

Censo Nacional de Población,  
Hogares y Viviendas

# Informe de evaluación de calidad cuantitativa de la segunda prueba piloto

Hacia la Ronda 2020

---

Documento de trabajo N° 33

Diciembre de 2020



Ministerio de Economía  
Argentina

Instituto Nacional de  
Estadística y Censos  
República Argentina



## Censo Nacional de Población, Hogares y Viviendas

### Informe de evaluación de calidad cuantitativa de la segunda prueba piloto. Hacia la ronda 2020

#### Documento de trabajo N° 33

Instituto Nacional de Estadística y Censos (INDEC)

Esta publicación ha sido realizada por:

**Dirección Nacional de Estadísticas Sociales y de Población:** Gladys Massé, Julio Rodríguez, Lía Pesaresi.

**Dirección de Estadísticas Poblacionales:** Cecilia Rodríguez Gauna, Alejandra Cuasnicu, Silvia Mario, Suyay Mura, Nicolás Zucco, Tatiana Santillán, Sofía Muñafra, Iván Williams, Facundo García Tarsia, Fernando Aparicio, Viviana Spinardi, Natalia Núñez, Matías Guzman, Flavia Oliari, Néstor Miño, Emilio Salgado, Carina David, Roxana Cuevas, Carlos Achigar, Miguel Reisenman, Nicolás Basso.

**Dirección de Capacitación y Carrera del Personal:** Javier Winokur, Mariana Goutman, María del Rosario Villamayor, Pilar Barrientos, Mariela Velázquez e Irene Valdivia.

**Coordinación de Clasificadores y Nomenclaturas:** Patricia Hoxter, Javier Domench, Marcelo Esses y Claudio La Rocca.

ISSN 2683-7919

ISBN 978-950-896-598-1

Instituto Nacional de Estadística y Censos - I.N.D.E.C.

Informe de evaluación de calidad cuantitativa de la segunda prueba piloto : hacia la ronda 2020. - 1a ed - Ciudad Autónoma de Buenos Aires : Instituto Nacional de Estadística y Censos - INDEC, 2020.

Libro digital, PDF - (Documentos de trabajo ; 33)

Archivo Digital: descarga y online

ISBN 978-950-896-598-1

1. Estadísticas. 2. Censos de Población. I. Título

CDD 310



Queda hecho el depósito que fija la ley nº 11.723

Libro de edición argentina

#### Responsabilidad editorial

Director del INDEC: Marco Lavagna

Director Técnico: Pedro Llunes

Directora Nacional de Difusión y Comunicación: Silvina Viazzi

Coordinador de producción gráfica y editorial: Marcelo Costanzo

Buenos Aires, diciembre de 2020

#### Publicaciones del INDEC

Las publicaciones editadas por el Instituto Nacional de Estadística y Censos están disponibles en [www.indec.gob.ar](http://www.indec.gob.ar) y en el Centro Estadístico de Servicios, ubicado en Av. Presidente Julio A. Roca 609 C1067ABB, Ciudad Autónoma de Buenos Aires, Argentina. También pueden solicitarse al teléfono +54 11 5031-4632 en el horario de atención al público de 9:30 a 16:00. Correo electrónico: [ces@indec.gob.ar](mailto:ces@indec.gob.ar)  
Calendario anual anticipado de informes: [www.indec.gob.ar/indec/web/Calendario-Fecha-0](http://www.indec.gob.ar/indec/web/Calendario-Fecha-0)



INDECArgentina

# Índice

<b>Introducción .....</b>	<b>7</b>
<b>1. Aspectos conceptuales .....</b>	<b>7</b>
<b>2. Organización de la base de datos .....</b>	<b>11</b>
2.1 Implementación del proceso de estructura jerárquica .....	16
2.2 Tipo de errores .....	16
2.3 Etapas del proceso de organización de la base de datos y ajustes realizados .....	16
<b>3. Evaluación de calidad de la lectura e interpretación de los datos .....</b>	<b>17</b>
3.1 Introducción .....	17
3.2 Metodología .....	17
3.3 Evaluación .....	18
3.4 Tipos de error de la interpretación y sus causas .....	21
3.5 Conclusiones .....	25
<b>4. Errores de calidad .....</b>	<b>26</b>
4.1 Tipo de vivienda, Condición de ocupación y Detección de hogares (V1 a V7) .....	27
4.1.1 Diseño de preguntas en el cuestionario .....	27
4.1.2 Objetivo y principales características de las preguntas .....	27
4.1.3 Resultados .....	28
4.2 Materiales de los pisos, las paredes y los techos (A: H9 a H13) (B: H9 a H11) .....	31
4.2.1 Diseño de preguntas en el cuestionario .....	31
4.2.2 Objetivo y principales características de las preguntas .....	31
4.2.3 Resultados .....	31
4.3 Condiciones sanitarias, Combustible y Agua caliente (A: H14 a H22) (B: H12 a H17) .....	34
4.3.1 Diseño de preguntas en el cuestionario .....	34
4.3.2 Objetivo y principales características de las preguntas .....	35
4.3.3 Resultados .....	35
4.4 Cantidad de habitaciones, Superficie cubierta y Régimen de tenencia de la vivienda (A: H23 a H25) (B: H18 a H19) .....	37
4.4.1 Diseño de preguntas en el cuestionario .....	37
4.4.2 Objetivo y principales características de las preguntas .....	38
4.4.3 Resultados .....	38
4.5 Climatización y Equipamiento (A: H26 a H28) .....	40
4.5.1 Diseño de preguntas en el cuestionario .....	40
4.5.2 Objetivo y principales características de las preguntas .....	42
4.5.3 Resultados .....	42
4.6 Relación o parentesco con la (el) jefa(e) o persona de referencia del hogar (P1) .....	43
4.6.1 Diseño de preguntas en el cuestionario .....	43
4.6.2 Objetivo y principales características de las preguntas .....	44
4.6.3 Resultados .....	44
4.7 Sexo, Edad y Fecha de nacimiento (P2 y P4) .....	45

4.7.1 Diseño de preguntas en el cuestionario .....	45
4.7.2 Objetivo y principales características de las preguntas .....	45
4.7.3 Resultados .....	45
4.7.4 Evaluación demográfica de la variable edad .....	46
4.8 Identidad de género (P3) .....	46
4.8.1 Diseño de preguntas en el cuestionario .....	46
4.8.2 Objetivo y principales características de las preguntas .....	49
4.8.3 Resultados .....	49
4.9 Educación (A: P5 a P10) (B: P5 a P8) .....	50
4.9.1 Diseño de preguntas en el cuestionario .....	50
4.9.2 Objetivo y principales características de las preguntas .....	51
4.9.3 Resultados .....	51
4.10 Migraciones (A: P11 a P13) (B: P9) .....	54
4.10.1 Diseño de preguntas en el cuestionario .....	54
4.10.2 Objetivo y principales características de las preguntas .....	55
4.10.3 Resultados .....	55
4.11 Discapacidad y Previsión social (A: P14 a P16) .....	58
4.11.1 Diseño de preguntas en el cuestionario .....	58
4.11.2 Objetivo y principales características de las preguntas .....	59
4.11.3 Resultados .....	59
4.12 Etnias (A: P17) (B: P10) .....	62
4.12.1 Diseño de preguntas en el cuestionario .....	62
4.12.2 Objetivo y principales características de las preguntas .....	62
4.12.3 Resultados .....	62
4.13 Número de documento nacional de identidad (DNI) (A: P18) (B: P11) .....	64
4.13.1 Diseño de preguntas en el cuestionario .....	64
4.13.2 Objetivo y principales características de las preguntas .....	64
4.13.3 Resultados .....	65
4.14 Religión (A: P19) .....	66
4.14.1 Diseño de preguntas en el cuestionario .....	66
4.14.2 Objetivo y principales características de las preguntas .....	66
4.14.3 Resultados .....	66
4.15 Situación conyugal (A: P20) .....	67
4.15.1 Diseño de preguntas en el cuestionario .....	67
4.15.2 Objetivo y principales características de las preguntas .....	67
4.15.3 Resultados .....	67
4.16 Actividad económica (A: P21 a P33) (B: P12 a P16) .....	68
4.16.1 Diseño de preguntas en el cuestionario .....	68
4.16.2 Objetivo y principales características de las preguntas .....	70
4.16.3 Resultados .....	71
4.17 Fecundidad (A: P34 a P36) .....	76
4.17.1 Diseño de preguntas en el cuestionario .....	76
4.17.2 Objetivo y principales características de las preguntas .....	76

4.17.3 Resultados .....	76
<b>5. Evaluación de calidad de la variable Número de DNI .....</b>	<b>79</b>
5.1 Introducción .....	79
5.2 Evaluación de calidad .....	79
5.3 Número de DNI declarado .....	82
5.4 Relación respecto a la edad .....	90
5.5 Casos aparentemente anómalos (C y D) .....	92
5.6 Comparación entre los casos con números de DNI aptos para ser pareables y el resto (sin información para parear) .....	92
5.7 Conclusiones .....	88
<b>6. Evaluación de la codificación de preguntas abiertas .....</b>	<b>89</b>
6.1 Introducción .....	89
6.2 Migraciones .....	90
6.2.1 Codificación de Provincia .....	90
6.2.2 Codificación de País .....	91
6.3 Pueblos originarios .....	92
6.4 Carrera universitaria y terciaria .....	94
6.5 Rama de actividad económica .....	96
6.5.1 Codificación con la Clasificación de Actividades Económicas para Encuestas Sociodemográficas (CAES) .....	97
6.5.2 Codificación con la Clasificación Nacional de Actividades Económicas (ClaNAE) .....	99
6.6 Ocupación .....	101
6.7 Conclusiones .....	103
<b>7. Evaluación de calidad de los datos provenientes de la capacitación de “emergencia” y de “estudiantes terciarios” .....</b>	<b>104</b>
7.1 Introducción .....	104
7.2 Bloque Vivienda (V1 a V7) .....	105
7.3 Bloque Características del Hogar y de la Vivienda (A: H9 a H28 y B: H9 a H19) .....	106
7.4 Bloque Población .....	109
7.4.1 Relación de parentesco, sexo, identidad de género, edad y fecha de nacimiento (P1 a P4) .....	109
7.4.2 Educación (A: P5 a P10 y B: P5 a P8) .....	109
7.4.3 Migraciones (A: P11 a P13 y B: P9) .....	110
7.4.4 Discapacidad y Previsión social (A: P 14 a P16) .....	111
7.4.5 Etnias (A: P17 y B: P10) .....	112
7.4.6 Número de Documento Nacional de Identidad (DNI) (A: P18 y B: P11) .....	113
7.4.7 Situación conyugal (A: P20) .....	113
7.4.8 Actividad económica (A: P21 a P33 y B: P12 a P16) .....	113
7.4.9 Fecundidad (A: P34 a P36) .....	115
7.5 Conclusiones .....	115
<b>8. Conclusiones y recomendaciones .....</b>	<b>116</b>
8.1 Recomendaciones relativas a organizar una base de datos con estructura jerárquica .....	116
8.2 Recomendaciones respecto del proceso de escaneo, digitalización y captura de los datos .....	117
8.3 Recomendaciones relativas al diseño conceptual .....	118

8.3.1 Tipo de vivienda, Condición de ocupación y Detección de hogares (V1 a V7) .....	118
8.3.2 Recomendaciones .....	118
8.3.3 Otras modificaciones .....	119
8.4 Materiales de los pisos, las paredes y los techos (A: H9 a H13) (B: H9 a H11) .....	119
8.4.1 Recomendaciones .....	120
8.5 Condiciones sanitarias, Combustible y Agua caliente (A: H14 a H22) (B: H12 a H17) .....	120
8.5.1 Recomendaciones .....	120
8.6 Cantidad de habitaciones, Superficie cubierta y Régimen de tenencia de la vivienda (A: H23 a H25) (B: H18 a H19) .....	121
8.6.1 Recomendaciones .....	121
8.7 Climatización y Equipamiento (A: H26 a H28) .....	121
8.7.1 Recomendaciones .....	121
8.8 Relación o parentesco con la (el) jefa (e) o persona de referencia del hogar (P1) .....	122
8.8.1 Recomendaciones .....	122
8.9 Sexo, Edad y Fecha de nacimiento (P2 y P4) .....	122
8.9.1 Recomendaciones .....	122
8.10 Identidad de género (P3) .....	122
8.10.1 Recomendaciones .....	123
8.11 Educación (A: P5 a P10) (B: P5 a P8) .....	123
8.11.1 Recomendaciones .....	123
8.12 Migraciones (A: P11 a P13) (B: P9) .....	123
8.12.1 Recomendaciones .....	124
8.13 Discapacidad y Previsión social (A: P14 a P16) .....	124
8.13.1 Recomendaciones .....	124
8.14 Etnias (A: P17) (B: P10) .....	125
8.14.1 Recomendaciones .....	125
8.15 Número de DNI (A: P18) (B: P11) .....	125
8.15.1 Recomendaciones .....	125
8.16 Religión (A: P19) .....	126
8.16.1 Recomendaciones .....	126
8.17 Situación conyugal (A: P20) .....	126
8.17.1 Recomendaciones .....	126
8.18 Actividad económica (A: P21 a P33) (B: P12 a P16) .....	126
8.18.1 Recomendaciones .....	127
8.19 Fecundidad (A: P34 a P36) .....	127
8.19.1 Recomendaciones .....	128
8.20 Recomendaciones sugeridas para el mejoramiento del proceso de codificación .....	128
8.21 Recomendaciones sugeridas en relación con las estrategias de capacitación y el reclutamiento de los recursos humanos .....	128
<b>9. Bibliografía .....</b>	<b>130</b>

# Introducción

Este informe complementa los resultados brindados en el *Informe de evaluación de la segunda prueba piloto. Hacia la Ronda 2020*,<sup>1</sup> difundido en diciembre de 2019.

En esta ocasión se presentan los resultados de la evaluación cuantitativa de los datos que, en conjunto con la evaluación cualitativa elaborada en el primer informe, permitirá detectar y analizar la magnitud y la naturaleza de los errores. Reflexionar acerca de los posibles factores invalidantes que afectan de manera particular y variable la calidad de los datos es lo que permite orientar el modo de corregirlos y diseñar una mejor medición hacia el Censo Experimental.

En el primer capítulo se exponen los aspectos conceptuales que orientan la evaluación de calidad de los datos censales. El segundo apartado presenta una síntesis del proceso de preparación de la base de datos requerida para realizar una correcta evaluación de estos. El tercer capítulo incluye un análisis de la lectura e interpretación de las marcas, los números y las letras que surgen del proceso de escaneo de los cuestionarios censales. El cuarto apartado examina los errores de calidad de cada una de las preguntas incluidas en los cuestionarios básico y ampliado, y avanza en delinear una interpretación de sus posibles causas. En el quinto capítulo se comenta el resultado de la evaluación de calidad de la variable Número de documento nacional de identidad. En el capítulo seis se puede observar el resultado de la evaluación de la codificación de las preguntas abiertas: País y Provincia, Carrera universitaria o terciaria, Pueblo indígena u originario, Rama de actividad y Ocupación. El séptimo capítulo refiere a la evaluación de calidad de los datos relevados específicamente en dos radios censales seleccionados, que estuvieron a cargo de censistas con un perfil o una capacitación diferente respecto del resto. El octavo capítulo sistematiza las conclusiones y recomendaciones derivadas del análisis cuantitativo que orientan el diseño conceptual del cuestionario del Censo Experimental. El Anexo contiene el detalle de cada uno de los tipos de error por área de prueba.

## 1.

## Aspectos conceptuales

¿A qué nos referimos con la evaluación de calidad de los datos? Según las *Buenas prácticas estadísticas que guían al INDEC*,<sup>2</sup> se trata de un conjunto de principios y buenas prácticas que tienen como finalidad mejorar la calidad de las estadísticas oficiales y fortalecer la credibilidad y confianza de las personas usuarias de la información. Este código está basado en los principios fundamentales de las estadísticas oficiales y las mejores prácticas internacionales. Estas recomendaciones se vinculan, entre otros aspectos, con las prácticas relacionadas con el entorno institucional, el proceso estadístico y la producción estadística. En particular esta última remite a tener en cuenta aspectos tales como la pertinencia, precisión y confiabilidad, oportunidad y puntualidad, coherencia y comparabilidad, y accesibilidad y claridad de los datos.

La presente publicación, por su parte, viene a cumplir con uno de estos criterios enmarcados en el entorno de la producción estadística, el de reflejar la realidad de forma precisa y confiable. De hecho, el relevamiento de información censal se realiza conforme con diseños conceptuales y metodologías generadas en el marco de las buenas prácticas estadísticas. Para ello, los datos originales, los resultados intermedios y los productos estadísticos se prueban, validan y documentan con la finalidad de garantizar la precisión y confiabilidad del producto estadístico final.

<sup>1</sup> Disponible en [https://www.indec.gob.ar/ftp/cuadros/publicaciones/censo\\_2020\\_informe\\_2PP.pdf](https://www.indec.gob.ar/ftp/cuadros/publicaciones/censo_2020_informe_2PP.pdf).

<sup>2</sup> Disponible en [https://www.indec.gob.ar/ftp/documentos/principios\\_indec.pdf](https://www.indec.gob.ar/ftp/documentos/principios_indec.pdf).

Es sabido que el proceso de producción de información cuantitativa está sometido a diferentes fuentes de invalidez que afectan la calidad del producto: el dato. En palabras de Chackiel y Mació (1978), Giusti y Massé (1997), CEPAL (2011) y NU (2017), es responsabilidad de los institutos de estadística detectar, desentrañar e interpretar los problemas que aquél presenta.

El objetivo de la evaluación de los datos en un censo de población refiere a la necesidad de ofrecer a los usuarios cierto nivel de confianza cuando los utilicen y explicar los errores existentes en el resultado censal (NU, 2017: p. 118). En el marco de una prueba piloto, el proceso de evaluación de los datos permite anticiparse de modo que los errores detectados puedan ser corregidos y evitados en el censo definitivo.

La producción de datos supone el desarrollo de diferentes etapas lógicas que, mediante la implementación previa de pruebas piloto, en la medida de lo posible, tienden a evitar las fuentes de invalidez, analizando cuidadosamente los errores cometidos con posterioridad a su ejecución y estableciendo para ellos ciertos umbrales de tolerancia.

Las pruebas piloto del censo resultan indispensables para tomar decisiones. Pueden organizarse con diversos objetivos y realizarse en cada una de las etapas del proyecto censal. Estas pruebas suministran información valiosa sobre la organización del operativo de campo, el plan de capacitación, la carga de trabajo de los recursos humanos que conforman la estructura operativa del censo, el plan de procesamiento de los datos y otros aspectos relevantes. En especial, resultan excepcionalmente útiles para descubrir las deficiencias del cuestionario y de las instrucciones impartidas que podrían afectar la calidad de los datos.

En general las estadísticas demográficas remiten a considerar los errores muestrales de manera diferenciada respecto de los errores no muestrales. Los primeros refieren a estadísticas basadas en encuestas por muestreo y requieren evaluar el componente de error estadístico derivado de las metodologías de inferencia de los resultados muestrales a los poblacionales. Ese análisis cuenta con sus propios procedimientos de cálculo y control (CEPAL-UNFPA, 2014: p. 34) y no es objeto de análisis en la presente publicación, que se dirige a evaluar los denominados errores no muestrales.

Los errores censales no muestrales se suelen clasificar en dos categorías: los errores de cobertura y los de contenido (NU, 2017: p. 118). El error de cobertura fue tratado en el primer *Informe de evaluación de la segunda prueba piloto. Hacia la Ronda 2020* y se relaciona con la omisión o duplicación de las unidades de observación, es decir, de individuos, hogares o viviendas. Por su parte, los denominados errores de contenido son los originados en fallas en el registro de las características de las personas, los hogares y las unidades de vivienda y se relacionan con cuestiones referidas a la validez<sup>3</sup> y a la integridad de la información relevada (Giusti y Massé, 1997: p. 13).

Antes de continuar avanzando en describir el tipo de error que es posible detectar a partir de analizar la base de datos, interesa reiterar que el universo de la segunda prueba piloto (II PP) fue el de las viviendas particulares.<sup>4</sup> La prueba se llevó a cabo en un día de operativo, utilizando la entrevista directa como método de empadronamiento, es decir, mediante entrevista cara a cara entre censista y censado. También se aplicó muestreo en el relevamiento, lo que implicó trabajar con dos cuestionarios, uno básico y otro ampliado. Se trató de cuestionarios en papel cuyos datos fueron escaneados para proceder a su digitalización y captura a través de un software de reconocimiento OCR/ICR/OMR<sup>5</sup> hasta conformar la correspondiente base de microdatos.

<sup>3</sup> “Una definición operacional es válida siempre que, cumpliendo las operaciones de medida indicadas en ella, se contempla todo aquello a lo que con su contenido significativo remite el concepto” (Mayntz, Holm y Hubner, 1993).

<sup>4</sup> Vivienda particular es la vivienda destinada al alojamiento de un hogar particular, esté o no habitada. Un hogar particular es la persona o el grupo de personas, parientes o no, que habita o habitan la misma vivienda bajo un régimen familiar y comparten los gastos de alimentación.

<sup>5</sup> Optical Character Recognition (OCR): para reconocimiento de caracteres.

Intelligent Character Recognition (ICR): para interpretación de números y letras manuscritas.

Optical Mark Reading (OMR): para reconocimiento de marcas.

En referencia a la técnica implementada de cuestionario papel y escaneo de los datos relevados, los errores que desarrollaremos pueden clasificarse según:<sup>6</sup>

**1. El momento en que se producen: antes, durante o después del relevamiento.**

- 1.1 Antes del relevamiento: por insuficiente claridad en los conceptos; operacionalización errónea de estos, redacción equívoca de las preguntas y de las categorías de respuesta; dificultades en la secuencia de las preguntas; problemas en el diseño gráfico del cuestionario; instrucciones inespecíficas de llenado; distorsión en las definiciones durante la etapa de capacitación de los censistas.
- 1.2 Durante el relevamiento: por una inexacta interpretación o disposición de la persona censada en el momento de responder una pregunta; una inadecuada interpretación o disposición del censista que genera un error de registro; dificultades en el modo de escribir del censista o deficiencia en el registro de las marcas.
- 1.3 Despues del relevamiento: por problemas en la manipulación física de los cuestionarios o durante el momento del escaneo, la lectura, la interpretación o el procesamiento de los datos.

**2. El tipo de invalidez que los hayan originado: al azar o sistemáticos.**

- 2.1 Al azar: afectan a todos los valores de las variables de manera uniforme, son de baja ocurrencia y no distorsionan las distribuciones de los datos obtenidos.
- 2.2 Sistemáticos: de repetirse el censo en las mismas condiciones, estos errores volverían a producirse muy probablemente en las mismas variables de los mismos registros. Suelen afectar a un grupo específico de variables, cuestionarios o registros, y ocasionan cierta distorsión en las distribuciones.

**3. La fuente de invalidez.**

- 3.1 Limitaciones conceptuales: la persona censada puede tener una idea diferente sobre un concepto o asumir un criterio diferente acerca de lo que se está preguntando e involuntariamente brindar una respuesta incorrecta. También se presenta cuando quien censa indaga preguntas que no corresponden a un determinado universo de población.
- 3.2 Interpretación errónea de la pregunta: la persona censada puede no comprender correctamente el dato que se le solicita e incurrir en una respuesta equivocada. Quien censa puede cometer este error cuando completa las respuestas a preguntas que requieren de mayor indagación.
- 3.3 Ignorar la respuesta correcta: en la mayoría de los casos, cuando la persona censada responde acerca de las características de una tercera persona.
- 3.4 Error deliberado: cuando la persona censada falsea u omite la información de manera deliberada y consciente.

<sup>6</sup> Para mayor detalle, ver Giusti y Massé (1997) y CEPAL (2011).

#### 4. Se detecten en un registro o entre registros de la base de datos.

4.1 Los errores a nivel de registro se analizan de manera independiente para cada una de las variables relativas a la vivienda, el hogar o la población y se clasifican como:

- No respuesta (NR): falta de respuesta en una variable o pregunta que debía ser investigada.
- Multimarca (MM): presencia de más de una marca en una variable o pregunta, es decir que una pregunta tiene marca en más de una categoría de respuesta.
- Error de secuencia (ES): marca que se presenta en una variable o pregunta que no debía ser investigada y que se produce debido a no haber seguido la secuencia o el camino indicado en el instrumento de recolección.
- Inválido (INV): error que se presenta en las preguntas que requieren de respuestas alfanuméricas. Se originan en un error de interpretación de los números y las letras transcriptos en el cuestionario.

Una medida de la calidad de los datos en término de análisis de contenido se podrá obtener mediante la cuantificación de los errores que presentan cada una de las variables, detectadas a partir de evaluar el nivel de no respuesta, multimarca o inconsistencia de los datos, de acuerdo con pautas de validación predeterminadas.

La magnitud del error surge de considerar el cociente entre los valores con error y los valores esperados de acuerdo con lo establecido en el instrumento de recolección de datos. Debe ser analizada a partir de los umbrales de error especificados históricamente para cada variable del cuestionario, a partir de la evaluación de calidad implementada en censos previos. Estos valores hacen de referencia, de modo que las variables o preguntas que superen los valores de error previamente establecidos deben ser sometidas a acciones correctivas o bien debe evaluarse su exclusión del futuro cuestionario a diseñar.

#### 4.2 Entre variables de un mismo registro o entre registros.

Inconsistencia: cuando se observan relaciones ilógicas o inconsistentes entre variables de un mismo registro o entre registros.

Para evaluar estos errores que afectan la validez y confiabilidad de la pregunta, se realiza el análisis de calidad de la información mediante el estudio de las estructuras de cada una de las variables y el diseño de un plan de tabulados de calidad de la información que permita analizar cada una de las inconsistencias posibles o las más comunes.

## 2.

# Organización de la base de datos

Con posterioridad al relevamiento censal, se envían todos los cuestionarios en papel a la sede responsable del escaneo, la digitalización y la interpretación de los datos para que genere la primera base de datos sin estructura jerárquica y sin consistir.

En este marco, el proceso que debe seguir cada cuestionario censal es el siguiente:

1. Ordenamiento: se realiza con el objetivo de que los cuestionarios ingresen a la siguiente etapa del proceso de manera metódica, según la ubicación geográfica a la que corresponde, la numeración consecutiva de las viviendas, la de los hogares dentro de cada vivienda y la de cuestionarios consecutivos<sup>7</sup> al interior de cada hogar.
2. Guillotinado: dado que los cuestionarios cuentan con un engomado que reduce la posibilidad de deterioro por su manipulación durante la entrevista, la tarea se ejecuta con el objetivo de separar cada una de las hojas del cuestionario para escanearlas.
3. Escaneo: se captura y digitaliza el cuestionario censal en papel para transferirlo a una computadora en formato de imagen o a una base de microdatos para su procesamiento.
4. Reconocimiento e interpretación: a través del *software* de reconocimiento OCR/ICR/OMR, se extraen los datos de las imágenes de los cuestionarios censales y se construye una base de datos que permite ser procesada.

La primera base de datos que resulta de estos procesos está representada por una matriz cuyas filas corresponden a cada hoja del cuestionario escaneada e interpretada, y en las columnas muestra cada una de las preguntas del cuestionario. En este sentido, para cada cuestionario censal, el primer renglón de la base de datos contiene las preguntas relativas a la vivienda y a las características habitacionales del hogar, y los siguientes, a los cuestionarios de población.

En esta base de datos inicial, cada fila es independiente de las otras, es decir, las filas no se encuentran asociadas ni como viviendas, ni como hogares al interior de la vivienda, ni como población dentro de cada hogar. Por lo tanto, para efectuar un correcto análisis de la información, primero es necesario realizar un ordenamiento que permita conformar una base de datos en la que estén diferenciadas las viviendas desocupadas respecto de aquellas ocupadas, es decir, con personas presentes al momento de llegar el censista a la vivienda. A su vez, al interior de las viviendas ocupadas se requiere organizar tanto los hogares que las conforman como las personas correspondientes a esos hogares. Asimismo, dado que todas las hojas de los cuestionarios son escaneadas y sujetas al *software* de reconocimiento, independientemente de que estén completas con datos o vacías, es preciso eliminar estas últimas, es decir, aquellas que efectivamente no constituyen un registro de vivienda, hogar o población.

A continuación, se detalla el proceso de estructuración y organización de la base de datos implementado en ocasión de la II PP.

<sup>7</sup> Debido a que el cuestionario ampliado contempla hasta 7 cuestionarios de población y el cuestionario básico, hasta 6 cuestionarios de población, los hogares con más personas requerían abrir cuestionarios consecutivos.

## 2.1

# Implementación del proceso de estructura jerárquica

El proceso de organización de la base de datos, llamado “estructura jerárquica”, consiste en ordenar las unidades de análisis del censo, es decir, ubicar a las personas en sus hogares, como así también los hogares en las viviendas en donde fueron censados; e identificar también las viviendas en las que no se encontraron personas al momento del relevamiento. Esta actividad se realizó mediante la aplicación de un programa informático diseñado con base en criterios que permitieron establecer la correcta correspondencia entre los registros de población, de hogares y de viviendas.<sup>8</sup>

## 2.2

# Tipo de errores

Si bien el programa informático que se aplicó para generar una base de datos estructurada jerárquicamente se diseñó con base en la experiencia de censos anteriores y fue probado en forma previa al trabajo de campo de la II PP, al testearse en bases de datos simuladas, se debieron realizar ajustes significativos<sup>9</sup> al momento de su implementación, debido a la presencia de diferentes tipos de errores no esperados en la captura de los datos.

A continuación, se presentan los principales errores detectados durante la implementación del proceso de estructura jerárquica en la base de datos y la procedencia de aquellos según la etapa en la que se considera se introdujeron:

## 1. Antes del relevamiento

- Impresión de códigos de barra repetidos para cuestionarios básico y ampliado.<sup>10</sup>
- Dificultades en la operacionalización de la identificación de hogares.<sup>11</sup>
- Falta de claridad en la instrucción acerca de cómo relevar los hogares de más de 6 o 7 personas y más de un hogar en la vivienda.<sup>12</sup>

## 2. Durante el relevamiento

- Completamiento incorrecto u omisión de campos del cuestionario, imprescindibles para generar una base de datos estructurada jerárquicamente. Así, las cédulas censales contenían errores introducidos por quienes censaron en los datos de ubicación geográfica, el número de vivienda o de hogar, la condición de ocupación y desocupación de la vivienda, o la detección de hogares en la vivienda.

<sup>8</sup> Cabe consignar que la actividad consistió en una instancia inicial de desarrollo conceptual de las pautas de detección de errores para proceder a su posterior programación en lenguaje informático desarrollada por un equipo de trabajo conjunto de la Dirección de Estadísticas Poblacionales y de la Dirección de Informática del INDEC.

<sup>9</sup> Si bien resulta esperable realizar ajustes a los programas informáticos previamente diseñados, una vez obtenida la base de datos resultante del escaneo, la lectura y el reconocimiento de datos, el volumen y tipo de errores detectados implicaron mayores ajustes y controles de los esperados, motivo por el cual el proceso de organización y estructuración de la base de datos demandó unos dos meses de trabajo.

<sup>10</sup> El código de barras es una imagen cuadrada o rectangular que consiste en una serie de líneas paralelas negras y espacios blancos con un ancho variable y que puede ser escaneada. En el caso de la II PP el código debía corresponder a un número único e irrepetible correspondiente al anverso o reverso de cada hoja del cuestionario censal.

<sup>11</sup> Refiere a las preguntas incluidas en el cuestionario censal con el objetivo de detectar la existencia de un hogar o más de un hogar en la vivienda que está siendo censada.

<sup>12</sup> El cuestionario ampliado contaba con espacio para censar 7 personas y el básico, para 6 personas. Cuando el hogar contaba con más personas, el censista debía abrir otro cuestionario, transcribir los mismos datos de la ubicación geográfica y de la pregunta V7 (número de hogar) y continuar censando a las personas en el segundo cuestionario.

### 3. Después del relevamiento

- Ordenamiento incorrecto de los cuestionarios censales en papel, en la sede del operativo o bien en el momento de preparación para el ingreso al escaneo de los datos, etapa que incluye también el proceso de guillotinado previamente mencionado.
- Escaneo de cuestionarios sobrantes con todas sus páginas en blanco.
- Incorrecta lectura e interpretación de marcas, letras o números por parte del *software* de reconocimiento.

## 2.3

## Etapas del proceso de organización de la base de datos y ajustes realizados

En este apartado se presentan cada una de las etapas del proceso de estructura jerárquica, los errores encontrados y las correcciones realizadas hasta generar una base de microdatos estructurada jerárquicamente.

La implementación del proceso de estructura jerárquica consistió en:

- 1. Identificar los cuestionarios duplicados.** Detectar cuestionarios escaneados dos o más veces a partir de un código de barra único que identifica un cuadernillo/cuestionario censal básico o ampliado.

**Error:** se hallaron cuestionarios básico y ampliado con el mismo código de barra preimpreso; cuestionarios sin código de barra; el mismo código de barra entre diferentes cuestionarios básicos; el mismo código de barra entre cuestionarios ampliados e idénticos hogares con diferentes códigos de barra. Asimismo, en la base de datos se detectaron registros correspondientes a cuestionarios utilizados en una prueba de lectura anterior y que no habían sido previamente eliminados al proceder a generar la base de microdatos correspondiente a la II PP.

**Corrección implementada:** se compararon los datos de ubicación geográfica de la vivienda y vivienda-hogar para identificar registros idénticos. También se detectaron los registros correspondientes a cuestionarios de prueba y se eliminaron de la base de datos.

- 2. Validar el registro “persona”.** Determinación del registro “persona válida”. El diseño del programa informático de estructura jerárquica determinaba que un registro correspondiente a la unidad “persona” sería considerado válido con la sola presencia de al menos una marca en el cuestionario de población.

**Error:** validación de personas que debían considerarse inválidas, debido a que la marca escaneada se originaba en la presencia de marcas únicas mal borradas por quien había censado o con errores de lectura.

**Corrección implementada:** se modificó la definición del registro “persona válida” por la presencia de al menos tres marcas detectadas en el conjunto de las variables Relación o parentesco, Sexo, Identidad de género, Edad, Asistencia escolar y Lugar de nacimiento.

**3. Validar el registro de “vivienda-hogar ocupado”.** Determinación de registro “vivienda-hogar válido ocupado”.

**Error:** escaneo y lectura de cuestionarios sobrantes vacíos y de cuestionarios de vivienda-hogar con marcas originadas en un error de lectura.

**Corrección implementada:** se analizó la existencia de al menos una persona en el hogar y en caso afirmativo se lo consideró vivienda-hogar ocupado.

**4. Validar el registro de “vivienda-hogar desocupado”.** Determinación del registro “vivienda-hogar válido desocupado”. El diseño del programa informático de estructura jerárquica determinaba que un registro de vivienda-hogar sería considerado válido con la sola presencia de al menos una marca en cualquier pregunta del cuestionario vivienda-hogar, aunque no existieran personas relevadas en él.

**Error:** validación de viviendas-hogar inválidas debido a la presencia de marcas únicas mal borradas por el censista o con errores de lectura.

**Corrección implementada:** se modificó la definición de registro “vivienda-hogar válido” por la presencia de al menos tres marcas en las variables ubicación geográfica, tipo de vivienda, vivienda principal y condición de ocupación de la vivienda.

**5. Asignar personas a hogares, hogares a viviendas y creación de un identificador único.<sup>13</sup>** Identificación y ubicación de las personas en los hogares y de los hogares en las viviendas en donde fueron censados. Este paso comprendió un proceso especial para los hogares de más de 6 personas para el cuestionario básico o 7 personas para el cuestionario ampliado.

Cada vivienda tiene una única identificación que surge de la concatenación de los datos de Ubicación geográfica y Número de vivienda; si a esta concatenación se adiciona el Número de hogar, se obtiene la clave única de vivienda-hogar.

**Error:**

- Problemas de completamiento, por parte de quien realizó el censo, de la Ubicación geográfica y/o del número del segundo hogar y sucesivos, que generaban claves de vivienda diferentes para hogares en la misma vivienda.
- Problemas en la lectura e interpretación de letras y números en la variable Ubicación geográfica que generaban claves de vivienda-hogar diferentes.

**Corrección implementada:** se redefinió el proceso y se adaptó para que abarque los diferentes errores encontrados, con lo que se detectaron diferentes hogares en la misma vivienda y se arribó a una clave de vivienda y una clave de vivienda-hogar que permitió identificar las viviendas, los hogares y las personas.

---

<sup>13</sup> Un identificador es un número único que permite ubicar de manera única cada uno de los registros de la base de datos.

**6. Unir hogares con más de 6 o 7 personas.** Unir cuestionarios con más de 6 o 7 personas pertenecientes al mismo hogar.

Error:

- Problemas en el completamiento del número de hogar.
- Desorden de los cuestionarios, que provocó que las personas número 7 u 8, pertenecientes a un mismo hogar, se registraran como hogares diferentes en la base de datos, dado que no aparecían en la base de manera consecutiva.
- Error de lectura e interpretación de la variable que indicaba el número de persona asignado por el censista (por ejemplo, la persona censada en tercer término –persona 03– aparecía relevada erróneamente con el número 08), lo que alteraba la posibilidad de identificar a las personas que pertenecían al mismo hogar.

**Corrección implementada:** se realizó un proceso manual de identificación de casos de hogares con más de 6 o 7 personas que no pudieron ser detectados por el programa informático diseñado y se los “unió” de manera manual.

**7. Determinar una sola persona jefa de hogar de 14 años o más:** determinación de una jefatura única de 14 años y más en cada hogar.<sup>14</sup>

Error: presencia de símbolos en la variable edad.

Corrección: creación de una variable intermedia de edad en la que se le imputó blanco.

**8. Determinar la condición de ocupación de vivienda/hogar.** Último paso del proceso de estructura jerárquica en el que se crea una variable nueva en la base de datos de acuerdo con la presencia o ausencia de registros de personas en ese hogar. Una vez que quedó determinado qué personas integraban efectivamente cada hogar y qué hogares integraban cada vivienda, se pudo analizar por cada clave de vivienda-hogar qué viviendas se encontraban acompañadas por registros de persona (viviendas ocupadas) y cuáles no (viviendas desocupadas).

A continuación, se presenta el total de registros de vivienda, hogar y persona por tipo de cuestionario incluido en la base de datos previa al proceso de estructura jerárquica y con posterioridad a la implementación de dicho proceso y sus ajustes.

---

<sup>14</sup> Para proceder a conformar una base de datos estructurada jerárquicamente, se requiere considerar una persona de referencia del hogar cuyo registro corresponde al primero de aquellos que integran el hogar. En general suele remitirse a aquel como la persona jefe/jefa del hogar, entre cuyas características se encuentra la edad mínima de 14 años.

**Cuadro 2.3.1 Total de registros preestructura jerárquica y posestructura jerárquica por tipo de cuestionario. II PP 2019**

Tipo de cuestionario	Tipo de registro	Base de datos inicial	Registros vacíos	Base de datos sin registros vacíos	Registros inválidos	Base de datos final
Cuestionario ampliado	Vivienda-hogar	6.424	89	6.335	76	<b>6.259</b>
	Persona	44.923	31.742	13.181	2.988	<b>10.193</b>
	<b>Total</b>	<b>51.347</b>	<b>31.831</b>	<b>19.516</b>	<b>3.064</b>	<b>16.452</b>
Cuestionario básico	Vivienda hogar	1.775	19	1.756	53	<b>1.703</b>
	Persona	10.650	7.071	3.579	2.007	<b>1.572</b>
	<b>Total</b>	<b>12.425</b>	<b>7.090</b>	<b>5.335</b>	<b>2.060</b>	<b>3.275</b>

**Fuente:** INDEC. Base de datos sin estructura jerárquica y base de datos con estructura jerárquica y sin consistir. II PP 2019.

Los totales de viviendas, hogares y personas de la base de datos final, que se presentan en el cuadro 2.3.2, corresponden a la base de datos con estructura jerárquica y sin consistir, insumo básico para llevar a cabo la evaluación de calidad cuantitativa de la información de la II PP.

**Cuadro 2.3.2 Total de viviendas, hogares y personas de la base de datos estructurada sin consistir. II PP 2019**

Unidad de observación	Tipo de cuestionario		
	Cuestionario básico	Cuestionario ampliado	Total
Viviendas	1.689	6.199	7.888
Hogares	1.703	6.259	7.962
Personas	1.572	10.193	11.765

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

### 3.

# Evaluación de calidad de la lectura e interpretación de los datos

## 3.1

### Introducción

Tal como se mencionó, en el marco de las tareas preparatorias del Censo de la ronda 2020, el INDEC implementó la II PP el domingo 8 de septiembre de 2019. Para ello realizó un censo de hecho en un día de operativo, asociado a una semana de supervisión y recuperación. El método de empadronamiento fue la entrevista directa y se aplicó muestreo en el relevamiento, lo que implicó trabajar con dos cuestionarios, uno básico (B) y otro ampliado (A), todos ellos en papel. El universo fue el de las viviendas particulares, es decir, aquellas destinadas al alojamiento de hogares particulares, estén o no habitadas, en las que las personas conviven bajo un régimen (INDEC, 2019).

