

レシピ66 変数値カラム名をmeltで整然化（P194）

- Pythonの大規模ライブラリではありがちだが、pandasでは同じ作業を行うのに様々な方法がある
- よくある違いは読みやすさと性能
- pandasには melt というDataFrameメソッドがあり、前のレシピ65の stack と同じ働きをするが柔軟性が高い
- このレシピでは、 melt メソッドを使って、変数値カラム名の簡単なDataFrameを整然化する

(1) state_fruit2を読み込み、どのカラムを変形し、どのカラムは必要ないかを定める

In [3]:

```
import pandas as pd
state_fruit2 = pd.read_csv('state_fruit2.csv')
state_fruit2
```

Out[3]:

	State	Apple	Orange	Banana
0	Texas	12	10	40
1	Arizona	9	7	12
2	Florida	0	14	190

(2) id_varsとvalues_varsパラメーターに適切なカラムを渡し、 melt メソッドを使う

melt

- id_vars は、変形せず保持するカラム名のリスト
- id_vars は、1つのカラムに変形するカラム名のリスト
- value_vars は、「識別変数」の略で、同じカラムにとどまり、value_varsに渡されたカラムごとに繰り返す
- melt はインデックスの値を削除し、RangeIndexに置き換える・・・（4）を参照
- 保持したいインデックスの値があるなら、meltを使う前に、インデックスをリセットする必要がある

In [6]:

```
# こちらはカラム名はデフォルトのまま
state_fruit2.melt(id_vars=['State'],
                  value_vars=['Apple', 'Orange', 'Banana'])
```

Out[6]:

	State	variable	value
0	Texas	Apple	12
1	Arizona	Apple	9
2	Florida	Apple	0
3	Texas	Orange	10
4	Arizona	Orange	7
5	Florida	Orange	14
6	Texas	Banana	40
7	Arizona	Banana	12
8	Florida	Banana	190

(3) 1ステップでカラム名まで変更できる

In [5]:

```
state_fruit2.melt(id_vars=['State'],
                  value_vars=['Apple', 'Orange', 'Banana'],
```

```
var_name='Fruit',
value_name='Weight')
```

Out[5]:

	State	Fruit	Weight
0	Texas	Apple	12
1	Arizona	Apple	9
2	Florida	Apple	0
3	Texas	Orange	10
4	Arizona	Orange	7
5	Florida	Orange	14
6	Texas	Banana	40
7	Arizona	Banana	12
8	Florida	Banana	190

(4) インデックスありの状態で、meltをそのまま適用するとそれを削除してしまう

In [7]:

```
state_fruit2.set_index('State').melt()
```

Out[7]:

	variable	value
0	Apple	12
1	Apple	9
2	Apple	0
3	Orange	10
4	Orange	7
5	Orange	14
6	Banana	40
7	Banana	12
8	Banana	190

(補足) 一部割愛