

---

# Appendix A. Census of Agriculture Methodology

---

# Apéndice A. Metodología del Censo Agrícola

---

## INTRODUCTION

The purpose of a census is to enumerate all members of a population with a defined characteristic. For the Puerto Rico Census of Agriculture, that goal is to account for “any place from which \$500 or more of agricultural products were produced and sold, or normally would have been sold, during the census year.” To do this, NASS creates a Census Mail List (CML) of agricultural operations that potentially meet the farm definition, collects agricultural information from those operations, reviews the data, and combines the data to provide information on the farm characteristics of farm operations and farm operators at the island and regional levels. In this appendix, these census processes are described.

## THE CENSUS POPULATION

### The Census Mail List

The National Agricultural Statistics Service (NASS) maintains a list of farmers and ranchers from which the CML is compiled. The goal is to build as complete a list as possible of agricultural places that meet the farm definition. This list is compiled prior to the census, using the list of active farms from the 2018 Census of Agriculture, lists of farmers from the Puerto Rico Department of Agriculture, and the Agricultural Extension Service of the University of Puerto Rico, plus names and addresses of farm operators identified through other sources. Each record on the list includes a name, address, telephone number, and email plus additional information that is used to efficiently administer the census of agriculture. These outside source lists are matched to the NASS list using record linkage programs. Most names on newly acquired sources are already on the NASS list. Records not on the NASS list are treated as potential farms and added to the CML.

List building activities for developing the 2022 Puerto Rico CML started in 2020 by updating list information from respondents to the 2018 Puerto Rico Census of Agriculture. Measures were taken

## INTRODUCCIÓN

El propósito de un censo es enumerar todos los miembros de una población con una característica definida. Para el Censo de Agricultura de Puerto Rico, ese objetivo es contabilizar "cualquier lugar desde el cual se produjeron y vendieron \$500 o más en productos agrícolas, o que normalmente se hubiesen vendido, durante el año censal". Para hacer esto, NASS crea el listado de correo del censo (CML, por sus siglas en inglés) de operaciones agrícolas que potencialmente cumplen con la definición de finca, recopila información agrícola sobre esas operaciones, revisa los datos y combina los datos para proporcionar información sobre las características de las operaciones agrícolas y los agricultores a nivel insular y regional. En este apéndice, se describen estos procesos censales.

## LA POBLACIÓN DEL CENSO

### La lista de Correo del Censo

El Servicio Nacional de Estadísticas Agrícolas (NASS) mantiene una lista de agricultores a partir de la cual se compila el listado de correo del censo (CML). El objetivo es construir un listado lo más completo posible de lugares agrícolas que cumplan con la definición de finca. Este listado se recopila antes del censo, utilizando el listado de fincas activas del Censo de Agricultura de 2018, listados del Departamento de Agricultura de Puerto Rico y del Servicio de Extensión Agrícola de la Universidad de Puerto Rico, además de los nombres y direcciones de operadores agrícolas identificados a través de otras fuentes. Cada registro en la lista incluye un nombre, dirección, número de teléfono y correo electrónico, además de información adicional que se utiliza para administrar eficientemente el censo de la agricultura. Las listas de fuentes externas se verifican con la lista de NASS utilizando programas de vinculación de registros. La mayoría de los nombres en fuentes recién adquiridas ya están en la lista de NASS. Los registros que no figuran en la lista NASS se tratan como posibles fincas y se agregan al CML.

Las actividades para desarrollar la CML para el Censo de Agricultura de Puerto Rico de 2022 comenzaron en el 2020 actualizando la información de la lista de los agricultores que respondieron al Censo de Agricultura de Puerto Rico de 2018. Se tomaron medidas para mejorar la calidad de los

to improve name and address quality. Additional record linkage programs were run to detect and remove duplicate records. The official CML for the 2022 Puerto Rico Census of Agriculture was established on September 2022. The list contained 17,875 records.

### **Not on the Mail List (NML)**

Extensive efforts are directed toward developing a CML that includes all farms in Puerto Rico. However, some farms are not on the list, and some agricultural operations on the list are not farms. NASS uses its Area Coverage Evaluation Survey (ACES) to quantify the number and types of farms not on the CML. The records in the ACES that are not on the CML are said to be in the Not-on-the-Mail List (NML) domain. If an ACES record in the NML domain is determined to be a farm during the census, it is an NML farm. The NML farms are used to measure coverage associated with the census.

The ACES is based on an area frame, which covers all land in Puerto Rico and includes all farms. The land in Puerto Rico is stratified by characteristics of the land. A probability sample of segments is drawn within each stratum for the ACES. Segments of approximately equal size are delineated within each stratum and designated on aerial photographs. The ACES sample of segments is allocated to strata to provide accurate measures of cuerdas (1 cuerda = 0.97 acres) planted to widely grown crops, farm numbers, and inventories of livestock. The ACES sample consisted of 300 segments, which are personally enumerated. Each operation identified within a segment boundary is known as a tract.

During the ACES enumeration process, each tract is identified as either agricultural or nonagricultural. Each ACES agricultural tract is identified as a farm or non-farm in 2022 based on the farm definition of \$500 of sales or potential sales of agricultural products. Non-agricultural tracts are further classified into categories: with farm potential, with unknown farm potential, or with no farm potential. The names and addresses collected in ACES were matched to the CML. Those names from the ACES that did not match were determined to be in the NML domain and sent a census report form. Instructions on the census report form directed any respondent who received duplicate forms to complete only one form and to mail all duplicate forms back together. Those who returned a CML and an NML form had been

nombres y direcciones. Se ejecutaron programas de vinculación de registros adicionales para detectar y eliminar duplicados. La CML oficial para el Censo Agrícola de Puerto Rico 2022 se estableció en septiembre de 2022. La lista contenía 17,875 registros.

### **No en la Lista de Correo (NML)**

Se realizan grandes esfuerzos para desarrollar una CML que incluya todas las fincas en Puerto Rico. Sin embargo, algunas fincas no están en la lista, y algunas operaciones agrícolas en la lista no son fincas. NASS utiliza la Encuesta de Evaluación de Cobertura de Área (ACES, por sus siglas en inglés) para cuantificar el número y los tipos de fincas que no están en la CML. Las fincas en el ACES que no están en la CML se asignaron a la lista de agricultores que no se encuentran en la lista de correo del censo (NML, por sus siglas en inglés). Si se determina durante el censo que un registro ACES en la NML es una finca, se identifica como una finca NML. Las fincas NML se utilizan para medir la cobertura asociada con el censo.

El ACES se basa en un marco de área, que cubre toda la superficie terrestre de Puerto Rico e incluye todas las fincas. La superficie terrestre de Puerto Rico se estratifica por las características del terreno. Para ACES, se escoge una muestra probabilística de segmentos dentro de cada estrato. Se delimitan segmentos de aproximadamente el mismo tamaño dentro de cada estrato y se designan en fotografías aéreas. La muestra de segmentos ACES se asigna a estratos para proporcionar medidas precisas de cuerdas (1 cuerda = 0.97 acres). La muestra de ACES consistió en 300 segmentos que se enumeraron personalmente. Cada operación identificada dentro de los límites de un segmento se conoce como tracto.

Durante el proceso de enumeración de ACES, cada tracto se identifica como agrícola o no agrícola. Cada tracto agrícola de ACES se identifica como una finca o no finca en 2022 basado en la definición de la finca de \$500 o más en venta de productos agrícolas o que potencialmente hubieran podido haberse vendido. Los tractos no agrícolas se clasifican en las siguientes categorías: con potencial agrícola, con potencial agrícola desconocido, o sin potencial agrícola. Los nombres y direcciones recopilados en ACES se comparan con los nombres en la CML. Los nombres de ACES que no se encontraron en la CML se consideraron como NML y se les envió un formulario censal. Las instrucciones en el formulario censal indicaban a cualquier encuestado que recibiera más de un formulario censal a completar solamente un formulario y a devolver todos los formularios duplicados junto con el formulario completado. Aquellos que devolvieron un formulario CML y NML habían sido clasificados erróneamente como NML y fueron eliminados del dominio NML. El envío inicial de

misclassified as NML and were removed from the NML domain. The initial NML mailout consisted of 1,383 records. A total of 1,367 NML records were analyzed of which 87 records were confirmed to be NML and in-scope.

The farm/nonfarm status of each NML domain operation was determined based on the reported data in the census form. An operation in the NML domain that was determined to be a farm is referred to as an NML farm. Characteristics of NML farms and their operators provided a measure of the undercoverage of farms on the CML. In general, NML farms tended to be small in acreage, production, and sales of agricultural products. Farm operations were missing from the CML for various reasons, including the possibility that the operation started after development of the CML, the operation was so small that it did not appear in any agriculture-related source list, or the operation was misclassified as a nonfarm prior to census mailout. The CML was used with the NML in a capture-recapture framework to represent all farming operations across Puerto Rico.

## **DATA COLLECTION OUTREACH AND PROMOTIONAL EFFORTS**

NASS developed a communication plan largely based on promotional materials that were shared with local outreach partners, including but not limited to the University of Puerto Rico and the Puerto Rico Department of Agriculture, for the island's operators, enumerators, and media. The goal with these products included:

- Encouraging participation in the Census of Agriculture
- Directing operators to the NASS website for online response
- Communicating how the census provides much needed data that are used by federal and local decision makers
- Explaining that response to the Census of Agriculture is required and that reported information is protected by federal law
- Increasing general awareness and perceived value of NASS, its products, and services

### **Partnership and Local-Level Outreach**

In addition to the contribution made by University of Puerto Rico Extension Service, which produced the Census brochure and poster, several USDA agencies in Puerto Rico, including the Natural

NML consistió en 1,383 registros. Se analizó un total de 1,367 registros NML, de los cuales se confirmó que 87 registros eran NML y eran fincas según la definición del censo.

La clasificación como finca o no finca de cada operación en el listado NML se determinó basado en los datos reportados en el formulario censal. Una operación en el listado NML que se determinó que es una finca se denomina como una finca NML. Las características de las fincas NML y sus operadores proporcionaron una medida de la cobertura insuficiente de las fincas en la CML. En general, las fincas NML tienden a ser pequeñas en tamaño, producción, y ventas de productos agrícolas. Algunas razones por la cual operaciones agrícolas no fueron incluidas en la CML incluyen: la posibilidad de que la operación comenzara después del desarrollo de la CML, la operación era tan pequeña que no aparecía en ninguna lista de fuentes relacionadas con agricultura, o la operación se clasificó erróneamente como no agrícola antes del envío del censo. La CML se utilizó con el NML en un marco de captura-recaptura para representar todas las operaciones agrícolas en Puerto Rico.

## **DIFUSIÓN DE RECOPIACIÓN DE DATOS Y ESFUERZOS PROMOCIONALES**

El NASS desarrolló un plan de comunicación basado principalmente en materiales de promoción que se compartieron con nuestros colaboradores locales, entre ellos, la Universidad de Puerto Rico y el Departamento de Agricultura de Puerto Rico, para ser distribuidos entre los agricultores, los enumeradores, y los medios de comunicación local. El objetivo con estos productos incluyó:

- Fomentar la participación en el Censo de Agricultura
- Dirigir a los agricultores a la página web de NASS para que respondieran en línea
- Comunicar cómo el censo provee datos muy necesarios que serán utilizados por el gobierno Federal y local para tomar decisiones
- Explicar que la respuesta al Censo de Agricultura es obligatoria y que la información reportada está protegida por la ley federal
- Aumentar la conciencia general y el valor percibido de NASS, sus productos, y servicios

### **Colaboración y Difusión a Nivel Local**

Además de la contribución realizada por el Servicio de Extensión Agrícola de la Universidad de Puerto Rico, que produjo el folleto y el afiche del Censo, varias agencias del USDA en Puerto Rico, incluyendo el Servicio de

Resources and Conservation Service (NRCS) and the Farm Service Agency (FSA), helped spread the word about data collection. Before data collection began, the local government held a proclamation ceremony for the media, and other stakeholders. Several dozen Census folders were stuffed with associated content and provided to partners. Talking points in both English and Spanish were printed and laminated for local leadership, partners, and enumerators.

## **Public Relations**

In the public relations arena, NASS worked with internal and external stakeholders to equip them with communications tools and resources to deliver the census communications message to their audiences.

Two news releases/stakeholder notices, two public service announcements, and a Frequently Asked Questions (FAQs) document were drafted in English and Spanish for local media and other stakeholder distribution. These materials were available both electronically and in hard copy. Other outreach tools also included Census swag (i.e. pens, notepads) as well as copies of the 2018 Puerto Rico Census of Agriculture Highlights and instructions for responding online in English and Spanish.

