为了方便你更加理解明白我想要的数据和图形是什么，以及这些数据和图形的用途。

所以我简单的把我研究内容和研究方案描述一下~

**标绿的地方，是需要你用代码做好的图或数据**

标黄的地方，是解释我为什么需要这样的图或数据，用来干嘛~

***楷体并斜体的地方，是我新增的需求，新增了2处。***

--------------------------分界线，以下是说明-------------------------

**1.研究内容（研究数据来源）：**疫情期间，人民网留言板中的公民诉求留言内容以及政府回应内容。

**2.研究目的：**通过对文本情感分析和LDA主题模型对疫情期间公民留言内容和政府回应内容进行分析，来理清公民留言和政府回应内容具有什么样的特点。

--------------------------分界线，以下是需求-------------------------

**3.文本情感分析**

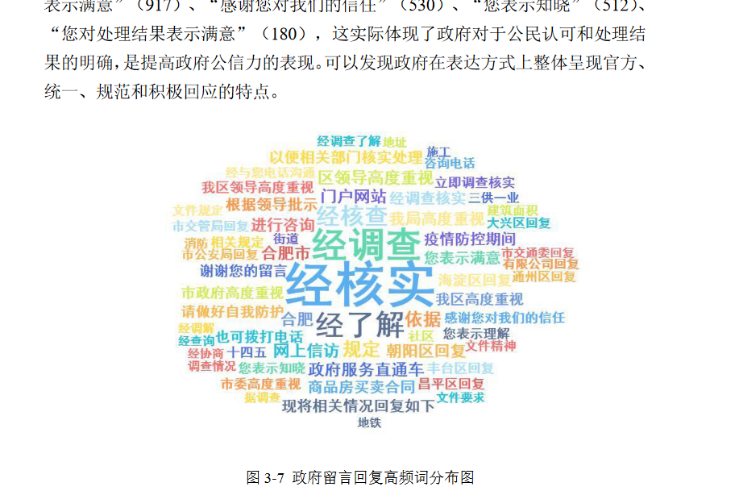
**3.1公民留言：**比如公民留言主题的偏好（问的最多的是医疗还是城建。。。），不同主题下公民留言的情感值，不同主题公民留言文字数量等；

