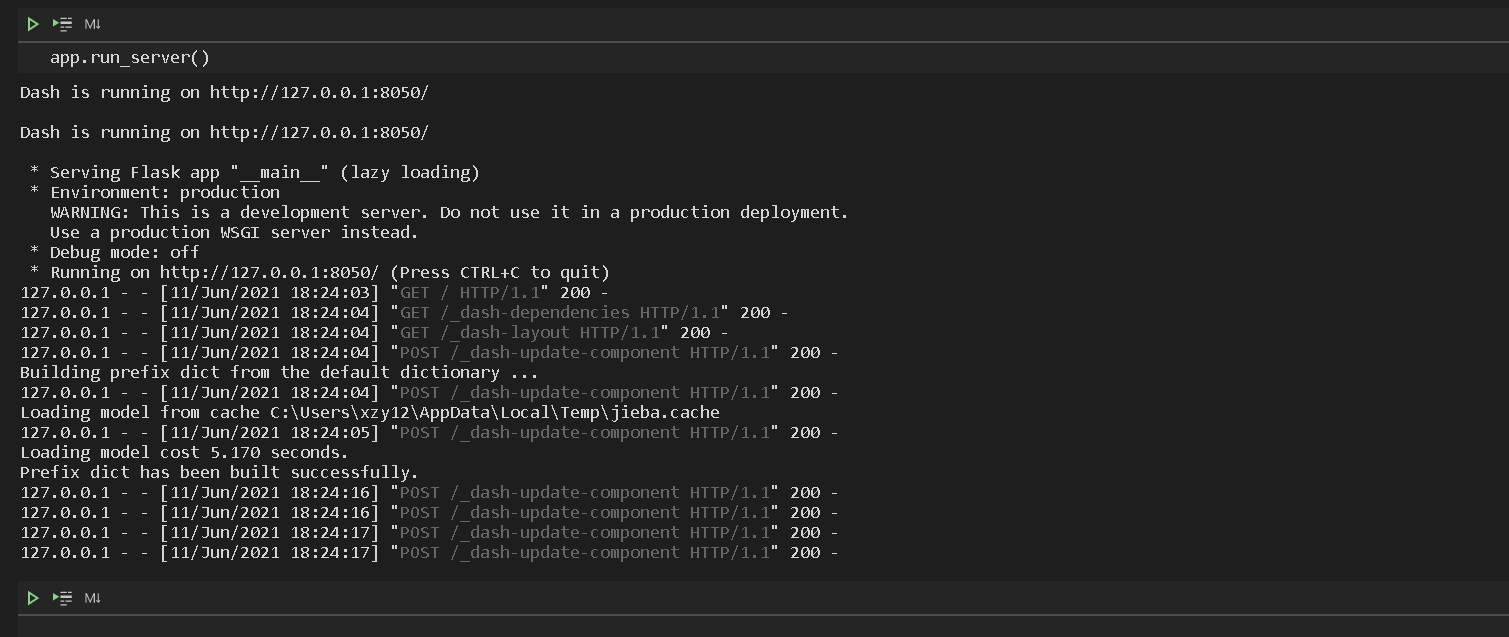
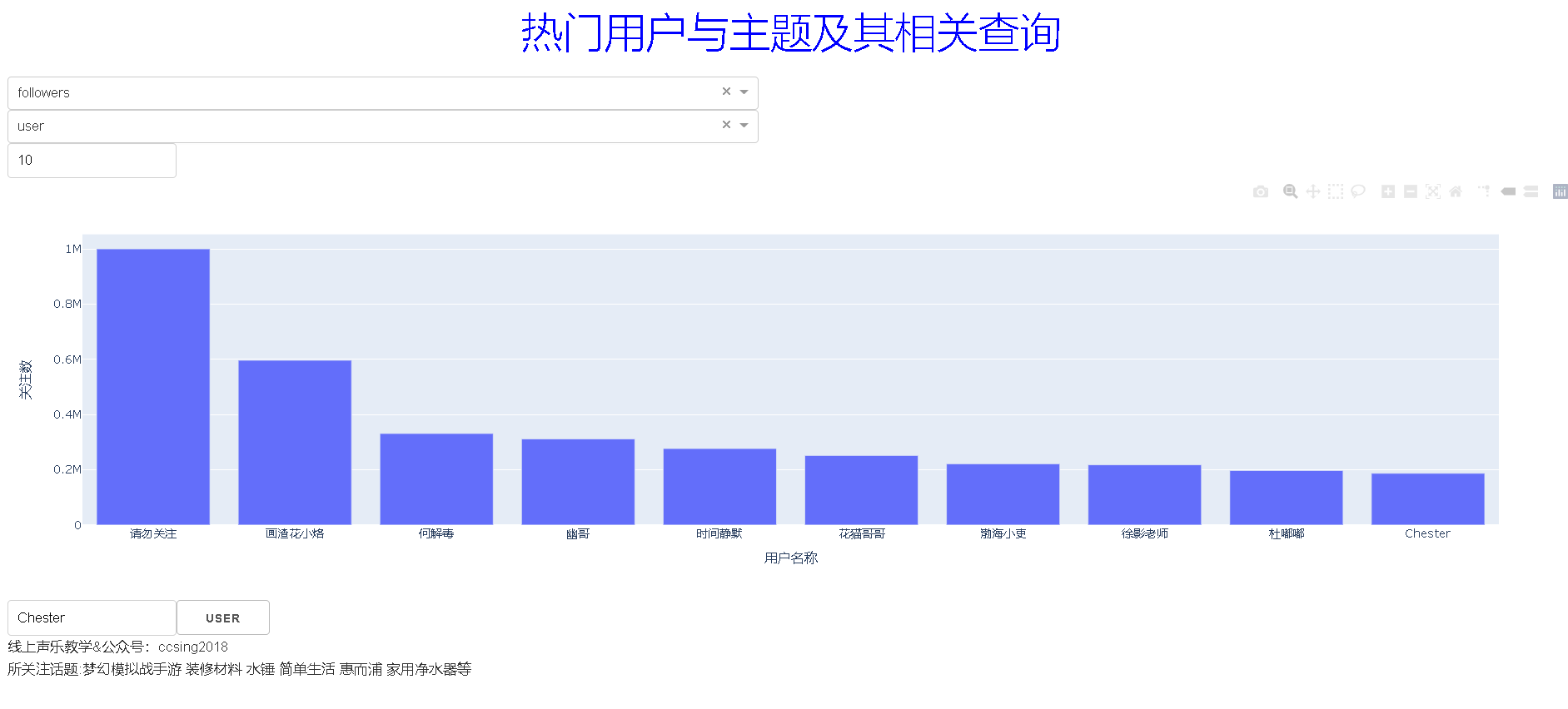
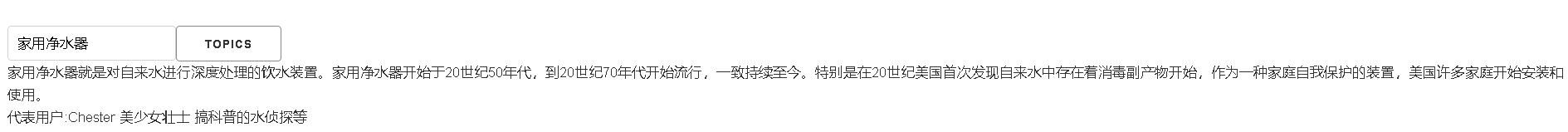
用Visual Studio Code打开testWEBPAGE中的main.ipynb文件,然后点击运行所有单元（假如有未安装的模块，执行pip install去安装相应模块,注意使用了Jupyter），执行到最后一个单元时，根据最后一个单元的指示打开<http://127.0.0.1:8050/> （推荐使用Chrome浏览器，**打开网页后稍微等待一下，全部渲染需要10秒钟左右**）



