

Trabajo Práctico I

Calidad de Datos

2C 2024

Dante Culaciati - 351/22 - mellidante@gmail.com
Ivo Lamonica - 66/22 - ivolamonicam@gmail.com
Nicolás Rozenberg - 565/21 - nikorozenberg@gmail.com



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Resumen Ejecutivo

Introducción y Objetivo

Este informe detalla la preparación y limpieza de datos obtenidos del archivo *Global Shark Attack File* (GSAF), que registra incidentes relacionados con ataques de tiburones. El objetivo principal fue depurar y estructurar la información para analizar con precisión la frecuencia y distribución de estos incidentes según el tipo de ataque y el país.

Para poder mantener la reproducibilidad y sencillez de análisis, importamos los datos originales con un script de Python y realizamos todos los cambios y modificaciones directamente desde el script, que nos devuelve los datos finales.

Decisiones sobre los datos

Para trabajar de forma adecuada, realizamos una limpieza de los datos:

- **Eliminación de Filas y Columnas Vacías:** Se eliminaron 2 filas y 231 columnas que estaban completamente vacías, reteniendo 24 columnas de interés.
- **Remoción de Filas Incompletas:** Se eliminaron 18 filas con valores faltantes en **Country**, **Type** y **Incident Type** (color) (valores claves para nuestro análisis).
- **Estandarización y Validación de la Columna *Country*:** Removimos caracteres especiales, espacios de más, y aplicamos nombres consistentes. Se incorporaron datos de Wikipedia que incluían 387 islas y 204 países para asociar islas a países, y posteriormente validar los países. Todos los casos que no son países, incluyendo casos de océanos, mares, nulos, o que no corresponden con ninguna de las reglas establecidas, fueron catalogados como casos no identificables.

Cabe destacar que tomamos un enfoque en el cual consideramos la **soberanía** de los países sobre los territorios del planeta para agrupar los incidente, a diferencia de, por ejemplo, priorizar la ubicación geográfica.

Resultados y conclusiones

La cantidad de incidentes por tipo fue la siguiente.

| Tipo | Ocurrencias |
|------------------------|-------------|
| Unprovoked Incidents | 5069 |
| Provoked Incidents | 632 |
| Questionable Incidents | 568 |
| Attacks on Boats | 360 |
| Air/Sea Disasters | 239 |
| MISMATCH | 73 |

Cuadro 1: Tabla de tipos y ocurrencias. MISMATCH implica una discrepancia entre la columna de tipo y el color de la fila (que tiene asociado el tipo de incidente).

Los tipos de incidentes por país pueden verse en el anexo C. A continuación los primeros diez países con más incidentes.

| País | Incidentes | A | B | C | D | E | F |
|------------------|------------|----|-----|----|-----|-----|------|
| USA | 2547 | 35 | 82 | 32 | 234 | 199 | 1965 |
| AUSTRALIA | 1485 | 24 | 128 | 10 | 143 | 125 | 1055 |
| SOUTH AFRICA | 596 | 5 | 54 | 3 | 58 | 64 | 412 |
| FRANCE | 184 | 4 | 5 | 2 | 10 | 6 | 157 |
| NEW ZEALAND | 145 | 0 | 17 | 0 | 21 | 13 | 94 |
| PAPUA NEW GUINEA | 139 | 5 | 3 | 1 | 12 | 2 | 116 |
| BAHAMAS | 136 | 5 | 1 | 2 | 18 | 7 | 103 |
| NON-IDENTIFIABLE | 110 | 31 | 8 | 0 | 8 | 9 | 54 |
| MEXICO | 106 | 3 | 3 | 0 | 12 | 14 | 74 |

Cuadro 2: Distribución de tipos de incidente para los primeros diez países con más casos (ver anexo C para significado de las letras).

Se puede observar una predominancia clara de los ataques no provocados (categoría **F**), y vemos como la distribución porcentual de los tipos de incidentes por país es relativamente pareja. Esto nos lleva a pensar que en un primer lugar el área geográfica no influye directamente en la forma en la que se desarrollan los ataques de tiburón, sino que la situación se puede generalizar a muchos casos.

