

# PRÁCTICA 1

**Pablo Rivas Castellanos**

**Diego Alberto López Herrera**

Tipología y ciclo de vida de los datos  
Máster universitario de Ciencia de Datos

12/Mar/2021

## Índice

1.	Contexto.....	2
2.	Título del dataset.....	2
3.	Descripción del dataset.....	2
4.	Representación gráfica .....	2
5.	Contenido.....	3
6.	Agradecimientos.....	5
7.	Inspiración.....	5
8.	Licencia.....	5

## 1. Contexto

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

Se ha escogido el sitio web <https://www.bolsamadrid.es/> por aportar gran cantidad de información detallada sobre valores cotizados en la bolsa de Madrid. Este sitio web no establece ninguna limitación en cuanto a la descarga automática de datos mediante *web scraping* o técnicas similares (se ha realizado esta revisión a 12 de marzo de 2021).

Se pretende, por un lado, obtener información general sobre las empresas cotizadas: nombre, sector, mercado, índices. Por otra parte, se recopilarán datos detallados sobre la cotización a nivel diario: valores de cierre, referencia, último, máximo, mínimo y medio, así como el volumen y efectivo en cuanto a los títulos negociados.

Como posible mejora futura, se podrá establecer una actualización de los datos con la frecuencia conveniente para mantener un listado de empresas actualizado y un histórico completo de cotización diaria.

## 2. Título del dataset

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

Bolsa de Madrid: empresas y detalles de cotización diaria

## 3. Descripción del dataset

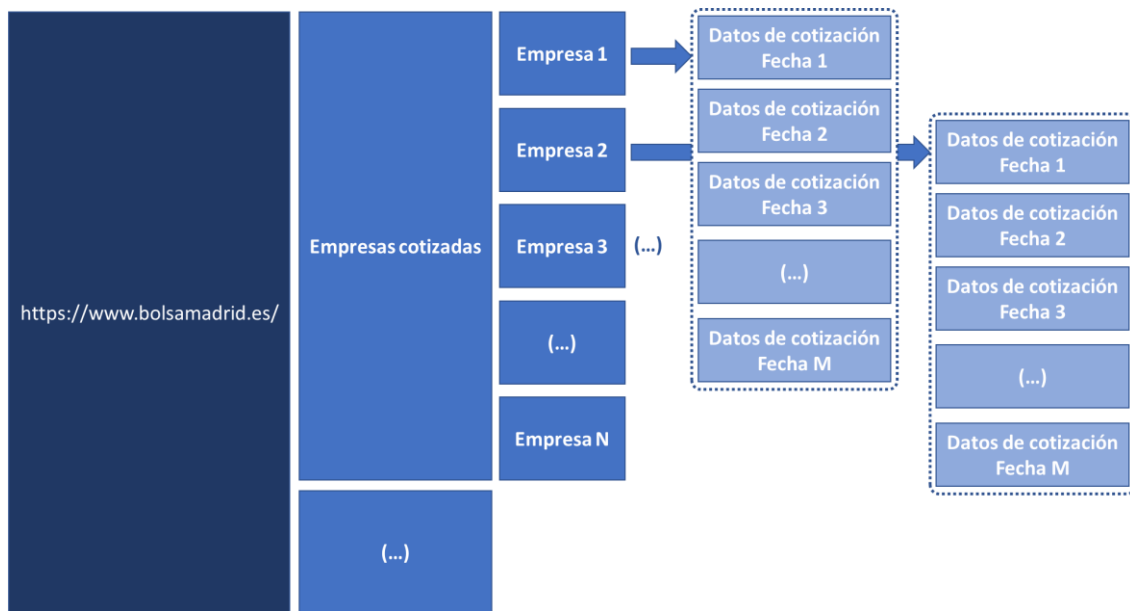
Descripción del dataset. Desarrollar una descripción breve del conjunto de datos que se ha extraído (es necesario que esta descripción tenga sentido con el título elegido).

Se dispone del listado de empresas cotizadas de la bolsa de Madrid, aportando detalles generales sobre ellas (sector, índice al que pertenecen, etc.) y el detalle de cotización diaria con varias métricas relevantes (valor de cierre, volumen, valor medio, etc.). El conjunto de datos se compone de dos tablas en las que el elemento común es el ISIN (<https://www.cnmv.es/portal/ANCV/CodigoISIN.aspx>), un identificador unívoco de los valores mobiliarios a nivel internacional, a través del cual se pueden relacionar los datos de ambas tablas.

