

Informes de Estadística **Sociodemográfica Aplicada**

Número 4

**Registro Estadístico Base Población
REBP 2018**



El futuro
es de todos

Gobierno
de Colombia

ISSN: 2805-6345 (en línea)

**DEPARTAMENTO ADMINISTRATIVO
NACIONAL DE ESTADÍSTICA
DANE**

JUAN DANIEL OVIEDO ARANGO

Director

RICARDO VALENCIA RAMÍREZ

Subdirector

MARÍA FERNANDA DE LA OSSA ARCHILA

Secretaria General

DIRECTORES

HORACIO CORAL DÍAZ

Dirección de Metodología y Producción Estadística

ANGELA PATRICIA VEGA LANDAETA

Dirección de Censos y Demografía

JULIETH ALEJANDRA SOLANO VILLA

Dirección de Regulación, Planeación,
Estandarización y Normalización

JUAN PABLO CARDOSO TORRES

Dirección de Síntesis y Cuentas Nacionales

SANDRA LILIANA MORENO MAYORGA

Dirección de Geoestadística

MAURICIO ORTIZ GONZÁLEZ

Dirección de Difusión, Mercadeo y Cultura
Estadística

EQUIPO DE TRABAJO

Dirección de Censos y Demografía

Lelio Alejandro Arias Vizcaino
Ana Berenice Heredia Gutiérrez
Ana Cecilia Olaya Cisa
Andryu Enrique Mendoza
Yenny Andrea Marín
Steven Cifuentes Rúgeles
Glen Harry Amaya Cruz
Angela Patricia Vega Landaeta

**Dirección de Difusión, Comunicación
y Cultura Estadística**

Edición de contenido y diseño
Brandon Steve Rojas
Gladys Adriana Quintero Hernández
Maria Fernanda Reyes Vargas

**Dirección de Geoestadística
Producción de mapas temáticos**

Juan Camilo Aros Castillo

© DANE, 2021

Prohibida la reproducción total o parcial sin permiso
o autorización del Departamento Administrativo
Nacional de Estadística, Colombia.

CONTENIDO

Introducción	5
01 Conformación del Registro Estadístico Base de Población - REBP	8
Integración de los registros administrativos	9
Fuentes de información para la conformación del REBP	10
Definición de variables demográficas básicas	11
Definición de la variable estado	12
02 Análisis de resultados	14
Emparejamiento del REBP 2018 con el CNPV 2018	15
Análisis nacional, comparación estructuras poblacionales CNPV y REBP	16
Indicadores nacionales REBP frente a CNPV	18
Análisis cobertura departamental	19
Análisis cobertura municipal	21
Indicadores departamentales calculados con REBP 2018 en comparación con CNPV 2018	23
03 Visualizaciones	26
Conclusiones	28
Bibliografía	30
Anexos	31
Anexo 1: Comparación de población por departamentos según sexo, CNPV2018 – REBP 2018	31
Anexo 2: Comparación de población por departamentos y capitales según sexo, CNPV2018 – REBP2018	33
Anexo 3. Orden de integración de los registros administrativos	35

LISTA DE TABLAS

Tabla 1. Censo de personas por estado REBP.....	13
Tabla 2. Censo de población ajustado por estado en el REBP.....	17

LISTA DE GRÁFICAS

Gráfico 1. Proceso de emparejamiento e integración de registros.....	9
Gráfico 2. Emparejamiento por estado REBP-CNPV.....	15
Gráfico 3. Pirámide de población, comparación CNPV - REBP 2018.....	16
Gráfico 4. Pirámide de población del CNPV 2018 contra el REBP 2018 después del ajuste del estado para las mujeres de 20 a 34 años.....	17
Gráfico 5. Indicadores demográficos obtenidos desde el REBP y su comparación con el CNPV.....	18
Gráfico 6. Diferencias relativas departamentales CNPV-REBP y ECV-REBP.....	19
Gráfico 7. Análisis de comparabilidad razón de dependencia REBP-CNPV 2018.....	23
Gráfico 8. Análisis de comparabilidad, índice de masculinidad CNPV 2018 - REBP.....	24
Gráfico 9. Análisis de comparabilidad, índice de envejecimiento CNPV 2018 - REBP.....	25

LISTA DE ILUSTRACIONES

Ilustración 1. Diferencias relativas departamentales CNPV-REBP y ECV-REBP.....	20
Ilustración 2. Diferencias relativas municipales CNPV-REBP.....	22
Ilustración 3. Captura de pantalla con la información del geovisor REBP y CNPV.....	27

INTRODUCCIÓN

Los países escandinavos (Suecia, Noruega, Finlandia, Dinamarca, Islandia), han sido pioneros en el uso de registros administrativos con fines estadísticos. La difusión de su experiencia en la región motivó la exploración de los modelos basados en registros, para la producción de estadísticas sociodemográficas, la enumeración y la caracterización de la población. Estos modelos ofrecen cualidades que los hacen atractivos: nuevas oportunidades de análisis, reducción de los costos que acarrearán las operaciones estadísticas tradicionales como censos y encuestas, menor carga a las fuentes de información, mayor cobertura, facilidad en los análisis longitudinales, entre otros. Las bondades de estos modelos son las diferencias metodológicas entre los dos sistemas de producción; los tipos de marcos (de áreas o de listas) y los métodos de recolección de datos.

Para citar un ejemplo, el Instituto Nacional de Estadística de Suecia, avanzado en la aplicación de estos modelos, considera el registro de población como el registro estadístico más importante, debido a que contiene a todas las personas registradas como residentes permanentes en Suecia, identificadas plenamente a través de sus números de identidad, lo que posibilita la vinculación con todos los registros disponibles.

El Departamento Administrativo Nacional de Estadística de Colombia, -DANE, no es ajeno a este tipo de ejercicios, desde hace un buen tiempo viene aprovechando los registros administrativos con fines estadísticos (entre ellos la medición de diferentes fenómenos socioeconómicos) para la producción de estadísticas vitales, estadísticas de educación formal, de comercio exterior y el directorio de empresas, entre otros usos alternativos. Aunque el propósito de los registros administrativos no es la producción de estadísticas, su cobertura nacional,

la información que proporcionan sobre la población residente en Colombia, las variables que permiten su caracterización sociodemográfica y dada la actualización de los reportes administrativos según la oferta y demanda de bienes y servicios de carácter esencial como el acceso a salud, a la educación y a la identificación ciudadana, los convierten en una fuente potencial de información.

Si bien el DANE no había incursionado en la consolidación de varias fuentes que permitieran la configuración de un gran registro informativo de las características de la población, motivado por las nuevas formas de hacer estadísticas de población, en 2010 propuso una estrategia encaminada a la conformación del registro de población a partir de la consolidación de diferentes fuentes. Este proyecto se fortalece posteriormente cuando se le identifica como insumo para solucionar la necesidad de actualizar la información poblacional en periodos intercensales el cual superó los 10 años, entre los dos últimos censos (2005 y 2018) y da inicio al Registro Estadístico Base de Población -REBP, el cual se presenta en este documento. Así mismo, el REBP se consolida a través de proyectos de cooperación técnica enfocados en desarrollar acciones metodológicas que permitan articular el Sistema Estadístico Nacional - SEN, para potenciar el aprovechamiento estadístico de la información captada a través de registros administrativos, usando la ciencia de datos como el lenguaje de comunicación entre diferentes entidades estatales.

El inicio de la estrategia encaminada al aprovechamiento de registros administrativos para el uso estadístico, específicamente en estadísticas demográficas, se probó en 2015 utilizándolo como insumo para la verificación de la cobertura y la calidad de Registro Estadístico de Población del Archipiélago de San Andrés,

Providencia y Santa Catalina (2012-2014), este ejercicio se fundamentaba en la necesidad de cuantificar y caracterizar a la población Raizal, identificar la apropiación y transmisión a las nuevas generaciones de sus tradiciones culturales (usos, costumbres y valores) y, contar con información base para la planificación y la toma de decisiones en política pública.

Las pruebas técnicas desarrolladas en este proceso, fueron compartidas con éxito en eventos regionales de los países andinos, posicionándolo como un ejercicio innovador y confirmando que la integración de datos administrativos para efectos estadísticos es un insumo de calidad de gran utilidad.

En 2016, utilizando registros administrativos, se obtuvo el registro estadístico de Jamundí (Departamento del Valle del Cauca), el cual se constituyó en insumo de contraste para los resultados del censo experimental que se llevó a cabo en ese municipio. En este caso, y dadas las características del registro, fue necesario implementar el uso de herramientas probabilísticas¹ para la identificación de personas duplicadas en el registro. En 2017, se conformó el Registro Estadístico de Población de Bogotá y Cundinamarca, que hasta ese momento, fue el ejercicio de mayor rigor técnico por el volumen de población del área geográfica analizada, lo cual implicó esfuerzos en recursos físicos (capacidad de máquinas y procesamiento) y de personal calificado.

Teniendo en cuenta las recomendaciones de las Naciones Unidas, sobre el uso de registros administrativos para la producción de estadísticas oficiales y el avance del proyecto BID “Estadística de población e inmuebles a partir del uso de registros administrativos oficiales en la comunidad Andina (BID, 2016)”²; los cuáles ofrecen el marco conceptual y metodológico a nivel regional, para el diseño, desarrollo e implementación de un Sistema Integrado de Registros Estadísticos de Población e Inmuebles - SIREPI,

en 2018 se conforma el Registro Estadístico Base de Población - REBP 2017 para el total nacional. Este insumo se conforma con información proveniente de los registros administrativos de la Base de Datos Única de Afiliados a Salud - BDUA, del Ministerio de Salud y Protección Social, las Entidades Promotoras de Salud - EPS; el Sistema de Selección de Beneficiarios para Programas Sociales - SISBEN producido por el Departamento Nacional de Planeación - DNP; el Sistema Integrado de Matriculas - SIMAT del Ministerio de Educación, el Registro de nacidos vivos y de defunciones la Planilla Integrada de Liquidación de Aportes - PILA del Ministerio de Salud y Protección Social, el Registro Único de Víctimas - RUV de la Unidad para la Atención y Reparación Integral a las Víctimas - UARIV y el Archivo Nacional de Identificación - ANI que es el centro de soporte del REBP.