Índice

| | |
|--|-----------|
| 1. Introducción | 4 |
| 2. Requerimientos de calidad | 4 |
| 3. Perfilado | 4 |
| 4. Limpieza | 5 |
| 5. Resultados del análisis y conclusion | 6 |
| A. Tablas de asignación | 9 |
| A.1. | 9 |
| A.2. | 9 |
| B. Asignación manual de islas a países | 10 |
| C. Tabla de incidentes por país | 11 |

1. Introducción

El sitio web Global Shark Attack File (GSAF¹) registra incidentes relacionados con ataques de tiburones. Este informe se centra en el análisis de los datos extraídos de dicho archivo, donde se especifican detalles sobre el incidente, como la fecha, ubicación, actividad, datos de la víctima, y especie de tiburón involucrado. Además, los incidentes están categorizados en tipos según el contexto del ataque, utilizando un sistema de colores en el archivo original para indicar si fueron incidentes no provocados, provocados, ataques a embarcaciones, desastres aéreos/marítimos o incidentes cuestionables.

Nuestro objetivo principal es saber la cantidad de incidentes por tipo, y la cantidad de incidentes por tipo para cada país.

La limpieza de este conjunto de datos es fundamental para realizar un análisis preciso de la frecuencia y distribución de los incidentes por tipo y país. La información sobre el país se encuentra en la columna *Country*, y la información del tipo de incidente se encuentra tanto en la columna *Type*, como en el color de fondo de cada fila. Trabajamos con un único script de Python para poder tener una forma sencilla y reproducible de obtener los datos finales y resultados intermedios a partir de la base de datos original.

2. Requerimientos de calidad

Para asegurar que el conjunto de datos cumpla con los estándares de calidad necesarios, es importante definir reglas que permitan detectar errores y establecer medios para medir la calidad de los datos actuales, y para que se pueda realizar un monitoreo continuo posteriormente. A continuación, enunciamos algunos puntos a considerar para cumplir con los objetivos postulados.

En primer lugar, consideramos importante priorizar la completitud de la información sobre el tipo de ataque y el país, pues serán la información necesaria para el análisis posterior. Además, la información sobre el tipo de incidente en la columna *Type* debe ser coherente con la información reportada en el color de fondo.

Otro aspecto a controlar es la unicidad de los registros: cada incidente debe ser único y no debería repetirse.

Asimismo, se requiere que los datos de ubicación sean consistentes. Utilizaremos un enfoque basado en la soberanía del territorio, y no en la ubicación geográfica, considerando que el país que ejerce soberanía sobre la ubicación donde se produjo el incidente es el principal responsable de atacar la problemática. Los casos en donde se desconozca el país correspondiente, deberán ser analizados por separado. Además, los incidentes que hayan ocurrido en un mismo país, deben tener el mismo valor en la columna que refiere al país; si un país es llamado de distintas formas, se debe elegir una denominación común para todos los casos.

Por último, para este trabajo, a priori, no es de interés evaluar la actualización de los datos o la correctitud de las fechas, pues el objetivo principal es obtener métricas históricas.

3. Perfilado

Para hacer los análisis que siguen, primero se realizó una extracción de los colores, como se describe en la sección 4. A pesar de “agregar” una columna nueva, esta información ya estaba disponible en la base original, únicamente la extrajimos a una columna.

En cuanto a la unicidad de los registros, se calculó la proporción conjunta de casos únicos con las columnas de *Date* (Fecha), *Year* (Año), *Country* (País) y *Name* (el nombre de la víctima). Se ha encontrado que un 99,25 % de las combinaciones eran únicas. Esto representa un número alto, pero para los próximos análisis, se decidió remover los duplicados, para así obtener métricas más exactas.

También analizamos la completitud de la información necesaria. Para eso, se calculó la proporción de valores no nulos sobre totales, para las columnas *Country* (99,03 % de los casos), *Type* (99,49 % de los casos) e *Incident Type (color)* (99,74 % de los casos). A priori, para responder las preguntas iniciales, parecen ser métricas favorables.

En cuanto a la coherencia de la información del tipo de incidente, se ha encontrado que un 1,10 % de los casos tiene información incoherente entre las columnas *Type* y el color.

Tenemos un total de 387 islas (cada una asociada a uno o varios países) y 204 países para validar los datos (ver sección 4). De un total de 6976 casos, 110 no tienen asociación.

¹ <https://www.sharkattackfile.net>

4. Limpieza

Comenzamos eliminando las filas y columnas completamente vacías. Únicamente habían 2 filas (sobre 6976) y nos quedamos con 24 columnas (sobre 255). Luego, eliminamos 18 filas donde los valores “importantes” estaban vacíos (país, tipo y color), ya que no aportan ningún dato para el análisis particular que queremos realizar.

“Normalizamos” la columna *Country*: quitamos caracteres especiales y espacios de más, y pasamos los nombres a mayúscula.

