# Dataset

## Fuente de datos

Descripción: Voy a trabajar con varias bases de datos. Los datos están relacionados con los desastres naturales: número de mortalidad, tipos de desastres naturales y su respectiva frecuencia. También analizaré qué relación tiene con el cambio climático, sobre todo con el cambio de la temperatura.

Origen: Los bases de datos serán tomadas de estas tres páginas principalmente.

Desastres naturales: <https://ourworldindata.org/natural-disasters> (actualizado en 2019-11) (desde 1900 hasta actualidad)

<https://www.kaggle.com/dataenergy/natural-disaster-data> (actualizado hace 1 año) (desde 1900 hasta actualidad)

Cambio climático(temperatura): <https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data#GlobalLandTemperaturesByCountry.csv> (actualizado hace 3 años) (desde 1743 hasta actualidad)

Formato: todos en csv.

Tamaño: suma de los csv utilizados (25851 KB 25 MB)

## Adquisición de datos:

Descargado de las páginas

## Modelo de datos

* GlobalLandTemperaturesByCountry.csv: contiene la temperatura media mensual de cada país desde noviembre de 1973 hasta 2013.

Formato de la tabla:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | dt | AverageTemperature | AverageTemperatureUncertainty | Country |
| formato | YYYY-MM-dd  (En Scala se parsea como tipo timestamp) | double | double | String |

* Number-of-natural-disaster-events.csv: Número de desastres naturales por año y por tipos.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | Number |
| Descripción | Tipo de desastres | En este caso Null | Año desde 1900 hasta 2018 | Número de desastres naturales reportados |
| Formato | String |  | Integer | Integer |

* economic-damage-from-natural-disasters.csv: El daño economico total causado por los desastres naturales por año y por tipo.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | money |
| Descripción | Tipo de desastres | En este caso Null | Año desde 1900 hasta 2018 | El daño económico total (US$) |
| Formato | String | Null | Integer | Long |

* number-of-deaths-from-natural-disasters.csv: Número de Muertos debido a los desastres naturales por año y por tipo.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | Deaths |
| Descripción | Tipo de desastres | En este caso Null | Año desde 1900 hasta 2019 | Muertes globales |
| Formato | String | Null | Integer | Integer |

* deaths-natural-disasters-ihme.csv: Número de Muertos debido a los desastres naturales por año y por país.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | Deaths |
| Descripción | País | Código del país en el formato ISO 3166-1 alpha-3 | Año desde 1990 hasta 2017 | Deaths - Exposure to forces of nature - Sex: Both - Age: All Ages (Number) (deaths) |
| Formato | String | String | String | Long |

* share-deaths-from-natural-disasters.csv: Porcentaje de Muertos debido a los desastres naturales por año y por país.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | Deaths(Percent) (%) |
| Descripción | País | Código del país en el formato ISO 3166-1 alpha-3 | Año desde 1990 hasta 2017 | Porcentaje de Muertos causado por los desastres naturales en dicho país |
| Formato | String | String | String | Double |

* significant-earthquakes.csc: número de terremotos significativos por país y año

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | Number of significant earthquakes (significant earthquakes) |
| Descripción | País | Código del país en el formato ISO 3166-1 alpha-3 | Año desde 2150 a.c. hasta 2017 | número de terremotos significativos |
| Formato | String | String | Integer | Integer |

* significant-volcanic-eruptions.csv:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Entity | Code | Year | Number of significant volcanic eruptions (NGDC-WDS) (significant eruptions) |
| Descripción | País | Código del país en el formato ISO 3166-1 alpha-3 | Año desde 1750 a.c. hasta 2017 | número de erupciones volcánicas significativos |
| Formato | String | String | Integer | Integer |

