**BIG DATA ANALYSIS**

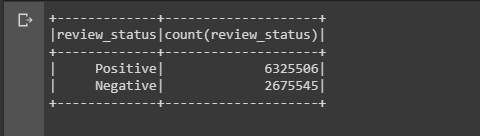
**Peter Pages**

**Amazon Review Data – Category Wireless**

**Dataset: url = "https://s3.amazonaws.com/amazon-reviews-pds/tsv/amazon\_reviews\_us\_Wireless\_v1\_00.tsv.gz"**

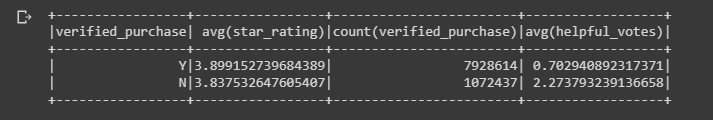
**What defines a positive review?**

A positive review is defined, for this analysis, as any review more then 3 stars. Anything less than 3 stars is considered a negative review.



Therefore, based on those metrics, the reviews were more positive.

The *average rating* grouped by verified purchased



With an average of 3.899 for Verified and 3.837 for Non-Verified. Those which were verified dwarf those not verified. The “non-verifed” reviews were deemed more helped by 3 times as much.

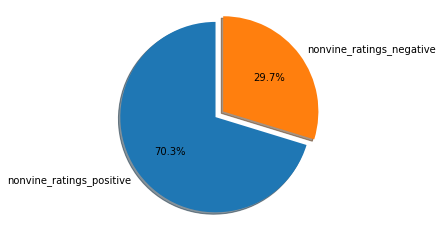
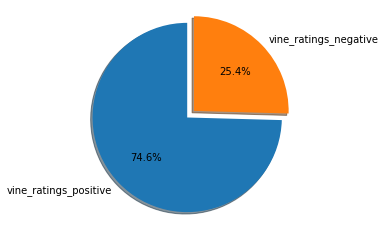
What are Vine Reviews:

*Vine reviews are the independent opinions of the Vine Voices. The vendor cannot influence, modify or edit the reviews. Amazon does not modify or edit Vine reviews, as long as they comply with our*[*posting guidelines*](https://www.amazon.com/review-guidelines/)*. A Vine review is identified with the green stripe Customer review from the Amazon Vine Program.(*[*Amazon*](https://www.amazon.com/gp/vine/help#:~:text=Vine%20reviews%20are%20the%20independent,from%20the%20Amazon%20Vine%20Program.)*)*

Breakdown

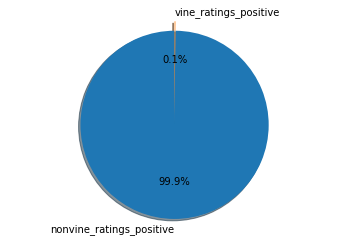
|  |  |
| --- | --- |
| Nonvine positive reviews account for 70.3%  Nonvine negative reviews acccount for 29.7% | Vine positive reviews account for 74.6%  Vine negative reviews account for 25.4% |

**Vine reviews tend to more positive by 4.3% generally but more measured. See 5 star breakdown.**



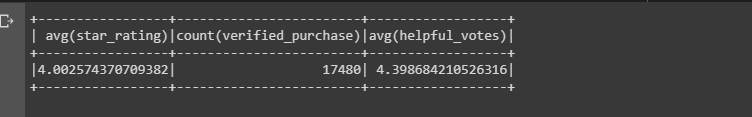
**5 Star Review Breakdown**

When you take a look at highest rating possible, 5 Stars, only 0.1 of vine reviews gave that high a score while the remaining 99.9% of 5-star reviews were given by non-viners.



So, the vine reviews were very hesitant to show too much approval to any product.

The average star rating for a vine review was 4.002  
Average helpful votes were 4.398



**Conclusion**

Viners while making up a very small percentage of reviews overall can a be trusted source of reviews eliminating outlier opinions, both excessively positive and negative.