



Centro Centroamericano de Población
Universidad de Costa Rica



Manual del usuario

Sistema de consultas a bases de datos estadísticas

Última versión
- Abril, 2015 -

TABLA DE CONTENIDO

PRESENTACIÓN	3
1. ACCESO AL SISTEMA DE CONSULTAS	4
1.1. Registrarse como nuevo usuario	5
1.2. Recuperar contraseña	6
1.3. Bases de datos disponibles	7
2. FORMULACIÓN DE CONSULTAS EN LÍNEA	10
2.1. Asistente para la definición de tablas	10
2.2. Asistente para la construcción de expresiones	14
2.3. Uso de operadores lógicos y aritméticos.....	15
2.4. Delimitar rangos	18
2.5. Diccionario de datos, cuestionarios y boletas	18
2.6. Cambiar bases de datos	19
2.7. Recodificación de variables	19
2.8. Recodificación por rangos	21
3. GENERACIÓN DE RESULTADOS	22
4. PROCESAMIENTO JERÁRQUICO	24
4.1. Tabulados que implican variables de diferentes jerarquias	24
4.2. Uso de la función de ponderación	25
4.3. Uso de funciones jerárquicas.....	26
5. EJEMPLOS PRACTICOS	28
Número de habitantes por cantón, según sexo	28
Número de habitantes por año, según sexo	30
Número promedio de hijos entre las mujeres, según zona de residencia.....	31
Índice de masculinidad o de relación entre sexos	33
Número de habitantes, por edad y sexo	34
Porcentaje de personas nacidas en el extranjero, por grupos de edad	35
Relación de dependencia demográfica.....	37
Número promedio de personas por vivienda independiente	39
Tasa bruta de mortalidad en el periodo 2000-2010	41
Tasa de mortalidad infantil en los años 2005 y 2010.....	42
Tasa de fecundidad por grupos quinquenales de edad, en 2010	44
Tasa global de fecundidad, en el año 2010	46

PRESENTACIÓN

El Centro Centroamericano de Población (CCP) de la Universidad de Costa Rica tiene a disposición de la comunidad un servicio gratuito para la consulta de las bases de datos de los Censos de Población y Vivienda, archivos de Nacimientos y Defunciones, Proyecciones de Población, entre otras. Bajo una interfaz amigable con el usuario, el *Sistema de Consulta a Bases de Datos Estadísticas* permite formular consultas en tiempo real y obtener los resultados en unos pocos segundos, en lugar de descargar tablas y gráficos prediseñados. El siguiente manual ha sido diseñado con el fin de ayudar al usuario a utilizar esta herramienta.

El acceso al sitio es completamente libre y gratuito, aunque para ingresar es necesario que el usuario introduzca un código de usuario y palabra clave. Estos datos los podrá obtener al registrarse en el sitio web. Para ingresar es necesario contar con un computador con conexión a internet y que cuente con un navegador web como Internet Explorer (versión 8 o superior), Mozilla Firefox (versión 9 o superior) o Google Chrome (versión 19 o superior).

Este manual pretende guiar al usuario de una manera amena, mostrando ejemplos gráficos de los procedimientos citados y brindando consejos o advertencias. De esta manera, los usuarios podrán aprovechar de una manera eficiente esta herramienta y hacer uso adecuado de los datos.

La primera sección de este manual describe los elementos más esenciales para acceder al Sistema de Consultas: página de inicio, cómo registrarse y cómo recuperar la contraseña. Esta sección incluye también una breve descripción de las bases de datos disponibles para Costa Rica. Para quienes ingresan por primera, esta sección puede ser de mucha utilidad.

El segundo capítulo muestra a los usuarios cómo establecer sus consultas y generar tabulaciones de una, dos o tres variables y además seleccionar un universo específico de casos. Se describen las diversas secciones de las que está compuesta la interfaz y que pueden ser utilizadas para la definición de la consulta. El usuario aprenderá como seleccionar las variables, a escribir expresiones lógicas y aritméticas para establecer filtros, así como a recodificar variables. En el tercer capítulo se describe el contenido de la ventana de resultados y en el cuarto se brindan instrucciones para realizar procesamientos jerárquicos.

En el último apartado el usuario encontrará ejercicios prácticos, que se describen y desarrollan paso a paso. Aquí se exponen algunos de los principales indicadores y algunas medidas demográficas muy utilizadas, y cuya estimación requiere la información de las bases de datos que se encuentran disponibles en el Sistema de Consultas.

Finalmente, es importante señalar que el Sistema de Consulta a bases de datos estadísticas ha sido creado para proteger la confidencialidad de los individuos o entrevistados. Ninguna de las bases de datos contiene información personal, o que permita identificar a los entrevistados o declarantes de la información que se procesa. Los archivos han sido creados únicamente para el análisis agregado del dato estadístico.

1. ACCESO AL SISTEMA DE CONSULTAS

Para acceder al sistema de consultas deberá ingresar al sitio web del Centro Centroamericano de Población: <http://ccp.ucr.ac.cr/> y dirigirse a “Censos y otras bases” en la opción “Datos en línea”, del menú “Contenidos” en la sección izquierda de la pantalla.

Figura 1. Página principal

The screenshot shows the main page of the Centro Centroamericano de Población website. The browser address bar displays www.ccp.ucr.ac.cr. The page header includes the logo of the Centro Centroamericano de Población (CCP) and the Universidad de Costa Rica, along with a 75th anniversary logo. A navigation menu at the top lists: Aula Virtual, Censos, InfoCensos, Tasas demográficas, CRELES, Biblioteca, Revista, LAPOP, Farmacoeconomía, and CNT/NTA. On the left, a 'Contenido' sidebar menu is open, with 'Datos en línea' selected, and a sub-menu showing 'Censos y otras bases' highlighted. The main content area features a 'Bienvenido' message and a photograph of the CCP building. A 'Novedades' section on the right lists two news items: 'Nueva ubicación del stand del CCP en la EXPO UCR' and 'Personajes Planificación Familiar en Costa Rica'.

Del mismo modo, al dirigirse a la dirección: <http://ccp.ucr.ac.cr/censos/> es posible acceder directamente a la página de inicio del Sistema de Consultas en línea (Ver Figura 2).

Figura 2. Ingresar al Sistema de Consultas

The screenshot displays the website's header with the logo and title "Centro Centroamericano de Población Sistema de Consulta a Bases de Datos Estadísticas". A navigation menu includes "Inicio", "Datos en Línea", "Información Demográfica", "Encuestas", "Actividades", "Publicaciones", "Biblioteca", and "Proyectos".

Enlaces

- > Datos en línea
- > Tasas demográficas
- > Información Demográfica
- > Encuestas
- > Proyectos
- > Publicaciones
- > Actividades
- > Bibliotecas
- Personal
- Correo Electrónico
- Enlaces
- Contáctenos
- Acerca de Nosotros
- Mapa del Sitio

En este sitio se pueden realizar consultas a grandes bases de datos estadísticos y obtener los resultados de forma casi inmediata. Estas bases de datos son archivos de información original de los individuos de censos, registros vitales y otros similares. Aquí es posible elaborar consultas con 1, 2 o 3 variables, una selección de casos y aplicar opciones de recodificación o procesamiento jerárquico. Para ingresar al sitio debe introducir su código de usuario y palabra clave o registrarse de manera libre y gratuita para obtenerlos.

Ingresar al Sitio

Código Usuario:

Contraseña:

[Recuperar contraseña](#)

Registrar Usuario

¿Nuevo en el sitio?

Novedades

Enero del 2015

Proyecciones Distritales de Población de Costa Rica 2000-2050

1 2 3

Posteriormente, deberá ingresar su código de usuario y contraseña en la sección correspondiente, a la derecha de la pantalla. Los usuarios que ingresan por primera vez deberán solicitar un código de usuario y contraseña completando la información del formulario que se despliega al presionar el botón "Registrar usuario" (ver Figura 2).

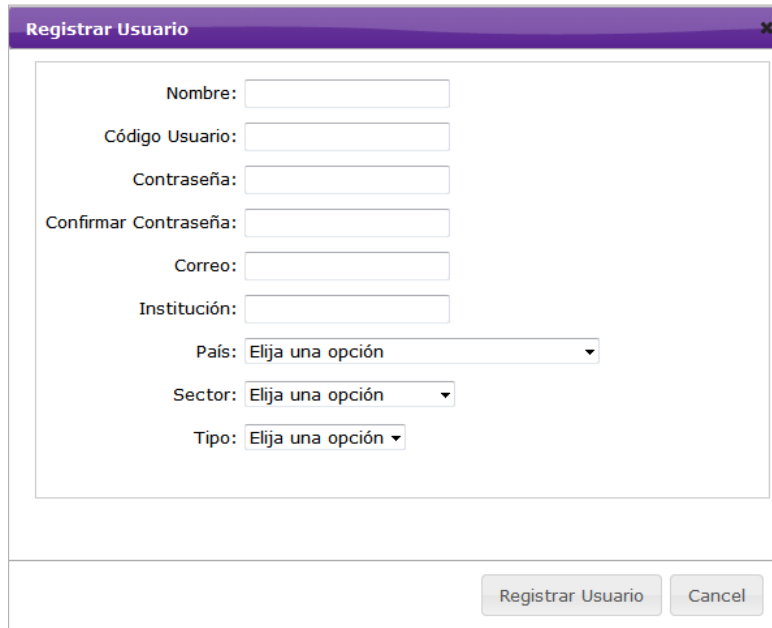
1.1. REGISTRARSE COMO NUEVO USUARIO

El formulario de inscripción requiere el ingreso de información básica como nombre y correo electrónico, además de registrar un código de usuario y contraseña (ver Figura 3).

Si el código de usuario ingresado ya existe, el sistema mostrará un mensaje indicando que el código ya está siendo utilizado por otra persona, por lo que deberá utilizar otro. No se han establecido restricciones para establecer la contraseña, aunque es recomendable que ésta contenga al menos 4 caracteres.

Actualmente el sistema no envía una confirmación de registro, por lo que una vez que complete exitosamente el formulario, podrá acceder automáticamente con el código y contraseña que registró. Es recomendable que utilice un código de usuario y contraseña que pueda recordar fácilmente o que anote estos datos, para que los pueda recuperar si los olvida.

Figura 3. Registrarse como nuevo usuario



Registrar Usuario

Nombre:

Código Usuario:

Contraseña:

Confirmar Contraseña:

Correo:

Institución:

País:

Sector:

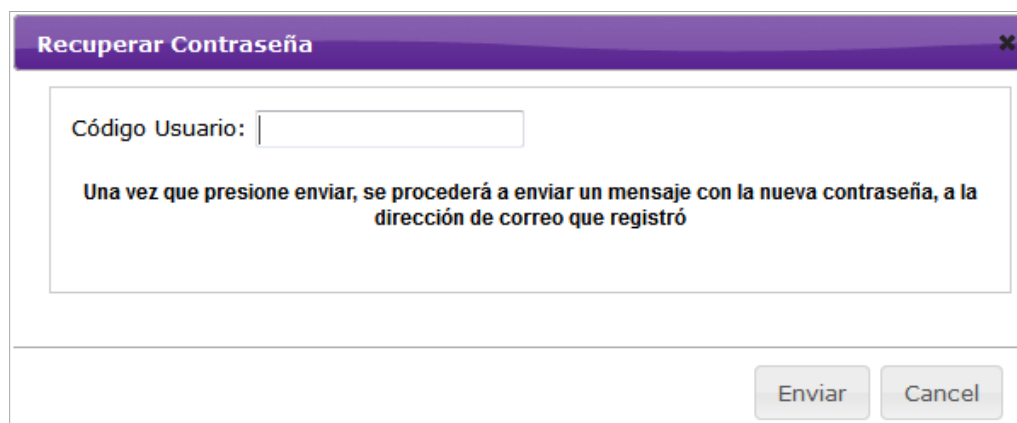
Tipo:

Registrar Usuario Cancel

1.2. RECUPERAR CONTRASEÑA

Si ya posee una cuenta, pero ha olvidado la contraseña, deberá dirigirse a la opción “*Recuperar contraseña*” que se encuentra en la página de ingreso al Sistema de Consultas (ver Figura 2). De manera automática, el sistema enviará un correo a la dirección electrónica que registró, brindándole una nueva contraseña, con la cual podrá ingresar nuevamente.

Figura 4. Recuperar contraseña



Recuperar Contraseña

Código Usuario:

Una vez que presione enviar, se procederá a enviar un mensaje con la nueva contraseña, a la dirección de correo que registró

Enviar Cancel

1.3. BASES DE DATOS DISPONIBLES

Después de ingresar el código de usuario y contraseña, y acceder al Sistema de Consultas, se listarán en pantalla las bases de datos disponibles (Ver Figura 5). Para iniciar la consulta, el usuario debe seleccionar una base de datos y presionar el botón “Procesa”, que se encuentra al fondo de la pantalla. Automáticamente será redireccionado al *Asistente para la definición de tablas*, mediante el cual se definen los criterios para generar tabulados. Esta herramienta se presentará con detalle en la siguiente sección.