**3.2政府回应：**不同主题政府的回应率、不同主题政府回应的时间间隔，政府回应时的情感值，不同主题政府回应的文字数量等等。然后利用以上数据做回归分析，分析诉求对于政府回应的影响，之间的关系。

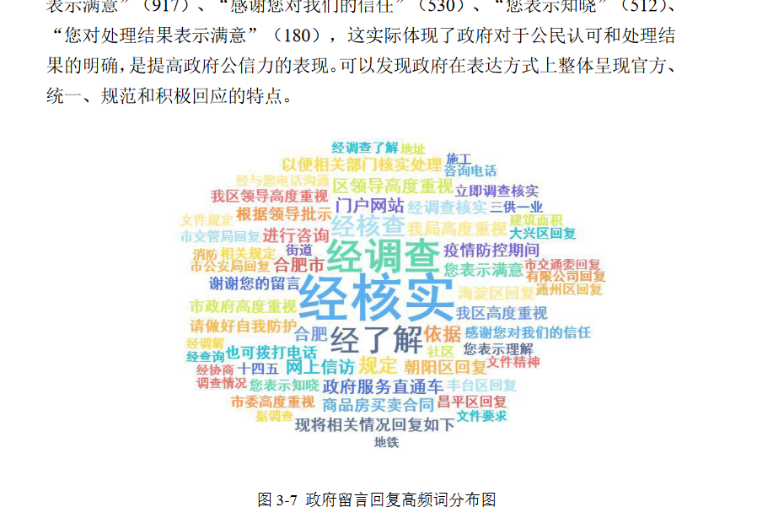
所以其中的情感值，需要用文本情感分析获得留言内容的情感值

所以在文本情感分析中，我需要的图有

**1.公民留言内容中高频词的词云图**

****

**2．政府回应内容中高频词的词云图**

****

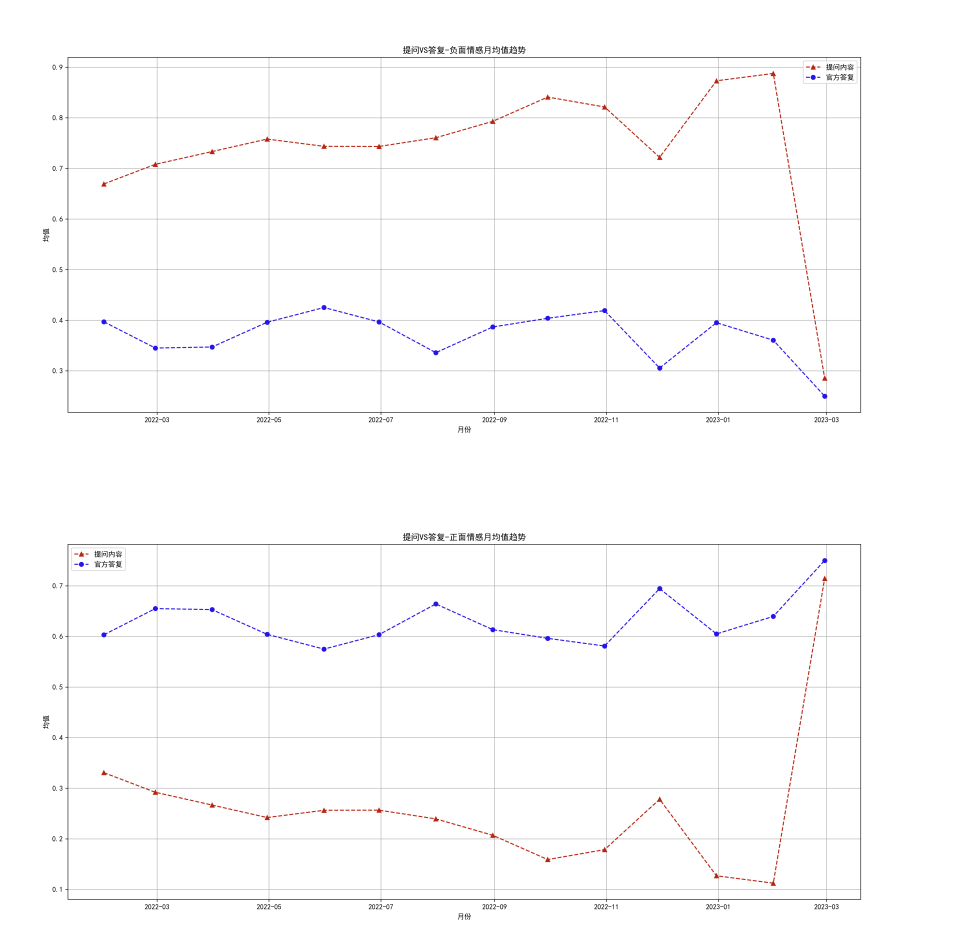
**3.公民留言和官方回复的情感特征，需要EXCEL表中每个留言内容增加一列：情感值**

