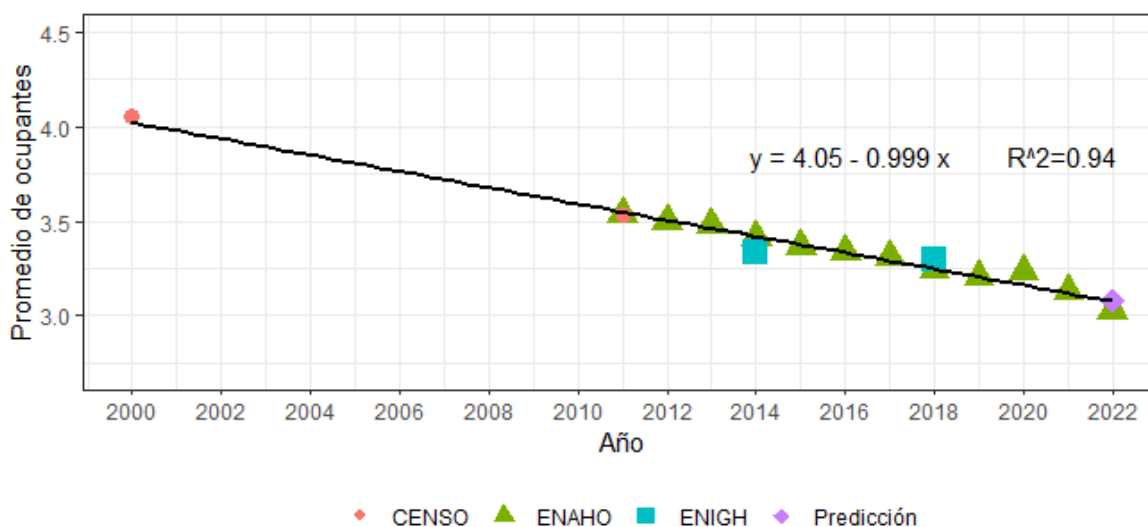
Para obtener los valores de los coeficientes de la ecuación del modelo de regresión, se aplicó la técnica de mínimos cuadrados ordinarios. Una vez obtenidos los coeficientes y sustituyendo el valor del año específico para el cual se deseaba realizar la predicción, se obtuvo la estimación del promedio de personas por vivienda.

Esta metodología se implementó a nivel nacional y para cada una de las regiones de planificación<sup>1</sup>, obteniendo como resultado 7 rectas de regresión: una para cada región y una estimación a nivel nacional para el 2022.

En el gráfico 4.1, se puede observar que, al aplicar el modelo a nivel nacional, este da como resultado un intercepto positivo y una pendiente negativa, es decir, en el periodo transcurrido el promedio de personas por vivienda disminuyó aproximadamente 0,999 personas por vivienda.

#### GRÁFICO 4.1

##### Promedio de ocupantes por vivienda según operación estadística, 2000 - 2022



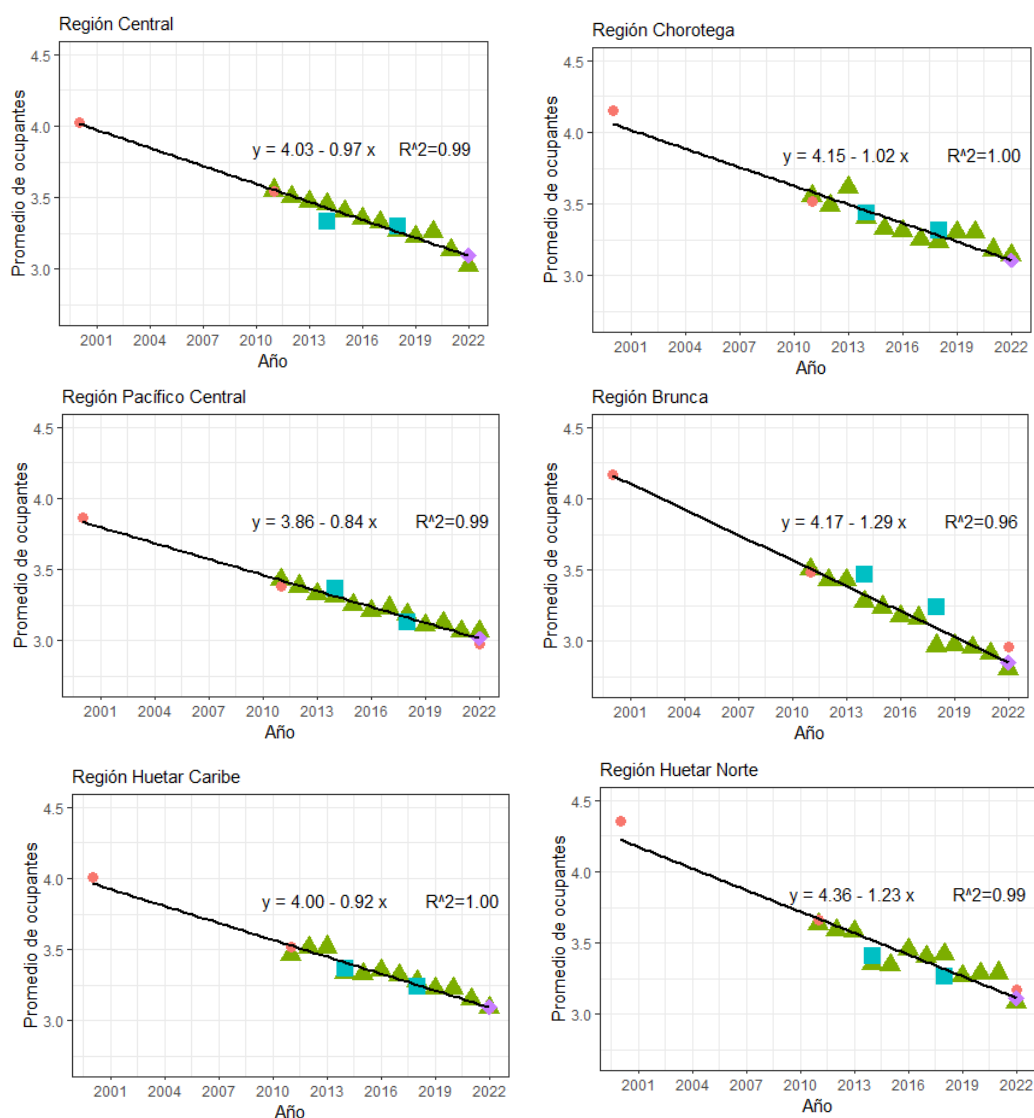
Fuente: INEC-Costa Rica. Encuesta Nacional de Hogares 2011 - 2022. Encuesta Nacional de Ingresos y Gastos de los Hogares 2014 y 2018. IX Censo Nacional de Población y V de Vivienda 2000. X Censo Nacional de Población y VI de Vivienda 2011. Proyecto Censal 2022.