## **DATA COLLECTION**

### **Method of Enumeration**

Data collection was accomplished primarily by mail. A Computer-Assisted Self-Interviewing (CASI) instrument was also available for operators who preferred to report online. A letter with a unique survey code and instructions for completing their census online was included in each mail package. Enumerators from the Puerto Rico Department of Agriculture and the Extension Service conducted field follow-up visits to enumerate operations that did not respond by mail.

### **Report Form**

A single version of the report form, in Spanish, was prepared by NASS, in cooperation with an inter-agency working group that included members of the Puerto Rico Department of Agriculture, the College of Agricultural Sciences at the University of Puerto Rico, Mayagüez Campus (RUM), the Extension Service, and other data users. The report forms targeted crops and livestock specifically grown or raised in Puerto Rico but it also allowed

Conservación y Recursos Naturales (NRCS, por sus siglas en inglés) y la Agencia de Servicio Agrícola (FSA, por sus siglas en inglés), ayudaron a difundir información sobre la recolección de datos. Antes del comienzo de la recopilación de datos, el gobierno local celebró una ceremonia de proclamación para los medios de comunicación y otros colaboradores. Se prepararon docenas de carpetas del Censo con materiales de promoción y se les proporcionaron a nuestros colaboradores locales. Los puntos de información importantes fueron impresos y laminados para distribución entre los supervisores y los enumeradores.

## **Relaciones Públicas**

En el ámbito de las relaciones públicas, NASS trabajó con partes interesadas internas y externas para equiparlos con herramientas y recursos de comunicación para difundir el mensaje de comunicación del censo a sus audiencias.

Se redactaron dos comunicados de prensa / avisos de partes interesadas, dos anuncios de servicio público y un documento de las preguntas más frecuentes (FAQ, por sus siglas en inglés) en inglés y español para los medios locales y la distribución de otras partes interesadas. Estos materiales estaban disponibles tanto en formato electrónico como impresos. Otros materiales de promoción incluyeron bolígrafos, libretas, etc., así como folletos con la información más importante del Censo de Agricultura de Puerto Rico de 2018 e instrucciones tanto en inglés como en español para responder en línea.

## **RECOPILACIÓN DE DATOS**

### **Método de Enumeración**

La recopilación de datos se realizó principalmente por correo. Un formulario de auto-entrevista asistida por computadora (CASI, por sus siglas en inglés) también estuvo disponible para los agricultores que prefirieron reportar en línea. En cada paquete de correo se incluyó una carta con el código de encuesta e instrucciones para completar el censo en línea. Enumeradores del Departamento de Agricultura de Puerto Rico y del Servicio de Extensión Agrícola realizaron visitas de seguimiento para encuestar los agricultores que no respondieron por correo.

### **Formulario Censal**

NASS preparó una sola versión del formulario censal, en español, en cooperación con el comité interagencial integrado por miembros del Departamento de Agricultura de Puerto Rico, el Colegio de Ciencias Agrícolas de la Universidad de Puerto Rico, Recinto Universitario de Mayagüez (RUM), el Servicio de Extensión Agrícola y otras entidades agrícolas. El formulario censal incluyó cultivos y animales característicos de Puerto Rico.

respondents to write in specific commodities that were not prelisted on the report form.

### **Data Collection Training Program**

Selected staff members from the Puerto Rico Department of Agriculture and the Extension Service received special training for the census in accordance with instructions prepared by NASS. The training included an overview of the census of agriculture program, and a detailed discussion of the enumerator's instructions manual and the report form.

### **Report Form Mailings**

Approximately 19,500 mail packets were mailed on December 2022. Each packet contained a cover letter, instruction sheet, a labeled report form, and a return envelope. The Census Bureau's National Processing Center (NPC) in Jeffersonville, IN was contracted to perform mail packet preparation, initial mailout, and one follow-up mailing to nonrespondents.

### **Nonresponse Follow-up**

NASS targeted selected groups for in-person enumeration. These groups included:

- Puerto Rico Department of Agriculture list - large records (\$50,000 or more in sales)
- Puerto Rico Department of Agriculture list - other records (with less than \$50,000 in sales)
- Extension Service list

These records were identified as large or unique operations, the absence of which could have significantly affected the accuracy of census results. Enumerators from the Puerto Rico Department of Agriculture and the Extension Service conducted field follow-up visits to enumerate operations that did not respond by mail. If an operation was no longer in business, its nonfarm status was documented.

## **REPORT FORM PROCESSING**

### **Data Capture**

The Census Bureau's National Processing Center (NPC) in Jeffersonville, IN was contracted to process returned mail packets. NASS staff on site at the NPC provided technical guidance and monitored NPC processing activities. Questionnaires returned to the NPC were

Además, ofreció a los encuestados poder incluir productos que no figuraban en las listas impresas en el formulario censal.

### **Programa de Capacitación en Recopilación de Datos**

Personal del Departamento de Agricultura de Puerto Rico y del Servicio de Extensión Agrícola recibieron capacitación especial para el censo de acuerdo con instrucciones preparadas por NASS. La capacitación incluyó una descripción general del programa del censo de agricultura y una discusión detallada del manual de instrucciones del enumerador y del formulario censal.

### **Envío de Formularios**

En diciembre de 2022 se enviaron aproximadamente 19,500 paquetes por correo. Cada paquete contenía una carta de presentación, una hoja de instrucciones, un formulario censal etiquetado y un sobre de devolución. El Centro Nacional de Procesamiento (NPC, por sus siglas en inglés) del Negociado del Censo en Jeffersonville, Indiana, fue contratado para preparar los paquetes de correo, el envío inicial, y un segundo envío a quienes no respondieron.

### **Seguimiento a la No Respuesta**

NASS seleccionó grupos determinados que no habían devuelto el formulario censal para ser visitados personalmente por un enumerador. Estos grupos incluyeron:

- Lista del Departamento de Agricultura de Puerto Rico de fincas con \$50,000 o más en ventas
- Lista del Departamento de Agricultura de Puerto Rico de fincas con menos de \$50,000 en ventas
- Lista de fincas del Servicio de Extensión Agrícola

Estos registros se identificaron como fincas grandes o únicas, cuya ausencia podría haber afectado significativamente la precisión de los resultados del censo. Enumeradores del Departamento de Agricultura de Puerto Rico y el Servicio de Extensión visitaron personalmente las fincas que no habían respondido por correo. Si una finca ya no estaba en operación, se documentaba su estado como no agrícola.

## **PROCESAMIENTO DEL FORMULARIO CENSAL**

### **Captura de los Datos**

El Centro Nacional de Procesamiento (NPC) del Negociado del Censo en Jeffersonville, IN, fue contratado para procesar los paquetes de correo devueltos. El personal de

immediately checked in, utilizing the bar codes printed on the mailing label. This ensured the case would be removed from follow-up report form mailings. All forms with data were scanned and an image was made of each page of the questionnaire. Optical Mark Recognition (OMR) was used to capture categorical responses and to identify entries in numeric and alpha-character answer zones.

Data entry operators keyed data from the scanned images. Answer zones with entries, identified in the earlier OMR analysis were presented to the data entry operators. The keyer evaluated the contents and captured pertinent responses. Ten percent of the identified answer zones were keyed a second time for independent quality control. If differences existed between the first keyed value and the second, an adjudicator handled resolution. The adjudicator identified the correct entry and identified the “cause” of the error in the other entry.

The adjudication provided feedback to the keyers to improve data capture skills, reward skilled keyers, and ensure that the error rate did not exceed the Acceptable Quality Limit (AQL) of 1%. The data capture error rate for the Puerto Rico Census of Agriculture was measured at 0.69%. The images and captured data were transmitted to NASS’s centralized network and became available to NASS analysts on a flow basis. The images were then available for use in all stages of review.

## **EDITING DATA**

Captured data were processed through a computer formatting program that verified that records were valid – that the record ID number was on the list of census records, that the reported municipios of operation and production were valid, and other related criteria. Rejected records were referred to analysts for correction. Accepted records were sent to a complex computer batch edit process. Each execution of the computer edit in batch mode flowed as the data were received from NPC.

The computer edit determined whether a reporting operation met the qualifying criteria to be counted as a farm (in-scope). The edit examined each in-scope record for reasonableness and completeness and determined whether to accept the recorded value for each data item or take corrective action. Such corrective actions included removing erroneously reported values, replacing an unreasonable value with one consistent with other reported data, or providing a value for an item

NASS adscrito al NPC proporcionó orientación técnica y supervisó las actividades de procesamiento que se llevaron a cabo. Los formularios devueltos al NPC se procesaban inmediatamente, utilizando los códigos de barras impresos en la etiqueta de correo. Esto aseguró que el caso se eliminara de los listados de correos designados a recibir un formulario de seguimiento. Se escanearon todos los formularios con datos y se realizó una imagen de cada página del formulario. El reconocimiento óptico de marcas (OMR, por sus siglas en inglés) se utilizó para capturar respuestas categóricas e identificar entradas en zonas de respuesta numéricas y de caracteres alfabéticos.

Los operadores de entrada de datos ingresaron los datos utilizando las imágenes escaneadas. Las zonas de respuesta con entradas, identificadas en el análisis OMR, se asignaron a los operadores de entrada de datos. El operador de entrada de datos evaluó el contenido y capturó las respuestas pertinentes. El diez por ciento de las zonas de respuesta identificadas fueron codificadas por segunda vez por un control de calidad independiente. Si existían diferencias entre el primer valor tecleado y el segundo, un supervisor se encargaría de la resolución. El supervisor identificó la respuesta correcta e identificó la "causa" del error en la respuesta incorrectas.

El proceso de adjudicación proporcionó información a los operadores de entrada de datos para mejorar sus habilidades de captura de datos, premiar a los más proficientes y garantizar que el Límite de Calidad Aceptable (AQL, por sus siglas en inglés) no superase el 1%. La tasa de error de captura de datos para el Censo Agrícola de Puerto Rico se midió en 0.69%. Las imágenes y los datos capturados se transmitieron a la red centralizada de NASS y se pusieron a disposición de los analistas de NASS según se iban procesando. Las imágenes estuvieron disponibles durante todas las etapas de revisión.

## **EDICIÓN DE DATOS**

Los datos recolectados se procesaron a través de un programa de formato de computadora que verificó que las respuestas fueran válidas: que el número de identificación asignado a cada cuestionario estaba en la lista de registros del censo, que los municipios de operación y producción informados eran válidos y otros criterios relacionados. Los formularios rechazados fueron remitidos a los analistas para su corrección. Los formularios aceptados fueron agrupados y enviados a un proceso de revisión informática. Cada ejecución de esta revisión informática se realizó según se fueron recibiendo los datos del Centro Nacional de Procesamiento (NPC).

El programa de revisión determinó si el formulario cumplía con los criterios establecidos para calificar como una finca

omitted by the respondent. To the extent possible, the computer edit determined a replacement value. Strategies for determining replacement values are discussed in the next section. Operations failing to meet the qualifying criteria for being classified as a farm were categorized as out-of-scope for the census. Records that NASS had reason to believe might have been erroneously classified as out-of-scope (indications of recent and/or significant agricultural activity reported on NASS surveys, for example) were referred to analysts for verification.

The edit systematically checked reported data section-by-section with the overall objective of achieving an internally consistent and complete report. NASS subject-matter experts had previously defined the criteria for acceptable data. Problems that could not be resolved within the edit were referred to an analyst for intervention. A group of analysts examined the scanned images, consulted additional sources of information, and determined an appropriate action.

### **Imputing Data**

The edit determined the best value to impute for reported responses that were deemed unreasonable and for required responses that were absent. For operators who had not changed in five years, demographics such as race and gender were taken from the previous census. Administrative data from the Farm Service Agency (FSA) were used for a few items, such as Conservation Reserve Program (CRP) acreage. When deterministic edit logic and previously reported data sources were unable to provide a current value, data from a reporting farm of similar type, size, and location were considered. In cases where automated imputation was unable to provide a consistent report, the record was referred to an analyst for resolution.

Separate system processes were established to efficiently provide data from a similar farm to the edit when donor imputation was required. The farm characteristics used to define similarity between a recipient record and its donor record were in part defined beforehand, and in part determined by the edit logic. Euclidean distance was used for similarity computations, with each contributing similarity characteristic scaled appropriately. The most similar farm based on this criterion (the “nearest neighbor”) was identified and returned to the edit for use as a donor. The calculated distance between the centroids of the principal municipios of production of the donor and recipient was always included as one of the measures of

para propósitos del censo. Para cada formulario el programa verificó que los datos reportados fuesen razonables y estuviesen completos. Además, determinó si aceptaba los valores reportados o si tomaba medidas correctivas. Las acciones correctivas incluyeron la eliminación de valores reportados erróneamente, la sustitución de un valor no razonable por uno consistente con los otros datos reportados, o el proveer valores adecuados para partidas dejadas en blanco. Según fuese necesario, el programa determinó valores sustitutos. Las estrategias para determinar los valores sustitutos se discuten en la siguiente sección. Las operaciones que no cumplían con los criterios de calificación para ser clasificadas como fincas se clasificaron como no fincas. Los cuestionarios que NASS creía que habían sido clasificados erróneamente como no fincas (por ejemplo, indicaciones de actividad agrícola reciente en otras encuestas de NASS) se remitieron a los analistas para su verificación.

