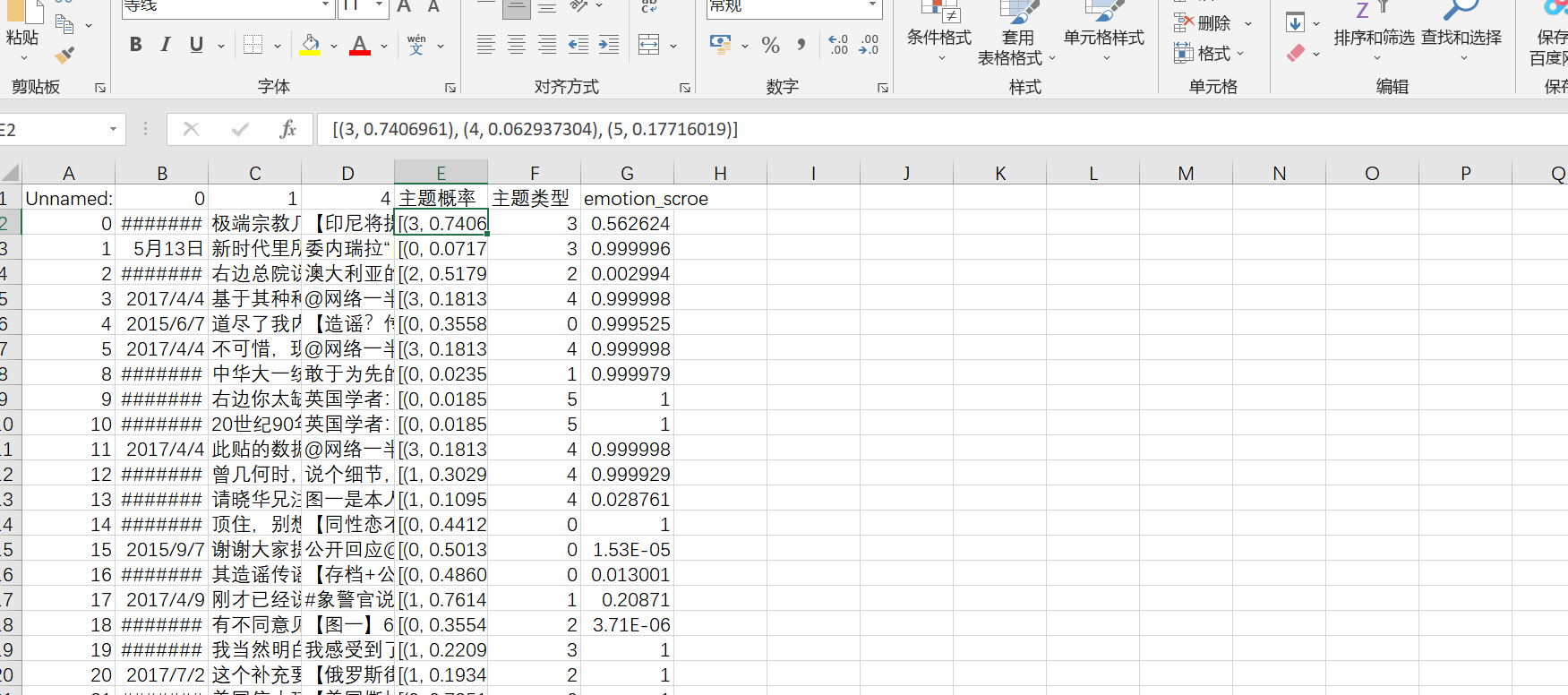
# 主题深度分析

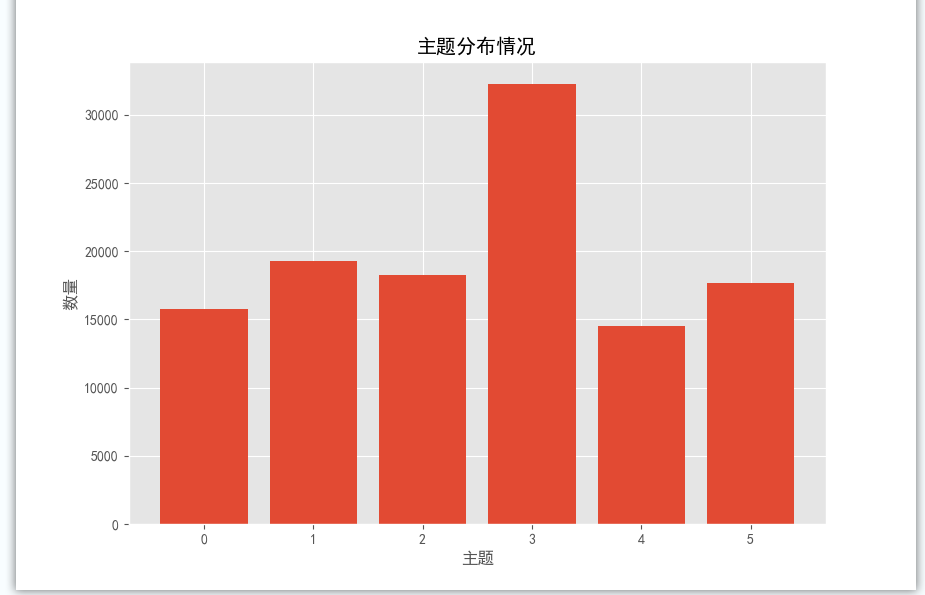


