



[VIGILADA MINEDUCACIÓN]



Epiverse TRACE LAC

INSTRUCTIVO PARA LA LIMPIEZA DE BASES DE DATOS APTAS PARA USO EN LENGUAJE R Y COLOPENDATA

Ing. María Camila Tavera

Supported by :



MRC Unit
The
Gambia



data.org

Funded by:



IDRC • CRDI

Canada

EPIVERSE TRACE-LAC

Investigadores principales

Zulma M. Cucunubá

Catalina González-Uribe

Juan Manuel Cordovéz

Natalia Niño-Machado

Coordinación de equipo Uniandes

Natalia Niño-Machado

Hawyn Díaz

Equipo de desarrolladores Uniandes

Juan D. Umaña

Julian Otero

Juan Montenegro-Torres

José Bocanegra

María Camila Tavera-Cifuentes

Mauricio Santos-Vega

Gerencia de proyecto

Mauricio Alberto Cortés-Cely

Diseño y comunicaciones

Denís Toledo-Nieto

Contenidos:

María Camila Tavera-Cifuentes

Diseño y revisión de textos:

Denís Toledo-Nieto

TRACE-LAC is a collaborative project co-led by Pontificia Universidad Javeriana and Universidad de Los Andes in Colombia for the region. Its aim is to enhance the response to infectious disease epidemics in Latin America and the Caribbean by providing robust data tools to inform decisionmaking effectively. TRACE-LAC is linked with the Epiverse-TRACE initiative of data.org, which is led by the London School of Hygiene and Tropical Medicine (LSHTM) in the United Kingdom, in collaboration with the Medical Research Council - The Gambia. This work is part of the TRACE-LAC research project funded by the International Research Centre (IDRC) Ottawa, Canada. [109848-001-]. The funders had no role in study design, data collection and analysis, decision to publish or preparation of manuscripts, reports or working papers. The views expressed herein do not, however, represent those of IDRC or its Board of Governors.

2024

¿Qué es EPIVERSE?

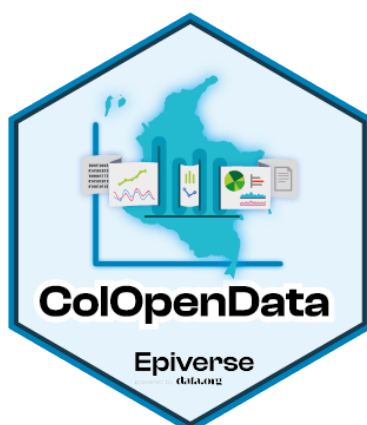
Epiverse es una iniciativa colaborativa global liderada por data.org, que pretende construir un ecosistema de análisis de datos confiable para dar respuesta oportuna a las crisis de salud pública. Esta iniciativa promueve el desarrollo de herramientas analíticas de alta calidad, transparentes y de código abierto, así como la creación de una comunidad activa de usuarios.

¿Qué es TRACE-LAC?

TRACE-LAC* es un proyecto coliderado por la Pontificia Universidad Javeriana y la Universidad de los Andes, que busca fortalecer la respuesta a epidemias por enfermedades infecciosas en América Latina y el Caribe por medio de herramientas de datos que permitan informar de manera adecuada y eficaz la toma de decisiones.

TRACE-LAC está articulado con la iniciativa Epiverse-TRACE que está liderada por el London School of Hygiene and Tropical Medicine (LSHTM) en el Reino Unido en colaboración con el Medical Research Council Unit de Gambia.

¿Qué es ColOpenData?



El equipo de desarrollo de software de TRACE-LAC Uniandes, produjo la herramienta de datos ColOpenData, una librería en lenguaje de programación R que permite consultar de manera fácil y confiable la información disponible en varias bases de datos oficiales de Colombia, como la del Departamento Administrativo Nacional de Estadística - DANE, para estudios de epidemiología y correlacionarlos para su análisis y toma de decisiones.

Esta herramienta facilita la búsqueda de información de varias bases de datos estratégicas simultáneamente y propone formas de manejo y uso.

