我们使用皮尔森模型来测量评级与评论的相关性，如果结果中相关性很强，则可以认为评级有可能会影响评论。如果结果中相关性很弱，则可以认为评级不会影响评论。我们将star\_rating作为y，经过处理后reviews的评分作为x输入到如下模型中:

 

最后得出结果如表1：

|  |  |  |
| --- | --- | --- |
| product | Pearson’s coefficient | p\_value |
| hair\_dryer | 0.4289077925901544 | 0.0 |
| microwave | 0.5201799364540123 | 1.1728092568045194e-112 |
| pacifier | 0.38983812064082785 | 0.0 |

表1 pearson结果

结果中三个产品的p\_value值均接近于0，根据pearson原理，结果均可靠，Pearson’s coefficient的值也都很小，所以可以认为用户在看到一些低星或者高星后并不会刺激跟多的评论。可以认为用户的评论是根据自己的使用情况来进行评论的，其他人的评论并不会影响。并且我们在reviews中发现部分如表2中的评论。

|  |  |
| --- | --- |
| Review\_id | reviews |
| R325YZPEPDZH0C | WHY IS THIS ITEM GETTING SUCH BAD REVIEWS ?!?!? I HAVE THIS BLOW DRYER IT'S THE ONLY BLOW DRYER THAT REALLY WORKS ON MY HAIR & DOESN'T BREAK OR BLOW OUT WHILE DEALING WITH MY THICK HAIR. I ONLY WISH THE COMB WAS MORE STRONGER & IT CAME IN PINK :) OTHER THAN THAT NO COMPLAINTS !!! |
| |  | | --- | | R33BLYBNLM87NM | | I've had this hair dryer for 3 years now and it works great. I'm very surprised by reading some negative reviews. I have never had any problems with my hair dryer. My hair is very fine so it gets the job done in 5 minutes. My only complaint is that it's a little loud for me cause I have a baby that takes naps all the time. Other than that it's a great basic hair dryer,nothing fancy. Although I did buy mine at Walmart cause it was cheaper |
| |  | | --- | | ROTACUL0CWK71 | | I don't understand the great reviews for this dryer |
| ⋯ | ⋯ |

表2 部分评论