

Practica 1 WEB SCRAPING

1. Contexto. Explicar en qué contexto se ha recolectado la información. Explique por qué el sitio web elegido proporciona dicha información.

Cambio político es un sitio web (CPNEWS), que nació en virtud de la necesidad de contar con un espacio donde se pueda leer columnas de opinión, análisis, entrevistas y reportajes de un sector de pensamiento que usualmente no es muy divulgado, se constituye en un esfuerzo por fortalecer el análisis, la reflexión y el estudio de temas políticos, económicos y sociales de la realidad costarricense y latinoamericana, mediante el desarrollo de un portal electrónico que contenga artículos, análisis, documentos y columnas en esa línea.

Este sitio web ofrece la información en línea que publica la OMS de las estadísticas globales del COVID 19 a nivel mundial.

2. Definir un título para el DATASET. Elegir un título que sea descriptivo

El nombre del set de datos es covid-19.csv, el mismo que será incluido en el envío de esta práctica.

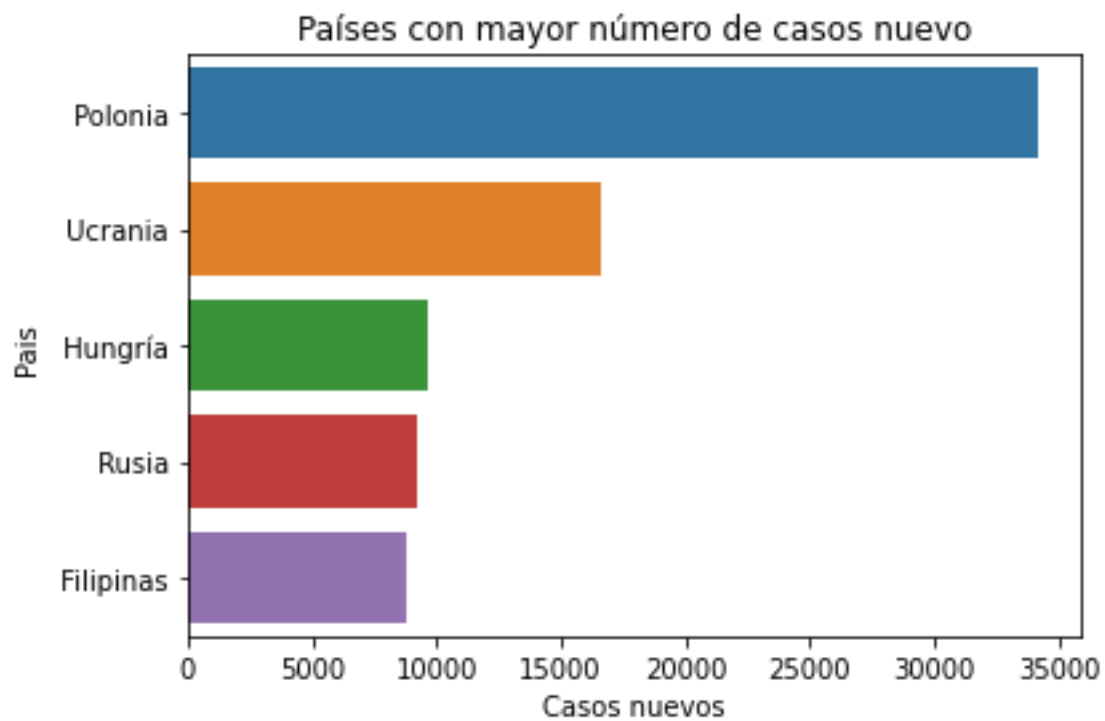
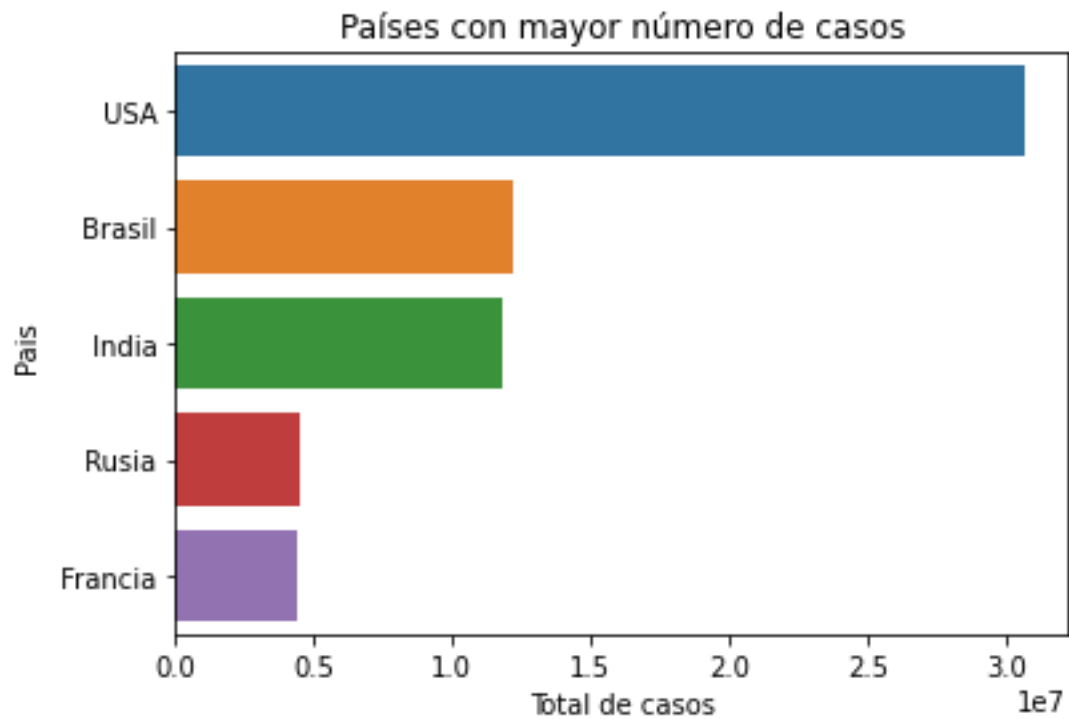
3. Descripción del DATASET.