Figura 5. Selección de la base de datos

Información

- Instrucciones de Uso
- Recomendaciones Técnicas

Enlaces

- > Datos en línea
- > Tasas demográficas
- > Información Demográfica
- > Encuestas
- > Proyectos
- > Publicaciones
- > Actividades
- > Bibliotecas

Personal

- Correo Electrónico

Enlaces

- Contactémos
- Acerca de Nosotros
- Mapa del Sitio

Selección de la base de datos

Costa Rica **Centroamérica y México**

Otras Bases de Datos Encuestas de Hogares Clusters

Costa Rica - Instituto Nacional de Estadística y Censos

- Proyecciones Distritales de Poblacion de Costa Rica 2000-2050
- Proyecciones Nacionales de Población
- Defunciones 1970 - 2013
- Nacimientos 1972 - 2013
- Censos de Población y Vivienda 2011
- Censos de Población y Vivienda 2000
- Censos de Población y Vivienda 1984
- Censos de Población y Vivienda 1973
- Muestra Censo de Población y Vivienda 1963

Caja Costarricense del Seguro Social

- Egresos Hospitalarios 1990-2003

Ministerio de Salud

- Registro Nacional de Tumores 1981-2005

Tribunal Supremo de Elecciones y Registro Civil

- Padrón electoral 2006
- Padrón electoral 2002
- Padrón electoral 1998
- Padrón electoral 1994

Centro de Investigaciones Históricas de América Central

- Censo de Población de 1927

Procesar

La pantalla “Selección de la base de datos” le mostrará la posibilidad de elegir alguna de las bases de datos disponibles. Al desplazarse por las pestañas correspondientes, ubicadas en la sección superior del

recuadro, podrá seleccionar entre alguna de las bases de datos disponibles para Costa Rica o para Centroamérica y México.

Para este último grupo, podrá disponer de los más recientes Censos de Población y Vivienda realizados en la región. También encontrará las bases de datos de algunas encuestas de Demografía y Salud. Éstas pueden ser analizadas al descargar directamente la base de datos (disponibles en ASCII, SPSS y Stata), por lo que no necesitará utilizar el Sistema de Consultas para analizarlas.

Para Costa Rica, podrá disponer de las bases de datos que se describen a continuación:

▪ **Censos de Población y Vivienda:**

Actualmente se encuentran disponibles los Censos de Población y Vivienda, efectuados por el Instituto Nacional de Estadística y Censos, desde el año 1973, incluyendo el más reciente Censo de Población y Vivienda realizado en el año 2011. La base de datos del Censo de Población y Vivienda del año 1963, corresponde a una muestra del 5% generada de manera aleatoria sobre del total de viviendas enumeradas en dicho Censo. Por tratarse de una muestra, el usuario deberá tener cuidado al analizar información en niveles desagregados o grupos pequeños.

Cada uno de los Censos de Población registra información para tres niveles o jerarquías: Vivienda, Hogar y Persona. El usuario deberá tener cautela al analizar las bases de datos censales, pues los resultados mostrarán el número de casos de la jerarquía de menor rango involucrada en la consulta. El manejo de la información entre los distintos niveles, se abordará con mayor detalle en la sección *Análisis Jerárquico*.

▪ **Registro de Nacimientos:**

Esta base de datos contiene información sobre los nacimientos registrados en Costa Rica desde 1972 y es actualizada anualmente, para sumar los nuevos casos que se van registrando año con año. Incluye variables que contemplan información sobre los recién nacidos, como el peso y la talla, y sobre sus padres como edad, nacionalidad y estado civil.

▪ **Registro de Defunciones:**

Esta base de datos contempla información sobre las defunciones registradas en el territorio nacional desde el año 1970. Es actualizada anualmente para añadir los nuevos casos de cada año.

Una de las variables más importantes registradas en esta base de datos es la causa de muerte. Para su codificación se usa la *Clasificación Internacional de las Enfermedades* (CIE), en sus últimas tres versiones. Se utilizó la octava versión CIE-8 para los casos registrados en el periodo 1970-1979 mientras que la clasificación CIE-9 se utilizó entre 1980 y 1996. La clasificación más reciente, CIE-10, se utiliza desde 1997 hasta el presente. Puede encontrar mayor información sobre la clasificación CIE en el sitio web de la Organización Mundial de la Salud: <http://www.who.int/classifications/icd/en/>.

▪ **Proyecciones Nacionales y Distritales de Población:**

Generalmente, estas bases de datos son actualizadas a partir de los resultados de cada Censo de Población. La más reciente actualización fue realizada por el CCP y el Instituto Nacional de Estadística y Censos (INEC), a partir del X Censo de Población y VI de Vivienda, del año 2011, posterior a la evaluación de la omisión censal y cálculo de una población base.

Con la base de datos de **Proyecciones Nacionales de Población**, es posible contabilizar el número de habitantes de Costa Rica (a nivel país) desde el año 1950 hasta el 2100, con la posibilidad de obtener cifras para cada año, por edad simple o quinquenal y sexo.

La base de datos de las **Proyecciones Distritales de Población** permite conocer el número estimado de habitantes a nivel distrital, y también para otras desagregaciones geográficas: cantonal, regional y según GAM. Las cifras pueden obtenerse por sexo, edad en grupos quinquenales y para el periodo comprendido entre el año 2000 y el 2050. El usuario podría encontrar algunas variaciones, si compara el número total de habitantes del país -obtenido a partir de la suma de las estimaciones de esta base de datos- con cifras obtenidas a partir de las Proyecciones Nacionales de Población. Estas diferencias se deben a efectos de redondeo.

▪ **Encuestas de hogares:**

Las Encuestas de Hogares, Encuestas de Hogares y de Propósitos Múltiples (EHPM: 1976-2009) o Encuesta Nacional de Hogares (ENAHG: 2010 al presente) se realizan de manera anual, generalmente en el mes de julio. A pesar de que incluyen algunos de los temas y variables que se analizan en los censos de población, estas encuestas se realizan con mayor periodicidad e incluyen información mucho más extensa, sin embargo no permiten realizar estimaciones para niveles geográficos menores (cantones o distritos) o grupos poblacionales pequeños (migrantes o extranjeros, entre otros). Contienen información sobre los hogares y sus miembros (edad, sexo, nivel educativo).

A diferencia de las bases de datos descritas anteriormente, que son procesadas mediante el sistema *Public Data Queries* (PDQ), las Encuestas de Hogares se procesan mediante el paquete estadístico Stata. En este caso, su análisis se desarrolla de manera muy similar a las otras bases de datos aunque el usuario encontrará algunas diferencias al interactuar con la herramienta.

2. FORMULACIÓN DE CONSULTAS EN LÍNEA

Esta sección, describe la forma en la que se deben establecer los criterios necesarios para generar tabulados. Se discutirá sobre cómo seleccionar las variables que desee analizar, cómo establecer expresiones lógicas o aritméticas, cómo recodificar variables, entre otros aspectos.

Los ejemplos que se desarrollarán para ilustrar el uso de las funciones, se realizarán con la base de datos del *X Censo de Población y VI de Vivienda del año 2011*. Las operaciones realizadas con esta base de datos pueden extenderse a otras, de manera análoga. Además, en la tercera sección de este informe se desarrollan algunos ejemplos con datos de dichas bases de datos.

2.1. ASISTENTE PARA LA DEFINICIÓN DE TABLAS

Después de seleccionar la base de datos, aparecerá en la pantalla un recuadro como el que se muestra en la Figura 6. Mediante el *Asistente para la definición de la tabla* podrá elaborar consultas con hasta tres variables, seleccionar casos, realizar recodificaciones o procesamientos jerárquicos.

Figura 6. Definición de la tabla

The image shows a web-based interface titled "Definición de la tabla". It features a central area with six input fields, each preceded by a button labeled "Selección", "Fila", "Columna", "Control", "Sumarización", and "Ponderación". Below these fields is a row of five buttons: "Procesar", "Limpiar", "Diccionario", "Ejemplo", and "Ayuda". At the bottom of the interface, there is a section titled "Cambiar Base de Datos" containing two dropdown menus: "País: Seleccione" and "Base de Datos: Seleccione". A help icon is visible in the top right corner of the window.

Los botones de la sección "*Definición de la tabla*" permiten seleccionar las variables y establecer algunas expresiones. La función de cada uno de estos elementos se detalla a continuación.

- **Selección:** esta opción permite seleccionar o filtrar casos, generando tabulados para un grupo de casos que cumpla con alguna característica de interés. Por ejemplo, seleccionar únicamente a la población femenina, a los residentes de la zona urbana del país, personas mayores de 65 años, etc.

En la opción de selección **obligatoriamente debe escribirse una expresión lógica**, que defina el conjunto de casos que participarán en la construcción del cuadro. Sin embargo, cuando se desea trabajar con la totalidad de los casos, este espacio no es requerido y puede permanecer en blanco.

La opción de selección puede admitir tanto expresiones simples que establezcan restricciones para una única variable, como expresiones complejas que involucren restricciones para dos o más variables y el uso de operadores lógicos o aritméticos.

EJEMPLO 1. Diversas formas de establecer filtros mediante la opción *Selección*, utilizando la base de datos del “Censo de Población y Vivienda 2011”.

a. Población de 15 años o más

Selección P03_EDAD>=15

b. Población de 15 a 64 años de edad

Selección P03_EDAD>=15&&P03_EDAD<=64

c. Población de 15 años o más, que se encuentra ocupada

Selección P03_EDAD>=15&&P42_CONDUCT=1

d. Población de 15 a 64 años de edad, que se encuentra ocupada

Selección (P03_EDAD>=15&&P03_EDAD<=64)&&P42_CONDUCT=1

Note el uso del operador AND (&&) en los ejemplos b, c y d mediante el cual se establecen dos condiciones simultáneamente. En el caso de los ejemplos c y d, la expresión utilizada involucra restricciones para dos variables. Note también el uso de paréntesis en el ejemplo d, que permite controlar que la restricción establecida para la variable “condición de actividad” se aplique tanto a las personas de 15 años o más, como a las personas de 64 años o menos.

En los ejemplos b y d, se estableció un rango para la variable edad: entre los 15 y 64 años, observe que a pesar de que la restricción se refiere a la misma variable, el nombre de ésta se escribe dos veces.

- **Fila:** indica cuál es la variable que se mostrará en la fila de la tabla. Cada una de las categorías o valores que toma la variable seleccionada representará una fila de la tabla.
- **Columna:** indica cuál es la variable que se mostrará en columna. Cada una de las categorías o valores de la variable seleccionada representará una columna de la tabla.
- **Control:** permite generar tablas independientes, una por cada categoría de la variable seleccionada.

Las opciones *Fila*, *Columna* y *Control* permiten generar tablas de una, dos y hasta tres vías, respectivamente. **La expresión más elemental es el propio nombre de una variable**, como se muestra en el Ejemplo 2. Basta con seleccionar una variable, de tal manera que su nombre aparezca en alguna de las casillas, para que se puedan obtener resultados.

EJEMPLO 2. Distribución de la población, según variables de interés, utilizando las opciones *Fila*, *Columna* y *Control*.

a. Número de habitantes, por edades simples

Selección	<input type="text"/>
Fila	P03_EDAD
Columna	<input type="text"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

b. Número de habitantes, por edades simples y sexo

Selección	<input type="text"/>
Fila	P03_EDAD
Columna	P02_SEXO
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

c. Número de habitantes, por provincia de residencia, sexo y estado conyugal

Selección	<input type="text"/>
Fila	P03_EDAD
Columna	P20_ESTADO_CONYUGAL
Control	P02_SEXO
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

Tenga presente que es más fácil leer una tabla que sea más “larga” que “ancha” y por ello, es preferible que aquellas variables que posean un mayor número de categorías o valores, se presenten en las filas de la tabla, y siempre que sea posible, se omita presentarlas en columna o control. Este es un principio, que se emplea para la construcción de cualquier tabla estadística.

En este ejemplo, la variable con el mayor número de categorías, es la variable de edad (P03_EDAD), seguida por la variable estado conyugal (P20_ESTADO_CONYUGAL), que posee seis categorías y sexo (P02_SEXO), que posee solo dos. Por esta razón, en el ejemplo 2c. la variable sexo se muestra como control y no como columna, tal como se mostró en 2b.

Observe que en la consulta no es necesario incluir ninguna instrucción que contabilice el número de habitantes, esta acción se logra por defecto.

- **Sumarización:** calcula y muestra el número de casos, el promedio, desviación estándar y los valores mínimo y máximo para una variable de interés. Es conveniente que la variable seleccionada sea métrica (edad, años de escolaridad, número de hijos, etc.). Esta opción puede combinarse con *Fila*, *Columna* o *Control* para obtener la distribución de la variable seleccionada en *Sumarización*, por cada combinación entre las categorías de las variables restantes.

