

Proyecto 2: Registro del núcleo básico de las víctimas de minas antipersonal

Análisis y visualización de datos con Python

víctimas de minas antipersonal(MAP)|

1 - Entendimiento del proyecto

1.1 - Objetivos del proyecto.

La situación de las minas antipersonal en Colombia ha representado una problemática humanitaria y de seguridad durante décadas, dejando un legado de víctimas y áreas afectadas en diversas regiones del país. Estos artefactos explosivos improvisados han tenido un impacto devastador en las comunidades locales, causando muertes, lesiones graves y desplazamientos forzados. El Registro del Núcleo Básico de las Víctimas de Minas Antipersonal (MAP) se ha establecido como una herramienta fundamental para recopilar información precisa y detallada sobre las víctimas y proporcionar asistencia integral, rehabilitación y atención psicosocial. El análisis de datos relacionado con este registro es crucial para entender la magnitud del problema, identificar patrones geográficos y demográficos, evaluar las necesidades específicas de las víctimas y diseñar estrategias efectivas de prevención, desminado y atención. A través de técnicas de análisis de datos, se busca generar información resumida que permitan a las autoridades y organizaciones pertinentes tomar decisiones informadas, asignar recursos de manera eficiente y trabajar hacia la erradicación completa de las minas antipersonal en Colombia, garantizando así un futuro más seguro y libre de riesgos para todas las comunidades afectadas.

1.2 - Objetivo del proyecto.

El objetivo principal es obtener la mayor cantidad de información de utilidad que no es visible a simple vista en la base de datos proporcionada, por lo que se abordara el cumplimiento del objetivo mediante las siguientes metas parciales:

- Limpieza y organización de las columnas de forma correcta validando entradas lógicas de datos.
- Presentación del informe del total de víctimas por año, mes (Independiente del año), departamento, actividad, género y estado.

2 - Entendimiento de los datos

En la etapa de entendimiento de los datos utilizaremos *Python* en análisis avanzado con estadística inferencial y modelado estadístico junto con las librerías *Matplotlib* y *Seaborn* en la construcción de visualizaciones de información. Utilizaremos el *dataset* de [Registro del núcleo básico de las víctimas de minas antipersonal \(MAP\)](#) dicho conjunto de datos se encuentra de manera pública en el portal.*

El conjunto de datos cuenta con las siguientes variables:

- departamento
- codigodanedepto
- municipio
- codigodanemunicipio
- tipoarea
- sitio
- anio
- mes
- rangoedad
- grupoetnico
- condicion
- estado
- genero
- latitudcabecera
- longitudcabecera
- tipoevento
- Ubicación
- Actividad

```
In [ ]: !pip install -U plotly
!pip install -U kaleido
```

Requirement already satisfied: plotly in c:\users\yuri ochoa\appdata\local\programs\python\python310\lib\site-packages (5.18.0)
Requirement already satisfied: tenacity>=6.2.0 in c:\users\yuri ochoa\appdata\local\programs\python\python310\lib\site-packages (from plotly) (8.2.3)
Requirement already satisfied: packaging in c:\users\yuri ochoa\appdata\local\programs\python\python310\lib\site-packages (from plotly) (23.2)
Requirement already satisfied: kaleido in c:\users\yuri ochoa\appdata\local\programs\python\python310\lib\site-packages (0.2.1)

```
In [ ]: import plotly
print(f'Plotly: {plotly.__version__}')
```

Plotly: 5.18.0

```
In [ ]: #Importar Librerías
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib as mpl
import seaborn as sns
from datetime import datetime
from scipy import stats
import statsmodels.api as sm
import unicodedata
import re
import calendar

%matplotlib inline
```

```
In [ ]: !python --version
print(f'pandas { pd.__version__}')
```

Python 3.10.5pandas 2.1.4
NumPy 1.26.3

Matplotlib 3.8.2
seaborn 0.13.1

Este proyecto fue realizado con las siguientes versiones:

- Python 3.10.5
- pandas 2.1.4
- NumPy 1.26.3
- Matplotlib 3.8.2
- seaborn 0.13.1
- Plotly: 5.18.0

2.1 - Adquisición e integración de datos.

Para este paso importaremos el archivo de datos aprovechando la utilidad de carga de conjuntos de datos comunes de Seaborn.

```
In [ ]: #Creamos Los datos
url = 'https://storage.googleapis.com/media-help-ots-geomarketing-odoo/Situaci_n_V_ctimas_Minas_Antipersonal_en_Colombia_20240103.csv'
registro_map = pd.read_csv(url)
registro_map.head()
```

	departamento	codigodanedepto	municipio	codigodanemunicipio	tipoarea	sitio	ano	mes	rangoedad	grupoetnico	condicion	estado	genero	latitudcabecera	longitudcabecera	tipoevento	Ubicación	Actividad
0	ANTIOQUIA	5	GRANADA	5,313	Rural	Sin información	2,006	1	Mayor de 18 años	No	Civil	Hombre	Herido	6.143703	-75.185183	Sin informacion	(6.143703, -75.185183)	Pasando, estando cerca
1	CAQUETA	18	MONTAÑITA	18,410	Rural	Sin información	2,009	12	Mayor de 18 años	No	Civil	Hombre	Herido	1.494944	-75.463090	Sin informacion	(1.49494421255, -75.4630895064)	Erradicador manual - PCI
2	PUTUMAYO	86	PUERTO ASÍS	86,568	Rural	Sin información	2,008	4	Mayor de 18 años	No	Civil	Hombre	Muerto	0.502569	-76.498180	Sin informacion	(0.502569, -76.49818)	Erradicador manual - PCI
3	CORDOBA	23	PUERTO LIBERTADOR	23,580	Rural	Sin información	2,009	11	Mayor de 18 años	No	Civil	Hombre	Herido	7.883545	-75.670508	Sin informacion	(7.883545, -75.670508)	Erradicador manual - PCI
4	CORDOBA	23	PUERTO LIBERTADOR	23,580	Rural	Sin información	2,009	11	Mayor de 18 años	No	Civil	Hombre	Herido	7.883545	-75.670508	Sin informacion	(7.883545, -75.670508)	Erradicador manual - PCI

2.2 - Análisis exploratorio de datos.

En las siguientes celdas se explora los diferentes tipos de variables (columnas) del conjunto de datos utilizando *pandas*.

```
In [ ]: registro_map.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12364 entries, 0 to 12363
Data columns (total 18 columns):
#   Column                Non-Null Count  Dtype
---  -
0   departamento          12364 non-null  object
1   codigodanedepto        12364 non-null  int64
2   municipio              12364 non-null  object
3   codigodanemunicipio    12364 non-null  object
4   tipoarea               12364 non-null  object
5   sitio                  12364 non-null  object
6   ano                    12364 non-null  object
7   mes                    12364 non-null  int64
8   rangoedad              12364 non-null  object
9   grupoetnico            12364 non-null  object
10  condicion               12364 non-null  object
11  estado                  12364 non-null  object
12  genero                  12364 non-null  object
13  latitudcabecera         12364 non-null  float64
14  longitudcabecera        12364 non-null  float64
15  tipoevento              12364 non-null  object
16  Ubicación               12364 non-null  object
17  Actividad               12364 non-null  object
dtypes: float64(2), int64(2), object(14)
memory usage: 1.7+ MB
```

A continuacion se identifican, que variables tienen **filas con valores faltantes**:

```
In [ ]: valores_faltantes = registro_map.columns[registro_map.isnull().any()]
valores_faltantes

Out[ ]: Index([], dtype='object')
```

2.3 - Descripciones generales.

Para una exploracion inicial, se imprime en pantalla los primeros y ultimos diez registros de datos junto con una muestra aleatoria, utilizando las funciones `head()`, `tail()` y `.sample()`:

```
In [ ]: registro_map.head(10)
```

Out[]:

	departamento	codigodan	departamento	municipio	codigodan	municipio	tipo	area	sitio	ano	mes	rango	edad	grupo	etnico	condicion	estado	genero	latitud	cabecera	longitud	cabecera	tipo	evento	Ubicación	Actividad
0	ANTIOQUIA	5		GRANADA	5,313		Rural		Sin información	2,006	1	Mayor de 18 años		No		Civil	Hombre	Herido	6.143703		-75.185183		Sin informacion	(6.143703, -75.185183)	Pasando, estando cerca	
1	CAQUETA	18		MONTAÑITA	18,410		Rural		Sin información	2,009	12	Mayor de 18 años		No		Civil	Hombre	Herido	1.494944		-75.463090		Sin informacion	(1.49494421255, -75.4630895064)	Erradicador manual - PCI	
2	PUTUMAYO	86		PUERTO ASÍS	86,568		Rural		Sin información	2,008	4	Mayor de 18 años		No		Civil	Hombre	Muerto	0.502569		-76.498180		Sin informacion	(0.502569, -76.49818)	Erradicador manual - PCI	
3	CORDOBA	23		PUERTO LIBERTADOR	23,580		Rural		Sin información	2,009	11	Mayor de 18 años		No		Civil	Hombre	Herido	7.883545		-75.670508		Sin informacion	(7.883545, -75.670508)	Erradicador manual - PCI	
4	CORDOBA	23		PUERTO LIBERTADOR	23,580		Rural		Sin información	2,009	11	Mayor de 18 años		No		Civil	Hombre	Herido	7.883545		-75.670508		Sin informacion	(7.883545, -75.670508)	Erradicador manual - PCI	
5	CORDOBA	23		PUERTO LIBERTADOR	23,580		Rural		Sin información	2,009	11	Mayor de 18 años		No		Civil	Hombre	Herido	7.883545		-75.670508		Sin informacion	(7.883545, -75.670508)	Erradicador manual - PCI	
6	CORDOBA	23		PUERTO LIBERTADOR	23,580		Rural		Sin información	2,009	11	Mayor de 18 años		No		Civil	Hombre	Herido	7.883545		-75.670508		Sin informacion	(7.883545, -75.670508)	Erradicador manual - PCI	
7	NARIÑO	52		BARBACOAS	52,079		Rural		Vereda Santa Lucia	2,010	1	Mayor de 18 años		No		Fuerza pública	Hombre	Herido	1.587500		-78.227500		Sin informacion	(1.5875, -78.2275)	Fuerza pública	
8	ANTIOQUIA	5		TARAZÁ	5,790		Rural		Quebrada Iraca	2,010	1	Mayor de 18 años		No		Fuerza pública	Hombre	Herido	7.372778		-75.204444		Sin informacion	(7.372777777777777, -75.20444444444445)	Fuerza pública	
9	ANTIOQUIA	5		VALDIVIA	5,854		Rural		Sin información	2,008	3	Mayor de 18 años		No		Civil	Hombre	Herido	7.165222		-75.438081		Sin informacion	(7.165222, -75.438081)	Erradicador manual - PCI	

