**词频统计**

## **需求分析**

* （1）程序可读入任意英文文本文件，该文件中英文词数大于等于1个。
* （2）程序需要很壮健，能读取容纳英文原版《哈利波特》10万词以上的文章。

## **可行性分析**

* （1）操作可行性：系统的操作方式能否在用户组织内行得通，词频统计所需要的资源有限。
* （2）经济可行性：从经济角度考虑，规划方案的实现性和开发的价值，本次是一个作业，不存在经济问题。
* （3）技术可行性：要考虑采用的硬件和软件能否满足用户需求。

## **基本功能**

* （1）指定单词词频统计功能：用户可输入从该文本中想要查找词频的一个或任意多个英文单词，运行程序的统计功能可显示对应单词在文本中出现的次数和柱状图。
* （2）高频词统计功能：用户从键盘输入高频词输出的个数k，运行程序统计功能，可按文本中词频数降序显示前k个单词的词频及单词。
* （3）统计该文本所有单词数量及词频数，并能将单词及词频数按字典顺序输出到文件result.txt。

## **环境需求**

* 1.测试机环境：Windows环境
* 2.JDK版本：jdk8u161
* 3.JRE版本：jre8u161

## **功能设计**

* 1.新建文本文件data.txt并且统计各个单词出现的次数
* 2.如果次数相同，安装单词的字典顺序排序，输出单词个数
* 3.输入要查询的单词，显示它出现的次数 ，用户从键盘输入高频词输出的个数k，运行程序统计功能，可按文本中词频数降序显示前k个单词的词频及单词。.统计该文本所有单词数量及词频数
* 4.输出结果并将其存储到result.txt文件中