El programa de revisión verificó sistemáticamente todos los datos reportados en el formulario, sección por sección, con el objetivo general de lograr un informe de datos que fueran internamente consistentes y completos. Los analistas de NASS definieron los criterios para datos aceptables. Los problemas que el programa no pudo resolver fueron remitidos a un analista para su intervención. Los analistas utilizaron fuentes adicionales de información, examinaron las imágenes escaneadas, y determinaron la acción apropiada.

### **Imputación de Datos**

El programa de revisión de datos determinó el mejor valor para imputar las respuestas informadas que se consideraron no eran razonables y las respuestas requeridas que estuvieron ausentes. Para operadores que no habían cambiado en cinco años, se tomaron los datos demográficos tales como raza y género del censo anterior. Los datos administrativos de la Agencia de Servicios Agrícolas (FSA, por sus siglas en inglés) se utilizaron para algunas partidas, tales como las cuerdas en el Programa de Reserva de Conservación (CRP, por sus siglas en inglés). Cuando la lógica de edición determinista y las fuentes de datos informados anteriormente no pudieron proporcionar un valor actual, se consideraron los datos de una finca similar en cuanto al tipo, el tamaño y la ubicación. En los casos en que la imputación automática no pudo proporcionar un informe coherente, el formulario se envió a un analista para su resolución.

Se establecieron procesos de sistema separados para proporcionar eficientemente datos de una finca similar al programa de edición cuando se requirió un donante para imputar un valor. Las características de finca utilizadas para definir la similitud entre un registro de destinatario y su

similarity.

To provide donors to the automated edit, a pool of successfully edited records was maintained for each section of the report form. These donor pools began with 2018 census data, reconfigured to emulate 2022 data and then edited using 2022 logic. As 2022 records were successfully processed, they were added to the donor pools, which maintained the most recent data for each farm. Donor pools were updated periodically, as determined by edit processing schedules. After each update, donor pool records were grouped into strata containing farms of similar type and size, using a data-driven algorithm to define strata.

In response to each donor request issued by the edit, a dedicated system process would search the appropriate stratum and respond with the most similar donor, while giving preference to more recent donors. In relatively rare instances where it was unable to provide a donor, the donor selection process issued an appropriate failure message to the edit. Imputation failures occurred for several different reasons. The requirement that an imputed value be positive could have ruled out all available donors, as could have the necessity for the donor record to satisfy a particular constraint – say, that the donor record has cattle, but no milk cows. In general, an imputation failure occurred if there were no satisfactory donors in the same profile as the report being edited. Records with imputation failures were either held until more records were available in the donor pool or referred to an analyst. In addition, when such a failure occurred in finding a donor for expenditure data, donor pool averages were provided in lieu of an individual donor, wherever possible. This “failover” utility was first introduced for the 2012 census imputation process, and significantly reduced the number of imputation failures among the expenditure and labor variables.

After receiving a donor's data, the edit substituted the values into the edited record. In many cases, the donor record's data value was scaled using another data field specified in the edit logic. In such cases, the size of the auxiliary field's value in the edited record, relative to its value in the donor record, was used to appropriately scale the donor record's value for the field to be imputed. The imputed data were then validated by the same edit logic to which reported data were subject. Since imputation was conducted independently for each occurrence, reports requiring multiple imputations may have drawn from multiple donors.

registro de donante se definieron en parte de antemano y en parte se determinaron mediante la lógica de edición. La distancia euclidiana se usó para cálculos de similitud, con cada característica de similitud contribuyente escalada adecuadamente. La finca más similar basada en este criterio (el "vecino más cercano") fue identificada y devuelta al programa de edición para su uso como donante. La distancia calculada entre los centroides de los principales municipios de producción del donante y el receptor siempre se incluyó como una de las medidas de similitud.

Para proporcionar donantes a la edición automatizada, se mantuvo un grupo de reportes de fincas editados con éxito para cada sección del cuestionario. Estos grupos de donantes comenzaron con los datos del censo de 2018, reconfigurados para emular los datos de 2022 y luego editados utilizando la lógica de 2022. A medida que los cuestionarios de 2022 se procesaron con éxito, se agregaron a los grupos de donantes, que mantuvieron los datos más recientes para cada finca. Los grupos de donantes se actualizaron periódicamente, según lo determinado por los horarios de procesamiento de edición. Después de cada actualización, los registros de la agrupación de donantes se agruparon en estratos que contenían fincas de tipo y tamaño similares, utilizando un algoritmo basado en datos para definir estratos.

En respuesta a cada solicitud de donante emitida por el programa de edición, un proceso de sistema dedicado buscaría el estrato apropiado y respondería con el donante más similar, al tiempo que daría preferencia a los donantes más recientes. En casos relativamente raros en los que no pudo proporcionar un donante, el proceso de selección de donantes emitió un mensaje de error apropiado para la edición. Las fallas de imputación ocurrieron por varias razones diferentes. El requisito de que un valor imputado sea positivo podría haber descartado a todos los donantes disponibles, ya que podría tener la necesidad de que el registro de donantes satisfaga una restricción particular; por ejemplo, que el reporte donante tenga ganado, pero no vacas lecheras. En este caso, se produjo un error de imputación si no había donantes satisfactorios en el mismo perfil que el cuestionario que se está editando. Los reportes con fallas de imputación se mantuvieron hasta que hubo más cuestionarios disponibles en el grupo de donantes o fueron remitidos a un analista. Además, cuando se produjo tal falla en la búsqueda de un donante para los datos de gastos, se proporcionaron promedios de grupo de donantes en lugar de un donante individual, siempre que fuera posible. Esta utilidad de "conmutación por error" se introdujo por primera vez para el proceso de imputación del censo de 2012, y redujo significativamente el número de fallas de imputación entre las variables de gastos y mano de obra.



## Data Analysis

The complex edit ensured the full internal consistency of the record. All substantial changes to the data generated by the computer edits were reviewed and verified by analysts whom were also provided an additional set of tools, in the form of listings and graphs, to review record-level data across farms. These examinations revealed extreme outliers, large and small, or unique data distribution patterns that were possibly a result of reporting, recording, or handling errors. Potential problems were investigated and, when necessary, corrections were made, and the record interactively edited again.

Prior to publication, tabulated totals were reviewed by statisticians to identify inconsistencies and potential coverage problems. Comparisons were made with previous census data, as well as other available data. Tallies of all selected data items for various sets of criteria which included, but were not limited to, geographic levels, farm types, and sales levels were reviewed. When necessary, data inconsistencies were resolved.

## ACCOUNTING FOR UNDERCOVERAGE, NONRESPONSE, AND MISCLASSIFICATION

Although much effort has been expended making the CML as complete as possible, it does not include all farms in Puerto Rico, resulting in list undercoverage. Additionally, some farm operators who were on the CML did not respond to the census, despite numerous attempts to contact them. Finally, although each operation was classified as a farm or a nonfarm based on the responses to the census report form, some were misclassified; that is, some nonfarms were classified as farms and some farms were classified as nonfarms. NASS's goal was to produce agricultural census totals for publication level that were fully adjusted for list undercoverage, nonresponse, and misclassification.

In 2018, NASS used a series of models based on a subset of the responding census and all ACES records in a capture-recapture framework to separately adjust for undercoverage, nonresponse, and misclassification. For the 2022 Puerto Rico Census of Agriculture, the capture-recapture methodology was extended to model the probability of capture with a single model, thereby allowing the utilization of all census responses and

Después de recibir los datos de un reporte donante, la edición sustituyó los valores en el formulario editado. En muchos casos, el valor de datos del reporte donante se escaló utilizando otra partida de datos especificado en la lógica del programa de edición. En tales casos, el tamaño del valor de la partida auxiliar en el formulario editado, en relación con su valor en el cuestionario donante, se utilizó para escalar adecuadamente el valor del formulario donante para la partida que se imputará. Los datos imputados fueron validados por la misma lógica de edición a la que estaban sujetos los datos informados. Dado que la imputación se realizó de forma independiente para cada caso, los formularios que requirieron múltiples imputaciones pueden haberse extraído de múltiples donantes.

## Análisis de los datos

La edición compleja garantizó la plena consistencia interna del formulario. Todos los cambios sustanciales en los datos generados por las ediciones por computadora fueron revisados y verificados por analistas a quienes también se les proporcionó un conjunto adicional de herramientas, en forma de listados y gráficas, para revisar datos a nivel de reporte individual de fincas. Esta revisión identificó valores extremos, grandes o pequeños, y patrones únicos de distribución de datos que posiblemente fueron el resultado de errores al reportar, codificar o procesar los formularios. Se investigaron los posibles problemas y, cuando fue necesario, se hicieron correcciones y el formulario se editó de forma interactiva nuevamente.

Antes de la publicación, los estadísticos revisaron los totales tabulados para identificar inconsistencias y posibles problemas de cobertura. Se hicieron comparaciones con datos de censos anteriores, así como con otros datos disponibles. Se revisaron los totales de datos seleccionados utilizando criterios que incluyeron, entre otros, niveles geográficos, tipos de finca y niveles de ventas. Cuando fue necesario, se resolvieron las inconsistencias de los datos.

## CONTABILIDAD POR SUBCUBRIMIENTO, NO RESPUESTA, Y CLASIFICACIÓN ERRONEA

Aunque se han realizado muchos esfuerzos para que la CML sea lo más completa posible, la misma no incluye todas las fincas en Puerto Rico, lo que resulta en una cobertura insuficiente de la lista. Además, algunos agricultores que estaban en la CML no respondieron al censo, a pesar de los numerosos intentos para tratar de contactarlos. Finalmente, aunque cada operación se clasificó como finca o no finca dependiendo de las respuestas al formulario censal, algunas fueron clasificadas erróneamente; es decir, algunas operaciones no agrícolas se clasificaron como fincas y algunas fincas se clasificaron

ACES records in the adjustments. To implement capture-recapture methods, two independent surveys were required. The 2022 Puerto Rico Census of Agriculture (based on the CML) and the 2022 Agricultural Coverage Evaluation Survey (ACES) (based on the area frame) were those two surveys. Historically, NASS has been careful to maintain the independence of the CML and the area frame. Thus, the Puerto Rico Census of Agriculture and ACES were assumed to be independent after accounting for heterogeneity in the capture probabilities based on characteristics of records.

For a farm to be identified as a farm, and thus captured by the census it must be on the CML, respond to the census report form and, based on the census response, be classified as a farm.

$$\pi_C = \pi(\text{CML, Responded, Farm on Census} | \text{Farm})$$

Two types of classification error can occur. First, a farm can be misclassified as a nonfarm. This type of misclassification is accounted for in determining the probability of capture  $\pi_C$ . The second type of classification error results when a response to the census is classified as a farm operation when it does not meet the definition of a farm. That is, some farms on the CML may be misclassified from their census report response and may be nonfarms. To account for the misclassification of nonfarms as farms, the probability of a farm on the census being classified correctly must be estimated; that is,

$$\pi_{CCFC} = \pi(\text{Farm} | \text{Farm on Census})$$

where *CCFC* represents Correct Census Farm Classification.

To adjust for undercoverage, nonresponse, and misclassification, each CML record classified as a farm based on its response to the census report form was given a weight of the ratio of the estimated probability of correct classification of a farm on the census and the estimated probability of capture ( $\hat{\pi}_C$  and  $\hat{\pi}_{CCFC}$ ) where the hat symbol (^) denotes an estimate. To estimate the number of farms with a given set of characteristics, the weights of CML records responding as farms on the census and having that set of characteristics were summed.

This estimator is referred to as the capture-recapture estimator (*CR*):

como operaciones no agrícolas. El objetivo de NASS era publicar totales para el censo agrícola ajustados por errores de cobertura en la lista, la no respuesta, y la clasificación errónea.

