

Trabajo: Práctica II

Estudiante: ANDRÉS FELIPE BELLO ZAPATA

Materia: Visualización de datos

1. Título de la visualización donde se presentan la visualización realizada. URL de la visualización y del código. Y descripción corta del documento y del que se presenta

1.1. Título de la Visualización:

Comportamiento del delito de tráfico de personas en el mundo.

1.2. URL de la visualización

En el siguiente URL se encuentran todos los archivos empleados durante la práctica

https://github.com/andresfb88/Practica2_Visualizacion.git

1.3. Descripción del documento:

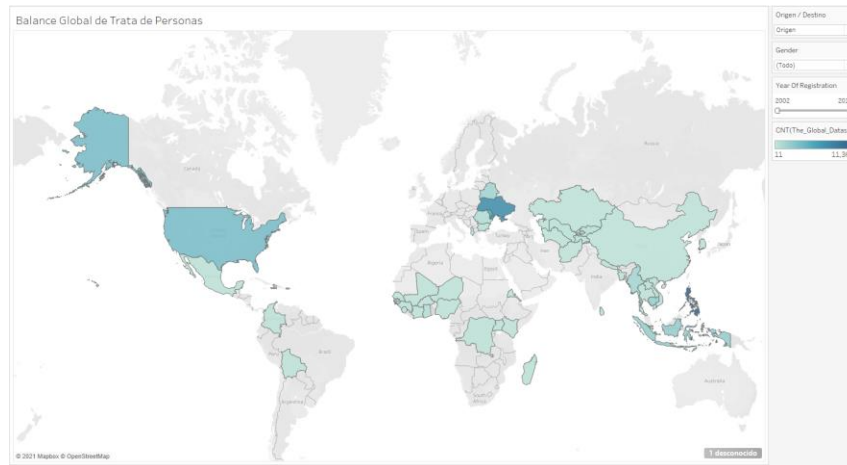
El presente documento contiene una serie de visualizaciones que de una manera interactiva permiten al usuario comprender cuál ha sido el comportamiento del delito de tráfico de personas en el mundo en el lapso del 2002 al 2019. La visualización parte inicialmente de un análisis global que da a conocer al usuario los países en los cuales es más común que se presente este delito desde dos ámbitos: por un lado, se analizan los países en los cuales es más común que se recluten personas, y por el otro se analizan los países en los cuales se es más común que las personas sean explotadas. Una vez identificado el impacto global, se analizan aspectos más particulares como el género de la persona afectada, el rango de edad, medio de coerción empleado por los criminales, la actividad que fue forzada a realizar, la relación que presentaba con su victimario, entre otros.

2. Título de la visualización donde se presentan la visualización realizada. URL de la visualización y del código. Y descripción corta del documento y del que se presenta

2.1. Que visualización se ha escogido y porque?

Con el fin de dar a conocer al usuario el comportamiento del crimen de trata de personas en el mundo, se emplean las siguientes visualizaciones

Balance global de trata de personas:

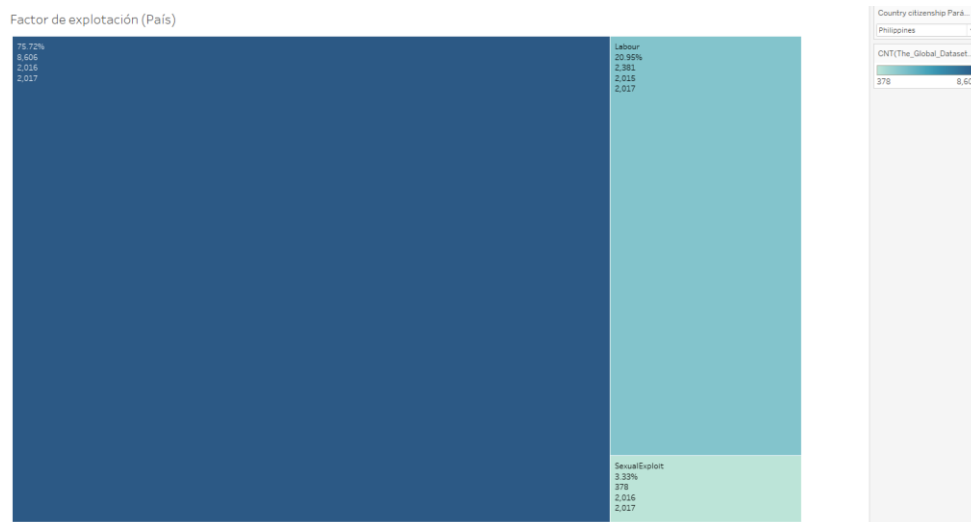


Link:

https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Visualizacin_Practica_Final/MundoOrigDest

