

## 20210318 作业

- 1、文件 `ChinaAirport.txt` 是一份存在一定格式问题的数据文件，其内容是 2018 年中国 235 个机场吞吐量。数据由三列组成，第一列是序号，第二列是机场名称，第三列是年度累计吞吐量，第四列是同比增速（同 2017 年相比）。另外，部分机场没有 2018 年数据（原因未知）。请根据数据完成作业。注意：数据中有千分位逗号且逗号也是数据列分隔符；
- 2、编写一个名为“学号\_DataArrage.py”的程序，用于消除第三列中的千分位(不能手工删除，要考虑大量数据时)，并删除缺失 2018 年数据的机场数据(不能手工删除，判断标志是该行第三列数据缺失)。将整理后的数据存储为“学号\_cnAirport.txt”，提交时一并提交该数据文件和程序。后续问题都是基于该数据；
- 3、编写一个名为“学号\_DataDeep.py”程序，用于完成如下功能：
  - 1) 读入第二问的数据文件，将数据整理成 list,每个 list 成员为 tuple，包含机场名称(字符串)，2018 吞吐量(浮点数)、2017 吞吐量(浮点数)、同比增量(浮点数)；
  - 2) 以 1)数据为基础，输出按 2017 年吞吐量降序排名的数据，每个机场一行，第一列为序号(按排序后的顺序编号，从 1 开始计数)、第二列为机场，第三列为 2018 年吞吐量、第四列为 2017 年吞吐量，没有标题行，列之间用英文逗号间隔；
  - 3) 以 1)数据为基础，输出按 2018 年增量(2018 年减去 2017 年吞吐量)降序排名的数据，每个机场一行，第一列为序号(按排序后的顺序编号，从 1 开始计数)、第二列为机场，第三列为 2018 年吞吐量、第四列为 2017 年吞吐量，没有标题行，列之间用英文逗号间隔；
  - 4) 以 1)数据为基础，输出按 2018 年增长率(2018 年减去

2017 年吞吐量除以 2017 年吞吐量)降序排名的数据，每个机场一行，第一列为序号(按排序后的顺序编号，从 1 开始计数)、第二列为机场，第三列为 2018 年吞吐量、第四列为 2017 年吞吐量，没有标题行，列之间用英文逗号间隔；

- 4、 作业提交内容：两个 Python 源代码（学号\_DataArrage.py、学号\_DataDeep.py）、一个数据文本文件（学号\_cnAirport.txt）。
- 5、 作业提交截止时间：2021-03-24 18:00 前。