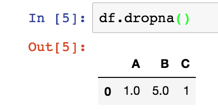
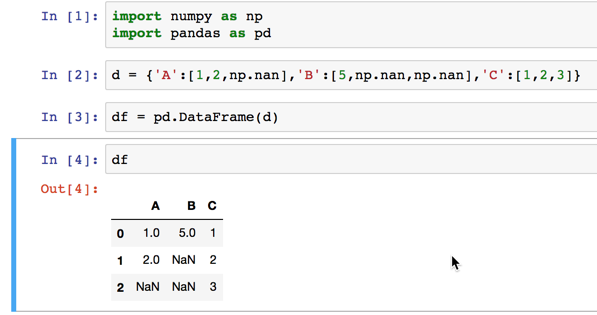
欢迎继续我们的学习。

在本课的内容中，我们将学习在Pandas中如何处理missing data（缺失数据）的问题。

首先还是打开Anaconda和JupyterNotebook，创建一个新的项目，将其命名为Missing Data。

使用代码创建一个有缺失数据的矩阵：



在以上代码中，我们创建了一个有缺失数据的DataFrame。

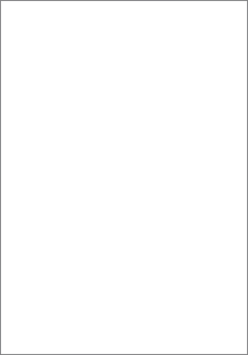
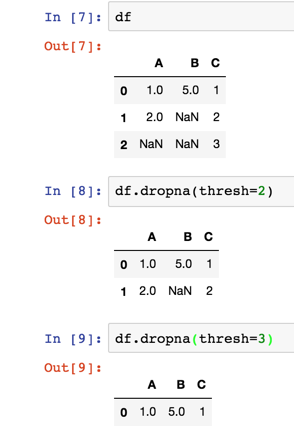
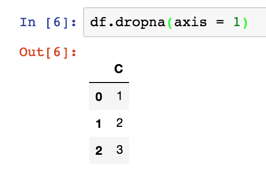
接下来看看如何处理缺失数据

首先是dropna函数。

可以看到，dropna函数直接删除了所有带缺失数据的行。

如果我们想按列删除数据，那么加上axis的参数就好了~

对于dropna函数来说，还有一个参数thresh也很有用，使用thresh参数可以指定保留有限个有效数据的行，比如：



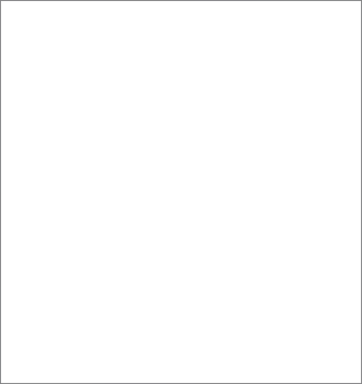
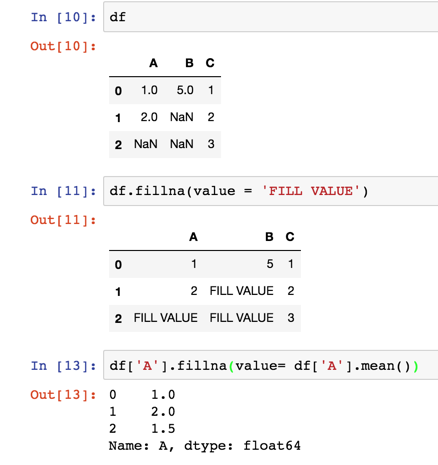
在以上代码中，我们通过设置thresh参数，就可以筛选出所需的数据行。

接下来看看如何替代缺失数据。

在以上代码中，我们使用fillna函数来替代所缺失的数据。

比如在cell11中，使用字符串替代缺失数据。

在cell13中，使用index为A的列的平均数来替代缺失数据。



讨论群-笨猫学编程QQ群：

375143733

知乎专栏：

<https://zhuanlan.zhihu.com/kidscoding>

新浪博客:

<http://blog.sina.com.cn/eseedo>

Github:

<https://github.com/eseedo>

个人网站：

<http://icode.ai/>