Esta primera visualización pretende brindar al usuario una idea de la cantidad de casos de trata de personas en el mundo, relacionando con colores más oscuros a los países que presentan una mayor cantidad de casos. La visualización al ser interactiva permite al usuario escoger si desea visualizar la criticidad del crimen en relación con el país en el cual se presenta la mayor cantidad de reclutamientos, o el país en el cual se realiza mayor cantidad de explotaciones a las víctimas. Otro aspecto que el usuario puede escoger es el género de la víctima, en caso de que desee conocer particularmente la afectación a las mujeres, hombres o ambos sexos. Finalmente se da la opción al usuario de escoger un rango de años en particular, en caso de que se desee analizar el comportamiento de este delito en un periodo de tiempo en particular.

Factor de Explotación (País).



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/FactordeexplotacinPas/CountryExploitR>

La siguiente visualización interactiva está diseñada de tal forma que una vez el usuario conozca el panorama global, se incline por determinado país. Seleccionado el país de referencia, mediante la siguiente visualización de mapa de árbol, se logra identificar los tipos de explotación que generan un impacto mayor, ilustrando la cantidad de casos reportados y el porcentaje del total. La visualización también permite identificar para cada factor el año del primer caso y del último, dando a conocer al usuario cual ha sido el delito que ha afectado por más tiempo a determinada nación. Dentro de los diferentes tipos de explotación se encuentran:

- Secuestro
- Trabajo forzado
- Matrimonio Forzado
- Mercenario
- Remoción de órganos
- Sexual

Relación con el reclutador:



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Relacinconelreclutador/RecruiterRelation>

Una vez el usuario ha seleccionado determinado país y ha identificado un factor de trata de personas basado en el impacto que tiene el mismo y en los intereses particulares del usuario, se procede a identificar para determinado factor cual fue la relación existente entre la víctima y el victimario. Es impactante identificar como en algunos casos, como en el que presenta la imagen de la visualización interactiva para el país Filipinas y el factor de Trabajo Forzado, los familiares son las personas que promueven que se presenten este tipo de delitos.

Modo de control

Modo de Control

Mean Ctr	Recuent o de Th..	% de total Re..
Nulo	2,326	97.69%
NoEspeci	14	0.59%
FPromes	12	0.50%
Deuda	11	0.46%
Robo	7	0.29%
APsicolo	5	0.21%
Amenaza	3	0.13%
AFisico	1	0.04%
ASexual	1	0.04%
Documentos	1	0.04%

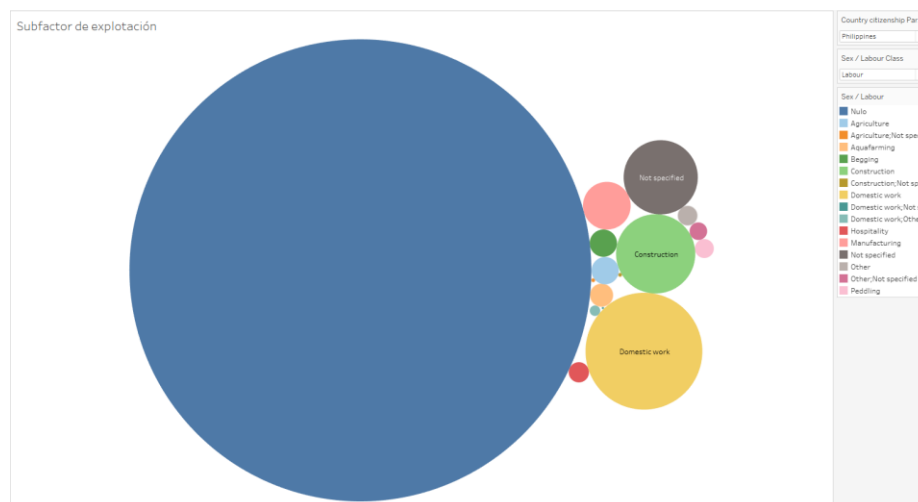
Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/MododeControl/ModoControl>

