作业说明：

这次的数据会更加复杂，数据来源于百度贴吧，主要是关于华为手机论坛的帖子信息和回帖内容。为了让每位同学更好地学习和训练，在课前希望大家做如下准备工作。

1. 到该贴吧观察帖子的特点和回复的数据内容http://tieba.baidu.com/p/4604377284
2. 综合前面几次课程的内容将txt数据读取到R中，这次就没有对大家向以前那么好了，帮大家把数据存入了csv，希望大家想方设法在程序中实现把txt文件读入到R中，而不是先存到csv中，然后在读取到R中。Txt文件中，每列之间的数据用table符’\t’进行分隔，大家在读取文件的时候需要注意。我只截取了近5千行数据作为练习，数据文件是sample\_data\_utf8.txt。
3. 每一列表示的意义我在excel文件中第一行已经为大家写清楚了，请根据读入的内容，看看dataframe中是否每一列与excel中的每一列对应。
4. 读入数据后，观察数据特别，由于每个帖子有多次回复，所以一个帖子对应了多行数据。
5. 根据前面课程学习的内容，找出谁发的帖子浏览量最多，谁在该贴吧中回复的帖子最多。

作业截止时间为本周六晚上12点，请发pdf文档（代码附在文档最后）到邮箱：[jasonyaopku@gmail.com](mailto:jasonyaopku@gmail.com)