



Centro Universitario de la Defensa en la Escuela Naval Militar

TRABAJO FIN DE GRADO

Estudio y evaluación de enfrentamientos armados registrados

Grado en Ingeniería Mecánica

ALUMNO: Gonzalo Díaz Sánchez-Brunete

DIRECTORES: María Álvarez Hernández

CURSO ACADÉMICO: 2023-2024

Universida_{de}Vigo



Centro Universitario de la Defensa en la Escuela Naval Militar

TRABAJO FIN DE GRADO

Estudio y evaluación de enfrentamientos armados registrados

Grado en Ingeniería Mecánica
Intensificación en Tecnología Naval
Infantería de Marina

Universida_deVigo

RESUMEN

Los conflictos armados son un permanente objeto de estudio de diferentes organizaciones, que generan informes de la situación de violencia mundial a partir de los datos recogidos. En este Trabajo de Fin de Grado, se utilizan las bases de datos creadas por dichas organizaciones con el fin de realizar un análisis multivariante, y proporcionar así una visión general del panorama global. Para todo ello, primeramente se realiza un análisis descriptivo del conjunto de datos en su totalidad, apoyándose en una serie de gráficos generados. Como consecuencia de estos, se hará uso de diversos procesos y herramientas estadísticas tales como contrastes de hipótesis o modelos de regresión, con el fin de corroborar algunos de los resultados observados. Posteriormente se plantean varias hipótesis relativas a conflictos actuales en las que se estudian posibles relaciones entre diferentes eventos aparentemente independientes, centrándose en el análisis de variables concretas de cada uno de ellos. En estos casos, se pudo establecer relaciones claras entre los enfrentamientos contrastados.

PALABRAS CLAVE

Registro, base de datos, conflicto armado, estadística, hipótesis

AGRADECIMIENTOS

Quiero aprovechar este apartado para agradecer a toda mi la promoción 424-154 haber sido motivación apoyo constante durante estos cinco años. También quiero recordar a mi familia, que ha sido siempre un pilar fundamental, y no ha sido menos durante estos años. Quiero por último agradecer a mi tutora, la Sra. Álvarez, por su gran implicación en el proyecto y su paciencia desde el primer momento.

CONTENIDO

Contenido	1
Índice de Figuras	3
Índice de Tablas.....	5
1 Introducción y objetivos	6
1.1 Motivación del trabajo	6
1.2 Objetivos	6
1.3 Estructura del trabajo	7
2 Estado del arte	8
2.1 Registros de conflictos bélicos a lo largo de la historia	10
2.2 Registro de conflictos bélicos en la actualidad	12
2.3 Estudios actuales sobre conflictos armados	14
2.4 Herramientas software para el estudio del Dataset	17
3 Desarrollo del TFG.....	19
3.1 Matrices de datos y principales variables	19
3.2 Principales hipótesis de estudio	20
3.3 Procesos estadísticos utilizados	21
3.3.1 Gráficos.....	22
3.3.2 Modelos de regresión.....	26
3.3.3 Tablas de contingencia	27
3.3.4 Comparación de medias/medianas	28
4 Resultados	29
4.1 Preprocesado de la base de datos	29
4.2 Primeros análisis	29
4.2.1 Estudio de variables cuantitativas.....	34
4.3 Hipótesis 1: Existe relación entre el surgimiento de nuevos conflictos en 2021-2022 y la guerra de Ucrania.	41
4.4 Hipótesis 2: A medida que desaparece la participación del ISIS en conflictos en Siria, aumenta en conflictos en África.....	42
4.5 Trabajo realizado en R	45
5 Conclusiones y líneas futuras	48
5.1 Conclusiones	48
5.2 Líneas futuras	49
6 Bibliografía.....	50
Anexo I: Implicaciones Sociales, y/o Económicas, y/o Ambientales	52

Anexo II: Reflexiones Éticas y Sociales 53

Anexo III: Código de R Project.....54

ÍNDICE DE FIGURAS

Figura 2-1: Gráfico explicativo de las etapas de un conflicto (Fuente: Elaboración propia basada en [1]).....	9
Figura 2-2: Libro “Historia de la Guerra del Peloponeso” por Tucídides (Fuente: https://www.casadellibro.com)	10
Figura 2-3: Cartografía militar de la Guerra de los Treinta Años (Fuente: https://www.juntadeandalucia.es)	11
Figura 2-4: Soldados fallecidos reconocidos por tecnología satélite (Fuente: https://as.com)	12
Figura 2-5: Consecuencias de la aparición de nuevos actores en el enfrentamiento Siria-Irak (Fuente: [12]).....	15
Figura 2-6: Distribución geográfica de ISIS, al-Qaeda y afiliados en 2010 y 2019 respectivamente (Fuente: [14])	16
Figura 2-7: Perfil espacial de la violencia contra personal civil por parte de grupos militantes (Fuente: [15])	16
Figura 2-8: Intensidad de conflictos en África (Fuente: [16]).....	17
Figura 2-9: Logotipo R Project (Fuente: [17])	17
Figura 3-1: Diagrama de barras (Fuente: Elaboración propia).....	23
Figura 3-2: Diagrama de sectores (Fuente: Elaboración propia).....	23
Figura 3-3: Histograma (Fuente: [24])	24
Figura 3-4: Gráfico temporal (Fuente: Elaboración propia)	25
Figura 3-5: Diagrama de caja (Fuente: [24]).....	26
Figura 3-6: Diagrama de dispersión (Fuente: [25]).....	26
Figura 4-1: Gráfico espacial de conflictos (Fuente: Elaboración propia)	30
Figura 4-2: Proporción de bajas causadas por tipo de violencia (Fuente: Elaboración propia)	31
Figura 4-3: Conflictos ocurridos por años (Fuente: Elaboración propia).....	33
Figura 4-4: Bajas por años (Fuente: Elaboración propia)	34
Figura 4-5: Muertes de cada tipo ocurridas anualmente (Fuente: Elaboración propia)	35
Figura 4-6: Muertes de cada tipo ocurridas anualmente sin datos atípicos (Fuente: Elaboración propia)	35
Figura 4-7: Gráficos de caja y bigotes de muertes de combatientes y de civiles (Fuente: Elaboración propia)	36
Figura 4-8: Muertes ocurridas anualmente en cada región sin datos atípicos (Fuente: Elaboración propia)	37
Figura 4-9: Conflictos acaecidos por cada tipo de violencia por años (Fuente: Elaboración propia)	38
Figura 4-10: Conflictos acaecidos por cada tipo de violencia por años (Fuente: Elaboración propia)	39
Figura 4-11: Conflictos en África y Europa por tipo de violencia (Fuente: Elaboración propia)....	40

Figura 4-12: Conflictos en Asia y Oriente Medio por tipo de violencia (Fuente: Elaboración propia)	40
Figura 4-13: Gráfico espacial de conflictos surgidos en 2021 y 2022 (Fuente: Elaboración propia)	41
Figura 4-14: Muertes en Siria causadas por el IS (Fuente: Elaboración propia)	43
Figura 4-15: Muertes en Siria y África causadas por el IS (Fuente: Elaboración propia)	43
Figura 4-16: Gráfico de dispersión de muertes en Siria y África por el IS incluyendo violencia tipo 1 (Fuente: Elaboración propia)	44
Figura 4-17: Gráfico de dispersión de muertes en Siria y África por el IS incluyendo violencia tipo 1 con línea de tendencia exponencial (2017-2022) (Fuente: Elaboración propia)	45
Figura 4-18: Trabajo realizado en R (parte 1)	46
Figura 4-19: Trabajo realizado en R (parte 2)	47

ÍNDICE DE TABLAS

Tabla 3-1: Explicación de las variables de interés de la base de datos de UCDP (Fuente: Elaboración propia basada en [5]).....	20
Tabla 3-2: Tabla de contingencia (Fuente: Elaboración propia)	27
Tabla 4-1: Número de conflictos registrados por regiones (Fuente: Elaboración propia)	30
Tabla 4-2: Número de conflictos registrados por tipos de violencia (Fuente: Elaboración propia).....	31
Tabla 4-3: Medidas descriptivas de los tipos de muertes (Fuente: Elaboración propia).....	32
Tabla 4-4: Número de conflictos por año (Fuente: Elaboración propia).....	32
Tabla 4-5: Número de bajas causadas por año (Fuente: Elaboración propia).....	33
Tabla 4-6: Número de conflictos por tipo de violencia y año (Fuente: Elaboración propia).....	38
Tabla 4-7: Proporción de conflictos de cada tipo por regiones (Fuente: Elaboración propia).....	39
Tabla 4-8: Conflictos surgidos en Siria en los años 2021 y 2022 (Fuente: Elaboración propia)	42
Tabla 4-9: Número de muertes por país y año en África de 2012 a 2022 (Fuente: Elaboración propia)	44

1 INTRODUCCIÓN Y OBJETIVOS

1.1 Motivación del trabajo

Los conflictos armados constituyen una de las más destructivas realidades de nuestra sociedad actual, cuya influencia abarca prácticamente todos los ámbitos de la sociedad, desde la política y la economía hasta su dimensión social y humana. En un mundo donde la paz es un ideal que cada vez se antoja más inalcanzable, ser capaces de comprender la naturaleza de estos enfrentamientos, así como la dinámica con la que se desarrollan, se ha convertido en una extrema necesidad para la prevención de crisis internacionales, y llegado el caso, para su resolución.

Es por ello, que un correcto y detallado estudio de los datos ayudado de procedimientos estadísticos, así como un análisis de los resultados obtenidos, es esencial para poder determinar tendencias y patrones seguidos por estos enfrentamientos. Todo ello deriva en la posibilidad de desarrollar estrategias para la prevención de conflictos basadas en las evidencias y conclusiones obtenidas de acontecimientos previos.

A partir de los datos recogidos se generan informes sobre la situación de violencia mundial, ofreciendo descripciones detalladas de los distintos tipos de conflictos y territorios involucrados. El aporte de conclusiones a fuentes públicas y medios de divulgación suele ser determinado a partir de un análisis que, normalmente, se limita a una observación superficial de la situación, sin profundizar en investigaciones que puedan establecer relaciones, similitudes y/o un patrón temporal (en caso de haberlo). Por tanto, en este trabajo de fin de grado se pretende indagar en estos datos con una mayor profundidad, de forma que se puedan determinar (o no) las conclusiones anteriores de forma exploratoria. Para ello, se llevará a cabo un análisis multivariante de la/s base/s de datos, considerando las diversas características estudiadas en cada conflicto. A este efecto se hará uso del software R, un programa especializado en el análisis estadístico tanto descriptivo como inferencial.

1.2 Objetivos

Este trabajo se centra en realizar un estudio estadístico de los principales aspectos de los conflictos armados registrados. En él se pretende hacer uso, en un primer momento, de la estadística descriptiva, mediante la cual poder obtener una visión general de los datos, para así establecer determinadas hipótesis con el fin de corroborarlas (o no) a través de procedimientos de inferencia estadística, ahondando un poco más en estos datos. La relación entre las características de los enfrentamientos permitirá obtener relaciones que ayuden a comprender las secuencias patrones de estos sucesos.

Es por ello, que otro de los objetivos de este trabajo es el correcto análisis estadístico de los datos registrados y el estudio de las variables que caracterizan (de manera genérica) los enfrentamientos

violentos. También se establece como meta entender e interpretar los resultados obtenidos, de forma que se puedan extraer conclusiones claras, que faciliten unas futuras líneas de investigación y dejen unas bases asentadas sobre las relaciones elementales entre los conflictos armados registrados. Se pretende, además, utilizar para todo ello el software R, específico para realizar análisis estadísticos, el cual requiere un aprendizaje y periodo de familiarización previos

1.3 Estructura del trabajo

La memoria se estructura comenzando con una introducción, en la cual se pretende abarcar la idea general del trabajo, proporcionando una explicación de la situación actual en lo referente a la recopilación y tratamiento de datos relativos a los conflictos armados, así como de las intenciones y objetivos del trabajo.

El trabajo consta de los siguientes capítulos:

- Estado del arte: Pretende proporcionar un contexto actualizado sobre el tema de estudio, introduciendo el registro de los conflictos bélicos a lo largo de la historia y en la actualidad. Se integrará así el tema en el panorama actual, y se indicarán trabajos y estudios realizados hasta el momento sobre los datos que se usarán también en el desarrollo de este trabajo.
Se hará una introducción a la base estadística sobre la que estará elaborado el proyecto, la cual nos permitirá sacar conclusiones sobre las hipótesis que se plantearán; así como al software que lleva implementados dichos procedimientos estadísticos.
Este capítulo, a su vez se divide en tres apartados. El primero de ellos en el que se hace un repaso histórico de los diferentes métodos de recopilación de datos relativos a conflictos bélicos en las diferentes épocas. El segundo se centra en exponer las organizaciones que han creado bases de datos que han sido empleadas para otros proyectos similares a este. En el tercero y último, se han seleccionado y sintetizado una serie de artículos acerca de la materia en cuestión, con el objetivo de obtener una línea general a seguir, en base a proyectos llevados a cabo en la actualidad.
- Desarrollo del trabajo: Se comentarán métodos de estadística tanto descriptiva como inferencial con la finalidad de establecer las relaciones, en caso de haberlas, entre la aparición de diferentes conflictos, así como de su distribución espacial y su desarrollo en el tiempo. Además, se plantearán hipótesis de estudio, que se abordarán con detalle en el siguiente capítulo.
- Resultados: Se presenta la secuencia de trabajo seguida, a la vez que se exponen los resultados y productos gráficos obtenidos en el proceso de estudio, tanto del conjunto de datos con carácter descriptivo, como de las dos hipótesis planteadas. También se analizan los resultados obtenidos, permitiendo extraer información que confirme o desmienta cada una de las hipótesis. Para ello se hará uso del lenguaje de programación R, herramienta que nos permitirá llevar a cabo esos estudios estadísticos, así como la realización de diferentes tablas y gráficos que faciliten la visualización y el entendimiento de los resultados.
- Conclusiones: Se expondrán los conceptos extraídos de la realización del proyecto, así como los objetivos alcanzados. Además, se proponen futuras líneas de investigación para ahondar en el estudio de los conflictos con mayor profundidad y con un carácter ya menos descriptivo.

2 ESTADO DEL ARTE

El registro de conflictos bélicos ofrece una visión en detalle de los enfrentamientos armados de diversas índoles en el transcurso de la historia. No solo nos proporciona una visión detallada de los diferentes conflictos a lo largo de los años, sino que también nos permite estudiar sus causas y consecuencias.

Estos registros no se limitan a recoger las fechas y los lugares de los conflictos, sino que además las bases de datos pueden ofrecer información sobre campos como las fuerzas involucradas¹, las armas y tácticas empleadas, el transcurso del enfrentamiento, o las bajas y daños causados, tanto muertes, como personas que resulten heridas como consecuencia de la batalla; además de los daños sufridos por el medio ambiente y las infraestructuras de la zona involucrada.

Además, facilitan la comprensión de los factores que terminan por derivar en eventos bélicos. En este aspecto podemos analizar las causas inmediatas de los conflictos, disputas o eventos concretos que desencadenan un determinado enfrentamiento, y los factores subyacentes que, aunque no de una manera tan directa y evidente como los anteriores, contribuyen al surgimiento del conflicto. También podemos examinar las consecuencias, tanto a corto plazo -la resolución de la guerra, o cambios políticos y/o territoriales producidos a raíz del evento-, como a largo plazo -la forma en la que las sociedades y las culturas involucradas se ven afectadas como consecuencia del enfrentamiento-.

Dicho lo cual, queda patente la importancia del registro de episodios bélicos como herramienta de comprensión histórica, así como una ayuda para la obtención de una visión y entendimiento profundos de los enfrentamientos armados, de sus causas y consecuencias.

Se ha mencionado en numerosas ocasiones el término “conflicto armado”, pero ¿qué significa realmente? Por un lado, el Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR) [1], establece que:

“Se entiende por conflicto armado todo enfrentamiento protagonizado por grupos armados regulares o irregulares con objetivos percibidos como incompatibles en el que el uso continuado y organizado de la violencia:

a) provoca un mínimo de 100 víctimas mortales en un año y/o un grave impacto en el territorio (destrucción de infraestructuras o de la naturaleza) y la seguridad humana (ej. población herida o desplazada, violencia sexual, inseguridad alimentaria, impacto en la salud mental y en el tejido social o disrupción de los servicios básicos);

¹ Se ofrece información como las partes involucradas, la entidad y la composición de estas

b) pretende la consecución de objetivos diferenciados de los de la delincuencia común y normalmente vinculados a: i) demandas de autodeterminación y autogobierno, o aspiraciones identitarias; ii) la oposición al sistema político, económico, social o ideológico de un Estado o a la política interna o internacional de un Gobierno, lo que en ambos casos motiva la lucha para acceder o erosionar al poder; iii) el control de los recursos o del territorio”

En el gráfico mostrado en la *Figura 2-1*, se diferencian las etapas de los conflictos basados en la información proporcionada por ACNUR [1]:

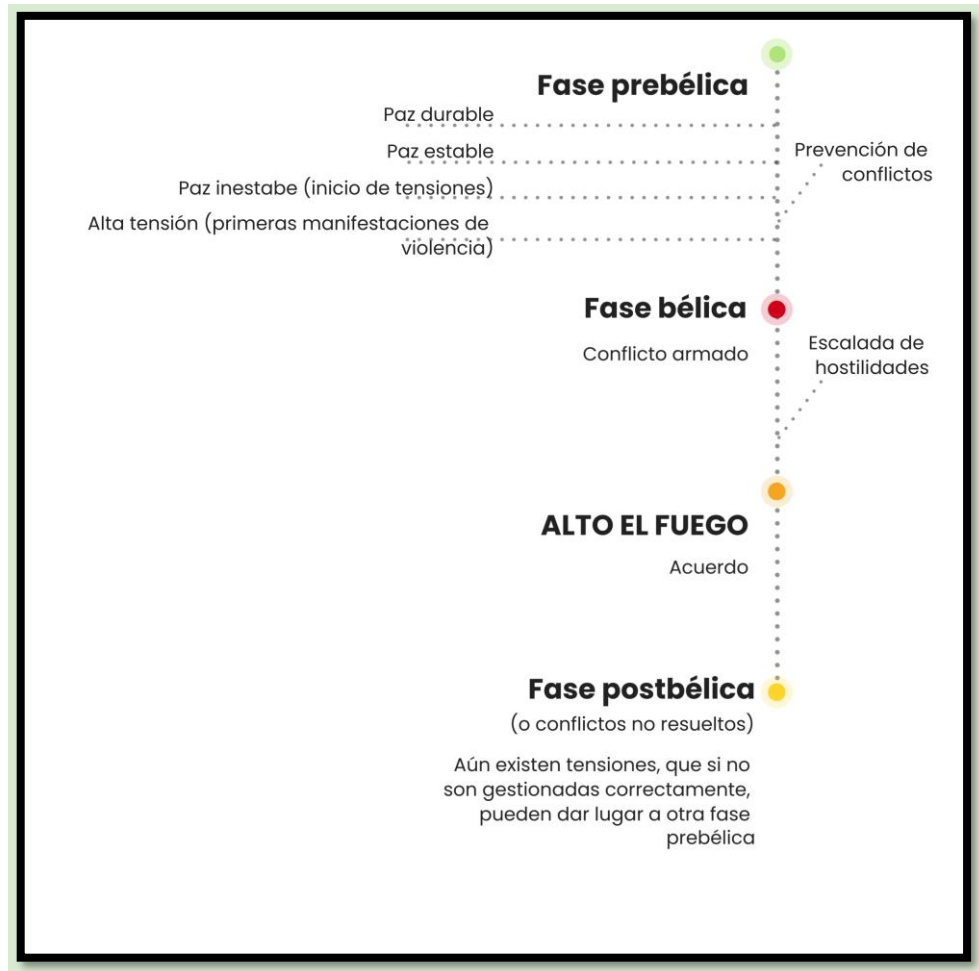


Figura 2-1: Gráfico explicativo de las etapas de un conflicto (Fuente: Elaboración propia basada en [1])

Por otro lado, el Derecho Internacional Humanitario (DIH) -rama del derecho internacional que rige los conflictos armados- no proporciona una única definición, sino que hace una distinción entre dos tipos de conflictos armados: conflictos armados internacionales, donde se enfrentan dos o más Estados, y conflictos armados no internacionales, entre fuerzas gubernamentales y grupos no gubernamentales, o entre varios grupos no gubernamentales.

De acuerdo con el artículo 2 común a los Convenios de Ginebra de 1949 [2], un conflicto armado internacional (CAI) será aquel en el que se enfrenten “Altas Partes contratantes”, entendidas estas como Estados, sin importar las razones o intensidad de dicho enfrentamiento. De acuerdo con el artículo 3 común a los Convenios de Ginebra de 1949 [2], un conflicto armado no internacional (CANI), se entiende como “conflicto armado que no sea de índole internacional y que surja en el territorio de una de las Altas Partes Contratantes” todo enfrentamiento armado en el que participen uno o más grupos armados no gubernamentales, ya sea la hostilidad solo entre estos grupos, o entre un grupo armado no gubernamental y fuerzas armadas gubernamentales.

Por su parte, las Fuerzas Armadas, en el Derecho Internacional Humanitario aplicado a las Fuerzas Armadas (FAS) [3] proporcionan la siguiente definición:

“... la amplia noción de conflicto debe completarse con la de enfrentamiento armado, que quiere decir tanto como lucha entre dos partes contendientes con utilización de las armas, sin que importe la existencia de una resistencia de gran entidad por una de ellas, ni naturalmente el requisito formal de una declaración de guerra. Existe conflicto armado cuando se inician los actos de fuerza o violencia armada con manifiesta voluntad hostil.”

2.1 Registros de conflictos bélicos a lo largo de la historia

El transcurso de la historia se ha visto alterado en numerosas ocasiones por diferentes conflictos bélicos, siendo estos uno de los principales timones de nuestro pasado. Han sido el principal motivo de los grandes cambios, así como instrumento de aprendizaje sobre errores cometidos, no solo a nivel bélico, sino a nivel humano. La recopilación de datos acerca de los conflictos armados acontecidos a lo largo de la historia es un asunto de gran importancia. Es esencial valorar la importancia de la memoria histórica, conocer lo sucedido en nuestro pasado, y ser capaces de identificar las causas de los diferentes enfrentamientos, así como las consecuencias derivadas de los mismos, para ser capaces de prevenir eventos similares en tiempos futuros.

Esta recopilación de datos ha ido evolucionando a lo largo de la historia. Aunque bien es cierto que el estudio de los métodos desarrollados en la Antigüedad presenta una dificultad añadida por la naturaleza limitada de las fuentes disponibles, se tiene constancia de que se llevaron a cabo registros escritos. Se escribían crónicas y anales -relatos de un carácter más oficial, donde con frecuencia se incluía información sobre guerras y batallas-, también en los textos religiosos se podían encontrar referencias a ciertos eventos bélicos. La literatura también aportó, aunque con un carácter menos oficial y con una fidelidad variable, descripciones de guerras y batallas. Buenos ejemplos de este registro de conflictos en la Antigüedad pueden ser los historiadores griegos Tucídides y Polibio. Tucídides, por su parte, documentó con buen grado de detalle la guerra entre Atenas y Esparta -Guerra del Peloponeso *Figura 2-2*-, mientras que Polibio, realizó escritos acerca de las guerras entre Roma y Cartago.

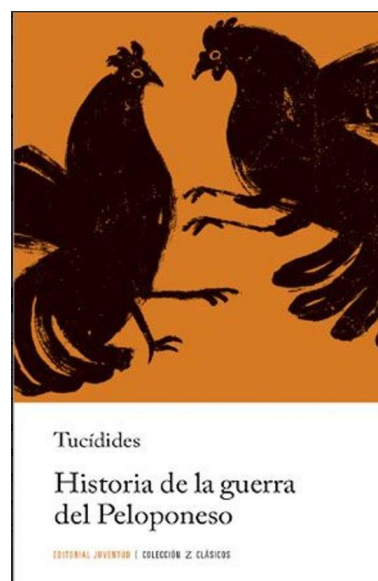


Figura 2-2: Libro “Historia de la Guerra del Peloponeso” por Tucídides (Fuente: <https://www.casadellibro.com>)

La documentación de esta época presenta un claro hándicap, que es la falta de datos, ya que la información que hay sobre las guerras en la Antigüedad suele ser incompleta y fragmentada. Además,

las fuentes suelen ser sesgadas, es decir, no acostumbran a ser fuentes con perspectiva objetiva, sino que suelen presentar una tendencia favorable a uno de los bandos. Y, aunque quizá no sea un aspecto crítico, la diferencia cultural y lingüística también puede llegar a dificultar el análisis y la comprensión de estos registros.

Si se sigue avanzando en la historia, se entra en la Edad Media, donde esta recopilación de datos sobre los eventos bélicos siguió evolucionando poco a poco. Aunque la mayoría de los formatos coincidían con los de la Antigüedad -crónicas, anales, textos religiosos, y literatura, que documentan episodios tan importantes como la Guerra de los Cien Años, las Cruzadas o la Reconquista-, fueron a la vez apareciendo nuevos registros en forma de cartas y documentos diplomáticos, que, aunque no contuvieran descripciones de las batallas como el caso de las crónicas, proporcionaban otro tipo de información que contribuía al entendimiento de las relaciones entre reinos, de las negociaciones políticas que se llevaban a cabo y otro tipo de aspectos que favorecen la comprensión de las batallas así como de sus causas y eventos detonantes. Entre estos nuevos tipos de registros se encuentran cartas reales, tratados de paz, acuerdos comerciales, ... Se puede destacar algún escrito fundamental de esta época como por ejemplo el Tratado de Verdún -año 843-, mediante el cual se dividió el imperio carolingio entre los tres hijos de Carlomagno, o la Magna Carta -año 1215-, que limitaba el poder del rey de Inglaterra.

