

Práctica 1: Web scraping

Asignatura: *M2.851 - Tipología y ciclo de vida de los datos aula 1*

1. Contexto:

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

Después de que en 2008 estallara la burbuja inmobiliaria en España dejando en evidencia el gran endeudamiento de familias y empresas, hemos considerado interesante poder analizar el comportamiento de la morosidad.

Si un acreedor impaga un préstamo de forma reiterada el deudor podrá recurrir a la justicia para recobrar el importe. La justicia podrá llevar a subasta los bienes que garantizaban en préstamo con el fin de que el acreedor pueda saldar su deuda. Los bienes subastados más conocidos son las viviendas, pero también pueden ser vehículos y otros bienes muebles.

Con esto queremos saber dónde se concentran los morosos, analizar las entidades acreedoras, la procedencia de la subasta, los precios con los que salen, como se cierran,...

El sitio web elegido contiene todas las subastas desde 2016 hechas en España con información sobre el acreedor, la autoridad gestora, información general de la subasta, los importes de apertura y condiciones de puja, la puja final,... En definitiva, todo lo necesario para analizar las ejecuciones de garantías mediante procedimiento judicial o administrativo en España.

2. Definir un título para el dataset:

Elegir un título que sea descriptivo.

Subastas judiciales y administrativas

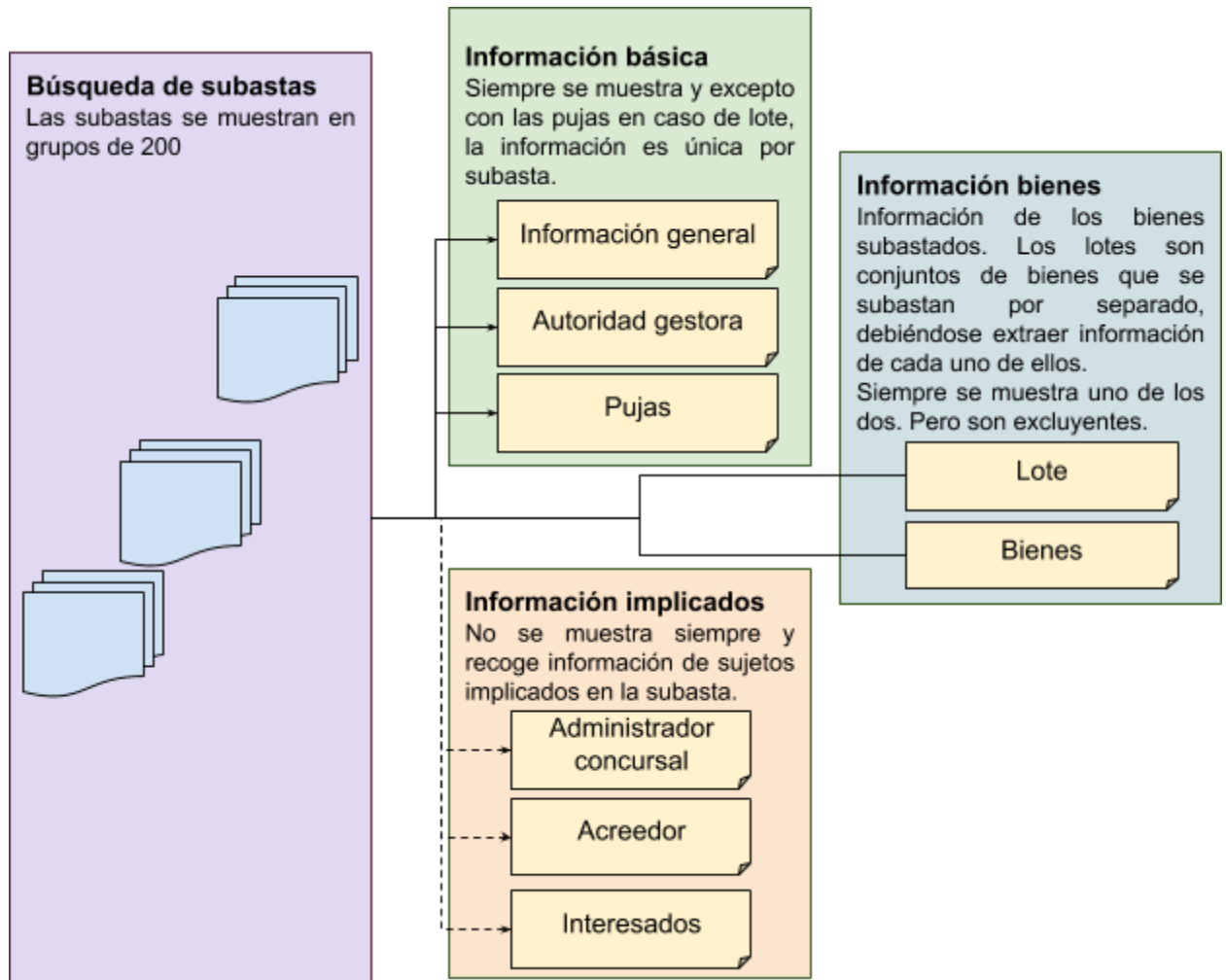
3. Descripción del dataset:

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

4. Representación gráfica:

Presentar una imagen o esquema que identifique el dataset visualmente

La web tiene la siguiente estructura:



La información que contiene cada una de las subastas es diferente por lo que se debe ser abrir el código a todas las casuísticas.

A nivel gráfico hay que destacar las siguientes ventanas:

- *Resultado búsqueda:*

Subastas. Búsqueda avanzada

Búsqueda

Resultados

Resultados 1 a 50 de 470

1 2 3 4 5 6 7 8 9 10 [Pág. siguiente ▶](#)

SUBASTA SUB-JA-2018-93973

JUZGADO 1ª INST E INSTRUCC. 1 - SEPULVEDA

Expediente: 0025/17

Estado: Celebrándose - [Conclusión prevista: 24/04/2019 a las 18:00:00]

APARTAMENTO EN CEREZO DE ARRIBA (SEGOVIA)

 [Más...](#)

SUBASTA SUB-AT-2019-19R4786001002

U.R. SUBASTAS CASTILLA-LEON - VALLADOLID

Estado: Celebrándose - [Conclusión prevista: 22/04/2019 a las 18:00:00]

VIVIENDA , PZ MAYOR-MORALEJA 1. 40233 - OLOMBRADA (SEGOVIA)

 [Más...](#)

SUBASTA SUB-AT-2019-19R4786001020

U.R. SUBASTAS CASTILLA-LEON - VALLADOLID

Estado: Celebrándose - [Conclusión prevista: 17/04/2019 a las 18:00:00]

- *Estructura de la información de cada subasta*

Subasta SUB-JC-2018-113482

Información general

Autoridad gestora

Lotes

Administrador concursal

Pujas



LA SUBASTA HA SIDO FINALIZADA POR LA AUTORIDAD GESTORA.



La subasta contiene varios lotes que se subastan de forma separada.

Lote 1

Lote 2

Lote 3

Lote 4

Lote 5

Lote 6

Lote 7

Lote 1

GARAJE

Datos relacionados con la subasta del lote 1

Valor Subasta	4.477,96 €
Importe del depósito	223,89 €
Puja mínima	1.343,38 €
Tramos entre pujas	150,00 €

5. Contenido:

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

6. Agradecimientos:

Presentar al propietario del conjunto de datos. Es necesario incluir citas de investigación o análisis anteriores (si los hay).

7. Inspiración:

Explique por qué es interesante este conjunto de datos y qué preguntas se pretenden responder.

8. 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*