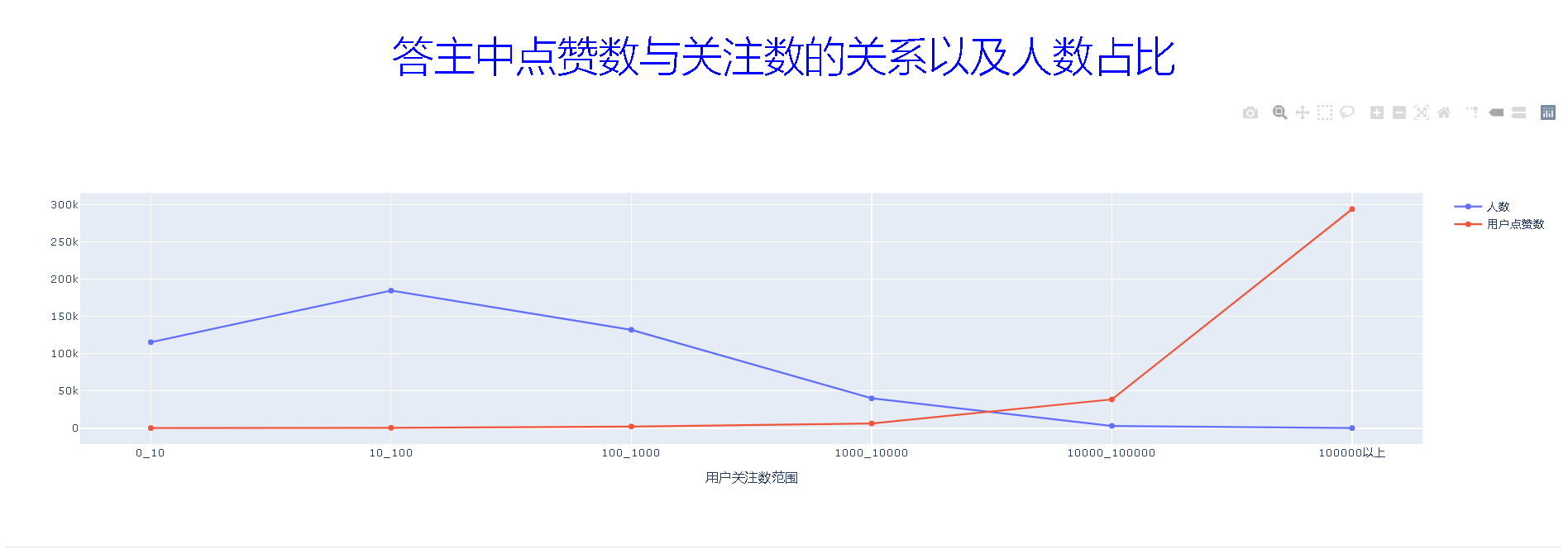
接下来介绍一下内部的交互式操作（详细内容见实验报告中的实验效果部分），



首先是热门用户与主题及其相关查询，上面有两个选择框，一个数字输入框。第一个选择框表示显示的是热门用户还是话题，第二个选择框的值则会根据第一个选择框的值变化，如果第一个选择框选择用户，那第二个选择框就是关注数和总点赞数，而且下方的搜索按钮也会显示为“USER”,如果第一个选择框选择话题，那第二个选择框就是关注数和提问数，下方的搜索按钮显示为“TOPICS”,数字输入框则表示推荐范围，默认选择前10个用户或者话题进行可视化展示，柱状图中y轴根据两个选择框的值联合确定，x轴的范围根据数字输入框确定。搜索部分可以特意搜索上面的推荐用户或者话题，也完全可以单独查询，结果为用户及其相关话题（或者话题及其相关用户），上图已经有一个用户的搜索结果，在介绍完该用户信息之后，下面会显示所关注的话题。如果调换为话题，就会显示话题信息及其相关用户，如图：



接下来是“答主中点赞数与关注数的关系以及人数占比”