Realizamos la asociación entre islas y países y su posterior validación (de todos los países) de esta manera:

Utilizamos información de Wikipedia² para obtener un listado de 387 islas con sus respectivos países, y convertimos estas tablas a formato CSV mediante una herramienta en línea³, además evaluando ciertos casos a mano, reportados en el anexo B. Recopilamos un conjunto de datos de 204 países soberanos. Estas listas pueden verse en los siguientes enlaces: [países](#), [islas](#)

Nos quedamos con un diccionario que asocia nombres de islas a países, y una tabla con nombres de países válidos. Decidimos mantener la columna original de *Country* (con sus modificaciones) y colocar el nombre “final” de cada país en una nueva columna *Matched Country*.

La cantidad de casos restantes sin asociar son el 1,57%. Debido a que no representan una parte importante de los casos decidimos catalogarlos con el valor **NON-IDENTIFIABLE** (No identificable), e ignorarlos para el análisis. Estos casos son:

- Valores nulos.
- Casos que refieran a un océano, a un golfo, a una bahía, a una costa, o a un mar específico.
- Más de dos países. Por ejemplo, donde aparezca “PAÍS1 / PAÍS2”.
- Todos los casos restantes que no sean países.

Finalmente, transformamos los colores de fondo a tipos de incidente de acuerdo con la tabla en el anexo A.1, y creamos una nueva columna llamada “*Incident Type (color)*” con los valores obtenidos.

Se determina cómo deberían estar apareados los valores de la columna *Type* con los obtenidos a partir de los colores (ver tabla del anexo A.2). Al igual que en el caso de *Country*, dejamos ambas columnas intactas y colocamos los resultados en una nueva columna: *Matched Type*. En caso de no coincidir ambas columnas, ponemos **MISMATCH** como valor.

Uno de los muchos problemas que observamos en la base fue que había dos tonos de amarillo, que asociamos al mismo tipo de incidente. También nos encontramos con diferentes nombres para los mismos países y casos donde no era tan fácil distinguir entre país, territorio, o demás definiciones. Volvemos a hacer énfasis en la utilidad que le dimos al script, donde en vez de ir trabajando con distintas versiones de los datos y tratar de replicar los cambios entre todos los integrantes, cada paso, caso especial y definición están definidos de forma objetiva, precisa y reproducible.

² https://en.wikipedia.org/wiki/List_of_islands_by_area
https://en.wikipedia.org/wiki/Territories_of_the_United_States
https://en.wikipedia.org/wiki/British_Overseas_Territories

³ <https://wikitable2csv.ggor.de/>

5. Resultados del análisis y conclusion

Los resultados principales pueden verse en estos gráficos:

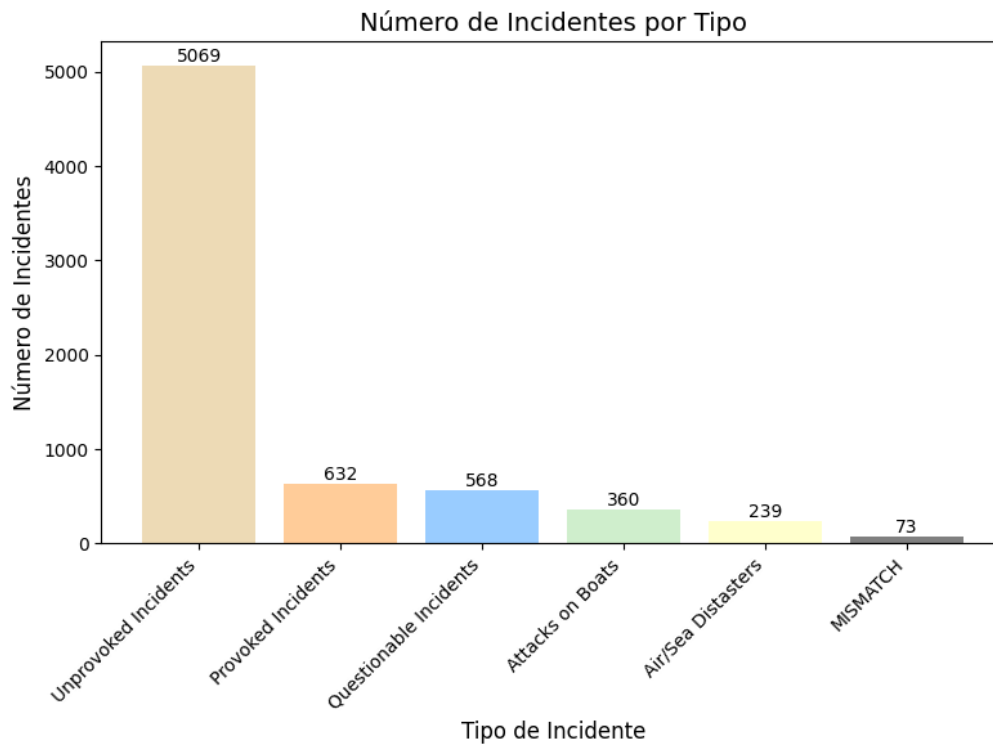


Figura 1: Cantidad de incidentes por tipo.

