

[About](#) [Citation Policy](#) [Donate a Data Set](#) [Contact](#)
☒ Repository ☐ Web [View ALL Data Sets](#)

Plants Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Data has been extracted from the USDA plants database. It contains all plants (species and genera) in the database and the states of USA and Canada where they occur.



Data Set Characteristics:	Multivariate	Number of Instances:	22632	Area:	Life
Attribute Characteristics:	Categorical	Number of Attributes:	70	Date Donated	2008-12-31
Associated Tasks:	Clustering	Missing Values?	Yes	Number of Web Hits:	141367

Source:

Original source:

USDA plants database: <http://plants.usda.gov/index.html>

Extracted and encoded by W. Hämmäläinen, Department of Computer Science, University of Helsinki, Finland. [whamalai '@' cs.helsinki.fi](mailto:whamalai@cs.helsinki.fi)

Data Set Information:

The data is in the transactional form. It contains the Latin names (species or genus) and state abbreviations.

Attribute Information:

Each row contains a Latin name (species or genus) and a list of state abbreviations.

Relevant Papers:

Hämäläinen, W. and Nykänen, M.: Efficient discovery of statistically significant association rules. Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), pp. 203-212. IEEE Computer Society 2008.

Citation Request:

Even if the data is processed, it is good to give a reference to the original source:

USDA, NRCS. 2008. The PLANTS Database ([Web Link](#)), 31 December 2008). National Plant Data Center, Baton Rouge, LA 70874-4490 USA.

Supported By:



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)