La presente visualización tiene como consideración el mismo país seleccionado en el Mapa de Árbol y el mismo método de explotación seleccionado en la visualización “Relacion con el reclutador”, con el fin de brindar al usuario información relacionada con los diferentes métodos empleados por parte de los victimarios para forzar a las víctimas a permanecer sometidos al delito. Los diferentes métodos disponibles son:

- No especificado
- Falsas Promesas
- Deudas
- Robos
- Acoso psicológico
- Amenazas
- Acoso físico
- Acoso Sexual
- Retención de documentos
- Uso de sustancias psicoactivas
- Retención física
- Restricción a accesos médicos
- Alta carga de trabajo
- Amenazas legales
- Privación de necesidades básicas

Subfactor de explotación:

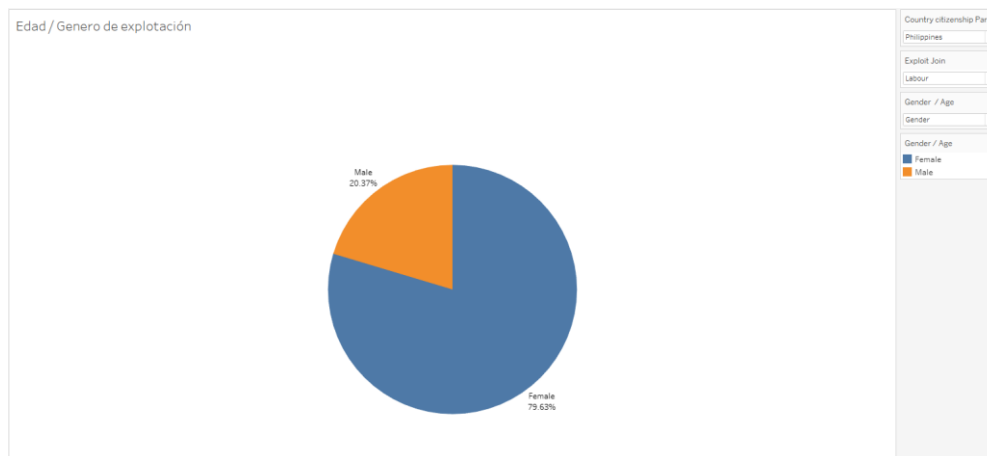


Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Subfactordeexplotacion/MainExploitFact>

Continuando con la secuencia con la que se vienen interpretando las diferentes visualizaciones, se procede ahora a determinar el subfactor de explotación, de acuerdo con el país y el factor escogido. Este subfactor esta disponible para los casos en los cuales el crimen de trata de personas está relacionado con explotación sexual y trabajo forzado. Esto se debe a que mencionados crímenes son los que se presentan de manera más común. La intención de esta gráfica radica en la relevancia de dar a conocer al usuario el tipo de trabajo forzado o de abuso sexual al que se ve sometido la víctima.

Edad / Genero de Explotación



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/EdadGenerodeExplotacin/ExplotAge>

La presente visualización interactiva, continúa con la secuencia que vienen describiendo las anteriores visualizaciones. Basado en el país y el tipo de explotación, la PIE CHART permite identificar 3 aspectos relevantes. El primero de ellos hace referencia a la proporción de género, de tal forma que el usuario este en capacidad de identificar cual de los dos es el más afectado (Male y Female). En la segunda opción de visualización se identifica el rango de edad de las personas en el momento en el cual fueron víctimas de mencionado delito. Finalmente, la última opción de visualización ilustra el rango de edad de las personas en el momento en el cual realizaron la denuncia.

País de destino:



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Pasdedestino/CountryofDestine>

Finalmente se realiza una visualización que permite identificar los países a los cuales son enviadas las personas para ser explotadas, teniendo como referencia el país de donde se reclutaron a las personas y el factor que se desea considerar (ambos elegibles por parte del usuario). De esta forma en el presente ejemplo se puede visualizar que las personas que son reclutadas en Filipinas suelen ser transportadas a los países de Arabia Saudita, Malasia y Japón, para ser explotadas en dichos territorios.