En 2018, NASS utilizó una serie de modelos basados en un subconjunto del censo que respondió y todos los registros de ACES en un marco de captura-recaptura para ajustar por separado la cobertura insuficiente, la falta de respuesta, y la clasificación errónea. Para el Censo Agrícola de Puerto Rico de 2022, se amplió la metodología de captura-recaptura para modelar la probabilidad de captura con un solo modelo, permitiendo así la utilización de todas las respuestas del censo y registros de ACES en los ajustes. Para implementar los métodos de captura-recaptura, se requirieron dos encuestas independientes. El Censo de Agricultura de Puerto Rico de 2022 (basado en la CML) y la Encuesta de Evaluación de Cobertura Agrícola de 2022 (ACES) (basado en un marco de área) fueron esas dos encuestas. Históricamente, NASS ha tenido cuidado de mantener la independencia de la CML y el marco de área. Por lo tanto, se asumió que el Censo Agrícola de Puerto Rico y la ACES eran independientes después de tomar en cuenta la heterogeneidad en las probabilidades de captura basadas en las características de los registros.

Para que una finca se identifique como finca y, por lo tanto, sea capturada por el censo, debe estar en la CML, responder al formulario censal, y según la respuesta al censo, clasificarse como una finca.

$$\pi_C = \pi(\text{CML, Respondió, Finca en el censo} | \text{Finca})$$

Se pueden producir dos tipos de errores de clasificación. Primero, una finca puede clasificarse erróneamente como no agrícola. Este tipo de clasificación errónea se tiene en cuenta para determinar la probabilidad de captura  $\pi_C$ . El segundo tipo de error de clasificación se produce cuando una respuesta al censo es clasificada como una finca cuando no cumple con la definición de finca. Es decir, algunas fincas en la CML pueden clasificarse erróneamente a partir de su respuesta al formulario censal, pero luego se determina que no son operaciones agrícolas. Para determinar la clasificación errónea de las no fincas como fincas, se debe estimar la probabilidad de que una finca en el censo se clasifique correctamente; esto es,

$$\pi_{CCFC} = \pi(\text{Finca} | \text{Finca en el censo})$$

donde *CCFC* representa la clasificación correcta de finca en el censo.

Para ajustar por la cobertura insuficiente, la falta de respuesta, y la clasificación errónea, cada registro en la CML clasificado como una finca basado en su respuesta al

$$CR = \sum_{i \in F} \frac{\hat{\pi}_{CCFC,i}}{\hat{\pi}_{C,i}}$$

where  $F$  is the set of all CML records classified as farms based on their responses to the census report form.

To estimate these probabilities ( $\hat{\pi}_C$  and  $\hat{\pi}_{CCFC}$ ), the records in the 2022 ACES sample were matched to the 2022 CML using probabilistic record linkage allowing the records only on the CML, ACES, and on both the CML and ACES to be identified. All CML records and ACES tracts were used to estimate the capture-recapture probabilities jointly.

### Capture Probabilities

Recall that, for a farm to be identified as a farm, and thus captured, by the census, it must be on the CML, respond to either the census or ACES report form and, based on that response, be classified as a farm. These adjustments are dependent. Therefore, the probability of capture  $\pi_C$  may be written as

$$\pi_C = \pi(\text{CML, Responded, Farm on Census}|\text{Farm}) = \pi(\text{CML}|\text{Farm})\pi(\text{Responded}|\text{CML, Farm})\pi(\text{Farm on Census}|\text{CML, Responded, Farm})$$

Terms in the probability of capturing a farm depend on the characteristics of the farm. Using all Census and ACES data, model variables were selected by applying a stepwise variable selection algorithm and expert opinion. Estimation was based on a conditional weighted likelihood. The events of a farm not being included in either ACES or the CML was excluded from the likelihood but was accounted for through the model's capture-recapture properties. Although the probability of capture is estimated for both CML and ACES records, only CML records with a census response are given a census weight; records with only an ACES response are not given a census weight or used further to produce census estimates.

### Misclassification

An operation is misclassified if: (1) it meets the definition of a farm but is classified as a nonfarm on the census or (2) it does not meet the definition of a farm but is classified as a farm on the census. The first type of misclassification is accounted for when modeling the probability of capture. An adjustment is still needed for the misclassification of nonfarms as farms. As with farm status and capture, the probability of this misclassification

formulario censal recibió un peso que es la proporción de la probabilidad estimada de clasificación correcta de una finca en el censo y la probabilidad estimada de captura ( $\hat{\pi}_C$  and  $\hat{\pi}_{CCFC}$ ) donde el símbolo del sombrero (^) denota un estimado. Para estimar el número de fincas con un conjunto dado de características, se sumaron los pesos de los registros de CML que respondieron como fincas en el censo y que tenían ese conjunto de características. Este estimador se conoce como el estimador de captura-recaptura (CR):

$$CR = \sum_{i \in F} \frac{\hat{\pi}_{CCFC,i}}{\hat{\pi}_{C,i}}$$

donde  $F$  es el conjunto de todos los registros CML clasificados como fincas en función de sus respuestas al formulario censal.

Para estimar estas probabilidades ( $\hat{\pi}_C$  and  $\hat{\pi}_{CCFC}$ ) los registros de la muestra de ACES de 2022 se compararon con la CML de 2022 mediante vinculación de registros probabilísticos, lo que permitió identificar los registros solo en la CML, la ACES y tanto en la CML como en la ACES. Todos los registros de CML y los tractos ACES se utilizaron para estimar las probabilidades de captura y recaptura de forma conjunta.

### Probabilidades de Captura

Recuerde que, para que una finca se identifique como una finca y, por lo tanto, sea capturada por el censo, debe estar en la CML, responder al censo o al formulario de ACES y, en base a esa respuesta, ser clasificado como una finca. Estos ajustes son dependientes. Por lo tanto, la probabilidad de captura  $\pi_C$  puede escribirse como

$$\pi_C = \pi(\text{CML, Respondió, Finca en censo}|\text{Finca}) = \pi(\text{CML}|\text{Finca})\pi(\text{Respondió}|\text{CML, Finca})\pi(\text{Finca en censo}|\text{CML, Respondió, Finca})$$

Los términos en la probabilidad de capturar una finca dependen de las características de la finca. Utilizando todos los datos del Censo y ACES, las variables del modelo se seleccionaron aplicando un algoritmo de selección de variables paso-a-paso y la opinión de expertos. La estimación se basó en una probabilidad ponderada condicional. Los eventos de una finca que no estuviera incluida ni en ACES ni en CML se excluyeron de la probabilidad, pero se contabilizaron a través de las propiedades de captura-recaptura del modelo. Aunque la probabilidad de captura se estima tanto para los registros CML como para ACES, sólo los registros CML con una respuesta censal reciben una ponderación censal; los registros con solo una respuesta ACES no reciben una ponderación censal ni se utilizan más para producir

depends on an operation's characteristics. Thus, a conditional logistic model was developed. Given that a farm on the CML was classified as a farm in the census, the probability of its being a farm was modeled based on its characteristics.

## **CALIBRATION**

Each operation identified as a farm on the CML was given a weight equal to the probability of misclassification divided by the probability of capture. This weight accounted for undercoverage, nonresponse, and both types of misclassification. These initial weights were used to calculate the farm operation coverage targets.

For calibration there were six farm operation targets: total number of farms, land in farms, three breaks for farms by value of agricultural sales, and female principal operators. These Island-level estimates were general purpose in that they did not provide any control over expected levels of commodity production of the individual farm operation. As a result of this limitation, the procedures could have over-adjusted or under-adjusted for commodity production. To address this, a second set of variables, known as commodity targets, was added to the calibration algorithm. These targets were commodity totals from administrative sources (e.g. Puerto Rico's Department of Agriculture).

The introduction of these commodity coverage targets strengthened the overall adjustment procedure by ensuring that major commodity totals remained within reasonable bounds of established benchmarks.

The calibration algorithm addressed commodity coverage. The algorithm was controlled by the 6 Island-level farm operation coverage targets and 7 Island-level commodity coverage targets. Because calibration targets are estimates subject to uncertainty, NASS allowed some tolerance in the determination of the adjusted weights. Rather than forcing the total for each calibration variable computed using the adjusted weights to equal a specific amount, NASS allowed the estimated total to fall within a tolerance range.

Tolerance ranges for the farm operation coverage targets were determined differently from the commodity targets. The tolerance range for the six Island-level farm operation coverage targets was the estimated Island total for the variable plus or minus one standard error of the capture-recapture

estimaciones censales.

## **Clasificación Errónea**

Una operación se clasifica erróneamente si: (1) cumple con la definición de finca, pero está clasificada como no agrícola en el censo o (2) no cumple con la definición de finca, pero está clasificada como finca en el censo. El primer tipo de clasificación errónea se estima al modelar la probabilidad de captura. Todavía se necesita un ajuste para la clasificación errónea de las no fincas como fincas. Al igual que la captura de la finca, la probabilidad de esta clasificación errónea depende de las características de la operación. Así, se desarrolló un modelo logístico condicional. Dado que una finca en la CML fue clasificada como finca en el censo, se modeló la probabilidad de que fuera una finca en función de sus características.

## **CALIBRACIÓN**

A cada operación identificada como finca en la CML se le asignó un peso igual a la probabilidad de clasificación errónea dividido por la probabilidad de captura. Este peso explicaba la falta de cobertura, la falta de respuesta, y ambos tipos de clasificación errónea. Estos pesos iniciales se usaron para calcular los objetivos de cobertura de las fincas.

El proceso de calibración se basó en seis objetivos de operación de la finca: número total de fincas, tierra en fincas, tres agrupamientos de las fincas basado en el valor de ventas agrícolas y principales operadoras. Estas estimaciones a nivel de la isla tenían un propósito general ya que no proporcionaban ningún control sobre los niveles esperados de producción de productos básicos de la operación de la finca individual. Como resultado de esta limitación, los procedimientos podrían haber sobre-ajustado o sub-ajustado la producción de productos básicos. Para mitigar esta limitación, se agregó un segundo conjunto de variables, conocidas como objetivos de productos básicos, al algoritmo de calibración. Estos objetivos fueron los totales de productos proveídos por fuentes administrativas (por ejemplo, el Departamento de Agricultura de Puerto Rico).

La introducción de estos objetivos de cobertura de productos básicos fortaleció el procedimiento de ajuste ya que garantizaron que los principales totales de productos permanecieran dentro de los límites razonables de los puntos de referencia establecidos.

El algoritmo de calibración mitigó la cobertura de productos básicos. El algoritmo fue controlado por los 6 objetivos de cobertura de operación de finca a nivel de isla y 7 objetivos de cobertura de productos básicos a nivel de la isla. Como los objetivos de calibración son estimaciones

estimate. Commodity coverage targets with acceptable ranges were established based on the administrative source for the Island. Ranges were not necessarily symmetric around the target value.

To ensure that all subdomains for which NASS publishes summed to their grand total, integer weights were produced by a discrete calibration algorithm. This eliminated the need for rounding individual cell values and ensured that marginal totals always added correctly to the grand total. If a weight was initially not in the interval [1,6], it was trimmed so that it was in that interval. That is, adjusted weights less than 1 were set to 1, and those greater than 6 were set to 6. The remaining non-integer weights were then rounded sequentially to reduce the distance of the estimated totals from the targets.

Calibration adjustments began with the computation of a priority index for each record. The priority index was the absolute value of the gradient of the relative error associated with increasing or decreasing a record's weight by one. The record with the highest priority index was then selected as a candidate to increase or decrease its weight by one to reduce the cumulative distance from the targets as measured by the relative error. If the new value produced an improvement and satisfied the range restrictions, the weight was updated and new priorities were assigned; otherwise, the record with the next highest priority index was processed. This process was iteratively performed until convergence was attained. Because census data collection was assumed to be complete for very large and unique farms, their weights were controlled to 1 during the calibration adjustment process. For all other farms, the final census record weights were forced to be an integer number in the interval [1, 6]. The calibration process considered all targets simultaneously through the priority index. Although calibration was seldom not able to adjust weights so that all Island-level targets were met, all targets were brought collectively as close to the targets as possible.

The proportions of selected census data items that were due to coverage, response, and classification adjustments are displayed in Table A.

## DISCLOSURE REVIEW

After tabulation and review of the aggregates, a comprehensive disclosure review will be conducted. NASS is obligated to withhold, under

sujetas a incertidumbre, NASS permitió cierta tolerancia en la determinación de los pesos ajustados. En lugar de forzar el total para cada variable de calibración calculada usando los pesos ajustados para igualar una cantidad específica, NASS permitió que el total estimado cayera dentro de un rango de tolerancia. Los rangos de tolerancia para los objetivos de cobertura de operación de la finca se determinaron de manera diferente a los objetivos de productos básicos. El rango de tolerancia para los seis objetivos de cobertura de operación de la finca a nivel de la isla fue el total estimado de la isla para la variable más o menos un error estándar del estimado de captura-recaptura. Se establecieron objetivos de cobertura de productos básicos con rangos aceptables basados en la fuente administrativa de la isla. Los rangos no eran necesariamente simétricos alrededor del valor objetivo.

