

## 题目二十八：共享单车租赁时天气数据分析

### 【数据说明】

该数据集是 2011、2012 年共享单车租借统计数据集。

字段说明：

- 1) instant: 租借记录序号。
- 2) dteday : 日期
- 3) season : 季节 (1:springer, 2:summer, 3:fall, 4:winter)
- 4) yr : 年(0: 2011, 1:2012)
- 5) mnth : 月 ( 1 to 12)
- 6) hr : 时 (0 to 23)
- 7) holiday : 是否是假期
- 8) weekday : 星期几
- 9) workingday : 是否工作日.
- 10) weathersit : 天气
  - 1: 晴天 (Clear)
  - 2: 雾、云 (Cloudy)
  - 3: 小雨小雪(Light Rain)
  - 4: 大雨(Heavy Rain)
- 11) temp : 温度
- 12) atemp: 归一化温度
- 13) hum: 归一化湿度。
- 14) windspeed: 归一化风速。
- 15) casual: 临时用户。
- 16) registered: 注册用户。

### 【任务】

- 1、用 pandas 库读取 “bike\_day.csv” 文件，查看前五、后两行，并将缺失值全部丢弃处理。
- 2、选择 instant、dteday、weathersit、casual、registered 共 5 列数据导出到文本文件 “bike\_weathersit\_user.txt”，要求数据之间用空格分隔，每行末尾包含换行符。
- 3、读取文本文件 “bike\_weathersit\_user.txt”，计算列 casual 与列 registered 和（即 casual+registered），并作为一个新的列 cnt 添加到原始数据，导出到新的 Excel 文件 “bike\_weathersit\_user\_cnt.xlsx” 中。

- 4、读取 Excel 文件 “bike\_weathersit\_user\_cnt.xlsx”，按照列 weathersit 分组，计算列 cnt 的均值，并将分类计算结果导出到新的文本文件 “bike\_weathersit\_user\_cnt\_mean.txt” 中。
- 5、读取文本文件 “bike\_weathersit\_user\_cnt\_mean.txt”，可视化显示上一步分组计算得出的 cnt 均值。要求以列 weathersit 为 X 轴，轴刻度分别为 “Clear”、“Cloudy”、“lightRain”、“heavyRain”，包括图例、图标题，填充颜色为蓝色，并保存为 “bike\_weathersit\_user\_cnt.png”，要求分辨率不低于 300dpi。

**【要求】**

- 1、根据以上数据处理任务，设计并编程实现 “数据分析与可视化系统”，要求
  - ① 各个任务选择用菜单实现（菜单可用字符串输出模拟，或者 Tkinter 形式实现）。
  - ② 各个任务名称自己定义，须由独立的函数实现，且每个任务执行成功与否须给出必要的文字提示。
  - ③ 数据输入和结果输出的文件名须由人工输入，且输出结果都要以文件形式保存。
  - ④ 为保持程序的健壮性，各个任务执行过程中需要进行必要的判断（如文件是否存在、输入是否合法等）、程序异常控制等。
- 2、根据以上统计结果，书写不少于 300 字的结果分析。