**是为了做这个图~（图就是为了告诉你我需要的原因~，我自己做就行啦~）**

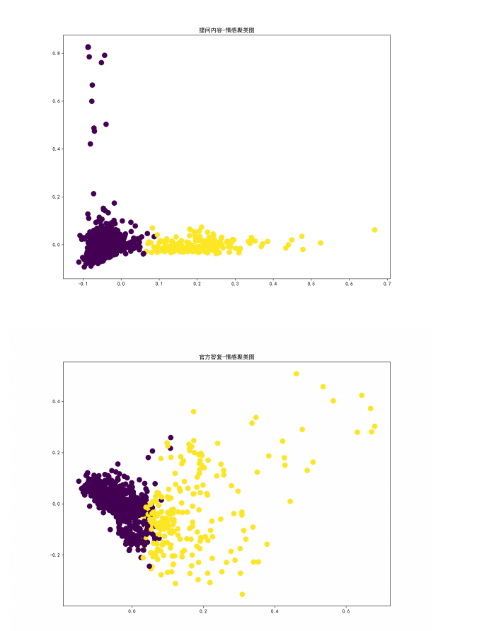
****

**4.提问VS答复 情感值趋势图，这个图就是你之前做的啦**

**就是标题改为：公民诉求VS政府回应 情感趋势图**



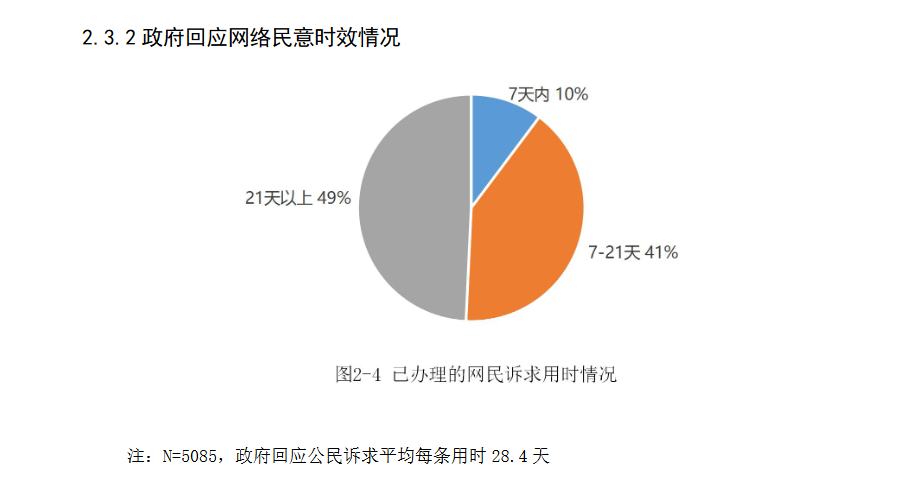
**5.情感聚类图，也是你之前做过的啦**

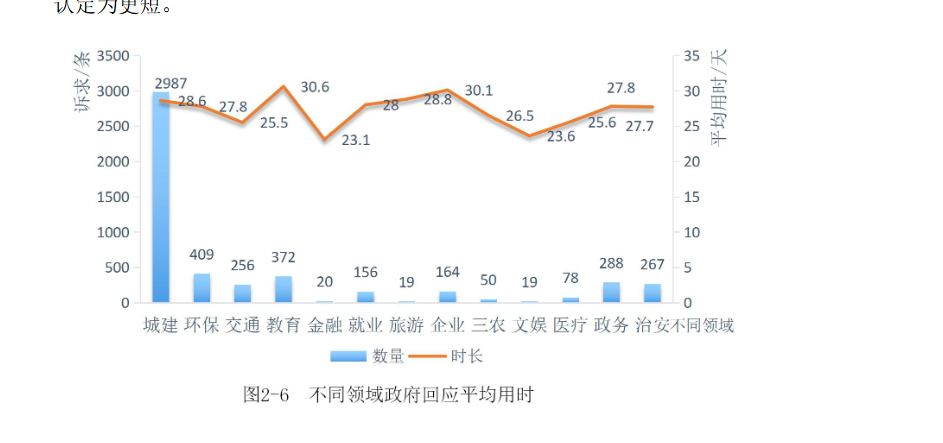
****

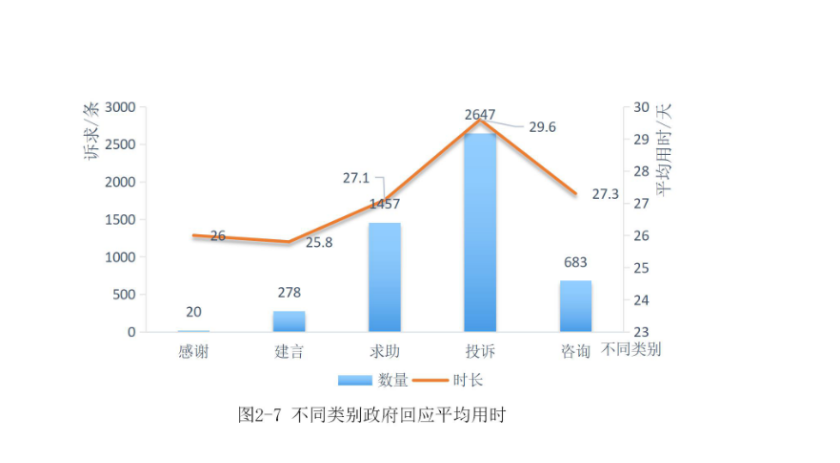
***6.需要EXCEL表最后增加一列：时间差（官方回复时间-公民提问时间，单位为天）***