El proceso de impresión de los cuestionarios y planillas y de digitalización y captura de datos mediante lectura por escáner (OCR/ICR) estuvo a cargo del Correo Argentino.

El presente apartado tiene como objetivo evaluar el nivel y los tipos de errores más comunes del proceso mencionado, tratando de identificar aquellos correspondientes a la lectura e interpretación de las imágenes.

## 3.2

### Metodología

La digitalización y captura de los datos se puede dividir en tres etapas o procesos:

1. el escaneo del soporte papel para obtener una imagen digital,
2. la lectura de la información digitalizada,
3. la interpretación de esta información para la generación de una base de datos.

Durante estos tres procesos se pueden originar una serie de errores. En la primera etapa estos están ligados a la integridad de los datos, es decir, a la omisión o duplicación de cuestionarios censales. En cambio, en la tercera etapa, los errores están dados por una interpretación incorrecta de la imagen del cuestionario, que luego se traduce en la base de datos.

Para la evaluación de los errores de interpretación de los datos se consideró una muestra de 310 cuestionarios, de un total de 6.323 de tipo A y 1.720 de tipo B<sup>15</sup> y se realizaron dos procedimientos:

1. Se evaluó si las imágenes escaneadas de cada una de las preguntas del cuestionario coincidían con la información registrada en la base de datos. De esta forma se pudo cuantificar la cantidad de respuestas donde no coincidían.

<sup>15</sup> Con el asesoramiento de la Dirección Nacional de Metodología Estadística, la Dirección de Estadísticas Poblacionales procedió a realizar una muestra aleatoria simple rondando el 3% de los cuestionarios con al menos una persona presente, respetando el balance entre tipo de cuestionarios (A y B).

2. Se seleccionaron las carillas correspondientes a las primeras dos personas listadas en la vivienda<sup>16</sup> y se evaluaron los tipos de error en las preguntas Número de Documento Nacional de Identidad (Número de DNI), Edad, Sexo, Provincia de nacimiento y Tarea que realiza en su ocupación principal.

### 3.3

## Evaluación

En la primera etapa, para los 310 cuestionarios de la muestra se verificó de manera visual si la información registrada en cada una de las preguntas de la base de datos coincidía con la imagen digital del cuestionario en papel.

Se consideró que había un error cuando el contenido de una pregunta difería entre la imagen del cuestionario y la base de datos, por ejemplo, cuando en el cuestionario la edad de una persona era “56” y en la base de datos figuraba “55”. Si en una misma pregunta había más de un error, se contabilizó solo una vez, por ejemplo, si en el cuestionario decía “JUJUY” y en la base de datos se registró “JYVUY”. Esto incluye aquellas situaciones en las que se detectó no respuesta en la imagen y respuesta en la base de datos o viceversa.

En el cuadro 3.3.1, se clasifican los cuestionarios de la muestra según tengan al menos un error de interpretación en los registros de vivienda, hogar y población. Para el cuestionario B, el porcentaje de registros vivienda-hogar con al menos un error de interpretación es del 70,5%, mientras que, para el cuestionario A, los registros con al menos un error alcanzan el 67,7%. En el caso de los registros de personas, aquellos con al menos un error representan el 53,7% en el cuestionario B y 66,1% en el cuestionario A. Es decir que aproximadamente 2 de cada 3 registros de vivienda-hogar y personas censadas poseen algún error en sus atributos relevados como resultado del proceso de interpretación de la lectura, independientemente de su gravedad y de la posibilidad de que sea posible recuperar su valor preciso.

**Cuadro 3.3.1 Registros por tipo de registro, tipo de cuestionario y cantidad de errores en la interpretación de los datos. II PP 2019**

Tipo de registro	Cuestionario	Sin errores		Con al menos un error		Total	
		N	%	N	%	N	%
Vivienda/ hogar	<b>Total</b>	<b>99</b>	<b>31,9</b>	<b>211</b>	<b>68,1</b>	<b>310</b>	<b>100,0</b>
	Básico	13	29,5	31	70,5	44	100
	Ampliado	86	32,3	180	67,7	266	100
Personas	<b>Total</b>	<b>320</b>	<b>35,4</b>	<b>585</b>	<b>64,6</b>	<b>905</b>	<b>100,0</b>
	Básico	50	46,3	58	53,7	108	100
	Ampliado	270	33,9	527	66,1	797	100

**Fuente:** INDEC. Base de datos sin estructura jerárquica y sin consistir. II PP 2019.

Una forma complementaria de evaluación es considerar la respuesta como unidad de análisis, independientemente de las personas o viviendas-hogares que les correspondan, para poder analizar la magnitud del error para cada tipo de pregunta.

<sup>16</sup> En las dos primeras carillas es donde se concentra la mayor parte de las respuestas de las personas del hogar. Si bien la cantidad de respuestas incluidas en las dos primeras carillas resulta óptima para el análisis descriptivo, esta selección deja por fuera el efecto cansancio por parte del censista cuando se trata de hogares con más de dos personas, cuyo porcentaje es de aproximadamente 55,9% del total de hogares.

Al analizar el error de interpretación a nivel de las respuestas, en el cuadro 3.3.2 se observa que el 3,5% de las respuestas registra algún tipo de error, siendo más alto para el cuestionario A que para el cuestionario B (3,6% y 2,2%, respectivamente). Sin embargo, el nivel de error muestra diferencias si se considera el tipo de pregunta: las preguntas abiertas (numéricas y alfanuméricas) muestran un nivel de error mayor que las preguntas cerradas (5,3% y 1,8%, respectivamente). Entre las preguntas abiertas, las que contienen letras presentan niveles de error más elevados que las que contienen números. En este sentido, las preguntas abiertas relativas a la descripción de las características del empleo son las que presentan el nivel de error más alto para el cuestionario A (12,3%), seguidas por las preguntas abiertas referidas a la Ubicación geográfica para ambos cuestionarios (8,2%).

En las preguntas numéricas el nivel de error es más bajo, entre el 2,8 % y 2,3% en el cuestionario A y el 1,2% y 0,3% en el cuestionario B. Por último, las preguntas cerradas tienen un nivel de error menor al 2% (1,8% para el cuestionario A y 0,6% para el cuestionario B), que aumenta en el caso de las preguntas en forma de grilla del cuestionario A (tenencia sobre bienes y climatización del hogar y discapacidad de la población), que muestra un porcentaje de error del 3,0%.

Se debe tener en cuenta que, si bien en las preguntas abiertas de letras,<sup>17</sup> los errores resultan más elevados que en el resto de las variables, muchas veces no tienen un impacto directo en la base de datos final porque pueden ser interpretados correctamente en el proceso de codificación de los datos. En cambio, los errores de lectura en las preguntas abiertas de números y en las preguntas cerradas tienen un impacto directo en la información final de la base, ya que no pasan por ningún proceso de codificación.

**Cuadro 3.3.2 Cantidad de respuestas, cantidad de errores de interpretación y porcentaje de error por tipo de cuestionario y tipo de pregunta. II PP 2019**

Tipo de pregunta	Cuestionario ampliado			Cuestionario básico			Total		
	Respuestas	Errores	%	Respuestas	Errores	%	Respuestas	Errores	%
<b>Total</b>	<b>89.602</b>	<b>3.233</b>	<b>3,6</b>	<b>6.436</b>	<b>142</b>	<b>2,2</b>	<b>96.038</b>	<b>3.375</b>	<b>3,5</b>
<b>Preguntas abiertas</b>	<b>43.599</b>	<b>2.382</b>	<b>5,5</b>	<b>3.344</b>	<b>123</b>	<b>3,7</b>	<b>46.943</b>	<b>2.505</b>	<b>5,3</b>
<b>Letras</b>	<b>24.188</b>	<b>1.892</b>	<b>7,8</b>	<b>1664</b>	<b>102</b>	<b>6,1</b>	<b>17.349</b>	<b>1.280</b>	<b>7,4</b>
Actividad económica	3.188	392	12,3	///	///	///	///	///	///
Ubicación geográfica	1.596	125	7,8	264	28	10,6	1.860	153	8,2
Migración	6.377	484	7,6	304	16	5,3	6.681	500	7,5
Pueblo indígena u originario	1.063	79	7,4	152	3	2,0	1.215	82	6,7
Grilla de personas	6.649	490	7,4	944	55	5,8	7.593	545	7,2
Dirección de trabajo	4.252	269	6,3	///	///	///	///	///	///
Carrera	1.063	53	5,0	///	///	///	///	///	///
<b>Números</b>	<b>19.411</b>	<b>490</b>	<b>2,5</b>	<b>1.680</b>	<b>21</b>	<b>1,3</b>	<b>21.091</b>	<b>511</b>	<b>2,4</b>
Ubicación geográfica	2.396	67	2,8	396	1	0,3	2.792	68	2,4
Otros números	9.308	244	2,6	676	13	1,9	9.984	257	2,6
Fechas	7.707	179	2,3	608	7	1,2	8.315	186	2,2

(continúa)

<sup>17</sup> Refiere a las preguntas: de actividad económica, migración, pueblos indígenas y carrera.

**Cuadro 3.3.2 (conclusión)**

Tipo de pregunta	Cuestionario ampliado			Cuestionario básico			Total		
	Respuestas	Errores	%	Respuestas	Errores	%	Respuestas	Errores	%
Preguntas cerradas	46.003	851	1,8	3.092	19	0,6	49.095	870	1,8
Categorías	36.163	551	1,5	3.092	19	0,6	39.255	570	1,5
Grilla	9.840	300	3,0	///	///	///	///	///	///

**Nota:** ver esquema 3.3.1.

**Fuente:** INDEC. Base de datos sin estructura jerárquica y sin consistir. Cuestionarios básico y ampliado. II PP 2019.

A modo de orientación, el esquema 3.3.1 detalla qué preguntas se incluyen en cada categoría del tipo de pregunta. Las preguntas se presentan según estén incluidas en el cuestionario de Vivienda, de Hogar o de Población a partir de transcribir las letras V, H y P, respectivamente. A continuación, se indica el número de pregunta según su inclusión en el cuestionario A o B.

**Esquema 3.3.1 Preguntas de los cuestionarios ampliado y básico según tipología de pregunta. II PP 2019**

Tipología	Ampliado	Básico
Pregunta abierta-letras-Actividad económica	P29; P30; P31	-
Pregunta abierta-letras-Carrera	P10	-
Pregunta abierta- letras-UG	hdepto; edificio; entrada; n_calle; hnpuerta; casa_lote	
Pregunta abierta-letras-Grilla de personas	nompers; V8_a1; V8_b1; V8_c1; V8_a2; V8_b2; V8_c2; V8_a3; V8_b3; V8_c3; V8_a4; V8_b4; V8_c4; V8_a5; V8_b5; V8_c5; V8_a6; V8_b6; V8_c6; V8_a7; V8_b7; V8_c7;	
Pregunta abierta- letras -indígenas	P17_2	P10_2
Pregunta abierta- letras -Migración	P11_2; P11_3; P12_2; P12_3; P13_2; P13_3	P9_2; P9_3
Pregunta abierta-números	hpiso; nviv; V6; V7; H18; H23; H24_1; H28_1; H28_3; H28_4; H28_5; nper; P4; P9_1; P32_3 P34_2; P34_3	V6; H18; nper; hpiso; nviv; V7; P4; P8_1
Pregunta abierta-números-Fechas	P11_4; H28_2; P4_1; P4_2; P4_3; P32_3; P35; P36_1; P36_2	P9_4; P4_1; P4_2; P4_3
Pregunta abierta- números-UG	prov; dpto; loc; frac; radio; seg; mza; lmza; ccalle	
Pregunta cerrada-Categorías	H24_2; V1; V2; V3; V4; V5; H09; H10; H11; H12; H13; H14; H15; H16; H17; H19; H20; H21; H22; H25; respuesta; P1; P2; P3; P5; P6; P7; P8; P9_2; P11_1; P12_1; P13_1; P15; P16; P17_1; P18_2; P19; P20; P21; P22; P23; P24; P25; P26; P27; P28; P32_1; P33; P34_1	V1; V2; V3; V4; V5; H09; H10; H11; H13; H14; H15; H16; H19; respuesta; P2; P3; P5; P6; P7; P8_2; P9_1; P10_1; P12; P13; P14; P15; P16; H12; H17; P1; P11_2
Pregunta cerrada-Grilla	H26_1; H26_2; H26_3; H26_4; H26_5; H26_6; H26_7; H27_1; H27_2; H27_3; H27_4; H27_5; H27_6; P14_1; P14_2; P14_3; P14_4; P14_5; P14_6	-

**Fuente:** INDEC. Cuestionarios Básico y Ampliado. II PP 2019.

### 3.4

## Tipos de error de la interpretación y sus causas

Como se mencionó antes, con el objetivo de analizar los tipos de error y sus causas, se seleccionaron los datos correspondientes a las dos primeras personas censadas en cada cuestionario y se analizaron algunas preguntas específicas: Número de DNI, Edad, Provincia de nacimiento, Sexo y Tarea que realiza. Se distinguieron 5 tipos de errores según posibles causas:

- 1. Base No-Imagen Sí:** inconsistencia entre información en la imagen del cuestionario y en la base de datos.
- 2. Caligrafía:** confusa caligrafía de quien realizó el censo.
- 3. Fuera de *check box*:** problemas de ubicación de la marca, letra o número en el *check box* (parte del trazo fuera del *check box*).
- 4. Intensidad:** débil intensidad del grafito en la marca, letra o número.
- 5. Interpretación:** interpretación del software (luego de descartadas las anteriores causas).

Se revisaron 2.675 respuestas en cuestionarios de 549 personas. Del total de personas, el 46,4% tuvo al menos un error en alguna de las cinco preguntas y el 14,0%, en al menos dos (cuadros 3.4.1 y 3.4.2).

**Cuadro 3.4.1 Frecuencia de personas según cantidad de errores en las respuestas. II PP 2019**

Cantidad de errores	Frecuencia	%	Frecuencia acumulada	%
5	1	0,2	1	0,2
4	2	0,4	3	0,5
3	17	3,1	20	3,6
2	57	10,4	77	14,0
1	178	32,4	255	46,4
0	294	53,6	549	100,0
<b>Total</b>	<b>549</b>	<b>100,0</b>		

**Fuente:** INDEC. Base sin estructura jerárquica y sin consistir. II PP 2019.

En el cuadro 3.4.2 se observa que, para la mayoría de las preguntas, el error más frecuente es el de interpretación de las letras o los números (7,4%); seguido por aquellos que tienen que ver con la caligrafía de quien realizó el censo (3,1%) y la intensidad del grafito (1,9%, relacionado con el tipo de lápiz utilizado y la calibración del escaneo). Un ejemplo del contraste observado en intensidad puede verse en la figura 1.

Es importante mencionar que, si bien el error que proviene de datos que no están en la base de datos, pero sí en la imagen del cuestionario, es muy bajo, este se concentra en la pregunta Sexo, lo cual lo convierte en un error de alto impacto para el procesamiento de la información demográfica.

Al igual que en el análisis de la cantidad de errores según tipo de pregunta, en el cuadro 3.4.2 también se observa que las preguntas de respuesta con formato alfanumérico son las que presentan mayor porcentaje de error (Tarea que realiza: 27,6%, Provincia de nacimiento: 22,8%) y también los errores de interpretación son mayores (Tarea que realiza: 11,7%, Provincia de nacimiento: 13,5%).

**Cuadro 3.4.2 Personas. Datos interpretados correctamente y tipo de error en la lectura, según pregunta. II PP 2019**

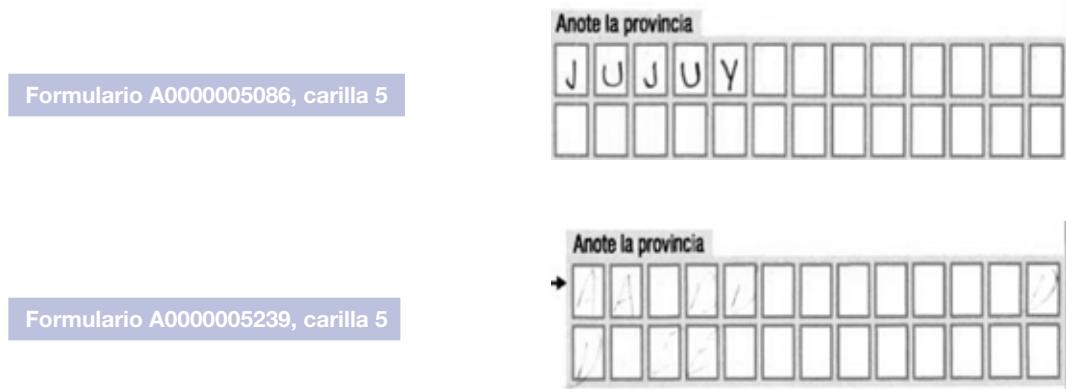
Pregunta	Correcto	Tipo de error						Total de personas
		Base No-Imagen Sí	Caligrafía	Fuera Box	Intensidad	Interpretación	Total	
<b>DNI</b>	483 (88,0%)	0 (0,0%)	7 (1,3%)	0 (0,0%)	9 (1,6%)	50 (9,1%)	66 (12,0%)	549 (100%)
<b>Edad</b>	525 (95,6%)	1 (0,2%)	4 (0,7%)	0 (0,0%)	2 (0,4%)	17 (3,1%)	24 (4,4%)	549 (100%)
<b>Provincia de nacimiento</b>	424 (77,2%)	424 (77,2%)	41 (7,5%)	7 (1,3%)	3 (0,4%)	74 (13,5%)	125 (22,8%)	549 (100%)
<b>Sexo</b>	540 (98,4%)	9 (1,6%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	9 (1,6%)	549 (100%)
<b>Tarea que realiza (1)</b>	347 (72,4%)	0 (0,0%)	32 (6,7%)	6 (1,2%)	38 (7,9%)	56 (11,7%)	132 (27,6%)	479 (100%)
<b>Total de respuestas</b>	2.319 (86,7%)	10 (0,2%)	84 (3,1%)	13 (0,5%)	52 (1,9%)	197 (7,4%)	356 (13,3%)	2.675 (100%)

(1) La pregunta Tarea que realiza totaliza menos personas por estar incluida solo en el cuestionario A.

**Nota:** en las filas se muestra la distribución del acierto/error en el total de personas.

**Fuente:** INDEC. Base de datos sin estructura jerárquica y sin consistir. II PP 2019.

**Figura 1. Ejemplos de respuesta para Provincia de nacimiento con grafitos de diferente intensidad. II PP 2019**



**Fuente:** INDEC. II PP 2019.

En el cuadro 3.4.3 se observa el porcentaje de error del Número de DNI según la posición del dígito. De los 169 dígitos clasificados como erróneos, los mayores errores se produjeron en las posiciones finales.

**Cuadro 3.4.3 Porcentaje de error en el número de DNI según posición del número. II PP 2019**

Posición	Frecuencia	Porcentaje
1	15	9,0
2	16	10,0
3	19	11,0
4	19	11,0
5	17	10,0
6	20	12,0
7	34	20,0
8	29	17,0
<b>Total</b>	<b>169</b>	<b>100,0</b>

**Fuente:** INDEC. Base de datos sin estructura jerárquica y sin consistir. II PP 2019.

En la matriz de errores del cuadro 3.4.4, se destacan el reemplazo del 8 por el 9, el 5 por el 9, el 7 por el 2 y el 7 por el 1 (celdas con predominancia roja). El número peor interpretado fue el 8 (27 errores, en totales de columna).

**Cuadro 3.4.4 Matriz de errores según carácter de número de DNI. II PP 2019**

Lectura	Carácter	Observación																		Total
		" "	0	1	2	3	4	5	6	7	8	9	b	e	F	g	I	I		
	" "	0	1	0	2	2	2	1	1	1	0	1	0	0	0	0	1	0	12	
	0	0	0	0	1	0	1	4	1	0	6	0	1	0	0	0	0	0	14	
	1	0	1	0	2	3	7	0	2	8	1	4	0	0	0	1	0	1	30	
	2	0	0	0	0	5	2	1	1	9	7	1	0	0	0	0	1	1	28	
	3	0	1	1	1	0	1	2	0	2	6	1	1	0	0	0	0	0	16	
	4	0	0	4	1	0	0	0	1	0	0	2	0	0	0	0	0	0	8	
	5	0	1	0	1	0	0	0	2	0	1	0	0	0	1	0	0	0	6	
	6	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	6	
	7	5	2	3	3	0	0	2	1	0	2	3	0	1	0	0	0	0	22	
	8	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
	9	0	0	0	1	0	2	9	0	0	10	0	0	0	0	0	0	0	22	
	A	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	N	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	
	<b>Total</b>	6	8	9	14	12	15	18	12	21	27	18	1	2	1	1	2	2	<b>169</b>	

**Nota:** las filas corresponden a lo registrado en la base de datos y las columnas a lo observado en la imagen.

**Fuente:** INDEC. Base de datos sin estructura jerárquica y sin consistir. II PP 2019.

Con el objetivo de profundizar el análisis en una pregunta con texto, se tomó la correspondiente a Provincia de nacimiento y se analizaron 205 letras erróneas, siendo los principales problemas el reemplazo de la J por la U y de la U por la V. La mayoría de estos errores se produjeron en la palabra “Jujuy” (65% de las respuestas donde hubo algún error).

**Cuadro 3.4.5 Matriz de errores según carácter. Provincia de nacimiento. II PP 2019**

		Observación																					
		Carácter	Carácter	'''	B	E	G	I	L	U	J	Y	A	T	S	V	H	N	R	D	O	Z	Total
Lectura	Carácter	A	1	0	1	1	1	2	0	1	0	0	0	0	0	1	0	2	0	0	0	10	
	N	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	
	O	0	1	0	2	1	0	10	8	0	1	0	0	0	0	0	0	0	0	3	0	26	
	S	0	0	2	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	9	
	D	0	0	0	0	0	0	0	0	12	0	1	0	0	0	0	0	0	0	0	1	14	
	L	1	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	1	0	0	0	6	
	V	1	0	0	0	0	0	0	19	4	2	0	0	0	0	0	0	0	0	0	0	26	
	X	0	0	0	0	0	0	0	2	3	0	1	0	0	0	0	0	0	0	0	0	6	
	B	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	2	
	U	0	0	0	0	0	1	0	31	0	0	0	0	0	0	3	0	0	0	1	0	36	
	I	3	0	1	0	0	0	5	10	6	0	0	0	0	0	0	0	0	0	0	0	25	
	Y	3	0	0	0	1	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	7	
	Z	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2	0	0	0	4	
	P	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	2	
	E	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	
	F	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	
	J	3	0	0	0	0	0	0	5	0	3	0	0	0	0	1	0	0	0	0	0	12	
	M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	2	
	H	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	
	R	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	
	T	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	
	C	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	4	
	K	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	G	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
Total		16	2	6	4	4	6	44	80	14	5	2	1	3	3	2	7	4	1	1	205		

Fuente: INDEC. Base de datos sin estructura jerárquica y sin consistir. II PP 2019.

### 3.5

## Conclusiones

- Del total de los registros de vivienda, el 68,1% tenía al menos un error proveniente de la lectura de los datos; y en los registros de población, el 64,6% tenía al menos un error de este tipo.
- Los niveles de error de interpretación muestran importantes diferencias en relación con el tipo de pregunta: las preguntas abiertas (numéricas y alfanuméricas) muestran un nivel de error mayor que las preguntas cerradas en especial, las que contienen letras.
- En la mayoría de las preguntas, los errores de interpretación son superiores en el cuestionario A que en el cuestionario B.
- En cuanto a los tipos de error, el más frecuente es el de interpretación de letras y números. En orden de importancia le siguen los relacionados con la caligrafía de quienes realizaron el censo y la intensidad del grafito (vinculado al tipo de lápiz utilizado, la intensidad al escribir y la calibración del escaneo).
- Del análisis de la pregunta de Número de DNI se desprende que los mayores errores se produjeron en las posiciones finales y que la combinación de números con mayor error fue el 8 por el 9, el 5 por el 9, el 7 por el 2 y el 7 por el 1.
- Para el caso de la pregunta sobre Provincia de nacimiento, los principales problemas fueron el reemplazo de la J por la U y de la U por la V.

## Errores de calidad

En el presente capítulo se analizan e interpretan los errores de contenido correspondientes a las preguntas de acuerdo con el error total, por área de prueba, por tipo de cuestionario (B o A) y el error de secuencia, y se proponen mejoras en vistas a la realización del Censo Experimental.

Es relevante contar con un análisis por tipo de cuestionario, dado que en la II PP 2019 se empleó una metodología de censo con muestra y se aplicó un cuestionario B, de 35 preguntas, y uno A, de 64 preguntas, en Mar del Plata, San Carlos de Bariloche y la Ciudad Autónoma de Buenos Aires (CABA). En Humahuaca, no se realizó muestra y se empleó el cuestionario A en todos los hogares. Entre los objetivos de prueba figuró la evaluación de una carga de trabajo diferencial para censistas a quienes se les había asignado uno u otro cuestionario. El tamaño de los segmentos a censar con cuestionario B en las localidades con muestra fue de 40 viviendas en promedio, con un límite inferior de 35 viviendas y uno superior de 45. Por su parte, los segmentos en los que se aplicó el cuestionario A se definieron tratando de incluir 20 viviendas (tamaño deseado) con límites de tolerancia de entre 15 y 25 viviendas. En los casos de villas y asentamientos se previó que los segmentos censales se redujeran a 10 viviendas, en función del mayor tamaño promedio de los hogares que allí habitan. En INDEC (2019: p. 15) se presentó la evaluación cualitativa del tamaño promedio de la carga de trabajo de los censistas y su impacto en los resultados de la cobertura censal alcanzados. En esta segunda parte del informe de evaluación, se analiza en qué medida los tamaños diferenciales de la carga de trabajo de censistas podrían impactar en la calidad de los datos relevados.

Cabe mencionar que las preguntas incluidas en el cuestionario B priorizan y aseguran:

- la mejora de la estratificación de niveles socioeconómicos a nivel de radio censal para la expansión de los datos del censo y para el diseño de la Muestra Maestra Urbana de Viviendas de la República Argentina (MMUVRA),
- la elaboración de las proyecciones y estimaciones de población 2020-2050,
- el cálculo de los indicadores componentes del índice de Necesidades Básicas Insatisfechas (NBI) (INDEC, 1985) y del Índice de Privación Material de los Hogares (IPMH) (INDEC, 2004),
- la medición de fenómenos de escasa relevancia cuantitativa a ser estudiados con posterioridad mediante la implementación de encuestas complementarias al censo,
- la medición de los indicadores concebidos en el marco del concepto de bienestar de la OCDE (2013) y de los correspondientes a los Objetivos de Desarrollo Sostenible –ODS– (2019),
- el aporte a las cuentas nacionales a partir de variables que contribuyan a la valorización del precio de la vivienda particular,
- la articulación de los datos censales con aquellos que provienen de registros administrativos.

Las preguntas incluidas en el cuestionario A se elaboraron tomando en cuenta los acuerdos de compatibilidad entre distintas fuentes de datos producidos por el INDEC, la comparabilidad con censos anteriores y los requerimientos de información de otros organismos del Estado.

Las preguntas del cuestionario se presentarán en los títulos de acuerdo con su unidad de análisis y su número de pregunta: las variables del cuestionario de vivienda contarán con una “V” delante de su número de pregunta; las variables del cuestionario de características habitacionales del hogar y la vivienda se mostrarán con una “H” delante de su número de pregunta y en las variables del cuestionario de población habrá una “P” delante de su número de pregunta.

Por último, el nivel del error se presenta comparado con los de los últimos tres censos (1991, 2001 y 2010) a fin de tener un parámetro. De todas maneras, la comparación debe realizarse con precaución, dado que la operacionalización de las preguntas, y en algunos casos los universos de análisis, pueden ser diferentes; y porque los resultados de los operativos censales que se presentan corresponden a un promedio ponderado de los diversos niveles de error por jurisdicción, los cuales cuentan con mejores condiciones de difusión y capacitación que las pruebas piloto, motivo por el cual es previsible que estas últimas muestren errores más altos.

## 4.1 Tipo de vivienda, Condición de ocupación y Detección de hogares (V1 a V7)

### 4.1.1 Diseño de preguntas en el cuestionario

#### Cuestionarios ampliado y básico

VIVIENDA																	
<p><b>1</b> Tipo de vivienda particular Anote por observación</p> <table style="margin-left: 20px;"> <tr><td>Casa</td><td><input type="checkbox"/> 1</td></tr> <tr><td>Rancho</td><td><input type="checkbox"/> 2</td></tr> <tr><td>Casilla</td><td><input type="checkbox"/> 3</td></tr> <tr><td>Departamento</td><td><input type="checkbox"/> 4</td></tr> <tr><td>Pieza en inquilinato, hotel familiar o pensión</td><td><input type="checkbox"/> 5</td></tr> <tr><td>Local no construido para habitación</td><td><input type="checkbox"/> 6</td></tr> <tr><td>Vivienda móvil</td><td><input type="checkbox"/> 7</td></tr> <tr><td>Persona(s) viviendo en la calle</td><td><input type="checkbox"/> 8</td></tr> </table> <p style="text-align: right;">Pase a la LISTA DE PERSONAS <b>8</b></p> <p><b>2</b> Hay personas presentes ¿Es la vivienda... principal? <input type="checkbox"/> 1 → <b>4</b> que se usa para vacaciones, fin de semana u otro uso temporal? <input type="checkbox"/> 2</p> <p><b>3</b> No hay personas presentes ¿Es una vivienda que... está en alquiler o venta? <input type="checkbox"/> 1 está en construcción? <input type="checkbox"/> 2 se usa como comercio, oficina o consultorio? <input type="checkbox"/> 3 se usa para vacaciones, fin de semana u otro uso temporal? <input type="checkbox"/> 4 está deshabitada por otra razón? <input type="checkbox"/> 5 No hay personas presentes, pero habitualmente viven personas <input type="checkbox"/> 6 → Deje aviso de visita</p>	Casa	<input type="checkbox"/> 1	Rancho	<input type="checkbox"/> 2	Casilla	<input type="checkbox"/> 3	Departamento	<input type="checkbox"/> 4	Pieza en inquilinato, hotel familiar o pensión	<input type="checkbox"/> 5	Local no construido para habitación	<input type="checkbox"/> 6	Vivienda móvil	<input type="checkbox"/> 7	Persona(s) viviendo en la calle	<input type="checkbox"/> 8	<p><b>4</b> ¿Cuántas personas pasaron la noche del sábado al domingo en esta vivienda, incluyendo a los bebés, niñas(os) pequeñas(os) y personas mayores? Una sola <input type="checkbox"/> 1 → <b>7</b> Dos o más <input type="checkbox"/> 2</p> <p><b>5</b> Esas personas ¿comparten los gastos para la comida? Sí <input type="checkbox"/> 1 → <b>7</b> No <input type="checkbox"/> 2</p> <p><b>6</b> ¿Cuántos grupos de personas mantienen gastos separados para la comida? Un hogar es la persona sola o grupo de personas que comparten los gastos para la comida y viven bajo el mismo techo. Cantidad de hogares: <input type="checkbox"/> <input type="checkbox"/></p> <p><b>7</b> Anote el número de hogar - Para el primer hogar anote <b>01</b>. - Para el segundo hogar y sucesivos donde haya personas presentes, abra un nuevo cuestionario, transcriba los mismos datos de la UBICACIÓN GEGRÁFICA y complete la pregunta <b>7</b> con el número <b>02, 03</b> (según corresponda). - Para el segundo hogar y sucesivos donde no hay personas presentes, pero habitualmente vive personas, abra un nuevo cuestionario, transcriba los mismos datos de la UBICACIÓN GEGRÁFICA, complete la pregunta <b>7</b> con el número <b>00</b> y deje aviso de visita. Hogar N°: <input type="checkbox"/> <input type="checkbox"/></p>
Casa	<input type="checkbox"/> 1																
Rancho	<input type="checkbox"/> 2																
Casilla	<input type="checkbox"/> 3																
Departamento	<input type="checkbox"/> 4																
Pieza en inquilinato, hotel familiar o pensión	<input type="checkbox"/> 5																
Local no construido para habitación	<input type="checkbox"/> 6																
Vivienda móvil	<input type="checkbox"/> 7																
Persona(s) viviendo en la calle	<input type="checkbox"/> 8																

### 4.1.2 Objetivo y principales características de las preguntas

- El objetivo de la pregunta sobre Tipo de vivienda es tener un insumo para caracterizar socioeconómicamente a la población, por ejemplo, mediante la construcción del índice de NBI.
- Esta pregunta brinda el *stock* de viviendas particulares requerido como dato por el Sistema de Cuentas Nacionales.<sup>18</sup>
- Las preguntas sobre el uso y la condición de ocupación de la vivienda permiten determinar la oferta del parque habitacional. Ambas se incorporaron para conocer si la unidad de relevamiento vivienda se encontraba ocupada o desocupada durante el operativo censal.

<sup>18</sup> El censo es una de las fuentes principales para las estimaciones de la cuenta de generación del ingreso, como así también para la estimación del stock de viviendas y sus características dentro del Sistema de Cuentas Nacionales. Para más detalles ver:

[https://www.indec.gob.ar/ftp/cuadros/economia/metodologia\\_21\\_cuentas\\_nacionales.pdf](https://www.indec.gob.ar/ftp/cuadros/economia/metodologia_21_cuentas_nacionales.pdf).

- La identificación de los hogares al interior de cada vivienda ocupada permite conocer cómo se organizan para vivir las personas que conviven bajo un mismo techo y así registrar otra de las unidades de relevamiento del censo, el hogar. Es un insumo fundamental para estudiar las condiciones habitacionales de estos, y las condiciones socioeconómicas de la población, entre ellas, la construcción de los indicadores de NBI y de los componentes del IPMH, como así también para conocer el déficit habitacional.
- Estas preguntas se encuentran presentes en ambos cuestionarios de la II PP.

#### 4.1.3

### Resultados

- La pregunta sobre Tipo de vivienda presentó un error considerablemente superior al de los censos anteriores.
- La indagación respecto al uso de la vivienda no fue respondida por la mitad de los hogares ocupados y la Causa de deshabitación de la vivienda no fue completada para un cuarto de las viviendas en las que no había personas presentes al momento del censo.
- El resto de las preguntas presentó un elevado nivel de error.

**Cuadro 4.1.3.1 Variables de vivienda (V1, V2, V3, V4, V5, V6 y V7). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto al Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Tipo de vivienda	0,8	1,0	3,4	8,5	Se agrupan categorías.
Uso de la vivienda	No se preguntó.			49,9	La pregunta se dividió en dos partes. La primera indaga sobre el tipo de uso de la vivienda (residencia primaria o secundaria) para las viviendas con personas presentes. La segunda busca establecer los motivos por los que la vivienda se encuentra deshabitada.
Causa por la que está deshabitada	4,5	2,2	3,8	25,2	
Cuántas personas pasaron la noche	No se preguntó.			13,5	
Gastos de alimentación				6,4	
Cantidad de hogares	0,4	No se preguntó.	1,5	20,1	
Número de hogar	No se preguntó.		1,5	17,2	

**Fuente:** Giusti (1997), archivos sin consistir de los censos 2001 y 2010 (INDEC, s.f.) y base de datos sin consistir de la II PP 2019.

- La pregunta sobre Tipo de vivienda (que debe ser completada por observación antes de ingresar a la vivienda) y la que indaga sobre si se comparten los gastos de alimentación son las que registran el menor nivel de error en todas las áreas.

- La pregunta sobre Uso de la vivienda presenta el valor más alto y homogéneo de error en todas las áreas.
- El resto de las preguntas que integran la identificación de hogares presentan un error promedio de entre 13% y 20%, con importantes variaciones entre las áreas.

**Cuadro 4.1.3.2 Variables de vivienda (V1 a , V2, V3, V4, V5, V6 y V7). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Tipo de vivienda	8,5	4,8	9,5	6,4	6,3
Uso de la vivienda	49,9	55,3	47,6	63,2	40,7
Causa por la que está deshabitada	25,2	48,0	17,5	35,7	41,0
Cuántas personas pasaron la noche	13,5	6,7	12,5	22,0	15,7
Gastos de alimentación	6,4	3,7	6,7	9,1	3,6
Cantidad de hogares	20,1	4,8	24,7	14,0	15,2
Número de hogar	17,2	13,4	15,7	28,7	17,8

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En las áreas donde se aplicaron los dos cuestionarios (excluye Humahuaca, donde se aplicó únicamente cuestionario A), se observa que las preguntas correspondientes a las viviendas habitadas tuvieron menor nivel de error en el cuestionario A que en el B, mientras que la que corresponde a viviendas deshabitadas muestra mayor error en el cuestionario A.

**Cuadro 4.1.3.3 Variables de vivienda (V1, V2, V3, V4, V5, V6 y V7). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Tipo de vivienda	6,5	4,9
Uso de la vivienda	61,6	40,6
Causa por la que está deshabitada	37,7	46,4
Cuántas personas pasaron la noche	20,8	9,5
Gastos de alimentación	6,0	5,1
Cantidad de hogares	14,5	8,3
Número de hogar	25,2	15,4

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Del cuadro 4.1.3.4 se deduce que las preguntas que presentan mayores niveles de error de secuencia son las relativas a la identificación de los hogares (preguntas V5, V6 y V7). Cerca de un 40% de las viviendas unipersonales o de aquellas en las que vivían más de dos personas compartiendo gastos de alimentación respondieron la pregunta siguiente cuando no correspondía hacerlo.
- No se observa un comportamiento sistemático por área de prueba.

**Cuadro 4.1.3.4 Variables de vivienda (V1, V2, V3, V4, V5, V6 y V7). Error de secuencia. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Tipo de vivienda	///				
Uso de la vivienda	10,0	5,2	9,4	10,6	17,6
Causa por la que está deshabitada	11,6	15,3	10,2	17,9	11,8
Cuántas personas pasaron la noche	5,0	15,2	4,7	3,2	4,2
Gastos de alimentación	38,7	33,3	41,5	36,2	27,0
Cantidad de hogares	36,0	63,0	32,0	50,9	34,1
Número de hogar	8,5	26,8	7,6	6,9	5,6

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.2 Materiales de los pisos, las paredes y los techos (A: H9 a H13) (B: H9 a H11)

### 4.2.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	Cuestionario básico
<p><b>9</b> ¿Cuál es el material predominante de los pisos?</p> <p>Cerámica, mosaico, baldosa, mármol, alfombra, madera, vinílico, microcemento o cemento alisado <input type="checkbox"/> 1 Carpeta, contrapiso o ladrillo fijo <input type="checkbox"/> 2 Tierra o ladrillo suelto <input type="checkbox"/> 3 Otro <input type="checkbox"/> 4</p>	<p><b>9</b> ¿Cuál es el material predominante de los pisos?</p> <p>Cerámica, mosaico, baldosa, mármol, alfombra, madera, vinílico, microcemento o cemento alisado <input type="checkbox"/> 1 Carpeta, contrapiso o ladrillo fijo <input type="checkbox"/> 2 Tierra o ladrillo suelto <input type="checkbox"/> 3 Otro <input type="checkbox"/> 4</p>
<p><b>10</b> ¿Cuál es el material predominante de las paredes exteriores?</p> <p>Ladrillo, piedra, bloque, hormigón o panel premoldeado <input type="checkbox"/> 1 Adobe <input type="checkbox"/> 2 Madera <input type="checkbox"/> 3 Chapa de metal o fibrocemento <input type="checkbox"/> 4 Chorizo, cartón, palma, paja sola o material de desecho <input type="checkbox"/> 5 Otro <input type="checkbox"/> 6</p>	<p><b>10</b> ¿Cuál es el material predominante de la cubierta exterior del techo?</p> <p>Baldosa, membrana, pintura asfáltica, pizarra o teja <input type="checkbox"/> 1 Losa o carpeta a la vista (sin cubierta) <input type="checkbox"/> 2 Chapa de metal <input type="checkbox"/> 3 Chapa de cartón, caña, palma, tabla con barro, paja con barro o paja sola <input type="checkbox"/> 4 Otro <input type="checkbox"/> 5</p>
<p><b>11</b> Las paredes exteriores ¿tienen revestimiento externo o revoque?</p> <p>Incluye terminación "ladrillo a la vista"</p> <p>Sí <input type="checkbox"/> 1 No <input type="checkbox"/> 2</p>	<p><b>11</b> El techo ¿tiene revestimiento interior o cielorraso?</p> <p>Sí <input type="checkbox"/> 1 No <input type="checkbox"/> 2</p>
<p><b>12</b> ¿Cuál es el material predominante de la cubierta exterior del techo?</p> <p>Si la vivienda forma parte de un edificio de departamentos, se considera el techo del último piso del edificio</p> <p>Baldosa, membrana, pintura asfáltica, pizarra o teja <input type="checkbox"/> 1 Losa o carpeta a la vista (sin cubierta) <input type="checkbox"/> 2 Chapa de metal <input type="checkbox"/> 3 Chapa de cartón, caña, palma, tabla con barro, paja con barro o paja sola <input type="checkbox"/> 4 Otro <input type="checkbox"/> 5</p>	<p><b>10</b> ¿Cuál es el material predominante de la cubierta exterior del techo?</p> <p>Si la vivienda forma parte de un edificio de departamentos, se considera el techo del último piso del edificio</p> <p>Baldosa, membrana, pintura asfáltica, pizarra o teja <input type="checkbox"/> 1 Losa o carpeta a la vista (sin cubierta) <input type="checkbox"/> 2 Chapa de metal <input type="checkbox"/> 3 Chapa de cartón, caña, palma, tabla con barro, paja con barro o paja sola <input type="checkbox"/> 4 Otro <input type="checkbox"/> 5</p>

### 4.2.2 Objetivo y principales características de las preguntas

- El objetivo de estas preguntas es conocer las condiciones de habitabilidad, calidad, aislamiento y durabilidad de las viviendas en las que habitan los hogares y las personas y disponer de un insumo para caracterizar socioeconómicamente a la población a partir, por ejemplo, de la construcción del IPMH.
- Las preguntas sobre Material predominante de las paredes y Revestimiento exterior de las paredes solamente se formulan en el cuestionario A.