In []:

registro_map.tail(10)

Out[]:

	departamento	codigodane	departamento	municipio	codigodane	municipio	tipo	area	sitio	ano	mes	rango	edad	grupo	etnico	condicion	estado	genero	latitud	cabecera	longitud	cabecera	tipo	evento	Ubicación	Actividad
12354	NARIÑO	52		OLAYA HERRERA	52,490		Rural		VEREDA SAN JOSE LOS ROBLES	2,023	6	Menor de 18 años		No		Civil	Mujer	Herido	2.314415		-78.334345		Campo		(2.31441546587, -78.3343454855)	Agricultura
12355	NORTE DE SANTANDER	54		EL TARRA	54,250		Rural		Vereda El Porvenir	2,023	6	Mayor de 18 años		No		Fuerza pública	Hombre	Herido	8.658333		-73.146944		Sin informacion		(8.658333333333333, -73.14694444444444)	Actividades fuerza pública
12356	NARIÑO	52		IPIALES	52,356		Rural		VEREDA EL ASUAY	2,023	6	Mayor de 18 años		No		Civil	Hombre	Herido	0.829810		-77.629760		Campo		(0.82981043391, -77.6297604828)	Agricultura
12357	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Mujer	Herido	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información
12358	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Hombre	Herido	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información
12359	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Mujer	Herido	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información
12360	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Mujer	Herido	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información
12361	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Hombre	Herido	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información
12362	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Hombre	Herido	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información
12363	NARIÑO	52		RICAURTE	52,612		Rural		RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA	2,023	6	Mayor de 18 años		Indigena		Civil	Hombre	Muerto	1.190763		-78.117007		Caminos		(1.19076305555556, -78.11700694444444)	Sin Información

In []:

registro_map.sample(10)

Out []:

	departamento	codigodanede	departamento	municipio	codigodanemunicipio	tipoarea	sitio	ano	mes	rangoedad	grupoetnico	condicion	estado	genero	latitudcabecera	longitudcabecera	tipoevento	Ubicación	Actividad			
	3377		ANTIOQUIA		5	ITUANGO		5,361	Rural	Vereda El Olivar	2,005	9	Mayor de 18 años	No	Fuerza pública	Hombre	Herido	7.231944	-75.864167	Sin informacion	(7.231944, -75.864167)	Fuerza pública
	2422		TOLIMA		73	RIOBLANCO		73,616	Rural	Vereda La Catalina, Finca Los Guayabos	2,007	1	Menor de 18 años	No	Civil	Hombre	Muerto	3.529701	-75.644823	Sin informacion	(3.529701, -75.644823)	Pasando, estando cerca
	1535		ARAUCA		81	ARAUQUITA		81,065	Rural	Sur Filipinas	2,009	1	Mayor de 18 años	No	Fuerza pública	Hombre	Herido	6.646667	-71.216944	Sin informacion	(6.646666666666665, -71.21694444444445)	Fuerza pública
	7514		ANTIOQUIA		5	ZARAGOZA		5,895	Rural	Sin información	2,001	4	Menor de 18 años	No	Civil	Mujer	Herido	7.490632	-74.867778	Sin informacion	(7.490632, -74.867778)	Pasando, estando cerca
	12156		CAUCA		19	EL TAMBO		19,256	Rural	CORREGIMIENTO HUISITÓ	2,022	2	Mayor de 18 años	No	Civil	Hombre	Herido	2.450573	-76.805205	Sin informacion	(2.45057270515, -76.8052045915)	Sin Información
	8953		PUTUMAYO		86	PUERTO ASÍS		86,568	Rural	Sin información	2,011	8	Mayor de 18 años	No	Fuerza pública	Hombre	Muerto	0.285278	-76.540833	Sin informacion	(0.2852777778, -76.5408333333)	Fuerza pública
	8116		NORTE DE SANTANDER		54	TEORAMA		54,800	Rural	Sin información	2,007	9	Mayor de 18 años	No	Fuerza pública	Hombre	Herido	8.441389	-73.268611	Sin informacion	(8.441389, -73.268611)	Fuerza pública
	7163		META		50	VISTAHERMOSA		50,711	Rural	Vereda La Reforma Sur	2,008	5	Mayor de 18 años	No	Fuerza pública	Hombre	Herido	3.125088	-73.752373	Sin informacion	(3.125088, -73.752373)	Fuerza pública
	10888		CAUCA		19	EL TAMBO		19,256	Rural	Huisito Santa Rita	2,010	6	Mayor de 18 años	No	Fuerza pública	Hombre	Herido	2.604722	-77.012500	Sin informacion	(2.604722222222222, -77.0125)	Fuerza pública
	11586		NORTE DE SANTANDER		54	SAN CALIXTO		54,670	Rural	VEREDA LA LAGUNA	2,018	8	Menor de 18 años	No	Civil	Hombre	Herido	8.404078	-73.208168	Caminos	(8.40407771429, -73.208167626)	Recogiendo comida, agua ó madera

Debido a que la base de información con la que se esta trabajando tiene en su mayoría variables categóricas, se quiere encontrar la información acerca de las categorías que cada una de ellas puede tener para posteriormente ser limpiada en el caso de que tengan caracteres especiales como tildes, acentos, ñ, entre otros. Para ello utilizaremos para identificar un arreglo con estas categoría (unique) y posteriormente bibliotecas como unicodedata y expresiones regulares (regex).

In []:

```
registro_map['departamento'].unique()
```

Out []:

```
array(['ANTIOQUIA', 'CAQUETA', 'PUTUMAYO', 'CORDOBA', 'NARIÑO', 'CAUCA',
      'GUAVIARE', 'META', 'NORTE DE SANTANDER', 'BOLIVAR', 'HUILA',
      'VALLE DEL CAUCA', 'VAUPES', 'ARAUCA', 'TOLIMA', 'CASANARE',
      'VICHADA', 'CHOCO', 'SANTANDER', 'CESAR', 'LA GUAJIRA',
      'CUNDINAMARCA', 'CALDAS', 'BOGOTA DC', 'MAGDALENA', 'SUCRE',
      'QUINDIO', 'BOYACA', 'GUAINIA', 'RISARALDA', 'ATLANTICO',
      'AMAZONAS'], dtype=object)
```

In []:

```
registro_map['codigodanededepartamento'].unique()
```

Out []:

```
array([ 5, 18, 86, 23, 52, 19, 95, 50, 54, 13, 41, 76, 97, 81, 73, 85, 99,
      27, 68, 20, 44, 25, 17, 11, 47, 70, 63, 15, 94, 66,  8, 91],
      dtype=int64)
```