Entre los múltiples usos que se le han dado, resalta la comparación y depuración de las variables de identificación incluidas en el Censo Nacional de Población y Vivienda - CNPV 2018, a través de algoritmos de comparación fonética. Desde ese mismo año, se avanza en la consolidación del REBP a nivel nacional, comparando los volúmenes y estructura resultante del CNPV 2018 analizando la inclusión de otras variables presentes en los registros administrativos que pueden ser comparables con la operación censal. También se conforma una tabla, denominada identificadores, en donde se almacenan todas las direcciones y teléfonos de las personas incluidas en los registros administrativos. Los datos que contiene esta tabla, así como los que se recopilan en la fase de recolección del proceso de producción de información estadística son confidenciales y su uso es exclusivo para fines estadísticos como lo establece la ley de Reserva Estadística (Ley 73 de 1993, artículo 4).

Igualmente, en el marco del aprovechamiento de registros administrativos para fines estadísticos, durante el operativo censal del CNPV 2018 y posteriormente para

¹ Ver (Christen, 2012)

² Ver Cooperación técnica regional BID

el análisis de consistencias y producción de indicadores demográficos, se utilizaron los registros administrativos de la Registraduría Nacional del Estado Civil - RNEC, la Base Única de Afiliados - BDUA y el Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales - SISBEN.

Además, en 2019 el REBP sirvió de insumo para el cálculo de la omisión censal a nivel de áreas menores, lo que lo convirtió en un referente técnico para la elaboración de las certificaciones oficiales de población.

Las múltiples soluciones que el REBP ha brindado, motivaron la conformación de un REBP año base 2018 (REBP2018), con una actualización periódica anual que se convierta en insumo para actualizar la enumeración y caracterización básica de la población, su estructura y distribución geográfica, así como también para la construcción de supuestos de la dinámica poblacional que permitan actualizar las proyecciones de población en el período intercensal.

En el presente documento se explica la conformación del REBP y su línea base REBP2018, los resultados obtenidos y la comparación con la información arrojada por el CNPV 2018 (información se recolectó de manera tradicional), como punto de referencia y validación.

Este Informe de Estadística Sociodemográfica Aplicada se distribuye en cinco apartados; en el primero, se describe el proceso de integración de registros administrativos REBP2018, el emparejamiento de los registros individuales y la construcción de las variables demográficas básicas. En el segundo apartado, se analizan los principales resultados de los indicadores demográficos obtenidos, usando como fuente el REBP2018, tanto a nivel nacional, como departamental y municipal; igualmente, se hace la comparación del REBP y el CNPV2018. En el cuarto y último apartado, se describen las herramientas propuestas para la visualización de resultados y finalmente, se presentan las conclusiones.



1

Conformación del

Registro Estadístico Base
de Población - REBP

Integración de los registros administrativos

El Registro Estadístico Base de Población - REBP, está compuesto por las variables que determinan las características principales de la población y que son básicas para la generación y análisis de estadísticas demográficas; así como la vinculación con otros registros estadísticos, en el marco de un sistema estadístico basado en registros.

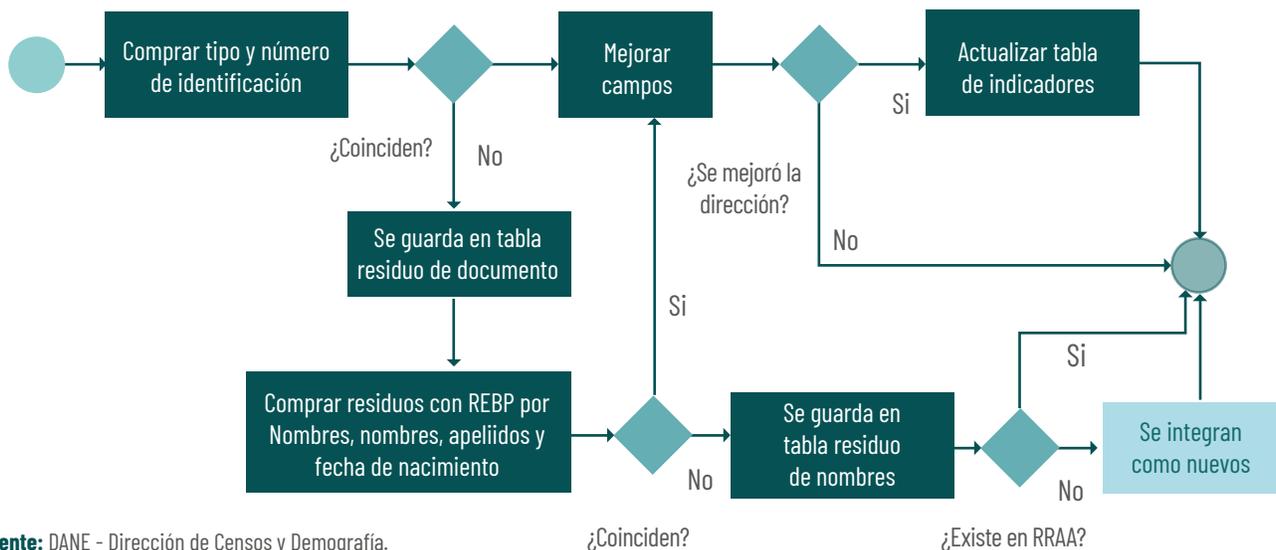
Para la construcción del REBP, es fundamental contar con los registros administrativos generados por cada institución del orden nacional y territorial. Inicialmente, se realiza la consulta en el inventario de registros administrativos del DANE con el fin de identificar los datos de población que satisfacen la necesidad de información. En caso de no contar con el registro, se identifica la entidad proveedora de información considerando las bases legales que garanticen el acceso a la información.

Así mismo, se revisan los conceptos y variables requeridas a través de una solicitud formal del DANE a las entidades responsables de los registros administrativos. Una vez se cuente con el acceso a las unidades y variables requeridas, estas se consolidan en una base de datos, donde se clasifica, codifica y estandariza

la información, de tal manera que permita: i) la correlación de tablas de referencia, ii) la homologación de nombres de variables y recodificación de categorías, iii) la creación de funciones para corrección de los caracteres especiales en los campos de identificación de las personas, iv) la separación de nombres y apellidos cuando son reportados en un mismo campo y v) el uso de expresiones fonéticas³ y conteo de la letras que contienen las variables de identificación.

Posteriormente, se lleva a cabo el emparejamiento determinístico de las fuentes de información, que consiste en identificar registros que corresponden a la misma persona en dos o más archivos distintos e integrar los cuando estos coinciden. Este proceso se realiza por medio de variables de vinculación, que es un término que engloba, tanto a variables de identificación, como a variables de ubicación (un vínculo entre dos registros consta de una o más variables de vinculación). El proceso completo de emparejamiento de los registros administrativos, para la conformación del Registro Estadístico Base de Población - REBP se puede observar con más detalle en el gráfico 1.

Gráfico 1. Proceso de emparejamiento e integración de registros



Fuente: DANE - Dirección de Censos y Demografía.

³ Debido a que los nombres y apellidos son susceptibles a presentar errores de escritura se hace relevante la construcción de una función que convierta una cadena de caracteres a su fonema (Unidad fonológica mínima que resulta de la abstracción o descripción teórica de los sonidos de la lengua)

La integración se realiza de forma ordenada, en donde el primer registro incorporado resulta ser el que contiene la mayor cantidad de registros (filas). La tabla 3 (ver anexos) muestra el orden de integración, así como el aporte de cada una de las fuentes empleadas y sus respectivas fechas de corte.

Fuentes de información para la conformación del REBP

Dadas las necesidades estadísticas del país, la Dirección de Censos y Demografía analiza los requisitos de información, identificando los registros administrativos útiles para la conformación del REBP. Para esto, se realiza un análisis previo de las posibles fuentes de datos administrativas (proveedores de información), la viabilidad y los mecanismos establecidos para su acceso. Como resultado de este proceso, a 2018 se cuenta con las siguientes fuentes de información:

Registros Administrativos - RRAA

- Archivo Nacional de Identificación – ANI de la Registraduría Nacional del Estado Civil - RNEC.
- Registro Civil de Nacimiento – RCN de la Registraduría Nacional del Estado Civil - RNEC.
- Registro Civil de Defunciones – RCD de la Registraduría Nacional del Estado Civil - RNEC
- Base de Datos Única de Afiliados al Sistema de Seguridad Social – BDUA, régimen contributivo y subsidiado del Ministerio de Salud y Protección Social - MSPS.
- Entidades Promotoras de Salud – EPS.
- Sistema de Selección de Beneficiarios para Programas Sociales – SISBEN del Departamento Nacional de Planeación - DNP.
- Sistema Integrado de Matriculas – SIMAT del Ministerio de Educación Nacional - MEN.
- Sistema Nacional de Información de la Educación Superior – SNIES del Ministerio de Educación Nacional - MEN.
- Sistema de Información de Educación para el Trabajo – SIET del Ministerio de Educación Nacional - MEN.
- Planilla Integrada de Liquidación de Aportes a Seguridad Social – PILA de cotizantes y pensionados del Ministerio de Salud y Protección Social - MSPS.
- Programa Ampliado de Inmunización Nacional – PAI del Ministerio de Salud y Protección Social - MSPS.
- Programa Ampliado de Inmunización Distrital – PAI de la Secretaría Distrital de Salud de Bogotá - SDS.
- Registro Único Tributario – RUT de la Dirección de Impuestos y Aduanas Nacionales - DIAN.
- Registros Individuales de Prestación de Servicios de Salud – RIPS del Ministerio de Salud y Protección Social - MSPS.
- Registro para la Localización y Caracterización de Personas con Discapacidad – RLCPD del Ministerio de Salud y Protección Social - MSPS.
- Registro Único de Víctimas de la Unidad de Atención y Reparación Integral – RUV.
- Registro Estadístico de Migración Internacional – REMI del DANE y Base de Cédulas de Extranjería de Migración Colombia.