2.2. Para qué y para quien puede servir esta visualización

La información presente en esta visualización puede llegar a tener múltiples propósitos, de acuerdo con el tipo de organización que se desee analizar. A continuación, se relacionan algunos tipos de instituciones y su posible interés en las presentes visualizaciones:

Medios de Comunicación:

Las visualizaciones relacionadas en el presente documento pueden ser empleadas por parte de los diferentes medios de comunicación con el fin de generar artículos y noticias que permitan dar a conocer al mundo la condición actual del delito de tráfico de personas. En muchas ocasiones, no se le da mucha relevancia al impacto que puede generar determinado tema en un medio de comunicación, sin embargo, es importante recordar que gracias a artículos como el del periodista Nicholas Kristof del New York Times sobre el estado del agua en algunos países africanos, se logró llamar la atención de la fundación GATES FOUNDATION, permitiendo obtener miles de millones de dólares para apoyar la solución de esta problemática mundial.

Organizaciones internacionales

Organizaciones como National Human Trafficking pueden emplear este tipo de visualizaciones con el fin de extraer información valiosa de todos los datos que ellos recolectan. Precisamente las visualizaciones empleadas en el presente documento fueron realizadas con datos de esta organización, los cuales se encuentran disponibles de manera pública precisamente con el fin de que pueda ser analizada por diferentes entes. A raíz de las visualizaciones generadas, este tipo de

organizaciones pueden adquirir conocimientos sobre como focalizar las campañas de prevención de tráfico de personas, identificando aspectos como en qué países se debe direccionar, a que rangos de edades, y para que tipos de servicios.

Redes Sociales

Las redes sociales han venido trabajando de la mano últimamente con organizaciones como el National Trafficking Resource Center, brindando información sobre personas denunciadas como posibles víctimas del tráfico de personas. Por otro lado, se ha evidenciado que las personas encargadas de reclutar personas para este delito están utilizando este tipo de plataformas para conseguir posibles víctimas de una manera más sencilla. Las redes sociales están trabajando también para lograr identificar y denunciar este tipo de delincuentes.

Basado en lo anterior, las redes sociales podrían emplear este tipo de visualizaciones con el fin de ingresar parámetros adicionales a sus algoritmos de identificación de tal forma que se logren identificar de manera más precisa y rápida tanto las víctimas del delito de tráfico de personas, como los victimarios. Identificando aspectos como por ejemplo la nacionalidad, el rango de edad, el género, entre otros; es posible identificar cuales usuarios son posibles blancos para ser reclutados y de esta forma fortalecer anuncios relacionados con medidas de prevención para caer bajo este tipo de delitos.

Organismos de defensa de los diferentes países

La responsabilidad última de la seguridad de los ciudadanos de un estado radica en los diferentes organismos de defensa. Por este motivo, considero que esas instituciones se pueden favorecer enormemente de esta visualización al brindar un entendimiento mayor sobre como funciona este delito en su país. A través de las visualizaciones pueden identificar también aspectos como los diferentes medios de coerción empleados por los criminales para doblegar a sus víctimas, de tal forma que se identifique un patrón en el modus operandi de las diferentes bandas criminales. Es posible también extraer información de el tipo de explotación que se lleva a cabo en el país, permitiendo con esto identificar que tipos de industrias se deben vigilar con mayor hincapié (construcción, minería, domésticos, entre otros), aportando al proceso de identificación de víctimas y judicialización de victimarios.

2.3. Que preguntas responde la visualización:

La presente visualización responde a las siguientes preguntas:

- ✓ ¿Cuál es el panorama general del delito de tráfico de personas en los Estados Unidos?
- ✓ ¿Cuál es el país en el mundo en el cual se presentan mayor cantidad de reclutamientos?
- ✓ ¿Cuál es el país en el mundo en el cual se presentan mayor cantidad de explotaciones a víctimas del delito de tráfico de personas?
- ✓ ¿A qué tipo de factor son forzados las víctimas de Colombia, cuando están siendo explotados por el delito de tráfico de personas?
- ✓ ¿Qué género es el más afectado por delitos relacionados con prostitución en España?
- ✓ ¿Quiénes son más vulnerables a ser víctimas del tráfico de personas en Filipinas, los jóvenes o la gente adulta?
- ✓ ¿A dónde son transportados las personas reclutadas en Ucrania, con el fin de que sean explotados en dichos países?