Permite la consulta y descarga de información socioeconómica, climática y de uso del suelo de diferentes bases de datos oficiales; también la integración de las diferentes variables asociadas a la transmisión de enfermedades según condiciones locales.

<https://epiverse-trace.github.io/ColOpenData>

¿Por qué este instructivo?

Durante la fase de desarrollo de ColOpenData se detectó que los datos que entrega el DANE sobre el censo 2018 tienen un formato que dificulta su manejo con el lenguaje R, por lo cual se requiere adaptar las bases de datos para poder cargar los datos a la herramienta.

Esta tarea no es fácil dada la diversidad y volumen de información que tiene esta base de datos, sin embargo, el equipo TRACE_LAC Uniandes diseñó un proceso que facilita la tarea mediante un código automatizado que modifica la presentación de los datos para que puedan ser manejados con herramientas desarrolladas en R, como ColOpenData, por parte de equipos investigadores que generan evidencia a tomadores de decisión.

Siguiendo los pasos de este instructivo, se optimizan los tiempos del proceso de ajuste respecto a hacerlo manualmente con uso exclusivo de Excel. Por el momento ha sido usado únicamente para modificar los datos del Censo pero es muy posible que el proceso se pueda adecuar a otros datos presentados por el DANE, o incluso se pueda extender a otras fuentes de datos.



1.




IDENTIFICACIÓN DE ELEMENTOS EN EL FORMATO

Lo primero es revisar qué elementos están incluidos en el formato de la hoja para identificar aquellos que deben ser eliminados. Para cada hoja se deben revisar:

a. IMÁGENES


Es necesario revisar si hay imágenes (logos, símbolos, ilustraciones, etc) en el formato y su ubicación. Usualmente estas están en las esquinas superior derecha e izquierda de la hoja, y debajo de ellas se encuentra una barra de colores que es una imagen también. Sin embargo, algunas veces las hojas presentan otras imágenes que deben ser identificadas.

Todas las imágenes, incluyendo las institucionales, deben ser eliminadas, una a una, en cada hoja.


  									
Censo Nacional de Población y Vivienda - CNPV 2018									
PERSONAS (Demográfico)									
Población total censada, por sexo e índices de masculinidad y feminidad, según departamentos, áreas (Total, Cabecera y Centros poblados y Rural disperso) y grupos de edad									
Cuadro 1PD Departamental									
Índice									
2018									
Grupos de edad, departamento y Áreas (Total, Cabecera y Centros poblados y Rural disperso)		Total	Sexo		Índice de Masculinidad		Índice de Feminidad		
			Hombre	Mujer					
Total Nacional	Total	Total	44.164.417	21.570.493	22.593.924	95,5	104,7		
		0 a 4	3.037.781	1.555.605	1.482.176	105,0	95,3		
		5 a 9	3.335.240	1.705.574	1.629.666	104,7	95,5		
		10 a 14	3.610.584	1.848.218	1.762.366	104,9	95,4		
		15 a 19	3.852.255	1.970.530	1.881.725	104,7	95,5		

b. ENCABEZADO

Identifique las filas que hacen parte del encabezado, es decir, las filas que no están directamente relacionadas con la tabla. Estas son un resumen y proveen la misma información que el índice. Usualmente, todas las hojas del mismo archivo ocupan las mismas filas para el encabezado.



DANE
INFORMACIÓN PARA TODOS



El futuro
es de todos

Gobierno de Colombia

Censo Nacional de Población y Vivienda - CNPV 2018

PERSONAS (Demográfico)

Población total censada, por sexo e índices de masculinidad y feminidad, según departamentos, áreas (Total, Cabecera y Centros poblados y Rural disperso) y grupos de edad