***这个要麻烦你看看python能不能计算，EXCEL是可以计算的，但是我之前一直没成功，所以想看看你放不方便直接代码算，如果不方便，我就还是自己再想想其他办法啦。***

***需要以上数据，是为了做下列这些图***

****

****

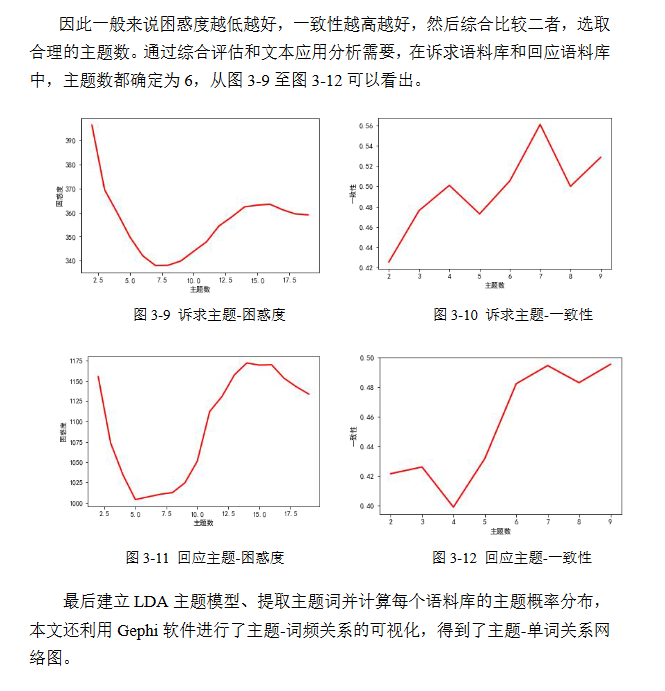
****

**4.LDA主题模型分析**

文本情感分析后，再用LAD主题模型对公民留言和政府留言进行分析。

1. ***诉求主题和回应主题的困惑度和一致性图***

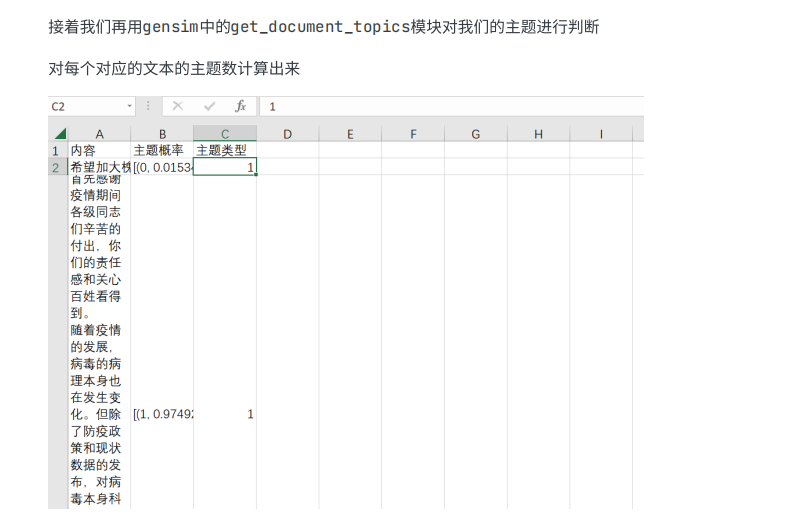
***因为我看其他用LDA主题模型分析的，确定主题数大多还是看困惑度，所以可以的话，还是希望用困惑度和一致性的图。确定主题数***