Todos estos documentos nos sumergen en los asuntos políticos, religiosos y económicos de la época, dejan entrever las relaciones entre reinos, alianzas, relaciones religiosas, situación del comercio y las finanzas del momento, lo que contribuye a un entendimiento general de la cultura de la época, y facilita la comprensión de los demás registros permitiéndonos leerlos con conciencia histórica.

Continuando a lo largo de la línea temporal, se llega al siglo XV y con él a la Edad Moderna. Una época en la que avanzaron los métodos de registro. Tuvo lugar el auge de la cartografía en el mundo militar *Figura 2-3* -en el siglo XVI se creó en Amberes, Bélgica, la primera escuela de cartografía militar- constituyendo un nuevo método de recopilación de datos, ya que sobre estos planos, en numerosas ocasiones se plasmaba una gran cantidad de información relativa al conflicto, lo que unido a los demás métodos de recopilación de datos desarrollados durante las épocas anteriores, iba ya conformando un buen sistema de registro.



Figura 2-3: Cartografía militar de la Guerra de los Treinta Años (Fuente: <https://www.juntadeandalucia.es>)

Se llega así a la *Época Contemporánea*, la cual comprende desde el siglo XIX hasta la actualidad. En esta época sucedieron un elevado número de cambios tanto en la forma de hacer la guerra como en la forma en la que se registraban y analizaban los datos de los conflictos armados, motivados, en gran medida, por los grandes avances tecnológicos. Centrándose en el registro de datos, en cuanto a su obtención, las nuevas tecnologías han permitido una mayor precisión, así como una mayor velocidad de recopilación. Los drones y satélites, muy en boga en la actualidad, a través de imágenes aéreas, como la *Figura 2-4*, y de una visión cenital del campo de batalla, permiten realizar un rápido y preciso conteo de bajas, de heridos, dimensionar los daños causados en las infraestructuras de la zona, e incluso seguir el enfrentamiento en tiempo real. Además, los avances en los sistemas de comunicación han permitido una mejora sustancial en el flujo de información en el campo de batalla, lo que conlleva una mayor precisión en los informes de bajas y una mayor brevedad en su transmisión a los puestos de mando, y posteriormente, a las bases de datos de las que se nutren las diferentes organizaciones para realizar los análisis y estudios de los datos.

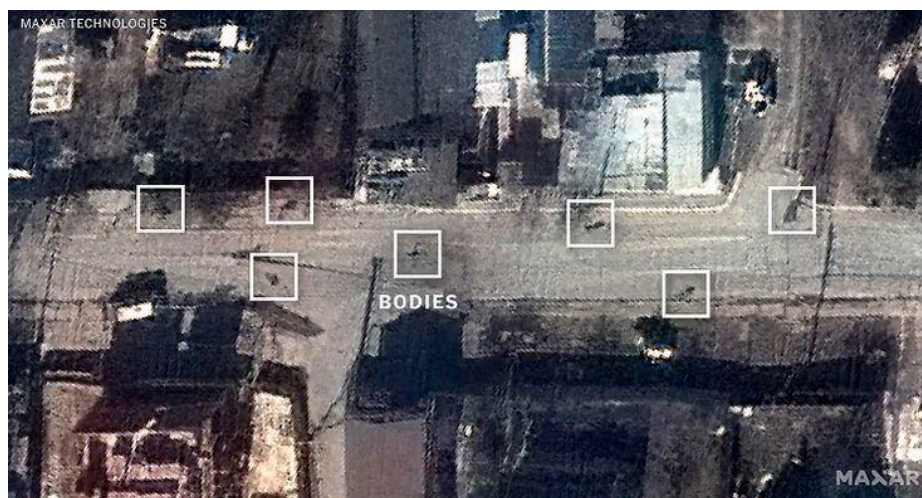


Figura 2-4: Soldados fallecidos reconocidos por tecnología satélite (Fuente: <https://as.com>)

Además, estas nuevas tecnologías han posibilitado la creación de bases de datos digitales, con la capacidad de albergar inmensas cantidades de información, así como el desarrollo de diferentes softwares que facilitan el tratamiento de estos datos.

2.2 Registro de conflictos bélicos en la actualidad

Dada la importancia del tema en cuestión, han sido muchas las asociaciones y organizaciones que se han adentrado en el registro y el estudio de los conflictos bélicos, especialmente con la finalidad de facilitar el análisis de sus causas y consecuencias, para así, mediante estudios estadísticos, poder extraer patrones y tendencias que permitan realizar previsiones y poder actuar en consecuencia.

- En 1959 se fundó el *Instituto Internacional de Investigación de la Paz de Oslo (PRIO)*, Noruega, una institución de investigación independiente especializada en el estudio de temas relacionados con la paz, la seguridad y los enfrentamientos armados. Este instituto cuenta con su base de datos propia y, además, contribuye a la creación de conjuntos de datos sobre la paz y enfrentamientos armados en todo el mundo. Aborda así cuestiones de gran importancia como la seguridad internacional, resolución de conflictos, y construcción de la paz entre otros.

La base de datos de *PRIO*, contiene información acerca de las partes involucradas, la localización de las guerras, la fecha en la que tuvieron lugar, y la duración entre otras. *PRIO* también colabora con otras instituciones y proyectos [4], tales como la Lista de finalistas del Premio Nobel de la Paz, la ONU, o el Banco Mundial.

- En 1963, J. David Singer, un reconocido investigador en Ciencias Políticas de la Universidad de Michigan, Estados Unidos, comenzó el proyecto “*Correlates of War Project*”. Proyecto en el que se han involucrado un gran número de académicos en el ámbito de las relaciones y conflictos internacionales. Además, ha contado con la colaboración de diversas instituciones académicas, cuyo compromiso ha hecho posible la continuidad del proyecto.
El proyecto se centró en recopilar y analizar datos de los conflictos internacionales desde el siglo XIX hasta la actualidad. En él se abordan diferentes aspectos específicos tales como la duración de los enfrentamientos, las alianzas entre estados, la magnitud de la violencia y demás factores propios de los eventos bélicos.
- En 1979, en la Universidad de Uppsala, en Suecia, se fundó el *Uppsala Conflict Data Program (UCDP)*. Se trata de un proyecto centrado en la recopilación y análisis de datos acerca de conflictos armados a nivel global. El *UCDP* cuenta con una extensa base de datos en constante actualización. Esta base de datos incluye una amplia variedad de variables con la finalidad de abordar los diferentes aspectos de los enfrentamientos, tales como la duración, las partes involucradas, las bajas tanto civiles como militares, entre otros.
El *UCDP* no consiste únicamente de la base de datos, sino que también ofrece herramientas de análisis y visualización que permiten examinar tendencias a lo largo del tiempo, explorar patrones, ... Además, publica de manera anual informes que analizan las tendencias tanto globales como regionales de los conflictos armados. Realiza también trabajos de forma conjunta con *PRIO*, organización con la que comparten relaciones estrechas por su cercanía. De hecho, en [5] se describe y explica el conjunto de datos usado por ambas instituciones en su proyecto conjunto.
Este programa utiliza en la creación de la base de datos, fuentes de información de fuente abierta de diferentes países. Siendo las más usadas Reuters News, de Reino Unido, BBC y Associated Press, de EE. UU., Agence France Presse, de Francia, o Xinhua, de China.
- En 2001, se fundó el *Global Terrorism Database (GTD)* [6] como una iniciativa liderada por el *Consorcio Nacional para el Estudio del Terrorismo y Respuestas al Terrorismo de la Universidad de Maryland*, Estados Unidos. Se trata de una de las bases de datos más completas en cuanto a terrorismo se refiere, y es muy utilizada para el estudio de las actividades terroristas en numerosos centros universitarios a nivel global.
Esta base de datos incluye información detallada de cada acto terrorista registrado, como la localización de los hechos, la fecha, los perpetradores del atentado/acto terrorista, y las víctimas de este, así como el objetivo del atentado. Además, incluye una serie de variables más complejas como los métodos empleados en el atentado, las motivaciones de los terroristas, así como la respuesta gubernamental.
El [6] *GTD* proporciona una cobertura global, es decir, incluye incidentes ocurridos en todo el mundo, aportando una visión general acerca del estado del terrorismo a gran escala. Incluye no solo actos de grupos terroristas, sino que también atentados perpetrados por actores individuales.
- En 2014, en Londres, Reino Unido, se fundó el *Armed Conflict Location & Event Data Project (ACLED)*, otra iniciativa centrada en la recopilación, análisis y publicación de datos sobre enfrentamientos bélicos en todo el mundo.
Cuenta también con una amplia base de datos que registra eventos relacionados con enfrentamientos bélicos, incluyendo ataques, protestas, enfrentamientos armados. Todas las entradas de la base de datos contienen información relevante relativa al suceso en cuestión, como la fecha, ubicación, y partes involucradas entre otros.
Con el objetivo de proporcionar una imagen más completa de la dinámica de la violencia política, *ACLED* presta también atención a los eventos de baja intensidad, que nutren el entendimiento del ambiente en el que se desarrollan el resto de los conflictos.

Esta iniciativa tiene un enfoque global ya que se asocia con organizaciones locales para recopilar información de difícil acceso, por lo que contiene datos de una enorme variedad de regiones del mundo. Con la finalidad de mejorar la calidad de los datos, [7] *ACLED* colabora con investigadores académicos, organizaciones no gubernamentales entre otras entidades.

2.3 Estudios actuales sobre conflictos armados

Son múltiples los artículos elaborados -de los cuales se habla posteriormente en el apartado- que se han servido de las bases de datos mencionadas en el apartado anterior, y es precisamente la riqueza de esa información la que permite realizar análisis centrados en diferentes aspectos de los enfrentamientos, obteniendo como resultado una amplia variedad de estudios y conclusiones. Cabe mencionar que las organizaciones comentadas estudian y analizan este tipo de eventos observándolos desde diferentes prismas. Hay, por tanto, artículos que, pese a poseer la misma temática, están-enfocados hacia aspectos concretos que se desvían de nuestro foco de estudio.

Algunos de estos estudios se focalizan en determinados aspectos de la sociedad, como el [8], desarrollado en la Universidad de Göttingen, Alemania, enfocado al estudio de los efectos de los conflictos en la educación, o el [9], desarrollado en la Universidad de Dublín, enfocado en la relación entre los acontecimientos bélicos y las crisis alimenticias. Otros, se centran en un análisis de carácter comparativo, como el [10], que se centra en la comparativa de potencias militares en función al material en posesión de diferentes ejércitos, o el [11], enfocado hacia la predicción de cómo posibles apoyos a un determinado grupo podría propiciar un levantamiento, y en función de dichos apoyos, estimar los niveles de violencia del levantamiento.

En nuestro proyecto tomaremos referencias de aquellos trabajos centrados en el estudio de las relaciones existentes entre conflictos, o de variables que intervengan directa y decisivamente en los desenlaces de los enfrentamientos. En definitiva, aquellos trabajos más enfocados a los conflictos en sí, en lugar de estudiar los efectos de los conflictos en otros ámbitos de la sociedad.

En 2024 se realizó el proyecto [12] para el *Global Security Programme* de la universidad de Pembroke de Reino Unido. En él se explican los cambios espaciales durante los conflictos. Se estudia cómo la aparición de un nuevo actor dominante² en un conflicto ya existente suele conllevar un desplazamiento geográfico del mismo. Esto estará generalmente propiciado por el intento de ese nuevo integrante de combatir en un territorio bajo su control, llevando a cabo la conocida como “atracción de bajo riesgo/ alta probabilidad”. Este mecanismo empleado con frecuencia se basa en trasladar el conflicto a un territorio en el que tenga una baja probabilidad de ser derrotado por otro actor del conflicto, y a la vez tener una alta probabilidad de potenciar sus capacidades y crecer como fuerza beligerante.

De este artículo se extrae la importancia de mantener registro de la localización de los conflictos, así como de sus participantes con el fin de poder predecir posibles desplazamientos del evento. Tal y como se muestra en la *Figura 2-5*, este trabajo presentó, entre otras, las consecuencias de la aparición de nuevos actores dominantes en el conflicto entre Siria e Irak entre 2010 y 2013, exponiendo el cambio espacial del evento que tuvo lugar.

² Actor con el primer o segundo número más alto de encuentros violentos con otras figuras combatientes

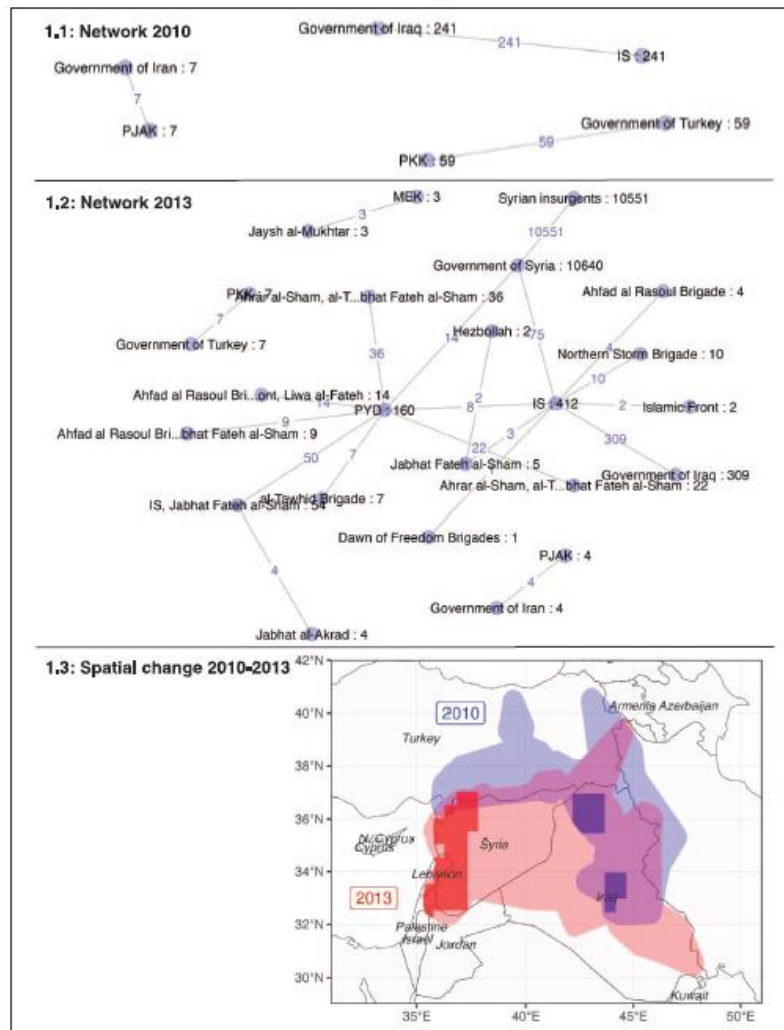


Figura 2-5: Consecuencias de la aparición de nuevos actores en el enfrentamiento Siria-Irak (Fuente: [12])

En 2021, personal del *Department of Peace and Conflict Research* de la Universidad de Uppsala, Suecia, realizó el proyecto [13], orientado analizar los patrones de los enfrentamientos armados en terreno urbano y rural. Para ello se sirvió del conjunto de datos "Ciudades y Eventos de Conflicto Armado", que amplía el Conjunto de Datos de Eventos Georreferenciados del UCDP.

En este trabajo se trata de establecer una relación entre los enfrentamientos armados y la población de las localidades en las que han tenido lugar. De esta manera, se pretende estudiar la distribución de eventos bélicos, tendencias temporales, categorías de violencia, y contraste de conflictos urbanos con los acaecidos en ambiente rural.

También personal del *Department of Peace and Conflict Research* de la Universidad de Uppsala en el año 2020, llevó a cabo el estudio [14]. Este estudio hace referencia al grupo terrorista ISIS, ya que desde que se derrotó a dicho conjunto, el número de muertes en Siria disminuyó hasta llegar al mínimo desde el estallido de la guerra. No obstante, la parte interesante del estudio viene después, pues los datos revelan que a medida que el ISIS fue siendo derrotado en Siria, se fue trasladando hacia África, donde relocalizó sus esfuerzos.



Figura 2-6: Distribución geográfica de ISIS, al-Qaeda y afiliados en 2010 y 2019 respectivamente (Fuente: [14])

La *Figura 2-6* muestra el traslado de la actividad yihadista en el periodo de 2010-2019. A pesar de las grandes represiones y su final derrota en Siria durante el 2019, ISIS, Al-Qaeda y sus afiliados llevaron a cabo ataques en 25 países diferentes, desafiando a 20 gobiernos.

El análisis de este artículo nos conduce a una posible hipótesis a desarrollar y comprobar en el desarrollo del proyecto: los grupos terroristas de una entidad notable, a medida que son derrotados en una zona, trasladan su influencia para apoderarse del control de otra zona, dando lugar a una migración de los conflictos bélicos.

Por otra parte, en Universidad de Essex, en Colchester, Inglaterra, se llevó a cabo el estudio [15] sobre los ataques de grupos militantes contra personal civil. Se llegó a la conclusión de que había dos factores principales que determinaban el nivel de violencia por grupos armados hacia población civil, uno de ellos era la presencia de competencia en la zona, y el otro el control, o no, de la zona por uno de los grupos. Pues, tal y como se puede ver en la *Figura 2-7*, en caso de que uno de los grupos tuviera el control territorial, los casos de violencia contra civiles aumentaban, especialmente en presencia de competencia de otro grupo en la zona, ya que no tenían que ganarse la opinión de nadie, simplemente demostrar su poder y mantenerse en control de la zona siendo más potente que sus adversarios o competidores. Sin embargo, si no tenían el control de dicho territorio, estas acciones eran más comedidas ya que necesitaban lograr un balance entre la demostración de poder y de autoridad, y el ganarse el favor de la población local para poder llegar a hacerse con el control territorial y legitimarse en el poder.

		Competition v. No competition		Likelihood of civilian targeting
		Competition	No competition	
Territorial control v. No territorial control	Control	<i>JNIM</i> <i>Est Region, Burkina Faso</i>	<i>JNIM</i> <i>Baraboulé, Burkina Faso</i>	Higher High Low
	No control	<i>Boko Haram</i> <i>Borno, Nigeria</i>	<i>al-Shabaab</i> <i>Nairobi, Kenya</i>	

Figura 2-7: Perfil espacial de la violencia contra personal civil por parte de grupos militantes (Fuente: [15])

Para la determinación de si un grupo se hallaba en control del territorio en estudio, el investigador de [15], se basó en los datos proporcionados por ACLED acerca de sucesos que comprendían batallas, establecimiento de cuarteles, desarrollo estratégico y transferencia no violenta del territorio.

También basado en la base global de datos ACLED, se llevó a cabo el proyecto [16] en el que trabajó en la identificación de zonas conflictivas y la relación entre diversos atributos y la intensidad de los conflictos. Ayudado de la tecnología GIS y usando un modelo estadístico, halló una fuerte correlación espacial y temporal en el estallido de múltiples conflictos en África, siendo las regiones del norte y el este las de mayor concentración de eventos violentos de alta intensidad, especialmente durante el año 2021 tal y como se aprecia en la *Figura 2-8*. Los eventos principales identificados incluyen batallas y violencia contra civiles, y se identificaron las zonas de Somalia, Sudán del Sur, Nigeria y Mali como las más conflictivas. [16]

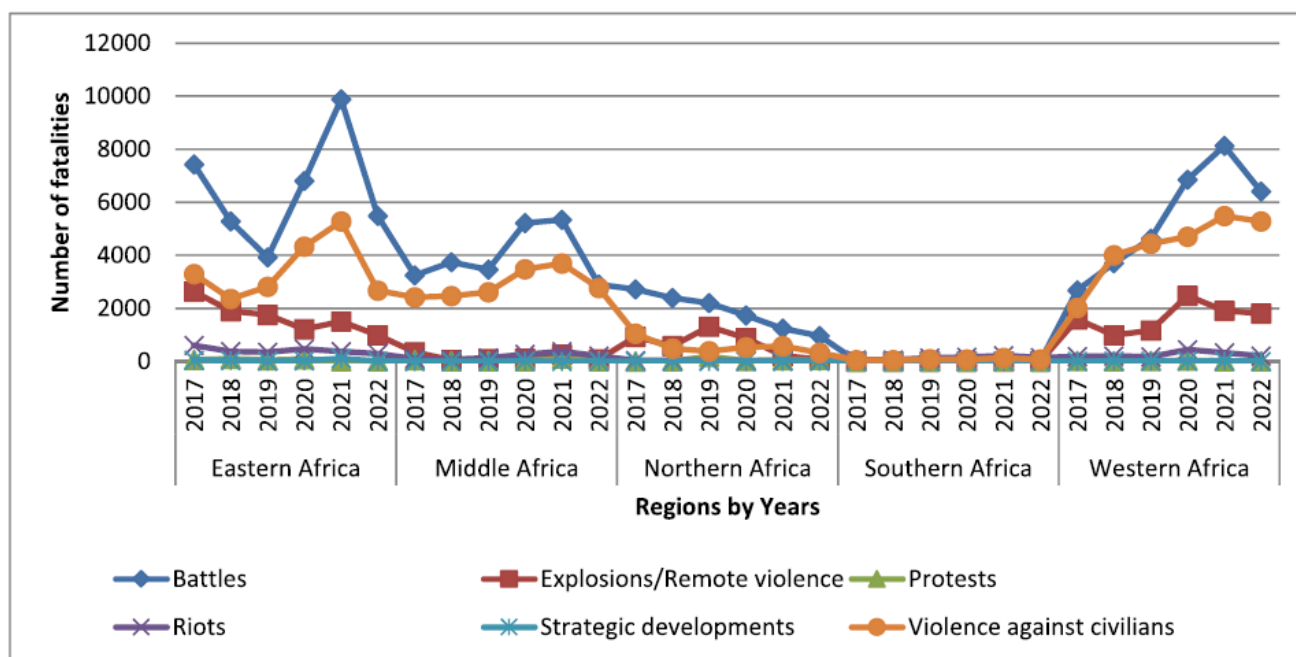


Figura 2-8: Intensidad de conflictos en África (Fuente: [16])

2.4 Herramientas software para el estudio del Dataset

Para el estudio del conjunto de datos del que nos serviremos en la elaboración del proyecto, nos serviremos de R Project -Figura 2-9-. Se trata de una iniciativa a nivel global de carácter colaborativo en la que están involucrados miles de usuarios y desarrolladores.



Figura 2-9: Logotipo R Project (Fuente: [17])

R Project es un lenguaje de programación completo que proporciona al usuario una gran variedad de técnicas, desde clasificación de datos, hasta modelado lineal y no lineal, pruebas estadísticas clásicas o análisis de series temporales. Además, proporciona una gran facilidad para la producción de gráficos. Facilita también el uso de librerías, que aportan funcionalidades extras, ampliando los márgenes de la capacidad creativa de R [17].

R, además de ser un lenguaje de programación es también un entorno de software completo. Es decir, que proporciona todo lo necesario para trabajar con un determinado conjunto de datos, desde la edición de código hasta el análisis de estos. Además, nos ofrece una amplia gama de funciones para el análisis estadístico de datos, así como herramientas para la creación de gráficos que faciliten la visualización y comprensión de los datos [18].

3 DESARROLLO DEL TFG

3.1 Matrices de datos y principales variables

Para la realización de este proyecto nos basaremos en la base de datos de UCDP, ya que nos aporta información sobre los principales aspectos de relevancia para nuestro estudio -año, localización, partes en conflicto, número de muertes por bando, o tipo de violencia- e incluye un margen temporal amplio que permite un correcto análisis de diferentes hipótesis. Esta base de datos previamente será depurada con el fin de obtener una únicamente con información de interés para nuestro análisis. Para ello se seleccionaron las zonas geográficas a tener en cuenta, descartando el continente americano debido a un menor número de grandes conflictos y a que ninguna de las principales hipótesis de estudio planteadas en un principio incluye esta zona. Además, se filtraron las variables, eliminando las que no eran relevantes para nuestros cálculos como la precisión de las coordenadas, diversos datos acerca de la fuente de la que provenían los datos, o la claridad de lo sucedido en cada evento, entre otros.

Cada fila constituye un enfrentamiento mientras que cada columna constituye un aspecto o característica de dicho conflicto, desde el nombre del conflicto, las partes involucradas, el tipo de violencia, o el número de bajas hasta las fechas de inicio y fin, las coordenadas de la localización donde tuvieron lugar los hechos, o el continente y la localidad -ver *Tabla 3-1*-.

