

UOC – MÁSTER EN CIENCIA DE DATOS

TIPOLOGÍA Y CICLO DE VIDA DE DATOS – PRÁCTICA 1

CONTEXTO.

La economía mundial actual tiene una componente financiera muy elevada comparada con otras épocas. En este contexto las valoraciones de los diferentes activos financieros cobran gran importancia, tanta, que en no pocas ocasiones la evolución de los mercados financieros (entre ellos las bolsas de valores) es utilizada para evaluar la fortaleza económica de un país, así como, otros aspectos no necesariamente financieros que influyen directamente en la vida de las personas.

Se ha seleccionado la página web de la Bolsa de Madrid por ser la página de referencia en España de lo que comúnmente se conoce como la Bolsa. En la actualidad, las bolsas son mercados electrónicos, que en el caso español componen el SIBE, un sistema único del que forman parte las cuatro bolsas españolas (la de Madrid, Barcelona, Bilbao y Valencia), así como otras empresas afines.

TÍTULO.

Evolución de los precios de los valores del IBEX35.

DESCRIPCIÓN.

Se trata de un *dataset* en el que se recogen los precios de cierre diarios de los valores que cotizan en el Sistema de Interconexión Bursátil Español (SIBE) y forman parte del índice IBEX 35.

El IBEX 35 es un índice elaborado por BME (operador de los mercados de valores de España) que aglutina a las 35 empresas con más liquidez de la Bolsa española. Es un índice ponderado en función de la capitalización bursátil de cada uno de los valores que lo componen.

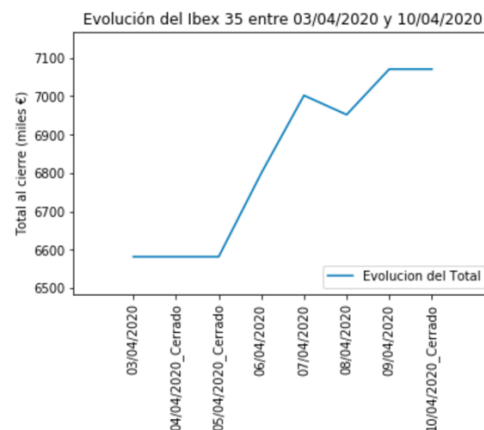
Todos los días al cierre de la sesión bursátil se publican en el diario de sesiones las cotizaciones de cierre de todos los valores. En este caso los datos han sido extraídos de la página web de la Bolsa de Madrid, pero los mismos también se publican en el resto de Bolsas españolas (Barcelona, Valencia y Bilbao).

La selección de las 35 empresas participantes en el índice puede variar. La revisión de su composición se realiza dos veces al año y es llevada a cabo por el Comité Asesor Técnico (CAT). De decidirse algún cambio, éste es efectivo el primer día hábil de julio o el primer día hábil de enero según proceda, por lo que el índice sólo varía dos veces al año, salvo casos excepcionales.

UOC – MÁSTER EN CIENCIA DE DATOS

TIPOLOGÍA Y CICLO DE VIDA DE DATOS – PRÁCTICA 1

REPRESENTACIÓN GRÁFICA.



IBEX 35®								
Nombre	Anterior	Último	% Dif.	Máximo	Mínimo	Fecha	Hora	% Dif. Año 2020
▲ IBEX 35®	6.951,80	7.070,80	1,71	7.116,80	6.918,70	09/04/2020	17:38:00	-25,96

Nombre	Últ.	% Dif.	Máx.	Min.	Volumen	Efectivo (miles €)	Fecha	Hora
▲ ACCIONA	97,2500	0,15	99,5000	96,2500	70,921	6.911,93	09/04/2020	Cierre
▲ ACERINOX	6,8340	4,34	6,8340	6,6480	719,099	4.862,77	09/04/2020	Cierre
▲ ACS	20,0000	1,55	20,5200	19,5600	2.746,445	55.130,41	09/04/2020	Cierre
▲ AENA	121,5000	5,84	121,5000	118,0000	449,733	54.419,56	09/04/2020	Cierre
▲ AMADEUS	49,7000	2,75	49,8000	47,9100	2.282,009	111.364,74	09/04/2020	Cierre
▲ ARCELORMIT.	9,5940	1,58	9,8570	9,2020	1.275,947	12.117,06	09/04/2020	Cierre
▼ B.SANTANDER	2,1650	-2,24	2,2645	2,1650	92.426,134	202.552,48	09/04/2020	Cierre
▼ BA.SABADELL	0,4436	-0,36	0,4599	0,4388	28.364,880	11.736,63	09/04/2020	Cierre
▲ BANKIA	1,0370	2,47	1,0390	0,9950	13.719,986	14.020,30	09/04/2020	Cierre
▲ BANKINTER	3,5040	6,18	3,5380	3,3410	3.463,636	11.993,88	09/04/2020	Cierre

CONTENIDO.

El *dataset* obtenido, recoge los datos de cierre de los valores que componen el IBEX35 al cierre de cada sesión. La primera de las columnas recoge el nombre de la acción / empresa y en las siguientes las diferentes fechas de sesión.

Para simplificar el formato de almacenaje de los datos se ha optado por considerar como fijas las 35 empresas de referencia, en caso de que el listado cambiase, se debería empezar un nuevo listado o realizar modificaciones en el código.

Adicionalmente, se ha incluido también el valor del índice (IBEX35) en cada fecha.

El periodo de tiempo abarcado por el *dataset* va desde la sesión del día 3 de abril de 2020 a la del 10 de abril de 2020.