### 4.2.3 Resultados

- Las preguntas sobre materiales constructivos de los pisos, las paredes y los techos presentan errores superiores a los de los censos anteriores, en especial las relativas al material predominante de las paredes exteriores, la cubierta exterior del techo y el cielorraso o revestimiento interior (cuadro 4.2.3.1).

**Cuadro 4.2.3.1 Materiales de piso, paredes y techo (A: H9 a H13) (B: H9 a H11). Error total. Censos 1991, 2001 y 2010 y II PP 2019, y modificaciones respecto al Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Material predominante de los pisos	1,2	1,2	1,3	4,7	Se incluyen nuevos materiales.
Material predominante de las paredes exteriores	1,2	1,4	1,3	7,0	Se incluyen nuevos materiales.
Revestimiento externo o revoque	No se preguntó.	0,6	1,7	3,5	La inscripción “incluye terminación ladrillo a la vista” deja de formar parte de la pregunta y se presenta como aclaración.
Material predominante de la cubierta exterior del techo	1,5	2,3	2,3	6,4	Se cambia el orden, se modifican y agrupan categorías.
Revestimiento interior o cielorraso	No se preguntó.	1,3	1,5	5,8	La pregunta no presenta modificaciones.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, 2010 y base de datos sin consistir de la II PP 2019.

- Humahuaca es la jurisdicción que concentra el mayor nivel de error en las preguntas referidas a los materiales de construcción de la vivienda, excepto en la pregunta sobre Revestimiento externo o revoque (cuadro 4.2.3.2).

**Cuadro 4.2.3.2 Materiales de piso, paredes y techo (A: H9 a H13) (B: H9 a H11). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (')
Material predominante de los pisos	4,7	3,5	5,6	3,1	0,0
Material predominante de las paredes exteriores	7,0	4,2	7,5	3,6	1/25
Revestimiento externo o revoque	3,5	2,8	3,4	4,8	1/19
Material predominante de la cubierta exterior del techo	6,4	4,6	7,0	6,2	3,3
Revestimiento interior o cielorraso	5,8	2,3	6,9	4,9	0,5

(') El nivel de error de San Carlos de Bariloche para las variables Material predominante de las paredes exteriores y Revestimiento externo o revoque, que solo están incluidas en el cuestionario A, se presentan en absolutos, dado que solo se cuenta en el denominador con 25 y 19 hogares ocupados respectivamente.

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Para las preguntas relativas a los Materiales constructivos de la vivienda, el cuestionario A muestra errores superiores al B (cuadro 4.2.3.3).

**Cuadro 4.2.3.3 Materiales de piso, paredes y techo (A: H9 a H13) (B: H9 a H11). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Material predominante de los pisos	1,0	3,4
Material predominante de la cubierta exterior del techo	3,8	5,9
Revestimiento interior o cielorraso	1,6	4,1

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- La pregunta sobre Revestimiento externo o revoque (cuadro 4.2.3.4) no debía ser respondida por aquellos hogares que declararon tener como material de las paredes exteriores: “madera”, “chapa de metal o fibrocemento”, “chorizo, cartón, palma, paja sola o material de desecho” u “otros”. Dada la escasa cantidad de viviendas con este tipo de paredes exteriores, los errores se expresan en valores absolutos en forma de fracción. La mayoría de los censistas no siguieron correctamente el pase y realizaron la pregunta en viviendas que no correspondían (27 casos sobre 38).

**Cuadro 4.2.3.4 Revestimiento externo o revoque. Error de secuencia. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos**

Pregunta	Total	CABA (')	Humahuaca (')	Mar del Plata (')	San Carlos de Bariloche (')
Revestimiento externo o revoque	27/38	3/3	18/24	5/6	1/5

(') Se presentan en absolutos dado que en todas las áreas de prueba el denominador es inferior a 25.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.3.1

## Diseño de preguntas en el cuestionario

## Cuestionario ampliado

## Cuestionario básico

**14** ¿Tiene agua...

por cañería dentro de la vivienda?  1  
fuera de la vivienda, pero dentro del terreno?  2  
fuera del terreno?  3

**15** El agua que usa este hogar para beber y cocinar, ¿proviene de...

red pública (agua corriente)?  1  
perforación con bomba a motor?  2  
perforación con bomba manual?  3  
pozo sin bomba?  4  
Otro  5

**16** Este hogar ¿tiene baño o letrina...

dentro de la vivienda?  1  
fuera de la vivienda, pero dentro del terreno?  2  
No tiene  3 → **20**

**17** El baño ¿tiene...

Si tiene más de un baño, se considera el baño principal  
inodoro con botón, mochila o cadena (arrastre de agua)?  1  
inodoro sin botón o sin cadena (a balde)?  2  
pozo?  3

**18** El desagüe del inodoro ¿es...

a red pública (cloaca)?  1  
a cámara séptica y pozo ciego?  2  
solo a pozo ciego?  3  
a hoyo, excavación en la tierra, etc.?  4

**19** El baño o letrina ¿es...

usado solo por este hogar?  1  
compartido con otros hogares?  2

**20** Este hogar ¿tiene un lugar o cuarto para cocinar...

Incluye cocina integrada  
con instalación de agua y desagüe?  1  
con instalación de agua sin desagüe?  2  
sin instalación de agua?  3  
No tiene lugar o cuarto para cocinar  4

**21** Para cocinar, ¿utiliza principalmente...

electricidad?  1  
gas de red?  2  
gas en garrafa, en tubo o a granel (zeppelin)?  3  
leña o carbón?  4  
Otro  5

**22** Este hogar ¿tiene agua caliente a través de...

calefón o termotanque a gas?  1  
calefón o termotanque eléctrico?  2  
caldera central?  3  
panel solar?  4  
No tiene agua caliente  5

**12** ¿Tiene agua...

por cañería dentro de la vivienda?  1  
fuera de la vivienda, pero dentro del terreno?  2  
fuera del terreno?  3

**13** El agua que usa este hogar para beber y cocinar, ¿proviene de...

red pública (agua corriente)?  1  
perforación con bomba a motor?  2  
perforación con bomba manual?  3  
pozo sin bomba?  4  
Otro  5

**14** Este hogar ¿tiene baño o letrina...

dentro de la vivienda?  1  
fuera de la vivienda, pero dentro del terreno?  2  
No tiene  3 → **17**

**15** El baño ¿tiene...

Si tiene más de un baño, se considera el baño principal  
inodoro con botón, mochila o cadena (arrastre de agua)?  1  
inodoro sin botón o sin cadena (a balde)?  2  
pozo?  3

**16** El desagüe del inodoro ¿es...

a red pública (cloaca)?  1  
a cámara séptica y pozo ciego?  2  
solo a pozo ciego?  3  
a hoyo, excavación en la tierra, etc.?  4

**17** Para cocinar, ¿utiliza principalmente...

electricidad?  1  
gas de red?  2  
gas en garrafa, en tubo o a granel (zeppelin)?  3  
leña o carbón?  4  
Otro  5

### 4.3.2

### Objetivo y principales características de las preguntas

- El objetivo de estas preguntas es conocer las condiciones sanitarias de los hogares, el acceso y la disponibilidad de agua corriente para la alimentación, la higiene personal y la limpieza del hogar, el acceso a las condiciones básicas de salubridad y a combustibles seguros y al recurso de agua caliente.
- La Tenencia de baño se utiliza para la construcción del índice de NBI y junto con la variable Descarga del inodoro aporta al cálculo del IPMH.
- Las preguntas sobre Uso compartido del baño, Tenencia de lugar para cocinar y Disponibilidad de agua caliente se formulan solo en el cuestionario A.

### 4.3.3

### Resultados

- Las preguntas sobre condiciones sanitarias muestran errores superiores, en especial a los últimos dos censos, a excepción de la disponibilidad de inodoro y del sistema utilizado para la descarga del inodoro.
- La pregunta relativa a la Disponibilidad de agua caliente, que se incorpora en esta prueba piloto, muestra una calidad similar al resto de las preguntas, con 5,7% de error total.

**Cuadro 4.3.3.1 Condiciones sanitarias, combustible y agua caliente (A: H14 a H22) (B: H12 a H17). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto al Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Provisión del agua	1,7	2,5	2,1	6,2	La pregunta no presenta modificaciones.
Procedencia del agua	1,2	1,3	1,4	4,9	Se especifica la categoría “pozo sin bomba” y “red pública (agua corriente)”. Se eliminan categorías y se agrega la categoría “otro”.
Disponibilidad y ubicación de baño o letrina	No se preguntó.	1,3	1,3	5,2	Se incorpora la indagación sobre la ubicación del baño.
Disponibilidad de inodoro con botón, mochila o cadena	3,4	2,3	1,2	2,0	Se incorpora la indagación por el tipo de descarga de inodoro.
Tipo de desagüe del inodoro	2,4	0,6	1,2	2,5	La pregunta no presenta modificaciones.
Baño de uso compartido o de uso exclusivo del hogar (¹)	3,4	2,3	2,1	4,4	La pregunta no presenta modificaciones.
Disponibilidad de cuarto de cocina	No se preguntó.	2,4	2,6	6,4	Se incorpora la tenencia de desagüe.
Combustible utilizado principalmente para cocinar	3,3	2,3	1,9	6,5	Se agrupan el gas en garrafa, tubo y a granel y se pone en primer lugar electricidad.
Disponibilidad de agua caliente	No se preguntó.			5,7	Se incorpora la pregunta.

(¹) La pregunta sobre baño compartido o de uso exclusivo se midió en el Censo 1991 dentro de la pregunta sobre disponibilidad de inodoro, por lo tanto, se consigna el mismo error total en ambas preguntas.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la segunda prueba piloto del Censo de la ronda 2020.

- Humahuaca es la jurisdicción que concentra el mayor nivel de error en la gran parte de las preguntas sobre las condiciones sanitarias y combustible usado para cocinar, mientras que San Carlos de Bariloche y CABA son las áreas con menor nivel de error.

**Cuadro 4.3.3.2 Condiciones sanitarias, combustible y agua caliente (A: H14 a H22) (B: H12 a H17). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (*)
Provisión del agua	6,2	3,1	7,4	4,4	0,8
Procedencia del agua	4,9	3,9	5,9	2,3	0,3
Disponibilidad y ubicación de baño o letrina	5,2	1,9	6,4	3,1	0,3
Disponibilidad de inodoro con botón, mochila o cadena	2,0	1,2	2,1	3,2	0,3
Tipo de desagüe del inodoro	2,5	3,2	2,7	2,4	0,8
Baño de uso compartido o de uso exclusivo del hogar	4,4	2,7	4,6	4,2	0/25
Disponibilidad de cuarto de cocina	6,4	4,2	6,7	5,8	0/25
Combustible utilizado principalmente para cocinar	6,5	4,6	7,6	5,1	1,4
Disponibilidad de agua caliente	5,7	6,3	5,8	4,0	0,0

(\*) El nivel de error de San Carlos de Bariloche para las variables Baño de uso compartido o de uso exclusivo del hogar y Disponibilidad de cuarto de cocina y Disponibilidad de agua caliente, que solo están incluidas en el cuestionario A, se presentan en absolutos, dado que solo se cuenta en el denominador con 25 hogares ocupados.

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En las preguntas sobre condiciones sanitarias, combustible y disponibilidad de agua caliente, el cuestionario A muestra errores considerablemente superiores al cuestionario B.

**Cuadro 4.3.3.3 Condiciones sanitarias, combustible y agua caliente (A: H14 a H18 y H21) (B: H12 a H17). Error total. Cuestionarios básico y ampliado. Áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Provisión del agua	1,7	4,1
Procedencia del agua	1,4	2,7
Disponibilidad y ubicación de baño o letrina	0,9	3,0
Disponibilidad de inodoro con botón, mochila o cadena	1,1	2,4
Tipo de desagüe del inodoro	0,7	3,8
Combustible utilizado principalmente para cocinar	2,6	5,0

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Las condiciones sanitarias del baño no debían ser respondidas por aquellos hogares que habían manifestado no disponer de baño. Casi no se observan hogares en esta situación en CABA, Mar del Plata y San Carlos de Bariloche. En Humahuaca, por su parte, se registraron 71 hogares en dichas condiciones, de los cuales cerca del 20,0% presenta errores en los pases de preguntas. Es decir que los cuestionarios correspondientes a estos cuentan con marcas en las preguntas sobre las características del baño (cuadro 4.3.3.4).

**Cuadro 4.3.3.4 Condiciones sanitarias del hogar (A: H17 a H19) (B: H15 y H16). Error de secuencia. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA (%)	Humahuaca	Mar del Plata (%)	San Carlos de Bariloche (%)
Disponibilidad de inodoro con botón, mochila o cadena	18,7	0/2	19,7	0/1	0/1
Tipo de desagüe del inodoro	17,3	0/2	18,3	0/1	0/1
Baño de uso compartido o de uso exclusivo del hogar	20,3	0/2	21,1	0/1	0/0

(1) Se presentan en absolutos dado que, en CABA, Mar del Plata y San Carlos de Bariloche, el denominador es inferior a 25.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.4 Cantidad de habitaciones, Superficie cubierta y Régimen de tenencia de la vivienda (A: H23 a H25) (B: H18 a H19)

### 4.4.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	Cuestionario básico
<p><b>23</b> En total, ¿cuántos ambientes, habitaciones o piezas tiene este hogar (sin contar baños ni cocinas)?</p> <p>Cantidad total de ambientes, habitaciones o piezas: <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/></p>	<p><b>18</b> En total, ¿cuántos ambientes, habitaciones o piezas tiene este hogar (sin contar baños ni cocinas)?</p> <p>Cantidad total de ambientes, habitaciones o piezas: <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/></p>
<p><b>24</b> ¿Cuál es la superficie cubierta de esta vivienda?</p> <p>Superficie: <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> <input style="width: 40px; height: 20px; border: 1px solid #ccc; border-radius: 5px;" type="text"/> m<sup>2</sup></p> <p>Ignorado <input type="checkbox"/> 99999</p>	
<p><b>25</b> La vivienda ¿es...</p> <p>propia? <input type="checkbox"/> 1 alquilada? <input type="checkbox"/> 2 cedida por trabajo? <input type="checkbox"/> 3 prestada? <input type="checkbox"/> 4 Otra situación <input type="checkbox"/> 5</p>	<p><b>19</b> La vivienda ¿es...</p> <p>propia? <input type="checkbox"/> 1 alquilada? <input type="checkbox"/> 2 cedida por trabajo? <input type="checkbox"/> 3 prestada? <input type="checkbox"/> 4 Otra situación <input type="checkbox"/> 5</p>

#### 4.4.2

#### Objetivo y principales características de las preguntas

- La pregunta que indaga sobre Cantidad total de habitaciones (pregunta abierta) permite el cálculo del indicador de hacinamiento, es decir, la cantidad de personas por cuarto de habitación.
- La Superficie cubierta de la vivienda (pregunta abierta) proporciona información relevante para el cálculo de la valorización de la vivienda, de utilidad para las cuentas nacionales.
- El Régimen de tenencia de la vivienda se encuentra en estrecha relación con el indicador de déficit habitacional y, en consecuencia, con las condiciones socioeconómicas de los hogares.
- La pregunta sobre Superficie cubierta de la vivienda se formula solo en el cuestionario A.

#### 4.4.3

#### Resultados

- La pregunta sobre Cantidad total de habitaciones presenta mayor porcentaje de error que en los censos anteriores, mientras que la pregunta sobre Régimen de tenencia de la vivienda manifiesta un nivel de error superior al de los últimos dos censos, pero inferior al Censo 1991.
- La pregunta sobre Superficie cubierta de la vivienda es la única del bloque que se ha puesto a prueba en esta oportunidad y es la que presenta mayor error respecto de las dos variables anteriormente mencionadas<sup>19</sup> (cuadro 4.4.3.1).

**Cuadro 4.4.3.1 Habitaciones, régimen de tenencia y superficie (A: H23 a H25) (B: H18 a H19). Error total censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto al Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Cantidad total de habitaciones	1,8	1,0	1,9	4,6	Se agrega la palabra "ambientes" y se elimina la pregunta sobre cuartos para dormir.
Superficie cubierta de la vivienda	No se preguntó.			7,9	Se incorpora la pregunta.
Régimen de tenencia de la vivienda	6,1	1,2	1,3	5,4	Se modifica el orden de las categorías.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- Humahuaca es la jurisdicción con mayor nivel de error en estas preguntas, mientras que San Carlos de Bariloche presenta errores muy bajos (cuadro 4.4.3.2).

<sup>19</sup> Las preguntas abiertas tienen dos tipos de error, no respuesta e inválidos, mientras que las variables precodificadas pueden tener no respuesta y multmarca. En los casos en los que no se menciona la composición del error, casi la totalidad se debe a no respuestas.

**Cuadro 4.4.3.2 Habitaciones, régimen de tenencia y superficie (A: H23 a H25) (B: H18 a H19). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (')
Cantidad total de habitaciones	4,6	4,2	5,1	4,4	1,6
Superficie cubierta de la vivienda	7,9	7,3	8,1	6,7	0/25
Régimen de tenencia de la vivienda	5,4	3,5	6,6	3,1	0,3

(') El nivel de error de San Carlos de Bariloche para la variable Superficie cubierta de la vivienda, que solo está incluida en el cuestionario A, se presenta en absolutos, dado que solo se cuenta en el denominador con 25 hogares ocupados.

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Para ambas preguntas, el cuestionario A muestra errores superiores (cuadro 4.4.3.3).

**Cuadro 4.4.3.3 Habitaciones y régimen de tenencia (A: H23 y H25) (B: H18 y H19). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Cantidad total de habitaciones	2,1	4,8
Régimen de tenencia de la vivienda	0,9	3,9

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Se destaca que el error total de la pregunta sobre Cantidad total de habitaciones se compone en partes similares por error de no respuesta y datos inválidos (cuadro 4.4.3.4).

**Cuadro 4.4.3.4 Cantidad total de habitaciones. Error según tipo. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Tipo de error	Total	Áreas de prueba			
			CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Cantidad total de habitaciones	Inválido	2,1	2,3	2,2	2,3	1,1
	No respuesta	2,5	1,9	2,9	2,1	0,5
	Error total	4,6	4,2	5,1	4,4	1,6

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En el cuadro 4.4.3.5 se observa que el principal problema de la pregunta relativa a Superficie cubierta de la vivienda es el desconocimiento de la información. Más de la mitad declara ignorar la respuesta.

**Cuadro 4.4.3.5 Superficie cubierta de la vivienda. Distribución de frecuencias. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Pregunta	Tipo de error	Total	Áreas de prueba			
			CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Superficie cubierta	Valor válido diferente a ignorado	40,5	55,0	39,2	42,2	56,0
	Ignorado	51,6	37,7	52,7	51,1	44,0
	<b>Error total</b>	<b>7,9</b>	<b>7,3</b>	<b>8,1</b>	<b>6,7</b>	<b>0,0</b>

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.5 Climatización y Equipamiento (A: H26 a H28)

### 4.5.1 Diseño de preguntas en el cuestionario

**Cuestionario ampliado**

**26** Este hogar ¿tiene...

	Sí	No
aire acondicionado frío-calor?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
aire acondicionado solo frío?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
calefacción central?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
estufa a gas?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
estufa eléctrica?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
estufa a leña, carbón o querosén?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
panel solar?	<input type="checkbox"/> 1	<input type="checkbox"/> 2

**27** Este hogar ¿tiene...

	Sí	No
heladera con freezer o freezer solo?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
heladera sin freezer?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
computadora (PC, notebook, tablet, etc.)?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
internet en la vivienda?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
al menos un teléfono móvil con acceso a internet?	<input type="checkbox"/> 1	<input type="checkbox"/> 2
al menos un teléfono móvil sin acceso a internet?	<input type="checkbox"/> 1	<input type="checkbox"/> 2

**28** Este hogar ¿cuántas(os)...

autos, camionetas y/o utilitarios tiene?	<input type="checkbox"/> <input type="checkbox"/>	→	Si tiene uno o más de uno, anote el año del último
motocicletas (100 c.c. o más) tiene?	<input type="checkbox"/> <input type="checkbox"/>		
ciclomotores (menos de 100 c.c.) tiene?	<input type="checkbox"/> <input type="checkbox"/>		
bicicletas tiene?	<input type="checkbox"/> <input type="checkbox"/>		

Año:

## 4.5.2

## Objetivo y principales características de las preguntas

- La pregunta sobre Equipamiento que posee el hogar para la calefacción y la refrigeración de la vivienda permite conocer cuál es la fuente de energía que utilizan los hogares para calefaccionar y refrigerar el espacio en el que conviven. En este bloque de preguntas también se indaga sobre equipamiento para la refrigeración de los alimentos y sobre los sistemas de tecnología de la información y de las comunicaciones (TIC) que poseen los hogares.
- Estas preguntas, en especial la relativa a Heladera con o sin freezer, pueden ser utilizadas como un indicador *proxy* de la tenencia de electricidad.
- La pregunta sobre Tenencia de vehículos aportaría datos para el diseño de políticas públicas relativas al transporte.
- Todas estas preguntas se formulan solo en el cuestionario A.

## 4.5.3

## Resultados

- Las preguntas sobre climatización, equipamiento y TIC presentan errores que oscilan entre el 10% y el 20%, mientras que aquellas que indagan sobre vehículos superan el 33%.
- Las preguntas que pueden ser comparadas con censos anteriores superan ampliamente el error que estos presentan (cuadro 4.5.3.1).

**Cuadro 4.5.3.1 Climatización y equipamiento (A: H26 a H28). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto al Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010	
	1991	2001	2010			
Aire acondicionado frío-calor	No se preguntó.		15,2	Se incorpora la pregunta.		
Aire acondicionado solo frío	No se preguntó.		17,6	Se incorpora la pregunta.		
Calefacción central	No se preguntó.		18,5	Se incorpora la pregunta.		
Estufa a gas	No se preguntó.		16,5	Se incorpora la pregunta.		
Estufa eléctrica	No se preguntó.		15,5	Se incorpora la pregunta.		
Estufa a leña, carbón o querosén	No se preguntó.		17,4	Se incorpora la pregunta.		
Panel solar	No se preguntó.		17,3	Se incorpora la pregunta.		
Heladera con <i>freezer</i> / <i>freezer</i> solo (1)	No se preguntó.	3,1	1,5	10,1	Se separa heladera con <i>freezer</i> y sin <i>freezer</i> .	
Heladera sin <i>freezer</i>	No se preguntó.	4,2	1,5	19,9		
Computadora (PC, <i>notebook</i> , <i>tablet</i> , etc.)	No se preguntó.	4,0	2,5	11,4	Se agregan otros tipos de dispositivos entre paréntesis.	
Internet en la vivienda	No se preguntó.	4,0	No se preguntó	12,7	Se incorpora la pregunta.	
Teléfono móvil con acceso a internet (2)	No se preguntó.		2,4	10,2	Se separa teléfono móvil con y sin acceso a internet.	
Teléfono móvil sin acceso a internet	No se preguntó.		16,4			

(continúa)

**Cuadro 4.5.3.1 (conclusión)**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Cantidad de autos, camionetas y/o utilitarios en el hogar		No se preguntó.		29,9	Se incorpora la pregunta.
Año del último vehículo		No se preguntó.		33,1	Se incorpora la pregunta.
Cantidad de motocicletas		No se preguntó.		43,2	Se incorpora la pregunta.
Cantidad de ciclomotores		No se preguntó.		48,1	Se incorpora la pregunta.
Cantidad de bicicletas		No se preguntó.		41,2	Se incorpora la pregunta.

(<sup>1</sup>) El Censo 2010 pregunta por tenencia de heladera, sin diferenciar si posee o no freezer.

(<sup>2</sup>) En el Censo 2010 se pregunta por tenencia de teléfono celular.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- San Carlos de Bariloche exhibe los errores más bajos, aunque deben considerarse con reparo debido a la poca cantidad de hogares ocupados censados con el cuestionario A. Las preguntas referidas a la climatización del hogar presentan un nivel mayor de error en CABA y en Mar del Plata. El nivel de error en las preguntas relativas al equipamiento del hogar es más homogéneo en las todas áreas, mientras que el error de las preguntas relativas a la Cantidad de vehículos es inferior en Mar del Plata.
- Las preguntas sobre cantidad de vehículos tienen el error más elevado en Humahuaca (cuadro 4.5.3.2).

**Cuadro 4.5.3.2 Climatización y equipamiento del hogar (A: H26 a H28). Error total. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes y porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Aire acondicionado frío-calor	15,2	16,2	14,9	18,4	1/25
Aire acondicionado solo frío	17,6	19,4	17,3	21,1	0/25
Calefacción central	18,5	22,5	18,0	21,5	2/25
Estufa a gas	16,5	22,0	16,0	18,4	2/25
Estufa eléctrica	15,5	13,1	15,4	21,1	1/25
Estufa a leña, carbón o querosén	17,4	22,0	17,5	13,5	0/25
Panel solar	17,3	23,0	16,8	21,1	0/25
Heladera con freezer/ freezer solo	10,1	5,8	10,4	10,8	1/25
Heladera sin freezer	19,9	24,1	20,1	15,7	2/25
Computadora (PC, notebook, tablet, etc.)	11,4	13,6	11,3	12,1	0/25
Internet en la vivienda	12,7	13,6	12,6	13,9	1/25
Teléfono móvil con acceso a internet	10,2	7,9	10,5	9,0	0/25
Teléfono móvil sin acceso a internet	16,4	20,4	16,3	15,2	2/25
Cantidad de autos, camionetas y/o utilitarios en el hogar	29,9	24,6	31,6	15,2	3/25
Año del último vehículo	33,1	25,4	33,8	32,4	4/17

(continúa)

**Cuadro 4.5.3.2 (conclusión)**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Cantidad de motocicletas	43,2	41,4	44,7	29,1	6/25
Cantidad de ciclomotores	48,1	40,3	49,9	34,1	7/25
Cantidad de bicicletas	41,2	36,6	42,8	26,9	4/25

(1) El nivel de error de San Carlos de Bariloche se presenta en absolutos, dado que solo se cuenta en el denominador con 25 hogares ocupados.

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- El año del último vehículo solo debía ser respondido por los que tenían al menos un auto, camioneta o utilitario. Menos de un 3% de los que respondieron no tenerlos tienen marca en esta pregunta (cuadro 4.5.3.3).

**Cuadro 4.5.3.3 Año del vehículo. Error de secuencia. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (1)
Año del último vehículo	2,8	1,6	2,9	2,5	1/8

(1) Se presentan en absolutos dado que, en San Carlos de Bariloche, el denominador es inferior a 25.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.6 Relación o parentesco con la (el) jefa(e) o persona de referencia del hogar (P1)

### 4.6.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	Cuestionario básico
<p><b>1</b> ¿Cuál es la relación o parentesco con la (el) jefa(e) o persona de referencia del hogar?</p> <p>Jefa(e) o persona de referencia <input checked="" type="checkbox"/> 1</p>	<p><b>1</b> ¿Cuál es la relación o parentesco con la (el) jefa(e) o persona de referencia del hogar?</p> <p>Jefa(e) o persona de referencia <input checked="" type="checkbox"/> 1</p>
<p><b>1</b> ¿Cuál es la relación o parentesco con la (el) jefa(e) o persona de referencia del hogar?</p> <p>Cónyuge o pareja <input type="checkbox"/> 2            Hija(o) solo de jefa(e) <input type="checkbox"/> 3            Hija(o) solo de cónyuge <input type="checkbox"/> 4            Hija(o) de ambos <input type="checkbox"/> 5            Nuera / yerno <input type="checkbox"/> 6            Nieta(o) <input type="checkbox"/> 7            Madre / padre / suegra(o) <input type="checkbox"/> 8            Otros familiares <input type="checkbox"/> 9            Otros no familiares <input type="checkbox"/> 10            Servicio doméstico y sus familiares <input type="checkbox"/> 11</p>	<p><b>1</b> ¿Cuál es la relación o parentesco con la (el) jefa(e) o persona de referencia del hogar?</p> <p>Cónyuge o pareja <input type="checkbox"/> 2            Hija(o) solo de jefa(e) <input type="checkbox"/> 3            Hija(o) solo de cónyuge <input type="checkbox"/> 4            Hija(o) de ambos <input type="checkbox"/> 5            Nuera / yerno <input type="checkbox"/> 6            Nieta(o) <input type="checkbox"/> 7            Madre / padre / suegra(o) <input type="checkbox"/> 8            Otros familiares <input type="checkbox"/> 9            Otros no familiares <input type="checkbox"/> 10            Servicio doméstico y sus familiares <input type="checkbox"/> 11</p>

## 4.6.2

## Objetivo y principales características de las preguntas

- El objetivo de esta pregunta es identificar a una persona de referencia del hogar de manera de relacionar a todos los miembros del hogar con esta, para luego poder reconstruir la composición del hogar.
- Esta pregunta se encuentra presente en ambos cuestionarios y cuenta con un diseño formal diferente para el caso de la primera persona relevada en el hogar respecto de los restantes miembros. La primera hoja del cuestionario de población solo contiene la opción de respuesta “Jefa(e) o persona de referencia del hogar” y las restantes incluyen el resto de relaciones o parentescos posibles, excluyendo la opción “Jefa(e) o persona de referencia”. Fundamenta la elección de este diseño el tratar de aportar datos válidos que permitan construir una base de datos jerárquicamente estructurada a partir de considerar a la persona número 1 como referente del hogar.

## 4.6.3

## Resultados

- La pregunta presentó un error superior al de los censos anteriores (cuadro 4.6.3.1).

**Cuadro 4.6.3.1 Relación o parentesco con la jefa(e) o persona de referencia del hogar (P1). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Relación o parentesco	0,8	0,4	0,5	1,6	Se incorpora el término “persona de referencia”. Se divide la categoría hija/o en si es hija/o de jefa/e, de cónyuge o de ambos.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- La pregunta no supera el 2,2% de error en ningún área, siendo CABA la que manifiesta el error más alto y San Carlos de Bariloche, el más bajo (cuadro 4.6.3.2).

**Cuadro 4.6.3.2 Relación o parentesco con la (el) jefa(e) o persona de referencia del hogar (P1). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Relación o parentesco	1,6	2,2	1,6	1,7	1,1

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En las áreas donde se aplicaron los dos cuestionarios, el error del cuestionario A triplica al del B, aunque sin superar el 2,4% (cuadro 4.6.3.3).

**Cuadro 4.6.3.3 Relación o parentesco con la (el) jefa(e) o persona de referencia del hogar (P1).**  
**Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019.**  
**En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Relación o parentesco	0,8	2,4

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.7 Sexo, Edad y Fecha de nacimiento (P2 y P4)

### 4.7.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	Cuestionario básico
<p><b>2</b> ¿Cuál es el sexo de [nombre]?</p> <p>Refiere al sexo al nacer</p> <p>Mujer <input type="checkbox"/> 1 Varón <input type="checkbox"/> 2</p>	<p><b>2</b> ¿Cuál es el sexo de [nombre]?</p> <p>Refiere al sexo al nacer</p> <p>Mujer <input type="checkbox"/> 1 Varón <input type="checkbox"/> 2</p>
<p><b>4</b> ¿Cuántos años tiene? (en años cumplidos)</p> <p>Si todavía no cumplió el año, anote 000</p> <p>Años: <input type="text"/> <input type="text"/> <input type="text"/></p> <p>Fecha de nacimiento: <input type="text"/> <input type="text"/> DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/> <input type="text"/></p>	<p><b>4</b> ¿Cuántos años tiene? (en años cumplidos)</p> <p>Si todavía no cumplió el año, anote 000</p> <p>Años: <input type="text"/> <input type="text"/> <input type="text"/></p> <p>Fecha de nacimiento: <input type="text"/> <input type="text"/> DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/> <input type="text"/></p>

### 4.7.2 Objetivo y principales características de las preguntas

- El objetivo de estas preguntas responde a la finalidad fundamental de los censos de población de brindar la estructura de la población por sexo y edad. Permite obtener diferenciales en una multiplicidad de análisis sociodemográficos, como así también aportar a la elaboración de las proyecciones y estimaciones de población, focalización de políticas públicas, entre otros usos.
- Existen dos métodos para conocer la edad de la población. Uno a partir de la fecha de nacimiento de la persona censada y otro a partir de su edad en años cumplidos.
- La indagación de la fecha de nacimiento permite también llevar a cabo estudios sociodemográficos por cohortes, de fundamental importancia en el estudio de la mortalidad, la fecundidad, etc., como así también estimar tasas de escolaridad al 30 de junio del año calendario. La edad en años cumplidos es de más sencilla captación, en especial en el caso de que el censo no sea autorrespondido, como en el caso de un eventual e-censo complementario. Por ello, para asegurar la captación del dato, resulta de utilidad incorporar ambas preguntas.
- Estas preguntas se encuentran presentes en ambos cuestionarios de la II PP.

## 4.7.3

## Resultados

- El nivel de error de la pregunta Sexo se encontró dentro de los parámetros observados en censos anteriores. La Edad en años cumplidos muestra un error más elevado, mientras que la Fecha de nacimiento, y en especial el año de este, arrojó un error considerablemente superior al registrado en el Censo 2010 (cuadro 4.7.3.1).

**Cuadro 4.7.3.1 Sexo, Edad y Fecha de nacimiento (P2 y P4). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentaje**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Sexo	1,7	2,2	1,4	1,5	Se modifica la pregunta y se indaga ¿Cuál es el sexo de [nombre]?
Edad	1,4	1,0	1,5	2,6	
Día de nacimiento			1,8	4,9	
Mes de nacimiento	No se preguntó.		1,8	4,9	
Año de nacimiento			2,7	9,1	

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- En el análisis por área de prueba se observa que San Carlos de Bariloche presenta menor magnitud de error que el resto de las áreas, mientras que Mar del Plata arroja los errores más elevados para las variables Edad y Fecha de nacimiento (cuadro 4.7.3.2).

**Cuadro 4.7.3.2 Sexo, edad y fecha de nacimiento (P2 y P4). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Sexo	1,5	1,7	1,6	1,5	0,4
Edad	2,6	2,2	2,6	4,8	0,5
Día de nacimiento	4,9	3,8	4,6	8,3	4,3
Mes de nacimiento	4,9	5,2	4,5	8,9	4,4
Año de nacimiento	9,1	9,4	8,7	14,3	6,8

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En relación con los tipos de error que prevalecen, se observa que una parte importante del error total está constituido por los valores inválidos.<sup>20</sup>

<sup>20</sup> Se consideraron valores inválidos en edad aquellos mayores a 120; en día de nacimiento, aquellos mayores a 31; en mes de nacimiento, aquellos mayores a 12; y en año de nacimiento, aquellos menores a 1900 y mayores a 2019.

- En las preguntas relativas a Edad declarada y Año de nacimiento, el error inválido es superior a la no respuesta; por lo tanto, se subsanaría con el mejoramiento de la calidad de la lectura e interpretación de los datos.
- Para las preguntas Día y Mes de nacimiento el error de no respuesta es superior al error inválido (cuadro 4.7.3.3).

**Cuadro 4.7.3.3 Edad declarada, Día, Mes y Año de nacimiento (P2 y P4). Error según tipo. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Tipo de error	Total	Áreas de prueba			
			CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Edad declarada	Inválido	1,4	1,7	3,9	1,2	0,3
	No respuesta	1,2	0,5	1,0	1,4	0,2
	<b>Error total</b>	<b>2,6</b>	<b>2,2</b>	<b>4,8</b>	<b>2,6</b>	<b>0,5</b>
Día de nacimiento	Inválido	1,3	1,1	2,0	1,2	0,8
	No respuesta	3,6	2,6	6,4	3,4	3,5
	<b>Error total</b>	<b>4,9</b>	<b>3,8</b>	<b>8,3</b>	<b>4,6</b>	<b>4,3</b>
Mes de nacimiento	Inválido	1,5	2,5	2,6	1,4	1,0
	No respuesta	3,4	2,6	6,3	3,1	3,4
	<b>Error total</b>	<b>4,9</b>	<b>5,2</b>	<b>8,9</b>	<b>4,5</b>	<b>4,4</b>
Año de nacimiento	Inválido	5,0	5,8	7,9	4,9	2,9
	No respuesta	4,1	3,6	6,5	3,8	3,9
	<b>Error total</b>	<b>9,1</b>	<b>9,4</b>	<b>14,3</b>	<b>8,7</b>	<b>6,8</b>

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- El nivel de error en el cuestionario B es ostensiblemente menor en las variables Sexo, Edad en años cumplidos y Año de nacimiento, mientras que para las variables Día y Mes de nacimiento esta diferencia se reduce (cuadro 4.7.3.4).

**Cuadro 4.7.3.4 Sexo, edad y fecha de nacimiento (P2 y P4). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Sexo	0,4	2,0
Edad	1,0	4,1
Día de nacimiento	5,2	6,0
Mes de nacimiento	5,4	7,1
Año de nacimiento	8,7	11,8

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

#### 4.7.4

#### Evaluación demográfica de la variable edad

- La evaluación de calidad y la exactitud en la declaración de la edad también se puede medir a través de los índices de preferencia de dígitos: el índice de Whipple y el índice de Myers.
- El índice de Whipple expresa la preferencia por las edades terminadas en 0 y en 5 y los valores se encuentran entre 100 y 500. Se alcanza el valor 100 cuando los datos son correctos (es decir, no hay atracción ni rechazo por ninguno de estos dígitos finales de la edad) y el valor 500 en el caso extremo en que la totalidad de la población del universo de estudio declare edades terminadas en los dígitos finales 0 y 5.
- El índice de Myers mide la intensidad en la preferencia o el rechazo de cada dígito final de la edad 0 a 9 y además proporciona un índice resumen que permite clasificar el poder de atracción de todos los dígitos finales en las estadísticas por edad. Los valores van de 0 (cuando toda la población declara correctamente la edad) a 180 (cuando toda la población declara edades terminadas en un solo dígito final). Adicionalmente, se puede medir el rechazo o la preferencia por cada dígito. Valores del índice entre 5,1 y 15,0 indican un nivel intermedio de error en la declaración de la edad.
- Los índices que resultan de Edad calculada (a partir de la fecha de nacimiento) y Edad en años cumplidos muestran resultados similares. Sin embargo, pareciera obtenerse un dato de mejor calidad cuando se pregunta por la Fecha de nacimiento.

**Cuadro 4.7.4.1 Índices de declaración de la edad según edad declarada y edad calculada a partir de la fecha de nacimiento. II PP 2019**

Variable	Índice de Whipple			Índice de Myers		
	Varones	Mujeres	Total	Varones	Mujeres	Total
Edad declarada	103,0	104,1	103,6	6,8	5,7	6,1
Edad calculada (fecha de nacimiento)	102,7	100,5	101,5	5,3	5,7	4,3

**Fuente:** Base de datos con estructura jerárquica y sin consistir. II PP 2019.

#### 4.8

#### Identidad de género (P3)

##### 4.8.1

##### Diseño de preguntas en el cuestionario

###### Cuestionario ampliado

###### Cuestionario básico

3 Y de acuerdo a la identidad de género, ¿[nombre] se considera...

- mujer?  1  
varón?  2  
mujer trans / travesti?  3  
varón trans?  4  
otro?  5  
Ignorado  9

3 Y de acuerdo a la identidad de género, ¿[nombre] se considera...

- mujer?  1  
varón?  2  
mujer trans / travesti?  3  
varón trans?  4  
otro?  5  
Ignorado  9

## 4.8.2

### Objetivo y principales características de las preguntas

- La pregunta se orienta a visibilizar una población demandante de políticas públicas específicas.
- Se indaga en ambos cuestionarios.