VARIABLE	DESCRIPCIÓN
year	Año del evento
type_of_violence	Tipo de conflicto según UCDP: 1: Conflicto relativo a un gobierno y/o territorio. Al menos uno de los actores es gubernamental. 2: Conflicto entre grupos organizados (sin participación gubernamental) 3: Violencia de un actor gubernamental/grupo organizado contra civiles
conflict_ident	Código de identificación propio de cada conflicto
conflict_name	Nombre del conflicto de UCDP al que pertenece el evento

side_a	Nombre del primero de los actores intervinientes en el conflicto. En violencia de tipo 1 siempre es un gobierno, y en violencia de tipo 3 es el actor perpetrador de los actos.
side_b	Nombre del segundo de los actores intervinientes en el conflicto. En violencia de tipo 1 siempre es el movimiento rebelde o la oposición, y en violencia de tipo 3 es el actor perpetrador de los actos.
latitude	Latitud (en grados decimales)
longitude	Longitud (en grados decimales)
country	País donde tuvo lugar el evento
region	Región donde tuvo lugar el evento. Una de las siguientes: {África, América, Asia, Europa, Oriente Medio}
date_start	Fecha de comienzo (DD/MM/AAAA)
date_end	Fecha de fin (DD/MM/AAAA)
deaths_a	La mejor estimación de bajas por parte del bando “a”. Siempre 0 en los casos de violencia tipo 3
deaths_b	La mejor estimación de bajas por parte del bando “b”. Siempre 0 en los casos de violencia tipo 3
deaths_civilians_unknown	Bajas civiles y sin confirmar
best	La mejor estimación de bajas totales del conflicto

Tabla 3-1: Explicación de las variables de interés de la base de datos de UCDP (Fuente: Elaboración propia basada en [5])

Al ordenar los conflictos por orden cronológico, se detectó la presencia de filas consecutivas con el mismo nombre de conflicto, mismos actores, y misma localización, lo que apuntaba a ser partes de un mismo conflicto. Para subsanar este hecho y unificarlo todo como parte de un único conflicto, se programó un código en R que permitiera detectar aquellas filas que, siendo consecutivas, tuvieran el mismo nombre de conflicto, y las unificara en una sola realizando las siguientes operaciones: se suman el número de muertes, se escoge la fecha de inicio más temprana y la fecha de fin más tardía, se hace la coordenada media (la media de las longitudes y la media de las latitudes), se coge el año de la fecha de fin. Para el resto de las variables, dado que deberían ser iguales para todas las filas en cuestión, se cogerán los valores de la última fila del grupo (si se trata de un año activo, el tipo de violencia, nombre de la pareja de actores, actor a, actor b, país y región).

3.2 Principales hipótesis de estudio

Previo al análisis específico de los datos orientado a las hipótesis planteadas, se pueden realizar una serie de apreciaciones de carácter general sobre el conjunto de datos. Se pueden apreciar determinadas zonas con tendencia conflictiva, de modo que se analizará, con la ayuda de representaciones visuales, una posible preponderancia de la zona central de África sobre el resto del continente, así como del sur de Asia sobre la zona del norte asiático, y en un plano global, de África y Asia sobre Europa.

Posteriormente se estudiarán dos hipótesis principales, las cuales se tratará de confirmar o desmentir basándose en la base de datos de UCDP y mediante el uso de los procedimientos estadísticos que se explican en el apartado 3.3.

La primera de las hipótesis es si existe relación entre el surgimiento de nuevos conflictos en 2021-2022 y la guerra de Ucrania. Para el estudio de esta hipótesis se plasmarán sobre el mapa los conflictos cuya primera aparición fuera en 2021 o 2022. Una vez representados los datos de interés, se realizará un estudio de los países en los que dichos conflictos tuvieron lugar, para estudiar si existen posibles relaciones con Rusia. Además, se evaluarán variables de esos conflictos tales como las muertes provocadas.

La segunda hipótesis consiste en un estudio de si, a medida que desaparece la participación del ISIS en conflictos en Siria, aumenta en conflictos en África. Para ello se estudiará la evolución temporal de los conflictos y muertes en Siria y en África, comprobando si existe, o no, una relación evidente entre la decadencia de la actividad terrorista del IS en Siria con su auge en el continente africano.

3.3 Procesos estadísticos utilizados

Si comenzamos haciendo una introducción a lo que es la estadística, según la Real Academia Española [19]: “Rama de la matemática que utiliza grandes conjuntos de datos numéricos para obtener inferencias basadas en el cálculo de probabilidades”. Trabajando con procesos estadísticos se hace amplio uso de los términos *población* y *muestra*: una *población* es el conjunto de elementos que se quiere estudiar, y a cada elemento de esa población se le denomina individuo; una *muestra* es un subconjunto limitado extraído de la población con la finalidad de reducir experiencias, normalmente será el conjunto de datos con el que trabajaremos.

Otro concepto de frecuente empleo es el de *variable*. Se define una *variable* como cualquier característica de un individuo que pueden tomar diferentes valores para distintos individuos. Además, denominamos *modalidad* a cada uno de los tipos de valores que puede tomar una variable estadística. Estas variables a su vez se pueden clasificar en categóricas -o cualitativas- y numéricas -o cuantitativas-. Las variables categóricas son aquellas que indican a qué grupo o categoría pertenece un individuo, mientras que las variables cuantitativas toman valores numéricos para los que tiene sentido hacer operaciones aritméticas.

La estadística a su vez se divide en dos ramas fundamentales: la estadística descriptiva y la inferencia estadística. En primer lugar, la estadística descriptiva [20] consiste en el recuento, ordenación y clasificación de los datos obtenidos en las observaciones. En esta rama de la estadística se realizan tablas y gráficos que permiten simplificar los datos y favorecen su entendimiento y análisis. Además, se calculan parámetros estadísticos que caracterizan la distribución. Es decir, se limita a realizar deducciones a partir de los datos y parámetros obtenidos.

Por otra parte, la inferencia estadística [20] es la rama de la estadística dedicada a obtener conclusiones generales sobre una o más variables aleatorias de la población a partir de los datos observados en la muestra. Es decir, trabajar con un conjunto reducido de datos (muestra) de modo que mediante métodos y algoritmos matemáticos se logre obtener una conclusión aplicable al total de la población, o incluso llegar a predecir el comportamiento de la variable estudiada. Esta rama de la estadística se sirve principalmente de dos procedimientos: estimación (puntual e intervalos de confianza) y contraste de hipótesis.

- Estimación puntual y estimación por intervalos de confianza:
La estimación puntual [21] consiste en el establecimiento de un valor concreto para el parámetro. El valor escogido es el suministrado por un estadístico³ concreto. Este estadístico suele llamarse *estimador*. De modo que, por ejemplo, se podría usar el estadístico “media aritmética de la muestra”

³ Función real medible de la muestra de una variable aleatoria

como el estimador del parámetro “media aritmética de la población”. Los intervalos de confianza [22] se emplean en ocasiones en las que la estimación puntual es poco informativa, y es preferible dar una estimación por intervalo, en el que se pretende hallar el valor del parámetro con una elevada probabilidad.

- El contraste de hipótesis [22] nos permite responder a preguntas concretas o verificar afirmaciones sobre los parámetros poblacionales de interés. Para ello se establece una hipótesis nula⁴ (H_0) y una hipótesis alternativa (H_1) contraria a la primera.

Se rechazará H_0 cuando los valores del estadístico sean “grandes” y la aceptaremos cuando sean “pequeños” en comparación con un estadístico teórico basado en el error de significación⁵ α que el investigador esté dispuesto a cometer (habitualmente $\alpha=0.05$). Se define así el p-valor como la probabilidad de que suceda lo observado o algo más raro (siendo esto último un suceso con una probabilidad menor que el observado).

- Si $p\text{-valor} \geq \alpha$ aceptamos la hipótesis nula y por tanto no hay evidencias de que la afirmación establecida en ella sea falsa. Por lo tanto, consideramos H_0 como válida.
- Si $p\text{-valor} < \alpha$ rechazaremos la hipótesis nula, aceptando la alternativa, lo que implica que hay evidencias en contra de la afirmación establecida en H_0 . Por lo tanto, consideramos que existe una elevada certeza de que H_1 es válida.

Si la/s variable/s de estudio siguen una distribución normal, se dice que el contraste de hipótesis se realiza mediante test paramétricos, mientras que, en otro caso, se emplean test no paramétricos.

3.3.1 Gráficos

El hecho de examinar los datos para describir sus características principales se llama análisis exploratorio de los datos, y esto constituye el primer paso de todo trabajo de estadística. En los trabajos estadísticos, los métodos gráficos cobran una especial importancia ya que ayudan a una fácil visualización, comprensión y asimilación tanto de los datos en una primera instancia como de los resultados una vez finalizado el estudio. Son varios los gráficos de uso recurrente en los estudios estadísticos:

- Diagramas de barras [23]: Cada modalidad aparece representada por un rectángulo, todos ellos con la misma base, y donde las alturas de cada uno de ellos corresponden a la frecuencia (absoluta o relativa), estableciéndose así una relación de proporcionalidad entre áreas y frecuencias. Los diagramas de barras se representan tal y como se indica en la *Figura 3-1*. Este tipo de gráficos es normalmente empleado con variables categóricas y variables cuantitativas discretas cuyas variables tomen un rango relativamente pequeño de valores.

⁴ Hipótesis que consideramos como verdadera hasta que haya evidencias de lo contrario

⁵ Probabilidad de que ocurra el error tipo I -que se rechace H_0 , siendo esta cierta-

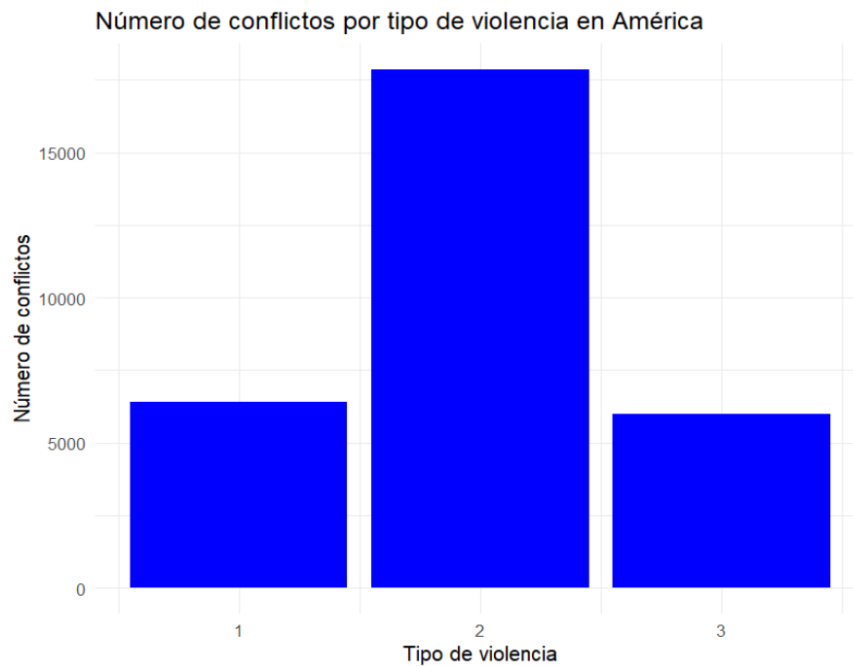


Figura 3-1: Diagrama de barras (Fuente: Elaboración propia)

- Diagrama de sectores [23]: Cada modalidad representa un sector, donde el ángulo es proporcional a la frecuencia (absoluta o relativa). Este tipo de diagramas nos ayuda a visualizar la importancia relativa de cada modalidad respecto al total, tal y como se observa en la *Figura 3*, facilitando en gran medida la comprensión de los datos. Los diagramas de sectores son empleados con frecuencia para el estudio de variables categóricas.

Distribución de conflictos por tipo en América

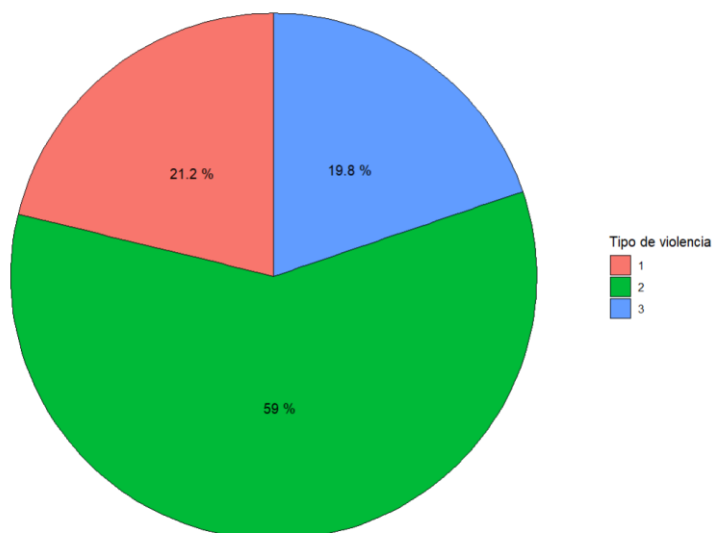


Figura 3-2: Diagrama de sectores (Fuente: Elaboración propia)

- Histograma [23]: Consiste en un conjunto de rectángulos en el que cada uno de ellos representa un intervalo de agrupación o clase. En este tipo de gráficos, los rectángulos tienen por altura la frecuencia (relativa o absoluta) de las observaciones del intervalo. Y, aunque todos los intervalos suelen tener la misma anchura, puede darse el caso de que no fuera así, en cuyo caso cobraría especial importancia el parámetro de *densidad de frecuencia* -calculado mediante la división de la frecuencia de cada intervalo entre el ancho de este- de modo que el área de cada

rectángulo represente la frecuencia relativa de los intervalos o clases. En caso de tratarse de frecuencias relativas, el valor total del área del histograma ha de ser 1. Además, las barras de un histograma han de cubrir todo el recorrido de una variable. Todo ello aparece reflejado en la *Figura 3-2*. Buena parte del valor aportado por los histogramas reside en su interpretación, se deberá prestar especial atención a la forma del histograma (a su centro y a su dispersión) así como a la aparición de posibles observaciones atípicas. Los histogramas se utilizan cuando se trabaja con variables numéricas o cuantitativas.

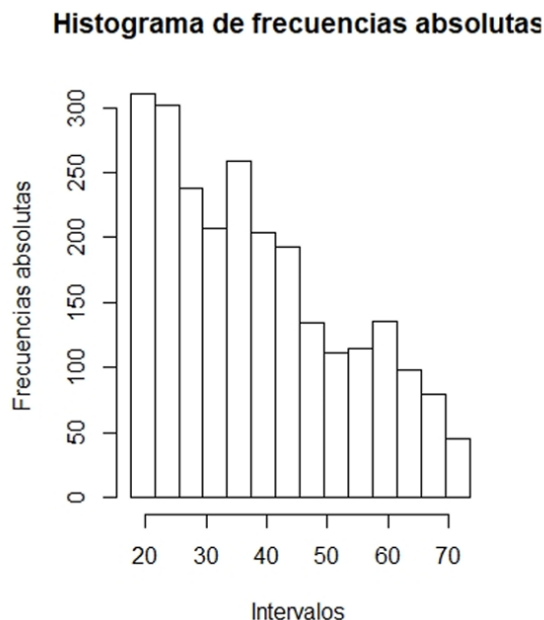


Figura 3-2: Histograma (Fuente: [24])

- Gráficos temporales [23]: Muchas variables se miden a lo largo del tiempo, y en estos casos es cuando los gráficos temporales cobran especial importancia ya que representan cada observación en relación con el momento en que se midió. En estos gráficos siempre se representa el tiempo en el eje de las abscisas mientras que el de las ordenadas se reserva para los valores obtenidos de las mediciones. La unión de los puntos contiguos con un trazo continuo facilita la visualización de los cambios a lo largo del eje temporal. En el análisis de este tipo de gráficos es importante fijarse, una vez más en su aspecto general y en las desviaciones atípicas. Un aspecto que aparece con frecuencia es la tendencia⁶. En la *Figura 3-3*, podemos observar un gráfico temporal donde se aprecia una tendencia de tipo creciente sin desviaciones notables.

⁶ Patrón de comportamiento de las muestras tomadas de una determinada población durante un periodo determinado

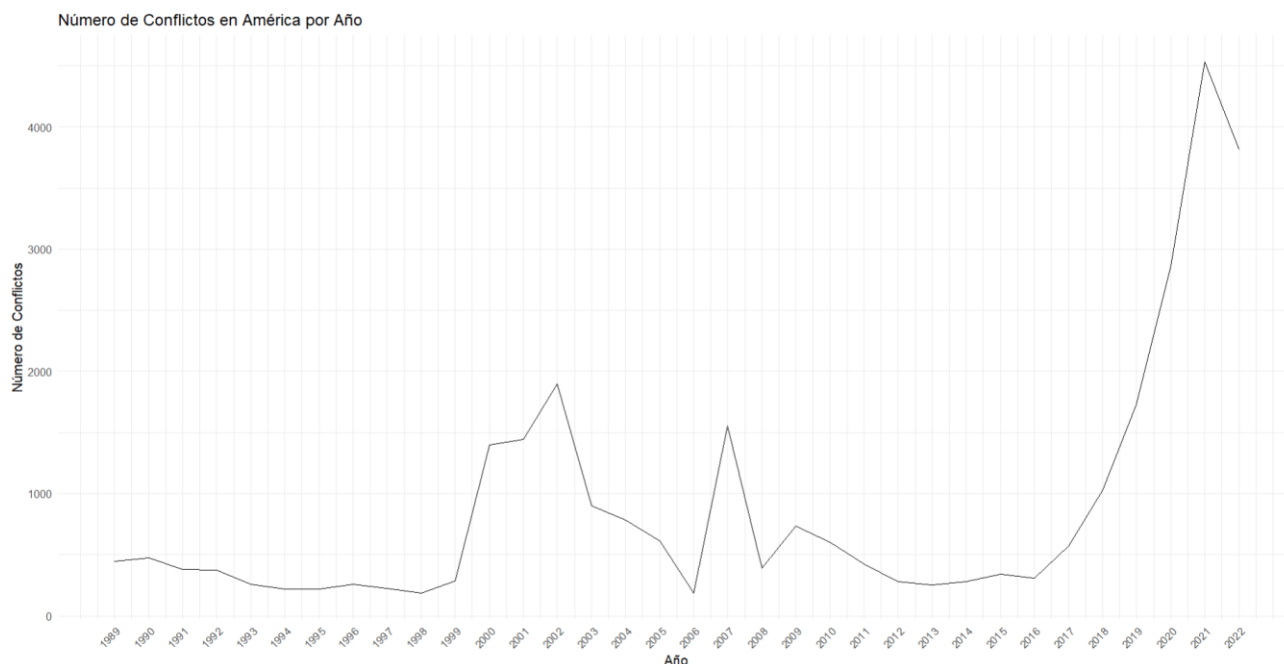


Figura 3-3: Gráfico temporal (Fuente: Elaboración propia)

- Diagramas de caja [23]: Muestran de una forma gráfica los cinco números resumen, tal y como se muestra en la *Figura 3-4* -observación mínima, primer cuartil⁷, mediana⁸, tercer cuartil⁹, observación máxima- de modo que:
 - Los lados superior e inferior de la caja corresponden a los cuartiles.
 - La línea horizontal representada en el interior de la caja corresponde a la mediana.
 - Los extremos superior e inferior de los segmentos verticales que salen de la caja corresponden a los valores máximos y mínimos respectivamente.
 - En otras ocasiones, cuando el muestreo presenta datos atípicos, estos aparecen por encima del extremo que representa el máximo y/o por debajo del extremo que representa el mínimo.

⁷ Representa el valor que separa el 25% inferior de los datos del 75% superior

⁸ Representa el valor que separa el 50% inferior de los datos del 50% superior

⁹ Representa el valor que separa el 75% inferior de los datos del 25% superior

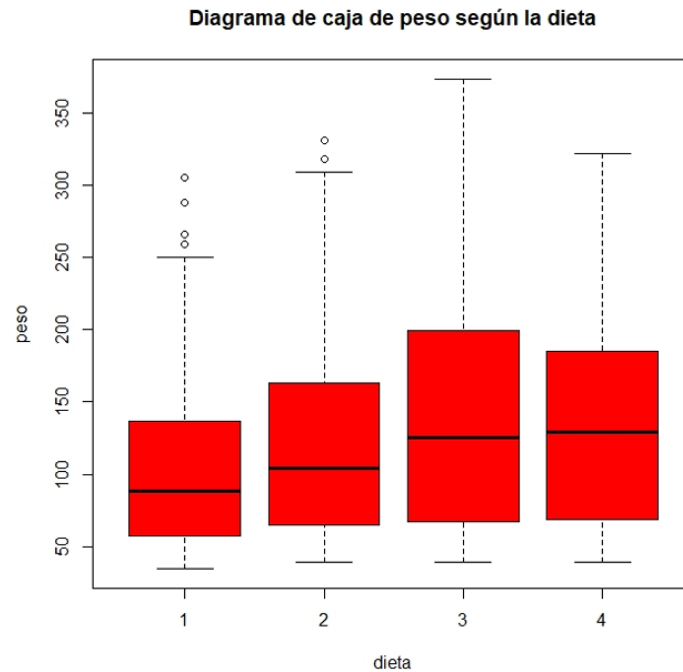


Figura 3-4: Diagrama de caja (Fuente: [24])

- Diagramas de dispersión [23]: Son la manera más común de representar la relación entre dos variables cuantitativas medidas en los mismos individuos. Se representan de modo que los valores de una de las variables aparecen en el eje de abscisas y los valores de la otra variable en el eje de ordenadas -ver *Figura 3-5*-. Cada individuo aparece como un punto del diagrama, cuya posición depende de los valores que tomen las variables para cada individuo. Como norma general, en caso de haber una variable explicativa, esta se colocará en el eje de las abscisas.

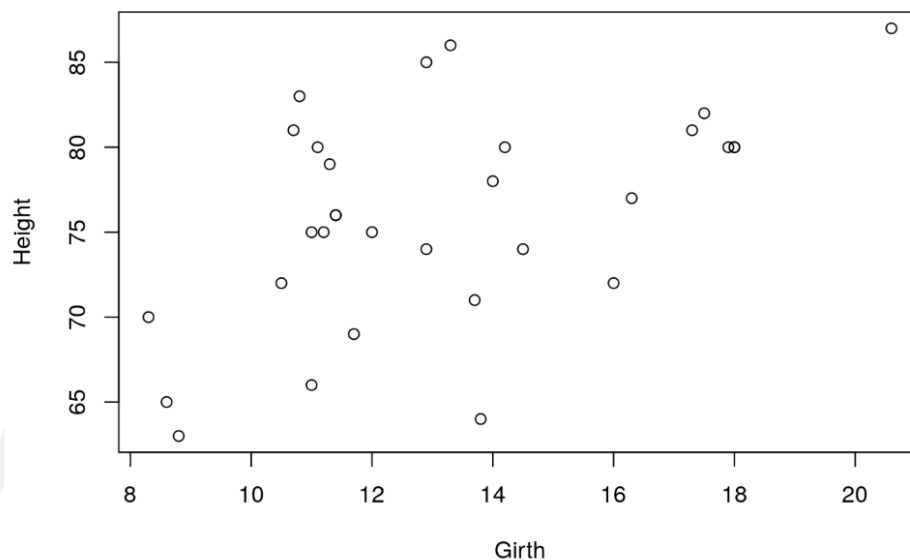


Figura 3-5: Diagrama de dispersión (Fuente: [25])

3.3.2 Modelos de regresión

La regresión es un concepto de especial relevancia en los estudios estadísticos, especialmente a partir del empleo de gráficos de dispersión. Existen varios tipos de regresión en función de la relación

existente entre los datos, donde habrá que tener en cuenta en cada caso los supuestos sobre las variables de estudios y los errores asociados al modelo.

- Modelo de regresión lineal o de regresión simple [26]: También conocido como recta de regresión, estudia la relación lineal entre una variable respuesta (variable independiente) ubicada en el eje Y, y una covariable (variable independiente) ubicada en el eje X según el modelo:

$$Y_i = \beta_0 + \beta_1 X_i + error_{aleatorio}$$

siendo β_0 y β_1 parámetros desconocidos que tendrán que ser estimados a partir de los datos.

- Modelos de regresión no lineal [26]: Se emplean cuando la función $f(X)$ de la expresión $Y = f(X) + error_{aleatorio}$ no es lineal (polinómica de grado igual o mayor que dos, exponencial, etc.). Existen modelos de regresión no lineales que pueden ser resueltos como una recta de regresión simplemente aplicando una transformación de los datos, que se denominan modelos de regresión linealizables.

Todo modelo de regresión tiene asociado un coeficiente de determinación (R^2) el cual indica el porcentaje de variabilidad explicada por el modelo. Así mismo para un modelo de regresión en una gráfica determinada, se obtiene un p-valor -a partir de la varianza de la variable Y debida al ajuste y la varianza debida al azar- que permitirá decidir si el modelo es adecuado o no.

3.3.3 Tablas de contingencia

Son tablas formadas por las frecuencias relativas para las dos clasificaciones y sus correspondientes categorías -ver *Tabla 3-2*-. En ellas se supone que las dos características son independientes, de modo que bajo la hipótesis nula de independencia se desea saber si existe una diferencia considerable entre las frecuencias observadas y las esperadas, tal que la hipótesis nula se rechace [27].

	Categorías	Característica B				Totales
		1	2	...	c	
Característica A	1	n_{11}	n_{12}	...	n_{1c}	$n_{1\cdot}$
	2	n_{21}	n_{22}	...	n_{2c}	$n_{2\cdot}$
	...	...	...	...	...	...
	...	...	...	...	...	...
	r	n_{r1}	n_{r2}	...	n_{rc}	$n_{r\cdot}$
	Totales	$n_{\cdot 1}$	$n_{\cdot 2}$	...	$n_{\cdot c}$	n

Tabla 3-2: Tabla de contingencia (Fuente: Elaboración propia)

En esta tabla encontramos los siguientes elementos:

- n_{ij} : Frecuencias observadas
- $n_{r\cdot}$: Frecuencia total de la fila
- $n_{\cdot c}$: Frecuencia total de la columna
- n: Número total de individuos

Con la finalidad de descubrir si las características son o no independientes, se realiza el ajuste chi-cuadrado, en el que se establece una hipótesis nula (H_0) y una hipótesis alternativa (H_1).

$$\begin{cases} H_0: X \text{ independiente de } Y \\ H_1: X \text{ dependiente de } Y \end{cases}$$

Para realizar el contraste, se realiza el ajuste chi-cuadrado en el que, basado en los valores de la *Tabla 3-2*, se obtiene la fórmula que define la frecuencia esperada:

$$e_{ij} = \frac{n_{i.} \times n_{.j}}{n}$$

De esta manera, podemos definir el estadístico de contraste, que nos proporciona el p-valor como:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

que nos proporciona el p-valor para decidir sobre las hipótesis planteadas.