3. Descripción técnica del proyecto: lenguajes, librerías, licencias, descripción técnica del

proyecto.

3.1. Que transformación de datos ha habido que hacer respecto al juego de datos inicial

La preparación de los datos se realiza implementando la herramienta Python. El primer paso consistente en verificar los tipos de datos y las valores que cada una de las variables pueda tomar, identificando de igual forma la cantidad de valores que puede asumir y la existencia de valores nulos. Este proceso se logra implementando el siguiente algoritmo, el cual recibe como parámetro de entrada la variable a analizar y el siguiente grupo de opciones:

- **Names_labels:** Este parámetro retorna el nombre de los diferentes valores que puede tomar la variable.
- **Num_labels:** Este parámetro retorna la cantidad de valores diferentes que puede tomar determinada variable.
- **Min_max:** Este parámetro retorna los valores mínimos y máximos de las diferentes variables que se pasan como argumento.

```
def data_structure(list_info, parameter):
    link = r"C:\Users\User\OneDrive\ESTUDIOS\MAESTRIA\MASTER_DATA_SCIENCE\5. Fifth Semester\Visualizacion de Datos\PR1 Seleccion Datos\Datos Explotacion\The_Global_Dataset_Apr_2020.csv'
    data = pd.read_csv(link, index_col=0)
    for info in list_info:
        if(info == 'num_labels'):
            a = list(set(data[parameter]))
            print('Num_Labels: {}'.format(len(a)))
        if(info == 'names_labels'):
            a = list(set(data[parameter]))
            print('Name_Labels: {}'.format(a))
        if(info == 'min_max'):
            print('Min_Value: {}'.format(data[parameter].min()))
            print('Max_Value: {}'.format(data[parameter].max()))

data_structure(['names_labels', 'recruiterRelationUnknown'], # 'num_labels', 'names_labels', 'min_max'
```

Valores Nulos

A raíz de la identificación de las diferentes variables que integran la base de datos se logra evidenciar que los valores nulos en la base de datos están representados por dos valores: el número -99 y el carácter '-99'. Considerando lo anterior, se toma la determinación de reemplazar mencionados valores por valores nulos, de tal forma que sean fácilmente procesados.

```
def null_values_replacement():
    link = r"C:\Users\User\OneDrive\ESTUDIOS\MAESTRIA\MASTER_DATA_SCIENCE\5. Fifth Semester\Visualizacion de Datos\PR1 Seleccion Datos\Datos Explotacion\The_Global_Dataset_Apr_2020.csv'
    data = pd.read_csv(link, index_col=0)
    data = data.replace(-99, np.nan)
    data = data.replace('-99, np.nan)
    data.to_csv(r"C:\Users\User\OneDrive\ESTUDIOS\MAESTRIA\MASTER_DATA_SCIENCE\5. Fifth Semester\Visualizacion de Datos\PR1 Seleccion Datos\Datos Explotacion\The_Global_Dataset_Apr_2020_nan.csv')

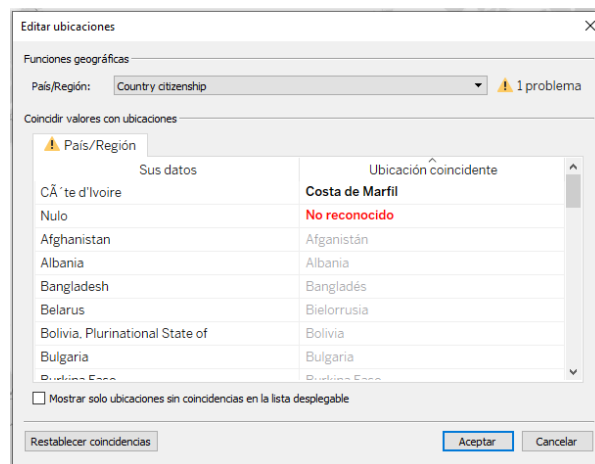
null_values_replacement()
```

Nombre de países:

