Untitled

April 12, 2023

[2]:
```
import numpy as np
import pandas as pd
```

[9]:
```
d0 = pd.read_csv("states0.csv")
d1 = pd.read_csv("states1.csv")
d2 = pd.read_csv("states2.csv")
df = pd.concat([d0, d1], axis=0)
df = pd.concat([df, d2], axis=0)
df
```

[9]:

	Unnamed: 0	State	TotalPop	Hispanic
0	0	Alabama	4830620	3.7516156462584975%
1	1	Alaska	733375	5.909580838323351%
2	2	Arizona	6641928	29.565921052631502%
3	3	Arkansas	2958208	6.215474452554738%
4	4	California	38421464	37.291874687968054%
5	5	Colorado	5278906	20.78438003220608%
0	0	Colorado	5278906	20.78438003220608%
1	1	Connecticut	3593222	15.604830917874388%
2	2	Delaware	926454	8.82476635514019%
3	3	District of Columbia	647484	9.165921787709499%
4	4	Florida	19645772	21.3385426653884%
5	5	Georgia	10006693	8.418242207460397%
0	0	Georgia	10006693	8.418242207460397%
1	1	Hawaii	1406299	9.186708860759486%
2	2	Idaho	1616547	11.505369127516781%
3	3	Illinois	12873761	15.601733547351516%
4	4	Indiana	6568645	6.536744186046501%
5	5	Iowa	3093526	5.30364520048603%

	White	Black	Native
0	61.878656462585%	31.25297619047618%	0.4532312925170065%
1	60.910179640718574%	2.8485029940119775%	16.39101796407186%
2	57.120000000000026%	3.8509868421052658%	4.35506578947368%
3	71.13781021897813%	18.968759124087573%	0.5229197080291965%
4	40.21578881677474%	5.677396405391911%	0.40529206190713685%
5	69.89557165861504%	3.546376811594201%	0.5738325281803548%

0	69.89557165861504%	3.546376811594201%	0.5738325281803548%
1	67.6770531400966%	10.34806763285027%	0.12620772946859898%
2	64.63271028037383%	20.743925233644834%	0.25981308411214965%
3	33.103910614525134%	51.77653631284915%	0.20055865921787713%
4	59.08374880153398%	15.165675934803444%	0.2104506232023015%
5	54.28630556974962%	32.08829841594277%	0.18758303525804798%
0	54.28630556974962%	32.08829841594277%	0.18758303525804798%
1	25.032278481012657%	2.052848101265823%	0.1449367088607596%
2	83.1362416107383%	0.5667785234899323%	1.468120805369128%
3	60.85980738362764%	17.108410914927717%	0.11842696629213499%
4	78.43189368770771%	11.18697674418606%	0.1940863787375415%
5	87.71968408262464%	3.2569866342648868%	0.2897934386391251%

	Asian	Pacific	Income	\
0	1.0502551020408146%	0.03435374149659865%	$43296.35860306644	
1	5.450299401197604%	1.0586826347305378%	$70354.74390243902	
2	2.876578947368419%	0.16763157894736833%	$54207.82095490716	
3	1.1423357664233578%	0.14686131386861315%	$41935.63396778917	
4	13.052234148776776%	0.35141038442336353%	$67264.78230266465	
5	2.661996779388082%	NaN	$64657.801787164906	
0	2.661996779388082%	NaN	$64657.801787164906	
1	4.021980676328502%	0.01859903381642512 3%	$76146.5605875153	
2	3.2686915887850483%	NaN	$61827.97663551402	
3	3.3832402234636865%	0.02960893854748603 4%	$75466.36363636363	
4	2.2831735378715257%	0.05151006711409391%	$50690.194986743794	
5	3.0976494634644895%	0.04660194174757282 4%	$50811.08205128205	
0	3.0976494634644895%	0.04660194174757282 4%	$50811.08205128205	
1	36.59208860759495%	8.758860759493672%	$73264.42628205128	
2	1.135906040268457%	0.1271812080536914%	$48017.31543624161	
3	4.475377207062604%	0.02003210272873195%	$59587.04887459807	
4	1.5782724252491687%	0.03262458471760798%	$48616.22784810127	
5	1.699392466585662%	0.05516403402187123 5%	$53017.75304136253	