<sup>1</sup> La región de planificación es la desagregación geográfica menor que permite la muestra de las encuestas analizadas.

Asimismo, en el gráfico 4.2 se puede observar que para cada región la pendiente es negativa, siendo la región Brunca la que presenta un mayor decrecimiento a lo largo de los años para el promedio de personas por vivienda.

## GRÁFICO 4.2

### Promedio de ocupantes por vivienda según región de planificación, 2000 – 2022



Fuente: INEC-Costa Rica. Encuesta Nacional de Hogares 2011 - 2022. Encuesta Nacional de Ingresos y Gastos de los Hogares 2014 y 2018. IX Censo Nacional de Población y V de Vivienda 2000. X Censo Nacional de Población y VI de Vivienda 2011. Proyecto Censal 2022.

Como se puede observar, los promedios obtenidos después del análisis de regresión se encuentran a nivel de región de planificación, debido a que es la desagregación geográfica menor que alcanzan las encuestas.

Para poder desagregar los promedios a nivel de cantón se realizó el cálculo de la población, las viviendas y el promedio de ocupantes por cantón, utilizando la tasa de crecimiento exponencial. Para ello, se tuvo que dividir los cantones en dos grupos, debido a la constante actualización de la división territorial administrativa. El primero grupo contenía a los cantones coincidentes entre 2000 y 2011 y el segundo a los que fueron creados después del año 2000.

Para el primer grupo, se aplicó la tasa de crecimiento exponencial, la cual supone que el crecimiento es en forma continua para cada año transcurrido, utilizando la siguiente fórmula:

$$r = \frac{1}{k} \ln \frac{N_f}{N_i}$$

Donde:

- $k$  es el tiempo transcurrido entre ambos censos (11 años)
- $N_f$  es la población final (Censo 2011)
- $N_i$  la población inicial (Censo 2000).

Posteriormente se realizó la proyección para las viviendas, la población y el promedio de personas por vivienda para 2022 utilizando la siguiente fórmula:

$$N_{2022} = N_{2011} * e^{(r*t)}$$

Donde:

- $N_{2022}$  es la cantidad proyectada para el año 2022
- $N_{2011}$  la obtenida del Censo 2011.

Para el segundo grupo, se asignó la tasa de crecimiento de los cantones que les dieron origen y así poder calcular la población, viviendas y promedio de personas actualizadas a 2022. Dado que estos no tenían registros de población y viviendas para ambos periodos, se calculó el porcentaje de población que representa cada uno en el proyecto censal y con base en esa cantidad, se procedió a segregar las viviendas y la población del o los cantones de origen con el fin de obtener la población y viviendas proyectadas.

Con dicha información se asumió que los cantones se comportan igual a la región de planificación a la que pertenecen y así se aplicó la siguiente fórmula:

$$f = \bar{X}_{(\text{regresión}_r)} / \bar{X}_{(\text{exponencial}_r)}$$

Donde:

- $\bar{X}_{(\text{regresión}_r)}$ : promedio de personas por vivienda obtenido de la regresión lineal.
- $\bar{X}_{(\text{exponencial}_r)}$ : promedio de personas por vivienda estimado a partir de la tasa de crecimiento exponencial.
- r: región de planificación

Seguidamente, se procedió a calcular el promedio de personas por vivienda final mediante la siguiente fórmula:

$$\bar{X}_{\text{final}} = \bar{X}_{(\text{exponencial}_c)} * f; \text{ donde } c \text{ corresponde al cantón.}$$

### 4.1.2 Ajuste de clasificación en las viviendas desocupadas

El Censo 2022 recopiló información de 795 860 viviendas ocupadas y contabilizó 476 000 viviendas pendientes o rechazadas, lo que daría un total de 1 271 860 viviendas con residentes habituales. Además, se identificaron 278 142 viviendas desocupadas, lo que representa un 17,9%. En el Censo 2011 el porcentaje de viviendas desocupadas fue de 10,8%.

La diferencia de las viviendas desocupadas entre 2011 y 2022 fue de casi un 40%, la cual podría ser mayor si se considera que entre las viviendas no visitadas también pudo haber viviendas desocupadas. Por tanto, se decidió realizar un análisis más minucioso de las viviendas desocupadas para validar el dato o bien, identificar si pudo haber existido una mala clasificación de esas viviendas.