(Lo ideal sería que el programa siguiera acumulando datos hasta que cambiaran los miembros del IBEX35, pero de momento, entendemos que una recogida de 8 días, ya incluye todo tipo de eventos normales).

La recogida de los datos que conforman el *dataset* se ha realizado mediante *scraping* de la página web de la Bolsa de Madrid donde se publica diariamente los datos de cada sesión del IBEX35.

UOC – MÁSTER EN CIENCIA DE DATOS

TIPOLOGÍA Y CICLO DE VIDA DE DATOS – PRÁCTICA 1

FUNCIONAMIENTO DEL PROGRAMA:

Este programa recoge los precios de cierre diarios del IBEX35 de la web:

<http://www.bolsamadrid.es/esp/aspx/Mercados/Precios.aspx?indice=ESI100000000>

cada día a las 6 de la tarde hora española, desde el momento en que se active hasta la fecha indicada al lanzarlo: *lanzarScraping('10/04/2020')*.

La fecha debe ser introducida entre comillas y con el formato dd/mm/aaaa.

A cada recarga, el programa devolverá un documento *CSV* con todos los datos recopilados hasta la fecha y una gráfica mostrando la evolución.

Una vez recolectados, los *string* originales con "." para separar los miles y "," para separar los decimales, fueron transformados en *floats* con "." para separar decimales. El resultado es convertido en una lista de las empresas y sus valores diarios preparado para ser incluidos dentro de un *dataframe* de pandas.

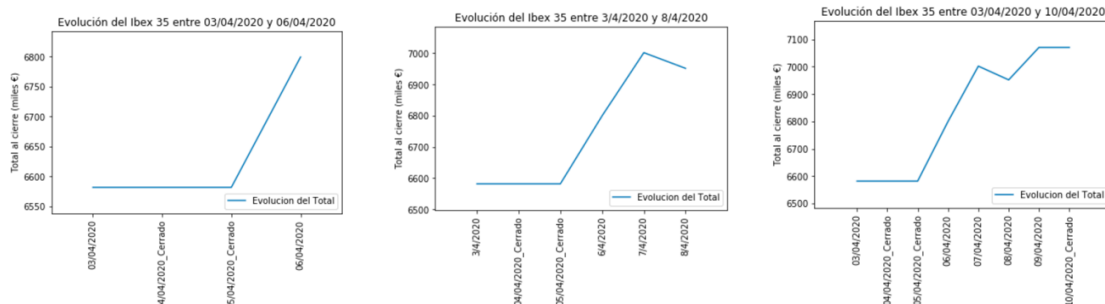
Una vez recopilados los datos, si no existe el documento para guardarlos lo crea y empieza una tabla desde cero. Si ya existía un documento previo, lo abre y lo edita, añadiendo una nueva columna con los datos del día.

Si al hacer la carga, a las 6 de la tarde, la fecha de cierre no ha cambiado respecto al día anterior, se entiende que la bolsa no ha abierto ese día, y aunque se cargan los datos igualmente, se anota en nombre de la columna como fecha junto con la anotación “Cerrada” (Esto sucede los fines de semana, pero también puede suceder en circunstancias especiales como vacaciones o catástrofes).

Una vez realizada esta acción el programa, que sigue abierto, se pone en pausa hasta que 24h después continúa con la ejecución del bucle desde el principio.

Cada vez que se ejecuta, el programa informa de la hora y muestra una gráfica en pantalla con la evolución del índice.

A continuación se muestran los resultados gráficos de varios días



UOC – MÁSTER EN CIENCIA DE DATOS

TIPOLOGÍA Y CICLO DE VIDA DE DATOS – PRÁCTICA 1

AGRADECIMIENTOS.

El propietario de los datos es BME (Bolsas y mercados españoles, S.A.) que es la empresa que gestiona todas las Bolsas y mercado financieros españoles. Así mismo es la responsable de la gestión de las órdenes de compra y ventas electrónicas que se cruzan en el mercado dando lugar a los precios de la acciones.

INSPIRACIÓN.

Con los datos recopilados mediante *scraping* se crea una base de datos que permitirá analizar la evolución del valor de las empresas más importantes del país de forma que se pueda utilizar en estudios y análisis de diversa índole.

Por ejemplo, permitirá comparar la evolución en bolsa del valor de diversas compañías como IAG o MELIÁ HOTELS en relación con el desarrollo de pandemia del Coronavirus.

Por otra parte, también podría usarse para seguir cómo evolucionan las distintas compañías en bolsa una vez superada la pandemia y dé comienzo la recuperación de la economía.

IDEA DE MEJORA DEL PROGRAMA.

Puesto que se tienen los datos diarios de todas las empresas del ibex35, el programa debería dar la opción de representar gráficamente la evolución de cualquier empresa o de varias de ellas para poder comparar visualmente.

LICENCIA.

En el caso de nuestro software optaríamos por una licencia de Creative Commons tipo CC BY-NC-SA.

Al elegir esta licencia permitimos a los usuarios copiar, distribuir, exhibir y representar la obra y hacer obras derivadas de la misma siempre que se cite la autoría (BY). Cualquier distribución de la obra por parte de los usuarios de la licencia tendrá que hacerse bajo una licencia idéntica a la elegida (SA).

Por otra parte, no se autorizaría el uso con fines comerciales salvo que se nos solicitara permiso para ello (NC).

Contribuciones	Firma
Investigación previa	XRV, GMD
Redacción de las respuestas	XRV, GMD
Desarrollo código	XRV, GMD