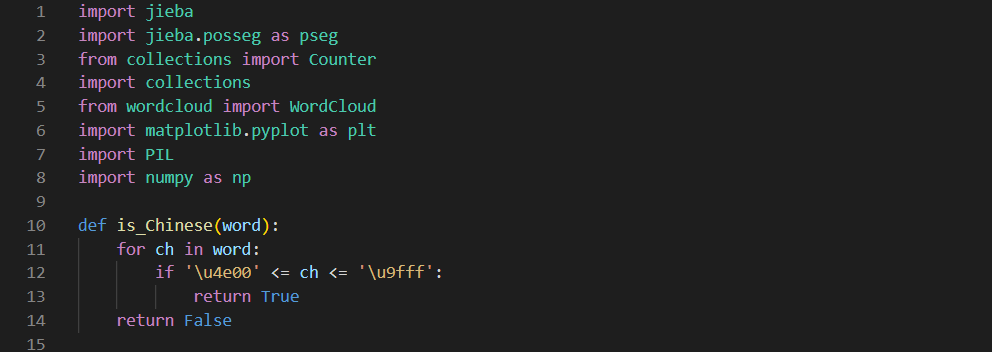
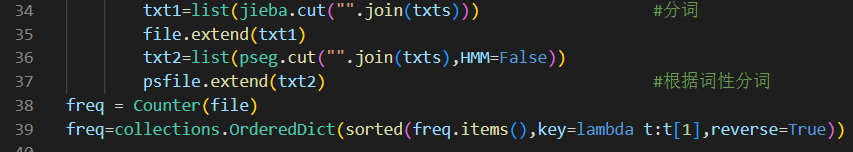
**21377245 黄煜旸 第二周作业**