Para garantizar que todos los subdominios para los que publica NASS se sumen a su gran total, los pesos enteros se produjeron mediante un algoritmo de calibración discreto. Esto eliminó la necesidad de redondear los valores de las celdas individuales y aseguró que los totales marginales siempre se agregaran correctamente al total general. Si un peso no estaba inicialmente en el intervalo [1,6], se recortó para que estuviera en ese intervalo. Es decir, los pesos ajustados menores de 1 se establecieron en 1, y los mayores de 6 se establecieron en 6. Los pesos no enteros restantes se redondearon secuencialmente para reducir la distancia de los totales estimados a los objetivos.

Los ajustes de calibración comenzaron con el cálculo de un índice de prioridad para cada registro. El índice de prioridad era el valor absoluto del gradiente del error relativo asociado con aumentar o disminuir el peso de un registro en uno. El registro con el índice de mayor prioridad se seleccionó como candidato para aumentar o disminuir su peso en uno para reducir la distancia acumulativa de los objetivos según lo medido por el error relativo. Si el nuevo valor producía una mejora y satisfacía las restricciones de rango, el peso se actualizaba y se asignaban nuevas prioridades. De lo contrario, se procesó el registro con el siguiente índice de mayor prioridad. Este proceso se realizó de forma iterativa hasta que se logró la convergencia. Debido a que se asumió que la recopilación de datos del censo estaba completa para fincas muy grandes y únicas, sus pesos se controlaron a 1 durante el proceso de ajuste de calibración. Para todas las otras fincas, los pesos finales del registro del censo se vieron obligados a ser un número entero en el intervalo [1, 6]. El proceso de calibración consideró todos los objetivos simultáneamente a través del índice de prioridad. Aunque la calibración rara vez pudo ajustar los pesos para que se cumplieran todos los objetivos a nivel de la isla, todos los objetivos se acercaron colectivamente a los objetivos lo más posible.

Title 7, U.S. Code, any total that would reveal an individual's information or allow it to be closely estimated by the public. Cell suppression will be used to protect the cells that are determined to be sensitive to a disclosure of information. Farm counts are not considered sensitive and are not subject to disclosure.

Based on agency standards, a data disclosure risk is determined to be present if a particular data cell being considered for publication violates either one of two criteria. First, the threshold rule is violated if the estimated number of farms in a data cell is less than three. For example, if the estimate of the number of farms producing hogs in a region is equal to exactly one farm, NASS could not publish the regional total for hog inventory without disclosing the reporting farm's information. Second, the dominance rule fails if the distribution of the data within the cell allowed a data user to estimate any respondent's data too closely. For example, if many farmers produce hogs in a region and some of them were large enough to dominate the cell total, NASS could not publish the regional total for hog inventory without risking disclosing an individual respondent's data. In both situations, the data would be suppressed and a "(D)" is placed in the cell in the publication table. These data cells are referred to as primary suppressions.

Since most items will be summed to marginal totals, primary suppressions within these summation relationships are protected by ensuring that there are additional suppressions within the linear relationship that provide adequate protection for the primary. A detailed computer routine selects additional data cells for suppression to ensure all primary suppressions are properly protected in all linear relationships in all tables. These data cells are referred to as complementary suppressions. These cells are not themselves sensitive to a disclosure but are suppressed to protect other primary suppressions. A "(D)" is placed in the cell of the publication table to indicate a complementary suppression.

NASS analysts will review all complementary suppressions to ensure no cells are withheld that are vital to the data users. In instances where complementary suppressions are deemed critically important to the Island, analysts can request an override and a different complement can be chosen.

Las proporciones de los elementos de datos censales seleccionados que se debieron a la cobertura, la respuesta, y los ajustes de clasificación se muestran en la Tabla A.

## REVISIÓN DE DIVULGACIÓN

Después de la tabulación y revisión de los agregados, se realizará una revisión exhaustiva de la divulgación. Conforme al Título 7, Código de los EE. UU., NASS está obligado a retener cualquier total que revele la información de un individuo o permita que el público lo pueda calcular. La supresión de celda se utilizará para proteger las celdas que se consideran sensibles a la divulgación de información. El número de fincas no se considera información sensible y no está sujeto a divulgación.

Según los estándares de la agencia, se determina que existe un riesgo de divulgación de datos si una celda de datos que se considera para publicación viola uno de dos criterios. Primero, se infringe la regla de umbral si el número estimado de fincas en una celda de datos es menor de tres. Por ejemplo, si la estimación del número de fincas que producen cerdos en una región es exactamente igual a una finca, NASS no podría publicar el inventario de cerdos para el total de la región sin revelar la información de la finca que informa. En segundo criterio, se infringe la regla de dominancia cuando la distribución de los datos dentro de una celda, permiten a un usuario de datos estimar los datos de cualquier encuestado con precisión. Por ejemplo, si muchos agricultores producen cerdos en una región y algunos de ellos son lo suficientemente grandes como para dominar el total de la celda, NASS no puede publicar el total regional para el inventario de cerdos sin arriesgarse a revelar los datos de un encuestado individual. En ambas situaciones, los datos se suprimirían y se colocaría una "(D)" en la celda en la tabla de publicación. Estas celdas de datos se denominan supresiones primarias.

Dado que la mayoría de los artículos se sumarán a totales marginales, las supresiones primarias dentro de estas relaciones de suma se protegen asegurando que haya supresiones adicionales dentro de la relación lineal que brinden protección adecuada para la primaria. Una rutina informática detallada selecciona celdas de datos adicionales para supresión que garantizan que todas las supresiones primarias estén protegidas adecuadamente en todas las relaciones lineales en todas las tablas. Estas celdas de datos se denominan supresiones complementarias. Estas celdas no son sensibles a una divulgación, pero se suprimen para proteger otras supresiones primarias. Se coloca una "(D)" en la celda de la tabla de publicación para indicar una supresión complementaria.

Los analistas del NASS revisarán todas las supresiones complementarias para garantizar que no se retengan celdas que sean vitales para los usuarios de datos. En los casos en

## CENSUS QUALITY

The purpose of the census of agriculture is to account for “any place from which \$500 or more of agricultural products were produced and sold, or normally would have been sold, during the census year.” To accomplish this, NASS develops a CML that contains identifying information for operations that have an indication of meeting the census definition, develops procedures to collect agricultural information from those records, establishes criteria for analyst review of the data, creates computer routines to correct or complete the requested information, and provides census estimates of the characteristics of farms and farm operators with associated measures of uncertainty.

It is not likely that either the CML includes all operations that meet the definition of a farm or that all those that do meet the definition of a farm respond to the census inquiry. The goal is to publish data with a high level of quality. The quality of a census may be measured in many ways. One of the first indicators used is a measure of the response to the census data collection as it has generally been thought that a high response rate indicates more complete coverage of the population of interest. This is a valid assumption if the enumeration list, the CML here, has complete coverage of the population of interest. In the case of the census of agriculture, the definition requiring advance knowledge of sales makes achieving a high level of coverage difficult. To ensure that the census of agriculture is as complete as possible, records are included that might not meet the definition of a farm. A second indicator of quality then is the coverage of the farm population by the CML. Other indicators of quality relate to the accuracy and completeness of the data, and the validity of the procedures used in processing the data.

In some cases, NASS was able to produce measures of quality – such as the response rate to the data collection, the coverage of the census mail list, and the variability of the final adjusted estimates. In other cases, measures were not produced but descriptions of procedures that NASS used to reduce errors from the procedures were subsequently provided.

### Census Response Rate

The response rate is one indicator of the quality of a data collection. It is generally assumed that if a response rate is close to a full participation level of

que las supresiones complementarias se consideran críticamente importantes para la Isla, los analistas pueden solicitar una anulación y se puede elegir un complemento diferente.

## CALIDAD DEL CENSO

El propósito del censo de agricultura es dar cuenta de "cualquier lugar desde el cual se produjeron y vendieron \$ 500 o más de productos agrícolas, o normalmente se habrían vendido, durante el año del censo". Para lograr esto, NASS desarrolla la CML que contiene información de identificación para operaciones que tienen una indicación de cumplir con la definición del censo, desarrolla procedimientos para recopilar información agrícola de esos registros, establece criterios para la revisión de los datos por parte de analistas, crea rutinas informáticas para corregir o completar la información solicitada, y proporciona estimaciones del censo de las características de las fincas y los operadores de fincas con medidas de incertidumbre asociadas.

No es probable que la CML incluya todas las operaciones que cumplan con la definición de finca o que todas las que cumplan con la definición de finca respondan a la consulta del censo. El objetivo es publicar datos con un alto nivel de calidad. La calidad de un censo puede medirse de muchas maneras. Uno de los primeros indicadores utilizados es una medida de la respuesta a la recopilación de datos del censo, ya que generalmente se ha pensado que una tasa de respuesta alta indica una cobertura más completa de la población de interés. Esta es una suposición válida si la lista de enumeración, la CML, en este caso, tiene una cobertura completa de la población de interés. En el caso del censo de agricultura, la definición que requiere un conocimiento anticipado de las ventas dificulta el logro de un alto nivel de cobertura. Para garantizar que el censo de la agricultura sea lo más completo posible, se incluyen registros que podrían no cumplir con la definición de finca. Un segundo indicador de calidad es la cobertura de la población agrícola por parte de la CML. Otros indicadores de calidad se relacionan con la precisión e integridad de los datos, y la validez de los procedimientos utilizados en el procesamiento de los datos.

En algunos casos, NASS pudo producir medidas de calidad, como la tasa de respuesta a la recopilación de datos, la cobertura de la lista de correo del censo y la variabilidad de las estimaciones finales ajustadas. En otros casos, no se produjeron medidas, pero posteriormente se proporcionaron descripciones de los procedimientos que NASS utilizó para reducir los errores de los procedimientos.



100 percent, the potential for nonresponse bias is small, although this is not always true. The response rate for the 2022 Puerto Rico Census of Agriculture CML was 42.0 percent, as compared with the 2018 Puerto Rico Census of Agriculture's response rate of 29.1 percent.

The 2022 Puerto Rico Census of Agriculture's response rate used the fourth response rate formula (RR4) from the American Association of Public Opinion Research's Response Rate Standard Definitions manual:

$$RR4 = \frac{C_{adj}}{C_{adj} + R + NC + O + Replicated + e(U)} (100)$$

where

$C_{adj}$  = number of fully and partially completed records, excluding replicated records

$R$  = number of explicit refusals

$NC$  = number of non-contacted operations known to be eligible

$O$  = number of other types of nonrespondents

$Replicated$  = number of replicated records

$U$  = number of operations of unknown eligibility

$e(U)$  = estimated number of operations of unknown eligibility assumed to be eligible

Records were classified into the above variables based on the combination of their active status (AS) codes, in-scope status, and replication status. Active status refers to the eligibility status of records for selection on the CML.

Certain active status classifications indicated records of unknown agricultural status. These classifications included records to be removed from the CML but had data from outside sources indicating agricultural activity, new records from outside data sources, and records with FSA or CRP on operations that are not owned by the principal operator. These records were stratified (grouped) based on their probabilities of being in-scope had they responded. The estimated number of in-scope nonrespondents was calculated for the  $h$ th stratum (group) by the following formula:

$$e(U_h) = \left( \frac{C_{in-scope,h}}{C_h} \right) U_h$$

where

## Tasa de Respuesta al Censo

La tasa de respuesta es un indicador de la calidad de una recopilación de datos. En general, se supone que si una tasa de respuesta está cerca de un nivel de participación total del 100 por ciento, el potencial de sesgo de no respuesta es pequeño, aunque esto no siempre es cierto. La tasa de respuesta para la CML del Censo de Agricultura de Puerto Rico de 2022 fue del 42.0 por ciento, en comparación con la tasa de respuesta del Censo de Agricultura de Puerto Rico de 2018 del 29.1 por ciento.

La tasa de respuesta del Censo de Agricultura de Puerto Rico de 2022 utilizó la cuarta fórmula de tasa de respuesta (RR4) del manual de definiciones estándar de la tasa de respuesta de la Asociación Americana de Investigación de Opinión Pública:

$$RR4 = \frac{C_{adj}}{C_{adj} + R + NC + O + Replicated + e(U)} (100)$$

dónde

$C_{adj}$  = cantidad de registros completados y parcialmente completados, excluyendo registros replicados

$R$  = cantidad de rechazos explícitos

$NC$  = cantidad de operaciones no contactadas que se sabe que son elegibles

$O$  = cantidad de otros tipos de no encuestados

$Replicated$  = cantidad de registros replicados

$U$  = número de operaciones de elegibilidad desconocida

$e(U)$  = se estima que el número de operaciones de elegibilidad desconocida es elegible

Los registros se clasificaron en las variables anteriores en función de la combinación de sus códigos de estado activo (AS, por sus siglas en inglés), estado de alcance y estado de replicación. El estado activo se refiere al estado de elegibilidad de los registros para la selección en la CML.