## 4.8.3

### Resultados

- El nivel de error de la pregunta Identidad de género se encuentra en valores cercanos a los de la variable Sexo (1,7% y 1,4%, respectivamente). Al ser la primera vez que se incorpora en una prueba piloto censal, esta pregunta no tiene valor de referencia para ser comparada (cuadro 4.8.3.1).

**Cuadro 4.8.3.1 Identidad de género (P3). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Identidad de género		No se preguntó.		1,7	Se incorpora la pregunta

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- En el análisis por área de prueba se observa que San Carlos de Bariloche presenta menor magnitud de error que el resto de las áreas, mientras que Humahuaca arroja el porcentaje de error más elevado (cuadro 4.8.3.2).

**Cuadro 4.8.3.2 Identidad de género (P3). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Identidad de género	1,7	1,4	1,9	1,6	0,5

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- El nivel de error en el cuestionario B es la mitad del registrado para el cuestionario A (cuadro 4.8.3.3).

**Cuadro 4.8.3.3 Identidad de género (P3). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Identidad de género	0,7	1,6

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Si bien la evaluación cualitativa concluyó que esta pregunta provocó incomodidad tanto en la persona censista como en la censada, en especial cuando se trataba de niñas y niños o personas mayores,<sup>21</sup> esto no impactó en la calidad de los datos (cuadro 4.8.3.4).

**Cuadro 4.8.3.4 Identidad de género (P3). Error total por grupos de edad. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Grupos de edad	Error total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
0 a 10 años	1,8	1,1	2,2	0,0	0,6
11 a 20 años	1,4	3,4	1,3	1,7	0,6
21 a 30 años	1,2	0,7	1,3	0,6	1,2
31 a 40 años	1,6	0,8	1,8	2,9	0,0
41 a 50 años	1,8	1,7	1,9	2,7	0,0
51 a 60 años	1,4	0,0	1,8	1,1	0,0
61 a 70 años	2,1	0,00	2,60	0,9	1,1
71 y más	1,8	0,00	2,31	2,1	0,0

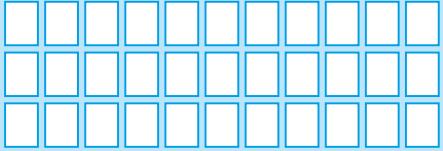
Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.9 Educación (A: P5 a P10) (B: P5 a P8)

### 4.9.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado		Cuestionario básico	
<b>5</b> ¿Sabe leer y escribir?	Sí <input type="checkbox"/> 1 No <input type="checkbox"/> 2	<b>5</b> ¿Cursa o cursó en algún establecimiento educativo (guardería, jardín, escuela, instituto, universidad)?	Incluye educación a distancia o virtual formal
<b>6</b> ¿Cursa o cursó en algún establecimiento educativo (guardería, jardín, escuela, instituto, universidad)?	Cursa en un establecimiento estatal <input type="checkbox"/> 1 Cursa en un establecimiento privado <input type="checkbox"/> 2 No cursa, pero cursó <input type="checkbox"/> 3 Nunca cursó <input type="checkbox"/> 4 → 11	Cursa en un establecimiento estatal <input type="checkbox"/> 1 Cursa en un establecimiento privado <input type="checkbox"/> 2 No cursa, pero cursó <input type="checkbox"/> 3 Nunca cursó <input type="checkbox"/> 4 → 9	Incluye educación a distancia o virtual formal
<b>7</b> ¿Qué nivel educativo cursa o cursó?	Jardín maternal, guardería, centro de cuidado, salas de 0 a 5, jardín de infantes o preescolar <input type="checkbox"/> 1 → 11 Primario <input type="checkbox"/> 2 EGB <input type="checkbox"/> 3 Secundario <input type="checkbox"/> 4 Polimodal <input type="checkbox"/> 5 Terciario no universitario <input type="checkbox"/> 6 Universitario de grado <input type="checkbox"/> 7 Posgrado (especialización, maestría o doctorado) <input type="checkbox"/> 8	<b>6</b> ¿Qué nivel educativo cursa o cursó?	Jardín maternal, guardería, centro de cuidado, salas de 0 a 5, jardín de infantes o preescolar <input type="checkbox"/> 1 → 9 Primario <input type="checkbox"/> 2 EGB <input type="checkbox"/> 3 Secundario <input type="checkbox"/> 4 Polimodal <input type="checkbox"/> 5 Terciario no universitario <input type="checkbox"/> 6 Universitario de grado <input type="checkbox"/> 7 Posgrado (especialización, maestría o doctorado) <input type="checkbox"/> 8

<sup>21</sup> Ver INDEC (2019b, p. 66).

<b>8</b> ¿Completó el nivel?	Sí <input type="checkbox"/> 1 No, está cursando actualmente <input type="checkbox"/> 2 No completó ese nivel <input type="checkbox"/> 3	<b>7</b> ¿Completó el nivel?	Sí <input type="checkbox"/> 1 No, está cursando actualmente <input type="checkbox"/> 2 No completó ese nivel <input type="checkbox"/> 3
<b>9</b> ¿Cuál es el grado o año que cursa o el último grado o año que aprobó?		Grado o año: <input type="text"/> ° Ignorado <input type="checkbox"/> 99	
		Terciario, universitario o posgrado pase a <b>10</b> Si no, pase a <b>11</b>	
<b>10</b> ¿Cuál es la carrera universitaria de grado o terciaria que completó?			
			

#### 4.9.2 Objetivo y principales características de las preguntas

- El objetivo de las preguntas es conocer el nivel de alfabetismo, construir el indicador máximo nivel de instrucción alcanzado por la población, establecer la asistencia de la población en edad escolar, y la carrera terciaria o universitaria finalizada para poder relacionar por, ejemplo, el título terciario o universitario obtenido y la inserción en el mercado de trabajo.
- Las preguntas sobre asistencia y máximo nivel de instrucción se indagan en ambos cuestionarios. La condición de alfabetismo y la carrera terciaria y universitaria solo se investigan en el cuestionario A.

#### 4.9.3 Resultados

- A nivel general, el error de todas las preguntas del bloque supera el de los censos anteriores.
- La pregunta sobre Carrera terciaria o universitaria no fue respondida por un tercio de las personas que cursaron esos niveles (cuadro 4.9.3.1).

**Cuadro 4.9.3.1 Educación (A: P5 a P10) (B: P5 a P8). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Condición de alfabetismo	2,5	1,1	1,3	4,7	Se indaga para todas las edades. Cambia de lugar el bloque.
Asistencia actual o pasada (y sector del establecimiento)	3,1	5,6 (Asiste) 1,3 (Asistió) 1,4 (Sector)	2,3	5,4	Se indaga para todas las edades. Cambia de lugar el bloque. Se modifica la palabra "asiste" por "cursa". Se incorpora aclaración educación a distancia o virtual. Se incluye el sector del establecimiento.

(continúa)

**Cuadro 4.9.3.1 (conclusión)**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Nivel educativo que cursa o cursó	4,4 (Cursa)	1,5 (Cursa)	2,4	3,4	Se elimina educación especial. Se elimina el desdoble de primario de seis o siete años. Se detallan los títulos de posgrado y se modifica la denominación de los niveles.
	1,9 (Cursó)	0,8 (Cursó)			
Completitud de nivel	1,4	1,7	1,2	2,4	Se incorpora en las categorías la opción de estar cursando actualmente.
Grado o año que está cursando o que aprobó	0,8 (Cursando)	1,4 (Cursando)	7,1	11,7	Se incorpora en la pregunta si es el grado o año que cursa o el último que aprobó.
	1,3 (Aprobó)	6,4 (Aprobó)			
Carrera universitaria de grado o terciaria que completó	No se preguntó.	s/d	No se preguntó.	34,6	Se incorpora la pregunta.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- San Carlos de Bariloche y Mar del Plata son las áreas en las que la magnitud del error en la variable Carrera universitaria de grado o terciaria que completó es más relevante (cuadro 4.9.3.2).

**Cuadro 4.9.3.2 Educación (A: P5 a P10) (B: P5 a P8). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Condición de alfabetismo	4,7	4,3	4,8	4,0	3,1
Asistencia actual o pasada	5,4	3,7	5,5	6,7	5,1
Nivel educativo que cursa o cursó	3,4	2,0	3,5	4,0	3,5
Completitud de nivel	2,4	2,1	2,6	2,1	1,5
Grado o año que está cursando o que aprobó	11,7	6,1	12,4	14,9	7,1
Carrera universitaria de grado o terciaria que completó	34,6	32,0	34,6	39,3	40,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Comparando los resultados según los datos correspondan a los cuestionarios A o B, el cuestionario A exhibe un nivel de error superior en las preguntas sobre Completitud de nivel y Grado o año que está cursando o que aprobó, mientras que para Asistencia actual o pasada y Nivel educativo que cursa o cursó, los errores son levemente superiores en el cuestionario B (cuadro 4.9.3.3).

**Cuadro 4.9.3.3 Educación (A: P5 a P10) (B: P5 a P8). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Asistencia actual o pasada	5,6	5,0
Nivel educativo que cursa o cursó	3,4	3,1
Completitud de nivel	1,4	2,4
Grado o año que está cursando o que aprobó	7,9	11,4

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019

- A excepción de la Carrera terciaria o universitaria, el error de secuencia oscila entre 7,4% y 18,4% cuando no se distingue por áreas para las preguntas del bloque, lo que indica dificultades del censista para seguir los pasos. Esto se observa con especial énfasis en CABA, mientras que para la pregunta sobre Completitud de nivel, en todas las áreas de prueba (cuadro 4.9.3.4).

**Cuadro 4.9.3.4 Educación (A: P5 a P10) (B: P5 a P8). Error de secuencia. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Nivel educativo que cursa o cursó	7,4	19,0	6,6	8,5	4,7
Completitud de nivel	18,4	30,8	16,7	24,8	15,2
Grado o año que está cursando o que aprobó	9,3	21,5	7,7	11,4	10,5
Carrera terciaria o universitaria que completó	2,4	0,8	2,7	0,0	3,8

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Tal como se menciona en el cuadro 4.9.3.1, del análisis de la implementación de la ampliación del universo de la pregunta Condición de alfabetismo, surge que la población de 0 a 2 años cuenta con un error total de 9,5% en dicha pregunta, y 24% en Asistencia actual o pasada. Si solo se analiza el error en la población de 3 años y más, se advierte que el error total disminuye a un 3,8% para la primera pregunta y a un 4% para la segunda. Es decir que una elevada proporción del error se debe al aporte de la población menor de 3 años (cuadro 4.9.3.5).

**Cuadro 4.9.3.5 Error total en las preguntas de alfabetización y asistencia escolar según edad agrupada. Cuestionario ampliado. II PP 2019. En porcentajes**

Pregunta	Edad agrupada		Total de población
	De 0 a 2 años	3 años y más	
Condición de alfabetismo	9,7	3,8	4,7
Asistencia actual o pasada	24,1	4,1	5,4

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir II PP 2019.

- La pregunta Carrera terciaria o universitaria que completó presenta niveles de error que superan el 30% en todas las áreas. Si se descompone el nivel de error de acuerdo con el Máximo nivel de instrucción alcanzado y su Completitud, se observa que disminuye considerablemente para los que completaron los niveles terciarios y universitarios, respecto de los que no lo hicieron (cuadro 4.9.3.6).

**Cuadro 4.9.3.6 Error total en la pregunta Carrera terciaria o universitaria según nivel cursado y completitud de nivel. Cuestionario ampliado. II PP 2019. En porcentajes**

Máximo nivel que cursa o cursó	Completitud de nivel	
	Completo	No completó
Terciario	9,3	63,6
Universitario	4,3	57,1
Posgrado	0,0	0,0

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.10 Migraciones (A: P11 a P13) (B: P9)

### 4.10.1 Diseño de preguntas en el cuestionario

**Cuestionario ampliado**

**11 ¿Dónde nació?**

En Ciudad Autónoma de Buenos Aires  1  
 En la provincia de Buenos Aires  2  
 En otra provincia argentina  3 → Anote la provincia  
 En otro país  4 → Anote el país  
 ↓ ¿En qué año llegó a Argentina?

**12 ¿Dónde vivía hace 5 años?**

En Ciudad Autónoma de Buenos Aires  1  
 En la provincia de Buenos Aires  2  
 En otra provincia argentina  3 → Anote la provincia  
 En otro país  4 → Anote el país  
 No había nacido  5

**13 ¿Dónde vive habitualmente?**

En Ciudad Autónoma de Buenos Aires  1  
 En la provincia de Buenos Aires  2  
 En otra provincia argentina  3 → Anote la provincia  
 En otro país  4 → Anote el país

9. ¿Dónde nació?

En Ciudad Autónoma de Buenos Aires	<input type="checkbox"/> 1	→	Añote la provincia											
En la provincia de Buenos Aires	<input type="checkbox"/> 2		Añote el país											
En otra provincia argentina	<input type="checkbox"/> 3		↓											
En otro país	<input type="checkbox"/> 4	¿En qué año llegó a Argentina? <span style="border: 1px solid black; padding: 2px; display: inline-block; width: 100px; height: 15px; vertical-align: middle;"></span>												

#### 4.10.2 Objetivo y principales características de las preguntas

- La pregunta sobre Lugar de nacimiento de la persona censada permite distinguir el stock de la población nativa y de la extranjera y aporta a la medición de la migración internacional e interna.
- El Año de llegada al país de la población no nativa brinda información sobre las distintas cohortes de migrantes y permite observar si sus características demográficas han variado a lo largo del tiempo, identificar si han cambiado los patrones de migración por lugar de origen y en qué período se produjo específicamente ese fenómeno.
- Con la información sobre Lugar de nacimiento y Lugar de residencia habitual se obtiene la migración absoluta. La combinación de las preguntas sobre Lugar de residencia habitual y Lugar de residencia habitual 5 años antes permite construir matrices de migración y estimar las tasas netas migratorias anuales para la segunda mitad del período intercensal. Esta información constituye un insumo relevante para la elaboración de las proyecciones de población.
- El proceso de operacionalización de las variables del bloque de migración previó incluir la distinción específica entre CABA y la provincia de Buenos Aires, ante el conocimiento acumulado respecto de que la población asume y responde “Buenos Aires” como única denominación en el caso de una y otra jurisdicción y ello genera, en el mejor de los casos, la necesidad de que el censista reitere la pregunta para precisar si la respuesta refiere a la ciudad o a la provincia.
- Las preguntas por Lugar de nacimiento y Año de llegada al país de los extranjeros se incluyeron en ambos cuestionarios, mientras que en el cuestionario A se incorporaron las relativas a la Residencia habitual 5 años antes y Residencia habitual al momento del censo.

#### 4.10.3 Resultados

- Para las preguntas cerradas (¿Dónde nació?, ¿Dónde vivía hace 5 años?, ¿Dónde reside habitualmente?) el nivel del error supera ampliamente a los de censos anteriores.
- La pregunta sobre Año de llegada al país (pregunta abierta) también presenta un nivel superior al de los censos previos (cuadro 4.10.3.1).

**Cuadro 4.10.3.1 Migración (A: P9 a P13) (B: P9). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En valores absolutos y porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Lugar de nacimiento	2,0	4,9	0,9	33,7	Se modifica la operacionalización y se cambia la ubicación del bloque.
Provincia	s/d	s/d	4,8	0,3	
País	s/d	s/d	0,7	1,1	
Año de llegada	s/d	17,5	2,1	26,6	
Residencia 5 años antes	2,5	4,3	5,1	38,2	
Provincia	s/d	s/d	s/d	0,5	
País	s/d	s/d	No se preguntó.	0,0	
Residencia habitual	2,3	3,3	3,5	39,3	
Provincia	s/d	s/d	s/d	0,5	
País	s/d	s/d	No se preguntó.	3/7	

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- En el análisis por área de prueba se observa que Humahuaca es la jurisdicción que contribuye al elevado nivel de error de las preguntas cerradas (¿Dónde nació?, ¿Dónde vivía hace 5 años?, ¿Dónde reside habitualmente?). Mar del Plata y CABA son las áreas con menor proporción de error en todas las preguntas del bloque (cuadro 4.10.3.2).

**Cuadro 4.10.3.2 Migración (A: P9 a P13) (B: P9). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total (1)	CABA	Humahuaca (1)	Mar del Plata (1)	San Carlos de Bariloche
Lugar de nacimiento	33,7	7,4	42,7	4,0	13,6
Provincia	0,3	1,9	0,3	1,5	0,5
País	1,1	0,3	0,0	0,0	5,6
Año de llegada	26,6	17,3	61,6	22,5	33,3
Residencia 5 años antes	38,2	2,2	44,3	4,0	12,5
Provincia	0,5	0,0	0,4	20,0 (3/15)	5,6
País	0,0	0,0	0,0	0,0	0,0
Residencia habitual	39,3	2,0	45,7	2,8	15,6
Provincia	0,5	0,0	0,3	3/6	5,6
País	3/7	0,0	3/7	0,0	0,0

(1) En determinadas variables se presentan en absolutos dado que, en Humahuaca y Mar del Plata el denominador es inferior a 25.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Las únicas preguntas que se incluyeron en ambos cuestionarios fueron Lugar de nacimiento y Año de llegada al país para los extranjeros. En estos casos, el cuestionario B mostró errores considerablemente más altos que el cuestionario A. Por su parte, para las dos preguntas abiertas destinadas a anotar específicamente el país y la provincia (pregunta abierta), para los que declararon haber nacido en otro país o en otra provincia, se observa el comportamiento inverso.

**Cuadro 4.10.3.3 Lugar de nacimiento (A: P11 y B: P9). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentaje**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Lugar de nacimiento	10,8	6,0
Provincia	0,5	1,8
País	0,0	0,3
Año de llegada	31,0	18,1

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Solo en Humahuaca los errores de secuencia son elevados, en especial para las preguntas sobre Provincia de nacimiento y Residencia habitual. Este dato resulta consistente con lo observado en la evaluación cualitativa, según la cual las personas que nacieron en Humahuaca escribían “Jujuy” en el lugar destinado a anotar la provincia sin haber marcado previamente la opción “En otra provincia argentina”<sup>22</sup> (cuadro 4.10.3.4).

**Cuadro 4.10.3.4 Migración (A: P11 a P13) (B: P9). Error de secuencia. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA	Humahuaca (%)	Mar del Plata	San Carlos de Bariloche (%)
Provincia de nacimiento	2,9	1,9	10,5	2,0	4,9
País de nacimiento	1,1	1,4	1,4	0,2	0,0
Año de llegada	0,2	2,1	0,1	0,0	0,1
Provincia de residencia 5 años antes	2,4	1,3	5,4	2,0	0/2
País de residencia 5 años antes	0,7	0,5	0,9	0,0	1,8
Provincia de residencia habitual	2,1	2,0	4/18	1,7	0/0
País de residencia habitual	0,7	0,9	0,8	0,1	1,9
Residencia habitual	39,3	2,0	45,7	2,8	15,6
Provincia	0,5	0,0	0,3	3/6	5,6
País	3/7	0,0	3/7	0,0	0,0

(1) En determinadas variables se presentan en absolutos dado que, en Humahuaca y San Carlos de Bariloche, el denominador es inferior a 25.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

<sup>22</sup> Ver INDEC (2019b, p. 67).

- Si bien los errores detectados son elevados, en el cuadro 4.10.3.5 se observa que la información sobre Lugar de nacimiento, Lugar de residencia habitual 5 años antes y Lugar de residencia habitual al momento del censo se obtiene mediante la combinación de las respuestas a las preguntas cerradas y las preguntas abiertas y el error se reduce sustancialmente.
- El error total de la pregunta Lugar de nacimiento se reduce al 1,4%; el de Residencia habitual 5 años antes lo hace al 3,5%; y el de la pregunta de Residencia habitual, al 2,5%.
- Humahuaca y Mar del Plata son las que mantienen mayores porcentajes de población de la que no se puede conocer a información, aunque sin superar el 3,6%.

**Cuadro 4.10.3.5 Migración (A: P11 a P13) (B: P9). Error total recalculado con respuestas abiertas. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Provincia de nacimiento	2,9	1,9	10,5	2,0	4,9
País de nacimiento	1,1	1,4	1,4	0,2	0,0
Año de llegada	0,2	2,1	0,1	0,0	0,1

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.11 Discapacidad y Previsión social (A: P14 a P16)

### 4.11.1 Diseño de preguntas en el cuestionario

#### Cuestionario ampliado

<b>14</b> ¿[Nombre] tiene dificultad o limitación...					
Lea las opciones de respuesta al finalizar cada pregunta		Sí, no puede	Sí, mucha dificultad	Sí, alguna dificultad	No, ninguna dificultad
para caminar o subir escaleras?		<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
para entender, recordar o concentrarse? Ej.: recordar la dirección de su casa		<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
para hablar o comunicarse, aun usando lengua de señas?		<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
para oír, aun cuando usa audífono?		<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
para ver, aun cuando usa anteojos o lentes de contacto? Ej.: ver la letra impresa en un diario		<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
para comer, bañarse o vestirse sola(o)?		<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<b>15</b> ¿Cobra jubilación o pensión?					
Sí <input type="checkbox"/> 1					
No <input type="checkbox"/> 2 → <b>17</b>					
<b>16</b> ¿Cobra...					
solo jubilación? <input type="checkbox"/> 1					
solo pensión por fallecimiento? <input type="checkbox"/> 2					
jubilación y pensión por fallecimiento? <input type="checkbox"/> 3					
solo pensión de otro tipo? <input type="checkbox"/> 4					

## 4.11.2 Objetivo y principales características de las preguntas

- El objetivo de las preguntas sobre Dificultades o limitaciones es cuantificar a la población que presenta dificultad en al menos una de las siguientes actividades: ver; oír; caminar o subir escaleras; comer, bañarse o vestirse solo/a; hablar o comunicarse; entender, recordar o concentrarse. El bloque de preguntas se basa en el set corto recomendado por el Grupo de Washington,<sup>23</sup> con algunas modificaciones que surgieron de la experiencia de la primera prueba piloto realizada en 2017 y del Estudio Nacional sobre el Perfil de las Personas con Discapacidad del 2018.<sup>24</sup> Además, el bloque tiene en cuenta el marco teórico de la Clasificación Internacional del Funcionamiento de la Discapacidad y de la Salud (CIF).<sup>25</sup>
- El objetivo de las preguntas de Previsión social es cuantificar a la población que recibe ingresos monetarios en concepto de jubilación o pensión, e identificar el tipo de prestación percibida.
- Este bloque de preguntas se incluyó solo en el cuestionario A.

## 4.11.3 Resultados

- A excepción del Tipo de jubilación, las restantes preguntas del bloque que se indagaron en censos anteriores presentan errores superiores en la II PP.
- Para las seis preguntas sobre Dificultades o limitaciones, el nivel de error más bajo se encuentra en la primera pregunta, referida a la dificultad para caminar o subir escaleras, mientras que los niveles de error más altos se encuentran en las tres últimas preguntas del bloque (cuadro 4.11.3.1).

**Cuadro 4.11.3.1 Dificultad o limitación y Previsión social (A: P14 a P16). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Dificultad o limitación para caminar o subir escaleras?	No se preguntó.	Se preguntó a nivel de hogar.	2,8	4,0	Se modifica el orden de las dimensiones, la redacción, se agrega un dominio de auto valimiento y de comprensión y entendimiento, se elimina la dificultad motriz superior y se reemplaza el sistema de categorías dicotómico por un gradiente de gravedad de la dificultad.
Dificultad o limitación para entender, recordar o concentrarse?			3,0	4,6	
Dificultad o limitación para hablar o comunicarse, aun usando lengua de señas?			No se preguntó.	4,6	
Dificultad o limitación para oír, aun cuando usa audífono?			2,8	5,0	

(continúa)

<sup>23</sup> El Grupo de Washington sobre Estadísticas de la Discapacidad es un grupo de trabajo de la Comisión Estadística de las Naciones Unidas creado en 2001. Su objetivo es promover y coordinar la cooperación internacional en el área de las estadísticas de salud, centrándose en el desarrollo de herramientas de recopilación de datos estadísticos sobre la población con discapacidad, satisfaciendo la necesidad de medir las principales deficiencias de la población a nivel nacional y que sean comparables internacionalmente. En su labor han generado distintos sets de preguntas, entre los que se destaca el set corto de 6 preguntas (*short set*), elaborado para medir discapacidad en censos o encuestas (ver: <http://www.washington-group-disability.com/washington-group-question-sets/short-set-of-disability-questions>).

Por otra parte, es importante mencionar que los principios y las recomendaciones para los censos de población y habitación de Naciones Unidas (2017) aconsejan utilizar el set corto de preguntas del Grupo de Washington para los censos de población.

<sup>24</sup> Ver INDEC (2019a).

<sup>25</sup> La Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud (CIF) constituye el marco conceptual de la Organización Mundial de la Salud (OMS) para la comprensión del funcionamiento, la discapacidad y la salud. Es una clasificación universal que establece un marco y lenguaje estandarizados para describir la salud y las dimensiones relacionadas con ella.

**Cuadro 4.11.3.1 (conclusión)**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Dificultad o limitación para ver, aun cuando usa anteojos o lentes de contacto?	No se preguntó.	Se preguntó a nivel de hogar.	2,0	5,0	Se modifica el orden de las dimensiones, la redacción, se agrega un dominio de auto valimiento y de comprensión y entendimiento, se elimina la dificultad motriz superior y se reemplaza el sistema de categorías dicotómico por un gradiente de gravedad de la dificultad.
Dificultad o limitación para comer, bañarse o vestirse sola(o)?			No se preguntó.	4,9	
¿Cobra jubilación o pensión?	3,5	3,2	4,3	6,4	Se modifica “recibe” por “cobra”.
Tipo de jubilación	No se preguntó.		6,7	5,4	Se modifica la redacción de las categorías. Se elimina “solo pensión no contributiva asistencial o graciable” y se agrega “solo pensión de otro tipo”.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos de la II PP 2019.

- Humahuaca registra las proporciones más elevadas de error en las preguntas sobre dificultad o limitación, mientras que San Carlos de Bariloche presenta los niveles de error más bajos.
- A excepción de San Carlos de Bariloche, el nivel de error se incrementa a medida que se indaga cada pregunta.
- En las preguntas de Previsión social, las jurisdicciones presentan errores similares, aunque para la primera de las preguntas es superior en Humahuaca y para la segunda (que deviene de un filtro de la pregunta anterior), en Mar del Plata (cuadro 4.11.3.2).

**Cuadro 4.11.3.2 Dificultad o limitación y Previsión social (A: P14 a P16). Error total. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Dificultad o limitación para caminar o subir escaleras?	4,0	3,1	1,7	4,3	1,6
Dificultad o limitación para entender, recordar o concentrarse?	4,6	4,3	2,9	4,8	3,1
Dificultad o limitación para hablar o comunicarse, aun usando lengua de señas?	4,6	4,5	3,3	4,7	1,6
Dificultad o limitación	5,0	4,9	3,6	5,1	3,1
Dificultad o limitación para ver, aun cuando usa anteojos o lentes de contacto?	5,0	4,6	4,0	5,1	3,1
Dificultad o limitación para comer, bañarse o vestirse sola(o)?	4,9	5,0	4,7	4,9	1,6
¿Cobra jubilación o pensión?	6,4	5,3	5,5	6,6	1,6
Tipo de jubilación	5,4	5,3	7,0	5,3	0,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En consonancia con lo observado en la evaluación cualitativa,<sup>26</sup> en el cuadro 4.11.3.3 se registra que el nivel de error es mayor en el caso de los niños de 0 a 5 años respecto del resto de la población. También se observa una tendencia al incremento del error en las personas de 65 años y más; especialmente en las preguntas que hacen referencia a las dificultades o limitaciones para oír y para ver, en donde el nivel de error alcanza el 5,9% y 5,8%, respectivamente.

**Cuadro 4.11.3.3 Dificultad o limitación (A: P14). Error total por grupo de edad. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Grupos de edad	dificultad o limitación para caminar o subir escaleras?	para entender, recordar o concentrarse?	para hablar o comunicarse, aun usando lengua de señas?	para oír, aun cuando usa audífono?	para ver, aun cuando usa anteojos o lentes de contacto?	para comer, bañarse o vestirse sola(o)?
0-5 años	9,8	10,8	10,8	10,9	10,9	11,1
6-14 años	3,0	3,5	2,8	3,5	3,6	3,4
15-39 años	2,7	3,0	3,0	3,4	3,5	3,2
40-64 años	3,7	4,1	4,2	4,5	4,2	4,5
65 años y más	4,1	4,7	5,4	5,9	5,8	5,3

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Con el objetivo de conocer el nivel del error para la elaboración del indicador prevalencia de personas con al menos una dificultad o limitación, en el cuadro 4.11.3.4 se presenta el error en el bloque completo de preguntas sobre dificultades o limitaciones. Se observa que el 91,3% de la población no registra errores en ninguna de las preguntas del bloque y que el 2,6% de las personas posee errores en el bloque completo.

**Cuadro 4.11.3.4 Dificultad o limitación (A: P14). Cantidad de preguntas con error en el conjunto de preguntas. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

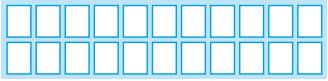
Cantidad de preguntas con error	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Bloque sin error	91,3	92,3	91,7	91,2	93,8
Bloque con una pregunta con error	3,2	2,0	4,3	3,1	4,7
Bloque con dos preguntas con error	1,1	1,3	1,1	1,1	0,0
Bloque con tres preguntas con error	0,7	1,1	0,7	0,7	0,0
Bloque con cuatro preguntas con error	0,6	0,6	0,7	0,6	0,0
Bloque con cinco preguntas con error	0,6	0,4	0,4	0,6	0,0
Bloque con todas las preguntas con error	2,6	2,4	1,2	2,8	1,6

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

<sup>26</sup> Ver INDEC (2019b, pp. 67-68).

## 4.12 Etnias (A: P17) (B: P10)

### 4.12.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	<p><b>17</b> ¿[Nombre] se reconoce...</p> <p>indígena o descendiente de pueblos indígenas u originarios? <input type="checkbox"/> 1 → ¿De qué pueblo?</p> <p>afrodescendiente o con antepasados negros? <input type="checkbox"/> 2</p> <p><input type="checkbox"/> 3</p> <p><input type="checkbox"/> 9</p> <p>Ninguna de las anteriores</p> <p>Ignorado</p> 
Cuestionario básico	<p><b>10</b> ¿[Nombre] se reconoce...</p> <p>indígena o descendiente de pueblos indígenas u originarios? <input type="checkbox"/> 1 → ¿De qué pueblo?</p> <p>afrodescendiente o con antepasados negros? <input type="checkbox"/> 2</p> <p><input type="checkbox"/> 3</p> <p><input type="checkbox"/> 9</p> <p>Ninguna de las anteriores</p> <p>Ignorado</p> 

### 4.12.2 Objetivo y principales características de las preguntas

- El objetivo de esta pregunta es identificar el Autorreconocimiento de pertenencia o de origen ancestral de la población indígena y afrodescendiente. Para aquellas personas que se reconocieron indígenas o descendientes de pueblos indígenas u originarios, se solicitó que se indicara el Pueblo de autorreconocimiento (pregunta abierta).
- Si bien se trata de dos preguntas independientes, ya que las personas pueden responder afirmativamente a ambos reconocimientos, debido a problemas de espacio del cuestionario, se las agrupó en una sola pregunta, lo cual permite una sola respuesta.
- Esta pregunta se encuentra presente en ambos cuestionarios de la II PP.

### 4.12.3 Resultados

La pregunta no resulta estrictamente comparable, ya que en los censos anteriores se indagó solo en el cuestionario del hogar, es decir, se consultó si había alguna persona en el hogar perteneciente o descendiente a pueblos indígenas (Censos 2001 y 2010) y afrodescendientes (solo Censo 2010), y en la II PP se investigó a cada persona del hogar. Sin embargo, se menciona que el nivel de error fue superior en la prueba piloto (cuadro 4.12.3.1).

**Cuadro 4.12.3.1 Etnias (A: P17) (B: P10). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentaje**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Etnias				10,7	
Pueblo indígena u originario	No se preguntó.	Se indagó a nivel de hogar.		6,1	Se incorpora a nivel de población. Se indaga en la misma pregunta autorreconocimiento “indígena” y “afrodescendiente”. Para la categoría afrodescendiente se agrega la expresión “negro”.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- El error en la pregunta sobre Pertenencia étnica alcanza al 10,7% de la población, siendo Humahuaca el área de prueba donde se registra el mayor nivel de error y CABA, el menor.
- Para la población que se autorreconoce indígena o descendiente de pueblos indígenas u originarios, la pregunta sobre Pueblo indígena u originario tiene un 6,1% de error, destacándose Mar del Plata por su alto nivel de error y Humahuaca por ser el área con menor nivel de error (cuadro 4.12.3.2).

**Cuadro 4.12.3.2 Etnias (A: P17) (B: P10). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Etnias	10,7	8,1	11,2	8,7	10,5
Pueblo indígena u originario	6,1	17,8	4,9	39,1	7,1

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Para las áreas de prueba donde se aplicaron cuestionarios B y A, la pregunta sobre Autorreconocimiento de pertenencia o descendencia indígena o afrodescendiente registra un error superior en el cuestionario B y lo contrario sucede con la pregunta abierta sobre Pueblo indígena u originario (cuadro 4.12.3.3).

**Cuadro 4.12.3.3 Etnias (A: P17) (B: P10). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Etnias	10,6	7,7
Pueblo indígena u originario	11,6	20,0

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Por su parte, se detectó una proporción muy baja de personas que no habiéndose reconocido pertenecientes o descendientes de un pueblo indígena u originario, respondieran algún nombre de pueblo (cuadro 4.12.3.4).

**Cuadro 4.12.3.4 Pueblo indígena u originario. Error de secuencia. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Cantidad de preguntas con error	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Pueblo indígena u originario	0,3	0,1	0,4	0,1	0,4

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- El 26,1% de las personas a las que se les aplicó el cuestionario A y el 12,7% de aquellas a las que se les aplicó el cuestionario B respondieron ignorar la respuesta. En ambos casos, esta categoría muestra un valor aproximado al de aquellos que se reconocen pertenecientes o descendientes de un pueblo indígena u originario y supera ampliamente a los que se reconocen afrodescendientes (cuadro 4.12.3.5).

**Cuadro 4.12.3.5 Etnias (A: P17) (B: P10). Distribución de frecuencias por tipo de cuestionario. II PP 2019. En porcentajes**

Cuestionario	Reconocimiento étnico				
	Indígena	Afro	Ninguna	Ignorado	Error total
Ampliado	30,1	0,1	33,0	26,1	10,7
Básico	16,0	0,8	59,9	12,7	10,6

**Nota:** el error para el cuestionario A difiere del presentado en el cuadro 4 porque incorpora a Humahuaca (área en que no se aplicó el cuestionario B).

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.13 Número de documento nacional de identidad (DNI) (A: P18) (B: P11)

### 4.13.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	<p><b>18</b> ¿Cuál es el número de DNI (Documento Nacional de Identidad)?</p> <p>DNI N°: <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; margin-right: 10px;" type="text"/></p> <p>No tiene DNI argentino <input type="checkbox"/> 1            No contesta <input type="checkbox"/> 2            Ignorado <input type="checkbox"/> 9</p>	
Cuestionario básico	<p><b>11</b> ¿Cuál es el número de DNI (Documento Nacional de Identidad)?</p> <p>DNI N°: <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; margin-right: 10px;" type="text"/></p> <p>No tiene DNI argentino <input type="checkbox"/> 1            No contesta <input type="checkbox"/> 2            Ignorado <input type="checkbox"/> 9</p>	

### 4.13.2 Objetivo y principales características de las preguntas

- La inclusión de la pregunta sobre Número de DNI se enmarca en el proyecto del Censo Nacional de Población, Hogares y Viviendas de la ronda 2020 que busca concretar la integración de información relevada en distintas fuentes de datos (censo, encuestas a hogares, registros administrativos) en un único sistema estadístico. En consecuencia, el objetivo de la inclusión de esta variable en el cuestionario de población es proporcionar una clave de identificación única que permita integrar la información censal con la que proviene de otras fuentes de datos.
- Esta pregunta se incluye en ambos cuestionarios de la II PP.

## 4.13.3 Resultados

- La pregunta sobre Número de DNI es la primera vez que se incorpora en un censo de población en la Argentina.
- Se observa un nivel de error más elevado en áreas urbanas de mayor tamaño (CABA y Mar del Plata) que en el resto de las áreas de prueba (cuadro 4.13.3.1).

**Cuadro 4.13.3.1 Número de DNI (A: P18) (B: P11). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
DNI	6,2	10,2	9,7	5,6	4,4

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- No se observan diferencias en el nivel de error entre los cuestionarios B y A (cuadro 4.13.3.2).

**Cuadro 4.13.3.2 Número de DNI (A: P18) (B: P11). Error total. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
DNI	7,8	8,1

Nota: las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- La no respuesta constituye la mayor fuente de error de la variable, aunque en CABA también se observa un nivel elevado de multimarca. Los casos inválidos son una proporción menor en todas las áreas de prueba (cuadro 4.13.3.3).

**Cuadro 4.13.3.3 DNI (A: P18) (B: P11). Tipo de error. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Tipo de error	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Error total	6,2	10,2	9,7	5,6	4,4
No respuesta	4,1	5,3	6,8	3,7	3,9
Multimarca	1,8	4,9	2,6	1,5	0,5
Inválido	0,3	0,0	0,3	0,3	0,1

Fuente: INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.14 Religión (A: P19)

### 4.14.1 Diseño de preguntas en el cuestionario

**Cuestionario ampliado**

19 ¿Cuál es su religión?

Católica	<input type="checkbox"/> 1
Evangélica	<input type="checkbox"/> 2
Judía	<input type="checkbox"/> 3
Islam	<input type="checkbox"/> 4
Testigos de Jehová	<input type="checkbox"/> 5
Iglesia de Jesucristo de los Santos de los Últimos Días	<input type="checkbox"/> 6

### 4.14.2 Objetivo y principales características de las preguntas

- Su propósito es identificar la pertenencia religiosa de la población.
- La pregunta sobre Religión forma parte del cuestionario A y está dirigida al total de mujeres y varones de 14 años y más.

### 4.14.3 Resultados

- La pregunta no fue indagada en los censos 1991, 2001 y 2010 y tiene un error total de 5,5%. CABA es la jurisdicción que presenta el nivel más bajo de error (cuadro 4.14.3.1).

**Cuadro 4.14.3.1 Religión (A: P19). Error total. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Religión	5,5	3,9	5,6	4,6	7,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- El 22,2% de la población de 0 a 13 años, que no debía responder esta pregunta, tiene respuesta en religión, es decir que presenta un elevado nivel de error de secuencia. Esto ocurrió con mayor frecuencia en CABA (cuadro 4.14.3.2).

**Cuadro 4.14.3.2 Religión (A: P19). Error de secuencia. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (1)
Religión	22,2	32,2	21,2	23,0	14,3

(1) El error de secuencia de San Carlos de Bariloche debe ser tomado con recaudos porque la población de 0 a 13 años es inferior a 10.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.15 Situación conyugal (A: P20)

### 4.15.1 Diseño de preguntas en el cuestionario

**Cuestionario ampliado**

**20** ¿Está...

unida(o) de hecho?  1

casada(o) legalmente?  2

en unión civil o unión convivencial?  3

separada(o) de hecho?  4

divorciada(o) legalmente?  5

viuda(o) de unión o matrimonio?  6

soltera(o) nunca unida(o)?  7

### 4.15.2 Objetivo y principales características de las preguntas

- El objetivo de la pregunta es conocer cuál es la situación de la población en relación con las uniones legales y consensuales.
- Esta pregunta está incluida solo en el cuestionario A y se indagó en las mujeres y varones de 14 años y más.

### 4.15.3 Resultados

- El error es superior al de los censos anteriores (cuadro 4.15.3.1).

**Cuadro 4.15.3.1 Situación conyugal (A: P19). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentaje**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Situación conyugal	1,6	3,5 (Convivencia en pareja)	2,2 (Convivencia en pareja)	6,1	Se agrupan las preguntas sobre estado civil y convivencia en pareja.
		1,5 (Estado civil)	1,4 (Estado civil)		

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- El error total alcanza un 6,1% de la población censada, siendo CABA la jurisdicción con el error más bajo para esta pregunta y San Carlos de Bariloche el área de prueba con el error más elevado (cuadro 4.15.3.2).