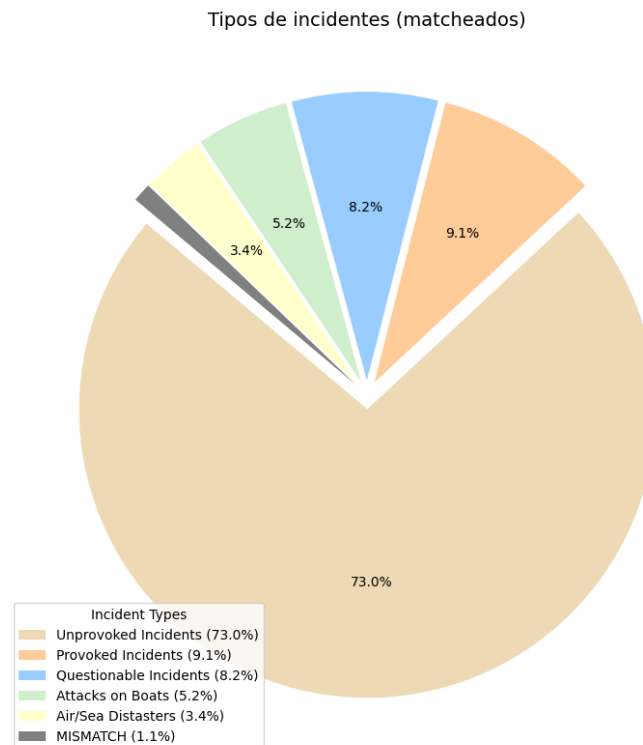


Figura 2: Proporción de incidentes por tipo.

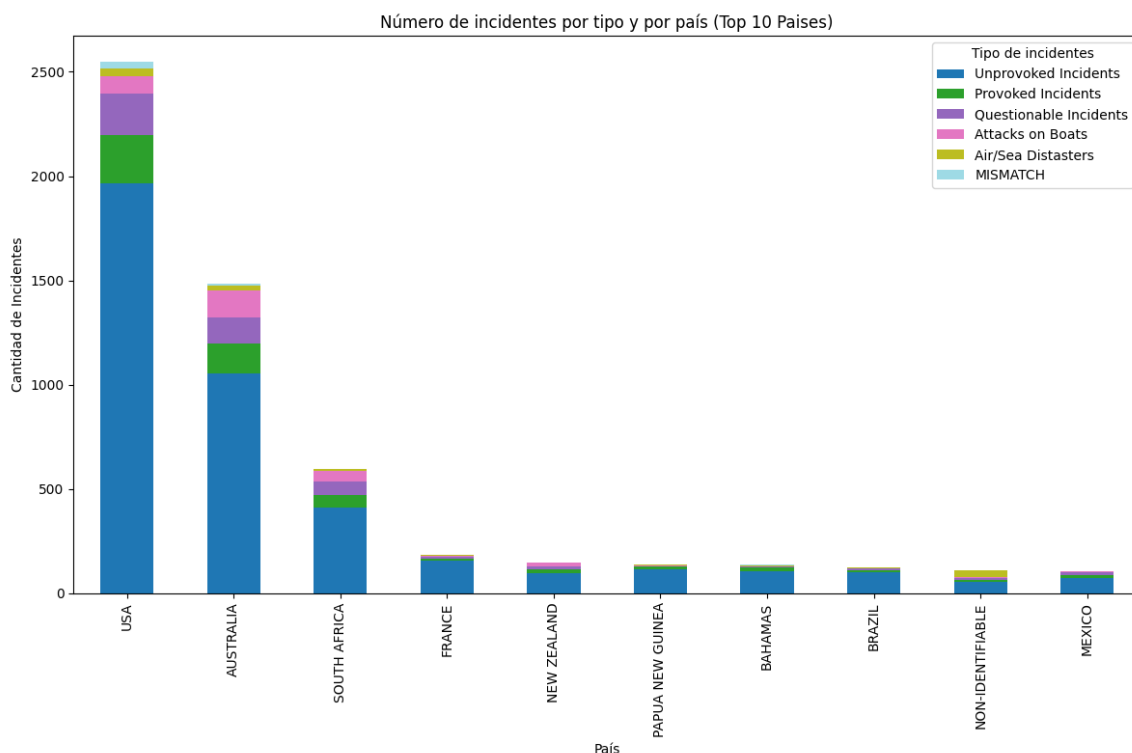


Figura 3: Tipos y cantidad de incidentes de los 10 países con más casos.