- Base de Datos Única de Afiliados a la Seguridad Social del Régimen de Excepción - BDEX del Ministerio de Salud y Protección Social - MSPS.
- Sistema de Información y Gestión del Empleo Público SIGEP del Departamento Administrativo de la Función Pública - DAFP.
- Caja de Retiro Militar de las Fuerzas Militares - CREMIL.
- Fuerzas Militares y sus familiares - FFMM.

Estadísticas vitales

- Nacidos vivos
- Defunciones No Fetales

Operaciones estadísticas por muestreo

- Gran Encuesta Integrada de Hogares - GEIH
- Encuesta de Calidad de Vida - ENCV

Definición de variables demográficas básicas

Con la finalidad de conocer la dinámica de la población residente que anualmente demanda servicios en Colombia a través del REBP, es necesario partir del análisis de las estructuras por edad y sexo según lugar de residencia. A continuación se describen las variables demográficas básicas incluidas en el REBP:

- **Sexo:** desde la perspectiva biológica, el sexo se refiere a las características genéticas, endocrinas y morfológicas del cuerpo al momento del nacimiento, las cuales identifican la capacidad reproductiva de las poblaciones humanas; es decir, desde un enfoque demográfico se busca medir los componentes del cambio poblacional. Las categorías utilizadas para clasificar estas características en los seres humanos son hombre, mujer e intersexual⁴. En general, a pesar que en la conformación del REBP se encontraron 1.041 personas sin información de sexo, esta variable tiene buena calidad en los registros administrativos.

- **Edad:** dentro de los límites biológicos, se acepta la sobrevivencia humana de los 0 años hasta los 115 años (Dong, 2016). El REPB contiene las edades de la población residente en Colombia con corte a 31 de diciembre de 2018; esta variable es calculada con la información de la fecha de nacimiento presente en los registros administrativos, las encuestas probabilísticas y el CNPV 2018⁵. Los valores fuera de rango pueden ser depurados a partir de criterios establecidos.

Los rangos de edades con mayor proporción de valores ausentes en la información de municipio de residencia se encuentran en las edades mayores a 85 años, aportando más del 25% de la inconsistencia. Las edades entre los 40 y 85 años aportan entre el 5% y 25% de las personas sin información de residencia. Las edades menores a 40 años presentan la mejor ubicación con porcentajes de faltantes inferiores al 5%, a excepción de la

⁴ En el CNPV 2018 y los RRAA solo se encuentran valores para las categorías hombre y mujer.

⁵ Si bien, el CNPV 2018, tiene un rango hasta los 125 años, se evidenció en los ejercicios de retroproyección que estos límites eran menores y que la sobrevivencia de esta población era poco probable y genera un sobre envejecimiento no real de la población.

población en edad cero, con un reporte del 9,2% de niños y 9,5% de niñas que no son localizados en el REBP.

- **Departamento y municipio:** teniendo en cuenta que en el CNPV 2018 estas variables, se relacionaron con el lugar de residencia habitual de las personas y en los registros administrativos captaron el lugar de afiliación o donde se reciben servicios; se incorporó al REBP la dirección reportada a las EPS con mayor número de usuarios, no obstante, es posible que en algunos casos no corresponda al lugar de residencia de todas las personas del

hogar como estándar idóneo, tal como se recoge en las operaciones estadísticas a hogares. En los registros administrativos estas variables tienen un alto porcentaje de información faltante que no permiten identificar la ubicación de las personas; así mismo, se evidencia la ausencia de nomenclaturas estandarizadas en varios registros administrativos frente a la identificación de las entidades territoriales demarcadas en la División Político Administrativa - DIVIPOLA del país. El 6,7% de registros no tiene ubicación de departamento, de los cuales el 3,6% son registros de hombres y el 3,1% de mujeres, lo que representa 3.333.324 individuos sin ubicación.

Definición de la variable estado

El REBP está compuesto por las variables que se determinan como características principales de la población, consideradas básicas para la elaboración de estadísticas demográficas y esenciales para la articulación con otros registros estadísticos. Adicionalmente, producto del proceso de integración, son generadas nuevas variables que caracterizan a los individuos presentes en el registro, según las características de la fuente que aporta a cada individuo. Esta nueva variable es denominada estado de la persona, la cual es descrita como:

- *Activo:* se asigna a las personas que son residentes vivos que se encuentran registradas en alguno de los registros administrativos.
- *Por emigración:* emigrante hacia el exterior basado en información de Migración Colombia.
- *Inactivo por proceso administrativo:* registros que se han detectados como no válidos, pero no se deben eliminar.

- *Inactivo por muerte:* personas registradas en la base de defunciones.
- *Inactivo por duplicado determinístico:* personas registrados dos o más veces.
- *Inactivo por cancelación en RNEC:* registros que tienen un documento en la base de la Registraduría en estado Cancelado.
- *Inactivo por inactividad en BDUA:* esta marca se realiza a los registros que en la BDUA tienen un estado de afiliación diferente a Activo o Suspendido.

La tabla 1, denominada conteo de personas por estado dentro del REBP, evidencia que el 78,3% de las personas se encuentran en estado activo y el 21,7% se encuentra inactivo en el restante de estados.

Tabla 1. Conteo de personas por estado REBP

Estado	Número de personas	%
Activo	49.291.079	78,3%
Inactivo por emigración	821.762	1,3%
Inactivo por proceso administrativo	0	0,0%
Inactivo por muerte	3.771.504	6,0%
Inactivo por duplicado determinístico	19.836	0,0%
Inactivo por cancelación en RNEC	3.574.574	5,7%
Inactivo por inactividad en BDUA	5.460.439	8,7%
Total	62.939.194	100%

Fuente: DANE - Dirección de Censos y Demografía.

2 **Análisis** de resultados

Emparejamiento del REBP 2018 con el CNPV 2018

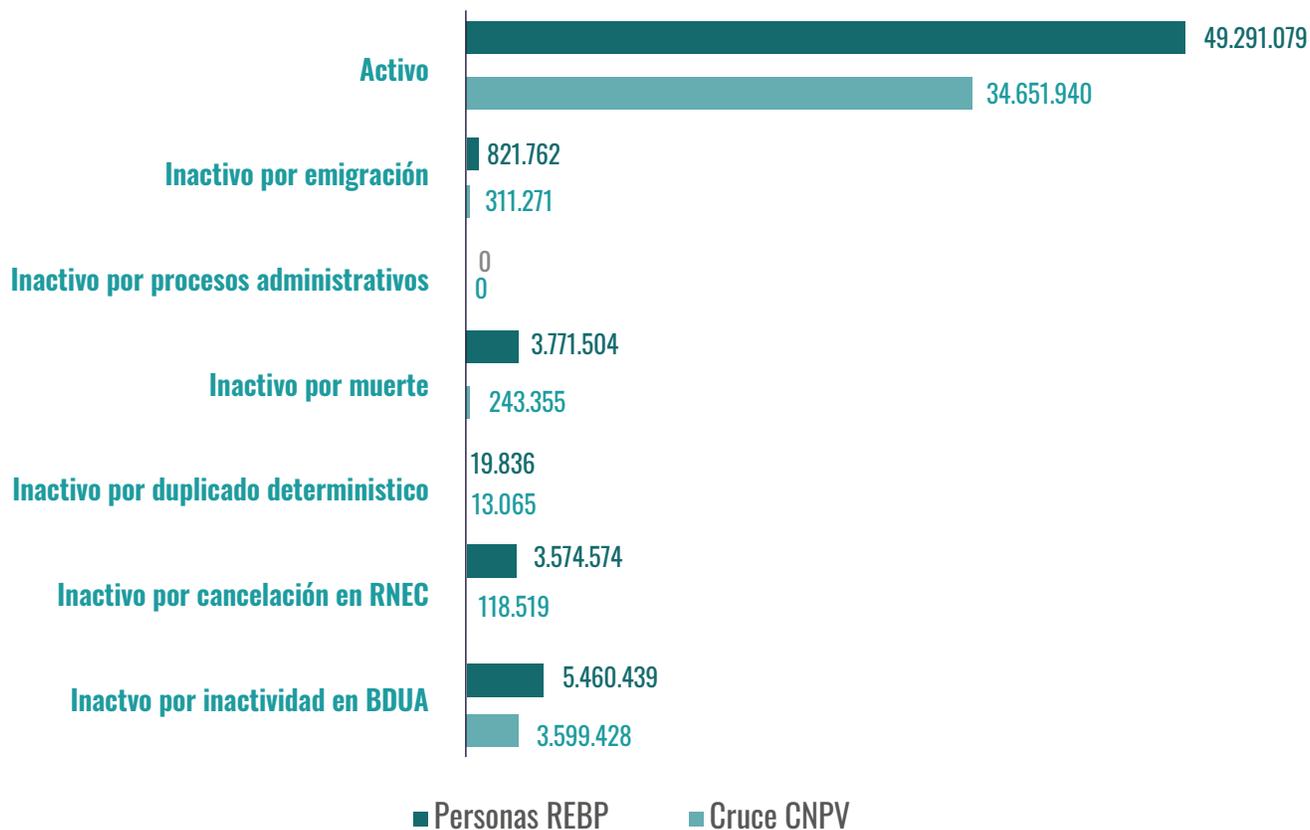
Con el fin de determinar el grado de integración que presenta el REBP con el CNPV 2018, e identificar discrepancias relevantes entre los resultados provenientes de las dos fuentes de información, se realizó el emparejamiento determinístico por tipo y número de documento de identificación entre los conjuntos de datos, a partir de la evaluación de diferencias en cobertura e indicadores demográficos.