Según los censos 2000 y 2011, la mayor proporción de viviendas desocupadas estaban disponibles para alquilar o vender, seguidas de las temporales para vacacionar, trabajar o estudiar. Si bien la categoría “Desocupada otra” considerada una categoría residual tenía un porcentaje importante en el 2000 (20,4%), para 2011 mostró una reducción (18,1%). Dicha categoría se refiere a otros motivos de desocupación no contemplados como las viviendas abandonadas.



Para el Censo 2022 la distribución de viviendas desocupadas mostró un comportamiento no esperado, donde las viviendas para alquilar o vender y las temporales disminuyen su peso relativo y aumentó el de la categoría “Otra”. Dicho comportamiento no es coherente con las tendencias inmobiliarias del país donde ha habido un incremento de construcciones tanto residenciales como vacacionales. En el cuadro 1 se presenta la distribución de los tipos de viviendas desocupadas para los últimos tres censos.

#### CUADRO 4.1

##### Distribución de las viviendas desocupadas por tipo, 2000 y 2011.

Tipo de desocupación	Censo 2000		Censo 2011		Censo 2022	
	Absoluto	Relativo	Absoluto	Relativo	Absoluto	Relativo
<b>Desocupadas</b>	<b>98 650</b>	<b>100,0</b>	<b>147 204</b>	<b>100,0</b>	<b>278 142</b>	<b>100,0</b>
Desocupada para alquilar o vender	31 559	32,0	51 426	34,9	92 797	33,4
Desocupada Temporal (Vacacionar, trabajadores)	30 404	30,8	50 081	34,0	78 842	28,3
Desocupada en construcción o reparación	16 541	16,8	19 004	12,9	31 319	11,3
Desocupada otra	20 146	20,4	26 693	18,1	75 184	27,0

Fuente: INEC-Costa Rica. Censos 2000, 2011 y 2022.

Ante eso, se decidió realizar un estudio más detallado para identificar posibles causas del incremento de la categoría “Otro” en las viviendas desocupadas. Para ello se realizó un análisis comparado de todas las categorías de desocupación utilizando los datos del Censo 2000 y 2011 y además el Censo Nacional Agropecuario 2014, el índice de urbanidad y la información de las referencias cartográficas con el objetivo de correlacionar los datos de viviendas desocupadas con alguna posible razón de existencia.

Los resultados obtenidos demostraron que hay concentración de viviendas desocupadas para trabajadores en zonas donde se esperaría que eso suceda por la predominancia de actividades agropecuarias o industriales como por ejemplo la zona de Los Santos. También, hay coherencia con la concentración de viviendas denominadas “desocupadas para estudiantes” en distritos cercanos a centros universitarios. Además, hay comportamientos esperados de viviendas para alquiler, venta o en construcción o reparación en distritos donde ha incremento del desarrollo inmobiliario residencial como Carmen en San José o San Mateo en Alajuela, por ejemplo.

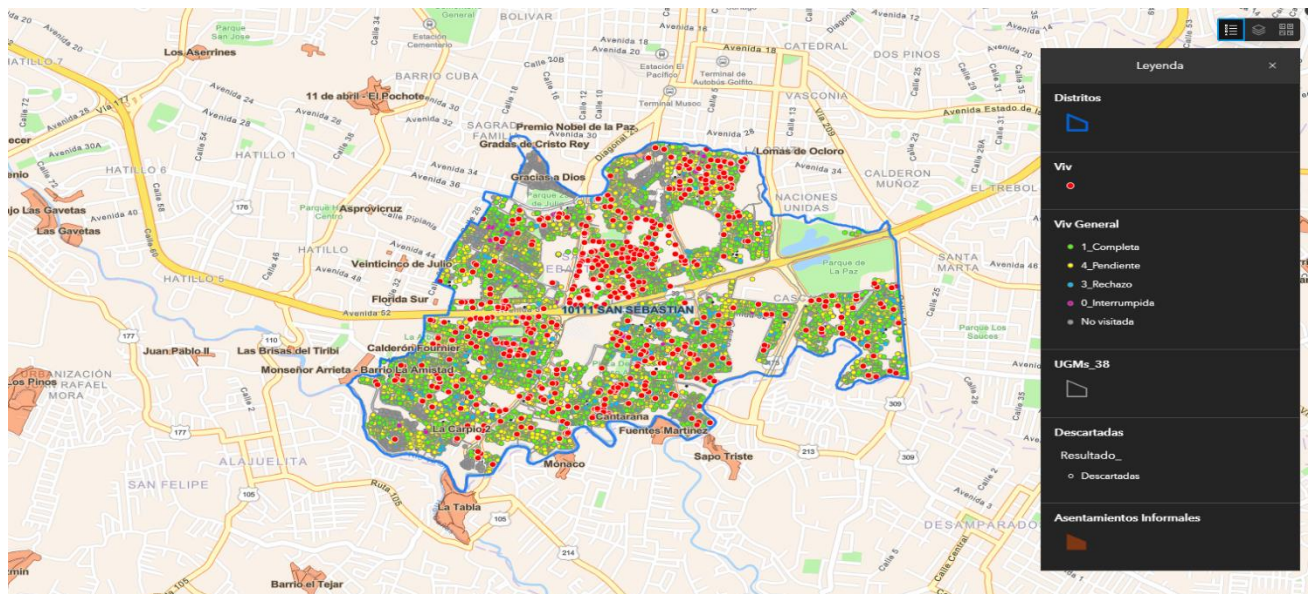
No obstante, la concentración de viviendas desocupadas por “otro motivo” no parece estar relacionada con alguna actividad económica en particular y, por el contrario, aparecen en zonas donde no debería haber concentración de viviendas desocupadas como asentamientos informales o territorios indígenas, por ejemplo, en Los Guido, Telire, Boruca o Brunka, donde más del 50% de sus viviendas desocupadas son por “Otro motivo”.