Una vez se realiza el reemplazo de los valores nulos de todas las columnas se procede a generar los nombres de los países, tanto para identificar la nacionalidad de las víctimas, como para identificar el nombre del país en donde fueron transportadas para ser explotadas. Lo anterior se realiza considerando que las dos variables que hacen relación a países (citizenship, CountryOfExploitation) están determinadas empleando el código el estándar ISO 3166-1 Alpha-2, el cual no es reconocido por el software Tableau. Durante el presente análisis se logró identificar que los códigos de los países presentan valores de '00' los cuales no relacionan ningún país, motivo por el cual dicho valor es reemplazado por valores nulos.


```
def country_code():
    # Se define la ubicación de los datos con el nombre de los países y el código.
    source = r"C:\Users\User\OneDrive\ESTUDIOS\MAESTRIA\MASTER_DATA_SCIENCE\5. Fifth Semester\Visualización de Datos\VPRAI Selección Datos\Datos Explotación\Country_code.txt'
    # Se cargan los datos de tipo csv.
    country_data = pd.read_csv(source, index_col = 0)
    # Se reinicia el index con el fin de poder disponer de
    country_data = country_data.reset_index()
    # Se cambian los nombres con el fin de dejar uno estandarizado junto con el que presenta el otro dataset
    country_data.columns = ['Country_citizenship', 'citizenship']
    # Se carga la base de datos globales
    source_2 = r"C:\Users\User\OneDrive\ESTUDIOS\MAESTRIA\MASTER_DATA_SCIENCE\5. Fifth Semester\Visualización de Datos\VPRAI Selección Datos\Datos Explotación\The_Global_Dataset_Apr_2020.csv'
    # Se cargan los datos de tipo CSV.
    global_dataset = pd.read_csv(source_2, index_col = 0)
    # Se reemplazan los valores '00' por nulos para ambas columnas con información de la ubicación
    global_dataset['citizenship'] = global_dataset['citizenship'].replace('00', np.nan)
    global_dataset['CountryOfExploitation'] = global_dataset['CountryOfExploitation'].replace('00', np.nan)
    # Se unen ambas bases de datos, usando el tipo left
    final_data = pd.merge(global_dataset, country_data, how='left', on='citizenship')
    # Se crea una nueva columna con el nombre de la otra columna que hace referencia a la ubicación
    country_data['CountryOfExploitation'] = country_data['citizenship']
    # Se realiza la modificación del nombre de la columna del código del país, con el fin de integrar la información con la columna que hace referencia al país en donde fue explotado.
    country_data = country_data[['Country_citizenship', 'CountryOfExploitation']]
    country_data.columns = ['Country_Exp', 'CountryOfExploitation']
    # Se realiza la segunda unión para el país de explotación
    final_data = pd.merge(final_data, country_data, how="left", on=['CountryOfExploitation'])
    # Ya que existe un país con código NaN, se cambia mencionado país por null ya que de acuerdo a lo verificado ese código no existe, pero si es interpretado como tal por los valores nulos
    final_data['Country_citizenship'] = final_data['Country_citizenship'].replace('nan', np.nan)
    final_data['Country_Exp'] = final_data['Country_Exp'].replace('nan', np.nan)
    # Se reemplaza el nombre del país considerando que no es reconocido por Tableau.
    final_data['Country_citizenship'] = final_data['Country_citizenship'].replace('Côte d'Ivoire', 'Ivory Coast')
    # Se almacena la información en un archivo de tipo csv
    final_data.to_csv(r"C:\Users\User\OneDrive\ESTUDIOS\MAESTRIA\MASTER_DATA_SCIENCE\5. Fifth Semester\Visualización de Datos\VPRAI Selección Datos\Datos Explotación\The_Global_Dataset_Country.csv")
```

Una vez se tiene la base de datos cargadas en Tableau, en el momento de generar las visualizaciones de tipo **Mapa** para la variable que relaciona el nombre del país donde se produce el reclutamiento de personas, se identifica que no se reconoce el nombre **CA te dIvoire**. Considerando lo anterior, se hace necesario reemplazar el nombre del país por **Costa de Marfil**, de tal forma que no se pierda la información relacionada con esa región.



De igual forma, en el momento de generar las visualizaciones de tipo **Mapa** para la variable que relaciona el nombre del país donde se explotan a las personas, se identifica que no se reconoce el nombre **Taiwan, Province of China**. Considerando lo anterior, se hace necesario reemplazar el nombre del país por **Taiwan**.