EJEMPLO 3. Obtener los años promedio de escolaridad de la población mediante la opción *Sumarización*, de acuerdo con las condiciones que se indican en cada caso.

a. Años promedio de escolaridad en la población

Selección	P03_EDAD>=5
Fila	
Columna	
Control	
Sumarización	P40_ANOS_ESCOLARIDAD
Ponderación	

b. Años promedio de escolaridad, por sexo

Selección	P03_EDAD>=5
Fila	P02_SEXO
Columna	
Control	
Sumarización	P40_ANOS_ESCOLARIDAD
Ponderación	

c. Años promedio de escolaridad, por sexo y zona de residencia

Selección	P03_EDAD>=5
Fila	P02_SEXO
Columna	ID_ZONA
Control	
Sumarización	P40_ANOS_ESCOLARIDAD
Ponderación	

Observe que la instrucción

Selección	P03_EDAD>=5
-----------	-------------

 establece que los resultados se muestren solamente para la población de 5 años y más, pues la pregunta sobre el nivel educativo solo se le formula a esta población.

- **Ponderación:** expande o multiplica cada caso o registro, por la frecuencia de la variable seleccionada. Su uso se discute más adelante en la sección 3.2 y el capítulo de *Ejemplos prácticos*.

Al presionar alguno de los botones descritos, se desplegará en pantalla el “Asistente para la construcción de expresiones” (ver Figura 7) que permite seleccionar las variables y establecer algunas condiciones o restricciones, como las que se han mostrado en los ejemplos anteriores.

Observe que en la sección inferior del *Asistente para la definición de tablas* (ver Figura 6) se encuentran los botones *Procesar*, *Limpiar*, *Diccionario*, *Ejemplo* y *Ayuda*. Después de que se hayan establecido todas las condiciones para generar el tabulado que desea, deberá presionar el botón *Procesar*, que le indica al sistema que calcule la tabla y muestre en pantalla los resultados de la consulta formulada.

La opción *Limpiar* le permitirá borrar todas las condiciones establecidas, y así realizar una nueva consulta. Al presionar el botón *Diccionario*, aparecerá en pantalla una nueva ventana que le mostrará el diccionario de datos de la base con la que se encuentra trabajando (ver sección “Diccionario de datos, cuestionarios y boletas” más adelante en este capítulo). El botón de *Ayuda* le redirigirá a la *Guía rápida para formular consultas*, para que pueda acceder a ella durante la consulta.

2.2. ASISTENTE PARA LA CONSTRUCCIÓN DE EXPRESIONES

Este módulo permite seleccionar, de manera interactiva, las variables requeridas y establecer algunas condiciones o restricciones para el análisis. Los usuarios más experimentados, podrían omitir este paso, escribiendo directamente las instrucciones en las casillas. Puede acceder a este módulo al presionar cualquiera de los botones del *Asistente para la definición de tablas*.

Figura 7. Asistente para la construcción de expresiones

El *Asistente para la construcción de expresiones* (ver Figura 7) permite construir expresiones lógicas, relacionales o aritméticas, para formular consultas complejas. Una expresión es una combinación de variables, operadores o constantes que permite establecer condiciones o restricciones que se aplicarán a un conjunto de datos.

A través de este asistente podrá construir la expresión que requiera, mediante la selección de variables, operadores o valores que deberán seguir una estructura lógica. La estructura de la expresión varía de acuerdo con las restricciones que se deseen establecer, no existe una regla general que defina una forma básica. Sin embargo, deben tenerse en cuenta los siguientes criterios:

1. En toda consulta es necesario que se seleccione al menos una variable.
2. La expresión más simple es la selección de la variable misma, si no hay restricciones, bastará con seleccionar el nombre de la variable.
3. El uso de operadores y valores será necesario solamente cuando se deba establecer una restricción o filtro para los casos que entrarán en análisis.
4. La definición de expresiones es particularmente útil en el caso de la opción *Selección*. Ésta no tendrá mucho sentido en *Sumarización* o *Ponderación*.
5. **Las expresiones que se establezcan en *Selección* permitirán filtrar o seleccionar casos.** Al establecer una expresión en *Fila*, *Columna* o *Control* el tabulado que se genere **indicará el número de casos que cumplen o no, con dicha expresión.**

Para seleccionar una variable, deberá presionar doble clic sobre alguna de las que se listan en el menú "*Variables*", en la sección superior izquierda de la ventana. La variable seleccionada se mostrará en color azul y su nombre aparecerá en el recuadro inferior, como se muestra en la imagen (ver Figura 7).

Note que en el recuadro "*Valores*", en la sección superior derecha de este asistente, se despliegan las categorías o valores asociados a la variable que se ha seleccionado. Estos también pueden ser seleccionados, presionando doble clic. Los operadores lógicos y aritméticos que se le presentan en esta ventana, son seleccionados tras presionar una única vez el operador requerido.

Conforme se seleccionan las variables, operadores y valores, la expresión se va escribiendo en el recuadro inferior de la ventana. Una vez que haya terminado de escribir la expresión, para continuar con el proceso, presione el botón "*Aceptar*", que le permitirá volver nuevamente al *Asistente para la definición de tablas*. Note que la expresión que ha establecido, aparece ahora en la casilla (*Selección*, *Fila*, *Columna* o *Control*) correspondiente.

2.3. USO DE OPERADORES LÓGICOS Y ARITMÉTICOS

A pesar de que el Sistema de Consulta solamente puede producir tabulaciones de dos o tres dimensiones, la posibilidad de emplear expresiones aritméticas en la definición de la tabla permite construir tablas sumamente complejas para satisfacer la mayoría de las necesidades de los usuarios.

Las expresiones aritméticas pueden contener nombres de variables, constantes numéricas, operadores aritméticos: +, -, *, /, operadores lógicos Y: && O: || y operadores de relación: =, !=, >, <, >=, <=, además de los paréntesis. A continuación, se describe la función de cada uno de ellos.

- =** **Igual que:** se utiliza para establecer una igualdad, de tal manera que solo se analizarán aquellos casos cuyo valor sea exactamente igual al seleccionado.
- !=** **Diferente a:** tiene la función inversa a “igual que”, con este operador se considerarán únicamente los casos cuyo valor es diferente al seleccionado.
- <** **Menor que:** permite considerar únicamente aquellos aquellos casos cuyo valor es inferior al seleccionado.
- <=** **Menor o igual que:** a diferencia “menor que” este operador permite seleccionar los casos cuyo valor es menor o igual al que se ha establecido en la expresión.
- >** **Mayor que:** permite analizar casos cuyo valor es estrictamente mayor al que se ha seleccionado.
- >=** **Mayor o igual que:** con este operador se podrán analizar casos cuyo valor sea mayor o igual al que se hace referencia.
- +** **Suma:** este operador puede ser utilizado para sumar una constante a cada valor de una variable, o bien, para sumar dos o más variables.
- **Resta:** con el este operador es posible restar una constante a una variable o restar los valores de dos o más variables.
- *** **Multiplicación:** permite multiplicar por una constante o por una variable.
- /** **División:** este operador permite la división por un valor constante o entre una variable.
- (** **Paréntesis:** permiten agrupar operaciones y establecer prioridad en la ejecución de las expresiones. Las operaciones que están dentro de paréntesis se evalúan primero. Si existen diferentes paréntesis anidados, las expresiones más internas se evalúan primero.
-)** Su correcta definición requiere tanto del paréntesis inicial **(** como del paréntesis final **)**.
- AND** **&& (Y: conjunción):** mediante este operador es posible asociar diferentes expresiones, permitiendo que las condiciones involucradas se cumplan **simultáneamente**.
- OR** **|| (ó: disyunción):** el operador **OR** permite establecer una relación entre dos expresiones, que indicará la selección de los casos que cumplan **con al menos una** de las restricciones.

Los operadores **=**, **!=**, **<**, **<=**, **>**, **>=** pueden ser empleados para establecer una relación entre una variable y un valor numérico, o bien, entre dos variables. Cuando se emplean para establecer una relación entre dos variables, los resultados mostrarán el número de casos que cumplen o no, con la condición establecida a través de las indicaciones “*true*” y “*false*”, respectivamente.

Es importante señalar que **todas las expresiones desplegarán resultados en valores numéricos enteros. Los valores decimales serán truncados al valor entero más próximo.** Algunas expresiones podrán devolver cifras redondeadas, particularmente aquellas que involucren el uso de los operadores *multiplicación* o *división*.

EJEMPLO 4. Uso de operadores lógicos y aritméticos para establecer las expresiones, que permitan generar información para las restricciones deseadas.

a. Población femenina

P02_SEXO=2

P02_SEXO!=1

b. Población femenina, residente en la zona rural

P02_SEXO=2 && ID_ZONA=2

c. Población menor de 65 años

P03_EDAD<65

P03_EDAD<= 64

d. Población de 65 años o menos

P03_EDAD<= 65

e. Población de 15 años o más

P03_EDAD> 14

P03_EDAD>= 15

f. Población entre 15 y 64 años de edad (inclusive)

P03_EDAD>= 15 && P03_EDAD<= 64

g. Población entre 15 y 64 años, que se encuentran ocupados

(P03_EDAD>= 15 && P03_EDAD<= 64) && P42_CONDUCT=1

h. Población menor de 15 años, o de 65 años o más

P03_EDAD< 15 || P03_EDAD>= 65

i. Número de hogares en los que habitan menos mujeres que hombres

H05B_TOTAL_MUJERES < H05A_TOTAL_HOMBRES

Note que al establecer una restricción, pueden existir una serie de relaciones equivalentes, como se muestra en a, c y e.

El uso de paréntesis en 4g permite establecer el orden en el que se procesará la expresión. En este caso, se seleccionará primero a la población según su edad. De tal manera que, entre las personas en el rango de 15 y 64 años de edad, se contabilizará a los que están ocupados.

En los ejemplos 4f y 4h es posible observar la diferencia entre los operadores *AND* (&&) y *OR* (||). Cada individuo podría cumplir las dos condiciones del ejemplo 4f (ser mayor de 15 y menor de 65), pero no pueden cumplir simultáneamente las condiciones establecidas en 4h (ser menor de 15 y a su vez, ser mayor de 65 años), razón por la cual se usa la expresión *OR*.

2.4. DELIMITAR RANGOS

En algunas ocasiones puede ser necesario delimitar el rango de códigos que tendrá una variable o expresión. Por ejemplo, seleccionar a la población que se encuentra entre los 15 y 64 años de edad. En esos casos, existe una función que permite establecer el rango deseado, sin la necesidad de utilizar los operadores relacionales. Observe el siguiente ejemplo:

Fila P03_EDAD:15..64

Note que después del nombre de la variable, el rango se define iniciando con dos puntos (:), después el valor inicial del rango, punto-punto (..) y el valor final del rango. La expresión anterior mostrará la distribución por edad de la población, pero solamente para las personas entre los 15 y 64 años de edad. Note que esta expresión es equivalente a la que se estableció en el ejemplo 4f.

Esta función para definir rangos **solo puede ser utilizada en las opciones *Fila* o *Columna*** de la tabla. La opción no es válida en las casillas *Selección*, *Sumarización* o *Ponderación*, y tampoco puede ser combinada con otras expresiones, sino que **debe ser utilizada en su forma más simple**, tal como se mostró en el ejemplo.

2.5. DICCIONARIO DE DATOS, CUESTIONARIOS Y BOLETAS

Una buena práctica antes de empezar a generar tabulados, es familiarizarse con la base de datos que se empleará. Para ello se puede consultar el diccionario de datos y el cuestionario o formulario a partir del cual se ha recopilado la información. En el *Asistente para la construcción de la tabla* al presionar el botón “*Diccionario*” (ver Figura 6) se abrirá una nueva ventana como la que se muestra en la Figura 8.

Figura 8. Diccionario de datos



Jerarquía: Vivienda			
Variable	Descripción	Rango	Códigos
ID_PROVINCIA	Provincia	0-9	1 San José
ID_CANTON	Cantón	101-706	101 San José
ID_DISTRITO	Distrito	10101-70605	10101 Carmen
ID_AE_CONSECU	área de empadronamiento (consecutivo)	0-999999	-
ID_VIVIENDA	Vivienda	0-999	-
ID_HOGAR	Hogar	0-9	-
ID_ZONA	Zona	0-9	1 Urbano
ID_REGION	Región	0-9	1 Gran Área Metropolitana
ID_TERR_INDIGENA	Territorio indígena	0-99	11 Salitre

En la parte superior de esta página, justo debajo del nombre de la base de datos, es posible encontrar un enlace que permite descargar la boleta empleada en el censo o encuesta.

La primera columna del recuadro muestra el nombre asignado a cada variable en la base de datos; **este es el nombre que se observa al seleccionar una variable para construir una expresión**. La segunda columna corresponde a la “*etiqueta de la variable*” pues muestra una descripción más específica de ella. Esta es la descripción que se observa en el menú para seleccionar variables.

Los menús desplegables bajo la columna “Códigos” permiten ver los posibles valores que toma cada variable. Esta información puede ser particularmente útil para conocer la base de datos y planear la recodificación de variables.

