

DR. ALVIN'S PUBLICATIONS

CONVERTING JSON TO CSV FROM SCIKIT LEARN DATASETS

IN PYTHON BY DR. ALVIN ANG



COPYRIGHTED BY DR ALVIN ANG
WWW.ALVINANG.SG



Converting JSON to CSV from Scikit Learn Datasets

in Python by Dr. Alvin Ang



Dr. Alvin Ang

1 min read · 23 hours ago



https://www.alvinang.sg/s/Converting_JSON_to_CSV_from_Scikit_Learn_Datasets_by_Dr_Alvin_Ang.ipynb

7.1. Toy datasets

scikit-learn comes with a few small standard datasets that do not require to download any file from some external...

scikit-learn.org

7.2. Real world datasets

scikit-learn provides tools to load larger datasets, downloading them if necessary. They can be loaded using the...

scikit-learn.org

1. Import and Load Dataset

[4.8, 3.4, 1.6, 0.2],
[4.8, 3. , 1.4, 0.1],
[4.3, 3. , 1.1, 0.1],
[5.8, 4. , 1.2, 0.2],
[5.7, 4.4, 1.5, 0.4],
[5.4, 3.9, 1.3, 0.4],
[5.1, 3.5, 1.4, 0.3],
[5.7, 3.8, 1.7, 0.3],
[5.1, 3.8, 1.5, 0.3],
[5.4, 3.4, 1.7, 0.2],
[5.1, 3.7, 1.5, 0.4],
[4.6, 3.6, 1. , 0.2],
[5.1, 3.3, 1.7, 0.5],
[4.8, 3.4, 1.9, 0.2],
[5. , 3. , 1.6, 0.2],
[5. , 3.4, 1.6, 0.4],
[5.2, 3.5, 1.5, 0.2],
[5.2, 3.4, 1.4, 0.2],
[4.7, 3.2, 1.6, 0.2],
[4.8, 3.1, 1.6, 0.2],
[5.4, 3.4, 1.5, 0.4],
[5.2, 4.1, 1.5, 0.1],
[5.5, 4.2, 1.4, 0.2],
[4.9, 3.1, 1.5, 0.2],
[5. , 3.2, 1.2, 0.2],
[5.5, 3.5, 1.3, 0.2],
[4.9, 3.6, 1.4, 0.1],
[4.4, 3. , 1.3, 0.2],
[5.1, 3.4, 1.5, 0.2],
[5. , 3.5, 1.3, 0.3],
[4.5, 2.3, 1.3, 0.3],
[4.4, 3.2, 1.3, 0.2],
[5. , 3.5, 1.6, 0.6],
[5.1, 3.8, 1.9, 0.4],
[4.8, 3. , 1.4, 0.3],
[5.1, 3.8, 1.6, 0.2],
[4.6, 3.2, 1.4, 0.2],
[5.3, 3.7, 1.5, 0.2],
[5. , 3.3, 1.4, 0.2],
[7. , 3.2, 4.7, 1.4],
[6.4, 3.2, 4.5, 1.5],
[6.9, 3.1, 4.9, 1.5],
[5.5, 2.3, 4. , 1.3],
[6.5, 2.8, 4.6, 1.5],
[5.7, 2.8, 4.5, 1.3],
[6.3, 3.3, 4.7, 1.6],
[4.9, 2.4, 3.3, 1.],
[6.6, 2.9, 4.6, 1.3],
[5.2, 2.7, 3.9, 1.4],
[5. , 2. , 3.5, 1.],
[5.9, 3. , 4.2, 1.5],
[6. , 2.2, 4. , 1.],
[6.1, 2.9, 4.7, 1.4],
[5.6, 2.9, 3.6, 1.3],
[6.7, 3.1, 4.4, 1.4],
[5.6, 3. , 4.5, 1.5],
[5.8, 2.7, 4.1, 1.],
[6.2, 2.2, 4.5, 1.5],
[5.6, 2.5, 3.9, 1.1],
[5.9, 3.2, 4.8, 1.8],
[6.1, 2.8, 4. , 1.3],
[6.3, 2.5, 4.9, 1.5],
[6.1, 2.8, 4.7, 1.2],
