

## 1.- Contexto

Esta práctica se ha realizado en el contexto de la asignatura Tipología y ciclo de vida de los datos, con código M12.851, del Máster de Ciencia de Datos de la Universitat Oberta de Catalunya.

En la práctica se solicitaba la elección de una web, para posteriormente identificar los datos relevantes y realizar una extracción de éstos a través de alguna de las diferentes técnicas explicadas en el Bloque 1 de la asignatura, indicando como método preferente el web scraping.

La web escogida ha sido la del Mercado de Metales de Londres, <https://www.lme.com>, en la cual se refleja de manera diaria las diferentes cotizaciones adquiridas por los metales que procesa.

Se ha revisado que no hay disponibilidad de una API propia libre y gratuita, si es posible adquiriendo una licencia anual de 18.000 (con un retraso de los datos de 30 minutos) o 3.000 dólares (con un retraso en los datos de 1 día).

La información, con un día de retraso, se puede obtener, sin histórico, a través de la página principal para los principales metales, por lo que la realización de un scrap que se programe diariamente y que vaya añadiendo en un dataset la información nos daría la misma información que la licencia de 3.000 dólares de una forma gratuita.

Se ha revisado el fichero robots.txt y no hay limitaciones en el mismo, así como las condiciones legales, que indican que los usuarios registrados no pueden sustraer información de la web, por lo que toda la labor de scrap se realizará en la web sin realizar login de la misma, para no incumplir las condiciones legales.

## 2.- Definición del Dataset

Valores diarios de los principales metales en el London Metal Exchange.

## 3.- Descripción del Dataset

Recolección diaria de valores obtenidos en el LME de los siguientes metales:

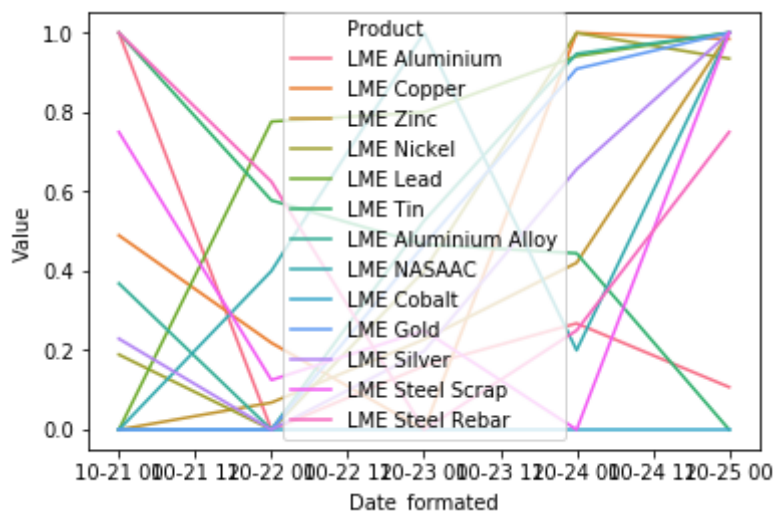
- Aluminium
- Copper
- Zinc
- Nickel

- Lead
- Tin
- Aluminium Alloy
- NASAAC
- Cobalt
- Gold
- Silver
- Steel Scrap
- Steel Rebar