**Cuadro 4.15.3.2 Situación conyugal (A: P19). Error total. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Situación conyugal	6,1	4,1	6,3	5,5	7,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Esta pregunta fue indagada por el censista al 18,3% de las personas de 0 a 13 años que no debían responderla, situación que se presentó con mayor frecuencia en CABA (cuadro 4.15.3.3).

**Cuadro 4.15.3.3 Situación conyugal (A: P19). Error de secuencia. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (%)
Situación conyugal	18,3	31,7	17,0	20,8	0,0

(1) El error de secuencia de San Carlos de Bariloche debe ser tomado con recaudos porque la población de 0 a 13 años es inferior a 10.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.16 Actividad económica (A: P21 a P33) (B: P12 a P16)

### 4.16.1 Diseño de preguntas en el cuestionario

Cuestionario ampliado	Cuestionario básico
<p><b>21</b> Durante la semana pasada, ¿trabajó por lo menos una hora? (sin contar las tareas de su hogar)</p> <p>Sí <input type="checkbox"/> 1 → 25 No <input type="checkbox"/> 2</p>	<p><b>12</b> Durante la semana pasada, ¿trabajó por lo menos una hora? (sin contar las tareas de su hogar)</p> <p>Sí <input type="checkbox"/> 1 → 16 No <input type="checkbox"/> 2</p>
<p><b>22</b> En esa semana, ¿hizo alguna changa, fabricó algo para vender afuera, ayudó a un familiar o amigo en su chacra o negocio?</p> <p>Sí <input type="checkbox"/> 1 → 25 No <input type="checkbox"/> 2</p>	<p><b>13</b> En esa semana, ¿hizo alguna changa, fabricó algo para vender afuera, ayudó a un familiar o amigo en su chacra o negocio?</p> <p>Sí <input type="checkbox"/> 1 → 16 No <input type="checkbox"/> 2</p>
<p><b>23</b> En esa semana, ¿tenía trabajo, pero estuvo de licencia por vacaciones o enfermedad, suspensión con pago, conflicto laboral, etc.?</p> <p>Sí <input type="checkbox"/> 1 → 25 No <input type="checkbox"/> 2</p>	<p><b>14</b> En esa semana, ¿tenía trabajo, pero estuvo de licencia por vacaciones o enfermedad, suspensión con pago, conflicto laboral, etc.?</p> <p>Sí <input type="checkbox"/> 1 → 16 No <input type="checkbox"/> 2</p>
<p><b>24</b> Durante las últimas 4 semanas, ¿buscó trabajo: contestó avisos, consultó amigos o parientes, puso carteles, hizo algo para ponerse por su cuenta?</p> <p>Incluye la búsqueda en forma personal, por medios de comunicación, por internet, etc.</p> <p>Sí <input type="checkbox"/> 1 → 34 No <input type="checkbox"/> 2</p>	<p><b>15</b> Durante las últimas 4 semanas, ¿buscó trabajo: contestó avisos, consultó amigos o parientes, puso carteles, hizo algo para ponerse por su cuenta?</p> <p>Incluye la búsqueda en forma personal, por medios de comunicación, por internet, etc.</p> <p>Sí <input type="checkbox"/> 1 → 16 No <input type="checkbox"/> 2 → Fin de la entrevista para esta persona</p>

## Cuestionario ampliado

## Cuestionario básico

**25** ¿Trabajó por un pago en dinero o en especie?

- Sí
- No
- Ignorado

## 26 El trabajo principal, en el que trabaja más horas...

¿es una actividad o empresa propia que emplea personas asalariadas?	<input type="checkbox"/>	1
¿es una actividad o empresa propia que no emplea personas asalariadas?	<input type="checkbox"/>	2
¿lo hace como servicio doméstico?	<input type="checkbox"/>	3
como obrero o empleado para un patrón, empresa o institución? (incluye agencia de empleo)	<input type="checkbox"/>	4
o hace para el negocio o actividad de un familiar?	<input type="checkbox"/>	5

**27** En ese trabajo, ¿le descuentan para la jubilación?

Sí	<input type="checkbox"/>	1	→ 29
No	<input type="checkbox"/>	2	
Ignorado	<input type="checkbox"/>	9	

**28** En ese trabajo, ¿aporta por sí mismo para la jubilación?

Sí  No  Ignorado

29 ¿Cuál es el nombre de la ocupación que realiza en ese trabajo? Ej.: quintero, profesora, jefe de mecánicos, vendedora, chofer, peón de campo, secretaria, etc.

30 ¿Qué tareas realiza? Ej.: manejo el camión de la fábrica, superviso a los mecánicos y arreglo maquinarias, vendo electrodomésticos, siembro y cosecho verduras, doy de comer a los animales y limpio los corrales, etc.

A large grid of 30 empty blue-outlined boxes for writing responses.

31) ¿A qué se dedica o qué produce la actividad, negocio, empresa o institución en la que trabajas? Ej.: fábrica de ropa, fábrica de tractores, clínica, cría de ovejas, venta de electrodomésticos, reparación de calzado, tintorería, alquiler de maquinarias, cultivo de verduras, servicios de fumigación aérea, etc.

**32** ¿Cuál es la dirección de ese lugar de trabajo?

<p>Esta vivienda</p> <p>No tiene lugar fijo de trabajo</p> <p>Otro país</p> <p>Otro lugar</p>	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4	<div style="border-left: 1px solid black; padding-left: 5px; margin-left: 10px;"> <p>33</p> <p>Anote dirección, localidad, municipio y provincia</p> </div>
---	--	---

Dirección:

Localidad:

A horizontal row of 20 empty square boxes, arranged in two rows of 10. The boxes are outlined in blue and are intended for children to draw on.

## Municipio

Provincia:

A horizontal row of 50 empty blue-bordered boxes, arranged in 5 rows and 10 columns, intended for handwritten responses.

**16** ¿Trabajó por un pago en dinero o en especie?

Sí  1  
No  2  
orado  9

Fin de la entrevista para esta persona

## Cuestionario ampliado

33 ¿Cuántas personas, incluyendo a [nombre] trabajan allí en total?

1 persona	<input type="checkbox"/> 1
De 2 a 5 personas	<input type="checkbox"/> 2
De 6 a 25 personas	<input type="checkbox"/> 3
De 26 a 40 personas	<input type="checkbox"/> 4
De 41 a 100 personas	<input type="checkbox"/> 5
Más de 100 personas	<input type="checkbox"/> 6
Ignorado	<input type="checkbox"/> 9

### 4.16.2 Objetivo y principales características de las preguntas

- Las preguntas relativas a las características económicas de las personas son de especial relevancia en los censos de población ya que es necesario contar con información acerca de la condición de actividad y características de esta, junto a otras características demográficas y sociales. Esto permite disponer de una “panorámica global de la situación económica” (NU, 2010: p. 163) de un territorio.
- La medición de la Condición de actividad tiene como principal objetivo determinar la situación laboral de una persona, en un breve período de referencia en relación con su condición de económicamente activa, empleada o desempleada o económicamente inactiva.
- La operacionalización de la Condición de actividad de la II PP se encuentra compuesta por una pregunta dirigida a conocer la situación laboral o de empleo de manera básica y otras dos preguntas complementarias para rescatar a aquellos trabajadores que no son fáciles de empadronar y, por último, una cuarta que determina la condición de desempleo o desocupación o de inactividad de las personas.
- Se incluye la pregunta que indaga sobre el trabajo pago en dinero o en especie, consistentemente con la 19° Conferencia Internacional de Estadísticos del Trabajo (CIET, 2013),<sup>27</sup> que reemplazó a las Recomendaciones Internacionales de 1982, adoptando un marco conceptual consistente con el Sistema de Cuentas Nacionales.
- La variable Categoría ocupacional tiene como objetivo conocer la relación existente entre la persona económicamente activa y su empleo y es de suma importancia para comprender la estructura y el funcionamiento de cualquier mercado laboral. Esta relación se encuentra dada a partir del tipo de contrato de trabajo implícito o explícito, la solidez de la relación entre la persona y el empleo, y el tipo de autoridad que la persona tiene sobre los establecimientos y otros trabajadores. Asimismo, la información que otorga esta variable, dependiendo de la definición que se tome acerca del trabajo informal, permite establecer la pertenencia de la persona a este grupo, mediante su pertenencia a ciertas categorías asociadas al trabajo informal.
- Las variables Descuento y Aporte jubilatorio son indicadores del empleo informal en asalariados, dado a partir de la ausencia de protección social.

<sup>27</sup> 19° Conferencia Internacional de Estadísticos del Trabajo, llevada a cabo entre el 2 y el 11 de octubre de 2013, en Ginebra (Suiza). Se introdujeron cambios importantes para medir las diferentes formas de trabajo, incluyendo definiciones de empleo como trabajo pago o en especie, para expandir la medición de la subutilización del trabajo más allá del desempleo tradicional. Esta conferencia sugiere que, para no romper la serie histórica de las mediciones de población económica activa, empleo, desempleo y subempleo, la participación de las siguientes actividades debe medirse de manera separada: producción de bienes para uso propio de los hogares, trabajo no pago realizado por aprendices, trabajo voluntario y trabajo voluntario para la producción de bienes para otros hogares. Asimismo, la CIET recomienda que, transicionalmente, se mantengan ambas mediciones en los sistemas estadísticos nacionales.

- La variable Ocupación se compone de dos preguntas (Nombre de la ocupación y Tarea que realiza). Tiene como objetivo conocer el tipo de trabajo que la persona efectúa con independencia de dónde y en qué circunstancias.
- La pregunta sobre Rama de actividad tiene el objetivo de conocer el tipo de producción o la actividad del establecimiento o unidad en que se emplea la persona, describe lo que hace el establecimiento y no lo que hace específicamente la persona que trabaja. La información que surge de esta pregunta es esencial para contar con estadísticas completas y detalladas de todo el empleo en un país desglosado por esta variable y aporta información estructural a las cuentas nacionales.
- La pregunta del Lugar de trabajo hace referencia a la naturaleza del lugar de trabajo distinguiéndolo entre la casa y otros lugares de trabajo, y si no tiene lugar fijo de trabajo. Para aquellos que cuentan con una dirección fija de trabajo, se recoge la dirección de este con vistas a la posibilidad de relacionar esta dirección con registros administrativos. Asimismo, permite realizar tabulaciones y cartografiar el lugar de trabajo según el lugar de residencia. Por último, permite, a partir de la primera distinción, realizar una aproximación a la identificación del sector informal.
- La pregunta sobre Tamaño del establecimiento se relaciona con la identificación del sector informal ya que un indicador de esta característica está dado a partir de establecimientos de un número pequeño de empleados. Si bien en un censo no es posible identificar de forma precisa a las personas que trabajan en el sector informal debido al limitado número de preguntas que pueden incluirse en esa fuente de datos y debido a que implica la respuesta de terceros, debe ser posible obtener algunos indicadores seleccionados, uno de ellos es el Tamaño del establecimiento.
- Cabe mencionar que este bloque de preguntas fue producto de un sostenido intercambio con el área económica del INDEC, de manera que el diseño conceptual del censo de población aportara a las necesidades de información económica requerida por el área.
- Las preguntas para determinar la Condición de actividad económica y Percepción de pago se incluyeron en ambos cuestionarios, mientras que en el cuestionario A se incorporaron las relativas a Categoría ocupacional, Descuento y Aporte jubilatorio, Rama de actividad económica, Ocupación, Dirección del lugar de trabajo y Tamaño del establecimiento.

#### 4.16.3 Resultados<sup>28</sup>

- Todas las preguntas referidas a las características laborales muestran errores superiores a los de censos anteriores.
- La pregunta relativa al Domicilio del lugar de trabajo, que se incorpora en esta prueba piloto, presentó los valores más altos, seguida por las variables Rama de actividad económica, Tarea que realiza, Tamaño del establecimiento y Categoría ocupacional (cuadro 4.16.3.1).

<sup>28</sup> En este apartado solo se evaluará la no respuesta (NR) y el error de secuencia.

**Cuadro 4.16.3.1 Actividad económica (A: P21 a P33) (B: P12 a P16). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentaje**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Trabajó	4,4	3,3	4,0	7,2	La pregunta no presenta modificaciones.
Hizo alguna changa	4,4	1,4	4,4	7,9	La pregunta no presenta modificaciones sustanciales, solo se modifica una “/” por la letra “o”.
Licencia	3,2	1,5	2,1	5,9	Se aclara que la suspensión es con pago.
Buscó trabajo	2,1	1,1	1,2	6,6	Se incorpora una aclaración que da cuenta de los medios por los cuales se puede buscar trabajo.
Percepción de pago	No se preguntó.			6,7	Se incorpora la pregunta.
Categoría ocupacional	3,5	3,6	4,8	10,1	<p>Se modifica el orden en el que aparece esta pregunta, apareciendo primera luego de condición de actividad y percepción de pago, y antes de las preguntas abiertas de ocupación y rama de actividad.</p> <p>Se modifica la formulación de la pregunta y las categorías de respuesta. En el Censo 2010, las categorías se autodefinen. En la II PP, se describen las características de cada una de las categorías ocupacionales. Se incorpora por separado la categoría: servicio doméstico.</p> <p>En 2010, se preguntaba: “En ese trabajo, ¿es...” en la II PP “El trabajo principal, en el que trabaja más horas...”</p>
Descuentos jubilatorios	4,6	2,0	2,3	3,8	La pregunta no presenta modificaciones.
Aportes jubilatorios	No se preguntó.	6,3	9,0	8,0	La pregunta no presenta modificaciones.
Ocupación	No se preguntó.	s/d	5,7	9,0	<p>En el Censo 2010, se preguntaba: “¿Cuál es el nombre de la ocupación?”, actualmente: “¿Cuál es el nombre de la ocupación que realiza en ese trabajo?”. Además, se incorpora un grisado con una lista de ejemplos de ocupaciones.</p> <p>Se modifica la posición dentro del bloque, indagándose antes que la rama de actividad.</p>
Tarea	s/d	s/d	No se preguntó.	11,1	Se incorpora la pregunta.

(continúa)

**Cuadro 4.16.3.1 (conclusión)**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Rama de actividad económica	s/d	s/d	6,8	20,6	Cambia su posición dentro del bloque. Se ubica luego de las preguntas que miden la variable de ocupación. La pregunta sigue siendo abierta, pero cambia su formulación. En el Censo 2010, se preguntaba: "¿A qué se dedica o que servicio presta la empresa o lugar donde trabaja más horas?". En la II PP, "¿A qué se dedica o qué produce la actividad, negocio, empresa o institución en la que trabaja?". Además, se incorpora un grisado con una lista de ejemplos de rama de actividad.
Dirección del lugar de trabajo precodificada	No se preguntó.			23,6	Se incorpora la pregunta.
Dirección				20,9	
Número				63,9	
Localidad				6,8	
Municipio				21,9	
Provincia				9,0	
Tamaño del establecimiento	8,8	6,8	9,0	11,4	En 2010 "¿Cuántas personas hay en total en la empresa o lugar donde trabaja?" En II PP "¿Cuántas personas incluyendo a [nombre] trabajan allí en total?" Se pasa de 4 a 7 categorías, desagregando más intervalos e incluyendo el ignorado. Se modifica la formulación.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- Según se observa en el cuadro 4.16.3.2, las preguntas referidas a Condición de actividad y Categoría ocupacional registran menores niveles de error en San Carlos de Bariloche y en CABA. En este bloque de preguntas, no se destaca ningún área de prueba con errores específicamente bajos.

**Cuadro 4.16.3.2 Actividad económica (A: P21 a P33) (B: P12 a P16). Error total. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (1)
Trabajó	7,2	5,3	7,9	5,4	6,0
Hizo alguna changa	7,9	3,4	8,5	9,9	4,3
Licencia	5,9	3,8	6,4	6,6	3,0
Buscó trabajo	6,6	5,3	7,0	8,5	2,2

(continúa)

**Cuadro 4.16.3.2 (conclusión)**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (1)
Percepción de pago	6,7	1,0	8,0	4,1	5,0
Categoría ocupacional	10,1	6,8	10,7	7,4	5,3
Descuentos jubilatorios	3,8	3,9	3,8	3,8	0,0
Aportes jubilatorios	8,0	5,5	8,5	3,1	18,5
Ocupación	9,0	5,8	9,8	3,3	7,9
Tarea	11,1	9,9	11,6	7,4	5,3
Rama de actividad	20,6	23,6	20,9	15,6	10,5
Dirección del lugar de trabajo precodificada	23,6	26,4	23,0	25,6	47,4
Dirección	20,9	25,4	20,5	21,7	1/9
Número	63,9	73,7	63,4	65,1	2/9
Localidad	6,8	28,1	4,6	12,0	2/9
Municipio	21,9	76,3	17,6	18,1	3/9
Provincia	9,0	74,6	3,0	16,9	3/9
Tamaño del establecimiento	11,4	9,6	11,9	8,5	2/38

(1) En determinadas variables se presentan en absolutos dado que, en San Carlos de Bariloche, el denominador es inferior a 40.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Cuando se comparan las preguntas que están incluidas en ambos cuestionarios, no se observa un comportamiento sistemático, y mientras que la pregunta sobre si Trabajó la semana pasada, si estaba de Licencia y si Buscó trabajo tienen error más elevado en el cuestionario A, las que indagan por si Hizo alguna changa y Percepción de pago lo muestran en el B (cuadro 4.16.3.3).

**Cuadro 4.16.3.3 Error total en las preguntas de la variable condición de actividad. Cuestionarios básico y ampliado para áreas de prueba con muestra. II PP 2019. En porcentajes**

Pregunta	Error total	
	Cuestionario básico	Cuestionario ampliado
Trabajó	5,5	5,8
Hizo alguna changa	8,6	3,8
Licencia	3,2	5,9
Buscó trabajo	5,0	6,0
Percepción de pago	3,8	3,3

**Nota:** las áreas de prueba con muestra fueron CABA, Mar del Plata y San Carlos de Bariloche.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- De acuerdo con el cuadro 4.16.3.4, el 5,5% de la población de 14 años o más no puede ser clasificada debido a los errores en alguna de las cuatro preguntas que conforman la variable Condición de actividad, siendo el error más alto en Humahuaca y San Carlos de Bariloche, y más bajo en CABA (cuadro 4.16.3.4).

**Cuadro 4.16.3.4 Error total en las preguntas que conforman la variable condición de actividad. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Condición de actividad	5,5	3,1	6,0	3,5	5,5

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- El cuadro 4.16.3.5 muestra que los censistas formularon la primera pregunta de Condición de actividad al 14,6% de la población menor de 14 años. Asimismo, se observa que se presentaron dificultades en los pasos de las preguntas 22, 23 y 24 de Condición de actividad y en la pregunta 25 referida a la Percepción de pago, ya que entre un 16,0% y un 12,0% de los cuestionarios presentan marca en preguntas que no debían ser respondidas. La variable sobre Descuento y Aporte jubilatorio también ostentan errores de secuencia relevantes.
- San Carlos de Bariloche es el área de prueba que registra menores niveles de error de secuencia, a excepción de las preguntas sobre Dirección del lugar de trabajo.
- El error de secuencia para el resto del bloque, es decir, la presencia de respuesta por parte de personas desocupadas e inactivas se observa con muy baja frecuencia.

**Cuadro 4.16.3.5 Actividad económica (A: P21 a P33) (B: P12 a P16). Error de secuencia. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Trabajó	14,6	20,0	14,7	17,0	5,8
Hizo alguna changa	16,0	20,1	16,8	15,8	6,8
Licencia	13,9	17,7	14,5	13,9	6,4
Buscó trabajo	12,9	17,1	13,5	12,0	6,0
Percepción de pago	12,6	15,3	12,6	17,4	5,0
Categoría ocupacional	2,8	1,7	3,0	1,8	3,8
Descuentos jubilatorios	10,1	11,4	10,0	8,9	19,6
Aportes jubilatorios	8,7	8,4	8,8	7,9	4,9
Ocupación	3,4	2,4	3,5	2,2	3,8
Tarea	3,2	2,2	3,4	2,4	3,8
Rama de actividad	2,5	1,2	2,7	1,1	3,8
Dirección del lugar de trabajo precodificada	2,4	1,7	2,5	1,5	0,0
Dirección	4,1	2,9	4,2	3,3	10,8
Número	2,3	2,1	2,3	1,9	10,8
Localidad	4,9	2,1	5,1	4,6	10,8
Municipio	3,8	1,0	4,0	4,1	10,8
Provincia	4,8	1,0	5,1	3,9	10,8
Tamaño del establecimiento	3,3	3,1	3,3	2,6	3,8

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 4.17 Fecundidad (A: P34 a P36)

### 4.17.1 Diseño de preguntas en el cuestionario

**Cuestionario ampliado**

**A partir de aquí contestan todas las mujeres de 14 años y más**

**34** ¿Cuántos hijas e hijos nacidos vivos tuvo?

No tuvo hijas o hijos nacidos vivos  1 → Fin de la entrevista para esta persona

Cantidad de hijas e hijos nacidos vivos:

Cantidad de hijas e hijos que están vivos actualmente:

**35** ¿En qué año tuvo la primera hija o hijo nacido vivo?

Año:

**36** ¿Cuál es la fecha de nacimiento de la última hija o hijo nacido vivo?

Mes:  Año:

Fin de la entrevista para esta persona

### 4.17.2 Objetivo y principales características de las preguntas

- La inclusión de las preguntas sobre Fecundidad y Mortalidad infantil en los censos de población posibilita no solo estimar el nivel de la mortalidad y la fecundidad, sino en particular vincular su comportamiento con dimensiones de índole social y económica.
- La Cantidad de hijas e hijos nacidos vivos (HNV) permite calcular el porcentaje de mujeres de 14 años y más que han tenido hijos al momento del censo, así como también el porcentaje de mujeres nulíparas. Es fundamental para el cálculo de los indicadores de fecundidad retrospectiva y actual y es insumo para calcular la paridez media a una edad determinada y para estimar la fecundidad actual a través de métodos indirectos.
- La Cantidad de hijas e hijos que están vivos actualmente (HVA) aporta al cálculo de la mortalidad infantil a través de métodos indirectos.
- Conocer en qué año tuvo la primera o el primer HNV constituye el insumo para la elaboración de las tasas de fecundidad adolescente y de la edad de la mujer al primer hijo.
- La información sobre la Fecha de nacimiento de la última o el último HNV se utiliza para calcular, junto con la cantidad de HNV, la tasa global de fecundidad por métodos indirectos y las tasas de fecundidad por edad.
- Estas preguntas se incluyen solo en el cuestionario A.

### 4.17.3 Resultados

- Todas las preguntas del bloque muestran errores superiores a los de censos anteriores, en especial a los últimos dos (cuadro 4.17.3.1).

**Cuadro 4.17.3.1 Fecundidad (A: P34 a P36). Error total. Censos 1991, 2001, 2010 y II PP 2019, y modificaciones respecto del Censo 2010. En porcentajes**

Pregunta	Error total del país en censos			Error total II PP. Cuestionarios ampliado y básico	Modificaciones respecto del Censo 2010
	1991	2001	2010		
Tenencia y cantidad de hijas o hijos nacidos vivos	No se preguntó. 6,2	4,3 0,5	4,0 0,4	10,3	Se agrupan bajo el mismo número de pregunta.
Cantidad de hijas e hijos vivos actualmente	1,5	0,4	0,4		
Año de nacimiento de la primera hija o hijo nacido vivo	No se preguntó.			13,9	Se incorpora la pregunta.
Fecha de nacimiento de su última hija o hijo nacido vivo Mes	13,0	4,6	3,3	12,0	La pregunta no presenta modificaciones.
Fecha de nacimiento de su última hija o hijo nacido vivo Año		4,6	1,8	13,2	La pregunta no presenta modificaciones.

**Fuente:** Giusti (1997), archivos sin consistir del Censo 2001, Censo 2010 y base de datos sin consistir de la II PP 2019.

- A excepción del error de la pregunta Cantidad de hijas e hijos vivos actualmente registrada en CABA y el Año de nacimiento de la primera o el primer HNV en Mar del Plata, no se encuentran diferencias relevantes entre las áreas de prueba (cuadro 4.17.3.2).

**Cuadro 4.17.3.2 Fecundidad (A: P34 a P36). Error total. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentaje**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche (¹)
Tenencia y cantidad de hijas o hijos nacidos vivos (²)	10,3	7,2	10,7	7,7	2/30
Cantidad de hijas e hijos vivos actualmente	8,2	20,0	7,3	7,7	3/21
Año de nacimiento de la primera hija o hijo nacido vivo	13,9	10,9	13,9	17,5	2/21
Fecha de nacimiento de su última hija o hijo nacido vivo Mes	12,0	8,5	11,9	13,7	6/21
Fecha de nacimiento de su última hija o hijo nacido vivo Año	13,2	11,5	13,3	14,2	2/21

(¹) El error total de San Carlos de Bariloche se presenta en absolutos porque las mujeres de 14 años y más que tuvieron al menos 1 hijo son 21.

(²) Debido a la operacionalización de la pregunta, se consideró no respuesta cuando ambas preguntas estaban en blanco. Si había marca en la categoría no tuvo HNV y, al mismo tiempo, un valor superior a cero en la cantidad de HNV se consideró la cantidad por sobre la marca.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- En el cuadro 4.17.3.3 se observa que los censistas continuaron indagando por la Tenencia y Cantidad de hijos nacidos vivos al 6,2% de la población (personas menores de 14 años y varones). El resto de los errores de secuencia para todas las preguntas es cercano al 5%, lo que muestra una disminución en las preguntas que solo debían ser respondidas para las mujeres que tuvieron HNV. En cuanto a la diferencia entre las áreas de prueba, San Carlos de Bariloche se destaca con errores de secuencia que van del 11,8% al 28,1%.

**Cuadro 4.17.3.3 Fecundidad (A: P34 a P36). Error de secuencia. Cuestionario ampliado. Total y áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Total	CABA	Humahuaca	Mar del Plata	San Carlos de Bariloche
Tenencia y cantidad de hijas o hijos nacidos vivos	6,2	12,6	5,9	2,4	11,8
Cantidad de hijas e hijos vivos actualmente	5,0	9,4	4,5	5,1	28,1
Año de nacimiento de la primera hija o hijo nacido vivo	4,5	9,6	3,9	4,1	22,0
Mes de nacimiento de su última hija o hijo nacido vivo	4,5	8,6	4,1	3,5	19,5
Año de nacimiento de su última hija o hijo nacido vivo	4,6	8,8	4,2	3,3	22,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

- Al analizar la composición del error de secuencia se puede observar que, para todas las preguntas de fecundidad, el mayor porcentaje de error de secuencia lo aporta el haber continuado realizando esas preguntas a los varones y, en menor medida, se detecta similar error si se tiene en cuenta la edad de la persona censada.

**Cuadro 4.17.3.4 Fecundidad (A: P34 a P36). Error de secuencia según universo. Cuestionario ampliado. Total de áreas de prueba. II PP 2019. En porcentajes**

Pregunta	Error de secuencia total	Varones	Mujeres 0 a 13 años	Mujeres de 14 años y más sin HNV
Tenencia y cantidad de hijas o hijos nacidos vivos	6,2	4,2	2,0	///
Cantidad de hijas e hijos vivos actualmente	5,0	4,0	0,3	0,7
Año de nacimiento de la primera hija o hijo nacido vivo	4,5	3,9	0,3	0,3
Fecha de nacimiento de su última hija o hijo nacido vivo Mes	4,5	3,8	0,4	0,3
Fecha de nacimiento de su última hija o hijo nacido vivo Año	4,6	3,9	0,4	0,3

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

# Evaluación de calidad de la variable Número de DNI

## 5.1

### Introducción

En este documento se realiza un análisis de la variable Número de DNI con el objetivo general de aportar elementos sobre la pertinencia de su inclusión en el cuestionario censal. Se tendrán en cuenta diversos aspectos de la calidad de la declaración y de la lectura e interpretación del dato en términos de su efecto en la validez de la información para ser pareada con registros administrativos, objetivo primordial de su inclusión, y se realizará una breve comparación y caracterización de ambos subconjuntos. Este apartado constituye un primer abordaje sobre esta variable, novedosa para los relevamientos censales argentinos.

Sin embargo, es preciso realizar algunas aclaraciones respecto al alcance del informe. Este es un análisis de tipo exploratorio y las descripciones realizadas se refieren a lo observado en la II PP. Ellas se llevan a cabo con un conocimiento limitado respecto a la forma de asignación del Número de DNI en la Argentina. En consecuencia, se considera necesario profundizar en la relación con los organismos sectoriales responsables de este registro para poder mejorar la comprensión y enriquecer el análisis de la validez de la información. En ese sentido, como ya se mencionó, en ningún caso se puede determinar la validez del Número de DNI –en términos de exactitud de la declaración y correspondencia con un número otorgado por el Registro Nacional de las Personas (RENAPER), organismo en la órbita del Ministerio del Interior que realiza el registro e identificación de todas las personas físicas que se domicilien en el territorio argentino–. Para ello se requiere parear<sup>29</sup> el número de DNI relevado en la prueba respecto del que figura en los registros administrativos. Debido a que la información es altamente sensible, se la presenta solo en forma agregada. Cuando el análisis exige identificar algún número, por ejemplo, para comparar la base con la imagen del cuestionario, se utilizaron mecanismos para impedir que se pueda reconocer el número de DNI. Esto implica que el lector no podrá reproducir el análisis, pero garantiza la protección de los datos personales involucrados. Por último, esta información no puede ser extrapolada al total del país o a totales provinciales, debido principalmente a que el operativo no fue diseñado metodológicamente para la expansión muestral de los atributos aquí tratados.

## 5.2

### Evaluación de calidad

La pregunta sobre Número de DNI (P11 del cuestionario B y P18 del cuestionario A) consta de dos partes: una abierta, donde debía consignarse el número del documento; y otra cerrada, donde se establecían los motivos de no respuesta a dicha pregunta.

<p>18. ¿Cuál es el número de DNI (Documento Nacional de Identidad)?</p> <p>DNI N°: <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px; margin-bottom: 5px;" type="text"/></p>	<p>No tiene DNI argentino <input type="checkbox"/> 1</p> <p>No contesta <input type="checkbox"/> 2</p> <p>Ignorado <input type="checkbox"/> 9</p>
---	---

<sup>29</sup> Nos referimos aquí a “pareo” como el cotejo único de números de DNI entre dos conjuntos de datos. Adicionalmente existen metodologías donde se consideran conjuntos de atributos para vincular individuos, buscando su similitud, estrategia usual ante posibles problemas de calidad en la información.

Para el análisis de la calidad y validez de esta variable, se consideró un dato en blanco o **no respuesta** cuando no se registró un número de DNI ni se marcó el motivo por el cual no había registro de este, es decir, la pregunta estaba completamente vacía. Se consideró **multimarca** cuando se había registrado un número de DNI y simultáneamente al menos una respuesta en las opciones precodificadas o cuando había más de una respuesta en las opciones precodificadas. Asimismo, el número de DNI se consideró **válido** para las respuestas que tuvieran entre 6 y 9 dígitos.

El cuadro 5.2.1 presenta la distribución de la pregunta de Número de DNI para el total de áreas de prueba y para cada una de ellas. Como se observa, la mayoría de las respuestas corresponden a un número de DNI válido (78,8%), siendo ligeramente superior en Humahuaca (82,5) y más bajo en CABA y Mar del Plata (62,2% para esta última área de prueba). Un 9,7% indicó que ignora la información, un 5,2% decidió no responder y un 0,2% no tiene DNI argentino. Solo 6,2% corresponden a errores.

En relación con las diferencias entre las áreas de prueba, se observa que Mar del Plata presenta los porcentajes más elevados de no respuesta (6,8%) e ignorado (20,4%), mientras que, en San Carlos de Bariloche, se registra el porcentaje más alto de quienes no quisieron responder (9,5%).

**Cuadro 5.2.1 Distribución de la pregunta de Número de DNI. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Tenencia y número de DNI	Total	CABA	Mar del Plata	Humahuaca	San Carlos de Bariloche
No respuesta	4,1% (486)	5,3% (46)	6,8% (76)	3,7% (321)	3,9% (43)
Multimarca	1,8% (209)	4,9% (43)	2,6% (29)	1,5% (132)	0,5% (5)
No tiene DNI argentino	0,2% (20)	1,1% (10)	0% (0)	0,1% (9)	0,1% (1)
No contesta	5,2% (613)	8,6% (75)	7,6% (85)	4% (348)	9,5% (105)
Ignorado	9,7% (1.138)	10,4% (91)	20,4% (228)	7,8% (679)	12,6% (140)
DNI válido	78,8% (9.265)	69,6% (608)	62,2% (694)	82,5% (7150)	73,4% (813)
DNI inválido	0,3% (34)	0% (0)	0,3% (3)	0,3% (30)	0,1% (1)
Total	100% (11.765)	100% (873)	100% (1.115)	100% (8.669)	100% (1.108)

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

Con respecto a las diferencias entre tipos de cuestionarios, debe tenerse en cuenta que el número de personas que respondieron el cuestionario A fue 10.193 frente a 1.512 que respondieron el cuestionario B. En el cuadro 5.2.2 se observa que el cuestionario A presentó una mayor proporción de DNI válidos (80,2%) que el cuestionario B (69,4%).

Entre las personas que respondieron el cuestionario B, es superior la proporción que no responde esta información (10,0%) y la que la ignora (12,7%), en relación con las que responden el cuestionario A (4,5% y 9,2%, respectivamente).

**Cuadro 5.2.2 Distribución de la pregunta de Número de DNI según tipo de cuestionario. Total de áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Tenencia y número de DNI	Total	Cuestionario ampliado	Cuestionario básico
No respuesta	4,1% (486)	3,9% (402)	5,3% (84)
Multimarca	1,8% (209)	1,7% (171)	2,4% (38)

**(continúa)**

**Cuadro 5.2.2 (conclusión)**

Tenencia y número de DNI	Total	Cuestionario ampliado	Cuestionario básico
No tiene DNI argentino	0,2% (20)	0,2% (18)	0,1% (2)
No contesta	5,2% (613)	4,5% (456)	10,0% (157)
Ignorado	9,7% (1.138)	9,2% (939)	12,7% (199)
DNI válido	78,8% (9.265)	80,2% (8.174)	69,4% (1.091)
DNI inválido	0,3% (34)	0,3% (33)	0,1% (1)
Total	100% (11.765)	100% (10.193)	100% (1.572)

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

En el cuadro 5.2.3 se presenta la distribución de la pregunta Número de DNI según variables seleccionadas. Se destaca un mayor porcentaje de respuesta válida del DNI en mujeres, menor en el grupo 0-14 (73,2%), con un efecto respondente.<sup>30</sup> Las personas que no son autorrespondentes declaran ignorar (19,0%) o no responden el número de DNI (7,2%) en mayor magnitud que quienes son autorrespondentes. Ello se refleja también en un porcentaje mayor en jefas(es)-personas de referencia/cónyuges (que en general responden el cuestionario por todos los miembros del hogar) que declaran un número de DNI válido respecto de otras relaciones o parentesco. Por último, las personas nacidas en otro país (no nativas) declararon en menor medida que las nacidas en la Argentina número de DNI válido, 71,5% y 79,1%, respectivamente.<sup>31</sup>

**Cuadro 5.2.3 Distribución de pregunta sobre DNI según variables de corte seleccionadas. Cuestionarios básico y ampliado. Total y áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Variable		No respuesta	Multimarca	No tiene DNI argentino	No contesta	Ignorado	DNI válido	DNI inválido	Total
Total		4,1% (486)	1,8% (209)	0,2% (20)	5,2% (613)	9,7% (1.138)	78,8% (9.265)	0,3% (34)	11.765
Sexo	Mujer	4,1% (258)	1,6% (101)	0,1% (9)	4,6% (288)	8,3% (516)	81,0% (5.053)	0,2% (10)	6.235
	Varón	3,9% (208)	1,9% (104)	0,2% (11)	5,9% (315)	11,1% (592)	76,8% (4.112)	0,3% (15)	5.357
	No respuesta	11,6% (20)	2,3% (4)	0,0% (0)	5,8% (10)	17,3% (30)	57,8% (100)	5,2% (9)	173
Edad	0-14 años	6,5% (191)	1,4% (41)	0,2% (6)	6,3% (187)	12,1% (356)	73,2% (2156)	0,3% (9)	2.946
	15-64 años	2,8% (206)	1,9% (137)	0,1% (7)	4,7% (350)	8,8% (648)	81,6% (6.023)	0,2% (14)	7.385
	65 años y más	3,4% (38)	2,4% (27)	0,4% (5)	5,1% (58)	7,4% (83)	80,9% (912)	0,4% (5)	1.128
	No respuesta	16,7% (51)	1,3% (4)	0,7% (2)	5,9% (18)	16,7% (51)	56,9% (174)	2,0% (6)	306

**(continúa)**

<sup>30</sup> Esta variable presenta un 26,8% de no respuesta.

<sup>31</sup> La variable lugar de nacimiento fue consistida para realizar este análisis.

Cuadro 5.2.3 (conclusión)

Variable		No respuesta	Multimarca	No tiene DNI argentino	No contesta	Ignorado	DNI válido	DNI inválido	Total
Autorrespondiente	No	5,7% (241)	1,1% (48)	0,2% (8)	7,2% (309)	19,0% (812)	66,4% (2.832)	0,3% (13)	4.263
	Sí	1,1% (50)	1,1% (49)	0,2% (7)	3,2% (140)	2,3% (101)	91,8% (3.991)	0,2% (10)	4.348
	No respuesta	6,2% (195)	3,6% (112)	0,2% (5)	5,2% (164)	7,1% (225)	77,4% (2.442)	0,3% (11)	3.154
Relación o parentesco	Jefa(e)/cónyuge	2,0% (111)	2,3% (124)	0,1% (7)	4,1% (221)	5,7% (309)	85,7% (4.658)	0,1% (8)	5.438
	Hija(o)/madre, padre/suegra(o)	5,3% (255)	1,3% (64)	0,2% (8)	6,3% (301)	11,7% (562)	74,9% (3.598)	0,4% (18)	4.806
	Otros	7,4% (99)	1,2% (16)	0,4% (5)	6,3% (84)	18,0% (241)	66,4% (887)	0,3% (4)	1.336
	No respuesta	11,4% (21)	2,7% (5)	0,0% (0)	3,8% (7)	14,1% (26)	65,9% (122)	2,2% (4)	185
Lugar de nacimiento	Native	3,3% (251)	1,4% (106)	0,1% (8)	5,5% (414)	10,3% (774)	79,1% (5.955)	0,3% (20)	7.528
	No native	2,7% (15)	4,4% (25)	1,6% (9)	6,0% (34)	13,7% (77)	71,5% (402)	0,0% (0)	562
	No respuesta	6,0% (220)	2,1% (78)	0,1% (3)	4,5% (165)	7,8% (287)	79,1% (2.908)	0,4% (14)	3.675

Fuente: INDEC. Base de datos sin consistir. II PP 2019.

## 5.3

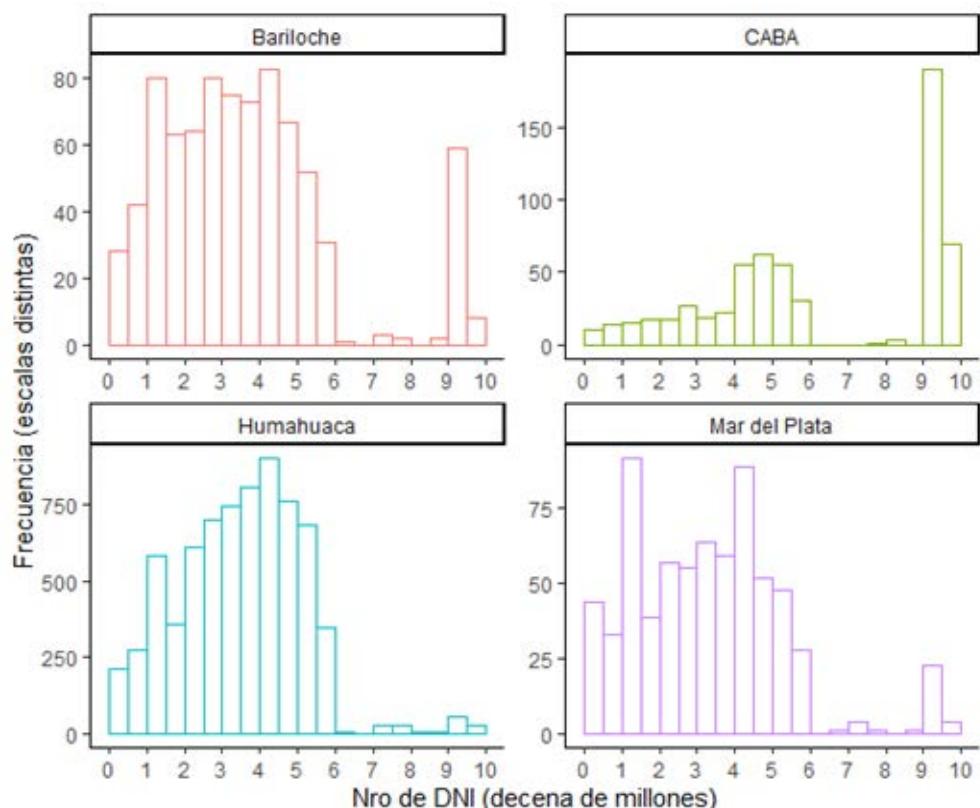
### Número de DNI declarado

El propósito principal para la inclusión de la pregunta Número de DNI en la II PP fue examinar la posibilidad de parear a la población a través del número de DNI con distintos registros administrativos, para enriquecer estos últimos con información censal con vistas a la construcción de un registro estadístico de población. Para esto, es necesario trabajar con el grupo con “DNI válido” del cuadro 5.2.1. En este subuniverso (9.265 casos), se detectaron 145 con número duplicado (1,6%), los cuales competirían en un proceso de pareo.<sup>32</sup> De estos, solo 26 corresponden a la misma vivienda (17,9%). Como caso anecdotico, se encontró en la base de datos una vivienda de Humahuaca con 6 personas que declararon el mismo número de DNI.

