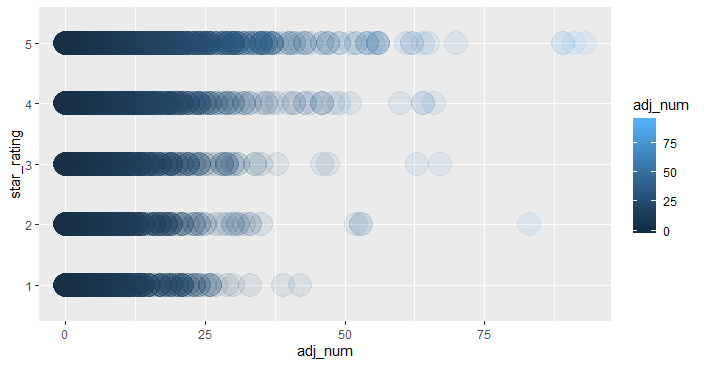
（e题展示数据是以pacifier为例，其他两个数据集规律相似）

通常情况下，形容词是带有强烈感情的词汇，为了判断Are specific quality descriptors of text-based reviews strongly associated with rating levels? 我们可以先看一看评论中出现形容词的数目与star\_rating的关系



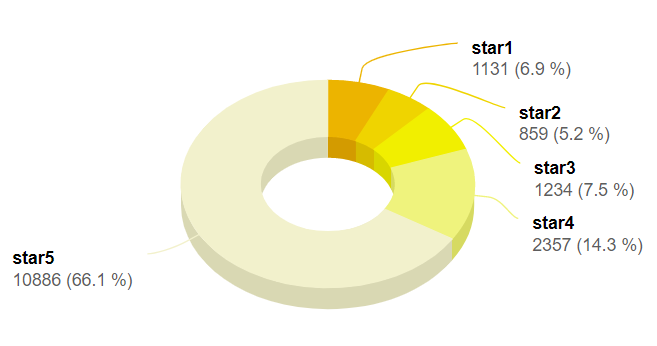
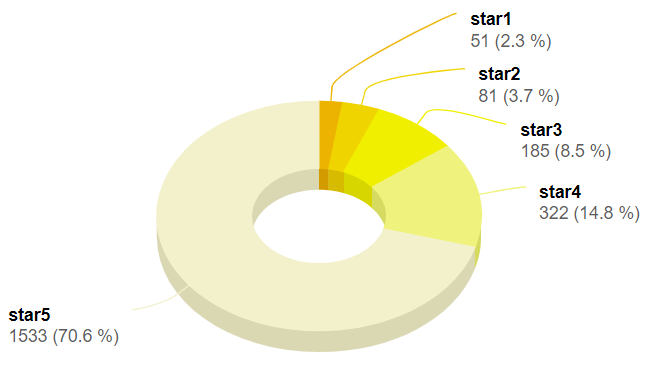
直观的，可以看到随着star\_rating的升高，形容词的数量也倾向于更高，跟我我们的假设：情感词汇大多为形容词，则可以初步认为，带有感情色彩的词汇与评级是strongly associated的。再对上述关系进行单变量回归，得到回归方程 p-value < 2.2e-16，现在对这个相关关系更有把握。

进一步，为了判断若评论里出现带有specific强烈情感的词汇，该评分是否会有显著不同。我们用词频统计出了出现频率较高的词。消极的：“dissapoint”,”bad”；积极的：”comfortable”,”cute”。筛选包含这些词的数据，消极312条，积极2172条(如下表所示)。因为评分机制是5分级，数据的量可能和客观实际不符。我们采用非参数方法中的Ridit检验。他的基本原理是去一个样本较多的组或将几组数据汇总成为参照组，根据参照组的样本结构将原来各组响应数变为参照得分——Ridit得分，利用变换后的Ridit得分进行各处理之间强弱的公平比较。

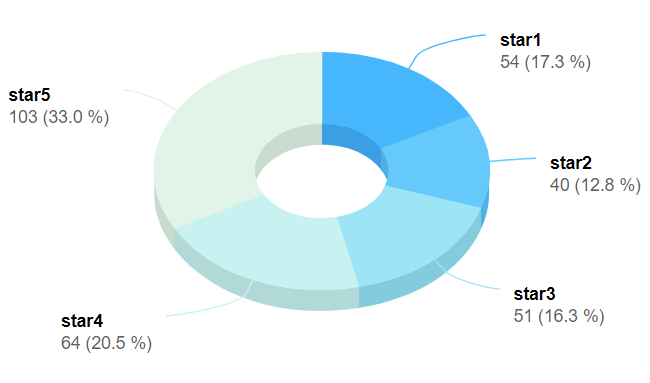
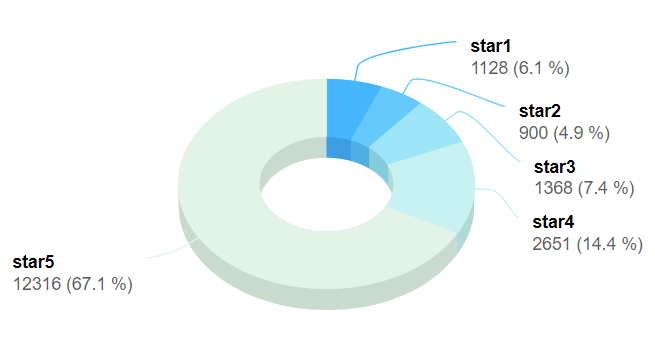
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1star | 2star | 3star | 4star | 5star |
| Disappoint & bad | 54 | 40 | 51 | 64 | 103 |
| Comfortable & cute | 51 | 81 | 185 | 322 | 1533 |

我们这里为了比较两种处理”消极词汇“和”积极词汇“的强弱是否有差异，用R中的Ridit.test()函数检验，检验结果如下表。根据p-value=4.101e-49 < 0.001和Mean Ridit(Comfortable & cute) > Mean Ridit(Disappoint & bad)，拒绝原假设，说明评论中出现Disappoint & bad这类消极词汇的评分会比出现Comfortable & cute这类积极词汇的评分低。进而说明了客户的体验的感受是直接影响对商品的打分的。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean Ridit(Comfortable & cute) | Mean Ridit(Disappoint & bad） | Chi-squared | df | p-value |
| 0.5272 | 0.3103 | 216.99 | 1 | 4.101e-49 |



没有bad&disappointed vs 有bad或disappointed



没有cute&comfortable vs 有cute或comfortable