Se consideró la no respuesta de las variables identificación, que en el CNPV del 2018 alcanzó el 7,1%; es decir, de las

44.164.417 personas censadas, 41.045.645 cuentan con esta información en la base de datos. Como resultado del emparejamiento se identifica que 34.651.490 de personas coincidieron, lo que indica un emparejamiento del 84,4% entre el REBP y el CNPV 2018.

El gráfico 2 muestra el total de personas dentro REBP por estado, así como el conteo de los individuos que cruzaron con el CNPV 2018. Los resultados muestran que, la mayor proporción de emparejamientos se encuentran activos dentro del registro.

Gráfico 2. Emparejamiento por estado REBP-CNPV



Análisis nacional, comparación estructuras poblacionales CNPV y REBP

El volumen de población total del registro, incluyendo aquellos sin municipio, corresponde a 49.291.079 personas activas en 2018 sin incluir aquellas que no cuentan con información de municipio de residencia, esta cifra desciende a 45.957.755 personas activas. Teniendo en cuenta que, en el CNPV 2018, el total de personas efectivamente censadas es de 44.164.417, evidenciando una diferencia de 4,1% entre los volúmenes poblacionales de las dos fuentes. Estos resultados son un avance importante en la difusión de estadísticas poblacionales basadas en los registros administrativos como fuentes alternativas. No obstante, es preciso señalar que los agregados nacionales reflejan una mirada parcial de los conteos que describen el tamaño poblacional, por lo cual es importante analizar la consistencia de estos resultados desagregados temática y geográficamente, teniendo en cuenta variables demográficas como la edad, el sexo y la distribución en el territorio.

De acuerdo con el REBP, la distribución poblacional según edad y sexo da cuenta de características asociadas a la evolución de los procesos de transición demográfica; dadas las características propias de la recolección de datos administrativos, la pirámide poblacional refleja una participación mayor de población en edades potencialmente activas (productivas). Por su parte, la distribución etaria de la declaración censal evidencia un comportamiento más simétrico y consistente en el seguimiento a las participaciones por edad, en especial en la población de jóvenes y en general de adultos menores de 40 años. Así mismo, se identifican diferencias importantes en la captación de la población de menores de un (1) año de edad, quienes presentan una subestimación significativa en los registros administrativos asociados a los rezagos en los procesos de captación de su identificación.

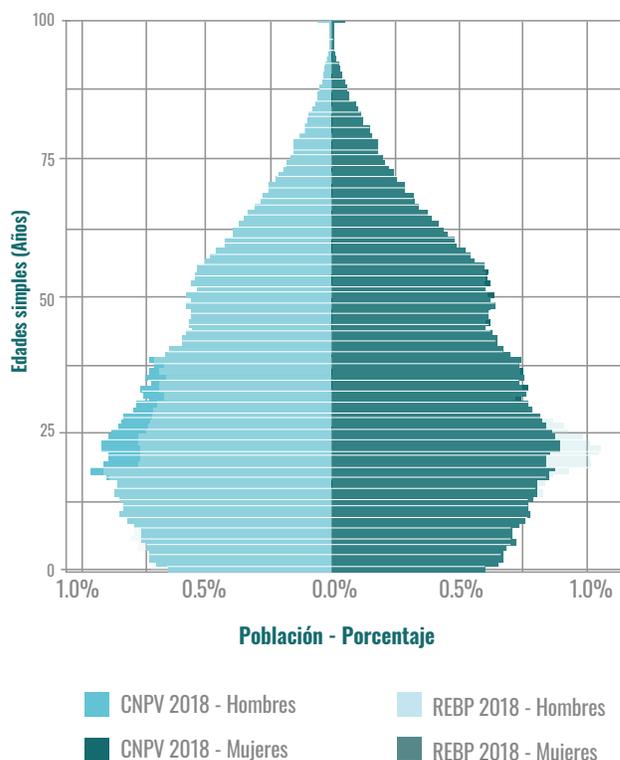
La comparación de las pirámides de población con información del CNPV 2018 y del REBP, muestra diferencias en algunos segmentos de edad para ambos sexos, principalmente en las

mujeres en edades entre 20 y 34 años. Lo anterior, requirió un análisis de consistencia para identificar las hipótesis asociadas a la declaración de edad en el REBP, con el fin de depurar los posibles factores que duplican los registros individuales de poblaciones de mujeres en edades jóvenes adultas.

Por lo demás, se observa una composición por edad en ambas fuentes, asociada a procesos de transición moderada con predominancia femenina, característica de poblaciones urbanizadas.

Gráfico 3. Pirámide de población, comparación CNPV - REBP 2018

Pirámide población en valores relativos - Colombia
No incluye registros sin departamento



Fuente: DANE - Dirección de Censos y Demografía.

Para verificar el predominio de mujeres entre 20 y 34 años en la estructura por edad y sexo, se revisó la información de todos los registros administrativos que aportan información al REBP con lo cual se pudo verificar que correspondían a registros duplicados, pertenecientes a niños recién nacidos que quedaban registrados con la identificación de la madre, mientras se generaba el correspondiente registro civil de nacimiento. A partir de esta información, se desactivan un total

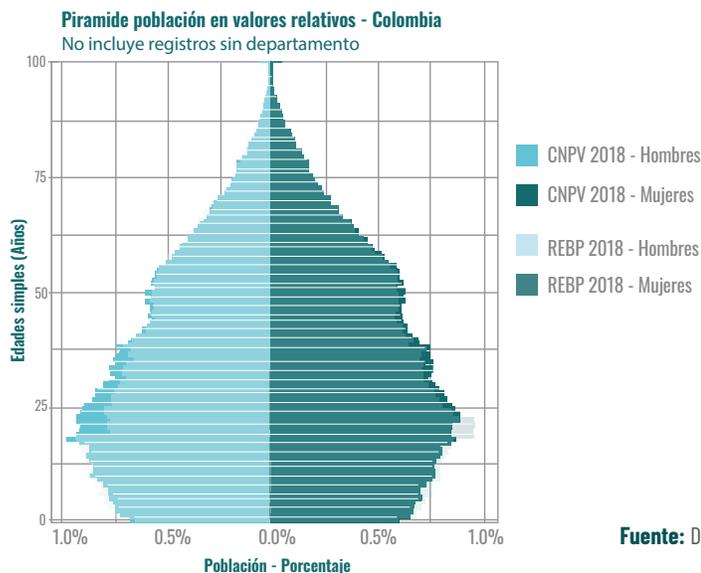
de 571.960 registros, de los cuales 415.598 son mujeres de 20 a 34 años, en estado activo con un municipio asignado. El impacto de esta decisión en el volumen poblacional del REBP por estado se observa en la tabla 2; las mujeres que cambiaron de estado activo a inactivo, por proceso administrativo, representan el 0,9% de las personas en el REBP. Adicionalmente, se observa un cambio en la estructura poblacional, que se muestra en el gráfico 4: Conteo de población ajustado por estado en el REBP.

Tabla 2. Conteo de población ajustado por estado en el REBP

Estado	Número de personas	%
Activo	48.719.119	77,4%
Inactivo por emigración	821.762	1,3%
Inactivo por proceso administrativo	571.960	0,9%
Inactivo por muerte	3.771.504	6,0%
Inactivo por duplicado determinístico	19.836	0,0%
Inactivo por cancelación ENMEC	3.574.574	5,7%
Inactivo por inactividad en BDUA	5.460.439	8,7%
Total	62.939.194	100%

Fuente: DANE - Dirección de Censos y Demografía.

Gráfico 4. Pirámide de población del CNPV 2018 contra el REBP 2018 después del ajuste del estado para las mujeres de 20 a 34 años



Fuente: DANE - Dirección de Censos y Demografía.

Indicadores nacionales REBP frente a CNPV

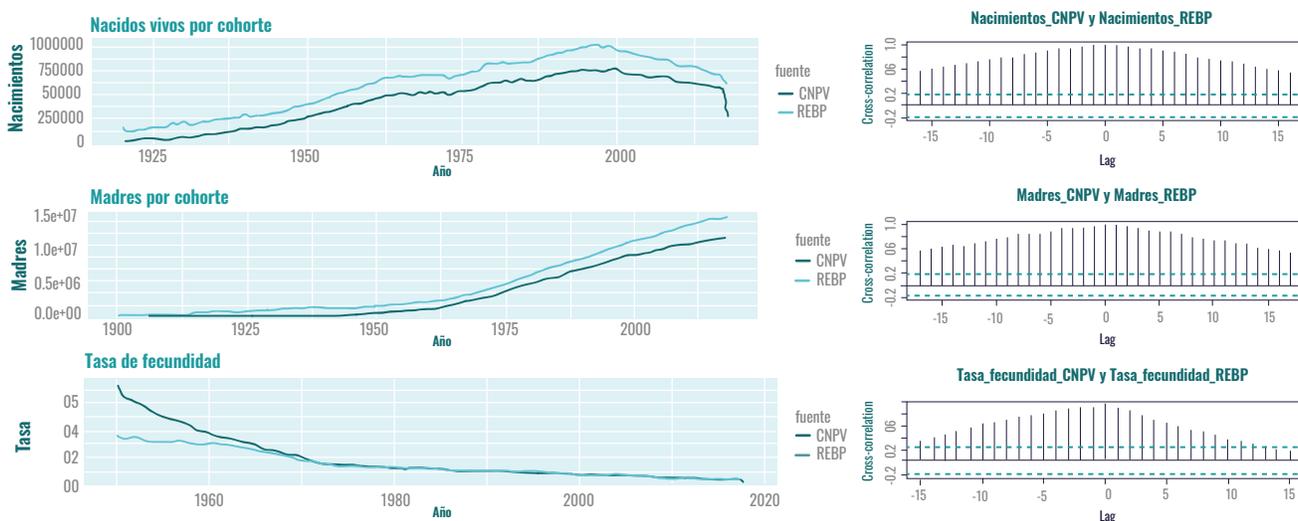
Para realizar la comparación de las dos fuentes de información, se obtienen los siguientes indicadores desagregados por cohorte:

- Número de nacidos vivos reportados.
- Número de mujeres en edad fértil (15 a 49 años).
- Tasa General de Fecundidad - TGF (la razón entre las dos anteriores y que representa el número de hijos por mujer).

