

第 7 周课外作业

实验素材：

1、2 万字以上的英文文档和中文文档各一篇，带有白色背景的轮廓图片二张，使用词频统计，去除无意义的词频数据，分别用词云图展示英文文档和中文文档的统计结果，以此分析词云图所表达文档的内涵。

2、从网上下载电子版水浒传文档，参考课程案例，使用中文分词，统计出 108 将每个人在书中出场的次数。

每个同学上交的作业包括使用的素材文件，PYTHON 的源代码，要求源代码在自己的文件夹环境中就可以运行。