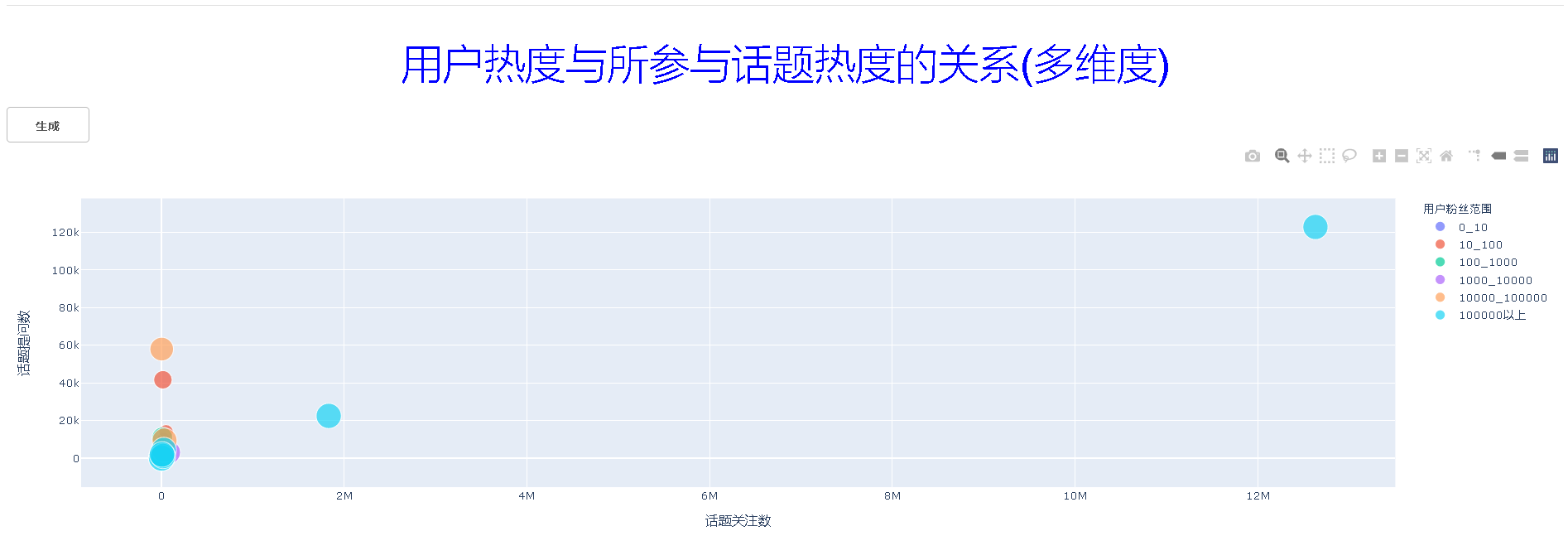
首先通过分箱为每个用户确定了关注人数的范围，然后分组计数，统计点赞数平均值，蓝色表示每个范围内的用户人数，因为爬取时使用的一位知名答主，并且递归的逻辑是爬取他所关注的人，这一方面防止爬虫暂停，另一方面也是对数据的清洗，防止爬到匿名小号，因此我们可以看到，相当一部分人都有关注用户，但头部的明星用户占比很少。红色代表点赞数的平均值，根据图表显示，关注人数在1000人以下的用户点赞数差别不大，在关注人数在1000人以上之后，随着关注人数的增加，用户受到的点赞数出现了指数式攀升。