[6.4, 2.9, 4.3, 1.3],
[6.6, 3. , 4.4, 1.4],
[6.8, 2.8, 4.8, 1.4],
[6.7, 3. , 5. , 1.7],
[6. , 2.9, 4.5, 1.5],
[5.7, 2.6, 3.5, 1.],
[5.5, 2.4, 3.8, 1.1],
[5.5, 2.4, 3.7, 1.],
[5.8, 2.7, 3.9, 1.2],
[6. , 2.7, 5.1, 1.6],
[5.4, 3. , 4.5, 1.5],
[6. , 3.4, 4.5, 1.6],
[6.7, 3.1, 4.7, 1.5],
[6.3, 2.3, 4.4, 1.3],


```
[5.6, 3. , 4.1, 1.3],
[5.5, 2.5, 4. , 1.3],
[5.5, 2.6, 4.4, 1.2],
[6.1, 3. , 4.6, 1.4],
[5.8, 2.6, 4. , 1.2],
[5. , 2.3, 3.3, 1. ],
[5.6, 2.7, 4.2, 1.3],
[5.7, 3. , 4.2, 1.2],
[5.7, 2.9, 4.2, 1.3],
[6.2, 2.9, 4.3, 1.3],
[5.1, 2.5, 3. , 1.1],
[5.7, 2.8, 4.1, 1.3],
[6.3, 3.3, 6. , 2.5],
[5.8, 2.7, 5.1, 1.9],
[7.1, 3. , 5.9, 2.1],
[6.3, 2.9, 5.6, 1.8],
[6.5, 3. , 5.8, 2.2],
[7.6, 3. , 6.6, 2.1],
[4.9, 2.5, 4.5, 1.7],
[7.3, 2.9, 6.3, 1.8],
[6.7, 2.5, 5.8, 1.8],
[7.2, 3.6, 6.1, 2.5],
[6.5, 3.2, 5.1, 2. ],
[6.4, 2.7, 5.3, 1.9],
[6.8, 3. , 5.5, 2.1],
[5.7, 2.5, 5. , 2. ],
[5.8, 2.8, 5.1, 2.4],
[6.4, 3.2, 5.3, 2.3],
[6.5, 3. , 5.5, 1.8],
[7.7, 3.8, 6.7, 2.2],
[7.7, 2.6, 6.9, 2.3],
[6. , 2.2, 5. , 1.5],
[6.9, 3.2, 5.7, 2.3],
[5.6, 2.8, 4.9, 2. ],
[7.7, 2.8, 6.7, 2. ],
[6.3, 2.7, 4.9, 1.8],
[6.7, 3.3, 5.7, 2.1],
[7.2, 3.2, 6. , 1.8],
[6.2, 2.8, 4.8, 1.8],
[6.1, 3. , 4.9, 1.8],
[6.4, 2.8, 5.6, 2.1],
[7.2, 3. , 5.8, 1.6],
[7.4, 2.8, 6.1, 1.9],
[7.9, 3.8, 6.4, 2. ],
[6.4, 2.8, 5.6, 2.2],
[6.3, 2.8, 5.1, 1.5],
[6.1, 2.6, 5.6, 1.4],
[7.7, 3. , 6.1, 2.3],
[6.3, 3.4, 5.6, 2.4],
[6.4, 3.1, 5.5, 1.8],
[6. , 3. , 4.8, 1.8],
[6.9, 3.1, 5.4, 2.1],
[6.7, 3.1, 5.6, 2.4],
[6.9, 3.1, 5.1, 2.3],
[5.8, 2.7, 5.1, 1.9],
[6.8, 3.2, 5.9, 2.3],
[6.7, 3.3, 5.7, 2.5],
[6.7, 3. , 5.2, 2.3],
[6.3, 2.5, 5. , 1.9],
[6.5, 3. , 5.2, 2. ],
[6.2, 3.4, 5.4, 2.3],
[5.9, 3. , 5.1, 1.8]],
'target': array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2]),
'frame': None,
'target_names': array(['setosa', 'versicolor', 'virginica'], dtype='<U10'),
'DESCR': '.. _iris_dataset:\n\nIris plants dataset\n-----\n\n**D
'feature_names': ['sepal length (cm)',
'sepal width (cm)',
'petal length (cm)',
'petal width (cm)'],
'filename': 'iris.csv',
```

```
'data_module': 'sklearn.datasets.data'}
```

