**任务1 数据预处理（100分）**

在做数据统计之前，需要对数据进行预处理，具体要求如下：

分别删除drugs\_sale\_data1.csv和 drugs\_sale\_data2.csv中所有有缺失值的数据行；

将删除缺失值后的2个数据合并为一个数据，并输出合并后的行数；

删除订单号和商品编码这2个无关列。

【答案】（80分）

6433

**任务2 数据统计（100分）**

对处理好的数据进行统计，具体要求如下：

按照销售数量对数据行进行降序排序；

找出销量前三的药品，添加到列表中，并输出列表；

找出销量倒数前三的的药品，添加到列表中，并输出列表；

计算实收金额这一列数据的中位数、平均值（保留2位小数）、总数，并输出。**任务**

【答案】（每行15分，全正确80分）

['缬沙坦氢氯噻嗪片(复代文)', '吲达帕胺片(寿比山)', '苯磺酸氨氯地平片(络活喜)']

['缬沙坦胶囊(代文)', '尼群地平片', '厄贝沙坦片(安博维)']

26.5

46.48

299014.26

**3 数据更新（100分）**

将已经完成预处理和筛选更新后的数据，保存为新的文件。具体要求如下：

* 将DataFrame类型的数据转为字典类型；
* 将最新的数据保存为new\_drugs\_sale.json文件；
* 查看new\_drugs\_sale.json文件的大小。

【答案】（80分）

256kb 说明：值在254kb到258kb范围内都正确