3.3.4 Comparación de medias/medianas

Cuando las distribuciones de las poblaciones no satisfacen las condiciones necesarias para otros test paramétricos, el test no paramétrico “Mann-Whitney-Wilcoxon” [28] nos permite contrastar si dos muestras proceden de poblaciones equidistribuidas. Es decir, suponiendo una variable X de estudio es medida en dos grupos distintos A y B, que a priori no son independientes, se quiere hacer inferencia sobre la afirmación de que la medida de tendencia central es igual en ambos grupos. Para ello se establecerán las siguientes hipótesis:

$$\begin{cases} H_0: Me_A = Me_B \\ H_1: Me_A \neq Me_B \end{cases}$$

En este test, primero se calcula la diferencia de cada par de observaciones. Se asigna al valor absoluto de cada diferencia un número de posición, siendo la posición 1 para el valor más pequeño. En caso de haber valores repetidos, fenómeno conocido como ligadura, se les asigna como valor de posición la media de las posiciones que ocupan. En la asignación de posiciones se ignoran las diferencias que sean 0. Posteriormente, se calcula el estadístico W :

$$W = \min(W+, W-)$$

- $W+$ = suma de los rangos con signo positivo
- $W-$ = suma de los rangos con signo negativo

Una vez calculado el estadístico W , se puede calcular cuál es la probabilidad de que adquiera valores iguales o más extremos que el observado, es decir, el p-valor.

4 RESULTADOS

Con el objetivo de facilitar la lectura y comprensión del trabajo realizado, este apartado se ha estructurado en una primera parte en la que se hace un análisis descriptivo de la base de datos, seguida de la segunda parte en la que se hace un análisis exploratorio y algún contraste de hipótesis con la finalidad de obtener una idea general de los datos contenidos en UCDP, posteriormente se plantean una serie de hipótesis que corroborar o desmentir basándonos en un estudio de los datos disponibles. Durante todo el proceso se trabajará con el margen temporal que nos proporciona UCDP, siendo este de 1989 a 2022.

4.1 Preprocesado de la base de datos

Esta fase comienza con la importación de la base de datos de UCDP, que contiene los conflictos bélicos de importancia significativa ocurridos entre 1989 y 2022, de la que nos hemos quedado con los campos mostrados en la *Tabla 3.1*, que son las variables que se han considerado relevantes a la hora de trabajar con los datos en R Project.

Lo siguiente a realizar fue la simplificación o fusión de la base con el objetivo de facilitar el manejo de esta. Se unificaron las filas correspondientes a un mismo conflicto, pero realizándolo anualmente, evitando así perder información, de modo que si un conflicto estallaba en un año determinado pero pasado un periodo de tiempo volvía a surgir, se mostrara en todos los años correspondientes en vez de unificarlo todo en una única fila en la que no se podría diferenciar este tipo de situaciones. En esta operación, se modificaron las variables de longitud y latitud (haciéndose la media de cada una de ellas), la de las fechas de inicio y fin (cogiendo el inicio más temprano y el fin más tardío), y todas las características referentes a muertes, tanto de un bando y otro, como de civiles, y el total de muertes (a las que se les aplicó la suma, quedando el sumatorio de todas las filas de ese año pertenecientes a dicho enfrentamiento). El resto de las variables, dado que eran comunes a todas las filas de un mismo conflicto, no sufrieron modificaciones.

4.2 Primeros análisis

La primera operación que se llevó a cabo con esta base de datos fue representar todos los conflictos sobre un mapamundi para ver la distribución de los enfrentamientos de forma espacial, tal y como se muestra en la *Figura 4-1*.

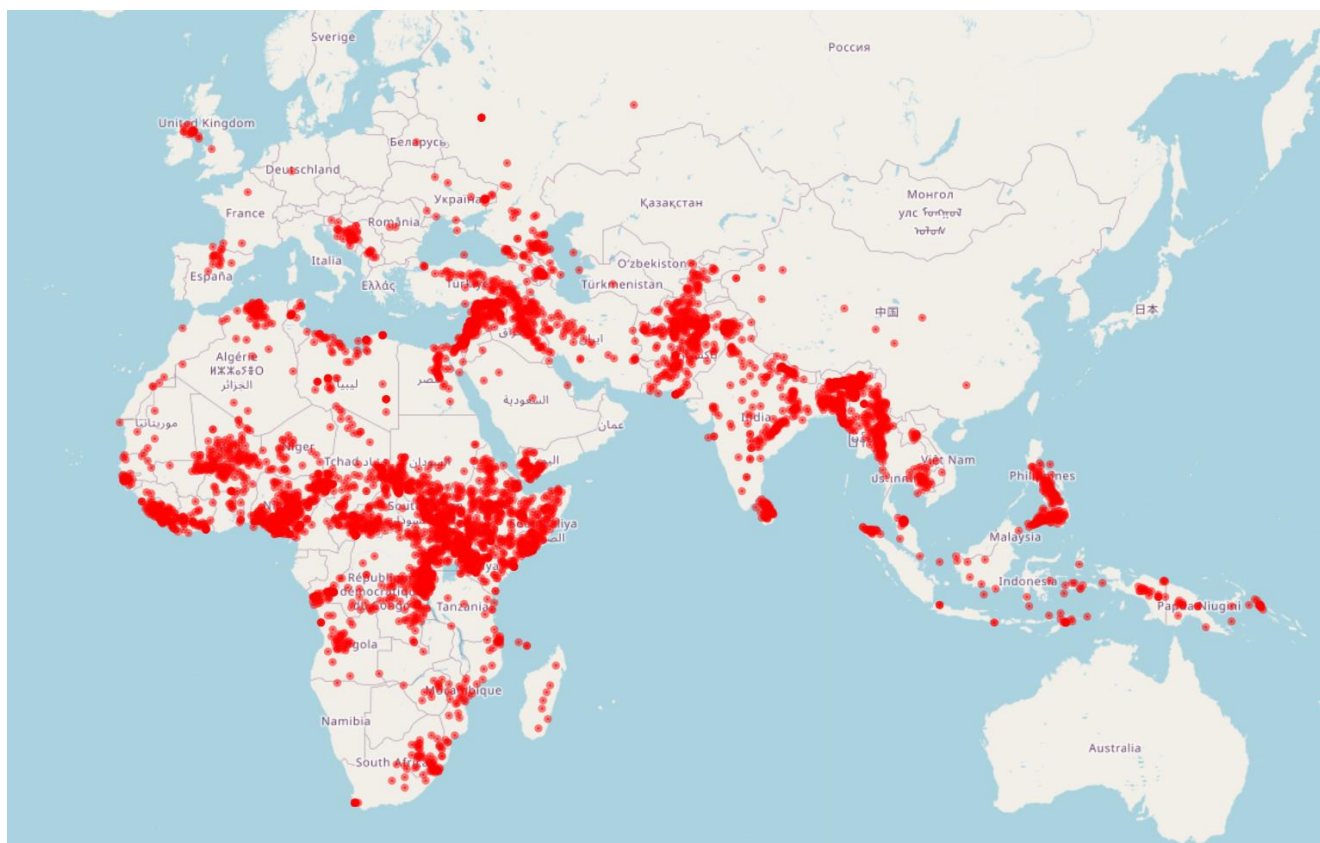


Figura 4-1: Gráfico espacial de conflictos (Fuente: Elaboración propia)

En este gráfico, podemos observar una serie de zonas en las que la concentración de conflictos es mayor. Destacan zonas como la franja central de África (con especial énfasis en la zona del cuerno de África y alrededores), la zona de Oriente Medio y la zona sur de Asia. No obstante, se observa una menor concentración de conflictos en Europa. Sin embargo, dado que se muestran conflictos desde 1989, cabe destacar los conflictos representados en España -correspondientes a los atentados perpetrados por el grupo terrorista ETA y grupo islamista ISIS- y en Irlanda -correspondientes en su mayoría a los conflictos con participación del IRA¹⁰-.

Con la finalidad de obtener una primera visión general y puramente descriptiva de los datos, se crearon dos tablas de frecuencias, la *Tabla 4-1* y la *Tabla 4-2*, en las que se recogen los datos de número de conflictos primero por cada región y posteriormente por cada tipo de violencia.

Region	Conflicts
Africa	3410
Asia	1813
Europe	254
Middle East	799

Tabla 4-1: Número de conflictos registrados por regiones (Fuente: Elaboración propia)

¹⁰ Grupo paramilitar de Irlanda que pide la reunificación de Irlanda y su total independencia de Reino Unido

Type of violence	Conflicts
1	1944
2	2133
3	2199

Tabla 4-2: Número de conflictos registrados por tipos de violencia (Fuente: Elaboración propia)

Podemos observar que en la región donde más número de conflictos tienen lugar es en África con un amplio margen respecto al resto de regiones, siendo Europa la región con menor número de conflictos. Por otro lado, también se aprecia que la distribución por tipo de violencia es bastante equitativa, siendo las tres cifras de número de conflictos muy similares.

Posteriormente, se contabilizaron las personas fallecidas correspondientes a los conflictos mostrados en la *Tabla 4-2* -un total de 3186068 fallecidos- y se elaboró el diagrama de sectores mostrado en la *Figura 4-2*, en el que aparecen representados los porcentajes correspondientes a las bajas provocadas por cada tipo de violencia respecto al total de bajas. Se puede observar que más de la mitad de las bajas son causadas por los conflictos de tipo 1 -conflictos relativos a un gobierno y/o territorio en el que al menos uno de los actores es gubernamental [5]-. Un claro ejemplo de este tipo de eventos es el conflicto en Siria, que comenzó como una serie de protestas antigubernamentales y pronto se convirtió en una guerra civil. Por el contrario, también se observa en el gráfico que los conflictos de tipo 2 -conflictos entre grupos organizados sin participación gubernamental- es de tan solo un 6.6%, constituyéndose como el tipo de violencia con menos bajas causadas.

Porcentaje de bajas por tipo de violencia

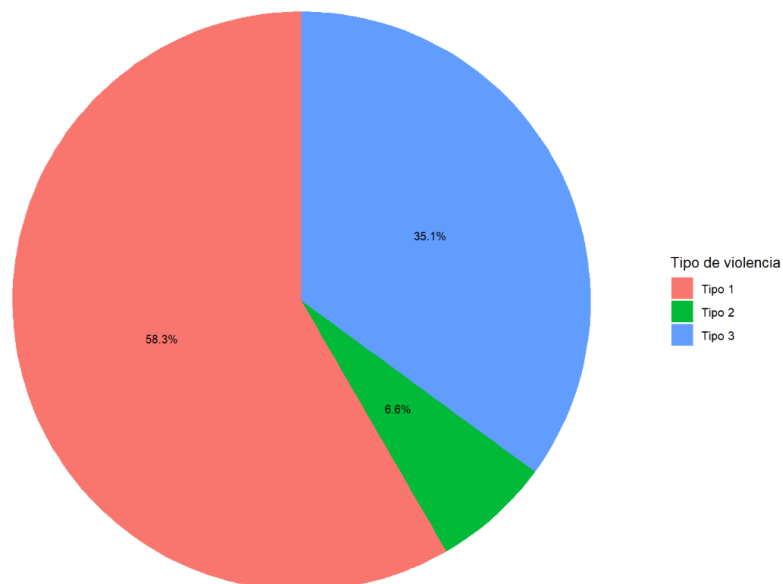


Figura 4-2: Proporción de bajas causadas por tipo de violencia (Fuente: Elaboración propia)

Se realizó también un estudio de las medidas descriptivas de los tres tipos de muertes, tanto las de ambos bandos como las del personal civil. Los resultados fueron los mostrados en la *Tabla 4-3*.

	Media	Mediana	Desviación estándar	Q1	Q3
deaths_a	91.55	0	941.22	0	13
deaths_b	112.39	0	876.58	0	16
deaths_civilians_unknown	303.71	14	9738.94	2	58

Tabla 4-3: Medidas descriptivas de los tipos de muertes (Fuente: Elaboración propia)

Podemos observar que la población civil, de media, es la que más bajas sufre, y la que cuenta con una mayor desviación estándar, lo que indica que, en condiciones normales, es el segmento poblacional más castigado en los conflictos. Este aspecto se desarrollará con más detalle al final del apartado 4.1.

Lo siguiente que se realizó fue una tabla de frecuencias con los años y el número de conflictos acaecidos en cada año, tal y como se muestra en la *Tabla 4-4*.

Año	N.º de conflictos	Año	N.º de conflictos
1989	108	2006	153
1990	131	2007	165
1991	141	2008	171
1992	178	2009	158
1993	172	2010	143
1994	164	2011	155
1995	150	2012	184
1996	164	2013	191
1997	149	2014	230
1998	165	2015	227
1999	161	2016	240
2000	177	2017	275
2001	171	2018	271
2002	171	2019	261
2003	180	2020	274
2004	162	2021	236
2005	157	2022	241

Tabla 4-4: Número de conflictos por año (Fuente: Elaboración propia)

Los datos de la *Tabla 4-4* se representaron en un gráfico de barras (mostrado en la *Figura 4-3*) con la finalidad de obtener una idea de la evolución temporal a lo largo de los años del número de conflictos.

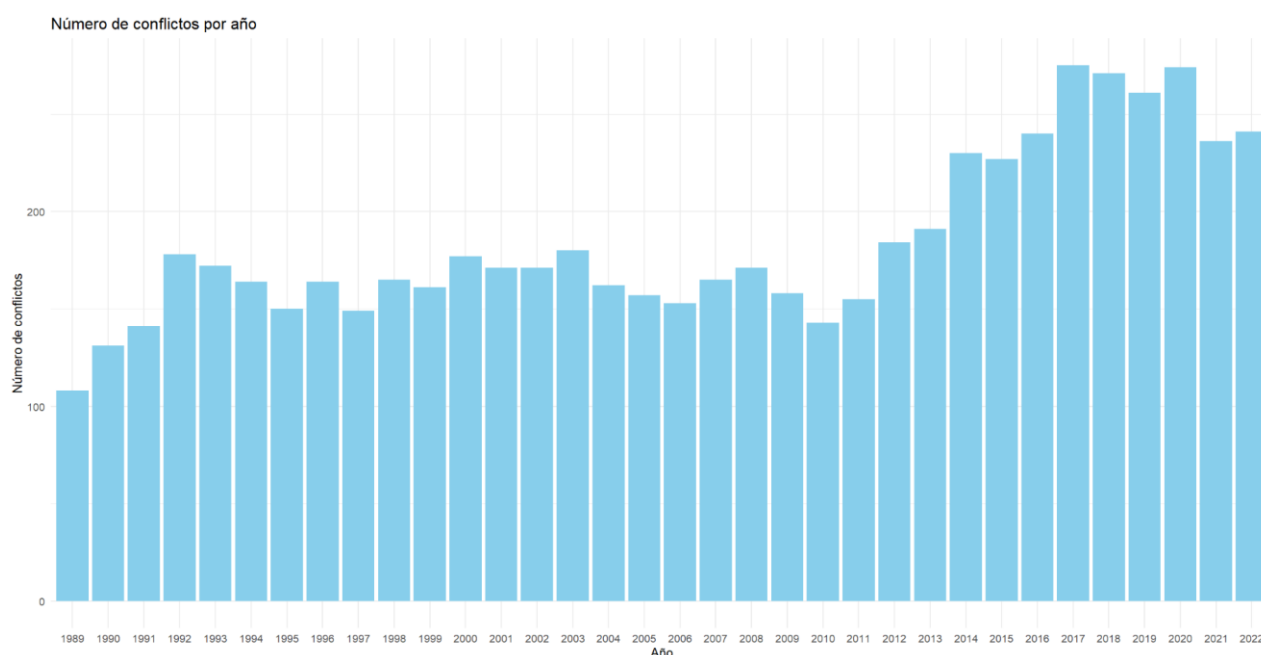


Figura 4-3: Conflictos ocurridos por años (Fuente: Elaboración propia)

Podemos observar cómo la tendencia es claramente al alza, es decir, la tendencia histórica es que aumente el número de conflictos. Se aprecia también una significativa subida en el número de conflictos en la segunda década del S. XXI.

Se procedió de la misma manera, pero esta vez con el número de muertes por año, creando primero la tabla de frecuencias mostrada en la *Tabla 4-5*.

Año	Bajas	Año	Bajas
1989	57922	2006	26321
1990	92059	2007	26310
1991	81346	2008	35338
1992	72757	2009	44120
1993	66598	2010	27376
1994	821229	2011	35841
1995	59656	2012	84103
1996	70333	2013	110393
1997	63340	2014	147740
1998	59324	2015	126171
1999	96863	2016	111039
2000	89268	2017	100210
2001	29764	2018	74560
2002	35443	2019	65474
2003	36079	2020	67986
2004	32641	2021	99336
2005	17617	2022	221511

Tabla 4-5: Número de bajas causadas por año (Fuente: Elaboración propia)

Los datos de la *Tabla 4-5* se representaron en un gráfico de barras (mostrado en la *Figura 4-4*) con la finalidad de obtener una idea de la evolución temporal a lo largo de los años del número de muertes.

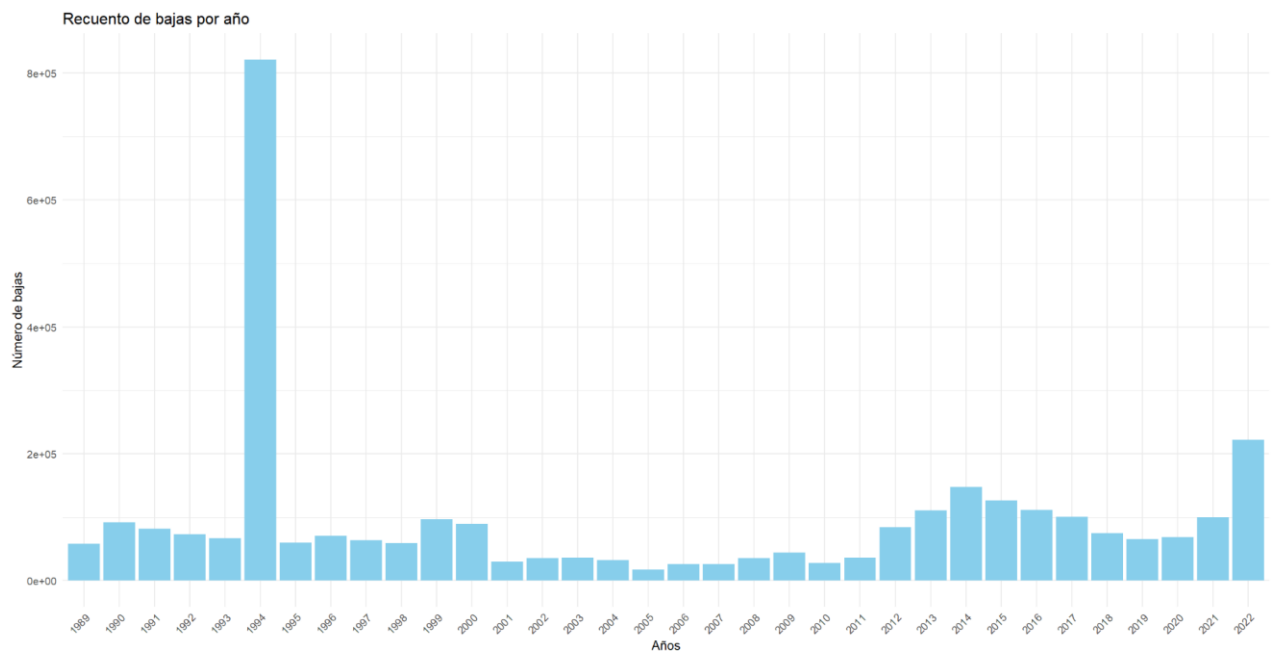


Figura 4-4: Bajas por años (Fuente: Elaboración propia)

En el gráfico se puede observar un pico prominente en el número de bajas del año 1994, que corresponde con el genocidio que perpetró el gobierno hegemónico hutu de Ruanda de aquel momento contra la población tutsi con la finalidad de exterminarla. El conflicto provocó cerca de 170.000 muertes, que supuso el exterminio del 70% de la población tutsi de Ruanda [29]. Contrastando la *Figura 4-4* con la *Figura 4-3*, se puede apreciar que no se sigue una relación directa entre el número de conflictos y el número de muertes, se puede observar claramente cómo el número de conflictos -ver *Figura 4-3*- aumenta a partir de 2014, pero, sin embargo, las bajas -ver *Figura 4-4*- disminuyen a partir de 2014.

4.2.1 Estudio de variables cuantitativas

Una vez analizado el conjunto de datos en su totalidad, pasamos a un análisis específico de los tres tipos de muertes: las muertes del bando A, las del bando B, las muertes de civiles o de personal desconocido, y el total de todas ellas. A este efecto se dividieron los datos por tipo de muerte y a su vez por años, de modo que se calculó el total de muertes de cada tipo para cada año, y el conjunto de datos resultantes se plasmó en el gráfico de líneas mostrado en la *Figura 4-5*.

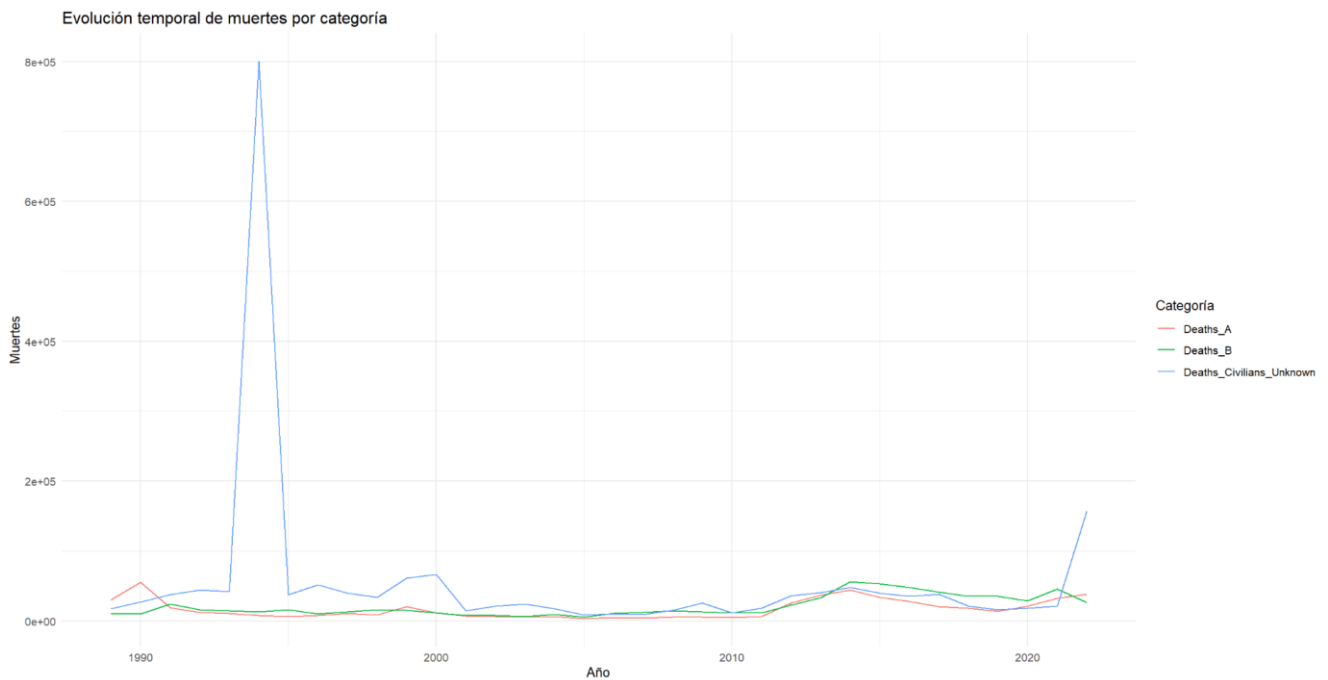


Figura 4-5: Muertes de cada tipo ocurridas anualmente (Fuente: Elaboración propia)

Observamos que hay un pico en las muertes de personal civil en 1994, correspondiente al previamente mencionado genocidio de Ruanda. También hay un valor atípico en 2022 que pertenece a un conflicto encuadrado en la Guerra de Tigray acaecida en Etiopía entre 2020 y 2022 [30]. Como estos picos en el gráfico hacen que no se vean el resto de las líneas con claridad, quitaremos el genocidio de Ruanda y el conflicto de la Guerra de Tigray para una mejor visualización de los datos. Para ello se suprimen los conflictos correspondientes a dichos eventos, repitiéndose de nuevo el gráfico sin datos atípicos, observándose el resultado en la *Figura 4-6*.