Cuadro 1PD Departamental

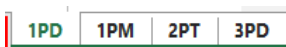
2018

Grupos de edad, departamento y Áreas (Total, Cabecera y Centros poblados y Rural disperso)		Total	Sexo		Índice de Masculinidad	Índice de Feminidad	
			Hombre	Mujer			
Total Nacional	Total	Total	44.164.417	21.570.493	22.593.924	95,5	104,7
	0 a 4		3.037.781	1.555.605	1.482.176	105,0	95,3
	5 a 9		3.335.240	1.705.574	1.629.666	104,7	95,5
	10 a 14		3.610.584	1.848.218	1.762.366	104,9	95,4
	15 a 19		3.852.255	1.970.530	1.881.725	104,7	95,5

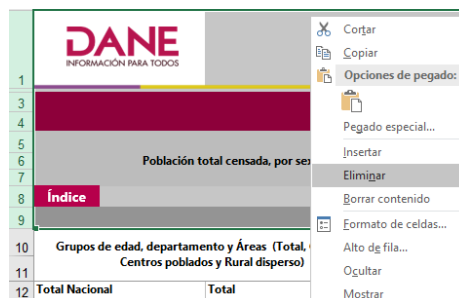
Todas estas filas deben ser eliminadas en todas las hojas. Si las hojas que se quieren modificar tienen la misma estructura, se puede hacer la eliminación de filas en todas a la vez al seleccionarlasy hacer los cambios deseados en solo una de ellas; las modificaciones se aplicarán en todas a la vez. El proceso es el siguiente:

- 1- Oprima la tecla **ctrl** y seleccione todas las hojas de interés.
- 2- Seleccione las filas objetivo y elimínelas haciendo clic en **Supr** o **Del**. Dependiendo de la cantidad de hojas y el equipo esto puede tomar tiempo.
- 3- Haga clic derecho en cualquiera de las hojas seleccionadas y seleccione la opción **Desagrupar hojas**.

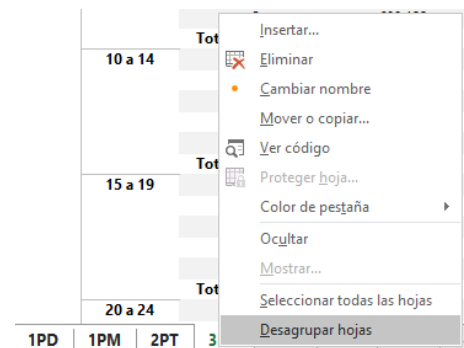
1. Seleccionar hojas



2. Eliminar filas



3. Desagrupar hojas



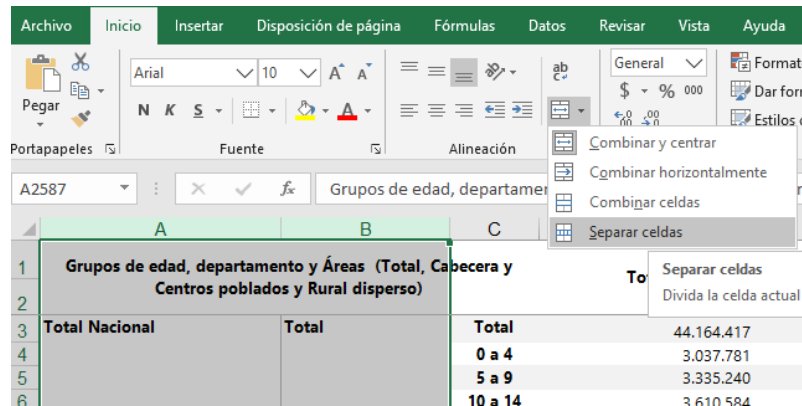
2.

SEPARACIÓN DE CELDAS

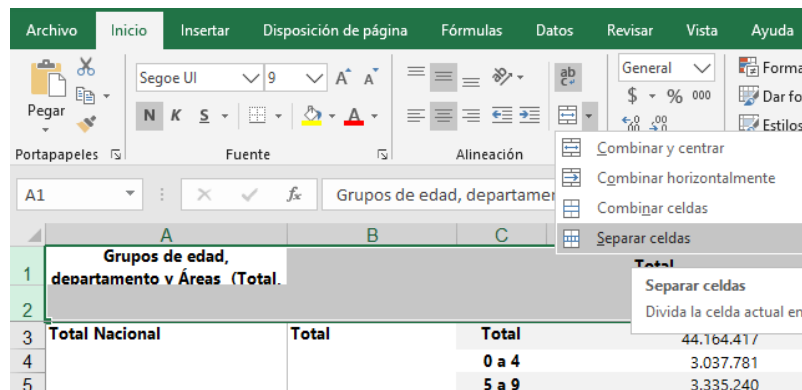
Para facilidad en el manejo de los datos, **ninguna celda debe estar combinada**, por lo cual tanto en columnas como en filas se deben separar. Las filas a separar son las que involucran el encabezado y las columnas son aquellas que presentan caracterización, es decir, no tienen conteos.