In []:

```
registro_map['municipio'].unique()
```

```
Out[ ]: array(['GRANADA', 'MONTAÑITA', 'PUERTO ASÍS', 'PUERTO LIBERTADOR',
              'BARBACOAS', 'TARAZÁ', 'VALDIVIA', 'TORIBÍO', 'SANTA CRUZ',
              'RICAURTE', 'EL RETORNO', 'SAN JUAN DE ARAMA', 'TIBÚ', 'SAN PABLO',
              'SAN MIGUEL', 'SUÁREZ', 'EL CARMEN DE BOLÍVAR', 'MAPIRIPÁN',
              'ALGECIRAS', 'ITUANGO', 'FLORIDA', 'BARAYA', 'TIERRALTA',
              'TALAIGUA NUEVO', 'CARURÚ', 'TEORAMA', 'HACARÍ', 'CÁCERES',
              'PUERTO RONDÓN', 'SAMANIEGO', 'TAME', 'MIRAFLORES', 'RIOBLANCO',
              'CIÉNAGA DE ORO', 'CALAMAR', 'SAN CARLOS', 'APARTADÓ', 'LA PLAYA',
              'VALLE DEL GUAMUEZ', 'SOLANO', 'SAN CALIXTO', 'SAN ANDRÉS',
              'AGUAZUL', 'ISNOS', 'ARAUQUITA', 'PUERTO RICO', 'FLORENCIA',
              'MEDELLÍN', 'COLOMBIA', 'EL TARRA', 'BALBOA', 'SANTA ROSA DEL SUR',
              'URIBE', 'SAN LUIS', 'SAN JOSÉ DEL FRAGUA', 'PRADERA', 'PALERMO',
              'ANORÍ', 'SAN JOSÉ DEL GUAVIARE', 'VISTAHERMOSA', 'ÁBREGO',
              'CORINTO', 'URRAO', 'TULUÁ', 'MESETAS', 'EL TAMBO', 'CUMARIBO',
              'CALOTO', 'SAN AGUSTÍN', 'BRICEÑO', 'EL CARMEN DEL DARIEN',
              'BUGALAGRANDE', 'CANTAGALLO', 'CONVENCIÓN', 'VIGÍA DEL FUERTE',
              'NECOCLÍ', 'SABANA DE TORRES', 'PAZ DE ARIPORO', 'DABEIBA',
              'PUERRES', 'NEIVA', 'GUADALUPE', 'SAN PEDRO', 'PLANADAS',
              'PUERTO CAICEDO', 'PUERTO LEGUIZAMO', 'SAN VICENTE DEL CAGUÁN',
              'LÓPEZ', 'LEJANÍAS', 'MELGAR', 'VALLEDUPAR', 'MITÚ', 'SANTA ROSA',
              'URIBIA', 'RIOHACHA', 'LOS ANDES', 'EL BAGRE', 'ANGOSTURA',
              'POLICARPA', 'SANTA BÁRBARA', 'TUMACO', 'CALDONO', 'ARGELIA',
              'PUERTO TRIUNFO', 'TELLO', 'ARAUCA', 'EL CASTILLO', 'EL DORADO',
              'ARENAL', 'PASCA', 'GUAMAL', 'IPIALES', 'SIMITÍ', 'CUBARRAL',
              'CAMPOALEGRE', 'AMALFI', 'SÁCAMA', 'CUMBITARA',
              'CARTAGENA DEL CHAIRÁ', 'VEGACHÍ', 'MILÁN', 'LA MACARENA',
              'RONCESVALLES', 'SARDINATA', 'BELÉN DE LOS ANDAQUÍES',
              'PUERTO GUZMÁN', 'CAMPAMENTO', 'TOLEDO', 'ORITO', 'SARAVENA',
              'LA LLANADA', 'SAN FRANCISCO', 'MONTELÍBANO', 'EL MOLINO',
              'NORCASIA', 'LANDÁZURI', 'LITORAL DEL SAN JUAN', 'HERRÁN',
              'PALMIRA', 'LA SIERRA', 'PINILLOS', 'GARZÓN', 'SAMANÁ', 'SONSÓN',
              'ROVIRA', 'PEQUE', 'MAGÚÍ', 'EL CHARCO', 'ZARAGOZA', 'CAUCASIA',
              'SAN JACINTO DEL CAUCA', 'CURUMANÍ', 'FONSECA', 'Bogotá DC',
              'BUENOS AIRES', 'ZAMBRANO', 'EL PAUJIL', 'LA JAGUA DE IBIRICO',
              'RIOSUCIO', 'SAN RAFAEL', 'PÁCORÁ', 'CHIRIGUANÁ', 'EL CERRITO',
              'SAN ANTONIO', 'EL DONCELLO', 'SANTA MARTA', 'LABATECA', 'COCORNÁ',
              'FORTUL', 'CISNEROS', 'VILLAVICENCIO', 'CHIGORODÓ', 'TIQUISIO',
              'TURBO', 'PÁEZ', 'OVEJAS', 'NECHÍ', 'CALIFORNIA', 'PELAYA',
              'URUMITA', 'PIJAO', 'DAGUA', 'SEGOVIA', 'MIRANDA', 'DOLORES',
              'VILLARRICA', 'BUENAVENTURA', 'ICONONZO', 'AGUACHICA', 'MUTATÁ',
              'FRONTINO', 'SAN JOSÉ DEL PALMAR', 'CAREPA', 'EL CARMEN',
              'VILLANUEVA', 'EL PEÑÓN', 'PIEDECUESTA', 'BARRANCABERMEJA',
              'GALÁN', 'NARIÑO', 'CHITA', 'SUCRE', 'MORALES',
              'SAN VICENTE DE CHUCURÍ', 'MATANZA', 'MANZANARES', 'PATÍA',
              'QUIBDÓ', 'CABRERA', 'BUCARAMANGA', 'TIMBÍO', 'YARUMAL', 'CÓRDOBA',
              'MALLAMA', 'CHAPARRAL', 'ROBERTO PAYÁN', 'OCAÑA', 'CURILLO',
              'CALIMA', 'OTANCHE', 'MARULANDA', 'MAJAGUAL', 'LABRANZAGRANDE',
              'LA PALMA', 'FÓMEQUE', 'SAN SEBASTIÁN', 'VALPARAÍSO', 'ORTEGA',
              'EL CARMEN DEL ATRATO', 'SANTIAGO', 'MURINDÓ', 'TOPAIPÍ',
              'POPAYÁN', 'ÚTICA', 'BECERRIL', 'LA BELLEZA', 'LA JAGUA DEL PILAR',
              'CARMEN DE VIBORAL', 'GUATAPÉ', 'VILLAHERMOSA', 'PENSILVANIA',
              'PAJARITO', 'GUAYABAL DE SÍQUIMA', 'ABEJORRAL', 'YOLOMBÓ',
              'CIMITARRA', 'ALEJANDRÍA', 'ARBOLEDAS', 'BUGA', 'YONDÓ',
              'SAN JUANITO', 'REMEDIOS', 'CAÑASGORDAS', 'SALAMINA', 'LA PEÑA',
              'BARRANCOMINAS', 'SUAZA', 'LA VEGA', 'RIONEGRO', 'TAURAMENA',
              'CONCEPCIÓN', 'LA CUMBRE', 'SESQUILÉ', 'INZÁ', 'ACEVEDO', 'YOPAL',
              'UNE', 'CARAMANTA', 'CRAVO NORTE', 'CURITÍ',
              'SANTANDER DE QUILICHAO', 'SANTUARIO', 'HATO COROZAL', 'HERVEO',
              'SALAZAR', 'PUEBLO BELLO', 'TOTORÓ', 'MONTENEGRO', 'VILLAGARZÓN',
              'SAN JACINTO', 'SANTO DOMINGO', 'EL ZULIA', 'SILVIA',
              'AGUSTÍN CODAZZI', 'NÓVITA', 'DUITAMA', 'MORELIA', 'TASCO',
              'MONTECRISTO', 'BOLÍVAR', 'SANTA ROSA DE OSOS', 'PEÑOL', 'GIRÓN',
              'AGUADAS', 'ANZOÁTEGUI', 'AQUITANIA', 'EL COCUY', 'PUERTO GAITÁN',
              'CALI', 'CARTAGENA DE INDIAS', 'BUCARASICA', 'GUTIÉRREZ', 'PRADO',
              'SURATÁ', 'SAN BERNARDO', 'SAN JUAN NEPOMUCENO', 'EL CALVARIO',
              'CAJAMARCA', 'SIMACOTA', 'JAMBALÓ', 'DIBULLA', 'PIENDAMÓ',
              'JARDÍN', 'ACACÍAS', 'FUNDACIÓN', 'CHAGUANÍ', 'EL COPEY',
              'CHÁMEZA', 'ATACO', 'CIÉNAGA', 'JURADÓ', 'VILLAGÓMEZ', 'GIGANTE',
              'DURANIA', 'ALBÁN', 'PISBA', 'PAUNA', 'PURACÉ', 'CAJIBÍO',
              'PUERTO LLERAS', 'CÚCUTA', 'CHALÁN', 'BAGADÓ', 'LA CRUZ',
              'SAN LORENZO', 'GUADUAS', 'MISTRATÓ', 'LA DORADA', 'PULÍ', 'YALÍ',
              'LINARES', 'CÁCHIRA', 'MANIZALES', 'ZAPATOCA', 'SUPÍA', 'MOCOA',
              'TADÓ', 'LÉRIDA', 'CARACOLÍ', 'NATAGAIMA', 'REGIDOR', 'UNGUÍA',
              'HATO', 'CHOCONTÁ', 'ACHÍ', 'PUERTO CONCORDIA',
              'SANTA FE ANTIOQUIA', 'GACHETÁ', 'CHISCAS', 'MONTEBELLO',
```

'GUAYABETAL', 'VENADILLO', 'PUERTO WILCHES', 'VIANÍ', 'SUAITA',
'MACHETÁ', 'BAHÍA SOLANO', 'QUETAME', 'CHARTA', 'BELLO', 'VÉLEZ',
'BARRANQUILLA', 'QUINCHÍA', 'ALVARADO', 'LA ESTRELLA', 'BOJAYÁ',
'QUEBRADANEGRA', 'EL PLAYÓN', 'ARACATACA', 'SOLITA', 'TÚQUERRES',
'PLATO', 'RÍOVIEJO', 'SAN VICENTE', 'PEREIRA', 'GAMA', 'SEVILLA',
'GÉNOVA', 'RIOFRÍO', 'LÍBANO', 'COLOSÓ', 'EL GUAMO', 'MACARAVITA',
'IBAGUÉ', 'SAN MARTÍN DE LOBA', 'NILO', 'SAN ROQUE', 'LA UNIÓN',
'TÁMARA', 'SAN ONOFRE', 'PUERTO CARREÑO', 'ANSERMA',
'SABANAGRANDE', 'SAN MARTÍN', 'LA SALINA', 'PAILITAS',
'VILLAPINZÓN', 'ALMAGUER', 'BURITICÁ', 'GUÁTICA', 'GARAGOA',
'SAN JUAN DEL CESAR', 'LEBRÍJA', 'BELALCÁZAR', 'BARRANCO DE LOBA',
'SASAIMA', 'PUEBLO RICO', 'COLÓN', 'CÁQUEZA', 'PANDI',
'SAN JUAN DE RIOSECO', 'CAPARRAPÍ', 'GALERAS', 'EL ÁGUILA',
'JAMUNDÍ', 'CUMBAL', 'SAN JERÓNIMO', 'LA PLATA', 'LOS PALMITOS',
'TÁMESIS', 'GUARNE', 'MORROA', 'PITALITO', 'PUERTO BERRÍO',
'SOCORRO', 'COYAIMA', 'MONTERÍA', 'ABRIAQUÍ', 'GÓMEZ PLATA',
'VERGARA', 'ANZÁ', 'FUENTE DE ORO', 'INIRIDA', 'RESTREPO',
'SAN ALBERTO', 'LA ARGENTINA', 'BELTRÁN', 'LIBORINA',
'SANTA ROSA DE CABAL', 'PAYA', 'SAN EDUARDO', 'GUATAVITA',
'SOGAMOSO', 'GUARANDA', 'QUÍPAMA', 'BELÉN', 'MEDIO ATRATO',
'LLORÓ', 'EL ROSARIO', 'COROZAL', 'GUACARÍ', 'SANTA MARÍA',
'LEIVA', 'CHIPAQUE', 'SABANALARGA', 'MANAURE BALCÓN DEL CESAR',
'NUQUÍ', 'ALBANIA', 'SIPÍ', 'CaRDOBA', 'CERTEGUI', 'SANDONÁ',
'MERCADERES', 'RIVERA', 'ISTMINA', 'BELÉN DE UMBRÍA', 'CUBARÁ',
'ACANDÍ', 'PUERTO SANTANDER', 'ALTO BAUDÓ', 'LA PAZ', 'FILADELFIA',
'ZONA BANANERA', 'RIO IRÓ', 'MACEO', 'NOROSÍ', 'SAN JUAN DE URABÁ',
'EL CANTÓN DEL SAN PABLO', 'RIO QUITO', 'BAJO BAUDÓ', 'IQUIRA',
'PIAMONTE', 'OLAYA HERRERA', 'MEDIO SAN JUAN', 'MEDIO BAUDÓ',
'GUAPI', 'PINCHOTE', 'EL DOVIO'], dtype=object)