La distribución de números válidos por provincia se observa en el gráfico 5.3.1. Se destaca un primer porcentaje de números menores a 60 millones en todas las localidades; también una mayor dispersión en San Carlos de Bariloche y CABA por el peso de los números mayores a 90 millones, y una mediana en esta última muy por encima de las restantes localidades. Medidas de la distribución se muestran en el cuadro 5.3.1, donde se destacan números reportados cercanos al mínimo impuesto como condición de validez en Humahuaca y Mar del Plata.

<sup>32</sup> Al haber dos personas con un mismo número de DNI en la base de datos censal, en el proceso de pareo con los registros administrativos, solo uno de estos casos podría ser identificado en la base de datos de los registros.

**Gráfico 5.3.1 Histograma de Número de DNI según área de prueba. Cuestionarios básico y ampliado. II PP 2019**



**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

**Cuadro 5.3.1 Medidas de la distribución (mínimo, máximo, promedio y percentiles seleccionados) de Número de DNI según área de prueba censal. Cuestionarios básico y ampliado. II PP 2019**

Área de prueba	Mínimo	p25	p50	p75	Máximo
San Carlos de Bariloche	259.000	18.606.000	33.293.000	46.799.000	99.357.000
CABA	1.463.000	41.299.000	54.244.000	94.269.000	99.596.000
Humahuaca	103.000	22.932.000	35.823.000	46.175.000	99.021.000
Mar del Plata	115.000	16.179.000	32.144.000	44.253.000	98.611.000

**Nota:** se redondearon los números a la unidad de mil más próxima con motivo de resguardar el dato estadístico y su individualización.

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

Se examinó la declaración del número de DNI en término preferencia de dígitos. En términos ideales, la frecuencia relativa en las últimas 3 posiciones debería ser relativamente uniforme, salvo existencia de un sesgo en la declaración (por ejemplo, reportar sistemáticamente un 0 al final del número de DNI). Esto podría detectarse en la frecuencia relativa de los números 0 a 9 en las últimas 3 posiciones (centena a unidades). De acuerdo con lo analizado, puede aceptarse que no se encontraron sesgos en la declaración del número de DNI y no existió un patrón de preferencia en la declaración.

**Cuadro 5.3.2 Frecuencia relativa de los dígitos de los números de DNI reportados, según su posición. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019. En porcentajes**

Posición	Dígitos										Total
	0	1	2	3	4	5	6	7	8	9	
6	11,4	11,2	9,7	9,7	8,7	9,8	10,1	8,5	10,1	10,8	100,0
7	9,8	10,7	10,5	10	9,1	10,1	9,8	10,5	10	9,5	100,0
8	10,4	11,2	10,7	10,1	9,7	9,6	9,3	10,4	8,8	10	100,0

Fuente: INDEC. Base de datos sin consistir. II PP 2019.

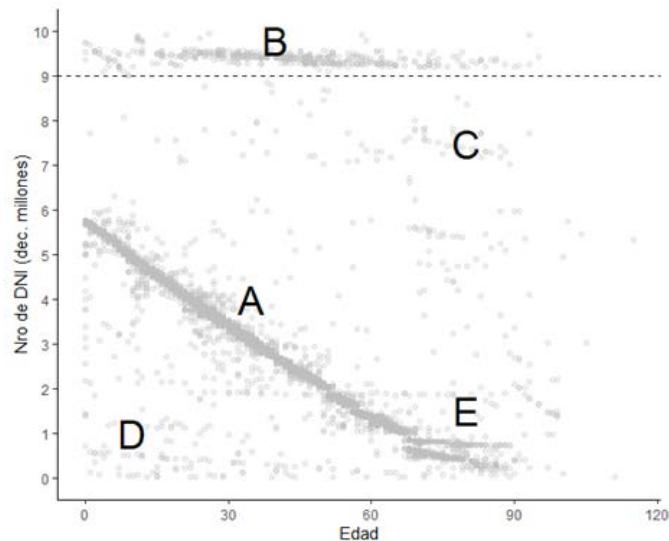
## 5.4

### Relación respecto a la edad

Si bien no existe una relación unívoca entre la edad de las personas y el número de DNI que se les asigna,<sup>33</sup> en la mayoría de los casos existe una relación entre ambas variables. En el siguiente gráfico pueden identificarse 5 grupos (ver gráfico 5.4.1). El grupo A se definió como aquellas personas con número de DNI menor a 90 millones y edad menor a 70 años, contenidas en la recta de regresión considerando rango de valores de más y menos 10 millones para cada edad; al grupo B, como aquellas con números de DNI mayor o igual a 90 millones; el grupo E, como aquellas con número de DNI menor a 20 millones y edad mayor o igual a 70 años.

Los grupos con más casos son el grupo A, que se caracteriza por presentar una correlación lineal alta ( $R^2 = 0,79$ ), con 7.833 casos (un 86,2%); luego el grupo B, por encima de los 90 millones, con 430 casos de números de DNI que probablemente pertenezcan a personas no nativas residentes; y luego el grupo E (474 casos), que presenta una relación más plana entre las variables. Los grupos C y D tienen menor densidad (con 179 y 175 casos, respectivamente) y se posicionan por fuera de los demás. En los próximos puntos se analizará algunos aspectos de B, C y D.

**Gráfico 5.4.1 Distribución número de DNI válido según Edad y Grupos sugeridos. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019**



Fuente: INDEC. Base de datos sin consistir. II PP 2019.

<sup>33</sup> El RENAPER informó que, por ejemplo, en el caso de la inscripción tardía de nacimientos el número de DNI asignado es el que se otorga a las cohortes que nacen en ese momento.

Debido a la asignación específica de la numeración de 90 millones y más para las personas que solicitan residencia en el país se esperaría que, a septiembre de 2019, los casos en ese rango correspondieran a personas no nativas. En el cuadro 5.4.1 se observa la relación entre estas variables.

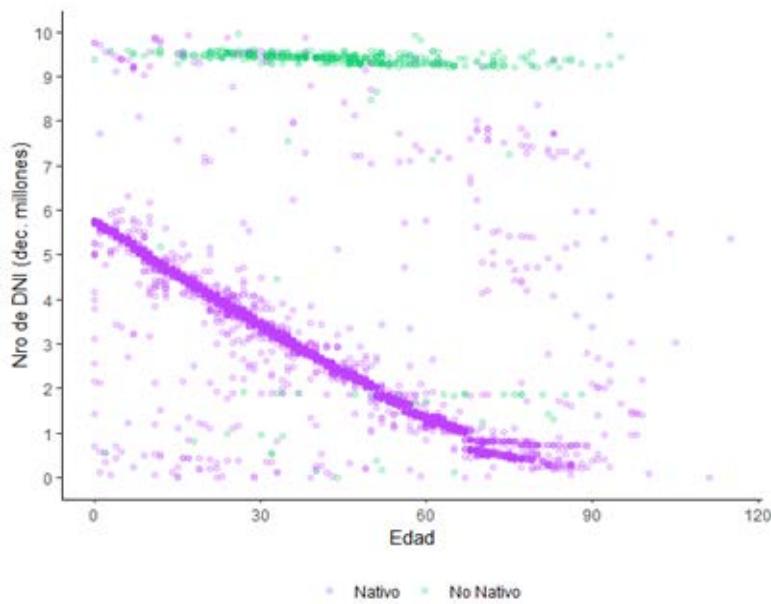
**Cuadro 5.4.1 Subconjunto con números de DNI mayor o igual a 90 millones según Lugar de nacimiento. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019. En valores absolutos y porcentajes**

¿Es nativo?	N	%
Sí	47	10,8
No	355	81,4
No respuesta	34	7,8
Total	436	100,0

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

Los resultados indican que un 10,8% de las personas con números de DNI de numeración 90 millones y más declaró haber nacido en la Argentina. En el gráfico 5.4.2, se observa que las personas nativas (puntos violetas) con número de números de DNI 90 millones y más se concentran en las edades jóvenes (lo que sugiere alguna imprecisión en la asignación de números de DNI a aquellas personas nativas descendientes de extranjeros(as); mientras que, para las personas no nativas (puntos verdes) la gran concentración se ubica en el rango esperado.

**Gráfico 5.4.2 Número de DNI según Edad y País de nacimiento. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019**



**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

Existen algunos casos de personas no nativas en la numeración de 20 millones, ello podría estar asociado a errores de lectura o al conjunto de excepciones de asignación del número de DNI, lo que no implica una imprecisión del número de DNI declarado. Más allá de estas observaciones, la concordancia entre País de nacimiento y Número de DNI parece aceptable.

## 5.5

## Casos aparentemente anómalos (C y D)

Continuando con el análisis por grupos, se realizó una inspección visual de una muestra aleatoria simple de 5 casos con edad mayor a 70 años y número de DNI declarado entre 70 millones y 90 millones (rincón derecho superior del grupo C), y 5 casos con edad menor a 10 años y número de DNI declarado menor a 10 millones (rincón izquierdo inferior del grupo D). El objetivo era verificar si estos casos anómalos se debían efectivamente a la declaración de números de DNI distintos a los esperados o a errores en la lectura e interpretación de los cuestionarios censales. Para ello se realizó una comparación entre la base de datos y la imagen del cuestionario escaneado. Para preservar la identidad de las personas se reemplazaron algunos dígitos de los números de DNI por el signo \*, dejando los dígitos que muestran diferencias de lectura. Los resultados de este análisis se presentan en el cuadro 5.5.1.

**Cuadro 5.5.1 Muestra en grupos C y D para inspección visual. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019**

Código barra	Carilla	Edad	Sexo	Número de DNI en la base	Número de DNI en el cuestionario
00000908	03	75	M	72513041	92513***
00005608	03	83	M	71888164	F1888***
00008120	03	71	V	78205049	M78205***
00008320	03	93	M	77310240	00820***
00001834	03	74	M	75246428	F5246***
00000456	07	2	M	05636182	54479***
00000814	07	1	V	1101610	53689***
00001401	13	5	-	5689144	57016***
00003878	05	4	M	04479260	95636***
00005916	11	6	M	5603193	***3193.

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

La inspección visual mostró que en algunos casos de edades avanzadas se antepuso la letra del sexo (“F” o “M”, relacionados a las libretas, documentos de emisión previa al DNI), pero en otras se leyó de manera incorrecta. En conclusión, aquellos extremos de los grupos C y D parecen producto de la lectura e interpretación incorrecta de los cuestionarios censales<sup>34</sup> o al incumplimiento de pauta por parte de quien censó al incorporar letras en los casilleros, o un punto como en el último caso.

## 5.6

## Comparación entre los casos con números de DNI aptos para ser pareables y el resto (sin información para parear)

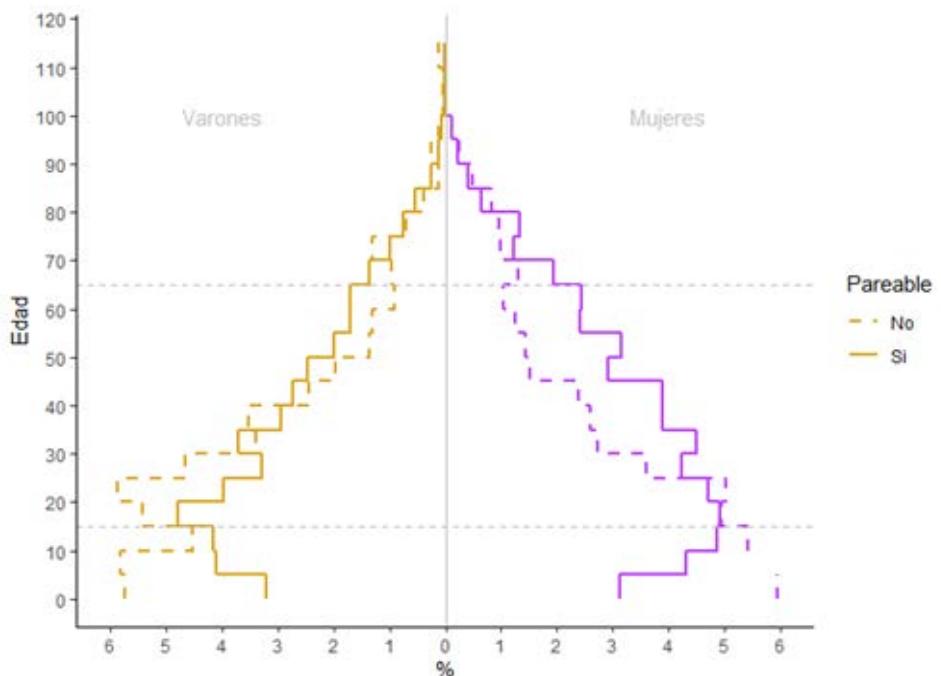
Se agruparon las respuestas en dos categorías, por una parte, los casos con número de DNI válido (definido de acuerdo con los criterios ya especificados) que presentan potencial para ser pareados con los registros administrativos y, en una segunda categoría, el resto de los casos no pareables.<sup>35</sup> Se compararon ambos grupos según algunas características de la población y de los hogares a los que pertenecen y se encontraron algunas diferencias.

<sup>34</sup> Para profundizar sobre las dificultades en la lectura e interpretación de los cuestionarios, ver el capítulo 3.

<sup>35</sup> “Blanco”, “Multimarca”, “No tiene DNI argentino”, “No contesta”, “Ignorado” y “DNI inválido”.

Como puede verse en la pirámide poblacional y la tabla de indicadores, el grupo de casos no pareables es más joven, y con menor presencia de mujeres, especialmente en el grupo de edad central de 15-64.

**Gráfico 5.6.1 Pirámide de población. Población pareable (línea llena) y no pareable (línea punteada). Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019**



**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

**Cuadro 5.6.1 Indicadores demográficos según validez de Número de DNI. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019.**

Indicador	No pareable	Pareables
Edad media	27,5	32,9
% 0-14	33,7	23,7
% Mujeres	49,1	55,2
% Mujeres 15-64	46,3	55,9

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

En el cuadro 5.6.2 se seleccionaron algunas variables socioeconómicas (presentes solo en el cuestionario A) para caracterizar a los dos grupos. Entre los casos no pareables, se observa una mayor presencia de personas inactivas, con menor nivel de instrucción superior y mayor tamaño del hogar (cantidad de personas). No se observan diferencias según la calidad del material de los pisos (variable que se puede considerar proxy de las condiciones de vida del hogar).

**Cuadro 5.6.2 Características de la vivienda o las personas, según validez de número de DNI. Total de áreas de prueba. Cuestionarios básico y ampliado. II PP 2019. En valores absolutos y porcentajes**

Variable		No pareables	Pareables	N
Condición de actividad	Desocupado	2,9% (18)	3,5% (119)	137
	Inactivo	41,4% (255)	37,6% (1.266)	1.521
	Ocupado	55,7% (343)	58,9% (1.986)	2.329
Nivel educativo	Desde terciario incompleto	16,4% (267)	21,8% (1.573)	1.840
	Hasta secundario completo	39,3% (641)	39,8% (2.873)	3.514
	Hasta primario completo	44,3% (722)	38,5% (2.780)	3.502
Material del piso	Carpeta, contrapiso o ladrillo fijo	20,9% (404)	18,4% (1.473)	1.877
	Cerámica, mosaico, baldosa	74,0% (1.432)	72,7% (5.822)	7.254
	Otro	0,3% (6)	1,2% (97)	103
	Tierra o ladrillo suelto	4,8% (93)	7,7% (614)	707
Tamaño del hogar	1	5,3% (109)	7,1% (588)	697
	2	9,4% (193)	15,2% (1.265)	1.458
	3	20,2% (415)	20,1% (1.675)	2.090
	4	20,1% (413)	21,9% (1.829)	2.242
	5	17,4% (357)	13,9% (1.163)	1.520
	6	11,4% (233)	10,1% (838)	1.071
	7	16,2% (332)	11,8% (980)	1.312

**Fuente:** INDEC. Base de datos sin consistir. II PP 2019.

## 5.7 Conclusiones

El análisis realizado permite extraer las siguientes conclusiones, que se consideran de valor como instancia previa al ejercicio de pareo:

- Mientras que las categorías “Ignorado” y “No contesta” representan un 9,7% (1.138) y 5,2% (613) de los casos respectivamente, un 79,1% (9.294) respondió un número de DNI (al menos 3 marcas en los 8 casilleros de la pregunta). Existen algunas diferencias según Área de prueba, Sexo, Condición de autorrespondiente y Lugar de nacimiento.
- El 78,8% (9.265) de los casos podrían considerarse candidatos para un proceso de pareo según la definición de válido que se precisó: números con cantidad de dígitos entre 6 y 8.
- De estos, 7.833 casos (un 86,2%) presentan una relación clara entre la numeración del DNI y la edad de la persona. Entre las personas mayores de 70 años se observa una distribución particular de numeración del DNI que probablemente obedezca a cuestiones administrativas del registro civil, quizás específicas de las localidades censales relevadas.
- Los casos con número de DNI mayor a 90 millones fueron 430 (4,7%). Entre estos, un 10,8% declaró haber nacido en la Argentina y está concentrado principalmente en las primeras edades.
- La inspección visual de los cuestionarios en aquellos casos donde la relación entre el número de DNI y la edad se presenta aparentemente anómala indicó que son producto de errores de lectura o del incumplimiento de pauta de trabajo por parte quien censó (como colocar letras o separar con punto los millares en las casillas donde se debían escribir solo números).

- Los casos en que no se podría realizar el pareo unívoco a través del Número de DNI (no responde/ignorado/número no válido/etc.) representan un 21,2% y corresponden a población comparativamente más joven y en las edades centrales presenta un predominio de varones; además, se trata de una población con mayor proporción de inactivos, de menor nivel educativo y que vive en hogares de mayor tamaño.
- Se recomienda realizar los ejercicios de pareo con los registros del RENAPER y otros registros administrativos para poder evaluar integralmente la calidad de la declaración.

## 6. Evaluación de la codificación de preguntas abiertas

### 6.1 Introducción

La codificación de las preguntas abiertas de la II PP se realizó con el Sistema de Codificación Informatizada (en adelante SiCI). Este sistema lee las bases a procesar y codifica de manera automática. La información que no es codificada de esta forma pasa a la instancia de codificación semimanual, que es la realizada en línea por codificadores.

Las preguntas codificadas provienen de algunos de los dos cuestionarios censales o de ambos cuestionarios:

- 1 Migraciones (cuestionario B y A)
- 2 Pueblos originarios (cuestionario B y A)
- 3 Carrera universitaria y terciaria (cuestionario A)
- 4 Rama de actividad económica (cuestionario A)
- 5 Ocupación (cuestionario A)

El objetivo principal de este proceso fue el de conocer los porcentajes de la codificación automática y poner a prueba el funcionamiento del SiCI con preguntas que se codificaban por primera vez.

Una vez codificadas las preguntas, se analizó la información no codificada para intentar, mediante distintos procedimientos, incorporar todos los casos posibles al diccionario de codificación, y de esta manera ir aumentando los niveles del procesamiento automático.

Al procesar la información se encontró un importante problema relacionado con la lectura de todas las preguntas abiertas de los cuestionarios censales, y otros relacionados con el relevamiento, que dificultaron el proceso de codificación. A continuación, se detallan los diferentes problemas que se presentaron:

- Palabras juntas, sin separación entre ellas. Por ejemplo, PROFESORADOENSENANZA PRIMARIA, MANTENIMIENTOIDEALUMBRADO.
- Confusión sistemática de letras, que en muchos casos imposibilita su lectura o “interpretación”. Por ejemplo, U x V; P x A; H x N; K x E; etc.
- Casos de ilegibilidad total de la información. Por ejemplo, VROTESORDEFTSICAYJUIMICA.
- Respuestas que no responden a lo que se pregunta.

A continuación, se presentan los resultados de los procesos de codificación de cada una de las preguntas.

## 6.2 Migraciones

Las respuestas abiertas que componen la temática migración corresponden a país y provincia de nacimiento, residencia cinco años antes y residencia habitual. Estas respuestas se codificaron de manera independiente entre sí, utilizando el “Código numérico de dos posiciones para divisiones político-administrativas de la República Argentina” y el código de países del INDEC.

### 6.2.1 Codificación de Provincia

El porcentaje de codificación automática registrado se debe a las distintas correcciones que debieron hacerse a la base para que el sistema pudiera procesarlas.

Realizada la codificación automática, con la información tal como fuera leída de campo, el SiCI codificó un 71,6% de los casos. A partir del análisis de la información no codificada se observaron los siguientes problemas para el procesamiento:

- Valores erróneos introducidos en la etapa de lectura. Por ejemplo, AUCUMAN, CALTA, JVJUY, MSNDOPA, MEUQUEN, etc., que el sistema no podía reconocer.
- Nombres de lugares geográficos (localidades, parajes, etc.) que no corresponden a provincias. Por ejemplo, Banfield, Bolsón, Lanús, Rosario, Mar del Plata, etcétera.

Con el propósito de incorporar esta información a los diccionarios de codificación del SiCI, cuando la lectura o interpretación de lo leído era factible de ser asignada o asociada a un nombre del código de manera inequívoca, fue corregida como sinónimo para su posterior codificación automática. Por ejemplo, CALTA fue incorporada como SALTA, o MEUQUEN como NEUQUEN.

Este procedimiento permite que, si se repite este mismo error en el censo definitivo, la lectura sea correcta y el caso codificado.

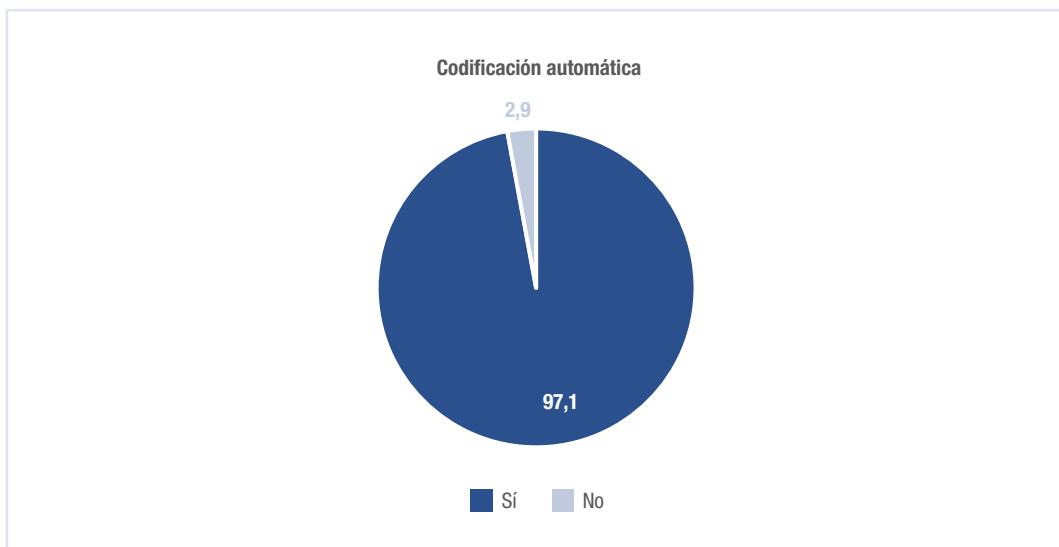
Cuando se encontraron nombres de otras ciudades que pertenecían claramente a una provincia se los incorporó como sinónimo de la provincia. Por ejemplo: Bolsón se consideró como sinónimo de Río Negro; o Mar del Plata, como sinónimo de Buenos Aires. Este procedimiento de corrección permitió aumentar la codificación automática de esta pregunta a un 87,5%. Se reiteraron estas tareas por segunda vez, y en el nuevo procesamiento se codificó un 97,1%.

**Cuadro 6.2.1.1 Provincia (A: P11, P12 y P13) (B: P9). Codificación de respuestas abiertas. Cuestionario básico y ampliado. Total de áreas de prueba. II PP 2019. En valores absolutos**

Tipo de cuestionario	Total
Básico	971
Ampliado	25.442
Total	26.413

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

**Gráfico 6.2.1.1 Provincia (A: P11, P12 y P13) (B: P9). Porcentaje de codificación automática. Cuestionario básico y ampliado. Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 6.2.2 Codificación de País

El procesamiento de las respuestas sobre País se realizó de igual manera que para provincia. Una vez codificada la base tal cual fuera recibida, el primer procesamiento codificó automáticamente un 44,9%, el segundo procesamiento arrojó un 58,8% automático. Y en el último, con todas las posibles correcciones incorporadas, se codificó un 91,8%.

En esta pregunta se encontraron situaciones similares a las descriptas en provincia:

- Errores en la lectura de los países. Por ejemplo, AKGENTINA, BOIIYIA, SANTIAGO DC CHILE.
- Nombres de provincias o ciudades en lugar de países. Por ejemplo, JUJUY, SANTIAGO, RIO NEGRO, PUERTO MONTT.

Se incorporaron los nombres de ciudades o provincias que pertenecieran inequívocamente a un país. Por ejemplo, si decía JUJUY o RIO NEGRO fue codificado como Argentina.

Sin embargo, se encontraron respuestas que no pudieron ser asignadas, por ejemplo, Santiago, ya que podía ser Santiago del Estero, Argentina o Santiago de Chile.

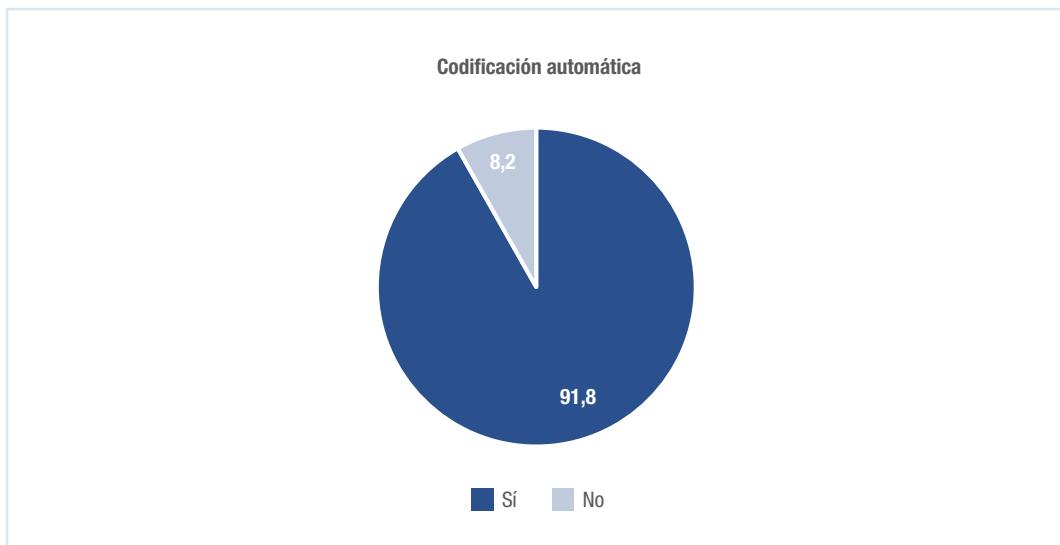
Asimismo, se encontró una amplia cantidad de respuestas de “Argentina” en la pregunta abierta referida a la de “otro país”, que no correspondía.

**Cuadro 6.2.2.1 País (A: P11, P12 y P13) (B: P9). Codificación de respuestas abiertas. Cuestionario básico y ampliado. Total de áreas de prueba. II PP 2019. En valores absolutos**

Tipo de cuestionario	Total
Básico	110
Ampliado	1.692
Total	1.802

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

**Gráfico 6.2.2.1 País (A: P11, P12 y P13) (B: P9). Porcentaje de codificación automática. Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 6.3

## Pueblos originarios

Para la codificación de estas preguntas (pregunta 10 del cuestionario B y pregunta 17 del cuestionario A) se utilizó el Código de Pueblos Originarios, conformado por una serie de listados de nombres de pueblos provenientes del Censo Nacional de Población, Hogares y Viviendas 2001(CNPhyV 2001), la Encuesta Complementaria de Pueblos Indígenas (ECPI) y el Censo Nacional de Población, Hogares y Viviendas 2010 (CNPhyV 2010).

En el procesamiento de esta pregunta se implementó la misma dinámica que la usada para las preguntas referidas a migraciones. Se procesó la información tal como llegó para su codificación automática, con lo que se obtuvo un resultado de 39,1%.

La formulación de las preguntas es la siguiente:

“¿(Nombre) se reconoce indígena o descendiente de pueblos indígenas u originarios?”, “¿De qué pueblo?”. Esta formulación da lugar a contestar por el pueblo en el que la persona censada vive y no del pueblo originario al que pertenece o desciende. Se sugiere que la segunda parte de esta pregunta se modifique por “¿De qué pueblo indígena u originario?”.

En esta pregunta se encontraron los problemas comunes al resto de las preguntas abiertas y otros específicos de la temática, que se detallan a continuación:

- Illegibilidad total de respuestas. Por ejemplo, AABB DD, AANCO CAPAC, OMAHUACB, CFFA GKKHDE.
- Respuestas de diferentes lugares geográficos en lugar de los nombres de pueblos originarios. Por ejemplo, Santa Rosa, Aguas Blancas y Bolivia, entre otros.
- Palabras que no se puede determinar qué significan. Por ejemplo, Cova, Doncellas, Nazareno, Cueva Indio, Unión y Fuerza, etcétera.
- Nombres de comunidades que no se puede establecer si pertenecen a algún pueblo particular. Por ejemplo, comunidad aborigen, comunidad de Varas, comunidad de la Poma, comunidad de Rodero, la comunidad Cacique Limpitay, comunidad de Ovara, etcétera.

En relación con el código de pueblo:

- Nombres de pueblos que no están en el listado y aparecieron en la base. Por ejemplo, INCA, QUILMES, CALCHAQUI, RANQUELES.
- Nombres de pueblos que serían sinónimos de pueblos que aparecen en el listado y que, de ser correctos, deberían incorporarse con el mismo código. Por ejemplo, MATACOS como sinónimo de WICHIS.
- Nombres de pueblos que no son del actual territorio argentino y habría que definir si se agregan al listado. Por ejemplo, PIPIL, que es un pueblo indígena de El Salvador.
- Nombres compuestos de pueblos que no sabemos si son tales o arrastran algún problema de lectura o relevamiento: GUARANI OMA; COYA/KOLLA CALETE; COYA/KOLLA CASPALA; INCA CUEVA.
- Se observaron nombres de pueblos que se pudieron asociar por interpretación a algunos pueblos y se incorporaron como sinónimos:

AYHARA por AYMARA

AYMAQA por AYMARA

GUARDNI por GUARANI

GUARMNY por GUARANI

OMABUACA por OMAGUACA

OMACUACAS por OMAGUACA

OMAGCACA por OMAGUACA

La inclusión de los casos relacionados al Código de Pueblos Originarios no tiene problemas de clasificación ni metodológicos; sin embargo, se requiere la confirmación de algunos criterios conceptuales por los especialistas en la temática.

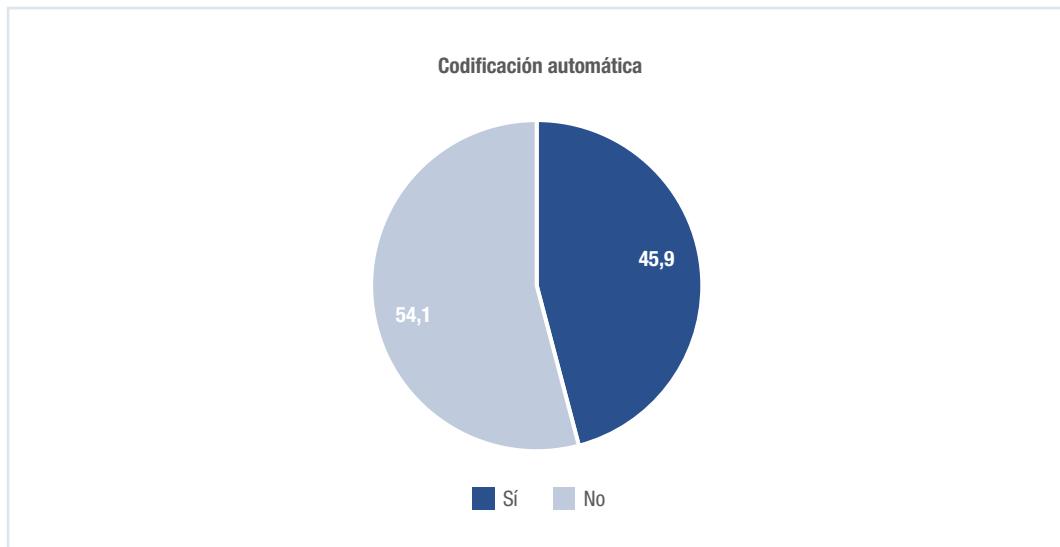
Una vez realizadas las correcciones relacionadas a los problemas de lectura, pero que permitieran la interpretación inequívoca de las palabras o nombres y la incorporación de las diversas formas de escritura de un pueblo, se llegó al resultado que se presenta en el gráfico 6.3.1.

**Cuadro 6.3.1 Pueblos originarios. (A: P17) (B: P10). Codificación de respuestas abiertas. Cuestionario básico y ampliado. Total de áreas de prueba. II PP 2019. En valores absolutos**

Tipo de cuestionario	Total
Básico	234
Ampliado	3.373
Total	3.607

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

**Gráfico 6.3.1 Pueblos originarios. (A: P17) (B: P10). Porcentaje de codificación automática. Cuestionario básico y ampliado. Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 6.4

## Carrera universitaria y terciaria

La población a la que va dirigida la pregunta sobre carrera terciaria o universitaria de grado que completó (pregunta 10 del cuestionario A) se compone del universo de los graduados universitarios y los graduados del nivel terciario que en conjunto permiten conocer a la población con educación superior del país.

Para su codificación se utilizó la Clasificación de Carreras Universitarias y Terciarias (CdeC-2020).<sup>36</sup> Este clasificador fue elaborado especialmente para el censo de población de la ronda 2020 y probado tanto en su estructura como en su procesamiento con el SiCI en la II PP.

El porcentaje de codificación automática fue del 18,9% (252 casos) y luego de analizar la información, se la corrigió y se la incorporó al clasificador. Al finalizar este proceso se alcanzó aproximadamente el 90% de la codificación.

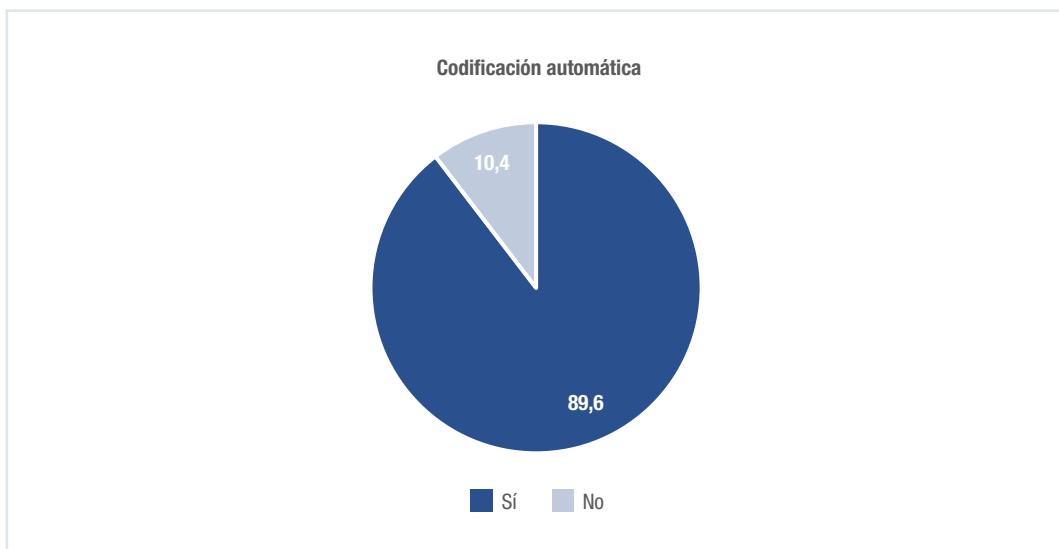
Esta variación en los niveles de codificación automática se dio por la corrección en gabinete de errores relacionados con la mala lectura de los cuestionarios. Además, permitió aumentar el diccionario de codificación con las nuevas formas de mencionar algunas de las carreras relevadas.

**Cuadro 6.4.1 Carrera universitaria y terciaria. (A: P31). Codificación de respuestas abiertas. Cuestionario ampliado. Total de áreas de prueba. II PP 2019. En valores absolutos**

Tipo de cuestionario	Total
Básico	234
Ampliado	3.373
Total	3.607

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

**Gráfico 6.4.1 Carrera universitaria y terciaria. (A: P31). Porcentaje de codificación automática. Cuestionario ampliado. Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

<sup>36</sup> Esta clasificación fue elaborada por la Coordinación de Clasificaciones y Nomenclaturas (CCN), con la participación de equipos específicos de especialistas de la Dirección Nacional de Estadísticas Sociales y Poblacionales (DNEStyP) y la colaboración del Ministerio de Ciencia, Tecnología e Innovación y el Ministerio de Educación de la Nación.

Más allá de los ajustes realizados en el SiCI, se observaron los siguientes inconvenientes en la codificación:

- Palabras sin separación entre ellas:

PROFESIONALGASTRONOMICO,

TECNICATURAEN HOTLLERIA,

PRQFESORADONIVEL PRIMARIO

TECNICATURAEN ASISTENCIA ODONTOLO

PROFESORADOENSENANZA PRIMARIA

PROFESORADOPARA LG CMSEMANZA PRIM

TECNICATURAAGENTE SANITARIO

- Frases cortadas, sin finalizar por falta de espacio:

PROFESOR DEARTES MANUALES Y LERAM

PROFKSORAOOELEMENTAL PARA LA ENS

PROFESORADDDE MATEMAT ICA Y COSMO

PROFESORADOPARA LA EN SENANZA PRI

TECNICATURAEN RECREACION EN DEPOR

PROFESORA OE CASTELLANO Y LITERAT

## 6.5

## Rama de actividad económica

Habitualmente la codificación de la pregunta referida a la rama de actividad económica (pregunta 31 del cuestionario A) en los relevamientos sociodemográficos,<sup>37</sup> se realiza con la Clasificación de Actividades Económicas para encuestas Sociodemográficas (CAES). Para el censo de población de la ronda 2020 se decidió implementar, por primera vez, la codificación con la Clasificación Nacional de Actividades Económicas (ClNAE). Si bien las dos clasificaciones parten de la Clasificación Internacional Única de Actividades Económica (CIIU Rev. 4), la CAES es una adaptación más agregada, pensada para relevamientos en hogares; por el contrario, la ClNAE es una adaptación más desagregada, con más especificidad de los grupos clasificatorios, formulada para relevamientos en unidades económicas.

La adaptación del SiCI para la ClNAE implica la recodificación de bases provenientes de distintos relevamientos sociodemográficos (que habían sido codificadas con la CAES). Esta recodificación no

<sup>37</sup> Encuesta Permanente de Hogares (EPH), Encuesta de Gastos de los Hogares (ENGHo) y los censos de población desde 2001.

siempre es posible a su máximo nivel de desagregación, ya que un relevamiento a hogares no suele ser respondido por cada integrante del hogar, sino que responde una persona por todos sus miembros, motivo por el cual se puede desconocer el detalle de la información referida a la actividad del resto.

Al momento del procesamiento de la II PP aún no se había decidido qué clasificación se utilizaría; por lo tanto, esta codificación se realizó con ambas clasificaciones, para analizar los posibles escenarios.