Ciertas clasificaciones de estado activo indicaron registros de estado agrícola desconocido. Estas clasificaciones incluían registros que se eliminarían de la CML, pero tenían datos de fuentes externas que indicaban actividad agrícola, nuevos registros de fuentes de datos externas, y registros con datos de FSA o CRP sobre operaciones que no son propiedad del productor principal. Estos registros se estratificaron (agruparon) en función de sus probabilidades de estar dentro del alcance si hubieran respondido. El número estimado de no respondedores dentro del alcance se calculó para el  $h$ ésimo estrato (grupo) mediante la siguiente fórmula:



$e(U_h)$  = estimated number of operations of unknown eligibility assumed to be eligible in the  $h$ th group

$C_{in-scope,h}$  = the number of completed and in-scope census records in the  $h$ th group

$C_h$  = the number of completed census records in the  $h$ th group

$U_h$  = number of operations of unknown eligibility in the  $h$ th group

## Census Coverage

As a side-product of the statistical adjustment used to account for undercoverage, nonresponse of farms on the CML, and misclassification of responses to the census, the proportion of the adjustments due to each of those factors can be derived. The percentages of final census estimates due to adjustments for undercoverage, nonresponse, and misclassification as well as the total percent adjustment for selected items are displayed in Tables A and C.

## MEASURED ERRORS IN THE CENSUS PROCESS

Uncertainty is introduced into the data in compiling the CML, in NASS's data collection procedures, in data editing and processing, and in compiling the final data. Additionally, NASS uses statistical procedures to both measure errors in the various processes and in making adjustments for those errors in the final data. One example is the statistical process used to account for undercoverage, nonresponse of farms on the CML, and misclassification of responses to the census. The basis of the undercoverage adjustment is the capture-recapture procedure that uses the ACES. The largest contributors to error in the census estimates are due to the adjustments for nonresponse, undercoverage, misclassification, and calibration.

## Variability in Census Estimates due to Statistical Adjustment

In conducting the 2022 Puerto Rico Census of Agriculture, efforts were initiated to measure error associated with the adjustments for farm operations that were not on the CML, for farm operations that were on the CML but did not respond to the census report form, and for farms and nonfarms that were misclassified as nonfarms and farms, respectively, and for integer calibration. These error measurements were developed from the standard error of the estimates at the Island, regional, and municipio levels and were expressed as

$$e(U_h) = \left( \frac{C_{in-scope,h}}{C_h} \right) U_h$$

dónde

$e(U_h)$  = se estima que el número de operaciones de elegibilidad desconocida es elegible en el grupo  $h$ th

$C_{in-scope,h}$  = la cantidad de registros censales completados y dentro del alcance en el grupo  $h$ th

$C_h$  = El número de registros censales completados en el grupo  $h$ th

$U_h$  = Número de operaciones de elegibilidad desconocida en el grupo  $h$ th

## Cobertura del Censo

Como un producto secundario del ajuste estadístico utilizado para dar cuenta de la cobertura insuficiente, la falta de respuesta de las fincas en la CML, y la clasificación errónea de las respuestas al censo, se puede derivar la proporción de los ajustes debidos a cada uno de esos factores. Los porcentajes de las estimaciones finales del censo debido a ajustes por subcobertura, falta de respuesta y clasificación errónea, así como el ajuste porcentual total para los elementos seleccionados se muestran en las Tablas A y C.

## ERRORES MEDIDOS EN EL PROCESO DE CENSO

Se introduce incertidumbre en los datos al compilar la CML, en los procedimientos de recopilación de datos de NASS, en la edición y procesamiento de datos, y en la compilación de los datos finales. Además, NASS utiliza procedimientos estadísticos tanto para medir errores en los diversos procesos como para realizar ajustes para esos errores en los datos finales. Un ejemplo es el proceso estadístico utilizado para dar cuenta de la cobertura insuficiente, la falta de respuesta de las fincas en la CML, y la clasificación errónea de las respuestas al censo. La base del ajuste de cobertura insuficiente es el procedimiento de captura-recaptura que utiliza el ACES. Los mayores contribuyentes al error en las estimaciones del censo se deben a los ajustes por falta de respuesta, cobertura insuficiente, clasificación errónea, y calibración.

## Variabilidad en las Estimaciones del Censo Debido al Ajuste Estadístico

Al realizar el Censo de Agricultura de Puerto Rico de 2022, se iniciaron esfuerzos para medir el error asociado con los ajustes para las operaciones agrícolas que no estaban en la CML, para las operaciones agrícolas que estaban en la CML pero que no respondieron al formulario censal, y para las fincas y no fincas que fueron clasificadas erróneamente como no fincas y fincas, respectivamente, y para calibración entera. Estas mediciones de error se desarrollaron a partir del error estándar de las estimaciones

coefficients of variation (CVs) at the Island, regional, and municipio levels.

The standard error of an estimate is an estimate of the standard deviation of the sampling distribution of the estimator. In each case, standard errors were computed using an approach based on a delete-a-group jackknife methodologies. To conduct the jackknifing,  $k = 10$  mutually exclusive and exhaustive groups of records were formed. The groups were selected using a stratified random design so that each group reflected capture status by the CML and ACES. Based on estimated weights for records in each group, a delete-a-group jackknife estimator of the variance would account for the uncertainty associated with modeling the capture-recapture probabilities and the uncertainty due to integer calibration. Therefore, the weights within each jackknife group were computed using the group-specific models and calibrated to match group-specific targets. For a given data item  $i$ , such as the number of farms, the estimate  $T_i$  was computed at the specified geographical level, such as Island, region, or municipio, using the  $(k - 1)$  groups remaining after deleting the calibration-adjusted jackknife group  $j$ . Estimates of the variance and standard error associated with the estimator  $T_i$  are then, respectively,

$$\sigma_i^2 = \frac{k-1}{k} \sum_{j=1}^k \left( T_i^{(j)} - \sum_{l=1}^k \frac{T_i^{(l)}}{k} \right)^2; \quad SE(T_i) = \sqrt{\sigma_i^2}$$

Increasing  $k$  improves the estimate of the variance but, as  $k$  increases, the observations become too sparse to reflect the survey design and to provide Island-wide coverage. Ten (10) calibration-adjusted jackknife groups were used to provide standard errors for island and regional estimates. For the estimate of the number of farms with a given set of characteristics, only the CML records with those characteristics were used to obtain the overall estimate as well as the estimates from each calibration-adjusted jackknife group.

Note that the calibrated jackknife groups are only constructed once, and different subsets of the records are used to compute estimates and standard errors for the data items.

The CV is a measure of the relative amount of error associated with the sample estimate:

a nivel de la isla, región, y municipio, y se expresaron como coeficientes de variación (CV) a nivel de la isla, región y municipio.

El error estándar de una estimación es una estimación de la desviación estándar derivada de una muestra particular usada para computar el estimador. En cada caso, los errores estándar se calcularon utilizando un enfoque basado en metodologías jackknife de eliminación de un grupo. Para realizar el jackknifing, se formaron  $k = 10$  grupos de registros exhaustivos y mutuamente excluyentes. Los grupos se seleccionaron mediante un diseño aleatorio estratificado para que cada grupo reflejara el estado de captura por parte de CML y ACES. Con base en los pesos estimados para los registros en cada grupo, un estimador de la varianza tipo jackknife de eliminación de un grupo tomaría en cuenta la incertidumbre asociada con el modelado de las probabilidades de captura-recaptura y la incertidumbre debida a la calibración de números enteros. Por lo tanto, las ponderaciones dentro de cada grupo de jackknife se calcularon utilizando los modelos específicos del grupo y se calibraron para que coincidan con los objetivos específicos del grupo. Para un elemento de datos dado  $i$ , como el número de granjas, la estimación  $T_i$  se calculó en el nivel geográfico especificado, como Isla, región, o municipio, utilizando los  $(k - 1)$  grupos restantes después de eliminar el grupo jackknife ajustado por calibración  $j$ . Las estimaciones de la varianza y el error estándar asociados con el estimador  $T_i$  son entonces, respectivamente,

$$\sigma_i^2 = \frac{k-1}{k} \sum_{j=1}^k \left( T_i^{(j)} - \sum_{l=1}^k \frac{T_i^{(l)}}{k} \right)^2; \quad SE(T_i) = \sqrt{\sigma_i^2}$$

Cuando se incrementa  $k$  mejora la estimación de la varianza, pero a medida que aumenta  $k$ , las observaciones se vuelven demasiado escasas para reflejar el diseño de la encuesta y para proporcionar cobertura en toda la isla. Se utilizaron diez (10) grupos de jackknife con ajustes de calibración para proporcionar errores estándar para las estimaciones regionales e insulares. Para la estimación del número de fincas con un conjunto dado de características, solo se usaron los registros de CML con esas características para obtener la estimación general, así como las estimaciones de cada grupo jackknife con ajustes de calibración.

Tenga en cuenta que los grupos jackknife calibrados solo se construyeron una vez, y se usaron diferentes subconjuntos de los registros para calcular las estimaciones y los errores estándar para los elementos de datos.

El CV es una medida de la cantidad relativa de error asociado con la estimación de la muestra:

$$CV_i = \frac{SE(T_i)}{T_i} 100\%$$

where  $SE(T_i)$  is the standard error of the capture recapture estimate for data item  $i$ . This relative measure allows the reliability of a range of estimates to be compared. For example, the standard error is often larger for large population estimates than for small population estimates, but the large population estimates may have a smaller CV, indicating a more reliable estimate.

Note: The standard errors and consequently the CVs, tend to be substantially smaller than those reported for the 2018 Puerto Rico Census of Agriculture. For 2018, the model of the probability of capture incorporated information from the 360 respondents to the 2018 ACES and the census records matched to an ACES record. In contrast, the models for the 2022 Puerto Rico Census of Agriculture relied on information from the approximately 3,800 responding CML records and the 2022 ACES, some of which were on both the CML and ACES. The large increase in the number of records used in the modeling process led to a major decrease in the measures of uncertainty (standard errors and CVs).

Table B presents the fully adjusted estimates with the coefficient of variation for selected items.

## NONMEASURED ERRORS IN THE CENSUS PROCESS

As noted in the previous section, sampling errors can be introduced from the coverage, nonresponse, and misclassification adjustment procedures. This is a measurable error. However, nonsampling errors are imbedded in the census process that cannot be directly measured as part of the design of the census but must be contained to ensure an accurate count. Extensive efforts were made to compile a complete and accurate mail list for the census, to elicit response to the census, to design an understandable report form with clear instructions, to minimize processing errors through the use of quality control measures, to reduce matching error associated with the capture-recapture estimation process, and to minimize error associated with identification of a respondent as a farm operation (referred to as classification error). The weight adjustment and tabulation processes recognize the presence of nonsampling errors; however, it is assumed that these errors are small and that, in total, the net effect is zero. In other words, the

$$CV_i = \frac{SE(T_i)}{T_i} 100\%$$

donde  $SE(T_i)$  es el error estándar de la estimación de recuperación de captura para el elemento de datos  $i$ . Esta medida relativa permite comparar la fiabilidad de un rango de estimaciones. Por ejemplo, el error estándar es a menudo mayor para estimaciones de población grande que para estimaciones de población pequeña, pero las estimaciones de población grande pueden tener un CV más pequeño, lo que indica una estimación más confiable.

Nota: Los errores estándar y, en consecuencia, los CV, tienden a ser sustancialmente más pequeños que los reportados para el Censo Agrícola de Puerto Rico de 2018. Para 2018, el modelo de probabilidad de captura incorporó información de los 360 encuestados de la ACES 2018 y los registros censales cotejados con un registro de ACES. Por contrato, los modelos para el Censo Agrícola de Puerto Rico de 2022 se basaron en información de aproximadamente 3,800 registros de CML que respondieron y del ACES de 2022, algunos de los cuales estaban tanto en el CML como en el ACES. El gran aumento en el número de registros utilizados en el proceso de modelado condujo a una disminución importante en las medidas de incertidumbre (errores estándar y CVs).

La Tabla B presenta las estimaciones totalmente ajustadas con el coeficiente de variación para los elementos seleccionados.