****

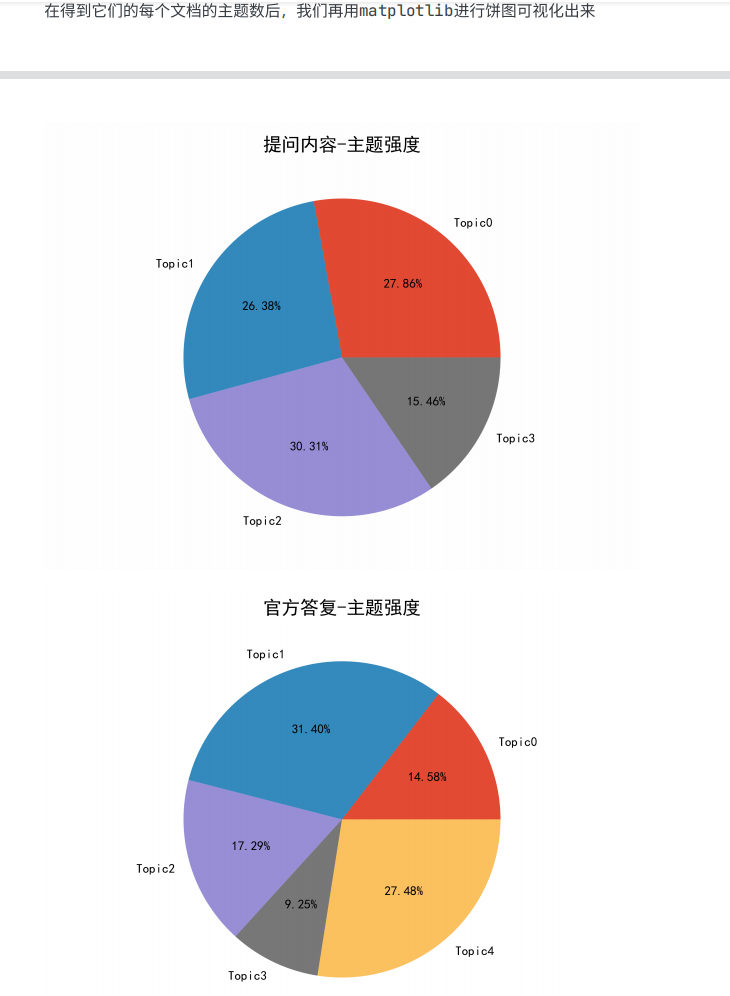
1. **主题可视化展示，也是你之前做滴~**

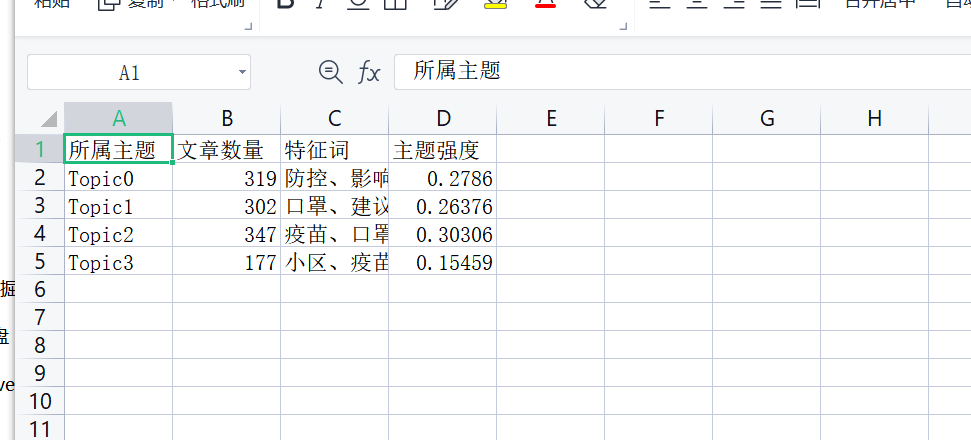
****

1. **主题数计算，也是你之前做滴~**

****

1. **主题强度饼图，也是你之前做滴**

****

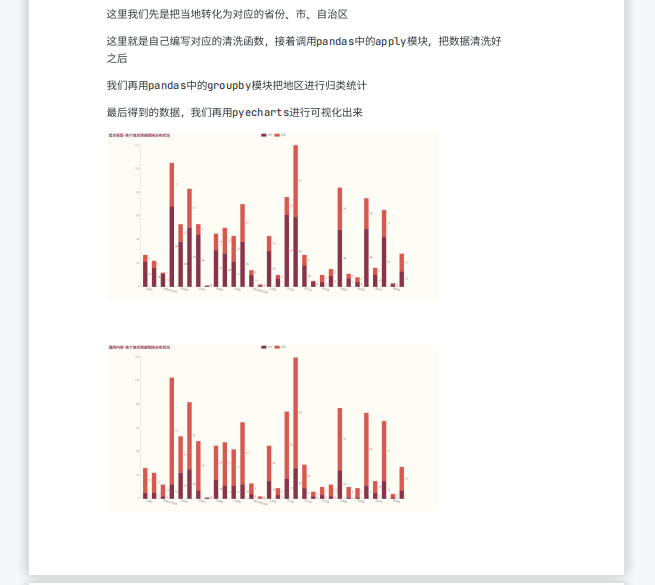
****

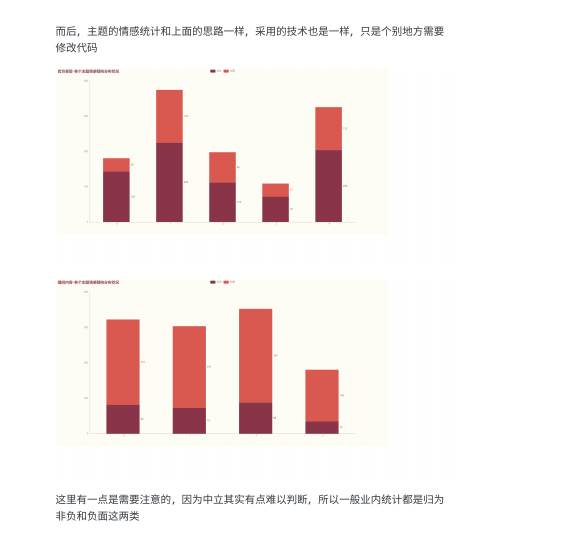
1. **确定主题数后，得到各主体特征词表格，也是你之前做滴**

****

**最后可视化展现出来的图，也继续保留~**

****

****

****