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

Queremos recolectar datos de internet para hacer un estudio. La mayoría de webs no proporcionan APIs para descargar esos datos por lo que utilizaremos la técnica del webscraping. Hemos elegido el sector de electrodomésticos para realizar el citado estudio y concretamente la cadena Expert ya que al analizar el archivo robots.txt hemos comprobado que no estaba prohibido el uso de robots.

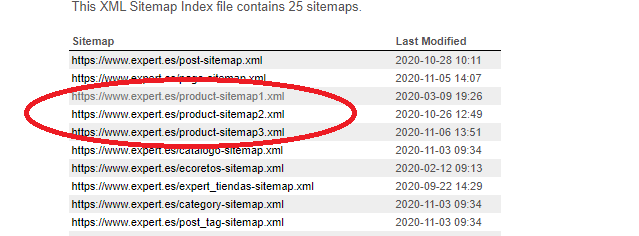
User-Agent: \*

Disallow: /wp-admin/

Allow: /wp-admin/admin-ajax.php

Sitemap: https://www.expert.es/sitemap\_index.xml

Nos interesa tener los datos de los productos que se ofertan en esta cadena y después de acceder al sitemap y de hacer un análisis manual vimos de donde podíamos sacar la información que nos interesaba.



**2. Definir un título para el dataset. Elegir un título que sea descriptivo.**

electrodata.csv ya que refiere a electrodomésticos y datos.

**3. Descripción del dataset. Desarrollar una descripción breve del conjunto de datos que se ha extraído (es necesario que esta descripción tenga sentido con el título elegido).**

Se ha extraído un dataset de la web https://www.expert.es en la que constan el nombre del producto, su precio, si está agotado o no, y su categoría.

**4. Representación gráfica. Presentar una imagen o esquema que identifique el dataset visualmente.**

A la espera de contestación de la consultora.

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

A la espera de contestación de la consultora.

**6. Agradecimientos. Presentar al propietario del conjunto de datos. Es necesario incluir citas de investigación o análisis anteriores (si los hay).**

Agradecemos a EXPERT ESPAÑA el uso de sus datos con fines educativos.

**7. Inspiración. Explique por qué es interesante este conjunto de datos y qué preguntas se pretenden responder.**

Este conjunto de datos puede ser interesante por si solo o como primer paso para un proyecto mayor. En el primer caso puede servir para la misma cadena Expert para analizar sus productos (aunque lo normal es que este análisis lo hagan desde sus bases de datos) y también para la competencia para comparar con sus productos. Consideramos que también puede ser útil para alimentar un comparador de precios en el ámbito de los electrodomésticos.

**8. Licencia. Seleccione una de estas licencias para su dataset y explique el motivo de su selección: ○ Released Under CC0: Public Domain License ○ Released Under CC BY-NC-SA 4.0 License ○ Released Under CC BY-SA 4.0 License ○ Database released under Open Database License, individual contents under Database Contents License ○ Other (specified above) ○ Unknown License**

Nosotros nos decantamos por Released Under CC0: Public Domain License que según la web oficial creative.commons.org “permite a los científicos, educadores, artistas y otros creadores y propietarios de contenido protegido por derechos de autor o bases de datos renunciar a esos intereses en sus obras y, por lo tanto, colocarlos lo más completamente posible en el dominio público, de modo que otros puedan construir libremente sobre ellos, mejorarlos y reutilizar las obras para cualquier propósito sin restricciones bajo derechos de autor o leyes de bases de datos.” Esta filosofía está en concordancia con la del repositorio Zenodo donde habrá que publicar el dataset.

**9. Código. Adjuntar el código con el que se ha generado el dataset, preferiblemente en Python o, alternativamente, en R.**

Cuando lo tengamos ;)

**10. Dataset. Publicación del dataset en formato CSV en Zenodo (obtención del DOI) con una breve descripción.**

A la espera de contestación de la consultora.