## ERRORES NO MEDIDOS EN EL PROCESO DE CENSO

Como se señaló en la sección anterior, se pueden introducir errores de muestreo a partir de los procedimientos de ajuste de cobertura, falta de respuesta, y clasificación errónea. Este error es medible. Sin embargo, los errores ajenos al muestreo están integrados en el proceso del censo que no pueden medirse directamente como parte del diseño del censo, sino que deben incluirse para garantizar un recuento preciso. Se hicieron grandes esfuerzos para compilar una lista de correo completa y precisa para el censo, para obtener una respuesta al censo, para diseñar un formulario censal comprensible con instrucciones claras, para minimizar los errores de procesamiento mediante el uso de medidas de control de calidad, para reducir los errores de correspondencia asociados con el proceso de estimación de captura-recaptura y para minimizar el error asociado con la identificación de un encuestado como una operación de finca (denominado error de clasificación). Los procesos de ajuste de peso y tabulación reconocen la presencia de errores que no son de muestreo; sin embargo, se supone que estos errores son pequeños y que, en total, el efecto neto es

positive errors cancel the negative errors.

### **Respondent and Enumerator Error**

Incorrect or incomplete responses to the census report form or to the questions posed by an enumerator can introduce error into the census data. Steps were taken in the design and execution of the census of agriculture to reduce errors from respondent reporting. Poor instructions and ambiguous definitions lead to misreporting. Respondents may not remember accurately, may estimate responses, or may record an item in the wrong cell. To reduce reporting and recording errors, the report form was tested prior to the census using industry accepted cognitive testing procedures. Detailed instructions for completing the report form were provided to each respondent. Questions were phrased as clearly as possible based on previous tests of the report form. In addition, each respondent's answers were checked for completeness and consistency by the complex edit and imputation system.

### **Processing Error**

Processing of each census report form was another potential source of nonsampling error. All mail returns that included multiple reports, respondent remarks, or that were marked out of business and report forms with no reported data were sent to an analyst for verification and appropriate action. Integrity checks were performed by the imaging system and data transfer functions. Standard quality control procedures were in place that required that randomly selected batches of data keyed from image to be re-entered by a different operator to verify the work and evaluate key entry operators. All systems and programs were thoroughly tested before going on-line and were monitored throughout the processing period.

Developing accurate processing methods is complicated by the complex structure of agriculture. Among the complexities are the many places to be included, the variety of arrangements under which farms are operated, the continuing changes in the relationship of operators to the farm operated, the expiration of leases and the initiation or renewal of leases, the problem of obtaining a complete list of agriculture operations, the difficulty of contacting and identifying some types of contractor/contractee relationships, the operator's absence from the farm during the data collection period, and the operator's opinion that part or all of the operation does not qualify and

cero. En otras palabras, los errores positivos cancelan los errores negativos.

### **Error de Encuestado y Enumerador**

Las respuestas incorrectas o incompletas al formulario censal o a las preguntas formuladas por un enumerador pueden introducir errores en los datos del censo. Se tomaron medidas en el diseño y la ejecución del censo de agricultura para reducir estos errores. Las instrucciones inadecuadas y las definiciones ambiguas conducen a informes erróneos. Los encuestados pueden no recordar con precisión, pueden estimar las respuestas o pueden escribir una respuesta en el lugar incorrecto. Para reducir los errores de respuesta y registro, el formulario se probó antes del censo utilizando procedimientos de prueba cognitiva aceptados por la industria. Se proporcionaron instrucciones detalladas para completar el formulario a cada encuestado. Las preguntas fueron formuladas de la manera más clara posible en base a pruebas previas del formulario. Además, el complejo sistema de edición e imputación verificó la integridad y consistencia de las respuestas de cada encuestado.

### **Error de Proceso**

El procesamiento de cada formulario censal fue otra fuente potencial de errores muestrales. Todas las devoluciones de correo que incluían múltiples informes, comentarios de los encuestados, o que se marcaron fuera de negocio y los formularios sin datos informados se enviaron a un analista para su verificación y la acción adecuada. Los controles de integridad fueron realizados por el sistema de imágenes y las funciones de transferencia de datos. Se implementaron procedimientos estándar de control de calidad que requerían que grupos diferentes de datos seleccionados al azar de la imagen fueran reingresados por un operador diferente para verificar el trabajo y evaluar los operadores clave de entrada. Todos los sistemas y programas se probaron exhaustivamente antes de conectarse y se monitorearon durante todo el período de procesamiento.

El desarrollo de métodos de procesamiento precisos se complica por la compleja estructura de la agricultura. Entre las complejidades se encuentran los muchos lugares a incluir, la variedad de arreglos bajo los cuales se operan las fincas, los cambios continuos en la relación de los agricultores con la finca operada, el vencimiento/inicio o la renovación de los arrendamientos, el problema de obtener una lista completa de las operaciones agrícolas, la dificultad de contactar e identificar algunos tipos de relaciones entre contratante / contratista, la ausencia del agricultor de la finca durante el período de recopilación de datos y la opinión del agricultor de que parte o la totalidad de la operación no califica y no debe ser incluida en el censo. Durante la recopilación de datos y el procesamiento del

should not be included in the census. During data collection and processing of the census, all operations underwent several quality control checks to ensure results were as accurate as possible.

### **Item Nonresponse**

All item nonresponse actions provide another opportunity to introduce measurement errors. Regardless of whether it was previously reported data, administrative data, the nearest neighbor algorithm, the fully conditional specification method, or manually imputed by an analyst, some risk exists that the imputed value does not equal the actual value. Previously reported and administrative data were used only when they related to the census reference period. A new nearest neighbor was randomly selected for each incident to eliminate the chance of a consistent bias.

### **Record Matching Error**

The process of building and expanding the CML involves finding new list sources and checking for names not on the list. An automated processing system compared each new name to the existing CML names and “linked” like records for the purpose of preventing duplication. New names with strong links to a CML name were discarded and those with no links were added as potential farms. Names with weak links, possible matches, were reviewed by staff to determine whether the new name should be added. Despite this thorough review, some new names may have been erroneously added or deleted. Additions could contribute to duplication (overcoverage) whereas deletions could contribute to undercoverage. As a result, some names received more than one report form, and some farm operators did not receive a report form.

Another opportunity for error came when comparing Agricultural Coverage Evaluation Survey (ACES) tract operator names to the CML. Area operators whose names were not found on the CML were part of the measure of list incompleteness, or NML. Mistakes in determining overlap status resulted in overcounts (including a tract whose operator was on the CML) or undercounts (excluding a tract whose operator was not on the CML). All tracts determined to not be on the list were checked to eliminate, or at least minimize, any error. In order to attempt to identify duplication, all respondents who received multiple

censo, todas las fincas se sometieron a una serie de controles de calidad para garantizar que los resultados fueran lo más precisos posible.

### **Partidas Sin Respuesta**

Todas las acciones de no respuesta a nivel de partidas brindan otra oportunidad para introducir errores de medición. Independientemente de si fueron datos previamente informados, datos administrativos, el algoritmo vecino más cercano, el método de especificación totalmente condicional, o imputado manualmente por un analista, existe algún riesgo de que el valor imputado no sea igual al valor real. Los datos administrativos y de informes anteriores solo se utilizaron cuando se relacionaban con el período de referencia del censo. Se seleccionó al azar un nuevo vecino más cercano para cada incidente para eliminando así la posibilidad de un sesgo constante.

### **Error de Fusión o Enlace de Reportes**

El proceso de construir y expandir la CML implica encontrar nuevas listas y buscar nombres que no estén en la lista de NASS. Un sistema de procesamiento automatizado comparó cada nuevo nombre con los nombres existentes en la CML y los registros similares “vinculados” con el fin de evitar la duplicación. Se descartaron nuevos nombres con enlaces fuertes a un nombre en la CML y aquellos sin enlaces se agregaron como fincas potenciales. Los nombres con enlaces débiles, posibles pareos, fueron revisados por el personal para determinar si se debe agregar el nuevo nombre a la CML. A pesar de esta revisión exhaustiva, algunos nombres nuevos pueden haberse agregado o eliminado por error. Las inclusiones podrían contribuir a la duplicación (sobrecobertura) mientras que las eliminaciones podrían contribuir a la subcobertura. Como resultado, algunos nombres recibieron más de un formulario, y algunos operadores de fincas no recibieron un formulario.

Otra oportunidad de error surgió al comparar los nombres de los agricultores de tractos agrícolas en la Encuesta de Evaluación de la Cobertura Agrícola (ACES) con la CML. Los agricultores en la muestra de área cuyos nombres no se encontraron en la CML formaban parte de la medida de cuan incompleta es la lista o NML. Los errores en la determinación del estado de superposición dieron como resultado un conteo excesivo (incluido un tracto cuyo operador estaba en la CML) o un conteo bajo (excluyendo un tracto cuyo operador no estaba en la CML). Se verificaron todos los tractos que se determinó que no estaban en la lista para eliminar, o al menos minimizar, cualquier error. Con el fin de intentar identificar la duplicación, se ordenó a todos los encuestados que recibieron múltiples formularios censales que completaran

report forms were instructed to complete one form and return all other forms so duplication could be removed.

Records in the 2022 ACES were matched to the 2022 census using probabilistic record linkage. The uncertainty associated with this estimate, with the exception of model uncertainty, was accounted for, but errors not found through this process were not.

un solo formulario y devolvieran todos los demás formularios para poder eliminar la duplicación.

Los registros en el ACES de 2022 se combinaron con el censo de 2022 utilizando el enlace de registros probabilísticos. La incertidumbre asociada con esta estimación, con la excepción de la incertidumbre del modelo, se tuvo en cuenta, pero los errores no encontrados a través de este proceso no lo fueron.

**Table A. Summary of Puerto Rico Coverage, Nonresponse, and Misclassification Adjustments: 2022**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
Farms ..... number	7,602	616	50.7	24.8	8.6	17.3
Land in farms ..... cuerdas	494,481	37,426	46.0	25.0	9.3	11.7
Farms by size:						
Less than 10 cuerdas ..... number	2,258	210	42.4	20.0	7.3	15.0
..... cuerdas	10,345	803	44.9	20.3	7.2	17.4
10 to 19 cuerdas ..... number	1,741	150	51.4	25.8	7.7	18.0
..... cuerdas	23,601	2,079	51.5	26.7	7.7	17.2
20 to 49 cuerdas ..... number	1,676	210	55.7	26.3	7.3	22.1
..... cuerdas	51,750	7,344	56.2	26.4	7.9	21.9
50 to 99 cuerdas ..... number	762	56	57.3	26.0	12.9	18.3
..... cuerdas	52,541	3,601	56.7	26.2	12.8	17.7
100 to 174 cuerdas ..... number	500	18	52.9	33.6	7.6	11.8
..... cuerdas	65,154	2,191	52.8	33.6	7.6	11.5
175 to 259 cuerdas ..... number	268	49	50.8	24.0	14.9	11.9
..... cuerdas	55,977	11,248	50.6	24.1	15.0	11.4
260 cuerdas or more ..... number	397	63	42.5	23.1	9.0	10.5
..... cuerdas	235,113	20,560	37.4	21.4	7.8	8.2
Irrigation:						
Land irrigated ..... farms	1,229	76	47.6	20.2	8.8	18.6
..... cuerdas	44,893	3,585	37.8	22.3	9.9	5.6
..... water (acre-feet)	25,017	1,584	45.9	23.6	12.9	9.4
Public system ..... farms	387	26	51.0	16.0	6.1	28.9
..... cuerdas	12,410	2,186	44.0	30.2	7.5	6.3
Private system ..... farms	899	55	46.2	23.0	10.4	12.8
..... cuerdas	32,484	2,506	35.4	19.7	10.4	5.3
Market value of agricultural products sold ..... dollars	703,255,498	28,181,328	29.4	19.9	6.4	3.1
Farms by value of sales:						
Less than \$1,000 ..... farms	1,474	180	48.1	41.7	2.3	4.1
..... dollars	238,963	31,842	33.5	22.2	6.0	5.3
\$1,000 to \$2,499 ..... farms	690	88	41.1	27.1	11.1	2.9
..... dollars	1,142,284	140,159	41.3	27.1	11.3	3.0
\$2,500 to \$4,999 ..... farms	810	64	46.2	24.6	12.8	8.8
..... dollars	2,907,427	211,920	46.2	23.8	12.4	10.0
\$5,000 to \$7,499 ..... farms	589	81	56.7	12.4	5.1	39.2
..... dollars	3,552,300	485,822	56.7	12.5	5.2	38.9
\$7,500 to \$9,999 ..... farms	466	81	57.5	10.5	5.9	41.1
..... dollars	4,003,489	706,783	57.5	10.4	6.1	41.0
\$10,000 to \$19,999 ..... farms	1,232	127	62.6	15.6	6.0	40.9
..... dollars	17,640,040	1,799,182	62.5	15.3	6.3	40.8
\$20,000 to \$24,999 ..... farms	247	38	48.5	29.0	13.8	5.7
..... dollars	5,400,768	860,866	48.2	28.7	13.8	5.7
\$25,000 to \$29,999 ..... farms	216	25	49.2	21.7	23.5	4.0
..... dollars	5,891,741	681,906	49.4	22.0	23.3	4.0
\$30,000 to \$39,999 ..... farms	298	28	50.5	27.6	18.3	4.6
..... dollars	10,244,881	953,982	50.5	27.2	18.8	4.4
\$40,000 to \$49,999 ..... farms	166	21	50.0	32.8	11.5	5.7
..... dollars	7,377,361	947,712	49.9	32.9	11.4	5.7
\$50,000 to 59,999 ..... farms	153	14	50.4	30.3	14.2	6.0
..... dollars	8,313,125	765,222	50.5	30.0	14.5	6.0
\$60,000 or more ..... farms	1,261	99	44.1	25.2	14.2	4.7
..... dollars	636,543,119	24,326,165	26.8	19.0	5.9	2.0
Farms by operation's legal status for tax purposes:						
Family or individual ..... farms	5,835	496	50.8	25.1	8.2	17.5
..... cuerdas	270,538	21,185	49.3	27.2	7.6	14.5
Partnership ..... farms	188	16	52.9	25.2	12.9	14.8
..... cuerdas	19,769	1,383	41.2	24.1	12.4	4.8
Corporation ..... farms	1,399	106	50.0	24.2	10.0	15.7
..... cuerdas	183,549	14,445	42.7	22.3	10.7	9.7
Other ..... farms	180	20	46.7	18.8	5.1	22.8
..... cuerdas	20,625	2,903	49.0	29.5	9.2	10.3
Tenure:						
Full owners ..... number	5,166	427	51.5	25.9	8.1	17.5
..... cuerdas	214,004	14,256	49.6	25.6	8.9	15.0
Part owners ..... number	855	76	50.4	23.2	10.6	16.7
..... cuerdas	131,334	10,071	40.2	23.5	9.6	7.1
Tenants ..... farms	1,581	127	50.4	23.4	9.3	17.8
..... cuerdas	149,144	14,864	47.5	25.9	9.7	11.9
Principal operator characteristics by -						
Sex of operators:						
Male ..... operators	6,658	531	50.6	24.4	8.8	17.4
Female ..... operators	944	89	49.9	27.0	6.9	16.0
Primary occupation:						
Agriculture ..... operators	4,562	343	50.1	23.1	9.7	17.3
Nonagriculture ..... operators	3,040	278	51.1	27.0	6.9	17.2
Hispanic, Latino or Spanish origin ..... operators	7,517	609	50.6	24.8	8.5	17.4
Not of Hispanic, Latino or Spanish origin ..... operators	85	14	41.0	23.4	9.8	7.8
Race:						
Black or African American ..... operators	879	63	50.8	22.9	7.7	20.2
White ..... operators	6,455	533	50.6	24.9	8.7	17.0
Other ..... operators	44	9	50.0	35.1	10.6	4.3
More than one race reported ..... operators	224	30	48.0	29.3	6.2	12.6
Military Service:						
Never served or only on active duty for training in the Reserves or National Guard ..... operators	7,201	584	50.5	24.7	8.5	17.3
Active duty now or in the past ..... operators	401	38	52.5	26.7	8.8	17.0
Age group:						
Under 25 years ..... operators	29	8	49.6	22.3	19.0	8.4
25 to 34 years ..... operators	332	33	47.5	18.7	12.0	16.8

