**实验四 Pandas数据处理基础**

**【实验目的】**

1. 掌握pandas库的使用；
2. 能够使用pandas进行简单的数据处理。

**【实验内容】**

1. pandas常规操作；
2. 使用pandas进行简单的数据处理；

**【预备知识】**

1. pandas库
2. Python编程基础

**【实验设备】**

1.硬件：PC机

2.软件：python3开发环境

**【实验步骤】**

1. 使用pandas读取‘实验4数据.csv’文件的数据，并保存在data中；
2. 分别查看data的属性：index、columns、values、dtypes等；
3. 将data中的前10条记录和后15条记录（前10行和后15行）打印出来；
4. 将data中的前3列和后3列打印出来；
5. 使用loc和iloc打印出第100~150行（index值为100~150）和第2、4、6、8列（下标为1,3,5,7）的数据；
6. 将与第一行的Id号具有相同Id号的行打印出来；
7. 将第一位用户的步数大于10000步的数据取出来；
8. 统计出一共有几位用户（Id是用户的唯一标识，统计不同Id数）；
9. 删除LoggedActivitiesDistance列的数据；
10. 为data添加一列VeryActiveSpeed，用于表示用户的VeryActiveDistance和VeryActiveMinutes的比值，表示活跃期间的平均速度，单位是米/秒。
11. 计算出每位用户的平均步长。
12. （选做）将时间列中的年月日改成月份和日期的形式，如4/21/2016改成421，5/1/2016改为501；
13. （选做）考虑VeryActiveDistance、ModeratelyActiveDistance和LightActiveDistance的值与当天消耗的Calories之间的关系，设计算法，分析这三种不同运动强度下消耗Calories的占比关系。