接下来是文章及其词云图查询，



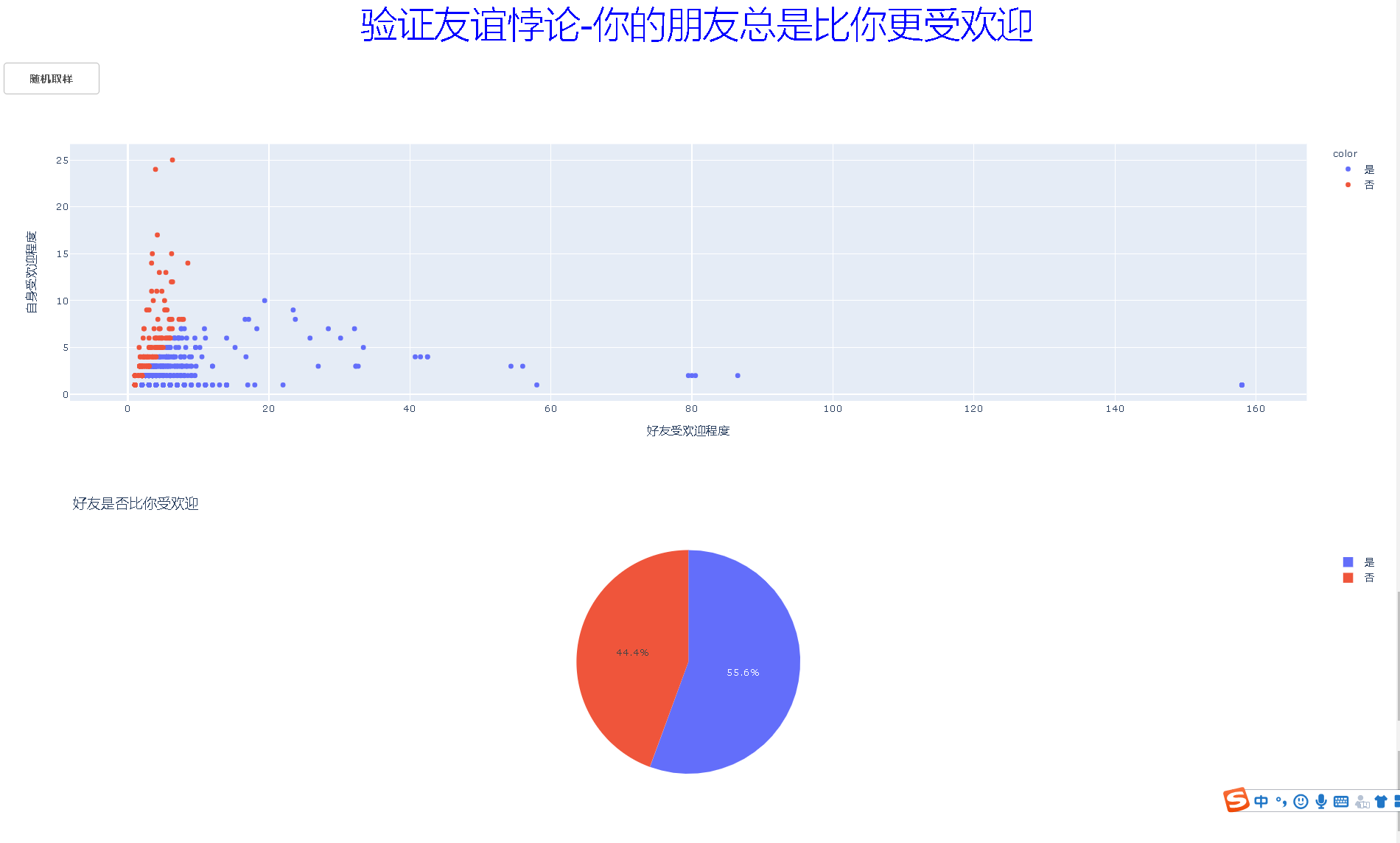
这里的文章是根据点赞数进行排序的，输入框输入文章次序，1代表点赞数最多的文章，其中最大值可达到9600，在点击按钮进行查询之后，就会得到相应的文章并且显示词云图。

接下来是用户热度与所参与话题热度的关系(多维度)



在这里，不同颜色表示不同的用户关注数范围，通过点击右侧颜色点可以选择是否只显示某一部分的颜色。大小表示用户点赞数，x轴代表话题关注数，y轴代表话题提问数。每点击一次生成按钮就会进行一次随机取样，我进行了多次随机取样发现总会有关注数在10000以上的用户关注一个热度特别高的话题，侧面体现了用户与话题的关系，而不同关注数范围的用户点赞数也往往不同，关注越多，点赞越多（为保证能够出现在同一散点图上，进行了取对数操作）。

接下来是验证友谊悖论：



关于如何生成社交网络图的问题已在实验报告的技术路线部分给出陈述，这里表述结果，每次点击随机取样按钮就会对1000个用户进行随机取样并生成社交网络图以及相应的可视化图表，蓝色表示朋友比自己受欢迎，根据散点图可以看出部分用户朋友相当受欢迎，而根据饼状图可以直观的看出认为朋友比自己受欢迎的人数占比更多，多次随机取样并验证后，发现认为朋友比自己受欢迎的人数占比大约在51%到56%之间，验证了友谊悖论。