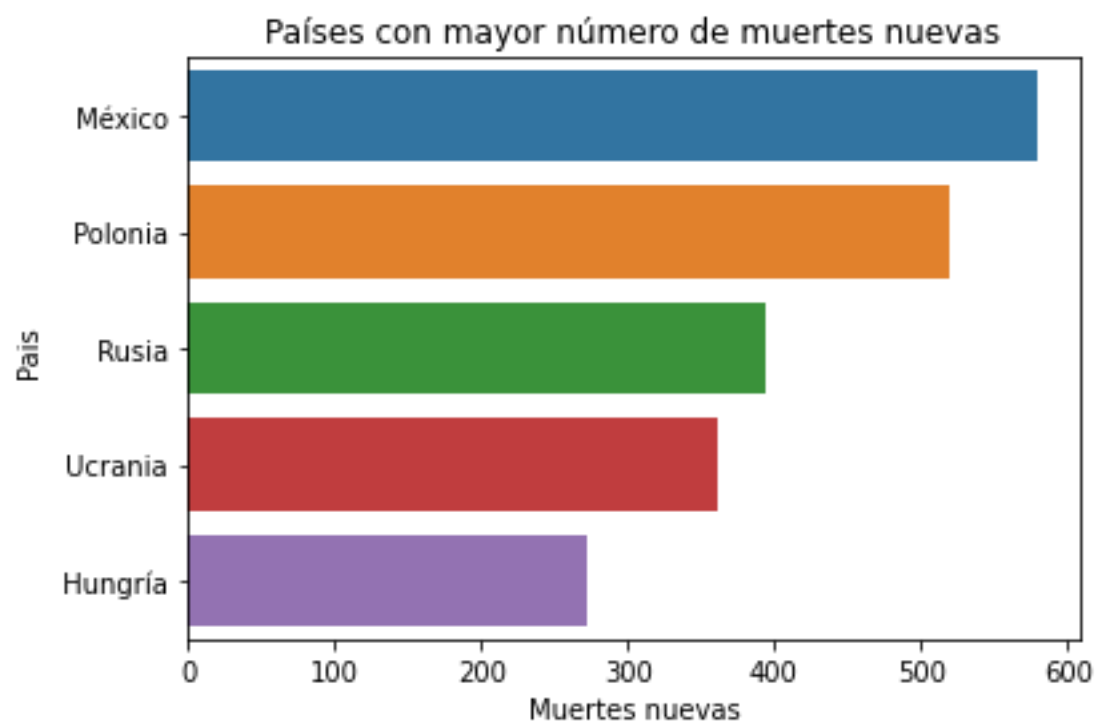
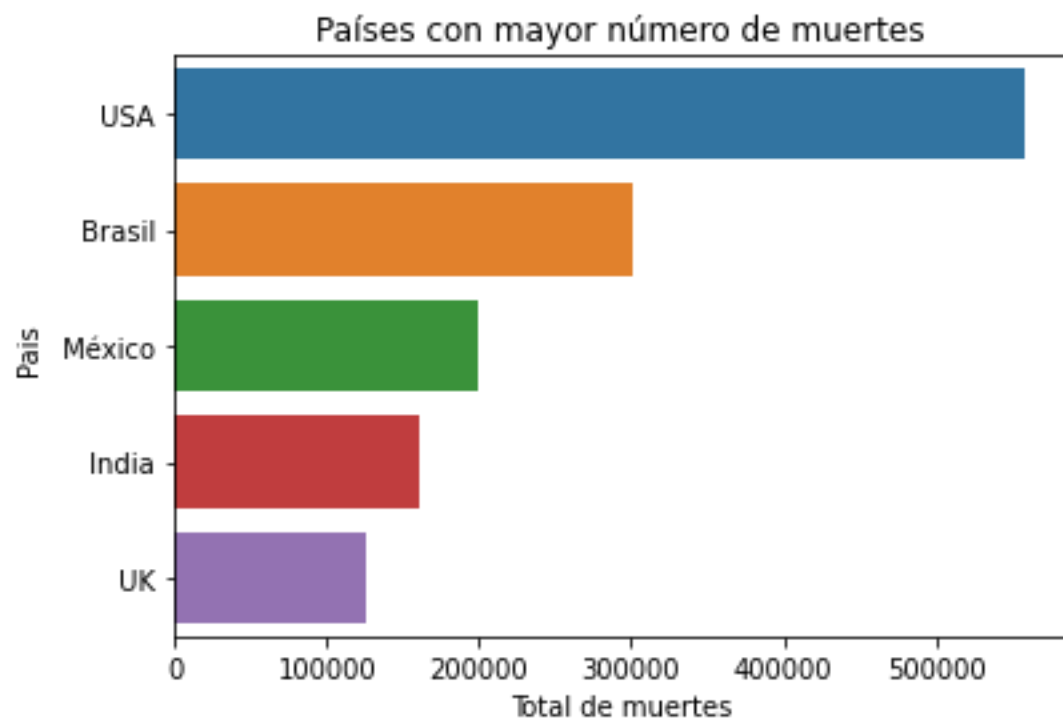
El set de datos contiene ocho columnas:

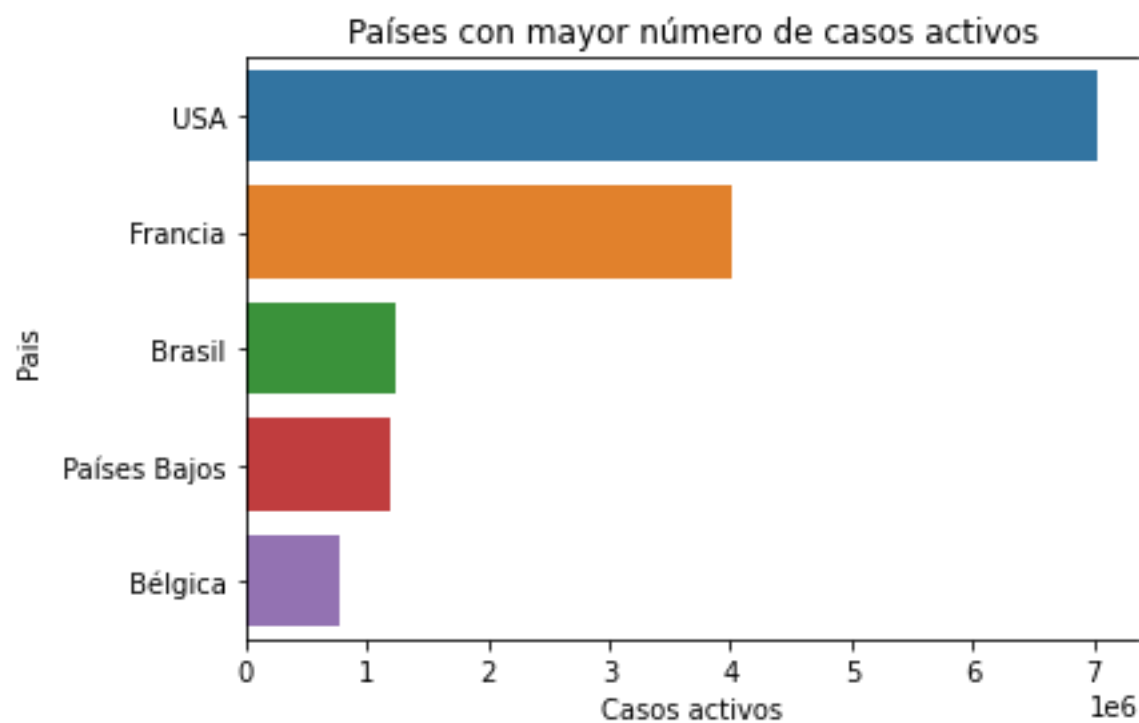
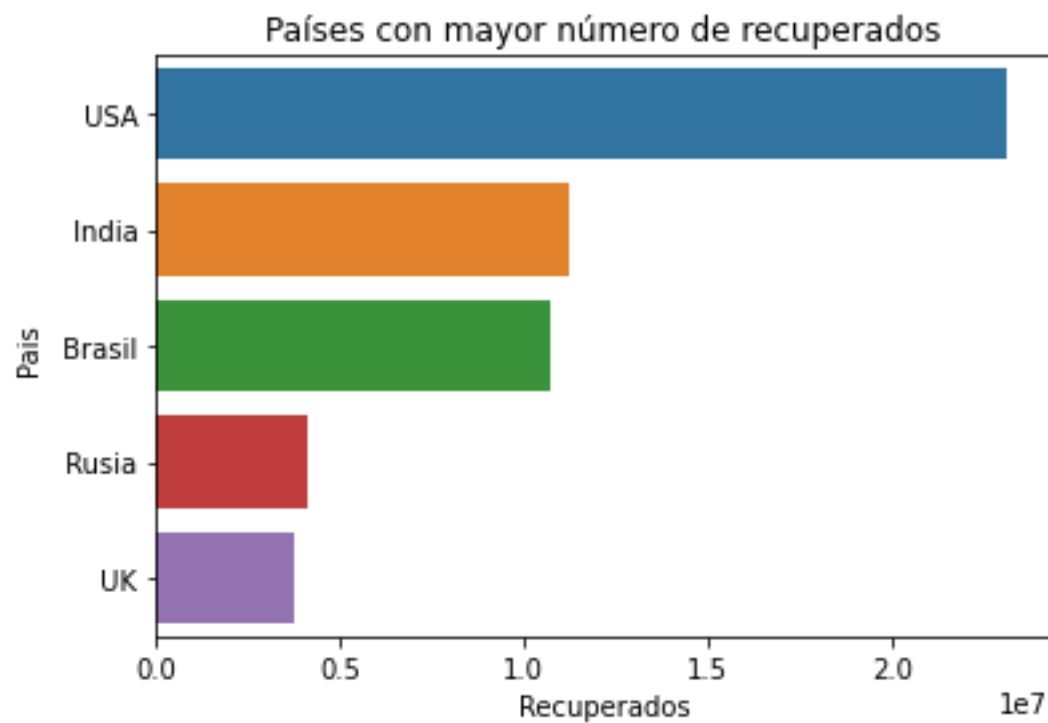
- País
- Total de casos
- Casos nuevos
- Total de muertes
- Muertes nuevas
- Recuperados
- Casos activos
- Casos críticos