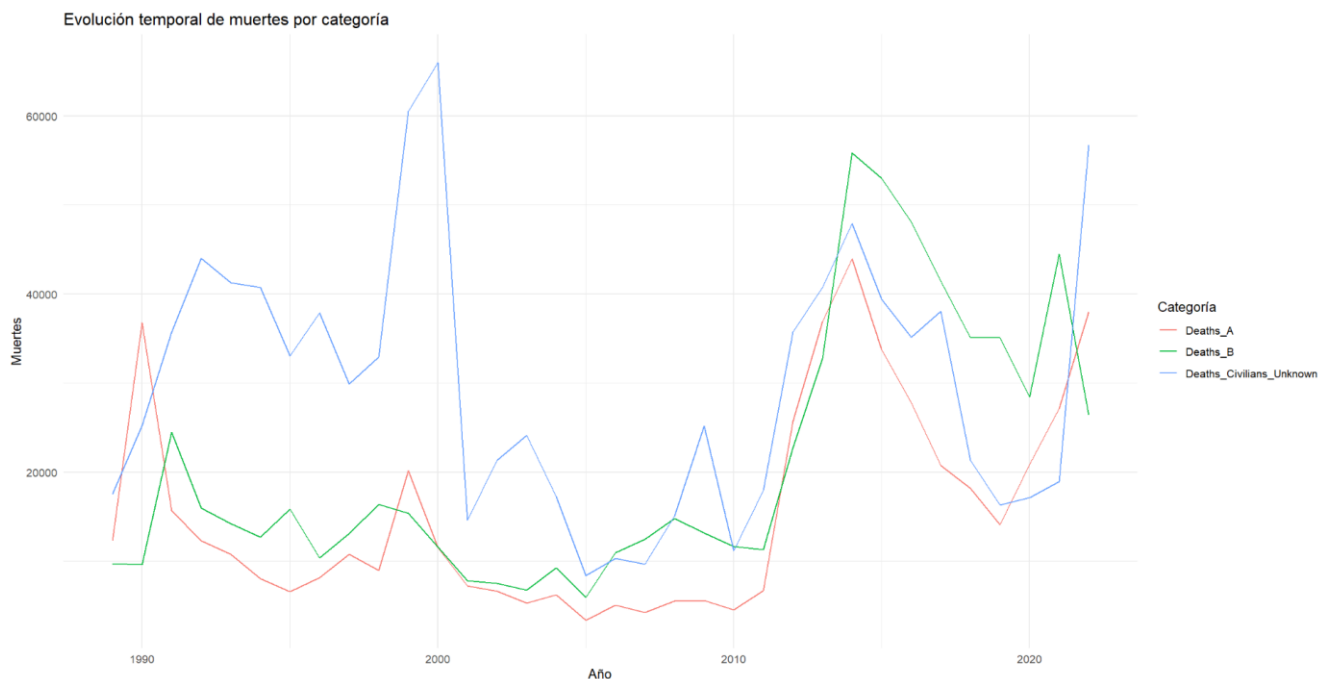


Figura 4-6: Muertes de cada tipo ocurridas anualmente sin datos atípicos (Fuente: Elaboración propia)

Observamos un pico en las muertes de civiles en los últimos años de la década de 1990, seguido de una bajada en estas al entrar en el siglo XXI. Se aprecia también una subida general de todos los tipos de muertes en la primera mitad de la década de 2010. Además, se observa una clara predominancia de las muertes de civiles sobre las de combatientes, siendo estas las más numerosas durante la mayor parte de la línea temporal, siendo superada por las muertes del bando B a mediados de la década de 2010.

Dado que en los gráficos generados hasta el momento se ve una predominancia de muertes civiles, se han generado una comparativa de gráficos de caja y bigotes para las muertes de combatientes -donde se han integrado las muertes del bando A con las del bando B- y las muertes de civiles, partiendo de los datos contenidos en la *Tabla 4-3*, tal y como se muestra en la *Figura 4-7*. Para una mejor visualización de los datos, se han considerado los datos con menos de 100 fallecidos, pues en ambas situaciones hay demasiados datos atípicos que dificultan la observación del gráfico.



Figura 4-7: Gráficos de caja y bigotes de muertes de combatientes y de civiles (Fuente: Elaboración propia)

Se puede apreciar que la distribución de las bajas en ninguno de los dos casos es simétrica, por ello, para comparar si los dos grupos son, o no, similares, y, en consecuencia, si tienen la misma tendencia, hay que realizar el test U-Mann-Whitney-Wilcoxon -ver 3.3.4- con las siguientes hipótesis de tipo unilateral:

$$H_0: Me_{combatientes} \geq Me_{civiles}$$

$$H_1: Me_{combatientes} < Me_{civiles}$$

El resultado obtenido fue de p-valor < 0.001 que es muy inferior a 0.05. Por tanto, nos confirma que ambos conjuntos no son similares, tendiendo el número de muertes a ser mayor en la población civil que entre el personal combatiente.

Por último, repetimos el proceso de sección de los datos, pero esta vez, se dividen en función de la región, es decir, se hacen dos divisiones del conjunto total, una por años, y otra por regiones, de modo que queda la información distribuida por regiones de manera anual, y todo ello se plasma en el gráfico de la *Figura 4-8*.

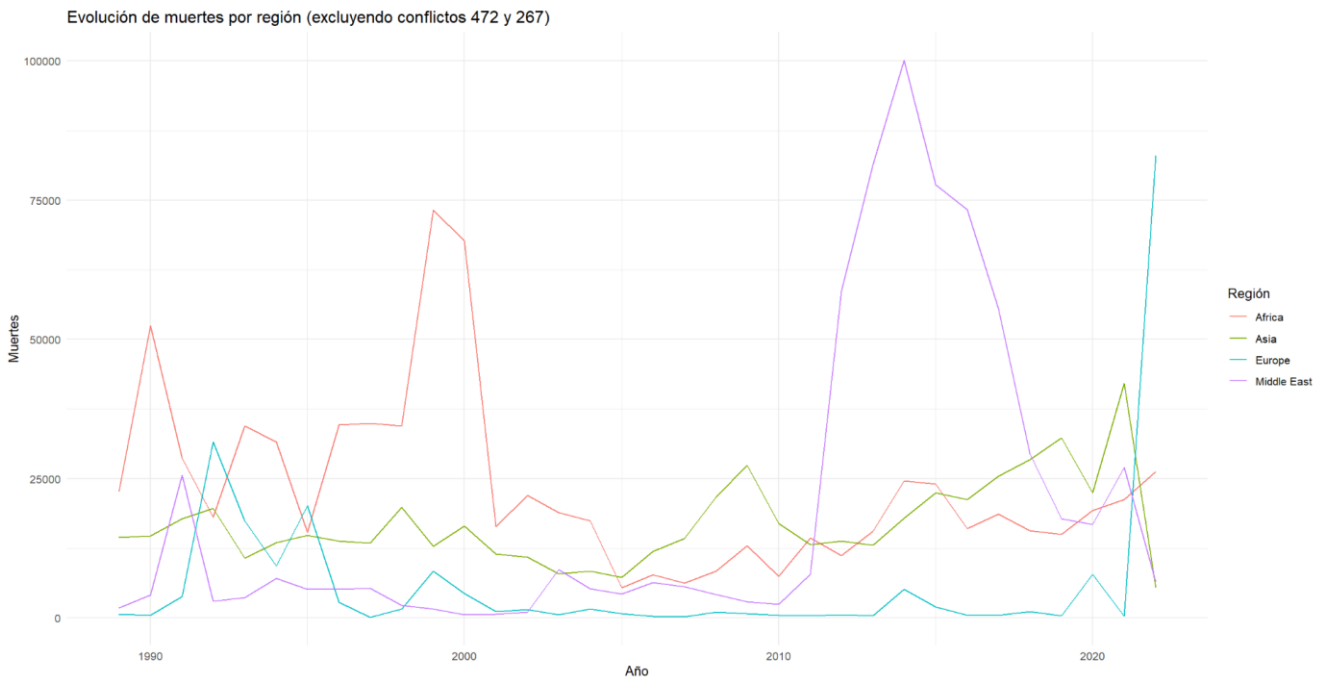


Figura 4-8: Muertes ocurridas anualmente en cada región sin datos atípicos (Fuente: Elaboración propia)

En este gráfico podemos observar un pico en África en los últimos años del siglo XX que coincide con el de las muertes de civiles observado en el gráfico anterior. Además, vemos un pico en Oriente Medio en la primera mitad de la década de 2010 que coincide con el aumento en las muertes de las tres categorías del gráfico anterior. Por último, se aprecia una subida drástica de las muertes en Europa en el 2022 correspondiente al conflicto entre Rusia y Ucrania [31].

Posteriormente, se clasificó anualmente el número de conflictos según el tipo de violencia, quedando plasmado en la *Tabla 4-6*.

Año	Tipo 1	Tipo 2	Tipo 3	Año	Tipo 1	Tipo 2	Tipo 3
1989	42	26	40	2006	45	47	61
1990	49	25	57	2007	47	44	74
1991	53	34	54	2008	48	58	65
1992	60	49	69	2009	49	54	55
1993	55	49	68	2010	47	47	49
1994	60	38	66	2011	55	48	52
1995	59	33	58	2012	52	68	64
1996	54	40	70	2013	52	80	59
1997	53	39	57	2014	64	111	55
1998	54	48	63	2015	74	93	60
1999	50	50	61	2016	75	104	61
2000	49	63	65	2017	79	124	72
2001	49	57	65	2018	75	122	74
2002	44	57	70	2019	82	99	80
2003	45	57	78	2020	82	106	86
2004	42	50	70	2021	77	83	76
2005	44	44	69	2022	79	86	76

Tabla 4-6: Número de conflictos por tipo de violencia y año (Fuente: Elaboración propia)

Estos datos también se representaron mediante el diagrama de barras apiladas mostrado en la *Figura 4-9*, en el que se ven los conflictos clasificados por tipos de violencia a lo largo de los años.

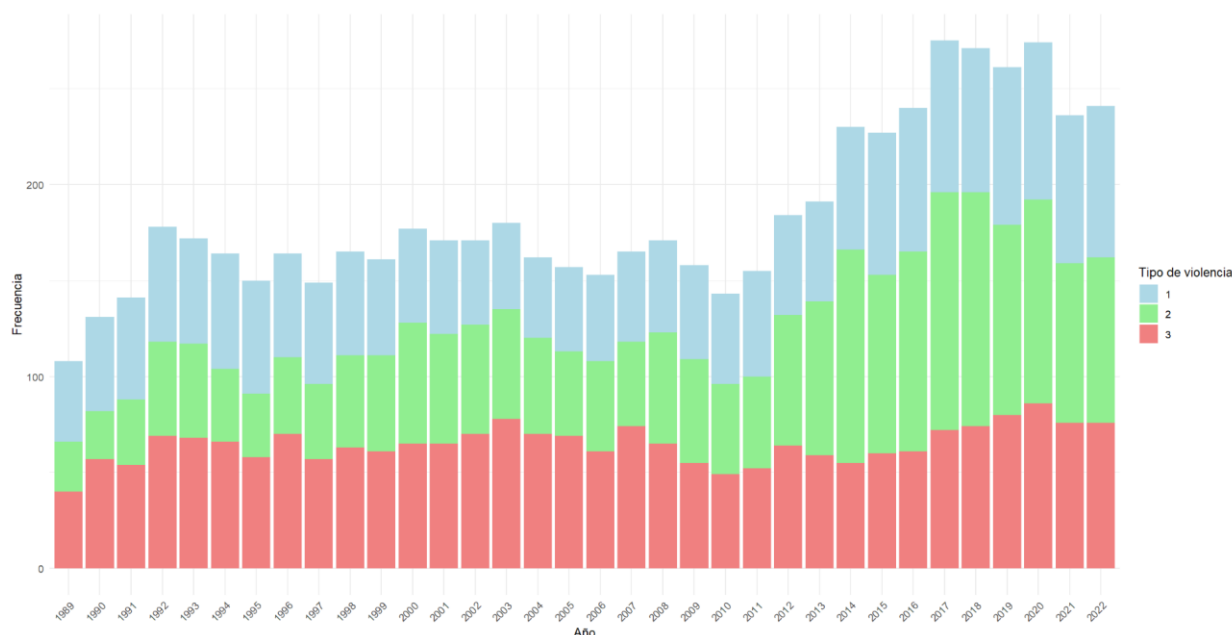


Figura 4-9: Conflictos acaecidos por cada tipo de violencia por años (Fuente: Elaboración propia)

En este gráfico podemos observar cómo las barras de los tres colores tienden a tener tamaños similares durante la mayoría del periodo de estudio, no obstante, se aprecia que, a partir de 2014, las barras verdes (correspondientes al tipo de violencia 2) comienzan a ser de mayor tamaño que el resto.

Cabe recordar que los conflictos tipo 1 son conflictos relativos a un gobierno y/o territorio en los que, al menos uno de los bandos, es un gobierno; los de tipo 2, aquellos entre grupos organizados; y los de tipo 3 representan la violencia gubernamental contra civiles.

También se realizó, con el mismo conjunto de datos, un gráfico de líneas en el que se ve por separado el número de conflictos de cada tipo a lo largo de los años tal y como se muestra en la *Figura 4-10*. En él se puede observar una tendencia creciente en el tipo 2 (conflictos entre grupos organizados) con un gran incremento a partir del 2011, mientras que los tipos 1 y 3, pese a aumentar ligeramente, presentan un aspecto más estable.



Figura 4-10: Conflictos acaecidos por cada tipo de violencia por años (Fuente: Elaboración propia)

Se realizó también una tabla de contingencia -*Tabla 4-7*- en la que se plasma la proporción de conflictos de cada tipo para cada una de las regiones.

Tipo de violencia	África	Asia	Europe	Middle East
1	0.2094	0.4286	0.6063	0.3742
2	0.4496	0.1715	0.0669	0.3404
3	0.3411	0.3999	0.3268	0.2854

Tabla 4-7: Proporción de conflictos de cada tipo por regiones (Fuente: Elaboración propia)

Partiendo de los datos de la *Tabla 4-7*, se realizó un contraste de hipótesis para validar la independencia o dependencia de las dos variables, haciendo uso de lo comentado en 3.3.3, obteniendo un $p\text{-valor} < 0.001$, lo que nos indica que sí que hay una relación entre la región y el tipo de violencia. Por ello, se procede a realizar un gráfico por cada región (*Figura 4-11* y la *Figura 4-12*) en el que se representa la evolución de los tres tipos de violencia.

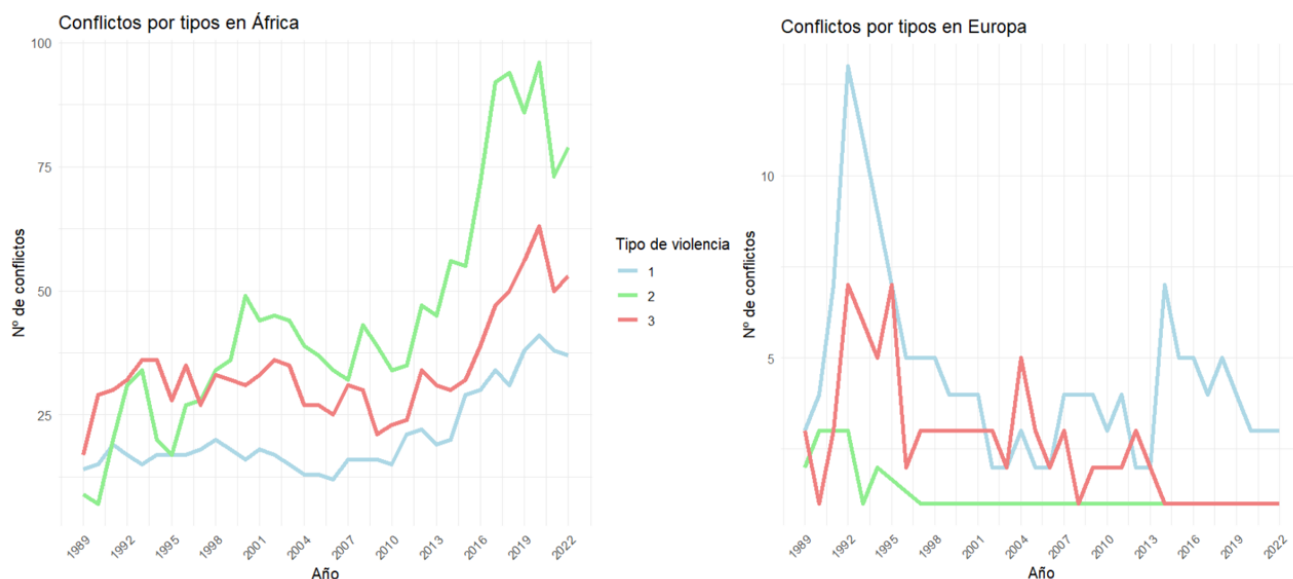


Figura 4-11: Conflictos en África y Europa por tipo de violencia (Fuente: Elaboración propia)

En la región africana se observa cómo claramente predominan los conflictos con violencia de tipo 2. Es cierto que los tres tipos de violencia tienen una tendencia al alza exponencial, sin embargo, vemos un drástico aumento del número de conflictos de tipo 2 entre los años 2015 y 2017. La región europea, sin embargo, presenta un pico acentuado de conflictos de tipo 1 durante los primeros años de la década de 1990, siendo este tipo de violencia el que prevalece la mayoría de los años. Durante todo el periodo de estudio, los conflictos tipo 2 en esta región suponen, con diferencia, la menor parte de los conflictos. Pese a todo ello, Europa es una región ciertamente estable, ya que, tal y como se observa en el gráfico, los valores son bajos en comparación con las demás regiones.

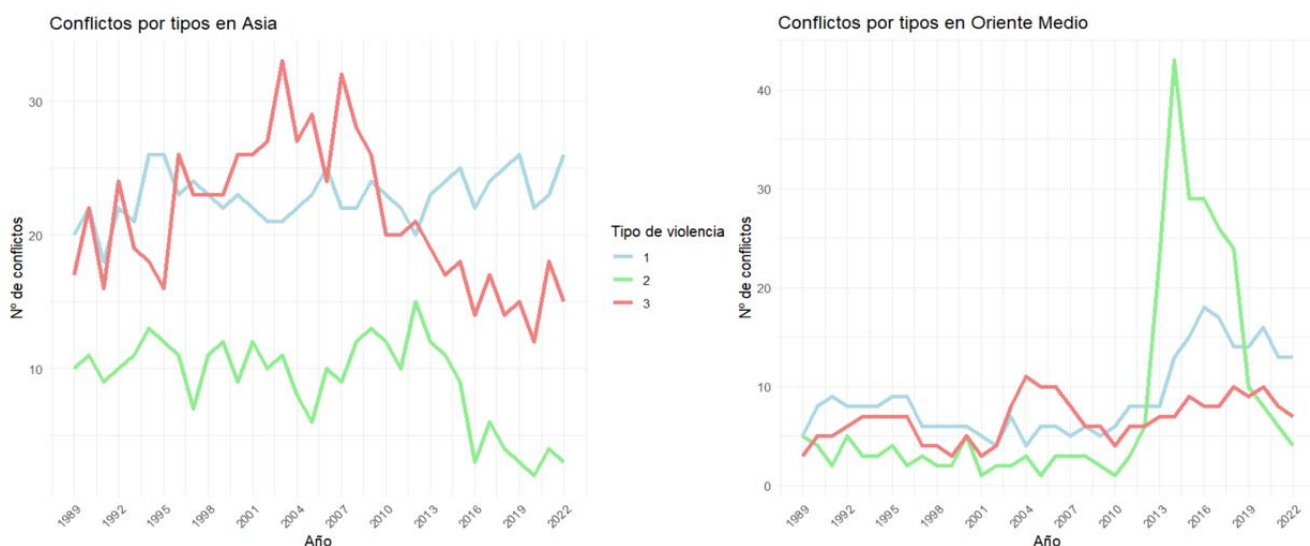


Figura 4-12: Conflictos en Asia y Oriente Medio por tipo de violencia (Fuente: Elaboración propia)

En la región asiática se observa una situación contraria a la de la región africana, ya que los conflictos tipo 2 suponen una minoría durante todo el periodo de estudio. Los conflictos de tipo 1 y 3 se mantienen medianamente estables a lo largo del tiempo con la excepción de la primera década del S. XXI, donde observamos un pico de los conflictos de tipo 3. En Oriente Medio podemos observar cierta

Una vez concluido un análisis descriptivo general del conjunto de datos con el que trabajaremos, se procedió al estudio de las hipótesis planteadas.

Para el análisis de esta hipótesis se han filtrado los conflictos activos en 2021 y 2022 (el año previo y el año de la guerra de Ucrania). De esta forma se han reducido los posibles conflictos relacionados con la guerra a los mostrados en el gráfico espacial de la *Figura 4-13*. En el gráfico se ha hecho distinción en función de si el conflicto estuvo activo solo en 2021, solo en 2022 o si estuvo activo tanto en 2021 como en 2022, representándolos en azul, rojo, y verde respectivamente.

Figura 4-13: Gráfico espacial de conflictos surgidos en 2021 y 2022 (Fuente: Elaboración propia)

La influencia rusa en Siria tiene como objetivos proteger su base naval en el Mediterráneo (Base naval de Tartús, actualmente la única base mediante la cual Rusia mantiene su presencia e influencia en el área mediterránea), hacer contrapeso al flanco Oriental de la OTAN, principalmente establecido en el Mar Báltico y el Mar Negro, y desviar el foco de atención occidental del conflicto de Ucrania, entre otros [32]. Es este último el que nos atañe en el estudio de la hipótesis, ya que sí que parece haber una posible relación entre conflictos surgidos en Siria en 2021 y 2022 con la Guerra de Ucrania.

Para estudiar más en detalle los conflictos surgidos en Siria en los años en estudio, se seleccionan únicamente aquellos conflictos ocurridos en Siria, de modo que el conjunto de conflictos resultantes solo contendrá aquellos que hayan tenido lugar en Siria durante los años 2021 y 2022 -ver *Tabla 4-8*-.

country	year	side_a	side_b	best
Syria	2021	Government of Syria	Syrian insurgents	674
Syria	2021	Government of Syria	Civilians	1
Syria	2021	Syrian insurgents	Civilians	5
Syria	2021	Government of Syria	IS	680
Syria	2021	IS	SDF	120
Syria	2021	SDF	Civilians	9
Syria	2021	Government of Syria	SDF	19
Syria	2021	IS	HTS	9
Syria	2021	HTS	Tanzim Hurras ad-Din	2
Syria	2021	Government of Iran	Government of Israel	162
Syria	2021	SNA	SDF	187
Syria	2022	Government of Syria	Syrian insurgents	530
Syria	2022	Government of Syria	Civilians	6
Syria	2022	Syrian insurgents	Civilians	3
Syria	2022	Government of Syria	IS	347
Syria	2022	IS	SDF	453
Syria	2022	SDF	Civilians	4
Syria	2022	Government of Syria	SDF	7
Syria	2022	Government of Iran	Government of Israel	55
Syria	2022	SNA	SDF	159

Tabla 4-8: Conflictos surgidos en Siria en los años 2021 y 2022 (Fuente: Elaboración propia)

Se puede observar en la tabla un elevado número de conflictos surgidos en Siria en el periodo de estudio, de los cuales hay una buena parte que supusieron un gran número de bajas. Se hizo el recuento de bajas por años, quedando un total de 1868 muertes en el año 2021 y 1564 muertes en 2022. Se puede concluir, que, aunque no se puede tener la certeza de algo así sin llevar a cabo un estudio geopolítico profundo, aparentemente sí que hay relación entre los conflictos en Siria durante este periodo de dos años y el conflicto ruso-ucraniano. No solo por las intenciones estratégicas de Rusia en Siria, sino también por el elevado número de bajas, pues la atención pública siempre tiende a ir donde hay un gran número de muertes.

4.4 Hipótesis 2: A medida que desaparece la participación del ISIS en conflictos en Siria, aumenta en conflictos en África.

En esta hipótesis se pretende estudiar si existe una relación entre la pérdida de poder del Estado Islámico¹¹ en Siria y su aparición en países africanos. Esta hipótesis se ha formulado ya que la tendencia es que cuando un grupo terrorista es derrotado en un territorio, con gran frecuencia se produce un traslado a otra zona sobre la que proyectar su poder y hacerse con el control de la zona[14].

¹¹ IS: Islamic State; ISIS: Islamic State of Iraq and Syria

Para ello, primero se seleccionaron aquellos conflictos en los que el IS estuviera involucrado, sacando dos subconjuntos de datos, uno de los conflictos con participación del IS localizados en Siria y otro con los localizados en África, ambos plasmados en la *Figura 4-14*.