Para tener una mayor profundidad en el análisis se mapeó la distribución de las viviendas desocupadas y en particular, para algunos distritos atípicos, se hizo un análisis específico para las desocupadas por otro motivo a nivel de punto georreferenciado. Además, para encontrar posibles patrones explicativos en los distritos con resultados inconsistentes se identificaron concentraciones de viviendas desocupadas por otro motivo.

En el siguiente mapa se observa el ejemplo de ese análisis para el distrito San Sebastián. Los puntos rojos representan las viviendas desocupadas por otro motivo. Como se puede observar hay áreas con concentraciones importantes de ese tipo de viviendas, lo cual no es esperable en un distrito tan poblado, siendo esa una categoría residual.

**FIGURA 4.1**

**Distribución de puntos de viviendas por resultado de entrevista. Distrito San Sebastián, 2022**



Fuente: INEC-Costa Rica. Censo 2022.

Imagen tomada del visor de viviendas desocupadas. Los puntos rojos corresponden a las viviendas desocupadas por otro motivo.

Se presume que en esas zonas se pudo haber hecho una mala interpretación del uso de esa categoría. Por tanto, se concluye que la categoría “desocupada por otro motivo” fue utilizada de forma errónea por las personas censistas para asignar a viviendas que realmente eran ocupadas, pero no se pudo concretar una entrevista, por lo que debieron clasificarse como “pendientes”.

A partir del análisis de viviendas desocupadas se decide ajustar todas las viviendas “desocupadas por otro motivo” de manera que se tomarían como viviendas “no visitadas” y, por tanto, se estimaría si corresponden a viviendas efectivamente desocupadas o bien, ocupadas.

Además, también se tomaron como sospechosas aquellas viviendas desocupadas que se encontraran en UGM que sobrepasaron los 20 días de recorrido, ya que después de ese periodo era menos probable que quedaran viviendas desocupadas sin visitar y podrían deberse a una aplicación errónea y, similar a las desocupadas por otro motivo, tratarse realmente de viviendas pendientes.

Finalmente, tanto las viviendas desocupadas como aquellas identificadas en UGM recorridas después de 20 días, se ajustaron aplicando como factor la proporción de viviendas desocupadas del cantón. Las viviendas obtenidas quedaban categorizadas como desocupadas y el restante se reclasificaron como viviendas ocupadas.

El siguiente cuadro muestra la distribución final de las viviendas reclasificadas:

#### Cuadro 4.2

##### Distribución de las viviendas desocupadas reclasificadas. 2022

Tipo de vivienda desocupada	Clasificación		
	Total	Ocupadas	Desocupadas
Desocupada otra	75 184	65 690	9 499
Viviendas desocupadas extremo 20 días	38 999	34 603	4 396

Fuente: INEC-Costa Rica. Censos 2000, 2011 y 2022.

Después de los ajustes, el dato final de viviendas desocupadas fue de 214 654, lo que representa el 11,7% de las viviendas del país.

### 4.1.3 Clasificación de viviendas no visitadas

Para la estimación era necesario clasificar las viviendas que no fueron visitadas ya sea como ocupadas o bien, como desocupadas. Para ello, se aplicó el mismo factor utilizado en la etapa anterior (ajuste de viviendas desocupadas). Es decir, se utilizó como factor la proporción de viviendas desocupadas del cantón para asignar las viviendas desocupadas, el restante se clasificó como viviendas ocupadas.

En total hubo 218 823 viviendas no visitadas, de las cuales, 193 155 se clasificaron como viviendas ocupadas y 25 668 como viviendas desocupadas.

### 4.1.4 Estimación de la población

Con los ajustes realizados y los promedios de ocupantes por vivienda validados, se procedió a realizar la estimación de población la cual consistió en sumar todas las viviendas pendientes y rechazadas, además de las viviendas reclasificadas como ocupadas y multiplicarlas por el promedio de ocupantes por vivienda de cada cantón.

Una vez estimada esa población, se sumó al total de personas efectivamente censadas tanto en vivienda individuales como colectivas, para obtener una población estimada por cantón. La población nacional estimada con este método fue 4 864 442.

## 4.2 Estimación demográfica de población basada en registros administrativos

Esta metodología se basó en la utilización de diferentes registros administrativos como el padrón electoral, el maestro de nacimientos<sup>2</sup>, el maestro de defunciones<sup>3</sup>, las bases de datos de nacimientos y defunciones del INEC y la base de entradas y salidas de la Dirección General de Migración y Extranjería (DGME).

---

<sup>2</sup> Base de datos que agrupa todos los nacimientos históricamente registrados en el país.

<sup>3</sup> Base de datos que agrupa todas las defunciones históricamente registrados en el país.

La estimación demográfica se realizó a partir de la construcción de una población base en diferentes etapas, las cuales se explican a continuación.

### 4.2.1 Estimación de la población nacional

La estimación de la población nacional se realizó en dos etapas, en la primera se estimó la población menor de 20 años y en la segunda la población de 20 años y más.