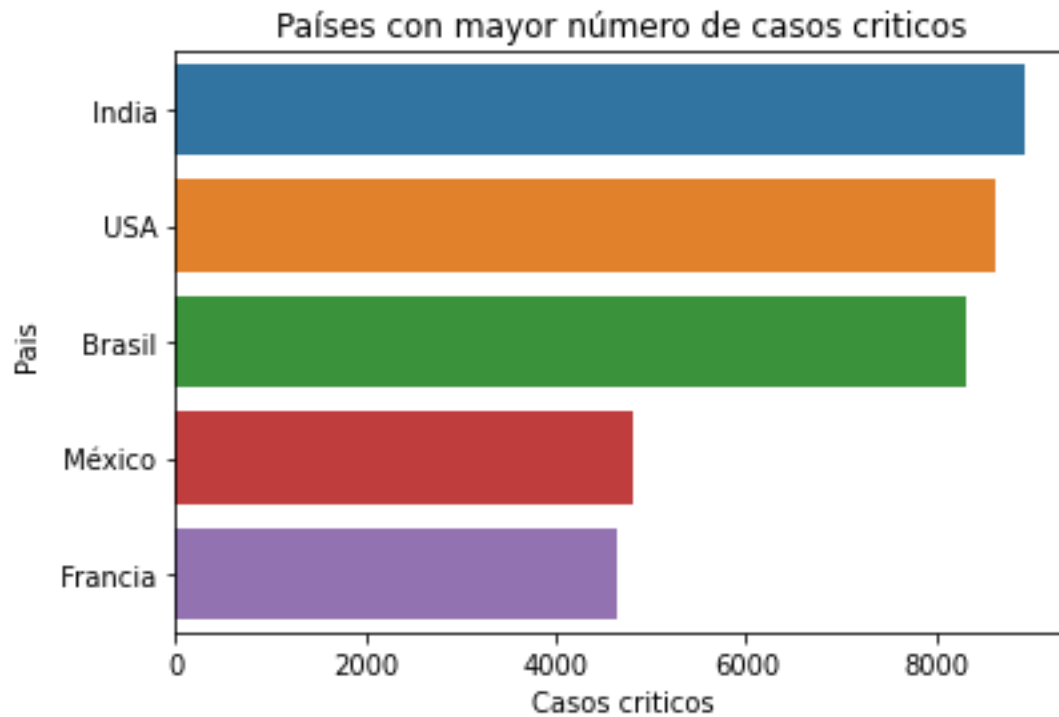
Las columnas tienen nombres descriptivos todos relacionados con el COVID-19

