说明文档

**text\_process.r**

R语言代码文档，目标是将评论原始数据进行切词分割，并计算TF-IDF值。输出 处理后评价数据.xlsx 文档

**处理后评价数据.xlsx**

共1754行，3423（=3400+3）列，作为SPSS模型的输入。

每一行代表一条评论数据

3400列代表切割后的词语，每一列的值，代表当前词语在评论的TF-IDF值

有3个列例外：

* Product\_name：商品名称，每条评论的评价对象
* Poor\_review：是否是差评，1=差评，0=好评
* Index：评论序号