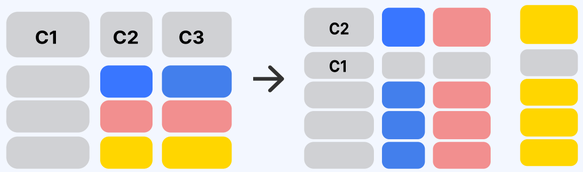
RESHAPING DATA

```
>>>df1=  pd.pivot_table(df3,
values='C4', index='C5',
columns='C6')
```

```
>>>df3=df7.pivot(index='C1',col
umns='C2', values='C3')
(Spread rows into columns)
```



MISSING DATA

```
>>>df1.dropna()
(Drop NaN values)
```

```
>>>df2.fillna(df2.mean())
(Fill NaN values with a
predetermined value)
```

```
>>>df3.replace(j,J)
(Replace values with
others)
```

DUPLICATE DATA

```
>>>df1.unique() (Return
unique values)
```

```
>>>df2.duplicated('C1') (Check
duplicates)
```

```
>>>df2.drop_duplicates('C2',
keep='last')
(Drop duplicates)
```

```
>>>df3.index.duplicated()
(Check index duplicates)
```

GROUPING DATA

Aggregation

```
>>>df2.groupby(by=['
C1', 'C2']).mean()
```

```
>>>df3.groupby(level
=0).sum()
```

```
>>>df4.groupby(level
=0).agg({'m':lambda
x:sum(x)/
len(x),'n':np.sum})
```

Transformation

```
>>>customSum = lambda x:
(x+x%2)
```

```
>>>df1.groupby(level=0).t
ransform(customSum)
```

COMBINING DATA

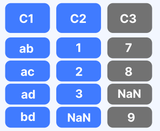
Merge


```
>>>pd.merge(df1,df2,how
='left',on= C1)
```


```
>>>pd.merge(df1,df2,how
='right',on= C3)
```


```
>>>pd.merge(df1,df2,how
='inner',on= C1)
```


```
>>>pd.merge(df1,df2,how
='outer',on= C1)
```



a. Vertical

```
>>>df1.append(df2)
```

b. Horizontal/Vertical

```
>>>pd.concat([df1,df2],axis=1,
join= 'inner')
```

```
>>>pd.concat([df1,df2]
,axis=1,
keys=['C1','C2')
```

```
>>>df1.join(d
f2,
how='right')
```

Jovian

DATA WRANGLING IN PANDAS