

# Dataset Plants Class:

*Nadia Nathaly Sánchez Pozo*

*5 de noviembre de 2019*

## Descripción

El conjunto de datos generado como parte de esta práctica reúne diferentes características de una variedad de plantas. Algunas de las variables que se recogen en el conjunto de datos son categoría, usos, longevidad. Además se tiene un conjunto de fotos de cada planta.

## Contexto

Los datos han sido recolectados de las plantas disponibles en un vivero de Athens, Ohio. Dicho vivero ofrece variedades de plantas de hierbas comunes y exóticas, incluidas plantas medicinales, culinarias, ceremoniales, aromáticas, mariposas, de fibra y colorantes. Cada planta cuenta con características específicas como son categoría, usos, longevidad, condiciones de iluminación, precio, foto de la planta. En este sitio encontramos un gran número de plantas y características con las que se pueden realizar futuras investigaciones.

## Imagen identificativa



## Contenido

Para cada planta, el cual se corresponde con un registro en el conjunto de datos, se recogen las siguientes características:

- **Name:** nombre científico de la planta.
- **Price:** valor de la planta en el vivero.
- **Category:** categoría correspondiente a la planta.
- **Lighting\_Conditions:** condición de luz.
- **Longevity:** tiempo de vida de la planta.
- **Average\_Height:** altura media.
- **imagen:** link de acceso a la foto de la planta.
- **Uses:** usos de la planta.

A (Aromatic) Aromático -La planta es adecuada para su uso en bolsitas y popurrí.

B (Butterflies) Mariposas - La planta atrae mariposas.

c (Culinary) Culinaria- La planta se utiliza en la preparación y aromatización de alimentos, o como té.

D (Dyes) Tintes - Útil en la fabricación de tintes.

E (Everlasting) Eterno - Bueno para coronas y otros arreglos florales.

H (Honey) Miel - La planta es particularmente atractiva para las abejas.

T (Traditional) Tradicional - Tiene usos ceremoniales tradicionales.

M (Medicinal) Medicinal -Se cree o se sabe que la planta tiene componentes medicinales útiles.

N (Native) Nativo -Nativo del este de EE. UU.

O (Ornamental) Ornamental -La planta es particularmente decorativa.

R (Repellent) Repelente- Repelente de insectos u otras plagas.

Los autores de la web *Companion Plants* llevan ofreciendo sus productos al mercado desde el año 1982.

## Agradecimientos

Los datos han sido recolectados desde el catalogo de datos online Companion Plants. Para ello, se ha hecho uso del lenguaje de programación Python y de técnicas de *Web Scraping*, haciendo uso de la librería *Selenium* para automatizar el navegador, y extraer la información de la web.

## Inspiración

El conjunto de datos de clases de plantas podría ser utilizado en diferentes ámbitos. Por ejemplo en el área comercial para el análisis de mercado del valor comercial de plantas. También puede ser usado en periodismo al tener una variedad de plantas con fotos, usos y otras características.

También podría ser de gran utilidad en el campo de la *minería de datos*, a la hora de elaborar modelos de clasificación.

## Licencia

La licencia escogida para la publicación de este conjunto de datos ha sido **CC BY-NC-SA 4.0** . Ya que este conjunto de datos cumple con los siguientes puntos contemplados en una licencia **CC BY-NC-SA 4.0** .

Proporciona la informacion de créditos correspondiente: nombre del creador, un aviso de licencia, enlace al material.

## Código fuente y dataset

Tanto el código fuente escrito para la extracción de datos como el dataset generado pueden ser accedidos a través de este enlace.

## Recursos

1. Subirats, L., Calvo, M. (2018). Web Scraping. Editorial UOC.
2. Masip, D. El lenguaje Python. Editorial UOC.