Podemos ver la tabla completa de los datos analizados en el anexo [C](#).

A simple vista podemos observar que, por una diferencia considerable, los incidentes no provocados son los más comunes. La distribución por país aparenta ser bastante pareja a lo largo de cada país. Veamos una lista más larga de los países con más casos:

| País | Incidentes |
|------------------|------------|
| USA | 2547 |
| AUSTRALIA | 1485 |
| SOUTH AFRICA | 596 |
| FRANCE | 184 |
| NEW ZEALAND | 145 |
| PAPUA NEW GUINEA | 139 |
| BAHAMAS | 136 |
| BRAZIL | 122 |
| MEXICO | 106 |
| ITALY | 72 |
| FIJI | 70 |
| PHILIPPINES | 64 |
| UNITED KINGDOM | 61 |
| CUBA | 49 |
| EGYPT | 49 |
| MOZAMBIQUE | 46 |
| SPAIN | 46 |
| JAPAN | 42 |
| INDIA | 41 |

En general, la tabla muestra una distribución de incidentes de ataques de tiburones en una variedad de países, con una clara concentración en países costeros. Los países con más casos son aquellos que tienen costas más largas o que están ubicados en zonas frecuentadas por turistas o pescadores, lo que aumenta la probabilidad de interacción entre humanos y tiburones.

El hecho de que Estados Unidos y Australia lideren la lista se puede atribuir a sus extensas costas, con playas populares y zonas de buceo donde los ataques pueden ser más frecuentes. En estos países, la presencia de

tiburones en aguas relativamente cercanas a la costa también podría ser un factor determinante. Además, el alto volumen de turistas en ciertas playas y actividades acuáticas aumenta las probabilidades de encuentros.

Además hay que notar que estos valores pueden estar “inflados”, ya que al asociar islas a países, tenemos territorios no necesariamente cercanos geográficamente que pasan a formar parte del mismo grupo. Lógicamente, este efecto se nota más mientras más territorios de ultramar tenga cada país.

Analizando la distribución de tipos de incidentes por cada país, podemos ver que se asemeja considerablemente al caso general. Esto podría indicar que las condiciones para que suceda un incidente no son tan dependientes de factores geográficos y/o locales a una zona, permitiéndonos sacar conclusiones para el caso general. De esta forma, se podrían llegar a identificar patrones para predecir o prevenir futuros incidentes.

A futuro, se podrían ver demás datos de esta base, como la proporción de incidentes fatales, un análisis con enfoque más geográfico, y una forma de poder estandarizar la entrada o el tratamiento de los datos para tener una fuente de información más consistente.

A. Tablas de asignación

A.1.

| Color (R, G, B) | Tipo de incidente |
|---------------------------------------|------------------------|
| (Nulo) | Blank |
| (255, 255, 204) (Amarillo) | Air/Sea Disasters |
| (255, 255, 197) (Amarillo, otro tono) | Air/Sea Disasters |
| (237, 218, 181) (Marrón) | Unprovoked Incidents |
| (255, 204, 153) (Naranja) | Provoked Incidents |
| (153, 204, 255) (Azul) | Questionable Incidents |
| (207, 238, 204) (Verde) | Attacks on Boats |

Cuadro 4: Asignaciones entre el color de fondo encontrado para cada fila en formato RGB, con el tipo de incidente asociado.

A.2.

| Tipo “original” | Tipo final |
|---------------------|------------------------|
| Unprovoked | Unprovoked Incidents |
| Provoked | Provoked Incidents |
| Questionable | Questionable Incidents |
| ? | Questionable Incidents |
| Unconfirmed | Questionable Incidents |
| Unverified | Questionable Incidents |
| Invalid | Questionable Incidents |
| Under Investigation | Questionable Incidents |
| Boat | Attacks on Boats |
| Watercraft | Attacks on Boats |
| Sea Disaster | Air/Sea Disasters |
| Blank | Blank |

Cuadro 5: Asignaciones elegidas entre los valores para tipos reportados en la columna Type, con los reportados en el color de fondo de las filas.

