

USO Y PROCESAMIENTO DE DATOS CENSALES PARA LA GENERACIÓN DE INDICADORES SOCIODEMOGRÁFICOS UTILIZANDO REDATAM

Buenos Aires, Argentina 24-27 julio 2018

Prof. Alejandra Silva
CELADE, División de Población
CEPAL

Temario:

Día 1 A: TEMA: PRESENTACION DE LA BASE DE DATOS CENSAL, TRABAJANDO CON LAS VARIABLES y EL CUESTIONARIO

- **OBJETIVOS**
- Entender la base de datos del censo y/o encuestas
- Analizar la potencialidad de la herramienta.
- Entender el concepto de estructura jerárquica
- Manejo de base de datos
- Entender la estructura geográfica.

El censo de población y vivienda
Como se crea una base de datos a partir del censo
Características de las variables
Como estructuro la geografía de un país en forma jerárquica
Características de Redatam y sus módulos

El censo de población y vivienda (aspecto operativo)

El método utilizado por los Censos de Población y Vivienda para obtener la información es la localización de las viviendas, en las que se identifica a los hogares y personas que las componen. Para cada hogar que compone la vivienda se dispone de un cuestionario censal donde el empadronador registrará los datos de la vivienda a la que pertenece el hogar; datos del hogar en cuestión e información para cada una de las personas que lo integran, todo esto a partir de las respuestas entregadas por el jefe de hogar o, en su defecto, otra persona calificada para responder a las preguntas que se le formulen.

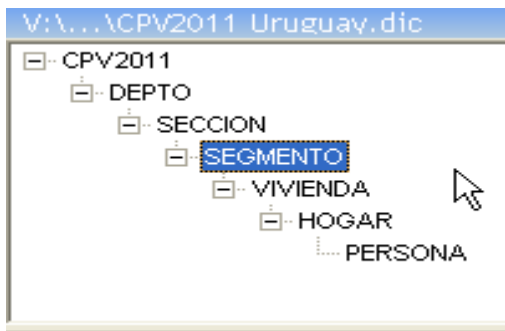
La información censal es relevada considerando a todos quienes pernoctaron en el hogar la noche anterior al censo (Censo de hecho) o a los residentes habituales (Censos de Derecho), modalidad, esta última, que permite extender el periodo de levantamiento censal y que corresponde a la adoptada en el Censo demográfico de 2011.

Código geográfico

Una información de vital importancia para la creación de la base de datos numérica que se construirá después de recolectar y validar toda la información del país y para cada una de

sus divisiones político administrativas (Departamentos) y las que corresponden a la organización operativa del evento censal (Secciones y Segmentos censales), es la que da cuenta de la localización en un espacio geográfico determinado, de cada una de las viviendas que fueron objeto de la aplicación del cuestionario. De esta manera es posible, que una vez construida la base de datos censal, se pueda recuperar la información recogida hasta el nivel más bajo de desagregación geográfica, que en el caso de la base censal de Uruguay 2011 corresponde al Segmento Censal. Por razones de ley y de protección del secreto estadístico que asegura la inviolabilidad de la intimidad de las personas, la base de datos no permite la individualización de una vivienda, hogar y personas en particular. Esta información compuesta de códigos numéricos es la que da origen al **código geográfico** que, según se verá más adelante, posee una relevancia especial al momento de trabajar con dicha base de dato

Estructura de la base de datos censal



Después de los procesos de validación pertinentes, la información se organiza de manera jerárquica dando origen a una estructura como la que se ilustra en la figura de la izquierda.