	GenderPop
0	2341093M_2489527F
1	384160M_349215F
2	3299088M_3342840F
3	1451913M_1506295F
4	19087135M_19334329F
5	2648667M_2630239F
0	2648667M_2630239F
1	1751607M_1841615F
2	448413M_478041F
3	306674M_340810F
4	9600009M_10045763F
5	4883331M_5123362F
0	4883331M_5123362F

```
1      709871M_696428F
2      810464M_806083F
3    6316899M_6556862F
4    3235263M_3333382F
5    1534595M_1558931F
```

[3]:
```
data = pd.DataFrame(np.arange(6).reshape((2, 3)),
index=pd.Index(['Ohio', 'Colorado'], name='state'),
columns=pd.Index(['one', 'two', 'three'],
name='number'))
```

[4]:
```
data
```

[4]:
```
number    one  two  three
state
Ohio        0    1      2
Colorado    3    4      5
```

[5]:
```
result = data.stack()
```

[6]:
```
result
```

[6]:
```
state     number
Ohio      one       0
          two       1
          three     2
Colorado  one       3
          two       4
          three     5
dtype: int64
```

[7]:
```
result.unstack()
```

[7]:
```
number    one  two  three
state
Ohio        0    1      2
Colorado    3    4      5
```

[8]:
```
result.unstack(0)
```

[8]:
```
state   Ohio  Colorado
number
one        0         3
two        1         4
three      2         5
```

[13]:
```
result = df.stack()
```

[14]:
```
result
```

[14]:
```
0  Unnamed: 0                          0
   State                         Alabama
   TotalPop                      4830620
   Hispanic          3.7516156462584975%
   White                61.878656462585%
                              …
5  Native             0.2897934386391251%
   Asian               1.699392466585662%
   Pacific          0.055164034021871235%
   Income               $53017.75304136253
   GenderPop             1534595M_1558931F
Length: 195, dtype: object
```

[16]:
```
result.unstack()
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
/tmp/ipykernel_5954/1434488353.py in <module>
----> 1 result.unstack()

~/anaconda3/lib/python3.9/site-packages/pandas/core/series.py in unstack(self,␣
 ↪level, fill_value)
   4155         from pandas.core.reshape.reshape import unstack
   4156
-> 4157         return unstack(self, level, fill_value)
   4158
   4159     #␣
 ↪---------------------------------------------------------------------------

~/anaconda3/lib/python3.9/site-packages/pandas/core/reshape/reshape.py in␣
 ↪unstack(obj, level, fill_value)
    489         if is_1d_only_ea_dtype(obj.dtype):
    490             return _unstack_extension_series(obj, level, fill_value)
--> 491         unstacker = _Unstacker(
    492             obj.index, level=level, constructor=obj.
 ↪_constructor_expanddim
    493         )

~/anaconda3/lib/python3.9/site-packages/pandas/core/reshape/reshape.py in␣
 ↪__init__(self, index, level, constructor)
    138             )
    139
--> 140         self._make_selectors()
    141
    142     @cache_readonly
```

```
~/anaconda3/lib/python3.9/site-packages/pandas/core/reshape/reshape.py in␣
 ↪_make_selectors(self)
    190
    191         if mask.sum() < len(self.index):
--> 192             raise ValueError("Index contains duplicate entries, cannot␣
 ↪reshape")
    193
    194         self.group_index = comp_index

ValueError: Index contains duplicate entries, cannot reshape
```

[17]:
```
result = result.reset_index()
```

[18]:
```
result.unstack()
```

[18]:
```
level_0  0                          0
         1                          0
         2                          0
         3                          0
         4                          0
                           …
0        190       0.2897934386391251%
         191        1.699392466585662%
         192    0.055164034021871235%
         193        $53017.75304136253
         194         1534595M_1558931F
Length: 585, dtype: object
```

[ ]:
```

```