B. Asignación manual de islas a países

Aclaramos algunos casos donde el conjunto de datos de islas con la información que teníamos de Wikipedia no funcionó pero que luego incorporamos para que cada vez que se haga un análisis sobre las islas se contemplen esos casos y se pueda hacer de forma automática:

- **MAYOTTE**: es un archipiélago que forma parte de Francia, cambiamos el país por el indicado.
- **ROATÁN**: es una isla de Honduras, cambiamos el país por el indicado.
- **CEYLON**: es el antiguo nombre de Sri Lanka, cambiamos el país por el indicado. (Aclaración: Los casos donde dice Ceylon ocurrieron en cierto año donde el país efectivamente se llamaba así, antes de independizarse y cambiarse el nombre).
- **ADMIRALTY ISLANDS**: es una isla de Papúa Nueva Guinea.
- **ANDAMAN ISLANDS**: pertenecen a la India.
- **ANTIGUA**: refiere a una isla que forma parte de Antigua y Barbuda.
- **AZORES**: es un archipiélago de islas que pertenecen a Portugal.
- **BRITISH NEW GUINEA**: es una isla de Papúa Nueva Guinea.
- **BRITISH WEST INDIES**: forma parte de Turks and Caicos, es decir, United Kingdom.
- **CENTRAL PACIFIC**: convertido en Océano Pacífico.
- **DIEGO GARCIA**: isla de United Kingdom.
- **FRENCH POLYNESIA**: pertenece a Francia.
- **GRAND CAYMAN**: es la más grande de las Islas Caimán, un territorio de United Kingdom, de ultramar en el Caribe.
- **MALDIVE ISLANDS**: refiere a MALDIVES.
- **NEVIS**: es una isla del país Saint Kitts and Nevis.
- **NEW CALEDONIA**: es un archipiélago perteneciente a Francia.
- **OKINAWA**: pertenece a Japón.
- **REUNION**: pertenece a Francia.
- **ST MARTIN**: refiere a una isla que es mitad de Francia y mitad de Países Bajos.
- **TOBAGO**: isla de Trinidad y Tobago.

Comentarios sobre las modificaciones al conjunto de datos de islas:

Observamos al menos cuatro casos donde se escribió correctamente el país, pero como hay más de una forma de llamar al país genera problemas. Por ejemplo, *Maldives* y *Maldiv Islands*. Una mejora para esta automatización sería que el dataset tenga todos los nombres de cada país, así si una persona pone cualquiera de los nombres del país al final cuenta como el mismo país.

No encontramos un registro web que contenga la totalidad de las islas que existen en el mundo con su correspondiente país al que pertenecen. Por eso tuvimos que combinar varias páginas de Wikipedia y sumar manualmente algunos casos como los mencionados.

Observamos que muchos errores al intentar automatizarlo ocurren por cómo la gente lo anota. Por ejemplo, en el conjunto de datos de islas tenemos *Mid Ocean Pacific*, pero se registró el valor *Central Pacific* en 2 casos. No es que esté mal la forma en la que lo denominó, sino que es difícil encontrar un criterio que englobe todo lo que la gente pueda poner y no falle. Otro caso similar es el de *Ceylon*, *Ceylon (SRI LANKA)* y *Sri Lanka*, donde los 3 refieren a lo mismo.

C. Tabla de incidentes por país

| Tipo de incidente | Letra asociada |
|------------------------|----------------|
| Air/Sea Disasters | A |
| Attacks on Boats | B |
| MISMATCH | C |
| Provoked Incidents | D |
| Questionable Incidents | E |
| Unprovoked Incidents | F |

