小组成员：谢安琪、金成志、李鑫、陈思宇、王宏宇、任勇鹏、徐娆、周晖

摘要：

在学校图书馆中的馆藏资源中进行搜索，并通过爬虫获取了150篇关于数据分析的论文。获取到论文之后，通过分析每篇上传论文中的文章来源、文章的引用文献数量、出版年份、作者人数以及文章高频词这五个功能来获取每篇论文的基本信息。

这五大信息能够检测出论文的质量以及时效性情况，帮助用户更快地进行识别和参考适合自己的论文。最终可以将分析论文得到的结果进行下载和展示。

数据描述：

来源：成都信息工程大学图书馆馆藏资源，搜索关键词为“数据分析”，单位为成都信息工程大学

规模：150个数据集。