Los indicadores demográficos obtenidos desde el REBP y su comparación con el CNPV muestra la calidad del REBP

para registrar la estructura en el tiempo, corroborada después con la prueba gráfica de correlación cruzada⁶ en la que se muestra la alta simetría de estas tres series de datos en el tiempo. Además, se ha usado el Error Medio Porcentual Absoluto (MAPE en inglés)⁷ para mostrar qué tanto error presenta la curva del REBP en comparación con el CNPV 2018. El MAPE obtenido para la serie de nacidos vivos es del 37,3%, para las madres por cohorte del 57,0% y para la TGF del 14,5%. Algo importante de observar es que la discrepancia generada en la TGF, sobre todo, en cohortes más viejas es dada por el hecho de que el censo no captura bien las mujeres en edad fértil de dichas cohortes, con lo que la información mejora a partir de 1970.

Gráfico 5. Indicadores demográficos obtenidos desde el REBP y su comparación con el CNPV



Fuente: DANE - Dirección de Censos y Demografía.

⁶ Ver Correlación Cruzada.

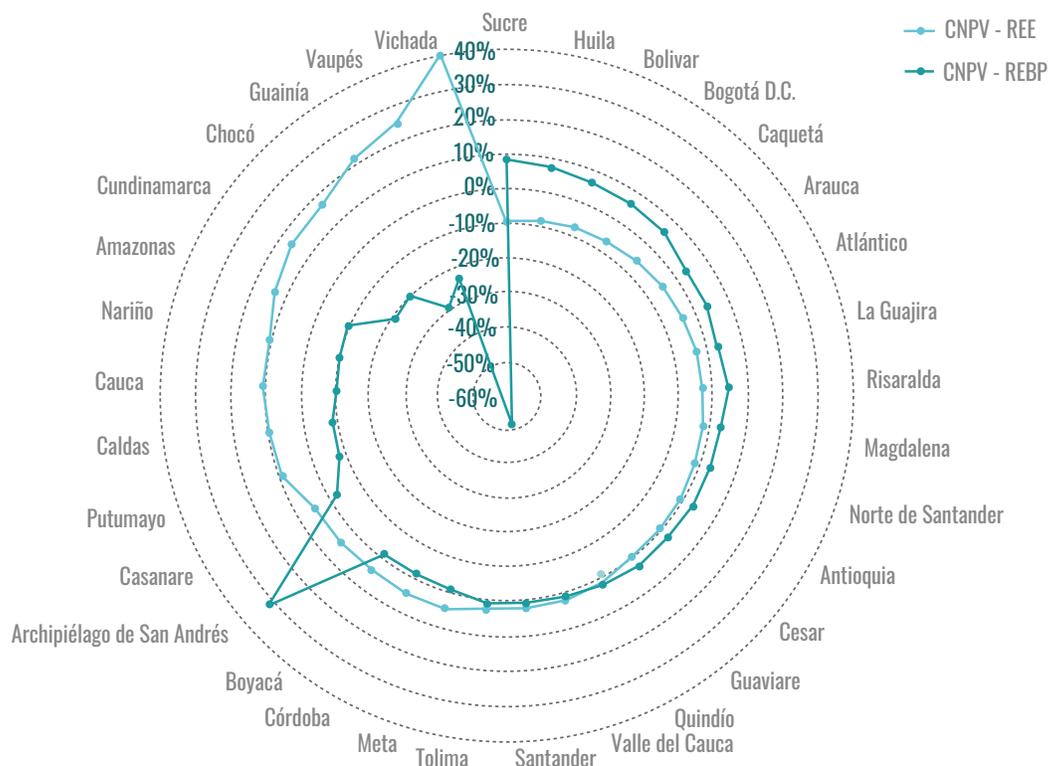
⁷ $MAPE = 100 * \frac{1}{n} \sum_{i=1}^n \frac{|Adjusted_i - Observed_i|}{Observed_i}$ Ver MAPE-Error-medio-de-porcentaje-absoluto

Análisis cobertura departamental

Para este análisis, se construyó la estructura de errores por departamento de los volúmenes poblacionales, realizando la comparación mediante el porcentaje de variación entre cada conjunto de datos como indicador principal⁸. Este indicador es importante, debido a que con él se tiene una primera aproximación a los departamentos con mayor desviación en caso de realizar una enumeración de población.

Con los volúmenes poblacionales del CNPV 2018, la Encuesta de Calidad de Vida 2018-ECV⁹ y el REBP, se aplicó la prueba K-S¹⁰ mostrando como resultado que la distribución de las diferencias relativas entre el REBP y el CNPV 2018 se encuentra dentro de un rango de -4,1% y 43,2%, destacando a los departamentos con predominancia de población en el área rural y con sub-registro de personas respecto al CNPV y los departamentos de la costa caribe y región andina como los de mayor sobre-registro. Las pruebas fueron realizadas en el software R, obteniendo un

Gráfico 6. Diferencias relativas departamentales CNPV-REBP y ECV-REBP



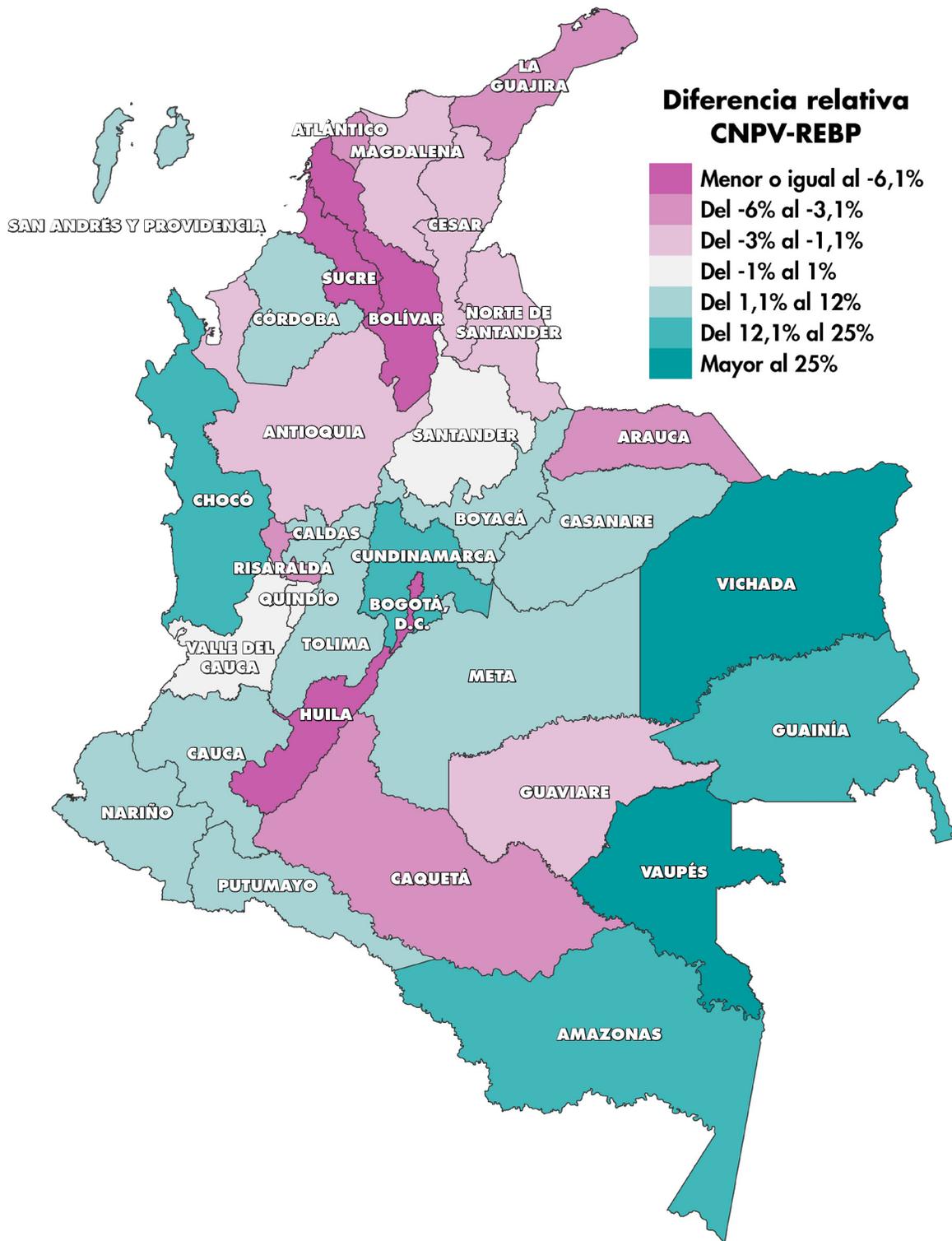
Fuente: DANE - Dirección de Censos y Demografía.

⁸ Diferencia relativa=1-Población REBP/Población CNPV

⁹ Encuesta de Calidad de Vida 2018 ECV 2018

¹⁰ Ver Kolmogorov-Smirnov Tests

Ilustración 1. Diferencias relativas departamentales CNPV-REBP y ECV-REBP



Fuente: DANE - Dirección de Censos y Demografía.

p-valor de 0,9995 para las dos pruebas; por lo tanto, con una significancia del 5% no existe suficiente evidencia estadística para rechazar que las distribuciones departamentales CNPV-REBP y ECV-REBP son diferentes. Las Diferencias relativas departamentales CNPV-REBP y ECV-REBP muestran los porcentajes de discrepancia

(usando las diferencias relativas) entre el REBP y el CNPV y entre el REBP y la ECV, en la cual se observa que 21 de los 33 departamentos presentan diferencias relativas menores al 10%. Adicionalmente se muestra la distribución geográfica de las diferencias relativas entre el CNPV y el REBP.