2.6. CAMBIAR BASES DE DATOS

Los usuarios pueden moverse entre las distintas bases de datos, cambiando la base de datos con la que está trabajando, sin la necesidad de volver al menú *Selección de la base de datos*. En la sección inferior del *Asistente para la construcción de la tabla* (ver Figura 6) se encuentran dos menús desplegables. El primero permite seleccionar el país al que pertenece la base de datos, mientras que el segundo permite seleccionar la nueva base de datos con la que se desea trabajar.

2.7. RECODIFICACIÓN DE VARIABLES

En ocasiones las categorías o valores preestablecidos para una variable no responden exactamente a las categorías con las que se desea trabajar. En ese caso, es posible utilizar la recodificación de variables para establecer una nueva codificación que permita agrupar dos o más categorías. Este puede ser el caso de variables como la edad, que en las diferentes bases de datos se registra en edades simples (0, 1, 2...) pero que en ocasiones se desea analizar en grupos quinquenales (0 a 4 años, de 5 a 9 años, etc.) o en grupos específicos.

Figura 9. Preparación de recodificaciones

Preparación de Recodificaciones

====Complete la recodificación requerida====

R: Fila

r: 0 14 1 Menores de 15 años

r: 15 64 2 De 15 a menos de 65 años

r: 65 112 3 De 65 años y más

Desde: 65 Hasta: 112 Vale: 3 Etiqueta: De 65 años y más

Incluir Excluir Cambiar Aceptar Cancelar

La recodificación de variables se establece desde el *Asistente para la construcción de expresiones* (ver Figura 7), seleccionando la variable con doble clic y presionando a continuación el botón *Recodifica* en la sección inferior derecha de la ventana. Al presionar este botón aparece automáticamente el asistente para la *Preparación de recodificaciones*, como se muestra en la Figura 9.

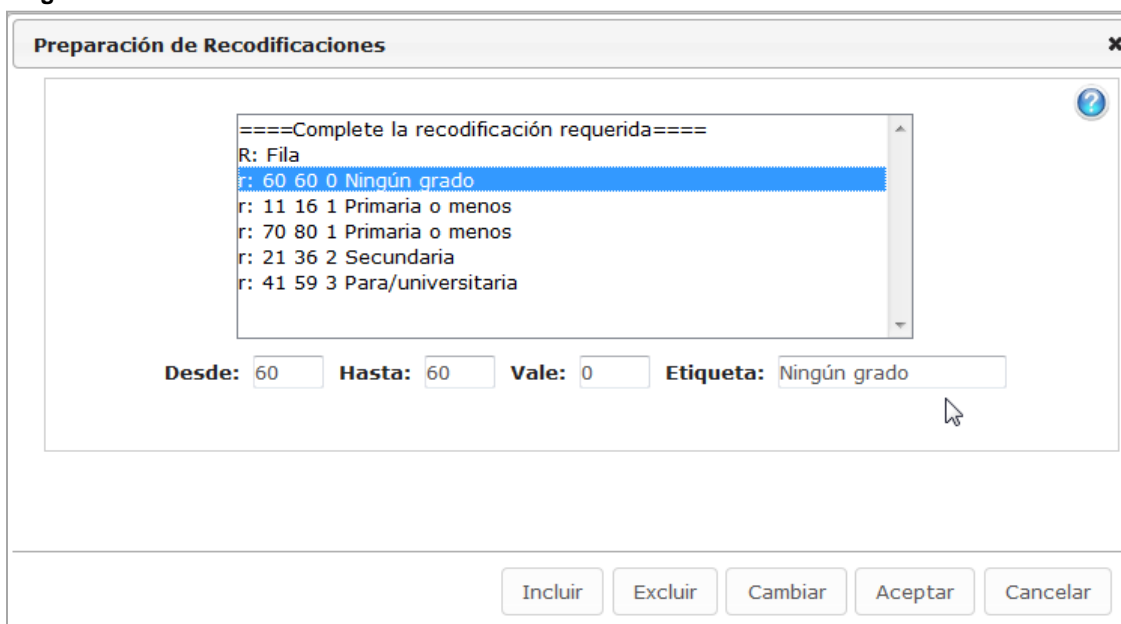
Complete la información de las casillas *Desde* y *Hasta* con los valores mínimo y máximo, respectivamente, que incluirá cada nueva categoría. En la casilla *Vale* se debe indicar el nuevo código que desea asignar. La casilla *Etiqueta* permite anotar una breve descripción de lo que representará la nueva categoría.

En el ejemplo que se observa en la Figura 9, se ha recodificado la variable edad (P03_EDAD) en tres nuevas categorías. La primera línea indica que aquellos casos cuyas edades se encuentren entre 0 y 14 años, serán asignados en una nueva categoría con el código "1" que recibirá el nombre "Menores de 15 años". De manera análoga, la segunda categoría incluye a las personas entre los 15 y 64 años, mientras que la tercera incluye a aquellas personas de 65 años o más.

Presione el botón *Incluir* al haber introducido los criterios de cada nueva categoría. Automáticamente aparecerán en el recuadro las nuevas expresiones definidas, como se muestra en la Figura 9. Si desea eliminar o modificar alguna de las expresiones deberá posicionarse sobre ella, de tal manera que se muestre en color azul, y presionar los botones *Excluir* o *Cambiar*, según corresponda. Al haber finalizado, deberá presionar el botón *Aceptar* para regresar nuevamente al *Asistente para la construcción de expresiones*.

Para la recodificación de variables **las categorías deben ser mutuamente excluyentes y es necesario analizar la codificación empleada inicialmente para determinar cuáles serán los nuevos rangos que tendrá cada nueva categoría.**

Figura 10. Recodificación de la variable nivel de instrucción



La Figura 10, muestra la codificación de la variable "nivel de instrucción" (P16_ULTIMO_GRADO_APROBADO) en las categorías 'Ningún nivel de instrucción', 'Primaria o menos', 'Secundaria' y 'Para/universitaria'. En

este ejemplo, los valores insertos en las casillas “Desde” y “Hasta” son iguales en la primera categoría, ya que se desea modificar un único valor y no un rango de valores. Observe también que la categoría ‘Primaria o menos’ aparece dos veces. Los códigos 70 (Enseñanza especial) y 80 (Kinder y preparatoria) fueron incluidos en esa misma categoría, pero como no poseen un rango próximo a los valores referentes a la educación primaria (del 11 al 16) deben insertarse de manera independiente.

La opción de recodificación no almacena la información en la base de datos, ni crea una nueva variable en ella, por lo que los usuarios no deben preocuparse de modificar la base de datos con la que se encuentran trabajando. En su lugar, el sistema de consultas genera temporalmente una nueva variable que se almacena en memoria. Al cerrar la sesión, cerrar la página o limpiar la pantalla se borrará la información que se ha ingresado.

2.8. RECODIFICACIÓN POR RANGOS

En algunas ocasiones, quizá sea necesario recodificar una variable de tipo métrico en una variable categórica, pero cuyas categorías posean un **rango estándar**. Tal puede ser el caso al recodificar la variable edad en grupos quinquenales, donde cada una de las categorías tendrá un rango de 5 años; rango que se mantendrá constante para cada grupo etario definido.

El Sistema de Consultas permite realizar el proceso de recodificación en este formato de manera simplificada, a través del uso del operador de división. Tras seleccionar la variable a recodificar, desde el *Asistente para la construcción de expresiones* (ver Figura 7), basta con presionar el botón de división y a continuación escribir el rango estándar que debe poseer cada categoría.

Para generar una variable de edad en grupos quinquenales, basta con escribir la siguiente instrucción en alguna de las casillas *fila*, *columna* o *control*, como se muestra a continuación.

Fila	P03_EDAD/5
------	------------

También es posible establecer otros rangos para la recodificación de variables, **siempre que el rango sea contante para cada categoría**. Para recodificar la variable edad en grupos decenales, bastará entonces con escribir la instrucción: “P03_EDAD/10”.

3. GENERACIÓN DE RESULTADOS

Tras establecer las características de la tabla, deberá presionar el botón “Procesar” en el *Asistente para la construcción de la tabla*. Tras esta acción, automáticamente se abrirá una nueva ventana, que mostrará los resultados, tal como se observa en la Figura 11.

Los resultados podrán ser exportados a un archivo *Microsoft Office Excel*, tras presionar el botón “Exportar Tabla” que se encuentra en la sección superior derecha de la ventana. Con la opción “Ver Gráfico”, podrá ver los resultados en forma de gráfico, como se muestra en la Figura 12. Desde allí podrá imprimir o descargar el gráfico que se presenta en pantalla, mediante los dos pequeños botones que se encuentran en la parte superior derecha.

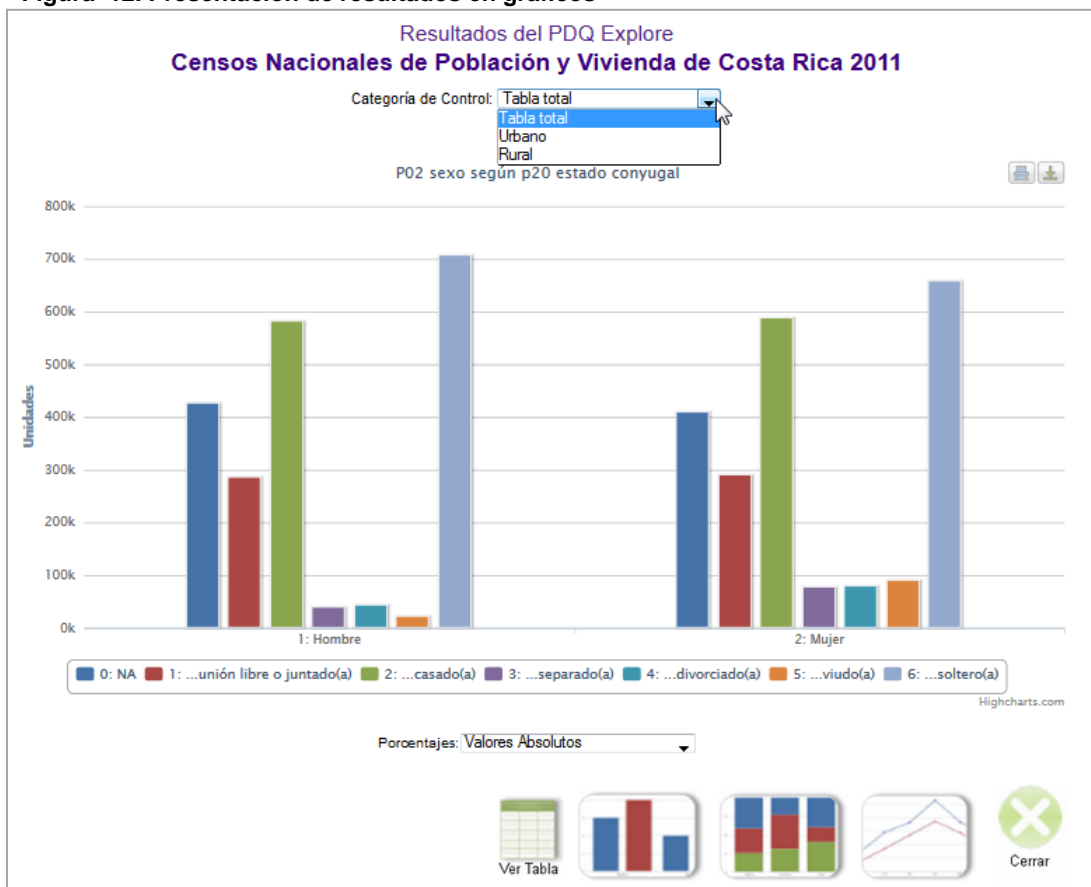
Figura 11. Ventana de resultados



La forma que toma el gráfico que se presenta en pantalla dependerá del tipo de variable que se está analizando y el número de categorías que posee. Por defecto, cuando la variable posee siete categorías o menos, las cifras se presentarán en un gráfico circular o de pastel. Para variables con un mayor número de categorías, se generará por defecto un gráfico de barras verticales. Cuando la consulta involucre simultáneamente dos o más variables continuas, con múltiples valores, como por ejemplo la edad o los años de educación, la información no podrá presentarse en forma gráfica, solamente en formato de tabla.

Cuando la opción esté disponible, se presentarán en la sección inferior derecha de la pantalla, otros tipos de gráficos para presentar la información, tales como los gráficos de líneas y gráficos de barras apiladas, como se muestra en la Figura 12. Al presionar el botón correspondiente, el usuario podrá seleccionar otro tipo de gráfico para mostrar la información.

Figura 12. Presentación de resultados en gráficos



Note que tanto en la ventana que muestra la tabla como en la que muestra el gráfico, aparecerá el menú “*Porcentajes*” en la parte inferior. Mediante este menú podrá estimar los porcentajes por fila, columna o con respecto al total de observaciones en la tabla.

Cuando la variable que se seleccionó en la casilla fila posee muchas categorías (por ejemplo, en variables como la edad o el distrito de residencia), además del menú *Porcentajes*, aparecerá también en la sección inferior de la pantalla el menú “*Número de filas por página*”, que le permitirá desplegar algunas secciones de la tabla, o bien, la totalidad de ella.

