



Universidad de **Nariño**



MinTIC
Ministerio de Tecnologías
de la Información y las Comunicaciones

IMPORTACIÓN DE LOS DATOS DE VIOLENCIA INTRAFAMILIAR 2017 SEGURIDAD Y DEFENSA

Integrantes:

Eddie Jean Pierre Alarcon Andrade

Jonathan Andrés Rodríguez Tello

Juan Camilo Vallejos Guerrero

Asesora:

Mg. Sandra Vallejo

Universidad de Nariño

Ministerio de Tecnologías de la información y las comunicaciones

Pasto - Nariño

Octubre del 2017



Universidad de Nariño



MinTIC
Ministerio de Tecnologías
de la Información y las Comunicaciones

DATOS VIOLENCIA INTRAFAMILIAR 2017 SEGURIDAD Y DEFENSA

Violencia Intrafamiliar (del 01 de enero al 31 de julio año 2017).

Fuente: Grupo Información de Criminalidad (GICRI) - DIJIN.

Información preliminar, datos extraídos el 02 de agosto de 2017, sujeto a variación.

Actualizado 22 de agosto de 2017

Datos ofrecidos por Policía Nacional

Propietario de conjunto de datos DIJIN

Departamento: Bogotá D.C.

Municipio: Bogotá D.C.

Importación

Formato y estandarización de los datos

1. Se descarga el archivo csv para Excel del link oficial
<https://www.datos.gov.co/Seguridad-y-Defensa/Violencia-intrafamiliar-2017/3sag-xzkg>
2. Se importan los datos del archivo en un documento de Excel.
3. Los datos contienen encabezado, se separan por comas, el texto está contenido entre comillas y dentro de este existen comas.
4. Se da formato a los siguientes datos (los demás en formato general):
 - a. Fecha, de tipo fecha en formato MM/DD/AAAA.
5. Al terminar la importación de los datos nos percatamos la fecha, tenía información adicional que no requeríamos por lo tanto se procede a eliminar.
 - a. La fecha contaba con una hora, las 0:00 o las 12:00:00 AM. Esta se eliminó.
6. Guardar cambios, para la importación dentro de la aplicación de la aplicación de Tableau, la cual se utilizara para la realización de la infografía.
7. Dentro de la aplicación de Tableau, se importa el archivo de Excel con los datos estandarizados, y de esta se elige la hoja que posee todos los datos.
8. Los datos en la aplicación se manejan con formatos, lo cuales se eligieron los siguientes (los demás en formato general, ABC dentro de la aplicación):
 - a. Fecha, de tipo fecha
 - b. Departamentos de tipo texto con rol geográfico (Estado/Provincia).
 - c. Municipio de tipo texto con rol geográfico (Provincias/Municipios).
 - d. Hora de tipo fecha y hora



Universidad de Nariño



MinTIC
Ministerio de Tecnologías
de la Información y las Comunicaciones

- e. Edad de tipo número entero
 - f. Cantidad de tipo número entero
9. Una vez realizados los cambios se procede a la realización de la hoja de trabajo en Tableau, donde se procede a ubicar los datos dentro del mapa de Tableau (el cual a su vez hace uso de la API de OpenStreetMap), donde nos encontramos con los siguientes problemas:
- a. El aplicativo reconoce de forma automática en qué país se encuentran los atributos de departamento y municipios, en ocasiones restaba que las ubicaciones las generaba en otras regiones del mundo, por lo que se plantea la adición a los datos de una columna que contenga el país donde suceden los delitos, en este caso Colombia, de tipo texto con rol geográfico (País/Región).
 - b. Los departamentos de Guajira, San Andrés y Valle no fueron reconocidos por el mapa, esto debido a que dentro de los mapas de OpenStreetMap, estos aparecen con los nombres de La Guajira, San Andrés y Providencia, y Valle del Cauca respectivamente. Dado esto se plantea corregir dentro del archivo de Excel los datos, con los nombres que les corresponden.
 - c. Existe gran cantidad de errores dentro de los municipios estos debido a dos razones, la primera porque muchos de ellos pertenecen a los departamentos anteriormente nombrados y la segunda por la inclusión de un complemento que acompaña el nombre del municipio, para el caso de la ciudades capitales se tiene el complemento (CT) y para el caso especial de Bogotá se tiene el complemento D.C. (CT), lo que resulta en el aplicativo, la no ubicación de los datos. Por lo que se plantea eliminar dichos complementos.
10. Para la realización de las estandarizaciones se procede a cerrar la fuente de datos dentro de Tableau y realizar las estandarizaciones pertinentes.
11. Se importan nuevamente los datos y se repite el numeral 6, 7 anteriormente descrito.
12. Al proceder nuevamente a la creación de la hoja de trabajo se encontraron los siguientes problemas:
- a. Los municipios Mariquita, Morales (Cauca), Morales (Bolívar), Salazar, San Luis de Sincé, Uribe, Villa de San Diego de Ubaté, Carolina, Cerro San Antonio y Providencia, los cuales no se pueden ubicar porque al igual que antes sus nombres aparecen diferente dentro de OpenStreetMap, por lo que se plantea estandarizar sus nombres, teniendo en cuenta como aparecen dentro de los mapas oficiales. Respectivamente San Sebastián de Mariquita, Morales, Morales, Salazar de Las Palmas, Sincé, La Uribe, Ubaté, Carolina del príncipe y Cerro de San Antonio. Por otra parte providencia no fue posible



Universidad de Nariño



MinTIC
Ministerio de Tecnologías
de la Información y las Comunicaciones

ubicarla, por lo que se procedió a añadir la posición geográfica en latitud y longitud las cuales se usó. En latitud: 13,348889 y en longitud: -81,374722. Obtenido de diferentes plataformas web de localización, tales como Google, Bing y el mismo OpenStreetMap.

- b. El municipio de Bogotá no se puede localizar debido a que dentro de los mapas de OpenStreetMap, este aparece como municipio del departamento de Bogotá, mientras que en la base de datos esta de Cundinamarca. Por lo que se plantea cambiar el departamento del municipio de Bogotá de Cundinamarca, por el de Bogotá.

- 13. Para ingresar a la hoja con todas las estandarizaciones realizadas por favor ingresar al siguiente link:

[http://www.mediafire.com/file/uu8r3ph8068cpw3/Violencia intrafamiliar 2017 Estandarizado.xlsx](http://www.mediafire.com/file/uu8r3ph8068cpw3/Violencia_intrafamiliar_2017_Estandarizado.xlsx)