In []: registro_map['codigodanemunicipio'].unique()

```
Out[ ]: array(['5,313', '18,410', '86,568', '23,580', '52,079', '5,790', '5,854',
              '19,821', '52,699', '52,612', '95,025', '50,683', '54,810',
              '13,670', '86,757', '19,780', '13,244', '50,325', '41,020',
              '5,361', '76,275', '41,078', '23,807', '13,780', '97,161',
              '54,800', '54,344', '5,120', '81,591', '52,678', '81,794',
              '95,200', '73,616', '23,189', '95,015', '5,649', '5,045', '54,398',
              '86,865', '18,756', '54,670', '5,647', '85,010', '41,359',
              '81,065', '18,592', '18,001', '5,001', '41,206', '54,250',
              '50,590', '19,075', '13,688', '50,370', '5,660', '18,610',
              '76,563', '41,524', '5,040', '95,001', '50,711', '54,003',
              '19,212', '5,847', '76,834', '50,330', '19,256', '99,773',
              '19,142', '41,668', '5,107', '27,150', '76,113', '13,160',
              '54,206', '5,873', '5,490', '68,655', '85,250', '5,234', '52,573',
              '41,001', '5,315', '76,670', '73,555', '86,569', '86,573',
              '18,753', '19,418', '50,400', '73,449', '20,001', '97,001',
              '19,701', '44,847', '44,001', '52,418', '5,250', '5,038', '52,540',
              '52,696', '52,835', '19,137', '19,050', '5,591', '41,799',
              '81,001', '50,251', '50,270', '13,042', '25,535', '50,318',
              '52,356', '13,744', '50,223', '41,132', '5,031', '85,315',
              '52,233', '18,150', '5,858', '18,460', '50,350', '73,622',
              '54,720', '18,094', '86,571', '5,134', '54,820', '86,320',
              '81,736', '52,385', '5,652', '23,466', '44,110', '17,495',
              '68,385', '27,250', '54,347', '76,520', '19,392', '50,313',
              '13,549', '41,298', '17,662', '5,756', '73,624', '5,543', '52,427',
              '52,250', '5,895', '5,154', '13,655', '20,228', '44,279', '11,001',
              '19,110', '13,894', '18,256', '20,400', '27,615', '5,667',
              '17,513', '20,178', '76,248', '73,675', '18,247', '47,001',
              '54,377', '5,197', '81,300', '5,190', '50,001', '5,172', '13,810',
              '5,837', '19,517', '70,508', '5,495', '68,132', '20,550', '44,855',
              '63,548', '76,233', '5,736', '19,455', '73,236', '73,873',
              '76,109', '73,352', '20,011', '5,480', '5,284', '27,660', '5,147',
              '68,235', '13,873', '68,250', '68,547', '68,081', '68,296',
              '5,483', '15,183', '68,773', '13,473', '68,689', '68,444',
              '17,433', '19,532', '27,001', '25,120', '68,001', '19,807',
              '5,887', '63,212', '52,435', '44,874', '5,055', '73,168', '52,621',
              '54,245', '54,498', '18,205', '76,126', '15,507', '17,446',
              '70,429', '15,377', '25,394', '25,279', '19,693', '18,860',
              '73,504', '27,245', '54,680', '5,475', '25,823', '17,614',
              '19,001', '25,851', '20,045', '68,377', '44,420', '5,148', '5,321',
              '73,870', '17,541', '68,686', '15,518', '25,328', '5,002', '5,890',
              '68,190', '70,771', '5,021', '54,051', '76,111', '5,893', '50,686',
              '5,604', '5,138', '17,653', '25,398', '94,343', '41,770', '19,397',
              '68,615', '85,410', '5,206', '76,377', '25,736', '19,355',
              '41,006', '85,001', '25,845', '5,145', '81,220', '68,229',
              '19,698', '5,697', '85,125', '73,347', '54,660', '20,570',
              '19,824', '63,470', '86,885', '13,654', '5,690', '54,261',
              '19,743', '20,013', '27,491', '15,238', '18,479', '15,790',
              '13,458', '19,100', '5,686', '5,541', '68,307', '17,013', '73,043',
              '15,047', '15,244', '50,568', '76,001', '13,001', '54,109',
              '25,339', '68,320', '73,563', '68,780', '25,649', '13,657',
              '50,245', '73,124', '68,745', '86,760', '19,364', '44,090',
              '19,548', '5,364', '50,006', '47,288', '25,168', '20,238',
              '85,015', '73,067', '47,189', '27,372', '25,871', '41,306',
              '54,239', '13,212', '25,019', '15,550', '15,531', '19,585',
              '19,130', '50,577', '54,001', '70,230', '25,258', '27,073',
              '52,378', '52,687', '25,320', '66,456', '17,380', '25,580',
              '5,885', '52,411', '54,128', '17,001', '68,895', '17,777',
              '86,001', '27,787', '73,408', '5,142', '73,483', '13,580',
              '27,800', '68,344', '25,183', '13,006', '66,687', '50,450',
              '5,042', '25,297', '15,180', '5,467', '25,335', '73,861', '68,575',
              '25,867', '68,770', '25,426', '27,075', '25,594', '68,169',
              '5,088', '68,861', '8,001', '66,594', '73,026', '5,380', '27,099',
              '25,592', '68,255', '47,053', '18,785', '52,838', '47,555',
              '13,600', '5,674', '66,001', '25,299', '76,736', '63,302',
              '76,616', '15,514', '73,411', '70,204', '13,248', '68,425',
              '73,001', '13,667', '25,488', '5,670', '70,400', '85,400',
              '70,713', '99,001', '17,042', '8,634', '20,770', '85,136',
              '20,517', '25,873', '19,022', '5,113', '66,318', '5,615', '15,299',
              '44,650', '68,406', '17,088', '13,074', '25,718', '66,572',
              '52,203', '25,151', '25,524', '25,662', '25,148', '70,235',
              '76,243', '76,364', '5,400', '52,227', '5,656', '41,396', '70,418',
              '5,789', '5,318', '70,473', '41,551', '5,579', '68,755', '73,217',
              '23,001', '5,004', '5,310', '25,862', '5,044', '50,287', '94,001',
              '19,290', '76,606', '20,710', '41,378', '25,086', '5,411',
              '66,682', '15,533', '15,660', '25,326', '15,759', '70,265',
```



```
'15,580', '52,693', '52,083', '27,425', '27,413', '52,256',
'70,215', '76,318', '41,676', '5,819', '52,405', '25,178', '5,628',
'20,443', '27,495', '44,035', '19,473', '27,745', '52,215',
'27,160', '52,683', '19,450', '41,615', '27,361', '66,088',
'15,223', '27,006', '91,669', '27,025', '20,621', '17,272',
'47,980', '27,580', '5,425', '13,490', '52,685', '5,659', '27,135',
'27,600', '27,077', '41,357', '19,533', '52,490', '27,450',
'27,430', '19,318', '68,549', '76,250'], dtype=object)
```

```
In [ ]: registro_map['tipoarea'].unique()
```

```
Out[ ]: array(['Rural', 'Urbana'], dtype=object)
```

```
In [ ]: registro_map['sitio'].unique()
```

```
Out[ ]: array(['Sin información', 'Vereda Santa Lucia', 'Quebrada Iraca', ...,
'VEREDA SAN JOSE LOS ROBLES', 'VEREDA EL ASUAY',
'RESGUARDO DE VEGAS CHAGÜI - CHIMBUZA'], dtype=object)
```

```
In [ ]: registro_map['ano'].unique()
```

```
Out[ ]: array(['2,006', '2,009', '2,008', '2,010', '1,993', '2,000', '2,003',
'2,004', '2,007', '2,005', '2,001', '2,011', '1,996', '2,002',
'1,999', '1,998', '1,994', '1,992', '1,997', '1,991', '1,995',
'1,990', '2,015', '2,013', '2,012', '2,014', '2,016', '2,017',
'2,018', '2,019', '2,020', '2,021', '2,022', '1,984', '2,023'],
dtype=object)
```

```
In [ ]: registro_map['mes'].unique()
```

```
Out[ ]: array([ 1, 12,  4, 11,  3,  8,  7, 10,  9,  2,  5,  6], dtype=int64)
```

```
In [ ]: registro_map['rangoedad'].unique()
```

```
Out[ ]: array(['Mayor de 18 años', 'Menor de 18 años'], dtype=object)
```

```
In [ ]: registro_map['grupoetnico'].unique()
```

```
Out[ ]: array(['No', 'Indigena', 'Afrodescendiente'], dtype=object)
```

```
In [ ]: registro_map['condicion'].unique()
```

```
Out[ ]: array(['Civil', 'Fuerza pública'], dtype=object)
```

```
In [ ]: registro_map['estado'].unique()
```

```
Out[ ]: array(['Hombre', 'Mujer', 'Sin Información'], dtype=object)
```

```
In [ ]: registro_map['genero'].unique()
```

```
Out[ ]: array(['Herido', 'Muerto'], dtype=object)
```

```
In [ ]: registro_map['latitudcabecera'].unique()
```

```
Out[ ]: array([6.143703 , 1.49494421, 0.502569 , ..., 8.65833333, 0.82981043,
1.19076306])
```

```
In [ ]: registro_map['longitudcabecera'].unique()
```

```
Out[ ]: array([-75.185183 , -75.46308951, -76.49818 , ..., -73.14694444,
-77.62976048, -78.11700694])
```

```
In [ ]: registro_map['tipoevento'].unique()
```

```
Out[ ]: array(['Sin informacion', 'En cerca a edificio residencial', 'Caminos',
'Cerros', 'en sobre cerca rivera de un rio', 'Canos', 'Quebradas',
'Sendero', 'Viviendas abandonadas', 'Campo', 'Fincas', 'Puentes',
'Campamentos ilegales', 'Infraestructura de comunicaciones',
'Carretera', 'Oleoductos', 'En cerca instalacion militar',
'Borde de la carretera', 'Escuelas', 'Infra torres electricas',
'Centro de salud', 'Infra gaseoducto',
'Centros deportivos recreacion', 'En cerca edificio gubernamental'],
dtype=object)
```

```
In [ ]: registro_map['Ubicación'].unique()
```

```
Out[ ]: array(['(6.143703, -75.185183)', '(1.49494421255, -75.4630895064)',
              '(0.502569, -76.49818)', ...,
              '(8.6583333333333, -73.146944444444)',
              '(0.82981043391, -77.6297604828)',
              '(1.19076305555556, -78.1170069444444)'], dtype=object)
```

```
In [ ]: registro_map['Actividad'].unique()
```

```
Out[ ]: array(['Pasando, estando cerca', 'Erradicador manual - PCI',
              'Fuerza pública', 'Sin Información', 'Transito por via usual',
              'Cuidado de animales', 'Trabajos de casa',
              'Transito por trocha atajo', 'Otra', 'Agricultura',
              'Recogiendo comida, agua ó madera',
              'Erradicacion de cultivos ilicitos', 'Jugando', 'Caza, pesca',
              'Viajando', 'Manipulación', 'Labores de desminado',
              'Mantenimiento de infraestructura', 'Actividades fuerza pública',
              'Labores de rescate', 'Erradicación - PONAL'], dtype=object)
```

3- Limpieza de la información.