La población menor de 20 años se realizó mediante la reconstrucción de cohortes de nacimientos que constituyen esta población en 2022, para ello, se utilizaron las bases de nacimientos y defunciones del INEC, corregidas por registro tardío, siguiendo el siguiente procedimiento:

1. Se elaboró un diagrama de Lexis, mediante el cual se estimó la población de 0 a 19 años al 01 de enero de 2022 y al 01 de enero de 2023.
2. Para reconstruir esta población al 01 de enero de 2022 mediante el diagrama de Lexis se utilizaron los nacimientos del periodo 2002 - 2021 y las defunciones de menores de 20 años para ese mismo periodo, en el caso de la población al 01 de enero de 2023 se utilizaron los nacimientos y defunciones de menores de veinte años del periodo 2003 – 2023.
3. Para obtener la proporción de las defunciones de cada año que correspondían a cada cohorte se utilizaron los factores de Glover.
4. Se elaboró una reconstrucción para hombres y otra para mujeres y el total se obtuvo por sumatoria.
5. Para obtener la población a mitad de periodo (30 de junio de 2022) se utilizó la fórmula de crecimiento exponencial

$$P_f = P_i * e^{rt}$$

Donde:

- $P_f$  es la población menor de veinte años al 1 de enero de 2023,
- $P_i$  es la población menor de veinte años al 1 de enero de 2022,
- $r$  es la tasa de crecimiento entre ambos periodos y

- $t$  es el periodo al cual se requiere trasladar la población, que en este caso es mitad de año.

6. Para aplicar esta fórmula primero debió calcularse  $r$ , el cual se obtiene despejando la misma fórmula

$$r = \text{LN}(P_f/P_i)/t$$

La población nacional de 20 años y más se estimó a partir del padrón electoral, el cual se ajustó eliminando las defunciones, para esto se realizó una vinculación entre el padrón y el maestro de defunciones por número de cédula.

Finalmente, para obtener el total de población nacional se realizó la sumatoria de la población de 0 a 19 años y de 20 años y más. La población nacional estimada fue de 4 732 177 personas.

#### 4.2.2 Estimación de la población extranjera

La población extranjera se estimó en dos etapas, primero se estimó la población nicaragüense, debido a que es la más numerosa en el país y después, el resto de población extranjera.

En primer lugar, se obtuvo la razón de población nicaragüense y población costarricense para cada grupo de edad y sexo, tanto la observada en el Censo 2011 como en el Censo 2022. Esas relaciones se promediaron para cada grupo de edad y sexo. Posteriormente se multiplicaron esas razones específicas por cada grupo de edades por sexo para obtener el total de población nicaragüense.

El mismo procedimiento se realizó con la razón de población extranjera de otros países para obtener ese total de población. Finalmente, se sumaron las dos poblaciones para obtener la población extranjera total, la cual se estimó en 461 547 personas.

### 4.2.3 Estimación de la población emigrante costarricense

Para estimar a la población emigrante costarricense se utilizó la base de entradas y salidas de la DGME. Mediante la metodología de ciclos<sup>4</sup>, se tomaron únicamente los ciclos coherentes donde el último movimiento fue una salida con 6 meses<sup>5</sup> o más de haberse efectuado.

Un aspecto importante a señalar es que se limitó la cantidad de salidas a tres, con la finalidad de filtrar casos de personas que, al tener muchos movimientos, sugieren un perfil de turista o empresario internacional, más que de una persona emigrante.

A esta metodología se agregaron los ciclos de un solo movimiento (salida) la cual tuvo que haberse realizado seis meses o más desde la fecha de corte.

Las salidas de personas mayores a 100 años fueron prorrateadas, así como los casos que no contaban con fecha de nacimiento.

Debido a que las bases de entradas y salidas no contienen a las personas menores de edad, se realizó una estimación mediante una razón de población emigrante sobre población costarricense residente, derivada del Censo 2022, dado que en el censo 2011 no se contaba con la información por sexo y edad de las personas emigrantes.

Para evaluar la razón del Censo 2011 se estimó la misma razón, pero con los datos de emigrantes mayores de 20 años y la población costarricense residente del padrón electoral. Esta estimación se comparó con la del censo y se calculó la diferencia, es decir, en cuanto el censo subestimaba esta relación.

Finalmente, para estimar la población emigrante menor de edad se usó la razón emigrantes/residentes, ajustada por el dato de la mujeres de 20 - 24 años, se utilizó por ser el grupo más próximo y partiendo del supuesto que la mayoría de niños y niñas migra con su madre. La población emigrante estimada fue de 149 527 personas.

---

<sup>4</sup> Un ciclo se define como una secuencia de dos movimientos, por ejemplo: entrada-salida, o salida-entrada

<sup>5</sup> Con respecto a la fecha de corte, que es el 30 de junio de 2022.

#### 4.2.4 Estimación de la población total

Para obtener la población total se aplicó la ecuación compensadora, mediante la siguiente fórmula:

$$\text{Población total} = \text{población nacional} + \text{población extranjera} - \text{población emigrante}$$

La población final estimada fue de 5 044 197 personas para junio 2022.

#### 4.2.5 Desagregación de la población por provincia y cantón

Para identificar la mejor forma de distribuir la población entre las provincias y cantones, se analizaron diferentes fuentes demográficas tales como: nacimientos, defunciones, padrón electoral, estimaciones y proyecciones de población vigentes; así como la distribución derivada de la aplicación del modelo bayesiano que se explicará más adelante.

Finalmente, se determinó que la estructura más adecuada surgía del promedio de las estructuras de los nacimientos y las defunciones para los años 2019 (pre pandemia) y 2022 (post pandemia) y el padrón al 30 de junio de 2022, así como la estructura derivada del modelo bayesiano, para obtener una estructura modelo.

Con esta estructura modelo se aplicó un factor a la población nacional para obtener la población desagregada por cantones. La población por provincia fue el resultado de la sumatoria de los cantones.