2. Extracting the Dataframe out from JSON

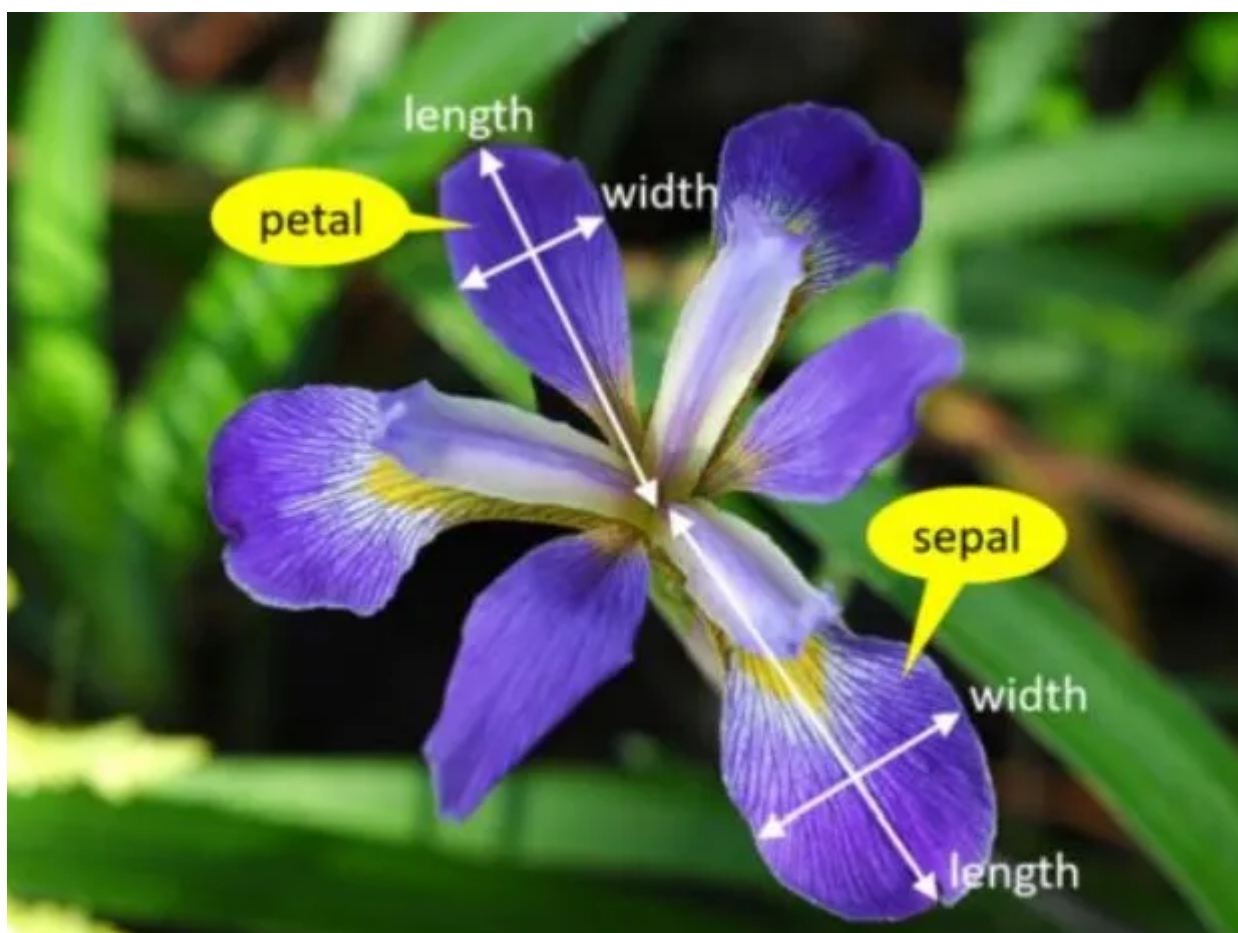
```
2. Extracting the Dataframe out from JSON
```

```
1 df = pd.DataFrame(data=iris.data, columns=iris.feature_names)
```

```
[ ] 1 df
```

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
...
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9
147	6.5	3.0	5.2	2.0
148	6.2	3.4	5.4	2.3

```
df = pd.DataFrame(data=iris.data, columns=iris.feature_names)  
df
```