Luego de identificar donde hay celdas combinadas: 1) seleccione una de las columnas o filas que tenga agrupamientos, dando clic en la letra o número correspondiente y luego oprimir **ctrl** para seleccionar las restantes; 2) en la parte superior en **Alineación**, seleccione **Separar celdas**.

Separación sobre columnas



Separación sobre filas



3.

RELLENO DE FILAS VACÍAS

Parte 1: Ya se eliminaron las filas del encabezado inicial de las hojas y en la parte superior de las columnas quedan los títulos. Al haber separado las celdas en el paso anterior, quedan espacios vacíos que deben llenarse siendo necesario revisar qué información alberga cada columna para ponerle el título adecuado a cada una; igualmente organizar la información para que cargue de manera sencilla en R y por lo tanto en ColOpenData.

EJEMPLO

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Municipios y edades simples				Total	Sexo		Cabecera		Centro Poblado		Rural disperso				
2						Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer
3	Total		0 a 4	0	560.377	286.536	273.841	410.654	210.110	200.544	47.548	24.335	23.213	102.175	52.091	50.084

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Departamento	Municipio	Grupo de edad	Edad	Total	Cabecera		Centro Poblado		Rural disperso						
2	Departamento	Municipio	Grupo de edad	Edad	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer
3	Total		0 a 4	0	560.377	286.536	273.841	410.654	210.110	200.544	47.548	24.335	23.213	102.175	52.091	50.084
4				1	602.549	308.380	294.169	438.701	224.544	214.157	50.452	25.911	24.541	113.396	57.925	55.471

En este ejemplo se tienen las columnas correspondientes a departamento, municipio, grupo de edad y edad. Estas primeras columnas, que son las que presentan las características, deben tener el título o encabezado replicado en las dos primeras filas para que no queden vacías. A la derecha de estas se encuentran los valores de la población censada por área (total, cabecera, centro y rural) y sexo.

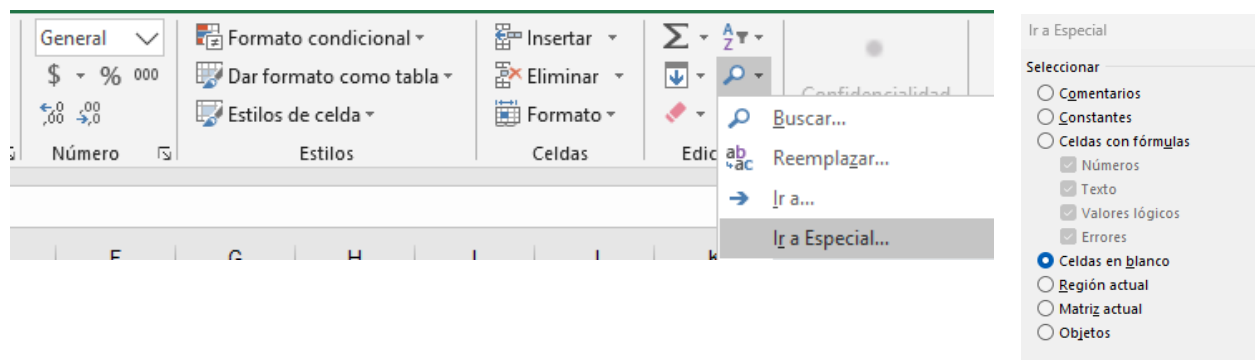
Parte 2: Una vez completos los encabezados, el siguiente paso es rellenar una a una las columnas con espacios vacíos. Para ello los pasos son:

1. **Clic** sobre la letra de la columna a rellenar.
2. **Ctrl + Shift + Flecha hacia arriba**. Si al hacer esto en la parte final se encuentra una fila con información no correspondiente a la tabla directamente, tal como:

Fuente: DANE - Censo Nacional de Población y Vivienda 2018

ésta se debe eliminar, para luego repetir los pasos 1 y 2.