“CPV2011” es el nombre que define la base de datos en REDATAM y representa a la entidad padre o “país”. Esta contiene todas la entidades de menor nivel jerárquico desde “Departamento” hasta “persona”. Ahora bien, dada esta relación de dependencia entre una entidad y otra, las variables que definen características de una entidad determinada, define al mismo tiempo las características de todos los elementos que se encuentran en un nivel de jerarquía menor. Por ejemplo, la variable “VIVIENDA.DV05” pertenece a la entidad “VIVIENDA” y los atributos que respecto a esta variable posea una vivienda o conjunto de viviendas en particular son también atributos que describen características de los elementos que le pertenecen. Si la vivienda posee agua proveniente de la red general de agua potable, entonces todos los hogares y personas que pertenecen a dicha vivienda comparten el mismo origen del agua declarado para la vivienda.

REDATAM hace un uso total de las propiedades que presenta la estructura jerárquica del Censo de Población y Vivienda, posibilitando por ejemplo determinar la distribución de la población según sexo, de acuerdo a características de la vivienda o del hogar, de tal manera que es muy rápido y sencillo conocer cuántos hombres y mujeres habitan en viviendas cuyo origen del agua que usan para beber corresponda a “pozo surgente no protegido”. El aprovechamiento de esta particular organización de los datos censales con fines de análisis estadísticos depende del conocimiento y manejo que los usuarios tengan de las variables y categorías que forman parte de cada entidad ubicada en la estructura jerárquica, de ahí que antes de continuar con el conocimiento de REDATAM resulte beneficioso detenerse un momento y “navegar” por la base de datos del Censo de 2011 ayudados por la boleta o cuestionario censal.

Ejercicios

Revise en el cuestionario censal las preguntas sobre Actividad Laboral. A quienes se pregunta? Cuantas posibles respuestas tiene la pregunta? Cuales son las categorías? Cómo podríamos determinar la población ocupada?

Día 1 B: TEMA: REDATAM COMO UNA HERRAMIENTA DE PROCESAMIENTO ESTADÍSTICO

- **OBJETIVOS**
- Adquirir la capacidad de analizar los datos censales a partir de tabulados : producir frecuencias, tablas cruzadas y promedios en forma simple
- Entender el uso básico de un programa en REDATAM+SP.
- Poder ejecutar programas simples

Tabulados fáciles en Redatam
Selecciones geográficas
Programación de procesos en Redatam
Comando TABLE

Lo primero que usted debe realizar en una sesión es abrir un proyecto (archivo .prjx) el cual automáticamente abre el diccionario (archivo .dicx) o metadata de la base de datos. Se abre una ventana con la estructura jerárquica de la base de datos y las variables.

Navegar a través de REDATAM para obtener tabulados fáciles

1. Teniendo activa la ventana de inicio de REDATAM, aún sin haber abierto un diccionario, el sistema ofrece una barra de menús y funciones relacionadas con su sesión de trabajo:
 - Icono para abrir un diccionario
 - Icono para abrir ventana de preferencias
 - Icono de ayuda
 - Menú archivo
 - Menú diccionario
 - Proceso de Estadístico
 - Menú de Herramientas
 - Menú de ventana
 - Menú de ayuda
2. Activar cada icono y desplegar las distintas opciones de menús haciendo un breve comentario de cada opción (No se requiere entrar en detalles por cuanto ello será abordado en los módulos subsiguientes)
3. Abrir el diccionario de la base de datos
4. Desplegar las opciones de cada menú haciendo un breve comentario sobre ellas.

ARCHIVOS QUE SE MANEJAN EN REDATAM

- **PROGRAMAS:** Conjunto de instrucciones que el usuario a empleado con el fin “pedirle” al sistema la obtención de estadísticas provenientes de la base de datos. Estos pueden ser escritos directamente por el usuario en el lenguaje exclusivo de REDATAM (formato libre), o ser generados por REDATAM utilizando el asistente de programas. Los programas se guardan en archivos con la extensión .spc.
- **SELECCIONES:** Para la ejecución de cualquier programa u obtención de tabulados, REDATAM requiere que el usuario defina, previamente, el área geográfica a ser procesada. Esta área puede corresponder a todo el país o a cualquier agrupación de entidades geográficas seleccionables, como ser Departamento, Secciones o Segmentos censales; que corresponde al nivel más desagregado para el procesamiento de la información censal a partir de la estructura de la base. De manera predeterminada, REDATAM tiene seleccionada toda la base de datos con la que se trabaje, el usuario puede hacer referencia a esta selección con la palabra clave “ALL”.