--continued

**Table A. Summary of Puerto Rico Coverage, Nonresponse, and Misclassification Adjustments: 2022 (continued)**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
Age group: - Con.						
35 to 44 years .....operators	820	61	48.3	22.5	10.1	15.7
45 to 54 years .....operators	1,318	115	50.4	24.9	9.0	16.4
55 to 64 years .....operators	1,930	163	51.2	25.8	8.7	16.7
65 to 74 years .....operators	1,932	154	51.0	24.5	7.2	19.3
75 years and over .....operators	1,241	123	51.7	27.0	7.3	17.4
Livestock and poultry inventory:						
Cattle and calves .....farms	1,821	118	51.0	22.8	11.7	16.5
.....number	202,621	11,628	32.3	21.4	6.1	4.7
Hogs and pigs .....farms	273	22	47.6	14.4	20.8	12.5
.....number	36,069	3,631	44.3	8.5	32.7	3.1
Horses .....farms	468	39	75.1	35.1	22.2	17.9
.....number	6,798	186	77.6	26.8	43.7	7.2
Sheep .....farms	375	34	66.8	21.1	24.9	20.8
.....number	16,604	2,115	73.2	18.1	39.9	15.2
Goats .....farms	275	18	67.3	20.4	18.0	28.9
.....number	3,446	292	71.3	20.2	27.0	24.1
Rabbits .....farms	142	13	67.7	23.8	13.6	30.4
.....number	28,203	4,387	76.2	14.4	34.7	27.0
Laying hens .....farms	284	32	51.0	14.4	14.7	21.8
.....number	424,268	122,760	12.6	9.0	2.7	1.0
Broilers and other chickens for meat production .....farms	73	7	35.4	10.7	11.4	13.3
.....number	3,946,440	528,260	40.3	33.1	5.2	2.0
Crops Harvested:						
Coffee grown in the shade .....farms	1,671	137	48.0	22.8	6.8	18.4
.....cuerdas	8,294	798	50.0	25.5	6.3	18.2
Coffee grown in the open .....farms	1,746	145	48.2	26.6	4.5	17.1
.....cuerdas	10,667	714	48.0	25.6	6.3	16.0
Pineapples .....farms	208	17	51.7	23.4	9.9	18.4
.....cuerdas	1,086	252	62.6	42.0	14.4	6.2
Plantains .....farms	2,741	207	51.1	21.9	7.1	22.0
.....cuerdas	15,874	1,425	48.9	30.1	6.3	12.4
Bananas .....farms	1,492	125	51.5	20.9	7.7	22.9
.....cuerdas	5,314	505	41.0	27.1	4.9	9.0



**Table B. Reliability Estimates of Puerto Rico Totals: 2022**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Coefficient of variation (percent)
Farms .....	7,602	8.1
Land in farms .....	494,481	7.6
Farms by size:		
Less than 10 cuerdas .....	2,258	9.3
..... cuerdas	10,345	7.8
10 to 19 cuerdas .....	1,741	8.6
..... cuerdas	23,601	8.8
20 to 49 cuerdas .....	1,676	12.5
..... cuerdas	51,750	14.2
50 to 99 cuerdas .....	762	7.3
..... cuerdas	52,541	6.9
100 to 174 cuerdas .....	500	3.7
..... cuerdas	65,154	3.4
175 to 259 cuerdas .....	268	18.2
..... cuerdas	55,977	20.1
260 cuerdas or more .....	397	15.8
..... cuerdas	235,113	8.7
Irrigation:		
Land irrigated .....	1,229	6.2
..... farms	44,893	8.0
..... cuerdas	25,017	6.3
..... water (acre-feet)	387	6.8
Public system .....	12,410	17.6
..... farms	899	6.1
..... cuerdas	32,484	7.7
Private system .....		
Market value of agricultural products sold .....		
..... dollars	703,255,498	4.0
Farms by value of sales:		
Less than \$1,000 .....	1,474	12.2
..... farms	238,963	13.3
..... dollars	690	12.7
\$1,000 to \$2,499 .....	1,142,284	12.3
..... farms	810	7.9
..... dollars	2,907,427	7.3
\$2,500 to \$4,999 .....	589	13.7
..... farms	3,552,300	13.7
..... dollars	466	17.3
\$5,000 to \$7,499 .....	4,003,489	17.7
..... farms	1,232	10.3
..... dollars	17,640,040	10.2
\$7,500 to \$9,999 .....	247	15.5
..... farms	5,400,768	15.9
..... dollars	216	11.4
\$10,000 to \$19,999 .....	5,891,741	11.6
..... farms	298	9.5
..... dollars	10,244,881	9.3
\$20,000 to \$24,999 .....	166	12.5
..... farms	7,377,361	12.8
..... dollars	153	9.3
\$25,000 to \$29,999 .....	8,313,125	9.2
..... farms	1,261	7.8
..... dollars	636,543,119	3.8
..... dollars		
Farms by operation's legal status for tax purposes:		
Family or individual .....	5,835	8.5
..... farms	270,538	7.8
..... cuerdas	188	8.7
Partnership .....	19,769	7.0
..... farms	1,399	7.6
..... cuerdas	183,549	7.9
Corporation .....	180	10.8
..... farms	20,625	14.1
..... cuerdas		
Other .....		
Tenure:		
Full owners .....	5,166	8.3
..... number	214,004	6.7
..... cuerdas	855	8.9
Part owners .....	131,334	7.7
..... number	1,581	8.0
..... cuerdas	149,144	10.0
..... farms		
..... cuerdas		
Principal operator characteristics by -		
Sex of operators:		
Male .....	6,658	8.0
..... operators	944	9.5
..... operators		
Primary occupation:		
Agriculture .....	4,562	7.5
..... operators	3,040	9.1
Nonagriculture .....		
..... operators		
Hispanic, Latino or Spanish origin .....	7,517	8.1
..... operators	85	17.0
Not of Hispanic, Latino or Spanish origin .....		
..... operators		
Race:		
Black or African American .....	879	7.1
..... operators	6,455	8.3
White .....	44	20.5
..... operators	224	13.6
Other .....		
More than one race reported .....		
..... operators		
Military Service:		
Never served or only on active duty for training		
in the Reserves or National Guard .....	7,201	8.1
..... operators	401	9.4
Active duty now or in the past .....		
..... operators		
Age group:		
Under 25 years .....	29	26.4
..... operators	332	10.0
25 to 34 years .....		
..... operators		

--continued



**Table C. Summary of Coverage, Nonresponse, and Misclassification Adjustments by Region: 2022**

[For meaning of abbreviations and symbols, see introductory text.]

Geographic area	Total (number)	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
<b>ALL FARMS (NUMBER)</b>						
<b>Puerto Rico Total</b>						
Puerto Rico .....	7,602	616	50.7	24.8	8.6	17.3
<b>Regions</b>						
Region 1 - Utuado .....	1,533	154	49.0	26.2	6.1	16.7
Region 2 - Arecibo .....	927	78	51.7	23.1	11.9	16.7
Region 3 - Mayagüez .....	590	57	51.1	23.2	12.0	15.9
Region 4 - Ponce .....	659	57	52.2	24.3	10.8	17.0
Region 5 - Caguas .....	1,042	70	53.1	28.4	8.5	16.3
Region 6 - Naranjito .....	738	81	53.6	25.8	9.1	18.6
Region 7 - Lares .....	1,299	111	47.7	24.7	5.6	17.4
Region 8 - San Germán .....	814	53	51.0	22.0	8.4	20.6
<b>LAND IN FARMS (CUERDAS)</b>						
<b>Puerto Rico Total</b>						
Puerto Rico .....	494,481	37,426	46.0	25.0	9.3	11.7
<b>Regions</b>						
Region 1 - Utuado .....	44,932	4,573	57.0	32.5	7.9	16.7
Region 2 - Arecibo .....	86,816	6,333	40.5	18.0	8.0	14.5
Region 3 - Mayagüez .....	37,204	4,005	36.4	19.1	8.1	9.2
Region 4 - Ponce .....	56,784	4,721	42.9	24.9	11.2	6.9
Region 5 - Caguas .....	98,353	6,968	48.5	33.2	7.9	7.4
Region 6 - Naranjito .....	32,265	3,616	54.0	29.7	9.1	15.2
Region 7 - Lares .....	41,902	3,996	48.8	30.1	5.9	12.8
Region 8 - San Germán .....	96,225	6,342	48.1	23.2	12.2	12.6
<b>SALES (DOLLARS)</b>						
<b>Puerto Rico Total</b>						
Puerto Rico .....	703,255,498	28,181,328	29.4	19.9	6.4	3.1
<b>Regions</b>						
Region 1 - Utuado .....	28,086,397	4,556,863	42.6	22.0	8.9	11.7
Region 2 - Arecibo .....	149,946,480	6,839,508	22.2	15.5	4.8	1.9
Region 3 - Mayagüez .....	40,723,528	3,624,008	26.9	17.9	5.3	3.7
Region 4 - Ponce .....	156,955,940	4,349,174	15.9	10.9	4.1	0.9
Region 5 - Caguas .....	117,859,011	13,066,467	42.6	30.8	7.9	4.0
Region 6 - Naranjito .....	93,078,342	9,604,205	35.7	26.1	5.9	3.7
Region 7 - Lares .....	49,647,630	6,255,296	40.7	24.5	8.0	8.2
Region 8 - San Germán .....	66,958,170	2,759,995	35.1	19.5	8.6	7.0