首先，pandas没有提供任何方法来操作mongodb数据库，需要使用pymongo的MongoClient类来操作mongodb数据库

pandas处理mongodb的步骤是：1.安装pymongo库，2.从pymongo库中导入MongoClient类，3.实例化一个MongoClient类并且用一个变量

来接收。查找这个数据库的所有数据，并且转换为list。4.利用pandas的DataFrame类从这个list中创建实例并且保存到一个变量中。然后就可以利用pandas来做相关处理。

简单实例：

创建一个数据库students（前提是你的机器安装了mongodb最好将其安装为服务，如果没有安装服务，需要运行mongod.exe）

打开一个cmd窗口

输入如下命令创建数据库students：

|  |
| --- |
| > use students  switched to db students |

创建一个my\_students集合,并且添加一些数据

|  |
| --- |
| > db.my\_students.insert({"name":"小美","age":"16","sex":"female"})  WriteResult({ "nInserted" : 1 })  …. |

现在可以读取数据了

|  |
| --- |
| import pandas as pd from pymongo import MongoClient # pandas不能直接读取mongodb，需要使用pymongo来读取然后pandas使用DataFrame类来处理读取的结果  client = MongoClient() collections = client["students"]["my\_students"] data = list(collections.find()) # print(data) df = pd.DataFrame(data=data) print(df) |

运行上面的程序，结果如下：

|  |
| --- |
| \_id name age sex  0 63261838db42b5a3ecaaa2ac 珍珍 18 female  1 632619c0db42b5a3ecaaa2ad LiLi 19 female  2 6326268bdb42b5a3ecaaa2ae 小明 18 male  3 632626a7db42b5a3ecaaa2af 小张 17 male  4 63262b05db42b5a3ecaaa2b0 小美 16 female |