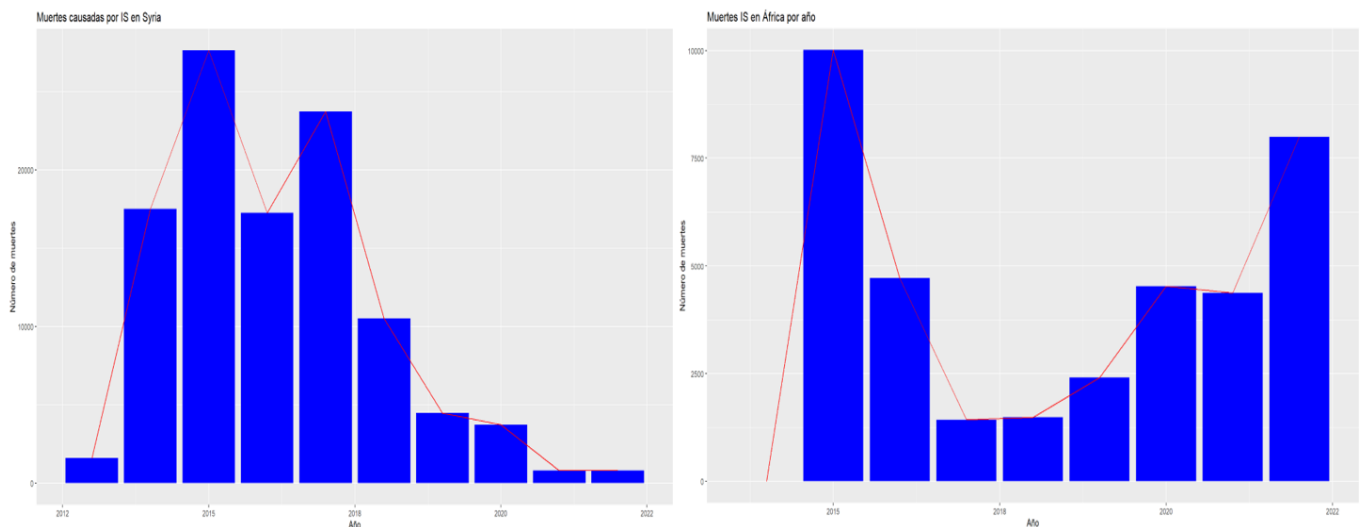


Figura 4-14: Muertes en Siria causadas por el IS (Fuente: Elaboración propia)

En Siria, podemos observar una clara tendencia decadente a partir del año 2017, coincidiendo con la paulatina expulsión del IS. Por otra parte, en África podemos observar cómo, pese a haber cierta presencia en 2015, esta no es continuada, sino que se trató de algo puntual, disminuyendo hasta 2017, donde comenzó la escalada definitiva que sí significó el traslado de influencia del IS al continente africano, donde vemos que el número de conflictos tiene una clara tendencia al alza a partir de ese año.

Los datos de ambas zonas se unificaron en un solo gráfico -Figura 4-15- donde en morado aparecen las muertes causadas por el IS en Siria, y en diferentes colores las causadas en los países africanos.

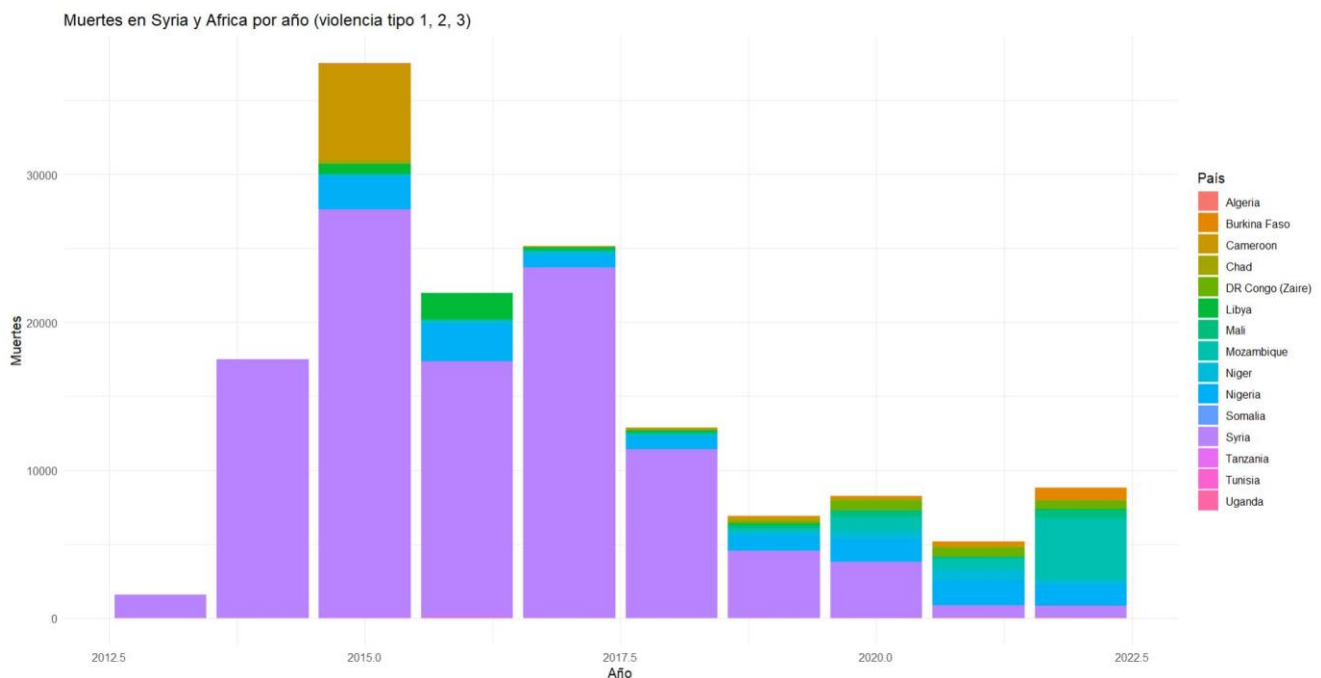


Figura 4-15: Muertes en Siria y África causadas por el IS (Fuente: Elaboración propia)

Además, se creó una tabla -Tabla 4-9- con los países africanos que fueron objeto de conflicto con participación del IS con las muertes causadas en cada uno de ellos.

	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Burkina Faso	0	0	0	0	42	19	176	1276	885	1447	3094
Cameroon	0	0	13	7383	38	233	921	800	662	829	769
Libya	378	36	1455	1481	2284	1095	664	1784	685	12	13
Mali	359	912	291	257	182	634	1077	1068	1514	1098	2104
Mozambique	2	45	33	0	68	3	173	459	1047	740	4187
Niger	1	0	15	171	274	241	179	285	655	833	488
Nigeria	2016	3383	11378	6904	3718	3749	3296	2420	2971	2802	2649
Somalia	3044	1131	1436	1691	2158	2097	2484	2126	2032	2266	3049

Tabla 4-9: Número de muertes por país y año en África de 2012 a 2022 (Fuente: Elaboración propia)

A esta tabla se le aplicó un chi-cuadrado -ver apartado 3.3.3- con la finalidad de comprobar la relación entre el año y el número de muertes. El resultado fue un p-valor < 0.001 por tanto, se reafirma que existe una relación entre el país y el intervalo temporal. De modo que se confirma que, en un mismo país, el tipo de violencia predominante varía en función del marco temporal.

Por último, se realizó un gráfico de dispersión partiendo de parejas de valores de muertes tanto en África como en Siria para cada uno de los años estudiados. Se observa cierta linealidad y relación entre las muertes en ambos territorios, tal y como se aprecia en la *Figura 4-16*, donde vemos que a medida que los conflictos en Siria toman valores más bajos, los valores de las muertes en África aumentan. Se pueden observar un pico de muertes en Camerún en el 2015, año en el que Boko Haram le juró lealtad al IS, constituyéndose este como un año sangriento que acabó con 6354 muertes civiles provocadas por el mencionado grupo terrorista [33]. Se ha estimado que se trata de un hecho concreto que no implica un asentamiento del IS en África ya que hasta el 2017 no comienza un aumento consolidado y constante de las muertes, aunque sí que es cierto que en Nigeria las muertes constantes comienzan desde el 2015, lo que indica que, pese a seguir centrado en Siria, el IS ya tenía una presencia relativamente estable en Nigeria desde 2015.

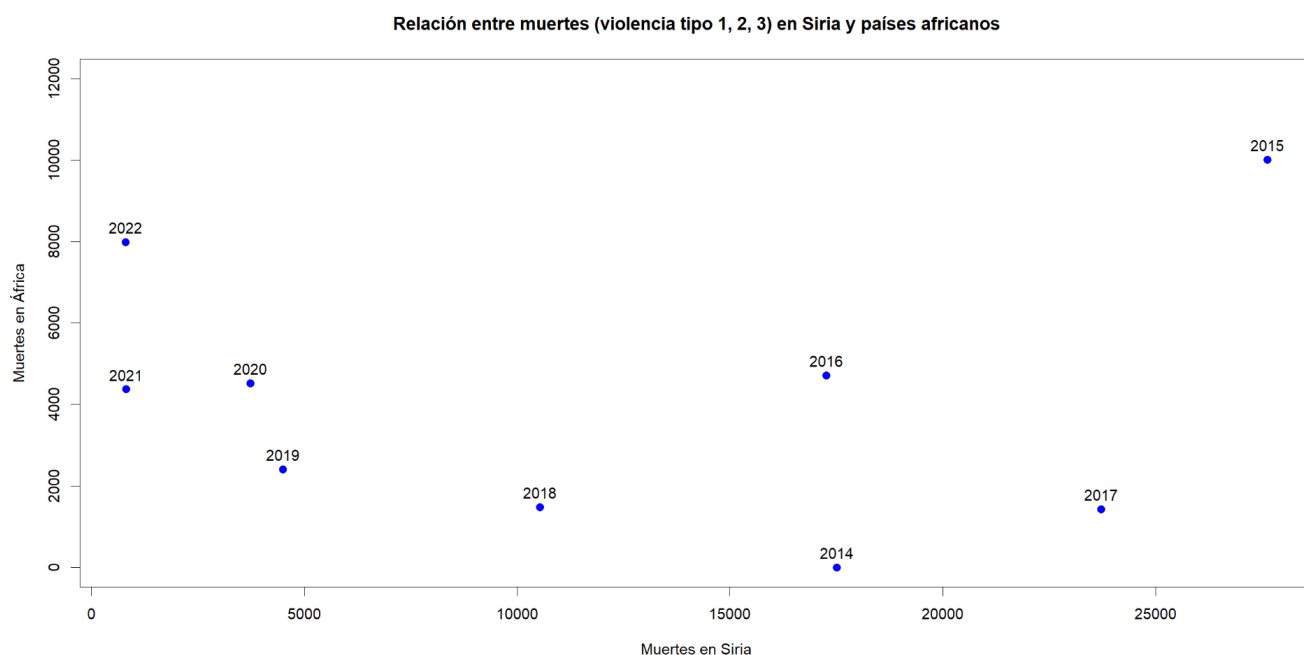


Figura 4-16: Gráfico de dispersión de muertes en Siria y África por el IS incluyendo violencia tipo 1 (Fuente: Elaboración propia)

Para comprobar de una forma más clara la linealidad de la relación entre las muertes en ambas regiones, se filtran los datos quedándonos únicamente con los años 2017-2022, periodo en el cual se da realmente la migración de influencia del IS. De esta forma, en la *Figura 4-17* sí que podemos observar una tendencia más clara y que a medida que las muertes disminuyen en Siria, aumentan en África.

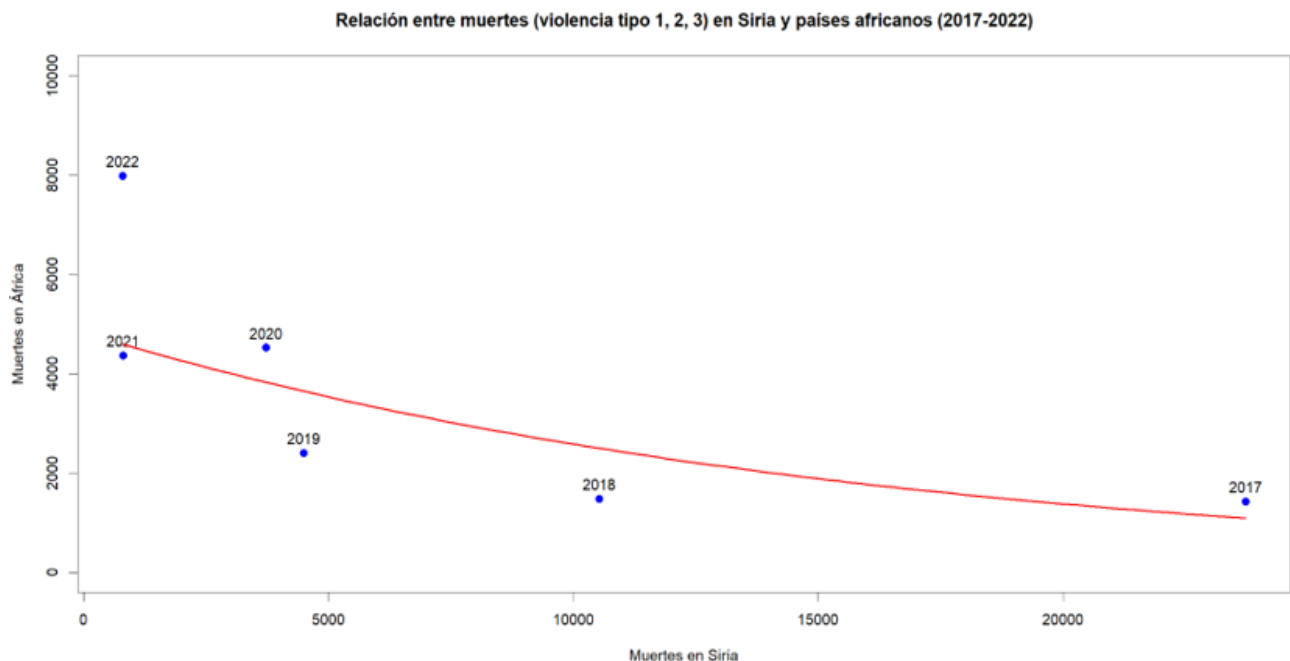


Figura 4-17: Gráfico de dispersión de muertes en Siria y África por el IS incluyendo violencia tipo 1 con línea de tendencia exponencial (2017-2022) (Fuente: Elaboración propia)

En el gráfico se ha incluido una línea de tendencia exponencial. Esta regresión – ver apartado 3.3.2- tiene un valor de $R^2=0.55$, lo que indica que el modelo es capaz de explicar el 55% de la variabilidad de la variable Y en función de X, quedando un 45% de la tendencia por explicar. Sin embargo, al realizar el contraste de hipótesis sobre si realmente el ajuste es bueno, el p-valor = 0.0561 > α , por lo que no podemos afirmar que los datos se ajusten al modelo. Bien es cierto que el p-valor es bastante cercano a α , por lo que hay evidencias de significación.

4.5 Trabajo realizado en R

Todo el análisis estadístico mostrado hasta el momento en el capítulo ha sido realizado con una base de programación en R Project, cuyo código se adjunta en el *Anexo III: Código de R Project*. Con el fin de familiarizar al lector con el proceso seguido, se han creado unos diagramas explicativos -*Figura 4-18* y *Figura 4-19*- con la secuencia seguida, así como con algunas de las funcionalidades del lenguaje de programación usadas. Por motivos de legibilidad del diagrama no se han incluido todas las librerías ni todo el detalle de lo realizado, todo ello se mostrará de forma más extensa en el ya citado Anexo III.

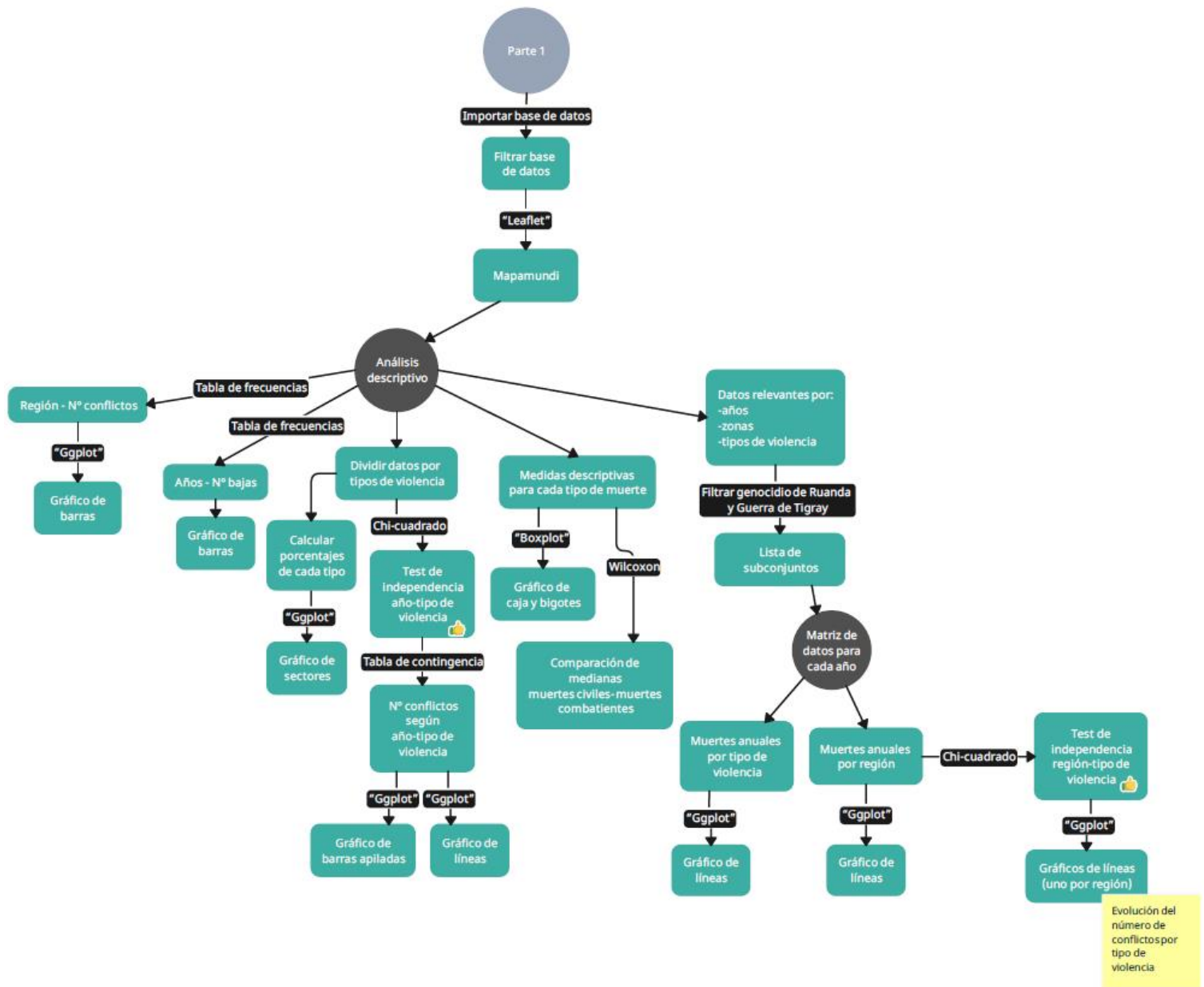


Figura 4-18: Trabajo realizado en R (parte 1)

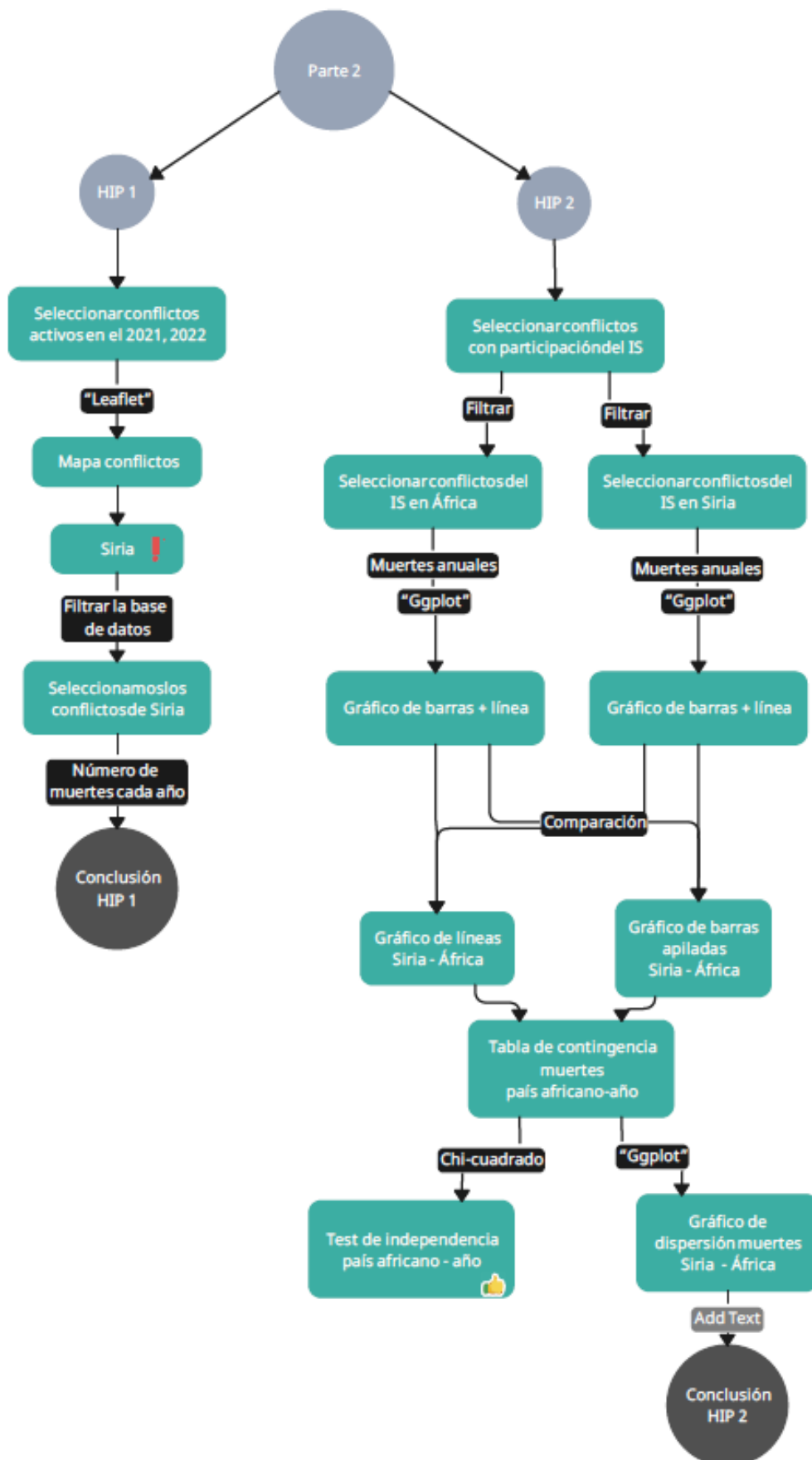


Figura 4-19: Trabajo realizado en R (parte 2)

5 CONCLUSIONES Y LÍNEAS FUTURAS

En este apartado se expondrán las conclusiones obtenidas a lo largo del trabajo realizado. Se explicará también si se alcanzaron o no los objetivos establecidos en el primer apartado de este TFG. Por último, se presentarán unas posibles líneas futuras de trabajo en las que profundizar con la intención de ampliar y mejorar este trabajo.

5.1 Conclusiones

A lo largo de este TFG se ha hecho uso de diferentes técnicas estadísticas -desde análisis descriptivo, representación de diferentes gráficos, hasta el empleo de modelos de regresión, tablas de contingencia o comparación de medias/medianas- que, empleadas en consonancia, han permitido llevar a cabo un estudio estadístico en busca de cumplir los objetivos principales de este TFG. Tanto del conjunto del estudio, como de sus subapartados, se pueden sacar las siguientes conclusiones:

- Se ha obtenido una idea general de la base de datos en su conjunto, proporcionando representaciones visuales del panorama mundial de conflictos, permitiendo al lector entrar en materia y tener un conocimiento general de la información con la que se trabaja durante el resto del estudio.
- Se ha conseguido estudiar y analizar las diferentes variables presentes en el conjunto de datos, presentando gráficos que ayuden a su visualización y realizando contrastes que permitan establecer relaciones entre variables, favoreciendo la comprensión del entorno bélico en estudio.
- En la hipótesis 1, se han estudiado los conflictos con potencial relación con la guerra entre Rusia y Ucrania, apoyándose en información básica de geopolítica para dar un sentido al proceso estadístico pudiendo llegar a conclusiones razonables.
- En la hipótesis 2, se ha hecho un estudio detallado de la influencia del Estado Islámico en diferentes regiones, y mediante el empleo de diferentes herramientas estadísticas, apoyadas por representaciones gráficas, se ha conseguido establecer una relación entre la disminución de la influencia del Estado Islámico en Siria y su aumento de presencia en los países africanos.

Teniendo en cuenta todas las conclusiones obtenidas durante las diferentes partes del trabajo, se concluye de manera global que se han cumplido los objetivos generales de este TFG: hacer uso de la estadística descriptiva para obtener una visión general de los datos, para poder así establecer hipótesis con el fin de corroborarlas -o desmentirlas- a través de procedimientos de inferencia estadística, sacando conclusiones que faciliten el entendimiento del panorama de conflictos bélicos y las relaciones existentes entre ellos.

5.2 Líneas futuras

Extraídas las conclusiones de la realización del trabajo, se proponen las siguientes líneas futuras de mejora y profundización en la materia:

- La primera línea futura es la inclusión de los conflictos sucedidos en el continente americano, pudiendo aportar datos descriptivos de la situación en América, así como una comparativa con el resto de las regiones de estudio.
- La segunda consiste en el uso de otras bases de datos, como puede ser la de ACLED, o incluso un estudio de la posibilidad de fusionar datos de ACLED y de UCDP, enriqueciéndose así mutuamente y aportando un espectro más amplio de completo de datos, ampliando el periodo de estudio hasta el 2024.
- Otra línea a considerar es el análisis de los conflictos que, siendo el mismo, surgen en diferentes momentos. En este trabajo han sido unificados, pero puede ser interesante ver si existen relaciones entre los conflictos que resurgen pasado un tiempo de su aparente resolución.
- Otro aspecto de potencial interés es el estudio de la participación de la OTAN en diferentes conflictos internacionales, buscando posibles patrones de intervención.

