**Daily Report**

6th March, 2018

1. **Sort\_index**

通过该方法可以将Series或Dataframe按照index进行排序，但index和value仍然一一对应;改方法默认的排序方法是上升的；若想改为降序，只需写明ascending=False

对于Dataframe，若存在多个columns时同时只想改变某几竖列的排序时，可以将这几列的名字赋给by，如：frame.sort\_index(by='b')

1. **Order**

对于Series，若要其按value值大小进行排序，则通过order（）方法即可；该方法中，任何缺失的value值默认都将位于末端

1. **Rank（）**

该方法能够获得元素所处的位置，默认的rank方法表示在这个数在原来的Series中排第几名，有相同的数，取其排名平均（默认）作为值

也可通过赋值，使其序号变为出现顺序，如：

In [185]: obj.rank(method='first')

Out[185]:

0 6

1 1

2 7

3 4

4 3

5 2

6 5

上述方法也可通过ascending=False使其变为降序

1. 对于Series而言，其index没有必要都是不同的，可以相同；可以通过is\_unique来获知某个属性是否为唯一的，如：obj.index.is\_unique
2. 但为默认值时，Dataframe中的sum或mean方法遇到NA值便跳过，不予考虑，但是当添加语句skipna=False后，则有NA值的结果将变为NA
3. Idmin（）和idmax（）函数能够获知最小值和最大值的索引，但不能直接显示数字；describe（）函数能够获知关于某个值的很多信息，包括平均值，最小值等
4. Unique（）能将数组中的唯一value输出；value\_counts能将value重复次数输出
5. 对于Series，dropna（）函数能将其中包含的NA值删除；对于Dataframe，可以选择对包含NA的行和列删除，或者都删除；默认形式删除行；当输入how='all'时，将删除元素都为NA的行；对于列，只需输入axis=1；当输入thresh=n时，将输出数组中至少包含n个非NA值的行
6. Fillna（）函数能将存在NA值的位置都代之为括号中的值；也可以有如下写法：

df.fillna({1: 0.5, 3: -1})，表示在df中索引为1的列中的NA值都代之为0.5，而索引为3的列中的NA值都代之为-1

当不改变默认值时，使用fillna的对象不会改变，只能返回一个修改后的对象；但当输入inplace=True后，对象将发生改变

对于fillna函数，也能对其method进行设置，如：df.fillna(method='ffill')，将使NA的值与该列最后一个数值相同；同时，也能限制插补的长度，通过limit=？即可

1. **阶层式索引**

通过对index赋予多个lists或者arrays，将使index具有不同阶层

可以通过unstack（）函数将阶层式变为Dataframe

相反，stack函数能将Dataframe变为阶层式