**Daily Report**

12th March, 2018

1. **Aggregate或者agg**

通过该函数能将任何函数功能作用于数组，如：grouped.agg(peak\_to\_peak)

对于不同列需要作用不同的函数功能时，可以将函数名作为string输入，该函数会按照列的顺序作用；也可以将列的name和函数名作为一个tuple输入，如：

grouped\_pct.agg(['mean', 'std', peak\_to\_peak])

grouped\_pct.agg([('foo', 'mean'), ('bar', np.std)])

对于不同列，需要作用相同的几个函数时，可以将函数作为一个list输入，同时，按需要作用函数的列进行分组，如：functions = ['count', 'mean', 'max']

result = grouped['tip\_pct', 'total\_bill'].agg(functions)

对于不同函数，作用于一列或者多列时，可以将相应的列名和函数名作为dict传入，如：grouped.agg({'tip' : np.max, 'size' : 'sum'})

grouped.agg({'tip\_pct' : ['min', 'max', 'mean', 'std'], 'size' : 'sum'})

1. **Transform**

transform会将一个函数应用到各个分组，然后将结果放在适当的位置. 如果各分组产生的标量值，则该标量值会被广播出去

transform也是有严格条件的特殊函数：传入的函数只能产生两种结果，要么产生一个可以广播的标量值(如：np.mean), 要么产生一个相同大小的结果数组