绘制用户画像是各大APP了解用户需求的重要工具，通过用户的发言与浏览偏好来进行绘制。其中词频统计是重要的一环，以下作业基于提供的微博数据进行词频统计的模拟。

**任务一：读取文件，分隔后去噪**

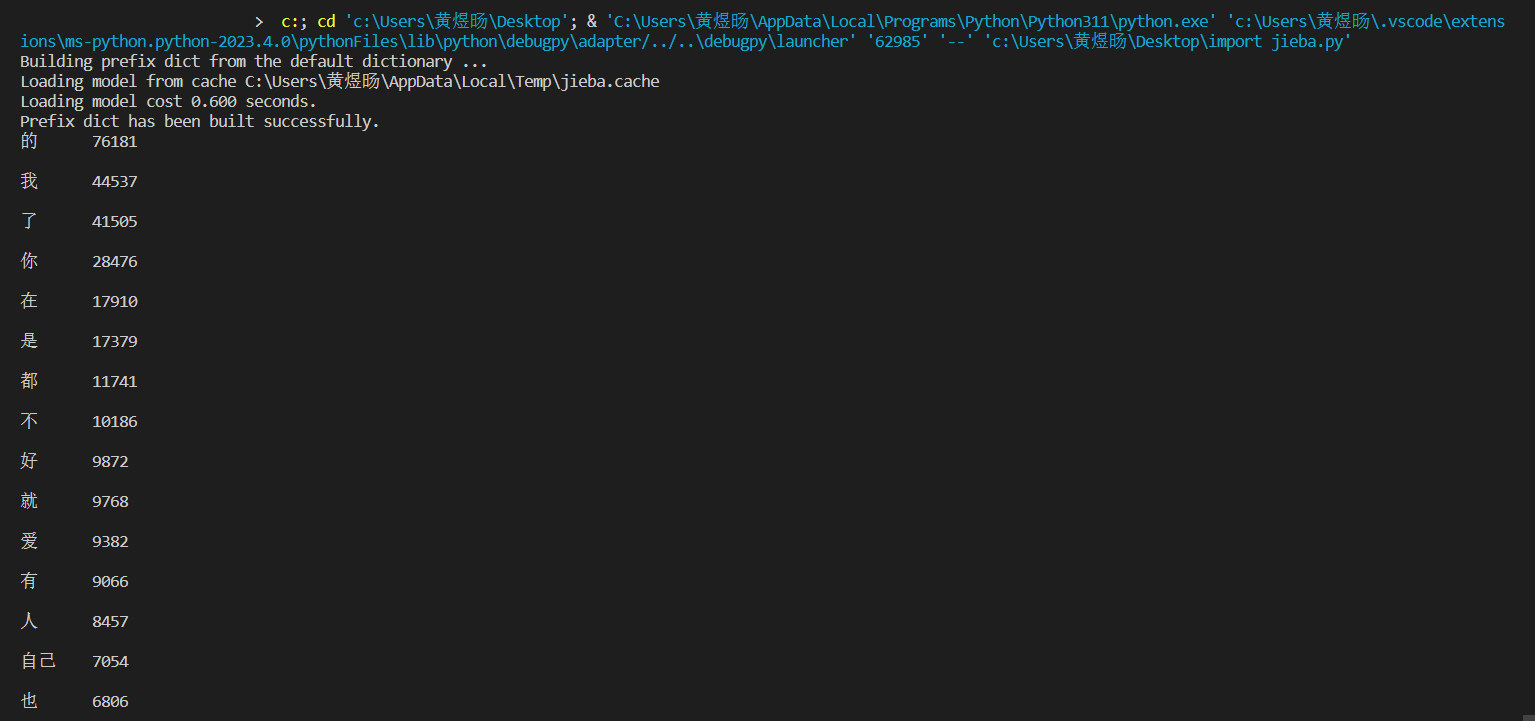
****

**任务二：分词，统计词频，由高到低排序**

****

因为后续任务中需要根据词性分词，为方便起见，在此也建立的另一个列表储存词性的分词

**结果展示：**

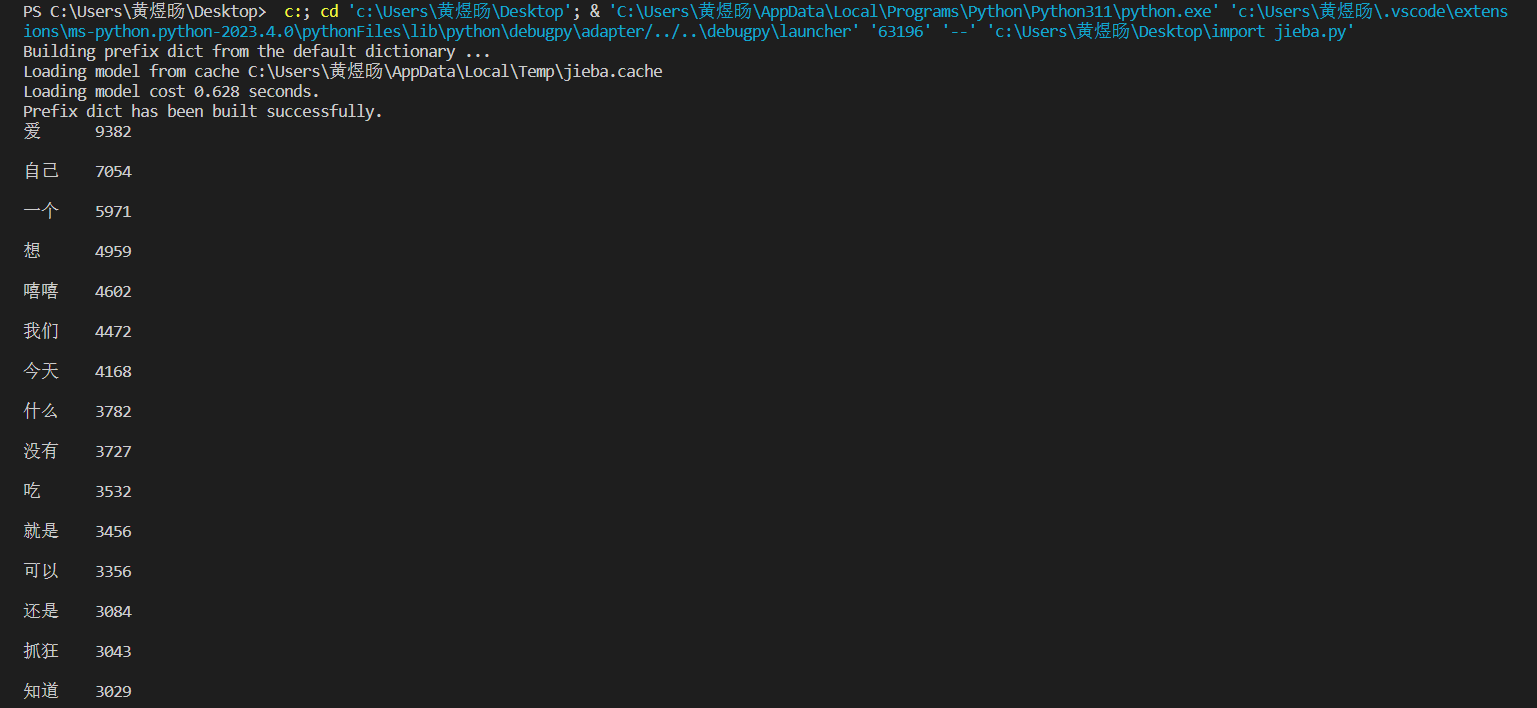