## 4.- Representación gráfica

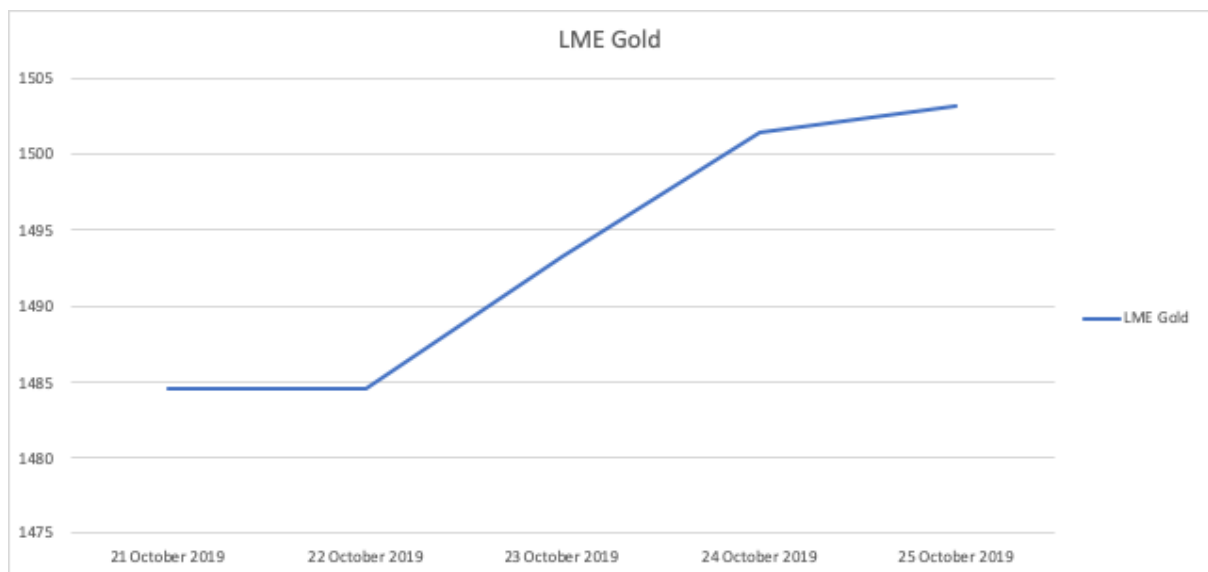
La gráfica principal se puede obtener desde el fichero, disponible en el repositorio de Github, AnalisisDatasetLME.ipynb.

Para esta representación gráfica, al ver que una representación conjunta de todos los datos tenía como inconveniente la gran diferencia de valor existente en los metales, se ha realizado una normalización de valores para pasarlos a una escala de 0 a 1, de forma que se puede evaluar, de un solo vistazo, la evolución de cada metal.



En el caso de querer tener información con valores económicos se ha añadido también al repositorio de Github la hoja de cálculo Gráfico\_dinámico.xlsx, en la cual se ha diseñado un gráfico conectado a un formulario dinámico, donde se puede filtrar por metal, o incluir varios metales, obteniendo diferentes representaciones gráficas.

Como ejemplo vamos a exponer la fluctuación del oro en el periodo analizado, con el gráfico obtenido a través de esta hoja de cálculo.



## 5.- Contenido

Los campos que incluye el Dataset son los siguientes:

- Date: Incluye la fecha de la captura.
- Product: Metal a valorar.
- Currency: Moneda por la cual se valora.
- Value: Valor económico en el mercado del producto a fecha.

Para este documento se han recogido los datos provenientes de la semana 43 de 2019, la cual incluye las valoraciones desde el día 21 al 27 de octubre.

El dataset que se incluye en el repositorio de Github tiene recopilados datos desde el 21 de octubre hasta la fecha en la que se entrega la práctica.

La ejecución ha sido diaria, el código va añadiendo al dataset los resultados diarios. Esta ejecución se puede realizar con una tarea programada desde el sistema operativo, para realizar la automatización de la misma.

## 6.- Agradecimientos

Agradecemos a The London Metal Exchange Private Unlimited Company<sup>1</sup>, propietarios de la web, ya que gracias a su servicio hemos podido realizar la práctica.

---

<sup>1</sup> <https://www.whois.com/whois/lme.com>

También a Juan Navalón, que nos explicó de primera mano el uso que hacían de esta información en su empresa de recuperación de metales ligeros y que nos ayudó a entender su utilidad en los negocios que operan en estos mercados.

## 7.- Inspiración

La bolsa de metales de Londres o *London Metal Exchange* es el mercado especializado en metales más importante del mundo. Los precios que se generan en sus plataformas de negociación son una referencia global y sirven de base para la compraventa física de metales.