### 6.3 Modelo bayesiano para estimación de población<sup>6</sup>

Mediante reuniones semanales, con el liderazgo y la participación activa del INEC, la División de Estadísticas de CEPAL, junto con CELADE y UNFPA-LACRO, realizaron una serie de estudios sobre poblaciones modeladas que dieron como resultado la definición de tres modelos estadísticos que hicieron uso integrado de diferentes fuentes de información; esto es, el pre-censo, el censo, registros administrativos e información geoespacial<sup>7</sup> que pueda ayudar a predecir la distribución espacial de la población y su densidad. La consecución de la información geoespacial se realizó respetando los límites de las unidades geoestadísticas mínimas (UGM), que fungieron como sectores censales y áreas de empadronamiento en este Censo. A continuación, se detallan los pormenores de los modelos:

1. En una primera instancia se definió un modelo estadístico para la predicción de la probabilidad de que una vivienda estuviese ocupada. Debido a que en el levantamiento censal hubo una gran cantidad de viviendas que nunca atendieron el llamado de las personas encuestadoras, o rechazaron ser entrevistadas, o a las que simplemente nunca se llegó, fue necesario asignar una probabilidad de ocupación sobre estas viviendas. El modelo utilizado fue un modelo bayesiano mixto con respuesta binomial y efectos aleatorios sobre las provincias, cantones y distritos, tal como se detalla a continuación:

$$U_{ij} \sim \text{Binomial}(p_{ij})$$

$$\text{logit}(p_{ij}) = x'_{ij}\beta + z'_{ij}\gamma$$

En donde  $U_{ij}$  es una variable dicotómica que refiere al estado de ocupación de la vivienda,  $p_{ij}$  es la probabilidad de que la  $i$ -ésima vivienda de la  $j$ -ésima UGM está ocupada,  $x'_{ij}\beta$  corresponde a los efectos fijos considerados en el modelo, mientras que  $z'_{ij}\gamma$  a los efectos aleatorios. Las distribuciones previas de los parámetros de los modelos bayesianos se consideraron no informativas. Con este modelo es posible

<sup>6</sup> La descripción de esta metodología se tomó de forma íntegra de la nota técnica elaborada por CEPAL, CELADE y UNFPA con el fin de evitar cualquier distorsión de los procedimientos ejecutados.

<sup>7</sup> Entre las variables consideradas en los modelos se encuentran zona (urbano, rural), peligrosidad, problemas de acceso, riesgos o amenazas, cobertura en telecomunicaciones, permisos de construcción, población adscrita al sistema de salud de la Caja Costarricense del Seguro Social, población matriculada en educación primaria, fracción de suelo urbano, fracción de suelo cultivado, asentamiento informal, elevación, territorios pertenecientes a grupos indígenas, áreas protegidas, tiempo promedio de traslado a las escuelas e iglesias, intensidad de luces nocturnas, entre otras.

definir una estimación sobre el número de viviendas desocupadas en cada desagregación de la malla geográfica del país.

2. En un segundo paso, es necesario realizar la predicción del número de personas que habitan en las viviendas que no se visitaron, que estaban ocupadas, pero con habitantes ausentes, o que rechazaron responder al cuestionario censal. El modelo bayesiano utilizado en esta asesoría se asemeja a los modelos del Proyecto WorldPop de la Universidad de Southampton en Inglaterra. El supuesto fundamental de este modelo es que el número de personas en una determinada UGM sigue una distribución de Poisson. A continuación, se presenta una descripción básica del modelo:

$$Y_{ij} \sim \text{Poisson}(N_{ij} D_j)$$

$$\log(D_j) = x'_j \beta + z'_j \gamma$$

En donde  $Y_{ij}$  es el número de personas que habitan en la  $i$ -ésima vivienda de la  $j$ -ésima UGM,  $N_{ij}$  es el número de viviendas en esa UGM (conocido desde el censo y el precenso),  $D_j$  es la densidad poblacional promedio en la UGM, la cual se relaciona por medio de la función logaritmo con los correspondientes efectos fijos  $x'_j \beta$  y los efectos aleatorios  $z'_j \gamma$ . Asimismo, las distribuciones previas de los parámetros de los modelos bayesianos se consideraron no informativas. Con este modelo es posible definir una estimación sobre el número de viviendas desocupadas en cada desagregación de la malla geográfica del país.

3. En el último paso, se realiza una modelación sobre la estructura demográfica de las UGM a nivel de provincia. Esta estructura está representada en 40 subgrupos poblacionales, resultantes del cruce entre sexo y quinquenios de edad. La inferencia sobre las viviendas no observadas se realiza mediante una predicción condicional sobre los resultados encontrados en el paso anterior. Por la naturaleza del fenómeno estudiado, se consideró apropiado utilizar una distribución multinomial como verosimilitud para el conteo sobre los 40 grupos. A continuación, se presenta una descripción del modelo:

$$G_i \sim \text{Multinomial}(\mathbf{p}_{G_i})$$

$$\text{logit}(\mathbf{p}_{G_i}) = \mathbf{x}'_i \boldsymbol{\beta}$$

En donde  $G_i$  es el conteo sobre cada uno de los cuarenta grupos poblacionales, mientras que  $\mathbf{p}_{G_i}$  representa un vector de tamaño 40 conteniendo las probabilidades de que en una vivienda habite una persona clasificada en cada grupo.