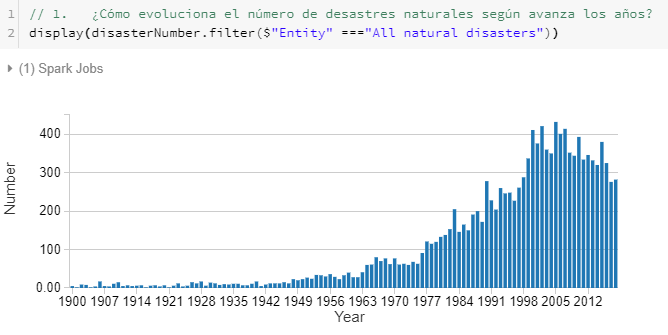
## Planteamiento inicial

1. ¿Cómo evoluciona el número de desastres naturales según avanza los años?
2. ¿Existe alguna relación entre el número de desastres naturales con el daño económico?
3. ¿Cuál es la temperatura anual de cada país?
4. ¿Existe alguna relación entre el número de desastres naturales con el número total de muertos anuales causados por los desastres naturales?
5. ¿Cómo evolucionan cada tipo de desastres naturales según avanza el tiempo?
6. ¿Cuál es el país con más muertes por los desastres naturales?
7. ¿Cómo evoluciona la temperatura media anual global?
8. ¿Existe alguna relación entre la temperatura media anual global con el número de desastres naturales?
9. ¿Qué tipo de desastres naturales provoca mayores muertos?
10. ¿Qué tipo de desastres naturales provoca mayores daños económicos?
11. ¿Cuándo hubo más concentración de terremotos significativos?
12. ¿Cuándo hubo más concentración de erupciones volcánicas?
13. ¿Cómo evoluciona el porcentaje de muertes?

## Primeras aproximaciones sobre el planteamiento

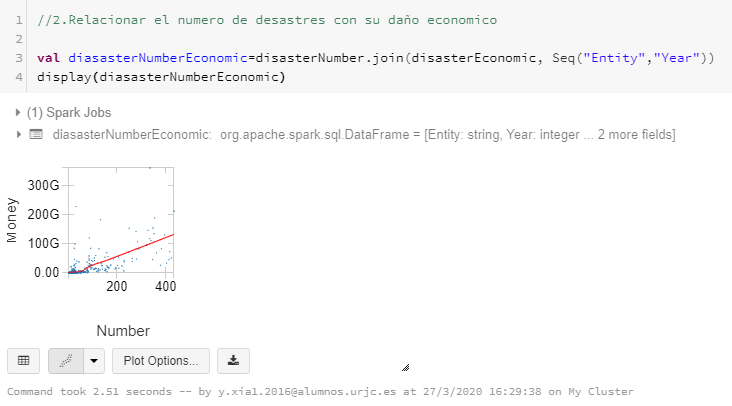
Con la ayuda de las gráficas en la plataforma Databricks, podemos dar las primeras conclusiones sobre las preguntas:

1. ¿Cómo evoluciona el número de desastres naturales según avanza los años?



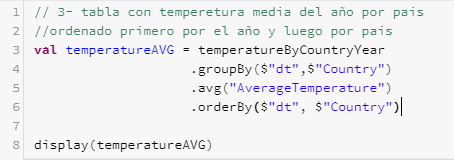
Podemos ver que, aunque hay fluctuaciones, la tendencia en general es que el número de desastres naturales es cada vez más.

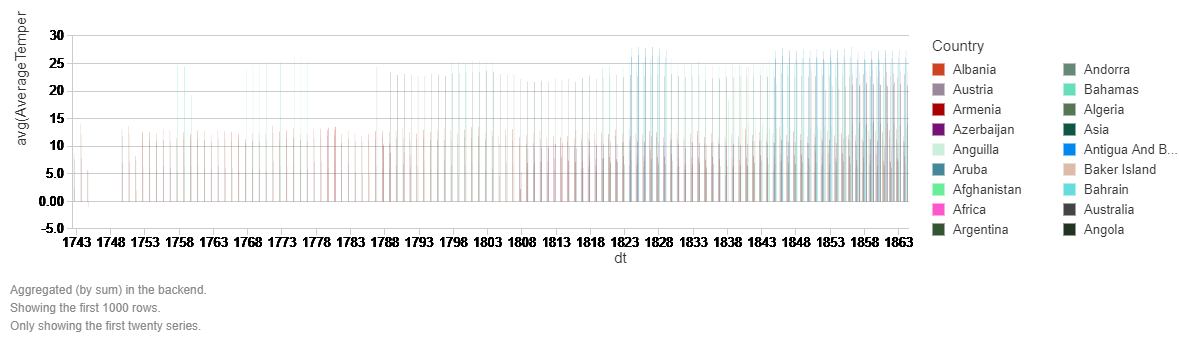
1. ¿Existe alguna relación entre el número de desastres naturales con el daño económico?



