大家好，我们小组的课程项目主题是B站历史弹幕词频统计及情感分析，分为以下四个部分......

1. 简介。首先是项目初衷......，其次是对视频剧情简要介绍，便于理解后续的情感分析工作，这个视频主要讲述的是主角从美好的梦境中醒来，不得不接受残酷的现实的过程。
2. 数据爬取。简要分为四个步骤，首先......其次......然后......最后......

我们爬取了视频发布六天内的6000条数据，可以根据需要扩大爬取范围，这是我们爬取到数据集的部分内容

1. 词频统计。首先是分词操作......其次是统计操作......

分词统计后先进行可视化预处理，然后为每个可视化函数设置对应的参数，调用得到结果。

词云如图所示，其背景形状自定义为右边的图片，可以看到，频率较高的词为纳西妲、草神、玩到关服等，呈现的效果较好，符合游戏的内容。

然后是饼图的可视化和柱状图的可视化。

1. 情感分析。

（1）调用 snowNLP库下的 snowNLP()函数，对弹幕进行逐句的情感赋分，得分为0-1内的值，其中靠近0为消极，靠近1为积极；

1. 鉴于snowNLP()给出的评分并非是0-1上的均匀分布，评分的统计结果往往呈现两端少，中间多的迹象，不利于体现情感变化，故而对情感得分进行分级处理,为两端分配更大概率空间；

（3）划分数据以平滑化弹幕情感的变化，降低噪音的影响：

例如此处我们对相邻的200条弹幕的得分进行求和处理，以此减小同一时间段内个别弹幕与其相邻弹幕得分差异较大而带来的噪音影响，使得评分更具可读性，充分反映整体的情绪变化；

（4）虽然我们已经对情感分析的得分数据进行了些许平滑化处理与降噪处理，但是这时直接对该数据进行绘制仍然称不上美观（得到的结果如同股票涨势图）

故而，我们借助 np.linespace()以及make\_interp\_spline()函数，利用插值法对数据做平滑化处理，从而绘制出平滑的情绪变化曲线

最终绘制情绪变化曲线图 “情感分析.png”如图，考虑前面对视频剧情的介绍，我们发现，刚进入视频，观众情绪额外高涨，夹杂着对新视频的好奇与兴奋，然后降低至正常值，随着剧情展开，观众的情绪逐渐被欢快的氛围调动到最高值。然而剧情走向倏尔改变，由原来的欢快变为悲伤，观众的情绪瞬间跌到低估，随着视频接近尾声，观众的情绪平复到正常值。总体来看，使用NLP对B站弹幕进行情绪分析的效果达到预期。

以上是我们小组展示的全部内容，感谢聆听！