Análisis cobertura municipal

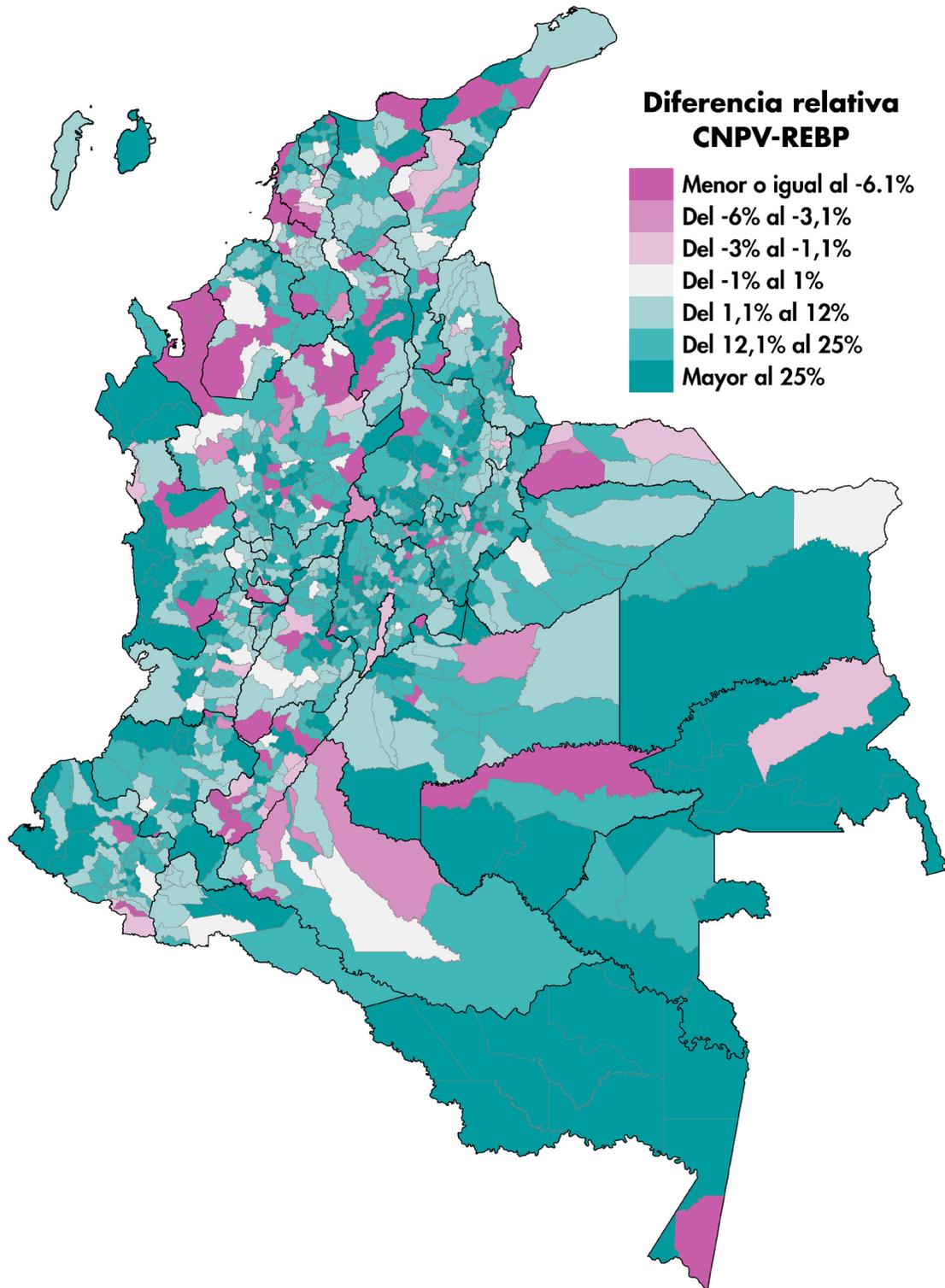
En contraste con la distribución departamental, el análisis de las diferencias relativas a nivel municipal muestra, que dichos valores se encuentran dentro de un rango de -58% y 91%, que intuitivamente son evidencia de la disparidad de las distribuciones poblaciones del REBP y el CNPV. Dado que la ECV 2018 no tiene representatividad municipal, se realiza únicamente la comparación de las distribuciones entre CNPV 2018 y REBP. Al aplicar el mismo procedimiento que a los datos departamentales, se obtiene un p-valor de 0,0001. Por lo

tanto, con una significancia del 5%, se concluye que existe suficientes evidencia estadística para afirmar que, las distribuciones municipales del REBP y CNPV son iguales, por lo que hasta el momento no es recomendable obtener indicadores demográficos a este nivel.

En el gráfico 7 se muestra dicha distribución, en la que destaca el subregistro de los municipios pertenecientes a Guaviare y Amazonas, en contraste con el sobre registro en los municipios de la región caribe.



Ilustración 2. Diferencias relativas municipales CNPV-REBP



Fuente: DANE - Dirección de Censos y Demografía.

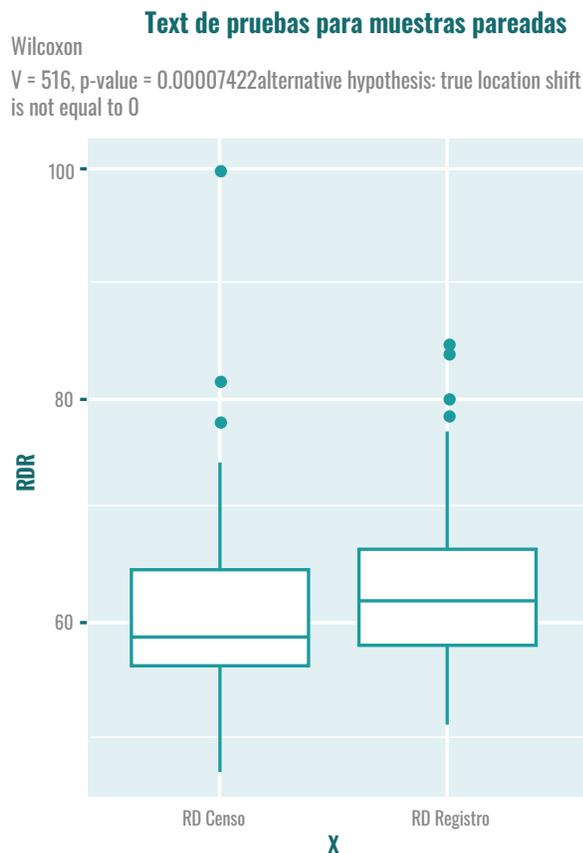
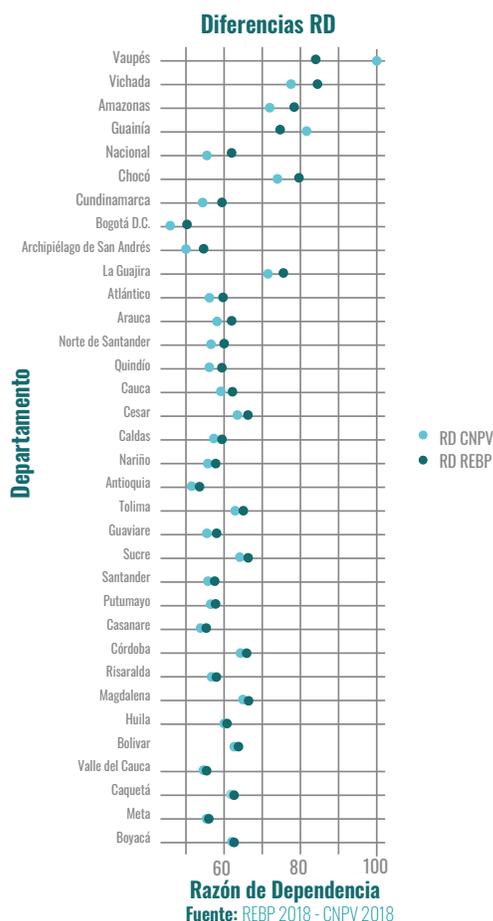
Indicadores departamentales calculados con REBP 2018 en comparación con CNPV 2018

Razón de Dependencia - RD

A partir de este análisis se busca establecer la relación entre toda la estructura poblacional al integrar los tres grandes grupos de edad (0-14, 15-59, 60 y más) en el indicador. Los resultados presentados en el gráfico 8 muestran que Boyacá, Meta, Caquetá, Bolívar y Huila presentan altas coincidencias en la RD, en donde posiblemente las estructuras no evidencian contrastes en los grupos de edad comparados entre las dos fuentes.

Por otra parte en Vaupés, Vichada, Guainía y Amazonas las brechas son más amplias, corroborando la hipótesis de la baja calidad de los registros fuente, en los territorios llamados antiguos territorios nacionales. Estas brechas generan que las diferencias sean significativas a la prueba de muestras pareadas de Wilcoxon¹¹ con un p-valor < 0,05.

Gráfico 7. Análisis de comparabilidad razón de dependencia REBP-CNPV 2018



¹¹ Diferencia relativa = 1 - Población REBP / Población CNPV

Fuente: DANE - Dirección de Censos y Demografía.