4. Representación gráfica. Presentar esquema o diagrama que identifique el DATASET visualmente y el proyecto elegido









5. Contenido. Explicar los campos que incluye el DATASET, el periodo de tiempo de los datos y cómo se ha recogido.

El set de datos contiene ocho columnas

- País. - Nombre del país al que se le aplica las distintas mediciones
- Total de casos. - Número total de casos registrados por país
- Casos nuevos. - Número total de casos nuevos (Último mes) registrados por país
- Total de muertes. – Número total de muertes registrada por país
- Muertes nuevas. - Número total de muertes nuevas (Último mes) registrados por país
- Recuperados. – Número totas de personas recuperadas
- Casos activos. – Número total de casos activos por país
- Casos críticos. - Número total de casos críticos por país
-

La información que proporciona el sitio Web es actualizada mes a mes con los datos de la OMS

6. Agradecimientos. Presentar al propietario del conjunto de datos. Es necesario incluir citas de análisis anteriores o, en caso de no haberlas, justificar esta búsqueda con análisis similares.

La información es publica proporcionada por la OMS a través del sitio web cambio político.

7. Inspiración

Debido a la pandemia que está azotando al mundo, es de relevancia importancia la información proporcionada, dado que así podemos ver y analizar como esta evolucionado el COVID 19 en los distintos países del mundo.

8. Licencia

La licencia escogida para este set de datos es CC BY-SA 4.0 LICENSE dado que esta licencia permite a otros remezclar, modificar y basarse en su trabajo incluso con fines comerciales, siempre y cuando le acrediten y os autoricen sus nuevas creaciones bajo los mismos términos. Esta licencia se compara a menudo con licencias de software gratuitas y de código abierto "copyleft". Todas las obras nuevas basadas en la suya llevarán la misma licencia, por lo que cualquier derivado también permitirá el uso comercial. Esta es la licencia utilizada por Wikipedia, y se recomienda para materiales que se beneficiarían de la incorporación de contenido de Wikipedia y proyectos con licencia similar.

<https://creativecommons.org/licenses/by-sa/4.0/>

<https://creativecommons.org/licenses/by-sa/4.0/legalcode>

9. Código fuente y DATASET

Tanto el código fuente escrito para la extracción de datos como el set de datos generado pueden ser accedidos a través de este enlace https://github.com/raulec1982/Practica_1_Raul_Sanchez_Campos.git

10. Publicación del DATASET en formato CSV en ZENODO (obtención del DOI) con una breve descripción.

<https://zenodo.org/record/4637936#.YFzx8K9KjIV>