6 BIBLIOGRAFÍA

En esta sección figurarán todas las referencias, sean recursos web, libros, artículos, etc., incluyendo la información de autores, título de la obra, nombre de la publicación, año, edición y enlace más fecha de último acceso en el caso de referencias a recursos online.

- [1] Escola de Cultura de Pau, «Alerta 2008! Informe sobre conflictos, derechos humanos y construcción de paz», 2008.
- [2] B. Darussalam, «Convenios de Ginebra del 12 de agosto de 1949 y Protocolos adicionales del 8 de junio de 1977-Ratificaciones y adhesiones».
- [3] N. Melzer y E. Kuster, *Derecho internacional humanitario*. Suiza: Comité Internacional de la Cruz Roja (CICR), 2019.
- [4] Peace Research Institute Oslo, «PRIO-about». [En línea]. Disponible en: <https://www.prio.org/about>
- [5] N. P. Gleditsch, P. Wallensteen, M. Eriksson, M. Sollenberg, y H. Strand, «UCDP/PRIO Armed Conflict Dataset Codebook Version 21.1».
- [6] «Global Terrorism Database (GTD)». [En línea]. Disponible en: <https://www.start.umd.edu/gtd/about/>
- [7] «Socios de investigación y recopilación de datos locales». [En línea]. Disponible en: <https://acleddata.com/about-acled/our-partners/local-data-collection-partners/>
- [8] K. Unfried y K. Kis-Katos, «The heterogeneous effects of conflict on education: A spatial analysis in Sub-Saharan Africa», *J. Peace Res.*, vol. 60, n.º 6, pp. 968-984, nov. 2023, doi: 10.1177/00223433221099557.
- [9] C. Dowd, «Food-related violence, hunger and humanitarian crises», *J. Peace Res.*, vol. 60, n.º 6, pp. 935-950, nov. 2023, doi: 10.1177/00223433221099309.
- [10] M. Souva, «Material military power: A country-year measure of military power, 1865–2019», *J. Peace Res.*, vol. 60, n.º 6, pp. 1002-1009, nov. 2023, doi: 10.1177/00223433221112970.
- [11] J. A. Jackson, B. San-Akca, y Z. Maoz, «International support networks and the calculus of uprising», *J. Peace Res.*, vol. 57, n.º 5, pp. 632-647, sep. 2020, doi: 10.1177/0022343319885181.
- [12] A. Idler y K. Tkacova, «Conflict shapes in flux: Explaining spatial shift in conflict-related violence», *Int. Polit. Sci. Rev.*, vol. 45, n.º 1, pp. 31-52, ene. 2024, doi: 10.1177/01925121231177445.
- [13] E. Elfversson, «Cities and armed conflict: A systematic urban-rural coding of UCDP conflict events data», *Data Brief*, vol. 39, p. 107554, dic. 2021, doi: 10.1016/j.dib.2021.107554.
- [14] T. Pettersson y M. Öberg, «Organized violence, 1989–2019», *J. Peace Res.*, vol. 57, n.º 4, pp. 597-613, jul. 2020, doi: 10.1177/0022343320934986.

- [15] B. Welsh, «Your space or mine? Competition, control, and the spatial profile of militant violence against civilians», *J. Peace Res.*, vol. 60, n.º 4, pp. 557-572, jul. 2023, doi: 10.1177/00223433221091597.
- [16] H. Tadesse, «Modelling Conflict Dynamics: Evidence from Africa: What Do the Data Show via Spatiotemporal Global ACLED Dataset?», *Appl. Spat. Anal. Policy*, vol. 16, n.º 4, pp. 1541-1559, dic. 2023, doi: 10.1007/s12061-023-09522-1.
- [17] «¿Qué es R?», R Project. [En línea]. Disponible en: <https://www.r-project.org/about.html>
- [18] Equipo de Expertos en Ciencia y Tecnología, «Lenguaje de programación R: qué es, características e importancia en el Big Data», Ciencia y Tecnología.
- [19] Real Academia Española, «Definición de estadística», *Real Academia Española*. [En línea]. Disponible en: <https://dle.rae.es/estad%C3%ADstico?m=form>
- [20] S. B. Del Pino, «Estadística descriptiva e inferencial», *Innov. Exp. Educ.*, pp. 2-10, 2008.
- [21] V. Manzano Arrondo, «Estimación estadística», 2014.
- [22] P. Botella-Rocamora, M. Alacreu-García, y M. A. Martínez-Beneito, «Inferencia estadística (intervalos de confianza y p-valor). Comparación de dos poblaciones (test t de comparación de medias, comparación de dos proporciones, comparación de dos varianzas).», p. 26.
- [23] D. S. Moore, *Estadística aplicada básica*, 2ª. Antoni Bosch, 2008.
- [24] Universidad Complutense de Madrid, «Curso de R», 2018.
- [25] A. Santana y Carmen Nieves Hernández, «Gráficos en R: la función plot», Introducción a R.
- [26] R. Walpole, R. Myers, y S. Myers, *Probabilidad y estadística para ingeniería y ciencias*. Pearson educación, 2012.
- [27] C. C. George, «Probabilidad y estadística, aplicaciones y métodos», *McGrawHillInteramericana México SA México*, 1988.
- [28] J. Amat Rodrigo, «Test de Wilcoxon Mann Whitney como alternativa al t-test».
- [29] H. Orozco, «El Genocidio de Ruanda, una introducción al conflicto», *Rev. Huellas Hist. Número 3 Año*, vol. 1, 2008.
- [30] S. Molfino, «Claves para entender el conflicto de Tigray», *Cuad. Política Exter. Argent.*, n.º 134, pp. 93-101, 2021.
- [31] C. A. Patiño, *Guerra en Ucrania*. DEBATE, 2022.
- [32] M. Milosevich-Juaristi, «La finalidad estratégica de Rusia en Siria y las perspectivas de cumplimiento del acuerdo de Astaná», *Real Instituto Elcano*, may 2017.
- [33] Carlos Echeverría, «Consecuencias de la enésima escisión de Boko Haram para el futuro del grupo terrorista», *Grup. Militantes Ideol. Radic. Carácter Violento Área África Subsahariana*, mar. 2017.

ANEXO I: IMPLICACIONES SOCIALES, Y/O ECONÓMICAS, Y/O AMBIENTALES

Este trabajo tiene una aplicación directa en cuestiones sociales, económicas y ambientales, siendo el nexo entre todas ellas la posibilidad de emplear la inferencia estadística sobre los datos recopilados para prever futuros conflictos y poder mitigar sus consecuencias.

En el ámbito social, ser capaces de identificar patrones de los enfrentamientos, así como los factores que contribuyen a una escalada de la violencia, facilitan una concienciación de la ciudadanía al respecto. El análisis de los sectores de la población más afectados durante los conflictos también permite proporcionar ayuda humanitaria a aquellos grupos sociales más perjudicados. El hecho de poner en conocimiento público las estadísticas de las devastadoras consecuencias de la guerra, también ayuda a crear una conciencia social al respecto, que favorezca la erradicación de las actividades violentas.

Los conflictos armados también tienen repercusiones en el ámbito económico. Ser capaces de estudiar y predecir casos futuros, nos permite calcular cuales serían las pérdidas aproximadas en términos de infraestructuras, gastos militares, y atención médica entre otros, permitiendo crear planes de contingencia acordes a las situaciones previstas que permitan a las naciones recuperarse de una forma mas rápida de los momentos de crisis que habitualmente siguen a las guerras.

El medioambiente es también un aspecto a tener en cuenta, ya que, durante los eventos bélicos, las explosiones, y el vertido de sustancias tóxicas provocan el deterioro de numerosos hábitats naturales. La salud humana a largo plazo también puede verse comprometida en los casos de empleo de municiones nucleares, tal y como se ha comprobado a lo largo de la historia.

Por todo ello, queda patente la importancia de un correcto estudio y análisis de los enfrentamientos bélicos para su aplicación en los diferentes ámbitos de la sociedad, mejorando la calidad de vida de los ciudadanos, y promoviendo el cuidado del medioambiente.

ANEXO II: REFLEXIONES ÉTICAS Y SOCIALES

El estudio e investigación acerca de los conflictos armados ha de ser llevada a cabo con una gran sensibilidad social y ética, ya que la información que se maneja, en muchas ocasiones se obtiene mediante procesos de carácter íntimo, como puede ser el recuento de víctimas, o la identificación de cadáveres, que son a su vez esenciales para un correcto estudio de los hechos.

Se ha de actuar en todo momento con un profundo respeto hacia las víctimas, así como hacia sus círculos cercanos. Es, por consiguiente, de gran importancia el prestar especial atención al manejo y difusión de información sensible, especialmente en la sociedad actual, donde la tecnología ha avanzado a pasos agigantados y no solo la obtención de imágenes de zonas de conflicto que puedan ser de contenido sensible, sino también su difusión está al alcance de todo el mundo.

Por otro lado, los investigadores han de ser conscientes de su deber moral de tratar la información de forma transparente a la par que objetiva, buscando siempre la imparcialidad y evitando sesgos políticos e ideológicos que puedan distorsionar los resultados. Tienen además la responsabilidad de utilizar sus hallazgos y estudios para la búsqueda de soluciones pacíficas para los conflictos, así como para destapar posibles crímenes de guerra, contribuyendo a la justicia de las personas afectadas.

La ética, por tanto, debe desempeñar un papel fundamental en las investigaciones y estudios acerca de los enfrentamientos bélicos, contribuyendo a la justicia social y el respeto hacia las víctimas.

ANEXO III: CÓDIGO DE R PROJECT

#Instalamos las librerías y paquetes necesarios

```
install.packages("tidyr")
install.packages("leaflet")
install.packages("ggrepel")
install.packages("reshape2")
install.packages("openxlsx")
install.packages("dplyr")
```

```
library(readxl)
library(dplyr)
library(openxlsx)
library(ggplot2)
library(leaflet)
library(ggrepel)
library(reshape2)
library(magrittr)
```

PARTE 1

*#Realizamos las operaciones de unificar conflictos por años para la
#base de datos UCDP_seleccion_123 que incluye los tres tipos de violencia*

```
UCDP_seleccion2 <- read_excel("UCDP_seleccion2.xlsx")
View(UCDP_seleccion2)
```

*# Vamos a fusionar las filas que pertenezcan a un mismo conflicto, pero por años,
de modo que no se pierda información.*

Dividimos la base de datos en subconjuntos por año y aplicamos las condiciones de fusión

```
lista_subconjuntos_2 <- lapply(split(UCDP_seleccion2, UCDP_seleccion2$year),
function(subconjunto) {
```

Fusionar las filas correspondientes a un mismo conflicto

```
grupo_fusionado <- subconjunto %>%
  group_by(conflict_id) %>%
  summarise(
```

```

latitude = mean(latitude),
longitude = mean (longitude),
date_start = min(date_start),
date_end = max(date_end),
deaths_a = sum(deaths_a),
deaths_b = sum(deaths_b),
deaths_civilians_unknown = sum(deaths_civilians_unknown),
best = sum(best),
# Mantener las columnas comunes sin modificaciones
country = first(country),
region = first(region),
type_of_violence = first(type_of_violence),
conflict_name = first(conflict_name),
side_a = first(side_a),
side_b = first(side_b),
year = first(year) ) %>%
ungroup()

# Asignamos un nombre al subconjunto fusionado y lo devolvemos
nombre_grupo <- paste0("grupo_", unique(subconjunto$year))
return(nombre_grupo = grupo_fusionado))
# Verificamos los nombres de los grupos creados
print(names(lista_subconjuntos_2))
# Fusionamos todas las matrices en una sola
UCDP_seleccion2_unificada <- do.call(rbind, lista_subconjuntos_2)

# Ahora, una vez unificada la base de datos por años, vamos a representar todos
# los conflictos sobre un mapamundi

# Creamos un mapa leaflet
mapa <- leaflet(UCDP_seleccion2_unificada) %>%
# Añadimos las capas del mapa base
addTiles() %>%

```

Añadimos los marcadores de los eventos con el icono personalizado

```
addCircleMarkers(lng = ~longitude, lat = ~latitude,  
  radius = 1, # Tamaño del marcador en píxeles  
  color = "red", # Color del borde del marcador  
  fill = TRUE, # Rellenar el marcador  
  fillOpacity = 0.7, # Opacidad del relleno  
  popup = ~paste("Evento:", conflict_ident, "<br>",  
    "Tipo:", type_of_violence, "<br>", "Conflicto:", conflict_name, "<br>",  
    "Año:", year))
```

Muestra el mapa interactivo

```
mapa
```

Creamos una tabla de frecuencias con años y número de conflictos

```
tabla_frecuencias_2 <- table(UCDP_seleccion2_unificada$year)
```

Convertimos la tabla de frecuencias en un data frame

```
df_frecuencias_2 <- as.data.frame(tabla_frecuencias_2)
```

Renombramos las columnas del data frame

```
colnames(df_frecuencias_2) <- c("Año", "Número de conflictos")
```

Mostrar el data frame

```
print(df_frecuencias_2)
```

Realizamos el mismo proceso con las tablas de frecuencia de nº de conflictos

por tipos de violencia y por regiones

Creamos una tabla de frecuencias con tipos de violencia y número de conflictos

```
tabla_tipo_violencia <- table(UCDP_seleccion2_unificada$type_of_violence)
```

Convertimos la tabla de frecuencias en un data frame

```
df_tipo_violencia <- as.data.frame(tabla_tipo_violencia)
```

Renombramos las columnas del data frame

```
colnames(df_tipo_violencia) <- c("Tipo de violencia", "Número de conflictos")
```

Mostrar el data frame

```
print(df_tipo_violencia)
```

Creamos una tabla de frecuencias con regiones y número de conflictos

```
tabla_region <- table(UCDP_seleccion2_unificada$region)
```

Convertimos la tabla de frecuencias en un data frame

```
df_region <- as.data.frame(tabla_region)
```

Renombramos las columnas del data frame

```
colnames(df_region) <- c("Región", "Número de conflictos")
```

Mostrar el data frame

```
print(df_region)
```

Creamos un gráfico de barras con ggplot2

```
conflictos_por_años <- ggplot(df_frecuencias_2, aes(x = Año, y = `Número de conflictos`)) +  
  geom_bar(stat = "identity", fill = "skyblue") +
```

```
  labs(title = "Número de conflictos por año", x = "Año", y = "Número de conflictos") +
```

```
  theme_minimal() + # Aplica un tema minimalista al gráfico
```

```
  theme(axis.text.x = element_text(angle = 45, hjust = 1, vjust = 1)) # Rotación de etiquetas eje X
```

Mostrar el gráfico

```
print(conflictos_por_años)
```

Creamos una tabla de frecuencias con años y número de bajas

Calculamos la suma de las bajas por año

```
suma_bajas_2 <- tapply(UCDP_seleccion2_unificada$best, UCDP_seleccion2_unificada$year,  
  sum)
```

Creamos la tabla de frecuencias con los años y la suma de bajas

```
tabla_frecuencias_2 <- data.frame(year = as.integer(names(suma_bajas_2)), bajas =  
  as.integer(suma_bajas_2))
```

Mostrar la tabla de frecuencias

```
print(tabla_frecuencias_2)
```

Crea el gráfico de barras

```
grafico_barras_2 <- ggplot(tabla_frecuencias_2, aes(x = factor(year), y = bajas)) +
```

```
  geom_bar(stat = "identity", fill = "skyblue") + # Crea las barras con el recuento de bajas
```

```
  labs(x = "Años", y = "Número de bajas", title = "Recuento de bajas por año") + # Etiqueta los ejes  
  # y el título
```

```
  theme_minimal() + # Aplica un tema minimalista al gráfico
```

```
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotación de etiquetas en el eje X
```

Muestra el gráfico

```
print(grafico_barras_2)
```

Creamos un diagrama de sectores en el que veamos la proporción de bajas causadas por cada tipo de violencia

Calcular los porcentajes

Filtrar los datos para cada tipo de violencia

```
datos_tipo_1 <- filter(UCDP_seleccion2_unificada, type_of_violence == 1)
```

```
datos_tipo_2 <- filter(UCDP_seleccion2_unificada, type_of_violence == 2)
```

```
datos_tipo_3 <- filter(UCDP_seleccion2_unificada, type_of_violence == 3)
```

Calcular los porcentajes para cada tipo de violencia

```
porcentaje_tipo_1 <- (sum(datos_tipo_1$best) / sum(UCDP_seleccion2_unificada$best)) * 100
```

```
porcentaje_tipo_2 <- (sum(datos_tipo_2$best) / sum(UCDP_seleccion2_unificada$best)) * 100
```

```
porcentaje_tipo_3 <- (sum(datos_tipo_3$best) / sum(UCDP_seleccion2_unificada$best)) * 100
```

Crear un nuevo dataframe para los porcentajes

```
df_porcentajes_2 <- data.frame(
```

```
  type_of_violence = factor(c("Tipo 1", "Tipo 2", "Tipo 3")),
```

```
  percent = c(porcentaje_tipo_1, porcentaje_tipo_2, porcentaje_tipo_3))
```

Crear el gráfico de sectores con porcentajes

```
ggplot(df_porcentajes_2, aes(x = "", y = percent, fill = type_of_violence, label =  
paste0(round(percent, 1), "%")))
```

```
  geom_bar(width = 1, stat = "identity") +
```

```
  geom_text(position = position_stack(vjust = 0.5), size = 3) + # Agregar etiquetas de porcentaje
```

```
  coord_polar("y", start = 0) + # Convertir en gráfico de sectores
```

```
  labs(fill = "Tipo de violencia", title = "Porcentaje de bajas por tipo de violencia") +
```

```
  theme_void() # Quitar ejes y etiquetas
```

#Ahora vamos a analizar la frecuencia con la que aparece cada tipo de violencia

Calculamos la tabla de frecuencias

```
tabla_frecuencias_tipo_violencia_2<-table(UCDP_seleccion2_unificada$year,  
UCDP_seleccion2_unificada$type_of_violence)
```

#Mostramos la tabla con países en las filas y tipo de violencia en las columnas

```
print(tabla_frecuencias_tipo_violencia_2)
```

#Hacemos chi-cuadrado para esta tabla de frecuencias

```
chisq.test(tabla_frecuencias_tipo_violencia_2)
```

#Creamos un gráfico de barras apiladas para ver, por años, qué proporción de
#los conflictos ocurridos son de cada tipo

Convertir las columnas "Año" y "Tipo de violencia" a factores

```
tabla_df_tipo_violencia_2$Año <- as.factor(tabla_df_tipo_violencia_2$Año)
```

```
tabla_df_tipo_violencia_2$`Tipo de violencia` <- as.factor(tabla_df_tipo_violencia_2$`Tipo de violencia`)
```

Crear un vector de colores para los tres tipos de violencia

```
colores <- c("lightblue", "lightgreen", "lightcoral")
```

Crear el gráfico de barras apiladas

```
ggplot(tabla_df_tipo_violencia_2, aes(x = Año, y = Freq, fill = `Tipo de violencia`)) +
```

```
  geom_bar(stat = "identity") +
```

```
  labs(x = "Año", y = "Frecuencia", fill = "Tipo de violencia") +
```

```
  scale_fill_manual(values = colores) + # Asignar colores
```

```
  theme_minimal() +
```

```
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotación de etiquetas en el eje X
```

Crear el gráfico de líneas

```
ggplot(tabla_df_tipo_violencia_2, aes(x = Año, y = Freq, color = `Tipo de violencia`, group = `Tipo de violencia`)) +
```

```
  geom_line(size = 1.5) +
```

```
  labs(x = "Año", y = "Nº de conflictos", color = "Tipo de violencia") +
```

```
  scale_color_manual(values = colores) + # Asignar colores
```

```
  theme_minimal() +
```

```
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotación de etiquetas en el eje X
```

Ahora calculamos las medidas descriptivas de las variables cuantitativas

"deaths_a", "deaths_b" y "deaths_civilians_unknown"

```
medias_descriptivas_2 <- data.frame(
```

```
  deaths_a = c(mean(UCDP_seleccion2_unificada$deaths_a),
    sd(UCDP_seleccion2_unificada$deaths_a), min(UCDP_seleccion2_unificada$deaths_a),
    max(UCDP_seleccion2_unificada$deaths_a), quantile(UCDP_seleccion2_unificada$deaths_a,
    probs = c(0.25, 0.5, 0.75)), median(UCDP_seleccion2_unificada$deaths_a)),
```

```
  deaths_b = c(mean(UCDP_seleccion2_unificada$deaths_b),
    sd(UCDP_seleccion2_unificada$deaths_b), min(UCDP_seleccion2_unificada$deaths_b),
    max(UCDP_seleccion2_unificada$deaths_b), quantile(UCDP_seleccion2_unificada$deaths_b,
    probs = c(0.25, 0.5, 0.75)), median(UCDP_seleccion2_unificada$deaths_b)),
```

```
  deaths_civilians_unknown = c(mean(UCDP_seleccion2_unificada$deaths_civilians_unknown),
    sd(UCDP_seleccion2_unificada$deaths_civilians_unknown),
```

```
min(UCDP_seleccion2_unificada$deaths_civilians_unknown),
max(UCDP_seleccion2_unificada$deaths_civilians_unknown),
quantile(UCDP_seleccion2_unificada$deaths_civilians_unknown, probs = c(0.25, 0.5, 0.75)),
median(UCDP_seleccion2_unificada$deaths_civilians_unknown)))
```

Asignar nombres a las filas

```
rownames(medias_descriptivas_2) <- c("Media", "Desviación estándar", "Mínimo", "Máximo",
"Cuartil 1", "Cuartil 2", "Cuartil 3", "Mediana")
```

Transponer el marco de datos

```
medias_descriptivas_2 <- t(medias_descriptivas_2)
```

Mostrar la tabla de medidas descriptivas

```
medias_descriptivas_2
```

Unificamos los datos las muertes de combatientes como la suma de deaths_a y deaths_b

```
combatientes= apply(UCDP_seleccion2_unificada[,6:7],1,sum)
```

```
civiles = UCDP_seleccion2_unificada$deaths_civilians_unknown
```

Vemos las medidas estadísticas de las muertes en ambos colectivos

```
summary(combatientes)
```

```
summary(civiles)
```

Las representamos en un boxplot cada una

```
boxplot(combatientes[combatientes<100], main = "Muertes de combatientes")
```

```
boxplot(civiles[civiles<100], main = "Muertes de civiles")
```

Realizamos un Wilcoxon test para corroborar que mueren más civiles que combatientes

H0: mediana_combatientes >= mediana_civiles

H1: mediana_combatientes < mediana_civiles

```
wilcox.test(combatientes,civiles,alternative="less",paired=TRUE)
```

Ahora vamos a calcular las medidas descriptivas de las variables cuantitativas

"deaths_a", "deaths_b" y "deaths_civilians_unknown" por años

Creamos una lista para almacenar los resúmenes de medidas descriptivas

```
resumenes_2 <- list()
```

Calculamos medidas descriptivas para cada año en la lista de subconjuntos

```
for (i in 1:length(lista_subconjuntos_2)) {
```

```
  resumen <- summary(lista_subconjuntos_2[[i]][c("deaths_a", "deaths_b",
"deaths_civilians_unknown")])
```

```
  resumenes_2[[i]] <- resumen}
```

Mostrar los resúmenes de medidas descriptivas para cada año

```
for (i in 1:length(resumenes_2)) {
  cat("Resumen para el año", names(lista_subconjuntos_2)[i], ":\n")
  print(resumenes_2[[i]])
  cat("\n")}
```

#Vamos a obtener representaciones gráficas de los datos relevantes estudiándolo

#por años, por zonas y por tipos de violencia

#Primero creamos matrices (una por año) con el subconjunto de datos únicamente

#de ese año

Iterar sobre cada año y seleccionar los datos de ese año

```
for (año in 1989:2022) {
```

Encontrar el índice de la matriz en la lista

```
indice_matriz <- which(names(lista_subconjuntos_2) == as.character(año))
```

Seleccionar la matriz por su nombre

```
datos_año <- lista_subconjuntos_2[[indice_matriz]]
```

Asignar la matriz seleccionada a una variable con un nombre único del año correspondiente

```
assign(paste0("datos_", año), datos_año)}
```

Ahora tenemos variables con nombres como datos_1989, datos_1990, ..., datos_2022,

cada una conteniendo los datos correspondientes a ese año.