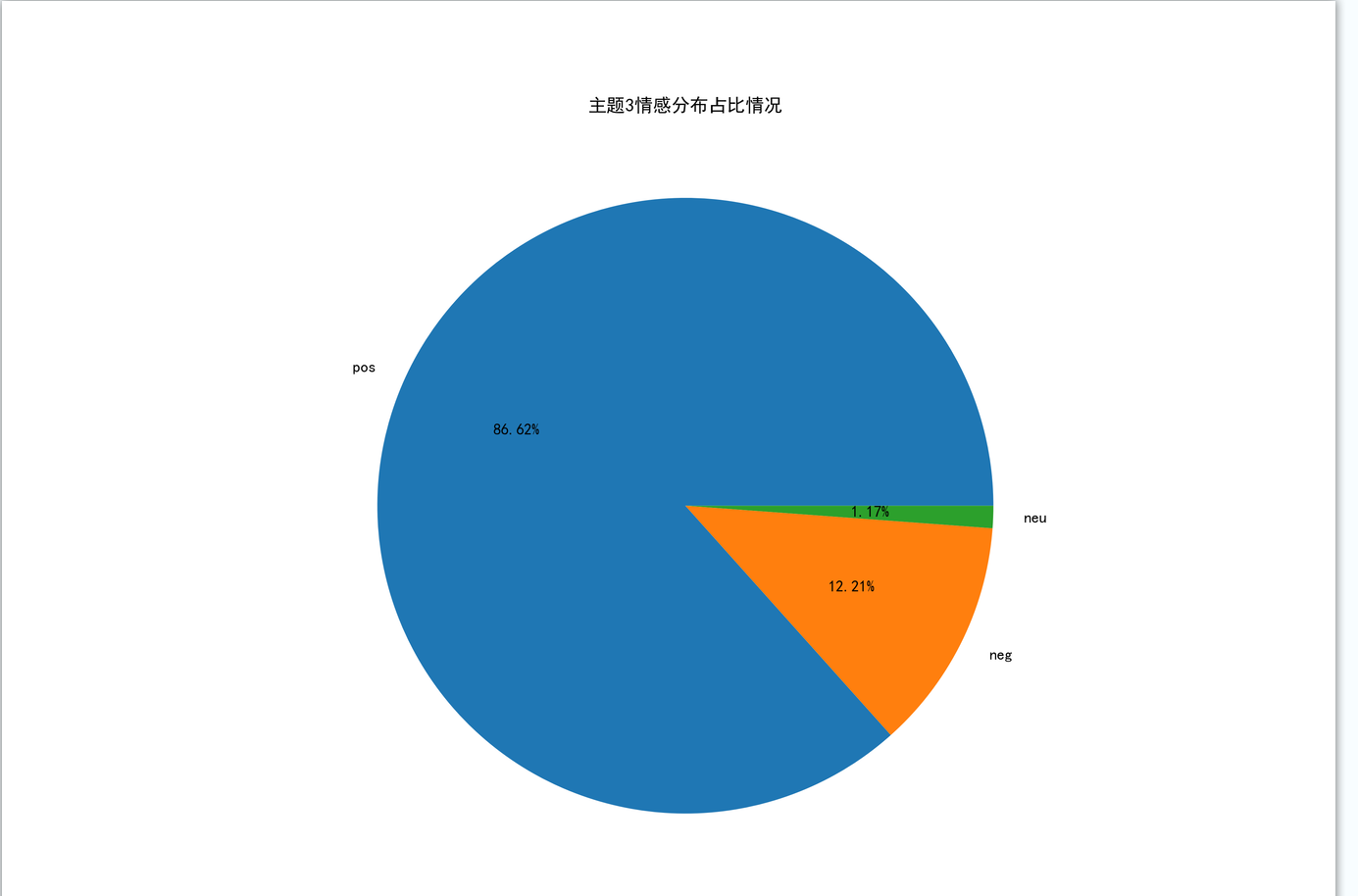
首先这里做了主题概率去计算每个文档的主题概率