Podemos observar que cuando el número de desastres naturales aumenta también aumenta el daño económico causado.

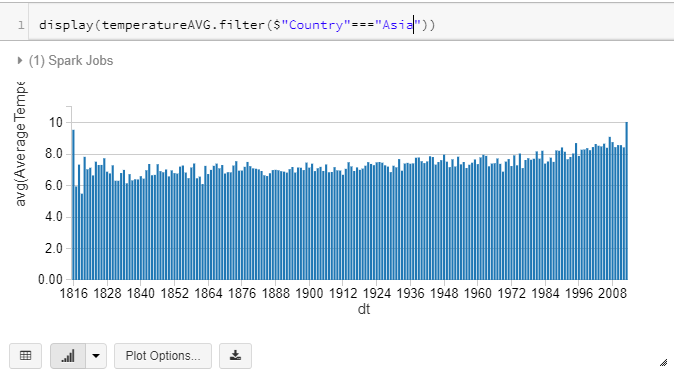
1. ¿Cuál es la temperatura anual de cada país?



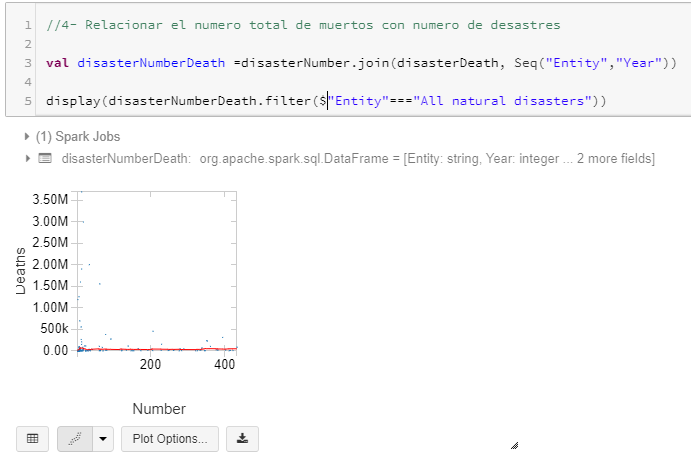


La temperatura media de los países fluctúa entre 15 y 25 grados centígrados, pero a partir de los años 1800 vemos que algunos países superan los 25 grados centígrados.

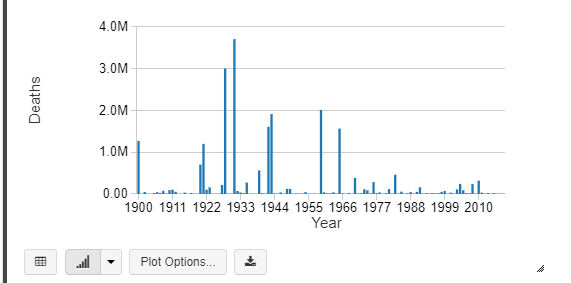
Debido que son muchos países y no se puedo ver bien, he cogido una gráfica de ejemplo con Asia y se puede ver que hay aumentos en los últimos años.



1. ¿Existe alguna relación entre el número de desastres naturales con el número total de muertos anuales causados por los desastres naturales?