Conocer las fluctuaciones e incluso ser capaces de predecirlas puede ser la clave para muchos negocios, cuyos costes de producción dependen directamente de la gestión del aprovisionamiento de estas materias primas. Saber en qué momento comprar o vender permite maximizar los beneficios de las empresas involucradas en esta cadena de suministro. En definitiva, el conocimiento de las fluctuaciones en los precios de los metales sirve para reducir el riesgo a las empresas que participan en el mercado de metales

## 8.- Licencia

### LICENCIA PARA EL DATASET

Hemos elegido la licencia de Creative Commons “**CC BY-SA 4.0 License**”. Esta licencia en concreto combina dos de las 4 limitaciones posibles: la condición de que se cite a los autores de forma expresa y la de que si se usa para crear otra obra, esta se distribuya bajo la misma licencia. Nuestro objetivo al elegir esta licencia es compartir el conocimiento y promover que otros compartan también sus creaciones, pero reconociendo siempre la labor de los autores.

[https://creativecommons.org/licenses/by-sa/4.0/deed.es\\_ES](https://creativecommons.org/licenses/by-sa/4.0/deed.es_ES)

### LICENCIA PARA EL CÓDIGO

Hemos elegido la licencia “**GNU General Public License v3.0**”. Esta licencia se usa habitualmente en los proyectos de software libre y código abierto. Permite que nuestro código se pueda usar, compartir y modificar, pero obliga al que lo utilice, modificado o ampliado, a seguir usando la misma licencia. De esta manera queremos promover el uso de código abierto para el beneficio de toda la comunidad.

<https://www.gnu.org/licenses/gpl-3.0.html>

## 9.- Código

```
import requests
from bs4 import BeautifulSoup
import csv
import os

def storeURL(str):

    page = requests.get(str)

    if page.status_code == 200:

        soup = BeautifulSoup(page.content)

        tag = soup.table
        head_tag = tag.thead
        body_tag = tag.tbody

        data = []
        separador = head_tag.get_text().find(":")
        moneda = head_tag.get_text()[0:separador].strip()
        fecha = head_tag.get_text()[separador + 1:].strip()

        trs = body_tag.findAll('tr')

        for tr in trs:
            metal = tr.th.get_text().replace("*", "").strip()
            valor = tr.td.get_text().strip()
            print("Fecha: {0}\nProducto: {1}\nMoneda: {2}\nCantidad:
{3}\n".format(fecha, metal, moneda, valor))
            data.append((fecha,metal,moneda,valor))

        with open("lme.csv", "a") as f:
            writer = csv.writer(f)
            if os.stat('lme.csv').st_size == 0:
                writer.writerow(["Date", "Product", "Currency", "Value"])
                for fecha,metal,moneda,valor in data:
                    writer.writerow([fecha,metal,moneda,valor])
            else:
                print ("Error code {}".format(page.status_code))

storeURL("https://www.lme.com")
```

## 10.- Dataset

Date	Product	Currency	Value
21 October 2019	LME Aluminium	US\$	1,735.50
21 October 2019	LME Copper	US\$	5,820.00
21 October 2019	LME Zinc	US\$	2,500.00
21 October 2019	LME Nickel	US\$	16,200.00
21 October 2019	LME Lead	US\$	2,202.00
21 October 2019	LME Tin	US\$	16,825.00
21 October 2019	LME Aluminium Alloy	US\$	1,310.00
21 October 2019	LME NASAAC	US\$	1,045.00
21 October 2019	LME Cobalt	US\$	35,500.00
21 October 2019	LME Gold	US\$	1,484.50
21 October 2019	LME Silver	US\$	17.575
21 October 2019	LME Steel Scrap	US\$	256.00
21 October 2019	LME Steel Rebar	US\$	417.00
22 October 2019	LME Aluminium	US\$	1,707.50
22 October 2019	LME Copper	US\$	5,794.00
22 October 2019	LME Zinc	US\$	2,503.00
22 October 2019	LME Nickel	US\$	16,025.00
22 October 2019	LME Lead	US\$	2,235.00
22 October 2019	LME Tin	US\$	16,730.00
22 October 2019	LME Aluminium Alloy	US\$	1,275.00