| País | Incidentes | A | B | C | D | E | F |
|--------------------------------|------------|----|-----|----|-----|-----|------|
| País | Incidentes | A | B | C | D | E | F |
| ALGERIA | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| ANGOLA | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| ANTIGUA AND BARBUDA | 2 | 1 | 0 | 0 | 0 | 1 | 0 |
| ARGENTINA | 2 | 0 | 0 | 0 | 0 | 1 | 1 |
| ARUBA | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| ATLANTIC OCEAN | 22 | 11 | 1 | 0 | 2 | 1 | 7 |
| AUSTRALIA | 1485 | 24 | 128 | 10 | 143 | 125 | 1055 |
| BAHAMAS | 136 | 5 | 1 | 2 | 18 | 7 | 103 |
| BANGLADESH | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| BARBADOS | 6 | 1 | 0 | 0 | 0 | 0 | 5 |
| BELIZE | 8 | 0 | 0 | 1 | 2 | 1 | 4 |
| BRAZIL | 122 | 5 | 0 | 0 | 6 | 10 | 101 |
| BURMA | 4 | 0 | 0 | 0 | 1 | 0 | 3 |
| Blank | 50 | 8 | 1 | 0 | 3 | 5 | 33 |
| CANADA | 14 | 2 | 3 | 0 | 2 | 1 | 6 |
| CAPE VERDE | 3 | 0 | 0 | 0 | 0 | 2 | 1 |
| CHILE | 8 | 0 | 0 | 0 | 1 | 0 | 7 |
| CHINA | 8 | 1 | 0 | 0 | 3 | 0 | 4 |
| COLOMBIA | 2 | 0 | 0 | 0 | 0 | 1 | 1 |
| COLUMBIA | 10 | 1 | 0 | 0 | 0 | 0 | 9 |
| COMOROS | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| COOK ISLANDS | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| COSTA RICA | 18 | 0 | 0 | 1 | 0 | 2 | 15 |
| CROATIA | 35 | 0 | 3 | 0 | 3 | 7 | 22 |
| CUBA | 49 | 6 | 1 | 2 | 5 | 3 | 32 |
| CYPRUS | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| "DANISH REALM | | | | | | | |
| (KINGDOM OF DENMARK)" | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| DJIBOUTI | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| DOMINICAN REPUBLIC | 7 | 2 | 0 | 0 | 1 | 0 | 4 |
| ECUADOR | 11 | 0 | 0 | 0 | 0 | 0 | 11 |
| EGYPT | 49 | 2 | 0 | 1 | 0 | 4 | 42 |
| EL SALVADOR | 4 | 0 | 1 | 0 | 1 | 0 | 2 |
| FEDERATED STATES OF MICRONESIA | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| FIJI | 70 | 3 | 1 | 0 | 3 | 2 | 61 |
| FRANCE | 184 | 4 | 5 | 2 | 10 | 6 | 157 |
| GABON | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| GEORGIA | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| GHANA | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| GREECE | 26 | 2 | 0 | 0 | 0 | 6 | 18 |
| GRENADA | 4 | 0 | 0 | 0 | 0 | 0 | 4 |
| GUATEMALA | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| GUINEA | 3 | 0 | 0 | 0 | 0 | 1 | 2 |
| GUYANA | 3 | 0 | 0 | 1 | 0 | 1 | 1 |
| HAITI | 3 | 0 | 0 | 0 | 0 | 1 | 2 |
| País | Incidentes | A | B | C | D | E | F |

| País | Incidentes | A | B | C | D | E | F |
|-----------------------|-------------------|----------|----------|----------|----------|----------|----------|
| HONDURAS | 7 | 1 | 0 | 2 | 0 | 1 | 3 |
| HONG KONG | 24 | 1 | 0 | 0 | 0 | 0 | 23 |
| ICELAND | 2 | 0 | 0 | 0 | 0 | 1 | 1 |
| INDIA | 41 | 1 | 1 | 0 | 1 | 1 | 37 |
| INDONESIA | 25 | 6 | 0 | 0 | 0 | 1 | 18 |
| IRAN | 29 | 0 | 0 | 0 | 1 | 0 | 28 |
| IRAQ | 8 | 0 | 0 | 0 | 0 | 0 | 8 |
| ISRAEL | 8 | 0 | 1 | 0 | 1 | 1 | 5 |
| ITALY | 72 | 4 | 16 | 0 | 13 | 12 | 27 |
| JAMAICA | 31 | 1 | 0 | 1 | 2 | 3 | 24 |
| JAPAN | 42 | 3 | 2 | 1 | 2 | 5 | 29 |
| JORDAN | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| KENYA | 10 | 0 | 0 | 0 | 0 | 0 | 10 |
| KIRIBATI | 6 | 2 | 0 | 0 | 0 | 0 | 4 |
| KUWAIT | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| LEBANON | 3 | 0 | 0 | 0 | 0 | 0 | 3 |
| LIBERIA | 3 | 0 | 0 | 0 | 0 | 1 | 2 |
| LIBYA | 6 | 1 | 0 | 0 | 0 | 0 | 5 |
| MADAGASCAR | 8 | 0 | 0 | 0 | 1 | 1 | 6 |
| MALAYSIA | 5 | 0 | 0 | 1 | 0 | 1 | 3 |
| MALDIVES | 7 | 1 | 0 | 2 | 1 | 0 | 3 |
| MALTA | 5 | 0 | 2 | 0 | 0 | 0 | 3 |
| MARSHALL ISLANDS | 13 | 0 | 0 | 0 | 0 | 2 | 11 |
| MAURITIUS | 10 | 1 | 0 | 0 | 2 | 0 | 7 |
| MEXICO | 106 | 3 | 3 | 0 | 12 | 14 | 74 |
| MICRONESIA | 3 | 0 | 0 | 0 | 0 | 0 | 3 |
| MONACO | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| MONTENEGRO | 3 | 0 | 1 | 0 | 0 | 1 | 1 |
| MOROCCO | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| MOZAMBIQUE | 46 | 2 | 2 | 0 | 0 | 4 | 38 |
| NAMIBIA | 2 | 0 | 0 | 0 | 1 | 0 | 1 |
| NEW ZEALAND | 145 | 0 | 17 | 0 | 21 | 13 | 94 |
| NICARAGUA | 7 | 0 | 0 | 0 | 1 | 0 | 6 |
| NIGERIA | 4 | 0 | 0 | 0 | 0 | 0 | 4 |
| NON-IDENTIFIABLE | 123 | 32 | 8 | 0 | 10 | 9 | 64 |
| NORWAY | 2 | 0 | 0 | 0 | 1 | 1 | 0 |
| PACIFIC OCEAN | 11 | 6 | 0 | 0 | 1 | 1 | 3 |
| PALAU | 5 | 0 | 0 | 0 | 1 | 0 | 4 |
| PANAMA | 32 | 0 | 0 | 0 | 4 | 1 | 27 |
| PAPUA NEW GUINEA | 139 | 5 | 3 | 1 | 12 | 2 | 116 |
| PARAGUAY | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| PERU | 2 | 0 | 0 | 1 | 0 | 0 | 1 |
| PHILIPPINES | 64 | 26 | 0 | 0 | 2 | 4 | 32 |
| PORTUGAL | 9 | 0 | 4 | 1 | 3 | 1 | 0 |
| RUSSIA | 4 | 0 | 0 | 0 | 0 | 0 | 4 |
| SAINT KITTS AND NEVIS | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| SAMOA | 10 | 1 | 1 | 1 | 1 | 0 | 6 |
| SAUDI ARABIA | 4 | 0 | 0 | 0 | 0 | 1 | 3 |
| SCOTLAND | 8 | 0 | 3 | 0 | 2 | 0 | 3 |
| SENEGAL | 11 | 1 | 0 | 0 | 2 | 2 | 6 |
| SEYCHELLES | 10 | 0 | 0 | 0 | 0 | 1 | 9 |
| SIERRA LEONE | 9 | 2 | 0 | 0 | 0 | 0 | 7 |
| SINGAPORE | 6 | 0 | 0 | 0 | 0 | 0 | 6 |
| SLOVENIA | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| SOLOMON ISLANDS | 30 | 8 | 0 | 0 | 1 | 0 | 21 |
| SOMALIA | 6 | 0 | 0 | 0 | 1 | 1 | 4 |
| País | Incidentes | A | B | C | D | E | F |