Índice de Masculinidad, razón por sexo (número de hombres por cada cien mujeres)

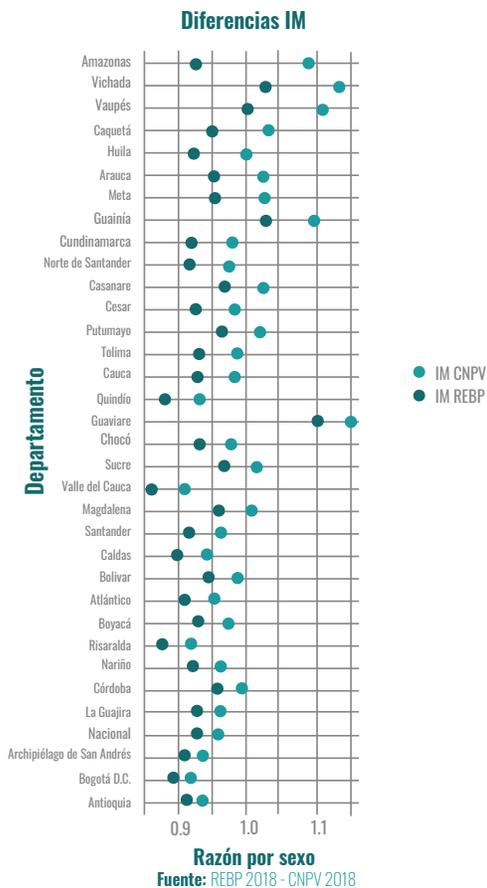
Este indicador es apropiado para identificar sesgos de sexo en las estructuras poblacionales, determinando si encuentran masculinizadas o feminizadas de acuerdo con el resultado del indicador. Para 2018 el Índice de Masculinidad - IM a nivel nacional obtenido por el CNPV fue de 95 hombres por cada 100 mujeres. Por otra parte, el REBP presenta un resultado de 92 hombres por cada 100 mujeres.

El gráfico 9 muestra que estas diferencias son mucho más amplias en departamentos como Amazonas, Vichada y Vaupés, con una brecha de más de 10 hombres. Por ejemplo, para Amazonas el REBP reporta 92 hombres por cada 100 mujeres y el CNPV 2018 reporta 108 hombres por cada 100 mujeres, con una discrepancia de 16 hombres más en la

información censal. En Antioquia, Bogotá y el Archipiélago de San Andrés, Providencia y Santa Catalina las divergencias son menores. Lo anterior se constata al aplicar la prueba de Wilcoxon, en este caso, un p-valor menor de 0,05 indica que las diferencias son significativas con una mediana mucho más alta para el CNPV 2018, como se observa en el diagrama de caja, que compara las dos fuentes.

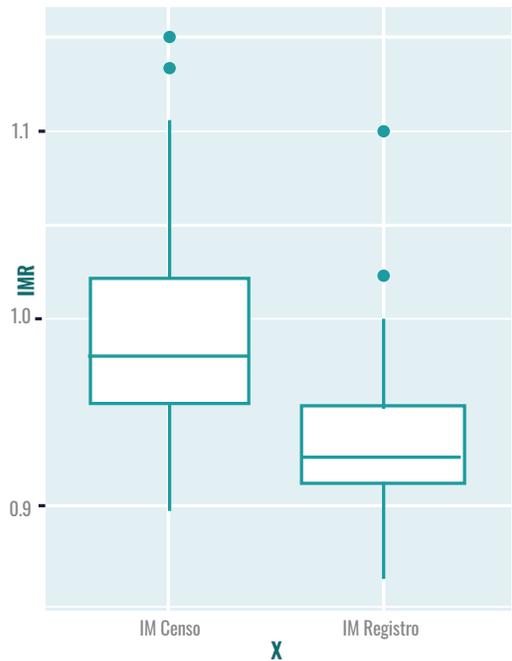
Esto indica que las estructuras del REBP son mucho más feminizadas que las del CNPV 2018, como se observa en el gráfico 9, en donde los resultados del CNPV siempre se encuentran a la derecha y los del REBP a la izquierda, manteniendo los mismos patrones, pero con una tendencia a la feminización de la estructura.

Gráfico 8. Análisis de comparabilidad, Índice de Masculinidad CNPV 2018 - REBP



Text de pruebas para muestras pareadas

Wilcoxon¹
V = 0, p-value = 0.000000001164
alternative hypothesis: true location shift is not equal to 0



Fuente: DANE - Dirección de Censos y Demografía.

Índice de Envejecimiento - IE

Este indicador mide el ensanchamiento de las cúspides de las estructuras poblacionales. El gráfico 10, muestra coincidencias muy altas en el IE de Caquetá, Santander, Meta, Putumayo, Antioquia, Casanare, Cauca y Chocó, estimados por las dos fuentes comparadas. Vaupés, Norte de Santander, Sucre, Guaviare y La Guajira presentan las mayores divergencias, con diferencias entre 7 y 10 personas adultas adicionales en el indicador, esto se traduce en un mayor envejecimiento de estos departamentos, con respecto a lo estimado en el CNPV 2018.

En cuanto a las menores discrepancias, en comparación con lo evaluado en la razón por sexo, la prueba de Wilcoxon arroja un p-valor menor a 0,05, indicando que las diferencias entre el IE calculado con las dos fuentes si es significativa. El diagrama de caja muestra también una mayor dispersión del indicador a nivel departamental, estimada a partir del registro, con una mediana mayor que la de CNPV 2018, como se observa en el gráfico 10.

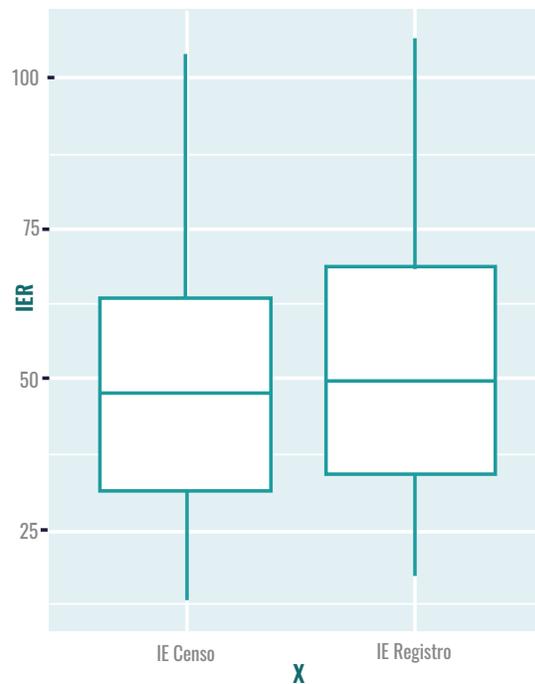
Gráfico 9. Análisis de comparabilidad, Índice de Envejecimiento CNPV 2018 - REBP



Text de pruebas para muestras pareadas Wilcoxon

V = 0, p-value = 0.00005627

alternative hypothesis: true location shift is not equal to 0

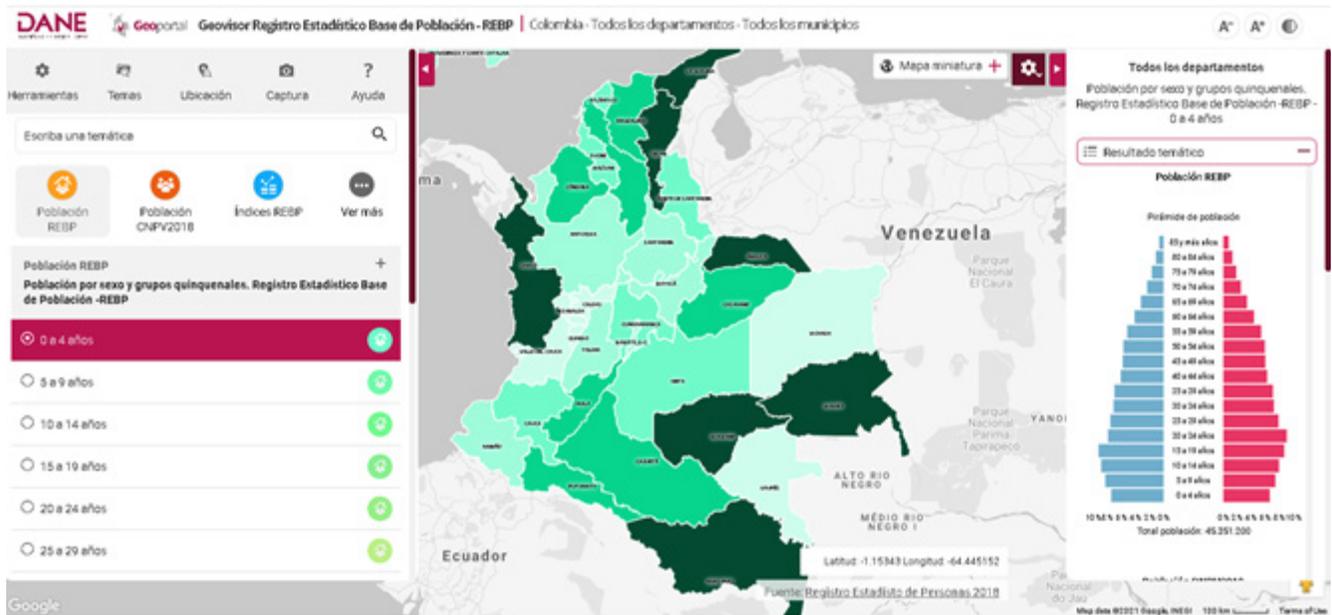


Fuente: DANE - Dirección de Censos y Demografía.

3 Visualizaciones

Para la visualización de los indicadores se cuenta con un geovisor, donde se observan las distribuciones espaciales de los indicadores. El geoportal que lo alberga se puede revisar en la siguiente URL: <https://geoportal.dane.gov.co/usuarios/registro-estadistico-personas/>

Ilustración 3. Captura de pantalla con la información del geovisor REBP y CNPV



Fuente: DANE - Geoportal.

4 Conclusiones

1. El proceso de integración se realizó de forma determinística, debido a la capacidad computacional disponible hasta el momento. Esto implica que, a medida que se disponga de mayores recursos y capacidad de procesamiento, será necesario actualizar el proceso de integración con la aplicación de métodos probabilísticos; aumentando la calidad del cruce en cuanto a duplicados, inclusión de nuevos registros y la posibilidad de integrar otras fuentes que no cuenten con la variable documento de identificación.
2. El acceso a las bases de registros administrativos se realizó por medio de solicitudes directas a los custodios de la información, lo que implica un trabajo adicional para estandarizar y homologar los conjuntos de datos para su posterior integración en el REBP. Estos procesos pueden ser más eficientes a través de la interoperabilidad gestionada en el marco del Sistema Estadístico Nacional - SEN.
3. Las características del REBP permiten proponer análisis longitudinales y la implementación de metodologías probabilísticas para la estimación de fenómenos demográficos. No

obstante, es importante implementar acciones de fortalecimiento de los RRAA en cuanto a cobertura y calidad; así como, mejoras continuas a los procesos de construcción del REBP.