El usuario tiene la posibilidad de guardar las selecciones que realice durante su trabajo, en cuyo caso REDATAM genera un archivo con la extensión “.slw” o “.sel” el cual se almacena donde el usuario defina, con el fin de que puedan ser utilizados cuando se requieran y sin necesidad de repetir el mismo proceso de selección.

- **COMPOSICIÓN DE MAPAS:** Aquí el sistema guarda todas las indicaciones que le entrega el usuario con el fin de utilizar la capacidad de REDATAM para vincular o empalmar la información de la base de datos con un mapa, permitiendo la obtención de mapas temáticos que faciliten la focalización de sectores o áreas geográficas que presenten características de interés para el investigador. La composición de mapas se guarda en un archivo con la extensión .mxp.
- **BASE DE DATOS:** Compuesta por varios archivos, primero los punteros (ptr) que definen la pertenencia geográfica de los registros en toda la base jerárquica y luego las variables (rbf) que se almacenan una por una en un archivo vectorial. La base de datos se trabaja a través del diccionario con los metadatos (dicx).
- **PROYECTO:** El proyecto es una ventana que localiza (y ordena según una referencia) todos los archivos antes mencionados, así de manera fácil y rápida uno puede crear nuevos archivos,. Abrir existentes o guardar otros. (prjx).

Día 2: TEMA: GENERACIÓN DE INDICADORES DE EDUCACIÓN Y DEMOGRÁFICOS

• OBJETIVOS

- Adquirir capacidad de producir indicadores básicos
- Comprender los conceptos básicos del procesamiento jerárquico.
- Comprender y adquirir la capacidad para realizar procesos de conteo de elementos.
- Comprender el proceso de generación y adición de nuevas variables al diccionario de la Base de Datos.

Comando DEFINE: Derivando Nuevas Variables – RECODE

Utilizando UNIVERSE y AREABREAK

Promoviendo Información – El Comando COUNT

Condicionar el conteo con un filtro - FOR

Resultados según Listas de Áreas

Haciendo uso del editor de programas de REDATAM avanzamos en la obtención de los siguientes indicadores:

- 1. Índice de masculinidad:** Consiste en la razón de varones entre mujeres en una determinada población o sub-población. Se considera como el principal indicador, para analizar la distribución por sexo en la población.
- 2. Proporción de masculinidad:** Como su nombre lo indica, corresponde a la proporción de varones entre la población total. Al igual que la medida anterior, permite conocer la distribución de la población por sexo.
- 3. Relación niños-mujer:** Esta relación se utiliza como un indicador aproximado de la fecundidad, especialmente cuando no se disponen datos de calidad sobre los nacimientos. Corresponde a la razón de los niños menores de 5 años, por cada 1.000 mujeres en edad reproductiva (edades cumplidas entre 15 y 49 años).
- 4. Tasa de alfabetismo:** Es un índice del nivel educativo de un pueblo. Aunque se le conoce como tasa, en realidad es una proporción. Se obtiene del cociente de la población alfabetizada de 10 años y más, entre el total de la población de 10 años y más.
- 5. Relación de dependencia:** Esta medida, desde el punto de vista demográfico, se define como el cociente de la población económicamente dependiente, que corresponde a los menores de 15 años y junto con las personas de 65 años o más, sobre la población con edades cumplidas entre 15 y 64 años. Sin embargo, existe otra definición de esta medida, que corresponde al concepto económico del término. En ella, se varían las definiciones de la población económicamente activa (PEA) y de la económicamente dependiente. No obstante, su fórmula de cálculo es similar.
- 6. Relación de reemplazo de la población en edad activa:** Esta medida al igual que la anterior, tiene algunos problemas en la definición de los términos. Generalmente, se considera como la relación entre el grupo de edades entrante a la población activa

(grupo de 15-19) con respecto al grupo saliente de esta población (grupo 60-64). Su valor representa la capacidad de población económicamente activa de reemplazarse.