Cuando la tabla que ha generado involucra una variable de *Control*, aparecerá en la sección superior de la ventana, justo debajo del título, el menú “*Categoría de Control*” mediante el cual podrá seleccionar la categoría bajo la que se desplegarán los resultados.

4. PROCESAMIENTO JERÁRQUICO

En una misma base de datos, puede registrarse información de diferentes entidades o jerarquías, algunas veces “anidadas” unas dentro de otras. Tal es el caso de los Censos de Población y Vivienda, que registran información para tres entidades: viviendas, hogares y personas.

De esta forma, hay algunas variables cuya información se refiere específicamente a alguno de estos niveles. Por ejemplo, la variable “sexo” es un atributo de la entidad *Persona*, la variable “total de miembros del hogar” pertenece a la entidad *Hogar*, mientras que la variable “cielo raso” registra información propia de una *Vivienda*.

Cuando se tabula información, para cualquiera de las variables de los censos de población, **los resultados contabilizarán el número de casos de la jerarquía de menor rango involucrada en la consulta**. Note por ejemplo, que si tabula únicamente alguna de las variables descritas en el párrafo anterior, el número total de casos será diferente en cada consulta. De las tres jerarquías que se manejan en los Censos, *Persona* es la de menor rango y *Vivienda* es la de mayor.

Sin embargo, en algunos casos, se busca información del tipo *¿cuántas personas residen en viviendas con acceso a agua por tubería?*. Si se trabaja únicamente con la variable “agua por tubería” la tabla resultante mostrará el número de viviendas que reciben agua por tubería, y no el número de personas que viven en ellas, como se quiere determinar. La tabla que responde a esta interrogante, involucraría información de dos jerarquías: *Personas* y *Vivienda*. Las secciones 4.1 y 4.2 describen dos procesos que permiten responder a interrogantes de este tipo.

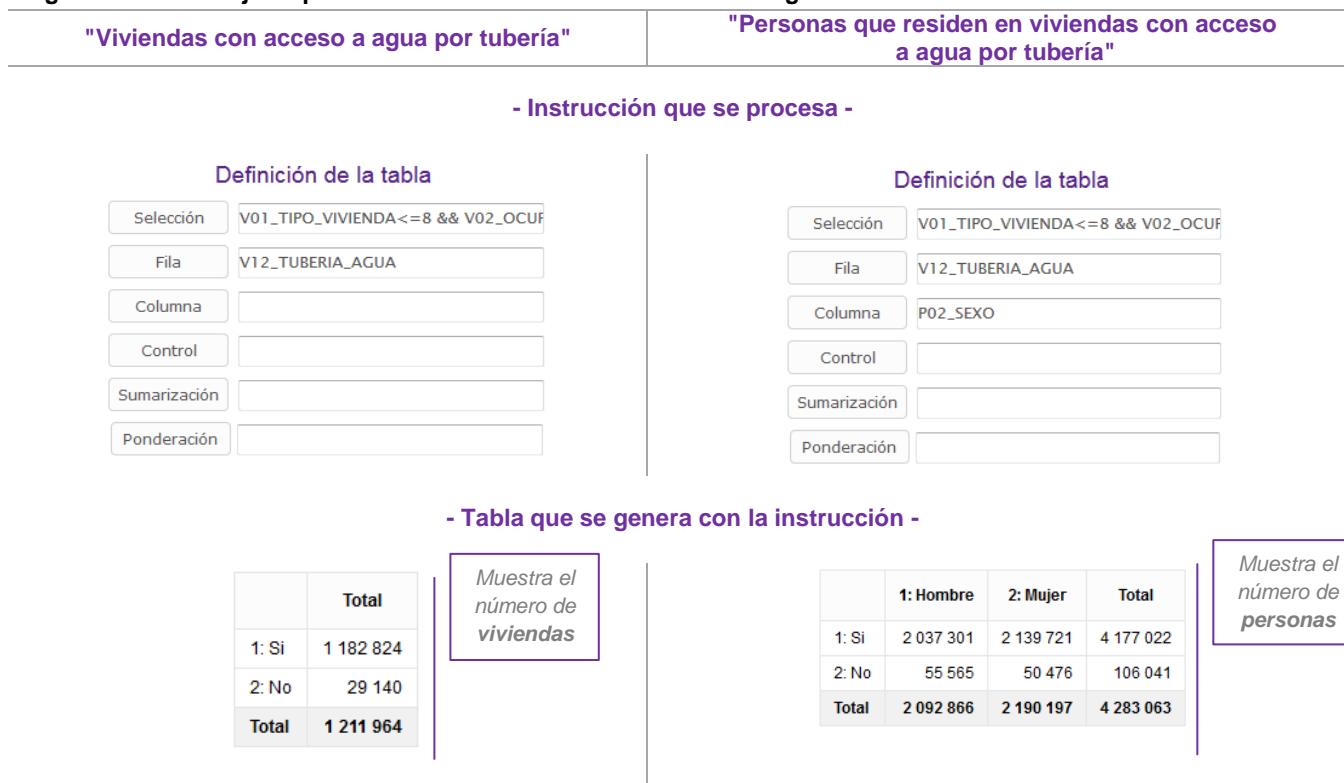
En otras ocasiones, por otra parte, será necesario considerar información de unidades de menor jerarquía pero los cuadros deben contabilizar las unidades de nivel jerárquico superior a las que pertenecen las primeras. Así por ejemplo, quizá se desee conocer *¿en cuántos hogares costarricenses habitan personas adultas mayores?*. Este tipo de operaciones se desarrollarán en la sección 4.3.

4.1. TABULADOS QUE IMPLICAN VARIABLES DE DIFERENTES JERARQUIAS

La primera forma, la más sencilla, para contabilizar unidades de menor jerarquía con información de variables de jerarquía superior, es involucrar en la consulta una variable que sea atributo de la entidad de menor rango. La Figura 13 muestra un ejemplo para obtener el número de personas que habitan en viviendas con acceso a agua por tubería. El panel de la izquierda, muestra la instrucción que se debe establecer para obtener el número de viviendas que tienen este beneficio, mientras que, la sección de la derecha muestra las instrucciones que permiten obtener el número de personas que habitan en viviendas que cuentan con acceso a agua por tubería. Nótese que en ese último caso, se incluyó en la consulta la variable “sexo”, que es un atributo de *Personas*.

Es importante destacar que en este ejemplo, el total de casos es diferente en cada procesamiento. En el primer panel, se contabilizan 1 211 964 viviendas, de las cuáles 1 182 824 sí reciben agua por tubería. Sin embargo, al referirse al número de personas, en el segundo panel, basta con observar la columna *Total* (si no se desea obtener el dato desagregado por sexo), que permite determinar que en Costa Rica 4 177 022 personas residían en viviendas que sí se abastecen de agua a través de tubería.

Figura 13. Análisis jerárquico involucrando variables de menor rango



En este mismo ejemplo (ver Figura 13), la instrucción que se incluye en la casilla *Selección*: V01_TIPO_VIVIENDA<=8 && V02_OCUPACION=1 filtra los resultados para contabilizar únicamente las viviendas individuales y ocupadas (o las personas que residen en ellas). La información referente a las características de las viviendas no se registra para aquellas viviendas desocupadas (disponibles para alquiler, en construcción o reparación, temporales para vacacionar) o colectivas (albergues infantiles, cárceles, hogares de ancianos o conventos, entre otras). Para obtener mayor información al respecto el lector puede consultar también la boleta censal.

4.2. USO DE LA FUNCIÓN DE PONDERACIÓN

La segunda posibilidad para trabajar con datos de tipo jerárquico es empleando la opción de *Ponderación*, en el asistente para la definición de la tabla (ver Figura 6). Esta opción permite multiplicar cada caso de acuerdo con la frecuencia de una variable seleccionada. Volviendo al ejemplo para determinar el número de personas que residen en viviendas con acceso a agua por tubería, utilizando la opción *Ponderación* basta con especificar la tabla de la siguiente manera para obtener la información requerida,

Definición de la tabla

Selección	VIVIENDA<=8 && V02_OCUPACION=1
Fila	V12_TUBERIA_AGUA
Columna	
Control	
Sumarización	
Ponderación	H05C_TOTAL_PERSONAS

En la opción de *Ponderación* se incluyó la variable “Total de personas que reside en cada vivienda” (H05C_TOTAL_PERSONAS) lo que permite multiplicar la frecuencia de la variable “agua por tubería” tantas veces como residentes habiten en cada vivienda. En la opción *Selección* nuevamente se han filtrado los resultados para que éstos se desplieguen únicamente entre las viviendas individuales y ocupadas. Estas instrucciones mostrarán los mismos resultados presentados en la Figura 13.

4.3. USO DE FUNCIONES JERÁRQUICAS

Una tercera posibilidad para trabajar con datos de tipo jerárquico es empleando funciones específicas para tal fin. El Sistema de Consultas cuenta con algunas funciones que se encuentran disponibles en el *Asistente para la construcción de expresiones* (ver Figura 7). Al definir la tabla, estas funciones se pueden emplear en las casillas selección, fila, columna o control de un tabulado.

Estas funciones permiten realizar operaciones que se representen a un nivel jerárquico superior. A continuación se describen las funciones disponibles para el procesamiento jerárquico.

- **SUM:** la opción *SUM* tiene dos modalidades definidas por el tipo de argumento. Cuando el argumento es una condición, esta función cuenta el número de casos que la cumplen y se crea una variable a un nivel jerárquico superior con este conteo como contenido (ver Ejemplo 5a).

Por el contrario, cuando el argumento de la función es una variable, *SUM* crea una variable a nivel jerárquico superior con la suma correspondiente. Suponga por ejemplo que desea conocer el ingreso total de cada hogar, a partir de la suma del ingreso que percibe cada uno de sus miembros. En ese caso, la instrucción que requiere tendrá la siguiente forma: *SUM(ingreso):0..99999999*

- **MIN Y MAX:** estas funciones permiten obtener el valor mínimo o el máximo, respectivamente, de la variable que se ha establecido como argumento. Este valor se expresa para el nivel jerárquico inmediato superior. Por ejemplo, permite determinar el nivel educativo más alto de los miembros del hogar o la edad del miembro más joven dentro del hogar (ver Ejemplos 5b y 5c, respectivamente).
- **PICK:** devuelve el valor de la variable para la primera ocurrencia que cumple con la condición establecida. Esta función se compone de dos argumentos:

PICK(condición, variable)

donde, *condición* permite identificar el primer caso que la cumpla y *variable* corresponde al nombre de la variable para la que se desea obtener la información. Con esta función es posible obtener, por ejemplo, el nivel educativo del jefe del hogar o el sexo del conyugue del jefe del hogar.

EJEMPLO 5. Uso de funciones para el procesamiento jerárquico.

a. Hogares con residentes de 65 años o más

Fila `sum(P03_EDAD >= 65):0..99`

b. Número de hogares según el nivel educativo más alto alcanzado por cualquier miembro

Fila `max(P40_ANOS_ESCOLARIDAD):0..99`

c. Número de hogares según edad del miembro más joven

Fila `min(P03_EDAD):0..112`

d. Nivel educativo del jefe del hogar

Fila `pick(P01_PARENTESCO=1, P16_ULTIMO_GRADO_APROBADO):0..99`

Observe que **en cada una de las funciones expuestas es necesario especificar el rango esperado (:00..99) del resultado de la función**. Este rango varía de acuerdo con el número de dígitos en los que se espera varíen los resultados, si cree que la variable resultante tendrá tres dígitos, entonces debe ingresarse el rango 0..999. Nótese que después de la función, el rango se establece iniciando en dos puntos (:), después el valor inicial del rango, punto-punto (..) y el valor final.

En los ejemplos 5a, 5b y 5c también es conveniente filtrar los resultados, a través de la opción selección, de tal manera que solo se contabilicen los residentes de viviendas individuales y ocupadas (`V01_TIPO_VIVIENDA<=8 && V02_OCUPACION=1`). De esta forma, se excluirán del conteo los hogares de ancianos y cárceles. Se excluirán también las viviendas desocupadas, para las que no se registra información sobre la edad o nivel educativo de las personas, y en esos casos, se emplea el código 0 (cero) para denotar “No aplica” (NA) el registro de esta información.

En el ejemplo 5a es posible determinar, por ejemplo, que existen 996 552 hogares en los cuáles ninguno de sus integrantes tiene 65 años o más y en 175 739 hogares habita tan solo un adulto mayor.

El ejemplo 5c permite conocer que en 64 172 hogares el menor de sus miembros era un niño menor de un año y hay 9 hogares en los que el miembro de menor edad tiene 100 años. Observe que en este ejemplo, dado que en el Censo de Población se registran personas de 100 años o más, el rango definido es más amplio. En este caso, como se conoce el rango de la edad (desde 0 hasta 112) es posible especificar el rango de manera más puntual. Si este valor no se conociera, podría emplearse un rango estándar, por ejemplo, 0..999.

Utilizando la función *PICK*, en el ejemplo 5d, es posible determinar que en 24 747 hogares el jefe había alcanzado hasta el primer grado de primaria, y en 54 449 el jefe del hogar no contaba con ningún grado de escolaridad (ver código 60, al final de la tabla), mientras que en 26 169 hogares el jefe había alcanzado hasta nueve años de estudios universitarios.

