

Práctica 1:

WEB SCRAPING

Rubén Herrera
Mariana Tolivar

Descripción e imagen identificativa

Como resultado de esta práctica, obtenemos un conjunto de datos, que nos muestra un resumen de la evolución de los distintos índices del IBEX 35, obtenidos de la web de la Bolsa de Madrid.

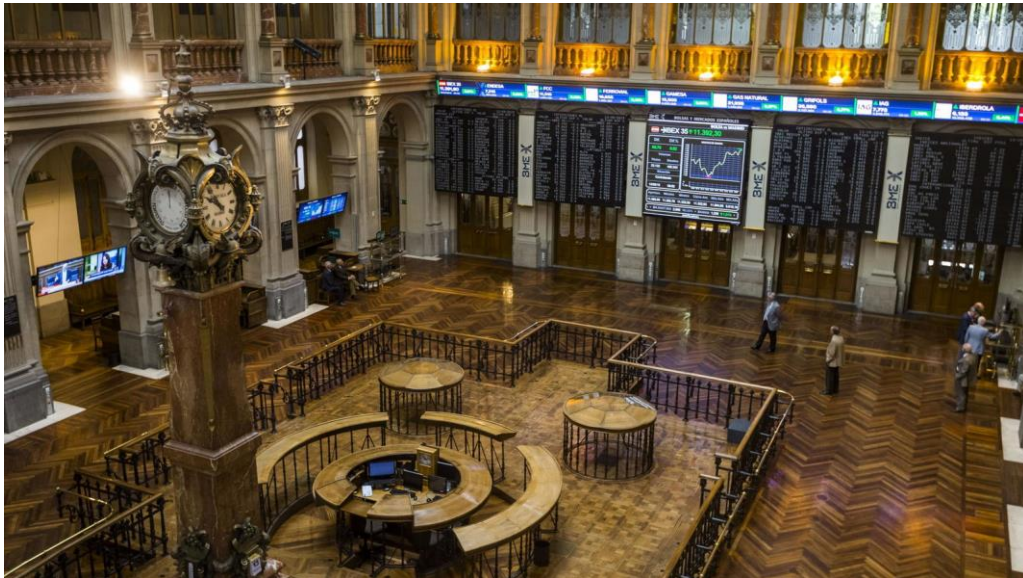


Figura 1: Imagen de la bolsa de Madrid

Fuente: https://blogs.elconfidencial.com/mercados/valor-anadido/2015-05-19/la-cruda-realidad-de-los-beneficios-de-la-bolsa-espanola_819494/

Contexto y contenido

Tal y como hemos comentado, lo que obtenemos es un conjunto de datos formado por los distintos índices del IBEX 35. La página web es actualizada cada 15 minutos, por lo que el código que obtenemos, obtendría un dataset distinto cada vez que realizáramos la petición (fuera del mismo rango de 15 minutos). De esta manera, se puede obtener un resumen del comportamiento de las cotizaciones de las empresas con mayor liquidez en la Bolsa de Madrid. Agrupa los 35 valores con mayor capitalización, negociación y frecuencia de cotización.

Los campos que obtenemos son los siguientes:

- 'Índice': Nos indica el nombre del índice del que obtenemos el valor
- 'Valor': Nos indica el valor del índice en esos 15 minutos en los que hemos obtenido el dataset
- 'Máximo': Nos indica el valor máximo que ha alcanzado el índice a lo largo del día
- 'Mínimo': Nos indica el valor mínimo que ha alcanzado el índice a lo largo del día

Inspiración

El conjunto de datos que obtenemos, puede ser útil sobre todo en el ámbito de la economía y las inversiones. Si realizamos el web scraping que hemos creado, durante distintos días, obtendremos la evolución de estos índices y de esta manera podemos hacernos una idea sobre la liquidez y sobre las cotizaciones de las empresas pertenecientes al IBEX 35.

Además, nos puede servir, para evaluar el riesgo de una determinada cartera de acciones y de esta manera hacernos mas fácil la creación de una cartera de inversiones.

Licencia

En este caso, la licencia escogida para nuestro conjunto de datos ha sido la conocida como **“Released Under CC0: Public Domain License”**.

Es decir, renunciamos a los derechos del dataset bajo las leyes de derechos autorales, en la medida permitida por la ley. Otorgamos el permiso de copiar, modificar, distribuir e interpretar la obra, sin necesidad de pedir ningún permiso, eximiéndonos de toda responsabilidad por los usos del dataset.

Agradecimientos

Los datos han sido recolectados de la página oficial de la Bolsa de Madrid.

Código fuente y dataset

Tanto el código como el dataset, se puede obtener en los ficheros correspondientes en este repositorio.

Recursos

- Subirats, L., Calvo, M. (2018). Web Scraping. Editorial UOC
- Tutorial de Github <https://guides.github.com/activities/hello-world>.
- Lawson, R. (2015). Web Scraping with Python. Packt Publishing Ltd. Chapter 2. Scraping the Data.
- https://creativecommons.org/publicdomain/zero/1.0/deed.es_ES
- Página web de la Bolsa de Madrid:
<https://www.bolsamadrid.es/esp/aspx/Portada/Portada.aspx>