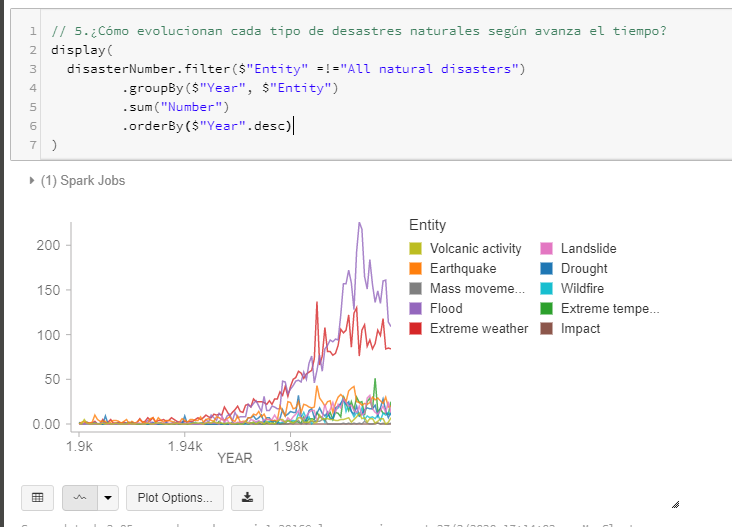


Al parecer no hay una relación entre ellas. En el apartado 1, hemos visto que el número de desastres naturales es cada vez más pero el de las muertes no ocurre lo mismo y podemos apreciarlo en la siguiente gráfica:



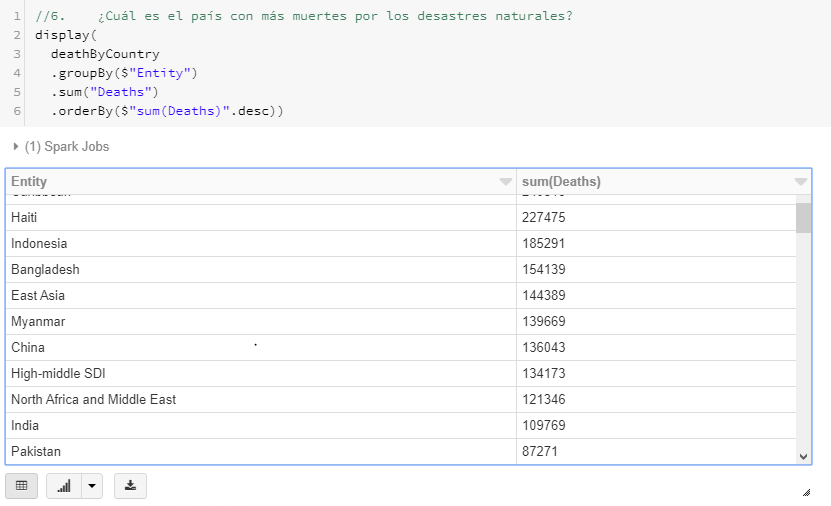
Y eso puede ser debido a los avances que hubo en los últimos años, lo que hace que el número de muertos sea menos.

1. ¿Cómo evolucionan cada tipo de desastres naturales según avanza el tiempo?



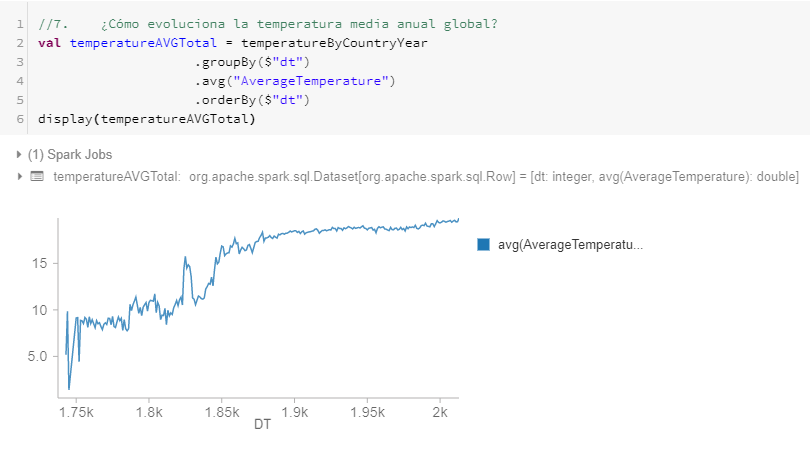
Como en el apartado 1 hemos visto el número de desastres naturales aumenta cada vez más, podemos observar que todos los tipos de desastres han aumentado, sobre todo las inundaciones y las meteorologías extremas.

1. ¿Cuál es el país con más muertes por los desastres naturales?