Realizamos una funcion que nos permita realizar la limpieza de los valores unicos de cada variable, las cuales se guardaran en un nuevo dataset el cual denominaremos como registro_map_limpia.

```
In [ ]: def limpiar_texto(texto):
        # Normalizar texto eliminando diacríticos (tildes, acentos)
        texto_normalizado = unicodedata.normalize('NFKD', texto).encode('ASCII', 'ignore').decode('ASCII')

        # Utilizar expresiones regulares para eliminar caracteres específicos
        texto_limpio = re.sub(r'^\w\s', '', texto_normalizado)

        return texto_limpio
```

```
In [ ]: def limpiar_tabla(tabla):
        """Función para limpiar cada valor de cada columna en una tabla."""
        tabla_limpia = {}

        for columna, valores in tabla.items():
            valores_limpio = [limpiar_texto(str(valor)) for valor in valores]
            tabla_limpia[columna] = valores_limpio

        # Convertir el diccionario a un DataFrame de pandas
        df = pd.DataFrame(tabla_limpia)

        return df
```

```
In [ ]: registro_map_limpia = limpiar_tabla(registro_map)
```

```
In [ ]: registro_map_limpia.head()
```

	departamento	codigodanede	departamento	municipio	codigodanemunicipio	tipoarea	síto	ano	mes	rangoedad	grupoetnico	condicion	estado	genero	latitudcabecera	longitudcabecera	tipoevento	Ubicación	Actividad
0	ANTIOQUIA	5		GRANADA	5313	Rural	Sin informacion	2006	1	Mayor de 18 anos	No	Civil	Hombre	Herido	6143703	75185183	Sin informacion	6143703 75185183	Pasando estando cerca
1	CAQUETA	18		MONTANITA	18410	Rural	Sin informacion	2009	12	Mayor de 18 anos	No	Civil	Hombre	Herido	149494421255	754630895064	Sin informacion	149494421255 754630895064	Erradicador manual PCI
2	PUTUMAYO	86		PUERTO ASIS	86568	Rural	Sin informacion	2008	4	Mayor de 18 anos	No	Civil	Hombre	Muerto	0502569	7649818	Sin informacion	0502569 7649818	Erradicador manual PCI
3	CORDOBA	23		PUERTO LIBERTADOR	23580	Rural	Sin informacion	2009	11	Mayor de 18 anos	No	Civil	Hombre	Herido	7883545	75670508	Sin informacion	7883545 75670508	Erradicador manual PCI
4	CORDOBA	23		PUERTO LIBERTADOR	23580	Rural	Sin informacion	2009	11	Mayor de 18 anos	No	Civil	Hombre	Herido	7883545	75670508	Sin informacion	7883545 75670508	Erradicador manual PCI

