**Daily Report**

9th March, 2018

1. **Join**

能将两个Dataframe按照index进行合并,也能使其按照Dataframe的某一列进行合并，如：left1.join(right1, on='key')

1. **Merge，join，concat区别**

1、当没有索引时：merge、join为按照一定条件合并

2、当有索引、并按照索引合并时，得到结果为两者混合到一起了，重新按照一定规则排序了

3、当没有索引时、concat不管列名，直接加到一起，可以加到后面、也可以加到右边，axis=0为加到后面，axis=1为加到右边，左边的数据结构没有变，变的是右边数据结构

4、当有索引、并按照索引合并时，得到结果两者混合到一起了

1. **Concat**

该函数能将两个Dataframe合并，当axis=0时，后一个Dataframe的元素加在前一个的后面；当axis=1时，则加在右边

对于合并后的新Dataframe，可以使用join\_axes来确定其index

1. **Combine\_first**

该方法即为‘补漏’作用，将需要作用的范围中出现某条件时，即使用另一个范围中的数值补漏

1. **Stack和unstack**

可以使用stack将Dataframe的列转化到行；使用unstack将其恢复

若使用unstack或stack的默认值，则最内层的level会发生转化；可以指定需要转化的level名或level层数

Unstack函数默认为在无对应数值时，添加NA；stack会自动删除NA值，可以通过dropna=False进行设置

1. **Duplicated和drop\_duplicates**

对于一个Dataframe，可以通过duplicated来获知是否其中元素重复，并且返回一个包含boolean类型的Series

通过drop\_duplicates可以将Dataframe中重复元素删除，返回剩余值