awk系列3：如何使用awk按模式帅选文本或者字符串

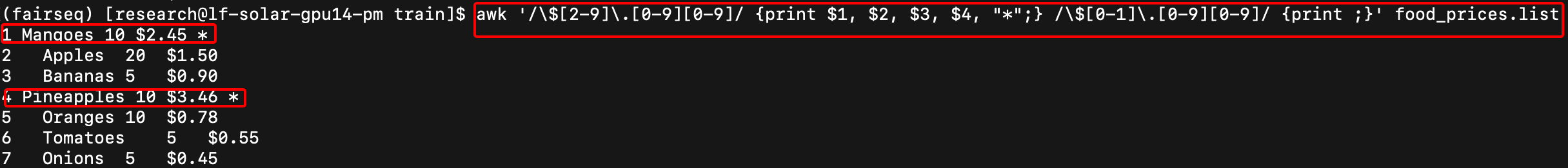
这一节主要是关于如何基于用户定义的特定模式来筛选文本或者字符串。

在筛选文本时，有时可能需要使用一个特定的模式去匹配文本中的某些内容，使用awk可以很方便的实现这个功能。

加入我们有下面的例子，比方说你有一个你想要购买的实物的购物清单，其名称为food\_prices.list，它包含所有的食物名称及相应的加个如下所示：



假如我们想使用一个(\*)符号去标记那些单价大于$2的食物，那么你可以通过运行下面的命令来达到此目的：



从上面的结果中可以看出含有Mangoes和pineapples的那行【他们的单价均大于$2】末尾都已经有了一个(\*)标记。

在这个例子中，我们使用了两个模式：

* 第一个模式：/\$[2-9]\.[0-9][0-9]/将会得到那些含有食物单价大于$2的行
* 第二个模式：/\$[0-1]\.[0-9][0-9] / 将会得到那些食物单价小于$2的那些行。

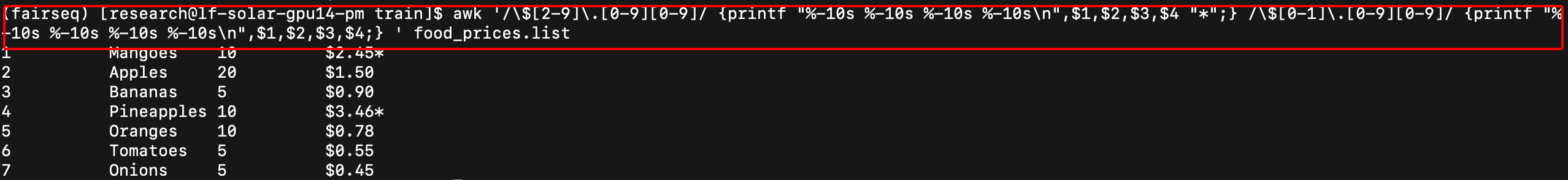
上面的命令具体做了什么呢：这个文件有四个字段，当模式一匹配到含有食物单价大于$2的行时，它便会输出所有的四个字段并在该行末尾加上一个(\*)符号来作为标记。

第二个模式只是简单地输出其他含有食物单价小于$2的行，按照现在输入文件food\_prices.list中的样子。

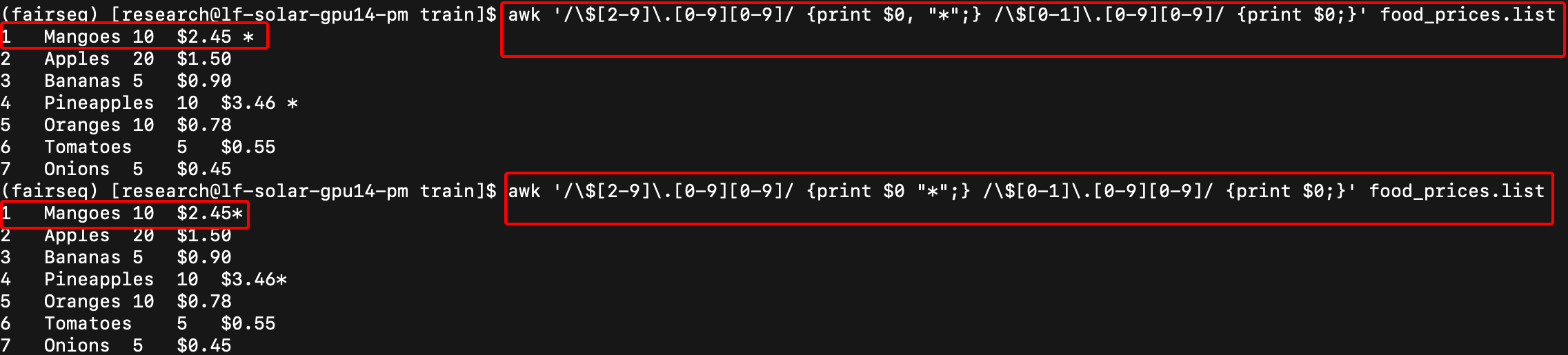
这样我们就可以使用模式来筛选那些价格超过$2的食物项目，尽管上面的输出还有些问题，带有(\*)符号的那些行并没有像其他行那样被格式化输出，这使得输出显得不够清晰。

下面我们可以使用下面的两种方式来解决：

1. 可以像下面这样使用printf命令，但这样使用又长又无聊：



1. 使用$0字段。Awk使用变量0来存储整个输入行。对于上面的问题，这种方式非常方便，并且它还简单、快速。



注意这里{print $0 "\*";}与{print $0, "\*"}的区别，使不使用","标志着是否在两个字段之间是否有分割符。