A continuación, se presentan los ejercicios realizados sobre los datos con los dos instrumentos de clasificación. Independientemente de la clasificación utilizada se presentan problemas comunes.

## 6.5.1

### Codificación con la Clasificación de Actividades Económicas para Encuestas Sociodemográficas (CAES)

En comparación con los resultados de la codificación obtenida al procesar la Encuesta Permanente de Hogares (EPH), donde la codificación automática de la pregunta de Rama de actividad económica promedia el 68%, el porcentaje de codificación automática obtenido en la II PP fue inferior. Es importante mencionar que las condiciones de relevamiento de una encuesta a hogares y de un censo no son las mismas, y la calidad de la información que se obtiene en un censo suele ser inferior a la de las encuestas (cuadro 6.5.1.1).

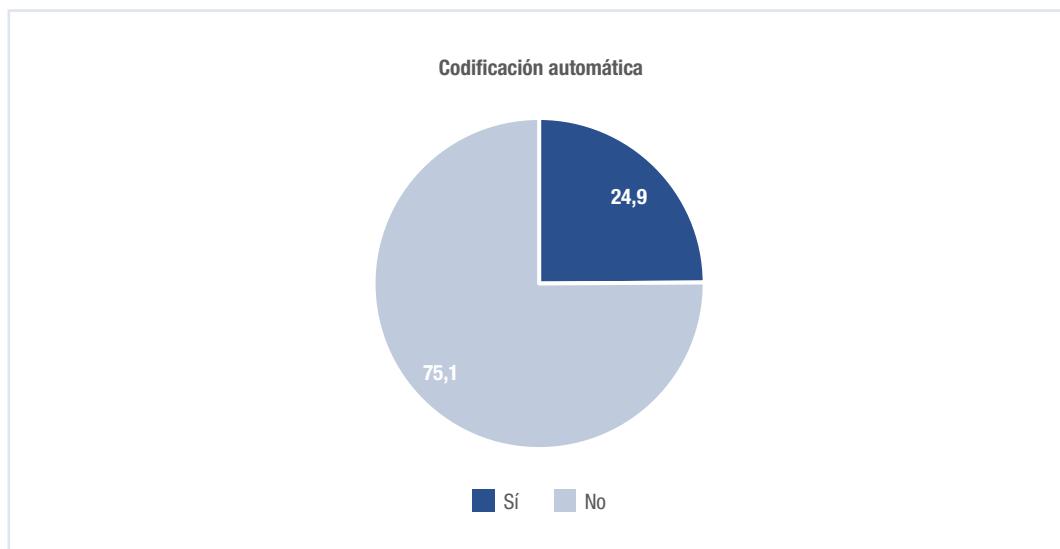
Una vez realizada la codificación automática se procedió a la codificación de forma semimanual y se incorporaron las frases resueltas manualmente al diccionario de codificación del sistema; aquí radica la importancia de este procedimiento, alimentar al sistema de codificación de información proveniente de campo en situación censal.

**Cuadro 6.5.1.1 Clasificación de Actividades Económicas para Encuestas Sociodemográficas. Codificación de respuestas abiertas. Cuestionario ampliado (A: P31). Total de áreas de prueba. II PP 2019. En valores absolutos**

Rama de actividad económica	Total
Total a codificar	4.214
Codificado automáticamente	1.048

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

**Gráfico 6.5.1.1 Clasificación de Actividades Económicas para Encuestas Sociodemográficas. Porcentaje de codificación automática. Cuestionario ampliado (A: P31). Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Un 45,1% del total a codificar lo fue a la mayor desagregación y un 12,3% se codificó a menor desagregación. Sin embargo, se registró un alto porcentaje de información en blanco, no codificable y casos ilegibles: 42,6% del total (cuadro 6.5.1.2).

**Cuadro 6.5.1.2 Clasificación de Actividades Económicas para Encuestas Sociodemográficas. Codificación de respuestas abiertas. Cuestionario ampliado (A: P31). Total de áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Rama de actividad económica	Total	Porcentaje
Codificado a 4 dígitos	1.901	45,1
Codificado a menos de 4 dígitos (¹)	517	12,3
No codificables (²)	1.796	42,6
Total a codificar	4.214	100,0

(¹) Incluye lo codificado a letra y a dos dígitos.

(²) Incluye respuestas en blanco, información insuficiente y casos ilegibles.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Si bien la formulación de la pregunta en la II PP fue igual a la de la EPH, lo que sí varió fue el orden de las preguntas de Rama de actividad y Ocupación. En esta prueba se invirtió el orden, preguntándose primero por Ocupación y luego por Rama de actividad, pero los resultados de la prueba piloto no reflejan una mejora con ese cambio, y contando con menor cantidad de espacio para registrar la respuesta.

## 6.5.2

## Codificación con la Clasificación Nacional de Actividades Económicas (ClNAE)

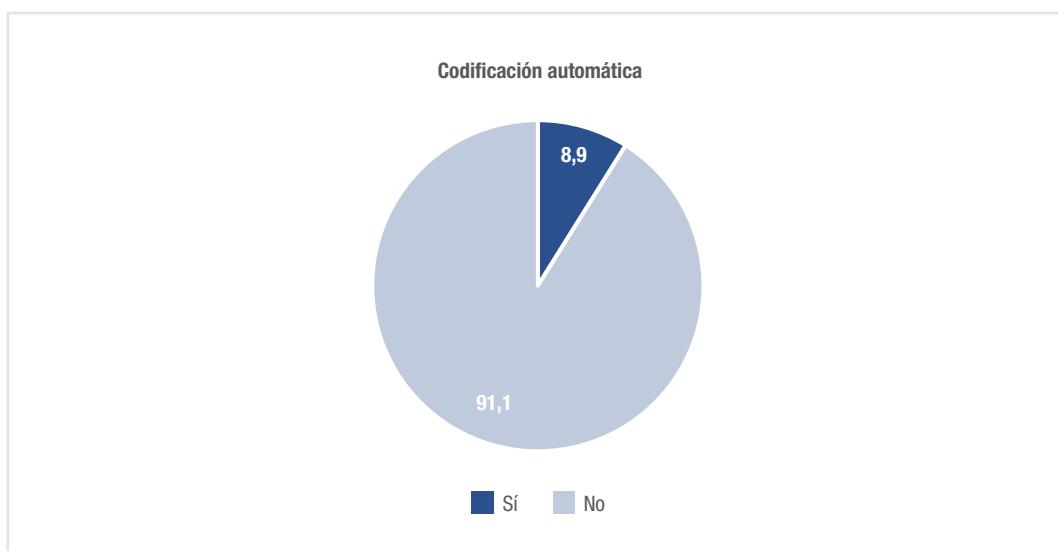
El procesamiento de codificación de esta pregunta no fue óptimo por diversos motivos, y alcanzó el 8,9% (gráfico 6.5.2.1). Por un lado, como se detallará más adelante, se observaron problemas en la lectura y la calidad de la información recibida; y por otro, el SiCI está en proceso de adaptación de sus diccionarios de codificación para esta Clasificación.

**Cuadro 6.5.2.1 Clasificación Nacional de Actividades Económicas. Codificación de respuestas abiertas. Cuestionario ampliado (A: P31). Total de áreas de prueba. II PP 2019. En valores absolutos**

Rama de actividad económica	Total
Total a codificar	4.214
Codificado automáticamente	375

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

**Gráfico 6.5.2.1 Clasificación Nacional de Actividades Económicas. Porcentaje de codificación automática. Cuestionario ampliado (A: P31). Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Realizada la codificación automática se completó la base con la instancia de codificación semimanual, con lo que se llegó a codificar el 74,3% de los casos. Es importante considerar los niveles de codificabilidad total o parcial de la información, ya que, como se explicó anteriormente, las respuestas pueden ser no tan detalladas para una codificación más desagregada, pero sí podrían llegar a codificarse a menor cantidad de dígitos.

Del total de registros a codificar se obtuvo un 32,1% codificado a 5 dígitos, máximo nivel de desagregación para esta Clasificación, y un 37,7% con codificabilidad parcial. El 30,2% de casos no pudieron ser codificados por falta de información o por ilegibilidad.

**Cuadro 6.5.2.2 Clasificación Nacional de Actividades Económicas Codificación de respuestas abiertas. Cuestionario ampliado (A: P31). Total de áreas de prueba. II PP 2019. En valores absolutos**

Rama de actividad económica	Total	Porcentaje
Codificado a 5 dígitos	1.353	32,1
Codificado a menos de 5 dígitos <sup>38</sup>	1.589	37,7
No codificable <sup>39</sup>	1.272	30,2
Total a codificar	4.214	100,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Las observaciones mencionadas en la codificación con la CAES aquí se repiten. A continuación, se detallan ejemplos de los problemas encontrados en el procesamiento, que son válidos independientemente de la clasificación que se utilice:

- Palabras sin separación y que no en todos los casos completan las frases:

ESTABLECIMIEUTDDKLYUTACION

FABRICACIONDEREJASPUERTAS

GESTIONACTIVIDADCOMUNITARIA

MANTENIMIENTOIDEALUMBRADO

PINTURAMACRAME

PRENDADEPUNTO

PRENDAINTERIOR

SERVICIOTURISTI

VENTADEROPA

VENTAMINORISTA

- Ilegibilidad por cambio de letras en las palabras:

AOHIMISTRATIVA

ARTESHSIAS

CONCTXIJCCJON

DISTRIBUCION

FPBRICACION

<sup>38</sup> Incluye codificación a letra, a 2/3 dígitos y 4 dígitos.

<sup>39</sup> Incluye respuestas en blanco o con información insuficiente o ilegibilidad de la información.

GASTKONOMIA

HELOERIA

HOTCLERIA

IASTITUCION

KIOSCM

LIMPILZJ

LOMERCIALISACION

REPAKFCIOJ

## 6.6 Ocupación

Para la codificación de las preguntas de ocupación (preguntas 29 y 30 del cuestionario A) se utilizó el Clasificador Nacional de Ocupaciones del INDEC. Este es el mismo instrumento que se utiliza en la EPH y también en la ENGHO 2017/2018. La indagación de esta temática se desarrolla en dos preguntas, que al momento de codificar se trabajan en forma conjunta.

Si bien la pregunta del Nombre de la ocupación y las Tareas que en ella se desarrollan están formuladas de la misma manera que en las encuestas a hogares mencionadas, en la II PP se quitó del cuestionario la pregunta acerca del uso o no de equipo/maquinarias o sistemas electromecánicos o informatizados, por lo cual solo se tendrán parámetros de esta dimensión clasificatoria si el uso de algún tipo de maquinaria se registra en las tareas.

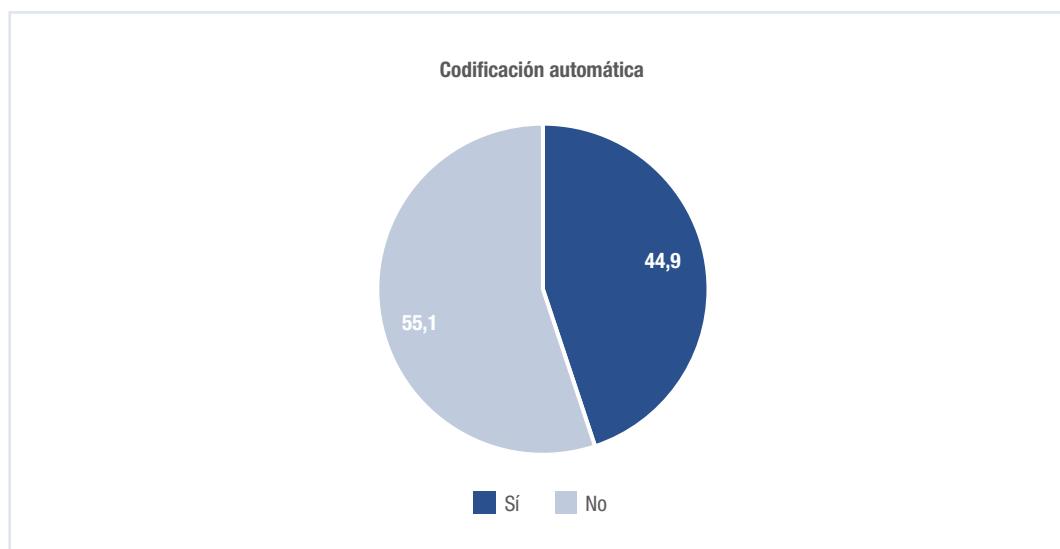
**Cuadro 6.6.1 Ocupación. Codificación de respuestas abiertas. Cuestionario ampliado (A: P29 y P30). Total de áreas de prueba. II PP 2019. En valores absolutos y porcentajes**

Ocupación	Total
Total a codificar	4.214
Codificado automáticamente	1.890

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Como se puede observar en el gráfico 6.6.1, la codificación automática alcanzó el 44,9%.

**Gráfico 6.6.1 Ocupación. Porcentaje de codificación automática. Cuestionario ampliado (A: P29 y P30). Total de áreas de prueba. II PP 2019**



**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Comparando estos resultados con los obtenidos habitualmente en la EPH, resultan ser menores, ya que en la encuesta los porcentajes de codificación automática son entre el 65% y 68%. Realizada la codificación automática, el resto de los casos fueron codificados de manera semimanual.

Si bien se pudo codificar aproximadamente el 83,0% de los registros, es importante señalar la cantidad de respuestas en blanco o ilegibles que se concentran en la categoría Sin información, aproximadamente el 15% (cuadro 6.6.2).

**Cuadro 6.6.2 Ocupación. Codificación de respuestas abiertas. Cuestionario ampliado (A: P29 y P30). Total de áreas de prueba. II PP 2019. En valores absolutos**

Ocupación	Total	Porcentaje
Con código de ocupación	3.482	83,0
Información no codificable	89	2,1
Inactivos	18	0,43
Sin información	625	14,8
Total a codificar	4.214	100,0

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En esta codificación se observaron los mismos problemas que en las anteriores.

- Palabras juntas sin separación entre sí:

AYUDANTEOEALBANIL

BORDADODECUERO

CONFKCCDONCAPACITACION

DOCENTEGASTRONS

PERSONALDOMESTICC

SERVICIOSDETJRISMO

VENDEDORATAMBULANTEI

- Cambios de letras:

AKTCSANO

BIBLMOTECARIA

CARVICERO

CCCINA

CHOFSR

COMCRCIPNTE

ELLCTRICISTA

PROSRAMADOR

QUIRURDICA

RECEPCIOHISTA

UENDEOORD

El cambio de secuencia en el cuestionario entre esta pregunta y la de Rama de actividad también impacta en la calidad de las respuestas de Ocupación a pesar de ser más notable en las de Rama de actividad.

También aquí resulta insuficiente el espacio asignado para completar las respuestas: la pregunta 29 (Nombre de la ocupación) dispone de 32 caracteres y la pregunta 30 (Tareas), de 48. Esta misma formulación, por ejemplo, en el Censo 2001 disponía de 44 y de 66 caracteres, respectivamente.

## 6.7 Conclusiones

Un proceso de codificación de relevamientos masivos, como lo es un censo de población, requiere del aumento de los niveles de codificabilidad automática de las preguntas abiertas. Eso permitirá reducir los tiempos para contar con la información y unificar los criterios de codificación para cada pregunta.

A modo de síntesis, el capítulo 8 relativo a las recomendaciones que emanen de la evaluación cuantitativa de los resultados de la II PP detalla el conjunto de acciones sugeridas para el mejoramiento del proceso de codificación.

# Evaluación de calidad de los datos provenientes de la capacitación de “emergencia” y de “estudiantes terciarios”

## 7.1

### Introducción

En el presente capítulo se analizan los errores de contenido correspondientes a las preguntas del cuestionario A en los radios 9 (R9) y 11 (R11) de la fracción 4 de Humahuaca.

El interés en focalizar la mirada en estos radios residió en que allí se probaron estrategias novedosas respecto al reclutamiento y la capacitación de censistas. En el caso del R9, la característica distintiva de sus censistas fue haber transitado por un proceso de capacitación diferente al resto, y en el caso del R11, su particularidad recayó en tratarse de un grupo de estudiantes terciarios.

Uno de los objetivos de la II PP fue explorar la posibilidad de desarrollar estas estrategias en el censo y obtener aprendizajes en torno a la planificación y capacitación de los recursos humanos, como así también detectar beneficios y obstáculos para su implementación (INDEC, 2019: p. 44).

En el R9, los censistas asignados no fueron convocados a los cursos presenciales, ni se les dio acceso al campus virtual, con el objetivo de simular una situación de emergencia. Tanto en el Censo de 2001 como en el de 2010, se observó la necesidad de convocar a censistas voluntarios sobre la fecha del Censo, como recurso ante eventuales bajas que no pudieran cubrirse con censistas suplentes. Este grupo no recibía capacitación por falta de tiempo (INDEC, 2019: pp. 51 y 52). Por esta razón, por primera vez se diseñó en nuestro país un dispositivo pedagógico que tuvo por objetivo formar censistas en la misma mañana del relevamiento, en una actividad de capacitación de una hora de duración.

Por otro lado, en el R11 se probó una estrategia novedosa de reclutamiento: se conformó una estructura censal integrada por estudiantes terciarios convocados y designados como censistas, con su docente como jefa de radio. El objetivo de prueba fue evaluar el desempeño de esta estructura completa, reconocida y organizada jerárquicamente, a modo de personal alternativo a los docentes primarios. Cabe aclarar que este grupo realizó con normalidad la capacitación censal destinada a los censistas, basada en la participación de dos instancias de formación: la capacitación virtual y la capacitación presencial.

El presente análisis se centra en comparar los errores obtenidos en las respuestas a las preguntas de los radios 9 y 11, con relación al total de errores observados en el resto de los radios de Humahuaca y en el total de las jurisdicciones de la prueba. La evaluación se realizó analizando los errores producidos por ausencias de respuesta, multimarcas o respuestas inválidas, observados en el registro de datos de las variables de vivienda, hogar y población del cuestionario. Los datos de Humahuaca (radios 9, 11 y resto de Humahuaca) corresponden a las preguntas del cuestionario A; los datos del total de jurisdicciones refieren a los cuestionarios ampliado y básico.

Esta evaluación cuantitativa de la calidad de los datos censales se complementó con el análisis cualitativo de las observaciones no participantes efectuadas durante la prueba piloto. Esta estrategia permite identificar problemas que pueden no ser visibilizados en el análisis cuantitativo de los resultados obtenidos, ser coincidentes o complementarlo.

Las preguntas se presentarán en los títulos de acuerdo con su unidad de análisis y su número de pregunta. Las variables del cuestionario del bloque Vivienda contarán con una “V” delante de su número de pregunta; las variables del cuestionario del bloque Características habitacionales del hogar y de la vivienda se mostrarán con una “H” y las variables del cuestionario del bloque Población se distinguirán con una “P”. La numeración de las preguntas corresponde al cuestionario A para facilitar la lectura y el análisis de los cuadros.

## 7.2 Bloque Vivienda (V1 a V7)

En los R9 y R11, para el Tipo de vivienda (V1) se registraron menores porcentajes de error que en el Resto de Humahuaca y el Total de jurisdicciones de la prueba.

Para el Uso de la vivienda, en el R11 se observan porcentajes altos de error acompañando la tendencia dada en el Resto de Humahuaca y el Total de jurisdicciones (que ronda el 50%). En el R9 se registra un porcentaje de error comparativamente menor respecto de las otras áreas de prueba.

Las preguntas que integran la identificación de hogares presentan niveles altos de error en todas las áreas. No obstante, los censistas del R11 ostentan valores inferiores para algunas variables respecto del Total de jurisdicciones y el Resto de Humahuaca.

El R9 evidencia valores de error total superiores en las preguntas de indagación de hogares, a excepción de la pregunta V5 que muestra valores similares al Resto de Humahuaca y el Total de jurisdicciones (6,0%).

**Cuadro 7.2.1 Variables de Vivienda (V1 a V7). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Tipo de vivienda (V1)	8,5	10,0	6,4	4,19
Uso de la vivienda (V2)	49,9	47,8	34,1	51,2
Causa por la que está deshabitada (V3)	25,2	17,5	30,0	12,5
Cuántas personas pasaron la noche (V4)	13,5	12,4	19,8	8,1
Gastos de alimentación (V5)	6,4	6,1	6,0	14,1
Cantidad de hogares (V6)	20,1	25,6	8/24	2/26
Número de hogar (V7)	17,2	15,4	31,0	8,8

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Para complementar el análisis cuantitativo, podemos destacar que en las observaciones no participantes del R9 se detectaron problemas en la administración de las preguntas V2 y V3 derivadas de dificultades en el registro conjunto de la condición de ocupación de la vivienda (“hay personas presentes”–“no hay personas presentes”) y el uso de esta (principal, vacaciones/fin de semana, etc.). Asimismo, se observaron dificultades en la administración de las preguntas relativas a la determinación de la cantidad de hogares que ocupan la vivienda (V4 a V7). Los censistas observados no siempre respetaban los pasos de la detección de hogares. Las preguntas con frecuencia se reformulaban o acortaban. Apelaron al concepto de “familia” o de “olla” compartida y faltó énfasis en la idea de “las personas que pasaron la noche”, ya que lo asimilaban a “las personas que viven acá”. Se observó que el casillero de cantidad de hogares se omitía con relativa frecuencia. En el R11 también se constataron dificultades o confusiones en la administración de las preguntas V1 y V2 de vivienda.

## 7.3

### Bloque Características del hogar y de la vivienda (A: H9 a H28 y B: H9 a H19)

Para las preguntas sobre materiales de construcción, condiciones sanitarias, cocina, combustible y cantidad de habitaciones de la vivienda, los porcentajes de error en los R9 y R11 fueron levemente menores o del mismo orden que en el Resto de Humahuaca y el Total de Jurisdicciones de la prueba.

De las 17 preguntas presentadas en el cuadro 7.3.1, el R9 tiene porcentajes de error menores que el Resto de Humahuaca en 12 de ellas; y el R11, en 13. Si la comparación es con el Total de jurisdicciones, tanto el R9 como el R11 cuentan con 10 variables con porcentajes de error menores.

**Cuadro 7.3.1 Variables de Características del hogar y de la vivienda (A: H9 a H25) (B: H9 a H19). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Material predominante de los pisos (H9)	4,7	5,7	5,4	4,9
Material predominante de las paredes exteriores (H10)	7,0	7,7	5,4	5,5
Revestimiento externo o revoque (H11)	3,5	3,6	1,0	2,6
Material predominante de la cubierta exterior del techo (H12)	6,4	6,9	7,1	7,9
Revestimiento interior o cielorraso (H13)	5,8	7,0	5,4	5,5
Provisión del agua (H14)	6,2	7,5	9,8	4,3
Procedencia del agua (H15)	4,9	6,2	1,8	5,5
Disponibilidad y ubicación del baño o letrina (H16)	5,2	6,6	4,5	5,5
Disponibilidad de inodoro con botón, mochila o cadena (H17)	2,0	2,1	0,9	3,3

(continúa)

**Cuadro 7.3.1 (conclusión)**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Tipo de desagüe del inodoro (H18)	2,5	2,6	1,9	4,6
Baño de uso compartido o de uso exclusivo del hogar (H19)	4,4	4,7	5,6	3,3
Disponibilidad de cuarto de cocina (H20)	6,4	6,8	4,5	6,1
Combustible utilizado principalmente para cocinar (H21)	6,5	7,5	7,1	9,2
Disponibilidad de agua caliente (H22)	5,7	5,9	5,4	5,5
Cantidad de habitaciones (H23)	4,6	5,3	5,4	1,8
Superficie cubierta de la vivienda (H24)	7,9	8,2	6,3	7,3
Régimen de tenencia de la vivienda (H25)	5,4	6,8	7,1	4,3

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B. La numeración de las preguntas corresponde al cuestionario A.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

De los registros de las observaciones no participantes del R9, se destaca que en la pregunta H23 existieron problemas de comprensión, tanto por parte de los censistas como de las personas censadas. En algunas entrevistas se advirtió que confundían el objetivo de la pregunta, ya que se declaraba el total de piezas para dormir y no el total de ambientes de la vivienda. En la pregunta H24, se observaron dificultades para determinar la superficie cubierta de la vivienda en metros cuadrados. Las personas censadas desconocen los metros cuadrados precisos de la vivienda o bien responden el ancho por el largo del terreno.

De todas maneras, los datos cuantitativos demuestran que para las preguntas H23 y H24 se han obtenido similares niveles de error en todas las áreas de prueba o incluso inferiores. Cabe destacar que en el R11 la determinación de la Cantidad de habitaciones (H23) registra el error total más bajo de todas las jurisdicciones, inferior al 2%.

En cuanto a los datos correspondientes a las preguntas sobre climatización y equipamiento del hogar, el cuadro 7.3.2 muestra que los R9 y R11 registraron porcentajes de error superiores al Resto de Humahuaca y al Total de jurisdicciones, en la mayoría de las preguntas. Sin embargo, los errores más elevados se observan en el R9.

**Cuadro 7.3.2 Climatización y equipamiento del hogar (H26 y H27). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Aire acondicionado frío-calor (H26.1)	15,2	13,9	25,9	23,8
Aire acondicionado solo frío (H26.2)	17,6	16,6	26,8	22,0

(continúa)

**Cuadro 7.3.2 (Conclusión)**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Calefacción central (H26.3)	18,5	17,2	25,9	25,0
Estufa a gas (H26.4)	16,5	15,6	25,0	16,5
Estufa eléctrica (H26.5)	15,6	14,6	22,3	22,6
Estufa a lecha, carbón o querosén (H26.6)	17,4	16,5	28,6	26,2
Panel solar (H26.7)	17,3	15,9	28,6	22,0
Heladera con <i>freezer</i> / <i>freezer</i> solo (H27.1)	10,1	10,5	9,8	7,9
Heladera sin <i>freezer</i> (H27.2)	19,9	19,7	25,0	22,5
Computadora (PC, <i>notebook</i> , <i>tablet</i> , etc.) (H27.3)	11,4	11,0	16,1	12,8
Internet en la vivienda (H27.4)	12,7	11,8	20,5	18,9
Teléfono móvil con acceso a internet (H27.5)	10,2	10,5	13,4	8,5
Teléfono móvil sin acceso a internet (H27.6)	16,4	15,6	33,0	16,5

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En relación con las preguntas sobre cantidad de vehículos, los datos cuantitativos de ambos radios arrojan porcentajes de error similares a los registrados para el Resto de Humahuaca y el Total de jurisdicciones.

**Cuadro 7.3.3 Cantidad de vehículos (H28). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Cantidad de vehículos (H28.1)	29,9	31,7	32,1	31,1
Año del último vehículo (H28.2)	33,1	35,0	28,9	22,4
Cantidad de motocicletas (H28.3)	43,2	44,2	40,2	55,5
Cantidad de ciclomotores (H28.4)	48,1	49,4	46,4	59,8
Cantidad de bicicletas (H28.5)	41,2	42,6	38,4	49,4

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En las observaciones de campo de los censistas del R9 se observó, para las preguntas que solicitan marcas en cada alternativa (preguntas H26, H27, H28), que a menudo se marcaban solamente las respuestas afirmativas, dejando en blanco las opciones restantes. Las preguntas sobre tenencia de vehículos (cuadro 7.3.3) con frecuencia no fueron leídas textualmente. Adicionalmente, en la pregunta H28.2 (Año del último vehículo) no se indagaba correctamente si era el año de compra o del modelo del vehículo.

De lo expuesto se deduce que los errores relevados en el desempeño de los censistas del R9 y R11 para las preguntas referidas a las Características del hogar y de la vivienda no presentan diferencias sustanciales respecto de censistas de las otras áreas de prueba.

Para las preguntas sobre materiales de construcción de la vivienda, condiciones sanitarias, combustible, agua caliente y cantidad de habitaciones, los porcentajes de error en los R9 y R11 fueron levemente menores o de calidad similar al Resto de Humahuaca y el Total de jurisdicciones. En la mayoría de las preguntas relativas a climatización, equipamiento y TIC, ambos radios registraron porcentajes de error más elevados.

## 7.4 Bloque Población

### 7.4.1 Relación de parentesco, sexo, identidad de género, edad y fecha de nacimiento (P1 a P4)

En este conjunto de preguntas, el R9 presenta niveles de error inferiores que el Resto de Humahuaca y el Total de jurisdicciones.

El R11 registra porcentajes de error también menores en comparación con las mismas áreas para las preguntas de Relación de parentesco, Edad y Fecha de nacimiento, pero mayores para Sexo e Identidad de género.

**Cuadro 7.4.1.1 Relación de parentesco, sexo, identidad de género, edad y fecha de nacimiento (P1 a P4). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Relación de parentesco (P1)	1,6	1,6	1,2	1,3
Sexo (P2)	1,5	1,5	0,6	2,7
Identidad de género (P3)	1,7	1,9	0,9	3,6
Edad (P4)	2,6	2,7	2,1	2,2
Día de nacimiento (P4.1)	4,9	4,7	4,2	3,6
Mes de nacimiento (P4.2)	4,9	4,5	4,2	3,8
Año de nacimiento (P4.3)	9,1	8,9	7,3	7,2

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

### 7.4.2 Educación (A: P5 a P10 y B: P5 a P8)

Con relación a las preguntas de educación, los porcentajes de error para el R9 son mayores respecto del Resto de Humahuaca y el Total de jurisdicciones en cuatro de las seis variables. Para el R11 solo una de las preguntas registró porcentajes menores al resto de las áreas. Por lo tanto, puede concluirse que existieron problemas de captación de estos indicadores para ambos radios, especialmente el R11.

**Cuadro 7.4.2.1 Educación (A: P5 a P10) (B: P5 a P8). Error total. Cuestionario básico y ampliado y. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Condición de alfabetismo (P5)	4,7	4,8	3,9	6,3
Asistencia actual o pasada (y Sector del establecimiento) (P6)	5,5	5,3	7,3	6,7
Nivel educativo que cursa o cursó (P7)	3,5	3,2	3,1	7,7
Completitud del nivel (P8)	2,4	2,5	5,0	2,3
Grado o año que está cursando o aprobó (P9)	11,7	11,9	14,7	27,8
Carrera universitaria de grado o terciaria que completó (P10)	34,6	36,4	14,4	32,5

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B. La numeración de las preguntas corresponde al cuestionario A.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En las observaciones no participantes del R9 se habían detectado problemas en la administración de la pregunta P6, centradas en la dificultad de determinar si debía privilegiarse la cursada actual o pasada en un establecimiento educativo. En las preguntas P7 y P8, se observaron situaciones donde los censistas privilegiaron la completitud del nivel y marcaban el nivel educativo correspondiente a la asistencia pasada completa, aunque se cursara actualmente. Asimismo, se observaron algunas situaciones donde se registraban todos los niveles educativos realizados (error de multimarca).

### 7.4.3

### Migraciones (A: P11 a P13 y B: P9)

Respecto de las preguntas sobre Lugar de nacimiento, Lugar de residencia 5 años antes y Lugar de residencia habitual (las tres preguntas cerradas), el R9 presenta errores notablemente mayores que el Resto de Humahuaca y que el Total de jurisdicciones. Por el contrario, en el R11 se registran porcentajes de error más bajos que el Total de jurisdicciones en cuatro preguntas; y porcentajes más bajos que el Resto de Humahuaca, en tres.

**Cuadro 7.4.3.1 Migración (A: P11 a P13) (B: P9). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En valores absolutos y porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Lugar de nacimiento (P11.1)	33,7	43,4	52,7	26,5
Provincia de nacimiento (P11.2)	0,3	0,2	1,3	-
País de nacimiento (P11.3)	1,1	6,4	1/3	-
Año de llegada (P11.4)	26,6	58,7	2/3	6/7

(continúa)

**Cuadro 7.4.3.1 (conclusión)**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Lugar de residencia 5 años antes (P12.1)	38,2	45,0	52,4	29,4
Provincia de residencia 5 años antes (P12.2)	0,5	0,3	2,0	-
Lugar de residencia habitual (P13.1)	39,3	46,5	55,2	28,7
Provincia de residencia habitual (P13.2)	0,5	0,2	4,1	0,3
País de residencia habitual (P13.3)	3/7	3/7	-	-

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B. La numeración de las preguntas corresponde al cuestionario A.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

La evaluación cualitativa del R9 es consistente con los datos cuantitativos que se muestran en el cuadro 7.4.3.1. En la observación no participante se detectó que algunos censistas escribían el nombre de la provincia, pero no marcaban previamente la opción “En otra provincia”.

#### 7.4.4 Discapacidad y Previsión social (A: P 14 a P16)

Para las preguntas sobre dificultad o limitación, el R9 y el R11 han registrado errores que en ciertos casos duplican y en otros hasta triplican los porcentajes del Resto de Humahuaca y el Total de jurisdicciones. Los porcentajes más altos correspondieron al R9.

**Cuadro 7.4.4.1 Dificultad o limitación (A: P14). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Dificultad o limitación para caminar o subir escaleras (P14.1)	4,0	3,6	13,0	8,8
Dificultad o limitación para entender, recordar o concentrarse (P14.2)	4,6	4,1	13,0	9,2
Dificultad o limitación para hablar o comunicarse, aun usando lengua de señas (P14.3)	4,6	4,1	12,4	9,6
Dificultad o limitación para oír, aun cuando usa audífono (P14.4)	5,0	4,4	13,3	9,2
Dificultad o limitación para ver, aun cuando usa anteojos o lentes de contacto (P14.5)	5,0	4,4	14,6	9,8
Dificultad o limitación para comer, bañarse o vestirse sola(o) (P14.6)	4,9	4,3	13,6	8,7

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En las observaciones no participantes se había constatado que las seis preguntas sobre dificultad o limitación no siempre se leían textualmente. Se observó que los censistas reformulaban y acortaban la pregunta por “¿tiene alguna dificultad?” o “¿tiene alguna discapacidad?”, sin mencionar el nombre de persona que se estaba censando. Se observó que frecuentemente no se indagaba el grado de dificultad y se registraban solamente las respuestas afirmativas y no las negativas.

En las preguntas sobre Previsión social, el R9 contó con porcentajes de error más altos que el Resto de Humahuaca y el Total de jurisdicciones; mientras que el R11 obtuvo un error porcentual más bajo para la P15 y más alto para la P16 respecto del resto de las áreas de prueba.

**Cuadro 7.4.4.2 Previsión social (A: P15 y P16). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
¿Cobra jubilación o pensión? (P15)	6,4	6,5	12,7	4,0
Tipo de jubilación (P16)	5,4	5,0	9,4	7,1

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

## 7.4.5 Etnias

El error en la pregunta sobre Pertenencia étnica es alto en todas las áreas de prueba. Alcanza su mayor valor en el Resto de Humahuaca (11,4% de la población censada), siendo los radios 9 y 11 donde se registran valores inferiores en comparación.

**Cuadro 7.4.5.1 Etnias (A: P17) (B: P10). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Etnias (P17.1)	10,7	11,4	9,1	9,6
Pueblo indígena u originario (P17.2)	6,1	5,0	4,5	3,4

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B. La numeración de las preguntas corresponde al cuestionario A.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En la pregunta que capta el nombre del Pueblo indígena u originario para la población que se autorreconoce indígena o descendiente de pueblos indígenas u originarios, también los radios 9 y 11 muestran niveles de error menores a las restantes áreas de prueba (4,5% y 3,4%, respectivamente).

#### 7.4.6

#### Número de Documento Nacional de Identidad (DNI) (A: P18 y B: P11)

**Cuadro 7.4.6.1 DNI (A: P18) (B: P11). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
DNI (P18)	6,2	5,8	3,9	3,6

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B. La numeración de las preguntas corresponde al cuestionario A.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

El nivel de error para la pregunta sobre el DNI en los radios 9 y 11 es inferior al 4% de la población censada, siendo las áreas con los valores más bajos en esta variable.

En las observaciones no participantes de los radios 9 y 11 también se destacó que los censistas no presentaron demasiados inconvenientes para indagar y registrar el DNI de la población censada. Las personas entrevistadas brindaban el número de documento sin dudas o incomodidad, y resolvían con facilidad cuando no recordaban o conocían la información de los restantes integrantes del hogar.

#### 7.4.7

#### Situación conyugal (A: P20)

Respecto a la situación de las personas con relación a las uniones legales y consensuales, el error total en el R9 alcanza el 4,4%, siendo el más bajo en comparación con todas las áreas de prueba. Asimismo, el R11 registra un porcentaje de error también menor con relación al Resto de Humahuaca y el Total de jurisdicciones.

**Cuadro 7.4.7.1 Situación conyugal (A: P20). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Situación conyugal (P20)	6,1	6,4	4,4	5,6

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En las observaciones de campo se registraron dificultades para comprender e identificarse con alguna de las opciones de respuesta, especialmente en el caso de las “uniones”, tanto por parte de las personas censadas como de los censistas.

#### 7.4.8

#### Actividad económica (A: P21 a P33 y B: P12 a P16)

En cuanto a las preguntas sobre la actividad económica de las personas de 14 años y más, en general se observa que el desempeño de los censistas del R9 fue mejor que en las otras áreas de prueba. En el R11, el rendimiento fue más alto solo en tres de las diez variables asociadas a la temática.

**Cuadro 7.4.8.1 Variables de actividad económica (A: P21 a P31) (B: P12 a P16). Error total. Cuestionario básico y ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Trabajó (P21)	7,2	8,1	4,0	6,8
Hizo alguna changa (P22)	7,9	8,4	5,4	12,0
Licencia (P23)	5,9	6,2	6,8	8,5
Buscó trabajo (P24)	6,6	6,5	12,6	9,3
Percepción de pago (P25)	6,7	8,3	4,2	7,1
Categoría ocupacional (P26)	10,1	10,8	7,2	11,9
Descuentos jubilatorios (P27)	3,8	4,0	1,7	2,2
Aportes jubilatorios (P28)	8,0	8,0	10,2	15,0
Ocupación (P29)	9,0	10,0	4,8	10,0
Tarea (P30)	11,1	11,8	6,0	12,9
Rama de actividad económica (P31)	20,6	21,2	15,1	21,4

**Nota:** en Humahuaca (radios 9, 11 y Resto de Humahuaca) se utilizó cuestionario A; en el Total de jurisdicciones se utilizaron cuestionarios A y B. La numeración de las preguntas corresponde al cuestionario A.

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

Con respecto a las preguntas del cuadro 7.4.8.2, los valores de error que se produjeron en los R9 y R11 en general no distaron de los registrados en el Resto de Humahuaca y en el Total de jurisdicciones.

**Cuadro 7.4.8.2 Dirección del lugar de trabajo y Tamaño del establecimiento (A: P32 y P33). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Dirección del lugar de trabajo precodificada (P32.1)	23,7	23,0	29,5	17,1
Dirección (P32.2)	20,9	21,5	4,9	21,4
Número (32.3)	63,9	62,6	65,9	69,9
Localidad (P32.4)	6,8	4,9	-	4,9
Municipio (P32.5)	21,9	16,6	32,9	16,5
Provincia (P32.6)	9,0	3,2	2,4	-
Tamaño del establecimiento (P33)	11,4	12,0	5,4	15,7

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En las observaciones no participantes, se detectaron dificultades en la administración de los pases y en la comprensión de las preguntas tanto por parte de los censistas como de las personas censadas, especialmente en la indagación sobre Condición de actividad (P21 a P24) y en la pregunta sobre Categoría ocupacional (P26). En la P31, se constataron problemas para describir la actividad del establecimiento; y en la P32, dificultades para declarar la información completa sobre el domicilio del lugar de trabajo.

## 7.4.9

### Fecundidad (A: P34 a P36)

En el cuadro 7.4.9.1 se observan mayores porcentajes de error para el R9 respecto al Resto de Humahuaca y al Total de jurisdicciones en la mayoría de las variables (con excepción de la P35). Se evidencian más dificultades para registrar la Cantidad de hijas o hijos vivos actualmente, ya que el error casi duplica al Resto de Humahuaca.

Por el contrario, el R11 muestra un menor nivel de error para todas las preguntas del bloque.

**Cuadro 7.4.9.1 Fecundidad (A: P34 a P36). Error total. Cuestionario ampliado. Total y Resto de Humahuaca, radios 9 y 11. II PP 2019. En porcentajes**

Pregunta	Total de jurisdicciones	Resto de Humahuaca	R9	R11
Tenencia y cantidad de hijas o hijos nacidos vivos (P34.1 y P34.2)	10,3	10,7	12,8	9,8
Cantidad de hijas o hijos vivos actualmente (P34.3)	8,2	7,4	13,6	1,4
Año de nacimiento de la primera hija(o) nacido vivo (P35)	13,9	14,1	12,6	11,7
Mes de nacimiento de su última hija o hijo nacido vivo (P36.1)	12,0	11,9	13,6	11,0
Año de nacimiento de su última hija o hijo nacido vivo (P36.2)	13,2	13,3	17,5	9,7

**Fuente:** INDEC. Base de datos con estructura jerárquica y sin consistir. II PP 2019.

En las observaciones no participantes también se advirtieron omisiones en el registro de estas preguntas, como así también errores de secuencia.