Los resultados que se obtienen de la comparación REBP 2018 y CNPV 2018, permiten conocer y formular acciones concretas para el mejoramiento del registro, mostrando el potencial de la información con la que se cuenta, configurándose en una fuente pertinente, oportuna y de gran utilidad para conocer la evolución de la población del país.

4. El DANE pone a disposición herramientas gráficas que permiten observar la distribución poblacional, estructuras y algunos de los principales indicadores demográficos provenientes del REBP comparados con el CNPV 2018 a niveles nacional, departamental y municipal. Se espera que este ejercicio tenga un exhaustivo aprovechamiento por las entidades encargadas del diseño y seguimiento de la política pública, la comunidad académica y el público en general.

BIBLIOGRAFÍA

Christen, P. (2012). *Data Matching Concepts and Techniques for Record Linkage Entity Resolution and Duplicate Detection*. New York: Springer.

Dong, X., Milholland, B., & Vijg, J. (2016). Evidence for a limit to human lifespan. *Nature*, 538(7624), 257-259.

UNECE (2007). *Register-based statistics in the Nordic countries. Review of best practices with focus on population and social statistics*. Naciones Unidas. Nueva York y Ginebra 2007.

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics* 1, 80-83.

ANEXOS

Anexo 1: Comparación de población por departamentos según sexo, CNPV2018 – REBP 2018

DPTO	Departamento	CNPV 2018		REBP 2018	
		Hombre	Mujer	Hombre	Mujer
5	Antioquia	2.885.383	3.089.405	2.967.147	3.191.267
8	Atlántico	1.140.357	1.201.908	1.196.080	1.296.447
11	Bogotá, D.C.	3.433.586	3.747.883	3.540.725	3.892.605
13	Bolívar	947.598	961.862	1.024.811	1.064.968
15	Boyacá	559.084	576.614	539.776	567.242
17	Caldas	447.499	475.973	408.446	446.082
18	Caquetá	182.378	177.224	196.931	200.055
19	Cauca	615.833	627.670	602.204	632.151
20	Cesar	543.838	554.739	558.491	591.214
23	Córdoba	774.448	781.148	803.409	826.168
25	Cundinamarca	1.379.858	1.413.019	1.114.954	1.182.995
27	Chocó	225.982	231.430	203.804	216.070
41	Huila	503.766	505.782	541.343	565.093
44	La Guajira	404.215	421.149	416.716	444.826
47	Magdalena	632.502	631.286	641.522	656.132
50	Meta	464.906	454.223	468.119	476.425
52	Nariño	653.754	681.767	667.291	710.469
54	Norte de Santander	664.057	682.749	696.233	744.145
63	Quindío	245.663	263.977	240.413	266.550
66	Risaralda	401.748	437.849	434.191	484.907

DPTO	Departamento	CNPV 2018		REBP 2018	
		Hombre	Mujer	Hombre	Mujer
68	Santander	983.439	1.025.402	986.306	1.049.189
70	Sucre	434.459	429.577	462.441	469.318
73	Tolima	609.017	619.746	602.268	632.630
76	Valle del Cauca	1.800.614	1.989.260	1.956.513	2.220.969
81	Arauca	121.060	118.443	127.146	128.368
85	Casanare	192.020	187.872	189.306	190.050
86	Putumayo	142.900	140.297	148.757	149.134
88	Archipiélago de San Andrés Providencia y Santa Catalina	23.326	24.973	26.773	28.929
91	Amazonas	34.422	31.634	30.289	31.286
94	Guainía	23.214	21.217	17.975	17.177
95	Guaviare	39.080	34.001	41.954	36.824
97	Vaupés	19.793	17.897	14.552	14.097
99	Vichada	40.694	35.948	30.931	29.601

Anexo 2: Comparación de población por departamentos y capitales según sexo, CNPV2018 - REBP2018

DPTO	Departamento	MPIO	Municipio	CNPV 2018		REBP 2018	
				Hombre	Mujer	Hombre	Mujer
5	Antioquia	05001	Medellín	1.115.273	1.257.057	1.223.590	1.373.450
8	Atlántico	08001	Barranquilla	536.790	583.313	676.222	739.651
11	Bogotá, D.C.	11001	Bogotá, D.C.	3.433.586	3.747.883	3.540.725	3.892.605
13	Bolívar	13001	Cartagena	426.981	460.965	523.626	569.023
15	Boyacá	15001	Tunja	79.244	88.747	85.244	96.199
17	Caldas	17001	Manizales	189.095	211.341	180.702	207.065
18	Caquetá	18001	Florencia	77.620	79.169	82.194	88.940
19	Cauca	19001	Popayán	133.282	143.988	135.438	150.351
20	Cesar	20001	Valledupar	223.735	235.614	316.488	343.924
23	Córdoba	23001	Montería	210.379	223.344	234.894	252.835
27	Chocó	27001	Quibdó	55.909	64.770	5.431	6.104
41	Huila	41001	Neiva	150.525	164.001	66.294	76.150
44	La Guajira	44001	Riohacha	86.766	90.807	184.319	203.725
47	Magdalena	47001	Santa Marta	232.340	247.513	119.753	127.706
50	Meta	50001	Villavicencio	222.585	228.627	254.959	272.889
52	Nariño	52001	Pasto	166.804	185.522	246.826	262.028
54	Norte de Santander	54001	Cúcuta	304.357	325.057	183.365	205.859
63	Quindío	63001	Armenia	130.090	145.551	365.695	398.038
66	Risaralda	66001	Pereira	193.060	216.610	142.205	160.428
68	Santander	68001	Bucaramanga	250.975	277.880	228.720	262.784
70	Sucre	70001	Sincelejo	133.008	141.614	306.623	344.432
73	Tolima	73001	Ibagué	239.289	261.397	163.622	173.864
76	Valle del Cauca	76001	Cali	852.852	970.017	249.942	280.890
81	Arauca	81001	Arauca	37.805	37.930	1.010.774	1.177.099
85	Casanare	85001	Yopal	77.972	78.970	41.161	43.343

DPTO	Departamento	MPIO	Municipio	CNPV 2018		REBP 2018	
				Hombre	Mujer	Hombre	Mujer
86	Putumayo	86001	Mocoa	23.870	24.552	82.048	84.073
88	Archipiélago de San Andrés, Providencia y Santa Catalina	88001	San Andrés	21.114	22.640	25.057	27.008
91	Amazonas	91001	Leticia	22.047	20.797	25.057	27.189
94	Guainía	94001	Inírida	15.612	14.492	24.987	25.850
95	Guaviare	5001	San José del Guaviare	24.865	23.221	16.144	15.513
97	Vaupés	7001	Mitú	14.494	13.215	28.711	27.269
99	Vichada	99001	Puerto Carreño	10.422	9.366	11.446	11.146

Anexo 3. Orden de integración de los registros administrativos

Orden de integración	RRAA	Observaciones	Registros Integrados	Acumulado
1	RNEC	Corte Diciembre 2018	60.062.094	60.062.094
2	BDOA	Corte Diciembre 2018	1.413.669	61.475.763
3	BDEX	Corte Octubre 2018	36.911	61.512.674
4	COMPENSAR EPS	Corte Diciembre 2018	7.692	61.520.366
5	SANITAS EPS	Corte Diciembre 2018	1.763	61.522.129
6	EMSSANAR ESS	Corte Diciembre 2018	25	61.522.154
7	CAPITAL SALUD	Corte Diciembre 2018	-	61.522.154
8	COOMEVA EPS	Corte Diciembre 2018	539	61.522.693
9	ALIANSALUD EPS	Corte Diciembre 2018	16	61.522.709
10	SAVIA SALUD EPS	Corte Diciembre 2018	916	61.523.625
11	CRUZ BLANCA	Corte Diciembre 2018	181	61.523.806
12	SALUD VIDA	Corte Diciembre 2018	1	61.523.807
13	SISBEN	Corte Diciembre 2018	231.643	61.755.450
14	NACIMIENTOS	Corte 2008-2018	576.839	62.332.289
15	DEFUNCIONES	Corte 2008-2018	10.052	62.342.341
16	PILA COTIZANTES	Corte Diciembre 2018	41.570	62.383.911
17	PILA PENSIONADOS	Corte Diciembre 2018	5.340	62.389.251
18	GEIH	Corte Diciembre 2018	7	62.389.258
19	ECV	Corte Diciembre 2018	4	62.389.262
20	FFMM	Corte Diciembre 2018	133	62.389.395
21	FFMM_FAMILARES	Corte Diciembre 2018	623	62.390.018
22	CREMIL	Corte Diciembre 2018	8	62.390.026
23	REMI	Corte Diciembre 2017	467	62.390.493
24	DISCAPACIDAD	Corte Diciembre 2018	19.550	62.410.043
25	SIGEP	Corte Diciembre 2017	51	62.410.094
26	RUV	Corte Febrero 2019	48.982	62.459.076
27	RUT	Corte Diciembre 2015	3.083	62.462.159

Orden de integración	RRAA	Observaciones	Registros Integrados	Acumulado
28	PAI NACIONAL	Corte 2012-2017	407.074	62.869.233
29	PAI DISTRITAL	Corte Diciembre 2018	6.160	62.875.393
30	RIPS	Corte Diciembre 2018	-	62.875.393
31	SIMAT	Corte Diciembre 2017	56.151	62.931.544
32	SNIES	Corte Diciembre 2017	3.618	62.935.162
33	SIET	Corte Diciembre 2017	4.029	62.939.191
34	EPS COMPARTA	Corte Diciembre 2018	3	62.939.194

DANE

INFORMACIÓN PARA TODOS



@DANE_Colombia



/DANEColombia



/DANEColombia



@DANEColombia

www.dane.gov.co