A continuación, realizaremos verificación de la limpieza de la información:

```
In [ ]: registro_map_limpia['departamento'].unique()
```

```
Out[ ]: array(['ANTIOQUIA', 'CAQUETA', 'PUTUMAYO', 'CORDOBA', 'NARINO', 'CAUCA',
              'GUAVIARE', 'META', 'NORTE DE SANTANDER', 'BOLIVAR', 'HUILA',
              'VALLE DEL CAUCA', 'VAUPES', 'ARAUCA', 'TOLIMA', 'CASANARE',
              'VICHADA', 'CHOCO', 'SANTANDER', 'CESAR', 'LA GUAJIRA',
              'CUNDINAMARCA', 'CALDAS', 'BOGOTA DC', 'MAGDALENA', 'SUCRE',
              'QUINDIO', 'BOYACA', 'GUAINIA', 'RISARALDA', 'ATLANTICO',
              'AMAZONAS'], dtype=object)

In [ ]: registro_map_limpia['municipio'].unique()
```

```
Out[ ]: array(['GRANADA', 'MONTANITA', 'PUERTO ASIS', 'PUERTO LIBERTADOR',
'BARBACOAS', 'TARAZA', 'VALDIVIA', 'TORIBIO', 'SANTA CRUZ',
'RICAUORTE', 'EL RETORNO', 'SAN JUAN DE ARAMA', 'TIBU', 'SAN PABLO',
'SAN MIGUEL', 'SUAREZ', 'EL CARMEN DE BOLIVAR', 'MAPIRIPAN',
'ALGECIRAS', 'ITUANGO', 'FLORIDA', 'BARAYA', 'TIERRALTA',
'TALAIGUA NUEVO', 'CARURU', 'TEORAMA', 'HACARI', 'CACERES',
'PUERTO RONDON', 'SAMANIEGO', 'TAME', 'MIRAFLORES', 'RIOBLANCO',
'CIENAGA DE ORO', 'CALAMAR', 'SAN CARLOS', 'APARTADO', 'LA PLAYA',
'VALLE DEL GUAMUEZ', 'SOLANO', 'SAN CALIXTO', 'SAN ANDRES',
'AGUAZUL', 'ISNOS', 'ARAUQUITA', 'PUERTO RICO', 'FLORENCIA',
'MEDELLIN', 'COLOMBIA', 'EL TARRA', 'BALBOA', 'SANTA ROSA DEL SUR',
'URIBE', 'SAN LUIS', 'SAN JOSE DEL FRAGUA', 'PRADERA', 'PALERMO',
'ANORI', 'SAN JOSE DEL GUAVIARE', 'VISTAHERMOSA', 'ABREGO',
'CORINTO', 'URRAO', 'TULUA', 'MESETAS', 'EL TAMBO', 'CUMARIBO',
'CALOTO', 'SAN AGUSTIN', 'BRICENO', 'EL CARMEN DEL DARIEN',
'BUGALAGRANDE', 'CANTAGALLO', 'CONVENCION', 'VIGIA DEL FUERTE',
'NECOCLI', 'SABANA DE TORRES', 'PAZ DE ARIPORO', 'DABEIBA',
'PUERRES', 'NEIVA', 'GUADALUPE', 'SAN PEDRO', 'PLANADAS',
'PUERTO CAICEDO', 'PUERTO LEGUIZAMO', 'SAN VICENTE DEL CAGUAN',
'LOPEZ', 'LEJANIAS', 'MELGAR', 'VALLEDUPAR', 'MITU', 'SANTA ROSA',
'URIBIA', 'RIOHACHA', 'LOS ANDES', 'EL BAGRE', 'ANGOSTURA',
'POLICARPA', 'SANTA BARBARA', 'TUMACO', 'CALDONO', 'ARGELIA',
'PUERTO TRIUNFO', 'TELLO', 'ARAUCA', 'EL CASTILLO', 'EL DORADO',
'ARENAL', 'PASCA', 'GUAMAL', 'IPIALES', 'SIMITI', 'CUBARRAL',
'CAMPOALEGRE', 'AMALFI', 'SACAMA', 'CUMBITARA',
'CARTAGENA DEL CHAIRA', 'VEGACHI', 'MILAN', 'LA MACARENA',
'RONCESVALLES', 'SARDINATA', 'BELEN DE LOS ANDAQUIES',
'PUERTO GUZMAN', 'CAMPAMENTO', 'TOLEDO', 'ORITO', 'SARAVENA',
'LA LLANADA', 'SAN FRANCISCO', 'MONTELIBANO', 'EL MOLINO',
'NORCASIA', 'LANDAZURI', 'LITORAL DEL SAN JUAN', 'HERRAN',
'PALMIRA', 'LA SIERRA', 'PINILLOS', 'GARZON', 'SAMANA', 'SONSON',
'ROVIRA', 'PEQUE', 'MAGUI', 'EL CHARCO', 'ZARAGOZA', 'CAUCASIA',
'SAN JACINTO DEL CAUCA', 'CURUMANI', 'FONSECA', 'Bogota DC',
'BUENOS AIRES', 'ZAMBRANO', 'EL PAUJIL', 'LA JAGUA DE IBIRICO',
'RIOSUCIO', 'SAN RAFAEL', 'PACORA', 'CHIRIGUANA', 'EL CERRITO',
'SAN ANTONIO', 'EL DONCELLO', 'SANTA MARTA', 'LABATECA', 'COCORNA',
'FORTUL', 'CISNEROS', 'VILLAVICENCIO', 'CHIGORODO', 'TIQUISIO',
'TURBO', 'PAEZ', 'OVEJAS', 'NECHI', 'CALIFORNIA', 'PELAYA',
'URUMITA', 'PIJAO', 'DAGUA', 'SEGOVIA', 'MIRANDA', 'DOLORES',
'VILLARRICA', 'BUENAVENTURA', 'ICONONZO', 'AGUACHICA', 'MUTATA',
'FRONTINO', 'SAN JOSE DEL PALMAR', 'CAREPA', 'EL CARMEN',
'VILLANUEVA', 'EL PENON', 'PIEDRECUESTA', 'BARRANCABERMEJA',
'GALAN', 'NARINO', 'CHITA', 'SUCRE', 'MORALES',
'SAN VICENTE DE CHUCURI', 'MATANZA', 'MANZANARES', 'PATIA',
'QUIBDO', 'CABRERA', 'BUCARAMANGA', 'TIMBIO', 'YARUMAL', 'CORDOBA',
'MALLAMA', 'CHAPARRAL', 'ROBERTO PAYAN', 'OCANA', 'CURILLO',
'CALIMA', 'OTANCHE', 'MARULANDA', 'MAJAGUAL', 'LABRANZAGRANDE',
'LA PALMA', 'FOMEQUE', 'SAN SEBASTIAN', 'VALPARAISO', 'ORTEGA',
'EL CARMEN DEL ATRATO', 'SANTIAGO', 'MURINDO', 'TOPAIPI',
'POPAYAN', 'UTICA', 'BECERRIL', 'LA BELLEZA', 'LA JAGUA DEL PILAR',
'CARMEN DE VIBORAL', 'GUATAPE', 'VILLAHERMOSA', 'PENSILVANIA',
'PAJARITO', 'GUAYABAL DE SIQUIMA', 'ABEJORRAL', 'YOLOMBO',
'CIMITARRA', 'ALEJANDRIA', 'ARBOLEDAS', 'BUGA', 'YONDO',
'SAN JUANITO', 'REMEDIOS', 'CANASGORDAS', 'SALAMINA', 'LA PENAS',
'BARRANCOMINAS', 'SUAZA', 'LA VEGA', 'RIONEGRO', 'TAURAMENA',
'CONCEPCION', 'LA CUMBRE', 'SESQUILE', 'INZA', 'ACEVEDO', 'YOPAL',
'UNE', 'CARAMANTA', 'CRAVO NORTE', 'CURITI',
'SANTANDER DE QUILICHAO', 'SANTUARIO', 'HATO COROZAL', 'HERVEO',
'SALAZAR', 'PUEBLO BELLO', 'TOTORO', 'MONTENEGRO', 'VILLAGARZON',
'SAN JACINTO', 'SANTO DOMINGO', 'EL ZULIA', 'SILVIA',
'AGUSTIN CODAZZI', 'NOVITA', 'DUISTAMA', 'MORELIA', 'TASCO',
'MONTECRISTO', 'BOLIVAR', 'SANTA ROSA DE OSOS', 'PENOL', 'GIRON',
'AGUADAS', 'ANZOATEGUI', 'AQUITANIA', 'EL COCUY', 'PUERTO GAITAN',
'CALI', 'CARTAGENA DE INDIAS', 'BUCARASICA', 'GUTIERREZ', 'PRADO',
'SURATA', 'SAN BERNARDO', 'SAN JUAN NEPOMUCENO', 'EL CALVARIO',
'CAJAMARCA', 'SIMACOTA', 'JAMBALO', 'DIBULLA', 'PIENDAMO',
'JARDIN', 'ACACIAS', 'FUNDACION', 'CHAGUANI', 'EL COPEY',
'CHAMEZA', 'ATACO', 'CIENAGA', 'JURADO', 'VILLAGOMEZ', 'GIGANTE',
'DURANIA', 'ALBAN', 'PISBA', 'PAUNA', 'PURACE', 'CAJIBIO',
'PUERTO LLERAS', 'CUCUTA', 'CHALAN', 'BAGADO', 'LA CRUZ',
'SAN LORENZO', 'GUADUAS', 'MISTRATO', 'LA DORADA', 'PULI', 'YALI',
'LINARES', 'CACHIRA', 'MANIZALES', 'ZAPATOCA', 'SUPIA', 'MOCOA',
'TADO', 'LERIDA', 'CARACOLI', 'NATAGAIMA', 'REGIDOR', 'UNGUIA',
'HATO', 'CHOCONTA', 'ACHI', 'PUERTO CONCORDIA',
'SANTA FE ANTIOQUIA', 'GACHETA', 'CHISCAS', 'MONTEBELLO',
```

```
'GUAYABETAL', 'VENADILLO', 'PUERTO WILCHES', 'VIANI', 'SUAITA',
'MACHETA', 'BAHIA SOLANO', 'QUETAME', 'CHARTA', 'BELLO', 'VELEZ',
'BARRANQUILLA', 'QUINCHIA', 'ALVARADO', 'LA ESTRELLA', 'BOJAYA',
'QUEBRADANEGRA', 'EL PLAYON', 'ARACATACA', 'SOLITA', 'TUQUERRES',
'PLATO', 'RIOVIEJO', 'SAN VICENTE', 'PEREIRA', 'GAMA', 'SEVILLA',
'GENOVA', 'RIOFRIO', 'LIBANO', 'COLOSO', 'EL GUAMO', 'MACARAVITA',
'IBAGUE', 'SAN MARTIN DE LOBA', 'NILO', 'SAN ROQUE', 'LA UNION',
'TAMARA', 'SAN ONOFRE', 'PUERTO CARRENO', 'ANSERMA',
'SABANAGRANDE', 'SAN MARTIN', 'LA SALINA', 'PAILITAS',
'VILLAPINZON', 'ALMAGUER', 'BURITICA', 'GUATICA', 'GARAGOA',
'SAN JUAN DEL CESAR', 'LEBRIJA', 'BELALCAZAR', 'BARRANCO DE LOBA',
'SASAIMA', 'PUEBLO RICO', 'COLON', 'CAQUEZA', 'PANDI',
'SAN JUAN DE RIOSECO', 'CAPARRAPI', 'GALERAS', 'EL AGUILA',
'JAMUNDI', 'CUMBAL', 'SAN JERONIMO', 'LA PLATA', 'LOS PALMITOS',
'TAMESIS', 'GUARNE', 'MORROA', 'PITALITO', 'PUERTO BERRIO',
'SOCORRO', 'COYAIMA', 'MONTERIA', 'ABRIAQUI', 'GOMEZ PLATA',
'VERGARA', 'ANZA', 'FUENTE DE ORO', 'INIRIDA', 'RESTREPO',
'SAN ALBERTO', 'LA ARGENTINA', 'BELTRAN', 'LIBORINA',
'SANTA ROSA DE CABAL', 'PAYA', 'SAN EDUARDO', 'GUATAVITA',
'SOGAMOSO', 'GUARANDA', 'QUIPAMA', 'BELEN', 'MEDIO ATRATO',
'LLORO', 'EL ROSARIO', 'COROZAL', 'GUACARI', 'SANTA MARIA',
'LEIVA', 'CHIPAQUE', 'SABANALARGA', 'MANAURE BALCON DEL CESAR',
'NUQUI', 'ALBANIA', 'SIPI', 'CARDOBA', 'CERTEGUI', 'SANDONA',
'MERCADERES', 'RIVERA', 'ISTMINA', 'BELEN DE UMBRIA', 'CUBARA',
'ACANDI', 'PUERTO SANTANDER', 'ALTO BAUDO', 'LA PAZ', 'FILADELFIA',
'ZONA BANANERA', 'RIO IRO', 'MACEO', 'NOROSI', 'SAN JUAN DE URABA',
'EL CANTON DEL SAN PABLO', 'RIO QUITO', 'BAJO BAUDO', 'IQUIRA',
'PIAMONTE', 'OLAYA HERRERA', 'MEDIO SAN JUAN', 'MEDIO BAUDO',
'GUAPI', 'PINCHOTE', 'EL DOVIO'], dtype=object)
```

```
In [ ]: registro_map_limpia['sitio'].unique()
```

```
Out[ ]: array(['Sin informacion', 'Vereda Santa Lucia', 'Quebrada Iraca', ...,
               'VEREDA SAN JOSE LOS ROBLES', 'VEREDA EL ASUAY',
               'RESGUARDO DE VEGAS CHAGUI CHIMBUZA'], dtype=object)
```

```
In [ ]: registro_map_limpia['mes'].unique()
```

```
Out[ ]: array([ 1, 12,  4, 11,  3,  8,  7, 10,  9,  2,  5,  6], dtype=int64)
```

```
In [ ]: registro_map_limpia['rangoedad'].unique()
```

```
Out[ ]: array(['Mayor de 18 anos', 'Menor de 18 anos'], dtype=object)
```

```
In [ ]: registro_map_limpia['condicion'].unique()
```

```
Out[ ]: array(['Civil', 'Fuerza publica'], dtype=object)
```

```
In [ ]: registro_map_limpia['estado'].unique()
```

```
Out[ ]: array(['Hombre', 'Mujer', 'Sin Informacion'], dtype=object)
```

```
In [ ]: registro_map_limpia['genero'].unique()
```

```
Out[ ]: array(['Herido', 'Muerto'], dtype=object)
```

```
In [ ]: registro_map_limpia['tipoevento'].unique()
```

```
Out[ ]: array(['Sin informacion', 'En cerca a edificio residencial', 'Caminos',
               'Cerros', 'en sobre cerca rivera de un rio', 'Canos', 'Quebradas',
               'Sendero', 'Viviendas abandonadas', 'Campo', 'Fincas', 'Puentes',
               'Campamentos ilegales', 'Infraestructura de comunicaciones',
               'Carretera', 'Oleoductos', 'En cerca instalacion militar',
               'Borde de la carretera', 'Escuelas', 'Infra torres electricas',
               'Centro de salud', 'Infra gaseoducto',
               'Centros deportivos recreacion', 'En cerca edificio gubernamental'],
               dtype=object)
```

```
In [ ]: registro_map_limpia['Actividad'].unique()
```

```
Out[ ]: array(['Pasando estando cerca', 'Erradicador manual PCI',
        'Fuerza publica', 'Sin Informacion', 'Transito por via usual',
        'Cuidado de animales', 'Trabajos de casa',
        'Transito por trocha atajo', 'Otra', 'Agricultura',
        'Recogiendo comida agua o madera',
        'Erradicacion de cultivos ilicitos', 'Jugando', 'Caza pesca',
        'Viajando', 'Manipulacion', 'Labores de desminado',
        'Mantenimiento de infraestructura', 'Actividades fuerza publica',
        'Labores de rescate', 'Erradicacion PONAL'], dtype=object)
```