Editar ubicaciones

Funciones geográficas

País/Región: Country Exp ⚠️ 1 problema

Coincidir valores con ubicaciones

⚠️ País/Región

Sus datos	Ubicación coincidente
Nulo	No reconocido
Taiwan, Province of China	Taiwán
Afghanistan	Afganistán
Albania	Albania
Argentina	Argentina
Austria	Austria
Bahrain	Bahréin
Bangladesh	Bangladés
Belize	Belize

☐ Mostrar solo ubicaciones sin coincidencias en la lista desplegable

Restablecer coincidencias Aceptar Cancelar

3.2. ¿Qué lenguaje, librería, software has usado y por qué?

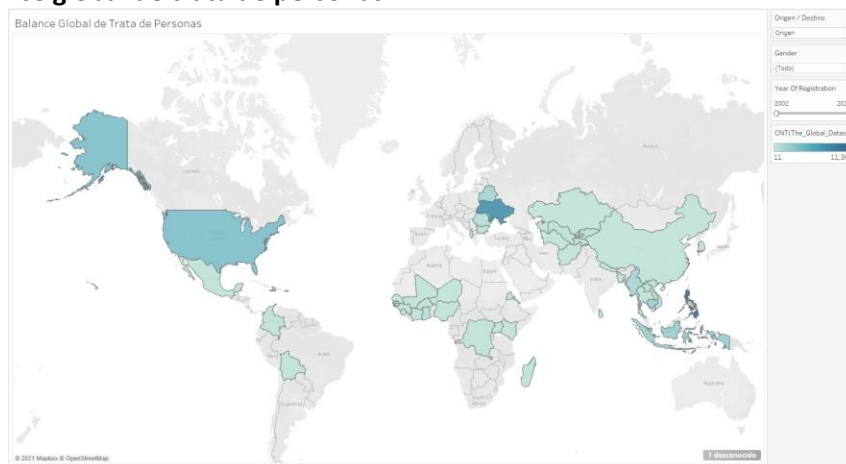
La preparación de los datos se realiza empleando el software **Python**, considerando que contiene una serie de bibliotecas que facilita el proceso de preparación. Del software Python se emplearon las siguientes librerías:

- **Pandas:** Pandas es una librería de tipo BSD-Licensed (Berkely Source Distribution) de fuente abierta que provee métodos de gran rendimiento para el análisis de datos que presentan una estructura.
- **Numpy:** Numpy es una librería de tipo BSD-Licensed (Berkely Source Distribution) de fuente abierta que permite realizar cálculos numéricos en Python. Se emplea principalmente para la creación y modificación de matrices.

4. Visualización de datos.

En el presente punto se ilustran las diferentes visualizaciones generadas empleando el software Tableau.

4.1. Balance global de trata de personas:

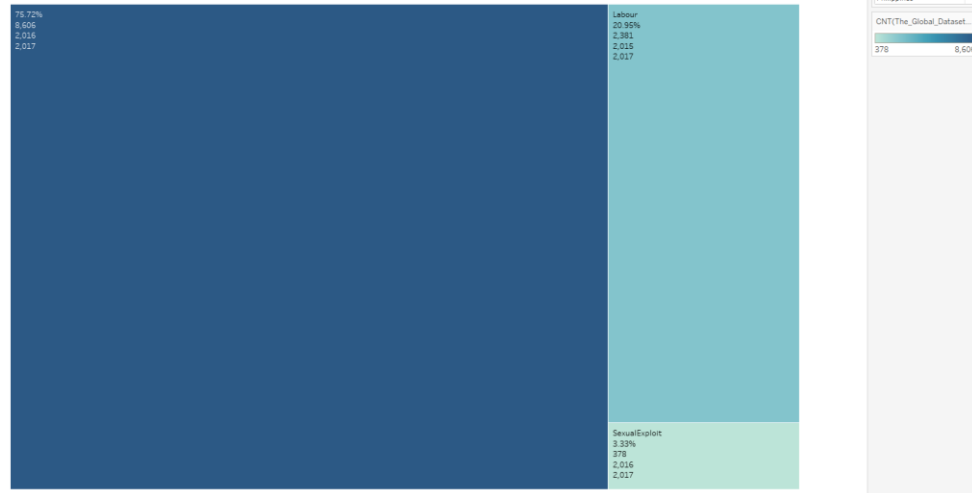


Link:

https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Visualizacin_Practica_Final/MundoOrigDest

4.2. Factor de Explotación (País).

Factor de explotación (País)



