## Some notes to the iris dataset

**Hinweis**: Nachfolgende Ausführung bezieht sich auf das *Scikit Learn*-basierte Laden des Datasets.

Grundsätzlich gestaltet sich der Aufbau des Datasets so, dass jede Zeile für genau eine Blume steht (siehe Abbildung 1). Doch wie gelangt man zu dieser Darstellung?

Fisher's Iris Data [hide]					
Dataset Order +	Sepal length +	Sepal width +	Petal length +	Petal width +	Species +
1	5.1	3.5	1.4	0.2	I. setosa
2	4.9	3.0	1.4	0.2	I. setosa
3	4.7	3.2	1.3	0.2	I. setosa
4	4.6	3.1	1.5	0.2	I. setosa
5	5.0	3.6	1.4	0.3	I. setosa
6	5.4	3.9	1.7	0.4	I. setosa
7	4.6	3.4	1.4	0.3	I. setosa
8	5.0	3.4	1.5	0.2	I. setosa
9	4.4	2.9	1.4	0.2	I. setosa
10	4.9	3.1	1.5	0.1	I. setosa

Abbildung 1: Iris Dataset (Auszug)

Führt man folgenden Code aus, ist ersichtlich, dass das Ergebnis der *Dictionary*-Form entspricht:

Es ist ersichtlich, dass zumindest der Key data existiert. Über diesen sind die Daten (siehe Abbildung 1, Spalten Sepal length,..., Petal width) verfügbar. Sämtliche Keys erfährt man, wenn man nach unten scrollt oder die Methode keys () anwendet:

```
dict_keys(['data', 'target', 'target_names', 'DESCR', 'feature_names',
'filename'])
```

Die jeweilige Spezies ist dem Key target und der dazugehörige Klartextname dem Key target\_name zu entnehmen. Spaltennamen liefert feature\_names, eine allgemeine Beschreibung DESCR.