En el anterior punto se observo que las columnas "estado" y "genero" se encuentran intercambiadas, se procede hacer la corrección pertinente cambiado el nombre de cada variable.

```
In [ ]: registro_map_limpia.rename(columns={'estado': 'genero_'}, inplace=True)
registro_map_limpia.rename(columns={'genero': 'estado'}, inplace=True)
registro_map_limpia.rename(columns={'genero_': 'genero'}, inplace=True)
registro_map_limpia.rename(columns={'ano': 'anio'}, inplace=True)
registro_map_limpia.head()
```

Out []:	departamento	codigodane	departamento	municipio	codigodane	municipio	tipoarea	sitio	anio	mes	rangoedad	grupoetnico	condicion	genero	estado	latitudcabecera	longitudcabecera	tipoevento	Ubicación	Actividad
0	ANTIOQUIA	5	GRANADA	5313	Rural	Sin informacion	2006	1	Mayor de 18 años	No	Civil	Hombre	Herido		6143703	75185183	Sin informacion	6143703 75185183	Pasando estando cerca	
1	CAQUETA	18	MONTANITA	18410	Rural	Sin informacion	2009	12	Mayor de 18 años	No	Civil	Hombre	Herido		149494421255	754630895064	Sin informacion	149494421255 754630895064	Erradicador manual PCI	
2	PUTUMAYO	86	PUERTO ASIS	86568	Rural	Sin informacion	2008	4	Mayor de 18 años	No	Civil	Hombre	Muerto		0502569	7649818	Sin informacion	0502569 7649818	Erradicador manual PCI	
3	CORDOBA	23	PUERTO LIBERTADOR	23580	Rural	Sin informacion	2009	11	Mayor de 18 años	No	Civil	Hombre	Herido		7883545	75670508	Sin informacion	7883545 75670508	Erradicador manual PCI	
4	CORDOBA	23	PUERTO LIBERTADOR	23580	Rural	Sin informacion	2009	11	Mayor de 18 años	No	Civil	Hombre	Herido		7883545	75670508	Sin informacion	7883545 75670508	Erradicador manual PCI	

Los datos de la columna año se encuentran como tipo object lo que puede dificultar trabajar posteriormente con estos, por lo cual se pasan como tipo de dato de tiempo.

```
In [ ]: registro_map_limpia['anio'] = pd.to_datetime(registro_map_limpia['anio'])

Debido a que los datos de la columna "mes" pueden existir algunos string, nos aseguramos de que 'mes' sea numérico (si no lo es, se convierte)

In [ ]: registro_map_limpia['mes'] = pd.to_numeric(registro_map_limpia['mes'], errors='coerce')
```

4- Reporte de resultados.

Se presenta el resultados de los reportes siguientes sobre las víctimas:

- total víctimas por año
- total víctimas por mes (Independiente del año)
- total víctimas por departamento
- total víctimas por actividad
- total de víctimas por género
- total de víctimas por estado

4.1- Total víctimas por año.

```
In [ ]: total_victimas_x_anio = pd.DataFrame(registro_map_limpia['anio'])
total_victimas_x_anio['anio'] = pd.to_datetime(total_victimas_x_anio['anio']).dt.year
total_victimas_x_anio = total_victimas_x_anio['anio'].value_counts().sort_index()
total_victimas_x_anio
```

```
Out[ ]: anio
1984      1
1990     23
1991     70
1992    153
1993     88
1994     86
1995    130
1996    120
1997     96
1998     59
1999     54
2000    141
2001    299
2002    631
2003    760
2004    900
2005   1176
2006   1224
2007    980
2008    851
2009    771
2010    679
2011    564
2012    589
2013    427
2014    293
2015    223
2016     90
2017     59
2018    179
2019    117
2020    176
2021    157
2022    139
2023     59
Name: count, dtype: int64
```

```
In [ ]: total_victimas_x_anio.plot(kind='bar')
plt.xlabel('Año')
plt.ylabel('Victimas por año')
plt.title('Total de victimas por año')
```

Out[]: Text(0.5, 1.0, 'Total de victimas por año')



De acuerdo con el informe y la gráfica de barras, se evidencia un aumento en el número de víctimas por minas antipersona desde el año 1990. Este aumento alcanzó su punto máximo en el año 2006, con un total de 1,224 víctimas registradas. A partir de ese año, se observa una disminución progresiva en el número de víctimas.

4.2 - Total víctimas por mes (Independiente del año).

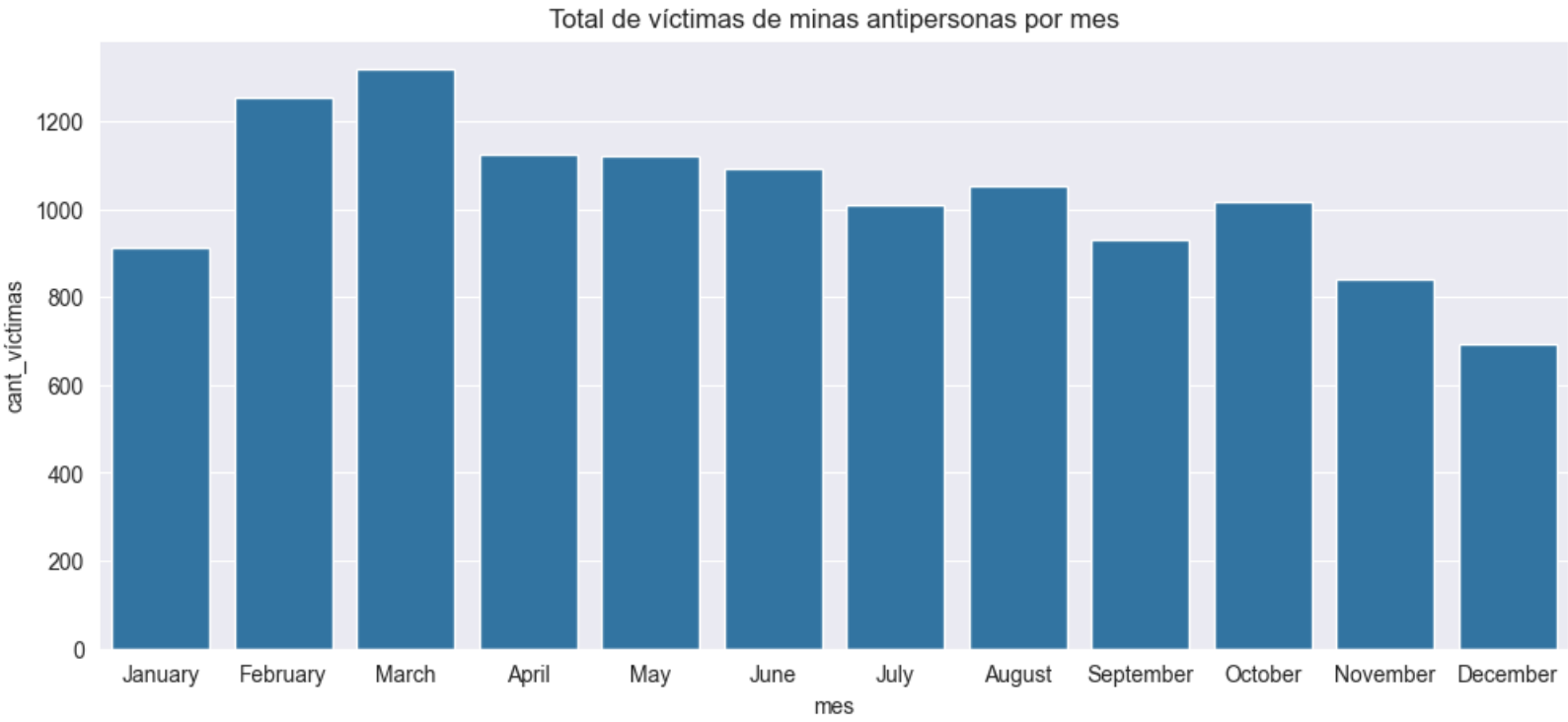
```
In [ ]: total_victimas_x_mes = pd.DataFrame(registro_map_limpia['mes'].value_counts()).reset_index()
total_victimas_x_mes.columns = ['mes', 'cant_victimas']
total_victimas_x_mes['mes'] = total_victimas_x_mes['mes'].astype(int).apply(lambda x: calendar.month_name[x])
meses_ordenados = [calendar.month_name[i] for i in range(1, 13)]
total_victimas_x_mes['mes'] = pd.Categorical(total_victimas_x_mes['mes'], categories=meses_ordenados, ordered=True)
total_victimas_x_mes = total_victimas_x_mes.sort_values(by='mes')
total_victimas_x_mes
```

Out[]:

	mes	cant_victimas
9	January	911
1	February	1255
0	March	1318
2	April	1124
3	May	1120
4	June	1093
7	July	1008
5	August	1054
8	September	930
6	October	1018
10	November	840
11	December	693

```
In [ ]: sns.set_style('darkgrid')
plt.figure(figsize=(12,5))
sns.barplot(x='mes', y='cant_victimas', data = total_victimas_x_mes)
plt.xticks( rotation=0)
plt.title('Total de víctimas de minas antipersonas por mes')
```

Out[]: Text(0.5, 1.0, 'Total de víctimas de minas antipersonas por mes')



