恭喜基本知识已经全部学完，开始进行期中考试

**时间安排**：这周5到下周日共计10天（越早完成，分数越高哦）

**难度**：不超纲（在这两周的参考答案中基本都有体现），但是偏难

**成绩规则**：

超时未完成：0蛋

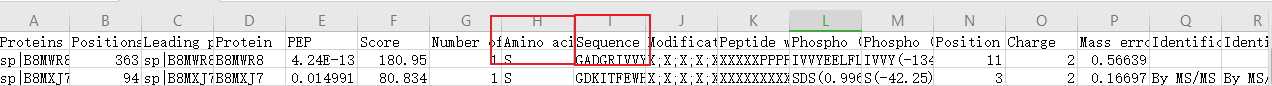
超时完成：还是0蛋，嘿嘿嘿

做出来：60分

60行以内（或者结构清晰）做出来：80分

惊艳到老师：100分（代码很优秀，或者提前完成任务等等都可以惊艳到老师哦）

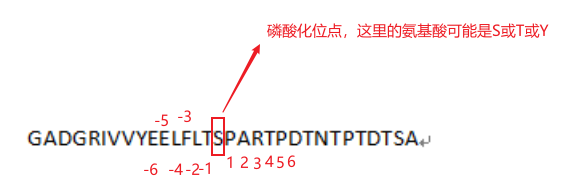
**解释说明：**有一个文件 Pho.csv



我们只需要其中的“Amino acid”和“Sequence window”列数据

Amino acid列是氨基酸残基S或T或Y，即Sequence window列数据的第16个位置的字母。

Sequence window列每个单元格都是31个氨基酸残基的肽段，肽段中间的是磷酸化位点氨基酸残基S或T或Y



**题目**：现在我们需要统计磷酸化位点左侧6个，右侧6个位置的每个氨基酸残基出现的个数，再由此计算出，每个位置出现每个氨基酸的概率是多少。S，T，Y分开计算，最后生成如下格式的3个csv文件。

示例结果如下（S）



这样算完之后，每列的数值相加正好为1（可能有计算误差，接近1）

恭喜，考试结束，开始逐步学习进阶知识

Python爬虫（request，json，html基础，BeatifuSoup，**pubmed API**，sci-hub下载）

其他（异常处理try，模块os等）

精髓完了，更难的我就不会了，嘤嘤嘤。