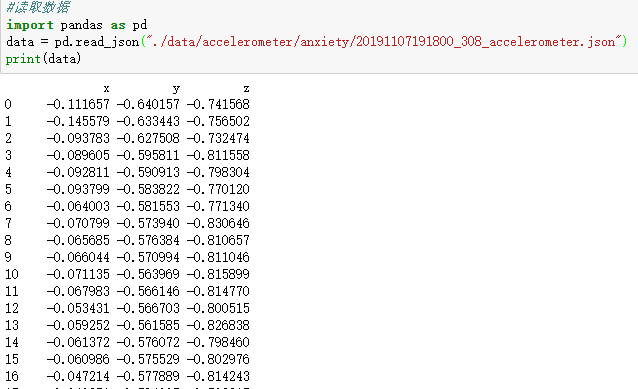
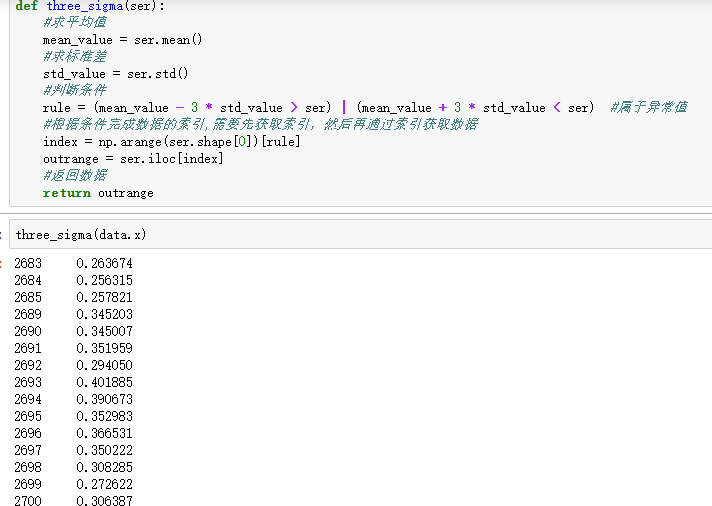
首先打开文件并且读取数据，发现数据量挺大



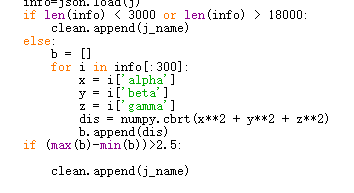
接着进行了简单的数据异常值处理：





发现用拉依达准则筛选出的数据很多，而且不知道筛选出的数据占整个文件的占比为多少算异常。

通过这种方式处理数据浮动异常，时间异常，文件异常，代码如下：



代码解析：

通过os.listdir获取路径data\_dir下的所有文件名称，具体代码如下：



打开json文件，具体代码如下：



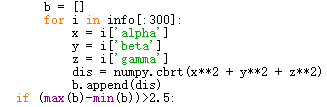
通过info = json.load(j)来读取json文件内容



根据采样频率，采样时间少于十分钟采集的点为3000个或者大于一小时采集的点为18000个，清理这样的的异常数据，并且清理空文件，代码如下：



根据三维坐标分析，开始时的波动异常只需将点的欧氏距离算在一个范围之外的都算异常。



然后进行数据处理即异常文件删除，先新创建文件夹以用来保存删除后的剩余文件。



打开并读取文件，json.load()只接收文件描述符，完成读取文件和反序列化。



保存文件，json.dump()是将info内容转换为JSON格式的字符串，将字符串写入到文件中，文件描述符f是必须要的参数。



数据处理后剩余文件：