5. EJEMPLOS PRACTICOS

Esta sección muestra paso a paso cómo generar los principales indicadores demográficos y sociales que se estiman a partir de las bases de datos disponibles en el Sistema de Consultas. Los ejemplos desarrollados en esta sección, permitirán aprender a utilizar esta herramienta de una manera más eficiente, pues se mostrarán algunos ejemplos con operaciones un tanto más complejas que las expuestas en los capítulos anteriores.


Con esta sección se pretende que los usuarios tomen los ejemplos como referencia, ya sea para calcular exactamente el mismo indicador, o bien, para llevar a cabo procedimientos análogos. En esta sección también podrá aprender algunas expresiones más avanzadas.

NÚMERO DE HABITANTES POR CANTÓN, SEGÚN SEXO

Este indicador es uno de los requeridos con mayor frecuencia. Permite determinar el número de hombres y mujeres que habitan en Costa Rica, según su cantón de residencia.

Empleando la base de datos del **Censo de Población y Vivienda**, se deben seguir estos procedimientos,

1. Tras insertar su código de usuario y contraseña, e ingresar al Sistema de Consultas, tendrá acceso a la página *Selección de la base de datos*, donde podrá seleccionar la base de datos del Censo de Población y Vivienda del año 2011.
2. Presione el botón *Procesar* al fondo de esta página. Tras esta acción será redirigido automáticamente al *Asistente para la definición de la tabla*.
3. Presione el botón *fila* para acceder al *Asistente para la construcción de expresiones*. Seleccione la variable “Cantón” (ID_CANTON) presionando doble clic sobre ella, desde el listado de variables que se muestra en la sección derecha de la ventana. A continuación presione el botón *Aceptar* para volver al *Asistente para la definición de la tabla*.
4. Seleccione la variable “Sexo” (P02_SEXO) en la casilla *Columna*, siguiente un procedimiento análogo al descrito en el punto anterior. La definición de la tabla deberá verse de la siguiente manera:


Definición de la tabla 

Selección	<input type="text"/>
Fila	<input type="text" value="ID_CANTON"/>
Columna	<input type="text" value="P02_SEXO"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

5. Finalmente, presione el botón *Procesar* desde el *Asistente para la definición de la tabla* para obtener los resultados. Éstos se desplegarán en una nueva ventana en su navegador web.

Esta misma información también puede generarse a partir de la base de datos de las **Proyecciones Distritales de Población**, para obtener información sobre un año de interés, inclusive para alguno futuro. Empleando esta base de datos, las acciones a seguir son las siguientes:

1. Desde la página *Selección de la base de datos* seleccione la opción “Proyecciones Distritales de Poblacion de Costa Rica 2000-2050”, que le permitirá tener acceso a esa base de datos. Presione el botón *Procesar* al fondo de la página, para ser redirigido al *Asistente para la definición de la tabla*.
2. Presione el botón *fila*, que le permitirá abrir el *Asistente para la construcción de expresiones* dónde podrá seleccionar la variable “Cantón” (Canton). Busque esta variable en el menú de la izquierda y presione doble clic sobre ella. Observe que ésta aparecerá en el recuadro que se muestra en el fondo de la ventana. Presione el botón *Aceptar*, para continuar con la definición de la tabla.
3. Siguiendo un procedimiento análogo seleccione la variable “Sexo” (sexo) en la casilla *Columna*.
4. La base de datos sobre las Proyecciones Distritales de Población, contiene información en el periodo comprendido entre el año 2000 y el 2050. **Si no se estableciera una restricción según año, los resultados desplegarían la suma del número de habitantes para cada uno de estos años.** Presione el botón *Selección*, en el *Asistente para la definición de la tabla* y construya la expresión: `ano=2015` utilizando los botones disponibles para tal fin, o tecleando la instrucción, en el *Asistente para la construcción de expresiones*. En este caso se obtendrá el número de habitantes por cantón según sexo, únicamente para el año 2015.
5. Las instrucciones deben haberse establecido como se muestra en la figura adjunta. Al finalizar la definición de la tabla, presione el botón *Procesar* para obtener la información deseada.

Definición de la tabla 


Selección	ano=2015
Fila	Canton
Columna	sexo
Control	
Sumarización	
Contenido	Población

La base de datos de Proyecciones Distritales de Población, también permite conocer el número de habitantes por provincia o distrito. Para obtener estas cifras, basta con seleccionar la variable “Provincia” (provincia) o “Distrito” (distrito) en lugar de la variable “Cantón” y seguir el resto del procedimiento de la misma manera descrita anteriormente.

NÚMERO DE HABITANTES POR AÑO, SEGÚN SEXO

Para obtener el número total de habitantes para Costa Rica de forma anual, se deben seguir las siguientes instrucciones:


1. Tras ingresar al Sistema de Consultas ingresando el código de usuario y contraseña, deberá seleccionar la base de datos “Proyecciones Nacionales de Población” y presionar el botón *Procesar*, en la página *Selección de la base de datos*.
2. La acción anterior lo llevará al *Asistente para la definición de la tabla*,
3. Presione el botón *Fila* y en el menú *Variables* que se muestra en la sección izquierda de la ventana emergente, seleccione con doble clic la variable “Año de trabajo” (anio). Presione el botón *Procesar* para continuar con la definición de la tabla.
4. Nuevamente desde el *Asistente para la definición de la tabla*, presione el botón *Columna* y seleccione la variable “Sexo” (sexo).
5. Al finalizar la selección de ambas variables presione el botón *Procesar*, para obtener los resultados deseados. Estos se desplegarán en una nueva ventana en el explorador web.
6. La definición de la tabla debe establecerse como se muestra en la figura siguiente,

Definición de la tabla 

Selección	<input type="text"/>
Fila	anio
Columna	sexo
Control	<input type="text"/>
Sumarización	<input type="text"/>
Contenido	Población ▼


La base de datos de Proyecciones Nacionales de Población, contiene información en el periodo comprendido entre el año 1950 y 2100. Si se deseara obtener información para un único año, o para un periodo menor puede establecerse esta restricción empleando la opción *Selección*.

Para seleccionar un único año, por ejemplo el año 2015, la tabla puede definirse de la siguiente manera:

Definición de la tabla 

Selección	anio=2015
Fila	sexo
Columna	
Control	
Sumarización	
Contenido	Población

Por otra parte, si se desea conocer el número de habitantes por sexo, para cada año en el periodo 2000-2015, la tabla puede definirse de la siguiente manera. Nótese en este ejemplo la definición de rangos para la variable año, la tal como se describió en la sección 2.4.

Definición de la tabla 

Selección	
Fila	anio:2000..2015
Columna	sexo
Control	
Sumarización	
Contenido	Población

NÚMERO PROMEDIO DE HIJOS ENTRE LAS MUJERES, SEGÚN ZONA DE RESIDENCIA

Para obtener esta información, se utilizará la base de datos del Censo de Población del año 2011. Al acceder al Sistema de Consultas ingresando el código de usuario y contraseña, los procedimientos que deben seguirse son los siguientes,


1. Desde la página *Selección de la base de datos* elija la base de datos “Censo de Población y Vivienda 2011” y presione a continuación el botón *Procesar* al final de esta página.
2. Obtenga una frecuencia simple de las variables que se van a analizar, estas son: sexo (P02_SEXO), zona (ID_ZONA) y número de hijos nacidos vivos que han tenidos las mujeres (P30_HIJOS_NACIDOS_VIVOS). Esto le permitirá conocer los posibles valores que toman estas variables y así identificar cualquier particularidad de los datos.

3. Para obtener la frecuencia simple de estas variables, basta con seleccionar tan solo una de ellas en la casilla *Fila* y presionar el botón *Procesar*. A continuación se abrirá una nueva pestaña en su explorador, donde se mostrará una tabla con los resultados. Restablezca el *Asistente para la construcción de las tabla*, presionando el botón *Limpiar*, cada vez que haya obtenido los resultados para una nueva variable.

Al analizar el número de hijos nacidos vivos, puede observarse un gran número de casos que indican que tienen 0 (cero) hijos. Si observa la boleta censal, podrá darse cuenta que el número de hijos tan solo se le pregunta a las mujeres de 12 años o más. Note que el código 0 (cero) se ha empleado como indicador de la condición "No aplica" el registro de ésta información cuando se trata de mujeres menores de 12 años y de la población masculina.

Note también que el código 99 corresponde a "Ignorado", que se asigna a todos aquellos casos para los que no se pudo obtener una respuesta.

4. En vista de esto, es preciso filtrar los datos, para que los resultados brinden información únicamente para las mujeres mayores de 12 años. Además, es importante omitir los casos para los cuáles se desconoce el número de hijos tenidos. Esto se logra escribiendo la siguiente expresión en la casilla *Selección*: P02_SEXO=2 && P03_EDAD>=12 && P30_HIJOS_NACIDOS_VIVOS<99
5. Finalmente, seleccione la variable zona de residencia en la casilla *Fila* y la variable de hijos nacidos vivos en la celda de *Sumarización*. La definición de la tabla se verá de la siguiente manera:

Definición de la tabla 

Selección	P02_SEXO=2 && P03_EDAD>=12 && P
Fila	ID_ZONA
Columna	
Control	
Sumarización	P30_HIJOS_NACIDOS_VIVOS
Ponderación	

Los resultados de esta tabla muestran, en primer lugar, el número de casos a partir de los cuáles se estimó el número promedio de hijos de las mujeres, según la zona en la que estas residían. El Censo de Población y Vivienda del año 2011 registró a 1 250 330 mujeres mayores de 12 años residentes en la zona urbana, y a 415 431, en la zona rural.

Los resultados generados permiten determinar que en promedio, las mujeres residentes de la zona rural del país, tienen más hijos que las mujeres que residen en la zona urbana. Las cifras son de 2.49 y de 2.06, respectivamente.

Resultados del PDQ Explore
Censos Nacionales de Población y Vivienda de Costa Rica 2011

Vivienda: p02 sexo igual a 2 mujer y p03 edad mayor o igual que 12 y p30 hijas e hijos nacidos vivos menor que 99 ignorado según zona




ID_ZONA	Frequency	Mean	Std. Dev.	Low	High
Urbano	1 250 330	2,06	2,24	0	30
Rural	415 431	2,49	2,73	0	30

Las últimas dos columnas de esta tabla muestran los valores mínimo y máximo de la variable "Número de hijos nacidos vivos". Esto quiere decir, que tanto en la zona urbana, como en la rural, se encontraron mujeres que no tenían ningún hijo, y que la(s) mujer(es) con el mayor número de hijos, indicaron haber tenido 30 hijos.

ÍNDICE DE MASCULINIDAD O DE RELACIÓN ENTRE SEXOS

El índice de masculinidad, también llamado índice de relación entre sexos, es uno de los principales indicadores para analizar la distribución por sexo de la población. Se estima como la razón de varones entre mujeres, y generalmente, se multiplica esta medida por cien para facilitar su interpretación. Este indicador se puede obtener de una manera sencilla, pues para su cálculo basta con obtener el número de mujeres y el número de hombres en la población de interés.

Definición de la tabla 

Selección	<input type="text"/>
Fila	P02_SEXO
Columna	<input type="text"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

1. Seleccione la base de datos "Censo de Población y Vivienda 2011" desde la página *Selección de la base de datos* y presione el botón *Procesar*, que se encuentra al final de ésta.
2. En el *Asistente para la definición de la tabla*, presione el botón *Fila*, que automáticamente lo llevará al *Asistente para la construcción de expresiones*.

3. Desde el menú *Variables*, en la sección izquierda de la ventana, seleccione la variable sexo (P02_SEXO) presionando doble clic sobre ésta. Note que la variable aparecerá automáticamente en el recuadro que se muestra en la parte inferior de la ventana.
4. Finalmente, presione el botón *Aceptar*, para volver al *Asistente para la definición de la tabla*.
5. Como se observa en la figura adjunta, el número de hombres y de mujeres se obtiene ingresando únicamente la variable sexo (P02_SEXO). No se requiere ninguna otra instrucción, y bastará con presionar a continuación el botón *Procesar*. Los datos necesarios para estimar el indicador se mostrarán en una nueva ventana del explorador web.

De esta manera, es posible determinar que en Costa Rica en el año 2011 había 2 106 063 hombres y 2 195 649 mujeres. El índice de masculinidad se calcula de la siguiente manera:

$$\text{Índice de relación entre sexos: } \frac{N^{\circ} \text{ de hombres}}{N^{\circ} \text{ de mujeres}} \times 100$$


$$\text{Índice de relación entre sexos, 2011} = \frac{2\ 106\ 063}{2\ 195\ 649} \times 100 = 95.9$$

Los resultados expresan que, en el año 2011, en el país habitaban alrededor de 96 varones por cada 100 mujeres.

NÚMERO DE HABITANTES, POR EDAD Y SEXO

La distribución por sexo y edad de la población puede obtenerse de la siguiente manera. Esta información es muy útil, pues permite generar las conocidas *Pirámides de Población*. En este ejemplo se ilustra el caso para la población costarricense, empleando la información del Censo de Población y Vivienda del año 2011.


