

Practica 1 de Tipologia y Ciclo de Vida de los Datos

<https://www.elcorteingles.es/robots.txt>

Descripción

 Se debería mostrar la imagen

Miembros del equipo

Esta practica ha sido realizada por Oscar Pedreño y Daniel Villalobos

1. Contexto

Trabajamos en el departamento de *Data* para unos grandes almacenes dedicados exclusivamnete a la moda de hombre. Se nos ha pedido que realicemos un estudio de los precios de la competencia para conocer nuestra situación de mercado respecto a nuestros competidores. Nos centraremos en **El corte Inglés** ya que por volumen de venta nos manejamos sobre las mismas cuotas de mercado.

La página web de donde obtendremos la información principalmente es:

<https://www.elcorteingles.es/moda-hombre/ropa/>

2. Título Dataset

El dataset final que generaremos a partir de los datos recolectados será:

precios_eci_ropa

3. Descripción

Los datos que nos interesa obtener principalmente son, el precio de la prenda, la marca, el nombre de la prenda, la categoría, los colores de la prenda, sus características. Los campos que extraeremos son:

- Fecha
- Precio
- Prenda
- Marca
- Categoría
- Colores
- Composición
- Mangas

4. Representación gráfica

Poner foto representativa de la info a extraer. Esta en el repositorio.

5. Contenido

Campos dataset

Los campos contenidos en el dataset son:

- Fecha
- Precio
- Prenda
- Marca
- Categoría
- Colores
- Composición
- Mangas

Periodo temporal

Cómo se ha recogido

6. Agradecimientos

La página web desde donde extraemos los datos es **El Corte Inglés**, unos grandes almacenes reconocidos mundialmente.

Otros trabajos serían: (No he encontrado nada, con python)

Se ha utilizado el lenguaje de programación *Python* y de técnicas de *Web Scraping* para extraer la información alojada en las páginas HTML.

7. Inspiración

La motivación principal para realizar este proyecto es responder la pregunta: ¿Que diferencia de precios existe entre la competencia y nosotros? De esta manera podemos realizar ajustes de precios para ser más competitivos y analizar si la competencia esta realizando prácticas ilegales al rebajar precios por debajo del precio de mercado.

8. Licencia

La licencia escogida para este trabajo es la licencia CC BY-SA 4.0. Los motivos principales que nos han llevado a elegir esta licencia son:

- Nos permite reconocer el trabajo realizado previamente y las aportaciones que se han realizado respecto el trabajo original.
- Se permite un uso comercial por lo que podría ser utilizado por terceros, pero siempre reconociendo el trabajo original del autor.
- Las contribuciones se deben publicar sobre la misma licencia por lo que se respetan los dos puntos anteriores. De esta manera, se sigue reconociendo al autor original.

9. Código y Dataset

10. Tabla de Firmas

Contribuciones	Firma
Investigación Previa	OP; DV
Redacción de las Respuesta	OP; DV
Desarrollo de código	OP; DV

11. Recursos

Poner la bibliografía del temario del curso)

- **Esta fuente nos puede ayudar con lo que queremos realizar**
https://www.hlaboa.com/post/Compra_tu_casa_de_forma_inteligente_1_web_scraping/