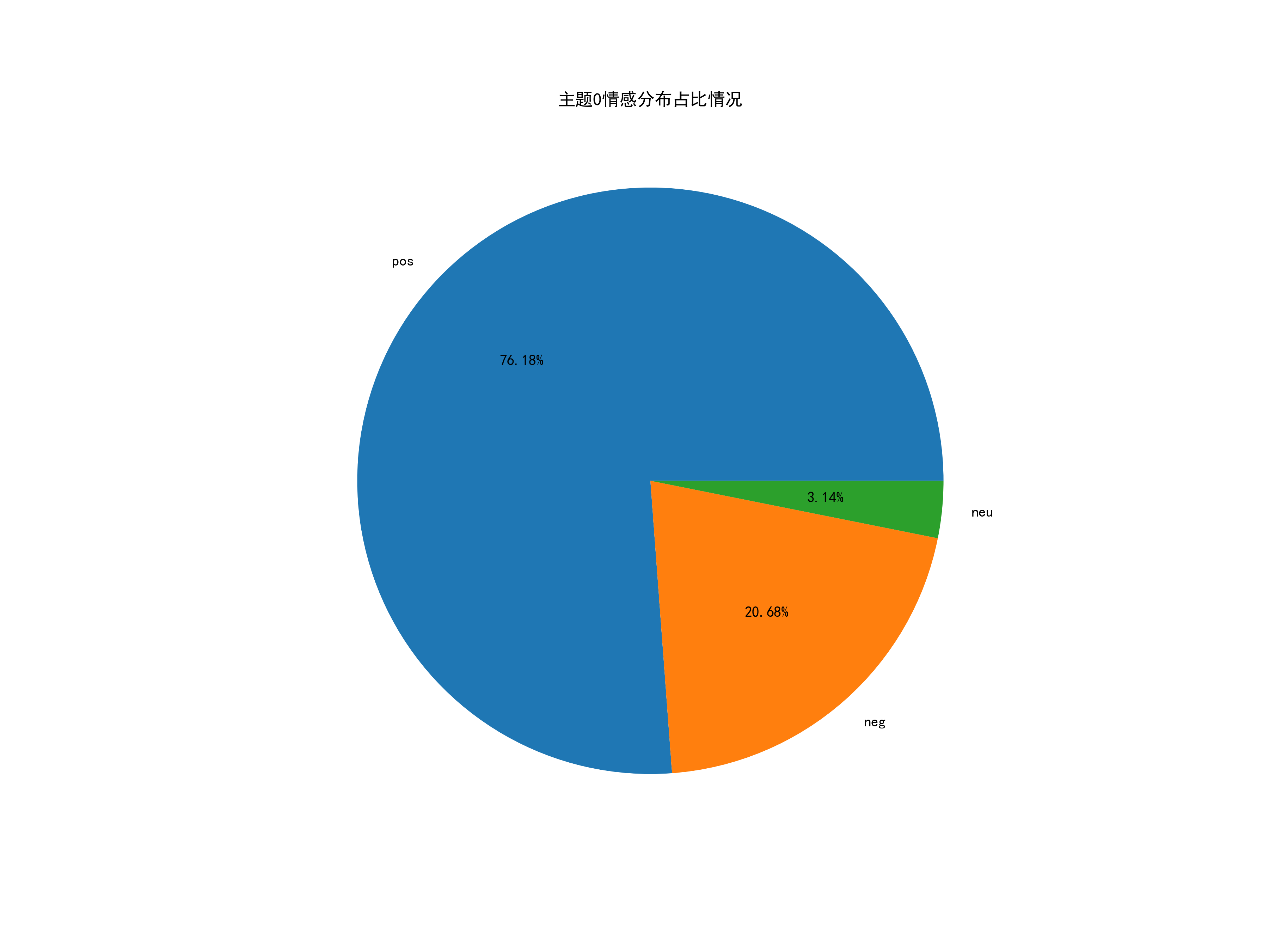
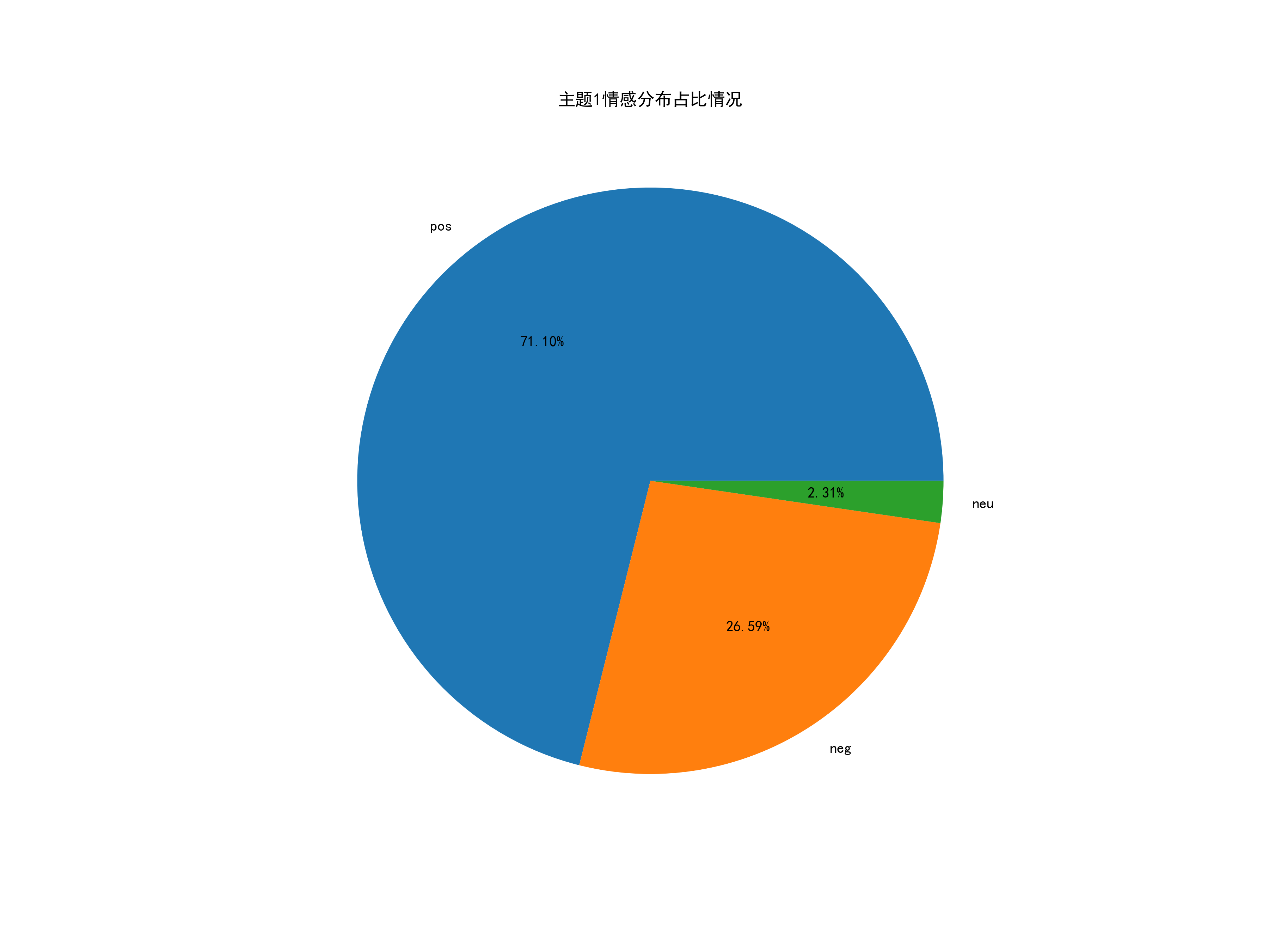
