**上机作业七**

1. **数据读取**

- PopulationData 文件夹中包含 251 个csv文件，分别存储了世界上 251 个国家和地区 1950 ~ 2021 年 72 年间的人口数据，同时包括了不同年龄段的人口数据

- 请你读取 251 个文件并将数据汇总到一个数据结构之中（推荐使用pandas库中的DataFrame结构）

- 请输出 1950 年和 2021 年的世界总人口

**2、异常值处理**

- 编写函数识别数据中的异常值并对异常值进行修复（提示：某国某年的总人口等于各个年龄段的人口之和）

**3、数据挖掘**

- 计算每个国家 72 年间的人口平均增长率，并列出增长率最高的10个国家，计算公式为：

人口增长率=（年末人口数-年初人口数）/年平均人口×1000‰

- 计算平均增长率与平均人口数之间的相关系数，计算公式为：



**4、数据可视化**

- 绘制 China 2021 年各个年龄段人口的频数分布直方图

- 绘制 China和India 1950 ~ 2021 年人口总数的变化情况（折线图）

**5、数据预测**

- 根据已有数据，对 China 未来三年（2022~2024）的总人口数进行预测（提示：使用线性平均法或指数平滑法）