需要新加的词（sentiment+comment）

\* 可再根据实际情况进行修改，加到sentiments中的词也要加到keyword excel中

1. positive\_sentiments

（暂时没有）

1. ~~neutral\_sentiments~~

~~防滑性、发货、物流、快递、客服、弹力、支撑力、抓力、配色、设计、版型、样式、价格、包裹感、支撑感~~

1. negative\_sentiments

响声、嘎吱声、声音、异响、塑料感、不舒适、开裂

1. positive\_comments

没有差别、非常可以、好多了**、【快、很快、超级快】形容发货物流（只能形容发货，不过目前keyword里没有发货类的，第一步要先在keyword里加上快递，物流这些词）**等、周到、细腻、不粗糙、好太多、不输、没失望、没有失望、不要太便宜、不要太舒适、不是盖的、无差异、没有差异、无偏差、不是一般的、不次于、远超、不逊色

1. ~~negative\_comments~~

~~不强烈、短板、并非、没有感觉到、不喜欢、差强人意、很丑、坏、不精细、牺牲、美中不足、天差地别、比不上、未、有出入~~

1. ~~需要连在一起的词（add\_words）~~

~~不好刷、不好洗、怀疑是不是、没穿几次、没穿几天、没几次、没几天、没有穿几次、没有穿几天、没有想到、从没想过、不要犹豫、不要太~~

特殊情况处理

\*需要再进一步探讨是否修改

1. negative\_comments + negative\_sentiments = neg的情况：【劣质塑料味】应该为负面，不应该是(-1)\*(-1)=1 （先把negative\_comments里的no words挑出来加到no words，然后negative\_comments 就不能再在前面形容negative sentiments；没有感觉到弹性，弹性没有感觉到 no words不能在后面形容sentiments）
2. 以~~【是不是、担心、怕】开头的为中性，将【是不是、担心、怕】加入neu\_begin~~
3. ~~【不是】加入no\_words~~
4. 【合适】从sentiment改到comment能解决不在说【脚型适配】却出现【合适】这个词的情况，但此时合适就不能作为【脚型适配】的keyword了。或者判断【合适】必须和有关【脚型适配】的词一起出现时才起作用。

**遗留问题：**在聚类时应该依旧会把出现【合适】的所有句子聚在一起，依旧无法区分是否为脚型适配，但如果情感分析结果好不需要聚类的话应该可以解决（此时聚类的目的可以只是挑出复购场景人群精神认同）

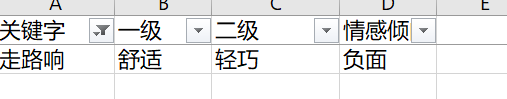
1. 【很快】从sentiment改到comment，或者【很快】的tag改成服务。目前看到说到【很快】的大多数是物流和客服回复，说性能的较少？
2. 【鞋底薄】不是透气，但是keyword【薄】是透气
3. 【不一样的脚感】不是在说脚感不好，【不一样】只能在后面形容sentiments？（不一样样现在是negative comments）
4. 【~~不舒服】应该只在sentiment而不在comment~~
5. ~~【不行】不能在后面形容neg\_sentiments (难看地不行 -1\*-1的问题)~~
6. 【踩屎感有但不强烈】的时候【但】不应该被断句？（不管了）
7. 【有】应该放到just\_before？或者不能在后面形容situation？

同时，对于【'暂时', '没', '发现', '有', '瑕疵'】被误判为负的问题，【有】可以不在前面形容neg\_sentiments？因为neg\_sentiments已经是负的分数，不用pos\_comments加强也是负的，依旧为负面。而【有】的绝大部分作用是在遇到neu\_sentiments时对其进行加强，从而表正面。

1. 【脏了不好洗】等问题：【不好】不能在后面形容【脏】或neg\_sentiemnts？

目前可先将【不好洗、不好刷】放到add\_words中让它不被分开

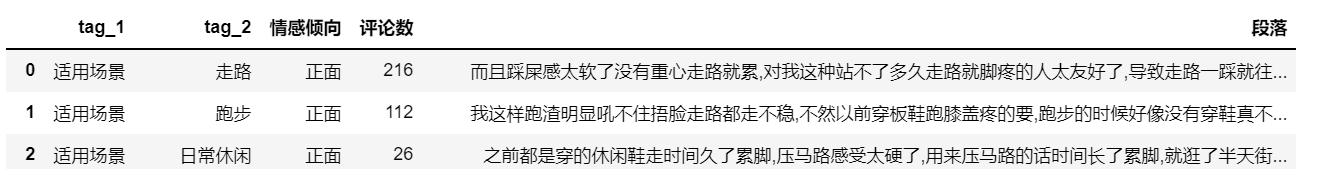
1. 【胶味】等形容胶和声音的加入neg\_after\_useless
2. 【包裹】无法区分是【包裹性】还是【快递包裹】
3. 【臭】无法区分是【透气】还是【做工】
4. 【不是一般的/得XX】问题：【一般】只能在后面形容sentiments
5. 特殊句式：【……就好了】【如果……就好了】【……就更好了】表示负面
6. 特殊句式：【XX不XX】表示中性 （正不正品，正品不正品；好不好过几天再说，）
7. 为什么【走路响】是舒适？



1. 也买、一直买 从keyword 删掉了，有可能是复购，有可能是鞋码

复购场景人群精神认同

1. situation判断正负需要改进，增加正负面判断的词（截图只是部分）



1. 【配速】中的【配】不能因为出现【配】而加到搭配 （搭配里很有多是配速）
2. 去叠字问题：若不去叠字，人群的判断可优化，特殊情况判断可删去
3. 场景里面的休闲，有可能是商品标题，‘休闲跑鞋’ 品牌+休闲鞋+数字（年份） 较多的【鞋】字出现