Creamos un data frame vacío para almacenar los totales de muertes

```
totales_muertes_2 <- data.frame(Año = 1989:2022,
                                Deaths_A = numeric(length = 34),
                                Deaths_B = numeric(length = 34),
                                Deaths_Civilians_Unknown = numeric(length = 34))
```

Calculamos los totales de muertes para cada año

```
for (i in 1:34) {
  año <- 1989 + i - 1
  matriz_año <- lista_subconjuntos_2[[i]]
  totales_muertes_2[i, "Deaths_A"] <- sum(matriz_año[, "deaths_a"])
  totales_muertes_2[i, "Deaths_B"] <- sum(matriz_año[, "deaths_b"])
  totales_muertes_2[i, "Deaths_Civilians_Unknown"] <-
  sum(matriz_año[, "deaths_civilians_unknown"])}
}
```

Convertimos los datos a formato largo (tidy)

```
totales_muertes_long_2 <- tidyr::gather(totales_muertes_2, "Categoria", "Muertes", -Año)
```

```
# Creamos el gráfico de líneas
```

```
ggplot(totales_muertes_long_2, aes(x = Año, y = Muertes, color = Categoria)) +  
  geom_line() +  
  labs(title = "Evolución temporal de muertes por categoría",  
        x = "Año",  
        y = "Muertes",  
        color = "Categoría") +  
  theme_minimal()
```

```
# Como hay un pico en "deaths_civilians_unknown" en 1994 y en 2022, no se ven
```

```
# bien las líneas con valores más bajos, así que, para un mejor análisis,
```

```
# vamos a eliminar el genocidio de Ruanda de 1994 y la Guerra de Tigray de
```

```
# Etiopía en 2022 para tener un mejor detalle de los demás años
```

```
# Filtramos la base de datos UCDP_seleccion_2 para quitarle los dos conflictos
```

```
UCDP_seleccion2_unificada_sin_Ruanda1994 <-
```

```
UCDP_seleccion2_unificada[!(UCDP_seleccion2_unificada$conflict_id == 472 | UCDP_seleccion2_unificada$conflict_id == 267), ]
```

```
# Ahora vamos a calcular las medias descriptivas de las variables cuantitativas
```

```
# "deaths_a", "deaths_b" y "deaths_civilians_unknown" por años excluyendo
```

```
# el genocidio de Ruanda en 1994 y la Guerra de Tigray de Etiopía en 2022
```

```
# Creamos una lista para almacenar los resúmenes de medidas descriptivas
```

```
resumenes_sin_Ruanda1994 <- list()
```

```
# Eliminamos el genocidio de Ruanda y la Guerra de Tigray de la lista de subconjuntos
```

```
lista_subconjuntos_sin_Ruanda1994 <- lapply(lista_subconjuntos_2, function(subconjunto) {  
  subconjunto <- subconjunto %>%  
    filter(conflict_id != 472, conflict_id != 267)  
  return(subconjunto)})
```

```
# Vamos a obtener representaciones gráficas de los datos relevantes estudiándolo
```

```
# por años, por zonas y por tipos de violencia, una vez descartados los datos atípicos
```

```
# Primero creamos matrices (una por año) con el subconjunto de datos únicamente de ese año
```

```

# Iterar sobre cada año y seleccionar los datos de ese año
for (año in 1989:2022) {
# Encontrar el índice de la matriz en la lista
indice_matriz <- which(names(lista_subconjuntos_sin_Ruanda1994) == as.character(año))
# Seleccionar la matriz por su nombre
datos_año <- lista_subconjuntos_sin_Ruanda1994[[indice_matriz]]
# Asignar la matriz seleccionada a una variable con un nombre único
assign(paste0("datos_", año), datos_año)}

# Ahora tenemos variables con nombres como datos_1989, datos_1990, ..., datos_2022, cada una
# conteniendo los datos correspondientes a ese año.

# Creamos un data frame vacío para almacenar los totales de muertes
totales_muertes_sin_Ruanda1994 <- data.frame(Año = 1989:2022,
      Deaths_A = numeric(length = 34),
      Deaths_B = numeric(length = 34),
      Deaths_Civilians_Unknown = numeric(length = 34))

# Calculamos los totales de muertes para cada año
for (i in 1:34) {
  año <- 1989 + i - 1
  matriz_año <- lista_subconjuntos_sin_Ruanda1994[[i]]
  totales_muertes_sin_Ruanda1994[i, "Deaths_A"] <- sum(matriz_año[, "deaths_a"])
  totales_muertes_sin_Ruanda1994[i, "Deaths_B"] <- sum(matriz_año[, "deaths_b"])
  totales_muertes_sin_Ruanda1994[i, "Deaths_Civilians_Unknown"] <-
sum(matriz_año[, "deaths_civilians_unknown"])
}

# Convertimos los datos a formato largo (tidy)
totales_muertes_long_sin_Ruanda1994 <- tidyr::gather(totales_muertes_sin_Ruanda1994,
"Categoría", "Muertes", -Año)

# Creamos el gráfico de líneas
ggplot(totales_muertes_long_sin_Ruanda1994, aes(x = Año, y = Muertes, color = Categoría)) +
  geom_line() +
  labs(title = "Evolución temporal de muertes por categoría", x = "Año", y = "Muertes",
color = "Categoría") +
  theme_minimal()

# Ahora vamos a analizarlo por regiones en vez de por tipos de violencia

# Lista de todas las regiones

```

```
todas_las_regiones <- unique(unlist(lapply(lista_subconjuntos_sin_Ruanda1994, function(matriz)
unique(matriz[, "region"]))))
```

```
# Crear un data frame vacío para almacenar los totales de muertes por región
```

```
totales_muertes_region_sin_Ruanda1994 <- data.frame(Año = integer(), Muertes = numeric(),
Region = character())
```

```
# Calcular los totales de muertes para cada año por región
```

```
for (i in 1:length(lista_subconjuntos_sin_Ruanda1994)) {
  matriz_año <- lista_subconjuntos_sin_Ruanda1994[[i]]
  año <- as.numeric(names(lista_subconjuntos_sin_Ruanda1994)[i])
  for (region in todas_las_regiones) {
    muertes_region <- sum(matriz_año[matriz_año$region == region, c("deaths_a", "deaths_b",
"deaths_civilians_unknown")])

    totales_muertes_region_sin_Ruanda1994 <- rbind(totales_muertes_region_sin_Ruanda1994,
data.frame(Año = año, Muertes = ifelse(is.na(muertes_region), 0, sum(muertes_region)), Region =
region)))} }
```

```
# Crear el gráfico de líneas
```

```
ggplot(totales_muertes_region_sin_Ruanda1994, aes(x = Año, y = Muertes, color = Region, group
= Region)) +
  geom_line() +
  labs(title = "Evolución de muertes por región (excluyendo conflictos 472 y 267)",
    x = "Año",
    y = "Muertes",
    color = "Región") +
  theme_minimal()
```

```
# Vamos a comprobar si la región y el tipo de violencia están relacionados
```

```
tabla2 = table(UCDP_seleccion2_unificada$region,
UCDP_seleccion2_unificada$type_of_violence)
```

```
prop.table(tabla2, 1)
```

```
chisq.test(tabla2)
```

```
# Como del test chi-cuadrado hemos obtenido que SÍ hay relación entre la
```

```
# región y el tipo de violencia, vamos a hacer un gráfico para cada región
```

```
# en el que se vea la evolución de los diferentes tipos de violencia en cada región
```

```
# Creamos ahora 4 gráficos (uno por región) en los que se represente la evolución
```

```
# de los tipos de violencia a lo largo de los años
```

Filtrar datos por regiones

```
datos_africa <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$region == "Africa", ]
datos_asia <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$region == "Asia", ]
datos_europa <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$region == "Europe", ]
datos_orient_medio <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$region == "Middle East", ]
```

Calculamos el número de conflictos por año y tipo de violencia

```
df_africa <- as.data.frame(datos_africa)
df_asia <- as.data.frame(datos_asia)
df_europa <- as.data.frame(datos_europa)
df_orient_medio <- as.data.frame(datos_orient_medio)
```

AFRICA

Calcular el número de conflictos por año y tipo de violencia

```
df_count <- df_africa %>%
  count(year, type_of_violence)
```

Convertir las columnas "year" y "type_of_violence" a factores

```
df_count$year <- as.factor(df_count$year)
df_count$type_of_violence <- as.factor(df_count$type_of_violence)
```

Convertir la variable year a numérica

```
df_count$year <- as.numeric(as.character(df_count$year))
```

Creamos el gráfico de líneas

```
ggplot(df_count, aes(x = year, y = n, color = type_of_violence, group = type_of_violence)) +
  geom_line(size = 1.5) +
  labs(x = "Año", y = "Nº de conflictos", color = "Tipo de violencia", title = "Conflictos por tipos en África") +
  scale_color_manual(values = colores) + # Asignar colores
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Inclinar etiquetas del eje x
  scale_x_continuous(breaks = seq(min(df_count$year), max(df_count$year), by = 3)) # Mostrar años cada tres
```

EUROPA

Calcular el número de conflictos por año y tipo de violencia

```
df_count <- df_europa %>%
  count(year, type_of_violence)
```

Convertir las columnas "year" y "type_of_violence" a factores

```
df_count$year <- as.factor(df_count$year)
df_count$type_of_violence <- as.factor(df_count$type_of_violence)
# Convertir la variable year a numérica
df_count$year <- as.numeric(as.character(df_count$year))
# Creamos el gráfico de líneas
ggplot(df_count, aes(x = year, y = n, color = type_of_violence, group = type_of_violence)) +
  geom_line(size = 1.5) +
  labs(x = "Año", y = "Nº de conflictos", color = "Tipo de violencia", title = "Conflictos por tipos
en Europa") +
  scale_color_manual(values = colores) + # Asignar colores
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Inclinar etiquetas del eje x
  scale_x_continuous(breaks = seq(min(df_count$year), max(df_count$year), by = 3)) # Mostrar
años cada tres

# ASIA
# Calcular el número de conflictos por año y tipo de violencia
df_count <- df_asia %>%
  count(year, type_of_violence)
# Convertir las columnas "year" y "type_of_violence" a factores
df_count$year <- as.factor(df_count$year)
df_count$type_of_violence <- as.factor(df_count$type_of_violence)
# Convertir la variable year a numérica
df_count$year <- as.numeric(as.character(df_count$year))
# Creamos el gráfico de líneas
ggplot(df_count, aes(x = year, y = n, color = type_of_violence, group = type_of_violence)) +
  geom_line(size = 1.5) +
  labs(x = "Año", y = "Nº de conflictos", color = "Tipo de violencia", title = "Conflictos por tipos
en Asia") +
  scale_color_manual(values = colores) + # Asignar colores
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Inclinar etiquetas del eje x
  scale_x_continuous(breaks = seq(min(df_count$year), max(df_count$year), by = 3)) # Mostrar
años cada tres
```

ORIENTE MEDIO**# Calcular el número de conflictos por año y tipo de violencia**`df_count <- df_oriente_medio %>%``count(year, type_of_violence)`**# Convertir las columnas "year" y "type\_of\_violence" a factores**`df_count$year <- as.factor(df_count$year)``df_count$type_of_violence <- as.factor(df_count$type_of_violence)`**# Convertir la variable year a numérica**`df_count$year <- as.numeric(as.character(df_count$year))`**# Creamos el gráfico de líneas**`ggplot(df_count, aes(x = year, y = n, color = type_of_violence, group = type_of_violence)) +``geom_line(size = 1.5) +``labs(x = "Año", y = "Nº de conflictos", color = "Tipo de violencia", title = "Conflictos por tipos en Oriente Medio") +``scale_color_manual(values = colores) + # Asignar colores``theme_minimal() +``theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Inclinar etiquetas del eje x``scale_x_continuous(breaks = seq(min(df_count$year), max(df_count$year), by = 3)) # Mostrar años cada tres`**PARTE 2****HIPÓTESIS 1****# Vamos a ver qué conflictos tienen lugar durante 2021 y 2022 (año de la guerra de****# Ucrania) y evaluar si pueden tener relación con la guerra de Ucrania****# Paso 1: Filtrar los conflictos que ocurrieron en el año 2021-2022**`conflictos_2022_2021 <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$year == 2022 | UCDP_seleccion2_unificada$year == 2021,]`**# Veamos ahora el listado de países en los que ha habido conflictos en****# 2021-2022 para comprobar si existe alguna posible relación con Rusia****# Extraemos la columna "country" del subconjunto de datos**`países_conflictos_2021_2022 <- conflictos_2022_2021$country`**# Creamos la tabla de frecuencias para ver qué países aparecen y con qué peso**

```
tabla_frecuencias_paises_conflictos_2021_2022 <- table(paises_conflictos_2021_2022)
```

```
# Convertimos la tabla de frecuencias en un data frame
```

```
tabla_df_paises_conflictos_2021_2022 <-  
as.data.frame(tabla_frecuencias_paises_conflictos_2021_2022)
```

```
# Renombramos las columnas
```

```
colnames(tabla_df_paises_conflictos_2021_2022) <- c("País", "Frecuencia")
```

```
# Mostramos el data frame resultante
```

```
print(tabla_df_paises_conflictos_2021_2022)
```

```
# Ahora vamos a representarlos en el mapa, haciendo distinción en función de si
```

```
# el conflicto solo estuvo presente en 2021, solo en 2022, o en ambos años
```

```
# Paso 1: Filtrar los conflictos que ocurrieron solo en 2021
```

```
conflictos_2021 <- conflictos_2022_2021[conflictos_2022_2021$year == 2021, ]
```

```
# Paso 2: Filtrar los conflictos que ocurrieron solo en 2022
```

```
conflictos_2022 <- conflictos_2022_2021[conflictos_2022_2021$year == 2022, ]
```

```
# Paso 3: Filtrar los conflictos que ocurrieron en ambos años
```

```
conflictos_ambos_años <-  
conflictos_2022_2021[duplicated(conflictos_2022_2021$conflict_ident), ]
```

```
# Paso 4: Crear un mapa base
```

```
mapa <- leaflet() %>%
```

```
  addTiles()
```

```
# Paso 5: Añadir marcadores para cada conjunto de conflictos con colores diferentes
```

```
for (i in 1:nrow(conflictos_2021)) {
```

```
  marcador <- conflictos_2021[i, ]
```

```
  mapa <- mapa %>% addCircleMarkers(
```

```
    lng = marcador$longitude, lat = marcador$latitude,
```

```
    popup = paste("Ubicación:", marcador$country, "<br>",
```

```
      "Fecha:", marcador$date_start, "<br>",
```

```
      "Tipo de conflicto:", marcador$type_of_violence),
```

```
    color = "blue", # Color azul para conflictos en 2021
```

```
    radius = 1 )} # Grosor del marcador
```

```
for (i in 1:nrow(conflictos_2022)) {
```

```
  marcador <- conflictos_2022[i, ]
```

```
  mapa <- mapa %>% addCircleMarkers(
```

```

lng = marcador$longitude, lat = marcador$latitude,
popup = paste("Ubicación:", marcador$country, "<br>",
              "Fecha:", marcador$date_start, "<br>",
              "Tipo de conflicto:", marcador$type_of_violence),
color = "red", # Color rojo para conflictos en 2022
radius = 1 )} # Grosor del marcador

for (i in 1:nrow(conflictos_ambos_años)) {
  marcador <- conflictos_ambos_años[i, ]
  mapa <- mapa %>% addCircleMarkers(
    lng = marcador$longitude, lat = marcador$latitude,
    popup = paste("Ubicación:", marcador$country, "<br>",
                  "Fecha:", marcador$date_start, "<br>",
                  "Tipo de conflicto:", marcador$type_of_violence),
    color = "green", # Color verde para conflictos en ambos años
    radius = 1 )} # Grosor del marcador

mapa # Mostrar el mapa

# Observamos varios conflictos en Siria, país en el cual Rusia tiene influencia e intereses

# Filtramos y mostramos los conflictos que ocurrieron en Siria
conflictos_syria <- conflictos_2022_2021[conflictos_2022_2021$country == "Syria", c("country",
"year", "side_a", "side_b", "best")]
print(conflictos_syria)

# Lo transformamos a formato Excel
# Escribir el conjunto de datos en un archivo Excel
write.xlsx(conflictos_syria, "conflictos_syria.xlsx", rowNames = FALSE)
# Obtenemos los datos de muertes totales en cada año para estos conflictos
# Para el año 2021
total_muertes_2021 <- sum(conflictos_syria$best[conflictos_syria$year == 2021])
# Para el año 2022
total_muertes_2022 <- sum(conflictos_syria$best[conflictos_syria$year == 2022])

# Imprimir los resultados

```

```
print(total_muertes_2021)
print(total_muertes_2022)
```

PARTE 3

HIPÓTESIS 2

Vamos a ver cómo evolucionan los conflictos en los principales países

africanos a medida que desaparece la influencia del IS en Siria

Primer paso: Comprobar cuando acaba la influencia del IS en Siria

Filtrar conflictos que tengan "IS" en las columnas "side_a" o "side_b"

```
UCDP_IS_123 <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$side_a == "IS" |
UCDP_seleccion2_unificada$side_b == "IS", ]
```

Como conflictos siempre sigue habiendo por "pequeños" que sean, vamos a

representar las muertes causadas por estos conflictos (solo en Siria) en el eje

temporal a ver cuándo disminuyen

Filtramos los datos para conflictos en Siria

```
UCDP_IS_Syria_123 <- UCDP_IS_123[UCDP_IS_123$country == "Syria", ]
```

Calculamos la suma de muertes por año

```
UCDP_IS_Syria_123_años <- aggregate(best ~ year, data = UCDP_IS_Syria_123, FUN = sum)
```

Mostramos los primeros registros de la nueva base de datos

```
head(UCDP_IS_Syria_123_años)
```

Creamos un gráfico en el que veamos tanto el gráfico de líneas como el gráfico de barras

ggplot() +

```
  geom_bar(data = UCDP_IS_Syria_123, aes(x = year, y = best), stat = "identity", fill = "blue") +
```

```
  geom_line(data = UCDP_IS_Syria_123_años, aes(x = year, y = best), color = "red") +
```

```
  labs(title = "Muertes causadas por IS en Siria", x = "Año", y = "Número de muertes")
```

Podemos observar cómo el IS empieza a perder influencia en Siria a partir de

2017 de una forma clara, será ahí donde centremos nuestro foco en África

Seleccionamos aquellos conflictos en los que intervino el IS en África y realizamos el mismo

gráfico que en Siria para ver la evolución de las muertes causadas por IS en África

```
UCDP_IS_Africa_123 <- UCDP_seleccion2_unificada[UCDP_seleccion2_unificada$region ==
"Africa" &
```

```
(UCDP_seleccion2_unificada$side_a == "IS" | UCDP_seleccion2_unificada$side_b == "IS"), ]
```

Calculamos la suma de muertes por año

```
UCDP_IS_Africa_123_años <- aggregate(best ~ year, data = UCDP_IS_Africa_123, FUN = sum)
```

Mostramos los primeros registros de la nueva base de datos

```
head(UCDP_IS_Africa_123_años)
```

Creamos un gráfico en el que veamos tanto el gráfico de líneas como el gráfico de barras

```
ggplot() +
```

```
geom_bar(data = UCDP_IS_Africa_123, aes(x = year, y = best), stat = "identity", fill = "blue") +
```

```
geom_line(data = UCDP_IS_Africa_123_años, aes(x = year, y = best), color = "red") +
```

```
labs(title = "Muertes IS en África por año", x = "Año", y = "Número de muertes")
```

Ahora mostramos los datos de Siria y África superpuestos para ver la evolución de ambos a la vez

Combinar las bases de datos UCDP_IS_Syria y UCDP_IS_Africa

```
UCDP_IS_combined_123 <- rbind(UCDP_IS_Syria_123, UCDP_IS_Africa_123)
```

Creamos un gráfico de líneas

```
ggplot() +
```

```
geom_line(data = UCDP_IS_Syria_123_años, aes(x = year, y = best, color = "Siria"), linetype = "solid") +
```

```
geom_line(data = UCDP_IS_Africa_123_años, aes(x = year, y = best, color = "África"), linetype = "solid") +
```

```
labs(x = "Año", y = "Muertes", color = "Región") +
```

```
scale_x_continuous(breaks = seq(min(min(UCDP_IS_Syria_123_años$year),
min(UCDP_IS_Africa_123_años$year)), max(max(UCDP_IS_Syria_123_años$year),
max(UCDP_IS_Africa_123_años$year)), by = 1)) +
```

```
scale_y_continuous(expand = c(0, 0)) +
```

```
theme_minimal()
```

Creamos también un gráfico de barras apiladas

```
ggplot(UCDP_IS_combined_123, aes(x = year, y = best, fill = country)) +
```

```
geom_bar(stat = "identity") +
```

```
labs(title = "Muertes en Syria y Africa por año (violencia tipo 1, 2, 3)", x = "Año", y = "Muertes",
fill = "País") +
```

```
theme_minimal()
```

Ahora vamos a plasmar en una tabla de contingencia los datos de muertes de cada uno

de los países africanos para cada uno de los años de 2012 a 2022

Definimos los países y años de interés

```
países <- c("Burkina Faso", "Cameroon", "Libya", "Mali", "Mozambique", "Niger", "Nigeria",  
"Somalia")
```

```
años <- 2012:2022
```

```
# Filtramos los datos de UCDP_seleccion_unificada para los países y años especificados
```

```
datos_filtrados_123 <- UCDP_seleccion2_unificada %>%
```

```
  filter(country %in% países & year %in% años)
```

```
# Creamos la tabla de contingencia de muertes por país africano y año
```

```
tabla_muertes_123 <- xtabs(best ~ country + year, data = datos_filtrados_123)
```

```
# Mostrar la tabla
```

```
tabla_muertes_123
```

```
# Realizamos ahora un test de independencia chi-cuadrado
```

```
resultado_test_123 <- chisq.test(tabla_muertes_123)
```

```
# Mostrar el resultado del test
```

```
resultado_test_123
```

```
# Por tanto SÍ hay relación entre el país y el periodo temporal
```

```
# Ahora haremos un gráfico de dispersión para ver de forma visual si hay
```

```
# alguna relación entre las muertes causadas por el IS en Siria y los
```

```
# países africanos, quedándonos con la suma de muertes en cada region para cada año
```

```
# Agrupamos y sumamos los datos de muertes por año para Siria
```

```
UCDP_IS_Syria_agregado_123 <- UCDP_IS_Syria_123 %>%
```

```
  group_by(year) %>%
```

```
  summarise(muertes_Syria_123 = sum(best))
```

```
# Agrupamos y sumamos los datos de muertes por año para África
```

```
UCDP_IS_Africa_agregado_123 <- UCDP_IS_Africa_123 %>%
```

```
  group_by(year) %>%
```

```
  summarise(muertes_Africa_123 = sum(best))
```

```
# Unimos los datos en un solo dataframe
```

```
datos_agregados_123 <- merge(UCDP_IS_Syria_agregado_123,  
UCDP_IS_Africa_agregado_123, by = "year", all = TRUE)
```

Creamos el gráfico de dispersión

```
plot(datos_agregados_123$muertes_Syria_123, datos_agregados_123$muertes_Africa_123,
     xlab = "Muertes en Siria", ylab = "Muertes en África",
     main = "Relación entre muertes (violencia tipo 1, 2, 3) en Siria y países africanos",
     col = "blue", pch = 20, cex = 1.75, ylim = c(0, 12000))
text(datos_agregados_123$muertes_Syria_123, datos_agregados_123$muertes_Africa_123,
     labels = datos_agregados_123$year, pos = 3)
```

Ahora nos centraremos en el periodo de 2017-2022 que es cuando el IS comienza

a perder influencia en Siria

Filtramos los datos para incluir solo los años 2017-2022

```
datos_filtrados_17_22 <- datos_agregados_123[datos_agregados_123$year >= 2017 &
datos_agregados_123$year <= 2022, ]
```

Creamos el gráfico de dispersión con los datos filtrados

```
plot(datos_filtrados_17_22$muertes_Syria_123, datos_filtrados_17_22$muertes_Africa_123,
     xlab = "Muertes en Siria", ylab = "Muertes en África",
     main = "Relación entre muertes (violencia tipo 1, 2, 3) en Siria y países africanos (2017-
2022)", col = "blue", pch = 20, cex = 1.75, ylim = c(0, 10000))
text(datos_filtrados_17_22$muertes_Syria_123, datos_filtrados_17_22$muertes_Africa_123,
     labels = datos_filtrados_17_22$year, pos = 3)
```

Ajustar un modelo exponencial

```
modelo_exponencial <- lm(log(muertes_Africa_123) ~ muertes_Syria_123, data =
datos_filtrados_17_22)
```

Generar valores de x para la línea de tendencia

```
x_valores <- seq(min(datos_filtrados_17_22$muertes_Syria_123),
max(datos_filtrados_17_22$muertes_Syria_123), length.out = 100)
```

Predecir valores de y a partir del modelo exponencial

```
y_valores <- predict(modelo_exponencial, newdata = data.frame(muertes_Syria_123 =
x_valores))
```

Transformar los valores de y de nuevo a la escala original

```
y_valores_exp <- exp(y_valores)
```

Agregar la línea de tendencia exponencial al gráfico

```
lines(x_valores, y_valores_exp, col = "red", lwd = 2)
```

Vemos la información sobre la aproximación de la línea de tendencia empleada

```
summary(modelo_exponencial)
```

PARTE 4*MAPA ANIMADO DE LOS CONFLICTOS*

Instalar y cargar paquetes necesarios

```
install.packages("ggplot2")
install.packages("gganimate")
install.packages("maps")
install.packages("tidyverse")
install.packages("ggrepel")
install.packages("ggtheme")
```

```
library(ggplot2)
library(gganimate)
library(maps)
library(ggrepel)
library(ggthemes)
library(tidyverse)
```

Obtener datos del mapa del mundo

```
mapa_mundo <- map_data("world")
mapa_gd <- mapa_mundo %>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group), fill = "grey20", color = "white") +
  geom_point(data = UCDP_seleccion2_unificada, aes(x = longitude, y = latitude), color =
"orange", alpha = 0.7, size = 0.2) +
  theme_minimal() +
  coord_fixed(xlim=c(-20,80), ylim=c(-15,60),ratio=1.3)
```

Generar la animación del mapa

```
mapa_animado <- mapa_gd+
  transition_states(UCDP_seleccion2_unificada$year, transition_length = 1, state_length = 1) +
  shadow_mark(past = TRUE) +
  labs(title = "Conflictos ocurridos desde 1989 hasta 2022 \n Year:
{UCDP_seleccion2_unificada$year}") +
  theme_map()

mapa_animado
```

```
anim_save("mapa_animado.gif",animation=last_animation())
```