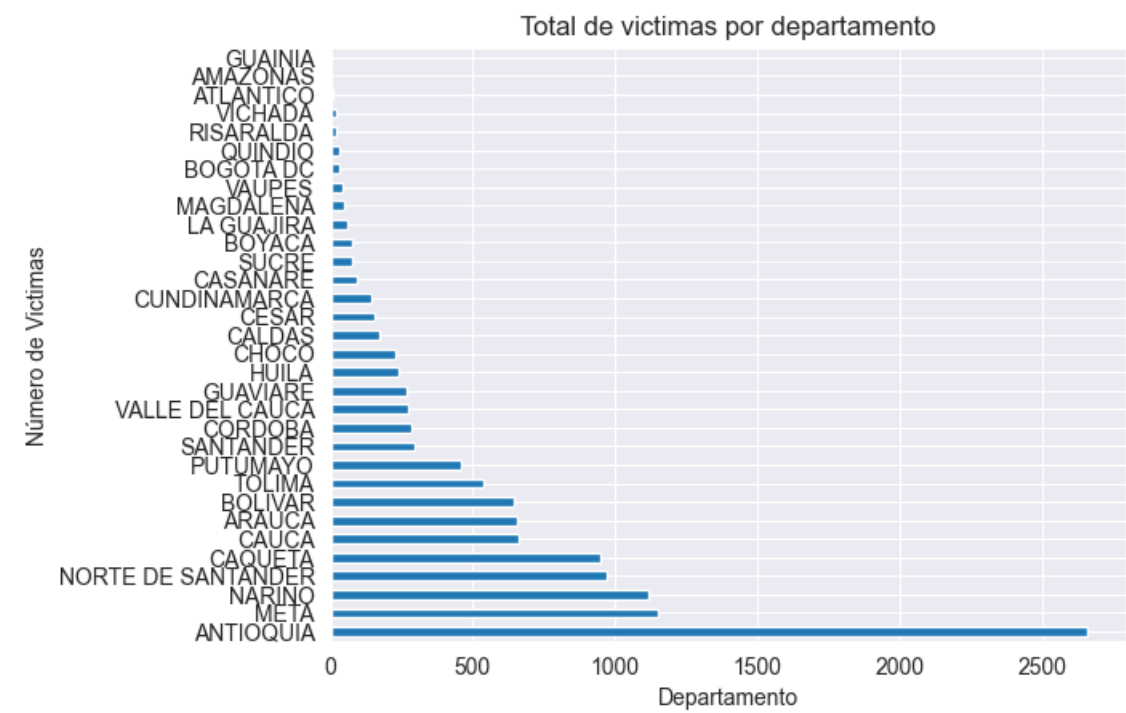
4.3 - Total víctimas por departamento.

```
In [ ]: total_victimas_x_dept = pd.DataFrame(registro_map_limpia['departamento'])
total_victimas_x_dept = total_victimas_x_dept['departamento'].value_counts()
total_victimas_x_dept
```

```
Out[ ]: departamento
ANTIOQUIA      2658
META           1152
NARINO         1115
NORTE DE SANTANDER  970
CAQUETA        946
CAUCA          662
ARAUCA         656
BOLIVAR        645
TOLIMA         539
PUTUMAYO       459
SANTANDER      297
CORDOBA        283
VALLE DEL CAUCA 271
GUAVIARE       267
HUILA          239
CHOCO          230
CALDAS         170
CESAR          156
CUNDINAMARCA   141
CASANARE       90
SUCRE          78
BOYACA         78
LA GUAJIRA     60
MAGDALENA      46
VAUPES         44
BOGOTA DC      30
QUINDIO        28
RISARALDA      22
VICHADA        17
ATLANTICO       8
AMAZONAS       4
GUAINIA        3
Name: count, dtype: int64
```

```
In [ ]: total_victimas_x_dept.plot(kind='barh')
plt.xlabel('Departamento')
plt.ylabel('Número de Víctimas')
plt.title('Total de víctimas por departamento')
```

```
Out[ ]: Text(0.5, 1.0, 'Total de víctimas por departamento')
```



4.4 - Total víctimas por actividad.

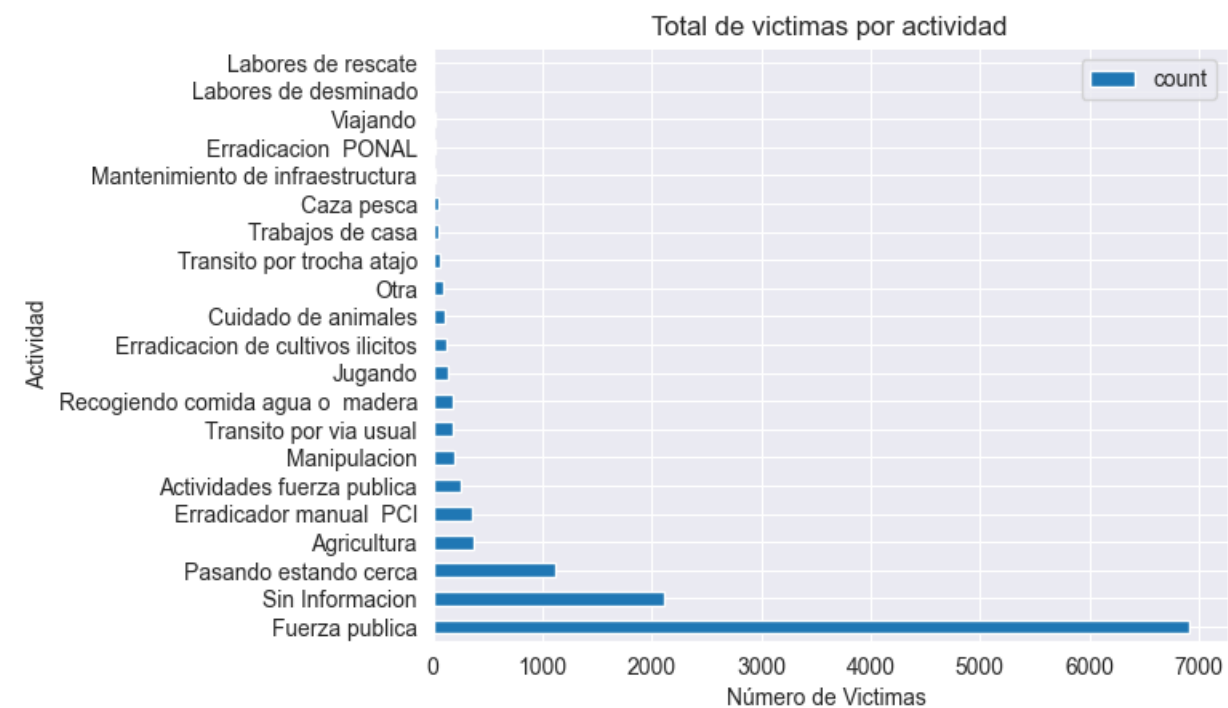
```
In [ ]: total_victimas_x_actividad = pd.DataFrame(registro_map_limpia['Actividad'])
total_victimas_x_actividad = pd.DataFrame(registro_map_limpia['Actividad'].value_counts())
total_victimas_x_actividad
```

Out[]:

Actividad	count
Fuerza publica	6913
Sin Informacion	2108
Pasando estando cerca	1112
Agricultura	366
Erradicador manual PCI	360
Actividades fuerza publica	254
Manipulacion	192
Transito por via usual	187
Recogiendo comida agua o madera	180
Jugando	136
Erradicacion de cultivos ilicitos	118
Cuidado de animales	103
Otra	98
Transito por trocha atajo	59
Trabajos de casa	56
Caza pesca	49
Mantenimiento de infraestructura	25
Erradicacion PONAL	24
Viajando	18
Labores de desminado	5
Labores de rescate	1

```
In [ ]: total_victimas_x_actividad.plot(kind='barh')
plt.xlabel('Número de Victimas ')
plt.ylabel('Actividad')
plt.title('Total de victimas por actividad')
```

Out[]: Text(0.5, 1.0, 'Total de victimas por actividad')



4.5 - Total de víctimas por género.

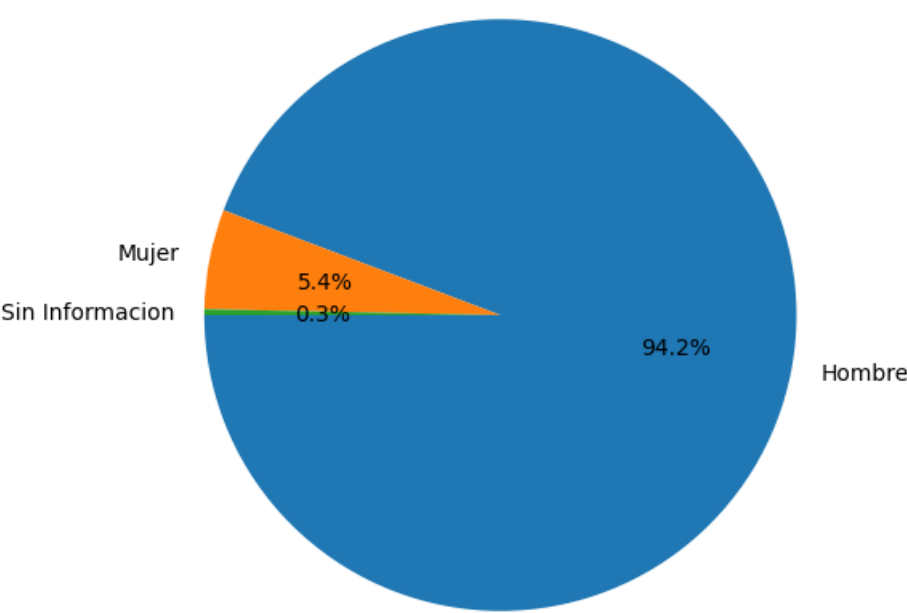
```
In [ ]: total_victimas_x_genero = pd.DataFrame(registro_map_limpia['genero'])
total_victimas_x_genero = total_victimas_x_genero['genero'].value_counts()
total_victimas_x_genero
```

```
Out[ ]: genero
Hombre      11651
Mujer        673
Sin Informacion  40
Name: count, dtype: int64
```

```
In [ ]: plt.figure(figsize=(8, 6))
total_victimas_x_genero.plot.pie(autopct='%1.1f%%', startangle=180, labels=total_victimas_x_genero.index)
plt.title('Distribución de víctimas por género')
plt.ylabel('')
```

```
Out[ ]: Text(0, 0.5, '')
```

Distribución de víctimas por género



4.6 - total de víctimas por estado.

```
In [ ]: total_victimas_x_estado = pd.DataFrame(registro_map_limpia['estado'])
total_victimas_x_estado = pd.DataFrame(registro_map_limpia['estado'].value_counts()).reset_index()
total_victimas_x_estado.columns = ['estado', 'cantidad de victimas']
total_victimas_x_estado
```

Out[]:

	estado	cantidad de victimas
0	Herido	10013
1	Muerto	2351

```
In [ ]: sns.set_style('darkgrid')
plt.figure(figsize=(12,5))
sns.barplot(x='estado', y='cantidad de victimas', data = total_victimas_x_estado)
plt.xticks( rotation=0)
plt.title('Estado de las víctimas de minas antipersonas')
```

Out[]: Text(0.5, 1.0, 'Estado de las víctimas de minas antipersonas')

Estado de las víctimas de minas antipersonas