3. Using Reset + Set_Index to Rename the Index Column to "Row No."

3. Using Reset + Set_Index to Rename the Index Column to "Row No."

```
[ ] 1 df = df.reset_index()  
    2 df
```

	index	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	0	5.1	3.5	1.4	0.2
1	1	4.9	3.0	1.4	0.2
2	2	4.7	3.2	1.3	0.2
3	3	4.6	3.1	1.5	0.2
4	4	5.0	3.6	1.4	0.2
...
145	145	6.7	3.0	5.2	2.3
146	146	6.3	2.5	5.0	1.9
147	147	6.5	3.0	5.2	2.0
148	148	6.2	3.4	5.4	2.3
149	149	5.9	3.0	5.1	1.8

```
df = df.reset_index()  
df
```

```
1 df = df.rename(columns={"index": "Row No."})  
2 df
```

	Row No.	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	0	5.1	3.5	1.4	0.2
1	1	4.9	3.0	1.4	0.2
2	2	4.7	3.2	1.3	0.2
3	3	4.6	3.1	1.5	0.2
4	4	5.0	3.6	1.4	0.2
...
145	145	6.7	3.0	5.2	2.3
146	146	6.3	2.5	5.0	1.9
147	147	6.5	3.0	5.2	2.0
148	148	6.2	3.4	5.4	2.3
149	149	5.9	3.0	5.1	1.8

150 rows × 5 columns

```
df = df.rename(columns={"index": "Row No."})  
df
```

```
1 df = df.set_index('Row No.')
```

```
2 df
```

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
Row No.				
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
...
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9
147	6.5	3.0	5.2	2.0
148	6.2	3.4	5.4	2.3
149	5.9	3.0	5.1	1.8

```
df = df.set_index('Row No.')
```

```
df
```

4. Extracting out Iris Class

4. Extracting out the "Iris Class"

```
[ ] 1 y = pd.DataFrame(data=iris.target, columns=["Iris Class"])
```

```
1 y
```

	Iris Class
0	0
1	0
2	0
3	0
4	0
...	...
145	2
146	2
147	2
148	2
149	2

```
y = pd.DataFrame(data=iris.target, columns=["Iris Class"])
```

```
y
```


5. Using Map to Create a New Column Z

5. Using Map to Create a New Column Z

To Map the Iris Type

- 0: Setosa
- 1: Versicolor
- 2: Virginica

```
1 z = y["Iris Class"].map({0: 'Setosa', 1: 'Versicolor', 2: 'Virginica'})
2 z
```

```
0      Setosa
1      Setosa
2      Setosa
3      Setosa
4      Setosa
...
145   Virginica
146   Virginica
147   Virginica
148   Virginica
149   Virginica
Name: Iris Class, Length: 150, dtype: object
```

```
z = y['Iris Class'].map({0: 'Setosa', 1: 'Versicolor', 2: 'Virginica'})
z
```

6. Combining to get the Final Table...

6. Combining to get the Final Table...

```
[ ] 1 df["Iris Type"] = z
```

```
[ ] 1 df
```

Row No.	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	Iris Type
0	5.1	3.5	1.4	0.2	Setosa
1	4.9	3.0	1.4	0.2	Setosa
2	4.7	3.2	1.3	0.2	Setosa
3	4.6	3.1	1.5	0.2	Setosa
4	5.0	3.6	1.4	0.2	Setosa
...
145	6.7	3.0	5.2	2.3	Virginica
146	6.3	2.5	5.0	1.9	Virginica

```
df["Iris Type"] = z
df
```

THE END

About Dr. Alvin Ang



www.AlvinAng.sg

Dr. Alvin Ang earned his Ph.D., Masters and Bachelor degrees from NTU, Singapore. Previously he was a Principal Consultant (Data Science) as well as an Assistant Professor. He was also 8 years SUSS adjunct lecturer. His focus and interest is in the area of real world data science. Though an operational researcher by study, his passion for practical applications outweigh his academic background. He is a scientist, entrepreneur, as well as a personal/business advisor.

More about him at www.AlvinAng.sg.

[Json](#)

[Scikit Learn](#)

[Data Conversion](#)

[Json To Csv](#)

[Google Colab](#)