3. En la parte superior en **Inicio -> Edición**, se debe dar clic en la lupa, en la opción **Ir a Especial...** y en la ventana emergente seleccionar **Celdas en blanco**, y dar clic en **Aceptar**.

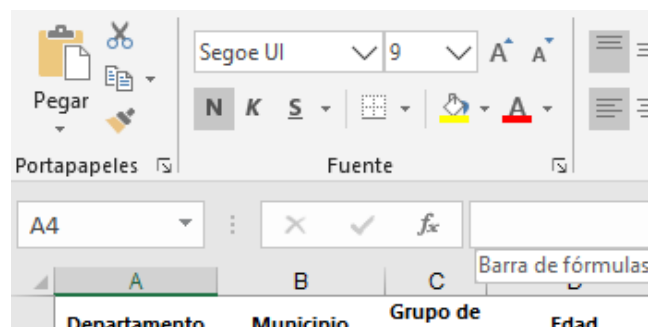


Una vez hecho esto, la columna seleccionada debe tener un formato similar al de la imagen abajo, donde las celdas con contenido no están seleccionadas (color blanco), y las vacías sí (color gris). Verifique revisando algunas filas con contenido para tener completa seguridad de que fueron reconocidas como tal. Si esto no sucede, dirigirse al *caso especial 2* al final del capítulo.

Solo celdas vacías seleccionadas

	A
1	Departamento
2	Departamento
3	Total
4	
5	
6	
7	
8	
9	

4. Después suba lo que más se pueda en la tabla sin dar clic en ninguna celda y luego **haga clic en la barra de fórmulas**.



5. Estando en la barra de fórmulas **escriba =, oprimir ctrl sostenido y** seleccionar la primera celda de la columna seleccionada que no hace parte del encabezado (revisar caso *especial 1* al final del capítulo, si la primera celda luego del encabezado está vacía). Luego **haga clic en Enter SIN SOLTAR CTRL**.

A3	X	✓	f _x	=A3
	A	B	C	D
1	Departamento	Municipio	Grupo de edad	Edad
2	Departamento	Municipio	Grupo de edad	Edad
3	Total		0 a 4	0
4	=A3			1
5				2
6				3

Las celdas vacías después de una celda de texto deben verse así:

	A
1	Departamento
2	Departamento
3	Total
4	Total
5	Total
6	Total
7	Total
8	Total
9	Total
10	Total

	A
1	Departamento
2	Departamento
146	Total
147	Total
148	Total
149	Total
150	05_Antioquia
151	05_Antioquia
152	05_Antioquia
153	05_Antioquia
154	05_Antioquia

	A
1	Departamento
2	Departamento
15707	05_Antioquia
15708	05_Antioquia
15709	05_Antioquia
15710	05_Antioquia
15711	08_Atlántico
15712	08_Atlántico
15713	08_Atlántico
15714	08_Atlántico
15715	08_Atlántico

6. Las celdas vacías debajo de la última celda con texto siguen vacías porque no se tuvieron en cuenta en la selección. Para llenarlas se debe **hacer clic en cualquier celda con texto de la columna y oprimir Ctrl + Flecha hacia abajo**. De esta manera se seleccionará la última celda con texto de la columna.
7. Esta celda se identifica porque tiene un recuadro verde, que en su esquina inferior derecha tiene un cuadrado pequeño. Para terminar

de rellenar se debe **hacer clic izquierdo dos veces seguidas sobre este cuadrado**. Para confirmar que se haya rellenado hasta el final repita el paso 6.