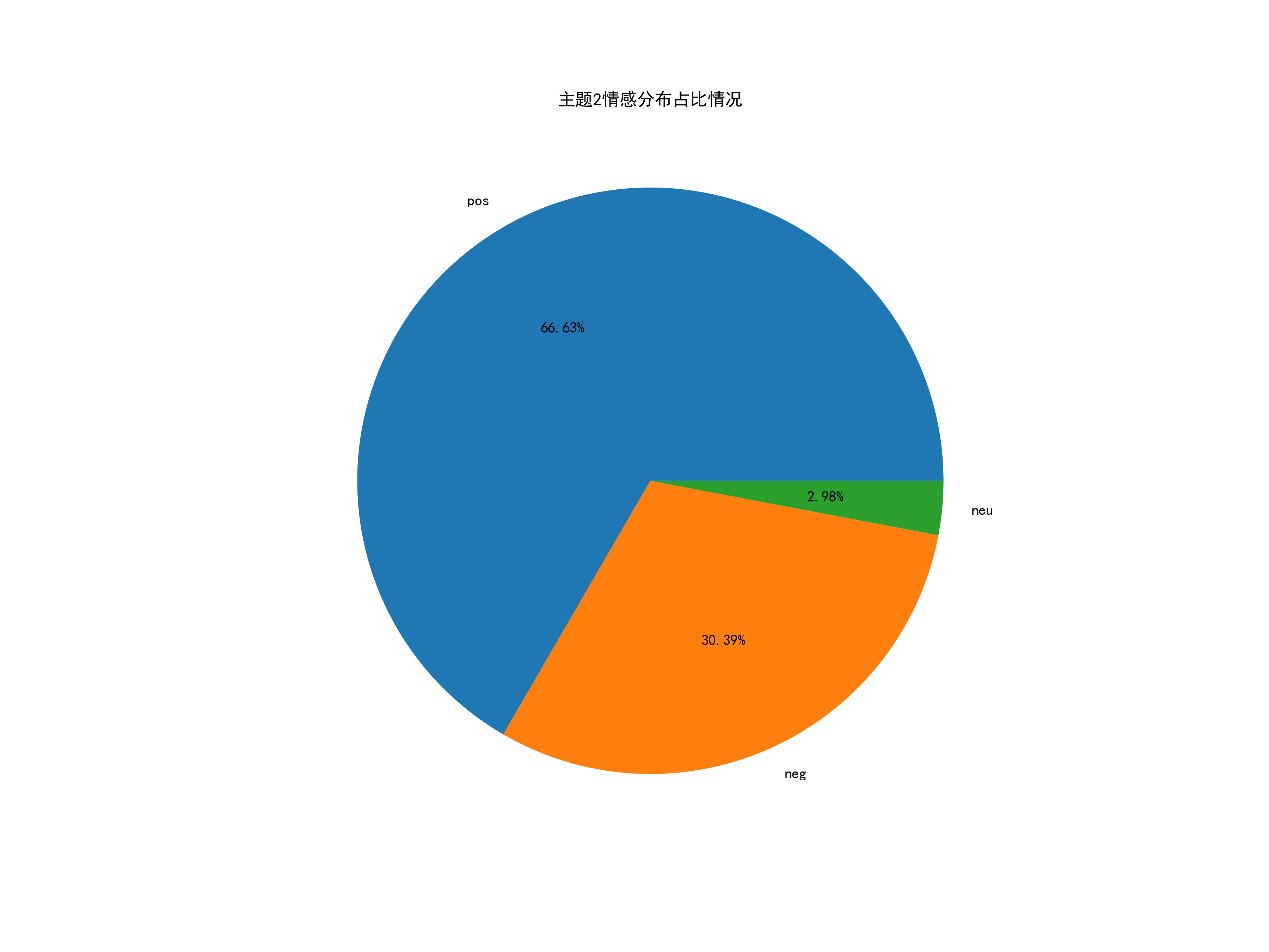
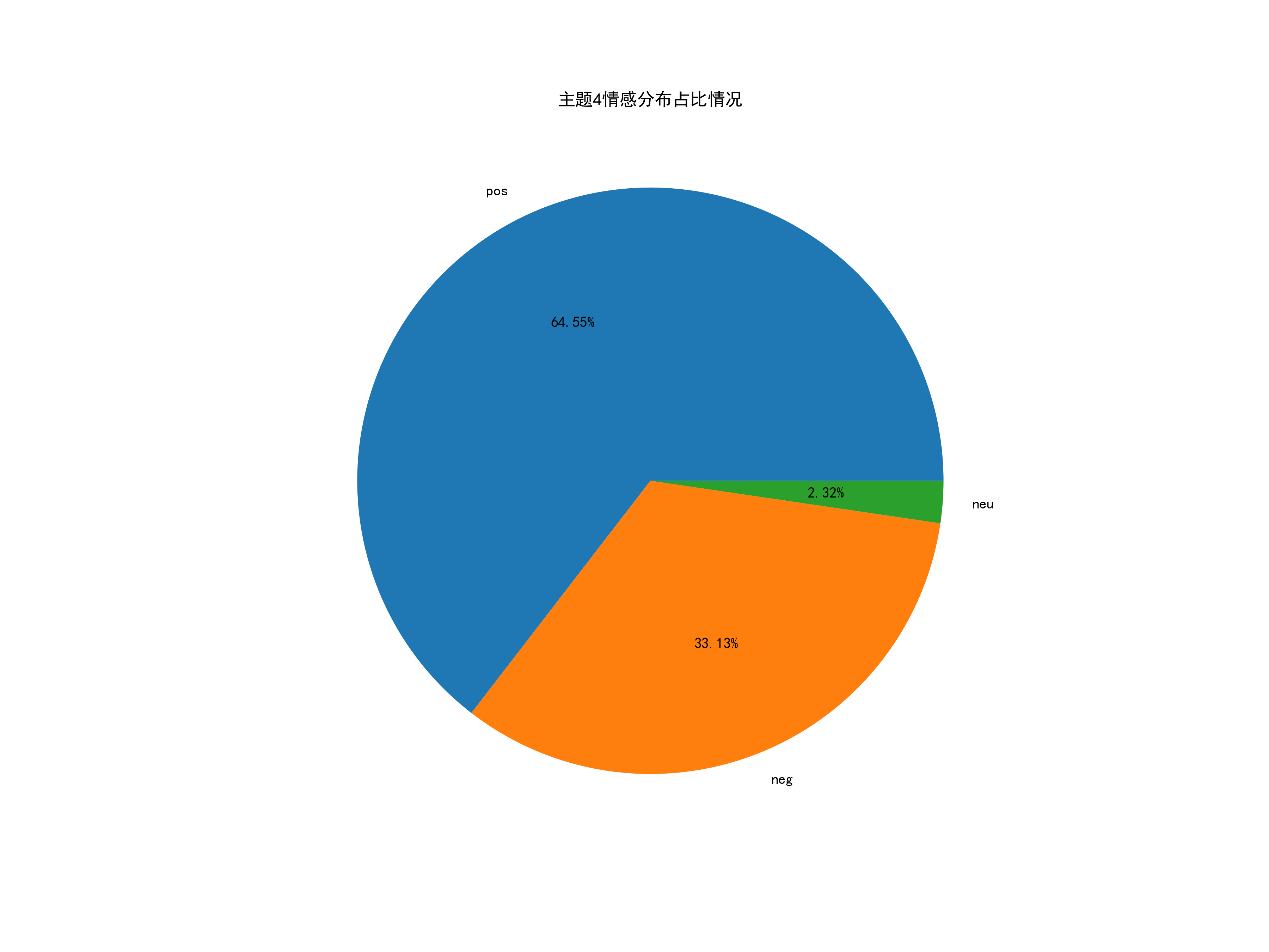
