

交通信息检测与处理 作业 1 R 基本操作

上交时间 10 月 19 日 上课时间

1 数据

FSTR_LOOPGROUPID	FDT_TIME	FINT_VOLUME	FINT_SPEED	FINT_OCCUPY
检测器编号	日期 时间	流量	速度	占有率
NHNX40(1)	2010-04-24 00:00:00	2	72	2
NHNX40(1)	2010-04-24 00:00:20	7	73	8

2 作业

2.1 数据框操作

- 1) 读取线圈检测器数据样例: `Detector_sample.csv`, 输出总记录数. (参考函数: `read.table()`, `read.csv()`)
- 2) 统计流量一列的缺失个数. (参考函数: `is.na()`)
- 3) 对包含缺失值的行进行剔除, 输出剔除后记录数. (参考函数: `na.omit()`)
- 4) 统计冗余记录数, 如有需剔除. (参考函数: `duplicated()`)
- 5) 计算流量、速度、占有率的平均值、方差、25%、50%和 75%分位数. (参考函数: `quantile()`)

2.2 日期时间处理

- 6) 对数据框按照时间列进行排序. (参考函数: `order()`)
- 7) 分别提取日期时间列的小时、分钟、秒, 新增为 `DAY`、`MINUTE`、`SECOND` 列. (参考包: `lubridate`, 其中的 `day()`, `minute()`, `second()`)

2.3 绘图

- 8) 绘制流量、速度、占有率的直方图. (参考绘图包: `ggplot2`)
- 9) 绘制流量-速度散点图.