## 7.5

### Conclusiones

El análisis cuantitativo de los datos de los radios 9 y 11, enriquecido por el resultado de las observaciones no participantes, permite extraer algunas conclusiones con respecto a las estrategias de reclutamiento y capacitación de los censistas.

Para el R9, constituido por censistas que fueron capacitados exclusivamente con el dispositivo de emergencia de 60 minutos de duración, si bien los errores registrados fueron superiores en la mayoría de las variables respecto del resto de las áreas de prueba, podría considerarse que su desempeño fue aceptable, en un contexto de niveles de error más altos de lo esperado en todas las áreas de prueba.

Con relación al R11, los datos relevados por estudiantes de nivel terciario, quienes participaron en la capacitación virtual y presencial, obtuvieron porcentajes de error que acompañan los registrados en el Resto de Humahuaca y en el Total de jurisdicciones. Si esta evidencia se integra con los registros obtenidos en la observación no participante, puede afirmarse que la estructura de nivel terciario resulta una opción de reclutamiento viable para completar la estructura tradicional de docentes primarios.

## 8.1

### Recomendaciones relativas a organizar una base de datos con estructura jerárquica

Se sugieren las siguientes acciones para organizar la base de datos a partir de realizar el proceso de estructura jerárquica en menos tiempo y con una mejor calidad:

- Reformular en el **diseño conceptual del cuestionario** tanto las preguntas relativas a la identificación de las viviendas con personas presentes y sin personas presentes como la identificación de hogares, de manera de simplificar la tarea de censistas.
- Reforzar la **capacitación de censistas y jefes de radio** en relación con la importancia del correcto completamiento de los datos de ubicación geográfica, el número de la vivienda en el orden de recorrido de censistas, el número de hogar al interior de la vivienda y el número de persona; recalcar el procedimiento a realizar cuando el hogar cuenta con más de 6 o 7 personas, y cuán imprescindible es mantener en orden los cuestionarios en el sobre y también en el de los materiales completados que se incluyen en la caja del jefe o la jefa de radio.
- Verificar, en el momento de realización de la tarea, la correcta **impresión de los cuestionarios censales** mediante un procedimiento de evaluación específica basado en una muestra representativa, de manera de asegurar que cada cuestionario en papel, sin importar su tipo (básico o ampliado), cuente con un **código de barra único e irrepetible**.
- En cuanto al **proceso de lectura e interpretación de los datos**:
  - Garantizar que los cuestionarios sean leídos en el orden establecido. Para ello se debe asegurar su correcto ordenamiento durante el trabajo de campo de los censistas y en particular al momento del cierre del operativo. Asimismo, interesa estar atento y monitorear el proceso de guillotinado previo a iniciar el proceso de escaneo, verificando también el orden en el que los cuestionarios vienen organizados en las cajas enviadas desde cada sede del operativo.
  - Respetar las posiciones<sup>40</sup> indicadas en el diseño de registro de las variables del cuestionario.
  - Incorporar la utilización de los diccionarios geográficos y la tabla de segmentación<sup>41</sup> generados por el INDEC para asegurar que los campos de ubicación geográfica se encuentren correctamente completados cuando corresponde y sean homogéneos.
  - Garantizar que las preguntas abiertas<sup>42</sup> tengan caracteres numéricos o alfanuméricos, según corresponda.
  - No permitir que la base de datos contenga símbolos en las variables numéricas.

<sup>40</sup> Refiere al lugar que ocupa cada campo dentro del diseño numerado. Si el campo está formado por un grupo de caracteres, se indica siempre la posición que ocupa el primer carácter del grupo.

<sup>41</sup> La segmentación es el proceso mediante el cual se subdividen los radios en unidades operativas menores denominadas “segmentos” (carga de trabajo de cada censista). Una adecuada segmentación permite cubrir operativamente la totalidad del radio, estimar la cantidad de censistas, distribuir equitativamente la carga de trabajo, calcular y distribuir los materiales a utilizar. Este proceso es realizado por el área de cartografía de cada dirección provincial con normativas establecidas y metodologías sugeridas por la Coordinación del Sistema Geostadístico del INDEC.

<sup>42</sup> El cuestionario censal contiene dos tipos de preguntas: cerradas y abiertas. Una pregunta está compuesta por opciones de respuesta establecidas previamente, entre las cuales la persona censada debe elegir. En una pregunta abierta se responde utilizando palabras propias.

- Implementar acciones tendientes a que un mismo segmento censal<sup>43</sup> cuente con exactamente los mismos datos en los campos Provincia, Departamento, Localidad, Fracción, Radio, Segmento y que no existan datos en blanco o inconsistentes.

## 8.2 Recomendaciones respecto del proceso de escaneo, digitalización y captura de los datos

- Establecer un límite de nivel del error tolerado por el INDEC en marcas, letras y números.
- Incorporar la utilización de diccionarios geográficos para disminuir el porcentaje de error en las preguntas abiertas.
- Incorporar la tabla de segmentación en el momento de la interpretación de los datos de ubicación geográfica.
- Realizar el escaneo de manera ordenada para que la tabla de segmentación se pueda utilizar para precisar los datos de la variable ubicación geográfica de la vivienda.
- Homogeneizar la intensidad del escaneo a través de la utilización del mismo lápiz por parte de quien censa. Este lápiz deberá ser sugerido por la empresa responsable del proceso de lectura e interpretación. Asimismo, se sugiere reforzar la capacitación respecto del completamiento de los cuestionarios y las planillas.
- Incorporar la experiencia de esta prueba piloto para el entrenamiento del software de detección de dígitos y letras/palabras. Se recomienda incorporar, también, la detección de letras del idioma castellano, como por ejemplo las “Ñ” (en la base de datos recibida, todas las “Ñ” fueron reemplazadas por “N”); y mayores variantes de caligrafía para cada letra y número, para mejorar la interpretación de los trazos.
- Mejorar la interpretación de frases en el mismo renglón y ante cambios de renglón. Por ejemplo, en el formulario A0000002139, carilla 5, en la pregunta por la Tarea que realiza, se puede leer en la imagen: “MANEJA COLECTIVO/ DE PASAJEROS / /” (las barras / señalan el cambio de línea). Mientras que la base de datos tiene “MANEJA COLECTIVO DE PASAJEROS”, sin respetar el espacio, probablemente interpretando palabras por línea (y, por lo tanto, borrando un espacio inicial sin aparente utilidad en la segunda línea).

---

<sup>43</sup> El segmento es la porción de territorio o área de trabajo asignada a cada censista que participa en el operativo. Los segmentos son subdivisiones de un radio. La característica de estas subdivisiones es que no se superponen. Por lo tanto, si se unieran todos los segmentos del país, reconstuiríamos el mapa de la República Argentina.

## 8.3

# Recomendaciones relativas al diseño conceptual

### 8.3.1

## Tipo de vivienda, Condición de ocupación y Detección de hogares (V1 a V7)

En consonancia con lo observado mediante la evaluación cualitativa implementada durante la II PP,<sup>44</sup> las preguntas con mayor nivel de error son las relativas al Uso de la vivienda (principal o secundaria) y la Condición de deshabitación de la vivienda. Las viviendas habitadas registran un alto nivel de error en las preguntas referidas a la captación de hogares, en particular la Cantidad de hogares cuando se trata de más de un hogar en la vivienda.

En cuanto a las áreas de prueba, se observa que el error se distribuye de diferente modo para cada una de las preguntas.

En aquellas en las que se aplicaron los dos cuestionarios, el básico presenta mayor nivel de error en la mayoría de las preguntas de la primera página. La única pregunta que presenta un nivel de error más alto en el cuestionario A es la referida a la Condición de deshabitación de la vivienda.

Respecto al error de secuencia y en consonancia con los problemas detectados en la operacionalización de la detección de hogares, estas son las preguntas que mayor nivel de error presentaron.

En síntesis, las preguntas correspondientes a la primera página del cuestionario registran mayor nivel de error total que el resto de los bloques de preguntas del cuestionario de vivienda-hogar.

### 8.3.2

## Recomendaciones

- Transformar la pregunta sobre el Tipo de vivienda, que se debía registrar por observación, en una instrucción “Anote por observación el tipo de vivienda”.
- Eliminar la pregunta relativa al Uso de la vivienda.
- Modificar la operacionalización de la pregunta que refiere a la Condición de deshabitación de la vivienda, a través de la indagación de dos preguntas: una pregunta sobre la presencia o ausencia de personas en la vivienda en el momento de la entrevista y una instrucción sobre el motivo por el que no se encuentran personas presentes.
- Modificar la operacionalización de la captación de hogares.
- Modificar el pase para evitar las preguntas sobre captación de hogares para los hogares unipersonales.
- Incluir un grisado con el significado de compartir los gastos de alimentación aludiendo a que “compartir” refiere al presupuesto común independientemente del aporte y a que los “gastos de alimentación” corresponden a la “olla compartida”.
- Modificar el pase para evitar registrar el Número de hogar (01) a las personas que comparten los gastos de alimentación.

<sup>44</sup> Ver INDEC (2019b, p. 63).

- Transformar la pregunta sobre la Cantidad de hogares en la vivienda por la instrucción “Anote la cantidad de grupos de personas que tienen gastos separados para la alimentación” e incluir la aclaración “Si en la vivienda pasaron la noche visitas considere que comparten los gastos para la alimentación”.
- Visibilizar el espacio para registrar el Número de hogar colocando las instrucciones al final.

### 8.3.3 Otras modificaciones

A continuación, se presentan recomendaciones generales que surgen de la necesidad de facilitar la tarea de quienes censan y obtener información más válida y confiable:

- Incorporar instrucciones en la primera página del cuestionario.
- Incluir, en las preguntas que tienen pases, también el que indica continuar en la pregunta siguiente.
- Modificar la captación de piezas en inquilinato, hotel familiar o pensión; local no construido para habitación y viviendas móviles, incorporando al final de cada categoría la condición de ocupación “ocupada”.
- Eliminar del cuestionario la indicación “Deje Aviso de Visita”, dado que el aviso de visita se eliminó del operativo.<sup>45</sup>
- Incorporar el recordatorio “incluya a los bebés, niñas(os) y personas mayores” en la pregunta que indaga por la cantidad de personas del hogar.
- Simplificar la pregunta de la lista de personas que pasaron la noche en el hogar e incluir una aclaración “incluya a los bebés, niñas(os) y personas mayores”.
- Disminuir las aclaraciones e instrucciones que se encuentran en la lista de personas del hogar.

## 8.4 Materiales de los pisos, las paredes y los techos (A: H9 a H13) (B: H9 a H11)

Las preguntas relativas a los materiales constructivos de la vivienda alcanzaron mayor porcentaje de error que lo esperado a partir de censos anteriores. Ello resulta consistente con la evaluación cualitativa realizada,<sup>46</sup> en la que se observó que la población solicitaba la lectura de las categorías a fin de orientar su respuesta.

Las preguntas que registran mayor nivel de error son las que corresponden a los Materiales del exterior de la vivienda: Cubierta exterior del techo y Revestimiento externo de las paredes.

<sup>45</sup> Para conocer los fundamentos de esta recomendación, ver INDEC (2019b).

<sup>46</sup> Ver INDEC (2019b, p. 64).

Si bien el nivel de error es homogéneo para todas las jurisdicciones, Humahuaca presenta valores superiores al resto de las áreas de prueba y el cuestionario A es el que presenta magnitudes más elevadas.

En cuanto al error de secuencia, se observa que la mayoría de los censistas continuaron indagando sobre la existencia de Revestimiento externo en materiales constructivos de paredes para las que no correspondía indagar.

#### 8.4.1 Recomendaciones

- Transformar el enunciado de las preguntas referidas a los materiales constructivos de la vivienda en una preguntas guiadas a través de la lectura de las opciones de respuesta. De este modo se pretende orientar a la persona censada en la respuesta.
- Incluir una aclaración en texto grisado con el objetivo de reforzar el concepto de “predominante” para que el censista pueda aplicarlo cuando la persona censada responde por más de un material constructivo.
- Incluir todos los pasos indicando la secuencia a seguir según cada una de las respuestas, de manera de orientar mejor al censista en la lectura de las preguntas que corresponde realizar.

### 8.5 Condiciones sanitarias, Combustible y Agua caliente (A: H14 a H22) (B: H12 a H17)

Las preguntas sobre condiciones sanitarias de la vivienda y del hogar alcanzaron mayor porcentaje de error que en los dos últimos censos. Humahuaca es la jurisdicción con mayor nivel de error y en las áreas donde se aplicó muestra en el relevamiento, el cuestionario A es el que muestra mayor nivel de error.

Por último, en aproximadamente 2 de cada 10 viviendas de Humahuaca los censistas indagaron las condiciones sanitarias del baño a hogares sin baño.

#### 8.5.1 Recomendaciones

- Incluir una aclaración en texto grisado con el concepto “principalmente” para que quien censa pueda aplicarlo cuando la persona censada responde por más de un combustible usado para cocinar.
- Incluir una aclaración en texto grisado que indique que el tipo de agua caliente al que se refiere la pregunta es la que se usa para la higiene personal.
- Incluir todos los pasos indicando la secuencia a seguir según cada una de las respuestas, de manera de orientar mejor a quienes censan en la lectura de las preguntas que corresponde realizar.

## 8.6 Cantidad de habitaciones, Superficie cubierta y Régimen de tenencia de la vivienda (A: H23 a H25) (B: H18 a H19)

Las preguntas sobre la Cantidad de habitaciones totales y el Régimen de tenencia de la vivienda por parte del hogar también alcanzaron mayor porcentaje de error que en los dos últimos censos. Humahuaca es la jurisdicción de la prueba piloto con mayor nivel de error y, en las áreas de prueba donde se aplicó muestra en el relevamiento, el cuestionario A concentra un error superior al doble en comparación con el cuestionario B.

En consonancia con la evaluación cualitativa que proviene de la observación no participante,<sup>47</sup> el principal problema que se registra en la pregunta Superficie cubierta de la vivienda refiere al error que se origina en el desconocimiento de la respuesta por parte de las personas censadas y que figura captada a partir de la marca en la categoría de respuesta ignorado.

### 8.6.1 Recomendaciones

- Mejorar la captación de la Cantidad de habitaciones totales a través de la inclusión de la pregunta sobre cantidad de cuartos para dormir. De este modo, además, se evitará el subregistro detectado en la evaluación cualitativa,<sup>48</sup> de cantidad de habitaciones totales del hogar, por parte de las personas censadas, al declarar solo los dormitorios y no la cantidad total de habitaciones.
- Modificar la enunciación de la pregunta Superficie cubierta de la vivienda indagando por la cantidad de metros cuadrados e introducir categorías de respuesta en intervalos que orienten a la persona censada. De este modo, no se requiere que responda con precisión la superficie cubierta exacta de la vivienda, sino la cantidad de metros cuadrados aproximados dentro de un intervalo determinado.
- Incluir una pregunta sobre la regularidad de la propiedad de la vivienda que mejore la confiabilidad de la respuesta de la persona censada en torno a las particularidades de la tenencia.

## 8.7 Climatización y Equipamiento (A: H26 a H28)

Este bloque de preguntas, junto a las correspondientes a la primera página del cuestionario, son las que presentan las magnitudes más elevadas de error total, en especial en las que indagan por la Cantidad de vehículos (con errores que van entre el 30% y el 50%). Las preguntas referidas al Equipamiento y la Climatización del hogar también concentran un elevado error total.

### 8.7.1 Recomendaciones

- Eliminar las preguntas sobre la Cantidad de vehículos que posee el hogar debido al elevado nivel de error detectado.
- Simplificar la formulación de las preguntas sobre Tenencia de heladera, Aire acondicionado y Teléfono.

<sup>47</sup> Ver INDEC (2019b, p. 65).

<sup>48</sup> Ver INDEC (2019b, pp.64 y 65).

## 8.8 Relación o parentesco con la(el) jefa(e) o persona de referencia del hogar (P1)

La pregunta, que no presenta complejidad conceptual, tuvo un error superior al de los censos anteriores.

### 8.8.1 Recomendaciones

- Si bien la evaluación cuantitativa no indica inconvenientes en esta variable, a partir a la evaluación cualitativa, se rescata que quienes censaron no indagaron si los hijos o las hijas eran solo del jefe o la jefa, del cónyuge o de ambos; por lo tanto, se recomienda eliminar la distinción entre hija(o) solo de jefa(e), solo de cónyuge o de ambos.<sup>49</sup>

## 8.9 Sexo, Edad y Fecha de nacimiento (P2 y P4)

El nivel de error de la pregunta Sexo se encontró dentro de los parámetros observados en censos anteriores. La Edad en años cumplidos y, en especial, la Fecha de nacimiento, muestran errores más elevados, constituidos en parte por los datos inválidos. Ello indica que el mejoramiento de la calidad de la lectura e interpretación disminuiría considerablemente los niveles de error.

### 8.9.1 Recomendaciones

- Si bien la evaluación cuantitativa no indica errores de magnitud en la variable Sexo, dado que la evaluación cualitativa señaló la incomodidad que manifiesta el censista al preguntar el sexo de la persona,<sup>50</sup> se recomienda modificar la formulación, evitando la palabra “sexo” y leyendo directamente las opciones “mujer” y “varón” e incorporando una leyenda que aclare a quien censa que la pregunta refiere al sexo al nacer.
- Se considera necesario, además, que la pregunta de Edad y la de Fecha de nacimiento se incluyan en forma separada a fin de disminuir los niveles de no respuesta.
- Por otro lado, eliminar la utilización del nombre de la persona en todas las preguntas en donde se lo incluyó, dado que mediante la evaluación cualitativa también se observó que produjo una mala formulación y confusiones.

## 8.10 Identidad de género (P3)

El nivel de error de la pregunta Identidad de género se encuentra en valores bajos y similares a los que presenta la variable Sexo (1,7% y 1,4%, respectivamente). Por su parte, la evaluación cualitativa presentada en el primer informe de esta prueba<sup>51</sup> concluyó que esta pregunta provocó incomodidad en la persona censada, en especial cuando se trataba de niñas y niños o personas mayores.

<sup>49</sup> Ver INDEC (2019b, pp. 65 y 66).

<sup>50</sup> Ver INDEC (2019b, p. 66).

<sup>51</sup> Ver INDEC (2019b, p. 66).

## 8.10.1 Recomendaciones

- Indagar la pregunta a la población de 14 años o más de manera de minimizar la incomodidad en quienes censan para formular la pregunta, en especial cuando se trataba de menores de edad o de adultos mayores.
- Modificar el orden de las categorías de respuesta, comenzando por “mujer trans/travesti”, “varón trans”, a fin de mejorar la captación.

## 8.11 Educación (A: P5 a P10) (B: P5 a P8)

A nivel general, el error de todas las preguntas del bloque supera a la de censos anteriores.

La pregunta de Carrera terciaria o universitaria no fue respondida por un tercio de las personas que cursaron esos niveles. Sin embargo, este error disminuye considerablemente en el universo de la población que terminó tanto uno u otro nivel.

A excepción de la Carrera terciaria o universitaria, el error de secuencia oscila entre 7,4% y 18,4% cuando no se distingue por áreas para las preguntas del bloque, lo que indica dificultades en quienes censaron para seguir los filtros y pases.

Del análisis de la implementación de la ampliación del universo a la población de 0 a 2 años surge que este grupo tiene no respuestas considerablemente superiores a las que presentan los mayores de 3 años.

## 8.11.1 Recomendaciones

- Preguntar por la asistencia actual en primer lugar y, posteriormente, en caso de que la persona censada no asista actualmente, preguntar si asistió en el pasado.
- Indagar acerca del Sector del establecimiento al que asiste de manera separada de la asistencia.
- Realizar las preguntas de Nivel educativo, Completitud de nivel y Grado o año de manera independiente para quienes asisten y para quienes asistieron, para otorgar claridad en la formulación de las preguntas y, por lo tanto, facilidad al respondiente a la hora de interpretar las preguntas.
- Efectuar la pregunta de Carrera terciaria o universitaria completada solo a personas graduadas de nivel terciario y universitario y a aquellas que asisten al nivel de posgrado o lo completaron.

## 8.12 Migraciones (A: P11 a P13) (B: P9)

Para las preguntas cerradas, el error supera ampliamente a los de censos anteriores, aunque al reconstruir cada una de las variables a partir de considerar las respuestas en las preguntas abiertas de manera integrada, los errores disminuyen considerablemente.

En Humahuaca los errores de secuencia son elevados, en especial para las preguntas sobre Provincia de nacimiento y Residencia habitual. Este dato resulta consistente con lo observado en la evaluación cualitativa, según la cual las personas que nacieron en Humahuaca escribían “Jujuy” en el lugar destinado a anotar la provincia sin haber marcado previamente la opción nació “En otra provincia argentina” diferente a CABA o provincia de Buenos Aires.

### 8.12.1 Recomendaciones

- Con base en estos resultados se sugiere simplificar los pasos y asignar un número de pregunta tanto a las cerradas como a las abiertas.
- Incorporar las opciones “en esta provincia” para el Lugar de nacimiento; “en esta localidad o paraje” y “en otra localidad o paraje de esta provincia” para el Lugar de residencia cinco años antes y “en esta vivienda”, “en otra vivienda de esta localidad o paraje” y “en otra localidad o paraje de esta provincia” para el Lugar de residencia habitual.
- Mantener la distinción específica entre CABA y la provincia de Buenos Aires, ante el conocimiento acumulado respecto de que la población asume y responde “Buenos Aires” como única denominación en el caso de una y otra jurisdicción.

## 8.13 Discapacidad y Previsión social (A: P14 a P16)

A excepción del Tipo de jubilación, las restantes preguntas del bloque que se indagaron en censos anteriores presentan errores superiores en la II PP. Para las seis preguntas sobre Dificultades y limitaciones, el nivel de error más bajo se encuentra en la primera subpregunta, referida a la dificultad para caminar o subir escaleras; mientras que los niveles de error más altos se encuentran en las tres últimas subpreguntas del bloque.

En la evaluación cualitativa ya se había detectado que las preguntas de Dificultades y limitaciones presentaban una particular complicación en el caso de tratarse de niños menores, lo cual fue corroborado cuantitativamente, ya que el nivel de error es significativamente mayor en el grupo etario de 0 a 5 años que en el resto.

### 8.13.1 Recomendaciones

- Reforzar durante la capacitación la indicación de que se deben realizar todas las preguntas de Dificultades y limitaciones a todas las personas. Asimismo, se recomienda revisar la redacción para hacerla más sencilla.
- Incorporar una aclaración respecto de que la pensión no refiere a la Asignación Universal por Hijo (AUH). El fundamento de esta recomendación radica en que, si bien la evaluación cuantitativa no presenta errores de relevancia en la captación, la evaluación cualitativa registró que a menudo la persona censada consideró que la percepción de una pensión hacía referencia a la recepción de algún plan de ayuda –como AUH, por ejemplo<sup>52</sup>.

<sup>52</sup> Ver INDEC (2019b, p. 68).

## 8.14

## Etnias (A: P17) (B: P10)

El error en la pregunta sobre Pertenencia étnica alcanza al 10% de la población, en Humahuaca se registra el mayor nivel de error y en CABA, el menor. El 26,1% de las personas a las que se les aplicó el cuestionario A, y el 12,7% de aquellas a las que se les aplicó el básico respondieron ignorar la respuesta. En ambos casos, esta categoría muestra un valor aproximado al de quienes se reconocen pertenecientes o descendientes de un pueblo indígena u originario y supera ampliamente a quienes se reconocen afrodescendientes.

Además del nivel del error observado en este bloque, de acuerdo a la evaluación cualitativa, el error más frecuente que se observó fue que las opciones de respuesta resultan ser excluyentes y no permiten registrar ambos autorreconocimientos en los casos manifestados por las personas censadas, aspecto que, se considera, ocasiona parte del bajo porcentaje de población que se reconoce afrodescendiente.

Otro error detectado al analizar la pregunta abierta relativa al Pueblo indígena u originario de pertenencia fue que las personas censadas respondieron la localidad de nacimiento en lugar del pueblo indígena u originario (acorde con el elevado porcentaje registrado en esta opción), que ignoraban la respuesta o que no entendieran el significado de “reconocimiento indígena”, lo que resulta acorde con la no respuesta detectada.

### 8.14.1

### Recomendaciones

- Indagar las dos variables en preguntas independientes. Una de ellas destinada a la identificación de la población indígena y otra a la identificación de la población afrodescendiente.
- Reemplazar la pregunta abierta ¿De qué pueblo...? por la forma más específica ¿De qué pueblo indígena...?

## 8.15

## Número de DNI (A: P18) (B: P11)

Se observa un nivel de error más elevado en áreas urbanas de mayor tamaño (CABA y Mar del Plata) que en Humahuaca y San Carlos de Bariloche.

La no respuesta constituye la mayor fuente de error de la variable, aunque en CABA también se observa un nivel elevado de multimarca. Los casos inválidos son una proporción menor en todas las áreas de prueba.

Por su parte, la evaluación cualitativa detectó como mayor dificultad el desconocimiento del número de DNI por parte de las personas censadas que respondían las preguntas relativas a otro integrante del hogar y la consecuente pérdida de tiempo que ocasionaba tener que buscar la documentación.<sup>53</sup>

### 8.15.1

### Recomendaciones

- Continuar incluyendo la pregunta en el cuestionario censal, pero incorporar ajustes para intentar mejorar su captación.
- Ubicar las categorías cerradas debajo de las casillas incluidas para anotar el número de DNI de manera de evitar las multimarcas.

<sup>53</sup> Ver INDEC (2019b, p. 68).

## 8.16 Religión (A: P19)

La pregunta tiene un error total de 5,5%. CABA es la jurisdicción con el nivel más bajo de error. Dos de cada diez personas de 0 a 13 años tienen respuesta en esta pregunta, a pesar de que solo debía ser respondida por aquellas de 14 años o más, lo cual fue observado también en la evaluación cualitativa.<sup>54</sup> En ella se observaron dificultades conceptuales por parte de la población en relación con la religión cristiana y a los ateos y agnósticos.

### 8.16.1 Recomendaciones

- No incluir esta pregunta en el cuestionario censal, dados los problemas conceptuales detectados durante la situación de entrevista, hecho que induce a concluir la baja confiabilidad de la respuesta en una fuente de datos como la censal.

## 8.17 Situación conyugal (A: P20)

El nivel del error es superior al de los censos anteriores y dos de cada diez menores de 0 a 13 años tienen respuesta en esta pregunta.

De acuerdo a la evaluación cualitativa, esta pregunta presenta problemas conceptuales respecto del Estado civil porque combina las preguntas sobre Situación conyugal y Estado civil,<sup>55</sup> mientras que los datos cuantitativos indican que el filtro de edad no fue respetado.

### 8.17.1 Recomendaciones

- Continuar incluyendo la pregunta, debido a que es una variable básica para desarrollar los estudios demográficos, e incorporar ajustes para mejorar su captación.
- Reforzar el filtro e indagar la situación legal y la situación de hecho en forma separada.

## 8.18 Actividad económica (A: P21 a P33) (B: P12 a P16)

Todas las preguntas del bloque muestran errores superiores a los obtenidos en censos anteriores. La pregunta relativa al Domicilio del lugar de trabajo, que se incorpora en esta prueba piloto, presentó los valores más altos, seguida de la Rama de actividad, la Tarea que realiza, el Tamaño del establecimiento y la Categoría ocupacional.

En cuanto al error de secuencia, la indicación de que no respondan los menores de 14 años fue la que mayor dificultad ocasionó.

<sup>54</sup> Ver INDEC (2019b, pp. 68 y 69).

<sup>55</sup> Ver INDEC (2019b, p. 69).

## 8.18.1 Recomendaciones

- Respecto a la pregunta sobre Licencia, se propone reformular de la siguiente manera: En esa semana ¿tenía trabajo y no concurrió? (por vacaciones, enfermedad, suspensión con pago, conflicto laboral, mal tiempo, etc.), para destacar que la persona tenía trabajo, pero no concurrió, independientemente de los motivos por los que no lo hizo.
- Para la pregunta que refiere a la búsqueda laboral, se propone la siguiente reformulación: Durante las últimas 4 semanas ¿buscó trabajo de alguna manera? (contestó avisos de diario o internet, consultó amigos o parientes, puso carteles, envió currículum, hizo algo para ponerse por su cuenta, etc.), para privilegiar que la persona está buscando trabajo independientemente del modo en que lo hace.
- En la pregunta sobre Pago en dinero o en especie, se propone modificar por una forma más sencilla: ¿Es un trabajo pago? *Incluye pago en dinero o especie.*
- Para la pregunta de Categoría ocupacional, volver a la manera tradicional de preguntar: obrero o empleado, cuenta propia, patrón o empleador, trabajador familiar (rescatando de la actual operacionalización, la categoría “servicio doméstico”, como categoría diferenciada de las demás). Respecto al ordenamiento de las categorías, anteponer la categoría “servicio doméstico”, como primera categoría, para que no se genere un registro incorrecto al escuchar primero las categorías “obrera(o)” o “empleada(o)” y “cuenta propia”. Se propone la siguiente formulación: El trabajo principal ¿lo hace como... *Refiere al trabajo que le lleva más horas.* servicio doméstico?; obrera(o) o empleada(o)?; cuenta propia?; patrón(a) o empleador (a)?; trabajador(a) familiar?; Ignorado.
- En relación a las preguntas abiertas, modificar el orden y preguntar en primer lugar la Rama de actividad económica.
- Eliminar la pregunta sobre Domicilio del lugar de trabajo, por el elevado nivel de error detectado y la baja confiabilidad de la respuesta obtenida, así como también reincorporar la pregunta sobre Sector institucional del empleo.
- Reformular la pregunta sobre Tamaño del establecimiento, con una indicación para quien censa que refiera a que en la suma se debe incluir a la persona censada y modificar los intervalos por: “1 persona”; “De 2 a 5 personas”; “De 6 a 40 personas”; “De 41 a 100 personas”; “Más de 100 personas”; “Ignorado”.

## 8.19 Fecundidad (A: P34 a P36)

Todas las preguntas del bloque muestran errores superiores a los de censos anteriores, en especial las últimas dos. Estas, sin embargo, fueron respondidas por un alto porcentaje de varones debido a que la persona que realizó el censo no respetó el filtro.

El elevado error del bloque de preguntas sobre la Fecundidad y Mortalidad infantil pone de manifiesto las dificultades que tuvo en la aplicación en campo la operacionalización propuesta para la II PP.

Unificar las preguntas de Tenencia, Cantidad de HNV e HVA en una sola pregunta se tradujo en un aumento del error (no respuesta y datos inválidos) del bloque.

## 8.19.1

### Recomendaciones

- Incorporar una primera pregunta sobre si la mujer tuvo o no HNV. Esta primera pregunta permite registrar de mejor forma a aquellas mujeres que no han tenido HNV.
- Desdoblar en dos preguntas la Cantidad de HNV y la Cantidad de HVA.
- Numerar independientemente las preguntas referidas a la Tenencia de HNV, Cantidad de HNV y Cantidad de HVA.
- Modificar el texto que indica el filtro de universo, indicando la acción que debe realizar el censista en vez de quienes deben ser censados y aclarando que allí termina el cuestionario para los varones: “A partir de aquí debe hacer las preguntas a todas las mujeres de 14 años y más. Para los varones fin de la entrevista”.

## 8.20

### Recomendaciones sugeridas para el mejoramiento del proceso de codificación

- Mejorar la calidad de la lectura de los cuestionarios censales para poder realizar de forma eficiente la codificación de las preguntas abiertas.
- Incorporar una instancia de control de la lectura de letras y números, en el momento en que esta se realiza, ya que es importante para garantizar la calidad de este proceso.
- Reforzar los mecanismos de capacitación para mejorar la recolección de la información, de modo que las respuestas sean las que corresponden a las preguntas.
- Reformular la pregunta de pueblos originarios, especificando que se trata del nombre de un pueblo originario y no geográfico y revisar el código de pueblo.
- Aumentar la cantidad de caracteres para las preguntas de pueblos originarios, carrera, rama de actividad y ocupación.
- Invertir el orden de la secuencia entre las preguntas de rama de actividad y ocupación.

## 8.21

### Recomendaciones sugeridas en relación con las estrategias de capacitación y el reclutamiento de los recursos humanos

A partir de lo analizado y teniendo en cuenta el objetivo final de prueba acerca de la posibilidad de desarrollar estrategias nuevas de capacitación y reclutamiento de los recursos humanos, se han obtenido aprendizajes que permiten arribar a las siguientes recomendaciones:

- Persistir en la convocatoria de bloques alineados jerárquicamente conformados por estudiantes terciarios en el rol de censistas y sus docentes como jefes de radio.
- Sostener la necesidad de brindar una capacitación de emergencia. Contar con censistas capacitados que ingresan a la estructura de relevamiento el mismo día del censo, como recurso ante eventuales bajas que no puedan cubrirse con censistas suplentes, produciría un impacto positivo en la cobertura y en la calidad de los datos recogidos. Implicaría, asimismo, no sobrecargar la etapa de supervisión.
- Tal como se sostuvo en el capítulo 5 del *Informe de la segunda prueba piloto. Hacia la Ronda 2020* (INDEC, 2019), se reafirma la necesidad de extender la actividad de capacitación de emergencia entre 15 y 20 minutos. El objetivo es enfatizar los aspectos del procedimiento de aplicación del cuestionario censal en general y por bloques, que en la evaluación cuantitativa y cualitativa evidenciaron las mayores deficiencias.
- Por lo antes enunciado podría afirmarse que, en el marco de un censo de hecho, las estrategias probadas impactan de manera favorable sobre la cobertura censal. Se propone su implementación en el Censo Experimental con las adecuaciones mencionadas.

# Bibliografía

Chackiel, J. y Maci  , G. (1978). Conceptos b  sicos. En *Evaluaci  n y correcci  n de datos demogr  ficos*. Santiago: CELADE.

Comisi  n Econ  mica para Am  rica Latina y el Caribe (2011). *C  digo Regional de Buenas Pr  cticas en Estad  sticas para Am  rica Latina y el Caribe*. Recuperado de: [https://repositorio.cepal.org/bitstream/handle/11362/16422/FILE\\_148023\\_es.pdf?sequence=1&isAllowed=y](https://repositorio.cepal.org/bitstream/handle/11362/16422/FILE_148023_es.pdf?sequence=1&isAllowed=y).

Comisi  n Econ  mica para Am  rica Latina y el Caribe - Fondo de Poblaci  n de Naciones Unidas (2014). *Los datos demogr  ficos. Alcances, limitaciones y m  todos de evaluaci  n*. Santiago. Recuperado de: [https://repositorio.cepal.org/bitstream/handle/11362/37145/1/S1420555\\_es.pdf](https://repositorio.cepal.org/bitstream/handle/11362/37145/1/S1420555_es.pdf).

Consejo Nacional de Coordinaci  n de Pol  ticas Sociales (2019). *Metada de los indicadores de seguimiento de los ODS. Agenda 2030 Argentina*. Recuperado de: [https://www.odsargentina.gob.ar/public/documentos/seccion\\_publicaciones/metadata\\_25\\_09\\_19.pdf](https://www.odsargentina.gob.ar/public/documentos/seccion_publicaciones/metadata_25_09_19.pdf).

Instituto Nacional de Estad  stica y Censos (1985). *La pobreza en Argentina. Indicadores de necesidades b  sicas insatisfechas a partir de los datos del Censo Nacional de Poblaci  n y Vivienda 1980*. Buenos Aires: INDEC.

(2004). *El estudio de la pobreza seg  n el Censo Nacional de Poblaci  n, Hogares y Viviendas 2001. Metodolog  a de elaboraci  n del ´ndice de Privaci  n Material de los Hogares (IPMH)*. Buenos Aires: INDEC.

(2016). *Cuentas Nacionales. Metodolog  a de estimaci  n. Base 2004 y serie a precios constantes y corrientes*. Metodolog  a INDEC N   21. Buenos Aires: INDEC Recuperado de: [https://www.indec.gob.ar/ftp/cuadros/economia/metodologia\\_21\\_cuentas\\_nacionales.pdf](https://www.indec.gob.ar/ftp/cuadros/economia/metodologia_21_cuentas_nacionales.pdf).

(2019a). *Identificaci  n de la poblaci  n con discapacidad en la Argentina: aprendizajes y desaf  os hacia la Ronda Censal 2020*. Documentos de Trabajo INDEC N   24. Buenos Aires. Recuperado de: [https://www.indec.gob.ar/ftp/cuadros/publicaciones/discapacidad\\_ronda\\_censal\\_2020.pdf](https://www.indec.gob.ar/ftp/cuadros/publicaciones/discapacidad_ronda_censal_2020.pdf).

(2019b). *Informe de evaluaci  n de la segunda prueba piloto: hacia la ronda 2020*. CABA. Recuperado de: [https://www.indec.gob.ar/ftp/cuadros/publicaciones/censo\\_2020\\_informe\\_2PP.pdf](https://www.indec.gob.ar/ftp/cuadros/publicaciones/censo_2020_informe_2PP.pdf).

(s.f.). *Buenas pr  cticas estad  sticas que gu  n al INDEC*. Buenos Aires. Recuperado de: [https://www.indec.gob.ar/ftp/documentos/principios\\_indec.pdf](https://www.indec.gob.ar/ftp/documentos/principios_indec.pdf).

(s.f.). *Evaluaci  n del Dise  o Conceptual del Censo 2010. Serie: Evaluaci  n de calidad. Vivienda y caracter  sticas habitacionales del hogar*. Buenos Aires: INDEC. (mimeo).

Giusti, A. (1997). Cuantificaci  n de los errores en las variables de los formularios b  sico y ampliado. En INDEC, *Los errores en el Censo de 1991. Primeros avances sobre su magnitud en Censo Nacional de Poblaci  n y Vivienda 1991*. Serie J n   2. Evaluaci  n de la calidad de datos y avances metodol  gicos. Buenos Aires: INDEC.

Giusti, A. y Mass  , G. (1997). Aspectos conceptuales relativos a la evaluaci  n de calidad. En INDEC, *Los errores en el Censo de 1991. Primeros avances sobre su magnitud en Censo Nacional de Poblaci  n y Vivienda 1991*. Serie J n   2. Evaluaci  n de la calidad de datos y avances metodol  gicos. Buenos Aires: INDEC.

Mayntz, R., Holm, K. y Hübner, P. (1993). *Introducción a los métodos de la sociología empírica*. Madrid: Alianza Universidad.

Naciones Unidas (2017). *Principles and Recommendations for Population and Housing Censuses*. Revisión 3. Recuperado de: [https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles\\_and\\_Recommendations/Population-and-Housing-Censuses/Series\\_M67rev3-E.pdf](https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf).

Organización Internacional del Trabajo (2010). *Medición de la población económicamente activa en los censos de población: manual*. Recuperado de: [http://www.ilo.org/global/publications/books/WCMS\\_172095/lang--es/index.htm](http://www.ilo.org/global/publications/books/WCMS_172095/lang--es/index.htm).

(2013). Informe II. Estadísticas del trabajo, el empleo y la subutilización de la fuerza de trabajo. 19.a Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de: [https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms\\_220537.pdf](https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_220537.pdf).

Organización Mundial de la Salud - Organización Panamericana de la Salud - Instituto de Migraciones y Servicios Sociales (2001). *Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud (CIF)*. Madrid: OMS - OPS - IMSERSO. Recuperado de: [https://apps.who.int/iris/bitstream/handle/10665/43360/9241545445\\_spa.pdf;jsessionid=104E3E70123A0913AA1E62F5C7619E4C?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/43360/9241545445_spa.pdf;jsessionid=104E3E70123A0913AA1E62F5C7619E4C?sequence=1).

Organización para la Cooperación y el Desarrollo Económicos (2013). *OECD Guidelines on Measuring Subjective Well-being*. OECD Publishing. Recuperado de: [https://read.oecd-ilibrary.org/economics/oecd-guidelines-on-measuring-subjective-well-being\\_9789264191655-en#page1](https://read.oecd-ilibrary.org/economics/oecd-guidelines-on-measuring-subjective-well-being_9789264191655-en#page1).

Organización para la Cooperación y el Desarrollo Económicos (2015). *Recomendación del Consejo de la OCDE sobre Buenas Prácticas Estadísticas*. Recuperado de: [https://www.oecd.org/statistics/good-practice-toolkit/Brochure-Recommendation%20of%20the%20OECD%20Council%20on%20Good%20Statistical%20Practice\\_SPANISH.pdf](https://www.oecd.org/statistics/good-practice-toolkit/Brochure-Recommendation%20of%20the%20OECD%20Council%20on%20Good%20Statistical%20Practice_SPANISH.pdf).