En la definición de cada uno de los modelos se tuvieron en cuenta los siguientes pasos:

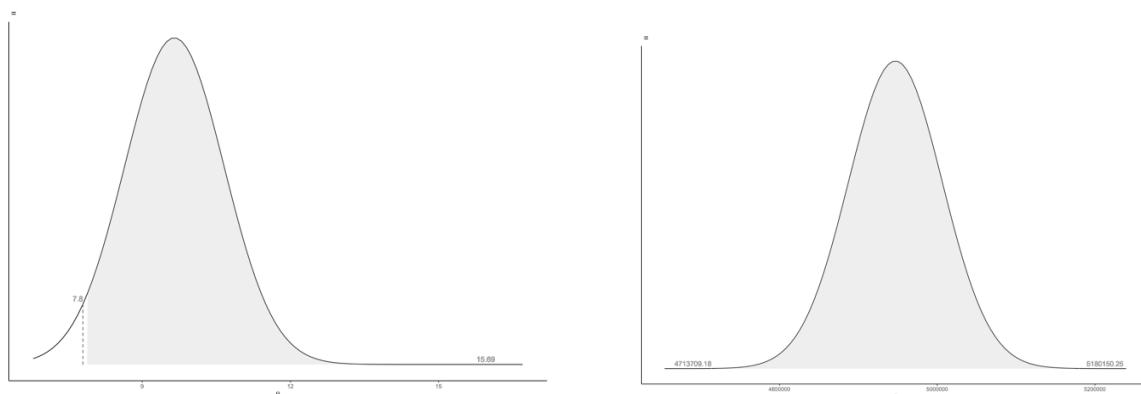
- Recopilación de datos: incluyendo el número de viviendas observadas en cada UGM en el precenso, los datos censales del operativo presencial y del operativo virtual, información geoespacial y de registros administrativos.
- Preparación de los datos: en esta etapa los datos recopilados se analizaron y se evaluó su calidad e idoneidad para su análisis. Esto incluyó la limpieza de los datos, la estandarización de formatos, la eliminación de datos faltantes o inconsistentes, la transformación y estandarización de los datos para su procesamiento posterior.
- Consolidación del tipo de información: en este paso se clasificó cada una de las viviendas en las siguientes seis categorías: desocupadas (7.36%), ocupadas (43.7%), desocupadas censadas en papel (0.08%), ocupadas censadas en papel (1.32%), ocupadas sin información sobre el número de habitantes (26.89%), sin ninguna clase de información (20.66%). La predicción sobre el número de habitantes (modelo Poisson) se realizó sobre estas dos últimas categorías, mientras que la predicción del modelo de desocupación de viviendas (modelo Binomial) se llevó a cabo sobre la primera categoría.
- Definición del modelo: a continuación, se definieron los modelos bayesianos que describen la relación entre las variables observadas, las covariables (efectos fijos y aleatorios). Esto implicó seleccionar distribuciones de probabilidad apropiadas que representan la incertidumbre y las relaciones entre las variables.
- Estimación de parámetros: una vez definido el modelo, se procede a estimar los parámetros del mismo utilizando los softwares R y STAN con técnicas de inferencia bayesiana (Monte Carlo Markov Chain, MCMC). Estas técnicas permiten obtener estimaciones de los parámetros y sus intervalos de credibilidad.

- Validación del modelo: en esta etapa se evaluó la calidad y la precisión del modelo desarrollado. Esto implicó realizar comparaciones de las estimaciones del modelo con datos observados independientes y realizar análisis de sensibilidad para evaluar cómo las diferentes fuentes de incertidumbre afectan los resultados.
- Generación de mapas de población y gráficos de incertidumbre: una vez validado el modelo, se puede proceder a generar mapas de estimación de población a partir de los datos y los parámetros estimados. Estos mapas representan la distribución espacial de la población y contienen gráficos que indican la incertidumbre del modelo de población.

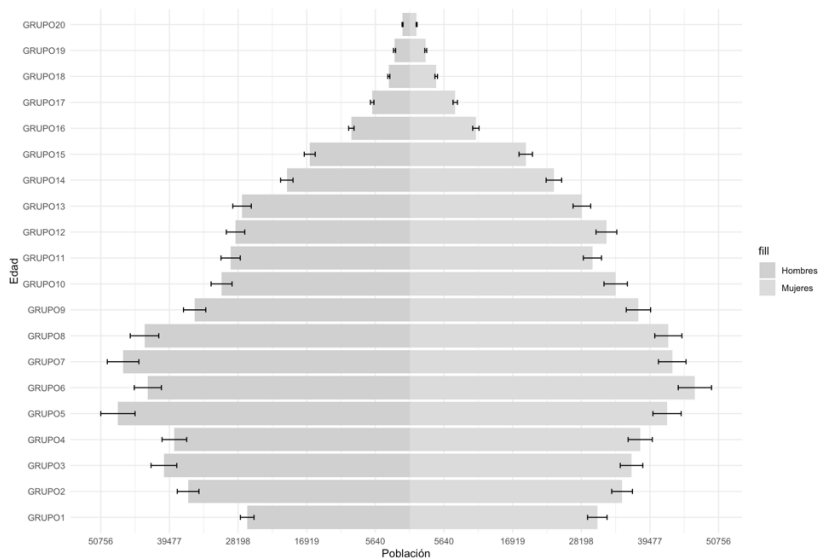
Utilizando este tipo de modelación se llegó a una estimación de 170.596 viviendas desocupadas, esto es 9.65% del total de viviendas en el territorio nacional con un intervalo de credibilidad de entre 7.80% y 15.7% (Figura 1, izquierda). Además, en este esfuerzo conjunto con el INEC, se estimó que el total de habitantes en Costa Rica asciende a 4.946.930 con un intervalo de credibilidad entre 4.713.709 y 5.180.150 (Figura 1, derecha). Por último, se muestra la distribución por edad y sexo, con sus respectivos intervalos de credibilidad en cada uno de los 40 grupos (Figura 2). Por medio de este acercamiento inferencial es posible desagregar las estimaciones a nivel provincial, cantonal y distrital.