| País | Incidentes | A | B | C | D | E | F |
|----------------------|-------------------|----------|----------|----------|----------|----------|----------|
| SOUTH AFRICA | 596 | 5 | 54 | 3 | 58 | 64 | 412 |
| SOUTH KOREA | 8 | 0 | 0 | 0 | 0 | 0 | 8 |
| SPAIN | 46 | 0 | 1 | 0 | 5 | 10 | 30 |
| SRI LANKA | 17 | 1 | 0 | 1 | 0 | 2 | 13 |
| SUDAN | 4 | 1 | 0 | 0 | 1 | 0 | 2 |
| SWEDEN | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SYRIA | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| TAIWAN | 9 | 1 | 0 | 0 | 2 | 0 | 6 |
| TANZANIA | 8 | 0 | 0 | 0 | 1 | 1 | 6 |
| THAILAND | 12 | 0 | 0 | 0 | 2 | 0 | 10 |
| TONGA | 18 | 2 | 0 | 0 | 2 | 2 | 12 |
| TRINIDAD & TOBAGO | 3 | 0 | 0 | 0 | 1 | 1 | 1 |
| TRINIDAD AND TOBAGO | 4 | 0 | 0 | 2 | 0 | 0 | 2 |
| TUNISIA | 3 | 0 | 0 | 0 | 0 | 0 | 3 |
| TURKEY | 12 | 0 | 6 | 0 | 1 | 1 | 4 |
| TUVALU | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| UNITED ARAB EMIRATES | 2 | 0 | 1 | 0 | 0 | 0 | 1 |
| UNITED KINGDOM | 61 | 3 | 4 | 1 | 18 | 9 | 26 |
| URUGUAY | 4 | 0 | 0 | 0 | 0 | 0 | 4 |
| USA | 2547 | 35 | 82 | 32 | 234 | 199 | 1965 |
| VANUATU | 14 | 0 | 1 | 0 | 1 | 0 | 12 |
| VENEZUELA | 11 | 0 | 0 | 0 | 0 | 0 | 11 |
| VIETNAM | 15 | 1 | 0 | 1 | 1 | 0 | 12 |
| YEMEN | 9 | 1 | 0 | 0 | 0 | 1 | 7 |