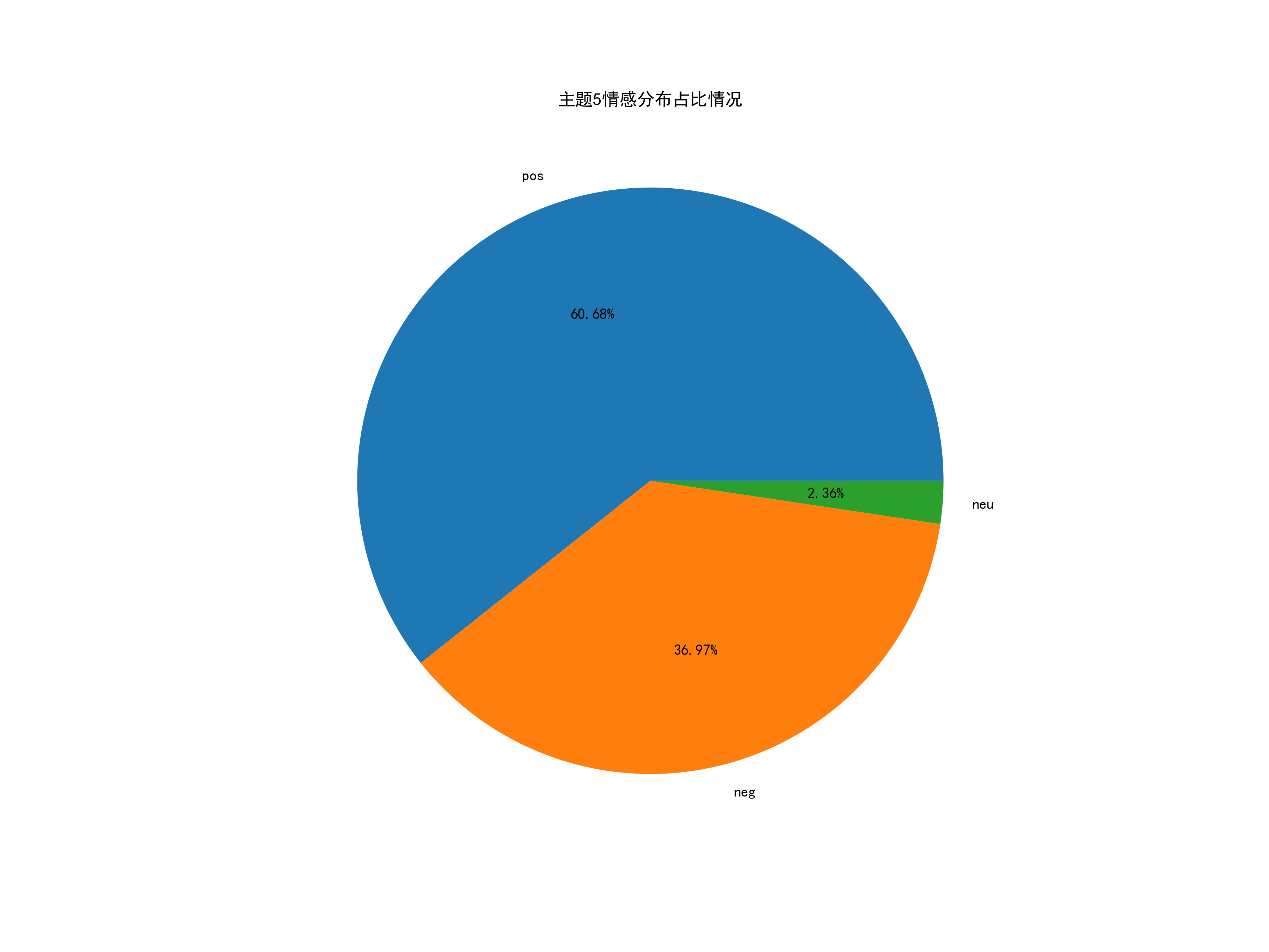
然后根据最大的概率来判断该文档是属于哪一个主题的，并且基于字典的情感分析使用snownlp算法去分析文档的情感属性

最后进行整合，去分析以下内容

通过计算去分析主题的出现频次

从这里看出，主题3数量是最多的，其他的主题基本处于用一个平均线上

而在主题3中，主题3的情感占比如图所示，主要以正面占比为主



面对其他主题的情感分布情况，大多数主题情感占比，正面占比均值60%以上，综合来看，主题3和其他主题差距略有不同