

pandas_pivot(8)

January 12, 2020

Pivot basics

```
[1]: import pandas as pd
import numpy as np
df = pd.read_csv("weather.csv")
df
```

```
[1]:      date      city  temperature  humidity
0  5/1/2017  new york           65         56
1  5/2/2017  new york           66         58
2  5/3/2017  new york           68         60
3  5/1/2017   mumbai           75         80
4  5/2/2017   mumbai           78         83
5  5/3/2017   mumbai           82         85
6  5/1/2017  beijing           80         26
7  5/2/2017  beijing           77         30
8  5/3/2017  beijing           79         35
```

```
[2]: df.pivot(index='city',columns='date')
```

```
[2]:      temperature      humidity
date      5/1/2017 5/2/2017 5/3/2017 5/1/2017 5/2/2017 5/3/2017
city
beijing           80         77         79         26         30         35
mumbai            75         78         82         80         83         85
new york          65         66         68         56         58         60
```

```
[3]: df.pivot(index='city',columns='date',values="humidity")
```

```
[3]: date      5/1/2017 5/2/2017 5/3/2017
city
beijing         26         30         35
mumbai          80         83         85
new york        56         58         60
```

```
[4]: df.pivot(index='date',columns='city')
```

```
[4]:
```

	temperature			humidity		
city	beijing	mumbai	new york	beijing	mumbai	new york
date						
5/1/2017	80	75	65	26	80	56
5/2/2017	77	78	66	30	83	58
5/3/2017	79	82	68	35	85	60

```
[5]: df.pivot(index='humidity',columns='city')
```

```
[5]:
```

	date			temperature		
city	beijing	mumbai	new york	beijing	mumbai	new york
humidity						
26	5/1/2017	NaN	NaN	80.0	NaN	NaN
30	5/2/2017	NaN	NaN	77.0	NaN	NaN
35	5/3/2017	NaN	NaN	79.0	NaN	NaN
56	NaN	NaN	5/1/2017	NaN	NaN	65.0
58	NaN	NaN	5/2/2017	NaN	NaN	66.0
60	NaN	NaN	5/3/2017	NaN	NaN	68.0
80	NaN	5/1/2017	NaN	NaN	75.0	NaN
83	NaN	5/2/2017	NaN	NaN	78.0	NaN
85	NaN	5/3/2017	NaN	NaN	82.0	NaN

Pivot Table

```
[6]: df = pd.read_csv("weather2.csv")
df
```

```
[6]:
```

	date	city	temperature	humidity
0	5/1/2017	new york	65	56
1	5/1/2017	new york	61	54
2	5/2/2017	new york	70	60
3	5/2/2017	new york	72	62
4	5/1/2017	mumbai	75	80
5	5/1/2017	mumbai	78	83
6	5/2/2017	mumbai	82	85
7	5/2/2017	mumbai	80	26

```
[7]: df.pivot_table(index="city",columns="date")
```

```
[7]:
```

	humidity		temperature	
date	5/1/2017	5/2/2017	5/1/2017	5/2/2017
city				
mumbai	81.5	55.5	76.5	81.0
new york	55.0	61.0	63.0	71.0

Margins

```
[8]: df.pivot_table(index="city",columns="date", margins=True,aggfunc=np.sum)
```

```
[8]:
```

	humidity			temperature		
date	5/1/2017	5/2/2017	All	5/1/2017	5/2/2017	All
city						
mumbai	163	111	274	153	162	315
new york	110	122	232	126	142	268
All	273	233	506	279	304	583

Grouper

```
[9]: df = pd.read_csv("weather3.csv")
df
```

```
[9]:
```

	date	city	temperature	humidity
0	5/1/2017	new york	65	56
1	5/2/2017	new york	61	54
2	5/3/2017	new york	70	60
3	12/1/2017	new york	30	50
4	12/2/2017	new york	28	52
5	12/3/2017	new york	25	51

```
[10]: df['date'] = pd.to_datetime(df['date'])
```

```
[11]: df.pivot_table(index=pd.Grouper(freq='M',key='date'),columns='city')
```

```
[11]:
```

	humidity	temperature
city	new york	new york
date		
2017-05-31	56.666667	65.333333
2017-12-31	51.000000	27.666667