1. Seleccione la base de datos “Censo de Población y Vivienda 2011” desde la página *Selección de la base de datos* y presione el botón *Procesar*, que se encuentra al final de esta página.
2. Seleccione la variable edad (P03_EDAD) en la casilla *Fila*. Tras presionar este botón, se abrirá el *Asistente para la construcción de expresiones*. Busque la variable edad en el menú *Variables*, ubicado del lado izquierdo de la ventana, y selecciónela presionando doble clic sobre ella. A continuación presione el botón *Aceptar* para volver al *Asistente para la definición de la tabla*.
3. Presione el botón *Columna* y seleccione la variable sexo (P02_SEXO) siguiendo el mismo procedimiento descrito en el punto anterior.
4. Presionar el botón *Procesar* para que se desplieguen los resultados. Estos se mostraran en una nueva ventana de su explorador web.
5. La definición de la tabla debe verse como se muestra en la siguiente figura.

Definición de la tabla 

Selección	<input type="text"/>
Fila	P03_EDAD
Columna	P02_SEXO
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

Estas instrucciones generarán una tabla con múltiples filas, una por cada edad simple. Esta variable mostrará valores desde 0 hasta 112; el valor cero representa a los niños menores de un año y el valor 112 corresponde a la persona de mayor edad registrada en Censo del año 2011.

En muchas ocasiones, quizá sea necesario obtener la información por grupos quinquenales de edad. Para obtener este dato de manera abreviada, sin la necesidad de utilizar el proceso de recodificación, basta con escribir la instrucción `P03_EDAD/5` en las casillas *Fila* o *Columna*, como se describió en la sección 2.8. En este ejemplo, la tabla se puede definir como se muestra a continuación:

Definición de la tabla 


Selección	<input type="text"/>
Fila	P03_EDAD/5
Columna	P02_SEXO
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

PORCENTAJE DE PERSONAS NACIDAS EN EL EXTRANJERO, POR GRUPOS DE EDAD

1. Seleccione la base de datos del Censo de Población desde la página *Selección de la base de datos* y presione el botón *Procesar*, al final de esta, para ser redirigido al *Asistente para la definición de la tabla*.
2. Seleccione la variable “Edad” (P03_EDAD) en la casilla *Fila*. Al presionar este botón se abrirá el *Asistente para la construcción de expresiones*. Busque la variable “Edad” en el menú *Variables* y selecciónela dando doble clic sobre ella. Para obtener la edad por grupos quinquenales modifique

la expresión que se muestra en el recuadro inferior de la ventana, escribiendo en su lugar P03_EDAD/5. Puede utilizar los botones que se muestran en pantalla o el teclado de su computadora para escribir la expresión. Finalmente presione el botón Aceptar, para volver a la definición de la tabla.

3. Seleccione en la casilla *Columna* la variable Lugar de Nacimiento (P05A_LUGAR_NACIMIENTO), siguiendo un procedimiento análogo al descrito en el punto anterior. Esta variable permite identificar si una persona nació en el mismo cantón en el cuál fue censado, en otro diferente al actual cantón de residencia, o bien, si el entrevistado nació en otro país. Las primeras dos categorías permiten identificar a las personas nacidas en Costa Rica mientras que la tercera identifica a quienes nacieron fuera del territorio nacional. Esta última es la que se desea analizar.
4. Tras seleccionar estas dos variables, deberá presionar el botón *Procesar* para generar los resultados, que se mostrarán en una nueva ventana.

Definición de la tabla 

Selección	<input type="text"/>
Fila	P03_EDAD/5
Columna	P05A_LUGAR_NACIMIENTO
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

5. Al pie del cuadro generado, se muestra el menú *Porcentajes*, mediante el cual se podrá seleccionar la modalidad bajo la que se desean estimar los porcentajes. En este caso se seleccionara la opción "*Porcentajes por columna*" que permite identificar la proporción de personas en un determinado grupo de edad, con respecto al total de personas en una población.

Si se desea obtener el porcentaje de personas nacidas en un país en específico, por ejemplo en Estados Unidos o Nicaragua, deberá seleccionarse la variable Código de país (P05C_CODIGO_PAIS) que identifica, para cada persona nacida en el extranjero, el país en el que nació. Note que para esta variable, el código 0 identifica a las personas nacidas en Costa Rica.

Para conocer la distribución por edad únicamente de las personas nacidas en Nicaragua, por ejemplo, se puede definir la tabla como se muestra en la siguiente figura. Utilizando la opción *Selección*, los casos fueron filtrados de acuerdo a la variable Código de país, de tal manera que la tabla generada muestre únicamente la información de las personas de esa nacionalidad (código 558).

Definición de la tabla ?

Selección	<input type="text" value="P05C_CODIGO_PAIS=558"/>
Fila	<input type="text" value="P03_EDAD/5"/>
Columna	<input type="text"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

RELACIÓN DE DEPENDENCIA DEMOGRÁFICA

La relación de dependencia muestra la relación entre la población económicamente inactiva o dependiente y la población en edades productivas. Es un indicador de dependencia económica potencial, ya que no todas las personas menores de 15 años o mayores de 65 están fuera del mercado laboral, y no todas las personas entre los 15 y 65 años de edad se encuentran activas.

Representa el cociente entre el total de la población de 0 a 14 años y la población de 65 y más, dividido entre el total de la población entre los 15 a 64 años de edad, comúnmente este cociente es multiplicado por 100 para facilitar la interpretación de los resultados.

1. Este ejemplo se desarrollará con los datos del Censo de Población y Vivienda 2011. Seleccione esta base de datos desde la página *Selección de la base de datos*. Tras presionar el botón *Procesar*, será redireccionado al *Asistente para la definición de la tabla*.
2. Seleccione la variable edad (P03_EDAD) en la casilla *Fila* y presione el botón *Procesar*. Tras esta acción se abrirá una nueva ventana en el explorador web, que mostrará un cuadro con la distribución por edad de toda la población. Esto le permitirá ver con detalle el rango de edades de la población costarricense, que va desde 0 hasta 112.
3. El siguiente paso será recodificar la variable edad, de tal manera que los resultados contabilicen el número de habitantes, pero clasificados en dos categorías: "Población en edades económicamente inactivas" y "Población en edades económicamente activas". Esta recodificación simplificará el proceso para generar el indicador deseado.
4. Vuelva al *Asistente para la definición de la tabla* y presione el botón *Limpiar* para restablecer el visor.
5. Seleccione nuevamente la variable edad en la casilla *Fila*.
6. Presione el botón *Recodificar*, que se ubica en el margen inferior derecho del *Asistente para la construcción de expresiones*.

7. Complete el cuadro para la *Preparación de Recodificaciones* como se muestra en la figura adjunta. Para ello deberá completar la información de las casillas *Desde* y *Hasta*, con los valores mínimo y máximo que tomará cada categoría, respectivamente. Deberá también asignarle a este rango de valores un nuevo código en la casilla *Vale*. Puede ingresar una breve descripción de esta nueva categoría, en la celda *Etiqueta*.

====Complete la recodificación requerida====

R: Fila

r: 0 14 1 Inactivo

r: 65 112 1 Inactivo

r: 15 64 2 Activo

Desde: 15 Hasta: 64 Vale: 2 Etiqueta: Activo


Incluir Excluir Cambiar Aceptar Cancelar

Como la población en edades económicamente inactivas no corresponde a una numeración continua, pues comprende a las personas menores a los 15 años y a aquellas de 65 años o más, se creará la categoría "Inactivo" en dos fases. Note que las primeras dos líneas en la figura adjunta, se refieren a una única nueva categoría; ambas con el código 1. El Sistema de Consultas de manera automática desplegará los resultados sumando a todas las personas en esos grupos de edades, en una única categoría.

Si se cometió algún error durante el proceso, cualquiera de las líneas podrá modificarse. Para ello deberá posicionarse sobre ella con el cursor y realizar los cambios correspondientes en las celdas *Desde*, *Hasta*, *Vale* o *Etiqueta* y presionar el botón *Cambiar* una vez que se haya realizado la modificación deseada. Podrá también eliminar una línea, tras seleccionarla y presionar el botón *Excluir*.

8. Tras completar las instrucciones presione el botón *Aceptar*. Note que ahora no se verá el nombre de la variable edad en la casilla sino el término "*Fila*". *Este* nuevo nombre se genera de manera automática e indica que esta variable ha sido recodificada.
9. Presione el botón *Aceptar* en el *Asistente para la construcción de expresiones*.
10. Finalmente, presione el botón *Procesar*, para obtener los resultados que le permitirán estimar el indicador. Deberá obtener resultados como los que se muestran en la figura adjunta.

Resultados del PDQ Explore
Censos Nacionales de Población y Vivienda de Costa Rica 2011


Exportar Tabla Ver

	Total
1: Inactivo	1 379 542
2: Activo	2 922 170
Total	4 301 712

Porcentajes:

11. Con esta información, la relación de dependencia demográfica se estima de esta manera:

$$\text{Relación de dependencia: } \frac{\text{Población de 0 a 14 años} + \text{Población de 65 años o más}}{\text{Población de 15 a 64 años}} \times 100$$

$$\text{Relación de dependencia: } \frac{1\,379\,542}{2\,922\,170} \times 100 = 47.2$$


Este valor de 47.2 indica que, en el año 2011, por cada 100 personas en edades productivas, en la población hay alrededor de 47 personas en edades económicamente inactivas.

NÚMERO PROMEDIO DE PERSONAS POR VIVIENDA INDEPENDIENTE

Para obtener el número promedio de personas por vivienda independiente, es necesario estimar el número de personas que residen en este tipo de vivienda y dividir esta cifra entre el número total de viviendas independientes. Las acciones a seguir se describen enseguida,


1. Seleccione la base de datos del Censo de Población y Vivienda 2011, desde la página *Selección de la base de datos*. Presione el botón *Procesar* que encontrará al fondo de esta página para ser redirigido al *Asistente para la definición de la tabla*.
2. Presione el botón *Fila* para seleccionar la variable “Tipo de vivienda” (V01_TIPO_VIVIENDA). Búsquela en el menú *Variables*, en la sección izquierda del *Asistente para la construcción de expresiones*, y selecciónela presionando doble clic sobre ella. Presione a continuación el botón *Aceptar*, para volver a la definición de la tabla.

- En la casilla *Ponderación*, seleccione la variable número de personas que residen en el hogar (H05C_TOTAL_PERSONAS). La opción de ponderación, permite multiplicar cada registro por la frecuencia de la variable seleccionada, en este caso contará tantas viviendas como personas habiten en ella.

Definición de la tabla 

Selección	<input type="text"/>
Fila	V01_TIPO_VIVIENDA
Columna	<input type="text"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	H05C_TOTAL_PERSONAS

- A continuación, presione el botón *Procesar* para obtener la primera tabla de resultados. La tabla generada presentará el número de personas que viven en cada tipo de vivienda. Nótese que el total de casos es de 4 301 712, el número total de habitantes registrados en el país.
- El número de viviendas independientes, se obtiene simplemente al tabular la variable tipo de vivienda (V01_TIPO_VIVIENDA). Note que al no ponderarse los resultados por el número de habitantes en una vivienda, los resultados que despliega esta tabla se refieren únicamente a viviendas.

Definición de la tabla 

Selección	<input type="text"/>
Fila	V01_TIPO_VIVIENDA
Columna	<input type="text"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

Con estos datos, podemos calcular el número promedio de personas por vivienda independiente:

$$N^{\circ} \text{ promedio de personas en viviendas independientes: } \frac{4\ 044\ 239}{1\ 250\ 183} = 3.2$$

Esto equivale a decir que en Costa Rica en el año 2011 vivían en promedio 3.2 personas en cada vivienda independiente. En algunas ocasiones, para generar este indicador también se consideran los residentes y las viviendas independientes ubicadas en condominios.

TASA BRUTA DE MORTALIDAD EN EL PERIODO 2000-2010

Para obtener la tasa bruta de mortalidad (TBM) será necesario trabajar tanto con la base de datos del Registro de defunciones, como con la base de datos de las Proyecciones Nacionales de Población. Esta tasa se obtiene a como el cociente del número de defunciones ocurridas a lo largo de un año y el número de habitantes a mitad de ese mismo año. La base de datos de Proyecciones de población, ya ofrece la estimación de la población a mitad de año, por lo que no es necesario realizar ningún ajuste.