#### Figura 4.2

**Estimación del número de viviendas desocupadas (izquierda) y del número de habitantes (derecha) en Costa Rica.**



**Figura 4.3**  
**Estimación de la estructura poblacional (hombres y mujeres en edades quinquenales) en Costa Rica.**



El trabajo y resultados de los modelos desarrollados representa insumos que el INEC incorpora en su proceso para verificación y comparación con las diversas estimaciones realizadas de forma independiente por parte de sus equipos técnicos.

## 6.4 Validación de resultados y definición de estimación de población final

El proceso de elaboración de las estimaciones de población se desarrolló de forma integrada por tres equipos de trabajo, los cuales contaron con el acompañamiento valioso del CELADE y el CCP que, a través de sus personas expertas, asesoraron al equipo del INEC en la implementación de los cálculos. Además, el modelo bayesiano se trabajó con los expertos de la División Estadística de la CEPAL y UNFPA. Los equipos del INEC estuvieron conformados por personas de la Unidad de Estadísticas Demográficas, Unidad de Cartografía y del proyecto Censo 2022. Por tanto, durante todo el proceso se dio una validación constante de los resultados parciales con sesiones de trabajo, tanto internamente como con los expertos internacionales.

Al finalizar la aplicación de las metodologías, se compararon las tres estimaciones y lo observado en el Censo 2022 para identificar posibles inconsistencias tanto en la distribución de la población por cantón, como por grupos de edad y sexo. También se revisaron los intervalos de confianza generados del modelo bayesiano, donde se valoraron, a nivel de cantón, cómo se comportaron las estimaciones.

Si bien las tres estimaciones se consideran robustas y establecen un rango de población estimada, se determina que la población final será la obtenida bajo la metodología demográfica basada en registros. Esto porque es la que se basa en datos observados de los registros administrativos, los cuales a su vez han sido evaluados previamente y se sabe de su completitud y calidad. Además, es la información que servirá de fundamento al cálculo de las proyecciones de población.

## **6.5 Estimación de viviendas**

Una vez definida la población estimada, se procedió a analizar las viviendas. Para ello se tomó como base los promedios de ocupantes por vivienda estimados y validados para la metodología basada en datos censales. Esto porque, como se comentó anteriormente, fueron analizados en profundidad, comparados con diversas fuentes recientes y del pasado y se pudo determinar que eran los más adecuados.

Siendo así, se procedió a estimar la cantidad de viviendas ocupadas necesarias para el total de población definido, con los promedios de ocupantes calculados. De esa manera se estimó un total de 1 621 381 viviendas ocupadas.

A partir de la validación de viviendas censadas y del análisis realizado con la distribución de las viviendas desocupadas y no visitadas, explicado en la metodología de estimación basada en datos censales, se estableció un total de 1 554 171 viviendas ocupadas en el Censo 2022, por tanto, hubo una sub cobertura de 67 210 viviendas ocupadas.

Para las viviendas desocupadas se tomaron las mismas viviendas analizadas y clasificadas en la metodología de estimación basada en datos censales, que incluía las viviendas identificadas en el Censo 2022, la reasignación de las viviendas desocupadas por otro motivo

y las viviendas no visitadas, así como de aquellas ubicadas en UGM que se recorrieron en más de 20 días.

Además, en nueve cantones<sup>8</sup> se mantuvieron únicamente las viviendas desocupadas identificadas en el Censo 2022, porque eran cantidades importantes que ya no requerían mayor ajuste. El total de viviendas desocupadas fue 214 654.

Finalmente, el total de viviendas del país se estimó en 1 836 291.

---

<sup>8</sup> Escazú, Santa Ana, Montes de Oca, Curridabat, Jiménez, Barva, Santo Domingo, Liberia y Carrillo.

## 7. Bibliografía

- Cebrián, E. (2011). Desafección, apatía y participación. La desafección activa como forma de participación política. *Revista de Gestión Pública y Privada*, 71-80.
- Contraloría General de la República. (Julio de 2022). *CGR.go.cr*. Obtenido de <https://sites.google.com/cgr.go.cr/ipp2020-situacionyperspectivas/implementaci%C3%B3n-de-la-regla-fiscal-en-costa-rica>
- Dirección General de Migración y Extranjería. (2022). Base de datos de entradas y salidas migratorias. Costa Rica.
- Hillygus, Nie, & al., e. (2006). *The Hard Count*. New York: Russell Sage Foundation.
- INEC. (2000). Censo 2000. Costa Rica.
- INEC. (2010-2022). Encuesta Continua de Empleo. Costa Rica.
- INEC. (2010-2022). Encuesta Nacional de Hogares. Costa Rica.
- INEC. (2011). Censo 2011. Costa Rica.
- INEC. (2014 y 2018). Encuesta Nacional de Ingresos y Gastos de los Hogas. Costa Rica.
- Registro Civil. (Junio de 2022). Padrón electoral. Costa Rica.
- Registro Civil. (Maestro a junio 2022). Base de datos de defunciones. Costa Rica.
- Registro Civil. (Maestro a junio 2022). Base de datos de nacimientos. Costa Rica.
- Romano, A. (2015). *La apatía política en democracia*. Ecuador: FLACSO.





[www.inec.cr](http://www.inec.cr)

INEC, de la rotonda de La Bandera 450 metros oeste, sobre calle Los Negritos,  
edificio Ana Lorena, Mercedes de Montes de Oca, Costa Rica.

Correo e.: [informacion@inec.go.cr](mailto:informacion@inec.go.cr) Apartado: 10163 - 1000 San José, C. R.

Teléfonos: 2527 - 1144, 2527 - 1145, 2527 - 1146 y 2527 - 1147