Día 3: TEMA: GENERACIÓN DE INDICADORES DE SITUACIÓN LABORAL

- **OBJETIVO:** Conocer la tasa de participación laboral y la tasa de desempleo que presenta la población por grupos de edad según características educativas y área de residencia

Comando DEFINE: Utilizando Expresiones Aritméticas (calculo de tasas y %)

Promoviendo Información - El comando SWITCH

Guardar la variable: SAVE

Agregación de un indicador según niveles geográficos

Definir variables para análisis de la situación laboral

Haciendo uso del editor de programas de REDATAM avanzamos en la obtención de los siguientes indicadores:

1.- Tasa de participación laboral en jóvenes de 19 a 29 años:

Definición: Mide la proporción de jóvenes que se encuentran trabajando o buscando trabajo, en relación al total de jóvenes. Proporciona un primer acercamiento respecto a la inserción juvenil en el ámbito laboral. Se calcula obteniendo el cociente entre los jóvenes que se encuentran en la PEA dividido por el total de Jóvenes, por cien.

Opcional desagregación: Total y nivel de instrucción por área Urbano - rural

2.- Tasa de desempleo en jóvenes de 19 a 29 años:

Definición: Mide la proporción de jóvenes que se encuentran sin trabajo. Permite un acercamiento más refinado a la situación laboral que afecta a la población joven. Se calcula como cociente entre los jóvenes desocupados sobre el total de jóvenes en la PEA, por cien.

Opcional desagregación: Total y nivel de instrucción por área Urbano – rural

Día 4: TEMA: CARENCIA DE VIVIENDAS (NBI)

- **OBJETIVO:** Conocer la tasa de participación laboral y la tasa de desempleo que presenta la población por grupos de edad según características educativas y área de residencia

Comando DEFINE: Utilizando Expresiones Aritméticas (calcula de tasas y %)

Promoviendo Información - El comando SWITCH

Guardar la variable: SAVE

Agregación de un indicador según niveles geográficos

Mapeo de los NBI

Haciendo uso del editor de programas de REDATAM avanzamos en la obtención de los siguientes indicadores:

1.- Viviendas en situación de hacinamiento

Definición: Mide el grado de hacinamiento de las viviendas particulares ocupadas. Se calcula como las viviendas con más de 3 personas por dormitorio dividido por el total de viviendas, multiplicado por 100. **Desagregación:** por área Urbano – rural

2.- Mala calidad de Techos, Paredes o Pisos

Definición: Es un indicador de falta de acceso vivienda de calidad; mide la proporción de viviendas construidas con materiales deficientes en sus paredes, techos o pisos. Se calcula como las viviendas con materiales inadecuados dividido por el total de viviendas, multiplicado por 100. **Desagregación:** por área Urbano – rural

3.- Porcentaje de viviendas sin acceso a agua potable:

Definición: Es un indicador de falta de acceso a agua potable; mide la proporción de viviendas que cuentan con abastecimiento de agua inadecuado. Se calcula como las viviendas con abastecimiento de agua inadecuado dividido por el total de viviendas, multiplicado por 100. **Desagregación:** por área Urbano – rural

4.- Porcentaje de viviendas sin acceso a sistema de alcantarillado:

Definición: Es un indicador de falta de acceso a algún sistema de saneamiento; mide la proporción de viviendas sin baño, sin sistema o inadecuado sistema de evacuación de excretas. Se calcula como las viviendas sin acceso a saneamiento dividido por el total de viviendas, multiplicado por 100.

Opcional desagregación: por área Urbano – rural