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/FactordeexplotacinPas/CountryExploitR>

4.3. Relación con el reclutador

Relación con el reclutador



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Relacinconelreclutador/RecruiterRelation>

4.4. Modo de control

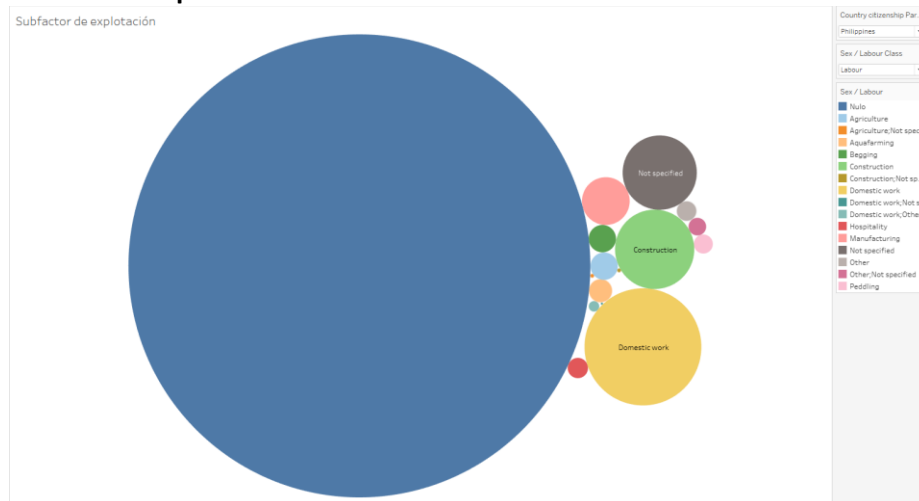
Modo de Control

Mean Ctr	Recuent o de Th..	% de total Re..
Nulo	2,326	97.69%
NoEspeci	14	0.59%
FPromes	12	0.50%
Deuda	11	0.46%
Robo	7	0.29%
APsicolo	5	0.21%
Amenaza	3	0.13%
AFisico	1	0.04%
ASexual	1	0.04%
Documentos	1	0.04%

Link:

[https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/MododeControl/Mo
doControl](https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/MododeControl/Mo
doControl)

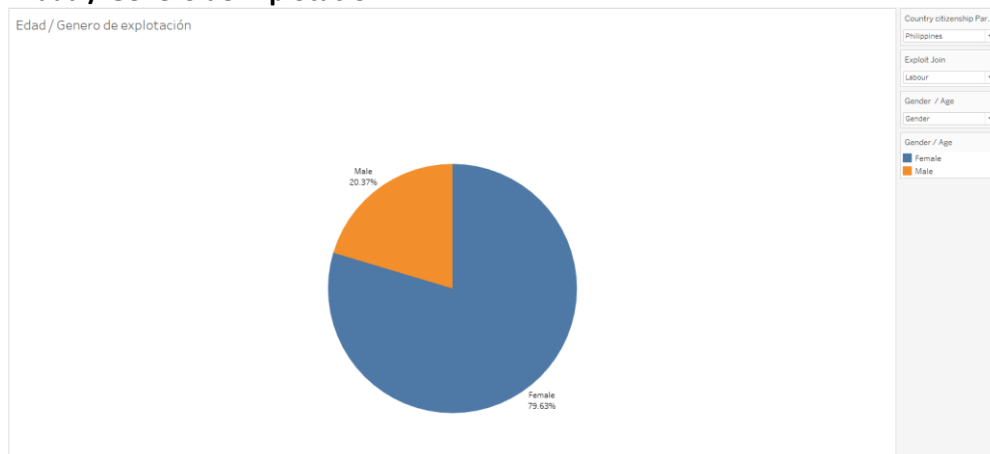
4.5. Subfactor de explotación:



Link:

[https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Subfactordeexplota
cion/MainExploitFact](https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Subfactordeexplota
cion/MainExploitFact)

4.6. Edad / Genero de Explotación



Link:

[https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/EdadGenerodeExplota
cion/ExplotAge](https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/EdadGenerodeExplota
cion/ExplotAge)

4.7. País de destino:



Link:

<https://public.tableau.com/app/profile/andres.felipe.bello.zapata2179/viz/Pasdedestino/CountryofDestine>