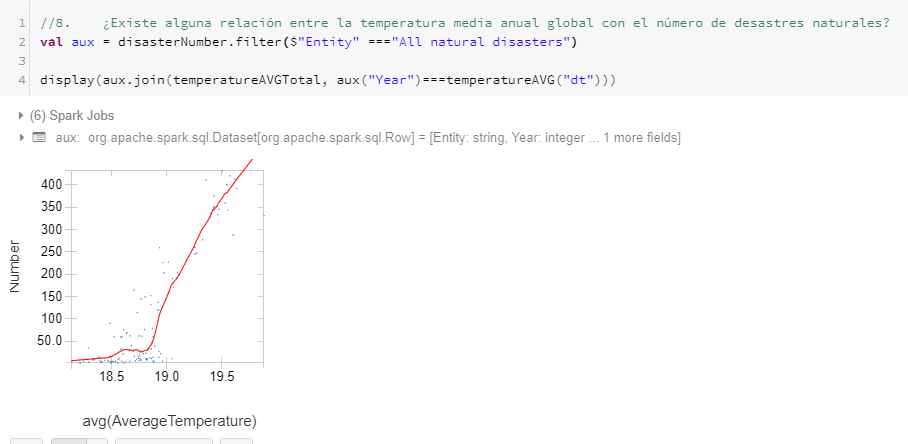
Según la tabla es Haití con un total de 227475.

1. ¿Cómo evoluciona la temperatura media anual global?



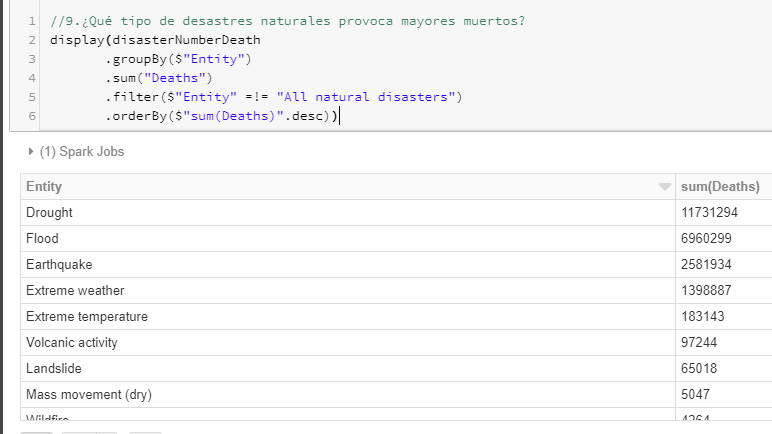
Podemos ver que la temperatura media global va aumentando.

1. ¿Existe alguna relación entre la temperatura media anual global con el número de desastres naturales?

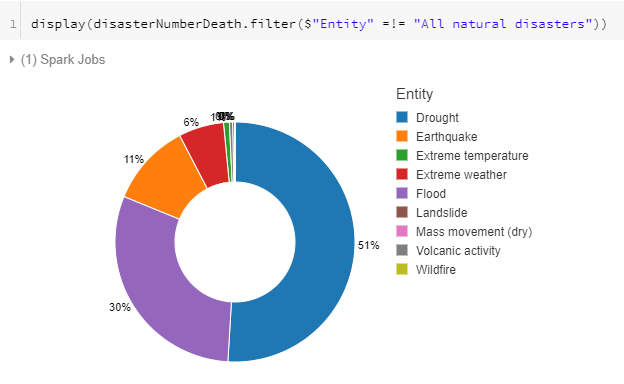


Podemos ver que hay una relación exponencial entre la temperatura media con el número de desastres naturales.