1. Seleccione en primer lugar la base de datos “Defunciones 1970 – 2013”, desde la página *Selección de la base de datos* y presione el botón *Procesar* al final de esta página.
2. Automáticamente será dirigido al *Asistente para la definición de la tabla*. Desde allí, presione el botón *Selección* y restrinja el rango de valores que desplegará la variable año de trabajo para el periodo entre 2000 y 2010. Esta variable identifica el año en que se registró cada defunción.
3. Busque esta variable (aatraba) en el menú *Variables* presionando doble clic sobre ella. Deberá modificar la expresión que se muestra en el recuadro inferior de la ventana. Escriba en su lugar `aatraba>=2000 && atraba<=2010`. Para construir la expresión puede valerse de los botones disponibles en el *Asistente de la construcción de expresiones*, o bien teclear esta información directamente en la ventana. El término && se consigue al presionar el botón *AND*. Presione el botón *Aceptar* para volver a la definición de la tabla.
4. Seleccione en la casilla *Fila*, la variable año de trabajo, para que los resultados se desplieguen anualmente. La definición de la tabla deberá ser igual a la que se muestra en la figura,

Definición de la tabla

Selección: `aatraba>=2000 && atraba<=2010`

Fila: `aatraba`

Columna:

Control:

Sumarización:

Ponderación:


Procesar Limpiar Diccionario Ejemplo Ayuda

Cambiar Base de Datos

País: Costa Rica Base de Datos: Seleccione

5. Presione *Procesar* para obtener la tabla de resultados. Puede descargarlos en un archivo Excel y utilizarlos posteriormente.
6. A continuación, deberá obtener el número de habitantes en cada año. Observe el menú *Cambiar base de datos* en la sección inferior del *Asistente para la definición de la tabla*. En la opción *País*, seleccione “Costa Rica” y en *Base de datos* seleccione “Proyecciones Nacionales de Población”. Automáticamente habrá cambiado de base de datos y podrá empezar a trabajar con ella.

- El procedimiento que deberá seguir ahora, es análogo al utilizado para generar el número de defunciones, descrito en el punto 3. La definición de esta tabla se muestra a continuación, en este caso la variable año de trabajo posee otro nombre (anio) pero su rango también se restringirá empleando la opción *Selección*.

Definición de la tabla 

Selección	anio>=2000 &&anio<=2010
Fila	anio
Columna	
Control	
Sumarización	
Contenido	Población ▼

Procesar Limpiar Diccionario Ejemplo Ayuda

- Tras presionar el botón *Procesar*, se abrirá una nueva ventana mostrando una tabla con el número anual de habitantes en el periodo 2000-2010.

La estimación de la tasa bruta de mortalidad para cada año se obtiene dividiendo ambas cifras. Por ejemplo, en el año 2000 la TBM es:

$$TBM_{2000} = \frac{\text{Total de defunciones 2000}}{\text{Población 30/06/2000}} \times 1000$$

$$TBM_{2000} = \frac{14\ 944}{3\ 872\ 349} \times 1000 = 3.86$$


Esta cifra indica que en el año 2000, murieron 3.86 personas por cada mil.

TASA DE MORTALIDAD INFANTIL EN LOS AÑOS 2005 Y 2010

La tasa de mortalidad infantil (TMI) se calcula como el cociente entre el número de defunciones de niños menores de un año, entre el total de nacimientos ocurridos en un año dado. Esta cifra suele ser multiplicada por mil para facilitar su interpretación. Para obtener este indicador, es necesario trabajar con la base de datos del Registro de Nacimientos y la Base de datos del Registro de defunciones. Los procedimientos a seguir son los siguientes:



- Como primer paso se obtendrán el número de defunciones de niños menores de un año ocurridas en los años de interés. Seleccione desde la página *Selección de la base de datos* la opción “Defunciones 1970-2013” para abrir esta base de datos. Presione el botón *Procesar* que se ubica al final de esta página. Automáticamente se encontrará en el *Asistente para la definición de la tabla*.

- Presione el botón *Fila*, y construya la siguiente expresión utilizando los botones correspondientes o bien, tecleando la instrucción: `edad=0 && (aatraba=2005 || atraba=2010)`. Los términos `&&` y `||` se obtienen al presionar los botones *AND* y *OR*, respectivamente. La expresión `edad=0` permite seleccionar las defunciones de las personas menores de un año y la expresión `(aatraba=2005 || atraba=2010)` permite seleccionar las defunciones registradas en el año 2005 o bien, aquellas registradas en el año 2010. Los paréntesis se utilizan con el fin de que la restricción establecida para la variable `edad` se aplique en ambos años.
- Seleccione la variable año de trabajo (`aatraba`) en la casilla *Fila*, para poder observar los resultados de cada año seleccionado. La definición de la tabla deberá verse de esta manera:

Definición de la tabla 

Selección	<code>edad=0 && (aatraba=2005 atraba=</code>
Fila	<code>aatraba</code>
Columna	
Control	
Sumarización	
Ponderación	

Cambiar Base de Datos

País:  Base de Datos: 

- Tras presionar el botón *Procesar*, se abrirá una nueva ventana en su navegador web mostrando la tabla de resultados. Podrá determinar que en el año 2005 fallecieron 701 niños menores de un año y en el 2010 fallecieron 671.
- Ahora que se ha identificado el número de defunciones, el siguiente paso es estimar el número de nacimientos ocurridos en esos mismos años. Utilice la opción *Cambiar base de datos*, en la sección inferior del Asistente para la definición de la tabla. En el menú *País* seleccione “Costa Rica” y en el menú *Base de datos* elija “Nacimientos 1972 - 2013”, automáticamente se cambiará la base de datos de trabajo, ahora estará utilizando la base de datos del Registro de Nacimientos. También puede abrir esta base de datos desde la página *Seleccionar base de datos*.
- Establezca una restricción utilizando la opción *Selección*, de tal manera que solo se muestren los resultados para los años 2005 y 2010. Presione este botón y busque la variable Año de trabajo (`aatraba`) en el menú *Variables*, podrá seleccionar esta variable presionando doble clic sobre ella.
- Modifique la expresión que se muestra en la sección inferior de la ventana, utilice los botones disponibles o tecléela directamente. Deberá escribir: `aatraba=2005 || atraba=2010`.
- A continuación presione el botón *Aceptar*, para volver al *Asistente para la definición de la tabla*.

9. Presione el botón *Fila* y seleccione esta casilla la variable año de trabajo. La definición de la tabla deberá ser igual a la que se muestra en la imagen.

Definición de la tabla ?

Selección	aatraba=2005 aatraba=2010
Fila	aatraba
Columna	
Control	
Sumarización	
Ponderación	

Procesar
Limpiar
Diccionario
Ejemplo
Ayuda

10. Finalmente obtenga los resultados presionando el botón *Procesar*. En la tabla de resultados podrá determinar que en el año 2005 nacieron un total de 71 548 niños y en 2010 nacieron 70 922.

La estimación de la tasa de mortalidad infantil en los años 2005 y 2010 se obtiene dividiendo ambas cifras.

$$TMI: \frac{\text{Total de defunciones año } i}{\text{Total de nacimientos año } i} \times 1000$$

Utilizando las cifras obtenidas a lo largo del ejercicio es posible determinar que la TMI en 2005 y 2010 es,

$$TMI_{2005}: \frac{701}{71\,548} \times 1000 = 9.8 \qquad TMI_{2010}: \frac{671}{70\,922} \times 1000 = 9.5$$

La TMI indica que en el año 2005 fallecieron 9.8 niños menores de un año, por cada mil nacimientos ocurridos en ese año. En el año 2010 esta cifra fue de 9.5.

TASA DE FECUNDIDAD POR GRUPOS QUINQUENALES DE EDAD, EN 2010

Las tasas específicas de fecundidad se obtienen del cociente entre el número de nacimientos de mujeres con edad cumplida entre “*i*” y “*i+5*”, ocurridas en el periodo de interés, y el número de mujeres en ese mismo rango de edad. Para estimar este indicador será necesario utilizar la base de datos del Registro de Nacimientos y la base de datos de Proyecciones Nacionales de Población. Los pasos a seguir son los siguientes,

1. Seleccione la base de datos con la información de nacimientos. Busque la opción “Nacimientos 1972 - 2013” en la página *Seleccionar base de datos* y presione el botón *Procesar* que se encuentra al final de ésta. Automáticamente será redirigido al *Asistente para la definición de la tabla*.

- Presione el botón *Selección* para establecer una restricción que limite el conteo para los nacimientos ocurridos en el año 2010, solamente. Esta acción lo llevará al *Asistente para la construcción de expresiones*, desde donde deberá construir la expresión `aatraba=2010`, ya sea utilizando los botones o tecleándola directamente. Al finalizar, presione el botón *Aceptar* para volver nuevamente a la definición de la tabla.
- En la casilla *Fila*, recodifique la variable edad de la madre en grupos quinquenales. Presione el botón y busque la variable Edad de la madre (`edmadre`) en el menú *Variables*, selecciónela presionando doble clic sobre ella.
- Modifique la expresión que se muestra en la ventana inferior del *Asistente para la construcción de expresiones*. Escriba en su lugar `edmadre/5` lo que le permitirá recodificar la variable de manera abreviada esta variable, tal como discutió en la sección 2.8. Al finalizar presione el botón *Aceptar*.
- La formulación de la consulta deberá ser igual a la que se muestra en la figura siguiente,


Definición de la tabla ?

Selección	<input type="text" value="aatraba=2010"/>
Fila	<input type="text" value="edmadre/5"/>
Columna	<input type="text"/>
Control	<input type="text"/>
Sumarización	<input type="text"/>
Ponderación	<input type="text"/>

- Presione el botón *Procesar*. Automáticamente se abrirá una nueva ventana en el navegador web, mostrando la tabla de resultados.
- Puede descargar estos resultados a un archivo Excel, presionando el botón *Exportar tabla*, ubicado en la sección superior derecha de la ventana. Para estimar este indicador, se suele utilizar la información sobre mujeres entre los 15 y 49 años, por lo que podrá obviar la información de los otros grupos de edad. En la categoría "0: edmadre:0..4" encontrará un total de 206 nacimientos. Esta información no corresponde a nacimientos de mujeres en ese rango de edad, sino que contabiliza los casos para los que se ignora la edad de la madre, identificados por el código 0.
- A continuación deberá estimar el número de mujeres entre los 15 y 49 años de edad, para lo que deberá utilizar la base de datos de las Proyecciones Nacionales de Población. Utilice la opción *Cambiar base de datos*, en la sección inferior del *Asistente para la definición de la tabla*, para utilizar esta base de datos. Seleccione "Costa Rica" en el menú *País* y "Proyecciones Nacionales de Población" en el menú *Base de datos*.
- Presione el botón *Selección* y establezca la restricción: `anio=2010 && sexo=2`. Esta le permitirá contabilizar únicamente el número de mujeres, en el año 2010.

10. Seleccione la variable edad en grupos quinquenales (edadq) en la casilla *Fila*. Esta variable ya ha sido creada y se encuentra disponible en la base de datos, por lo que no deberá realizar ningún procedimiento adicional.

11. La definición de la tabla deberá verse como se presenta en la figura.

Definición de la tabla 

Selección	anio=2010&&sexo=2
Fila	edadq
Columna	
Control	
Sumarización	
Contenido	Población

12. Presione el botón *Procesar* para obtener la tabla de resultados. Nuevamente puede obviar las cifras referentes a las mujeres en los grupos de edad que no son de interés.

Las tasas específicas de fecundidad por grupos de edad se estiman de la siguiente manera,

$$TF_{i;i+4} = \frac{\text{Total de nacimientos edades } i ; i + 4}{\text{Total de mujeres en edades } i ; i + 4} \times 1000$$

Así por ejemplo, la TF para el grupo de mujeres entre los 15 y 19 años de edad se obtiene de esta forma,

$$TF_{15-19} = \frac{12\ 828}{204\ 402} \times 1000 = 62.8$$

Esta cifra indica que por cada mil mujeres entre los 15 y 19 años de edad, ocurrieron 62.8 nacimientos. La estimación de las restantes tasas de fecundidad se genera de manera equivalente, las cifras obtenidas entre las mujeres de 20 a 24 años, y entre aquellas de 40 a 44, son respectivamente 99.3 y 8.4.

TASA GLOBAL DE FECUNDIDAD, EN EL AÑO 2010

La tasa global de fecundidad (TGF) es otro de los indicadores comúnmente utilizados en demografía. La TGF puede estimarse a partir de las tasas específicas de fecundidad por grupos quinquenales de edad que se desarrollaron en el punto anterior.

El indicador se obtiene multiplicando por cinco la suma de las tasas específicas de fecundidad. Para el año 2010, este indicador se estima entonces de la siguiente manera. Considere que las tasas específicas que

no están expresadas en miles.

$$TGF_{2010} = 0.360 * 5 = 1.8$$

La TGF representa el número promedio de hijos que tendría una mujer de una cohorte hipotética de mujeres que durante su vida fértil tuviera a sus hijos de acuerdo a las tasas de fecundidad por edad de un determinado período de estudio y que no estuviera expuesta al riesgo de mortalidad desde el nacimiento hasta el término del período fértil.

¿Cómo obtener ayuda?

Si desea hacer alguna consulta, o simplemente brindar sus opiniones o recomendaciones sobre este servicio, por favor envíe un e-mail a: censos@ccp.ucr.ac.cr

Elaboración del manual:
Alejandra González, CCP-UCR

Revisión del contenido:
Daniel Antich y Alejandra Escalante, CCP-UCR



Centro Centroamericano de Población
Universidad de Costa Rica
Abril 2015