1	Departamento
2	Departamento
137588	97_Vaupés
137589	97_Vaupés
137590	97_Vaupés
137591	97_Vaupés
137592	97_Vaupés
137593	97_Vaupés
137594	99_Vichada
137595	
137596	
137597	
137598	
137599	
137600	



1	Departamento
2	Departamento
137588	97_Vaupés
137589	97_Vaupés
137590	97_Vaupés
137591	97_Vaupés
137592	97_Vaupés
137593	97_Vaupés
137594	99_Vichada
137595	99_Vichada
137596	99_Vichada
137597	99_Vichada
137598	99_Vichada
137599	99_Vichada
137600	99_Vichada

1	Departamento
2	Departamento
138065	99_Vichada
138066	99_Vichada
138067	99_Vichada
138068	99_Vichada
138069	99_Vichada
138070	99_Vichada
138071	99_Vichada
138072	99_Vichada
138073	
138074	
138075	
138076	
138077	
138078	

Todos los pasos de la parte 2 se deben repetir en cada una de las columnas que involucren caracterización de la población, tales como departamento, municipio, área, sexo, grupo de edad y edad. De esta manera quedan rellenas todas las filas que antes estaban vacías para evitar inconvenientes en la carga en lenguaje R.

CASOS ESPECIALES:

CASO ESPECIAL 1

PRIMERA CELDA LUEGO DEL ENCABEZADO VACÍA:

Se pueden tener columnas, por ejemplo "Municipio", que no tenga una celda con texto luego del encabezado ya que usualmente el primer departamento que aparece es "Total" que corresponde al total nacional. Esto quiere decir que se están tomando todos los municipios, por lo que no hay un municipio específico asociado. En este caso, se puede seleccionar la celda "Municipio" del encabezado para rellenar y seguir los demás pasos. **Es importante que sea la celda inmediatamente anterior a la primera celda vacía de la columna.** Al finalizar, el primer grupo de celdas vacías estará ahora relleno con "Municipio". Lo más sencillo es seleccionarlas y hacer clic en **supr**. Siga estos pasos para casos similares.

CASO ESPECIAL 2

CELDA VACIAS RECONOCIDAS COMO CELDAS DE TEXTO:

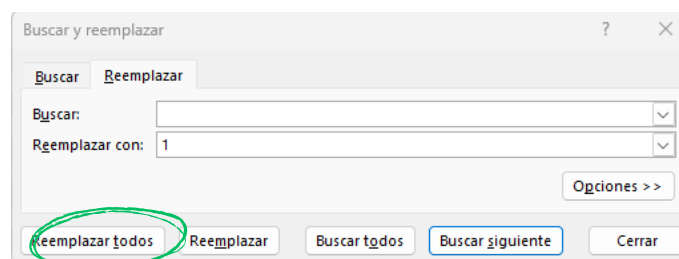
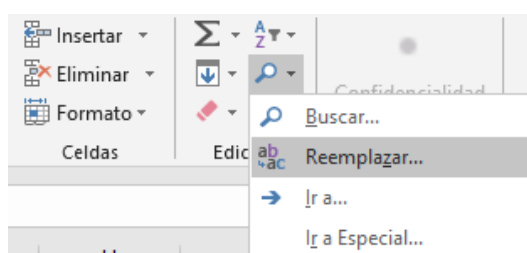
Puede pasar que al llegar al paso 3 de la parte 2 la columna de interés se vea algo así, lo que quiere decir que se están identificando la mayoría de celdas vacías como celdas con texto.



	A
1	Departamento
2	Departamento
3	Total Nacional
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	05_Antioquia
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	08_Atlántico
28	
29	
30	
31	

En este caso, para rellenar las celdas, se debe:

- Seguir los pasos 1 y 2 de la parte 2.
- En la parte superior en **Inicio -> Edición**, hacer **click en la lupa**, en la opción **Reemplazar...**
- En la ventana emergente dejar la casilla de **Buscar** vacía y en la de **Reemplazar** escribir **1** o algún número de un solo carácter. Haga clic en **Reemplazar todos**.