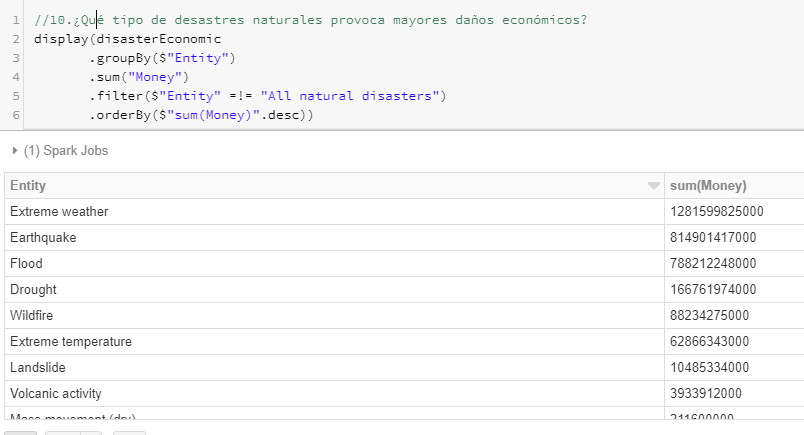
1. ¿Qué tipo de desastres naturales provoca mayores muertos?



Según la query, son las inundaciones, eso también concuerda con que las inundaciones son las más frecuentes.

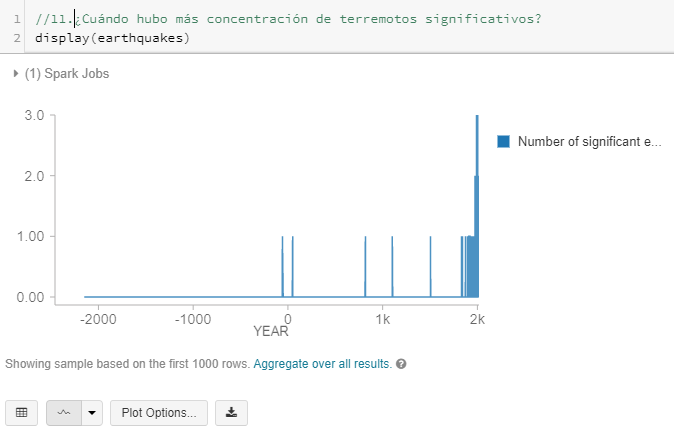


1. ¿Qué tipo de desastres naturales provoca mayores daños económicos?



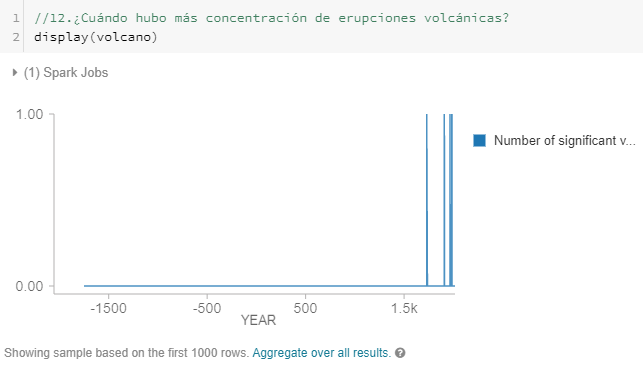
Según los datos es la meteorología extrema.

1. ¿Cuándo hubo más concentración de terremotos significativos?



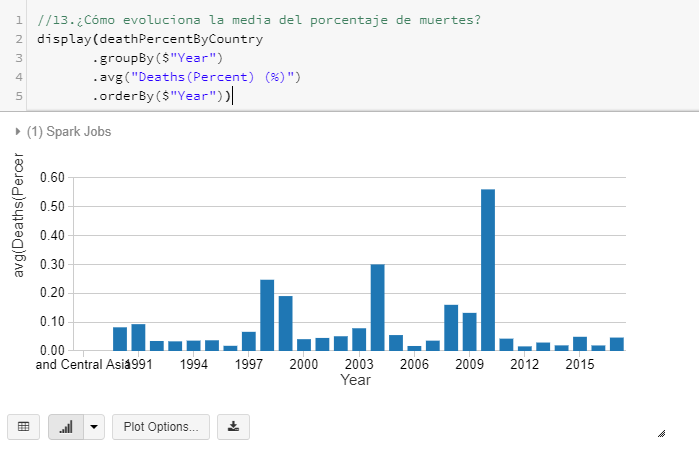
Se puede observar que se concentra en los últimos años.

1. ¿Cuándo hubo más concentración de erupciones volcánicas?



Igual que en el apartado anterior, en los últimos años.

1. ¿Cómo evoluciona la media del porcentaje de muertes?



En general no ha sobrepasado 0.6% de la población.