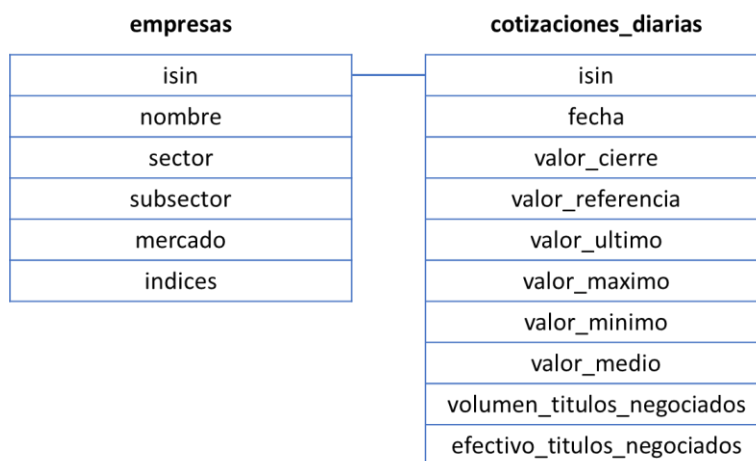
## 4. Representación gráfica

Representación gráfica. Presentar esquema o diagrama que identifique el dataset visualmente y el proyecto elegido.

La información se recoge tal y como se muestra en el siguiente esquema:



En cuanto al dataset, en la siguiente figura se detallan las tablas y campos que los conforman, junto con las relaciones correspondientes:



## 5. Contenido

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

En la siguiente tabla se incluye todo el detalle de los campos del dataset:

Tabla	Campo	Descripción	Tipo	Ejemplo
<b>empresas</b>	isin	ISIN. Identificador unívoco de los valores mobiliarios a	Texto	<i>ES0125220311</i>

		nivel internacional.		
<b>Empresas</b>	nombre	Nombre de la empresa	Texto	<i>ACCIONA, S.A.</i>
<b>Empresas</b>	sector	Sector al que pertenece la empresa	Texto	<i>Mat.Basicos, Industria y Construcción</i>
<b>Empresas</b>	subsector	Subsector al que pertenece la empresa	Texto	<i>Construcción</i>
<b>Empresas</b>	mercado	Mercado en el que cotiza la empresa	Texto	<i>Mercado Continuo</i>
<b>Empresas</b>	indices	Índice bursátil al que pertenece la compañía	Texto	<i>IBEX 35®</i>
<b>cotizaciones_diarias</b>	isin	ISIN. Identificador unívoco de los valores mobiliarios a nivel internacional.	Texto	<i>ES0125220311</i>
<b>cotizaciones_diarias</b>	fecha	Fecha de la cotización	Fecha	<i>02/03/2021</i>
<b>cotizaciones_diarias</b>	valor_cierre	Valor de cierre de cotización	Numérico (decimal)	<i>133,1000</i>
<b>cotizaciones_diarias</b>	valor_referencia	Valor de referencia de cotización	Numérico (decimal)	<i>135,0000</i>
<b>cotizaciones_diarias</b>	valor_ultimo	Valor último de cotización	Numérico (decimal)	<i>133,1000</i>
<b>cotizaciones_diarias</b>	valor_maximo	Valor máximo de cotización	Numérico (decimal)	<i>135,3000</i>
<b>cotizaciones_diarias</b>	valor_minimo	Valor mínimo de cotización	Numérico (decimal)	<i>132,7000</i>
<b>cotizaciones_diarias</b>	valor_medio	Valor medio de cotización	Numérico (decimal)	<i>133,7140</i>
<b>cotizaciones_diarias</b>	volumen_titulos_negociados	Volumen de títulos negociados	Numérico (entero)	<i>71.918</i>
<b>cotizaciones_diarias</b>	efectivo_titulos_negociados	Volumen de efectivo en títulos negociados	Numérico (decimal)	<i>9.616.442,00</i>

En cuanto a los sitios webs de obtención de los datos, se tienen los siguientes para cada dataset:

- empresas: <https://www.bolsamadrid.es/esp/asp/Empresas/Empresas.aspx> , que requiere navegación vía *javascript* para visualizar todas las compañías en base a la letra inicial de su nombre (se agrupan alfabéticamente).
- cotizaciones\_diarias: [https://www.bolsamadrid.es/esp/asp/Empresas/InfHistorica.aspx?ISIN=\[ISIN\]](https://www.bolsamadrid.es/esp/asp/Empresas/InfHistorica.aspx?ISIN=[ISIN]) , donde [ISIN] es el ISIN de la empresa cotizada. La información que proporciona por defecto es la de los últimos 30 días.

## 6. Agradecimientos

Agradecimientos. Presentar al propietario del conjunto de datos. Es necesario incluir citas de análisis anteriores o, en caso de no haberlas, justificar esta búsqueda con análisis similares.

## 7. Inspiración

Inspiración. Explique por qué es interesante este conjunto de datos y qué preguntas se pretenden responder. Es necesario comparar con los análisis anteriores presentados en el apartado 6.

## 8. Licencia

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

Se ha elegido la licencia “CC BY-NC-SA 4.0 License” para que se requiera citación de los autores, solo se permita el uso no comercial y todas las adaptaciones se distribuyan bajo los mismos términos. De esta forma, se pretende que quien use el material no persiga beneficios económicos y, en caso de que se enriquezca / adapte el material, pueda accesible por cualquier persona / compañía con las mismas condiciones que el material compartido.