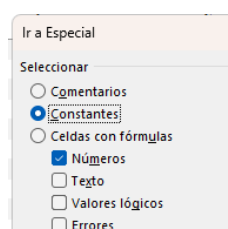
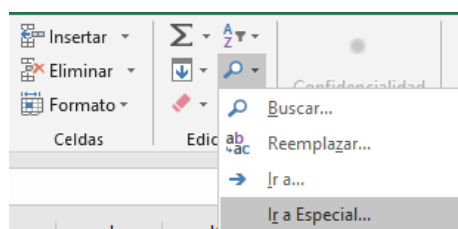


Ahora las celdas que estaban vacías deben tener un número **1** o el carácter escogido.



	A
1	Departamento
2	Departamento
3	Total Nacional
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	05_Antioquia
16	1
17	1
18	1
19	1
20	1

Compruebe que todas las celdas de la columna aparecen seleccionadas. Si esto no es así, se deben seguir los pasos 1 y 2 de la parte 2. Luego, en la parte superior en **Inicio -> Edición**, se debe dar **click en la lupa**, en la opción **Ir a Especial...** y en la ventana emergente **seleccionar Constantes**, y solo debe quedar seleccionada adicionalmente la casilla de **Números**. Haga clic en **Aceptar**.



Siga los pasos 4 y 5 de la parte 2. El paso 4 tiene una diferencia mínima y es que se debe borrar el 1 que se muestra en la barra de fórmulas para escribir lo que pide el paso. Luego se debe seguir el paso 6 para verificar que todas las celdas vacías estén llenas con la información correspondiente.

4.

ESTANDARIZACIÓN DE
ENCABEZADO

Luego de llevar a cabo todos los pasos anteriores, es necesario terminar la limpieza de los datos ajustando también el encabezado para que tenga un formato que permita pasar el archivo por R. Al hablar de encabezado se hace referencia a los “títulos” de las columnas que albergan los datos. Por ejemplo, en la siguiente imagen estas columnas corresponden a todas aquellas que van luego de “Edad”.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Departamento	Área	Sexo	Edad	Total personas menores de 5 años	Dónde o con quién permanece el menor							
2	Departamento	Área	Sexo	Edad		Asiste a un hogar comunitario, jardín, centro de desarrollo infantil o colegio	Con su padre o madre en la vivienda	Con su padre o madre en el trabajo	En la vivienda donde vive el niño al cuidado de un pariente o una persona de 18 años o más	En la vivienda donde vive el niño al cuidado de un pariente o una persona menor de 18 años	Al cuidado de un pariente o de otra persona en otro lugar	En la vivienda, solo	Sin Información
3	Total Nacional	Total	Hombre	0	286.111	15.769	235.039	4.459	17.975	1.854	5.835	599	4.581

Para que el archivo sea apto para pasar por R es necesario que la división global “Dónde o con quién permanece el menor” esté al lado de la última columna de caracterización “Edad” y que la segunda fila no tenga espacios vacíos. Este ajuste se realiza de forma que el total de menores encuestados hace parte de la división global, así:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Departamento	Área	Sexo	Edad	Dónde o con quién permanece el menor								
2	Departamento	Área	Sexo	Edad	Total personas menores de 5 años	Asiste a un hogar comunitario, jardín, centro de desarrollo infantil o colegio	Con su padre o madre en la vivienda	Con su padre o madre en el trabajo	En la vivienda donde vive el niño al cuidado de un pariente o una persona de 18 años o más	En la vivienda donde vive el niño al cuidado de un pariente o una persona menor de 18 años	Al cuidado de un pariente o de otra persona en otro lugar	En la vivienda, solo	Sin Información
3	Total Nacional	Total	Hombre	0	286.111	15.769	235.039	4.459	17.975	1.854	5.835	599	4.581

La información y cómo está almacenada en las diferentes hojas de un mismo archivo -o de diferentes archivos- puede ser muy diferente, por lo que esta última parte de estandarización de encabezado es muy subjetiva.

