挑战题：data1 文件里记录是一些命令的操作记录，现在需要你从里面找出出现频率次数前3的命令并保存在 /home/shiyanlou/result。

目标：1处理文本文件 /home/shiyanlou/data1

2将结果写入 /home/shiyanlou/result

3结果包含三行内容，每行内容都是出现的次数和命令名称，如“100 ls”

提示：

1. cut 截取 (参数可以使用 -c 8-，使用 man cut 可以查看含义)
2. uniq -dc 去重
3. sort 的参数选择 -k1 -n -r
4. 操作过程使用管道，

答案：cat data1 |cut -c 8-|sort|uniq -dc|sort -rn -k1 |head -3 > /home/shiyanlou/result

解释：这个打开之后会发现他是1000条命令，前面是数字后面是命令，题中卓要找出频率前3的命令。

首先我们要截取出命令，所以要用cut，cut –c 8-就可以把前面的数字去掉。Sort排序后将同样的命令连在了一起，再使用uniq去重。-d:只输出重复的行 -c:标出出现的次数 sort再排一次，-n是通过前面的数值排序，-r是逆序排 -k1，就是选择第一列， head -3 ，就是取最前面的三条，然后输出到/home/shiyanlou/result

/////////////////////////////////////////////////////////////////////

正则表达式

* +表示前面的字符必须出现至少一次(1次或多次)，例如，"goo+gle",可以匹配"gooogle","goooogle"等；
* ?表示前面的字符最多出现一次(0次或1次)，例如，"colou?r",可以匹配"color"或者"colour";
* \*星号代表前面的字符可以不出现，也可以出现一次或者多次（0次、或1次、或多次），例如，“0\*42”可以匹配42、042、0042、00042等。

| **字符** | **描述** |
| --- | --- |
| \ | **将下一个字符标记为一个特殊字符、或一个原义字符。**例如，“n”匹配字符“n”。“\n”匹配一个换行符。序列“\\”匹配“\”而“\(”则匹配“(”。 |
| ^ | **匹配输入字符串的开始位置。** |
| $ | **匹配输入字符串的结束位置。** |
| {n} | n是一个非负整数。**匹配确定的n次**。例如，“o{2}”不能匹配“Bob”中的“o”，但是能匹配“food”中的两个o。 |
| {n,} | n是一个非负整数。**至少匹配n次**。例如，“o{2,}”不能匹配“Bob”中的“o”，但能匹配“foooood”中的所有o。“o{1,}”等价于“o+”。“o{0,}”则等价于“o\*”。 |
| {n,m} | m和n均为非负整数，其中n<=m。**最少匹配n次且最多匹配m次。**例如，“o{1,3}”将匹配“fooooood”中的前三个o。“o{0,1}”等价于“o?”。请注意在逗号和两个数之间不能有空格。 |
| \* | **匹配前面的子表达式零次或多次**。例如，zo\*能匹配“z”、“zo”以及“zoo”。\*等价于{0,}。 |
| + | **匹配前面的子表达式一次或多次**。例如，“zo+”能匹配“zo”以及“zoo”，但不能匹配“z”。+等价于{1,}。 |
| ? | **匹配前面的子表达式零次或一次**。例如，“do(es)?”可以匹配“do”或“does”中的“do”。?等价于{0,1}。 |
| ? | 当该字符紧跟在任何一个其他限制符（\*,+,?，{n}，{n,}，{n,m}）后面时，匹配模式是非贪婪的。非贪婪模式尽可能少的匹配所搜索的字符串，而默认的贪婪模式则尽可能多的匹配所搜索的字符串。例如，对于字符串“oooo”，“o+?”将匹配单个“o”，而“o+”将匹配所有“o”。 |
| . | **匹配除“\n”之外的任何单个字符**。要匹配包括“\n”在内的任何字符，请使用像“(.｜\n)”的模式。 |
| (pattern) | **匹配pattern并获取这一匹配的子字符串**。该子字符串用于向后引用。要匹配圆括号字符，请使用“\(”或“\)”。 |
| x｜y | **匹配x或y**。例如，“z｜food”能匹配“z”或“food”。“(z｜f)ood”则匹配“zood”或“food”。 |
| [xyz] | 字符集合（character class）。**匹配所包含的任意一个字符**。例如，“[abc]”可以匹配“plain”中的“a”。其中特殊字符仅有反斜线\保持特殊含义，用于转义字符。其它特殊字符如星号、加号、各种括号等均作为普通字符。脱字符^如果出现在首位则表示负值字符集合；如果出现在字符串中间就仅作为普通字符。**连字符 - 如果出现在字符串中间表示字符范围描述；如果出现在首位则仅作为普通字符。** |
| [^xyz] | 排除型（negate）字符集合。**匹配未列出的任意字符。**例如，“[^abc]”可以匹配“plain”中的“plin”。 |
| [a-z] | 字符范围。**匹配指定范围内的任意字符。**例如，“[a-z]”可以匹配“a”到“z”范围内的任意小写字母字符。 |
| [^a-z] | 排除型的字符范围。**匹配任何不在指定范围内的任意字符**。例如，“[^a-z]”可以匹配任何不在“a”到“z”范围内的任意字符。 |

优先级：从上到下依次降低

| **运算符** | **说明** |
| --- | --- |
| \ | 转义符 |
| (), (?:), (?=), [] | 括号和中括号 |
| \*、+、?、{n}、{n,}、{n,m} | 限定符 |
| ^、$、\任何元字符 | 定位点和序列 |
| ｜ | 选择 |