****

**任务三：引入停用词表，重新排列**

****

在网站上面下载停用词表并且引入，筛选后重新得出词频

结果展示：

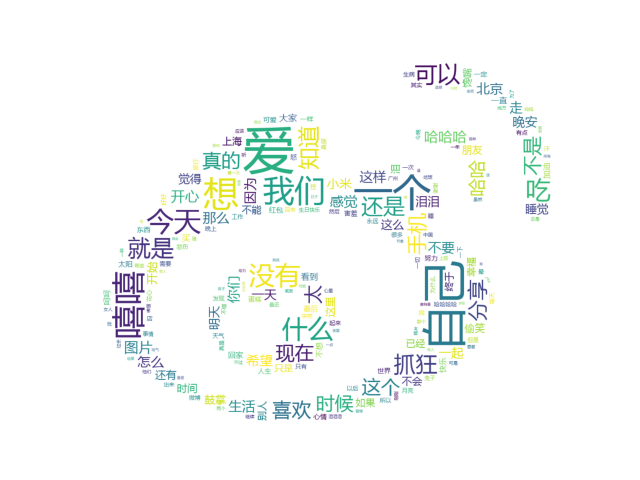


**任务四：可视化词云**

****

因为是基于微博数据，所以导入了一张如下的图片作为背景

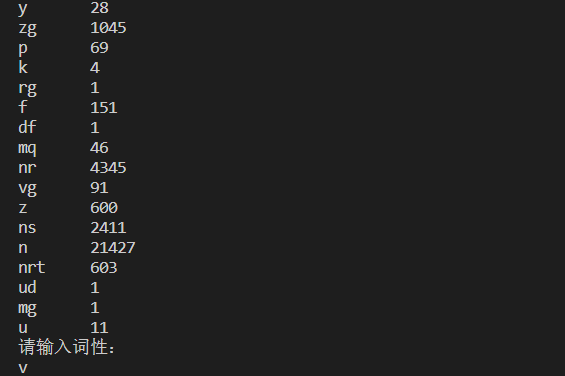
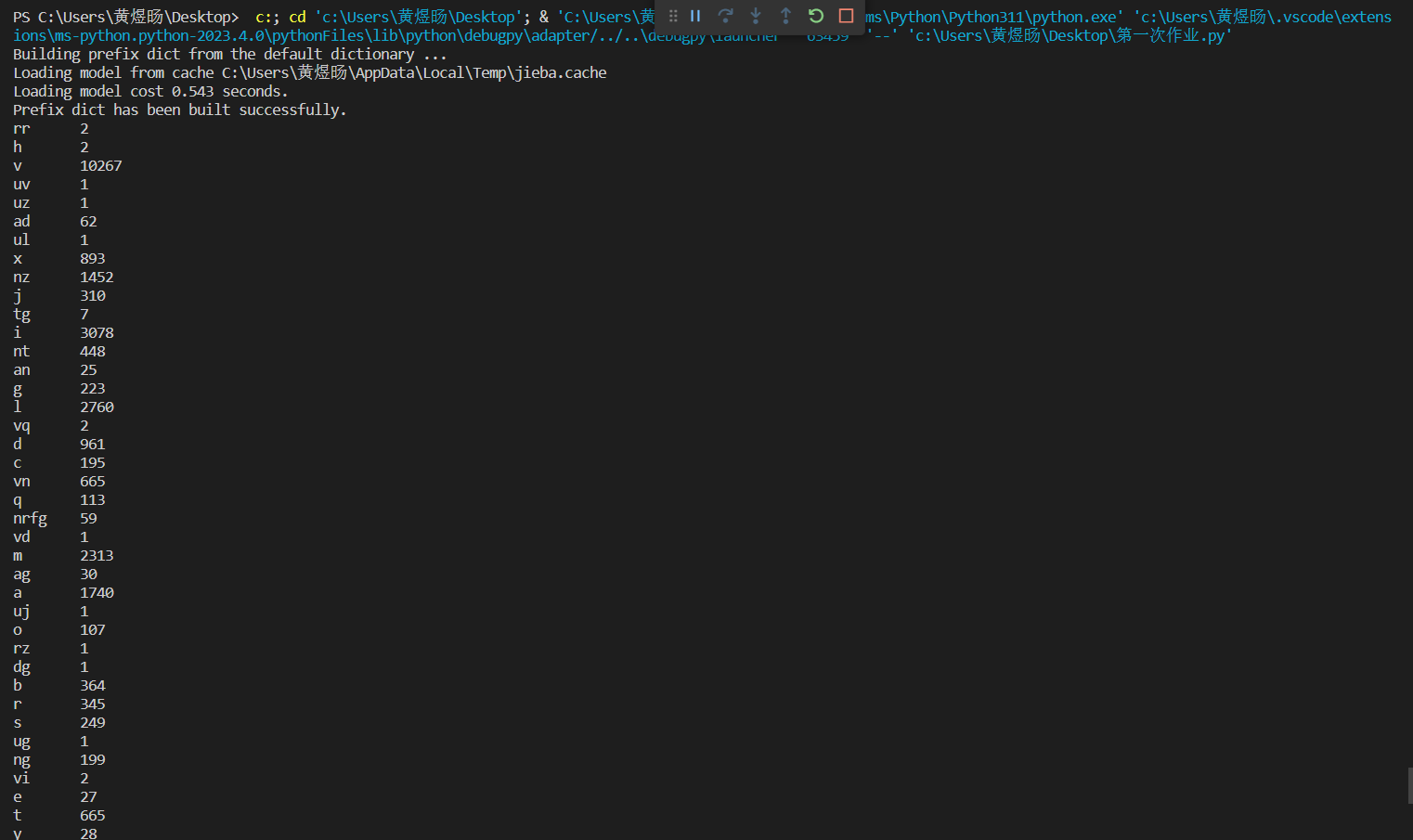
结果展示：



**任务五：词性分析，根据不同词性绘制词云**

****

结果展示：



输入词性v后获得了动词词性的词云