Sin embargo, aunque no sea estándar se deben cumplir con ciertas características, según el caso:

1. Si la hoja cuenta con 3 filas de encabezado:

- La primera fila debe tener una casilla única que presente la división o factor de lo presentado en la segunda fila y debe estar al lado de la última casilla de caracterización.
- La segunda fila debe tener los niveles de esta primera fila.
- La tercera fila debe tener una subdivisión que se repite para los diferentes niveles presentados en la segunda fila.

EJEMPLO

ANTES

C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Grupo de edad	Sabe leer y escribir														
Grupo de edad	Total	Si					No			Sin información					
Grupo de edad	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer
De 5 años y más	40.802.017	19.763.734	21.038.283	37.872.479	18.279.901	19.592.578	2.432.568	1.236.330	1.196.238	496.970	247.503	249.467			

DESPUÉS

C	D	E	F	G	H	I	J	K	L	M	N	O
Grupo de edad	Sabe leer y escribir											
Grupo de edad	Total	Si					No			Sin información		
Grupo de edad	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombre	Mujer
De 5 años y más	40.802.017	19.763.734	21.038.283	37.872.479	18.279.901	19.592.578	2.432.568	1.236.330	1.196.238	496.970	247.503	249.467

2. Si la hoja cuenta con 2 filas de encabezado:

- La primera fila debe tener una única casilla que presente la división o factor de lo presentado en la segunda fila, y debe estar de lado de la última casilla de caracterización.
- La segunda fila debe tener los niveles de esta primera fila.

Nota: Si la primera fila tiene más de una casilla, es necesario añadir o quitar casillas -pueden ser filas- de tal forma que se cuente con toda la información en el formato deseado.

EJEMPLO

ANTES

C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Grupo de edad	2018			2005			1993			1985			1973		
Grupo de edad	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombres	Mujeres	Total	Hombres	Mujeres	Total	Hombres	Mujeres
Total	44,164,417	21,570,493	22,593,924	41,468,384	20,336,117	21,132,267	33,109,839	16,296,538	16,813,301	27,853,436	13,785,523	14,067,913	20,666,920	10,124,394	10,542,526

DESPUÉS

C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Grupo de edad	2018			2005			1993			1985			1973		
Grupo de edad	Total	Hombre	Mujer	Total	Hombre	Mujer	Total	Hombres	Mujeres	Total	Hombres	Mujeres	Total	Hombres	Mujeres

RESULTADO

Una vez seguidos todos los pasos de este instructivo, la base de datos tendrá un formato que permite un manejo más sencillo por medio de R con un patrón más estandarizado. Esto permitirá que se generen códigos automatizados para organizar toda la información de una manera similar a lo que serían factores y niveles, es decir, que en cada columna se encuentra un encabezado y los demás valores de las filas tienen relación con este. Así, los datos se podrán manejar por medio de data frames y será más sencillo obtener información relevante para el investigador.

EJEMPLO

	codigo_departamento	departamento	codigo_municipio	municipio	area	condicion	total
1	0	NA	0	NA	total	total_hogares	14243223
2	0	NA	0	NA	total	hogares_con_menores_de_15_anos	6216640
3	0	NA	0	NA	total	total_menores_de_15_anos	9932653
3370	0	Total nacional	0	NA	cabecera	total_hogares	11128226
3371	0	Total nacional	0	NA	cabecera	hogares_con_menores_de_15_anos	4706422
3372	0	Total nacional	0	NA	cabecera	total_menores_de_15_anos	7213825
6739	0	Total nacional	0	NA	centro_poblado	total_hogares	947943
6740	0	Total nacional	0	NA	centro_poblado	hogares_con_menores_de_15_anos	484939
6741	0	Total nacional	0	NA	centro_poblado	total_menores_de_15_anos	850521
10108	0	Total nacional	0	NA	rural_disperso	total_hogares	2167054



Pontificia Universidad
JAVERIANA
Bogotá

| VIGILADA MINEDUCACIÓN |

 **Universidad de
los Andes**
Colombia

Epiverse TRACE LAC

Supported by :



MRC Unit
The
Gambia



data.org

Funded by:



IDRC • CRDI

Canada