M2.851 - Tipología y ciclo de vida de los datos - Práctica 1  
 43287170C - Betancor Sánchez, Manuel - Aula 2  
 52774328T - Navalón Hernández, María Dolores - Aula 1

22 October 2019	LME NASAAC	US\$	1,055.00
22 October 2019	LME Cobalt	US\$	35,500.00
22 October 2019	LME Gold	US\$	1,484.50
22 October 2019	LME Silver	US\$	17.465
22 October 2019	LME Steel Scrap	US\$	253.50
22 October 2019	LME Steel Rebar	US\$	415.50
23 October 2019	LME Aluminium	US\$	1,712.00
23 October 2019	LME Copper	US\$	5,773.00
23 October 2019	LME Zinc	US\$	2,510.00
23 October 2019	LME Nickel	US\$	16,395.00
23 October 2019	LME Lead	US\$	2,236.00
23 October 2019	LME Tin	US\$	16,705.00
23 October 2019	LME Aluminium Alloy	US\$	1,325.00
23 October 2019	LME NASAAC	US\$	1,070.00
23 October 2019	LME Cobalt	US\$	35,500.00
23 October 2019	LME Gold	US\$	1,493.20
23 October 2019	LME Silver	US\$	17.560
23 October 2019	LME Steel Scrap	US\$	254.00
23 October 2019	LME Steel Rebar	US\$	413.00
24 October 2019	LME Aluminium	US\$	1,715.00
24 October 2019	LME Copper	US\$	5,869.00
24 October 2019	LME Zinc	US\$	2,518.50
24 October 2019	LME Nickel	US\$	16,950.00

M2.851 - Tipología y ciclo de vida de los datos - Práctica 1  
 43287170C - Betancor Sánchez, Manuel - Aula 2  
 52774328T - Navalón Hernández, María Dolores - Aula 1

24 October 2019	LME Lead	US\$	2,242.00
24 October 2019	LME Tin	US\$	16,700.00
24 October 2019	LME Aluminium Alloy	US\$	1,365.00
24 October 2019	LME NASAAC	US\$	1,050.00
24 October 2019	LME Cobalt	US\$	35,500.00
24 October 2019	LME Gold	US\$	1,501.50
24 October 2019	LME Silver	US\$	17.780
24 October 2019	LME Steel Scrap	US\$	253.00
24 October 2019	LME Steel Rebar	US\$	414.00
25 October 2019	LME Aluminium	US\$	1,710.50
25 October 2019	LME Copper	US\$	5,867.50
25 October 2019	LME Zinc	US\$	2,544.00
25 October 2019	LME Nickel	US\$	16,890.00
25 October 2019	LME Lead	US\$	2,244.50
25 October 2019	LME Tin	US\$	16,600.00
25 October 2019	LME Aluminium Alloy	US\$	1,370.00
25 October 2019	LME NASAAC	US\$	1,070.00
25 October 2019	LME Cobalt	US\$	35,500.00
25 October 2019	LME Gold	US\$	1,503.20
25 October 2019	LME Silver	US\$	17.945
25 October 2019	LME Steel Scrap	US\$	257.00
25 October 2019	LME Steel Rebar	US\$	416.00



M2.851 - Tipología y ciclo de vida de los datos - Práctica 1  
43287170C - Betancor Sánchez, Manuel - Aula 2  
52774328T - Navalón Hernández, María Dolores - Aula 1

## 11.- Contribuciones

Contribuciones	Firma
Investigación previa	MBS, MDNH
Redacción de las respuestas	MBS, MDNH
Desarrollo código	MBS, MDNH